The Role of Clinical, Genetic and Segregation Evaluation in Sudden Infant Death

Authors
Oscar Campuzano PhD\textsuperscript{1}, Catarina Allegue PhD\textsuperscript{1}, Georgia Sarquella-Brugada MD\textsuperscript{2}, Monica Coll MSc\textsuperscript{1}, Jesus Mates MSc\textsuperscript{1}, Mireia Alcalde MSc\textsuperscript{1}, Carles Ferrer-Costa, PhD\textsuperscript{3}, Anna Iglesias MSc\textsuperscript{1}, Josep Brugada MD, PhD\textsuperscript{2}, Ramon Brugada MD, PhD\textsuperscript{1}

Affiliations
\textsuperscript{1} Cardiovascular Genetics Center, University of Girona-IdIBGi, Girona (Spain)
\textsuperscript{2} Arrhythmias Unit, Hospital Sant Joan de Déu, University of Barcelona, Barcelona (Spain)
\textsuperscript{3} Gendiag SL, Barcelona (Spain)

Corresponding author
Oscar Campuzano, BSc, PhD
Cardiovascular Genetics Center / School of Medicine
Parc Cientific i Tecnologic, Universitat de Girona-IDIBGI
17003 Girona (Spain) oscar@brugada.org

Highlights
Abstract

Sudden infant death syndrome (SIDS) is the leading cause of death in the first year of life. Several arrhythmogenic genes have been associated with cardiac pathologies leading to infant sudden cardiac death (SCD). Our aim was to take advantage of Next Generation Sequencing (NGS) technology to perform a thorough genetic analysis of a SIDS case.

A SIDS case was referred to our institution after negative autopsy. We performed a genetic analysis of 104 SCD-related genes using a custom panel. Confirmed variants in index case were also analyzed in relatives. Clinical evaluation of first-degree family members was performed.

Relatives did not show pathology. NGS identified seven variants. Two previously described as pathogenic. Four previously catalogued without clinical significance. The seventh variation was novel. Familial segregation showed that the index case’s mother carried all same genetic variations except one, which was inherited from the father. The sister of the index case carried three variants.

We believe that molecular autopsy should be included in current forensic protocols after negative autopsy. In addition to NGS technologies, familial genetic testing should be also performed to clarify potential pathogenic role of new variants and to identify genetic carriers at risk of SCD.

Key Words

SIDS, molecular autopsy, next generation sequencing, custom panel, genetics
1.- Introduction

Sudden infant death syndrome (SIDS) is defined as the death of an apparently healthy infant of less than one year of age. The death usually occurs during sleep and remains unexplained after an exhaustive investigation including complete autopsy and medical history [1]. Despite SIDS rates differ significantly across countries, ethnic groups and gender [2], SIDS is the main cause of death in Europe and North-America in infants less than one year of age [3]. A large number of pathophysiological mechanisms have been suggested but the etiology of SIDS still remains to be clarified. SIDS is considered a multifactorial disorder, with several intrinsic and extrinsic risk factors resulting in or predisposing to the death of the infant. Among them, genetic defects associated with inheritable arrhythmias play a role in this entity [4]. To date, 10%-15% of the SIDS cases are thought to be caused by cardiac channelopathies [5].

The ‘molecular autopsy’ enables genetic analysis to identify the defect that might be associated with a certain disease [6, 7]. Until now, few molecular autopsy series have been reported [8]. In addition, these studies have included the analyses of only the major long QT syndrome (LQTS) and catecholaminergic polymorphic ventricular tachycardia (CPVT) genes using a candidate gene approach [4, 9]. Financial limitations have impaired the use of this technology beyond the research area. The advent of massive parallel DNA sequencing platforms, named next-generation sequencing (NGS) technology, has revolutionized the field of medical genomics, allowing fast and cost-effective generation of genetic data [10]. The massive genetic screening has yet to fully enter the clinical field, hampered by the excess of generated genetic data, and specially the clinical phenotype interpretation.

The purpose of this study was to identify the genetic defect that could explain the cause of death in a SIDS case. Due to the amount of genes associated with lethal arrhythmogenic syndromes, we used a NGS custom panel. To our knowledge, no genetic report has been performed in SIDS cases using custom panel technology.

2.- Methods
A complete autopsy examination was performed according to current international regulations for unexpected death [11-13]. All relatives included in our study were clinically evaluated at Hospital Josep Trueta of Girona (Girona, Spain), and Hospital Sant Joan de Déu (Barcelona, Spain). Complete clinical evaluation, including electrocardiogram (ECG) and echocardiogram (ECHO), was performed in index case’s parents and sister. The study was approved by the ethics committee of the Hospital Josep Trueta (Girona, Spain), followed the Helsinki II declaration and informed consent was obtained from all participants. All patients were Caucasian and native of Spain.

DNA sample
Genomic DNA was extracted with Chemagic MSM I from whole blood (Chemagic human blood). DNA samples were checked in order to assure quality and quantify before processing to get the 3µg needed for the NGS strategy. DNA integrity was assessed on a 0,8% agarose gel. Spectrophotometric measurements are also performed to assess quality ratios of absorbance; dsDNA concentration is determined by fluorometry (PicoGreen assay). DNA sample was fragmented by Adaptive Focused Acoustics (Covaris). Library preparation was performed according to the manufacturer’s instructions (SureSelect XT Custom 0.5-2.9Mb library, Agilent Technologies, Inc). Indexed libraries enter finally in the sequencing path; pooled captures (up to 13 samples per lane) were sequenced on Illumina HiSeq2000 instrument using 2x75 bp reads length.

Custom Resequencing panel
We selected the most prevalent 104 genes involved in SCD-related pathologies, accordingly to available scientific literature focus on SCD, so far. The genomic coordinates corresponding to these 104 genes (Table 1) were designed by Ferrer inCode using the tool eArray (Agilent Technologies, Inc.). All the isoforms described at the University of California, Santa Cruz (UCSC) browser were included at the design. The biotinylated cRNA probe solution was manufactured for Ferrer inCode.
by Agilent Technologies and provided as capture probes. The final size was 680 Kbp of encoding regions and UTR boundaries. The coordinates of the sequence data is based on National Center for Biotechnology Information (NCBI) build 37 (UCSC hg19).

Bioinformatics

The bioinformatic approach includes a first step trimming of the FAST-Q files with a Ferrer inCode developed method. The trimmed reads are then mapped with GEM II and output is joined and sorted and uniquely and properly mapping read pairs are selected. Finally, variant call over the cleaned BAM file is performed with SAMtools v.1.18, GATK v2.4 together with a Ferrer inCode developed method to generate the first raw VCF files. Variants are annotated with dbSNP IDs, Exome Variant Server and the 1000 Genomes browser, in-home database IDs and Ensembl information, if available.

Tertiary analysis is then developed. For each genetic variation identified, allelic frequency was consulted in Exome Variant Server -EVS- (http://evs.gs.washington.edu/EVS/) and 1000 genomes database (http://www.1000genomes.org/). In addition, Human Gene Mutation Database -HGMD- (http://www.hgmd.cf.ac.uk/ac/index.php) was also consulted to identify pathogenic mutations previously reported. In silico pathogenicity of novel genetic variations were consulted in CONDEL software (CONsensus DELeteriousness scores of missense SNVs) (http://bg.upf.edu/condel/), and PROVEAN (PROtein Variation Effect ANalyzer) (http://provean.jcvi.org/index.php). Alignment among species was also performed for these novel variations using Uniprot database (http://www.uniprot.org/).

Genetic confirmation

Pathogenic known mutations and rare genetic variants were confirmed by Sanger method. First, polymerase chain reaction (PCR) was performed. PCR products were purified using ExoSAP-IT (USB Corporation, Cleveland, OH, USA), and the analysis of the exonic and intron-exon regions
was performed by direct sequencing (Genetic Analyzer 3130XL, Applied Biosystems) with posterior SeqScape Software v2.5 (Life Technologies) analysis comparing obtained results with the reference sequence from hg19. Each sample underwent a genetic study of corresponding genes (NCBI -National Center for Biotechnology Information-, http://www.ncbi.nlm.nih.gov/)(TTN NM_133378, EN1 NM_001426, AKAP9 NM_005751, VCL NM_014000, KCNE3 NM_005472, PKP2 NM_004572). Familial cosegregation of rare genetic variants was also performed using Sanger technology.

3.- Results

Forensic and clinical data

The 11 month-old male was born full term with uneventful antenatal and perinatal history. During his months of life, no anomalous clinical events were reported, including syncopes, infections,
metabolic disorders, or epileptic episodes. The death occurred at night, during sleep. Scene investigation did not reveal any relevant detail. A comprehensive autopsy was performed, revealing that all organs were normal in size and structure, with no evidence of trauma, malignancy, malformation, infection or metabolic disease. Toxicological test and histological study did not reveal any anomalous substance or microscopic alteration, respectively. The forensic conclusion was unexpected death of unknown cause after a thorough investigation, in concordance to San Diego classification of SIDS [14]. The parents were non-consanguineous and neither them of their families showed any previous history of any pathology associated to SD. All relatives were clinically assessed (father, mother and sister). The clinical tests were completely normal, including ECG (Figure 1 and 2), echocardiography (in all of them) and cardiac Magnetic Resonance Imaging –MRI- (in the mother) (data not shown). When the genetic results came out, beta-blockers were started in both mother (bisoprolol) and sister (propranolol) accordingly to recent HRS/EHRA/APHRS guidelines on the diagnosis and management of patients with inherited primary arrhythmia syndromes [15].

Genetics
We analyzed 104 genes previously associated with SCD. After the NGS process and the application of Ferrer inCode bioinformatics pipeline, only a 0.17% of the total bases (379076) were not covered. The exons coverage over the 30X threshold is 99%. We selected the NS variants with a MAF (Minimum Allele Frequency) lower than 1% in the EVS for its conventional Sanger sequencing for confirmation, accordingly to published reports [16]. A total of 7 single nucleotide variants (SNV) identified in 6 different genes were considered as a cause of death in index case. Sanger sequencing confirmed all of them (Table 2 and 3).

Of these 7 SNV, 6 were previously cataloged in international databases and 1 was novel (p.S341R_EN1). Of six reported, two were previously considered pathogenic (p.R83H_KCNE3 and
p.S140F_{PKP2}), three predicted in silico as pathogenic (p.H636R_{VCL}, p.S11996T_{TTN} and p.T21743A_{TTN}), and one predicted as neutral (p.I1643L_{AKAP9}) (Table 2).

Two known pathogenic mutations were reported in HGMD as disease-associated. The first one, p.R83H_{KCNE3} (rs17215437), causes a change of G>A at nucleotide 248 (c.248G>A) in the KCNE3 gene. The MAF was 0.6755/0.1818/0.5082 (European-American/Afro-American/Global respectively). The second one, p.S140F_{PKP2} (rs150821281), causes a change of C>T at nucleotide 419 (c.419C>T) in the PKP2 gene. The MAF was 0.2907/0.0227/0.1999 (European-American/Afro-American/Global respectively) (Figure 3 and 4).

Four more SNV were reported previously but with unknown disease effect. The first one was p.H636R_{VCL} (rs71579374) which causes a change A>G at nucleotide 1907 (c.1907A>G) in the VCL gene. The MAF for this SNV was 0.1279/0.0227/0.0923 (European-American/Afro-American/Global respectively). The second known SNV was p.I1643L_{AKAP9} (rs141990258). This SNV is a change from A>C in the position 4927 (c.4927A>C) of the AKAP9 gene. The MAF for this SNV was 0.0116/0.0/0.0077 (European-American/Afro-American/Global respectively). Both SNV showed a position highly conserved between species. However, p.H636R_{VCL} was predicted in silico as deleterious and p.I1643L_{AKAP9} was predicted as neutral (Figure 3 and 4).

The other two reported SNV were identified in the TTN gene (p.S11996T -rs181189778- and p.T21743A -rs56201325-). The p.S11996T is due to a nucleotide change of T to A in the position 35986. The second SNV, p.T21743A is due to nucleotide change of A to G in the position 65227. The MAF for these two SNV were 0.0/0.0805/0.0251 and 0.5194/0.0523/0.3718 respectively (European-American/Afro-American/Global). Both SNVs were predicted as pathogenic and showed high level of conservation between species (Figure 3 and 4).

Finally, novel SNV (p.S341R_{EN1}) is due to an A>C nucleotide change at position 1021 (c.1021C>A) in the EN1 gene. The genetic variation was predicted in silico as neutral. Alignment between species showed a high level of conservation (Figure 3 and 4).
Family segregation

All 7 genetic variants identified in the index case (II.1) were evaluated in relatives (father –I.1-, mother –I.2- and sister-II.2-)(Figure 3). Genetic analysis showed that index case’s father only carry one of the SNV (p.T21743A\_TTN). Index case’s mother carried all same genetic variants that proband except p.T21743A\_TTN. Finally, index case’ sister carried three SNV also carries by index case (p.I1643L\_AKAP9, p.R83H\_KCNE3, and p.S11996T\_TTN).

4.- Discussion

To the best of our knowledge, the present work is the first study using NGS technology in a family affected by SIDS. Our study identified several potentially genetic variants after NGS genetic analysis of 104 SCD-related genes.

The San Diego classification define SIDS as “death of an infant minor than 1 year of age with onset of the fatal episode apparently occurring during sleep, that remains unexplained after a thorough investigation including performance of a complete autopsy and review of the circumstances of death and the clinical history” [14]. Our case is in concordance with this definition. The genetic analysis in our index case identified 7 genetic variations in 6 different genes that could explain his death. Of them, 2 variants were previously associated with pathologies. Thus, p.R83H\_KCNE3 has been previously associated with a susceptibility to thyrotoxic periodic paralysis [17] but our index case did not show any symptom of paralysis before SD. In addition, familial segregation showed that both his mother and sister also carried this variation but no clinical symptom of paralysis was identified in any of them. This fact suggests no robust association of this genetic variation with SIDS, at least in our family. The other pathogenic variation identified was p.S140F\_PKP2, previously associated with arrhythmogenic right ventricular cardiomyopathy (ARVC) [18] and dilated cardiomyopathy (DCM) [19]. Our index case did not show any clinical sign of ARVC or DCM. It is though accepted that diseases like ARVC may cause SD in very incipient forms, and
pathogenic variations associated with hypertrophic cardiomyopathy (HCM) have been described in SIDS without any microscopic alteration in the myocardium [20]. In our case, autopsy identified no anatomic or microscopic alterations in the myocardium that could suggest a structural disease. In addition, familial segregation showed that the mother carried this pathogenic variation in PKP2 but clinical assessment, which included cardiac MRI, reported no structural alterations. Therefore, we believe that this genetic variation is not the responsible for our SIDS case.

The same argument applies to the two SNV identified in the TTN gene (p.S11996T and p.T21743A). This gene codifies the largest protein in humans, called titin, a crucial protein in the myocytes cytoskeleton [21]. Pathogenic mutations in this gene have been associated with cardiac pathologies with structural alterations, mainly DCM [22]. Both SNV were conserved between species, and predicted in silico as pathogenic. However, and as mentioned before, no structural modifications were identified after autopsy. Family segregation reveals that only the index case carried both SNV in TTN, suggesting that the combination of both SNV could be responsible for the death. However, we believe that it is not a consistent explanation because of the lack of structural alteration in the case and family members. All these facts suggest that none of the SNV identified in TTN protein are responsible for the death of the index case.

Two previously reported SNV were identified in the index case. The p.H636R_VCL is a missense variation localized in a highly conserved position between species, and predicted in silico as damaging. These facts suggest a pathogenic role, supported by reduced MAF. However, the VCL gene encodes the vinculin protein, a structural protein that induces structural alterations in myocardium diseases, such as HCM and DCM [23]. However, no anatomical signs of structural disease were observed in autopsy, thus we could discard this variation as a potential cause of SIDS.

The second known SNV identified in index case was p.I1643L_AKAP9. Pathogenic variations in this gene have been associated with LQTS, an inherited heart disease associated with SCD [24]. This SNV was previously reported in NHLBI Exome Sequencing Project (ESP) with unknown clinical effect, but a much reduced MAF ratio and highly conserved position between species
suggest a potential pathogenic role. A recent report recommend a cautious interpretation of these genetics variants in SIDS families [25]. In addition, the index case’s mother and sister carried p.I1643L_\textit{AKAP9} but remained asymptomatic at the moment of study. This fact agrees with the incomplete penetrance, a hallmark of LQTS families. It is important to note that SCD may be the first manifestation of the pathology. Therefore, family members carrying potentially pathologic variants associated with LQTS, even being asymptomatic, are at high risk of sudden death. Taking all this into account, we believe that this is the genetic variation potentially related to the child’s death. In consequence, as a preventive measure of SCD and following the current guidelines [15], betablocking agents (bisoprolol in mother and propranolol in sister), were given in both mutation carriers.

The last SNV identified in index case (p.S341R_\textit{EN1}) was a novel genetic variation predicted \textit{in silico} as neutral. However, aminoacid position with a high level of conservation between species suggesting a potential pathogenic role. Familial segregation showed that index case’s mother also carried the same variation. Previous studies in SIDS identified pathogenic mutations in the \textit{EN1} gene, suggesting that p.S341R_\textit{EN1} could also have a potential pathogenic role for SIDS, at least in our family. However, \textit{in vitro} studies should be also performed to clarify its potential pathogenic role.

There are some limitations in our study that should be noted. Hence, the main limitation is that our study only focuses on one family. Further NGS studies should be performed to improve the identification of genetic causality in SIDS cases. Another limitation is that our index case could carry genetic variations localized in other genes not analyzed in our study that could explain the cause of death. As mentioned before, \textit{in vitro} studies of genetic variants of unknown significance (GVUS) should be performed in order to clarify their pathogenic role. In addition, the death could be also due to copy number variation (CNV), previously reported in SIDS cases [26]. Recently, organ specific miRNA dysregulation has been suggested to pathogenesis of SIDS [27], despite further studies should be performed to elucidate the role of miRNA in SIDS. In other study, it has
been reported that high levels of cytokine could lead to SIDS, mainly due to inflammatory mechanisms [28]. All these facts support SIDS as a multifactorial disorder focus on association of genetic variants with intrinsic and extrinsic risk factors as cause of death, supporting the use of NGS technologies in SIDS cases. Taking all these limitations into account, pathogenic interpretation of novel genetic variants and translation into clinical practice should be analyzed and consensus with extremely careful by a team of specialist in different areas.

Despite that an important portion of SIDS cases have a genetic origin, molecular autopsy is not yet included as standard of care in current forensic protocols. In the present report, we show the challenges of NGS in SIDS, as a paradigm of unexplained death or even as paradigm of unexplained syncope. The combination of several genetic variations together with the characteristic incomplete penetrance shows the current challenge for geneticists and clinicians to undertake clinical decisions. Familial genotyping is crucial to clarify pathogenic role of unknown genetic variants and to identify other genetic carriers at risk of SCD. Thus, a multidisciplinary team is essential to perform a correct interpretation of all genetic data obtained by NGS technology, and provide a helpful genetic counseling for families affected by SIDS.
References


Figure 1.-Twelve-lead ECG of mother’s index case. The ECG does not show any alteration.
Figure 2.- Twelve-lead ECG of sister’s index case. The ECG does not show any alteration.

Figure 3.- Pedigree of the family. Index case is II.1 Slash represents SD. White Round/squares indicates healthy status after clinical evaluation. Below each family member, all genetic variations identified by NGS technology and confirmed by Sanger method. Plus sign indicates carrier of genetic variation. Minus sign indicates non-carrier of the genetic variation.

Figure 4.- Conservation and taxonomy between species. Asterisk indicates the position of each genetic variation (p.H636R_ VCL, p.I1643L_ AKAP9, p.S341R_ EN1, p.11996T_ TTN and p.21743A_ TTN).

Tables Legend

Table 1.-

ABCC9, ACTA2, ACTC1, ACTN2, AKAP9, ANK2, CACNA1C, CACNA1G, CACNA1H, CACNA1I, CACNB2, CASQ2, CAV3, CHRM2, CRYAB, CSRP3, CTF1, DES, DMD, DSC2, DSG2, DSP, ECE1,
EMD, EN1, EYA4, FBN1, FHL2, FKTN, GJC1 (GJA7), GLA, GPD1L, HCN1, HCN2, HCN4, ILK, JPH2, JUP, KCNA4, KCNA5, KCND2, KCND3, KCNE1, KCNE2, KCNE3, KCNH2, KCNJ2, KCNJ3, KCNK4, KCNQ1, LAMA4, LAMP2, LDB3, LMNA, MYBPC3, MYH6, MYH7, MYL2, MYL3, MYLK2, MYOZ2, MYPN, NEXN, NOS1AP, NPPA, NUP155, PDLIM3, PHOX2A, PHOX2B, PKP2, PLN, PRKAG2, PSEN1, PSEN2, RBM20, RET, RYR2, SCN1B, SCN2B, SCN3B, SCN4B, SCN5A, SGCA, SGCB, SGCD, SIRT3, SLC25A4, SLC6A4 (5HTT), SLC8A1, SNTA1, TAZ, TCAP, TGFB3, TGFB1, TGFB2, TLX3, TMEM43, TMPO, TNNC1, TNNT2, TPM1, TTN, VCL

Table 1.- List of the 104 SCD-related/suspicious genes included in our panel.

Table 2.-

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Nucleotide</th>
<th>Protein</th>
<th>MAF% (EA, AA, all)</th>
<th>dbSNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKAP9</td>
<td>A-kinase anchor protein 9</td>
<td>c.4927A&gt;C</td>
<td>p.I1643L</td>
<td>0.0116/0.0/0.0077</td>
<td>rs141990258</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Variant</td>
<td>Allele</td>
<td>MAF</td>
<td>rsID</td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
<td>----------</td>
<td>---------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>EN1</td>
<td>Engrailed homeobox 1</td>
<td>c.1021A&gt;C</td>
<td>p.S341R</td>
<td>NI</td>
<td>Novel</td>
</tr>
<tr>
<td>KCNE3</td>
<td>MiRP2</td>
<td>c.248G&gt;A</td>
<td>p.R83H</td>
<td>0.6755/0.1818/0.5082</td>
<td>rs17215437</td>
</tr>
<tr>
<td>PKP2</td>
<td>Plakoglobin</td>
<td>c.419C&gt;T</td>
<td>p.S140F</td>
<td>0.2907/0.0227/0.1999</td>
<td>rs150821281</td>
</tr>
<tr>
<td>VCL</td>
<td>Vinculin</td>
<td>c.1907A&gt;G</td>
<td>p.H636HR</td>
<td>0.1279/0.0227/0.0923</td>
<td>rs71579374</td>
</tr>
<tr>
<td>TTN</td>
<td>Titin</td>
<td>c.35986T&gt;A</td>
<td>p.S11996T</td>
<td>0.5194/0.0523/0.3718</td>
<td>rs181189778</td>
</tr>
<tr>
<td>TTN</td>
<td>Titin</td>
<td>c.65227A&gt;G</td>
<td>p.T21743A</td>
<td>0.0/0.0805/0.0251</td>
<td>rs56201325</td>
</tr>
</tbody>
</table>

Table 2.- Genetic data of index case. MAF, Minor Allele Frequency (%). EA, European-American. AA, Afro-American. NI, No identified.
Fig1_ECG_Mother.
Fig2_ECG_Sister
Figure 3.