Research Article

Structural and Functional Impact of Genetic Variations in FOXC2: A Computational Study

Mukund Madhav[1], VineetVaibhav[1], Chandan Kumar Gautam[1], Rajasekaran R[2], Ramanathan K[1], Shanthi V[1]*
[1]Industrial Biotechnology Division, School of Bio Sciences and Technology, VIT University, Vellore, Tamil Nadu, India.
[2]Bioinformatics Division, School of Bio Sciences and Technology, VIT University, Vellore, Tamil Nadu, India.
***Corresponding author's E-mail:** shanthi.v@vit.ac.in

Accepted on: 15-12-2013; **Finalized on:** 28-02-2014.

ABSTRACT

Forkhead Box C2 gene plays an important role in the development of mesenchymal tissues. A mutation caused in FOXC2 gene leads to lymphedema-distichiasis. These variations occur mainly due to nsSNPs. Computational analysis of the SNPs occurring in this gene was done using SIFT, PolyPhen 2.0, I MUTANT 3.0, PANTHER, PhD SNP and SNPs & Go. The analysis unveils that out of 38 variants of this gene, 5 (I98V, I98M, P140L, R121H and S125L) were found deleterious by all the 6 servers. T147I was found damaging except by SNP&GO. Further the interaction between the different variants and protein was done using the STRING 9.05. The current study investigated the possible deleterious variations of the FOX C2 gene. Pre diagnosis of these variants would help predict the possible onset of a disease and further developing personalized drug to treat a particular disease.

Keywords: Forkhead Box C2, I-MUTANT 3.0, Lymphedema-distichiasis, nsSNP, PolyPhen 2.0, SIFT, SNPs & GO, STRING 9.05.

INTRODUCTION

Foxc2gene belongs to the Forkhead family of transcription factors which is characterized by a distinct DNA-binding forkhead domain. It is located on chromosome 16q24.1.[1] The specific function of this gene has not yet been determined; however, it may play a role in the development of mesenchymal tissues. Finegold[2] identified a mutation in the FOXC2 gene in affected members of a family with lymphedema-distichiasis syndrome which shows phenotypic overlap with lymphedema-yellow nail syndrome. Lymphedema-distichiasis is an autosomal dominant disorder. It is classically represented by distichiasis and lymphedema of the limbs. The complications of this disorder includes problem in cornea, cardiac defects, varicose veins, ptosis, cleft palate, spinal extradural cysts, and photophobia.[3-5]

Genetic variation is the main reason for triggering the phenotypic diversity in an organism.[6] The recent research investigation presents that, a single organism has more than 3.5 million genetic variations in its genome, roughly corresponding to 1,000 alterations per mega–base pair.[7, 8] The most common type of variation is SNP (Single Nucleotide Polymorphism), which occurs once in every 300 base pairs of sequence with minor allele frequency (MAF)>1%.[9-11] The two databases Online Mendelian Inheritance in Man (OMIM) and Human Gene Mutation Database (HGMD) contains lots of examples of diseases induced due to change in variants which depicts the significance of nonsense SNP.[12-14] It has been reported that 67,000- 200,000 non synonymous SNPs are common in the human population whereas only 24,000-40,000 nsSNPs are heterozygous.[15-17] Considering the vastness of the reported SNPs, it would be difficult, expensive and time consuming to analyze the influence of each nsSNP on translation and protein function. Since these single amino acid changes can have a huge impact on protein expression, it's important to investigate the nsSNP to predict the protein function or expression of deleterious genes.[14] Using computational analysis for investigating the occurrence of SNP can help in studying the possible genetic variation and related diseases. With the advancement of world towards pharmacogenomics, genome mapping and drug discovery; the SNP is in lime light and has a crucial role to play.[18]

METHODS

SIFT tool for analysis of the deleterious nature of the missense mutation of variant

The tool is accessed at (http://sift.jcvi.org) for analysis of the deleterious nature of the missense mutation of variant. Sorting Intolerant from Tolerant (SIFT) tool, forecasts the consequences of change in coding variant of the protein function.[19] SIFT explores the protein database by using the PSI-BLAST and collects functionally related protein sequences. Subsequently by performing the homologous alignment of sequence, it computes the probability of the existence of the amino acid at a particular place.[20] The scores <0.05 are considered as intolerated whereas scores ≥0.05 are took as tolerated.[21]

Analysis the potential effect of the amino acid substitution using the Polyphen 2.0

Polymorphism Phenotyping (PolyPhen 2) tool accessed at (http://genetics.bwh.harvard.edu/pph2/), tells about the potential effect of the amino acid substitution on the constancy and work of protein utilizing structural and evolutionary characters.[22] Protein sequence, substituted amino acid and its position was provided to server as input. The obtained output is tabulated in table 1. Prediction with the score of "probably damaging",

"possibly damaging" and "benign" are considered as affecting the structure or function, may or may not affecting the structure or function and do not affect the structure and function respectively.[23]

Analysis of the stability of the protein using the I-MUTANT 3.0

I-mutant 3.0 is a web server accessed at (http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi), used for predicting the stability of the protein in case of single change in amino acid. This tool has ability to deduce information regarding data from the ProTherm,[24] which is an important database providing the values of experimental free energy change of protein stability due to mutation.[25] Protein Fasta sequence is provided along with the new residue and position for obtaining the free energy change of protein stability (DDG) with sign. DDG value with the positive sign is the indication of the mutated protein with high stability

Relating the disease and mutation by the SNPs & GO

SNPs & GO is a web server tool accessible at (http://snps.biofold.org/snps-and-go/snps-and-go.html), predicts the relation between disease and the mutation by using the protein functional annotation. To predict more accurately this tool includes; local sequence surrounding the mutation, characteristics deduced from sequence alignment, prediction data furnished from PANTHER (Protein Analysis THrough Evolutionary Relationships) assortment system and operational based log-odds hit computed for the GO (Gene Ontology) classification.[26]

The PANTHER prediction method of pathogenicity is based on alignment of evolutionary related protein and then computation of substitution position specific evolutionary conservation score.[27] PANTHER subPSEC results differ from 0 (neutral) to about -10 (most likely to be diseased). Protein sequences having subPSEC Value <-3 is said to be diseased.[28]

In the PhD SNP, the single sequence Support Vector Machine (SVM) categories the variants into diseased or neutral grounded on the type of the substitution and attributes of the neighboring sequence surrounding.

Proteins interaction analysis using STRING 9.05 database

Interaction analysis is done using the STRING 9.05 accessed at (http://string-db.org/newstring_cgi/show_input_page.pl?UserId=pzmTa_Bsb7hY&sessionId=YcNoJnby604v). STRING is a database of known and predicted protein interactions. The interactions included are direct (physical) and indirect (functional) associations; they are derived from four sources; genomic context, high through output experiments, conserved co-expression and previous knowledge.

RESULTS AND DISCUSSION

The computational analysis is commenced in 38 resides fetched from the dbSNP for MFH1 genes in *Homo sapiens*. Sequence homology based SIFT tool, a structure based PolyPhen 2.0 and I-MUTANT 3.0 tools, protein function based SNPs & GO tool were taken into the account for the analysis of a deleterious nature of the missense mutation.

Categorization Using Homology Based Sift Tool

SIFT tool is used to categorize 38 variants into tolerating or damaging based on their tolerance indices obtained. After submission of input in SIFT server, 26.31% nsSNPs are obtained damaging with the between scores of 0.00-0.04 and the rest were non- damaging or tolerated; values ranging from 0.05-1. The obtained tolerated index score is tabulated in the table 1 and graph showing % contribution can be seen in figure 1. Out of 38 variants 9 variants viz. I98V, I120V, N87Y, I98M, P140L, T147I, Y488C, R121H, S125L are highly intolerant with the score of 0.00. Based on the tolerance index score one can sjudge the severity of mutation and will be able to pay heed on lower score first.

Table 1: Analysis using SIFT, POLYPHEN 2.0, I-MUTANT 3.0, PhD SNP, Panther and SNP & GO

rs ID'S	Nucleotide change	Amino acid change	SIFT	POLYPHEN 2.0	I-mutant 3.0	PhD-SNP	Panther	SNP & GO
rs377509198	A/G	I98V	0	0.968	-0.98	Disease	Disease	Disease
rs375555433	A/G	S36N	0.3	0.003	-0.47	Neutral	NA	Neutral
rs375500879	A/G	I120V	0	0.997	-0.93	Neutral	Disease	Neutral
rs375372556	G/T	S429I	0.24	0.232	0.5	Disease	Disease	Disease
rs374000899	A/T	N87Y	0	0.997	0.48	Disease	Disease	Neutral
rs371359613	A/G	A422T	0.98	0.001	-0.46	Neutral	Neutral	Neutral
rs371246110	C/G	I98M	0	1	-1.57	Disease	Disease	Disease
rs370541542	C/G	E179Q	0.11	0.068	-0.42	Neutral	Disease	Neutral
rs369622966	C/T	P64L	0.76	0.078	-0.3	Disease	NA	Neutral
rs369098719	C/G	R54G	0.14	0.377	-1.16	Neutral	NA	Neutral

Table 1: Analysis using SIFT, POLYPHEN 2.0, I-MUTANT 3.0, PhD SNP, Panther and SNP & GO (Continued.....)

rs ID'S	Nucleotide change	Amino acid change	SIFT	POLYPHEN 2.0	I-mutant 3.0	PhD-SNP	Panther	SNP & GO
rs368190981	C/T	S8P	0.07	0.969	-0.29	Neutral	NA	Neutral
rs202085650	C/G	L183V	0.29	0.06	-1.52	Neutral	Disease	Neutral
rs201924901	C/G	H182D	0.57	0.001	-0.27	Neutral	Disease	Neutral
s201895173	A/G	V212M	0.15	0.944	-0.38	Neutral	Disease	Neutral
rs201833900	C/T	P140L	0	0.999	-0.71	Disease	Disease	Disease
rs201456476	A/C	D202A	0.75	0.019	0.32	Neutral	Disease	Neutral
rs201393006	C/G/T	P45R	0.35	0.811	-0.86	Disease	N/A	Neutral
rs201189554	C/T	T147I	0	0.992	-0.36	Disease	Disease	Neutral
rs201027957	G/T	L183R	0.54	0.014	-1.27	Neutral	Disease	Neutral
rs200766961	A/T	K205M	0.1	0.913	0.14	Neutral	Disease	Neutral
rs200751941	C/G	P195A	0.96	0	-0.59	Neutral	Neutral	Neutral
rs200483763	A/C	A189E	1	0.007	-0.27	Neutral	Disease	Neutral
rs200408083	C/G	A422G	0.32	0.155	-0.88	Disease	Disease	Neutral
rs200301728	A/C	P220Q	0.15	0.487	-1.38	Disease	Disease	Disease
rs200039004	A/G	S232N	0.37	0.967	-0.44	Neutral	Disease	Neutral
rs199924880	A/G	K177R	0.08	0.059	-0.22	Neutral	Disease	Neutral
rs199862001	C/T	P45S	0.38	0.031	-1.46	Neutral	NA	Neutral
rs199772307	A/G	Q117R	1	1	-0.23	Disease	Disease	Neutral
rs147258453	A/G	Q444R	1	0.981	-0.06	Disease	Neutral	Neutral
rs145316350	A/G	Y488C	0	0.999	-1.17	Disease	NA	Neutral
rs144326380	A/T	K92M	0.05	0.264	0.14	Neutral	Disease	Neutral
rs140209790	C/T	P58S	0.1	0.084	-1.21	Neutral	NA	Neutral
rs138612549	A/G	N455S	0.17	0	-0.71	Disease	Neutral	Neutral
rs121909107	A/G	R121H	0	1	-1.23	Disease	Disease	Disease
rs121909106	C/T	S125L	0	1	-0.23	Disease	Disease	Disease
rs78018668	C/T	S191F	0.04	0.36	0.35	Neutral	Disease	Neutral
rs61753346	C/T	C498R	0.21	0.998	-0.14	Disease	NA	Disease
rs7187073	G/T	Q413H	0.12	0.001	-0.1	Neutral	Disease	Neutral

Table 2: Effect on the variants' solubility, charge, polarity

Variant	I-MUTANT 3.0	Solubility		Charge		Polarity	
		Before	After	Before	After	Before	After
I98V	Decrease	Hydrophobic	Hydrophobic	Uncharged	Uncharged	Nonpolar	Nonpolar
I120V	Decrease	Hydrophobic	Hydrophobic	Uncharged	Uncharged	Nonpolar	Nonpolar
A422T	Decrease	Hydrophobic	Neutral	Uncharged	Uncharged	Nonpolar	Polar
I98M	Decrease	Hydrophobic	Hydrophobic	Uncharged	Uncharged	Nonpolar	Nonpolar
E179Q	Decrease	Hydrophilic	Hydrophilic	Negative	Uncharged	Polar	Polar
P64L	Decrease	Neutral	Hydrophobic	Uncharged	Uncharged	Nonpolar	Nonpolar
R54G	Decrease	Hydrophilic	Neutral	Positive	Uncharged	Polar	Nonpolar
L183V	Decrease	Hydrophobic	Hydrophobic	Uncharged	Uncharged	Nonpolar	Nonpolar
H182D	Decrease	Neutral	Hydrophilic	Positive	Negative	Polar	Polar
V212M	Decrease	Hydrophobic	Hydrophobic	Uncharged	Uncharged	Nonpolar	Nonpolar
P140L	Decrease	Neutral	Hydrophobic	Uncharged	Uncharged	Nonpolar	Nonpolar

Table 2: Effect on the variants' solubility, charge, polarity (Continued.....)

Variant	I-MUTANT 3.0	Solubility		Charge		Polarity	
		Before	After	Before	After	Before	After
P45R	Decrease	Neutral	Hydrophilic	Uncharged	Positive	Nonpolar	Polar
T147I	Decrease	Neutral	Hydrophobic	Uncharged	Uncharged	Polar	Nonpolar
L183R	Decrease	Hydrophobic	Hydrophilic	Uncharged	Positive	Nonpolar	Polar
P195A	Decrease	Neutral	Hydrophobic	Uncharged	Uncharged	Nonpolar	Nonpolar
A189E	Decrease	Hydrophobic	Hydrophilic	Uncharged	Negative	Nonpolar	Polar
A422G	Decrease	Hydrophobic	Neutral	Uncharged	Uncharged	Nonpolar	Nonpolar
P220Q	Decrease	Neutral	Hydrophilic	Uncharged	Uncharged	Nonpolar	Polar
K177R	Decrease	Hydrophilic	Hydrophilic	Positive	Positive	Polar	Polar
P45S	Decrease	Neutral	Neutral	Uncharged	Uncharged	Nonpolar	Polar
Q117R	Decrease	Hydrophilic	Hydrophilic	Uncharged	Positive	Polar	Polar
Q444R	Decrease	Hydrophilic	Hydrophilic	Uncharged	Positive	Polar	Polar
Y488C	Decrease	Neutral	Hydrophobic	Uncharged	Uncharged	Polar	Polar
P58S	Decrease	Neutral	Neutral	Uncharged	Uncharged	Nonpolar	Polar
N455S	Decrease	hydrophilic	Neutral	Uncharged	Uncharged	Polar	Polar
R121H	Decrease	hydrophilic	Neutral	Positive	Positive	Polar	Polar
S125L	Decrease	Neutral	Hydrophobic	Uncharged	Uncharged	Polar	Nonpolar

Table 3: Damaging Variants by all the servers ss

Variant	SIFT	Polyphen 2.0	i-mutant 3.0	PhD-SNP	Panther	SNP & GO
I98V	Damaging	Probably Damaging	Decrease	Disease	Disease	Disease
I98M	Damaging	Probably Damaging	Decrease	Disease	Disease	Disease
P140L	Damaging	Probably Damaging	Decrease	Disease	Disease	Disease
R121H	Damaging	Probably Damaging	Decrease	Disease	Disease	Disease
S125L	Damaging	Probably Damaging	Decrease	Disease	Disease	Disease
T147I	Damaging	Probably Damaging	Decrease	Disease	Disease	Neutral

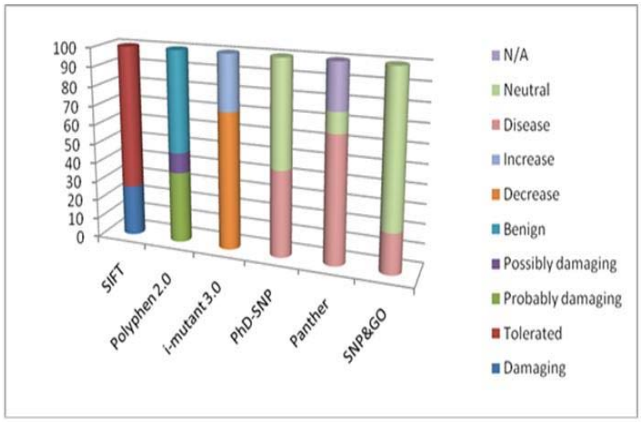

Figure 1: Distribution % of nsSNP in the FOXC2 gene

Identification of nonsense mutation using PolyPhen 2.0

This structure based analysis tool accepted FASTA sequence along with the position and substitution in variants. In output, the results obtained are 36.84% probably damaging, 10.52% possibly damaging and 52.63% benign with the score varying from 0 to 1. This % contribution can be seen in fig1 and scores can be checked in table1. Out of 38 variants, the 14 variants viz., I98V, I120V, N87Y, I98M, S8P, P140L, T147I, S232N, Q117R, Q444R, Y448C, R121H, S125L, C498R variants fall under the category of probably damaging with a score range of 0.967-1 and the 4 variants viz., V212M, P45R, K205M, P220Q come under the possibly damaging category with a score range of 0.487-. 944. The rests of the variants are benign with a score varying from 0 to 0.366. 7 variants are commonly identified as damaging or deleterious by both the tools SIFT and PolyPhen 2.0. This significant value will help in the assessment of severity on the protein function.

Investigation of nsSNPs using the I-MUTANT 3.0 Server

I-MUTANT 3.0 based on support vector machine, accepted protein sequence as input. Output showed that 71.05 % of the variant depicted the decrease in stability whereas 28.94 % expressed the increase in stability. All the related data are tabulated in the table 1 and % composition bar graph can be seen in the fig1.Variants I98V, I120V, A422T, I98M, E179Q, P64L, R54G, L183V, H182D, V212M, P140L, P45R, T147I, L183R, P195A, A189E, A422G, P220Q, K177R, P45S, Q177R, Q444R,

Y488C, P58S, N455S, R121H, S125L with the values of -.98, -.93, -.46, -1.57, -0.42, -.3, -1.16, -1.52, -.27, -.38, -.71, -.86, -.36, -1.27, -.59, -.27, -.88, -1.38, -.22, -1.46, -.23, -.06, -1.17, -1.21, -0.7, -1.23, -.23 respectively showed decrease in stability. The above variants showing decrease in stability are further interpreted for change in other properties such as polarity, solubility and charge, which is tabulated in table 2. 6 Variants are favoring deleterious character in common according to SIFT, PolyPhen 2.0 and I-MUTANT 3.0.

Validation of results using PhD-SNP, PANTHER and SNPs & GO

The results obtained from the SIFT, PolyPhen 2.0, I-MUTANT 3.0 servers are endorsed by PhD-SNP, PANTHER and SNPs & GO tools for the protein function upon nonsense mutation. The 17 variants viz., I98V, S429I, N87Y, I98M, P64L, P140L, P45R, T147I, A422G, P220Q, Q117R, Q444R, Y488C, N455S, R121H, S125L, C498R due to probability score > 0.5 are diseased and have a reliability score lying between 0-10. According to PANTHER, the 25 variants viz., I98V, I120V, S429I, N87Y, I98M, E179Q, L183V, H182D, V212M, P140L, D202A, T147I, L183R, K205M, A189E, A422G, P220Q, S232N, K177R, Q117R, K92M, R121H, S125L, S191F, Q413H are detected diseased with score and reliability index lying between 0.516 - 0.981, 0-10 respectively. Finally SNPs & GO identified 8 variants viz., I98V, S429I, I98M, P140L, P220Q, R121H, S125L, C498R as diseased with the score > 0.5 and reliability score lying in range of 1 to 10. According to PhD-SNP, PANTHER and SNPs & GO, 44.736%, 65.789% and 21.05% variants respectively are procured diseased. % contribution can be viewed in figure 1 and all the data retrieved from the tools are indexed in the table 1. Total 5 variants are detected deleterious or diseased by all the tools used above.

Prediction of Protein Interaction using STRING 9.05

Interaction analysis divulged that MFH1 is related to mesoderm posterior 1 /Mesp1(plays role in the epithelialization of somitic mesoderm and in the development of cardiac mesoderm), glycosyl phosphatidylinositol anchor attachment protein 1/GPAA1(essential for GPI-anchoring of precursor proteins but not for GPI synthesis), delta-like 4/DLL4 (plays role in the Notch signaling pathway), angiopoietin 2/ ANGPT2 (Can induce tyrosine phosphorylation of TIE2), neuropilin 1/NRP1 (a receptor involved in the development of the cardiovascular system, in angiogenesis, in the formation of certain neuronal circuits and in organogenesis outside the nervous system), chemokine (C-X-C motif) receptor 4/CXCR4 (involved in haematopoiesis and in cardiac ventricular septum formation, essential role in vascularization of the gastrointestinal tract), insulin/INS (decreases blood glucose concentration), myogenic differentiation 1/MYOD1 (involved in muscle differentiation), sonic hedgehog homolog /SHH (binds to the patched (PTC) receptor) and vascular endothelial growth factor A /VEGFA (induces endothelial cell proliferation, promotes cell migration, inhibits apoptosis, and induces permeability of blood vessels). The information regarding the interaction is shown in figure 2.

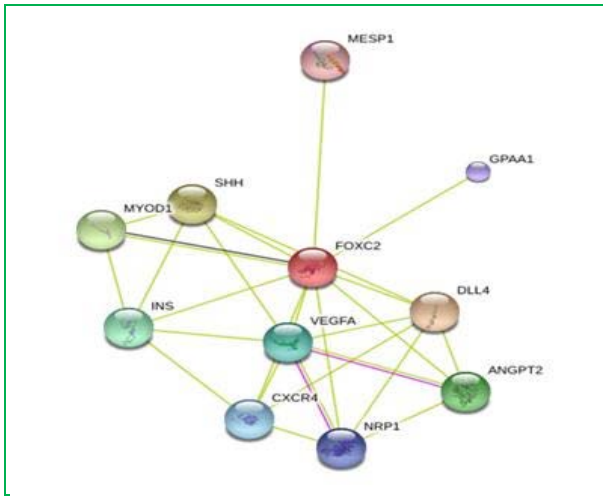

Figure 2: FOXC2interactions with the other proteins

CONCLUSION

38 variants which were studied, 10 and 28 were found damaging and tolerated respectively by SIFT analysis. By Polyphen 2.0 14 probably damaging, 4 possibly and 20 benign are reported. Stability of variant was analyzed by I MUTENT 3.0, in 38 variant the stability of 27 was found to be decreased and for 11 it got increased. 17 disease causing variant is reported by PHD-SNP, 25 by panther, 8 by SNPs & GO. Five variants I98V, I98M, P140L, R121H and S125L had been found damaging by all tools used for this work and is tabulated in table 3. T147I has been found damaging by all the tools except SNP&GO. Two variants viz., R121H, S125L are also confirmed damaging by wet lab in literature. According to literature when above two variants R121H, S125L were tested biochemically, both mutations showed reduction in DNA binding and transcriptional capacity. In addition R121H had effect on the nuclear localization of FOXC2.[29]

REFERENCES

1. Gene names, Retrieved 21 sept 2013. http: // www. Gene names. Org /gene families/FOX.
2. Finegold DN, Kimak MA, Lawrence, EC, Levinson KL, Cherniske EM, Pober BR, Dunlap JW, Ferrell RE, Truncating mutations in FOXC2 cause multiple lymphedema syndromes, Human Molec Genet, 10, 2001, 1185-1189.
3. Fang J, Dagenais SL, Erickson RP, Arlt MF, Glynn MW, Gorski JL, Seaver LH, Glover TW, Mutations in FOXC2 (MFH-1), a forkhead family transcription factor, are responsible for the hereditary lymphedema-distichiasis syndrome, Am J Hum Genet, 67, 2000, 1382-1388.
4. Brice G, Mansour S, Bell R, Collin JRO, Child AH, Brady AF, Sarfarazi M, Burnand KG, Jeffery S, Mortimer P, Murday VA, Analysis of the phenotypic abnormalities in lymphoedema-distichiasis syndrome in 74 patients with FOXC2 mutations or linkage to 16q24, J Med Genet, 39, 2002, 478-483.

5. Johnson SM, Kincannon JM, Horn TD, Lymphedema-distichiasis syndrome: report of a case and review, JAMA Dermatol, 135(3), 1999, 347-348.
6. Bromberg Y, Rost B, SNAP: predict effect of nonsynonymous polymorphisms on function, Nucleic Acids Research, 35, 2007, 3823–3835.
7. Cooper GM, Shendure J, Needles in stacks of needles: Ending disease-causal variants in a wealth of genomic data, Nature Reviews Genetics, 12, 2011, 628–640.
8. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D, MutationTaster evaluates disease-causing potential of sequence alterations, Nature Methods, 7, 2010, 575–576.
9. Kruglyak L, Nickerson DA, Variation is the spice of life, Nat Genet, 27**,** 2001, 234–236.
10. Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han JH, Haplotype variation and linkage disequilibrium in 313 human genes. Science, 293, 2001, 489–493.
11. MR, Marnellos G, Kammerer S, Hoyal CR, Shim MM, Cantor CR, Braun A, Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions, Genome Re, 14, 2004, 1664-1668.
12. Hamosh 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 Res, 33, 2005, D514–D517.
13. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Human Gene Mutation Database (HGMD), Hum Mutat, 21, 2003, 577–581.
14. Ng PC, Henikoff S, Predicting the Effects of Amino Acid Substitutions on Protein Function, Annu. Rev. Genomics Hum Genet, 7, 2006, 61–80.
15. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, et al., Characterization of singlenucleotide polymorphisms in coding regions of human genes, Nat Genet, 22, 1999, 231–238.
16. Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, et al., Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis, Nat. Genet., 22, 1999, 239–247.
17. Livingston RJ, Niederhausern AV, Jegga AG, Crawford DC, Carlson CS, et al., Pattern of sequence variation across 213 environmental response genes, Genome Res, 14, 2004, 1821–1831.
18. Bell PA, Chaturvedi S, Gelfand CA, Huang CY, Kochersperger M, Kopla R, Modica F, SNP stream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery, Biotechniques, 32, 2002, 70-72.
19. Ngak-Leng S, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC, SIFT web server: predicting effects of amino acid substitutions on proteins, Nucleic acids research, 40, 2012, W452-W457.
20. Ming Z, Chen HM, Wang YP, Missense mutations of MLH1 and MSH2 genes detected in patients with gastrointestinal cancer are associated with exonic splicing enhancers and silencers, Oncology letters, 5, 2013, 1710-1718.
21. Ng PC, Henikoff S, Predicting deleterious amino acid substitutions, Genome research, 11, 5, 2001, 863-874.
22. Ivan A, Jordan DM, Sunyaev SR, Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2, Current Protocols in Human Genetics, 2013, 7-20.
23. Melissa JM, Houck J, Chen C, Screening for deleterious nonsynonymous single-nucleotide polymorphisms in genes involved in steroid hormone metabolism and response, Cancer Epidemiology Biomarkers & Prevention 14, 5, 2005, 1326-1329.
24. Abdulla BK, Gromiha MM, Uedaira H, Kitajima K, Sarai A, ProTherm, version 4.0: thermodynamic database for proteins and mutants, Nucleic acids research 32, 1, 2004, D120-D121.
25. Rajasekaran R, Sethumadhavan R, Exploring the cause of drug resistance by the detrimental missense mutations in KIT receptor: computational approach, Amino acids 39, 3, 2010, 651-660.
26. Remo C, Capriotti E, Fariselli P, Martelli PL, Casadio R, Functional annotations improve the predictive score of human disease-related mutations in proteins, Human mutation 30, 8, 2009, 1237-1244.
27. Janita T, Olatubosun A, Vihinen M, Performance of mutation pathogenicity prediction methods on missense variants, Human mutation, 32, 4, 2011, 358-368.
28. Priya DCG, NagaSundaram N, Investigating the structural impacts of I64T and P311S mutations in APE1-DNA complex: a molecular dynamics approach, PloS one 7, 2, 2012, e31677.
29. Berry FB , Tamimi Y, Carle MV, Lehmann OJ, Walter MA, The establishment of a predictive mutational model of the forkhead domain through the analyses of FOXC2 missense mutations identified in patients with hereditary lymphedema with distichiasis, Human Molecular Genetics, 14, 18, 2005, 2619–2627.

Source of Support: Nil, **Conflict of Interest:** None.