Indian Journal of Biochemistry & Biophysics
Vol. 61, May 2024, pp. 279-289
DOI: 10.56042/ijbb.v61i5.6227

Computational strategies for identifying high-risk SNP of PTEN in prostate cancer: A Mutational profiling study

Vharshini Sivaraj, Deetchika Sekar, Zoubiya Afshan Ansar, Keerthana Karunakaran & Rajiniraja Muniyan*

School of BioSciences and Technology, Vellore Institute of Technology, Vellore-632 014, Tamil Nadu, India

Received 15 October 2023; *revised 12 February 2024*

The fifth most prevalent cause of cancer-related mortality and the most often diagnosed malignancy in males is prostate cancer. The PTEN tumor suppressor gene is one of the most often altered genomes in prostate cancer. PTEN is a potentially helpful genetic marker to discriminate between indolent and aggressive illness in individuals with clinically localised tumors because loss of PTEN function activates the PI3K-AKT pathway and is highly related with unfavourable oncological outcomes. Moreover, research has suggested that PTEN inactivation due to deletion or mutation influence tumor formation by modifying the immune system and the tumor microenvironment, in addition to its known roles in the PI3K/AKT pathway. Hence, the study aimed to screen all PTEN associated SNPs of prostate cancer for analyzing its structural and functional impact through various computational tools. The results showed C124S and G129R as the most pathogenic variant and are highly conserved causing protein destabilization in prostate cancer. Further, structural analysis through molecular simulation showed that the mutant G129R caused huge instability, high residue fluctuation and loss of compactness in the PTEN protein. The study's findings shed light on the structural and functional consequences of specific PTEN mutations in prostate cancer.

Keywords: Biomarker, GROMACS, Missense mutation, Molecular dynamics simulation, Tumor suppressor

The second leading cause of mortality around the globe is cancer and the prevalence varies between males and females. Prostate, colon, lungs and bronchus cancer are common in the former and breast, thyroid, rectum, uterine corpus is common among females[1]. Prostate cancer is the fifth leading cause of cancer related death and is one of the deadliest types of cancer. In 2020, there were 1,414,259 new cases and mortality of 375,304 was found and is diagnosed with MRI scans, prostate specific antigens test (PSA) and tissue biopsy[2]. Androgen-deprivation therapy, surgery, and a combination of chemotherapy and radiotherapy are the current treatments for prostate cancer. The survival rate of patients had been extended in chemotherapy in comparison to androgen deprivation therapy. Androgen deprivation therapy (ADT) is the primary systemic treatment prescribed for advanced prostate cancer[3]. However, the depth and length of response to ADT are quite different, and most people eventually stop responding to it and develop metastatic castration-resistant prostate cancer (mCRPC)[4]. Radiotherapy prescribed for the localized cancer caused post treatment threats like sexual dysfunction that can affect the quality of life.

Prostate cancer is a heterogenous disease which involves several factors like environmental and genetic variations. Pathogenic variants such as BRCA1, BRCA2, HOXB13 and various mismatch repair genes are found to be the major driver for prostate carcinogenesis[5]. Comprehensive genomic analysis of prostate cancer has found frequent alterations in androgen signaling, DNA repair, and PI3K pathway like PTEN[6]. Recurrent somatic mutations, copy number alterations, and carcinogenic structural DNA rearrangements (chromosomal abnormalities) in PTEN, have been discovered in primary prostate cancer, metastatic prostate cancer, and metastatic, castration-resistant prostate cancer (mCRPC)[7]. The proliferation and aggressive development of prostate cancer cells are fuelled by metabolic reprogramming brought on by PTEN loss[8]. In about 9% of all tumors, PTEN is either deleted or somatically altered[9]. Mutation in PTEN has been found in ~20% of prostate tumors and in 50 % of castration resistant prostate tumors. About <5% of initial stage of prostate cancer and 60%-70% metastatic cancers typically include missense mutation in PTEN[10].

*Correspondence:
E-mail: rajiniraja.m@vit.ac.in

PTEN mutation leads to the constituent activation of P13K-AKT pathway, which regulates proliferation, migration, and survival of cell. It serves as a genomic biomarker in clinically localized tumors[11]. PTEN loss is associated with a poor prognosis, including a restricted progression-free survival time, a higher chance of relapse, and metastasis[12]. PTEN mutation leads to the growth of hamartomas which are non-cancerous tumors. A hamartoma is usually a harmless mass of jumbled tissue that grows only in a particular part of body such as lungs, breast, colon, hypothalamus, and prostate[1]. PTEN mutations frequently co-occur with mutations in other CRC-related genes, including as KRAS, APC, TP53, and PIK3CA, in non-identical ways[13].

A mutation named non-synonymous SNP (nsSNP) results in a single amino acid change that affect the protein's structure and consequently affecting its function. While certain nsSNPs have an impact on structural characteristics, others have functional repercussions. Furthermore, while some nsSNPs can be related to a medical condition, others might not be related to any pathological phenotype and are therefore regarded as neutral[14-16]. The prediction of useful nsSNPs for PTEN gene by *in silico* tools would be highly beneficial owing to its difficulty in experimental prediction. Hence, the goal of the current study was to forecast SNPs that may have an impact on PTEN stability and function in prostate cancer and the complete workflow is given in (Fig. 1).

Materials and Methods

Dataset collection

The protein sequence of human PTEN was retrieved from Uniprot (Uniprot ID - P60484). The list of mutations reported for human PTEN in prostate cancer was collected from COSMIC database (catalogue of somatic mutations in cancer) https://cancer.sanger.ac.uk/cosmic from the following entries PTEN, TPTE_ENST00000427445, TPTE, TPTE_ENST00000622113, TPTE_ENST00000612746. We collected 62 missense mutations associated with prostate cancer from cosmic server. COSMIC server was utilized since it has large collection of datasets from both large-scale sequencing projects and individual research studies

Pathogenicity analysis

Meta-SNP (https://snps.biofold.org/meta-snp/) was used to predict the pathogenicity of the missense mutations obtained from COSMIC server. It utilizes five servers for prediction. PhD-SNP uses sequence information to estimate disease associated with nsSNVs. SIFT makes predictions regarding nsSNVs functional effects from evolutionary information's. The biochemical features of substitutions along with predicted structural proteins and functional characteristics to classify neutral and non-neutral variants are performed by a neutral network-based method called SNAP. The effect of nsSNVs is predicted by PANTHER[17]. Mutation predicted as disease variant in all the five tools were taken for further studies.

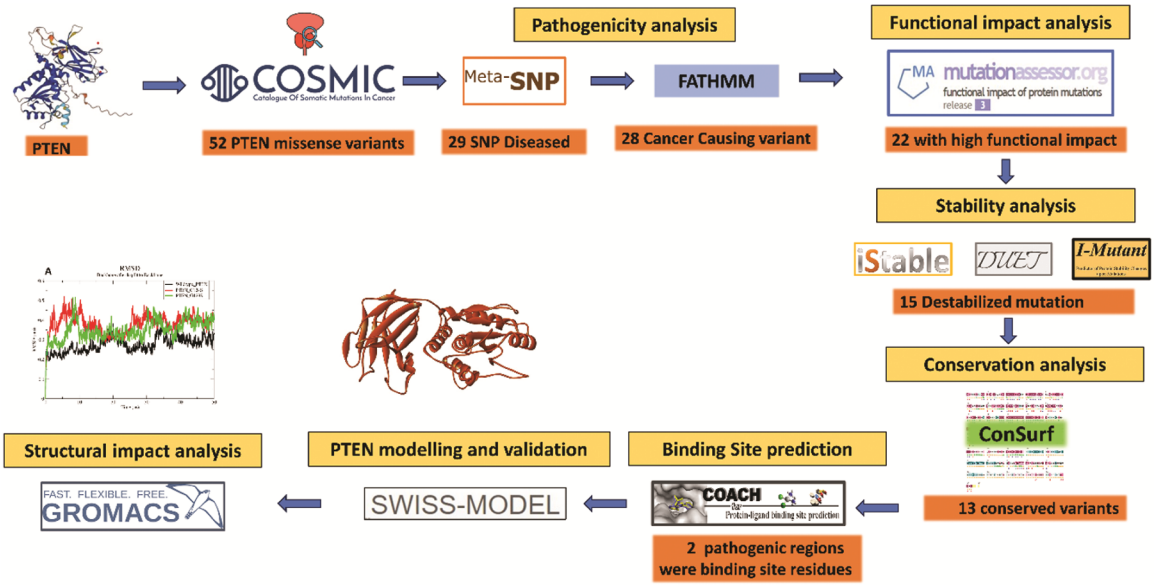


Fig. 1 — Schematic diagram of computational prediction of PTEN mutation and stability analysis in prostate cancer

FATHMM (functional analysis through hidden Markov models) is a high-throughput web server that can estimate the functional effects of both coding variants like single nucleotide variants (nsSNVs), non-coding variants and non-synonymous in the human genome. It also utilizes hidden Markov models algorithm for predicting the mutational impact on structural phenotype of protein (https://fathmm.biocompute.org.uk/). The resolution for predicting the cancer-causing mutations were set as -3.0 since the false positive results is eliminated. The Uniprot ID of PTEN along with its amino acid mutation is submitted as input. The mutations predicted to cause cancer are selected for further functional annotation.

Functional impact analysis

To classify high, medium, low, and neutral we used mutation server assessor (http://mutationassessor.org/r3/). This server estimates the functional influence of missense mutation. This server also depends on multiple sequence alignment and residues of amino acids that are conserved evolutionarily.

The list of mutation predicted as disease variant and Uniprot protein accession ID (P60484) is submitted as input.

Stability analysis

To determine the stability and functional consequences of the mutation we used different kinds of servers like MuPro (https://mupro.proteomics.ics.uci.edu/), iStable, I-mutant 2.0 (http://predictor.nchu.edu.tw/istable/indexSeq.php) and mCSM, SDM, DUET (http://marid.bioc.cam.ac.uk/sdm2/prediction). MuPro assess the protein stability based on primary sequence data distinctly where the tertiary structures are not revealed. I-mutant2.0 is an online web server that recklessly detects the changes in stability of protein in point mutations. It also assesses severity effect of mutations on the stability of folded protein. Using the information from sequence and data predictions, iStable was constructed. Sequence and structure are the two different types of input required for iStable. mCSM predicts the impact of missense mutations on the stability of the protein, protein-DNA interaction, and protein-protein binding using information from sequence and data predictions. SDM is a statistical method that studies amino acid substitutions in different structural conditions and generates probability tables for substitutions. DUET is an online web server that estimates the impact of point mutations on protein stability and relies on additional mutation information, including secondary structure and pharmacophore vector. For MuPro, the mutation name, position, original amino acid, and substitute amino acid for each mutation, along with protein structure, were submitted as input for MuPro, the mutation name, position, original amino acid, and substitute amino acid for each mutation, along with protein structure, were submitted as input. For iStable and SDM, amino acid sequence of PTEN wild-type, mutant residue and their position were submitted as input.

Conservation analysis

Further, consurf webserver (https://consurf.tau.ac.il/consurf_index.php) was utilized to analyse the SNP's occurring at the conserved sites. To run consurf, we specified a PDB ID, the gene mutation and the chain identifier as the query.

Binding site prediction

COACH D is a consensus algorithm that is developed to predict protein ligand binding sites[18].

Structure prediction and validation

The crystal structure of PTEN had missing residues in the region (281-311). Hence the structure of both wild PTEN and mutant PTEN was predicted with Swiss model. Further, the predicted models were evaluated using SAVES server (https://saves.mbi.ucla.edu/) for Ramachandran plot and stereochemical quality.

Molecular dynamics simulation

Wild type PTEN and mutants were subjected to simulation in Webgro to analyse structural and functional impact of mutants[19]. It is an automated tool which uses Gromacs for molecular dynamics simulation. Protein topology was created by Gromos96 54a7. The protein was placed in the cubic box of 1 cm and solvated with SPC (simple point charge). The simulation was set in the presence of 0.15 M NaCl with 1.0 bar pressure and 300 K constant temperature. Energy minimization was done for 50,000 steps and temperature, pressure was stabilized by NVT and NPT (300 K and 1 bar pressure). RMSD (Root mean square deviation), RMSF (Root mean square fluctuation), SASA (Solvent accessible surface area), Rg (Radius of gyration) and hydrogen bond interaction were analysed by Xmgrace after 50 ns.

Results

Data collection

Due to the fact that PTEN is linked to cancer, the analysis of mutational impact in PTEN structure and

function are of significant clinical importance. Therefore, all missense mutations of PTEN were collected from COSMIC server. 52 mutations were analyzed with different *in silico* tools to find its impact on PTEN protein. The prediction analysis of PTEN has represented as graph in (Fig. 2). Pathogenicity analysis revealed that 29 mutations are disease and damaging variant since they were predicted to cause disease in all the five tools. Mutations predicted as disease variants by all five tools were selected for further analysis since the ensemble strategy leads to more robust and accurate predictions (Table 1). Further, FATHMM server predicted 28 missense variants as cancer causing and one variant (G129W) as passenger (Table 2). Fathmm was used to obtain the cancer causing/driver mutation since it includes features for estimating the impact of mutation in the context of cancer.

The functional impact of missense variants was assessed by mutation assessor server which works

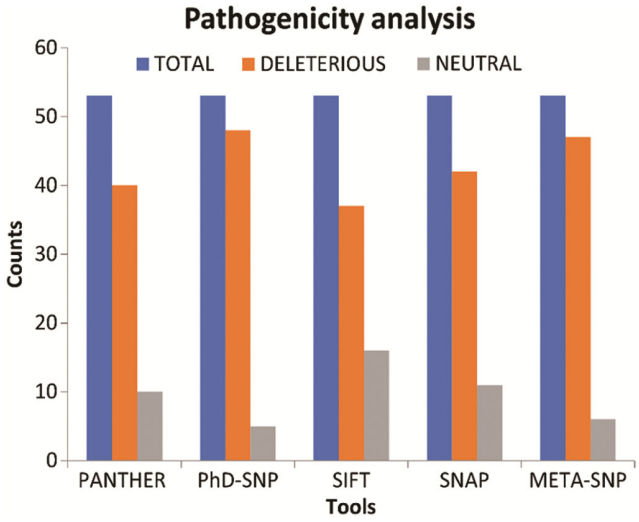


Fig. 2 — Deleterious variant: Variants predicted by Meta-SNP to be associated with disease phenotypes. Neutral variant: Variants predicted by Meta-SNP to have no significant association with disease phenotypes

Table 1 — List of missense mutations predicted from Meta-SNP shows disease or neutral variants

S. No	Mutation	PANTHER	PhD-SNP	SIFT	SNAP	Meta-SNP
1	I5T	NA	Neutral	Neutral	Disease	Neutral
2	K6E	NA	Disease	Disease	Disease	Disease
3	K13E	NA	Disease	Disease	Disease	Disease
4	D24N	Disease	Disease	Disease	Disease	Disease
5	Y27N	Disease	Disease	Disease	Disease	Disease
6	I28T	Neutral	Disease	Disease	Disease	Disease
7	M35V	Disease	Disease	Neutral	Disease	Disease
8	G36R	Disease	Disease	Disease	Disease	Disease
9	P38T	Disease	Disease	Disease	Disease	Disease
10	R47G	Disease	Disease	Disease	Disease	Disease
11	N48K	Disease	Disease	Disease	Disease	Disease
12	R55G	Neutral	Disease	Disease	Disease	Disease
13	Y68C	Disease	Disease	Disease	Disease	Disease
14	Y68H	Neutral	Disease	Disease	Disease	Disease
15	F81C	Disease	Disease	Disease	Disease	Disease
16	D92E	Disease	Disease	Disease	Disease	Disease
17	D92Y	Disease	Disease	Disease	Disease	Disease
18	P95S	Disease	Neutral	Neutral	Neutral	Neutral
19	P96R	Disease	Disease	Disease	Disease	Disease
20	I101N	Disease	Disease	Disease	Neutral	Disease
21	V119F	Disease	Disease	Disease	Disease	Disease
22	C124S	Disease	Disease	Disease	Disease	Disease
23	C124Y	Disease	Disease	Disease	Disease	Disease
24	K128N	Neutral	Disease	Disease	Disease	Disease
25	G129R	Disease	Disease	Disease	Disease	Disease
26	R130Q	Disease	Disease	Disease	Disease	Disease
27	T131I	Disease	Disease	Disease	Disease	Disease
28	T131N	Disease	Disease	Disease	Disease	Disease
29	G132V	Disease	Disease	Disease	Disease	Disease
30	M134I	Disease	Neutral	Neutral	Neutral	Disease
31	M134L	Neutral	Disease	Neutral	Neutral	Neutral
32	M134R	Disease	Disease	Disease	Disease	Disease
33	C136R	Neutral	Disease	Disease	Disease	Disease
34	E150G	Disease	Disease	Disease	Disease	Disease

(*Contd.*)

Table 1 — List of missense mutations predicted from Meta-SNP shows disease or neutral variants (*Contd*.)

S. No	Mutation	PANTHER	PhD-SNP	SIFT	SNAP	Meta-SNP
35	V166G	Disease	Disease	Disease	Disease	Disease
36	Q171R	Disease	Disease	Disease	Disease	Disease
37	R173C	Disease	Disease	Disease	Disease	Disease
38	R173H	Disease	Disease	Disease	Disease	Disease
39	Y174D	Disease	Disease	Disease	Disease	Disease
40	Y174N	Disease	Disease	Disease	Disease	Disease
41	Y176C	Disease	Disease	Neutral	Neutral	Neutral
42	F200C	Disease	Disease	Neutral	Neutral	Disease
43	P213R	Disease	Disease	Neutral	Disease	Disease
44	V217D	Disease	Disease	Neutral	Disease	Disease
45	F241S	Disease	Disease	Neutral	Neutral	Disease
46	I253N	Disease	Disease	Neutral	Disease	Disease
47	V255A	Neutral	Neutral	Neutral	Disease	Neutral
48	D326G	Disease	Disease	Disease	Disease	Disease
49	F341V	Neutral	Disease	Disease	Disease	Disease
50	K344R	Neutral	Disease	Neutral	Neutral	Disease
51	V369G	Disease	Disease	Neutral	Neutral	Neutral
52	T382S	Neutral	Neutral	Disease	Neutral	Neutral

Table 2 — Prediction of cancer causing PTEN mutants using FATHMM server shows the medium and high functional impacts mutants

S. NO	Mutation	FATHMM	Functional impact
1	D24N	Cancer	high
2	Y27N	Cancer	high
3	G36R	Cancer	high
4	P38T	Cancer	high
5	R47G	Cancer	high
6	N48K	Cancer	Medium
7	Y68C	Cancer	Medium
8	F81C	Cancer	high
9	D92E	Cancer	Medium
10	D92Y	Cancer	high
11	P96R	Cancer	high
12	V119F	Cancer	high
13	C124S	Cancer	high
14	C124Y	Cancer	high
15	G129R	Cancer	high
16	R130Q	Cancer	high
17	T131I	Cancer	Medium
18	T131N	Cancer	high
19	G132V	Cancer	high
20	M134R	Cancer	high
21	E150G	Cancer	Medium
22	V166G	Cancer	high
23	Q171R	Cancer	high
24	R173C	Cancer	high
25	R173H	Cancer	high
26	Y174D	Cancer	high
27	Y174N	Cancer	high
28	D326G	Cancer	Medium

based on the evolutionary conservation of mutated amino acid in the homologs of protein. Mutation with high functional impact is grouped a high, less functional impact as medium and non-functional impact is classified as low. In our study, among the 28 mutations predicted as disease causing, 22 missense mutations had high functional impact (Table 3).

Stability analysis

The stability of proteins is crucial for their optimal functioning. To enhance prediction accuracy, various algorithms such as I-Mutant 2.0, istable, mCSM, SDM, DUET, and MuPro were used to identify missense mutations that may destabilize the protein. Mutations that lead to a decrease in Gibbs free energy

Table 3 — Stability analysis of PTEN missense mutants shows destabilized mutants

S. No	Mutation	INSPD	Imutant	Istable	mCSM	SDM	DUET	MuPro
1	Y27N	-2.505805	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
2	I28T	-3.156435	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
3	G36R	-0.781273	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
4	P96R	-0.597214	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
5	I101N	-3.20988	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
6	V119F	-2.000435	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
7	C124S	-2.684075	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
8	G129R	-0.876976	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
9	M134R	-1.342843	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
10	C136R	-1.99884	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
11	V166G	-3.442145	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
12	R173C	-0.650608	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
13	R173H	-1.31921	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
14	Y174D	-2.445645	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease
15	Y174N	-2.047785	Decrease	Decrease	Destabilizing	Destabilizing	Destabilizing	Decrease

of less than 2 kcal mol^{-1} were considered to cause protein malfunction, and those occurring at the binding site or active site were predicted to cause disease. The stability of the protein is measured by the Gibbs free energy in the above-mentioned algorithms. Sequence based algorithms namely I-mutant 2.0, istable and MuPro, predicted 15 missense mutations with $\Delta\Delta G < 0$ as the protein destabilizers (Table 3).

Conservation analysis

Consurf, a conservation analysis tool was utilized to find the conservation status of the destabilizing mutations. Among the 15 missense variants all the variants excluding Y27, I101 were predicted to be highly conserved since their conservation score was 9. The 14 variants were highly conserved and are in buried and exposed position which is denoted by s and f in (Fig. 3). All these results indicate that the amino acid position predicted as pathogenic and cause destabilization could affect the function of protein.

Binding site prediction

The estimation of ligand binding site plays an important role in protein regulation. Therefore, PTEN was subjected to predict the binding site using the COACH-D server. Among the 10 residues of binding pocket, two amino acid position C124S and G129R were predicted to cause disease, highly conserved and destabilize the protein in the above analysis of our study. Hence, from our study we describe C124S and G129R as the pathogenic, deleterious and destabilizing variant of PTEN in prostate cancer (Table 4). Further, the conformational changes in presence of mutations were analyzed through molecular dynamics simulation.

Structure prediction

The crystal structure of PTEN available in PDB lacked the middle part of structure ranging from the aminoacids 280-311. Hence, the structure was modelled in swissmodel using 7JVX as a template. The template had the query coverage and identity of 80% and 99.75%. Wild type, C124S and G129R modelled was further evaluated for its Ramachandran plot. All the models had a favoured regions of 85.2%, additionally allowed regions of 12% and disallowed regions of 1% (Fig. 4).

Molecular Dynamics Simulation

The modelled structures were minimized in Gromacs and further the simulation was set up for 50ns. The simulation parameters such as RMSD, RMSF, SASA, Rg were employed in this work to compare the stability and behaviour of molecules in wild and mutated type. Low and stable RMSD over the simulation in general depicts greater stability. The average RMSD values for 1D5R, C124S and G129R are 0.46 nm, 0.42 nm and 0.47 nm. The RMSD graph of wildtype PTEN remained to be in the straight line but minor fluctuations were seen throughout the simulation process. The relatively stable RMSD were seen at initial 20 ns. C124S system showed stable RMSD for 10 ns to 30 ns and thereafter the value increased and followed the straight trendline (Fig. 5A). However, the average RMSD of C124S was significantly low in comparison to the mutant. Mutant G129R system had gradual increase in RMSD for 0 ns to 20 ns and thereafter the RMSD value decreased from 0.52 nm to 0.47 nm at 25 ns. For 25 to 30 ns the RMSD value remained as 0.47 nm and thereafter RMSD value again raised high to 0.57 nm

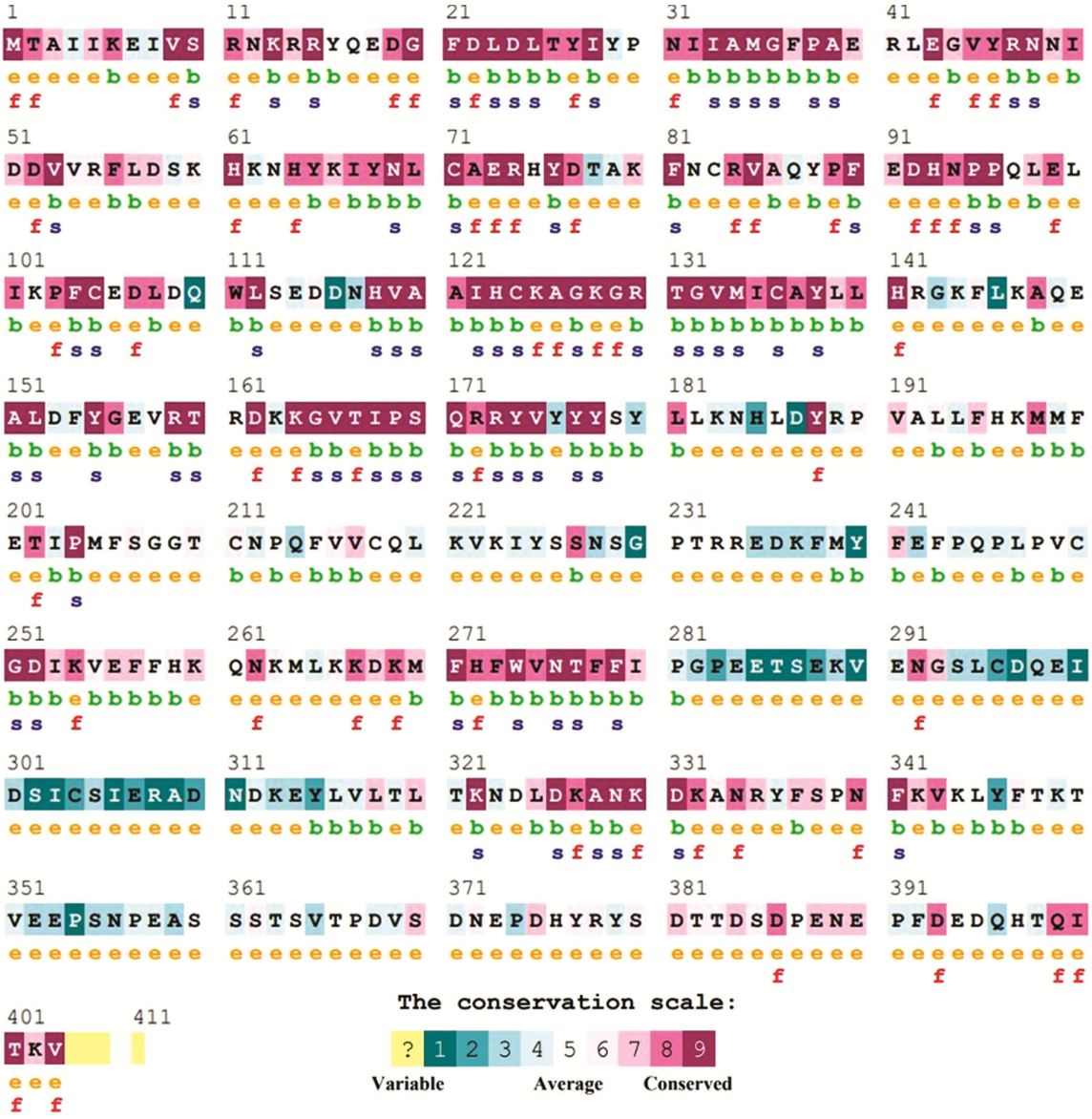


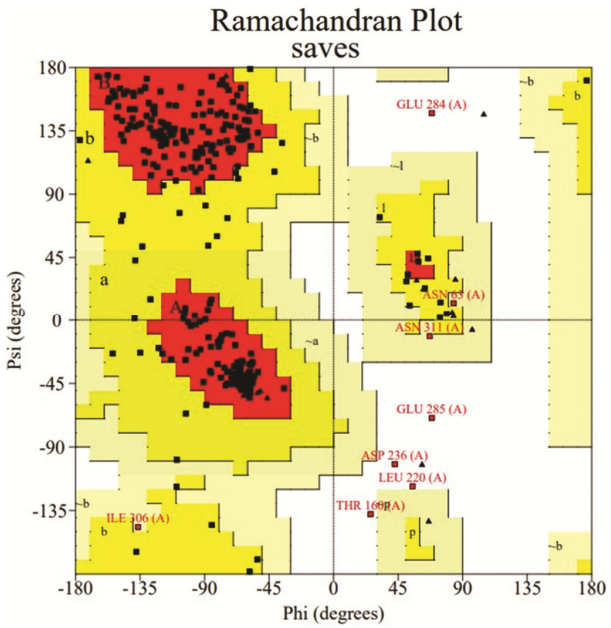
Fig. 3 — Conservation analysis of PTEN in Consurf. The conservation scale shows 15 variants (14 s and 1 f) were highly conserved and buried region, respectively

Table 4 — Binding site prediction of PTEN protein (PDB ID: 1D5R)

Rank	C-score	Cluster size	Energy	Predicted binding residue number
1	0.53	41	-3.0	92,93,124,125,126,127,128,129,130,168
2	0.03	3	-5.0	134,138,141,142
3	0.02	2	-1.9	42,47,74,125
4	0.02	2	0.2	70,88,90,96,98,101,104,134,137,174,177
5	0.01	1	-2.8	90,124,130,171

showing the instability of the mutant. The major deviation in RMSD graph of mutants gives us an insight that G129R mutation in the active site led to different confirmation changes resulting in instability.

The average C-alpha atom mobility is studied by RMSF to assess the flexibility of proteins. Plots of residue mobility against the amino acid residues are shown for wild and mutant types. The average RMSF value for native PTEN protein was 0.22 nm and the value ranged to 0.8 nm in the loop regions since loop regions exhibit more fluctuations normally. C124S mutant system showed high fluctuations including in the regions of helices and sheets in comparison to native. The average RMSF produced was 0.22 nm. In



Plot statistics

Residues in most favoured regions [A,B,L]	265	85.2%
Residues in additional allowed regions [a,b,l,p]	38	12.2%
Residues in generously allowed regions [~a,~b,~l,~p]	3	1.0%
Residues in disallowed regions	5	1.6%
Number of non-glycine and non-proline residues	311	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	15	
Number of proline residues	17	
Total number of residues	345	

Fig. 4 — Ramachandran plot of modeled PTEN PROCHECK shows that 85.2% (red) are in the most favored region. Yellow areas indicate additionally allowed region (12.2%) and the white region represent disallowed region (0.5%)

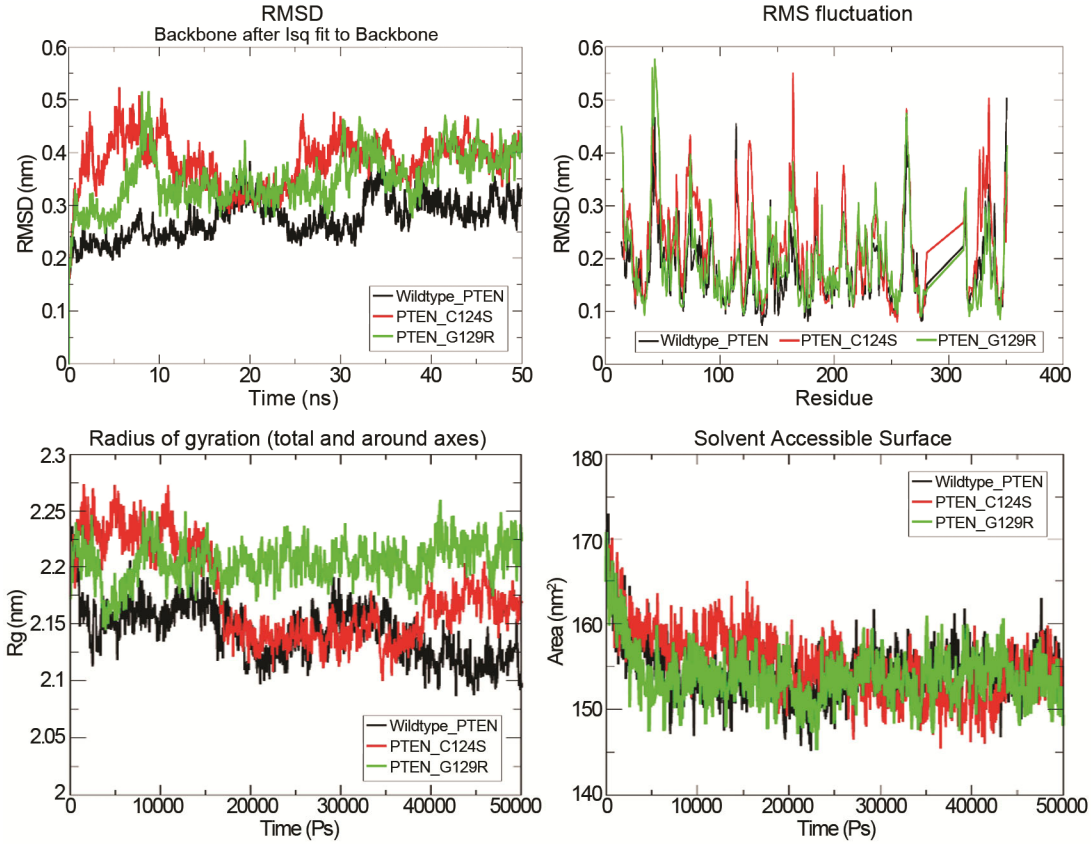


Fig. 5 — Stability analysis of wildtype PTEN and mutants. (A) RMSD of wildtype PTEN (black) mutant C124S (red) and mutant G129R (green); (B) RMSF of wildtype PTEN (black) mutant C124S (red) and mutant G129R (green); (C) Rg of wildtype PTEN (black) mutant C124S (red) and mutant G129R (green); and (D) SASA of wildtype PTEN (black) mutant C124S (red) and mutant G129R (green)

G129R system, major fluctuations were seen in all the residues and produced an average RMSF of 0.25 nm. This implies that mutant C124S system exhibited minor conformational change whereas major conformational change was seen in G129R resulting in the system's instability (Fig. 5B). The significant difference in the average RMSF and in the graph peaks between native and PTEN_G129R supports the notion of instability in mutant models and demonstrates the impact of SNP in the binding site of PTEN. This is in accordance with the RMSD.

The radius of gyration reveals the proteins compactness and determines protein's folding and rigidity. A lower Rg value denotes stable folding, whereas an elevated Rg value with significant fluctuation denotes unstable folding and less compactness. Native PTEN displayed an Rg of 2.23 nm throughout the simulation process and produced an average Rg of 2.23 nm. Mutant C124S showed a low Rg value of 2.21nm for 10 ns – 30 ns and thereafter the value raised to 2.24 nm indicating the minor loss of compactness in the mutant's tertiary structure (Fig. 5C). The average Rg produced was equal as native PTEN but the graph showed minor fluctuation and is in accordance with RMSD plot. The G129R system produced an average of 2.25 nm showing high Rg value from the point of simulation to 30 ns. Followed by, the Rg value decreased gradually to 2.18 nm till 40 ns and thereafter the Rg value raised to 2.25 nm at the end of 50 ns. The increase and decrease in Rg throughout the simulation process implies that SNP G129R resulted in the loss of compactness in the tertiary protein structure and low structural stability. Mutants graph plot is in accordance with RMSD and Rg results.

SASA (Solvent Accessible Surface area) is a measurement of a protein's surface area that is accessible to water molecules and the changes that occur in the protein as a result of ligand binding or mutation. Lower SASA values imply strong compactness of the complex, while higher SASA values suggest poor compactness and high flexibility. The average SASA value of 1D5R, G129R and C124S are 173.61 nm^2, 173.71 nm^2 and 172.49 nm^2 (Fig. 5D).

Discussion

In silico mutational analysis plays a pivotal role in advancing our understanding of disease mechanisms, guiding personalized medicine approaches, and accelerating drug discovery and development efforts. Its efficiency, scalability, and predictive power make it an indispensable tool in modern biomedical research and clinical practice [20-24]. Prostate cancer (PCa) represents a major health burden globally, being the most prevalent malignancy and a leading cause of cancer-related morbidity and mortality among men. Despite advances in early detection and treatment modalities, the prognosis for patients with advanced or metastatic PCa remains poor[25]. In recent years, research efforts have increasingly focused on unraveling the molecular underpinnings of PCa pathogenesis to identify novel therapeutic targets and improve patient outcomes[26]. One of the most significant molecular aberrations implicated in PCa is the loss of the tumor suppressor gene phosphatase and tensin homolog (PTEN). PTEN is a critical regulator of cellular processes such as proliferation, survival, and metabolism, exerting its tumor-suppressive effects primarily through its lipid phosphatase activity, which negatively regulates the PI3K/AKT/mTOR signaling pathway[27]. Due to the significance of PTEN in prostate cancer it is important to standpoint to examine the effects of mutations on PTEN structure and function. A study showing the structural and functional impact of PTEN in computational view has been published in endometrial cancer and ovarian cancer, and confirmed R173L/Q and R130C/H as highly pathogenic in those cancer types[28]. However, no studies have been focussed on in-depth analysis of SNPs occurring in PTEN of prostate cancer cases. Hence, *in silico* mutational analysis was performed in this study using various computational tools to predict the potential pathogenicity of genetic variants by assessing their effects on protein structure and function,

The study focused on collecting the 62 missense mutations of PTEN in prostate cancer from COSMIC server since it has large collection of datasets from both large-scale sequencing projects and individual research studies. A total of 5 webservers such as Ph-D-SNP, SIFT, SNAP, PAHTHER and metaSNP was used to analyse the pathogenicity of all SNPS's since no single tool would ensure high prediction. Mutations (29 SNPs) predicted as disease variants by all five tools were selected for further analysis since the ensemble strategy leads to more robust and accurate predictions. Further, Fathmm was used to obtain the cancer causing/driver mutation since it includes features for estimating the impact of

mutation in the context of cancer. Mutation accessor was utilized to prioritize the mutation causing the functional impact from the filtered 28 cancer causing variants. Multiple servers, including MuPro, iStable, I-mutant 2.0, mCSM, SDM, and DUET, were used to assess the stability of the missense variants. Stability analysis tools predict changes in protein stability caused by mutations. Disease-causing mutations often destabilize the protein and this information aids in assessing the impact of mutations on the protein structure. Among the 25 shortlisted mutations, 15 SNPs were predicted as destabilizing and may affect the protein folding.

Mutations in conserved residues are more likely to disrupt these critical functions, leading to protein dysfunction and potentially contributing to disease pathogenesis. Mutations in the binding site are also of particular interest due to their potential to disrupt protein-protein or protein-ligand interactions critical for PTEN's tumor-suppressive functions. In accordance with this, 13 destabilizing mutants were highly conserved. Among them two (C124S and G129R) mutants present in the binding site of PTEN were selected for further detailed structural analysis through simulation.

The RMSD of G129R showed a major deviation throughout the simulation period of 50 ns. Similarly, RMSF data shows that residue level variations were higher for mutant protein G129R than both the wildtype and other mutant type C124S. According to a recent study, this higher deviation of the mutant protein indicates that it had a strong destabilising effect on the protein. It was also demonstrated that mutants that have higher RMSF had a significant impact on protein structure, resulting in increased protein flexibility. In accordance with the RMSD and RMSF, Rg plot also showed less compactness in G129R in comparison to wildtype and C124S making it unstable. Based on all of these *in silico* predictions, this study demonstrated that specific mutations (G129R) have a destabilising and robust effect on the PTEN protein. In correlation with our study, C124S in OVCAR-3 cells lacked the lipid phosphatase activity and had high amount of Akt in its nucleus[29]. Similarly, catalytically dead mutant G129R in prostate cancer cell line resulted in the increased activation of cyclin A in comparison to wild PTEN[30].

Conclusion

Our study demonstrated that the predicted PTEN mutants C124S and G129R were highly pathogenic to cause protein dysfunction and induce destabilization in prostate cancer. In addition, these mutant were present in the binding site of PTEN also added the value. Structural analysis among PTEN mutants through simulation demonstrated G129R occurring in the functional domain cause significant change in stability and compactness of the protein. The RMSF analysis provides us an insight that both the mutants cause structural instability, residue fluctuations and cause huge impact on the function of the protein. The study's comprehensive analysis of PTEN-associated SNP's sheds light on the potential role of G129R in prostate cancer. The experimental prediction of structural and functional insights provided by computational tools might offer potential avenues for improved diagnostics and therapeutics.

Conflict of interest

All authors declare no conflict of interest.

References

1 Hassanpour SH & Dehghani M, Review of cancer from perspective of molecular. *J Cancer Res Pract*, 4 (2017) 127.

2 Yu T, Nantasenamat C, Kachenton S, Anuwongcharoen N & Piacham T, Cheminformatic Analysis and Machine Learning Modeling to Investigate Androgen Receptor Antagonists to Combat Prostate Cancer. *ACS Omega*, 8 (2023) 6729.

3 Guo Z, Lu X, Yang F, Qin L, Yang N, Wu J & Wang H, Docetaxel chemotherapy plus androgen-deprivation therapy in high-volume disease metastatic hormone-sensitive prostate cancer in Chinese patients: an efficacy and safety analysis. *Eur J Med Res*, 27 (2022) 148.

4 Perlmutter MA & Lepor H, Androgen deprivation therapy in the treatment of advanced prostate cancer. *Rev Urol,* 9 (2007) S3.

5 Vietri MT, Elia G, Caliendo G, Resse M, Casamassimi A, Passariello L, & Molinari AM, Hereditary prostate cancer: Genes related, target therapy and prevention. *Int J Mol Sci* 22 (2021) 3753

6 Armenia J, Wankowicz SA, Liu D, Gao J, Kundra R, Reznik E, Chatila WK, Chakravarty D, Han GC, Coleman I & Montgomery, The long tail of oncogenic drivers in prostate cancer. *Nat Genet,* 50 (2018) 645.

7 Cooper CS, Eeles R, Wedge DC, Van Loo P, Gundem G, Alexandrov LB, Kremeyer B, Butler A, Lynch AG, Camacho N & Massie CE, Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat Genet,* 47 (2015) 367.

8 Zhou X, Yang X, Sun X, Xu X, Li XA, Guo Y, Wang J, Li X, Yao L, Wang H & Shen, Effect of PTEN loss on metabolic reprogramming in prostate cancer cells. *Oncol Lett,* 17 (2019) 2856.

9 Hasle N, Matreyek KA & Fowler DM, The Impact of Genetic Variants on PTEN Molecular Functions and Cellular Phenotypes. *Cold Spring Harb Perspect Med*, 9 (2019) 11.

10 Álvarez-Garcia V, Tawil Y, Wise HM & Leslie NR, Mechanisms of PTEN loss in cancer: It's all about diversity. *Semin Cancer Biol,* 59 (2019) 66.

11 Jamaspishvili T, Berman DM, Ross AE, Scher HI, De Marzo AM, Squire JA & Lotan TL, Clinical implications of PTEN loss in prostate cancer. *Nat Rev Urol,* 15 (2018) 222.

12 Blumenthal GM & Dennis PA, PTEN hamartoma tumor syndromes. *Eur J Hum Genet*, 16 (2008) 1289.

13 Serebriiskii IG, Pavlov V, Tricarico R, Andrianov G, Nicolas E, Parker MI, Newberg J, Frampton G, Meyer JE & Golemis EA, Comprehensive characterization of PTEN mutational profile in a series of 34,129 colorectal cancers. *Nat Commun*, 13 (2022) 1618.

14 Dakal TC, Kala D, Dhiman G, Yadav V, Krokhotin A & Dokholyan NV, Predicting the functional consequences of non-synonymous single nucleotide polymorphisms in IL8 gene. *Sci Rep,* 6 (2017) 6525.

15 Katturajan R, Medha T, Karra S & Prince SE, A comparative computational approach on the most deleterious missense variant in Connexin 43 protein and its potent inhibitor analysis. *Indian J Biochem Biophys,* 60 (2023) 7.

16 Janani DM, Poornima G & Usha B, *In silico* characterization of structural and functional impact of the deleterious snps on fshr gene. *Indian J Biochem Biophys,* 56 (2019) 492.

17 Capriotti E, Altman RB & Bromberg Y, Collective judgment predicts disease-associated single nucleotide variants. *BMC Genomics*, 14 (2013) 1.

18 Wu Q, Peng Z, Zhang Y & Yang J, COACH-D: Improved protein-ligand binding sites prediction with refined ligand-binding poses through molecular docking. *Nucleic Acids Res*, 46 (2018) W438.

19 Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B & Lindahl E, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Software X,* 1 (2015) 19.

20 Janani DM & Usha B, *In silico* analysis of functional non-synonymous and intronic variants found in a polycystic ovarian syndrome (PCOS) candidate gene: *DENND1A*. *Indian J Biochem Biophys*, 57 (2020) 584.

21 Dey P & Kumar P, Mutational analysis of resveratrol-cleaving dioxygenase towards enhancement of vanillin synthesis. *Indian J Biochem Biophys,* 58 (2021) 284.

22 Erdogan T, Computational evaluation of 2-arylbenzofurans for their potential use against SARS-CoV-2: A DFT, molecular docking, molecular dynamics simulation study. *Indian J Biochem Biophys,* 59 (2022) 59.

23 Ganeshpurkar A, Chaturvedi A, Shrivastava A, Dubey N, Jain S, Saxena N, Gupta P & Mujariya R, *In silico* interaction of Berberine with some immunomodulatory targets: A docking analysis. *Indian J Biochem Biophys*, 59 (2022) 848.

24 Latha V, Gomathi V & Rajeshkanna A, Generating a potent inhibitor against MCF7 breast cancer cell through artificial intelligence based virtual screening and molecular docking studies. *Indian J Biochem Biophys,* 60 (2023) 844.

25 Kotrikadze N, Alibegashvili M, Ramishvili L, Mikaia N, Khazaradze A, Sepiashvili B, Nakashidze I, Gordeziani M & Ahmad S, The role of lipids and fatty acid metabolism in the development of prostate cancer. *Indian J Biochem Biophys*, 59 (2022) 873.

26 Ferraldeschi R, Rodrigues DN, Riisnaes R, Miranda S, Figueiredo I, Rescigno P, Ravi P, Pezaro C, Omlin A, Lorente D & Zafeiriou Z, PTEN protein loss and clinical outcome from castration-resistant prostate cancer treated with abiraterone acetate. *Eur Urol*, 67 (2015) 795.

27 Turnham DJ, Bullock N, Dass MS, Staffurth JN & Pearson HB, The PTEN Conundrum: How to Target PTEN-Deficient Prostate Cancer. *Cells*, 9 (2020) 2342

28 Smith IN & Briggs JM, Structural mutation analysis of PTEN and its genotype-phenotype correlations in endometriosis and cancer. *Proteins,* 84 (2016) 1625.

29 Phadngam S, Castiglioni A, Ferraresi A, Morani F, Follo C & Isidoro C, PTEN dephosphorylates AKT to prevent the expression of GLUT1 on plasma membrane and to limit glucose consumption in cancer cells. *Oncotarget,* 7 (2016) 84999.

30 Wegiel B, Bjartell A, Culig Z & Persson JL, Interleukin-6 activates PI3K/Akt pathway and regulates cyclin A1 to promote prostate cancer cell survival. *Int J Cancer*, 122 (2008) 1521.