RESEARCH ARTICLE

Editorial Process: Submission:07/24/2023 Acceptance:11/16/2023

Frequencies of Diagnostically Significant Polymorphisms of Hereditary Breast Cancer Forms in *BRCA1* and *BRCA2* Genes in the Kazakh Population

Gulnara Svyatova[1], Galina Berezina[1]*, Gulfairuz Urazbayeva[2], Alexandra Murtazaliyeva[1]

Abstract

Objective: Breast cancer is the most common form of cancer in women in the world with more than 400,000 deaths each year worldwide. The aim of this study is to compare population frequencies of alleles and genotypes of polymorphic variants of *BRCA1* and *BRCA2* genes associated with breast cancer risk in an ethnically homogenous Kazakh population with previously studied world populations. The material of the study was DNA isolated from peripheral blood of the enrolled population control group, represented by 1800 conditionally healthy individuals of Kazakh ethnicity. **Methods:** The DNA extraction was possible with the use of M-PVA magnetic particle separation method on the Prepito (PerkinElmer) automatic analyser for extraction of Chemagic Prepito (Wallac, Finland) nucleic acids using the PrepitoDNACytoPure reagent kit. Statistical calculations of allele and genotype frequencies, significance tests, and non-parametric $\chi 2$ analysis were carried out using PLINK software. **Results:** The results favour for the high genetic heterogeneity of the studied polymorphisms, which reflects the specifics of the Kazakh population structure resulted from complex evolutionary and migration processes, as well as the median geographic location between the populations of Asia and Europe. **Conclusion:** Knowledge of the spectrum and frequency of mutations in BRCA1 and BRCA2 genes predisposing to breast cancer, which are present in varying frequencies in the Kazakh population, will provide a more effective approach to the screening protocol and allow for a faster, less expensive and more accessible genetic testing strategy for the Kazakhstan citizens.

Keywords: DNA- polymorphisms- screening- genes- mutation

Asian Pac J Cancer Prev, **24 (11)**, 3899-3907

Introduction

According to the World Health Organization (WHO) in 2020, there were 2.3 million women diagnosed with breast cancer worldwide and 685,000 deaths globally. (World Health Organization, 2023). In 2020, there were 4,390 new cases of breast cancer in Kazakhstan (The Global Cancer Observatory, 2023). About 50-60% of hereditary forms of breast cancer result from highly penetrant mutations in *BRCA1* and *BRCA2* genes, which are involved in the repair of double-stranded DNA (deoxyribonucleic acid) breaks (Mahdavi et al., 2018). The remainder is due to mutations in genes that encode tumour suppressor proteins and are responsible for genomic stability of the cell (*TP53, PTEN, STK11, CDH1*) and for repairing of the damaged DNA (*ATM, CHEK2, PALB2, BARD1, BRIP1, MRE11A, NBN, BLM*) (Urbina-Jara et al., 2019; Maslii et al., 2021).

Terminal mutations (mutations that occur in gametes or in the cells from which they are formed) in *BRCA1* and *BRCA2* genes with high penetrance inherited under an autosomal dominant pattern increase the lifetime risk of developing breast cancer by up to 80%. In addition to familial forms, the genetic contribution of pathogenic mutations in *BRCA1* and *BRCA2* genes takes part in the development of sporadic breast cancer (Gou et al., 2017; Ghafouri-Fard et al., 2017; Alyahri et al., 2019). The prevalence of *BRCA1* and *BRCA2* pathogenic mutations in the general population is about 0.2% (1 per 500) but can vary significantly between countries and ethnic groups. According to a meta-analysis, the frequency of mutations in BRCA genes in non-family breast cancer was quite low, ranging from 0.02% to 10%: 5.1% in Asia, 0.8% in Japan, up to 8.0% in Singapore and 0.7% in the United Kingdom (Antoniou et al., 2003). The burden of breast cancer is high in Kazakhstan, so finding out the genetic susceptibility in Kazakhs is important for public health. It is also worth adding that previous studies of *BRCA* mutations in ethnic Kazakhs are limited. Different ethnic and geographical

*[1]Republican Medical Genetic Consultation, Scientific Center of Obstetrics, Gynecology and Perinatology, Almaty, Republic of Kazakhstan. [2]Scientific Center of Obstetrics, Gynecology and Perinatology, Almaty, Republic of Kazakhstan. *For Correspondence: gberezina@yahoo.com*

regions have their own prevalence and spectrum of *BRCA1* and *BRCA2* mutations. However, current strategies for selecting subjects for genetic testing seem undeveloped in most countries due to insufficient knowledge of the genetic structure of specific populations (Pestun, 2021). The unique profile of *BRCA1/2* gene polymorphisms in Kazakhs probably reflects the peculiar genetic history of this population at the crossroads of Europe and Asia. Factors such as the founder effect, genetic drift, natural selection, and intermixing of different ethnic groups could have shaped the specific spectrum of *BRCA1/2* variations in this population. Further research is needed on the evolutionary mechanisms and migration processes that led to the specific distribution of *BRCA1/2* polymorphisms alleles among Kazakhs.

With the foregoing in mind, the objective of this study was to evaluate the population frequencies of alleles and genotypes of polymorphic variations of the *BRCA1* and *BRCA2* genes linked with breast cancer risk in an ethnically homogenous Kazakh population with previously studied world populations. This appears to be one of the first large-scale studies analyzing the spectrum and frequencies of *BRCA1/2* mutations specifically in an ethnically homogeneous Kazakh population sample. Previous studies on BRCA mutations in Kazakhstan had limitations like small sample sizes and lack of ethnic stratification. The task of the research of the study were to characterize the spectrum and frequencies of functionally significant *BRCA1/2* mutations in the Kazakh population that predispose to hereditary susceptibility to breast cancer. To identify mutations with high population frequencies in Kazakhs that may have greater pathogenic contribution to breast cancer risk in this ethnic group. To compare mutation frequencies in Kazakhs to other global populations to understand population-specific differences reflecting evolutionary/migration history and geographic location between Asia and Europe.

Materials and Methods

The material of the study was DNA isolated from peripheral blood of the enrolled population control group, represented by 1,800 conditionally healthy individuals of Kazakh ethnicity. The criteria for selection to the control group were as follows: Kazakh ethnicity, including grandparents, age 18 years or older and capacity of the subject to make an independent decision to participate in the project. All those enrolled were aware of the aims of the project and signed an informed consent for participation. The DNA of the study subjects is stored in the "Miras" Biobank of Scientific Center of Obstetrics, Gynecology and Perinatology, established under the project "Genetic Research on Pre-eclampsia in Central Asian and European Populations" (InterPregGen) 7th Framework Programme of the European Commission Grant Agreement No. 282540.

To select significant polymorphisms in *BRCA1* and *BRCA2* genes associated with the risk of breast cancer, the authors used the databases of population mutation frequencies according to their clinical significance: National Library of Medicine, Pharmacogenomics Knowledgebase (PharmGKB), Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA), International Genome Sample Resource (IGSR), Genome Aggregation Database (gnomAD) and results of GWAS (genome-wide association studies). The DNA extraction was possible with the use by M-PVA magnetic particle separation method on the Prepito (PerkinElmer) automatic analyser for extraction of Chemagic Prepito (Wallac, Finland) nucleic acids using the PrepitoDNACytoPure reagent kit. Genotyping of each individual for ~2.5 million SNPs was carried out using IlluminaOmniChip 2.5 – 8 chips at DECODE Iceland Genome Centre as part of the InterPregGen project. Genotyping quality control was conducted with exclusion of SNPs with MAF (minimal allele frequency) below 1%, call rate<98% with deviation from Hardy-Weinberg equilibrium (P<0.05) (Berndt, 2013).

Statistical calculations of allele and genotype frequencies, significance tests, and non-parametric χ^2 (chi-square) analysis were performed using PLINK software (Purcell, 2007). The assessment of the consistency of the genotype frequencies obtained with Hardy-Weinberg equilibrium (HWE) law was calculated using the HWE-test function of the PLINK software. The results of the analysis in all cases deem statistically significant at the P<0.05 level. The SNPs with a minimum population frequency in the Kazakh population (MAF) of less than 0.01 and with significant deviations from Hardy-Weinberg equilibrium (P<0.05) were excluded from the study. Twelve single-nucleotide polymorphisms made entry from GWAS studies with possible association with breast cancer for further comparative analysis of population frequencies of significant *BRCA1* and *BRCA2* gene polymorphisms. In the *BRCA1* gene – *rs1799950, rs4986850, rs16942, rs1800709, rs4986852, rs1799966, rs1799949, rs799917* and in the *BRCA2* gene – *rs766173, rs144848, rs4987117, rs1801426.*

The local ethical commission of the Asfendiyarov Kazakh National Medical University (Almaty, Kazakhstan) has approved the study under Application No. 1189 dated 28 September 2021. All patients enrolled were aware of the aims of the project and signed an informed consent to participate in it.

Results

The *BRCA1* gene contains 22 exons with mapping on chromosome 17q217. *BRCA1* plays a critical role in DNA repair, cell cycle control and maintenance of genome stability by forming several different complexes through association with different adaptor proteins. The *BRCA1* gene expression takes place in many tissues such as lymph nodes, skin, bladder, cervix, liver, uterus, prostate, pancreas, lung, kidney, bone, and brain and is associated with various types of cancer, including breast, ovarian, endometrial, pancreatic, prostate and colorectal cancers. The *BRCA2* gene is localised on chromosome 13q12- q13 at position 13.1 and contains 27 exons. The BRCA2 belongs to a class of tumour suppressor genes that regulate cell division, keeping cells from multiplying too quickly. Both *BRCA1* and *BRCA2* genes are involved in maintaining genome stability, especially in the homologous recombination pathway for double-stranded DNA repair.

Table 1. Genetic Characteristics of 12 Polymorphisms in *BRCA1* and *BRCA2* Genes Associated with the Development of Breast Cancer

No.	Name of the gene		Chromosome	rs	Position
BRCA1					
1	*BRCA1*	Breast And Ovarian Cancer Susceptibility Protein 1	17	rs4986850	43093454
2				rs1799950	43094464
3				rs16942	43091983
4				rs1799966	43071077
5				rs1800709	43093010
6				rs4986852	43092412
7				rs1799949	43093449
8				rs799917	43092919
BRCA2					
9	*BRCA2*	Breast And Ovarian Cancer Susceptibility Protein 2	13	rs766173	32332343
10				rs144848	32332592
11				rs4987117	32340099
12				rs1801426	32398747

rs, polymorphism identifier (SNP Identifier).

The mutations in the *BRCA2* gene are associated with an increased risk of breast cancer (31-56%), ovarian cancer (10-27%), prostate cancer, colorectal cancer, pancreatic cancer, gall bladder cancer and bile duct cancer, stomach cancer and malignant melanoma (Debniak et al., 2018).

The *BRCA1* and *BRCA2* genes have many polymorphisms; 3,517 SNPs in *BRCA1* and 3,902 in *BRCA2* are present in the Human Gene Mutation Database (HGMD), a regularly updated list. However, the database consisted mainly of Caucasian patients with hereditary breast cancer and ovarian cancer. Hence, the frequencies and spectrum of polymorphisms in *BRCA1* and *BRCA2* genes in sporadic BC forms have been largely unexplored, especially in Asian populations. The findings support recent proposals for genetic testing of *BRCA1* and *BRCA2* not only in familial but also in sporadic forms of breast cancer (Akilzhanova et al., 2013) carrying pathogenic mutations in these genes, which is associated with a high adverse prognosis of choosing effective therapy (Pashayan et al., 2018).

There is no consensus to date on the genetic contribution of common polymorphisms and often-unspecified pathogenicity in *BRCA1* and *BRCA2* genes to the risk of breast cancer. This issue is being studied extensively by the Consortium of Investigators of Modifiers of *BRCA1/2* (2022), which aims to identify genetic modifiers of breast cancer risk in *BRCA1* and *BRCA2* carriage. The genome-wide associative (GWAS) studies have identified a number of common polymorphisms with low penetrance but their contribution to breast cancer risk differs between ethnic populations (Paulo et al., 2017). Associative GWAS studies have shown that more than 50 common polymorphisms in *BRCA1* and *BRCA2* genes are associated with a moderate increase in the risk of breast cancer and multiply the shared polygenic risk of BC in carriers of mutations in the BRCA genes (Pashayan et al.,

Table 2. Allele and Polymorphic Genotype Frequencies in *BRCA1* and *BRCA2* Genes associated with the Development of Breast Cancer in the Kazakh Population

Name of the gene	rs	MAF	N	A1	A2	GENO
BRCA1	rs4986850	0.02721	1,801	G	A	1703/98/0
	rs1799950	0.02417	1,800	A	G	1714/85/1
	rs16942	0.3077	1,799	A	G	857/777/165
	rs1799966	0.3104	1,801	A	G	851/782/168
	rs1800709	0.001388	1,801	C	T	1796/5/0
	rs4986852	0.006941	1,801	G	A	1776/25/0
	rs1799949	0.2869	1,797	G	A	930/703/164
	rs799917	0.3111	1,800	C	T	848/784/168
BRCA2	rs766173	0.1214	1,800	T	G	1392/379/29
	rs144848	0.2489	1,800	T	G	1024/656/120
	rs4987117	0.01194	1,801	C	T	1758/43/0
	rs1801426	0.0399	1,801	A	G	1660/138/3

Note: rs, polymorphism identifier (SNP Identifier); MAF, minor allele population frequency; N, number of genotyped; wild type A1 allele and minor A2 allele; GENO, number of genotypes identified.

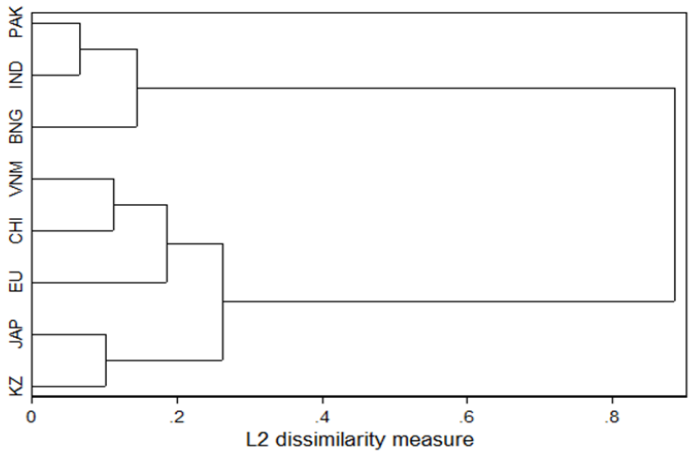

Figure 1. Dendrogram of Genetic Distances of Kazakh and Global Populations by Frequency of SNPs and GWAS Associated with Breast Cancer

2018; Zhang, 2020).

A total of 14 of the 25 polymorphisms associated with the development of breast cancer are localised in the BRCA genes: 5 – in *BRCA1 – rs1799950 (Q356R), rs4986850 (D693N), rs2227945 (S1140G), rs16942 (K1183R), rs1799966 (S1613G); 9 SNP in the gene BRCA2 – rs766173 (N289H), rs144848 (N372H), rs4987117 (T1915M), rs1799954 (R2034C), rs11571746 (S2835P), rs11571747 (E2856A), rs4987047 (I2944F), rs11571833 (K3326), rs1801426 (I3412V)*.

Discussion

Analysis of minor allele frequencies of 14 polymorphisms in *BRCA1* and *BRCA2* genes in the Kazakh population according to the genomic database of the "Miras" DNA Biobank found low (less than 0.01) MAF for *rs2227945* polymorphisms in the *BRCA1* gene, and for 5 polymorphisms in the *BRCA2* gene – *rs1799954, rs11571746, rs11571747, rs4987047 and rs11571833*. This prevented these six SNPs from being included in further analysis. Additionally included are *rs1800709 (R841W), rs4986852 (S1040N), rs1799949 (S694S)* and *rs799917 (P871L)* polymorphisms in the *BRCA1* gene, which have been shown to have significant associations with breast cancer risk in different ethnic populations according to GWAS. Table 1 shows the genetic characteristics of 12 polymorphisms of *BRCA1* and *BRCA2* genes, indicating the SNP Identifier, the location of the polymorphism on the chromosome, the physical distance at the paired ends (base-pair position – bp) and gene name.

As shown in Table 2, the lowest minor allele population frequency is 0.01%, which was for *rs1800709* polymorphism of the *BRCA1* gene and the highest minor allele population frequency was for *rs799917* polymorphism of the *BRCA1* gene, which is 31.1% in the Kazakh population studied. The correspondence of genotype distribution according to Hardy-Weinberg equilibrium in 12 polymorphisms of *BRCA1* and *BRCA2* genes in the Kazakh population, obtained by statistical processing in PLINK software using HWE-test function, is presented in Table 3.

Table 3 shows that genotype distribution for all studied

Table 3. The Correspondence of Genotype Frequency Distribution According to Hardy-Weinberg Equilibrium in 12 Polymorphisms of *BRCA1* and *BRCA2* Genes in the Kazakh Population.

Name of the gene	rs	N	GENO	O(HET)	E(HET)	p
BRCA1	rs4986850	1,801	1703/98/0	0.05441	0.05293	0.6408
	rs1799950	1,800	1714/85/1	0.04722	0.04717	1
	rs16942	1,799	857/777/165	0.4319	0.426	0.5802
	rs1799966	1,801	851/782/168	0.4342	0.4281	0.582
	rs1800709	1,801	1796/5/0	0.002776	0.002772	1
	rs4986852	1,801	1776/25/0	0.01388	0.01378	1
	rs1799949	1,797	930/703/164	0.3912	0.4091	0.06502
	rs799917	1,800	848/784/168	0.4356	0.4286	0.5098
BRCA2	rs766173	1,800	1392/379/29	0.2106	0.2133	0.5803
	rs144848	1,800	1024/656/120	0.3644	0.3739	0.2841
	rs4987117	1,801	1758/43/0	0.02388	0.02359	1
	rs1801426	1,801	1660/138/3	0.07662	0.07676	0.7616

rs, polymorphism identifier (SNP Identifier); MAF, minor allele population frequency; N, number genotyped; GENO, number of genotypes detected; O (HET), expected heterozygosity by Hardy-Weinberg equilibrium; E (HET), observed heterozygosity by Hardy-Weinberg equilibrium; p, observed significance of differences

Table 4. Comparative Analysis of Allelic Frequencies of 12 Polymorphisms of *BRCA1* and *BRCA2* Genes Associated with Breast Cancer in Global Populations

Population	N	MAF	χ^2	p
BRCA1				
BRCA1/rs4986850				
Kazakhstan	1,801	0.02721		
Europe*	503	0.083	32.842	<0.001**
East Asia*	504	0	14.010	<0.001**
South Asia*	489	0.030	0.170	0.680
BRCA1/rs1799950				
Kazakhstan	1,800	0.02417		
Europe*	503	0.060	15.661	<0.001**
East Asia*	504	0	12.560	<0.001**
South Asia*	489	0.013	2.668	0.103
BRCA1/rs16942				
Kazakhstan	1,799	0.3077		
Europe*	503	0.355	4.159	0.042**
East Asia*	504	0.371	7.179	0.008**
South Asia*	489	0.498	61.774	<0.001**
BRCA1/rs1799966				
Kazakhstan	1,801	0.3104		
Europe*	503	0.360	4.411	0.036**
East Asia*	504	0.372	6.617	0.011**
South Asia*	489	0.498	60.074	<0.001**
BRCA1/rs1800709				
Kazakhstan	1,801	0.001388		
Europe*	503	0.005	2.797	0.095
East Asia*	504	0	0.841	0.360
South Asia*	489	0.002	0.032	0.859
BRCA1/rs4986852				
Kazakhstan	1,801	0.006941		
Europe*	503	0.029	16.732	<0.001**
East Asia*	504	0	3.659	0.056
South Asia*	489	0.003	1.694	0.194
BRCA1/rs1799949				
Kazakhstan	1,797	0.2869		
Europe*	503	0.300	0.325	0.569
East Asia*	504	0.371	13.054	<0.001**
South Asia*	489	0.497	76.280	<0.001**
BRCA1/rs799917				
Kazakhstan	1,800	0.3111		
Europe*	503	0.363	4.997	0.026**
East Asia*	504	0.371	6.453	0.012**
South Asia*	489	0.530	79.927	<0.001**
BRCA2				
BRCA2/rs766173				
Kazakhstan	1,800	0.1214		
Europe*	503	0.035	31.409	<0.001**
East Asia*	504	0.096	2.684	0.102
South Asia*	489	0.133	0.448	0.504
BRCA2/rs144848				
Kazakhstan	1,800	0.2489		
Europe*	503	0.295	4.214	0.041**
BRCA2/rs144848				
East Asia*	504	0.285	2.797	0.095
South Asia*	489	0.354	21.401	<0.001**
BRCA2/rs4987117				
Kazakhstan	1,801	0.01194		
Europe*	503	0.025	4.882	0.028**
East Asia*	504	0	6.216	0.013**
South Asia*	489	0.004	4.001	0.046**
BRCA2/rs1801426				
Kazakhstan	1,801	0.0399		
Europe*	503	0.003	16.394	<0.001**
East Asia*	504	0.021	3.738	0.050**
South Asia*	489	0.002	17.931	<0.001**

Note: χ^2, chi-square; *, ensembl genome database project; **, p<0.05.

polymorphisms in the Kazakh population is in accordance with Hardy-Weinberg equilibrium, as the differences between expected and observed heterozygosity for all 12 polymorphisms were not significant (p>0.05). Table 4 shows the results of a comparative analysis of 12 minor allele population frequencies of *BRCA1* and *BRCA2* gene polymorphic variants associated with breast cancer in the Kazakh population compared to previously studied world populations. The minor allele population frequencies are presented worldwide from the 1,000 Phase III Genome database (International Genome Sample Resource, 2022), as well as electronic resources Genome Aggregation Database (2021), publication data from global databases and GWAS analyses. It is noteworthy that the sample size of the population studied, 1,800 Kazakhs, was the highest, which demonstrates the significance of the results and the possibility of extrapolating them to the entire Kazakh population.

Analysis of minor allele frequency of *rs4986850* polymorphism in the *BRCA1* gene in Kazakhs (2.7%) showed differences compared to Europeans (8.3%) and Asians (0%). This may reflect variations in triple negative breast cancer frequency. The minor allele frequency of *rs1799950 BRCA1* polymorphism in Kazakhs (2.4%) was higher than in Asians (0%) but lower than Europeans (6%). This polymorphism associates with increased breast cancer risk. Therefore, *rs1799950 BRCA1* polymorphism may make a significant contribution to breast cancer susceptibility among Kazakhs and play a role as a marker in population genetic testing of this ethnic group. Further large-scale genotyping studies specifically in Kazakh patients are essential to clarify the role of these variants in hereditary breast cancer predisposition within this unique population. The *rs1799966 BRCA1* polymorphism located in exon 16 is in linkage disequilibrium with *rs799917* and *rs16941* in most populations (International Genome Sample Resource, 2022).

Association studies of r*s1799966* with breast cancer risk remain inconsistent. A meta-analysis found no significant association in either Asian or Caucasian populations (Xu et al., 2018; Pekur et al., 2023).

However, other studies reported higher frequencies of unfavorable rs1799966 genotypes in breast cancer cases (Akkuzu et al., 2019). The heterogeneity may stem from limited sample sizes and methodological differences. Given the low minor allele frequency of *rs1799966* in Kazakhs (31.0%) versus Europeans (36.0%) and Asians (37.2%-49.8%), and the ambiguous data on its pathogenicity, this polymorphism likely contributes little to breast cancer risk in this population. Conversely, the rare *rs1800709 BRCA1* variant impairs protein function and associates with cancer (Fernandes et al., 2019). Despite an extremely low frequency (0.13% in Kazakhs), its mechanistic pathogenicity warrants inclusion. The homogeneous low carriage rates globally allow use of *rs1800709* as a negative prognostic marker for breast cancer severity in diverse populations including Kazakhs. In summary, large-scale studies specifically investigating relationships between *BRCA1* polymorphisms and breast cancer incidence in ethnically stratified Kazakh patients are essential to clarify their population-specific influences on disease risk.

Due to the rare population frequency, the *rs1800709* polymorphism went unidentified in genomic studies of 1,000 healthy individuals (Carter et al., 2018) and undescribed in the Exome Variant Server – ESP. This mutation demonstrated to inhibit *BRCA1* interaction with helicase and transcription corepressor and showed to decrease stability and to impair severely the function of the *BRCA1*-encoded protein (Fernandes et al., 2019), allowing the ClinVar expert panel to classify the rs1800709 polymorphism as pathogenic (polymorphism identifier: 17694). Although the population frequency of the minor T allele of *rs1800709* polymorphism of the *BRCA1* gene is extremely low and may only explain about 1% of all cases of breast and ovarian cancer (95% CI: 0-1.7%) (Fernandes et al., 2019; Badalova et al., 2020), the proven mechanism of pathogenicity of this polymorphism in the development of breast cancer was the reason for its inclusion in this study.

Low prevalence of *rs4986852* polymorphism of the *BRCA1* gene in the Kazakh population and conflicting scientific views on the pathogenicity of this mutation suggest that this polymorphism may be of low genetic significance in the risk of breast cancer in the Kazakh population. As shown in Table 4, minor A allele population frequency of *rs1799949* polymorphism in the *BRCA1* gene in Kazakhs was 28.7%, which did not differ from its frequency in European populations – 30.0% ($p>0.05$), but was significantly lower than in the East Asian population at 37.1% and in South Asia at 49.8% ($p<0.001$). Most studies do not find a pathogenic effect of this mutation on the risk of breast cancer but the specialists suggest not ruling out the possible potential risk of BC carrying *rs1799949* polymorphism of the *BRCA1* gene (Schapovalova et al., 2022). Significantly lower minor A allele population frequency of *rs1799949* polymorphism of the *BRCA1* gene in Kazakhs in comparison with the studied world populations suggests a relatively low frequency of breast cancer in the Kazakh population with a low genetic contribution of this polymorphism to BC risk.

Some studies have suggested an association between the *rs766173* polymorphism of the *BRCA2* gene and familial breast cancer. For example, a study of 230 women with breast cancer in Jordan found a significant association between *rs766173* polymorphism carriage and the risk of breast cancer related to breastfeeding status ($p=0.002$). However, this study did not establish a clear link between these polymorphisms and overall breast cancer risk (Al-Eitan et al., 2019). Similarly, genotyping results from 12 polymorphisms in the *BRCA2* gene in 1,109 Cypriot breast cancer patients and 1,177 healthy women indicated that carriers of the *rs766173* polymorphism in the *BRCA2* gene had an increased risk of breast cancer (OR=1.41, 95% CI: 1.08-1.83, $p=0.01$). These associations highlight the potential significance of the *rs766173* polymorphism in the *BRCA2* gene in breast cancer risk, especially in certain populations like Kazakhs.

Regarding the *rs144848* polymorphism in the *BRCA2* gene, it has a minor G allele frequency of 24.9% in the Kazakh population, which is similar to the East Asian population (28.5%, $p>0.05$) but lower than in the European population (29.5%, $p<0.05$) and the South Asian population (35.4%, $p<0.001$). This polymorphism is located in exon 10 of the *BRCA2* gene and has been the subject of multiple association studies examining its potential association with cancer risk. A meta-analysis of 40 studies, including 34,911 breast cancer cases and 48,329 controls, did not find significant associations between the *rs144848* polymorphism and breast cancer risk. However, there are contrasting results, with some studies suggesting an increased risk of breast cancer for homozygotes of the minor G/G allele of rs144848 compared to A/A genotypes. Additionally, a meta-analysis focusing on various cancer types found significant associations between the *rs144848* polymorphism and breast cancer risk in all genetic models.

As shown in Table 4, minor G allele population frequency of *rs1801426* polymorphism of the *BRCA2* gene in Kazakhs was 4.0% and was the highest compared with populations of Europe – 0.3%, South Asia – 0.2% ($p<0.001$) and East Asia – 0% ($p<0.05$). Most researchers classify SNPs of *rs1801426* polymorphism as a benign polymorphism not significantly associated with cancer risk. The researchers Gonzalez et al., (2021) studied 39 Nicaraguan women with histopathological diagnosis of breast cancer and genotyping mutations in *BRCA1*, *BRCA2, Tp53, PALB2, CDH1, PTEN* and *CHEK2* genes in 10.2% of patients and found abnormal mutations in the *BRCA2* gene. A total of 13 *BRCA2* mutations were diagnosed, of which 1% were benign (*N372H – rs144848, N289H – rs766173, N991D – rs1799944, I2490T, A2951T, I2944F* and *I3412V – rs1801426*), 10.3% with pathogenicity conflict (*Y3417H, D156G* and *L2512F*), 3.4% were indeterminate (Y600F), 3.4% were unreported (*V2446A*) and another 3.4% were pathogenic (c.2808_2811del4 (p. Ala938Profs). Using automated direct sequencing the researchers assessed the frequency and spectrum of mutations in *BRCA1* and *BRCA2* genes in 156 Kazakh women with sporadic breast cancer and 112 control patients of the same age. The authors identified twenty-two different polymorphisms with nine missense mutations and three synonymous polymorphisms detected in the *BRCA1* gene.

In *BRCA2*, missense mutations of the *c.865A>C (N289H), c.10234A>G (I3412V)* gene were detected with a higher frequency in 15.2% and 30.8% of breast cancer cases compared to controls (10.1%, 18.7%, respectively). It was observed that the frequency of almost all detected polymorphisms including c.10234A>G in the *BRCA2* gene was significantly higher in the European vs. the Asian breast cancer patients ($p<0.05$). The *BRCA2* gene mutations (c.2127T>C, c.2410G>A) have been classified as neutral polymorphisms due to high population frequencies of their minor alleles (>30%), while possible pathogenicity of these polymorphisms in certain ethnic populations cannot be excluded (Akilzhanova et al., 2013; Suleymanov et al., 2022). The results of this study of the frequency and spectrum of mutations in *BRCA1* and *BRCA2* genes in the Kazakh population have a number of limitations for extrapolation to the entire Kazakh population. This is primarily due to insufficient sample size and lack of ethnic stratification as only race, not nationality, was taken into account. Consequently, the frequency and spectrum of mutations in *BRCA1* and *BRCA2* genes in the Kazakh population remain unstudied to date (Nuruev et al., 2023).

Minor T allele population frequency of *rs4987117* polymorphism of the *BRCA2* gene in the investigated sample of 1,800 Kazakhs was 1.1%, which fully matched the similar frequency in the control group of Kazakhs – 1.1% ($p>0.05$). The minor G allele carriage frequency of *rs144848* polymorphism of the *BRCA2* gene in Kazakhs was 25.0%, which had no significant difference with its frequency in Kazakh control group – 25.0% and in women with breast cancer of Kazakh nationality – 25.0% ($p>0.05$). The genetic distances in minor allele frequency of studied SNPs of *GWAS* gene polymorphisms associated with breast cancer in ethnically homogenous Kazakh population underwent calculation in Plink software and cluster analysis with dendrogram constructed. The authors carried out cluster analysis of *BRCA1* and *BRCA2* gene polymorphism frequencies in eight populations using the Ward method. The researchers counted the mean values of individual variables in clusters for all available observations with calculations of the squares of the Euclidean distances from individual observations in each cluster to the cluster mean. Those clusters, which give the smallest increase in the total sum of distances, accumulated into one new cluster.

The resulting dendrogram (Figure 1) shows two main clusters, India (IND) and Pakistan (PAK), grouped in the first cluster, forming a cluster with a minimum Euclidean distance of 0.06, followed by Bangladesh (BNG) at a distance of 0.144 to the India-Pakistan cluster. Kazakhstan (KZ) and Japan (JAP) entered the second cluster into a sub-cluster with a distance of 0.102. A second sub-cluster took place between China (CHI) and Vietnam (VNM) at 0.112 then joined by the Europe Union (EU) cluster at 0.186. The second cluster as a whole covers a distance of 0.263 and the union of the two clusters is at a distance of 0.887.

The findings of cluster analysis indicate that minor allele frequencies of the studied SNPs associated with breast cancer in the Kazakh population are closest to those in populations of Japan and further Europe and East Asia. The results favour for the high genetic heterogeneity of the studied polymorphisms, which reflects the specifics of the Kazakh population structure resulted from complex evolutionary and migration processes, as well as the median geographic location between the populations of Asia and Europe. The findings of the comparative analysis of the population frequencies of alleles and genotypes of polymorphisms in *BRCA1* and *BRCA2* genes, associated with the risk of breast cancer in the Kazakh population with the previously studied populations of the world have shown genetically heterogeneous mutational spectrum and variable distribution of polymorphisms. That reflects the specifics of population structure and evolutionary-genetic mechanisms of the Kazakh populations. To select significant polymorphisms in *BRCA1* and *BRCA2* genes associated with breast cancer risk, the authors used international database of population frequency and clinical significance of mutations, results of genome-wide associative GWAS studies, panels of known significant pathogenic mutations in *BRCA1* and *BRCA2* genes and official genotyping panels of European and Russian populations to identify BC risk groups.

The SNPs with a minimum population frequency in the Kazakh population (MAF) of less than 0.01 and with significant deviations from Hardy-Weinberg equilibrium ($P<0.05$) were excluded from the study. Twelve single-nucleotide polymorphisms made selection for further comparative analysisof population frequencies: In the *BRCA1* gene – *rs1799950, rs4986850, rs16942, rs1800709, rs4986852, rs1799966, rs1799949, rs799917,* in the *BRCA2* gene – *rs766173, rs144848, rs4987117* and *rs1801426*. The results of comparative analysis of population frequencies of alleles and genotypes for 12 polymorphisms of *BRCA1* and *BRCA2* genes associated with the risk of breast cancer indicate high genetic heterogeneity of studied polymorphisms. This reflects the specifics of the Kazakh population structure formed by complex evolutionary and migration processes as well as the median geographic location between populations of Asia and Europe.

In conclusion, the genotype distribution of 12 studied *BRCA1* and *BRCA2* gene polymorphisms in the Kazakh population was analyzed using PLINK HWE-test software, revealing no significant deviations from Hardy-Weinberg equilibrium ($p>0.05$). This suggests genetic stability in these polymorphisms among Kazakh individuals. Most of these polymorphisms showed intermediate or lower minor allele frequencies in Kazakh individuals compared to global populations, indicating a distinct breast cancer risk profile in the Kazakh population, potentially associated with triple-negative and hereditary breast cancer. However, the genetic contribution of these polymorphisms to breast cancer risk in Kazakh individuals appears limited when compared to European and Asian populations. The inconsistency in the results of studies of the associations of *BRCA1/2* polymorphisms with breast cancer risk in different populations suggests that their effects are specific to certain ethnic groups. This makes it difficult to develop a universal genetic testing panel and requires caution when interpreting the results of screening for polymorphism

carriers among Kazakh women. Large-scale association studies in the Kazakh population are needed to determine the contribution of specific *BRCA1/2* polymorphisms to hereditary breast cancer susceptibility among Kazakh women. This will allow us to develop an optimal strategy for genetic counseling regarding individual risk. Some polymorphisms exhibited higher minor allele frequencies in Kazakh individuals, such as *rs766173* in *BRCA2*, and these may have a greater impact on breast cancer risk among Kazakh women, suggesting their relevance in genetic testing for this population. Certain polymorphisms had unique allele frequencies in Kazakh individuals compared to other populations, underscoring the need for further genotyping in Kazakh breast cancer patients to better understand their potential role in disease development. Comprehensive knowledge of *BRCA1* and *BRCA2* gene mutations within the Kazakh population will enhance breast cancer screening strategies, offering a cost-effective and accessible genetic testing approach for Kazakhstan's citizens.

Author Contribution Statement

All authors contributed equally in this study.

Acknowledgements

The work was fulfilled within the Scientific and Technical Program "National Program for Introduction of Personalised and Preventive Medicine in the Republic of Kazakhstan", OR12165486, ID NCT05121415 and research project "Genetic Research on Pre-eclampsia in Central Asian and European Populations" (InterPregGen) 7th Framework Programme of the European Commission Grant Agreement No. 282540.

References

Akilzhanova A, Nyshanbekkyzy B, Nurkina Zh, Shtephanov I (2013). *BRCA1* and *BRCA2* Gene Mutations Screening in Sporadic Breast Cancer Patients in Kazakhstan. *Cent Asian J Glob Health*, **2**, 29.

Akkuzu, MZ, Küçüköner M, İrtegün S, et al (2019). Presence of Common Polymorphism Mutations in *BRCA1* and *BRCA2* in Breast Cancer in Our Region. *Dicle Med J*, **46**, 623-31.

AL-Eitan LN, Rababa'h DM, Alghamdi MA, Khasawneh RH (2019). Correlation between Candidate Single Nucleotide Variants and Several Clinicopathological Risk Factors Related to Breast Cancer in Jordanian Women: A Genotype-Phenotype Study. *J Cancer*, **10**, 4647-54.

Alyahri N, Abdi S, Khan W, et al (2019). Novel Associations between BRCA1 Variants C.181 T>G (Rs28897672) and Ovarian Crisk in Saudi Females. *J Med Biochem*, **38**, 13-21.

Antoniou A, Pharoah PDP, Narod S, et al (2003). Average risks of breast and ovarian cancer associated with *BRCA1* or *BRCA2* mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet*, **72**, 1117-30.

Badalova AT, Aliyev EM, Gasimova ASh, et al (2020). ANtioxidant and lymphostimul1ng activities of blackberry extract (Rubus caucasicus F.) in experimental diabetes. Azerbaijan. *Pharm Pharmacother J*, **20**, 19-23.

Berndt SI (2013). Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet*, **45**, 501-12.

Carter NJ, Marshall ML, Susswein LR, Zorn KK (2018). Germline pathogenic variants identified in women with ovarian tumors. *Gynecol Oncol*, **151**, 481-8.

Consortium of Investigators of Modifiers of BRCA1/2. 2022. http://cimba.ccge.medschl.cam.ac.uk/.

Debniak T, Scott R, Górski B (2018). BRCA1/2 mutations are not a common cause of malignant melanoma in the Polish population. *PLoS One*, **13**, e0204768.

Fernandes VC, Golubeva VA, Di Pietro G, Shields C (2019). Impact of amino acid substitutions at secondary structures in the BRCT domains of the tumor suppressor BRCA1: Implications for clinical annotation. *J Biol Chem*, **294**, 5980-92.

Genome Aggregation Database (2022). https://gnomad.broadinstitute.org.

Ghafouri-Fard S, Dianatpour A, Faramarzi S (2018). Meta-analysis of BRCA1 polymorphisms and breast cancer susceptibility. *Klin Onkol*, **31**, 330-8.

Gonzalez J, Gaitán M, Molina M (2021). Genetic mutations associated with hereditary breast cancer in Nicaraguan women. *Revista Torreón Universitario*, **29**, 112-8.

Gou W, Wang Y, Zhang W (2017). Polymorphism analysis of rs16941 and rs16942 loci of BRCA1 gene Uighur and Han sporadic breast cancer. *Chin J Clin Exp Pathol*, **12**, 46-50.

Mahdavi M, Nassiri M, Kooshyar MM, Vakili-Azghand M (2018). Hereditary breast cancer; Genetic penetrance and current status with BRCA. *J Cell Physiol*, **234**, 5741-50.

Maslii Y, Garmanchuk L, Ruban O, et al (2021). Preclinical study of anti-proliferative effects of ascorbic acid in combination with lysozyme hydrochloride on cultured cells. *Azerbaijan Pharm Pharmacother J*, **21**, 70-9.

Nuruev M, Sakibaev K, Dzholdosheva G, Maksimova K, Kanymgul AK (2023). Features of circumference sizes in women of different constitutional groups. *Clin Physiol Funct Imaging*, **43**, 40-6.

Pashayan N, Morris S, Gilbert FJ, Pharoah PDP (2018). Cost-effectiveness and benefit-to-harm ratio of risk-stratified screening for breast cancer: a life-table model. *JAMA Oncol*, **4**, 1504-10.

Paulo CM, Lyra-Junior G, Tessarollo IS, Guimarães TB, Henriques DZ (2017). GWAS in Breast Cancer. In: Breast Cancer – From Biology to Medicine. London: Intech. pp 99-117.

Pekur DV, Khmil DN, Bacherikov YYu, et al (2023). Investigation of gamma-ray sensitivity of YAG:Ce based scintillation structures. *Semicond Phys Quantum Electron*, **26**, 89-96.

Pestun İV (2021). Research of self-medication among population in Ukraine. *Azerbaijan Pharm Pharmacother J*, **21**, 34-41.

Purcell S (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, **81**, 559.

Schapovalova O, Gorlova A, de Munter J, et al (2022). Immunomodulatory effects of new phytotherapy on human macrophages and TLR4- and TLR7/8-mediated viral-like inflammation in mice. *Front Med*, **9**, 952977.

Suleymanov TA, Balayeva EZ, Mammadov FI, Huseynova GH (2022). Development and validation of a method for the determination of the amount of immunosuppressants in the blood by high-performance liquid chromatography. *Azerbaijan Pharm Pharmacother J*, **22**, 5-12.

The Global Cancer Observatory (2023). https://www.who.int/news-room/fact-sheets/detail/breast-cancer.

Urbina-Jara L, Rojas-Martinez A, Martinez-Ledesma E (2019). Land scape of Germline Mutations in DNA Repair Genes for

Breast Cancer in Latin America: Opportunities for PARP-Like Inhibitors and Immunotherapy. *Genes*, **10**, 768.
World Health Organization (2023). https://www.who.int/news-room/fact-sheets/detail/breast-cancer.
Xu GP, Zhao Q, Wang D, Xie WY (2018). The association between BRCA1 gene polymorphism and cancer risk: a meta-analysis. *Oncotarget*, **9**, 8681-94.
Zhang H (2020). Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat Genet*, **52**, 572-81.