

The impact of damaging epilepsy and cardiac genetic variant burden in sudden death in the young

Megan J. Puckelwartz, PhD^{1,2,*}; Lorenzo L. Pesce, PhD^{1,2,3}; Edgar J. Hernandez, PhD^{3,4}; Gregory Webster, MD⁵, MPH; Lisa M. Dellefave-Castillo, MS²; Mark W. Russell, MD⁶, Sarah S. Geisler, MS⁶, Samuel D. Kearns, BS²; Felix K Karthik²; Susan P. Etheridge, MD⁷; Tanner O. Monroe², PhD; Tess D. Pottinger², PhD; Prince J. Kannankeril, MD⁸; M. Benjamin Shoemaker, MD⁹; Darlene Fountain, BSN⁸; Dan M. Roden¹⁰, MD; Heather MacLeod, MS¹¹; Kristin M. Burns, MD¹²; Mark Yandell, PhD¹³; Martin Tristani-Firouzi, MD^{7,14}; Alfred L. George, MD^{1,14}; Elizabeth M. McNally, MD, PhD^{2,14}

³The authors contributed equally
¹⁴These authors contributed equally

¹ Department of Pharmacology, Feinberg School of Medicine, Northwestern University, Chicago, IL
² Center for Genetic Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL
⁴ Biomedical Informatics, University of Utah, Salt Lake City, Utah
⁵ Division of Cardiology, Department of Pediatrics, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL
⁶ Department of Pediatrics, University of Michigan
⁷ Division of Pediatric Cardiology, University of Utah, Salt Lake City, Utah
⁸ Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN
⁹ Department of Medicine, Division of Cardiovascular Medicine, Vanderbilt University Medical Center, Nashville, TN
¹⁰ Departments of Medicine and Pharmacology, Vanderbilt University Medical Center, Nashville, TN
¹¹ Data Coordinating Center, SDY Case Registry
¹² Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland
¹³ Department of Human Genetics, University of Utah, Salt Lake City, Utah

*Correspondence:
Megan J. Puckelwartz PhD
303 E. Superior
SQ 5- 511
Chicago, IL 60611
P: 312-503-6232
F: 312 503 6210
m.puckelwartz@northwestern.edu
Word count: 2973

ABSTRACT

Background: Sudden unexpected death in children is a tragic event. Understanding the genetics of sudden death in the young (SDY) enables family counseling and cascade screening. The objective of this study was to characterize genetic variation in an SDY cohort using whole genome sequencing.

Methods: The SDY Case Registry is a National Institutes of Health/Centers for Disease Control surveillance effort to discern the prevalence, causes, and risk factors for SDY. The SDY Case Registry prospectively collected clinical data and DNA biospecimens from SDY cases <20 years of age. SDY cases were collected from medical examiner and coroner offices spanning 13 US jurisdictions from 2015-2019. The cohort included 211 children (mean age 1 year; range 0-20 years), determined to have died suddenly and unexpectedly and in whom DNA biospecimens and next-of-kin consent were ascertained. A control cohort consisted of 211 randomly sampled, sex- and ancestry-matched individuals from the 1000 Genomes Project. Genetic variation was evaluated in epilepsy, cardiomyopathy and arrhythmia genes in the SDY and control cohorts. American College of Medical Genetics/Genomics guidelines were used to classify variants as pathogenic or likely pathogenic. Additionally, genetic variation predicted to be damaging was identified using a Bayesian-based artificial intelligence (AI) tool.

Results: The SDY cohort was 42% European, 30% African, 17% Hispanic, and 11% with mixed ancestries, and 39% female. Six percent of the cohort was found to harbor a pathogenic or likely pathogenic genetic variant in an epilepsy, cardiomyopathy or arrhythmia gene. The genomes of SDY cases, but not controls, were enriched for rare, damaging variants in epilepsy, cardiomyopathy and arrhythmia-related genes. A greater number of rare epilepsy genetic variants correlated with younger age at death.

Conclusions: While damaging cardiomyopathy and arrhythmia genes are recognized contributors to SDY, we also observed an enrichment in epilepsy-related genes in the SDY cohort, and a correlation between rare epilepsy variation and younger age at death. These findings emphasize the importance of considering epilepsy genes when evaluating SDY.

Keywords: sudden death in the young, genome sequencing, epilepsy, arrhythmia, cardiomyopathy, gene burden

BACKGROUND

Sudden death in children has immediate and sustained medical and emotional consequences for affected families. Elucidating the etiology of sudden death can inform risk for surviving family members. The Sudden Death in the Young (SDY) Case Registry was created as a joint initiative between the National Institutes of Health (NIH) and the Centers for Disease Control and Prevention (CDC) to enable a population-based surveillance of SDY and facilitate research through the collection of clinical data and biospecimens (1). The SDY Case Registry engaged public health agencies in thirteen diverse US sites to collect information on children, ages 0-20 years, who died suddenly and unexpectedly. Data from the SDY Registry indicates that sudden death mortality rates are greater for infants less than 1 year (120/100000 live births) than for children 1-17 years (1.9/100000 children) (2). SDY can be attributed to primary cardiac causes including arrhythmias, cardiomyopathies, congenital heart defects, and vascular disease, as well as noncardiac causes such as epilepsy. While extensive work has established social and environmental factors that contribute to sudden infant death syndrome (SIDS), the genetic contributors to sudden unexpected infant death (SUID), sudden cardiac death, and sudden unexpected death in epilepsy (SUDEP) remain incompletely described.

There is increasing evidence that SDY has a genetic component and postmortem genetic screening can aid in identifying the cause of death. Testing of genes implicated in cardiac rhythm and function identifies a pathogenic or likely pathogenic (P/LP) variant in 10-25% of individuals with sudden unexplained death <40 years old (3, 4). Approximately 4% of SIDS may be due to clinically actionable genetic cardiac causes

(5). In a cohort of sudden unexpected death in pediatrics (SUPD), contributory genetic variants were identified in 11% of decedents, and the authors also found an excess burden of rare, damaging SUDEP gene variants compared to a control group (6). We previously used genome sequencing (GS) to analyze an SDY cohort (n=103) covering a larger age range (1-44 years) and found ~13% carried a pathogenic/likely pathogenic (P/LP) variant in an arrhythmia or cardiomyopathy gene. Younger decedents carried an excess of suspicious variants of uncertain significance (VUS) and P/LP variants compared to a control population. Furthermore, for decedents >2 years of age at death, a younger age was associated with harboring more rare cardiac variants (7). Here, we interrogated the genomes of 211 decedents (≤ 19 years old) ascertained through the NIH/CDC SDY Registry with a median age at death of 1 year, with an analysis of genes associated with epilepsy, cardiomyopathies and arrhythmias.

METHODS

SDY Case Registry. The SDY Case Registry was previously described (1). Briefly, the SDY Case Registry aggregated information from 13 participating states and jurisdictions in collaboration with public health agencies, including medical examiner offices. Clinical data were collected via the US National Child Death Review Case Reporting System. The Institutional Review Board (IRB) at the Office of Research Integrity at the Michigan Public Health Institute (MPHI) provided ethical approval for the study (initial approval 2015 with subsequent annual renewals). The MPHI was the central IRB, and for those jurisdictions who did not rely on the Central IRB, local IRBs provided approval for the study. Parents/next of kin provided written informed consent for research use of the

decedent's DNA and linked information with consent to broad, future, unspecified research on their child's DNA and linked information to create a research repository. The MPHI also served as the Data Coordinating Center (DCC) for the SDY Case Registry. Biospecimens for DNA extraction were obtained at autopsy. All data was deidentified for the analysis performed in this work.

Genome Sequence (GS) Analysis. Sequencing methods are detailed in the Supplement. Variants were called with either the Genome Analysis Tool Kit (GATK v3.3.0) using the MegaSeq pipeline or Sentieon joint variant calling software (8, 9, 10). Variant call files were annotated using SnpEff (11). Variants were annotated using ClinVar (accessed March 2022) and global and ancestry-specific allele frequencies using the Genome Aggregation Database (gnomAD) (12). M-CAP was used to annotate variant pathogenicity (13).

Genetic Evaluation of Ancestry. Ancestry was determined using principal component analysis (PCA) conducted using singular-value decomposition of ~5 million biallelic variants across the genome. PCAs were generated using the first 2 components for global ancestry estimates. Asian ancestry was determined using the 3rd PC. Analyses were performed using PLINK v1.9 and R v4.1.

Gene Panel Analysis. Four previously-established gene panels were analyzed (two epilepsy panels and two arrhythmia and cardiac function panels, eTables 2 and 3 in the Supplement): 1) The Early Infantile Epileptic Encephalopathy (EIEE)- Online Mendelian

Inheritance in Man (OMIM) panel (82 genes) derived from the OMIM phenotypic series for early infantile epileptic encephalopathy, a curated epilepsy list; 2) The Epilepsy gene panel (191 genes, overlaps with the EIEE panel), and is a curated list derived from the Invitae Epilepsy Panel (Invitae, San Francisco, CA); 3) CMAR1, CardioMyopathy and Arrhythmia (gene panel including 118 arrhythmia and cardiac genes previously described (7); and 4) CMAR2, a Pan-Cardiomyopathy Arrhythmia and Cardiomyopathy Comprehensive gene panel including 143 genes with overlap with the CMAR1 panel (14). Control genes were selected from uniformly distributed housekeeping genes (15, 16). To create burden-matched control gene lists, a burden ratio measurement was calculated for each gene in RefSeq by counting the number of rare variants ($MAF \leq 0.005$) found in the gnomAD database over the largest coding transcript of every gene divided by transcript length (12). The mean burden ratio and standard deviation from each list were matched by randomly sampling from the whole genome (excluding genes in the test list) (15, 17).

Artificial-Intelligence (AI)-Based Variant Prioritization. AI-based prioritization and scoring of candidate disease genes and diagnostic conditions were performed using GEM (18) in the commercially available Fabric Enterprise platform (Fabric Genomics, Oakland, CA). Briefly, GEM aggregates data from multiple sources and clinical data sets to identify pathogenic variants relevant to phenotype terms input by the user. The GEM pipeline requires a variant call file, affection status, and human phenotype ontology (HPO) terms for each individual. A \log_{10} Bayes-factor score (GEM score) was generated that calculated the extent of support for a given model using multiple lines of

evidence from open-source tools and databases. A GEM score ≥ 0.69 was used to define a likely deleterious genetic variant, based on the observation this threshold recovered 95% of true positive cases in a cohort of critically ill newborns undergoing rapid genome sequencing (18).

Control Cohort. The control cohort of 211 genome samples derived from the 1000 Genomes Project (<https://www.internationalgenome.org/>) that were variant called using the Sentieon pipeline described above against the GRCh37/hg19 human reference genome(19). This control cohort matched the ancestry and sex composition as the SDY cohort.

Enrichment Analysis. Since detailed clinical information was lacking from the SDY cohort, we used HPO terms expected to be associated with SDY: Sudden Cardiac Death (HP:0001645), Sudden Death (HP:0001699), Cardiac arrest (HP:0001695), Cardiomyopathy (HP:0001638), Abnormal QT interval (HP:0031547), and Seizure (HP:0001250). We also performed the analysis using a control root phenotype, Phenotypic Abnormality (HP:0000118) which was used as a normalization factor to control for ontology biases and was expected to behave similarly across different cohorts.

GEM analysis was performed for each HPO term with each iteration identifying potentially causal variants for the three phenotypes described. To assess enrichment between cases and the controls a resampling analysis was performed as previously described (17). Briefly, for each investigated gene list (CMAR1, CMAR2, EIEE-OMIM,

and Epilepsy Panels) of size N, 100,000 random samples of equal size were drawn from the 18,876 RefSeq genes and intersected with the corresponding GEM list of damaged genes producing the number of damaged genes in the resampled list. Fisher exact test was used to determine if there were significant differences in the number of genes between cases and controls using the permutation test to normalize the data across runs and produce an empirical p-value (17). Multiple comparison correction was performed using the False Discovery Rate (FDR) approach (20). Each null distribution was scaled and centered producing Z-scores to directly compare the p values of different gene sets.

Age at death burden analysis. Age at death dependence was correlated with the cumulative number of nonsynonymous variants in the Epilepsy or CMAR gene panels. Variants were binned by global gnomAD frequency [<0.001 , $0.001-0.01$, $0.01-0.1$, $0.1-0.25$, $0.25-0.5$]. Multivariate linear models adjusting for ancestry (PCs 1-6) using custom code based on the *lm* function in R 4.1 were regressed to test each bin. P-values were adjusted for multiple comparisons. An extreme value sensitivity test was performed for decedents at the ends of the rare variant distribution (gnomAD <0.001); p values were adjusted for ancestry PCs 1-6.

Variants of Uncertain Significance (VUS) Analysis. The number of nonsynonymous, rare variants (<0.001) found in gnomAD (V2) for each gene in the Epilepsy and CMAR1 gene panels was compiled. To determine expected number of variants in each gene for the gene set used (Epilepsy/CMAR), we created a multinomial model using the

observed number of missense variants in the same gene set in gnomAD. We estimated the probability of observing a variant in each gene in the gene set as the observed number of variants in gnomAD by dividing the total number of variants in the gene set in gnomAD.

RESULTS

The Sudden Death in the Young (SDY) Cohort. In total, 230 decedents from the SDY registry had full genome sequencing (**Figure 1**). Of the 230, 19 were excluded; 4 due to failed inclusion criteria and 15 due to failed sequencing where no additional sample was available. Of the remaining 211, 152 (72%) had partial phenotype data available from the SDY Case Registry (**Table 1**). The median age at death was 1 year [0.33, 0.96] (median, IQR), indicating an enriched sample of cases who died in the first year of life. Genetic ancestry analysis was used to classify SDY decedents as: 42% European, 30% African, 17% Hispanic, and 11% with mixed ancestries. The cohort was 61% male and 39% female (**Table 1**). The SDY Case Registry provided cause of death information by standardized case report forms. The cause of death was included for 152 decedents, including 104 whose death remained unexplained and 48 with a known cause of death reported. Detailed cause of death is noted in **eTable 1** in the Supplement with sudden unexplained death as the most common cited cause (43%).

Pathogenic/likely pathogenic variants in cardiac and epilepsy genes. To identify variants that may have contributed to SDY, we employed gene panels to filter for rare, high impact variants. We hypothesized that the most likely candidates for SDY would be epilepsy, and cardiomyopathy and arrhythmia gene variants. We identified

nonsynonymous genetic variants in 161 out of 191 genes for the Epilepsy panel and in 99 out of 118 of the CMAR1 genes (**eTable 2** and 3 in the Supplement). Variants were annotated using ACMG classification and ClinVar interpretations to designate variants as pathogenic (P), likely pathogenic (LP), or variant of uncertain significance (VUS). Within the two gene sets we identified P/LP variants in ~6% of decedents (11 of 211). In these 11 individuals, the same *TTR* variant appeared in 4 decedents (**eTable 4** in the Supplement). In the epilepsy panel, we identified 4 variants in 3 decedents, but these genes generally associate with autosomal recessive inheritance. Six individuals had at least one P or LP variant in genes from the CMAR1 panel. One variant was identified in the *FKRP* gene which associates with recessive disease. One individual had a variant in the *KCNH2* gene, which is found in both the Epilepsy and CMAR1 panels (**eTable 4** in the Supplement).

To identify potentially pathogenic variants in a broader set of genes without constraining to the Epilepsy and CMAR panels, we deployed GEM, an AI tool designed for whole genome-based diagnosis of Mendelian conditions (18), using HPO terms related to seizures and sudden cardiac death. Thirty-eight decedents (18%) harbored ClinVar-designated P/LP variants in genes linked to a Mendelian disorder (13 related to the HPO term seizure and 25 related to sudden cardiac death; **eTable 5** in the Supplement). Four of these decedents harbored compound heterozygous damaging variants in genes associated with recessive disorders, including *CFTR* (Cystic Fibrosis), *BCDHA* (Maple Syrup Urine Disease), *MPDZ* (Congenital Hydrocephalus-2), and *GDAP1* (Charcot-Marie-Tooth Disease, Type 4A).

Damaging Genetic Variation in Cardiac and Epilepsy Genes was Enriched in the SDY cohort. We compared the burden of damaging cardiac and epilepsy genes in SDY cases versus controls, using two CMAR and two Epilepsy gene panels (see Methods). A GEM score ≥ 0.69 was used to define a likely deleterious genetic variant (see Methods). We also used a known housekeeping gene list and a random sample of genes matched for similar variant burden (16). **Figure 2** provides the distributions for the random gene samples (gray bars) and the number of genes identified for enrichment in the gene lists (dark arrows) and the burden-matched controls (light arrows). This analysis revealed an increased burden of genes with damaging variants for both Epilepsy and CMAR1 panels when compared with the control gene set (**Figure 2A and B**, respectively). A similar pattern was seen when we analyzed the EIEE-OMIM and CMAR2 gene panels (**eFigure 1** in the Supplement). The Fisher exact-test p-values FDR adjusted are Epilepsy $p < 0.027$, EIEE-OMIM $p < 0.019$, CMAR1 $p < 0.001$, CMAR2 $p < 0.001$ (**eTable 6** in the Supplement).

To compare the findings among the distributions shown in **Figure 2A and B**, each distribution was normalized by Z-score transformation (**Figure 2C**). The standardized Z-score values of the enrichment observed in the gene panels (**Figure 2C**) exceeded the expected normal distribution (shown in black), indicating a strong enrichment for damaging variation in Epilepsy and CMAR genes in the SDY cohort (+8.0-+13.0 standard deviations).

Rare Variants in Epilepsy Genes Correlated with Age at Death. The enrichment of predicted damaging variation in Epilepsy and CMAR genes and the few decedents with

P/LP variants, led us to determine if there is an association between rare Epilepsy and CMAR variants and age at death. We aggregated variants across five gnomAD allele frequency bins (<0.001 , $0.001-0.01$, $0.01-0.1$, $0.1-0.25$, $0.25-0.5$) for the Epilepsy and CMAR1 gene panels. We found that having more rare, nonsynonymous variants in Epilepsy genes was associated with younger age at death. Nonsynonymous Epilepsy variants with an allele frequency <0.001 were significantly associated with age at death, ($p=0.0053$, adjusted for ancestry) (**Figure 3A** and **eTable 7** in the Supplement). Results remained significant ($p=0.026$) following correction for multiple hypothesis testing for all 5 frequency bins. To confirm that these data were not driven by a few outliers, we performed a sensitivity analysis that supported the results of the main analysis (**eTable 7** in the Supplement). A similar analysis of rare, nonsynonymous variants in the CMAR1 gene panel did not show an association for age at death in any frequency bin (**Figure 3B**).

Rare VUS in Epilepsy and Cardiac Genes were enriched in the SDY cohort. We next considered if variants of uncertain significance (VUS) might associate with SDY. VUS are rare variants about which there is insufficient information as to whether they are pathogenic or benign, and the major determinant of VUS status is rare population frequency. **Figure 4** plots the number of rare, nonsynonymous variants (<0.001 gnomAD allele frequency) not reported as pathogenic or benign in ClinVar. Variants were scored using the *in silico* predictor M-CAP and scored as damaging using an M-CAP score >0.025 . We plotted total variants, damaging variants, expected number of variants based on gnomAD distribution and age at death (3rd quartile) of each decedent,

and we found that genes in both Epilepsy and CMAR1 panels had greater than expected numbers of variants. Some genes had very few damaging variants, while others had more than expected. Those genes with greater than expected variation may potentially drive or modify risk for SDY.

DISCUSSION

This SDY cohort is unique for its very young age (mean age 1 year) and its diverse ancestry. In the SDY cohort, we identified P/LP variants in Epilepsy and CMAR genes in ~6% of the cohort. An unrestrained AI-driven analysis identified 42 P/LP variants in 38 decedents suggesting that a purely gene panel-based approach may be diagnostically limited. Previous studies identified that the prevalence of P/LP gene variants is less in younger cohorts (3, 4, 5, 21), consistent with findings in our young cohort. In adults, the prevalence is likely 15-20%; whereas, in infants, our data agrees with others that the prevalence is < 10% when cardiomyopathy and arrhythmia genes are considered with a panel approach. Analysis of 278 SIDS cases for ultrarare variants in noncardiac, SIDS-susceptibility genes did not identify a monogenic basis for SIDS, even when using pathway burden analyses (22).

A recent study of SDY examined decedents >1 to 44 years of age, and specifically excluded the group under 1 year of age (7). A similar burden analysis found more CMAR1 gene variants and an association between gene variant burden and younger age at death (7). There is evidence that epilepsy-related mechanisms associate with SDY pathogenesis (23, 24). This younger SDY cohort, with a median age of 1 year at time of death, displayed enrichment of both epilepsy and cardiac

damaging variants, but only the epilepsy gene variant burden correlated with younger age at death. These data are in line with another study showing an increased burden in epilepsy, cardiac and metabolic genes with an excess of rare variants in all three gene classes in SUDP cases ≤ 1 year old (6).

CONCLUSIONS

Together, these findings support epilepsy etiologies for SDY, particularly in infants with a broader genetic contribution including cardiac and epilepsy variants in those >1 year of age. Even more relevant, the genetic associations for SDY appear to derive less from single gene P/LP variants and instead correlate with an aggregation of potentially damaging variants within an individual genome that predispose to SDY.

LIST OF ABBREVIATIONS

SDY: Sudden death in the young
NIH: National Institutes of Health
CDC: Center for Disease Control
SIDS: Sudden infant death syndrome
SUID: Sudden unexpected infant death
SUDEP: Sudden unexplained death in epilepsy
P/LP: Pathogenic or likely pathogenic
SUDP: Sudden unexpected death in pediatrics
GS: Genome sequencing
VUS: Variant of unknown significance

gnomAD: Genome aggregation database

PCA: Principal component analysis

EIEE: Early infantile epileptic encephalopathy

OMIM: Online Mendelian inheritance in man

CMAR1: CardioMyopathy and Arrhythmia panel 1

CMAR2: CardioMyopathy and Arrhythmia panel 2

AI: Artificial intelligence

HPO: Human phenotype ontology

FDR: False discovery rate

DECLARATIONS

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

We attest that data included in this manuscript was conducted in a manner consistent with principles of research ethics per the standards of the Belmont Report. This research was performed with the voluntary, informed consent of the next of kin of the decedent, free from coercion. Registry activities involving biospecimen collection and consent of surviving family members for research were approved by the institutional review boards at the Data Coordinating Center and participating states/jurisdictions.

CONSENT FOR PUBLICATION

NA

AVAILABILITY OF DATA AND MATERIALS

Both phenotype (where available) and sequence data will be available via dbGAP after publication or by reasonable request to corresponding author. Relevant data and scripts will be made available upon request to the corresponding author.

COMPETING INTERESTS

MY serves as consultant to Fabric Genomics Inc. and has received consulting fees and stock grants from Fabric Genomics Inc. EMM is a consultant for Amgen, Avidity, AstraZeneca, Cytokinetics, PepGen, Pfizer, Tenaya Therapeutics, Stealth BioTherapeutics, Invitae, and is the founder of Ikaika Therapeutics.

FUNDING

This work was supported by grants from the National Institutes of Health (U01HL131914, U01HL131698, HL128075, and U01HL131911), the American Heart Association Strategically Focused Research Network on Arrhythmia and Sudden Cardiac Death and Career Development Award (MJP). This study was made possible by the Sudden Death in the Young Case Registry, a collaboration between the National Heart, Lung, and Blood Institute and National Institute of Neurologic Disorders and Stroke of the National Institutes of Health and the Centers for Disease Control and Prevention, and its Data Coordinating Center at the Michigan Public Health Institute and biorepository at the University of Michigan.

AUTHOR CONTRIBUTIONS

MJP, GW, LMDC, KMB, MTF, ALG, and EMM conceptualized the project. MJP, LMM, EJH, GW, SDK, MY developed methods including scripts for data analysis. MJP, LLP, GW, LMDC, SDK, FKF, HM, KMB participated in data curation. MJP, LLP, EJH, performed formal analysis of the data. MJP, LLP, EJH, GW, LMDC, FKF, HM, KMB conducted experiments through data collection and in silico analysis. TOM, TDP validated data and experiments. MJP was responsible for the first draft of the manuscript. LLP, EJH, GW, LMDC, SPE, PJK, MBS, DF, DMR, KMB, HM, MY, MTF, ALG and EMM carefully reviewed data and experiments and edited the manuscript. PJK, HM, KMB coordinated collection of decedent samples and information. MJP, PJK, MTF, ALG and EMM secured funding.

ACKNOWLEDGEMENTS

The views expressed in this manuscript are those of the authors and do not reflect official positions of the National Heart, Lung, and Blood Institute, the National Institutes of Health, or the United States Department of Health and Human Services.

This work used Stampede2 at The Texas Advanced Computing Center (TACC) through allocation XRAC - BIO220012 (PI MJP) from the Extreme Science and Engineering Discovery Environment (XSEDE), which was supported by National Science Foundation grant number #1548562 (25). We thank Victor EijKhout for his assistance with porting and optimization, which was made possible through the XSEDE Extended Collaborative Support Service (ECSS) program (26). We also thank Justin Wozniak for his support in installing and optimizing our swift-t workflow (27).

449

450 *The information that is the basis of this presentation and/or publication was provided by*

451 *the NCRPCD/DCC, which is funded in part by the US Centers for Disease Control and*

452 *Prevention Division of Reproductive Health (CDC) and the Health Resources and*

453 *Services Administration. The data is part of the Sudden Death in the Young (SDY) Case*

454 *Registry, which is funded by the National Heart, Lung, and Blood Institute and the*

455 *National Institute of Neurologic Disorders and Stroke of the National Institutes of Health*

456 *(NIH) and the Centers for Disease Control and Prevention. The contents are solely the*

457 *responsibility of the authors and do not necessarily represent the official views of*

458 *NCRPCD, CDC, NIH or the participating states.*

459

460

461

462

463

464

465

466

467

468

469

470

471

472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513

REFERENCES

1. Burns KM, Cottengim C, Dykstra H, Faulkner M, Lambert ABE, MacLeod H, et al. Epidemiology of Sudden Death in a Population-Based Study of Infants and Children. *J Pediatr X*. 2020;2:100023.
2. Burns KM, Cottengim C, Dykstra H, Faulkner M, Lambert ABE, MacLeod H, et al. Epidemiology of Sudden Death in a Population-Based Study of Infants and Children. *J Pediatr X*. 2020;2.
3. Lahrouchi N, Raju H, Lodder EM, Papatheodorou E, Ware JS, Papadakis M, et al. Utility of Post-Mortem Genetic Testing in Cases of Sudden Arrhythmic Death Syndrome. *J Am Coll Cardiol*. 2017;69(17):2134-45.
4. Bagnall RD, Weintraub RG, Ingles J, Duflou J, Yeates L, Lam L, et al. A Prospective Study of Sudden Cardiac Death among Children and Young Adults. *N Engl J Med*. 2016;374(25):2441-52.
5. Tester DJ, Wong LCH, Chanana P, Jaye A, Evans JM, FitzPatrick DR, et al. Cardiac Genetic Predisposition in Sudden Infant Death Syndrome. *J Am Coll Cardiol*. 2018;71(11):1217-27.
6. Koh HY, Haghighi A, Keywan C, Alexandrescu S, Plews-Ogan E, Haas EA, et al. Genetic Determinants of Sudden Unexpected Death in Pediatrics. *Genet Med*. 2022;24(4):839-50.
7. Webster G, Puckelwartz MJ, Pesce LL, Dellefave-Castillo LM, Vanoye CG, Potet F, et al. Genomic Autopsy of Sudden Deaths in Young Individuals. *JAMA Cardiol*. 2021;6(11):1247-56.
8. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-303.
9. Puckelwartz MJ, Pesce LL, Nelakuditi V, Dellefave-Castillo L, Golbus JR, Day SM, et al. Supercomputing for the parallelization of whole genome analysis. *Bioinformatics*. 2014;30(11):1508-13.
10. Freed D, Aldana R, Weber JA, Edwards JS. The Sentieon Genomics Tools - A fast and accurate solution to variant calling from next-generation sequence data. *bioRxiv*. 2017:115717.
11. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6(2):80-92.
12. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-43.

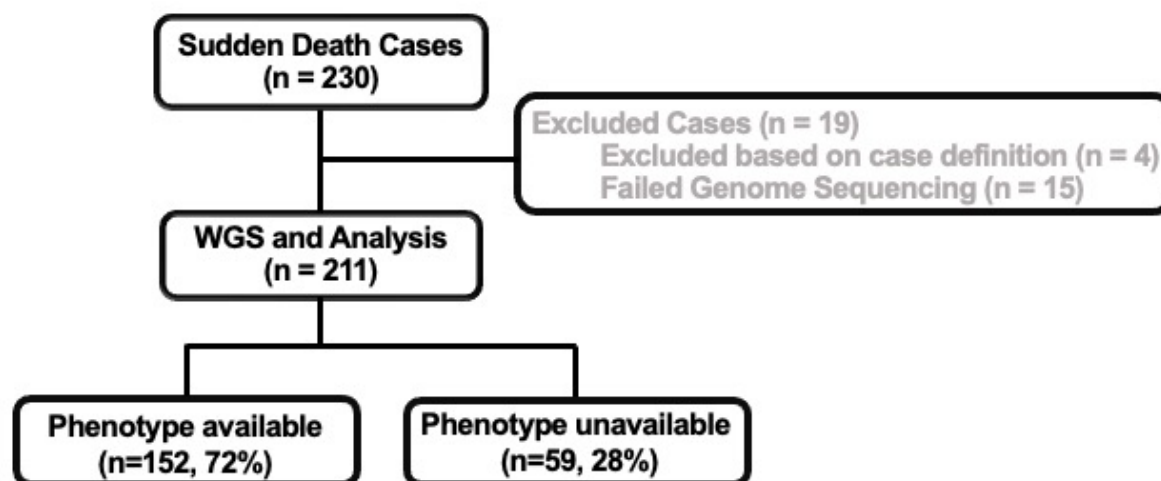
13. Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN, et al. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat Genet.* 2016;48(12):1581-6.
14. Dellefave-Castillo LM, Cirino AL, Callis TE, Esplin ED, Garcia J, Hatchell KE, et al. Assessment of the Diagnostic Yield of Combined Cardiomyopathy and Arrhythmia Genetic Testing. *JAMA Cardiol.* 2022;7(9):966-74.
15. Eilbeck K, Quinlan A, Yandell M. Settling the score: variant prioritization and Mendelian disease. *Nat Rev Genet.* 2017;18(10):599-612.
16. Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet.* 2013;29(10):569-74.
17. Watkins WS, Hernandez EJ, Wesolowski S, Bisgrove BW, Sunderland RT, Lin E, et al. De novo and recessive forms of congenital heart disease have distinct genetic and phenotypic landscapes. *Nat Commun.* 2019;10(1):4722.
18. De La Vega FM, Chowdhury S, Moore B, Frise E, McCarthy J, Hernandez EJ, et al. Artificial intelligence enables comprehensive genome interpretation and nomination of candidate diagnoses for rare genetic diseases. *Genome Med.* 2021;13(1):153.
19. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74.
20. Benjamini Y, Hochberg Y. On the Adaptive Control of the False Discovery Rate in Multiple Testing with Independent Statistics. *Journal of Educational and Behavioral Statistics.* 2000;25(1):60-83.
21. Steinberg C, Padfield GJ, Champagne J, Sanatani S, Angaran P, Andrade JG, et al. Cardiac Abnormalities in First-Degree Relatives of Unexplained Cardiac Arrest Victims: A Report From the Cardiac Arrest Survivors With Preserved Ejection Fraction Registry. *Circ Arrhythm Electrophysiol.* 2016;9(9).
22. Perrone S, Lembo C, Moretti S, Prezioso G, Buonocore G, Toscani G, et al. Sudden Infant Death Syndrome: Beyond Risk Factors. *Life (Basel).* 2021;11(3).
23. Brownstein CA, Goldstein RD, Thompson CH, Haynes RL, Giles E, Sheidley B, et al. SCN1A variants associated with sudden infant death syndrome. *Epilepsia.* 2018;59(4):e56-e62.
24. Halvorsen M, Petrovski S, Shellhaas R, Tang Y, Crandall L, Goldstein D, et al. Mosaic mutations in early-onset genetic diseases. *Genet Med.* 2016;18(7):746-9.
25. Towns J, Cockerill T, Dahan M, Foster I, Gaither K, Grimshaw A, et al. XSEDE: Accelerating scientific discovery. *Computing in Science and Engineering.* 2014;16(5):62-74.
26. Wilkins-Diehr N, Sanielevici S, Alameda J, Cazes J, Crosby L, Pierce M, et al., editors. An overview of the XSEDE extended collaborative support program. *Communications in Computer and Information Science*; 2016.
27. Wozniak JM, Armstrong TG, Wilde M, Katz DS, Lusk E, Foster IT, editors. Swift/T: Scalable data flow programming for many-task applications. *Proceedings of the ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, PPOPP*; 2013.
28. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-60.

FIGURES

Table 1. Demographics of study cohort (n=211)

Phenotype Information	Yes	No	
n (%)	152 (72)	59 (28)	
Age at death (median [IQR])	1 [0.33,0.96]	NA	
Genetic Ancestry (%)			
African	50 (32.9)	13 (22)	
European	68 (44.7)	21(35.6)	
Hispanic	21 (13.8)	15(25.4)	
Mixed	13 (8.6)	10(16.9)	
Genetic Sex (% male)	92 (60.5)	36 (61.0)	

594



595

596

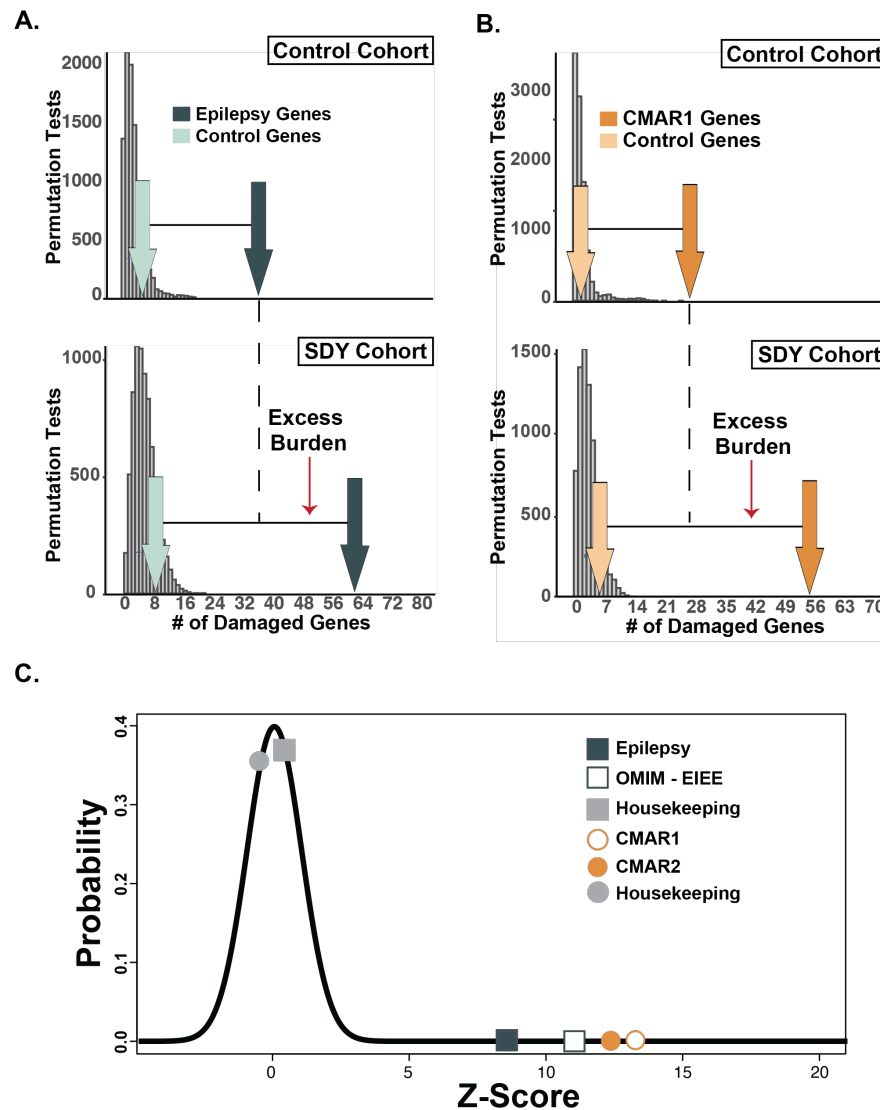
597

598

599

Figure 1. CONSORT diagram of decedents selected for genome sequencing. GS = genome sequencing.

600



601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

Figure 2. Relative enrichment of damaging variants in epilepsy and Cardiomypopathy/Arrhythmia (CMAR) genes. The SDY cohort has enriched damaging (A) epilepsy and (B) cardiac gene burden compared to a sex- and ancestry-matched control cohort. Histograms (gray bars) represent distributions of damaged genes (GEM Score > 0.69) sampled randomly from RefSeq genes using a root phenotype from the SDY cohort (n=211) (bottom panels) and 1000 Genomes Project Cohort (control) matched for sex and ancestry (n=211) (top panels). Damaged genes identified in the Epilepsy (n=191 genes, green) (A) and CMAR1 (n=118 genes, orange) (B) gene lists were significantly different between the SDY and control cohorts (dark arrows, epilepsy, $p=0.027$; cardiac $p<0.001$). Light arrows represent the number of damaged genes identified in a control housekeeping gene set. (C) The results in Figure 1A and B and eFigure 1 were Z-score transformed to make the gene list findings comparable. The plot reveals a considerable enrichment in both cardiac and epilepsy gene lists (+8.0 - +13.0 SDs). Black line represents the expected normal distribution.

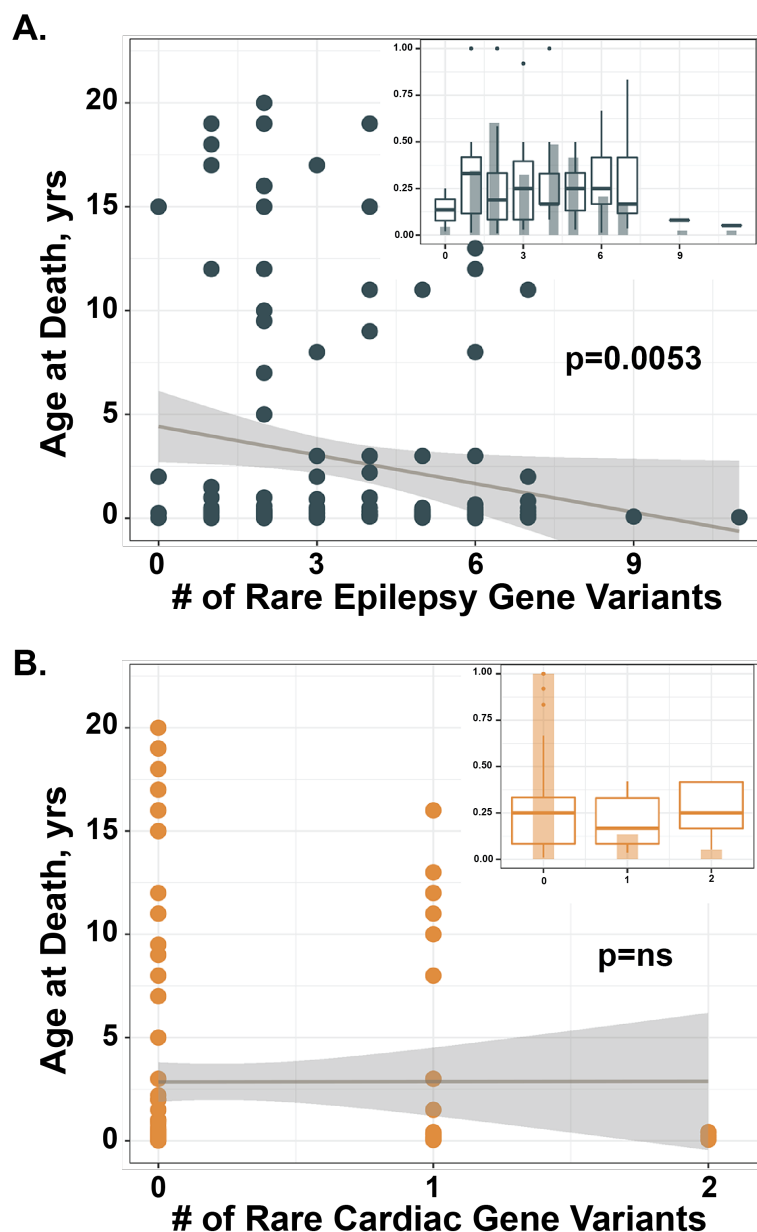


Figure 3. Number of rare, nonsynonymous, epilepsy gene variants correlated with younger age at death. The age at death was plotted against against the number of rare variants (gnomAD allele frequency <0.001) in either the (A) Epilepsy or (B) CMAR1 gene list. Number of variants in the epilepsy gene list significantly correlated with age at death, (p=0.0053, adjusted for the first 6 principal components of ancestry). A similar analysis of rare, predicted damaging variation in the CMAR1 genes did not show association with age at death (p=0.85, p value adjusted for ancestry). Inset is a histogram showing number of decedents <1 year old by variant burden, box plots (black) represent decedent age distribution for each variant burden. Dark gray line = regression line of unadjusted model; dark gray shading represents 95% confidence intervals.

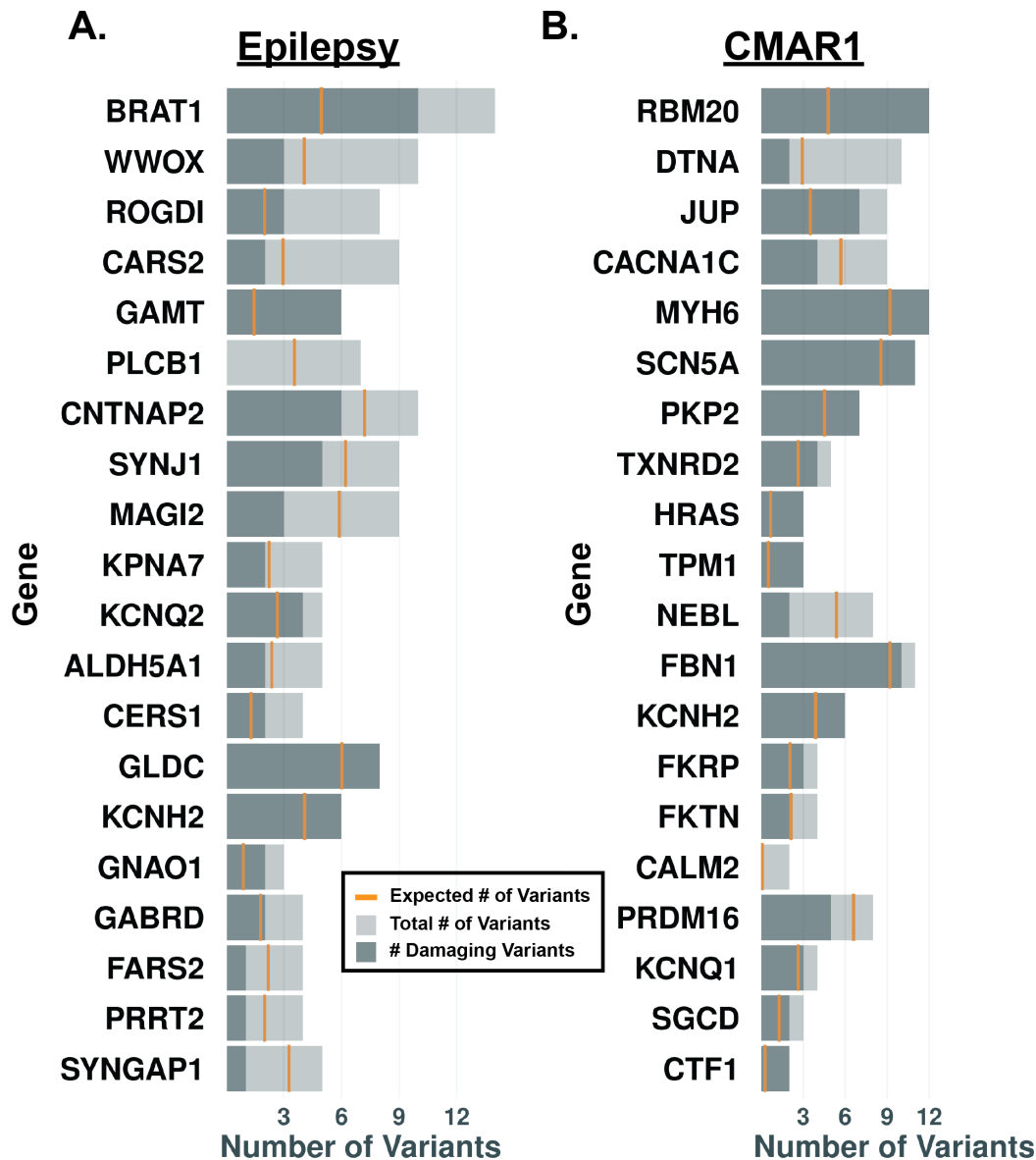


Figure 4. Rare VUS in Epilepsy and CMAR1 Genes. The number of rare (<0.001 gnomAD allele frequency) variants not classified as pathogenic or benign in ClinVar are reported for genes from either the (A) Epilepsy or (B) CMAR1 gene list (light gray bars). Variants were scored using the *in silico* tool M-CAP. The number of damaging VUS for each gene (>0.025 M-CAP score) is shown by the dark gray bars. The gold line indicates the expected number of variants based on gnomAD allele frequency data. Gene lists are ranked by the difference between the observed and expected number of variants, when ratios are approximately the same, ranking is determined by number of damaging gene variants.

SUPPLEMENT

Supplemental Methods

WGS was performed on DNA obtained from the SDY Case Registry using an Illumina XTen sequencer (Garvan Institute of Medical Research NSW, Australia) with a yield of >100 GB sample, correlating to >30-fold coverage across the genome. The Burrows-Wheeler Aligner was used to align reads to human reference sequence GRCh37/hg19 (28).

Supplemental Table 1. Detailed Cause of Death

Cause of Death	Number of Decedents (%)
Unexplained (n=104, 49%)	
Sudden unexplained death	90 (43)
Possible SUDEP	6 (3)
Cardiac findings of uncertain significance	5 (2)
Concerning family history	3 (1)
Explained (n=48, 23%)	
Infant suffocation	15 (7)
Systemic viral or bacterial infection	13 (6)
Asthma	3 (1)
Cardiomyopathy	3 (1)
Drowning	3 (1)
Multisystem failure in complex medical patient	3 (1)
Major vascular malformation	2 (1)
Myocarditis	2 (1)
Congenital heart disease	2 (1)
Diabetic ketoacidosis	1 (<1)
Pulmonary embolism	1 (<1)
No phenotype information available	59 (28)
TOTAL	211

SUDEP = Sudden Unexplained Death in Epilepsy. Cardiac findings of uncertain significance included findings of cardiomyopathy without sufficient criteria for diagnosis (2 decedents), possible coronary artery disease (2), and aortic valve abnormality with left ventricular hypertrophy (1). Concerning family history included long QT syndrome, sudden unexplained death in a sibling, and family history of heart attack and stroke < 50 years of age, each in 1 decedent. Congenital heart disease included hypoplastic left heart syndrome (1) and anomalous left coronary artery (1).

663 **Supplemental Table 2. Epilepsy Gene panels**

Gene	Epilepsy n=191	OMIM EIEE n=82
<i>AARS1</i>		X
<i>ABAT</i>	X	
<i>ACTL6B</i>		X
<i>ADAM22</i>		X
<i>ADSL</i>	X	
<i>ALDH5A1</i>	X	
<i>ALDH7A1</i>	X	
<i>ALG13</i>	X	X
<i>AMT</i>	X	
<i>AP3B2</i>		X
<i>ARHGEF9</i>	X	X
<i>ARHGEF15</i>	X	X
<i>ARV1</i>		X
<i>ARX</i>	X	X
<i>ATP1A2</i>	X	
<i>ATP1A3</i>	X	
<i>ATP6AP2</i>	X	
<i>ATRX</i>	X	
<i>BRAT1</i>	X	
<i>C12orf57</i>	X	
<i>CACNA1A</i>	X	X
<i>CACNA1E</i>		X
<i>CACNA1H</i>	X	
<i>CACNA2D2</i>	X	
<i>CACNB4</i>	X	
<i>CAD</i>		X
<i>CARDS2</i>	X	
<i>CASK</i>	X	
<i>CASR</i>	X	
<i>CDKL5</i>	X	X
<i>CERS1</i>	X	
<i>CHD2</i>	X	
<i>CHRNA2</i>	X	
<i>CHRNA4</i>	X	
<i>CHRNA7</i>	X	
<i>CHRNA2</i>	X	
<i>CLCN4</i>	X	
<i>CLN2</i>	X	
<i>CLN3</i>	X	
<i>CLN5</i>	X	

<i>CLN6</i>	X	
<i>CLN8</i>	X	
<i>CNTN2</i>	X	
<i>CNPY3</i>		X
<i>CNTNAP2</i>	X	
<i>CPA6</i>	X	
<i>CPLX1</i>		X
<i>CSTB</i>	X	
<i>CTSD</i>	X	
<i>CUX2</i>		X
<i>CYFIP2</i>		X
<i>DENND5A</i>		X
<i>DEPDC5</i>	X	
<i>DIAPH1</i>	X	
<i>DNAJC5</i>	X	
<i>DNM1</i>	X	X
<i>DOCK7</i>	X	X
<i>DYRK1A</i>	X	
<i>EEF1A2</i>	X	X
<i>EFHC1</i>	X	
<i>EHMT1</i>	X	
<i>EPM2A</i>	X	
<i>FARS2</i>	X	
<i>FASN</i>	X	
<i>FGF12</i>		X
<i>FLNA</i>	X	
<i>FOLR1</i>	X	
<i>FOXG1</i>	X	
<i>FRRS1L</i>	X	X
<i>GABBR2</i>	X	X
<i>GABRA1</i>	X	X
<i>GABRA2</i>	X	X
<i>GABRB3</i>	X	X
<i>GABRA5</i>		X
<i>GABRB1</i>		X
<i>GABRB3</i>		X
<i>GABRD</i>	X	
<i>GABRG2</i>	X	X
<i>GAL</i>	X	
<i>GAMT</i>	X	
<i>GATM</i>	X	
<i>GCSH</i>	X	

<i>GLDC</i>	X	
<i>GLRA1</i>	X	
<i>GLS</i>		X
<i>GNAO1</i>	X	X
<i>GOSR2</i>	X	
<i>GPHN</i>	X	
<i>GRIN1</i>	X	
<i>GRIN2A</i>	X	
<i>GRIN2B</i>	X	X
<i>GRIN2D</i>		X
<i>GUF1</i>		X
<i>HCN1</i>	X	X
<i>HNRNPU</i>	X	X
<i>IER3IP1</i>	X	
<i>IQSEC2</i>	X	
<i>ITPA</i>	X	X
<i>JMJD1C</i>	X	
<i>KANSL1</i>	X	
<i>KCNA1</i>	X	
<i>KCNA2</i>	X	X
<i>KCNB1</i>	X	X
<i>KCNC1</i>	X	
<i>KCND2</i>	X	
<i>KCNH2</i>	X	
<i>KCNH5</i>	X	
<i>KCNJ10</i>	X	
<i>KCNMA1</i>	X	
<i>KCNQ2</i>	X	X
<i>KCNQ3</i>	X	
<i>KCNT1</i>	X	
<i>KCNT2</i>		X
<i>KCTD7</i>	X	
<i>KPNA7</i>	X	
<i>LGI1</i>	X	
<i>LIAS</i>	X	
<i>LMNB2</i>	X	
<i>MAGI2</i>	X	
<i>MBD5</i>	X	
<i>MDH2</i>		X
<i>MECP2</i>	X	
<i>MEF2C</i>	X	
<i>MFSD8</i>	X	

<i>MTOR</i>	X	
<i>NECAP1</i>	X	X
<i>NEDD4L</i>	X	
<i>NEUROD2</i>		X
<i>NEXMIF</i>	X	
<i>NGLY1</i>	X	
<i>NHLRC1</i>	X	
<i>NPRL3</i>	X	
<i>NR2F1</i>	X	
<i>NRXN1</i>	X	
<i>NTRK2</i>		X
<i>PACS1</i>	X	
<i>PACS2</i>		X
<i>PARS2</i>		X
<i>PCDH19</i>	X	X
<i>PHACTR1</i>		X
<i>PIGA</i>	X	X
<i>PIGB</i>		X
<i>PIGG</i>	X	
<i>PIGN</i>	X	
<i>PIGO</i>	X	
<i>PIGP</i>		X
<i>PIGQ</i>	X	X
<i>PIGV</i>	X	
<i>PIK3AP1</i>	X	
<i>PLCB1</i>	X	X
<i>PNKD</i>	X	
<i>PNKP</i>	X	X
<i>PNPO</i>	X	
<i>POLG</i>	X	
<i>PPT1</i>	X	
<i>PRDM8</i>	X	
<i>PRICKLE1</i>	X	
<i>PRICKLE2</i>	X	
<i>PRIMA1</i>	X	
<i>PRRT2</i>	X	
<i>PTEN</i>	X	
<i>PURA</i>	X	
<i>QARS</i>	X	
<i>RANBP2</i>	X	
<i>RBFOX1</i>	X	
<i>RBFOX3</i>	X	

<i>RELN</i>	X	
<i>RHOBTB2</i>		X
<i>RNF13</i>		X
<i>ROGD1</i>	X	
<i>RYR3</i>	X	
<i>SATB2</i>	X	
<i>SCARB2</i>	X	
<i>SCN1A</i>	X	X
<i>SCN1B</i>	X	X
<i>SCN2A</i>	X	X
<i>SCN3A</i>	X	X
<i>SCN8A</i>	X	X
<i>SCN9A</i>	X	X
<i>SERPINI1</i>	X	
<i>SETD2</i>	X	
<i>SGCE</i>	X	
<i>SIK1</i>	X	X
<i>SCL12A5</i>	X	X
<i>SLC13A5</i>	X	X
<i>SLC19A3</i>	X	
<i>SLC1A2</i>		X
<i>SLC25A12</i>	X	X
<i>SLC25A22</i>	X	X
<i>SLC2A1</i>	X	
<i>SLC35A2</i>	X	X
<i>SLC35A3</i>	X	
<i>SLC6A1</i>	X	
<i>SLC6A8</i>	X	
<i>SLC9A6</i>	X	
<i>SMC1A</i>	X	
<i>SNAP25</i>	X	
<i>SNX27</i>	X	
<i>SPATA5</i>	X	
<i>SPTAN1</i>	X	X
<i>SRPX2</i>	X	
<i>ST3GAL3</i>	X	X
<i>ST3GAL5</i>	X	
<i>STRADA</i>	X	
<i>STX1B</i>	X	
<i>STXBP1</i>	X	X
<i>SYN1</i>	X	
<i>SYNGAP1</i>	X	X

<i>SYNJ1</i>	X	X
<i>SZT2</i>	X	X
<i>TBC1D24</i>	X	X
<i>TBL1XR1</i>	X	
<i>TCF4</i>	X	X
<i>TPK1</i>	X	
<i>TPP1</i>	X	
<i>TRAK1</i>		X
<i>TSC1</i>	X	
<i>TSC2</i>	X	
<i>UBA5</i>		X
<i>UBE3A</i>	X	
<i>WDR45</i>	X	
<i>WWOX</i>	X	X
<i>YWHAG</i>		X
<i>ZDHHC9</i>	X	
<i>ZEB2</i>	X	

664
665
666
667
668
669

670 **Supplemental Table 3. Cardiomyopathy and Arrhythmia Gene panels**

Gene	CMAR1 n=118	CMAR2 n=143
<i>A2ML1</i>	X	
<i>ABCC9</i>	X	X
<i>ACTC1</i>	X	X
<i>ACTN2</i>	X	X
<i>AGL</i>		X
<i>AKAP9</i>	X	X
<i>ALMS1</i>	X	
<i>ANK2</i>	X	X
<i>ANKRD1</i>	X	X
<i>BAG3</i>	X	X
<i>BRAF</i>	X	X
<i>CACNA1C</i>	X	X
<i>CACNA2D1</i>	X	X
<i>CACNB2</i>	X	X
<i>CALM1</i>	X	X
<i>CALM2</i>	X	X
<i>CALM3</i>	X	X
<i>CALR3</i>		X
<i>CASQ2</i>	X	X
<i>CAV3</i>	X	X
<i>CBL</i>	X	
<i>CHRM3</i>	X	X
<i>CRYAB</i>	X	X
<i>CSRP3</i>	X	X
<i>CTF1</i>	X	X
<i>CTNNA3</i>	X	X
<i>DEPDC5</i>		X
<i>DES</i>	X	X
<i>DMD</i>	X	X
<i>DOLK</i>	X	X
<i>DSC2</i>	X	X
<i>DSG2</i>	X	X
<i>DSP</i>	X	X
<i>DTNA</i>	X	X
<i>EMD</i>	X	X
<i>EYA4</i>	X	X
<i>FBN1</i>	X	
<i>FHL1</i>	X	X
<i>FHL2</i>	X	X
<i>FKRP</i>	X	X

<i>FKTN</i>	<i>X</i>	<i>X</i>
<i>FLNC</i>	<i>X</i>	<i>X</i>
<i>GAA</i>		<i>X</i>
<i>GATA4</i>	<i>X</i>	<i>X</i>
<i>GATA6</i>	<i>X</i>	
<i>GATAD1</i>	<i>X</i>	<i>X</i>
<i>GJA5</i>		<i>X</i>
<i>GLA</i>	<i>X</i>	<i>X</i>
<i>GPD1L</i>	<i>X</i>	<i>X</i>
<i>HCN4</i>	<i>X</i>	<i>X</i>
<i>HRAS</i>	<i>X</i>	<i>X</i>
<i>ILK</i>	<i>X</i>	<i>X</i>
<i>JPH2</i>	<i>X</i>	<i>X</i>
<i>JUP</i>	<i>X</i>	<i>X</i>
<i>KCNA5</i>		<i>X</i>
<i>KCND3</i>	<i>X</i>	<i>X</i>
<i>KCNE1</i>	<i>X</i>	<i>X</i>
<i>KCNE2</i>	<i>X</i>	<i>X</i>
<i>KCNE3</i>	<i>X</i>	<i>X</i>
<i>KCNE5</i>	<i>X</i>	<i>X</i>
<i>KCNH2</i>	<i>X</i>	<i>X</i>
<i>KCNJ2</i>	<i>X</i>	<i>X</i>
<i>KCNJ5</i>	<i>X</i>	<i>X</i>
<i>KCNJ8</i>	<i>X</i>	<i>X</i>
<i>KCNK3</i>		<i>X</i>
<i>KCNQ1</i>	<i>X</i>	<i>X</i>
<i>KCNQ2</i>		<i>X</i>
<i>KCNQ3</i>		<i>X</i>
<i>KCNT1</i>		<i>X</i>
<i>KRAS</i>	<i>X</i>	<i>X</i>
<i>LAMA4</i>	<i>X</i>	<i>X</i>
<i>LAMP2</i>	<i>X</i>	<i>X</i>
<i>LDB3</i>	<i>X</i>	<i>X</i>
<i>LMNA</i>	<i>X</i>	<i>X</i>
<i>LRRC10</i>		<i>X</i>
<i>MAP2K1</i>	<i>X</i>	<i>X</i>
<i>MAP2K2</i>	<i>X</i>	<i>X</i>
<i>MOG1</i>	<i>X</i>	
<i>MTND1</i>		<i>X</i>
<i>MTND5</i>		<i>X</i>
<i>MTND6</i>		<i>X</i>
<i>MTTD</i>		<i>X</i>

<i>MTTG</i>		X
<i>MTTH</i>		X
<i>MTTI</i>		X
<i>MTTK</i>		X
<i>MTTL1</i>		X
<i>MTTL2</i>		X
<i>MTTM</i>		X
<i>MTTQ</i>		X
<i>MTTS1</i>		X
<i>MTTS2</i>		X
<i>MYBPC3</i>	X	X
<i>MYH6</i>	X	X
<i>MYH7</i>	X	X
<i>MYL2</i>	X	X
<i>MYL3</i>	X	X
<i>MYL4</i>	X	X
<i>MYLK2</i>	X	X
<i>MYOM1</i>	X	X
<i>MYOZ2</i>	X	X
<i>MYPN</i>	X	X
<i>NEBL</i>	X	X
<i>NEXN</i>	X	X
<i>NKX2-5</i>	X	X
<i>NPPA</i>		X
<i>NRAS</i>	X	X
<i>PCDH19</i>		X
<i>PDLIM3</i>	X	X
<i>PKP2</i>	X	X
<i>PLEKHM2</i>		X
<i>PLN</i>	X	X
<i>PRDM16</i>	X	X
<i>PRKAG2</i>	X	X
<i>PRRT2</i>		X
<i>PTPN11</i>	X	X
<i>RAF1</i>	X	X
<i>RANGRF</i>		X
<i>RASA1</i>	X	
<i>RBM20</i>	X	X
<i>RIT1</i>	X	
<i>RYR2</i>	X	X
<i>SCN10A</i>	X	X
<i>SCN1A</i>		X

<i>SCN1B</i>	<i>X</i>	<i>X</i>
<i>SCN2B</i>	<i>X</i>	
<i>SCN3B</i>	<i>X</i>	<i>X</i>
<i>SCN4B</i>	<i>X</i>	<i>X</i>
<i>SCN5A</i>	<i>X</i>	<i>X</i>
<i>SCN8A</i>		<i>X</i>
<i>SCN9A</i>		<i>X</i>
<i>SDHA</i>	<i>X</i>	
<i>SGCD</i>	<i>X</i>	<i>X</i>
<i>SHOC2</i>	<i>X</i>	
<i>SLC22A5</i>		<i>X</i>
<i>SLC2A1</i>		<i>X</i>
<i>SLMAP</i>		<i>X</i>
<i>SNTA1</i>	<i>X</i>	<i>X</i>
<i>SOS1</i>	<i>X</i>	<i>X</i>
<i>SPRED1</i>	<i>X</i>	
<i>TAZ</i>	<i>X</i>	<i>X</i>
<i>TCAP</i>	<i>X</i>	<i>X</i>
<i>TGFB3</i>	<i>X</i>	<i>X</i>
<i>TMEM43</i>	<i>X</i>	<i>X</i>
<i>TMPO</i>	<i>X</i>	<i>X</i>
<i>TNNC1</i>	<i>X</i>	<i>X</i>
<i>TNNI3</i>	<i>X</i>	<i>X</i>
<i>TNNT2</i>	<i>X</i>	<i>X</i>
<i>TPM1</i>	<i>X</i>	<i>X</i>
<i>TRDN</i>	<i>X</i>	<i>X</i>
<i>TRPM4</i>		<i>X</i>
<i>TTN</i>	<i>X</i>	<i>X</i>
<i>TTR</i>	<i>X</i>	<i>X</i>
<i>TXNRD2</i>	<i>X</i>	<i>X</i>
<i>VCL</i>	<i>X</i>	<i>X</i>

671
672

673
674 **Supplemental Table 4. Pathogenic and Likely Pathogenic Variants Identified in Epilepsy and Cardiac**
675 **Genes**

ID*	Age at Death (yrs)	Cause of Death	ClinVar	GENE	gDNA	AA Change	gnomAD
<u>Epilepsy Panel Gene Variants</u>							
1010	<1	Cardiac, congenital	LP	<i>QARS*</i>	3:49137655 G>A	p.Arg367Cys	3.19E-05
1172	<1	Infant Suffocation	LP	<i>CLN8*</i>	8:1719428 C>T	p.Arg70Cys	0
1193	10-20	Venous Malformation	P	<i>PPT1*</i>	1:40558081 T>G	p.Thr25Pro	0
	10-20	Venous Malformation	P/LP	<i>CSTB*</i>	21:45194641 C>G	n.104G>C	0.0003
<u>CMAR1 Panel Gene Variants</u>							
1001	<1	Unexplained	P	<i>FKRP*</i>	19:47259533 C>A	p.Leu276Ile	0.0014
1012	<1	Unexplained	P	<i>TTR</i>	18:29178618 G>A	p.Val142Ile	0.0048
1028	<1	Unexplained	P	<i>TTR</i>	18:29178618 G>A	p.Val142Ile	0.0048
1058	>1- <10	Unexplained	LP	<i>CALM3</i>	19:47112212 A>G	p.Asp96Gly	0
1117	<1	Infant Suffocation	P	<i>TTR</i>	18:29178618 G>A	p.Val142Ile	0.0048
1127	<1	Unexplained	P	<i>TTR</i>	18:29178618 G>A	p.Val142Ile	0.0048
1132	NA	NA	P	<i>DSP</i>	6:7571745 C>T	p.Gln611*	0
<u>Epilepsy and CMAR1 Panel Gene Variants</u>							
1214	NA	NA	P	<i>KCNH2</i>	7:150654393 C>CAG	p.Thr371_Glu372fs	0

676*All IDs are created for WGS purposes and not to reidentify subjects. Variants in this table were
677annotated with population frequency, protein effect, literature review, and *in silico* modeling.
678Pathogenicity was determined by methods similar to ACMG standards (see Methods). NA= Not
679reported in case review. P=Pathogenic; LP=Likely Pathogenic. *Gene is generally associated with
680autosomal recessive inheritance and decedent is heterozygous. AA= amino acid
681

Supplemental Table 5. Pathogenic and Likely pathogenic, Mendelian variants as ranked by GEMS

ID*	Age (yrs)	Cause of Death	CV	GENE	gDNA	AA Change	gnomAD Freq	MIM
<u>HPO Term: Seizure</u>								
1098	>1- <10	SUDEP	P	ANKRD11	16:89345623 G>A	Gln2443Ter	0	148050
1079	<1	UnE	P	BCKDHA	19:41916543 AC>A	Arg40GlyfsTer23	6.50E-05	248600
			VUS	BCKDHA	19:41916645 A>G	Asn71Ser	6.49E-05	248600
1199	<1	E	P	CFTR	7:117199645 TCTdel	Phe508del	0.0019518	219700
			P	CFTR	7:117188877 G>T	Lys464Asn	0.007261	219700
1048	10-20	E	P	DHDDS	1:26764719 A>G	Lys42Glu	0	617836
1172	<1	I.S>	P	GJB2	13:20763626 C>T	Arg32His	0	148210
1189	NA	NA	P	KISS1R	19:920746 T>A	Ter399Argext*?	0	176400
1139	>1- <10	E	P	MC4R	18:58038600 A>T	Leu328Ter	0	618406
1152	<1	NA	P	MPDZ	9:13126766 CT>C	Gln1490Arg fsTer19	0.000688	615219
			VUS	MPDZ	9:13175799 C>T	Glu1003Lys	0.000574	615219
1190	<1	E	LP	MYOCD	17:12655857 A>G	Ile418Val	0.002642	618719
1069	<1	I.S.	P	NEK1	4:170482671 C>T	Trp409Ter	6.48E-05	617892
1035	<1	UnE	P	SRPX2	X:99917224 A>C	Tyr72Ser	0	300643
1023	10-20	UnE	P	TOR1A	9:132576340 TCTCdel	Glu303Ter	0	128100
1107	NA	NA	LP	TREX1	3:48508721 G>A	Ala223Thr	0	192315
<u>HPO Term: SDY</u>								
1209	NA	NA	P	ANO5	11:22242647- -A	Asn64LysfsTer15	0	166260
1019	NA	NA	P	ANO5	11:22257823 G>A	splice donor	6.48E-05	166260
1166	NA	NA	P	BARD1	2:215595215 G>A	Arg641Ter	0	114480
1180	NA	NA	P	BARD1	2:215645402 A>C	Leu399Ter	0	114480
1058	10-20	UnE	LP	CALM3	19:47112212 A>G	Asp132Gly	0	618782
1203	NA	NA	P	CAPN3	15:42695962 TG>T	Asp591IlefsTer4	0	618129
1059	NA	NA	P	CAPN3	15:42703497 C>A	Ala798Glu	0	618129

1161	<1	UnE	P	CAPN3	15:42693950 G>A	Arg489Gln	0.000344	618129
1132	NA	NA	P	DSP	6:7571745 C>T	Gln611Ter	0	607450
1085	NA	NA	LP	F11	4:187201717 T>C	Leu373Ser	0	612416
1119	NA	NA	P	FGA	4:155508672 G>A	Arg168Ter	0	105200
1050	NA	NA	P	G6PD	X:153761337 C>T	Val321Met	0	300908
1026	<1	UnE	P	GDAP1	8:75275175 C>G	.Ser194Ter	0	214400
			VUS	GDAP1	8:75263650 A>G	Ile87Val	0	214400
1140	<1	I.S.	LP	ITGB3	17:45362012 C>T	Pro189Ser	0	187800
1214	NA	NA	P	KCNH2	7:150654393 C>CAG	Glu372Leu fsTer63	0	613688
1077	<1	UnE	LP	LDLR	19:11230828 G>A	Gly636Ser	0	143890
1133	NA	NA	P	LDLR	19:11218160 G>A	Asp304Asn	0	143890
1018	<1	UnE	P	NALCN	13:10194442 3 A>G	Ile322Thr	0	616266
1034	NA	NA	P	RYS1	19:38983277 G>A	splice donor	0	255310
1107	NA	NA	LP	SCN4A	17:62026868 A>T	Tyr958Ter	0	168300
1065	<1	NA	P	SPG7	16:89598370 G>GC	Gly352Arg fsTer44	0.000324	607259
1028	<1	UnE	P	TTR	18:29178618 G>A	Val142Ile	0.017218	105210
1117	<1	I.S.	P	TTR	18:29178618 G>A	Val142Ile	0.017218	105210
1038	<1	I.S>	P	VIM	10:17271434 TC>T	Val6CysfsTer26	0	116300
1010	<1	E, cardiac	LP	VWF	12:6080863 C>T	Val2484Ile	6.49E-05	193400

* All IDs are created for WGS purposes and not to reidentify subjects. NA= Not reported in case review. I.S. = infant suffocation; UnE = unexplained; E= explained; CV=ClinVar; AA = amino acid.

Supplemental Table 6. Enrichment of variants in epilepsy or cardiac genes in the SDY cohort compared to a ancestry and sex matched 1000 genomes cohort.

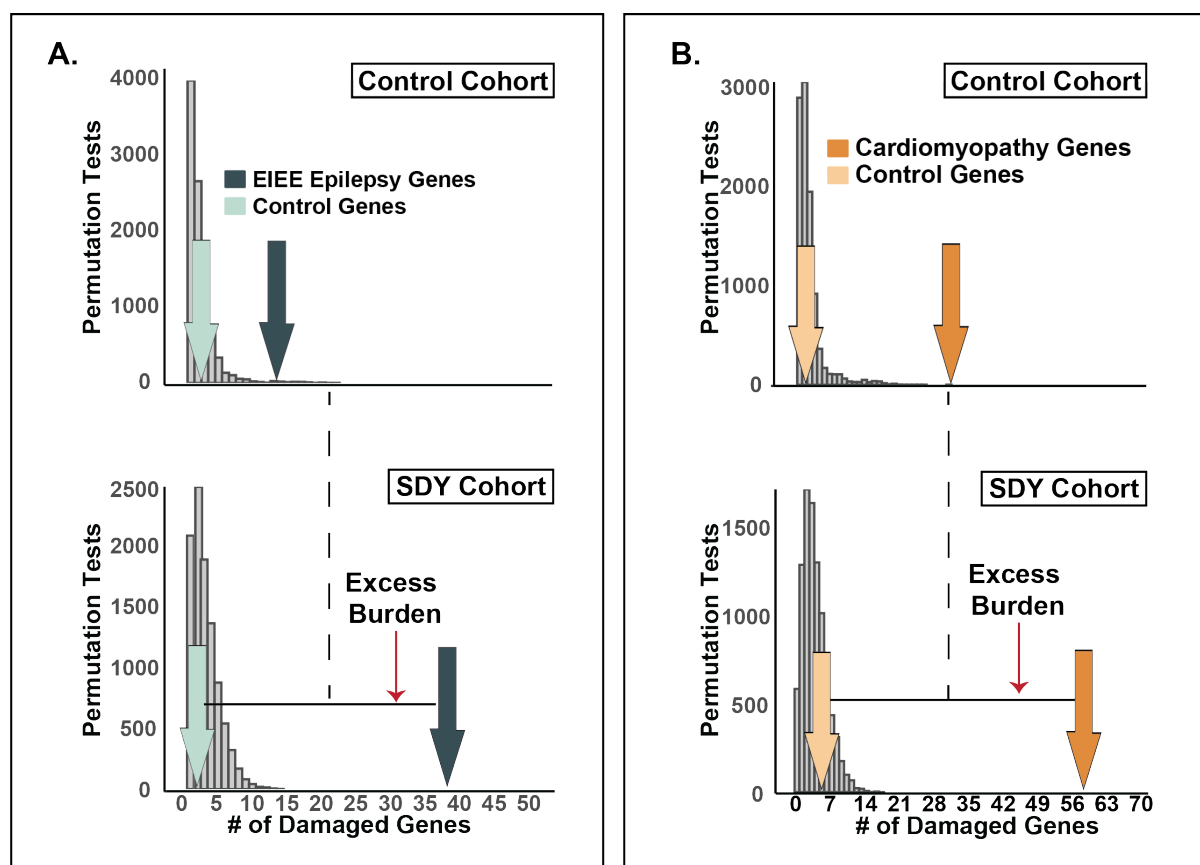
GENE LIST	PHENOTYPE TERMS			
	Seizure			
	SDY	1000 Genomes Project	Genes	Adjusted p-Value*
Epilepsy	62	36	191	0.027
EIEE-OMIM	38	16	82	0.019
	Cardiac			
CMAR1	55	27	118	<0.001
CMAR2	58	31	143	<0.001

*P values adjusted for multiple comparisons using the FDR adjustment.

Supplemental Table 7. Linear regression of age at death against number of rare epilepsy variants*.

Model	Coefficient	P value
Adjusted (Ancestry PCs 1-6)	-0.647	0.0053
Unadjusted	-0.458	0.0387
Adjusted log	-0.159	0.0494
<i>Extreme value sensitivity analysis</i>		
Removal of cases with more than 9 variants (Adjusted PCs)	-0.636	0.0113
Removal of cases with more than 9 variants and less than 1 (Adjusted PCs)	-0.638	0.0137
Removal of cases with 0 rare epilepsy(<0.001) variants (Adjusted PCs)	-0.650	0.0063

*Rare epilepsy variants = nonsynonymous variants identified in the Epilepsy gene panel with an allele frequency <0.001 in gnomAD.



Supplemental Figure 1. The SDY cohort had enriched damaging (A) epilepsy and (B) CMAR2 gene burden compared to a sex- and ancestry-matched control cohort.

Histograms (gray bars) represent distributions of damaged genes (GEM Score > 0.69) sampled randomly from RefSeq genes using a root phenotype from the SDY cohort (n=211) (bottom panels) and 1000 Genomes Project Cohort (control) matched for sex and ancestry (n=211) (top panels). Damaged genes identified in the Epilepsy EIEE (n=82 genes, green) (A) and CMAR2 (n=143 genes, orange) (B) gene lists were significantly different between the SDY and control cohorts (dark arrows, epilepsy, $p=0.019$; cardiac $p<0.001$). Light arrows represent the number of damaged genes identified in a control gene set.