Germline variation in BRCA1/2 is highly ethnic-specific: Evidence from over 30,000 Chinese hereditary breast and ovarian cancer patients

Shanmuga Priya Bhaskaran1, Khyati Chandratre2, Hemant Gupta1, Li Zhang1, Xiaoyu Wang1, Jian Cui2, Yeong C. Kim7, Siddharth Sinha1, Luhan Jiang1, Boya Lu1, Xiaobing Wu1, Zixin Qin1, Teng Huang2 and San Ming Wang1,8

1Cancer Centre and Institute of Translational Medicine, Faculty of Health Sciences, University of Macau, Macau, China
2Eppley Cancer Institute, University of Nebraska Medical Center, Omaha, NE

BRCA1 and BRCA2 play essential roles in maintaining the genome stability. Pathogenic germline mutations in these two genes disrupt their function, lead to genome instability and increase the risk of developing breast and ovarian cancers. BRCA mutations have been extensively screened in Caucasian populations, and the resulting information are used globally as the standard reference in clinical diagnosis, treatment and prevention of BRCA-related cancers. Recent studies suggest that BRCA mutations can be ethnic-specific, raising the question whether a Caucasian-based BRCA mutation information can be used as a universal standard worldwide, or whether an ethnicity-based BRCA mutation information system need to be developed for the corresponding ethnic populations. In this study, we used Chinese population as a model to test ethnicity-specific BRCA mutations considering that China has one of the latest numbers of breast cancer patients therefore BRCA mutation carriers. Through comprehensive data mining, standardization and annotation, we collected 1,088 distinct BRCA variants derived from over 30,000 Chinese individuals, one of the largest BRCA data set from a non-Caucasian population covering nearly all known BRCA variants in the Chinese population (https://dbBRCA-Chinese.fhs.umac.mo). Using this data, we performed multi-layered analyses to determine the similarities and differences of BRCA variation between Chinese and non-Chinese ethnic populations. The results show the substantial differences of BRCA data between Chinese and non-Chinese ethnicities. Our study indicates that the current Caucasian population-based BRCA data is not adequate to represent the BRCA status in non-Caucasian populations. Therefore, ethnic-based BRCA standards need to be established to serve for the non-Caucasian populations.

Introduction

Approximately 10–15% of breast cancer cases are caused by hereditary genetic mutations.1 The most penetrating mutations are those in the BRCA1 and BRCA2 (BRCA) genes,2,3 which are essential for maintaining the genome stability. Women carrying pathogenic mutations in BRCA1 have a 72% lifetime risk of developing breast cancer, while those with BRCA2 mutations have a 69% risk.4 Mutations in BRCA also increase the risk for ovarian cancer, prostate cancer, melanoma and pancreatic cancer. Identification of BRCA mutation carriers before the development of cancer is crucial in order to protect them from developing cancer by taking preventive measures of early cancer surveillance, chemoprevention and preventive surgery.5–9 Extensive efforts have identified a large number of BRCA mutations, mainly in the Caucasian populations of Europe and North America. BRCA databases with well-documented, annotated and freely accessible BRCA information have been developed and used worldwide as the references for diagnosis, treatment and prevention of BRCA associated cancers.

Studies have revealed that human BRCA1 and BRCA2 are rapidly evolving under positive selection to effectively protect genome stability.10,11 Recent studies have further suggested that the variation in human BRCA could be ethnic-specific in different ethnic populations. For example, BRCA variants within Latin American populations are highly heterogeneous,12 and BRCA variants in Asian populations differ substantially from those in other populations.13 Understanding the ethnic-specificity of BRCA
variation is important, as it can provide a precise genetic basis to study the relationship between human evolution and diseases. Furthermore, it will determine whether the Caucasian-based BRCA data is adequate to serve as a universal reference to determine BRCA status in non-Caucasian populations around the world, or whether the ethnicity-based BRCA mutation data should be developed instead. However, the answer remains elusive owing to the lack of BRCA data from most of the non-Caucasian populations. For example, in the recently completed CIMBA study that collected BRCA data from 49 countries across six continents, the data available from single, non-Caucasian ethnic populations remain very limited. To fully prove the existence of ethnic-specific BRCA mutations, a comprehensive data from particular ethnic populations will be required.

The Chinese population is the largest one in the world. Breast cancer is the most common cancer among Chinese women, with 260,000 new breast cancer cases diagnosed and 70,000 mortalities annually. With nearly two decades of BRCA studies in China, in particular in the recent years owing largely to the adoption of next-generation sequencing technologies, BRCA data exclusively derived from the Chinese population are increasingly reported. Thus, the Chinese population can serve as an ideal model to test the presence of ethnic-specific BRCA mutations. Through a comprehensive data mining, standardization and annotation, we collected nearly all BRCA variant data currently available for the Chinese population and developed the data into a public database dbBRCA-Chinese (https://dbBRCA-Chinese.fhs.umac.mo). Using this rich data set, we studied the similarities and differences in BRCA variation between Chinese and non-Chinese populations. The data from our study provides a convincing evidence for the existence of ethnic-specific BRCA mutation. Here, we report detailed information from the study.

**Materials and Methods**

**Variant data collection and analysis**

We searched for resources reporting BRCA variation data from individuals of Chinese ethnicity, including publications in PubMed, the China National Knowledge Infrastructure (CNKI) database (http://oversea.cnki.net/kns55/default.aspx) and WanFang (http://www.wanfangdata.com/CJ/IN/asp#China Online), as well as Chinese-derived BRCA variants in existing BRCA databases. For the collected BRCA variants, we performed extensive standardization and reannotation, following HGVS and ACMG guidelines. The reference sequences used for BRCA analysis were: cDNA NM_007294.3, protein NP_009225.1, genome hg19, BIC cDNA: U14680.1 and BIC protein: AAA73985.1; those used for BRCA2 were: cDNA NM_000059.3, protein NP_000050.2, genome hg19, BIC cDNA: U43746.1 and BIC protein: AAB07231.1, respectively (https://doi.org/10.1093/nar/gkw1070). The transcript-oriented position of each variant was converted to its respective genome position in hg19 using the Position Converter tool in Mutalyzer, and the consistency with HGVS nomenclature was confirmed using the Name Checker DNA tool. The variants were annotated using ANNOVAR with eight reference databases namely: RefGene, dbsNP (version 150), 1000genome, ESP5600, ExAC, ClinVar, InterVar and DBNSFP. Following BRCA databases were used for the comparative analysis: BIC (https://research.nhgri.nih.gov/bic/, accessed February 20, 2018), ClinVar(http://www.ncbi.nlm.nih.gov/clinvar/, accessed February 20, 2018), BRCA Exchange (http://brcaexchange.org, accessed February 20, 2018), ENIGMA (downloaded from BED database, http://brcaexchange.org, accessed February 20, 2018), BMD (http://www.arup.utah.edu/database/BRCA/Home/BRCA1_landing.php,http://www.arup.utah.edu/database/BRCA/Home/BRCA2_landing.php)(https://www.aruplab.com/topics/breast-cancer/brccasebase.php, accessed February 20, 2018), LOVD (http://www.lowd.nl/3.0/home, accessed February 20, 2018) and CIMBA. The Chinese variants present in these databases were classified as known variants; those absent were classified as novel variants and deposited in ClinVar database (accession number nstd165). Five categories based on ACMG guidelines were used: pathogenic, likely pathogenic, uncertain significance, likely benign, and benign for the variant classification of known variants using ClinVar database. For those variants not matched with the variant list using the above-mentioned resources, their classifications were predicted from InterVar database using ANNOVAR annotation tool. For these variants not being able to classified, they were included as “Unclassified” group. BRCA variants in Latin population were extracted from Villarreal-Garza et al., BRCA variants in Asian populations were extracted from, and BRCA variants in the Indian population were extracted from, respectively.
Statistical data analysis
Fisher’s exact test was used to analyze the differences between the Chinese and non-Chinese variant data. \( p < 0.05 \) was considered as significant difference.

Results
Collection of Chinese BRCA data
The majority of Chinese BRCA variation data was gathered from mainland China (90.2%), and the remainder were from Hong Kong (6.4%), Taiwan (1.6%), Singapore (0.9) and Malaysia (0.4%) (Table S1A, S1B, Supporting Information). In total, 31,680 cases with Chinese ethnicity were tested for BRCA mutations and their results were reported between 1999 and 2017; of these, 69.3% were reported between 2016 and 2017. Nearly all BRCA variants were from breast and/or ovarian cancer patients under different clinical criteria, except for four variants that were derived from 1,043 healthy control individuals.32

We collected and summarized the clinical information reported from each of the reference studies (Fig. S1 and Table S2, Supporting Information). The data showed several unique features as follows: 1). BMI is considered as one of the risk factors for breast cancer. However, according to our data, 83.7% of Chinese patients were within the normal range of 18.5–22.9 BMI and 10.5% were seen to have even lower than 18.5 BMI, indicating that obesity had no significant role in increasing risk for breast and ovarian cancer in these Chinese patients; 2). Family history is considered as a high-risk factor for hereditary breast cancer. However, 72.3% of patients from our study did not report any family history of cancer, suggesting that family history was not an essential factor in this disease cohort; 3). Stage II breast cancer cases accounted for 66.2% whereas stage I cases only 6.3%, indicating the lack of earlier diagnosis.

A total of 3,791 BRCA mutation carriers were identified (12%, 2,123 BRCA1 and 1,688 BRCA2), of which 1,978 carriers (52.2%, 990 BRCA1 and 988 BRCA2) were within the clinically reported mutation categories of pathogenic or likely pathogenic. By standardizing and re-annotating all the variants following Human Genome Variation Society (HGVS)-recommended nomenclature, we identified a total of 1,088 distinct BRCA variants (557 in BRCA1 and 531 in BRCA2) of which 519 (47.7%, 278 in BRCA1 and 241 in BRCA2) were recurrent, and 50% were either detected or validated by Sanger sequencing (Tables S4–S5, Supporting Information). Except the 26 variants specifically from Uygur ethnic population, all variants were from Han Chinese population. We developed the dBBRCA-Chinese database as an open source to host the entire set of BRCA variants and their annotation information (https://dBBRCA-Chinese.fhs.umac.mo).

General features of Chinese BRCA data
Age and abundance of variants. The 1,088 BRCA variants were distributed with different frequencies among the 3,791 BRCA variant carriers. The age distribution data show that 52.3% of the cancer developed at early age (<40 years) in these BRCA variant carriers (Table S1C, Supporting Information). About half of the variants were detected only once in single individuals (46.5% in BRCA1, 54.6% in BRCA2), and the rest variants were distributed between 2 to 100 individuals with increased frequencies (Table S1D, Supporting Information).

Pathogenic and nonpathogenic variants. We classified the BRCA variants into five classes—pathogenic, likely pathogenic, uncertain significance, likely benign, and benign—following the American College of Medical Genetics and Genomics (ACMG) standards and guidelines. Pathogenic and likely pathogenic variants, which are clinically reportable, accounted for 46% of BRCA1 variants and 52% of BRCA2 variants. Such higher rates do not necessarily imply high BRCA variation rate in Chinese breast and ovarian cancer patient population but reflect the fact that the cancer patients included in many of the studies were selected from high-risk patients of strong family history or early age of cancer development. Importantly, 13% and 8% of BRCA1 and BRCA2 were classified as of uncertain significance, and 30% of both BRCA1 and BRCA2 variants remained as unclassified variants (Fig. 1a). This fact indicates that much effort needs to be made in order to determine the function of these BRCA variants existing in Chinese population.

Ethnic origins of current BRCA data
We compiled ethnic origins of currently existing BRCA variation data to know the status of BRCA study across ethnic human populations. By combining the BRCA variants from the Breast Cancer Information Core (BIC), ClinVar, BRCA Exchange Database (BED), BRCA1 and BRCA2 Mutation Database (BMD), Leiden Open Variation Database (LOVD) and ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles), we generated a single BRCA variation data set containing 6,343 distinct BRCA1 and 8,884 distinct BRCA2 variants (S6A). Classification of the ethnic origins of these variants showed that 62% were from Caucasian populations and 15% from the Ashkenazi Jewish population. The remaining 23% were originated from non-Chinese Asian (13%), Latino (5%), African (3%), and Chinese (2%) populations (Fig. 1b and Tables S6B, S6C, Supporting Information). Of the 5,925 CIMBA BRCA variants data with defined ethnicities, 80.2% were from Caucasian, 2.4% from Ashkenazi Jews, 10.8% from Asian, 3.8% from Hispanic and 2.8% from African American.15 The analysis shows that the current BRCA variant data contains very limited information from non-Caucasian populations.

We performed the multi-layer analyses to investigate the similarities and differences in BRCA variation between Chinese and non-Chinese populations.

Comparison with existing BRCA data
In order to determine the similarities and differences of BRCA variation between Chinese and non-Chinese populations, we made a comprehensive comparison between Chinese and any available non-Chinese BRCA data as represented below. The rationale for comparing with each database are:
1. GnomAD: It contains extensive normal population variation data collected from human population by the largest exome and whole-genome sequencing projects. The comparison aimed to determine the similarities and differences of normal BRCA variants present between Chinese and non-Chinese populations;

2. BIC, BED, BMD, ClinVar, ENIGMA and LOVD: These are the major BRCA databases, with the data mostly derived from Caucasians as indicated by our analysis (Fig. 1b). Comparison with these databases aimed to determine the similarities and differences of BRCA mutation between Chinese and Caucasian (mostly) populations;

3. CIMBA data: It contains BRCA data collected from 49 countries across six continents. The comparison aimed to determine the similarities and differences between Chinese and worldwide non-Chinese populations including more non-Caucasian data;

Figure 1. BRCA data. (a) Clinical classification of Chinese BRCA data as pathogenic, likely pathogenic, uncertain significance, likely benign, benign and unclassified. The pathogenic and likely pathogenic variants accounted for 49.5%. (b) Relationship between population sizes and their contribution to current BRCA data. The proportions of different human ethnic populations were from the database (http://www.worldometers.info/world-population/), the ethnic origins of BRCA data were from different BRCA databases. It shows that the current BRCA data is not proportional to the human ethnic populations. (c) Comparison of BRCA data between Chinese and non-Chinese populations. A total of 557 BRCA1 and 531 BRCA2 Chinese variants were compared to 6,344 BRCA1 and 8,886 BRCA2 non-Chinese variants compiled from all existing BRCA databases. The results show that 38% of Chinese BRCA variants were present only in the Chinese population.
4. Latin America and the Caribbean data: The data were from Latin American population of Argentina, Bahamas, Brazil, Chile, Colombia, Costa Rica, Cuba, Mexico, Peru, Puerto Rico, Uruguay, Venezuela and the Hispanic population in the United States. The comparison aimed to determine the similarities and differences between Chinese and Latin America populations;

5. Non-Chinese Asian populations: BRCA data are available from Bangladeshi, Filipino, Iranian, Israeli, Japanese, Korean, Lebanese, Malay, Oman, Pakistani, Sri Lankan, Thai and Turkish populations. These populations were genetically and geographically closer to the Chinese population than other non-Chinese populations. The comparison aimed to determine the similarities and differences between Chinese and the non-Chinese Asian populations;

6. Indian population: India has the 2nd largest population in the world, with highly diversified genetic background. Several large-scale BRCA studies were reported recently with substantial BRCA data collected from the Indian patients. The comparison aimed to determine the similarities and differences between Chinese and Indian populations, the two largest populations in the world.

GnomAD. Matching the Chinese BRCA data with those in GnomAD from the largest exome and whole-genome sequence data collection showed that only 76 of the 557 (13.6%) Chinese BRCA1 had matches in 2,476 (3%) BRCA1 variants in GnomAD and 97 of the 531 (18.2%) Chinese BRCA2 had matches in 3,674 variants (2.6%) in GnomAD. The results indicate that the vast majority of the Chinese BRCA variants were not present in the population data provided by current exome and whole-genome sequencing studies (Table 1A and Table S7A, Supporting Information). For those with the matches, their abundance as judged by the East Asia population frequencies were mostly at lower levels [62/76 (81.5%) BRCA1 variants and 73/97 (72.3%) BRCA2 variants <0.001], highlighting their pathogenic potential.

| Origin | BRCA1 variants | BRCA2 variants |
|--------|----------------|----------------|
|        | Total 1 Matched Proportion 2 | Total Matched Proportion |
| A. GnomAD database | 2,496 76 0.136 | 3,674 97 0.182 |
| B. Multiple BRCA databases 3 | 22,095 371 0.666 | 28,648 305 0.574 |
| Total BED BIC BMD Clinvar ENIGMA LOVD | 7,810 347 0.601 | 10,378 277 0.522 |
| 1,702 207 0.359 | 1,916 159 0.299 |
| 1,271 179 0.310 | 1,321 122 0.230 |
| 5,537 355 0.601 | 7,688 286 0.539 |
| 2,712 206 0.357 | 3,442 178 0.335 |
| 3,063 164 0.284 | 3,903 131 0.247 |
| C. CIMBA BRCA variants | 1,651 91 0.163 | 1,731 71 0.134 |
| D. Latin American BRCA variants | 75 13 0.023 | 76 11 0.021 |
| E. Asian BRCA variants | 276 121 0.217 | 266 111 0.209 |
| F. Indian BRCA variants 4 | 89 23 0.041 | 41 2 0.004 |

1Total refers to the numbers in each reference database.
2Proportion = Shared variants / total variants (557 in BRCA1 or 531 in BRCA2) in Chinese population.
3Variants can be overlapped among different databases and populations.
4Indian BRCA variants in Refs. 13 and 22 were combined for the comparison.

CIMBA. Comparing the Chinese BRCA data with the recent CIMBA data enriching non-Caucasian data shows that 17.6% of the 1,088 Chinese BRCA variants had matches [BRCA1: 106/557 (19%) and BRCA2: 86/531 (16.2%). There are 15 Chinese BRCA1 variants and 16 Chinese BRCA2 variants included in both our Chinese data and the CIMBA data. These shared
Chinese variants were removed in order to know the similarity and differences between Chinese and non-Chinese populations included in the CIMBA data. After the removal, the overall matched rate decreased to 14.9% [BRCA1: 91/557 (16.3%) and BRCA2: 71/531 (13.4%)] (Table 1C and Table S7B, Supporting Information).

Latin America and the Caribbean data. Comparing the Chinese BRCA data with those from Latin America and the Caribbean found that 97.8% of total variants [544 (97.7%) BRCA1 variants and 520 (97.9%) BRCA2] were specific only to the Chinese population (Table 1D and Table S7C, Supporting Information).

Non-Chinese Asian populations. Comparing the Chinese BRCA data with those in non-Chinese Asian populations found that 78.6% of total variants [436 (78.2%) of BRCA1 and 420 (79.2%) of BRCA2] were present only in the Chinese population (Table 1E and Table S7D, Supporting Information).

Indian population. Comparison shows that only 23 (4.1%) BRCA1 and two (0.4%) BRCA2 variants were shared between the Chinese and Indian populations (Table 1F and Table S7E, Supporting Information).

Comparison in exon distribution
We compared the variant distribution across BRCA1 and BRCA2 exons between Chinese and non-Chinese populations using the data from BIC database as a testing model. For both BRCA1 and BRCA2, the distribution of variants differed significantly in multiple exons between the two data sets (Fig. S2, Supporting Information). For BRCA1, the proportions of variants in exons 2, 11D, 16, 20 and 24 were higher in the BIC data than in the Chinese data, whereas the proportions in exons 11B and 11C in the Chinese data were significantly higher than in the BIC data; in BRCA2, the proportions in exons 11A, 25 and 27 were higher in the BIC data than in the Chinese data, whereas the proportions in exons 2, 11F, 14, 21 and 22 were higher in the Chinese data than in the BIC data, respectively. The results show the presence of differences of exon distribution between Chinese and non-Chinese BRCA variants.

Comparison in base changes and variant types
We compared single-base changes and other variant types between Chinese and BIC data. The results show the significant differences present in multiple types of base changes between the two data sets, including G > A, C > T and delT in BRCA1, and A > C, delA, delG and delC in BRCA2. In BRCA1, delT had higher frequency in the Chinese population than in the BIC data (6.8% versus 3.5%, p < 0.0034); in BRCA2, delA, delG and delC were more frequent in the Chinese data than in the BIC data (11.5% versus 4.1%, 6.7% versus 1.6% and 5.1% versus 2.4%, respectively, p < 0.000, 0.004, accordingly) (Table 2A). Significant differences were also present in the missense, nonsense, stop gain, splice variants and intronic variant types in both BRCA1 and BRCA2, and in frameshifts in BRCA2 (Table 2B).

Comparison in clinical categories
We compared the clinical categories: pathogenic, likely pathogenic, uncertain significance, likely benign, and benign between Chinese BRCA data and BIC data. The results showed significant differences in multiple categories between the two data sets. For example, 12.7% of BRCA1 variants in Chinese were variants of uncertain significance, which was much higher than the value of 0.57% in the BIC data (p < 0.0000); the proportion of BRCA2 pathogenic variants in the Chinese population was also significantly higher than that in the BIC data (p < 0.0005), and the proportions of unclassified variants in both BRCA1 and BRCA2 were much higher in the BIC data than in the Chinese variants (Table 3).

Comparison in founder mutations
Firstly, we checked if the Chinese variant data contained any BRCA founder mutations known in other populations, including BRCA1 c.66_67delAG (185delAG), c.5263_5264insC (5382insC), and BRCA2 c.5946delET (6174delIT) in Ashkenazi Jews;34 BRCA1 c.-58C > G (C61G), c.4153delA (c.4035delA), and c.5263_5264insC (5382insC) in Poles;35 BRCA1 c.1303 T > G, c.1623dupG, c.4122_4123delTG and c.5324 T > G in Africans;14 BRCA1 ex9-12del in Mexicans;36 BRCA1 390C > A in Koreans and Japanese; and BRCA1 c.470_471delCT, BRCA2 c.7480C > T in Koreans.37 We observed that many of these founder mutations were either absent or present at low prevalence, hence they could not be considered as founder mutations in the Chinese population. Secondly, we searched for potential BRCA founder mutation candidates in the Chinese population by referring to 1) the abundance as calculated by the total number of variant carriers divided by the total number of individuals tested although haplotype data will be required to finally determine the true founder mutations; and 2) additional criteria for removing the variants unlikely to be founder mutations in order to focus on the variants as potential candidates: more than 100 tested individuals (a founder mutation should have a reasonable prevalence in a given population. 100 was set as a minimal cut-off for the population size); at least two detected variant carriers (as a founder mutation, it cannot be only present in a single individual. Therefore, 2 cases were set as the minimal number of mutation carriers. In this way, all variants detected only in single individuals will be eliminated); carrier frequency > 1% (a precondition for founder mutation as pathogenic one is its lower population frequency in population. We set 1% of mutation carrier as the minimal cut-off to eliminate these with high population frequency, which are mostly normal
### Table 2. Comparison of BRCA variation types between Chinese and BIC data

| Change                  | BRCA1                  |         | BRCA2                  |         |
|-------------------------|------------------------|---------|------------------------|---------|
|                         | Chinese Proportion     | BIC     | Proportion             | p       |
|                         |                        |         |                        |         |
| A \( \rightarrow \) G   | 36                     | 0.085   | 155                    | 0.107   | 0.540 |
| A \( \rightarrow \) C   | 10                     | 0.024   | 39                     | 0.027   | 0.867 |
| A \( \rightarrow \) T   | 28                     | 0.066   | 54                     | 0.037   | 0.062 |
| G \( \rightarrow \) A   | 46                     | 0.108   | 204                    | 0.141   | 0.026 |
| G \( \rightarrow \) C   | 12                     | 0.028   | 57                     | 0.039   | 0.479 |
| G \( \rightarrow \) T   | 39                     | 0.092   | 117                    | 0.081   | 0.697 |
| C \( \rightarrow \) A   | 11                     | 0.026   | 57                     | 0.039   | 0.200 |
| C \( \rightarrow \) G   | 19                     | 0.045   | 67                     | 0.046   | 0.436 |
| C \( \rightarrow \) T   | 61                     | 0.144   | 167                    | 0.115   | 0.042 |
| T \( \rightarrow \) A   | 11                     | 0.026   | 43                     | 0.030   | 0.752 |
| T \( \rightarrow \) G   | 13                     | 0.031   | 59                     | 0.041   | 0.262 |
| T \( \rightarrow \) C   | 23                     | 0.054   | 93                     | 0.064   | 1.000 |
| insA                    | 13                     | 0.031   | 57                     | 0.039   | 0.479 |
| insG                    | 4                      | 0.009   | 19                     | 0.013   | 0.815 |
| insC                    | 3                      | 0.007   | 11                     | 0.008   | 0.745 |
| insT                    | 2                      | 0.005   | 24                     | 0.017   | 0.271 |
| delA                    | 25                     | 0.059   | 86                     | 0.059   | 0.910 |
| delG                    | 24                     | 0.056   | 49                     | 0.034   | 0.052 |
| delC                    | 16                     | 0.038   | 41                     | 0.028   | 0.529 |
| delT                    | 29                     | 0.068   | 50                     | 0.035   | 0.003 |

### Table 3. Comparison of Clinical classification between Chinese and BIC BRCA variants

| Class                | BRCA1                  |         | BRCA2                  |         |
|----------------------|------------------------|---------|------------------------|---------|
|                      | Chinese Proportion     | BIC     | Proportion             | p       |
|                      |                        |         |                        |         |
| Pathogenic           | 233                    | 0.418   | 729                    | 0.417   | 0.248 |
| Likely pathogenic    | 25                     | 0.045   | -                      | -       | -     |
| Uncertain signi.     | 71                     | 0.127   | 10                     | 0.006   | 0.000 |
| Likely benign        | 25                     | 0.045   | -                      | -       | -     |
| Benign               | 37                     | 0.066   | 23                     | 0.013   | 0.000 |
| Unclassified         | 166                    | 0.298   | 988                    | 0.565   | 0.000 |
| Total                | 557                    | 1,750   | 531                    | 1,994   |
| Variant impact          | Cases tested$^1$ | Carrier number | Proportion | Exon | HGVS cDNA                  | Reported             | MAF       | 1,000 g EAS | Mutation Type |
|------------------------|------------------|----------------|------------|------|----------------------------|----------------------|-----------|-------------|---------------|
|                        |                  |                |            |      |                           |                      |           |             |               |
| **BRCA1**              |                  |                |            |      |                           |                      |           |             |               |
| Pathogenic             | 266              | 15             | 0.056      | 16   | c.5154G>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Stop gain    |
| Pathogenic             | 118              | 3              | 0.025      | 11D  | c.4258C>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Stop gain    |
| Pathogenic             | 124              | 3              | 0.024      | 11C  | c.3296delC                | B|C|L|B|C|L|V|B|L|O|V|D|E|N|G|M|A | -         | Frameshift deletion |
| Pathogenic             | 125              | 3              | 0.024      | 21   | c.5533_5540delATGGGCA/delTACAGTG | -                  | -        | -           | Frameshift deletion |
| Pathogenic             | 313              | 6              | 0.019      | 11D  | c.3640G>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Stop gain    |
| Pathogenic             | 291              | 5              | 0.017      | 11B  | c.1945G>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Stop gain    |
| Pathogenic             | 118              | 2              | 0.017      | 2    | c.21312A>G                | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           |               |
| Pathogenic             | 257              | 4              | 0.016      | 16   | c.5161C>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Stop gain    |
| Pathogenic             | 130              | 2              | 0.015      | 11A  | c.1066C>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Nonsense     |
| Pathogenic             | 517              | 7              | 0.014      | 18   | c.5332+1G>C               | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Splice site  |
| Pathogenic             | 739              | 10             | 0.014      | 11B  | c.2275C>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Stop gain    |
| Pathogenic             | 1910             | 25             | 0.013      | 6    | c.470_471delCT             | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Frameshift deletion |
| Pathogenic             | 643              | 7              | 0.011      | 15   | c.4986+1G>A               | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Splice site  |
| Pathogenic             | 190              | 2              | 0.011      | 2    | c.68_69delAG               | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Frameshift deletion |
| Pathogenic             | 496              | 5              | 0.010      | 17   | c.5267_5268insC            | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Frameshift insertion |
| Likely pathogenic      | 172              | 3              | 0.017      | 11A  | c.1036C>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Nonsynonymous SNV  |
| Likely pathogenic      | 310              | 3              | 0.010      | 11C  | c.2952delT                | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Frameshift deletion |
| Uncertain significance | 127              | 3              | 0.024      | 11C  | c.3432G>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | Nonsynonymous SNV  |
| Uncertain significance | 193              | 4              | 0.021      | 2    | c.-2A>T                   | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           |               |
| Uncertain significance | 214              | 3              | 0.014      | 11C  | c.2941C>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Nonsynonymous SNV  |
| Uncertain significance | 310              | 4              | 0.013      | 11B  | c.1934C>A                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Nonsynonymous SNV  |
| Uncertain significance | 591              | 6              | 0.010      | 11D  | c.3488C>T                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | 0.000199681  | 0.001 | Nonsynonymous SNV  |
| Unclassified           | 836              | 45             | 0.054      | 11C  | c.2790delT                | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Frameshift deletion |
| Unclassified           | 935              | 29             | 0.031      | 11C  | c.3232C>G                 | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Nonsynonymous SNV  |
| Unclassified           | 135              | 3              | 0.022      | 11C  | c.3180insA                | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Frameshift insertion |
| Unclassified           | 139              | 3              | 0.022      | 11B  | c.2010_2011insTG          | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Frameshift insertion |
| Unclassified           | 141              | 3              | 0.021      | 2    | c.-7G>A                   | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | -           |
| Unclassified           | 320              | 5              | 0.016      | 11C  | c.3420_3421insT           | B|C|L|B|C|L|V|B|E|N|E|G|M|A | -         | -           | Frameshift insertion |

(Continues)
### Table 4. High frequent BRCA variants in Chinese population (Continued)

| Variant impact | Cases tested | Carrier number | Proportion | Exon | HGVS cDNA | Reported | MAF | dbSNP150 | 1,000 g EAS | Mutation Type |
|----------------|--------------|----------------|------------|------|-----------|----------|-----|----------|-------------|---------------|
| Unclassified   | 133          | 2              | 0.015      | 11D  | c.3780_3781delAG/c.3780_3781delAT | -        |     |          |             | Frameshift deletion |
| Unclassified   | 139          | 2              | 0.014      | 3    | c.302-66 T > A | -        |     |          |             |                |
| Unclassified   | 139          | 2              | 0.014      | 7    | c.548-32A > G | BED|LOVD  | -        |          |                |
| Unclassified   | 139          | 2              | 0.014      | 7    | c.548-37A > T | -        |     |          |             |                |
| Unclassified   | 139          | 2              | 0.014      | 11D  | c.4096+122G > T | -        |     |          |             |                |
| Unclassified   | 495          | 7              | 0.014      | 11D  | c.3694_3695insAA | -        |     |          |             | Frameshift insertion |
| Unclassified   | 214          | 3              | 0.014      | 11C  | c.2939 T > A | -        |     |          |             |                |
| Unclassified   | 430          | 6              | 0.014      | 11B  | c.1846_1847insT | -        |     |          |             |                |
| Unclassified   | 430          | 6              | 0.014      | 11C  | c.3182 T > G | -        |     |          |             |                |
| Unclassified   | 430          | 6              | 0.014      | intron 2 | IVS2-55insG | -        |     |          |             |                |
| Unclassified   | 430          | 6              | 0.014      | intron 2 | IVS2-55insTG | -        |     |          |             |                |
| Unclassified   | 238          | 3              | 0.013      | 11C  | c.3072C > G | ClinVar|BED|ENIGMA | -        | Nonsynonymous SNV |
| Unclassified   | 837          | 10             | 0.012      | 11B  | c.2073delA | -        |     |          |             |                |
| Unclassified   | 179          | 2              | 0.011      | 11A  | c.1010delA | BIC      | -        |          |             | Frameshift deletion |
| Unclassified   | 360          | 4              | 0.011      | 17   | c.5277+75_5,277+76insC | -        |     |          |             |                |
| Unclassified   | 274          | 3              | 0.011      | 11B  | c.2252_2253delTG | -        |     |          |             | Frameshift deletion |
| Unclassified   | 403          | 4              | 0.010      | 2    | c.43A > G | -        |     |          |             |                |
| **BRCA2**      |              |                |            |      |            |          |     |          |             |                |
| Pathogenic     | 107          | 4              | 0.037      | 14   | c.7655_7658delTTAA | BIC|ClinVar|BED|ENIGMA | -        | Frameshift deletion |
| Pathogenic     | 180          | 4              | 0.022      | 11A  | c.2636_2637delCT | BIC|BMD|ClinVar|BED|ENIGMA | -        | Frameshift deletion |
| Pathogenic     | 99           | 2              | 0.020      | 11A  | c.2339C > G | BM|D|ClinVar|BED|LOVD|ENIGMA | -        | Frameshift deletion |
| Pathogenic     | 180          | 3              | 0.017      | 11F  | c.6715G > T | BM|D|ClinVar|BED|LOVD|ENIGMA | -        | Frameshift deletion |
| Pathogenic     | 250          | 4              | 0.016      | 9    | c.956_957insA | BIC|ClinVar|BED|BMD|ENIGMA | -        | Frameshift insertion |
| Pathogenic     | 589          | 9              | 0.015      | 22   | c.9098_9099insA | BIC|BMD|ClinVar|BED|LOVD|ENIGMA | -        | Frameshift insertion |
| Pathogenic     | 133          | 2              | 0.015      | 7    | c.755_755delA | BIC|ClinVar|BED|ENIGMA | -        | Frameshift deletion |
| Pathogenic     | 525          | 6              | 0.011      | 23   | c.9253delA | BM|D|ClinVar|BED|ENIGMA | -        | Frameshift deletion |
| Pathogenic     | 471          | 5              | 0.011      | 11F  | c.6449_6450insT | BIC|ClinVar|BED|LOVD|ENIGMA | -        | Frameshift insertion |
| Pathogenic     | 496          | 5              | 0.010      | 11E  | c.5682C > A | BIC|ClinVar|BED|LOVD|ENIGMA | -        | Frameshift deletion |
| Likely pathogenic | 214         | 2              | 0.009      | 20   | c.8800C > T | -        |     |          |             | Stop gain |
| Unclassified   | 119          | 2              | 0.017      | 15   | c.7806-9 T > G | BIC|BMD|ClinVar|BED|LOVD | -        | Intron Variant |
| Unclassified   | 133          | 2              | 0.015      | 11F  | c.6645delC | -        |     |          |             | Frameshift deletion |
| Unclassified   | 496          | 7              | 0.014      | 13   | c.7178_7179delTG | -        |     |          |             | Frameshift deletion |
| Unclassified   | 149          | 2              | 0.013      | 11A  | c.2188_2189insC | -        |     |          |             | Frameshift insertion |
polymorphism); and variants in the categories of pathogenic, likely pathogenic, uncertain significance, or unclassified (Founder mutations must be pathogenic. Restricted the candidates to these classes will narrow down the founder mutation candidates by eliminating the benign and likely benign variants as they do not increase cancer risk). Using these conditions, we tested whether the data could support the Chinese BRCA founder mutations proposed by previous studies including BRCA1 c.981_982delAT (1100delAT), BRCA1 1081delT(1081delG), BRCA1c.5154G > A and BRCA1c.5468-1del8, and BRCA2 c.3109C > T, BRCA2 c.7436_7805del370 and BRCA2c.9097_9098insA. With an exception for BRCA1c.5154G > A variant, our data do not support the above-mentioned variants as the founder mutations. Next, we searched for high-frequency variants meeting the same criteria as above and identified a total of 16 pathogenic, two likely pathogenic, 22 unclassified variants, and five of uncertain significance in BRCA1; and ten pathogenic, one likely pathogenic, and 11 unclassified variants in BRCA2 (Table 4 and Table S8, Supporting Information), respectively. The higher prevalence and clinical pathogenicity of these variants supported them as potential candidates for BRCA1 and BRCA2 founder mutations in the Chinese population. Despite of the higher prevalence, the unclassified variants or those of uncertain significance cannot be regarded as potential founder mutations unless their pathogenicity is determined.

The most significant variants found were the following:

- **BRCA1** c.5154G > A. This variant had the highest prevalence of 5.6% (15 out of 266 detected by five studies). It is a stop-gain pathogenic mutation, present in the BIC, BMD, LOVD, ClinVar, BED databases and was reported as a Chinese founder mutation by a previous study (31).
- **BRCA1** c.4258C > T. This variant had a prevalence of 2.5% (3 out of 118), is pathogenic, and is present in the BIC, BMD, LOVD, ClinVar and BED databases.
- **BRCA1** c.3296delC. This variant had a prevalence of 2.4% (3 out of 124), is pathogenic, and is present in the BIC, BMD, ClinVar, BED and LOVD databases.
- **BRCA1** c.5533_5540delATTGGGCA/delTACCAGTG. This variant had a prevalence of 2.5% (3 out of 125), is pathogenic, and is absent from other BRCA databases.
- **BRCA2** c.7655_7658delTTAA. This variant had a prevalence of 3.7% (4 out of 107, reported by four studies), is a pathogenic frameshift deletion and is present in the BIC, ClinVar and BED databases.
- **BRCA2** c.2636_2637delCT. This variant had a prevalence of 2.2% (4 out of 180), is pathogenic, and is present in the BIC, BMD, ClinVar and LOVD databases.
- **BRCA2** c.2339C > G. This variant had a prevalence of 2% (2 out of 99), is pathogenic, and is present in the BIC, BMD, ClinVar, BED and LOVD databases.

Although these high-frequency BRCA mutations suggest the presence of certain potential founder mutations in Chinese population, it is also obvious from the data that there are unlikely to
be high-prevalence founder mutations in the Chinese population as these in the Ashkenazi Jewish population. However, the Chinese population is composed of highly heterogeneous ethnic groups with different genetic features, and even the dominant Han ethnic group is not homogeneous. Therefore, there remains a possibility for the presence of certain high-prevalence founder mutations in certain specific ethnic groups, and in certain populations located at specific geographical locations.

Comparison within the Chinese population
It is of interest to know whether BRCA ethnic-specificity exists within the Chinese population, given the fact that it has 56 ethnic groups with divergent genetic backgrounds.41 We tested this possibility by using Uygur group as a model. Uygur group is the largest minority group in Xinjiang, northwestern China, with its unique genetic features.41 A series of BRCA studies have been carried out in Uygur group, with the identification of 70 BRCA variants. Of these, 20 BRCA1 variants and 6 BRCA2 variant were present only in the Uygur group. Of the 26 Uygur-specific BRCA1 variants, one was likely pathogenic, one was uncertain significance, one was likely benign, and 23 remain unclassified (Table S9, Supporting Information). The results indicate the presence of Uygur-specific BRCA variants within the Chinese population.

Discussion
By using the rich Chinese BRCA variation data as a representative of non-Caucasian populations, our study provides solid evidence to conclude the presence of ethnic-specific BRCA mutation. This likely reflects the human evolutionary history of genetic diversity and environmental adaptation.11 The variants shared between different ethnic populations were likely originated before their diversification, whereas the ethnic-specific variants were likely generated after their diversification. Since BRCA reference data plays key roles in identifying the mutation carriers, lack of ethnic-specific data in the present references implies that they have inadequate power in locating the mutation carriers with non-Caucasian ethnic background. This is vividly exemplified by the presence of only 16 BRCA variants derived from mainland Chinese among the 3,791 BRCA variants in the BIC database.19 In order to identify the mutation carriers with various ethnic background, ethnic-specific BRCA references need to be developed. Combined usage of both ethnic-specific and existing BRCA reference databases should provide comprehensive identification of BRCA mutation carriers in different ethnic populations, a critical step towards precision medicine. Developing ethnic-specific BRCA references will certainly be a challenge both scientifically and financially, but this task needs to be completed sooner or later for the sake of prevention of BRCA-related cancers in the non-Caucasian populations. The issue of ethnic-specific germline mutation could also exist in other cancer predisposition genes. Experiences from developing ethnic-specific BRCA references should provide a valuable example to address the same issue in these genes.

Acknowledgements
This work was supported by a grant from The University of Macau (SRG2017-00097-FHS), and a startup fund from The Faculty of Health Sciences, University of Macau (SMW). The funders had no roles in study design, data collection and analysis, and decision to publish the study.

References
1. Stratton MR, Rahman N. The emerging landscape of breast cancer susceptibility. Nat Genet. 2008;40:17–22.
2. Miki Y, Swensen J, Shattuck-Edens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science. 1994;266:66–71.
3. Wooster R, Neuhausen SL, Mangion J, et al. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science. 1994; 265:2088–90.
4. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. JAMA. 2017;317:2402–16.
5. Kulkarni A, Carley H. Advances in the recognition and management of hereditary cancer. Br Med Bull. 2016;120:123–38.
6. Thomson E, Varricchio C. Recommendations for follow-up care of individuals with an inherited predisposition to cancer. II. BRCA1 and BRCA2. Cancer genetics studies consortium. JAMA. 1997; 277:997–1003.
7. Pujade-Lauraine E, Ledermann JА, Selle F, et al. Olaparib tablets as maintenance therapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-Orv21): a double-blind, randomised, placebo-controlled, phase 3 trial. Lancet Oncol. 2017;18:1274–84.
8. Knoff ND, Domchek SM, Friebel TM, et al. Risk-reducing salpingo-oophorectomy for the prevention of BRCA1- and BRCA2-associated breast and gynecologic cancer: a multicenter, prospective study. J Clin Oncol. 2008;26:1331–7.
9. King MC, Levy-Lahad E, Lahad A. Population-based screening for BRCA1 and BRCA2: 2014 Lasker award. JAMA. 2014;312:1091–2.
10. Fleming MA, Potter JD, Ramirez C, et al. Understanding missense mutations in the BRCA1 gene: an evolutionary approach. Proc Natl Acad Sci USA. 2003;100:1151–6.
11. Lou DI, McBee RM, Le UQ, et al. Rapid evolution of BRCA1 and BRCA2 in humans and other primates. BMC Evol Biol. 2014;14:155.
12. Dutli J, Golubeva VA, Pacheco-Torres AL, et al. The spectrum of BRCA1 and BRCA2 alleles in Latin America and the Caribbean: a clinical perspective. Breast Cancer Res Treat. 2015;154: 441–53.
13. Kwong A, Shin VY, Ho JC, et al. Comprehensive spectrum of BRCA1 and BRCA2 deleterious mutations in breast cancer in Asian countries. J Med Genet. 2016;53:15–23.
14. Zhang J, Fackenthal JD, Zheng Y, et al. Recurrent BRCA1 and BRCA2 mutations in breast cancer patients of African ancestry. Breast Cancer Res Treat 2012;134:889–94.
15. Rebbeck TR, Friebel TM, Friedman E, et al. Muta- tional Spectra in a worldwide study of 29,700 families with BRCA1 or BRCA2 mutations. Hum Mutat. 2018;39:593–620.
16. Qi J. How China is rewriting the book on human origins. Nature. 2016;535:22–5.
17. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China. 2015. CA Cancer J Clin. 2016;66:115–32.
18. Kim YC, Zhao L, Zhang H, et al. Prevalence and spectrum of BRCA germline variants in mainland Chinese familial breast and ovarian cancer patients. Oncotarget 2016;7:9600–12.
19. den Dunnen JT, Dalgleish R, Maglott DR, et al. HGVS recommendations for the description of sequence variants: 2016 update. Hum Mutat. 2016;37:S64–9.
20. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence vari- ants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–24.
21. Benson DA, Cavanaugh M, Clark K, et al. Genbank. Nucleic Acids Res. 2017;45:D37–42.
23. Wildeman M, van Ophuizen E, den Dunnen JT, et al. Improving sequence variant descriptions in mutation databases and literature using the mutalyzer sequence variation nomenclature checker. *Hum Mutat.* 2008;29:6–13.

24. Liu X, Wu C, Li C, et al. dbNSFP v3.0: a one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Human Mutat.* 2016;37:235–41.

25. Szabo C, Masiello A, Ryan JF, et al. The breast cancer information core: database design, structure, and scope. *Hum Mutat.* 2000;16:123–31.

26. Li Q, Wang K. InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. *Am J Hum Genet.* 2017;100:267–80.

27. Spurdle AB, Healey S, Devereau A, et al. ENIGMA—evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. *Hum Mutat.* 2011;33:2–7.

28. Fokkema IF, Taschner PE, Schaafsma GC, et al. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat.* 2011;32:557–63.

29. Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 2013;42(Database issue):D980–5.

30. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from next-generation sequencing data. *Nucleic Acids Res.* 2010;38:e164.

31. Singh J, Thota N, Singh S, et al. Screening of over 1000 Indian patients with breast and/or ovarian cancer with a multi-gene panel: prevalence of BRCA1/2 and non-BRCA mutations. *Breast Cancer Res Treat.* 2018;170:189–96. https://doi.org/10.1007/s10549-018-4726-x.

32. Yang X, Wu J, Lu J, et al. Identification of a comprehensive spectrum of genetic factors for hereditary breast cancer in a Chinese population by next-generation sequencing. *PLoS One.* 2015;10:e0125571.

33. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536:285–91.

34. Struwing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *N Engl J Med.* 1997;336:1401–8.

35. Górski B, Jakubowska A, Huzarzki T, et al. A high proportion of founder BRCA1 mutations in polish breast cancer families. *Int J Cancer.* 2004;110:683–6.

36. Villarreal-Garza C, Alvarez-Gómez RM, Pérez-Plasencia C, et al. Significant clinical impact of recurrent BRCA1 and BRCA2 mutations in Mexico. *Cancer.* 2015;121:752–8.

37. Kang E, Seong MW, Park SK, et al. The prevalence and spectrum of BRCA1 and BRCA2 mutations in Korean population: recent update of the Korean hereditary breast cancer (KOHBBRA) study. *Breast Cancer Res Treat.* 2015;151:157–68.

38. Li WF, Hu Z, Zhang B, et al. BRCA1 1100delAT is a recurrent mutation in Chinese women with familial breast cancer. Zhonghua Yi Xue Za Zhi. 2007;87:76–80.

39. Kwong A, Ng EK, Wong CL, et al. Identification of BRCA1/2 founder mutations in southern Chinese breast cancer patients using gene sequencing and high resolution DNA melting analysis. *PLoS One.* 2012;7:e3994.

40. Khoo US, Chan KY, Cheung AN, et al. Recurrent BRCA1 and BRCA2 germline mutations in ovarian cancer: a founder mutation of BRCA1 identified in the Chinese population. *Hum Mutat.* 2002;19:307–8.

41. Feng Q, Lu Y, Ni X, et al. Genetic history of Xinjiang’s Uyghurs suggests bronze age multiple-way contacts in Eurasia. *Mol Biol Evol.* 2017;34:2572–82.