Cancer Epidemiology, Biomarkers & Prevention

AACR

Association of Cancer Susceptibility Variants with Risk of Multiple Primary Cancers: The Population Architecture using Genomics and Epidemiology Study

S. Lani Park, Christian P. Caberto, Yi Lin, et al.

Updated version	Access the most recent version of this article at: doi:10.1158/1055-9965.EPI-14-0129
Supplementary Material	Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2014/08/21/1055-9965.EPI-14-0129.DC1.html

Cited Articles	This article cites by 97 articles, 19 of which you can access for free at: http://cebp.aacrjournals.org/content/23/11/2568.full.html#ref-list-1

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.

Published OnlineFirst August 19, 2014; DOI: 10.1158/1055-9965.EPI-14-0129

Cancer
Epidemiology,
Biomarkers
& Prevention

Research Article

Association of Cancer Susceptibility Variants with Risk of Multiple Primary Cancers: The Population Architecture using Genomics and Epidemiology Study

S. Lani Park[1], Christian P. Caberto[2], Yi Lin[3], Robert J. Goodloe[4], Logan Dumitrescu[4,5], Shelly-Ann Love[6], Tara C. Matise[7], Lucia A. Hindorff[8], Jay H. Fowke[9], Fredrick R. Schumacher[1], Jennifer Beebe-Dimmer[10,11], Chu Chen[3], Lifang Hou[12], Fridtjof Thomas[13], Ewa Deelman[14], Ying Han[1], Ulrike Peters[3], Kari E. North[6,15], Gerardo Heiss[6], Dana C. Crawford[16], Christopher A. Haiman[1], Lynne R. Wilkens[2], William S. Bush[16], Charles Kooperberg[3], Iona Cheng[17], and Loïc Le Marchand[2]

Abstract

Background: Multiple primary cancers account for approximately 16% of all incident cancers in the United States. Although genome-wide association studies (GWAS) have identified many common genetic variants associated with various cancer sites, no study has examined the association of these genetic variants with risk of multiple primary cancers (MPC).

Methods: As part of the National Human Genome Research Institute (NHGRI) Population Architecture using Genomics and Epidemiology (PAGE) study, we used data from the Multiethnic Cohort (MEC) and Women's Health Initiative (WHI). Incident MPC (IMPC) cases ($n = 1,385$) were defined as participants diagnosed with more than one incident cancer after cohort entry. Participants diagnosed with only one incident cancer after cohort entry with follow-up equal to or longer than IMPC cases served as controls (single-index cancer controls; $n = 9,626$). Fixed-effects meta-analyses of unconditional logistic regression analyses were used to evaluate the associations between 188 cancer risk variants and IMPC risk. To account for multiple comparisons, we used the false-positive report probability (FPRP) to determine statistical significance.

Results: A nicotine dependence–associated and lung cancer variant, *CHRNA3* rs578776 [OR, 1.16; 95% confidence interval (CI), 1.05–1.26; $P = 0.004$], and two breast cancer variants, *EMBP1* rs11249433 and *TOX3* rs3803662 (OR, 1.16; 95% CI, 1.04–1.28; $P = 0.005$ and OR, 1.13; 95% CI, 1.03–1.23; $P = 0.006$), were significantly associated with risk of IMPC. The associations for rs578776 and rs11249433 remained ($P < 0.05$) after removing subjects who had lung or breast cancers, respectively ($P \leq 0.046$). These associations did not show significant heterogeneity by smoking status ($P_{heterogeneity} \geq 0.53$).

Conclusions: Our study has identified rs578776 and rs11249433 as risk variants for IMPC.

Impact: These findings may help to identify genetic regions associated with IMPC risk. *Cancer Epidemiol Biomarkers Prev; 23(11); 2568–78. ©2014 AACR.*

Introduction

In the past 50 years, the 5-year relative survival rate for all cancers has dramatically increased from 49% to 67%

(1). As a result, the number of cancer survivors is growing by 2% per year (2). Compared with the general population, cancer survivors are at a greater risk for a number of

[1]Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, California. [2]Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii. [3]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington. [4]Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee. [5]Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee. [6]Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina. [7]Department of Statistics and Biostatistics, Rutgers University, Piscataway, New Jersey. [8]Division of Genomic Medicine, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland. [9]Vanderbilt Epidemiology Center, Vanderbilt University, Nashville, Tennessee. [10]School of Medicine, Wayne State University, Detroit, Michigan. [11]Karmanos Cancer Institute, Detroit, Michigan. [12]Department of Preventative Medicine, The Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois. [13]Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, Tennessee. [14]USC Information Sciences Institute, University of Southern California, Marina del Rey, California. [15]Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, California. [16]Department of Epidemiology and Biostatistics, Institute for Computational Biology, Case Western Reserve University, Cleveland, Ohio. [17]Cancer Prevention Institute of California, Fremont, California.

American Association for *Cancer Research*

comorbidities, including the occurrence of subsequent primary cancers (3), which is one of the leading causes of death for long-term cancer survivors (>5 years; refs. 4, 5). In 2002, Surveillance Epidemiology and End Results (SEER) data indicated that approximately 16% of all new cancers in the United States were second or higher-order primary cancers (1). Risk factors for developing multiple primary cancers (MPC) include younger age at index (first) cancer, treatment modality for the index cancer, shared etiologic factors (e.g., smoking), and genetic susceptibility (3).

Genome-wide association studies (GWAS) have identified several hundred genetic variants associated with various types of cancers (6). Pleiotropy, defined as when a single variant is associated with more than one phenotype, has been identified for some of these cancer risk loci (7, 8). For instance, variants in the chromosome 8q24 locus have been associated with risk of prostate, breast, colorectal, and other cancers (8) and those near the telomerase reverse transcriptase gene (*TERT*) have been associated with many cancer sites (9–13), including lung (14–16), estrogen receptor (ER)–negative breast (17), ovary (18), and pancreatic (19) cancers. Given that some cancer risk variants are associated with more than one cancer site, it is possible that they may also be associated with risk of developing MPC. Identification of such risk variants may elucidate common cancer etiologies and pathways, improve our understanding of treatment-related effects or identify populations at risk for developing multiple cancers.

With the exception of familial cancer syndromes, the literature examining genetic variants and risk of MPC is sparse, and is usually limited to risk after a specific index cancer site. Furthermore, prior studies are rarely based on prospective data and there has been no systematic investigation of the relationship between common cancer susceptibility variants and MPC risk. Thus, we used prospective data from the Population Architecture using Genomics and Epidemiology (PAGE) study to investigate the associations of 188 cancer risk variants with risk of incident MPC (IMPC).

Materials and Methods

Study populations

The PAGE study (20) was initiated in 2008 by the National Human Genome Research Institute (NHGRI). Two PAGE studies in collaboration with the NHGRI and the coordinating center (20) participated in this analysis, the Multiethnic Cohort (MEC; ref. 21) and the Women's Health Initiative (WHI; ref. 22). Informed consent was obtained from all study participants. This investigation was approved by each study's respective Institutional Review Boards.

Briefly, the MEC was initiated in 1993 to investigate the impact of dietary and environmental factors on major chronic diseases, particularly cancer, in ethnically diverse populations in Hawaii and California (21). The study

recruited 96,810 men and 118,441 women of ages 45 to 75 years between 1993 and 1996. MEC subjects recontacted mostly from 1995 to 2001 for blood collection included incident cases with breast, prostate, or colorectal cancers, as well as a random sample of cohort participants to serve as controls in genomic nested case–control studies (participation rate 72% and 63%, respectively). The median interval between diagnosis and blood draw was 14 months (interquartile range, 10–19 months). From 2001 to 2006, blood was also collected, prospectively, without regard for cancer diagnosis, from willing cohort participants (participation rate 43%). Incident cancers are identified through annual linkage to the Hawaii Tumor Registry, the Los Angeles County Cancer Surveillance Program, and the California Cancer Registry. All three registries are members of SEER (23). At the time of case selection for this analysis, follow-up was complete as to December 31, 2007.

The WHI is a long-term health study that recruited 161,808 postmenopausal women of ages 50 to 79 years between 1992 and 1998 at 40 clinical centers throughout the United States. WHI comprises a clinical trial arm, an observational study arm, and several extension studies, which are studies with continued follow-up of consenting participants from the original two study arms. The details of WHI have been previously described (22, 24), and are available online (25). Blood draw was taken at time of study recruitment on all study participants. During the follow-up period of January 1, 1992, to September 30, 2013, incident cancers were identified among participants in the observational study and clinical trial arm. Potential cancer status was identified from self-reported and/or hospital record data and confirmation of status was determined through review of the participant's medical records and pathology reports by trained cancer coders.

Case and control definitions

Details regarding the study design and case–control definitions are presented in Supplementary Table S1. All subjects for this analysis were cancer-free at baseline and had to have developed at least one cancer during study follow-up. MPC cases were defined as those who were diagnosed with more than one invasive primary cancer (nonmelanoma skin cancers were excluded) during study follow-up (26), and are referred herein as IMPC cases. Controls were defined as subjects who developed only one invasive primary cancer during study follow-up and are referred to as incident single-index cancer (ISC) controls. To test the association of cancer risk variants with risk of subsequent cancers, cancer-free controls were not used in this analysis. In the MEC, which identifies cancer cases by linkage to the SEER registries, IMPC cases could have a subsequent primary cancer of the same site, provided that the subsequent cancer met the SEER definition for same site MPCs [e.g., a subsequent breast cancer following a first breast cancer defined as primary given the following conditions: (i) contralateral breast cancer that did not result from metastases, (ii) breast cancer of a

different histologic cell type, or (iii) breast cancer 5+ years after initial diagnosis; ref. 26). No second primaries were allowed for prostate cancer, leukemia, and non-Hodgkin lymphoma, also based on the SEER rules. In the WHI, which adjudicated cancer cases by medical record review, it was not possible to systematically determine whether the cancer of the same site was a subsequent primary cancer or metastases. Thus, a repeat diagnosis of the same cancer site during follow-up was not counted as IMPC in WHI. IMPC cases for each study were frequency-matched to all eligible ISC controls by age of index cancer diagnosis (within 5-year intervals), length of follow-up (where controls had follow-up time that was equal to or greater than that of the cases), race, and sex. This analysis included a total of 1,385 IMPC cases and 9,626 ISC controls. In preliminary analyses, we had additionally matched for first cancer site. As the findings were similar, we matched only on age, sex, and race/ethnicity to increase our sample size and adjusted for index cancer in the statistical model.

SNP selection and genotyping

A total of 188 cancer risk variants associated with 18 cancers and nicotine dependence (nicotine dependence SNPs were considered herein as cancer risk variants given that smoking is a major risk factor for several cancer sites; ref. 27) were identified from the NHGRI catalog of GWAS studies (6) as of January 2010, followed by a review of the original reports and the fine mapping literature (12, 13–16, 19, 28–83). The risk allele for each SNP was determined on the basis of the prior literature and was defined as the allele associated with an increased risk of cancer in the first GWAS report.

Genotyping for the 188 cancer variants (175 SNPs in WHI and 156 SNPs in MEC; 143 SNPs with overlap) was performed using the TaqMan OpenArray platform (MEC) and Illumina BeadXpress (WHI). To control for population stratification, a panel of 128 ancestry informative markers (84) were genotyped. Principal components analysis (85) was performed and the main principal components were included in the regression model to adjust for genetic ancestry in each study.

Standard quality-assurance and quality-control measures were applied to ensure genotyping quality (20). Samples and SNPs were included on the basis of call rates ($\geq$90%), concordance of blinded replicates (>98%), and no clear departure from Hardy–Weinberg equilibrium ($P > 0.001$).

Statistical analysis

For each study, we tested the association between each SNP and risk of IMPC using unconditional logistic regression. SNPs were coded additively, with 0, 1, and 2 referring to the number of risk alleles. Models were adjusted for age at diagnosis for the index cancer (continuous), sex (MEC only), study design (WHI only: clinical trial vs. observational study), the most significant principal components of genetic ancestry [to account for race/ethnicity (five principal components for the MEC and three prin-

cipal components for WHI)], and index cancer site (17 categories), stage (local, regional, and distant), and diagnosis year (to account for treatment cohort effects). We adjusted for index cancer site and stage to account for differences in cancer survival.

The regression estimates were combined across studies using inverse-variance weighted, fixed-effect meta-analysis as implemented in METAL (http://www.sph.umich.edu/csg/abecasis/Metal/). We calculated the cross-study and cross-race heterogeneity P values based on Cochran Q statistic. Stratified analysis by race/ethnicity was conducted for SNP associations with $P < 0.05$. To account for the multiple testing of 188 SNPs, we used the false-positive report probability (FPRP) introduced by Wacholder and colleagues (86). We set a stringent FPRP threshold of 0.20 and assigned a prior probability range of 0.01 to 0.10 to detect an OR of 1.2 or 0.83. Variants that were found to be significantly associated with IMPC after correction for multiple comparisons were further investigated in sensitivity analyses in which subjects with the cancer site (either as index cancer or IMPC) corresponding to the variant's first known GWAS association were removed, as well as after adjusting for or stratifying by smoking status, obtained by self-report at baseline.

Results

The main characteristics of the 1,385 IMPC cases and 9,626 ISC controls are presented in Table 1. Subjects were of six different ethnic/racial populations: European American, African American, Hispanic, Asian, Pacific Islander, and Native American. Forty-three percent of subjects were European American and 59% were female (the MEC ascertained both sexes, while WHI ascertained only women). The most common index cancer sites were prostate and breast for the MEC and colorectal and breast for WHI. In both studies, among all IMPC cases, breast cancer was the most common incident second cancer (for the MEC, $n = 164$ and for WHI, $n = 133$). The distribution of index and second (subsequent) incident cancers by study is presented in Table 2 and Supplementary Table S2 (which includes latency intervals between cancer sites). The most common index and subsequent cancer combination in the MEC study was an index breast and subsequent breast cancers and in the WHI study index endometrial and subsequent breast cancers. Less common cancers were collapsed as "other"; the enumeration of these cancer sites can be found in Supplementary Table S3 (WHI) and Supplementary Table S4 (MEC). There were 115 MEC IMPC cases and 15 WHI IMPC cases with three or more cancers (Supplementary Table S5).

We examined a total of 188 cancer risk variants (Table 3 for associations with $P < 0.05$ and Supplementary Table S6 for associations with $P \geq 0.05$). Prostate and breast cancer risk variants represented the greatest proportions of variants (28% and 15%, respectively). Among the 188 SNPs, 12 were found to be associated with risk of IMPC at $P < 0.05$ (Table 3). These 12 associations were with three

Table 1. Study characteristics for IMPC cases and first index cancer controls

	MEC		WHI	
	Cases	**Controls**	**Cases**	**Controls**
	($n = 965$)	**($n = 6,969$)**	**($n = 420$)**	**($n = 2,657$)**
Age of index cancer; mean (SD)	68.3 (7.8)	68.0 (8.3)	70.1 (7.0)	69.7 (9.0)
Age of second cancer; mean (SD)	72.6 (8.0)	NA	74.2 (7.3)	NA
Follow-up time, y; mean (SD)	9.4 (3.3)	13.7 (1.9)	9.6 (3.7)	12.7 (2.3)
Sex; n (%)				
Male	549 (56.9)	3,930 (56.4)		
Female	416 (43.1)	3,039 (43.6)	420 (100)	2,657 (100)
Race/ethnicity; n (%)				
European American	254 (26.3)	1,534 (22.0)	372 (88.6)	2,088 (78.6)
African American	235 (24.4)	1,419 (20.4)	37 (8.8)	412 (15.5)
Hispanic	194 (20.1)	1,581 (22.7)	6 (1.4)	96 (3.6)
Asian	232 (24.0)	2,045 (29.3)	4 (0.95)[b]	58 (2.2)[b]
Pacific Islander	50 (5.2)	390 (5.6)		
Indian/Native American	NA	NA	1 (0.24)	3 (0.11)
Smoking status; n (%)				
Never	352 (36.9)	2,843 (41.4)	174 (41.4)	1,292 (48.6)
Former	421 (44.1)	3,064 (44.6)	200 (47.6)	1,123 (42.3)
Current	182 (19.1)	959 (14.0)	42 (10.0)	212 (8.0)
Missing, n	10	103	4	30
Pack-years[a]; mean (SD)	21.9 (17.0)	18.0 (15.7)	15.3 (22.0)	11.0 (19.1)
Missing, n	31	260	14	90
Index cancer site; n (%)				
Breast	265 (27.5)	1,970 (28.3)	82 (19.5)	906 (34.1)
Prostate	278 (28.8)	2,906 (41.7)	NA	NA
Colorectal	164 (17.0)	987 (14.2)	67 (16.0)	433 (16.3)
Lung	24 (2.5)	175 (2.5)	43 (10.2)	291 (11.0)
Other[c]	234 (24.2)	931 (13.4)	228 (54.3)	1,027 (38.7)
Secondary cancer site; n (%)				
Breast	164 (17.0)	NA	133 (31.7)	NA
Prostate	111 (11.5)	NA	NA	NA
Colorectal	158 (16.4)	NA	49 (11.7)	NA
Lung	120 (12.4)	NA	65 (15.5)	NA
Other[c]	412 (42.7)	NA	173 (41.2)	NA
Stage of index cancer; n (%)				
Localized	683 (70.8)	5,153 (73.9)	288 (68.6)	1,651 (61.1)
Regional	215 (22.3)	1,346 (19.3)	86 (20.1)	640 (24.1)
Distant	67 (6.9)	470 (6.7)	46 (11.0)	366 (13.8)
Diagnosis year for index cancer, mean (SD)	1998 (3.4)	2000 (3.7)	2002 (3.5)	2003 (3.7)
Diagnosis year for second cancer, mean (SD)	2002 (3.2)	NA	2006 (3.7)	NA

Abbreviation: NA, not applicable.
[a]Pack-years for ever smokers only.
[b]Counts for Asian and Pacific Islanders.
[c]The four most common other cancer sites include bladder, endometrium, leukemia, melanoma, and non-Hodgkin lymphoma; enumeration for these sites can be found in Table 2.

prostate cancer variants, three lung cancer or nicotine dependence variants, two breast cancer variants, one pancreatic cancer variant, one melanoma variant, one thyroid variant, and one colorectal variant. These associations were not found to be heterogeneous by study or race/ethnicity ($P \geq 0.05$; Table 3 and Supplementary Table S7).

After correction for multiple comparisons, we found that the associations of *CHRNA3* rs578776, *EMBP1* rs11249433, and *TOX3* rs3803662 with IMPC remained statistically significant at a FPRP threshold of 0.20 and a prior probability of 0.01 (Table 4). The *CHRNA3* rs578776 risk allele "C" previously associated with nicotine

2571

Table 2. Distribution of incident index and second cancer site among IMPC cases, stratified by study

Index cancer site	Breast MEC	Breast WHI	Colorectal MEC	Colorectal WHI	Lung MEC	Lung WHI	Prostate MEC	Prostate WHI	Urinary bladder MEC	Urinary bladder WHI	Melanoma MEC	Melanoma WHI	Kidney MEC	Kidney WHI	Non-Hodgkin lymphoma MEC	Non-Hodgkin lymphoma WHI	Endometrial MEC	Endometrial WHI	Leukemia MEC	Leukemia WHI	Other[a] MEC	Other[a] WHI	Total MEC	Total WHI	ALL
Breast	111	0	21	16	25	17	0	NA	2	3	6	2	4	2	7	6	30	7	16	8	43	21	265	82	347
Colorectal	15	29	58	0	7	10	36	NA	2	2	4	3	4	4	7	3	3	1	7	0	21	15	164	67	231
Lung	3	16	4	3	6	0	6	NA	1	0	0	1	0	0	1	3	0	4	0	3	3	13	24	43	67
Prostate	2	NA	45	NA	62	NA	0	NA	10	NA	14	NA	17	NA	17	NA	0	NA	29	NA	82	NA	278	0	278
Urinary bladder	1	1	2	1	3	0	18	NA	0	0	0	0	1	0	0	0	0	NA	0	0	0	0	25	2	27
Melanoma	3	16	2	4	1	6	9	NA	1	1	16	0	1	2	0	2	1	5	2	3	2	6	38	45	83
Kidney	2	3	1	2	1	1	5	NA	2	0	1	0	4	0	2	0	0	0	2	0	2	1	22	7	29
Non-Hodgkin lymphoma	4	15	3	5	2	5	6	NA	1	3	0	2	1	2	0	0	0	2	1	2	4	4	23	40	63
Endometrial	7	30	5	8	1	13	6	NA	0	1	0	2	0	0	1	5	2	0	1	2	6	14	23	75	98
Leukemia	3	2	3	6	1	3	7	NA	0	0	1	0	0	1	2	1	0	0	0	0	2	2	20	13	33
Other[a]	13	21	14	4	11	10	24	NA	0	1	1	1	2	1	5	5	1	1	2	1	10	1	83	46	129
Total	164	133	158	49	120	65	111	0	19	11	44	11	35	11	42	25	37	20	60	19	175	76	965	420	1,385

[a]Detailed cancer distribution for other cancers can be found in Supplementary Table S3 (WHI) and Supplementary Table S4 (MEC).

dependence and lung cancer was associated with an increased risk of IMPC [OR, 1.15; 95% confidence interval (CI) = 1.05–1.26; P = 0.004; Table 2]. The risk allele "C" for the breast cancer variant *EMBP1* rs11249433 was also associated with an increased risk of IMPC (OR, 1.16; 95% CI, 1.04–1.28; P = 0.005). The risk allele "T" for the breast cancer variant *TOX3* rs11249433, located at 16q12.1, was also associated with an increased risk of IMPC (OR, 1.13; 95% CI, 1.03–1.23; P = 0.006).

When removing from the analysis all subjects with lung cancer for the *CHRNA3* variant and with breast cancer for the *EMBP1* and *TOX3* variants, either as index cancer or IMPC, the associations for *CHRNA3* rs578776 and *EMBP1* rs11249433 remained significant ($P \leq 0.046$; Table 5). The association for rs3803662 had a similar OR as its main effects; however, the association did not quite reach statistically significant (OR, 1.13; P = 0.005 vs. OR, 1.11; P = 0.063, respectively).

Because smoking is a risk factor for many cancers, we additionally adjusted our models for smoking status and pack-years and found similar associations (data not shown). In our analyses stratified by smoking status, no heterogeneity in the effects of the genetic variants was detected between never and ever smokers ($P_{\text{heterogeneity}} \geq$ 0.53; Table 5). Because of a limitation in sample size, we were unable to restrict this analysis to cancers that are unrelated to smoking.

Discussion

To our knowledge, this is the first study to examine the association of cancer risk variants with risk of IMPC. We tested 188 established cancer risk variants among 1,385 IMPC cases and 9,626 ISC controls and found that the lung cancer/nicotine dependence risk allele "C" of *CHRNA3* rs578776 and the breast cancer risk alleles "C" of *EMBP1* rs11249433 and "T" of *TOX3* of rs3803662 were associated with an increased risk of IMPC. The associations for rs578776 and rs11249433 remained after accounting for multiple hypothesis testing, were not restricted to participants with the cancer previously associated with those variants, and were consistent with regard to smoking status.

The 15q25.1 region includes the *CHRNA5–CHRNA3–CHRNB4* cluster of cholinergic nicotine receptor subunit genes. Variants within this region have been associated with increased risk of lung cancer primarily due to their relationships with smoking behavior and nicotine dependence (14, 80, 81, 87). In addition, risk variants in the *CHRNA5–CHRNA3–CHRNB4* region have been positively associated with chronic obstructive pulmonary disease (88), serum albumin levels (89), pulmonary function (90), and childhood obesity–related traits in Hispanics (91). In our study, we found that both rs578776 and rs8042374, which are in high linkage disequilibrium (LD) in whites and Asians [Utah residents with ancestry from northern and western Europe (CEU): r^2 = 0.83, Yoruba in Ibadan, Nigeria (YRI): r^2 = 0.16, Japanese in Tokyo, Japan and Han

Table 3. Meta-analysis of the significant associations ($P < 0.05$) between risk variants for cancer and IMPC risk

SNP	First GWAS citation	First GWAS site	Region	Position	Mapped gene	Ca	Co	Risk allele/ ref. allele	RAF	Meta-OR[a] (95% CI)	P	Direction[b]	WHI MEC Direction[b]	P_{het} for study
rs578776	Saccone et al., 2009 (83)	Nicotine dependence	15q25.1	76,675,455	CHRNA3	1,377	9,559	C/T	0.49	1.15 (1.05–1.26)	0.004	+	+	0.26
rs11249433	Thomas et al., 2009 (68)	Breast cancer	1p11.2	120,982,136	EMBP1	1,347	9,406	C/T	0.26	1.16 (1.04–1.28)	0.005	+	+	0.59
rs3803662	Thomas et al., 2009 (68)	Breast cancer	16q12.1	51,143,842	TOX3	1,377	9,577	T/C	0.40	1.13 (1.03–1.23)	0.006	+	+	0.50
rs4975616	Broderick et al., 2009 (15)	Lung cancer	5p15.33	1,368,660	Intergenic	1,378	9,576	A/G	0.61	0.89 (0.82–0.98)	0.014	–	–	0.55
rs8042374	Wang et al., 2008 (14)	Lung cancer	15q25.1	76,695,087	CHRNA3	1,368	9,550	A/G	0.57	1.12 (1.02–1.24)	0.015	+	+	0.28
rs4857841	Gudmundsson et al., 2009 (46)	Prostate cancer	3q21.3	129,529,333	EEFSEC	1,376	9,534	A/G	0.44	0.90 (0.82–0.98)	0.020	–	–	0.34
rs4785763	Bishop et al., 2009 (31)	Melanoma	16q24.3	88,594,437	AFG3L1P	1,370	9,551	A/C	0.28	1.11 (1.01–1.22)	0.023	+	+	0.62
rs965513	Gudmundsson et al., 2009 (46)	Thyroid cancer	9q22.33	99,595,930	Intergenic	1,377	9,562	A/G	0.24	0.89 (0.80–0.99)	0.026	–	–	0.42
rs5945619	Eeles et al., 2008 (42)	Prostate cancer	Xp11.22	51,258,412	Intergenic	1,378	9,566	C/T	0.27	1.09 (1.01–1.18)	0.035	+	+	0.73
rs3790844	Petersen et al., 2010 (19)	Pancreatic cancer	1q32.1	198,274,055	NR5A2	420	2,657	T/C	0.77	1.23 (1.01–1.49)	0.035	+	×	1.00
rs10086908	Al Olama et al., 2009 (29)	Prostate cancer	8q24.21	128,081,119	Intergenic	957	6,904	T/C	0.74	1.13 (1.01–1.27)	0.040	×	+	1.00
rs4939827	Tenesa et al., 2008 (67)	Colorectal cancer	18q21.1	44,707,461	SMAD7	1,374	9,532	T/C	0.39	1.09 (1.00–1.19)	0.047	+	+	0.18

Abbreviation: RAF, risk allele frequency.

[a]Models were adjusted for age at diagnosis for the index cancer (continuous), sex (MEC only), study design (WHI only: clinical trial vs. observational study), the most significant principal components.

[b]+, positive association; –, negative association; ×, not tested.

Table 4. FPRP values for the 10 SNPs associated with IMPC risk

SNP	First GWAS site	Meta-OR (95% CI)[a]	Power[b]	Reported P	FPRP based on prior: 0.10	0.01	0.001
rs578776	Nicotine dependence	1.15 (1.05–1.26)	0.82	0.004	**0.03**	**0.25**	0.77
rs11249433	Breast cancer	1.16 (1.04–1.28)	0.75	0.005	**0.04**	**0.29**	0.81
rs3803662	Breast cancer	1.13 (1.03–1.23)	0.92	0.006	**0.44**	**0.34**	0.84
rs4975616	Lung cancer	0.89 (0.82–0.98)	0.93	0.014	**0.15**	0.66	0.95
rs8042374	Lung cancer	1.12 (1.02–1.24)	0.91	0.015	**0.12**	0.61	0.94
rs4857841	Prostate cancer	0.90 (0.82–0.98)	0.96	0.020	**0.13**	0.61	0.94
rs4785763	Melanoma	1.11 (1.01–1.22)	0.95	0.023	**0.22**	0.76	0.97
rs965513	Thyroid cancer	0.89 (0.80–0.99)	0.89	0.026	**0.24**	0.78	0.95
rs5945619	Prostate cancer	1.09 (1.01–1.18)	0.99	0.035	**0.21**	0.75	0.97
rs3790844	Pancreatic cancer	1.23 (1.01–1.49)	0.40	0.035	**0.44**	0.89	0.99
rs10086908	Prostate cancer	1.13 (1.01–1.27)	0.84	0.040	**0.30**	0.83	0.98
rs4939827	Colorectal cancer	1.09 (1.00–1.19)	0.98	0.047	**0.33**	0.85	0.98

NOTE: The values in bold indicate that the associations remained significant at an FPRP threshold of 0.20 and a prior probability range of 0.1 through 0.001 of detecting an OR of 1.2 or 0.83.
[a]ORs from Table 2, models were adjusted for age at diagnosis for the index cancer (continuous), sex (MEC only), study design (WHI only: clinical trial vs. observational study), the most significant principal components of genetic ancestry (five principal components for the MEC and three principal components for WHI), index cancer site, stage, and diagnosis year (to account for treatment cohort effects).
[b]Statistical power calculated using the recessive model, except where noted, is the power to detect an OR of 1.2 or 0.83 at a level of 0.05.

Chinese in Beijing, China (JPT+CHB): $r^2 = 0.72$], were positively associated with risk of IMPCs when compared with those with only one incident cancer. Only the association with rs578776 remained significant after correction for multiple testing. Smoking is an established risk factor for at least 10 different cancer sites (27); therefore, a reasonable explanation for the association of these variants with IMPC would be the result of their known association with smoking. However, when stratifying by smoking status, we found that the association was somewhat stronger, albeit not significantly ($P_{\text{heterogeneity}}$ by

smoking ≥ 0.47), in never smokers (OR, 1.21; $P = 0.01$) than in ever smokers (OR, 1.13; $P = 0.05$). This association in never smokers is unexpected and could be due to chance. Further study of this risk variant with IMPC in never smokers is warranted.

The rs11249433 variant is located in the pericentromeric region at 1p11.2 and within the embigin pseudogene, *EMBP1*. This SNP maps near many other pseudogenes in a SNP desert region. *EMBP1* rs11249433 was originally found to be associated with breast cancer ($P = 7 \times 10^{-10}$; ref. 68). Additional studies found that rs11249433 was

Table 5. Associations of two cancer risk variants with IMPC risk, after subjects with the cancer site corresponding to the known association have been removed, or after stratifying by smoking status

SNP	Breast cancer SNP rs11249433				Breast cancer SNP rs3803662				Lung cancer SNP rs578776			
	Cases	Control	OR (95% CI)[b]	P	Cases	Control	OR (95% CI)[b]	P	Cases	Control	OR (95% CI)[b]	P
Cancer site removed[a]	822	6,596	1.15 (1.00–1.31)	**0.046**	841	6,712	1.11 (0.99–1.24)	0.063	1,117	9,072	1.18 (1.06–1.30)	**0.0015**
Stratify by smoking[c]												
Never smokers	509	4,027	1.11 (0.94–1.32)	0.20	523	4,111	1.10 (0.95–1.26)	0.20	524	4,118	1.16 (1.00–1.36)	**0.049**
Ever smokers	796	5,034	1.19 (1.04–1.36)	**0.01**	809	5,117	1.15 (1.03–1.29)	0.02	808	5,093	1.12 (0.99–1.26)	0.06
$P_{\text{heterogeneity}}$ by smoking status			0.53				0.63				0.73	

[a]Subjects with the cancer site corresponding to the known association with cancer have been removed.
[b]Models were adjusted for age at diagnosis for the index cancer (continuous), sex (MEC only), study design (WHI only: clinical trial vs. observational study), the most significant principal components.
[c]Additionally adjusted for pack-years.

specifically associated with ER-positive breast cancer (68, 92) and lower-grade breast tumors (92), two characteristics that are associated with improved breast cancer survival (93). It is possible that the association for rs11249433 with increased risk of MPC may be related to survivorship such that a longer survival after breast cancer diagnosis may lead to a higher likelihood of developing a subsequent primary cancer. When removing subjects with breast cancer as the index cancer or an IMPC from our analysis, the association of rs11249433 with IMPC persisted, although not as significant, suggesting that this association was not exclusively due to a better breast cancer survival. This SNP may play a role in increased overall cancer survival or risk of cancer development. The closest located genes to rs11249433 are a transmembrane coding gene, NOTCH2, and highly paralogous low-affinity Fc gamma receptor family 1B gene, FCGR1B. SNPs in NOTCH2 have been associated with type II diabetes (94), a condition that has been associated with an increased risk of breast, colorectal, pancreatic, liver, urinary tract, and endometrial cancers (95, 96). In addition, a study that investigated the expression of five genes within 1 Mb of rs11249433 found an increased expression of NOTCH2 in patients with ER-positive breast cancer without TP53 mutations compared with those with TP53 mutations (97). NOTCH2 receptors are involved in regulation of cell communication, proliferation, differentiation, and death (98), processes that are all influential in cancer development.

The rs3803662 variant is located at 16q12.1 in the Cancer Susceptibility Candidate 16 gene, CASC16, which is an RNA gene. TOX3 and LOC643714 are other genes located close to this polymorphism. This variant has been previously associated with breast cancer ($P = 6 \times 10^{-19}$; ref. 64), with slightly stronger associations in ER-positive breast cancer than ER-negative disease (99). A prior study found that the risk allele for rs3803662 is correlated with lower mRNA expression of TOX3 in ER-positive tumors (100). Increased TOX3 mRNA expression has been previously found to be predictive of breast cancer metastases (101) and lower overall survival among patients with breast cancer (100). It is possible that the association for rs3803662 with increased risk of MPC may be related to survival from breast cancer. When removing subjects with breast cancer, the association was not quite statistically significant, suggesting that this association may be due to the risk variants prior association with breast cancer.

This study had a number of strengths, including the prospective design that allowed us to focus on incident cancers and minimize survival bias, while investigating MPC risk in an adult population. We also had a long follow-up period to assess subsequent cancers (MEC, 14 years and WHI, 19 years), a large sample size, and well-characterized study populations enabling adjustment for multiple potential confounders. Study limitations include limited power to detect effects for rare SNPs and the lack of detailed information on treatment, especially, on radiotherapy, a known risk factor for multiple primaries (102).

In the MEC, where information on first course of cancer therapy was available, similar results were observed after adjusting for radiation treatment (yes/no). In addition, our meta-analyzed findings did not change when adjusting for index cancer stage, suggesting that our findings are not strongly dependent on the treatment received, because treatment is relatively well standardized by stage. Our study oversampled the more common cancer sites as a result of the blood collection design in the earlier years of the MEC. Also, for WHI, the study undersampled breast and colorectal cancer cases to be genotyped for the PAGE analysis. Therefore, the more common cancer sites, with the exception of lung cancer, may have been overrepresented in the MEC and endometrial cancer was overrepresented in WHI. In addition, cases with subsequent primary cancers of the same site were not available in the WHI study and were, thus, underrepresented in that study.

In conclusion, our findings, if reproduced, may contribute toward identifying common cancer etiologic pathways, common treatment-related effects, or populations at risk for developing more than one cancer. We found the cancer risk variant rs578776 in CHRNA3 to be associated with occurrence of MPCs. Our analysis suggests that the associations for TOX3 rs3803662 may have been driven by the variants known association with breast cancer. However, the association for EMBP1 rs11249433 and CHRNA3 rs578776 appears to be independent of their known cancer association. These findings should be confirmed in other study populations.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Disclaimer

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Authors' Contributions

Conception and design: S.L. Park, F.R. Schumacher, U. Peters, L.R. Wilkens, C. Kooperberg, L. Le Marchand
Development of methodology: S.L. Park, C. Kooperberg
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): F.R. Schumacher, U. Peters, G. Heiss, D.C. Crawford, L.R. Wilkens, W.S. Bush, C. Kooperberg, L. Le Marchand
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S.L. Park, Y. Lin, R.J. Goodloe, L. Dumitrescu, J.H. Fowke, L. Hou, Y. Han, L.R. Wilkens, W.S. Bush, I. Cheng, L. Le Marchand
Writing, review, and/or revision of the manuscript: S.L. Park, L. Dumitrescu, T.C. Matise, L.A. Hindorff, J.H. Fowke, F.R. Schumacher, J. Beebe-Dimmer, C. Chen, L. Hou, F. Thomas, Y. Han, U. Peters, K.E. North, D.C. Crawford, C.A. Haiman, L.R. Wilkens, W.S. Bush, C. Kooperberg, I. Cheng, L. Le Marchand
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S.L. Park, C.P. Caberto, S.-A. Love, T.C. Matise, L.A. Hindorff, E. Deelman, K.E. North, G. Heiss, L. Le Marchand
Study supervision: K.E. North, G. Heiss, D.C. Crawford, C. Kooperberg, L. Le Marchand

References

1. Ries LAG, Melbert D, Krapcho M, Mariotto A, Miller BA, Feuer EJ, et al. SEER cancer statistics review, 1975–2004. Bethesda, MD: National Cancer Institute.
2. Cancer survivors: living longer, and now, better. The Lancet 2004; 364:2153–4.
3. Travis LB, Rabkin CS, Brown LM, Allan JM, Alter BP, Ambrosone CB, et al. Cancer survivorship—genetic susceptibility and second primary cancers: research strategies and recommendations. J Natl Cancer Inst 2006;98:15–25.
4. Aleman BMP, van den Belt-Dusebout AW, Klokman WJ, van't Veer MB, Bartelink H, van Leeuwen FE. Long-term cause-specific mortality of patients treated for Hodgkin's Disease. J Clin Oncol 2003;21:3431–9.
5. Reulen RC, Winter DL, Frobisher C, Lancashire ER, Stiller CA, Jenney ME, et al. Long-term cause-specific mortality among survivors of childhood cancer. JAMA 2010;304:172–9.
6. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 2009;106:9362–7.
7. Mocellin S, Verdi D, Pooley KA, Landi MT, Egan KM, Baird DM, et al. Telomerase reverse transcriptase locus polymorphisms and cancer risk: a field synopsis and meta-analysis. J Natl Cancer Inst 2012; 104:840–54.
8. Ghoussaini M, Song H, Koessler T, AlOlama AA, Kote-Jarai Z, Driver KE, et al. Multiple loci with different cancer specificities within the 8q24 gene desert. J Natl Cancer Inst 2008;100:962–6.
9. Rothman N, Garcia-Closas M, Chatterjee N, Malats N, Wu X, Figueroa JD, et al. A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. Nat Genet 2010;42: 978–84.
10. Kote-Jarai Z, Olama AA, Giles GG, Severi G, Schleutker J, Weischer M, et al. Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study. Nat Genet 2011;43:785–91.
11. Turnbull C, Rapley EA, Seal S, Pernet D, Renwick A, Hughes D, et al. Variants near DMRT1, TERT and ATF7IP are associated with testicular germ cell cancer. Nat Genet 2010;42:604–7.
12. Stacey SN, Sulem P, Masson G, Gudjonsson SA, Thorleifsson G, Jakobsdottir M, et al. New common variants affecting susceptibility to basal cell carcinoma. Nat Genet 2009;41:909–14.
13. Shete S, Hosking FJ, Robertson LB, Dobbins SE, Sanson M, Malmer B, et al. Genome-wide association study identifies five susceptibility loci for glioma. Nat Genet 2009;41:899–904.
14. Wang Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakidou A, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet 2008;40:1407–9.
15. Broderick P, Wang Y, Vijayakrishnan J, Matakidou A, Spitz MR, Eisen T, et al. Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. Cancer Res 2009;69: 6633–41.
16. Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet 2009;85:679–91.
17. Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, Millikan RC, et al. A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. Nat Genet 2011;43:1210–4.
18. Beesley J, Pickett HA, Johnatty SE, Dunning AM, Chen X, Li J, et al. Functional polymorphisms in the *TERT* promoter are associated with risk of serous epithelial ovarian and breast cancers. PLoS ONE 2011;6:e24987.
19. Petersen GM, Amundadottir L, Fuchs CS, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet 2010;42: 224–8.
20. Matise TC, Ambite JL, Buyske S, Carlson CS, Cole SA, Crawford DC, et al. The Next PAGE in understanding complex traits: design for the analysis of population architecture using genetics and epidemiology (PAGE) study. Am J Epidemiol 2011;174:849–59.
21. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J 2000;151:346–57.
22. Design of the women's health initiative clinical trial and observational study. Control Clin Trials 1998;19:61–109.
23. SEER. Surveillance, Epidemiology, and End Results. 2009. [cited 2014 Jan. 14]. Available from: http://seer.cancer.gov/.
24. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, et al. The women's health initiative recruitment methods and results. Ann Epidemiol 2003;13:S18–77.

25. Women's Health Initiative. [cited 2014 Jan. 14]. Available from: http://www.whi.org/.

26. 2007 Multiple Primary and Histology Coding Rules. Surveillance Research Program. Bethesda, MD: National Cancer Institute. 2013 [cited 2013 Jan 1]. Available from: http://seer.cancer.gov/tools/mphrules/.

27. IARC. Monographs on the evaluation of carcinogenic risks to humans. Tobacco smoke and involuntary smoking. Lyon, France: IARC; 2004.

28. Ahmed S, Thomas G, Ghoussaini M, Healey CS, Humphreys MK, Platte R, et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 2009;41:585–90.

29. Al Olama AA, Kote-Jarai Z, Giles GG, Guy M, Morrison J, Severi G, et al. Multiple loci on 8q24 associated with prostate cancer susceptibility. Nat Genet 2009;41:1058–60.

30. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, Fuchs CS, Petersen GM, Arslan AA, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. Nat Genet 2009;41:986–90.

31. Bishop DT, Demenais F, Iles MM, Harland M, Taylor JC, Corda E, et al. Genome-wide association study identifies three loci associated with melanoma risk. Nat Genet 2009;41:920–5.

32. Brown KM, Macgregor S, Montgomery GW, Craig DW, Zhao ZZ, Iyadurai K, et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. Nat Genet 2008;40:838–40.

33. Capasso M, Devoto M, Hou C, Asgharzadeh S, Glessner JT, Attiyeh EF, et al. Common variations in BARD1 influence susceptibility to high-risk neuroblastoma. Nat Genet 2009;41:718–23.

34. Conde L, Halperin E, Akers NK, Brown KM, Smedby KE, Rothman N, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nat Genet 2010;42:661–4.

35. Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, Pooley KA, et al. A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet 2007;39:352–8.

36. Crowther-Swanepoel D, Broderick P, Di Bernardo MC, Dobbins SE, Torres M, Mansouri M, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. Nat Genet 2010;42:132–6.

37. Cui R, Kamatani Y, Takahashi A, Usami M, Hosono N, Kawaguchi T, et al. Functional variants in ADH1B and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk. Gastroenterology 2009;137:1768–75.

38. Di Bernardo MC, Crowther-Swanepoel D, Broderick P, Webb E, Sellick G, Wild R, et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet 2008;40:1204–10.

39. Duggan D, Zheng SL, Knowlton M, Benitez D, Dimitrov L, Wiklund F, et al. Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. J Natl Cancer Inst 2007;99:1836–44.

40. Easton DF, Pooley KA, Dunning AM, Pharoah PDP, Thompson D, Ballinger DG, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007;447:1087–93.

41. Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 2009;41:1116–21.

42. Eeles RA, Kote-Jarai Z, Giles GG, Olama AA, Guy M, Jugurnauth SK, et al. Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet 2008;40:316–21.

43. Falchi M, Bataille V, Hayward NK, Duffy DL, Bishop JA, Pastinen T, et al. Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. Nat Genet 2009;41:915–9.

44. Fernandez LP, Milne RL, Pita G, Aviles JA, Lazaro P, Benitez J, et al. SLC45A2: a novel malignant melanoma-associated gene. Hum Mutat 2008;29:1161–7.

45. Gold B, Kirchhoff T, Stefanov S, Lautenberger J, Viale A, Garber J, et al. Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. Proc Natl Acad Sci U S A 2008;105:4340–5.

46. Gudmundsson J, Sulem P, Gudbjartsson DF, Blondal T, Gylfason A, Agnarsson BA, et al. Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. Nat Genet 2009;41:1122–6.

47. Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 2007;39:631–7.

48. Haiman CA, Patterson N, Freedman ML, Myers SR, Pike MC, Waliszewska A, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet 2007;39:638–44.

49. Houlston RS, Webb E, Broderick P, Pittman AM, Di Bernardo MC, Lubbe S, et al. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. Nat Genet 2008;40:1426–35.

50. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007;39:870–4.

51. Kanetsky PA, Mitra N, Vardhanabhuti S, Li M, Vaughn DJ, Letrero R, et al. Common variation in KITLG and at 5q31.3 predisposes to testicular germ cell cancer. Nat Genet 2009;41:811–5.

52. Kibriya MG, Jasmine F, Argos M, Andrulis IL, John EM, Chang-Claude J, et al. A pilot genome-wide association study of early-onset breast cancer. Breast Cancer Res Treat 2009;114:463–77.

53. Kiemeney LA, Thorlacius S, Sulem P, Geller F, Aben KK, Stacey SN, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat Genet 2008;40:1307–12.

54. Kilpivaara O, Mukherjee S, Schram AM, Wadleigh M, Mullally A, Ebert BL, et al. A germline JAK2 SNP is associated with predisposition to the development of JAK2(V617F)-positive myeloproliferative neoplasms. Nat Genet 2009;41:455–9.

55. Maris JM, Mosse YP, Bradfield JP, Hou C, Monni S, Scott RH, et al. Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. N Engl J Med 2008;358:2585–93.

56. Murabito JM, Rosenberg CL, Finger D, Kreger BE, Levy D, Splansky GL, et al. A genome-wide association study of breast and prostate cancer in the NHLBI's Framingham heart study. BMC Med Genet 2007;8:S6.

57. Ng CC, Yew PY, Puah SM, Krishnan G, Yap LF, Teo SH, et al. A genome-wide association study identifies ITGA9 conferring risk of nasopharyngeal carcinoma. J Hum Genet 2009;54:392–7.

58. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1006–10.

59. Prentice RL, Huang Y, Hinds DA, Peters U, Pettinger M, Cox DR, et al. Variation in the FGFR2 gene and the effects of postmenopausal hormone therapy on invasive breast cancer. Cancer Epidemiol Biomarkers Prev 2009;18:3079–85.

60. Rapley EA, Turnbull C, Al Olama AA, Dermitzakis ET, Linger R, Huddart RA, et al. A genome-wide association study of testicular germ cell tumor. Nat Genet 2009;41:807–10.

61. Skibola CF, Bracci PM, Halperin E, Conde L, Craig DW, Agana L, et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. Nat Genet 2009;41:873–5.

62. Song H, Ramus SJ, Tyrer J, Bolton KL, Gentry-Maharaj A, Wozniak E, et al. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. Nat Genet 2009;41:996–1000.

63. Stacey SN, Gudbjartsson DF, Sulem P, Bergthorsson JT, Kumar R, Thorleifsson G, et al. Common variants on 1p36 and 1q42 are associated with cutaneous basal cell carcinoma but not with melanoma or pigmentation traits. Nat Genet 2008;40:1313–8.

64. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 2007;39:865–9.

65. Sun J, Zheng SL, Wiklund F, Isaacs SD, Li G, Wiley KE, et al. Sequence variants at 22q13 are associated with prostate cancer risk. Cancer Res 2009;69:10–5.

66. Sun J, Zheng SL, Wiklund F, Isaacs SD, Purcell LD, Gao Z, et al. Evidence for two independent prostate cancer risk-associated loci in the HNF1B gene at 17q12. Nat Genet 2008;40:1153–5.

67. Tenesa A, Farrington SM, Prendergast JG, Porteous ME, Walker M, Haq N, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nat Genet 2008;40:631–7.

68. Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, et al. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). Nat Genet 2009;41:579–84.

69. Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet 2008;40:310–5.

70. Tomlinson IP, Webb E, Carvajal-Carmona L, Broderick P, Howarth K, Pittman AM, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. Nat Genet 2008;40:623–30.

71. Trevino LR, Yang W, French D, Hunger SP, Carroll WL, Devidas M, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1001–5.

72. Wrensch M, Jenkins RB, Chang JS, Yeh R-F, Xiao Y, Decker PA, et al. Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. Nat Genet 2009;41:905–8.

73. Wu X, Ye Y, Kiemeney LA, Sulem P, Rafnar T, Matullo G, et al. Genetic variation in the prostate stem cell antigen gene PSCA confers susceptibility to urinary bladder cancer. Nat Genet 2009;41:991–5.

74. Yeager M, Chatterjee N, Ciampa J, Jacobs KB, Gonzalez-Bosquet J, Hayes RB, et al. Identification of a new prostate cancer susceptibility locus on chromosome 8q24. Nat Genet 2009;41:1055–7.

75. Zanke BW, Greenwood CM, Rangrej J, Kustra R, Tenesa A, Farrington SM, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet 2007;39:989–94.

76. Zheng SL, Stevens VL, Wiklund F, Isaacs SD, Sun J, Smith S, et al. Two independent prostate cancer risk-associated Loci at 11q13. Cancer Epidemiol Biomarkers Prev 2009;18:1815–20.

77. Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 2009;41:324–8.

78. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet 2008;40:616–22.

79. Bierut LJ, Madden PAF, Breslau N, Johnson EO, Hatsukami D, Pomerleau OF, et al. Novel genes identified in a high-density genome wide association study for nicotine dependence. Hum Mol Genet 2007;16:24–35.

80. Saccone NL, Culverhouse RC, Schwantes-An T-H, Cannon DS, Chen X, Cichon S, et al. Multiple independent loci at chromosome 15q25.1 affect smoking quantity: a meta-analysis and comparison with lung cancer and COPD. PLoS Genet 2010;6:e1001053.

81. Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 2008;452:638–42.

82. McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet 2008;40:1404–6.

83. Saccone NL, Wang JC, Breslau N, Johnson EO, Hatsukami D, Saccone SF, et al. The CHRNA5-CHRNA3-CHRNB4 nicotinic receptor subunit gene cluster affects risk for nicotine dependence in African-Americans and in European-Americans. Cancer Res 2009;69:6848–56.

84. Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Human Mutat 2009;30:69–78.

85. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–9.

86. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.

87. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 2008;452:633–7.

88. Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, Need AC, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. PLoS Genet 2009;5:e1000421.

89. Franceschini N, van Rooij Frank JA, Prins Bram P, Feitosa Mary F, Karakas M, Eckfeldt John H, et al. Discovery and fine mapping of serum protein loci through transethnic meta-analysis. Am J Human Genet 2012;91:744–53.

90. Hancock DB, Artigas MS, Gharib SA, Henry A, Manichaikul A, Ramasamy A, et al. Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. PLoS Genet 2012;8:e1003098.

91. Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS ONE 2012;7:e51954.

92. Figueroa JD, Garcia-Closas M, Humphreys M, Platte R, Hopper JL, Southey MC, et al. Associations of common variants at 1p11.2 and 14q24.1 (RAD51L1) with breast cancer risk and heterogeneity by tumor subtype: findings from the Breast Cancer Association Consortium. Hum Mol Genet 2011;20:4693–706.

93. Dunnwald L, Rossing M, Li C. Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients. Breast Cancer Res 2007;9:R6.

94. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638–45.

95. Shikata K, Ninomiya T, Kiyohara Y. Diabetes mellitus and cancer risk: review of the epidemiological evidence. Cancer Sci 2013; 104:9–14.

96. Onitilo A, Engel J, Glurich I, Stankowski R, Williams G, Doi S. Diabetes and cancer I: risk, survival, and implications for screening. Cancer Causes Control 2012;23:967–81.

97. Fu Y-P, Edvardsen H, Kaushiva A, Arhancet J, Howe T, Kohaar I, et al. NOTCH2 in breast cancer: association of SNP rs11249433 with gene expression in ER-positive breast tumors without TP53 mutations. Mol Cancer 2010;9:113.

98. Fiúza U-M, Arias AM. Cell and molecular biology of Notch. J Endocrinol 2007;194:459–74.

99. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, et al. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. PLoS Genet 2008;4:e1000054.

100. Gudmundsdottir ET, Barkardottir R, Arason A, Gunnarsson H, Amundadottir LT, Agnarsson B, et al. The risk allele of SNP rs3803662 and the mRNA level of its closest genes TOX3 and LOC643714 predict adverse outcome for breast cancer patients. BMC Cancer 2012; 12:621.

101. Smid M, Wang Y, Klijn JGM, Sieuwerts AM, Zhang Y, Atkins D, et al. Genes associated with breast cancer metastatic to bone. J Clin Oncol 2006;24:2261–7.

102. Travis L. The epidemiology of second primary cancers. Cancer Epidemiol Biomarkers Prev 2006;15:2020–6.