Author/s:
Bagnall, RD;Crompton, DE;Petrovski, S;Lam, L;Cutmore, C;Garry, SI;Sadleir, LG;Dibbens, LM;Cairns, A;Kivity, S;Afawi, Z;Regan, BM;Duflou, J;Berkovic, SF;Scheffer, IE;Semsarian, C

Title:
Exome-based analysis of cardiac arrhythmia, respiratory control, and epilepsy genes in sudden unexpected death in epilepsy

Date:
2016-04-01

Citation:
Bagnall, R. D., Crompton, D. E., Petrovski, S., Lam, L., Cutmore, C., Garry, S. I., Sadleir, L. G., Dibbens, L. M., Cairns, A., Kivity, S., Afawi, Z., Regan, B. M., Duflou, J., Berkovic, S. F., Scheffer, I. E. & Semsarian, C. (2016). Exome-based analysis of cardiac arrhythmia, respiratory control, and epilepsy genes in sudden unexpected death in epilepsy. Annals of Neurology, 79 (4), pp.522-534. https://doi.org/10.1002/ana.24596.

Exome-based analysis of cardiac arrhythmia, respiratory control and epilepsy genes in sudden unexpected death in epilepsy

Richard D. Bagnall[1, 2], Douglas E. Crompton[3, 4], Slavé Petrovski, [4,5] Lien Lam[1], Carina Cutmore[1], Sarah I. Garry[4], Lynette G. Sadleir[6], Leanne M. Dibbens[7], Anita Cairns[8], Sara Kivity[9], Zaid Afawi[10], Brigid M. Regan[4], Johan Duflou[2, 11], Samuel F. Berkovic[4], Ingrid E. Scheffer[4, 12, 13, 14], Christopher Semsarian[1, 2, 15]

[1] Agnes Ginges Centre for Molecular Cardiology, Centenary Institute, Sydney, Australia

[2] Sydney Medical School, University of Sydney, Sydney, Australia

[3] Neurology Department, Northern Health, Melbourne, Australia

[4] Epilepsy Research Centre, Department of Medicine, University of Melbourne, Austin Health, Melbourne, Australia

[5] Institute for Genomic Medicine, Columbia University, New York, United States of America

[6] Department of Paediatrics and Child Health, School of Medicine and Health Sciences, University of Otago, Wellington, New Zealand

[7] Epilepsy Research Program, School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, SA, Australia

[8] Neurosciences Department, Lady Cilento Children's Hospital, Brisbane, Australia

[9] Epilepsy Unit, Schneider Children's Medical Center of Israel, Petach Tikvah, Israel

[10] Tel-Aviv Sourasky Medical Center, Tel Aviv, Israel

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version record. Please cite this article as doi:10.1002/ana.24596.

[11] Department of Forensic Medicine, Sydney, Australia

[12] Department of Neurology, The Royal Children's Hospital, Parkville, Melbourne, Victoria, Australia

[13] Florey Institute of Neurosciences and Mental Health, Melbourne, Australia

[14] Department of Paediatrics, University of Melbourne, Royal Children's Hospital, Melbourne, Australia

[15] Department of Cardiology, Royal Prince Alfred Hospital, Sydney, Australia

Word Count: 4968 (excluding refs)

Title character count: 121

Running title (34 characters): Exome sequencing analysis of SUDEP

Address for Correspondence:

Professor Christopher Semsarian

Agnes Ginges Centre for Molecular Cardiology

Centenary Institute

Locked Bag 6, Newtown, NSW, 2042, Australia

Email: c.semsarian@centenary.org.au

ABSTRACT

Objective The leading cause of epilepsy-related premature mortality is sudden unexpected death in epilepsy (SUDEP). The cause of SUDEP remains unknown. To search for genetic risk factors in SUDEP cases, we performed an exome-based analysis of rare variants.

Methods Demographic and clinical information of 61 SUDEP cases were collected. Exome sequencing and rare variant collapsing analysis with 2,936 control exomes were performed to test for genes enriched with damaging variants. Additionally, cardiac arrhythmia, respiratory control and epilepsy genes were screened for variants with frequency of <0.1% and predicted to be pathogenic with multiple *in silico* tools.

Results: The 61 SUDEP cases were categorised as 'definite SUDEP' (n= 54), 'probable SUDEP' (n=5), and 'definite SUDEP plus' (n=2). We identified *de novo* mutations, previously reported pathogenic mutations, or candidate pathogenic variants in 28/61 (46%) cases. Four SUDEP cases (7%) had mutations in common genes responsible for the cardiac arrhythmia disease, long QT syndrome. Nine cases (15%) had candidate pathogenic variants in dominant cardiac arrhythmia genes. Fifteen cases (25%) had mutations or candidate pathogenic variants in dominant epilepsy genes. No gene reached genome-wide significance with rare variant collapsing analysis, however, *DEPDC5* (p=0·00015) and *KCNH2* (p=0·0037) were among the top 30 genes, genome-wide.

Interpretation: A sizeable proportion of SUDEP cases have clinically relevant mutations in cardiac arrhythmia and epilepsy genes. In cases with a long QT syndrome

gene mutation, SUDEP may occur due to a predictable and preventable cause. Understanding the genetic basis of SUDEP may inform cascade testing of at-risk family members.

INTRODUCTION

People with epilepsy have an increased risk of sudden death and the most common cause of epilepsy-related premature mortality is sudden unexpected death in epilepsy (SUDEP).[1] SUDEP is characterised by a sudden, unexpected, witnessed or unwitnessed, non-traumatic and non-drowning death in people with epilepsy, with or without evidence for a seizure and excluding documented status epilepticus, and in whom post-mortem examination does not reveal a cause of death.[2] In children with epilepsy who were followed for 40 years, 24% died, with 38% of deaths due to SUDEP.[3] Recent analysis of several population-based studies of SUDEP showed that the incidence is highest in the third and fourth decades of life.[4] When comparing potential years of life lost from SUDEP with other neurologic conditions, SUDEP ranks second only to stroke in the United States.[4] Some reported risk factors for SUDEP include chronic uncontrolled epilepsy, generalised tonic-clonic seizures, young age at seizure onset, and long duration of epilepsy.[5] Despite increasing appreciation of the burden and risk factors of SUDEP by patients, doctors, and the global community, the underlying causes remain unknown.

A commonly suggested mode of death in SUDEP is cardiac arrhythmia. The inherited cardiac arrhythmias, such as long QT syndrome (LQTS) and catecholaminergic polymorphic ventricular tachycardia (CPVT) are known to cause syncope, arrhythmias, and sudden unexplained death.[6,7] Cardiac rhythm abnormalities recorded during epileptic seizures include prolongation of the QT interval, bradycardia, and tachycardia.[8,9] Transgenic mice with a point mutation in the most common LQTS gene, *KCNQ1*, have cardiac arrhythmias and epileptic seizures, with sudden death recorded

during electrocardiogram and electroencephalogram (EEG) monitoring.[10] Post-mortem genetic testing of three genes that account for 70% of all LQTS cases (*KCNH2*, *KCNQ1*, and *SCN5A*), is important in autopsy-negative sudden unexpected death.[6,7] Our post-mortem genetic studies of the three common LQTS genes in a retrospective coronial series of SUDEP cases found rare variants associated with LQTS.[11] Another proposed factor contributing to SUDEP is respiratory dysfunction. Inpatient video EEG monitoring of patients with SUDEP showed that terminal seizures are followed by respiratory and cardiac abnormalities with terminal apnea and cardiac arrest.[12] Thus, mechanisms altering cardiac or respiratory function, or both, may underpin SUDEP. Some genetic epilepsies may have an increased risk of SUDEP. In >80% of patients, Dravet Syndrome is caused by mutations in *SCN1A*, and there is a high mortality rate with about half of deaths due to SUDEP.[13] *KCNA1* and *SCN8A* have been suggested as SUDEP genes from mouse and human studies.[14-17] The contribution of epilepsy gene mutations to SUDEP risk in an unselected series of cases is unknown.

We performed comprehensive exome sequencing based analysis of single nucleotide variants (SNVs), small insertions and deletions (indels), and copy number variants (CNVs) in 61 SUDEP cases. We performed a genome-wide gene collapsing analysis of rare variant enrichment, and a focused screen of cardiac arrhythmia, respiratory control, and epilepsy genes.

SUBJECTS AND METHODS

Subjects

Sixty-one SUDEP cases were recruited from three sources: 27 (44%) had participated in the epilepsy genetics research program in Melbourne, Australia, during life and had SUDEP on follow-up; 15 (25%) prospective coronial SUDEP cases were collected from 2010 to 2012 by the Departments of Forensic Medicine (DOFM) in New South Wales, Victoria, Queensland, and South Australia; and 19 (31%) retrospective coronial SUDEP cases were collected from a review of autopsy reports over a 17-year period from 1993 to 2010 at the DOFM in Sydney. The latter group are part of a larger SUDEP cohort previously reported,[11] and were included in the current study following an additional review of autopsy and neuropathology reports by a neurologist.

SUDEP cases were classified as definite SUDEP, definite SUDEP plus, probable SUDEP, or possible SUDEP.[2] The Human Research Ethics Committees of Austin Health, Tel Aviv Sourasky Medical Centre, and Royal Prince Alfred Hospital approved the study. For patients recruited through the Epilepsy Research Centre, Melbourne, all patients, or their parent, next-of-kin, or legal guardian in the case of children or patients with intellectual disability, gave informed consent for epilepsy genetics research during life of the patient. The senior next of kin gave further consent for ongoing research after SUDEP occurred. Informed consent was not required for the de-identified retrospective coronial SUDEP cases.

For gene collapsing analysis, we compared SUDEP cases to 2,936 European ancestry control exomes, sequenced for various studies at the Institute for Genomic Medicine, Columbia University – formerly, the Center for Human Genome Variation (CHGV), Duke University. The control samples were not recruited for sudden unexplained death, epilepsy or other neuropsychiatric or neurodevelopmental disorder studies. Cryptic relatedness testing confirmed no duplicates, first or second degree related pairs.

Methods

An overview of all gene screens is shown in Figure 1. Whole blood for DNA extraction was collected at autopsy and stored at -20°C, or collected during life of the patient. DNA was isolated using a QIAmp DNA blood mini kit (Qiagen, Limburg, NL). Exome sequencing was performed as described previously (Macrogen Inc, Seoul, ROK).[18] SNVs and indels were annotated using SeattleSeq Annotation tool v8.07 (http://snp.gs.washington.edu/SeattleSeqAnnotation137/index.jsp) and compared against the 1000 Genomes Project, phase 3, data (http://www.1000genomes.org/), the NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [Nov, 2013 release], the Exome Aggregate Consortium data [r0.3] (http://exac.broadinstitute.org/), and in-house exome sequences of 118 unrelated individuals.

Focused variant filtration and prioritisation

Target genes for a focused assessment of sequence variants were defined by a review of recent literature and searching the OMIM database [Nov 01, 2014] for diseases

containing the terms 'epilepsy', 'epileptic encephalopathy', 'long QT', 'short QT', 'atrial fibrillation', 'Brugada syndrome', 'heart block', 'catecholaminergic polymorphic ventricular tachycardia' or 'central hypoventilation syndrome' (Supplementary Table 1). The genes include *KCNA1* and *KCNQ1*, which have been suggested as SUDEP genes from mice studies.[10,14] We searched for SNVs and indels of target genes causing a change in protein sequence and with a frequency of <0.1% across all ethnic sub-populations of the 1000 Genomes Project data, or in 6,503 exomes of the ESP data. Candidate pathogenic variants were variants at canonical splice site dinucleotides, nonsense variants and indels; and missense variants predicted to be probably damaging or possibly damaging by at least 2 of 3 *in silico* pathogenicity prediction tools (SIFT, PolyPhen-2 and Mutation Taster).[19-21]

Gene-based Collapsing Test

To assess enrichment of rare qualifying variants in the SUDEP cases, we used a similar gene-based collapsing test applied to a recent amyotrophic lateral sclerosis (ALS) sequencing paper.[22] The exomes of 57 SUDEP cases of European ancestry were compared to 2,936 control exomes with matching European ancestry assessed using the first five principal components from EIGENSTRAT.[23] Whole-exome sequence data from the 57 SUDEP cases and 2,936 controls were processed using the same bioinformatics pipeline and were found to have at least 10-fold read coverage for 91.4% and 94.9% of the 33.3 Mbps of consensus coding sequence (CCDS, release 14), respectively. To alleviate the confounding issues with differences in exon coverage, for each exon in each sample, we determined the percentage of CCDS sites with at least 10-fold coverage. An exon was excluded from the analysis if there was greater

than 22% difference in the average percentage of sites covered with at least 10-fold coverage between the case and control populations. This exon-pruning resulted in 8.3% of CCDS bases being excluded, thus, the gene collapsing tests were performed on the pruned 30.5 Mbps of CCDS.

Variants were required to have: i) at least 10-fold coverage, ii) a quality score (QUAL) of at least 50, iii) a genotype quality (GQ) score of at least 20, iv) a quality by depth (QD) score of at least 2, v) a mapping quality (MQ) score of at least 40, vi) a read position rank sum (RPRS) score greater than -8 and vii) mapping quality rank sum (MQRS) score greater than -12. Indels were also required to have a maximum strand bias (FS) of 200. Variants were further screened according to VQSR tranche calculated using the known SNV sites from HapMap v3.3, dbSNP, and the Omni chip array from the 1000 Genomes Project. To “PASS”, variants were required to be located in the consensus coding sequence (CCDS release 14) regions and achieve a tranche of 99.9% for exonic SNVs in genomes, 99% for SNVs in exomes and 95% for exonic indels in genomes. Finally, variants were excluded if they were among a predefined list of known sequencing artifacts or if they were marked by EVS as being problematic variants. For the collapsing analyses all variants were annotated to Ensembl 73 using SnpEff.

To search for genes that confer risk of SUDEP, we implemented a genic collapsing test. For each CCDS gene, we indicate when a subject has no qualifying variants in the gene (0) or when it has one or more qualifying variants in the gene (1). Thus, for each gene, each subject was indicated as carrying or not carrying a “qualifying” variant. At the allele frequency level, a qualifying variant was defined as a variant with a minor allele frequency (MAF) cut-off of less than 0.05% among the remaining 2,992 case plus

control subjects, EVS European and African-American controls, and ExAC's six ethnically stratified populations.[24] At the variant effect prediction level, qualifying variants were defined as all non-synonymous, canonical splice and frameshift variants, except those missense variants predicted by PolyPhen-2 HumDiv[25] to be benign. To identify qualifying variants we used Analysis Tools for Annotated Variants (http://redmine.igm.cumc.columbia.edu/projects/atav/news).

Once qualifying variants were collected for all genes, a two-tailed Fisher's exact test was performed for each gene comparing the rates of qualifying variants observed among the cases to the control population. Only genes with >1 qualifying variant among the combined case-control population were tested as assessable genes. The adjusted alpha after correcting for the number of assessable genes tested was $p<3.27x10^{-6}$

Mitochondrial genome sequencing and analysis

Mitochondrial genome sequences were reconstructed from off-target exome sequence reads using GSNAP.[26] Mitochondrial genomes having <90% sequence read coverage were excluded. Mitochondrial SNVs were annotated using Mitotool v1.1.2.[27] Non-synonymous variants were compared against a database of mitochondrial DNA variants collated from 26,850 human mitochondrial genomes (http://www.mitomap.org/MITOMAP [Nov 06, 2014]).

Copy number variant analysis

CNVs were searched for using eXome Hidden Markov Model (XHMM) software[28] in 41 SUDEP cases and 56 exomes from individuals with cardiomyopathy to normalise read

depth, and all were processed using the Illumina TruSeq exome enrichment method. The remaining 20 SUDEP cases, processed using the Illumina NextEra exome enrichment method, were not assessed for CNVs as comparable control exomes were not available. In brief, XHMM software uses principal component analysis (PCA) to normalize exome read depth across multiple samples, and a hidden Markov model to detect and genotype copy number variation from exome sequencing experiments. PCA on a sample versus target-depth matrix finds the main modes in which depth varies across multiple samples and targets, and the largest of such effects are removed. Exome depth of coverage was calculated using GATK and exome targets with (i) length <10 bp or >10 kb, (ii) GC content >90% or <10%, or (iii) >10% of bases masked by RepeatMasker were excluded. Common CNVs and CNVs overlapping in more than 5% of all samples were removed. We retained CNVs with a quality score ≥60, and spanning ≥3 exome targets, with length ≥1 kb, and overlapping target genes or loci previously associated with epilepsy (1q21, 15q11.2, 15q13.3, 16p11.2, 16p13.11 and 22q11.2).[29-32]

Variant validation and co-segregation analysis

Genomic regions up to 500 bp surrounding variants for validation were PCR amplified and Sanger DNA sequenced (Macrogen). Sequencing electropherograms were manually inspected using Sequencher v5.1 (Gene Codes Corp, MI, USA).

RESULTS

The demographic details of the SUDEP cohort (n=61) are summarised in Table 1 and Table 2. SUDEP was categorised as '*definite SUDEP*' in 54 (89%) cases, '*probable SUDEP*' in 5 (8%) cases without an autopsy report, and '*definite SUDEP plus*' in 2 (3%) cases: EP15, who had a 75% stenosis of the left anterior descending coronary artery, and EP67 who had pneumonia. The role of these co-morbidities in the cause of death is difficult to determine, but were not felt severe enough to be the sole cause of death; therefore both cases were included in all analyses.

The mean age at epilepsy onset was 10·3 ± 8·2 years (range 0-34 years) and the mean age at SUDEP was 28·1 ± 12·0 years (range 1-53 years). Epilepsy phenotyping of SUDEP cases during life in the Melbourne Epilepsy Research Centre cases revealed a range of epilepsies including temporal lobe epilepsy, epileptic encephalopathies, and genetic generalised epilepsies such as juvenile myoclonic epilepsy. Diagnoses of specific epilepsy syndromes were not available for the coronial-recruited cases due to lack of prospective electro-clinical data. The diagnosis of epilepsy and death by SUDEP were made using available information at the time of autopsy by the forensic pathologist. We performed a rigorous re-review of available information; ethical considerations did not allow contact with the families. No SUDEP case had a diagnosis of cardiac disease. ECGs were available in six cases (EP19, EP35, EP39, EP41, EP66 and EP67). All showed sinus rhythm. The corrected QT interval was normal in five cases and borderline in EP19 (459 ms). EP19 had refractory mesial temporal lobe epilepsy with hippocampal sclerosis and was seizure free for 13 years after anterior temporal lobectomy prior to seizure recurrence 3 years before

death. EP19 was found to carry a pathogenic variant in the LQT2 gene *KCNH2* (Table 3). SUDEP occurred in bed in 40 (66%) cases, with the position noted as prone in 27 and supine in one. SUDEP was witnessed in one case, EP10; a male aged 34 years who was found breathing very slowly on the floor next to his bed and could not be revived by the emergency services.

Exome sequencing of the 61 SUDEP cases yielded an average of 76,968,829 sequence reads per person, of which 57% mapped uniquely to the target exome (Supplementary Table 2). When considering only the intersect with CCDS release 14 protein-coding sequence, 86% of target regions were covered by 20 or more reads. Eigenvector-based predictions of ethnicity identified 57 (93%) European and four non-European patients. Analysis of exome variants did not reveal hidden relatedness among any pair of SUDEP cases and controls used in the gene collapsing analysis. Evidence for genes enriched with rare non-synonymous protein-coding variants was sought by comparing the 57 European SUDEP cases to 2,936 European controls. Frameshift and non-synonymous SNVs, including missense predicted to be possibly or probably damaging by PolyPhen-2, with a minor allele frequency <0·05% among the 2,992 remaining case plus control subjects, and located in 30·5 Mbp (91·7%) of the consensus coding sequence (CCDS release 14), were compared in SUDEP cases and controls. Supplementary Table 3 shows the top 30 genes with the greatest enrichment of variants, ranked by their Fisher's exact test two-tailed p-value. No gene reached genome-wide significance (adjusted alpha = $p<3.27\times10^{-6}$); however, among the top 30 genes of 15,305 assessable CCDS genes were the familial focal epilepsy gene, *DEPDC5* ($p=1.48\times10^{-4}$, ranked 2nd), the LQT2 gene, *KCNH2* (p=0·0037, ranked 24th),

and *NOS1AP* (p=2.17x10^{-4}, ranked 3rd), which has common variants that are associated with sudden cardiac death.[33] When comparing variants that appear only once in the entire case and control data set, (i.e. singleton variants), and that are either predicted to have a loss-of-function effect or be "probably damaging" by PolyPhen-2, the *DEPDC5* gene ranked first (p=1·5x10^{-6}), *NOS1AP* ranked 7th (p=0·0021), *SCN2A* ranked 13th (p=0·0051) and *KCNH2* ranked 18th (p=0·007) genome-wide (Supplementary Table 4).

In a focused review of known disease genes, we first screened 32 cardiac arrhythmia-related disease genes for rare variants with a frequency of <0.1% among over 60,000 publically available exomes and predicted to be damaging by multiple *in silico* pathogenicity prediction tools (Table 3). Among four (7%) SUDEP cases, we found a Sanger-verified *de novo* mutation or previously pathogenic-reported mutation among the three common LQTS genes (Table 3; Figure 2). These four mutations are regarded as highly likely to be pathogenic for LQTS. Three SUDEP cases with LQTS mutations were from the Melbourne Epilepsy Research Centre cohort, and one was from the coronial cohorts. The *KCNH2* Arg744* nonsense mutation in EP11 was previously reported in five unrelated patients with LQTS, and segregated with LQTS in seven affected members of one family.[34,35] The *KCNH2* Gly924Ala missense mutation in EP19, and the *KCNQ1* Tyr662* nonsense mutation in EP40 have been previously described in patients with LQTS.[35] A *de novo* Ile397Val *SCN5A* mutation was found in EP41. The Ile397 residue is located in a highly evolutionarily conserved transmembrane domain (Figure 2). Nine (15%) SUDEP cases had Sanger-verified *candidate* pathogenic variants in cardiac arrhythmia genes implicated with dominant

disorders (Table 3). One novel candidate pathogenic variant was found in the LQTS2 gene, *KCNH2* Gly749Ala in EP55, and one candidate pathogenic rare variant was found in the CPVT gene, *RYR2*, resulting in a Cys1489Arg change in EP14. The amino acid residue Val223 of *SCN5A* was substituted with a glycine in SUDEP case EP12 and with a leucine in a patient with Brugada syndrome.[36] EP43, who is of Ashkenazi Jewish descent, had three candidate pathogenic variants in cardiac arrhythmia genes: a novel Arg2607Gly variant in *AKAP9*, and rare Ser841Leu and Arg248Cys variants in *HCN4* and *TRPM4,* respectively.

The MORTEMUS study highlighted dysregulation of respiratory control preceding SUDEP.[12] We next analysed five genes with plausible roles in central control of ventilation (*ASCL1*, *BDNF*, *EDN3*, *GDNF,* and *RET*), however, no candidate variants with a general population frequency of <0.1% were found among our 61 SUDEP cases.

We finally focused on 72 epilepsy-related disease genes for previously reported mutations and candidate pathogenic variants. Among 15 (25%) of the SUDEP cases, we found 16 candidate pathogenic variants: four *de novo* mutations, two previously reported pathogenic mutations, and ten candidate pathogenic variants in epilepsy genes responsible for dominant disorders (Table 4). Nine of these 15 cases came from the Melbourne Epilepsy Research Centre cohort and six from the coronial cohorts (Fisher's exact test two-tailed *p*-value=0·23). Of the 72 epilepsy related genes, gene-collapsing analysis ranked *DEPDC5* the highest (p=1·48x10^{-4}). In *DEPDC5*, the Arg843* mutation found in EP09 is a recurrently reported truncation mutation predisposing to focal epilepsy with variable foci (FFEVF) in a French-Canadian

population.[37] Also found in *DEPDC5* were three other nonsense variants, Arg286*, Gln1016*, and Arg1332*, and two missense variants, Ser19Thr and Arg347His (Figure 3a). The *DEPDC5* nonsense variants were regarded as highly likely to be pathogenic. The *DEPDC5* Arg347His variant found in EP10 occurs in a protein domain of unknown function, which is highly evolutionarily conserved to *Anopheles gambiae* (mosquito) (Figure 3b). In *SCN2A*, a *de novo* Arg1882Gln mutation was reported in EP67 prior to her SUDEP,[38] and a *de novo* Asn976Lys mutation was found in EP73. The Asn976 residue of SCN2A occurs in a transmembrane domain that is conserved to *Anopheles gambiae.* A *de novo SCN1A* mutation (Gly1480Val) was reported in EP37 prior to his SUDEP [39], a *de novo* Tyr182Phe *GABRB3* mutation was found in a patient EP29 who has an epileptic encephalopathy, and a rare candidate pathogenic variant, Arg96Gln in *SCN1B* was found in EP51. We also found a novel Ala306Val variant in a highly conserved transmembrane domain of the potassium channel gene *KCNQ2,* as previously reported.[38] The Gly162Ser mutation in *PAFAH1B1* (*LIS1*), found in EP38, was also reported in an unrelated individual with focal epilepsy and occipito-parietal pachygyria.[40]

Variant co-segregation analysis was possible in six SUDEP cases where parental DNA was available (EP23, EP29, EP37, EP41, EP67 and EP73). Genotyping of the proband and available parents revealed *de novo* mutations arising in the proband in 5/6 cases; *GABRB3* Tyr182Phe in EP29, *SCN1A* Gly1480Val in EP37 as previously reported in this patient,[39] *SCN2A* Asn976Lys in EP73 and Arg1882Gln in EP67, as previously reported in EP67,[38] and *SCN5A* Ile397Val in EP41. All five *de novo* mutations are absent among over 60,000 individuals of the Exome Aggregate Consortium data. A

DEPDC5 nonsense variant Arg1332* in EP23 was inherited from his unaffected father.

CNVs make a significant contribution to the genetic basis of epilepsy.[32,41] In 41 SUDEP cases, 136 putative CNVs were identified with strict quality filters. Three CNVs overlapped with loci previously implicated in epilepsy; a 15q11·2 deletion in EP23 and EP43, and a 16p11·2 duplication in EP06 (Supplementary Table 5). The 15q11·2 deletions do not encompass target genes and overlap with CNVs found in controls, whereas the 16p11·2 duplication is associated with typical and atypical rolandic epilepsies,[32] and spans a ~575 kb region containing 27 genes (Figure 4). EP06 was ascertained as a de-identified coronial case and details of their epilepsy syndrome were not available.

Mitochondrial genome mutations are known to sometimes cause epilepsy and abnormal cardiac rhythms.[42,43] The sequence data from EP09 and EP24 covered only 70.1% and 89.7% of the mitochondrial genome and were excluded from further analysis. On average, 63% of the mitochondrial protein-coding regions were covered with at least 10 sequence reads. While of unknown significance to the patient's clinical characteristics, three novel Sanger-verified non-synonymous mitochondrial genome variants were identified across 59 SUDEP cases; *ND1* Ile13Val in EP22, *ATP8* His3Tyr in EP01, and *ND6* Val10Met in EP29.

DISCUSSION

Our study of exome sequencing in a large series of SUDEP cases has revealed a surprising number of pathogenic or likely pathogenic variants. Mutations known to cause long QT syndrome were found in 7% of cases and a further 15% had candidate variants in genes potentially predisposing to malignant cardiac arrhythmia. Finally, 25% had mutations in known epilepsy genes and *DEPDC5* stood out with variants identified in six cases, suggesting further work is required to understand whether patients with epilepsy due to *DEPDC5* may be at heightened risk of SUDEP. These findings provide key insights into the aetiology and mechanisms likely to increase risk of SUDEP.

LQTS is an inherited arrhythmia disorder known to cause sudden cardiac death. In the common LQTS genes, we found four known or *de novo* mutations and two novel candidate pathogenic variants. In one SUDEP patient with a family history of epilepsy, we found a *de novo SCN5A* missense mutation in the highly conserved sodium channel transmembrane domain. Patients with mutations in *SCN5A* transmembrane regions are at higher risk of sudden death, particularly during sleep, even those with a normal QT interval.[44] There were two mutations and one candidate pathogenic variant in the LQTS2 gene, *KCNH2*, including one nonsense mutation reported in multiple LQTS patients and a family.[34,35] We found one nonsense mutation in the LQTS1 gene *KCNQ1*, which has also been suggested as a SUDEP gene in mice studies.[10] This suggests that, in some cases, SUDEP may result from an unfortunate combination of epilepsy and LQTS genetic determinants. Transcripts from all three of the major LQTS genes (*KCNQ1*, *KCNH2* and *SCN5A*) have been detected in human brain. *KCNQ1* transcripts are found in the adult human brain, and kcnq1 protein in the mouse is found

in pyramidal neurons in CA1 to CA3, granule cells of the dentate gyrus, and hilar interneurons.[10] Full-length *KCNH2* transcripts, and a primate-specific brain isoform, are found in human hippocampus, and KCNH2 protein expression was confirmed in human hippocampus and frontal cortex using western analysis.[45] *SCN5A* transcripts are found in human brain, and scn5a protein expression in the rat is detected in the ventral medial, dorsal medial and posterior hypothalamic nuclei.[46] The finding of LQTS mutations in SUDEP cases has important implications for the surviving family members, both with and without epilepsy, who risk inheriting the mutation, and therefore being at increased risk of arrhythmias and sudden death.[47] Patients with epilepsy who have LQTS mutations may have a heightened risk of SUDEP. If an LQTS mutation is found in a patient with epilepsy during life, there may be significant clinical implications, including avoidance of medications that may prolong the QT interval, selection of antiarrhythmic drugs such as beta-blockers, and interventions for potentially lethal cardiac arrhythmias, e.g. implantable cardioverter defibrillator and pacemaker therapy.

The MORTEMUS study identified respiratory dysfunction preceding SUDEP.[12] We did not find any rare variants (classified as variants with a minor allele frequency of <0.1%) in five genes potentially involved in congenital central hypoventilation syndrome (CCHS), although it is likely that additional candidate genes for respiratory dysfunction exist. Similarly, we previously did not find *PHOX2B* mutations causing CCHS in a larger SUDEP cohort.[48] SUDEP patients of the MORTEMUS study represent a selected refractory group undergoing presurgical evaluation, with the added aspect that the dose of their antiepileptic drugs had been reduced by more than 50%, or completely withdrawn. In ours as in other cohorts, SUDEP victims were more likely to

be found in a prone position.[49] The underpinnings of respiratory dysfunction following terminal seizures in SUDEP are likely to be a complex interplay of genetic and non-genetic factors.

Our focused analysis of cardiac arrhythmia, respiratory control, and epilepsy genes was restricted to rare variants predicted to be pathogenic by multiple *in silico* tools. These criteria may have overlooked additional clinically relevant variants in the sequence data. Among the top genes highlighted by gene collapsing analysis was *NOS1AP* encoding the nitric oxide synthase 1 (neuronal) adaptor protein. It is well established that common variants in *NOS1AP* are associated with sudden death, drug-induced QT prolongation, and regulation of the QT interval.[33,50] Any association between rare functional *NOS1AP* variants and SUDEP risk would be a novel finding that requires larger collections of these special cohorts.

Genome-wide unbiased analysis of 15,305 CCD genes for enrichment of rare damaging variants in 57 SUDEP cases did not securely implicate any individual gene with SUDEP. However, the high ranks we observe for known LQTS and epilepsy genes suggest that there is a considerable genetic contribution from rare functional variation to these two clinical ascertainments. Gene collapsing analysis ranked the LQTS gene, *KCNH2,* among the top 30 genes, and our data suggest that, assuming identical mutation prevalence, an attainable cohort of less than 200 SUDEP cases may be sufficient for *KCNH2* to achieve genome-wide statistical significance. Only a minority of the top ranked genes in our gene collapsing analysis have been implicated in SUDEP or epilepsy. With larger collections of cases, such collapsing analyses could

provide the most promising avenue to identify novel SUDEP-associated genes.

Our gene collapsing analysis and focused study of epilepsy genes suggest that *DEPDC5* mutations may contribute to an increased risk of SUDEP. We found one previously described pathogenic mutation and five candidate pathogenic variants in *DEPDC5,* encoding dishevelled, Egl-10 and plekstrin domain containing protein 5, which causes the FFEVF syndrome and is an important gene in both non-lesional and lesional focal epilepsy.[37,51] Four (7%) SUDEP cases had *DEPDC5* nonsense variants. The recurrent *DEPDC5* Arg843* mutation was found in a de-identified coronial case, EP09, whose neuropathological examination showed microdysgenesis in the right medial temporal region and mild bilateral amygdala gliosis, and microglial activation. An Arg1332* nonsense variant was found in EP23, who had structural fronto-temporal epilepsy.[52] He underwent resective surgery for focal cortical dysplasia type IIa with significant seizure improvement prior to SUDEP. He had a strong family history of epilepsy through his maternal grandmother; however, the *DEPDC5* nonsense variant was inherited from his father. A Gln1016* variant was found in EP64, who had focal epilepsy and a strong family history of epilepsy. Two *DEPDC5* missense variants were found in two patients with limited clinical details. In one male with a rare *DEPDC5* Arg347His variant, epilepsy reportedly developed following head trauma. *DEPDC5* is a negative regulator of the mammalian target of rapamycin complex 1 (mTORC1), and nonsense mutations in *DEPDC5* lead to increased mTORC1 activity.[53] mTOR signalling regulates diverse processes including cell growth, metabolism and homeostasis. Future studies comparing larger cohorts of SUDEP cases to living patients with epilepsy will determine whether mutations in *DEPDC5* influence SUDEP

risk, or merely reflect that *DEPDC5* is a common cause of focal epilepsy.[26]

Mutations in sodium channels are highly associated with epilepsy.[54] We found three *de novo* mutations and one rare missense variant in sodium channel genes in four cases. Mutations of *SCN1A*, encoding the sodium channel alpha 1 subunit, are found in >80% of patients with Dravet syndrome in whom there is a high risk of SUDEP.[13] *SCN1A* mutations are much less common in the syndrome of epilepsy with myoclonic-atonic seizures.[55] Patient EP37 with the syndrome of epilepsy with myoclonic-atonic seizures had a *de novo SCN1A* mutation. The sodium channel alpha 1 subunit is expressed in brain, and at a lower level in the sino-atrial node, which regulates heart rate. Patients with Dravet syndrome have reduced heart rate variability,[56] and this may increase susceptibility to cardiac conduction disease and SUDEP. Similarly, changes in autonomic tone have been shown to influence RR and QT interval variability leading to onset of ventricular arrhythmias, as recently demonstrated in LQTS1 patients.[57] One SUDEP case had a missense mutation in *SCN1B,* encoding the sodium channel beta 1 subunit. *SCN1B* mutations have been described in patients with genetic epilepsy with febrile seizures plus, and are a rare cause of Brugada syndrome.[58,59] Similar to *SCN1A*, *SCN1B* is expressed in brain and heart.[60] *Scn1b* null mice recapitulate features of Dravet syndrome, including premature death,[61] and QT interval prolongation.[60] *De novo SCN2A* mutations are associated with epileptic encephalopathies and we found two *SCN2A de novo* mutations among our 61 SUDEP cases. We speculate that mutations in sodium channels co-expressed in the brain and heart may contribute to neuronal and cardiac dysfunction in SUDEP.

CONCLUSIONS

SUDEP is a tragic event and failure to identify a cause of death has major clinical, emotional, and psychological effects on the surviving family. Based on comprehensive exome sequencing analysis, a high proportion of SUDEP ascertained cases were found to have clinically relevant variants in cardiac arrhythmia and epilepsy genes, including four cases with LQTS mutations. These findings have relevance to understanding the causes of SUDEP and evaluation of at-risk family members. Moreover, SUDEP in patients with LQTS mutations may be predictable and preventable.

ACKNOWLEDGMENTS

We thank the patients and their families for participating in our research program, and the medical staff around Australia and New Zealand who provide their care. Dr Katherine B. Howell and Dr Mark Newton assisted with SUDEP case collection. CS and IES are the recipients of National Health and Medical Research Council (NHMRC) Practitioner Fellowships (#1059156 and #1006110). This work is supported by project grants from NHMRC (#1046441), Health Research Council of New Zealand, and CURE, and an NHMRC program grant (#628952). Acknowledgement of individuals for the contributions of control samples is provided in the supplementary information.

AUTHOR CONTRIBUTIONS

RDB, DEC, SP, SFB, IES and CS contributed to the concept and study design. RDB, DEC, SP, LL, CC, SIG, LGS, LMD, AC, SK, ZA, BMR, JD, SFB, IES and CS

contributed to the data acquisition and analysis. RDB, DEC, SP, LL, SFB, IES and CS contributed to drafting the manuscript and figures.

CONFLICTS OF INTEREST

Dr. Dibbens, Dr Scheffer and Dr. Berkovic hold a patent, international publication number WO2014/110628, which is directed to the identification of mutations in *DEPDC5*, Egl-10 or Dishevelled genes or their products to identify and treat specific causes of epilepsy more effectively. Dr Scheffer and Dr. Berkovic are each one of the inventors listed on a patent held by Bionomics Inc on methods of treatment and diagnosis of epilepsy by detecting mutations in the SCN1A gene, international publication number WO2006/133508.

REFERENCES

1. Neligan A, Bell GS, Johnson AL, Goodridge DM, Shorvon SD, Sander JW. The long-term risk of premature mortality in people with epilepsy. *Brain* 2011; **134**(Pt 2): 388-95.
2. Nashef L, So EL, Ryvlin P, Tomson T. Unifying the definitions of sudden unexpected death in epilepsy. *Epilepsia* 2012; **53**(2): 227-33.
3. Sillanpaa M, Shinnar S. Long-term mortality in childhood-onset epilepsy. *N Engl J Med* 2010; **363**(26): 2522-9.
4. Thurman DJ, Hesdorffer DC, French JA. Sudden unexpected death in epilepsy: Assessing the public health burden. *Epilepsia* 2014.
5. Hesdorffer DC, Tomson T, Benn E, et al. Combined analysis of risk factors for SUDEP. *Epilepsia* 2011; **52**(6): 1150-9.
6. Napolitano C, Bloise R, Monteforte N, Priori SG. Sudden cardiac death and genetic ion channelopathies: long QT, Brugada, short QT, catecholaminergic polymorphic ventricular tachycardia, and idiopathic ventricular fibrillation. *Circulation* 2012; **125**(16): 2027-34.
7. Skinner JR, Crawford J, Smith W, et al. Prospective, population-based long QT molecular autopsy study of postmortem negative sudden death in 1 to 40 year olds. *Heart Rhythm* 2011; **8**(3): 412-9.
8. Seyal M, Pascual F, Lee CY, Li CS, Bateman LM. Seizure-related cardiac repolarization abnormalities are associated with ictal hypoxemia. *Epilepsia* 2011; **52**(11): 2105-11.
9. Surges R, Scott CA, Walker MC. Enhanced QT shortening and persistent tachycardia after generalized seizures. *Neurology* 2010; **74**(5): 421-6.
10. Goldman AM, Glasscock E, Yoo J, Chen TT, Klassen TL, Noebels JL. Arrhythmia in heart and brain: KCNQ1 mutations link epilepsy and sudden unexplained death. *Sci Transl Med* 2009; **1**(2): 2ra6.
11. Tu E, Bagnall RD, Duflou J, Semsarian C. Post-mortem review and genetic analysis of sudden unexpected death in epilepsy (SUDEP) cases. *Brain Pathol* 2011; **21**(2): 201-8.
12. Ryvlin P, Nashef L, Lhatoo SD, et al. Incidence and mechanisms of cardiorespiratory arrests in epilepsy monitoring units (MORTEMUS): a retrospective study. *Lancet Neurol* 2013; **12**(10): 966-77.
13. Dravet C, Bureau M, Oguni H, Cokar O, Guerrini R. Dravet syndrome (Severe myoclonic epilepsy in infancy). *In: Bureau M, Genton P, Dravet C, Delgado-Escueta A, Tassinari CA, Thomas P, Wolf P, editors Epileptic syndromes in Infancy, Childhood and Adolescence* 2012; **5th edn.**(Montrouge, France: John Libbey Eurotext): 125-56.
14. Glasscock E, Yoo JW, Chen TT, Klassen TL, Noebels JL. Kv1.1 potassium channel deficiency reveals brain-driven cardiac dysfunction as a candidate mechanism for sudden unexplained death in epilepsy. *J Neurosci* 2010; **30**(15): 5167-75.
15. Wagnon JL, Korn MJ, Parent R, et al. Convulsive seizures and SUDEP in a mouse model of SCN8A epileptic encephalopathy. *Hum Mol Genet* 2015; **24**(2): 506-15. doi: 10.1093/hmg/ddu470. Epub 2014 Sep 16.

16. Klassen TL, Bomben VC, Patel A, et al. High-resolution molecular genomic autopsy reveals complex sudden unexpected death in epilepsy risk profile. *Epilepsia* 2014; **55**(2): e6-12.
17. Veeramah KR, O'Brien JE, Meisler MH, 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**(3): 502-10. doi: 10.1016/j.ajhg.2012.01.006. Epub Feb 23.
18. Bagnall RD, Das KJ, Duflou J, Semsarian C. Exome analysis-based molecular autopsy in cases of sudden unexplained death in the young. *Heart Rhythm* 2014; **11**(4): 655-62.
19. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7**(4): 248-9. doi: 10.1038/nmeth0410-248.
20. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009; **4**(7): 1073-81. doi: 10.38/nprot.2009.86. Epub Jun 25.
21. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* 2014; **11**(4): 361-2. doi: 10.1038/nmeth.2890.
22. Cirulli ET, Lasseigne BN, Petrovski S, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science* 2015; **19**.
23. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**(8): 904-9. Epub 2006 Jul 23.
24. ExomeAggregationConsortium(ExAC). Cambridge, MA. URL: http://exac.broadinstitute.org.
25. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nature methods* 2010; **7**(4): 248-9.
26. Picardi E, Pesole G. Mitochondrial genomes gleaned from human whole-exome sequencing. *Nat Methods* 2012; **9**(6): 523-4.
27. Fan L, Yao YG. An update to MitoTool: using a new scoring system for faster mtDNA haplogroup determination. *Mitochondrion* 2013; **13**(4): 360-3.
28. Fromer M, Moran JL, Chambert K, et al. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am J Hum Genet* 2012; **91**(4): 597-607.
29. Cook EH, Jr., Scherer SW. Copy-number variations associated with neuropsychiatric conditions. *Nature* 2008; **455**(7215): 919-23.
30. de Kovel CG, Trucks H, Helbig I, et al. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* 2010; **133**(Pt 1): 23-32.
31. Helbig I, Mefford HC, Sharp AJ, et al. 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nat Genet* 2009; **41**(2): 160-2.
32. Reinthaler EM, Lal D, Lebon S, et al. 16p11.2 600 kb Duplications confer risk for typical and atypical Rolandic epilepsy. *Hum Mol Genet* 2014.
33. Eijgelsheim M, Newton-Cheh C, Aarnoudse AL, et al. Genetic variation in NOS1AP is associated with sudden cardiac death: evidence from the Rotterdam Study. *Hum Mol Genet* 2009; **18**(21): 4213-8. doi: 10.1093/hmg/ddp356. Epub 2009 Jul 30.

34. Ko YL, Tai DY, Chen SA, Lee-Chen GJ, Chu CH, Lin MW. Linkage and mutation analysis in two Taiwanese families with long QT syndrome. *J Formos Med Assoc* 2001; **100**(11): 767-71.
35. Kapplinger JD, Tester DJ, Salisbury BA, et al. Spectrum and prevalence of mutations from the first 2,500 consecutive unrelated patients referred for the FAMILION long QT syndrome genetic test. *Heart Rhythm* 2009; **6**(9): 1297-303.
36. Kapplinger JD, Tester DJ, Alders M, et al. An international compendium of mutations in the SCN5A-encoded cardiac sodium channel in patients referred for Brugada syndrome genetic testing. *Heart Rhythm* 2010; **7**(1): 33-46.
37. Dibbens LM, de Vries B, Donatello S, et al. Mutations in DEPDC5 cause familial focal epilepsy with variable foci. *Nat Genet* 2013; **45**(5): 546-51.
38. Carvill GL, Heavin SB, Yendle SC, et al. Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat Genet* 2013; **45**(7): 825-30. doi: 10.1038/ng.2646. Epub 013 May 26.
39. Harkin LA, McMahon JM, Iona X, et al. The spectrum of SCN1A-related infantile epileptic encephalopathies. *Brain* 2007; **130**(Pt 3): 843-52.
40. Leventer RJ, Cardoso C, Ledbetter DH, Dobyns WB. LIS1 missense mutations cause milder lissencephaly phenotypes including a child with normal IQ. *Neurology* 2001; **57**(3): 416-22.
41. Mefford HC, Yendle SC, Hsu C, et al. Rare copy number variants are an important cause of epileptic encephalopathies. *Ann Neurol* 2011; **70**(6): 974-85.
42. Bates MG, Bourke JP, Giordano C, d'Amati G, Turnbull DM, Taylor RW. Cardiac involvement in mitochondrial DNA disease: clinical spectrum, diagnosis, and management. *Eur Heart J* 2012; **33**(24): 3023-33. doi: 10.1093/eurheartj/ehs275. Epub 2012 Aug 30.
43. Khurana DS, Valencia I, Goldenthal MJ, Legido A. Mitochondrial dysfunction in epilepsy. *Semin Pediatr Neurol* 2013; **20**(3): 176-87. doi: 10.1016/j.spen.2013.10.001. Epub Oct 9.
44. Kapa S, Tester DJ, Salisbury BA, et al. Genetic testing for long-QT syndrome: distinguishing pathogenic mutations from benign variants. *Circulation* 2009; **120**(18): 1752-60.
45. Huffaker SJ, Chen J, Nicodemus KK, et al. A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia. *Nat Med* 2009; **15**(5): 509-18. doi: 10.1038/nm.962. Epub 2009 May 3.
46. Donahue LM, Coates PW, Lee VH, Ippensen DC, Arze SE, Poduslo SE. The cardiac sodium channel mRNA is expressed in the developing and adult rat and human brain. *Brain Res* 2000; **887**(2): 335-43.
47. Semsarian C IJ, Wilde AA. Sudden cardiac death in the young: the molecular autopsy and a practical approach to surviving relatives. *Eur Heart J* 2015; ((in press)).
48. Bagnall RD, Crompton DE, Cutmore C, et al. Genetic analysis of PHOX2B in sudden unexpected death in epilepsy cases. *Neurology* 2014.
49. Liebenthal JA, Wu S, Rose S, Ebersole JS, Tao JX. Association of prone position with sudden unexpected death in epilepsy. *Neurology* 2015; **84**(7): 703-9. doi: 10.1212/WNL.0000000000001260. Epub 2015 Jan 21.
50. Tomas M, Napolitano C, De Giuli L, et al. Polymorphisms in the NOS1AP gene modulate QT interval duration and risk of arrhythmias in the long QT syndrome. *J Am Coll Cardiol* 2010; **55**(24): 2745-52. doi: 10.1016/j.jacc.2009.12.065.

51. Scheffer IE, Heron SE, Regan BM, et al. Mutations in mammalian target of rapamycin regulator DEPDC5 cause focal epilepsy with brain malformations. *Ann Neurol* 2014; **75**(5): 782-7.
52. Crompton DE, Scheffer IE, Taylor I, et al. Familial mesial temporal lobe epilepsy: a benign epilepsy syndrome showing complex inheritance. *Brain* 2010; **133**(11): 3221-31. doi: 10.1093/brain/awq251. Epub 2010 Sep 23.
53. van Kranenburg M, Hoogeveen-Westerveld M, Nellist M. Preliminary Functional Assessment and Classification of DEPDC5 Variants Associated with Focal Epilepsy. *Hum Mutat* 2014; **4**(10): 22723.
54. Brunklaus A, Ellis R, Reavey E, Semsarian C, Zuberi SM. Genotype phenotype associations across the voltage-gated sodium channel family. *J Med Genet* 2014; **51**(10): 650-8.
55. Wallace RH, Scheffer IE, Barnett S, et al. Neuronal sodium-channel alpha1-subunit mutations in generalized epilepsy with febrile seizures plus. *Am J Hum Genet* 2001; **68**(4): 859-65. Epub 2001 Mar 13.
56. Delogu AB, Spinelli A, Battaglia D, et al. Electrical and autonomic cardiac function in patients with Dravet syndrome. *Epilepsia* 2011; **52 Suppl 2**: 55-8.
57. Porta A, Girardengo G, Bari V, et al. Autonomic control of heart rate and QT interval variability influences arrhythmic risk in long QT syndrome type 1. *J Am Coll Cardiol* 2015; **65**(4): 367-74.
58. Scheffer IE, Harkin LA, Grinton BE, et al. Temporal lobe epilepsy and GEFS+ phenotypes associated with SCN1B mutations. *Brain* 2007; **130**(Pt 1): 100-9. Epub 2006 Oct 4.
59. Watanabe H, Koopmann TT, Le Scouarnec S, et al. Sodium channel beta1 subunit mutations associated with Brugada syndrome and cardiac conduction disease in humans. *J Clin Invest* 2008; **118**(6): 2260-8. doi: 10.1172/JCI33891.
60. Lopez-Santiago LF, Meadows LS, Ernst SJ, et al. Sodium channel Scn1b null mice exhibit prolonged QT and RR intervals. *J Mol Cell Cardiol* 2007; **43**(5): 636-47. Epub 2007 Aug 10.
61. Reid CA, Leaw B, Richards KL, et al. Reduced dendritic arborization and hyperexcitability of pyramidal neurons in a Scn1b-based model of Dravet syndrome. *Brain* 2014; **137**(Pt 6): 1701-15. doi: 10.093/brain/awu077. Epub 2014 Apr 17.

FIGURE LEGENDS

Figure 1. Overview of genetic analysis of SUDEP cases. All SUDEP cases (n=61) underwent exome sequencing and analysis of 3 common LQTS genes, 29 cardiac arrhythmia genes, 5 central hypoventilation genes and 72 epilepsy genes. CNV analysis of TruSeq exome data was performed on 41 patients. Analysis of mitochondrial genome was performed on 59 patients.

Figure 2. LQTS mutations and candidate pathogenic variants in SUDEP cases. KCNQ1, KCNH2, and SCN5A proteins are shown with positively charged transmembrane segments in green. Positions of previously reported pathogenic mutations (blue circles), *de novo* missense mutation (red circle), and candidate pathogenic variants (green circles) are shown. Multiple protein alignments of regions surrounding the SCN5A missense sites, with species specific numbering on the right are shown at the bottom. Amino acid residues identical to the top sequence are indicated by '.'.

Figure 3. Location and validation of *DEPDC5* mutations in SUDEP cases. A. DEPDC5 protein showing location of conserved protein domains and mutations found in SUDEP cases, with Sanger sequencing electropherograms. B. Multiple protein alignments of the Ser19Thr and Arg347His missense regions, with species specific numbering on the right. Amino acid residues identical to the top sequence are indicated by a '.'.

Figure 4. XHMM plot of 16p11·2 duplication in EP06. Normalized read depths at each exome target (black dots) in the region are connected by grey lines. Predicted EP06 duplication is shown (green line). Lower panel shows the 16p11·2 genomic region (black bar) annotated with RefSeq gene names (blue bars) and duplicated region (green bar).

Table 1. Characteristics of the SUDEP cohort

Characteristic	Overall	Male	Female
Number of subjects	61	34 (56%)	27 (44%)
Age at epilepsy onset (y) mean, ± SD, (range)	10·3 ± 8·2 (0-34)	10·9 ± 8·8 (0-34)	9·6 ± 7·5 (0-24)
Age at SUDEP (y) mean, ± SD, (range)	28·1 ± 12·0 (1-53)	31·0 ± 11·7 (9-53)	24·4 ± 11·6 (1-40)

Table 2 Characteristics of SUDEP Occurrence

Characteristic	SUDEP Cohort
SUDEP category	
Definite SUDEP	54 (89%)
Definite SUDEP plus	2 (3%)
Probable SUDEP	5 (8%)
Witnessed death	
No	57 (93%)
Yes	1 (2%)
Not recorded	3 (5%)
Position at SUDEP	
Prone	27 (44%)
Supine	1 (2%)
Unknown	33 (54%)
Number of prescribed AED	
0	6 (10%)
1	32 (52%)
>1	20 (33%)
Not recorded	3 (5%)

AED - anti epileptic drugs

Table 3. *De novo* Mutation and Pathogenic Variants in Cardiac Arrhythmia Genes

***De novo* Mutations and Previously Reported Pathogenic Mutations**

ID/Gene Amino acid (ExAC AC)	Variant	Epileptiform EEG/ Brain MRI	Epilepsy syndrome or coronial evidence of epilepsy
EP11/*KCNH2* R744* (0)	LQT2 pathogenic	NA / NA	History of epilepsy; infarct of left basal ganglia.
EP19/*KCNH2* G924A (0)	LQT2 pathogenic	Yes / hippocampal sclerosis	Mesial temporal lobe epilepsy
EP40/*KCNQ1* Y662* (0)	LQT1 pathogenic	Yes / normal	Dravet syndrome
EP41/*SCN5A* I397V (0)	LQT3 *de novo* mutation	Yes / normal	Juvenile myoclonic epilepsy
Candidate Pathogenic Variants in Dominant Cardiac Arrhythmia Genes			
EP63/*ANK2* A1027D (0)	LQT4 novel	NA / NA	History of absence seizures and tonic clonic seizure
EP35/*ANK2* S2440N (0)	LQT4 novel	NA / normal	Temporal lobe epilepsy
EP59/*ANK2* I3903N (0)	LQT4 rare	NA / NA	History of seizures
EP39/*AKAP9* I1749T (109)	LQT11 rare	Yes / temporal heterotopia	Structural temporal lobe epilepsy
EP43/*AKAP9* R2607G (0)	LQT11 novel	Yes / NA	Juvenile myoclonic epilepsy
EP21/*HCN4* E1193Q (76)	BrS8 rare	NA / NA	History of nocturnal seizures; focal cortical dysplasia
EP55/*KCNH2* G749A (0)	LQT2 novel	NA / NA	History of epilepsy
EP14/*RYR2* C1489R (20)	CPVT1 rare	NA / NA	History of nocturnal seizures
EP12/*SCN5A* V223G (0)	LQT3 novel	NA / NA	Unspecified diagnosis of epilepsy

ExAC AC - Exome Aggregate Consortium allele count; LQT - Long QT syndrome; BrS1 - Brugada syndrome type 1; BrS8 - Brugada syndrome type 8; CPVT1 - Catecholaminergic polymorphic ventricular tachycardia type 1; NA - not available.

Table 4. *De novo* Mutations and Pathogenic Variants in Epilepsy Related Genes

De novo mutations and Previously Reported Pathogenic Mutations

ID/Gene Amino acid (ExAC AC)	Variant classification	Abnormal EEG/ Brain MRI	Epilepsy syndrome or coronial evidence of epilepsy
EP09/*DEPDC5* R843* (0)	FFEVF pathogenic	NA / NA	History of epilepsy; temporal lobe pathology
EP29/*GABRB3* Y182F (0)	EE *de novo* mutation	Yes / normal	Epileptic encephalopathy
EP38[$]/*PAFAH1B1* G162S (0)	Lissencephaly pathogenic	Yes / NA	Structural focal epilepsy
EP37/*SCN1A* G1480V (0)	GEFS+2 *de novo* mutation	Yes / normal	Epilepsy with myoclonic-atonic seizures
EP67/*SCN2A* R1882Q (0)	EE11 *de novo* mutation	Yes / normal	Epileptic encephalopathy
EP73/*SCN2A* N976K (0)	EE11 *de novo* mutation	Yes / microcephaly	Epileptic encephalopathy
Candidate Pathogenic Variants in Dominant Epilepsy Genes			
EP13/*CHRNA4* F66L (0)	NFLE1 novel	NA / NA	History of epilepsy
EP62/*DEPDC5* S19T (0)	FFEVF novel	NA / NA	History of epilepsy
EP70/*DEPDC5* R286* (0)	FFEVF novel	NA / NA	History of nocturnal epilepsy
EP10/*DEPDC5* R347H (1)	FFVEF rare	NA / NA	History of epilepsy
EP64/*DEPDC5* Q1016* (0)	FFVEF novel	Yes / parietal dysplasia	Structural Parietal Lobe Epilepsy
EP23/*DEPDC5* R1332* (0)	FFVEF novel	Yes / cortical dysplasia	Structural fronto-temporal lobe epilepsy
EP66/*KCNQ2* A306V (0)	EE7 novel	Yes / normal	Ohtahara Syndrome
EP32/*PCDH19* N509S (0)	EE9 novel	Yes / normal	Juvenile myoclonic epilepsy
EP51/*SCN1B* R96Q (11)	GEFS+1 rare	NA / NA	History of epilepsy
EP38[$]/*SPTAN1* Q425R (6)	EE5 rare	Yes / NA	Structural focal epilepsy

ExAC AC - Exome Aggregate Consortium allele count; '$' - Patient has two variants in epilepsy related genes; NFLE1 - Nocturnal Frontal Lobe Epilepsy; FFEVF - Familial focal epilepsy with variable foci; GEFS+1 - Genetic epilepsy with febrile seizures plus type 1; GEFS+2 - Genetic epilepsy with febrile seizures plus type 2; EE - Early infantile epileptic encephalopathy; NA - not available.

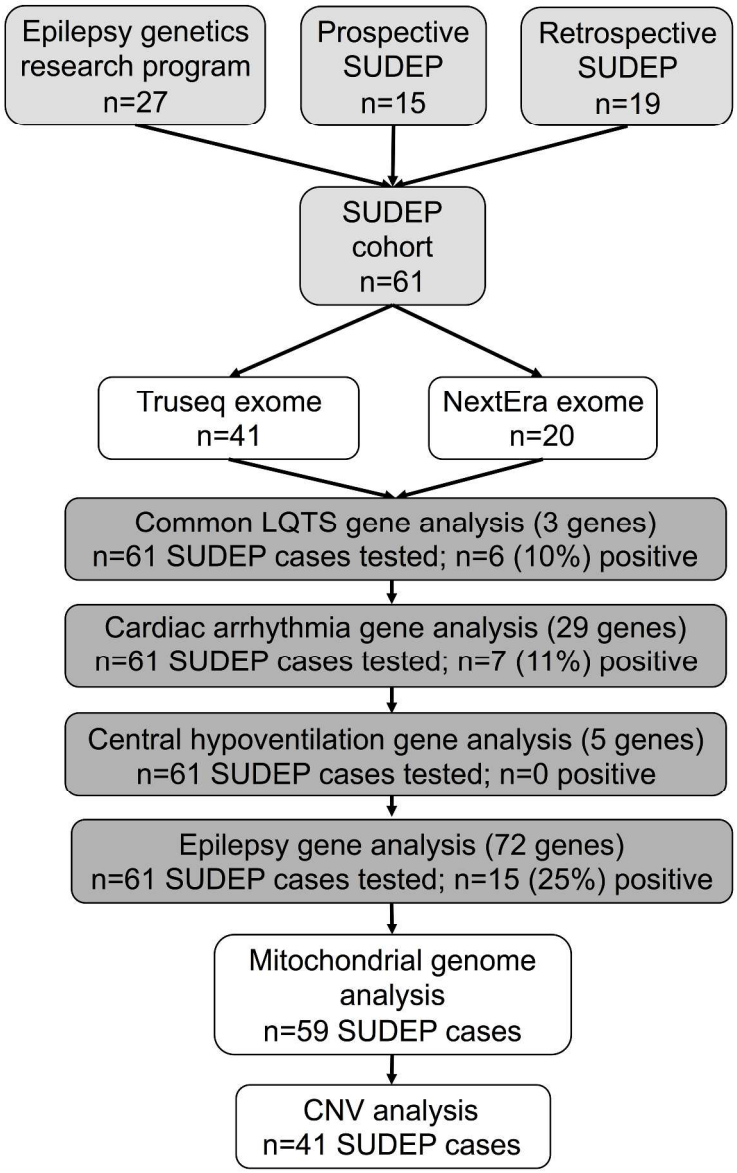

Figure 1
260x398mm (300 x 300 DPI)

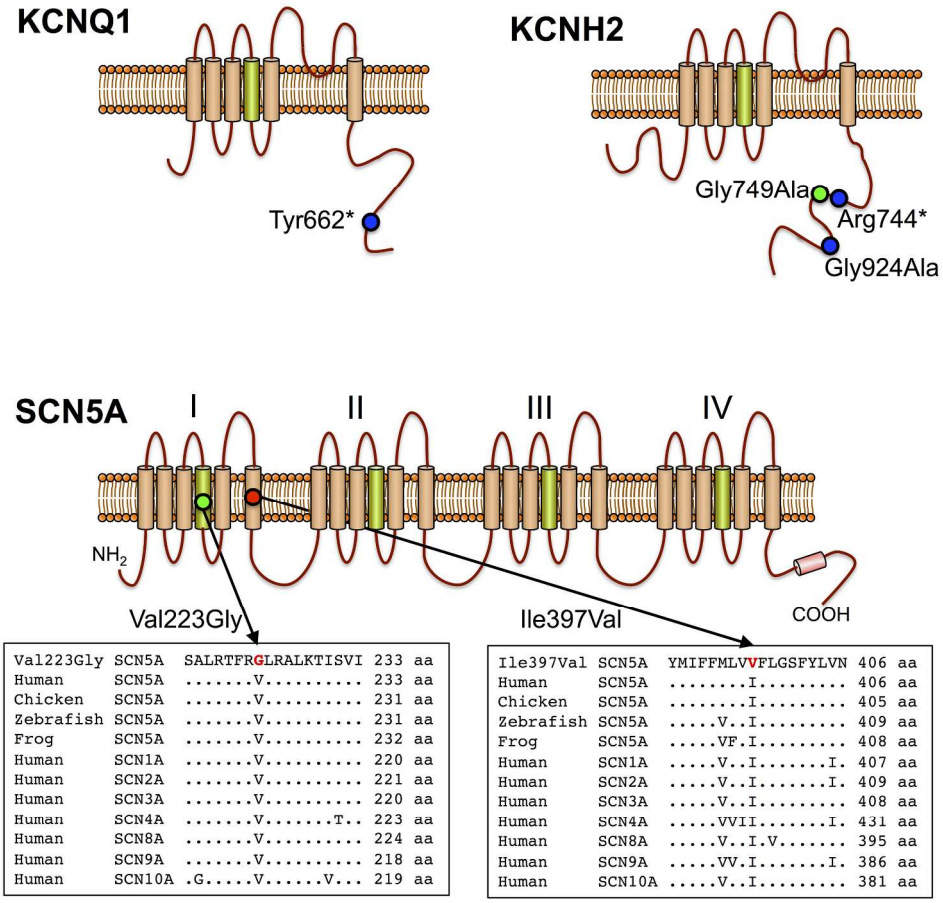

Figure 2
240x227mm (300 x 300 DPI)

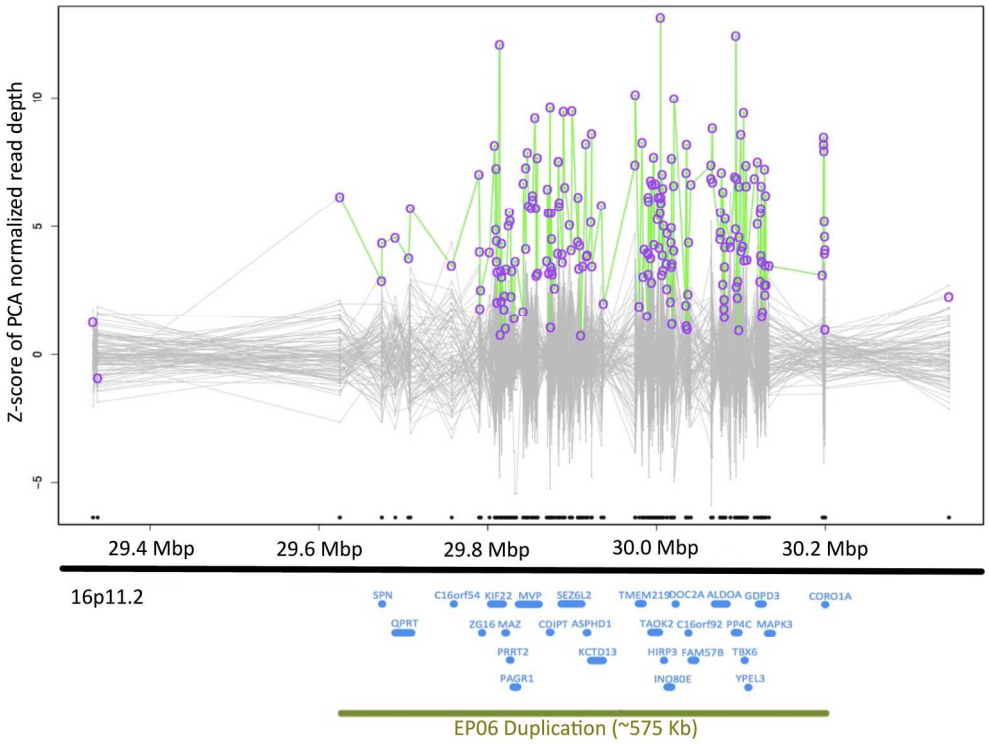

Figure 4
190x142mm (300 x 300 DPI)

Supplementary Information

Supplementary Table 1. Target Genes for Focused Analysis of Mutations and Candidate Pathogenic Variants

Gene Symbol	Gene category	OMIM Disease
ALDH7A1	Epilepsy	Epilepsy, pyridoxine-dependent, 266100 (3)
ARHGEF9	Epilepsy	Epileptic encephalopathy, early infantile, 8, 300607 (3)
ARX	Epilepsy	Epileptic encephalopathy, early infantile, 1, 308350 (3)
ASAH1	Epilepsy	Spinal muscular atrophy with progressive myoclonic epilepsy, 159950 (3)
CACNA1H	Epilepsy	{Epilepsy, childhood absence, susceptibility to, 6}, 611942 (3); {Epilepsy, idiopathic generalized, susceptibility to, 6}, 611942 (3)
CACNB4	Epilepsy	{Epilepsy, idiopathic generalized, susceptibility to, 9}, 607682 (3); {Epilepsy, juvenile myoclonic, susceptibility to, 6}, 607682 (3)
CASR	Epilepsy	{Epilepsy idiopathic generalized, susceptibility to, 8}, 612899 (3)
CARS2	Epilepsy	Neurology. 2014 Oct 31; PMID: 25361775
CDKL5	Epilepsy	Epileptic encephalopathy, early infantile, 2, 300672 (3)
CHD2	Epilepsy	Epileptic encephalopathy, childhood-onset, 615369 (3)
CHRNA2	Epilepsy	Epilepsy, nocturnal frontal lobe, type 4, 610353 (3)
CHRNA4	Epilepsy	Epilepsy, nocturnal frontal lobe, 1, 600513 (3)
CHRNB2	Epilepsy	Epilepsy, nocturnal frontal lobe, 3, 605375 (3)
CLCN2	Epilepsy	{Epilepsy, idiopathic generalized, susceptibility to, 11}, 607628 (3); {Epilepsy, juvenile absence, susceptibility to, 2}, 607628 (3);{Epilepsy, juvenile myoclonic, susceptibility to, 8}, 607628 (3)
CLN8	Epilepsy	Ceroid lipofuscinosis, neuronal, 8, Northern epilepsy variant, 610003 (3)
CNTN2	Epilepsy	?Epilepsy, familial adult myoclonic, 5, 615400 (3)
CNTNAP2	Epilepsy	Cortical dysplasia-focal epilepsy syndrome, 610042 (3)
CPA6	Epilepsy	Epilepsy, familial temporal lobe, 5, 614417 (3)
CSTB	Epilepsy	Epilepsy, progressive myoclonic 1A (Unverricht and Lundborg), 254800 (3)
DEPDC5	Epilepsy	Epilepsy, familial focal, with variable foci, 604364 (3)
DOCK7	Epilepsy	Epileptic encephalopathy, early infantile, 23, 615859 (3)
DNM1	Epilepsy	Am J Hum Genet. 2014 Oct 2;95(4):360-70; PMID: 25262651
EFHC1	Epilepsy	{Epilepsy, juvenile absence, susceptibility to, 1}, 607631 (3); {Myoclonic epilepsy, juvenile, susceptibility to, 1}, 254770 (3)
EPM2A	Epilepsy	Epilepsy, progressive myoclonic 2A (Lafora), 254780 (3)
GABRA1	Epilepsy	Epileptic encephalopathy, early infantile, 19, 615744 (3); {Epilepsy, childhood absence, susceptibility to, 4}, 611136 (3); {Epilepsy, juvenile myoclonic, susceptibility to, 5}, 611136 (3)
GABRB3	Epilepsy	{Epilepsy, childhood absence, susceptibility to, 5}, 612269 (3)
GABRD	Epilepsy	{Epilepsy, generalized, with febrile seizures plus, type 5, susceptibility to}, 613060 (3); {Epilepsy, idiopathic generalized, 10}, 613060 (3); {Epilepsy, juvenile myoclonic, susceptibility to}, 613060 (3)
GABRG2	Epilepsy	Epilepsy, generalized, with febrile seizures plus, type 3, 611277 (3); {Epilepsy, childhood absence, susceptibility to, 2}, 607681 (3)
GNAO1	Epilepsy	Epileptic encephalopathy, early infantile, 17, 615473 (3)
GOSR2	Epilepsy	Epilepsy, progressive myoclonic 6, 614018 (3)
GRIN2A	Epilepsy	Epilepsy, focal, with speech disorder and with or without mental retardation, 245570 (3)
HCN1	Epilepsy	Epileptic encephalopathy, early infantile, 24, 615871 (3)
IER3IP1	Epilepsy	Microcephaly, epilepsy, and diabetes syndrome, 614231 (3)
KCNB1	Epilepsy	Epileptic encephalopathy, early infantile, 26, 616056 (3)

KCNA1	Epilepsy	Epilepsia. 2014 Nov;55(11):1808-16 (PMID: 25377007)
KCNMA1	Epilepsy	Generalized epilepsy and paroxysmal dyskinesia, 609446 (3)
KCNQ2	Epilepsy	Epileptic encephalopathy, early infantile, 7, 613720 (3)
KCNT1	Epilepsy	Epilepsy, nocturnal frontal lobe, 5, 615005 (3); Epileptic encephalopathy, early infantile, 14, 614959 (3)
KCTD7	Epilepsy	Epilepsy, progressive myoclonic 3, with or without intracellular inclusions, 611726 (3)
LGI1	Epilepsy	Epilepsy, familial temporal lobe, 1, 600512 (3)
ME2	Epilepsy	{Epilepsy, idiopathic generalized, susceptibility to}, 600669 (3)
MEF2C	Epilepsy	Mental retardation, stereotypic movements, epilepsy, and/or cerebral malformations, 613443 (3)
NECAP1	Epilepsy	?Epileptic encephalopathy, early infantile, 21, 615833 (3)
NHLRC1	Epilepsy	Epilepsy, progressive myoclonic 2B (Lafora), 254780 (3)
PAFAH1B1	Epilepsy	Lissencephaly 1, 607432 (3)
PCDH19	Epilepsy	Epileptic encephalopathy, early infantile, 9, 300088 (3)
PLCB1	Epilepsy	Epileptic encephalopathy, early infantile, 12, 613722 (3)
PNKP	Epilepsy	Nat Genet. 2010 Mar;42(3):245-9 (PMID: 20118933)
PRICKLE1	Epilepsy	Epilepsy, progressive myoclonic 1B, 612437 (3)
PRICKLE2	Epilepsy	Epilepsy, progressive myoclonic 5, 613832 (3)
PRRT2	Epilepsy	Neurology. 2012 Nov 20;79(21):2104-8 (PMID: 23077018)
SCARB2	Epilepsy	Epilepsy, progressive myoclonic 4, with or without renal failure, 254900 (3)
SCN1A	Epilepsy	Epilepsy, generalized, with febrile seizures plus, type 2, 604403 (3)
SCN1B	Epilepsy/cardiac arrhythmia	Epilepsy, generalized, with febrile seizures plus, type 1, 604233 (3)
SCN2A	Epilepsy	Epileptic encephalopathy, early infantile, 11, 613721 (3)
SCN8A	Epilepsy	Epileptic encephalopathy, early infantile, 13, 614558 (3)
SCN9A	Epilepsy	Epilepsy, generalized, with febrile seizures plus, type 7, 613863 (3)
SIAT9	Epilepsy	Amish infantile epilepsy syndrome, 609056 (3)
SLC13A5	Epilepsy	Epileptic encephalopathy, early infantile, 25, 615905 (3)
SLC25A22	Epilepsy	Epileptic encephalopathy, early infantile, 3, 609304 (3)
SLC2A1	Epilepsy	{Epilepsy, idiopathic generalized, suscpetibility to, 12}, 614847 (3)
SNIP1	Epilepsy	Psychomotor retardation, epilepsy, and craniofacial dysmorphism, 614501 (3)
SPTAN1	Epilepsy	Epileptic encephalopathy, early infantile, 5, 613477 (3)
SRPX2	Epilepsy	?Rolandic epilepsy, mental retardation, and speech dyspraxia, 300643 (3)
ST3GAL3	Epilepsy	Epileptic encephalopathy, early infantile, 15, 615006 (3)
STRADA	Epilepsy	Polyhydramnios, megalencephaly, and symptomatic epilepsy, 611087 (3)
STXBP1	Epilepsy	Epileptic encephalopathy, early infantile, 4, 612164 (3)
SYN1	Epilepsy	Epilepsy, X-linked, with variable learning disabilities and behavior disorders, 300491 (3)
SYNGAP1	Epilepsy	Nat Genet. 2013 Jul;45(7):825-30; PMID: 23708187
SZT2	Epilepsy	Epileptic encephalopathy, early infantile, 18, 615476 (3)
TBC1D24	Epilepsy	Epileptic encephalopathy, early infantile, 16, 615338 (3); Myoclonic epilepsy, infantile, familial, 605021 (3)
TNK2	Epilepsy	Ann Neurol. 2013 Sep;74(3):496-501 (PMID: 23686771)
ABCC9	Cardiac arrhythmia	Atrial fibrillation, familial, 12, 614050 (3)
AKAP9	Cardiac arrhythmia	Long QT syndrome 11, 611820 (3)
ANK2	Cardiac arrhythmia	Long QT syndrome 4, 600919 (3)

CACNA1C	Cardiac arrhythmia	Brugada syndrome 3, 611875 (3)
CACNB2	Cardiac arrhythmia	Brugada syndrome 4, 611876 (3)
CALM1	Cardiac arrhythmia	Ventricular tachycardia, catecholaminergic polymorphic, 4, 614916 (3)
CASQ2	Cardiac arrhythmia	Ventricular tachycardia, catecholaminergic polymorphic, 2, 611938 (3)
CAV3	Cardiac arrhythmia	Long QT syndrome 9, 611818 (3)
CPVT3	Cardiac arrhythmia	Ventricular tachycardia, catecholaminergic polymorphic, 3 (2)
GJA5	Cardiac arrhythmia	Atrial fibrillation, familial, 11, 614049 (3)
GNAI2	Cardiac arrhythmia	Ventricular tachycardia, idiopathic, 192605 (3)
GPD1L	Cardiac arrhythmia	Brugada syndrome 2, 611777 (3)
HCN4	Cardiac arrhythmia	Brugada syndrome 8, 613123 (3) / Sick sinus syndrome 2, 163800 (3)
KCNA5	Cardiac arrhythmia	Atrial fibrillation, familial, 7, 612240 (3)
KCNE1	Cardiac arrhythmia	Long QT syndrome 5, 613695 (3)
KCNE2	Cardiac arrhythmia	Long QT syndrome 6, 613693 (3); Atrial fibrillation, familial, 4, 611493 (3)
KCNE3	Cardiac arrhythmia	Brugada syndrome 6, 613119 (3)
KCNH2	Cardiac arrhythmia	Long QT syndrome 2, 613688 (3); Short QT syndrome 1, 609620 (3)
KCNJ2	Cardiac arrhythmia	Short QT syndrome 3, 609622 (3); Atrial fibrillation, familial, 9, 613980 (3)
KCNJ5	Cardiac arrhythmia	Long QT syndrome 13, 613485 (3)
KCNQ1	Cardiac arrhythmia	Long QT syndrome 1, 192500 (3); Short QT syndrome 2, 609621 (3)
NPPA	Cardiac arrhythmia	Atrial fibrillation, familial, 6, 612201 (3)
NUP155	Cardiac arrhythmia	Atrial fibrillation, familial, 15, 615770 (3)
RYR2	Cardiac arrhythmia	Ventricular tachycardia, catecholaminergic polymorphic, 1, 604772 (3)
SCN1B	Cardiac arrhythmia epilepsy	Atrial fibrillation, familial, 13, 615377 (3); Brugada syndrome 5, 612838 (3)
SCN2B	Cardiac arrhythmia	Atrial fibrillation, familial, 14, 615378 (3)
SCN3B	Cardiac arrhythmia	Brugada syndrome 7, 613120 (3)
SCN4B	Cardiac arrhythmia	Long QT syndrome 10, 611819 (3)
SCN5A	Cardiac arrhythmia	Long QT syndrome 3, 603830 (3); Brugada syndrome 1, 601144 (3)
SNTA1	Cardiac arrhythmia	Long QT syndrome 12, 612955 (3)
TRDN	Cardiac arrhythmia	Ventricular tachycardia, catecholaminergic polymorphic, 5, with or without muscle weakness, 615441 (3)
TRPM4	Cardiac arrhythmia	Heart block, progressive, type IB, 604559 (3)
ASCL1	Central hypoventilation syndrome	Central hypoventilation syndrome, congenital, 209880 (3)
BDNF	Central hypoventilation syndrome	Central hypoventilation syndrome, congenital, 209880 (3)
EDN3	Central hypoventilation syndrome	Central hypoventilation syndrome, congenital, 209880 (3)
GDNF	Central hypoventilation syndrome	Central hypoventilation syndrome, 209880 (3)
RET	Central hypoventilation syndrome	Central hypoventilation syndrome, congenital, 209880 (3)

Supplementary Table 2. Exome Sequencing Mapping Metrics

ID	Exome Kit	Total reads	Mapped reads	Reads mapped to genome (%)	Reads mapped uniquely to target exome	Reads mapped uniquely to target exome (%)	Mean coverage at target bases	Coverage >10 reads (%)
EP01	TruSeq	65,262,816	65,162,409	99.85	41,058,946	63.0	53.1	91.0
EP02	TruSeq	78,864,690	78,741,405	99.84	48,310,194	61.4	62.7	91.0
EP03	TruSeq	82,612,436	82,407,125	99.75	48,861,796	59.3	63.6	91.2
EP04	TruSeq	74,704,442	74,560,542	99.81	47,501,596	63.7	61.4	91.5
EP05	TruSeq	69,417,724	69,296,198	99.82	43,334,518	62.5	56.2	91.0
EP06	TruSeq	62,666,666	62,556,939	99.82	39,332,954	62.9	51.0	90.3
EP07	TruSeq	69,159,554	69,028,812	99.81	42,307,352	61.3	54.7	91.0
EP08	TruSeq	70,025,432	69,850,523	99.75	45,122,302	64.6	58.9	90.2
EP09	TruSeq	63,487,834	63,285,757	99.68	39,322,994	62.1	51.1	91.2
EP10	TruSeq	64,237,510	64,095,857	99.78	40,737,473	63.6	52.9	90.5
EP11	TruSeq	71,833,752	71,672,532	99.78	44,677,470	62.3	58.1	91.2
EP12	TruSeq	61,851,110	61,723,486	99.79	37,952,498	61.5	49.4	90.2
EP13	TruSeq	69,657,012	69,517,295	99.80	43,018,741	61.9	55.8	91.2
EP14	TruSeq	67,630,836	67,516,103	99.83	42,111,077	62.4	54.5	90.9
EP15	TruSeq	75,356,108	75,211,736	99.81	46,936,869	62.4	60.8	91.7
EP16	TruSeq	67,260,564	67,132,322	99.81	41,547,852	61.9	54.0	90.8
EP17	TruSeq	72,871,150	72,687,717	99.75	44,485,058	61.2	57.8	91.4
EP19	TruSeq	67,199,306	67,046,523	99.77	41,358,706	61.7	53.8	91.0
EP20	TruSeq	66,707,472	66,564,635	99.79	41,877,627	62.9	54.3	91.0
EP21	TruSeq	62,102,156	61,920,411	99.71	38,793,714	62.7	50.4	91.1
EP22	TruSeq	71,489,552	71,357,059	99.81	44,250,201	62.0	57.3	91.1
EP23	TruSeq	73,299,232	73,147,922	99.79	45,703,164	62.5	59.5	91.0
EP24	TruSeq	86,522,636	86,237,193	99.67	51,820,910	60.1	67.3	92.4
EP25	TruSeq	64,544,142	64,413,897	99.80	39,911,033	62.0	51.7	90.5
EP26	TruSeq	66,872,354	66,725,583	99.78	41,073,644	61.6	53.4	90.9
EP27	TruSeq	64,701,826	64,558,392	99.78	40,399,859	62.6	52.3	90.4
EP29	TruSeq	67,640,030	67,495,347	99.79	42,298,682	62.7	54.9	90.9
EP30	TruSeq	72,483,720	72,356,366	99.82	45,086,925	62.3	58.2	91.2
EP31	TruSeq	72,640,648	72,491,894	99.80	44,138,485	60.9	57.3	90.8
EP32	TruSeq	62,616,476	62,498,540	99.81	39,416,126	63.1	51.0	90.2
EP33	TruSeq	70,386,930	70,223,500	99.77	42,236,391	60.1	55.0	91.0
EP34	TruSeq	61,917,048	61,787,353	99.79	38,559,487	62.4	50.1	90.4
EP35	TruSeq	80,800,982	80,611,814	99.77	48,561,705	60.2	63.1	91.9
EP37	TruSeq	73,534,120	73,382,128	99.79	44,436,933	60.6	57.8	90.9
EP38	TruSeq	68,605,854	68,470,231	99.80	42,231,157	61.7	54.8	90.4
EP39	TruSeq	74,287,450	74,116,115	99.77	45,443,473	61.3	58.8	91.7

EP40	TruSeq	66,279,314	66,136,478	99.78	40,473,449	61.2	52.5	90.8
EP41	TruSeq	69,857,920	69,714,223	99.79	43,464,614	62.3	56.3	91.4
EP42	TruSeq	70,190,838	70,039,990	99.79	43,547,172	62.2	56.5	91.0
EP43	TruSeq	81,864,676	81,668,680	99.76	50,948,028	62.4	66.2	91.8
EP44	TruSeq	76,327,248	76,175,749	99.80	47,521,088	62.4	61.5	91.5
EP51	NextEra	124,456,259	124,325,564	99.89	36,243,677	29.2	46.9	80.6
EP52	NextEra	65,118,764	64,938,748	99.72	36,331,767	55.9	47,2	83.1
EP53	NextEra	80,772,362	80,616,090	99.81	43,707,824	54.2	57.0	86.3
EP55	NextEra	124,456,260	124,327,477	99.90	40,810,955	32.8	53.0	86.0
EP56	NextEra	72,642,582	72,463,705	99.75	39,564,622	54.6	51.5	84.6
EP57	NextEra	64,332,508	64,229,148	99.84	36,388,682	56.7	47.2	82.7
EP58	NextEra	124,456,261	124,259,815	99.84	43,669,522	35.1	56.5	83.6
EP59	NextEra	78,695,618	78,577,379	99.85	42,434,404	54.0	55.1	84.8
EP60	NextEra	80,158,384	80,037,806	99.85	44,963,928	56.2	58.4	84.0
EP61	NextEra	65,637,810	65,521,877	99.82	36,601,665	55.9	47.6	83.0
EP62	NextEra	124,456,257	124,328,667	99.90	36,348,923	29.2	47.3	82.7
EP63	NextEra	70,561,036	70,436,235	99.82	38,914,313	55.2	50.1	80.0
EP64	NextEra	64,332,834	64,227,193	99.84	36,123,672	56.2	46.9	82.4
EP65	NextEra	77,091,408	76,964,097	99.83	42,385,787	55.1	55.1	84.5
EP66	NextEra	124,456,258	124,334,113	99.90	39,353,353	31.7	51.0	83.6
EP67	NextEra	81,315,240	81,175,020	99.83	42,657,425	52.5	55.5	83.6
EP69	NextEra	79,540,784	79,396,822	99.82	41,916,762	52.8	54.4	83.0
EP70	NextEra	124,456,255	124,362,597	99.92	33,796,039	27.2	43.8	80.7
EP72	NextEra	73,216,016	72,666,214	99.25	40,407,490	55.6	52.4	82.9
EP73	NextEra	124,456,256	124,323,224	99.89	46,963,885	37.8	61.0	85.7

Supplementary Table 3. Gene Collapsing Test of Rare Variants Predicted to be Damaging

Rank	Gene Symbol	RVIS %	Qualifying Cases (frequency) n=57	Qualifying Controls (frequency) n=2936	Fisher's Exact Test p-value	Gene Assessed (%)
1	*ZNF358*	77.8	3 (0.0526)	2 (0.0007)	6.38×10^{-5}	100.00
2	*DEPDC5*	6.62	5 (0.0877)	23 (0.0078)	1.48×10^{-4}	92.62
3	*NOS1AP*	55.45	3 (0.0526)	4 (0.0014)	2.17×10^{-4}	100.00
4	*NEFM*	69.62	2 (0.0351)	0	3.56×10^{-4}	43.90
5	*SAMD11*	N/A	3 (0.0526)	8 (0.0027)	9.7×10^{-4}	90.32
6	*PRCC*	23.25	3 (0.0526)	8 (0.0027)	9.7×10^{-4}	100.00
7	*RRAGB*	42.88	2 (0.0351)	1 (0.0003)	0.0011	92.47
8	*TLDC2*	79.25	2 (0.0351)	1 (0.0003)	0.0011	100.00
9	*SESN3*	31.46	2 (0.0351)	1 (0.0003)	0.0011	94.60
10	*DDX55*	45.57	3 (0.0526)	9 (0.0031)	0.0013	100.00
11	*KTN1*	9.41	4 (0.0702)	25 (0.0085)	0.002	99.79
12	*SH2D3C*	16.71	3 (0.0526)	11 (0.0037)	0.0021	74.31
13	*TBX15*	65.96	3 (0.0526)	11 (0.0037)	0.0021	100.00
14	*KLF10*	75.43	2 (0.0351)	2 (0.0007)	0.0021	78.79
15	*RSPO2*	53.51	2 (0.0351)	2 (0.0007)	0.0021	64.36
16	*TUFM*	26.73	2 (0.0351)	2 (0.0007)	0.0021	100.00
17	*IFT74*	92.98	3 (0.0526)	12 (0.0041)	0.0025	96.46
18	*NAGK*	22.36	3 (0.0526)	13 (0.0044)	0.0031	100.00
19	*OR2AT4*	78.46	2 (0.0351)	3 (0.001)	0.0034	100.00
20	*STARD7*	50.01	2 (0.0351)	3 (0.001)	0.0034	100.00
21	*CLEC4F*	99.29	2 (0.0351)	3 (0.001)	0.0034	100.00
22	*SPO11*	78.46	2 (0.0351)	3 (0.001)	0.0034	96.70
23	*OR1N1*	79.38	2 (0.0351)	3 (0.001)	0.0034	100.00
24	*KCNH2*	3.81	3 (0.0526)	14 (0.0048)	0.0037	73.76
25	*EML6*	17.5	4 (0.0702)	31 (0.0106)	0.004	96.57
26	*MYBPC2*	98.45	4 (0.0702)	33 (0.0112)	0.0049	96.72
27	*CST7*	53.19	2 (0.0351)	4 (0.0014)	0.0051	83.70
28	*PHF20*	19.86	2 (0.0351)	4 (0.0014)	0.0051	96.68
29	*GPR26*	NA	2 (0.0351)	4 (0.0014)	0.0051	34.50
30	*PSMC6*	44.89	2 (0.0351)	4 (0.0014)	0.0051	90.62

RVIS - residual variation intolerance score. Gene Assessed (%): Percentage of gene analysed after excluding exons with high variability in coverage between cases and controls.

Supplementary Table 4. Gene Collapsing Analysis of Singleton Variants Predicted to be Damaging

Rank	Gene Symbol	RVIS %	Qualifying Cases	Frequency Qualifying Cases (n=57)	Qualifying Controls	Frequency Qualifying Controls (n=2936)	Fisher's Exact Test p-value	Gene Assessed (%)
1	DEPDC5	6.62	5	0.0877	7	0.0024	1.50E-06	92.62
2	RSPO2	53.51	2	0.0351	0	0	3.56E-04	64.36
3	NEFM	69.62	2	0.0351	0	0	3.56E-04	43.90
4	NFE2L2	57.31	2	0.0351	0	0	3.56E-04	97.33
5	PGK2	77.06	2	0.0351	1	3.41E-04	0.0011	100.00
6	CCDC112	38.82	2	0.0351	1	3.41E-04	0.0011	100.00
7	NOS1AP	55.45	2	0.0351	2	6.81E-04	0.0021	100.00
8	PADI3	97.15	2	0.0351	2	6.81E-04	0.0021	100.00
9	CPSF7	17.03	2	0.0351	2	6.81E-04	0.0021	95.07
10	HNF1B	20.26	2	0.0351	2	6.81E-04	0.0021	100.00
11	ETV3	30.07	2	0.0351	2	6.81E-04	0.0021	100.00
12	IFT74	92.98	2	0.0351	3	0.001	0.0034	96.46
13	SCN2A	1.77	2	0.0351	4	0.0014	0.0051	89.69
14	LRWD1	55.83	2	0.0351	4	0.0014	0.0051	100.00
15	CCDC88B	71.08	2	0.0351	4	0.0014	0.0051	77.81
16	TELO2	76.68	2	0.0351	4	0.0014	0.0051	87.86
17	WDR96	3.76	2	0.0351	4	0.0014	0.0051	96.56
18	KCNH2	3.81	2	0.0351	5	0.0017	0.007	73.76
19	TIE1	1.83	2	0.0351	6	0.002	0.0093	97.83
20	NINL	94.94	2	0.0351	6	0.002	0.0093	100.00
21	AGBL1	N/A	2	0.0351	6	0.002	0.0093	96.17
22	AKAP12	89.55	2	0.0351	7	0.0024	0.0118	99.46
23	SCAPER	84.18	2	0.0351	7	0.0024	0.0118	96.66
24	INADL	98.24	2	0.0351	7	0.0024	0.0118	98.28
25	MUC17	99.98	2	0.0351	7	0.0024	0.0118	99.66
26	NRXN1	2.25	2	0.0351	7	0.0024	0.0118	81.06
27	CARD14	89.08	2	0.0351	8	0.0027	0.0145	100.00
28	ZNF467	N/A	2	0.0351	8	0.0027	0.0145	100.00
29	PTCHD2	35	2	0.0351	8	0.0027	0.0145	97.23
30	SIPA1L3	2.95	2	0.0351	9	0.0031	0.0176	92.70

RVIS - residual variance intolerance score.

Supplementary Table 5. CNVs at Target Gene Loci or Recurrent CNV Loci.

ID	Cytoband	CNV Iinterval (hg19/Build37)	Estimated size (KB)	CNV Type	Protein Coding Genes in CNV Region	Exome Targets	Q_Exact Score	Q_Some Score
EP42	15q11.2	chr15:25417782-25429534	11.75	Deletion	*IPW, SNORD115-2, SNORD115-3, SNORD115-4, SNORD115-5, SNORD115-6, SNORD115-7, SNORD115-8*	6	24	96
EP23	15q11.2	chr15:25420074-25429534	9.46	Deletion	*PAR4, SNORD115-11, SNORD115-15, SNORD115-16, SNORD115-17SNORD115-22, SNORD115-24, SNORD115-25, SNORD115-26, SNORD115-27, SNORD115-28, SNORD115-30, SNORD115-31, SNORD115-32, SNORD115-33, SNORD115-35, SNORD115-37, SNORD115-38, SNORD115-39, SNORD115-40, SNORD115-41, SNURF-SNRPN*	5	41	96
EP06	16p11.2	chr16:29624424-30199570	575.15	Duplication	*SPN, QPRT, C16orf54, ZG16, KIF22, MAZ, PRRT2, PAGR1, MVP, CDIPT, SEZ6L2, ASPHD1, KCTD13, TMEM219, TAOK2, HIRIP3, INO80E, DOC2A, C16orf92, FAM57B, ALDOA, PPP4C, TBX6, YPEL3, GDPD3, MAPK3, CORO1A*	204	36	97

Supplementary acknowledgements.

We would like to acknowledge the following individuals for the contributions of control samples: Dr. Julie Hoover-Fong, Dr. Nara L. Sobreira and Dr. David Valle; Dr. Annapurna Poduri; Dr. Scott M. Palmer; Mr. David H. Murdock and The MURDOCK Study Community Registry and Biorepository; Dr. Gianpiero Cavalleri, Dr. Norman Delanty, and Dr. Chantal Depondt; The Washington University Neuromuscular Genetics Project: Dr. Stanley Appel, Dr. Ericka Simpson, Ms. Sharon Halton, and Mr. Luis Lay; Dr. Robert Brown; Dr. William B. Gallentine, Dr. Erin L. Heinzen, Dr. Aatif M. Husain, Ms. Kristen N Linney, Dr. Rodney A. Radtke, Dr. Saurabh R. Sinha, and Ms. Nicole M. Walley; Dr. Vandana Shashi and Ms. Kelly Schoch; Ms. Erica Nading and Dr. Stefanie Schuman; Dr. Doug Marchuk; Dr. Zvi Farfel, Dr. Doron Lancet, and Dr. Elon Pras; Dr. Manal Abdelmalek, Dr. Anna Mae Diehl, Yiping Pan, Ms. Dawn Piercy; Dr. Demetre Daskalakis; Dr. Rasheed Gbadegesin and Dr. Michelle Winn; Dr. Joseph McEvoy, Dr. Anna Need, Mr. Jordan Silver and Ms. Marlyne Silver; Mr. Arthur Holden and Dr. Elijah Behr; Dr. Deborah Koltai Attix, Dr. Ornit Chiba-Falek; Dr. Elizabeth T. Cirulli, Ms. Vicki Dixon and Ms. Jill McEvoy; Dr. Kenneth Schmader, Dr. Shelley McDonald, Dr. Heidi K. White, Dr. Mamata Yanamadala, and the Carol Woods and Crosdaile Retirement Communities.

The collection of control samples and data was funded in part by NIAID Grant UO1AIO67854 (Center for HIV/AIDS Vaccine Immunology ("CHAVI")); The Epilepsy Phenome/Genome Project U01NS053998; Epi4K Project 1 – Epileptic Encephalopathies U01NS077364; Epi4K Sequencing, Biostatistics and Bioinformatics Core U01NS077303; NIMH Grant RC2MH089915; NIMH Grant K01MH098126; an award from SAIC-Frederick, Inc. (M11-074); and an award from Biogen Idec. Some control samples were sequenced under an award to Dr. David Goldstein (R01MH097971, R01MH099216) and Dr. Gerald Nestadt (R01MH097993). Clinical materials were collected under awards R01MH50214 (GN) and R01MH071507 (GN), and a collaborative research study funded by NIMH Grants R01MH071507, R01MH079494, R01MH079487, R01MH079489, and R01MH079494. The latter study includes the following collaborative centers: Johns Hopkins University [Gerald Nestadt, MB.BCh., M.P.H. (PI), Joseph Bienvenu, MD, PhD, Fernando Goes, MD, Marco Grados, MD MPH, Brion Maher, PhD, Mark Riddle, MD, Jack Samuels, PhD, Yin Yao, PhD (NIMH)]; Brown University/Butler Hospital [Benjamin D. Greenberg, M.D., Ph.D]; Columbia University, [Abby Fyer, M.D.]; Harvard/Massachusetts General Hospital [Dan Geller, M.D., David Pauls, PhD]; National Institute of Mental Health (NIMH) [Dennis Murphy, M.D.]; University of California at Los Angeles (UCLA) [James McCracken, M.D.], and University of Southern California [James Knowles MD PhD].