




ORIGINAL ARTICLE

The impact of *CYP19A1* variants and haplotypes on breast cancer risk, clinicopathological features and prognosis

Ahmad Mohammed Alwan¹  | Fahimeh Afzaljavan^{2,3}  | Jalil Tavakol Afshari¹ |
Fatemeh Homaei Shandiz⁴ | Matineh Barati Bagherabad²  | Elham Vahednia² |
Nahid Kheradmand² | Alireza Pasdar^{2,5} 

¹Immunology Research Group, Immunogenetic Section, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

²Department of Medical Genetics and Molecular Medicine, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

³Student Research Committee, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

⁴Cancer research Center, Mashhad University of Medical Sciences, Mashhad, Iran

⁵Division of Applied Medicine, Medical School, University of Aberdeen, Foresterhill, Aberdeen, UK

Correspondence

Alireza Pasdar, Department of Medical Genetics and Molecular Medicine, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
Email: a.pasdar@abdn.ac.uk; pasdara@mums.ac.ir

Funding information

Mashhad University of Medical Sciences

Abstract

Background: Different genetic variants in hormone-regulating pathways have been identified to influence the risk of breast cancer. This study aimed to evaluate the association of *CYP19A1* *rs10046* and *rs700519* polymorphisms with the risk, clinicopathological factors and prognosis of breast cancer.

Methods: In a case-control study, *rs10046* and *rs700519* polymorphisms were genotyped using ARMS-PCR and high-resolution melting (HRM), respectively, in a total of 702 females. Statistical analysis and evaluation of haplotypes and linkage disequilibrium were performed using SPSS v16, PHASE and 2LD.

Results: Although no association of *rs700519* with breast cancer was observed, *rs10046* in different genetic models as well as C-C/C-T and C-C/C-C diplotypes, revealed the association with the risk of breast cancer ($p < 0.05$). Moreover, the *rs700519*-C allele was shown to be associated with longer overall survival. In contrast, the T-T haplotype conferred a shorter overall survival. *rs700519*-C allele was also significantly associated with menarche age.

Conclusion: Based on the identified independent association between *CYP19A1* diplotypes and *rs700519*-C allele with the risk and prognosis of the disease, the gene region and its genetic variants may have a diagnostic and prognostic role in breast cancer development. Further confirmation using other variants in this locus can validate these findings.

KEYWORDS

biomarker, breast neoplasm, *CYP19A1*, diagnosis, genetic variation, overall survival, *rs10046*, *rs700519*

Ahmad Mohammed Alwan, Fahimeh Afzaljavan and Jalil Tavakol Afshari have equal contribution.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals LLC.

1 | INTRODUCTION

Breast cancer is one of the most globally prevalent malignancy among females and the second cancer-caused mortality in Asian ethnic women. In addition to the lower age of onset (Bagherabad et al., 2019), the incidence of the disease is increasing in Asian countries alongside the socioeconomic growth of nations (Bray et al., 2018). It may necessitate identifying the risk factors that may contribute to the development of cancer in this population.

Because breast cancer is a multifactorial disorder, different genetic and environmental risk factors are involved in tumorigenesis. Previous studies have illustrated that exposure to a higher amount of endogenous estrogens during life has a critical role in increased breast cancer risk (Fortner et al., 2013; Moore et al., 2016; Travis & Key, 2003). Circulating concentrations of estrogen are related to the *CYP19A1* gene (OMIM: * 107910), encoding the aromatase enzyme, which has a significant role in increasing the risk of breast cancer (Friesenhengst et al., 2018). Aromatase, locating on chromosome 15q21.2, is expressed in ovaries and different extragonadal tissues such as the subcutaneous fat, liver, bone, brain, vascular endothelial tissues as well as mesenchymal cells in the breast's adipose tissue (Artigalás et al., 2015).

Different genetic variants in *CYP19A1* gene region were identified to be associated with various diseases (Chace et al., 2012; Ma et al., 2005; Wang et al., 2016; Yang et al., 2010). In this regard, single-nucleotide substitutions such as *rs10046*, located in the 3' untranslated region (3'UTR) (Zins et al., 2014) (Zhang et al., 2009) (Fasching et al., 2008), and *rs700519* (Arg264Cys) in exon seven codon 264 have been considered in various molecular studies (Chattopadhyay et al., 2014; Khvostova et al., 2012; Yang et al., 2015). These variants may influence *CYP19A1* gene expression and its related activities with susceptibility to cancer development (Warsy et al., 2017).

The potential diagnostic and prognostic role of *CYP19A1* variants and their haplotypes have not been studied in Iranian population. Therefore, the present study aimed to assess the association of *rs10046* and *rs700915* polymorphisms with the risk of breast cancer as well as histopathological characteristics and prognosis in a group of north-eastern Iranian population.

2 | MATERIALS AND METHODS

2.1 | Ethical compliance

Written informed consent was obtained from all subjects. The study was approved by the Mashhad University of Medical Sciences ethics committee (ethical approval number: IR.MUMS.MEDICAL.REC.1397.166).

2.2 | Study population

A total of 399 patients with confirmed breast cancer who were referred to the oncology departments in teaching hospitals of Mashhad University of medical sciences, Mashhad, Iran and two private oncology clinics between the years 2015 and 2018 were recruited. A total of 303 healthy individuals who referred for screening and their health status was confirmed by clinicians were also enrolled as the control group. Demographic parameters including age, weight, height, BMI and family history of other cancer types were collected using a questionnaire. Clinical data were also extracted from medical records for all patients.

2.3 | Genotyping

Five ml of peripheral blood was collected in tubes containing EDTA. DNA extraction was performed using the salting-out method, and the extracted DNA was quantified by Epoch™ Microplate Spectrophotometer (BioTek Instruments Inc.).

rs10046 polymorphism was genotyped by Amplification Refractory Mutation-Polymerase Chain Reaction (ARMS-PCR). The sequences of primers used for genotyping were as following: Forward outer: 3'GACAGTGTGTTGAGAGCATAACAGAAG5', Reverse outer: 3'CTTTTCTCTTGTAGCCTGGTTCTCT5', Forward inner: 3'AACACTAGAGAAGGCTGGTCAGTAAG5', Reverse inner: 3'TACTGATGAGAAATGCTCCAGATTG5'. Two outer primers produced a sequence of 328 bp. This sequence was considered as an internal control. The forward inner primer and the reverse outer primer produced a sequence with 218 bp that was indicating the *rs10046T* allele. The reverse inner primer and the forward outer primer produced a sequence with 162 bp that was indicating the C allele. PCR amplification for *rs10046* was performed in a 10 µl reaction volume containing 4 µl Taq 2x master mix (AMPLIQON), 1 µl of each outer primer with the concentration of 10 µM and 1.5 µl of each inner primers with the concentration of 15 µM and 1 µl (150 ng) genomic DNA. The ARMS-PCR condition was as follows: initial denaturation at 95°C for 5 min, followed by 35 cycles of 95°C for 15 s, 95°C for 15 s, 72°C for 20 s, and 72°C for 5 min as the final extension step. To further validate ARMS-PCR findings, some samples were randomly sequenced using Sanger sequencing.

The high-resolution melting (HRM) analysis was performed in the Rotor-Gene 6000™ real-time analyser (Applied Biosystem) using the HRM curve analysis assay. Real-time PCR for HRM was conducted using the Eva Green saturating dye (Type-it HRM PCR Kit, Qiagen). Sequences of the forward primer (5'GACAGTGTGTTGAGAGCATAACAGAAG3') and the reverse primer (5'CTTTTCTCTTGTAGCC

TGGTTCTCT3') were applied in each reaction. PCR was performed in 20 µl reaction volumes containing 5 µl of 2× HRM PCR Master Mix, 10 µl of distilled water, 2 µl of the primer mix (final concentration of 0.7 µM each) and 1 µl of DNA (150 ng/reaction). Amplification was performed with an initial denaturation at 95°C for 10 min, followed by 40 cycles at 95°C for 20's and 45°C for 15 s and 72°C for 10's and a final extension of 72°C for 7 min. HRM ramps were generated by acquiring fluorescence data at the temperature ramp of 74°C to 78°C at 0.05°C intervals. To further validate, some samples were randomly sequenced using Sanger sequencing.

2.4 | Statistical analysis

The data were analysed using SPSS version 16. Statistical significance was considered at a p -value < 0.05. Hardy–Weinberg analysis was performed to assess the genotype distributions using the Pearson χ^2 distribution with one degree of freedom. The odds ratio (OR) with 95% confidence intervals (95%CI) was estimated for the measured risk factors. Multivariate logistic regression analysis was used to assess the variables that were independently associated with the risk of breast cancer. Independent sample t -test and one-way ANOVA tests were used to evaluate normally distributed variables and were expressed as means with their standard deviation.

Haplotype frequency distributions were obtained from genotypes using the PHASE software version 2 (Stephens & Donnelly, 2003; Stephens et al., 2001). 2LD program version 1.00 was used to estimate linkage disequilibrium (LD) (Zhao, 2004). Association between haplotypes and the risk of breast cancer was shown using odds ratios with 95%CI and related p -value. A p -value of less than 0.05 was considered significant.

3 | RESULTS

3.1 | Demographic and histopathologic characteristics of the study population

Demographic parameters of the studied population are shown in Table 1. Although the frequency of patients with an age greater than 40 years is higher compared with controls ($p = 0.001$), there was no significant difference between the mean age of breast cancer cases (47.32 ± 10.65) and controls (43.32 ± 11.96) ($p = 0.053$).

Regarding the tumour characteristics of breast cancer cases in Table 2, the incidence of invasive ductal carcinoma was 72.7% as the most frequent type of breast tumours. Moreover, 41.6% of tumours were grade II, and 38.6% of patients were in stage II. Evaluation of the tumour markers showed ER and/or PR positive types were the most frequent tumour marker-based category.

TABLE 1 Demographic characteristics of study population

Characteristic	Breast cancer	Control	p -value*	OR (95%CI)
Age of diagnosis ^a	47.32±10.65	45.65±11.74	0.053	1.01 (1.00–1.03)
Age of diagnosis <40	90 (23.3%)	106 (35.0%)	Reference	
Age of diagnosis ≥40	297 (76.7%)	197 (65.0%)	0.001	1.78 (1.27–2.48)
Age of menarche ^a	13.08±1.64	13.24±1.62	0.209	0.94 (0.85–1.04)
Age of menopause ^{a,b}	47.44±5.52	48.45±5.43	0.195	0.97 (0.92–1.02)
Menopause status				
Pri and pre-menopause	216 (59.0%)	215 (73.4%)	Reference	
Post-menopause	150 (41.0%)	78 (26.6%)	<0.001	1.91 (1.37–2.67)
BMI ^a	27.77±4.91	25.40±4.20	<0.001	1.12 (1.088–1.16)
BMI <25	99 (27.1%)	145 (49.3%)	Reference	
BMI ≥25	266 (72.9%)	149 (50.7%)	<0.001	2.62 (1.89–3.62)

Data are presented as mean ± SD or n (%).

Abbreviations: BMI, body mass index; CI, confidence interval; OR, odds ratio.

^aMean ± SD.

^bThe age of menopause in individuals with natural menopause.

*Significant p -values have been shown in bold.

TABLE 2 Tumour characteristics of breast cancer cases

Characteristics	Number	Percent (%)
Tumour type		
Invasive ductal carcinoma	290	72.7
Invasive lobular carcinoma	11	2.8
Others	45	11.3
Unreported	53	13.3
Grade		
I	49	12.3
II	166	41.6
III	80	20.1
Unreported	104	26.1
Tumour size		
T1	77	19.3
T2	170	42.6
T3	53	13.3
T4	19	4.8
Unreported	80	20.1
Lymph node		
Negative	122	30.6
I	106	26.6
II	70	17.5
III	15	3.8
Unreported	86	21.6
Metastasis		
Negative	292	73.2
Positive	20	5.0
Unreported	87	21.8
Stage		
I	36	9.0
II	154	38.6
III	101	25.3
IV	20	5.0
Unreported	88	22.1
ER status		
Negative	88	22.1
Positive	266	66.7
Unreported	45	11.3
PR status		
Negative	99	24.8
Positive	255	63.9
Unreported	45	11.3
HER2		
Negative	226	56.6
Positive	93	23.3

(Continues)

TABLE 2 (Continued)

Characteristics	Number	Percent (%)
Equivocal	24	6.0
Unreported	56	14.0
Receptor status		
ER/PR+ HER2±	275	68.9
ER/PR– HER2+	35	8.8
Triple negative (TNBC)	39	9.8
Unreported	50	12.5

Abbreviations: ER; estrogen receptor, PR; human epidermal growth factor receptor 2; progesterone receptor and HER2.

3.2 | The association of *CYP19A1* *rs10046* and *rs700519* genotypes, haplotypes and diplotypes with breast cancer risk

Hardy–Weinberg equilibrium was established between the genotypes distribution of *rs10046* ($p = 0.137$) and *rs700519* ($p = 0.385$) in the control group.

Considering the *rs10046*, allele frequencies illustrated the lack of association of *rs10046* alleles with the risk of breast cancer in the population ($p = 0.790$). While the association was present in recessive ($p = 0.040$), additive ($p = 0.005$) and co-dominant ($p = 0.001$) models, the dominant model did not indicate the difference in the distribution of T allele carriers (CT + TT) compared with CC genotype between the groups ($p = 0.055$). Adjustment for confounding factors including age, BMI and menopause status did not alter the findings. Results are shown in Table 3.

The distribution of *rs700519* alleles and genotypes were similar between cases and controls ($p \geq 0.05$). Therefore, there was no association between *rs700519* and breast cancer risk in our studied population. Adjustment for confounding factors including age, BMI and menopause status did not alter the findings. Results are shown in Table 3.

The identified *rs10046*-*rs700519* haplotypes and diplotypes with a frequency greater than 10% are shown in Table 4. The distribution of haplotypes in our population was similar between the groups ($p \geq 0.05$). C-C haplotype with the frequency of 52.8% and 53% in cases and controls, respectively, was the most frequent haplotype. These data pointed out the lack of association between *rs10046*-*rs700519* haplotypes and breast cancer risk ($p \geq 0.05$). Furthermore, according to linkage disequilibrium analysis, these variants are not in a tight LD (D' coefficient = 0.63).

According to the genotype data, out of nine recognized diplotypes, three including C-C/T-C, C-C/C-C and T-C/T-C were identified to have frequencies more than 10%. The comparison of diplotype distribution between cases and controls

TABLE 3 Distribution of genotypes and allele frequency of *CYP19A1* *rs10046* and *rs700519* polymorphisms in breast cancer cases and controls

Genetic model	Breast cancer	Control	<i>p</i> -value _{Adj} *	OR (95% CI) _{Adj}
<i>rs10046</i>				
CC	130 (32.6%)	120 (39.6%)	Reference	
CT	219 (54.9%)	128 (42.2%)	0.006	1.66 (1.16–2.39)
TT	50 (12.5%)	55 (18.2%)	0.423	0.82 (0.50–1.34)
Multiplicative				
C	479 (60.0%)	368 (60.7%)	Reference	
T	319 (40.0%)	238 (39.3%)	0.811	1.03 (0.81–1.30)
Dominant				
CC	130 (32.6%)	120 (39.6%)	Reference	
CT + TT	269 (67.4%)	183 (60.4%)	0.053	1.40 (0.99–1.96)
Recessive				
TT	50 (12.5%)	55 (18.2%)	Reference	
CC + CT	349 (87.5%)	284 (81.8%)	0.034	1.64 (1.04–2.58)
Additive				
TT	50 (18.6%)	55 (30.1%)	Reference	
CT	219 (81.4%)	128 (69.9%)	0.003	2.09 (1.28–3.40)
Co-dominant				
CC + TT	180 (45.1%)	175 (58.7%)	Reference	
CT	219 (54.9%)	128 (42.2%)	0.001	1.77 (1.27–2.45)
<i>rs700519</i>				
CC	340 (85.2%)	256 (84.5%)	Reference	
TC	56 (14.0%)	43 (14.2%)	0.639	1.12 (0.69–1.81)
TT	3 (0.8%)	4 (1.3%)	0.506	0.55 (0.09–3.19)
Multiplicative				
C	736 (92.2%)	555 (91.6%)	Reference	
T	62 (7.8%)	51 (8.4%)	0.928	0.98 (0.64–1.51)
Dominant				
CC	340 (85.2%)	256 (84.5%)	Reference	
CT + TT	59 (14.8%)	47 (15.5%)	0.772	0.93 (0.59–1.48)
Recessive				
TT	3 (0.8%)	4 (1.3%)	Reference	
CC + CT	396 (99.2%)	299 (98.7%)	0.495	1.84 (0.32–10.64)
Additive				
CT	56 (14.1%)	43 (14.4%)	Reference	
CC	340 (85.9%)	256 (85.6%)	0.643	1.12 (0.69–1.81)
Co-dominant				
CC + TT	343 (86.0%)	260 (85.8%)	Reference	
CT	56 (14.0%)	43 (14.2%)	0.620	0.89 (0.55–1.43)

Abbreviations: CI, confidence interval; OR, odds ratio.

*Significant *p*-values have been shown in bold.

pointed out that frequency of CC-TC diplotype was higher in cases [$p < 0.001$, OR = 1.77, 95% CI (1.31–2.41)] however, C-C/C-C diplotype was more frequent in controls [$p = 0.019$,

OR = 0.67, 95% CI (0.48–0.94)]. The distribution of other diplotypes was found to be similar between the groups ($p \geq 0.05$). Results are shown in Table 4.

TABLE 4 Association of *CYP19A1* *rs10046* and *rs700519* haplotypes and risk of breast cancer

	Breast cancer	Control	<i>p</i> -value*	OR (95% CI)
Haplotype				
<i>rs10046-rs700519</i>				
C-C vs. Others	421 (52.8%)	321 (53.0%)	0.937	0.99 (0.80–1.22)
T-C vs. Others	234 (38.6%)	315 (39.5%)	0.744	1.04 (0.83–1.29)
Diplotype				
<i>rs10046-rs700519/</i> <i>rs10046-s700519</i>				
C-C/T-C vs. Others	196 (48.9%)	106 (35.0%)	<0.001	1.77 (1.31–2.41)
C-C/C-C vs. Others	97 (24.3%)	98 (32.3%)	0.019	0.67 (0.48–0.94)
T-C/T-C vs. Others	48 (12.0%)	52 (17.2%)	0.055	0.66 (0.43–1.01)

D' coefficient = 0.63, $p < 0.001$

*Significant *p*-values have been shown in bold.

3.3 | The association of *CYP19A1* *rs10046* and *rs700519* genotypes, haplotypes and diplotypes with breast cancer risk factors

Breast cancer risk factors including age at diagnosis, menarche and menopause age, menopause status and BMI were assessed in association with *rs10046* and *rs700519* genotypes, haplotypes and diplotypes in the breast cancer group. Results revealed the mean age of menarche was lower in *rs700519* C allele carriers in allelic, dominant and additive models with a significant difference. Moreover, the C-T haplotype of *rs10046-rs700519* was associated with higher age of menarche ($p = 0.011$). No significant difference was observed between other risk factors and haplotypes and diplotypes ($p > 0.05$). The results are shown in the supplementary file.

3.4 | The association of *CYP19A1* *rs10046* and *rs700519* genotypes, haplotypes and diplotypes with breast cancer tumour characteristics

The regression analysis showed a lack of association between the polymorphisms, haplotypes and diplotypes with the tumour characteristics of breast cancer including grade, stage, hormone receptors and HER2 status. Results can be found in the supplementary file.

3.5 | The association of *CYP19A1* *rs10046* and *rs700519* genotypes, haplotypes and diplotypes with breast cancer overall survival

Kaplan–Meier analysis displayed no significant association between alleles and genotypes of *rs10046* and overall

survival ($p \geq 0.05$). Evaluation of overall survival in relationship with *rs700519* specified the significant association between C allele with longer survival [$p = 0.006$, HR = 0.38, 95%CI (0.19–0.76)]. In addition, the recessive model (TT vs. CC + CT) [$p < 0.001$, HR = 21.92, 95%CI (4.80–100.23)] indicated a significant association with overall survival. Results are shown in Figure 1.

According to the survival plots of *rs10046* and *rs700519* haplotypes, a significant association was found for T-T haplotype [$p < 0.001$, HR = 16.34, 95% CI (3.84–69.56)] as shown in Figure 1. Other haplotypes, as well as diplotypes, did not indicate any association with the overall survival.

4 | DISCUSSION

In recent years, patients with breast cancer have increased among different populations, and the rate of cancer mortality and morbidity in women surges consequently. Due to the hormonal influence on the disease, there is some evidence about the higher expression level of aromatase (cytochrome P450 enzyme) in breast tumours than normal tissues (Chen, 1998). Furthermore, the association of *CYP19* genetic variants with enzyme activity has been documented (Ma et al., 2005). In this regard, the hypothesis of the association of *CYP19A1* polymorphisms with the risk of cancer development has been proposed. To validate the previous findings in our population, in this study the association of *CYP19A1* *rs10046* and *rs700519* polymorphisms with the risk of breast cancer was assessed in an Iranian population. Furthermore, these variants and their haplotypes and diplotypes were evaluated in association with the risk factors, pathological characteristics and prognosis of breast cancer. Results indicated the association of *rs10046* in different genetic models and the most frequent C-C/T-C (*rs10046-rs700519*) diplotype with the risk of breast cancer. Moreover, the *rs700519*-C allele was associated with longer survival.

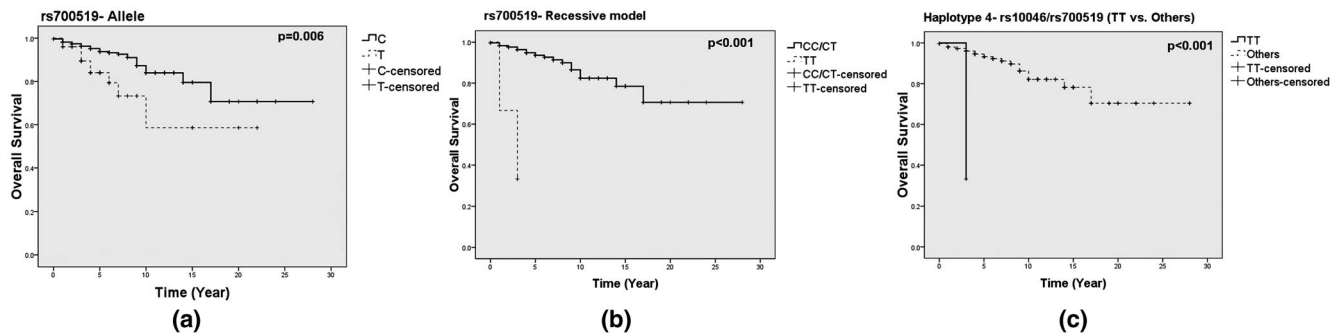


FIGURE 1 Association of *CYP19A1* polymorphism with the overall survival of breast cancer. (a) Association of *rs700519*-C allele with overall survival [C allele (42 events out of 550) versus T allele (10 events out of 50)]; (b) Association of *rs700519* recessive model with overall survival [TT genotype (2 events out of 3) versus CC + CT genotypes (24 events out of 297)]; (c) Association of T-T haplotype of *rs10046* and *rs700519* with overall survival [T-T haplotype (2 events out of 3) versus others (50 events out of 597)]

Based on the evaluation of *rs10046* in the northeast Iranian population, the frequency of the ancestral C allele (60.7% in the control group) was similar to the previous report (Farzaneh et al., 2016). The distribution of alleles has been reported to be ranged between 36% and 87% in different ethnicities ("Reference SNP (refSNP) Cluster Report: *rs10046*", 2018) which can modify the impact of this allele in association with cancer pathogenesis. According to the results of the present study, the *rs10046*-CT genotype is associated with an elevated risk of breast cancer. The impact of this genotype was observed in recessive, additive and codominant models with the risk ranged between 66% up to more than two-fold. However, we acknowledge that due to the higher rate of heterozygosity for this marker, a larger sample size may provide a better understanding of the allelic architecture. A previous study on an Iranian population represented this finding for the T allele with the increased risk of 60% and between 0.7 up to 2.1-fold risk for recessive, dominant and additive models. Although consistent with our results, a significant association was observed in two Chinese studies, with conflicting results in terms of affecting the risk (Pan et al., 2016; Yang et al., 2015). Similar to a Caucasian cohort, a significant association was not observed between allele frequencies and the risk of the disease (Hefler et al., 2004). Moreover, a meta-analysis study explored no statistically significant association between *rs10046* and breast carcinoma in all populations (Wu et al., 2018). On the other hand, several studies of the *rs700519* effect on breast cancer found that this variant is associated with the reduced risk of breast neoplasm in North Indian (Chattopadhyay et al., 2014) and Han Chinese (Sun et al., 2015) population. However, in this study, the association of *rs700519* with breast cancer risk in our Iranian population was not confirmed. Consistent with our results, a meta-analysis study demonstrated no significant association of *rs700519* with the overall risk of breast cancer as well as in Asian and Caucasian ethnicities (Wu et al., 2018).

In further stratified analysis, we did not find a significant association between *rs10046* and clinicopathological factors. Similar to our results, studies on Chinese and Caucasian patients did not report any impact of *rs10046* genotypes on risk factors of the disease (Sun et al., 2015; Zins et al., 2014). However, a significant association was found between *rs10046*-CC genotype and pre-menopausal status in Chinese breast cancer patients (Pan et al., 2016). A Caucasian cohort of breast cancer patients from Germany also reported the association of TT genotype with estrogen receptor-positive tumours (Fasching et al., 2008). Considering *rs700519*, the results of the present study represented the CC genotype carriers tended to have lower age of menarche. The evaluation of *rs700519* polymorphism in the North Indian population with the clinicopathological features reported the association with menopausal status and clinical stage (Chattopadhyay et al., 2014). Furthermore, the association of this variant with the age of diagnosis and menopausal status has been demonstrated in the Chinese population (Pan et al., 2016). Comparing these findings, genotype frequencies of *rs700519* were similar between different clinicopathological factors in Caucasians (Fasching et al., 2008; Khvostova et al., 2012).

Analysis of prognosis based on overall survival represented the longer survival in association with the *rs700519*-C allele in the studied population. The prognostic role of *rs700519* has been previously observed in the Chinese population specially in premenopausal women as the TT genotype carriers had a higher hazard of death and lower disease-free survival compared with CC genotype ones (Long et al., 2006). In contrast with these findings, a previous Caucasian breast cancer cohort investigation discovered no association of *rs700519* with disease-free survival or overall survival (Fasching et al., 2008). In vitro analysis of aromatase activity has confirmed the contribution of the T allele (Cys) with the increased enzyme activity (Wang et al., 2011). Elevated enzyme function leads to a higher level of estrogen and links to poor prognosis (Friesenhengst et al., 2018). Therefore, the

prognostic effect of *rs700519* may be explained by the level of the protein function. However, the demonstration of this hypothesis needs a well-designed functional analysis in breast cancer patients. In addition, similar to our results indicating no effective role of *rs10046* on overall survival, a previous research reported a similar finding in the Caucasian population (Fasching et al., 2008). However, carriers of the *rs10046*-C allele along with *rs4646*-A allele have been reported to have better survival (Johansson et al., 2019). Although the finding indicates a lack of the prognostic value of *rs10046* in breast cancer, more investigation in other ethnicities may be needed to confirm this.

According to haplotype and diplotype analysis, *rs10046*-*rs700519* C-C/C-T and C-C/C-C diplotypes were recognized to increase and decrease the risk of breast cancer up to 77% and 33%, respectively. Furthermore, C-T haplotype was more observed in patients with higher age of menarche. However, a previous report assessed *rs10046*, *rs700519* and other polymorphisms of *CYP19A1* and did not identify any effective haplotype in association with cancer development (Wang et al., 2011). Although it was found carriers of T-T haplotype have better survival, because of the small number of samples in each group, this finding should be replicated to confirm with high statistical power. Another investigation revealed a better survival for the T-C haplotype along with five other polymorphisms of hormone-related genes including *ESR1*, *COMT*, *SHBG* and *GSTP1* (Pan et al., 2016). To our knowledge, it is the first study evaluating haplotypes and diplotypes in relationship with the tumour characteristics and the disease risk factors. Therefore, to extend understanding of the effects of haplotypes on clinicopathological properties of the disease, more research is required.

In conclusion, *rs10046* and *rs700519* may have a diagnostic and prognostic role, respectively. Our findings help to enrich the literature about the genetic basis of breast cancer. However, more investigation of *CYP19A1* variant in association with cancer development, clinical aspects of the disease and histopathological tumour characteristics should be performed to validate these findings. Since different tissues have specific aromatase promoters, functional analysis is suggested for breast tumours in association with different *CYP19A1* variants.

ACKNOWLEDGEMENTS

The authors thank all participants in this research. We would also like to thank Mashhad University of Medical Sciences and Omid hospital (Mashhad, Iran) for supporting the project. This work was financially supported by Mashhad University of Medical Sciences under grant No. 960691.

CONFLICT OF INTEREST

No potential conflict of interest was reported by the authors.

AUTHOR'S CONTRIBUTIONS

Ahmad Mohammed Alwan, Fahimeh Afzaljavan, Jalil Tavakol Afshari, Fatemeh Homaei Shandiz and Alireza Pasdar contributed to designing of the study. Ahmad Mohammed Alwan, Fahimeh Afzaljavan, Matineh Barati Bagherabad, Elham Vahednia and Nahid Kheradmand collected the data and helped with laboratory work. Fahimeh Afzaljavan and Alireza Pasdar performed statistical analyses and interpreted data. All authors read and approved the final manuscript.

DATA AVAILABILITY STATEMENT

The analysed datasets generated during the present study are not available publicly due to the university rules. However, subjected to approval, they may be available upon request.

ORCID

Ahmad Mohammed Alwan  <https://orcid.org/0000-0003-2998-2338>

Fahimeh Afzaljavan  <https://orcid.org/0000-0001-7413-6945>

Matineh Barati Bagherabad  <https://orcid.org/0000-0003-4646-6968>

Alireza Pasdar  <https://orcid.org/0000-0002-7864-9729>

REFERENCES

- Artigalás, O., Vanni, T., Hutz, M. H., Ashton-Prolla, P., & Schwartz, I. V. (2015). Influence of *CYP19A1* polymorphisms on the treatment of breast cancer with aromatase inhibitors: a systematic review and meta-analysis. *BMC Medicine*, *13*(1), 139. <https://doi.org/10.1186/s12916-015-0373-9>
- Bagherabad, M. B., Afzaljavan, F., Vahednia, E., Rivandi, M., Vakili, F., Sadr, S. S. H., Shandiz, F. H., & Pasdar, A. (2019). Association of Caspase 8 Promoter Variants and Haplotypes with the Risk of Breast Cancer and Its Molecular Profile in an Iranian Population: A case-control Study. *Journal of Cellular Biochemistry*, *120*(10), 16435–16444. <https://doi.org/10.1002/jcb.28781>
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424. <https://doi.org/10.3322/caac.21492>
- Chace, C., Pang, D., Weng, C., Temkin, A., Lax, S., Silverman, W., Zigman, W., Ferin, M., Lee, J. H., Tycko, B., & Schupf, N. (2012). Variants in *CYP17* and *CYP19* cytochrome P450 genes are associated with onset of Alzheimer's disease in women with down syndrome. *Journal of Alzheimer's Disease*, *28*(3), 601–612. <https://doi.org/10.3233/jad-2011-110860>
- Chattopadhyay, S., Siddiqui, S., Akhtar, M. S., Najm, M. Z., Deo, S. V., Shukla, N. K., & Husain, S. A. (2014). Genetic polymorphisms of *ESR1*, *ESR2*, *CYP17A1*, and *CYP19A1* and the risk of breast cancer: A case control study from North India. *Tumour Biology*, *35*(5), 4517–4527. <https://doi.org/10.1007/s13277-013-1594-1>
- Chen, S. (1998). Aromatase and breast cancer. *Frontiers in Bioscience*, *3*, d922–d933. <https://doi.org/10.2741/a333>

- Farzaneh, F., Noghabaei, G., Barouti, E., Pouresmaili, F., Jamshidi, J., Fazeli, A., Taghavi, S., Emamalizadeh, B., & Darvish, H. (2016). Analysis of CYP17, CYP19 and CYP1A1 gene polymorphisms in Iranian women with breast cancer. *Asian Pacific Journal of Cancer Prevention, 17*(S3), 23–26. <https://doi.org/10.7314/apjcp.2016.17.s3.23>
- Fasching, P. A., Loehberg, C. R., Strissel, P. L., Lux, M. P., Bani, M. R., Schrauder, M., Geiler, S., Ringleff, K., Oeser, S., Weibrecht, S., Schulz-Wendland, R., Hartmann, A., Beckmann, M. W., & Strick, R. (2008). Single nucleotide polymorphisms of the aromatase gene (CYP19A1), HER2/neu status, and prognosis in breast cancer patients. *Breast Cancer Research and Treatment, 112*(1), 89–98. <https://doi.org/10.1007/s10549-007-9822-2>
- Fortner, R. T., Eliassen, A. H., Spiegelman, D., Willett, W. C., Barbieri, R. L., & Hankinson, S. E. (2013). Premenopausal endogenous steroid hormones and breast cancer risk: Results from the Nurses' Health Study II. *Breast Cancer Research, 15*(2), R19. <https://doi.org/10.1186/bcr3394>
- Friesenhengst, A., Pribitzer-Winner, T., Miedl, H., Prostling, K., & Schreiber, M. (2018). Elevated aromatase (CYP19A1) expression is associated with a poor survival of patients with estrogen receptor positive breast cancer. *Hormones and Cancer, 9*(2), 128–138. <https://doi.org/10.1007/s12672-017-0317-2>
- Hefler, L. A., Tempfer, C. B., Grimm, C., Lebrecht, A., Ulbrich, E., Heinze, G., Leodolter, S., Schneeberger, C., Mueller, M. W., Muendlein, A., & Koelbl, H. (2004). Estrogen-metabolizing gene polymorphisms in the assessment of breast carcinoma risk and fibroadenoma risk in Caucasian women. *Cancer, 101*(2), 264–269. <https://doi.org/10.1002/cncr.20361>
- Johansson, H., Aristarco, V., Gandini, S., Gjerde, J., Macis, D., Guerrieri-Gonzaga, A., Serrano, D., Lazzeroni, M., Rajasekaran, A., Williard, C. V., Mellgren, G., DeCensi, A., & Bonanni, B. (2019). Prognostic impact of genetic variants of CYP19A1 and UGT2B17 in a randomized trial for endocrine-responsive postmenopausal breast cancer. *The Pharmacogenomics Journal, 20*(1), 19–26. <https://doi.org/10.1038/s41397-019-0087-z>
- Khvostova, E. P., Pustylnyak, V. O., & Gulyaeva, L. F. (2012). Genetic polymorphism of estrogen metabolizing enzymes in Siberian women with breast cancer. *Genetic Testing and Molecular Biomarkers, 16*(3), 167–173. <https://doi.org/10.1089/gtmb.2011.0131>
- Long, J. R., Kataoka, N., Shu, X. O., Wen, W., Gao, Y. T., Cai, Q., & Zheng, W. (2006). Genetic polymorphisms of the CYP19A1 gene and breast cancer survival. *Cancer Epidemiology, Biomarkers & Prevention, 15*(11), 2115–2122. <https://doi.org/10.1158/1055-9965.Epi-06-0464>
- Ma, C. X., Adjei, A. A., Salavaggione, O. E., Coronel, J., Pelleymounter, L., Wang, L., Eckloff, B. W., Schaid, D., Wieben, E. D., Adjei, A. A., & Weinshilboum, R. M. (2005). Human aromatase: gene resequencing and functional genomics. *Cancer Research, 65*(23), 11071–11082. <https://doi.org/10.1158/0008-5472.Can-05-1218>
- Moore, S. C., Matthews, C. E., Ou Shu, X., Yu, K., Gail, M. H., Xu, X., Ji, B.-T., Chow, W.-H., Cai, Q., Li, H., Yang, G., Ruggieri, D., Boyd-Morin, J., Rothman, N., Hoover, R. N., Gao, Y.-T., Zheng, W., & Ziegler, R. G. (2016). Endogenous estrogens, estrogen metabolites, and breast cancer risk in postmenopausal Chinese women. *JNCI: Journal of the National Cancer Institute, 108*(10), djw103. <https://doi.org/10.1093/jnci/djw103>
- Pan, Z., Fu, Z., Song, Q., Cao, W., Cheng, W., & Xu, X. (2016). Genetic polymorphisms and haplotype of hormone-related genes are associated with the risk of breast cancer in Chinese women. *Genetics and Molecular Research, 15*(2). <https://doi.org/10.4238/gmr.15028640>
- Reference SNP (refSNP) Cluster Report: rs10046. (2018). https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?do_not_redirect&rs=rs10046
- Stephens, M., & Donnelly, P. (2003). A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics, 73*(5), 1162–1169. <https://doi.org/10.1086/379378>
- Stephens, M., Smith, N. J., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics, 68*(4), 978–989. <https://doi.org/10.1086/319501>
- Sun, M. Y., Du, H. Y., Zhu, A. N., Liang, H. Y., de Garibay, G. R., Li, F. X., Li, M., & Yang, X. X. (2015). Genetic polymorphisms in estrogen-related genes and the risk of breast cancer among Han Chinese women. *International Journal of Molecular Sciences, 16*(2), 4121–4135. <https://doi.org/10.3390/ijms16024121>
- Travis, R. C., & Key, T. J. (2003). Oestrogen exposure and breast cancer risk. *Breast Cancer Research, 5*(5), 239–247. <https://doi.org/10.1186/bcr628>
- Wang, B., Fu, Z.-Y., Ma, Y.-T., Huang, D., Liu, F., Dong, C.-L., Wang, T., & Meng, Y.-J. (2016). Identification of a CYP19 gene single-nucleotide polymorphism associated with a reduced risk of coronary heart disease. *Genet Test Mol Biomarkers, 20*(1), 2–10. <https://doi.org/10.1089/gtmb.2015.0157>
- Wang, H., Li, Q., Wang, T., Yang, G., Wang, Y., Zhang, X., Sang, Q., Wang, H., Zhao, X., Xing, Q., Shi, J., He, L., & Wang, L. (2011). A common polymorphism in the human aromatase gene alters the risk for polycystic ovary syndrome and modifies aromatase activity in vitro. *Molecular Human Reproduction, 17*(6), 386–391. <https://doi.org/10.1093/molehr/gar007>
- Warsy, A. S., Almkaynizi, F. B., Aldaihan, S., Alam, S., & Daghestani, M. (2017). Genetic polymorphisms in aromatase (CYP19) gene and cancer. In N. Reddy Parine (Ed.), *Genetic polymorphisms* (pp. 27–56). : InTech.
- Wu, Y., Qu, X., Xia, J., Gu, Y., & Qian, Q. (2018). Four CYP19A1 polymorphisms and breast cancer risk: A meta-analysis. *Journal of Biochemistry and Physiology, 1*(1), 2.
- Yang, H. P., Gonzalez Bosquet, J., Li, Q., Platz, E. A., Brinton, L. A., Sherman, M. E., Lacey, J. V. Jr., Gaudet, M. M., Burdette, L. A., Figueroa, J. D., Ciampa, J. G., Lissowska, J., Peplonska, B., Chanock, S. J., & Garcia-Closas, M. (2010). Common genetic variation in the sex hormone metabolic pathway and endometrial cancer risk: pathway-based evaluation of candidate genes. *Carcinogenesis, 31*, 827–833.
- Yang, L., Wang, X. Y., Li, Y. T., Wang, H. L., Wu, T., Wang, B., Zhao, Q., Jinsihan, D., & Zhu, L. P. (2015). CYP19 gene polymorphisms and the susceptibility to breast cancer in Xinjiang Uigur women. *Genetics and Molecular Research, 14*(3), 8473–8482. <https://doi.org/10.4238/2015.July.28.15>
- Zhang, L., Gu, L., Qian, B., Hao, X., Zhang, W., Wei, Q., & Chen, K. (2009). Association of genetic polymorphisms of ER-alpha and the estradiol-synthesizing enzyme genes CYP17 and CYP19 with breast cancer risk in Chinese women. *Breast Cancer Research and Treatment, 114*(2), 327–338. <https://doi.org/10.1007/s10549-008-9998-0>
- Zhao, J. H. (2004). 2LD, GENECOUNTING and HAP: Computer programs for linkage disequilibrium analysis. *Bioinformatics, 20*(8), 1325–1326. <https://doi.org/10.1093/bioinformatics/bth071>

Zins, K., Mogg, M., Schneeberger, C., Abraham, D., & Schreiber, M. (2014). Analysis of the rs10046 polymorphism of aromatase (CYP19) in premenopausal onset of human breast cancer. *International Journal of Molecular Sciences*, 15(1), 712–724. <https://doi.org/10.3390/ijms15010712>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Alwan AM, Afzaljavan F, Tavakol Afshari J, et al. The impact of *CYP19A1* variants and haplotypes on breast cancer risk, clinicopathological features and prognosis. *Mol Genet Genomic Med.* 2021;9:e1705. <https://doi.org/10.1002/mgg3.1705>