

RESEARCH

Open Access



Exploring the prognostic significance of PKCε variants in cervical cancer

Sameen Zafar¹ , Khushbukhat Khan¹ , Yasmin Badshah^{1*} , Kanza Shahid¹ , Janeen H. Trembley^{2,3,4}, Amna Hafeez¹ , Naeem Mahmood Ashraf⁵ , Hamid Arslan⁶, Maria Shabbir¹ , Tayyaba Afsar⁷, Ali Almajwal⁷ and Suhail Razak^{7*}

Abstract

Background Protein Kinase C-epsilon (PKCε) is a member of the novel subfamily of PKCs (nPKCs) that plays a role in cancer development. Studies have revealed that its elevated expression levels are associated with cervical cancer. Previously, we identified pathogenic variations in its different domains through various bioinformatics tools and molecular dynamic simulation. In the present study, the aim was to find the association of its variants rs1553369874 and rs1345511001 with cervical cancer and to determine the influence of these variants on the protein-protein interactions of PKCε, which can lead towards cancer development and poor survival rates.

Methods The association of the variants with cervical cancer and its clinicopathological features was determined through genotyping analysis. Odds ratio and relative risk along with Fisher exact test were calculated to evaluate variants significance and disease risk. Protein-protein docking was performed and docked complexes were subjected to molecular dynamics simulation to gauge the variants impact on PKCε's molecular interactions.

Results This study revealed that genetic variants rs1553369874 and rs1345511001 were associated with cervical cancer. Smad3 interacts with PKCε and this interaction promotes cervical cancer angiogenesis; therefore, Smad3 was selected for protein-protein docking. The analysis revealed PKCε variants promoted aberrant interactions with Smad3 that might lead to the activation of oncogenic pathways. The data obtained from this study suggested the prognostic significance of PRKCE gene variants rs1553369874 and rs1345511001.

Conclusion Through further in vitro and in vivo validation, these variants can be used at the clinical level as novel prognostic markers and therapeutic targets against cervical cancer.

Keywords PKCε, Cervical cancer, Genotyping, Protein-protein interactions, Molecular docking, Molecular dynamics simulations

*Correspondence:

Yasmin Badshah
yasmeenb_1982@yahoo.com
Suhail Razak
smarazi@ksu.edu.sa

¹Department of Healthcare Biotechnology, Atta-Ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad, Pakistan

²Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN, USA

³Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA

⁴Minneapolis VA Health Care System Research Service, Minneapolis, MN, USA

⁵School of Biochemistry and Biotechnology, University of the Punjab, Lahore, Pakistan

⁶University of Bonn, LIMES Institute (AG-Netea), Carl-Troll-Str. 31, 53115 Bonn, Germany

⁷Department of Community Health Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia



© The Author(s) 2023, corrected publication 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Cervical cancer has the fourth highest incidence rate and remains the fourth highest cause of cancer-related deaths in females around the world, with 342,000 deaths and 604,000 new cases reported globally in 2020. According to WHO cancer statistics in 2020 cervical cancer is the one of the most prevalent cancer in Pakistani female population with third highest mortality rate and 3,197 estimated deaths [1, 2]. Chronic infections with high-risk human papillomavirus (HPV) are responsible for 99.7% of cases of cervical cancer [3]. However, progression to cervical cancer is a gradual process and involves three stages of pre-cancer lesions (i.e., cervical intraepithelial neoplasia: CIN-I, CIN-II, CIN-III) prior to the development of invasive cervical cancer [4, 5]. Factors other than HPV infection, including reproductive and sexual factors, oral contraceptives, family history, tobacco smoking, and socioeconomic status. Moreover, various studies have proved that genetic factors of hosts such as SNPs, also contribute to the development of cervical lesions and cervical cancer [6, 7].

PKCs belong to AGC-kinase family of serine/threonine kinases, and they are known to act as intersection points in various cellular processes and signal-transduction pathways [8, 9]. PKC ϵ (alternative name KPCE) is an isoform in the calcium-independent novel PKC sub-family [10]. It is a multi-functional protein involved in various cellular processes, encoded by PRKCE gene [11, 12]. Studies have shown that it is a transforming oncogene and its elevated levels regulate various signal transducers in oncogenic processes in various malignancies including cervical cancer [13–17]. Various studies have reported that expression of PKC isoforms is elevated in cervical cancer [18, 19].

Single nucleotide polymorphisms (SNPs) are the most common form of genetic variants occurring in human [20]. SNPs in the coding sequence of genes might have deleterious effects on the structure and function of the protein that may lead to or contribute to oncogenesis [21, 22]; as such, SNPs can be regarded as the genetic basis and as potential genetic markers for cancer susceptibility of an individual [23]. There are various SNPs identified in genes IL10, TP53, XRCC1, PAX8, CLPTM1L, HLA-A, and TGFBR2 which are associated with cervical carcinoma in different populations globally [4, 24, 25]. Genetic variants in PRKCE alter the structure and functions of the protein, ultimately leading towards cancer [26]. A few SNPs in PKC isoforms (rs454006 of PRKCG; rs546950 and rs4955720 of PRKCI; rs9907521 of PRKCA) have been reported to be associated with various cancers [27–31]. Recently, Khan et al. reported that several missense variants present in different domains of PKC ϵ are responsible for altering the normal functioning of the protein and might play roles in disease contribution. The

C2-like regulatory domain of PKC ϵ is involved in regulating interactions with proteins, which are essential for controlling PKC ϵ activation. Variants present in the regulatory domain tend to affect PKC ϵ 's flexibility, leading to deviation in the ability of PKC ϵ to make protein-protein interactions that might be associated with disease progression [32]. In a recent study it was indicated that the PRKCE variant rs1553369874 was associated with HCV-induced HCC [33]. However, so far, no published data has reported the association of these variants with the risk of cervical cancer in *in vivo* or *in vitro* studies.

In this study, two missense variants rs1553369874 and rs1345511001 of PRKCE, resulting in amino acid substitutions E14K and D39H, respectively, in the C2-like regulatory domain of the protein were selected. Their genotype frequencies were determined in cervical cancer patients and healthy control groups to reveal their possible association with the risk of cervical cancer in Pakistani population. Additionally, the findings of this study can be extended beyond the Pakistani population to the global context in the future, as these PKC ϵ variants might be associated with the risk of cervical cancer in other populations as well, hence holding the potential to be used as a novel biomarker for the prognosis of cervical cancer.

Protein interactions studies have proven to be helpful in the identification of cancer biomarkers by providing the knowledge about the physiological functioning of these biomolecules and their interactions occurring in an organism [34, 35]. Therefore, to evaluate the functional impact of these variants on the protein-protein interactions of PKC ϵ , molecular docking was performed using HADDOCK [36] with Smad3, an essential component in TGF- β signaling pathway known to be regulated by PKC ϵ [16, 37]. Smad3 is also known to mediate carcinogenic processes, particularly in cervical cancer [38, 39]. Lastly, the predicted docked structures were evaluated through molecular-dynamic simulations for further analysis and to attain improved insight to the molecular interactions of PKC ϵ .

Methodology

Sample collection and processing

Prior to the initiation of current study, ethical approval from institutional review board (IRB No. 10-2021-01/01) was obtained from parent department Atta-ur-Rahman School of Applied Biosciences (ASAB) of National University of Sciences and Technology (see online source 1). Written and oral informed consent were attained from each participant before the collection of blood samples (see online source 2). All methods were carried out in accordance with relevant guidelines and regulations.

In this study, 95 female cervical cancer patients were ascertained from combined military hospital (CMH)

after a thorough process of histological diagnosis. An equal number of control samples were also randomly collected. Patients suffering from HPV-induced cervical cancer were included for this study, while patients suffering with co-morbidity leading to cervical cancer and HIV co-infections were excluded from this study. Blood samples were obtained from the participants and stored in anticoagulating EDTA tubes for genotype analysis.

Genotype analysis

DNA from whole blood samples of study subjects was extracted through organic (phenol-chloroform) extraction protocol [40]. The primers for variants analysis were designed using Primer 1 software [41]. Four primers were designed for genotyping of each genetic variant. Two outer primers and two internal primers were designed in such a manner that they must amplify that region of the gene which carrying the targeted variants. Primers were also validated by UCSC *In silico* PCR [42]. The genotyping analysis for variants rs1553369874 (E14K) and rs1345511001 (D39H) was performed through Tetra-ARMS PCR in Veriti™ 96-Well Thermal Cycler of Applied Biosystems™. Reaction mixture of 25 µl was prepared for every sample using Solis BioDyne FIREPol® Master Mix with 7.5 mM MgCl₂. The primer sequences and parameters for PCR-reaction are given in Table 1. The amplified product was analyzed by agarose electrophoresis (2% W/V) under UV-transilluminator.

Table 1 Sequences and parameters of primers used for genotyping of rs1553369874 and rs1345511001

Variant rs IDs	Primer sequences	Tm°C	Ta°C	Product size
rs1553369874(G/A)	CTTCTTAAGAT- CAAAATCTTCA			G-allele – 196
	CTTCAAGCTCACGGCATC	53.9 °C	72 °C	A-allele – 169
	CACAAGGTGTAGGGAGTGT			Outer-band – 325
	GCTGTTGGTCTTCTGCTT CGCAGACTTTCCTTACC			G-allele – 192
rs1345511001 (G/C)	AGGGCAATGTAGGGCTC	55.6 °C	72 °C	C-allele – 212
	TTCTTCATTCTGCCCTC			Outer-band – 369
	TCAAAGTGGATGGTGCAG			

Statistical analysis

Statistical analysis on obtained genotyping data was performed using GraphPad prism 9.0 software (California, USA). Fisher exact test was applied on both cervical cancer and control samples. The analysis of Odds ratios and relative risk was also performed and their respective confidence-intervals were also identified. The P-value ≤ 0.05 was regarded as statistically significant.

Predicting the effect of variants on secondary structures of mRNA

To predict and study the effect of genetic variants on the secondary structure of PKCε mRNA an online mRNA structure prediction bioinformatics tool RNA-fold was used (<http://rna.tbi.univie.ac.at/>) [43, 44]. The minimum free energy (MFE) for wildtype and variants mRNA was predicted. MFE and secondary structures of wild-type and variants mRNAs were compared and analysed for the structural stability.

Protein-protein interaction analysis

To determine the impact of amino acid variants in PKCε on its interactions with target proteins, molecular docking was performed. The 3D structure of Smad3 used in molecular docking simulation was predicted from I-TASSER <https://zhanglab.ccmb.med.umich.edu/I-TASSER/> that use fold recognition/threading approach for 3D modeling of protein structures [45]. 3D structure of Smad3 was selected based on the highest C-score value depicting the global accuracy of the predicted model as well as low RMSD and local error values. The protein structures were further validated using Ramachandran analysis using PROCHECK [46]. Smad3 was chosen after the verification from literature that it is a target protein of PKCε and it is involved in the regulation of cancer cell metabolism and proliferation [16, 37, 47]. The variant structures of PKCε were attained through *In silico* mutagenesis using PyMol v4.0.4 [48]. For first variant glutamic acid at residue position 14 was substituted with lysine and aspartic acid at residue 39 was mutated to histidine in the second variant structure and they were subsequently used for molecular docking and molecular-dynamic simulations. Three Molecular docking simulations (WT-Smad3, E14K-Smad3, D39H-Smad3) were performed through HADDOCK SERVER 2.4 [36]. The cluster with lowest Z-value was selected from each docking simulation and docked structure from each cluster was chosen for further analysis.

Interaction dynamic analysis

The dynamics of molecular interactions between native PKCε and its variants with Smad3 were studied through MD simulations using GROMACS 2016 [49], with OPLS-AA force field [50] that was used for the simulation of

wildtype and variants PKC ϵ /Smad3 complexes. A cubic box was formed around each complex for solvation by adding SPC216 water molecules, and subsequent neutralization by incorporating Na⁺ / Cl⁻ ions. Initial energy minimization of MD-simulations was performed for 50,000 steps, followed by NVT and NPT equilibration for 100ps, through steepest descent. From the same random seed, trajectories of MD complexes were initiated. MD-simulations for wildtype and variants complexes were performed for the production run of 10ns. Dynamic trajectories were constructed using in-built programs of GROMACS 2016 (gmx_ trjconv). The visualization of docked structure was performed through LIGPLOT [51] and contact analysis of MD simulations was mostly performed using VMD [52]. Several structural analyses of complexes were also performed. Calculation of root mean square deviation (RMSD for protein backbone) was performed using command gmx_ rms, radius of gyration (Rg for protein, backbone) was calculating using gmx_ gyrate command, analysis of solvent accessibility surface area (SASA) number of hydrogen bonds was performed using gmx_ sasa and gmx_ hbond commands respectively. All MD analyses were plotted as scatter line plots.

Pathway construction

In order to construct a molecular pathway illustrating the impact of amino acid variations in PKC ϵ and its subsequent effect on cellular signalling, Kyoto Encyclopedia of Genes and Genomes (KEGG: hsa:351) [53] and GeneMANIA [54] were employed. Furthermore, STRING database [55] was used to predict and analyze genetic interactions while functional annotation of genes and the genetic pathway was attained through DAVID database [56].

Results

Clinical features of cervical cancer in patients

The current study involved 95 cervical cancer patients. The general features of the patients are presented in Table 2. The ages of cervical cancer patients range from 35 years to 66 years, whereas their median age was 51 years. Furthermore, the median age of the control group was 41 years, with the range of ages from 30 years to 64 years.

Table 2 Clinical features of cervical cancer patients and control

Clinical characteristics		Cervical cancer	Control
Age	≤ 50 years	47	81
	> 50 years	48	14
Clinical stage	Stage I-II	50	
	Stage III-IV	45	
Metastasis stage	Metastatic	35	
	Non-Metastatic	60	

Association of prkce genetic variants with cervical cancer and control

DNA extracted from the samples was genotyped for the presence of missense variants rs1553369874 (g. 45,652,140 G>A) and rs1345511001 (g. 45652215G>C) in the PRKCE gene; these variants result in the substitution of Glutamic acid (E) to Lysine (K) at residue 14 and Aspartic acid (D) to Histidine (H) at position 39. Tetra ARMS-PCR was used for the process of genotyping. This technique utilizes four primers that amplify the targeted gene sequence and result in an outer control band and inner genotype bands.

The frequency distributions for genotypes of both PKC ϵ genetic variants rs1553369874, and rs1345511001 for control and cervical cancer samples are presented in Table 3. It was found that for variant rs1553369874 genotype GG was associated with elevated risk for cervical cancer (Odds Ratio (OR)=2.571, Relative Risk (RR)=1.609, P=0.0022); whereas, genotype AA was found to have protective role in this regard (OR=0.1566, RR=0.3080, P<0.0001). For variant rs1345511001, genotype GG was found to be significantly associated with increased risk of cervical cancer (OR=2.363, RR=1.548, P<0.0056); whereas, genotype CC was statistically associated with a protective role in cervical cancer (OR=0.5128, RR=0.7031, P=0.0456).

Allele frequencies of both PKC ϵ variants are described in Table 3. Analysis revealed that the frequency of the G-allele in variant rs1553369874 was significantly higher in the cervical cancer group compared to the control group, (OR=0.3153, RR=0.5914, P=0.0004). The A-allele was also found to be significantly different between the two groups, and the OR and RR indicated that a pathogenic role (OR=3.171, RR=1.691, P-value=0.0004). Similarly, the frequency of the G-allele in variant rs1345511001 was higher in the cervical cancer group compared to the control group, and the G allele associated with a protective role (OR=0.4487, RR=0.6809, P-value=0.0116). The frequency of the C-allele in this variant was increased in the control samples and has role in elevated risk of cervical cancer (OR=2.228, RR=1.469, P-value=0.01).

Association of prkce genetic variants with clinical characteristics of cervical cancer patients

Frequency distribution of genotypes for both variants in different age groups are given in Table 4. Genotypes GG and GA of variant rs1553369874 were not statistically significant in either age group; however, genotype AA of the variant showed statistical significance for subjects of both age groups (≤50 years: OR=3.596, RR=1.439, P=0.0118; >50 years: OR=26.11, RR=5.185, P=0.0016). The CC-genotype of variant rs1345511001 showed significance for the age group of ≤50 years (OR=2.368,

Table 3 Genotype and allele distribution and association of PRKCE variants rs1553369874 and rs1345511001 with cervical cancer

SNP	Genotype	Frequency of distribution		OR		RR		P-value
		Control	Patients	OR	95% CI	RR	95% CI	
rs1553369874	GG	38 40.00%	60 63.16%	2.571	1.401–4.677	1.609	1.197–2.204	0.002
	AA	32 33.68%	7 7.37%	0.1566	0.06131–0.3646	0.3080	0.1523–0.5708	<0.0001
	GA	25 26.32%	28 29.47%	1.170	0.6229 to 2.227	1.080	0.7786–1.442	0.7465
	G	50 52.63%	74 77.89%	0.3153	0.1730–0.5988	0.5914	0.4509–0.7780	0.0004
	A	45 47.37%	21 22.11%	3.171	1.670–5.781	1.691	1.285–2.218	0.0004
rs1345511001	GG	41 43.16%	61 64.21%	2.363	1.327–4.281	1.548	1.149–2.127	0.0056
	CC	39 41.05%	25 26.32%	0.5128	0.2788–0.9541	0.7031	0.4897–0.9713	0.0456
	GC	15 15.79%	9 9.47%	0.5581	0.2345–1.335	0.7238	0.4018–1.142	0.2747
	G	48 50.53%	66 69.47%	0.4487	0.2547–0.8036	0.6809	0.5146–0.9016	0.0116
	C	47 49.47%	29 30.53%	2.228	1.244–3.927	1.469	1.109–1.943	0.0116

Table 4 Genotype distribution and association of PRKCE variants rs1553369874 and rs1345511001 with cervical cancer in age groups of ≤50 and >50

SNP	Genotype	Frequency of distribution		OR		RR		P-value
		Control	Patients	OR	95% CI	RR	95% CI	
rs1553369874	GG (≤50 YEARS)	32 39.51%	26 55.32%	0.5275	0.2602–1.115	0.7882	0.5867–1.031	0.0989
	GG (>50 YEARS)	6 42.86%	34 70.83%	0.3088	0.09686–0.9917	0.4125	0.1683–1.015	0.0657
	AA (≤50 YEARS)	27 33.33%	6 12.77%	3.596	1.339–8.788	1.439	1.103–1.816	0.0118
	AA (>50 YEARS)	5 35.71%	1 2.08%	26.11	2.638–311.4	5.185	2.267–10.11	0.0016
	GA (≤50 YEARS)	22 27.16%	15 32.61%	0.7706	0.3477–1.644	0.9070	0.6453–1.198	0.5466
	GA (>50 YEARS)	3 21.43%	13 27.08%	0.7343	0.1955–2.795	0.7841	0.2507–2.161	>0.9999
rs1345511001	GG (≤50 YEARS)	35 43.21%	29 61.70%	0.4723	0.2265–1.000	0.7609	0.5726–0.9923	0.0662
	GG (>50 YEARS)	6 42.86%	32 66.67%	0.3750	0.1156–1.181	0.4737	0.1921–1.168	0.1287
	CC (≤50 YEARS)	34 41.98%	11 23.40%	2.368	1.081–5.020	1.334	1.024–1.717	0.0368
	CC (>50 YEARS)	5 35.71%	14 29.17%	1.349	0.4126–4.377	1.257	0.4830–3.053	0.7444
	GC (≤50 YEARS)	12 14.81%	7 14.89%	0.9938	0.3598–2.566	0.9977	0.6362–1.350	>0.9999
	GC (>50 YEARS)	3 18.75%	2 4.17%	5.308	0.9671–31.28	2.723	0.9578–5.467	0.0949

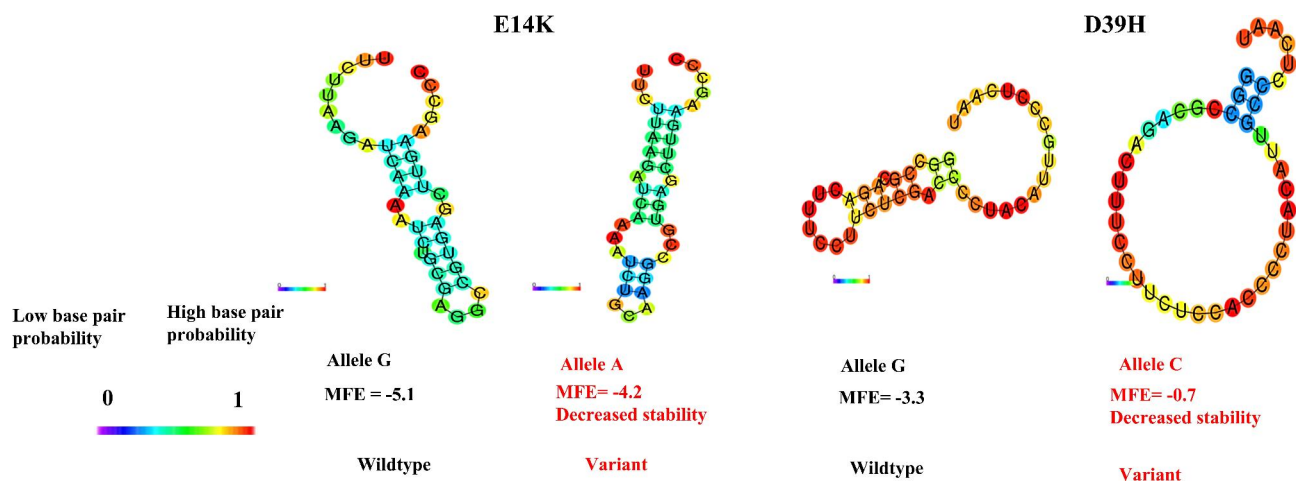
RR=1.334, P=0.0368). All other genotypes of this variant were non-significant in both age groups.

Genotype distribution frequencies among non-metastatic and metastatic patients were also determined

in Table 5. In variant rs1553369874, genotypes AA and GA were shown to be statistically correlated with cervical cancer metastasis; genotype AA has a damaging role whereas genotype GA was shown to have a protective

Table 5 Genotype distribution and association of PRKCE variants rs1553369874 and rs1345511001 with secondary site metastatic and stages of cervical cancer

SNP	Genotype	Frequency distribution		OR	RR	P-Value	Frequency distribution		OR	RR	P-Value
		Metastatic	Non-metastatic				Stage I-II	Stage III-IV			
rs1553369874	GG	24 68.57%	36 60.00%	1.455	1.273	0.5095	31 62.00%	29 64.44%	0.9002	0.9518	0.8342
	AA	6 17.14%	1 1.67%	12.21	2.601	0.0094	1 2.00%	6 13.33%	7.538	1.934	0.0500
	GA	5 14.29%	23 38.33%	0.2681	0.3988	0.0189	18 36.00%	10 22.22%	1.969	1.346	0.1784
rs1345511001	GG	20 57.14%	41 68.33%	0.6179	0.7432	0.3750	36 72.00%	25 55.56%	2.057	1.433	0.1334
	CC	11 31.43%	14 23.33%	1.506	1.283	0.4704	11 22.00%	14 31.11%	0.6245	0.7897	0.3566
	GC	4 11.43%	5 8.33%	1.419	1.233	0.7211	3 6.00%	6 13.33%	0.4149	0.6099	0.3000

**Fig. 1** Predicted structures of mRNAs using RNAfold (a) rs1553369874 effect on mRNA structure due to the substitution of reference allele G (left) with variant allele A (right) (b) rs1345511001 effect on mRNA structure due to the substitution of reference allele G (left) with variant allele C (right) Darker colors represent high probability for base-pairing whereas lighter colors indicate that these sequences do not form base-pairs at their alignment positions

role. None of the genotypes in variant rs1345511001 showed any statistical correlation with metastasis of cervical cancer. Lastly, the distribution of genotypes among the cervical cancer patients of stage I-II and stage III-IV were determined (Table 5). It has been reported here that only genotype AA of variant rs1553369874 showed significant correlation as a risk factor in patients with stage-III or stage-IV of cervical cancer. Genotype distribution of variants with cervical cancer patient's clinical features is given (see online source 3).

Impact of genetic variants rs1553369874 and rs1345511001 on prkce mRNA structure

In silico analysis was performed to predict the secondary structures of mRNAs of wildtype PKCε and its selected variants. Minimal free energy (MFE) for the wildtype and variants was also determined and analyzed. The secondary structures of mRNA for rs1553369874 and

rs1345511001 showed a dramatic change when compared to their wildtype counterparts, demonstrating the significant effect of variant alleles on the overall structure of mRNAs (Fig. 1). MFE calculations for alleles of the variant rs1553369874 predicted that MFE for the reference G allele was -5.1 Kcal/mol, whereas MFE was elevated for variant allele A with the value of -4.2 Kcal/mol. The variant rs1345511001 showed a decrease in MFE value for the wildtype allele G (MFE = -3.3 Kcal/mol) compared to the variant allele C with MFE value -0.7 Kcal/mol (see online source 4). The decreased values of MFE for the reference alleles in both variants depict increased structural stability [57]. Whereas the altered alleles for both variants had increased MFE values compared to their wildtype counterparts and thus were structurally less stable. The altered alleles in both variants are associated with cervical cancer risk (see Table 3) which might be the result of decreased mRNA structural stability.

Analysing intramolecular interactions between PKC ϵ and Smad3 complexes

Molecular interactions of PKC ϵ with Smad3 were also determined and molecular docking was performed between PKC ϵ (wildtype, E14K, D39H) and Smad3 to predict and compare the nature of interactions of PKC ϵ prior to and after the occurrence of missense variations. 3-dimensional structures of Smad3 and PKC ϵ were predicted through I-TASSER and the models with highest C-score values were selected which indicated the quality of the predicted 3D models. Three docking simulations were carried out between wildtype PKC ϵ and Smad3, E14K and Smad3, and D39H and Smad3 using HADDOCK SERVER 2.4 and clusters of docked models were generated. The cluster with lowest value of Z-score was selected from each docking simulation, which indicates that those clusters have lowest values of standard deviation, and the best protein interaction model from each cluster was selected that corresponds to the highest score structure [58, 59]. It was particularly noteworthy that in the computational structures obtained from docking, PKC ϵ (wildtype, E14K, D39H) approached Smad3 mostly through its catalytic domain. This finding is exactly in accordance with the previous findings that Smad3 is a phosphorylation substrate of PKC ϵ [16, 37]. While analyzing the intramolecular interactions between native PKC ϵ and its variants with Smad3, it was determined that variants E14K and D39H interacted with Smad3 via catalytic domain residues by forming increased numbers of hydrogen bonds and hydrophobic interactions. Whereas native PKC ϵ interacts with Smad3 by making fewer hydrogen bonds and hydrophobic interactions. It was also observed that the wildtype PKC ϵ and its variants are involved in strong interactions with the following residues of Smad3: Ser213, Pro214, Ala215, His216, Asn217, Asn218, Asp220. The visualization of intermolecular interactions between wildtype and variants PKC ϵ with Smad3 in the docked structures was performed through LigPlot+v.2.2.4 as shown in online source 5. To further validate the molecular interactions of the protein-protein interaction complexes, PDBsum server was used which presents the interaction-statistics as well as depicts the protein interacting sites and interacting residues [60]. Furthermore, this tool also provides the information about the types of the molecular interactions present in protein-protein interaction complexes which includes hydrogen binds, salt-bridges, disulfide bonds, and non-bonded interactions. It was revealed that the linker domain residues of Smad3 made interactions via salt bridges, hydrogen-bonds, and non-bonding interactions with the kinase domain of PKC ϵ that resulted in Smad3 activation (Fig. 2).

Interactions dynamic analysis of wildtype and variant PKC ϵ - Smad3 complexes

In order to understand the consequences of amino acid variants on the structure of PKC ϵ and dynamics of its interactions with Smad3, MD-simulations were performed that predict the molecular motion and behavior of protein complexes in four dimensions [61]. GROMACS tool was used to carry out simulations between wildtype PKC ϵ and its variants with Smad3 for a time period of 10 ns and data for structural analysis of all three complexes was obtained (see online source 6).

RMSD (root mean square deviation) analysis was performed to analyze the average residual deviations in the protein structures over the course of simulation [58]. RMSD values for PKC ϵ and its variants are given in the figure (Fig. 3a). RMSD values increased to 0.3 nm for all three structures during the 0 ns – 1 ns period of simulation. This value was further increased to 0.5 nm for all structures at 4 ns. After that, RMSD became more or less stable for the wildtype PKC ϵ structure till the end of simulation. However, RMSD for both variants started to increase from the time interval of 5 ns and reached the value of 0.63 nm till the end of simulations, though during this period RMSD values for both variants remained constant.

To evaluate the variation in accessibility of surface area of the protein complex imparted by amino acid variants, SASA (solvent accessibility surface area) analysis was carried out. The data presented in Fig. 2 demonstrates that values of SASA were significantly increased from 540 nm² to 620 nm² during 0 ns – 1 ns of simulations (Fig. 3b). SASA values were further increased to 640 nm² for the native PKC ϵ till 2 ns; however, after that SASA values gradually decreased and at 10 ns this value was 620 nm² for wildtype PKC ϵ . SASA values for both variants showed dramatic fluctuations in their values. Particularly, during the time frame of 6 ns – 8 ns SASA for E14K was significantly reduced to 590 nm², whereas SASA for D39H was recorded to be 633 nm². However, during the interval of 8 ns – 10 ns SASA for all three complex remained stable around 610 nm².

To determine the effect of variants on the compactness of the protein complexes, the Rg (radius of gyration) was evaluated for MD-complexes. Data from the simulations is presented (Fig. 3c). The variant E14K has Rg value of 3.6 nm at the beginning of the simulations that reached value of 3.7 nm at 2 ns. This value remained somewhat constant till 4.8 ns. However, Rg value was slightly decreased to 3.65 nm during the time frame of 5 ns – 9 ns, but by the end of simulation, its value was approximately 3.7 nm. Rg for wildtype has initial value of 3.25 nm whereas for D39H its value is 3.35 nm. During 0 ns -2 ns, Rg for D39H remained higher compared to the wildtype but this value was approximately equal to

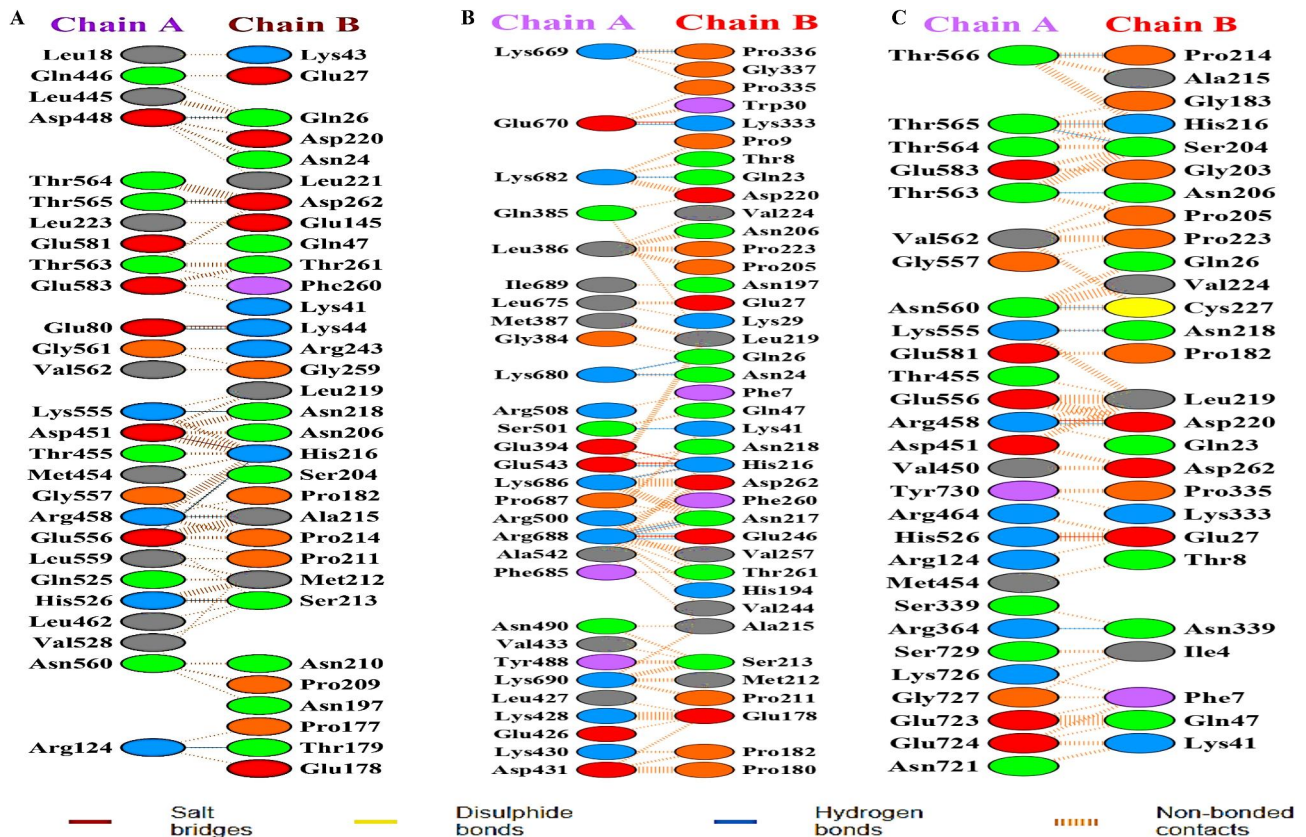


Fig. 2 PDBsum illustration depicting list of interacting amino acid residues of chain A representing PKC ϵ and chain B representing Smad3 (a) List of interacting residues in wildtype PKC ϵ -Smad3 complex (b) List of interacting residues in E14K-Smad3 complex (c) List of interacting residues in D39H-Smad3 complex

3.4 nm during the time period of 2 ns – 8 ns; however, its value decreased once more to approximately 3.36 nm during the last 2 ns.

Hydrogen bonds plays a critical role in the molecular interactions and maintenance of 3D structures of proteins. Number of hydrogen bonds that are formed in protein structures during the course of simulations are determined to analyze the effect of variants on protein interactions. It can be seen that during the period of 0 ns – 2 ns, the number of hydrogen bonds decreased from 700 to 600, but after that interval the number of hydrogen bonds remained stable throughout the course of simulations (Fig. 3d). Interactions dynamic analysis that the number of hydrogen bonds formed in E14K and D39H complexes are slightly higher compared to that in the wildtype PKC ϵ , showing that the molecular interactions which are formed between variants with Smad3 are more potent and stable compared to that of the wildtype.

The distances between interacting residues of PKC ϵ (wildtype, E14K, D39H) and residues of Smad3 were also determined. The residues selected for this purpose were involved in intramolecular interactions by forming hydrogen bonds. Three pairs of such interacting residues were randomly selected and their distances were

measured at three different time intervals (0 ns, 5 ns, 10 ns) of MD simulations. In the wildtype PKC ϵ -Smad3 complex, the distance between the interacting residues were determined at 0 ns and 10 ns. It was shown that at 10ns the distance between Asp451 of PKC ϵ and His216 of Smad3 was decreased to 13.51Å which was 15.14Å at 0ns. The distance between His525 of PKC ϵ and Pro211 of Smad3 was also significantly decreased to 5.68Å at the end of simulations, which was 13.63Å at 0ns. While the distance between Gly561 and Asp193 was increased to 12.92Å from 8.06Å at the end of simulations (Fig. 4a). Similarly, distances between the interacting residues of E14K-Smad3 depicted that the bond length between the residues Asp699 (PKC ϵ) and Asn197 (Smad3) as well as Tyr468 (PKC ϵ) and Asp220 (Smad3) were decreased to 14.57Å and 7.08Å at 10ns compared to 17.85Å and 12.76Å at 0ns, respectively. However, the bond length between the interacting pair Lys437 (PKC ϵ) and Pro209 (Smad3) was increased to 14.09Å from 6.68Å at the end of simulations (Fig. 4b). Lastly, bond distances between the interacting residues of D39H and Smad3 were also determined which also exhibited similar trends in their bond lengths. The distance between interacting residues Thr566 (PKC ϵ) and Tyr226 (Smad3) was 4.94Å at 10ns,

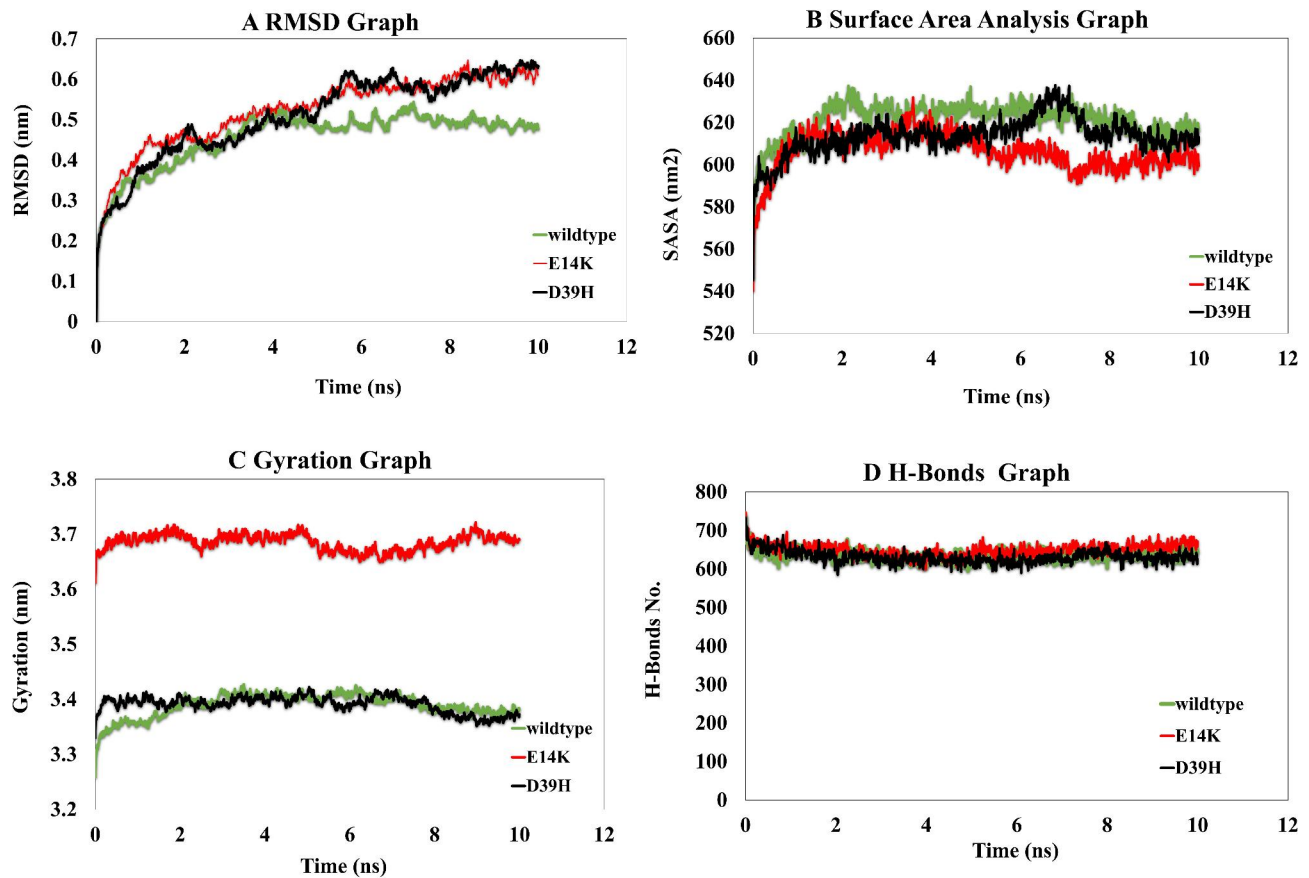


Fig. 3 Molecular dynamic analysis of wildtype and variants PKCε-Smad3 **(a)** Root mean square deviation (RMSD) **(b)** Solvent accessibility surface area (SASA) **(c)** Radius of gyration (Rg) **(d)** Number of Hydrogen bonds

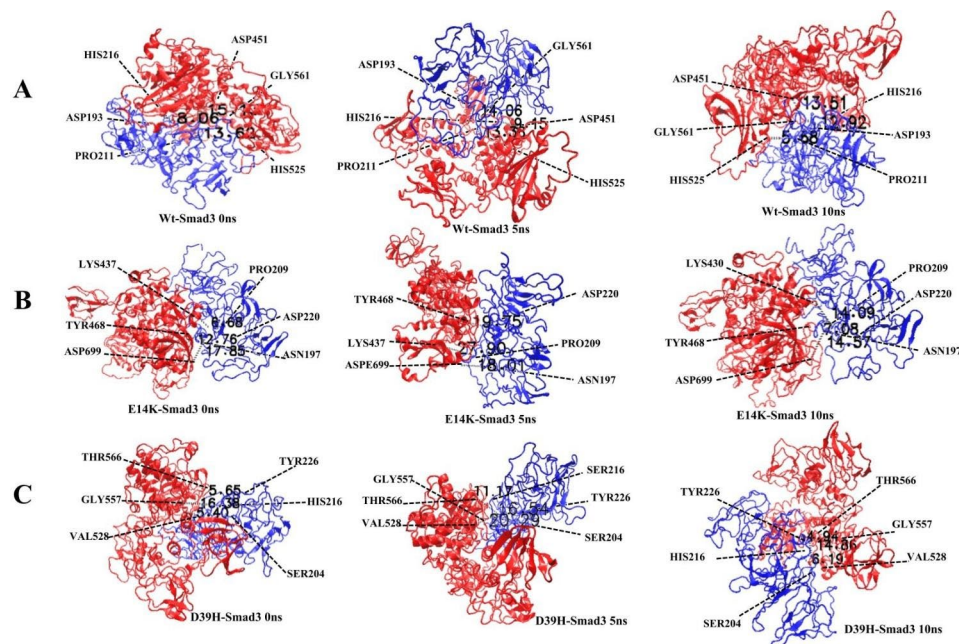


Fig. 4 Molecular distances between PKCε and Smad3 at three random pairs of interacting residues. Red colour represents PKCε/variants protein and Blue colour represents Smad3 protein in PKCε-Smad3 complex **(a)** Molecular distances recorded for wildtype PKCε-Smad3 complexes at 0ns, 5ns, 10ns **(b)** Molecular distances recorded for E14K-Smad3 at 0ns, 5ns, 10ns **(c)** Molecular distances recorded for D39H-Smad3 at 0ns, 5ns, 10ns

while at 0ns the bond distance was recorded to be 5.65Å. Similarly, the bond distance between Gly557 (PKCε) and His216 (Smad3) was reduced to 14.86Å at 10ns from 16.38Å recorded at the start of the MD simulations (Fig. 4c). Distances between the interacting residue pairs for all three complexes are described in Table 6.

Discussion

Several studies have demonstrated that genetic variants in PKCε and its isoforms are associated with several diseases, particularly cancers [62–65] [66]. The present study has explored the role of missense variants in PRKCE in association with cervical cancer in Pakistani female population. Studies have demonstrated that PRKCE non-synonymous variants rs1553369874 and rs1345511001 might have a damaging role leading to cancer susceptibility [32]. The results of this study have suggested that variants rs1553369874 and rs1345511001 in PRKCE are associated with risk of cervical cancer susceptibility in Pakistani population.

It was suggested from the tetra ARMS-PCR data that genotypes GG of rs1553369874 and GG of rs1345511001 were associated with increased risk of cervical cancer, whereas their counterpart genotypes AA and CC were most prevalent in healthy controls and imparted protective effect in Pakistani female population. This result was consistent with reported studies where variant genotypes were not associated with disease progression, rather they were imparting a protective effect in this regard [67–70]. Though allele A of rs1553369874 and allele C of rs1345511001 had high distribution frequency in control samples, their ORs, and RRs were higher compared to that of reference alleles of both variants, respectively. This discrepancy in data might have occurred because of small cohort size.

The cervical cancer incidence rates are highest in woman of age 35–50 years, whereas they are low for females of age 51–65 years [71]. It was found that genotype AA of rs1553369874 was acting as a risk factor in women of both age groups, while only genotype CC for the variant rs1345511001 was found to have detrimental role in females of age ≤50 years. Association of genetic variants with tumor metastasis state and tumor stage was also determined and genotype AA of rs1553369874 was shown to be correlated with advanced stage cervical cancer and secondary site metastasis.

The findings of this research have the potential to be implicated in cervical cancer studies beyond the studied population. These PRKCE variants may contribute towards the risk of cervical cancer in various population, as similar SNPs have been corelated with the risk of cancers in previous studies. For instance, previously studies have identified genetic variants rs546950, rs4955720 in PRKCI gene which are associated with the risk of prostate cancer in Han Chinese [30] and Iranian populations [67]. Similarly variant rs1801270 in CDKN1A gene are associated with development of cervical cancer in Han Chinese [69], Brazilian [72], and Iranian populations [73]. Nonetheless, further research is required in diverse population with large cohort size to validate the finding of this study as well as to assess the global significance of these variants with cervical cancer.

Secondary structures of mRNAs of both variants and their corresponding wildtype were predicted and the difference of mRNA structures and MFE values between wildtype and variants revealed that the structural stability of wildtype mRNA was higher compared to the variant mRNAs, which could lead to the onset of diseases [57, 74]. This analysis could be verified from the fact that altered alleles A and C of variants rs1553369874

Table 6 Bond distance between PKCε and Smad3 interacting residues

Complex	Residue pair 1	Distance	Residue pair 2	Distance	Residue pair 3	Distance
Wt.PKCε-Smad3	Asp451(PKCε) - His216(Smad3)	15.14 Å ⁰	Gly561(PKCε) - Asp193(Smad3)	8.06 Å ⁰	His525(PKCε) - Pro211(Smad3)	13.63 Å ⁰
		(0ns)		(0ns)		(0ns)
		9.15 Å ⁰		14.06 Å ⁰		13.36 Å ⁰
		(5ns)		(5ns)		(5ns)
		13.51 Å ⁰		12.92 Å ⁰		5.69 Å ⁰
		(10ns)		(10ns)		(10ns)
E14K-Smad3	Lys437(PKCε) - Pro209(Smad3)	6.68 Å ⁰	Tyr468(PKCε) - Asp220(Smad3)	12.76 Å ⁰	Asp699(PKCε) - Asn197(Smad3)	17.85 Å ⁰
		(0ns)		(0ns)		(0ns)
		27.90 Å ⁰		19.75 Å ⁰		18.01 Å ⁰
		(5ns)		(5ns)		(5ns)
		14.09 Å ⁰		7.08 Å ⁰		14.57 Å ⁰
		(10ns)		(10ns)		(10ns)
D39H-Smad3	Thr566(PKCε) -Tyr226(Smad3)	5.65 Å ⁰	Gly557(PKCε) - Ser216(Smad3)	16.38 Å ⁰	Val528(PKCε) - Ser204(Smad3)	5.40 Å ⁰
		(0ns)		(0ns)		(0ns)
		16.57 Å ⁰		11.17 Å ⁰		20.29 Å ⁰
		(5ns)		(5ns)		(5ns)
		4.94 Å ⁰		14.86 Å ⁰		6.19 Å ⁰
		(10ns)		(10ns)		(10ns)

and rs1345511001 were associated with increased cervical cancer risk. Furthermore, secondary structures of mRNAs are responsible for the modulation of translation rates and any changes in mRNA structures can lead to protein misfolding and structural instability [75–77], that can lead to severe illnesses including cancer.

Such variants that cause protein structure instability can lead to cancer progression. Destabilizing variants in E-cadherin have been reported to be associated with the development of gastric cancer at early ages compared to the variants that do not alter the protein stability [78]. Similarly, missense variants in c-src destabilized the structure of the protein and caused oncogenic transformation [76, 79]. Moreover, such structural destabilizations can also alter the protein-protein interactions [80]. Cervical cancer associated variants in PKC ϵ may also affect its interactions with other proteins, leading towards oncogenic progression.

The effect of PKC ϵ 's non-synonymous variants (E14K, D39H) on functions of the protein was studied along with the effect of amino acid variants on interactions of PKC ϵ with Smad3 that are responsible for tumor cell growth [66] [81]. 3D structure of PKC ϵ and Smad3 was predicted through I-TASSER that uses threading approach for prediction of protein structures [45]. *In silico* mutagenesis was performed on the predicted 3D structure of PKC ϵ and the two variant proteins were attained. The variants E14K and D39H were shown to be pathogenic in previous studies [32] and this study evaluated their impact on PPI of PKC ϵ with Smad3.

Several studies have reported the role of protein-protein interactions in identification of tumor biomarkers [34]. The integration of PPIs network analysis with computational bioinformatics techniques can be considered valuable for getting a deep insight into the intricated biochemical interactions occurring among the proteins present in a physiological system and which can give valuable information for designing potential prognostic biomarkers for cervical cancer [35, 82].

PKC ϵ and its variants were than docked with Smad3 to determine the alteration in protein-protein interactions resulting from the single amino acid variation. The interaction parameters included score of desolvation, electrostatic and hydrophobic interactions, and hydrogen bonding [83] which were increased for the variants than in wild type. These results clearly indicate that interactions of E14K and D39H with Smad3 were stronger than that of wild type (online source 7).

The dynamics of molecular interactions of PKC ϵ and its variants with Smad3 were analyzed through MD simulations. The number of hydrogen bonds that are formed within the complex over the course of MD simulations were also slightly higher for variants compared to wild type. This shows that intramolecular interactions

between PKC ϵ variants and Smad3 are stronger as hydrogen bonds plays a crucial role in stabilizing protein structure and its activity [84, 85]. The distance between interacting residues of all three complexes was also analyzed at three random sites. It was determined that those interacting residues of native PKC ϵ /variants and Smad3 were linked through hydrogen bonds. During the course of simulations, the distance between those interacting residues showed alterations. The molecular distances for some residues were increased during the middle of simulations, however, their distances showed shrinkage at the end of simulations. Interactions between variant PKC ϵ -Smad3 complexes were stronger due to extensive hydrogen bonding and shorter length of hydrogen bonds which are critical for the maintenance of stable protein interactions [86].

In this study analysis regarding the impact of non-synonymous PKC ϵ variants on the crosstalk with Smad3 and the downstream molecular actors was performed. For this purpose, several genetic and protein-interactions databases were used. The information regarding gene-functional annotation and genetic linkage was attained through an extensive literature review and this information was also obtained from several databases including KEGG, STRING, geneMANIA, and DAVID. It was revealed from the studies that PKC ϵ phosphorylates Smad3 at its linker region serine residue resulting in activation and nuclear translocation in a Tgf- β independent manner [81, 87]. Inside the nucleus activated Smad3 binds the promoter regions of several target genes and activates their transcription. It has been revealed through docking simulations and interaction dynamics analysis that variant PKC ϵ proteins interact with Smad3 in an aberrant manner compared to their native counterpart. The binding affinity between the variant PKC ϵ proteins and Smad3 is significantly higher compared to that in the wildtype-Smad3 complex. Since the binding interactions between both variants and Smad3 are much stronger, it will increase the binding energy of the kinase and Smad3 complex which in turn will lead towards the lowering of activation energy. This lowering in activation energy will assist in the elevated speed of the phosphorylation reaction. When the rates of Smad3 phosphorylation are increased, its activation levels will be aberrantly increased. These abnormally high levels of phosphorylated Smad3 will then mobilize to nucleus where they will frequently bind the promoter regions of target genes including HKII, MCT4, and Hif-1 α , resulting in their increased rates of transcription. Elevated levels of these genes are known to regulate key oncogenic processes including Warburg's effect. Moreover, Smad3-dependent increased activation of Hif-1 α will further lead to the activation of several other oncogenes, especially VEGF which is a key activator of angiogenesis [88]. Activation

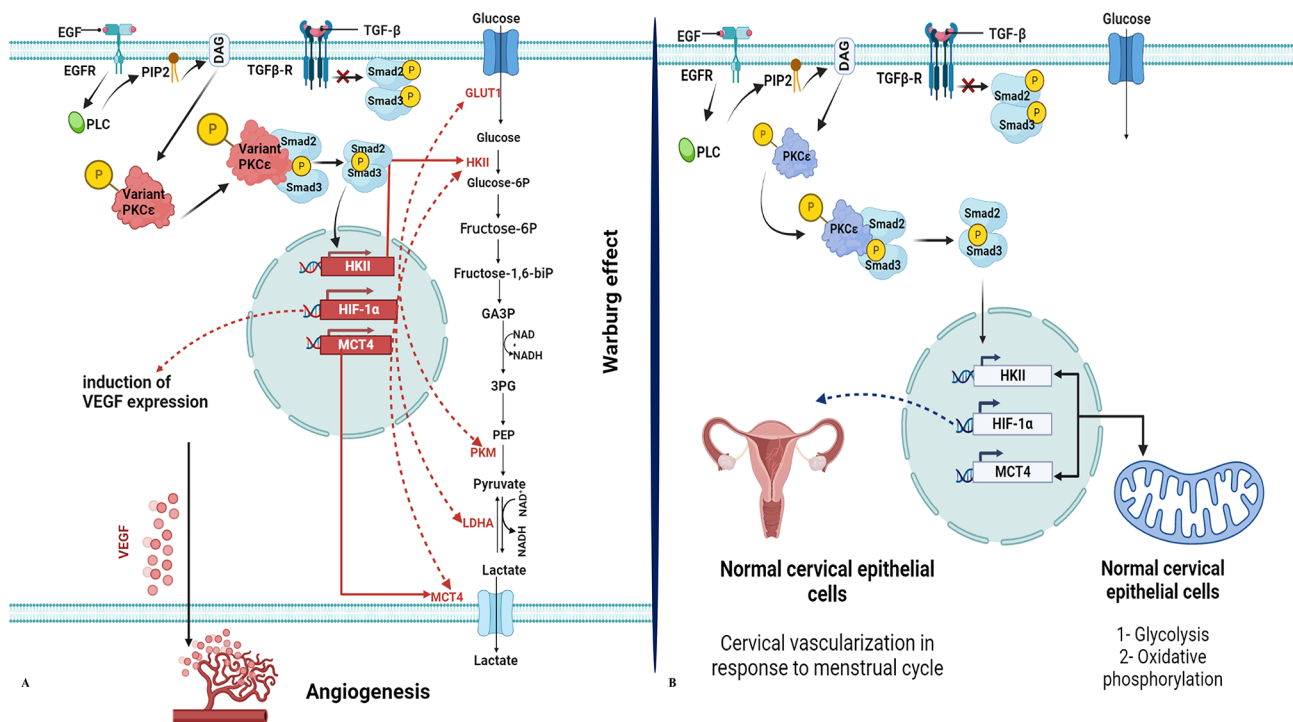


Fig. 5 Effect of non-synonymous variants on molecular interplay between PKC ϵ and Smad3 **(a)** Schematic illustration for crosstalk between variant PKC ϵ and Smad3 **(b)** Schematic illustration for crosstalk between variant PKC ϵ and Smad3

of both these carcinogenic processes result in rapid proliferation of cancer cells, enhanced tumour survival and poor prognosis in patients (Fig. 5a). Contrarily, when wildtype PKC ϵ forms a complex with Smad3, it does not bind with a high affinity and the binding energy between the native PKC ϵ -Smad3 complex is substantially low. This will result in higher activation energy to accomplish the phosphorylation reaction. Hence, the PKC ϵ induced activation of Smad3 will not increase aberrantly and typical levels of activated Smad3 will localize to the nucleus to activate gene transcription in a normal fashion, which is essential for the regulation of biological processes in the normal cervix-uterine environment including vascularization of cervix-uteri after menstruation and normal biochemical processes which are essential for the generation of energy by cervix epithelial cells (Fig. 5b).

Conclusion

Genotype analysis of PRKCE non-synonymous variants was performed which indicated that the altered genotype-AA of variant E14K with P-value 0.0001, and altered genotype-CC of variant D39H with P-value 0.045, were associated significantly with the risk of the development of cervical cancer. The wet lab analysis also revealed that the variant genotype-AA of E14K is associated with the risk of advanced stage cervical cancer with a P-value of 0.05 and Odds ratio of 7.538. The *In silico* molecular docking and interaction dynamics analysis of

wildtype and variants PKC ϵ suggested the potential role of variants in altered structural stability as well as protein-protein interactions that can have oncogenic implications and can serve as potential biomarkers for cancer. Therefore, these variants hold the potential to be further explored via *in vitro* and *in vivo* experimentations. Lastly, functional studies of these variants on large cohort-size should be performed to attain in-depth information regarding the molecular functions of these, so that these variants could be used as new therapeutic targets for the treatment of cervical cancer as well as novel biomarkers for better prognosis and early diagnosis of cervical cancer.

Acknowledgements

The authors extended their appreciation to the Researchers Supporting project number (RSP2023R502), King Saud University, Riyadh Saudi Arabia, for funding this project.

Authors' contributions

Conceptualization, S.Z, K.K. K.S, A.H, N.M.A and Y.B.; methodology, M.S., HA, SZ, K.K., and N.M.A.; experimentation, A.H., K.K.; validation .S.R; formal analysis, A.H., K.S, T.A., A.A., J.H.T. and K.K.; investigation, S.Z., J.H.T., K.K. and K.S.; resources, M.S., S.R.; data curation, N.M.A. S.R.; writing—original draft preparation, J.H.T.; writing—review and editing, K.K., S.R, T.A, J.H.T., Y.B. and M.S.; visualization, K.S., A.A., M.S. and N.M.A.; supervision, M.S.; project administration, M.S.; funding acquisition, S.R., J.H.T. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding

The authors are grateful to the Researchers Supporting Project number (RSP2023R502) king Saud University, Riyadh, Saudi Arabia for funding this project.

Data Availability

All the used and/or analyzed data during the current study are available from the corresponding author upon reasonable request.

Declarations**Ethics approval and consent to participate**

Ethical approval from the institutional review board (IRB No. 10-2021-01/01) was obtained from the parent department Atta-ur-Rahman School of Applied Biosciences (ASAB) of the National University of Sciences and Technology (see online source 1). Written and oral informed consent was attained from each participant before the collection of blood samples (see online source 2). All methods were carried out under relevant guidelines and regulations.

Consent for publication

Not relevant.

Competing interests

The authors affirm that they have no competing interests.

Received: 7 June 2023 / Accepted: 29 July 2023

Published online: 04 September 2023

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin*. 2021;71(3):209–49.
- Organization WH. WHO framework for strengthening and scaling-up of services for the management of invasive cervical cancer. 2020.
- Brusselselaers N, Shrestha S, Van De Wijgert J, Verstraelen H. Vaginal dysbiosis and the risk of human papillomavirus and cervical cancer: systematic review and meta-analysis. *Am J Obstet Gynecol*. 2019;221(1):9–18. e18.
- Bowden SJ, Bodinier B, Kalliala I, Zuber V, Vuckovic D, Doulgeraki T, Whitaker MD, Wielscher M, Cartwright R, Tsilidis KK. Genetic variation in cervical pre-invasive and invasive disease: a genome-wide association study. *Lancet Oncol*. 2021;22(4):548–57.
- Yang J, Yan Z, Wang Y, Xu J, Li R, Li C, Liu S, Shi L, Yao Y. Association study of relationships of polymorphisms in the miR-21, miR-26b, miR-221/222 and miR-126 genes with cervical intraepithelial neoplasia and cervical cancer. *BMC Cancer*. 2021;21(1):1–9.
- Muhammad SB, Hassan F, Bhowmik KK, Millat MS, Sarwar MS, Aziz MA, Barek MA, Uddin MS, Ferdous M, Islam MS. Detection of association of IL1 β , IL4R, and IL6 gene polymorphisms with cervical cancer in the bangladeshi women by tetra-primer ARMS-PCR method. *Int Immunopharmacol*. 2021;90:107131.
- Sharma P, Pattanshetty SM. A study on risk factors of cervical cancer among patients attending a tertiary care hospital: a case-control study. *Clin Epidemiol Global Health*. 2018;6(2):83–7.
- Tovell H, Newton AC. PHLPPing the balance: restoration of protein kinase C in cancer. *Biochem J*. 2021;478(2):341–55.
- Newton AC. Protein kinase C: poised to signal. *Am J Physiology-Endocrinology Metabolism* 2010.
- Pan Q, Bao LW, Kleer CG, Sabel MS, Griffith KA, Teknos TN, Merajver SD. Protein kinase C ϵ is a predictive biomarker of aggressive breast cancer and a validated target for RNA interference anticancer therapy. *Cancer Res*. 2005;65(18):8366–71.
- Isakov N. Protein kinase C (PKC) isoforms in cancer, tumor promotion and tumor suppression. *Seminars in cancer biology*: 2018: Elsevier; 2018:36–52.
- Newton AC. Protein kinase C: perfectly balanced. *Crit Rev Biochem Mol Biol*. 2018;53(2):208–30.
- Basu A. Regulation of autophagy by protein kinase C- ϵ in breast cancer cells. *Int J Mol Sci*. 2020;21(12):4247.
- Jain K, Basu A. Protein kinase C- ϵ promotes EMT in breast cancer. *Breast cancer: basic and clinical research*. 2014;8:BCBCR. S13640.
- Sadeghi MM, Salama MF, Hannun YA. Protein kinase C as a therapeutic target in Non-Small Cell Lung Cancer. *Int J Mol Sci*. 2021;22(11):5527.
- Xu W, Zeng F, Li S, Li G, Lai X, Wang QJ, Deng F. Crosstalk of protein kinase C ϵ with Smad2/3 promotes tumor cell proliferation in prostate cancer cells by enhancing aerobic glycolysis. *Cell Mol Life Sci*. 2018;75(24):4583–98.
- Guo Q-Z, Li L, Bo Q-M, Chen L, Sun L, Shi H-T. Expressions of PKC α , PKC ϵ and PKC δ in cervical cancer and their relationships with chemosensitivity. *J Hainan Med Univ*. 2019;25(2):43–8.
- Yu L, Lv J, Jin L, Ding S, Ma X, Wang J, Zhu X. Over-expression of protein kinase C isoforms (α , δ , θ and ζ) in squamous cervical cancer. *Neoplasma*. 2011;58(6):491–8.
- Wang Q, Xu H, Zhao X. Baicalin inhibits human cervical cancer cells by suppressing protein kinase C/signal transducer and activator of transcription (PKC/STAT3) signaling pathway. *Medical science monitor: international medical journal of experimental and clinical research* 2018;24:1955.
- El-Fattah A, Ahmed A, Sadik NAH, Shaker OG, Mohamed Kamal A. Single nucleotide polymorphism in SMAD7 and CH13L1 and colorectal cancer risk. *Mediators of inflammation* 2018.
- Deng N, Zhou H, Fan H, Yuan Y. Single nucleotide polymorphisms and cancer susceptibility. *Oncotarget*. 2017;8(66):110635.
- Gao L-B, Rao L, Wang Y-Y, Liang W-B, Li C, Xue H, Zhou B, Sun H, Li Y, Lv M-L. The association of interleukin-16 polymorphisms with IL-16 serum levels and risk of colorectal and gastric cancer. *Carcinogenesis*. 2009;30(2):295–9.
- Li H, Guo J, Cheng G, Wei Y, Liu S, Qi Y, Wang G, Xiao R, Qi W, Qiu W. Identification and validation of SNP-containing genes with prognostic value in gastric cancer via integrated bioinformatics analysis. *Front Oncol*. 2021;11:1434.
- Liu G-C, Zhou Y-F, Su X-C, Zhang J. Interaction between TP53 and XRCC1 increases susceptibility to cervical cancer development: a case control study. *BMC Cancer*. 2019;19:1–9.
- Guo C, Wen L, Song J-K, Zeng W-J, Dan C, Niu Y-M, Shen M. Significant association between interleukin-10 gene polymorphisms and cervical cancer risk: a meta-analysis. *Oncotarget*. 2018;9(15):12365.
- He S, Li Q, Huang Q, Cheng J. Targeting protein kinase C for cancer therapy. *Cancers*. 2022;14(5):1104.
- Kubo M, Hata J, Ninomiya T, Matsuda K, Yonemoto K, Nakano T, Matsu-shita T, Yamazaki K, Ohnishi Y, Saito S. A nonsynonymous SNP in PRKCH (protein kinase C η) increases the risk of cerebral infarction. *Nat Genet*. 2007;39(2):212–7.
- Sultana N, Rahman M, Myti S, Islam J, Mustafa MG, Nag K. A novel knowledge-derived data potentiating method revealed unique liver cancer-associated genetic variants. *Hum Genomics*. 2019;13(1):1–10.
- Lin H-Y, Wang X, Tseng T-S, Kao Y-H, Fang Z, Molina PE, Cheng C-H, Berglund AE, Eeles RA, Muir KR. Alcohol intake and Alcohol-SNP interactions Associated with prostate Cancer aggressiveness. *J Clin Med*. 2021;10(3):553.
- Li Q, Gu C, Zhu Y, Wang M, Yang Y, Wang J, Jin L, Zhu ML, Shi TY, He J. Two novel PRKCI polymorphisms and prostate cancer risk in an Eastern Chinese Han population. *Mol Carcinog*. 2015;54(8):632–41.
- Shah H, Khan K, Khan N, Badshah Y, Ashraf NM, Shabbir M. Impact of deleterious missense PRKCI variants on structural and functional dynamics of protein. *Sci Rep*. 2022;12(1):1–17.
- Khan K, Shah H, Rehman A, Badshah Y, Ashraf NM, Shabbir M. Influence of PRKCE non-synonymous variants on protein dynamics and functionality. *Human Molecular Genetics*; 2022.
- Rehman A, Shabbir M, Badshah Y, Khan K, Trembley JH, Ashraf NM, Afsar T, Almajwal A, Alruwaili NW, Alshamari A. Elucidating the role of missense SNP of protein kinase C epsilon in HCV-induced hepatocellular carcinoma. *BMC Cancer*. 2023;23(1):1–7.
- Ashok G, Ramaiah S. A critical review of datasets and computational suites for improving cancer theranostics and biomarker discovery. *Med Oncol*. 2022;39(12):206.
- Ashok G, Miryala SK, Saju MT, Anbarasu A, Ramaiah S. FN1 encoding fibronectin as a pivotal signaling gene for therapeutic intervention against pancreatic cancer. *Mol Genet Genomics*. 2022;297(6):1565–80.
- Roel-Touris J, Don CG, Honorato V, Rodrigues R, Bonvin JP. Less is more: coarse-grained integrative modeling of large biomolecular assemblies with HADDOCK. *J Chem Theory Comput*. 2019;15(11):6358–67.
- Yakymovych I, Dijke Pt, Heldin CH, Souchelnytskyi S. Regulation of smad signaling by protein kinase C. *FASEB J*. 2001;15(3):553–5.
- Luo L, Wang M, Li X, Luo C, Tan S, Yin S, Liu L, Zhu X. A novel mechanism by which ACTA2-AS1 promotes cervical cancer progression: acting as a ceRNA of mir-143-3p to regulate SMAD3 expression. *Cancer Cell Int*. 2020;20(1):1–13.
- Fan Q, Qiu M-T, Zhu Z, Zhou J-H, Chen L, Zhou Y, Gu W, Wang LH, Li Z-N, Xu Y. Twist induces epithelial-mesenchymal transition in cervical

- carcinogenesis by regulating the TGF- β /Smad3 signaling pathway. *Oncol Rep.* 2015;34(4):1787–94.
40. Ghatak S, Muthukumar RB, Nachimuthu SK. A simple method of genomic DNA extraction from human samples for PCR-RFLP analysis. *J Biomol techniques: JBT.* 2013;24(4):224–31.
41. Collins A, Ke X. Primer1: primer design web service for tetra-primer ARMS-PCR. *Open Bioinf J* 2012;6(1).
42. Yu B, Zhang C. In silico PCR analysis. In *Silico Tools for Gene Discovery*. edn.: Springer; 2011:91–107.
43. Langdon WB, Petke J, Lorenz R. Evolving better RNAfold structure prediction. In: *European Conference on Genetic Programming: 2018*. Springer; 2018:220–236.
44. Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. ViennaRNA Package 2.0. *Algorithms for molecular biology.* 2011;6(1):1–14.
45. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: protein structure and function prediction. *Nat Methods.* 2015;12(1):7–8.
46. Laskowski R, MacArthur M, Thornton J. Procheck: validation of protein-structure coordinates. 2006.
47. Xu W, Zhang Z, Zou K, Cheng Y, Yang M, Chen H, Wang H, Zhao J, Chen P, He L. MiR-1 suppresses tumor cell proliferation in colorectal cancer by inhibition of Smad3-mediated tumor glycolysis. *Cell Death Dis.* 2017;8(5):e2761–1.
48. Anuar NFSK, Wahab RA, Huyop F, Halim KBA, Hamid AAA. In silico mutation on a mutant lipase from *Acinetobacter haemolyticus* towards enhancing alkaline stability. *J Biomol Struct Dynamics.* 2020;38(15):4493–507.
49. 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. *SoftwareX.* 2015;1:19–25.
50. Kulig W, Pasenkiewicz-Gierula M, Róg T. Topologies, structures and parameter files for lipid simulations in GROMACS with the OPLS-aa force field: DPPC, POPC, DOPC, PEPC, and cholesterol. *Data in brief.* 2015;5:333–6.
51. Laskowski RA, Swindells MB. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. In: *ACS Publications*; 2011.
52. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph.* 1996;14(1):33–8.
53. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
54. Franz M, Rodriguez H, Lopes C, Zuberi K, Montojo J, Bader GD, Morris Q. GeneMANIA update 2018. *Nucleic Acids Res.* 2018;46(W1):W60–4.
55. Mering Cv, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* 2003;31(1):258–61.
56. Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* 2003;4(9):1–11.
57. Bakshi D, Nagpal A, Sharma V, Sharma I, Shah R, Sharma B, Bhat A, Verma S, Bhat G, Abrol D. MassARRAY-based single nucleotide polymorphism analysis in breast cancer of north indian population. *BMC Cancer.* 2020;20(1):1–8.
58. Parate S, Rampogu S, Lee G, Hong JC, Lee KW. Exploring the binding Interaction of Raf kinase inhibitory protein with the N-Terminal of C-Raf through Molecular Docking and Molecular Dynamics Simulation. *Front Mol Biosci.* 2021;8:496.
59. Galeazzi R, Laudadio E, Falconi E, Massaccesi L, Ercolani L, Mobbili G, Minnelli C, Scirè A, Cianfruglia L, Armeni T. Protein–protein interactions of human glyoxalase II: findings of a reliable docking protocol. *Org Biomol Chem.* 2018;16(28):5167–77.
60. Laskowski RA, Jablonska J, Pravda L, Vařeková RS, Thornton JM. PDBsum: structural summaries of PDB entries. *Protein Sci.* 2018;27(1):129–34.
61. Gopalakrishnan C, Al-Subaie AM, Yeh HY, Tayubi IA, Kamaraj B. Prioritization of SNPs in y + LAT-1 culpable of lysinuric protein intolerance and their mutational impacts using protein–protein docking and molecular dynamics simulation studies. *J Cell Biochem.* 2019;120(10):18496–508.
62. Chen Y, Xiao Z, Wang HEM, WU Z. Y: Study on the relationship between single nucleotide polymorphisms of PRKCG gene (rs2547362 and rs3745406) and susceptibility to osteosarcoma in a chinese Han population. *Chin J Orthop* 2013;1223–30.
63. Newton AC. Protein kinase C as a tumor suppressor. *Seminars in cancer biology: 2018*. Elsevier; 2018: 18–26.
64. Ma RC, Tam CH, Wang Y, Luk AO, Ho C, Yang X, Lam V, Chan AW, Ho JS, Chow C-C. Genetic variants of the protein kinase C- β 1 gene and development of end-stage renal disease in patients with type 2 diabetes. *Jama.* 2010;304(8):881–9.
65. Riera P, Virgili AC, Salazar J, Sebío A, Tobeña M, Sullivan I, Páez D. Genetic variants in the VEGF pathway as prognostic factors in stages II and III colon cancer. *Pharmacogenomics J.* 2018;18(4):556–64.
66. Gorin MA, Pan Q. Protein kinase C ϵ : an oncogene and emerging tumor biomarker. *Mol Cancer.* 2009;8(1):1–8.
67. Hashemi M, Shahkar G, Simforoosh N, Basiri A, Ziaee S, Narouie B, Taheri M. Association of polymorphisms in PRKCI gene and risk of prostate cancer in a sample of Iranian Population. *Cell Mol Biol.* 2015;61(5):16–21.
68. Tan H. On the protective effects of gene SNPs against human cancer. *EBio-Medicine.* 2018;33:4–5.
69. Wang N, Wang S, Zhang Q, Lu Y, Wei H, Li W, Zhang S, Yin D, Ou Y. Association of p21 SNPs and risk of cervical cancer among chinese women. *BMC Cancer.* 2012;12(1):1–6.
70. Campa D, Hüsing A, Stein A, Dostal L, Boeving H, Pischon T, Tjønneland A, Roswall N, Overvad K, Østergaard JN. Genetic variability of the mTOR pathway and prostate cancer risk in the european prospective investigation on Cancer (EPIC). *PLoS One.* 2011;6(2):e16914.
71. Au WW, Abdou-Salama S, Sierra-Torres CH, Al-Hendy A. Environmental risk factors for prevention and molecular intervention of cervical cancer. *Int J Hyg Environ Health.* 2007;210(6):671–8.
72. Vargas-Torres SL, Portari EA, Silva AL, Klumb EM, da Rocha Guillobel HC, de Camargo MJ, Santos-Rebouças CB, Russomano FB, Macedo JMB. Roles of CDKN1A gene polymorphisms (rs1801270 and rs1059234) in the development of cervical neoplasia. *Tumor Biology.* 2016;37:10469–78.
73. Salimi S, Hajizadeh A, Yaghmaei M, Rezaie S, Shahrakypour M, Teimoori B, Parache N, Naghavi A, Mokhtari M. The effects of p21 gene C98A polymorphism on development of uterine leiomyoma in southeast iranian women. *Tumor Biology.* 2016;37:12497–502.
74. Li S, Hua Y, Jin J, Wang H, Du M, Zhu L, Chu H, Zhang Z, Wang M. Association of genetic variants in lncRNA H19 with risk of colorectal cancer in a chinese population. *Oncotarget.* 2016;7(18):25470.
75. Faure G, Ogurtsov AY, Shabalina SA, Koonin EV. Role of mRNA structure in the control of protein folding. *Nucleic Acids Res.* 2016;44(22):10898–911.
76. Nagaraj NS, Singh OV, Merchant NB. Proteomics: a strategy to understand the novel targets in protein misfolding and cancer therapy. *Expert Rev Proteomics.* 2010;7(4):613–23.
77. Southey MC, Park DJ, Nguyen-Dumont T, Campbell I, Thompson E, Trainer AH, Chenevix-Trench G, Simard J, Dumont M, Soucy P. COMPLEXO: identifying the missing heritability of breast cancer via next generation collaboration. *Breast Cancer Res.* 2013;15(3):1–2.
78. Simoes-Correia J, Figueiredo J, Lopes R, Stricher F, Oliveira C, Serrano L, Seruca R. E-cadherin destabilization accounts for the pathogenicity of missense mutations in hereditary diffuse gastric cancer. *PLoS ONE.* 2012;7(3):e33783.
79. Falsone SF, Leptihn S, Osterauer A, Haslbeck M, Buchner J. Oncogenic mutations reduce the stability of SRC kinase. *J Mol Biol.* 2004;344(1):281–91.
80. Nishi H, Tyagi M, Teng S, Shoemaker BA, Hashimoto K, Alexov E, Wuchty S, Panchenko AR. Cancer missense mutations alter binding properties of proteins and their interaction networks. *PLoS One.* 2013;8(6):e66273.
81. Kamato D, Do BH, Osman N, Ross BP, Mohamed R, Xu S, Little PJ. Smad linker region phosphorylation is a signalling pathway in its own right and not only a modulator of canonical TGF- β signalling. *Cell Mol Life Sci.* 2020;77(2):243–51.
82. Debroy R, Ramaiah S. Translational protein RpsE as an alternative target for novel nucleoside analogues to treat MDR Enterobacter cloacae ATCC 13047: network analysis and molecular dynamics study. *World J Microbiol Biotechnol.* 2023;39(7):187.
83. Sang P, Hu W, Ye Y-J, Li L-H, Zhang C, Xie Y-H, Meng Z-H. In silico screening, molecular docking, and molecular dynamics studies of SNP-derived human P5CR mutants. *J Biomol Struct Dynamics.* 2017;35(11):2441–53.
84. Pace CN, Fu H, Lee Fryar K, Landua J, Trevino SR, Schell D, Thurkill RL, Imura S, Scholtz JM, Gajiwala K. Contribution of hydrogen bonds to protein stability. *Protein Sci.* 2014;23(5):652–61.
85. Carlsson A-C, Scholfield MR, Rowe RK, Ford MC, Alexander AT, Mehl RA, Ho PS. Increasing enzyme stability and activity through hydrogen bond-enhanced halogen bonds. *Biochemistry.* 2018;57(28):4135–47.
86. Perrin CL. Are short, low-barrier hydrogen bonds unusually strong? *Acc Chem Res.* 2010;43(12):1550–7.
87. Ooshima A, Park J, Kim SJ. Phosphorylation status at Smad3 linker region modulates transforming growth factor- β -induced epithelial-mesenchymal transition and cancer progression. *Cancer Sci.* 2019;110(2):481–8.
88. Li J, Li SX, Gao XH, Zhao LF, Du J, Wang TY, Wang L, Zhang J, Wang HY, Dong R. HIF1A and VEGF regulate each other by competing endogenous RNA

mechanism and involve in the pathogenesis of peritoneal fibrosis. *Pathology-Research and Practice*. 2019;215(4):644–52.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.