Forensic Science International xxx (2015) e1–e6

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Case Report

Using whole exome sequencing and bioformatics in the molecular autopsy of a sudden unexplained death syndrome (SUDS) case

Chun Wang [a], Shan Duan [b], Guoli Lv [c], Xiaoping Lai [d], Rui Chen [d], Hanguang Lin [d], Shengyuan Qiu [d], Jianpin Tang [d], Wenjian Kuang [d], Chuanchao Xu [d,*]

[a] Chengdu Sport University, Chengdu, Sichuan 610041, PR China
[b] Shenzhen Research Institute of Population and Family Planning, Shenzhen, Guangdong 518040, PR China
[c] Guangzhou institute of Forensic Science, Guangzhou, Guangdong 510030, PR China
[d] Guangdong Medical College, Dongguan, Guangdong 523808, PR China

ARTICLE INFO

Article history:
Received 30 March 2015
Received in revised form 27 July 2015
Accepted 28 August 2015
Available online xxx

Keywords:
Forensic science
Forensic bioinformatics
Sudden unexplained death syndrome (SUDS)
Second-generation sequencing
Exome
Disease mutations

ABSTRACT

Whole exome sequencing (WES) and bioinformatics analysis were used to investigate potential disease-causing gene mutations in a sudden unexplained death syndrome (SUDS) case after autopsy and pathology tests failed to suggest an obvious disease mechanism. Following whole exome sequencing, a 3-step bioinformatics filtering procedure was carried out to identify possible pathogenic genomic features. Single nucleotide variations (SNVs) were analyzed and ranked by likely mutation impact using various open online tools. After screening, we identified G643S as a putative causative heterozygous mutation in the KCNQ1 gene. This mutation has been reported in abnormalities consistent with SUDS, such as IKs in cardiac myocytes, a condition that predisposes for arrhythmias. Our work demonstrates the application of sequencing technology at the whole exome level for determining potential causes of an otherwise unexplained death.

1. Introduction

Sudden unexplained death syndrome (SUDS) [1], or Sudden manhood death syndrome [2] is the sudden unexpected death of adolescents and adults up to 35–40 years old. In postmortem autopsy and pathological examinations, no significant or lethal morphological changes can be found in major organs or tissues [3,4]. For many years, most SUDS cases have been attributed to an unknown type of cardiomyopathy, based on the characteristics of sudden death and the non-specific changes of acute cardiac dysfunction. However, the underlying pathogenic mechanism of SUDS remains unknown.

In recent decades, studies have shown that the pathological characteristics of SUDS are similar to two kinds of inherited cardiomyopathy: channelopathy and cardiomyopathy [5,6]. Examples of cardiac ion channel diseases are long QT syndrome (LQTS), Brugada syndrome (BS), catecholaminergic ventricular tachycardia arrhythmia (CPVT) [7–9] and short QT syndrome (SQTS) [10,11]. Types of cardiomyopathy include hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), and arrhythmogenic right ventricular cardiomyopathy (ARVC) [12–14]. Patients with these conditions are often asymptomatic and sudden cardiac death can occur either during exercise or physical activity or while sleeping.

With the rapid development of genomics and molecular genetics technology, the mechanisms of ion channelopathy and cardiomyopathy are becoming clearer. Most of these diseases have a genetic component, for example, over a thousand mutation variants are related to HCM [15]. To date, about ninety inherited genes have been associated with cardiac diseases [5].

Usually, the genetic testing of a patient with a clear clinical diagnosis has candidate genes to screen for the pathogenic mutations. For example, genetic testing of a patients having LQTS will focus on genes encoding ion channel proteins in cardiac myocytes, including KCNE1, -2, -3, KCNQ1, KCNH2, KCNJ2, SCN5A, etc. [2,5–8]. However, clinical history that would guide the analysis

* Corresponding author at: Department of Forensic Science, Guangdong Medical College, No. 1, Xin Cheng Road, Songshan Lake District, Dongguan, Guangdong 523808, PR China. Tel.: +86 769 22896443; fax: +86 769 22896076.
E-mail address: chuanchaoxu_GDMC@hotmail.com (C. Xu).

http://dx.doi.org/10.1016/j.forsciint.2015.08.022
0379-0738/

of a specific candidate gene or gene family in SUDS cases is often lacking. Therefore, researchers have to screen approximately 90 of the inherited cardiomypathy candidate genes to find the possible disease-causing mutation(s). It is very expensive and time-consuming to use conventional molecular biology techniques to assay multiple suspect mutations. Consequently, genetic testing for SUDS is not yet widely used.

Exome sequencing is a powerful, high-throughput genomic analysis technology based on the second-generation sequencing platform that is being applied to define the extent to which rare genetic features explain the heritability of complex diseases and health-related traits [16,17]. Compared to conventional techniques, exome sequencing is currently the most powerful and cost-effective method for the screening of single nucleotide polymorphisms (SNPs) and mutations in the coding genome region. Therefore, this method is also referred to as the "molecular autopsy" [18–20]. Here, we used second-generation exome sequencing technology for molecular investigation of one SUDS case and screened for suspicious mutations in 90 inherited cardiac diseases-associated genes [5,18,20]. As the cost of WES is rapidly decreasing [21,22], it will eventually become affordable to use WES in the forensic examination of the cause of death of SUDS victims.

2. Materials and methods

2.1. Subject

A 34-year-old male factory worker died in the dormitory while sleeping. Family members stated that the patient had no family history of heart disease.

2.2. Autopsy and pathology testing

A systematic autopsy was performed by a forensic scientist. The heart, brain, liver, kidney, lung, and other major organ tissues were collected for histopathological examination. Blood from the heart was collected for toxicological analyses.

2.3. Exome sequencing

50 mg gDNA was isolated from skeletal muscle using the QIAGEN Genomic DNA isolation kit (QIAGEN, Germany) according to the manufacturer's protocol. DNA purity was assessed using the ND-1000 Nanodrop (Nanodrop Technologies Inc., USA). DNA sample had an A260:A280 ratio above 1.8 and A260:A230 ratio above 2.0. DNA integrity was evaluated using agarose gel electrophoresis. Briefly, DNA sample was subjected to fragmentation, end-repair and dA tailing following by adaptor ligation and enrichment with a low-cycle according to instructions of TruSeq® DNA LT/HT Sample Prep Kit (Illumina, USA). Then, up to six DNA libraries were pooled together and captured by exon probes using the TruSeq® Exome Enrichment Kit (Illumina, USA). The purified library products were evaluated using the Agilent 2200 TapeStation and Qubit® 2.0 (Life Technologies, USA) and then diluted to 10 pM for cluster generation in situ on the HiSeq2500 (Illumina, USA) pair-end flow cell followed by sequencing (2 × 100 bp) on HiSeq 2500.

2.4. Bioinformatics analyses

FASTQ files (raw data) were filtered to remove the sequences of linker, *N*-value (invalid nucleotide data), and low quality reads. The remaining high-quality filtered reads were aligned to the human reference sequence (hg19) using a Burrows–Wheeler aligner (BWA) to obtain BAM files for bioinformatics analyses (Table 1). After the variants were annotated (Table 2), a 3-step variant filtration procedure was carried out (Fig. 1), which involved the exclusion of all noncoding regions and synonymous variants (i.e. DNA nucleotide alterations that do not alter the amino acid sequence of the coded protein) and the exclusion of none-90 channelopathy and cardiomyopathy gene-specific variants [5,18]. Next, to determine if a mutation may have contributed to the sudden death, variants within the cardiomyopathy/channelopathy-associated gene subset were screened in three publicly available exome databases for minor allele frequency (MAF). The three databases used were: the Single Nucleotide Polymorphism Database (dbSNP), the 1000Genomes Project (http://www.1000genomes.org/home, $n = 1094$ subjects; 381 Caucasian, 246 African–American, 286 Asian, and 181 Hispanic subjects), National Heart, Lung, and Blood Institute (NHLBI) Grand Opportunity Exome Sequencing Project ($n = 6503$ subjects; 2203 African–American and 4300 European–American unrelated individuals (13,006 chromosomes total). The single-nucleotide variants (SNVs) with an allele frequency less than and around 1% were identified.

Selected mutations were subsequently verified by Sanger sequencing (primers shown in Table 4). For validated mutations,

Table 1
Summary of alignment results.

Exon capture statistics	Sample data
Target size (bp)	62,085,295
Clean reads	60,206,990
Aligned reads	59,408,854
% Aligned reads	98.67
Reads in target region	31,722,915
% Reads in target region	53.4
Mean depth of target region	51.1
Coverage of target region	94.25
Reads in flanking region	18,800,058
% Reads in flanking region	31.65
Mean depth of flanking region	26.16
Coverage of flanking region	87.36
% Targeted region covered at depths of at least 4×	93.99
% Targeted region covered at depths of at least 10×	93.14
% Targeted region covered at depths of at least 20×	90.31
% Flanking region covered at depths of at least 4×	86.11
% Flanking region covered at depths of at least 10×	80.19
% Flanking region covered at depths of at least 20×	63.86
Non-duplicated reads	39,584,226
% Non-duplicated reads	66.63
Uniq. Mapped reads	39,584,226
% Uniq. Mapped reads	66.63

Table 2
Summary of identified SNPs.

SNPs	Properties	Sample data
Known or new SNPs site	Known SNPs	185,491
	New SNVs	8778
Heterozygosity	Homozygous	94,923
	Heterozygous	99,346
Classification and annotation of functional gene elements	UTR3	19,676
	UTR5	3247
	Downstream	2831
	Exon	21,461
	Intron	75,996
	ncRNA	12,483
	Cleavage site	112
	Upstream	3236
	Upstream and downstream*	199
Type of mutation	Missense	10,272
	Stop codon acquisition	90
	Stop codon missing	12
	Synonymous	10,940

* Some genes have multiple transcripts, an SNP/SNV here means that it is located at the upstream side of one transcript, while also at the downstream side of another one.

Please cite this article in press as: C. Wang, et al., Using whole exome sequencing and bioformatics in the molecular autopsy of a sudden unexplained death syndrome (SUDS) case, Forensic Sci. Int. (2015), http://dx.doi.org/10.1016/j.forsciint.2015.08.022