Published in final edited form as:
Gut. 2011 December ; 60(12): 1703–1711. doi:10.1136/gut.2011.237727.

Type-2 Diabetes Risk Variants and Colorectal Cancer Risk: The Multiethnic Cohort and PAGE Studies

Iona Cheng[1], **Christian P. Caberto**[1], **Annette Lum-Jones**[1], **Ann Seifried**[1], **Lynne R. Wilkens**[1], **Fredrick R. Schumacher**[2], **Kristine R Monroe**[2], **Unhee Lim**[1], **Maarit Tiirikainen**[1], **Laurence N. Kolonel**[1], **Brian E. Henderson**[2], **Daniel O. Stram**[2], **Christopher A. Haiman**[2], and **Loïc Le Marchand**[1]

[1]Epidemiology Program, University of Hawai`i Cancer Center, University of Hawai`i, Honolulu, HI 96813, USA

[2]Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA

Abstract

Background—Diabetes has been positively associated with the risk of colorectal cancer. We investigated whether recently established risk variants for diabetes also have effects on colorectal cancer.

Methods—Nineteen SNPs associated with type-2 diabetes (T2D) in genome-wide association studies were tested in a case-control study of 2,011 colorectal cancer cases and 6,049 controls nested in the Multiethnic Cohort as part of the Population Architecture using Genomics and Epidemiology (PAGE) initiative. Odds ratios (OR) and 95% Confidence Intervals (CI) were estimated by unconditional logistic regression to evaluate the association between SNPs and colorectal cancer risk, adjusting for age, sex, and race/ethnicity. Permutation testing was conducted to correct for multiple hypothesis testing.

Results—Four type 2 diabetes SNPs were associated with colorectal cancer risk: rs7578597 (*THADA*), rs864745 (*JAZF1*), rs5219 (*KCNJ11*), and rs7961581 (*TSPAN8, LGR5*). The strongest association was for the rs7578597 (*THADA*) Thr1187Ala missense polymorphism (P_{trend} = 0.004 adjusted for multiple testing) with the high risk allele for colorectal cancer being the low risk allele for diabetes. Similar patterns of associations were seen with further adjustment for diabetes status and body mass index. The association of diabetes status with colorectal cancer risk was somewhat weakened after adjustment for these SNPs.

Conclusion—Our findings suggest that diabetes risk variants also influence colorectal cancer susceptibility, possibly through different mechanisms than for diabetes.

Keywords

colorectal cancer; SNPs; type 2 diabetes

Corresponding Author: Iona Cheng, PhD, University of Hawaii Cancer Center, 1236 Lauhala Street, Suite 407, Honolulu, HI 96817, icheng@crch.hawaii.edu, Phone: 808.441.3483, Fax: 808.586.2982.

Introduction

Colorectal cancer and type 2 diabetes (T2D) are common chronic diseases that share several risk factors. Colorectal cancer is one of the leading cancers in the United States, representing 9% of all new cancer cases in 2010 [1]. By the end of 2021 approximately 13.5% of U.S. adults — or more than 32 million adults — will be living with diabetes [2]. The growing epidemic of diabetes is of relevance to future rates of colorectal cancer as epidemiologic studies have demonstrated an increased risk of colorectal cancer among diabetics [3]. A meta-analysis of 15 studies (nine cohort studies and six case-control studies) reported that diabetics had a significant 30% increased risk of developing colorectal cancer compared with non-diabetic individuals [3]. In a recent study of the Multiethnic Cohort (MEC), a similar finding was observed such that a 19% increased risk of colorectal cancer was found among diabetics in comparison to those without the disease (95% CI: 1.09-1.29; $P<0.001$) [4].

Shared risk factors for diabetes and colorectal cancer include aging, diet, obesity, and physical inactivity. Whether the underlying link between T2D and colorectal cancer is through these shared risk factors or whether T2D serves as a marker for biological states that influence cancer risk (e.g. hyperinsulinemia and/or hyperglycemia) has been debated [5]. The significant residual association between diabetes and colorectal cancer after adjusting for shared risk factors [6, 7, 8, 9] would suggest that the two diseases may also share intrinsic etiologic factors, such as genetic factors.

The few studies that have explored the relationship between T2D susceptibility variants and colorectal cancer risk have focused on a small number of genetic variants in transcription factor 7-like 2 (*TCF7L2*) [10, 11, 12], a key susceptibility loci identified by genome-wide association studies of T2D [13, 14, 15, 16, 17, 18, 19, 20]. Findings from these initial studies have reported inconsistent results between *TCF7L2* polymorphisms, rs7903146 and rs1255372, and colorectal cancer risk [10, 11, 12]. In addition, genetic association studies of *INSR*, which encodes for the insulin receptor, a key receptor in diabetes pathogenesis, have reported two *INSR* polymorphisms (rs1051690 and rs10426094) to be associated with colorectal cancer risk [21, 22, 23].

Genome-wide association studies (GWAS) of T2D have identified numerous risk loci [13, 14, 15, 16, 17, 18, 19, 20, 24, 25, 26, 27, 28, 29] and the impact of these genetic variants on colorectal cancer risk has yet to be fully examined. Thus, we tested 19 T2D risk variants for their association with colorectal cancer within a large case-control study nested in the Multiethnic Cohort Study as part of the Population Architecture using Genomics and Epidemiology (PAGE) initiative. Furthermore, we evaluated whether adjustment for these risk variants explained at least partially the association between diabetes status and colorectal cancer risk.

Methods

Study Subject

The Multiethnic Cohort Study is a large population-based cohort study of more than 215,000 men and women from Hawaii and Los Angeles. The cohort is composed predominantly of individuals from the following five racial/ethnic groups: African Americans, Native Hawaiians, Japanese, Latinos, and Whites. Participants between the ages of 45 and 75 years were recruited from March 1993 through May 1996. Participants completed a 26-page self-administered questionnaire that included information regarding medical history, family history of cancer, diet, dietary supplements and medication use, and physical activity. Further details about this cohort are provided elsewhere [30].

Incident colorectal cancers in MEC participants were identified (up to December 5, 2009) by cohort linkage to population-based Surveillance, Epidemiology and End Results (SEER) cancer registries covering Hawaii and California. Information on stage of disease at the time of diagnosis was also collected from the cancer registries. Blood samples were collected from incident colorectal, breast and prostate cancer cases after their diagnosis, as well as a random sample of cohort members to serve as controls, from 1996 through 2001, and prospectively from all surviving participants from 2002 through 2007. Informed consent was obtained at blood draw. Among the colorectal cancer cases used in this analysis, 69.4% had their blood drawn after diagnosis and 30.6% prior to diagnosis.

Control subjects were men and women without colorectal cancer before entry into the cohort and without a colorectal cancer diagnosis up to December 5, 2009. This case-control study consisted of 2,011 case patients with colorectal cancer and 6,049 control subjects. We had substantial power (80%) to detect modest effects (odds ratios of 1.18 and 1.32 for allele frequencies of 20% and 5%, respectively) under log-additive models (alpha=0.05, two-sided) [31]. In this study population, we had previously examined the generalizability of 11 colorectal cancer risk loci in multiple populations, largely confirming associations observed from GWAS of European Americans while also identifying variation in genetic associations across non-European populations [32]. Thus, we were able to test whether any new associations with gene variants were independent from established ones. This study was approved by the institutional review boards at the University of Hawaii and the University of Southern California.

SNP Selection and Genotyping

We selected 19 established risk variants for T2D reported by genome-wide association studies through September, 2009 (Table 1) [13, 14, 15, 16, 17, 18, 19, 20, 24, 25, 26, 27, 28, 29]. All SNPs were genotyped using either the TaqMan OpenArray or standard Taqman allelic discrimination assay (14 SNPs and 5 SNPs, respectively). All assays were undertaken by laboratory personnel blinded to colorectal case-control status. The average genotyping success rate across the 19 SNPs was 99.6% and the average genotype concordance rate for QC duplicates (>10% of all samples) was 100%. All SNPs were in Hardy-Weinberg equilibrium among controls for each racial/ethnic group (defined as $P > 0.01$ for at least 4 of the 5 racial/ethnic groups).

Statistical Analysis

To examine the association between each T2D SNP and colorectal cancer risk, odds ratios (OR) and 95% confidence intervals (CI) were estimated by unconditional logistic regression, adjusting for age, sex, and race/ethnicity. Each SNP was examined by co-dominant and log-additive genetic models. In addition, analyses were adjusted for diabetes status and body mass index (BMI) to examine possible confounding effects of these variables. A positive diagnosis of diabetes was based on any of the following: (a) a self-report of diabetes either at baseline questionnaire or follow-up questionnaires; (b) self-report of taking medication for T2D at the time of blood draw; (c) diabetes diagnosis from the Office of Statewide Health Planning and Development (California Residents); (d) diabetes diagnosis from Kaiser Permanente or Hawaii Medical Service Association (Hawaii residents). Non-diabetics were defined as having none of all of the aforementioned criteria. Additional adjustment for known risk factors for colorectal cancer (family history of colorectal cancer; dietary intake of fiber, calcium, folate, alcohol; vigorous physical activity; and smoking) did not notably alter results; thus, these factors were not included in our final multivariable models. Tests of heterogeneity for genetic effects were conducted for race/ethnicity, diabetes status, and BMI by including interaction terms between genotype and race/ethnicity, diabetes status, and BMI, respectively, in our regression models. Stratified analysis by race/ethnicity, diabetes status, and BMI are presented to evaluate consistency of effects across groups. In addition, heterogeneous effects by age-group, sex, anatomical site, and stage were evaluated for the four associated SNPs. All statistical significance levels (*P* values) presented are two-sided. Analyses were performed using SAS 9.1 (SAS Institute, Inc. Cary, North Carolina).

To correct for potential population stratification within our study population, genetic ancestry was estimated by principal component analysis using R software and included in our regression models [33]. Specifically, 109 ancestry informative markers reported by Kosoy et al. [34] that distinguish between the major continental groups found in America were genotyped on all subjects by the TaqMan Openarray assay. The first four principal components delineated individuals of African, Asian, European, Latino, and Native Hawaiian descent and were used as estimates of genetic ancestry.

To guide interpretation of nominally statistically significant associations, we conducted permutation testing, which can be used to obtain an estimate of statistical significance that is corrected for multiple-hypothesis testing [35]. Case-control status within strata of sex and racial/ethnic group was randomly permuted 10,000 times for the 19 SNPs. Permutation P values were determined by examining where the nominal P value for an “associated” SNP fell in relation to the distribution of minimal P values generated from the permuted data. For example, if a nominal P value of 0.05 marked the 25th percentile of this distribution, then the permutation P value would be 0.25.

Results

Main characteristics of the 8,060 subjects (cases/controls = 2,011/6,049) in our study are presented by case-control status in Table 2. Japanese Americans represented the largest racial/ethnic group followed by African Americans, Latinos, Whites, and Native Hawaiians.

Colorectal cancer cases were older than controls (mean = 70 years versus 67 years) with a higher proportion of males (54%) than females (46%). In addition, cases were more likely to have a BMI > 25 kg/m^2, a history of diabetes, a family history of colorectal cancer, lower mean intakes of calcium and folate, higher mean intake of alcohol, and greater mean pack-years of cigarette smoking than controls. The overall prevalence of diabetes among controls was 23.3% with the highest prevalence observed among Japanese Americans (28.4%) followed by African Americans (25.8%), Latinos (20.5%), Whites (12.8%), and Native Hawaiians (12.4%).

Four of the 19 T2D SNPs were nominally associated with colorectal cancer risk, adjusting for age, sex, and race/ethnicity: rs7578597 (*THADA*), rs864745 (*JAZF1)*, rs5219 (*KCNJ11)*, rs7961581 (*TSPAN8*) (Table 3). Interestingly, only the T2D risk allele of rs5219 (*KCNJ11)* conferred an increased risk of colorectal cancer, while for rs7578597 (*THADA*), rs864745 (*JAZF1)* and rs7961581 (*TSPAN8*), the T2D risk alleles were associated with decreased risk of colorectal cancer. Because, as in other populations, diabetes and BMI are risk factors for colorectal cancer in our study ($OR_{diabetes}$ = 1.26; 95% CI: 1.12-1.42 and $OR_{BMI>25}$ = 1.33; 95% CI: 1.19-1.48), we examined the associations between T2D SNPs and colorectal cancer risk with further adjustment for diabetes status and BMI (Table 3). Similar associations were observed for all 19 T2D SNPs with adjustment for diabetes and BMI. In particular, risk estimates for the top associations, rs7578597 (*THADA*) and rs5219 (*KCNJ11*), remained virtually unchanged ($OR_{rs7578597}$ = 0.84; 95% CI: 0.75-0.95 and OR_{rs5219} = 1.10; 95% CI: 1.01-1.19) with these adjustments. Permutation analysis revealed that a similar significant association as observed between rs7578597 and colorectal cancer would occur by chance <1% of the time ($P_{permutation}$ =.004).

Two of the four associated SNPs (rs7578597 and rs5219), which are both missense SNPs, demonstrated strong associations with colorectal cancer in racial/ethnic stratified analyses (Table 4). For rs7578597 (*THADA*), a significant inverse association with the diabetes risk (T) allele was observed for all groups except for African Americans and Native Hawaiians, with the strongest association seen among Japanese Americans (OR=0.52; 95% CI: 0.36-0.75; P = 0.0005). This SNP also displayed evidence of heterogeneity in effects across racial/ethnic groups (P_{het} = 0.008). For rs5219 (*KCNJ11*), a consistent positive association with the diabetes risk (T) allele was seen across all racial/ethnic groups with the exception of Japanese Americans. However, the test for heterogeneity was not significant for rs5219, as well as for the two other remaining SNPs (P_{het} > 0.19).

To further investigate how these SNPs operate in light of diabetes status and BMI, we conducted stratified analysis by these two risk factors and tested for heterogeneity (Table 5). Both rs7578597 (*THADA; P_{het}* = 0.13) and rs5219 (*KCNJ11; P_{het}* = 0.31) displayed the strongest associations among non-diabetics (n=5,727) ($OR_{rs7578597}$ = 0.78; P_{trend} = 9.0 x 10^{-4} and OR_{rs5219}=1.11; P_{trend} = 0.03) and no association was observed among diabetics (n=1,866) ($OR_{rs7578597}$=1.05; P_{trend} = 0.73 and OR_{rs5219}=0.95; P_{trend} = 0.58). The rs7578597 (*THADA*) variant was negatively associated with colorectal cancer risk among both individuals with a BMI < 25 and those with a BMI > 25, but more strongly so among non-overweight subjects (P_{het}= 0.06). Similar patterns of associations were observed when adjusting for BMI in diabetes stratified analysis and likewise with adjustment of diabetes

status in BMI stratified analysis. After adjustment for all 19 SNPs, we noted that BMI (P<0.0001) and diabetes status (P=0.053) remained significant predictors in the model for colorectal cancer risk.

The four associated loci (*THADA*, *JAZF1*, *KCNJ11*, *TSPAN8*) demonstrated similar associations across age-groups (<67 years vs. >67 years), anatomical site (colon vs. rectum), and stage (localized vs. advanced). However, heterogeneous effects by sex were observed for *THADA* and *KCNJ11* (P_{het}< 0.05). For *THADA*, rs7578597 was significantly associated with colorectal cancer risk among females (OR=0.81; 95% CI: 0.69-0.97) and a similar association, although not statistically significant, was observed among males (OR=0.90; 95% CI: 0.76-1.06) (P_{het}= 0.01). For *KCNJ11*, rs5219 was associated with colorectal cancer risk among males (OR=1.18; 95% CI: 1.05-1.31) and no association was seen among females (OR=1.01; 95% CI: 0.89-1.14) (P_{het}< 0.0001).

To examine the independent genetic effects of *THADA* (rs7578597) and *KCNJ11* (rs5219) on colorectal cancer risk, we utilized data from a previous report [32] of 11 colorectal cancer GWAS hits tested in the MEC and adjusted for these established risk variants in our statistical analysis. Similar significant associations were observed with adjustment for these SNPs for *THADA* (rs7578597; OR=0.83; 95% CI: 0.70-0.99) and *KCNJ11* (rs5219; OR=1.16; 95% CI: 1.03-1.31), demonstrating that the association with these diabetes risk variants and colorectal cancer are independent from those of the GWAS hits.

Furthermore, whereas diabetes remained significantly associated with colorectal cancer after adjustment for all 19 T2D variants, there was a slight attenuation in the risk estimate for diabetes (OR: before adjustment for these SNPs: 1.20, 95% CI: 1.06-1.36; after adjustment: 1.15, 1.00-1.32), suggesting that these genetic variants may confound the relationship between diabetes status and colorectal cancer risk.

In addition, similar associations were observed for *THADA* (rs7578597; OR=0.84; 95% CI: 0.72-0.98) and *KCNJ11* (rs5219; OR=1.15; 95% CI: 1.031-1.28) with adjustment of genetic ancestry to control for population structure (i.e. adjustment for the leading principal components that can distinguish between African, Asian, European, Latino, and Native Hawaiian ancestry), indicating that population stratification was not a source of bias.

Discussion

Our study examined 19 T2D established risk variants for their association with colorectal cancer in a large multiethnic population. We found that 4 of 19 risk variants were associated with colorectal cancer, even after accounting for diabetes status and body mass index. Moreover, two missense variants in the *THADA* and *KCNJ11* loci (rs7578597 and rs5219, respectively) demonstrated the strongest associations with colorectal cancer. The associations of colorectal cancer with the *THADA* and *KCNJ11* variants were most evident among subjects, who did not report a diagnosis of diabetes, and the *THADA* variant was stronger for normal weight subjects, although it was observed among overweight subjects as well.

THADA (thyroid adenoma associated protein) maps to a chromosomal cluster at 2p21 and is a target gene of frequent chromosomal rearrangements in thyroid adenomas [36, 37]. While its biological effect is not well known, *THADA* is hypothesized to be involved in the death receptor pathway and apoptosis, as truncation of *THADA* may facilitate the proliferation and/or development of thyroid adenomas [37]. A common polymorphism in *THADA* (rs1465618), uncorrelated with rs7578597 (r^2=0.02 for European Americans), has also been associated with prostate cancer in a genome-wide association study [38]. The rs7578597 missense variant is located in exon 24 of the *THADA* gene and results in a threonine to alanine (T>C) amino acid change. In a meta-analysis of three genome-wide associations studies of T2D among 10,128 individuals of European descent, the T allele of rs7578597 (risk allele frequency = 0.902) was associated with a 15% increased risk of T2D (95% CI: 1.10-1.20; $P = 1.1 \times 10^{-9}$) [20]. In contrast, the T allele in our study was associated with a reduced risk of colorectal cancer ($P = 0.006$) in our multiethnic population (risk allele frequency range = 0.751-0.984). Interestingly, the T allele of rs7578597 has been associated with lower insulin levels during oral glucose tolerance testing in a Chinese population [39]. The difference in the direction of the association for T2D and colorectal cancer may relate to differences in the biological effects of thyroid adenoma associated protein, whose biological activity remains poorly characterized. Furthermore, the heterogeneity in effects across racial/ethnic groups (especially the lack of association in African Americans, a group with a shorter linkage disequilibrium) would suggest that rs7578597 may not be the causal variant, as similar associations and biological consequences would be expected for a functional SNP regardless of race/ethnicity. It is possible that rs7578597 is in linkage disequilibrium with another polymorphism that is more relevant for colorectal cancer. However, we do recognize that our study had <80% power to detect a similar association in African Americans as observed among the other racial/ethnic groups for rs7578597.

KCNJ11 encodes for the potassium inwardly-rectifying channel, superfamily J, member 11, a major subunit of the ATP-sensitive potassium channel complex, which plays a key role in glucose-stimulated insulin secretion [40, 41]. The rs5219 polymorphism (C>T), also a missense variant, results in a glutamate to lysine amino acid change in exon 1, and the T allele was identified in a genome-wide association study of T2D to be associated with a 14% increased risk of diabetes (95% CI: 1.10-1.19; $P = 6.7\times10^{-11}$) [15]. The T allele has also been associated with higher insulin sensitivity and early-phase insulin release [42]. The similar positive association we observed between the rs5219 T allele and colorectal cancer risk supports the possible role of hyperinsulinemia in promoting colorectal carcinogenesis [5].

The strong associations observed for *THADA* (rs7578597) and *KCNJ11* (rs5219) with colorectal cancer risk among non-diabetics in contrast to diabetics may reflect the reduced study power among the diabetic sub-group (colorectal cancer cases/controls=538/1,378). Alternatively, for *THADA* (rs7578597), the observation that the high diabetes risk allele (T) was associated with a reduced risk of colorectal cancer may suggest that this polymorphism has separate effects for diabetes and colorectal cancer. Clearly, additional studies are needed among well-characterized study populations with sufficient power to carefully discern the association of *THADA* with diabetes, and colorectal cancer.

Although *JAZF1* (rs864745) and *TSPAN8/LGR5* (rs7961581) displayed only weak associations with colorectal risk, these genes may still have important biological implications for colorectal cancer development. *JAZF1* encodes a nuclear protein with three C_2H_2-type zinc fingers, and functions as a transcriptional repressor [43, 44]. In addition, the *JAZF1* gene demonstrates evidence of pleiotropy, being associated with multiple traits in genome-wide association studies, including diabetes [20, 45], prostate cancer [46], and height [47, 48]. These pleiotropic effects, including possibly colorectal cancer, suggest that *JAZF1* may be a central node in important biological pathways linking these common phenotypes. *TSPAN8/LGR5* encodes for protein in the transmember 4 superfamily and mediates signal transduction events involved in cell development, growth, and motility [49]. Recently, a facilitating effect was described for *TSPAN8* on cell migration and adhesion in colon carcinoma, suggesting an important role for *TSPAN8* in colon cancer progression and metastasis [50].

The significant associations observed between rs7578597 (*THADA*), rs5219 (*KCNJ11*), rs864745 (*JAZF1*), rs7961581 (*TSPAN8*) and colorectal cancer that remained unchanged after adjustment for diabetes status indicate that diabetes does not confound the relationship between these T2D variants and colorectal cancer risk.

In a small nested case-control study of the Atherosclerosis Risk in Communities (ARIC) Study, the *TCF7L2* rs7903146 polymorphism was associated with a significant increased risk of colon cancer (P trend = 0.009). In contrast, in both a prior colon cancer case-control study [12] and our study, no overall association between rs7903146 and colon and colorectal cancer, respectively, was observed. This discrepancy may reflect a chance finding given the small number of cases in the ARIC study (128 colon cancer cases) [10], in contrast to our larger study of 2,011 colorectal cancer cases and the prior case-control study of 1,578 colon cancer cases [12], both demonstrating no association.

Diabetes may influence colorectal cancer development through multiple inter-related pathways, such as hyperglycemia, hyperinsulinemia and chronic inflammation. We note that diabetes status remained an influential predictor of colorectal cancer in models that included the SNPs considered here and BMI, indicating that additional factors need to be investigated in order to understand the effect of T2D on the risk of this cancer. Our findings of associations between T2D variants and colorectal cancer risk after adjustment for both diabetes status and BMI, however, does suggest that these variants have independent effects on the carcinogenic process. Although some of the observed effects may be through unrecognized hyperinsulinemia (*KCNJ11*), others (e.g., *THADA*) reflect the effect of the opposite allele and, thus, that of different mechanisms and, possibly, different functional alleles. Clearly, functional studies are needed to clarify the biological effects of these variants.

To our knowledge, our study is the first to test multiple T2D variants (19 SNPs) in a large multiethnic case-control study of colorectal cancer. The robust association observed for the *THADA* polymorphism and colorectal cancer with correction for multiple hypothesis testing through permutation testing suggests that this finding is unlikely to be due to chance. These results warrant additional studies in other populations to confirm these results. Furthermore,

the similar genetic associations observed for *THADA* and *KCNJ11* with adjustment for genetic ancestry argues against bias in our findings due to population stratification. Consistent replications in multiple well-powered studies will be essential to establish the associations between these T2D variants and colorectal cancer risk.

In summary, our study suggests that selected established T2D risk variants contribute to the risk of colorectal cancer. This finding builds upon prior epidemiologic studies demonstrating an association between diabetes and colorectal cancer, and provides new information on the complexity of the pathways shared between these two common diseases. Research into the biological mechanisms by which inherited T2D variants influence colorectal cancer risk will further our understanding of the key contributors to colorectal cancer development.

Acknowledgments

The Population Architecture Using Genomics and Epidemiology (PAGE) program is funded by the National Human Genome Research Institute (NHGRI), supported by U01HG004803 (CALiCo), U01HG004798 (EAGLE), U01HG004802 (MEC), U01HG004790 (WHI), and U01HG004801 (Coordinating Center). The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The complete list of PAGE members can be found at http://www.pagestudy.org.

The Multiethnic Cohort study (MEC) characterization of epidemiological architecture is funded through the NHGRI PAGE program (U01HG004802). The MEC study is funded through the National Cancer Institute (R37CA54281, R01 CA63, P01CA33619, U01CA136792, and U01CA98758).

Assistance with phenotype harmonization, SNP selection and annotation, data cleaning, data management, integration and dissemination, and general study coordination was provided by the PAGE Coordinating Center (U01HG004801-01). The National Institutes of Mental Health also contributes to the support for the Coordinating Center.

We thank the participants of the Multiethnic Cohort who have contributed to a better understanding of the lifestyle and genetic contributions to colorectal cancer. We also would like to thank Lauren Hu and Hansong Wang for their assistance with analysis and manuscript preparation.

References

1. American Cancer Society. Cancer Facts & Figures 2010. Atlanta: American Cancer Society; 2010.
2. Mainous AG 3rd, Baker R, Koopman RJ, Saxena S, Diaz VA, Everett CJ, et al. Impact of the population at risk of diabetes on projections of diabetes burden in the United States: an epidemic on the way. Diabetologia. 2007; 50:934–40. [PubMed: 17119914]
3. Larsson SC, Orsini N, Wolk A. Diabetes mellitus and risk of colorectal cancer: a meta-analysis. J Natl Cancer Inst. 2005; 97:1679–87. [PubMed: 16288121]
4. He J, Stram DO, Kolonel LN, Henderson BE, Le Marchand L, Haiman CA. The association of diabetes with colorectal cancer risk: the Multiethnic Cohort. Br J Cancer. 2010; 103:120–6. [PubMed: 20531412]
5. Giovannucci E, Harlan DM, Archer MC, Bergenstal RM, Gapstur SM, Habel LA, et al. Diabetes and cancer: a consensus report. CA Cancer J Clin. 2010; 60:207–21. [PubMed: 20554718]
6. Will JCGD, Vinicor F, Calle EE. Colorectal cancer: another complication of diabetes mellitus? Am J Epidemiol. 1998; 147:816–25. [PubMed: 9583711]
7. Nilsen TI, Vatten LJ. Prospective study of colorectal cancer risk and physical activity, diabetes, blood glucose and BMI: exploring the hyperinsulinaemia hypothesis. Br J Cancer. 2001; 84:417–22. [PubMed: 11161410]
8. Seow A, Yuan JM, Koh WP, Lee HP, Yu MC. Diabetes mellitus and risk of colorectal cancer in the Singapore Chinese Health Study. J Natl Cancer Inst. 2006; 98:135–8. [PubMed: 16418516]

9. Vinikoor LCLM, Keku TO, Martin CF, Galanko JA, Sandler RS. The association between diabetes, insulin use, and colorectal cancer among Whites and African Americans. Cancer Epidemiol Biomarkers Prev. 2009; 18:1239–42. [PubMed: 19336553]
10. Folsom AR, Pankow JS, Peacock JM, Bielinski SJ, Heiss G, Boerwinkle E. Variation in TCF7L2 and increased risk of colon cancer: the Atherosclerosis Risk in Communities (ARIC) Study. Diabetes Care. 2008; 31:905–9. [PubMed: 18268068]
11. Hazra A, Fuchs CS, Chan AT, Giovannucci EL, Hunter DJ. Association of the TCF7L2 polymorphism with colorectal cancer and adenoma risk. Cancer Causes Control. 2008; 19:975–80. [PubMed: 18478343]
12. Slattery ML, Folsom AR, Wolff R, Herrick J, Caan BJ, Potter JD. Transcription factor 7-like 2 polymorphism and colon cancer. Cancer Epidemiol Biomarkers Prev. 2008; 17:978–82. [PubMed: 18398040]
13. Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proenca C, et al. Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. Nat Genet. 2009; 41:1110–5. [PubMed: 19734900]
14. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science. 2007; 316:1331–6. [PubMed: 17463246]
15. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science. 2007; 316:1341–5. [PubMed: 17463248]
16. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature. 2007; 445:881–5. [PubMed: 17293876]
17. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet. 2007; 39:770–5. [PubMed: 17460697]
18. Takeuchi F, Serizawa M, Yamamoto K, Fujisawa T, Nakashima E, Ohnaka K, et al. Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. Diabetes. 2009; 58:1690–9. [PubMed: 19401414]
19. Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwehand WH, Strachan DP, et al. Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. Diabetes. 2009; 58:505–10. [PubMed: 19056611]
20. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet. 2008; 40:638–45. [PubMed: 18372903]
21. Pechlivanis S, Pardini B, Bermejo JL, Wagner K, Naccarati A, Vodickova L, et al. Insulin pathway related genes and risk of colorectal cancer: INSR promoter polymorphism shows a protective effect. Endocr Relat Cancer. 2007; 14:733–40. [PubMed: 17914103]
22. Landi D, Moreno V, Guino E, Vodicka P, Pardini B, Naccarati A. Polymorphisms affecting micro-RNA regulation and associated with the risk of dietary-related cancers: A review from the literature and new evidence for a functional role of rs17281995 (CD86) and rs1051690 (INSR), previously associated with colorectal cancer. Mutat Res.
23. Leroy EC, Moore JH, Hu C, Martinez ME, Lance P, Duggan D, et al. Genes in the insulin and insulin-like growth factor pathway and odds of metachronous colorectal neoplasia. Hum Genet.
24. Salonen JT, Uimari P, Aalto JM, Pirskanen M, Kaikkonen J, Todorova B, et al. Type 2 diabetes whole-genome association study in four populations: the DiaGen consortium. Am J Hum Genet. 2007; 81:338–45. [PubMed: 17668382]
25. Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, Andersen G, et al. SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. Nat Genet. 2008; 40:1098–102. [PubMed: 18711366]
26. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet. 2008; 40:1092–7. [PubMed: 18711367]

27. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 2007; 316:1336–41. [PubMed: 17463249]

28. Pare G, Chasman DI, Parker AN, Nathan DM, Miletich JP, Zee RY, et al. Novel association of HK1 with glycated hemoglobin in a non-diabetic population: a genome-wide evaluation of 14,618 participants in the Women's Genome Health Study. PLoS Genet. 2008; 4:e1000312. [PubMed: 19096518]

29. Sandhu MS, Weedon MN, Fawcett KA, Wasson J, Debenham SL, Daly A, et al. Common variants in WFS1 confer risk of type 2 diabetes. Nat Genet. 2007; 39:951–3. [PubMed: 17603484]

30. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol. 2000; 151:346–57. [PubMed: 10695593]

31. Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. Am J Epidemiol. 2002; 155:478–84. [PubMed: 11867360]

32. He J, Wilkens LR, Stram DO, Kolonel LN, Henderson BE, Wu AH, et al. Generalizability and epidemiologic characterization of eleven colorectal cancer GWAS hits in multiple populations. Cancer Epidemiol Biomarkers Prev. 20:70–81. [PubMed: 21071539]

33. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–9. [PubMed: 16862161]

34. Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat. 2009; 30:69–78. [PubMed: 18683858]

35. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 2005; 6:95–108. [PubMed: 15716906]

36. Drieschner N, Belge G, Rippe V, Meiboom M, Loeschke S, Bullerdiek J. Evidence for a 3p25 breakpoint hot spot region in thyroid tumors of follicular origin. Thyroid. 2006; 16:1091–6. [PubMed: 17123335]

37. Rippe V, Drieschner N, Meiboom M, Murua Escobar H, Bonk U, Belge G, et al. Identification of a gene rearranged by 2p21 aberrations in thyroid adenomas. Oncogene. 2003; 22:6111–4. [PubMed: 12955091]

38. Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet. 2009; 41:1116–21. [PubMed: 19767753]

39. Hu C, Zhang R, Wang C, Wang J, Ma X, Lu J, et al. PPARG, KCNJ11, CDKAL1, CDKN2A-CDKN2B, IDE-KIF11-HHEX, IGF2BP2 and SLC30A8 are associated with type 2 diabetes in a Chinese population. PLoS One. 2009; 4:e7643. [PubMed: 19862325]

40. Cook DL, Hales CN. Intracellular ATP directly blocks K+ channels in pancreatic B-cells. Nature. 1984; 311:271–3. [PubMed: 6090930]

41. Inagaki N, Gonoi T, Clement JPt, Namba N, Inazawa J, Gonzalez G, et al. Reconstitution of IKATP: an inward rectifier subunit plus the sulfonylurea receptor. Science. 1995; 270:1166–70. [PubMed: 7502040]

42. Stancakova A, Kuulasmaa T, Paananen J, Jackson AU, Bonnycastle LL, Collins FS, et al. Association of 18 confirmed susceptibility loci for type 2 diabetes with indices of insulin release, proinsulin conversion, and insulin sensitivity in 5,327 nondiabetic Finnish men. Diabetes. 2009; 58:2129–36. [PubMed: 19502414]

43. Koontz JI, Soreng AL, Nucci M, Kuo FC, Pauwels P, van Den Berghe H, et al. Frequent fusion of the JAZF1 and JJAZ1 genes in endometrial stromal tumors. Proc Natl Acad Sci U S A. 2001; 98:6348–53. [PubMed: 11371647]

44. Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, et al. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci U S A. 2002; 99:16899–903. [PubMed: 12477932]

45. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet. 2010; 42:579–89. [PubMed: 20581827]

46. Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet. 2008; 40:310–5. [PubMed: 18264096]

47. Johansson A, Marroni F, Hayward C, Franklin CS, Kirichenko AV, Jonasson I, et al. Common variants in the JAZF1 gene associated with height identified by linkage and genome-wide association analysis. Hum Mol Genet. 2009; 18:373–80. [PubMed: 18952825]

48. Soranzo N, Rivadeneira F, Chinappen-Horsley U, Malkina I, Richards JB, Hammond N, et al. Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. PLoS Genet. 2009; 5:e1000445. [PubMed: 19343178]

49. Szala S, Froehlich M, Scollon M, Kasai Y, Steplewski Z, Koprowski H, et al. Molecular cloning of cDNA for the carcinoma-associated antigen GA733-2. Proc Natl Acad Sci U S A. 1990; 87:3542–6. [PubMed: 2333300]

50. Greco C, Bralet MP, Ailane N, Dubart-Kupperschmitt A, Rubinstein E, Le Naour F, et al. E-cadherin/p120-catenin and tetraspanin Co-029 cooperate for cell motility control in human colon carcinoma. Cancer Res. 2010; 70:7674–83. [PubMed: 20858717]

51. Tsai FJ, Yang CF, Chen CC, Chuang LM, Lu CH, Chang CT, et al. A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. PLoS Genet. 2010; 6:e1000847. [PubMed: 20174558]

Summary Box

What is already known about this subject

- Epidemiologic studies have implicated type-2 diabetes to play a role in colorectal cancer.
- Previous genetic association studies of colorectal cancer have examined only a few T2D risk variants for their impact on colorectal cancer in mostly small study populations.

What are the new findings

- In examining 19 established T2D risk variants in a large case-control study of colorectal cancer nested in a multiethnic cohort study, our study identifies genetic variants for T2D that also impact the risk of colorectal cancer.
- Since similar patterns of associations were observed with adjustment for diabetes status and body mass index, the underlying effects of these type-2 diabetes risk variants may operate through separate pathways than for diabetes.

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

- If these variants are confirmed to be associated with colorectal cancer, they could be used to improve risk prediction models and, ultimately, personalize indications for colorectal cancer screening.

Table 1

Nineteen T2D established risk variants identified by genome-wide association studies

Chromosome Region	SNP (reference)	Locus	T2D Risk Allele	Alleles	Risk Allele Frequency				
					African American	Native Hawaiian	Japanese American	Latino	White
1p12	rs10923931 [20]	*NOTCH2*	T	G, T	0.326	0.054	0.023	0.094	0.108
2p21	rs7578597 [20]	*THADA*	T	C, T	0.751	0.949	0.984	0.928	0.896
3p25.2	rs1801282 [14, 15, 27]	*PPARG*	C	G, C	0.972	0.937	0.966	0.900	0.881
3p14.1	rs4607103 [20]	*ADAMTS9*	C	T, C	0.715	0.737	0.636	0.687	0.741
3q27.2	rs4402960 [14, 15, 18, 20, 27]	*IGF2BP2*	T	G, T	0.509	0.287	0.313	0.262	0.310
4p16	rs10010131 [29]	*WFS1*	G	A, G	0.653	0.810	0.980	0.706	0.610
6p22.3	rs7754840 [14, 15]	*CDKAL1*	C	G, C	0.539	0.533	0.434	0.322	0.313
7p15.1	rs864745 [20]	*JAZF1*	T	C, T	0.728	0.732	0.778	0.617	0.519
8q24.11	rs13266634 [13, 14, 15, 16, 17, 18, 19, 27]	*SLC30A8*	C	T, C	0.894	0.621	0.606	0.749	0.712
9p21.3	rs2383208 [18]	*CDKN2A-2B*	A	G, A	0.809	0.748	0.564	0.855	0.815
10p13	rs12779790 [20]	*CDC123,CAMK1D*	G	A, G	0.143	0.190	0.162	0.178	0.168
10q23.33	rs1111875 [14, 15, 16, 18]	*HHEX*	G	A, G	0.748	0.298	0.272	0.628	0.588
10q25.2	rs7903146 [13, 14, 15, 16, 17, 18, 19, 20]	*TCF7L2*	T	C, T	0.282	0.143	0.044	0.234	0.302
11p15.1	rs5219 [14, 15, 19]	*KCNJ11*	T	C, T	0.104	0.371	0.381	0.386	0.360
11p15.4	rs2237897 [25]	*KCNQ1*	C	T, C	0.919	0.781	0.614	0.787	0.948
11p15.5	rs2074196 [26]	*KCNQ1*	G	T, G	0.892	0.770	0.546	0.784	0.952
11p15.5	rs2237892 [18, 26]	*KCNQ1*	C	T, C	0.907	0.766	0.596	0.778	0.930
11p15.5	rs2237895 [51]	*KCNQ1*	C	A, C	0.792	0.644	0.659	0.569	0.582
12q21.1	rs7961581 [20]	*TSPAN8, LGR5*	C	T, C	0.224	0.289	0.203	0.213	0.286

Table 2
Study characteristics of 8,060 subjects by colorectal cancer case-control status

		Cases n=2011	Controls n=6049
Age, mean (SD)		70.0 (8.5)	66.9 (8.3)
Sex, n (%)			
	Male	1095 (54.5)	3272 (54.1)
	Female	916 (45.6)	2777 (45.9)
Race/Ethnicity, n (%)			
	African American	398 (19.8)	1495 (24.7)
	Native Hawaiian	108 (5.4)	489 (8.1)
	Japanese American	675 (33.6)	1717 (28.4)
	Latino	457 (22.7)	1195 (19.8)
	White	373 (18.5)	1153 (19.1)
BMI; n (%)*			
	> 25 kg/m^2	1266 (63.7)	3609 (60.2)
Diabetes, n (%)			
	Yes	538 (26.8)	1328 (22.0)
	Missing	113 (5.6)	354 (5.9)
Family history of colorectal cancer, n (%)			
	Yes	216 (12.6)	535 (10.1)
Dietary Intake*			
	Fiber; g/kcal/day, mean (SD)	11.67 (4.41)	11.81 (4.23)
	Calcium (foods + supplements); mg/day, mean (SD)	962.87 (588.78)	1011.11 (644.25)
	Dietary Folate equivalents (foods + supplements) ug/day, mean (SD)	908.90 (634.55)	930.67 (683.63)
	Alcohol; g/day, mean (SD)	11.44 (30.07)	9.33 (24.96)
Vigorous physical activity; hours/day, mean (SD)*		0.36 (0.80)	0.39 (0.80)
Smoking; pack-years, mean (SD)*		12.06 (16.2)	10.4 (14.7)
Anatomical sub-site, n (%)*			
	Colon	1409 (17.5)	-
	Rectum	462 (5.7)	-

* missing do not add to 8,060 due to missing information

Table 3

Association between 19 T2D established risk variants and colorectal cancer

			All*					
Locus	**SNP**	**Genotype**	**Cases, n (%)**	**Controls, n (%)**	**OR (95% CI)***	**p**	**OR (95% CI)****	**p**
NOTCH2	rs10923931	GG	1578 (80.0)	4652 (77.2)	1.00		1.00	
		GT	342 (17.3)	1178 (19.6)	0.96 (0.83 - 1.11)	0.55	0.94 (0.81 - 1.09)	0.44
		TT	53 (2.7)	197 (3.3)	0.99 (0.71 - 1.37)	0.95	1.01 (0.72 - 1.40)	0.96
		trend			0.97 (0.86 -1.09)	0.64	0.97 (0.86 - 1.09)	0.60
THADA	rs7578597	CC	63 (3.3)	124 (2.1)	1.00		1.00	
		TC	297 (15.4)	1004 (16.7)	**0.55 (0.39 - 0.77)**	**0.0004**	**0.55 (0.39 - 0.77)**	**0.0005**
		TT	1574 (81.4)	4878 (81.2)	**0.52 (0.38 - 0.72)**	**6.92×10^{-5}**	**0.52 (0.38 - 0.73)**	**0.0001**
		trend			**0.84 (0.74 - 0.95)**	**0.0041**	**0.84 (0.75 - 0.95)**	**0.0063**
PPARG	rs1801282	GG	8 (0.4)	32 (0.5)	1.00		1.00	
		CG	237 (11.9)	705 (11.7)	1.34 (0.60 - 3.00)	0.47	1.37 (0.61 - 3.08)	0.44
		CC	1738 (87.6)	5274 (87.7)	1.33 (0.60 - 2.94)	0.48	1.37 (0.62 - 3.04)	0.44
		trend			1.01 (0.87 - 1.18)	0.88	1.02 (0.87 - 1.19)	0.80
ADAMTS9	rs4607103	TT	207 (10.5)	575 (9.6)	1.00		1.00	
		CT	850 (43.2)	2536 (42.1)	0.95 (0.80 - 1.14)	0.61	0.96 (0.80 - 1.15)	0.66
		CC	911 (46.3)	2907 (48.3)	0.92 (0.77 - 1.11)	0.39	0.95 (0.79 - 1.13)	0.56
		trend			0.96 (0.89 - 1.04)	0.37	0.98 (0.90 - 1.06)	0.59
IGF2BP2	rs4402960	GG	850 (43.8)	2621 (43.7)	1.00		1.00	
		GT	861 (44.4)	2581 (43.0)	1.08 (0.97 - 1.21)	0.16	1.08 (0.97 - 1.21)	0.16
		TT	228 (11.8)	799 (13.3)	1.01 (0.85 - 1.20)	0.89	1.04 (0.87 - 1.24)	0.66
		trend			1.03 (0.95 - 1.11)	0.48	1.04 (0.96 - 1.13)	0.35
WFS1	rs10010131	AA	130 (6.8)	480 (8.1)	1.00		1.00	
		GA	566 (29.5)	1882 (31.7)	1.10 (0.88 - 1.37)	0.39	1.11 (0.89 - 1.38)	0.37
		GG	1221 (63.7)	3571 (60.2)	1.14 (0.91 - 1.41)	0.25	1.16 (0.93 - 1.44)	0.20
		trend			1.05 (0.96 - 1.16)	0.28	1.06 (0.97 - 1.17)	0.20

			All*					
Locus	**SNP**	**Genotype**	**Cases, n (%)**	**Controls, n (%)**	**OR (95% CI)***	**p**	**OR (95% CI)****	**p**
CDKAL1	rs7754840	GG	703 (35.8)	2048 (34.2)	1.00		1.00	
		CG	904 (46.0)	2817 (47.1)	0.96 (0.85 - 1.08)	0.49	0.97 (0.86 - 1.09)	0.57
		CC	357 (18.2)	1118 (18.7)	1.01 (0.87 - 1.18)	0.90	1.02 (0.87 - 1.19)	0.85
		trend			1.00 (0.93 - 1.08)	0.95	1.00 (0.93 - 1.08)	0.98
JAZF1	rs864745	CC	246 (12.5)	651 (10.8)	1.00		1.00	
		TC	769 (39.2)	2541 (42.3)	**0.82 (0.69 - 0.97)**	**0.024**	**0.83 (0.70 - 0.99)**	**0.038**
		TT	947 (48.2)	2821 (46.9)	0.92 (0.77 - 1.09)	0.32	0.93 (0.78 - 1.10)	0.39
		trend			1.00 (0.93 - 1.09)	0.93	1.01 (0.93 - 1.09)	0.85
SLC30A8	rs13266634	TT	213 (10.9)	544 (9.1)	1.00		1.00	
		CT	707 (36.2)	2183 (36.5)	0.87 (0.72 - 1.04)	0.12	0.87 (0.72 - 1.04)	0.13
		CC	1031 (52.8)	3256 (54.4)	0.90 (0.75 - 1.08)	0.25	0.90 (0.75 - 1.09)	0.29
		trend			0.98 (0.90 - 1.06)	0.58	0.98 (0.90 - 1.07)	0.67
CDKN2A-2B	rs2383208	GG	173 (8.9)	476 (8.0)	1.00		1.00	
		AG	669 (34.6)	2085 (35.0)	0.94 (0.77 - 1.15)	0.58	0.95 (0.77 - 1.16)	0.59
		AA	1092 (56.5)	3396 (57.0)	0.99 (0.81 - 1.21)	0.92	0.99 (0.81 - 1.21)	0.89
		trend			1.02 (0.93 - 1.11)	0.74	1.01 (0.93 - 1.10)	0.79
CDC123,CAMK1D	rs12779790	AA	1314 (68.0)	4192 (70.0)	1.00		1.00	
		AG	558 (28.9)	1628 (27.2)	1.09 (0.97 - 1.23)	0.13	1.10 (0.98 - 1.23)	0.12
		GG	61 (3.2)	167 (2.8)	1.18 (0.87 - 1.61)	0.28	1.17 (0.86 - 1.60)	0.32
		trend			1.09 (0.99 - 1.20)	0.08	1.09 (0.99 - 1.20)	0.08
HHEX	rs1111875	AA	533 (28.2)	1565 (26.2)	1.00		1.00	
		GA	757 (40.1)	2594 (43.5)	0.92 (0.80 - 1.06)	0.23	0.91 (0.79 - 1.05)	0.19
		GG	600 (31.8)	1810 (30.3)	1.10 (0.94 - 1.29)	0.25	1.10 (0.93 - 1.29)	0.27
		trend			1.05 (0.97 - 1.14)	0.20	1.05 (0.97 - 1.14)	0.22
TCF7L2	rs7903146	CC	1312 (67.7)	3937 (65.5)	1.00		1.00	
		CT	529 (27.3)	1771 (29.5)	0.98 (0.86 - 1.11)	0.74	0.98 (0.87 - 1.12)	0.81
		TT	98 (5.1)	301 (5.0)	1.06 (0.82 - 1.35)	0.67	1.06 (0.82 - 1.36)	0.67

Locus	SNP	Genotype	All*					
			Cases, n (%)	Controls, n (%)	OR (95% CI)*	p	OR (95% CI)**	p
		trend			1.00 (0.91 - 1.11)	0.95	1.01 (0.91 - 1.11)	0.91
KCNJ11	rs5219	CC	841 (43.6)	2922 (48.7)	1.00		1.00	
		CT	872 (45.2)	2444 (40.8)	**1.18 (1.05 - 1.32)**	**0.006**	**1.19 (1.06 - 1.34)**	**0.0033**
		TT	215 (11.2)	629 (10.5)	1.13 (0.94 - 1.35)	0.19	1.13 (0.94 - 1.35)	0.21
		trend			**1.10 (1.01 - 1.19)**	**0.027**	**1.10 (1.01 - 1.19)**	**0.025**
KCNQ1	rs2237897	TT	127 (6.6)	368 (6.1)	1.00		1.00	
		CT	580 (30.3)	1658 (27.6)	1.11 (0.89 -1.40)	0.35	1.11 (0.88 - 1.39)	0.39
		CC	1208 (63.1)	3984 (66.3)	1.10 (0.88 - 1.38)	0.39	1.08 (0.86 - 1.36)	0.50
		trend			1.02 (0.93 - 1.13)	0.61	1.01 (0.92 - 1.12)	0.78
KCNQ1	rs2074196	TT	179 (9.3)	482 (8.0)	1.00		1.00	
		GT	620 (32.0)	1759 (29.2)	1.03 (0.84 - 1.26)	0.77	1.02 (0.84 - 1.25)	0.81
		GG	1137 (58.7)	3774 (62.7)	0.99 (0.81- 1.22)	0.95	0.98 (0.80 - 1.20)	0.86
		trend			0.99 (0.90 - 1.08)	0.75	0.98 (0.89 - 1.07)	0.66
KCNQ1	rs2237892	TT	143 (7.4)	398 (6.7)	1.00		1.00	
		CT	620 (31.9)	1754 (29.3)	1.07 (0.86 -1.33)	0.56	1.06 (0.85 - 1.32)	0.61
		CC	1183 (60.8)	3828 (64.0)	1.03 (0.83 - 1.28)	0.78	1.02 (0.82 - 1.27)	0.85
		trend			1.00 (0.91 - 1.09)	0.92	0.99 (0.90 - 1.09)	0.86
KCNQ1	rs2237895	AA	884 (45.1)	2648 (44.2)	1.00		1.00	
		AC	818 (41.7)	2594 (43.3)	0.90 (0.81-1.01)	0.08	0.90 (0.80 - 1.01)	0.07
		CC	259 (13.2)	752 (12.6)	0.96 (0.81 - 1.14)	0.65	0.94 (0.79 - 1.11)	0.48
		trend			0.96 (0.89 - 1.04)	0.29	0.95 (0.88 - 1.03)	0.20
TSPAN8, LGR5	rs7961581	TT	1200 (62.0)	3521 (59.0)	1.00		1.00	
		CT	634 (32.7)	2111 (35.4)	**0.88 (0.79 - 0.99)**	**0.032**	**0.88 (0.78 - 0.98)**	**0.022**
		CC	103 (5.3)	338 (5.7)	0.92 (0.73 - 1.17)	0.50	0.92 (0.72 - 1.16)	0.47
		trend			0.92 (0.84 - 1.00)	0.06	**0.91 (0.83 - 1.00)**	**0.046**

*adjusted for age, sex, and racial/ethnic group

** adjusted for age, sex, racial/ethnic group, BMI, and diabetes

Table 4

Association between *THADA, JAZF1, KCNJ11,* and *TSPAN8* variants and colorectal cancer by racial/ethnic group*

			African Americans		Native Hawaiians		Japanese Americans		Latinos		Whites		
			ca/co=398/1495		ca/co=108/489		ca/co=675/1717		ca/co=457/1195		ca/co=373/1153		P het
Locus	**SNP**	**Genotype**	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**	**race**
THADA	rs7578597	CC	1.00		1.00		1.00		1.00		1.00		
		TC	0.97 (0.60 - 1.57)	0.89	--	--	**0.09 (0.02 - 0.42)**	**0.0024**	**0.27 (0.09 - 0.81)**	**0.02**	**0.26 (0.12 - 0.57)**	**0.0006**	
		TT	0.95 (0.59 - 1.52)	0.83	--	--	**0.07 (0.02 - 0.31)**	**0.0005**	**0.25 (0.09 - 0.72)**	**0.01**	**0.26 (0.13 - 0.54)**	**0.0002**	
		trend	0.98 (0.81 - 1.18)	0.82	1.49 (0.67 - 3.30)	0.33	**0.52 (0.36 - 0.75)**	**0.0005**	0.79 (0.59 - 1.04)	0.09	**0.75 (0.59 - 0.96)**	**0.02**	**0.008**
JAZF1	rs864745	CC	1.00		1.00		1.00		1.00		1.00		
		CT	0.99 (0.65 - 1.52)	0.97	0.68 (0.28 - 1.63)	0.39	0.77 (0.51 - 1.17)	0.23	0.87 (0.63 - 1.2)	0.40	**0.73 (0.54 - 0.98)**	**0.04**	
		TT	1.01 (0.67 - 1.53)	0.97	0.88 (0.38 - 2.04)	0.77	0.82 (0.55 - 1.23)	0.35	0.94 (0.67 - 1.32)	0.72	1.03 (0.74 - 1.44)	0.86	
		trend	1.01 (0.85 - 1.20)	0.91	1.09 (0.76 - 1.56)	0.64	0.99 (0.85 - 1.15)	0.88	0.99 (0.84 - 1.17)	0.93	1.03 (0.86 - 1.22)	0.77	0.94
KCNJ11	rs5219	CC	1.00		1.00		1.00		1.00		1.00		
		CT	1.16 (0.87 - 1.53)	0.31	**2.03 (1.22 - 3.37)**	**0.01**	1.15 (0.94 - 1.40)	0.17	1.10 (0.86 - 1.41)	0.45	1.11 (0.86 - 1.45)	0.42	
		TT	1.40 (0.41 - 4.73)	0.59	1.47 (0.71 - 3.08)	0.30	0.77 (0.57 - 1.05)	0.10	**1.50 (1.08 - 2.08)**	**0.02**	1.19 (0.81 - 1.77)	0.38	
		trend	1.16 (0.90 - 1.50)	0.26	1.35 (0.97 - 1.86)	0.07	0.95 (0.83 - 1.09)	0.49	**1.20 (1.02 - 1.41)**	**0.02**	1.10 (0.92 - 1.32)	0.31	0.19
TSPAN8	rs7961581	TT	1.00		1.00		1.00		1.00		1.00		
		CT	1.10 (0.86 - 1.41)	0.45	0.92 (0.58 - 1.44)	0.70	0.85 (0.69 - 1.04)	0.10	0.83 (0.65 - 1.06)	0.14	0.80 (0.62 - 1.04)	0.09	
		CC	0.80 (0.46 - 1.41)	0.44	0.62 (0.26 - 1.48)	0.28	1.37 (0.89 - 2.10)	0.16	0.91 (0.53 - 1.53)	0.71	0.73 (0.45 - 1.20)	0.21	
		trend	1.00 (0.83 - 1.22)	0.97	0.84 (0.60 - 1.19)	0.33	0.97 (0.83 - 1.14)	0.75	0.88 (0.73 - 1.07)	0.21	0.83 (0.68 - 1.01)	0.06	0.70

* adjusted for age and sex

Table 5

Association between *THADA*, *JAZF1*, *KCNJ11*, and *TSPAN8* variants and colorectal cancer by diabetes status and body mass index*

			non-Diabetic ca/co=1360/4367		Diabetic ca/co=538/1378		P het	BMI < 25 ca/co=722/2388		BMI > 25 ca/co=1266/3609		P het
Locus	**SNP**	**Genotype**	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**	**diabetes**	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**	**BMI**
THADA	rs7578597	CC	1.00		1.00			1.00		1.00		
		TC	**0.48 (0.32 - 0.73)**	**0.0005**	0.79 (0.38 - 1.63)	0.52		**0.38 (0.20 - 0.71)**	**0.0023**	**0.60 (0.40 - 0.91)**	**0.01**	
		TT	**0.43 (0.29 - 0.64)**	**2.81×10^{-5}**	0.89 (0.44 - 1.79)	0.75		**0.31 (0.17 - 0.57)**	**0.0001**	**0.62 (0.42 - 0.92)**	**0.02**	
		trend	**0.78 (0.68 - 0.90)**	**0.0009**	1.05 (0.81 - 1.35)	0.73	0.13	**0.70 (0.56 - 0.87)**	**0.0012**	**0.91 (0.79 - 1.06)**	**0.22**	0.06
JAZF1	rs864745	GG	1.00		1.00			1.00		1.00		
		AG	0.84 (0.68 - 1.03)	0.09	0.74 (0.52 - 1.06)	0.10		0.79 (0.59 - 1.06)	0.12	0.84 (0.68 - 1.04)	0.11	
		AA	0.91 (0.74 - 1.12)	0.39	0.84 (0.59 - 1.2)	0.34		0.96 (0.71 - 1.28)	0.76	0.89 (0.71 - 1.10)	0.27	
		trend	0.99 (0.90 - 1.09)	0.89	0.99 (0.84 - 1.16)	0.88	0.90	1.05 (0.91 - 1.20)	0.51	0.97 (0.88 - 1.08)	0.62	0.68
KCNJ11	rs5219	CC	1.00		1.00			1.00		1.00		
		CT	**1.16 (1.01 - 1.34)**	**0.04**	1.08 (0.86 - 1.37)	0.50		**1.36 (1.12 - 1.65)**	**0.0016**	1.10 (0.95 - 1.27)	0.22	
		TT	1.20 (0.96 - 1.49)	0.11	0.82 (0.58 - 1.18)	0.28		1.05 (0.77 - 1.43)	0.75	1.16 (0.92 - 1.46)	0.21	
		trend	**1.11 (1.01 - 1.23)**	**0.03**	0.95 (0.81 - 1.12)	0.58	0.44	1.12 (0.98 - 1.28)	0.10	1.08 (0.97 - 1.20)	0.14	0.13
TSPAN8	rs7961581	TT	1.00		1.00			1.00		1.00		
		CT	**0.84 (0.74 - 0.97)**	**0.01**	1.06 (0.85 - 1.32)	0.63		0.88 (0.73 - 1.06)	0.18	0.87 (0.76 - 1.01)	0.07	
		CC	0.98 (0.75 - 1.30)	0.91	0.75 (0.45 - 1.25)	0.27		1.00 (0.67 - 1.49)	0.98	0.88 (0.65 - 1.18)	0.38	
		trend	0.91 (0.82 - 1.01)	0.09	0.97 (0.81 - 1.16)	0.73	0.31	0.93 (0.80 - 1.08)	0.36	0.90 (0.81- 1.01)	0.08	0.88

* adjusted for age, sex, and racial/ethnic group