PROJECT REPORT

DEVELOPMENT OF A NEXT GENERATION SEQUENCING-BASED DIAGNOSTIC METHOD FOR INHERITED CARDIAC DISEASES CAUSING SUDDEN CARDIAC DEATH IN ATHLETES

Robert Sepp MD, PhD

2nd Department of Internal Medicine and Cardiology Center, University of Szeged

1. BACKGROUND

Sudden cardiac death (SCD), which is defined as a natural death due to cardiac causes, preceded by abrupt loss of consciousness within one hour of the onset of acute symptoms, is a devastating consequence of cardiac diseases. While the majority of SCD cases are due to coronary artery disease, there are a considerable number of patients where genetic causes lead to SCD. These factors are especially important in the young, where coronary artery disease is very rare. The occurrence of SCD in young people or elite athletes, who have no previously known cardiac abnormality, is always shocking and frequently raise great media attention.

Sudden cardiac death may occur as a consequence of an inherited genetic abnormality affecting key proteins of the heart. Diseases such as ion channel diseases or cardiomyopathies are the best known examples of monogenic diseases predisposing to SCD.[1]Apart form these monogenic diseases there is evidence for supporting the existence of a genetic 'susceptibility factor' predisposing to SCD. The latter has emerged from large-scale epidemiological studies where familial association of SCD was demonstrated. These studies provide a major departure from the conventional view that biochemical and clinical markers are the most important predictors of SCD.

Ion channel diseases, or channelopathies, are diseases arising from genetic mutations affecting ion channels of the heart.[1]As the normal generation and propagation of the cardiac action potential relies on the highly ordered flow of ionic currents, provided by the fine tuned opening and closure of ion channels, any disturbance of ion channel function is likely to lead rhythm disturbances, and, in unfortunate cases, to fatal arrhythmias causing sudden cardiac death. Ion channel diseases of the heart are best exemplified by the long QT syndrome (LQTS), Brugada syndrome (BrS) or catecholaminergic polymorphic ventricular tachycardia (CPVT). The long QT syndrome (LQTS) characterized by an abnormally prolonged QT interval and, usually, by stress-mediated life-threatening ventricular arrhythmias. Since the discovery of the primary LQTS-causative genes of potassium (*KCNQ1* and *KCNH2*)[2, 3] and

sodium (*SCN5A*)[4] channels at least 13 LQTS genes have been reported, all encoding different ion channels. The Brugada syndrome (BrS)[5, 6] is an arrhythmogenic disorder associated with high risk of SCD. It has been demonstrated that BrS is an allelic disorder to LQT3 presenting mutations on the cardiac sodium channel (*SCN5A*).[7] Later research showed that at least eight other ion channel genes may be causally involved.[1] Although the *SCN5A* gene accounts for the vast majority (>75%) of BrS genotype positive cases, the yield of *SCN5A* genetic testing for clear clinical cases of BrS is approximately 25%. Thus, the majority (>65%) of BrS cases remain genetically elusive. Catecholaminergic polymorphic ventricular tachycardia (CPVT)[8] is characterized by adrenergically mediated ventricular arrhythmias causing syncope, cardiac arrest, and sudden cardiac death (SCD).. Approximately 65% of CPVT index cases have a mutation in the gene encoding for the cardiac ryanodine receptor (*RYR2*)[9] while there are cases resulting from mutations in the gene for cardiac calsequestrin (*CASQ2*),[10] with a prevalence estimated at approximately 3% to 5%. Genetic screening allows the identification of mutations in up to 65% of patients with a clinical diagnosis. Other cardiac ion channel diseases include the Andersen-Tawil syndrome, Timothy syndrome, cardiac conduction disease (CCD) and short QT syndrome (SQTS). There is evidence that a portion of sudden unexplained death syndrome (SUDS) and sudden infant death syndrome (SIDS) are also due to ion channel mutations.[1]

Enormous progress has recently been made in identifying the genetic causes of different forms of cardiac disease causing sudden cardiac death, which, in turn, has enabled greater understanding of molecular mechanisms underlying each disease. This progress has also increased the probability of establishing specific genetic diagnoses, thereby providing new opportunities to use this genetic information. Although molecular genetics provided fundamental insights into the pathogenesis of ion channel diseases, many open questions remained. Despite extensive interrogation by Sanger (capillary) sequencing of previously identified causative genes, only 25-70% of patients with ion channel diseases have identifiable mutations, even in familial cases. This relatively low sensitivity is a result of

several factors. First, not all potentially affected causative genes have been identified and second, not all underlying mutations in the commonly affected ion channel genes have been identified. There is also a significant problem testing for ion channel mutations in patients that carry multiple disease causing mutations, whereby more severe disease may be present and lead to complex phenotypes.[9]. Lastly, current methods do not test known disease genes for deep intronic mutations, large insertions and deletions, or gross genomic rearrangements, which can lead to the reduced sensitivity of genetic testing.

The 2nd Department of Internal Medicine and Cardiology Center, University of Szeged has a long-standing interest in familial cardiac diseases, in ion channel diseases and cardiomyopathies in particular.[11-13] With regard to ion channel diseases the group described the first congenital long QT syndrome in Hungary.[14, 15] Molecular genetics to investigate ion channel diseases has been first applied at this institution in Hungary. We were also the first in Hungary in establishing a molecular genetic laboratory for successful identification of disease causing mutations in long QT syndrome (*KCNQ1, KCNH2, KCNE1* genes)[16-18] and Andersen syndrome (*KCNJ2* gene).[19] Links have been established aiming the functional characterization of mutations by means of expression of mutations in vitro and in transgenic mice. Parallel to this, foundations of a biobank have been laid down as we have been collecting blood samples from patients with different forms of ion channel diseases and cardiomyopathies. Up to now, the biobank contains blood samples from about 100 families with LQTS or other ion channel diseases. Samples from several peculiar families, characterized by multiple cases of SCD and conduction diseases in the same family or a family with familial bradycardia are also stored in the biobank.

2. AIMS OF THE PROJECT

Our aim was to develop a next generation sequencing-based diagnostic method for inherited cardiac diseases causing sudden cardiac death in athletes

3. PATIENTS AND METHODS

PATIENT SELECTION FROM THE EXISTING IN-HOUSE BIOBANK

The research group of 2nd Department of Internal Medicine and Cardiology Center, University of Szeged, lead by Prof. Miklós Csanády and involving senior researchers Prof. Tamás Forster and Dr. Róbert Sepp has built up a special expertise for characterization of patients with channelopathies. They have also established a clinically well characterized, followed up and ever growing Biobank, fully authorized, for ion channel diseases. Patient cohort selection utilized the resources provided by the Biobank: due to the ongoing biobanking activities blood samples from patients were available, thus no extra time was needed for sample collection. Importantly, all patients from whom biological material is stored within the Biobank gave informed consent for genetic screening. We planned to analyse 16 unrelated patients with ion channel disease, including the long QT syndrome (LQTS) as positive controls for the test method we developed. Long QT patients with already known mutations were also included.

TARGETED RESEQUENCING

We have decided to use the targeted resequencing approach, in which we aim to enrich only a fraction of the human genome, by targeting coding regions of selected genes. Hence, essential for the successful execution of the project was the efficient targeting approach. Currently there were several available technologies on the market that utilizes either RNA or DNA baits for capturing. Because DNA-RNA hybrids are yielding increased specificity of the hybridization, we used the Agilent's SureSelect technology in which 120-mer RNA baits specific to the target region are delivered by the supplier. The Sequencing Platform operating at the Biological Research Centre of the Hungarian Academy of Sciences (BRC HAS) has gained the necessary expertise in using this technology (applied, among others, in the pilot project). Thus, targeted resequencing was performed on a SOLiD 5500xl System (Life

Technologies). The SOLiD System offered the combination of high throughput and accuracy necessary for assaying large numbers of candidate regions.

Regions of interest: A single targeted resequencing platform has been established which contained all the causative genes for ion channel diseases (LQTS, Brugada syndrome, CPVT) and was used to screen the whole patient/control cohort. The platform included the full coding sequences of those genes that have already been associated with ion channel diseases. The platform also included "giant" genes - never screened before in full length - e.g. ryanodine receptor *(RYR2)*.

Probe design: we have designed long (120-mer) RNA oligonucleotides (probes) for capturing the regions of interest. Two different bait design strategies have been applied: **a)** centered baits - for large target regions - have even coverage across entire region (2x tiling) and **b)** justified baits - for short target regions - have uneven coverage (2-3x tiling). Importantly, repetitive elements have been masked out.

Sample processing: intact, high-molecular weight DNA was purified from blood and was fragmented by sonication; fragments of ~150bp were purified from gels, end-repaired and specific adaptors - containing barcodes - were ligated onto both ends of the fragments. In solution target enrichment was performed using the RNA oligonucleotides coupled to magnetic beads (Agilent's SureSelect technology) according to the supplier's protocol. Finally, sequencing beads was prepared from the enriched fragments using the emulsion PCR (ePCR) technology. Throughout the processing quality of the samples was monitored using lab-on-a-chip technology (Agilent Bioanalyzer).

NGS sequencing: sequencing was performed using the double-base encoded ligation technology of SOLiD System yielding 50 nucleotide sequence tag from a single template.

<u>BIOINFORMATICS</u>

Analysis of targeted resequencing data: As first, the low quality sequence reads were trimmed or removed. The high quality sequence tags were mapped to the reference human genome

using the color-space output of the SOLiD System, ensuring high accuracy, and the SNPs existing in the sequenced samples were identified, the location information (exon, intron, UTR, promoter, intergenic, etc.) of the SNPs were collected. At the end of this process a list of SNPs located in the selected channelopathy target regions for all of the investigated controls and affected patients were generated. Next, statistical analysis of the genotype and phenotype data sets were performed to assess the effects of the different polymorphisms on the appearance the disease phenotypes. The systematic collection of diagnostic evidences containing the multiparameter output of sophisticated medical instruments, allowed us to statistically analyze the genetic associations not only with the diagnosed disease itself, but with any of the qualitative or quantitative data element recorded.

4. RESULTS

The results of the experiment can be summarized as follows.

1) The targeted resequencing platform, we developed, is capable to detect known or novel variants with good specificity and sensitivity.

Coverage data showed an exceptionally rich coverage of the target regions with an average coverage of 1977±715 reads, and both >10 read coverage or >20 read coverage was achieved in 98±2 % and 97±6 % of the target regions, respectively.

2) The platform is capable to detect reported or novel variants of known disease causing ion channel genes.

As a validation process, we included some samples as “positive controls”, known to carry mutations in multiple ion channel genes (KCNQ1, KCNH2 and KCNJ2 genes), identified previously by capillary sequencing. All mutations, expect one, was detected by the platform, proving an excellent sensitivity and specificity of the method. Apart of detecting mutations in major LQTS genes (i.e. KCNQ1, KCNH2 and SCN5A) some of the genetic variants we

identified include some real genuine novelties. We found mutations in some exceptionally rare LQTS genes, e.g. in syntrophin (SNTA1) -reported only in 7 families with LQTS or sudden infant death syndrome world-wide to date- or in A kinase anchor protein (yotiao, AKAP9) reported only in three families with LQTS so far. We also identified a canonical Gly406Arg mutation of the CACNA1C gene in an exceptionally rare genetic disorder, Timothy syndrome. The mutation affects exon 8A, defining the case as Timothy syndrome 1 (TS1) on genetic grounds, however, the case is not associated with syndactyly, which is a rule in TS1. These findings raise the possibility for identifying a novel clinical entity.

4) Most importantly, we identified cases, with an unequivocal clinical diagnosis of long QT syndrome, which doesn't carry any variants in the target genes.

These cases provide proof for our concept about the existence of novel disease causing genes and represent ideal candidates for exome sequencing in the later phases of the study.

5) Our preliminary results provide evidence that many patients carry more than one rare variant of multiple ion channel genes which may alter clinical phenotype.

Analyzing genotype-phenotype correlations, we found trend like differences indicating that the presence of multiple variants or variants with a dominant effect (i.e. definite causative mutation) may lead to more severe form of the disease in LQTS patients. The average corrected QT interval (QTc) showed a trend to increase in mutation carriers with multiple variants and the magnitude of QTc prolongation tended to correlate with the number of identified variants. Also, the average age at the time of first symptoms were lower in carriers of a causative mutation or if they carried multiple variants.

5. DISCUSSION

Hence, this pilot experiment highlighted the enormous potential of this approach to identify potential new disease causing genes in channelopathies and provided proof that the platform we developed is capable to find new genetic variants with good specificity and sensitivity.

6. REFERENCES:

1. 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). Europace. 2011 Aug;13(8):1077-109. doi: 10.1093/europace/eur245.

2. Curran ME, Splawski I, Timothy KW et al. A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. Cell 1995;80:795–798.

3. Wang Q, Curran ME, Splawski I et al. Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. Nat Genet 1996; 12:17–23.

4. Wang Q, Shen J, Splawski I et al. SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. Cell 1995;80:805–811.

5. Brugada P, Brugada J. Right bundle branch block persistent ST segment elevation sudden cardiac death: a distinct clinical electrocardiographic syndrome. J Am Coll Cardiol 1992;20:1391–1396.

6. Antzelevitch C, Brugada P, Borggrefe M et al. Brugada syndrome. Report of the Second Consensus Conference. Heart Rhythm 2005;2:429–440.

7. Antzelevitch C, Pollevick GD, Cordeiro JM et al. Loss-of-function mutations in the cardiac calcium channel underlie a new clinical entity characterized by ST-segment elevation, short QT intervals, and sudden cardiac death. Circulation 2007;115:442–449.

8. Coumel P, Fidelle J, Lucet V et al. Catecholaminergic-induced severe ventricular arrhythmias with Adams-Stokes syndrome in children: report of four cases. Br Heart J 1978;40:28–37.

9. Priori SG, Napolitano C, Tiso N et al. Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia. Circulation 2001;103:196–200.

10. Lahat H, Pras E, Olender T et al. A missense mutation in a highly conserved region of CASQ2 is associated with autosomal recessive catecholamine-induced polymorphic ventricular tachycardia in Bedouin families from Israel. Am J Hum Genet 2001;69:1378–1384.

11. Csanády, M. and K. Szász, Familial cardiomyopathy. Cardiology, 1976. 61: p. 122-130.

12. Csanády, M., et al., Familial dilated cardiomyopathy: a worse prognosis compared with sporadic forms. British Heart Journal, 1995. 74: p. 171-173.

13. Csanády, M., M. Hőgye, and T. Forster, Hypertrophic cardiomyopathy associated with congential deaf-mutism. European Heart Journal, 1987. 8: p. 528-534.

14. Csanády, M. and Z. Kiss, Heritable Q-T prolongation without congential deafness (Romano- Ward syndrome). Chest, 1973. 64: p. 359-362.

15. Csanády, M. and Z. Kiss, Az elektrokardiogram QT-távolságának örökletes megnyúltsága, veleszületett süketség nélkül (Romano-Ward-syndroma). Orv Hetil, 1972. 47: p. 2840-3.

16. Sepp, R., et al., Hypertrophiás cardiomyopathiát okozó génmutáció azonosítása a béta myozin nehéz lánc génben. Az első molekuláris genetikai analízissel igazolt magyar család leírása. Cardiologica Hungarica, 2001. 1: p. 65-70.

17. Sepp, R., et al., Az első KCNQ1-génmutáció azonosítása hosszú QT-szindrómás magyar betegben. Cardiologia Hungarica, 2006. 36: p. 11-6.

18. Sepp, R., et al., Az első hosszú-QT szindrómát okozó génmutáció azonosítása magyar betegben. Cardiologica Hungarica, 2004. 34: p. 184-
19. Hategan L, Katona M, Kornyei L, Racz K, Csanady M, Forster T, Sepp R. Identification of a novel KCNJ2 mutation causing Andersen-Tawil syndrome in a Hungarian patient. Cardiology Letters 2012; 21(3): 266-7.