Novel *mutations within PRSS1* Gene that could potentially cause hereditary pancreatitis: Using Comprehensive in silico Approach

Mujahed I. Mustafa^{1,2*}, Abdelrahman H. Abdelmoneim², Nafisa M. Elfadol², Soada A. osman², Tebyan A. Abdelhameed², Mohamed A. Hassan²

1-Department of Biochemistry, University of Bahri, Sudan

2-Department of Biotechnology, Africa city of Technology, Sudan

*Corresponding author: Mujahed I. Mustafa: email: mujahedimustafa@gmail.com

Abstract:

Background: Hereditary pancreatitis (HP) is an autosomal dominant disorder with incomplete penetrance characterized by recurring episodes of severe abdominal pain often presenting in childhood. The comprehensive in silico analysis of coding SNPs, and their functional impacts on protein level, still remains unknown. In this study, *we aimed to identify the pathogenic SNPs in PRSS1 gene by computational analysis approach.* **Materials and Methods:** *We carried out in silico analysis of structural effect of each SNP using different bioinformatics tools to predict Single-nucleotide polymorphisms influence on protein structure and function.* **Result:** *Two novel mutations out of 339 nsSNPs that are found be deleterious effect on the PRSS1 structure and function.* **Conclusion:** *This is the first in silico analysis in PRSS1 gene, which will be a valuable resource for future targeted mechanistic and population-based studies.*

Keywords: Hereditary pancreatitis, autosomal dominant disorder, comprehensive computational analysis, Single-nucleotide polymorphisms, *PRSS1*.

1. Introduction:

Hereditary pancreatitis (HP) is an autosomal dominant disorder with incomplete penetrance characterized by recurring episodes of severe abdominal pain often presenting in childhood (1-5) First symptoms begin since childhood, mainly before 10 years old. Main symptoms are pancreatic pain and acute pancreatitis (>70%). CP morphological changes as pancreatic calcifications are diagnosed at a median age of 22-25 years. Exocrine and endocrine pancreatic insufficiency occurred in 34% and 26% at a median age of 29 and 38 years.(6, 7) The first family of hereditary pancreatitis was described in 1952.(8, 9) About 100 families have been documented worldwide.(4) which has been screened in different countries (3, 10-16)

HP is a progressive inflammatory disease in which pancreatic secretory parenchyma is destroyed and replaced by fibrous tissue(17), eventually leading to malnutrition and diabetes.(18) several studies shows that, Patients with HP have a markedly increased risk of pancreatic cancer compared with the general population up to 40-50% .(1, 16, 19-22) Also cigarette Smoking increases the risk factor for developing pancreatic cancer in Patients with Hereditary Pancreatitis.(23, 24)

Currently, there is no treatment for HP but it can be managed by pancreatic enzyme replacement therapy and analgesics are offered to control pain. In addition, endoscopic retrograde cholangiopancreatography (ERCP) and surgical intersections are reserved for all relevant complications. Patients should avoid tobacco use and alcohol exposure. Unfortunately, surgical removal of affected pancreatic tissue does not necessarily guarantee the patient's long healthy life.(25-31)

The term hereditary pancreatitis has primarily been associated with mutations in the serine protease 1 gene (*PRSS1*) which has been mapped to chromosome 7 q35.(5, 20) *PRSS1* is the most reported gene of hereditary pancreatitis.(1, 5-7, 15, 26, 30, 32-36) which encoding human cationic trypsinogen has been conclusively associated with autosomal dominant hereditary pancreatitis and sporadic nonalcoholic chronic pancreatitis. Most high-penetrance *PRSS1* variants increase intrapancreatic trypsin activity by stimulating trypsinogen autoactivation and/or by inhibiting chymotrypsin C-dependent trypsinogen degradation. (37) Interesting, mutations in the gene may protect one against the disease.(38, 39) In several studies, (R122H) mutation is strongly associated with hereditary pancreatitis.(6, 25, 32, 40, 41)

The main objective of this study, to identify functional SNPs within dbSNP located in coding nonsynonymous, 3-untranslated and 5-untranslated regions of *PRSS1* gene. Also to identify the most deleterious SNPs that could be used as diagnostic markers, using bioinformatics approach.(42-44)

2. Material and Methods:

2.1. Data mining:

The data on human *PRSS1* gene was collected from National Center for Biological Information (NCBI) web site. (45) The SNP information (protein accession number and SNP ID) of the *PRSS1* gene was retrieved from the NCBI dbSNP (<http://www.ncbi.nlm.nih.gov/snp/>) and the protein sequence was collected from Swiss Prot databases.(46) (<http://expasy.org/>)

2.2. Non-Synonymous SNPs Analysis:

2.2.1. Assessment of the Functional Impacts of Deleterious nsSNPs:

SIFT (Sorting intolerant from tolerant):

Predicts whether an amino acid substitution affects protein function based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences.(47) SIFT scores <0.05 are predicted by the algorithm to be intolerant or deleterious amino acid substitutions, whereas scores >0.05 are considered tolerant. It is available at (<http://sift.bii.a-star.edu.sg/>).

Polyphen-2:

It is an online tool (48) to predict possible impact of an amino acid substitution on both structure and function of a human protein by analysis of multiple sequence alignment and protein 3D structure, in addition it calculates position-specific independent count scores (PSIC) for each of two variants, and then calculates the PSIC scores difference between two variants. The higher a PSIC score difference, the higher the functional impact a particular amino acid substitution is likely to have. Prediction outcomes could be classified as probably damaging, possibly damaging or benign according to the value of PSIC as it ranges from (0_1); values closer to zero considered benign while values closer to 1 considered probably damaging and also it can be indicated by a vertical black marker inside a color gradient bar, where green is benign and red is damaging. nsSNPs that predicted to be intolerant by Sift has been submitted to Polyphen as protein sequence in FASTA format that obtained from UniprotKB /Expasy after submitting the relevant ensemble protein (ESNP) there, and then we entered position of mutation, native amino acid and the new substituent for both structural and functional predictions. PolyPhen version 2.2.2 is available at (<http://genetics.bwh.harvard.edu/pph2/index.shtml>).

Provean:

Provean is an online tool (49) that predicts whether an amino acid substitution has an impact on the biological function of a protein grounded on the alignment-based score. The score measures the change in sequence similarity of a query sequence to a protein sequence homolog between without and with an amino acid variation of the query sequence. If the PROVEAN score ≤ -2.5 , the protein variant is predicted to have a “deleterious” effect, while if the PROVEAN score is > -2.5 , the variant is predicted to have a “neutral” effect. It is available at (<https://roslab.org/services/snap2web/>).

SNAP2:

Functional effects of mutations are predicted with SNAP2 (50) is a trained classifier that is based on a machine learning device called "neural network". It distinguishes between effect and neutral variants/non-synonymous SNPs by taking a variety of sequence and variant features into account. The most important input signal for the prediction is the evolutionary information taken from an automatically generated multiple sequence alignment. Also structural features such as predicted secondary structure and solvent accessibility are considered. If available also annotation (i.e. known functional residues, pattern, regions) of the sequence or close homologs are pulled in. In a cross-validation over 100,000 experimentally annotated variants, SNAP2 reached sustained two-state accuracy (effect/neutral) of 82% (at an AUC of 0.9). In our hands this constitutes an important and significant improvement over other methods. It is available at (<https://roslab.org/services/snap2web/>).

2.2.2. Prediction of Disease Related Mutations SNPs:

SNPs&GO:

SNPs&GO (51) is a support vector machine (SVM) based on the method to accurately predict the mutation related to disease from protein sequence. The input is the FASTA sequence of the whole protein, the output is based on the difference among the neutral and disease related variations of the protein sequence. The RI (reliability index) with value of greater than 5 depicts the disease related effect caused by mutation on the function of parent protein. The PHD SNP and PANTHER, algorithms were also used in the display of output.

P-Mut:

PMUT a web-based tool (52) for the annotation of pathological variants on proteins, allows the fast and accurate prediction (approximately 80% success rate in humans) of the pathological

character of single point amino acidic mutations based on the use of neural networks. It is available at (<http://mmb.irbbarcelona.org/PMut>).

3. Protein Stability Analysis:

I-Mutant 3.0:

I-Mutant 3.0 Is a neural network based tool for the routine analysis of protein stability and alterations by taking into account the single-site mutations. The FASTA sequence of protein retrieved from UniProt is used as an input to predict the mutational effect on protein stability.(53) It is available at (<http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>).

MUpro:

MUpro is a support vector machine-based tool for the prediction of protein stability changes upon nonsynonymous SNPs. The value of the energy change is predicted, and a confidence score between -1 and 1 for measuring the confidence of the prediction is calculated. A score <0 means the variant decreases the protein stability; conversely, a score >0 means the variant increases the protein stability.(54) It is available at (

Identification of Functional SNPs in Conserved Regions by using ConSurf server:

ConSurf web server provides evolutionary conservation profiles for proteins of known structure in the PDB. Amino acid sequences similar to each sequence in the PDB were collected and multiply aligned using CSI-BLAST and MAFFT, respectively. The evolutionary conservation of each amino acid position in the alignment was calculated using the Rate 4Site algorithm, implemented in the ConSurf web server. The algorithm takes explicitly into account the phylogenetic relations between the aligned proteins and the stochastic nature of the evolutionary process. Rate 4 Site assigns a conservation level for each residue using an empirical Bayesian inference. Visual inspection of the conservation patterns on the 3-dimensional structure often enables the identification of key residues that comprise the functionally-important regions of the protein.(55, 56) It is available at (<http://consurf.tau.ac.il/>).

BioEdit:

BioEdit is an easy-to-use biological sequence alignment editor. This free software is intended to supply a single program that can handle most simple sequence and alignment editing and manipulation functions as well as a few basic sequences analyses.(57) It is available for download at (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>).

4. GeneMANIA:

We submitted genes and selected from a list of data sets that they wish to query. GeneMANIA approach to know protein function prediction integrate multiple genomics and proteomics data sources to make inferences about the function of unknown proteins.(58) It is available at (<http://www.genemania.org/>)

5. Structural Analysis:

5.1. Detection of nsSNPs Location in Protein Structure:

Mutation3D is a functional prediction and visualization tool for studying the spatial arrangement of amino acid substitutions on protein models and structures. Mutation3D is able to separate functional from nonfunctional mutations by analyzing a combination of 8,869 known inherited disease mutations and 2,004 SNPs overlaid together upon the same sets of crystal structures and homology models. Further, it present a systematic analysis of whole-genome and whole-exome cancer datasets to demonstrate that mutation3D identifies many known cancer genes as well as previously underexplored target genes(59). It is available at (<http://mutation3d.org>).

5.2. Modeling nsSNP locations on protein structure:

Project hope is a new online web-server to search protein 3D structures (if available) by collecting structural information from a series of sources, including calculations on the 3D coordinates of the protein, sequence annotations from the UniProt database, and predictions by DAS services. Protein sequences were submitted to project hope server in order to analyze the structural and conformational variations that have resulted from single amino acid substitution corresponding to single nucleotide substitution. It is available at (<http://www.cmbi.ru.nl/hope>).

5.3. Modeling Amino Acid Substitution:

UCSF Chimera is a highly extensible program for interactive visualization and analysis of molecular structures. Chimera (version 1.8) software was used to scan the 3D (three-dimensional) structure of specific protein, and hence modifies the original amino acid with the mutated one to see the impact that can be produced. The outcome is then a graphic model depicting the mutation.(60) Chimera (version 1.8) available at (<http://www.cgl.ucsf.edu/chimera/>).

-	G49V	Damaging	0	damaging probably damaging	1	Deleterious	-8.061	effect	87
rs1368323821	G50S	Damaging	0	damaging probably damaging	0.999	Deleterious	-5.385	effect	70
rs1166804108	G50D	Damaging	0	damaging probably damaging	0.999	Deleterious	-6.281	effect	85
rs1415566919	S51Y	Damaging	0.01	damaging probably damaging	0.99	Deleterious	-5.092	effect	80
-	L52F	Damaging	0	damaging probably damaging	1	Deleterious	-3.628	effect	32
rs149246646	I53N	Damaging	0	damaging probably damaging	1	Deleterious	-6.266	effect	68
rs778570468	W57G	Damaging	0	damaging probably damaging	1	Deleterious	11.764	effect	83
rs1338646513	W57C	Damaging	0	damaging probably damaging	1	Deleterious	11.773	effect	63
rs1219784204	V58M	Damaging	0.01	damaging probably damaging	1	Deleterious	-2.61	effect	51
rs1192452565	V58G	Damaging	0	damaging probably damaging	1	Deleterious	-6.083	effect	73
rs370761165	A61E	Damaging	0	damaging probably damaging	1	Deleterious	-4.397	effect	80
rs776806418	H63R	Damaging	0	damaging probably damaging	0.997	Deleterious	-6.813	effect	86
rs745551657	V71M	Damaging	0	damaging probably damaging	0.996	Deleterious	-2.611	effect	26
rs1468857521	H76P	Damaging	0	damaging probably damaging	0.985	Deleterious	-8.063	effect	74
rs1254878660	H76Q	Damaging	0.01	damaging probably damaging	0.979	Deleterious	-6.192	effect	42
rs372411481	G83R	Damaging	0	damaging probably damaging	1	Deleterious	-7.011	effect	40
-	G83R	Damaging	0	damaging probably damaging	1	Deleterious	-7.011	effect	40
-	G83W	Damaging	0	damaging probably damaging	1	Deleterious	-7.169	effect	5
rs1383806868	I88N	Damaging	0.01	damaging probably damaging	1	Deleterious	-6.199	effect	58
rs1461511291	I94T	Damaging	0	damaging probably damaging	0.996	Deleterious	-4.602	effect	59
rs545007137	I94M	Damaging	0	damaging probably damaging	0.999	Deleterious	-2.719	effect	35
rs1268805560	H96Y	Damaging	0	damaging probably damaging	1	Deleterious	-5.652	effect	72
rs1209409723	H96Q	Damaging	0	damaging probably damaging	1	Deleterious	-7.515	effect	78
rs765315575	P97H	Damaging	0	probably	0.999	Deleterious	-7.798	effect	23

-	P97L	Damaging	0	damaging probably damaging	0.997	Deleterious	-8.746	effect	45
rs1454816504	Y99C	Damaging	0	damaging probably damaging	1	Deleterious	-8.117	effect	72
rs1323769980	D107Y	Damaging	0	damaging probably damaging	1	Deleterious	-8.448	effect	94
rs758965932	D107G	Damaging	0.01	damaging probably damaging	1	Deleterious	-6.581	effect	93
rs1482612529	L109T	Damaging	0	damaging probably damaging	0.998	Deleterious	-2.871	effect	32
rs1426710453	L113H	Damaging	0	damaging probably damaging	1	Deleterious	-6.567	effect	79
rs606231348	N120K	Damaging	0.05	damaging probably damaging	0.998	Deleterious	-4.991	effect	67
rs144403091	V123M	Damaging	0	damaging probably damaging	1	Deleterious	-2.664	effect	21
-	V123L	Damaging	0.01	possibly damaging probably	0.497	Deleterious	-2.558	effect	58
rs749518244	L128Q	Damaging	0	damaging probably damaging	1	Deleterious	-5.698	effect	49
-	L128P	Damaging	0	damaging probably damaging	1	Deleterious	-6.649	effect	74
rs768673799	P129H	Damaging	0	damaging probably damaging	1	Deleterious	-8.254	effect	45
rs117497341	T137R	Damaging	0.01	damaging probably damaging	0.98	Deleterious	-4.659	effect	77
-	T137M	Damaging	0	damaging probably damaging	0.997	Deleterious	-4.653	effect	42
rs748208676	I141N	Damaging	0	damaging probably damaging	1	Deleterious	-6.392	effect	89
-	I141T	Damaging	0	possibly damaging probably	0.905	Deleterious	-4.224	effect	76
rs772256079	I141M	Damaging	0	damaging probably damaging	0.991	Deleterious	-2.516	effect	67
rs1164996242	S142P	Damaging	0	damaging probably damaging	1	Deleterious	-4.777	effect	86
rs1164331073	W144C	Damaging	0	damaging probably damaging	1	Deleterious	-	effect	80
rs1221038304	G145R	Damaging	0	damaging probably damaging	1	Deleterious	-7.632	effect	91
rs769724683	T147N	Damaging	0.01	possibly damaging possibly	0.781	Deleterious	-4.615	effect	60
rs1355136604	P155T	Damaging	0.01	damaging probably damaging	0.785	Deleterious	-6.372	effect	63
rs1333246826	P155R	Damaging	0.01	damaging probably damaging	0.999	Deleterious	-7.093	effect	75
rs1172272446	C160R	Damaging	0	probably damaging	1	Deleterious	-	effect	28

				damaging			11.072		
-	C160G	Damaging	0	possibly	0.889	Deleterious	-	effect	40
rs778796800	C160Y	Damaging	0	damaging	1	Deleterious	11.071	effect	22
-	C160F	Damaging	0	probably	1	Deleterious	-	effect	41
rs777441424	L161R	Damaging	0.01	damaging	1	Deleterious	10.147	effect	29
rs202228418	V165A	Damaging	0	probably	0.834	Deleterious	-5.365	effect	50
rs776120935	L166P	Damaging	0.01	damaging	1	Deleterious	-3.478	effect	79
-	S167I	Damaging	0.01	possibly	0.909	Deleterious	-6.111	effect	19
rs763453065	S167R	Damaging	0.02	damaging	0.952	Deleterious	-5.101	effect	31
rs200973660	C171Y	Damaging	0	probably	1	Deleterious	-4.259	effect	80
-	C171S	Damaging	0	damaging	1	Deleterious	-	effect	78
rs1473604680	S174P	Damaging	0.03	probably	0.981	Deleterious	-9.272	effect	44
rs746669985	S174F	Damaging	0	damaging	0.995	Deleterious	-3.077	effect	29
rs756821075	Y175N	Damaging	0	probably	1	Deleterious	-4.013	effect	75
-	Y175H	Damaging	0	damaging	1	Deleterious	-8.298	effect	66
rs745420768	T180I	Damaging	0	probably	0.999	Deleterious	-4.616	effect	59
rs200070487	M183V	Damaging	0	damaging	0.983	Deleterious	-5.291	effect	64
rs762123204	M183K	Damaging	0	probably	0.997	Deleterious	-3.563	effect	85
-	M183T	Damaging	0	damaging	0.997	Deleterious	-5.395	effect	73
rs11553848	F184S	Damaging	0.02	probably	0.989	Deleterious	-5.361	effect	74
rs1439241122	L189F	Damaging	0.05	damaging	0.997	Deleterious	-4.165	effect	15
rs1366653442	K193E	Damaging	0.01	probably	0.962	Deleterious	-3.427	effect	79
rs1217657614	D194V	Damaging	0	damaging	1	Deleterious	-3.421	effect	85
rs763907908	C196G	Damaging	0	probably	1	Deleterious	-8.36	effect	93
rs1201496061	G198D	Damaging	0.02	damaging	0.97	Deleterious	-	effect	88

rs1412477456	G201R	Damaging	0	damaging	1	Deleterious	-7.194	effect	87
rs1288010897	G201V	Damaging	0	damaging	1	Deleterious	-8.04	effect	83
rs199531292	G202A	Damaging	0.02	damaging	0.997	Deleterious	-5.297	effect	57
rs1289842951	P203H	Damaging	0	damaging	1	Deleterious	-8.323	effect	65
rs747422004	V205D	Damaging	0	damaging	1	Deleterious	-5.836	effect	81
rs189270875	G208A	Damaging	0.02	damaging	0.998	Deleterious	-5.479	effect	53
rs1366495669	S215P	Damaging	0	damaging	1	Deleterious	-4.618	effect	93
rs1310865247	G217S	Damaging	0	damaging	0.985	Deleterious	-5.299	effect	84
rs1481112469	C220W	Damaging	0	damaging	1	Deleterious	10.127	effect	90
rs1164573795	P226S	Damaging	0	damaging	1	Deleterious	-7.318	effect	63
rs1309672836	Y229H	Damaging	0	damaging	1	Deleterious	-4.582	effect	85
-	V232F	Damaging	0.01	damaging	1	Deleterious	-4.531	effect	72

***SUB:** Substitutions (variants)

Table (2): List of non-synonymous SNPs analyzed for disease association by SNP&GO, PHD-SNP, PANTHER and P-mut:

dbSNP rs#	Mutation	PHD-SNP Prediction	RI	Probability	SNP&GO Prediction	RI	Probability	PANTHER Prediction	RI	Probability	P mut prediction	Score
rs1366278558	M1R	Disease	0	0.514	Disease	3	0.668	Disease	9	0.95	Disease	0.78 (88%)
rs767036052	L4P	Disease	3	0.656	Disease	1	0.54	Disease	3	0.626	Disease	0.70 (86%)
rs755628827	L5P	Disease	4	0.704	Neutral	1	0.433	Neutral	1	0.475	Neutral	0.43 (85%)
rs1222840668	D21N	Neutral	0	0.5	Disease	6	0.786	Disease	5	0.736	Neutral	0.05 (97%)
-	D21Y	Disease	6	0.781	Disease	7	0.871	Disease	9	0.929	Neutral	0.14 (94%)
rs564368252	V25A	Disease	3	0.648	Disease	1	0.575	Disease	2	0.62	Neutral	0.02 (98%)
rs766579839	G26R	Disease	7	0.853	Disease	6	0.795	Disease	8	0.911	Neutral	0.10 (96%)
rs1432168030	C30S	Disease	3	0.649	Disease	5	0.774	Disease	9	0.935	Neutral	0.02

												(99%)
rs780969708	C30W	Disease	6	0.825	Disease	7	0.848	Disease	10	0.99	Neutral	0.15 (94%)
rs769459903	P36R	Disease	4	0.708	Disease	1	0.558	Disease	5	0.734	Neutral	0.01 (99%)
rs775259109	Q38K	Disease	5	0.749	Disease	0	0.507	Neutral	2	0.378	Neutral	0.21 (92%)
-	S40C	Disease	4	0.716	Disease	2	0.576	Disease	6	0.822	Neutral	0.19 (93%)
-	S40F	Disease	5	0.742	Disease	2	0.59	Disease	1	0.573	Neutral	0.02 (98%)
rs1437198196	S43F	Disease	1	0.552	Disease	2	0.606	Disease	5	0.738	Neutral	0.04 (98%)
rs772636606	G44S	Disease	4	0.683	Disease	3	0.647	Disease	1	0.566	Neutral	0.02 (98%)
rs773739445	G44D	Disease	7	0.829	Disease	6	0.813	Disease	5	0.758	Neutral	0.02 (98%)
-	G44A	Disease	3	0.669	Disease	3	0.632	Disease	0	0.519	Neutral	0.04 (98%)
rs138464021	G49D	Disease	8	0.917	Disease	7	0.862	Disease	8	0.88	Neutral	0.11 (95%)
-	G49V	Disease	8	0.913	Disease	7	0.852	Disease	8	0.902	Neutral	0.11 (95%)
rs1368323821	G50S	Disease	8	0.895	Disease	6	0.808	Disease	0	0.507	Neutral	0.20 (92%)
rs1166804108	G50D	Disease	8	0.925	Disease	7	0.827	Disease	0	0.506	Neutral	0.20 (92%)
rs1415566919	S51Y	Disease	7	0.861	Disease	7	0.836	Disease	7	0.856	Neutral	0.18 (93%)
-	L52F	Disease	7	0.841	Disease	6	0.783	Disease	9	0.954	Neutral	0.18 (93%)
rs149246646	I53N	Disease	8	0.888	Disease	7	0.844	Disease	8	0.891	Neutral	0.18 (93%)
rs778570468	W57G	Disease	8	0.897	Disease	6	0.787	Disease	10	0.98	Neutral	0.18 (93%)
rs1338646513	W57C	Disease	9	0.942	Disease	7	0.849	Disease	10	0.993	Neutral	0.18 (93%)
rs1219784204	V58M	Disease	5	0.733	Disease	1	0.56	Disease	4	0.685	Neutral	0.09 (96%)
rs1192452565	V58G	Disease	6	0.818	Disease	4	0.695	Disease	7	0.848	Neutral	0.18 (93%)
rs370761165	A61E	Disease	8	0.891	Disease	5	0.762	Disease	6	0.779	Neutral	0.04 (98%)
rs776806418	H63R	Disease	7	0.836	Disease	5	0.771	Disease	10	1	Neutral	0.04 (98%)
rs745551657	V71M	Disease	3	0.642	Neutral	0	0.492	Disease	5	0.727	Neutral	0.07 (97%)
rs1468857521	H76P	Disease	8	0.904	Disease	6	0.815	Disease	4	0.693	Neutral	0.13 (94%)
rs1254878660	H76Q	Disease	4	0.71	Disease	1	0.529	Neutral	0	0.491	Neutral	0.03 (98%)
rs372411481	G83R	Disease	3	0.67	Disease	3	0.667	Disease	4	0.715	Neutral	0.01

												(93%)
rs1221038304	G145R	Disease	7	0.866	Disease	6	0.812	Disease	9	0.966	Neutral	0.04 (98%)
rs769724683	T147N	Disease	4	0.702	Disease	3	0.625	Disease	3	0.667	Neutral	0.09 (96%)
rs1355136604	P155T	Neutral	0	0.477	Neutral	5	0.261	Disease	0	0.521	Neutral	0.03 (98%)
rs1333246826	P155R	Disease	3	0.626	Neutral	2	0.408	Disease	1	0.533	Neutral	0.06 (97%)
rs1172272446	C160R	Disease	5	0.757	Disease	8	0.876	Disease	10	0.987	Neutral	0.07 (96%)
-	C160G	Disease	3	0.649	Disease	7	0.866	Disease	9	0.966	Neutral	0.04 (98%)
rs778796800	C160Y	Disease	3	0.674	Disease	7	0.834	Disease	10	0.992	Neutral	0.03 (98%)
-	C160F	Disease	5	0.76	Disease	7	0.866	Disease	10	0.99	Neutral	0.09 (96%)
rs777441424	L161R	Disease	5	0.739	Disease	7	0.846	Disease	7	0.864	Neutral	0.11 (95%)
rs202228418	V165A	Neutral	4	0.288	Neutral	7	0.167	Neutral	3	0.362	Neutral	0.06 (97%)
rs776120935	L166P	Disease	6	0.8	Disease	6	0.807	Disease	8	0.901	Neutral	0.12 (95%)
-	S167I	Disease	1	0.564	Neutral	2	0.399	Disease	4	0.696	Neutral	0.04 (98%)
rs763453065	S167R	Disease	1	0.568	Neutral	1	0.459	Disease	3	0.628	Neutral	0.04 (98%)
rs200973660	C171Y	Disease	8	0.925	Disease	7	0.827	Disease	10	1	Neutral	0.04 (98%)
-	C171S	Disease	8	0.895	Disease	5	0.772	Disease	10	1	Neutral	0.04 (98%)
rs1473604680	S174P	Disease	4	0.675	Disease	4	0.724	Disease	4	0.689	Neutral	0.04 (98%)
rs746669985	S174F	Disease	0	0.515	Disease	2	0.599	Disease	7	0.832	Neutral	0.02 (99%)
rs756821075	Y175N	Disease	7	0.835	Disease	4	0.698	Disease	7	0.873	Neutral	0.19 (93%)
-	Y175H	Disease	3	0.661	Disease	2	0.589	Disease	6	0.806	Neutral	0.02 (98%)
rs745420768	T180I	Disease	6	0.788	Disease	3	0.657	Disease	5	0.761	Neutral	0.10 (95%)
rs200070487	M183V	Disease	8	0.884	Disease	7	0.843	Disease	2	0.607	Neutral	0.07 (97%)
rs762123204	M183K	Disease	9	0.944	Disease	8	0.897	Disease	7	0.828	Neutral	0.18 (93%)
-	M183T	Disease	8	0.876	Disease	8	0.89	Disease	6	0.782	Neutral	0.00 (99%)
rs11553848	F184S	Disease	5	0.766	Disease	2	0.615	Disease	2	0.575	Neutral	0.11 (95%)
rs1439241122	L189F	Disease	0	0.524	Neutral	1	0.472	Disease	2	0.586	Neutral	0.03 (98%)
rs1366653442	K193E	Disease	4	0.692	Disease	5	0.772	Disease	2	0.603	Neutral	0.07

												(97%)
rs1217657614	D194V	Disease	8	0.904	Disease	7	0.872	Disease	8	0.898	Neutral	0.20
rs763907908	C196G	Disease	6	0.816	Disease	5	0.743	Disease	10	0.999	Neutral	(92%)
rs1201496061	G198D	Disease	8	0.881	Disease	5	0.743	Disease	0	0.519	Neutral	0.11
rs1412477456	G201R	Disease	8	0.882	Disease	7	0.839	Disease	7	0.849	Neutral	(95%)
rs1288010897	G201V	Disease	8	0.906	Disease	7	0.855	Disease	7	0.84	Neutral	0.00
rs199531292	G202A	Disease	5	0.775	Disease	4	0.696	Neutral	1	0.471	Neutral	(99%)
rs1289842951	P203H	Disease	7	0.835	Disease	4	0.701	Disease	10	0.979	Neutral	0.04
rs747422004	V205D	Disease	7	0.831	Disease	6	0.775	Disease	8	0.903	Neutral	(98%)
rs189270875	G208A	Disease	4	0.714	Disease	3	0.649	Neutral	0	0.476	Neutral	0.04
rs1366495669	S215P	Disease	8	0.878	Disease	7	0.827	Disease	6	0.804	Neutral	(93%)
rs1310865247	G217S	Disease	8	0.886	Disease	7	0.842	Disease	6	0.8	Neutral	0.04
rs1481112469	C220W	Disease	9	0.928	Disease	7	0.855	Disease	9	0.967	Neutral	(98%)
rs1164573795	P226S	Disease	7	0.871	Disease	6	0.787	Disease	9	0.933	Neutral	0.11
rs1309672836	Y229H	Disease	6	0.795	Disease	1	0.573	Disease	5	0.742	Neutral	(95%)
-	V232F	Disease	8	0.902	Disease	7	0.838	Disease	6	0.789	Neutral	0.18

***RI:** Reliability Index

Table (3): Protein Stability Analysis predicted by I-Mutant v3.0 and Mupro (also Show the two novel mutations):

dbSNP rs#	Amino Acid change	SVM2 Prediction Effect	RI	DDG Value Prediction	Mupro Prediction	Mupro Score
rs1366278558	M1R	Decrease	1	-0.62	Decrease	-1.018261
rs767036052	L4P	Decrease	8	-1.66	Decrease	-1.623945

***RI:** Reliability Index

Table (4): Conservation profile of amino acids in *PRSS1* that coincide in location with high-risk nsSNPs:

Residues position	Residues	CS Score normalized	Color	Residues Variety	B/E
1	M	0.211	7.3	M,N,P	e
4	L	0.361	6.3	M,S,L,I,H,F	b

CS Score normalized (1-4= variable, 5= average, 6-9= conserved).

B/E: Buried (b) or Exposed (e) residue.

Table (5): The *PRSS1* gene functions and its appearance in network and genome:

Function	FDR	Genes in network	Genes in genome
serine-type endopeptidase inhibitor activity	0.001356361	4	44
renal system process involved in regulation of systemic arterial blood pressure	0.001356361	3	13
regulation of hemostasis	0.001356361	4	54
regulation of coagulation	0.001356361	4	56
regulation of blood coagulation	0.001356361	4	54
regulation of wound healing	0.003496886	4	74
platelet activation	0.006856052	5	211
endopeptidase inhibitor activity	0.007599469	4	99
endopeptidase regulator activity	0.007599469	4	102
peptidase inhibitor activity	0.007599469	4	101
blood microparticle	0.008676071	4	108
regulation of endopeptidase activity	0.009340491	5	251
regulation of peptidase activity	0.009856636	5	258
peptidase regulator activity	0.013382385	4	128
renal system process	0.013666888	3	45
endocrine process	0.013666888	3	42
negative regulation of endopeptidase activity	0.013666888	4	140
regulation of blood vessel size	0.013666888	3	44
regulation of tube size	0.013666888	3	44
negative regulation of peptidase activity	0.013666888	4	142
regulation of systemic arterial blood pressure	0.013666888	3	43
platelet alpha granule lumen	0.01586832	3	48

vascular process in circulatory system	0.028313203	3	59
platelet alpha granule	0.029997483	3	61
secretory granule lumen	0.030241005	3	62
serine-type endopeptidase activity	0.033116705	3	67
killing of cells of other organism	0.033116705	2	10
exocytosis	0.033116705	4	191
disruption of cells of other organism	0.033116705	2	10
glomerular filtration	0.037840129	2	11
renal filtration	0.037840129	2	11
vesicle lumen	0.040192639	3	76
enzyme inhibitor activity	0.040192639	4	214
protein kinase C-activating G-protein coupled receptor signaling pathway	0.040192639	2	12
cytoplasmic membrane-bounded vesicle lumen	0.040192639	3	76
positive regulation of collagen metabolic process	0.044298773	2	13
positive regulation of collagen biosynthetic process	0.044298773	2	13
regulation of blood pressure	0.044298773	3	81
platelet degranulation	0.044771122	3	82
positive regulation of ERK1 and ERK2 cascade	0.04525757	3	83
negative regulation of multicellular organismal process	0.045315882	4	230
positive regulation of coagulation	0.048108433	2	15
negative regulation of hydrolase activity	0.048108433	4	242
collagen metabolic process	0.048108433	3	87
positive regulation of hemostasis	0.048108433	2	15
positive regulation of blood coagulation	0.048108433	2	15
positive regulation of multicellular organismal metabolic process	0.048108433	2	15
multicellular organismal macromolecule metabolic process	0.049591958	3	91
regulation of endocrine process	0.051022071	2	16
regulation of collagen biosynthetic process	0.051022071	2	16
tissue homeostasis	0.051393559	3	94
zymogen activation	0.052534202	3	97
regulation of collagen metabolic process	0.052534202	2	17
multicellular organismal metabolic process	0.052534202	3	97
positive regulation of release of sequestered calcium ion into cytosol	0.052534202	2	17
serine-type peptidase activity	0.052961891	3	98
serine hydrolase activity	0.055058507	3	101
cellular metal ion homeostasis	0.055058507	4	264
positive regulation of Ras protein signal transduction	0.055058507	2	18
collagen biosynthetic process	0.057743853	2	19
multicellular organismal homeostasis	0.057743853	3	105
regulation of multicellular organismal metabolic process	0.057743853	2	19
positive regulation of small GTPase mediated signal transduction	0.057743853	2	19
cobalamin metabolic process	0.0619817	2	20
fibrinolysis	0.0619817	2	20

cellular cation homeostasis	0.063339568	4	284
regulation of ERK1 and ERK2 cascade	0.063981924	3	111
metal ion homeostasis	0.064852029	4	288
cellular ion homeostasis	0.066330871	4	292
extracellular matrix organization	0.066330871	4	292
extracellular structure organization	0.066330871	4	293
ERK1 and ERK2 cascade	0.07491573	3	120
positive regulation of calcium ion transport into cytosol	0.094036985	2	26
positive regulation of cytosolic calcium ion concentration	0.094313684	3	131

***FDR:** false discovery rate is greater than or equal to the probability that this is a false positive.

Table (6): the gene co-expressed, share domain and Interaction with *MEFV* gene network:

Gene 1	Gene 2	Weight	Network group
F2R	PRSS1	0.025958601	Co-expression
PRSS2	PRSS1	0.03579626	Co-expression
PRSS2	F2R	0.025450852	Co-expression
KRT1	KRT2	0.021397633	Co-expression
PRSS3	PRSS1	0.033738803	Co-expression
PRSS3	PRSS2	0.033134628	Co-expression
ABCA4	PRSS1	0.020958284	Co-expression
ABCA4	PRSS2	0.021833802	Co-expression
ABCA4	PRSS3	0.020297663	Co-expression
F2R	F2RL1	0.01723061	Co-expression
SERPINF2	KRT1	0.005162225	Co-expression
PRSS3	PRSS2	0.030308522	Co-expression
ABCA4	PARD3	0.018756393	Co-expression
SERPINI1	F2RL1	0.018079963	Co-expression
PRSS2	PRSS1	0.010663025	Co-expression
KRT10	PRSS1	0.004751287	Co-expression
KRT1	KRT2	0.017388888	Co-expression
PRSS3	PRSS2	0.016354393	Co-expression
ALB	PRSS1	0.00978991	Co-expression
TMPRSS15	PRSS1	0.009836646	Co-expression
TMPRSS15	PRSS2	0.02349308	Co-expression
ONECUT2	PRSS1	0.011885054	Co-expression
ONECUT2	PRSS2	0.026482629	Co-expression
ONECUT2	ALB	0.023143126	Co-expression

ONECUT2	TMPRSS15	0.023315145	Co-expression
KRT2	PARD3	0.00407993	Co-expression
KRT2	F2RL3	0.002086929	Co-expression
KRT1	PARD3	0.008348665	Co-expression
KRT1	KRT2	0.001395255	Co-expression
SERPINF2	PRSS2	0.020101048	Co-expression
PRSS3	PRSS2	0.037385598	Co-expression
KRT2	F2RL1	0.004147036	Co-expression
KRT2	PARD3	0.004057049	Co-expression
PRSS2	SERPINB8	0.013921468	Co-expression
KRT1	PARD3	0.007636325	Co-expression
KRT1	F2RL3	0.001890512	Co-expression
KRT1	KRT2	0.000903638	Co-expression
SERPINF2	KRT2	0.001538189	Co-expression
PRSS3	SERPINB8	0.012880304	Co-expression
PRSS3	PRSS2	0.01267024	Co-expression
PRSS3	PRSS2	0.008046081	Co-expression
ALB	SERPINF2	0.012912146	Co-expression
PRSS3	PRSS2	0.059784386	Co-expression
ABCA4	PRSS1	0.021043219	Co-expression
PRSS2	PRSS1	0.017638749	Co-expression
PRSS3	PRSS1	0.010804671	Co-expression
PRSS3	PRSS2	0.008761909	Co-expression
ABCA4	TMPRSS15	0.014334416	Co-expression
KRT10	KRT2	0.013258844	Co-expression
KRT1	KRT2	0.012602167	Co-expression
KRT1	KRT10	0.010437634	Co-expression
PRSS3	PRSS2	0.020688048	Co-expression
APP	F2R	0.007284051	Co-expression
KRT1	SERPINI1	0.011799906	Co-expression
PRSS3	KRT2	0.01028535	Co-expression
PRSS3	PRSS2	0.056584958	Co-expression
SERPINB8	F2RL1	0.007467608	Co-expression
PRSS2	F2RL1	0.009436603	Co-expression
PRSS3	F2RL1	0.0091086	Co-expression
PRSS3	PRSS2	0.006767724	Co-expression
PRSS2	SERPINB8	0.005437602	Co-expression
CEP250	KRT1	0.007838975	Co-expression
ALB	APP	0.015414801	Co-expression
PRSS2	F2RL1	0.011420856	Co-expression
KRT10	KRT2	0.031353086	Co-expression
KRT1	SERPINB8	0.017290482	Co-expression
KRT1	KRT2	0.030572487	Co-expression

ALB	SERPINF2	0.011540175	Co-localization
F2R	F2RL3	0.004708032	Co-localization
PRSS2	PRSS1	0.025921002	Co-localization
KRT1	KRT10	0.025425598	Co-localization
PRSS3	PRSS1	0.023983192	Co-localization
PRSS3	PRSS2	0.023682926	Co-localization
ALB	SERPINF2	0.016614223	Co-localization
TMPRSS15	PRSS1	0.02293591	Co-localization
TMPRSS15	PRSS2	0.022705158	Co-localization
TMPRSS15	PRSS3	0.022577295	Co-localization
ONECUT2	PRSS1	0.022678863	Co-localization
ONECUT2	PRSS2	0.022438433	Co-localization
ONECUT2	PRSS3	0.02270408	Co-localization
ONECUT2	TMPRSS15	0.026521208	Co-localization
F2RL1	PRSS1	0.31954107	Pathway
PARD3	PRSS1	0.15361263	Pathway
F2RL3	PRSS1	0.1483753	Pathway
F2R	PRSS1	0.12154161	Pathway
F2R	F2RL3	0.03525233	Pathway
PRSS2	F2RL1	0.31954107	Pathway
PRSS2	PARD3	0.15361263	Pathway
PRSS2	F2RL3	0.1483753	Pathway
PRSS2	F2R	0.12154161	Pathway
F2RL1	PRSS1	0.7365974	Pathway
CEP250	PRSS1	0.08479679	Physical Interactions
CEP250	DDX5	0.021656184	Physical Interactions
ALB	PRSS1	0.0910383	Physical Interactions
ALB	KRT10	0.0910383	Physical Interactions
ALB	KRT1	0.04481957	Physical Interactions
ALB	PRSS3	0.0910383	Physical Interactions
DDX5	PRSS1	0.3375174	Physical Interactions
KRT10	KRT2	0.37723482	Physical Interactions
KRT1	KRT2	0.38693836	Physical Interactions
KRT1	KRT10	0.20642798	Physical Interactions
KRT2	PRSS1	0.28754097	Physical Interactions
KRT10	PRSS1	0.24892193	Physical Interactions
KRT10	KRT2	0.15666465	Physical Interactions
KRT1	PRSS1	0.24077669	Physical Interactions
KRT1	KRT2	0.15153825	Physical Interactions
KRT1	KRT10	0.13118546	Physical Interactions
SERPINB8	PRSS1	0.4951709	Physical Interactions
SERPINF2	PRSS1	0.28534	Physical Interactions
APP	PRSS1	0.41637108	Physical Interactions

PRSS3	APP	0.13214332	Physical Interactions
TMEM150B	PRSS1	0.6281027	Physical Interactions
ALB	APP	0.0232307	Physical Interactions
CEP250	APP	0.0232307	Physical Interactions
SERPINB8	PRSS1	0.18267727	Physical Interactions
APP	PRSS2	0.030401716	Physical Interactions
SERPINF2	PRSS1	0.0573582	Physical Interactions
PRSS3	APP	0.05769704	Physical Interactions
F2RL3	F2RL1	0.5863869	Predicted
F2R	F2RL1	0.17300643	Predicted
KRT1	KRT2	0.15433232	Predicted
SERPINI1	PRSS1	1	Predicted
			Shared protein domains
F2RL3	F2RL1	0.09312839	Shared protein domains
F2R	F2RL1	0.09312839	Shared protein domains
F2R	F2RL3	0.09312839	Shared protein domains
SERPINB8	SERPINI1	0.028828265	Shared protein domains
PRSS2	PRSS1	0.008043572	Shared protein domains
KRT10	KRT2	0.008737453	Shared protein domains
KRT1	KRT2	0.024351321	Shared protein domains
SERPINF2	SERPINI1	0.028828265	Shared protein domains
SERPINF2	SERPINB8	0.028828265	Shared protein domains
PRSS3	PRSS1	0.010362946	Shared protein domains
PRSS3	PRSS2	0.010437737	Shared protein domains
TMPRSS15	PRSS1	0.006456545	Shared protein domains
TMPRSS15	PRSS2	0.006503143	Shared protein domains
F2RL3	F2RL1	0.00354377	Shared protein domains
F2R	F2RL1	0.00354377	Shared protein domains
F2R	F2RL3	0.00354377	Shared protein domains
SERPINB8	SERPINI1	0.027777778	Shared protein domains

PRSS2	PRSS1	0.009871521	Shared protein domains
KRT10	KRT2	0.010121632	Shared protein domains
KRT1	KRT2	0.018882168	Shared protein domains
KRT1	KRT10	0.011125098	Shared protein domains
SERPINF2	SERPINI1	0.027777778	Shared protein domains
SERPINF2	SERPINB8	0.027777778	Shared protein domains
PRSS3	PRSS1	0.009871521	Shared protein domains
PRSS3	PRSS2	0.011186602	Shared protein domains

Figure 1: A graphical representation of the distribution of frame shift, missense, nonsense and synonymous SNPs for *PRSS1* gene in coding region (based on the dbSNP database).

***Note: the numbers represent in the chart, is numbers of SNPs.**

Figure2: Graphical representation of Disease variations by PHD-SNP, SNP&GO, PANTHER and P-mut.

Figure 3: Interaction between PRSS1 and its related genes.

Figure4: The amino acid Methionine change to Arginine at position 1.

Figure 5: The amino acid Leucine change to Proline at position 4.

Figure 6: Alignment of 10 amino acid sequences of *PRSS1* demonstrating that the residues predicted to be mutated (indicated by red arrow) are evolutionarily conserved across species. Sequences Alignment were done by BioEdit (version 7.2.5).

Figure 7: Structural models for wild type *PRSS1* and (L4P) high-risk nsSNP in its domain.

7. Discussion:

Two novel mutations have been found (*see table 3*) that effect on the stability and function of the *PRSS1* gene using bioinformatics tools. The methods used were based on different aspects and parameters describing the pathogenicity and provide clues on the molecular level about the effect of mutations. It was not easy to predict the pathogenic effect of SNPs using single method. Therefore, multiple methods were used to compare and rely on the results predicted. In this study we used different in silico prediction algorithms: SIFT, PolyPhen-2, Provean, SNAP2, SNP&GO, PHD-SNP, PANTHER, P-MUT and I-Mutant 3.0, MUpro, ConSurf and Mutation3D.

Single-nucleotide polymorphism (SNP) association studies have become crucial in revealing the genetic correlations of genomic variants with complex diseases. In silico analysis has been done for many disorders for cancer related genes and other disorders. (61-63) Therefore, we used deep comprehensive in silico analysis to study nsSNP in Homo sapiens located in coding region of *PRSS1* gene, were investigate in dbSNP/NCBI Database, out of 911 there are 339 nsSNPs (missense mutations),(see figure1) were submitted to SIFT server, PolyPhen-2 server, Provean sever and SNAP2 respectively, 141 SNPs were predicted to be deleterious in SIFT server. In PolyPhen-2 server, the result showed that were found to be damaging (54 possibly damaging and 116 probably damaging showed deleterious). In Provean server our result showed that 203 SNPs were predicted to be deleterious. While in SNAP2 server the result showed that 172 SNPs were predicted to be Effect. The differences in prediction capabilities refer to the fact that every prediction algorithm uses different sets of sequences and alignments. In table (2) we were submitted four positive results from SIFT, PolyPhen-2, Provean and SNAP2 (*see table 1*) to observe the disease causing one by SNP&GO, PHD-SNP, P-Mut and MUpro servers.

In SNP&GO, PHD-SNP, PANTHER and P-Mut servers were used to predict the association of SNPs with disease. According to SNP&GO, PHD-SNP, PANTHER and P-Mut (85, 95, 90 and 2 SNPs respectively) were found to be disease related SNPs. We selected the four positive results (disease related SNPs) for further analysis by I-Mutant 3.0 and MUpro, Table (3) (figure 2) While I-Mutant and MUpro results revealed that the protein stability decreased which destabilize the amino acid interaction, (see Table 3).

GeneMANIA revealed that *PRSS1* has many vital functions: blood microparticle, cobalamin metabolic process, extracellular matrix organization, extracellular structure organization, serine hydrolase activity, serine-type endopeptidase activity, serine-type peptidase activity. The genes co-expressed with, share similar protein domain, or participate to achieve similar function were illustrated by GeneMANIA and shown in figure 3, Tables (5&6).

Project HOPE server was used to submit the two deleterious nsSNPs: (rs1366278558): (M1R): Methionine residue changed to Arginine at position 1. (Figure 4) shows the schematic structures and homology modeling of the original and mutant amino acid. The backbone, which is the same for each amino acid, is colored red. The side chain, unique for each amino acid, is colored black. Each amino acid has its own specific size, charge, and hydrophobicity value. The mutation is located within the signal peptide. This sequence of this peptide is important because it is recognized by other proteins and often cleaved of to generate the mature protein. The new residue that is introduced in the signal peptide differs in its properties from the original one. It is possible that this mutation disturbs recognition of the signal peptide. The mutant residue is smaller than the wild-type, and this will cause a possible loss of external interactions. The wild-type residue is more hydrophobic than the mutant residue and this mutation might cause loss of hydrophobic interactions with other molecules on the surface of the protein and thus may cause function loss or change.

(rs767036052):(L4P): Leucine residue changed to Proline at position 4. (Figure 5) shows the schematic structures of the original and the mutant amino acid. The backbone, which is the same for each amino acid, is colored red. The side chain, unique for each amino acid, is colored black. The mutant residue is smaller and more hydrophobic than the wild-type residue. The mutation is located within the signal peptide. This sequence of this peptide is important because it is recognized by other proteins and often cleaved of to generate the mature protein. While the new residue that is introduced in the signal peptide differs in its properties from the original one. It is possible that this mutation disturbs recognition of the signal peptide. The wild-type and mutant amino acids differ in size. The mutant residue is smaller; this might lead to loss of interactions.

We also used ConSurf web server; the nsSNPs that are located at highly conserved amino acid positions tend to be more deleterious than nsSNPs that are located at non-conserved sites. Our ConSurf analysis revealed that (L4P) mutation was found in the highly conserved region and predicted to have a potential impact on *PRSS1* protein.(see table 4) To confirm our findings (M1R & L4P) mutations we used BioEdit (version 7.2.5) Alignment of 10 amino acid sequences of *PRSS1* confirm that, the residues predicted to be mutated are evolutionarily conserved across species. (See figure 6) Additionally, we performed analysis by Mutation3D, our result show that: (L4P) located in the domain. (See figure 7)

In the light of our study we recommend to perform whole exome sequencing on suspicious patients(3, 16, 17) and their parents for pathogenic variant (M1R & L4P) besides the ones that has been screened in other studies (6, 64) to allow early diagnosis and prevention and/or management of symptoms.

8. Conclusion:

A total of two nsSNPs (M1R & L4P) were predicted to be responsible for the structural and functional modifications of *PRSS1* protein. It is evident from the deep comprehensive in silico analysis that, these two nsSNPs might serve as, a novel biomarkers for the prognosis of

HP. These findings may also facilitate the development of novel therapeutic elements for associated diseases.

Conflict of interest:

The authors have declared that no competing interest exists

Data Availability:

All relevant data are within the paper and its supporting information files.

Acknowledgment:

The authors wish to acknowledgment the enthusiastic cooperation of Africa City of Technology - Sudan.

References:

1. Charnley RM. Hereditary pancreatitis. *World journal of gastroenterology*. 2003;9(1):1-4.
2. Jackson WD. Pancreatitis: etiology, diagnosis, and management. *Current opinion in pediatrics*. 2001;13(5):447-51.
3. Dai LN, Chen YW, Yan WH, Lu LN, Tao YJ, Cai W. Hereditary pancreatitis of 3 Chinese children: Case report and literature review. *Medicine*. 2016;95(36):e4604.
4. Le Bodic L, Bignon JD, Ragues O, Mercier B, Georgelin T, Schnee M, et al. The hereditary pancreatitis gene maps to long arm of chromosome 7. *Human molecular genetics*. 1996;5(4):549-54.
5. Dyrła P, Nowak T, Gil J, Adamiec C, Bobula M, Saracyn M. [Hereditary pancreatitis]. *Polski merkuriusz lekarski : organ Polskiego Towarzystwa Lekarskiego*. 2016;40(236):113-6.
6. Rebours V, Levy P, Ruszniewski P. An overview of hereditary pancreatitis. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*. 2012;44(1):8-15.
7. Howes N, Lerch MM, Greenhalf W, Stocken DD, Ellis I, Simon P, et al. Clinical and genetic characteristics of hereditary pancreatitis in Europe. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. 2004;2(3):252-61.
8. Lee SK. [Hereditary pancreatitis]. *The Korean journal of gastroenterology = Taehan Sohwagi Hakhoe chi*. 2005;46(5):358-67.
9. Raphael KL, Willingham FF. Hereditary pancreatitis: current perspectives. *Clinical and experimental gastroenterology*. 2016;9:197-207.
10. Masamune A, Kikuta K, Hamada S, Nakano E, Kume K, Inui A, et al. Nationwide survey of hereditary pancreatitis in Japan. *Journal of gastroenterology*. 2018;53(1):152-60.
11. Mora J, Comas L, Ripoll E, Goncalves P, Queralto JM, Gonzalez-Sastre F, et al. Genetic mutations in a Spanish population with chronic pancreatitis. *Pancreatology : official journal of the International Association of Pancreatology (IAP) [et al]*. 2009;9(5):644-51.
12. Raty S, Piironen A, Babu M, Pelli H, Sand J, Uotila S, et al. Screening for human cationic trypsinogen (PRSS1) and trypsinogen inhibitor gene (SPINK1) mutations in a Finnish family with hereditary pancreatitis. *Scandinavian journal of gastroenterology*. 2007;42(8):1000-5.
13. Sibert JR. Hereditary pancreatitis in England and Wales. *Journal of medical genetics*. 1978;15(3):189-201.
14. Goebell H, Ammann R, Creutzfeldt W. History of the European pancreatic club: the first 40 years 1965-2005. The development of the European pancreatic club as a scientific society. *Pancreatology : official journal of the International Association of Pancreatology (IAP) [et al]*. 2005;5 Suppl 1:1-15.
15. Applebaum-Shapiro SE, Finch R, Pfitzer RH, Hepp LA, Gates L, Amann S, et al. Hereditary pancreatitis in North America: the Pittsburgh-Midwest Multi-Center Pancreatic Study Group Study. *Pancreatology : official journal of the International Association of Pancreatology (IAP) [et al]*. 2001;1(5):439-43.
16. Joergensen MT, Brusgaard K, Cruger DG, Gerdes AM, Schaffalitzky de Muckadell OB. Genetic, epidemiological, and clinical aspects of hereditary pancreatitis: a population-based cohort study in Denmark. *The American journal of gastroenterology*. 2010;105(8):1876-83.
17. Atlas AB, Orenstein SR, Orenstein DM. Pancreatitis in young children with cystic fibrosis. *The Journal of pediatrics*. 1992;120(5):756-9.
18. Masamune A. Genetics of pancreatitis: the 2014 update. *The Tohoku journal of experimental medicine*. 2014;232(2):69-77.
19. Langner C. [Hereditary gastric and pancreatic cancer]. *Der Pathologe*. 2017;38(3):164-9.

20. Lowenfels AB, Maisonneuve P, DiMagno EP, Elitsur Y, Gates LK, Jr., Perrault J, et al. Hereditary pancreatitis and the risk of pancreatic cancer. International Hereditary Pancreatitis Study Group. *Journal of the National Cancer Institute*. 1997;89(6):442-6.
21. Malka D, Hammel P, Maire F, Rufat P, Madeira I, Pessione F, et al. Risk of pancreatic adenocarcinoma in chronic pancreatitis. *Gut*. 2002;51(6):849-52.
22. Weiss FU. Pancreatic cancer risk in hereditary pancreatitis. *Frontiers in physiology*. 2014;5:70.
23. Lowenfels AB, Maisonneuve P, Whitcomb DC, Lerch MM, DiMagno EP. Cigarette smoking as a risk factor for pancreatic cancer in patients with hereditary pancreatitis. *Jama*. 2001;286(2):169-70.
24. Lowenfels AB, Maisonneuve P, Whitcomb DC. Risk factors for cancer in hereditary pancreatitis. International Hereditary Pancreatitis Study Group. *The Medical clinics of North America*. 2000;84(3):565-73.
25. Mastoraki A, Tzortzopoulou A, Tsela S, Danias N, Sakorafas G, Smyrniotis V, et al. Hereditary pancreatitis: dilemmas in differential diagnosis and therapeutic approach. *Journal of gastrointestinal cancer*. 2014;45(1):22-6.
26. Patel MR, Eppolito AL, Willingham FF. Hereditary pancreatitis for the endoscopist. *Therapeutic advances in gastroenterology*. 2013;6(2):169-79.
27. Kargl S, Kienbauer M, Duba HC, Schofl R, Pumberger W. Therapeutic step-up strategy for management of hereditary pancreatitis in children. *Journal of pediatric surgery*. 2015;50(4):511-4.
28. LaRusch J, Solomon S, Whitcomb DC. Pancreatitis Overview. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. *GeneReviews*((R)). Seattle (WA): University of Washington, Seattle

University of Washington, Seattle. *GeneReviews* is a registered trademark of the University of Washington, Seattle. All rights reserved.; 1993.

29. Rivera Rivera ED. [Pancreatitis, genes and islet cells auto transplant; updates and new horizons]. *Revista de gastroenterologia del Peru : organo oficial de la Sociedad de Gastroenterologia del Peru*. 2017;37(2):156-61.
30. Solomon S, Whitcomb DC, LaRusch J. PRSS1-Related Hereditary Pancreatitis. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. *GeneReviews*((R)). Seattle (WA): University of Washington, Seattle

University of Washington, Seattle. *GeneReviews* is a registered trademark of the University of Washington, Seattle. All rights reserved.; 1993.

31. LaRusch J, Barmada MM, Solomon S, Whitcomb DC. Whole exome sequencing identifies multiple, complex etiologies in an idiopathic hereditary pancreatitis kindred. *JOP : Journal of the pancreas*. 2012;13(3):258-62.
32. Kim JY, Choi SH, Ihm JS, Kim SJ, Kim IJ, Kim CM. [A case of R122H mutation of cationic trypsinogen gene in a pediatric patient with hereditary pancreatitis complicated by pseudocyst and hemosuccus pancreaticus]. *The Korean journal of gastroenterology = Taehan Sohwagi Hakhoe chi*. 2005;45(2):130-6.
33. Jamer T, Iwanczak B. [Genetic mutations as a cause of acute recurrent pancreatitis in children - case report and literature review]. *Developmental period medicine*. 2016;20(3):228-34.
34. Lal A, Lal DR. Hereditary pancreatitis. *Pediatric surgery international*. 2010;26(12):1193-9.
35. Le Marechal C, Masson E, Chen JM, Morel F, Ruszniewski P, Levy P, et al. Hereditary pancreatitis caused by triplication of the trypsinogen locus. *Nature genetics*. 2006;38(12):1372-4.
36. Simon P, Weiss FU, Sahin-Toth M, Parry M, Nayler O, Lenfers B, et al. Hereditary pancreatitis caused by a novel PRSS1 mutation (Arg-122 --> Cys) that alters autoactivation and autodegradation of cationic trypsinogen. *The Journal of biological chemistry*. 2002;277(7):5404-10.

37. Nemeth BC, Sahin-Toth M. Human cationic trypsinogen (PRSS1) variants and chronic pancreatitis. *American journal of physiology Gastrointestinal and liver physiology*. 2014;306(6):G466-73.
38. Chen JM, Le Marechal C, Lucas D, Raguenes O, Ferec C. "Loss of function" mutations in the cationic trypsinogen gene (PRSS1) may act as a protective factor against pancreatitis. *Molecular genetics and metabolism*. 2003;79(1):67-70.
39. Whitcomb DC, Gorry MC, Preston RA, Furey W, Sossenheimer MJ, Ulrich CD, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nature genetics*. 1996;14(2):141-5.
40. Lee TY, Oh HC, Kim MH, Kwon S, Lee SS, Seo DW, et al. [Three cases of hereditary pancreatitis in two households in the same family associated with R122H mutation in cationic trypsinogen gene]. *The Korean journal of gastroenterology = Taehan Sohwagi Hakhoe chi*. 2007;49(6):395-9.
41. Hanck C, Schneider A, Whitcomb DC. Genetic polymorphisms in alcoholic pancreatitis. *Best practice & research Clinical gastroenterology*. 2003;17(4):613-23.
42. Chen X, Chang JT. Planning bioinformatics workflows using an expert system. *Bioinformatics (Oxford, England)*. 2017;33(8):1210-5.
43. Bajard A, Chabaud S, Cornu C, Castellan AC, Malik S, Kurbatova P, et al. An in silico approach helped to identify the best experimental design, population, and outcome for future randomized clinical trials. *Journal of clinical epidemiology*. 2016;69:125-36.
44. Mustafa MI, Abdelhameed T, Abdelrhman F, Osman S, Hassan M. Novel deleterious nsSNPs within MEFV gene that could be used as Diagnostic Markers to Predict Hereditary Familial Mediterranean Fever: Using bioinformatics analysis. *bioRxiv*. 2018:424796.
45. Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic acids research*. 2012;40(Database issue):D48-53.
46. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al. ExPASy: SIB bioinformatics resource portal. *Nucleic acids research*. 2012;40(Web Server issue):W597-603.
47. Sim NL, 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(Web Server issue):W452-7.
48. Capriotti E, Altman RB. Improving the prediction of disease-related variants using protein three-dimensional structure. *BMC bioinformatics*. 2011;12 Suppl 4:S3.
49. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PloS one*. 2012;7(10):e46688.
50. Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants. *BMC genomics*. 2015;16 Suppl 8:S1.
51. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Human mutation*. 2009;30(8):1237-44.
52. Lopez-Ferrando V, Gazzo A, de la Cruz X, Orozco M, Gelpi JL. PMut: a web-based tool for the annotation of pathological variants on proteins, 2017 update. *Nucleic acids research*. 2017;45(W1):W222-w8.
53. 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(Web Server issue):W306-10.
54. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins*. 2006;62(4):1125-32.
55. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic acids research*. 2010;38(Web Server issue):W529-33.
56. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research*. 2016;44(W1):W344-50.

57. Oni OO, Owoade AA, Adeyefa CAO. Design and evaluation of primer pairs for efficient detection of avian rotavirus. *Tropical animal health and production*. 2018;50(2):267-73.
58. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic acids research*. 2010;38(Web Server issue):W214-20.
59. Meyer MJ, Lapcevic R, Romero AE, Yoon M, Das J, Beltran JF, et al. mutation3D: Cancer Gene Prediction Through Atomic Clustering of Coding Variants in the Structural Proteome. *Human mutation*. 2016;37(5):447-56.
60. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*. 2004;25(13):1605-12.
61. George Priya Doss C, Sudandiradoss C, Rajasekaran R, Choudhury P, Sinha P, Hota P, et al. Applications of computational algorithm tools to identify functional SNPs. *Functional & integrative genomics*. 2008;8(4):309-16.
62. Chandrasekaran G, Hwang EC, Kang TW, Kwon DD, Park K, Lee JJ, et al. In silico analysis of the deleterious nsSNPs (missense) in the homeobox domain of human HOXB13 gene responsible for hereditary prostate cancer. 2017;90(2):188-99.
63. Kosaloglu Z, Bitzer J, Halama N, Huang Z, Zapatka M, Schneeweiss A, et al. In silico SNP analysis of the breast cancer antigen NY-BR-1. *BMC cancer*. 2016;16(1):901.
64. Oracz G, Kolodziejczyk E, Sobczynska-Tomaszewska A, Wejnarska K, Dadalski M, Grabarczyk AM, et al. The clinical course of hereditary pancreatitis in children - A comprehensive analysis of 41 cases. *Pancreatology : official journal of the International Association of Pancreatology (IAP)* [et al]. 2016;16(4):535-41.