Clinica Chimica Acta 453 (2016) 80–85

Contents lists available at ScienceDirect

Clinica Chimica Acta

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

A fast and cost-effective molecular diagnostic tool for genetic diseases involved in sudden cardiac death

Valérie Chanavat [a,b], Alexandre Janin [a,c,d], Gilles Millat [a,b,c,d,*]

[a] Laboratoire de Cardiogénétique Moléculaire, Centre de Biologie et Pathologie Est, Hospices Civils de Lyon, Lyon, France
[b] NGS Sequencing Platform for Molecular Diagnosis, Hospices Civils de Lyon, Lyon, France
[c] Université de Lyon, Lyon F-69003, France
[d] Université Lyon 1, Lyon, France

ARTICLE INFO

Article history:
Received 29 September 2015
Received in revised form 8 December 2015
Accepted 9 December 2015
Available online 10 December 2015

Keywords:
Mutations
Sudden cardiac death
Cardiomyopathies
Arrhythmia syndromes
NGS sequencing
Molecular diagnosis

ABSTRACT

Background: Cardiomyopathies and arrhythmia syndromes are common genetic cardiac diseases that account for a significant number of sudden cardiac death (SCD) cases.
Methods: NGS workflow based on a panel of 95 genes was developed on Illumina NextSeq500™ sequencer for sequencing prevalent SCD-causing genes. A cohort of 90 patients (56 genotype-positive, 27 genotype-negative and 7 new cases) was screened to evaluate this strategy in terms of sensitivity, specificity, practicability and cost. *In silico* analysis were performed using a pipeline based on Next*GENe*® software and a personalized Sophia Genetics pipeline.
Results: Using our panel custom, 100% of targeted sequences were efficiently covered and all previously identified genetic variants were readily detected. Applied to 27 genotype-negative patients, this molecular strategy allowed the identification of pathogenic or likely pathogenic variants into 12 cases. It confirmed the involvement of *HCN4* mutations in the combined bradycardia–myocardial non-compaction phenotype, and also suggested, for the first time, the involvement of *PKP2*, usually associated with arrhythmogenic right ventricular dysplasia, in ventricular non-compaction.
Conclusion: This NGS approach is a fast, cheap, sensitive and high-throughput mutation detection method that is ready to be deployed in clinical laboratories and would provide new insights on physiopathology of SCD, more particularly of cardiomyopathies and arrhythmia syndromes.

1. Introduction

Sudden cardiac death (SCD), a major cause of death in developed countries, is defined as the unexpected death occurring within 1 h from the onset of symptoms in a subject with no known prior condition that would appear fatal [1]. In the general population, the incidence ranges between 300,000 and 400,000 cases per year in the USA and a similar number in Europe [2]. The pathophysiology is complex. SCD can occur at any age, and with or without detectable ("structural") heart disease. Coronary artery disease is the most common substrate underlying SCD (~40–60% of SCD cases), more particularly in patients aged 50 years and above, but primary hereditary myocardial diseases such as cardiomyopathies (dilated, hypertrophic, left ventricular noncompaction, and arrhythmogenic right ventricular cardiomyopathy) and arrhythmia syndromes (long QT syndrome, Brugada syndrome, and catecholaminergic polymorphic ventricular tachycardia) account for most of the of the residual SCD cases [3,4].

Molecular analysis of SCD patients is however challenging owing to the large cohort of patients to investigate, the number of putative disease-causing genes, and the presence of a large spectrum of private mutations. Because of this genetic and allelic heterogeneity, molecular testing could only be envisaged with Next Generation Sequencing approaches. Previous studies have already reported the efficiency of NGS approaches to sequence gene panels on benchtop NGS sequencers [5,6]. Thus, the aim of our study was to evaluate the efficiency of a fast protocol based on a large custom panel (95 genes) and NextSeq500 Sequencer for sequencing most prevalent SCD-causing genes. Besides to give the opportunity to establish a likely cause of death for SCD cases, more particularly in a forensic setting, this panel could also be used to investigate cases with either cardiomyopathies and/or arrhythmia syndromes.

A cohort of 90 cases was screened in order to evaluate this strategy in terms of sensibility, specificity, practicability and cost. *In silico* analysis was performed using either Next*GENe*® software as previously reported [5,6] or a personalized pipeline provided by Sophia Genetics.

Abbreviation: DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy; NGS, Next-Generation Sequencing; SCD, sudden cardiac death.

* Corresponding author at: Laboratoire de Cardiogénétique Moléculaire, Centre de Biologie Est, F-69677 Bron Cedex, France.
E-mail address: gilles.millat@chu-lyon.fr (G. Millat).

http://dx.doi.org/10.1016/j.cca.2015.12.011
0009-8981/

The diagnosis is important for the mourning process of those left behind. In addition to complete clinical history, molecular genetic testing is a crucial additional tool for the diagnosis of these syndromes. It can help to establish or confirm a clinical or presumptive diagnosis, to study relatives, and to make appropriate genetic counseling. It is also useful to stratify the risk associated with specific mutations and therefore, to offer appropriate treatment to family members.

2. Materials and methods

2.1. Patients

Genomic DNA samples from cardiomyopathies or arrhythmia syndromes patients were used to determine the sensitivity of this NGS mutation detection approach. The clinical diagnostic criteria were established according to international criteria [http://www.escardio.org/Guidelines-&-Education]. The cohort was established with 56 genotype-positive patients (previously screened on prevalent disease-causing genes either for cardiomyopathies [*LMNA*, *MYBPC3*, *MYH7*, *TNNT2*, *TNNI3*, and *TPM1*] or for arrhythmia syndromes [*KCNH2*, *KCNQ1*, *KCNE1*, *KCNE2* and *SCN5A*]), 27 previously genotype-negative patients, and 7 new patients. The study was conducted in accordance with the principles of the Declaration of Helsinki. Informed consent was obtained for all cases.

2.2. Library preparation and NGS sequencing

Genomic DNAs were tested by NGS sequencing using a custom design based on a SeqCap EZ Solution-Based Enrichment strategy (Roche NimbleGen Madison, Wisconsin, USA). Targeted sequencing capture probes were custom designed by Roche NimbleGen. The panel was designed to identify disease-causing mutations in 95 genes associated with sudden cardiac death (Table 1). The total length of the target regions was 863 kb. Target regions included coding exons (padding: +/− 30 bp), 5′- and 3′-UTR regions.

For each sample, 1 μg of high quality genomic DNA was fragmented with a Covaris M220 instrument (Covaris, Woburn, MA, USA). Library preparation was performed with the Kapa Library Preparation Kit for Illumina platforms (Kapa Biosystems, London, UK). The manufacturer's DNA sample preparation protocol for Roche NimbleGen SeqCap EZ Library (Roche, SeqCap-EZ_UGuide_v5p0) was followed, using single-index adapters (SeqCap Adapter Kit A et B). To reduce experimental costs, pre-enrichment pools were performed by equimolarly pooling 12 samples. After performing the SeqCap EZ enrichment protocol, the 2 pools of 12 samples were equimolarly pooled to a single pool that was sequenced on a NextSeq500 sequencer using the NextSeq500 Mid Output Kit v2 (300 cycles) chemistry (Illumina, San Diego, CA, USA). Each run allowed the targeted resequencing of 24 patients. Validation of enrichment and quantification of enriched target DNA were performed both on the Caliper LabChip GX using the High Sensitivity assay Kit (Caliper LifeSciences Waltham, Massachusetts, USA) and on Qubit Fluorometric Quantitation using Qubit DSDNA HS Assay kit (ThermoFisher Scientific, Illkirch, France).

2.3. Bioinformatic analyses

Fastq files (8 fastq files/patient) were uploaded from BaseSpace® (Illumina). Bioinformatics analyses were further performed using either a private pipeline developed by Sophia Genetics (Sophia Genetics, Lausanne, Switzerland) or a home-made pipeline based on Next*GENe*® v.2.3.4.2 (SoftGenetics, State College, PA, USA), and Alamut® 2.7.1 (Interactive Biosoftware, Rouen, France) softwares as previously reported [5,6].

The home-made pipeline allowed us to perform primary to tertiary analysis, including optimized signal processing, base calling, sequence alignment, and variant analysis. The CNV Tools proposed by Next*GENe*® (based either on the "SNP-Based Normalization with smoothing" algorithm or on the "Dispersion and HMM with RPKM") were also used to look for copy-number variations (CNVs) in the patients.

Putative identified pathogenic variants (SNV, indels) were verified by conventional dideoxy sequencing using BigDye® Terminator v.3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad, CA, USA). Putative identified CNVs were verified either using MLPA or CGH array methodologies. According to the recently reported guidelines, specific standard terminology ["pathogenic," "likely pathogenic," "uncertain significance," "likely benign," and "benign"] were further used to interpretate the putative pathogenicity of variants identified in studied genes [7].

3. Results

Four runs, containing 24 DNA samples each, were performed for development and validation of our NGS assay. Ninety samples were corresponded to genomic DNA samples from patients with cardiomyopathies and/or arrhythmia syndromes. Among the 6 remaining DNAs, 2 of them were used as intra- and inter-run replicates and 2 genomic DNAs were provided by Sophia Genetics for the development of a personalized bio-informatic pipeline. The coverage statistics were comparable between each run. The cluster density range, critical metric that impacts run quality, was between 198 K (+/− 12) clusters/mm^2, generating 48.7 Gb (+/− 3.6) by run with an average of 89.2% (+/− 1.2%) Q30 bases per sample sequenced. Coverage analysis allowed us to observe that 100% of target regions were covered with a minimum of 30× coverage, >99.3% of target regions were covered with 100× coverage, and >85% of target regions were covered with 1000× coverage. According to previous studies, only exons (and their adjacent boundaries sequences) with a read depth above 30 reads (30×) for each targeted nucleotide were considered as correctly covered [8,9]. Thus, using these settings, our custom panel allowed us to efficiently explore all targeted regions and do not require additional Sanger sequencing to explore uncovered regions.

A cohort of 56 genotype positive patients was used to evaluate our workflow. Globally, with selected genotype-positive patients, we expected to identify 58 pathogenic variations: 7 CNVs, 9 substitutions, and 42 short indels, some of them located in homopolymers regions (Table 2). Alignment of fasta files against hg19 reference genome was performed with Next*GENe*® software using adjusted alignment settings [5,6]. To achieve adequate specificity while maintaining satisfactory sensitivity for identification of pathogenic mutations (which are mostly private), gene variants from all patients present in a same flow cell were compiled on a unique Excel sheet file. Using the Excel's concatenate function (concatenation of "Mutation Call" and "chromosome position" columns), we further only took attention to variants identified in less than 3 patients or recurrent variants identified as different in terms of homogeneity. Using this strategy, all previously identified genomic variants (common SNPs, missense mutations and short indels) were easily identified. Similar results were obtained using the personalized Sophia Genetics pipeline. More interestingly, for CNV identification, the best results were obtained with the personalized Sophia Genetics pipeline as it allowed to detect them without being bothered by false-positives. Using the CNV tool of the Next*GENe*® software, the 7 known CNVs were identified but some false-positives were also detected.

This strategy was further applied to 27 previously genotype-negative patients and to 7 new index cases. Molecular analysis of this cohort allowed identification of 31 mutated alleles: 21 missense variations, 3 nonsense mutations, 2 splice mutations and 5 CNVs (Table 3). The 5 CNVs were highlighted by the Sophia Genetics pipeline. Among the 27 previously genotype-negative patients, putative pathogenic variants were detected in 17 cases (63%). More specifically, pathogenic variations (2 deletions and 2 substitutions) were identified for 4 cases (2 LVNC, 1 HCM and 1 LQT cases) and likely-pathogenic variations (8 substitutions and 1 substitution) were identified for 8 cases (4