In-silico analysis predicting effects of some Protein conding SNPs of human tumor suppressor protein53 (*TP53* gene) by using different Bioinformatic tools

Zainab Abdullah Waheed (Zainab.Abdullah@altoosi.edu.iq)
Altoosi university college https://orcid.org/0000-0002-6356-8133
Nadia Habeeb Sarhan
Basic Science Department - Faculty of Nursing
Ahmed Ibrahim Hamzah
Altoosi University College
Mustafa Mohammed Shaker
Altoosi University College

Research Article

Keywords: TP53, SNP, I-MUTANT 3.0, PhD-SNP, SNP&GO, PROVEAN

Posted Date: July 25th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1821578/v1

Abstract

Background

Tp53 gene encodes a tumor suppressor protein containing transcriptional activation, DNA binding, and oligomerization domains. The encoded protein responds to diverse cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. Mutations in this gene are associated with a variety of human cancers, including hereditary cancers such as Li-Fraumeni syndrome. Alternative splicing of this gene and the use of alternate promoters result in multiple transcript variants and isoforms

Aim of study

The study is aimed to study the effect of some protein coding SNPs on human health and if it benign or neutral by using different bioinformatics tools (I-MUTANT 3. 0,PROVEAN,and SNPS&GO; PhD-snp)

Results

The result shows that the protein sequence identification was as the following:

MAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDC

TTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRDRRTEEENLRKKGEPHHEL

PPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQMLLDLRWCYFLINSS

analysis of (10) SNPs shows that 5 SNPs(50%) having a nuclic acid change which result in a new amino acid formation and this change was considerd large increase stability according to I-Mutant while just 1SNP (1%)have neutral stability in the same tool but only 4(40%)SNPs showed large decrease stability in the analysis at the same tool which can be considered having effect on human health .

While as in PhD-SNP,SNP&GO,PROVEAN showed 0(0%) increase stability but prdected that (50%,50%,40%)have neutral affect in the protein modification while(40%,40%,50%) had Large decrease stability according to PhD-SNP,SNP&GO,PROVEAN respectively which may lead to human disease.

Introduction

TP53

TP53 gene encodes a tumor suppressor protein containing transcriptional activation, DNA binding, and oligomerization domains. The encoded protein responds to diverse cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. Mutations in this gene are associated with a variety of human cancers, including hereditary cancers (1).

Many cancer researchers concentrated their study during the 1960s and 1970s on tumor viruses and oncogenes, but multiple groups reported on the discovery of a cellular protein that was overexpressed in many tumors. This protein was known as p53 and had a molecular weight of about 53 kDa (protein 53). In one of the earliest experiments, p53 was discovered in cells that had been transformed by the simian virus 40 (SV40) large-T oncoprotein. **(2)**. Later, it was shown that p53 may bind to a number of additional oncoproteins made by various tumor viruses, such as the human papillomavirus (HPV) E6 protein and the adenovirus E1B 55K protein(3).

The capacity of p53 to alter rat embryo fibroblasts in conjunction with the well-known oncogene ras further demonstrated p53's "oncogenic" nature ,With the use of these data, it was determined that p53 might be a model cellular oncogene, whose overexpression is linked to cellular transformation when coexpressed with viral oncoproteins and mutant Ras.

The accumulation of mutant p53 in tumors was later demonstrated to be the cause of overexpressed p53, and missense mutations in the p53 gene did in fact impart substantial transforming potential.

(4)Additionally, a region critical for the bioactivities of the wild-type (wt) p53 protein was affected by dominant negative missense mutations in the initial cloned p53 cDNAs utilized in the early research. The concept of p53 as an oncogene was changing, and the wild-type form of p53 was now viewed as a crucial tumor suppressor. In addition, it has been demonstrated that binding of p53 to viral oncoproteins results in the inactivation or degradation of p53 (5).

Genomic Locations for TP53 Gene:

chr17:7,661,779-7,687,538(GRCh38/hg38),Size:25,760 basesOrientation:Minus strand (figure:1)

Quaternary structure

Forms homodimers and homotetramers, Binds DNA as a homotetramer. Interacts with AXIN1. Probably part of a complex consisting of TP53, HIPK2 and AXIN1 (By similarity), Interacts with histone acetyltransferases EP300 and methyltransferases HRMT1L2 and CARM1, and recruits them to promoters. Interacts (via C-terminus) with TAF1; when TAF1 is part of the TFIID complex. Interacts with ING4; this interaction may be indirect. Found in a complex with CABLES1 and TP73. Interacts with HIPK1, HIPK2, and TP53INP1. Interacts with WWOX. May interact with HCV core protein. Interacts with USP7 and SYVN1. Interacts with HSP90AB1. Interacts with CHD8; leading to recruit histone H1 and prevent transactivation activity (By similarity). Interacts with ARMC10, CDKN2AIP, NUAK1, STK11/LKB1, UHRF2 and E4F1. Interacts with YWHAZ; the interaction enhances TP53 transcriptional activity. Phosphorylation of YWHAZ on 'Ser-58' inhibits this interaction. Interacts (via DNA-binding domain) with MAML1 (via N-terminus). Interacts with MKRN1. Interacts with PML (via C-terminus). Interacts with MDM2; leading to ubiquitination and proteasomal degradation of TP53. Directly interacts with FBXO42; leading to ubiquitination and degradation of TP53. Interacts (phosphorylated at Ser-15 by ATM) with the phosphatase PP2A-PPP2R5C holoenzyme; regulates stress-induced TP53-dependent inhibition of cell proliferation. Interacts with PPP2R2A. Interacts with AURKA, DAXX, BRD7 and TRIM24. Interacts (when

monomethylated at Lys-382) with L3MBTL1. Isoform 1 interacts with isoform 2 and with isoform 4. Interacts with GRK5. Binds to the CAK complex (CDK7, cyclin H and MAT1) in response to DNA damage. Interacts with CDK5 in neurons. Interacts with AURKB, SETD2, UHRF2 and NOC2L. Interacts (via N-terminus) with PTK2/FAK1; this promotes ubiquitination by MDM2. Interacts with PTK2B/PYK2; this promotes ubiquitination by MDM2. Interacts with PRKCG. Interacts with PPIF; mediated ubiquitination and inhibits SNAI1-induced cell invasion(7). (Fig. 2A,B)

Definition and the origin of SNP

A nucleotide location that has a high rate of substitution among individual samples in a population is referred to as an SNP. The more precise definition provided by Brookes states that SNPs are single base pair sites in genomic DNA, at which many sequence alternatives (alleles) occur in healthy individuals in some population(s), where the abundance of the least frequent allele is 1% or more.

SNPs can be used as genetic markers because of the constraint on frequency that separates them from uncommon point mutations. According to this definition, single-nucleotide insertions/deletions (indels) are not considered to be part of SNP markers, even though some publications group these two polymorphism types with single-nucleotide substitutions under the overarching concept SNPs (simple nucleotide polymorphisms). Since there are four different types of nucleotides, there are theoretically four alleles that can exist at each place. However, in actuality, only two variants are typically found. As a result of the differential frequency of nucleotide transitions (A G, T C) and transversions, SNP markers are thus biallelic (A C, A T, G C, G T). Despite the fact that there are two times as many nucleotide transversion variations (10)

SNP Classification

According to their alignment to the structural element of genomic DNA or their functional effect, different types of SNPs are identified. SNPs are divided into exon (synonymous and non-synonymous), intron, and promoter types based on where they are located in a gene. Oligonucleotide changes that affect regulatory regions that regulate gene expression are known as regulatory SNPs or rSNPs. They might contain some inton SNPs as well as promoter SNPs. Single-nucleotide substitutions are divided into three categories depending on whether information about their functional impact is available: anonymous SNPs (functional impact is unknown), candidate SNPs (possibly having a functional impact), and protein SNPs (single-nucleotide substitutions causing a change in the protein function or expression). (11)

Methodology

Gene selection and gene sequence identification

The *Tp53* gene was selected from the databases of NCBI then the gene sequence were identified according to Ref sequences from NCBI(12)

SNP selection

Ten SNPS collected in-silico from NCBI SNPS then we choosed the SNPs that located on exosomes(having role in gene expression) then the nucleotide change were investigated and the amino acid modification and the location of the new amino acid ,after that the SNPs were investigate by different bioinformatics tools

in-silico SNP analysis

The SNPs were analysed by located the protein sequence and the aminoacid modification or change then submitting the informations in to the tool site which included:(,I-MUTANT 3. 0,PROVEAN,and SNPS&GO; PhD-snp)

Study Findings And Discussion

Gene selection and protein identification

The gene information were collected from (13)

This gene encodes a tumor suppressor protein containing transcriptional activation, DNA binding, and oligomerization domains. The encoded protein responds to diverse cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. Mutations in this gene are associated with a variety of human cancers, including hereditary cancers such as Li-Fraumeni syndrome. Alternative splicing of this gene and the use of alternate promoters result in multiple transcript variants and isoforms (13)

protein sequence identification

>NP_001263628.1 cellular tumor antigen p53 isoform I [Homo sapiens]

MAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDC
TTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRDRRTEEENLRKKGEPHHEL
PPGSTKRALPNNTSSSPQPKKKPLDGEYFTLQMLLDLRWCYFLINSS (1)

analysis of SNPs
rs1483922249 SNP:

the nuclic acid change was C > T was result in a new amino acid at position 359 which called Leucine and this change was considerd neutral according to PhD-SNP,SNP&GO,PROVEAN while had Large increase stability according to I-Mutant(14;15;16;17)

rs 909643864 SNP:

the nuclic acid change at C > T result in anew amino acid (Leucine) at position

300 and this change was considerd neutral according to PhD-SNP,SNP&GO,PROVEAN while had neutral stability according to I-Mutant. (14;15;16;17)

Rs1367492395 Snp:

the nuclic acid change at C > A result in anew amino acid (Arginie) at position

313 and this change was considerd neutral according to SNP&GO,PROVEAN while had Deleterious affect according to PROVEAN decrease stability according to I-Mutant. (14;15;16;17)

Rs 1177881399 Snp:

the nuclic acid change at A > G result in anew amino acid (G) glycine at position

324 and this change was considerd Deleterious, Disease, Disease according to PhD-SNP,SNP&GO,PROVEAN decrease stability according to I-Mutant (14;15;16;17)

Rs774269719 Snp:

the nuclic acid change was A > T was result in a new amino acid at position 377 which called Serine and this change was considerd neutral according to PhD-SNP,SNP&GO,PROVEAN while had Large increase large stability according to I-Mutant (14;15;16;17)

Rs 866775781 Snp:

the nuclic acid change was G > T was result in a new amino acid at position 132 which called Argenine and this change was considerd deleterious ,disease,disease according to PhD-SNP,SNP&GO,PROVEAN while had Large increase stability according to I-Mutant (14;15;16;17)

Rs1260903787 Snp:

the nuclic acid change was A > G was result in a new amino acid at position 204 which called G (glycine) and this change was considerd deleterious ,disease,disease according to PhD-SNP,SNP&GO,PROVEAN while had Large decrease large stability according to I-Mutant(14;15;16;17)

Rs864622237 Snp:

the nuclic acid change was T > G was result in a new amino acid at position 234 which called D and this change was considerd deleterious ,disease,disease according to PhD-SNP,SNP&GO,PROVEAN while had decrease large stability according to I-Mutant(14;15;16;17)

Rs 864622115 Snp:

the nuclic acid change was A > G was result in a new amino acid at position 174 which called G(glycine) and this change was considerd deleterious ,disease,disease according to PhD-SNP,SNP&GO,PROVEAN

while had Large increase large stability according to I-Mutant(14;15;16;17)

Rs 863224687 Snp:

the nuclic acid change was C > T was result in a new amino acid at position 359 which called Leucine and this change was considerd neutral according to PhD-SNP,SNP&GO,PROVEAN while had Large increasestability I-Mutant (14;15;16;17)

Recommendations And Conclusions

1-there many types of SNPs which result in gen modification but still having neutral affect

2-there are some SNPs in the exone location which have disease effect and ,may result in agene damage and cancer

3-more researchs about SNPs that causing gene damage

Table (1) bioinformatics tools results:

RsID	Nucleotide change	Amino acid change	I-Mutant	PROVEAN	SNPs&GO	PhD-snp
1483922249	359C > T	P359L	Large increase stability	Neutral	Neutral	Neutral
909643864	782C > T	P300L	Neutral stability	Neutral	Neutral	Neutral
1367492395	822C > A	S313R	Large decrease stability	Deleterious	Neutral	Neutral
1177881399	854A > G	D324G	Large decrease stability	Deleterious	Disease	Disease
774269719	1012A > T	T377S	Large increase stability	Neutral	Neutral	Neutral
866775781	279G > T	K132N	Large increase stability	Deleterious	Disease	Disease
1260903787	494A > G	E204G	Large decrease stability	Deleterious	Disease	Disease
864622237	583T > G	Y234D	Large decrease stability	Deleterious	Disease	Disease
864622115	403A > G	R174G	Large increase stability	Deleterious	Disease	Disease
863224687	847C > T	P322S	Large increase stability	Neutral	Neutral	Neutral

Abbreviations

Number	Abbreviations	Meaning
1	SNPs	Single nucleotide polymorphisms
2	*Tp53*	tumor suppressor protein
3	PolyPhen-2	Polymorphism Phenotyping v2
4	PROVEN	protein variation effect analyzer

Declarations

Competing interests:

The authors declare no competing interests.

Consent for publication

'Not applicable'

Availability of data and material

'Not applicable'

Competing interests

'Not applicable'

Funding

'Not applicable'

Authors' contributions

'Not applicable'

Acknowledgements

'Not applicable'

References

(1)- 1-https://www.ncbi.nlm.nih.gov/

(2)- Tornesello, M. L., Annunziata, C., Tornesello, A. L., Buonaguro, L., & Buonaguro, F. M. (2018). Human oncoviruses and p53 tumor suppressor pathway deregulation at the origin of human cancers. Cancers, 10(7), 213.

(3)- Querido, E., Blanchette, P., Yan, Q., Kamura, T., Morrison, M., Boivin, D., ... & Branton, P. E. (2001). Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. *Genes & development*, *15*(23), 3104-3117.

(4)_ Rivlin, N., Brosh, R., Oren, M., & Rotter, V. (2011). Mutations in the p53 tumor suppressor gene: important milestones at the various steps of tumorigenesis. *Genes & cancer*, *2*(4), 466-474.

(5)- Greenblatt, M. S. (1994). Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res.*, *54*, 4855-4878.

(6)- https://www.genecards.org/cgi-bin/carddisp.pl?gene=TP53

(7)- https://www.genecards.org/cgi-bin/carddisp.pl?gene=TP53&keywords=p53

(8)- https://www.genecards.org/cgi-bin/carddisp.pl?gene=TP53&keywords=p53

(9)- McCoy, M., Stavridi, E. S., Waterman, J. L., Wieczorek, A. M., Opella, S. J., & Halazonetis, T. D. (1997). Hydrophobic side-chain size is a determinant of the three-dimensional structure of the p53 oligomerization domain. *The EMBO journal*, *16*(20), 6230-6236.

(10)- https://www.ncbi.nlm.nih.gov/

(11)- https://www.ncbi.nlm.nih.gov/

(12)- https://www.ncbi.nlm.nih.gov/

(13)- https://www.genecards.org/Search/Keyword?queryString=p53

(14)- http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi

(15)-http://provean.jcvi.org/seq_submit.php

(16)-https://snps.biofold.org/snps-and-go/snps-and-go.html

(17)-https://snps.biofold.org/phd-snp/phd-snp.html

Figures

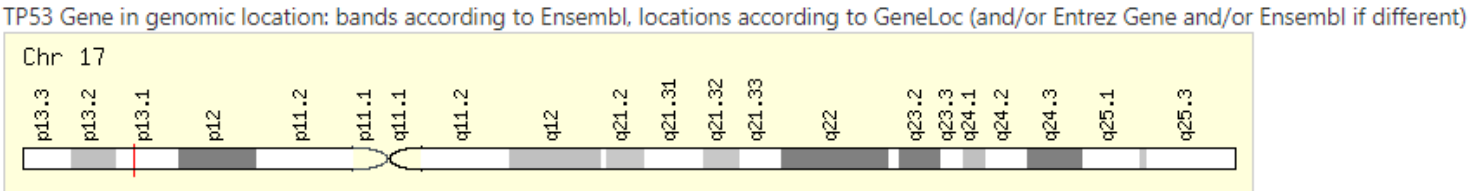

Figure 1

gene location (6)

Figure 2

A,B(8) interaction of the protein with other proteins

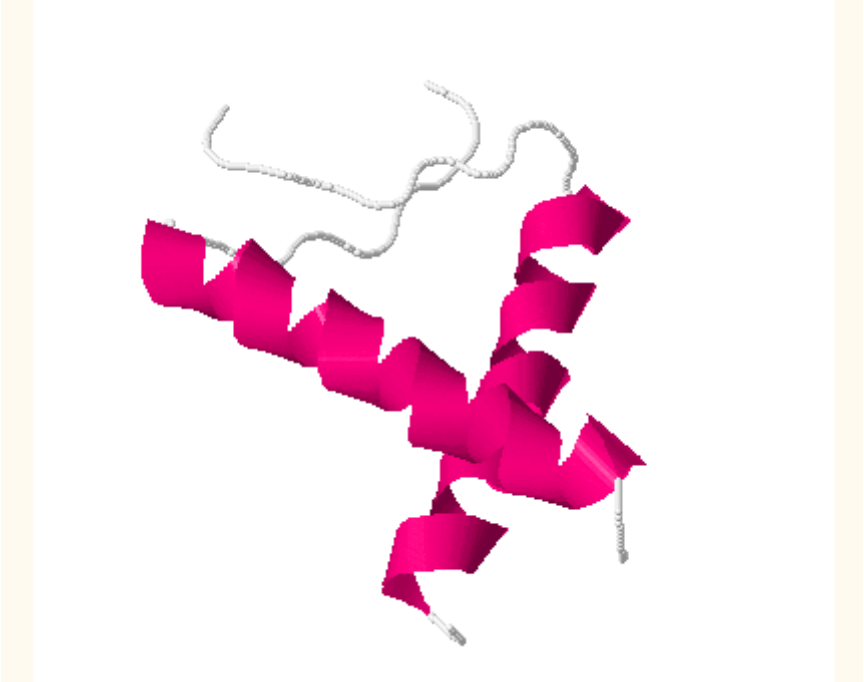

Figure 3

solution Structure Determination Of A P53 Mutant Dimerizatio Nmr, Minimized Average Structure(9)

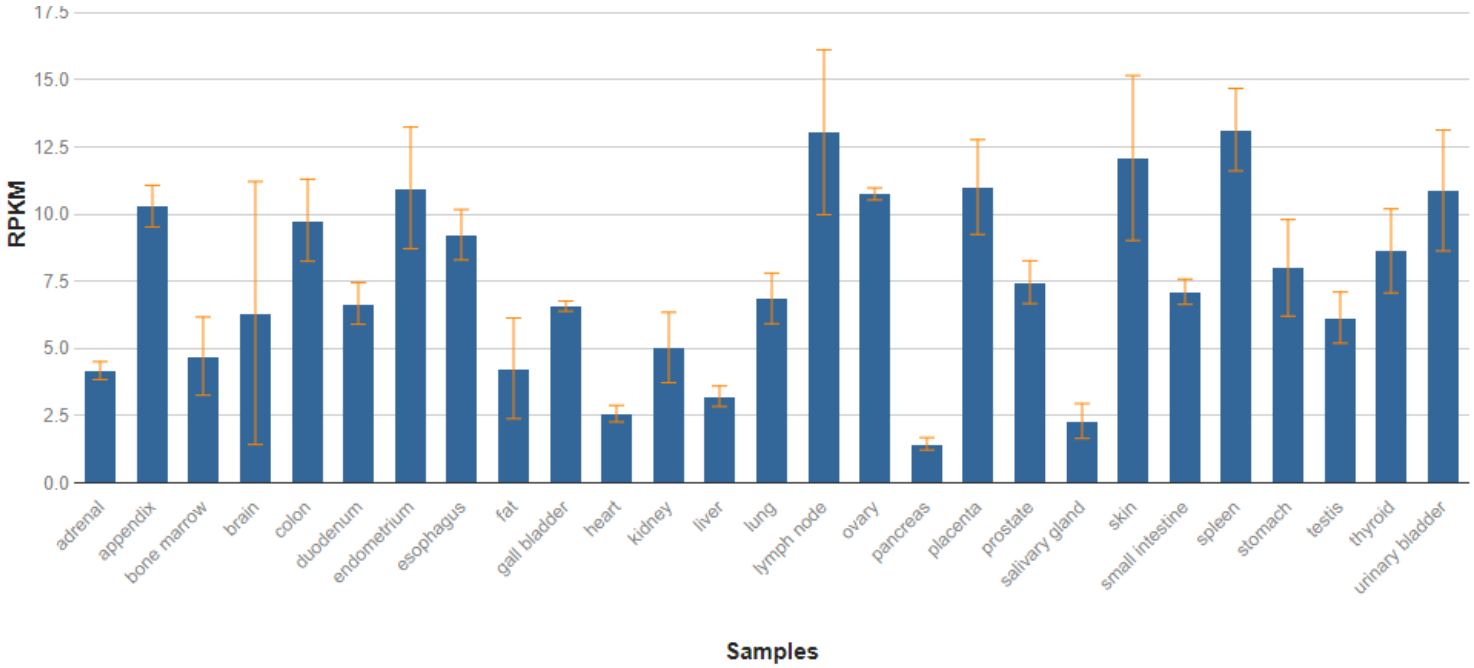

Figure 4

gene P53 expression in different organs(1)

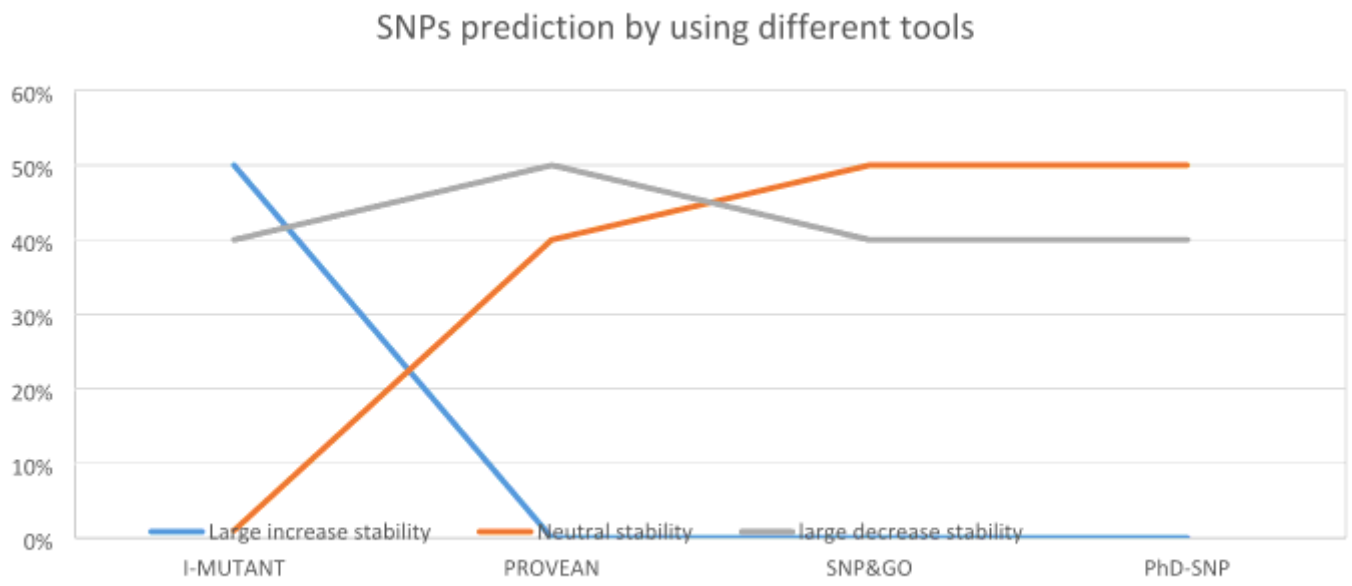

Figure 5

SNPs prediction by using different tools