International Journal of Epidemiology, 2016, 186–205
doi: 10.1093/ije/dyv185
Advance Access Publication Date: 7 October 2015
Original article

Genetic Epidemiology

Systematic meta-analyses and field synopsis of genetic association studies in colorectal adenomas

Zahra Montazeri,[1†] Evropi Theodoratou,[2†] Christine Nyiraneza,[1] Maria Timofeeva,[3] Wanjing Chen,[2] Victoria Svinti,[3] Shanya Sivakumaran,[2] Gillian Gresham,[1] Laura Cubitt,[2] Luis Carvajal-Carmona,[4] Monica M Bertagnolli,[5] Ann G Zauber,[6] Ian Tomlinson,[7] Susan M Farrington,[3] Malcolm G Dunlop,[3] Harry Campbell[2,3] and Julian Little[1]*

[1]School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, Ottawa, Canada, [2]Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK, [3]Colon Cancer Genetics Group and Academic Coloproctology, Institute of Genetics and Molecular Medicine, University of Edinburgh and MRC Human Genetics Unit Western General Hospital, Edinburgh, UK [4]Biochemistry and Molecular Medicine, Genome and Biomedical Sciences Facility, UC Davis School of Medicine, University of California Davis, Davis, CA, USA, [5]Department of Surgery, Brigham and Women's Hospital, Boston, MA, USA [6]Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA and [7]Wellcome Trust Centre for Human Genetics, Oxford, UK

*Corresponding author. School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Canada. E-mail: jlittle@uottawa.ca
[†]Joint first authors.

Accepted 20 August 2015

Abstract

Background: Low penetrance genetic variants, primarily single nucleotide polymorphisms, have substantial influence on colorectal cancer (CRC) susceptibility. Most CRCs develop from colorectal adenomas (CRA). Here we report the first comprehensive field synopsis that catalogues all genetic association studies on CRA, with a parallel online database [http://www.chs.med.ed.ac.uk/CRAgene/].

Methods: We performed a systematic review, reviewing 9750 titles, and then extracted data from 130 publications reporting on 181 polymorphisms in 74 genes. We conducted meta-analyses to derive summary effect estimates for 37 polymorphisms in 26 genes. We applied the Venice criteria and Bayesian False Discovery Probability (BFDP) to assess the levels of the credibility of associations.

Results: We considered the association with the rs6983267 variant at 8q24 as 'highly credible', reaching genome-wide statistical significance in at least one meta-analysis model. We identified 'less credible' associations (higher heterogeneity, lower statistical power, BFDP $>$ 0.02) with a further four variants of four independent genes: *MTHFR* c.677C$>$T p.A222V

© The Author 2015; all rights reserved. Published by Oxford University Press on behalf of the International Epidemiological Association

(rs1801133), *TP53* c.215C>G p.R72P (rs1042522), *NQO1* c.559C>T p.P187S (rs1800566), and *NAT1* alleles imputed as fast acetylator genotypes. For the remaining 32 variants of 22 genes for which positive associations with CRA risk have been previously reported, the meta-analyses revealed no credible evidence to support these as true associations.

Conclusions: The limited number of credible associations between low penetrance genetic variants and CRA reflects the lower volume of evidence and associated lack of statistical power to detect associations of the magnitude typically observed for genetic variants and chronic diseases. The CRA gene database provides context for CRA genetic association data and will help inform future research directions.

Key Messages

What is already known about this subject?

- Most colorectal cancers (CRC) develop from preneoplastic asymptomatic lesions known as colorectal adenomas (CRA).
- A recent original study found that eight common SNPs associated with CRC, identified through genome-wide association studies (GWAS), also increase the risk of CRA.
- We have previously summarized the associations between common genetic variants and CRC in a field synopsis of genetic association and GWAS, but the genetic basis of CRA is less well documented.

What are the new findings?

- We present here the first synthesis of all published genetic association data for CRAs and the results of meta-analyses to summarize risk estimates.
- Five variants out of 37 meta-analysed SNPs (approximately 14%) are likely to be associated with CRA.
- For the 32 variants of 22 genes for which positive associations with CRA risk have been previously reported, the meta-analyses revealed no credible evidence to support these as true associations.

How might it impact on clinical practice in the foreseeable future?

- The identification of genetic variants for which there is robust evidence of influence on CRA risk may provide new insights into the fundamental biological mechanisms involved in early CRC development.
- Improving our understanding of CRA risk factors may help inform the development of improved strategies for prevention of CRC.
- Findings from this study should help focus further clinical research on understanding the role of gene-gene and gene-environment interactions in the development of colorectal neoplasia.

INTRODUCTION

Colorectal cancer (CRC) constitutes a major public health challenge, with over 1.3 million cases estimated to have been newly diagnosed in 2012, and almost 700 000 deaths from the disease.[1] Most CRCs develop from preneoplastic asymptomatic lesions known as adenomatous polyps. The malignant potential of colorectal adenomas (CRAs) depends on their size, histological characteristics, degree of dysplasia and multiplicity.[2] In addition, serrated lesions, particularly sessile serrated adenomas/polyps (SSA/P), previously thought not to have malignant potential, are also associated with an increased risk of CRC.[3]

Several risk factors have been reported to be associated with risk of developing CRAs. These include an increased risk associated with cigarette smoking,[4] alcohol consumption[5] and obesity[6–9] and a decreased risk associated with regular aspirin intake.[10–12] Improving our understanding of these adenoma risk factors may help inform the development of new strategies for the prevention of CRC.[13]

Although the majority of CRCs arise sporadically, the disease has a clear genetic component as shown by segregation of highly penetrant mutations in genes such as *APC* in families affected by the familial adenomatous polyposis (FAP) syndrome, and mutations in DNA mismatch repair

genes (*MLH1*, *MSH2*, *MSH6* and *PMS2)* in families affected by Lynch syndrome (Hereditary Non Polyposis Colorectal Cancer—HNPCC).[14] However, whereas highly penetrant mutations account for less than 10% of CRC susceptibility, an expanding number of low penetrance genetic variants have been increasingly recognized to influence the risk of colorectal neoplasia. We summarized the contribution of these alleles in a field synopsis of genetic association and genome-wide association studies (GWAS) in CRC.[15] The genetic basis of CRA is less well documented. The risk of colorectal neoplasia in first-degree relatives of a patient with adenomas [relative risk (RR) for advanced adenoma 1.68, 95% confidence interval (CI) 1.29-2.18] compared with controls[16,17] is reported to be of a similar magnitude to the risk of CRC in first-degree relatives of patients with CRC.[18,19] In addition, a recent study investigated whether CRC single nucleotide polymorphisms (SNPs) identified through GWAS also increased the CRA risk, and found that 8 of 18 known CRC-associated SNPs (rs10936599, rs6983267, rs10795668, rs3802842, rs4444235, rs1957636, rs4939827 and rs961253) were over-represented in CRC-free patients with adenomas, compared with controls.[20]

The main objective of the present study was to identify and interpret associations between common genetic variants and CRA risk. The identification of genetic variants for which there is robust evidence of influence on CRA risk may provide new insights into the fundamental biological mechanisms involved in early CRC development and help to inform future research. Further, identification of CRA risk-associated variants may also show utility in contributing to future risk scores for accurate population risk stratification, which could be of potential value in targeting primary prevention and CRC screening modalities. We have previously undertaken a comprehensive review of genetic factors associated with CRC using published guidelines for the assessment of cumulative evidence on genetic association studies[21,22] following a format similar to published overview meta-analyses[23–25] and utilizing an inference framework to aid transparent and objective interpretation of data.[26] We now report the results of a similar exercise for CRA. This represents the first attempt to synthesize all published genetic association data for CRAs and conduct meta-analyses to summarize risk estimates. The search strategy and the results of meta-analyses are publicly available on a regularly updated internet database (CRAgene; http://www.chs.med.ed.ac.uk/CRAgene/).

METHODS

Literature search and data collection

We undertook a comprehensive systematic review of published data on genetics and colorectal polyps using the Medline database via the Ovid gateway. The search strategy is shown in Supplementary Box 1 (available as Supplementary data at *IJE* online). We cross-checked these findings against those listed in the HuGENet phenopedia™. Review articles and meta-analyses on genetic associations of colorectal polyps were also considered so that the references they used could be screened for eligibility, in case they had been missed in the Medline search. The abstracts and if necessary the full texts were screened for eligibility using the following inclusion and exclusion criteria. The paper must have evaluated the association between a polymorphic genetic variant (one with a MAF ≥ 0.01 in the general population based on the data on the reference panel of the 1000 Genomes; Table 1) and sporadic colorectal polyps. Papers studying only CRCs were not included. All studies needed to relate to human participants. Case-control, cohort and GWA studies were included. The study had to be published in English (one Chinese and one Spanish study were also included) in a peer-reviewed journal before 31 March 2014. Any research that had only been reported in abstracts, eg presented in scientific conferences but not yet fully published, was excluded and 14 family-based studies were also excluded. A list with all variants to be summarized using meta-analysis was generated. A second list with all variants with two or more studies was compared with a list of variants that were included in two GWAS (CORGI and APC[27]). If a variant was found to be included in either of these GWAS, then genotype counts were included in the meta-analysis of this variant. Descriptions of the CORGI and APC GWAS are presented in Supplementary Box 2 (available as Supplementary data at *IJE* online).

Data entry, management and abstraction

Once the search was completed, the references of the papers in the search were entered into a web-based database, 'RefWorks' [http://www.refworks.com/]. Data from all studies that met final inclusion and exclusion criteria were abstracted into two standardized tables (Supplementary Table 1, available as Supplementary data at *IJE* online). We abstracted key variables with regard to the study identifiers and context, study design and limitations, intervention specifics (such as whether the cases were ascertained as a result of CRC screening) and outcome details (type of polyp and information on size, histology, dysplasia and multiplicity, if recorded).

Statistical analysis

Statistical analysis was conducted in R, 3.1.0.[28] Meta-analysis was performed for all variants with case-control

Table 1. List of genes and variants that were selected for meta-analysis

Gene	Variant	rs number	Cases of polyps vs controls (number of samples)	Ref allele[a]	Ref allele frequency cases	Ref allele frequency controls	Attributable familial risk	Result of most recent meta-analysis; cases/controls (samples) (reference)	Other meta-analyses
APC	c.5465T>A p.Val1822Asp	rs459552	2805 vs 4369 (4*)	A	0.78	0.77	0.42%	No meta-analyses in colorectal polyps or adenomas only; Liang J *et al.* 2012 meta-analysis on colorectal cancer and adenomas[93]	n/a
BHMT	c.716G>A p.Arg239Gln	rs3733890	3691 vs 12 174 (4*§)	G	0.71	0.70	0.01%	n/a	n/a
PTGS1	c.22T>C p.Trp8Arg	rs1236913	2551 vs 4342 (4*§)	C	0.93	0.93	3.74%	n/a	n/a
PTGS1	c.382C>A p.Leu237Met	rs5789	1337 vs 1928 (3§)	C	0.97	0.97		n/a	n/a
PTGS2	c.-765G>C	rs20417	1507 vs 2042 (4)	C	0.87	0.86	0.70%	n/a	n/a
PTGS2	c.*427T>C	rs5275	2643 vs 4616 (6*§)	C	0.44	0.45	0.18%	n/a	n/a
PTGS2	c.640-275T>G	rs20432	2492 vs 4423 (5*§)	T	0.76	0.76	0.01%	n/a	n/a
CRP	c.*1082G>A	rs1205	2016 vs 3465 (3*)	C	0.59	0.61	0.42%	n/a	n/a
CYP1A1	c.1384A>G p.Ile462Val	rs1048943	1846 vs 1853 (3)	T	0.95	0.96		n/a	n/a
CYP1A1	c.*1189T>C	rs4646903	3126 vs 5079 (5*§)	A	0.87	0.85	0.49%	n/a	n/a
CYP1A2	c.-9-154C>A	rs762551	2689 vs 4373 (4*)	A	0.72	0.73	0.12%	n/a	n/a
GSTM1	Null variant		3952 vs 3514 (7)	Present				No association; 2658/2196 (5)[94]	n/a
GSTP1	c.313A>G p.Ile105Val	rs1695	3945 vs 5320 (6*§)	A	0.64	0.64	0.00%	No association; 2475/1991 (4)[94]	n/a
GSTP1	c.341C>T p.Ala114Val	rs1138272	1316 vs 1973 (3§)	C	0.92	0.91	0.22%	No association; 1008/1010 (2)[94]	n/a
GSTT1	Null variant		4345 vs 3964 (8)	Present				No association; 3119/2719 (6) [94]	n/a
MEH [EPHX1]	c.337T>C p.Tyr113His	rs1051740	7387 vs 8774 (12*§)	T/imputed phenotype Fast	0.69	0.68	0.07%	No association; 4940/5021 (9)[94]	n/a
MEH [EPHX1]	c.416A>G p.His139Arg	rs2234922	7424 vs 8845 (12*§)	A/imputed phenotype Fast	0.80	0.80	0.10%	No association; 4899/4954 (9)[94]	n/a
MTHFR	c.677C>T p.Ala222Val	rs1801133	11362 vs 23 006 (24*§)	C	0.68	0.68	0.01%	No association; 9232/12 676 (11)[95]	n/a
MTHFR	c.1286A>C p.Glu429Ala	rs1801131	6760 vs 15 498 (10*)	A	0.7	0.72	0.08%	No association; 2544/3368 (3)[95]	n/a

(Continued)

Table 1. Continued

Gene	Variant	rs number	Cases of polyps vs controls (number of samples)	Ref allele[a]	Ref allele frequency cases	Ref allele frequency controls	Attributable familial risk	Result of most recent meta-analysis; cases/controls (samples) (reference)	Other meta-analyses
MTRR	c.66A>G p.Ile22Met	rs1801394	2911 vs 9342 (3)	A	0.5	0.51	0.13%	n/a	n/a
MTR	c.2756A>G p.Asp919Gly	rs1805087	4730 vs 13 710 (7*§)	A	0.81	0.81	0.00%	No association; 4872/11 680 (6)[96]	n/a
NAT1[2]			2347 vs 3143 (5)	Imputed slow phenotype	0.54	0.39	1.00%	No association; 1553/2587 (4)[97]	n/a
NAT2[3]			4092 vs 4731 (9)	Imputed Slow phenotype	0.68	0.42	0.18%	No association; 3683/5109 (7)[97]	n/a
NQO1	c.559C>T p.Pro187Ser	rs1800566	4097 vs 5967 (6*§)	C	0.81	0.82	0.51%	Positive association; 2637/2638 (4)[98]	[94]
TP53	c.215C>G p.Arg72Pro	rs1042522	2135 vs 3738 (3*)	C	0.73	0.77	0.05%	n/a	n/a
PPAR-γ	c.34C>G p.Pro12Ala	rs1801282	1730 vs 2953 (3*)	C	0.87	0.88	0.06%	n/a	n/a
SLC19A1	c.80G>A p.His27Arg	rs1051266	3614 vs 11 467 (4*§)	G but T in most cases	0.56	0.56	0.09%	n/a	n/a
TGFB1	c.29C>T p.Pro10Leu	rs1982073	2840 vs 4218 (4*)	T	0.57	0.58	0.01%	No association; 4940/5021 (9)[94]	n/a
TGFB1	c.-1347T>C	rs1800469	2395 vs 3902 (4*)	C	0.63	0.62	0.04%	No association; 1515/1895 (3)[99]	n/a
TGFB1	c.74G>C p.Arg25Pro	rs1800471	1011 vs 1329 (3)	G	0.95	0.96	1.60%	No association; 1246/1539 (3)[99]	n/a
TYMS	TS tandem repeat	rs34743033	1633 vs 2034 (3)	3R (3 repeats)	0.53	0.52		No association; 1054/1391 (3)[99]	n/a
VDR	c.1024+283G>A	rs1544410	2403 vs 4356 (5*§)	G	0.58	0.59	0.00%	n/a	n/a
XPD	c.2251A>C p.Lys751Gln	rs13181	3363 vs 3523 (5§)	T	0.62	0.63	0.03%	No association; 1740/2216 (5; includes studies with first and recurrent adenoma)[100]	n/a
XRCC1	c.580C>T p.Arg194Trp	rs1799782	4821 vs 5972 (5*)	C	0.92	0.93	0.28%	n/a	n/a
XRCC1	c.1196A>G p.Arg399Gln	rs25487	5124 vs 6927 (6*§)	G	0.60	0.59	0.32%	n/a	n/a

(Continued)

Table 1. Continued

Gene	Variant	rs number	Cases of polyps vs controls (number of samples)	Ref allele[a]	Ref allele frequency cases	Ref allele frequency controls	Attributable familial risk	Result of most recent meta-analysis; cases/controls (samples) (reference)	Other meta-analyses
XRCC3	c.722C>T p.Thr241Met	rs861539	3183 vs 4514 (4*$)	C	0.65	0.64	0.04%	n/a	n/a
8q24.21	region 3	rs6983267	3559 vs 9586 (8)	T	0.44	0.49	2.02%	Positive association; 8148/17 065 (7)[101]	[102]

*Includes unpublished data from the APC trial.
§Includes unpublished data from CORGI.
[a]The reference allele for each SNP was selected according to what was reported in the original studies. If there was any conflict between studies we referred to the Single Nucleotide Polymorphism database of NCBI (dbSNP) and checked the results from the 1000-Genome project, which presents the reference allele for each SNP for different populations.
[b]Genotypes containing the NAT1*10 allele are associated with high NAT1 enzymatic activity, thus representing the fast acetylator phenotype.
[c]The proportion of slow and rapid metabolizers is known to differ between different ethnic populations. In general, the slow metabolizer phenotype is most prevalent (> 80%) in Northern Africans and Scandinavians, and lowest (5%) in Canadian Eskimos and Japanese. Intermediate frequencies are seen in Chinese populations (around 20% slow metabolizers), whereas 40–60% of African-Americans and most non-Scandinavian Caucasians are slow metabolizers.

data available from three or more independent samples. The reference allele for each SNP was selected according to what was reported in the original studies. If there was any conflict between studies we referred to the Single Nucleotide Polymorphism database of NCBI (dbSNP) and checked the results from the 1000-Genome project, which presents the reference allele for each SNP for different populations. We then selected the most frequent allele as the reference allele. We obtained summary crude odds ratios (ORs) and 95% confidence intervals (95% CI) for two genotypic models (var/wt vs wt/wt and var/var vs wt/wt), one recessive (var/var vs var/wt and wt/wt) and one dominant model (var/var and var/wt vs wt/wt). We applied either the fixed effect model (Mantel-Haenszel method) or, in case of heterogeneity the random effect model (DerSimonian-Laird method). Between-study heterogeneity was quantified by calculating the Q statistic with a P-value less than 0.10 being the threshold. We also calculated the I^2 heterogeneity metric and its 95% CI. Although in some cases we summarized studies that were very heterogeneous, it is recognized that, due to the variation in study methods and outcome definitions, the meta-estimates should be interpreted cautiously. To assess for any small-study effects, we performed funnel plot analysis and tested for significance using the Harbord modification of the Egger test. A negative result for small-study effects testing does not entirely exclude publication bias. In addition, the test for small-study effects may be underpowered with ~ 10 or less studies and may be inappropriate in the presence of large heterogeneity.[29] We also estimated the power that each meta-analysis had in order to detect a statistical significant effect, using the Power and Sample Size Programme[30] and based on a level of significance $\alpha = 0.05$ and the effect sizes and allele frequencies estimated from the meta-analyses (integral component of the Bayesian False Discovery Probability [BFDP] analysis). Finally, we tested whether any of the examined SNPs were in linkage disequilibrium by using the SNP Annotation and Proxy Search (SNAP) tool from the Broad Institute.[31]

The sibling relative risk attributable to a given SNP was calculated using the following formula [32,33]:

$$\lambda^* = \frac{p(pr_2 + qr_1)^2 + q(pr_1 + q)^2}{(p^2 r_2 + 2pqr_1 + q^2)^2}$$

where p is the population frequency of the referent allele, $q = 1 - p$, and $r1$ and $r2$ are the relative risks [estimated as odds ratios (ORs) from the meta-analyses] for heterozygotes and variant homozygotes, relative to wild type (wt) homozygotes. Assuming a multiplicative interaction, we calculated the proportion of the familial risk attributable to an SNP as $\log(\lambda^*)/\log(\lambda_0)$, where λ_0 is the overall

familial relative risk estimated from epidemiological studies, assumed to be 1.7.[34] Finally, when information on polyp type was available, we repeated the analysis for CRA only.

Credibility of genetic association

In assessing the credibility of genetic associations, we considered the BFDP [35] and the Venice criteria.[21,22] The BFDP assesses the noteworthiness of an observed association and was estimated using the Excel Calculation Spreadsheet [http://faculty.washington.edu/jonno/cv.html]. The BFDP threshold for noteworthiness was set up to be equal to 0.20, based on the assumption that a false discovery would be four times more costly than a false non-discovery. We chose to calculate BFDP values for two levels of prior probabilities: at a medium/low prior level ($0.05\text{-}10^{-3}$) that would be close to what would be expected for a candidate gene, and at a very low prior level (10^{-4} to 10^{-6}) that would be close to what would be expected for a random SNP. For the volume of evidence, replication and protection against bias Venice criteria, we used the same strategy as in the CRC field synopsis.[15] With regard to the Venice criteria, we operationalized the criterion of volume of evidence on the basis of statistical power to detect an association of the desired magnitude: A, 80% or more; B, 50–79%; or C, less than 50%. For replication, we used the I^2 criterion proposed by Ioannidis *et al.*[21]: A, $I^2 < 25\%$; B, I^2 25–50%; C, $I^2 > 50\%$. For protection against bias, we considered that the completeness of reporting was problematic. The phenotype definition was addressed by our inclusion criterion—namely, that case subjects would have colorectal polyps or adenomas, in the latter instance histologically confirmed. In general, genotyping error rates are low,[36] and the criterion of replication across studies in part addressed potential concern about variation in genotyping quality between studies. Whereas population stratification may impact on gene discovery,[37,38] the effect on the magnitude of association in general appears to be small.[39,40] A priori, we sought to classify the genetic associations into one of three categories according to the findings of the BFDP analysis and the application of the Venice criteria. First, associations were to be classified as of 'high credibility' if they fulfilled the following criteria: (i) they were statistically significant at a *P*-value level of 0.05 in at least two of the genetic models; (ii) they had a BFDP less than 0.20 at least at the *P*-value level of 0.05; (iii) they had a statistical power greater than 80%; and (iv) they had an I^2 less than 50%. Second, a 'less-credible' association was: (i) statistically significant at a *P*-value threshold of 0.05 in at least one of the genetic models; but (ii) its BFDP was greater than 0.20; and (iii) its statistical power was 50–79% (I^2 ranged from 0% to 48% for this category, but this criterion was not taken into account for this category). Third, all other associations were classified as negative.

RESULTS

Study characteristics

After screening 9750 titles and abstracts, 1750 publications were identified as potentially eligible, of which 130 articles met the inclusion criteria. Data were extracted from these 130 articles, reporting on 181 polymorphisms in 74 genes (Supplementary Table 2). The 130 articles related to 84 independent studies; 29 (35%) of these studies were published in the period 1995–2004, 42 (50%) in 2005–09 and 13 (15%) since 2010. Of the 84 independent studies, 49 (58%) were population-based, 30 (36%) hospital-based studies and for 5 (6%) this was unclear. Overall 67 (80%) studies related to populations of European origin, 7 (8%) Asian, 1 (1%) Mexican, 2 (2%) African and 5 (6%) related to more than one of these groups; in two studies, the population was not specified. In almost all studies, the polyps were confirmed histologically. In 9 (11%) the data on adenomas and hyperplastic polyps were reported separately but this was not done in the other studies. More information on the characteristics of these studies is presented in the CRAgene database.

Meta-analysis results

Separate meta-analysis was undertaken for variants for which data were available from at least three case-control studies. Thus, meta-analyses are reported for 37 polymorphisms in 26 different genes, with a mean of 3501 cases [median = 2911; interquartile range (IQR) = 4092-2347 = 1745] and 5982 controls (median = 4373; IQR = 6927-3514 = 3413) for each variant. Individual meta-analysis was based on a mean of 6 case-control studies (median = 4; IQR = 6−4 = 2; Table 1). In addition, unpublished data from APC trial for 23 SNPs and from CORGI for 18 SNPs, were included in these analyses. Overall summary results including crude odd ratios (ORs), their 95% confidence interval (95% CI) and *P*-value along with measures of heterogeneity (I^2, 95% confidence interval and *P*-value of the Q test) are presented in Tables 2 and 3. Table 2 shows the results of the meta-analyses based on two genotypic models (variant/ wild type vs wild type/ wild type and variant/variant vs wild type/wild type); and Table 3 presents the results of meta-analyses based on the recessive model (variant/variant vs wild type/wild type and wild type/variant) and the dominant model (wild type/variant and variant/variant vs wild type/wild type). Individual study results and the overall summary results, (OR; 95%

Table 2. Summary crude odds ratios (ORs) and 95% confidence intervals (95% CIs) of genotypic model of association between CRA and genetic variants identified for meta-analysis, with credibility factors

Gene/variant	Cases vs controls (number of samples)	var/wt VS wt/wt						var/var VS wt/wt						Credibility	
		N	Effect size		Heterogeneity			*N*	Effect size		Heterogeneity				
			OR (95% CI)	*P*-value	I^2 (95% CI)	*P*-value	Power		OR (95% CI)	*P*-value	I^2 (95% CI)	*P*-value	Power	BFDP[a]	Venice criteria grade[b]
APC/c.5465T>A p.Val1822Asp	2805 vs 4369 (4*)	4	0.99 (0.89 −1.09)	0.79	0 (0 −89)	0.64	0.05	4	0.86 (0.68 −1.07)	0.18	0 (0 −90)	0.59	0.26	0.99	CAB
BHMT/c.716G>A p.Arg239Gln	3691 vs 12 174 (4*§)	4	0.94 (0.87 −1.02)	0.16	11 (0 −97)	0.34	0.34	4	0.94 (0.81 −1.07)	0.34	34 (0 −98)	0.21	0.14	0.97	CBB
PTGS1/c.22T>C p.Trp8Arg	2551 vs 4342 (4*§)	4	1.06 (0.91 −1.23)	0.48	0 (0 −95)	0.45	0.13	4	0.64 (0.28 −1.48)	0.30	0 (0 −47)	0.93	0.19	0.98	CAB
PTGS1/c.382C>A p.Leu237Met	1337 vs 1928 (3§)	2	n/a	n/a	n/a	n/a	n/a	2	n/a	n/a	n/a	n/a	n/a	n/a	
PTGS2/c.-765G>C	1507 vs 2042 (4)	4	0.88 (0.73 −1.05)	0.16	31 (0 −96)	0.22	0.28	3	0.74 (0.44 −1.23)	0.24	0 (0 −98)	0.40	0.20	0.95	CBB
PTGS2/c.*427T>C	2643 vs 4616 (6*§)	6	1.02 (0.88 −1.19)	0.79	0 (0 −86)	0.45	0.06	6	1.14 (0.97 −1.33)	0.11	0 (0 −75)	0.51	0.41	0.98	CCB
PTGS2/c.640-275T>G	2492 vs 4423 (5*§)	5	0.95 (0.85 −1.07)	0.41	41 (0 −98)	0.15	0.15	5	0.99 (0.74 −1.34)	0.97	0 (0 −81)	0.59	0.05	0.98	CBB
CRP/c.*1082G>A	2016 vs 3465 (3)	3	0.96 (0.85 −1.08)	0.53	15 (0 −99)	0.31	0.10	3	1.15 (0.77 −1.71)	0.49	77 (9 −99)	0.02	0.36	0.98	CCB
CYP1A1/c.1384A>G p.Ile462Val	1846 vs 1853 (3)	3	1.05 (0.85 −1.30)	0.66	22 (0 −99)	0.28	0.07	2	n/a	n/a	n/a	n/a	n/a	0.97	CAB
CYP1A1/c.*1189T>C	3126 vs 5079 (5*§)	5	0.90 (0.79 −1.02)	0.11	0 (0 −89)	0.5	0.41	5	0.78 (0.52 −1.16)	0.21	0 (0 −0)	0.99	0.25	0.95	CAB
CYP1A2/c.-9-154C>A	2689 vs 4373 (4*)	4	1.00 (0.90 −1.11)	0.96	0 (0 −96)	0.5	0.05	4	1.08 (0.89 −1.31)	0.42	0 (0 −79)	0.86	0.13	0.99	CAB
GSTM1	3952 vs 3514 (7)	n/a	n/a	n/a	n/a	n/a	n/a	n	n/a	n/a	n/a	n/a	n/a	n/a	
GSTP1/c.313A>G p.Ile105Val	3945 vs 5320 (6*§)	6	0.99 (0.91 −1.09)	0.89	0 (0 −83)	0.43	0.06	6	1.00 (0.87 −1.15)	0.99	0 (n/a −n/a)	0.98	0.05	0.99	CAB
GSTP1/c.341C>T p.Ala114Val	1316 vs 1973 (3§)	3	0.86 (0.70 −1.05)	0.14	0 (0 −82)	0.85	0.32	3	0.77 (0.28 −2.07)	0.60	0 (0 −96)	0.38	0.09	0.95	CAB
GSTT1	4345 vs 3964 (8)	n/a	n/a	n/a	n/a	n/a	n/a	n	n/a	n/a	n/a	n/a	n/a	n/a	
MEH/c.337T>C p.Tyr113His	7387 vs 8774 (12*§)	11	0.96 (0.89 −1.03)	0.24	0 (0 −57)	0.74	0.22	11	0.92 (0.82 −1.03)	0.15	16 (0 −77)	0.30	0.34	0.98	CAB
MEH/c.416A>G p.His139Arg	7424 vs 8845 (12*§)	12	1.00 (0.93 −1.07)	0.97	0 (0 −56)	0.61	0.05	12	1.10 (0.93 −1.29)	0.26	14 (0 −72)	0.30	0.23	0.99	CAB
***MTHFR*/c.677C>T p.Ala222Val**	11 362 vs 23 006 (24*§)	22	**0.95 (0.90 −1.00)**	**0.04**	0 (0 −24)	0.92	0.69	21	0.95 (0.88 −1.03)	0.23	0 (0 −27)	0.92	0.25	0.96	BAB
MTHFR/c1286A>C p.Glu429Ala	6760 vs 15 498 (10*)	9	1.02 (0.95 −1.09)	0.65	0 (0 −60)	0.79	0.09	9	1.08 (0.97 −1.20)	0.16	0 (0 −92)	0.65	0.36	0.99	CAB
MTRR/c.66A>G p.Ile22Met	2911 vs 9342 (3)	3	0.94 (0.84 −1.06)	0.31	0(0 −94)	0.60	0.21	3	0.90 (0.79 −1.03)	0.12	0 (0 −97)	0.59	0.44	0.98	CAB

(Continued)

Table 2. Continued

Gene/variant	Cases vs controls (number of samples)	var/wt VS wt/wt						var/var VS wt/wt						Credibility	
		N	Effect size		Heterogeneity			*N*	Effect size		Heterogeneity				
			OR (95% CI)	*P*-value	I^2 (95% CI)	*P*-value	Power		OR (95% CI)	*P*-value	I^2 (95% CI)	*P*-value	Power	BFDP[a]	Venice criteria grade[b]
MTR/c.2756A>G p.Asp919Gly	4730 vs 13 710 (7*§)	7	1.02 (0.95 −1.10)	0.53	0 (0 −65)	0.93	0.08	7	1.00 (0.84 −1.20)	0.97	0 (0 −77)	0.51	0.05	0.99	CAB
NAT1	2347 vs 3143 (5)	5	1.14 (0.98 −1.34)	0.10	20 (0 −94)	0.29	0.51	4	**1.35 (1.02 −1.78)**	**0.04**	22 (0 −95)	0.28	0.98	0.95	BAB
NAT2	4092 vs 4731 (9)	9	0.99 (0.92 −1.08)	0.84	31 (0 −78)	0.17	0.06	6	0.86 (0.74 −1.00)	0.06	0 (0 −72)	0.72	0.53	0.99	BBB
***NQO1*/c.559C>T p.Pro187Ser**	4097 vs 5967 (6*§)	6	1.08 (0.95 −1.23)	0.25	48 (0 −93)	0.09	0.41	6	**1.26 (1.00 −1.59)**	**0.05**	0 (0 −77)	0.72	0.51	0.97	BBB
***TP53*/c.215C>G p.Arg72Pro**	2135 vs 3738 (3*)	3	**1.16 (1.04 −1.30)**	**0.008**	27 (0 −99)	0.25	0.74	3	1.17 (0.94 −1.47)	0.17	0 (0 −99)	0.17	0.29	0.77	BBB
PPARG/c.34C>G p.Pro12Ala	1730 vs 2953 (3*)	3	1.08 (0.94 −1.25)	0.29	0 (0 −76)	0.87	0.19	3	1.14 (0.73 −1.80)	0.56	0 (0 −99)	0.37	0.09	0.97	CAB
SLC19A1/c.80G>A p.His27Arg	3614 vs 11 467 (4*§)	4	0.99 (0.90 −1.08)	0.79	0 (0 −95)	0.40	0.06	4	1.08 (0.85 −1.35)	0.53	73 (22 −99)	0.01	0.32	0.99	CCB
TGFB1/c.29C>T p.Pro10Leu	2840 vs 4218 (4*)	4	1.06 (0.95 −1.18)	0.29	0 (0 −88)	0.52	0.19	4	1.04 (0.89 −1.20)	0.63	17 (0 −95)	0.30	0.09	0.98	CAB
TGFB1/c.-1347T>C	2395 vs 3902 (4*)	4	1.05 (0.94 −1.17)	0.38	47 (0 −96)	0.13	0.14	4	1.07 (0.90 −1.28)	0.45	39 (0 −96)	0.18	0.12	0.98	CBB
TGFB1/c.74G>C p.Arg25Pro	1011 vs 1329 (3)	3	1.04 (0.78 −1.38)	0.80	0 (0 −95)	0.48	0.06	3	1.97 (0.72 −5.37)	0.18	0 (0 −95)	0.57	0.27	0.97	CAB
TYMS/TS tandem repeat	1633 vs 2034 (3)	3	0.90 (0.77 −1.06)	0.20	0 (0 −97)	0.45	0.27	3	0.94 (0.77 −1.13)	0.49	0 (0 −95)	0.60	0.10	0.96	CAB
VDR/c.1024+283G>A	2403 vs 4356 (5*§)	5	1.01 (0.91 −1.14)	0.80	0 (0 −82)	0.61	0.05	5	1.00 (0.86 −1.15)	0.96	0 (0 −88)	0.53	0.05	0.98	CAB
XPD/c.2251A>C p.Lys751Gln	3363 vs 3523 (5§)	5	1.06 (0.96 −1.18)	0.26	0 (0 −87)	0.46	0.20	5	1.07 (0.82 −1.39)	0.63	64 (0 −96)	0.03	0.15	0.98	CCC
XRCC1/c.580C>T p.Arg194Trp	4821 vs 5972 (5*)	5	1.09 (0.98 −1.23)	0.12	0 (0 −80)	0.54	0.33	5	1.29 (0.74 −2.23)	0.37	0 (0 −84)	0.42	0.16	0.96	CAB
XRCC1/c.1196A>G p.Arg399Gln	5124 vs 6927 (6*§)	6	0.98 (0.90 −1.07)	0.61	0 (0 −69)	0.77	0.08	6	0.86 (0.72 −1.03)	0.11	53 (0 −92)	0.06	0.52	0.99	BCB
XRCC3/c.722C>T p.Thr241Met	3183 vs 4514 (4*§)	4	0.92 (0.83 −1.02)	0.10	0 (0 −77)	0.82	0.39	4	0.92 (0.79 −1.07)	0.29	0 (0 −55)	0.93	0.21	0.96	CAB
8q24.21/rs6983267	3559 vs 9586 (8)	8	**1.20 (1.08 −1.33)**	**0.0006**	0 (0 −78)	0.53	0.99	8	**1.51 (1.27 −1.79)**	**$2.1x10^{-6}$**	45 (0 −88)	0.08	1.00	0.35	ABB

Note: The associations which were considered to be "positives" or "less credible positives" are indicated in bold type.

*Includes unpublished data from the APC trial.

§Includes unpublished data from CORGI.

[a]BFDP analysis for the model of var/wt vs wt/wt for the 0.05 level.

[2b]Venice criteria grade for the model of var/wt vs wt/wt. For the third criterion (protection from bias), all meta-analyses of candidate gene association studies were scored with B. There was no obvious bias in the studies included, but there was considerable missing information on the generation of evidence.

Table 3. Summary crude odds ratios (ORs) and 95% confidence intervals (95% CIs) for recessive and dominant models of association between CRA and variants identified for meta-analysis

Gene/variant	Cases vs controls (number of samples)	RECESSIVE MODEL: var/var vs wt/wt & wt/var						DOMINANT MODEL: wt/var & var/var vs wt/wt					
		N	Effect size		Heterogeneity		Power	*N*	Effect size		Heterogeneity		Power
			OR (95% CI)	*P*-value	I^2 (95% CI)	*P*-value			OR (95% CI)	*P*-value	I^2 (95% CI)	*P*-value	
APC/c.5465T>A p.Val1822Asp	2805 vs 4369 (4*)	4	0.86 (0.69 −1.07)	0.18	0 (0 −93)	0.44	0.25	4	0.97 (0.88 −1.07)	0.53	0 (0 −67)	0.90	0.09
BHMT/c.716G>A p.Arg239Gln	3691 vs 12 174 (4*§)	4	0.97 (0.77 −1.22)	0.80	60 (0 −99)	0.06	0.07	4	0.94 (0.87 −1.02)	0.13	0 (0 −84)	0.84	0.38
PTGS1/c.22T>C p.Trp8Arg	2551 vs 4342 (4*§)	4	0.64 (0.28 −1.47)	0.29	0 (0 −45)	0.94	0.18	4	1.04 (0.90 −1.21)	0.60	0 (0 −95)	0.44	0.09
PTGS1/c.382C>A p.Leu237Met	1337 vs 1928 (3§)	2	n/a	n/a	n/a	n/a	n/a	3	1.03 (0.74 −1.42)	0.87	0 (0 −91)	0.79	0.06
PTGS2/c.-765G>C	1507 vs 2042 (4)	3	0.76 (0.46 −1.27)	0.30	4 (0 −98)	0.35	0.21	4	0.86 (0.72 −1.03)	0.10	21 (0 −96)	0.28	0.46
PTGS2/c.*427T>C	2643 vs 4616 (6*§)	6	1.08 (0.90 −1.30)	0.39	54 (0 −94)	0.05	0.31	6	1.06 (0.92 −1.22)	0.44	0 (0 −78)	0.57	0.16
PTGS2/c.640-275T>G	2492 vs 4423 (5*§)	5	1.12 (0.85 −1.49)	0.43	25 (0 −90)	0.25	0.2	5	0.96 (0.86 −1.08)	0.50	4 (0 −95)	0.39	0.13
CRP/c.*1082G>A	2016 vs 3465 (3)	3	1.14 (0.84 −1.55)	0.40	67 (0 −99)	0.05	0.44	3	1.03 (0.82 −1.29)	0.80	65 (0 −99)	0.06	0.08
CYP1A1/c.1384A>G p.Ile462Val	1846 vs 1853 (3)	2	n/a	n/a	n/a	n/a	n/a	3	1.04 (0.84 −1.28)	0.73	35 (0 −99)	0.22	0.07
CYP1A1/c.*1189T>C	3126 vs 5079 (5*§)	5	0.79 (0.54 −1.15)	0.22	0 (0 −0)	0.99	0.27	5	0.89 (0.99 −1.01)	0.06	0 (0 −89)	0.50	0.57
CYP1A2/c.-9-154C>A	2689 vs 4373 (4*)	4	1.08 (0.89 −1.30)	0.45	0 (0 −80)	0.88	0.14	4	1.01 (0.92 −1.12)	0.79	0 (0 −96)	0.51	0.06
GSTM1	3952 vs 3514 (7)	n/a	n/a	n/a	n/a	n/a	n/a	7	1.03 (0.94 −1.13)	0.50	26 (0 −88)	0.23	0.10
GSTP1/c.313A>G p.Ile105Val	3945 vs 5320 (6*§)	6	1.00 (0.88 −1.15)	0.95	0 (n/a −n/a)	0.99	0.05	6	1.00 (0.91 −1.09)	0.91	0 (0 −80)	0.52	0.06
GSTP1/c.341C>T p.Ala114Val	1316 vs 1973 (3§)	3	0.78 (0.29 −2.12)	0.63	0 (0 −96)	0.39	0.09	3	0.85 (0.70 −1.04)	0.12	0 (0 −91)	0.71	0.35
GSTT1	4345 vs 3964 (8)	n/a	n/a	n/a	n/a	n/a	n/a	8	1.06 (0.95 −1.19)	0.31	0 (0 −68)	0.32	0.21
MEH/c.337T>C p.Tyr113His	7387 vs 8774 (12*§)	11	0.95 (0.85 −1.05)	0.31	15 (0 −77)	0.30	0.16	12	0.96 (0.90 −1.03)	0.23	0 (0 −58)	0.69	0.25
MEH/c.416A>G p.His139Arg	7424 vs 8845 (12*§)	12	1.10 (0.94 −1.29)	0.25	14 (0 −72)	0.31	0.23	12	1.01 (0.94 −1.08)	0.78	0 (0 −57)	0.54	0.06
***MTHFR*/c.677C>T p.Ala222Val**	11362 vs 23 006 (24*§)	23	0.98 (0.91 −1.06)	0.62	0 (0 −24)	0.93	0.08	22	**0.95 (0.90 −1.00)**	**0.04**	0 (0 −13)	0.95	0.77
MTHFR/c1286A>C p.Glu429Ala	6760 vs 15 498 (10*)	10	1.07 (0.97 −1.18)	0.20	0 (0 −88)	0.73	0.28	9	1.03 (0.97 −1.10)	0.37	0 (0 −73)	0.72	0.17
MTRR/c.66A>G p.Ile22Met	2911 vs 9342 (3)	3	0.93 (0.84 −1.02)	0.13	0 (0 −97)	0.58	0.31	3	0.92 (0.83 −1.02)	0.12	0 (0 −96)	0.6	0.41
MTR/c.2756A>G p.Asp919Gly	4730 vs 13 710 (7*§)	7	1.00 (0.83 −1.19)	0.96	0 (0 −76)	0.52	0.05	7	1.03 (0.96 −1.11)	0.43	0 (0 −78)	0.75	0.13

(Continued)

Table 3. Continued

Gene/variant	Cases vs controls (number of samples)	RECESSIVE MODEL: var/var vs wt/wt & wt/var						DOMINANT MODEL: wt/var & var/var vs wt/wt					
		N	Effect size		Heterogeneity		Power	*N*	Effect size		Heterogeneity		Power
			OR (95% CI)	*P*-value	I^2 (95% CI)	*P*-value			OR (95% CI)	*P*-value	I^2 (95% CI)	*P*-value	
NAT1	2347 vs 3143 (5)	4	1.10 (0.95 −1.26)	0.20	0 (0 −79)	0.79	0.30	4	1.12 (0.96 −1.32)	0.16	41 (0 −98)	0.18	0.50
NAT2	4092 vs 4731 (9)	8	0.88 (0.76 −1.02)	0.09	0 (0 −57)	0.88	0.42	6	0.97 (0.90 −1.05)	0.48	41 (0 −83)	0.11	0.11
***NQO1*/c.559C>T p.Pro187Ser**	4097 vs 5967 (6*§)	6	1.24 (0.98 −1.56)	0.07	0 (0 −78)	0.69	0.59	6	**1.10 (1.01 −1.20)**	**0.03**	39 (0 −92)	0.15	0.61
***TP53*/c.215C>G p.Arg72Pro**	2135 vs 3738 (3*)	3	1.10 (0.88 −1.37)	0.39	0 (0 −97)	0.63	0.15	3	**1.16 (1.05 −1.30)**	**0.005**	38 (0 −99)	0.20	0.78
PPARG/c.34C>G p.Pro12Ala	1730 vs 2953 (3*)	3	1.12 (0.71 −1.77)	0.61	0 (0 −99)	0.38	0.07	3	1.09 (0.94 −1.25)	0.25	0 (0 −90)	0.77	0.23
SLC19A1/c.80G>A p.His27Arg	3614 vs 11 467 (4*§)	4	1.07 (0.87 −1.31)	0.53	74 (24 −99)	0.009	0.30	4	1.01 (0.93 −1.09)	0.87	41 (0 −98)	0.16	0.06
TGFB1/c.29C>T p.Pro10Leu	2840 vs 4218 (4*)	4	1.02 (0.83 −1.25)	0.87	55 (0 −97)	0.08	0.06	4	1.04 (0.94 −1.15)	0.42	0 (0 −72)	0.88	0.12
TGFB1/c.-1347T>C	2395 vs 3902 (4*)	4	1.06 (0.90 −1.26)	0.46	35 (0 −96)	0.20	0.13	4	1.06 (0.95 −1.18)	0.28	40 (0 −96)	0.17	0.19
TGFB1/c.74G>C p.Arg25Pro	1011 vs 1329 (3)	3	1.97 (0.72 −5.37)	0.19	0 (0 −94)	0.58	0.20	3	1.09 (0.83 −1.43)	0.55	14 (0 −97)	0.31	0.1
TYMS/TS tandem repeat	1633 vs 2034 (3)	3	1.00 (0.85 −1.17)	0.98	0 (0 −75)	0.90	0.05	3	0.91 (0.79 −1.06)	0.22	0 (0 −97)	0.42	0.25
VDR/c.1024+283G>A	2403 vs 4356 (5*§)	5	0.99 (0.87 −1.13)	0.85	0 (0 −80)	0.78	0.05	5	1.01 (0.91 −1.12)	0.85	0 (0 −88)	0.47	0.05
XPD/c.2251A>C p.Lys751Gln	3363 vs 3523 (5§)	5	1.06 (0.96 −1.17)	0.26	37 (0 −92)	0.18	0.14	5	1.03 (0.82 −1.30)	0.78	60 (0 −95)	0.04	0.09
XRCC1/c.580C>T p.Arg194Trp	4821 vs 5972 (5*)	5	1.27 (0.74 −2.20)	0.39	0 (0 −84)	0.40	0.18	5	1.10 (0.98 −1.23)	0.09	0 (0 −75)	0.66	0.41
XRCC1/c.1196A>G p.Arg399Gln	5124 vs 6927 (6*§)	6	0.88 (0.73 −1.05)	0.16	64 (6 −95)	0.02	0.70	6	0.96 (0.88 −1.04)	0.28	0 (0 −75)	0.68	0.19
XRCC3/c.722C>T p.Thr241M	3183 vs 4514 (4*§)	4	0.97 (0.84 −1.11)	0.63	0 (0 −64)	0.92	0.07	4	0.92 (0.83 −1.01)	0.08	0 (0 −72)	0.85	0.43
8q24.21/rs6983267	3559 vs 9586 (8)	8	**1.31 (1.14 −1.52)**	**0.0003**	55 (0 −90)	0.03	1.00	8	**1.30 (1.18 −1.44)**	**$1.9x10^{-7}$**	14 (0 −83)	0.32	1.00

Note: The associations which were considered to be "positives" or "less credible positives" are indicated in bold type.

*Includes unpublished data from the APC trial.

§Includes unpublished data from CORGI.

CI) for each variant and each model are shown in forest plots in Supplementary Figures 1–37 (available as Supplementary data at *IJE* online). Funnel plots and the results of the associated Egger test for possible small-study effects for each gene are displayed in Supplementary Figures 38–74 (available as Supplementary data at *IJE* online). The results of two additional assessments of the credibility of genetic associations, the Venice criteria[21,22] and BFDP,[35] are shown in Supplementary Table 3 (available as Supplementary data at *IJE* online). None of the examined variants was in high linkage disequilibrium (LD) with one other ($r^2 > 0.8$). The majority of the r^2 values were less than 0.4 and the highest reported r^2 value (0.62) was between the *TGFB1* c.-1347T>C (rs1800469) and the *TGFB1* c.74G>C p.Arg25Pro (rs1800471) variants. Strong heterogeneity ($I^2 > 50\%$) among studies was observed for six variants, namely *BHMT* c.716G>A p.Arg239Gln (rs3733890), *CRP* c.*1082G>A (rs1205), *SLC19A1* c.80G>A p.His27Arg (rs1051266), *TGFB1* c.29C>T p.Pro10Leu (rs1982073), *XPD* c.2251A>C p.Lys751Gln (rs13181) and *XRCC1* c.1196A>G p.Arg399Gln (rs25487).

We considered the association with the rs6983267 variant at 8q24.21 as 'highly credible', reaching genome-wide statistical significance in at least one meta-analysis model. We also identified four other variants in four genes—*MTHFR* c.677C>T p.A222V (rs1801133), *NAT1* (genotypes containing the NAT1*10 allele associated with high NAT1 enzymatic activity, thus representing the fast acetylator phenotype),[41,42] *NQO1* c.559C>T p.Pro187Ser (rs1800566), and *TP53* c.215C>G p.Arg72Pro (rs1042522) as 'less credible', since they were significant in at least one model with statistical power between 50% and 79% and the BFDP was greater than 0.2. These results are based on 2135 to 11 362 cases with median of 3559 cases per study; the range for number of controls is 3143 to 23 006 with median of 5967 controls per study. Thus, five variants out of 37 meta-analysed SNPs (approximately 14%) are likely to be associated with CRA (Tables 2 and 3).

DISCUSSION

To the best of our knowledge, this study represents the first systematic overview of genetic association data for CRA. We systematically analysed data on 37 variants of 26 independent genes. Of the 37 variants analysed, genotypes data had been reported in at least three studies for 23 SNPs. For the remaining 14 SNPs, genotype data were reported in two studies and additional data were obtained from one GWAS for seven variants and two GWAS for the other seven variants.

We compared the results of CRA and CRC field synopses for the two genotypic models (var/wt vs wt/wt and var/var vs wt/wt) (Table 4). The strength of evidence of genetic risk factors associated with the occurrence of CRA seems substantially lower compared with CRC, generally because the volume of evidence is lower (Table 4). The CRA meta-analyses included a median of 2911 cases and 4373 controls originating from six studies, compared with the CRC meta-analyses which included a median of 5281 cases and 6484 controls originating from eight studies.[15] Five SNP variants previously linked with CRC risk or already known gene targets involved in cancer pathways were found to be associated with CRA risk with 'high or low credibility'.

The rs69832687 variant, mapped to the 8q24.21 locus previously identified in CRC GWAS, was classified as 'highly credible', with moderate heterogeneity. By comparing 3559 case patients and 9586 control individuals from eight studies, a positive association between the heterozygosity and homozygosity for the G allele of rs69832687 and CRA risk was observed, in all examined models. This finding parallels the highly credible association observed for this variant in our previous meta-analyses in CRC.[15] The chromosome region 8q24.21 is recognized as a potential susceptibility locus for various cancers.[43,44] This locus is a gene desert region but harbours a number of susceptibility SNPs spanning about 800 kb. The nearest protein-coding gene in this region is the *MYC* proto-oncogene (the rs6983267 SNP mapped 335 kb downstream),[45,46] a target gene of Wnt signalling, a pathway constitutively activated in early development of most CRCs.[47] Although the causative mechanisms conferring the rs69832687 SNP increased cancer risk remain to be fully elucidated, the region harbouring this variant has been shown to contain elements enhancer of the transcription factor 7-like 2 (TCF7L2) and to have physical interaction with the *MYC* proto-oncogene in an allele-specific manner.[46] Therefore, the rs69832687 variant is thought to participate directly in CRC pathogenesis through enhancement of responsiveness of an important component of Wnt signalling.[45,46,48] The recent discovery that CCAT2, a long non-coding RNA transcript encompassing the rs6983267 SNP, up-regulates *MYC* through TCF7L2-mediated transcription, thus activating Wnt signalling, supports the involvement of the rs6983267 G risk allele in CRC pathogenesis.[49]

An intensively investigated SNP variant, c.677C>T p.Ala222Val (rs1801133) in the gene encoding the MTHFR (5,10-methylenetetrahydrofolate reductase) enzyme known to play a key role in one carbon metabolism,[50,51] showed 'less credible' evidence for association with CRA risk. By comparing 11 362 case patients and 23 006 control individuals from 24 studies including the unpublished data from the APC and CORGI trials, we

Table 4. Comparison of summary crude odd ratios (ORs) and 95% confidence intervals (95% CIs) for CRA and CRC associated with variants identified for CRA meta-analysis; genotypic models

Gene/Variant	Colorectal adenomas					Colorectal cancer				
	Cases vs controls (number of samples)	var/wt vs wt/wt		var/var vs wt/wt		Cases vs controls (number of samples)	var/wt vs wt/wt		var/var vs wt/wt	
		OR (95% CI)	*P*-value	OR (95% CI)	*P*-value		OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
APC/c.5465T>A p.Val1822Asp	2805 vs 4369 (4)	0.99 (0.89 −1.09)	0.79	0.86 (0.68 −1.07)	0.18	6282 vs 7038 (6)	0.99 (0.92 −1.07)	0.83	**0.84 (0.71 −0.98)**	**0.03**
BHMT/c.716G>A p.Arg239Gln	3691 vs 12 174 (4)	0.94 (0.87 −1.02)	0.16	0.94 (0.81 −1.07)	0.34	–	–	–	–	–
PTGS1/c.22T>C p.Trp8Arg	2551 vs 4342 (4)	1.06 (0.91 −1.23)	0.48	0.64 (0.28 −1.48)	0.30	–	–	–	–	–
PTGS1/c.382C>A p.Leu237Met	1337 vs 1928 (3)	–	–	1.03 (0.74 −1.42)	0.87	–	–	–	–	–
PTGS2/c.-765G>C	1507 vs 2042 (4)	0.88 (0.73 −1.05)	0.16	0.74 (0.44 −1.23)	0.24	5459 vs 7272 (11)	1.03 (0.95 −1.13)	0.45	1.21 (0.93 −1.57)	0.15
PTGS2/c.*427T>C	2643 vs 4616 (6)	1.02 (0.88 −1.19)	0.79	1.14 (0.97 −1.33)	0.11	4745 vs 5756 (7)	1.01 (0.93 −1.09)	0.87	1.03 (0.91 −1.17)	0.65
PTGS2/c.640-275T>G	2492 vs 4423 (5)	0.95 (0.85 −1.07)	0.41	0.99 (0.74 −1.34)	0.97	–	–	–	–	–
CRP/c.*1082G>A	2016 vs 3465 (3)	0.96 (0.85 −1.08)	0.53	1.15 (0.77 −1.71)	0.49	–	–	–	–	–
***CYP1A1*/c.1384A>G p.Ile462Val**	1846 vs 1853 (3)	1.05 (0.85 −1.30)	0.66	–	–	10274 vs 11 978 (13)	**1.28 (1.01 −1.63)**	**0.05**	**1.47 (1.17 −1.85)**	**0.001**
CYP1A1/c.*1189T>C	3126 vs 5079 (5)	0.90 (0.79 −1.02)	0.11	0.78 (0.52 −1.16)	0.21	4897 vs 6559 (7)	0.94 (0.86 −1.04)	0.23	0.84 (0.56 −1.27)	0.42
CYP1A2/c.-9-154C>A	2689 vs 4373 (4)	1.00 (0.90 −1.11)	0.96	1.08 (0.89 −1.31)	0.42	3051 vs 5326 (9)	1.13 (0.95 −1.34)	0.18	1.07 (0.92 −1.26)	0.40
GSTM1	3952 vs 3514 (7)	–	–	1.03 (0.94 −1.13)	0.50	18845 vs 26 662 (43)	–	–	**1.07 (1.01 −1.13)**	**0.02**
GSTP1/c.313A>G p.Ile105Val	3945 vs 5320 (6)	0.99 (0.91 −1.09)	0.89	1.00 (0.87 −1.15)	0.99	9267 vs 12 902 (22)	1.05 (0.99 −1.12)	0.11	0.95 (0.86 −1.05)	0.32
GSTP1/c.341C>T p.Ala114Val	1316 vs 1973 (3)	0.86 (0.70 −1.05)	0.14	0.77 (0.28 −2.07)	0.60	5183 vs 5457 (6)	1.02 (0.91 −1.13)	0.77	0.87 (0.55 −1.37)	0.55
GSTT1	4345 vs 3964 (8)	–	–	1.06 (0.95 −1.19)	0.31	13410 vs 20 455 (35)	–	–	**1.18 (1.07 −1.31)**	**0.002**
MEH/c.337T>C p.Tyr113His	7387 vs 8774 (12)	0.96 (0.89 −1.03)	0.24	0.92 (0.82 −1.03)	0.15	–	–	–	–	–
MEH/c.416A>G p.His139Arg	7424 vs 8845 (12)	1.00 (0.93 −1.07)	0.97	1.10 (0.93 −1.29)	0.26	–	–	–	–	–
***MTHFR*/c.677C>T p.Ala222Val**	11362 vs 23 006 (24)	**0.95 (0.90 −1.00)**	**0.04**	0.95 (0.88 −1.03)	0.23	27372 vs 39 867 (52)	1.00 (0.94 −1.06)	0.92	**0.87 (0.81 −0.94)**	**<0.0005**
MTHFR/c1286A>C p.Glu429Ala	6760 vs 15 498 (10)	1.02 (0.95 −1.09)	0.65	1.08 (0.97 −1.20)	0.16	1 7178 vs 24 792 (34)	1.01 (0.97 −1.06)	0.51	0.94 (0.87 −1.01)	0.09
MTRR/c.66A>G p.Ile22Met	2911 vs 9342 (3)	0.94 (0.84 −1.06)	0.31	0.90 (0.79 −1.03)	0.12	6170 vs 8732 (9)	0.98 (0.90 −1.07)	0.66	1.04 (0.94 −1.14)	0.47

(Continued)

Table 4. Continued

Gene/Variant	Colorectal adenomas					Colorectal cancer				
	Cases vs controls (number of samples)	var/wt vs wt/wt		var/var vs wt/wt		Cases vs controls (number of samples)	var/wt vs wt/wt		var/var vs wt/wt	
		OR (95% CI)	*P*-value	OR (95% CI)	*P*-value		OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
MTR/c.2756A>G p.Asp919Gly	4730 vs 13 710 (7)	1.02 (0.95 −1.10)	0.53	1.00 (0.84 −1.20)	0.97	1 1829 vs 15 975 (14)	0.97 (0.92 −1.02)	0.27	1.00 (0.83 −1.20)	0.98
NAT1	2347 vs 3143 (5)	1.14 (0.98 −1.34)	0.10	***1.35 (1.02 −1.78)***	**0.04**	4791 vs 6628 (15)	***0.80 (0.68 −0.93)***	**0.003**	*0.98 (0.79 −1.22)*	*0.97*
NAT2	4092 vs 4731 (9)	0.99 (0.92 −1.08)	0.84	0.86 (0.74 −1.00)	0.06	1 2908 vs 16 483 (26)	1.01 (0.83 −1.22)	0.94	0.95 (0.76 −1.20)	0.68
***NQO1*/c.559C>T p.Pro187Ser**	4097 vs 5967 (6)	1.08 (0.95 −1.23)	0.25	***1.26 (1.00 −1.59)***	**0.05**	5084 vs 5932 (8)	1.14 (0.96 −1.35)	0.12	1.10 (0.76 −1.59)	0.63
***TP53*/c.215C>G p.Arg72Pro**	2135 vs 3738 (3)	***1.16 (1.04 −1.30)***	**0.008**	*1.17 (0.94 −1.47)*	*0.17*	7414 vs 9872 (27)	1.01 (0.89 −1.14)	0.90	1.04 (0.82 −1.31)	0.77
PPARG/c.34C>G p.Pro12Ala	1730 vs 2953 (3)	1.08 (0.94 −1.25)	0.29	1.14 (0.73 −1.80)	0.56	1 5091 vs 18 690 (17)	0.98 (0.86 −1.11)	0.72	0.91 (0.73 −1.12)	0.37
SLC19A1/c.80G>A p.His27Arg	3614 vs 11 467 (4)	0.99 (0.90 −1.08)	0.79	1.08 (0.85 −1.35)	0.53	–	–	–	–	–
TGFB1/c.29C>T p.Pro10Leu	2840 vs 4218 (4)	1.06 (0.95 −1.18)	0.29	1.04 (0.89 −1.20)	0.63	–	–	–	–	–
TGFB1/c.-1347T>C	2395 vs 3902 (4)	1.05 (0.94 −1.17)	0.38	1.07 (0.90 −1.28)	0.45	*994 vs 2335 (5)*	1.12 (0.91 −1.37)	>0.05	**1.62 (1.30 −2.02)**	**<0.05**
TGFB1/c.74G>C p.Arg25Pro	1011 vs 1329 (3)	1.04 (0.78 −1.38)	0.80	1.97 (0.72 −5.37)	0.18	–	–	–	–	–
***TYMS*/TS tandem repeat**	1633 vs 2034 (3)	0.90 (0.77 −1.06)	0.20	0.94 (0.77 −1.13)	0.49	*3519 vs 5289 (5)*	***0.86 (0.78 −0.95)***	**0**	**0.83 (0.73 −0.94)**	**0.004**
***VDR*/c.1024+283G>A**	2403 vs 4356 (5)	1.01 (0.91 −1.14)	0.80	1.00 (0.86 −1.15)	0.96	*5607 vs 6202 (7)*	*0.77 (0.58 −1.02)*	*0.07*	**0.51 (0.28 −0.90)**	**0.02**
XPD/c.2251A>C p.Lys751Gln	3363 vs 3523 (5)	1.06 (0.96 −1.18)	0.26	1.07 (0.82 −1.39)	0.63	–	–	–	–	–
XRCC1/c.580C>T p.Arg194Trp	4821 vs 5972 (5)	1.09 (0.98 −1.23)	0.12	1.29 (0.74 −2.23)	0.37	6635 vs 8488 (11)	0.96 (0.87 −1.07)	0.5	1.10 (0.82 −1.48)	0.52
***XRCC1*/c.1196A>G p.Arg399Gln**	5124 vs 6927 (6)	0.98 (0.90 −1.07)	0.61	0.86 (0.72 −1.03)	0.11	*7247 vs 8786 (12)*	*0.99 (0.92 −1.06)*	*0.72*	**0.88 (0.79 −0.97)**	**0.02**
XRCC3/c.722C>T p.Thr241M	3183 vs 4514 (4)	0.92 (0.83 −1.02)	0.10	0.92 (0.79 −1.07)	0.29	4484 vs 5235 (10)	0.92 (0.84 −1.01)	0.09	0.95 (0.72 −1.24)	0.68
8q24.21/rs6983267	3559 vs 9586 (8)	**1.20 (1.08 −1.33)**	**0.0006**	**1.51 (1.27 −1.79)**	**$2.1x10^{-6}$**	4 0604 vs 42 672 (19)	**1.22 (1.16 −1.28)**	**<0.0005**	**1.45 (1.39 −1.51)**	**<0.0005**

Note: The associations which were considered to be "positives" or "less credible positives" are indicated in bold type.

observed an inverse association between the *MTHFR* c.677C>T p.Ala22Val (rs1801133) variant and CRA risk in both the dominant and the genotypic models (var/wt vs wt/wt and var/var vs wt/wt). The observed association between the *MTHFR* c.677C>T p.Ala22Val (rs1801133) variant and CRA risk is in line with our previous findings in CRC[15] (Table 4). However, the evidence for association with CRC was 'highly credible', probably due in large part to the much larger sample size (52 studies providing data on more than 27 000 cases and 40 000 controls).

Both homozygosity and heterozygosity for the c.677C>T p.Ala22Val (rs1801133) variant lead to the synthesis of a thermolabile MTHFR enzyme with depressed activity, hence affecting the biological level and distribution of folate.[50,51] Accumulating evidence suggests interactions between the *MTHFR* c.677C>T p.Ala22Val (rs1801133) variant and dietary factors in modulating the risk of colorectal neoplasia. In particular, the *MTHFR* 677TT genotype appears to be protective for individuals with adequate folate status but, under conditions of impaired folate status, the homozygous *TT* genotype is reported to result in increased CRC risk.[52,53] However, whereas higher folate intake and blood folate levels have been repeatedly inversely correlated with CRC risk, the role of the *MTHFR* c.677C>T p.Ala22Val (rs1801133) variant and folate on CRA risk still controversial.[54–56] Experimental studies in conjunction with epidemiological investigations indicate that folate may have a dose- and time-dependent effect on development of colorectal neoplasia.[57–59] Thus, low levels of folate may have an inhibitory effect, whereas folic acid fortification could promote the progression of established colorectal neoplastic lesions.[57–59] Currently there is no conclusive evidence supporting the use of folate supplementation as a chemopreventive measure for colorectal neoplasia. As our findings extend and confirm earlier studies on the impact of the *MTHFR* c.677C>T p.Ala22Val (rs1801133) variant on colorectal neoplasia risk,[60] we suggest that a well-designed large epidemiological study to investigate gene-gene and gene-environment interaction would help to clarify the role of the *MTHFR* c.677C>T variant and folate in the pathogenesis of colorectal neoplasia.

Our meta-analyses revealed a 'less credible' association between CRA risk and the c.215C>G p.Arg72Pro (rs1042522) variant of *TP53* gene. Based on three studies aggregating data on 2135 cases and 3738 controls, including unpublished data from the APC trial, this SNP showed a positive association in both genotypic (var/wt vs wt/wt) and dominant models with little heterogeneity. The tumour suppressor gene *TP53* encodes for a transcription factor identified as a master regulator of various signalling pathways controlling critical cellular processes, and the gene is generally referred to as the guardian of the genome.[61,62] *TP53* gene alteration is a hallmark of various human diseases, and its role in human neoplasia is unequivocal since somatic mutations in *TP53* occur in approximately 50% of human cancers.[63] Although both structural forms of the p53 p.Arg72Pro protein show no abnormalities in their DNA binding activities, there is convincing evidence for biochemical and functional differences between them, possibly underlying differential susceptibility to various cancers.[64,65] The protein associated with the Arg variant has been reported to induce apoptosis more efficiently than the Pro variant, and both variant forms have also been shown to differ in their vulnerability to degradation by the human papilloma virus E6 protein.[64,65] Consideration of the influence of the c.215C>G p.Arg72Pro (rs1042522) variant on colorectal neoplasia risk has been primarily focused on CRC rather than CRA. The results relating to CRC have been inconsistent, with a positive association with the Pro72 allele variant having been reported in both population-based and hospital-based studies from different ethnic groups,[66–72] and a positive association with the Arg72 variant in two studies,[73,74] although others did not detect any association.[75–77] Indeed, due to a small sample size and limited relevant studies with adequate design, our previous systematic meta-analysis did not suggest any credible association between the c.215C>G p.Arg72Pro variant and CRC risk[15] (Table 4). This finding was in line with a previous meta-analysis conducted by Dahabreh *et al.* including 23 studies published before 2009.[78] Since the *TP53* Pro72 variant is reported to induce the cell cycle more efficiently than the *TP53* Arg72 variant, but with lower potential to trigger apoptosis,[64] the observed association between the Pro72 allele variant and CRA risk could suggest that *TP53*-induced apoptosis could be critical during early CRC development.

Our meta-analyses also found 'less credible' evidence of association between CRA risk and the SNP rs1800566 known as c.559C>T p.Pro187Ser variant of the *NQO1* (NADP(H): quinine oxidoreductase 1) gene. With accumulated data from 4097 cases and 5967 controls from six studies, we identified a positive association for the *NQO1* c.559C>T p.Pro187Ser (rs1800566) variant in both the dominant and genotypic models (var/wt vs wt/wt and var/var vs wt/wt) without heterogeneity. The enzyme encoded by the *NQO1* gene plays a pivotal role in detoxification of various mutagenic and carcinogenic compounds such as quinones derived from diet or tobacco smoke.[79] In addition, NQO1 protein prevents generation of free radicals and reactive oxygen species, thereby protecting cells from oxidative damage.[79] The c.559C>T p.Pro187Ser (rs1800566) variant is reported to decrease NQO1 protein

enzymatic activity, and so has been suggested to increase metabolic activation of pro-carcinogenic compounds.[80] The impact of the *NQO1* c.559C>T p.Pro187Ser (rs1800566) variant on colorectal neoplasia risk has been previously investigated in a several case-control studies, but the results have so far remained inconclusive.[81–84] In a previous meta-analysis, both the *NQO1* c.559C>T and TT genotype showed modest increased risks for CRC in populations of European origin, whereas an inverse association was reported in Asians.[85] However, in our systematic meta-analyses of CRC no credible association with the *NQO1* c.559C>T p.Pro187Ser (rs1800566) variant was identified[15] (Table 4), based on a similar volume of evidence. Interestingly, interactions between *NQO1* c.559C>T p.Pro187Ser (rs1800566) variant and tobacco and alcohol have been recently reported both in CRA and in CRC.[86,84]

Polycyclic aromatic hydrocarbons (PAHs), heterocyclic amines (HAs) and arylamines (AA) are known potential carcinogenic compounds found in tobacco smoke, cooked and processed meat, petroleum products, coal and vehicle emissions.[87,90] The metabolism of these compounds is complex and involves activation and detoxiffcation steps catalyzed by several polymorphic enzymes, including the arylamine N-acetyltransferase enzymes (NAT1 and NAT2).[88] Polymorphic variants of these genes have been reported to decrease the stability and activity of the encoded enzymes and thus modify the association between environmental exposure such as cigarette smoking, cooked and processed meat consumption and colorectal neoplasia.[89] In this study we found a less credible association with CRA risk for the *NAT1* genotypes representing the fast acetylator phenotype (genotypes containing the *NAT1**10 allele associated with high NAT1 enzymatic activity).[41,42] We observed a positive association for the genotypic model (var/var vs wt/wt) in accumulated data on 2347 cases and 3143 controls from five studies, without heterogeneity. Although there have been some controversies among previous studies assessing the role of *NAT1* on colorectal neoplasia risk, the NAT1 fast acetylator phenotype has been associated with adenoma and CRC in various studies, where some of these also reported interaction with meat and smoking.[89–92] The association between *NAT1* genotypes associated with a rapid acetylator phenotype and CRA risk observed in the present study contrasts with the lack of association for CRC in our previous meta-analysis[15] (Table 4), for which there was at least double the volume of evidence.

Our meta-analyses did not find any credible evidence for association for the other 33 variants previously reported to be associated with CRA risk. With regard to these variants, meta-analyses included a minimum of 1011 cases and 1329 controls but there was low statistical power to detect any significant association for any of the three models of inheritance considered in this study [genotypic (var/wt vs wt/wt and var/var vs wt/wt), recessive and dominant models]. Further epidemiological investigations with bigger sample size and adequate design could be helpful in addressing the impacts of these variants on CRA risk.

The potential limitations of this study include mainly the relatively small sample size that could have contributed to the lack of sufficient statistical power to detect any association that may genuinely exist for some variants. In addition, for six variants for which strong heterogeneity between studies was observed, one could hypothesize different risk estimates arising due to ethnic variations. However, because of lack of sufficient information on ethnicity, inadequate sample size and methodological differences between studies, we were unable to conduct population stratification analyses to address this issue. This study was also limited to the main effect of SNP variations on the overall risk of CRA, and we were unable to undertake subpopulation analysis taking into account different types of polyps (conventional adenomatous polyps vs hyperplastic polyps including serrated polyps) or different colon localization. Furthermore, some studies used hospital-based rather than population-based controls, resulting in potentially different vulnerabilities to selection bias. We conducted population-based and hospital-based stratification analyses for the *MTHFR* c.677C>T p.Ala222Val (rs1801133) variant. Results were similar for population- and hospital-based studies and also similar to the whole sample analysis [var/wt vs var/var OR (95% CI): whole sample: 0.95 (0.90 –1.00); hospital-based: 0.95 (0.87 –1.03); population-based: 0.95 (0.89 –1.01)]. We could not do this stratification analysis for any other variant due to the lack of information or inadequate sample size.

Different studies included in our analyses also investigated gene-gene and gene-environment interaction, and significant interactions were reported for several variants. The overall risk estimates for some variants, including those for which our meta-analyses found no credible evidence for association with CRA, could be affected by potential gene-gene or gene-environment interactions. Nevertheless, compared with the main effects of low penetrance genetic variations on the CRA risk, very much larger sample sizes are still required to investigate gene-gene and gene-environment interactions adequately.

In conclusion, the number of common genetic variants likely to be associated with CRA is much less than that observed for CRC. Among the 74 candidate susceptibility genes for CRA investigated so far, our findings suggest 'high credibility' of association with the rs6983267 variant

at 8q24.21, and 'less credible' evidence of association with a further four variants, namely, *MTHFR* c.677C>T p.Ala222Val (rs1801133), *TP53* c.215C>G p.Arg72Pro (rs1042522), *NQO1* c.559C>T p.Pro187Ser (rs1800566) and *NAT1* genotypes associated with the fast acetylator phenotype. For some of these SNPs, interactions with environmental factors have been suggested, with the evidence for the *MTHFR* c.677C>T p.Ala222Val (rs1801133) being the most documented. Large-scale molecular epidemiological studies, designed to investigate the role of these variants, in combination with established colorectal neoplasia risk factors, will characterize the exact relationships between these SNPs and CRA susceptibility. Thus, the present findings should help focus further research on understanding the role of gene-gene and gene-environment interactions for identified genetic variants and CRA risk. This will allow identification of true causative factors associated with CRA occurrence and provide opportunities for improved surveillance and prevention strategies for CRC.

Supplementary Data

Supplementary data are available at *IJE* online.

Conflict of interest: The author declares no conflict of interest.

References

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J *et al*. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 2013;**49:**1374–403.
2. Eide TJ. Risk of colorectal cancer in adenoma-bearing individuals within a defined population. *Int J Cancer* 1986;**38:**173–76.
3. O'Brien MJ. Hyperplastic and serrated polyps of the colorectum. *Gastroenterol Clin North Am* 2007;**36:**947,68, viii.
4. Botteri E, Iodice S, Raimondi S, Maisonneuve P, Lowenfels AB. Cigarette smoking and adenomatous polyps: a meta-analysis. *Gastroenterology* 2008;**134:**388–95.
5. Zhu JZ, Wang YM, Zhou QY, Zhu KF, Yu CH, Li YM. Systematic review with meta-analysis: alcohol consumption and the risk of colorectal adenoma. *Aliment Pharmacol Ther* 2014;**40:**325–37.
6. Ben Q, An W, Jiang Y *et al*. Body mass index increases risk for colorectal adenomas based on meta-analysis. *Gastroenterology* 2012;**142:**762–72.
7. Hong S, Cai Q, Chen D, Zhu W, Huang W, Li Z. Abdominal obesity and the risk of colorectal adenoma: a meta-analysis of observational studies. *Eur J Cancer Prev* 2012;**21:**523–31.
8. Okabayashi K, Ashrafian H, Hasegawa H *et al*. Body mass index category as a risk factor for colorectal adenomas: a systematic review and meta-analysis. *Am J Gastroenterol* 2012;**107:**1175,85; quiz 1186.
9. Omata F, Deshpande GA, Ohde S, Mine T, Fukui T. The association between obesity and colorectal adenoma: systematic review and meta-analysis. *Scand J Gastroenterol* 2013;**48:** 136–46.
10. Cole BF, Logan RF, Halabi S *et al*. Aspirin for the chemoprevention of colorectal adenomas: meta-analysis of the randomized trials. *J Natl Cancer Inst* 2009;**101:**256–66.
11. Gao F, Liao C, Liu L, Tan A, Cao Y, Mo Z. The effect of aspirin in the recurrence of colorectal adenomas: a meta-analysis of randomized controlled trials. *Colorectal Dis* 2009;**11:**893–901.
12. Psaty BM, Potter JD. Risks and benefits of celecoxib to prevent recurrent adenomas. *N Engl J Med* 2006;**355:**950–52.
13. Chan AT, Giovannucci EL. Primary prevention of colorectal cancer. *Gastroenterology* 2010;**138:**2029,2043.e10.
14. de la Chapelle A. Genetic predisposition to colorectal cancer. *Nat Rev.Cancer* 2004;**4:**769–80.
15. Theodoratou E, Montazeri Z, Hawken S *et al*. Systematic meta-analyses and field synopsis of genetic association studies in colorectal cancer. *J Natl Cancer Inst* 2012;**104:**1433–57.
16. Tuohy TM, Rowe KG, Mineau GP, Pimentel R, Burt RW, Samadder NJ. Risk of colorectal cancer and adenomas in the families of patients with adenomas: a population-based study in Utah. *Cancer* 2014;**120:**35–42.
17. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol* 2001;**96:**2992–3003.
18. Samadder NJ, Curtin K, Tuohy TM *et al*. Increased risk of colorectal neoplasia among family members of patients with colorectal cancer: a population-based study in Utah. *Gastroenterology* 2014;**147:**814,821.e5.
19. Thun MJ, Calle EE, Namboodiri MM *et al*. Risk factors for fatal colon cancer in a large prospective study. *J Natl Cancer Inst* 1992;**84:**1491–500.
20. Carvajal-Carmona LG, Zauber AG, Jones AM *et al*. Much of the genetic risk of colorectal cancer is likely to be mediated through susceptibility to adenomas. *Gastroenterology* 2013;**144:**53–55.
21. Ioannidis JP, Boffetta P, Little J *et al*. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 2008;**37:**120–32.
22. Khoury MJ, Bertram L, Boffetta P *et al*. Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases. *Am J Epidemiol* 2009;**170:**269–79.
23. Allen NC, Bagade S, McQueen MB *et al*. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 2008;**40:**827–34.
24. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 2007;**39:**17–23.
25. Chatzinasiou F, Lill CM, Kypreou K *et al*. Comprehensive field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma. *J Natl Cancer Inst* 2011;**103:**1227–35.
26. Campbell H, Rudan I. Interpretation of genetic association studies in complex disease. *Pharmacogenomics J* 2002;**2:** 349–60.
27. Wang J, Carvajal-Carmona LG, Chu JH *et al*. Germline variants and advanced colorectal adenomas: adenoma prevention with celecoxib trial genome-wide association study. *Clin Cancer Res* 2013;**19:**6430–37.

28. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2014.
29. Ioannidis JP, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *CMAJ* 2007;**176:**1091–96.
30. Dupont WD, Plummer WD Jr. Power and sample size calculations. A review and computer program. *Control Clin Trials* 1990;**11:**116–28.
31. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008;**24:**2938–39.
32. Houlston RS, Ford D. Genetics of coeliac disease. *QJM* 1996;**89:**737–43.
33. Cox A, Dunning AM, Garcia-Closas M *et al*. A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet* 2007;**39:**352–58.
34. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol* 2001;**96:**2992–3003.
35. Wakefield J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet* 2007;**81:**208–27.
36. Pompanon F, Bonin A, Bellemain E, Taberlet P. Genotyping errors: causes, consequences and solutions. *Nat Rev Genet* 2005;**6:**847–59.
37. Marchini J, Cardon LR, Phillips MS, Donnelly P. The effects of human population structure on large genetic association studies. *Nat Genet* 2004;**36:**512–17.
38. Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet* 2006;**7:**781–91.
39. Wacholder S, Rothman N, Caporaso N. Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. *J Natl Cancer Inst* 2000;**92:**1151–18.
40. Ioannidis JP, Ntzani EE, Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. *Nat Genet* 2004;**36:**1312–18.
41. Bell DA, Stephens EA, Castranio T *et al*. Polyadenylation polymorphism in the acetyltransferase 1 gene (NAT1) increases risk of colorectal cancer. *Cancer Res* 1995;**55:**3537–42.
42. Hein DW, Doll MA, Fretland AJ *et al*. Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. *Cancer Epidemiol, Biomarkers Prev* 2000;**9:**29–42.
43. Bertucci F, Lagarde A, Ferrari A *et al*. 8q24 Cancer risk allele associated with major metastatic risk in inflammatory breast cancer. *PLoS One* 2012;**7:**e37943.
44. Ghoussaini M, Song H, Koessler T *et al*. Multiple loci with different cancer specificities within the 8q24 gene desert. *J Natl Cancer Inst* 2008;**100:**962–66.
45. Tuupanen S, Turunen M, Lehtonen R *et al*. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. *Nat Genet* 2009;**41:**885–90.
46. Pomerantz MM, Ahmadiyeh N, Jia L *et al*. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat Genet* 2009;**41:**882–84.
47. Bienz M, Clevers H. Linking colorectal cancer to Wnt signaling. *Cell* 2000;**103:**311–20.
48. Tuupanen S, Niittymaki I, Nousiainen K *et al*. Allelic imbalance at rs6983267 suggests selection of the risk allele in somatic colorectal tumor evolution. *Cancer Res* 2008;**68:**14–17.
49. Ling H, Spizzo R, Atlasi Y *et al*. CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. *Genome Res* 2013;**23:**1446–61.
50. Frosst P, Blom HJ, Milos R *et al*. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet* 1995;**10:**111–13.
51. Goyette P, Rozen R. The thermolabile variant 677C–>T can further reduce activity when expressed in cis with severe mutations for human methylenetetrahydrofolate reductase. *Hum Mutat* 2000;**16:**132–38.
52. Chen J, Giovannucci EL, Hunter DJ. MTHFR polymorphism, methyl-replete diets and the risk of colorectal carcinoma and adenoma among U.S. men and women: an example of gene-environment interactions in colorectal tumorigenesis. *J Nutr* 1999;**129:**560–64S.
53. Botto LD, Yang Q. 5,10-Methylenetetrahydrofolate reductase gene variants and congenital anomalies: a HuGE review. *Am J Epidemiol* 2000;**151:**862–77.
54. Murphy G, Sansbury LB, Cross AJ *et al*. Folate and MTHFR: risk of adenoma recurrence in the Polyp Prevention Trial. *Cancer Causes Control* 2008;**19:**751–58.
55. Kim YI. Folate and colorectal cancer: an evidence-based critical review. *Mol Nutr Food Res* 2007;**51:**267–92.
56. Kim YI. Folic acid fortification and supplementation – good for some but not so good for others. *Nutr Rev* 2007;**65:**504–11.
57. Kim YI. Folate, colorectal carcinogenesis, and DNA methylation: lessons from animal studies. *Environ Mol Mutagen* 2004;**44:**10–25.
58. Lin YW, Wang JL, Chen HM *et al*. Folic acid supplementary reduce the incidence of adenocarcinoma in a mouse model of colorectal cancer: microarray gene expression profile. *J Exp Clin Cancer Res* 2011;**30:**116,9966-30-116.
59. Lawrance AK, Deng L, Rozen R. Methylenetetrahydrofolate reductase deficiency and low dietary folate reduce tumorigenesis in Apc min/+ mice. *Gut* 2009;**58:**805–11.
60. Hubner RA, Houlston RS. MTHFR C677T and colorectal cancer risk: A meta-analysis of 25 populations. *Int J Cancer* 2007;**120:**1027–35.
61. Aloni-Grinstein R, Shetzer Y, Kaufman T, Rotter V. p53: The barrier to cancer stem cell formation. *FEBS Lett* 2014;**588:**2580–89.
62. Lin CP, Choi YJ, Hicks GG, He L. The emerging functions of the p53-miRNA network in stem cell biology. *Cell Cycle* 2012;**11:**2063–72.
63. Vousden KH, Prives C. Blinded by the Light: The Growing Complexity of p53. *Cell* 2009;**137:**413–31.
64. Dumont P, Leu JI, Della Pietra AC 3rd, George DL, Murphy M. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat Genet* 2003;**33:**357–65.
65. Storey A, Thomas M, Kalita A *et al*. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. *Nature* 1998;**393:**229–34.

66. Song HR, Kweon SS, Kim HN *et al*. p53 codon 72 polymorphism in patients with gastric and colorectal cancer in a Korean population. *Gastric Cancer* 2011;**14:**242–48.
67. Singamsetty GK, Malempati S, Bhogadhi S *et al*. TP53 alterations and colorectal cancer predisposition in south Indian population: A case-control study. *Tumour Biol* 2014;**35:** 2303–11.
68. Gemignani F, Moreno V, Landi S *et al*. A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene* 2004;**23:**1954–56.
69. Perfumo C, Bonelli L, Menichini P *et al*. Increased risk of colorectal adenomas in Italian subjects carrying the p53 PIN3 A2-Pro72 haplotype. *Digestion* 2006;**74:**228–35.
70. Goodman JE, Mechanic LE, Luke BT, Ambs S, Chanock S, Harris CC. Exploring SNP-SNP interactions and colon cancer risk using polymorphism interaction analysis. *Int J Cancer* 2006;**118:**1790–97.
71. Zhu ZZ, Wang AZ, Jia HR *et al*. Association of the TP53 codon 72 polymorphism with colorectal cancer in a Chinese population. *Jpn J Clin Oncol* 2007;**37:**385–90.
72. Cao Z, Song JH, Park YK *et al*. The p53 codon 72 polymorphism and susceptibility to colorectal cancer in Korean patients. *Neoplasma* 2009;**56:**114–48.
73. Dakouras A, Nikiteas N, Papadakis E *et al*. P53Arg72 homozygosity and its increased incidence in left-sided sporadic colorectal adenocarcinomas, in a Greek-Caucasian population. *Anticancer Res* 2008;**28:**1039–43.
74. Perez LO, Abba MC, Dulout FN, Golijow CD. Evaluation of p53 codon 72 polymorphism in adenocarcinomas of the colon and rectum in La Plata, Argentina. *World J Gastroenterol* 2006;**12:**1426–29.
75. Koushik A, Tranah GJ, Ma J *et al*. p53 Arg72Pro polymorphism and risk of colorectal adenoma and cancer. *Int J Cancer* 2006;**119:**1863–68.
76. Polakova V, Pardini B, Naccarati A *et al*. Genotype and haplotype analysis of cell cycle genes in sporadic colorectal cancer in the Czech Republic. *Hum Mutat* 2009;661–68.
77. Joshi AM, Budhathoki S, Ohnaka K *et al*. TP53 R72P and MDM2 SNP309 polymorphisms and colorectal cancer risk: the Fukuoka Colorectal Cancer Study. *Jpn J Clin Oncol* 2011;232–38.
78. Dahabreh IJ, Linardou H, Bouzika P, Varvarigou V, Murray S. TP53 Arg72Pro polymorphism and colorectal cancer risk: a systematic review and meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2010;**19:**1840–47.
79. Ross D, Kepa JK, Winski SL, Beall HD, Anwar A, Siegel D. NAD(P)H:quinone oxidoreductase 1 (NQO1): chemoprotection, bioactivation, gene regulation and genetic polymorphisms. *Chem Biol Interact* 2000;**129:**77–97.
80. Traver RD, Siegel D, Beall HD *et al*. Characterization of a polymorphism in NAD(P)H: quinone oxidoreductase (DT-diaphorase). *Br J Cancer* 1997;**75:**69–75.
81. Hamajima N, Matsuo K, Iwata H *et al*. NAD(P)H: quinone oxidoreductase 1 (NQO1) C609T polymorphism and the risk of eight cancers for Japanese. *Int J Clin Oncol* 2002;**7:**10–0-8.
82. Tijhuis MJ, Visker MH, Aarts JM *et al*. NQO1 and NFE2L2 polymorphisms, fruit and vegetable intake and smoking and the risk of colorectal adenomas in an endoscopy-based population. *Int J Cancer* 2008;**122:**1842–48.
83. Su XL, Yan MR, Yang L, Qimuge-Suyila. NQO1 C609T polymorphism correlated to colon cancer risk in farmers from western region of Inner Mongolia. *Chin J Cancer Res* 2012;**24:** 317–22.
84. Peng XE, Jiang YY, Shi XS, Hu ZJ. NQO1 609C>T polymorphism interaction with tobacco smoking and alcohol drinking increases colorectal cancer risk in a Chinese population. *Gene* 2013;**521:**105–10.
85. Chao C, Zhang ZF, Berthiller J, Boffetta P, Hashibe M. NAD(P)H:quinone oxidoreductase 1 (NQO1) Pro187Ser polymorphism and the risk of lung, bladder, and colorectal cancers: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;**15:**979–87.
86. Tijhuis MJ, Visker MH, Aarts JM *et al*. NQO1 and NFE2L2 polymorphisms, fruit and vegetable intake and smoking and the risk of colorectal adenomas in an endoscopy-based population. *Int J Cancer* 2008;**122:**1842–48.
87. Lewtas J. Air pollution combustion emissions: characterization of causative agents and mechanisms associated with cancer, reproductive, and cardiovascular effects. *Mutat Res* 2007;**636:**95–133.
88. Xue W, Warshawsky D. Metabolic activation of polycyclic and heterocyclic aromatic hydrocarbons and DNA damage: a review. *Toxicol Appl Pharmacol* 2005;**206:**73–93.
89. Lilla C, Verla-Tebit E, Risch A *et al*. Effect of NAT1 and NAT2 genetic polymorphisms on colorectal cancer risk associated with exposure to tobacco smoke and meat consumption. *Cancer Epidemiol Biomarkers Prev* 2006;**15:**99–107.
90. Goode EL, Potter JD, Bamlet WR, Rider DN, Bigler J. Inherited variation in carcinogen-metabolizing enzymes and risk of colorectal polyps. *Carcinogenesis* 2007;**28:**328–41.
91. Lin HJ, Probst-Hensch NM, Hughes NC *et al*. Variants of N-acetyltransferase NAT1 and a case-control study of colorectal adenomas. *Pharmacogenetics* 1998;**8:**269–81.
92. Tiemersma EW, Bunschoten A, Kok FJ, Glatt H, de Boer SY, Kampman E. Effect of SULT1A1 and NAT2 genetic polymorphism on the association between cigarette smoking and colorectal adenomas. *Int J Cancer* 2004;**108:**97–103.
93. Liang J, Lin C, Hu F *et al*. APC polymorphisms and the risk of colorectal neoplasia: a HuGE review and meta-analysis. *Am J Epidemiol* 2013;**177:**1169–79.
94. Zhao ZQ, Guan QK, Yang FY, Zhao P, Zhou B, Chen ZJ. System review and metaanalysis of the relationships between five metabolic gene polymorphisms and colorectal adenoma risk. *Tumour Biol* 2012;**33:**523–35.
95. Huang Y, Han S, Li Y, Mao Y, Xie Y. Different roles of MTHFR C677T and A1298C polymorphisms in colorectal adenoma and colorectal cancer: a meta-analysis. *J Hum Genet* 2007;**52:**73–85.
96. Ding W, Zhou DL, Jiang X, Lu LS. Methionine synthase A2756G polymorphism and risk of colorectal adenoma and cancer: evidence based on 27 studies. *PLoS One* 2013;**8:**e60508.
97. Liu J, Ding D, Wang X *et al*. N-acetyltransferase polymorphism and risk of colorectal adenoma and cancer: a pooled

analysis of variations from 59 studies. *PLoS One* 2012;**7:**e42797.

98. Chen J, Lin Y, Zhang R, Huang ZJ, Pan XG. Contribution of NAD(P)H quinone oxidoreductase 1 (NQO1) Pro187Ser polymorphism and risk of colorectal adenoma and colorectal cancer in Caucasians: a meta-analysis. *Arch Med Res* 2012; **43:**58–66.
99. Liu Y, Zhou W, Zhong DW. Meta-analyses of the associations between four common TGF-1 genetic polymorphisms and risk of colorectal tumor. *Tumour Biol* 2012;**33:**1191–99.
100. Lee JE. Circulating levels of vitamin D, vitamin D receptor polymorphisms, and colorectal adenoma: a meta-analysis. *Nutr Res Pract* 2011;**5:**464–70.
101. Wang YP, Zhang J, Zhu HY *et al*. Common variation rs6983267 at 8q24.1 and risk of colorectal adenoma and cancer: evidence based on 31 studies. *Tumour Biol* 2014;**35:**4067–75.
102. Li M, Zhou Y, Chen P *et al*. Genetic variants on chromosome 8q24 and colorectal neoplasia risk: a case-control study in China and a meta-analysis of the published literature. *PLoS One* 2011;**6:**e18251.