

## Polymorphisms of the DNA repair genes XPD (Lys751Gln) and XRCC1 (Arg399Gln and Arg194Trp): relationship to breast cancer risk and familial predisposition to breast cancer

A.M. Brewster<sup>1</sup>, T.J. Jorgensen<sup>2</sup>, I. Ruczinski<sup>3</sup>, H.Y. Huang<sup>4</sup>, S. Hoffman<sup>4</sup>, L. Thuita<sup>4</sup>, C. Newschaffer<sup>4</sup>, R.M. Lunn<sup>5</sup>, D. Bell<sup>6</sup>, and K.J. Helzlsouer<sup>4,7</sup>

<sup>1</sup>Department of Clinical Cancer Prevention, University of Texas M.D. Anderson Cancer Center, Houston, TX, USA; <sup>2</sup>Department of Radiation Medicine, Georgetown University Medical Center, Washington, DC, USA; <sup>3</sup>Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University School of Medicine, Baltimore, MD, USA; <sup>4</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; <sup>5</sup>Toxicology Operations Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA; <sup>6</sup>Environmental Genomics Section, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA; <sup>7</sup>The Prevention and Research Center, Mercy Medical Center, Baltimore, MD, USA

**Key words:** breast cancer risk, DNA repair genes, family history of breast cancer, imputation methods, nested case-control, XPD polymorphisms, XRCC1 polymorphisms

### Summary

Family history is a risk factor for breast cancer and could be due to shared environmental factors or polymorphisms of cancer susceptibility genes. Deficient function of DNA repair enzymes may partially explain familial risk as polymorphisms of DNA repair genes have been associated, although inconsistently, with breast cancer. This population based case-control study examined the association between polymorphisms in XPD (Lys751Gln) and XRCC1 (Arg399Gln and Arg194Trp) genes, and breast cancer. Breast cancer cases ( $n=321$ ) and controls ( $n=321$ ) were matched on age and menopausal status. Conditional logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI). The analysis was conducted omitting observations with missing data, and by using imputation methods to handle missing data. No significant association was observed between the XPD 751Gln/Lys (OR 1.37, 95% CI 0.96–1.96) and Gln/Gln genotypes (OR 1.08, 95% CI 0.62–1.86) (referent Lys/Lys), XRCC1 399Arg/Gln (OR 1.48, 95% CI 0.92–2.38) and Gln/Gln genotypes (1.11, 95% CI 0.67–1.83) (referent Arg/Arg) or the XRCC1 Arg/Trp and Trp/Trp genotypes (OR 1.12, 95% CI 0.69–1.83) (referent Arg/Arg) and breast cancer. In multivariate analysis, the adjusted odds ratios for the XPD and XRCC1 399 polymorphisms increased and became statistically significant, however, were attenuated when imputation methods were used to handle missing data. There was no interaction with family history. These results indicate that these polymorphisms in XPD and XRCC1 genes are only weakly associated with breast cancer. Without imputation methods for handling missing data, a statistically significant association was observed between the genotypes and breast cancer, illustrating the potential for bias in studies that inadequately handle missing data.

### Introduction

Family history of breast cancer is a particularly important risk factor for breast cancer, however, the inheritance of highly penetrance gene mutations accounts for only 5–9% of breast cancer cases [1]. The increased risk of breast cancer for the majority of women with a family history likely reflects shared minor low penetrant genetic factors and, to a lesser extent, shared environmental factors. Even among women without a family history,

twin studies suggest that inheritance of low penetrant genes, rather than environment, is the major breast cancer risk factor [2].

Single nucleotide polymorphisms (SNPs) of genes that have a significant effect on DNA repair capacity have been investigated as breast cancer susceptibility genes [3]. Of particular interest are polymorphisms of the xeroderma pigmentosum group D (XPD) and X-ray repair cross-complementing group 1 (XRCC1) genes, which are involved in the nucleotide excision (NER) and

the base excision (BER) DNA repair pathways, respectively. Despite evidence of functional significance of selected SNPs of the XPD [4,5] and XRCC1 DNA repair genes [6–10], data from hospital and population-based case–control studies have shown inconsistent findings of an association between the XPD Lys751Gln and XRCC1 Arg399Gln polymorphisms and risk of sporadic breast cancer [11–14].

Interestingly, a population-based case–control study from the Ontario Breast Cancer Family Registry found that while there was no evidence for a main effect of the XRCC1 399Gln variant allele on breast cancer risk, there was a statistically significant interaction observed with having a family history of breast cancer [14]. This finding suggests that SNPs of selected DNA repair genes may contribute to familial susceptibility to breast cancer or that studying families with a history of cancer increases the efficacy of identifying candidate-susceptibility genes [14].

A nested case–control study was conducted within an ongoing cohort study based in Washington County Maryland, CLUE II, to evaluate the association between the XPD Lys751Gln, XRCC1 Arg399Gln, and XRCC1 Arg194Trp genotypes and breast cancer risk. The effect modification of family history was examined among this population of women to assess the contribution of these DNA repair SNPs to familial breast cancer risk. To further explore the relationship between the DNA repair SNPs and breast cancer development, statistical methods to account for missing data were used in the analysis to both ensure validity of the findings and reduce any bias potentially introduced by dropping observations with missing data.

## Methods

A nested case–control study was conducted using the population based CLUE II cohort. The CLUE II cohort was established in 1989. Individuals residing in Washington County, Maryland, and surrounding regions were invited to donate blood for cancer and heart disease research (campaign slogan ‘Give us a clue to cancer and heart disease’). The CLUE II cohort consists of 32,892 individuals including approximately 30% of county residents [15]. Collected blood specimens were centrifuged, separated as buffy coat, red blood cells, and plasma, and then stored at  $-70^{\circ}\text{C}$ . Cancers that develop among cohort participants were ascertained through linkage to both the Washington County and, since 1992, the Maryland State Cancer Registries. In 1996, and about every two years afterwards, participants were asked to complete a follow-up questionnaire asking about health events, medication use, and cancer risk factors.

For this study, 321 incident cases of breast cancer that occurred after blood donation to CLUE II in 1989 were identified. Incident cases were defined as women with a first-time diagnosis of breast cancer (International

Classification of Disease-8 174 and ICD-9 174). Cases were excluded if they had a diagnosis of any other cancer except for non-melanoma skin cancer and cervical cancer *in situ*, or were under 18 at the time of blood donation. Three hundred and twenty one controls were individually matched to cases by gender (female sex), age at blood donation (within one year), and menopausal status at blood donation. Selected controls were cancer-free, except possibly for non-melanoma skin cancer and cervical cancer *in situ*, and were not known to be deceased up to the date of diagnosis of the case.

Information on breast cancer risk factors was obtained from several sources including a questionnaire that was sent in 1995 to a portion of the CLUE II breast cancer cases and controls who were part of a case–control study on organochlorine compounds and a 1996 follow-up questionnaire that was sent to all CLUE II participants. The overall response rate to the 1996 questionnaire was 79%, with an 81% response rate among CLUE II female participants. The 1995 questionnaire response rates were 89% for breast cancer cases and 79% for controls [15]. The questionnaires contained detailed information on family history of breast cancer, reproductive history, medication history and selective dietary intake. Information on breast cancer risk factors including age at menarche, age at first birth and family history of breast cancer among first degree relatives and grandparents was missing for approximately 16% of cases and 20% of controls. There was missing genotype information for XPD Lys751Gln (12 cases, 3 controls) XRCC1 Arg399Gln (16 cases, 11 controls) and XRCC1 Arg194Trp (12 cases, 4 controls). The study was approved by the Committee of Human Research of the Johns Hopkins Bloomberg School of Public Health.

## Genotyping

Genomic DNA was extracted from peripheral buffy coat using the alkaline lysis method [16]. Extracted DNA samples were resuspended in 10 mM Tris–HCl/1 mM EDTA (TE) and the DNA concentration quantified by absorbance at 260 nm ( $A_{260}$ ). DNA concentration was set at 100  $\mu\text{g}/\text{ml}$ . The XPD Lys751Gln, XRCC1 Arg399Gln and XRCC1 Arg194Trp genotypes were assessed using the patented fluorogenic method for nucleic acid analysis commonly known as the Taqman<sup>®</sup> or 5' nuclease assay (Applied Biosystems Division, Perkin-Elmer, Foster City, CA). For 115 of the cases and 115 controls, the XPD codon 751 and XRCC1 codons 399 and 194 genotyping was also conducted using a polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) technique with primers as previously described [4]. The PCR program consisted of a 4 min denaturation step at  $94^{\circ}\text{C}$  followed by 30 cycles of 30 s at  $94^{\circ}\text{C}$ , 30 s at  $60^{\circ}\text{C}$ , and 60 s at  $72^{\circ}\text{C}$ . The PCR products were digested for 2 h at  $37^{\circ}\text{C}$  with the MboII restriction enzyme. The digestive products were separated using 2% agarose gel electrophoresis. The

concordance was 94% for the study samples genotyped with both the Taqman and PCR techniques. When there was a discrepancy with the Taqman assay, the PCR genotype results were used for the analysis.

### Statistical analysis

Differences in the distribution of breast cancer risk factors between breast cancer cases and controls were compared using chi square tests. These included: education (years, <12, ≥12), history of first or second degree relative with breast cancer (no, yes), age at menarche (years, <12, 12–13, >13), age at first birth (years, <20, 20–24, 25–29 and ≥30), history of oral contraceptive use (never, former, current), history of other hormone use (never, current estrogen, current progesterone, former estrogen or progesterone) and smoking exposure (never, former and current).

Hardy Weinberg equilibrium of the observed frequency of the XPD and XRCC1 genotypes among controls was tested with the goodness of fit chi-square.

Conditional logistic regression was used to calculate the odds ratio (OR) and 95% confidence interval (CI) for the association between the XPD Lys751Gln, XRCC1 Arg399Gln and XRCC1 Arg194Trp genotypes, and breast cancer. The breast cancer risk factors considered as possible confounders were age at menarche, age at first birth, family history of breast cancer, hormone use, duration of breastfeeding, and smoking history. With the exception of family history, none of these risk factors was statistically significantly associated with breast cancer and none changed the parameter estimates by greater than 10% when entered in the multivariate conditional logistic model. Therefore, only family history of breast cancer was adjusted for in the multivariate conditional logistic model.

Because of the missing family history of breast cancer and genotype data, we approached the conditional logistic analysis in two different ways. First, the analysis was carried out omitting pairs of observations that had missing data. Second, the data analysis was carried out using complete data sets generated by multiple imputation [17,18]. We generated probability distributions for the missing values, given the observed data, using decision trees. Independent samples from those probability distributions were used to generate ten replicates of imputed complete data sets. Parameter estimates and standard errors were then combined from the analyses of the individual complete data sets, taking the variability between the parameter estimates from the ten data sets as well as the standard error in each parameter estimate into account [17,18]. Analyses were performed using Stata Statistical Software version 8.0 (Stata Corporation, College Station, TX, 2004), and R version 2.01 (The R Project for Statistical Computing, <http://www.r-project.org/>).

The association between the XPD and XRCC1 genotypes and breast cancer development was stratified by family history of breast cancer to examine family

history of breast cancer as a potential effect modifier. To allow for stratification by family history, the case-control sets were unmatched and analyzed using unconditional logistic regression with imputation methods for missing data. A conditional regression analysis was also performed with an interaction term (genotype × family history) included in the model.

### Results

Cases and controls were nearly identical on age at blood donation and menopausal status in 1989, suggesting that the individual matching was successful (Table 1). The majority of cases in the study were postmenopausal. Cases were slightly more likely than controls to be nulliparous, to have a history of current or former oral contraception use and current or former estrogen and/or progesterone use, but none of those differences was statistically significant. However, cases were much more likely than controls to have a family history of breast cancer ( $p = <0.001$ ). In univariate analysis, family history was associated with a statistically significant increased risk of breast cancer of approximately 2.4-fold. Imputation of missing family history data reduced the width of the confidence interval but had little effect on the odds ratio estimate (OR 2.43, 95% CI 1.54–3.84 versus OR 2.32, 95% CI 1.35–3.97).

Table 2 shows the association between the XPD Lys751Gln, and XRCC1 Arg399Gln and XRCC1 Arg194Trp genotypes and breast cancer. In the univariate conditional logistic model there was no statistically significant association between the XPD and XRCC1 genotypes and risk of breast cancer. The univariate odds ratios with imputation of missing genotype data as shown in Table 2 were not substantially different from the odds ratios obtained when case-control pairs with missing genotype data were excluded from the analysis (data not shown). Interestingly, the multivariate model excluding missing observations (Model 1) resulted in statistically significant estimates for genotypes Lys751Gln and Gln751Gln of 1.90 and 2.18, respectively (Lys751Lys referent genotype). In contrast, with imputation of missing data (Model 2), the odds ratios were attenuated with narrower confidence intervals that were no longer statistically significant, for both XPD genotypes. Although not as substantial, the findings for the two XRCC1 codons were similar. In each case, without imputation of missing values into the multivariate conditional logistic model, the effect was an increase in the point estimates for the genotypes. Using imputed data kept the odds ratio point estimates lower, and narrowed the confidence intervals.

The distribution of genotypes in cases/controls matched pairs with some missing data (i.e. the observations that would normally be dropped from conditional logistic regression analysis), and the distribution between matched pairs with no missing data (i.e. observations retained in analysis) is shown in Table 3.

*Table 1.* Distribution of selected characteristics and risk factors in breast cancer cases and matched population controls

	% Cases	% Controls	<i>p</i> -value
<i>Age at enrollment (years)</i>			NA
≤ 40	10.9	10.0	
41–50	23.1	24.3	
51–60	23.1	21.5	
≥61	43.0	44.2	
<i>Menopausal status at enrollment</i>			NA
Post	70.7	70.4	
Pre	24.6	26.2	
Peri	1.6	2.5	
Missing	3.1	0.9	
<i>Education (years)</i>			0.21
< 12	24.3	29.0	
≥12	75.7	71.0	
<i>Family history of breast cancer</i>			0.0002
Yes	20.6	9.0	
No	62.6	70.7	
Missing	16.8	20.2	
<i>Age at menarche (years)</i>			0.17
< 12	12.8	15.0	
12–13	47.7	38.9	
≥13	18.7	21.8	
Missing	20.9	24.3	
<i>Age at first birth (years)</i>			0.37
< 20	18.1	17.4	
20–24	31.5	31.5	
25–29	14.6	13.7	
≥30	3.4	4.7	
Nulliparous	12.1	7.8	
Missing	20.2	24.9	
<i>Oral contraception use at baseline</i>			0.15
Never	73.5	75.4	
Former	24.9	20.9	
Current	1.6	2.8	
Missing	0.0	0.9	
<i>Other hormone use at baseline</i>			0.19
Never	78.5	77.9	
Former	9.0	11.5	
Current estrogen	6.9	6.2	
Current estrogen/progesterone	3.1	1.9	
Current progesterone	1.2	0.0	
Missing	1.2	2.5	
<i>Cigarette smoking</i>			0.25
Never	62.3	64.5	
Former	25.2	20.2	
Current	12.5	15.3	

The XPD 751Gln and XRCC1 194Trp distribution in the cases/controls with missing data were different from the distribution in the cases/controls with non-missing data. Multiple imputation allowed the use of data from the case/control pairs with missing data, and this resulted in a lower estimate and narrower confidence interval for the odds ratio estimates of the association between the genotypes and breast cancer.

The association between XPD Lys751Gln, and XRCC1 Arg399Gln and XRCC1 Arg194Trp genotypes and breast cancer stratified by family history of breast cancer is shown in Table 4. In the group of women with a positive family history of breast cancer, the XRCC1 399Gln and XRCC1 194 Trp alleles were associated with a nonstatistically significant decrease in the risk of breast cancer. The *p*-values for interaction between XPD 751 and XRCC1 399 and 194 genotypes and family history in the conditional regression model were not statistically significant (data not shown).

## Discussion

In this study, we evaluated the role of three DNA repair gene SNPs in breast cancer within a well-defined large cohort of mostly Caucasian women whom have been followed for several decades in Washington County, Maryland, and for whom both genetic and family history data were available. The design was a nested case-control study, matched on age and menopausal status – two strong determinants of risk. We found that family history of breast cancer was strongly and significantly associated with breast cancer risk in this population. In the univariate analysis, there was no significant breast cancer risk associated with the XPD 751Gln and XRCC1 399Gln and XRCC1 194Trp variant alleles and no evidence of an effect modification for family history. In the multivariate model adjusted for family history with missing observations excluded, there was a statistically significant association observed between the XPD 751Gln and XRCC1 399Gln alleles and breast cancer. However, the associations were attenuated and became nonstatistically significant after imputation methods were used to handle missing genotype and family history data. This occurred because variation in genotype distributions between case-control pairs with and without missing family history data was not adequately accounted for when paired observations with missing family history data were dropped from analysis. In contrast, imputation of family history data allowed full incorporation of all genotype data, by restoring all observations with genotype information back into the analysis. This prevented the introduction of genotype information bias, maintained the validity of the findings, and presumably produced more accurate and reliable measures of association.

An earlier population-based case-control study from the Ontario Breast Cancer Family Registry of one of the SNPs studied here (XRCC1 Arg399Gln), found an association among women with a family history of breast cancer, suggesting that the XRCC1 gene contributed to the familial breast cancer in that population. Our study, in a similar population of predominantly Caucasian women, does not support this conclusion. The Ontario Breast Cancer Family Registry study, however, was limited to women under the age of 55 at time of diagnosis and included breast cancer cases whom

Table 2. Breast cancer risk associated with XPD codon 751 and XRCC1 codons 399 and 194 genotypes using conditional logistic regression models either without (Model 1) or with (Model 2) the imputation of missing data

Genotypes	Univariate OR (95% CI)	Multivariate adjusted <sup>a</sup> OR (95% CI) Model 1	Multivariate adjusted <sup>a</sup> OR (95% CI) Model 2
<b>XPD 751</b>			
No. Of case/control pairs	321	202	321
Lys/Lys	1.00 (Referent)	1.00 (Referent)	1.00 (Referent)
Lys/Gln	1.37 (0.96–1.96)	1.90 (1.20–3.00)	1.45 (1.00–2.10)
Gln/Gln	1.08 (0.62–1.86)	2.18 (1.08–4.40)	1.31 (0.74–2.34)
<b>XRCC1 399</b>			
No. Of case/control pairs	321	199	321
Arg/Arg	1.00 (Referent)	1.00 (Referent)	1.00 (Referent)
Arg/Gln	1.48 (0.92–2.38)	1.91 (1.05–3.46)	1.64 (1.00–2.69)
Gln/Gln	1.11 (0.67–1.83)	1.33 (0.69–2.57)	1.22 (0.73–2.06)
<b>XRCC1 194</b>			
No. Of case/control pairs	321	204	321
Arg/Arg	1.00 (Referent)	1.00 (Referent)	1.00 (Referent)
Arg/Trp & Trp/Trp	1.12 (0.69–1.83)	1.40 (0.70–2.80)	1.20 (0.72–2.00)

<sup>a</sup>Conditional logistic model adjusted for family history of breast cancer.

OR = odds ratio.

CI = confidence interval.

Table 3. Distribution of XPD codon 751 and XRCC1 codons 399 and 194 genotypes among cases and controls in matched pairs with missing family history data, versus cases and controls in matched pairs where family history data is complete

Genotypes	Missing family history data (n = 107)		Without missing family history data (n = 214)	
	Case	Control	Case	Control
<b>XPD 751</b>				
Lys/Lys	5	12	25	22
Lys/Gln	54	57	121	102
Gln/Gln	43	35	61	90
No data	5	3	7	0
<b>XRCC1 399</b>				
Arg/Arg	41	42	67	84
Arg/Gln	49	45	110	90
Gln/Gln	12	13	26	36
No data	5	7	11	4
<b>XRCC1 194</b>				
Arg/Arg	90	89	179	190
Arg/Trp	13	14	24	22
Trp/Trp	1	1	2	1
No data	3	3	9	1

were selected for having a family history of breast cancer [14]. Genetic susceptibility to reduced DNA capacity may be more apparent in younger breast cancer cases who have also been shown to have a lower capacity to repair UV damage compared to age matched controls [19]. Smith et al. evaluated the frequency of the XRCC1 194Trp allele among breast cancer cases and controls selected for having a family history of breast cancer and found a similar genotype distribution [20]. In this study of women unselected for family history of breast cancer, the distribution of the XPD and XRCC1 SNPs were similar among breast cancer cases and controls

indicating no association with breast cancer among women in the general population.

Case-control studies of candidate genes are frequently used to identify cancer-susceptibility alleles. Since some highly penetrant breast cancer genes (BRCA1, BRCA2, and ATM) have known DNA repair functions, and because DNA repair genes have been implicated in other germline cancer genes (e.g. mismatch repair genes in familial colon cancer, and nucleotide excision repair genes in familial skin cancer), it has been suspected that polymorphisms in other DNA repair pathways may represent low penetrance risk alleles.

Table 4. Effect of XPD codon 751 and XRCC1 codons 399 and 194 genotypes and family history on breast cancer risk

Family History	Genotypes	Cases (n)	Controls (n)	Multivariate adjusted <sup>a</sup> OR (95% CI)
<b>XPD751</b>				
Negative	Lys/Lys	81	108	1.00 (referent)
	Lys/Gln	138	143	1.06 (0.96–1.17)
	Gln/Gln	23	33	0.99 (0.85–1.16)
Positive	Lyn/Lyn	26	18	1.00 (referent)
	Lyn/Gln	44	17	1.13 (0.93–1.37)
	Gln/Gln	8	2	1.24 (0.86–1.80)
<b>XRCC1399</b>				
Negative	Arg/Arg	26	45	1.00 (referent)
	Arg/Gln	129	123	1.15 (1.00–1.32)
	Gln/Gln	88	116	1.06 (0.92–1.22)
Positive	Arg/Arg	14	6	1.00 (referent)
	Arg/Gln	38	17	0.97 (0.74–1.27)
	Gln/Gln	27	14	0.94 (0.71–1.24)
<b>XRCC1194</b>				
Negative	Arg/Arg	206	250	1.00 (referent)
	Arg/Trp & Trp/Trp	37	34	1.08 (0.95–1.22)
Positive	Arg/Arg	73	32	1.00 (referent)
	Arg/Trp & Trp/Trp	6	5	0.84(0.59–1.19)

<sup>a</sup>Unconditional logistic model adjusted for age and menopausal status with imputation of missing data.

OR = odds ratio.

CI = confidence interval.

Also, there have been numerous epidemiological studies that report associations of DNA repair gene variants with common cancers [3]. One cited problem with the epidemiological evidence has been that most reported associations come from small populations that fail to be confirmed by subsequent studies [21]. It is feasible that some of the conflicting results in epidemiological studies investigating the association between low penetrance genes and breast cancer risk may be due to inadequate handling of missing genotype and epidemiological data, such as seen in this study where dropping pairs with missing genotype and family history leads to a different conclusion than dealing with the missing data using multiple imputation.

Epidemiological data from well-designed population based case-control studies that indicate no relationship between polymorphisms of the XPD and XRCC1 genes and breast cancer risk could potentially conflict with the findings of functional studies. Within the BER pathway, selected SNPs of the *XRCC1* gene have been linked to measures of suboptimal DNA repair capacity. For example, the variant Gln allele of the XRCC1 Arg399Gln polymorphism is associated with increased DNA adduct levels [7,8], increased sister chromatid exchanges [9,10], and hypersensitivity to ionizing radiation [22]. In a prior study of 31 cancer free women, the number of chromatid aberrations present in lymphocytes after exposure to ionizing radiation was used to measure DNA repair proficiency. DNA repair proficiency was higher in women homozygous for the common XPD codon 751 Lys allele compared to those with the Gln allele [4]. When stratified by familial breast cancer risk status, the difference was significant in

women with a family history of breast cancer but not in those without a family history of breast cancer. Other functional studies measuring the influence of the XPD Lys751Gln polymorphism on DNA repair proficiency in cancer-free patients and patients with lung cancer have shown decreased DNA repair associated with the wild type Gln allele [5,23,24]. However, detailed mechanistic studies comparing function of the polymorphic alleles in XRCC1 and XPD have not been published. The small sample size, different assays used to measure DNA repair proficiency, and varying populations of cancer and cancer-free patients included in functional studies, further complicates the interpretation of their results [25].

The apparent discrepancy between negative population genotype studies and positive functional studies may be related to the heterogeneous nature of exposure to breast cancer carcinogens within study populations. DNA lesions and their respective repair pathways are carcinogen-dependent. For example, the NER pathway primarily repairs bulky adducts in DNA, such as those produced by ultraviolet radiation or cigarette smoke. In fact, individuals with germline mutations in NER genes are typically prone to skin cancers, eye cancers, tongue cancers, and probably smoking-related lung cancers [26], as would be expected based on the type of DNA damage caused by these carcinogens and the tissues exposed. Environmental factors such as ionizing radiation exposure have been shown to be necessary for full phenotypic expression of the XRCC1 Arg399Gln polymorphism on breast cancer susceptibility among Caucasian women [10]. BER repairs damage produced through free radical generating carcinogens, such as ionizing radiation and inducers of oxidative stress. Since

the nature of the carcinogens contributing to breast cancer risk within this populations are unknown and likely to be heterogeneous, the gene-environmental interactions between DNA repair genes and breast cancer carcinogens within the population cannot be directly assessed. It is likely that the influence of DNA repair genes in breast cancer risk would only be seen when combined with well-defined and discrete exposures to DNA damaging agents, since the effect of the specific DNA repair gene may only be important among the subpopulation with the relevant DNA damaging agent exposure.

The results of this study support the predominance of data from population-based case-control studies that indicate a weak or no role for these polymorphisms as breast cancer susceptibility genes among Caucasian women. These findings may not be applicable, however, to subpopulations of women with elevated exposures to DNA damaging agents, where significant gene-environmental interactions may occur. Further population-based case-control studies that have the advantage of reduced bias in case ascertainment and that use sophisticated statistical methods for handling missing breast cancer risk factor and genotype data are needed to evaluate the association between polymorphisms of XPD and XRCC1 and breast cancer among other ethnic populations of women.

### Acknowledgements

Grant Support: The National Cancer Institute Breast Spore P50CA88843, and the National Institute of Aging 5U01AG018033. A.M.B was supported by Minority Supplement of Clinical Oncology Research Career Development Program (K12-CA01709) from the National Cancer Institute. T.J.J. was funded by a Ruth L. Kirschstein Senior Fellow Award (NCI F33 CA09817-01). I.R. was supported, in part, by the Maryland Cigarette Restitution Fund Research Grant to the Johns Hopkins Medical Institutions and NIH Grant CA 074841.

### References

1. Ford D, Easton DF: The genetics of breast and ovarian cancer. *Br J Cancer* 72: 805-812, 1995
2. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K: Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 343(2): 78-85, 2000
3. Goode EL, Ulrich CM, Potter JD: Polymorphisms in DNA repair genes and associations with cancer risk. *Cancer Epidemiol Biomarkers Prev* 11(12): 1513-1530, 2002
4. Lunn RM, Helzlsouer KJ, Parshad R, Umbach DM, Harris EL, Sanford KK, Bell DA: XPD polymorphisms: effects on DNA repair proficiency. *Carcinogenesis* 21(4): 551-555, 2000
5. Spitz MR, Wu X, Wang Y, Wang LE, Shete S, Amos CI, Guo Z, Lei L, Mohrenweiser H, Wei Q: Modulation of nucleotide excision repair capacity by XPD polymorphisms in lung cancer patients. *Cancer Res* 61(4): 1354-1357, 2001
6. Hu JJ, Smith TR, Miller MS, Mohrenweiser HW, Golden A, Case LD: Amino acid substitution variants of APE1 and XRCC1 genes associated with ionizing radiation sensitivity. *Carcinogenesis* 22(6): 917-922, 2001
7. Lunn RM, Langlois RG, Hsieh LL, Thompson CL, Bell DA: XRCC1 polymorphisms: effects on aflatoxin B1-DNA adducts and glyco-phorin A variant frequency. *Cancer Res* 59(11): 2557-2561, 1999
8. Matullo G, Guarrera S, Carturan S, Peluso M, Malaveille C, Davico L, Piazza A, Vineis P: DNA repair gene polymorphisms, bulky DNA adducts in white blood cells and bladder cancer in a case-control study. *Int J Cancer* 92(4): 562-567, 2001
9. Abdel-Rahman SZ, El-Zein RA: The 399Gln polymorphism in the DNA repair gene XRCC1 modulates the genotoxic response induced in human lymphocytes by the tobacco-specific nitrosamine NNK. *Cancer Lett* 159(1): 63-71, 2000
10. Duell EJ, Wiencke JK, Cheng TJ, Varkonyi A, Zuo ZF, Ashok TD, Mark EJ, Wain JC, Christiani DC, Kelsey KT: Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers of DNA damage in human blood mononuclear cells. *Carcinogenesis* 21(5): 965-971, 2000
11. Duell EJ, Millikan RC, Pittman GS, Winkel S, Lunn RM, Tse CK, Eaton A, Mohrenweiser HW, Newman B, Bell DA: Polymorphisms in the DNA repair gene XRCC1 and breast cancer. *Cancer Epidemiol Biomarkers Prev* 10(3): 217-222, 2001
12. Shu XO, Cai Q, Gao YT, Wen W, Jin F, Zheng W: A population-based case-control study of the Arg399Gln polymorphism in DNA repair gene XRCC1 and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 12(12): 1462-1467, 2003
13. Justenhoven C, Hamann U, Pesch B, Harth V, Rabstein S, Baisch C, Vollmert C, Illig T, Ko YD, Brüning T, et al. ERCC2 genotypes and a corresponding haplotype are linked with breast cancer risk in a German population. *Cancer Epidemiol Biomarkers Prev* 13(12): 2059-2064, 2004
14. Figueiredo JC, Knight JA, Briollais L, Andrulis IL, Ozcelik H: Polymorphisms XRCC1-R399Q and XRCC3-T241M and the risk of breast cancer at the Ontario site of the Breast Cancer Family Registry. *Cancer Epidemiol Biomarkers Prev* 13(4): 583-591, 2004
15. Helzlsouer KJ, Alberg AJ, Huang HY, Hoffman SC, Strickland PT, Brock JW, Burse VW, Needham LL, Bell DA, Lavigne JA, et al. Serum concentrations of organochlorine compounds and the subsequent development of breast cancer. *Cancer Epidemiol Biomarkers Prev* 8(6): 525-532, 1999
16. Klientschar M, Neuhuber F: Evaluation of an alkaline lysis method for the extraction of DNA from whole blood and forensic stains for STR analysis. *J Forensic Sci* 45(3): 669-673, 2000
17. Rubin DB: Multiple imputation after 18+ years. *J Am Statistical Assoc* 91(434): 473-489, 1996
18. Schafer JL: *Analysis of Incomplete Multivariate Data*, vol. 72. Chapman & Hall, New York, 1997
19. Ramos JM, Ruiz A, Colen R, Lopez ID, Grossman L, Matta JL: DNA repair and breast carcinoma susceptibility in women. *Cancer* 100(7): 1352-1357, 2004
20. Smith TR, Miller MS, Lohman K, Lange EM, Case LD, Mohrenweiser HW, Hu JJ: Polymorphisms of XRCC1 and XRCC3 genes and susceptibility to breast cancer. *Cancer Lett* 190(2): 183-190, 2003
21. Pharoah PD, Dunning AM, Ponder BA, Easton DF: Association studies for finding cancer-susceptibility genetic variants. *Nat Rev Cancer* 4(11): 850-860, 2004
22. Angelini S, Kumar R, Carbone F, Maffei F, Forti GC, Violante FS, Lodi V, Curti S, Hemminki K, Hrelia P: Micronuclei in humans induced by exposure to low level of ionizing radiation: influence of polymorphisms in DNA repair genes. *Mutat Res* 570(1): 105-117, 2005
23. Shi Q, Wang LE, Bondy ML, Brewster A, Singletary SE, Wei Q: Reduced DNA repair of benzo[a]pyrene diol epoxide-induced adducts and common XPD polymorphisms in breast cancer patients. *Carcinogenesis* 25(9): 1695-1700, 2004

24. Qiao Y, Spitz MR, Shen H, Guo Z, Shete S, Hedayati M, Grossman L, Mohrenweiser H, Wei Q: Modulation of repair of ultraviolet damage in the host-cell reactivation assay by polymorphic XPC and XPD/ERCC2 genotypes. *Carcinogenesis* 23(2): 295–299, 2002
25. Benhamou S, Sarasin A: ERCC2/XPD gene polymorphisms and cancer risk. *Mutagenesis* 17(6): 463–469, 2002
26. Kraemer KH, Lee M-M, Andrews AD, Lambert WC: The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer: the xeroderma pigmentosum paradigm. *ArchDermatol* 130: 1018–1021, 1994

*Address for offprints and correspondence:* A.M. Brewster, Department of Clinical Cancer Prevention, Unit 1360, University of Texas, M.D. Anderson Cancer Center, P.O. Box 301439, Houston, TX, 77230-1439, USA; *Tel.:* +713-745-4928; *Fax:* +713-563-5746; *E-mail:* abrewster@mdanderson.org