Characterization and haplotype study of 6 novel STR markers related to the KCNQ1 gene in heterogeneous cardiovascular disorders in the Iranian population

Azam AMIRIAN1, Zahra ZAFARI1,2, Zohreh SHARIFI3, Alireza KORDAFSHARI1, Morteza KARIMPOOR1, Sirous ZEINALI1,3,*

1Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran
2Department of Biology, Shahed University, Tehran, Iran
3Medical Genetics Laboratory, Kawsar Human Genetics Research Center, Tehran, Iran

Received: 06.05.2018 • Accepted/Published Online: 15.09.2018 • Final Version: 18.04.2019

Abstract: Background/aim: The KCNQ1 gene has a significant role in long QT syndrome, Jervell and Lange-Nielsen syndrome, familial atrial fibrillation, and short QT syndrome. Analyzing such heterogeneous disorders, six novel short tandem repeat (STR) markers around the KCNQ1 gene were found and evaluated in a healthy population, and other statistical traits of the markers were detected.

Materials and methods: Using Tandem Repeats Finder (TRF) and Sequence-Based Estimation of Repeat Variability (SERV) software, STR markers were detected with valid tetra- and pentanucleotide repeats. The markers were investigated for a total of 60 unrelated Iranian healthy individuals and analyzed using GenAlEx 6.502 and Cervus 3.0.7.

Results: A total of 77 haplotypes was detected, of which 25 haplotypes were unique and the others occurred at least two times. The number of haplotypes per locus ranged from 7 to 18 with the highest frequency of 69.2%, and the observed heterozygosity was calculated as 0.589. The power of discrimination ranged from 0.70 to 0.96. Five of the markers meet Hardy–Weinberg equilibrium.

Conclusion: A novel panel of STR markers was described with high allele heterozygosity and segregation in every locus, which may lead to faster and more credible recognition of the disease-inducing KCNQ1 gene and allow it to be used for human identity testing and prenatal diagnosis.

Keywords: Allele heterozygosity, cardiac disease, haplotype, Iran, STR marker

1. Introduction
The KCNQ1 gene is related to a huge family of genes that encode heart potassium channel protein, the most diverse group of ion channels, that is responsible for the repolarization phase of the cardiac action potential in the voltage dependence of activation (1). The gene product is assumed to be capable of forming a heteromultimer with the other potassium channel protein, mink (encoded by KCNE1). Loss of function mutations in the KCNQ1 gene, which induce type 1 long QT syndrome (LQTS1, MIM#192500) (2), the most common type of LQTS, cause delayed rectifier potassium current (IKs) in the cardiomyocytes (3) and inner ear (4).

In addition to LQTS, mutations in this gene are also associated with other forms of inherited arrhythmias such as Jervell and Lange-Nielsen syndrome (JLNS), familial atrial fibrillation (AF), and short QT syndrome (SQTS) (5). The gene is located in a region of chromosome 11 that contains a large number of contiguous genes, consisting of 16 coding exons spanning approximately 400 kb (6). Hundreds of different mutations with variable effects on Kv7.1 function have been reported (7). This reveals the considerable clinical importance in LQTS, so linkage studies are the best approach to detection of the mutation for such genetic heterogeneous diseases.

Short tandem repeats (STRs) are highly polymorphic markers and are found in most genomes used in linkage studies (8–11). STRs could also be used for forensic applications, phylogenetic reconstruction, preimplantation genetic diagnosis, and prenatal genetic diagnosis (6,12).

This study set out to explore the utility of STRs in mutated gene diagnosis for LQTS by presenting the six novel tetra- or pentanucleotide STR markers surrounding the KCNQ1 gene. Heterozygosity and frequency evaluation of these markers has been carried out in the Iranian population.
2. Materials and methods

2.1. DNA extraction

Sixty unrelated healthy individuals were selected from the Iranian population. After obtaining informed consent, blood samples were collected in tubes containing ethylenediaminetetraacetic acid (EDTA). The research was approved by the Pasteur Institute of Iran and the local ethics committee. DNA was extracted from blood samples using the KBC Blood-DNA Extraction Kit.

2.2. STR marker preparation

The University of California Santa Cruz (UCSC) genome browser, Tandem Repeats Finder (TRF) (13), and SERV (Sequence-Based Estimation of Repeat Variability) sequence programs (14) were used for finding the STR markers. A pentanucleotide, D11SD8.3, and four tetranucleotide (D11SU10.9, D11SU2.2, D11SU0.6, D11SD13.6) tandem repeat markers flanking the KCNQ1 gene, plus one tetranucleotide, D11SI, located inside the gene, were selected.

2.3. Marker primer design and amplification

Primers were designed using Gene Runner software. We used the multiplex PCR method, which amplifies multiple DNA fragments in one polymerase chain reaction (PCR). Each forward primer was labeled with fluorescent-dye labels (either FAM, VIC, or NED dye). STR markers, primer sequences, and their fluorescent-dye labels are shown in Table 1.

Table 1. Characteristics of 6 STR markers developed for KCNQ1 gene.

| STR loci  | Primer sequences (5′–3′) | Dye label | Repeat motifs | Product sizes |
|-----------|--------------------------|-----------|---------------|--------------|
| D11SU10.9 | F: CACCCGTCTCTGTCGATTC  | VIC       | (TCCA)₅      | 324–341      |
|           | R: TGTGGAGAAGTTAAGTGATTGTG |           |               |              |
| D11SU2.2* | F: TAGATAGATGCACAGAGATCAAAG | NED       | (ATGG)₁₁     | 303–346      |
|           | R: GGTCTTTCTCTCTAGTTTCCCCCTCC |           |               |              |
| D11SU0.6  | F: TCTCCAGCTTGACAGACAGACAG | FAM       | (TCCA)₁₂     | 165–189      |
|           | R: GTAAGAGATACTGAGTGAGATCGATC |           |               |              |
| D11SI     | F: GTACAGAAAGTGCTCCCATCCAC | VIC       | (TGGA)₁₃     | 82–123       |
|           | R: GAGGCTTTGGTGAGATCGATC |           |               |              |
| D11SD8.3**| F: AACACACCAGGGATGAGTTTCC | NED       | (TCATO)₁₀    | 279–320      |
|           | R: ATTTGTCTCGGGCTCAGTTGAC |           |               |              |
| D11SD13.6 | F: ATCTTTGTAGTATATTGTTTGGAG | VIC       | (TCCA)₁₁     | 206–247      |
|           | R: ACAGACTGACTGAAATGGAAGATGG |           |               |              |

*The number shows the distance from the gene as 2.2 = 220 kb from the 5′ end of the gene or D8.3 being 830 kb from the 3′ end of the gene.

**STRs downstream from the genes were given a name beginning with D, those upstream were given a name beginning with U, and those that were intragenic were given a name beginning with I.
D11SU2.2, D11SU10.9) and downstream (D11SD13.6, D11SD8.3) of the gene; thus, an intragenic marker (D11SI) was selected.

The Iranian population’s allele frequencies and genetic analysis data (statistical characteristics) for STR markers are shown in Table 3 and Table 4, respectively.

4. Discussion
In total, 77 alleles were detected and the average number of alleles per locus was 12.8. For the D11SU10.9, D11SU2.2, D11SU0.6, D11SI, D11SD8.3, and D11SD13.6 STR markers, fragment assay showed 11, 18, 7, 12, 16 and 13 alleles, respectively. The D11SU2.2 locus with a

Table 2. Multiplex amplification conditions (cycling and optimized PCR conditions).

| Optimized PCR conditions          | Cycling and temperature condition |
|-----------------------------------|-----------------------------------|
| MgCl₂ (50 mM)                     | 3.94 mM 95 °C 5 min              |
| 10X PCR buffer                    | 2.8 µL 95 °C 1 min               |
| dNTP (10 mM)                      | 3.2 mM 63 °C 90 s                |
| Distilled water                   | 8 µL 70 °C 1 min                 |
| Forward & reverse primer mix      | 1 µL (1 µL of each 6 pair primers) 35 cycle |
| Bovine serum albumin              | 1 µL 70 °C 17 min                |
| SmarTaq DNA polymerase            | 1 µL Hold 4 °C*                  |
| Genomic DNA                       | 1 µL (200 ng/µL)                 |

* Temperature range (°C).

Table 3. Distribution of observed allele frequencies of 6 STR loci located in the flanking regions of the KCNQ1 gene (n = 60).

| Allele number | D11SU10.9 | D11SU2.2 | D11SU0.6 | D11SI | D11SD8.3 | D11SD13.6 |
|---------------|-----------|----------|----------|-------|----------|-----------|
|               | Allele size | Frequency | Allele size | Frequency | Allele size | Frequency | Allele size | Frequency | Allele size | Frequency | Allele size | Frequency |
| 1             | 326        | 0.058    | 295       | 0.017 | 165       | 0.025     | 82         | 0.117     | 279        | 0.025     | 120         | 0.008     |
| 2             | 330        | 0.033    | 302       | 0.008 | 173       | 0.175     | 85         | 0.042     | 284        | 0.225     | 206         | 0.017     |
| 3             | 331        | 0.067    | 303       | 0.008 | 176       | 0.008     | 86         | 0.025     | 287        | 0.017     | 216         | 0.300     |
| 4             | 332        | 0.008    | 304       | 0.050 | 177       | 0.692     | 104        | 0.017     | 289        | 0.250     | 219         | 0.142     |
| 5             | 333        | 0.017    | 305       | 0.008 | 181       | 0.067     | 102        | 0.008     | 292        | 0.117     | 220         | 0.208     |
| 6             | 335        | 0.517    | 306       | 0.008 | 185       | 0.025     | 105        | 0.008     | 294        | 0.100     | 221         | 0.008     |
| 7             | 336        | 0.017    | 308       | 0.025 | 189       | 0.008     | 107        | 0.008     | 297        | 0.042     | 223         | 0.083     |
| 8             | 339        | 0.250    | 310       | 0.008 | 111       | 0.200     | 299        | 0.008     | 224        | 0.092     | 224         | 0.092     |
| 9             | 342        | 0.017    | 311       | 0.017 | 115       | 0.400     | 302        | 0.100     | 227        | 0.100     | 231         | 0.008     |
| 10            | 343        | 0.008    | 312       | 0.075 | 119       | 0.150     | 306        | 0.050     | 231        | 0.008     | 316         | 0.008     |
| 11            | 352        | 0.008    | 314       | 0.200 | 123       | 0.017     | 307        | 0.017     | 247        | 0.017     | 316         | 0.008     |
| 12            | 315        | 0.375    |           |       | 215       | 0.008     | 308        | 0.008     | 316        | 0.008     | 320         | 0.008     |
| 13            | 317        | 0.008    |           |       |           |           | 311        | 0.008     | 320        | 0.008     |               |
| 14            | 319        | 0.083    |           |       |           |           | 312        | 0.008     |               |
| 15            | 321        | 0.017    |           |       |           |           | 316        | 0.017     |               |
| 16            | 323        | 0.042    |           |       |           |           | 320        | 0.008     |               |
| 17            | 342        | 0.008    |           |       |           |           |               |
| 18            | 345        | 0.042    |           |       |           |           |               |
frequency of 18 alleles was the most polymorphic STR marker. The observed heterozygosity (Ho) was between 0.48 at D11SD8.3 and 0.70 at the D11SI marker. The most frequently detected alleles were 177, 115, and 315 bp for the D11SU0.6, D11SI, and D11SU2.2 markers. The highest PI was 0.30 at D11SU0.6, and the lowest was 0.04 at the D11SD8.3 locus. GenALEX 6.502 showed that the PE was 0.872 and 0.450 at the D11SD8.3 and D11SU0.6 loci, respectively. The power of discrimination was 0.70 at D11SU0.6 and more than 0.92 at all of the other loci. Although 5 out of 6 loci had a PIC above 0.731, the D11SD8.3 marker with content of 0.831 was the most informative locus. According to Cervus 3.0, except for the D11SD8.3 locus, all loci were relevant to HWE, with a probability of less than 0.0002. This revealed that none of the loci except for D11SD8.3 had significant deviations from HWE. This may be due to either probable genotyping and laboratory slips or population stratum (20).

Our study’s conclusion revealed the novel STRs that show a high rate of informativity and high degree of variability that make STR markers very efficient for haplotype analysis and human identity testing. This method is easy to use for detecting the multiplex reproduction pattern, and also a cost-effective method for detection of disease-causative genes and the prenatal diagnosis of heterogeneous cardiovascular diseases compared to direct sequencing, which is usually time consuming and too expensive.

Acknowledgment
This research was supported by the Pasteur Institute of Iran (Grant Number 824 partially and PhD grant from Education Office, Pasteur Institute of Iran).

### Table 4. Statistical characteristics of 6 STR loci located in the flanking regions of the KCNQ1 gene (n = 60).

|     | Na | He  | Ho  | PI  | PD  | PIC | PE  | HWE expected P-value |
|-----|----|-----|-----|-----|-----|-----|-----|---------------------|
| D11SU10.9 | 11 | 0.661 | 0.60 | 0.15 | 0.85 | 0.620 | 0.635 | 0.014 |
| D11SU2.2  | 18 | 0.799 | 0.60 | 0.06 | 0.94 | 0.780 | 0.830 | 0.20  |
| D11SU0.6  | 7  | 0.485 | 0.533 | 0.30 | 0.70 | 0.450 | 0.450 | 0.88  |
| D11SI     | 12 | 0.761 | 0.70 | 0.09 | 0.95 | 0.731 | 0.757 | 0.55  |
| D11SD8.3  | 16 | 0.847 | 0.48 | 0.04 | 0.96 | 0.851 | 0.872 | <0.0002 |
| D11SD13.6 | 13 | 0.820 | 0.617 | 0.05 | 0.95 | 0.799 | 0.832 | 0.04  |

He: expected heterozygosity, Ho: observed heterozygosity, PI: probability of identity, PD: power of discrimination, PIC: polymorphic information content, PE: power of exclusion, HWE: Hardy–Weinberg equilibrium.

### References

1. Deal KK, England SK, Tamkun MM. Molecular physiology of cardiac potassium channels. Physiological Reviews 1996; 76: 49-67.
2. Larsen LA, Fosdal I, Andersen PS, Kanters JK, Vuust J et al. Recessive Romano-Ward syndrome associated with compound heterozygosity for two mutations in the KVLQT1 gene. European Journal of Human Genetics 1999; 7: 724-728.
3. Ning L, Moss AJ, Zareba W, Robinson J, Rosero S et al. Novel compound heterozygous mutations in the KCNQ1 gene associated with autosomal recessive long QT Syndrome (Jervell and Lange-Nielsen syndrome). Annals of Noninvasive Electrocardiology 2003; 8: 246-250.
4. Priori SG, Schwartz PJ, Napolitano C, Bianchi L, Dennis A et al. A recessive variant of the Romano-Ward long-QT syndrome? Circulation 1998; 97: 2420-2425.
5. Hedley PL, Jørgensen P, Schlamowitz S, Wangari R, Moolman-Smook J et al. The genetic basis of long QT and short QT syndromes: a mutation update. Human Mutation 2009; 30: 1486-1511.
6. Harnois A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Research 2005; 33: D514-D517.
7. Landrum MJ, Lee JM, Benson M, Brown G, Chao C et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Research 2015; 44: D862-D868.
8. Jeffreys AJ. Complex gene conversion events in germline mutation at human minisatellites. Trends in Genetics 1994; 10: 151.
9. Butler JM. Forensic DNA Typing: Biology & Technology Behind STR Markers. San Diego, CA, USA: Academic Press; 2001.

10. Agrawal S, Khan F, Talwar S, Nityanand S. Short tandem repeat technology has diverse applications: individual identification, phylogenetic reconstruction and chimerism based post haematopoietic stem cell transplantation graft monitoring. Indian Journal of Medical Sciences 2004; 58: 297-304.

11. Haas J, Frese KS, Peil B, Kloos W, Keller A et al. Atlas of the clinical genetics of human dilated cardiomyopathy. European Heart Journal 2014; 36: 1123-1135.

12. Hearne CM, Ghosh S, Todd JA. Microsatellites for linkage analysis of genetic traits. Trends in Genetics 1992; 8: 288-294.

13. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research 1999; 27: 573.

14. Legendre M, Pochet N, Pak T, Verstrepen KJ. Sequence-based estimation of minisatellite and microsatellite repeat variability. Genome Research 2007; 17: 1787-1796.

15. Peakall R, Smouse PE. GENALEX 6: Genetic Analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Resources 2006; 6: 288-295.

16. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics 1980; 32: 314.

17. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 1978; 89: 583-590.

18. Kalinowski ST, Taper ML, Marshall TC. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 2007; 16: 1099-1106.

19. Butler JM. Short tandem repeat typing technologies used in human identity testing. Biotechniques 2007; 43: 2-5.

20. Zafari Z, Amirian A, Nejad FR, Akbari V, Akbari MT et al. Development and diversity of a novel panel of short tandem repeat markers encompassing the SCN5A gene in Iranian population. Journal of Genetics 2017; 97: 1-4.