FULL PAPER

BJC

British Journal of Cancer (2013) 108, 1378–1386 | doi: 10.1038/bjc.2013.7

Keywords: common genetic variants; *CDKN2A*; 9p21.3

Common genetic variants in the 9p21 region and their associations with multiple tumours

F Gu[1], R M Pfeiffer[1], S Bhattacharjee[1], S S Han[1], P R Taylor[1], S Berndt[1], H Yang[1], A J Sigurdson[1], J Toro[1], L Mirabello[1], M H Greene[1], N D Freedman[1], C C Abnet[1], S M Dawsey[1], N Hu[1], Y-L Qiao[2], T Ding[3], A V Brenner[1], M Garcia-Closas[1,4], R Hayes[1], L A Brinton[1], J Lissowska[5], N Wentzensen[1], C Kratz[1], L E Moore[1], R G Ziegler[1], W-H Chow[1], S A Savage[1], L Burdette[1], M Yeager[1], S J Chanock[1], N Chatterjee[1], M A Tucker[1], A M Goldstein[1] and X R Yang*[,1]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20852, USA; [2]Cancer Institute and Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China; [3]Shanxi Cancer Hospital, Shanxi 30013, China; [4]Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton SM2 5NG, UK and [5]Department of Cancer Epidemiology and Prevention, the M. Sklodowska-Curie Cancer Center and Institute of Oncology, 02-781 Warsaw, Poland

Background: The chromosome 9p21.3 region has been implicated in the pathogenesis of multiple cancers.

Methods: We systematically examined up to 203 tagging SNPs of 22 genes on 9p21.3 (19.9–32.8 Mb) in eight case–control studies: thyroid cancer, endometrial cancer (EC), renal cell carcinoma, colorectal cancer (CRC), colorectal adenoma (CA), oesophageal squamous cell carcinoma (ESCC), gastric cardia adenocarcinoma and osteosarcoma (OS). We used logistic regression to perform single SNP analyses for each study separately, adjusting for study-specific covariates. We combined SNP results across studies by fixed-effect meta-analyses and a newly developed subset-based statistical approach (ASSET). Gene-based *P*-values were obtained by the minP method using the Adaptive Rank Truncated Product program. We adjusted for multiple comparisons by Bonferroni correction.

Results: Rs3731239 in cyclin-dependent kinase inhibitors 2A (*CDKN2A*) was significantly associated with ESCC ($P = 7 \times 10^{-6}$). The *CDKN2A*-ESCC association was further supported by gene-based analyses ($P_{gene} = 0.0001$). In the meta-analyses by ASSET, four SNPs (rs3731239 in *CDKN2A*, rs615552 and rs573687 in *CDKN2B* and rs564398 in *CDKN2BAS*) showed significant associations with ESCC and EC ($P < 2.46 \times 10^{-4}$). One SNP in *MTAP* (methylthioadenosine phosphorylase) (rs7023329) that was previously associated with melanoma and nevi in multiple genome-wide association studies was associated with CRC, CA and OS by ASSET ($P = 0.007$).

Conclusion: Our data indicate that genetic variants in *CDKN2A*, and possibly nearby genes, may be associated with ESCC and several other tumours, further highlighting the importance of 9p21.3 genetic variants in carcinogenesis.

The chromosome 9p21.3 region has been identified as a genetic susceptibility locus for multiple disease phenotypes including coronary artery disease, diabetes and cancer (Pasmant *et al*, 2011). This region encompasses several tumour suppressor genes including cyclin-dependent kinase inhibitors 2A (*CDKN2A*), *CDKN2B*, a non-coding RNA (*CDKN2BAS*, or *ANRIL*) and methylthioadenosine phosphorylase (*MTAP*). The *CDKN2A/2B* loci are well recognised as tumour-suppressor genes that are involved in the regulation of cell cycle, aging, senescence and apoptosis (Yang *et al*, 2010b). *CDKN2A* confers susceptibility to familial melanoma and germline mutations in *CDKN2A* occur in about 20% of melanoma families (Goldstein, 2004). The *CDKN2A* encodes both p16 (*INK4A*), a negative regulator of cyclin-dependant kinases, and p14 (*ARF*), an activator of p53. The exact

*Correspondence: Dr XR Yang; E-mail: royang@mail.nih.gov

Received 23 August 2012; revised 13 December 2012; accepted 18 December 2012; published online 29 January 2013

function of *CDKN2BAS* is unknown, but it has been shown to regulate gene expression of *CDKN2A/2B* and SNPs in this locus have been associated with cardiovascular disease, cancer and other diseases in genome-wide association studies (GWAS) (Yap *et al*, 2010; Pasmant *et al*, 2011). The *MTAP*, identified by GWAS as a naevus- and melanoma-associated gene (Bishop *et al*, 2009; Falchi *et al*, 2009), encodes an enzyme that has a role in polyamine metabolism. Loss of MTAP expression can exert a tumour-promoting effect, and has been observed in a variety of other tumours (Stevens *et al*, 2009), suggesting that *MTAP* may function as a tumour suppressor gene. The 9p21.3 region also includes a cluster of type I interferon (*IFN*) genes, which encode pleiotropic cytokines that exhibit strong antiviral, antiproliferative and immunomodulatory effects (Stark *et al*, 1998).

In addition to its well established role in melanoma, deletions of the 9p21.3 region have been observed in a variety of cancers (van der Riet *et al*, 1994; Okami *et al*, 1997; Waber *et al*, 1997; Nakanishi *et al*, 1999; Perinchery *et al*, 1999; Schmid *et al*, 2000; Sanchez-Cespedes *et al*, 2001; Hu *et al*, 2004; Hustinx *et al*, 2005; Bartoletti *et al*, 2007; Gu *et al*, 2008), and SNPs in 9p21.3 have been associated with breast cancer, melanoma and glioma by GWAS (Bishop *et al*, 2009; Shete *et al*, 2009; Wrensch *et al*, 2009; Turnbull *et al*, 2010).These findings are consistent with a broad role for 9p21.3 genes in carcinogenesis. However, whether genetic polymorphisms in 9p21.3 confer susceptibility to other cancers remains unclear. The goal of the current study was to systematically evaluate variants in 9p21.3 with the risk of multiple cancers/tumours.

MATERIALS AND METHODS

Study population. This study sample included data from eight studies that participated in iSelect, a jointly conducted project in the Division of Cancer Epidemiology and Genetics of the National Cancer Institute (NCI), with a goal to evaluate common genetic variants in selected genes and pathways in multiple tumours, especially rare cancers (Gao *et al*, 2009; Yang *et al*, 2010a; Gao *et al*, 2011; Mirabello *et al*, 2011; Han *et al*, 2012; Neta *et al*, 2012). The study samples comprised seven cancers (renal cell carcinoma (RCC), endometrial cancer (EC), thyroid cancer (ThC), colorectal cancer (CRC), oesophageal squamous cell carcinoma (ESCC), gastric cardia adenocarcinoma (GCA) and osteosarcoma (OS)) and one benign condition (colorectal adenoma (CA)). Study participants were Chinese (for ESCC and GCA studies) or whites (all other studies). The design of these studies included nested case–control (RCC (1994); Prorok et al, 2000; Han et al, 2012), CRC, CA (Gao *et al*, 2011)), population-based case–control (EC (Yang *et al*, 2010a)), hospital-based case–control (OS (Troisi *et al*, 2006; Mirabello *et al*, 2011)), and case–control studies of mixed design (ThC (Neta *et al*, 2012), ESCC and GCA (Blot *et al*, 1993; Gao *et al*, 2009)). After excluding subjects with a low genotyping completion rate (<80%), the final analysis for each tumour outcome included 437 RCC cases and 1603 controls; 417 EC cases and 407 controls; 344 ThC cases and 452 controls; 393 CRC cases and 434 controls; 1234 CA cases and 1368 controls; 1027 ESCC cases and 1452 controls; 753 GCA cases and 1452 controls; 96 OS cases and 1428 controls. We pooled controls for ESCC and GCA (1452 controls total), as these cases were drawn from the same underlying studies. The RCC and OS shared a subset (1170 and 1363, respectively) of PLCO controls with CA. Detailed information for each study is summarised in Table 1 and Supplementary Table S1.

After correction for multiple testing, ESCC, GCA and CA had adequate power (94%, 87% and 98%, respectively) to detect an association for a SNP with minor allele frequency (MAF) = 0.35 and an odds ratio (OR) of 1.4, while all other studies were underpowered (power < 80%). However, our aim was to identify genetic variants in the 9p21 region associated with multiple cancer/tumour outcomes using combined data across studies.

SNP selection, genotyping and quality control. SNP selection, genotyping and quality control have been described previously (Yang *et al*, 2010b). In brief, 252 tag SNPs for 22 genes located at the chromosome 9p21.3 region (19.9–32.8 Mb) were genotyped at the NCI Core Genotyping Facility (Advanced Technology Center, Gaithersburg, MD; http://snp500cancer.nci.nih.gov) using a custom-designed iSelect Infinium assay (Illumina, www.illumina.com). From telomere to centromere, these genes included: *IFNB1*, *IFNW1*, *IFNA21*,*10*,*16*,*17*,*14*,*5*, *KLHL9*, *IFNA6*, *2*, *8*, *1*, *IFNE1*, *MTAP*, *CDKN2A*, *CDKN2B*, *CDKN2BAS*, *TUSC1*, *PLAA*, *IFNK*, *ACO1*. For each gene, SNPs spanned 20 kb 5′ of the transcription start point (exon 1) to 10 kb 3′ of the last exon. Tag SNPs were selected using a MAF criterion of MAF >5% based upon HapMap data for whites (CEU) and Yoruba (YRI) samples using a Tagging

Table 1. Description of the study samples

Study	Cases	Controls	Ethnicity	Country	Study design	Covariates
RCC	437	1603	Caucasian	US, Finland	Nested case–control within the PLCO Cancer Screening Trial (Prorok *et al*, 2000) and ATBC (1994)	Age, gender, study centre (11)
EC	417	407	Caucasian	Poland	Population-based case–control (Yang *et al*, 2010a)	Age, site (2)
ThC	344	452	Caucasian	US	Cases from the USRT cohort, or University of Texas M D Anderson Cancer Center. Controls from USRT (Neta *et al*, 2012)	Age, gender, birth year category
CRC	393	434	Caucasian	US	Nested case–control within screening arm of PLCO	Age
CA	1234	1368	Caucasian	US	Nested case–control within screening arm of PLCO (Gao *et al*, 2011)	Age
ESCC	1027	1452	Asian	China	Neighborhood-based case–control from the UGI Cancer Genetics Project (Gao *et al*, 2009) and nested case–control from NITs (Blot *et al*, 1993)	Age, gender, study region (2)
GCA	753	1452	Asian	China	Same as ESCC	Age, gender, study region (2)
OS	96	1428	Caucasian	US	Hospital-based case–control (63 controls); 1365 additional controls from PLCO (Mirabello *et al*, 2011; Troisi *et al*, 2006)	Gender

Abbreviations: ATBC = alpha-tocopherol, Beta-Carotene Cancer Prevention; CA = colorectal adenoma; CRC = colorectal cancer; EC = endometrial cancer; ESCC = oesophageal squamous cell carcinoma; GCA = gastric cardia adenocarcinoma; NITs = nutrition intervention trials; PLCO = prostate, lung, colorectal and ovarian; RCC = renal cell carcinoma; OS = osteosarcoma; ThC = thyroid cancer; UGI = upper gastrointestinal; USRT = US radiologic technologists.

algorithm (Carlson *et al*, 2004). Selected SNPs are listed in Supplementary Table 2.

The iSelect panel was validated using all three HapMap populations (CEU, YRI, Japanese and Chinese). The SNPs with low (<90%) genotyping completion rate, low (<95%) concordance rate or deviation ($P<0.001$) from Hardy–Weinberg equilibrium among controls were excluded from each participating study. The number of SNPs included in the final analyses were: 170 SNPs for RCC, 202 SNPs for EC, 195 SNPs for ThC, 193 SNPs for CRC, 203 SNPs for CA, 139 SNPs for ESCC and GCA and 200 SNPs for OS. In the ESCC and GCA study, a larger number of SNPs were excluded due to low MAF (<5%), likely reflecting differences between white and Asian populations.

Statistical analyses. We first assessed the association between each SNP and each cancer outcome separately. Unconditional logistic regression was used to estimate ORs and 95% confidence intervals (CIs) and *P*-values for trend, using additive coding for genotypes (0,1,2 minor alleles). The homozygote of the common allele served as the reference group. Heterozygous and homozygous rare genotypes were combined when the number of subjects with homozygous minor alleles was <5, and a dominant genetic model was used. Appropriate covariates adjustment was performed for each tumour outcome per discussion with principal investigators of each study (Table 1).

To examine whether 9p21 variants were associated with multiple cancer/tumour outcomes, we conducted meta-analyses combining data from the eight studies. To combine SNP results across studies, we first used a standard fixed-effect meta-analysis and then a newly developed subset-based statistical approach (ASSET) (Bhattacharjee *et al*, 2012). ASSET is a modified fix-effect meta-analysis approach that allows for heterogeneity of SNP effects on different outcomes by exhaustively exploring subsets of the studies for the presence of association signals. The ASSET test statistic $Z(S)$ for a given subset S of k studies is a weighted sum of the k study-specific test statistics, $Z(S) = a_1 Z_1 + \ldots + a_k Z_k$, where the a_j is the proportion of the sample size for the jth study relative to the total sample size for the studies in the given subset S. The overall evidence of the association of the SNP is then based on evaluation of $Z_{max} = \max_S |Z(S)|$, i.e. the maximum of the subset-specific test statistics over all possible subsets of the studies. Under the null hypothesis the vector of values $Z(S)$ has a multivariate normal distribution with mean zero and variances equal to one. The correlation between $Z(A)$ and $Z(B)$ for two different subsets A and B is given in (Bhattacharjee *et al*, 2012). We computed a two-sided version of the test that also allows the detection of effects in opposite directions. Both fixed-effect and subset-based meta-analyses were performed using the 'ASSET' R package, which can take into account shared controls across studies. Because the SNP minor alleles may differ across studies, we standardised the effects before combining the data by multiplying the beta-coefficients of SNPs by 1 or -1.

Gene-based analyses were performed on the 22 genes to assess the significance of the joint effect of multiple SNPs in each gene on each outcome separately. Gene-based *P*-values (P_{gene}) were computed using the minP method by Adaptive Rank Truncated Product (ARTP) program (Yu *et al*, 2009). The minimum *P*-value of each gene was used as the test statistic and its significance was assessed using a permutation test with 10 000 permutations, taking into account the number of SNPs genotyped in each gene and their linkage disequilibrium (LD) structure.

We used Bonferroni correction to account for the number of SNPs or genes and studies tested, therefore P for SNP $< 3.1 \times 10^{-5}$ ($0.05/(203 \times 8)$) and $P_{gene} < 2.8 \times 10^{-4}$ ($0.05/(22 \times 8)$) were used to define SNP-based and gene-based statistical significance. In meta-analysis, P for combined analysis $< 2.46 \times 10^{-4}$ (0.05/203) was considered statistically significant after Bonferroni correction for numbers of SNPs. All statistical analyses were performed using the R software.

RESULTS

When analysing each study separately, we found one SNP in *CDKN2A* (rs3731239) that was significantly associated with ESCC after Bonferroni correction ($P = 7 \times 10^{-6}$) (Table 2a). The minor allele (G, MAF = 0.12) of this SNP was associated with increased ESCC risk (OR = 1.51, 95% CI = 1.25, 1.84, for AG *vs* AA; OR = 1.88, 95% CI = 1.04, 3.41, for GG *vs* AA; Table 2a). Figure 1 shows that the LD pattern and genotype frequencies among controls were different in the Chinese and Caucasian samples.

We also found 18 additional SNPs that were associated with at least one tumour outcome at $P<0.01$ (Table 2b), although the associations were not significant after Bonferroni correction. Among them, one SNP in *MTAP* (rs7023329) that was previously associated with melanoma and nevi in several GWAS (Bishop *et al*, 2009; Barrett *et al*, 2011), was associated with CA ($P = 0.0005$). Another previously identified SNP (rs4977756) in *CDKN2BAS* from a GWAS for glioma (Shete *et al*, 2009), was associated with EC ($P = 0.009$) and ESCC ($P = 0.002$).

In fixed-effect meta-analyses, only rs7023329 in *MTAP* showed marginal association (fixed effect $P<0.05$) before correction for multiple testing (Table 2c). When using the subset-based approach (ASSET), rs7023329 showed suggestive association with multiple tumours (positive effect $P = 0.007$), with the strongest signal obtained from the subset combining data from CRC, CA and OS studies (Table 2c). In addition, the subset approach identified significant associations between rs3731239 in *CDKN2A*, rs615552 and rs573687 in *CDKN2B*, and rs564398 in *CDKN2BAS*, and EC and ESCC after Bonferroni correction (positive effect $P < 2.46 \times 10^{-4}$, Table 2c), although these associations seemed to be mainly driven by ESCC based on sensitivity analyses that excluded ESCC. The effects of all SNPs with $P \leqslant 0.01$ in the subset analyses showed the same direction (positive effect) across contributing study outcomes (Table 2b and c).

Gene-based analyses showed that the *CDKN2A* gene was significantly associated with ESCC ($P_{gene} = 0.0001$) and the association remained significant after adjusting for multiple testing (Table 3). Other genes in the nearby region, *MTAP* ($P_{gene} = 0.015$), *CDKN2B* ($P_{gene} = 0.01$) and *CDKN2BAS* ($P_{gene} = 0.009$), also showed suggestive associations with ESCC (Table 3). In addition, *MTAP* showed a suggestive association with CA ($P_{gene} = 0.006$).

Table 2a. Association between rs3731239 and ESCC in a Chinese population[a]

Genotype	Case, $n = 1027$ (%)[b]	Control, $n = 1452$ (%)[b]	OR	95% CI	*P*-trend
AA	712 (69.3)	1093 (75.3)	1.00	—	
AG	272 (26.5)	294 (20.1)	1.51	1.25–1.84	
GG	25 (2.4)	22 (1.5)	1.88	1.04–3.41	
Per G allele			1.47	1.24–1.75	7×10^{-6}

Abbreviations: CI = confidence interval; ESCC = oesophageal squamous cell carcinoma; OR = odds ratio.
[a]Results were obtained from unconditional logistic regression, adjusting for age, gender and study region.
[b]Percentage does not sum up to 1 owing to missing values.

Table 2b. Selected SNP-based results, with *P*-value < 0.01 for at least one tumour outcome

			Renal cell carcinoma		Endometrial cancer		Thyroid cancer		Colorectal cancer		Colorectal adenoma		Oesophagus squamous cell carcinoma		Gastric cardia adenocarcinoma		Osteosarcoma	
SNP	**Gene**	**A1**[a]	**OR**[b]	***P***[b]	**OR**	***P***	**OR**	***P***	**OR**	***P***	**OR**	***P***	**OR**	***P***	**OR**	***P***	**OR**	***P***
rs17692502	*IFNW1*	G	0.9	0.23	0.84	0.11	1.18	0.18	0.72	**0.006**	1.01	0.93					1.11	0.53
rs10964862	*IFNW1*	A	0.98	0.8	0.99	0.89	1.04	0.7	0.74	**0.007**	0.99	0.88	1.01	0.87	0.82	0.05	1.15	0.35
rs10119678	*IFNA6*	A	1.15	0.19	1.12	0.39	0.84	0.25	0.95	0.7	1.22	**0.009**	1.05	0.47	1.06	0.36	0.84	0.48
rs10757257	*MTAP*	A	0.95	0.54	1.15	0.15	0.97	0.76	0.83	0.08	0.85	**0.005**	0.95	0.36	1.11	0.12	0.8	0.15
rs2039971	*MTAP*	T	0.92	0.65					1.28	0.31	1.15	0.29	1.25	**0.002**	0.95	0.48		
rs7023329	*MTAP*	A	1.02	0.79	0.97	0.73	1.04	0.72	1.19	0.1	1.22	**0.0005**	1.06	0.35	0.91	0.13	1.26	0.12
rs7027989	*MTAP*	A	1.15	0.07	0.95	0.6	1	0.98	1.15	0.19	1.17	**0.006**	0.89	0.08	0.9	0.16	1.25	0.14
rs7874112	*MTAP*	G	0.94	0.61	0.78	0.17	1.19	0.34	1.01	0.94	1.02	0.84	1.24	**0.002**	0.9	0.21	1.08	0.75
rs10811629	*MTAP*	G	0.98	0.75	1.15	0.15	0.92	0.44	0.9	0.29	0.86	**0.008**	0.94	0.34	1.1	0.15	0.78	0.1
rs10757261	*CDKN2A*	A	0.98	0.75	0.97	0.76	1.19	0.12	0.96	0.67	0.97	0.63	0.84	**0.007**	0.98	0.8	0.87	0.37
rs3731239	*CDKN2A*	G	1.03	0.73	1.22	0.05	0.85	0.15	1.05	0.64	0.97	0.56	1.47	$\mathbf{7 \times 10^{-6}}$[c]	1.16	0.13	0.84	0.27
rs1063192	*CDKN2B*	G	1.01	0.85	1.31	**0.009**	0.89	0.27	0.93	0.46	0.91	0.09					1.01	0.96
rs573687	*CDKN2B*	A	0.96	0.61	1.28	0.015	0.87	0.19	0.89	0.27	0.95	0.36	1.32	**0.002**	1	0.99	0.88	0.41
rs518394	*CDKN2B*	C	0.94	0.48	1.35	**0.003**					0.91	0.09					0.86	0.34
rs615552	*CDKN2B*	C	0.94	0.42	1.32	**0.005**	0.83	0.07	0.94	0.58	0.91	0.11	1.32	**0.002**	1	0.98	0.89	0.45
rs564398	*CDKN2BAS*	C	0.97	0.74	1.35	**0.003**	0.87	0.21	0.87	0.18	0.92	0.13	1.34	**0.001**	1.01	0.92	0.95	0.75
rs11790231	*CDKN2BAS*	A	1.06	0.64	0.89	0.55	1.63	**0.006**	1.02	0.92	1.04	0.65	0.86	0.06	0.92	0.31	0.76	0.34
rs4977756	*CDKN2BAS*	G	1.02	0.79	1.3	**0.009**	0.88	0.26	0.93	0.5	0.9	0.06	1.25	**0.002**	1.08	0.32	0.91	0.55
rs10757274	*CDKN2BAS*	G	0.98	0.81	0.76	**0.006**	1.12	0.28	0.99	0.89	1.03	0.6	0.94	0.31	1.05	0.49	1.03	0.86
Total SNP no.	170		202		195		193		203		139		139		200			

Abbreviations: CI = confidence interval; OR = odds ratio; SNP = single-nucleotide polymorphism.

[a]A1 is the effect allele (minor allele of Colorectal Adenoma study population).

[b]ORs and trend *P*-values for each SNP-tumour association were obtained by unconditional logistic regression with the adjustment of study-specific covariates listed in Table 1. *P*-value < 0.01 was shown in boldface.

[c]Significant after Bonferroni correction for number of SNPs and studies ($P < 0.05/(203 \times 8) = 3.1 \times 10^{-5}$).

Table 2c. Meta-analyses of selected SNPs with two-sided subset search $P \leqslant 0.01$[a]

			Two-side subset search[c]			
SNP	**Gene**	**Fixed-effect P[b]**	**Combined *P***	***P* for positive effect**	**Subsets of studies with strongest signal with positive effect**	***P* for negative effect**
rs7023329	*MTAP*	0.04	0.01	0.007	CRC, CA, OS	0.23
rs3731239	*CDKN2A*	0.06	3.13×10^{-4}	8.7×10^{-5}[d]	EC, ESCC	0.32
rs615552	*CDKN2B*	0.91	7.3×10^{-4}	1.3×10^{-4}[d]	EC, ESCC	0.53
rs573687	*CDKN2B*	0.82	2.0×10^{-3}	2.4×10^{-4}[d]	EC, ESCC	1
rs4977756	*CDKN2BAS*	0.38	3.2×10^{-3}	0.001	EC, ESCC	0.25
rs564398	*CDKN2BAS*	0.83	8.3×10^{-4}	2.0×10^{-4}[d]	EC, ESCC	0.39

Abbreviations: CA = colorectal adenoma; CRC = colorectal cancer; EC = endometrial cancer; ESCC = oesophageal squamous cell carcinoma; OS = osteosarcoma; SNP = single-nucleotide polymorphism.
[a]Results were from ASSET, a subset-based association analysis for combining SNP-based results across eight tumour outcomes.
[b]Fixed effect *P* was calculated by standard fixed-effect meta-analyses.
[c]Two-sided subset search allowed for opposite directions of allele effects across different outcomes.
[d]Significant after Bonferroni correction for number of SNPs ($0.05/203 = 2.46 \times 10^{-4}$).

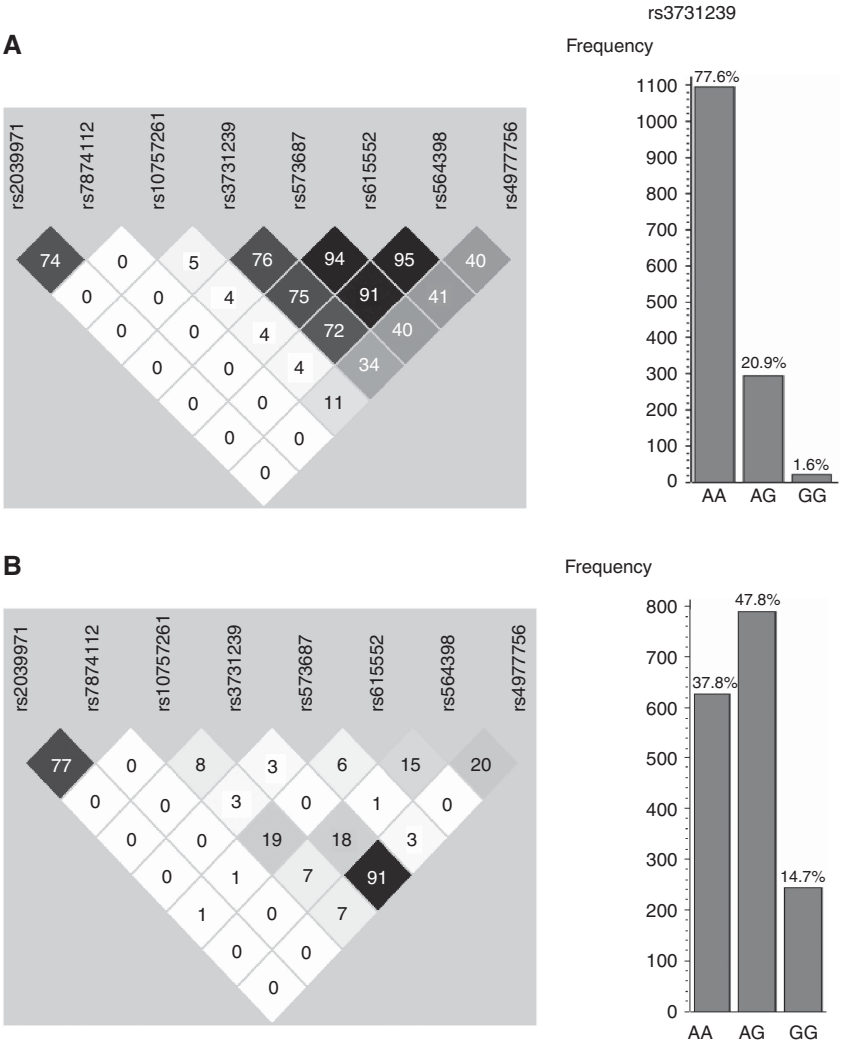

Figure 1. Linkage disequilibrium structures and genotype frequencies of rs3731239 among controls of Chinese (**A**) and Caucasian (**B**) samples. The LD (indicated by r^2) maps were drawn using the Haploview software, based on the genotyping data of control samples for ESCC (**A**) and CA (**B**). LD patterns in other Caucasian studies were similar to that in CA.

Table 3. Gene-based *P*-values for 9p21.3 genes in association with eight tumour outcomes[a]

Gene[b]	Renal cell carcinoma	Endometrial cancer	Thyroid cancer	Colorectal cancer	Colorectal adenoma	Oesophageal squamous cell carcinoma	Gastric cardia adenocarcinoma	Osteosarcoma
	(473/1603)[c]	(417/407)	(344/452)	(393/434)	(1234/1368)	(1027/1452)	(753/1452)	(96/1426)
IFNB1	0.15	0.94	0.74	0.79	0.39	0.27	0.91	0.99
IFNW1	0.62	0.63	0.56	0.06	0.68	0.82	0.3	0.97
IFNA21	0.65	0.65	0.97	0.78	0.28	0.78	0.9	0.74
IFNA10	0.51	0.76	0.8	0.88	0.07	0.87	0.26	0.9
IFNA16	0.42	0.71	0.35	0.53	0.29	0.95	0.05	0.52
IFNA17	0.58	0.95	0.99	0.28	0.78	0.32	0.86	0.89
IFNA14	0.4	0.47	0.66	0.98	0.51	0.7	0.56	0.65
IFNA5	0.61	0.88	0.82	0.48	0.08	0.77	0.22	0.17
KLHL9	0.31	0.7	0.45	0.46	0.038	0.68	0.16	0.75
IFNA6	0.4	0.57	0.51	0.74	0.026	0.56	0.32	0.81
IFNA2	0.75	0.11	0.73	0.87	0.1	0.66	0.38	0.89
IFNA8	0.76	0.65	0.88	0.94	0.19	0.26	0.31	0.6
IFNA1	0.23	0.66	1	0.99	0.14	0.18	0.56	0.65
IFNE1	0.07	0.14	0.53	0.29	0.37	0.89	0.25	0.57
MTAP	0.37	0.12	0.76	0.31	**0.006**	0.015	0.52	0.63
CDKN2A	0.88	0.32	0.65	0.47	0.77	**0.0001**[d]	0.43	0.19
CDKN2B	0.35	0.02	0.26	0.54	0.41	**0.01**	0.67	0.21
CDKN2BAS	0.99	0.04	0.07	0.44	0.45	**0.009**	0.84	0.43
TUSC1	0.75	0.77	0.18	0.14	0.1	0.19	0.79	0.32
PLAA	0.98	0.73	0.21	0.46	0.58	0.36	0.88	0.64
IFNK	0.07	0.82	0.6	0.92	0.3	0.84	0.49	0.86
ACO1	0.46	0.73	0.27	0.14	0.74	0.53	0.3	0.44

[a]Gene-based *P* values were computed using the minP method, based on 10 000 permutations; *P*-value $\leqslant$ 0.01 was shown in boldface.
[b]Genes are ordered by location from telomere to centromere.
[c]Number of cases and controls.
[d]Significant after Bonferroni correction for number of genes and studies ($0.05/(22 \times 8) = 2.8 \times 10^{-4}$).

DISCUSSION

In this study, we evaluated associations of up to 203 SNPs in 22 genes located on chromosome 9p21.3 with the risk of eight tumour outcomes in data from eight case–control studies. When analysing each tumour outcome separately, we identified a single SNP in *CDKN2A* (rs3731239) that was significantly associated with the risk of ESCC, after correction for multiple comparisons. Gene-based analyses also suggested that the *CDKN2A* gene was significantly associated with ESCC. In the subset-based meta-analyses, four SNPs (rs3731239 in *CDKN2A*, rs615552 and rs573687 in *CDKN2B*, and rs564398 in *CDKN2BAS*) showed significant associations with ESCC and EC. Two previously identified GWAS SNPs, rs7023329 in *MTAP* for melanoma and nevi and rs4977756 in *CDKN2A* for glioma, showed suggestive associations with CA (for rs7023329) and EC and ESCC (for rs4977756), respectively, in our study. Our findings further highlight the importance of 9p21.3, in particular the *MTAP-CDKN2A/2B/CDKN2BAS* region, in the pathogenesis of multiple tumours.

Rs3731239 previously demonstrated weak associations with breast cancer (Driver *et al*, 2008; Mavaddat *et al*, 2009) and ovarian cancer (Goode *et al*, 2009) in predominantly Caucasian populations. In our study, this SNP was significantly associated with ESCC in Chinese and only weakly associated with EC in Caucasians. The minor allele of this SNP is more common in Caucasians (0.39 among controls in our study) than in Chinese (0.12 among controls in ESCC). In addition, the two ethnic populations showed distinct LD patterns in the region flanking this SNP, which may also contribute to the differences in the association observed.

Recent studies have suggested that the 9p21.3 region was enriched in regulatory sequences such as enhancers that regulate the expression of genes in this region (*MTAP-CDKN2A/2B/CDKN2BAS*) and downstream (such as *IFNA21*), thereby establishing a functional link between 9p21 genetic variation and immune signalling pathways (Harismendy *et al*, 2011). Interestingly, rs564398 in *CDKN2BAS*, which showed suggestive associations with both EC and ESCC in our study (see Tables 2b and c), was located within a predicted enhancer sequence. The most significant SNP in our study, rs3731239 in *CDKN2A*, is located adjacent to the promoter region of *CDKN2A* (about 500 bp away from a CpG island and predicted transcription binding and DNase I sites based on ENCODE data, http://www.genome.ucsc.edu/ENCODE/). A previous study correlating 9p21 SNPs with gene expression found that rs3731239 was significantly associated with allele-specific expression of *CDKN2BAS* ($P = 10^{-25}$) (Cunnington *et al*, 2010). Three other SNPs (rs1063192, rs564398 and rs11790231) in the *CDKN2B/CDKN2BAS* locus that showed suggestive associations with EC (and/or ESCC, ThC) were also significantly associated with

allele-specific expression of *CDKN2BAS*. *CDKN2BAS* is a non-coding RNA within the *CDKN2A/2B* locus, which has been identified by GWAS of multiple diseases; its expression showed the strongest association with the multiple phenotypes (coronary disease, stroke, diabetes, melanoma and glioma) that were associated with the 9p21.3 region, as compared with the three other genes of the cluster (*MTAP*, *CDKN2A*, *CDKN2B*) (Pasmant *et al*, 2011). The *CDKN2BAS* is involved in regulating *CDKN2A/2B* expression through a *cis*-acting mechanism as well as by regulating cell proliferation and senescence through pathways independent from *CDKN2A/2B* (Visel *et al*, 2010; Congrains *et al*, 2012). In addition to *CDKN2BAS*, two SNPs in *MTAP* (rs10757257 and rs7027989), which were suggestively associated with CA in our study, were also found to be expression quantitative trait loci for *MTAP* (Zeller *et al*, 2010). These data, combined with previous publications, indicate that common genetic variants in this region may influence disease risk by regulating gene expression through a *cis*-effect. With rapid progress in mapping regulatory elements and the growing availability of cell and tissue-specific gene expression data, future studies should be able to evaluate the functional relevance of genetic variants at 9p21.3.

Somatic 9p21 deletions frequently occur in human cancers such as bladder cancer, pancreatic cancer, oesophageal cancer, glioma and melanoma (Schmid *et al*, 2000; Hu *et al*, 2004; Hustinx *et al*, 2005; Bartoletti *et al*, 2007; Gu *et al*, 2008; Rakosy *et al*, 2008). In a previous study conducted in the same Chinese population from which the ESCC cases in the current study were obtained, the majority (73%) of ESCC tumour specimens analysed were found to have LOH at 9p21–22, and 25% (14 of 56) of tumours had *CDKN2A* mutations (point mutations, deletions, insertions) (Hu *et al*, 2004). In addition, promoters in *CDKN2A* are typically methylated in ESCC tumours (Roth *et al*, 2006).

Ours is the first systematic evaluation of genetic variation in the 9p21.3 region in relation to multiple tumour outcomes. The strengths of our study include the careful and comprehensive selection of genes in the entire 9p21.3 region, the application of a newly developed subset analysis method to combine SNP data across multiple studies, and use of a gene-based permutation analysis method to comprehensively evaluate variation in genes with cancer risk. In addition, SNPs were genotyped for all studies using the same platform and quality control procedures. Our findings suggest that combining data from multiple cancer outcomes may provide additional information in understanding disease associations with GWAS variants.

There are several limitations in our study. First, studies included in this analysis varied by study design, population ethnicity and sharing controls in some studies, which posed challenges for combining data as well as generalising the findings. We therefore applied a new statistical approach, which was specifically designed to handle heterogeneity across studies. Second, our sample size was in general small, which may limit statistical power for identifying significant associations in the smaller individual studies. In fact, most associations were not significant after correcting for multiple testing, with the noted exception of rs3731239 in *CDKN2A*, with ESCC, which was among the largest studies. However, the Bonferroni test is conservative, especially for previously identified GWAS SNPs, and therefore the observed associations in our study warrant future investigation in larger samples.

In conclusion, our data indicated that genetic variants in the 9p21.3 region, particularly near the *MTAP-CDKN2A/2B/CDKN2BAS*, may be associated with ESCC and possibly several other tumours. Our findings further highlight the importance of the 9p21.3 region in disease susceptibility and cancer aetiology. Future studies are needed to further investigate the role of this chromosomal region in cancer pathogenesis. Further, data on somatic alterations of this region (in tumour tissue), such as gene expression, will be particularly helpful to identify the mechanisms underlying the observed associations.

ACKNOWLEDGEMENTS

We thank Xiaoqin Xiong for providing help with the data analyses. We thank iSelect principal investigators of ThC, EC, TGCT, RCC, CRC, CA, ESCC, GCA and OS studies for kindly providing data for this project. We thank Dr. Paula Hyland for her help in reviewing ENCODE data. This research was supported by the Intramural Research Program of the NIH, NCI, DCEG.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

ATBC Cancer Prevention Study Group (1994) The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. *Ann Epidemiol* **4**(1): 1–10.

Barrett JH, Iles MM, Harland M, Taylor JC, Aitken JF, Andresen PA, Akslen LA, Armstrong BK, Avril MF, Azizi E, Bakker B, Bergman W, Bianchi-Scarra G, Bressac-de Paillerets B, Calista D, Cannon-Albright LA, Corda E, Cust AE, Debniak T, Duffy D, Dunning AM, Easton DF, Friedman E, Galan P, Ghiorzo P, Giles GG, Hansson J, Hocevar M, Hoiom V, Hopper JL, Ingvar C, Janssen B, Jenkins MA, Jonsson G, Kefford RF, Landi G, Landi MT, Lang J, Lubinski J, Mackie R, Malvehy J, Martin NG, Molven A, Montgomery GW, van Nieuwpoort FA, Novakovic S, Olsson H, Pastorino L, Puig S, Puig-Butille JA, Randerson-Moor J, Snowden H, Tuominen R, Van Belle P, van der Stoep N, Whiteman DC, Zelenika D, Han J, Fang S, Lee JE, Wei Q, Lathrop GM, Gillanders EM, Brown KM, Goldstein AM, Kanetsky PA, Mann GJ, Macgregor S, Elder DE, Amos CI, Hayward NK, Gruis NA, Demenais F, Bishop JA, Bishop DT (2011) Genome-wide association study identifies three new melanoma susceptibility loci. *Nat Genet* **43**(11): 1108–1113.

Bartoletti R, Cai T, Nesi G, Roberta Girardi L, Baroni G, Dal Canto M (2007) Loss of P16 expression and chromosome 9p21 LOH in predicting outcome of patients affected by superficial bladder cancer. *J Surg Res* **143**(2): 422–427.

Bhattacharjee S, Rajaraman P, Jacobs KB, Wheeler WA, Melin BS, Hartge P, Yeager M, Chung CC, Chanock SJ, Chatterjee N (2012) A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am J Hum Genet* **90**(5): 821–835.

Bishop DT, Demenais F, Iles MM, Harland M, Taylor JC, Corda E, Randerson-Moor J, Aitken JF, Avril MF, Azizi E, Bakker B, Bianchi-Scarra G, Bressac-de Paillerets B, Calista D, Cannon-Albright LA, Chin AWT, Debniak T, Galore-Haskel G, Ghiorzo P, Gut I, Hansson J, Hocevar M, Hoiom V, Hopper JL, Ingvar C, Kanetsky PA, Kefford RF, Landi MT, Lang J, Lubinski J, Mackie R, Malvehy J, Mann GJ, Martin NG, Montgomery GW, van Nieuwpoort FA, Novakovic S, Olsson H, Puig S, Weiss M, van Workum W, Zelenika D, Brown KM, Goldstein AM, Gillanders EM, Boland A, Galan P, Elder DE, Gruis NA, Hayward NK, Lathrop GM, Barrett JH, Bishop JA (2009) Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet* **41**(8): 920–925.

Blot WJ, Li JY, Taylor PR, Guo W, Dawsey S, Wang GQ, Yang CS, Zheng SF, Gail M, Li GY, Yu Y, Liu BQ, Tangrea J, Sun YH, Liu FS, Fraumeni JF, Zhang YH, Li B (1993) Nutrition intervention trials in Linxian, China: supplementation with specific vitamin/mineral combinations, cancer incidence, and disease-specific mortality in the general population. *J Natl Cancer Inst* **85**(18): 1483–1492.

Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA (2004) Selecting a maximally informative set of single-nucleotide polymorphisms

for association analyses using linkage disequilibrium. *Am J Hum Genet* **74**(1): 106–120.

Congrains A, Kamide K, Katsuya T, Yasuda O, Oguro R, Yamamoto K, Ohishi M, Rakugi H (2012) CVD-associated non-coding RNA, ANRIL, modulates expression of atherogenic pathways in VSMC. *Biochem Biophys Res Commun* **419**(4): 612–616.

Cunnington MS, Santibanez Koref M, Mayosi BM, Burn J, Keavney B (2010) Chromosome 9p21 SNPs associated with multiple disease phenotypes correlate with ANRIL expression. *PLoS Genet* **6**(4): e1000899.

Driver KE, Song H, Lesueur F, Ahmed S, Barbosa-Morais NL, Tyrer JP, Ponder BA, Easton DF, Pharoah PD, Dunning AM (2008) Association of single-nucleotide polymorphisms in the cell cycle genes with breast cancer in the British population. *Carcinogenesis* **29**(2): 333–341.

Falchi M, Bataille V, Hayward NK, Duffy DL, Bishop JA, Pastinen T, Cervino A, Zhao ZZ, Deloukas P, Soranzo N, Elder DE, Barrett JH, Martin NG, Bishop DT, Montgomery GW, Spector TD (2009) Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. *Nat Genet* **41**(8): 915–919.

Gao Y, Hayes RB, Huang WY, Caporaso NE, Burdette L, Yeager M, Chanock SJ, Berndt SI (2011) DNA repair gene polymorphisms and tobacco smoking in the risk for colorectal adenomas. *Carcinogenesis* **32**(6): 882–887.

Gao Y, Hu N, Han X, Giffen C, Ding T, Goldstein A, Taylor P (2009) Family history of cancer and risk for esophageal and gastric cancer in Shanxi, China. *BMC cancer* **9**: 269.

Goldstein AM (2004) Familial melanoma, pancreatic cancer and germline CDKN2A mutations. *Hum Mutat* **23**(6): 630.

Goode EL, Fridley BL, Vierkant RA, Cunningham JM, Phelan CM, Anderson S, Rider DN, White KL, Pankratz VS, Song H, Hogdall E, Kjaer SK, Whittemore AS, DiCioccio R, Ramus SJ, Gayther SA, Schildkraut JM, Pharaoh PP, Sellers TA (2009) Candidate gene analysis using imputed genotypes: cell cycle single-nucleotide polymorphisms and ovarian cancer risk. *Cancer Epidemiol Biomarkers Prev* **18**(3): 935–944.

Gu J, Horikawa Y, Chen M, Dinney CP, Wu X (2008) Benzo(a)pyrene diol epoxide-induced chromosome 9p21 aberrations are associated with increased risk of bladder cancer. *Cancer Epidemiol Biomarkers Prev* **17**(9): 2445–2450.

Han SS, Yeager M, Moore LE, Wei MH, Pfeiffer R, Toure O, Purdue MP, Johansson M, Scelo G, Chung CC, Gaborieau V, Zaridze D, Schwartz K, Szeszenia-Dabrowska N, Davis F, Bencko V, Colt JS, Janout V, Matveev V, Foretova L, Mates D, Navratilova M, Boffetta P, Berg CD, Grubb 3rd RL, Stevens VL, Thun MJ, Diver WR, Gapstur SM, Albanes D, Weinstein SJ, Virtamo J, Burdett L, Brisuda A, McKay JD, Fraumeni Jr. JF, Chatterjee N, Rosenberg PS, Rothman N, Brennan P, Chow WH, Tucker MA, Chanock SJ, Toro JR (2012) The chromosome 2p21 region harbors a complex genetic architecture for association with risk for renal cell carcinoma. *Hum Mol Genet* **21**(5): 1190–1200.

Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, Heintzman N, Ren B, Fu XD, Topol EJ, Rosenfeld MG, Frazer KA (2011) 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature* **470**(7333): 264–268.

Hu N, Wang C, Su H, Li WJ, Emmert-Buck MR, Li G, Roth MJ, Tang ZZ, Lu N, Giffen C, Albert PS, Taylor PR, Goldstein AM (2004) High frequency of CDKN2A alterations in esophageal squamous cell carcinoma from a high-risk Chinese population. *Genes Chromosomes Cancer* **39**(3): 205–216.

Hustinx SR, Leoni LM, Yeo CJ, Brown PN, Goggins M, Kern SE, Hruban RH, Maitra A (2005) Concordant loss of MTAP and p16/CDKN2A expression in pancreatic intraepithelial neoplasia: evidence of homozygous deletion in a noninvasive precursor lesion. *Mod Pathol* **18**(7): 959–963.

Mavaddat N, Dunning AM, Ponder BA, Easton DF, Pharoah PD (2009) Common genetic variation in candidate genes and susceptibility to subtypes of breast cancer. *Cancer Epidemiol Biomarkers Prev* **18**(1): 255–259.

Mirabello L, Richards EG, Duong LM, Yu K, Wang Z, Cawthon R, Berndt SI, Burdett L, Chowdhury S, Teshome K, Douglass C, Savage SA (2011) Telomere length and variation in telomere biology genes in individuals with osteosarcoma. *Int J Mol Epidemiol Genet* **21:** 19–29.

Nakanishi H, Wang XL, Imai FL, Kato J, Shiiba M, Miya T, Imai Y, Tanzawa H (1999) Localization of a novel tumor suppressor gene loci on chromosome 9p21-22 in oral cancer. *Anticancer Res* **19**(1A): 29–34.

Neta G, Yu CL, Brenner A, Gu F, Hutchinson A, Pfeiffer R, Sturgis EM, Xu L, Linet MS, Alexander BH, Chanock S, Sigurdson AJ (2012) Common genetic variants in the 8q24 region and risk of papillary thyroid cancer. *Laryngoscope* **122**(5): 1040–1042.

Okami K, Cairns P, Westra WH, Linn JF, Ahrendt SA, Wu L, Sidransky D, Jen J (1997) Detailed deletion mapping at chromosome 9p21 in non-small cell lung cancer by microsatellite analysis and fluorescence *in situ* hybridization. *Int J Cancer* **74**(6): 588–592.

Pasmant E, Sabbagh A, Vidaud M, Bieche I (2011) ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J* **25**(2): 444–448.

Perinchery G, Bukurov N, Nakajima K, Chang J, Li LC, Dahiya R (1999) High frequency of deletion on chromosome 9p21 may harbor several tumor-suppressor genes in human prostate cancer. *Int J Cancer* **83**(5): 610–614.

Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, Fogel R, Gelmann EP, Gilbert F, Hasson MA, Hayes RB, Johnson CC, Mandel JS, Oberman A, O'Brien B, Oken MM, Rafla S, Reding D, Rutt W, Weissfeld JL, Yokochi L, Gohagan JK (2000) Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* **21**(6 Suppl): 273S–309S.

Rakosy Z, Vizkeleti L, Ecsedi S, Begany A, Emri G, Adany R, Balazs M (2008) Characterization of 9p21 copy number alterations in human melanoma by fluorescence in situ hybridization. *Cancer Genet Cytogenet* **182**(2): 116–121.

Roth MJ, Abnet CC, Hu N, Wang QH, Wei WQ, Green L, D'Alelio M, Qiao YL, Dawsey SM, Taylor PR, Woodson K (2006) p16, MGMT, RARbeta2, CLDN3, CRBP and MT1G gene methylation in esophageal squamous cell carcinoma and its precursor lesions. *Oncol Rep* **15**(6): 1591–1597.

Sanchez-Cespedes M, Decker PA, Doffek KM, Esteller M, Westra WH, Alawi EA, Herman JG, Demeure MJ, Sidransky D, Ahrendt SA (2001) Increased loss of chromosome 9p21 but not p16 inactivation in primary non-small cell lung cancer from smokers. *Cancer Res* **61**(5): 2092–2096.

Schmid M, Sen M, Rosenbach MD, Carrera CJ, Friedman H, Carson DA (2000) A methylthioadenosine phosphorylase (MTAP) fusion transcript identifies a new gene on chromosome 9p21 that is frequently deleted in cancer. *Oncogene* **19**(50): 5747–5754.

Shete S, Hosking FJ, Robertson LB, Dobbins SE, Sanson M, Malmer B, Simon M, Marie Y, Boisselier B, Delattre JY, Hoang-Xuan K, El Hallani S, Idbaih A, Zelenika D, Andersson U, Henriksson R, Bergenheim AT, Feychting M, Lonn S, Ahlbom A, Schramm J, Linnebank M, Hemminki K, Kumar R, Hepworth SJ, Price A, Armstrong G, Liu Y, Gu X, Yu R, Lau C, Schoemaker M, Muir K, Swerdlow A, Lathrop M, Bondy M, Houlston RS (2009) Genome-wide association study identifies five susceptibility loci for glioma. *Nat Genet* **41**(8): 899–904.

Stark GR, Kerr IM, Williams BR, Silverman RH, Schreiber RD (1998) How cells respond to interferons. *Annu Rev Biochem* **67**: 227–264.

Stevens AP, Spangler B, Wallner S, Kreutz M, Dettmer K, Oefner PJ, Bosserhoff AK (2009) Direct and tumor microenvironment mediated influences of 5'-deoxy-5'-(methylthio)adenosine on tumor progression of malignant melanoma. *J Cell Biochem* **106**(2): 210–219.

Troisi R, Masters MN, Joshipura K, Douglass C, Cole BF, Hoover RN (2006) Perinatal factors, growth and development, and osteosarcoma risk. *Br J Cancer* **95**(11): 1603–1607.

Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, Maranian M, Seal S, Ghoussaini M, Hines S, Healey CS, Hughes D, Warren-Perry M, Tapper W, Eccles D, Evans DG, Hooning M, Schutte M, van den Ouweland A, Houlston R, Ross G, Langford C, Pharoah PD, Stratton MR, Dunning AM, Rahman N, Easton DF (2010) Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat Genet* **42**(6): 504–507.

van der Riet P, Nawroz H, Hruban RH, Corio R, Tokino K, Koch W, Sidransky D (1994) Frequent loss of chromosome 9p21-22 early in head and neck cancer progression. *Cancer Res* **54**(5): 1156–1158.

Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, Blow MJ, Cohen JC, Rubin EM, Pennacchio LA (2010) Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* **464**(7287): 409–412.

Waber P, Dlugosz S, Cheng QC, Truelson J, Nisen PD (1997) Genetic alterations of chromosome band 9p21-22 in head and neck cancer are not restricted to p16INK4a. *Oncogene* **15**(14): 1699–1704.

Wrensch M, Jenkins RB, Chang JS, Yeh RF, Xiao Y, Decker PA, Ballman KV, Berger M, Buckner JC, Chang S, Giannini C, Halder C, Kollmeyer TM, Kosel ML, LaChance DH, McCoy L, O'Neill BP, Patoka J, Pico AR, Prados M, Quesenberry C, Rice T, Rynearson AL, Smirnov I, Tihan T, Wiemels J, Yang P, Wiencke JK (2009) Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. *Nat Genet* **41**(8): 905–908.

Yang HP, Gonzalez Bosquet J, Li Q, Platz EA, Brinton LA, Sherman ME, Lacey Jr JV, Gaudet MM, Burdette LA, Figueroa JD, Ciampa JG, Lissowska J, Peplonska B, Chanock SJ, Garcia-Closas M (2010a) Common genetic variation in the sex hormone metabolic pathway and endometrial cancer risk: pathway-based evaluation of candidate genes. *Carcinogenesis* **31**(5): 827–833.

Yang XR, Liang X, Pfeiffer RM, Wheeler W, Maeder D, Burdette L, Yeager M, Chanock S, Tucker MA, Goldstein AM (2010b) Associations of 9p21 variants with cutaneous malignant melanoma, nevi, and pigmentation phenotypes in melanoma-prone families with and without CDKN2A mutations. *Fam Cancer* **9**(4): 625–633.

Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM (2010) Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol cell* **38**(5): 662–674.

Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, Caporaso N, Kraft P, Chatterjee N (2009) Pathway analysis by adaptive combination of P-values. *Genet Epidemiol* **33**(8): 700–709.

Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, Maouche S, Germain M, Lackner K, Rossmann H, Eleftheriadis M, Sinning CR, Schnabel RB, Lubos E, Mennerich D, Rust W, Perret C, Proust C, Nicaud V, Loscalzo J, Hubner N, Tregouet D, Munzel T, Ziegler A, Tiret L, Blankenberg S, Cambien F (2010) Genetics and beyond–the transcriptome of human monocytes and disease susceptibility. *PloS one* **5**(5): e10693.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)