ONLINE RESOURCES

Development and diversity of a novel panel of short tandem repeat markers encompassing the *SCN5A* gene in Iranian population

Zahra Zafari [a] Azam Amirian[b], Faezeh Rahimi Nejad[c], Vahid Akbari[d], Mohammad Taghi Akbari [d], Sirus Zeinali [b,c]

[a] Ph.D. Student, Department of Medical Genetics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran

[b] Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran

[c] Medical Genetics Lab, Kawsar Human Genetics Research Center, Tehran, Iran

[d] Department of Medical Genetics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran

Correspondence: Dr. Sirus Zeinali, Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran.

E-mail: Zeinalipasteur@yahoo.com

Abstract

Background: The *SCN5A* gene plays a key role in a variety of heterogeneous cardiac diseases such as congenital Long-QT syndrome, Brugada syndrome and sudden cardiac death. The substantial utility of highly polymorphic STR markers in forensic and diagnosis purposes prompted us to develop and validate a panel of six novel STR markers encompassing the *SCN5A* gene.

Methods: Allele frequencies and forensic statistics of six tetra nucleotide tandem repeat markers identified by TRF and SERV programs and amplified

in a six-plex PCR system were calculated in 60 unrelated Iranian healthy individuals.

Results and discussion: Fragment analysis revealed 6-10 alleles for six STR markers with the observed heterozygosity of greater than 0.667 for five markers. The power of discrimination was more than 0.83 for the panel.

Conclusion: This novel panel of six polymorphic STR markers with the high level of heterozygosity and discrimination in each locus may help to establish the rapid and more reliable identification of the disease causative *SCN5A* gene and provide aid to forensic purposes and prenatal diagnosis.

Keywords: Cardiac disease, Forensic, STR marker, Allele frequency, Population data, Iran

Introduction

The *SCN5A* gene plays a key role in propagation of cardiac action potential and is responsible for rapid depolarization of inward sodium current in cardiomyocytes by encoding the α-subunit of the voltage-gated sodium channel, $Na_v1.5$ (Hamosh et al. 2005).

Gain or loss of function mutations in *SCN5A* gene lead to a variety of heterogeneous cardiac diseases (Zaklyazminskayaa and Dzemeshkevicha 2016) such as congenital Long-QT syndrome (LQTS) in which the *SCN5A* gene is the third major gene among 16 genes accounting for the genetic-positive LQTS (Tester and Ackerman 2014; J.Reed et al. 2015). Brugada syndrome (BrS) is another arrhythmic disorder in which 18 causative genes were reported and the percentage of *SCN5A* gene variants was attributed to 11-28% (Sieira et al. 2016).

SCN5A is also included in the most commonly mutated genes in dilated cardiomyopathy (DCM) with over 60 linked genes (Haas et al. 2014).Autopsy negative sudden cardiac death (SCD) accompanied by no previous manifestations is resulted from the *SCN5A* gene mutations as a major component (Tester et al. 2012; Kauferstein et al. 2013; Hertz et al. 2015). Atrial fibrillation (Darbar et al. 2008), Sick sinus syndrome (Benson et al. 2003) and idiopathic ventricular fibrillation (Chen et al. 1998) are some other overlapping syndromes associated with SCN5A gene mutations.

SCN5A consists of 27 coding Exons spanning approximately 80 kb (Hamosh et al. 2005). More than 500 pathogenic and likely pathogenic variants have been submitted for the cardiac sodium channel gene in ClinVar (Landrum et al. 2016); therefore mutated gene diagnosis for such genetic heterogeneous diseases is time consuming, unprofitable and additionally requires much DNA sample.

Haplotype analysis with short tandem repeat (STR) markers flanking the *SCN5A* gene can help to overcome the limitations. Furthermore, the utility of highly polymorphic markers in forensic purposes, homozygosity mapping, the prenatal genetic diagnosis (PND) and pre-implantation genetic diagnosis (PGD) approaches are previously advocated (Butler 2007; Zupanič Pajnič et al. 2010).

In the current study, we identified six novel tetranucleotide STR markers amplifiable in one multiplex-PCR reaction to rapidly identification of the disease causative *SCN5A* gene in a family represented with the above cardiac diseases. We also presented the heterozygosity and frequency assessments of these STR markers in Iranian population.

Materials and methods

DNA Preparation

Blood samples were collected from 60 unrelated healthy individuals from Iranian population. Informed consents were obtained from all participants. This study was approved by the ethics committees of Pasteur Institute of Iran (adopted from the 1975 Helsinki Declaration). Genomic DNA was isolated from peripheral blood according to the standard salting out protocol (Miller et al. 1988)

Markers identification

Five tetra-nucleotide tandem repeat markers flanking the *SCN5A* gene and one located inside the gene (Fig. 1) were identified and selected by USCS genome browser and TRF (G. Benson 1999) and SERV programs (Legendre M et al. 2007).

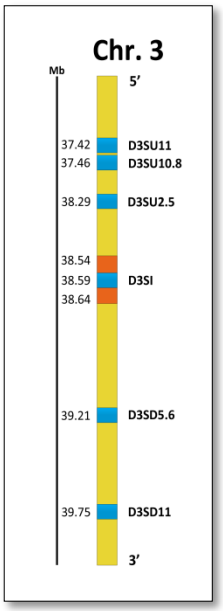

Figure 1. Relative positions of six STR loci encompassing the SCN5A gene on chromosome 3. The *SCN5A* gene (orange color) is located on the short region of chromosome 3 from 38.54 Mb to 38.64 Mb and D3SI (blue color) is an Intragenic marker. The numbers in the markers names show the distance between the gene and the flanking marker divided by 10^5 (*i.e.* 11 in D3SU11 means a distance of 1'100'000 bp from the gene). "U" in the markers names stand for upstream of the gene, "I": inside the gene and "D": downstream of the gene.

Markers amplification and visualization

Six primer pairs were designed and fluorescently labelled with either FAM, VIC or PET dyes to amplify these STR markers in one multiplex-PCR reaction (Table 1). PCR amplification was performed in a 17-µL reaction volume containing 2.8 µl 10X Buffer, 3.94 mM $MgCl_2$, 3.2 mM dNTP, 1µl Bovin Serum Albumin, 5 unit/µl KBC Taq-Plus DNA Polymerase in addition to about 200 ng of genomic DNA.

Thirty five amplification cycles were carried out in 95 ºC for 5 min, 95 ºC for 1 min, 63 ºC for 1 min and 30 s, 70ºC for 2 min, 70 ºC for 17 min. The fragments were separated on ABI 3130 Genetic Analyzer and GS500 LIZ was used as size standard. Data were analysed by GeneMapper ID software.

Statistical assessments

The allelic frequencies, observed and expected heterozygosity, probability of identity (PI) and the power of exclusion (PE) were calculated by the utility of GenAlEx 6.502 software. The power of discrimination (PD) for each locus was calculated by direct counting method using the formula PD=1-PI. The deviation from Hardy-Weinberg equilibrium (HWE) along with the

polymorphism information content (PIC) was evaluated by Cervus 3.0.7 program.

Results and discussion

All the previously reported STR markers for the SCN5A gene were dinucleotide repeats but we used TRF and SERV programs for identification of tetranucleotide tandem repeats to eliminate the minor bands and to reduce the stutter bands during PCR amplification (Butler 2007; Rabbani et al. 2008). The maximum interval of 1.1 mega bases around the SCN5A gene was specified to the tetranucleotide STR markers selection to minimize the meiosis recombination occurrence.

Fragment analysis of six novel STR markers revealed 6, 7, 7, 5, 10 and 8 alleles for D3SU11, D3SU10.8, D3SU2.5, D3SI, D3SD5.6 and D3SD11 respectively. Allele frequencies and the Iranian population data for STR markers are shown in Table 2. The mean number of alleles over all loci was 7.72. Five out of six Loci had the observed heterozygosity (Ho) of greater than 0.667. The highest Ho was 0.917 at D3SD5.6 and the lowest Ho was 0.433 at D3SU10.8 marker. The PD for all loci was more than 0.83. Despite the great extent of consanguineous marriages in Iranian population, the mean number of PIC for the panel was 0.68 indicating that all loci were highly informative. There were no significant deviations from HWE expectations for Loci examined within the Iranian population except for D3SU10.8 and D3SD5.6. These two markers displayed the probability of less than 0.0002 when the deviations from HWE were re-examined for them by GeneAlEx software. Such deviations can be explained by either probable genotyping and laboratory errors or population stratification.

In conclusion, this novel panel of six polymorphic STR markers with the high level of heterozygosity and discrimination in each locus may help to establish the rapid and more reliable *SCN5A* gene identification in heterogeneous cardiovascular diseases by haplotype analysis provided that it is the disease causative gene. The panel is further an efficient tool in forensic, prenatal diagnosis, preimplantation of genetic diagnosis (PGD) and homozygosity mapping especially in the Autosomal recessive diseases and in endogamous populations with the high rate of the consanguineous marriages.

Acknowledgement:

This research was supported by Pasteur Institute of Iran (grant number 824).

Conflict of interest

There is no conflict of interest.

Benson, D.W., Wang, D.W., Dyment, M., Knilans, T.K., Fish, F.A., Strieper, M.J., Rhodes, T.H., and George, A.L. 2003. Congenital sick sinus syndrome caused by recessive mutations in the cardiac sodium channel gene (SCN5A). *Journal of Clinical Investigation* **112**(7): 1019-1028.

Butler, J.M. 2007. Short tandem repeat typing technologies used in human identity testing. *BioTechniques* **43**(4): Sii-Sv.

Chen, Q., Kirsch, G.E., Zhang, D., Brugada, R., Brugada, J., Brugada, P., Potenza, D., Moya, A., Borggrefe, M., Breithardt, G. et al. 1998. Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. *Nature* **392**(6673): 293-296.

Darbar, D., Kannankeril, P.J., Donahue, B.S., Kucera, G., Stubblefield, T., Haines, J.L., George, A.L., and Roden, D.M. 2008. Cardiac Sodium Channel Variants Associated with Atrial Fibrillation. *Circulation* **117**(15): 1927.

G. Benson. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**(2): 573-580.

Haas, J., Frese, K.S., Peil, B., Kloos, W., Keller, A., Nietsch, R., Feng, Z., Müller, S., Kayvanpour, E., Vogel, B. et al. 2014. Atlas of the clinical genetics of human dilated cardiomyopathy. *European Heart Journal* **36**(18): 1123.

Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A. 2005. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Research* **33**(Database Issue): D514-D517.

Hertz, C.L., Christiansen, S.L., Ferrero-Miliani, L., Fordyce, S.L., Dahl, M., Holst, A.G., Ottesen, G.L., Frank-Hansen, R., Bundgaard, H., and Morling, N. 2015. Next-generation sequencing of 34 genes in sudden unexplained death victims in forensics and in patients with channelopathic cardiac diseases. *International Journal of Legal Medicine* **129**(4): 793-800.

J.Reed, G., J.Boczek, N., P.Etheridge, S., and J.Ackerman, M. 2015. CALM3 mutation associatedwithlongQTsyndrome *Heart Rhythm* **12**(2): 419-422.

Kauferstein, S., Kiehne, N., Peigneur, S., Tytgat, J., and Bratzke, H.r. 2013. Cardiac channelopathy causing sudden death as revealed by molecular autopsy. *International Journal of Legal Medicine* **127**(1): 145-151.

Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., and Hoover, J. 2016. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Research* **44**(D1): D862-D868.

Legendre M, Pochet N, Pak T, and Verstrepen KJ. 2007. Sequence-based estimation of minisatellite and microsatellite repeat variability. *Genome Res* **17**(12): 1787-1796.

Miller, S.A., Dykes, D.D., and Polesky, H.F. 1988. A simple salting out procedure forextracting DNA from human nucleated cells. *Nucleic acids Res* **16**(3): 1215.

Rabbani, B., Khanahmad, H., Bagheri, R., Mahdieh, N., and Zeinali, S. 2008. Characterization of minor bands of STR amplification reaction of FVIII gene by PCR cloning. *Clinica Chimica Acta* **394**(1-2): 114-115

Sieira, J., Dendramis, G., and Brugada, P. 2016. Pathogenesis and management of Brugada syndrome. *Nat Rev Cardiol* **advance online publication**.

Tester, D.J. and Ackerman, M.J. 2014. Genetics of long QT syndrome. *Methodist DeBakey Cardiovascular Journal* **10**(1): 29-33.

Tester, D.J., Medeiros-Domingo, A., Will, M.L., Haglund, C.M., and Ackerman, M.J. 2012. Cardiac Channel Molecular Autopsy: Insights From 173 Consecutive Cases of Autopsy-Negative Sudden Unexplained Death Referred for Postmortem Genetic Testing. *Mayo Clinic Proceedings* **87**(6): 524-539.

Zaklyazminskayaa, E. and Dzemeshkevicha, S. 2016. The role of mutations in the SCN5A gene in cardiomyopathies. *Biochimica et Biophysica Acta* **1863**(7): 1799–1805.

Zupanič Pajnič, I., Gornjak Pogorelc, B., and Balažic, J. 2010. Molecular genetic identification of skeletal remains from the Second World War Konfin I mass grave in Slovenia. *International Journal of Legal Medicine* **124**(4): 307-317.

Unedited Version Published Online 11 January 2018

Table 1. Six STR markers linked to the SCN5A gene

Marker	Consensus sequence	Repeat Structure	PCR Primer Sequences	Dye label	Size Range
D3SU11	GATA	$[GATA]_2[GAC]_1[GATA]_{12}[G]_1$	5'- TGACTTTGGCTTTACGGAGGTAG -3' 5'- GCATGTCCTTAAGGAACTTGAGTTA -3'	VIC	299-318
D3SU10.8	AAAT	$[AAAT]_{12}[AA]_1$	5'- AGCAGTGCAGCAAGAATTTCTC -3' 5'- CAATAGATAACTGAAACACTTGGCAT -3'	PET	134-161
D3SU2.5	ATCT	$[ATCT]_{12}[AATCT]_1[ATCT]_1[ATC]_1$	5'- TGGACAGAGGTAGAAAATAGACTTGC -3' 5'- CCACAGTCTTCATTCTAGTCTTCCC -3'	VIC	143-167
D3SI	ATCC	$[ATCC]_2[ATCT]_1[GTCC]_1[ATCC]_9[CTCC]_1[ATCC]_1[TTCC]_1[CTCC]_1[TTCT]_1[ATCC]_1[AACC]_1[ATCC]2[A]_1$	5'- GCTCTGCCTGATTTACTTACTACACC -3' 5'- CTCCACCTACAAATTACAAAATCTCAA -3'	PET	220-236
D3SD5.6	AAAG	$[AAAG]_{18}[AAG]_1[AAAG]_3[AAA]_1[AAAG]_{12}$	5'- ACAAATTTGGAGCTGATTCTTAACTG -3' 5'- CATCAGAAAGAGAACTAGGAGGAATCT -3'	FAM	239-281
D3SD11	CTAT	$[CTAT]_{11}$	5'- TTATCTTCTTTAAGACAATAGTGCCATG -3' 5'- CCAGATAGCCGTTAGTATGTTAGATAATG -3'	PET	334-357

Table 2: Population data of six novel STR markers

Allele number	D3SU11		D3SU10.8		D3SU2.5		D3SI		D3SD5.6		D3SD11	
	Allele Size	Frequency	Allele Size	Frequency	Allele Size	Frequency	Allele Size	Frequency	Allele Size	Frequency	Allele Size	Frequency
1	299	0.033	134	0.042	143	0.008	220	0.100	239	0.008	334	0.008
2	303	0.325	140	0.458	147	0.017	224	0.258	247	0.083	337	0.158
3	307	0.358	145	0.042	151	0.117	228	0.333	252	0.167	341	0.467
4	310	0.167	149	0.017	155	0.500	232	0.283	256	0.133	345	0.275
5	314	0.083	153	0.200	159	0.267	236	0.025	260	0.167	348	0.008
6	318	0.033	157	0.225	163	0.083			264	0.225	349	0.067
7			161	0.017	167	0.008			269	0.125	353	0.008
8									273	0.067	357	0.008
9									277	0.017		
10									281	0.008		
Na	6		7		7		5		10		8	
He	0.729		0.695		0.658		0.731		0.849		0.677	
Ho	0.750		0.433		0.667		0.683		0.917		0.683	
PI	0.12		0.14		0.17		0.12		0.04		0.16	
PD	0.88		0.86		0.83		0.88		0.96		0.84	
PIC	0.684		0.651		0.609		0.682		0.830		0.626	
PE	0.678		0.647		0.599		0.660		0.864		0.612	
HWE Expected P-value	0.59		<0.0002		0.88		0.94		<0.0002		0.86	