Published in final edited form as:
Gastroenterology. 2010 November ; 139(5): 1677–1685.e8. doi:10.1053/j.gastro.2010.07.038.

Genetic Heterogeneity in Colorectal Cancer Associations in Americans of African vs. European Descent

Sonia S. Kupfer[1], **Jeffrey R. Anderson**[1], **Stanley Hooker**[2], **Andrew Skol**[2], **Rick A. Kittles**[2], **Temitope O. Keku**[3], **Robert S. Sandler**[3], and **Nathan A. Ellis**[1]

[1]Department of Medicine, Section of Gastroenterology University of Chicago Medical Center, Chicago, IL

[2]Department of Medicine, Section of Genetic Medicine University of Chicago Medical Center, Chicago, IL

[3]Department of Medicine, Division of Gastroenterology and Hepatology, University of North Carolina, Chapel Hill, NC

Abstract

Background & Aims—Genome-wide association studies of colorectal cancer (CRC) have identified risk variants in 10 genomic regions. None of these studies included African Americans, who have the highest incidence and mortality from CRC in the US. For the 10 genomic regions, we performed an association study of Americans of African and European descent.

Methods—We genotyped 22 single nucleotide polymorphisms (SNPs) in DNA samples from 1194 patients with CRC (795 African Americans and 399 European Americans) and 1352 controls (985 African Americans and 367 European Americans). At chromosome 8q24.21 region 3, we analyzed 6 SNPs from 1000 African American cases and 1393 controls. Association testing was done using multivariate logistic regression controlling for ancestry, age, and sex.

Results—Sizes and directions of association for all SNPs tested in European Americans were consistent with previously published studies, but for 9 of 22 SNPs tested in African Americans, they were of an opposite direction. Among African Americans, the SNP rs6983267 at 8q24.21 was not associated with CRC (odds ratio [OR]=1.18; P=0.12); instead, the 8q24.21 SNP rs7014346 (OR=1.15; p=0.03) was associated with CRC in this population. At 15q13.3, rs10318 was associated with CRC in both populations. At 10p14, the opposite allele of rs10795668 was associated with CRC in African Americans (OR=1.35; P=0.04). At 11q23.1, rs3802842 was significantly associated with rectal cancer risk only among African Americans (OR 1.34; P=0.01); this observation was made in previous studies. Among European Americans, SNPs at 8q24.21, 11q23.1, and 16q22.1 were associated with CRC, in agreement with previous reports.

Correspondence to: Nathan A. Ellis, Department of Medicine, Section of Gastroenterology, University of Chicago Medical Center, 900 East 57th Street, 9th floor, Chicago, IL 60637, (773) 834-2271 office, (773) 702-2281, naellis@uchicago.edu.

Author participation:
Study concept and design (S.S.K. and N.A.E.)
Technical and material support (J.R.A., S.H., T.O.K., R.S.S.)
Statistical analysis (S.S.K. and A.S.)
Analysis and interpretation (S.S.K. and N.A.E.)
Manuscript preparation (S.S.K., N.A.E., R.A.K., T.O.K., R.S.S.)

Conclusion—There is genetic heterogeneity in CRC associations in Americans of African vs. European descent.

Keywords

colon cancer; rectal cancer; ethnicity; genetic polymorphism

Background

Recent genome-wide association studies (GWAS) in colorectal cancer (CRC) have shown strong evidence for common single nucleotide polymorphism (SNP) associations in a number of genes and chromosome regions (Table 1)(1). These studies have identified associations in regions of interesting candidate genes (*SMAD7, GREM1, CDH1, BMP4*, and *RHPN2*) as well as relative gene deserts (on chromosomes 8q23.3, 8q24.21, 10p14, 11q23.1, and 20p12.3). To date, SNPs in region 3 of chromosome 8q24.21 have been the most widely replicated CRC associations. Recent work has suggested that rs6983267 in region 3 of 8q24.21 may be the functional variant, as it displays enhancer function and interacts with *MYC* (2, 3).

Two genetic models can explain CRC associations detected in recent GWAS (4). The common disease-common variant model posits that susceptibility alleles are frequent (frequencies >5%) and exert modest effects, whereas the common disease-rare variant model posits that susceptibility alleles are more rare (frequencies <5%) and exert stronger effects. If the source of an association with a gene is a rare allele, population genetic theory suggests that, in the absence of selection, the allele will have established itself in the population more recently than most common alleles. If the origination occurred after the separation of two ethnic populations, then the two populations may demonstrate associations with different SNPs in the same gene and possibly associations with different genes. Consequently, comparisons of association results from different ethnic populations can help distinguish these two genetic models.

While the published GWAS reports have included thousands of subjects, they have been performed almost exclusively in individuals of European descent. Replication of genome-wide SNPs in African Americans is an important step in elucidating the genetic mechanisms underlying these associations. In addition, because the physical distance over which linkage disequilibrium (LD) extends in African Americans is smaller than in European Americans, these studies also permit more accurate fine mapping of common susceptibility alleles (5).

Determining genetic risk factors for CRC in African Americans is especially important because this population has the highest CRC incidence and mortality rates of any US population (6). Even when controlling for tumor stage, socioeconomic status and co-morbidities, disparities persist between African and European Americans (7), suggesting that biological factors including genetics and environmental exposures play a role in the disparities. Genetic studies in CRC including sample sizes over 400 cases of African Americans are limited. In the present study, we sought to replicate all previously reported GWAS SNP associations, including chromosome 8q24.21 SNPs, using a large sample of African and European Americans.

Materials and Methods

Cases and Controls

Cases and controls were obtained from the University of Chicago (UC) and the University of North Carolina (UNC). In total, we included DNA from 1194 CRC cases (803 UC and

391 UNC) and 1352 controls (935 UC and 417 UNC). Samples from the UC were obtained from formalin-fixed, paraffin-embedded archived surgical specimens (n=1209) and blood specimens (n=523). The details of sample collection from archived surgical specimens have been validated and described previously (8). Blood samples from UC cases and controls were also obtained prospectively in the oncology and other non-cancer clinics since 2006. The control subjects also included individuals found to have a normal screening colonoscopy or cancer-free individuals obtained from the UC Translational Research Initiative in the Department of Medicine (T.R.I.D.O.M). T.R.I.D.O.M is an ongoing, large-scale clinic-based sample repository, assembled to investigate the relationships of biomarkers with health status, disease status, and disease progression. Subjects over the age of 18 were recruited and consented from various UC outpatient clinics beginning in 2005. Consented individuals had 10 cc of peripheral blood drawn and deidentified samples were banked. We included DNA from control subjects only if they were cancer-free. The age at time of sample collection was used as the age for each control. Cases and controls from archived surgical specimens were matched by age, gender, and ancestry. Cases and controls from blood samples were not matched by any clinical characteristics. Clinical characteristics including age, gender, and ancestry were used in regression models to control for differences between UC cases and controls.

Samples from UNC were obtained through a large-scale, population-based case-control study of colon and rectal cancer, conducted in a 33 county area in central and eastern North Carolina. Cases were drawn at random from all CRC cases reported to the North Carolina Central Cancer Registry. Controls were randomly selected from North Carolina Division of Motor Vehicle records, based on sampling probabilities within blocks defined by 5-year age group, sex, and ancestry, using the technique of randomized recruitment (9). The details of this study have been published previously (10). The UC and UNC studies were approved by their respective institutional review boards, and where appropriate, subjects provided written informed consent.

Genotyping

Germline DNA from normal tissue was prepared from both archived surgical specimens and from blood specimens as described previously (8,10). A total of 22 SNPs previously associated with CRC (1) and 100 ancestry informative markers (11) were genotyped using the Sequenom MassARRAY platform as described previously (8). We tested for departures from Hardy-Weinberg equilibrium (HWE) in cases and controls separately. Because all the SNPs genotyped in the two ethnic control groups had HWE p-values > 0.01 (Supplementary Table 1), we included all these SNPs in the statistical analysis. We note that HWE tests yielded similar p-values in tests of genotypes obtained from DNAs prepared from archived surgical specimens and from blood specimens. Seven additional SNPs from 8q24.21 region 3 were selected for genotyping based on a chi-square > 3.0 in African Americans in the MEC study (12). Two of these SNPs failed HWE or quality control measures and were not included in subsequent analyses. Genotyping quality control for all SNPs was assessed using blinded duplicate genotyping for 24 DNA samples. A genotype concordance rate of 100% was observed for all markers. Genotyping call rates exceeded 99.4% for all individuals included in the analyses.

Genetic ancestry estimation

The genomes of admixed populations such as African Americans are comprised of different genetic segments arising from different “parental” populations (e.g., West Africans and Europeans). Genetic association studies in admixed populations can be confounded by population stratification in which false-positive disease associations arise due to ancestry differences in cases and controls. In order to control for such confounding, West African

ancestry was estimated in cases and controls using genetic variants called ancestry informative markers (AIMs). AIMs are markers selected based on their frequency differences between populations from different geographic regions. In the present study, "global" individual ancestry was determined for each individual using 100 AIMs selected from regions across the entire genome to estimate West African and European ancestry (11). Global individual ancestry (% West African and % European) was calculated from the genotype data using the Bayesian Markov Chain Monte Carlo (MCMC) method implemented in the program STRUCTURE 2.1 (13). STRUCTURE 2.1 assumes an admixture model using prior population information and independent allele frequencies. The MCMC model was run using K=3 populations (58 Europeans, 67 Native Americans, and 62 West Africans) and a burn-in length of 30,000 iterations followed by 70,000 replications. These ancestry estimates were used as covariates in the regression models. Additional detail on genetic association studies in admixed populations can be found in reference 14.

Statistical Analysis

We tested the 22 SNPs from GWAS for association with CRC in both the combined African American study group, individual UC and UNC African American study groups as well as UC European Americans. We calculated odds ratios and 95% confidence intervals using logistic regression assuming an additive effect (on the log scale) of allele dosage. We controlled for individual admixture by including West African ancestry estimates as a covariate in the logistic regression model. In addition, we controlled for age and gender in the logistic regression model. For stratified association testing by anatomic site, we considered colon cases to be cancer in the proximal to the sigmoid colon and rectal cases to be cancer in the rectum or rectosigmoid junction. For tests of odds ratio heterogeneity, we used the Breslow-Day test. All analyses were done using the program PLINK (15).

Measures of LD, including r^2 and D', LD plots and haplotype association analyses were performed using Haploview (16). Clinical characteristics were compared between cases and controls by ancestry. Two-sided t-tests were used to compare continuous variables including age and ancestry estimates. Pearson chi-square tests of independence were used to compare categorical variables.

Results

The clinical characteristics of CRC cases and controls are shown in Table 2. In total, we included 795 African American and 399 European American cases and 985 African American and 367 European American controls. Because our study included samples from several different studies and geographical locations, we found differences in age and gender between African American cases and controls. Given these differences, we controlled for age, gender, and ancestry in our association analysis.

Results of association testing for the combined and individual African American study groups are shown in Table 3 and Supplementary Table 2, respectively. On chromosome 10p14, significant association with CRC in African Americans was detected with rs10795668 (p=0.04). Association with CRC risk was detected with the A allele, whereas in previous reports in Europeans association was detected with the G allele. On chromosome 15q13.3, significant association was detected with rs10318 (p=0.04). On chromosome 8q24.21, associations were not detected with rs6983267 or rs7837328, but a trend for association was noted with rs7014346 (p=0.06). On chromosome 16q22.1, a trend for association was noted with rs1862748 (p=0.07). Comparison of odds ratios for association obtained in the combined African American study groups and the odds ratios obtained in the European meta-analysis (1) are presented in Supplementary Table 3.

In the UC African American study group, the strongest association signals after adjusting for covariates were found for rs9929218 on chromosome 16q22.1 (p=0.008), rs10318 on chromosome 15q13.3 (p=0.03) and rs10795668 on chromosome 10p14 (p=0.05) (Supplementary Table 2). We note that the association with CRC risk in the UC study group at rs9929218 was with the opposite allele compared to the association in European CRC (Table 1; reference 1). In the UNC study group, a trend for association was noted for rs3802842 on chromosome 11q23.1 (p=0.06) and for rs1862748 on chromosome 16q22.1 (p=0.07) (Supplementary Table 2). Tests for heterogeneity are presented in Supplementary Table 2. We note that the lowest P_{het} (p=0.01) in the Table was detected at rs9929218 on chromosome 16p22.1; however, we have no statistically significant evidence that genetic heterogeneity between the two African American populations explains differences in odds ratios after taking into account multiple tests.

Results of association testing for the European American study group are shown in Table 4. On chromosome 8q24.21, a significant association with CRC in European Americans was detected with rs7837328 (p=0.03), and trends for association were noted with the other two 8q24.21 SNPs tested rs6983267 and rs7014346 (p=0.08 and 0.09, respectively). On chromosome 11q23.1, significant association was detected for rs3802842 (p=0.02) and rs10749971 (p=0.0002). A trend was noted for a third chromosome 11q23.1 SNP rs11213809 (p=0.09). On chromosome 16q22.1, significant association was detected with rs1862748 (p=0.04). Trends for association were noted with rs16892766 on chromosome 8q23.3 (p=0.08) and rs10411210 on chromosome 19q13.11 (p=0.07).

We also analyzed associations stratified by anatomic site (colon versus rectum). The most significant results in the African American study group are presented in Table 5. In the colon sub-group, we found a significant association on chromosome 10p14 (rs10795668; p=0.01) and a trend for association on chromosome 15q13.3 (rs10318; p=0.06). In the rectal sub-group, we found significant associations on chromosomes 8q24.21 (rs6983267; p=0.04), 11q23.1 (rs3802842; p=0.01), and 20p12.3 (rs355527; p=0.02). Complete association results by anatomic site and ancestry are shown in Supplementary Tables 4A and 4B.

Because SNPs in 8q24.21 region 3 have exhibited associations in multiple CRC case-control studies, we genotyped additional SNPs from the 8q24.21 region that previously had displayed evidence of association in African Americans in the Multi-Ethnic Cohort (12). Combining the results on six SNPs for which we had genotype information from all three study groups (Table 6), we found that the association p-value for rs7014346 was significant (p=0.03), and trends for association were noted for the SNPs rs12682374 (p=0.08) and rs10808556 (p=0.09). P-values calculated for this set of SNPs conditioned on rs7014346 provided no evidence for multiple independent risk factors (data not shown). Haplotype analysis for five 8q24.21 SNPs that formed a haplotype block showed trends for associations in the combined analysis of African Americans, with one haplotype associated with risk (CCGTA, p=0.06) and two haplotypes associated with protection (CCGTG and GTTCG, p=0.07 and 0.08, respectively) (Supplementary Table 5). In European Americans, two of these haplotypes also showed trends for association (CCGTA, p=0.08, associated with risk; and GTTCG, p=0.06, associated with protection) (Supplementary Table 5).

Discussion

GWAS in CRC have unveiled a number of genetic associations in various candidate genes and gene deserts (Table 1). These discoveries hold promise for elucidation of disease pathogenesis and eventually may prove useful clinically as biomarkers for risk stratification. However, these studies have been limited to European populations. We present here the first study to test genetic risk factors, identified in GWAS, in a large African American and

European American case-control study. While the association results in our European American study group were consistent overall with the meta-analysis results (1), some of the results in our African American study group were not consistent with the results in European populations, pointing to differences in population structure and possibly also in the underlying basis of genetic susceptibility.

Because patterns of LD differ between different ethnic populations, even in the absence of an association with a particular SNP tested, other SNPs in a region could be associated with the same common, risk-causing allele. The CRC-associated SNPs in the 8q24.21 region have been widely replicated in persons of European descent (12, 17-21), and we obtained evidence for association with all three 8q24.21 SNPs tested (rs6983267, rs7837328, and rs7014346) in our European American study group. However, in our African American study group, in which we had twice the numbers of cases and controls, rs6983267 was not significantly associated with CRC even after combining our data with data from a previously published study of African American CRCs (12). On the other hand, we detected a significant association with rs 7014346 (p=0.03) and trends for association with rs12682374 and rs10808556 (p=0.08 and 0.09, respectively) in the combined analysis (Table 6). The analysis of cases stratified by site did reveal a significant association with rs6983267 confined to rectal cancer cases in African Americans (Table 5); however, we would interpret this finding cautiously due to the small number of rectal cancer cases in the sample (Table 2).

One explanation for the 8q24.21 results is that rs7014346 is a better marker for the true risk allele than rs6983267 in African American CRC. Whereas the correlation coefficient (r^2) between rs6983267 and rs7014346 in the European American study group is 0.55, the correlation coefficient in the African American study group is 0.08; consequently, there is more power to discriminate between the effects associated with these two SNPs in the African American study group than in person of European ancestry. Although two studies concluded that rs6983267 is a functional SNP (2, 3), we note that two more recent studies have reached opposite conclusions about the effects of rs6983267 on *MYC* gene transcription (22, 23). Overall, replication of associations on chromosome 8q24.21 in African Americans provides evidence that CRC associations can be shared across different populations, possibly association with the same common, ancestral risk allele.

We also replicated an association on 15q13.3 in African American CRC (Table 3). The 15q13.3 region harbors a CRC candidate gene *GREM1*, which is an antagonist of bone morphogenic proteins (members of the TGF-β superfamily of growth factors)(24). Previous linkage studies have identified this region as associated with hereditary mixed polyposis syndrome and familial CRC (25, 26), further supporting its role in CRC pathogenesis. Our results show that the association in African Americans is with rs10318, which localizes to the 3′ untranslated region of *GREM1*, whereas in European Americans we found association with rs4779584, which is more than 12 kb proximal to *GREM1.* rs4779584 and rs10318 are more strongly correlated in persons of European ancestry (r^2=0.53) than in African Americans (r^2=0), suggesting again that rs10318 may be a better marker in the chromosome 15q13.3 region (Supplementary Figure 1). Similarly, although association with SNPs on chromosomes 8q23.3 and 19q13.11 in African Americans were not significant, whilst in European Americans they were, the associations were in the same direction and the confidence intervals were overlapping, suggesting that these SNPs are not good markers for CRC susceptibility in African Americans. Further fine mapping studies are needed to confirm these results.

On chromosome 11q23.1, in a gene desert, we have evidence for SNP associations in both populations. In the combined African American study group, the association with rs3802842

was more significant in rectal compared to colon cancer cases in African Americans (Table 5). This site-specific association was also noted in two previous studies (18, 27). In the UNC African American study group, there was a trend with rs3802842 that was not detected in the UC African American study group, while associations were not detected with the other two chromosome 11q23.1 SNPs in either African American study group. There was evidence for association with all three chromosome 11q23.1 SNPs (which are in strong LD in Europeans) in our European American study group.

On chromosome 10p14, also in a gene desert, we found that the opposite allele—the A allele—of rs10795668 was significantly associated with CRC risk in African Americans compared to the G allele associated with risk in previous studies in Europeans (28). The A allele was associated with risk in both the UC and UNC study groups, separately, but these individual results did not reach statistical significance. Although it is possible that these results are explained by different patterns of LD in the 10q14 region (Supplementary Figure 2), it is also possible that different, low-frequency risk-causing alleles originated in the two ethnic populations, and these are associated with different SNPs.

On chromosome 16q22.1, where the gene for E-cadherin (*CDH1*) is localized, we detected associations in both populations. In our European American study group, the A allele of rs1862748 was found to be a protective allele as previously reported (1). In African Americans, we found the same allele for this SNP showed a trend for a protective association. However, the other chromosome 16p22.1 SNP tested, rs9929218, was not associated in African Americans. In fact, the effect was in the opposite direction, and in the UC African American study group the confidence intervals did not overlap (compare Table 1 and Supplementary Table 2). These two chromosome 16p22.1 SNPs rs9929218 and rs1862748 are in the same LD block in the European population, but they are not correlated in the Yoruban population (Supplementary Figure 3). rs1862748 may be a better marker for CRC risk than rs9929218 in African Americans, or there could be different risk alleles of *CDH1* in different populations.

We did not find evidence for association overall on chromosomes 14q22.2, 18q21.1, or 20p12.3 in either population. In our sub-group analysis by anatomic site, significant associations were noted for the two chromosome 14q22.2 SNPs in European American rectal cancer cases and one chromsome 20p12.3 SNP in African Americans rectal cancer cases (Supplementary Table 5). Interestingly, when comparing the odds ratios and confidence intervals in African Americans in these regions with those reported previously in Europeans (1), with the exception of rs12953717 on chromosome 18, we found the associations were in opposite directions with no overlap in the confidence intervals, strongly suggesting that these SNPs are not associated with CRC risk in African Americans. These results do not rule out the possibility that other SNPs in these or other regions around the genome will be found associated with CRC in African Americans.

In summary, we have replicated a number of CRC SNP associations in a large group of African Americans and European Americans. We anticipated that estimation of effect sizes might not be precise in our European American study group due to small sample size, but we were surprised by the overall lack of association obtained in African Americans despite having two times the sample size. Indeed, our results provide strong evidence that differences in LD patterns between African and European Americans could explain the lack of associations detected in our African Americans (e.g., the 8q24.21, 15q13.3, and possibly other SNPs). In addition, based on the overall lack of associations in African American CRC, we suspect that rare, population-specific risk alleles may explain some of the associations in CRC (e.g., the 10p14 SNP), pointing to genetic heterogeneity in susceptibility alleles. Future fine-mapping and deep sequencing studies are needed to

determine whether or not other SNPs can be found associated in African Americans as well as to identify both common or rare risk-causing alleles in the associated regions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Christopher Haiman and the MEC investigators for providing us with raw genotype data for the combined 8q24.21 analyses. We thank the University of Chicago Genotyping Core for assistance with genotyping, Dr. Joe G.N. Garcia and Dr. Karen Kim for providing control DNA samples, and Mr. Cory Henderson for his help on the pilot association study at the University of Chicago. We thank Ms. Amber McCoy and Mr. Kevin Smith of the University of North Carolina Center for Gastrointestinal Biology and Disease for their assistance with DNA preparation and shipment.

Grant Support: This work was supported by The Cancer Research Foundation, the Department of Medicine at the University of Chicago, the Digestive Disease Research Core Center (P30 DK42086), American Cancer Society Institutional Research Grant, the University of Chicago Comprehensive Cancer Research Center and Clinical Translational Science Award, the University of North Carolina Digestive Disease Research Core Center (P30 DK34987), and a grant from the National Institutes of Health (R01 CA 66635).

Abbreviations

GWAS	genome wide association studies
CRC	colorectal cancer
SNP	single nucleotide polymorphism
LD	linkage disequilibrium

References

1. Houlston RS, Webb E, Broderick P, et al. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. Nat.Genet. 2008; 40:1426–1435. [PubMed: 19011631]
2. 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; 41:882–884. [PubMed: 19561607]
3. 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; 41:885–890. [PubMed: 19561604]
4. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat.Genet. 2008; 40:695–701. [PubMed: 18509313]
5. Tenesa A, Dunlop MG. New insights into the aetiology of colorectal cancer from genome-wide association studies. Nat.Rev.Genet. 2009; 10:353–8. [PubMed: 19434079]
6. American Cancer Society. Cancer Facts & Figures. 2008.
7. Alexander DD, Waterbor J, Hughes T, Funkhouser E, Grizzle W, Manne U. African-American and Caucasian disparities in colorectal cancer mortality and survival by data source: an epidemiologic review. Cancer.Biomark. 2007; 3:301–313. [PubMed: 18048968]
8. Kupfer SS, Torres JB, Hooker S, et al. Novel single nucleotide polymorphism associations with colorectal cancer on chromosome 8q24 in African and European Americans. Carcinogenesis. 2009; 30:1353–1357. [PubMed: 19520795]
9. Weinberg CR, Sandler DP. Randomized recruitment in case-control studies. Am.J.Epidemiol. 1991; 134:421–432. [PubMed: 1877602]

10. Sansbury LB, Millikan RC, Schroeder JC, et al. COX-2 polymorphism, use of nonsteroidal anti-inflammatory drugs, and risk of colon cancer in African Americans (United States). Cancer Causes Control. 2006; 17:257–266. [PubMed: 16489533]
11. Robbins C, Torres JB, Hooker S, et al. Confirmation study of prostate cancer risk variants at 8q24 in African Americans identifies a novel risk locus. Genome Res. 2007; 17:1717–1722. [PubMed: 17978284]
12. Haiman CA, Le Marchand L, Yamamato J, et al. A common genetic risk factor for colorectal and prostate cancer. Nat.Genet. 2007; 39:954–956. [PubMed: 17618282]
13. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003; 164:1567–1587. [PubMed: 12930761]
14. Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M. Genome-wide association studies in diverse populations. Nat.Rev.Genet. 2010; 11:356–366. [PubMed: 20395969]
15. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am.J.Hum.Genet. 2007; 81:559–575. [PubMed: 17701901]
16. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21:263–265. [PubMed: 15297300]
17. Tomlinson I, Webb E, Carvajal-Carmona L, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat.Genet. 2007; 39:984–988. [PubMed: 17618284]
18. Tenesa A, Farrington SM, Prendergast JG, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nat.Genet. 2008; 40:631–637. [PubMed: 18372901]
19. Li L, Plummer SJ, Thompson CL, et al. A common 8q24 variant and the risk of colon cancer: a population-based case-control study. Cancer Epidemiol.Biomarkers Prev. 2008; 17:339–342. [PubMed: 18268117]
20. Berndt SI, Potter JD, Hazra A, et al. Pooled analysis of genetic variation at chromosome 8q24 and colorectal neoplasia risk. Hum.Mol.Genet. 2008; 17:2665–2672. [PubMed: 18535017]
21. Schafmayer C, Buch S, Volzke H, et al. Investigation of the colorectal cancer susceptibility region on chromosome 8q24.21 in a large German case-control sample. Int.J.Cancer. 2009; 124:75–80. [PubMed: 18839428]
22. Sotelo J, Esposito D, Duhagon MA, et al. Long-range enhancers on 8q24 regulate c-Myc. Proc.Natl.Acad.Sci.U.S.A. 2010; 107:3001–3005. [PubMed: 20133699]
23. Wright JB, Brown SJ, Cole MD. Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. Mol.Cell.Biol. 2010; 30:1411–1420. [PubMed: 20065031]
24. Wordinger RJ, Zode G, Clark AF. Focus on molecules: gremlin. Exp.Eye Res. 2008; 87:78–79. [PubMed: 18201700]
25. Tomlinson I, Rahman N, Frayling I, et al. Inherited susceptibility to colorectal adenomas and carcinomas: evidence for a new predisposition gene on 15q14-q22. Gastroenterology. 1999; 116:789–795. [PubMed: 10092300]
26. Daley D, Lewis S, Platzer P, et al. Identification of susceptibility genes for cancer in a genome-wide scan: results from the colon neoplasia sibling study. Am.J.Hum.Genet. 2008; 82:723–736. [PubMed: 18313025]
27. Pittman AM, Webb E, Carvajal-Carmona L, et al. Refinement of the basis and impact of common 11q23.1 variation to the risk of developing colorectal cancer. Hum.Mol.Genet. 2008; 17:3720–3727. [PubMed: 18753146]
28. Tomlinson IP, Webb E, Carvajal-Carmona L, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. Nat.Genet. 2008; 40:623–630. [PubMed: 18372905]
29. Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat.Genet. 2008; 40:638–645. [PubMed: 18372903]

Table 1

Single nucleotide polymorphisms associated with colorectal cancer in persons of European ancestry*

Region (Gene)†	rs number	Position (bp)	Minor allele	Allelic OR [95% CI]§
8q23.3 (*EIF3H*)	rs16892766	117630683	C	1.32 [1.21,1.44]
	rs11986063	117640315	T	1.29 [1.19,1.40]
	rs6983626	117802148	T	1.21 [1.11,1.31]
8q24.21 (*MYC*)	rs6983267	128413305	G	1.20 [1.15,1.27]
	rs7837328	128423127	A	1.17 [1.12,1.23]
	rs7014346	128424792	A	1.20 [1.14,1.26]
10p14	rs10795668	8701219	A	0.91 [0.86,0.96]
11q23.1 (*POU2AF1*)	rs11213809	111135745	A	1.20 [1.14,1.26]
	rs3802842	111171709	C	1.21 [1.15,1.27]
	rs 10749971	111189158	G	1.13 [1.08,1.19]
14q22.2 (*BMP4*)	rs4444235	54410919	C	1.12 [1.07,1.18]
	rs17563¶	54417522	C	--
15q13.3 (*GREM1*)	rs4779584	32994756	T	1.19 [1.12,1.26]
	rs10318	33025979	T	1.18 [1.11,1.25]
16q22.1 (*CDH1*)	rs9929218	68820946	A	0.88 [0.83,0.92]
	rs1862748	68832943	T	0.88 [0.84,0.93]
18q21.1 (*SMAD7*)	rs4939827	46453463	C	0.85 [0.81,0.89]
	rs12953717	46453929	T	1.19 [1.13,1.25]
19q13.11 (*RHPN2*)	rs10411210	33532300	T	0.79 [0.72,0.86]
	rs7259371	33534641	A	0.86 [0.81,0.92]
20p12.3 (*BMP2*)	rs355527	6388068	A	1.13 [1.08,1.19]
	rs961253	6404281	A	1.13 [1.08,1.19]

*All odds ratios shown in this table were taken from a meta-analysis of 21 of these single nucleotide polymorphisms as reported in reference 1; no data from the present study were included.

†Cytogenetic chromosome position is given. In some cases, a polymorphism in the chromosomal region associated with colorectal cancer is within the indicated gene (14q22.2, 15q13.3, 16q22.1, 18q21.1, and 19q13.11) and in other cases the polymorphism is within 200 kb of the indicated gene (8q23.3, 8q24.21, 11 q23.1, 20p12.3).

§OR, odds ratio; CI, confidence interval.

¶This additional single nucleotide polymorphism in the 14q22.2 region, which was reported in reference 1, was included because it is non-synonymous. An OR was not reported for rs17563.

Gastroenterology. Author manuscript; available in PMC 2013 July 24.

Table 2

Clinical characteristics of African American and European American colorectal cancer cases and controls by center

		African Americans		European Americans	
	Study group	**Cases**	**Controls**	**Cases**	**Controls**
Number of participants	Combined	795	985	399	367
	UC	*404*	*568*		
	UNC	*391*	*417*		
Mean age in years (SD)	Combined	64.5 (11.7) *	62.3 (13.2) *	64.6 (13.1) *	61.1 (12.7) *
	UC	*67.3 (12.7) **	*60.2 (15.8) **		
	UNC	*61.8 (10.0) **	*65.2 (9.6) **		
Gender (female/male)	Combined	423/372 **	570/412 **	170/229	181/183
	UC	*230/172 †*	*375/190 †*		
	UNC	*193/200*	*195/222*		
% Mean West African ancestry (SD)	Combined	84.1 (14.0)	85.5 (14.3)	1.0 (1.5)	0.9 (1.3)
	UC	*85.6 (14.8) ††*	*87.8 (13.7) ††*		
	UNC	*82.7 (12.9)*	*82.4 (14.6)*		
Anatomic site					
Colon (%)	Combined	605 (76.4)		248 (62.3)	
	UC	*337 (84.0)*			
	UNC	*268 (68.5)*			
Rectum (%)	Combined	187 (23.6)		150 (37.7)	
	UC	*64 (16.0)*			
	UNC	*123 (31.5)*			

all other p-values ≥ 0.05; UC = University of Chicago; UNC = University of North Carolina; colon=proximal to and including sigmoid colon; rectum=rectal & rectosigmoid cancers

* p-value for heterogeneity < 0.0001;

** p-value= 0.04;

† p-value=0.004;

†† p-value=0.02;

Table 3

Odds ratios, 95% confidence intervals and p-values for association between the GWAS SNPs and colorectal cancer in the combined African American cohort

			Allele frequencies			
Chr	**SNP**	**Minor allele***	**Cases**	**Controls**	**OR [95% CI]†**	**P**
8q23.3	rs16892766	C	0.14	0.13	1.15[0.93,1.41]	0.19
	rs11986063	T	0.20	0.19	1.09[0.92,1.30]	0.33
	rs6983626	T	0.40	0.39	1.05[0.90,1.21]	0.55
8q24.21	rs6983267	G	0.90	0.88	1.15[0.92,1.45]	0.21
	rs7837328	A	0.65	0.63	1.09[0.94,1.27]	0.24
	rs7014346	A	0.42	0.39	1.15[1.00,1.32]	0.06
10p14	rs10795668	A	0.08	0.06	1.35[1.01,1.81]	0.04
11q23.1	rs11213809	A	0.22	0.22	0.99[0.83,1.17]	0.87
	rs3 802842	C	0.36	0.34	1.13[0.98,1.30]	0.10
	rs 10749971	G	0.13	0.13	0.92[0.75,1.14]	0.44
14q22.2	rs4444235	C	0.31	0.33	0.89[0.76,1.04]	0.13
	rs17563	C	0.26	0.25	1.02[0.87,1.20]	0.79
15q13.3	rs4779584	T	0.56	0.55	1.05[0.92,1.22]	0.45
	rs10318	T	0.05	0.03	1.45[1.02,2.07]	0.04
16q22.1	rs9929218	A	0.30	0.29	1.06[0.91,1.24]	0.44
	rs1862748	T	0.16	0.18	0.84[0.69,1.01]	0.07
18q21.1	rs4939827	C	0.69	0.68	1.08[0.92,1.25]	0.39
	rs12953717	T	0.29	0.27	1.13[0.96,1.32]	0.14
19q13.11	rs10411210	T	0.41	0.42	0.97[0.83,1.12]	0.64
	rs7259371	A	0.69	0.69	0.99[0.85,1.15]	0.90
20p12.3	rs355527	A	0.18	0.20	0.87[0.73,1.05]	0.14
	rs961253	A	0.36	0.37	0.93[0.80,1.07]	0.30

P values were adjusted for West African ancestry, age and gender. GWAS, genome-wide association studies; Chr, cytogenetic chromosome position; SNP, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval.

*Minor allele per reference 1.

†Per-allele odds ratios were calculated assuming an additive model.

Table 4

Odds ratios, 95% confidence intervals and p-values for association between the GWAS SNPs and colorectal cancer in the European American cohort

			Allele frequencies			
Chr	**SNP**	**Minor allele***	**Cases**	**Controls**	**OR[95%CI]†**	**P**
8q23.3	rs16892766	C	0.09	0.07	1.43[0.96,2.12]	0.08
	rs11986063	T	0.11	0.10	1.09[0.77,1.54]	0.63
	rs6983626	T	0.12	0.09	1.26[0.90,1.77]	0.17
8q24.21	rs6983267	G	0.52	0.47	1.20[0.98,1.471]	0.08
	rs7837328	A	0.43	0.37	1.27[1.03,1.571]	0.03
	rs7014346	A	0.38	0.33	1.21[0.97,1.511]	0.09
10p14	rs10795668	A	0.27	0.30	0.90[0.71,1.131]	0.36
11q23.1	rs11213809	A	0.35	0.31	1.21[0.97,1.521]	0.09
	rs3802842	C	0.34	0.29	1.30[1.04,1.631]	0.02
	rs 10749971	G	0.42	0.33	1.54[1.23,1.931]	2×10^{-4}
14q22.2	rs4444235	C	0.51	0.48	1.12[0.91,1.381]	0.29
	rs17563	C	0.51	0.55	0.85[0.69,1.051]	0.15
15q13.3	rs4779584	T	0.19	0.19	0.96[0.74,1.231]	0.73
	rs10318	T	0.17	0.18	0.91[0.69,1.191]	0.47
16q22.1	rs9929218	A	0.27	0.31	0.83[0.67,1.041]	0.10
	rs1862748	T	0.27	0.33	0.80[0.64,0.991]	0.04
18q21.1	rs4939827	C	0.47	0.48	0.93[0.75,1.141]	0.47
	rs12953717	T	0.45	0.43	1.07[0.87,1.321]	0.53
19q13.11	rs10411210	T	0.09	0.13	0.74[0.54,1.021]	0.07
	rs7259371	A	0.18	0.22	0.83[0.64,1.081]	0.16
20p12.3	rs355527	A	0.33	0.30	1.14[0.91,1.421]	0.25
	rs961253	A	0.35	0.34	1.05[0.84,1.311]	0.67

P values were adjusted for West African ancestry, age and gender. GWAS, genome-wide association studies; Chr, cytogenetic chromosome position; SNP, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval.

*Minor allele per reference 1.

†Per-allele odds ratios were calculated assuming an additive model.

Table 5

Selected associations in combined African American study group stratified by anatomic site_

			Colon		Rectum	
Chr	**SNP**	**Allele**	**OR [95% CI]**	**P-value**	**OR [95% CI]**	**P-value**
8q24.21	rs6983267	G	1.06 [0.84, 1.35]	0.59	1.56 [1.02, 2.38]	0.04
10p14	rs10795668	A	1.48 [1.09, 2.02]	0.01	0.93 [0.55, 1.56]	0.77
11q23.1	rs3802842	C	1.05 [0.90, 1,23]	0.52	1.34 [1.06, 1.69]	0.01
15q13.3	rs10318	T	1.44 [0.98, 2.12]	0.06	1.42 [0.80, 2.53]	0.23
20p12.3	rs355527	A	0.95 [0.78, 1.15]	0.60	0.67 [0.49, 0.93]	0.02

Chr, cytogenetic chromosome position.

Table 6

Odds ratios and p-values for associations between 8q24.21 SNP and colorectal cancer in African Americans from three study groups

		OR				p-value			
SNP	**Position**	**UC**	**UNC**	**MEC**	**Combined**	**UC**	**UNC**	**MEC**	**Combined**
rs10505476	128408116	0.97	0.93	0.77	0.90	0.75	0.51	0.06	0.11
rs12682374	128410948	1.03	0.84	0.74	0.89	0.77	0.15	0.04	0.08
rs10808556	128413147	1.04	0.81	0.68	0.86	0.76	0.20	0.04	0.09
rs6983267	128413305	1.04	1.23	1.28	1.18	0.75	0.24	0.20	0.12
rs7013278	128414892	1.04	0.98	0.78	0.95	0.70	0.89	0.07	0.40
rs7014346	128424792	1 .1 7	1.09	1.21	1.15	0.13	0.43	0.15	0.03

P-values were adjusted for West African ancestry and gender. The p-values were combined by the method of Zeggini et al. (29). P_{het} was >0.3 for all SNPs, as determined by Breslow-Day test. OR, odds ratio; UC, University of Chicago; UNC, University of North Carolina; MEC, Multi-Ethnic cohort.