ORIGINAL ARTICLE

XPC as breast cancer susceptibility gene: evidence from genetic profiling, statistical inferences and protein structural analysis

Saima Shakil Malik[1,4] · Ayisha Zia[2] · Sajid Rashid[2] · Sumaira Mubarik[3] · Nosheen Masood[1] · Mubashar Hussain[4] · Azra Yasmin[1] · Razia Bano[5]

Abstract

Background Gene polymorphisms that affect nucleotide excision repair (NER) pathway may link with higher susceptibility of breast cancer (BC); however, the significance of these associations may vary conferring to the individual ethnicity. Xeroderma pigmentosum complementation gene (XPC) plays a substantial role in recognizing damaged DNA during NER process.

Objective and methods To estimate the relationship among XPC polymorphisms and breast cancer (BC) risk, we carried out a case–control-association study with 493 BC cases and 387 controls using TETRA–ARMS-PCR. Distributional differences of clinical features, demographic factors and XPC polymorphisms among BC cases and controls were examined by conditional logistic regression model. Kaplan–Meier test was applied to predict survival distributions and protein structure was predicted using computational tools.

Results Obesity, consanguinity, positive marital status and BC family history were associated ($P \leq 0.01$) with higher BC risk. Genotyping revealed significant involvement ($P \leq 0.01$) of two XPC polymorphisms rs2228001–A > C (OR = 3.8; CI 1.9–7.6) and rs2733532–C > T (OR = 2.6; CI 1.4–5.03) in BC development, asserting them potential risk factors for increased BC incidence. However, no association ($P > 0.05$) was detected for overall or progression free survival for both XPC polymorphisms possibly due to shorter follow-up time (45 months). As compared to normal XPC structure, pronounced conformational changes have been observed in the C-terminus of XPC^{Q939K}, bearing rs2228001–A > C substitution. In XPC^{Q939K}, two additional α-helices were observed at A292-E297 and Y252-R286, while L623-M630 and L649-L653 helices were converted into loop conformation.

Conclusion In conclusion, both XPC polymorphisms confer significant association with increased BC risk. rs2228001 substitution may change the structural and functional preferences of XPC C-terminus, while rs2733532 may have regulatory role thereby leading to potential BC risk.

Keywords XPC polymorphisms · Breast cancer · Protein structure · Overall survival · Progression free survival

Ayisha Zia, Sajid Rashid and Sumaira Mubarik contributed equally to this work.

✉ Saima Shakil Malik
saimamalik25@yahoo.com

1 Fatima Jinnah Women University, The Mall Rawalpindi, Rawalpindi, Pakistan

2 National Centre for Bioinformatics, Quaid-i-Azam University, Islamabad, Pakistan

3 Department of Epidemiology and Biostatistics, School of Health Sciences, Wuhan University, Wuhan, China

4 Department of Zoology, University of Gujrat, Gujrat, Pakistan

5 Breast Clinic, Combined Military Hospital, Rawalpindi, Pakistan

Introduction

Breast cancer (BC) is an intricate multifaceted disease with undistinguishable aetiology. BC potential risk can be determined by the extent of DNA damage and genetic instability caused by various environmental factors [1]. However, we are blessed by birth with a very efficient system to guard our

genome from several type of DNA damages and repair those damages if anyone occurred with the help of transcription—coupled repair (TCR), nucleotide excision repair (NER), base excision repair (BER), mismatch repair (MMR), and double—strand DNA break repair systems [2].

NER is meant to repair DNA damages induced by ultraviolent (UV) radiations, intrastranded DNA cross-links, products of organic combustion, oxidative stress and heavy metals. NER proteins involved in the repair of damages caused by aforementioned sources and to maintain genomic integrity by carcinogenesis prevention includes XPA, ERCC1, XPB, XPC, XPD, XPF, XPG and DDB1 [3]. NER process encompasses of various definite steps such as recognition of DNA damage, demarcation of damaged DNA, incision of damaged DNA, repair process (patch synthesis) followed by DNA ligation [4]. Xeroderma pigmentosum complementation gene (XPC) is meant to form a complex with HR23B and responsible for recognizing DNA damages predominantly cyclo-butane-pyrimidine dimers formed as a result of exposure to UV radiations. This complex plays major role in recognizing damages by binding with DNA lesion and lead to its unwinding, offering place for XPA so, that it can bind and unwound DNA to initiate the whole NER process. Most of the XPC genetic alterations (95%) are frameshift, deletion, splice site or non-sense mutations and produces truncated proteins [5, 6]. Numerous NER polymorphisms have been reported in literature specifically with reference to their involvement in carcinogenesis [4].

Researchers have demonstrated low NER levels in the peripheral blood lymphocytes of BC cases and their relatives [7]. Besides, some studies have shown association of NER gene polymorphisms with BC risk in different ethnicities [1, 8–10]. But the results have been unpredictable and inconsistent due to smaller sample size, lack of clinical details and follow up together with different sociocultural and genetic backgrounds [11]. Among the Pakistani population, we have previously reported the association of XPG polymorphisms with breast cancer and survival [10] but to date, no systematic examination is available on the susceptibility of XPC or other NER genes in breast cancer. XPC rs2228001 A > C polymorphism at exon 15 results in Gln substitution into Lys at 939 position, whereas rs2733532 lies in the intronic region. rs2228001 is one of the frequently studied XPC polymorphism with highly reported global minor allele frequency (MAF) i.e., 0.31 whereas rs2733532 have second highest MAF, 0.30 [12]. Therefore, current study was designed to offer a more comprehensive understanding of association between XPC polymorphisms (rs2228001–A > C & rs2733532–C > T) and BC carcinogenesis and to predict possible conformational changes in protein with different computational tools. Likewise, substantial survival hindrance as a result of XPC polymorphisms among BC cases in overall and progression free survival was explored.

Methodology

Study subjects and ethical considerations

For case–control study, from September 2014 to May 2018 ethical approval was obtained from "IRB—Institutional Review Board and Ethical Committees of Fatima Jinnah Women University, Rawalpindi and hospitals including Benazir Bhutto Shaheed Hospital, District Headquarter Hospital, Holy Family Hospital, Combined Military Hospital, Rawalpindi and Shifa International Hospital Islamabad". Study sample size was calculated with sample size calculator provided by World Health Organization (WHO) and manually confirmed with formula $\{n = Z_{\frac{\alpha}{2}}^2 \frac{PQ}{(MOE)^2}\}$. Using disease prevalence of 15% [13] and 5% margin of error (MOE) calculated sample size turns out to be 195, where $Z_{\frac{\alpha}{2}}$, MOE and p is statistical constant, relative precision and disease prevalence respectively. Study parameters were designated by following the recommendations of tumor marker prognostic studies [14]. Blood samples were collected from 493 histopathologically diagnosed BC cases and age-matched 387 controls of same geographical region following written informed consent. Clinical and demographic details were taken in a detailed questionnaire with the assistance of patient, their attendees, medical records and respective oncologist.

XPC genotyping

XPC genotyping was performed as previously described [10] via Tetra ARMS–PCR "tetra-primer amplification refractory mutation system–polymerase chain reaction". Primers were designed with a program established by Ye et al. with default settings [15]. Reaction conditions includes initial denaturation at 94 °C (5 min), followed by 30 cycles comprising of 94 °C (45 s), 74 °C (1 min), and 72 °C (45 s) and final extension at 72 °C (10 min). Results were validated by repeat analysis of randomly selected twenty-five samples each from cases and controls. PCR products were recorded with gel documentation system after resolving on agarose gel through electrophoresis.

Data set

SNPs (rs2228001 and rs2733532) were witnessed through dbSNP (https://www.ncbi.nlm.nih.gov/projects/SNP) [16] in the XPC gene. SNP rs2733532 C > T lies in the intronic region, while rs2228001 A > C polymorphism was evident in the exon 15. The primary protein sequence (ID: Q01831) was retrieved through UniProtKB/Swiss-Prot database (http://www.uniprot.org). 3D structures of XPC^WT and

XPC^{Q939K} were modelled using Phyre2 [17] and I-TASSER [18]. Energy minimization was carried out by GROMACS 5.1.4 through Amber force field [19]. The predicted structures were visualized by UCSF Chimera ver. 1.11.2, which is an extensive program for interactive visualization and analysis of molecular structures and related data (Fig. 2). Structure validation was carried out by RAMPAGE [20] and ProQ [21] tools, followed by refinement using WinCoot [22]. Structure joining, editing and analyses were carried out by UCSF Chimera ver. 1.11.2 [23] as shown in Fig. 3.

Statistical analysis

For the case–control study, goodness of fit χ^2 was tested by Hardy–Weinberg equilibrium. Distributional differences of several clinical features, demographic factors and XPC polymorphisms between BC cases and healthy controls were examined by Chi square test. Risk of XPC polymorphisms (rs2228001–A > C & rs2733532–C > T) and breast carcinogenesis was assessed by odds ratios (OR) and 95% confidence intervals (CIs) using a conditional logistic regression model.

A total of 347 BC cases participated in survival analysis due to their availability for follow-up till the end of the study period. All the BC cases were followed-up after every 6 months via telephonic calls or hospital visits to maintain a record of disease progression, patient's health status, disease free survival and deaths. Overall survival (OS) was defined "as the time of disease diagnosis (randomization) to death due to any cause whereas progression free survival (PFS) was the time of diagnosis to last clinical follow up (45 months for current study)". Kaplan–Meier test and log rank test were applied to predict survival distributions and differences among XPC polymorphisms and breast carcinogenesis respectively. Association of XPC polymorphisms with OS and PFS in BC was demonstrated by hazard ratio (HR) and 95% CI by univariate cox proportional hazards model. Homozygous normal genotype (wild type) was taken as reference for both XPC polymorphisms. P value ≤ 0.05 was considered as statistically significant. SPSS version 24.0 and MedCalc version 15.0 were used for respective statistical analysis.

Results

Demographic and clinical details

Mean age of BC cases and controls was 50.4 ± 14.1 years and 46.5 ± 14.9 years, respectively. No significant differences were observed in the age and menarche (OR = 0.88; CI 0.7–1.07) among BC cases and controls, whereas early menopause was evaluated as a risk factor (OR = 3.3;

CI 2.2–4.8) for BC development. Mean number of live births and mean ages at first pregnancy were 4.7 ± 1.8 years and 20.8 ± 3.4 years, respectively. Results showed that only 3% BC cases were nulliparous and 3.4% parous women had no lactating history. Obesity, consanguinity, positive marital status and BC family history was significantly associated ($P \leq 0.01$) with BC risk. In current study, most of BC cases were diagnosed with invasive ductal carcinoma (76%), whereas 24% cases had invasive lobular carcinoma illustrating that invasive ductal carcinoma was the most common BC type. None of the cases were reported with grade-I tumor, whereas 66% had grade-II tumor. Among 29% BC cases, tumor was metastasized to different body parts.

### XPC genotyping and clinical details

Genotyping revealed that rs2228001 (A > C) had 57 heterozygous (C/A) and 13 homozygous (A/A) polymorphisms and heterozygous (C/A) polymorphism was found to be significantly associated with BC risk as shown in Fig. 1a. XPC polymorphism, rs2733532 (C > T) had statistically significant involvement towards BC development (Fig. 1b) with 58 heterozygous (C/T) and 11 homozygous (T/T) alterations (Table 1). Additionally, most of the cases with XPC polymorphisms were married, had consanguinity and invasive ductal carcinoma with grade II tumor. Family history of cancer was found to be pivotal factor in possession of breast cancer and almost 42% BC cases with XPC polymorphisms had family history of cancer. A total of 69% cases
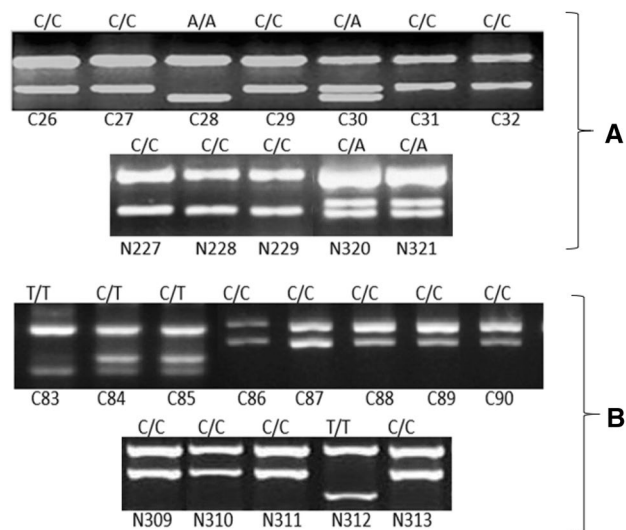
Fig. 1 Electropherogram of Tetra–ARMS PCR amplified product for **a** XPC rs2228001 (A > C) where; C/C homozygous wild type, C/A heterozygous mutant and A/A homozygous mutant. **b** XPC rs2733532 (C > T); C/C homozygous wild type, C/T heterozygous mutant and T/T homozygous mutant. C, Breast cancer; N, Control; numerical figures represents number of cases and controls

Table 1 Association between XPC variant and breast cancer risk

XPC polymorphisms	Cases (n = 493)		Control (n = 387)		OR (95% CI)	P value
	n	%	n	%		
rs2228001 (A > C)						
C/C	423	85.8	372	96.1	1.0 (Ref.)	< 0.0001
C/A	57	11.6	12	3.1	3.8 (1.89–7.66)	0.53
A/A	13	2.6	3	0.8	1.5 (0.40–5.67)	
rs2733532 (C > T)						
C/C	427	86.6	370	95.6	1.0 (Ref.)	0.005
C/T	58	11.8	14	3.6	2.6 (1.34–5.03)	0.17
T/T	11	1.3	3	0.8	2.8 (0.62–12.41)	

Fig. 2 Structure of XPC protein. **a** Domain organization in XPC. **b** Ribbon diagram of human XPC. Transglutaminase-like domain (TGD) is indicated in spring green color and three β-hairpin domains (BHD 1–3) are shown in purple, yellow and blue colors, respectively
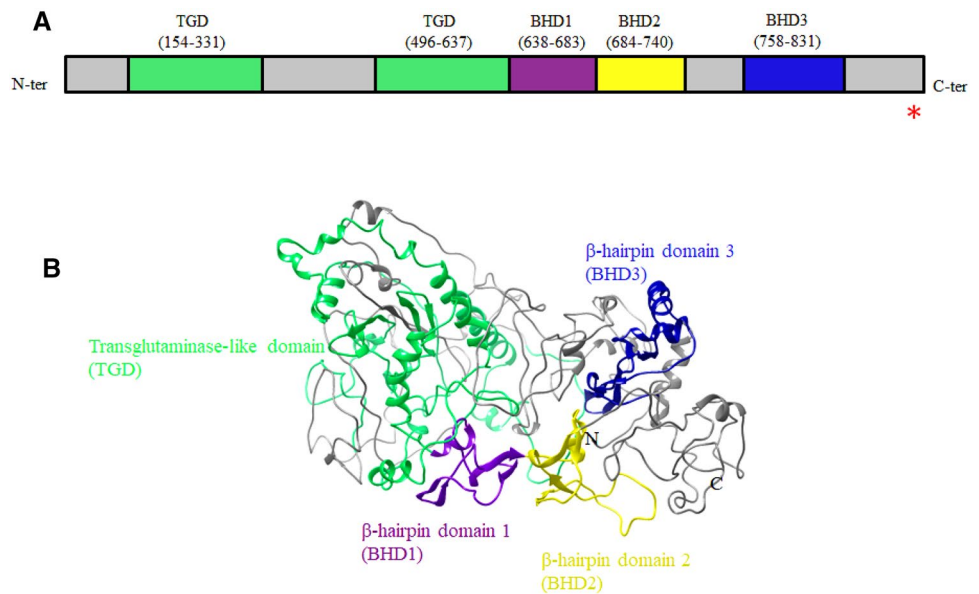
with rs2228001 polymorphism showed lymph node positivity whereas 34% had metastasis to different body organs whereas 76% and 32% cases were reported to have positive lymph nodes and metastasis with rs2733532 polymorphism. For, rs2733532 polymorphism, significant association was detected among age and delayed menopause ($P \leq 0.01$). Results showed statistically significant association between positive family history of cancer ($P = 0.02$) with higher tumor grade in BC cases with rs2733532 polymorphism. Therefore, results showed that XPC polymorphisms are linked with predisposition to BC and different lifestyle along with clinical factors contributes towards disease severity and progression.

Structural analysis

Human XPC comprises 4 domains named as Transglutaminase like domain, β-hairpin domain 1, β-hairpin domain 2 and β-hairpin domains 3, respectively (Fig. 2) [24]. Ramachandran analysis for XPC^{WT} and XPC^{Q939K}

Table 2 Structural evaluation of XPC^{WT} and XPC^{Q939K}

Proteins	RMSD (Å)	MolProbity	ProQ	
		Ramachandran favoured	LGscore	MaxSub
XPC^{WT}	0	96.80%	2.853	0.133
XPC^{Q939K}	2.1 Å	92.3%	2.163	0.129

indicated presence of 96.8% and 92.3% residues in sterically allowed region (fig S1) (Table 2). Through superimposition of XPC^{WT} and XPC^{Q939K} structures, an RMSD value of 2.1Å was detected, indicating significant change at structural level. Quite interestingly, by comparison to XPC^{WT}, predominant conformational changes were witnessed at the C-terminal region of XPC^{Q939K} (Fig. 3), particularly in α-helical pattern. Two additional α-helices (L845-K849 and E859-R864) were observed in XPC^{Q939K}, resulting in a shifting of T923-A930 helix. Similarly, another significant change was witnessed at the helical region of XPC^{Q939K}
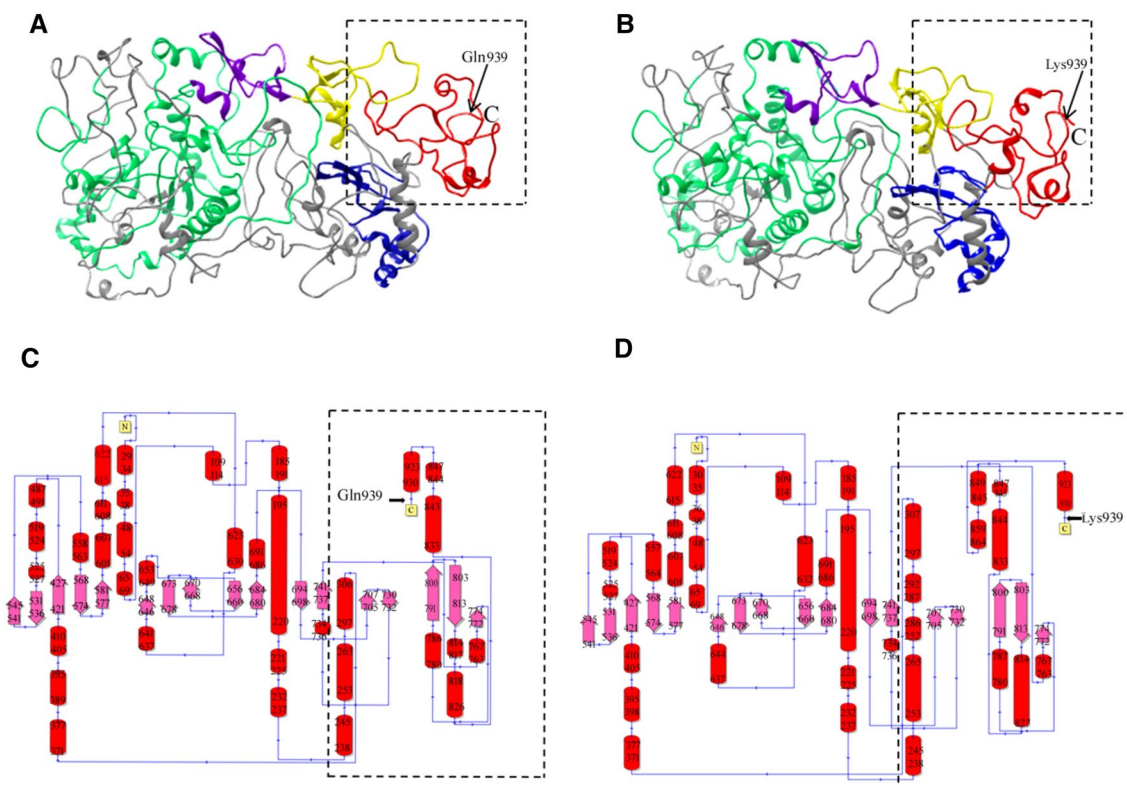
Fig. 3 Comparative 3D and 2D structure of XPC[WT] and XPC[Q939K]. Ribbon diagrams of **a** XPC[WT] and **b** XPC[Q939K]. 2D structures of **c** XPC[WT] and **d** XPC[Q939K]. The regions exhibiting conformational changes are shown in dotted box

encompassing P238-R307 residues at the vicinity of C-terminus and a helix H487-H491 position has been shifted into loop (Fig. 3).

## XPC genotyping, survival analysis and breast cancer prognosis

Relationship of XPC polymorphisms with OS and PFS was measured among BC cases with the help of Kaplan–Meier Survival curve (Fig. 4) and Cox proportional hazards model, respectively. Even though heterozygous genotypes of both polymorphisms (rs2733532 and rs2228001) were associated with BC risk, they exhibited no association ($P \geq 0.05$) with OS and PFS (Table 3). Though noticeable differences were observed among wild type and altered genotypes of rs2733532 with PFS, findings were not significant.

Discussion

Breast cancer is not only the most frequently identified malignancy amongst females, but also a leading cause for cancer-related deaths around the globe [25]. Demographic factors suggest that mean age of BC cases is 50.4 years and

ranges from 18 to 85 years, illustrating that BC incidence increases with age [26, 27]. Menarche and menopause are the hallmarks accountable for the onset and offset of ovarian and endocrine activity associated with reproduction. Current study demonstrated 3% rise in BC risk among females with early menopause, consistent with earlier reports [28–31]. In contrast to previous finding [32], where menarche role remains inconsistent, we did not observe any association of menarche with BC risk, rather consanguinity and familial cancer history were BC associated risk factors, as described elsewhere [10, 31, 33, 34]. The positive marital status had a substantial impact on BC risk signifying that married females tend to have better health than unmarried ones, conceivably due to reproductive and psychosocial factors as demonstrated by National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) 18 regions database [35].

Several studies presented contradictory results concerning the association of XPC polymorphisms and risk of cancer. Individuals having XPC polymorphisms were interrelated with 1000 folds higher skin carcinoma risk [36]. XPC and XPG cofactors stimulate OGG1, a DNA glycosylase activity and competent working of NTH1, whereas XPA supports in the repair of oxidized bases [37]. The current
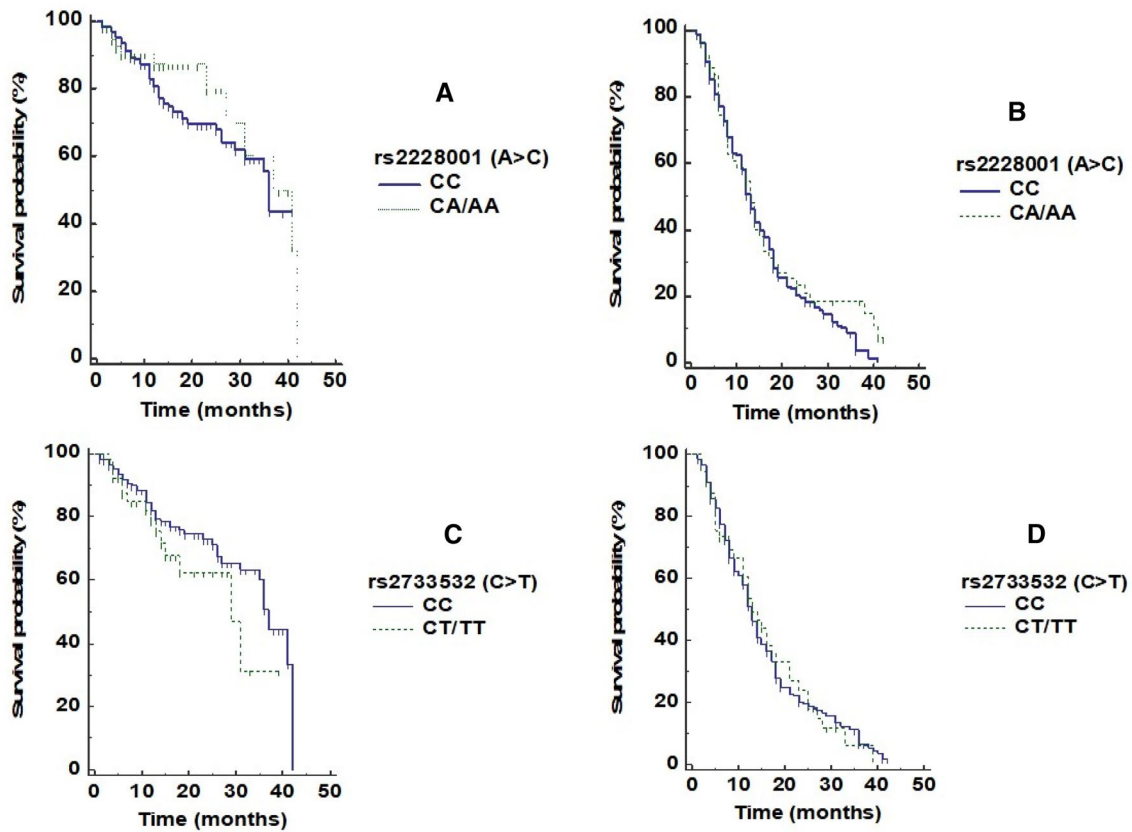
Fig. 4 Kaplan–Meier estimates representing association of XPC polymorphisms with OS and PFS in BC cases. **a**, **b** No association of rs2228001 polymorphism with PFS and OS among BC cases. **c**, **d** No association of rs2733532 polymorphism with PFS and OS among BC cases, respectively

case–control–association study presented a statistically significant association of two XPC polymorphisms (rs2228001 & rs2733532) with BC risk. XPC rs2228001 polymorphism results in the substitution glutamine into lysine at 939 position, while, another polymorphism, rs2733532 was present in the intronic region. These polymorphisms were selected due to their highest MAF values and lack of data with respect to their BC association in Pakistani population. Thus, XPC codon 939 polymorphic locus (rs2228001) may be considered as a potential marker for BC diagnosis. In a meta-analysis, no association has been detected between rs2228001 polymorphism and BC risk [38]. As long term follow-up is required for improved understandings of OS, PFS is a better option in such condition [39]. In the present study, OS and PFS analysis of 347 BC cases showed no association with both XPC polymorphisms. Though evident difference was found among wild-type and altered genotypes of rs2733532 with PFS, but outcomes were statistically non-significant that may be due to shorter follow-up time and limited study size. The visible differences in the outcomes of current and prior studies may possibly be due to diverse genetic backgrounds and sample population types.

Acquaintance to dissimilar socio-cultural backgrounds, environmental and life style factors in consort with varying study sample sizes may be the conceivable explanation of opposing results [11, 12].

Recent studies potentiate that XPC 939 substitution may be associated with higher frequency of p53 mutations and [40] Aflatoxin B1 (AFB1)-induced hepatocellular carcinoma [41], as XPC plays a major role in the reduction of AFB1-induced toxic effects. Intriguingly, in XPC^{Q939K}, presence of two additional α-helices at L845-K849 and E859-R864 positions induced conformational switching in the helix (T923-A930 AA) symmetry at the C-terminus. Another significant change was witnessed at the helical region of XPC^{Q939K} encompassing P238-R307 residues at the vicinity of C-terminus, that may contribute in its abnormal activity. XPC C-terminus is crucial for TFIIH recruitment, hHR23B and damaged DNA binding. XPC stable binding with NER protein hHR23B recognizes DNA damage sites, formed by exogenous carcinogens such as AFB1. Despite having normal NER activity, XPC^{Q939K} is associated with higher cancer susceptibility [42], and is linked with abnormal p53 degradation due to disruption in its interaction with Mdm2 ubiquitin

Table 3 Association between XPC variants and breast cancer prognosis using Kaplan Meier survival analysis and Cox proportional hazards model

SNPs	Genotype	cases (n=347)	PFS					OS				
			events (n=77)	Median (months)	Log-rank (P-value)	HR (95% CI)	P-val	events (n=270)	Median (months)	Log rank (P-value)	HR (95% CI)	P-val
rs2228001 (A>C)	CC	292	65	36	0.299	1.0 (Ref.)		227	13	0.2138	1.0 (Ref.)	
	CA+AA	55	12	37		1.35 (0.75–2.48)	0.72	43	13		1.21 (0.88–1.70)	0.42
rs2733532 (C>T)	CC	292	62	37	0.199	1.0 (Ref.)		230	13	0.98	1.0 (Ref.)	
	CT+TT	55	15	29		0.69 (0.34–1.24)	0.34	40	13		1.02 (0.70–1.42)	0.53

ligase [43]. These findings add substantial evidence in the non-NER activity of XPC^{Q939K}, possibly via impairing p53 proteolysis.

However, present study has some limitations. It was conducted in different hospitals of Rawalpindi and Islamabad but collecting samples from representative hospitals of all provinces could give more generalized results. Although individuals from all over the country visits these hospitals due to their better diagnosis and treatment modalities as they were situated in the capital of the country. Genotyping was carried out via ARMS-PCR while more accurate findings could be available by means of sequencing. Additionally, onset of clinical diseases is not the manifestation of disruptions in single or more genes. Moreover, genetic disturbances are entrenched in the entire genome and affected by various environmental factors as well. Hence, some other genes and polymorphisms could also involve in breast carcinogenesis.

In conclusion, present study reveals that XPC polymorphisms (rs2228001–A > C and rs2733532–C > T) are possible risk factors associated with increased BC incidence. XPC polymorphism, rs2228001 may change the structural and functional preferences of XPC C-terminus, while rs2733532 may have regulatory role thereby leading to potential BC risk. Additional studies with larger sample size and longer follow-up may provide better outcomes in future.

Compliance with ethical standards

Consent for publication The experiments were undertaken with the understanding and written consent of each subject, and that the study conforms with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Conflict of interest All authors declare that they have no conflict of interest.

References

1. He B-s, Xu T, Pan Y-q, Wang H-j, Cho WC, Lin K, Sun H-l, Gao T-y, Wang S-k. Nucleotide excision repair pathway gene polymorphisms are linked to breast cancer risk in a Chinese population. Oncotarget. 2016;7(51):84872.
2. Grand RJA, Reynolds JJ. DNA repair and replication: mechanisms and clinical significance. Boca Raton: CRC Press; 2018.
3. Pongsavee M, Wisuwan K. ERCC5 rs751402 polymorphism is the risk factor for sporadic breast cancer in Thailand. Int J Mol Epidemiol Genet. 2018;9(4):27.
4. Winczura A, Reynolds JJ. The repair of DNA single-strand breaks and DNA adducts: mechanisms and links to human disease. DNA repair and replication: mechanisms and clinical significance, 2018.
5. Lehmann J. Functional relevance of spontaneous alternative splice variants of xeroderma pigmentosum genes: Prognostic

marker for skin cancer risk and disease outcome?, Georg-August-Universität Göttingen, 2017.

6. Schäfer A. Clinical, functional, and genetic analysis of NER defective patients and characterization of five novel XPG mutations. Niedersächsische Staats-und Universitätsbibliothek Göttingen, 2012.

7. Latimer JJ, Johnson JM, Kelly CM, Miles TD, Beaudry-Rodgers KA, Lalanne NA, Vogel VG, Kanbour-Shakir A, Kelley JL, Johnson RR. Nucleotide excision repair deficiency is intrinsic in sporadic stage I breast cancer. Proc Natl Acad Sci. 2010;107(50):21725–30.

8. Cui J, Tan H, Jiang L, Yuan W, Guan Q. Association between XPC rs2228000 (C/T) polymorphism and the susceptibility of breast cancer: a Meta-analysis. J Int Oncol. 2016;43(10):752–7.

9. Özgöz A, Öztürk KH, Yükseltürk A, Şamlı H, Başkan Z, İçduygu FM, Bacaksız M. Genetic variations of DNA repair genes in breast cancer. Pathol Oncol Res. 2019;25(1):107–14.

10. Malik SS, Mubarik S, Masood N, Khadim MT. An insight into clinical outcome of XPG polymorphisms in breast cancer. Mol Biol Rep. 2018;45(6):2369–75.

11. Romanowicz H, Strapagiel D, Słomka M, Sobalska-Kwapis M, Kępka E, Siewierska-Górska A, Zadrożny M, Bieńkiewicz J, Smolarz B. New single nucleotide polymorphisms (SNPs) in homologous recombination repair genes detected by microarray analysis in Polish breast cancer patients. Clin Exp Med. 2017;17(4):541–6.

12. Aizat AAA, Nurfatimah MSS, Aminudin MM, Ankathil R. XPC Lys939Gln polymorphism, smoking and risk of sporadic colorectal cancer among Malaysians. World J Gastroenterol. 2013;19(23):3623.

13. DeSantis CE, Bray F, Ferlay J, Lortet-Tieulent J, Anderson BO, Jemal A. International variation in female breast cancer incidence and mortality rates. Cancer Epidemiol Biomarkers Prev. 2015;24(10):1495–506. https://doi.org/10.1158/1055-9965.EPI-15-0535.

14. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst. 2005;97(16):1180–4.

15. Ye S, Dhillon S, Ke X, Collins AR, Day IN. An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic Acids Res. 2001;29(17):e88.

16. Bhagwat M. Searching NCBI's dbSNP database. Curr Protoc Bioinform. 2010;32(1):1–19.

17. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10(6):845.

18. 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.

19. Abraham M, van der Spoel D, Lindahl E, Hess B. The GROMACS development team GROMACS User Manual version 5.1. SoftwareX. 2016;1:19.

20. Wang W, Xia M, Chen J, Deng F, Yuan R, Zhang X, Shen F. Data set for phylogenetic tree and RAMPAGE Ramachandran plot analysis of SODs in *Gossypium raimondii* and *G. arboreum*. Data Brief. 2016;9:345–8.

21. Wallner B, Elofsson A. Can correct protein models be identified? Protein Sci. 2003;12(5):1073–86.

22. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66(4):486–501.

23. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12.

24. Zhang ET, He Y, Grob P, Fong YW, Nogales E, Tjian R. Architecture of the human XPC DNA repair and stem cell coactivator complex. Proc Natl Acad Sci. 2015;112(48):14817–22.

25. Siegel RL, Miller KD. Jemal A (2018) Cancer statistics. CA Cancer J Clin. 2018;68(1):7–30. https://doi.org/10.3322/caac.21442.

26. Rojas K, Stuckey A. Breast cancer epidemiology and risk factors. Clin Obstet Gynecol. 2016;59(4):651–72.

27. Mubarik S, Malik SS, Wang Z, Li C, Fawad M, Yu C. Recent insights into breast cancer incidence trends among four Asian countries using age-period-cohort model. Cancer Manag Res. 2019;11:8145.

28. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A. Breast cancer before age 40 years. Semin Oncol. 2009;36(3):237–49. https://doi.org/10.1053/j.seminoncol.2009.03.001.

29. Butt Z, Haider SF, Arif S, Khan MR, Ashfaq U, Shahbaz U, Bukhari MH. Breast cancer risk factors: a comparison between pre-menopausal and post-menopausal women. J Pak Med Assoc. 2012;62(2):120.

30. Surakasula A, Nagarjunapu GC, Raghavaiah KV. A comparative study of pre- and post-menopausal breast cancer: risk factors, presentation, characteristics and management. J Res Pharm Pract. 2014;3(1):12–8. https://doi.org/10.4103/2279-042X.132704.

31. Aldabal BK, Koura MR. Risk factors of breast cancer among the primary health-care attendees in Eastern Saudi Arabia. Int J Med Sci Public Health. 2016;5(2):276–81.

32. Zahmatkesh BH, Keramat A, Alavi N, Khosravi A, Chaman R. Role of menopause and early menarche in breast cancer: a meta-analysis of Iranian studies. Nurs Midwifery Stud. 2017;6(1):e37712.

33. Pervaiz R, Tosun Ö, Besim H, Serakinci N. Risk factor assessment for breast cancer in North Cyprus: a comprehensive case–control study of Turkish Cypriot women. Turkish J Med Sci. 2018;48(2):293–304.

34. Malik SS, Baig M, Khan MB, Masood N (2019) Survival analysis of breast cancer patients with different treatments: a multicentric clinicopathological study. JPMA.

35. Hinyard L, Wirth LS, Clancy JM, Schwartz T. The effect of marital status on breast cancer-related outcomes in women under 65: a SEER database analysis. Breast. 2017;32:13–7.

36. Cadet J, Douki T. Formation of UV-induced DNA damage contributing to skin cancer development. Photochem Photobiol Sci. 2018;17(12):1816–41.

37. Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, Fan H, Shen H, Way GP, Greene CS. Genomic and molecular landscape of DNA damage repair deficiency across The Cancer Genome Atlas. Cell Rep. 2018;23(1):239–54.

38. Zheng W, Cong X-F, Cai W-H, Yang S, Mao C, Zou H-W. Current evidences on XPC polymorphisms and breast cancer susceptibility: a meta-analysis. Breast Cancer Res Treat. 2011;128(3):811–5.

39. Hernandez-Villafuerte K, Fischer A, Latimer N. Challenges and methodologies in using progression free survival as a surrogate for overall survival in oncology. Int J Technol Assess Health Care. 2018;34(3):300–16.

40. Yang S, Jin T, Su H-X, Zhu J-H, Wang D-W, Zhu S-J, Li S, He J, Chen Y-H. The association between NQO1 Pro187Ser polymorphism and bladder cancer susceptibility: a meta-analysis of 15 studies. PLoS ONE. 2015;10(1):e0116500.

41. Long X-D, Huang H-D, Huang X-Y, Yao J-G, Xia Q. XPC codon 939 polymorphism is associated with susceptibility to DNA

damage induced by aflatoxin B1 exposure. Int J Clin Exp Med. 2015;8(1):1197.

42. Gangwar R, Mandhani A, Mittal RD. XPC gene variants: a risk factor for recurrence of urothelial bladder carcinoma in patients on BCG immunotherapy. J Cancer Res Clin Oncol. 2010;136(5):779–86.

43. Krzeszinski JY, Choe V, Shao J, Bao X, Cheng H, Luo S, Huo K, Rao H. XPC promotes MDM2-mediated degradation of the p53 tumor suppressor. Mol Biol Cell. 2014;25(2):213–21.