A Genetic Locus within the FMN1/GREM1 Gene Region Interacts with Body Mass Index in Colorectal Cancer Risk


ABSTRACT

Colorectal cancer risk can be impacted by genetic, environmental, and lifestyle factors, including diet and obesity. Gene-environment interactions (GxE) can provide biological insights into the effects of obesity on colorectal cancer risk. Here, we assessed potential genome-wide G x E interactions between body mass index (BMI) and common SNPs for colorectal cancer risk using data from 36,415 colorectal cancer cases and 48,451 controls from three international colorectal cancer consortia (CCFR, CORECT, and GECCO). The G x E tests included the conventional logistic regression using multiplicative terms (one degree of freedom, 1DF test), the two-step EDGE method, and the joint 3DF test, each of which is powerful for detecting G x E interactions under specific conditions. BMI was associated with higher colorectal cancer risk. The two-step approach revealed a statistically significant G x E interaction located within the Formin 1/Gremlin 1 (FMN1/GREM1) gene region (rs58349661). This SNP was also identified by the 3DF test, with a suggestive statistical significance in the 1DF test. Among participants with the CC genotype of rs58349661, overweight and obesity categories were associated with higher colorectal cancer risk, whereas null associations were observed across BMI categories in those with the TT genotype. Using data from three large international consortia, this study discovered a locus in the FMN1/GREM1 gene region that interacts with BMI on the association with colorectal cancer risk. Further studies should examine the potential mechanisms through which this locus modifies the etiologic link between obesity and colorectal cancer.

Significance: This gene-environment interaction analysis revealed a genetic locus in FMN1/GREM1 that interacts with body mass index in colorectal cancer risk, suggesting potential implications for precision prevention strategies.
Introduction

Colorectal cancer is a multifactorial disease that results from many genetic and behavioral/lifestyle risk factors, including diet and obesity (1, 2). General obesity, usually defined in adult populations as a body mass index (BMI) equal to or above 30 kg/m², has been consistently associated with higher risk of colorectal cancer and is estimated to account for 5% to 14% of all colorectal cancer diagnosed cases (3, 4). Mechanistically, obesity affects colorectal cancer carcinogenesis through different pathways, including adipose tissue-associated inflammation, oxidative stress, impairment of lipid metabolism, alterations of the microbiome, and through its causal association with comorbidities such as type 2 diabetes (5). Obesity is a worldwide health issue, mainly attributed to the ubiquitously expanded obesogenic environment (6), with a worldwide prevalence that ranged from less than 5% in the 1970s to over 13% in 2016 (i.e., more than 650 million adults worldwide; ref. 7).

Understanding how inherited germline genetic variation can influence colorectal cancer risk according to obesity status is important for developing better disease risk assessment tools. In recent years, genome-wide association studies (GWAS) have identified over 200 genetic risk variants associated with colorectal cancer development (8–12). Overall, the genetic heritability of colorectal cancer is estimated to be up to 20% (13). Given the complexity of colorectal carcinogenesis, gene-environment interactions (G × E) may be particularly well suited for identifying novel susceptibility loci and biologically plausible interactions that could further elucidate...
Materials and Methods

Study participants

Our study sample consisted of individual level genomic and epidemiological data from 34 studies (50% prospective cohort studies) included in the Colorectal Cancer Family Registry (CCFR), Colorectal Cancer Transdisciplinary Study (CORECT), and Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO). Control participants were matched by age, sex, genetically defined ancestry, and enrollment date/trial group, when applicable. Cases were de-identified by sigmoidoscopy or colonoscopy. Each study was approved by the relevant ethics committees or review boards pertaining to the institutions. All participants provided written informed consent during recruitment and all studies were conducted in accordance with ethical guidelines relevant to their geographic location and calendar year of enrollment (e.g., Declaration of Helsinki, CIOMS, Belmont Report, U.S. Common Rule).

Data harmonization and BMI assessment

Data harmonization consisted of a multistep procedure performed at the GECCO consortium coordinating center at the Fred Hutchinson Cancer Center as previously described (19). In brief, common data elements (CDE) were defined a priori for data harmonization. These CDEs represent common variables such as age and sex, or similar variables such as smoking or dietary intake. CDEs ensure that each variable defined is similar and comparable across different studies, hence allowing statistical analysis across a combined dataset. Study questionnaires and data dictionaries were examined, and elements were mapped to these CDEs through an iterative process of communication with data contributors. Definitions, permissible values, and standardized coding were implemented in a single database via SAS (RRID:SCR_008567) and T-SQL. The resulting data were checked for errors and outliers within and between the studies.

Demographic, anthropometric, and lifestyle variables such as sex, age, smoking, and self-reported or measured weight and height were collected via in-person interviews or structured self-administered questionnaires. In the cohort studies, standing height and body weight were ascertained at the time of blood collection or buccal swabs. In case-control studies, standing height was recalled at enrollment for cases and controls, except in REACH, where height was recalled a year prior to the time of interview in cases and controls, and in DACHS, where it was recalled at the time of enrollment in controls and diagnosis in cases. Body weight was recalled 1 to 2 years prior to enrollment of controls or diagnosis of cases in most case-control studies to avoid bias from illness-associated weight loss, except for DACHS and DALS where prediagnostic (in cases) and pre-enrollment (in controls) recall times were 5 to 14 years and 2 to 5 years, respectively. In CRCGEN, weight was recalled at the age of 45 years in cases and controls or 10 years before diagnosis in cases. BMI was calculated as the weight (kg) of each participant divided by the square of the height (m²). We scaled the BMI to reflect a 5 kg/m² increment as our main analytical variable. In addition, for categorical analyses, we used the World Health Organization (WHO) predefined BMI cut-off points for normal weight (18.5 < <25 kg/m²), overweight (≥25.0 to <30 kg/m²), and obesity (≥30 kg/m²). We excluded a small number of participants with a BMI below 18.5 kg/m² (n = 677) because of reported nonlinear associations between BMI and colorectal cancer risk at this end of the BMI continuum (20).

Genotyping and imputation

DNA was extracted from blood or buccal samples. Genotype data were generated from germline DNA on the Affymetrix Axiom (CORECT), UK Biobank Axiom (UK Biobank), Illumina 1M, 1M-Duo, or Omni1 (CCFR), Illumina 300 K, Illumina OmniExpress, Illumina 550K/610 K, or Affymetrix 100K/500 K (GECCO), and Illumina Omni 2.5 in the Molecular Epidemiology of Colorectal Cancer Study (MECC) as detailed elsewhere (8, 10) and summarily presented in Supplementary Table S1. High-density genotype array data were cleaned by applying standard quality control filters at both individual subject and SNP levels. Genotype data for case and control participants were phased and imputed together by an array platform, thus avoiding any potentially differential imputation error between case and control participants.

Participants were excluded based on genotyping call rates (<97%), heterozygosity, duplicates or close kinship, and inconsistencies between self-reported and genotypic sexes. We limited the analyses to individuals of European ancestry (3,586 participants of non-European descent were excluded) as determined by genetically defined ancestry and principal components (PC) clustering results with 1000 Genomes EUR populations. In terms of markers, we excluded SNPs based on missing call rates (>2 to 5%), departure from Hardy-Weinberg Equilibrium (P < 1 x 10^-6), and discordant genotype calls within duplicate samples. Genotypes were imputed to the Haplotype Reference Consortium (version r1.1) using the University of Michigan Imputation Server (21). To facilitate data management and analysis, genotypes were converted into a binary format using the BinaryDosage R package (https://cran.r-project.org/web/packages/BinaryDosage). We filtered imputed SNPs based on imputation accuracy of R^2 > 0.8 and minor allele frequency (MAF) > 1%. Over 7.2 million SNPS were retained after imputation and quality control, among which ~1 million SNPs were evaluated in our analysis because of the correlation between SNPs. PC analysis for population stratification assessment was performed using PLINK1.9 (RRID:SCR_001757) on 30,000 randomly selected imputed SNPs with MAF and R^2 over 5% and 0.99, respectively. colorectal cancer molecular subtypes, that is, BRAF and KRAS mutation status, CpG island methylator phenotype (CIMP), and microsatellite instability (MSI), were analyzed in the subset with available tumor samples (6,565 cases; Supplementary Table S2). BRAF, CIMP, KRAS, and MSI were determined using specific markers assessed using PCR, sequencing, or IHC.

Statistical analyses

Logistic regression models were run to examine associations between BMI and colorectal cancer risk in each study, and the results for all the studies were summarized using random-effects meta-analysis methods (Hartung-Knapp) to obtain summary ORs and
95% confidence intervals (CI). We calculated the heterogeneity $P$ values using Cochran’s Q statistics, while funnel plots identified studies with potential outlying ORs. We fitted the models overall and further stratified them by study design, sex, and tumor site in the colon and rectum. All meta-analyses of BMI main effects were performed using the R package Meta (RRID:SCR_019055; ref. 22).

We performed genome-wide interaction scans using the R package GxEScanR (23), which implements several interaction testing methods. Our analyses included conventional logistic regression with multiplicative interaction terms (1 degree of freedom test, 1DF), two-step EDGE method (17), and 3DF analyses (24), as detailed in the Supplementary methods. Compared with the 1DF, the joint 3DF test has higher power to detect G×E interactions when they exist, while accommodating the main effect of each variant on risk of colorectal cancer (D|G) and the effect of the environment on colorectal cancer risk (E|G) associations (24). The two-step method reduces the burden of multiple testing by preserving the statistical power, mainly through the initial filtering step that incorporates the D|G and E|G associations.

We presented the 1DF analysis and the two-step method, followed by the 3DF test. The rationale for presenting the results in this order is that both the 1DF and two-step method test directly for G×E interactions, the first as a classic test and the latter as a generally more powerful test. The 3DF analysis is presented last because it tests interaction indirectly and is particularly powerful in the presence of association between G and D or between G and E induced by interaction.

To identify novel loci, we filtered out previously known GWAS hits and SNPs identified in linkage disequilibrium (LD) with them in the 1000 Genomes EUR dataset. All models were adjusted for study, sex, age at diagnosis or enrollment, and the first three PCs of genetic ancestry to account for population substructure. For SNPs that reached statistical significance ($P = 5.0 \times 10^{-3}$), we accounted for three main tests or showed a trend towards that threshold ($P < 5 \times 10^{-5}$), we performed stratified G×E analyses by sex, tumor site (proximal colon, distal colon, rectum), and study design. We also added smoking status (never/ever), an established risk factor for colorectal cancer, to the multivariate models. We did not adjust for other lifestyle or dietary variables associated with colorectal cancer, such as alcohol consumption, physical inactivity, intake of red and processed meats, or calcium, because they do not interact with the SNPs and their inclusion in previous genetic analyses in our consortium (25) did not substantially change the associations with colorectal cancer.

We calculated ORs stratified by BMI categories (normal, overweight, obese) and genotype to examine the patterns of stratum-specific associations. These models were fitted using imputation posterior probabilities and visualized with a plot of predicted log-odds versus genotype by BMI category. As the positive association between obesity and colorectal cancer risk was consistently stronger in males than in females, we also explored the G×BMI interactions separately by sex. We used logistic regression (cases/controls; mutated/nonmutated) and nominal polytomous logistic regression (mutated/nonmutated/controls) to explore the interactions between the significant SNPs and BMI on colorectal tumor molecular subtypes, that is, BRAF and KRAS mutation status, CIMP, and MSI.

Functional annotations were examined for the significant findings, as described in the Supplementary methods. In brief, the magnitude of the association, the extent of the association explained by LD, as well as chromosomal position and neighboring SNPs and genes were investigated. Regional plots were generated using the command line version (Standalone) of LocusZoom v1.3 (26). The putative functional role of the SNPs with significant interactions and those in LD ($r^2 > 0.2$) at 500 kb flanking regions were explored relative to their potential contribution to regulate gene expression by their physical location in regions of chromatin accessibility or histone modifications (variant enhancer loci). Genes where expression in colon tissue samples were regulated by functional SNPs ($P$ values below $5 \times 10^{-8}$) were identified using the colon transverse tissue samples from GTEx v8 dataset, and the colon transcriptome coexpresser (CoTEx) (https://barcuvaseq.org/ cotrex, accessed January 2023) of the University of Barcelona and University of Virginia genotyping and RNA sequencing (BarcUVASeq) project dataset (27). This dataset is comprised of 445 epithelium-enriched healthy colon biopsies from ascending, transverse, and descending colon. We retrieved for each significant SNP, the genes with available expression and located within one million base pairs. The gene expressions were fit in a gene only model, and subsequently in a model including the interaction between the continuous gene expression and BMI, and a model with the gene expression in categories.

All statistical analyses were performed using R version 3.6.0. (RRID:SCR_001905, Foundation for Statistical Computing).

**Data and code availability**

The datasets and code supporting the current study have not been deposited in a public repository because they are part of an international consortium but are available from the corresponding author upon request.

**Results**

A total of 36,415 colorectal cancer cases (17,139 females) and 48,451 control (23,717 females) participants were included in this analysis. Mean age (SD) of the participants was 63.8 (10.4) and 62.4 (9.7) years in case and control participants, respectively (Table 1).

In a meta-analysis of all the participating studies, each 5 kg/m² increase in BMI was associated with 17% higher risk of colorectal cancer (OR = 1.17; 95% CI, 1.12–1.21; Fig. 1). The positive association was consistent across study designs, sex, and tumor anatomical subtypes, although the estimated BMI effect was higher in males (OR = 1.22; 95% CI, 1.16–1.28) compared with females (OR = 1.13; 95% CI, 1.08–1.18). In the distal colon (OR = 1.23; 95% CI, 1.17–1.29) compared with the proximal colon (OR = 1.15; 95% CI, 1.10–1.20), and in the rectum (OR = 1.16; 95% CI, 1.05–1.26). Table 1 summarizes the most prominent SNP×BMI interactions. The quantile-quantile (Q–Q) plot for G×BMI interaction for the 1DF analysis did not reveal residual stratification in the study population (Supplementary Fig. S1). In the 1DF exploration using multiplicative interaction terms in logistic regression models, we did not observe any genome-wide statistically significant loci interacting with BMI on colorectal cancer risk (Supplementary Fig. S2). Nine SNPs with $P$ values below $5 \times 10^{-8}$ in the 1DF are presented in Supplementary Table S3.

In the two-step EDGE approach, we observed a statistically significant interaction between variants located in the form1/GREM1 (FMN1/GREM1) gene region and BMI in colorectal cancer risk (Supplementary Fig. S3). The lead SNP with the lowest $P$ value was rs58349661 ($\text{P}_{\text{3DF}} = 3.68 \times 10^{-10}$, Supplementary Table S3). Consistent with these findings, we observed suggestive findings for SNPs in this locus in the 1DF test (rs58349661 $\text{P}_{\text{1DF}} = 4.97 \times 10^{-7}$) as well as strong
Table 1. Selected characteristics of the participants.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cases (n = 36,415)</th>
<th>Controls (n = 48,451)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females, %</td>
<td>47.1</td>
<td>49.0</td>
</tr>
<tr>
<td>Age, yrs; mean ± SD</td>
<td>63.8 ± 10.4</td>
<td>62.4 ± 9.7</td>
</tr>
<tr>
<td>Anthropometry, mean ± SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height, cm</td>
<td>169 ± 9.6</td>
<td>169 ± 9.6</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>27.5 ± 4.8</td>
<td>27.1 ± 4.6</td>
</tr>
<tr>
<td>Socio-demographic and lifestyle, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education (highest completed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>26.9</td>
<td>21.7</td>
</tr>
<tr>
<td>High school/GED</td>
<td>18.8</td>
<td>14.7</td>
</tr>
<tr>
<td>Some college</td>
<td>22.5</td>
<td>24.7</td>
</tr>
<tr>
<td>College/graduate school</td>
<td>27.3</td>
<td>32.2</td>
</tr>
<tr>
<td>Smoking, ever</td>
<td>44.5</td>
<td>49.0</td>
</tr>
<tr>
<td>Yes</td>
<td>55.1</td>
<td>51.0</td>
</tr>
<tr>
<td>Medical information, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 2 diabetes (ever diagnosed)</td>
<td>81.7</td>
<td>85.0</td>
</tr>
<tr>
<td>Yes</td>
<td>11.8</td>
<td>8.4</td>
</tr>
<tr>
<td>Any postmenopausal HRT use</td>
<td>20.3</td>
<td>21.8</td>
</tr>
<tr>
<td>No</td>
<td>11.1</td>
<td>13.3</td>
</tr>
<tr>
<td>Regular aspirin or NSAID use</td>
<td>53.9</td>
<td>51.2</td>
</tr>
<tr>
<td>No</td>
<td>27.9</td>
<td>34.2</td>
</tr>
<tr>
<td>Dietary intake, Mean ± SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy, kcal/day</td>
<td>1,950 ± 760</td>
<td>1,890 ± 717</td>
</tr>
<tr>
<td>Calcium, mg/day</td>
<td>790 ± 442</td>
<td>809 ± 440</td>
</tr>
<tr>
<td>Folate, mcg/day</td>
<td>373 ± 217</td>
<td>389 ± 222</td>
</tr>
<tr>
<td>Fiber, g/day</td>
<td>19.5 ± 9.8</td>
<td>19.8 ± 9.7</td>
</tr>
<tr>
<td>Red meat, servings/day</td>
<td>0.7 ± 0.6</td>
<td>0.6 ± 0.6</td>
</tr>
<tr>
<td>Processed meat, servings/day</td>
<td>0.4 ± 0.4</td>
<td>0.3 ± 0.4</td>
</tr>
<tr>
<td>Fruit, servings/day</td>
<td>2.0 ± 1.7</td>
<td>2.2 ± 1.8</td>
</tr>
<tr>
<td>Vegetable, servings/day</td>
<td>2.9 ± 2.5</td>
<td>2.6 ± 2.5</td>
</tr>
</tbody>
</table>

Note: Frequencies may not add up to 100 due to missing values. Abbreviations: GED, general educational development; HRT, hormone replacement therapy; NSAID, nonsteroidal anti-inflammatory drugs.

Discussion

In this study from three large international colorectal cancer consortia, we discovered a new locus located within the FMN1/GREM1 gene region (rs58349661) that interacts with BMI on the association with colorectal cancer risk. Analyses stratified by genotypes of rs58349661 showed that obesity and overweight were associated with a higher risk of colorectal cancer in those with the CC genotype but not in those with the TT genotype, which showed consistently null associations across BMI categories.

Our novel G×BMI finding at the FMN1/GREM1 locus is close to a region with multiple known GWAS loci for colorectal cancer risk. Among the ~200 SNPs associated with colorectal cancer in GWAS, an increasing number of independent loci have been located within the broader SCG5-GREM1-FMN1 region (rs17816465, rs10318, rs1919364, rs16969681, rs7779584; refs. 8, 30, 31). The lead SNP (rs58349661), as well as other correlated SNPs in our interaction analysis, has not been previously observed in GWAS analyses for colorectal cancer risk. Moreover, none of these genetic variants were in strong LD with previously known loci, suggesting that this interaction locus was independent of previously known loci. Our finding of a higher colorectal cancer risk associated with obesity and the CC genotype, and null associations with the TT genotype, suggests that T and C alleles may participate in specific mechanisms that influence the effect of obesity on colorectal cancer risk.

The mechanisms by which obesity affects colorectal cancer risk are not fully understood, but several plausible hypotheses have been proposed, including the colonic microenvironment and related systemic inflammation, oxidative stress, and major changes in other metabolic activities, including hyperinsulinemia and diabetes mellitus (32). At the histologic and cellular levels, obesity originates from uncontrolled hyperplasia and/or hypertrophy of the adipocytes. GWAS studies on adiposity traits, including BMI, have not reported any associations with SNPs within FMN1 gene (33–37). This is consistent with gene expression studies that did not report the major activity of FMN1 in matured adipocytes (38). In contrast to adipocytes, FMN1 is expressed in the colon and rectal mucosal cells (38). FMN1 is known to interact with alpha-catenin in the production of adherens...
junctions and polymerization of actin monomers (39). Hence, this phenomenon is mainly observed during the formation of epithelial sheets in adipocyte stem cells when linear actin cables and microtubules are formed at adherens junctions (40). This may explain why FMN1 activity is not observed in fully developed adipocytes. FMN1 has been reported as an epigenomic region associated with early childhood adiposity, a period in which adipose tissue is still in development (41).

Although very little is known about FMN1 and colorectal cancer, there are extensive data on formins, more generally, their role in developmental biology, WNT signaling, and their association with colorectal cancer. Humans have 15 Formin genes including FMN1 (38). These genes are defined by a 400 amino acid formin homology-2 (FH2) domain, which is the actin nucleation apparatus responsible for eukaryotic actin filament assembly and elongation (42, 43). FH2 influences actin dynamics and is common in all Formins across multiple eukaryotic species (44). Formins were originally identified at the mouse limb deformity locus (45) and FMN1 particularly is central to one of the human syndactyly disorders (46), another one of which involves a novel splicing mutation in APC that results in an ~80% reduction in the wildtype transcript (47). Several Formins (including DAAM1 and DAAM2) are key components of canonical WNT signaling in cancer development (48, 49), thus they are involved in APC mutation and colon polyp formation. A variety of Formins but not FMN1, have been implicated in invasion and metastasis of colorectal cancer (38). Additional studies are necessary to fully understand the possible contribution of FMN1 to colorectal tumor development according to obesity status.

In a previous G × BMI analysis, we observed a locus located within SMAD7 that interacts with BMI on colorectal cancer risk. Our current

**Table 2.** Summary of G × BMI analyses using 1DF, two-step, and 3DF analyses.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1DF</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-step</td>
<td>15 SNPs (E in LD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDGE</td>
<td>rs58349661</td>
<td>33122966</td>
<td>FNM1</td>
<td>15</td>
<td>C T</td>
<td>0.21</td>
<td>4.3 × 10⁻⁷</td>
<td>3.7 × 10⁻¹⁰</td>
<td>5.0 × 10⁻⁶</td>
<td>2.22 × 10⁻⁶</td>
<td>4.97 × 10⁻⁶</td>
</tr>
<tr>
<td>3DF</td>
<td>12 SNPs (E in G driven)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: DF, degrees of freedom; EDGE, elastic data-driven genetic encoding.

*Position based on NCBI Build37.

+Association between genetic variant and colorectal cancer.

+Association between genetic variant and BMI.

+Two-step EDGE filtering P value.

+Two-step EDGE second step P value.

Five SNPs, all located on FNM1 and LD (presented in Supplementary Tables). The SNP with the lowest P value was rs58349661.

Twelve SNPs were significant for G × BMI in the 3DF analyses (presented in Supplementary Tables). None of these SNPs had a marginal G effect, except for rs58349661 (previously observed in the two-step approach) and rs7313400. The latter was not significant in the 1DF analysis (P_{1DF} = 0.07).
Figure 2.
LocusZoom plots for the interaction between rs58349661 (rs1975678 used as perfect LD proxy) and colorectal cancer. The genomic position is shown on the x-axis whereas the y-axis reports the −log_{10} of the P value of the interaction with BMI. Purple dot, rs1975678. The colors of the SNPs are based on their correlation with rs1975678. Previously known GWAS variants, with their references, are also included: rs17816465, rs1919364, rs12708491.

Figure 3.
BMI in relation to colorectal cancer risk, stratified by genotype. Using normal BMI as the reference, risk for colorectal cancer was estimated for overweight and obesity across the genotypes of rs58349661.

### rs58349661xBMI stratified by genotype

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Cases/controls</th>
<th>OR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>7,382/11,095</td>
<td>1.00 (Reference)</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>9,678/13,327</td>
<td>1.13 (1.08–1.18)</td>
<td>7.7e–09</td>
</tr>
<tr>
<td>Obese</td>
<td>5,812/6,626</td>
<td>1.45 (1.38–1.52)</td>
<td>3.9e–50</td>
</tr>
<tr>
<td>CT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>3,939/5,459</td>
<td>1.00 (Reference)</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>5,115/6,525</td>
<td>1.12 (1.05–1.18)</td>
<td>2.0e–4</td>
</tr>
<tr>
<td>Obese</td>
<td>2,924/3,436</td>
<td>1.28 (1.19–1.37)</td>
<td>1.5e–12</td>
</tr>
<tr>
<td>TT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>549/691</td>
<td>1.00 (Reference)</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>666/827</td>
<td>1.02 (0.86–1.19)</td>
<td>0.85</td>
</tr>
<tr>
<td>Obese</td>
<td>350/465</td>
<td>1.00 (0.83–1.21)</td>
<td>0.98</td>
</tr>
</tbody>
</table>
analytical strategy excluded SMAD7 and other known GWAS loci for colorectal cancer and BMI. TGFβ in cellular proliferation and inflammation has been proposed as a mechanism of action through which SMAD7 interacts with colorectal cancer risk (50). In contrast to the tumor-promoting profile of SMAD7 in obesity, FMIN1 does not have a clear relationship with inflammation. It is possible that FMIN1 acts in tandem with neighboring genes, such as GREM1 and SCSG because the catena SCSG-GREM1-FMIN1 constitutes a hotspot for several colorectal cancer–related SNPs. Animal studies have shown that GREM1 may be activated by a cis-regulatory region within FMIN1 (51). Given that the SNPs interacting with BMI are located within the enhancer region, as indicated by high H3K27ac activity, it might also be possible that these SNPs impact GREM1. Moreover, GREM1 has been associated with cancer fibroblasts in colonic and rectal smooth muscles and colonic crypt bases, and may participate in colorectal tumorigenesis through bone morphogenetic protein (BMP) signaling (52). Mutations in GREM1 are associated with the development of hereditary mixed polyposis syndrome (HMPs), a rare condition associated with an increased development of colon polyps and higher colorectal cancer risk, often beginning in childhood. Overexpression of GREM1, as observed in HMPs, acts as a BMP antagonist, thus allowing the cells to conserve stem properties and develop into neoplasia (53). Experimental studies using animal and human models have shown that increased expression of GREM1 can induce epithelial dedifferentiation through BMP signaling and initiate gut cells neoplasia (54, 55). Future studies should specifically aim to unveil the potential mechanisms by which this SNP interacts with obesity on the association with colorectal cancer risk.

The strength of our study was the sample size, the largest ever to have examined G×BMI associations. The use of several complementary statistical approaches was also a strength because it allowed clear characterization of a specific locus within FMIN1/GREM1 that was consistently captured in the two-step and 3DF analyses, with a suggestive association indicated by the 1DF test. One limitation of our study is that it included only participants of European ancestry; hence, our findings may not be generalizable to other populations. Notably, the T allele of rs8349661 is common in other populations, with frequencies of 0.186, 0.518, 0.210, 0.307, 0.216 in African, East Asian, European, South Asian, and American populations, respectively. The A allele (rs8349661 is triallelic) is rare in all populations (frequency in Asian, European, and American populations <0.05), except in the African population where it has a frequency of 0.059. Another limitation is that a recall time of 1 to 2 years predates in some (but not all) case–control studies may not be enough to rule out weight loss due to the tumor. Such weight loss may also have affected the BMI at recruitment in colorectal cancer cases occurring during the early years of follow-up in cohort studies.

In conclusion, we identified a new locus in the FMIN1/GREM1 gene region that interacts with BMI on the association with colorectal cancer risk. This locus has not been previously described in relation to obesity or colorectal cancer, and additional investigation is required to elucidate the potential mechanisms by which it may modify the detrimental effects of obesity in promoting colorectal carcinogenesis.

Authors' Disclosures

J. Morrison reports grants from NCI during the conduct of the study. E.L. Barry reports personal fees from BioRealm LLC outside the submitted work. S.A. Bien reports being currently employed by and owning stock in Adaptive Biotechnologies. G. Casey reports grants from NIH/NCI during the conduct of the study and from NIH/NCI outside the submitted work. A.T. Chan reports grants and personal fees from Pfizer Inc., Freenome, Zee Ltd., Bayer Pharma AG, and Boehringer Ingelheim outside the submitted work. D.D. Buchanan reports grants from Genentech during the conduct of the study. G. Hapel reports personal fees from Genome Medical, Natera, Promega, Gil OxidDemand, 23andMe, and Carolien during the conduct of the study. S. Ogino reports grants from NIH during the conduct of the study. J. Chang-Claude: reports grants from Carlos III and Spanish Association Against Cancer (AECC) Scientific Foundation during the conduct of the study. J.C. Figueiredo: reports grants from National Heart, Lung and Blood Institute during the conduct of the study. L.C. Sakoda reports grants from NCI during the conduct of the study. V. Moreno reports grants from Instituto de Salud Carlos III and Spanish Association Against Cancer (AECC) Scientific Foundation during the conduct of the study. S. Ogino reports grants from NIH during the conduct of the study. R.K. Pae reports personal fees from Alimentive Inc., Allergan, Eli Lilly, and Verily outside the submitted work. E.A. Platz reports grants from NIH, AICR during the conduct of the study and personal fees from AACR outside the submitted work.

Disclaimer

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government.

Authors’ Contributions

Aglago et al.

review and editing. E.S. Kwagga: Resources, writing-review and editing. T.G. Kuku: Resources, writing-review and editing. A. Kundaje: Resources, writing-review and editing. S.C. Larson: Resources, writing-review and editing. L. Le Marchand: Resources, writing-review and editing. J.P. Lewinger: Resources, writing-review and editing. E. Li: Resources, writing-review and editing. B.M. Lynch: Resources, writing-review and editing. B. Mahawar: Resources, writing-review and editing. M. Mandic: Resources, writing-review and editing. M. Obin-Santacana: Resources, writing-review and editing. V. Moreno: Resources, writing-review and editing. N. Murphy: Resources, writing-review and editing. H. Nan: Resources, writing-review and editing. N. Nüssli: Resources, writing-review and editing. B. Pardamean: Resources, writing-review and editing. E.A. Platz: Resources, writing-review and editing. J.D. Potter: Resources, writing-review and editing. R.L. Prentice: Resources, writing-review and editing. G. Rennert: Resources, writing-review and editing. E. Ruiz-Narvaez: Resources, writing-review and editing. L.C. Sakoda: Resources, writing-review and editing. P.G. Scacheri: Resources, writing-review and editing. S.I. Schmit: Resources, writing-review and editing. R.E. Schoen: Resources, writing-review and editing. A. Scherlacha: Resources, writing-review and editing. M.L. Slattery: Resources, writing-review and editing. M.C. Stern: Resources, writing-review and editing. Y.-R. Su: Resources, writing-review and editing. C.M. Tangen: Resources, writing-review and editing. S.N. Thibodeau: Resources, writing-review and editing. D.C. Thomas: Resources, writing-review and editing. Y. Tian: Resources, writing-review and editing. C.M. Ulrich: Resources, writing-review and editing. F.J.B. van Duijn: Resources, writing-review and editing. B. Van Guelpen: Resources, writing-review and editing. K. Visvanathan: Resources, writing-review and editing. L. Hsu: Conceptualization, resources, methodology, writing-review and editing. W.J. Gauderman: Conceptualization, methodology, writing-review and editing. J. Wang: Resources, writing-review and editing. E. White: Resources, writing-review and editing. A. Wolk: Resources, writing-review and editing. O.M. Woods: Resources, writing-review and editing. A.H. Wu: Resources, writing-review and editing. N. Zemlianikova: Resources, methodology, writing-review and editing. L. Hsu: Conceptualization, resources, methodology, writing-review and editing. W.J. Gauderman: Conceptualization, resources, formal analysis, supervision, visualization, methodology, project administration, writing-review and editing. U. Peters: Conceptualization, supervision, funding acquisition, project administration, writing-review and editing. K.K. Tsilidis: Conceptualization, resources, supervision, funding acquisition, methodology, project administration, writing-review and editing. P.T. Campbell: Conceptualization, resources, supervision, funding acquisition, methodology, project administration, writing-review and editing.

Acknowledgments

E.K. Aglago’s association was supported by a Cancer Research UK (CRUK) grant awarded to K.K. Tsilidis (PPRCPJT [1000005]). GECCO: NCI, NIH, U.S. Department of Health and Human Services (U01 CA137088, R01 CA059045, U01 CA146930, R01CA164930). Additional funding for the OFCCR/ARCTIC was through award GL201-043 from the Ontario Research Fund, award 112746 from the Canadian Institutes of Health Research, through a Cancer Risk Evaluation (CaRE) Program grant from the Canadian Cancer Society (T. G. Gallinger et al.). Sample generation support from the Ontario Ministry of Research and Innovation. The SFOCR Illumina HumanCytoSNP array was supported in part through NCI/NIH awards U01/24 CA074794 and R01 CA076366 (to P.A. Newcomb). The content of this manuscript does not necessarily reflect the views or policies of the NCI, NIH or any of the collaborating centers in the Colon Cancer Family Registry (CCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government, any cancer registry, or the CCFR.

Colonrectal Cancer Transdisciplinary (CCFR) Study: The CCFR Study was supported by the NCI/NIH, U.S. Department of Health and Human Services (grant numbers U19 CA144107, R01 CA148488, P30 CA104898, R01CA197350, P01CA196659, R01 CA210407) and National Institutes of Environmental Health Sciences, NIH (grant number T32 ES013678). CORSA: The CORSA study was funded by Austrian Research Funding Agency (FFG) BRIDGE (grant 829675 to Andrea Gsur), the “Herzfelder’sche Foundation” (to Andrea Gsur), and was supported by COST Action BM1206.

CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study II (CPS-II) cohort. This study was conducted with Institutional Review Board approval.

CRCGEN: Colonrectal Cancer Genetics & Genomics, Spanish study was supported by Instituto de Salud Carlos III, co-funded by FEDER funds—a way to build Europe—(grants PI14-613 and PI09-1286), Agency for Management of University and Research Grants (AGAUR) of the Catalan Government (grant 2017SGR12073) and a personal associateship was supported by a Cancer Research UK (CRUK) grant (RO 2270/8). The collection and availability of cancer registry data is also supported by the American Institute for Cancer Research.

The Colon Cancer Family Registry (CCFR, www.coloncfr.org) is supported in part by funding from the NCI, NIH (award U01 CA167551). Support for case ascertainment was provided in part from the Surveillance, Epidemiology, and End Results (SEER) Program and the following U.S. cancer registries: AZ, CO, MN, NC, NH, and by the Victoria Cancer Registry (Australia) and Ontario Cancer Registry (Canada). The CCFR Set-1 (Illumina 1M/1M-Duo) and Set 2 (Illumina Omni-Quad) scans were supported by NHLBI awards U01 CA122859 and R01 CA143247 (to G. Casey). The CCFR Set-3 (Affymetrix Axxom CORRECT Set array) was supported by NIH award U19 CA148107 and R01 CA184188 (to S.B. Gruber). The CCFR Set-4 (Illumina OncoArray 600K SNP array) was supported by NIH award U19 CA148107 (to S.B. Gruber) and by the Center for Inherited Disease Research (CIDR), which is funded by the NIH to the Johns Hopkins University, contract number HHSN268201200008E. Additional funding for the OPCRARCARTIC was through award GL201-043 from the Ontario Research Fund, award 112746 from the Canadian Institutes of Health Research, through a Cancer Risk Evaluation (CaRE) Program grant from the Canadian Cancer Society (T. G. Gallinger et al.). Sample generation support from the Ontario Ministry of Research and Innovation. The SFOCR Illumina HumanCytoSNP array was supported in part through NCI/NIH awards U01/24 CA074794 and R01 CA076366 (to P.A. Newcomb). The content of this manuscript does not necessarily reflect the views or policies of the NCI, NIH or any of the collaborating centers in the Colon Cancer Family Registry (CCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government, any cancer registry, or the CCFR.

Cancer Res; 83(15) August 1, 2023

2580

CANCER RESEARCH

Downloaded from http://aacrjournals.org/cancerres/article-pdf/83/15/2572/3351936/2572.pdf by Wageningen University user on 10 September 2023
Gene by BMI Interactions and Risk of Colorectal Cancer

C8221/A91017 to EPIC-Oxford, Medical Research Council (1000413 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford, United Kingdom).

ESTHER/VERDI. This work was supported by grants from the Baden-Württemberg Ministry of Science, Research and Arts and the German Cancer Aid. Harvard cohorts: HPFS is supported by the NIH (P01 CA50507, UM1 CA167532, U10 CA167552, R01 CA137178, R01 CA151993, and R35 CA197735), NHLBI by the NIH (P01 CA87969, UM1 CA181607, R01 CA137178, R01 CA151993, and R35 CA197735), and PHS by the NIH (R01 CA42182).

Hawaii Adenoma Study: NCI grants R01 CA72520.

Kentucky: This work was supported by the following grant support: Clinical and Investigator Award from Damon Runyon Cancer Research Foundation (CI-8); NCI R01/CA136726.

LCCS: The Leids Colorectal Cancer Study was funded by the Food Standards Agency and Cancer Research UK Programme Award (C588/A19167).

MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NEMRC grants 508448, 209057, 251553, and 504711 and by infrastructure provided by Cancer Council Victoria.

Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database.

MEC: NIH (R37 CA54281, P01 CA036314, and R01 CA063444).

MECC: This work was supported by the NIH, U.S. Department of Health and Human Services (R01 CA81488, R01 CA197350).

NCCLS I & II: The authors acknowledge funding support for this project from the NIH, R01 CA66635 and P30 DK034987.

NFCCR: This work was supported by an Interdisciplinary Health Research Team award from the Canadian Institutes of Health Research (CRT 438211), the NIH, U.S. Department of Health and Human Services (U01 CA74783), and NCI of Canada grants (18223 and 18226). The authors wish to acknowledge the contribution of Alexandre Bellefeuille and the genotyping team of the McGill University and Gémeaux Québec Innovation Centre, Montréal, Canada, for genotyping the Sequenom panel in the NFPCR samples. Funding was provided to Michael O. Woods by the Canadian Cancer Society Research Institute.

PLCO: Intramural Research Program of the Division of Cancer Epidemiology and Genomics supported by contracts from the Division of Cancer Prevention, NCI, NIH, DHHS. Funding was provided by NIH, Genes, Environment and Health Initiative (GEI) Z01 CP 010200, NIH U01 HG004446, and NIH GEI U01 HG004438.

SELECT: Research reported in this publication was supported in part by the NCI of the NIH under award numbers U10 CA75429 (CD Blanke) and UM1 CA186107 (CM Tangen/IM Thompson). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

SMS: This work was supported by the NCI (grant P61 CA074184 to J.D. Potter and P.A. Newcomb, grants R01 CA97325, R03 CA53323, and K05 CA152715 to P.A. Newcomb, and the National Center for Advancing Translational Sciences at the NIH, grant KL2 TR000421 to K.A. Reilly-H.)

REACH: This work was supported by the NCI (grant P01 CA704184 to J.D. Potter and P.A. Newcomb, grants R01 CA97325, R03 CA53323, and K05 CA152715 to P.A. Newcomb, and the National Center for Advancing Translational Sciences at the NIH, grant KL2 TR000421 to K.A. Reilly-H.)

UK Biobank: This research has been conducted using the UK Biobank Resource under Application Number 8614.

VITAL: NIH (R05 CA153437).

WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, NIH, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100046C, HHSN268201100001C, HHSN268201100059C, HHSN268201100005C, HHSN268201100004C, and HHSN271201100004C.

CCFR: The Colon CCR graciously thanks the generous contributions of their study participants, dedication of study staff, and the financial support from the U.S. NCI, without which this important registry would not exist. The authors would like to thank the study participants and staff of the Seattle Colorectal Cancer Family Registry and the Hormones and Colon Cancer study (CORE Studies).

CLUE II: The authors thank the participants of CLUE II and appreciate the continued efforts of the staff at the Johns Hopkins George W. Comstock Center for Public Health Research and Prevention in the conduct of the CLUE II Cohort Study.

CORSA: The authors kindly thank all individuals who agreed to participate in the CORSA study. Furthermore, the authors thank all cooperating physicians and students and the Biobank Graz of the Medical University of Graz.

CPS-II: The authors thank the CPS-II participants and Study Management Group for their invaluable contributions to this research. The authors would also like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, and cancer registries supported by the NCI Surveillance Epidemiology and End Results program.

Czech Republic CCS: The authors are thankful to all clinicians in major hospitals in the Czech Republic, without whom the study would not be practicable. The authors are also sincerely grateful to all patients participating in this study.

DAHCH: The authors thank all participants and cooperating clinicians, and everyone who provided excellent technical assistance.

EDRN: The authors acknowledge all contributors to the development of the resource at University of Pittsburgh School of Medicine, Department of Gastroenterology, Department of Pathology, Hepatology and Nutrition and Biomedical Informatics.

EPIC: Where authors are identified as personnel of the International Agency for Research on Cancer/WHO, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/WHO.

Harvard cohorts (HPFS, NHS, PHS): The study protocol was approved by the Institutional Review Boards of the Brigham and Women’s Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required. The authors acknowledge Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital as home of the NHS. The authors would like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention’s National Program of Cancer Registries (NPCR) and/or the NCI’s Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries include the following: Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Maine, Maryland, Michigan, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, Seattle SEER Registry, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia, Wyoming. The authors assume full responsibility for analyses and interpretation of these data.

Kentucky: The authors would like to acknowledge the contribution to the study from central cancer registries supported through the Centers for Disease Control and Prevention’s National Program of Cancer Registries (NPCR) and/or the NCI’s Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries include the following: Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Maine, Maryland, Michigan, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, Seattle SEER Registry, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia, Wyoming. The authors assume full responsibility for analyses and interpretation of these data.

LCCS: The authors acknowledge the contributions of Jennifer Barrett, Robin Waxman, Gillian Smith, and Emma Northwood in conducting this study.

NCCLS I & II: The authors would like to thank the study participants and the NC Colorectal Cancer Study staff.

PLCO: The authors thank the PLCO Cancer Screening Trial screening center investigators and the staff from Information Management Services Inc and Westat Inc. Most importantly, the authors thank the study participants for their contributions that made this study possible.

Cancer incidence data have been provided by the Department of Columbia Cancer Registry, Georgia Cancer Registry, Hawaii Cancer Registry, Minnesota Cancer Surveillance System, Missouri Cancer Registry, Nevada Central Cancer Registry, Pennsylvania Cancer Registry, Texas Cancer Registry, Virginia Cancer Registry, and Wisconsin Cancer Reporting System. All are supported in part by funds from the Center for Disease Control and Prevention, National Program for Central Registries, local states or by the NCI, Surveillance, Epidemiology, and End Results program. The results reported here and the conclusions derived are the sole responsibility of the authors.

SELECT: The authors thank the research and clinical staff at the sites that participated on SELECT study, without whom the trial would not have been successful. The authors are also grateful to the 35,533 dedicated men who participated in SELECT.

WHI: The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20Written%20Paper/WHI%20Investigators%20Short%20List.pdf.

The publication costs of this article were defrayed in part by the payment of publication fees. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

Note

Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Received November 25, 2022; revised February 25, 2023; accepted May 24, 2023, published first May 30, 2023.


