
Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
10.1016/j.jpeds.2018.08.011

Link to publication record in Explore Bristol Research

PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at https://www.sciencedirect.com/science/article/pii/S0022347618311211 . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms
Exome-Wide Rare Variant Analyses in Sudden Infant Death Syndrome

Short Title: Tester – Whole Exome Sequencing in SIDS

David J. Tester BS1*, Leonie CH Wong MBChir2,3*, Pritha Chanana MS4, Belinda Gray MBBS PhD2,3,5,6, Amie Jaye MSc7, Jared M. Evans MS4, Margaret Evans MBChB8, Peter Fleming PhD9, Iona Jeffrey MBChB10,11, Marta Cohen MD12, Jacob Tfelt-Hansen MD DMSc13,14, Michael A. Simpson PhD7#, Elijah R. Behr MD2,3#, Michael J. Ackerman MD PhD1#

1 Departments of Cardiovascular Medicine (Division of Heart Rhythm Services), Pediatrics (Division of Pediatric Cardiology), and Molecular Pharmacology & Experimental Therapeutics (Windland Smith Rice Sudden Death Genomics Laboratory), Mayo Clinic, Rochester, Minnesota, USA
2Molecular and Clinical Sciences Research Institute, St George's University of London, London, United Kingdom
3 Cardiology Clinical Academic Group, St George's University Hospitals’ NHS Foundation Trust, London, United Kingdom
4Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA
5Agnes Ginges Centre for Molecular Cardiology, Centenary Institute, Sydney, Australia
6Sydney Medical School, University of Sydney, Australia
7 Medical and Molecular Genetics, Guy's Hospital, King's College London, London, United Kingdom.
8Royal Infirmary of Edinburgh, Edinburgh, United Kingdom
9 Centre for Child and Adolescent Health, Bristol Medical School, University of Bristol, Bristol, United Kingdom
10Department of Cellular Pathology, St George's University of London, London, United Kingdom
11Department of Cellular Pathology, St George's University Hospitals NHS Foundation Trust, London, United Kingdom
12Histopathology Department, Sheffield Children's Hospital NHS FT, Sheffield, United Kingdom
13 Department of Cardiology, The Heart Centre, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark.
14Department of Forensic Medicine, Faculty of Medical Sciences, University of Copenhagen, Copenhagen, Denmark

* Joint First Authors
# Joint Senior Authors

Funding Sources: 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. BG is the
recipient of a National Health and Medical Research Council, Australia (NHMRC) Early Career Fellowship (#1122330).

ERB is supported by the Higher Education Funding Council for England and receives funds from the Robert Lancaster Memorial Fund sponsored by McColl’s Retail Group Ltd. The authors also acknowledge support from the UK Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy’s and St. Thomas’ National Health Service (NHS) Foundation Trust in partnership with King’s College London and King’s College Hospital NHS Foundation Trust. The authors also wish to acknowledge Newlife research funding for the establishment of the Edinburgh SIDS DNA cohort. There are no relevant conflicts of interest to this work.

**Disclosure:** MJA is a consultant for Audentes Therapeutics, Boston Scientific, Gilead Sciences, Invitae, Medtronic, MyoKardia, and St. Jude Medical. MJA and Mayo Clinic have potential equity/royalty relationships (without remuneration so far) with AliveCor, Blue Ox Health, and StemoniX. However, none of these entities were involved in this study in any way. David J. Tester wrote the first draft of the manuscript. No honorarium, grant, or other form of payment was given to anyone to produce the manuscript.

**Corresponding Authors:**

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

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

**LIST OF ABBREVIATIONS**

BrS – Brugada syndrome  
CADD – Combined annotation dependent depletion  
CPVT – Catecholaminergic polymorphic ventricular tachycardia  
EPACTS – Efficient and parallelizable association container toolbox  
ExAC – Exome aggregation consortium  
HCM – Hypertrophic cardiomyopathy  
IPA – Ingenuity pathway analysis  
LQTS – Long QT syndrome  
MAF - Minor allele frequency  
SIDS – Sudden infant death syndrome  
WES – Whole exome sequencing
Abstract

Objective: To determine the underlying monogenic basis for sudden infant death syndrome (SIDS) with an exome-wide focus.

Study Design: A cohort of 427 unrelated SIDS cases (257 males; average age = 2.7 ± 1.9 months) underwent whole exome sequencing (WES). Exome-wide rare variant analyses were carried out with 278 European ancestry SIDS cases (173 males; average age = 2.7 ± 1.98 months) and 973 ethnic-matched controls based on six genetic models. Ingenuity Pathway Analysis was also performed. The cohort was collected in collaboration with coroners, medical examiners, and pathologists by St George's University of London, UK and Mayo Clinic, Rochester, Minnesota, USA. WES was performed at the Genomic Laboratory, Kings College London, UK or Mayo Clinic’s Medical Genome Facility, Rochester, Minnesota.

Results: While no exome-wide significant (p<2.5x10^{-6}) difference in burden of ultra-rare variants was detected for any gene, 405 genes had a higher prevalence (p<0.05) of ultra-rare non-synonymous variants among cases with seventeen genes at p<0.005. Some of these potentially overrepresented genes may represent biologically plausible novel candidate genes for the monogenic basis for some of SIDS. The top canonical pathway identified was glucocorticoid biosynthesis (p=0.01).

Conclusions: The lack of exome-wide significant genetic associations indicates an extreme heterogeneity of etiologies underlying SIDS. Our approach to understanding the genetic mechanisms of SIDS has far reaching implications for the SIDS research community as a whole and may catalyze new evidence-based SIDS research across multiple disciplines. Perturbations in glucocorticoid biosynthesis may represent a novel SIDS-associated biological pathway for future SIDS investigative research.

Keywords: inherited cardiac conditions, molecular autopsy, sudden infant death syndrome, 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. It is 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, also referred to as sudden unexplained infant death (SUID) 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 occurring in 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] Potentially lethal inherited genetic heart diseases 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 (< 10%) of SIDS cases.[10, 13, 15-27] [28]

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 approaches that recognize established environmental risk factors for SIDS, with an average cohort size of just 125 SIDS cases.[13] Although one may postulate that the genetics of SIDS is most likely multigenic and complex, no studies have attempted to elucidate the underlying genetic basis in a single SIDS cohort across multiple biological pathways in an unbiased manner. Here, using
whole exome sequencing (WES), we conducted an exome-wide analysis of rare protein-altering variation followed by biological Ingenuity Pathway Analysis (IPA) in a cohort of nearly 300 unrelated Caucasian SIDS cases.

**Methods**

**Study Population**

As previously described, 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.[28] Because of the lack of uniformity in procedures and reporting between medical examiner offices in the US, 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, clearly demonstrated to have experienced 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.

**Control Population**

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

**Whole Exome Sequencing**

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 as described previously.[28] 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. To avoid potential confounding due to population stratification, ancestry estimation was undertaken using the first two dimensions of Multidimensional Scaling (MDS) using Euclidean distance undertaken with the King software package. Quality control metrics excluded 7 cases.[28] The resulting SIDS cases (N = 278) and European controls (N = 973) that formed a homogeneous cluster on the first two principal components were included in the case-control rare variant analysis.[28]

**Exome-Wide Rare Variant Analysis**

The 278 European Caucasian SIDS cases and 973 ethnic-matched controls underwent a case-control simple collapsing exome-wide rare variant association analysis using the Efficient and Parallelizable Association Container Toolbox (EPACTS).[30] Cases and controls were coded with a binary indicator according to the presence/absence of a “qualifying” variant in each gene based on six genetic models: dominant non-synonymous (amino acid altering), dominant non-synonymous-combined annotation dependent depletion (CADD ≥ 20: equivalent to a probability of 0.99 that the variant has a functional impact,[31]), dominant putative loss-of-function (i.e. nonsense, frame-shift, and canonical splice-site errors), recessive non-synonymous, recessive non-synonymous-CADD ≥ 20, and recessive putative loss-of-function.

For all models, only “ultra-rare” variants with a minor allele frequency (MAF) < 0.00005 (1:20,000 alleles) derived from the Exome Aggregation Consortium (ExAC)[32] were included. Deviation from the null model of no excess of rare variation in cases was evaluated with a one-tailed Fisher’s exact test. Despite being too stringent, a Bonferroni-adjusted p-value threshold of < 2.5x10^-6 was considered as exome-wide significant.

**Ingenuity Pathway Analysis**

An Ingenuity Pathway Analysis (IPA) was used to analyze a list of 405 genes (p-value<0.05)
obtained by performing an exome-wide rare variant comparative analysis between European SIDS cases and ethnic-matched controls. IPA is a web-based tool that is used for discovery and visualization of significant pathways and networks, and close correlations with disease and functional classes. Data were analyzed with QIAGEN’s Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity) using default analysis settings.[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). The significance of the association between the data set (i.e. 405 “SIDS genes”) and the IPA canonical pathway was measured in 2 ways: 1) A ratio of the number of genes from the data set that map to the pathway divided by the total number of genes that map to the canonical pathway. 2) Fisher’s exact test right-tailed was used to calculate a \( p \)-value determining the probability that the association between the genes in the dataset and the canonical pathway is explained by chance alone.[33]

**Results**

**Demographics**

The cohort consisted of 278 cases (173 male infants, 105 female infants; average age = 2.7 ± 1.98 months). The epidemiologically high risk age group of 2–4 months (55.4%) and male gender (62.9%) accounted for the majority of the cases. Sleep characteristics were known in 60% of the cohort (Table 1).

**Exome-Wide Rare Variant Analysis**

Exome-wide rare variant analyses were performed with the 278 European ancestry cases and 973 matched controls. None of the estimated 20,000 genes within the human genome were associated with SIDS when considering an exome-wide significant \( p \)-value threshold of \( < 2.5 \times 10^{-6} \) with any of the 6 aforementioned inheritance models.
There were 405 genes with potential case:control, ultra-rare (MAF < 0.00005), non-synonymous variant over-representation at p < 0.05 threshold (Table 2; Online) with 17 genes reaching p<0.005 (Table 3). Online Mendelian Inheritance in Man (OMIM) disease[34] and PathCards SuperPathway [35] associations for each gene is listed in Table 4; Online. The prevalence of ultra-rare, non-synonymous variants with a CADD score > 20 was higher (p<0.005) in cases versus controls in 10 genes (Table 5; Online). Some of these genes may represent novel, biologically plausible candidate genes for SIDS. For example, 6/278 (2.2%) SIDS cases hosted an NR3C2 (MIM: 600983) variant compared to only 2/973 (0.21%) controls (p=0.000778). Variants in the CHRM3 (MIM: 118494)-encoded cholinergic receptor muscarinic-3 were over-represented in SIDS cases (4/278 [1.4%]) compared to controls (0/973, p=0.002), and LHX9 (MIM: 606066) variants were also over-represented in SIDS cases (5/278 [1.8%]) compared to controls (0/973, p=0.000527).

Ingenuity Pathway Analysis (IPA)

IPA was performed on the aforementioned 405 genes from our dominant non-synonymous ultra-rare variants analysis with a p-value <0.05. IPA identified 5 canonical pathways with a p-value < 0.05 (Table 6). The top canonical pathway involved glucocorticoid biosynthesis (p=0.0107, Table 6). Of the 8 genes within the glucocorticoid biosynthesis pathway, 2 genes (25%) were present among the 405 candidate genes. Ultra-rare CYP17A1 ([MIM: 609300], 4/278 [1.4%] vs 1/973 [0.1%], p=0.0099), and CYP11B2 ([MIM: 124080], 6/278 [2.2%] vs 3/973 [0.3%], p=0.0052) non-synonymous variants were both over-represented in the cases versus controls (p=0.0001 when considered together).

Discussion

This manuscript details the genetic insights gleaned from WES in the largest cohort of unrelated SIDS cases with the first ancestry matched case-control burden analysis of ultra-rare variants involving nearly 300 cases and 1000 healthy controls to identify potential novel monogenic causes for SIDS. While no genes reached the very strict Bonferroni-corrected exome-wide significance
threshold of \( p < 2.5 \times 10^{-6} \), 405 genes showed a potential over-representation of ultra-rare, non-synonymous variants. However, variants were observed in less than 2% of the SIDS cohort for 90% of these 405 genes, thus confirming prior speculation that SIDS is a highly heterogeneous condition and likely results from a complex interplay of genetic and environmental factors rather than from a small number of highly penetrant monogenic disorders. In addition, at the p-value threshold of \( p < 0.05 \), one might expect 1000 genes to surface as seemingly over-represented in cases just by chance (i.e. 20,000 genes divided by 20). In other words, this study strongly suggests that the vast majority of SIDS is NOT monogenetically-mediated at least by ultra-rare genetic variation localizing to the 2% of the genetic code called the exome. Although it remains to be confirmed or refuted, much of SIDS may not have any genetic component at all.

Nevertheless, some of these genes identified are appealing pathobiologically and could represent novel, monogenic causes for a small minority of SIDS cases. For example, there was potential over-representation (\( p < 0.005 \)) of \( NR3C2 \), \( CHRM3 \), and \( LHX9 \) variants in SIDS cases compared to controls. These three genes may represent biologically plausible candidates for SIDS pathogenicity that could account independently for a small subset of SIDS cases (~2%).

The \( NR3C2 \) gene encodes for a mineralocorticoic receptor that mediates aldosterone activity and is responsible for autosomal dominant pseudohypoaldosternism Type 1 (adPHA1) which is a rare disorder characterized by renal resistance to aldosterone, with salt-wasting, hyperkalemia, and metabolic acidosis[36]. If undetected and untreated, adPHA1 may present with neonatal lethality owing to sever hyponatremia, metabolic acidosis, or hyperkalemia-mediated arrhythmia and cardiopulmonary arrest.[37, 38] Interestingly, conditional targeted cardiomyocyte over-expression of mineralocorticoid receptor in mice lead to early sudden death without cardiac structural alteration as a result of prolonged ventricular repolarization and spontaneous and triggered ventricular arrhythmias associated with ion channel remodeling.[39]

The \( CHRM3 \) gene encodes for cholinergic receptor muscarinic 3 (M3-receptor). Muscarinic acetylcholine receptors are G-protein-coupled receptors that play a key role in parasympathetic
acetylcholine neurotransmission for a variety of physiological functions within the brain, heart, and respiratory system.[40] In the respiratory system, muscarinic M₃-receptors mediate smooth muscle contraction for respiration.[40] In the heart, M₃-receptors may play a role in the regulation and maintenance of cardiac function.[41, 42] The M₃-receptors activate the delayed rectifying potassium current $I_{KM3}$ to participate in the regulation of heart rate, cardiac resting membrane potential, and cardiac repolarization.[43] The M₃-receptor also interacts with gap-junctional channel connexin 43 (Cx43) to maintain cell-cell communication and excitation propagation in the heart.[44] Interestingly, Cx43 mutations leading to gap junction loss were implicated previously in some cases of SIDS.[24] In 2010, cardiac muscarinic receptor over-expression in left ventricular heart samples was shown to be associated with SIDS.[45] In 1995, Kinney and colleagues provided evidence that deficiency of muscarinic receptor binding in the arcuate nucleus of the medulla may be associated with an increase in the probability of death from SIDS.[46]

**LHX9-encoded Lhx9** is a member of the apterous group of the LIM-homeodomain family of evolutionary conserved transcription factors crucial for the correct development of many organs including the heart and the nervous system, including the forebrain, thalamus, hypothalamus, and pineal gland.[47-51] Sparks and Hunsaker reported a significant reduction in the size of the pineal gland in SIDS cases compared to controls.[52] In mice, Lhx9 expression is maintained after birth in the hippocampal formation and restricted to the dentate gyrus.[51] Recently, Kinney identified dentate gyrus abnormalities as a potential developmental vulnerability in sudden unexplained death in infants that may lead to autonomic/respiratory instability or autonomic seizures, and sleep-related sudden death of the vulnerable infant when exposed to homeostatic stressors.[53, 54]

Whether **NR3C2**, **CHRM3**, **LHX9** or any of the other genes that may have an over-representation of ultra-rare, nonsynonymous variants in SIDS are indeed responsible for the pathogenic basis for SIDS will require substantial functional analyses and replication studies before being considered as a new pathogenic substrate for some infant deaths.

The Ingenuity Pathway Analysis has suggested a potential association between the
glucocorticoid biosynthesis signaling pathways and SIDS. Glucocorticoids and mineralocorticoid, such as aldosterone and cortisol, are hormones that regulate numerous physiological processes, including intermediate metabolism, immune function, and cardiovascular function.[55, 56] They are synthesized and released from the adrenal gland in a circadian manner and in response to stress. Nearly 3.5% of the European Caucasian cases had a heterozygote, ultra-rare, non-synonymous variant within either the CYP17A1 or CYP11B2 compared to 0.4% of controls. While the CYP17A1-encoded steroid 17-alpha-hydroxylase is involved in the synthesis of cortisol, the CYP11B2-encoded aldosterone synthase is involved in the synthesis of aldosterone.[55] Aldosterone, through increased activation of the mineralocorticoid receptor, increases calcium influx in ventricular myocytes, prolongs ventricular repolarization, and leads to triggered cardiac arrhythmias.[57] Additional investigation and replication will be necessary before invoking genetic perturbations in glucocorticoid biosynthesis signaling as a potentially novel pathogenic substrate for SIDS.

Conclusions

From the perspective of ultra-rare genetic variation in the exome, the lack of significant genetic associations suggests an extreme heterogeneity of etiologies underlying SIDS. Until or unless some of the genes identified herein emerge as true SIDS-susceptibility genes with sufficient monogenic penetrance, then monogenic SIDS, stemming from the cardiac channelopathies and other established sudden death predisposing pathways, comprises less than 10% of all SIDS. In other words, the vast majority of SIDS is NOT monogenic, at least within the exome. Whether oligo/polygenic models of genetic variation accounting for the “vulnerable infant” will emerge remains to be determined.

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 their sudden death.
References


[23] Tester DJ, Tan B-H, Medeiros-Domingo A, Song C, Makielski JC, Ackerman MJ. Loss-of-


[34] Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online


