Breast Cancer Research and Treatment (2020) 181:465–473
https://doi.org/10.1007/s10549-020-05643-0

EPIDEMIOLOGY

Evaluation of pathogenetic mutations in breast cancer predisposition genes in population-based studies conducted among Chinese women

Chenjie Zeng[1] · Xingyi Guo[1] · Wanqing Wen[1] · Jiajun Shi[1] · Jirong Long[1] · Qiuyin Cai[1] · Xiao-Ou Shu[1] · Yongbin Xiang[2] · Wei Zheng[1,3]

Received: 19 February 2020 / Accepted: 10 April 2020 / Published online: 21 April 2020
© The Author(s) 2020

Abstract

Purpose Limited studies have been conducted to evaluate pathogenetic mutations in breast cancer predisposition genes among Chinese women. To fully characterize germline mutations of these genes in this population, we used the whole-exome sequencing data in a population-based case–control study conducted in Shanghai, China.

Methods We evaluated exonic, splicing, and copy number variants in 11 established and 14 candidate breast cancer predisposition genes in 831 invasive breast cancer cases and 839 controls. We identified 55 pathogenic variants, including 15 newly identified in this study.

Results Approximately 8% of the cases and 0.6% of the cancer-free controls carried these pathogenetic variants ($P = 3.05 \times 10^{-15}$). Among cases, 3.7% had a *BRCA*2 pathogenic variant and 1.6% had a *BRCA1* pathogenic variant, while 2.5% had a pathogenic variant in other genes including *ATM, CHEK2, NBN, NF1, CDH1, PALB2, PTEN, TP53* as well as *BARD1, BRIP,* and *RAD51D*. Patients with *BRCA*1/2 pathogenic variants were more likely to have a family history of breast cancer and hormone receptor negative tumors compared with patients without pathogenic variants.

Conclusions This study highlighted the importance of hereditary breast cancer genes in the breast cancer etiology in this understudied population. Together with previous studies in East Asian women, this study suggested a relatively more prominent role of *BRCA2* compared to *BRCA1*. This study also provides additional evidence to design cost-efficient genetic testing among Chinese women for risk assessment and early detection of breast cancer.

Keywords Clinical genetic testing · Breast cancer risk · Chinese women · Hereditary breast cancer syndromes

Introduction

Breast cancer is the most commonly diagnosed cancer and the leading cause of cancer deaths in Chinese women [1]. Approximately 170,000 new invasive breast cancer patients are diagnosed and 45,000 patients die of this cancer in China each year [1]. The average age at breast cancer diagnosis in among Chinese women is 45–55 years, which is approximately 10–15 years younger than that in the United States. Additionally, our previous study and other studies in Chinese patients found a higher proportion of hormone receptor-negative tumors compared with those in the United States [2, 3]. These together suggested differences in the distributions of risk factors between Chinese and their American counterparts.

Breast cancer risk is strongly influenced by genetic factors. To date, multiple breast cancer predisposition genes have been identified, mainly from studies of women of

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10549-020-05643-0) contains supplementary material, which is available to authorized users.

✉ Wei Zheng
wei.zheng@vanderbilt.edu

1 Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA

2 State Key Laboratory of Oncogene and Related Genes &, Department of Epidemiology, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China

3 Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, 2525 West End Ave, Suite 800, Nashville, TN 37203, USA

European ancestry [4–9], including *BRCA1, BRCA2, PALB2*, *CHEK2,* and *ATM*, which together accounted for 25% of the familial risk [5–9]. In addition, variants in the genes *TP53, CDH1*, *PTEN*, *STK11,* and *NF1* that cause multiple types of cancers also documented in some breast cancer patients [4]. It is clinically important to identify patients carrying pathogenic variants of breast cancer predisposition genes as it informs breast cancer risk management strategies in patients [10] and enables cascade genetic testing to identify high-risk family members. It may also inform treatment decisions. Recent studies found that BRCA status predicted response to platinum-based chemotherapy [11] and poly (ADP-ribose) polymerase inhibitors [12, 13]. The selection criteria for genetic testing in patients are in general well defined in the clinic guidelines used in the United States and other developed countries. For example, the National Comprehensive Cancer Network guideline suggested that the selection be based on age of onset, family history of relevant malignancies, hormone receptor and human epidermal growth factor receptor 2 (HER2) status, and race/ethnical groups [14]. A number of methods predicting probabilities of carrying *BRCA1/2* pathogenic variants in patients were also developed [15–18]. Currently, guidelines for genetic testing in breast cancer patients have not been established in China. Developing new guidelines or adapting existing guidelines that were developed in other populations to the Chinese population requires sufficient knowledge on the prevalence of pathogenic variants in breast cancer predisposition genes and predictors of these variants in this population. However, such information is limited, particularly for genes other than *BRCA1/2*.

To better understand the impact of pathogenic variants in breast cancer predisposition genes on breast cancer risk and their association with clinical factors, we used whole-exome sequencing data obtained from participants included in a population-based case–control study conducted in Shanghai, China to detect germline variants including single-nucleotide polymorphisms (SNPs), insertions and deletions (indels), and copy number variants (CNVs) in 11 established and 14 candidate breast cancer predisposition genes. We aimed to describe the variant spectrum of breast cancer predisposition genes and estimate the percentage of pathogenic variant carriers among cases and controls and to determine clinical factors that are associated with pathogenic variants in patients in this population.

Materials and methods

Study populations

This study comprised 839 invasive breast cancer cases and 839 cancer-free controls. Both cases and controls came from population-based studies conducted in urban Shanghai, China, including the Shanghai Breast Cancer Studies (SBCS-I and SBCS-II) and the Shanghai Women's Health Study (SWHS). The details of these studies have been previously described [19–22]. Briefly, for the SBCS-I, which was a case–control study, participants were recruited between 1996 and 1998. Breast cancer patients were ascertained through a rapid case-ascertainment system and the population-based Shanghai Cancer Registry. Controls were randomly selected from the general population using the Shanghai Resident Registry, a population registry containing demographic information for all residents of urban Shanghai. The inclusion criteria for controls were identical to those for cases, with the exception of a breast cancer diagnosis. Using a protocol similar to that of the SBCS-I, the SBCS-II recruited incident breast cancer cases and community controls between 2002 and 2005. The SWHS study was a population-based cohort study conducted in Shanghai with baseline surveys conducted from 1996 to 2000. Breast cancer cases were ascertained through a combination of record linkage with data from the Shanghai Cancer Registry or home visits conducted every 2 to 4 years. Medical charts and pathology slides from diagnostic hospitals were also reviewed to further verify the cancer diagnosis. The protocols for these three studies were approved by their relevant institutional review boards, and all participants provided written informed consent.

For the SBCS-I and SBCS-II, we selected cases from those who were diagnosed with breast cancer at age 58 years or younger. We also included all bilateral breast cancer cases at diagnosis ($n = 13$). Controls were selected from those without a first-degree family history of breast cancer or ovarian cancer. The age at diagnosis of cases ranged from 28 to 58 years old and the age at interview for controls ranged from 29 to 65 years old. A total of 570 cases and 570 controls were selected for the SBCS. For the SWHS, which is a prospective cohort, we selected cases with age at diagnosis not older than 55 years. We selected controls that were on average 5 years older than cases at the time of last follow-up to reduce the possibility of including mutation carriers in the group. The age at diagnosis of cases ranged from 40 to 55 years old and the age at the last follow-up for controls ranged from 42 to 70 years old. A total of 269 cases and 269 controls were selected from the SWHS.

Whole-exome sequencing and quality control

Whole-exome sequencing of study participants was conducted using Illumina GAII sequencing platforms. The median sequence read depth was ~ 50X. The processing of the sequencing data has been previously described [23–25]. Briefly, paired-end reads were aligned to the human genome reference hg19 using BWA (version 0.75). Base quality

recalibration and variant calling were performed using the Picard and the GATK (version 3.8) tools according to the GATK best practice guideline. Copy number variants were predicted according to our in-house pipeline [26, 27]. Briefly, we called CNV using tool including xHMM [28] and CoNIFER [29], and combined results from both tools. We removed variants with low depth of coverage (average read depth < 8) and high rate of missingness (> 2%). We conducted principal components analyses (PCAs) to remove genetic outliers using EIGINSTART based on approximately 1000 ancestry information markers as previously described [27]. We also estimated pair-wise proportion of identity by-descent (IBD) to exclude genetically identical samples, unexpected duplicate samples, and close relatives from the study. After removing 8 samples that did not pass the QC criteria, a total of 831 cases and 839 controls remained for analyses.

Based on the clinical criteria of the breast cancer diagnosis and treatment guidelines and specifications by the Chinese Cancer Society (V2015) and several previous studies, we defined patients at a high familial risk meeting any of the following criteria: (1) age at diagnosis < 35 years old; (2) bilateral breast cancer and age at diagnosis < 50 years old; (3) first-degree family history of breast or ovarian cancer.

Functional annotation and classification of variants

Functional annotations of the identified variants were performed relative to the full set of Refseq genes, as obtained from the UCSC Genome Browser (March 2019). Genes evaluated in this study included 11 established breast cancer genes (*BRCA1, BRCA2, TP53, CDH1, PTEN, STK11, NF1, PALB2, CHEK2, ATM,* and *NBN*) and 14 candidate breast cancer genes (*ATK1, BARD1, BRIP1, CHEK1, FAM175A, FANCM, GEN1, MRE11A, RAD51B, RAD51C, RAD51D, RECQL, RINT1,* and *XRCC2*). We identified all exonic, splicing, and copy number variants in these genes. We evaluated these variants in large genomic databases including the1000 Genomes and the The Genome Aggregation Database (gnomAD) projects (version 2.1.1) if reported. Common variants identified in this study (minor allele frequency (MAF) > 1%) or reported in any public genomic database in any population with a MAF > 1% were excluded. We defined known pathogenic variants according to the following criteria: 1) pathogenic or likely pathogenic (P/LP) variants with ClinVar (last accessed in March 2019) star 2 + (i.e., multiple submitters with assertion criteria and no conflicts, expert panel or practice guideline); 2) pathogenic variants in the BRCA exchange databases (version 23, March 2019); 3) pathogenic or likely pathogenic in breast cancer by at least two submitters according to the LOVD database (version 3.0, last accessed in March 2019). We searched literature on association analyses in family or population-based studies and experimental studies on functional impact of variants in genes of interest. Additionally, in silico analyses using algorithms including CADD [30], DANN [31], SIFT [32], PolyPhen2 [33], Variant Taster [34], Variant assessor [35, 36], FATHMM-MKL [37], GERP + + [38], PhyloP [39], and SiPhy [40] were performed. All the annotations were performed using ANNOVAR [41, 42]. According to the ACMG guideline [43, 44], based on a scoring system including minor allele frequencies across different populations, prior reports of disease association/pathogenicity, experimental studies, and number of carriers in cases and controls, we classified variants into three categories: pathogenic variants (including pathogenic and likely pathogenic variants according to the guideline), benign (benign and likely benign), and variants of unknown significance (VUS).

Statistical analysis

We calculated the percentage of pathogenic variant carriers in cases and controls. We combined the risk alleles of pathogenic variants each gene and performed a burden test evaluating association of a mutated gene of interest with breast cancer risk. We estimated 2-sided *P* values using the Fisher's exact test.

Results

Demographic characteristics of cases and controls and clinical characteristics of cases are presented in Table 1. The mean age at diagnosis of cases was 46.3 years old (standard deviation (SD) = 5.2 years), and the mean age at reference of controls was 48.9 years old (SD = 7.9 years). Eight percent of patients and 2% of controls reported a family history of breast cancer. Approximately 25% of the patients had hormone receptor-negative tumors.

We identified 767 rare exonic, splicing variants, and large deletions in the 25 genes of interest. Among these variants, we identified 55 pathogenic variants in established breast cancer genes including *ATM*, *BRCA1*, *BRCA2*, *CDH1*, *CHEK2*, *NBN, NF1, PALB2, PTEN,* and *TP53* and 5 pathogenic variants in candidate breast cancer genes including *BARD1*, *BRIP1,* and *RAD51D* (Supplementary Table S1), including 31 frameshift indels, 18 nonsense variants, 6 missense variants, 2 splicing variants, and 3 large deletions (Fig. 1). Fifteen of these variants were newly identified in this study. Six variants were recurrent variants, including the known founder variants in Han Chinese, *BRCA1* c.5470_5477del ($n = 3$) and *BRCA2* c.C3109T ($n = 2$). We did not identify any pathogenic variant in the remaining 12 genes including *STK11, ATK1, CHEK1, FAM175A, FANCM, GEN1, MRE11A, RAD51B, RAD51C, RECQL, RINT1,* and *XRCC2*. Additionally, we identified 291 variants of unknown

Table 1 Demographic and clinical characteristics of study participants in the Shanghai Breast Cancer Genetic Study

Characteristics	Cases	Controls
Number of participants	831	839
Age		
Mean (SD)	46.3 (5.2)	48.9 (7.9)
< 35	17 (2%)	15 (2%)
35–44	339 (41%)	309 (37%)
45–54	448 (54%)	303 (36%)
55–	27 (3%)	212 (3%)
Family history of breast cancer*		
Yes	70 (8%)	16 (2%)
No	761 (92%)	823 (98%)
Family history of any cancer		
Yes	555 (68%)	501 (60%)
No	175 (21%)	235 (25%)
Molecular subtypes		
ER + /PR +	322 (39%)	NA
ER + /PR− or ER-/PR +	131 (16%)	NA
ER-/PR-	188 (23%)	NA
ER-/PR-/Her 2 +	21 (3%)	NA
ER-/PR-/Her 2-	49 (6%)	NA

SD standard deviation; *ER* estrogen receptor; *PR* progesterone receptor; *HER2* human epidermal growth factor receptor 2

*Among first-degree relatives

significance (VUS), 212 of which were newly identified in this study.

Among 831 patients with breast cancer, 65 (7.8%) carried a pathogenic variant, while among 839 cancer-free controls, 5 (0.6%) carried a pathogenic variant (Fig. 1, Table 2, Table S2). The gene contributing most to the inherited breast cancer risk was *BRCA2* ($P = 3 \times 10^{-10}$), accounting for 3.7% of the patients, followed by the gene *BRCA1* that accounted for an additional 1.6% of the patients ($P = 0.01$). Furthermore, 21 patients (2.5%) had a pathogenic variant in other genes including *ATM* ($n = 4$), *CHEK2* ($n = 2$), *NF1* ($n = 2$), NBN ($n = 1$), *CDH1* ($n = 1$), *PALB2* ($n = 1$), *PTEN* ($n = 1$), and *TP53* ($n = 1$) as well as *BARD1* ($n = 2$), *BRIP1* ($n = 2$), and *RAD51D* ($n = 4$) (Table 2).

We next evaluated whether clinical factors including age at diagnosis, family history of breast cancer or any cancer, and hormone receptor status of the tumor were associated with carrying a pathogenic variant in patients (Table 3). Factors that statistically significantly associated with carrying *BRCA1* pathogenic variants included a family history of breast cancer ($P = 0.001$) and hormone receptor-negative tumors ($P = 0.0004$) including triple-negative diseases ($P = 0.01$), while the only factor that was statistically significantly associated with carrying *BRCA2* pathogenic variants was a family history of breast cancer ($P = 0.0002$). No statistically significant association was found between age at diagnosis and *BRCA1* or *BRCA2* mutations. No factor

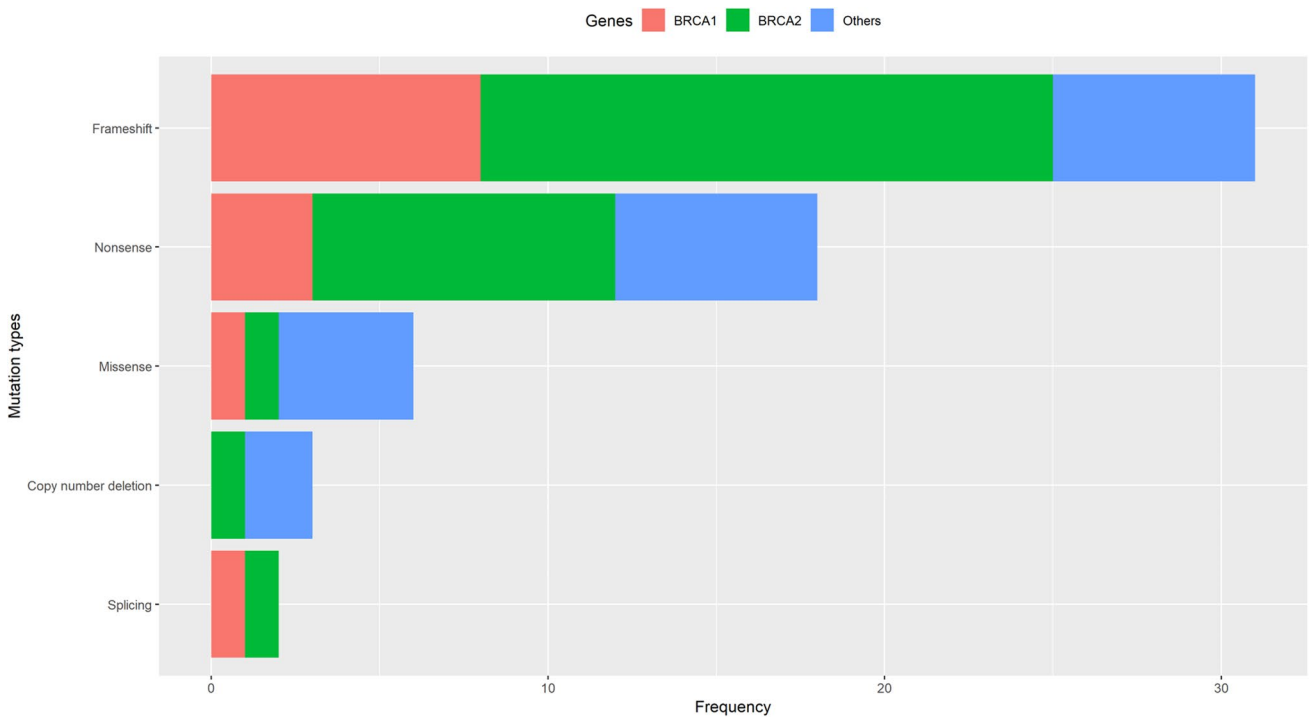

Fig. 1 Distribution of types of pathogenic mutations in the Shanghai Breast Cancer Genetics Study

Table 2 Frequencies of pathogenic variants in established and candidate breast cancer genes identified in cases ($n=831$) and controls ($N=839$) in the Shanghai Breast Cancer Genetic Study

Breast cancer predisposition genes*	Cases		Controls		*P*
	No. of carriers	%	No. of carriers	%	
Total	65	7.8	5	0.6	3.05×10^{-15}
Established genes	57	6.6	3	0.4	1.06×10^{-14}
BRCA1	13		3		
BRCA2	31		0		
PALB2	1		0		
PTEN	1		0		
TP53	1		0		
CDH1	1		0		
ATM	4		0		
CHEK2	2		0		
NBN	1		0		
NF1	2		0		
Candidate genes	8	1.2	2	0.2	0.12
BARD1	2		0		
BRIP1	2		0		
RAD51D	4		2		

*No pathogenic variant was identified in genes: *STK11, ATK1, CHEK1, FAM175A, FANCM, GEN1, MRE11A, RAD51B, RAD51C, RECQL, RINT1,* and *XRCC2*

was statistically significantly associated with carrying pathogenic variant in other breast cancer predisposition genes.

Discussion

By analyzing germline variations in breast cancer predisposition genes in population-based case–control studies using whole-exome sequencing, we found that 7.8% of cases and 0.6% of controls carried at least one pathogenic variant in 10 established and 3 candidate breast cancer predisposition genes. Specifically, among 831 cases, we found that 3.7% had a *BRCA2* pathogenic variant, 1.5% had a *BRCA1* pathogenic variant, and 2.5% had a pathogenic variant in any of the 11 other genes. Clinical factors associated with *BRCA1* pathogenic variants included a family history of breast cancer and hormone receptor status of the tumor, while for BRCA2, the only statistically significant clinical factor was a family history of breast cancer. Results from this study suggested an important role of pathogenetic variants in breast cancer risk in the general population in China. Our findings should be helpful in developing guidelines for identifying high-risk women for genetic risk evaluation of breast cancer.

Although the percentage of overall pathogenic variant carriers in our study was in general comparable to studies conducted in women of European or African ancestry [45], there were remarkable differences in contributions to breast cancer risk by each breast cancer predisposition gene among populations. For example, in our study, the gene contributing most to the breast cancer risk was *BRCA2* while in studies of European ancestry or African ancestry, it was *BRCA*1. This discrepancy has been observed in previous studies in both selected and unselected populations in China [46, 47] and in other East Asian countries including Japan and Korea [48, 49]. These indicated differences in the distribution of genetic risk factors across racial groups.

In our study, more than 2% of the patients carried pathogenic variants in genes other than *BRCA*1/2, supporting the need of using multigene panel testing in breast cancer patients in China. Currently, most of the recommendations for genetic testing for non-BRCA1/2 genes are based on specific cancer syndromes with well-studied clinical features, such as the Li-Fraumeni syndrome and Cowden syndrome. Without multigene panel testing, variants in these genes could have been neglected. For example, in this study, we identified one patient with a *TP53* pathogenic variant, whom could have been missed according to the current guidelines for genetic testing in China.

We also identified 291 VUS in nearly one-third of cases in this study, which were comparable to other previous studies. Some of these VUS are likely to be reclassified as pathogenic variants with more data in the future [50]. Given the clinical implications of genetic testing, accurate variant interpretation of VUS are critical for personalized management of cancer patients and informing cascade genetic testing among family members.

Our study has several strengths. First, the population-based study design was less susceptible to ascertainment biases that were commonly found in clinics-based studies that recruited high-risk patients. Second, whole-exome sequencing enabled us to detect large CNVs that were not routinely detected in target sequencing that is commonly used in clinical setting.

The limitations of this study included incomplete data on breast cancer subtypes. Nearly three-quarter of patients did not have data on HER2 status and nearly one-quarter did not have data on hormone receptor status. Therefore, we were not able to fully evaluate associations of breast cancer subtypes with pathogenic variants in this study. All patients included in this study were diagnosed with breast cancer at age 58 years or younger, and thus the prevalence of pathogenetic variants reported in this study is for a relatively young patient population. There were some differences in selecting controls between the SWHS and the SBCS studies. We selected older controls from the SWHS, a prospective study, to reduce the possibility of including mutation carriers in the

Table 3 Clinical factors associated with carrying pathogenic variants in *BRCA1/2* and other breast cancer predisposition genes in breast cancer cases (n = 831) in the Shanghai Breast Cancer Genetic Study

Characteristics	No. of patients	None[a]	*BRCA1* carriers	*BRCA2* carriers	Other carriers	*P*		
						BRCA1 carriers vs none	*BRCA2* carriers vs none	Other carriers vs none
Age at diagnosis						0.98	0.39	0.18
Mean (SD)	–	46.3 (5.2)	46.3 (4.2)	45.3 (5.2)	47.9 (5.4)			
Range	–	28–58	39–52	33–55	34–56			
< 35	17 (2%)	15 (2%)	0 (0%)	1 (3%)	1 (5%)			
35–45	342 (41%)	317 (41%)	5 (38%)	15 (48%)	5 (23%)			
> 45	472 (56%)	434 (56%)	8 (62%)	15 (48%)	15 (71%)			
Family history of breast cancer						0.001	0.0002	0.05
Yes	70 (8%)	52 (7%)	5 (38%)	9 (29%)	4 (19%)			
No	761 (92%)	714 (93%)	8 (62%)	22 (71%)	17 (81%)			
Family history of any cancer*						1.00	0.04	0.58
Yes	555 (76%)	509 (75%)	8 (80%)	25 (93%)	13 (68%)			
No	175 (24%)	165 (25%)	2 (20%)	2 (7%)	6 (32%)			
Hormone receptor status*						0.0004	0.50	0.16
Positive	456 (71%)	428 (72%)	3 (23%)	16 (67%)	9 (56%)			
Negative	188 (29%)	163 (28%)	10 (77%)	8 (33%)	7 (43%)			
Triple-negative diseases*						0.01	0.67	0.60
Yes	49 (10%)	42 (8%)	3 (50%)	2 (11%)	2 (18%)			
No	456 (90%)	428 (92%)	3 (50%)	16 (89%)	9 (82%)			

*Among patients with available data only

[a]No pathogenic variants identified

group, while we were not able to do so in the SBCSs. However, this discrepancy did not affect our conclusion, given that 0.6% of the controls carried a pathogenic variant.

In conclusion, this is so far the largest study using whole-exome sequencing to detect the full-variant spectrum of breast cancer predisposition genes in population-based case–control studies in Chinese women. We found that pathogenic variants in breast cancer predisposition genes were highly prevalent in Chinese breast cancer patients. The gene that contributed most to inherited breast cancer risk in this population was *BRCA2,* followed by *BRCA1*, together accounting for more than 5% of the cases. Patients with *BRCA1*/2 pathogenic variants were more likely to have a family history of breast cancer and hormone receptor-negative tumors. Other predisposition genes included *ATM, CHEK2, PALB2, NBN, NF1, CDH1, PTEN, TP53*, accounting for an additional 1.6% of the cases. Additional three candidate genes including *BRAD1, BRIP1*, and *RAD51D* potentially accounted for an additional 1.0% of the cases. This study provides significant data that should be useful in designing cost-efficient genetic testing of breast cancer predisposition genes for risk assessment and early detection among Chinese women.

Acknowledgements The authors thank the study participants and research staff for their contributions and support for this project. We thank Ms. Jing He for whole-exome sequencing data processing, and Mr. Marshal Younger for assistance with editing and manuscript preparation. The data analyses were conducted using the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University.

Funding This study was funded by the grants R01CA158473 and T32 CA160056 from the National Cancer Institute.

Data availability The WES data from 831 breast cancer cases and 839 controls and their clinical characterization in our study have been uploaded to the database of Genotypes and Phenotypes (dbGaP) under Sequence Read Archive (SRA) accession numbers "PRJNA560925" and "PRJNA557488" for sharing the data with the research community.

Compliance with ethical standards

Conflict of interest All authors declare no conflict of interest.

Ethical approval All procedures performed in this study were in accordance with the ethical standards of the Vanderbilt and national research committee and with the 1964 Helsinki declaration and its later amendments.

Informed consent Informed consent was obtained from all individual participants included in the study.

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

References

1. Fan L, Strasser-Weippl K, Li JJ, St Louis J, Finkelstein DM, Yu KD, Chen WQ, Shao ZM, Goss PE (2014) Breast cancer in China. Lancet Oncol 15(7):e279–289. https://doi.org/10.1016/S1470-2045(13)70567-9
2. Sivasubramaniam PG, Zhang BL, Zhang Q, Smith JS, Zhang B, Tang ZH, Chen GJ, Xie XM, Xu XZ, Yang HJ, He JJ, Li H, Li JY, Fan JH, Qiao YL (2015) Breast cancer disparities: a multicenter comparison of tumor diagnosis, characteristics, and surgical treatment in China and the U.S. Oncologist 20(9):1044–1050. https://doi.org/10.1634/theoncologist.2014-0290
3. Su Y, Zheng Y, Zheng W, Gu K, Chen Z, Li G, Cai Q, Lu W, Shu XO (2011) Distinct distribution and prognostic significance of molecular subtypes of breast cancer in Chinese women: a population-based cohort study. BMC Cancer 11:292. https://doi.org/10.1186/1471-2407-11-292
4. Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, Goldgar DE, Evans DG, Chenevix-Trench G, Rahman N, Robson M, Domchek SM, Foulkes WD (2015) Gene-panel sequencing and the prediction of breast-cancer risk. N Engl J Med 372(23):2243–2257. https://doi.org/10.1056/NEJMsr1501341
5. Couch FJ, Shimelis H, Hu CL, Hart SN, Polley EC, Na J, Hallberg E, Moore R, Thomas A, Lilyquist J, Feng BJ, McFarland R, Pesaran T, Huether R, LaDuca H, Chao EC, Goldgar DE, Dolinsky JS (2017) Associations between cancer predisposition testing panel genes and breast cancer. Jama Oncol 3(9):1190–1196. https://doi.org/10.1001/jamaoncol.2017.0424
6. Lu HM, Li SW, Black MH, Lee S, Hoiness R, Wu ST, Mu WB, Huether R, Chen J, Sridhar S, Tian Y, McFarland R, Dolinsky J, Davis BT, Mexal S, Dunlop C, Elliott A (2019) Association of breast and ovarian cancers with predisposition genes identified by large-scale sequencing. Jama Oncol 5(1):51–57. https://doi.org/10.1001/jamaoncol.2018.2956
7. Madanikia SA, Bergner A, Ye XB, Blakeley JO (2012) Increased risk of breast cancer in women with NF1. Am J Med Genet A 158a(12):3056–3060. https://doi.org/10.1002/ajmg.a.35550
8. Antoniou AC, Casadei S, Heikkinen T, Barrowdale D, Pylkas K, Roberts J, Lee A, Subramanian D, De Leeneer K, Fostira F, Tomiak E, Neuhausen SL, Teo ZL, Khan S, Aittomaki K, Moilanen JS, Turnbull C, Seal S, Mannermaa A, Kallioniemi A, Lindeman GJ, Buys SS, Andrulis IL, Radice P, Tondini C, Manoukian S, Toland AE, Miron P, Weitzel JN, Domchek SM, Poppe B, Claes KBM, Yannoukakos D, Concannon P, Bernstein JL, James PA, Easton DF, Goldgar DE, Hopper JL, Rahman N, Peterlongo P, Nevanlinna H, King MC, Couch FJ, Southey MC, Winqvist R, Foulkes WD, Tischkowitz M (2014) Breast-cancer risk in families with mutations in PALB2. New Engl J Med 371(6):497–506. https://doi.org/10.1056/NEJMoa1400382
9. Weischer M, Nordestgaard BG, Pharoah P, Bolla MK, Nevanlinna H, Van't Veer LJ, Garcia-Closas M, Hopper JL, Hall P, Andrulis IL, Devilee P, Fasching PA, Anton-Culver H, Lambrechts D, Hooning M, Cox A, Giles GG, Burwinkel B, Lindblom A, Couch FJ, Mannermaa A, Grenaker Alnaes G, John EM, Dork T, Flyger H, Dunning AM, Wang Q, Muranen TA, van Hien R, Figueroa J, Southey MC, Czene K, Knight JA, Tollenaar RA, Beckmann MW, Ziogas A, Christiaens MR, Collee JM, Reed MW, Severi G, Marme F, Margolin S, Olson JE, Kosma VM, Kristensen VN, Miron A, Bogdanova N, Shah M, Blomqvist C, Broeks A, Sherman M, Phillips KA, Li J, Liu J, Glendon G, Seynaeve C, Ekici AB, Leunen K, Kriege M, Cross SS, Baglietto L, Sohn C, Wang X, Kataja V, Borresen-Dale AL, Meyer A, Easton DF, Schmidt MK, Bojesen SE (2012) CHEK2*1100delC heterozygosity in women with breast cancer associated with early death, breast cancer-specific death, and increased risk of a second breast cancer. J Clin Oncol 30(35):4308–4316. https://doi.org/10.1200/JCO.2012.42.7336
10. Terry MB, Liao Y, Whittemore AS, Leoce N, Buchsbaum R, Zeinomar N, Dite GS, Chung WK, Knight JA, Southey MC, Milne RL, Goldgar D, Giles GG, McLachlan SA, Friedlander ML, Weideman PC, Glendon G, Nesci S, Andrulis IL, John EM, Phillips KA, Daly MB, Buys SS, Hopper JL, MacInnis RJ (2019) 10-year performance of four models of breast cancer risk: a validation study. Lancet Oncol 20(4):504–517. https://doi.org/10.1016/S1470-2045(18)30902-1
11. Isakoff SJ, Mayer EL, He L, Traina TA, Carey LA, Krag KJ, Rugo HS, Liu MC, Stearns V, Come SE, Timms KM, Hartman AR, Borger DR, Finkelstein DM, Garber JE, Ryan PD, Winer EP, Goss PE, Ellisen LW (2015) TBCRC009: a multicenter phase II clinical trial of platinum monotherapy with biomarker assessment in metastatic triple-negative breast cancer. J Clin Oncol 33(17):1902–1909. https://doi.org/10.1200/JCO.2014.57.6660
12. Gunjur A (2018) Talazoparib for BRCA-mutated advanced breast cancer. Lancet Oncol 19(10):e511. https://doi.org/10.1016/S1470-2045(18)30650-8
13. Litton JK, Rugo HS, Ettl J, Hurvitz SA, Goncalves A, Lee KH, Fehrenbacher L, Yerushalmi R, Mina LA, Martin M, Roche H, Im YH, Quek RGW, Markova D, Tudor IC, Hannah AL, Eiermann W, Blum JL (2018) Talazoparib in patients with advanced breast cancer and a germline BRCA mutation. N Engl J Med 379(8):753–763. https://doi.org/10.1056/NEJMoa1802905
14. Daly MB, Pilarski R, Berry M, Buys SS, Farmer M, Friedman S, Garber JE, Kauff ND, Khan S, Klein C, Kohlmann W, Kurian A, Litton JK, Madlensky L, Merajver SD, Offit K, Pal T, Reiser G, Shannon KM, Swisher E, Vinayak S, Voian NC, Weitzel JN, Wick MJ, Wiesner GL, Dwyer M, Darlow S (2017) NCCN guidelines (R) insights genetic/familial high-risk assessment: breast and ovarian, version 2.2017 featured updates to the NCCN guidelines. J Natl Compr Canc Ne 15(1):9–19. https://doi.org/10.6004/jnccn.2017.0003
15. Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B, Gumpper KL, Scholl T, Tavtigian SV, Pruss DR, Critchfield GC (2002) Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. J Clin Oncol 20(6):1480–1490. https://doi.org/10.1200/JCO.2002.20.6.1480

16. Couch FJ, DeShano ML, Blackwood MA, Calzone K, Stopfer J, Campeau L, Ganguly A, Rebbeck T, Weber BL, Jablon L, Cobleigh MA, Hoskins K, Garber JE (1997) BRCA1 mutations in women attending clinics that evaluate the risk of breast cancer. New Engl J Med 336(20):1409–1415. https://doi.org/10.1056/Nejm199705153362002
17. Antoniou AC, Durocher F, Smith P, Simard J, Easton DF, Program IB (2006) BRCA1 and BRCA2 mutation predictions using the BOADICEA and BRCAPRO models and penetrance estimation in high-risk French-Canadian families. Breast Cancer Res 8(1):R3. https://doi.org/10.1186/bcr1365
18. Berry DA, Iversen ES, Gudbjartsson DF, Hiller EH, Garber JE, Peshkin BN, Lerman C, Watson P, Lynch HT, Hilsenbeck SG, Rubinstein WS, Hughes KS, Parmigiani G (2002) BRCAPRO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. J Clin Oncol 20(11):2701–2712. https://doi.org/10.1200/Jco.2002.05.121
19. Zheng W, Long JR, Gao YT, Li C, Zheng Y, Xiang YB, Wen WQ, Levy S, Deming SL, Haines JL, Gu K, Fair AM, Cai QY, Lu W, Shu XO (2009) Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 41(3):324–328. https://doi.org/10.1038/ng.318
20. Kilfoy BA, Zhang Y, Shu XO, Gao YT, Ji BT, Yang G, Li HL, Rothman N, Chow WH, Zheng W (2008) Family history of malignancies and risk of breast cancer: prospective data from the Shanghai women's health study. Cancer Causes Control 19(10):1139–1145. https://doi.org/10.1007/s10552-008-9181-2
21. Zheng W, Wen WQ, Gao YT, Shyr Y, Zheng Y, Long JR, Li GL, Li C, Gu K, Cai QY, Shu XO, Lu W (2010) Genetic and clinical predictors for breast cancer risk assessment and stratification among Chinese women. J Natl Cancer I 102(13):972–981. https://doi.org/10.1093/jnci/djq170
22. Zheng W, Chow WH, Yang G, Jin F, Rothman N, Blair A, Li HL, Wen W, Ji BT, Li Q, Shu XO, Gao YT (2005) The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. Am J Epidemiol 162(11):1123–1131. https://doi.org/10.1093/aje/kwi322
23. Zhang Y, Li B, Li C, Cai Q, Zheng W, Long J (2014) Improved variant calling accuracy by merging replicates in whole-exome sequencing studies. Biomed Res Int 2014:319534. https://doi.org/10.1155/2014/319534
24. Guo X, Long J, Chen Z, Shu XO, Xiang YB, Wen W, Zeng C, Gao YT, Cai Q, Zheng W (2019) Discovery of rare coding variants in OGDHL and BRCA2 in relation to breast cancer risk in Chinese women. Int J Cancer. https://doi.org/10.1002/ijc.32825
25. Guo X, Lin W, Bai M, Li H, Wen W, Zeng C, Chen Z, He J, Chen J, Cai Q, Long J, Jia WH, Shu XO, Zheng W (2019) Discovery of a pathogenic variant rs139379666 (p. P2974L) in ATM for breast cancer risk in Chinese populations. Cancer Epidemiol Biomarkers Prev 28(8):1308–1315. https://doi.org/10.1158/1055-9965.EPI-18-1294
26. Guo X, Shi J, Cai Q, Shu XO, He J, Wen W, Allen J, Pharoah P, Dunning A, Hunter DJ, Kraft P, Easton DF, Zheng W, Long J (2018) Use of deep whole-genome sequencing data to identify structure risk variants in breast cancer susceptibility genes. Hum Mol Genet 27(5):853–859. https://doi.org/10.1093/hmg/ddy005
27. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. PLoS Genet 2(12):e190. https://doi.org/10.1371/journal.pgen.0020190
28. Fromer M, Purcell SM (2014) Using XHMM software to detect copy number variation in whole-exome sequencing data. Curr Protoc Hum Genet 8(1):7–23. https://doi.org/10.1002/0471142905.hg0723s81
29. Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP, Project NES, Quinlan AR, Nickerson DA, Eichler EE (2012) Copy number variation detection and genotyping from exome sequence data. Genome Res 22(8):1525–1532. https://doi.org/10.1101/gr.138115.112
30. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M (2018) CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. https://doi.org/10.1093/nar/gky1016
31. Quang D, Chen Y, Xie X (2015) DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics 31(5):761–763. https://doi.org/10.1093/bioinformatics/btu703
32. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC (2016) SIFT missense predictions for genomes. Nat Protoc 11(1):1–9. https://doi.org/10.1038/nprot.2015.123
33. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods 7(4):248–249. https://doi.org/10.1038/nmeth0410-248
34. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 7(8):575–576. https://doi.org/10.1038/nmeth0810-575
35. Reva B, Antipin Y, Sander C (2011) Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res 39(17):E118–U185. https://doi.org/10.1093/nar/gkr407
36. Reva B, Antipin Y, Sander C (2007) Determinants of protein function revealed by combinatorial entropy optimization. Genome Biol 8(11). doi: 10.1186/gb-2007–8–11-r232
37. Shihab HA, Rogers MF, Gough J, Mort M, Cooper DN, Day IN, Gaunt TR, Campbell C (2015) An integrative approach to predicting the functional effects of non-coding and coding sequence variation. Bioinformatics 31(10):1536–1543. https://doi.org/10.1093/bioinformatics/btv009
38. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S (2010) Identifying a high fraction of the human genome to be under selective constraint using GERP plus. Plos Comput Biol 6(12). doi:10.1371/journal.pcbi.1001025
39. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A (2010) Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 20(1):110–121. https://doi.org/10.1101/gr.097857.109
40. Garber M, Guttman M, Clamp M, Zody MC, Friedman N, Xie XH (2009) Identifying novel constrained elements by exploiting biased substitution patterns. Bioinformatics 25(12):I54–I62. https://doi.org/10.1093/bioinformatics/btp190
41. Yang H, Wang K (2015) Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc 10(10):1556–1566. https://doi.org/10.1038/nprot.2015.105
42. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38(16):e164. https://doi.org/10.1093/nar/gkq603
43. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, Committee ALQA (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 17(5):405–424. https://doi.org/10.1038/gim.2015.30
44. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho YY, Kobayashi Y, Patil N, Thusberg J, Westbrook M, Topper S, Grp ICG (2017) Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria. Genet Med 19(10):1105–1117. https://doi.org/10.1038/gim.2017.37
45. Tung N, Lin NU, Kidd J, Allen BA, Singh N, Wenstrup RJ, Hartman AR, Winer EP, Garber JE (2016) Frequency of germline mutations in 25 cancer susceptibility genes in a sequential series

of patients with breast cancer. J Clin Oncol 34(13):1460. https://doi.org/10.1200/Jco.2015.65.0747

46. Li JY, Jing RL, Wei HY, Wang MH, Qi XW, Liu HX, Liu J, Ou JH, Jiang WH, Tian FG, Yuan S, Li HY, Hong X, Zhang RS, Guan AH, Ke L, Jiang HC, Yu R, He JJ, Huang WW, Ning L, Cai XJ, Jia M, Rui L, Yan X, Hu CY, Zhang JG, Guo BL, Ouyang LZ, Ping S, Liu ZZ, Ling Z, Zhen Z, Ting Z, Xuan ZL, Tan XN, Liang JB, Pan QW, Li C, Fan Z, Fan LJ, Yi Z, Yang XH, Jing BL, Chen CJ, Jun J (2019) Germline mutations in 40 cancer susceptibility genes among Chinese patients with high hereditary risk breast cancer. Int J Cancer 144(2):281–289. https://doi.org/10.1002/ijc.31601
47. Sun J, Meng H, Yao L, Lv M, Bai J, Zhang JG, Wang LT, Ouyang T, Li JF, Wang TF, Fan ZQ, Fan T, Lin BY, Xie YT (2017) Germline mutations in cancer susceptibility genes in a large series of unselected breast cancer patients. Clin Cancer Res 23(20):6113–6119. https://doi.org/10.1158/1078-0432.Ccr-16-3227
48. Choi DH, Lee NH, Bale AE, Carter D, Haffty BG (2004) Incidence of BRCA1 and BRCA2 mutations in young Korean breast cancer patients. J Clin Oncol 22(9):1638–1645. https://doi.org/10.1200/Jco.2004.04.179
49. Karami F, Mehdipour P (2013) A comprehensive focus on global spectrum of BRCA1 and BRCA2 mutations in breast cancer. Biomed Res Int. https://doi.org/10.1155/2013/928562
50. Mersch J, Brown N, Pirzadeh-Miller S, Mundt E, Cox HC, Brown K, Aston M, Esterling L, Manley S, Ross T (2018) Prevalence of variant reclassification following hereditary cancer genetic testing. JAMA 320(12):1266–1274. https://doi.org/10.1001/jama.2018.13152

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