OXFORD

Carcinogenesis, 2016, Vol. 37, No. 12, 1138–1143

doi:10.1093/carcin/bgw107
Advance Access publication October 8, 2016
Original Manuscript

ORIGINAL MANUSCRIPT

A genome-wide analysis of gene–caffeine consumption interaction on basal cell carcinoma

Xin Li[1], Marilyn C.Cornelis[2], Liming Liang[1], Fengju Song[3,4], Immaculata De Vivo[1,5], Edward Giovannucci[1,5,6], Jean Y.Tang[7] and Jiali Han[8,9,10*]

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston MA 02115, USA, [2]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago IL 60611, USA, [3]Department of Epidemiology and Biostatistics, Tianjin Medical University Cancer Hospital and Institute, Tianjin 300060, China, [4]National Clinical Research Center for Cancer and Key Laboratory of Cancer Prevention and Therapy, Tianjin 300060, China, [5]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA, [6]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA, [7]Department of Dermatology, Stanford University School of Medicine, Redwood City, CA 94063, USA, [8]Department of Epidemiology, Fairbanks School of Public Health, Indiana University, Indianapolis, IN 46202, USA, [9]Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, IN 46202, USA and [10]Center for Pharmacoepidemiology, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN 46202, USA

*To whom correspondence should be addressed. Department of Epidemiology, Richard M. Fairbanks School of Public Health, Indiana University, 1050 Wishard Boulevard, RG 5112, Indianapolis, IN 46202-2872, USA. Tel: 317-278-0370; Fax: 317-278-0370; Email: jialhan@iu.edu

Abstract

Animal models have suggested that oral or topical administration of caffeine could inhibit ultraviolet-induced carcinogenesis via the ataxia telangiectasia and rad3 (ATR)-related apoptosis. Previous epidemiological studies have demonstrated that increased caffeine consumption is associated with reduced risk of basal cell carcinoma (BCC). To identify common genetic markers that may modify this association, we tested gene–caffeine intake interaction on BCC risk in a genome-wide analysis. We included 3383 BCC cases and 8528 controls of European ancestry from the Nurses' Health Study and Health Professionals Follow-up Study. Single nucleotide polymorphism (SNP) rs142310826 near the *NEIL3* gene showed a genome-wide significant interaction with caffeine consumption ($P = 1.78 \times 10^{-8}$ for interaction) on BCC risk. There was no gender difference for this interaction ($P = 0.64$ for heterogeneity). *NEIL3*, a gene belonging to the base excision DNA repair pathway, encodes a DNA glycosylase that recognizes and removes lesions produced by oxidative stress. In addition, we identified several loci with *P* value for interaction $<5 \times 10^{-7}$ in gender-specific analyses (*P* for heterogeneity between genders < 0.001) including those mapping to the genes *LRRTM4*, *ATF3* and *DCLRE1C* in women and *POTEA* in men. Finally, we tested the associations between caffeine consumption-related SNPs reported by previous genome-wide association studies and risk of BCC, both individually and jointly, but found no significant association. In sum, we identified a DNA repair gene that could be involved in caffeine-mediated skin tumor inhibition. Further studies are warranted to confirm these findings.

Introduction

Basal cell carcinoma (BCC), a major histological type of non-melanoma skin cancer, is the most common malignancy among populations of European ancestry (1). Like other common disorders, BCC is thought to have both environmental and genetic components and to involve their interactions. Some known risk factors for BCC include exposure to ultraviolet (UV) radiation, family history of skin cancer and lighter pigmentation (2,3). Habitual coffee, tea and total caffeine consumption has also been associated with lower incidence of non-melanoma skin cancers according to several epidemiological studies (4–7). In a recent prospective study using data from the Nurses' Health Study (NHS) and the Health Professionals Follow-up Study

Received: June 30, 2016; **Revised:** September 16, 2016; **Accepted:** October 6, 2016

Abbreviations

ATR	ataxia telangiectasia and rad3
BCC	basal cell carcinoma
GWAS	Genome-wide association studies
HPFS	Health Professionals Follow-up Study
MAF	minor allele frequency
NHS	Nurses' Health Study
SNP	single nucleotide polymorphism
UV	ultraviolet

(HPFS), researchers found a significant inverse association between total caffeine intake and risk of BCC (7). These observations are supported by experimental work, whereby oral or topical administration of caffeine inhibits UV-induced carcinogenesis in mice (8,9). Though the mechanisms underlying the protective effect of caffeine on BCC development are unclear, some hypothesize that caffeine might augment apoptosis in UV-damaged keratinocytes through ataxia telangiectasia and rad3 (ATR)-related kinase and its downstream effector, checkpoint kinase-1 (Chk1) (10–12).

Genome-wide association studies (GWASs) have identified several genetic loci that confer susceptibility to BCC (13–15); however, they tested only for associations between individual genetic markers and risk of BCC without taking interactions into consideration. Testing for gene–environment interactions on a genome-wide scale may uncover novel loci and pathways implicated in BCC. This approach has been successful for colorectal cancer (16), Parkinson's disease (17) and type 2 diabetes (18), but to the best of our knowledge has not been applied in the context of BCC. In this study, we conducted genome-wide analyses of gene–caffeine consumption interactions among men and women of the HPFS and NHS, respectively. Our study may provide new biologic insights into caffeine's role in BCC development and lead to discovery of BCC-related genes that have been missed in conventional GWASs.

Methods

Study population, inclusion and exclusion

Eighteen case–control studies nested within the NHS and HPFS with cleaned genotype data were included in our study. We treated 1986 as the baseline year because caffeine intake was first measured in 1984 in the NHS and 1986 in the HPFS, and excluded participants who had other common cancers before 1986. BCC cases who had other common cancers before diagnosis of BCC were excluded. Eligible controls were free of BCC and other common cancers. Participants with identical genetic information but different cohort IDs were removed, and participants sampled in more than one of the 18 case–control studies were included only once. In total, 3383 BCC cases and 8528 controls of European ancestry were included. See Supplementary Materials, available at *Carcinogenesis* Online, for more detailed descriptions of the NHS, the HPFS and the 18 nested case–control studies. The study protocol was approved by the Institutional Review Boards of Brigham and Women's Hospital and the Harvard T.H. Chan School of Public Health.

Genotyping, quality control and imputation

Detailed descriptions of quality control and imputation are provided in Supplementary Materials, available at *Carcinogenesis* Online. Samples from the 18 nested case–control studies were genotyped using a variety of platforms. We combined these datasets into three compiled datasets based on their genotype platform type: Affymetrix (Affy), Illumina HumanHap series (Illumina) or Illumina Omni Express (Omni) (Supplementary Table 1, available at *Carcinogenesis* Online). Quality control on SNP completion rate, sample completion rate, ancestry consistency, deviation from Hardy–Weinberg equilibrium, Mendelian consistency, minor allele frequency (MAF) and duplication samples were conducted within each of the three combined datasets. We then imputed the complied datasets using the 1000 Genomes Project ALL Phase I Integrated Release Version 3 Haplotypes excluding monomorphic and singleton sites (2010–2011 data freeze, 2012-03-14 haplotypes) as a reference panel. We included genetic markers with imputation *R*-square ≥ 0.3 and MAF ≥ 1% in further analysis. The numbers of such markers in the three combined datasets are presented in Supplementary Table 2, available at *Carcinogenesis* Online.

Caffeine intake

Information on consumption of coffee and other common caffeine-containing foods and beverages, including tea, cola and chocolate, was collected by validated food-frequency questionnaire. The questionnaires were completed in 1984, 1986, 1990, 1994, 1998, 2002 and 2006 for the NHS, and in 1986, 1990, 1994, 1998, 2002 and 2006 for the HPFS. On all questionnaires, participants were asked how many times on average during the previous year they had consumed each food and beverage. The participants could choose from nine frequency responses (never, 1–3 per month, 1 per week, 2–4 per week, 5–6 per week, 1 per day, 2–3 per day, 4–5 per day and ≥6 per day). Based on information obtained from the food-frequency questionnaire, the total intake of caffeine was calculated by multiplying the reported frequency of each food by the caffeine content of one serving of that food (1 cup for coffee or tea, one 12-ounce bottle or can for carbonated beverages and 1 ounce for chocolate). According to the U.S. Department of Agriculture food composition sources, caffeine content is 137 mg per cup of caffeinated coffee, 47 mg per cup of tea, 46 mg per bottle or can of cola beverage and 7 mg per serving of chocolate candy. Food and nutrient intakes assessed by this dietary questionnaire, including caffeine, have been validated previously against two 1-week diet records. The observed correlation between the questionnaire and the diet record was about 0.9 for caffeine consumption (19,20). We used daily caffeine intake (mg) measured in 1986 in this study.

BCC ascertainment

Disease follow-up procedures are identical for the NHS and the HPFS. Self-reported BCC case–control status has been updated every 2 years starting in 1984 in the NHS and 1986 in the HPFS without further pathological confirmation. The latest updates were made in 2008 in the NHS and 2010 in the HPFS. Colditz *et al.* (21) evaluated the validity of self-reported illnesses including skin cancer in the NHS. Among 33 random samples of women who had reported non-melanoma skin cancer, medical records indicated that 30 (91%) had correctly reported their skin cancer. Also, Hunter *et al.* (22) previously examined risk factors for BCC in the NHS using self-reported cases. As expected, they found that lighter pigmentation and higher tendency to sunburn were associated with an increased risk of BCC. In addition, using the self-reported BCC cases, our group identified the previously well-documented genetic variant in the MC1R gene as the top risk locus in our GWAS for BCC (15). These data support the validity of self-report of BCC in our study.

Statistical analysis

To account for gender differences (cohort differences), we divided each of the combined datasets into two parts and conducted genome-wide gene–environment (G–E) interaction analysis in the six datasets (Illumina_NHS, Illumina_HPFS, Affy_NHS, Affy_HPFS, Omni_NHS and Omni_HPFS). We used standard logistic regression with a product term to test the interaction between caffeine consumption and genetic markers in relation to BCC risk, adjusted for age in 1986, and the first three principal components (PCs) from EIGENSTRAT (23) to account for population substructure. The formula for logistic regression model was: Logit(Pr(D=1)) = b0 + b1*quartiles of caffeine consumption + b2*E(G) + b3*quartiles of caffeine consumption*E(G) + b4*age in 1986 + b5*PC1 + b6*PC2 +b7*PC3, where E(G) was expected dosage for imputed SNPs and dosage for genotyped SNPs. We assumed additive genetic effects for all analyses. Quartiles of caffeine intake were defined within each of the six datasets using the full range from zero to maximum intake among controls. We coded the quartiles as an ordinal variable (1st quartile = 1, 2nd quartile=2, 3rd quartile =3, 4th quartile = 4) in the main analysis, and used the median values of each quartile to represent the corresponding intake levels (1st quartile = median intake of 1st quartile, 2nd quartile=median intake of 2nd quartile, etc.) for sensitivity analysis. We combined results from Illumina_NHS, Affy_NHS

and Omni_NHS using inverse variance-weighted meta-analyses in METAL software (24). The same procedure was implemented for the three HPFS datasets to obtain combined results for men. We calculated *P* value for heterogeneity between men and women with the Cochran Q test, and performed the third meta-analysis for all six datasets if no significant difference was found between genders. All analyses were conducted using the ProbABEL package (25) and R-3.0.2 (https://www.r-project.org).We mapped identified SNPs to their nearest genes (physical distance on the same chromosome) using the UCSC GRCh37/hg19 assembly.

Results

The characteristics of participants within each of the six subsets are provided in Table 1. Participants in the NHS (women) consumed more caffeine compared to those in the HPFS (men). BCC was more prevalent among males.

For our main analysis, in which quartiles of caffeine intake were modeled as ordinal variables and data from men and women were combined, the *P* value for interaction between each genetic marker and caffeine are shown in the Manhattan plot and quantile–quantile (Q–Q) plot (Supplementary Figures 1 and 2, available at *Carcinogenesis* Online). On chromosome 4, we determined that SNP rs142310826 (MAF = 1.9%) had a genome-wide significant interaction with caffeine consumption ($P = 1.78 \times 10^{-8}$ for interaction). Results of this SNP's interaction with caffeine intake in each of the six datasets and in meta-analyses are summarized in Table 2. The estimate for interaction between this SNP and caffeine intake in relation to BCC among women was not significantly different from that among men ($P = 0.62$ for gender heterogeneity).

In the gender-specific analysis, 19 genetic markers and 3 genetic markers with *P* value for interaction less than 5×10^{-7} were identified among women and men, respectively. For correlated markers, only the most significant ones are presented in Table 3. The top significant marker identified among females was 2:76738900:TTAGA ($P = 2.51 \times 10^{-8}$ for interaction), which was mapped on gene *LRRTM4*. Eleven genetic variants located very close to this top marker were also identified. As expected, the beta estimates, MAF and imputation quality of these related genetic markers are very similar. The other two regions identified in NHS were mapped to gene *ATF3* on chromosome 1 and gene *DCLRE1C* on chromosome 10, respectively. In men, three correlated SNPs on chromosome 8 (*POTEA* gene) were reported. However, the *P* value for interaction of the most significant one, rs77868414, failed to reach genome-wide significance ($P = 5 \times 10^{-8}$). All 22 gender-specific markers reported above (19 in NHS and 3 in HPFS) had *P* values for heterogeneity <0.001 or 0.001. Therefore, we did not calculate the combined estimates for them. Manhattan plots and Q–Q plots for gender-specific analysis are shown in Supplementary Figures 3–6, available at *Carcinogenesis* Online. Markers that interact with caffeine consumption in relation to BCC risk at the significance level of 5×10^{-6} are presented in Supplementary Tables 3–5, available at *Carcinogenesis* Online.

In sensitivity analysis, we used the median values of each quartile to represent the corresponding caffeine intake levels (1st quartile = median intake of 1st quartile, 2nd quartile = median intake of 2nd quartile, etc.), which did not change the results materially. Supplementary Table 6, available at *Carcinogenesis* Online, summarizes sensitivity analysis results for the independent markers discovered in our main analysis.

Genome-wide association studies have identified several genetic loci associated with caffeine/coffee consumption (26–29). We extracted results for these SNPs from our genome-wide G–E interaction analysis to better illustrate their potential

Table 1. Descriptive characteristics of study population

Datasets	*n* (%)		Gender	Age in 1986, mean	Quartiles of caffeine intake, mg/day, Median			
	BCC cases	BCC controls			First	Second	Third	Fourth
Illumina_NHS	544 (28.6%)	1355 (71.4%)	Female	54	45	150	353	555
Illumina_HPFS	302 (34.5%)	573 (65.5%)	Male	53	15	81	201	458
Affymetrix_NHS	785 (23.5%)	2556 (76.5%)	Female	54	39	160	354	630
Affymetrix _HPFS	818 (31.9%)	1748 (68.1%)	Male	55	11	115	348	630
OmniExpress_NHS	524 (25.7%)	1513 (74.3%)	Female	54	42	150	356	627
OmniExpress_HPFS	410 (34.4%)	783 (65.6%)	Male	54	17	120	348	542

Table 2. Results of rs142310826 (CHR:BP-4:179402856) in each of the six datasets and meta-analyses

Datasets	MAF ('A' allele)	Imputation *R*-square	Beta (SE) for interaction	*P* for interaction	*P* for heterogeneity	*P* for heterogeneity between genders
NHS (female)						
Illumina_NHS	0.0189	0.556	−0.65 (0.38)	8.64E-02	7.16E−01 (3 NHS sets)	6.17E−01 (compare combined NHS results versus combined HPFS results)
Affymetrix_NHS	0.0181	0.587	−0.78 (0.24)	1.51E-03		
OmniExpress_NHS	0.0183	0.598	−1.06 (0.36)	3.53E-03		
All NHS sets (meta-analysis)	0.0184[a]	0.580[a]	−0.82 (0.18)	4.96E-06		
HPFS (male)						
Illumina_HPFS	0.0189	0.556	−0.86 (0.47)	6.91E−02	6.39E−01 (3 HPFS sets)	
Affymetrix _HPFS	0.0181	0.587	−0.50 (0.28)	7.89E−02		
OmniExpress_HPFS	0.0183	0.598	−0.90 (0.38)	1.77E−02		
All HPFS sets (meta-analysis)	0.0184[a]	0.580[a]	−0.68 (0.20)	8.55E−04		
All six sets (meta-analysis)	0.0184[a]	0.580[a]	−0.76 (0.13)	1.78E−08	8.74E−01 (6 sets)	

[a]Average MAF and imputation R-square quality metric of Illumina, Affymetrix and OmniExpress datasets.

Table 3. Genetic markers with *P* value for interaction < 5×10^{-7} in gender-specific analysis

SNP	CHR:BP	A1	A2	Frequency for A1	Average imputation R-square[a]	NHS (female)		HPFS (male)		*P* for heterogeneity between genders	Mapped gene
						Beta	*P* value	Beta	*P* value		
NHS (female)											
NA	2:76738900:TTAGA	Deletion	Insertion	0.186	0.97	−0.27	2.51E−08	−0.02	6.77E−01	5.50E-04	LRRTM4
rs6694870	1:212728988	a	g	0.153	0.96	0.27	1.64E−07	−0.03	5.53E−01	5.50E-05	ATF3
rs191976747	10:15021717	a	t	0.011	0.63	−1.32	4.51E−07	−0.04	8.80E−01	5.48E-04	DCLRE1C
HPFS (male)											
rs77868414	8:43285229	c	g	0.034	0.38	0.06	7.19E−01	−0.97	3.43E−07	3.50E-05	POTEA

[a]Average imputation R-square quality metric of Illumina, Affymetrix and OmniExpress datasets.

functions. We also tested their individual and combined association with risk of BCC, but no significant association was found. Supplementary Table 7, available at *Carcinogenesis* Online, shows analytic methods and results for these additional analyses.

Discussion

Using data from the NHS and the HPFS, our group previously reported an inverse association between dietary caffeine intake and risk of BCC (7): compared with the lowest quintile, the highest quintile had the lowest risk (RR = 0.82 in women, 95%CI: 0.76–0.86; RR = 0.87 in men, 95% CI: 0.81–0.94; $P_{trend} < 0.0001$ in both). Here, we conducted a meta-analysis of genome-wide G–E interaction studies among participants in the NHS and HPFS. We determined that caffeine consumption was differentially associated with BCC risk according to genetic variation at SNP rs142310826, which is located on chromosome 4q34.3. We also found several loci that modify the caffeine–BCC association differently in men and women (*P* for heterogeneity between genders < 0.001): genetic markers at chromosome 2p12, 1q32.2 and 10p13 in women, and SNPs at chromosome 8p11.21 in men. These genetic markers and their mapped genes may prove important in BCC etiology, especially when caffeine consumption is considered.

UV-induced DNA damage in the skin can be caused by direct UV radiation or by indirect stress via reactive oxygen species (30). The SNP rs142310826 identified in our study is about 1Mb upstream of *NEIL3-AGA* locus. *NEIL3* encodes a DNA glycosylase that recognizes and removes lesions produced by oxidative stress, such as spiroiminodihydantoin (Sp), guanidinohydantoin (Gh) and 8-oxoguanine (8-oxoG), primarily in single-stranded DNA (ssDNA) (31). NEIL3 has been shown to be an important facilitator of cell proliferation in neural stem/progenitor cells and tumor cells, suggesting its possible role in replication-associated DNA repair (32,33). Some studies have reported that polymorphisms of DNA glycosylases may possess altered enzymatic activity, increasing the risk of inflammation-related cancers (34,35). Though our understanding of the mechanism of caffeine's inhibitory effect on BCC development is quite limited, previous studies suggested the involvement of the ATR-Chk1 signaling pathway (10–12). Similar to targets of NEIL3, ATR is primarily recruited to the chromosome at the site of ssDNA damage (36). Activation of ATR-Chk1 pathway normally results in cell-cycle checkpoint. Biological studies have demonstrated that caffeine could either directly disrupt the ATR-Chk1 pathway (12) or inhibit ATR-mediated DNA repair (37), thus promoting damaged cells to prematurely undergo apoptosis rather than to process through the cell cycle. We therefore propose that caffeine may inhibit ATR-mediated DNA repair by NEIL3, which may enhance the elimination of mutated basal cells in skin. In addition, caffeine could exert antioxidant effects that may neutralize oxidative stress in cells (38), which may decrease oxidative DNA damage and alter the behavior of related DNA repair genes, such as *NEIL3*. The other gene presented in the reported region is AGA, a gene encodes a amidase which is involved in the degradation of glycoproteins within lysosomes (39). However, no evidence has indicated its role in cancer development.

The genetic markers identified among women mapped to *LRRTM4*, *ATF3* and *DCLRE1C*. The leucine-rich repeat transmembrane neuronal 4 (LRRTM4) may play a role in the development and maintenance of the excitatory synapse in the vertebrate nervous system (40). According to the GWAS catalog (http://www.ebi.ac.uk/gwas/home), SNPs mapped on this gene are associated with phenotypes such as verbal declarative memory, sporadic amyotrophic lateral sclerosis, and immunoglobulin G (IgG) glycosylation. Changes in IgG glycosylation have been linked to gastric cancer and ovarian cancer in previous studies (41,42), indicating other possible functions of this gene. The activating transcription factor 3 (*ATF3*) gene has been demonstrated to play opposite roles (oncogene or tumor suppression) in cancer development depending on the cell type and context (43). Though upregulation of *ATF3* appears to enhance tumor formation in keratinocytes (44), Atf3 protein levels decreased when caffeine was administered in a mouse model (45). Moreover, ATF3 was discovered to be related to the ATR-Chk1 pathway as well (46). The DNA cross-link repair 1C (*DCLRE1C*) gene encodes a nuclear protein that has single-strand-specific exonuclease activity and also functions in the regulation of the cell cycle in response to DNA damage (47).

SNPs that showed the greatest interaction with caffeine in BCC development in males mapped to *POTEA*. *POTE* is a highly homologous gene family located on numerous chromosomes and expressed in a wide variety of human cancers (colon, lung, breast, ovary and pancreas) (48). In normal tissue, its expression is restricted to testis, ovary and prostate, with the highest expression in testis (49). Little is known about the biological function of this gene family, but there is evidence for its role in inducing programmed cell death (49).

We specifically extracted results for caffeine consumption-related loci identified by previous GWAS analysis, but none showed significant interaction with caffeine intake in relation to BCC risk. We additionally tested the individual and combined associations between caffeine SNPs and risk of BCC, but none reached statistical significance. These results suggest that the inverse association between caffeine intake and risk of BCC is not due to interaction or association with established caffeine-related loci.

Our study has several strengths. First, we used high-quality cohort data, among which information on both caffeine intake and genetic markers is available for studying G–E interactions. The relatively large sample size facilitated detection of potential interaction, even using a conventional logistic regression approach and a stringent genome-wide significance level. Second, we took gender difference into consideration in our analysis, because men and women may have different caffeine consumption habits, and caffeine may interact with sex hormones when exerting its biological effects (50). This study design helped identify several loci that are specific to men or women, despite only one of them has reached genome-wide significance. Because none of the gender-specific SNPs (or their mapped genes) identified in this study has been shown to related to sex hormone level in large bioinformatics databases or published literatures, we propose the observed heterogeneity may be explained by different caffeine consumption habits between genders. It is possible that some of the genes could only be turned on or off at certain levels of caffeine consumption. Further epidemiological studies of larger sample sizes and biological studies are needed to verify our findings and better illustrate the underlying mechanisms. Third, we modeled caffeine consumption in a different way for the purpose of sensitivity analysis, and the results did not change materially.

We also acknowledge some limitations. First, we used caffeine consumption in 1986 rather than cumulative average intake in our analysis. To study G–E interaction, environmental exposure should be measured at appropriate time points, because many genes are expressed only during specific developmental periods, and some exposures may have greater impact on specific stages. However, we understand very little about the biological mechanisms and induction period of caffeine's effects on the development of BCC. Given that the induction period of cancer is relatively long, and computing cumulative average caffeine intake among BCC controls is not straightforward, caffeine intake in 1986 is an acceptable option. Second, SNP rs142310826 is less common and of only moderate imputation quality in our datasets. However, when we directly genotyped this SNP among 335 participants in our GWAS datasets, the correlation between imputed dosage and directly genotyped allele count was 0.7 ($P < 0.0001$). Third, because all the participants in this study are health professionals of European ancestry, our results may not be generalizable to other ethnic or socioeconomic groups. Finally, we did not split our data into discovery and replication sets, because combined analysis across all studies is the most powerful analytical strategy (51). However, the results of the top SNP rs142310826 were largely consistent among the six datasets, which may reflect the reliability of our finding.

In conclusion, in this genome-wide G–E interaction meta-analysis, the association of caffeine intake with BCC risk differed according to genetic variation of SNP rs142310826. Genetic markers at chromosomes 2p12, 1q32.2, 10p13 and 8p11.21 modified the caffeine-BCC association differently in women and men. Further G–E interaction analyses are warranted to verify our findings, and additional biological studies are needed to better elucidate the roles of these genetic variants and their mapped genes.

Supplementary material

Supplementary Tables 1–7 and Figures 1–4 can be found at http://carcin.oxfordjournals.org/

Funding

This work is supported by NIH R01 CA49449, P01 CA87969, UM1 CA186107 R03 CA188568 and UM1 CA167552.

Acknowledgements

We are indebted to the participants in the NHS and HPFS for their dedication to this research. We thank the following state cancer registries for their help: Alabama, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Nebraska, New Hampshire, New Jersey, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Virginia, Washington and Wyoming. The authors assume full responsibility for analyses and interpretation of these data.
Conflict of Interest Statement: None declared.

References

1. Wong, C. et al. (2003) Basal cell carcinoma. BMJ, 327, 794.
2. Gallagher, R.P. et al. (1995) Sunlight exposure, pigmentary factors, and risk of nonmelanocytic skin cancer: I. Basal cell carcinoma. Arch. Dermatol., 131, 157.
3. Han, J. et al. (2006) Risk factors for skin cancers: a nested case-control study within the Nurses' Health Study. Int. J. Epidemiol., 35, 1514–1521.
4. Abel, E.L. et al. (2007) Daily coffee consumption and prevalence of nonmelanoma skin cancer in Caucasian women. Eur. J. Cancer Prev., 16, 446–452.
5. Rees, J.R. et al. (2007) Tea consumption and basal cell and squamous cell skin cancer: results of a case-control study. J. Am. Acad. Dermatol., 56, 781–785.
6. Stensvold, I. et al. (1994) Coffee and cancer: a prospective study of 43,000 Norwegian men and women. Cancer Causes Control, 5, 401–408.
7. Song, F. et al. (2012) Increased caffeine intake is associated with reduced risk of basal cell carcinoma of the skin. Cancer Res., 72, 3282–3289.
8. Lu, Y.-P. et al. (2002) Topical applications of caffeine or (–)-epigallocatechin gallate (EGCG) inhibit carcinogenesis and selectively increase apoptosis in UVB-induced skin tumors in mice. Proc. Natl. Acad. Sci., 99, 12455–12460.
9. Kerzendorfer, C. et al. (2009) UVB and caffeine: inhibiting the DNA damage response to protect against the adverse effects of UVB. J. Invest. Dermatol., 129, 1611–1613.
10. Han, W. et al. (2011) Caffeine promotes ultraviolet B-induced apoptosis in human keratinocytes without complete DNA repair. J. Biol. Chem., 286, 22825–22832.
11. Heffernan, T.P. et al. (2009) ATR-Chk1 pathway inhibition promotes apoptosis after UV treatment in primary human keratinocytes: potential basis for the UV protective effects of caffeine. J. Invest. Dermatol., 129, 1805–1815.
12. Lu, Y.P. et al. (2008) Effect of caffeine on the ATR/Chk1 pathway in the epidermis of UVB-irradiated mice. Cancer Res., 68, 2523–2529.
13. Stacey, S.N. et al. (2008) Common variants on 1p36 and 1q42 are associated with cutaneous basal cell carcinoma but not with melanoma or pigmentation traits. Nat. Genet., 40, 1313–1318.
14. Stacey, S.N. et al. (2009) New common variants affecting susceptibility to basal cell carcinoma. Nat. Genet., 41, 909–914.
15. Nan, H. et al. (2011) Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. Hum. Mol. Genet., 20, 3718–3724.
16. Nan, H. et al.; CCFR; GECCO. (2015) Association of aspirin and NSAID use with risk of colorectal cancer according to genetic variants. JAMA, 313, 1133–1142.
17. Hamza, T.H. et al. (2011) Genome-wide gene-environment study identifies glutamate receptor gene GRIN2A as a Parkinson's disease modifier gene via interaction with coffee. PLoS Genet., 7, e1002237.

18. Cornelis, M.C. et al. (2012) Gene-environment interactions in genome-wide association studies: a comparative study of tests applied to empirical studies of type 2 diabetes. Am. J. Epidemiol., 175, 191–202.
19. Rimm, E.B. et al. (1992) Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am. J. Epidemiol., 135, 1114–1127.
20. Feskanich, D. et al. (1993) Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. J. Am. Diet. Assoc., 93, 790–796.
21. Colditz, G.A. et al. (1986) Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. Am. J. Epidemiol., 123, 894–900.
22. Hunter, D.J. et al. (1990) Risk factors for basal cell carcinoma in a prospective cohort of women. Ann. Epidemiol., 1, 13–23.
23. Price, A.L. et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet., 38, 904–909.
24. Willer, C.J. et al. (2010) METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics, 26, 2190–2191.
25. Aulchenko, Y.S. et al. (2010) ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics, 11, 134.
26. Sulem, P. et al. (2011) Sequence variants at CYP1A1-CYP1A2 and AHR associate with coffee consumption. Hum. Mol. Genet., 20, 2071–2077.
27. Amin, N. et al.; kConFab Investigators. (2012) Genome-wide association analysis of coffee drinking suggests association with CYP1A1/CYP1A2 and NRCAM. Mol. Psychiatry, 17, 1116–1129.
28. Cornelis, M.C. et al. (2011) Genome-wide meta-analysis identifies regions on 7p21 (AHR) and 15q24 (CYP1A2) as determinants of habitual caffeine consumption. PLoS Genet., 7, e1002033.
29. Cornelis, M. et al. (2014) Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. Mol. Psychiatry. 20.5:647–665.
30. Ichihashi, M. et al. (2003) UV-induced skin damage. Toxicology, 189, 21–39.
31. Krokeide, S.Z. et al. (2013) Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimindiohydantoin and guanidinohydantoin. DNA Repair, 12, 1159–1164.
32. Liu, M. et al. (2012) Expression and purification of active mouse and human NEIL3 proteins. Protein Expr. Purif., 84, 130–139.
33. Liu, M. et al. (2013) Neil3, the final frontier for the DNA glycosylases that recognize oxidative damage. Mutat. Res., 743, 4–11.
34. Nohmi, T. et al. (2005) Modulation of oxidative mutagenesis and carcinogenesis by polymorphic forms of human DNA repair enzymes. Mutat. Res., 591, 60–73.
35. Tudek, B. (2007) Base excision repair modulation as a risk factor for human cancers. Mol. Aspects Med., 28, 258–275.
36. Batista, L.F. et al. (2009) How DNA lesions are turned into powerful killing structures: insights from UV-induced apoptosis. Mutat. Res., 681, 197–208.
37. Selby, C.P. et al. (1990) Molecular mechanisms of DNA repair inhibition by caffeine. Proc. Natl. Acad. Sci., 87, 3522–3525.
38. Devasagayam, T. et al. (1996) Caffeine as an antioxidant: inhibition of lipid peroxidation induced by reactive oxygen species. Biochimica et Biophysica Acta (BBA) Biomembr., 1282, 63–70.
39. Saito, S. et al. (2008) Structural basis of aspartylglucosaminuria. Biochem. Biophys. Res. Commun., 377, 1168–1172.
40. Song, Y.S. et al. (2013) Presynaptic proteoglycans: sweet organizers of synapse development. Neuron, 79, 609–611.
41. Saldova, R. et al. (2007) Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. Glycobiology, 17, 1344–1356.
42. Ruhaak, L.R. et al. (2015) The serum immunoglobulin G glycosylation signature of gastric cancer. EuPA Open Proteom., 6, 1–9.
43. Thompson, M.R. et al. (2009) ATF3 transcription factor and its emerging roles in immunity and cancer. J. Mol. Med., 87, 1053–1060.
44. Wu, X. et al. (2010) Opposing roles for calcineurin and ATF3 in squamous skin cancer. Nature, 465, 368–372.
45. Jia, H. et al. (2014) Coffee intake mitigated inflammation and obesity-induced insulin resistance in skeletal muscle of high-fat diet-induced obese mice. Genes Nutr., 9, 1–10.
46. Demidova, A.R. et al. (2009) Dual regulation of Cdc25A by Chk1 and p53-ATF3 in DNA replication checkpoint control. J Biol Chem., 284, 4132–4139.
47. Felgentreff, K. et al. (2015) Functional analysis of naturally occurring DCLRE1C mutations and correlation with the clinical phenotype of ARTEMIS deficiency. J. Allergy Clin. Immunol. 136.1:140–150.
48. Bera, T.K. et al. (2006) POTE paralogs are induced and differentially expressed in many cancers. Cancer Res., 66, 52–56.
49. Liu, X.F. et al. (2009) A primate-specific POTE-actin fusion protein plays a role in apoptosis. Apoptosis, 14, 1237–1244.
50. Ascherio, A. et al. (2003) Caffeine, postmenopausal estrogen, and risk of Parkinson's disease. Neurology, 60, 790–795.
51. Skol, A.D. et al. (2006) Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat. Genet., 38, 209–213.