



BioBacta



Journal of Medical and Life Science

www.jmals.journals.ekb.eg



In silico study of gene-gene interaction of PIK3R1 and PIGF deleterious SNPs in correlation to the preeclampsia risk and its haplotype effect on antihypertensive treatment

Mohammed Y. Basher^{1*}, Asia M. Elrashied², Sahar G. Elbager³, Safinaz I. Khalil⁴

¹Faculty of Veterinary Medicine, University of Khartoum, Khartoum, Sudan

²Faculty of Science, University of Khartoum, Khartoum, Sudan

³Faculty of Medical Laboratory Sciences, University of Medical Sciences and Technology (UMST), Khartoum, Sudan.

⁴Faculty of Medicine, Al-Fajr College for Science and Technology, Khartoum, Sudan.

*Corresponding author: mohammedyousifbasher@gmail.com

DOI: [10.21608/jmals.2021.179658](https://doi.org/10.21608/jmals.2021.179658)

Abstract:

Introduction: The PIK3R1 (phosphoinositide-3-kinase regulatory subunit 1) and PIGF (Placenta growth factor) genes share some common pathways with eNOS, (Endothelial nitric oxide synthase) that plays a vital role in angiogenesis of blood vessels and regulating endothelial function. The PIK3R1 gene encodes the receptor for the p85 α regulatory subunit of the phosphoinositide-3-kinase (PI3K), that involved in endothelial cell migration. The PIGF gene encodes for the placenta growth factor, a homolog of vascular endothelial growth factor (VEGFA) that is involved in angiogenesis. PIK3R1 and/or PIGF mutations may cause dysregulation of eNOS contributing to endothelial dysfunction in preeclampsia. This study aims to analyze the effect of mutation of the PIK3R1 and PIGF genes on the structure and function of PIK3R1 and PGF protein that may have an important role in pathogenesis of preeclampsia.

Methodology: The data on human PIK3R1 and PGF genes were retrieved from dbSNP/NCBI. Ten prediction algorithms; SIFT, PROVEAN, Polyphen, SNAP2, SNPs&GO, PANTHER PhD-SNP, I-Mutant, Mutpred, and Hope were used to analyzing the effect of nsSNPs on functions and structure of the PGF and PIK3R1 protein. STRING and KEGG databases were used for PGF and PIK3R1 protein-protein interaction.

Results and Discussion: As per the dbSNP database, the human PIK3R1 gene contained 365 missense mutations. A total 3 nsSNPs (T239M, S229W, E47K) and 2 nsSNPs (H125Y, V59G) were predicted to have the most damaging effects on the structure and function of PIK3R1 and PGF respectively. STRING and KEGG revealed that PIK3R1 and PGF had strong interactions with proteins involved in the VEGF signaling pathway and PI3K-Akt signaling pathway. PIK3R1 is confirmed to be linked to important diseases like preeclampsia and can affect its treatment response. **Conclusion:** Gene-gene interaction is an important factor in preeclampsia treatment, the effect of mutation of the PIK3R1 and PIGF genes on the structure and function of PIK3R1 and PGF protein may have an important role in pathogenesis of preeclampsia. Also, these are linked to its treatment effect. This document gives formatting instructions for authors preparing papers for publication in the journal. Authors are encouraged to prepare manuscripts directly using this template. This template demonstrates the format requirements for the Journal.

Keywords: PGF, PIK3R1, Preeclampsia, Antihypertensive

1. Introduction

Preeclampsia is a multisystem pregnancy disorder that occurs in, diagnosed after the 20th gestation week and causes problems for both maternal and fetal, complicating 6% to 8% of pregnancies. In most cases, it may lead to maternal and fetal morbidity and mortality [1-3]. In preeclampsia (PE), cytotrophoblasts (CTBs) infestations are restricted, reducing placental perfusion and oxidative stress, and may lead to intermittent flow [4-7]. It can occur due to both fetoplacental and maternal factors, and also from genetic factors. In the worst cases, severe forms of PE could lead to systemic endothelial dysfunction, eclampsia, and maternal and/or fetal death.[7, 8]. Angiogenesis is a compound biological marvel decisive for correct embryonic development and post-natal growth. Amongst the several molecular companies elaborate in angiogenesis, some members of the VEGF family, VEGF-A, VEGF-B and placenta growth factor (PlGF), PIK3R1, and the related receptors VEGF receptor 1 and VEGF receptor 2 have a decisive role [9-10]. In this appraisal, we designate the discovery of the effect of single polymorphism mutation of the PIK3R1 and PGF genes on the structure and function of PIK3R1 and PGF protein that may have an important role in pathogenesis of preeclampsia.

PIK3R1 gene is known to be associated with preeclampsia-related biological processes [11]. The phosphoinositide-3-kinase regulatory subunit 1 (PIK3R1) gene, located on chromosome five (5q13.1) contains 22 exons, encodes for p85 α , p50 α , and p55 α isoforms [12]. This gene encodes the regulatory subunit of the phosphoinositide-3-kinase (PI3K) [13]. Since the PI3K gene is involved in endothelial cell migration, PIK3R1 mutations can cause dysregulation of the PI3K gene may contributing to endothelial dysfunction in preeclampsia.

The regulatory subunit is an 85 kD has an amino acid length of 724 consists of five domains, including SH3 domain, BH domain, and SH2(nSH2, iSH2, and cSH2) domains [14] Figure 1. The SH3 domain Interacts with LYN; this enhances enzyme activity. The proline-rich regions

(BH) domain mediates binding to XB-1, Rac, Cdc42, Rab5, PTEN. The nSH2 domain interacts with the helical domain of catalytic subunit alpha (p110 α) [14,15]. The SH2 domains interact with CCDC88A/GIV (tyrosine-phosphorylated form); the interaction enables recruitment of PIK3R1 to the EGFR receptor, enhancing PI3K activity and cell migration. The primary functions of the PIK3R1 are to regulate PI3K signaling, the pi3k-akt signal transduction pathway, receptor trafficking, and regulation of the unfolded protein response. The alternatively spliced of pik3r1 gene, resulting in four major protein-encoding transcripts (isoforms) p85 β , p55 α , p50 α , or p55 γ [15].

Placenta growth factor (PlGF) has been the second member of the VEGF family discovered. The human plgf gene mapped to chromosome 14q24 [16]. It contains different isoforms due to alternative splicing are encoded by the human plgf gene. It encodes four isoforms, PlGF 1-4 (composed of 131, 152, 203, and 224 amino acids after the elimination of signal peptide (18 amino acid residues in length), correspondingly [16,17]. Placental growth factor (PlGF), an angiogenic factor, The level of PlGF is reduced in patients who will subsequently develop preeclampsia, particularly in early-onset preeclampsia [18].

Although there is a relationship between preeclampsia and the amount of pik3r1 regulatory subunit and level of PlGF, there is not enough structural analysis study, to analyze the effects of mutations on the 3D structure of the subunit, and the consequence of this alteration on enzyme pathways. Hence this study aims to analyze the effect of mutation of the PIK3R1 and PGF genes on the structure and function of PIK3R1 and PGF protein that may have an important role in pathogenesis of preeclampsia.



Figure 1. Schematic primary structure of PIK3R1 protein. SH3 domain, BH domain, and SH2 domains are illustrated.

Pharmacogenomics may aid to guide the tailored treatment for non-responsive hypertensive patients.

There is an indication for the association of genetic variants with variable responses to the most frequently used antihypertensive drugs. Particularly, gene-gene interaction analyses have been applied to pharmacogenetic studies, including antihypertensive drug response[19].

2. Material and method

2.1 Retrieval of databases:

The data on the human PIK3R1 gene (Gene ID: 5295) was retrieved from the Entrez Gene databases from National Center for Biological Information (NCBI) database on 2 Jun 2020. The PIK3R1 protein sequence (accession ID: P27986) and SNPs information of the PIK3R1 gene were obtained from UniProtKB databases (<http://www.uniprot.org>) and NCBI dbSNP (<http://www.ncbi.nlm.nih.gov/snp/>) respectively.

2.2. Identification of deleterious nsSNPs

To determine the functional impact (deleterious, damaging, or natural), coding SNPs were analyzed using five different tools (SIFT, PROVEAN, Polyphen -2, SNAP2). nsSNPs predicted to be deleterious by at least three of these tools were categorized as high-risk nsSNPs, and subjected for further analysis like an association with disease, stability analysis, and structural effect, and protein-protein interaction using different tools.

2.2.1 SIFT (Sorting intolerant from tolerant) server was used to identify the tolerated and deleterious SNPs. The effect of amino acid substitution on protein structure was assessed based on the degree of conservation of amino acids using sequence homology Substitution of an amino acid at each position with probability < 0.05 is predicted to be deleterious and intolerant, while probability ≥ 0.05 is considered as tolerant, nsSNPs within dbSNP retrieved data were selected as an input for SIFT [20].

2.2.2 PROVEAN (Protein Variation Effect Analyzer) is used to predict the possible impact of substituted amino acid and indels on protein structure and biological function. It analyses the nsSNPs as deleterious or natural, if the final score was below the threshold score of -2.5 was considered deleterious; scores above this threshold were considered neutral

[21]. The input query is a protein FASTA sequence along with amino acid substitutions

2.2.3 Polyphen -2 (polymorphism and phenotype) server was used to predict the functional impact of amino acid substitution on protein structure and function based on sequence-based characterization. The Prediction outcome was obtained in the form of a probability score which classifies the variations as ‘probably damaging’, ‘possibly damaging’, and ‘benign’ [22]. A protein FASTA sequence along with the position and name of wild type and variant amino acids of screened nsSNPs were submitted as a query.

2.2.4 SNAP2 (Screening of Non-acceptable Polymorphism 2) is a tool, developed based on a neural network classification method that is freely available. It predicts the effect of nsSNPs on protein function [23]. The input query submitted is the protein FASTA sequence and lists of mutants which provided scores of each substitution that can then be translated into binary predictions neutral or non-neutral effect natural.

2.3 Prediction of Disease-Associated SNPs

The SNPs occurring in the protein-coding region may lead to deleterious consequences in its 3D structure and thus may lead to disease-associated phenomena. Here we used SNPs&GO, PANTHER, and Phd-SNP is tools to examine the disease-associated nsSNP occurring in the PIK3R1 protein-coding region.

2.3.1 SNPs&GO (Single Nucleotide Polymorphism Database (SNPs) & Gene Ontology (GO) is a support vector machine (SVM) based on the method to accurately predict the disease-related mutations from protein sequence. FASTA sequence of whole protein is considered to be an input option and output will be the prediction results based on the discrimination among disease-related and neutral variations of a protein sequence. A probability score higher than 0.5 reveals the disease-related effect of mutation on the parent protein function [24].

2.3.2 PANTHER (Protein Analysis Through Evolutionary Relationships) is a protein family and subfamily database that predicts whether the SNPs will cause a deleterious functional effect on the protein [25]. PANTHER output is subPSEC (substitution position-specific evolutionary

conservation) score which is the negative logarithm of the probability ratio of the wild-type and mutant amino acid at a particular position. The subPSEC scores are continuous values from 0 (neutral) to about 10 (most likely to be deleterious).

2.3.3 PhD SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms) webserver was used to predict if a given single point protein mutation can be classified as a disease-related or as neutral polymorphism. This server was mainly based on the support vector machines which can corroborate all the information regarding variations from the existing databases [26]. The input FASTA sequences of protein along with the residues change were submitted to the PhD-SNP server for analysis.

2.4 Prediction of nsSNPs Impact on the Protein Stability by I-Mutant2.0

I-Mutant2.0 is a tool used for the prediction of changes in protein stability due to single-site mutations under different conditions. It is a web server based on a support vector machine that worked on a dataset derived from Protherm, a database of experimental records on protein mutations. It can predict the stability changes in protein with 80% accuracy based on its structure and with 77% accuracy based on its sequence [27]. The input can be submitted either in the form of a protein sequence or on a structured basis. For the present study, input was submitted in the form of a protein FASTA sequence.

2.5 Analyzing the Effect of nsSNPs on Physiochemical Properties by Mutpred

Mutpred was used to predict structural and functional changes as a consequence of amino acid substitution [28]. These changes were expressed as probabilities of gain or loss of structure and function. In addition, it predicts the molecular cause of disease. The MutPred output contains a general score (g), i.e., the probability that the AAS is deleterious/ disease-associated and top five property scores (p), where p is the P-value that certain structural and functional properties are impacted. A missense mutation with a MutPred (g) score > 0.5 could be considered as "harmful," while a (g) score > 0.75 should be considered high confidence "harmful" prediction. The input was submitted in the form of protein

FASTA sequence along with an amino acid substitution

2.6 Prediction of structural effect of point mutation on the protein sequence using Project HOPE

Have Our Protein Explained (HOPE) is an easy-to-use web service that analyzes the structural effects of a point mutation in a protein sequence. FASTA sequence of whole protein and selection of mutant variants were submitted to project hope server. HOPE server predicts the output in the form of structural variation between mutant and wild-type residues [29].

2.7 Prediction of Protein-Protein Interactions by STRING and KEGG

Protein-protein interactions are important to assess all functional interactions among cell proteins.

2.7.1 STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) [30]. The STRING database gives a protein-protein interaction either it is direct or indirect associations. The input option we use is the protein name and the organism.

2.7.2 KEGG (Kyoto Encyclopedia of Genes and Genomes) [31] is a knowledge base for systematic analysis of gene functions in terms of the networks of genes and molecules, including metabolic pathways, regulatory pathways, and molecular complexes for biological systems

3. Results

3.1. SNP dataset from dbSNP

As per the dbSNP database, the human PIK3R1 gene investigated in this work contained a total of 15674 SNPs: 4 SNPs are initiator codon variant, 334 SNPs in 3' UTR region, 59 SNPs in 5' UTR region, 14817 SNPs in intron region, 185 SNPs in coding synonymous regions and 365 missense mutations SNPs (nsSNPs). We selected missense nsSNPs for our investigation.

3.2. Prediction of functional mutations

A total number of 365 nsSNPs PIK3R1 genes were submitted as a batch to the SIFT program. According to SIFT analysis, 30 SNPs were predicted to be deleterious, 256 SNPs were predicted to be tolerated and 79 nsSNPs were not found. Deleterious SNPs

were submitted to PROVEAN, Polyphen-2, and SNAP.

Out of 13 nsSNPs predicted to be deleterious by SIFT in PIK3R1, PROVEAN analysis predicted that 6nsSNPs were deleterious and 7nsSNPs were neutral. Polyphen-2 results analysis predicted that 4nsSNPs were benign, whereas 8nsSNPs were predicted to be probably damaging and the remaining one nsSNPs was predicted to be possibly damaging. The SNAP2 Analyzer predicted that 5 nsSNPs were affecting the function of protein and 8nsSNPs were neutral (Table 1).

In PGF, out of 9 nsSNPs predicted to be deleterious by SIFT, PROVEAN analysis predicted that 6 nsSNPs were deleterious and 3 nsSNPs were neutral. Polyphen-2 results analysis predicted that 2 nsSNPs were benign, whereas the remaining 7nsSNPs were predicted to be probably damaging The SNAP2 Analyzer predicted that 7nsSNPs were affecting the function of protein and 2nsSNPs were neutral (Table 1).

SNPs were considered high risk as predicted damaging by at least three tools. For PIK3R1, we classified 6 nsSNPs to be high deleterious while inPGF 7 nsSNPs were classified to be highly

deleterious. All the 11 nsSNPs were subjected to further analysis.

3.3 Prediction of Disease-Associated SNPs

SNPs&GO, PANTHER, and PhD-SNP were used to validate the results obtained from four previous tools. SNPs were considered to be associated with a disease if two or more algorithms predicted It as a disease associated. Out of 11 nsSNPs that were predicted to be deleterious, SNPs&GO predicted 2 nsSNP to be associated with the disease while PANTHER predicted 6 nsSNP and PhD-SNP predicted 7 nsSNP to be associated with the disease. In total Out of 11 high-risk nsSNPs, 5 nsSNP were predicted to be associated with disease (Table 2).

Alternative splicing of the PIK3R1 gene results in 5 transcript variants encoding different isoforms. Out of the 5 nsSNP associated with diseases, 3 nsSNPs (T239M, S229W, E47K) are located in isoform1 the canonical sequence, V87L in isoform 2, and V57L in isoform 3.

Table 1. List of nsSNP analysis by SIFT, PROVEAN, PolyPhen-2, SNAP2

SNP ID	AA CHANGE	SIFT PREDICTION	PROVEAN PREDICTION	POLY-2 PREDICTION	SNAP2 PREDICTION
PIK3R1					
rs141974044	R90W	Deleterious	Deleterious	Possibly Damaging	Effect
rs141981005	T239M	Deleterious	Deleterious	Probably Damaging	Effect
rs147795824	V357L	Deleterious	Deleterious	Benign	Neutral
rs150689648	I22T	Deleterious	Neutral	Probably Damaging	Neutral
rs188617536	K103T	Deleterious	Neutral	Benign	Neutral
rs188617536	K430T	Deleterious	Deleterious	Probably Damaging	Neutral
rs193921079	L347I	Deleterious	Neutral	Benign	Neutral
rs193921079	L647I	Deleterious	Neutral	Probably Damaging	Neutral
rs201598843	S229W	Deleterious	Neutral	Probably Damaging	Effect
rs370778654	G41A	Deleterious	Neutral	Probably Damaging	Neutral
rs371778244	P305A	Deleterious	Deleterious	Benign	Neutral
rs373560990	E47K	Deleterious	Neutral	Probably Damaging	Effect
rs55896055	R274G	Deleterious	Deleterious	Probably Damaging	Effect
PGF					
rs371967102	A54V	Deleterious	Neutral	Probably Damaging	Neutral
rs151333373	H125Y	Deleterious	Deleterious	Probably Damaging	Effect
rs370924022	Q124R	Deleterious	Deleterious	Probably Damaging	Effect
rs141182822	R110G	Deleterious	Deleterious	Benign	Effect
rs147722227	R114W	Deleterious	Deleterious	Probably Damaging	Effect
rs142448465	R206W	Deleterious	Neutral	Benign	Neutral
rs202127516	R220W	Deleterious	Neutral	Probably Damaging	Effect
rs371087083	T121M	Deleterious	Deleterious	Probably Damaging	Effect
rs200471804	V59G	Deleterious	Deleterious	Probably Damaging	Effect

Table 2. nsSNP predicted as a disease associated by SNPs&GO, PANTHER, and PhD-SNP

SNP ID	AA Change	SNPs&GO	PANTHER	PhD-SNP
PIK3R1				
rs141974044	R90W	Neutral	Disease	Neutral
rs141981005	T239M	Neutral	Disease	Disease
s188617536	K430T	Neutral	Neutral	Disease
rs201598843	S229W	Neutral	Disease	Disease
rs373560990	E47K	Disease	Neutral	Disease
rs55896055	R274G		Neutral	Neutral
PGF				
rs151333373	H125Y	Disease	Disease	Disease
rs370924022	Q124R	Neutral	Neutral	Disease
rs141182822	R110G	Neutral	Neutral	Neutral
rs147722227	R114W	Neutral	Disease	Neutral
rs202127516	R220W	Neutral	Neutral	Neutral
rs371087083	T121M	Neutral	Disease	Neutral
rs200471804	V59G	Neutral	Disease	Disease

3.4 Prediction of nsSNPs Impact on the Protein Stability

All SNPs predicted to be associated with the disease were submitted to the I- mutant server to predict the impact of these SNPs on the protein stability. All SNPs were predicted to decrease the stability of the protein (Table 3).

Table 3: Prediction result of I-Mutant software.

SNP ID	Amino acid Change	Stability Prediction	Reliability Index (RI)	DDG Value Prediction
PIK3R1				
rs141981005	T239M	Decrease	1	-0.29
rs201598843	S229W	Decrease	5	-1.21
rs373560990	E47K	Decrease	7	-1.05
PGF				
rs151333373	H125Y	Decrease	7	-0.67
rs200471804	V59G	Decrease	10	-2.35

3.5 Analyzing the Effect of nsSNPs on Physiochemical Properties Mutpred

The result showed (T239M, and S229W) were harmful E47K was nonharmful. Additionally, the

possible molecular mechanism disrupted were predicted; Loss of catalytic residue at T239 (P = 0.0438) for the mutation T239M Gain of MoRF binding (P = 0.0345) and Loss of disorder (P = 0.039) for the mutation T239M. Table 4 summarizes the result obtained from the MutPred server.

Table 4. Analysis of the effect of nsSNPs in structure, function, and evolution by MutPred server.

AA Change	Probability of deleterious mutation	Actionable Hypotheses	Top 5 features
PIK3R1			
T239M	0.665 (Harmful)	Loss of catalytic residue at T239 (P = 0.0438)	Loss of catalytic residue at T239 (P = 0.0438) Gain of MoRF binding (P = 0.1732) Gain of helix (P = 0.1736) Loss of loop (P = 0.3664) Loss of phosphorylation at T239 (P = 0.597)
S229W	0.566 (Harmful)	Gain of MoRF binding (P = 0.0345) Loss of disorder (P = 0.039)	Gain of MoRF binding (P = 0.0345) Loss of disorder (P = 0.039) Loss of phosphorylation at S229 (P = 0.116) Loss of relative solvent accessibility (P = 0.1807) Loss of solvent accessibility (P = 0.1952)
E47K	0.439 (Nonharmful)		Gain of ubiquitination at E47 (P = 0.0094) Gain of glycosylation at E47 (P = 0.0728) Loss of disorder (P = 0.1541) Gain of helix (P = 0.1736) Loss of loop (P = 0.2237)
PGF			
H125Y	0.895		Gain of catalytic residue at L120 (P = 0.0757) Gain of phosphorylation at H125 (P = 0.0829) Gain of MoRF binding (P = 0.1165) Loss of disorder (P = 0.1175) Loss of sheet (P = 0.302)
V59G	0.802	Loss of stability (P = 0)	Loss of stability (P = 0) Gain of disorder (P = 0.0551) Loss of sheet (P = 0.0817) Loss of MoRF binding (P = 0.1293) Gain of helix (P = 0.1736)

3.6 Analyzing the structural effect of point mutation on the protein Structure

The project HOPE was used to investigate the structural effects of these amino acid substitutions. Its results showed that the common effect of this nsSNPs was an alteration in charge, size, and hydrophobicity between wild-type and mutant residues which might lead to loss of interactions with other molecules and loss of hydrogen bonds hence, these changes could affect the ability of PIK3R1 to interact the other proteins. Table 5

Table 5. Schematic structures of the wild-type residue (left) and mutant residue (right) amino acid for each Mutation.

Residue	Structure	Properties
PIK3R1 T239M		<ul style="list-style-type: none"> ➤ The mutant residue is bigger than the wild-type residue. ➤ The mutant residue is more hydrophobic than the wild-type residue ➤ The wild-type residue forms a hydrogen bond with Glutamine at position 235 ➤ The mutation is located within a domain, annotated in UniProt as Rho-GAP ➤ The wild-type residue is very conserved and buried in the core of a domain.
S229W		<ul style="list-style-type: none"> ➤ The mutant residue is bigger than the wild-type residue. ➤ The mutant residue is more hydrophobic than the wild-type residue. ➤ The wild-type residue forms a hydrogen bond with Serine at position 231. ➤ The mutation is located within a domain, annotated in UniProt as Rho-GAP. ➤ The mutated residue is located on the surface of the protein.
E47K		<ul style="list-style-type: none"> ➤ The mutant residue is bigger than the wild-type residue. ➤ The wild-type residue charge was NEGATIVE, the mutant residue charge is POSITIVE. ➤ The wild-type residue forms a hydrogen bond with Glycine at position 45 and a salt bridge with Arginine at position 79 ➤ The mutation is located within a domain, annotated in UniProt as SH3 that is important for Molecular Function and binding of other molecules.
PGF H125Y		<ul style="list-style-type: none"> ➤ The mutant residue is bigger than the wild-type residue. ➤ The mutant residue is more hydrophobic than the wild-type residue. ➤ The wild-type residue forms a hydrogen bond with Cysteine at position 83 ➤ The wild-type residue is very conserved and located in a domain that is important for the activity of the protein and in contact with residues in another domain.
V59G		<ul style="list-style-type: none"> ➤ The mutant residue is smaller than the wild-type residue. ➤ The wild-type residue is more hydrophobic than the mutant residue. ➤ The wild-type residue is very conserved and located in a domain that is important for the activity of the protein and in contact with residues in another domain.

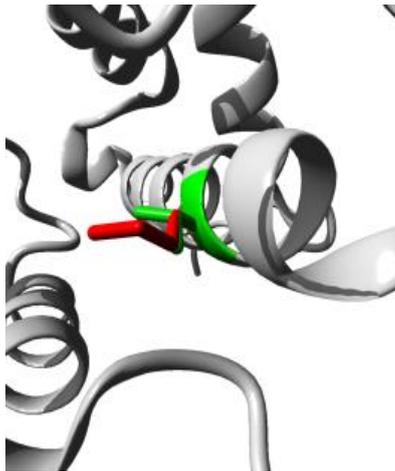


Figure 1: Showing 3D structure of T239M substitutions. The wild-type (green) and the mutant residue (red). SNP ID: rs141981005, protein position 239 changed from Threonine to Methionine.

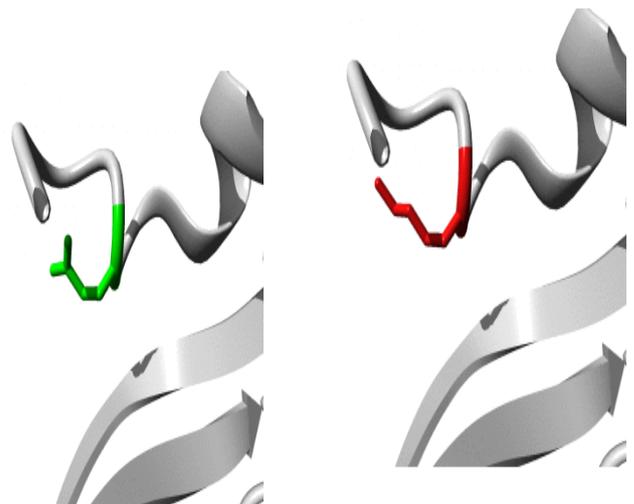


Figure 4: Showing 3D structure of E47K substitution. The wild-type (green) and the mutant residue (red). SNP ID: rs373560990, protein position 47 changed from Glutamic Acid to Lysine.

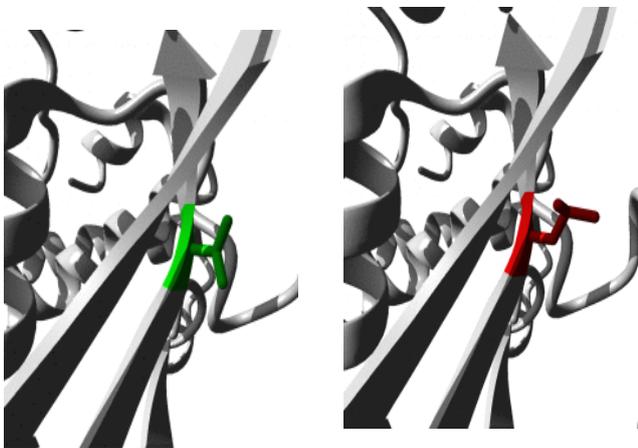


Figure 2: Showing 3D structure of V57L/ V87L substitutions. The wild-type (green) and the mutant residue (red). SNP ID: rs147795824, protein position 57/ 87 changed from Valine to Lysine.



Figure 5: Showing 3D structure of H125Y. The wild-type (green) and the mutant residue (red). SNP ID: rs151333373, protein position 125 changed from Histidine to a Tyrosine.

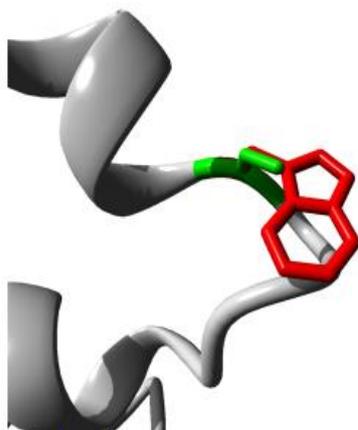


Figure 3: Showing 3D structure of S229W. The wild-type (green) and the mutant residue (red). SNP ID: rs141981005, protein position 229 changed from Serine to Tryptophan.

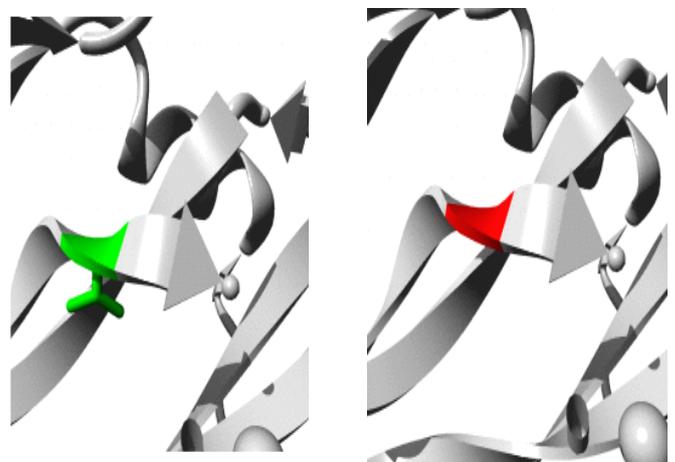


Figure 6: Showing 3D structure of V59G substitution. The wild-type (green) and the mutant residue (red). SNP ID: rs200471804, protein position 59 changed from Valine to Glycine.

3.7 Protein-protein interactions analysis

STRING interaction analysis revealed that PIK3R1 and PGF gene involved in many molecular and biological process and has high-confidence interactions with eNOS, FLT4, KDR, KIT, PDGFRB, PGF, PIK3CA, PIK3R1, TEK. (fig.7).

KEGG analysis showed that the same set of proteins are involved in the VEGF signaling pathway and PI3K-Akt signaling pathway (fig8.).

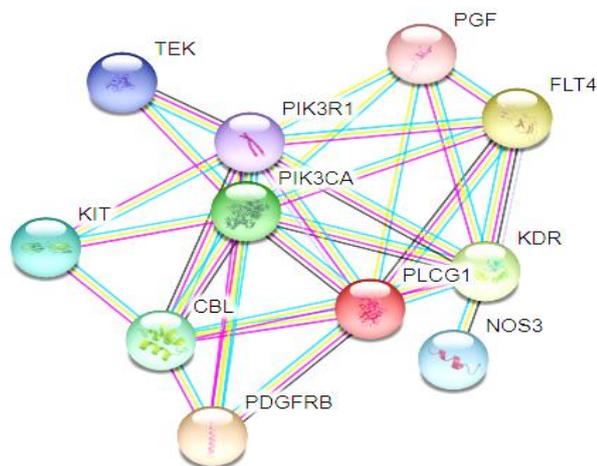


Figure7: Functional interaction between PIK3R1/PGF and its related proteins.

●	FLT1	Vascular endothelial growth factor receptor 1; Tyrosine-protein kinase that acts as a cell-surface receptor for VEGFA, VEGFB a...	● ● ●	0.999
●	KDR	Vascular endothelial growth factor receptor 2; Tyrosine-protein kinase that acts as a cell-surface receptor for VEGFA, VEGFC a...	● ● ●	0.984
●	VEGFA	Vascular endothelial growth factor A; Growth factor active in angiogenesis, vasculogenesis and endothelial cell growth. Induc...	● ● ● ● ●	0.954
●	VEGFC	Vascular endothelial growth factor C; Growth factor active in angiogenesis, and endothelial cell growth, stimulating their prolif...	● ● ● ● ●	0.927
●	VEGFB	Vascular endothelial growth factor B; Growth factor active in angiogenesis, lymphangiogenesis and endothelial cell growth, sti...	● ● ● ● ●	0.926
●	VEGFR	Vascular endothelial growth factor R; Growth factor for endothelial cells. VEGF-B167 binds heparin and neuropilin-1 whereas L...	● ● ● ● ●	0.922
●	PIK3CA	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform; Phosphoinositide-3-kinase (PI3K) that phosph...	● ● ●	0.915
●	FLT4	Vascular endothelial growth factor receptor 3; Tyrosine-protein kinase that acts as a cell-surface receptor for VEGFC and VEG...	● ● ● ● ●	0.910
●	PIK3R1	Phosphatidylinositol 3-kinase regulatory subunit alpha; Binds to activated (phosphorylated) protein-Tyr kinases, through its SH...	● ● ● ● ●	0.906
●	PLCG1	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1; Mediates the production of the second messenger mol...	● ● ● ● ●	0.902

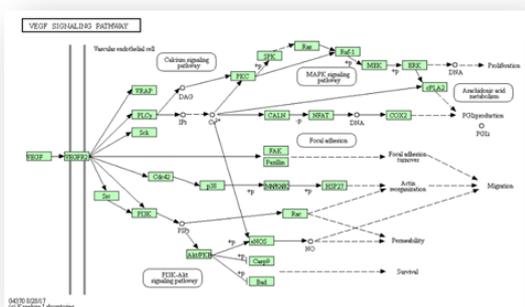


Figure 8: VEGF signaling pathway (KEGG ID: hsa04370)

4. Discussion

Over the past time, genome-wide association studies (GWAS) have permitted significant advances in the understanding of the genetic basis of many common diseases. GWAS is established on the idea that the

genetic foundation of many common diseases is complex and polygenic with many alternatives, each with modest effects that contribute to disease risk. Using this approach in preeclampsia, a large genome-wide association study (GWAS) recently recognized and simulated the first robust fetal genomic region associated with excess risk [32]. The current study undertakes a systematic in silico approach to screen the functional mutation in the human PIK3R1 gene linked to preeclampsia for a better understanding of how do these mutations affect the protein function and structure and hence affect the treatment outcome in such a serious condition

Ten prediction algorithms; SIFT, PROVEAN, Polyphen, SNAP2, SNPs&GO, PANTHER PhD-SNP, I-Mutant, Mutpred, and Hope were select for the screening of function mutation in the PGF gene, by comparing the scores of 10 tools, 2 nsSNPs with positions H125Y and V59G were found highly deleterious and decrease effective stability of PGF.

With regards to, PIK3R1 gene by comparing the scores of all 10 tools, 5 nsSNPs with positions T239M, S229W, E47K, V87L, and V57L were found to be highly deleterious and decrease effective stability of the PIK3R1.

The common physiochemical effect of this nsSNPs was an alteration in charge, size, and hydrophobicity between wild-type and mutant residues which might lead to loss of interactions with other molecules and loss of hydrogen bonds hence, these changes could affect the ability of PIK3R1 and PGF to interact the other proteins.

The majority of post-translational modifications effects this nsSNPs were altered disorders in protein structure, alteration in phosphorylation sites, gain or loss of MoRF activity, gain or loss glycosylation. loss of catalytic residue, a gain of ubiquitination, Loss of relative solvent accessibility gain.

If the mutation is predicted to altered disorders in protein structure, the mutation will have considered likely to cause disease [33]. Altered protein disorder was observed for substitution of H125Y and V59G for the PGF gene. Altered protein disorder was observed for substitution T239M, S229W, E47K for the PIK3R1 gene.

Phosphorylation of proteins is an important regulatory mechanism as it acts as their molecular switch to perform various functions during signal transduction pathways [34]. Variation in the phosphorylation sites can the enzymatic activity of a protein, disrupt the interactions between two or more proteins [35]. There is data in the literature documents that both gain and loss of a phosphorylation site may represent a molecular cause of disease for some inherited and somatic

F

mutations[36]. Gain of phosphorylation at H125 ($P = 0.0829$) and Gain of disorder ($P = 0.0551$), for PGF gene. Loss of phosphorylation was observed for substitutions T239M and S229W, for the PIK3R1 gene.

Molecular recognition features (MoRFs) are short interaction prone segments intrinsically disordered regions (IDPs) in protein sequences that fold upon binding to their interaction partners [37]. Intrinsically disordered proteins (IDPs) have been implicated in many human diseases, including cancer, diabetes [38]. MoRFs are involved in protein-protein interactions, which serve as the first step in molecular recognition and binding functions[39]. H125Y Gain of MoRF binding ($P = 0.1165$), V59G loss of MoRF binding ($P = 0.1293$) for PGF gene. Gain of MoRF binding activity was observed for substitution T239M and S229W for PIK3R1.

Protein glycosylation is known as one of the major post-translational modifications, with significant effects on protein folding, conformation, distribution, stability, and enzyme activity [40]. Point mutations that generating glycosylation sites have a greater impact on function[41]. Alterations in glycosylation are common to contribute to disease [42]. Gain of glycosylation was observed for substitution E47K for PIK3R1.

Catalytic residues are typically defined as amino acid residues directly involved in the chemistry of catalysis and involved in substrate binding and protein stability[43]. Therefore, the loss of catalytic residues may change the rate of enzymatic activity and catalytic reaction. There is evidence in the literature suggest that the loss and the gain of catalytic residues, may be actively involved in human-inherited disease [44]. loss of catalytic residue was observed for substitution T239M for PIK3R1.

Ubiquitylation is a post-translational modification that regulates several cellular mechanisms such as protein degradation, cell cycle division, the immune response [45]. Variation in the ubiquitination sites affects the degradation, changes the protein activity, and alters protein interactions. The abnormality of ubiquitylation is linked to human pathologies varying from inflammatory neurodegenerative diseases to different forms of cancers [46]. Gain of ubiquitination was observed for substitution E47K making the region prone to degradation.

Here we present the results upon each residue and discuss the conformational variations.

PIK3R1 SNPs

rs201598843 (S229W) this mutation resulted in a change of the Serine to Tryptophan at position 229. The Serine residue forms a hydrogen bond with

Serine at position 231. The Tryptophan residue is bigger and more hydrophobic than the Serine residue, this might lead to loss of hydrophobic interactions in the core of the protein. This substitution results in loss of disorder, phosphorylation, and relative solvent accessibility but the gain of MoRF binding. The mutation is located in the surface of the BH domain, involved in a wide variety of functions that usually require the interaction of the intact domain with another protein. S229W mutation might disturb this function.

rs373560990 (E47K) this substitution resulted in a change of Glutamic Acid to Lysine at position 47. The Glutamic Acid residue forms a hydrogen bond with Glycine at position 45 and a salt bridge with Arginine at position 79. The Glutamic Acid residue charge was negative; the lysine residue charge is positive, loss of hydrophobic interactions in the core of the protein. The majority of physiochemical properties of this substitution were Gain of ubiquitination and glycosylation.

This residue is located within the SH3 domain that is important for molecular function and binding of other molecules. E47K mutation might disturb this as such affect the function of the protein.

The main signaling networks downstream of Tpo/TPO-R include the JAK-STAT, MAPK, and PI3K-Akt pathways. These pathways control essential cellular processes like differentiation, proliferation, cell survival, and apoptosis with well-established roles in the initiation and progression of cancer.

PGF SNPs

rs151333373 (H125Y) this mutation resulted in a change of the Histidine to Tyrosine at position 125. The Histidine residue forms a hydrogen bond with Cysteine at position 83. The Histidine residue is bigger and more hydrophobic than the Tyrosine residue; this might lead to loss of hydrophobic interactions in the core of the protein. This substitution result in loss of disorder and sheet but the gain of catalytic residue, phosphorylation, and MoRF binding. The mutation is located in a domain that is important for the activity of the protein and in contact with residues in another domain. It is possible that this interaction is important for the correct function of the protein. (H125Y) mutation might disturb this function.

rs200471804 (V59G) this mutation resulted in a change of Valine to Glycine at position 59. The Valine residue is bigger and more hydrophobic than the Glycine residue; this might lead to loss of hydrophobic interactions in the core of the protein. This substitution result in loss of stability, sheet, and MoRF binding but the gain of disorder and helix. The mutation is located in the surface of the BH domain, involved in a wide variety of functions that

usually require the interaction of the intact domain with another protein. V59G mutation might disturb this function.

A study done in 2015, inspected differences in gene expression between preeclampsia patients and controls. Network analysis identified PIK3R1, VEGFA, and ITGB1, are known to be linked with preeclampsia or preeclampsia-related biological processes [47]. Our protein interactions analysis showed that the PIK3R1 and PGF protein PIK3R1 genes share common pathways with eNOS i.e VEGF signaling pathway and PI3K-Akt signaling pathway. eNOS enzyme is responsible for nitric oxide (NO) synthesis in the cardiovascular system. NO plays a vital role in the angiogenesis of blood vessels and regulating endothelial function. Reduced expression of eNOS, therefore, results in reduced NO bioavailability which plays an important role in the endothelial dysfunction associated with preeclampsia. Thus, eNOS is deliberated as a potential target for therapy for preeclampsia. So further studies can be done to confirm that PIK3R1 and PGF can affect the antihypertensive response in preeclampsia.

6. Conclusion

Gene- Gene interaction is an important factor in preeclampsia treatment, the effect of mutation of the PIK3R1 and PIGF genes on the structure and function of PIK3R1 and PGF protein that may have an important role in pathogenesis of preeclampsia. Also, these are linked to its treatment effect.

In this study, 2 SNPs are confirmed to be deleterious and associated with disease in the PGF gene while 5 SNPs are confirmed to cause disease in PIK3R1. eNOS is an enzyme accountable for nitric oxide (NO) synthesis in the cardiovascular system. NO plays a vital role in the angiogenesis of blood vessels and regulating endothelial function. Reduced expression of eNOS, therefore, results in reduced NO bioavailability which plays an important role in the endothelial dysfunction associated with preeclampsia. Thus, eNOS is deliberated as a potential target for therapy for preeclampsia and it has a common pathway with PIK3R1 and PGF. So further studies can be done to confirm that PIK3R1 and PGF can affect the antihypertensive response in preeclampsia.

Abbreviations:

PIK3R1(phosphoinositide-3-kinase regulatory subunit 1), **PIGF** (Placenta growth factor), **eNOS**(Endothelial nitric oxide synthase), **PROVEAN** (Protein Variation Effect Analyzer), **SIFT** (Sorting intolerant from tolerant), **Polyphen** - **2** (polymorphism and phenotype), **SNAP2** (Screening of non-acceptable Polymorphism 2),

SNPs&GO (Single Nucleotide Polymorphism Database (SNPs) & Gene Ontology (GO)), **PANTHER** (Protein Analysis Through Evolutionary Relationships), **PhD SNP** (Predictor of human Deleterious Single Nucleotide Polymorphisms), **HOPE**(Have Our Protein Explained), **STRING** (Search Tool for the Retrieval of Interacting Genes/Proteins) and **KEGG** (Kyoto Encyclopedia of Genes and Genomes).

References

- [1] Redman, C.W., and I.L.J.S. Sargent, *Latest advances in understanding preeclampsia*. 2005. **308**(5728): p. 1592-1594
- [2] Roberts, J.M., et al., *Preeclampsia: an endothelial cell disorder*. 1989. **161**(5): p. 1200-1204.
- [3] Sibai, B.M., et al. What we have learned about preeclampsia. in *Seminars in perinatology*. 2003. Elsevier.
- [4] Redman, C. and I.J.P. Sargent, *Placental stress and pre-eclampsia: a revised view*. 2009. **30**: p. 38-42.
- [5] Roberts, J.M. and C.J.P.H.A.I.J.o.W.s.C.H. Escudero, *The placenta in preeclampsia*. 2012. **2**(2): p. 72-83.
- [6] Zhou, Y., et al., Reversal of gene dysregulation in cultured cytotrophoblasts reveals possible causes of preeclampsia. 2013. **123**(7): p. 2862-2872.
- [7] Tsai, S., et al., Transcriptional profiling of human placentas from pregnancies complicated by preeclampsia reveals dysregulation of sialic acid acetyltransferase and immune signaling pathways. 2011. **32**(2): p. 175-182.
- [8] Cross, J.J.C.g., The genetics of pre-eclampsia: a fetoplacental or maternal problem? 2003. **64**(2): p. 96-103.
- [9] 9. eNOS haplotypes affect the responsiveness to antihypertensive therapy in preeclampsia but not in gestational hypertension.
- [10] Sandrim V C, Palei A C T, Luizon M R, Izidoro-Toledo T C, Cavalli R C, Tanus-Santos J E in *The pharmacogenomics journal* (2010)
- [11] Chudasama, K.K., et al., SHORT syndrome with partial lipodystrophy due to impaired phosphatidylinositol 3 kinase signaling. 2013. **93**(1): p. 150-157.
- [12] Conley, M.E., et al., Agammaglobulinemia and absent B lineage cells in a patient lacking the p85 α subunit of PI3K. 2012. **209**(3): p. 463-470.
- [13] Simoncini, T., et al., Interaction of oestrogen receptor with the regulatory subunit of phosphatidylinositol-3-OH kinase. 2000. **407**(6803): p. 538.

- [14] Panchamoorthy G, Fukazawa T, Stolz L, et al. Physical and functional interactions between SH2 and SH3 domains of the Src family protein tyrosine kinase p59fyn. *Mol Cell Biol.* 1994;14(9):6372-6385. doi:10.1128/mcb.14.9.6372
- [15] Groveman BR, Xue S, Marin V, et al. Roles of the SH2 and SH3 domains in the regulation of neuronal Src kinase functions. *FEBS J.* 2011;278(4):643-653. doi:10.1111/j.1742-4658.2010.07985.x
- [16] Sandro De Falco. et al 2012. The discovery of placenta growth factor and its biological activity. *Exp Mol Med.* 44(1):1-9.
- [17] Obinnaya Odigboegwu, MD, Lu J. Pan, MD, and Piyali Chatterjee, PhD. et al 2018 Use of Antihypertensive Drugs During Preeclampsia. *Frontiers in Cardiovascular Medicine.*
- [18] Catherine M. Brown and Vesna D. Garovic. et al 2011. Mechanisms and Management of Hypertension in Pregnant Women. *Curr Hypertens Rep* 338-346
- [19] Pharmacogenomics of Hypertension and Preeclampsia: Focus on Gene-Gene Interactions. et al 2018. *Front. Pharmacol.*, | <https://doi.org/10.3389/fphar.2018.00168>.
- [20] P.C. Ng SH. Predicting the effects of amino acid substitutions on protein function. *Annu Rev Genomics Hum Genet.* 2006;7 61-80.
- [21] Hoi Y and Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics.* 2015; 31(16): 2745-2747.
- [22] I.A. Adzhubei SS, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010; 7: 248-9.
- [23] Hecht L, Wass J, Kelly L, Clevenger-Firley E, Dunn C. SNAP-Ed Steps to Health Inspires Third Graders to Eat Smart and Move More. *Journal of Nutrition Education and Behavior.* 2013;45(6):800-2.
- [24] 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 10.1186/1471-2164-14-S3-S6
- [25] The PANTHER database of protein families, subfamilies, functions, and pathways.
- [26] *Nucleic Acids Res.*, 33 (Suppl. 1) (2005), pp. D284-D288
- [27] Capriotti E CRaCR. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics.* 2006; 22: 2729-34.
- [28] 27. E. Capriotti PF, R. Casadio I-Mutant 2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res:* 2005; 33: W306-W310.
- [29] Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics.* 2009; 25(21): 2744-50.
- [30] 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: 548.
- [31] Mering V. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res* 2005; 33(Suppl. 1): D433-D7.
- [32] Kanehisa, Minoru and Susumu Goto. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic acids research.* 1999; 27 (1): 29-34.
- [33] .
- [34] Vacic V., Markwick P. R. L., Oldfield C. J., Zhao X., Haynes C., Uversky V. N., and Iakoucheva L. M. Disease-associated mutations disrupt
- [35] Nishi H, Fong JH, Chang C, Teichmann SA, Panchenko AR. Regulation of protein-protein binding by coupling between phosphorylation and intrinsic disorder: analysis of human protein complexes. *Mol Biosyst.* 2013; 9:1620-1626.
- [36] Michael J. Lee and Michael B. Yaffe. Protein Regulation in Signal Transduction. *Cold Spring Harb Perspect Biol.* 2016; 8(6): a005918.
- [37] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011; 144:646-674.
- [38] Cheng Y., LeGall T., Oldfield C.J., Dunker A.K., Uversky V.N. Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry.* 2006; 45:10448-10460.
- [39] 40. Kotta-Loizou ITsaousis GN, Hamodrakas SJ. Analysis of Molecular Recognition Features (MoRFs) in membrane proteins. *Biochim Biophys Acta.* 2013 Apr;1834(4):798-807.
- [40] . Iakoucheva L.M., Brown C.J., Lawson J.D., Obradovic Z., Dunker A.K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* 2002; 323:573-584.

- [41] Skropeta, D. The effect of individual N-glycans on enzyme activity. *Bioorg. Med. Chem.* 2009; 17, 2645–2653.
- [42] Hassan Lemjabbar-Alaoui, Andrew McKinney, Yi-Wei Yang, Vy M. Tran, and Joanna J. Phillips. Glycosylation alterations in lung and brain cancer. *Adv Cancer Res.* 2015;126:305-44.
- [43] Hakomori SI, & Cummings RD. Glycosylation effects on cancer development. *Glycoconjugate Journal.* 2012; 29(8-9), 565–566.
- [44] Bartlett GJ, et al. Analysis of catalytic residues in enzyme active sites. *J. Mol. Biol.* 2002; 324:105–121.
- [45] Fuxiao Xin, Steven Myers, Yong Fuga Li, David N Cooper, Sean D Mooney, Predrag Radivojac. Structure-based kernels for the prediction of catalytic residues and their involvement in human inherited disease. *Bioinformatics.* 2010; 26, 1975–1982.
- [46] Elsasser S, Gali RR, Schwickart M, Larsen CN, Leggett DS, et al. Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat. Cell Biol.* 2002; 4:725–30.
- [47] Schnell JD, Hicke L (September 2003). "Non-traditional functions of ubiquitin and ubiquitin-binding proteins". *The Journal of Biological Chemistry.* 278 (38): 35857–60.
- [48] Marcelo R. Luizon, Daniela A. Pereira, and Valeria C. Sandrim et al 2018. Pharmacogenomics of Hypertension and Preeclampsia: Focus on Gene–Gene Interactions. *Front Pharmacol* 9:168