Comprehensive breast cancer risk analysis with whole exome sequencing and the prevalence of *BRCA1* and *ABCG2* mutations and oncogenic HPV

SUREEWAN BUMRUNGTHAI[1-3], SUREEWAN DUANGJIT[4], SUPAPORN PASSORN[5], SUTIDA PONGPAKDEESAKUL[5], SIRIWOOT BUTSRI[1,3] and SOMWANG JANYAKHANTIKUL[1,3]

[1]Division of Biopharmaceutical Sciences, Faculty of Pharmaceutical Sciences, Ubon Ratchathani University, Ubon Ratchathani 34190, Thailand; [2]Division of Microbiology and Parasitology, School of Medical Sciences, University of Phayao, Phayao 56000, Thailand; [3]Center for Pharmacogenomics and Clinical Translational Research, Faculty of Pharmaceutical Sciences, Ubon Ratchathani University, Ubon Ratchathani 34190, Thailand; [4]Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmaceutical Sciences, Ubon Ratchathani University, Ubon Ratchathani 34190, Thailand; [5]Division of Biotechnology, School of Agriculture and Natural resources, University of Phayao, Phayao 56000, Thailand

Received April 11, 2024; Accepted July 2, 2024

DOI: 10.3892/br.2024.1832

Abstract. Breast cancer is the most prevalent cancer and also the leading cause of cancer death in women worldwide. A comprehensive understanding of breast cancer risk factors and their incidences is useful information for breast cancer prevention and control planning. The present study aimed to provide information on single nucleotide polymorphisms (SNPs) and copy number variations (CNVs) in breast cancer, the allele frequency of two SNPs in breast cancer-related genes BRCA1 DNA repair associated (*BRCA1;* rs799917) and ATP binding cassette subfamily G member 2 (*ABCG2;* rs2231142), and the prevalence of human papillomavirus (HPV) infections in a normal population living in Phayao Province, Northern Thailand. One breast cancer and 10 healthy samples were investigated by whole exome sequencing (WES) and compared for genetic variation. The WES data contained SNPs in genes previously implicated in breast cancer and provided data on CNVs. The allele frequencies for SNPs rs799917 and rs2231142 were also examined. The SNP genotype frequencies were 35.88% CC, 46.54% CT, and 17.58% TT for rs799917 and 33.20% CC, 46.88% CA, and 19.92% AA for rs2231142. A total of 825 human whole blood samples were examined for HPV infection by PCR, and the pooled DNA was tested for HPV infection using metagenomic sequencing. No HPV infections were detected among all 825 samples or the pooled blood samples. The incidence of breast cancer among the tested samples was estimated based on acceptable breast cancer risk factors and demographic data and was 1.47%. The present study provided data on SNPs and CNVs in breast cancer-related genes. The associations between SNPs rs2231142 and rs799917 and breast cancer should be further investigated in a case-control study since heterozygous and homozygous variants are more common. Based on the detection of HPV infection in the blood samples, HPV may not be associated with breast cancer, at least in the Northern Thai population.

Correspondence to: Dr Somwang Janyakhantikul, Division of Biopharmaceutical Sciences, Faculty of Pharmaceutical Sciences, Ubon Ratchathani University, 85 Sathonlamark Road, Warin Chamrap, Ubon Ratchathani 34190, Thailand
E-mail: somwang.j@ubu.ac.th

Key words: breast cancer, whole exome sequencing, copy number variation, *BRCA1*, *ABCG2*, human papillomavirus, Thailand

Introduction

Breast cancer is the most common cancer in females. The global cancer statistics indicate that in 2020, breast cancer was the leading newly diagnosed cancer worldwide, accounting for 2,261,419 new cases (11.7%), and the leading cause of cancer-related mortality worldwide, accounting for 684,996 deaths (6.9%) (1). Xu *et al* (2) showed that the global incidence of breast cancer increased by 123% between 1990 and 2017 and it is expected to continue increasing yearly. Furthermore, the authors estimated that there will be 4,781,849 cases of breast cancer and 1,503,694 breast cancer-related mortality worldwide in 2050. In Thailand, breast cancer is a common malignancy, ranking third after liver and lung cancer in new cases and mortality annually (3). Therefore, a comprehensive understanding of breast cancer risk factors and their incidences is valuable for breast cancer prevention, including screening, diagnosis, prognosis, personal lifestyle modification, and effective therapy strategies, such as precision medicine, although the latter will require collecting data on breast cancer genetics, environmental factors, and lifestyle behaviors.

Numerous non-modifiable (e.g., female sex, older age, family history and genetic mutations) and modifiable (e.g., obesity, alcohol consumption and smoking) risk factors have been reported for breast cancer (4-6). Single nucleotide polymorphisms (SNPs) in the breast cancer-related genes BRCA1 DNA

repair associated gene (*BRCA1*; rs799917) and ATP-binding cassette subfamily G member 2 (*ABCG2*; rs2231142) are some examples of genetic mutations that have been reported. *BRCA1* is a tumor suppressor. The mutations in *BRCA1* may cause BRCA1 dysfunction and affect cancer risk, especially breast and ovarian cancer (7). However, recent meta-analysis studies have revealed no significant correlation between SNP rs799917 and breast cancer (8,9) or overall cancer (10,11) risk. Moreover, SNP rs799917 in *BRCA1* could be a protective factor for non-breast cancer in Asian populations (10,11). ABCG2, or breast cancer resistance protein (BCRP), is an efflux transporter that functions as a xenobiotic transporter. Homozygosity for the A allele of SNP rs2231142 has recently been reported to be associated with an increased risk of breast cancer (12,13) and axillary lymph node status (14).

One recent additional and intriguing potential risk factor for breast cancer is human papillomavirus (HPV) infection, which has been reported to increase the risk of breast carcinoma in a large-scale systemic review and meta-analysis of case-control studies (15). The mechanism by which HPV causes breast carcinogenesis is also not well studied, but some reports reveal that HPV DNA in serum-derived extracellular vesicles was found in breast cancer patients, where E6 and E7 oncoproteins act as stimulators of host cell proliferation and inhibit apoptosis (16,17). However, the mode of HPV transmission underlying breast cancer remains unclear (18). There is a report that HPV was identified in 1.7% of the blood samples or peripheral blood mononuclear cells (PBMCs) of healthy individuals (19) but the association between HPV transmission and breast cancer in blood samples in Thailand has not yet been verified.

The present study aimed to explore genetic mutations [i.e., SNPs and copy number variations (CNVs)] in breast cancer and, in particular, the allele frequencies for two SNPs (*BRCA1*; rs799917) and (*ABCG2*; rs2231142), as well as the prevalence of HPV infections in individuals living in Phayao Province, Thailand.

Materials and methods

Participants. The present study collected 825 human whole blood samples from normal individuals aged 3-90 years (mean age 51.25±19.20 years old, with 236 males and 589 females, sex ratio: 0.40) living in Phayao Province, Thailand, along with demographic data on their sex, body mass index (BMI), family history of cancer, exercise, alcohol consumption, smoking, secondhand smoke and cleanliness of drinking water. The participants were randomly selected between March 2020 and July 2022 by researchers at the Division of Microbiology and Parasitology, School of Medical Sciences, The University of Phayao, Thailand. The present study was approved by the Committee on Human Research Ethics in Health Sciences and Science and Technology at the University of Phayao (1.3/023/63 and 1.3/013/65) and the Ubonratchathani University (UBU-REC-68/2567).

Blood collection and DNA extraction. Human whole blood samples were randomly collected from 825 individuals aged 3-90 years living in Phayao Province, Thailand. DNA was extracted from these samples using the Genomic DNA Isolation Kit (cat. no. PDC11-0100; Bio-Helix Co., Ltd.) as in previous studies (20,21).

Whole exome sequencing (WES). WES was used to examine point mutations in 11 human samples, consisting of one female patient with breast cancer and 10 healthy individuals of both sexes who were born and live in Phayao Province, Thailand. The 10 healthy samples were aged 11-20, 21-30, 31-40, 41-50, and 51-60 years. The details of all samples, including sex, age, and family history of cancer, are provided in Table I. The DNA was extracted as previously described (20,21) and eluted in TE buffer. DNA quality control, library preparation, library quality control, cluster generation, and sequencing were performed as previous studies (20,21). Briefly, a qualified DNA sample was fragmented using an ultrasonicator for 20-40 kHz, 360 sec, at 4˚C. Then, the interrupted DNA fragment was constructed into a high-throughput sequencing library through the steps of terminal repair, adding a base A tail, adapter ligation, purification and pre-amplification, quantitative, exon capture and PCR enrichment. After the completion of library preparation, the size and concentration of each sample were determined and the Qubit fluorometer (Thermo Fisher Scientific, Inc.) was used for accurate measurement of DNA concentration. The accuracy and concentration of the sequencing library were assessed. Finally, to ensure the accuracy of the library concentration and data output, the effective concentration of the library mixture was measured. The optical signal under the four fluorescent channels scanned by the built-in software Illumina Real Time Analysis (RTA) software (version 1.17.28; Illumina, Inc.) and was converted to base calling files in real time. After the base calling, Illumina's official software, bcl2fastq (v2.17; Illumina, Inc.), was used to demultiplex the data according to the sample index sequence and convert it into FASTQ format. The primary analysis was conducted using the built-in High-Content Screening (HCS) sequencer software to determine whether the reads would pass filter (PF; the first 25 cycles have ≤2 bases, whose chastity value is <0.6) based on the purity of the first 25 cycles of the read signal or not. The PF clusters stored in the FASTQ format after a conversation are called PF data, or raw data. For paired-end data, sequence data consists of two FASTQ files that hold each end of the sequence read.

SNPs genotyping

BRCA1 (rs799917) detection by high-resolution melting (HRM) analysis. *BRCA1* SNP rs799917 was genotyped in 825 samples by HRM as previously described (22) using the forward and reverse primers listed in Table SI, which produced a PCR product of 47 bp. The quantitative (q)PCR reaction used the 5X FiREPOL Eva Green HRM Mix Plus (Solis BioDyne OÜ). The qPCR conditions were as follows: Initial denaturation at 95˚C for 12 min, then 40 cycles of denaturation at 95˚C for 15 sec, annealing/elongation at 60˚C for 20 sec and melting at 65-95˚C for 5 sec/step. A DNA sample known to have the *BRCA1* SNP rs799917 (C>T) was used as a positive control. DNase- and RNase-free water was used as a negative control.

Following HRM analysis, 60 samples were randomly selected to repeat for verification of SNP genotyping results by Sanger sequencing. Sanger sequencing was conducted

Table I. Characteristics of a female patient with breast cancer and 10 healthy individuals.

ID	Sex	Age, years	Family history of cancer
1	Female	16	No disease
2	Male	13	No disease
3	Female	29	Mother with breast cancer
4	Male	22	No disease
5	Female	32	Mother with breast cancer
6	Male	38	No disease
7	Female	43	No disease
8	Male	49	Father with liver cancer
9	Female	59	Sister with colon cancer
10	Male	56	Mother with liver cancer and sister with breast cancer
11	Female	63	Patient with breast cancer

to confirm the *BRCA1* SNP rs799917 (C>T) with primers for *BRCA1* amplifying a 719 bp fragment (Table SI). The sequences were computationally compared to the reference sequence in the GenBank database using the BioEdit (version 7.2) biological sequence alignment editor, which was developed by Tom Hall and downloaded from https://bioedit.software.informer.com/7.2/.

ABCG2 (rs2231142) detection by PCR-restriction fragment length polymorphism (PCR-RFLP) analysis. *ABCG2* SNP rs2231142 was genotyped in 256 samples, randomly selected from 825 samples, using the PCR-RFLP method described by Wu *et al* (12), with some modifications. Briefly, the 25 μl PCR mixture comprised 12.5 μl of One PCR (GeneDireX, Inc.), 1 μl (10 μM) of both the forward and reverse primers (Table SI), and 10-100 ng of genomic DNA. The PCR amplification was conducted with the following parameters: Initial denaturation step of 5 min at 94˚C; followed by 35 cycles of 30 sec at 94˚C, 30 sec at 58˚C, and 1 min at 72˚C and then a final elongation of 5 min at 72˚C. The PCR products (302 bp) were digested with *TaaI* (Thermo Fisher Scientific, Inc.) at 65˚C for 2 h. The A allele was uncut, and the C allele was cut into 252 and 50 bp fragments.

DNA sequencing of the ABCG2 promoter. Males and females in 18 samples were randomly selected for DNA sequencing in the *ABCG2* promoter region (chr4:89,079,995-89,080,518; hg19), according to Eclov *et al* (23) with some modifications. The PCR was conducted using the forward primer and reverse primers listed in Table SI and the following conditions: 95˚C for 2 min; followed by 35 cycles of 30 sec at 95˚C, 30 sec at 60˚C and 1 min at 68˚C and then a final extension of 10 min at 72˚C. The 524 bp PCR product was confirmed by gel electrophoresis and sequenced.

HPV DNA detection by PCR. HPV DNA was detected in 825 samples by PCR using the primers listed in Table SI, which produced a PCR product of 154 bp (24). Caski cell line (kindly provided by HPV & EBV and Carcinogenesis Research Group, Khon Kaen University, Thailand) was used as a positive control. The PCR reaction used 5X FiREPOL Eva Green HRM Mix Plus (Solis BioDyne OÜ). The hemoglobin subunit β gene was used as the housekeeping control gene, producing a PCR product of 268 bp, as in a previous study (20).

Metagenomics. The pooled DNA extracted from the 825 samples was analyzed by shotgun metagenomic sequencing (next-generation sequencing) as in a previous study (20) to detect oncogenic HPV DNA. According to WES in the material and methods.

CNV analysis. CNVs are a type of structural variation that represents repeated sections of the genome, and the number of repeats varies among individuals. They include gains and losses. Studies on CNVs have been seminal in rare genetic diseases. The present study used Control-FREEC to detect CNVs based on the change in read depth across the genome, annotated them using Annovar and counted their number (some examples are shown in Fig. 1).

Statistical analysis. The data were analyzed using the SPSS software (version 16; IBM Corp.). Pearson's χ^2 test was used to compare categorical variables between groups. An independent Student's t-test was used to compare means ± standard deviation (SD) between pairs of groups. $P<0.05$ was considered to indicate a statistically significant difference.

Results

Study population characteristics. The demographic characteristics of participants (i.e., sex, age, family history of cancer, BMI, exercise, cleanliness of water, alcohol consumption, smoking and secondhand smoke) are shown in Table II. Their age ranged from 3-90 years, with a mean of 51.25 years (SD 19.20 years). Notably, nearly 25% of the participants had a family history of cancer; 15.03 and 20.48% were at risk of cancer from smoking and secondhand smoke, respectively.

Genotype frequencies of SNPs rs799917 and rs2231142 and the DNA sequencing of the ABCG2 promoter. The genotyped SNPs (rs799917 and rs2231142) were both in the Hardy-Weinberg equilibrium (P=1.119 and P=0.534,

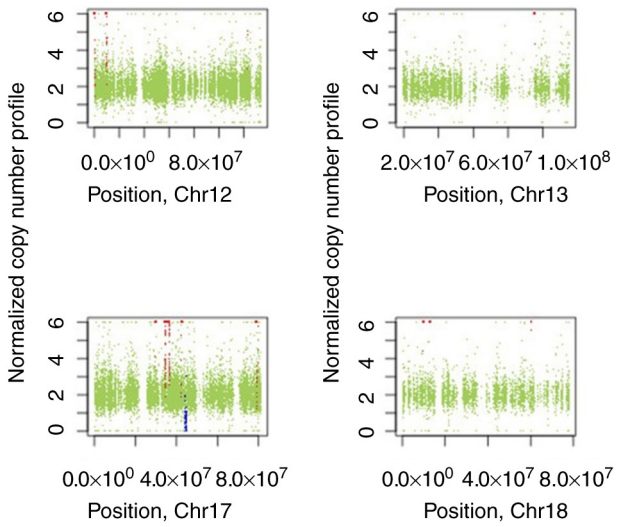

Figure 1. Copy number variants identified by whole exome sequencing in one breast cancer sample. The χ-coordinate shows the chromosome's position (in X.X x 10^x). The y-coordinate shows the copy number, with the normal copy number shown in green, duplications shown in red, and deletions shown in blue.

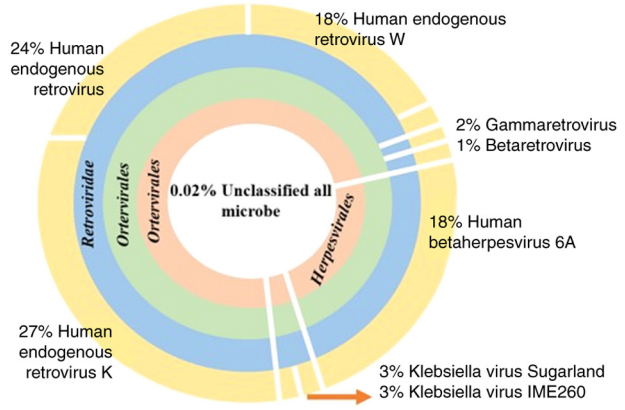

Figure 2. Metagenomic sequencing data. The white circle is an unclassified microbe. The orange circlet is *Ortervirales* and *Herpesvirales*, separated by a white line. The green circlet is *Ortervirales*. The blue circlet is *Retroviridae*. The yellow circlet is a group of viruses that are separated by a white line.

respectively) and had a minor allele frequency (MAF) of 0.408 and 0.434, respectively. The allele frequencies of these SNPs in the present study and other populations are compared in Table SII. The genotype frequencies were 35.88% CC, 46.54% CT, and 17.58% TT for rs799917, and 33.20% CC, 46.88% CA, and 19.92% AA for rs2231142. The genotype frequencies for SNP rs799917 were not associated with sex or family history of cancer (Table II). However, notably, the genotype frequencies for SNP rs2231142 differed between males and females, with the most common genotype being homozygous wide-type (CC) in males and heterozygous (CA) in females (P<0.001; Table II).

The present study investigated rare SNPs in the *ABCG2* promoter using DNA sequencing (~524 bp) in 18 samples, finding that one (5.56%) had a heterozygous (CT) at the SNP rs76656413 (C>T).

HPV detection by PCR and metagenomic sequencing. Blood was collected from 825 Thai donors, and HPV DNA was detected by PCR. However, HPV DNA was not detected in any examined sample (0/825; 0%). The metagenomic sequencing of pool blood DNA samples also detected no HPV DNA in the samples (Fig. 2).

The predicted number of breast cancer cases in the tested samples. The number of breast cancer cases among 68 tested samples (of 256 samples) from females aged ≥50 years who were tested for SNP rs2231142 was estimated based on all collected data on breast cancer risk factors, including demographic factors (i.e., female and aged ≥50 years), genetic factors (i.e., SNP rs2231142 and family history of cancer), environmental factors (i.e., HPV infection and secondhand smoke) and lifestyle behaviors (i.e., alcohol consumption, smoking, obesity and overweight, and physical activity). One of the 68 samples (1.47%) was predicted to have a high risk of breast cancer.

WES analysis. WES identified SNPs in the following genes (Fig. 3 and Table III): *ADAM, ADCY, ANKRD, APOBEC, ARHGEF, ASTN2, NUAK, ATAD, ATAD, BCL, BCL2L, BABAM, CNTROB, CASP, CCDC, CDCA, CDKAL, CMSS, MYC, CNTNAP, COX, CREB, CUX, DCLRE, DNAH, DNAJC, SREBF, PERM, ESRP, MRPS, FGFR, FLJ, GRHL, KANSL, L3MBTL, LSP1, MYEOV, OTUD, PDE, MAP3K, MCM, PEX, PIK3C, PLA2G, JMJD, PLA2G4F, PPRC, RAD51, RAD54, RGPD, RBL, RNF, SETBP, SLC, TNRC, TP53BP1, WDR, ZBTB, ZC3H*, and *ZNF*. Notably, SNPs were identified in the *NBPF9, NBPF1, SLC9B1, PABPC3, SLC9B1, PABPC3, PPIAL4G, PPIAL4H, H2BFS* and *ZC3H11B* genes in all 11 samples, including a patient with breast cancer.

Discussion

Breast cancer risk factors and their incidences are helpful for breast cancer prevention. In the present study, WES was used to detect SNPs and CNVs in one patient with breast cancer and 10 healthy individuals, and then the SNPs were compared. It also investigated the genotype frequencies of two SNPs (rs799917 and rs2231142) in breast cancer-related genes and HPV infection. Finally, all data on breast cancer risk factors, including genetic factors, environmental factors, and lifestyle behaviors, were gathered to predict the number of breast cancer cases among the tested samples.

CNVs are associated with breast cancer risk and diagnosed for breast cancer subtypes (36,37). Dennis *et al* (38) found that 0.5% of 86,788 cases had a deletion in one of the known breast cancer susceptibility genes. In the present study, CNVs were detected in a Thai patient with breast cancer (Fig. 1). While there are numerous duplication and deletion sites on chromosome 17 where several breast cancer-related genes are located (e.g., *HER2, TOP2A, TAU, p53, BRCA1,* and *HIC-1*) (39), the present study did not detect CNVs in these genes in this patient. While various germline and somatic CNVs associated with breast cancer risk have been reported, such as in *BRCA1, CHEK2, ATM, BRCA2, ERBB2, MYC, NBN, CCND1*, and *MCL1* (38,40), The present study also did not detect them in its Thai patient with breast cancer. It was hypothesized that the main reason may be the

Table II. Demographic characteristics of participants and genotype frequencies for SNPs rs799917 and rs2231142.

		Genotype frequencies of rs799917				Genotype frequencies of rs2231142			
Demographical factors	Total (n=825) (%)	CC (296)	CT (384)	TT (145)	P-value	CC (85)	CA (120) (%)	AA (51)	P-value
Sex					0.596				<0.001
Male	589 (71.39%)	205 (34.80%)	279 (47.37%)	105 (17.83%)		51 (43.59%)	30 (25.64%)	36 (30.77%)	
Female	236 (28.61%)	91 (38.56%)	105 (44.49%)	40 (16.95%)		34 (24.46%)	90 (64.75%)	15 (10.79%)	
Age (years)									
Mean	51.25								
SD	19.20								
Family history of cancer					0.488				0.748
Yes	186 (22.55%)	66 (35.48%)	82 (44.09%)	38 (20.43%)		20 (33.33%)	30 (50.00%)	10 (16.67%)	
No	639 (77.45%)	230 (35.99%)	302 (47.26%)	107 (16.75%)		65 (33.16%)	90 (45.92%)	41 (20.92%)	
BMI (kg/m^2)									
Low (<18.50)	102 (12.37%)								
Normal (18.50-22.90)	489 (59.27%)								
High (≥23.00)	234 (28.36%)								
Exercise									
Yes	418 (50.67%)								
No	407 (49.33%)								
Alcohol consumption									
Yes	326 (39.52%)								
No	499 (60.48%)								
Smoking									
Yes	124 (15.03%)								
No	701 (84.97%)								
Secondhand smoke									
Yes	169 (20.48%)								
No	656 (79.52%)								

SD, standard deviation; BMI, body mass index.

Figure 3. Comparison of single nucleotide polymorphisms between one patient with breast cancer and 10 healthy individuals of different ages. M, man; W, woman; BC, breast cancer.

difference in ethnic ancestry. However, the present study detected a CNV in *RAD51D* associated with breast cancer risk that had been published in the Thai population (41,42). *RAD51D* is involved in homologous recombinant DNA repair, and it carries mutations that are known to be pathogenic. While it is not the most commonly mutated gene, it is mutated more often in Thai

Table III. Breast cancer-related genes containing SNPs in 10 healthy individuals and one patient with breast cancer that have been reported in published articles (25-35).

Gene from whole exome sequencing	Genes reported in published articles	SNP rsID	Gene from whole exome sequencing	Genes reported in published articles	SNP rsID
ADAMTSL4	*ADAM29*	rs587742902	*L3MBTL1*	*L3MBTL3*	rs772502710
ADAM15		rs113878254	*L3MBTL1*		rs372050022
ADAM15			*LSP1*	*LSP1*	rs552802699
ADAMTS12		rs759606612	*LSP1*		rs148262402
ADAMTS6		rs368191265	*MAP3K11*	*MAP3K1*	rs11227236
ADAMDEC1		rs200134300	*MAP3K9*		rs34322726
ADAM32			*MAP3K10*		
ADAMTS13			*MCM10*	*MCM8*	rs187685058
ADAMTS14		rs147256643	*MCM8*		rs760412395
ADAMTS14		rs376614311	*MCM5*	*MDM4*	rs751091748
ADAM20		rs567945895	*MDM2*		
ADAMTSL3			*MYEOV*	*MYEOV*, *CCNDL*	rs148448631
ADAMTS17		rs371570653	*OTUD4*	*OTUD7B*	rs369626183
ADAMTS1		rs751984218	*PDE4DIP*	*PDE4D*	rs151058495
ADCY10	*ADCY3*	rs200816878	*PDE4DIP*		rs782064349
ADCY4		rs532277226	*PDE4DIP*		rs140993521
AKAP9			*PDE4DIP*		rs201403178
AKAP9		rs746860114	*PDE4DIP*		rs142679243
AKAP9		rs77447750	*PDE4DIP*		
AKAP13			*PDE4DIP*		rs573724
AKAP4			*PDE2A*		rs561445982
ANKRD34A	*ANKRD16*	rs781900571	*PDE8A*		rs189073229
ANKRD65		rs574552814	*PDE4C*		rs149614671
ANKRD36C		rs773442285	*PDE4C*		rs202177222
ANKRD36C		rs768682466	*PDE9A*		
ANKRD39		rs528356666	*PEX6*	*PEX14*	rs200115671
ANKRD36		rs751998840	*PIK3CG*	*PIK3R3*	
ANKRD36		rs745923584	*PLA2G2F*	*PLA2G6*	rs368567704
ANKRD36			*PLA2G4A*		rs188911054
ANKRD36		rs375602706	*JMJD7-PLA2G4B*, *PLA2G4B*		rs200327143
ANKRD36		rs768585370	*PLA2G4F*		rs530370813
ANKRD30BL			*PPRC1*	*PRC1*	rs565951388
ANKRD17		rs534030909	*PPRC1*		rs145446235
ANKRD31			*RAD51AP2*	*RAD51C*	rs575308098
ANKRD31		rs776341649	*RAD54L*	*RAD51D*	
ANKRD31		rs550567797	*RGPD2*	*RANBP9*	
ANKRD18B			*RGPD3*		
ANKRD26		rs200775533	*RANBP2*		
ANKRD52		rs758974930	*RANBP2*		
APOBEC1	*APOBEC3B*	rs61753204	*RANBP2*		rs774306184
APOBEC3G		rs183180481	*RBL2*	*RBL2*, *TOX3*, *TNRC9*	rs555878756
ARHGEF10L	*ARHGEF6*	rs544006964	*RNF180*	*RNF115*	rs774684839
ARHGEF4		rs373534611	*RNF8*		rs542382214
ARHGEF12			*RNF26*		rs145592178
FARP1		rs192616500	*RNF10*		
ASTN2	*ASTN2*		*RNF31*		
NUAK2	*AT*	rs539524885	*RNF31*		rs201497992
ATAD3B	*ATAD5*		*RNF111*		rs747518713

Table III. Continued.

Gene from whole exome sequencing	Genes reported in published articles	SNP rsID	Gene from whole exome sequencing	Genes reported in published articles	SNP rsID
ATAD5		rs769083129	*RNF40*		rs769686628
BCL11A	*BCL2L11*		*RNFT1*		rs138794420
BCL2L11			*RNF213*		rs371441113
BABAM2	*BRCA1*	rs368517485	*RNF225*		rs187505845
CNTROB	*BRCA2*	rs772562149	*SETBP1*	*SETBP1*	rs748289164
CASP7	*CASP8*	rs376949404	*SETBP1*		rs529611461
CCDC24	*CCDC88C*	rs187616405	*SLC45A1*	*SLC4A7*	rs373884958
CCDC121		rs767258483	*SLC4A4*		rs377031010
CCDC141		rs77071759	*SLC44A4*		rs375793445
CCDC150		rs201013091	*SLC45A4*		
CCDC136			*SLC4A1*		rs757478694
CCDC171			*SLC44A2*		rs145954566
CCDC65		rs142550817	*STXBP5L*	*STXBP4*	rs186768873
CCDC63		rs116032516	*STXBP2*		rs142105943
CCDC60		rs141367042	*TGIF2LX*	*TGFBR2*	
CCDC169, *CCDC169-SOHLH2*			*TNRC18*	*TNRC9*	
CCDC168			*TNRC18*		rs748968031
CCDC168		rs200872789	*TP53BP1*	*TP53*	rs548813580
CCDC168		rs540907577	*WDR49*	*WDR43*	
CCDC168		rs201134938	*WDR17*		rs142589281
CCDC168		rs116890855	*WDR41*		rs77995935
CCDC88C	*CCDC88C*		*WDR60*		rs748375887
CCDC33	*CCDC88C*	rs373900085	*WDR97*		rs543467041
CCDC33		rs538344642	*WDR38*		
CCDC189		rs182267845	*WDR5*		rs778304067
CCDC40		rs775239960	*WDR37*		rs144312898
CCDC40		rs760407697	*WDR89*		
CCDC105		rs553990912	*WDR89*		
CCDC114		rs377507314	*WDR89*		rs201238690
CCDC116			*WDR89*		
CDCA7	*CDCA7*	rs138353896	*WDR89*		
CDKAL1	*CDKAL1*	rs553804984	*WDR89*		
CMSS1	*CMSS1*	rs145645351	*WDR89*		
MYCBP2	*cMyC*		*WDR89*		
MYCBP2			*WDR89*		rs74383752
CNTNAP5	*CNTNAP1*	rs541671672	*WDR89*		rs200681506
CNTNAP2			*WDR24*		rs775417149
CNTNAP3B			*WDR24*		
CNTNAP3B		rs1755755	*WDR81*		rs780532350
COX7A2	*COX11*	rs138092231	*WDR62*		rs564143230
COX15		rs201703572	*WDR13*		
COX7A1		rs755756129	*WDR44*		rs200615882
CREB3L1	*CREB5*	rs376081099	*ZBTB48*	*ZBTB38*	rs554036434
CUX1	*CUX1*	rs139293638	*ZBTB40*		rs148301324
DCLRE1C	*DCLRE1B*	rs376186052	*ZBTB39*		rs182966445
DNAH14	*DNAH11*		*ZC3H11B*	*ZC3H11A*	rs2653989
DNAH14			*ZC3H14*		rs80289104
DNAH6		rs375106276	*ZC3H11B*		rs571704621

Table III. Continued.

Gene from whole exome sequencing	Genes reported in published articles	SNP rsID	Gene from whole exome sequencing	Genes reported in published articles	SNP rsID
DNAH6			*ZC3H12B*		
DNAH7		rs764776065	*ZNF687*	*ZNF365*	rs151299620
DNAH12			*ZNF638*		
DNAH8		rs575069902	*ZNF638*		
DNAH8			*ZNF142*		rs756225038
DNAH11		rs781560218	*ZNF860*		rs572072037
DNAH11		rs369849556	*ZNF501*		
DNAH11		rs183489539	*ZNF80*		rs572397764
DNAH11			*ZNF595*		rs146070291
DNAH10		rs556641156	*ZNF732*		rs150738695
DNAH3		rs200676672	*ZNF141*		rs201791423
DNAH3			*ZNF141*		rs79227679
DNAH3		rs777262918	*ZNF141*		rs79869819
DNAH3		rs539288270	*ZNF141*		
DNAJC10	*DNAJC1*	rs372447298	*ZNF518B*		rs185283370
DNAJC15		rs115128267	*ZNF330*		rs548266148
SREBF1	*EBF1*	rs775175384	*ZNF622*		rs200470817
PERM1	*ESR*	rs528106044	*ZNF354A*		rs575302139
ESRP1		rs528521502	*ZNF165*		
ESRRA		rs373399001	*ZNF316*		
ESRP2		rs777034822	*ZNF479*		rs201001924
MRPS25	*FGF10/MRPS30*	rs377446402	*ZNF679*		rs375602152
MRPS18A		rs750387982	*ZNF680*		rs188006471
MRPS23		rs369458033	*ZNF107*		rs184622647
MRPS12		rs147007310	*ZNF107*		
FGFR4	*FGFR2*		*ZNF107*		
FLJ44635	*FLJ43663*	rs199561699	*ZNF107*		
GRHL3	*GRHL1*	rs754006408	*ZNF138*		rs551808152
GRHL2		rs746616786	*ZNF804B*		
HSPA6	*HSPA*	rs200790521	*ZNF425*		rs781577739
HSPA9			*ZNF862*		rs371880003
HSPA8			*ZNF705G*		rs376157799
KANSL1L	*KANSL1*	rs374376792	*ZNF7*		rs75052405
KANSL1			*ZNF16*		rs139521477

SNP, single nucleotide polymorphism.

patients with breast cancer (41,42), supporting the inclusion of *RAD51D* in breast cancer genetic testing (43).

In the present study, SNPs were also characterized in one patient with breast cancer and 10 healthy individuals by WES. Notably, no SNPs were detected in the patient with breast cancer in the top genes previously reported in the Thai population (41,42). However, the results revealed that some SNPs were in previously reported breast cancer genes (Table III). A comparison of SNPs between the breast cancer patient and 10 healthy individuals (Fig. 3) identified several that were common in the tested samples.

BRCA1 is a tumor suppressor located on chromosome 17q21. Its protein has multiple functions in cell homeostasis during the cell cycle, including DNA replication and apoptosis. Therefore, mutations in *BRCA1* may cause BRCA1 dysfunction and affect cancer risk through its intracellular functions, as has been reported for various types of cancer, especially breast and ovarian (7). However, the association between SNP rs799917, a missense mutation (p.P871L), and breast cancer susceptibility remains unclear. Nicoloso *et al* (44) found inconsistent results in which the carrier (T) allele of this SNP was associated with a weaker microRNA 638-dependent *BRCA1* reduction. However, while it was shown to be associated with breast cancer risk in their case-control study and other studies (45), meta-analysis studies have revealed no significant correlation between SNP rs799917 and breast cancer (8,9) or overall cancer (10,11) risk. Moreover, SNP rs799917 in *BRCA1* could be a protective factor for non-breast cancer in

Asian populations (10,11). One study reported that this SNP was a neutral variant found in 17.6% of 190 healthy Thai individuals (46). Unexpectedly, in the present study, the frequency of this SNP was 17.6%, the same as in Ahmad *et al* (46). Due to its high prevalence in the Thai population, this protective factor for breast cancer should be investigated further in a large case-control study to confirm this hypothesis.

ABCG2, formerly called breast cancer resistance protein (BCRP), is an efflux transporter that functions as a xenobiotic transporter. While *ABCG2* contains numerous SNPs, one of the most important is rs2231142 (C421A), which affects ABCG2 protein expression and function with clinical significance in diseases (e.g., gout) (47,48) and drug pharmacokinetics (e.g., statins) (49). Homozygosity for the A allele of SNP rs2231142 has recently been reported to be associated with an increased risk of breast cancer (12,13) and axillary lymph node status (14). The frequency of this SNP differs among populations; it is more common in Asians than Caucasians and is rarely found in Africans (50,51). In the present Thai study, its MAF was 0.434 and its most common genotype was the heterozygote (CA). Therefore, it could be pathogenic and a risk factor for breast cancer in the samples of the present study. However, the relationship between SNP rs2231142 and breast cancer risk requires further study to be confirmed. Moreover, it was also found that one of 18 samples (5.56%) had a polymorphism in the *ABCG2* promoter (SNP rs76656413), which has been reported to markedly reduce promoter activity by 70% (23). Notably, this volunteer also had a homozygous variant (AA) genotype for SNP rs2231142. Therefore, the allele frequency for SNP rs76656413 should be investigated further for its prevalence and clinical relevance in the Thai population, especially in those who carry both SNPs.

The association between HPV infection and risk factors for breast cancer is controversial since the reported frequency of HPV infection in patients with breast cancer varies widely from 0-86.2% and there is a lack of evidence to support HPV transmission to the breast (18). However, a recent meta-analysis of 3,607 breast cancer cases and 1,728 controls confirmed that being HPV-positive increases breast cancer risk (15). In Thailand, one study reported a low HPV-positive frequency, detecting HPV DNA in 25/700 (3.57%) of Thai women with breast cancer and benign breast tumors (52). Our previous study also detected HPV DNA at a low HPV-positive frequency, detecting HPV DNA in 3.7% of oral rinses from 594 healthy participants in Northern Thailand (53). The present study is the first to detect HPV in human blood samples from Thai individuals, including the pool meta-genomic results, detecting no HPV DNA. Therefore, it appeared that HPV infection in a blood sample HPV might not be associated with breast cancer, at least in the Northern Thai population.

Notably, an analysis considering all data on breast cancer risk factors, including genetic factors, environmental factors and lifestyle behaviors, predicted a relatively high number of breast cancer cases among the 68 tested samples of 1.47%. Therefore, all causes should be further investigated and modifiable risk factors for breast cancer should be used to prevent and control breast cancer risk to reduce the number of cases in the future.

However, the present study also had limitations, such as the following: First, the present study is not a case-control study; all participants in the present study were healthy individuals, excluding one breast cancer patient. Therefore, a large case-control study still needs to be further investigated. Second, all the volunteers in the present study were individuals who live in Phayao Province, which is located in the northern part of Thailand. Thus, in some respects, the results of the present study might not be properly representative of the Thai population. Third, the present study revealed only information about WES in one breast cancer patient and some participants among 10 healthy individuals had a family history of cancer. Consequently, the interpretation of results must be based on awareness, and more information about breast cancer patients is still needed to confirm results in the future.

In conclusion, the present study provided information on SNPs and CNVs in breast cancer-related genes. The associations between SNPs rs2231142 of *ABCG2,* rs76656413 in the promoter of *ABCG2* and rs799917 in *BRCA1* and breast cancer should be further investigated in case-controlled studies since their variant genotypes are more common in the Thai population. Moreover, based on the predicted number of breast cancer cases, strategies for modifiable breast cancer risk factors should be applied to prevent and control breast cancer risk. The oncogenic HPV in the whole blood might not be a breast cancer-associated risk factor in the Northern Thai population. However, a larger sample size might be needed for further investigation to confirm this.

Acknowledgements

Not applicable.

Funding

The present study was funded by the University of Phayao (grant no. FF64-RIB005) and the Faculty of Pharmaceutical Sciences, Ubon Ratchathani University (grant nos. 0604.11-3/2566 and grant no. 0604.11-6/2567).

Availability of data and materials

The data generated in the present study may be requested from the corresponding author. The WES of the breast cancer patient was submitted to the SRA database (BioProject ID PRJNA1126023; https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA1126023).

Authors' contributions

SuB and SJ were responsible for conceptualization, methodology, validation, data curation, and writing, reviewing, and editing the manuscript. SuB and SupP was responsible for sample collection. SuB, SD, SiB, SutP and SJ were responsible for the investigation. SuB, SD, SiB, SupP and SJ were responsible for funding acquisition. SuB and SJ confirm the authenticity of all the raw data. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The present study was approved by the Committee on Human Research Ethics in Health Sciences and Science and

Technology, University of Phayao (approval nos. UP-HEC 1.3/023/63 and 1.3/013/65) and the Ubonratchathani University (approval no. UBU-REC-68/2567), Thailand. Informed consent was obtained from all subjects involved in the present study. All procedures involving human participants performed in the study were in accordance with the ethical standards of the Declaration of Helsinki, the Belmont Report, the Council for International Organizations of Medical Sciences guidelines, and the International Conference on Harmonization in Good Clinical Practice.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71: 209-249, 2021.
2. Xu Y, Gong M, Wang Y, Yang Y, Liu S and Zeng Q: Global trends and forecasts of breast cancer incidence and deaths. Sci Data 10: 334, 2023.
3. Ferlay J, Ervik M, Lam F, Laversanne M, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I and Bray F: Global cancer observatory: Cancer today. Lyon, France: International Agency for Research on Cancer, 2024. Available from: https://gco.iarc.who.int/today, accessed (15 February 2024).
4. Łukasiewicz S, Czeczelewski M, Forma A, Baj J, Sitarz R and Stanisławek A: Breast cancer-epidemiology, risk factors, classification, prognostic markers, and current treatment strategies-an updated review. Cancers (Basel) 13: 4287, 2021.
5. Youn HJ and Han W: A review of the epidemiology of breast cancer in Asia: Focus on risk factors. Asian Pac J Cancer Prev 21: 867-880, 2020.
6. Momenimovahed Z and Salehiniya H: Epidemiological characteristics of and risk factors for breast cancer in the world. Breast Cancer (Dove Med Press) 11: 151-164, 2019.
7. Takaoka M and Miki Y: BRCA1 gene: Function and deficiency. Int J Clin Oncol 23: 36-44, 2018.
8. Qin TT, Chen T, Zhang Q, Du HN, Shu YQ, Luo K and Zhu LJ: Association between BRCA1 rs799917 polymorphism and breast cancer risk: A meta-analysis of 19,878 subjects. Biomed Pharmacother 68: 905-910, 2014.
9. Yang M, Du X, Zhang F and Yuan S: Association between BRCA1 polymorphisms rs799917 and rs1799966 and breast cancer risk: A meta-analysis. J Int Med Res 47: 1409-1416, 2019.
10. Zhao H and Kong Z: Lack of association of BRCA1 gene rs799917 polymorphism with cancer risk. Int J Clin Exp Med 10: 15886-15895, 2017.
11. Xu GP, Zhao Q, Wang D, Xie WY, Zhang LJ, Zhou H, Chen SZ and Wu LF: The association between BRCA1 gene polymorphism and cancer risk: A meta-analysis. Oncotarget 9: 8681-8694, 2018.
12. Wu H, Liu Y, Kang H, Xiao Q, Yao W, Zhao H, Wang E and Wei M: Genetic variations in ABCG2 gene predict breast carcinoma susceptibility and clinical outcomes after treatment with anthracycline-based chemotherapy. Biomed Res Int 2015: 279109, 2015.
13. Ghafouri H, Ghaderi B, Amini S, Nikkhoo B, Abdi M and Hoseini A: Association of ABCB1 and ABCG2 single nucleotide polymorphisms with clinical findings and response to chemotherapy treatments in Kurdish patients with breast cancer. Tumour Biol 37: 7901-7906, 2016.
14. Al-Eitan LN, Rababa'h DM, Alghamdi MA and Khasawneh RH: Role of Four ABC transporter genes in pharmacogenetic susceptibility to breast cancer in Jordanian patients. J Oncol 2019: 6425708, 2019.
15. Ren C, Zeng K, Wu C, Mu L, Huang J and Wang M: Human papillomavirus infection increases the risk of breast carcinoma: A large-scale systemic review and meta-analysis of case-control studies. Gland Surg 8: 486-500, 2019.
16. Graham R, Gazinska P, Zhang B, Khiabany A, Sinha S, Alaguthurai T, Flores-Borja F, Vicencio J, Beuron F, Roxanis I, *et al*: Serum-derived extracellular vesicles from breast cancer patients contribute to differential regulation of T-cell-mediated immune-escape mechanisms in breast cancer subtypes. Front Immunol 14: 1204224, 2023.
17. Khodabandehlou N, Mostafaei S, Etemadi A, Ghasemi A, Payandeh M, Hadifar S, Norooznezhad AH, Kazemnejad A and Moghoofei M: Human papilloma virus and breast cancer: The role of inflammation and viral expressed proteins. BMC Cancer 19: 61, 2019.
18. Kudela E, Kudelova E, Kozubík E, Rokos T, Pribulova T, Holubekova V and Biringer K: HPV-associated breast cancer: Myth or fact? Pathogens 11: 1510, 2022.
19. Chen ACH, Keleher A, Kedda MA, Spurdle AB, McMillan NAJ and Antonsson A: Human papillomavirus DNA detected in peripheral blood samples from healthy Australian male blood donors. J Med Virol 81: 1792-1796, 2009.
20. Pongpakdeesakul S, Ekalaksananan T, Pientong C, Iamchuen N, Buddhisa S, Mahingsa K, Pingyod A, Sangsrijun W, Passorn S, Chopjitt P, *et al*: Human oncogenic epstein-barr virus in water and human blood infection of communities in Phayao Province, Thailand. Water 15: 323, 2023.
21. Sumala S, Ekalaksananan T, Pientong C, Buddhisa S, Passorn S, Duangjit S, Janyakhantikul S, Suktus A and Bumrungthai S: The association of HHV-6 and the TNF-α (-308G/A) promotor with major depressive disorder patients and healthy controls in Thailand. Viruses 15: 1898, 2023.
22. Słomka M, Sobalska-Kwapis M, Wachulec M, Bartosz G and Strapagiel D: High resolution melting (HRM) for high-throughput genotyping-limitations and caveats in practical case studies. Int J Mol Sci 18: 2316, 2017.
23. Eclov RJ, Kim MJ, Smith R, Ahituv N and Kroetz DL: Rare variants in the ABCG2 promoter modulate in vivo activity. Drug Metab Dispos 46: 636-642, 2018.
24. de Roda Husman AM, Walboomers JM, van den Brule AJ, Meijer CJ and Snijders PJ: The use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. J Gen Virol 76: 1057-1062, 1995.
25. Lilyquist J, Ruddy KJ, Vachon CM and Couch FJ: Common genetic variation and breast cancer risk-past, present, and future. Cancer Epidemiol Biomarkers Prev 27: 380-394, 2018.
26. Behravan H, Hartikainen JM, Tengström M, Kosma VM and Mannermaa A: Predicting breast cancer risk using interacting genetic and demographic factors and machine learning. Sci Rep 10: 11044, 2020.
27. Sartor H, Brandt J, Grassmann F, Eriksson M, Czene K, Melander O and Zackrisson S: The association of single nucleotide polymorphisms (SNPs) with breast density and breast cancer survival: The Malmö diet and cancer study. Acta Radiol 61: 1326-1334, 2020.
28. Yang H, Ting X, Geng YH, Xie Y, Nierenberg JL, Huo YF, Zhou YT, Huang Y, Yu YQ, Yu XY, *et al*: The risk variant rs11836367 contributes to breast cancer onset and metastasis by attenuating Wnt signaling via regulating NTN4 expression. Sci Adv 8: eabn3509, 2022.
29. Sellers TA: Genetic factors in the pathogenesis of breast cancer: Their role and relative importance. J Nutr 127 (5 Suppl): 929S-932S, 1997.
30. De Silva S, Tennekoon KH and Karunanayake EH: Overview of the genetic basis toward early detection of breast cancer. Breast Cancer (Dove Med Press) 11: 71-80, 2019.
31. Sokolova A, Johnstone KJ, McCart Reed AE, Simpson PT and Lakhani SR: Hereditary breast cancer: Syndromes, tumour pathology and molecular testing. Histopathology 82: 70-82, 2023.
32. Wendt C and Margolin S: Identifying breast cancer susceptibility genes-a review of the genetic background in familial breast cancer. Acta Oncol 58: 135-146, 2019.
33. Sun YS, Zhao Z, Yang ZN, Xu F, Lu HJ, Zhu ZY, Shi W, Jiang J, Yao PP and Zhu HP: Risk factors and preventions of breast cancer. Int J Biol Sci 13: 1387-1397, 2017.
34. Hill AD, Doyle JM, McDermott EW and O'Higgins NJ: Hereditary breast cancer. Br J Surg 84: 1334-1339, 1997.
35. Collins A and Politopoulos I: The genetics of breast cancer: Risk factors for disease. Appl Clin Genet 4: 11-19, 2011.

36. Kumaran M, Cass CE, Graham K, Mackey JR, Hubaux R, Lam W, Yasui Y and Damaraju S: Germline copy number variations are associated with breast cancer risk and prognosis. Sci Rep 7: 14621, 2017.
37. Pan X, Hu X, Zhang YH, Chen L, Zhu L, Wan S, Huang T and Cai YD: Identification of the copy number variant biomarkers for breast cancer subtypes. Mol Genet Genomics 294: 95-110, 2019.
38. Dennis J, Tyrer JP, Walker LC, Michailidou K, Dorling L, Bolla MK, Wang Q, Ahearn TU, Andrulis IL, Anton-Culver H, *et al*: Rare germline copy number variants (CNVs) and breast cancer risk. Commun Biol 5: 65, 2022.
39. Zhang W and Yu Y: The important molecular markers on chromosome 17 and their clinical impact in breast cancer. Int J Mol Sci 12: 5672-5683, 2011.
40. Chen C, Lin CJ, Pei YC, Ma D, Liao L, Li SY, Fan L, Di GH, Wu SY, Liu XY, *et al*: Comprehensive genomic profiling of breast cancers characterizes germline-somatic mutation interactions mediating therapeutic vulnerabilities. Cell Discov 9: 125, 2023.
41. Lertwilaiwittaya P, Roothumnong E, Nakthong P, Dungort P, Meesamarnpong C, Tansa-Nga W, Pongsuktavorn K, Wiboonthanasarn S, Tititumjariya W, Thongnoppakhun W, *et al*: Thai patients who fulfilled NCCN criteria for breast/ovarian cancer genetic assessment demonstrated high prevalence of germline mutations in cancer susceptibility genes: Implication to Asian population testing. Breast Cancer Res Treat 188: 237-248, 2021.
42. Sukpan P, Sangkhathat S, Sriplung H, Laochareonsuk W, Choochuen P, Auseng N, Khoonjan W, Salaeh R, Thangnaphadol K, Wanawanakorn K and Kanokwiroon K: Exome sequencing reveals novel germline variants in breast cancer patients in the southernmost region of Thailand. J Pers Med 13: 1587, 2023.
43. Manchanda R, Patel S, Gordeev VS, Antoniou AC, Smith S, Lee A, Hopper JL, MacInnis RJ, Turnbull C, Ramus SJ, *et al*: Cost-effectiveness of population-based BRCA1, BRCA2, RAD51C, RAD51D, BRIP1, PALB2 mutation testing in unselected general population women. J Natl Cancer Inst 110: 714-725, 2018.
44. Nicoloso MS, Sun H, Spizzo R, Kim H, Wickramasinghe P, Shimizu M, Wojcik SE, Ferdin J, Kunej T, Xiao L, *et al*: Single-nucleotide polymorphisms inside microRNA target sites influence tumor susceptibility. Cancer Res 70: 2789-2798, 2010.
45. Smolarz B, Bryś M, Forma E, Zadrożny M, Bieńkiewicz J and Romanowicz H: Data on single nucleotide polymorphism of DNA repair genes and breast cancer risk from Poland. Pathol Oncol Res 25: 1311-1317, 2019.
46. Ahmad J, Le Calvez-Kelm F, Daud S, Voegele C, Vallée M, Ahmad A, Kakar N, McKay JD, Gaborieau V, Léoné M, *et al*: Detection of BRCA1/2 mutations in breast cancer patients from Thailand and Pakistan. Clin Genet 82: 594-598, 2012.
47. Stiburkova B, Pavelcova K, Zavada J, Petru L, Simek P, Cepek P, Pavlikova M, Matsuo H, Merriman TR and Pavelka K: Functional non-synonymous variants of ABCG2 and gout risk. Rheumatology (Oxford) 56: 1982-1992, 2017.
48. Toyoda Y, Pavelcová K, Bohatá J, Ješina P, Kubota Y, Suzuki H, Takada T and Stiburkova B: Identification of two dysfunctional variants in the ABCG2 urate transporter associated with pediatric-onset of familial hyperuricemia and early-onset gout. Int J Mol Sci 22: 1935, 2021.
49. Keskitalo JE, Zolk O, Fromm MF, Kurkinen KJ, Neuvonen PJ and Niemi M: ABCG2 polymorphism markedly affects the pharmacokinetics of atorvastatin and rosuvastatin. Clin Pharmacol Ther 86: 197-203, 2009.
50. Phan L, Jin H, Zhang W, Qiang E, Shekhtman D, Shao D, Revoe R, Villamarin E, Ivanchenko M, Kimura ZY, *et al*: 'ALFA: Allele frequency aggregator'. national center for biotechnology information, U.S. National Library of Medicine, 10 Mar, 2020. www.ncbi.nlm.nih.gov/snp/docs/gsr/alfa/.
51. Bircsak KM, Moscovitz JE, Wen X, Archer F, Yuen PYS, Mohammed M, Memon N, Weinberger BI, Saba LM, Vetrano AM, *et al*: Interindividual regulation of the breast cancer resistance protein/ABCG2 transporter in term human placentas. Drug Metab Dispos 46: 619-627, 2018.
52. Ngamkham J, Karalak A, Chaiwerawattana A, Sornprom A, Thanasutthichai S, Sukarayodhin S, Mus-u-Dee M, Boonmark K, Phansri T and Laochan N: Prevalence of human papillomavirus infection in breast cancer cells from Thai women. Asian Pac J Cancer Prev 18: 1839-1845, 2017.
53. Bumrungthai S, Ekalaksananan T, Duangchai D, Lanpol P, Panya P, Kattiwong F, Acharya S and Pientong C: Prevalence of human papillomavirus in oral rinse samples from healthy individuals in northern Thailand. J Oral Pathol Med 48: 159-165, 2019.