



Article

Multiple Gene-Environment Interactions on the Angiogenesis Gene-Pathway Impact Rectal Cancer Risk and Survival

Noha Sharafeldin ^{1,2,*}, Martha L. Slattery ³, Qi Liu ¹, Conrado Franco-Villalobos ¹, Bette J. Caan ⁴, John D. Potter ^{5,6,7}  and Yutaka Yasui ^{1,8,*}

¹ School of Public Health, University of Alberta, Edmonton, AB T6G 2R3, Canada; ql3@ualberta.ca (Q.L.); conradofranco@ualberta.ca (C.F.-V.)

² Department of Medicine, School of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA

³ Department of Internal Medicine, University of Utah Health Sciences Center, Salt Lake City, UT 84132, USA; marty.slattery@hsc.utah.edu

⁴ Division of Research, Kaiser Permanente Medical Care Program, Oakland, CA 94612, USA; bette.caan@kp.org

⁵ Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA; jpotter@fredhutch.org

⁶ Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA 98195, USA

⁷ Centre for Public Health Research, Massey University, P.O. Box 756, Wellington 6140, New Zealand

⁸ Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, TN 38105, USA

* Correspondence: Sharafeldin.n@ualberta.ca (N.S.); yyasui@ualberta.ca (Y.Y.); Tel.: +1-205-638-2144 (N.S.); +1-901-595-5893 (Y.Y.)

Received: 30 July 2017; Accepted: 23 September 2017; Published: 28 September 2017

Abstract: Characterization of gene-environment interactions (GEIs) in cancer is limited. We aimed at identifying GEIs in rectal cancer focusing on a relevant biologic process involving the angiogenesis pathway and relevant environmental exposures: cigarette smoking, alcohol consumption, and animal protein intake. We analyzed data from 747 rectal cancer cases and 956 controls from the Diet, Activity and Lifestyle as a Risk Factor for Rectal Cancer study. We applied a 3-step analysis approach: first, we searched for interactions among single nucleotide polymorphisms on the pathway genes; second, we searched for interactions among the genes, both steps using Logic regression; third, we examined the GEIs significant at the 5% level using logistic regression for cancer risk and Cox proportional hazards models for survival. Permutation-based test was used for multiple testing adjustment. We identified 8 significant GEIs associated with risk among 6 genes adjusting for multiple testing: *TNF* (OR = 1.85, 95% CI: 1.10, 3.11), *TLR4* (OR = 2.34, 95% CI: 1.38, 3.98), and *EGR2* (OR = 2.23, 95% CI: 1.04, 4.78) with smoking; *IGF1R* (OR = 1.69, 95% CI: 1.04, 2.72), *TLR4* (OR = 2.10, 95% CI: 1.22, 3.60) and *EGR2* (OR = 2.12, 95% CI: 1.01, 4.46) with alcohol; and *PDGFB* (OR = 1.75, 95% CI: 1.04, 2.92) and *MMP1* (OR = 2.44, 95% CI: 1.24, 4.81) with protein. Five GEIs were associated with survival at the 5% significance level but not after multiple testing adjustment: *CXCR1* (HR = 2.06, 95% CI: 1.13, 3.75) with smoking; and *KDR* (HR = 4.36, 95% CI: 1.62, 11.73), *TLR2* (HR = 9.06, 95% CI: 1.14, 72.11), *EGR2* (HR = 2.45, 95% CI: 1.42, 4.22), and *EGFR* (HR = 6.33, 95% CI: 1.95, 20.54) with protein. GEIs between angiogenesis genes and smoking, alcohol, and animal protein impact rectal cancer risk. Our results support the importance of considering the biologic hypothesis to characterize GEIs associated with cancer outcomes.

Keywords: angiogenesis; candidate gene-pathway; gene-environment interactions; rectal cancer; cancer risk; cancer survival

1. Introduction

Studying genetic variants in epidemiologic studies is of great value for identifying disease risk and outcomes. Genetic associations are less sensitive to bias and represent valid, time-independent, biologically representative markers of disease [1]. In colorectal cancer, genome-wide association studies (GWAS) have identified multiple susceptibility loci in European populations [2–5]. Only a few of the identified single nucleotide polymorphisms (SNPs) have clear functional roles, possibly since SNP selection in GWAS is guided by linkage disequilibrium (LD) rather than functionality and it is difficult to determine whether the identified SNPs are causal or merely surrogates of the true causal variants [6]. Identified SNPs are typically of low-penetrance risk and, despite being common, their effects are usually small and of limited clinical impact. An analysis approach that aims to detect only significant marginal effects of an individual SNP on a disease would be successful if that individual SNP's function is in some way biologically critical to acquire the disease. However, in complex diseases such as cancer, the relationship between the phenotype and genotype is argued to fundamentally depend on the interaction between disease susceptibility loci and environmental exposures, referred to as gene-environment interactions (GEIs), which may well explain an important component of the “missing heritability” in cancer risk [7–10]. Furthermore, it is possible that individual loci are contributing to risk through a multi-gene model best approached using a pathway of biologic relevance to cancer based on a priori hypothesized mechanisms.

One of the critical cancer-related biologic processes necessary for tumor proliferation and progression in rectal cancer is angiogenesis: the fundamental process of sprouting and expansion of blood vessels from preexisting vessels [11]. Induction of angiogenesis seems to be an early event important for conversion of normal epithelium into cancer cells that influence risk of developing the disease while sustained angiogenesis is essential for tumor expansion ultimately influencing survival [12,13]. In this study, we focused on angiogenesis-related genes and aimed to construct a working pathway that captures a network of genes potentially influencing rectal cancer susceptibility and survival. Specific environmental exposures have been identified in the etiology of colorectal cancer. Evidence suggests diet low in fiber, fruit and vegetables, and high in calories, refined grains, fat content and red and processed meat is associated with an increased risk of colorectal cancer. Lifestyle factors have also been suggested to increase risk including smoking and alcohol consumption, while physical activity, use of non-steroidal anti-inflammatory drugs, increased intake of vitamin D and calcium have a reduced risk of colorectal cancer [14]. Our approach to testing pathway GEI effects on rectal cancer risk and survival was based on the biologic hypothesis involving selection of candidate genes in the angiogenesis pathway and three environmental exposures relevant to angiogenesis: smoking, alcohol consumption, and dietary protein.

The evidence of association of cigarette smoking and alcohol consumption with rectal cancer has been inconclusive [15–23]. Nicotine in tobacco smoke and ethanol has been shown to stimulate angiogenesis under ischemic conditions [24,25], thus examining these associations in genetically susceptible individuals carries the potential of strengthening these associations. In addition, certain dietary patterns, specifically those that contain high consumption of red and processed meat, are associated with a moderate increased risk of rectal cancer [26–30]. Plausible mechanisms are related to the content of meat (protein, iron) [31–33] not absorbed by the small intestine and transferred to the large intestine lumen and, when in excess, may have toxic effects on the large intestine mucosa [34]. The higher intake of protein and a decrease in its digestibility leads to more undigested proteins fermented by colonic bacteria which may promote DNA damage and loss of large intestine epithelial cell homeostasis and ultimately tumor growth [34]. Protein fermentation mainly occurs in the distal parts of the colon and rectum and previously reported associations of meat intake have been generally stronger for distal colon and rectal cancer [35,36] while animal protein intake has been previously associated with colorectal adenoma [37]. Diet is a complex mixture of many nutrients and characterization of GEI could help determine the specific nutrients affecting the cancer risk and cancer-related mortality. We focused on animal protein as a nutrient based on its documented stimulatory effect on angiogenesis [38,39].

In this study, we hypothesized that high animal/vegetable protein intake ratio and prolonged intense pattern of cigarette smoking influence hypoxia (oxygen deprivation) and long-term alcohol intake influences hypoglycemia (glucose deprivation). Both hypoxia and hypoglycemia are ischemic conditions that enhance angiogenesis [40–42]. We carried the logic of the biologic hypothesis to the analysis through applying a novel stepped approach to GEI testing. We have previously applied our method to colon cancer data and identified multiple significant GEIs for colon cancer risk [43]. In this study we report on our results investigating effects of GEI on rectal cancer risk and survival.

2. Materials and Methods

2.1. Study Population

This analysis was based on a multicenter, population-based, case-control study of rectal cancer (The Diet, Activity and Lifestyle as a Risk Factor for Rectal Cancer) conducted in two geographic areas in the United States: Utah and Northern California [44]. A rapid-reporting system was used to identify rectal cancer cases during the period between May 1997 and May 2001, with the majority of cases being interviewed within 4 months of diagnosis. Cases were identified directly through local tumor registries. Case eligibility was determined according to the Surveillance Epidemiology and End Results (SEER) Cancer Registries in Northern California and Utah. Eligibility criteria were: being 30 to 79 years of age at time of diagnosis; speaking English; and being mentally and physically competent to complete the interview. Cases with history of previous colorectal cancer or known familial adenomatous polyposis (as indicated on pathology reports), ulcerative colitis, or Crohn's disease were not eligible.

Criteria for eligibility for controls were the same as for cases. Controls were frequency matched to cases by sex and 5-year age groups in each geographical area. Controls were randomly selected at Kaiser Permanente Medical Care Program (KPMCP) of Northern California from health maintenance organization membership lists; in Utah controls aged 65 years or more from Health Care Financing Administration lists, and controls aged less than 65 years from driver's license lists [45]. Of rectal cancer study subjects contacted, 65.2% cases and 65.3% controls were interviewed. The response rates from the study were not greatly different than those reported in other epidemiologic studies [45].

Ethical statement: All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki. This analysis only included data from participants who agreed to use their information for further studies (roughly 99%) and received research ethics approval from the University of Alberta Health Research Ethics Board (HREB), No. Pro00026736. All previously conducted study procedures were approved by ethics committees at their respective study locations.

2.2. Interview Data

Trained and certified interviewers conducted a detailed computerized in-person interview that took approximately 2 to 3 h to complete [46]. Participants completed two questionnaires: (a) the health and lifestyle questionnaire (among data collected were demographic characteristics, medical history, meal patterns, smoking and alcohol consumption information); and (b) a diet-history questionnaire on dietary intakes. Dietary intake was ascertained using an adaptation of the CARDIA diet history [47–49]. A period of 2–3 calendar years prior to rectal cancer diagnosis in cases or 2–3 calendar years prior to enrollment in controls was used as the referent time period for study questionnaires.

2.3. Tumor Registry Data

Tumor registry data were obtained from local tumor registries to determine disease stage at diagnosis, months of survival after diagnosis, and vital status. Disease stage was categorized using the SEER staging criteria (in-situ, local, regional, distant, and unknown) [50]. Follow-up was obtained for all study participants and was terminated at the end of the year 2007. All study participants had over five years of follow-up.

2.4. Candidate Gene-Pathway

Genetic markers were genotyped using a multiplexed bead-array assay format based on Golden Gate chemistry (Illumina Human Hap550k, San Diego, CA, USA) and TaqMan assay from Applied Biosystems (Foster City, CA, USA). Genotyping details are provided in Supplementary Materials.

We constructed a working figure of the angiogenesis gene-pathway relevant to rectal cancer to guide the analysis (Figure 1). The process involved extracting information from the standard pathway maps and pathway text descriptions from three recognized web-based resources: The BioCarta organization, KEGG (Kyoto Encyclopedia of Genes and Genomes), and Cell Signaling Technologies (CST). We specifically searched these resources using the keyword “angiogenesis” and extracted information from the “VEGF, Hypoxia, and Angiogenesis Pathway” from the BioCarta Pathways (http://www.biocarta.com/pathfiles/h_vegfPathway.asp); the “VEGF Signaling Pathway” available from the KEGG Pathway database (http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=map04370&keyword=angiogenesis); and the “Angiogenesis Signaling Pathway” from the CST pathways (<http://www.cellsignal.com/common/content/content.jsp?id=pathways-angiogenesis>). We also conducted supplementary searches of online gene databases and PubMed for information on biologic functions of the candidate genes and experimental observations of biologic activities of genes in relation to tumor angiogenesis. Examination of the molecular interactions as illustrated in the pathway maps along with their descriptions, and the information on biologic activity of the genes guided the candidate-gene selection and provided rationale for grouping genes in specific sub-pathways. Genes were included in the working pathway figure as either key components of the angiogenesis pathway or secondary interacting genes (Figure 1).

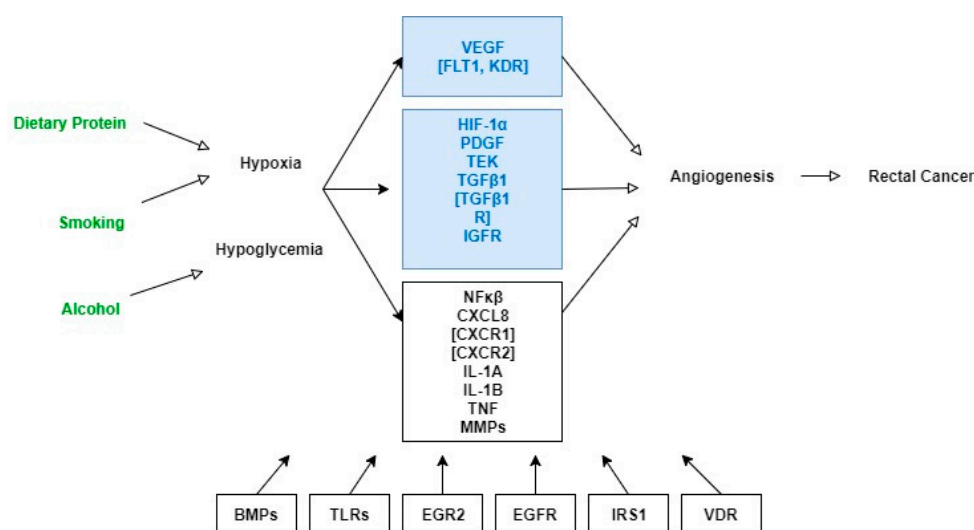


Figure 1. Working Figure of the Angiogenesis Pathway Genes. Key gene components of pathway in blue frames: *VEGF* = vascular endothelial growth factor; *FLT1* = vascular endothelial growth factor receptor 1; *KDR* = vascular endothelial growth factor receptor 2; *HIF-1* = hypoxia-inducible factor 1; *PDGF* = platelet-derived growth factor; *TEK* = TEK receptor tyrosine kinase; *TGFβ* = Transforming growth factor, beta; *TGFβR* = Transforming growth factor, beta receptor; *IGFR* = insulin-like growth factor receptor. Secondary interacting genes of pathway in black frames: *NF-kB* = nuclear factor of kappa light polypeptide gene enhancer in B-cells; *CXCL8* = C-X-C motif chemokine ligand 8; *CXCR1* = C-X-C motif chemokine receptor 1; *CXCR2* = C-X-C motif chemokine receptor 2; *IL-1A* = interleukin-1, alpha; *IL-1B* = interleukin-1, beta; *TNF* = tumor necrosis factor; *MMPs* (*MMP1*, *MMP3*, *MMP7*, *MMP9*) = matrix metalloproteinases; *BMPs* (*BMP1*, *BMP2*, *BMP4*, *BMPR1A*, *BMPR1B*, *BMPR2*, *GDF10*) = bone morphogenetic proteins; *TLRs* (*TLR2*, *TLR3*, *TLR4*) = toll-like receptors; *EGR2* = early growth response 2; *EGFR* = epidermal growth factor receptor; *IRS1* = insulin receptor substrate 1; *VDR* = Vitamin D Receptor. Environmental factors in green text.

2.5. Environmental Variables

2.5.1. Cigarette Smoking

An individual was considered a regular cigarette smoker if they smoked at least 100 cigarettes during their lifetime; otherwise they were classified as never having smoked. For smokers, pack-years of cigarettes smoked was determined by multiplying the usual number of cigarettes smoked per day by total years of smoking cigarettes (determined by taking into account start- and stop-dates of smoking), and dividing by 20. For this analysis, subjects were categorized using a cut-off of 20 pack-years (20 or more pack-years, less than 20 pack-years, and never smoked).

2.5.2. Alcohol

Participants were asked to report usual amounts consumed separately during weekdays and weekend days to better capture total alcohol consumption. Additionally, participants were asked about alcohol consumption 10 and 20 years ago as part of the health-and-lifestyle questionnaire. Alcoholic beverages were defined as beer, wine, and hard liquor including alcoholic cocktails, whiskey, gin, vodka, scotch, bourbon, or rum. Participants who responded with “no” to the question, “Did you ever drink an average of one or more alcoholic beverages a month for a year or longer?” were considered as never having drunk alcohol. Participants who responded “yes” to this question were then asked the usual number of 12-ounce bottles of beer, 4-ounce glasses of wine, and 1.5-ounce shots of hard liquor consumed 10 and 20 years ago. Long-term exposure to alcohol, based on consumption of any type of alcoholic beverage 10 and 20 years prior to the referent year, was categorized into two levels of consumption (none to moderate; high; cut-off was 20 gms/week for men and 10 gms/week for women).

2.5.3. Dietary Protein

Nutrient information was obtained by converting food-intake data into nutrient data using the Minnesota Nutrition Coordinating Center nutrient database [51]. Total protein intake included animal proteins (meats, poultry, fish, dairy, and eggs) and vegetable proteins (legumes, tofu). We calculated an animal/vegetable protein intake ratio (animal protein proportion of total protein intake). Using a cut-off equivalent to 1.5 animal/vegetable protein intake ratio we divided it in two categories (low and high animal/vegetable protein intake ratio).

2.6. Statistical Analysis

We applied a 3-step analysis framework to modeling candidate pathway GEIs consisting of the following steps (Table 1) [43]:

Step 1, developing a summary profile for each gene on the candidate pathway referred to as a *gene-specific tree* (GST). We used logic regression [52] to search for SNP-set interactions within each gene. Details on the method are provided in Supplementary Materials. The GSTs, rather than individual SNPs, were used as building blocks for the next two steps.

Step 2, modelling gene-set interactions across the full pathway referred to as *pathway tree(s)* by searching for GST-set interactions using logic regression. Pathway trees are adjusted for in the GEI models of the next step.

Step 3, modelling pathway GEIs between the GSTs and the three environmental exposures. Guided by the pathway figure (Figure 1), we first divided the full pathway into nine sub-pathways (grouped in boxes in the figure) and summarized GEIs in each sub-pathway using backward selection that eliminated the least significant interaction term(s) in a stepwise fashion. The GEIs that remained in the sub-pathway summary models at the 5% significance level were jointly tested in the final GEI model for the entire pathway. We fitted logistic regression models for rectal cancer risk and Cox proportional hazards regression models for rectal cancer survival. All models, in addition to adjusting for the pathway trees, were adjusted for age at diagnosis or selection, sex, race (white, Hispanic, or African American), and study center (University of Utah or the KPMCP of Northern

California). Baseline hazards for Cox proportional hazards models were stratified by rectal cancer stage at diagnosis. The GEI models were fitted using Stata version 12. For pathway-level significance of GEIs, a p -value threshold of $p \leq 0.05$ was applied. To correct for occurrence of false positives, we performed a permutation-based test. We performed a 1000 permutations of the 0/1 values of the GSTs and the three environmental exposures and repeated Step 3 of our analysis to correct for multiple testing. The permutation-based test at Step 3 is justified by the independence of the GEI testing from the prior two steps [53] and provided sets of random GSTs and random exposures while maintaining a similar prevalence to the original GSTs/exposures.

Table 1. Summary of the 3-step candidate pathway gene-environment interaction approach.

Analysis Step	Interaction of Interest	Variable of Interest	Model	Specific Procedures	Product
Step 1: Summarize gene effects	SNP-set interaction within gene	SNPs on each gene separately	Logic regression with logit link/fitting exponential survival models	Cross-validation to determine optimal model size	Gene-specific trees (GSTs)
Step 2: Summarize pathway effects	Gene-set interaction within pathway	All GSTs on the pathway	Logic regression with logit link/fitting exponential survival models	Cross-validation to determine optimal model size	Pathway Trees
Step 3: Test gene-environment interaction	Gene-environment interaction within pathway	a. Sub-pathway specific GSTxE * b. Full pathway GSTxE * [§]	Logistic regression model [§] /Cox Proportional Hazards model [¶]	Statistical significance testing	Pathway GEIs

* GSTxE, gene-specific tree—environment interaction; [§] Models adjusted for age, sex, race, study center, pathway trees; [¶] Models adjusted for age, sex, race, study center, pathway tree, stratified by cancer stage.

3. Results

We analyzed data from 747 rectal cancer cases and 956 controls. There were no statistical differences between cases and controls with respect to age, sex, race, education level, marital status and annual income (Table 2). Univariate analysis show significant associations between rectal cancer and cigarette smoking ($p = 0.001$), alcohol ($p = 0.02$) and dietary protein ($p = 0.02$) (Table 2). The angiogenesis candidate gene-pathway included a total of 257 SNPs belonging to 34 angiogenesis-related genes (Figure 1). Results from the first two analysis steps are provided in supplementary materials (Supplementary Materials Tables S1 and S2 and Figures S1 and S2). The third and final step of the analysis modeled the pathway GEIs; statistically significant GEI results are displayed in Table 3 for rectal cancer risk and Table 4 for rectal cancer survival. For most GEIs, we observed a positive gradient in the magnitude of the main GST effects with increasing levels of animal protein intake, smoking, and alcohol consumption.

Eight significant GEIs were associated with rectal cancer risk involving six genes. Two genes were among the major drivers of angiogenesis: *PDGFB* rs4821877 with high animal/vegetable protein intake (interaction odds ratio (OR_{INT}) = 1.75, 95% confidence interval (CI) (1.04, 2.92), $p = 0.034$) and *IGF1R* rs2139924 with long-term alcohol consumption (OR_{INT} = 1.69, 95% CI (1.04, 2.72), $p = 0.033$). Other statistically significant GEIs were: *TNF* rs1800630 (OR_{INT} = 1.85, 95% CI (1.10, 3.11), $p = 0.021$) with ≥ 20 pack-years of smoking and *MMP1* rs470215 (OR_{INT} = 2.44, 95% CI (1.24, 4.81), $p = 0.010$). Both complementary GSTs for *TLR4* and *EGR2* genes were interacting with both smoking and alcohol consumption. *TLR4* rs1927911 AND rs11536889 with ≥ 20 pack-years of smoking (OR_{INT} = 2.34, 95% CI (1.38, 3.98), $p = 0.002$), *TLR4* rs1927911 OR rs11536889 with long-term alcohol consumption (OR_{INT} = 2.10, 95% CI (1.22, 3.60), $p = 0.007$); *EGR2* rs2295814 with ≥ 20 pack-years of smoking (OR_{INT} = 2.23, 95% CI (1.04, 4.78), $p = 0.040$) with long-term alcohol consumption (OR_{INT} = 2.12, 95% CI (1.01, 4.46), $p = 0.048$). Adjustment for multiple testing showed none of the 1000 permutation runs resulted in eight or more significant GEIs (only two runs had a maximum number of six significant GEIs), providing evidence that our GEIs associated with rectal cancer risk are unlikely to be attributed to chance alone.

Table 2. Demographic characteristics of study participants.

Characteristic	Rectal Cancer Cases (n = 747)	Controls (n = 956)	p-Value
Age (Mean, SD)	61.2 (10.8)	62.1 (10.6)	0.09
Sex			
Male	447 (59.8%)	542 (56.7%)	0.19
Female	300 (40.2%)	414 (43.3%)	
Race			
White non-Hispanic	617 (82.6%)	821 (85.9%)	0.06
Other	130 (17.4%)	135 (14.1%)	
Education			
≤High School	266 (35.6%)	324 (33.9%)	0.46
>High School	481 (64.4%)	632 (66.1%)	
Marital Status			
Married	556 (74.4%)	730 (76.4%)	0.36
Other	191 (25.6%)	226 (23.6%)	
Annual Income			
≤30 K	206 (29.9%)	238 (27.4%)	0.27
>30 K	483 (70.1%)	632 (72.6%)	
Cigarette Smoking			
Non-smoker	346 (46.3%)	485 (50.7%)	0.001
≤20 pack-years	158 (21.2%)	240 (25.1%)	
>20 pack-years	243 (32.5%)	231 (24.2%)	
Alcohol			
Non/Moderate	556 (74.4%)	759 (79.4%)	0.02
Heavy	191 (25.6%)	197 (20.6%)	
Animal/Vegetable Protein Ratio			
Low	242 (32.4%)	363 (38.0%)	0.02
High	505 (67.6%)	593 (62.0%)	
Center			
Utah	270 (36.1%)	366 (38.3%)	0.37
Northern California	477 (63.9%)	590 (61.7%)	
Cancer Stage			
In-situ	20 (2.7%)		
Local	395 (52.9%)		
Regional	255 (34.1%)		
Distant	63 (8.4%)		
Unknown	14 (1.9%)		

Five GEIs were associated with survival, four of which were interactions with high animal/vegetable protein intake: *KDR* rs6838752 (interaction hazard ratio $HR_{INT} = 4.12$, 95% CI (1.52, 11.13), $p = 0.005$), *TLR2* rs7656411 ($HR_{INT} = 8.69$, 95% CI (1.09, 69.12), $p = 0.041$), *EGR2* rs224082 ($HR_{INT} = 2.41$, 95% CI (1.40, 4.15), $p = 0.002$), and *EGFR* rs17151957 ($HR_{INT} = 5.84$, 95% CI (1.80, 18.94), $p = 0.003$). The fifth significant interaction was *CXCR1* rs1008562 with ≥ 20 pack-years of smoking ($HR_{INT} = 2.05$, 95% CI (1.12, 3.76), $p = 0.019$). The GEIs associated with rectal cancer survival, however, were not statistically significant after adjustment for multiple testing. The permutation-based test showed 114 of the 1000 permutation runs resulted in five or more statistically significant GEIs.

Table 3. Effects of gene-environment interactions significant at 5% level between gene-specific trees and environmental factors on rectal cancer risk.

GST	Gene	Chr.	Cases (%)	Control (%)	Gene OR ^a (95% CI)	Env. Factor	Category	N (%) ^b	Gene OR by Env. Factor (95% CI)	OR _{INT} ^c (95% CI)	P _{INT} ^d
rs4821877 (CC or CT)	<i>PDGFB</i>	22q13.1	610 (80.7%)	746 (77.6%)	1.21 (0.95, 1.54)	Protein ^e	Low High	612 (35.6%) 1106 (64.4%)	0.85 (0.57, 1.26) 1.47 (1.08, 2.00)	Ref 1.75 (1.04, 2.92)	0.034
rs2139924 (AA)	<i>IGF1R</i>	15q26.3	243 (30.4%)	287 (28.5%)	0.93 (0.77, 1.00)	Alcohol	Non/Moderate Heavy	1396 (77.4%) 408 (22.6%)	0.82 (0.66, 1.02) 1.36 (0.91, 2.05)	Ref 1.69 (1.04, 2.72)	0.033
rs1800630 (CA or AA)	<i>TNF</i>	6p21.33	240 (31.8%)	267 (27.8%)	1.19 (0.96, 1.47)	Smoking	Non <20 PY ≥20 PY	834 (48.7%) 400 (23.4%) 477 (27.9%)	0.95 (0.70, 1.30) 1.13 (0.71, 1.81) 1.68 (1.11, 2.54)	Ref 1.14 (0.65, 2.01) 1.85 (1.10, 3.11)	0.644 0.021
rs470215 (TT or TC)	<i>MMP1</i>	11q22.2	715 (90.3%)	880 (87.9%)	1.23 (0.89, 1.69)	Protein ^e	Low High	640 (35.7%) 1153 (64.3%)	0.67 (0.40, 1.14) 1.78 (1.18, 2.70)	Ref 2.44 (1.24, 4.81)	0.010
rs1927911 (CC) rs11536889 (GG)	<i>TLR4</i>	9q33.1	396 (52.4%) 546 (72.2%)	495 (51.5%) 684 (71.1%)	0.93 (0.76, 1.15)	Smoking	Non <20 PY ≥20 PY	834 (48.7%) 400 (23.4%) 477 (27.9%)	0.80 (0.59, 1.08) 0.65 (0.41, 1.04) 1.33 (0.90, 1.98)	Ref 0.99 (0.57, 1.74) 2.34 (1.38, 3.98)	0.980 0.002
rs1927911 (CT or TT) rs11536889 (GC or CC)	<i>TLR4</i>	9q33.1	360 (47.6%) 210 (27.8%)	467 (48.5%) 278 (28.9%)	1.07 (0.87, 1.32)	Alcohol	Non/Moderate Heavy	1326 (77.2%) 391 (22.8%)	0.95 (0.75, 1.21) 1.58 (1.01, 2.47)	Ref 2.10 (1.22, 3.60)	0.007
rs2295814 (GA or AA)	<i>EGR2</i>	10q21.3	106 (14.0%)	115 (12.0%)	1.11 (0.83, 1.49)	Smoking	Non <20 PY ≥20 PY	834 (48.7%) 400 (23.4%) 477 (27.9%)	0.90 (0.58, 1.37) 1.21 (0.65, 2.28) 1.53 (0.89, 2.65)	Ref 1.84 (0.83, 4.09) 2.23 (1.04, 4.78)	0.130 0.040
rs2295814 (GG)	<i>EGR2</i>	10q21.3	650 (86.0%)	847 (88.0%)	0.90 (0.67, 1.20)	Alcohol	Non/Moderate Heavy	1326 (77.2%) 391 (22.8%)	0.81 (0.57, 1.14) 1.21 (0.69, 2.11)	Ref 2.12 (1.01, 4.46)	0.048

Abbreviations: GST, Gene-Specific Tree; Chr., Chromosome; Env., Environmental; PY, pack-years; OR, odds ratio; CI, confidence interval. ^a Gene odds ratios were adjusted for age, sex, race, study center, pathway trees; ^b N (%) frequency of the environmental variable within subjects with the GST; ^c OR_{INT}: Interaction Odds Ratio; ^d P_{INT}: Interaction *p*-value; ^e Animal/Vegetable Protein Ratio.

Table 4. Effects of gene-environment interactions significant at 5% level between gene-specific trees and environmental factors on rectal cancer survival.

GST	Gene	Chr.	Cases (%)	Gene HR ^a (95% CI)	Env. Factor	Category	N (%) ^b	Gene OR by Env. Factor ^a (95% CI)	HR _{INT} ^c (95% CI)	P _{INT} ^d
rs6838752 (TT or TC)	KDR	4q12	705 (93.6%)	0.89 (0.55, 1.45)	Protein ^e	Low	258 (32.4%)	0.44 (0.21, 0.91)	Ref	0.005
						High	538 (67.6%)	1.43 (0.73, 2.83)	4.12 (1.52, 11.13)	
rs1008562 (GG)	CXCR1	2q35	211 (27.9%)	1.17 (0.89, 1.53)	Smoking	Non	348 (46.2%)	1.04 (0.68, 1.60)	Ref	0.905
						<20 PY	160 (21.2%)	0.88 (0.44, 1.75)	0.96 (0.46, 1.98)	
						≥20 PY	245 (32.5%)	1.88 (1.20, 2.95)	2.05 (1.12, 3.76)	0.019
rs7656411 (GG)	TLR2	4q31.3	61 (8.1%)	0.83 (0.48, 1.44)	Protein ^e	Low	244 (32.3%)	0.13 (0.02, 0.98)	Ref	0.041
					High	512 (67.7%)	1.33 (0.74, 2.38)	8.69 (1.09, 69.12)		
rs224082 (GA or AA)	EGR2	10q21.3	455 (60.2%)	0.72 (0.56, 0.92)	Protein ^e	Low	244 (32.3%)	0.39 (0.25, 0.62)	Ref	0.002
					High	512 (67.7%)	0.93 (0.68, 1.23)	2.41 (1.40, 4.15)		
rs17151957 (AA)	EGFR	7p11.2	41 (6.5%)	1.82 (1.16, 2.88)	Protein ^e	Low	244 (32.3%)	0.54 (0.19, 1.53)	Ref	0.003
					High	512 (67.7%)	3.37 (1.95, 5.82)	5.84 (1.80, 18.94)		

Abbreviations: GST, Gene-Specific Tree; Chr., Chromosome; Env., Environmental; PY, pack-years; HR, hazard ratio; CI, confidence interval. ^a Gene hazard ratios were adjusted for age, sex, race, study center, pathway tree, baseline hazard stratified by cancer stage; ^b N (%) frequency of the environmental variable within subjects with the GST; ^c HR_{INT}: Interaction hazards ratio; ^d P_{INT}: Interaction *p*-value; ^e Animal/Vegetable Protein Ratio.

4. Discussion

Using a stepped approach to testing GEIs at the gene-pathway level, we identified eight statistically significant interactions between angiogenesis genes and smoking, alcohol, and animal protein intake on rectal cancer risk adjusting for multiple testing. Our approach emphasized the biologic hypothesis through construction of a working pathway figure of select angiogenesis-related genes and relevant environmental exposures and used it to guide the analysis. This is in contrast to the common approach to evaluating GEI in cancer through investigation of interactions between known common susceptibility loci (i.e., strong and statistically significant GWAS or candidate-gene findings) and established risk factors for the cancer. Despite the large size of these studies combining case-control and/or nested case-control samples, they have provided to-date limited evidence of GEI in colorectal cancer [54–57]. Difficulties in identifying GEIs may be attributed to the choice, measurement, and modeling of environmental exposures [58]. Furthermore, modelling SNP genotypes as indicator variables for one and two minor alleles rather than the number of SNP minor alleles potentially strengthens the detected GEI [59]. Certain considerations embedded in the framework of our approach to GEI testing addresses some of these issues. The choice of environmental variables was based on the hypothesis that protein intake, smoking, and alcohol are enhancing angiogenesis and interacting with the angiogenesis genes in the state of tumor ischemia. Lifestyle factors such as smoking and alcohol consumption could be considered more strictly “environmental”, having less genetic influence compared to other complex risk factors with more pronounced genetic influence (e.g., body mass index). We focused on intense and long-term patterns of smoking and alcohol consumption and used indicators of SNP genotypes to develop the gene-specific trees potentially enhancing the capacity of our approach to detect GEI in rectal cancer. Moreover, using RegulomeDB [60,61], there is bioinformatics evidence that SNPs identified in our results play regulatory roles in rectal mucosa: however, expression quantitative trait loci (eQTL) data showing the association with specific gene expressions were not available. The tissue specificity and chromatin states where these SNPs are located provide additional support for the plausibility of the biological function of the identified GEIs.

Studies that differentiated cancer by site show that the association of smoking is stronger for rectal cancer risk and mortality compared to colon cancer [62,63]. Recent genome-wide interaction studies, however, found no significant GEIs between smoking and colorectal cancer [64,65]. This may be attributed to combining colon and rectal cancer data and/or limited characterization of smoking variables including duration of smoking or time since quitting smoking. In our study, we considered amount, duration of exposure, and start- and stop-dates of smoking in participants compared to never-exposed participants, which maximizes power of detecting an interaction with the gene and avoids dilution of risk by ‘short-term’ and/or ‘long-since-quit’ former users. The genome-wide analysis for interaction between genetic variants and alcohol consumption using data from 14 studies identified significant interactions between 11 SNPs at the 9q22.32/*HIATL1* locus and light-to-moderate drinking with no evidence of heterogeneity across studies [65]. In our study, we similarly observed significant interactions on chromosome 9 in the 9q32-q33/*TLR4* locus with smoking (rs1927911 AND rs11536889) and alcohol (rs1927911 OR rs11536889) on rectal cancer risk. We have previously reported on associations of *TLR2* and *TLR4* SNPs with colon cancer risk and survival [66] and identified significant GEI of alcohol and *TLR2* gene with colon cancer risk [43]. Toll-like receptors (TLRs) play a key role in the innate immune system and are important mediators of inflammation in the gut, potentially modulating colorectal cancer risk. Tissue expression studies suggest involvement of *TLR2* [67] and *TLR4* [68] in colorectal carcinogenesis. Functional polymorphisms in both genes were also found in association with colorectal cancer risk, and their effects were modified by obesity and smoking [69]. The inflammatory response of TLRs is mediated through NF- κ B pathway shown to be activated by cigarette smoke [70]. Triggering an inflammatory response that leads to tumor promotion provides support to the biologic plausibility of these interactions. Such finding could provide insights into new drug targets for rectal cancer similar to the inhibition of NF- κ B dependent intestinal inflammation

attained by targeting an enteroglia-specific protein/TLR4 axis that demonstrated therapeutic effects in ulcerative colitis [71].

A recent genome-wide analysis identified an interaction between a SNP on chromosome 10p14 near the *GATA3* gene and processed meat that modified colorectal cancer risk [72], suggesting that *GATA3* transcription triggers a pro-tumorigenic inflammatory response to processed meat. In our results, we identified interactions of high animal protein intake on rectal cancer risk with *PDGFB* and *MMP1* genes previously implicated in inflammation-mediated pathologic processes including tumor progression [73]. The *GATA3* transcription factor was found to potentially mediate different expression levels of *MMP1* [74], and our observed *MMP1*-animal protein GEI is thus providing further characterization of the *GATA3*/processed-meat interaction. In addition, four of the five observed GEIs on rectal cancer survival were with high animal protein intake including with the *KDR* gene, a *VEGF* receptor that mediates *VEGF-A* induced production of Nitric Oxide (NO) by endothelial cells [75]. A high protein diet leads to a high amine concentration (due to the excess intake and increased fermentation) which in the presence of NO yields the potentially carcinogenic nitrosoamines (Nitric Oxide (NO) added to the amine). Other GEIs involved *TLR2* and *EGR2* genes, both related to the same signaling pathway, where evidence has shown *TLR* expression and signaling mediates the response of intestinal epithelial cells to bacterial antigens possibly increasing the rate of protein fermentation [76].

Our candidate approach compared to a pure empirical approach to examining GEI was able to detect an appreciable number of novel GEIs, nonetheless with some limitations. The GEI models were large, involving many GST-environment interactions across the pathway; accordingly we limited the adjustment variables to the most relevant rectal cancer risk and survival predictors (such as age, sex, race, study center, and cancer stage). Although it is possible that we missed important angiogenesis genes when developing the working pathway figure, our candidate pathway included major genes implicated in rectal carcinogenesis. We used cross-validation to specify model size for the logic regression models and, as such, summarization of the gene effects was limited by the specified model size in addition to the number of tagSNPs on each gene. Our candidate associations were biologically hypothesized a priori, however, GEI effects on rectal cancer survival did not remain statistically significant after multiple testing adjustment and need to be interpreted with caution. There were multiple strengths to our analysis based on a population-based case-control study of rectal cancer. Data were collected through a standardized interview process to minimize interviewer bias; long-term exposure information for smoking and alcohol were collected; and data on confounding variables were available. The interviewer-administered questionnaires were extensive and captured more detailed exposure information than is available from self-administered questionnaires. The major strength of the analysis, however, was our integration of the relevant biologic information in the construction of the pathway that was carried throughout the analysis process.

5. Conclusions

Our approach to pathway analysis provided a powerful tool to elucidate the overall effects of the angiogenesis pathway genes and their interaction with three environmental exposures on rectal cancer risk: cigarette smoking, alcohol consumption and animal protein intake. The angiogenesis pathway is one of the hallmarks of cancer, and findings could be potentially informative for other solid tumors. The diet and lifestyle factors are, in theory modifiable and, given the magnitude of the detected GEIs, this provides essential insights for preventive strategies, identifying drug targets, and opens avenues for personalized preventive and treatment strategies.

Supplementary Materials: The following are available online at www.mdpi.com/1660-4601/14/10/1146/s1, Supplementary Methods, Figure S1: Gene-pathway tree in association with rectal cancer risk, Figure S2: Gene-pathway tree in association with rectal cancer survival, Table S1: Rectal cancer gene-specific trees and rectal cancer risk, Table S2: Rectal cancer gene-specific trees and rectal cancer survival.

Acknowledgments: This work was supported by an operating grant from the Canadian Institutes of Health Research to Y.Y. The study was funded by NCI grants CA48998 and CA615757. This research also was supported by

the Utah Cancer Registry funded by Contract #N01-PC-67000 from the National Cancer Institute, with additional support from the State of Utah Department of Health, the Northern California Cancer Registry, and the Sacramento Tumor Registry. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official view of the National Cancer Institute. Y.Y. is also supported by the Canada Research Chair Program, and Y.Y. and N.S. are supported by Alberta Innovates-Health Solutions. N.S. is supported by Alberta Cancer Foundation and the University of Alberta Dissertation Fellowship. We would like to acknowledge the contributions of Sandra Edwards, Roger Edwards, Leslie Palmer, Donna Schaffer, Kristin Anderson, and Judy Morse for data management and collection.

Author Contributions: Authors contributed equally to this work. Y.Y. and N.S. developed the gene-environment interaction analysis framework. N.S. constructed the candidate gene-pathway. M.L.S. provided critical input for selection of genes and environmental variables and for the design and analysis strategy. M.L.S. and B.J.C. provided the data. N.S., Q.L. and C.F.-V. programmed and conducted the analysis. J.D.P., M.L.S., and B.J.C. provided critical input to interpretation and reporting of the analysis results. N.S. and Y.Y. drafted the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Le Marchand, L.; Wilkens, L.R. Design considerations for genomic association studies: Importance of gene-environment interactions. *Cancer Epidemiol. Biomark. Prev.* **2008**, *17*, 263–267. [[CrossRef](#)] [[PubMed](#)]
2. Whiffin, N.; Houlston, R.S. Architecture of inherited susceptibility to colorectal cancer: A voyage of discovery. *Genes* **2014**, *5*, 270–284. [[CrossRef](#)] [[PubMed](#)]
3. Schumacher, F.R.; Schmit, S.L.; Jiao, S.; Edlund, C.K.; Wang, H.; Zhang, B.; Hsu, L.; Huang, S.C.; Fischer, C.P.; Harju, J.F.; et al. Genome-wide association study of colorectal cancer identifies six new susceptibility loci. *Nat. Commun.* **2015**, *6*, 7138. [[CrossRef](#)] [[PubMed](#)]
4. Peters, U.; Bien, S.; Zubair, N. Genetic architecture of colorectal cancer. *Gut* **2015**, *64*, 1623–1636. [[CrossRef](#)] [[PubMed](#)]
5. Zeng, C.; Matsuda, K.; Jia, W.H.; Chang, J.; Kweon, S.S.; Xiang, Y.B.; Shin, A.; Jee, S.H.; Kim, D.H.; Zhang, B.; et al. Identification of Susceptibility Loci and Genes for Colorectal Cancer Risk. *Gastroenterology* **2016**, *150*, 1633–1645. [[CrossRef](#)] [[PubMed](#)]
6. Hindorff, L.A.; Sethupathy, P.; Junkins, H.A.; Ramos, E.M.; Mehta, J.P.; Collins, F.S.; Manolio, T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9362–9367. [[CrossRef](#)] [[PubMed](#)]
7. Thomas, D. Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. *Annu. Rev. Public Health* **2010**, *31*, 21–36. [[CrossRef](#)] [[PubMed](#)]
8. Thomas, D. Gene—Environment-wide association studies: Emerging approaches. *Nat. Rev. Genet.* **2010**, *11*, 259–272. [[CrossRef](#)] [[PubMed](#)]
9. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753. [[CrossRef](#)] [[PubMed](#)]
10. Marian, A.J. Elements of ‘missing heritability’. *Curr. Opin. Cardiol.* **2012**, *27*, 197–201. [[CrossRef](#)] [[PubMed](#)]
11. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674. [[CrossRef](#)] [[PubMed](#)]
12. Folkman, J.; Watson, K.; Ingber, D.; Hanahan, D. Induction of angiogenesis during the transition from hyperplasia to neoplasia. *Nature* **1989**, *339*, 58–61. [[CrossRef](#)] [[PubMed](#)]
13. Ross, R. Angiogenesis. Successful growth of tumours. *Nature* **1989**, *339*, 16–17. [[CrossRef](#)] [[PubMed](#)]
14. Johnson, C.M.; Wei, C.; Ensor, J.E.; Smolenski, D.J.; Amos, C.I.; Levin, B.; Berry, D.A. Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control* **2013**, *24*, 1207–1222. [[CrossRef](#)] [[PubMed](#)]
15. Potter, J.D. Colorectal cancer: Molecules and populations. *J. Natl. Cancer Inst.* **1999**, *91*, 916–932. [[CrossRef](#)] [[PubMed](#)]
16. Ferrari, P.; Jenab, M.; Norat, T.; Moskal, A.; Slimani, N.; Olsen, A.; Tjønneland, A.; Overvad, K.; Jensen, M.K.; Boutron-Ruault, M.C.; et al. Lifetime and baseline alcohol intake and risk of colon and rectal cancers in the European prospective investigation into cancer and nutrition (EPIC). *Int. J. Cancer* **2007**, *121*, 2065–2072. [[CrossRef](#)] [[PubMed](#)]
17. Slattery, M.L.; Potter, J.D.; Friedman, G.D.; Ma, K.N.; Edwards, S. Tobacco use and colon cancer. *Int. J. Cancer* **1997**, *70*, 259–264. [[CrossRef](#)]

18. Cheng, J.; Chen, Y.; Wang, X.; Wang, J.; Yan, Z.; Gong, G.; Li, G.; Li, C. Meta-analysis of prospective cohort studies of cigarette smoking and the incidence of colon and rectal cancers. *Eur. J. Cancer Prev.* **2015**, *24*, 6–15. [[CrossRef](#)] [[PubMed](#)]
19. Poynter, J.N.; Haile, R.W.; Siegmund, K.D.; Campbell, P.T.; Figueiredo, J.C.; Limburg, P.; Young, J.; Le Marchand, L.; Potter, J.D.; Cotterchio, M.; et al. Associations between smoking, alcohol consumption, and colorectal cancer, overall and by tumor microsatellite instability status. *Cancer Epidemiol. Biomark. Prev.* **2009**, *18*, 2745–2750. [[CrossRef](#)] [[PubMed](#)]
20. Gong, J.; Hutter, C.; Baron, J.A.; Berndt, S.; Caan, B.; Campbell, P.T.; Casey, G.; Chan, A.T.; Cotterchio, M.; Fuchs, C.S.; et al. A pooled analysis of smoking and colorectal cancer: Timing of exposure and interactions with environmental factors. *Cancer Epidemiol. Biomark. Prev.* **2012**, *21*, 1974–1985. [[CrossRef](#)] [[PubMed](#)]
21. Wei, E.K.; Giovannucci, E.; Wu, K.; Rosner, B.; Fuchs, C.S.; Willett, W.C.; Colditz, G.A. Comparison of risk factors for colon and rectal cancer. *Int. J. Cancer* **2004**, *108*, 433–442. [[CrossRef](#)] [[PubMed](#)]
22. Cho, E.; Smith-Warner, S.A.; Ritz, J.; van den Brandt, P.A.; Colditz, G.A.; Folsom, A.R.; Freudenheim, J.L.; Giovannucci, E.; Goldbohm, R.A.; Graham, S.; et al. Alcohol intake and colorectal cancer: A pooled analysis of 8 cohort studies. *Ann. Intern. Med.* **2004**, *140*, 603–613. [[CrossRef](#)] [[PubMed](#)]
23. Fedirko, V.; Tramacere, I.; Bagnardi, V.; Rota, M.; Scotti, L.; Islami, F.; Negri, E.; Straif, K.; Romieu, I.; La Vecchia, C.; et al. Alcohol drinking and colorectal cancer risk: An overall and dose-response meta-analysis of published studies. *Ann. Oncol.* **2011**, *22*, 1958–1972. [[CrossRef](#)] [[PubMed](#)]
24. Heeschen, C.; Chang, E.; Aicher, A.; Cooke, J.P. Endothelial progenitor cells participate in nicotine-mediated angiogenesis. *J. Am. Coll. Cardiol.* **2006**, *48*, 2553–2560. [[CrossRef](#)] [[PubMed](#)]
25. Gu, J.W.; Elam, J.; Sartin, A.; Li, W.; Roach, R.; Adair, T.H. Moderate levels of ethanol induce expression of vascular endothelial growth factor and stimulate angiogenesis. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2001**, *281*, R365–R372. [[PubMed](#)]
26. Larsson, S.C.; Wolk, A. Meat consumption and risk of colorectal cancer: A meta-analysis of prospective studies. *Int. J. Cancer* **2006**, *119*, 2657–2664. [[CrossRef](#)] [[PubMed](#)]
27. Gonzalez, C.A.; Riboli, E. Diet and cancer prevention: Contributions from the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Eur. J. Cancer* **2010**, *46*, 2555–2562. [[CrossRef](#)] [[PubMed](#)]
28. Chan, D.S.; Lau, R.; Aune, D.; Vieira, R.; Greenwood, D.C.; Kampman, E.; Norat, T. Red and processed meat and colorectal cancer incidence: Meta-analysis of prospective studies. *PLoS ONE* **2011**, *6*, e20456. [[CrossRef](#)] [[PubMed](#)]
29. Alexander, D.D.; Miller, A.J.; Cushing, C.A.; Lowe, K.A. Processed meat and colorectal cancer: A quantitative review of prospective epidemiologic studies. *Eur. J. Cancer Prev.* **2010**, *19*, 328–341. [[CrossRef](#)] [[PubMed](#)]
30. Alexander, D.D.; Cushing, C.A. Red meat and colorectal cancer: A critical summary of prospective epidemiologic studies. *Obes. Rev.* **2011**, *12*, e472–e493. [[CrossRef](#)] [[PubMed](#)]
31. Reddy, B.S.; Narisawa, T.; Weisburger, J.H. Effect of a diet with high levels of protein and fat on colon carcinogenesis in F344 rats treated with 1,2-dimethylhydrazine. *J. Natl. Cancer Inst.* **1976**, *57*, 567–569. [[CrossRef](#)] [[PubMed](#)]
32. Sugimura, T.; Sato, S. Mutagens-carcinogens in foods. *Cancer Res.* **1983**, *43*, 2415s–2421s. [[PubMed](#)]
33. Sun, Z.; Zhu, Y.; Wang, P.P.; Roeborn, B.; Zhao, J.; Zhao, J.; Dicks, E.; Cotterchio, M.; Buehler, S.; Campbell, P.T.; et al. Reported intake of selected micronutrients and risk of colorectal cancer: Results from a large population-based case-control study in Newfoundland, Labrador and Ontario, Canada. *Anticancer Res.* **2012**, *32*, 687–696. [[PubMed](#)]
34. Kim, E.; Coelho, D.; Blachier, F. Review of the association between meat consumption and risk of colorectal cancer. *Nutr. Res.* **2013**, *33*, 983–994. [[CrossRef](#)] [[PubMed](#)]
35. Silvester, K.R.; Cummings, J.H. Does digestibility of meat protein help explain large bowel cancer risk? *Nutr. Cancer* **1995**, *24*, 279–288. [[CrossRef](#)] [[PubMed](#)]
36. Chao, A.; Thun, M.J.; Connell, C.J.; McCullough, M.L.; Jacobs, E.J.; Flanders, W.D.; Rodriguez, C.; Sinha, R.; Calle, E.E. Meat consumption and risk of colorectal cancer. *JAMA* **2005**, *293*, 172–182. [[CrossRef](#)] [[PubMed](#)]
37. Yang, S.Y.; Kim, Y.S.; Song, J.H.; Chung, S.J.; Lee, I.H.; Hong, K.J.; Lee, E.J.; Kim, D.H.; Yim, J.Y.; Park, M.J.; et al. Dietary risk factors in relation to colorectal adenoma. *Korean J. Gastroenterol.* **2012**, *60*, 102–108. [[CrossRef](#)] [[PubMed](#)]

38. Speliotes, E.K.; Willer, C.J.; Berndt, S.I.; Monda, K.L.; Thorleifsson, G.; Jackson, A.U.; Allen, H.L.; Lindgren, C.M.; Luan, J.; Magi, R.; et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **2010**, *42*, 937–948. [[CrossRef](#)] [[PubMed](#)]
39. Grandison, R.C.; Piper, M.D.; Partridge, L. Amino-acid imbalance explains extension of lifespan by dietary restriction in *Drosophila*. *Nature* **2009**, *462*, 1061–1064. [[CrossRef](#)] [[PubMed](#)]
40. Dor, Y.; Porat, R.; Keshet, E. Vascular endothelial growth factor and vascular adjustments to perturbations in oxygen homeostasis. *Am. J. Physiol. Cell Physiol.* **2001**, *280*, C1367–C1374. [[PubMed](#)]
41. Barral, M.C.; Jimenez-Aparicio, R.; Priego, J.L.; Royer, E.C.; Urbanos, F.A.; Amador, U. Diruthenium(II,III) Carboxylate Compounds: Existence of both Polymeric and Ionic Forms in Solution and Solid State. *Inorg. Chem.* **1998**, *37*, 1413–1416. [[CrossRef](#)] [[PubMed](#)]
42. Harris, A.L. Hypoxia—A key regulatory factor in tumour growth. *Nat. Rev. Cancer* **2002**, *2*, 38–47. [[CrossRef](#)] [[PubMed](#)]
43. Sharafeldin, N.; Slattery, M.L.; Liu, Q.; Franco-Villalobos, C.; Caan, B.J.; Potter, J.D.; Yasui, Y. A Candidate-Pathway Approach to Identify Gene-Environment Interactions: Analyses of Colon Cancer Risk and Survival. *J. Natl. Cancer Inst.* **2015**, *107*. [[CrossRef](#)] [[PubMed](#)]
44. Slattery, M.L.; Caan, B.J.; Benson, J.; Murtaugh, M. Energy balance and rectal cancer: An evaluation of energy intake, energy expenditure, and body mass index. *Nutr. Cancer* **2003**, *46*, 166–171. [[CrossRef](#)] [[PubMed](#)]
45. Slattery, M.L.; Edwards, S.L.; Caan, B.J.; Kerber, R.A.; Potter, J.D. Response rates among control subjects in case-control studies. *Ann. Epidemiol.* **1995**, *5*, 245–249. [[CrossRef](#)]
46. Edwards, S.; Slattery, M.L.; Mori, M.; Berry, T.D.; Caan, B.J.P.; Palmer, P.; Potter, J.D. Objective system for interviewer performance evaluation for use in epidemiologic studies. *Am. J. Epidemiol.* **1994**, *140*, 1020–1028. [[CrossRef](#)] [[PubMed](#)]
47. Slattery, M.L.; Caan, B.J.; Duncan, D.; Berry, T.D.; Coates, A.; Kerber, R. A computerized diet history questionnaire for epidemiologic studies. *J. Am. Diet Assoc.* **1994**, *94*, 761–766. [[CrossRef](#)]
48. Liu, K.; Slattery, M.; Jacobs, D., Jr.; Cutter, G.; McDonald, A.; Van Horn, L.; Hilner, J.E.; Caan, B.; Bragg, C.; Dyer, A.; et al. A study of the reliability and comparative validity of the cardia dietary history. *Ethn. Dis.* **1994**, *4*, 15–27. [[PubMed](#)]
49. McDonald, A.; Van Horn, L.; Slattery, M.; Hilner, J.; Bragg, C.; Caan, B.; Jacobs, D., Jr.; Liu, K.; Hubert, H.; Gernhofer, N.; et al. The CARDIA dietary history: Development, implementation, and evaluation. *J. Am. Diet Assoc.* **1991**, *91*, 1104–1112. [[PubMed](#)]
50. Young, J.L.J.; Roffers, S.D.; Ries, L.A.G.; Fritz, A.G.; Hurlbut, A.A. *SEER Summary Staging Manual—2000: Codes and Coding Instructions*; National Cancer Institute: Rockville, MD, USA, 2001.
51. Dennis, B.; Ernst, N.; Hjortland, M.; Tillotson, J.; Grambsch, V. The NHLBI nutrition data system. *J. Am. Diet Assoc.* **1980**, *77*, 641–647. [[PubMed](#)]
52. Ruczinski, I.; Kooperberg, C.; LeBlanc, M.L. Logic regression. *J. Comput. Graph. Stat.* **2003**, *12*, 475–511. [[CrossRef](#)]
53. Dai, J.Y.; Kooperberg, C.; Leblanc, M.; Prentice, R.L. Two-stage testing procedures with independent filtering for genome-wide gene-environment interaction. *Biometrika* **2012**, *99*, 929–944. [[CrossRef](#)] [[PubMed](#)]
54. Siegert, S.; Hampe, J.; Schafmayer, C.; von Schonfels, W.; Egberts, J.H.; Forsti, A.; Chen, B.; Lascorz, J.; Hemminki, K.; Franke, A.; et al. Genome-wide investigation of gene-environment interactions in colorectal cancer. *Hum. Genet.* **2013**, *132*, 219–231. [[CrossRef](#)] [[PubMed](#)]
55. Hutter, C.M.; Chang-Claude, J.; Slattery, M.L.; Pflugeisen, B.M.; Lin, Y.; Duggan, D.; Nan, H.; Lemire, M.; Rangrej, J.; Figueiredo, J.C.; et al. Characterization of gene-environment interactions for colorectal cancer susceptibility loci. *Cancer Res.* **2012**, *72*, 2036–2044. [[CrossRef](#)] [[PubMed](#)]
56. Figueiredo, J.C.; Lewinger, J.P.; Song, C.; Campbell, P.T.; Conti, D.V.; Edlund, C.K.; Duggan, D.J.; Rangrej, J.; Lemire, M.; Hudson, T.; et al. Genotype-environment interactions in microsatellite stable/microsatellite instability-low colorectal cancer: Results from a genome-wide association study. *Cancer Epidemiol. Biomark. Prev.* **2011**, *20*, 758–766. [[CrossRef](#)] [[PubMed](#)]
57. Kantor, E.D.; Hutter, C.M.; Minnier, J.; Berndt, S.I.; Brenner, H.; Caan, B.J.; Campbell, P.T.; Carlson, C.S.; Casey, G.; Chan, A.T.; et al. Gene-environment interaction involving recently identified colorectal cancer susceptibility loci. *Cancer Epidemiol. Biomark. Prev.* **2014**, *23*, 1824–1833. [[CrossRef](#)] [[PubMed](#)]
58. Prentice, R.L. Empirical evaluation of gene and environment interactions: Methods and potential. *J. Natl. Cancer Inst.* **2011**, *103*, 1209–1210. [[CrossRef](#)] [[PubMed](#)]

59. Prentice, R.L.; Huang, Y.; Hinds, D.A.; Peters, U.; Pettinger, M.; Cox, D.R.; Beilharz, E.; Chlebowski, R.T.; Rossouw, J.E.; Caan, B.; et al. Variation in the FGFR2 gene and the effects of postmenopausal hormone therapy on invasive breast cancer. *Cancer Epidemiol. Biomark. Prev.* **2009**, *18*, 3079–3085. [[CrossRef](#)] [[PubMed](#)]
60. Boyle, A.P.; Hong, E.L.; Hariharan, M.; Cheng, Y.; Schaub, M.A.; Kasowski, M.; Karczewski, K.J.; Park, J.; Hitz, B.C.; Weng, S.; et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **2012**, *22*, 1790–1797. [[CrossRef](#)] [[PubMed](#)]
61. Xie, D.; Boyle, A.P.; Wu, L.; Zhai, J.; Kawli, T.; Snyder, M. Dynamic trans-acting factor colocalization in human cells. *Cell* **2013**, *155*, 713–724. [[CrossRef](#)] [[PubMed](#)]
62. Liang, P.S.; Chen, T.Y.; Giovannucci, E. Cigarette smoking and colorectal cancer incidence and mortality: Systematic review and meta-analysis. *Int. J. Cancer* **2009**, *124*, 2406–2415. [[CrossRef](#)] [[PubMed](#)]
63. Botteri, E.; Iodice, S.; Bagnardi, V.; Raimondi, S.; Lowenfels, A.B.; Maisonneuve, P. Smoking and colorectal cancer: A meta-analysis. *JAMA* **2008**, *300*, 2765–2778. [[CrossRef](#)] [[PubMed](#)]
64. Jiao, S.; Peters, U.; Berndt, S.; Bezieau, S.; Brenner, H.; Campbell, P.T.; Chan, A.T.; Chang-Claude, J.; Lemire, M.; Newcomb, P.A.; et al. Powerful Set-Based Gene-Environment Interaction Testing Framework for Complex Diseases. *Genet. Epidemiol.* **2015**, *39*, 609–618. [[CrossRef](#)] [[PubMed](#)]
65. Gong, J.; Hutter, C.M.; Newcomb, P.A.; Ulrich, C.M.; Bien, S.A.; Campbell, P.T.; Baron, J.A.; Berndt, S.I.; Bezieau, S.; Brenner, H.; et al. Genome-Wide Interaction Analyses between Genetic Variants and Alcohol Consumption and Smoking for Risk of Colorectal Cancer. *PLoS Genet.* **2016**, *12*, e1006296. [[CrossRef](#)] [[PubMed](#)]
66. Slattery, M.L.; Herrick, J.S.; Bondurant, K.L.; Wolff, R.K. Toll-like receptor genes and their association with colon and rectal cancer development and prognosis. *Int. J. Cancer* **2012**, *130*, 2974–2980. [[CrossRef](#)] [[PubMed](#)]
67. Nihon-Yanagi, Y.; Terai, K.; Murano, T.; Matsumoto, T.; Okazumi, S. Tissue expression of Toll-like receptors 2 and 4 in sporadic human colorectal cancer. *Cancer Immunol. Immunother.* **2012**, *61*, 71–77. [[CrossRef](#)] [[PubMed](#)]
68. Tchorzewski, M.; Lewkowicz, P.; Dziki, A.; Tchorzewski, H. Expression of toll-like receptors on human rectal adenocarcinoma cells. *Arch. Immunol. Ther. Exp.* **2014**, *62*, 247–251. [[CrossRef](#)] [[PubMed](#)]
69. Pimentel-Nunes, P.; Teixeira, A.L.; Pereira, C.; Gomes, M.; Brandao, C.; Rodrigues, C.; Goncalves, N.; Boal-Carvalho, I.; Roncon-Albuquerque, R., Jr.; Moreira-Dias, L.; et al. Functional polymorphisms of Toll-like receptors 2 and 4 alter the risk for colorectal carcinoma in Europeans. *Dig. Liver Dis.* **2013**, *45*, 63–69. [[CrossRef](#)] [[PubMed](#)]
70. Anto, R.J.; Mukhopadhyay, A.; Shishodia, S.; Gairola, C.G.; Aggarwal, B.B. Cigarette smoke condensate activates nuclear transcription factor-kappaB through phosphorylation and degradation of I-kappaB(alpha): Correlation with induction of cyclooxygenase-2. *Carcinogenesis* **2002**, *23*, 1511–1518. [[CrossRef](#)] [[PubMed](#)]
71. Esposito, G.; Capoccia, E.; Turco, F.; Palumbo, I.; Lu, J.; Steardo, A.; Cuomo, R.; Sarnelli, G.; Steardo, L. Palmitoylethanolamide improves colon inflammation through an enteric glia/toll like receptor 4-dependent PPAR-alpha activation. *Gut* **2014**, *63*, 1300–1312. [[CrossRef](#)] [[PubMed](#)]
72. Figueiredo, J.C.; Hsu, L.; Hutter, C.M.; Lin, Y.; Campbell, P.T.; Baron, J.A.; Berndt, S.I.; Jiao, S.; Casey, G.; Fortini, B.; et al. Genome-wide diet-gene interaction analyses for risk of colorectal cancer. *PLoS Genet.* **2014**, *10*, e1004228. [[CrossRef](#)] [[PubMed](#)]
73. Brinckerhoff, C.E.; Rutter, J.L.; Benbow, U. Interstitial collagenases as markers of tumor progression. *Clin. Cancer Res.* **2000**, *6*, 4823–4830. [[PubMed](#)]
74. Affara, M.; Dunmore, B.J.; Sanders, D.A.; Johnson, N.; Print, C.G.; Charnock-Jones, D.S. MMP1 bimodal expression and differential response to inflammatory mediators is linked to promoter polymorphisms. *BMC Genom.* **2011**, *12*, 43. [[CrossRef](#)] [[PubMed](#)]
75. Kroll, J.; Waltenberger, J. VEGF-A induces expression of eNOS and iNOS in endothelial cells via VEGF receptor-2 (KDR). *Biochem. Biophys. Res. Commun.* **1998**, *252*, 743–746. [[CrossRef](#)] [[PubMed](#)]
76. Singh, J.C.; Cruickshank, S.M.; Newton, D.J.; Wakenshaw, L.; Graham, A.; Lan, J.; Lodge, J.P.; Felsburg, P.J.; Carding, S.R. Toll-like receptor-mediated responses of primary intestinal epithelial cells during the development of colitis. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2005**, *288*, G514–G524. [[CrossRef](#)] [[PubMed](#)]

