European Journal of Human Genetics (2017) 25, 404–409

www.nature.com/ejhg

ARTICLE

Post-mortem whole-exome analysis in a large sudden infant death syndrome cohort with a focus on cardiovascular and metabolic genetic diseases

Jacqueline Neubauer*[,1], Maria Rita Lecca[2], Giancarlo Russo[2], Christine Bartsch[3], Argelia Medeiros-Domingo[4], Wolfgang Berger[5,6,7] and Cordula Haas[1]

Sudden infant death syndrome (SIDS) is described as the sudden and unexplained death of an apparently healthy infant younger than one year of age. Genetic studies indicate that up to 35% of SIDS cases might be explained by familial or genetic diseases such as cardiomyopathies, ion channelopathies or metabolic disorders that remained undetected during conventional forensic autopsy procedures. Post-mortem genetic testing by using massive parallel sequencing (MPS) approaches represents an efficient and rapid tool to further investigate unexplained death cases and might help to elucidate pathogenic genetic variants and mechanisms in cases without a conclusive cause of death. In this study, we performed whole-exome sequencing (WES) in 161 European SIDS infants with focus on 192 genes associated with cardiovascular and metabolic diseases. Potentially causative variants were detected in 20% of the SIDS cases. The majority of infants had variants with likely functional effects in genes associated with channelopathies (9%), followed by cardiomyopathies (7%) and metabolic diseases (1%). Although lethal arrhythmia represents the most plausible and likely cause of death, the majority of SIDS cases still remains elusive and might be explained by a multifactorial etiology, triggered by a combination of different genetic and environmental risk factors. As WES is not substantially more expensive than a targeted sequencing approach, it represents an unbiased screening of the exome, which could help to investigate different pathogenic mechanisms within the genetically heterogeneous SIDS cohort. Additionally, re-analysis of the datasets provides the basis to identify new candidate genes in sudden infant death.
European Journal of Human Genetics (2017) **25,** 404–409; doi:10.1038/ejhg.2016.199; published online 11 January 2017

INTRODUCTION

Sudden infant death syndrome (SIDS) is defined as the sudden and unexpected death of an infant younger than one year of age, with the onset of the fatal episode apparently occurring during sleep.[1] The cause of death remains unexplained after a thorough investigation, including performance of a complete autopsy,[2] review of the circumstances of death and the clinical history. Although the incidence rate of SIDS cases drastically decreased in the last years, SIDS is still one of the leading causes of postneonatal infant death in developed countries with a prevalence between 0.1 and 0.8 deaths per 1000 live births.[3] The occurrence of SIDS is described by a triple risk model involving (1) a critical developmental period in the first months after birth, (2) a vulnerable infant and (3) exogenous stress factors.[4] Environmental risk factors such as the prone sleeping position, sharing bed with parents, or smoking exposure during pregnancy are widely accepted stressors to expose a vulnerable infant at risk for cardiorespiratory failure or other homeostatic imbalance.[5] However, the pathophysiological mechanisms responsible for SIDS still remain poorly understood.[6,7] Genetic studies in SIDS cohorts collectively suggest that up to 15% of SIDS cases might be explained by inherited cardiac diseases not detectable during conventional forensic autopsy investigations.[8–10] Ion channelopathies such as Brugada syndrome (BrS), long QT syndrome (LQTS), short QT syndrome or catecholaminergic polymorphic ventricular tachycardia (CPVT), are described as disrupted channel functions causing disturbed ion current flow and lethal cardiac arrhythmias.[11] Cardiomyopathies are characterized by structural abnormalities in the heart, such as hypertrophic cardiomyopathy, dilated cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy (ARVC) and left ventricular non-compaction cardiomyopathy.[12] Additionally, undiagnosed inherited metabolic diseases such as medium-chain-acyl-CoA dehydrogenase (MCAD) deficiency or glucose metabolism deficiency might contribute to the cause of death in another 1% of the SIDS infants.[13]

Post-mortem genetic testing by using massive parallel sequencing (MPS) approaches represents an efficient and rapid strategy to investigate potential disease-causing mechanisms that remained undetected during conventional autopsy and may help to identify the cause of death in some of the SIDS infants and to detect families at risk for further sudden deaths.[14,15] A first MPS-based genetic investigation in 104 genes associated with sudden cardiac death had identified likely pathogenic variants in two cardiomyopathy-associated genes (*PKP2* and *VCL*) in one representative SIDS case.[10] A second MPS-based targeted sequencing study in 47 Danish cases of sudden unexpected death in infancy reported likely causative variants in cardiac disease-

[1]Zurich Institute of Forensic Medicine, University of Zurich, Zurich, Switzerland; [2]Functional Genomics Center Zurich (FGCZ), University of Zurich/ETH Zurich, Zurich, Switzerland; [3]University of Zurich, Zurich, Switzerland; [4]Department of Cardiology, Inselspital, University Hospital Bern, Bern, Switzerland; [5]Institute of Medical Molecular Genetics, University of Zurich, Schlieren, Switzerland; [6]Center for Integrative Human Physiology (ZIHP), University of Zurich, Zurich, Switzerland; [7]Neuroscience Center (ZNZ), University and ETH Zurich, Zurich, Switzerland
*Correspondence: J Neubauer, Zurich Institute of Forensic Medicine, University of Zurich, Forensic Genetics, Winterthurerstrasse 190/52, Zurich 8057, Switzerland. Tel: +41 44 635 56 57; Fax: +41 44 635 68 58; E-mail: jacqueline.neubauer@irm.uzh.ch
Received 6 July 2016; revised 18 November 2016; accepted 14 December 2016; published online 11 January 2017

associated genes in 16 cases (34%), demonstrating the potential of performing a molecular autopsy in sudden death cases.[16] As whole-exome sequencing (WES) is not substantially more expensive than a targeted sequencing approach, this sequencing strategy allows extended data analysis in cases of a negative result within a predefined gene list. Therefore, the aim of this study was to perform a WES analysis in our large SIDS cohort of 161 infant cases with focus on 192 genes associated with cardiovascular and metabolic disorders.

MATERIALS AND METHODS

SIDS study population

Our study population consisted of 161 SIDS cases collected between 1985 and 2014 at the Zurich Institute of Forensic Medicine, Zurich, Switzerland. Most of the SIDS cases were examined by the same forensic pathologist, ensuring a high level of uniformity in autopsy procedures and case reporting. The classification of SIDS cases has always been performed according to the generally accepted international definitions of SIDS, including a complete autopsy, review of the circumstances of death, and examination of the clinical history.[1]

Forty-one infants were determined as genuine SIDS cases belonging to SIDS category I, including infants with normal clinical history, normal growth and development, no similar deaths among siblings and found in a safe sleeping environment with no evidence of accidental death. The remaining 120 infants were classified into SIDS category II due to slight infections before death, preterm birth or other deviations to category I requirements. The median age of the 161 SIDS infants was 15.03 ± 8.3 weeks (range 0.6–48.1 weeks) and 60.2% were boys (97 males/64 females). All of the SIDS infants were Europeans, most of them Swiss. A targeted MPS approach (HID-Ion AmpliSeq Ancestry Panel, Thermo Fisher, Rotkreuz, Switzerland) using the Ion Torrent PGM platform (Thermo Fisher) was applied to verify the geographical origin of SIDS cases where no information on the family origin was available. Eleven SIDS cases were excluded from our original cohort of 172 individuals, because of non-European ethnicities. Even small amounts of population admixture can shift the results toward an association, and therefore, it is important to have a well-defined study population in regard to ethnic and geographic background.[17] Additional epidemiological data of the SIDS cohort are illustrated in Supplementary Table 1.

Ethical approval for this study was provided by the local ethics committee (KEK-ZH-Nr. 2013–0086), and the study was conducted in full conformance with Swiss laws and regulations. Family members were not available for co-segregation analysis.

DNA extraction and quantification

Genomic DNA of the SIDS infants was obtained from tissues stored in alcohol or from alcohol-fixed and paraffin-embedded tissue blocks.[18] In most of the cases kidney or tongue was used (otherwise heart, muscle or brain) because of reported good post-mortem DNA stability in these tissues.[19] DNA extractions were performed using the QIAamp DNA Mini Kit (Qiagen, Hombrechtikon, Switzerland) according to the manufacturer's protocol. All DNA quantities were determined with a Qubit 1.0 fluorometric quantification device (Thermo Fisher).

Exome sequencing and bioinformatics

DNA library preparation and exome capture were performed with the SureSelect target enrichment and SureSelect All Exon V5+UTR's kits (Agilent Technologies AG, Basel, Switzerland), using the protocol for 200 ng input amount of genomic DNA. Sequencing as well as sequence alignment and variant calling were performed at the Functional Genomics Center Zurich, Switzerland. Sequencing was done on the Illumina HiSeq2500 platform (Illumina, San Diego, USA), generating 2×100 bp paired-end reads. Sequences were aligned to the reference genome (GRCh37/hg19) using BWA.[20] MEM algorithm with default setting and quality control of the exome coverage was performed with Bioconductor package TecQC.[21] A filter was set so that a sample was required to have at least 80% of exome covered at $\geq 20\times$ read depth. Variants discovery was performed by means of GATK,[22] following the GATK best practices workflow.[23] In particular, groups were reassigned using PicardTools,[24] duplicate reads were removed using Samtools[25] and local realignment, variants discovery and filtering (minimum $20\times$ coverage, minimum 20% alternate allele frequency) was done with GATK walker and the dbSNP database.[26]

Data analysis

Within our WES results, we focused on a gene panel of 192 genes associated with cardiovascular or metabolic disease-associated genes (Supplementary Table 2). Variants with $<50\times$ bidirectional coverage and/or an alternate allele frequency ratio <0.4 were additionally confirmed by standard Sanger sequencing methods. Annotation of the variants was performed with the Software Alamut Batch version 1.4.2 (Interactive Biosoftware, Rouen, France). Output results were reported in an Excel-sheet for data analysis. Variants in *TTN* were not further evaluated due to reported difficulties in sequencing and variant interpretation.[27] We adapted our previously published filter strategy[28] as follows: (1) a global minor allele frequency value (MAF) of ≤ 0.01 based on NCBI dbSNP,[26] (2) focus on exonic and splice site variants and (3) the exclusion of synonymous variants (Supplementary Figure 1). In addition, Human Gene Mutation Database (HGMD, Qiagen)[29] was consulted to check already reported variants in the literature. Alamut Visual Version 2.7.1 (Interactive Biosoftware) was used to visualize coverage of variants and to review conservation of the variants across a variety of species. Pathogenicity of variants was assigned according to an adapted scoring scheme originally described by van Spaendonck-Zwarts *et al.*[30] and Hertz *et al.*[31] Our scoring scheme was based on the assessment of variant types (null-variants, splice site variants, missense variants), *in silico* protein predictions, and MAF in three European control populations namely Exome Sequencing Project (ESP), 1000 Genomes Project and Exome Aggregation consortium (ExAC) database (Table 1).[26,32–39] A small proportion of the 60 706 individuals in the ExAC database is originally coming from smaller databases possibly leading to an overlap of the European individuals.[40] ESP is not well-powered to filter at 0.1% allele frequency without removing many genuinely rare variants, however the majority of ESP European singletons are not seen a second time in ExAC. Therefore, we still used the European MAF of all three databases. Based on the scoring scheme, variants were classified into five separate subcategories, designated as variants of unknown significance VUS0-4 (Table 2). Co-segregation and functional analyses would have been required to classify a variant as pathogenic;[30] therefore, the highest score a variant in our study could get was VUS4. DNA variants were numbered according to reference sequences using HGVS nomenclature (http://varnomen.hgvs.org). Variants in subcategories VUS4/VUS3 have been submitted to the Leiden Open Variation Database (Individual IDs: 00064759 to 00064769/00065124 to 00065157) (http://databases.lovd.nl/shared/diseases/02087).

Variant confirmation

Potential disease-causing variants not reported in the mentioned databases were confirmed by Sanger sequencing. Additionally, allele frequencies were checked in an in-house exome database of 118 European patients with eye diseases.

RESULTS

Whole-exome sequencing and data analysis was successfully completed for 155 SIDS cases. DNA library preparation failed for six SIDS samples due to low DNA quantities or poor DNA qualities resulting from decayed post-mortem tissues and/or the fixation method.

Overall, 72.3% of the bases had a coverage of ≥ 20 reads and the average on-target coverage was 90.2% at ≥ 20 reads. The average depth within our gene panel was 100.64 ± 34.25 (Supplementary Table 3). By focusing on 192 genes of interest, an average of 1960 ± 467 variants per case were obtained for further data analysis. After the filtering steps, an average of 14.6 ± 9.8 variants per sample were manually checked with the Alamut Visual v2.7 software and evaluated according to our scoring scheme. Sanger sequencing confirmed all VUS4/VUS3 variants with $<50\times$ coverage and/or not reported in the databases (Supplementary Table 4). The majority of these variants were missense

Table 1 Scoring scheme for variant evaluation (adapted from Hertz *et al.*[31])

Parameter	*Score*
Coding effect	
Nonsense	24
Frameshift	24
Splice sites	
±1/±5	24
±2/±3	12
All others	0
Missense	0
In silico protein prediction	
AGVGD	
C65	4
C55	3.2
C35	2.4
C25	1.6
C15	0.8
C0	0
NA	0
Grantham distance	
>140	4
70–140	2
<70	0
NA	0
SIFT	
Deleterious	4
Good	0
NA	0
MAPP	
Bad	4
Good	0
NA	0
Polyphen2	
Probably damaging	4
Possibly damaging	2
Benign	0
MutationTaster	
Disease causing	4
Polymorphism	0
NA	0
MAF	
ESP EAMAF	
≤0.001	8
0.001<freq≤0.002	5.3
0.002<freq≤0.01	2.6
>0.01	0
NA	0
ExAC EURNFMAF	
≤0.001	8
0.001<freq≤0.002	5.3
0.002<freq≤0.01	2.6
>0.01	0
NA	0
1000 Genomes Project EURMAF	
≤0.001	8
0.001<freq≤0.002	5.3
0.002<freq≤0.01	2.6
>0.01	0
NA	0

Abbreviations: AGVGD, align Grantham variation and Grantham deviation[32]; ESP EAMAF, MAF in European American population in NHLBI GO Exome Sequencing Project[37]; ExAC EURNFMAF, MAF in European (non-Finnish) population in exome aggregation consortium[39]; MAF, minor allele frequency; MAPP, multivariate analysis of protein polymorphism prediction[34]; NA, not available; Polyphen2, polymorphism phenotyping v2 (ref. 35); SIFT, sorting intolerant from tolerant prediction[33]; MutationTaster[36]; 1000 Genomes Project EURMAF, MAF in European population in 1000 Genomes Project.[38]

Table 2 Subcategories of variants based on scoring scheme

Percentage (%)	*Score*	*Subclass*	
≥90	43.2–48.0	VUS4	Probably pathogenic
80–89	38.4–43.1	VUS3	Likely pathogenic
70–79	33.6–38.3	VUS2	Unclear
60–69	28.8–33.5	VUS1	Unlikely pathogenic
<59	≤28.7	VUS0	Not pathogenic (neutral variant or weak modifier)

Abbreviation: VUS, variant of unknown significance.

variants (97.1%) followed by splice site variants (1.5%), nonsense variants (0.9%), and frameshift variants (0.6%). Based on our scoring scheme, 11 variants (1.3%) were classified into sub-category VUS4 and 24 variants (2.7%) into VUS3 indicating variants with probably pathogenic effects.

Thirty-one (20%) out of the 155 SIDS cases had at least one variant with likely pathogenic functional effects (Table 3). Four of these cases had two likely causative variants. Details of the variants are available in Supplementary Table 4. Of the 31 SIDS cases, 17 (54.8%) were males and the median age of death was 4 months (range 1–9 months). Most of the variants were detected in SIDS category II infant cases (67.7%).

Among the 31 SIDS infants with likely causative variants, 14 infants (9%) carried putative pathogenic variants in genes associated with ion channelopathies and 11 SIDS infants (7%) had disease-causing variants in genes associated with cardiomyopathies (Figure 1). Additional four SIDS infants (2.5%) had variants in genes associated with mitral valve prolapse, aortic valve disease, Marfan syndrome or Ehlers-Danlos syndrome. Variants in genes associated with metabolic diseases were found in two SIDS cases (1%), in glycogen storage disease and systemic primary carnitine deficiency.

Most of the variants were detected in genes associated with BrS (2.5%), followed by dilated cardiomyopathy (2.1%), hypertrophic cardiomyopathy (1.4%), LQTS (1.4%), ARVC (1.0%) and CPTV (1.0%).

DISCUSSION

High-throughput sequencing provides a comprehensive and time-efficient sequencing strategy to identify rare DNA sequence variants in the genome/exome of patients with complex disorders or to discover underlying genetic causes in large heterogeneous study populations as for example in SIDS cases.

Starting with tissue collection of SIDS infants already in the early 1980s, we have a valuable and well-defined SIDS cohort of 161 infant cases at our institute. Although alcohol-fixed and paraffin-embedded tissue blocks do not provide optimal DNA qualities and quantities, exome sequencing was successfully completed for 155 out of 161 SIDS cases. By focusing on a gene list of 192 genes associated with cardiovascular or metabolic diseases, we identified potentially disease-causing variants in 20% of the 155 SIDS cases. The majority of these cases had variants in genes associated with channelopathies (9%) and cardiomyopathies (7%).

The main cardiac genes reported in other SIDS studies are *CAV3*, *GJA1*, *GPD1-L*, *KCNE2*, *KCNJ8*, *KCNQ1*, *KCNH2*, *MYBPC3*, *RYR2*, *SCN5A* and *TNNI3*.[6] We detected potentially causative variants in two of our SIDS cases in *SCN5A* p.(Arg1897Trp) and *RYR2* (c.2907-1G>C). *SCN5A* is primarily expressed in the cell membrane of cardiac tissue where it encodes sodium-gated channels. Although *SCN5A* p.(Arg1897Trp) has been reported in patients with LQTS and

Table 3 Variants with likely disease-causing effects in the SIDS cohort

Case	Sex	Age (mo)	rs-Nr.	Gene	HGVS genomic RefSeq-Nr.	HGVS RefSeq-Nr.	Coding Effect	c.DNA	Protein change	dbSNP ALLMAF	ESP EAMAF	ExAC EURNFMAF	1000 Genomes EURMAF	in-house database[a]	AGVGD	SIFT	MAPP	MutationTaster	Grantham distance	Polyphen2	Disease association	Subclass based on scoring scheme
SIDS001	F	9	rs17848485	*CPT2*	NG_008035.1	NM_000098.2	missense	c.1634A>C	p.(Glu545Ala)	0.0010	0.0000	0.0000	0.0030	NA	C65	deleterious	bad	disease-causing	107	probably damaging	SPCD	VUS3
SIDS004	M	7	rs199583774	*PKP2*	NG_009000.1	NM_004572.3	missense	c.2083C>T	p.(Arg695Cys)	0.0000	0.0000	0.0001	0.0010	NA	C25	deleterious	bad	disease-causing	180	probably damaging	ARVC	VUS4
SIDS005	M	4	rs142890619	*COL5A1*	NG_008030.1	NM_000093.4	missense	c.598G>A	p.(Asp200Asn)	0.0000	0.0010	0.0000	0.0010	NA	C0	deleterious	bad	disease-causing	23	probably damaging	EDS	VUS3
SIDS007	F	8	rs143641133	*DOLK*	NG_017009.1	NM_014908.3	missense	c.1324G>A	p.(Ala442Thr)	0.0000	0.0000	0.0000	0.0010	NA	C0	deleterious	bad	disease-causing	58	probably damaging	DCM	VUS3
SIDS010	F	3	rs200473206	*DSP*	NG_008803.1	NM_004415.3	splice site	c.273+5G>A	p.(?)	0.0000	0.0005	0.0005	NA	NA	NA	NA	NA	NA	NA	NA	ARVC	VUS3
SIDS011	M	2	rs71352737	*TRPM4*	NG_027551.1	NM_017636.3	nonsense	c.1575G>A	p.(Trp525*)	0.0010	0.0011	0.0018	0.0010	NA	NA	NA	NA	disease-causing	NA	NA	BrS	VUS4
SIDS015	M	1	rs143788120	*MRPL3*	NG_029207.1	NM_007208.3	missense	c.862T>C	p.(Ser288Pro)	0.0000	0.0000	0.0006	0.0030	NA	C65	deleterious	bad	disease-causing	74	probably damaging	HCM	VUS3
SIDS021	M	4	~	*TSFM*	NG_016971.1	NM_001172696.1	frameshift	c.971_972del	p.(Gln324Argfs*11)	NA	0.0000	NA	NA	0.0000	NA	NA	NA	NA	NA	NA	DCM	VUS3
SIDS024	M	2	rs140470576	*AKAP9*	NG_011623.1	NM_005751.4	missense	c.10118C>A	p.(Ser3373Tyr)	0.0000	0.0020	0.0000	0.0000	NA	C65	deleterious	bad	disease-causing	144	probably damaging	LQTS	VUS4
SIDS031	F	4	rs141131535	*MYLK*	NG_029111.1	NM_053025.3	missense	c.344G>A	p.(Arg115His)	0.0000	0.0002	0.0001	0.0010	NA	C25	deleterious	bad	disease-causing	29	possibly damaging	HCM	VUS3
SIDS037	F	4	rs139965373	*AKAP9*	NG_011623.1	NM_005751.4	missense	c.3580G>A	p.(Ala1194Thr)	0.0010	0.0000	0.0000	0.0000	NA	C55	deleterious	bad	disease-causing	58	probably damaging	LQTS	VUS4
SIDS041	F	4	rs185947256	*AGL*	NG_012865.1	NM_000028.2	missense	c.3290G>A	p.(Arg1097His)	0.0000	0.0000	0.0000	0.0000	NA	C0	deleterious	bad	disease-causing	29	probably damaging	GSD	VUS3
SIDS045	M	5	rs369588692	*SCN1B*	NG_013359.1	NM_199037.3	missense	c.673C>T	p.(Arg225Cys)	0.0000	0.0000	0.0001	0.0000	NA	C0	deleterious	bad	disease-causing	142	probably damaging	BrS	VUS4
SIDS050	M	2	rs199732064	*TRPM7*	NG_021363.1	NM_017672.4	missense	c.4084C>T	p.(Pro1362Ser)	0.0000	0.0009	0.0005	0.0010	NA	C0	deleterious	bad	disease-causing	74	possibly damaging	DCM	VUS3
			~	*ACTN2*	NG_009081.1	NM_001103.2	nonsense	c.2671G>T	p.(Glu891*)	NA	0.0000	NA	NA	0.0000	NA	NA	NA	disease-causing	NA	NA	DCM	VUS4
SIDS061	M	4	rs142217269	*SCN10A*	NG_031891.2	NM_006514.3	missense	c.4568G>A	p.(Cys1523Tyr)	0.0010	0.0027	0.0018	0.0010	NA	C65	deleterious	bad	disease-causing	194	probably damaging	BrS	VUS3
			rs143767864	*DCHS1*	NG_033858.1	NM_003737.3	missense	c.704G>A	p.(Arg235Gln)	0.0000	0.0000	0.0000	0.0010	NA	C0	deleterious	bad	disease-causing	43	probably damaging	MVP	VUS3
SIDS069	M	3	rs200893930	*NOTCH1*	NG_007458.1	NM_017617.4	missense	c.7645C>T	p.(Arg2549Cys)	0.0000	0.0001	0.0000	0.0000	NA	C25	deleterious	bad	disease-causing	180	possibly damaging	AOVD	VUS4
SIDS073	M	3	rs200473206	*DSP*	NG_008803.1	NM_004415.2	splice site	c.273+5G>A	p.(?)	0.0000	0.0005	0.0005	NA	NA	NA	NA	NA	NA	NA	NA	ARVC	VUS3
SIDS075	F	3	rs121434557	*GJA5*	NG_009369.2	NM_005266.6	missense	c.286G>T	p.(Ala96Ser)	0.0000	0.0000	0.0001	0.0010	NA	C65	deleterious	NA	disease-causing	99	probably damaging	Afib	VUS3
			rs149272010	*SYNE1*	NG_012855.1	NM_182961.3	missense	c.19534G>T	p.(Ala6512Thr)	NA	0.0000	0.0000	0.0000	NA	C55	deleterious	bad	NA	58	probably damaging	EDMD	VUS3
SIDS087	M	4	~	*MYLK*	NG_029111.1	NM_053025.3	missense	c.1490C>T	p.(Ser497Phe)	NA	0.0000	NA	NA	0.0000	C65	deleterious	bad	disease-causing	155	probably damaging	HCM	VUS3
SIDS089	F	5	rs199800922	*MYH11*	NG_009299	NM_001040114.1	missense	c.3624C>G	p.(His1208Gln)	0.0000	0.0001	0.0000	0.0010	NA	C15	deleterious	bad	disease-causing	24	possibly damaging	HCM	VUS3
SIDS091	M	2	rs200899037	*CASQ2*	NG_008802.1	NM_001232.3	missense	c.1052A>G	p.(Asp351Gly)	0.0000	0.0000	0.0000	0.0010	NA	C0	deleterious	bad	disease-causing	94	probably damaging	CPVT	VUS3
SIDS096	F	2	rs71352737	*TRPM4*	NG_027551.1	NM_017636.3	nonsense	c.1575G>A	p.(Trp525*)	0.0010	0.0011	0.0018	0.0010	NA	NA	NA	NA	disease-causing	NA	NA	BrS	VUS4
SIDS100	F	4	rs373347787	*SCN10A*	NG_031891.2	NM_006514.3	missense	c.4780C>T	p.(Arg1594Cys)	0.0000	0.0000	0.0000	NA	NA	C65	deleterious	bad	disease-causing	180	probably damaging	BrS	VUS3
SIDS101	M	6	rs200899037	*CASQ2*	NG_008802.1	NM_001232.3	missense	c.1052A>G	p.(Asp351Gly)	0.0000	0.0000	0.0000	0.0010	NA	C0	deleterious	bad	disease-causing	94	probably damaging	CPVT	VUS3
SIDS119	F	1	~	*TRPM4*	NG_027551.1	NM_017636.3	frameshift	c.247dup	p.(Ala83Glyfs*13)	NA	0.0000	NA	NA	0.0000	NA	NA	NA	NA	NA	NA	BrS	VUS3
			rs140148322	*FHL2*	NG_008844.2	NM_001450.3	missense	c.337C>T	p.(Arg113Cys)	0.0000	0.0005	0.0005	NA	NA	C65	deleterious	bad	disease-causing	180	probably damaging	DCM	VUS3
SIDS142	M	3	rs45465995	*SCN5A*	NG_008934.1	NM_001099404.1	missense	c.5689C>T	p.(Arg1897Trp)	0.0020	0.0004	0.0000	0.0020	NA	C65	deleterious	bad	disease-causing	101	probably damaging	BrS	VUS4
SIDS156	F	2	rs139170018	*SYNE1*	NG_012855.1	NM_182961.3	missense	c.15337G>A	p.(Val5113Ile)	0.0000	0.0012	0.0010	0.0000	NA	C25	deleterious	bad	disease-causing	29	probably damaging	EDMD	VUS3
SIDS161	M	1	rs377061868	*TPM1*	NG_007557.1	NM_000366.5	splice site	c.375-5T>C	p.(?)	0.0000	0.0005	0.0004	NA	NA	NA	NA	NA	NA	NA	NA	DCM	VUS3
SIDS172	M	8	rs111843122	*RYR2*	NG_008799.2	NM_001035.2	splice site	c.2907-1G>C	p.(?)	0.0000	0.0000	0.0000	0.0000	NA	NA	NA	NA	NA	NA	NA	CPVT	VUS3
SIDS194	F	5	rs72544141	*ANK2*	NG_009006.2	NM_001148.4	missense	c.4373A>G	p.(Glu1458Gly)	0.0010	0.0002	0.0006	0.0010	NA	C65	deleterious	bad	disease-causing	98	possibly damaging	LQTS	VUS4
SIDS223	F	2	rs148323523	*DCHS1*	NG_033858.1	NM_003737.3	missense	c.6467T>A	p.(Val2156Glu)	0.0000	0.0006	0.0002	0.0000	NA	C65	deleterious	bad	disease-causing	121	possibly damaging	MVP	VUS4

Abbreviations: Afib, atrial fibrillation; AGVGD, align Grantham variation and Grantham deviation; ALLMAF, minor allele frequency in all populations based on NCBI dbSNP; AOVD, aortic valve disease; ARVC, arrhythmogenic right ventricular cardiomyopathy; BrS, Brugada syndrome; CPVT, catecholaminergic polymorphic ventricular tachycardia; DCM, dilated cardiomyopathy; EDMD, Emery-dreifuss muscular dystrophy; EDS, Ehlers-Danlos syndrome; ESP EAMAF, MAF in European American population in NHLBI GO Exome Sequencing Project; ExAC EURNFMAF, MAF in European (non-Finnish) population in exome aggregation consortium; F, female; GSD, glycogen storage disease; HCM, hypertrophic cardiomyopathy; LQTS, long QT syndrome; M, male; MAF, minor allele frequency; MAPP, multivariate analysis of protein polymorphism prediction; MVP, mitral valve prolapse; MS, marfan syndrome; NA, not available; SIFT, sorting intolerant from tolerant prediction; SPCD, systemic primary carnitine deficiency; 1000 Genomes Project EURMAF, MAF in European population in 1000 Genomes Project.

~ No dbSNP rs-no. available.

[a]Based on 118 in-house control exoms.

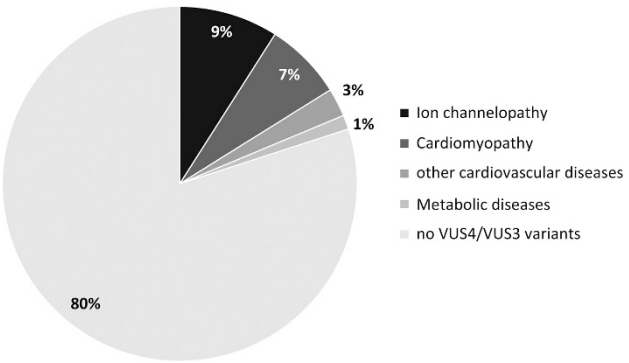

Figure 1 Percentage of SIDS infants with likely causative variants in genes associated with cardiomyopathies, ion channelopathies, other cardiovascular diseases, and metabolic diseases. VUS, variant of unknown significance.

atrial fibrillation, functional studies have indicated no effect on QTc intervals, syncope propensity, and overall mortality assuming that this variant is less likely associated with a dominant monogenic form of the disease.[41] *RYR2* encodes a ryanodine receptor found in cardiac sarcoplasmic reticulum and causative variants are described in stress-induced CPTV and ARVC.[42] *RYR2* (c.2907-1G>C) is located at the 5′-end of exon 26 and causes an altered acceptor site, however the variant was not described in ARVC or CPTV patients so far. Further variants with less likely functional effects were detected in *KCNE2*, *CAV3*, *RYR2* and *MYBPC3* (Supplementary Table 4), but no variants in *GJA1*, *GPD1-L*, *KCNH2*, *KCNJ8* and *TNNI3*.

Additional variants with likely functional effects were found in genes related to different cardiac diseases or sudden death such as *ANK2* p.(Glu148Gly), *ACTN2* p.(Glu891*), *DSP* (c.273+5G>A), *TRPM4* p.(Ala83Glfs*13), *TRPM4* p.(Trp525*), and *TSFM* p.(Gln324Argfs*11). Interestingly, *ANK2* p.(Glu148Gly) and *TRMP4* p.(Trp525*) were already described in a Danish SIDS cohort.[16] *ANK2* encodes ankyrin-B, which has an essential role in the localization and membrane stabilization of ion transporters and ion channels in cardiomyocytes.[43] The same variant was first reported in a large French family with LQTS including sinus node dysfunction and episodes of atrial fibrillation and one individual who suffered sudden death,[44] but *ANK2* p.(Glu148Gly) was also detected in eight Danish control individuals with normal mean QTc interval.[41] *TRPM4* belongs to the melastatin-related transient receptor channel family and encodes calcium-permeable cation channels localized predominantly in the plasma membrane.[45] Variants in *TRPM4* were described in patients with progressive familial heart block and BrS. Both alterations represent interesting candidate variants involved in the sudden death event of SIDS cases.

The most investigated gene with regard to metabolic diseases in SIDS is *ACADM*, which catalyzes the first step in the beta-oxidation of fatty acids.[13] The most prevalent variation causing MCAD deficiency is *ACADM* p.(Lys329Glu), which is present in 80% of individuals who clinically are diagnosed with MCAD.[7] The only variant detected in our SIDS cohort with regard to *ACADM* was *ACADM* p.(Arg53Cys) (Supplementary Table 4). Although this variant has been reported in one MCAD-patient in combination with the most common *ACADM* p.(Lys329Glu) pathogenic variant,[46] our scoring scheme predicts little functional effect for this variant.

Altered ion channel functions causing lethal arrhythmias may represent the most plausible and comprehensible cause in infant death cases.[10] Many channelopathies are characterized by incomplete penetrance and variable expressivity where sudden cardiac death is often the first manifestation of the disease.[47] In contrast, cardiomyopathies are mainly caused by variants in genes encoding desmosomal cell adhesion proteins or in sarcomeric proteins involved in heart contraction inducing structural heart abnormalities. However, a growing number of studies have established links between desmosomes and components of cardiac electrical machinery.[48] Consequently, variants in cardiomyopathy-associated genes may contribute more generally to cardiac diseases and might be involved in the cause of death in some of the SIDS infants even in absence of morphological abnormalities in the heart.

Today, exome sequencing is not substantially more expensive compared to targeted gene panels, but represents a more efficient and comprehensive sequencing method to investigate sudden unexplained death cases in absence of a specific phenotype. To our knowledge, this is the first WES study in a large SIDS cohort. Although we exclusively report the findings within predefined genes of interest, one major advantage of exome sequencing is the alternative of extended data analysis in cases without any results providing a chance to identify new candidate genes in SIDS. The underlying cause of death in the majority of SIDS cases still remains elusive and might be explained by a multifactorial etiology due to a combination of different genetic and environmental risk factors. Therefore, further analyses could focus on SIDS-related predisposing genetic factors in genes involved in early brain development, respiratory regulation, nicotine response, immune system, metabolic and energy production, thermoregulation and mitochondrial activity.[6,7]

The main current challenge in exome sequencing studies is the clinical interpretation of genetic variants identified. The categorization of variants in our study was based on a stringent scoring scheme involving different population-specific databases and *in silico* protein prediction tools. Nevertheless, the different VUS categories included single variants with a higher allele frequency than expected for specific cardiovascular diseases meaning that a part of the here reported VUS might not be severe enough to cause death in infancy; still some of them could act as predisposing risk factors whereas others might be reclassified as benign based on prospective findings. Therefore, further assessments such as functional studies are required and strongly recommended for an evidence-based classification of the pathogenicity.[49] Recently, mutations in genes previously associated with SIDS were identified in exome data from population studies indicating that many variants might have some pathological influence, but are most likely not the exclusive genetic cause of SIDS.[50] Therefore, caution is needed when translating such exome sequencing results from research to diagnostic applications. Genetic counseling of first-degree relatives should be based on a multidisciplinary approach, involving forensic pathologists, geneticists and cardiologists, to inform the family in case of positive genetic findings and to discuss further steps regarding genetic testing of family members and/or to monitor the affected person.[2]

A limitation of this study is the lack of functional assays in order to verify the potentially pathogenic role of detected variants, in particular amino acid substitutions. Also, family members were not available for co-segregation analyses due to the sample anonymity required by the ethical committees. This would be necessary to determine the mode of inheritance, to classify variants into the pathogenic category,[30] and to identify other genetic carriers at risk for sudden cardiac death. Our case reports only included clinical records on sudden death cases in siblings but not in other family members, which would be an important point to consider. Finally, exome sequencing data reveal lower sequencing coverage compared to targeted gene panels

potentially leading to a loss of important low-coverage variants and more false negative/positive calls.

Additional MPS studies combined with functional assessment in large SIDS cohorts are inevitable to better understand the etiology of SIDS and to identify additional pathophysiologic mechanisms involved in this tragic death event.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This project was supported by the Swiss National Science Foundation (SNF, project-No. 320030-149456). Special thanks to Corinne Moser for excellent technical assistance, to Mario Gysi for the ancestry panel sequencing on the Ion Torrent PGM platform, and to Amit Tiwari and Samuel Koller for access to the in-house exome database at the Institute of Medical Molecular Genetics (Switzerland).

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**: 234–238.
2 Wilhelm M, Bolliger SA, Bartsch C *et al*: Sudden cardiac death in forensic medicine - Swiss recommendations for a multidisciplinary approach. *Swiss Med Wkly* 2015; **145**: 1–6.
3 Hunt CE, Hauck FR: Sudden infant death syndrome. *CMAJ* 2006; **174**: 1861–1869.
4 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**: 194–197.
5 Kinney HC, Thach BT: The sudden infant death syndrome. *N Engl J Med* 2009; **361**: 795–805.
6 Courts C, Madea B: Genetics of the sudden infant death syndrome. *Forensic Sci Int* 2010; **203**: 25–33.
7 Opdal SH, Rognum TO: Gene variants predisposing to SIDS: current knowledge. *Forensic Sci Med Pathol* 2011; **7**: 26–36.
8 Ackerman MJ, Priori SG, Willems S *et al*: HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies: this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Heart Rhythm* 2011; **8**: 1308–1339.
9 Brion M, Allegue C, Santori M *et al*: Sarcomeric gene mutations in sudden infant death syndrome (SIDS). *Forensic Sci Int* 2012; **219**: 278–281.
10 Campuzano O, Allegue C, Sarquella-Brugada G *et al*: The role of clinical, genetic and segregation evaluation in sudden infant death. *Forensic Sci Int* 2014; **242**: 9–15.
11 Abriel H, Zaklyazminskaya EV: Cardiac channelopathies: genetic and molecular mechanisms. *Gene* 2013; **517**: 1–11.
12 Maron BJ, Towbin JA, Thiene G *et al*: Contemporary definitions and classification of the cardiomyopathies: an American Heart Association Scientific Statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention. *Circulation* 2006; **113**: 1807–1816.
13 Pryce JW, Weber MA, Heales S, Malone M, Sebire NJ: Tandem mass spectrometry findings at autopsy for detection of metabolic disease in infant deaths: postmortem changes and confounding factors. *J Clin Pathol* 2011; **64**: 1005–1009.
14 Loporcaro CG, Tester DJ, Maleszewski JJ, Kruisselbrink T, Ackerman MJ: Confirmation of cause and manner of death via a comprehensive cardiac autopsy including whole exome next-generation sequencing. *Arch Pathol Lab Med* 2013; **138**: 1083–1089.
15 Santori M, Blanco-Verea A, Gil R *et al*: Broad-based molecular autopsy: a potential tool to investigate the involvement of subtle cardiac conditions in sudden unexpected death in infancy and early childhood. *Arch Dis Child* 2015; **100**: 952–956.
16 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* 2015; **24**: 817–822.
17 Marchini J, Cardon LR, Phillips MS, Donnelly P: The effects of human population structure on large genetic association studies. *Nat Genet* 2004; **36**: 512–517.
18 Pikor LA, Enfield KS, Cameron H, Lam WL: DNA extraction from paraffin embedded material for genetic and epigenetic analyses. *J Vis Exp* 2011; **26**: 2763.
19 Bär W, Kratzer A, Mächler M, Schmid W: Postmortem stability of DNA. *Forensic Sci Int* 1988; **39**: 59–70.
20 Li H, Durbin R: Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010; **26**: 589–595.
21 Bioconductor package TecQC 2015, Available at https://www.bioconductor.org/packages/3.3/bioc/vignettes/TEQC/inst/doc/TEQC.pdf.
22 McKenna A, Hanna M, Banks E *et al*: The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; **20**: 1297–1303.
23 DePristo MA, Banks E, Poplin R *et al*: A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011; **43**: 491–498.
24 Picard tool by Broad Institute 2015, Available at http://broadinstitute.github.io/picard/.
25 Li H, Handsaker B, Wysoker A *et al*: The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
26 NCBI National Center for Biotechnology Information 2016, Available at http://www.ncbi.nlm.nih.gov/.
27 Lopes LR, Zekavati A, Syrris P *et al*: Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. *J Med Genet* 2013; **50**: 228–239.
28 Neubauer J, Haas C, Bartsch C, Medeiros-Domingo A, Berger W: Post-mortem whole-exome sequencing (WES) with a focus on cardiac disease-associated genes in five young sudden unexplained death (SUD) cases. *Int J Legal Med* 2016; **130**: 1011–1021.
29 Stenson PD, Mort M, Ball EV, Shaw K, Philips A, Cooper DN: The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet* 2014; **133**: 1–9.
30 van Spaendonck-Zwarts KY, van Rijsingen IA, van den Berg MP *et al*: Genetic analysis in 418 index patients with idiopathic dilated cardiomyopathy: overview of 10 years' experience. *Eur J Heart Fail* 2013; **15**: 628–636.
31 Hertz CL, Christiansen SL, Ferrero-Miliani L *et al*: Next-generation sequencing of 34 genes in sudden unexplained death victims in forensics and in patients with channelopathic cardiac diseases. *Int J Legal Med* 2015; **129**: 793–800.
32 Align GVGD 2016, Available at http://agvgd.iarc.fr/.
33 SIFT 2016, Available at http://sift.jcvi.org/.
34 MAPP (multivariate analysis of protein polymporthims) 2016, Available at http://www.ngrl.org.uk/Manchester/page/mapp-multivariate-analysis-protein-polymorphism.
35 PolyPhen-2 prediction of functional effects of human 2016, Available at http://genetics.bwh.harvard.edu/pph2/.
36 Mutationtaster 2016, Available at http://www.mutationtaster.org/.
37 NHLBI Exome Sequencing Project (ESP) 2016, Available at http://evs.gs.washington.edu/EVS/.
38 1000 Genomes Project Data 2016, Available at http://browser.1000genomes.org/Homo_sapiens/Info/Index.
39 Exome Aggregation Consortium (ExAC) 2016, Available at http://exac.broadinstitute.org/about.
40 Lek M, Karczewski KJ, Minikel EV *et al*: Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016; **536**: 285–291.
41 Ghouse J, Have CT, Weeke P *et al*: Rare genetic variants previously associated with congenital forms of long QT syndrome have little or no effect on the QT interval. *Eur Heart J* 2015; **36**: 2523–2529.
42 Tiso N, Stephan DA, Nava A *et al*: Identification of mutations in the cardiac ryanodine receptor gene in families affected with arrhythomogenic right ventricular cardiomyopathy type 2 (ARVD2). *Hum Mol Genet* 2001; **10**: 189–194.
43 Wu HC, Yamankurt G, Luo J *et al*: Identification and characterization of two ankyrin-B isoforms in mammalian heart. *Cardiovasc Res* 2015; **107**: 466–477.
44 Niven JE, Vahasoyrinki M, Kauranen M, Hardie RC, Juusola M, Weckstrom M: The contribution of Shaker K+ channels to the information capacity of Drosophila photoreceptors. *Nature* 2003; **421**: 630–634.
45 Farooqi AA, Javeed MK, Javed Z *et al*: TRPM channels: same ballpark, different players, and different rules in immunogenetics. *Immunogenetics* 2011; **63**: 773–787.
46 Derks TG, Touw CM, Ribas GS *et al*: Experimental evidence for protein oxidative damage and altered antioxidant defense in patients with medium-chain acyl-CoA dehydrogenase deficiency. *J Inherit Metab Dis* 2014; **37**: 783–789.
47 Campuzano O, Beltran-Alvarez P, Iglesias A, Scornik F, Perez G, Brugada R: Genetics and cardiac channelopathies. *Genet Med* 2010; **12**: 260–267.
48 Patel DM, Green KJ: Desmosomes in the heart: a review of clinical and mechanistic analyses. *Cell Commun Adhes* 2014; **21**: 109–128.
49 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**: 405–424.
50 Andreasen C, Refsgaard L, Nielsen JB *et al*: Mutations in genes encoding cardiac ion channels previously associated with sudden infant death syndrome (SIDS) are present with high frequency in new exome data. *Can J Cardiol* 2013; **29**: 1104–1109.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)