EUROPEAN JOURNAL OF CANCER 44 (2008) 710–717

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Xeroderma pigmentosum complementation group C genotypes/diplotypes play no independent or interaction role with polycyclic aromatic hydrocarbons-DNA adducts for breast cancer risk

Jing Shen[a,*], *Marilie D. Gammon*[b], *Mary Beth Terry*[c], *Susan L. Teitelbaum*[d], *Sybil M. Eng*[e], *Alfred I. Neugut*[c], *Regina M. Santella*[a]

[a]*Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, 630 West 168th Street, P&S 19-418, New York, NY 10032, USA*
[b]*Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, NC 27599, USA*
[c]*Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032, USA*
[d]*Department of Community and Preventive Medicine, Mt. Sinai School of Medicine, New York, NY 10029, USA*
[e]*Global Epidemiology, Safety and Risk Management, Pfizer Inc., New York, NY 10017, USA*

ARTICLE INFO

Article history:
Received 11 July 2007
Received in revised form
4 October 2007
Accepted 29 October 2007
Available online 28 November 2007

Keywords:
XPC
PAH-DNA adducts
Interaction
Breast cancer
Risk

ABSTRACT

Xeroderma pigmentosum complementation group C (XPC) is an important DNA nuclear excision repair (NER) gene that recognises the damage caused by a variety of bulky DNA adducts. We evaluated the association of two common non-synonymous polymorphisms in *XPC* (Ala499Val and Lys939Gln) with breast cancer risk in the Long Island Breast Cancer Study Project (LIBCSP), a population-based case–control study. Genotyping of 1067 cases and 1110 controls was performed by a high throughput assay with fluorescence polarisation. There were no overall associations between *XPC* polymorphisms and breast cancer risk. A diplotype CC–CC was significantly associated with increased breast cancer risk compared with diplotype CA–CA (OR = 1.4, 95%CI: 1.0–1.9), but was not significant when compared with all other diplotypes combined (OR = 1.22, 95%CI: 0.97–1.53). No modification effects were observed for *XPC* genotypes by cigarette smoking status, smoking pack-years or polycyclic aromatic hydrocarbons (PAH)-DNA adducts. The increase in breast cancer risk was slightly more pronounced among women with detectable PAH-DNA adducts and carrying the diplotype CC–CC (OR = 1.6, 95%CI: 1.1–2.2) compared to women with non-detectable PAH-DNA adducts carrying other diplotypes combined, but no statistically significant interaction was observed ($P_{interaction}$ = 0.69). These data suggest that *XPC*s have neither independent effects nor interactions with cigarette smoking and PAH-DNA adducts for breast cancer risk. Further studies with multiple genetic polymorphisms in NER pathway are warranted.

* *Corresponding author:* Tel.: +1 212 305 8158; fax: +1 212 305 5328.
E-mail address: js2182@columbia.edu (J. Shen).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2007.10.027

1. Introduction

DNA repair plays an important role in eliminating DNA damage and maintaining genetic integrity. Xeroderma pigmentosum complementation group C (XPC) is an important DNA damage recognition protein involved in global genome DNA repair (GGR), a subclass of nucleotide excision repair (NER). The XPC protein recognises a variety of bulky DNA adducts formed by exogenous carcinogens including those found in tobacco smoke (benzo[*a*]pyrene (B[*a*]P) and other polycyclic aromatic hydrocarbons (PAHs)), and endogenous carcinogens, all of which may play a role in the pathogenesis of breast cancer.[1–5] It binds tightly with another protein, HR23B, to form a stable XPC-HR23B complex. Recent studies identified the XPC-HR23B complex as the first protein component that recognises and binds to the damaged sites. Interestingly, defects in XPC have been associated with many types of cancer.[6–8] Transgenic mice studies also revealed predisposition to many types of cancer in XPC gene knockout mice.[9,10] These results suggest that the XPC protein plays an important role in the prevention of the DNA damage-mediated cancer occurrence. Recently, polymorphisms in *XPC* have been studied for associations with squamous cell carcinoma of the head and neck (SCCHN), lung, bladder, colorectal and breast cancer.[11,12] The heterozygous *XPC* genotype, Lys939Gln, but not the homozygous Gln939Gln genotype, was significantly associated with breast cancer susceptibility in Finnish[11] and Chinese populations.[13] But in Caucasians, African Americans or mixed population, no similar results were observed.[12,14] Overall, the relationship between *XPC* polymorphisms and breast cancer risk remains unclear.

Exposure to cigarette smoke and breast cancer risk has been investigated in numerous epidemiological studies with controversial outcomes. A recently published meta-analysis found that cigarette smoking had little or no independent effect on the risk of developing breast cancer.[15] This is consistent with findings in the Long Island Breast Cancer Study Project (LIBCSP).[16] However, several studies have suggested that cigarette smoking may increase breast cancer risk among women who have smoked for a long time or before their first full-term pregnancy.[17,18] An increasing number of studies are also indicating that a positive association between smoking and breast cancer risk may be stronger among or limited to women with certain genotypes.[19] PAH have been shown to induce mammary tumours in laboratory-based studies.[20] The major exposure sources include cigarette smoke, consumption of PAH-containing foods (such as grilled and smoked foods) and air pollutants (from industrial manufacturing, fires and motor vehicle traffic). PAH-DNA adducts have been associated with a modest increase in breast cancer risk in several epidemiologic studies,[21–25] although no apparent dose–response and substantial variations associated with sources of PAH, such as cigarette smoking or consumption of grilled and smoked foods, were observed.[21,22] A previous study indicated that workers with lower DNA repair capacity (carrying XPC-PAT+/+ genotype) had significantly increased anti-B[*a*]P diol epoxide-DNA adduct levels, which was significantly related with PAH exposure.[26] *XPC* polymorphisms (Arg492His, Ala499Val and Lys939Gln) were shown to modify smoking-related risk of advanced colorectal adenoma, but not bladder or breast cancer.[14] Whether *XPC* polymorphisms modify the risk of breast cancer by interaction with exogenous carcinogens (cigarette smoking and PAH-DNA adducts) is unknown.

In the present study, we describe the distribution of genotype for two non-conservative variants of *XPC* at codons 499 (Ala499Val) and 939 (Lys939Gln) in a large population-based case–control study, the LIBCSP, with the aim of evaluating associations between *XPC* genotypes and haplotypes/diplotypes and breast cancer risk, and the modifying effects of cigarette smoking and PAH-DNA adducts.

2. Materials and methods

2.1. Study design

The study methods of the LIBCSP had been described in detail previously.[27] In brief, case eligibility included adult female residents of Nassau and Suffolk counties on Long Island, NY, who were of any age or race, who spoke English, and were newly diagnosed with in situ or invasive breast cancer between 1st August 1996, and 31st July 1997. Potentially eligible controls were frequency matched to the age distribution of the cases and identified through random digit dialing for women under age 65 years, and through the Center for Medicare and Medicaid Services (CMS) rosters for women 65 and over. Eligible controls were defined as women who spoke English, and resided in the same Long Island counties as the cases, but with no personal history of breast cancer. This study was conducted with approval from participating institutional review boards, and in accordance with an assurance filed with and approved by the United States Department of Health and Human Services. In-person interviews were completed for 82.1% of cases ($n = 1508$) and 62.8% of controls ($n = 1556$). Of those who completed the main questionnaire, 73.1% of cases (1102) and 73.3% of controls (1141) donated a blood sample, and 1067 cases and 1110 controls with enough DNA for genotyping were analysed in the present study. The constituent of ethnicity and age distribution were similar between respondents who completed the main questionnaire and the subjects recruited in present analyses (data not shown).

2.2. Laboratory methods

Genomic DNA was extracted by standard RNase/proteinase K and phenol/chloroform treatment and genotyped by a fluorescence polarization (FP) method using a commercial AcycloPrime™-FP SNP Detection Kit obtained from Perkin-Elmer Life Sciences. This technique distinguishes the polymorphic base of a SNP by the template-directed incorporation of a dye-labelled dideoxynucleotide onto an oligonucleotide primer that anneals just 5′ to the polymorphic base.[28] The forward and reverse primers, designed according to the human *XPC* gene sequence (GeneBank accession number AH009651), were, respectively, 5′-GCC TCT GAT CCC TCT GAT GA-3′, 5′-CAT CGC TGC ACA TTT TCT TG-3′ for Ala499Val (NCBI rs2228000) and 5′-GCC TCA AAA CCG AGA AGA TG-3′, 5′-CTG CCT CAG TTT GCC TTC TC-3′ for Lys939Gln (NCBI rs2228001). The sequences of TDI probes were designed using Primer 3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/

primer3www.cgi) as forward 5′-GTA AGG ACC CAA GCT TGC CAG-3′ for Ala499Val and reverse 5′-GGG CGC TCA GCT CAC AGC T-3′ for Lys939Gln. The two allele-specific dye terminators used for each of the two SNPs are C/T for Ala499Val and G/T for Lys939Gln. The assay was validated by sequencing subjects with the three genotypes and these known samples were used as positive controls on each plate. Finally, the fluorescence polarization was measured with a Perkin Elmer Victor instrument. All the assays were performed with laboratory personnel blinded to the subject's disease status. In addition to assay-specific quality control samples, 10% of samples were re-assayed after re-labelling to keep laboratory staff blinded to its identity. Concordance for re-assayed samples was 100%. Genotyping data were available for 99.5% of breast cancer cases and 98.9% of controls with enough DNA.

The PAH-DNA adducts levels, determined by competitive ELISA, were available from the parent study.[21,22] Adduct levels were determined for 873 (79.2%) breast cancer cases and 941 (82.5%) controls who donated sufficient blood for the assay, which were 57.9% (873/1508) and 60.5% (941/1556) of the eligible cases and controls, respectively.[22] For the analyses presented here, PAH-DNA adduct levels were dichotomised into a single binary variable: detectable (>15% inhibition in the ELISA) versus non-detectable adducts. Full technical details are available on request.

2.3. *Questionnaire data*

The main questionnaire assessed the information on known and suspected risk factors for breast cancer, including PAH exposure sources over the life course as cigarette smoking. A current cigarette smoker was defined as smoking within the last 12 months prior to the reference date; a former smoker was defined as a smoker who reported quitting more than 12 months prior to the reference date.[16] Light and moderate or heavy smokers were categorised by using the 50th percentile pack-year [(cigarettes per day/20) × (years smoked)] values of the controls as the cutoff points (i.e. <15, and ⩾15 pack-years). Other factors considered as potential confounders (see below) were also assessed as part of the in-person interview.

2.4. *Statistical methods*

Hardy–Weinberg equilibrium (HWE)[29] was tested to compare the observed and expected genotype frequencies among cases and controls, respectively. χ^2 tests for categorical variables were used to assess case–control differences.[30] Haplotype and diplotype frequencies were estimated from genotype data by PHASE version 2.1.1 based on a Bayesian algorithm and the maximum likelihood model linked with the Stochastic-EM algorithm.[31,32] Haplotypes and diplotypes were selected according to the corresponding occurring probabilities with a higher likelihood (>0.95 as cut-point).[31–33] The most common haplotype CA was selected as the reference in the analysis. The risk of breast cancer was estimated for each diplotype compared to the reference (CA–CA) or all other diplotypes combined, adjusted for age at reference. Diplotypes were treated as categorical variables and were incorporated as dummy variables in the logistic regression models.

Unconditional logistic regression with SAS version 9.0 was used to estimate odds ratios (ORs) and corresponding 95% confidence intervals (CI).[34] The potential confounding variables were assessed by examining the percent changes (more than 10%) in the ORs of the main effects of genotyping. Variables found not to confound the associations of interest included age at menarche, parity, lactation, months of lactation, age at first birth, number of miscarriages, history of fertility problems, alcohol drinking, race, education, religion and marital status (as previously defined).[21,22,27] All models were adjusted for the frequency matching factor, age at reference (defined as age at diagnosis for cases and age at identification for controls). Effect modification was assessed on a multiplicative scale by calculating ORs for *XPC* genotypes stratified by indicators of cigarette smoking and PAH-DNA adducts status as previously defined,[21,22,27] running separate models for each exposure category and by including multiplicative interaction terms in the logistic regression model. The *XPC* genotypes were dichotomised by assuming a dominant genetic model based on previous studies of cancer risk.[35–37] All statistical tests were two-sided with a significant level of 0.05.

3. Results

The frequencies of the *XPC* polymorphisms (Ala499Val and Lys939Gln) did not deviate significantly from the Hardy–Weinberg equilibrium (*P* values were 0.45 and 0.18, respectively). For the codon 499 locus, the frequencies of the CC, CT and TT genotypes were 57.2%, 37.7% and 5.1%, respectively, among the control subjects (Table 1). The corresponding frequencies among the cases were 57.9%, 36.3% and 5.8%, respectively. The frequencies of the AA, AC, and CC genotypes for the codon 939 locus were 33.2%, 51.2% and 15.6% among controls, and 32.5%, 49.1% and 18.5% among cases. Overall, the distribution of *XPC* genotypes did not vary significantly between cases and controls when separately analysing the two variant alleles, either among pre- or post-menopausal women (data not shown).

Haplotypes CC and TA did not have any relationships with breast cancer risk compared with the reference haplotype, CA (Table 1). Common diplotypes (CC–TA, CC–CA and CA–TA) were not associated with statistically significant increased breast cancer risk, except diplotype (CC–CC) that displayed a significantly increased breast cancer risk, when compared with the diplotype CA–CA (OR = 1.4, 95%CI: 1.0–1.9), but not when compared with all other diplotypes combined as the reference (OR = 1.22, 95%CI: 0.97–1.53, *P* = 0.09).

Overall, there were no modification effects observed between *XPC* genotypes and breast cancer risk by cigarette smoking status, neither in former nor in current smokers (Table 2). In the multivariate analysis with adjustment for age at reference, we observed increased ORs in subgroups with <15 or ⩾15 pack-years smoking and carrying either *XPC* 939 or *XPC* 499 variant alleles compared with never smokers carrying wild-type genotype, but there was no significant heterogeneity across these cells. No significant associations were observed between *XPC* genotypes and breast cancer risk by passive smoking status (data not shown).

Table 1 – Estimated odds ratio (ORs) and 95% confidence intervals (CIs) for the association between XPC polymorphisms and breast cancer risk, Long Island Breast Cancer Study Project, 1996–1997

XPC genotype or haplotype/diplotype	Cases, *n* (%)	Controls, *n* (%)	OR (95%CI)[a]
Ala499Val genotype			
CC	614 (57.9)	632 (57.2)	1.0 (reference)
CT	385 (36.3)	417 (37.7)	0.9 (0.8–1.1)
TT	62 (5.8)	56 (5.1)	1.1 (0.8–1.6)
CT/TT	447 (42.1)	473 (42.8)	1.0 (0.8–1.1)
Lys939Gln genotype			
AA	344 (32.5)	366 (33.2)	1.0 (reference)
AC	520 (49.1)	565 (51.2)	1.0 (0.8–1.2)
CC	196 (18.5)	172 (15.6)	1.2 (0.9–1.5)
AC/CC	716 (67.6)	737 (66.8)	1.0 (0.9–1.2)
Haplotypes[b]			
CA	706 (33.3)[c]	769 (34.9)[c]	1.0 (reference)
CC	905 (42.7)	908 (41.2)	1.1 (0.9–1.2)
TA	502 (23.7)	528 (23.9)	1.0 (0.9–1.2)
Diplotypes (haplotype pairs)			
CA–CA	117 (11.0)	142 (12.9)	1.0 (reference)
CC–TA	215 (20.3)	249 (22.6)	1.0 (0.8–1.4)
CC–CA	300 (28.3)	316 (28.6)	1.2 (0.9–1.6)
CC–CC	196 (18.5)	172 (15.6)	1.4 (1.0–1.9)[d]
CA–TA	170 (16.0)	168 (15.2)	1.2 (0.9–1.7)
Other haplotype pairs	62 (5.9)	56 (5.1)	1.3 (0.9–2.1)
Total	1060 (100.0)	1103 (100.0)	

a Adjusted for age at reference.
b XPC haplotypes composed of two polymorphic sites: Ex9-377C > T and Ex16 + 211A > C.
c No. of chromosomes.
d OR = 1.22 (95%CI: 0.97–1.53) when all other haplotype pairs combined as reference, $P = 0.09$.

Statistically significant increased breast cancer risks were found among women with detectable PAH-DNA adducts, in carriers of the *XPC* 499CC genotype and *XPC* 939AC/CC or AA genotypes (Table 2). No significant multiplicative interactions were observed between any *XPC* variant allele and detectable PAH-DNA adducts. There was no evidence of a dose–response relationship between *XPC* genotypes and higher level of PAH-DNA adducts for breast cancer risk, although some subgroups showed significant differences (Table 2).

Based on the potential relationship between the diplotype CC–CC and breast cancer risk observed above, we further analysed the potential effect modifications of cigarette smoking and PAH-DNA adducts (Table 3). Overall, we found no statistically significant modifications for *XPC* CC–CC diplotype either by cigarette smoking status or smoking pack-years. However, a statistically significant increased breast cancer risk was observed in the subgroup carrying the CC–CC diplotype and exposed to moderate or heavy smoking (OR = 3.7, 95%CI: 1.8–7.6) compared to never smokers carrying other diplotypes combined, but with a small sample size. The increase in breast cancer risk was slightly more pronounced among women with detectable PAH-DNA adducts and carrying diplotype CC–CC (OR = 1.6, 95%CI: 1.1–2.2) compared to women with non-detectable PAH-DNA adducts carrying all other diplotypes combined, but no statistically significant interaction was observed ($P_{interaction} = 0.69$). There was no evidence of a modifying effect for the CC–CC diplotype and breast cancer risk by a higher level of PAH-DNA adducts.

4. Discussion

Although some significant results were observed in several subgroups, the present study does not support the hypothesis that genotypes or haplotype/diplotype of *XPC* Ala499Val and Lys939Gln have an independent aetiological role in breast cancer nor do they interact with cigarette smoking or PHA-DNA adducts. This is consistent with a recent epidemiological study indicating no relationship between *XPC* polymorphisms and breast cancer risk in Caucasians, African Americans or mixed population,[12,14] but is in conflict with reports in Finnish[11] and Chinese[13] populations. Of course, the different ethnic groups included in various studies may partly account for the conflicting results on whether or not *XPC* polymorphisms have independent effects and interaction with cigarette smoking for breast cancer risk, but other explanations should also be considered.

Methodological issues of inadequate study design should be considered, such as non-random sampling, limited sample size, lack of statistical power and precision in subgroups.[19] Most previous studies with positive findings may have been underpowered, which may contribute to spurious findings. According to one estimate, approximately 68% of stratum-specific analyses had less than 80% power to detect an OR of 2.0 between genotypes and breast cancer risk modified by cigarette smoking status.[19] Thus, earlier findings of a positive association between *XPC* polymorphisms and cancer risk in select subgroups with small sample sizes are likely false

Table 2 – The association between XPC genotypes and breast cancer risk by cigarette smoking, PAH-DNA adducts status, Long Island Breast Cancer Study Project, 1996–1997

Exposure status	XPC Ala499Val				XPCLys939Gln			
	CC		CT/TT		AA		AC/CC	
	Case–control	OR (95%CI)[a]	Case–control	OR (95%CI)[a]	Case–control	OR (95%CI)[a]	Case–control	OR (95%CI)[a]
Never smokers	295/293	1.0 (reference)	191/206	0.9 (0.7–1.3)	150/169	1.0 (reference)	336/329	1.2 (0.9–1.5)
Former smokers	201/226	0.9 (0.7–1.1)	170/170	1.0 (0.8–1.3)	116/133	1.0 (0.7–1.3)	254/263	1.1 (0.8–1.4)
Current smokers	118/111	1.2 (0.8–1.6)	86/97	0.9 (0.7–1.3)	78/63	1.5 (1.0–2.3)	126/144	1.1 (0.8–1.5)
Never smokers	295/293	1.0 (reference)	191/206	0.9 (0.7–1.2)	150/169	1.0 (reference)	336/329	1.2 (0.9–1.5)
< 15 pack-years smokers[d]	77/68	1.2 (0.8–1.7)	64/39	1.7 (1.1–2.6)	43/29	1.7 (1.0–2.9)	98/78	1.5 (1.0–2.1)
⩾15 pack-years smokers[d]	104/56	1.9 (1.3–2.7)	68/54	1.2 (0.8–1.8)	59/37	1.9 (1.2–3.0)	113/73	1.7 (1.2–2.5)
Non-detectable PAH-DNA adducts	121/160	1.0 (reference)	103/132	1.0 (0.7–1.5)	72/107	1.0 (reference)	151/183	1.2 (0.8–1.8)
Detectable PAH-DNA adducts	382/379	1.3 (1.0–1.8)	257/267	1.3 (0.9–1.7)[b]	202/203	1.5 (1.0–2.1)	437/443	1.5 (1.1–2.0)[c]
Non-detectable PAH-DNA adducts	121/160	1.0 (reference)	103/132	1.0 (0.7–1.5)	72/107	1.0 (reference)	151/183	1.2 (0.8–1.8)
Detectable (below median)	185/201	1.2 (0.9–1.7)	144/121	1.6 (1.1–2.2)	109/98	1.7 (1.1–2.5)	220/224	1.5 (1.0–2.1)
Detectable (median and above)	197/178	1.5 (1.1–2.0)	113/146	1.0 (0.7–1.4)	93/105	1.3 (0.9–2.0)	217/219	1.5 (1.0–2.1)

a Adjusted for age at reference.

b $P_{interaction} = 0.71$.

c $P_{interaction} = 0.33$.

d Pack-years less than that of former and current smokers are due to missing data for cigarettes per day or smoked years.

Table 3 – The association between XPC diplotypes and breast cancer risk by cigarette smoking, PAH-DNA adducts status, Long Island Breast Cancer Study Project, 1996–1997

Exposure status	XPC diplotypes			
	All other diplotypes combined		CC–CC diplotype	
	Case–control	OR (95%CI)[a]	Case–control	OR (95%CI)[a]
Never smokers	383/418	1.0 (reference)	103/80	1.4 (1.0–1.9)
Former smokers	307/333	1.0 (0.8–1.2)	63/63	1.1 (0.7–1.6)
Current smokers	174/179	1.1 (0.9–1.5)	30/28	1.2 (0.7–2.1)
Never smokers	383/419	1.0 (reference)	103/80	1.4 (1.0–1.9)
< 15 pack-years smokers[b]	121/91	1.5 (1.1–2.0)	20/16	1.4 (0.7–2.7)
⩾15 pack-years smokers[b]	138/100	1.5 (1.1–2.0)	34/10	3.7 (1.8–7.6)
PAH-DNA adducts				
Non-detectable	185/244	1.0 (reference)	38/46	1.1 (0.7–1.8)
Detectable	521/548	1.3 (1.0–1.6)	118/98	1.6 (1.1–2.2)[c]
Non-detectable	185/244	1.0 (reference)	38/46	1.1 (0.7–1.8)
Detectable (below median)	262/269	1.3 (1.0–1.7)	67/53	1.7 (1.1–2.5)
Detectable (median and above)	259/279	1.2 (0.9–1.6)	51/45	1.5 (0.9–2.3)

a Adjusted for age at reference.
b Pack-years less than that of former and current smokers are due to missing data for cigarettes per day or smoked years.
c $P_{interaction} = 0.69$.

positives due to chance. In the present study, we estimated the power of detecting an OR of 1.5 at a significance level of 0.05 based on our sample size and the minor allele frequency. The power for *XPC* codon 499 and 939 variant alleles were, respectively, 97% and 99%. The lack of an association between *XPC* polymorphisms and breast cancer risk in the current study is thus not due to insufficient statistical power.

Another explanation for the inconsistent results across studies may be related to the opposing effects of smoking on breast cancer risk.[18,38] Oestrogen is a known risk factor for breast cancer. The carcinogenic effects of DNA-damaging agents such as B[a]P and other aromatic hydrocarbons in cigarette smoke may have an anti-oestrogenic effect by acting as non-steroidal oestrogens[39] or activate the oestrogen receptor α.[40] A recent study[41] that found a significant inverse trend for cumulative pack-years of cigarette smoking and breast cancer risk among parous women supports our observations. In this recent report, cigarette smoking was found to have a harmful effect among women who began smoking before the birth of their first child, but among women who began smoking after the first live birth, the risk for breast cancer was reduced. These results suggest that the risk of breast cancer from cigarette smoking may vary with the timing of the exposure over the life course.[41] Although these observations have not been consistently reported across investigations,[16,42–44] it is possible that the lack of consideration of a possible interaction with oestrogen exposure, or for the effect of cigarette smoking during different periods, may mask or dilute real associations.

A third possible explanation is that most previous studies did not consider the complex gene–environment interactions. As part of the parent LIBCSP study,[21,22] participants were characterised in terms of PAH-DNA adducts, enhancing the precision of defining individual PAH body burden. Most previous studies lacked an exposure biomarker, possibly biasing (overestimating) the independent or interaction effects of *XPC* polymorphisms and cigarette smoking for breast cancer risk due to the inaccurate categorisation of cigarette smoke exposure.

Finally, only two polymorphisms in *XPC* gene involved in the recognition of DNA damage were explored in the present study. There were more than thirty other genes involved in the NER pathway were not considered. Crew and colleagues observed that a combined effect of polymorphisms in NER genes revealed a modest, statistically significant positive association between number of putative high-risk alleles and breast cancer risk in LIBCSP.[45] A previous study conducted in the same population observed that a polymorphism in *XPD* codon 751, also functioning in the NER pathway, statistically significantly increased breast cancer risk from PAH-DNA adducts and cigarette smoking exposure.[46] Moreover, *XPC* has a much broader role in affecting other DNA repair pathways.[47] But we did not observe any gene–gene interactions between the two *XPC* polymorphisms and *XPD* codon 751 in the present study (data not shown). These results suggest that the two *XPC* polymorphisms may not independently play a critical role in breast tumourigenesis, or at least, have a less important role than *XPD* polymorphisms. Even if the *XPC* polymorphisms affect breast cancer risk, it may not be through interaction with exposure to cigarette smoking or PAH-DNA adducts. Although increased breast cancer risk was observed in subgroups carrying the CC–CC diplotype and exposed to moderate/heavy smoking or detectable PAH-DNA adducts, the lack of *a priori* functional data as well as the limitation of haplotype/diplotype imputation limits the interpretation of the data.[48]

The relatively large sample size, population-based study design, and both genotyping and exposure biomarker (PAH-DNA adducts) available, are the strengths of this study. Recall bias is one limitation that is common to case–control studies; but because neither cases nor controls were aware of their genotype and the exposure biomarker at the time of the interview, the accuracy of the response is not likely to be related to genotype or exposure status. Although the

exposure information on cigarette smoking may be differentially recalled by cases and controls due to the cases' cancer diagnosis, this misclassification would not differ by genotype and PAH-DNA adducts status. Although our results were obtained from a mixed population, most of them (>93%) were Caucasians. The potential influence of ethnic heterogeneity might not bias the conclusion. Extensive information collected in the study for evaluating confounding factors and effect modifiers allows us to control for confounding in data analysis. The OR for *XPC* polymorphism and breast cancer risk changed less than 10% when considering potential confounding factors individually or in multivariate models indicating that confounding did not dramatically affect the results. Age at reference included in logistic models, minimised any possible residual confounding effect of age.

In conclusion, we did not find any independent main effects of *XPC* polymorphisms for breast cancer risk or interaction with cigarette smoking/PAH-DNA adducts, although several significant differences were observed for subgroups with *XPC* variant alleles or diplotypes and specific carcinogenic exposures. There is evidence suggesting that the two *XPC* polymorphisms are linked with a functional SNP in 3′ of intron 11 splice acceptor site and may decrease DNA repair activity of single strand breaks.[49,50] However, the biological effects of the two SNPs under conditions of carcinogenic exposure remain to be determined.

Conflict of interest statement

None declared.

Acknowledgements

This work is funded by U01 CA/ES66572, P30ES09089 and P30ES10126 Grants from the National Cancer Institute and the National Institute of Environmental Health Sciences, and an award from the Breast Cancer Research Foundation and gifts from private citizens.

REFERENCES

1. Wang A, Gu J, Judson-Kremer K, et al. Response of human mammary epithelial cells to DNA damage induced by BPDE: involvement of novel regulatory pathways. *Carcinogenesis* 2003;**24**:225–34.
2. Hartman AR, Ford JM. BRCA1 and p53: compensatory roles in DNA repair. *J Mol Med* 2003;**81**:700–7.
3. Emmert S, Kobayashi N, Khan SG, Kraemer KH. The xeroderma pigmentosum group C gene leads to selective repair of cyclobutane pyrimidine dimers rather than 6-4 photoproducts. *Proc Natl Acad Sci USA* 2000;**97**:2151–6.
4. Janicijevic A, Sugasawa K, Shimizu Y, et al. DNA bending by the human damage recognition complex XPC-HR23B. *DNA Repair (Amst)* 2003;**2**:325–36.
5. Chen Z, Xu XS, Yang J, Wang G. Defining the function of XPC protein in psoralen and cisplatin-mediated DNA repair and mutagenesis. *Carcinogenesis* 2003;**24**:1111–21.
6. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. *Cancer Epidemiol Biomark Prev* 2002;**11**:1513–30.
7. Miyashita H, Mori S, Tanda N, et al. Loss of heterozygosity of nucleotide excision repair factors in sporadic oral squamous cell carcinoma using microdissected tissue. *Oncol Rep* 2001;**8**:1133–8.
8. Takebayashi Y, Nakayama K, Kanzaki A, et al. Loss of heterozygosity of nucleotide excision repair factors in sporadic ovarian, colon and lung carcinomas: implication for their roles of carcinogenesis in human solid tumors. *Cancer Lett* 2001;**174**:115–25.
9. Cheo DL, Burns DK, Meira LB, Houle JF, Friedberg EC. Mutational inactivation of the xeroderma pigmentosum group C gene confers predisposition to 2-acetylaminofluorene-induced liver and lung cancer and to spontaneous testicular cancer in Trp53–/– mice. *Cancer Res* 1999;**59**:771–5.
10. Friedberg EC, Cheo DL, Meira LB, Reis AM. Cancer predisposition in mutant mice defective in the XPC DNA repair gene. *Prog Exp Tumor Res* 1999;**35**:37–52.
11. Forsti A, Angelini S, Festa F, et al. Single nucleotide polymorphisms in breast cancer. *Oncol Rep* 2004;**11**:917–22.
12. Jorgensen TJ, Visvanathan K, Ruczinski I, et al. Breast cancer risk is not associated with polymorphic forms of xeroderma pigmentosum genes in a cohort of women from Washington County, Maryland. *Breast Cancer Res Treat* 2007;**101**:65–71.
13. Zhang L, Zhang Z, Yan W. Single nucleotide polymorphisms for DNA repair genes in breast cancer patients. *Clin Chim Acta* 2005;**359**:150–5.
14. Mechanic LE, Millikan RC, Player J, et al. Polymorphisms in nucleotide excision repair genes, smoking and breast cancer in African Americans and whites: a population-based case–control study. *Carcinogenesis* 2006;**27**:1377–85.
15. Hamajima N, Hirose K, Tajima K, et al. Alcohol, tobacco and breast cancer–collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *Br J Cancer* 2002;**87**:1234–45.
16. Gammon MD, Eng SM, Teitelbaum SL, et al. Environmental tobacco smoke and breast cancer incidence. *Environ Res* 2004;**96**:176–85.
17. Band PR, Le ND, Fang R, Deschamps M. Carcinogenic and endocrine disrupting effects of cigarette smoke and risk of breast cancer. *Lancet* 2002;**360**:1044–9.
18. Terry PD, Rohan TE. Cigarette smoking and the risk of breast cancer in women: a review of the literature. *Cancer Epidemiol Biomark Prev* 2002;**11**:953–71.
19. Terry PD, Goodman M. Is the association between cigarette smoking and breast cancer modified by genotype? A review of epidemiologic studies and meta-analysis. *Cancer Epidemiol Biomark Prev* 2006;**15**:602–11.
20. el Bayoumy K, Chae YH, Upadhyaya P, et al. Comparative tumorigenicity of benzo[*a*]pyrene, 1-nitropyrene and 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine administered by gavage to female CD rats. *Carcinogenesis* 1995;**16**:431–4.
21. Gammon MD, Santella RM, Neugut AI, et al. Environmental toxins and breast cancer on Long Island. I. Polycyclic aromatic hydrocarbon DNA adducts. *Cancer Epidemiol Biomark Prev* 2002;**11**:677–85.
22. Gammon MD, Sagiv SK, Eng SM, et al. Polycylic aromatic hydrocarbon (PAH)-DNA adducts and breast cancer: a pooled analysis. *Arch Envir Health* 2004;**59**:640–9.
23. Li D, Wang M, Dhingra K, Hittelman WN. Aromatic DNA adducts in adjacent tissues of breast cancer patients: clues to breast cancer etiology. *Cancer Res* 1996;**56**:287–93.

24. Li D, Walcott FL, Chang P, et al. Genetic and environmental determinants on tissue response to in vitro carcinogen exposure and risk of breast cancer. *Cancer Res* 2002;**62**:4566–70.
25. Rundle A, Tang D, Hibshoosh H, et al. The relationship between genetic damage from polycyclic aromatic hydrocarbons in breast tissue and breast cancer. *Carcinogenesis* 2000;**21**:1281–9.
26. Pavanello S, Pulliero A, Siwinska E, Mielzynska D, Clonfero E. Reduced nucleotide excision repair and GSTM1-null genotypes influence anti-B[a]PDE-DNA adduct levels in mononuclear white blood cells of highly PAH-exposed coke oven workers. *Carcinogenesis* 2005;**26**:169–75.
27. Gammon MD, Neugut AI, Santella RM, et al. The Long Island Breast Cancer Study Project: description of a multi-institutional collaboration to identify environmental risk factors for breast cancer. *Breast Cancer Res Treat* 2002;**74**:235–54.
28. Chen X, Levine L, Kwok PY. Fluorescence polarization in homogeneous nucleic acid analysis. *Genome Res* 1999;**9**:492–8.
29. Wellek S. Tests for establishing compatibility of an observed genotype distribution with Hardy–Weinberg equilibrium in the case of a biallelic locus. *Biometrics* 2004;**60**:694–703.
30. Hosmer DW, Lemenshow S. *Applied logistic regression*. New York: John Wiley & Sons; 1989.
31. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001;**68**:978–89.
32. Tregouet DA, Escolano S, Tiret L, Mallet A, Golmard JL. A new algorithm for haplotype-based association analysis: the Stochastic-EM algorithm. *Ann Hum Genet* 2004;**68**:165–77.
33. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 2003;**73**:1162–9.
34. Rothman KJ, Greenland S. *Modern epidemiology*. New York: Lippcott-Raven; 1998. pp. 329–42.
35. Garcia-Closas M, Malats N, Real FX, et al. Genetic variation in the nucleotide excision repair pathway and bladder cancer risk. *Cancer Epidemiol Biomark Prev* 2006;**15**:536–42.
36. Huang WY, Berndt SI, Kang D, et al. Nucleotide excision repair gene polymorphisms and risk of advanced colorectal adenoma: XPC polymorphisms modify smoking-related risk. *Cancer Epidemiol Biomark Prev* 2006;**15**:306–11.
37. Sak SC, Barrett JH, Paul AB, Bishop DT, Kiltie AE. Comprehensive analysis of 22 XPC polymorphisms and bladder cancer risk. *Cancer Epidemiol Biomark Prev* 2006;**15**:2537–41.
38. Palmer JR, Rosenberg L. Cigarette smoking and the risk of breast cancer. *Epidemiol Rev* 1993;**15**:145–56.
39. Johnson MD, Kenney N, Stoica A, et al. Cadmium mimics the in vivo effects of estrogen in the uterus and mammary gland. *Nat Med* 2003;**9**:1081–4.
40. Meek MD, Finch GL. Diluted mainstream cigarette smoke condensates activate estrogen receptor and aryl hydrocarbon receptor-mediated gene transcription. *Environ Res* 1999;**80**:9–17.
41. Ha M, Mabuchi K, Sigurdson AJ, et al. Smoking cigarettes before first childbirth and risk of breast cancer. *Am J Epidemiol* 2007;**166**:55–61.
42. Gammon MD, Schoenberg JB, Teitelbaum SL, et al. Cigarette smoking and breast cancer risk among young women. *Cancer Cause Control* 1998;**9**:583–90.
43. Lash TL, Aschengrau A. A null association between active or passive cigarette smoking and breast cancer risk. *Breast Cancer Res Treat* 2002;**75**:181–4.
44. Lash TL, Aschengrau A. Active and passive cigarette smoking and the occurrence of breast cancer. *Am J Epidemiol* 1999;**149**:5–12.
45. Crew KD, Gammon MD, Terry MB, et al. Polymorphisms in nucleotide excision repair genes, polycyclic aromatic hydrocarbon-DNA adducts, and breast cancer risk. *Cancer Epidemiol Biomark Prev16* 2007;**16**:2033–41.
46. Terry MB, Gammon MD, Zhang FF, et al. Polymorphism in the DNA repair gene XPD, polycyclic aromatic hydrocarbon-DNA adducts, cigarette smoking and breast cancer risk. *Cancer Epidemiol Biomark Prev* 2004;**13**:2053–8.
47. Wang G, Dombkowski A, Chuang L, Xu XX. The involvement of XPC protein in the cisplatin DNA damaging treatment-mediated cellular response. *Cell Res* 2004;**14**:303–14.
48. Cheng R, Ma JZ, Elston RC, Li MD. Fine mapping functional sites or regions from case–control data using haplotypes of multiple linked SNPs. *Ann Hum Genet* 2005;**69**:102–12.
49. Khan SG, Muniz-Medina V, Shahlavi T, et al. The human XPC DNA repair gene: arrangement, splice site information content and influence of a single nucleotide polymorphism in a splice acceptor site on alternative splicing and function. *Nucleic Acids Res* 2002;**30**:3624–31.
50. Vodicka P, Kumar R, Stetina R, et al. Genetic polymorphisms in DNA repair genes and possible links with DNA repair rates, chromosomal aberrations and single-strand breaks in DNA. *Carcinogenesis* 2004;**25**:757–63.