

Concurrency of Mutations, Microsatellites and Predicted Domains in *kcnq1*, *kcnh2* and *scn5a* Genes Causing Long qt Syndrome Disease

Allam Appa Rao\*, G Venkata Swamy, B L V Vinay Kumar, Ch S U Ravi Kumar

Department of Computer Sciences and System Engineering, Andhra University, Visakhapatnam, India.

\*Corresponding author: Prof. Allam Appa Rao, Andhra University College of Engineering (Autonomous), Visakhapatnam-530003, INDIA; Fax: +91-891-2747969; Phone: +91-891-2844204; E-mail: principal@aucevizag.ac.in; Web: http://aucevizag.ac.in

Received April 20, 2008; Accepted May 15, 2008; Published May 25, 2008

**Citation:** Allam AR, Swamy GV, Kumar BLVV, Kumar Ch SUR (2008) Concurrency of Mutations, Microsatellites and Predicted Domains in *kcnq1*, *kcnh2* and *scn5a* Genes Causing Long qt Syndrome Disease. J Proteomics Bioinform S1: S012-S016. doi:10.4172/jpb.s1000003

**Copyright:** © 2008 Allam AR, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## Abstract

Bioinformatics is the field of science in which biology, computer science, and information technology merge to form a single discipline. The revolutionary growth in the computation speed and memory storage capability has fueled a new era in the analysis of biological data. The ultimate goal of the field is to enable the discovery of new biological insights as well as to create a global perspective from which unifying principles in biology can be discerned. Bioinformatics, also known as genomics, computational genomics, or computational molecular biology is a rate. We have used the analogy of genome analysis and VIRUS (vital information recourse under siege) and analyzed KCNQ1, KCNH2 and SCN5A genes, which are playing an important role in LQTS disease. We tried to find out whether the presence of microsatellites or simple sequence repeats in the KCNQ1, KCNH2 and SCN5A genes, are having any significance in the generation of these mutations and checked whether these mutations are fallen in the regions of those microsatellites and if so, is there any significance of these microsatellites in the functional domains of the each gene?. Our analysis revealed that 24 of the 26 mutations of the KCNQ1 gene, 19 of the 21 mutations of the KCNH2 gene and 3 of the 7 mutations of the SCN5A gene, which are existing in the microsatellite regions are fallen in the domain regions of the respective genes and thus indicating a positive role of microsatellites in mutagenesis.

## Introduction

Long QT syndrome (LQTS) is a congenital disorder characterized by a prolongation of the QT interval on ECG and a propensity to ventricular tachyarrhythmia, which may lead to syncope, cardiac arrest, or sudden death. Congenital LQTS is usually inherited. It is caused by an abnormality in the gene that forms the ion channels, slowing the recovery phase of the heartbeat. LQTS is caused by mutations of the genes for cardiac potassium and sodium or calcium ion channels; 8 genes have been identified. On the basis of this genetic background, 6 types of Romano-Ward syndrome, 1 type of Andersen syndrome and 1 type of Timothy syndrome those genes are KVLQT1, or KCNQ1, KCNH2, SCN5A, ANK2, KCNE1, KCNE2, KCNJ2, and CACNA1C.

Apart from genes, the human genome also consists of a large number of nucleotide repeat units of size 1-6 bp repeated tandemly called Micro satellites or Simple Sequence Repeats (SSRs) or Short Tandem Repeats (STRs) (Schlotterer, C, 2000) Micro satellites are found in all the known genomes, spanning from prokaryotes, eukaryotes and viruses and are widely distributed both in coding and non-coding regions (Toth, G et al, 2000; Sreenu et al, 2007). Mutations in these microsatellite regions occur at much higher rate when compared with those in the rest of the genome (Ellegren, H, 2000).

Micro satellites are known to be highly polymorphic due to the high rate of mutations in their tracts (Jarne, P. and Lagoda, P.J.L, 1996). These mutations can be either in the

form of increase / decrease of repeat units or in the form of single nucleotide substitutions/deletions/insertions and other events (Fan, H. and Chu, J.Y, 2007). Increase or decrease of repeat units of micro satellites in coding regions might lead to shift in reading frames there by causing changes in protein product (Li, Y.C et al, 2004) and in non-coding re-

gions are known to effect the gene regulation (Martin P et al, 2005). Point mutations (Substitutions and Indels) are also found to occur at a higher rate in micro satellites than elsewhere (Sibly, R.M et al, 2003).

Micro satellite mutations with in or near certain genes are

### For the Gene KCNQ1

Sl. No	Microsatellits	Microsatellite region	Codon Change	Aminoacid	Codon number	Domain □
1	CGG	501-508	GGG-AGG	Gly-Arg	168	Pfam00520
2	CGG	501-508	GGG-CGG	Gly-Arg	168	Pfam00520
3	TGGTC	515-530	CGC-TGC	Arg-Cys	174	Pfam00520
4	TGGTC	515-530	CGC-CAC	Arg-His	174	Pfam00520
5	CCG	530-542	GCC-ACC	Ala-Thr	178	Pfam00520
6	CCG	530-542	GCC-CCC	Ala-Thr	178	Pfam00520
7	CCG	530-542	GGC-AGC	Gly-Ser	179	Pfam00520
8	GGC	563-570	GGG-AGG	Gly-Arg	189	Pfam00520
9	GC	567-576	CGG-CAG	Arg-Gln	190	Pfam00520 □
10	GC	567-576	CTG-CCG	Leu-Pro	191	Pfam00520
11	TCC	689-697	CGC-TGC	Arg-Cys	231	Pfam00520
12	CCA	771-779	CGC-CTC	Arg-Leu	259	Pfam00520
13	CCA	771-779	CGC-TGC	Arg-Cys	259	Pfam00520
14	TCT	818-829	TTC-TCC	Phe-Ser	275	Pfam00520
15	TG	908-916	TGG-TAG	Trp-Term	305	Pfam00520, Pfam07885
16	G	914-922	GGG-GTG	Gly-Val	306	Pfam00520, Pfam07885
17	G	914-922	GGG-AGG	Gly-Arg	306	Pfam00520, Pfam07885
18	CAC	926-938	GTC-ATC	Val-Ile	310	Pfam00520, Pfam07885
19	CAC	926-938	ACC-ATC	Thr-Ile	311	Pfam00520, Pfam07885
20	CAC	926-938	ACC-ATC	Thr-Ile	312	Pfam00520, Pfam07885
21	AGG	1681-1688	AGG-ATG	Arg-Met	562	Pfam03520
22	GC	1765-1773	GGC-GAC	Gly-Asp	589	Pfam03520
23	GC	1765-1773	GCC-ACC	Ala-Thr	590	Pfam03520

**For the Gene KCNH2**

sl.no	Microsatellites	Microsatellite Region	Codon Change	Aminoacid	Codon number	Domain □
1	ACG	134-142	GGC-GTC	Gly-Val	47	PAS
2	GCAGG	207-222	CAC-CGC	His-Arg	70	PAS
3	GC	213-216	CCG-CAG	Pro-Gln	72	PAS
4	CGC	226-241	GCT-CCT	Ala-Pro	78	PAS
5	TCA	365-382	ATG-AGG	Met-Arg	124	PAC
6	TCA	1403-1411	AAC-GAC	Asn-Asp	470	Pfam: Ion trans □
7	CAC	1416-1424	ACC-ATC	Thr-Ile	474	Pfam: Ion trans □
8	ACT	1475-1483	TAC-TGC	Tyr-Cys	493	Pfam: Ion trans □
9	GCT	1576-1601	CGG-CAG	Arg-Gln	531	Pfam: Ion trans □
10	TGCT	1646-1658	TTG-TCG	Leu-Ser	552	Pfam: Ion trans □
11	GGCT	1750-1757	GGC-AGC	Gly-Ser	584	Pfam: Ion trans □
12	GGC	1807-1815	GGC-AGC	Gly-Ser	604	Pfam: Ion trans, Pfam: Ion trans2
13	GGC	1876-1884	GGC-GTC	Gly-Val	626	Pfam: Ion trans, Pfam: Ion trans2
14	GGC	1876-1884	GGC-AGC	Gly-Ser	626	Pfam: Ion trans, Pfam: Ion trans2
15	GGC	1876-1884	TTC-TTG	Phe-Leu	627	Pfam: Ion trans, Pfam: Ion trans2
16	GGC	1876-1884	GGC-AGC	Gly-Ser	628	Pfam: Ion trans, Pfam: Ion trans2
17	CCA	1895-1903	AAC-AGC	Asn-Ser	633	Pfam: Ion trans, Pfam: Ion trans2
18	TCT	1916-1927	TTC-TTA	Phe-Leu	640	Pfam: Ion trans, Pfam: Ion trans2
19	TCA	1931-1939	ATG-GTG	Met-Val	645	Pfam: Ion trans, Pfam: Ion trans2

known to be responsible for some human neurodegenerative diseases. So, we made a brief study to check whether the mutations in KCNQ1, KCNH2 and SCN5A genes, have any relation with these microsatellites repeats and the study revealed interesting results.

**Methods**

All the experimental proved mutations of the genes KCNQ1, KCNH2 and SCN5A that are falling inside the coding region and are eventually leading to phenotypic differences were collected from the Human Gene Mutation Database (HGMD) (Stenson, P.D et al, 2003). Micro satellites are obtained from the Imperfect Micro satellite Extract (IMEX) tool (Mudunuri, S.B et al, 2007) using intermediate mode with default values 6 for single, 5 for di, 3 for

**For the Gene SCN5A**

sl.no	Microsatellites	Microsatellite region	Codon Change	Aminoacid	Codon number	Domain <input type="checkbox"/>
1	GGC	3333-3349	GAC-AAC	Asp-Asn	1114	Pfam: Na trans assoc
2	CTGCG	3901-3914	ACG-ATG	Thr-Met	1304	Pfam: Ion trans
3	TCA	4997-5005	GTC-ATC	Val-Ile	1667	Pfam: Ion trans <input type="checkbox"/>

tri, 2 for tetra, 2 for penta and 2 for hexa and obtained 74,129,196 microsatellites in KCNQ1, KCNH2 and SCN5A respectively. Since micro satellites are drawn from the nucleotide sequence and HGMD mutations are given for protein sequence we have used DNA to Amino Acid translator. We compared the microsatellite regions with the mutations whether they have mutations in those regions and found some of the microsatellites have occurred in those regions. Now we analyzed whether these mutations and microsatellites have fallen in the functional domains of those genes by using Simple Modular Architecture Research Tool (SMART) (Letunic, I et al, 2004) and the results are as follows.

**Results and Discussions**

Long QT Syndrome can be acquired or congenital disorder. Here we have discussed about congenital since it is a inherited disease. We have taken KCNQ1, KCNH2 and SCN5A out of eight disease causing genes since LQT1, LQT2 and LQT3 account for most cases of LQTS, with estimated prevalences of 45%, 45% and 7% respectively. We have calculated the microsatellites for these three genes found 24, 19 and 3 mutations in those genes which falls in the microsatellite regions and also falls in the different domains. And thus, we can state that Concurrency of Mutations, Microsatellites and Predicted Domains in KCNQ1, KCNH2 and SCN5A genes may leads to Long QT Syndrome disease.

**Conclusion**

Microsatellites are known for their higher rate of mutations and are known to be associated with various diseases. So, we analyzed the KCNQ1, KCNH2 and SCN5A gene mutations and their possible association with the micro satellites. These mutations from HGMD database are mapped on to the micro satellite tracts and the results seem to indi-

cate that micro satellites play an important role in mutagenesis and by mapping the same with the functional domains we can say that these can cause functionality changes of those genes. Extending this work on a large scale by analyzing large number of genes might give a better evidence of the role of micro satellites in generating mutations.

**Acknowledgement**

The authors are thankful for partial financial support from IIT up gradation grants of AUCE (A).

**References**

1. Ellegren, H (2000) Heterogeneous mutation processes in human microsatellite DNA sequences. *Nat Genet* 24: 400-402. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)
2. Fan H, Chu JY (2007) A brief review of short tandem repeat mutation. *Genomics Proteomics Bioinformatics* 5: 7-14. » [CrossRef](#) » [Google Scholar](#)
3. Jarne P, Lagoda P (1996) Microsatellites, from molecules to populations and back. *Trends Ecol Evol* 11: 424-429. » [CrossRef](#) » [Google Scholar](#)
4. Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, et al. (2004) SMART 4.0: towards genomic data integration. *Nucleic Acids Res* 32: D142-4. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)
5. Li YC, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: structure, function, and evolution. *Mol Biol Evol* 21: 991-1007. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)
6. Martin P, Makepeace K, Hill SA, Hood DW, Moxon ER (2005) Microsatellite instability regulates transcription factor binding and gene expression. *PNAS* 102: 3800-3804. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)

7. Schlotterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109: 365-371.  
» [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)
8. Mudunuri SB, Nagarajaram HA (2007) IMEx: Imperfect Microsatellite Extractor. *Bioinformatics* 23: 1181-1187. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)
9. Sibly RM, Meade A, Boxall N, Wilkinson MJ, Corne DW, etal. (2003) The structure of interrupted human AC microsatellites. *Mol Biol Evol* 20: 453-9. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)
10. Splawski I, Shen J, Timothy KW, Vincent GM, Lelumann MH, etal. (2005) Genomic Structure of three long QT syndrome genes: KVLQT1, HERG, and KCNE1. Cardiology Division University of Utah Salt Lake City Utah 84112 USA. » [CrossRef](#) » [Pubmed](#)
11. Sreenu VB, Kumar P, Nagarajaram HA (2007) Simple sequence repeats in mycobacterial genomes. *J Biosci* 32: 3-15. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)
12. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, etal. (2003) The Human Gene Mutation Database (HGMD®): 2003 Update. *Hum Mutat* 21: 577-581. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)
13. Toth G, Gaspari Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res* 10: 967-981. » [CrossRef](#) » [Pubmed](#) » [Google Scholar](#)