Cardiac Genetic Predisposition in Sudden Infant Death Syndrome

Short Title: Tester – Cardiac Genetic Analyses in SIDS

David J. Tester BS[1]*, Leonie CH Wong MBBChir[2,7*], Pritha Chanana MS[1], Amie Jaye MSc[3], Jared M. Evans MS[1], David FitzPatrick MD[4], Margaret Evans MBChB[5], Peter Fleming PhD[6], Iona Jeffrey MBChB[2,7], Marta Cohen MD[8], Jacob Tfelt-Hansen MD DMSc[9,10], Michael A Simpson PhD[3#], Elijah R Behr MD[2,7#], Michael J Ackerman MD PhD[1#]

[1] Mayo Clinic, Rochester, Minnesota, USA
[2] St George's University of London, London, United Kingdom
[3] King's College London, London, United Kingdom
[4] University of Edinburgh, Edinburgh, United Kingdom
[5] Royal Infirmary of Edinburgh, Edinburgh, United Kingdom
[6] University of Bristol, Bristol, United Kingdom
[7] St George's University Hospitals' NHS Foundation Trust, London, United Kingdom
[8] Sheffield Children's Hospital, Sheffield, United Kingdom
[9]The Heart Centre, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark;
[10]Department of Medicine and Surgery, Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark

[*] Joint First Authors
[#] Joint Last Authors

Funding: This work was supported by Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health [R01HD042569 to MJA] and by the British Heart Foundation ([LW and ERB] BHF Clinical Research Training Fellowship FS/13/78/30520). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. DJT and MJA are also supported by the Mayo Clinic Windland Smith Rice Comprehensive Sudden Cardiac Death Program. LW was also supported by additional funds from Biotronik and Cardiac Risk in the Young. ERB is supported by the Higher Education Funding Council for England. There are no relevant conflicts of interest to this work.
Disclosures: MJA is a consultant for Audentes Therapeutics, Boston Scientific, Gilead Sciences, Invitae, Medtronic, MyoKardia, and St. Jude Medical. MJA and Mayo Clinic have an equity/royalty-based relationship with AliveCor, Blue Ox Health Corporation, and StemoniX. However, none of these entities were involved in this study in any way. DJT, LCHW, PC, AJ, JME, DF, ME, PF, IJ, MC, JT, MAS, and ERB have nothing to disclose.

Corresponding Authors:
Dr Elijah R. Behr
St George's University of London
Cranmer Terrace, London SW17 0RE
Email: ebehr@sgul.ac.uk
Telephone:+442087252994
Fax: +442087253416

and

Dr Michael J. Ackerman
Mayo Clinic
200 First St. SW, Rochester, MN 55905
Telephone: 507-284-8900
Fax: 507-507-284-3757
Email: ackerman.michael@mayo.edu

Acknowledgments: The authors gratefully acknowledge both the medical examiners and coroners for referring the sudden death victims to our program in an effort to find an explanation for the infant's sudden death.

Abstract
Background: Sudden infant death syndrome (SIDS) is a leading cause of post-neonatal mortality. Genetic heart diseases (GHDs) underlie some of SIDS.
Objectives: We aimed to determine the spectrum and prevalence of GHD-associated mutations as a potential monogenic basis for SIDS.
Methods: A cohort of 419 unrelated SIDS cases (257 males; average age = 2.7 ± 1.9 months) underwent whole exome sequencing and a targeted analysis of 90 GHD-susceptibility genes. The yield of "potentially informative", ultra-rare variants (MAF<0.00005) in GHD-associated genes was assessed.
Results: Overall, 53/419 (12.6%) SIDS cases had ≥ 1 "potentially informative", GHD-associated variant. The yield was 14.9% (21/141) for mixed-European ancestry cases and 11.5% (32/278) for European ancestry SIDS cases. Infants older than 4 months were more likely to host a "potentially informative" GHD-associated variant. There was a significant over-representation of ultra-rare non-synonymous variants in European SIDS cases (18/278, 6.5%) versus European controls (30/973, 3.1%, p=0.013) when combining all 4 major cardiac channelopathy genes (*KCNQ1*, *KCNH2*, *SCN5A*, and *RYR2*). According to the American College of Medical Genetics guidelines, only 17/419 (4.1%) SIDS cases hosted a "pathogenic" or "likely pathogenic" variant.
Conclusion: Less than 15% of over 400 SIDS cases had a "potentially informative" variant in a GHD-susceptibility gene, predominantly in the 4-12 month age group. Only 4.1% of cases possessed immediately clinically actionable variants. Consistent with previous studies, ultra-rare, non-synonymous variants within the major cardiac channelopathy-associated genes were over-represented in SIDS cases of European ethnicity. These findings have major implications for the investigation of SIDS cases and families.

Condensed Abstract: Sudden infant death syndrome (SIDS) is a leading cause of post-neonatal mortality. Genetic heart diseases (GHDs) underlie some of SIDS. Here, using a cohort of 419 unrelated SIDS cases, we aimed to determine the spectrum and prevalence of GHD-associated mutations as a possible monogenic basis for SIDS using whole exome sequencing and a targeted analysis of 90 GHD-susceptibility genes. The yield of "potentially informative", ultra-rare variants (MAF<0.00005) in GHD-associated genes was assessed. Less than 15% of cases had a "potentially informative" variant, predominantly in the 4-12 month age group. Only 4.1% of cases possessed immediately clinically actionable variants.

Keywords: Genetic heart diseases; molecular autopsy; sudden infant death syndrome; whole exome sequencing

Abbreviations
ACMG – American College of Medical Genetics
BrS – Brugada Syndrome
CPVT – Catecholaminergic Polymorphic Ventricular Tachycardia
ExAC – Exome Aggregation Consortium
GHD – Genetic Heart Diseases
HCM – Hypertrophic Cardiomyopathy
LQTS – Long QT Syndrome
NSV – Nonsynonymous Variant
SIDS – Sudden Infant Death Syndrome

WES – Whole Exome Sequencing

Introduction

Sudden infant death syndrome (SIDS) is the sudden unexpected death of an infant less than 1 year of age, which remains unexplained despite comprehensive clinical and pathological investigations (1). SIDS represents 70-80% of all sudden unexpected infant deaths with an incidence of 0.4/1000 live births in the UK and 0.5/1000 live births in the USA.(2,3) The peak incidence occurs between 2 – 4 months of age and is more common in males. Such infant deaths are associated commonly with environmental risk factors such as co-sleeping or prone sleeping position.(4) Despite successful targeted risk reduction campaigns, the number of SIDS cases have plateaued, and SIDS remains the leading cause of post-neonatal mortality(4).

A triple-risk model for SIDS suggest the convergence of the vulnerable infant in the setting of exogenous stressors during a critical development period (5). Although many pathophysiologic theories have been proposed, decisive pathogenic substrates/mechanisms triggering an infant's sudden demise remain unclear (6-9). Several studies have implicated both common and rare genetic variants involved in autonomic function, neurotransmission, energy metabolism, response to infection, and cardiac repolarization (10-14). Also, potentially lethal genetic heart diseases (GHDs) including long QT syndrome (LQTS), Brugada syndrome (BrS), catecholaminergic polymorphic ventricular tachycardia (CPVT), and hypertrophic cardiomyopathy (HCM) have been implicated as monogenic causes for a small proportion of SIDS cases (10,13,15-27).

However, less than 100 investigations of genetic variation in population-based SIDS cohorts have been published to date, largely based on hypothesis-driven, candidate gene/pathway-based approaches that recognize established pathobiological risk factors for SIDS, with an average cohort size of just 125 SIDS cases (13). Here, using whole exome sequencing

(WES), we conducted a GHD-associated gene-specific analysis on a cohort of over 400 unrelated SIDS cases.

Methods

Study Population

The SIDS cohort (N=427) consisted of 95 coroners' cases from the United Kingdom (UK; London, Sheffield, Edinburgh and Bristol) and 332 coroner/medical examiner/forensic pathologist-referred cases collected from six ethnically and geographically diverse United States (US) populations. Because of the lack of uniformity in procedures and reporting between medical examiner offices in the US, minor differences in protocols may exist. Nonetheless, both gross and histological examinations of all major organs were performed and all cases satisfied our enrolment criteria that included 1) sudden unexplained death of an infant < 1 year of age, 2) European descent, and 3) a comprehensive negative medico-legal autopsy including a negative toxicology screen and death scene investigation. Infants with asphyxia or specific disease causing death were excluded. Ethnicity was self-reported by the referring coroner/medical examiner. This anonymous necropsy study only had limited medical information such as the sex, ethnicity age at the time of death, and sleep position available. This study complies with the Declaration of Helsinki; locally appointed ethics committees including Mayo Clinic's Institutional Review Board have approved the research protocol. Some of the 332 samples from the US have been included in previous publications that involved hypothesis-driven, specific candidate gene mutational analysis (10,18,19,23-28). Of the 332 cases, 58 had been analyzed previously for variants in *SCN5A* (10), *KCNQ1* (18), *KCNH2* (18), *RYR2* (19), *SNTA1* (23), *KCNJ8* (24), *Cx43* (25), *GPD1L* (26), *CAV3* (27), *SCN1B* (28), *SCN2B* (28), *SCN3B* (28), and *SCN4B* (28). An additional 25 of the 332 were also analyzed for *RYR2* (19) and an additional

145 of the 332 cases were also analyzed for *SNTA1* (23), *KCNJ8* (24), *Cx43* (25), *GPD1L* (26), *CAV3* (27), *SCN1B* (28), *SCN2B* (28), *SCN3B* (28), and *SCN4B* (28). None of the 95 cases from the UK have been published previously.

Control Population

973 control exomes (509 females, 464 males) from the ICR1000 UK exome series and the 1958 Birth Cohort study were included for case-control analysis.(29) As previously reported, exome sequencing was performed using the Illumina TruSeq and Illumina instruments.(29).

Whole Exome Sequencing (WES)

Genomic DNA isolated from each SIDS case underwent WES at the KCL-GSTT Biomedical Research Centre Genomics Platform, London, UK or Mayo Clinic's Medical Genome Facility, Rochester, Minnesota.

Paired-end libraries were prepared following the manufacturer's protocol (Agilent) using the Bravo liquid handler from Agilent. Briefly, 1-3 ug of genomic DNA was fragmented to 150-200 bp using the Covaris E210 sonicator. The ends were repaired and an "A" base was added to the 3' ends. Paired end Index DNA adaptors (Agilent) with a single "T" base overhang at the 3' end were ligated and the resulting constructs were purified using AMPure SPRI beads (Agencourt). The adapter-modified DNA fragments were enriched by 4 cycles of PCR using SureSelect forward and SureSelect ILM Pre-Capture Indexing reverse (Agilent) primers. The concentration and size distribution of the libraries was determined on an Agilent Bioanalyzer DNA 1000 chip.

Whole exon capture was performed using the protocol for Agilent's Sure SelectXT Human All Exon V5+UTR kit. Briefly, 750 ng of the prepped library was incubated with whole exon biotinylated RNA capture baits supplied in the kit for 24 hours at 65 °C. The captured

DNA:RNA hybrids were recovered using Dynabeads MyOne Streptavidin T1 (Dynal). The DNA was eluted from the beads and purified using Ampure XP beads (Agencourt). The purified capture products were then amplified using the SureSelect Post-Capture Indexing forward and Index PCR reverse primers (Agilent) for 12 cycles.

Libraries were pooled at equimolar concentrations and loaded onto paired end flow cells at concentrations of 7-8 pM to generate cluster densities of 600,000-800,000/mm^2 following Illumina's standard protocol using the Illumina cBot and HiSeq Paired end cluster kit version 3. Each lane of a HiSeq flow cell produced 21-39 Gbases of sequence. The level of sample pooling was controlled by the size of the capture region and the desired depth of coverage.

The flow cells were sequenced as 101 X 2 paired end reads on an Illumina HiSeq 2000 using TruSeq SBS sequencing kit version 3 and HiSeq data collection version 2.0.12.0 software. Base-calling was performed using Illumina's RTA version 1.17.21.3.

The FASTQ files underwent quality control checks using FASTQC. The Illumina paired end reads were aligned to the GRCh37 (hg19) human reference genome using Novoalign (http://novocraft.com). Single-sample variant calling with the Genome Analysis Toolkit (GATK, Version 3.2-2)(30) and the resulting gVCFs subsequently underwent multi-sample genotyping and variant quality score recalibration. Genotypes were excluded if the QC was < 15 or there were fewer than 4 reads supporting the call. Further filtering of variant sites was performed to exclude sites with missingness > 0.1 in cases or controls. Variants were annotated with respect to the genes in which they reside with Annovar, allele frequencies were obtained from the Exome Aggregation Consortium (ExAC) database and combined annotation dependent depletion (CADD) scores from the CADD server (http://cadd.gs.washington.edu).

Quality Control Coverage Analysis and Principal Component Analysis (PCA) for Relatedness

and Ethnicity

Coverage across the exome was assessed using the Bedtools package cases were excluded from further analysis if < 75% of the Gencode defined protein coding exome was covered by < 20 reads. A set of 3847 common variants located outside of regions of the genome where there is extensive linkage disequilibrium were used to estimate relatedness within the study cohort and ethnic ancestry alongside the control group.(31) Estimation was undertaken using the first two dimensions of a Multidimensional Scaling (MDS) using Euclidean distance undertaken with the King software package.

Ancestry Confirmation

To avoid potential confounding due to population stratification resulting from genetic admixture, a principal component analysis (PCA) was performed (Online Supplement). The PCA served only for the rare variant analysis between European SIDS cases and European controls. The PCA data was not used for attributing causality to identified variants where ethnic matched controls would not be necessary for variant adjudication. SIDS cases and controls forming a homogeneous cluster on the first two components were included in the case-control rare variant analysis.

Genetic Heart Disease (GHD)-Gene Specific Variant Analysis

Known cardiac channelopathy- (LQTS, CPVT, BrS) and cardiomyopathy- (HCM, DCM, ACM) susceptibility genes (N=90, Online Table 1) were evaluated for the presence of “ultra-rare” nonsynonymous variants (NSVs) with a minor allele frequency (MAF) < 0.00005 (1:20,000 alleles) derived from the Exome Aggregation Consortium (ExAC)(32). A comparison of yield of was undertaken for ultra-rare NSVs in SIDS cases of PCA-determined European ancestry versus European controls across all 90 GHD-susceptibility genes and the 4 “major”

channelopathy genes (*KCNQ1*, *KCNH2*, *SCN5A*, and *RYR2*).

All putative loss of function variants (i.e. a "radical" variant: frame-shift, nonsense, and essential splice-site variants) or missense variants with a previously established abnormal in vitro function characterization that resided within any of the 90 GHD-associated genes and all ultra-rare, missense variants residing in any of the 4 "major" channelopathy genes were considered to be "potentially informative" variants that would be appropriate for investigation of their significance in a family. Such variants were confirmed using standard Sanger sequencing techniques. The American College of Medical Genetics and Genomics (ACMG) standards and guidelines for the interpretation of sequence variants was used to further assist in the classification of our genetic findings among all ultra-rare (MAF < 0.00005) NSV identified across the 90 GHD-associated genes (33).

Statistics

Categorical variables were expressed as absolute numbers and percentage, and compared with Fisher's exact or Chi-square tests. Probability values were based on two-sided tests considered significant at $P<0.05$. Analysis was conducted with SPSS version 18.0 software (SPSS Chicago III).

Results

Demographics

WES was performed in 427 SIDS cases. However, quality control metrics excluded 7 cases due to insufficient exome coverage and one individual from a half-sibling pair (Online Figure 1). The cohort therefore consisted of 419 cases (257 males, 162 females; average age = 2.7 ± 1.9 months) with a skewed bell-shaped distribution of age (Online Figure 2)**.** The epidemiologically higher risk age group of 2–4 months (58.9%) and male gender (61%)

accounted for the majority of the cases. The PCA demonstrated a wide distribution of ancestral origins with 278 cases (173 males, 105 females) considered European ancestry and 141 cases (84 males, 57 females) considered mixed-European (Online Figure 3, **Table 1**). Sleep characteristics were known in 54% of the cohort (**Table 1**). There were no significant differences in demographics and sleep characteristics between the European and mixed-European ancestry cases.

Genetic Heart Disease (GHD) Gene-Specific Analysis

Overall, a total of 285 unique, ultra-rare NSVs (256 missense, 23 putative loss of function [12 frame-shift, 8 splice-errors, 3 nonsense], and 6 in-frame-indels) were identified in 194/419 (46.3%) SIDS cases overall (**Figure 1)**. Further, 45/278 (16.2%) European ancestry cases and 25/141 (17.7%) mixed-European ancestry cases hosted > 1, ultra-rare NSVs.

These ultra-rare NSVs resided in 68 of the 90 GHD-associated genes (21/31 channelopathy-associated; 47/59 cardiomyopathy-associated genes). The gene-specific yields for the overall cohort and the European and mixed European subsets are shown in Online Table 2.

Of the 285 unique, ultra-rare NSVs identified, 57 (20%) were considered "potentially informative". Twenty-five/57 [43.9%] were missense variants in the 4 major channelopathy genes, 23/57 [40.3%] were putative loss of function variants, and 10/57 [17.5%] were variants previously reported in the literature as having an abnormal in vitro functional phenotype (**Table 2**). Overall, 53/419 (12.6%) SIDS cases hosted at least one "potentially informative" variant (**Figure 1**). Four out of 419 cases (0.95%) had two "potentially informative" variants. The yield was 14.9% (21/141) for mixed-European ancestry cases and 11.5% (32/278) for European ancestry SIDS cases (**Figure 1**).

There were no significant differences in the yield of either ultra-rare NSVs among all 90 GHD genes or "potentially informative" variants when comparing gender, sleep position (supine vs prone), or co-sleeping (yes vs no) in either the overall or stratified populations (**Table 3**). However, there was a significantly higher yield of "potentially informative" GHD-associated genetic variants in those infants that died at greater than 4 months of age (15/65, 23.1%) compared to those younger than 4 months of age (37/354, 10.4%, p=0.0075, **Figure 2**).

Following further vetting using the strict ACMG guidelines, only 17 of the 285 ultra-rare NSVs achieved a "pathogenic" or "likely pathogenic" designation and were identified in 17 (4.1%) of the 419 SIDS cases (**Table 2, Figure 1**). There was no difference in yield of "pathogenic" or "likely pathogenic" variants between the European (11/278 [4.0%]) and the mixed-European cohorts (6/141[4.3%]).

Case-Control Analysis

Consistent with previous studies, there was a significant over-representation of ultra-rare NSVs in European SIDS cases (18/278, 6.5%) versus European controls (30/973, 3.1%, p=0.013) when combining all 4 major cardiac channelopathy genes (*KCNQ1*, *KCNH2*, *SCN5A*, and *RYR2*, **Figure 3**). However, there was no significant difference in yield between cases and controls for any specific gene.

Discussion

This manuscript reports results derived from a whole exome molecular autopsy with GHD gene-specific analysis of the largest cohort of unrelated SIDS cases. Previous post-mortem genetic studies have implicated pathogenic mutations in cardiac channelopathy-associated genes as a monogenic cause for up to 15% of SIDS (13,16-19,23). Furthermore, rare HCM-associated sarcomere gene variants were implicated recently in 3.5% of SIDS (21). However, based on their

prevalence in gnomAD, only 1.4% of these SIDS cases hosted variants rare enough to be considered pathogenic.

Because of the rarity of SIDS and any likely causative disorders, we used a strict minor allele frequency cut-off equivalent to a heterozygous frequency of less than 1 in 10,000 individuals in gnomAD. Despite using this very conservative rarity filter, 46% of our SIDS cohort harbored novel or ultra-rare, protein-altering genetic variants within 68 of the 90 GHD-susceptibility genes. Unfortunately, despite their rarity, the vast majority of these ultra-rare variants still remain variants of uncertain significance (VUS) stuck in genetic purgatory.(34,35) In fact, about 25% of these 90 GHD-associated genes have a negative Z-score suggesting they tolerate variation (32,36).

Due to ambiguity surrounding the pathogenic nature of many GHD genes that play a "minor" role in their respective diseases, we examined the yield of ultra-rare NSVs with the highest likelihood of being true pathogenic mutations. Overall, about 13% of our SIDS cases hosted "potentially informative" variants regarded as having the greatest probability of being causative for the infant's sudden death and having potential clinical utility for assessing a family for genetic risk. Unfortunately, the anonymous nature of the cohort prevents us from verifying their presence among family members for the purpose of potential genotype-phenotype co-segregation analysis or to determine the frequency of de novo status of the variants of interest. Importantly, not all of these variants have been characterized functionally and great caution must still be exercised, even when interpreting ultra-rare variants residing within the major channelopathy genes.

Recently, Hertz reported a 34% yield of "variants with likely functional effects" following a genetic analysis of GHD-associated genes in only 47 sudden unexpected deaths in

infancy cases.(22) However, given the rarity of GHDs in the general population, we believe that their definition of rarity (MAF < 1%) was unacceptably and erroneously high; thus causing an overestimated burden of potentially pathogenic variants in their SIDS cohort. In fact, of their 16 "pathogenic" variants, only 1 novel *RYR2* variant would have been deemed 'potentially informative' by our robust criteria.

In 2017, Neubauer reported a yield of "potentially causative" variants in 20% of their 155 European SIDS cases following WES and genetic interrogation of their 192 gene focused panel that comprise both cardiovascular- and metabolic disorder-associated genes.(37) The majority of their seemingly genotype positive infants had a variant with "likely functional effects" in genes associated with a cardiac channelopathy (9%) or cardiomyopathy (7%). However, the vast majority of these variants represent missense variants within "minor" genes.(37) In fact, only 2.6% of their cases hosted what we would consider a "potentially informative" variant by our strict definition.

While our study supports the utility of WES to identify potential sudden death-causing variants within established or potential sudden death-susceptibility genes, the challenge of the WES-based molecular autopsy does not lie in the identification of variants, but rather, in the adjudication of their potential pathogenicity. Accurate variant classification is crucial to enable proper counseling of surviving family members. Erroneously or prematurely adjudicating ambiguous variants as pathogenic has the potential to harm patients and their families. Tragically, this became a reality for one family described by Ackerman and colleagues recently, as they dealt with the disastrous consequences of unnecessary treatment based on an erroneously interpreted variant in *KCNQ1* (35). Thus, over attribution of SIDS deaths to GHDs has significant implications for the immediate family and we urge extreme caution in variant

interpretation. When such cautionary advice is heeded, less than 5% of over 400 SIDS cases had either a “pathogenic” or “likely pathogenic” variant in one of 90 GHD-susceptibility genes which is substantially lower than previous extrapolations of the prevalence of either channelopathic- or cardiomyopathic-SID. This parallels our experience of the ‘molecular autopsy’ in unexplained sudden death where stringent variant evaluation results in a significant reduction of numbers of ‘likely pathogenic’ and ‘pathogenic’ variants of clinical utility (38).

Using a similarly stringent variant analysis, we observed previously a 13% (40/302) yield of ultra-rare “pathogenic” or “likely pathogenic” variants within sudden death-susceptibility genes among 302 autopsy-negative sudden arrhythmic death syndrome (SADS) cases that died at an age greater than 1 year (median age 24 years), compared with a significantly ($p=0.00002$) lower yield of 4.3% (18/419) in our SIDS cohort (38). This data suggests that the majority of SIDS stems from pathobiological bases that are largely different genetically and mechanistically from sudden death occurring after the age of 1 year.

Several risk factors for SIDS have been established. One might hypothesize that vulnerable infants dying of SIDS without the presence of additional risk factors are more likely to host a highly penetrant monogenic cause for their death compared to infants exposed to additional environmental risk factors. Yet, no significant differences in the yield of “potentially informative” GHD gene variants associated with sex, sleep position, or bed sharing were observed. However, a significant age-effect on the yield, where 23% of those infants older than 4 months of age hosted a “potentially informative” GHD-associated variant compared to only 10% of the infants younger than 4 months of age was observed. This data supports potential stratification of those SIDS cases that may benefit most from post-mortem genetic testing of the major channelopathy/cardiomyopathy-associated genes.

The significant over-representation of ultra-rare NSVs within the 4 major channelopathy genes associated with either inheritable LQTS (*KCNQ1*, *KCNH2*, *SCN5A*) or CPVT (*RYR2*) observed in our European Caucasian SIDS compared to ethnic matched controls (6.5% vs 3.1%, p=0.013) supports that cardiac channelopathies may represent the underlying pathogenic basis for some of SIDS and that post-mortem genetic testing of the 4 major channelopathy associated genes may be warranted in cases of SIDS.

In our study, we used a strict MAF cut-off 0.005% (i.e. 1 in 20,000 alleles or 1 in 10,000 individuals). While using a stringent threshold could reduce the possibility of identifying variants that would be deemed too common in the population to cause a rare disease such as LQTS, it could also prevent the identification of potentially important functionally significant variants that could play a role in SIDS pathogenesis. For example, the p.R176W-KCNH2 variant was identified in a single European SIDS case in our cohort. This variant has been associated with LQTS previously, it has been demonstrated to have a functional effect by in vitro assays, and has been considered a founder LQT2 mutation in the Finnish population(39,40). While this variant meets the current ACMG guideline classification of 'pathogenic' variant, its heterozygous frequency in gnomAD (44/53,551 overall, 31/22,031 European, and 10/4,022 Finnish European individuals) exceeds our stringent cut-off and was therefore not included in our analysis.

Conclusions

A whole exome molecular autopsy followed by a cardiac gene specific focus reveals that less than 15% of over 400 SIDS cases had a "potentially informative" variant in one of 90 GHD-susceptibility genes. Furthermore, less than 5% of these infant deaths possessed variants that are immediately useful in a family for further cascade testing. This represents a substantial reduction of the perceived importance of monogenic cardiac genetic disease in SIDS compared to previous

studies. Interpretation of GHD-associated rare variants must therefore be stringent and careful given the implications of misattribution in families. This has important clinical implications for the community managing SIDS cases and their relatives.

Perspectives

Competency in Medical Knowledge: Post-mortem genetic testing using whole exome sequencing (aka, the whole exome molecular autopsy or WEMA) may identify ultra-rare non-synonymous variants within genetic heart disease-susceptibility genes that may underlie the pathogenic basis for a significant number of sudden infant death syndrome (SIDS) cases, predominantly in the 4-12 month age group. However, the yield of ultra-rare non-synonymous variants in all but the 'major' channelopathy genes is very similar in a healthy population, suggesting that single gene cardiac disorders are not the major cause of SIDS. In fact, less than 15% of SIDS cases possessed a 'potentially informative' variant regarded as having the greatest probability of being causative for the infant's sudden death and having potential clinical utility for assessing the potential risk in family members left behind. Furthermore, clinically relevant variants that are immediately useful in a family for cascade/predictive testing were identified in only 4% of infant deaths.

Translational Outlook 1: While some of the ultra-rare 'potentially informative' variants may be disease-causing and contributing to SIDS pathogenesis, future research involving functional studies are necessary to determine which variants are truly pathogenic and whether cardiac genetic disease is a more significant contributor than our data currently suggest.

Translational Outlook 2: SIDS is likely multigenic and complex, with no single predominant genetic pathway for risk. Thus, future research studies are necessary to elucidate intrinsic genetic vulnerability for SIDS across multiple biological pathways in an unbiased manner.

References

1. Krous HF, Beckwith JB, Byard RW, et al. Sudden infant death syndrome and unclassified sudden infant deaths: a definitional and diagnostic approach. Pediatrics 2004;114:234-8.
2. Matthews T, MacDorman MF. Infant mortality statistics from the 2010 period linked birth/infant death data set. Natl Vital Stat Rep 2013;18:1-26.
3. Unexplained deaths in infancy: England and Wales: 2009. Office for National Statistics 2011. http:www.ons.gov.uk/ons/dcp171778_227450.pdf.
4. Moon RY, Horne RSC, Hauck FR. Sudden infant death syndrome. The Lancet 2007;370:1578-87.
5. Filiano JJ, Kinney HC. A perspective on neuropathologic findings in victims of the sudden infant death syndrome: the triple-risk model. Biol Neonate 1994;65:194-7.
6. Valdes-Dapena MA. Editorial: Sudden, unexpected and unexplained death in infancy--a status report--1973. N Engl J Med 1973;289:1195-7.
7. Schwartz PJ. Cardiac sympathetic innervation and the sudden infant death syndrome: a possible pathogenic link. Am J Med 1976;60:167-172.
8. Guntheroth WG. Theories of cardiovascular causes in sudden infant death syndrome. J Am Coll Cardiol 1989;14:443-7.
9. Yun AJ, Lee PY. Sudden death among infants and adults: companion disorders of maladaptive sympathetic bias. Med Hypotheses 2004;62:857-60.
10. Ackerman MJ, Siu BL, Sturner WQ, et al. Postmortem molecular analysis of SCN5A defects in sudden infant death syndrome. JAMA 2001;286:2264-9.
11. Kinney HC, Richerson GB, Dymecki SM, Darnall RA, Nattie EE. The brainstem and serotonin in the sudden infant death syndrome. Ann Rev Pathol 2009;4:517-550.

12. Weese-Mayer DE, Ackerman MJ, Marazita ML, Berry-Kravis EM. Sudden Infant Death Syndrome: Review of implicated genetic factors. Am J Med Genet 2007;143:771-788.

13. Van Norstrand DW, Ackerman MJ. Genomic risk factors in sudden infant death syndrome. Genome Med 2010;2:86.

14. Salomonis N. Systems-level perspective of sudden infant death syndrome. Pediatr Res 2014;76:220-229.

15. Schwartz PJ, Stramba-Badiale M, Segantini A, et al. Prolongation of the QT interval and the sudden infant death syndrome. N Engl J Med 1998;338:1709-14.

16. Glengarry JM, Crawford J, Morrow PL, Stables SR, Love DR, Skinner JR. Long QT molecular autopsy in sudden infant death syndrome. Arch Dis Child 2014;99:635-40.

17. Arnestad M, Crotti L, Rognum TO, et al. Prevalence of long-QT syndrome gene variants in sudden infant death syndrome. Circulation 2007;115:361-7.

18. Tester DJ, Ackerman MJ. Sudden infant death syndrome: How significant are the cardiac channelopathies? Cardiovasc Res 2005;67:388-96.

19. Tester DJ, Dura M, Carturan E, et al. A mechanism for sudden infant death syndrome (SIDS): Stress-induced leak via ryanodine receptors. Heart Rhythm 2007;4:733-9.

20. Van Norstrand DW, Ackerman MJ. Sudden infant death syndrome: Do ion channels play a role? Heart Rhythm 2009;6:272-8.

21. Brion M, Allegue C, Santori M, et al. Sarcomeric gene mutations in sudden infant death syndrome (SIDS). Forensic Sci Int 2012;219:278-81.

22. Hertz CL, Christiansen SL, Larsen MK, et al. Genetic investigations of sudden unexpected deaths in infancy using next-generation sequencing of 100 genes associated with cardiac diseases. Eur J Hum Genet 2016;24:817-22.

23. Cheng J, Van Norstrand DW, Medeiros-Domingo A, et al. Alpha1-syntrophin mutations identified in sudden infant death syndrome cause an increase in late cardiac sodium current. Circ Arrhythm Electrophysiol 2009;2:667-76.

24. Tester DJ, Tan B-H, Medeiros-Domingo A, Song C, Makielski JC, Ackerman MJ. Loss-of-Function Mutations in the KCNJ8-Encoded Kir6.1 KATP Channel and Sudden Infant Death Syndrome. Circ Cardiovasc Genet 2011;4:510-15.

25. Van Norstrand DW, Asimaki A, Rubinos C, et al. Connexin43 mutation causes heterogeneous gap junction loss and sudden infant death. Circulation 2012;125:474-81.

26. Van Norstrand DW, Valdivia CR, Tester DJ, et al. Molecular and Functional Characterization of Novel Glycerol-3-Phosphate Dehydrogenase 1 Like Gene (GPD1-L) Mutations in Sudden Infant Death Syndrome. Circulation 2007;116:2253-9.

27. Cronk LB, Ye B, Kaku T, et al. Novel mechanism for sudden infant death syndrome: Persistent late sodium current secondary to mutations in caveolin-3. Heart Rhythm 2007;4:161-6.

28. Tan B-H, Pundi KN, Van Norstrand DW, et al. Sudden infant death syndrome-associated mutations in the sodium channel beta subunits. Heart Rhythm 2010;7:771-8.

29. Ruark E, Münz M, Renwick A, et al. The ICR1000 UK exome series: a resource of gene variation in an outbred population [version 1; referees: 3 approved], 2015.

30. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 2010;20:1297-303.

31. Price AL, Weale ME, Patterson N, et al. Long-Range LD can confound genome scans in admixed populations. Am J Hum Genet 2008;83:132-5.

32. Exome Aggregation Consortium (ExAC), Cambridge, MA (URL: http://exac.broadinstitute.org) [date (December, 2017) accessed].

33. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015:17:405-24.

34. Ackerman MJ. Genetic Purgatory and the Cardiac Channelopathies: Exposing the Variants of Uncertain/Unknown Significance (VUS) Issue. Heart Rhythm 2015;12:2325-31.

35. Ackerman JP, Bartos DC, Kapplinger JD, Tester DJ, Delisle BP, Ackerman MJ. The promise and peril of precision medicine: Phenotyping still matters most. Mayo Clinic Proceedings 2016;91:1606-16.

36. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-91.

37. Neubauer J, Lecca MR, Russo G, et al. Post-mortem whole-exome analysis in a large sudden infant death syndrome cohort with a focus on cardiovascular and metabolic genetic diseases. Eur J Hum Genet 2017;25:404-9.

38. Lahrouchi N, Raju H, Lodder EM, et al. Utility of Post-Mortem Genetic Testing in Cases of Sudden Arrhythmic Death Syndrome. J Am Coll Cardiol 2017;69:2134-45.

39. Marjamaa A, Salomaa V, Newton-Cheh C, et al. High prevalence of four long QT syndrome founder mutations in the Finnish population. Ann Med 2009;41:234-40.

40. Lahti AL, Kujala VJ, Chapman H, et al. Model for long QT syndrome type 2 using human iPS cells demonstrates arrhythmogenic characteristics in cell culture. Dis Model

Mech 2012;5:220-30.

41. Roux-Buisson N, Rendu J, Denjoy I, et al. Functional analysis reveals splicing mutations of the CASQ2 gene in patients with CPVT: implication for genetic counselling and clinical management. Hum Mut 2011;32:995-9.

42. Fressart V, Duthoit G, Donal E, et al. Desmosomal gene analysis in arrhythmogenic right ventricular dysplasia/cardiomyopathy: spectrum of mutations and clinical impact in practice. EP Europace 2010;12:861-8.

43. 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;6:1297-303.

44. Napolitano C, Priori SG, Schwartz PJ, et al. Genetic testing in the long QT syndrome: development and validation of an efficient approach to genotyping in clinical practice.[see comment]. JAMA 2005;294:2975-80.

45. Crotti L, Lundquist AL, Insolia R, et al. KCNH2-K897T is a genetic modifier of latent congenital long-QT syndrome. Circulation 2005;112:1251-8.

46. Haghighi K, Kolokathis F, Gramolini AO, et al. A mutation in the human phospholamban gene, deleting arginine 14, results in lethal, hereditary cardiomyopathy. Proc Natl Acad Sciences of the United States of America 2006;103:1388-1393.

47. Medeiros-Domingo A, Bhuiyan ZA, Tester DJ, et al. The RYR2-encoded ryanodine receptor/calcium release channel in patients diagnosed previously with either catecholaminergic polymorphic ventricular tachycardia or genotype negative, exercise-induced long QT syndrome: a comprehensive open reading frame mutational analysis. J Am Coll Cardiol 2009;54:2065-74.

Figure Legends

Central Illustration. Whole Exome Sequencing and a Targeted Analysis of 90 GHD-Susceptibility Genes – Illustrated is the definition of sudden infant death syndrome (SIDS), the triple risk hypothesis for SIDS highlighting genetic heart disease as a potential explanation for infant vulnerability, and our whole exome sequencing strategy to detect ACMG guideline predicated "pathogenic" or "likely pathogenic" variants in SIDS cases.

Figure 1. Yield of Ultra-Rare and "Potentially Informative" Genetic Heart Disease (GHD)-Associated Gene Variants – Shown is a bar graph depicting the percent yield of ultra-rare (minor allele frequency < 0.00005), non-synonymous variants, the "potentially informative" variants, and ACMG guideline-specified "pathogenic" or "likely pathogenic" variants that were identified among the 90 genetic heart disease (GHD)-associated genes for the overall, European Caucasian, and Mixed-European Ancestry cohorts.

Figure 2. The Effect of Age at the Time of Death on the Yield of "Potentially Informative" Genetic Heart Disease (GHD)-Associated Gene Variants – Depicted is a bar graph comparing the percent yield of "potentially informative" variants between those infants that died at an age younger than or older than 4 months.

Figure 3. Yield of Ultra-Rare Major Cardiac Channelopathy Gene Variants in European SIDS Cases Versus European Controls - Shown is a bar graph depicting the percent yield of ultra-rare (minor allele frequency < 0.00005) variants in major cardiac channelopathy genes (*KCNQ1*, *KCNH2*, *SCN5A*, and *RYR2*).

Table 1. Summary of the Sudden Infant Death Syndrome Cohort Demographics

Demographics		Overall (n=419)	European Ancestry (n=278)	Mixed European Ancestry (n=141)
Sex	**Male**	257 (61.3%)	173 (62.2%)	84 (59.6%)
	Female	162 (38.7%)	105 (37.8%)	57 (40.4%)
Age	**Average (months)**	2.7 ± 1.9	2.7 ± 1.98	2.8 ± 2.3
	Range (months)	0.1 - 12	0.1 -12	0.25 – 9.7
Age Group	**< 2 months**	117 (27.9%)	81 (29.1%)	36 (25.5%)
	2 – 4 months	237 (56.6%)	154 (55.4%)	83 (58.9%)
	> 4 months	65 (15.5%)	43 (15.5%)	22 (15.6%)
Gender	**Male**	257 (61.3%)	173 (62.2%)	84 (59.6%)
	Female	162 (38.7%)	105 (37.8%)	57 (40.4%)
Sleep Position	**Supine**	113 (27%)	85 (30.6%)	28 (19.8%)
	Prone	73 (17.4%)	52 (18.7%)	21 (14.9%)
	Side	35 (8.4%)	29 (10.4%)	6 (4.2%)
	Seated	5 (1.2%)	2 (0.72%)	3 (2.1%)
	Unknown	193 (46%)	110 (39.6%)	83 (58.9%)
Co-sleeping	**Yes**	95 (22.7%)	66 (23.7%)	29 (20.6%)
	No	141 (33.6%)	106 (38.1%)	35 (24.8%)
	Unknown	183 (43.7%)	106 (38.1%)	77 (54.6%)

Table 2. Summary of "Potentially Informative", Ultra-Rare NSVs (MAF<0.00005) Identified in SIDS.

Case	Ethnicity	Sex	Age (Months)	Sleep Position	Co-Sleeping	Gene	Nucleotide	Amino Acid Change	Novel / Reference	ACMG Variant Interpretation
1	European	F	4.2	-	-	*CALR3*	c.179_180delAA	p.K60Rfs*4	Novel	Likely Pathogenic
2	European	F	0.1	-	-	*CASQ2*	c.546delT	p.F182Lfs*28	Reference (41)	Pathogenic
3	European	M	4	SIDE	NO	*CRYAB*	c.325-2A>G	Splice error		VUS
4	European	F	5.8	SUPINE	-	*CSRP3*	c.415-1G>T	Splice error	Novel	Likely Pathogenic
5	Mixed-European	F	4.1	PRONE	NO	*CTNNA3*	c.843+1G>T	Splice error	Novel	VUS
6	European	M	3	-	-	*DSG2*	c.523+2T>C	Splice error	Reference (42)	Likely Pathogenic
7	Mixed-European	M	2	SUPINE	YES	*DSP*	c.6540delG	p.V2181Sfs*6	Novel	Likely Pathogenic
8	Mixed-European	M	1	-	-	*GPD1L*	c.817C>T	p.R273C	Reference (26)	VUS
9	European	M	3	SUPINE	YES	*KCNE2*	c.369_370delCT	p.P123fs*14	Reference (43)	Likely Pathogenic
10	Mixed-European	F	1	SUPINE	NO	*KCNH2*	c.865G>A	p.E289K	Novel	VUS
11	Mixed-European	F	0.75	SUPINE	YES	*KCNH2*	c.2903C>T	p.P968L	Reference (44)	VUS
12	European	F	5	PRONE	NO	*KCNH2*	c.3013C>T	p.R1005W	Novel	VUS
13	European	M	1.1	SUPINE	NO	*KCNH2*	c.3347C>T	p.A1116V	Reference (45)	VUS
14	European	M	5	-	-	*KCNJ8*	c.995_997delAAG	p.E332del	Reference (24)	Likely Pathogenic
15	European	F	3	-	-	*KCNQ1*	c.230C>T	p.S77F	Novel	VUS
16	European	M	6	SIDE	NO	*KCNQ1*	c.1553G>A	p.R518Q	Reference (43)	VUS
17	European	M	6.1	SIDE	NO	*LAMA4*	c.2174-	Splice error	Novel	Likely Pathogenic

							8_2182delGTGTAAAGGGGATGCCC			
18	Mixed-European	M	1.73	-	-	*MIB1*	c.2245delA	p.K749Rfs*18	Novel	Likely Pathogenic
19	European	M	3.5	PRONE	NO	*MTO1*	c.1966delA	p.K656Nfs*19	Novel	Likely Pathogenic
20	Mixed-European	F	0.5	SUPINE	YES	*NEBL*	c.1560+1G>A	Splice error	Novel	Likely Pathogenic
21	Mixed-European	M	1	-	NO	*NEBL*	c.2148+1G>A	Splice error	Novel	VUS
22	Mixed-European	F	3.2	PRONE	NO	*PDLIM3*	c.650-2A>G	Splice error	Novel	Likely Pathogenic
23	European	M	5.5	SUPINE	NO	*PLN*	c.36_38delAAG	p.R14del	Reference (46)	Likely Pathogenic
24	European	M	4.9	PRONE	-	*RYR2*	c.950T>C	p.M317T	Novel	VUS
25	European	F	3.1	PRONE	NO	*RYR2*	c.2626C>A	p.P876T	Novel	VUS
26	Mixed-European	M	6	SIDE	NO	*RYR2*	c.3245G>A	p.G1082E	Novel	VUS
27	Mixed-European	F	5	PRONE	YES	*RYR2*	c.6203G>A	p.R2068Q	Novel	VUS
28	European	F	7.6	PRONE	YES	*RYR2*	c.6252G>A	p.M2084I	Novel	VUS
29	European	M	11	PRONE	NO	*RYR2*	c.6490G>A	p.A2164T	Novel	VUS
30	European	M	3.1	SUPINE	YES	*RYR2*	c.7528A>G	p.T2510A	Reference (47)	VUS
31	European	M	2	SUPINE	YES	*RYR2*	c.7946T>C	p.F2649S	Novel	VUS
32	Mixed-European	M	1	-	-	*RYR2*	c.9626C>T	p.P3209L	Novel	VUS
33	European	F	2	SUPINE	NO	*RYR2*	c.12692C>T	p.P4231L	Novel	VUS
34	European	M	2	SUPINE	YES	*RYR2*	c.12713T>G	p.I4238S	Novel	VUS
35	European	M	3	-	-	*RYR2*	c.13624G>A	p.A4542T	Novel	VUS
36	Mixed-European	F	1.5	-	-	*SCN3B*	c.106G>A	p.V36M	Reference (28)	VUS
37	Mixed-European	M	6	SUPINE	NO	*SCN3B*	c.161T>G	p.V54G	Reference (28)	Likely Pathogenic
38	European	M	3.1	PRONE	YES	*SCN5A*	c.997G>C	p.G333A	Novel	VUS
39	Mixed-	M	4	-	-	*SCN5A*	c.2251G>A	p.V751I	Novel	VUS

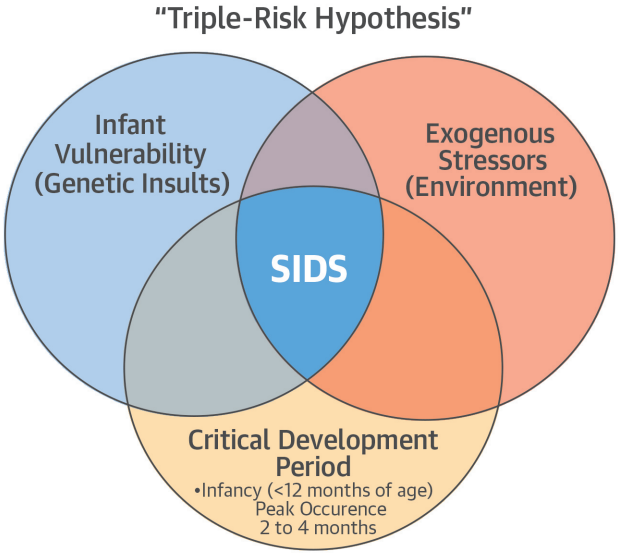
	European									
40	European	F	3	-	-	*SCN5A*	c.4870G>A	p.V1624I	Novel	VUS
41	Mixed-European	M	3	-	-	*SCN5A*	c.5359A>G	p.S1787G	Novel	VUS
42	European	M	4.1	-	-	*SDHA*	c.98_107delTTCACTTCAC	p.F33Lfs*22	Novel	Likely Pathogenic
43	Mixed-European	F	0.75	-	-	*SNTA1*	c.861C>G	p.S287R	Reference (23)	Pathogenic
44	Mixed-European	M	2	-	-	*SNTA1*	c.1115C>T	p.T372M	Reference (23)	VUS
45	European	F	1.2	-	-	*SNTA1*	c.1378G>A	p.G460S	Reference (23)	VUS
46	Mixed-European	M	3	-	-	*TMPO*	c.1244_1245insT	p.K416fs*0	Novel	VUS
47	European	M	0.6	PRONE	NO	*TTN*	c.13774G>T	p.E4592*	Novel	VUS
48	European	M	0.5	SUPINE	NO	*TTN*	c.26011C>T	p.R8671*	Novel	VUS
49	Mixed-European	M	2	-	-	*TTN*	c.47653delA	p.R15885Efs*9	Novel	VUS
50	Mixed European	M	7	-	-	*TTN*	c.37432_37433insGTGGTTACTACAGCCTC	p.N12478Sfs*19	Novel	VUS
						TTN	c.37437_37438insG	p.S12480Vfs*24	Novel	VUS
51	European	M	2.1	-	-	*GJA1*	c.124G>A	p.E42K	Reference (25)	Likely Pathogenic
						KCNH2	c.2654G>A	p.R885H	Novel	VUS
52	European	F	2	SUPINE	NO	*CSRP3*	c.286_287delCC	p.P96Kfs*35	Novel	Likely Pathogenic
						SCN5A	c.3079C>T	p.R1027W	Novel	VUS
53	European	M	1.5	-	-	*NEBL*	c.1639C>T	p.R547*	Novel	VUS
						SCN5A	c.2989G>T	p.A997S	Reference (10)	VUS

EUR = European Ancestry, MIX-EUR = Mixed-European Ancestry, M=Male, F=Female, VUS = Variant of Uncertain Significance

Table 3. The Effect of Various Demographics on the Yield of Genetic Heart Disease-Associated Gene Variants

Demographics		All Ultra-Rare, Non-Synonymous Variants						"Potentially Informative" Variants					
		Overall N= 419	**P value**	**European N=278**	**P value**	**Mixed European N=141**	**P value**	**Overall N= 419**	**P value**	**European N=278**	**P value**	**Mixed European N=141**	**P value**
Age	**2 - 4m**	108/237 (45.6%)	0.76	71/154 (46.1%)	0.63	37/83 (44.6%)	0.17	**22/237 (9.3%)**	**0.03**	15/154 (9.7%)	0.35	**7/83 (8.4%)**	**0.015**
	Other	86/182 (47.3%)		53/124 (42.7%)		33/58 (56.9%)		**31/182 (17.0%)**		17/124 (13.7%)		**14/58 (24.1%)**	
Gender	**Male**	118/257 (45.9%)	0.84	78/173 (45.1%)	0.90	40/84 (47.6%)	0.61	34/257 (13.2%)	0.76	21/173 (12.1%)	0.85	13/84 (15.5%)	1
	Female	76/162 (46.9%)		46/105 (43.8%)		30/57 (52.6%)		19/162 (11.7%)		11/105 (10.5%)		8/57 (14.0%)	
Sleep Position	**Prone**	31/73 (42.5%)	0.25	24/52 (46.2%)	0.65	7/21 (33.3%)	0.30	10/73 (13.7%)	0.98	8/52 (45.4%)	0.67	2/21 (9.5%)	0.91
	Supine	48/113 (42.5%)		36/85 (42.3%)		12/28 (42.9%)		15/113 (13.3%)		10/85 (11.8%)		5/28 (17.9%)	
	Side	14/35 (40.0%)		10/29 (34.5%)		4/6 (66.7%)		4/35 (11.4%)		3/29 (10.3%)		1/6 (16.7%	
	Seated	1/5 (20%)		0/2 (0%)		1/3 (33.3%)		1/5 (20%)		0/2 (0%)		0/3 (0%)	
	Unknown	100/193 (51.8%)		54/110 (49.1%)		46/83 (55.4%)		24/193 (12.4%)		11/110 (10%)		13/83 (15.7%)	
Co-sleeping	**Yes**	38/95 (40%)	0.16	25/66 (37.9%)	0.21	13/29 (44.8%)	0.8	9/95 (9.5%)	0.57	6/66 (9.1%)	0.78	3/29 (10.3%)	0.73
	No	62/141 (44%)		45/106 (42.5%)		17/35 (48.6%)		19/141 (13.5%)		13/106 (12.3%)		6/35 (17.1%)	
	Unknown	94/183 (51.4%)		54/106 (50.9%)		40/77 (51.9%)		25/183 (13.7%)		13/106 (12.3%)		12/77 (15.6%)	

Bold indicates a significant p value (<0.05)



Next-generation Whole-exome Sequencing of 419 SIDS cases

↓

Mapping, Calling, and Annotation of Variants

↓

Principle Component Analysis with European Caucasian Controls

↓ 278 European Ancestry

↓ 141 Mixed European Ancestry

"Ultra-rare" (MAF <0.005%) Non-synonymous Variant Analysis

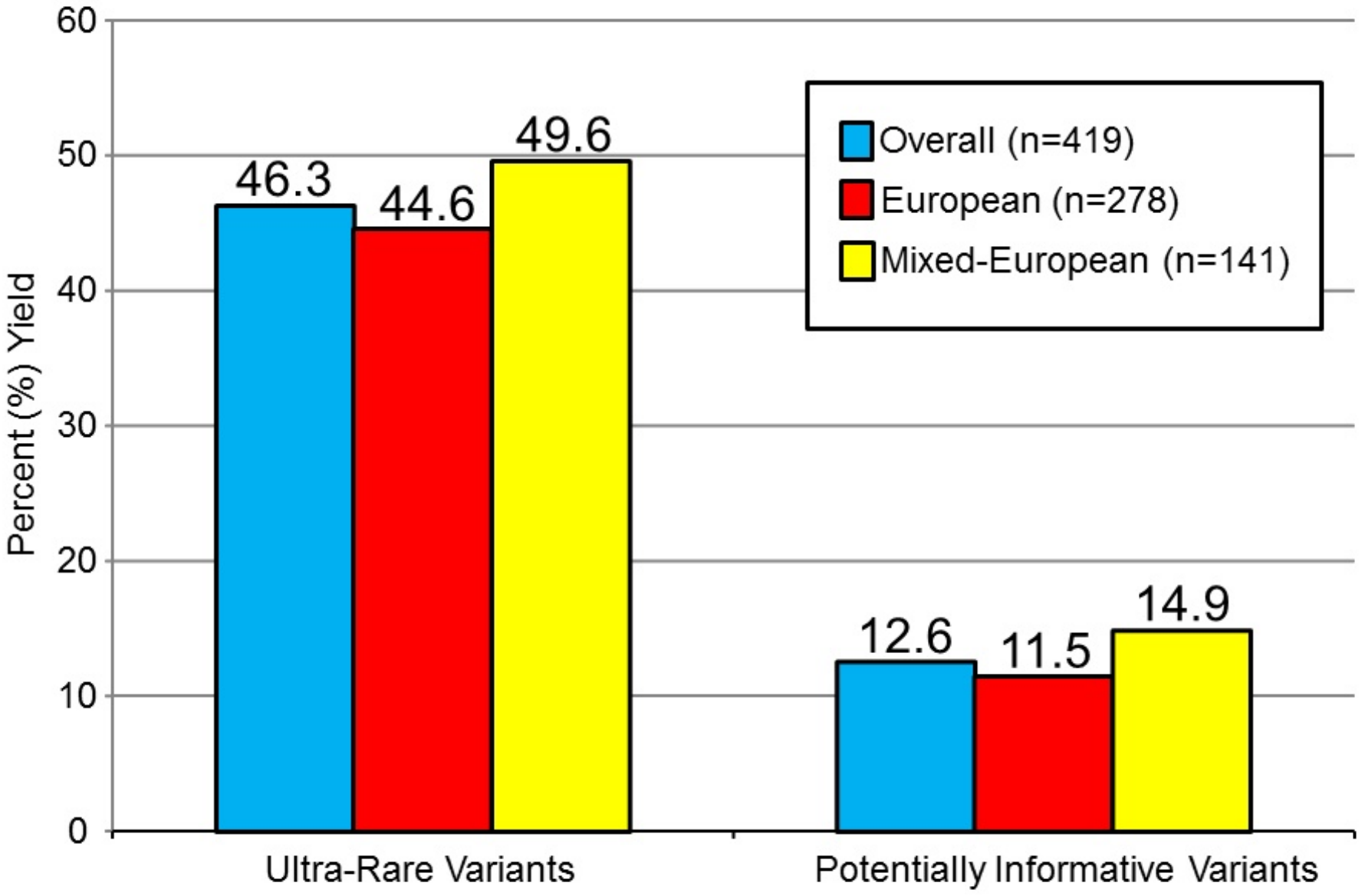
↓

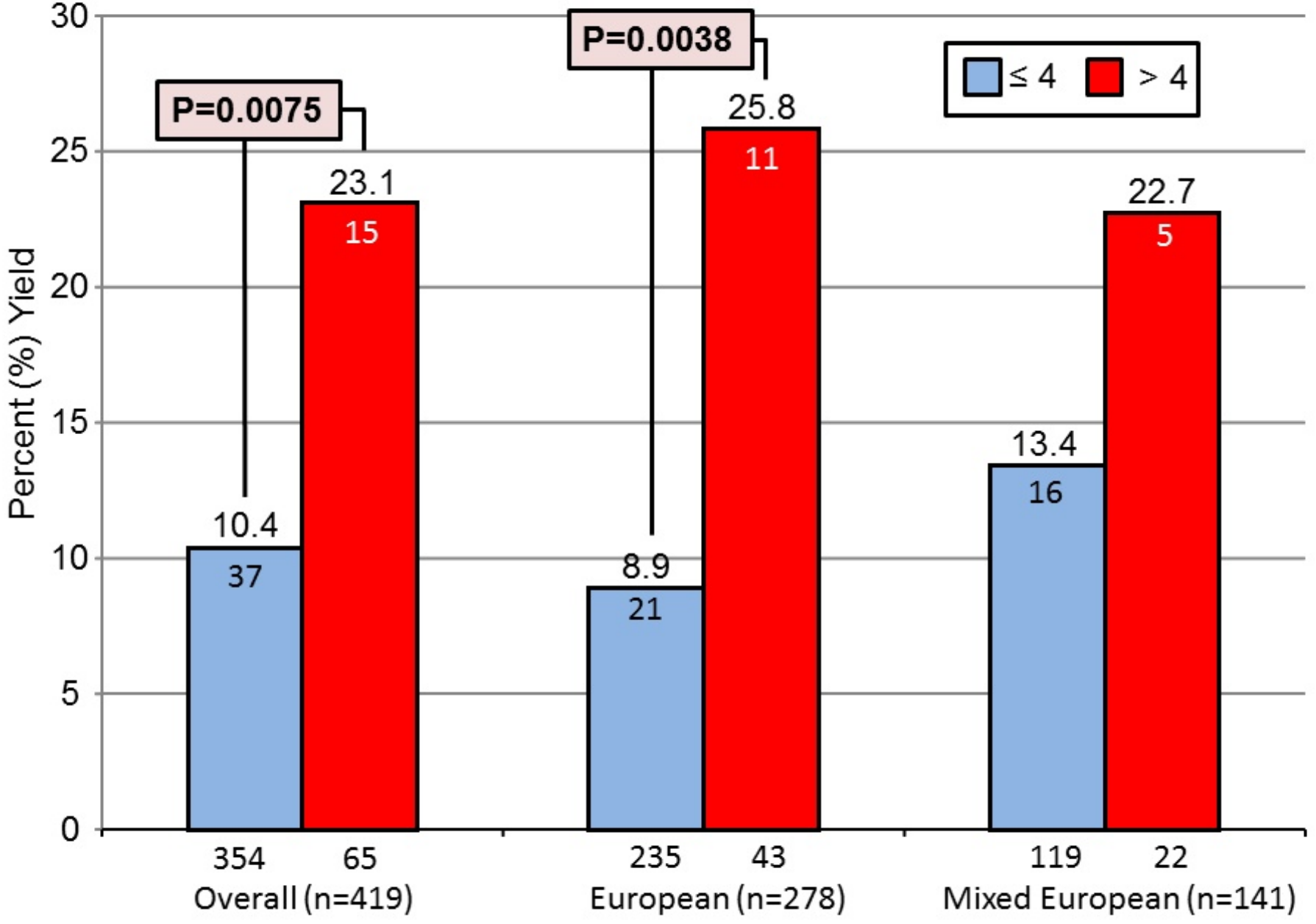
Genetic Heart Disease (GHD) Gene Specific Analysis (n = 90 Genes)

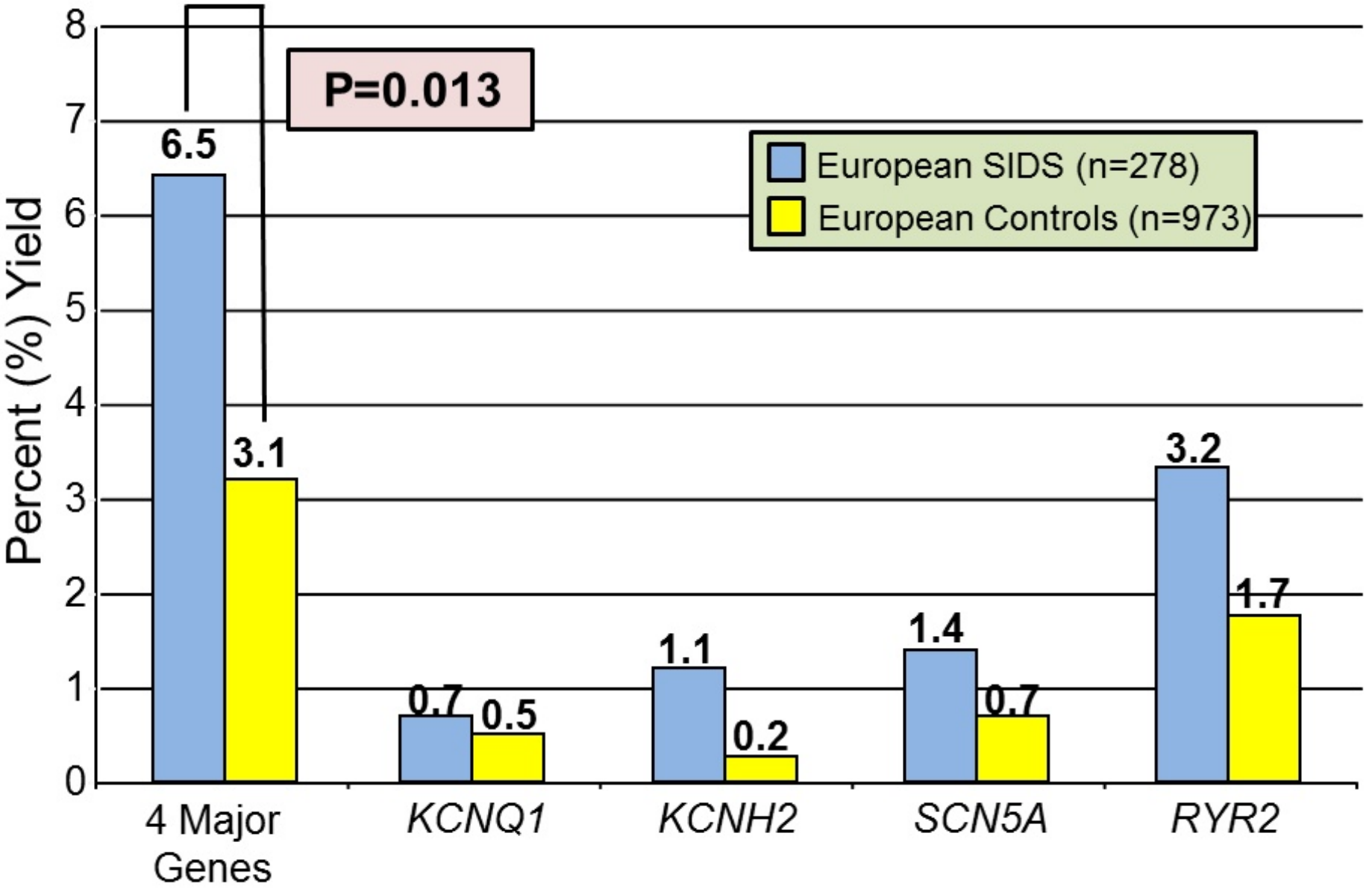
↓

ACMG Guidelines- Pathogenic or Likely Pathogenic Variants

4.1% of SIDS Host a Pathogenic/Likely Pathogenic Variant in a GHD Gene


60
50
40
30
20
10
0
Percent (%) Yield
46.3
44.6
49.6
12.6
11.5
14.9
Overall (n=419)
European (n=278)
Mixed-European (n=141)
Ultra-Rare Variants
Potentially Informative Variants


P=0.0075
P=0.0038
≤ 4
> 4
Percent (%) Yield
30
25
20
15
10
5
0
10.4
37
23.1
15
8.9
21
25.8
11
13.4
16
22.7
5
354
65
235
43
119
22
Overall (n=419)
European (n=278)
Mixed European (n=141)


P=0.013
European SIDS (n=278)
European Controls (n=973)
Percent (%) Yield
0
1
2
3
4
5
6
7
8
6.5
3.1
0.7
0.5
1.1
0.2
1.4
0.7
3.2
1.7
4 Major Genes
KCNQ1
KCNH2
SCN5A
RYR2

Cardiac Genetic Predisposition in Sudden Infant Death

David J. Tester BS[1*], Leonie CH Wong MBBChir[2,7*], Pritha Chanana MS[1], Amie Jaye MSc[3], David FitzPatrick MD[4], Margaret Evans MBChB[5], Peter Fleming PhD[6], Iona Jeffrey MBChB[7], Marta Cohen MD[8], Jacob Tfelt-Hansen MD DMSc[9,10], Michael A Simpson PhD[3#], Elijah R Behr MD[2,7#], Michael J Ackerman MD PhD[1#]

[1] Mayo Clinic, Rochester, Minnesota, USA

[2] St George's University of London, London, United Kingdom

[3] King's College London, London, United Kingdom

[4] University of Edinburgh, Edinburgh, United Kingdom

[5] Royal Infirmary of Edinburgh, Edinburgh, United Kingdom

[6] University of Bristol, Bristol, United Kingdom

[7] St George's University Hospitals' NHS Foundation Trust, London, United Kingdom

[8] Sheffield Children's Hospital, Sheffield, United Kingdom

[9]The Heart Centre, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; [10]Department of Medicine and Surgery, Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark

[*] Joint First Authors

[#] Joint Last Authors

SUPPORTING MATERIAL

eTable 1. Gene List

GENE	CHROMSOME POSITION	PROTEIN NAME
GENETIC HEART DISEASE (GHD)-ASSOCIATED GENES (n=90)		
AARS2	6:44267391-44281063	alanyl-TRNA synthetase 2, mitochondrial
ABCC9	12:21958108-22089608	ATP-binding cassette, sub-family C (CFTR/MRP), member 9
ACTC1	15:35080297-35088340	actin, alpha, cardiac muscle 1
ACTN2	1:236849808-236927931	actinin, alpha 2
AKAP9	7:91570181-91739987	A kinase (PRKA) anchor protein (yotiao) 9
ANK2	4:113970832-114304896	ankyrin 2
ANKRD1	10:92671853-92681033	ankyrin repeat domain 1 (cardiac muscle)
BAG3	10:121410882-121437331	Bcl2-associated athanogene 3
CACNA1C	12:2162464-2802108	calcium channel, voltage-dependent, L type, alpha 1C subunit
CACNA2D1	7:81575760-82073114	calcium channel, voltage-dependent, alpha 2/delta subunit 1

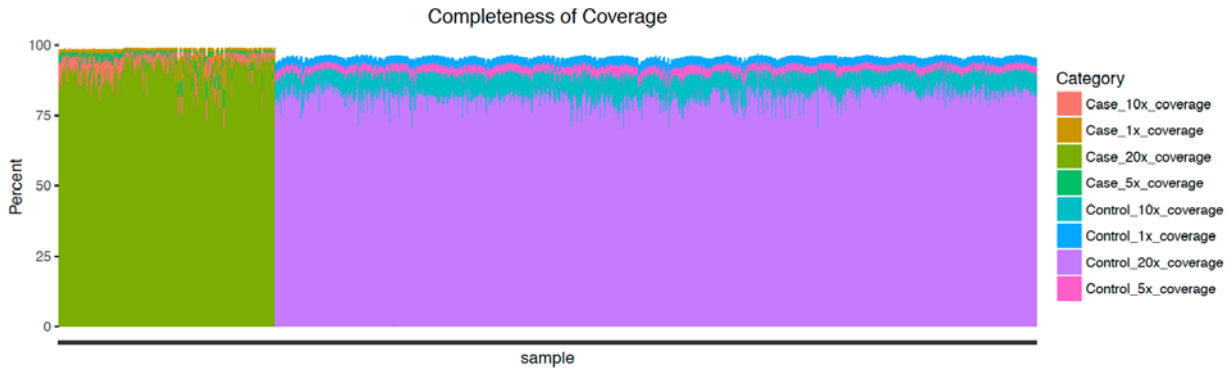
CACNB2	10:18629666-18830038	calcium channel, voltage-dependent, beta 2 subunit
CALM1	14:90863327-90874605	calmodulin 1
CALM2	2:47387221-47403740	calmodulin 2
CALM3	19:47104493-47114050	calmodulin 3
CALR3	19:16589868-16607003	calreticulin 3
CASQ2	1:116242628-116311402	calsequestrin 2 (cardiac muscle)
CAV3	3:8775486-8788451	caveolin 3
CRYAB	11:111779289-111782738	crystallin, alpha B
CSRP3	11:19203587-19232120	cysteine and glycine-rich protein 3 (cardiac LIM protein)
CTNNA3	10:67672276-69455927	catenin alpha 3
DES	2:220283099-220291461	desmin
DPP6	7:153749765-154685161	dipeptidyl peptidase like 6
DSC2	18:28646007-28682378	desmocollin 2
DSG2	18:29078006-29128971	desmoglein 2
DSP	6:7541808-7586950	desmoplakin
DTNA	18:32335940-32471808	dystrobrevin alpha
EYA4	6:133562489-133853258	eyes absent homolog 4 (Drosophila)
FHL2	2:105977283-106015508	four and a half LIM domains 2
FKTN	9:108320411-108403399	fukutin
GATAD1	7:92076767-92088150	GATA zinc finger domain containing 1
GJA1	6:121756838-121770873	gap junction protein, alpha 1, 43kDa (connexin 43)
GPD1L	3:32148003-32210205	glycerol-3-phosphate dehydrogenase 1-like
HCN4	15:73612200-73661605	hyperpolarization activated cyclic nucleotide-gated potassium channel 4
JPH2	20:42740335-42816218	junctophilin 2
JUP	17:39910856-39942950	junction plakoglobin
KCND3	1:112318431-112531777	potassium voltage gated channel, Shal-related family, member 3
KCNE1	21:35818989-35883613	potassium voltage-gated channel, Isk-related family, member 1
KCNE2	21:35736323-35743688	potassium voltage-gated channel, Isk-related family, member 2
KCNE3	11:74165886-74178673	potassium voltage-gated channel, Isk-related family, member 3
KCNH2	7:150642049-150675403	potassium voltage-gated channel, subfamily H (eag-related), member 2
KCNJ2	17:68165676-68176189	potassium inwardly-rectifying channel, subfamily J, member 2
KCNJ5	11:128761251-128790930	potassium inwardly-rectifying channel, subfamily J, member 5
KCNJ8	12:21917889-21927755	potassium inwardly-rectifying channel, subfamily J, member 8
KCNQ1	11:2466221-2870339	potassium voltage-gated channel, KQT-like subfamily, member 1
LAMA4	6:112429963-112575917	laminin, alpha 4
LDB3	10:88428428-88495825	LIM binding domain 3 (ZASP)
LMNA	1:156084498-156109880	lamin A/C
MIB1	18:19321281-19450918	mindbomb E3 ubiquitin protein ligase 1
MTO1	6:74171488-74211175	mitochondrial TRNA translation optimization 1
MYBPC3	11:47352957-47374253	myosin binding protein C, cardiac
MYH6	14:23851199-23877486	myosin, heavy chain 6, cardiac muscle, alpha
MYH7	14:23881947-23904927	myosin, heavy chain 7, cardiac muscle, beta

MYL2	12:111348623-111358381	myosin, light chain 2, regulatory, cardiac, slow
MYL3	3:46899362-46904973	myosin, light chain 3, alkali; ventricular, skeletal, slow
MYLK2	20:30407176-30422492	myosin light chain kinase 2
MYOM1	18:3066805-3220106	myomesin 1, 185kDa
MYOZ2	4:120056939-120108944	myozenin 2
MYPN	10:69868994-69971774	myopalladin
NEBL	10:21068902-21186531	nebulette
NEXN	1:78354313-78408909	nexilin (F actin binding protein)
PDLIM3	4:186422905-186456766	PDZ and LIM domain 3
PKP2	12:32943788-33049690	plakophilin 2
PLN	6:118869461-118881893	phospholamban
PRDM16	1:2985775-3355185	PR Domain 16
PRKAG2	7:151253197-151574210	protein kinase, AMP-activated, gamma 2 non-catalytic subunit
PSEN1	14:73603155-73690399	presenilin 1
PSEN2	1:227058264-227083799	presenilin 2
RANGRF	17:8191815-8193410	RAN guanine nucleotide release factor (MOG1)
RBM20	10:112404155-112599227	RNA binding motif protein 20
RYR2	1:237205505-237997288	ryanodine receptor 2 (cardiac)
SCN1B	19:35521588-35531352	sodium channel, voltage-gated, type I, beta
SCN3B	11:123499897-123525312	sodium channel, voltage-gated, type III, beta
SCN4B	11:118004092-118023535	sodium channel, voltage-gated, type IV, beta
SCN5A	3:38589557-38691119	sodium channel, voltage-gated, type V, alpha
SDHA	5:218356-256815	Succinate Dehydrogenase Complex Flavoprotein Subunit A
SGCD	5:155753756-156194799	sarcoglycan, delta (dystrophin-associated glycoprotein)
SNTA1	20:31995761-32031698	syntrophin, alpha 1
TAZ	X:153639892-153650065	tafazzin
TCAP	17:37820440-37822808	titin-cap (telethonin)
TGFB3	14:76424545-76447534	transforming growth factor, beta 3
TMEM43	3:14166440-14185179	transmembrane protein 43
TMPO	12:98909408-98929412	thymopoietin
TNNC1	3:52485118-52488086	troponin C type 1
TNNI3	19:55663138-55669100	troponin I type 3 (cardiac)
TNNT2	1:201328142-201342382	troponin T type 2 (cardiac)
TPM1	15:63334831-63358292	tropomyosin 1 (alpha)
TRDN	6:123785398-123958054	triadin
TTN	2:179390716-179672150	titin
TXNRD2	22:19863045-19929333	thioredoxin reductase 2
VCL	10:75757872-75879910	vinculin

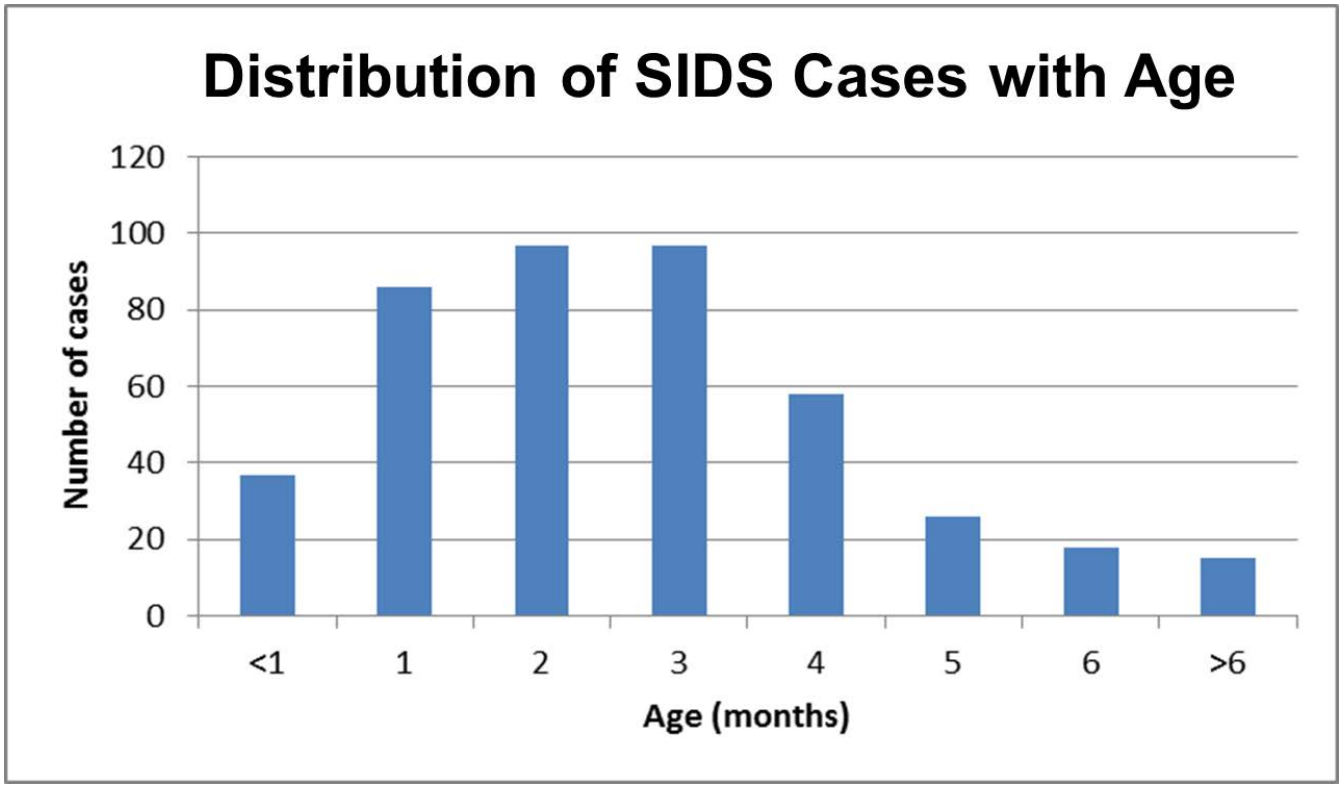
Genes are listed in alphabetical order.

SUPPLEMENTAL RESULTS

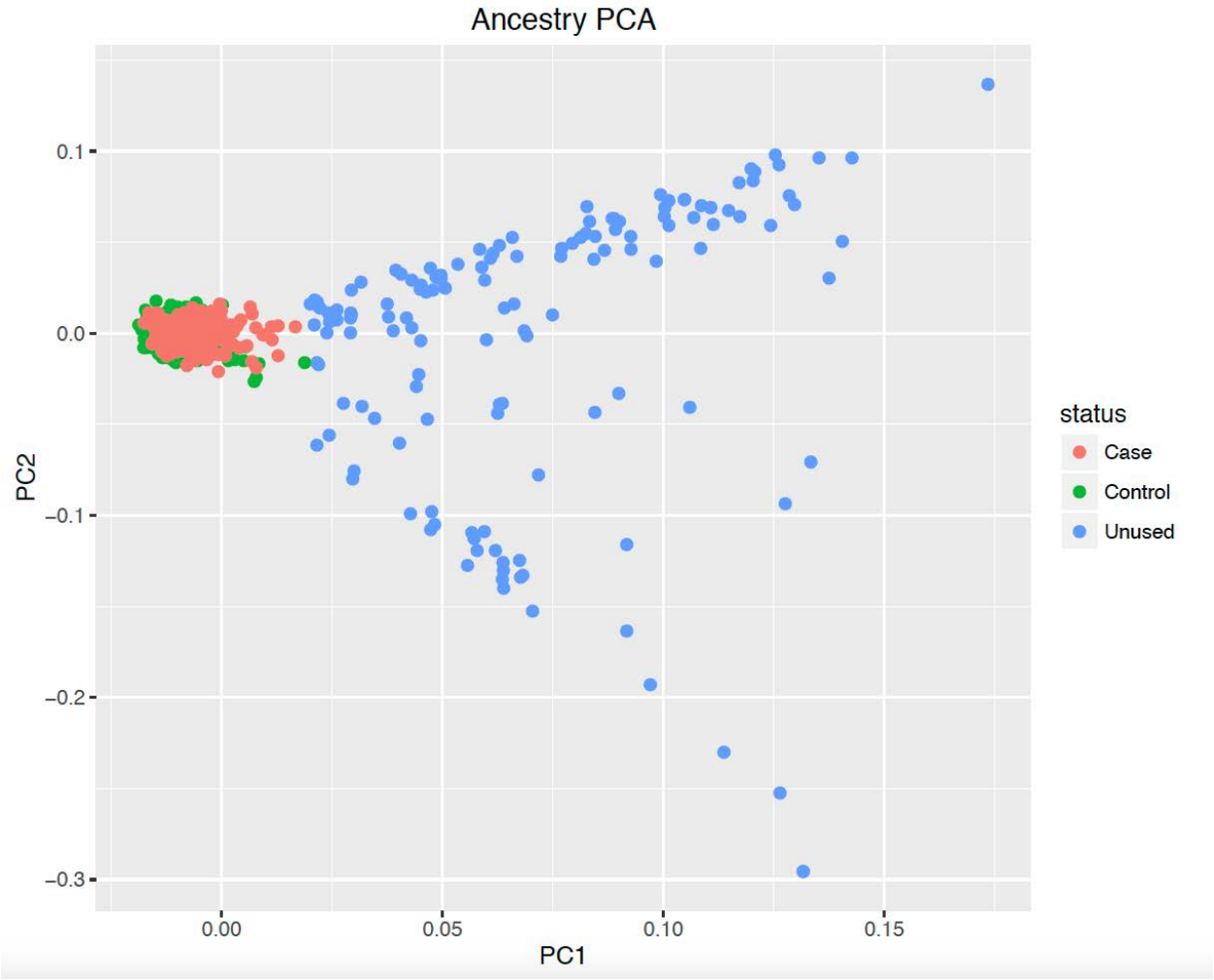
eFigure 1. Proportion of the Genocode-defined exome covered by more than 1x, 5x, 10x and 20x.



eFigure 2. Age Distribution of the SIDS Cases



eFigure 3. Components 1 and 2 from a Multidimensional Scaling Analysis of 3847 Common Variants Located Outside of Regions of the Genome Where There is Extensive Linkage Disequilibrium across the Entire Case and Control Cohorts


Ancestry PCA
0.1
0.0
−0.1
−0.2
−0.3
PC2
0.00
0.05
0.10
0.15
PC1
status
Case
Control
Unused

eTable 2. GHD-Associated Gene-Specific Yields of Ultra-Rare Nonsynonymous Variants".

Gene	NCBI mRNA Ref_Seq	Missense Z-Score	pLI	Overall Yield (n=419)	European Ancestry Yield (n=278)	Mixed-European Ancestry Yield (n=141)
Major Channelopathy Genes						
KCNQ1	NM_000218	2.73	0	2 (0.5%)	2 (0.7%)	0
KCNH2	NM_000238	4.81	1	5 (1.2%)	3 (1.1%)	2 (1.4%)
SCN5A	NM_198056	2.53	1	6 (1.4%)	4 (1.4%)	2 (1.4%)
RYR2	NM_001035.2	5.21	1	12 (2.9%)	9 (3.2%)	3 (2.1%)
Minor Channelopathy Genes						
AKAP9	NM_005751	-2.75	0	21 (5.0%)	11 (4.0%)	10 (7%)
ANK2	NM_001148	1.06	1	13 (3.1%)	8 (2.9%)	5 (3.5%)
CACNA1C	NM_000719	6.41	1	5 (1.2%)	3 (1.1%)	2 (1.4%)
CACNA2D1	NM_000722	3.1	1	6 (1.4%)	5 (2.8%)	1 (0.7%)
CACNB2	NM_201590	0.36	0	0	0	0
CALM1	NM_006888	3.21	0.89	0	0	0
CALM2	NM_001743	2.79	0.86	0	0	0
CALM3	NM_005184	2.92	0.58	0	0	0
CASQ2	NM_001232.3	-1.08	0	1 (0.2%)	1 (0.36%)	0
CAV3	NM_001234	1.19	0.34	0	0	0
DPP6	NM_130797	1.81	0.97	7 (1.7%)	2 (0.7%)	5 (3.5%)
GJA1	NM_000165	1.57	0.22	2 (0.5%)	2 (0.7%)	0
GPD1L	NM_015141	1.2	0.01	2 (0.5%)	1 (0.36%)	1 (0.7%)
HCN4	NM_005477	4.83	0.23	5 (1.2%)	5 (1.8%)	0
KCND3	NM_004980	5.35	0.8	2 (0.5%)	1 (0.36%)	1 (0.7%)
KCNE1	NM_001270402	-0.46	0	0	0	0
KCNE2	NM_172201	-0.23	0	1 (0.2%)	1 (0.36%)	0
KCNE3	NM_005472	-0.18	0.44	0	0	0

KCNJ2	NM_000891	3.02	0.82	1 (0.2%)	1(0.36%)	0
KCNJ5	NM_000890	1.46	0.31	0	0	0
KCNJ8	NM_004982	3.66	0.26	1 (0.2%)	1 (0.36%)	0
RANGRF	NM_016492	0.25	0.02	0	0	0
SCN1B	NM_001037	2.12	0.21	5 (1.2%)	2 (0.7%)	3 (2.1%)
SCN3B	NM_018400	1.12	0.35	2 (0.5%)	0	2 (1.4%)
SCN4B	NM_001142349	-0.13	0	0	0	0
SNTA1	NM_003098	0.77	0.43	3 (0.7%)	1 (0.36%)	2 (1.4%)
TRDN	NM_001256021	-1.76	0	1 (0.2%)	1 (0.36%)	0
Cardiomyopathy Genes						
AARS2	NM_020745	0.72	0	7 (1.7%)	2 (0.7%)	5 (3.5%)
ABCC9	NM_005691	4.89	0	3 (0.7%)	3 (1.1%)	0
ACTC1	NM_005159	5.25	0.95	0	0	0
ACTN2	NM_001103	1.76	1	2 (0.5%)	1 (0.36%)	1 (0.7%)
ANKRD1	NM_014391	-0.1	0	3 (0.7%)	0	3 (2.1%)
BAG3	NM_004281	-1.01	0.53	4 (1.0%)	1 (0.36%)	3 (2.1%)
CALR3	NM_145046	-0.1	0	3 (0.7%)	3 (1.1%)	0
CRYAB	NM_001885	0.38	0.01	3 (0.7%)	2 (0.7%)	1 (0.7%)
CSRP3	NM_003476	0.66	0	2 (0.5%)	2 (0.7%)	0
CTNNA3	NM_013266	-2.57	0	7 (1.7%)	5 (1.8%)	2 (1.4%)
DES	NM_001927	2.34	0	1 (0.2%)	0	1 (0.7%)
DSC2	NM_024422	-1.02	0	6 (1.4%)	5 (1.8%)	1 (0.7%)
DSG2	NM_001943	-1.2	0	3 (0.7%)	3 (1.1%)	0
DSP	NM_004415	0.91	1	20 (4.8%)	11 (4.0%)	9 (6.4%)
DTNA	NM_001390	1.17	0.92	1 (0.2%)	1 (0.36%)	0
EYA4	NM_004100	0.64	0.13	1 (0.2%)	0	1 (0.7%)
FHL2	NM_001450	0.35	0	0	0	0
FKTN	NM_006731	-0.64	0	2 (0.5%)	1 (0.36%)	1 (0.7%)
GATAD1	NM_021167	0.51	0.51	1 (0.2%)	0	1 (0.7%)
JPH2	NM_020433	3.93	0.01	2 (0.5%)	2 (0.7%)	0
JUP	NM_002230	0.93	0.04	3 (0.7%)	2 (0.7%)	1 (0.7%)
LAMA4	NM_001105206	-0.67	0	11 (2.6%)	8 (2.9%)	3 (2.1%)
LDB3	NM_007078	0.32	0	8 (1.9%)	5 (1.8%)	3 (2.1%)

LMNA	NM_170707	3.37	0.99	1 (0.2%)	1 (0.36%)	0
MIB1	NM_020774	3.51	0	2 (0.5%)	1 (0.36%)	1 (0.7%)
MTO1	NM_133645	0.68	0	3 (0.7%)	3 (1.1%)	0
MYBPC3	NM_000256	0.69	0	7 (1.7%)	5 (1.8%)	2 (1.4%)
MYH6	NM_002471	2.87	0	10 (2.4%)	6 (2.2%)	4 (2.8%)
MYH7	NM_000257	6.54	0	4 (1.0%)	3 (1.1%)	1 (0.7%)
MYL2	NM_000432	0.86	0.02	0	0	0
MYL3	NM_000258	0.75	0.89	0	0	0
MYLK2	NM_033118	-0.73	0.22	1 (0.2%)	0	1 (0.7%)
MYOM1	NM_003803	-0.35	0	8 (1.9%)	6 (2.2%)	2 (1.4%)
MYOZ2	NM_016599	0.03	0.02	2 (0.5%)	2 (0.7%)	0
MYPN	NM_032578	-0.35	0.07	5 (1.2%)	3 (1.1%)	2 (1.4%)
NEBL	NM_006393	-2.45	0	7 (1.7%)	4 (1.4%)	3 (2.1%)
NEXN	NM_144573	-1.32	0	3 (0.7%)	3 (1.1%)	0
PDLIM3	NM_001114107	-0.4	0	5 (1.2%)	2 (0.7%)	3 (2.1%)
PKP2	NM_004572	-0.8	0	3 (0.7%)	3 (1.1%)	0
PLN	NM_002667	0.57	0.11	1 (0.2%)	1 (0.36%)	0
PRDM16	NM_022114	2.06	1	4 (1.0%)	2 (0.7%)	2 (1.4%)
PRKAG2	NM_016203	1.85	0.98	0	0	0
PSEN1	NM_000021	1.81	1	0	0	0
PSEN2	NM_000447	0.53	0.03	1 (0.2%)	1 (0.36%)	0
RBM20	NM_001134363	NA (poor coverage)	NA	5 (1.2%)	3 (1.1%)	2 (1.4%)
SDHA	NM_004168	2.32	0	3 (0.7%)	3 (1.1%)	0
SGCD	NM_001128209	-0.23	0	2 (0.5%)	1 (0.36%)	1 (0.7%)
TAZ	NM_000116	2.42	0.97	1 (0.2%)	1 (0.36%)	0
TCAP	NM_003673	0.45	0.08	0	0	0
TGFB3	NM_003239	2.2	0.92	0	0	0
TMEM43	NM_024334	-0.88	0	5 (1.2%)	4 (1.4%)	1 (0.7%)
TMPO	NM_003276	-0.99	0	3 (0.7%)	2 (0.7%)	1 (0.7%)
TNNC1	NM_003280	2.22	0.51	0	0	0
TNNI3	NM_000363	1.88	0.17	0	0	0
TNNT2	NM_000364	1.54	0.01	1 (0.2%)	0	1 (0.7%)

TPM1	NM_000366	3.42	0.8	0	0	0
TTN*	NM_003319	-5.48	0	5 (1.2%)	2 (0.7%)	3 (2.1%)
TXNRD2	NM_006440	0.32	0	0	0	0
VCL	NM_014000	3.11	0.99	3 (0.7%)	3 (1.1%)	0

*only radical (i.e. nonsense, frameshift, splice-error) variants were considered. Z-score and pLI represent variant constraint scores based on the ExAC data for missense and radical variants respectively.[1] Positive Z-scores represent intolerability. pLI scores range from 0 to 1, with 1 representing the highest intolerability. NA = not available.

REFERENCES

1. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285-291.