J Nanostruct 15(1): 122-133, Winter 2025

RESEARCH PAPER

Analysis the Structure and Function of nsSNP of Interleukin 4 Gene by In-Silico Research

*Hamsa Faisal Najm [1], Osama Abdulmunem K. [2], Dunya Jawad Ridha [3], Mohammed I. Jameel [4,5]**

[1] College of Dentistry, Al-Bayan University, Baghdad, Iraq

[2] Department of Medical Laboratory Technology, College of Health and Medical Techniques, AL-Bayan University, Baghdad, Iraq

[3] Department of Medical Laboratory Techniques, University of Dijlah, Baghdad, Iraq

[4] Department of Medical Microbiology, Faculty of Science and Health, Koya University, Koya KOY45, Kurdistan Region, Iraq

[5] Department of Biomedical Sciences, College of Applied Sciences, Cihan University-Erbil, Erbil, Kurdistan Region, Iraq

ARTICLE INFO

Article History:
Received 16 October 2024
Accepted 28 December 2024
Published 01 January 2025

Keywords:
Cancer
HOPE
IL-4
nsSNP
Protein

ABSTRACT

Interleukin-4 is a type II inflammatory response cytokine. It is an essential component of the inflammatory response caused by an invasive allergen or parasite. Also, IL-4plays a key role in the up growth of inflammation and asthma by boost RI expression of the CF epsilon in B cells, mast cells and basophils, promoting the survival and proliferation of mast cells, and inducing chemotaxis of mast cells, basophils and eosinophils. We used six bioinformatics tools in this study (SIFT, pMut, PANTHER, Polyphen-2, PHDSNP, and SNPs&GO) prediction of disease-sensitive non-synonymous SNPs of IL4, IL-4's seven nsSNPs (V53A, A118G, M144T, G2D, L110R, N113Y, and C123R) are predicted to be potentially harmful. The 3D structure of the abnormal protein was modelled by (HOPE), and the interaction between protein/protein was evaluated by STRING. While, Kaplan–Meier Plotter showed that the deregulation of IL4 expression affects the survival rate of ovarian cancer patients. Thus, IL4 may be significant as a gene marker for certain cancers. Study resolved harmful nsSNPs of IL-4, which could engage to the degradation of IL4 proteins and eventually lead to IL4 related diseases.

How to cite this article
Najm H, Abdulmunem K. O., Ridha D, Jameel M. Analysis the Structure and Function of nsSNP of Interleukin 4 Gene by In-Silico Research. J Nanostruct, 2025; 15(1):122-133. DOI: 10.22052/JNS.2025.01.012

INTRODUCTION

The immune-system consists of various immune cells answerable for oversight and obtaining cells Eliminating unknown pathogens and foray microorganisms. Immune cells can work promptly or through synthesis molecules that can induce B-cells, NK-cells, T-cells and rest immune cells [1, 2]. Many immune cells depend on activation and differentiation. Various types of Interleukins (ILs). Interleukin is a subset of a big group Naturally show cytokines are released mainly by some immune cells have a function to endogenous threats, heat also inflammation. Acting as cell messengers by binding them. High-affinity receptors on the cell surface [3, 4]. play IL Important role in both

* Corresponding Author Email: *mohammed.isam@koyauniversity.org*

adaptive immune system the adaptive and innate also modulate cell behavior [3, 5]. Interleukin-4 is a characteristic cytokine of 2nd type inflammatory response. It plays a role in the inflammatory reaction caused by invading parasites or allergens. For cell source of IL-4 has been studied in depth and with CD4 (Tcells), basophils, eosinophils, the appropriate stimulants of ILC2 cells produce IL-4 [6, 7]. The IL-4 (and IL-5) genetic locus is known as Th2 (cytokine) locus, lying in the 5th chromosome in homo sp. and in the 11th in mice, controlled by (LCR) of the radon gene [8, 9]. The LCR in CD4 Tcells is requisite for the production of IL4*in-vivo* [10]. However, the production of the two cytokines is not the same: the production of IL-4 depends on calcineurin. When cells are stimulated appropriately, the T2 cytokine locus' LCR is epigenetically modified so that transcription factors can be accessed to DNA and then converted to these cytokines. The complex regulations have recently been reviewed in detail. [8] Interestingly, the multimorphism in(humans' DNA)methylation and gene expression5q31 is influenced by the polymorphism of DNase I hypersensitive sites (RHS)7, which correlates with mice's findings, and IgE levels are subsequently reported at the population level. [11]. IL4is a type I glycosylated cytokine with (3) sulfide-bridges in the chain, with a binding structure of 4alpha-helix. It is produced mainly by T cells, natural killer T cells and eosinophiles. IL-4 initiates signal transmission by two different receptor complexes: the first hematopoietic cell receptor or the second non-hematopoietic cell receptor [12].

MATERIALS AND METHODS

Extract of IL4 nsSNPs

The total SNPs for IL4and its protein sequence (UNIPROT-IDP05112) are obtained from the dbSNP database of NCBI (ncbi.nlmnih.gov/snp) and the Uniprot Knowledge Base database (UNIPROT.org). A total of4293 SNPsof different functional classes were mapped in the IL4gene sequence. Of the 4293 SNPs, 152 are (nsSNPs) appear in the encoder region, leading to misunderstood or non-remarkable mutations, which have an impact on protein annotation. Our study revealed nsSNP in the coding area of the IL4protein.

Prediction of the deleterious nsSNPs

A 7 different tools to predict the deleterious effects of nsSNPs Sorting Intolerant From Tolerant (http://sift.bii.astar.edu.sg) [13], Predictor of human Deleterious Single Nucleotide Polymorphisms (PhD-SNP; http://snps.biofold.org/ phd-snp/phd-snp.html) [14] PMut(http://mmb.irbbarcelona.orgPMut)Polymorphism Phenotypingv2(PolyPhen-2;http://genetics.bwh.Harvard.edu/pph/) [15], Protein analysis through assessment relationship (http://www.pantherdb.org/tools/csnpScoreFormjsp) [16] and SNPs and GO (http://snps.biofold.org/snps-and- go/snps- and-go.html) [17]. SIFT predicts the effect of substitution of amino acids on the protein function based on sequence homology and physical properties of substituted amino acids [13]. Polyphenol-2predicts the effects of substitute amino acids on protein annotation depend on physical properties and comparative properties [15]. PROVEAN is a support vector machine server that predicts whether substituted amino acids affect protein functions. Tools such as PANTHER, SNPs and GO, PHD-SNP, and PMut have been used to predict whether a single nucleotide polymorphism is associated with disease [18, 19]. nsSNPs predicted to be harmful by at least four of the above-mentioned tools in the silico tool were considered to be high-risk nsSNPs and selected for analysis.

Analyzing protein stability due to mutations

I-Mutant3.0 (gpcr2.biocomp.unibo.itcgi/predictors/I-Mutant3.0I-Mutant3.0cgi), Mupro (Mupro.proteomic.icuci. edu), also INPS-MD (inpsmdbiocomp. the uniboit) tool was used to assess the stability ofIL-4 proteins after mutation. I-Mutant3.0 is a support vector machine-based prediction that knows the degree of protein instability and determine the G value0(kcal/mol). The G-value is the difference between the Gibbs free energy value of a mutated and the wild protein. If the G value is less than 0 and the G value is greater than 0 indicates that the modifications are caused by reduced protein stability, and if the G value is greater than 0 and the protein is more stable [20]. While MUpro, a large number uses of mutation data sets, based on SVM, neural network and machine learning methods. The 3rd tool is INPS-MD (negative impact of non-synonymous mutations on protein stability-multidimension), which uses sequence descriptors to calculate G values using support vector regression (SVR). Both MUpro and INPS-MD measure G to estimate stability, and the G cut

value is identical to I-Mutant30 [4, 21]. IL-4 protein sequences, wild-type amino acids, and alternative amino acids have been used as inputs to predict mutation effects on protein stability in the above-mentioned tools [20].

Identifying the effects of mutations on the structural and functional characteristics of proteins

To sort out the substitutions of amino acids associated with diseases or neutral amino acids in protein sequences, MutPred2's website server (http://mutpr ed.mutdb.org) further investigated the commonly predicted mutations. It is a machine learning method that combines genetic and molecular data to foretell if substituted amino acids would be harmful. Additionally, it foretells the disease's molecular origins. [22].

Estimating how high risk nsSNPs will affect protein structure molecularly

An automatic mutant analysis service called Project HOPE was created to examine the structural and biochemical impacts of point mutations in protein sequences [23]. The main structure of IL-4 proteins from seven SNPs (rs IDs) from the Protein Data Bank (https://www.rcsb.org/pdb/) has been submitted to HOPE. HOPE collects structural information from a number of sources and predicts the 3D structure of mutated proteins and explains such changes (both in protein structure and function).

Protein–protein interaction prediction

The interactions between proteins and proteins are studied to identify and explain all functional interactions between cell proteins. The online STRING0database (STRING, http://stringdb.org/) uses to predict the interactions between 2 proteins. [24].

Kaplan–Meier plotter analysis

The European genome sequences (EGA), the cancer genome sequences (TCGA), and the genome expression omnibus (GEO) datasets are used by the Kaplan-Meier plotter database (http://kmplot.com/analysis) to provide non-relapse and overall survival (OS) information as well as meta-analysis-based discovery and biomarker assessment for cancer patients. To estimate the death period is the goal of this analysis, an event that occurs in everyone, and when used to inform clinical decisions, health policies, and resource allocations [25]. The algorithm investigates the potential effects of genes (mRNAs, miRNAs, proteins) on cancer survivors (including breast, lung, gastric and ovarian cancers) by microarray0gene expression data from 21cancers [26]. Using the IL-4 gene Affymetrix ID, a complete survival analysis of all cancer patients was performed. The hazard ratio (HR) of the 95% confidence intervals and the log-

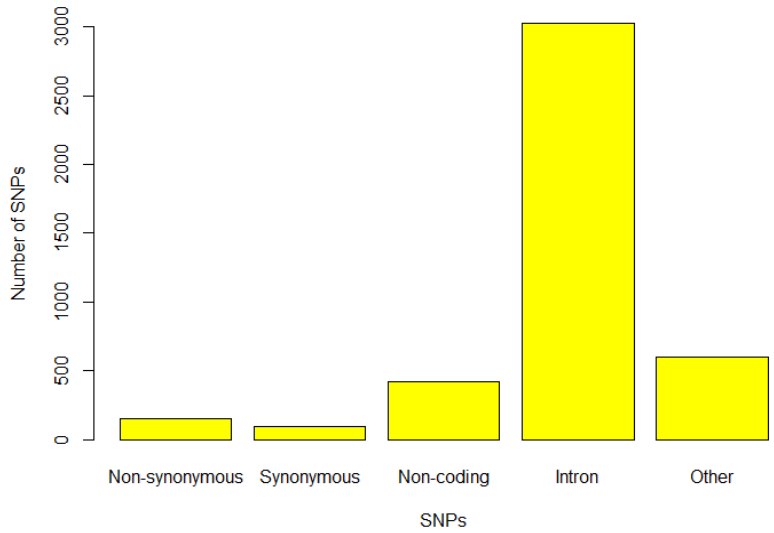

Fig. 1. bar plot showing the different numbers of SNPs in IL4gene based on db-SNP. (Missense-SNPs: 152, Synonymous-SNPs: 95, non-coding-SNPs: 425, Intron-SNPs: 3023, Other-SNPs: 598)

rank P-value were listed and displayed in the Table 1.

RESULTS AND DISCUSSION

nsSNPs retrieved from dbSNP database

According to the Db SNP database, the human IL-4 gene has 4293 SNPs, 152 of which are nsSNPs/ missenses (4%), 3023 are intronic SNPs (70%), 95 are synonymous SNPs (2%), 425 are not coding (10%), and the rest are other types (Fig. 1). We chose only nsSNPs for our investigation.

Prediction and analysis of deleterious nsSNPs

The functional impact of nsSNP was evaluated by evaluating the importance of amino acids it changes. An analysis dataset of 152 polymorphic inputs was used. The structural and functional effects of harmful SNPs on IL-4 proteins were tested by various computational tools. A graphical representation of harmful nsSNP predicted by six different computational tools. From 152, we found only 36 data information. Then 36 nsSNPs of IL-4 were submitted to the SIFT algorithm. According to the SIFT result, 10 nsSNPs with TI scores of 0.05 are predicted to be intolerant. The PhD-SNP and PMut tools proposed 11 nsSNP and 0 nsSNP as "diseases". In addition, PolyPhen-2 predicted 14 nsSNPs to be "possibly harmful" and 22 nsSNPs to be "possibly harmful". In addition, PANTHER_ PSEP described 18 SNPs as "dangerous". 10 SNPs are likely to cause damage, while the remaining 8 SNPs are likely to cause damage. The SNPs and GO tool predicts that five nsSNPs are associated with various types of diseases (Fig. 1). Finally, at least four of the instruments analyzed in silicate predict harmful/damaging/associated diseases that may

Table 1. list of nsSNP with different computation tools

SNP ID	Amino acid-Change	SIFT	PANTHER	PhD-SNP	PMut	PolyPhen-2	SNPs&Go
rs4986964	C27R	TOLERATED	D	D	N	N	D
rs55743996	A59G	TOLERATED	D	N	N	N	N
rs56279116	V53I	TOLERATED	D	N	N	N	N
rs71645915	L120S	TOLERATED	N	N	N	N	N
rs79908535	R109Q	TOLERATED	N	N	N	D	N
rs139863211	V53A	DELETERIOUS	D	N	N	D	D
rs143377308	A92A	TOLERATED	N	N	N	N	N
rs145068648	T142R	TOLERATED	N	N	N	N	N
rs146713238	D111N	TOLERATED	N	N	N	N	N
rs148396041	G19D	TOLERATED	N	N	N	N	N
rs149147538	K150M	DELETERIOUS	N	N	N	D	D
rs149950065	A118G	DELETERIOUS	D	N	N	D	D
rs199591873	C48A	TOLERATED	D	D	N	D	N
rs199760133	E50K	TOLERATED	D	D	N	N	N
rs199819530	H82D	TOLERATED	D	D	N	N	N
rs199921056	Q6Q	TOLERATED	D	N	N	N	N
rs199929962	M144T	TOLERATED	D	D	N	D	D
rs199936687	K85K	TOLERATED	N	N	N	N	N
rs200011844	S131R	TOLERATED	N	N	N	N	N
rs200140092	K26E	TOLERATED	N	N	N	N	N
rs200172265	K85K	TOLERATED	N	N	N	N	N
rs200393431	K26N	TOLERATED	N	N	N	N	N
rs200433079	L117R	DELETERIOUS	N	D	N	N	D
rs200549061	G2D	DELETERIOUS	D	D	N	D	D
rs201145789	T132M	TOLERATED	D	N	N	D	D
rs201446269	H100Q	TOLERATED	N	N	N	D	N
rs201594156	L110R	DELETERIOUS	D	D	N	D	D
rs201689675	V22D	DELETERIOUS	N	D	N	N	D
rs201694257	A58I	TOLERATED	D	N	N	D	D
rs201843425	H23H	TOLERATED	D	N	N	N	N
rs202011365	A73V	TOLERATED	N	N	N	N	N
rs202231191	N113Y	DELETERIOUS	D	D	N	D	D
rs373334025	R109W	DELETERIOUS	N	N	N	D	D
rs376367511	C123R	DELETERIOUS	D	D	N	D	D
rs377174102	H83H	TOLERATED	D	N	N	N	N
rs377297730	R88R	TOLERATED	N	N	N	N	N

be further investigated (Table 2).

Computer analysis using the six mentioned instruments revealed seven highly harmful nsSNPs in the IL-4gene.0Of the seven nsSNPs, four are nsSNPs, four are nsSNPs. (e.g., rs200549061 G2D, rs201594156 L110R, rs202231191 N113Y, and rs376367511 C123R) were predicted deleterious unanimously by at least 5 of the employed tools, and other three nsSNPs (rs139863211 V53A, rs149950065 A118G, and rs199929962 M144T) were predicted deleterious by the at least four computational tools.

Identification of functional and structural modifications of IL-4 predicted by MutPred2

The seven selected nsSNPs predicted to be harmful from previous steps have been submitted to MutPred2's web server. The resulting probability scores, g and p values are shown in Table 3. It helps predict the cause of molecular change that could affect the phenomenon. The annotation alterations predicted include- Altered Disordered interface, Loss of Acetylation at K1 8; Altered DNA binding, Altered Stability, altered ordered interface, Loss of Disulfide linkage at C123; and Altered Transmembrane Protein. The output of the MutPred2 tool is a general score(g), which represents the average score of all neurons in the MutPred2. The threshold value of the g score is 0.50. For some mutations, a value of g-score greater than 0.50(g > 0.50) indicates pathogenicity. The scores with a value g > 0.5 and a value p 0.05 are called action hypotheses, while the scores with a value g > 0.75 and a value p 0.05

Table 2. high risk nsSNP predict in*silico*programs

SNP ID	Amino acid change	SIFT	PANTHER	PhD-SNP	PMut	PolyPhen	SNPs&Go
rs139863211	V53A	DELETERIOUS	D	N	N	D	D
rs149950065	A118G	DELETERIOUS	D	N	N	D	D
rs199929962	M144T	TOLERATED	D	D	N	D	D
rs200549061	G2D	DELETERIOUS	D	D	N	D	D
rs201594156	L110R	DELETERIOUS	D	D	N	D	D
rs202231191	N113Y	DELETERIOUS	D	D	N	D	D
rs376367511	C123R	DELETERIOUS	D	D	N	D	D

Table 3. Functional0&0structural0modifications of IL-4 predicted by0MutPred2.

SNPs	Actionable	g-value	p-value	Probability
V53A	-	0.378	-	-
A118G	-	0.149	-	-
M144T	-	0.190	-	-
G2D	-	0.267	-	-
L110R	Altered Disordered interface	0.589	4.0e-03	0.39
	Loss of Acetylation at K108		0.05	0.19
	Altered DNA binding		0.04	0.16
	Altered Stability		0.03	0.12
N113Y	-	0.372	-	-
C123R	Altered ordered interface	0.733	0.02	0.30
	Loss of Disulfide linkage at C123		1.1e-03	0.30
	Altered0Transmembrane0Protein		0.04	0.10

Table 4. Protein stability change0prediction using I-Mutant 3.0, MUpro and INPS-MD.

AA Mutation	I-Mutant 3.0		MUpro		IPNS-MD	
	Stability	ΔΔG(kcal/mol)	Stability	ΔΔG(kcal/mol)0	Stability	ΔΔG(kcal/mol)
V53A	Decrease	-0.62	Decrease	-1.30	Decrease	-2.60
A118G	Decrease	-2.39	Decrease	-1.06	Decrease	-1.40
M144T	Decrease	-0.47	Decrease	-1.64	Decrease	-1.39
G2D	Decrease	-0.37	Decrease	-0.63	Decrease	-0.91
L110R	Decrease	-1.20	Decrease	-2.11	Decrease	-1.32
N113Y	Decrease	-0.05	Decrease	-1.42	Increase	0.09
C123R	Decrease	-1.21	Decrease	-1.85	Decrease	-1.36

are called confident hypotheses. In MutPred2 predictions, the replacement of L110R and C123R has a g value of greater0than 0.5 and a p value of less than 0.05(Table 3). These predicted data provide solid evidence that several nsSNPs could be involved in structural and functional changes in the IL-4 protein.

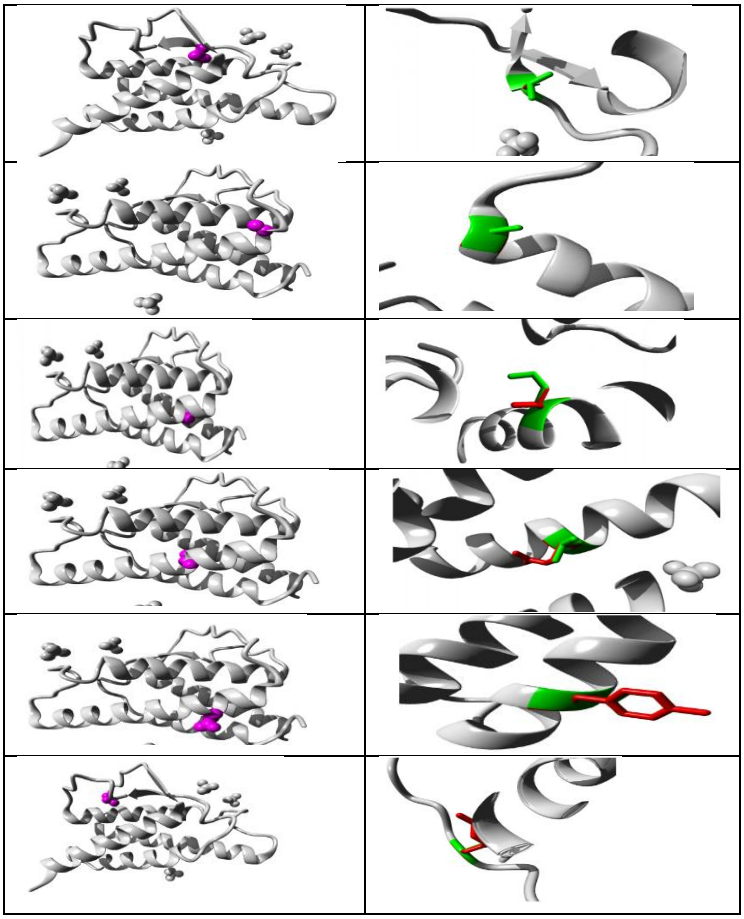

Fig. 2. Project HOPE predicts the structure of 3D models of IL-4proteins. In this case, the violet color of the ribbon diagram represents the mutation site. The green and red represent mutated native and mutated amino acids, respectively.

Table 5. The results of the properties of wild and mutant amino acids obtained from the Project Hope software

Amino Acid change	Wild type	Amino acid		Mutant type	Amino acid	
	Size	Charge	Hydrophobicity	Size	Charge	Hydrophobicity
V53A	Larger	-	-	Smaller	-	-
A118G	Larger	-	more hydrophobic	Smaller	-	Less hydrophobic
M144T	Larger	-	more hydrophobic	Smaller	-	Less hydrophobic
G2D	Small	Neutral	More hydrophobic	Big	Negative	Less hydrophobic
L110R	Small	Neutral	more hydrophobic	Big	Positive	Less hydrophobic
N113Y	Small	-	Less hydrophobic	Big	-	more hydrophobic
C123R	Small	Neutral	more hydrophobic	Big	Positive	Less hydrophobic

Table 6. Effects of amino acid changes on IL-4protein from Project Hope

Amino acid Change	Structure	Domain	Conservation	Loss H bond or other
V53A		The mutated residue is located in a domain that is important for binding of other molecules. The mutated residue is incontact with residues in another domain. It is possible that the mutation disturbs these contacts.	Highly conserved region	
A118G	The mutation introduces a glycine at this position. Glycines are very flexible and can disturb the required rigidity of the protein at this position	The mutated residue is located in a domain that is important for binding of other molecules and in contact with residues in a domain that is also important for binding. The mutation might disturb the interaction between these two domains and as such affect the function of the protein.	Highly conserved region	loss of hydrophobic interactions in the core of the protein
M144T	In the 3D-structure can be seen that the wild-type residue is located in an α-helix. The mutation converts the wild-type residue in a residue that does not prefer α-helices as secondary structure.	The mutated residue is located in a domain that is important for binding of other molecules and in contact with residues in a domain that is also important for binding. The mutation might disturb the interaction between these two domains and as such affect the function of the protein.	Highly conserved region	Loss of hydrophobic interactions in the core of the protein
G2D	The mutation is located within the signal peptide. This sequence of this peptide is important because it is recognized by other proteins and often cleaved of to generate the mature protein.	The mutated residue is located in a domain that is important for binding of other molecules. Mutation of the residue might disturb this function.	Highly conserved region	
L110R	In the 3D-structure can be seen that the wild-type residue is located in an α-helix. The mutation converts the wild-type residue in a residue that	The mutated residue is located in a domain that is important for binding of other molecules. Mutated residues are in contact with residues from another region. Mutations may interfere with these contacts.	Highly conserved region	loss of hydrophobic interactions in the core of the protein

	does not prefer α-helices as secondary structure.			
N113Y		Mutated residues are located in important areas for binding other molecules, and are in contact with residues in important areas for binding. Mutations may disturb the interaction between the two fields and therefore affect the function of the proteins.	Highly conserved region	
C123R		The mutated residue is located in a domain that is important for binding of other molecules and in contact with residues in a domain that is also important for binding. The mutation might disturb the interaction between these two domains and as such affect the function of the protein.	Highly conserved region	loss of hydrophobic interactions in the core of the protein

The impact of predicted deleterious mutations on IL-4protein stability

The seven expected nsSNPs0were0further analyzed by I-Mutant 3.0, INPS- MD, and Mupro using free energy comparison tools to analyze protein stability. The structure stability of six of the seven nsSNPs (V53A, A118G, M144T, G2D, L110Y and G123R) was completely reduced with the three analysis tools. The three variants of A118G, L110Y, G123R, and I168T showed a single-sided decrease in the delta G value of G to -1kcal/mol. The remaining three variants, V53A, M144T, and G2D, have shown a G value below zero, which is expected to change the structure and function of proteins by reducing their stability (Table 4).

Protein structure analysis

The three-dimensional models of seven mutant IL-4 proteins were created by the Hope Project (see Fig. 2). The HOPE project simulates the structural characteristics of amino acid residue replacement in native proteins. Furthermore, the HOPE project showed that physical chemical properties such as size, charge, and hydrophobicity were different between wild and mutant amino acids, as shown in (Table 5). All seven predicted nsSNPs caused change is in amino acid size. In addition to G2D mutations, L110R mutations, N113Y mutations, and C123R mutations, the remaining three mutant amino acids are smaller than the wild-type mutations. In seven nsSNPs, three nsSNPs (G2D, L110R, and C123R) altered amino acid load in mutant variants. Six nsSNPs (A118G, M144T, G2D, L110R, N113Y, and C123R) also cause changes in the water resistance of amino acids (Table 5).

Further analysis with the HOPE project showed that all seven mutations occurred in the domain area. In addition, it was found that four mutations (A118G, M144T, L110R and C123R) caused the loss of the hydrogen bond interaction, and three mutations caused the loss of the hydrophobic interaction. It is interesting to note that all seven mutations are located in preserved regions that may affect the structure and function of the IL-4 protein (Table 6).

Protein–protein interaction analysis

The STRING server result showed that Interleukin-4protein interacts with 10 proteins including, interleukin-13receptor alpha subunit-1(IL-13RA1), interleukin-4 receptor alpha subunit (IL-4R), tumor necrosis factor (TNF), interleukin-6 (IL-6), Interleukin-8(CXCL8), C–C motif chemokine-2(CCL2), signal0transducer and activator of transcription- 6 (STAT6), interleukin-1 beta (IL-1B), interleukin-1 alpha (IL-1A), and cytokines receptor common subunit gamma (IL-2RG), (Fig. 3).

The clinical correlation between IL-4deregulation and survival rates of patients with different types of cancers

In this phase, we tried to link IL-4 gene deregulation to a clinical database in order to infer the potential functional consequences of IL-4deregulation in cancer patients. Kaplan–Meier Plotter was used to obtain prediction information for the IL-4gene and to analyze the survival rate of patients with gastric, lung, breast, and ovarian cancer. Graphical analysis revealed that the IL-4 deregulation had different effects for different types of cancer. In ovarian cancer, increased levels of IL-4expression predict a reduction in the risk of patients (higher survival rates). The HR ratio and the P value of ovarian cancer are (0,77 HR (0.67–0.88), P = 0.000095, 1,08–1,39) and P (Fig. 4). In addition, low levels of IL-4expressions are associated with high risk patients (low survival rate) in breast cancer (HR 0.87 (0.78–0.96, P 0.0054), lung cancer (HR 1.02 (0.09–1.14, P 0.079), and gastric cancer (HR 1.6 (0.34–1.92, P 0.000016). Controlling the expression of IL-4 genes is something that is not expected of healthy people. Errors in the transcription of IL-4gene can lead to various types of cancer. Consequently, IL-4 genes may be useful as a potentially predictive marker of some cancers. Since nsSNPs affect the structure and function of IL-4 proteins, we believe that the seven nsSNPs identified in this study are likely to have almost the same functional effect on IL-4deregulation.

Thousands of polymorphisms have been reported in the coding and noncoding regions of the IL-4gene. Molecular approach is costly and takes time to identify functionally important SNPs in pools that contain harmful and neutral SNPs. Many computational approaches play a major role in predicting and identifying important changes that have adverse effects on protein annotation. [27-29]. However, the current silico method has some shortcomings in the prediction of harmful nsSNPs, as each algorithm uses different parameters for the prediction. Thus, it is not necessary to consider single algorithms to properly predict harmful nsSNPs. To predict harmful nsSNPs, different algorithms with different parameters and aspects must be implemented. A consensus outcome obtained from most tools can provide a reliable outcome. We examined genetic variations in the IL-4locus. In addition to 152 reported SNPs, six different computational tools have identified seven high-risk SNPs. The seven filtered nsSNPs were analyzed with I-Mutant3.0 Mupro, and IPNS-MD to study the protein stability effect. It was found that nsSNP sex causes a decrease in stability, while it is expected that the rigidity of the IL-4protein will increase (Table 4). The reason of molecular alternation0that potentially affects the annotation of the IL4protein were examined using

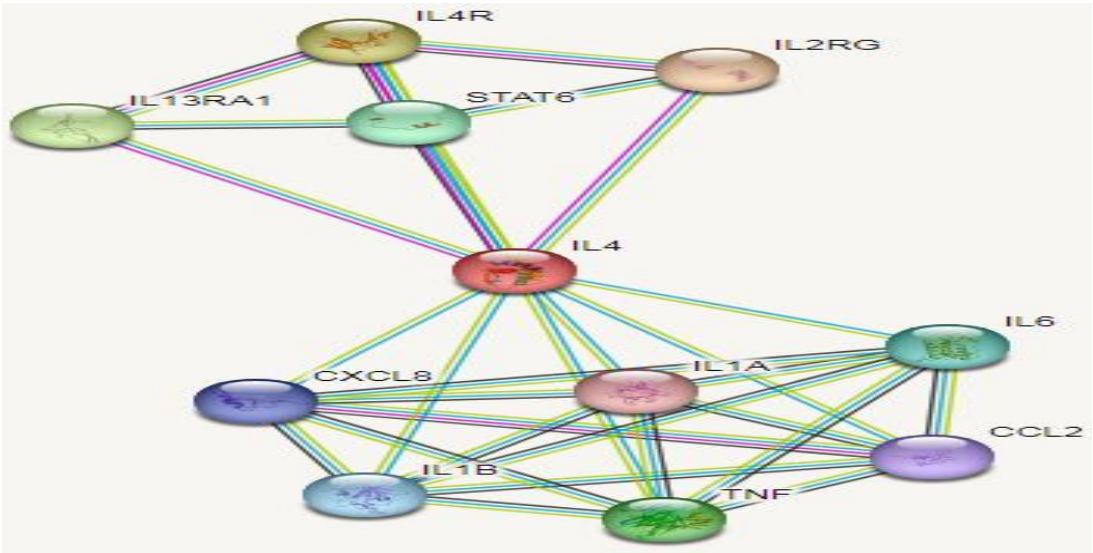

Fig.3. Protein–protein interaction network of IL-4protein using STRING.

MutPred2 web server (Table 3). The change in protein stability affects the conformation structure and thus determines the function of the protein [30]. nsSNPs mentioned can affect the stability of proteins and have the strongest harmful effects on their annotation. Reduction in protein stability can alter protein folding mechanisms, and can lead to protein degradation or abnormal aggregation. [31, 32]. Project HOPE software results have provided important information about the possible effects of missense0SNPs of *IL-4*gene. The polymorphisms (rs139863211, rs149950065, rs199929962, rs200549061, rs201594156, rs202231191, and rs376367511) result in V53A, A118G, M144T, G2D, L110R, N113Y, and C123R amino acid substitutions, respectively. These substituted amino acids have different physiological and chemical properties that can interrupt the structure of IL-4 proteins. Because of the polymorphism, the N113Y mutated residue was more water-resistant than the wild-type residue, causing the loss of hydrogen bonds with other molecules and disrupting the correct folding of the protein. On the contrary, wild amino acid residues were more hydrophobic than those of wild amino acids. (A118G, M144T, G2D, L110R, and C123R) mutation, resulting in loss of hydrophobic interactions with other molecules on the surface of the protein. From the analysis of our HOPE project, we found that most mutations cause loss of hydrophobic interactions. This finding indicates that mutations may interfere with the interconnection of two subunits, thus hindering

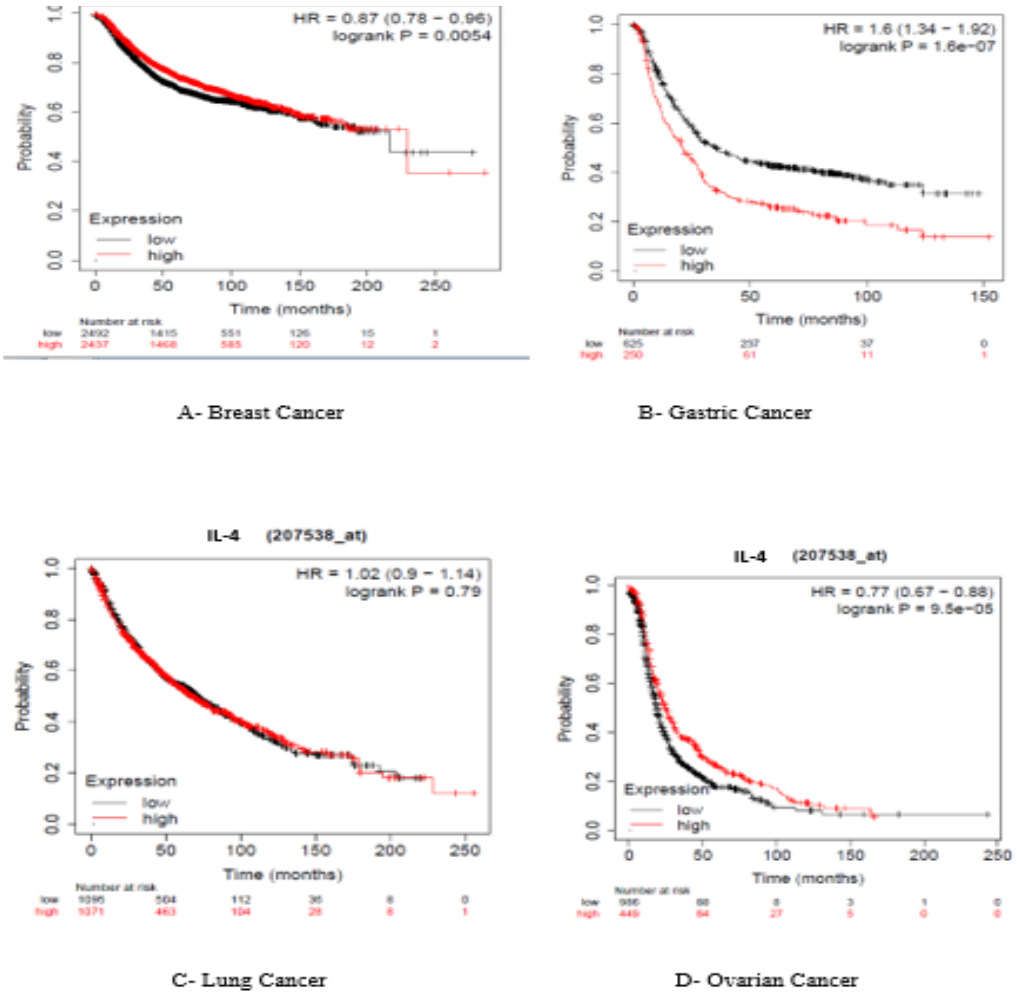

Fig. 4. IL-4 expression data-based (microarray) association study in the survival rate of patients with different types of cancers.

the dimerization process of IL-4. In addition, all mutations are located in the protein catalytic area and are crucial for the catalytic function of proteins. The mutation of these residues may disrupt the catalytic activity of IL-4. In IL-4wild proteins, residues of the M144 amino acids generate a helix structure (notated by UniProt). However, the M144T polymorphism of IL-4 does not support the alpha-helix as a secondary structure at each position. Another mutation, A118G, introduced glycine residues at this location. Glycine is very flexible and can disturb the protein rigidity required in this position. Overall, the results showed that the modeled mutated protein (Fig. 2) Different from wild IL-4 proteins, it causes instability in the protein and may cause the IL-4 to be incompatible with the receptor. Our findings show that it is located mainly in the binding site regions. In several studies, the functional effects of this SNP on the binding of the IL-4receptor complex to the IL-4receptor complex and downstream signaling were studied. [33] showed that this SNP not only extended IL-4binding to receptors, but also extended STAT6activation [34]. Thus, mutations in the binding site region of IL-4 can be speculated to interfere with their interactions with their respective receptors and ultimately prevent IL-4 from transmitting downstream signals. STRING analysis data reveal that IL-4 proteins have a number of essential functions, causing IL-2RG, IL-6, IL-1B, TNF, and CXCL8 to be synthesized by inflammatory macrophages and T cells. (Fig. 3). IL-4 also has tumor promoter and tumor inhibitor properties. The increase in IL-4 levels is associated with increased tumor growth and poor predictions and drug resistance. Again, increased expression of IL-4regulates Class I and other cytokines, thereby controlling tumor accumulation and inhibiting tumor formation. Previous studies have shown that IL-4 contributes to gastric cancer pathogenesis. Similarly, in ovarian cancer, high levels of IL-4 expression have been reported to inhibit the growth of ovarian cancer cells due to the decrease in inflammatory cell growth. The double effects of IL-4 may be the result of protein concentrations. The study showed that elevated IL-4 gene expression has a positive impact on the overall survival of patients with gastric and ovarian cancer. (Fig. 4). However, further research is needed to verify the correlation between IL-4protein defects and different types of cancer development.

CONCLUSION

We identified nine potentially harmful IL-4 nsSNPs using several insilico tools. We believe that the identification of these nsSNPs should help to develop cost-effective and rapid screening methods for the diagnosis of diseases related to the expression of IL-4. Furthermore, it will greatly facilitate the approach to experimental design for future laboratory research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests regarding the publication of this manuscript.

REFERENCES

1. Piazzon MC, Savelkoul HFJ, Pietretti D, Wiegertjes GF, Forlenza M. Carp Il10 Has Anti-Inflammatory Activities on Phagocytes, Promotes Proliferation of Memory T Cells, and Regulates B Cell Differentiation and Antibody Secretion. The Journal of Immunology. 2014;194(1):187-199.
2. Teresa Gonzalez-Garza M, Elva Cruz-Vega D, Maldonado-Bernal C. IL10 as Cancer Biomarker. Translational Research in Cancer: IntechOpen; 2021.
3. Hupe M. StatPearls. Journal of Electronic Resources in Medical Libraries. 2024;21(3):147-152.
4. Adiba M, Das T, Paul A, Das A, Chakraborty S, Hosen MI, et al. In silico characterization of coding and non-coding SNPs of the androgen receptor gene. Informatics in Medicine Unlocked. 2021;24:100556.
5. Rahim M, Gibbon A, Collins M, September AV. Genetics of musculoskeletal soft tissue injuries: Current status, challenges, and future directions. Sports, Exercise, and Nutritional Genomics: Elsevier; 2019. p. 317-339.
6. Mosmann TR, Cherwinski H, Bond MW, Giedlin MA, Coffman RL. Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. The Journal of Immunology. 1986;136(7):2348-2357.
7. Jameel MI, Al-Imam MJ, Mohammed IH. Finding the effect of missense-SNPs on protein structures and functions of HPRT1 using different tools. Applied Nanoscience. 2023;13(8):5379-5387.
8. Zhu J. T helper 2 (Th2) cell differentiation, type 2 innate lymphoid cell (ILC2) development and regulation of interleukin-4 (IL-4) and IL-13 production. Cytokine. 2015;75(1):14-24.
9. Ansel KM, Djuretic I, Tanasa B, Rao A. Regulation of TH2 Differentiation and Il4 Locus Accessibility. Annu Rev Immunol. 2006;24(1):607-656.
10. Koh BH, Hwang SS, Kim JY, Lee W, Kang M-J, Lee CG, et al. Th2 LCR is essential for regulation of Th2 cytokine genes and for pathogenesis of allergic asthma. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(23):10614-10619.
11. Schieck M, Sharma V, Michel S, Toncheva AA, Worth L, Potaczek DP, et al. A polymorphism in the TH2 locus control region is associated with changes in DNA methylation and

gene expression. Allergy. 2014;69(9):1171-1180.
12. Bernstein ZJ, Shenoy A, Chen A, Heller NM, Spangler JB. Engineering the IL-4IL-13 axis for targeted immune modulation. Immunol Rev. 2023;320(1):29-57.
13. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(7):1073-1081.
14. Capriotti E, Fariselli P. PhD-SNPg: a webserver and lightweight tool for scoring single nucleotide variants. Nucleic Acids Res. 2017;45(W1):W247-W252.
15. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Current protocols in human genetics. 2013;Chapter 7:Unit7.20-Unit27.20.
16. Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, Muruganujan A, et al. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. Nucleic Acids Res. 2006;34(Web Server issue):W645-W650.
17. Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, Casadio R. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. BMC Genomics. 2013;14 Suppl 3(Suppl 3):S6-S6.
18. López-Ferrando V, Gazzo A, de la Cruz X, Orozco M, Gelpí JL. PMut: a web-based tool for the annotation of pathological variants on proteins, 2017 update. Nucleic Acids Res. 2017;45(W1):W222-W228.
19. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics. 2006;22(22):2729-2734.
20. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res. 2005;33(Web Server issue):W306-W310.
21. Zhang W, Zhao Y, Zhang F, Wang Q, Li T, Liu Z, et al. The use of anti-inflammatory drugs in the treatment of people with severe coronavirus disease 2019 (COVID-19): The Perspectives of clinical immunologists from China. Clinical immunology (Orlando, Fla). 2020;214:108393-108393.
22. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H-J, et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nature communications. 2020;11(1):5918-5918.
23. Tanshee RR, Mahmud Z, Nabi AHMN, Sayem M. A comprehensive in silico investigation into the pathogenic SNPs in the RTEL1 gene and their biological consequences. PLoS One. 2024;19(9):e0309713-e0309713.
24. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43(Database issue):D447-D452.
25. Lacny S, Wilson T, Clement F, Roberts DJ, Faris P, Ghali WA, et al. Kaplan–Meier survival analysis overestimates cumulative incidence of health-related events in competing risk settings: a meta-analysis. J Clin Epidemiol. 2018;93:25-35.
26. Nagy Á, Lánczky A, Menyhárt O, Győrffy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. Sci Rep. 2018;8(1):9227-9227.
27. Hasnain MJU, Shoaib M, Qadri S, Afzal B, Anwar T, Abbas SH, et al. Computational analysis of functional single nucleotide polymorphisms associated with SLC26A4 gene. PLoS One. 2020;15(1):e0225368-e0225368.
28. Rangasamy N, Senthil Kumar N, K.S S. Computational analysis of missense variants in MMP2 gene linked with Winchester syndrome and Nodulosis-Arthropathy-Osteolysis reveals structural shift in protein-protein and protein-ligand complexes. Meta Gene. 2021;29:100931.
29. Yadav AK, Singh TR. Novel structural and functional impact of damaging single nucleotide polymorphisms (SNPs) on human SMYD2 protein using computational approaches. Meta Gene. 2021;28:100871.
30. Deller MC, Kong L, Rupp B. Protein stability: a crystallographer's perspective. Acta crystallographica Section F, Structural biology communications. 2016;72(Pt 2):72-95.
31. Witham S, Takano K, Schwartz C, Alexov E. A missense mutation in CLIC2 associated with intellectual disability is predicted by in silico modeling to affect protein stability and dynamics. Proteins. 2011;79(8):2444-2454.
32. Singh SM, Kongari N, Cabello-Villegas J, Mallela KMG. Missense mutations in dystrophin that trigger muscular dystrophy decrease protein stability and lead to cross-beta aggregates. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(34):15069-15074.
33. Ford AQ, Heller NM, Stephenson L, Boothby MR, Keegan AD. An atopy-associated polymorphism in the ectodomain of the IL-4R(alpha) chain (V50) regulates the persistence of STAT6 phosphorylation. Journal of immunology (Baltimore, Md : 1950). 2009;183(3):1607-1616.
34. Yabiku K, Hayashi M, Komiya I, Yamada T, Kinjo Y, Ohshiro Y, et al. Polymorphisms of interleukin (IL)-4 receptor alpha and signal transducer and activator of transcription-6 (Stat6) are associated with increased IL-4Ralpha-Stat6 signalling in lymphocytes and elevated serum IgE in patients with Graves' disease. Clinical and experimental immunology. 2007;148(3):425-431.