

## Polymorphisms in three base excision repair genes and breast cancer risk in Thai women

Suleeporn Sangrajrang · Peter Schmezer · Iris Burkholder · Peter Waas · Paolo Boffetta · Paul Brennan · Helmut Bartsch · Surapon Wiangnon · Odilia Popanda

Received: 18 September 2007 / Accepted: 18 September 2007 / Published online: 6 October 2007  
© Springer Science+Business Media, LLC. 2007

**Abstract** DNA repair plays an important role in tumor development. The base excision repair (BER) pathway mainly removes DNA damage caused by ionizing radiation and reactive oxidative species. Here, we examined possible associations between polymorphisms in three important BER genes (*OGG1* Ser326Cys, *APEX1* Asp148Glu, *XRCC1* Arg194Trp, *XRCC1* Arg280His, *XRCC1* Arg399Gln) and breast cancer incidence in Thai women. The study population consisted of 507 breast cancer cases and 425 controls. Odds ratios (OR) were adjusted by multivariate logistic regression analysis for age, body mass index, age at menarche, family history of breast cancer, menopausal status, reproduction parameters, use of contraceptives, tobacco smoking, involuntary tobacco smoking, alcohol drinking, and education. For homozygous carriers of the Glu allele in *APEX1*, a significant protective effect was found when compared to Asp/Asp carriers (odds

ratio (OR) = 0.60, 95% confidence interval (CI) = 0.38–0.94). Subgroup analysis based on menopausal status revealed increased breast cancer risk in postmenopausal women and *OGG1* (OR = 2.05, 95% CI 1.14–3.69). Reconstructed diplotypes for *XRCC1* showed that CGA/CGA carriers had an increased risk of breast cancer compared with carriers of the wild type diplotype CGG/CGG (OR = 2.56, 95% CI 1.28–5.15). When the joint effects of *XRCC1*, *APEX1* and *OGG1* polymorphisms were evaluated, individuals homozygous for two or three risk alleles were at increased risk (OR = 1.88, 95% CI 1.26–2.82). In conclusion, our data suggest that Thai women with a certain *XRCC1* diplotype or homozygous for two or three variant alleles of *XRCC1*, *OGG1*, and *APEX1* are likely to have an increased susceptibility to breast cancer.

**Keywords** APEX1 · Case-control study · Menopausal status · OGG1 · Oxidative DNA damage · XRCC1

S. Sangrajrang (✉)  
Research Division, National Cancer Institute, 286/1 Rama VI  
road Ratchathewe, Bangkok 10400, Thailand  
e-mail: sulee@health.moph.go.th; suleesang@yahoo.com

P. Schmezer · P. Waas · H. Bartsch · O. Popanda  
Division of Toxicology and Cancer Risk Factors, German  
Cancer Research Center (DKFZ), Heidelberg, Germany

I. Burkholder  
Division of Biostatistics, German Cancer Research Center  
(DKFZ), Heidelberg, Germany

P. Boffetta · P. Brennan  
International Agency for Research on Cancer (IARC), Lyon,  
France

S. Wiangnon  
Faculty of Medicine, Khon Kaen University, Khon Kaen,  
Thailand

### Introduction

Breast cancer is the second most common cancer among women in Thailand and its incidence is still increasing [1]. Although there are well established risk factors such as age at first child's birth, nulliparity, and family history of breast cancer, the aetiology of breast cancer is still not completely known [2]. There is, however, strong evidence that insufficient repair of DNA damage plays an important role in the development of breast cancer [2–5]. Besides DNA double-strand break repair, efficiency of base excision repair (BER) is suggested to be a major determinant of breast cancer risk as BER is involved in the repair of oxidative DNA lesions which are induced by free radicals produced

e.g., during cellular estrogen metabolism or by exogenous exposure to chemicals and ionizing radiation [6, 7].

The BER pathway consists of at least 11 DNA damage specific glycosylases and more than 20 further proteins [8], three of the most important ones being 8-oxoguanine DNA glycosylase (*hOGGI*), apurinic/apyrimidine endonuclease 1 (*APEX1*), and X-ray repair cross-complementing group 1 protein (*XRCC1*) [9, 10]. BER proteins are cooperating at the damaged site in a strongly coordinated way, e.g., an oxidized 8-oxoguanine base is recognized and removed by the OGG1 glycosylase leaving behind an apurinic site which is excised by APEX1 and subsequently re-polymerized by DNA polymerase  $\beta$  and ligase 3. XRCC1 acts as a scaffolding protein and can regulate time and strength of interaction between the multiple proteins of the BER machinery.

Imbalances in this tightly regulated process can be caused by single nucleotide polymorphisms and will result in insufficient DNA repair and an increase in DNA breaks [8, 9]. There are several common polymorphisms known affecting the amino acid sequence in BER genes, among them Arg194Trp (rs1799782), Arg280His (rs25489) and Arg399Gln in *XRCC1* (rs25487), Asp148Glu in *APEX1* (rs3136820), and Ser326Cys in *OGGI* (rs1052133). The *XRCC1* 399Gln allele and the *OGGI* 326Cys allele were shown to be associated with reduced DNA repair capacity as assessed by the persistence of DNA bulky adducts [11–13] and  $\gamma$ -irradiation induced oxidative DNA damage [14]. Decreased DNA repair capacity after exposure to oxidative agents was also found when cells carrying the 280His allele were compared with isogenic cells carrying the 280Arg allele [15]. In addition, an increasing number of variant alleles of *APEX1* Asp148Glu and *XRCC1* Arg399Gln was significantly associated with prolonged cell cycle delay in G2 phase and decreased DNA repair capacity after  $\gamma$ -irradiation [14, 16]. In contrast, the 194Trp variant of *XRCC1* was reported to be associated with lower bleomycin and benzo(a)pyrene diol-epoxide sensitivity in vitro [17], suggesting a protective role for the Trp-Allele in the development of cancer potentially by increasing BER activity.

Associations of these polymorphisms and breast cancer risk have been examined by epidemiologic studies and results have been summarized in three meta-analyses [18–20]. Only a few studies have investigated associations between breast cancer risk and the genetic variants *OGGI* Ser326Cys, *APEX1* Asp148Glu and *XRCC1* Arg280His. Reasonable numbers of studies are available only for the *XRCC1* Arg194Trp and Arg399Gln polymorphisms. The summary estimates for the 399 polymorphism presented an increased risk in Asian and African homozygous carriers of the 399 Gln allele but not in Caucasians [20]. Data on Asian women are however based on small sample size. Studies on Thai women are missing. Therefore, it is unclear

whether the observed differences by race have a biological basis and further evidence is needed to make a conclusion about the potential differences in relative risk by race.

Here, we present data on the association between these polymorphisms and breast cancer risk in a case-control study of Thai women. In addition, we determined the effects of the *XRCC1* haplotypes and analyzed whether the combined occurrence of these polymorphisms in genes of the BER pathway affects breast cancer risk.

## Material and methods

### Study population

Cases were all new incident breast cancer patients histologically diagnosed at the National Cancer Institute in Bangkok and at the hospital in Khon Kaen province of North Eastern of Thailand during the period of May 2002 to March 2004 [21], with a participation rate of 99.6% (554/556). Controls were randomly selected from healthy women who visited patients admitted to the same hospitals for diseases other than breast or ovarian cancer. The participation rate among visitors who were asked to participate was 98.7% (572/579). Informed consent was obtained from all participants and a structured questionnaire was administered by trained interviewers to collect information on demographic and anthropometric data, reproductive and medical history, residential history, physical activity and occupation as well as diet. Lifestyle exposure parameters were reported as follows: tobacco smoking: less than or equal to 50 cigarettes (non-smoker) versus more than 50 cigarettes over a 6-month period (smoker); involuntary tobacco smoking: less versus more than or equal to 1 h of exposure per day during the last 2 years (sum of three sources: from the spouse, at the workplace or in public settings); alcohol consumption: less versus more than or equal to once a week for at least 6 months. Approximately 7 ml of blood were collected from participants, but 45 cases and 149 controls refused to give blood samples. In total, blood samples of 507 cases and 425 controls were included in the genotype analysis, resulting in a participation rate of 91.2% (cases) and 73.4% (controls). Genomic DNA was isolated from buffy coats using a QIAmp DNA blood kit (Qiagen, Hilden, Germany). The study was approved by the ethical review committee for research in human subjects, Ministry of Public Health, Thailand.

### Genotyping analysis

Detection of polymorphisms was performed by rapid capillary PCR followed by melting curve analysis using

fluorescence labeled hybridization probes in a LightCycler (Roche Diagnostics, Mannheim, Germany) as described [22]. Primers and probes were designed with the help of Tib Molbiol (Berlin, Germany), and are given in Table 1.

Analysis was performed in 10- $\mu$ l volumes in glass capillary tubes (Roche Diagnostics) with Qiagen reagents: 1x PCR Buffer, 2.5 mM total  $MgCl_2$ , 0.2 mM dNTPs, 0.1% bovine serum albumin in 2 mM Tris-HCl/2.5 % glycerol, 1x Q-solution (not for *APEX1* Asp148Glu), 0.1 U Taq DNA polymerase, 0.5  $\mu$ M of each primer, 0.15  $\mu$ M of each probe and 10 ng of DNA. For *APEX1* Asp148Glu, 2  $\mu$ M of anchor and sensor were used. Reaction conditions were as follows: initial denaturation at 95°C for 3 min, then 40 cycles of denaturation at 95°C for 15 s, annealing at 58°C (*OGGI* Ser326Cys, *APEX1* Asp148Glu, *XRCC1* Arg194Trp and Arg280His), or 55°C (*XRCC1* Arg399Gln) for 10 s and elongation at 72°C for 12 s. Melting curve analysis was performed with an initial denaturing step at 95°C for 30 s, 20 s at 40°C, slow heating to 75°C with a ramping rate 0.2°C/s and continuous fluorescence detection. Genotypes were assigned by comparing melting

curves of unknown samples with those of samples with known (sequenced) genotype. A negative control containing all reagents but with water instead of the DNA template was included to each amplification. Melting points were evaluated by two independent observers who were unaware to the clinical diagnosis. In addition, 15% of randomly selected samples were repeated independently to verify genotyping results and 100% concordance was found. Because of an inadequate amount of DNA, a few samples did not generate complete information for all polymorphisms.

#### Data analysis

Breast cancer patients were compared with controls for basic demographic and lifestyle characteristics. Genotypes of both cases and controls were tested as to whether they were in Hardy-Weinberg equilibrium using the  $\chi^2$ -test of goodness of fit with one degree of freedom, with respect to the distribution of the considered allele groups. Because the

**Table 1** Primers and probes for genotyping of single-nucleotide polymorphisms

Gene and polymorphism (dbSNP ID)	Sequence of primer and probes <sup>a</sup>
<i>XRCC1</i> Arg194Trp (C26304T) (rs1799782)	F: 5'-TACTGACCTTGCGGGACCTTA-3' R: 5'-TACCCTCAGACCCACGAGTCTA-3' S: 5'-TGTCTTGTTGATCCAGCTGAAGAAG- FITC -3' A: 5'-LCRed640-AGCCCCGGCCTCAGAGAGTTGGp-3'
<i>XRCC1</i> Arg280His (G27466A) (rs25489)	F: 5'-GCTGGGGCCTGGATTG-3' R: 5'-GCACCACTACCACACCTGAA-3' S: 5'-TGCCAGCTCCAACCTCATACCCC- FITC -3' A: 5'-LCRed640-CCACAGCCCCAGTCCCTGCCCGp-3'
<i>XRCC1</i> Arg399Gln (G28152A) (rs25487)	F: 5'-CCCCAAGTACAGCCAGGTC-3' R: 5'-TGTCCCGCTCCTCTCAGTAG-3' S: 5'-CCCTCCCAGAGGTAAGGCC- FITC -3' A: 5'-LCRed640-CACACGCCAACCTGCTCCTTATp-3'
<i>APEX1</i> Asp148Glu (T2197G) (rs3136820)	F: 5'-CTTGATTGCTTTCCTTTTCTTA-3' R: 5'-CGCTGCCGGTACTCCA-3' S: 5'-LCRed640-TGCTCCTCTCGCCTATAGAAATGAp-3' A: 5'-CACAATCACCCGGCCTTCTGTATC-FITC-3'
<i>OGGI</i> Ser326Cys (C9072G) (rs1052133)	F: 5'-CCCAACTGTCACTAGTCTCA-3' R: 5'-TTGGGAATTTCTTTGTCCA-3' S: 5'-CGCCAATCCGCCATGCTCA- FITC -3' A: 5'-LCRed640-CCACCAGCAAAGCGCAGAAAGGGp-3'

<sup>a</sup> F, forward primer; R, reverse primer; S, sensor; A, anchor; mutation site is shaded

*XRCC1* 208His allele is uncommon in this population, individuals with genotypes Arg/His and His/His were combined into one group as 208His allele carriers and compared with Arg/Arg homozygotes as the reference.

For analyzing the association of breast cancer risk with polymorphisms, odds ratios (ORs) and their 95% confidence intervals (CI) were calculated and assessed for statistical significance according to [23], both as crude ORs and as adjusted ORs. Multivariate unconditional logistic regression analysis was performed to assess the association between occurrence of breast cancer and prevalence of polymorphisms and to adjust for potential confounders. Covariates were selected when either being significant in the univariate analysis at the level of 5% or when being considered as relevant factor for the occurrence of breast cancer in general. Family history of breast cancer (FH), menopausal status, use of contraceptives, and education ( $\leq 9$  years,  $> 9$  years) were incorporated in the model as binary predictors. Alcohol consumption, active and involuntary smoking were combined into one binary variable “hazardous lifestyle” (0 = no regularly alcohol consumption and no active and no involuntary smoking, 1 = else). In the analysis of all subjects, reproduction was a combination of pregnancy and breast-feeding in five categories: non-pregnant, age at 1st pregnancy  $\leq 22$  years and non breast-feeding, age at 1st pregnancy  $\leq 22$  years and breast-feeding, age at 1st pregnancy  $> 22$  years and non breast-feeding, age at 1st pregnancy  $> 22$  years and breast-feeding. To deal with possible non-linearity, continuous predictors (age, body mass index [BMI] and age at menarche) were modeled by using fractional polynomials [24]. Finally, ORs resulted from multivariate logistic regression model including age (non-linear), BMI, FH, age at menarche, reproduction parameters (5 classes), menopausal status, use of contraceptives, hazardous lifestyle, and education.

Possible interaction effects for age  $\times$  BMI, polymorphism  $\times$  hazardous lifestyle, polymorphism  $\times$  BMI, and polymorphism  $\times$  reproduction were tested by introducing an interaction term into the logistic regression model using the standard Wald test. Significance level alerting was set to 0.1 in accounting for the lower power to test for interaction compared to testing for single covariate's effects.

In addition, subgroups based on menopausal status were analysed. Here, reproduction was categorized as non-pregnant, age at 1st pregnancy  $\leq 22$  years, and age at 1st pregnancy  $> 22$  years. Thus, ORs in subgroups were adjusted by age (linear), BMI, FH, age at menarche, reproduction parameters (three classes), use of contraceptives, hazardous lifestyle, and education. Statistical analyses were performed using the statistical packages SAS (SAS Institute, Cary, NC) for Windows Version 9.

Linkage disequilibrium between the different markers in *XRCC1* was estimated using the “Estimating Haplotypes” program [25, 26]. Haplotypes pairs of *XRCC1* were reconstructed using the PHASE V2.1 online software [27]. The haplotype specific risks were investigated by univariate and multivariate logistic regression analysis as described for the single nucleotide polymorphisms using the most probable haplotype pairs yielded by the PHASE software.

To evaluate the potential combined effects of *XRCC1* Arg399Gln, *APEX1* Asp148Glu, and *OGGI* Ser326Cys on breast cancer risk, the number of homozygous variant alleles per individual was introduced in the model and crude as well as adjusted ORs were calculated.

## Results

Characteristics of the study population were compared by case-control status, as shown in Table 2. The mean age of controls ( $45.3 \pm 12.2$  years) was significantly lower ( $P < 0.01$ ) than that of breast cancer patients ( $48.0 \pm 10.0$  years). Pregnancy, breast feeding, BMI, involuntary tobacco smoking, FH, and education were different between cases and controls. However, as to oral contraceptive use, menopausal status, smoking, and alcohol consumption, no significant differences were found between cases and controls.

Frequencies of the variant alleles were 0.31 (*XRCC1* 194Trp), 0.07 (*XRCC1* 280His), 0.23 (*XRCC1* 399Gln), 0.34 (*APEX1* 148Glu) and 0.50 (*OGGI* 326Cys) in controls (Table 3). Genotype frequencies were in Hardy-Weinberg equilibrium in both cases and controls. In *XRCC1* Arg280His, only 3 cases and 2 controls were found to be homozygous for the His allele. Thus, carriers of Arg/His and His/His were combined for further analysis and compared to Arg/Arg carriers as the reference.

Univariate and multivariate logistic regression analysis of the overall effect of the polymorphisms (Table 3) did not give evidence that the *XRCC1* 194Trp, 280His, and 399Gln alleles were associated with breast cancer (OR = 0.70, 95% CI 0.42–1.14; OR = 1.30, 95% CI 0.88–1.93; and OR = 1.80, 95% CI 0.99–3.29). However, the *APEX1* 148Glu allele showed a significantly protective effect on breast cancer risk (OR = 0.60, 95% CI 0.38–0.94) when adjusted for age, BMI, family history of breast cancer, age at menarche, reproduction parameters (five classes), menopausal status, use of contraceptives, hazardous lifestyle, and education. For the *OGGI* Cys/Cys genotype, an adjusted OR of 1.42 was calculated (95% confidence limits 0.97–2.09). No interaction effects were observed between age and BMI, as well as between polymorphisms and hazardous lifestyle, BMI, or reproduction ( $P > 0.85$ ).

**Table 2** Selected characteristics of study population

Characteristics	Cases	Controls	P-value
Age, <i>n</i> (mean ± SD)	507, (48.0 ± 10.0)	425, (45.3 ± 12.2)	<0.01
Age at menarche, <i>n</i> (mean ± SD)	501, (14.9 ± 1.9)	425, (14.8 ± 1.8)	0.41
Age at menopause, <i>n</i> (mean ± SD)	235, (47.1 ± 5.3)	177, (46.8 ± 5.1)	0.49
Age at 1st pregnancy, <i>n</i> (mean ± SD)	391, (23.0 ± 5.4)	295, (22.8 ± 5.0)	0.67
Pregnancy ( <i>n</i> = 932), <i>n</i> (%)			
No	116 (22.9%)	130 (30.6%)	<0.01
Yes	391 (77.1%)	295 (69.4%)	
Breast feeding ( <i>n</i> = 686), <i>n</i> (%)			
No	45 (11.5%)	20 (6.8%)	0.04
Yes	346 (88.5%)	275 (93.2%)	
Oral contraceptive use ( <i>n</i> = 932), <i>n</i> (%)			
No	283 (55.8%)	256 (60.2%)	0.17
Yes	224 (44.2%)	169 (39.8%)	
Menopausal status ( <i>n</i> = 932), <i>n</i> (%)			
Premenopausal	268 (52.9%)	245 (57.6%)	0.14
Postmenopausal	239 (47.1%)	180 (42.4%)	
Body mass index (kg/m <sup>2</sup> )			
<i>n</i> (mean ± SD)	506 (24.2 ± 4.0)	425 (23.1 ± 4.0)	<0.01
Tobacco smoking ( <i>n</i> = 932), <i>n</i> (%)			
No	493 (97.2%)	419 (98.7%)	0.18
Yes	14 (2.8%)	6 (1.4%)	
Involuntary smoking ( <i>n</i> = 912), <i>n</i> (%)			
No	440 (89.3%)	395 (94.3%)	<0.01
Yes	53 (10.7%)	24 (5.7%)	
Alcohol consumption ( <i>n</i> = 932), <i>n</i> (%)			
No	471 (92.9%)	403 (94.8%)	0.28
Yes	36 (7.1%)	22 (5.2%)	
Family history of breast cancer in first-degree relatives ( <i>n</i> = 932), <i>n</i> (%)			
No	488 (96.3%)	419 (98.6%)	0.03
Yes	19 (3.7%)	6 (1.4%)	
Education ( <i>n</i> = 931), <i>n</i> (%)			
≤9 years	374 (73.9%)	238 (56.0%)	<0.01
>9 years	132 (26.1%)	187 (44.0%)	
Study area ( <i>n</i> = 932), <i>n</i> (%)			
Bangkok	292 (57.6%)	251 (59.1%)	0.65
Khon Kaen	215 (42.2%)	174 (40.9%)	

In the subgroup analysis based on menopausal status (Table 4), no effect of the polymorphisms was seen in premenopausal women, whereas there was a significant increase in breast cancer risk among postmenopausal women homozygous for the 326Cys variant of *OGGI* (OR = 2.05, 95% CI 1.14–3.69). The other polymorphisms did not contribute significantly to breast cancer risk in postmenopausal women.

A strong association (linkage disequilibrium) was found among the three *XRCCI* polymorphisms in codon 194, 280 and 399 (*P*-value < 0.0001) in cases and in controls. Reconstruction of haplotypes yielded four haplotypes

which were called CGG (Arg194/Arg280/Arg399), CGA (Arg194/Arg280/Gln399), CAG (Arg194/His280/Arg399), and TGG (Trp194/Arg280/Arg399). Haplotype frequencies in the overall study population were 36% for CGG (*n* = 673), 26% for CGA (*n* = 572), 7% for CAG (*n* = 140), and 31% for TGG (*n* = 475). The haplotypes with the variant sequence in all three loci (TAA) or in two loci (TAG, TGA, or CAA) did not occur. In addition, haplotype pairs were calculated for each individual (Table 5) and the association between haplotype-pair (diplotype) and breast cancer risk was analyzed. In total, 10 diplotypes occurred, the diplotype CAG/CAG was



**Table 3** Associations between *XRCCI*, *APEXI*, and *OGGI* polymorphisms and breast cancer

Genotype	Cases <i>n</i> (%)	Controls <i>n</i> (%)	Crude OR (95%CI)	<i>P</i> -value	Adjusted OR <sup>a</sup> (95%CI)	<i>P</i> -value
<i>Subjects analyzed</i>						
<i>XRCCI</i> Arg194Trp <sup>b</sup>						
Arg/Arg	239 (47.2)	202 (47.5)	1.00 (reference)		1.00 (reference)	
Arg/Trp	228 (45.1)	179 (42.1)	1.08 (0.82–1.41)	0.594	1.08 (0.81–1.44)	0.594
Trp/Trp	39 (7.7)	44 (10.4)	0.75 (0.47–1.20)	0.228	0.70 (0.42–1.14)	0.152
Trp allele frequency	0.30	0.31				
Hardy-Weinberg <sup>c</sup>	0.13	0.64				
<i>XRCCI</i> Arg280His						
Arg/Arg	431 (85.0)	366 (86.1)	1.00 (reference)		1.00 (reference)	
Arg/His + His/His	76 (15.0)	59 (13.9)	1.09 (0.76–1.58)	0.633	1.30 (0.88–1.93)	0.190
His allele frequency	0.08	0.07				
Hardy-Weinberg <sup>c</sup>	0.96	0.89				
<i>XRCCI</i> Arg399Gln <sup>b</sup>						
Arg/Arg	268 (52.9)	246 (58.0)	1.00 (reference)		1.00 (reference)	
Arg/Gln	201 (39.6)	158 (37.3)	1.17 (0.89–1.53)	0.262	1.20 (0.90–1.60)	0.211
Gln/Gln	38 (7.5)	20 (4.7)	1.74 (0.99–3.10)	0.055	1.80 (0.99–3.29)	0.056
Gln allele frequency	0.27	0.23				
Hardy-Weinberg <sup>c</sup>	0.97	0.40				
<i>APEXI</i> Asp148Glu						
Asp/Asp	250 (49.3)	194 (45.7)	1.00 (reference)		1.00 (reference)	
Asp/Glu	208 (41.0)	176 (41.4)	0.92 (0.70–1.21)	0.537	0.93 (0.69–1.25)	0.630
Glu/Glu	49 (9.7)	55 (12.9)	0.69 (0.45–1.06)	0.091	0.60 (0.38–0.94)	0.026
Glu allele frequency	0.30	0.34				
Hardy-Weinberg <sup>c</sup>	0.55	0.13				
<i>OGGI</i> Ser326Cys <sup>b</sup>						
Ser/Ser	112 (22.1)	104 (24.5)	1.00 (reference)		1.00 (reference)	
Ser/Cys	232 (45.9)	217 (51.2)	0.99 (0.72–1.37)	0.965	0.95 (0.67–1.34)	0.763
Cys/Cys	162 (32.0)	103 (24.3)	1.46 (1.02–2.10)	0.041	1.42 (0.97–2.09)	0.073
Cys allele frequency	0.55	0.50				
Hardy-Weinberg <sup>c</sup>	0.10	0.63				

<sup>a</sup> OR adjusted for age (non-linear using fractional polynomial functions), body mass index, age at menarche, family history of breast cancer, reproduction parameters (five classes), use of contraceptives, menopausal status, hazardous lifestyle, and education. Only patients with complete information were used for unconditional logistic regression analysis

<sup>b</sup> For missing samples, no PCR product was obtained

<sup>c</sup> *P*-value from Chi-square test

combined with CAG/TGG because of the small sample size of these two groups. For regression analysis, the diplotype CGG/CGG which consists of the wild type sequence in all loci was selected as reference. A significantly increased risk for breast cancer was found for individuals carrying the CGG/TGG, CGA/CGA, CGA/CAG and the CAG/CAG combined with CAG/TGG diplotype. The strongest effect was seen for CGA/CGA (Arg194/Arg280/Gln399) with an OR of 2.56 (95% CI 1.28–5.15; *P* = 0.008).

As already mentioned, the three genes analyzed in this study are involved in the BER pathway. We therefore explored the combined effects of variant alleles in these

genes on breast cancer risk. The analysis concentrated on individuals that were homozygous for the variant alleles because these individuals exhibited stronger effects of these alleles compared with heterozygous individuals for whom ORs were close to 1 (Table 3). As the results of the *XRCCI* diplotype analysis showed the strongest effect for the haplotype with the wild type sequence at the 194 and 280 loci, and the 399Gln variant at the third locus, the variant allele of *XRCCI* 399Gln was selected for combination analysis. For *APEXI* Asp148Glu, the 148Asp allele was designed as the risk allele because the Glu variant showed a protective effect. The reference population consisted of all individuals who were not homozygous for any

**Table 4** Associations between *XRCCI*, *APEXI*, and *OGGI* polymorphisms and breast cancer stratified by menopausal status

Genotype	Premenopausal women			Postmenopausal women		
	Cases/controls	OR <sup>a</sup> (95% CI)	<i>P</i> -value	Cases/controls	OR <sup>a</sup> (95%CI)	<i>P</i> -value
<i>Subjects analyzed</i>	268/245			239/180		
<i>XRCCI</i> Arg194Trp						
Arg/Arg	124/122	1.00 (reference)		115/80	1.00 (reference)	
Arg/Trp	121/100	1.23 (0.83–1.82)	0.302	107/79	0.92 (0.60–1.42)	0.716
Trp/Trp	23/23	0.88 (0.45–1.73)	0.705	16/21	0.55 (0.26–1.16)	0.116
<i>XRCCI</i> Arg280His						
Arg/Arg	232/207	1.00 (reference)		199/159	1.00 (reference)	
Arg/His + His/His	36/38	0.99 (0.58–1.70)	0.977	40/21	1.71 (0.94–3.10)	0.077
<i>XRCCI</i> Arg399Gln						
Arg/Arg	150/142	1.00 (reference)		118/104	1.00 (reference)	
Arg/Gln	101/92	1.05 (0.71–1.55)	0.810	100/66	1.30 (0.84–2.00)	0.236
Gln/Gln	17/10	1.45 (0.60–3.52)	0.414	21/10	1.84 (0.79–4.28)	0.157
<i>APEXI</i> Asp148Glu						
Asp/Asp	135/114	1.00 (reference)		115/80	1.00 (reference)	
Asp/Glu	110/105	0.91 (0.61–1.35)	0.632	98/71	1.00 (0.64–1.56)	0.992
Glu/Glu	23/26	0.62 (0.32–1.21)	0.158	26/29	0.58 (0.31–1.10)	0.091
<i>OGGI</i> Ser326Cys						
Ser/Ser	59/59	1.00 (reference)		53/45	1.00 (reference)	
Ser/Cys	125/120	0.92 (0.58–1.48)	0.741	107/97	0.97 (0.58–1.61)	0.900
Cys/Cys	84/66	1.13 (0.67–1.91)	0.636	78/37	2.05 (1.14–3.69)	0.016

<sup>a</sup> Only patients with complete information were used for unconditional logistic regression analysis. OR were adjusted for age (linear), body mass index, age at menarche, family history of breast cancer, reproduction parameters (three classes), use of contraceptives, hazardous lifestyle, and education

**Table 5** Associations between diplotypes of *XRCCI* and breast cancer

Diplotype <sup>a</sup>	Cases n	Controls n	Crude OR(95% CI)	<i>P</i> -value	Adjusted OR <sup>b</sup> (95% CI)	<i>P</i> -value
CGG/CGG	52	60	1.00 (reference)		1.00 (reference)	
CGG/CGA	94	78	1.39 (0.86–2.24)	0.176	1.66 (1.00–2.76)	0.049
CGG/CAG	26	27	1.11 (0.58–2.14)	0.752	1.45 (0.72–2.92)	0.300
CGG/TGG	126	98	1.48 (0.94–2.34)	0.090	1.65 (1.02–2.68)	0.042
CGA/CGA	38	20	2.19 (1.14–4.23)	0.019	2.56 (1.28–5.15)	0.008
CGA/CAG	26	15	2.00 (0.96–4.18)	0.065	2.20 (1.02–4.77)	0.045
CGA/TGG	81	65	1.44 (0.88–2.36)	0.150	1.64 (0.97–2.76)	0.065
CAG/CAG and CAG/TGG	24	17	1.63 (0.79–3.36)	0.186	2.78 (1.25–6.20)	0.013
TGG/TGG	39	44	1.02 (0.58–1.81)	0.938	1.09 (0.60–1.99)	0.779

<sup>a</sup> CGG (Arg194/Arg280/Arg399), CGA (Arg194/Arg280/Gln399), CAG (Arg194/His280/Arg399), TGG (Trp194/Arg280/Arg399)

<sup>b</sup> OR were adjusted for age (non-linear using fractional polynomial functions), body mass index, age at menarche, family history of breast cancer, reproduction parameters (five classes), use of contraceptives, menopausal status, hazardous lifestyle, and education. Only patients with complete information were used for unconditional logistic regression analysis

of the three risk alleles. Table 6 summarizes the joint effects of the *XRCCI* 399Gln, *APEXI* 148Asp and *OGGI* 326Cys alleles on breast cancer risk when carriers are homozygous for one and more of these variants. Both crude and adjusted ORs were significantly enhanced for carriers of two or three homozygous risk alleles (OR = 1.88, 95% CI 1.26–2.82; *P* = 0.002).

## Discussion

This study investigated possible associations between five polymorphisms in BER genes (*OGGI* Ser326Cys, *APEXI* Asp148Glu, *XRCCI* Arg194Trp, *XRCCI* Arg280His, and *XRCCI* Arg399Gln) and breast cancer risk for the first time in Thai women. Frequencies obtained among controls for

**Table 6** Breast cancer risk in individuals homozygous for more than one variant allele: Joint effects of variants in base excision repair (*XRCC1* 399Gln, *APEX1* 148Asp, and *OGG1* 326Cys)

Homozygous variant alleles per individual	Cases	Controls	Crude odds ratio			Adjusted odds ratio <sup>a</sup>		
	<i>n</i> (%)	<i>n</i> (%)	OR	95% CI	<i>P</i> -value	OR	95% CI	<i>P</i> -value
0	174 (34.4)	173 (40.9)	1.00	(reference)		1.00	(reference)	
1	225 (44.5)	191 (45.2)	1.17	0.88–1.56	0.278	1.17	0.87–1.60	0.308
2	99 (19.6)	58 (13.7)	1.70	1.15–2.50	0.007	1.76	1.17–2.66	0.007
3	8 (1.6)	1 (0.2)	7.93	0.98–63.99	0.052	8.80	1.05–73.9	0.045
2 or 3	107 (21.2)	59 (14.0)	1.80	1.23–2.64	0.002	1.88	1.26–2.82	0.002

<sup>a</sup> OR were adjusted for age (non-linear using fractional polynomial functions), body mass index, age at menarche, family history of breast cancer, reproduction parameters (five classes), use of contraceptives, menopausal status, hazardous lifestyle, and education. Only patients with complete information were used for unconditional logistic regression analysis

the rare alleles of all polymorphisms were consistent with frequencies found among other Asian populations [19]. Our data also confirmed that the *XRCC1* 194Trp and the *OGG1* 326Cys alleles are more frequent among Asians than among Caucasians. These differences may account for a different contribution of these polymorphisms to breast cancer in Asians, and especially in a Thai population.

All polymorphisms investigated were suggested recently to modify BER activity by correlating various cellular endpoints with genotype (see Introduction). Reduced BER capacity elongates the persistence of oxidative DNA damage in breast epithelial cells and may thus increase the mutational load in these cells and contribute to breast cancer risk.

We found no association between *XRCC1* Arg194Trp and Arg280His and breast cancer risk among Thai women. Although studies on Arg194Trp had contradictory outcomes, our result corresponds to risk estimates obtained by meta-analyses for Arg194Trp [19, 20] which showed that this polymorphism did not modify breast cancer risk regardless of ethnic origin. The Arg280His polymorphism is a very rare allele in both Caucasian and Asian populations. Therefore, only few studies have determined this polymorphism and can be compared with our data. For Caucasians, an 80% increased risk for carriers of the 280His allele was reported [28] but this increase was not confirmed by two other studies [20, 29]. For Asians, two small studies, one from India (123 cases/123 controls) and one from China (84 cases/252 controls), did not find an association between Arg280His and breast cancer risk [30, 31].

Studies on the role of *XRCC1* Arg399Gln in breast cancer risk have been summarized by meta-analyses [19, 20]. An effect of this polymorphism was found only when different ethnicities were analyzed separately. Summarizing 4 studies with Asian populations (1567 cases/1643 controls) [20] revealed a significant increase of risk among Asian carriers of the homozygous Gln variant of 60%

(OR = 1.6, 95% CI 1.1–2.4). In the Thai population, the Gln/Gln genotype was also more frequent in cases than in controls (*P* = 0.056). Although our data are not significant, they support that *XRCC1* Arg399Gln can modulate breast cancer risk in Asian women. This effect was still emphasized when we assessed the associations between the three polymorphisms determined in *XRCC1* and breast cancer risk based on diplotypes. Among the 10 diplotypes present in our population, we found the strongest effect, a 2.5-fold increase in risk, in individuals with the CGA/CGA diplotype. These individuals carry the wild type sequence in positions 194 and 280, but the variant Gln in position 399.

For *OGG1* Ser326Cys, we found a two-fold, significantly increased risk for breast cancer in postmenopausal women with the Cys/Cys genotype. This is in contrast to three studies among Caucasians [20, 32, 33] and two studies among Asians [34, 35] which did not find an association of this polymorphism with breast cancer risk in the overall population and in the subgroup analysis according to menopausal status. There is however strong evidence that risk factors differ in pre- and post-menopausal women and it is plausible that factors interacting with this polymorphism in determining breast cancer risk also have a different role in the subgroups [2]. The increased risk we observed for the *OGG1* Cys/Cys genotype is supported by detailed biochemical studies with the two variant enzymes. They found that the 326Cys-Ogg1 protein has a sevenfold lower activity for repairing 8-oxoguanine, altered substrate specificity and anomalous DNA binding conformation [36–38]. These functional alterations indicate a reduced repair activity for the variant allele and may contribute to the association of the polymorphism with breast cancer risk. Effects might become visible mainly in postmenopausal women who are older and lived with the *OGG1*-related repair insufficiency for a longer time. The differences to other studies may be related to the heterogeneous nature of exposure to breast carcinogens within study populations. As we have however no



clear data in our study that can be used to assess internal and external exposure to reactive oxygen species, our results need further validation.

To our knowledge, only one recent study examined the role of *APEX1* Asp148Glu and risk of breast cancer [20]. This study did not find an association with breast cancer risk among non-Hispanic white Americans. In contrast, a significant protective effect of *APEX1* 148Glu allele was observed in our Thai study which is the first report on Asian women. Little is known about how *APEX1* Asp148Glu affects DNA repair function. Characterization of the isolated variant proteins showed that *APEX1* Asp148Glu had no impact on endonuclease and DNA binding activities [39]. However, an increasing number of variant alleles of *APEX1* Asp148Glu and *XRCC1* Arg399Gln was significantly associated with prolonged cell cycle delay in G2 phase and decreased DNA repair capacity after  $\gamma$ -irradiation [14, 16]. Both conditions impede correct mitosis, and cells might thus favor cell death by apoptosis instead of proliferating with insufficiently repaired genome and potential accumulation of mutations. Therefore, the *APEX1* 148Glu allele might reduce breast cancer susceptibility as observed in our study. Nevertheless, our result needs further confirmation by additional case-control studies and functional investigations.

An individual's BER activity is characterized by more than one genetic variant. Although the functional impact of a single gene variant with low penetrance might be very limited, the interaction of several variant proteins with slightly reduced functional activity might sum up to a significant decrease in repair activity and an increased cancer risk. The analysis of the joint effects of the risk variants in the three BER genes *XRCC1*, *APEX1*, and *OGG1* revealed in fact a nearly two-fold increased risk for breast cancer in individuals homozygous for two or more variant alleles. It would therefore appear reasonable to hypothesize that those BER gene variants may be risk alleles for breast cancer and simultaneously contribute to higher cancer susceptibility. We are nevertheless aware that the polymorphisms studied here still do not give the complete information about variability in gene function. To study the relevance of BER on breast cancer risk in more detail, comprehensive analyses should include polymorphisms in all known BER genes.

Our risk estimates took into consideration several risk factors known from literature to affect the gene-risk association including age, reproductive parameters, hormonal use, and hazardous lifestyle [2]. Some of these risk factors were considerably different between cases and controls in our study (Table 1). This is in accordance with previous studies of breast cancer [2], with the exception of pregnancy and education: More cases had children than controls, but pregnancy is estimated as being protective,

and controls were more highly educated than cases. As both criteria were not used in patient recruitment as inclusion criteria, these differences might have happened by chance. They were, however, considered in our multivariate regression analysis as adjustment factors together with other covariates differing in our population or being especially important risk factors. One example is age that was different among cases and controls in our study. Therefore, we adjusted for this difference by modeling age as an exponential variable in the logistic regression model. The multivariate regression analysis for age only, for the complete model including all covariates, and the interaction analysis revealed that they had only a minor effect on the risk estimate for the polymorphisms. Thus, our results were robust to different model adjustments indicating that the risk contribution of genetic variants is moderate but not affected by the known strong breast cancer risk factors such as pregnancy, breast feeding and family history of breast cancer.

In conclusion, we have demonstrated that amino acid substitution variants of *OGG1*, *APEX1* and *XRCC1* genes, particularly in combination, are associated with increased susceptibility to breast cancer among Thai women. Although our study is larger in sample size than other studies among Asian populations, further epidemiological studies are warranted to confirm the role of genetic variation of BER in breast cancer susceptibility.

**Acknowledgments** The authors wish to thank S. Pummai and M. Rienkijarn for their technical assistance. We also acknowledge the contributions of all participants of the study. This work was supported by the National Center for Genetic Engineering and Biotechnology (Thailand) and by a Guest Scientist fellowship from the German Cancer Research Center (DKFZ) to S. Sangrajrang.

## References

1. Sriplung H, Sontipong S, Martin N et al (2005) Cancer incidence in Thailand, 1995–1997. *Asian Pac J Cancer Prev* 6:276–281
2. Dumitrescu RG, Cotarla I (2005) Understanding breast cancer risk – where do we stand in 2005? *J Cell Mol Med* 9:208–221
3. Bau DT, Mau YC, Ding SL et al (2007) DNA double-strand break repair capacity and risk of breast cancer. *Carcinogenesis* 28:1726–1730
4. Garcia-Closas M, Egan KM, Newcomb PA et al (2006) Polymorphisms in DNA double-strand break repair genes and risk of breast cancer: two population-based studies in USA and Poland, and meta-analyses. *Hum Genet* 119:376–388
5. Parshad R, Price FM, Bohr VA et al (1996) Deficient DNA repair capacity, a predisposing factor in breast cancer. *Br J Cancer* 74:1–5
6. Nock NL, Cicek MS, Li L et al (2006) Polymorphisms in estrogen bioactivation, detoxification and oxidative DNA base excision repair genes and prostate cancer risk. *Carcinogenesis* 27:1842–1848
7. Roy D, Cai Q, Felty Q et al (2007) Estrogen-induced generation of reactive oxygen and nitrogen species, gene damage, and

- estrogen-dependent cancers. *J Toxicol Environ Health B Crit Rev* 10:235–257
8. Mohrenweiser HW, Wilson DM III, Jones IM (2003) Challenges and complexities in estimating both the functional impact and the disease risk associated with the extensive genetic variation in human DNA repair genes. *Mutat Res* 526:93–125
  9. Memisoglu A, Samson L (2000) Base excision repair in yeast and mammals. *Mutat Res* 451:39–51
  10. Nilsen H, Krokan HE (2001) Base excision repair in a network of defence and tolerance. *Carcinogenesis* 22:987–998
  11. Duell EJ, Millikan RC, Pittman GS et al (2001) Polymorphisms in the DNA repair gene XRCC1 and breast cancer. *Cancer Epidemiol Biomarkers Prev* 10:217–222
  12. Lunn RM, Langlois RG, Hsieh LL et al (1999) XRCC1 polymorphisms: effects on aflatoxin B1-DNA adducts and glycoprotein A variant frequency. *Cancer Res* 59:2557–2561
  13. Matullo G, Guarrera S, Carturan S et al (2001) DNA repair gene polymorphisms, bulky DNA adducts in white blood cells and bladder cancer in a case-control study. *Int J Cancer* 92:562–567
  14. Vodicka P, Stetina R, Polakova V et al (2007) Association of DNA repair polymorphisms with DNA repair functional outcomes in healthy human subjects. *Carcinogenesis* 28:657–664
  15. Pachkowski BF, Winkel S, Kubota Y et al (2006) XRCC1 genotype and breast cancer: functional studies and epidemiologic data show interactions between XRCC1 codon 280 His and smoking. *Cancer Res* 66:2860–2868
  16. Hu JJ, Smith TR, Miller MS et al (2002) Genetic regulation of ionizing radiation sensitivity and breast cancer risk. *Environ Mol Mutagen* 39:208–215
  17. Wang Y, Spitz MR, Zhu Y et al (2003) From genotype to phenotype: correlating XRCC1 polymorphisms with mutagen sensitivity. *DNA Repair (Amst)* 2:901–908
  18. Breast Cancer Association Consortium (2006) Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. *J Natl Cancer Inst* 98:1382–1396
  19. Hung RJ, Hall J, Brennan P et al (2005) Genetic polymorphisms in the base excision repair pathway and cancer risk: a HuGE review. *Am J Epidemiol* 162:925–942
  20. Zhang Y, Newcomb PA, Egan KM et al (2006) Genetic polymorphisms in base-excision repair pathway genes and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 15:353–358
  21. Sangrajrang S, Schmezer P, Burkholder I et al (2007) The XRCC3 Thr241Met polymorphism and breast cancer risk: a case-control study in a Thai population. *Biomarkers* 12:523–532
  22. Popanda O, Schattenberg T, Phong CT et al (2004) Specific combinations of DNA repair gene variants and increased risk for non-small cell lung cancer. *Carcinogenesis* 25:2433–2441
  23. Breslow NE, Day NE (1980) Statistical methods in cancer research. Volume I – The analysis of case-control studies. *IARC Sci Publ* 32:5–338
  24. Sauerbrei W, Royston P (1999) Building multivariable prognostic and diagnostic models; transformation of the predictors by using fractional polynomials. *Journal of the Royal Statistical Society: Series A* 162:71–94
  25. Terwilliger J, Ott J (1994) Handbook of human genetic linkage. Johns Hopkins University Press, Baltimore
  26. Xie X, Ott J (1993) Testing linkage disequilibrium between a disease gene and marker loci. *Am J Hum Genet* 53(Suppl):313
  27. Stephens M, Donnelly P (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 73:1162–1169
  28. Moullan N, Cox DG, Angele S et al (2003) Polymorphisms in the DNA repair gene XRCC1, breast cancer risk, and response to radiotherapy. *Cancer Epidemiol Biomarkers Prev* 12:1168–1174
  29. Metsola K, Kataja V, Sillanpaa P et al (2005) XRCC1 and XPD genetic polymorphisms, smoking and breast cancer risk in a Finnish case-control study. *Breast Cancer Res* 7:R987–R997
  30. Chacko P, Rajan B, Joseph T et al (2005) Polymorphisms in DNA repair gene XRCC1 and increased genetic susceptibility to breast cancer. *Breast Cancer Res Treat* 89:15–21
  31. Jin MJ, Chen K, Zhang SS et al (2006) [Association of single nucleotide polymorphisms and haplotypes in DNA repair gene XRCC1 with susceptibility of breast cancer]. *Zhejiang Da Xue Xue Bao Yi Xue Ban* 35:370–376
  32. Rossner P Jr, Terry MB, Gammon MD et al (2006) OGG1 polymorphisms and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 15:811–815
  33. Vogel U, Nexo BA, Olsen A et al (2003) No association between OGG1 Ser326Cys polymorphism and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 12:170–171
  34. Cai Q, Shu XO, Wen W et al (2006) Functional Ser326Cys polymorphism in the hOGG1 gene is not associated with breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 15:403–404
  35. Choi JY, Hamajima N, Tajima K et al (2003) hOGG1 Ser326Cys polymorphism and breast cancer risk among Asian women. *Breast Cancer Res Treat* 79:59–62
  36. Hill JW, Evans MK (2006) Dimerization and opposite base-dependent catalytic impairment of polymorphic S326C OGG1 glycosylase. *Nucleic Acids Res* 34:1620–1632
  37. Kohno T, Shinmura K, Tosaka M et al (1998) Genetic polymorphisms and alternative splicing of the hOGG1 gene, that is involved in the repair of 8-hydroxyguanine in damaged DNA. *Oncogene* 16:3219–3225
  38. Weiss JM, Goode EL, Ladiges WC et al (2005) Polymorphic variation in hOGG1 and risk of cancer: a review of the functional and epidemiologic literature. *Mol Carcinog* 42:127–141
  39. Hadi MZ, Coleman MA, Fidelis K et al (2000) Functional characterization of Ape1 variants identified in the human population. *Nucleic Acids Res* 28:3871–3879