Long-read Sequence Confirmed a Large Deletion Including *MYH6* and *MYH7* in an Infant of Atrial Septal Defect and Atrial Arrhythmias.

First author: Sonoda,

Short title: Deletion in *MYH6* and *MYH7*confirmed by LRS

Keiko Sonoda MD, PhD,[1,2,3] Haruko Ishihara, MD,[4] Hisanori Sakazaki, MD, PhD,[4] Tsugutoshi Suzuki, MD, PhD,[5] Minoru Horie MD, PhD,[2] Seiko Ohno MD, PhD,[1,2]

1. Department of Bioscience and Genetics, National Cerebral and Cardiovascular Center, Suita, Japan
2. Department of Cardiovascular Medicine, Shiga University of Medical Science, Otsu, Japan.
3. Department of Cardiovascular Biology and Medicine, Niigata University Graduate School of Medical and Dental Sciences
4. Hyogo Prefectural Amagasaki General Medical Center, Amagasaki, Japan
5. Department of Pediatric Electrophysiology, Osaka City General Hospital, Osaka, Japan.

Corresponding author:

Seiko Ohno MD, PhD,

Department of Bioscience and Genetics, National Cerebral and Cardiovascular Center,

Address: 6-1 Kishibe-Shimmachi, Suita, Japan

ZIP:564-8565,

Tel: +81-6-6170-1070, FAX: +81-6-6170-1659

E-mail: sohno@ncvc.go.jp

Total word count:799, Article type: Research letter

Journal Subject Terms: Genomic and Precision Medicine

Key words: long-read sequence, deletion, MYH6, MYH7, atrial septal defect, atrial tachycardia, atrial flatter

Nonstandard Abbreviations and Acronyms

ASD, atrial septal defect

SNVs, single nucleotide variants

SVs, structural variants

SRS, short-read sequencing

LRS, long-read sequencing

AFL, atrial flatter

MYH6, encoding α-cardiac myosin heavy chain, is one of the responsible genes for atrial septal defect (ASD) [1]. Although single nucleotide variants (SNVs) in *MYH6* have been reported as the cause of ASD, there are few literatures describing DNA structural variants (SVs) which are defined as more than 50 bp regions of DNA showing deletions, insertions, duplications, inversions or translocations. Short-read sequencing (SRS) can detect SNVs with high accuracy but SVs with lower sensitivity[2], therefore, it might overlook SVs. The read length feasible by long-read sequencing (LRS) is tens to thousands of kilobases, and LRS can detect the breakpoints of complex SVs [3] .In this study, we confirmed a large deletion extending from *MYH6* to *MYH7* including the breakpoints in a family with inherited ASD, using both SRS and LRS.

The proband was a five-month-old boy with ASD detected by echocardiography (Fig.A), atrial flutter (AFL) and tachycardia. He received catheter ablation therapy for AFL at two. Although his voltage map of both atriums showed no myocardial damages, the right atrium was remarkably enlarged (Fig. B). His mother had ASD which was closed spontaneously. There were no findings indicating cardiomyopathy in the proband and his mother.

After obtaining the consent from his parents for the genetic analysis approved by our institutional review board, we performed targeted SRS for 58 genes related to inherited cardiac arrhythmias. The proband's SRS did not identify any responsible SNVs. To detect deletions and duplications, we compared the read depth between the proband and control sample (pair analysis) using SureCall software (v. 2.1.2.11) (Agilent, Santa, CA, USA). The log ratios from exon 26 to exon 3 in *MYH6* were lower than that of the normal range, which meant the deletion of these exons (Fig.C). Because a deletion between *MYH6* intron 25 and *MYH7* intron 26 was reported as esv2748480 and nsv4228245 on dbVar (https://www.ncbi.nlm.nih.gov/dbvar/), we performed long-

range PCR from *MYH6* intron 26 to *MYH7* exon 26 and found an abnormal PCR product with the size of 1.5K bases in the proband and his mother. The sequence from chr14:23390033 to 23390448 including *MYH6* exon 26 is identical to chr14:23419821 to 23420236 including *MYH7* exon 27 (Fig. D, yellow box indicating the homologous part). Although we designed primers at the specific sequence in each gene, it is difficult to read the sequence over the homologous part at once and we could not confirm the deletion.

To resolve this problem, we performed whole-genome LRS with the Nanopore sequencer GridION X5 (Oxford Nanopore Technologies, Oxford, UK). The long reads were aligned to the human reference genome sequence (build hg38) by NGMLR (v.0.2.7) [4]. Sniffles (v.1.0.11) [4] , calling any types of SVs, detected a deletion encompassing from chr14: 23390037 to 23419824 (hg38) (Fig. D and Fig. E). We reconfirmed the result of Sanger sequencing to determine the precise positions connecting the homologous part and the specific part of each gene. Based on 3' rule, we concluded that a range of the deletion was from chr14:23390449 to 23420236, which produced the same DNA rearrangement as the deletion detected by Sniffles (Fig. D).

This is the first report showing a large deletion between *MYH6* and *MYH7* in a family with ASD. Several missense and truncating variants in *MYH6* have been reported as responsible for ASD. In animal models, morpholino knock-down of expression of the chick *MYH6* homolog eliminates the formation of the atrial septum without overtly affecting atrial chamber formation. The variants in *TBX5* and *GATA4*, which are associated with expression of *MYH6*, have been reported as causative for ASD. These literatures suggested *MYH6* downregulation would be associated with ASD. In our patient, a new hybrid protein consisting of *MYH6* and *MYH7* might be produced because this deletion is in-frame, and hybrid protein was possibly under control of the

MYH7 promoter unaffected by the deletion. It may lead the expressions of reduced normal *MYH6* and of unbalanced *MYH6*/*MYH7* in the atria. Unfortunately, the cardiac tissue of the proband was not available and we could not confirm our hypothesis.

Nsv4228245 and esv2748480 are almost same as our deletion. Although esv2748480 was found in 4 participants in 100 Malays, the cohort was small, and 4 might be relatives. In gnomAD, nsv4228245 was found in only 2 of 20310 alleles (MAF= 9.8e-005) [5] and is enough rare to be considered as a clinically relevant variant. Moreover, most patients with ASD are asymptomatic, and ASD sometimes close spontaneously, therefore the participants with these SVs might have ASD latently.

In this study, we identified a large deletion using both SRS and LRS methods. As in our case, SVs including high homology sequence would be one of the suitable targets of Nanopore sequencing. The combination of SRS and LRS is useful to confirm the detail of SVs in patients with suspected inherited diseases but carrying no causative SNVs.

Sources of Funding:

This work was supported in part by grant from the Naito foundation to S.O.

Disclosures:

None.

Data, Materials, and Code Disclosure

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics approval:

The study was approved by the ethics committee of Shiga Medical University and followed the principles of the Declaration of Helsinki (Article II).

References

1. Geva T, Martins JD and Wald RM. Atrial septal defects. *Lancet*. 2014;383:1921-32.

2. Tattini L, D'Aurizio R and Magi A. Detection of Genomic Structural Variants from Next-Generation Sequencing Data. *Front Bioeng Biotechnol*. 2015;3:92.

3. Mitsuhashi S and Matsumoto N. Long-read sequencing for rare human genetic diseases. *J Hum Genet*. 2020;65:11-19.

4. Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, Haeseler Av and Schatz MC. Accurate detection of complex structural variations using single molecule sequencing. 2017.

5. Collins RL, Brand H, Karczewski KJ, Zhao X, Alfoldi J, Francioli LC, Khera AV, Lowther C, Gauthier LD, Wang H, et al. A structural variation reference for medical and population genetics. *Nature*. 2020;581:444-451.

Figure legends

Figure. Clinical data and the results of genetic analysis of the patient.

A. Transthoracic echocardiography showing atrial septal defect, which was categorized to ostium secundum (orange arrow). His pulmonary blood flow/systemic blood flow ratio (Qp/Qs) was 2.3.

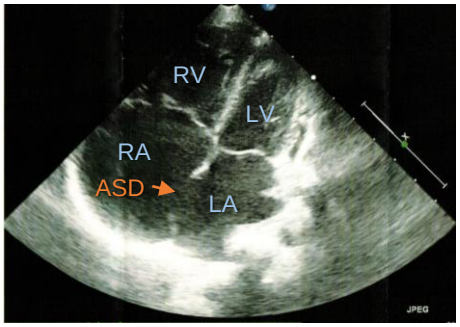
B. The result of CARTO electroanatomical mapping in the both atriums. There was no significant low voltage area. The right atrium was enlarged.

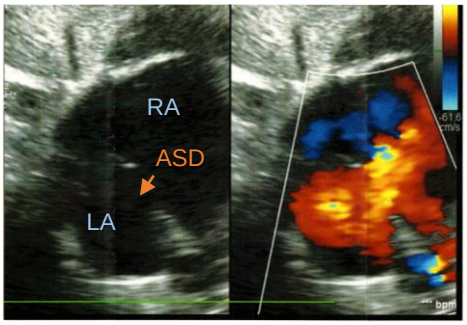
C. The result of targeted short-read sequencing analyzed by pair analysis (SureCall software). The blue boxes indicated by blue arrow show regions with lower log ratios than that of the normal range. The log ratio means log value of the ratio of mapped read counts in the proband to those in the control. Gray columns surrounded by light green represent read depths of the proband and the control. The regions with lower log ratio and decreased read depths are found in between exon 26 and exon 3 in *MYH6*.

D. The result of whole-genome sequencing performed by nanopore sequencer. IGV screen shot of reads at *MYH6* and *MYH7*. The gray box shows a range of the deletion, from chr14:23390037 to 23419824.

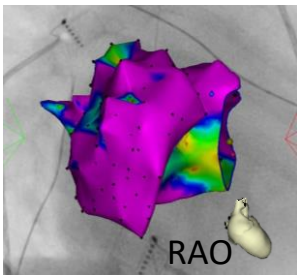
E. The detail of the large deletion. Upper is control and lower is the patient. Yellow box indicates the homologous part.

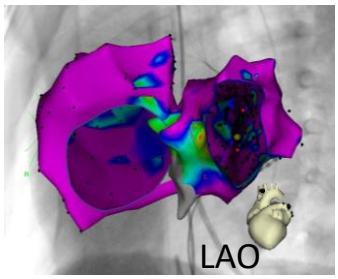
A

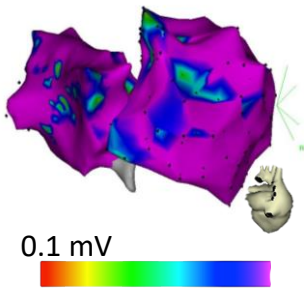

RV
LV
RA
ASD
LA


RA
ASD
LA

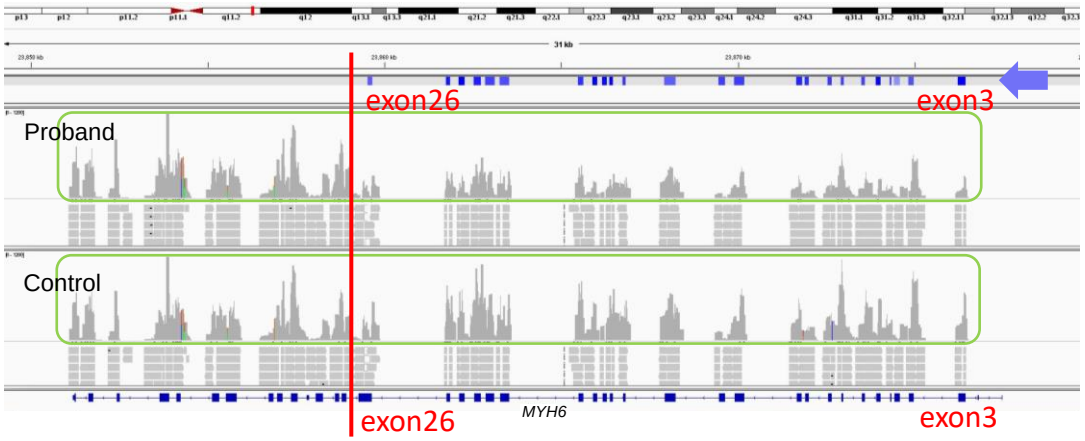
B


RAO

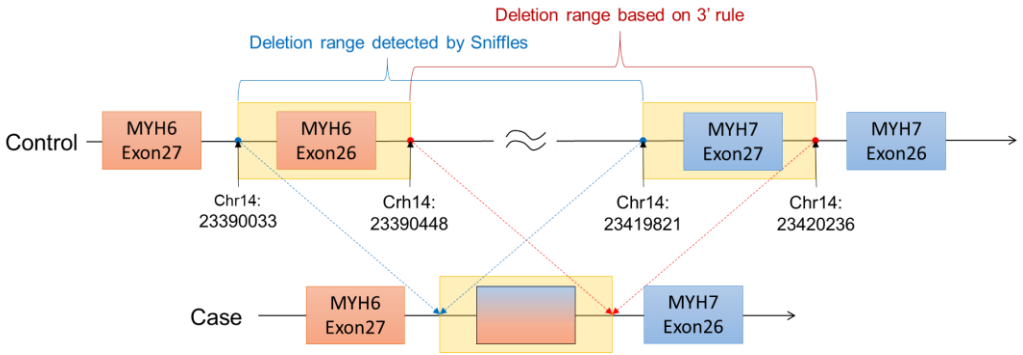

LAO


0.1 mV

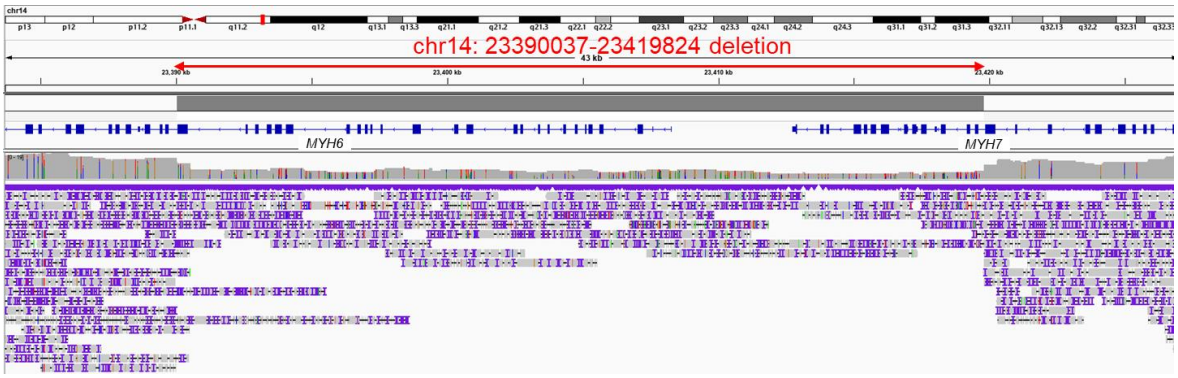
C


exon26
exon3
Proband
Control
exon26
MYH6
exon3

D


Deletion range based on 3' rule
Deletion range detected by Sniffles
Control
MYH6 Exon27
MYH6 Exon26
MYH7 Exon27
MYH7 Exon26
Chr14: 23390033
Crh14: 23390448
Chr14: 23419821
Chr14: 23420236
Case
MYH6 Exon27
MYH7 Exon26

E


chr14: 23390037-23419824 deletion
MYH6
MYH7