

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

Gemma L Carvill¹, Sinéad B Heavin², Simone C Yendle², Jacinta M McMahon², Brian J O’Roak³, Joseph Cook¹, Adiba Khan¹, Michael O Dorschner^{4,5}, Molly Weaver^{4,5}, Sophie Calvert⁶, Stephen Malone⁶, Geoffrey Wallace⁶, Thorsten Stanley⁷, Ann M E Bye⁸, Andrew Bleasel⁹, Katherine B Howell¹⁰, Sara Kivity¹¹, Mark T Mackay^{10,12,13}, Victoria Rodriguez-Casero¹⁴, Richard Webster¹⁵, Amos Korczyn¹⁶, Zaid Afawi¹⁷, Nathanel Zelnick¹⁸, Tally Lerman-Sagie^{19,20}, Dorit Lev¹⁹, Rikke S Møller²⁰, Deepak Gill¹⁵, Danielle M Andrade²¹, Jeremy L Freeman^{10,12}, Lynette G Sadleir⁷, Jay Shendure³, Samuel F Berkovic², Ingrid E Scheffer^{2,10,13,22} & Heather C Mefford¹

Epileptic encephalopathies are a devastating group of epilepsies with poor prognosis, of which the majority are of unknown etiology. We perform targeted massively parallel resequencing of 19 known and 46 candidate genes for epileptic encephalopathy in 500 affected individuals (cases) to identify new genes involved and to investigate the phenotypic spectrum associated with mutations in known genes. Overall, we identified pathogenic mutations in 10% of our cohort. Six of the 46 candidate genes had 1 or more pathogenic variants, collectively accounting for 3% of our cohort. We show that *de novo* *CHD2* and *SYNGAP1* mutations are new causes of epileptic encephalopathies, accounting for 1.2% and 1% of cases, respectively. We also expand the phenotypic spectra explained by *SCN1A*, *SCN2A* and *SCN8A* mutations. To our knowledge, this is the largest cohort of cases 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 rate of 3%. Epileptic encephalopathies are a devastating group of epilepsies characterized by refractory seizures and cognitive arrest or regression, associated with ongoing

epileptic activity, that typically carry a poor prognosis¹. *De novo* mutations in several genes are known to be responsible for some epileptic encephalopathies². 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, genetics-based diagnoses of patients can be challenging, as there is both genetic heterogeneity within a given epilepsy syndrome and phenotypic heterogeneity associated with mutations in a specific gene.

The full phenotypic spectrum associated with mutations in genes known to be involved in epileptic encephalopathy is not known. Very few studies have investigated the role of any given gene across a wide spectrum of epileptic encephalopathy syndromes. As a result, serial gene testing, an inefficient and expensive process, is carried out in the clinical setting, with the vast majority of cases remaining unexplained. Furthermore, it is clear that the discovery of additional genes that cause epileptic encephalopathies is needed to facilitate genetics-based diagnosis. Here we take advantage of a high-throughput targeted sequencing approach to perform comprehensive sequence analysis of 65 genes (19 known and 46 candidate disease-related genes) (**Supplementary Fig. 1**) in 500 cases with a range of epileptic encephalopathy phenotypes (**Table 1**). Candidate genes were selected from epilepsy-associated CNVs ($n = 33$) and mutations that caused neurodevelopmental disorders or other epilepsy syndromes ($n = 13$). Using this approach, we (i) identified new genes involved in

¹Department of Pediatrics, Division of Genetic Medicine, University of Washington, Seattle, Washington, USA. ²Department of Medicine, Epilepsy Research Centre, University of Melbourne, Austin Health, Melbourne, Victoria, Australia. ³Department of Genome Sciences, University of Washington, Seattle, Washington, USA. ⁴Department of Psychiatry & Behavioral Sciences, University of Washington, Seattle, Washington, USA. ⁵Veteran Affairs Puget Sound Health Care System, Seattle, Washington, USA. ⁶Neurosciences Children’s Health Queensland, Royal and Mater Children’s Hospitals, Brisbane, Queensland, Australia. ⁷Department of Paediatrics, School of Medicine and Health Sciences, University of Otago, Wellington, New Zealand. ⁸Department of Paediatric Neurology, University of New South Wales, Sydney Children’s Hospital, Sydney, New South Wales, Australia. ⁹Department of Neurology, Westmead Hospital, University of Sydney, Sydney, New South Wales, Australia. ¹⁰Department of Neurology, The Royal Children’s Hospital, Parkville, Melbourne, Victoria, Australia. ¹¹Epilepsy Unit, Schneider Children’s Medical Center of Israel, Petach Tikvah, Israel. ¹²Critical Care & Neurosciences Theme, Murdoch Childrens Research Institute, Melbourne, Victoria, Australia. ¹³Department of Paediatrics, The University of Melbourne, The Royal Children’s Hospital, Melbourne, Victoria, Australia. ¹⁴Paediatric Neurology, Monash Medical Centre, Melbourne, Victoria, Australia. ¹⁵TY Nelson Department of Neurology, The Children’s Hospital at Westmead, Sydney, New South Wales, Australia. ¹⁶Department of Neurology, Tel-Aviv University, Tel-Aviv, Israel. ¹⁷Tel-Aviv University Medical School, Tel-Aviv, Israel. ¹⁸Department of Pediatrics, Carmel Medical Center, Technion Faculty of Medicine, Haifa, Israel. ¹⁹Metabolic-Neurogenetic Service, Wolfson Medical Center, Holon, Israel. ²⁰Danish Epilepsy Centre, Dianalund, Denmark. ²¹Department of Medicine, Division of Neurology, University of Toronto, Toronto Western Hospital, Krembil Neurosciences Program, Toronto, Ontario, Canada. ²²Florey Institute, Melbourne, Victoria, Australia. Correspondence should be addressed to H.C.M. (hmefford@uw.edu) or I.E.S. (scheffer@unimelb.edu.au).

Received 28 December 2012; accepted 1 May 2013; published online 26 May 2013; doi:10.1038/ng.2646

Table 1 Epileptic encephalopathy cohort screened for mutations in 65 new and known disease-relevant genes

Syndrome	Number of cases	Pathogenic or likely pathogenic variants (%)
ABPE	6	0
Dravet	19	4 (21) ^a
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
Infantile spasms	81	4 (5)
LKS	3	0
LGS	40	5 (13)
MAE	81	3 (4)
Ohtahara syndrome	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; LKS, Landau-Kleffner syndrome; LGS, Lennox-Gastaut syndrome; MAE, myoclonic atonic epilepsy.

^aIn four of the Dravet syndrome cases included, *SCN1A* testing had not yet been undertaken.

epileptic encephalopathy and (ii) delineated the phenotypic spectra and mutation frequencies for both known and new genes in epileptic encephalopathies.

Overall, 91% of the target genes (65 genes) were sequenced at >25× coverage, a sequencing depth 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 cases; these cases were not included in the discovery cohort.

We detected 1 or more pathogenic or likely pathogenic mutations in 6 of our 46 candidate genes, with multiple individuals carrying mutations in either of the 2 new genes for epileptic encephalopathy: *CHD2* (NM_001271.3; NP_001262.3) and *SYNGAP1* (NM_006772.2; NP_006763.2) (Fig. 1 and Tables 1 and 2).

Notably, we detected six *de novo* variants in the candidate gene *CHD2* (Figs. 1 and 2), selected from within the interval shared by the 15q26.1 deletions detected in individuals with a range of epileptic encephalopathies (Supplementary Fig. 3)^{5,6}. Four mutations led to premature truncation of *CHD2* (Table 2). Two *de novo* missense variants disrupted highly conserved residues in the SNF2-related helicase/ATPase domain (p.Trp548Arg and p.Leu823Pro) and were predicted to be damaging by both PolyPhen-2 and SIFT. *CHD2* encodes a member of the chromodomain helicase DNA-binding (CHD) family of proteins. All CHD family members are characterized by a chromatin-remodeling domain termed the chromodomain (chromatin organization modifier) and an SNF2-related helicase/ATPase domain. These domains suggest that these proteins function as chromatin remodelers⁷. Although functional studies of *CHD2* have been limited, studies of another CHD protein, *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* showed that mutations affecting the helicase domain, which cause CHARGE (coloboma, heart defects, atresia of the choanae, retardation of growth and/or development, genital and/or urinary abnormalities and ear abnormalities and deafness) syndrome, resulted in decreased remodeling ability⁸. These results suggest that the two

de novo missense mutations described here may disrupt *CHD2* function in a similar manner, whereas truncating mutations probably result in haploinsufficiency.

The six cases with *CHD2* mutations had distinctive features, including median age at seizure onset of 18 months (range of 1–3 years; Table 2), myoclonic seizures (all six), photosensitivity (three) and intellectual disability ranging from moderate to severe (all six). A *de novo* *CHD2* frameshift mutation was reported in a proband with intellectual disability and absence seizures⁹, and a *de novo* missense mutation was identified in an individual with autism spectrum disorder (ASD)¹⁰. These results suggest that mutations in *CHD2* contribute to a broad spectrum of neurodevelopmental disorders. There was no clear genotype-phenotype correlation with respect to either the type of mutation or phenotype-specific intragenic localization (Fig. 2), motivating further *CHD2* mutation screening in phenotypically diverse cohorts. Notably, recent studies have implicated *de novo* mutations in *CHD8* in individuals with ASD¹¹. Three genes of the chromodomain family (*CHD2*, *CHD7* and *CHD8*) have now been implicated in disorders that affect the neurodevelopmental system. Further studies of this nine-member family will determine the role of alterations in each protein across the spectrum of neurodevelopmental disorders and will provide new avenues of research focused on the role of this chromatin-remodeling family in epileptogenesis.

We identified 9 pathogenic or likely pathogenic variants in 4 of the 13 epilepsy-associated genes (Fig. 1). We found five truncating variants in *SYNGAP1* (Fig. 2). Cases with *SYNGAP1* mutations had median age at seizure onset of 14 months (mean of 14 months, range of 6 months to 3 years) (Table 2). They had multiple seizure types, early developmental delay and subsequent regression. Outcomes were poor, with moderate to severe intellectual disability. *SYNGAP1* mutations have been associated with intellectual disability, and, although most cases also have epilepsy, seizures are typically well controlled^{9,12–18}. We observed no genotype-phenotype correlation with respect to *SYNGAP1* mutations, suggesting that the nature of the mutation does not underlie phenotypic heterogeneity. Rather, we hypothesize that alternative neurobiological conditions and mechanisms, genetic or otherwise, underlie this heterogeneity. Our study identifies 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.

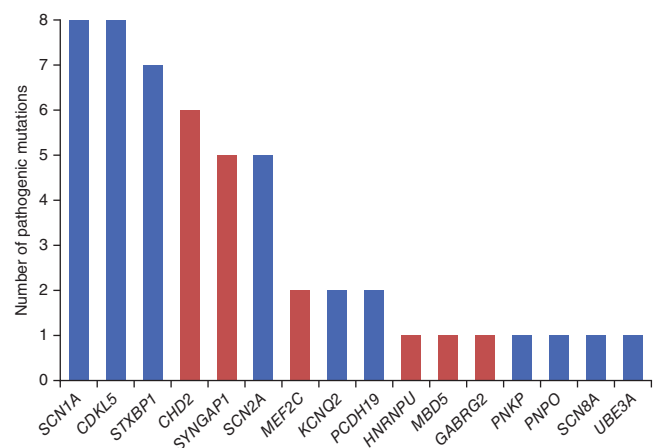


Figure 1 Pathogenic and likely pathogenic mutations identified in 500 cases of epileptic encephalopathies in new (red) and known (blue) disease-related genes.

Table 2 Pathogenic and likely pathogenic variants in new genes for epileptic encephalopathy

Proband	Sex, study age (years)	Gene	Protein change (Polyphen, SIFT) ^a	Diagnosis	Seizures ^b (age of onset)	EEG	Development before seizure onset	Cognitive outcome (regression)
Pathogenic variants (<i>de novo</i>)								
T38	M, 17	<i>CHD2</i>	p.Glu1412Glyfs*64	MAE	A (1 year) FS, absence, MJ-A, MJ, TC	3.8-Hz GSW	Mild delay, behavioral problems	Moderate intellectual disability, ASD (no)
T18697	F, 12	<i>CHD2</i>	p.Arg121*	EE	MJ (1 year) NCS, T, TC, MA	GPSW, MFD, GPFA, SSW	Normal	Severe intellectual disability (yes)
T2608	F, 29	<i>CHD2</i>	p.Gly491Valfs*13	LGS	Atypical absence (1 year) A, MJ, NCS, SE, T, TC	SSW, MFD, DS, GPSW, PPR-triggered MJ	Delayed	Severe intellectual disability (yes)
T20240	M, 12	<i>CHD2</i>	p.Arg1644Lysfs*22	MAE	A (2 years) MJ, SE, TC	DS, GPSW, 2.5-Hz GSW	Normal	Severe intellectual disability (yes)
T17756	M, 15	<i>CHD2</i>	p.Trp548Arg (1, 0)	EE	TC (3 years) FDS, H, MJ	GSW, MFD, DS	Delayed	Moderate intellectual disability (yes)
T18431	M, 2.5	<i>CHD2</i>	p.Leu823Pro (0.999, 0)	EE	FDS, MJ (2.5 years) MA, T	GSW, GPSW, MFD	Delayed	Severe intellectual disability, ASD (yes)
T15923	F, 26	<i>SYNGAP1</i>	p.Trp267*	EE	Atypical absence (3 years) A, aura, FDS, MJ	SSW, MFD	Delayed	Severe intellectual disability, ASD (yes)
T22387	F, 7	<i>SYNGAP1</i>	p.Arg143*	EE	Absence (10 months) MJ	GSW	Delayed	Severe intellectual disability, ASD (yes)
T1898	F, 20	<i>MBD5</i>	p.Thr157Glnfs*4	EE	TC (6 months) Absence, FDS, focal, T	GPSW, MFD, DS	Delayed	Severe intellectual disability, ASD (no)
T20719	M, 2.5	<i>GABRG2</i>	p.Arg323Gln (0.998, 0)	MAE	FS (8 months) Absence. A, MJ, TC	Normal	Normal	Normal (yes)
T23549	F, 3.5	<i>MEF2C</i>	p.Cys39Arg (0.998, 0)	EE	FS (13 months) Absence, SE, TC	MFD, DS	Delayed	Severe intellectual disability, ASD (yes)
T18044	M, 4	<i>MEF2C</i>	p.*464SerExt*?	EE	H (4 months) Absence, A, focal, IS, MJ, TC	MH, MFD	Normal	Intellectual disability (no)
Likely pathogenic								
T19988 ^c	M, 18	<i>SYNGAP1</i>	p.Lys108Valfs*25	EE	Unknown (in foster care) FDS	MFD, DS	Unknown	Moderate intellectual disability, ASD (unknown)
T15924 ^d	M, 11	<i>SYNGAP1</i>	Unknown (c.389–2A>T)	EE	Absence (6 months) TC	GSW, GPSW, MFD	Delayed	Severe intellectual disability, ASD (yes)
T2528 ^d	M, 26	<i>SYNGAP1</i>	p.Gln702*	EE	FS (18 months) Absence, aura, FDS, MJ, NCS, TC	SSW, bioccipital ED, DS	Delayed	Moderate intellectual disability (yes)
T162 ^d	M, 33	<i>HNRNPU</i>	p.Tyr805*	LGS	A (2 years) Atypical absence, MJ, NCS, T, TC	GSW, GPSW, DS, SSW, GPFA	Delayed	Severe intellectual disability (yes)

M, male; F, female; A, 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; EEG, electroencephalography; DS, diffuse slowing; ED, epileptiform discharge; GPFA, generalized paroxysmal fast activity; GPSW, generalized polyspike wave; GSW, generalized spike wave; MFD, multifocal discharges; PPR, photoparoxysmal response; MA, myoclonic absence; SSW, slow spike wave; MH, modified hypsarrhythmia.

^aPolyPhen-2 and SIFT scores are given only for missense variants. ^bInitial seizure listed first (age of onset) followed by subsequent seizure types. ^cDNA from parents unavailable. ^dDNA from father unavailable.

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 (p.Cys39Arg) and a stop-loss variant (p.*464SerExt*?). Furthermore, we found *de novo* pathogenic variants in *MBD5* (NM_018328.4; NP_060798.2) (p.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-identified candidate gene *HNRNPU* (NM_031844.2; NP_114032.2). The p.Tyr805* change arose as a result of two adjacent single-nucleotide changes, c.471T>C and c.472A>T (Supplementary Fig. 4), that occurred two amino acids upstream of the stop codon. Neither mutation was maternally inherited, and paternal DNA was not available for analysis. A recent report identified *HNRNPU* as a candidate for the intellectual disability and seizure phenotypes of probands with 1q44 microdeletions¹⁹. In addition, a *de novo* splice-site variant was identified in a proband with a complex neurodevelopmental phenotype including epilepsy²⁰. Collectively these data suggest that haploinsufficiency for *HNRNPU* is associated with epileptic

encephalopathy as well as with intellectual disability, although further genotype-phenotype correlation will improve understanding of the phenotypic spectrum corresponding to *HNRNPU* mutations.

We identified 31 variants fulfilling our criteria for pathogenicity and an additional 5 variants that are likely pathogenic in 10 of 19 known genes involved in epileptic encephalopathy (Fig. 1 and Tables 1 and 3). We identified multiple cases with mutations in *STXBPI*, *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 6 of the 19 known genes but 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 cases with mutations in genes known to be involved in disease are listed (Table 3), and, for some of these, we have expanded 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

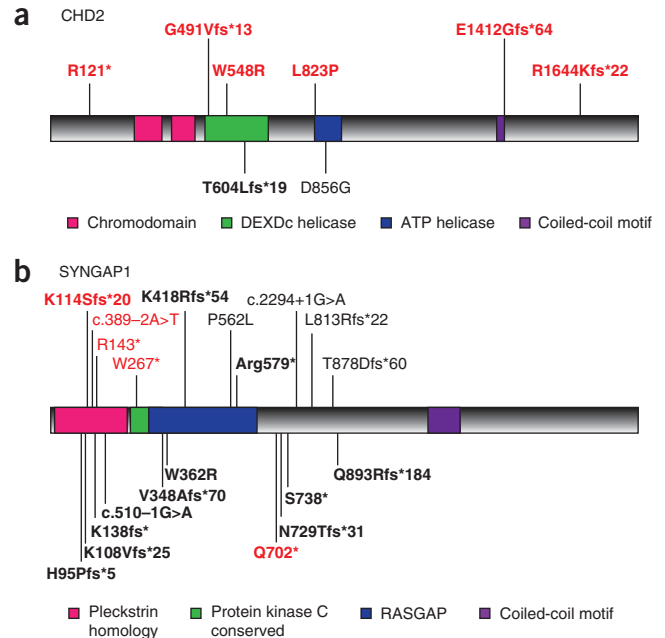
Figure 2 Protein alterations encoded by *de novo* mutations in new genes for epileptic encephalopathies. (a,b) Alterations are shown for CHD2 (a) and SYNGAP1 (b). Alterations shown in red were identified in this study; black entries denote previously reported variants of CHD2 in intellectual disability (Thr604Leufs*19)⁹ and autism (Asp856Gly)¹⁰ and of SYNGAP1 in intellectual disability and/or autism^{9,14–18}. Bold entries indicate pathogenic variants found in individuals with epilepsy. No evident genotype-phenotype correlations exist for mutations in either *CHD2* or *SYNGAP1*.

comprising microcephaly, early-onset intractable seizures and developmental delay²¹. By contrast, our case did not have microcephaly (head circumference was in the 50th percentile) or developmental delay but had normal cognition, despite refractory epilepsy with multiple seizure types. Also, three cases with *SCN1A* mutations presented with an epilepsy-aphasia phenotype, with two also having febrile seizures plus (FS+). *SCN1A* mutations are well known to be associated with genetic epilepsy with febrile seizures plus (GEFS+) but have not previously been reported in epilepsy-aphasia syndromes^{22,23}. It is possible that the *SCN1A* mutation is not responsible for the epilepsy-aphasia syndrome, but it could be a genetic modifier predisposing the individual to this group of epileptic encephalopathies (Supplementary Fig. 5). Further work is warranted to clarify this association, and exome sequencing of these cases could represent an effective approach.

We detected five variants in *SCN2A*, which encodes the $\alpha 2$ subunit of a voltage-gated sodium channel. Until now, 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 were reported in individuals with epileptic encephalopathies^{25,26}. Our five cases notably showed similar variability in the range of age at onset as seen in BFNIS, with three having seizures beginning in the neonatal period (11 h to 2 d after birth) and two having seizures beginning in infancy (at 6 weeks and 13 months of age). In two cases, seizures stopped relatively early at 5 weeks and 7 months of age. The refractory nature of seizures did not correlate with intellectual outcome, which ranged from mild (two probands) to severe (three probands) 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 (encoding p.Leu1290Val) in *SCN8A*. Before now, only a single *de novo* *SCN8A* mutation (encoding p.Asn1768Asp) had been described in a proband with severe epileptic encephalopathy and sudden unexplained death in epilepsy²⁷. Here we describe a second case presenting with epileptic encephalopathy beginning at 18 months of age. Notably, the identified variant was paternally inherited, although the father was shown to have somatic mosaicism (with 13% of cells with the mutant allele), supporting the interpretation of this variant as pathogenic as seen in other genetic encephalopathies with parental mosaicism²⁸.

The findings in this large series of cases with hitherto unsolved epileptic encephalopathies allow 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 found in 16 genes. However, this mutation rate is probably an underestimation of the true contribution of each gene to the overall burden of epileptic encephalopathies. Our cohort excluded individuals with previously identified mutations, and we were unable to conduct segregation analysis for a subset of variants that we identified, some of which are likely to be pathogenic. Furthermore, as larger numbers of cases with mutations of specific genes are identified, distinctive epileptic encephalopathy phenotypes are likely to emerge.



Considering these results together with the up to 8% prevalence rate of rare CNVs in cases with epileptic encephalopathy identified in an earlier analysis of a subset of the current series³, we can now ascribe causal variants to ~18% of all epileptic encephalopathies of previously unknown cause.

The genetic heterogeneity of epileptic encephalopathies is considerable; likely pathogenic variants were found in nine known or new disease-associated genes (Fig. 2). Even the most commonly mutated genes in our study each accounted for only up to 1.6% of cases. Notably, we identified new genes that are commonly mutated in epileptic encephalopathies, with *CHD2*, *SYNGAP1* and *SCN2A* each accounting for 1–1.2% of cases, a frequency similar to that seen for mutations in *SCN1A*, *STXBP1* and *CDKL5* in our cohort. However, no mutations were seen in 9 other known disease-related genes (*ARX*, *FOXG1*, *KCNT1*, *MECP2*, *PLCB1*, *SLC25A22*, *SLC2A1*, *SPTAN1* and *ARHGEP9*) in 500 cases. These results suggest that pathogenic mutations in these genes, although important, are rare causes of epileptic encephalopathies (accounting for <0.2% each in our cohort) or cause only very distinct syndromes that were not prevalent in our cohort. These findings support the notion of 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 individual.

Notably, mutations in *SYNGAP1* and *CHD2* have now been described in probands with epileptic encephalopathy, intellectual disability 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 cases with *de novo* mutations in the same gene^{9,18}. Our results show the power of targeted resequencing to screen large numbers of affected individuals in a high-throughput and cost-effective manner. This approach is critical to identify additional individuals 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

Table 3 Pathogenic variants in known genes for epileptic encephalopathy

Gene	Proband (sex)	Inheritance	Inferred effect	cDNA change	Protein change	Diagnosis
<i>SCN1A</i> (NM_001165963.1, NP_001159435.1)	T23445 (F)	<i>De novo</i>	Pathogenic	c.4836delC	p.Ile1613Phefs*5	Dravet
	T1639 (M)	Segregates	Pathogenic	c.5962G>A	p.Arg1988Trp	Epilepsy aphasia, FS+
	T18466 (M)	<i>De novo</i>	Pathogenic	c.4033G>A	p.Pro1345Ser	EOEE
	T18594 (M)	Segregates ^a	Pathogenic	c.133C>T	p.Asp45Asn	Epilepsy aphasia, FS+
	T19875 (F)	<i>De novo</i>	Pathogenic	c.3977G>A	p.Ala1326Val	Epilepsy- aphasia
	T18775 (M)	Segregates ^a	Pathogenic	c.1076T>G	p.Asn359Thr	Dravet (ICEGTC) ²⁹
	T18997 (M)	Unknown, father unavailable	Likely pathogenic	c.1209delA	p.Phe403Leufs*12	Dravet
	T19963 (M)	<i>De novo</i>	Pathogenic	c.4453T>C	p.Asn1485Asp	Dravet
<i>SCN2A</i> (NM_021007.2, NP_066287.2)	T20632 (F)	Unknown, parents unavailable	Likely pathogenic	c.408G>T	p.Met136Ile ^b	EOEE evolving to infantile spasms
	T21005 (F)	<i>De novo</i>	Pathogenic	c.2715G>C	p.Lys905Asn	EE
	T22816 (F)	Unknown, father unavailable	Likely pathogenic	c.2783T>G	p.Phe928Cys ^b	EE
	T20340 (M)	<i>De novo</i>	Pathogenic	c.3057delA	p.Ile1021Tyrfs*16	LGS
	T24127 (F)	<i>De novo</i>	Pathogenic	c.5645G>A	p.Arg1882Gln	EE
<i>PCDH19</i> (NM_001184880.1, NP_001171809.1)	T23579 (F)	X linked, female restricted	Pathogenic	c.1681G>A	p.Pro561Ser	EE
	T23305 (F)	X linked, female restricted	Pathogenic	c.2873C>T	p.Arg958Gln	LGS
<i>CDKL5</i> (NM_001037343.1, NP_001032420.1)	T20819 (M)	<i>De novo</i>	Pathogenic	c.464-2A>G	Unknown	EOEE
	T22724 (M)	Inherited from unaffected mother, X linked	Pathogenic	c.433C>T	p.His145Tyr	EE
	T22954 (F)	<i>De novo</i>	Pathogenic	c.545T>C	p.Leu182Pro	EOEE
	T897 (F)	<i>De novo</i>	Pathogenic	c.2564C>G	p.Ser855*	Infantile spasms
	T23057 (M)	<i>De novo</i>	Pathogenic	c.1926delT	p.Leu642Argfs*16	Infantile spasms
	T23951 (M)	<i>De novo</i>	Pathogenic	c.533G>A	p.Arg178Gln	EOEE
	T23234 (F)	<i>De novo</i>	Pathogenic	c.620G>A	p.Gly207Glu	EE
	T24139 (M)	Unknown, parents unavailable	Likely pathogenic	c.1926delT	p.Leu642Argfs*16	EOEE
<i>STXBP1</i> (NM_001032221.3, NP_001027392.1)	T22595 (M)	<i>De novo</i>	Pathogenic	c.1154delC	p.Met387Tyrfs*17	Ohtahara syndrome
	T1266 (M)	Unknown, mother unavailable	Likely pathogenic	c.1630G>T	p.Gly544Cys ^b	LGS
	T23151 (F)	<i>De novo</i>	Pathogenic	c.125C>T	p.Ser42Phe	EOEE
	T23553 (F)	<i>De novo</i>	Pathogenic	c.238T>C	p.Ser80Pro	EE
	T23122 (M)	<i>De novo</i>	Pathogenic	c.568C>T	p.Arg190Trp	EOEE
	T22856 (M)	<i>De novo</i>	Pathogenic	c.1060T>C	p.Cys354Arg	Ohtahara syndrome
	T23289 (M)	<i>De novo</i>	Pathogenic	c.1708A>G	p.Thr570Ala	EE
<i>UBE3A</i> (NM_000462.3, NP_000453.2)	T23859 (F)	Inherited from unaffected mother, affected sibling also had mutation	Pathogenic	c.1585G>A	p.Arg506Cys ^c	EE (features suggestive of Angelman syndrome)
<i>SCN8A</i> (NM_001177984.2, NP_001171455.1)	T3929 (M)	Inherited from somatic mosaic father ^d	Pathogenic	c.3868C>G	p.Leu1290Val	EE
<i>KCNQ2</i> (NM_004518.4, NP_004509.2)	T24158 (M)	<i>De novo</i>	Pathogenic	c.587G>A	p.Ala196Val	EOEE
	T23919 (F)	<i>De novo</i>	Pathogenic	c.602C>T	p.Arg201His	EOEE, IS
<i>PNPO</i> (NM_018129.3, NP_060599.1)	T23451 (M)	Homozygous recessive	Pathogenic	c.686G>A	p.Arg229Gln ^e	EME
<i>PNKP</i> (NM_007254.3, NP_009185.2)	T23141 (M)	Homozygous recessive	Pathogenic	c.58G>A	p.Pro20Ser	EE

EE, epileptic encephalopathy not otherwise specified; ICEGTC, intractable childhood epilepsy with generalized tonic clonic seizures.

^aVariant segregates with the disorder (pedigrees shown in **Supplementary Fig. 5**). ^bThree missense variants that are likely pathogenic (Online Methods); ^cKnown pathogenic variant³⁰.

^dFather is somatic mosaic, with 13% of cells carrying an alternate, pathogenic allele. ^eKnown pathogenic variant in dbSNP, rs104894629 (ref. 31).

additional mutation-positive cases for a specific gene and enhance understanding of disease mechanisms.

URLs. Exome Variant Server, National Heart, Lung, and Blood Institute (NHLBI) GO Exome Sequencing Project (ESP) (accessed October 2012), <http://evs.gs.washington.edu/EVS/>; Genome Analysis Toolkit (GATK, version 2.2, <http://www.broadinstitute.org/gatk/>; SeattleSeq (version 134, <http://snp.gs.washington.edu/SeattleSeqAnnotation134/>;

Burrows-Wheeler Aligner (BWA, version 0.5.9, <http://bio-bwa.sourceforge.net/>; Picard tools (version 1.82), <http://picard.sourceforge.net/>; PolyPhen-2, <http://genetics.bwh.harvard.edu/pph2/>; SIFT (Sorting Intolerant from Tolerant), <http://sift.bii.a-star.edu.sg/>.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the individuals with epileptic encephalopathies and their families for participating in our research. H.C.M. is supported by a grant from the US National Institutes of Health (NIH; National Institute of Neurological Disorders and Stroke (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. and I.E.S., practitioner fellowship 1006110 to I.E.S.) and a Health Research Council of New Zealand project grant to L.G.S.

AUTHOR CONTRIBUTIONS

G.L.C., H.C.M. and I.E.S. designed the study and wrote the manuscript. H.C.M. and I.E.S. supervised the study. G.L.C. constructed libraries, developed the variant calling pipeline (with assistance from J.C.) and analyzed the sequence data. B.J.O. and J.S. developed the molecular inversion probe (MIP) methodology and analysis pipeline. S.B.H., S.C.Y., J.M.M., S.C., S.M., G.W., T.S., A.M.E.B., A.B., K.B.H., S.K., M.T.M., V.R.-C., R.W., A. Korczyn, Z.A., N.Z., T.L.-S., D.L., R.S.M., D.G., D.M.A., J.L.F., L.G.S., S.F.B. and I.E.S. performed phenotypic analysis. S.B.H., J.M.M., S.F.B. and I.E.S. critically reviewed the manuscript. G.L.C. and A. Khan performed segregation analysis experiments. M.O.D. and M.W. performed Illumina HiSeq sequencing.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Berg, A.T. *et al.* Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005–2009. *Epilepsia* **51**, 676–685 (2010).
- Kamien, B.A., Cardamone, M., Lawson, J.A. & Sachdev, R. A genetic diagnostic approach to infantile epileptic encephalopathies. *J. Clin. Neurosci.* **19**, 934–941 (2012).
- Mefford, H.C. *et al.* Rare copy number variants are an important cause of epileptic encephalopathies. *Ann. Neurol.* **70**, 974–985 (2011).
- Heinzen, E.L. *et al.* Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. *Am. J. Hum. Genet.* **86**, 707–718 (2010).
- Capelli, L.P. *et al.* Deletion of the *RMGA* and *CHD2* genes in a child with epilepsy and mental deficiency. *Eur. J. Med. Genet.* **55**, 132–134 (2012).
- Dhamija, R. *et al.* Microdeletion of chromosome 15q26.1 in a child with intractable generalized epilepsy. *Pediatr. Neurol.* **45**, 60–62 (2011).
- Marfella, C.G. & Imbalzano, A.N. The Chd family of chromatin remodelers. *Mutat. Res.* **618**, 30–40 (2007).
- Bouazoune, K. & Kingston, R.E. Chromatin remodeling by the CHD7 protein is impaired by mutations that cause human developmental disorders. *Proc. Natl. Acad. Sci. USA* **109**, 19238–19243 (2012).
- Rauch, A. *et al.* Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* **380**, 1674–1682 (2012).
- Neale, B.M. *et al.* Patterns and rates of exonic *de novo* mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012).
- O’Roak, B.J. *et al.* Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* **338**, 1619–1622 (2012).
- Hamdan, F.F. *et al.* Mutations in *SYNGAP1* in autosomal nonsyndromic mental retardation. *N. Engl. J. Med.* **360**, 599–605 (2009).
- Hamdan, F.F., Gauthier, J., Rouleau, G.A. & Michaud, J.L. *De novo* mutations in *SYNGAP1* associated with non-syndromic mental retardation. *Med. Sci. (Paris)* **26**, 133–135 (2010).
- Hamdan, F.F. *et al.* *De novo SYNGAP1* mutations in nonsyndromic intellectual disability and autism. *Biol. Psychiatry* **69**, 898–901 (2011).
- Hamdan, F.F. *et al.* Excess of *de novo* deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am. J. Hum. Genet.* **88**, 306–316 (2011).
- Berryer, M.H. *et al.* Mutations in *SYNGAP1* cause intellectual disability, autism and a specific form of epilepsy by inducing haploinsufficiency. *Hum. Mutat.* **34**, 385–394 (2013).
- Vissers, L.E. *et al.* A *de novo* paradigm for mental retardation. *Nat. Genet.* **42**, 1109–1112 (2010).
- de Ligt, J. *et al.* Diagnostic exome sequencing in persons with severe intellectual disability. *N. Engl. J. Med.* **367**, 1921–1929 (2012).
- 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* **158A**, 1633–1640 (2012).
- Need, A.C. *et al.* Clinical application of exome sequencing in undiagnosed genetic conditions. *J. Med. Genet.* **49**, 353–361 (2012).
- Shen, J. *et al.* Mutations in *PNKP* cause microcephaly, seizures and defects in DNA repair. *Nat. Genet.* **42**, 245–249 (2010).
- Tsai, M.H. *et al.* Clinical genetic study of the epilepsy-aphasia spectrum. *Epilepsia* **54**, 280–287 (2013).
- Escayg, A. *et al.* Mutations of *SCN1A*, encoding a neuronal sodium channel, in two families with GEFS+2. *Nat. Genet.* **24**, 343–345 (2000).
- Oliva, M., Berkovic, S.F. & Petrou, S. Sodium channels and the neurobiology of epilepsy. *Epilepsia* **53**, 1849–1859 (2012).
- Ogiwara, I. *et al.* *De novo* mutations of voltage-gated sodium channel α II gene *SCN2A* in intractable epilepsies. *Neurology* **73**, 1046–1053 (2009).
- Kamiya, K. *et al.* A nonsense mutation of the sodium channel gene *SCN2A* in a patient with intractable epilepsy and mental decline. *J. Neurosci.* **24**, 2690–2698 (2004).
- Veeramah, K.R. *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.* **90**, 502–510 (2012).
- Weckhuysen, S. *et al.* KCNQ2 encephalopathy: emerging phenotype of a neonatal epileptic encephalopathy. *Ann. Neurol.* **71**, 15–25 (2012).
- Fujiwara, T. *et al.* Mutations of sodium channel α subunit type 1 (*SCN1A*) in intractable childhood epilepsies with frequent generalized tonic-clonic seizures. *Brain* **126**, 531–546 (2003).
- 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* **149A**, 343–348 (2009).
- Mills, P.B. *et al.* Neonatal epileptic encephalopathy caused by mutations in the *PNPO* gene encoding pyridox(am)ine 5'-phosphate oxidase. *Hum. Mol. Genet.* **14**, 1077–1086 (2005).

ONLINE METHODS

Subjects. This study was approved by the Human Research Ethics Committees of Austin Health and the University of Washington. Proband with epileptic encephalopathies were recruited after obtaining informed consent from the epilepsy clinic at Austin Health and the practices of the investigators and by referral for epilepsy genetics research from around Australia and internationally. The cohort consisted of 500 individuals with a diverse range of epileptic encephalopathy phenotypes. An epileptic encephalopathy was defined by refractory seizures and cognitive slowing or regression associated with frequent, ongoing epileptiform activity¹. Detailed epilepsy and medical histories were obtained together with the results of investigations, including EEG and magnetic resonance imaging (MRI) studies. Epilepsy syndromes were classified according to the Organization of the International League Against Epilepsy Commission on Classification¹ (Table 1). Some subjects had already undergone mutation screening for specific epilepsy-related genes involved in encephalopathy; none had been screened for all genes with known involvement. Subjects with a previously identified disease-causing mutation were excluded from this study. Some subjects with known mutations were included as mutation-positive controls but were not included in the 500 cases in the discovery cohort. Furthermore, 369 of 500 cases had previously been screened for pathogenic CNVs, either in the research setting ($n = 257$)³ or by clinical testing ($n = 112$). Proband with pathogenic CNVs were not included in this study.

Gene selection. We selected 65 genes for sequence analysis (Supplementary 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 included 13 epilepsy-associated genes, more commonly implicated in individuals with non-epileptic encephalopathy forms of epilepsy (*CHRNA7*, *KCNQ3*, *GABRD*, *GABRG2*, *PRICKLE1*, *CACNB4* and *SCN1B*) or a related neurodevelopmental disorder with epilepsy as a comorbid feature (*GRIN2B*, *MBD5*, *MEF2C*, *SYNGAP1*, *SYN1* and *ATPIA2*)^{35–39}. None of these genes have mutations that 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 the Supplementary Note and Supplementary Table 2 for details on candidate gene selection).

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

Target capture and sequencing. We used molecular inversion probes (MIPs) to capture all exon and intron-exon boundaries (and 5-bp flanking sequences) of target genes (RefSeq, hg19 build) (Supplementary Table 3). Detailed methodology is described elsewhere¹¹. Briefly, pooled MIPs were used to capture target exons from 100 ng of DNA from each proband. PCR was performed using universal primers, with the introduction of unique 8-base barcodes on the tagged reverse primer. Pooled libraries were subjected to massively parallel sequencing using a 101 paired-end protocol on the HiSeq platform. Libraries were prepared and sequenced in 2 batches, comprising a total of 30 (batch 1) and 35 (batch 2) genes.

Data analysis and variant calling. Raw read data processing and mapping with Burrows-Wheeler Aligner (BWA; see URLs) was performed as described¹¹. SNV and indel calling and filtering were performed using the Genome Analysis Toolkit (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 < 25× and clustered variants (window size of 10).

Variants were annotated with SeattleSeq (see URLs), and the Exome Sequencing Project (ESP6500) data set (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 nonsynonymous, splice-site and frameshift variants were assessed further. The GATK Depth of Coverage tool was used to calculate overall depth of coverage for each sample with a threshold of 25×, as well as the mean percentage (across all samples) of bases covered at >25× for each gene.

Rare variant segregation analysis. Where DNA from family members was available, segregation analysis was carried out for all rare (not present in ESP6500 controls) and possibly damaging (nonsynonymous, essential splice-site and frameshift) variants for all 65 target genes. This analysis was performed using a MIP-pick strategy. We selected and pooled only the MIPs that captured the genomic sequence harboring the rare variant of interest and performed target-enrichment PCR and sequencing as described for all relevant probands and family members. This approach allowed us to sequence variants at very high depth and to 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 new pathogenic variant. Pathogenic variants had to (i) arise *de novo*, (ii) segregate with the disorder, (iii) be inherited from a parent with somatic mosaicism or (iv) adhere 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 mechanism for a rare variant owing to the unavailability of DNA from one or both parents. It is probable that a subset of these variants also causes disease, although we report here as likely pathogenic only those variants that probably lead to protein truncation (splice-site, nonsense, frameshift and stop-loss changes). Additionally, three missense mutations in known genes (*STXBPI* and *SCN2A*) were interpreted to be likely pathogenic on the basis of the high incidence of pathogenic missense mutations in these genes, and the interpretation was further supported by the lack of the variant in the parent from whom DNA was available. 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.

32. Heron, S.E. *et al.* Missense mutations in the sodium-gated potassium channel gene *KCNT1* cause severe autosomal dominant nocturnal frontal lobe epilepsy. *Nat. Genet.* **44**, 1188–1190 (2012).
33. Marco, E.J. *et al.* *ARHGEF9* disruption in a female patient is associated with X linked mental retardation and sensory hyperarousal. *BMJ Case Rep.* **2009**, bcr06.2009.1999 (2009) Epub 2009 Jul 2.
34. Mastrangelo, M. & Leuzzi, V. Genes of early-onset epileptic encephalopathies: from genotype to phenotype. *Pediatr. Neurol.* **46**, 24–31 (2012).
35. Tao, H. *et al.* Mutations in prickle orthologs cause seizures in flies, mice, and humans. *Am. J. Hum. Genet.* **88**, 138–149 (2011).
36. Endeley, S. *et al.* Mutations in *GRIN2A* and *GRIN2B* encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat. Genet.* **42**, 1021–1026 (2010).
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.* **83**, 345–351 (2013).
38. Talkowski, M.E. *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.* **89**, 551–563 (2011).
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.* **31**, 722–733 (2010).
40. Mefford, H.C. *et al.* Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. *PLoS Genet.* **6**, e1000962 (2010).