PREDICTION AND STRUCTURAL IMPACTS OF THE PATHOGENIC MISSENSE SINGLE NUCLEOTIDE POLYMORPHISMS OF *IL-6* WITH *IN SILICO* METHODS

IN SILICO YÖNTEMLER İLE *IL-6*'NIN PATOJENİK TEK NÜKLEOTİT POLİMORFİZMLERİNİN TAHMİNİ VE YAPISAL ETKİLERİ

Nazlı Irmak GİRİTLİOĞLU

Molecular Biologist and Specialist in Bioengineering, İstanbul/TURKEY

Gizem KÖPRÜLÜLÜ KÜÇÜK

Lecturer, İstanbul Şişli Vocational School, Radiotherapy Program, İstanbul/TURKEY

Geliş Tarihi / Received: 28.08.2020
Kabul Tarihi / Accepted: 20.09.2020

Araştırma Makalesi/Research Article
DOI: 10.38065/euroasiaorg.239

ABSTRACT

Single nucleotide polymorphisms (SNPs) are known as single nucleotide changes in DNA among individuals. The most common type of SNPs is non-synonymous SNPs. Interleukin-6 (IL-6) is a pleiotropic cytokine produced by lymphoid and non-lymphoid cells with a wide range of biological functions in defense mechanism, immune response, hematopoiesis, and acute phase reactions. In this study, SNPs of *IL-6* were searched in NCBI dbSNP, and there are 2530 SNPs belong to *IL-6*, and 227 of them are classified as missense in this database. PANTHER, SIFT, and PROVEAN bioinformatic tools were used for determining the deleterious (D), probably damaging (PRD), or possibly damaging (POD) SNPs and the HOPE tool was utilized for the detection of the structural impacts caused by these SNPs on amino acids of *I-L6* protein products. SIFT and PROVEAN scores and PANTHER preservation times showed that 22 of these 227 SNPs have a category of D, PRD, or POD. According to HOPE tool results, rs11544633 (UniProt protein ids: P05231, B5MCZ3, B4DNV3, P05231, and B5MC21), rs199990564 (UniProt protein ids: P05231, B5MCZ3, B4DNV3, P05231, and B5MC21), rs201371019 (UniProt Protein ids: P05231, B5MCZ3, B4DNV3, B4DNV3, P05231, and B5MC21) causes the smaller amino acids while rs370818995 (UniProt Protein ids: P05231, B5MCZ3, and P05231) causes the bigger aminoacids in mutant types. Also, the mutant types of rs199990564 (UniProt Protein ids: P05231, B5MCZ3, B4DNV3, P05231, and B5MC21), and rs201371019 (UniProt Protein ids: P05231, B5MCZ3, B4DNV3, P05231, and B5MC21) cause more hydrophobic structure while rs370818995 (UniProt Protein ids: P05231, B5MCZ3, and P05231) causes less hydrophobic form than their wild types. In the study, this is the first time the pathogenic missense SNPs of *IL-6* of its was shown with *in silico* methods. We think the results of this study can help for future studies about the effects of *IL-6* pathogenic polymorphisms.

Keywords: SNP, polymorphism, IL-6, in silico

ÖZET

Tek nükleotid polimorfizmleri (SNPler) bireyler arasında; DNA'daki tek nükleotid değişiklikleri olarak bilinir. En yaygın SNP türü eş anlamlı olmayan SNPlerdir. İnterlökin-6 (IL-6) savunma mekanizması, bağışıklık tepkisi, hematopoez ve akut faz reaksiyonlarında çok çeşitli biyolojik fonksiyonlara sahip lenfoid ve lenfoid olmayan hücreler tarafından üretilen bir pleiotropik sitokindir. Bu çalışmada, *IL-6*'nın SNPleri NCBI dbSNP'de araştırılmıştır. Bu veritabanında *IL-6*'ya ait 2530 SNP vardır ve 227 tanesi yanlış anlamlı olarak sınıflandırılmıştır. PANTHER, SIFT ve PROVEAN biyoinformatik araçları, zararlı (D), muhtemelen zarar veren (PRD) veya mümkün olarak zarar veren (POD) SNPlerin belirlenmesi için kullanıldı ve bu SNPlerin *IL-6* protein ürünlerinin amino asitleri üzerindeki yapısal etkilerinin tespiti için Hope aracından yararlanıldı. SIFT ve PROVEAN skorları ile PANTHER koruma süreleri, bu 227 SNPnin 22'sinin D, PRD veya POD kategorisine sahip olduğunu göstermiştir. HOPE aracı sonuçlarına göre, mutant tiplerde rs370818995 (UniProt Protein id'leri: P05231, B5MCZ3 ve P05231) daha büyük boyutta aminoasitlere yol açmaktayken;

rs11544633 (UniProt protein id'leri: P05231, B5MCZ3, B4DNV3, P05231 ve B5MC21), rs199990564 (UniProt protein id'leri: P05231, B5MCZ3, B4DNV3, P05231 ve B5MC21), rs201371019 (UniProt Protein id'leri: P05231, B5MCZ3, B4DNV3, B4DNV3, P05231 ve B5MC21) daha küçük boyutta aminoasitlere neden olmaktadır. Ayrıca, rs370818995 (UniProt Protein id'leri: P05231, B5MCZ3 ve P05231) vahşi tipinden daha az hidrofobik forma neden olurken; rs199990564 (UniProt Protein id'leri: P05231, B5MCZ3, B4DNV3, P05231 ve B5MC21) ve rs201371019 (UniProt Protein id'leri: P05231, B5MCZ3, B4DNV3, P05231 ve B5MC21)'nin mutant tipleri ise daha hidrofobik yapılara yol açmaktadır. Çalışmada ilk kez *IL-6*'nın patojenik yanlış anlamlı SNPleri *in silico* yöntemlerle gösterilmiştir. Bu çalışmanın sonuçlarının, *IL-6*'nın patojenik polimorfizmlerinin etkileri ile ilgili gelecekteki çalışmalara yardımcı olabileceğini düşünüyoruz.

Anahtar Kelimeler: SNP, polimorfizm, IL-6, in silico

INTRODUCTION

The occurrence of a variation in more than 1% of a normal population is called genetic polymorphism. Although these variants are of low frequency, they are very important in some cases. Polymorphisms differ from mutations when they are found in the population as higher frequency variant alleles. A mutation is an event that creates a change in DNA or chromosome structure. Mutations can be causing diseases but polymorphisms are not (Das D. Et al., 2014). The most common polymorphism is Single nuclide polymorphism (SNP). SNPs are commonly referred to as single nucleotide changes in DNA among individuals (Wang DG. Et al., 1998). Different types of SNPs can alter the function or regulation and expression of a protein. The most common type is non-synonymous SNPs. The amino acid in the protein product is different in alleles of this type. Some SNPs are polymorphisms in the splice region and their exon content results in different variant proteins. Some SNPs are in the promoter region and affect the regulation and expression of the protein (Martines-Ariaz R. Et al., 2001). According to the data, SNPs are seen 1 time on an average of 1331 bases. Therefore, a person is heterozygous at the rate of 2.4 million bases (3.2 billion base x 1 difference / 1331bases =2.4 million). These rates vary between individuals and populations, as well as between chromosomes and loci. SNPs are seen at the rate of 1/1307 in Autosomal chromosomes, 1/2132 in the X chromosome, 1/6625 in the Y chromosome. There are two SNP types. Silent SNPs do not affect gene function and inherited traits. Most SNPs are included in this group (non-coding region SNPs). SNPs affect the protein function (coding region SNPs): It has two important effects: Direct effect (It changes amino acid sequence.) and indirect effect (It changes the function of the regulator sequence.).

Application areas of SNP data are diagnosis and risk profiling, candidate gene determination and mapping, polymorphism tests and epidemiological studies, pharmacogenetic and physiological genomics, environmental stimuli, and forensic genetics.

Cancer, stroke, heart disease, diabetes, and psychological problems are general diseases and they are affected by environmental factors as well as many genes. The purpose of finding the genes causing the disease; new approaches and preventive measures for therapeutic applications. In epidemiological and biomedical researches, SNP determination and comparison are performed in patients and healthy control groups from different populations. Cancer types, cardiovascular diseases, Alzheimer's, migraine, etc. SNP profiles specific to various diseases were made. By using these profiles, it is possible to scan for susceptibility to diseases. A method that will increase the effectiveness of using SNP is haplotype determination. Recent studies show that the vast majority of the genome is in the form of haplotype blocks. Only a few SNPs will be sufficient to mark these haplotype blocks and determine whether they are associated with a disease (Picoult-Newberg L. et al., 1999).

As a result of mutations, changes occur in the structure of DNA and chromosomes. The protein to be coded by these changes is also affected. As a result, diseases appear. One of the mutation types frequently seen in diseases is the base pair displacement mutations. Base pair displacement mutations appear as the replacement of one nucleotide to another nucleotide. These are called transitions and

transversions. Substitution mutations in genes encoding the protein can also be classified according to their effects on the amino acid sequence that creates the protein: If it results in the same amino acid production, it is called silent mutation. If a protein with the same effect is produced due to the production of different amino acid with the same chemical properties, it is called a neutral mutation. If it causes the coding of different amino acids and changes the protein and function, it is called missense mutation. If the codon, which normally expresses an amino acid, turns into a stop code, it is called a nonsense mutation.

Intercellular active substances, some of which are formed by lymphocytes (lymphokine), and some of them are formed by monocytes and macrophages, are called interleukins. Interleukin-6 (IL-6) is a pleiotropic cytokine produced by lymphoid and non-lymphoid cells with a wide range of biological activities such as acute phase response, inflammation, oncogenesis, hematopoiesis, and regulation of immune activity (Kishimoto, 1989). This cytokine initiates leukocyte infiltration; however, prolonged inflammation can turn into a devastating response that causes tissue fibrosis. IL-6 acts through the IL-6 receptor (IL-6R), which is detected in some cell types, such as hepatocytes or leukocytes. (Braunwald E., et al., 2018). IL-6 gene is found in chromosome 7p15-p21 in humans (Bowcock et al., 1988, Hirano et al., 1986), contains 5 exons and 4 introns (Yasukawa et al., 1987). IL-6 plays an important role in defense mechanisms, immune response, hematopoiesis, and acute phase reactions. On the other hand, expression of the IL-6 gene is particularly effective in the pathogenesis of various diseases such as autoimmune diseases, plasmacytoma-myeloma, and various chronic proliferative inflammatory diseases. Irregularity in IL-6 production plays a role in the formation of various diseases such as inflammation, autoimmune diseases, and malignancy (Weissenbach et al., 1980)

In the study, we used the online web-based databases and functional or structural effect detection tools: NCBI dbSNP (Sherry et al., 2001), SIFT (Vaser et al., 2016), PROVEAN (Choi & Chan, 2015), PANTHER coding SNPs (Tang & Thomas, 2016) and HOPE (Venselaar et al., 2010). These tools help to predict pathogenic missense SNPs and show us structural alterations for proteins of interest via the computational way.

MATERIALS AND METHODS

Extraction of The Nonsynonymous Missense SNPs of IL-6 From NCBI dbSNP

NCBI is a database that provides information about biomedical and genomic studies. dbSNP contains SNPs and data about their variation classes, clinical significances, etc. (https://www.ncbi.nlm.nih.gov/snp/). The first step of the study is the determination of the nonsynonymous missense SNPs of *IL-6*. There are 2530 SNPs for *IL-6*, and 227 of them are categorized as missense in this database. 227 SNPs were saved into a TXT file for further steps of the study.

Using The Web-Based Tools (SIFT, PROVEAN, and PANTHER) To Detect Pathogenic Missense SNPs

SIFT (https://sift.bii.a-star.edu.sg/), PROVEAN (http://provean.jcvi.org/index.php), and PANTHER (http://www.pantherdb.org/tools/index.jsp) online website tools provide predictions for the pathogenicity level of the SNPs of interest. SIFT dbSNP utilizes NCBI dbSNP data. In the study, the TXT file that contains missense SNPs of *IL-6* was uploaded to dbSNP rsIDs (Genome Tools, SIFT4G predictions). SIFT shows information about the species, chromosome, allele and amino acid changes, gene name, gene id (Ensembl, https://www.ensembl.org/index.html), transcript id, protein id, region, SIFT score, and SIFT prediction, etc. for SNPs. PROVEAN predicts the impacts of an amino acid substitution and indel on a biological function of a protein. PANTHER differs from SIFT and PROVEAN. This tool calculates the preservation time (million years) of an amino acid in the lineage which leads to the specific protein. The higher time of preservation creates a higher possible functionality. SIFT uses the list of rs ids (We used the TXT file mentioned before.) while PROVEAN and PANTHER need the information of protein sequences and amino acid variations (Figure 1).

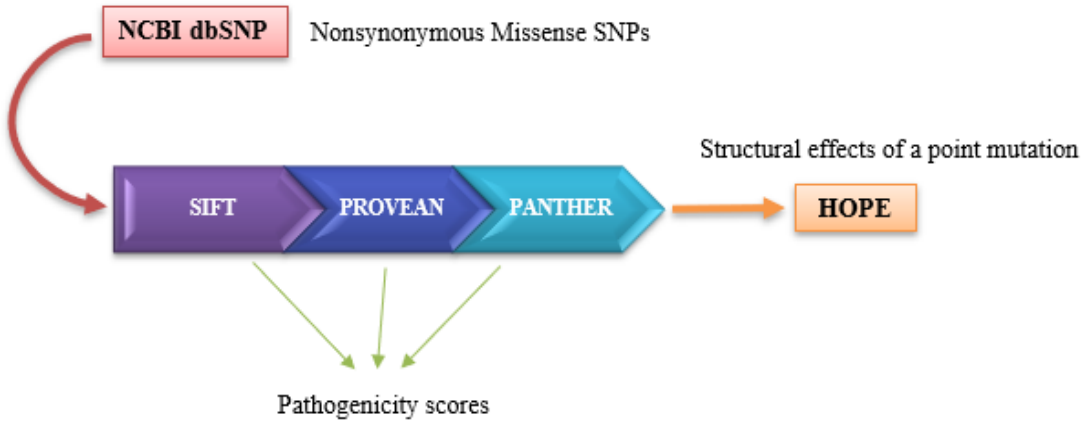


Figure 1. Online web-based tools for the analysis of pathogenic missense SNPs that were utilized in the study. The first step is the extraction of the rs ids of missense SNPs from the NCBI dbSNP, next steps lead to the prediction tools such as SIFT, PROVEAN, and PANTHER for the determining the pathogenicity level of the alteration. The last step is using HOPE for the detection of the structural alterations and functional impacts in the protein of interest.

Utilizing HOPE for the detection of the structural alterations and functional impacts

HOPE is the last step for the determining structural effects of the nonsynonymous missense SNPs. In this step, only the D, POD, and PRD SNPs (or we can call mutations) were searched on the web site (https://www3.cmbi.umcn.nl/hope/). HOPE uses the information of protein sequence, residue to mutate, and the mutation; and contain a collection of data coming from web services and databases which helps to create a document that serves animations, results, and figures for the mutation of interest.

RESULTS AND DISCUSSION

Evaluation of the data from SIFT, PROVEAN, and PANTHER online web-tools

Rs ids, protein ids, nucleotide (nt) changes, amino acid (aa) changes, regions (coding-CDS), PANTHER preservation times; and SIFT, PROVEAN scores of only D, POD, and PRD SNPs were shown in Table 1. SIFT, PROVEAN, and PANTHER indicated that 22 of 227 *IL-6* SNPs are likely pathogenic. In these web-based online tools detected that 10 of 227 SNPs are PRD, and D; also 12 of 227 SNPs are POD and D.

Table 1. Prediction of the pathogenic missense SNPs of *IL-6* using the SIFT, PROVEAN, and PANTHER web-based tools.

SNP ID	PROTEIN ID	NT CHANGE	AA CHANGE	REGION	PANTHER PRESERVATION TIME	SCORE		PREDICTION		
						SIFT	PROVEAN	PANTHER cSNPs	SIFT	PROVEAN
rs11544633	ENSP00000258743	T>C	L119P	CDS	455	0	-6.056	PRD	D	D
rs11544633	ENSP00000384928	T>C	L96P	CDS	455	0	-6.356	PRD	D	D
rs11544633	ENSP00000385718	T>C	L43P	CDS	455	0	-6.810	PRD	D	D
rs11544633	ENSP00000385675	T>C	L119P	CDS	455	0	-6.056	PRD	D	D
rs11544633	ENSP00000385227	T>C	L119P	CDS	455	0	-5.922	PRD	D	D
rs11544633	ENSP00000385043	T>C	L43P	CDS	455	0	-6.489	PRD	D	D

rs199990564	ENSP00000258743	A>G	E87G	CDS	220	0.003	-2.980	POD	D	D
rs199990564	ENSP00000384928	A>G	E64G	CDS	220	0.003	-3.246	POD	D	D
rs199990564	ENSP00000385718	A>G	E11G	CDS	220	0.003	-3.478	POD	D	D
rs199990564	ENSP00000385675	A>G	E87G	CDS	220	0.003	-2.946	POD	D	D
rs199990564	ENSP00000385227	A>G	E87G	CDS	220	0.003	-2.946	POD	D	D
rs199990564	ENSP00000385043	A>G	E11G	CDS	220	0.003	-3.280	POD	D	D
rs201371019	ENSP00000258743	A>G	Y128C	CDS	324	0.001	-7.005	POD	D	D
rs201371019	ENSP00000384928	A>G	Y105C	CDS	324	0.001	-7.105	POD	D	D
rs201371019	ENSP00000385718	A>G	Y52C	CDS	324	0.001	-7.827	POD	D	D
rs201371019	ENSP00000385675	A>G	Y128C	CDS	324	0.001	-7.005	POD	D	D
rs201371019	ENSP00000385227	A>G	Y128C	CDS	324	0.001	-6.938	POD	D	D
rs201371019	ENSP00000385043	A>G	Y52C	CDS	324	0.001	-7.338	POD	D	D
rs370818995	ENSP00000258743	T>C	C72R	CDS	455	0	-8.468	PRD	D	D
rs370818995	ENSP00000384928	T>C	C49R	CDS	455	0	-8.418	PRD	D	D
rs370818995	ENSP00000385675	T>C	C72R	CDS	455	0	-8.468	PRD	D	D
rs370818995	ENSP00000385227	T>C	C72R	CDS	455	0	-8.068	PRD	D	D

Structural Impacts of The Mutations by HOPE

Project HOPE software has shown the possible effects of missense SNP in the IL-6 gene. While 3 wild type proteins become less hydrophobic form, 10 mutant type proteins become more hydrophobic form. 10 wild type residues which Uniport IDs are P05231, B5MCZ3, B4DNV3, P05231, B5MC21, P05231, B5MCZ3, B4DNV3, P05231, B5MC21 became from less hydrophobic to more hydrophobic and 5 wild type residues which Uniport IDs are P05231, B5MCZ3, B4DNV3, P05231, B5MC21, their charges change from negative to neutral. Three SNPs (UniProt IDs: P05231, B5MCZ3, P05231) became more hydrophobic to less hydrophobic and their charges change from neutral to positive (Table 2).

According to the results, we determined that 12 mutant residues are located near a highly conserved position. This situation may cause the pathogenic effects of SNPs. It has been observed that loss of the cysteine bond in 2 proteins. Together with the loss of the cysteine bond, the differences between the old and new residue can cause destabilization of the structure.

In the wild-type residue of rs370818995 (P05231) and rs370818995 (B5MCZ3) are annotated in UniProt to be involved in a cysteine bridge, which is important for the stability of the protein. (Table 2).

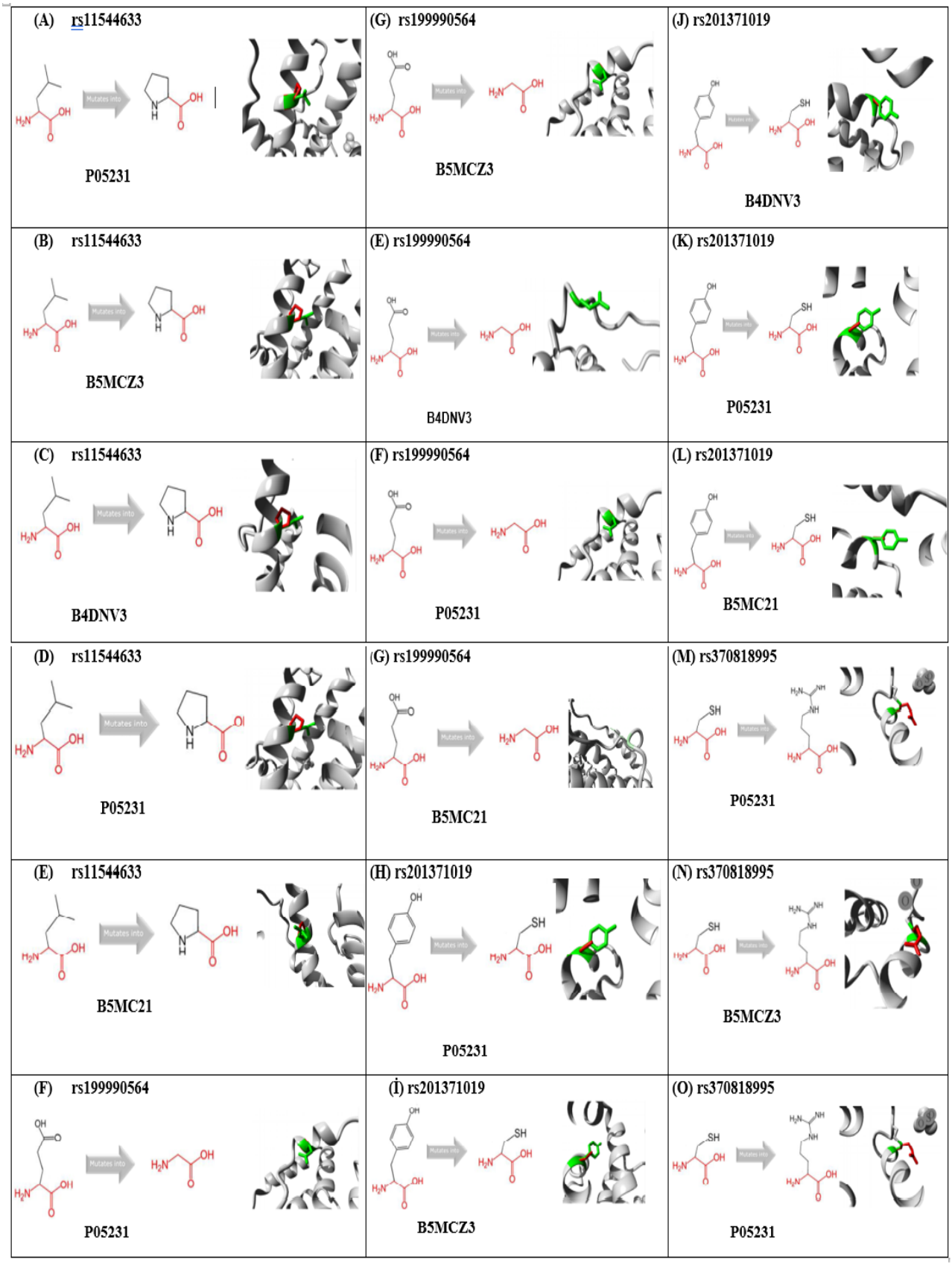
Table 2. The total extracted results from the HOPE web-based tool.

			WILD-TYPE PROPERTIES			**MUTANT TYPE PROPERTIES**		
SNP ID	**UNIPROT ID**	**AMINO ACID CHANGE**	**SIZE**	**CHARGE**	**PHYSICAL STATE**	**SIZE**	**CHARGE**	**PHYSICAL STATE**
rs11544633	P05231	L119P	>	No information	No information	<	No information	No information
rs11544633	B5MCZ3	L96P	>	No information	No information	<	No information	No information
rs11544633	B4DNV3	L43P	>	No information	No information	<	No information	No information
rs11544633	P05231	L119P	>	No information	No information	<	No information	No information

rs11544633	B5MC21	L43P	>	No information	No information	<	No information	No information
rs199990564	P05231	E87G	>	Negative	Less hydrophobic	<	Neutral	More hydrophobic
rs199990564	B5MCZ3	E64G	>	Negative	Less hydrophobic	<	Neutral	More hydrophobic
rs199990564	B4DNV3	E11G	>	Negative	Less hydrophobic	<	Neutral	More hydrophobic
rs199990564	P05231	E87G	>	Negative	Less hydrophobic	<	Neutral	More hydrophobic
rs199990564	B5MC21	E11G	>	Negative	Less hydrophobic	<	Neutral	More hydrophobic
rs201371019	P05231	Y128C	>	No information	Less hydrophobic	<	No information	More hydrophobic
rs201371019	B5MCZ3	Y105C	>	No information	Less hydrophobic	<	No information	More hydrophobic
rs201371019	B4DNV3	Y52C	>	No information	Less hydrophobic	<	No information	More hydrophobic
rs201371019	P05231	Y128C	>	No information	Less hydrophobic	<	No information	More hydrophobic
rs201371019	B5MC21	Y52C	>	No information	Less hydrophobic	<	No information	More hydrophobic
rs370818995	P05231	C72R	<	Neutral	More hydrophobic	>	Positive	Less hydrophobic
rs370818995	B5MCZ3	C49R	<	Neutral	More hydrophobic	>	Positive	Less hydrophobic
rs370818995	P05231	C72R	<	Neutral	More hydrophobic	>	Positive	Less hydrophobic

Only cysteines can make these type of bonds, the mutation causes loss of this interaction and will have a severe effect on the 3D-structure of the protein. Together with the loss of the cysteine bond, the differences between the old and new residue can cause destabilization of the structure. 12 mutant residues which are rs11544633, rs11544633, rs11544633, rs11544633, rs11544633, rs199990564, rs199990564, rs199990564, rs199990564, rs199990564, rs201371019, rs201371019 and their Uniport IDs are respectively P05231, B5MCZ3, B4DNV, P05231, B5MC21, P05231, B5MCZ3, B4DNV3, P05231, B5MC21, P05231, B5MCZ3, are located near a highly conserved position. This situation may cause the pathogenic effects of SNPs. All situations that we mentioned before are affecting intra- and inter-molecular interactions for example protein and protein or protein and other molecules (Table 3).

Table 3. The HOPE tool results that are based on structural alterations in the related amino acids.



CONCLUSION

The online web-based tools that predict the impact of SNPs of a specific gene are very important for the *in silico* analysis of mutations and variations. These tools save time, money and provide the almost accurate results and pre-thoughts before the *in vitro* or *in vivo* studies. Today, in all world, one of the

most deadly threats for people is cancer. IL-6 polymorphisms can give us information about cancer prediction and prognosis. Associations between the polymorphisms and diseases are important for early detection and improving the personalized treatments.

ACKNOWLEDGMENTS

The study was presented in EURO ASIA 7th INTERNATIONAL CONGRESS ON APPLIED SCIENCES (August 2020 on 21-22; Trabzon, Turkey). No financial support was received from any institution or organization by the authors.

REFERENCES

Bowcock, AM., Ray, A., Erlich, H., Sehgal, PB. 1989. "Rapid detection sequencing of alleles in the 3' flanking region of the interleukin-6 gene", Nucleic Acids Res., 17, 6855-64.

Braunwald, E. (2018). Cell-based therapy in cardiac regeneration: An overview. Circulation Research. Lippincott Williams and Wilkins. https://doi.org/10.1161/CIRCRESAHA.118.313484

Das D and Haloi A. Vitamin B12 Gene Polymorphisms and Chronic Diseases. J Nutr Disorders Ther. 2014; 4:149.

Hirano, T. 2014. "Revisiting the 1986 molecular cloning of interleukin 6. Frontiers in Immunology, 5(SEP)", https://doi.org/10.3389/fimmu.2014.00456

Kishimoto T.1989. The Biology of Interleukin-6 "The biology of interleukin-6. Blood", https://doi.org/10.1182/blood.v74.1.1.bloodjournal7411

Martinez-Arias R., Calafell F., Mateu E., Comas D., Andres A., Bertranpetit J., Sequence variability of a human pseudogene, Genome Res. 11 (2001) 1071– 1085

Picoult-Newberg L., Ideker T.E., Pohl M.G., Taylor S.L., Donaldson M.A., Nickerson D.A., Boyce-Jacino M., Mining SNPs from EST databases, Genome Res. 9 (1999) 167–174

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001 Jan 1;29(1):308-11.

Vaser, R., Adusumalli, S., Leng, S. *et al.* SIFT missense predictions for genomes. *Nat Protoc* **11,** 1–9 (2016). https://doi.org/10.1038/nprot.2015.123

Tang H, Thomas PD. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics*. 2016;32(14):2230-2232. doi:10.1093/bioinformatics/btw222

Wang D.G., Fan J.B., Siao C.J., Berno A., Young P., Sapolsky R., Ghandour G., Perkins N., Winchester E., Spencer J., Kruglyak L., Stein L., Hsie L., Topaloglou T., Hubbell E., Robinson E., Mittmann M., Morris M.S., Shen N., Kilburn D., Rioux J., Nusbaum C., Rozen S., Hudson T.J., Lander E.S., Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome, Science 280 (1998) 1077–1082.

Weissenbach, J., Chernajovsky, Y., Zeevi, M., Shulman, L., Soreq, H., Nir, U., … Revel, M. (1980). Two interferon mRNAs in human fibroblasts: In vitro translation and Escherichia coli cloning studies. Proceedings of the National Academy of Sciences of the United States of America, 77(12 II), 7152–7156. https://doi.org/10.1073/pnas.77.12.7152

Yasukawa, K., Hirano, T., Watanabe, Y., Muratani, K., Matsuda, T., Kishimoto, T. 1987. "Structure and expression of human B cell stimulatory factor 2 (BSF-2/IL-6) gene", 6, 2939–45.

Yongwook Choi, Agnes P. Chan, PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels, *Bioinformatics*, Volume 31, Issue 16, 15 August 2015, Pages 2745–2747, https://doi.org/10.1093/bioinformatics/btv195