Original Article *Iran J Ped Hematol Oncol. 2021, Vol 11, No 2, 114-133*

In-silico study to identify the pathogenic single nucleotide polymorphisms in the coding region of *CDKN2A* gene

Farzaneh Ghasemi[1], Mehri Khatami[1,*], Mohammad Mehdi Heidari[1], Reyhane Chamani[1], Hadi Zare-Zardini[2,3]

1. Department of Biology, Faculty of Science, Yazd University, Yazd, Iran
2. Hematology and Oncology Research Center, Shahid Sadoughi University of Medical Sciences, Yazd, Iran
3. Department of Sciences, Farhangian University, Isfahan, Iran
*Corresponding author: Dr Mehri Khatami, Department of Biology, Faculty of Science, Yazd University, Pajoohesh Street, Yazd, Iran. Email: m.khatami@yazd.ac.ir. ORCID ID: 0000-0002-5840-5399

Received: 13 August 2020 **Accepted:** 20 Decemebr 2020

Abstract

Background: *CDKN2A*, encoding two important tumor suppressor proteins p16 and p14, is a tumor suppressor gene. Mutations in this gene and subsequently the defect in p16 and p14 proteins lead to the downregulation of RB1/p53 and cancer malignancy. To identify the structural and functional effects of mutations, various powerful bioinformatics tools are available. The aim of this study is the identification of high-risk non-synonymous single nucleotide variants in the *CDKN2A* gene via bioinformatics tools.

Materials and Methods: Among the identified polymorphisms in this gene, 353 missense variants are retrieved from the national center for biotechnology information/single nucleotide polymorphism database (NCBI/dbSNP). Then, the pathogenicity of missense variants are considered using different bioinformatics tools. The stability of these mutant proteins, conservation of amino acids and the secondary and tertiary structural changes are analyzed by bioinformatics tools. After the identification of high-risk mutations, the changes in the hydrophobicity of high-risk amino acid substitutions are considered.

Results: Deleterious single nucleotide polymorphisms (SNPs) were screened step by step using the bioinformatics tools. The results obtained from the set of bioinformatics tools identify high-risk mutations in *CDKN2A* gene.

Conclusion: 18 high-risk mutations including L16R/Q, G23D/R/S, L32P, N42K, G55D, G67D/R, P81R, H83R, G89D/S, A102E, G101R, G122R, and V126D were identified. According to the previous experimental studies, the association of L16R, G23D/R/S, L32P, G67R, H83R, G89D, G101R, and V126D amino acid substitutions with various cancers has been confirmed.

Keywords: Computational biology, *CDKN2A*, Gene, SNP, Tumor suppressor protein.

Introduction

CDKN2A, localized on 9p21.3 and encoding different tumor suppressor proteins by alternate reading frame mechanism, is an important cell cycle regulator (1). The tumor suppressor p16 is a low molecular weight protein that contains 156 amino acids including four ankyrin (Ank) repeats (Ank1: codons 11-40; Ank2: codons 44-72; Ank3: codons 77-106; Ank4: codons 110-139) (2). The p16 binds to a cyclin-dependent kinase (CDK) 4/6 and inhibits the CDK4/6 and cell cycle at the G1 phase (3, 4). CDK4/6-cyclin D and CDK2-cyclin E complexes inactivate RB1 by phosphorylation mechanism. So, E2F1 inhibited by RB1 is released and leads to cell growth and division. Therefore, p16 activates RB1 by inhibition of CDK4/6 (5). The tumor suppressor p14 containing 132 amino acids activate p53 through binding to MDM2. MDM2 as an antagonist of p53 protein downregulates p53 through its ubiquitination (6). Downregulation or inactivation of the mentioned tumor suppressor proteins has been considered in several cancers such as colon cancer (7), lung cancer (8), melanoma (9), pancreatic cancer (10), head and neck squamous cell carcinoma (HNSCC) (11), glioma (12), and leukemia (13-15). Different mechanisms such as promoter hypermethylation (7), sequence deletion

[DOI: 10.18502/ijpho.v11i2.5844]

(16), and point mutation (4) lead to the decreased expression and dysfunction of p16 and p14.

The single nucleotide variants (SNVs) are responsible for about 90% of human variability (17). Also, due to the existence of SNVs, there exist alterations in protein structure and subsequently its function. The size, charge, and hydrophobicity value in amino acids are unique traits. The original wild type and mutant residues often differ in these properties (18). Therefore, the amino acid replacement may impair the function, structure, and stability of protein, and ultimately protein-protein interaction. Eventually, loss of function and modified function (or deficiency) of tumor suppressor proteins lead to abnormal proliferation and overgrowth. But, some SNVs are benign polymorphism, and these variants have no effects on protein function (19). The experimental methods, as the most reliable approaches, are expensive and time-consuming processes. Also, in some cases, the methods of mutagenesis and extraction of mutant protein are impossible *in vitro* and *in vivo* (20). So, bioinformatics tools can help researchers to predict the effects of mutations. But, there is no single tool for this purpose and to get reliable results, various tools can be used. There are powerful bioinformatics tools to evaluate the alterations of protein structure and function. Using these tools, the distance of atoms (20-22) and pathogenesis of mutations (23-28) as well as changes in protein structure and polar contacts (22) can be predicted. Also, the stability of mutant proteins can be investigated by the evaluation of the total energy of proteins (27). The determination of protein structure in the presence of the mutation is the most important challenge in biology.

This study included a comprehensive investigation to identify the pathogen nsSNVs in p16 protein using bioinformatics tools. Also, the high-risk missense variants were introduced via an *in-silico* study. To confirm the obtained results, the changes in structure and hydrophobicity of mutant amino acids were compared with native residues. The conservation of the native residue was also considered. This computational approach can be used as a prelude to planning a targeted molecular method to prove the obtained results from the bioinformatics study.

This paper organizes as follows: In the next section, the used methods have been introduced to consider the deleterious mutations. The related results obtained from the bioinformatics tools were collected in section 3. Also, the correlation of the high-risk mutations with various cancers was gathered in section 4. Finally, the paper was ended with conclusions in section 5.

Materials and Methods

Materials

Data collection

The identified nsSNVs in *CDKN2A* gene were retrieved from the national center for biotechnology information/single nucleotide polymorphism database (NCBI/dbSNP) (https://www.ncbi.nlm.nih.gov/snp/). Also, the structure of p16 (PDB ID: 1DC2) was retrieved from the protein databank (PDB), (https://www.rcsb.org/).

Amino acid substitution effects

Sorting Intolerant From Tolerant (SIFT)

Using SIFT analyzer (https://sift.bii.a-star.edu.sg/), it is possible to predict the effects of amino acid substitution on protein function through considering the sequence homology and the physical properties of amino acids (29). The mutation is introduced as an affecting protein function if the score evaluated by SIFT is lower than 0.05. 'Seq Rep' is a fraction of sequences that contain one of the basic amino acids. The low fraction indicates insufficient information in this position. The presence of severely gapped or unalignable in this position can cause low confidence prediction.

[DOI: 10.18502/ijpho.v11i2.5844]

Polymorphism Phenotyping v2 (PolyPhen-2)

PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), as a web-based tool, predicts the effect of an amino acid substitution on the structure and function of the protein, based on the multiple sequence alignment of the 3D protein structure. The score of position-specific independent count (PSIC) from 0 to 1 can be calculated using this tool. PolyPhen-2 reports the result as a benign (with 0-0.15 score), possibly damaging (with 0.15-0.85 score), and probably damaging (with 0.85-1 score) (30). This tool provides two values of "sensitivity" and "specificity" for Confidence predictions.

Protein Variation Effect ANalyzer (PROVEAN)

PROVEAN (http://provean.jcvi.org/index.php) web server is used to analyze protein variants via an alignment-based score approach. This online tool predicts the effect of missense variants and indel on the protein function (31). If the calculated score of the amino acid substitution is lower than -2.5, this mutation is deleterious.

Predicting Human Deleterious SNPs in the human genome (PHD-SNP^g)

PHD-SNP^g (https://snps.biofold.org/phd-snpg/) is a machine learning method that depends on sequence-based features. This tool considers the impact of SNVs in the coding and non-coding regions. The SNV is identified as a pathogenic or benign mutation. A probabilistic score is from 0 to 1. If the score is >0.5, the variants are predicted to be pathogenic mutations (32).

SNPs&GO

SNPs&GO (http://snps.biofold.org/snps-and-go/) is a web server tool that predicts the effect of amino acid substitution as a disease-associated variation or neutral variation effects based on protein structure, and sequence (33). The evaluated score is from 0 to 1 that SNV with a score >0.5 is identified as a disease association variation.

Protein Analysis through Evolutionary Relationships (PANTHER)

The PANTHER (http://www.pantherdb.org/) classification system is a genomic analysis system based on gene function, ontology, pathways, and statistical analysis tools (34).

I-Mutant2.0

I-Mutant2.0 (http://folding.biofold.org/i-mutant/i-mutant2.0.html) is a support vector machine (SVM) that predicts protein stability changes due to the single point mutations based on protein structure or sequence. Using empirical thermodynamic data, I-Mutant2.0 calculates the free energy changes of the protein, i.e., delta delta G (DDG) (35). Accordingly, protein stability is decreased or increased if DDG is lower or upper than 0, respectively.

Structural consideration

NetSurfP-2.0

NetSurfP-2.0 (http://www.cbs.dtu.dk/services/NetSurfP/), as a sequence-based tool, predicts the secondary structure, surface accessibility, structural disorder, and backbone dihedral angles (Phi and Psi angels) for each residue of the protein sequence. This tool predicts 2-class relative solvent accessibility (RSA) for an amino acid (buried or exposed) with a threshold of 25%. Also, absolute solvent accessibility (ASA) output is calculated by multiplying RSA and ASA^{max} (36).

PyMOL software

PyMOL written in Python is a molecular visualization system for the evaluation of structural biology. This software is used to create mutations in p16 and consider the polar contacts and hydrogen bond (H-bond) length in the native and mutant proteins.

Conservation Surface mapping (ConSurf)

ConSurf web server (https://consurf.tau.ac.il/) evaluates the conservation level of the amino acids, based on the evolutionary relations between the protein and its homologs (37).

[DOI: 10.18502/ijpho.v11i2.5844]

ConSurf maps the 3D structure of the protein with a color scale extending from 1 to 9 that 1 (9) is related to a hypervariable (highly conserved) amino acid (38).

Project HOPE

HOPE (https://www3.cmbi.umcn.nl/hope/) is a web service that analyzes the structural effects of the point mutation in a protein sequence by combining the available information obtained from a series of web services and databases (39).

Analyzing hydrophobicity changes

The hydrophobic changes are analyzed using the web-based PEPTIDE 2.0 (https://www.peptide2.com/N_peptide_hydrophobicity_hydrophilicity.php) and ExPASy/ProtScale with Kyte & Doolittle amino acid scale (https://web.expasy.org/protscale/). Based on the chemical and physical properties of the amino acids, ExPASy/ProtScale predicts the hydrophobic or hydrophilic scale of the protein structure parameters (40).

Results

Data collection

The 8405 SNPs (single nucleotide polymorphisms) have been identified in the *CDKN2A* gene but, there were 465 SNPs in the coding region of this gene. From 465 missense variants, the 353 nsSNVs have been identified in p16 transcript and these 353 nsSNVs were selected for bioinformatics analysis.

Amino acid substitution effects

Using SIFT, out of 353 amino acid substitutions, the 143 mutations were predicted as an "affected protein function", and 210 amino acid substitutions were predicted as a "tolerated substitution". The SIFT 'Seq Rep' score for all nsSNPs except 2 positions was 1.00, but for 2 positions including 125, 126, this score was 0.94. These high Seq Rep scores indicated high confidence prediction. Then, all of the 143 damaging amino acid substitutions identified by SIFT were screened by PolyPhen-2. According to the results of PolyPhen-2, the 119 amino acid substitutions were predicted as the "damaging" mutations. The "sensitivity" and "specificity" obtained from this tool for all of the predictions have been gathered in Table I. One can see that the evaluated sensitivity for the majority of predictions was very low, so these results had low confidence. But, this problem was solved using the combination of various bioinformatics tools. Using the PROVEAN analyzer, it was found that 19 amino acid substitutions had neutral effects, but 100 amino acid substitutes had deleterious effects. This tool used 166 related sequences classified in 30 clusters as the supporting sequence set for these predictions. In the next step, the stability of 100 mutant proteins was studied via the I-Mutant2.0 tool. I-Mutant2.0 calculated the DDG value of 100 amino acid substitutions, also this tool predicted a decrease in the stability of 79 mutations (see Table I).

Also, the deleterious or neutral effects of the amino acid substitutions shown in Table II were considered by SNPs&GO, PANTHER, and PHD-SNP^{g}. Accordingly, the 48 amino acid substitutions produced damaging effects on protein function.

Structural consideration

The secondary structure, RSA, ASA, and class assignment of the 48 damaging amino acid replacements were studied via NetSurfP-2.0 web service. Also, the conservation of residues was considered by ConSurf. Out of the 48 amino acid substitutions, the 18 amino acid substitutions (L16R/Q, G23D/R/S, L32P, N42K, G55D, G67D/R, P81R, H83R, G89D/S, A102E, G101R, G122R, and V126D) displayed a huge increase in RSA and also these amino acid substitutions were highly conserved. So, these mutations were identified as high-risk mutations. It is necessary to mention that the mutant protein stability was decreased when the solvent accessibility was increased (41). As shown in Table III, the class assignment, RSA, ASA, and the secondary structure of the native and

[DOI: 10.18502/ijpho.v11i2.5844]

mutant amino acids were considered. According to the NetSurfP-2.0 results, Asn at position 42 and Gly at positions 23, 55, 67, 89, 101, and 122 were located on the surface of the protein. But, Leu at positions 16 and 32, Pro at position 81, His at position 83, Ala at position 102, and Val at position 126 are buried in the core of the protein.

The structural alterations of the 18 amino acid substitutions were considered using the HOPE web service. According to the HOPE results, in the L16R/Q, G23D/R/S, N42K, G55D, G67D/R, P81R, H83R, G89D/S, G101R, A102E, G122R, and V126D amino acid substitutions, the mutant residues were bulkier or larger than the wild type residues. The helix structure might be unstable in the L16R due to the placement of two large amino acids next to each other (amino acids sequence at positions 15 and 16: WL to WR). About L32P, substituted Pro was localized in the α-helix structure and disrupted α-helix due to the missing H-bond.

In L16R, G23D/R, N42K, G55D, G67D/R, P81R, H83R, G89D, G101R, A102E, G122R, and V126D, the wild type residues were neutral, but the mutant residues were charged (either positive or negative). In L16R, P81R, H83R, A102E, and V126D, the mutant residues introduced a charge in buried residues leading to probable defects in protein folding.

Gly at positions 23, 55, 67, 89, 101, and 122 was wild type residue. Among the amino acids, Gly was the simplest and more flexible one that played an important role in the secondary structures specially β-turns. So, the mutation of this residue might lead to protein dysfunction or decreased protein stability (42). According to the ConSurf server, the wild type residues at positions 16, 32, 101, and 102 were completely conserved. So, the alteration of these positions was probably damaging to the protein function (37, 43).

Evaluation of 3D structure by PyMOL

The investigated structure of p16 has been retrieved from the PDB databank with ID: 1DC2. In the study of mutations using PyMOL, no change in polar contacts was observed in L16Q, G55D, P81R, and G101R mutations. As shown in Figures 1, 2, and 3, the polar contacts of mutant residues in the mutations L16R, G23D/R/S, L32P, N42K, G67D/R, H83R, G89D/S, A102E, G122R, and V126D in comparison with wild type residues were changed. According to Figure 1, new H-bonds have been created in the Arg substituted Leu at position 16 (between R16 and L63, length=1.9 Å), Arg substituted His at position 83 (between R83 and T77 with length=2.5 Å, between R83 and D108 with length=0.9 Å), Ser substituted Gly at position 89 (between S89 and G122, length=2.9 Å), and Glu substituted Ala at position 102 (between E102 and L97, length=2.1 Å). H-bond lengths between G89 with A85 (H-bond length=2.2 Å) and G89 with A86 (H-bond length=1.8 Å) were changed with replacing Ser at position 89. From Figure 1c one can see that H-bond length between S89 and A85 (S89 and A86) was equal to 2.3 (1.7) Å. According to Figure 1d, there was a polar contact between A102 and L104 (length=2.1 Å), this bond has been destroyed by replacing Glu at position 102. Figure 2 shows that H-bonds length was changed in G23D/R/S, G67R, and G89D amino acid substitutions. There was a polar contact between G23 and A20 with a length of 2.0 Å that by substitution Gly to Asp/Arg/Ser this distance was reduced to 1.9 Å (Figure 2a). As shown in Figure 2b, in G67D, the H-bond length between G67 and L63 was 1.7 Å and between G67 and L64 was 2.0 Å. But, in the mutant forms (G67R), the H-bond length between R67 and L63 (between R67 and L64) was equal to 1.4 (2.6) Å. Gly as a wild type residue at position 89, created polar contacts between G89 and A85 with a length of 2.2 Å and between G89 and A86 with a length of 1.8 Å. Placement of Asp at position 89 leds to an increase in distance between D89 and A85 to 2.3 Å. Also, as shown in Figure 2c,

[DOI: 10.18502/ijpho.v11i2.5844]

the distance between D89 and A86 was reduced to 1.7 Å.

Figure 3 shows that L32P, N42K, G67D, G122R, and V126D amino acid substitutions led to the loss of H-bonds. Also, in mutations L32P and G67D, H-bonds length was changed. Leu at position 32 as a wild type residue had polar contacts with G35 (length=2.6 Å) and V28 (length=1.5 Å). From Figure 3a one can see that by changing the Leu to Pro at position 32 the polar contact with G35 would be destroyed and the length of H-bond with V28 was noticeably changed (length=2.5 Å). According to Figure 3b, there was an H-bond between N42 and S43 with a length of 1.9 Å. While by the substitution of Asn to Lys at position 42, the polar contact with S43 has been destroyed. As shown in Figure 3c, Gly 67 formed an H-bond with Asn 39 (H-bond length=2.4 Å), an H-bond with L63 (H-bonds length=1.7 Å), and an H-bond with L64 (H-bond length=2.0 Å). The H-bond with L64 has been destroyed by changing the Gly to Asp at position 67. Also, the H-bond length between D67 and L63 was changed and was equal to 1.4 Å and a new H-bond has been created between D67 and N39 with a length of 2.5 Å. There was a polar contact between G122 and A118 (H-bond length=2.6 Å) that this polar contact has been missed by replacing Arg at position 122 (see Figure 3d). Figure 1e shows that at position 126, Val had H-bonding with H123 (H-bond length=1.9 Å) in the wild type form, but this polar contact was destroyed by replacing ASP at position 126.

Analyzing hydrophobicity changes

As shown in Table IV, the wild type residues at L16R/Q, G23D/R, L32P, G55D, G67D/R, G89D, G101R, A102E, G122R, and V126D amino acid substitutions were more hydrophobic than the mutant residues. So, because of the reduction in hydrophobicity values, the probability of hydrophobic interactions was reduced. So, these mutations can disrupt the structure of the protein. Changes in hydrophobicity calculated by ExPASy were noticeable for L16R/Q, G23R, L32P, G67R, G101R, A102E, G122R, and V126D, but can be ignored in G23S, N42K, and G89S mutations.

Table I. The summary of pathogenicity predictions of nsSNVs obtained via SIFT, PolyPhen-2, I-Mutant 2.0, and PROVEAN.

SNP ID	Amino acid change in p16	SIFT SCORE[1]	POLYPHEN-2 SCORE[2]	POLYPHEN-2 Sensitivity/Specificity	I-MUTANT2.0 DDG[3]	PROVEAN SCORE[4]
rs864622263	p.L16R	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.67	Deleterious -4.818
	p.L16P	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.04	Deleterious -5.55
	p.L16Q	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -2.07	Deleterious -4.842
rs760065045	p.A20P	APF 0.02	Probably Damaging 1.000	0.00/1.00	Decrease -0.16	Deleterious -4.06
rs864622484	p.A20G	APF 0.03	Probably Damaging 0.966	0.78/0.95	Decrease -1.20	Deleterious -3.21
rs1329324238	p.A21D	APF 0.02	Probably Damaging 1.000	0.00/1.00	Decrease -1.40	Deleterious -4.90
rs1064794292	p.G23D	APF 0.04	Probably Damaging 1.000	0.00/1.00	Decrease -2.41	Deleterious -5.899
rs1131691186	p.G23R	APF 0.01	Probably Damaging 1.000	0.00/1.00	Decrease -1.22	Deleterious -6.72
	p.G23S	APF 0.03	Probably Damaging 0.998	0.27/0.99	Decrease -1.40	Deleterious -5.029
	p.G23C	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.70	Deleterious -7.535
rs748780473	p.V25G	APF	Possibly Damaging	0.87/0.91	Decrease	Deleterious

[DOI: 10.18502/ijpho.v11i2.5844]

		0.00	0616		-3.29	-3.84
rs775176191	p.V28G	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -3.46	Deleterious -5.482
rs1554656382	p.R29W	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.68	Deleterious -5.26
rs878853650	p.L32P	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -2.03	Deleterious -5.743
rs745827714	p.L32V	APF 0.00	Probably Damaging 0.997	0.41/0.98	Decrease -2.18	Deleterious -2.50
rs746834149	p.G35V	APF 0.02	Probably Damaging 1.000	0.00/1.00	Decrease -0.17	Deleterious -5.95
rs200382984	p.A36G	APF 0.01	Probably Damaging 0.937	0.80/0.94	Decrease -1.30	Deleterious -2.76
rs752731682	p.N39H	APF 0.01	Probably Damaging 0.998	0.27/0.99	Decrease -1.45	Deleterious -3.93
rs1554656306	p.N39S	APF 0.03	Possibly Damaging 0.864	0.83/0.93	Decrease -0.69	Deleterious -3.80
rs864622439	p.N39K	APF 0.02	Probably Damaging 0.995	0.68/0.97	Decrease -1.04	Deleterious -4.66
rs1060501264	p.N42S	APF 0.00	Probably Damaging 0.999	0.14/0.99	Decrease -0.92	Deleterious -4.31
rs1587339638	p.N42K	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.03	Deleterious -5.17
rs1587339662	p.N42D	APF 0.00	Probably Damaging 0.999	0.14/0.99	Decrease -1.00	Deleterious -4.31
rs1328708469	p.G45S	APF 0.03	Probably Damaging 1.000	0.00/1.00	Decrease -0.16	Deleterious -4.66
rs1563892916	p.R46W	APF 0.01	Probably Damaging 1.000	0.00/1.00	Decrease -0.66	Deleterious -6.69
rs763804037	p.P48R	APF 0.04	Probably Damaging 1.000	0.00/1.00	Decrease -0.10	Deleterious -4.84
rs199907548	p.I49T	APF 0.01	Probably Damaging 1.000	0.00/1.00	Decrease -3.20	Deleterious -3.773
	p.I49S	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -2.94	Deleterious -4.666
rs587778189	p.Q50P	APF 0.02	Probably Damaging 1.000	0.00/1.00	Decrease -0.67	Deleterious -5.185
rs561034503	p.G55D	APF 0.04	Probably Damaging 1.000	0.00/1.00	Decrease -0.09	Deleterious -6.49
rs104894099	p.V59G	APF 0.00	Probably Damaging 0.998	0.27/0.99	Decrease -4.72	Deleterious -5.983
	p.V59E	APF 0.00	Probably Damaging 0.999	0.14/0.99	Decrease -2.28	Deleterious -5.242
rs36204594	p.A60E	APF 0.01	Probably Damaging 1.000	0.00/1.00	Decrease -1.24	Deleterious -4.617
rs769382085	p.A60P	APF 0.02	Probably Damaging 1.000	0.00/1.00	Decrease -0.39	Deleterious -4.68
	p.A60T	APF 0.04	Probably Damaging 1.000	0.00/1.00	Decrease -1.38	Deleterious -3.73
rs758389471	p.G67R	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.96	Deleterious -7.56
	p.G67S	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.35	Deleterious -5.70
rs863224605	p.G67D	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -2.26	Deleterious -6.70
rs1060501260	p.A68G	APF 0.00	Probably Damaging 0.998	0.27/0.99	Decrease -0.82	Deleterious -3.842
rs559848002	p.N71T	APF 0.02	Probably Damaging/ 0.993	0.70/0.97	Decrease -0.54	Deleterious -5.500
rs1554654113	p.T79P	APF 0.03	Probably Damaging 0.983	0.74/0.96	Decrease -0.58	Deleterious -3.81
rs11552823	p.P81R	APF 0.02	Probably Damaging 1.000	0.00/1.00	Decrease -0.63	Deleterious -8.15
rs34968276	p.H83Q	APF	Probably Damaging	0.00/1.00	Decrease	Deleterious

[DOI: 10.18502/ijpho.v11i2.5844]

		0.00	1.000		-0.97	-7.51
rs1057519881	p.H83R	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.47	Deleterious -7.494
rs121913385	p.H83D	APF 0.00	Probably Damaging 0.999	0.14/0.99	Decrease -1.62	Deleterious -8.475
rs1064796336	p.A85F	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.14	Deleterious -5.350
rs878853646	p.A85S	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.28	Deleterious -2.74
	p.A85T	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.87	Deleterious -3.53
rs1190283873	p.A86T	APF 0.01	Probably Damaging 1.000	0.00/1.00	Decrease -0.67	Deleterious -3.71
rs749714198	p.R87W	APF 0.01	Possibly Damaging 0.675	0.86/0.92	Decrease -0.48	Deleterious -7.439
rs137854597	p.G89S	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.04	Deleterious -5.600
rs137854599	p.G89A	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.84	Deleterious -5.650
	p.G89D	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.34	Deleterious -6.592
rs1563889362	p.L91Q	APF 0.00	Probably Damaging 0.999	0.14/0.99	Decrease -2.65	Deleterious -5.39
rs34886500	p.R99W	APF 0.01	Probably Damaging 1.000	0.00/1.00	Decrease -0.36	Deleterious -4.76
rs104894094	p.G101W	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.78	Deleterious -6.134
	p.G101R	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.64	Deleterious -5.890
rs35741010	p.A102T	APF 0.01	Probably Damaging 0.995	0.68/0.97	Decrease -0.53	Deleterious -3.59
rs137854598	p.A102E	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.38	Deleterious -4.47
rs767642535	p.R103W	APF 0.02	Probably Damaging 1.000	0.00/1.00	Decrease -0.21	Deleterious -5.10
rs1554654028	p.V106E	APF 0.04	Probably Damaging 1.000	0.00/1.00	Decrease -1.93	Deleterious -3.55
rs1339792331	p.D108V	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.39	Deleterious -8.24
rs121913381	p.D108Y	APF 0.01	Probably Damaging 1.000	0.00/1.00	Decrease -1.46	Deleterious -8.201
rs778971134	p.G111R	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.05	Deleterious -5.43
	p.G111S	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.20	Deleterious -3.67
rs876660436	p.R112C	APF 0.05	Probably Damaging 1.000	0.00/1.00	Decrease -0.59	Deleterious -4.05
rs104894104	p.P114S	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.39	Deleterious -7.479
	p.P114T	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.86	Deleterious -7.495
rs121913386	p.P114L	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -0.66	Deleterious -9.361
	p.P114H	APF 0.00	Probably Damaging 1.000	0.00/1.00	Decrease -1.83	Deleterious -8.407
rs750655995	p.V115G	APF 0.01	Probably Damaging 1.000	0.00/1.00	Decrease -3.55	Deleterious -6.16
	p.V115E	APF 0.01	Probably Damaging 0.999	0.14/0.99	Decrease -2.20	Deleterious -5.32
rs1060501270	p.A118G	APF 0.05	Probably Damaging 0.996	0.55/0.98	Decrease -1.03	Deleterious -3.73
rs1554653960	p.A118T	APF 0.02	Probably Damaging 1.000	0.00/1.00	Decrease -1.07	Deleterious -3.71
rs113798404	p.G122R	APF	Probably Damaging	0.00/1.00	Decrease	Deleterious

[DOI: 10.18502/ijpho.v11i2.5844]

		0.03	1.000		-0.61	-6.347
rs146179135	p.D125Y	APF 0.04*	Possibly Damaging 0.743	0.85/0.92	Decrease -0.31	Deleterious -4.222
rs104894098	p.V126D	APF 0.00*	Probably Damaging 1.000	0.00/1.00	Decrease -2.62	Deleterious -6.223
rs1350305259	p.V126L	APF 0.03*	Probably Damaging 0.898	0.82/0.94	Decrease -1.10	Deleterious -2.57
rs1563888826	p.R128W	APF 0.01	Probably Damaging 0.992	0.70/0.97	Decrease -0.42	Deleterious -3.24

*There is low confidence in this prediction. [1]Score<0.05= Affected Protein Function (APF), [2]Score 0.15-0.85= possibly damaging and score: 0.85-1= probably damaging, [3]DDG<0= Decrease stability, [4]Score<-2.5= deleterious.

Table II. Disease probability by PHD-SNP[g], PANTHER, and SNPs&GO.

SNP ID	Amino acid change in p16	PHD-SNP[g]/SCORE*	PANTHER	SNPs&GO/SCORE**
rs864622263	p.L16R	Pathogenic/0.979	Probably Damaging	Disease/0.893
	p.L16Q	Pathogenic/0.969	Probably Damaging	Disease/0.858
	p.L16P	Pathogenic/0.971	Probably Damaging	Disease/0.679
rs760065045	p.A20P	Pathogenic/0.972	Possibly Damaging	Disease/0.531
rs1329324238	p.A21D	Pathogenic/0.955	Possibly Damaging	Disease/0.645
rs1064794292	p.G23D	Pathogenic/0.946	Probably Damaging	Disease/0.903
rs1131691186	p.G23S	Pathogenic/0.932	Probably Damaging	Disease/0.813
	p.G23C	Pathogenic/0.973	Probably Damaging	Disease/0.864
	p.G23R	Pathogenic/0.963	Probably Damaging	Disease/0.662
rs775176191	p.V28G	Pathogenic/0.940	Possibly Damaging	Disease/0.826
rs878853650	p.L32P	Pathogenic/0.973	Probably Damaging	Disease/0.861
rs1587339638	p.N42K	Pathogenic/0.767	Possibly Damaging	Disease/0.571
rs1328708469	p.G45S	Pathogenic/0.916	Probably Damaging	Disease/0.544
rs199907548	p.I49S	Pathogenic/0.796	Possibly Damaging	Disease/0.801
rs587778189	p.Q50P	Pathogenic/0.974	Possibly Damaging	Disease/0.903
rs561034503	p.G55D	Pathogenic/0.963	Probably Damaging	Disease/0.657
rs36204594	p.A60E	Pathogenic/0.967	Possibly Damaging	Disease/0.747
rs758389471	p.G67R	Pathogenic/0.971	Probably Damaging	Disease/0.643
	p.G67S	Pathogenic/0.973	Probably Damaging	Disease/0.594
rs863224605	p.G67D	Pathogenic/0.963	Probably Damaging	Disease/0.696
rs559848002	p.N71T	Pathogenic/0.994	Possibly Damaging	Disease/0.688
rs11552823	p.P81R	Pathogenic/0.984	Probably Damaging	Disease/0.521
rs34968276	p.H83Q	Pathogenic/0.980	Probably Damaging	Disease/0.753
rs1057519881	p.H83R	Pathogenic/0.947	Probably Damaging	Disease/0.905
rs121913385	p.H83D	Pathogenic/0.974	Probably Damaging	Disease/0.895
rs1064796336	p.A85F	Pathogenic/1.000	Probably Damaging	Disease/0.832
rs878853646	p.A85S	Pathogenic/0.986	Probably Damaging	Disease/0.649
	p.A85T	Pathogenic/0.982	Probably Damaging	Disease/0.617
rs1190283873	p.A86T	Pathogenic/0.969	Probably Damaging	Disease/0.597
rs749714198	p.R87W	Pathogenic/0.905	Probably Damaging	Disease/0.893
rs137854597	p.G89S	Pathogenic/0.973	Probably Damaging	Disease/0.886
rs137854599	p.G89A	Pathogenic/0.979	Probably Damaging	Disease/0.872
	p.G89D	Pathogenic/0.981	Probably Damaging	Disease/0.933
rs1563889362	p.L91Q	Pathogenic/0.832	Possibly Damaging	Disease/0.774
rs104894094	p.G101W	Pathogenic/0.937	Probably Damaging	Disease/0.873
	p.G101R	Pathogenic/0.806	Probably Damaging	Disease/0.810
rs35741010	p.A102T	Pathogenic/0.598	Probably Damaging	Disease/0.553

[DOI: 10.18502/ijpho.v11i2.5844]

rs137854598	p.A102E	Pathogenic/0.705	Probably Damaging	Disease/0.762
rs767642535	p.R103W	Pathogenic/0.612	Possibly Damaging	Disease/0.567
rs1339792331	p.D108V	Pathogenic/0.981	Possibly Damaging	Disease/0.806
rs121913381	p.D108Y	Pathogenic/0.986	Possibly Damaging	Disease/0.902
rs104894104	p.P114S	Pathogenic/0.987	Probably Damaging	Disease/0.819
	p.P114T	Pathogenic/0.990	Probably Damaging	Disease/0.788
rs121913386	p.P114L	Pathogenic/0.993	Probably Damaging	Disease/0.751
	p.P114H	Pathogenic/0.995	Probably Damaging	Disease/0.840
rs1554653960	p.A118T	Pathogenic/0.972	Probably Damaging	Disease/0.526
rs113798404	p.G122R	Pathogenic/0.941	Possibly Damaging	Disease/0.822
rs104894098	p.V126D	Pathogenic/0.970	Possibly Damaging	Disease/0.932

*, **Amino acid substitution with the score>0.5 is identified as a pathogenic or disease association variation.

Table III. The secondary structure analysis of high-risk mutations using NetsurfP-2.0 and the conservation study by ConSurf.

SNP ID	Amino acid position and mutation	Conservation score*	Amino acid change	RSA[1] (%)	ASA[2] (Å)	Class assignment	Secondary structure
rs864622263	L16R/Q	9	L	4	8	Buried	α Helix
			R	22	51	Buried	α Helix
			Q	18	33	Buried	α Helix
rs1131691186	G23R/S	8	G	30	24	Exposed	Turn
			R	48	110	Exposed	Turn
			S	37	43	Exposed	Turn
rs1064794292	G23D	8	G	30	24	Exposed	Turn
			D	51	74	Exposed	Turn
rs878853650	L32P	9	L	12	21	Buried	α Helix
			P	25	36	Buried	α Helix
rs1587339638	N42K	7	N	15	22	Exposed	Coil
			K	24	50	Exposed	Coil
rs561034503	G55D	8	G	28	22	Exposed	Turn
			D	48	69	Exposed	Turn
rs863224605	G67D	8	G	69	54	Exposed	Turn
			D	85	122	Exposed	Turn
rs758389471	G67R	8	R	77	177	Exposed	Turn
rs11552823	P81R	7	P	2	3	Buried	α Helix
			R	12	27	Buried	α Helix
rs1057519881	H83R	8	H	4	8	Buried	α Helix
			R	11	25	Buried	α Helix
rs137854599	G89D/S	8	G	29	23	Exposed	Turn
			D	51	74	Exposed	Turn
			S	38	45	Exposed	Turn
rs104894094	G101R	9	G	65	51	Exposed	Turn
			R	76	175	Exposed	Turn
rs137854598	A102E	9	A	6	6	Buried	Coil
			E	25	44	Buried	Coil
rs113798404	G122R	7	G	56	44	Exposed	Turn
			R	75	172	Exposed	Turn
rs104894098	V126D	7	V	7	11	Buried	α Helix
			D	22	32	Buried	α Helix

[DOI: 10.18502/ijpho.v11i2.5844]

*The score is from 1 to 9. Amino acid substitution with the score=1 is identified as a hypervariable residue and with the score=9 is identified as a highly conserved residue. [1]RSA: Related solvent accessibility (the threshold for exposed or buried residue is 25%). [2]ASA: Absolute solvent accessibility.

Table IV. Evaluation of the hydrophobicity changes via the PEPTIDE 2.0 tool and ExPASy resource portal.

SNP ID	Amino acid change	PEPTIDE-2 prediction/Hydrophobicity index*(%)	Change of the nature of the amino acid	Hydrophobicity change in substituted position by ExPASy resource portal
rs864622263	p.L16R	Hydrophobic: 49.36/Basic:14.74	Hydrophobic to Hydrophilic	-0.923
	p.L16Q	Hydrophobic: 49.36/Neutral: 22.44	Hydrophobic to Neutral	-0.811
rs1064794292	p.G23D	Acidic: 14.74/Neutral: 21.15	Neutral to Hydrophilic	-0.344
rs1131691186	p.G23S	Not change	Neutral to Neutral	-0.044
	p.G23R	Basic: 14.74/Neutral: 21.15	Neutral to Hydrophilic	-0.455
rs878853650	p.L32P	Not change	Hydrophobic to Hydrophobic	-0.600
rs1587339638	p.N42K	Basic: 14.74/Neutral: 21.15	Hydrophilic to Hydrophilic	-0.045
rs561034503	p.G55D	Acidic: 14.74/Neutral: 21.15	Neutral to Hydrophilic	-0.344
rs863224605	p.G67D	Acidic: 14.74/Neutral: 21.15	Neutral to Hydrophilic	-0.344
rs758389471	p.G67R	Basic: 14.74/Neutral: 21.15	Neutral to Hydrophilic	-0.455
rs11552823	p.P81R	Hydrophobic: 49.36/Basic:14.74	Hydrophobic to Hydrophilic	-0.322
rs1057519881	p.H83R	Not change	Hydrophilic to Hydrophilic	-0.145
rs137854597	p.G89S	Not change	Neutral to Neutral	-0.044
rs137854599	p.G89D	Acidic: 14.74/Neutral: 21.15	Neutral to Hydrophilic	-0.344
rs104894094	p.G101R	Acidic: 14.74/Neutral: 21.15	Neutral to Hydrophilic	-0.456
rs137854598	p.A102E	Hydrophobic: 49.36/Acidic: 14.74	Hydrophobic to Hydrophilic	-0.589
rs113798404	p.G122R	Basic: 14.74/Neutral: 21.15	Neutral to Hydrophilic	-0.455
rs104894098	p.V126D	Hydrophobic: 49.36/Acidic: 14.74	Hydrophobic to Hydrophilic	-0.856

*Hydrophobicity index for wild type: Hydrophobicity: 50%, Acidic: 14.1%, Basic: 14.1%, Neutral: 21.79%.

Table V. Experimental studies on the correlation of high-risk mutations with various cancers.

Amino acid change in p16	Type of cancer
p.L16R	Hereditary cutaneous melanoma (46)
p.G23D	Multiple primary melanoma (47-49), Familial pancreatic cancer (50), Melanoma (51)
p.G23R	Melanoma prone family (47), Multiple primary melanoma (48)
p.G23S	Familial melanoma (52)
p.L32P	Primary familial melanoma (53), Pancreatic cancer (47), Melanoma (51), Familial melanoma (54)
p.G67R	Familial melanoma (55)
p.H83R	Pancreatic cancer (56)
p.G89D	Melanoma (57), Melanoma, HNSCC, and pancreatic cancer (58), Familial melanoma (54)
p.G101R	Cutaneous melanoma (59), Melanoma prone family (59), Pancreatic cancer (60)
p.V126D	Melanoma (61), Pancreatic cancer (60)
L16Q, N42K, G55D, G67D, P81R, G89S, A102E, G122R	There is no study about the association of these mutations with diseases.

[DOI: 10.18502/ijpho.v11i2.5844]

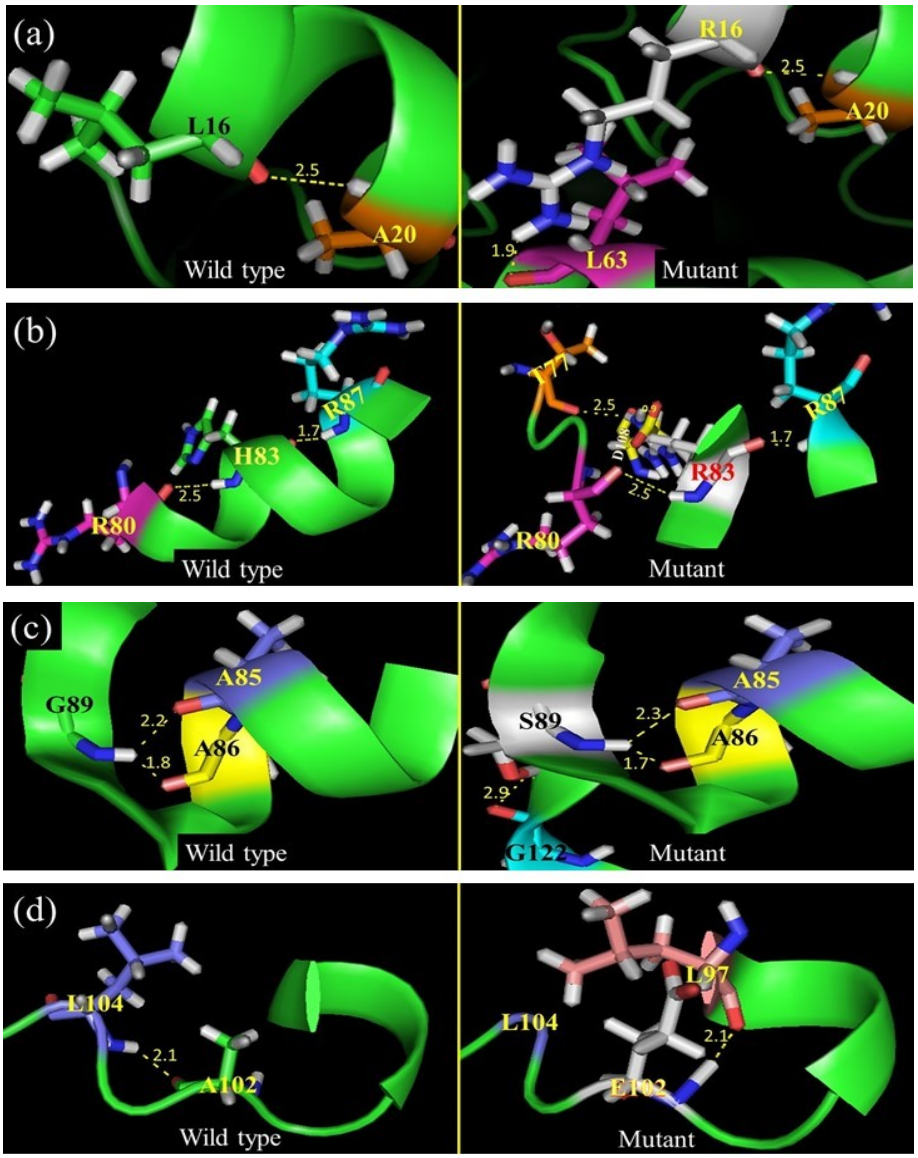

Figure 1. The created H-bonds in mutant variants (right column) while these bonds do not exist in the wild types (left column) (PDB ID=1DC2). (a): The created H-bond between R16 (mutant) and L63 in the L16R (H-bond length=1.9 Å). (b): The created H-bonds between R83 (mutant) and T77 (H-bond length=2.5 Å), between R83 and D108 (H-bond length=0.9 Å) in the H83R. (c): The created H-bond between S89 (mutant) and G122 in the G89S (H-bond length=2.9 Å). In the wild type, the H-bond length between G89 and A85 is 2.2 Å. By substituting Ser in the mutant form, the length of polar contact between S89 and A85 is equal to 2.3 Å. Also, in the wild type, the H-bond length between G89 and A86 is 1.8 Å, but between S89 and A86 is equal to 1.7 Å. (d): There was a polar contact between A102 and L104 with a length of 2.1 Å, but this bond has been destroyed by substituting of Glu and a new H-bond has been created between E102 and L97 with a length of 2.1 Å.

[DOI: 10.18502/ijpho.v11i2.5844]

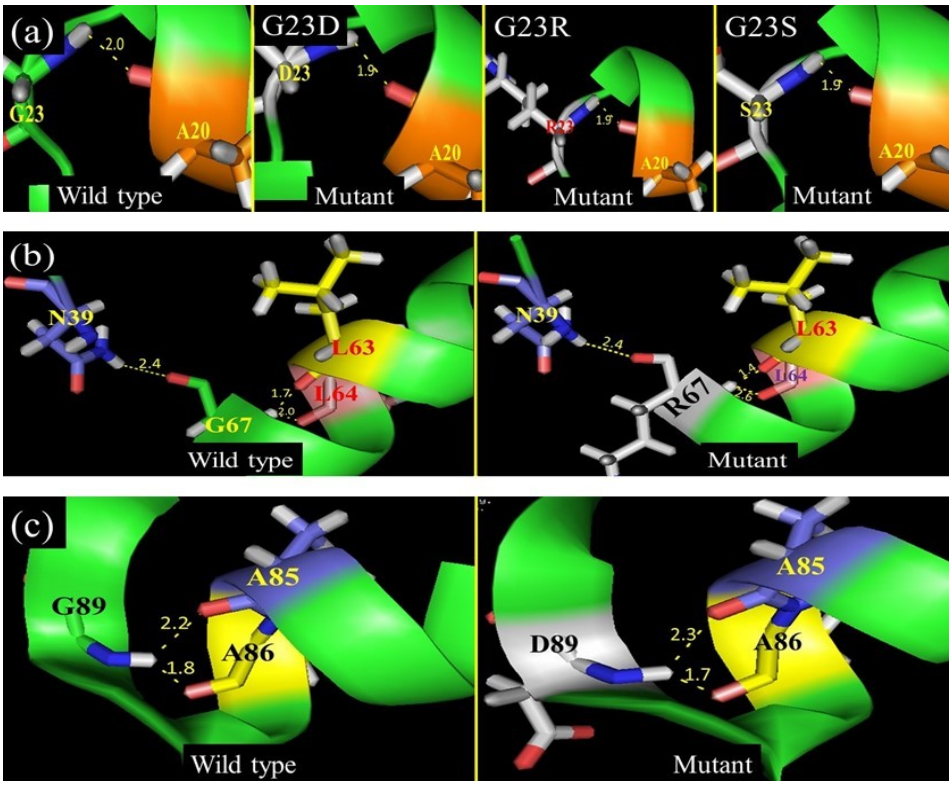

Figure 2. The change in the length of polar contact in G23D/R/S and G89D (PDB ID=1DC2) in the mutant form in comparison with the wild type protein. (a): In G23 (wild type), the H-bond length between G23 and A20 is 2.0 Å, but in the mutant forms, the H-bond length between D/R/S23 and A20 is 1.9 Å. (b): In G67 (wild type), the H-bond length between G67 and L63 is 1.7 Å and the H-bond length between G67 and L64 is 2.0 Å. But, in the mutant forms, the H-bond length between R67 and L63 is 1.4 Å and the H-bond length between R67 and L64 is 2.6 Å. (c): In G89 (wild type), the H-bond length between G89 and A85 is 2.2 Å, but the H-bond length between D89 and A85 is 2.3 Å. Also, in the wild type, the H-bond length between G89 and A86 is 1.8 Å, but the H-bond length between D89 and A86 in the mutant form is 1.7 Å.

[DOI: 10.18502/ijpho.v11i2.5844]

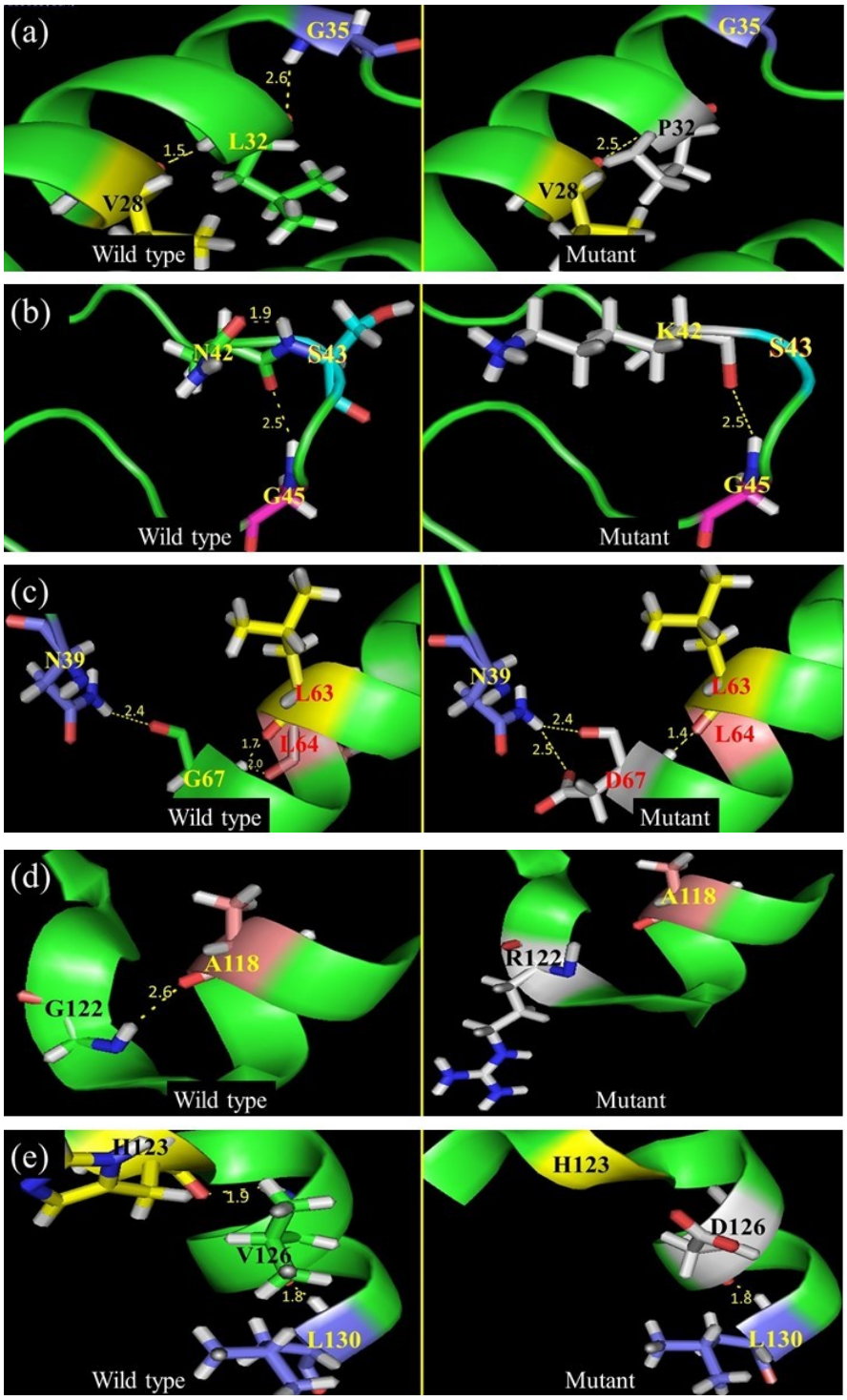

Figure 3. The loss of polar contact in the mutant form (right column) in comparison with the wild type protein (left column) (PDB ID=1DC2). (a): The H-bond between P32 and G35 has been missed while this bond exists between L32 and G35 in the wild type form with a length of 2.6 Å. Also, the H-bond length between L32 and V28 is 1.5 Å, but the H-bond length between P32 and V28 in the mutant form is 2.5 Å. (b): At position 42, Asn has H-bonding with S43 (H-bond length=1.9 Å) in the wild type form, but this polar contact is destroyed by replacing Lys at position 42. (c): Gly at position 67 has been H-bonding with N39 (H-bond length=2.4 Å), L63 (H-bond length=1.7 Å) and L64 (H-bond length=2.0 Å). The H-bond with L64 has been missed by replacing Asp. Also, the new H-bond has

[DOI: 10.18502/ijpho.v11i2.5844]

been created between D67 and N39 with a length of 2.5 Å, while there is not this bond in wild type. Also, the distance of H-bond with L63 is changed to 1.4 Å by replacing Asp at position 67. (d): There is a polar contact between G122 and A118 (H-bond length=2.6 Å) that this polar contact has been missed by replacing Arg at position 122. (e): At position 126, Val has H-bonding with H123 (H-bond length=1.9 Å) in the wild type form, but this polar contact is destroyed by replacing ASP at position 126.

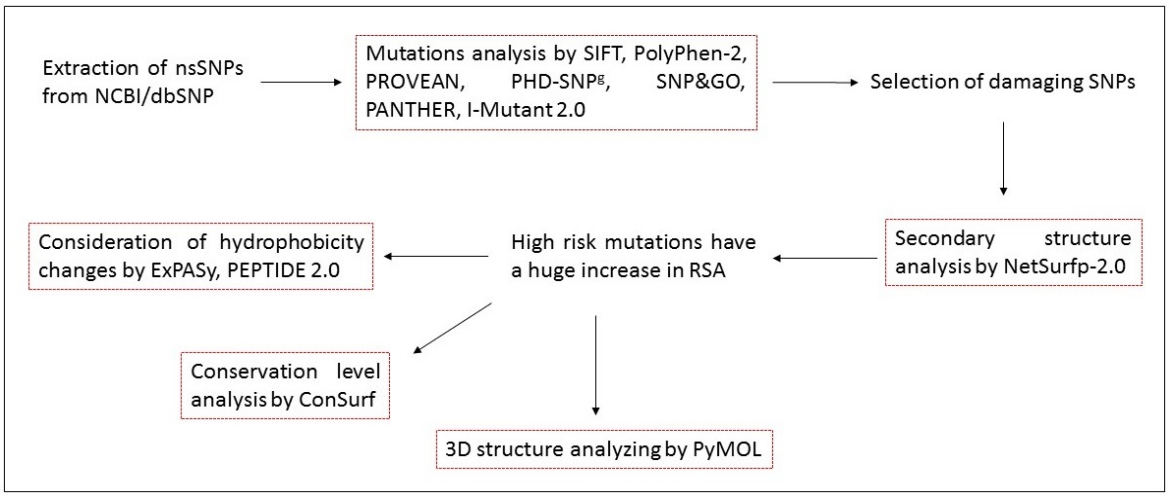

Figure 4. The flowchart of used methods for high-risk mutations identification.

Discussion

Computational approaches can be extremely useful to plan the targeted molecular methods because identifying and studying the SNPs are quite expensive and time-consuming. Also, molecular procedures such as mutagenesis or protein extraction are occasionally impossible. Ou et al. showed that the high-risk SNPs can be identified using the combined bioinformatics tools (with 94% sensitivity and 80% specificity) (27). Rajasekaran et al. considered 118 SNPs (in coding and untranslated region) in the *CDKN2A* gene in malignant melanoma via bioinformatics tools (44). In their study, the pathogenicity of nsSNVs was considered using only PolyPhen. They concluded that the missense variant with dbSNP ID rs11552822 (D84Y) could be the most deleterious SNP that leads to malignant melanoma (44). This study included a comprehensive investigation of the identified pathogen nsSNVs in p16 protein using multiple bioinformatics tools with different approaches regardless of the specific disease. Also, conservation, hydrophobicity changes, and structural alterations of these high-risk mutations were considered. The used tools in this study were SIFT, PolyPhen-2, PROVEAN, I-Mutant2.0, SNPs&GO, PANTHER, PHD-SNPg, ConSurf, NetSurfP-2.0, PyMOL, PEPTIDE 2.0, ExPASy, and HOPE. The flowchart of used bioinformatics tools in this study has been shown in Figure 4. In the present consideration, according to the results of the bioinformatics tools, out of 353 amino acid substitutions, the 18 amino acid substitutions (L16R/Q, G23D/R/S, L32P, N42K, G55D, G67D/R, P81R, H83R, G89D/S, A102E, G101R, G122R, and V126D) have been identified as the high-risk mutations. It is worth mentioning that a constraint on the formation of the α-helix is the presence of Pro residue, which has the least proclivity to form α-helixes. In

[DOI: 10.18502/ijpho.v11i2.5844]

Pro, the nitrogen atom is a part of a rigid ring, and rotation about the N—C α bond is not possible. Thus, a Pro residue introduces a destabilizing kink in the α-helix. In addition, the nitrogen atom of a Pro residue in a peptide linkage has no substituent hydrogen to participate in H-bonds with other residues. For these reasons, Pro is only rarely found in the α-helix (45). So, L32P amino acid substitution can have severe effects on the disruption of the protein structure. About V126D amino acid substitution, because of placement of two negatively amino acids next to each other in the V126D (amino acid sequence at positions 125 and 126: DV to DD) and placement of two positively amino acids next to each other in the P81R (amino acid sequence at positions 80 and 81: RP to RR), the helix structures might be unstable.

Correlation of the mutations of the *CDKN2A* with various cancers such as colon cancer (7), lung cancer (8), melanoma (9), pancreatic cancer (10), leukemia (13-15), glioma (12), and HNSCC (11) have been determined. As shown in Table V, the association of L16R, G23D/R/S, L32P, G67R, H83R, G89D, G101R, and V126D amino acid substitutions with melanoma, pancreatic cancer, and HNSCC has been confirmed in previous studies (the related references have been gathered in Table V). It should be mentioned that the UTR regions in regulating protein expression are very important. So, it is suggested that mutations in these regions should be considered with appropriate tools.

Conclusions

In this study, the nsSNVs of the *CDKN2A* gene have been identified from NCBI/dbSNP databank. Also, the pathogenicity of these nsSNVs was considered using powerful bioinformatics tools. The high-risk mutations were screened step by step via SIFT, PolyPhen-2, PROVEAN, I-Mutant2.0, PHD-SNP^{g}, SNPs&GO, and PANTHER. Then, the secondary structure, amino acid conservation, and feature of amino acids including hydrophobicity, size, and polar contacts of 18 amino acid substitutions in protein were investigated via NetSurfP-2.0, ConSurf, HOPE, ExPASy, PEPTIDE 2.0, and PyMOL. Out of 353 missense variants, the 18 amino acid replacements including L16R/Q, G23D/R/S, L32P, N42K, G55D, G67D/R, P81R, H83R, G89D/S, A102E, G101R, G122R, and V126D were determined as the high-risk mutations. According to the previous studies, there is an association between ten amino acid replacements (L16R, G23D/R/S, L32P, G67R, H83R, G89D, G101R, and V126D) and some diseases including melanoma, pancreatic cancer, and HNSCC.

Conflict of interests

The authors declare no conflict of interest.

References

1. Mahjoub S, Ghali R, Chaeib V, Achour B, Megdich F. CDKN2A and CDKN2B Gene Variants in Acute Lymphoblastic Leukemia in Tunisian Population. J Leuk. 2018;6(247):2.
2. Lin Y, Diccianni M, Kim Y, Lin H, Lee C, Lin R, et al. Human p16 γ, a novel transcriptional variant of p16 INK4A, coexpresses with p16 INK4A in cancer cells and inhibits cell-cycle progression. Oncogene. 2007;26(49):7017-27.
3. Li M, Yang J, Liu K, Yang J, Zhan X, Wang L, et al. p16 promotes proliferation in cervical carcinoma cells through CDK6-HuR-IL1A axis. Journal of Cancer. 2020;11(6):1457-67.
4. Seif AA, Aly HH, Elzoghby DM, Elbreedy AM, Lotfy M. Aberrant p16 methylation as an early diagnostic marker in blood of hepatocellular carcinoma patients. Egyptian Journal of Medical Human Genetics. 2019;20(1):27.
5. Carrasco Salas P, Fernandez L, Vela M, Bueno D, Gonzalez B, Valentin J, et al. The role of CDKN2A/B deletions in pediatric acute lymphoblastic leukemia.

[DOI: 10.18502/ijpho.v11i2.5844]

Pediatric hematology and oncology. 2016;33(7-8):415-22.
6. Jia BY, Yang RH, Jiao WJ, Tian KH. Investigation of the effect of P14 promoter aberrant methylation on the biological function of human lung cancer cells. Thoracic cancer. 2019;10(6):1388-94.
7. Shima K, Nosho K, Baba Y, Cantor M, Meyerhardt JA, Giovannucci EL, et al. Prognostic significance of CDKN2A (p16) promoter methylation and loss of expression in 902 colorectal cancers: Cohort study and literature review. International journal of cancer. 2011;128(5):1080-94.
8. Mounawar M, Mukeria A, Le Calvez F, Hung RJ, Renard H, Cortot A, et al. Patterns of EGFR, HER2, TP53, and KRAS mutations of p14arf expression in non–small cell lung cancers in relation to smoking history. Cancer research. 2007;67(12):5667-72.
9. Ghiorzo P, Gargiulo S, Pastorino L, Nasti S, Cusano R, Bruno W, et al. Impact of E27X, a novel CDKN2A germ line mutation, on p16 and p14ARF expression in Italian melanoma families displaying pancreatic cancer and neuroblastoma. Human molecular genetics. 2006;15(18):2682-9.
10. Vasen HF, Wasser M, Van Mil A, Tollenaar RA, Konstantinovski M, Gruis NA, et al. Magnetic resonance imaging surveillance detects early-stage pancreatic cancer in carriers of a p16-Leiden mutation. Gastroenterology. 2011;140(3):850-6.
11. Zhou C, Shen Z, Ye D, Li Q, Deng H, Liu H, et al. The association and clinical significance of CDKN2A promoter methylation in head and neck squamous cell carcinoma: a meta-analysis. Cellular Physiology and Biochemistry. 2018;50(3):868-82.
12. Robertson LB, Armstrong GN, Olver BD, Lloyd AL, Shete S, Lau C, et al. Survey of familial glioma and role of germline p16 INK4A/p14 ARF and p53 mutation. Familial cancer. 2010;9(3):413-21.
13. Kohno T, Yamada Y, Tawara M, Takasaki Y, Kamihira S, Tomonaga M, et al. Inactivation of p14ARF as a key event for the progression of adult T cell leukemia/lymphoma. Leukemia research. 2007;31(12):1625-32.
14. Kim M, Yim S-H, Cho N-S, Kang S-H, Ko D-H, Oh B, et al. Homozygous deletion of CDKN2A (p16, p14) and CDKN2B (p15) genes is a poor prognostic factor in adult but not in childhood B-lineage acute lymphoblastic leukemia: a comparative deletion and hypermethylation study. Cancer genetics and cytogenetics. 2009;195(1):59-65.
15. Naji P, Heidari MM, Khatami M, Zare-Zardini H, Chamani R. MicroRNAs as a New Molecular Biomarker for Diagnosis and Prognosis of T-cell Acute Lymphoblastic Leukemia (T-ALL): A Systematic Review. Iranian Journal of Pediatric Hematology and Oncology. 2020;10(3):184-99.
16. Mistry M, Zhukova N, Merico D, Rakopoulos P, Krishnatry R, Shago M, et al. BRAF mutation and CDKN2A deletion define a clinically distinct subgroup of childhood secondary high-grade glioma. Journal of clinical oncology. 2015;33(9):1015.
17. 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 research. 2017;45(W1):W222-W8.
18. Won A, Pripotnev S, Ruscito A, Ianoul A. Effect of point mutations on the secondary structure and membrane interaction of antimicrobial peptide anoplin. The Journal of Physical Chemistry B. 2011;115(10):2371-9.
19. Shastry BS. SNPs: impact on gene function and phenotype. Single Nucleotide Polymorphisms: Springer; 2009. p. 3-22.
20. Porto WF, Franco OL, Alencar SA. Computational analyses and prediction of

[DOI: 10.18502/ijpho.v11i2.5844]

guanylin deleterious SNPs. Peptides. 2015;69:92-102.
21. Pires AS, Porto WF, Franco OL, Alencar SA. In silico analyses of deleterious missense SNPs of human apolipoprotein E3. Scientific reports. 2017;7(1):1-9.
22. Wang D, Song L, Singh V, Rao S, An L, Madhavan S. SNP2Structure: a public and versatile resource for mapping and three-dimensional modeling of missense SNPs on human protein structures. Computational and structural biotechnology journal. 2015;13:514-9.
23. Azevedo L, Mort M, Costa AC, Silva RM, Quelhas D, Amorim A, et al. Improving the in silico assessment of pathogenicity for compensated variants. European Journal of Human Genetics. 2017;25(1):2-7.
24. Castellana S, Rónai J, Mazza T. MitImpact: an exhaustive collection of pre-computed pathogenicity predictions of human mitochondrial non-synonymous variants. Human mutation. 2015;36(2):E2413-E22.
25. Rajith B, Chakraborty C. Predicting the impact of deleterious mutations in the protein kinase domain of FGFR2 in the context of function, structure, and pathogenesis—a bioinformatics approach. Applied biochemistry and biotechnology. 2013;170(8):1853-70.
26. Tekcan A. In silico analysis of FMR1 gene missense SNPs. Cell biochemistry and biophysics. 2016;74(2):109-27.
27. Ou L, Przybilla MJ, Whitley CB. Phenotype prediction for mucopolysaccharidosis type I by in silico analysis. Orphanet journal of rare diseases. 2017;12(1):125.
28. Afkhami E, Heidari M, Khatami M, Ghadamyari F, Dianatpour S. Detection of novel mitochondrial mutations in cytochrome C oxidase subunit 1 (COX1) in patients with familial adenomatous polyposis (FAP). Clinical and Translational Oncology. 2019:1-11.
29. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic acids research. 2012;40(W1):W452-W7.
30. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nature methods. 2010;7(4):248-9.
31. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. 2015;31(16):2745-7.
32. Capriotti E, Fariselli P. PhD-SNPg: a webserver and lightweight tool for scoring single nucleotide variants. Nucleic acids research. 2017;45(W1):W247-W52.
33. 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(S3):S6.
34. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. Nature protocols. 2013;8(8):1551.
35. Capriotti E, Fariselli P, Casadio R. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic acids research. 2005;33(suppl_2):W306-W10.
36. Klausen MS, Jespersen MC, Nielsen H, Jensen KK, Jurtz VI, Sønderby CK, et al. NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. Proteins: Structure, Function, and Bioinformatics. 2019;87(6):520-7.
37. Arshad M, Bhatti A, John P. Identification and in silico analysis of functional SNPs of human TAGAP protein: A comprehensive study. PloS one. 2018;13(1):e0188143.
38. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology

[DOI: 10.18502/ijpho.v11i2.5844]

to estimate and visualize evolutionary conservation in macromolecules. Nucleic acids research. 2016;44(W1):W344-W50.
39. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC bioinformatics. 2010;11(1):1-10.
40. Xuan C, Jia K-G, Wang B-B, Bai X-Y, Gao G, Yang Q, et al. Identification of two novel mutations of the HOMEZ gene in Chinese patients with isolated ventricular septal defect. Genetic testing and molecular biomarkers. 2013;17(5):390-4.
41. Saraboji K, Gromiha MM, Ponnuswamy MN. Relative importance of secondary structure and solvent accessibility to the stability of protein mutants.: A case study with amino acid properties and energetics on T4 and human lysozymes. Computational Biology and Chemistry. 2005;29(1):25-35.
42. Trevino SR, Schaefer S, Scholtz JM, Pace CN. Increasing protein conformational stability by optimizing β-turn sequence. Journal of molecular biology. 2007;373(1):211-8.
43. Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, Mayrose I, et al. ConSurf: using evolutionary data to raise testable hypotheses about protein function. Israel Journal of Chemistry. 2013;53(3-4):199-206.
44. Rajasekaran R, Doss CGP, Sudandiradoss C, Ramanathan K, Sethumadhavan R. In silico analysis of structural and functional consequences in p16INK4A by deleterious nsSNPs associated CDKN2A gene in malignant melanoma. Biochimie. 2008;90(10):1523-9.
45. Nelson D, Cox M. Lehninger Principles of Biochemistry, 7th Edn New York. WH Freeman & Co Ltd.[Google Scholar]; 2017.
46. Hyland PL, Pfeiffer RM, Rotunno M, Hofmann JN, Liu C-S, Cheng W-L, et al. Constitutive mitochondrial DNA copy number in peripheral blood of melanoma families with and without CDKN2A mutations. Journal of carcinogenesis & mutagenesis. 2014;2014(Suppl 4).
47. Tsao MAT, Whitaker L, Yakobson E. Features associated with germline CDKN2A. J Med Genet. 2007;44:99-106.
48. Bruno W, Pastorino L, Ghiorzo P, Andreotti V, Martinuzzi C, Menin C, et al. Multiple primary melanomas (MPMs) and criteria for genetic assessment: MultiMEL, a multicenter study of the Italian Melanoma Intergroup. Journal of the American Academy of Dermatology. 2016;74(2):325-32.
49. Scaini MC, Rossi E, de Siqueira Torres PLA, Zullato D, Callegaro M, Casella C, et al. Functional impairment of p16INK4A due to CDKN2A p. Gly23Asp missense mutation. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2009;671(1-2):26-32.
50. Poi MJ, Knobloch TJ, Yuan C, Tsai M-D, Weghorst CM, Li J. Evidence that P12, a specific variant of P16INK4A, plays a suppressive role in human pancreatic carcinogenesis. Biochemical and biophysical research communications. 2013;436(2):217-22.
51. Harland M, Cust AE, Badenas C, Chang Y-M, Holland EA, Aguilera P, et al. Prevalence and predictors of germline CDKN2A mutations for melanoma cases from Australia, Spain and the United Kingdom. Hereditary cancer in clinical practice. 2014;12(1):20.
52. Gensini F, Sestini R, Piazzini M, Vignoli M, Chiarugi A, Brandani P, et al. The p. G23S CDKN2A founder mutation in high-risk melanoma families from Central Italy. Melanoma research. 2007;17(6):387-92.
53. Jovanovic B, Egyhazi S, Eskandarpour M, Ghiorzo P, Palmer JM, Scarrà GB, et al. Coexisting NRAS and BRAF mutations in primary familial melanomas with specific CDKN2A germline alterations. The Journal of

[DOI: 10.18502/ijpho.v11i2.5844]

investigative dermatology. 2010;130(2):618.
54. Hansson J. Familial melanoma. Surgical Clinics of North America. 2008;88(4):897-916.
55. Lang J, Boxer M, MacKie R. CDKN2A mutations in Scottish families with cutaneous melanoma: results from 32 newly identified families. British Journal of Dermatology. 2005;153(6):1121-5.
56. Lowder CY, Dhir T, Goetz AB, Thomsett HL, Bender J, Tatarian T, et al. A step towards personalizing next line therapy for resected pancreatic and related cancer patients: A single institution's experience. Surgical Oncology. 2020;33:118-25.
57. Goldstein AM, Stacey SN, Olafsson JH, Jonsson GF, Helgason A, Sulem P, et al. CDKN2A mutations and melanoma risk in the Icelandic population. Journal of medical genetics. 2008;45(5):284-9.
58. Cabanillas R, Astudillo A, Valle M, de la Rosa J, Álvarez R, Durán NS, et al. Novel germline CDKN2A mutation associated with head and neck squamous cell carcinomas and melanomas. Head & neck. 2013;35(3):E80-E4.
59. Nikolaou V, Kang X, Stratigos A, Gogas H, Latorre M, Gabree M, et al. Comprehensive mutational analysis of CDKN2A and CDK4 in Greek patients with cutaneous melanoma. British Journal of Dermatology. 2011;165(6):1219-22.
60. Zhen DB, Rabe KG, Gallinger S, Syngal S, Schwartz AG, Goggins MG, et al. BRCA1, BRCA2, PALB2, and CDKN2A mutations in familial pancreatic cancer: a PACGENE study. Genetics in Medicine. 2015;17(7):569-77.
61. Burgstaller-Muehlbacher S, Marko M, Müller C, Wendt J, Pehamberger H, Okamoto I. Novel CDKN2A mutations in Austrian melanoma patients. Melanoma research. 2015;25(5):412-20.

[DOI: 10.18502/ijpho.v11i2.5844]