European Journal of Clinical Nutrition (2011) 65, 683–689

0954-3007/11
www.nature.com/ejcn

ORIGINAL ARTICLE

Single nucleotide polymorphisms in uracil-processing genes, intake of one-carbon nutrients and breast cancer risk

C Marian[1], M Tao[2], JB Mason[3], DS Goerlitz[1], J Nie[2], A Chanson[3], JL Freudenheim[2] and PG Shields[1]

[1]*Carcinogenesis, Biomarkers and Epidemiology Program, Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC, USA;* [2]*Department of Social and Preventive Medicine, School of Public Health and Health Professions, State University of New York at Buffalo, Buffalo, NY, USA and* [3]*Vitamins & Carcinogenesis Laboratory, Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, Boston, MA, USA*

Background/Objectives: The misincorporation of uracil into DNA leads to genomic instability. In a previous study, some of us identified four common single nucleotide polymorphisms (SNPs) in uracil-processing genes (rs2029166 and rs7296239 in *SMUG1*, rs34259 in *UNG* and rs4775748 in *DUT*) that were associated with significantly altered levels of uracil in human DNA. We investigated whether any of these SNPs are associated with an altered risk of developing breast cancer and if one-carbon nutrients intake can modify their effects.
Subjects/Methods: We genotyped the four SNPs in 1077 cases of incident breast cancer and 1910 age and race-matched controls in the Western New York Exposures and Breast Cancer (WEB) Study and examined associations with breast cancer risk and interactions with intake of folate, vitamins B6 and B12.
Results: After adjustment for known risk factors for breast cancer, there was increased risk of breast cancer among postmenopausal women who were heterozygous for either of the two *SMUG1* SNPs (odds ratio (OR) 1.29, 95% confidence interval (CI) 1.07–1.56) and OR 1.29, 95% CI 1.07–1.55, respectively). Among premenopausal women, increased risk associated
with the *SMUG1* rs2029166 genotype was limited to those with low folate intake. There were no other interactions with vitamins B_6 or B_{12} intake.
Conclusions: Our study suggests that the four selected SNPs are not robust determinants of breast cancer risk, but that the two SNPs in *SMUG1* might modestly alter the risk of breast cancer. However, the increase in risk among heterozygotes in the two SNPs in *SMUG1*, which is thought to be the most active glycosylase *in vivo*, raises the possibility that subtle 'heterosis' effects on cancer risk might be produced by these SNPs.
European Journal of Clinical Nutrition (2011) **65,** 683–689; doi:10.1038/ejcn.2011.29; published online 23 March 2011

Keywords: uracil-processing genes; single nucleotide polymorphisms; breast cancer

Introduction

Uracil misincorporation into DNA occurs spontaneously at low levels as a result of cytosine deamination or misincorporation of dUMP during DNA replication (Kavli *et al*., 2007; Sousa *et al*., 2007). Under normal conditions, such lesions are rapidly repaired by the base excision repair mechanism initiated by uracil-DNA glycosylase enzymes (Krokan *et al*., 2002; Visnes *et al*., 2009). In humans, four uracil-DNA glycosylases have been identified, encoded by the *UNG*, *SMUG1*, *MBD4* and *TDG* genes (Mohrenweiser and Jones, 1998; Barnes and Lindahl, 2004; Hung *et al*., 2005). In addition, mammalian cells have enzymes (dUTPases) that prevent the incorporation of uracil into DNA by de-phosphorylating dUTP into dUMP. In humans the only known UTPase is encoded by the *DUT* gene (Barnes and Lindahl, 2004; Sousa *et al*., 2007).

Intact mechanisms for the removal of uracil from DNA and prevention of its incorporation are essential for maintaining DNA integrity. It has been shown that failure to remove misincorporated uracil can lead to mutations after

Correspondence: Dr C Marian, Carcinogenesis, Biomarkers and Epidemiology Program, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, 3800 Reservoir Road, NW LL(s) Level, Room 116A, Washington, DC 20057-1465, USA.
E-mail: cm444@georgetown.edu
Received 14 July 2010; revised 13 December 2010; accepted 7 February 2011; published online 23 March 2011

DNA replication or to double-strand DNA breaks leading to chromosomal aberrations, both hallmarks of the cancer phenotype (Fenech, 2001; Kapiszewska *et al*., 2005; Sousa *et al*., 2007; Mashiyama *et al*., 2008).

Thus, a major factor influencing DNA integrity is the deoxyribonucleotide pool homeostasis, an equilibrium that is directly dependent on the adequate availability of certain elements of the B-vitamin family of micronutrients. Folate deficiency leads to low 5,10-methylene-THF concentrations, the essential co-factor for the conversion of dUMP to dTMP and may thereby increase the level of uracil incorporation into DNA (Blount *et al*., 1997). The chromosomal fragility and breaks that accompany folate depletion under some circumstances are thought to be caused by increased uracil misincorporation and attempted repair followed by the development of double-strand breaks (Ames, 1999; Jang *et al*., 2005). Similarly, vitamins B_{12} and B_6 appear to be crucial factors in maintaining genomic stability of human cells as well, as each have integral roles in recycling folate into the co-enzymatic form used for thymidine synthesis. (Fenech, 2001; Ames and Wakimoto, 2002; Mashiyama *et al*., 2008). Epidemiologic studies of humans have shown increased risk, although not consistently, of cancers including breast cancer, associated with low intake or low blood concentrations of folate, vitamin B12 and B6 (Choi and Mason, 2000; Giovannucci, 2002; Zhang *et al*., 2003). It has therefore been postulated that the DNA damage caused by uracil misincorporation may explain the increased cancer risk associated with inadequate availability of these vitamins (Ames and Wakimoto, 2002).

Very little is known about the functional effects of uracil-processing gene polymorphisms. It is thought that they can lead to altered enzyme activity, contributing to uracil concentrations, increased uracil misincorporation and therefore to human disease (Sousa *et al*., 2007). One study detected novel germline sequence variations in *TDG*, *UNG* and *SMUG1* in colorectal cancer patients with familial aggregation, suggesting that these variants may have a role in disease susceptibility (Broderick *et al*., 2006). Other studies have described an association of a single nucleotide polymorphism (SNP) in *MBD4* with increased risk of lung (Shin *et al*., 2006; Miao *et al*., 2008) and esophageal cancer (Hao *et al*., 2004). Finally, in a previous study, some of us performed an extensive screening of SNPs in *UNG*, *SMUG1*, *MBD4*, *TDG* and *DUT*, and discovered that two SNPs in *SMUG1* and one SNP in *UNG* were associated with increased blood uracil DNA concentration in carriers of the variant genotype, whereas one SNP in the *DUT* gene was associated with decreased uracil DNA concentration (Chanson *et al*., 2009).

The association of the four SNPs in *SMUG1*, *UNG* and *DUT* with altered uracil DNA levels and subsequent DNA damage led to the hypothesis that these SNPs may influence cancer risk and may affect DNA mutations in tumors. In order to test our hypothesis we sought out a cohort of subjects whose cancer was both a common one and one in which a reasonably extensive set of epidemiologic observations had shown it to be linked to the intake of 1-carbon nutrients. Breast cancer fulfilled both of these criteria. Moreover, we chose to study *p53* mutations, as these mutations are among the most frequent genetic events in human cancer including 25–40% of breast cancers, and uracil misincorporation contributes to mutagenesis, and based on our previous observation that plasma uracil concentrations are correlated with strand breaks in the coding region of the human *p53* gene (Chanson *et al*., 2009).

Therefore, we investigated the association of these four SNPs with breast cancer risk in a population-based breast cancer case control study, the Western New York Exposures and Breast Cancer Study (WEB). Moreover, we investigated interactions with dietary folate, vitamin B_6 and B_{12} and examined whether there was increased likelihood of mutations in the *p53* gene in tumors associated with these SNPs.

Subjects and methods

Subjects

Subjects were participants in the Western New York Exposures and Breast Cancer (WEB) Study, a population based case–control study described in detail previously (Bonner *et al*., 2005; Tao *et al*., 2009). Briefly, 1170 cases (35–79 years old) accrued between 1996–2001, residents of two counties in upstate New-York, were incident, primary, histologically confirmed breast cancer cases. Controls ($n = 2115$) were randomly selected from the New York State Department of Motor Vehicles driver's license list age $\leqslant 65$ years) and the Health Care Finance Administration rolls (age > 65 years), and were frequency matched to cases on age (5-year intervals) and race. Both cases and controls had no previous history of cancer other than non-melanoma skin cancer. The study protocol was approved by the Institutional Review Boards of the University at Buffalo and all participating institutions, and written informed consent was obtained from all participants.

Data on participant demographics, medical history, reproductive history, lifetime alcohol consumption and other breast cancer risk factors were collected by extensive in-person interviews. A modified version of the self-administered Health Habits and History food frequency questionnaire (FFQ) was used for cases and controls to collect dietary intake data 12–24 months before the interview, and the DietSys nutrient analysis software (HHHQ-DietSys Analysis Software, Version 4.02, National Cancer Institute, 1999) was used to calculate nutrient intakes, adjusted for total energy (Block *et al*., 1986; Willett *et al*., 1997).

Laboratory methods

Blood or saliva samples were obtained from 1099 (94%) cases and 1945 (92%) controls, and genomic DNA was extracted from buffy coat fractions or buccal cells (17% of cases and 8% of controls). Genotyping for the four SNPs investigated was

performed using an allelic discrimination method by real time PCR with TaqMan probes, in an ABI 7900HT real time PCR instrument (Applied Biosystems, Foster City, CA, USA), according to the manufacturer's instructions. As quality control measures, 10% duplicate samples were included in all measurements and all technical personnel was blinded to the case-control status of samples. The concordance rates between blinded duplicates were 98.1% for rs2029166, 98.7% for rs7296239, 97.4% for rs34259 and 97.4% for rs4775748. The call rates were 98.4% for rs2029166, 97.4% for rs7296239, 98.1% for rs34259 and 98.8% for rs4775748, hence in part the reason for the slightly different number of women analyzed for each SNP.

Archived tumor blocks were obtained from 920 (78.60%) of cases, and clinical and pathological characteristics were abstracted from medical records. Estrogen receptor status was determined by immunohistochemistry and slides were scored by a certified pathologist as previously described (Tao *et al*., 2009). Screening for p53 mutation status was performed on all tumor blocks as previously described (Tennis *et al*., 2006). Briefly, DNA was extracted from the tumor tissue on slides after micro-dissection using a pathologically confirmed HE stained slide as a guidance. Screening for *P53* mutations was performed using the Affymetrix*p53*GeneChip system (Affymetrix, Santa Clara, CA, USA), and mutation status was confirmed by bi-directional sequencing.

Statistical analyses

All statistical analyses were performed using SAS v.9. 1. (SAS Institute, Cary, NC, USA). Hardy–Weinberg expectations were tested using the exact χ^2 goodness-of-fit tests. Unconditional logistic regression models were used to calculate odds ratios (OR) and 95% confidence intervals (CI) for associations with breast cancer risk according to SNP genotype status. We stratified women in categories defined by the median dietary intake of controls for folate, vitamin B_6 and vitamin B_{12} intake, and tested the association with genotype within these strata for each vitamin. Interactions between genotype and folate, vitamins B_6 and B_{12} intake were investigated by the evaluation of a multiplicative term in the regression model. Unconditional logistic regression was used in case–case comparisons to estimate ORs and 95% CIs according to p53 mutations status. All models were adjusted for age and education. In addition models were also adjusted for total energy intake when we tested the associations of genotype with breast cancer risk stratified by dietary intake of folate, vitamin B_6 and B_{12}. The case-case comparisons were additionally adjusted for estrogen receptor status. Other potential confounders including demographic factors and breast cancer risk factors (body mass index, age at first birth, age at menarche, age at menopause (for postmenopausal women), parity, family history of breast cancer, history of benign breast disease), alcohol intake and smoking status were examined as potential confounders; estimates of ORs were changed by less than 10% (data not shown).

Results

Characteristics of the study participants

Descriptive characteristics of the WEB study cases and controls have been presented elsewhere (Bonner *et al*., 2005; Tao *et al*., 2009). The *SMUG1* rs7296239 SNP was not in Hardy Weinberg equilibrium for the whole group ($p<0.05$) whereas it was in Hardy Weinberg equilibrium for Caucasians only. Therefore we limited further analyses for all SNPs to Caucasians, who constitute more than 90% of the study population, with 1077 cases and 1910 controls. Briefly, Table 1 presents the main characteristics among Caucasians and the allelic distribution of the investigated SNPs.

Associations between uracil processing genes variants and breast cancer risk

ORs and 95% CIs for breast cancer risk by genotype for the four SNPs were investigated, overall and stratified according to menopausal status, the results of which are shown in Table 2.

The genotypes for *UNG* rs34259 and *DUT* rs4775748 were not associated with breast cancer risk regardless of menopausal status for all of the genetic models tested (dominant, recessive and codominant; results not shown for the different genetic models). Among premenopausal women, there was no association with risk for either of the *SMUG1* genotypes. Among postmenopausal women, cancer risk was increased among those with the CT genotype of the rs7296239 SNP (OR 1.29, 95% CI 1.07–1.55) but not the CC (OR 1.03, 95% CI 0.74–1.43) genotype. Similarly there was an increased risk associated with the heterozygous CT genotype only for the rs2029166 genotype (OR 1.29, 95% CI 1.07–1.56).

Interactions between folate, vitamin B_6, vitamin B_{12} intakes and SNP genotypes, and the effect on breast cancer risk

As shown in Table 3, an association of the SMUG1 rs2029166 genotype with premenopausal breast cancer risk was observed, but limited to participants with lower intake of folate (p-interaction, 0.01). For the SMUG1 rs7296239 SNP, there was a similar association limited to those with low folate although the association was not statistically significant. Among postmenopausal women, there was no interaction with folate intake. There was no evidence on interaction of genotype by intake of either vitamins B_6 or B_{12} (data not shown).

Case-only analysis for associations with p53 mutations, and interactions between dietary measures and genetic polymorphisms across p53 mutation status

The *p53* mutation rate was 27% among cases, the majority of these mutations (94%) being found in exons 5–8 of the *p53* gene. In case-only analysis (Table 4), there were no

Table 1 Allelic frequencies for the *SMUG1*, *UNG* and *DUT* SNPs and distribution of selected breast cancer risk factors among Caucasian cases and controls, WEB Study

	Cases (n = *1077*)	*Controls* (n = *1910*)	P-*value*
Age, years[a]	58.2 ± 11.1	57.3 ± 11.7	0.04[b]
Education level, years			
<12	78 (7.2%)	167 (8.7%)	
12	422 (39.2%)	731 (38.3%)	
>12	577 (53.6%)	1012 (53.0%)	0.35[c]
Postmenopausal	778 (72.2%)	1331 (69.7%)	0.14[c]
Hormone replacement therapy[d]	425 (39.9%)	665 (36.3%)	0.05[c]
Alcohol consumption			
Never	169 (15.8%)	263 (13.9%)	
Ever	900 (84.2%)	1631 (86.1%)	0.15[c]
Smoking status			
Never	485 (45.1%)	928 (48.7%)	
Former	462 (43.0%)	697 (36.6%)	
Current	128 (11.9%)	279 (14.7%)	<0.01[c]
Age at menarche[a]	12.6 ± 1.6	12.7 ± 1.6	0.06[b]
Age at menopause[a]	48.4 ± 5.3	47.6 ± 6.0	<0.01[b]
BMI[a]	28.3 ± 6.3	28.0 ± 6.2	0.16[b]
Breast cancer among first-degree relatives	210 (21.0%)	238 (13.3%)	<0.01[c]
SMUG1 (rs2029166)			
C	72%	74%	
T	28%	26%	0.08[c]
p (HWE)	0.38	0.11	
SMUG1 (rs7296239)			
T	67%	70%	
C	33%	30%	0.06[c]
p (HWE)	0.04	0.12	
UNG (rs34259)			
G	80%	78%	
C	20%	22%	0.21[c]
p (HWE)	0.15	0.86	
DUT (rs4775748)			
T	83%	84%	
G	17%	16%	0.16[c]
p (HWE)	0.83	0.42	

Abbreviations: BMI, body mass index; HWE, Hardy–Weinberg equilibrium; SNP, single nucleotide polymorphism; WEB, Western New York Exposures and Breast Cancer Study.
[a]Mean ± s.d.
[b]Student's *t*-test was used to compare continuous variables.
[c]χ^2-test was used to compare categorical variables.
[d]Among postmenopausal women.

Table 2 *SMUG1*, *UNG* and *DUT* SNP genotypes and risk of breast cancer among Caucasians (WEB Study)

	Premenopausal			*Postmenopausal*		
	Cases	*Controls*	*Adjusted OR*[a]	*Cases*	*Controls*	*Adjusted OR*[a,b]
SMUG1 (rs2029166)						
CC	142	292	1.00	369	694	1.00
CT	111	194	1.15 (0.86–1.56)	295	447	1.29 (1.07–1.56)
TT	21	41	1.06 (0.61–1.86)	49	86	1.05 (0.72–1.52)
CT/TT	132	235	1.14 (0.85–1.52)	344	533	1.22 (1.01–1.47)
SMUG1 (rs7296239)						
TT	120	252	1.00	310	611	1.00
CT	126	218	1.19 (0.89–1.60)	331	488	1.29 (1.07–1.55)
CC	26	54	1.01 (0.61–1.69)	64	118	1.03 (0.74–1.43)
CT/CC	152	272	1.16 (0.86–1.56)	395	606	1.29 (1.07–1.56)
UNG (rs34259)						
GG	160	333	1.00	456	735	1.00
CG	104	170	1.27 (0.94–1.73)	232	433	0.87 (0.71–1.05)
CC	11	26	0.85 (0.41–1.77)	23	57	0.65 (0.40–1.07)
CG/CC	115	196	1.22 (0.90–1.64)	255	490	0.85 (0.70–1.03)
DUT (rs4775748)						
TT	191	387	1.00	476	850	1.00
GT	70	123	1.14 (0.82–1.60)	205	321	1.14 (0.93–1.40)
GG	7	13	1.11 (0.43–2.82)	20	33	1.10 (0.62–1.94)
GT/GG	77	136	1.16 (0.83–1.61)	225	354	1.15 (0.94–1.40)

Abbreviations: OR, odds ratio; SNP, single nucleotide polymorphism; WEB, Western New York Exposures and Breast Cancer Study.
[a]Adjusted for age and education.
[b]Also adjusted for age at menopause for postmenopausal women.

significant associations between the four genetic polymorphisms in uracil processing genes and p53 mutation status, although there were some trends towards increased risk of *p53* mutations in premenopausal women with the CT/TT and the CT/CC genotypes for *SMUG1* rs2029166 and rs7296239. Interestingly, in each of these two trends the heterozygotes demonstrated a more robust OR compared with the homozygotes, although this most likely could be due to the fact that there are much less rare homozygotes than heterozygotes.

Discussion

We examined the association with breast cancer risk of four single nucleotide polymorphisms in three uracil processing genes (two SNPs in *SMUG1*, one SNP in *UNG* and one SNP in *DUT*), which have been previously shown to be associated with altered blood uracil levels in humans (Chanson *et al*., 2009). Given their association with uracil levels, we hypothesized that these SNPs might be associated with *p53* mutations, and that folate and related B vitamin intake would modify associations with breast cancer risk. To our knowledge this is the only study investigating uracil processing gene variants in relation to breast cancer risk.

We found associations between breast cancer risk and the heterozygous genotypes for *SMUG1* rs2029166 and rs7296239 among postmenopausal women, but not the *UNG* rs34259 or *DUT* rs4775748 genotypes.

To date, only one study investigated sequence alterations in *UNG* and *SMUG1* and their involvement in cancer risk by performing a mutation screening in these genes and others in colorectal cancer patients with familial aggregation (Broderick *et al*., 2006). They did not detect the same sequence variants that we investigated herein; however, they detected both novel and known sequence variants in these genes that were present only in the cancer patients and

Table 3 Risk of breast cancer associated with *SMUG1*, *UNG* and *DUT* SNP genotypes by folate intake, among Caucasians (WEB Study)

Folate[a]	*Premenopausal*				*Postmenopausal*			
	High		*Low*		*High*		*Low*	
	Cases/controls	*OR (95% CI)*[a]	*Cases/controls*	*OR (95% CI)*[a]	*Cases/controls*	*OR(95% CI)*[a]	*Cases/controls*	*OR (95% CI)*[a]
SMUG1 (rs2029166)								
CC	80/116	1.00	62/176	1.00	199/360	1.00	170/334	1.00
CT	54/102	0.72 (0.47–1.10)	57/92	1.76 (1.14–2.70)	160/222	1.30 (1.00–1.69)	135/225	1.14 (0.87–1.50)
TT	9/18	0.72 (0.31–1.69)	12/23	1.52 (0.71–3.22)	29/40	1.33 (0.80–2.22)	20/46	0.82 (0.47–1.42)
CT/TT	63/120	0.74 (0.49–1.14)	69/115	1.65 (1.08–2.50)	189/262	1.34 (1.04–1.73)	155/271	1.12 (0.85–1.46)
p interaction				0.01				0.34
SMUG1 (rs7296239)								
TT	65/103	1.00	55/149	1.00	170/316	1.00	140/295	1.00
CT	64/109	0.84 (0.55–1.29)	62/109	1.57 (1.03–2.40)	176/248	1.32 (1.02–1.71)	155/240	1.26 (0.96–1.65)
CC	14/26	0.81 (0.40–1.66)	12/28	1.20 (0.57–2.50)	40/51	1.49 (0.95–2.33)	24/67	0.69 (0.41–1.13)
CT/ CC	78/135	0.89 (0.58–1.37)	74/137	1.42 (0.93–2.17)	216/299	1.37 (1.06–1.78)	179/307	1.22 (0.93–1.60)
p interaction				0.13				0.52
UNG (rs34259)								
GG	84/149	1.00	76/184	1.00	241/375	1.00	215/360	1.00
CG	54/79	1.16 (0.75–1.79)	50/91	1.41 (0.92–2.16)	132/219	0.94 (0.72–1.23)	100/214	0.78 (0.59–1.04)
CC	6/9	1.05 (0.36–3.06)	5/17	0.74 (0.26–2.08)	15/28	0.88 (0.46–1.68)	8/29	0.45 (0.20–1.00)
CG/ CC	60/88	1.21 (0.79–1.86)	55/108	1.24 (0.81–1.90)	147/247	0.94 (0.72–1.22)	108/243	0.75 (0.57–1.00)
p interaction				0.93				0.25
DUT (rs4775748)								
TT	97/170	1.00	94/217	1.00	259/427	1.00	217/423	1.00
GT	40/63	1.04 (0.65–1.65)	30/60	1.21 (0.73–1.99)	110/162	1.11 (0.84–1.48)	95/159	1.14 (0.85–1.54)
GG	5/2	3.87 (0.73–20.5)	2/11	0.46 (0.10–2.11)	10/17	0.99 (0.45–2.19)	10/16	1.20 (0.54–2.70)
GT/ GG	45/65	1.19 (0.75–1.88)	32/71	1.09 (0.67–1.78)	120/179	1.10 (0.83–1.46)	105/175	1.18 (0.88–1.58)
p interaction				0.79				0.75

Abbreviations: CI, confidence interval; OR, odds ratio; SNP, single nucleotide polymorphism; WEB, Western New York Exposures and Breast Cancer Study.
ORs and 95% CIs adjusted for age, education and total energy intake.
[a]Folate cutoff for high and low categories at the median intake of total Caucasian controls, 239 μg per day.

not in normal controls, although at very low frequency. Moreover, *in silico* analysis of the effect of these variants on protein structure suggested that they may have a deleterious effect. The low frequency of these variants in their population suggests that these variants may have a limited role as low penetrance genes in colorectal cancer susceptibility (Broderick *et al*., 2006).

In our study, carriers of the variant alleles for these SNPs had 15–30% increased risk of breast cancer, especially among postmenopausal women. Interestingly, these associations were limited to the heterozygous genotype carriers for both SNPs, even though we had previously found associations of increased blood uracil with the TT genotype for the rs2029166 SNP and with the CC genotype for the rs7296239 SNP (Chanson *et al*., 2009). Our results suggest a possible molecular heterosis effect; however we cannot exclude the possibility that the associations with the homozygote variant genotypes were not statistically significant due to fewer subjects carrying this genotype in our population. Molecular heterosis takes place when heterozygous individuals for a specific locus show a significant increased or decreased effect for a quantitative or dichotomous trait compared with homozygous subjects, and is a rather controversial concept in population genetics, although well documented in plant, animal and human studies (Comings and MacMurray, 2000; Crow, 2008). Several well-documented examples of molecular heterosis in human gene association studies are presented in a review on the subject by Comings and MacMurray 2000 (Comings and MacMurray, 2000). More recently, a meta-analysis of 19 case–control and three genome wide association studies of type 2 diabetes revealed that molecular heterosis still remains a plausible scenario explaining the effect of certain SNPs associated with disease (Salanti *et al*., 2009). We therefore contend that the present observations do not justify a definitive conclusion for a null effect of these SNPs.

We found that there was an interaction of genotype and folate intake for premenopausal breast cancer for the *SMUG1* rs2029166 SNP. The association with genotype was limited to participants with low folate intake. For the *SMUG1* rs7296239 SNP, results were similar but the interaction was not statistically significant. There was no evidence of interaction for postmenopausal breast cancer for these two genotypes. In addition, there was no interaction for the other SNPs examined or for vitamin B_6 and B_{12}. In our previous study of these SNPs in a Puerto Rican population

Table 4 Likelihood of *p53* mutations in tumors by *SMUG1*, *UNG* and *DUT* SNP genotypes among Caucasians (WEB Study)

	p53 mutation					
	(+)	(−)	*OR (95% CI)*[a]	(+)	(−)	*OR (95% CI)*[a]
	Premenopausal			*Postmenopausal*		
SMUG1 (rs2029166)						
CC	35	64	1.00	57	177	1.00
CT	18	59	1.74 (0.89–3.40)	47	121	0.93 (0.60–1.45)
TT	4	12	1.30 (0.38–4.39)	12	22	0.65 (0.30–1.39)
CT/TT	22	71	1.78 (0.93–3.39)	59	143	0.80 (0.52–1.22)
SMUG1 (rs7296239)						
TT	30	52	1.00	51	153	1.00
CT	21	68	1.84 (0.95–3.57)	49	137	1.10 (0.71–1.70)
CC	5	15	1.51 (0.50–4.59)	14	30	0.81 (0.40–1.63)
CT/CC	26	83	1.89 (0.99–3.60)	63	167	0.91 (0.59–1.40)
UNG (rs34259)						
GG	34	81	1.00	73	209	1.00
CG	19	55	1.19 (0.61–2.29)	38	100	0.99 (0.63–1.55)
CC	5	2	0.09 (0.01–0.86)	5	11	0.73 (0.24–2.19)
CG/ CC	24	57	1.00 (0.53–1.90)	43	111	0.89 (0.57–1.39)
DUT (rs4775748)						
TT	39	92	1.00	80	223	1.00
GT	14	37	1.16 (0.56–2.39)	29	81	1.10 (0.67–1.79)
GG	1	4	1.91 (0.20–17.86)	2	13	2.42 (0.53–10.96)
GT/ GG	15	41	1.18 (0.58–2.40)	31	94	1.08 (0.67–1.76)

Abbreviations: CI, confidence interval; OR, odds ratio; SNP, single nucleotide polymorphism; WEB, Western New York Exposures and Breast Cancer Study.
[a]ORs and 95% CIs for case-case comparisons, adjusted for age, education and ER status.

including both men and women (Chanson *et al*., 2009), we did not see an effect of plasma folate and B vitamin levels on the association with altered uracil levels. The association of the *SMUG1* rs2029166 genotype with breast cancer risk among premenopausal women with low intake of folate implies that a higher intake of these nutrients may be necessary in order to maintain genomic stability of human cells, as previously suggested (Fenech, 2001).

The fact that all positive associations were limited to those with the heterozygous genotype possibly implies the presence of molecular heterosis in our study. However, caution should be used when interpreting these results due to the low magnitude of the associations. We had previously found associations with plasma uracil that were limited to homozygous genotypes for both SNPs. It may be that multiple factors are involved in the complex association between genes involved in uracil and one carbon metabolism and breast cancer risk, also explaining the inconsistency in epidemiology studies on the matter (Xu and Chen, 2009).

In this study, there was a suggestion of an increased number of p53 mutations among premenopausal breast cancers from women with the heterozygous genotype for both *SMUG1* SNPs. These findings, although not statistically significant, are consistent with uracil misincorporation contributing to mutagenesis and with our previous observation that plasma uracil concentrations are correlated with strand breaks in the coding region of the human *p53* gene (Chanson *et al*., 2009).

The strengths of this study include the population-based study design and the relatively large number of subjects. The detailed exposure information enabled an investigation of gene-environment interaction. However, the statistical power in subgroups of our study remained limited due to low frequencies of the variant alleles, which limited our ability to identify weak associations. We studied four polymorphisms that had been shown in another population to be associated with blood uracil concentrations but with unknown functional effects on the enzymatic activity of the encoded proteins, so we cannot rule out the possibility that other unidentified SNPs tightly linked to the ones described herein might instead be the functional SNPs that are associated with breast cancer risk. Although there are concerns with the FFQ as a measure of usual diet and of absolute intakes, there is evidence that this measure is able to provide information regarding relative intakes, that is, provide information regarding high and low intakes such as in this study. This study was conducted in 1996–2001, during the period of implementation of folate fortification of flour in the United States in 1998. Although clearly the fortification impacted intakes, it is unlikely that it affected carcinogenesis of study participants because of the long period of that process. For these purposes, we were interested in the usual diet in period before diagnosis for cases and before the interview for controls; we did not include fortification in the calculation of intakes. Another concern is recall bias. Compared with controls, cases might differentially recall exposures of dietary because of their cancer diagnosis. However, this bias would likely not affect our comparisons, in that the participants would not be aware of their genotypes.

In conclusion, our study suggests that the four selected SNPs are not robust determinants of breast cancer risk, but that the two SNPs in *SMUG1* might modestly alter the risk of breast cancer, especially in postmenopausal women and in premenopausal women with lower intakes of folate and that these variants may contribute to p53 mutations. Future studies should be directed towards confirming these associations and probing for the presence of molecular heterosis.

Conflict of interest

The authors declare no conflict of interest.

References

Ames BN (1999). Micronutrient deficiencies. A major cause of DNA damage. *Ann N Y Acad Sci* **889**, 87–106.
Ames BN, Wakimoto P (2002). Are vitamin and mineral deficiencies a major cancer risk? *Nat Rev Cancer* **2**, 694–704.
Barnes DE, Lindahl T (2004). Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet* **38**, 445–476.

Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, Gardner L (1986). A data-based approach to diet questionnaire design and testing. *Am J Epidemiol* **124**, 453–469.

Blount BC, Mack MM, Wehr CM, MacGregor JT, Hiatt RA, Wang G *et al.* (1997). Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proc Natl Acad Sci USA* **94**, 3290–3295.

Bonner MR, Han D, Nie J, Rogerson P, Vena JE, Muti P *et al.* (2005). Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. *Cancer Epidemiol Biomarkers Prev* **14**, 53–60.

Broderick P, Bagratuni T, Vijayakrishnan J, Lubbe S, Chandler I, Houlston RS (2006). Evaluation of NTHL1, NEIL1, NEIL2, MPG, TDG, UNG and SMUG1 genes in familial colorectal cancer predisposition. *BMC Cancer* **6**, 243.

Chanson A, Parnell LD, Ciappio ED, Liu Z, Crott JW, Tucker KL *et al.* (2009). Polymorphisms in uracil-processing genes, but not one-carbon nutrients, are associated with altered DNA uracil concentrations in an urban Puerto Rican population. *Am J Clin Nutr* **89**, 1927–1936.

Choi SW, Mason JB (2000). Folate and carcinogenesis: an integrated scheme. *J Nutr* **130**, 129–132.

Comings DE, MacMurray JP (2000). Molecular heterosis: a review. *Mol Genet Metab* **71**, 19–31.

Crow JF (2008). Mid-century controversies in population genetics. *Annu Rev Genet* **42**, 1–16.

Fenech M (2001). The role of folic acid and vitamin B12 in genomic stability of human cells. *Mutat Res* **475**, 57–67.

Giovannucci E (2002). Epidemiologic studies of folate and colorectal neoplasia: a review. *J Nutr* **132**, 2350S–2355S.

Hao B, Wang H, Zhou K, Li Y, Chen X, Zhou G *et al.* (2004). Identification of genetic variants in base excision repair pathway and their associations with risk of esophageal squamous cell carcinoma. *Cancer Res* **64**, 4378–4384.

Hung RJ, Hall J, Brennan P, Boffetta P (2005). Genetic polymorphisms in the base excision repair pathway and cancer risk: a HuGE review. *Am J Epidemiol* **162**, 925–942.

Jang H, Mason JB, Choi SW (2005). Genetic and epigenetic interactions between folate and aging in carcinogenesis. *J Nutr* **135**, 2967S–2971S.

Kapiszewska M, Kalemba M, Wojciech U, Milewicz T (2005). Uracil misincorporation into DNA of leukocytes of young women with positive folate balance depends on plasma vitamin B12 concentrations and methylenetetrahydrofolate reductase polymorphisms. A pilot study. *J Nutr Biochem* **16**, 467–478.

Kavli B, Otterlei M, Slupphaug G, Krokan HE (2007). Uracil in DNA—general mutagen, but normal intermediate in acquired immunity. *DNA Repair (Amst)* **6**, 505–516.

Krokan HE, Drablos F, Slupphaug G (2002). Uracil in DNA—occurrence, consequences and repair. *Oncogene* **21**, 8935–8948.

Mashiyama ST, Hansen CM, Roitman E, Sarmiento S, Leklem JE, Shultz TD *et al.* (2008). An assay for uracil in human DNA at baseline: effect of marginal vitamin B6 deficiency. *Anal Biochem* **372**, 21–31.

Miao R, Gu H, Liu H, Hu Z, Jin G, Wang H *et al.* (2008). Tagging single nucleotide polymorphisms in MBD4 are associated with risk of lung cancer in a Chinese population. *Lung Cancer* **62**, 281–286.

Mohrenweiser HW, Jones IM (1998). Variation in DNA repair is a factor in cancer susceptibility: a paradigm for the promises and perils of individual and population risk estimation? *Mutat Res* **400**, 15–24.

Salanti G, Southam L, Altshuler D, Ardlie K, Barroso I, Boehnke M *et al.* (2009). Underlying genetic models of inheritance in established type 2 diabetes associations. *Am J Epidemiol* **170**, 537–545.

Shin MC, Lee SJ, Choi JE, Cha SI, Kim CH, Lee WK *et al.* (2006). Glu346Lys polymorphism in the methyl-CpG binding domain 4 gene and the risk of primary lung cancer. *Jpn J Clin Oncol* **36**, 483–488.

Sousa MM, Krokan HE, Slupphaug G (2007). DNA-uracil and human pathology. *Mol Aspects Med* **28**, 276–306.

Tao MH, Shields PG, Nie J, Millen A, Ambrosone CB, Edge SB *et al.* (2009). DNA hypermethylation and clinicopathological features in breast cancer: the Western New York Exposures and Breast Cancer (WEB) Study. *Breast Cancer Res Treat* **114**, 559–568.

Tennis M, Krishnan S, Bonner M, Ambrosone CB, Vena JE, Moysich K *et al.* (2006). p53 Mutation analysis in breast tumors by a DNA microarray method. *Cancer Epidemiol Biomarkers Prev* **15**, 80–85.

Visnes T, Doseth B, Pettersen HS, Hagen L, Sousa MM, Akbari M *et al.* (2009). Uracil in DNA and its processing by different DNA glycosylases. *Philos Trans R Soc Lond B Biol Sci* **364**, 563–568.

Willett WC, Howe GR, Kushi LH (1997). Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, 1220S–1228S.

Xu X, Chen J (2009). One-carbon metabolism and breast cancer: an epidemiological perspective. *J Genet Genomics* **36**, 203–214.

Zhang SM, Willett WC, Selhub J, Hunter DJ, Giovannucci EL, Holmes MD *et al.* (2003). Plasma folate, vitamin B6, vitamin B12, homocysteine, and risk of breast cancer. *J Natl Cancer Inst* **95**, 373–380.