Int J Mol Epidemiol Genet 2016;7(1):45-57
www.ijmeg.org /ISSN:1948-1756/IJMEG0012975

Original Article

Genes, environment and gene expression in colon tissue: a pathway approach to determining functionality

Martha L Slattery, Daniel F Pellatt, Roger K Wolff, Abbie Lundgreen

Department of Internal Medicine, 383 Colorow, University of Utah School of Medicine, Salt Lake, USA

Received July 16, 2015; Accepted August 19, 2015; Epub March 23, 2016; Published March 30, 2016

Abstract: Genetic and environmental factors have been shown to work together to alter cancer risk. In this study we evaluate previously identified gene and lifestyle interactions in a candidate pathway that were associated with colon cancer risk to see if these interactions altered gene expression. We analyzed non-tumor RNA-seq data from 144 colon cancer patients who had genotype, recent cigarette smoking, diet, body mass index (BMI), and recent aspirin/non-steroidal anti-inflammatory use data. Using a false discovery rate of 0.1, we evaluated differential gene expression between high and low levels of lifestyle exposure and genotypes using DESeq2. Thirteen pathway genes and 17 SNPs within those genes were associated with altered expression of other genes in the pathway. BMI, NSAIDs use and dietary components of the oxidative balance score (OBS) also were associated with altered gene expression. SNPs previously identified as interacting with these lifestyle factors, altered expression of pathway genes. NSAIDs interacted with 10 genes (15 SNPs) within those genes to alter expression of 28 pathway genes; recent cigarette smoking interacted with seven genes (nine SNPs) to alter expression of 27 genes. BMI interacted with FLT1, KDR, SEPN1, TERT, TXNRD2, and VEGFA to alter expression of eight genes. Three genes (five SNPs) interacted with OBS to alter expression of 12 genes. These data provide support for previously identified lifestyle and gene interactions associated with colon cancer in that they altered expression of key pathway genes. The need to consider lifestyle factors in conjunction with genetic factors is illustrated.

Keywords: Gene expression, inflammation, BMI, NSAIDs, diet, colon cancer

Introduction

Interpreting the functionality between single-nucleotide polymorphisms (SNPs) based on associations with outcomes, such as cancer, is often difficult. Associations with tagSNPs can result from being in LD with other disease-causing SNPs or can be the result of chance findings. To help guard against erroneous associations various methods taking into account multiple comparisons are employed. Likewise, in silico programs are available to help predict functionality based on their involvement in splicing, transcription, translation, and post-translation [1, 2], although studies have found that the prediction made by these programs do not correspond with associations observed in analytical studies [3]. We and others have utilized gene expression techniques to help determine potential functionality. However, gene expression associated with specific genotypes can help provide an indication of functionality if it exists, but lack of an association does not rule out functionality [4-6]. In our previous work, we shown that candidate genes and their associated SNPs influence expression of pathway genes to a greater extent than they do their own expression [4].

The examination of the interaction between genetic variants and lifestyle-related factors has been a hallmark of our ability to comprehend the complicated factors contributing to disease. It has long been hypothesized that genes and environment work together in creating disease risk and that consideration of one without the other leaves a void in our understanding of the carcinogenic process. Our examination of The Convergence of Hormones, Inflammation, and Energy-Related Factors (CHIEF) Pathway has illustrated both the risk associated with genetic variation [7] and that lifestyle factors further modify the risk associated with these genes [8, 9]. Important lifestyle fac-

tors that interact with genes in the pathway are dietary components such as antioxidants that interact with MAPK [9, 10], angiogenesis [8], and TLR [11] genes; aspirin and/or ibuprofen use with angiogenesis [8], cytokines [12, 13], MAPK [9], JAK/STAT [14], TGFβ-signaling pathway [15], TLR [11], and selenoprotein genes [16]; cigarette smoking with angiogenesis [8], cytokines [13, 17], MAPK [9], JAK/STAT [14], TGFB1 [15], and selenoprotein 16 pathway-related genes; and BMI with angiogenesis [8], estrogen-related genes [18], VDR [19], and cytokines [13, 17].

In this paper, we build on our previous work with the CHIEF pathway, which is composed of genes associated with inflammation, angiogenesis, energy-related factors, and hormones [20]. Previously we examined associations between pathway genes and expression of select genes within the pathway [4]. In this study, we evaluate expression of all pathway-related genes with genotypes of pathway genes. We test our previously identified genes and SNPs that were associated with colon cancer using ARTP method to evaluate significance over the pathway [7]. We further evaluated the functionality of the pathway, by assessing how genotypes combined with lifestyle factors impact gene expression. We focus our assessment of lifestyle factors on variables associated with oxidative stress (i.e. cigarette smoking, recent use of aspirin and/or ibuprofen referred to as NSAIDs, body mass index (BMI), and our oxidative balance score (OBS) and its related dietary components). Based on our previous work, we hypothesize that genetic variation in the pathway influences expression of other genes in the pathway; the altered expression is influenced by the combination of genetic and lifestyle-related factors.

Methods

Total RNA was available from colonic non-tumor tissue for 175 colon cancer cases who were part of the Diet, Activity, and Lifestyle study, an incident, population-based, case-control study of colon cancer from Utah and the Kaiser Permanente Medical Research Program (KPMRP). Cases had tumor registry verification of a first primary adenocarcinoma of the colon and were diagnosed between October 1991 and September 1994. Tumor tissue blocks were obtained for 97% of all Utah cases and for 85% of all KPMRP cases [21] and included those who signed informed consent and those retrieved by local tumor registries and sent to study investigators without personal identifiers. Individuals with known adenomatous polyposis coli (APC), Crohn's disease, or inflammatory bowel disease were not eligible for the study. The study was approved by the Institutional Review Board of the University of Utah and at KPMRP and all study participants signed informed consent.

Diet and lifestyle data

Data were collected by trained and certified interviewers using laptop computers. All interviews were audio-taped as previously described and reviewed for quality control purposes [22]. The referent period for the study was two years prior to diagnosis for cases and selection for controls. Dietary information was obtained for the referent year from an extensive diet history questionnaire adapted from the validated CARDIA diet history [23]. As part of the study questionnaire, information was collected on regular use and current use of aspirin and non-steroidal anti-inflammatory drugs and cigarette smoking history including start and stop dates for smoking. Measured height at the time of interview and self-reported weight from two years prior to diagnosis were used to calculate body mass index (kg/m^2).

TagSNPs and genetic assessment

TagSNPs were selected using the following parameters: r^2=0.8 defined LD blocks using a Caucasian LD map, minor allele frequency (MAF) >0.1, range =-1500 bps from the initiation codon to +1500 bps from the termination codon, and 1 SNP/LD bin. All markers were genotyped using a custom multiplexed bead array assay format based on Golden Gate chemistry (Illumina, San Diego, California). A genotyping call rate of 99.85% was attained. Blinded internal replicates represented 4.4% of the sample set. The duplicate concordance rate was 100.00%. Previous analysis using Adaptive Rank Truncation Product (ARTP) identified several genes as significant within the pathway [7]. We focus our genotype analysis with pathway gene expression on results of those 155 genes and their 1246 SNPs and the previously identified inheritance model (Supplemental Table 1 includes pathway genes

evaluated along with their location and full name). In our previous analysis, genes were considered significant if the Adaptive Rank Truncation Product had a gene *p* value of 0.05 or less and a SNP *p* value of 0.05 or less. When evaluating genotype and lifestyle interactions with pathway gene expression we focused on those SNPs previously identified as interacting with lifestyle factors to modify colon cancer risk (Supplemental Table 2 contains a list of SNP and lifestyle interactions tested).

RNA processing

RNA was extracted from formalin-fixed paraffin embedded tissues. We assessed slides containing thin section of colon tissue and blocks that were prepared over the duration of the study prior to the time of RNA isolation to determine their suitability. Older slides produced comparable RNA quality as more recent slides and were not correlated with time lapse between slide preparation and RNA preparation. The study pathologist reviewed slides to delineate tumor and non-tumor tissue. Cells from non-tumor colon tissue were dissected from 1-4 sequential sections on aniline blue stained slides using an H&E slide for reference. Total RNA was extracted, isolated, and purified using the RecoverAll Total Nucleic Acid isolation kit (Ambion), RNA yields were determined using a NanoDrop spectrophotometer.

Sequencing library preparation

Library construction was performed using the Illumina TruSeq Stranded Total RNA Sample Preparation Kit with Ribo-Zero. Briefly, Ribosomal RNA was removed from 100 ng total RNA using biotinylated Ribo-Zero oligos attached to magnetic beads that are complimentary to cytoplasmic rRNA. Following purification, the rRNA-depleted sample is fragmented with divalent cations under elevated temperatures and primed with random hexamers in preparation for cDNA synthesis. First strand reverse transcription is accomplished using Superscript II Reverse Transcriptase (Invitrogen). Second strand cDNA synthesis is accomplished using DNA polymerase I and Rnase H under conditions in which dUTP is substituted for dTTP, yielding blunt-ended cDNA fragments in which the second strand contains dUTP. An A-base is added to the blunt ends as a means to prepare the cDNA fragments for adapter ligation and block concatamer formation during the ligation step. Adapters containing a T-base overhang were ligated to the A-tailed DNA fragments. Ligated fragments were PCR-amplified (13 cycles) under conditions in which the PCR reaction enables amplification of the first strand cDNA product, whereas attempted amplification of the second strand product stalls at dUTP bases and therefore is not represented in the amplified library. The PCR-amplified library was purified using Agencourt AMPure XP beads (Beckman Coulter Genomics). The concentration of the amplified library was measured with a NanoDrop spectrophotometer and an aliquot of the library is resolved on an Agilent 2200 Tape Station to define the size distribution of the sequencing library.

Sequencing and data processing

Sequencing libraries (18 pM) were chemically denatured and applied to an Illumina TruSeq v3 single read flow cell using an Illumina cBot. Hybridized molecules were clonally amplified and annealed to sequencing primers with reagents from an Illumina TruSeq SR Cluster Kit v3-cBot-HS. Following transfer of the flowcell to an Illumina HiSeq instrument, a 50 cycle single-read sequence run was performed using TruSeq SBS v3 sequencing reagents. The single-end 50-base reads from the Illumina HiSeq-2500 were aligned to a sequence database containing the human genome (build GRCh37/hg19, February 2009, from genome.ucsc.edu) plus all splice junctions generated using the USeq MakeTranscriptome application (version 8.8.1, available here: http://useq.sourceforge.net/). Alignment was performed using novoalign version 2.08.01 available from novocraft.com, which also trimmed any adapter sequence. Following alignment, genome alignments to splice junctions were translated back to genomic coordinates using the US eq Sam Transcriptome Parser application. The resulting alignments were sorted and indexed using the Picard Sort Sam application (version 1.100, available here: http://broadinstitute.github.io/picard/). Aligned read counts for each gene were calculated using pysam (https://code.google.com/p/pysam/) and samtools (http://samtools.sourceforge.net/). A python script using the pysam library was given a list of the genome coordinates for each gene, and counts to the exons and UTRs of those genes were calculated. Gene coordinates were downloaded

from http://genome.ucsc.edu. We focused our analysis on differential gene expression of 236 genes related to the CHIEF pathway.

Statistical methods

Of the 197 initial tumor/non-tumor tissue pairs, 22 subjects failed quality control (QC) based on low number of sequence leaving 175 subjects with high quality expression data. Of these, 144 had questionnaire data for diet and lifestyle data, and 138 had genotype data for inclusion in the analysis. For each genetic, dietary, and lifestyle factor, our analysis centered on contrasting gene expression levels of individuals with lower intake or exposure levels to those of individuals with higher intake or exposure levels. For genotype associated we focused on the previously identified inheritance model (i.e. dominant, recessive, or additive) that was associated with colon cancer to define categories for contrast. To summarize risk associated with multiple exposures, we utilized an oxidative balance score (OBS) that consisted of 13 diet and lifestyle factors that were pro-oxidants (dietary iron and polyunsaturated fat and cigarette smoking) and anti-oxidants (vitamin C, vitamin E, selenium, beta carotene, lycopene, lutein/zeaxanthin, vitamin D, calcium, and folic acid and NSAID use) [8]. This score has been shown be associated with colon cancer [8]. To create the OBS, these diet and lifestyle factors were assigned values of 2 for low levels of exposure for each pro-oxidants or high exposure to anti-oxidants (low-risk), one for intermediate levels of exposure, and zero for high levels of exposure to pro-oxidants and low exposure to anti-oxidants (high-risk). The individual scores for the 13 variables were then combined to obtain the OBS. Higher summary score corresponded to greater oxidative balance; individual's OBSs were categorized as low, intermediate, or high based on tertiles associated with the empirical distribution of the OBSs. Additionally, we evaluated specific components of the OBS. Dietary data were evaluated using nutrients per 1000 calories and then categorized using quartiles of intake based on sex-specific distributions, combining the second and third quartiles to form the intermediate group. Cigarette smoking was categorized as never, former, or current smoker. Use of NSAIDs (which included aspirin and/or nonsteroidal anti-inflammatory drugs) was categorized as either being a recent user (i.e. using NSAIDs during the referent period) or a non-user. We evaluated BMI categorized as <25 kg/m^2, 25-30 kg/m^2, or >30 kg/m^2.

For each variable of interest (genotype, specific dietary factors, recent NSAIDS use, recent cigarette smoking, BMI, and OBS), we assessed which genes displayed statistically significant differential expression between low and high categories of gene, diet, and lifestyle factor using the Bioconductor package DESeq2 written for the R statistical programming environment. DESeq2 assumes the RNA-seq counts are distributed according to negative binomial distributions. It utilizes generalized linear modeling to test individual null hypotheses of zero log2 fold changes between high and low categories (i.e. no differential expression); for each gene it employs both an independent-filtering method and the Benjamini and Hochberg (BH) [24] procedure to improve power and control the false discovery rate (FDR). The default DESeq2 options were applied, including the replacement of outliers, as defined by Cook's distance, and the use of the Wald test. For further details regarding DESeq2, see Love et al. [25]. In identifying genes with differential expression, an FDR of 0.10 was used. To help describe the data we report the average DESeq2-adjusted gene expression levels (size factor adjusted counts) among individuals in the high and low categories of pathway genes. The fold change calculations associated with these genes was determined by DESeq2 and represents the log2 change in expression level (i.e. counts) for the high category compared to the low category.

Evaluation of gene and environment interactions on pathway gene expression also was done in DESeq2 as described above, with the exception of using the likelihood ratio test (LRT) in place of the Wald test in order to test all levels of the interaction term. This analysis focused on previously reported statistically significant interactions between genetic variants and lifestyle factors with expression levels of 236 pathway-related genes. We report mean normalized count values along with unadjusted and BH-adjusted *p* values where the FDR was set at <0.1.

Results

The mean age of the study population was 64.5 (**Table 1**). There were more men than women,

Table 1. Description of study population

		Mean	STD
Age		64.5	9.8
Oxidative balance score			
	Low	9.7	2
	High	17.7	1.7
		N	%
Sex			
	Male	80	55.60%
	Female	64	44.40%
Smoking status			
	Never	60	41.70%
	Former	65	45.10%
	Current	19	13.20%
Recent NSAID use			
	No	89	62.70%
	Yes	53	37.30%
BMI (kg/m^2)			
	<25	50	34.7%
	25-30	53	36.8%
	>30	41	28.4%

more people who never smoked than were recent smokers, and more people who did not recently use NSAIDS (aspirin and/or ibuprofen) than used NSADIs. Almost 35% had a BMI <25, 36.8% were considered overweight (BMI 25-30), and 28.4% were considered obese (BMI >30).

No pathway genes significantly altered their own gene expression (data not shown in table). However, several pathway genes altered expression of other genes in the pathway (**Table 2** shows all de-regulated genes with FDR of <0.1). Thirteen pathway genes and 17 SNPs within those genes were associated with altered expression of other genes in the pathway. The most significant associations (adjusted *p* values less than 0.05) were for *BMPR1B* rs17616243 and rs1863652 altering expression of ALDH1A1 and FGF1 respectively, *MAP3K9* rs11628333 altering expression of TYK2 and AKT2, *SMAD3* rs1498506 altering expression of NFκB1, *TERT* rs2853668 altering expression of TYK2 and IRF3, *TGFB1* rs480-3455 altering expression of IRS2, and *TNF* rs1800630 altering expression of TEK.

BMI, NSAIDs use, and dietary components of the OBS were associated with altered gene expression (**Table 3**), although neither cigarette smoking nor the summary OBS altered expression of any pathway genes with a FDR of 0.1. Most factors altered only one gene, although folic acid, vitamin D, and lutein/zeaxanthin each altered expression of two pathway genes. SOD2 expression was altered by both folic acid and vitamin C, while RUNX2 expression was altered by both folic acid and beta carotene.

Focusing on genes and their associated SNPs that have previously been shown to interact with NSAIDs, cigarette smoking, BMI, and OBS to alter colon cancer risk, we observed that several of these interactions were associated with significant differential expression in pathway genes (**Tables 4** and **5**). NSAIDs interacted with 10 genes and 15 SNPs within those genes to alter expression of 28 pathway genes (**Table 4**). Three SNPs in *KDR* altered gene expression, while two SNPs in *IRF2*, *TXNRD1*, and *TNXRD2* altered expression of multiple genes. Other groups of gene families that altered expression were interferon regulatory factors (IRF2, IRF5, and IRF6) altering SMAD6, EPX, NPY2R, MAP3K10, SOD2, PCK1, TEP1, and MYO15B. *KDR*, a *VEGFA* receptor, altered expression of CXCR2, NFAT5, AKT1, PTEN, CALM3, VDR, and DUSP4. Recent cigarette smoking interacted with seven genes and nine SNPs to alter expression of 27 genes. *JAK2* rs10815160 altered expression of 13 of these genes, included altered expression of MTOR, BMP1 and 4, SEPN1, MAP3K3, TLR2, PDGF, SOD2, and PTGIS. Expression of MAPK1 was altered by two SNPs in NOS2A and VEGFA interacting with cigarette smoking. FLT1, a receptor for VEGFA, interacted with smoking to alter expression of VEGFA.

There were fewer interactions between BMI and OBS and genes to alter gene expression than observed for either NSAIDs or cigarette smoking (**Table 5**). BMI interacted with FLT1, KDR, SEPN1, TERT, TXNRD2, and VEGFA to alter expression of eight genes. Interestingly, both KDR and SEPN1 altered expression of AR, while TXNRD2, a selenoprotein like SEPN1, altered expression of SOD2. VEGFA altered expression of IL6R. Three genes and five SNPs interacted with OBS to alter expression of 12 genes. Two of these genes, FLT1 and KDR, are VEGFA receptors, while the other gene, inducible nitric oxide synthase (NOS2A), is induced by cytokines and has been reported to play a role in oxidative stress-induced inflammation. FLT1 also interacted with OBS to alter AR expression.

Table 2. Pathway gene expression altered by other pathway genes

Gene	SNP rs# (Inheritance model)	Gene	Mean expression		Log2 fold change	*p* value	Adjusted *p* value
		Altered	Genotype 0	Genotype 1			
BMPR1A	rs7088641 (Dominant)	IL6R	57.87	42.91	-0.35	7.50E-04	9.75E-02
BMPR1B	rs17616243 (Additive)	ALDH1A1	20.66	2.55	-1.4	1.34E-04	2.68E-02
		NFAM1	10.19	0.44	-1.29	7.14E-04	7.14E-02
		DDIT4	16.21	3.76	-1.16	1.42E-03	9.49E-02
	rs1863652 (Recessive)	FGF1	1.46	6.08	0.59	6.05E-05	1.28E-02
		EIF4EBP3	7.74	19.65	0.63	9.05E-04	9.59E-02
IL6R	rs4845623 (Recessive)	AR	39.12	87.89	0.47	7.16E-04	9.36E-02
		TCF7L2	111.08	80.34	-0.35	1.22E-03	9.36E-02
	rs7549250 (Dominant)	STAT3	127.92	109.64	-0.18	7.56E-03	9.18E-02
		TCF7L2	89.02	111.63	0.25	7.65E-03	9.18E-02
IL8	rs4073 (Additive)	VEGFA	125.18	166.75	0.33	6.80E-03	8.16E-02
IRF3	rs2304204 (Recessive)	NADSYN1	124.94	99.08	-0.33	1.96E-03	9.40E-02
MAP3K3	rs11658329 (Dominant)	AR	30.76	64.29	0.4	1.35E-03	9.58E-02
MAP3K9	rs11628333 (Recessive)	TYK2	70.2	106.66	0.53	1.11E-04	7.90E-03
		AKT2	62.48	84.71	0.37	1.01E-03	3.57E-02
		TSC2	73.43	103.79	0.41	4.02E-03	9.50E-02
	rs17176971 (Recessive)	PIK3CB	43.12	21.03	-0.99	4.04E-04	6.96E-02
		AKT2	64.09	116.65	0.74	8.69E-04	6.96E-02
		NFAT5	200.65	109.9	-0.73	1.22E-03	6.96E-02
		TYK2	73.26	144.3	0.87	1.73E-03	6.96E-02
		SMAD2	104.07	57.18	-0.71	1.75E-03	6.96E-02
		MAP3K11	64.72	115.11	0.76	3.08E-03	9.97E-02
		LEPR	7.23	23.15	1.04	4.60E-03	9.97E-02
		DUSP7	6.93	20.9	1.01	4.69E-03	9.97E-02
		SMAD4	65.76	39.2	-0.74	4.77E-03	9.97E-02
		PRRX1	6.29	23.93	1.06	5.01E-03	9.97E-02
SMAD3	rs1498506 (Additive)	NFKB1	34.45	50.44	0.43	1.12E-04	2.65E-02
SMAD7	rs12953717 (Additive)	RAF1	71.17	91.1	0.3	9.88E-04	9.38E-02
STAT5B	rs6503691 (Additive)	CALM2	146.84	74.06	-0.7	9.09E-04	9.59E-02
		MAPK1	88.42	43.75	-0.59	2.02E-03	9.59E-02
		AR	40.4	130.94	0.83	3.59E-03	9.59E-02
		MAPK3	42.49	86.92	0.69	3.88E-03	9.59E-02
		PTEN	81.58	42.12	-0.64	4.06E-03	9.59E-02
TERT	rs2736100 (Recessive)	ERBB2	110.38	150.32	0.37	4.00E-03	9.58E-02
		IRF1	85.86	62.54	-0.38	6.90E-03	9.58E-02
		STAT1	135.39	93.11	-0.41	7.98E-03	9.58E-02
	rs2853668 (Additive)	TYK2	69.98	135.4	0.73	1.26E-04	1.95E-02
		IRF3	26.39	53.36	0.76	1.74E-04	1.95E-02
		PPARG	43.67	25.14	-0.72	9.25E-04	6.91E-02
		TSC2	73.63	126.51	0.62	1.50E-03	7.93E-02
		MAP3K11	61.11	97	0.56	2.04E-03	7.93E-02
		MMP2	58.24	116.28	0.74	2.12E-03	7.93E-02
		TRAF2	14.29	25.32	0.68	3.74E-03	9.98E-02
		STAT5A	16	31.23	0.63	4.13E-03	9.98E-02
		TLR1	7.4	22.42	0.73	4.36E-03	9.98E-02
		DUSP4	7.47	15.92	0.72	4.46E-03	9.98E-02

TGFB1	rs4803455 (Additive)	IRS2	51.95	84.95	0.5	3.07E-04	2.55E-02
		JUNB	62.6	90.06	0.39	2.11E-03	8.78E-02
TNF	rs1800630 (Dominant)	TEK	6.09	12.27	0.47	1.99E-04	4.70E-02

Genotype 0 refers to the homozygote common for the additive or dominant model and homozygote common/heterozygote for the recessive model. Genotype 1 refers to the homozygote variant for the additive or recessive model and heterozygote/homozygote variant for the dominant model. Log2 fold change = difference between coefficients (i.e. genotype1/2 coeff.-genotype 0 coeff.), on log2 scale.

Table 3. Associations between lifestyle factors and gene expression

Variable	Gene	Average adjusted count		Log2 fold change[1]	*P* value	Adjusted *p* value
	Altered	Low tertile	High tertile			
BMI	*MAP3K7*	42.15	28.15	-0.39	2.50E-04	5.90E-02
NSAIDs	*IL6R*	56.85	42.20	-0.34	7.90E-05	1.87E-02
Folic acid	*SOD2*	201.42	154.47	-0.32	1.45E-03	9.43E-02
	RUNX2	26.16	38.40	0.44	1.75E-04	2.27E-02
Vitamin D	*STAT1*	96.41	133.40	0.37	2.93E-03	8.59E-02
	IRF1	67.82	86.66	0.32	4.77E-03	8.59E-02
Beta carotene	*RUNX2*	23.00	33.68	0.41	8.05E-05	1.90E-02
Lutein	*MAPK14*	72.92	88.79	0.26	4.10E-03	9.84E-02
	JUNB	97.22	77.05	-0.31	3.92E-03	9.84E-02
Vitamin C	*SOD2*	185.23	149.01	-0.27	5.29E-03	6.35E-02

[1]Log2 fold Change *p* value from DESeq 2 is the difference between coefficients (i.e. high tertile coeff.-low tertile coeff.) on log2 scale.

Discussion

Our assessment of the CHIEF pathway genes and diet and lifestyle factors related to oxidative stress and inflammation supported our hypothesis that genes in the pathway influence genes expression of pathway genes and interact with lifestyle factors to alter gene expression of pathway genes. These findings add to the body of literature that recognize the importance of including lifestyle factors to more fully understand how genetic factors contribute to risk. The major changes in gene expression came from the combination of genetic and lifestyle factors rather than either genetic or lifestyle factor independently.

The genetic variants that we identified as altering gene expression did not alter their own gene expression, but altered expression of other genes in the pathway. The TGFβ-signaling pathway has repeatedly been associated with colon cancer [26, 27]. In our data of the 13 genes influencing expression of other genes, five were in the TGFβ-signaling pathway, BMPR1A, BMPR1B, SMAD3, SMAD7, and TGFB1. Several studies suggest the importance of the BMP receptors, given that BMPs signal through their type I and II receptors 28. BMPR1A and BMPR1B are the two best characterized type I receptors. Substrates for these receptors include Smad proteins that play a central role in BMP signaling. Smad7 also is involved in inflammation-related pathways and has been shown to modulate TGF-β and wnt-signaling [29]. Genetic variation in the Smad7 gene on 8q21 has been identified through numerous genome-wide association studies (GWAS) as being associated with colorectal cancer [30]. *SMAD3* rs1498506 previously associated with colon cancer after multiple comparison adjustment, was associated with NFκB1 expression; *SMAD7* rs12953717, which we have previously identified as influencing colon cancer and has been identified through numerous genome-wide association studies (GWAS) as being associated with colorectal cancer (CRC) [30], altered expression of RAF1, a member of the MAP3K family. We previously reported that *TGFB1* rs4803455 altered colon cancer risk by interacting with SMAD2 [15]. Two MAP3K family members, MAP3K3 and MAP3K9, altered expression of several genes, with MAP3K9 altering expression of 13 genes. TERT also altered expression of 13 genes, including ERBB2, STAT1, PPARG, TSC2, STAT5A, TLR1, and MAP3K11. One of the TERT SNPs that we found to influence expression levels of 10 genes, was rs2853668, which associated with colon cancer risk in our data and was identified in a genome-wide association study (GWAS) to be significantly associated with colorectal cancer risk [31]. These findings provide support for previously identified associations.

Table 4. Interactions between recent aspirin/NSAID use and smoking with pathway genes on expression of genes in pathway

Gene SNP	Expression gene	Means of normalized counts						*P* value	Adjusted *p* value
		No recent aspirin/NSAID use			Recent aspirin/NSAID use				
		Homozygote common	Heterozygote	Homozygote variant	Homozygote common	Heterozygote	Homozygote variant		
*EPX*_rs10853004	BMPR1A	10.17	7.84	0.00	6.27	6.39	6.91	5.43E-06	1.28E-03
	PTGS1	18.68	13.60	3.67	15.56	12.93	21.62	1.83E-04	2.15E-02
	MMP2	70.90	54.11	13.08	70.80	81.80	83.79	7.55E-04	5.94E-02
*IL15*_rs17461269_Rec	IRF2	21.13		12.71	17.44		33.25	1.10E-04	2.59E-02
*IRF2*_rs3756093_Dom	SMAD6	7.62	14.58		12.11	5.65		4.58E-04	9.72E-02
*IRF2*_rs9684244	EPX	1.01	0.17	0.61	0.25	2.44	0.00	2.74E-04	3.89E-02
	NPY2R	0.95	0.09	1.09	0.01	0.84	0.00	3.30E-04	3.89E-02
*IRF5*_rs1874328	MAP3K10	11.26	6.26	3.98	4.89	14.39	8.66	3.25E-04	7.67E-02
*IRF6*_rs2013196_Dom	SOD2	160.71	131.18		138.05	179.08		9.76E-04	1.17E-02
	PCK1	137.18	112.33		99.75	182.62		9.37E-03	4.39E-02
	TEP1	151.71	162.52		166.69	120.52		1.10E-02	4.39E-02
	CALM2	143.26	123.86		146.08	183.63		2.26E-02	5.97E-02
	MYO15B	224.71	250.99		316.52	227.45		2.49E-02	5.97E-02
*KDR*_rs12502008	CXCR2	1.28	0.38	1.33	0.01	3.28	0.62	3.72E-04	8.77E-02
*KDR*_rs12505758	NFAT5	217.15	163.37	277.80	174.75	225.63	224.75	1.23E-04	7.27E-03
	AKT1	63.07	83.02	50.67	80.30	62.72	169.60	3.63E-04	1.07E-02
	PTEN	85.89	71.04	59.42	66.62	86.68	110.31	6.87E-03	8.41E-02
	CALM3	56.09	65.32	15.69	68.04	53.95	66.18	6.94E-03	8.41E-02
	VDR	65.99	66.13	31.90	67.61	67.89	108.93	7.13E-03	8.41E-02
*KDR*_rs2071559	DUSP4	6.37	11.62	8.42	42.94	7.94	11.78	2.61E-05	6.16E-03
*SMAD3*_rs7173811	MAP3K11	78.03	58.13	61.42	51.75	71.07	64.89	1.36E-03	9.62E-02
*TERT*_rs2853668_Rec	TEP1	155.26		137.59	144.04		276.95	3.57E-03	9.24E-02
	RAF1	76.37		100.04	73.28		50.24	6.52E-03	9.24E-02
	CALM2	144.25		135.69	146.86		48.94	7.35E-03	9.24E-02
	HIF1A	70.84		86.23	68.33		29.58	7.70E-03	9.24E-02
*TXNRD1*_rs4523760_Dom	TYK2	68.29	63.29		69.72	118.63		3.94E-04	9.30E-02
*TXNRD1*_rs4964778_Dom	MAP3K10	7.91	6.72		8.37	21.96		8.74E-04	9.80E-02
	RNF146	14.77	20.44		15.95	8.39		1.11E-03	9.80E-02
*TXNRD2*_rs3788314	DUSP4	7.98	9.42	9.76	40.06	8.09	8.39	6.12E-04	9.36E-02
*TXNRD2*_rs756661	GC	0.15	2.74	0.00	0.08	0.07	2.43	3.11E-04	7.34E-02

		Non smoker/non recent smoker			Recent smoker				
*FLT1*_rs17086609	KLF6	229.24	239.11	221.55	308.38	177.89	131.96	9.97E-03	7.59E-02
	SOD2	148.05	147.92	138.79	210.38	136.58	251.93	1.99E-02	7.59E-02
	MYO15B	275.77	227.96	219.43	215.22	345.29	227.94	2.26E-02	7.59E-02
	VEGFA	134.10	138.03	120.01	211.52	123.18	149.96	2.53E-02	7.59E-02
*JAK2*_rs10815160_Rec	MTOR	88.98		191.21	102.96		66.33	6.57E-05	1.31E-02
	BMP4	10.15		22.21	11.79		1.47	3.01E-04	3.01E-02
	SEPN1	45.62		62.67	46.34		19.27	9.51E-04	3.68E-02
	ACVR1	13.75		22.36	14.09		4.53	9.86E-04	3.68E-02
	TEK	8.62		10.10	5.36		0.00	1.04E-03	3.68E-02
	MAP3K3	24.18		30.74	28.39		8.78	1.10E-03	3.68E-02
	TLR4	18.33		10.04	21.24		3.28	2.69E-03	7.69E-02
	PDGFB	5.95		8.17	6.80		0.80	3.87E-03	8.41E-02
	PTGIS	4.54		9.98	3.89		0.00	3.92E-03	8.41E-02
	SOD2	149.71		92.69	168.85		224.31	4.21E-03	8.41E-02
	MMP2	63.81		109.97	86.44		32.06	5.75E-03	9.87E-02
	NADSYN1	121.43		172.70	118.70		100.38	6.07E-03	9.87E-02
	BMP1	27.57		59.13	33.62		21.48	6.41E-03	9.87E-02
*MAP3K11*_rs7116712_Rec	TLR3	6.66		9.41	7.17		0.00	2.35E-04	5.55E-02
*NFAM1*_rs13055337_Dom	IL6R	44.87		59.81	65.26		38.78	1.06E-03	8.84E-02
*NOS2A*_rs2274894	MYO15B	252.75	249.64	251.49	424.95	234.31	126.57	2.21E-03	7.94E-02
*NOS2A*_rs3729508_Rec	DUSP6	37.98		20.44	31.95		58.25	3.01E-05	7.07E-03
	MAPK1	86.36		62.87	81.36		123.68	1.53E-04	1.80E-02
*NOS2A*_rs944725	MAPK1	79.50	87.45	82.33	117.18	74.90	51.83	4.58E-04	7.00E-02
	RPS6KA2	28.15	25.33	28.98	20.39	42.41	18.08	1.80E-03	9.71E-02
	PTGS1	17.50	13.84	14.55	8.80	28.80	35.35	1.90E-03	9.71E-02
*TLR2*_rs3804099_Rec	DUSP1	71.70		48.59	48.11		99.72	1.94E-03	9.30E-02
*VEGFA*_rs3025033	CALM2	145.27	146.44	0.00	160.76	131.48	143.56	2.97E-04	1.75E-02
	MAPK1	86.43	80.47	0.00	95.45	93.55	96.92	1.55E-03	4.56E-02
	DUSP1	69.86	62.91	0.00	42.41	80.78	50.29	4.29E-03	8.43E-02

Table 5. Interactions between BMI and OBS and pathway genes on expression of genes in pathway

		Means of normalized counts										
		Normal (BMI <25)			Overweight (BMI 25 to <30)			Obese (BMI ≥30)				
	Expressed	Homozygote		Homozygote			Homozygote			Homozygote		
Gene SNP	Gene	Common	Heterozygote	Variant	Common	Heterozygote	Variant	Common	Heterozygote	Variant	*P* value	Adjusted *p* value
*FLT1*_rs3936415	NRG2	0.27	5.91	0.71	0.92	0.96	2.69	1.25	0.28	1.81	1.30E-04	3.07E-02
	MMP9	2.77	6.69	2.20	2.32	4.12	6.72	5.46	1.54	17.18	4.62E-04	5.45E-02
*KDR*_rs2219471_Dom	AR	21.70	47.64		50.87	27.88		93.02	31.83		8.89E-04	9.42E-02
*SEPN1*_rs11247735	AR	45.74	35.82	25.08	17.82	38.66	67.67	9.33	84.12	119.79	1.43E-04	3.37E-02
	TERT	8.72	0.86	2.50	1.08	1.78	2.03	0.14	1.03	0.95	5.27E-04	6.22E-02
*TERT*_rs10069690	PCK1	163.78	88.71	125.30	112.00	136.08	136.84	102.53	171.90	74.87	3.97E-03	9.52E-02
TERT_rs2242652_Dom	MMP9	6.20	1.64		4.06	4.20		3.18	7.71		4.53E-04	9.60E-02
*TXNRD2*_rs1044732_Dom	SOD2	145.68	197.59		152.60	153.28		154.07	138.55		6.91E-03	8.30E-02
*VEGFA*_rs25648_Rec	IL6R	52.55		112.12	48.57		40.46	51.10		11.01	7.18E-04	9.33E-02
		Low oxidative balance score			Intermediate oxidative balance score			High oxidative balance score				
*FLT1*_rs12429309_Dom	AR	63.42	23.92		71.81	26.58		27.77	73.42		9.33E-04	9.89E-02
*KDR*_rs1531290	SDR16C5	2.16	1.91	17.58	0.75	1.22	2.48	3.62	1.82	0.15	8.56E-05	2.02E-02
*KDR*_rs2305948_Dom	FGF1	0.78	3.92		1.11	2.08		3.46	0.12		1.24E-04	1.15E-02
	PIK3CG	12.40	30.55		23.17	17.26		19.50	8.93		1.42E-04	1.15E-02
	MAP3K1	36.61	63.92		43.50	45.74		47.40	30.29		1.46E-04	1.15E-02
	TYK2	72.83	116.78		78.62	62.03		76.74	45.71		2.06E-04	1.22E-02
	TGFB2	1.35	8.86		2.26	1.64		2.68	4.20		3.91E-04	1.85E-02
	TLR3	5.69	15.19		8.06	5.35		7.36	2.39		1.83E-03	7.19E-02
	MMP2	56.10	108.50		65.09	74.35		73.00	46.60		2.24E-03	7.54E-02
	DHCR7	5.54	12.53		6.29	7.92		9.03	3.93		3.00E-03	8.85E-02
*NOS2A*_rs4795067	KLF6	231.23	204.66	303.46	188.84	236.28	211.24	276.26	251.45	128.71	8.10E-03	9.72E-02
*NOS2A*_rs8072199_Rec	MMP2	59.82		109.87	73.42		32.12	74.12		33.76	5.97E-04	9.85E-02

There have been few studies that have evaluated diet and lifestyle factors with gene expression [32-34]. Challenges of evaluating diet and other lifestyle factors with gene expression stem from the fact that the gene expression profile is time dependent and therefore relevant to current exposure. Exposure from current diet and lifestyle factors may not be recent enough to maintain an altered gene expression even if they do have an effect. Thus, although we assessed current smokers, current NSAID users, and diet close to the time of diagnosis, when the tissue would have been biopsied, the exposure may have been too far removed to impact expression. This could explain why few genes were de-regulated for NSAIDs specifically or for dietary factors and the OBS which was comprised mainly of dietary factors. BMI, which is a more constant exposure, was only associated with MAP3K7 expression. Likewise, nutrients evaluated were only associated with one or two genes when applying a FDR of <0.1.

Our previous analyses have shown that multiple lifestyle factors interact with pathway genes to influence colon cancer risk. Important lifestyle factors previously identified as interacting with genes in the pathway are diet and MAPK [9, 10], angiogenesis [8], and TLR [11] genes; NSAIDs use with angiogenesis [8], estrogen-related genes [18], cytokines [12, 13], MAPK [9], JAK/STAT [14], TGFβ-signaling pathway [15], TLR [11], TERT [35], and selenoproteins [16]; cigarette smoking with angiogenesis [8], cytokines [13, 17], MAPK [9], JAK/STAT [14], TGFβ [15], and selenoproteins [16] sub-pathways; BMI with angiogenesis [8], estrogen-related genes [18], VDR [19], TERT [35], and cytokines [13, 17]. In these analyses, we have shown that many of the previously identified interactions alter gene expression in the CHIEF pathway. We demonstrated that many of the interactions between specific SNPs and lifestyle factors altered gene expression of important pathway genes, lending support for the previously identified associations. The SNPs within these genes that we previously reported significant interactions with, are also the SNPs previously identified as interacting with lifestyle factors. We identified altered gene expression associated with 50% of the previously identified interactions for NSAIDS and OBS, 29% of the previously identified interactions with smoking, and 35% of those identified with BMI. For instance, with NSAIDs we detected significant associations for 2 of the 4 previously identified IRF2 SNPs that altered colon cancer risk; we observed significant differential gene expression with the only significant interaction between NSAIDs and IRF5, and for one of the two IRF6 SNPs. For FLT1, KDR, and VEGF, we observed significant interaction for all previously identified SNP interacting with NSAID use, cigarette smoking, and BMI. Likewise for TERT we saw differential gene expression associated with the previously identified interaction between rs2853668 and NSAIDs use and *TERT* rs10069690 and rs2242652 with BMI. For TXNRD1 and TXNRD2 we were able to replicate significant findings with NSAIDs for four of the previously identified six SNPs that interacted with NSAIDs to alter colon cancer risk.

While we have provided data that support a biological basis for previously identified interactions, our study has limitations. As we previously stated, gene expression data represent expression at a given point in time, in this case at the time of surgery to remove the colon tumor. Various variables are impacted by this limitation. Genotypes do not change over time and therefore should not vary by the timing of the biopsy, while variables such as NSAID use and diet would have variability and likely inhibit our ability to detect associations that might exist. Likewise, the associations which we did detect could have a larger change if the measurement had been closer to the time of tissue extraction. Additionally, we have utilized colonic non-tumor tissue, so genes would have to be expressed in colon tissue for detection. These diet and lifestyle factors could influence other genes in other tissue sources. Utilizing colonic tissue that is located close to a tumor could alter gene expression, however, it would alter it is a manner that was not dependent on diet or lifestyle exposure. While we did not show altered gene expression with all previously identified interactions, it should be kept in mind that altered gene expression is one aspect of functionality and failure to see an association does not mean that genes couldn't alter protein levels or other aspects of functionality we were unable to evaluate.

Our data illustrate the importance of evaluating broader pathways to determine functionality of genes and SNPs. Genes within pathways were more likely to have expression altered than were the genes themselves. The importance of evaluating the interaction between genes and

environment in terms of functionality also has been demonstrated. In many instances, the combination of factors altered expression of pathway genes that neither gene nor lifestyle factor altered alone. This is one of the first, if not the first study, to illustrate this point. However, since these findings are from only one small study, replication is needed in other similar studies.

Acknowledgements

This study was funded by NCI grants CA48998. 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. We would like to acknowledge the contributions of Dr. Bette Caan, Judy Morse and Donna Schaffer and the Kaiser Permanente Medical Research Program, and Sandra Edwards for data collection and organization, Jennifer Herrick for data management, Erica Wolff and Michael Hoffman for RNA extraction, Wade Samowitz for slide review, and Brett Milash at the Bioinformatics Core Facility at the University of Utah.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Martha L Slattery, Department of Internal Medicine, 383 Colorow, University of Utah School of Medicine, Salt Lake, USA. Tel: 801-585-6955; E-mail: Marty.slattery@hsc.utah.edu

References

[1] Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. Nucleic Acids Res 2008; 36: D820-4.

[2] Yuan HY, Chiou JJ, Tseng WH, Liu CH, Liu CK, Lin YJ, Wang HH, Yao A, Chen YT, Hsu CN. FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. Nucleic Acids Res 2006; 34: W635-41.

[3] Penney RB, Lundgreen A, Yao-Borengasser A, Koroth-Edavana V, Williams S, Wolff R, Slattery ML, Kadlubar S. Lack of correlation between in silico projection of function and quantitative real-time PCR-determined gene expression levels in colon tissue. Pharmacogenomics pers med 2013; 6: 99-103.

[4] Slattery ML, Lundgreen A, Mullany LE, Penney RB, Wolff RK. Influence of CHIEF pathway genes on gene expression: a pathway approach to functionality. Int J Mol Epidemiol Genet 2014; 5: 100-11.

[5] Gonzalez-Moreno O, Calvo A, Joshi BH, Abasolo I, Leland P, Wang Z, Montuenga L, Puri RK, Green JE. Gene expression profiling identifies IL-13 receptor alpha 2 chain as a therapeutic target in prostate tumor cells overexpressing adrenomedullin. Int J Cancer 2005; 114: 870-8.

[6] Hoffstedt J, Eriksson P, Mottagui-Tabar S, Arner P. A polymorphism in the leptin promoter region (-2548 G/A) influences gene expression and adipose tissue secretion of leptin. Horm Metab Res 2002; 34: 355-9.

[7] Slattery ML, Wolff RK, Lundgreen A. A pathway approach to evaluating the association between the CHIEF pathway and risk of colorectal cancer. Carcinogenesis 2015; 36: 49-59.

[8] Slattery ML, Lundgreen A, Welbourn B, Wolff RK, Corcoran C. Oxidative balance and colon and rectal cancer: Interaction of lifestyle factors and genes. Mutat Res 2012; 734: 30-40.

[9] Slattery ML, Lundgreen A, Wolff RK. MAP kinase genes and colon and rectal cancer. Carcinogenesis 2012; 33: 2398-408.

[10] Slattery ML, Lundgreen A, Wolff RK. Dietary Influence on MAPK-Signaling Pathways and Risk of Colon and Rectal Cancer. Nutr Cancer 2013; 65: 729-38.

[11] Slattery ML, Herrick JS, Bondurant KL, Wolff RK. Toll-like receptor genes and their association with colon and rectal cancer development and prognosis. Int J Cancer 2012; 130: 2974-80.

[12] Slattery ML, Lundgreen A, Bondurant KL, Wolff RK. Tumor necrosis factor-related genes and colon and rectal cancer. Int J Mol Epidemiol Genet 2011; 2: 328-38.

[13] Bondurant KL, Lundgreen A, Herrick JS, Kadlubar S, Wolff RK, Slattery ML. Interleukin genes and associations with colon and rectal cancer risk and overall survival. Int J Cancer 2013; 132: 905-15.

[14] Slattery ML, Lundgreen A, Kadlubar SA, Bondurant KL, Wolff RK. JAK/STAT/SOCS-signaling pathway and colon and rectal cancer. Mol Carcinog 2013; 52: 155-66.

[15] Slattery ML, Lundgreen A, Wolff RK, Herrick JS, Caan BJ. Genetic variation in the transforming growth factor-beta-signaling pathway, lifestyle factors, and risk of colon or rectal cancer. Dis Colon Rectum 2012; 55: 532-40.

[16] Slattery ML, Lundgreen A, Welbourn B, Corcoran C, Wolff RK. Genetic variation in selenoprotein genes, lifestyle, and risk of colon and rectal cancer. PLoS One 2012; 7: e37312.

[17] Slattery ML, Lundgreen A, Bondurant KL, Wolff RK. Interferon-signaling pathway: associations with colon and rectal cancer risk and subsequent survival. Carcinogenesis 2011; 32: 1660-7.
[18] Slattery ML, Lundgreen A, Herrick JS, Kadlubar S, Caan BJ, Potter JD, Wolff RK. Variation in the CYP19A1 gene and risk of colon and rectal cancer. Cancer Causes Control 2011; 22: 955-63.
[19] Slattery ML, Murtaugh M, Caan B, Ma KN, Wolff R, Samowitz W. Associations between BMI, energy intake, energy expenditure, VDR genotype and colon and rectal cancers (United States). Cancer Causes Control 2004; 15: 863-72.
[20] Slattery ML, Fitzpatrick FA. Convergence of hormones, inflammation, and energy-related factors: a novel pathway of cancer etiology. Cancer Prev Res (Phila) 2009; 2: 922-30.
[21] Slattery ML, Edwards SL, Palmer L, Curtin K, Morse J, Anderson K, Samowitz W. Use of archival tissue in epidemiologic studies: collection procedures and assessment of potential sources of bias. Mutat Res 2000; 432: 7-14.
[22] Edwards S, Slattery ML, Mori M, Berry TD, Caan BJ, Palmer P, Potter JD. Objective system for interviewer performance evaluation for use in epidemiologic studies. Am J Epidemiol 1994; 140: 1020-8.
[23] Liu K, Slattery M, Jacobs D Jr, Cutter G, McDonald A, Van Horn L, Hilner JE, 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.
[24] Benjamini YaH Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 1995; 57: 289-300.
[25] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15: 550.
[26] Slattery ML, Herrick J, Lundgreen A, Wolff RK. Genetic variation in the TGF{beta}-signaling pathway and colon and rectal cancer risk. Cancer Epidemiol Biomarkers Prev 2011; 20: 57-69.
[27] Slattery ML, Lundgreen A, Herrick JS, Kadlubar S, Caan BJ, Potter JD, Wolff RK. Genetic variation in bone morphogenetic protein and colon and rectal cancer. International journal of cancer. J Int Du Cancer 2012; 130: 653-64.
[28] Chen D, Zhao M, Mundy GR. Bone morphogenetic proteins. Growth Factors 2004; 22: 233-41.
[29] ten Dijke P, Hill CS. New insights into TGF-beta-Smad signalling. Trends in biochemical sciences 2004; 29: 265-73.
[30] Broderick P, Carvajal-Carmona L, Pittman AM, Webb E, Howarth K, Rowan A, Lubbe S, Spain S, Sullivan K, Fielding S, Jaeger E, Vijayakrishnan J, Kemp Z, Gorman M, Chandler I, Papaemmanuil E, Penegar S, Wood W, Sellick G, Qureshi M, Teixeira A, Domingo E, Barclay E, Martin L, Sieber O; CORGI Consortium, Kerr D, Gray R, Peto J, Cazier JB, Tomlinson I, Houlston RS. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. Nat Genet 2007; 39: 1315-7.
[31] Peters U, Hutter CM, Hsu L, Schumacher FR, Conti DV, Carlson CS, Edlund CK, Haile RW, Gallinger S, Zanke BW, Lemire M, Rangrej J, Vijayaraghavan R, Chan AT, Hazra A, Hunter DJ, Ma J, Fuchs CS, Giovannucci EL, Kraft P, Liu Y, Chen L, Jiao S, Makar KW, Taverna D, Gruber SB, Rennert G, Moreno V, Ulrich CM, Woods MO, Green RC, Parfrey PS, Prentice RL, Kooperberg C, Jackson RD, Lacroix AZ, Caan BJ, Hayes RB, Berndt SI, Chanock SJ, Schoen RE, Chang-Claude J, Hoffmeister M, Brenner H, Frank B, Bézieau S, Küry S, Slattery ML, Hopper JL, Jenkins MA, Le Marchand L, Lindor NM, Newcomb PA, Seminara D, Hudson TJ, Duggan DJ, Potter JD, Casey G. Meta-analysis of new genome-wide association studies of colorectal cancer risk. Hum Genet 2011; 131: 217-34.
[32] Hu R, Kong AN. Activation of MAP kinases, apoptosis and nutrigenomics of gene expression elicited by dietary cancer-prevention compounds. Nutrition 2004; 20: 83-8.
[33] Ornish D, Magbanua MJ, Weidner G, Weinberg V, Kemp C, Green C, Mattie MD, Marlin R, Simko J, Shinohara K, Haqq CM, Carroll PR. Changes in prostate gene expression in men undergoing an intensive nutrition and lifestyle intervention. Proc Natl Acad Sci U S A 2008; 105: 8369-74.
[34] Landi MT, Dracheva T, Rotunno M, Figueroa JD, Liu H, Dasgupta A, Mann FE, Fukuoka J, Hames M, Bergen AW, Murphy SE, Yang P, Pesatori AC, Consonni D, Bertazzi PA, Wacholder S, Shih JH, Caporaso NE, Jen J. Gene expression signature of cigarette smoking and its role in lung adenocarcinoma development and survival. PLoS One 2008; 3: e1651.
[35] Pellatt AJ, Wolff RK, Herrick J, Lundgreen A, Slattery ML. TERT's role in colorectal carcinogenesis. Mol Carcinog 2013; 52: 507-13.

Supplemental Table 1. List of genes, aliases, and chromosomal location

Official gene name	Common aliases	Chromosome
AKT1 (V-AKT MURINE THYMOMA VIRAL ONCOGENE HOMOLOG 1)	AKT, MGC99656, PKB, PRKBA, RAC, RAC-ALPHA	14q32.32
AR (ANDROGEN RECEPTOR)	DHTR, NR3C4	Xq12
BMP1 (BONE MORPHOGENETIC PROTEIN 1)	FLJ44432, PCOLC, PCP, TLD	8p21
BMP2 (BONE MORPHOGENETIC PROTEIN 2)	BMP2A	20p12
BMP4 (BONE MORPHOGENETIC PROTEIN 4)	BMP2B, BMP2B1, ZYME	14q22-q23
BMPR1A (BONE MORPHOGENETIC PROTEIN RECEPTOR, TYPE IA)	ACVRLK3, ALK3, CD292	10q22.3
BMPR1B (BONE MORPHOGENETIC PROTEIN RECEPTOR, TYPE IB)	ALK-6, ALK6, CDw293	4q22-q24
BMPR2 (BONE MORPHOGENETIC PROTEIN RECEPTOR, TYPE II)	BMPR-II, BMPR3, BMR2	2q33-q34
C11orf31 (CHROMOSOME 11 OPEN READING FRAME 31)	C17orf10, SELH	11q12.1
CYP19A1 (CYTOCHROME P450, FAMILY 19, SUBFAMILY A, POLYPEPTIDE 1)	AROMATASE, CYP19	15q21.1
DUSP1 (DUAL SPECIFICITY PHOSPHATASE 1)	MKP-1, MKP1, PTPN10	5q34
DUSP2 (DUAL SPECIFICITY PHOSPHATASE 2)	PAC-1, PAC1	2q11
DUSP4 (DUAL SPECIFICITY PHOSPHATASE 4)	HVH2, MKP-2, MKP2, TYP	8p12-p11
DUSP6 (DUAL SPECIFICITY PHOSPHATASE 6)	MKP3, PYST1	12q22-q23
DUSP7 (DUAL SPECIFICITY PHOSPHATASE 7)	MKP-X, MKPX, PYST2	3p21
EGFR (EPIDERMAL GROWTH FACTOR RECEPTOR)	ERBB, ERBB1, HER1	7p12
EGR2 (EARLY GROWTH RESPONSE 2)	CMT1D, CMT4E, KROX20	10q21.1
EIF4E (EUKARYOTIC TRANSLATION INITIATION FACTOR 4E)	CBP, EIF4E1, EIF4EL1, EIF4F	4q21-q25
EIF4EBP2 (EUKARYOTIC TRANSLATION INITIATION FACTOR 4E BINDING PROTEIN 2)	4EBP2	10q21-q22
EIF4EBP3 (EUKARYOTIC TRANSLATION INITIATION FACTOR 4E BINDING PROTEIN 3)	4E-BP3	5q31.3
EPX (EOSINOPHIL PEROXIDASE)	EPO, EPP, EPX-PEN	17q23.1
ESR1 (ESTROGEN RECEPTOR 1)	ESR, ER, ESRA	6q21.5
ESR2 (ESTROGEN RECEPTOR 2)	ESRB, ESR-Beta, ER-Beta	14q23.2-q23.3
FLT1 (FMS-RELATED TYROSINE KINASE 1)	FLT, VEGFR1	13q12
GDF10 (GROWTH DIFFERENTIATION FACTOR 10)	BMP-3b, BMP3B	10q11.22
HIF1A (HYPOXIA-INDUCIBLE FACTOR 1, ALPHA SUBUNIT)	HIF-1alpha, HIF1, HIF1-ALPHA, MOP1	14q21-q24
IFNG (INTERFERON, GAMMA)	IFG, IFI	12q14
IFNGR1 (INTERFERON GAMMA RECEPTOR 1)	CD119, FLJ45734, IFNGR	6q23-q24
IFNGR2 (INTERFERON GAMMA RECEPTOR 2)	AF-1, IFGR2, IFNGT1	21q22.11
IGF1 (INSULIN-LIKE GROWTH FACTOR I)	IGF I	12q23.2
IGF1R (INSULIN-LIKE GROWTH FACTOR RECEPTOR I)		15q26.3
IGFPB3 (INSULIN-LIKE GROWTH FACTOR-BINDING PROTEIN 3)	IBP3	7p12.3
IKBKB (INHIBITOR OF KAPPA LIGHT POLYPEPTIDE GENE ENHANCER IN B-CELLS, KINASE BETA)	FLJ40509, IKK-beta, IKK2, IKKB, NFKBIKB	8p11.2
IL10 (INTERLEUKIN 10)	CSIF, IL-10, IL10A, TGIF	1q31-q32
IL15 (INTERLEUKIN 15)	IL-15, MGC9721	4q31
IL17A (INTERLEUKIN 17A)	CTLA8, IL-17, IL-17A, IL17	6p12
IL1A (INTERLEUKIN 1, ALPHA)	IL-1A, IL1, IL1-ALPHA, IL1F1	2q14
IL1B (INTERLEUKIN 1, BETA)	IL-1, IL1-BETA, IL1F2	2q14
IL1RN (INTERLEUKIN 1 RECEPTOR ANTAGONIST)	ICIL-1RA, IL-1ra3, IL1F3, IL1RA, IRAP	2q14.2
IL2 (INTERLEUKIN 2)	IL-2, TCGF, lymphokine	4q26-q27
IL23R (INTERLEUKIN 23 RECEPTOR)		1p31.3

IL2RA (INTERLEUKIN 2 RECEPTOR, ALPHA)	CD25, IDDM10, IL2R, TCGFR	10p15-p14
IL3 (INTERLEUKIN 3)	IL-3, MULTI-CSF	5q31.1
IL4 (INTERLEUKIN 4)	BSF1, IL-4	5q31.1
IL6 (INTERLEUKIN 6 (INTERFERON, BETA 2))	BSF2, HGF, HSF, IFNB2, IL-6	7p21
IL6R (INTERLEUKIN 6 RECEPTOR)	CD126, IL-6R-1, IL-6R-alpha, IL6RA	1q21
IL8 (INTERLEUKIN 8)	AMCF-I, CXCL8, GCP-1, GCP1, K60, NAF, NAP1	4q13-q21
IL8RA (INTERLEUKIN 8 RECEPTOR, ALPHA)	CD128, CD181, CDw128a, CKR-1, IL8R1, IL8RBA	2q35
IL8RB (INTERLEUKIN 8 RECEPTOR, BETA)	CD182, CDw128b, CXCR2, IL8R2, IL8RA	2q35
IRF1 (INTERFERON REGULATORY FACTOR 1)	IRF-1, MAR	5q31.1
IRF2 (INTERFERON REGULATORY FACTOR 2)	DKFZp686F0244, IRF-2	4q34.1-q35.1
IRF3 (INTERFERON REGULATORY FACTOR 3)		19q13.3-q13.4
IRF4 (INTERFERON REGULATORY FACTOR 4)	LSIRF, MUM1	6p25-p23
IRF5 (INTERFERON REGULATORY FACTOR 5)		7q32
IRF6 (INTERFERON REGULATORY FACTOR 6)	LPS, OFC6, PIT, PPS, VWS	1q32.3-q41
IRF7 (INTERFERON REGULATORY FACTOR 7)	IRF-7H, IRF7A	11p15.5
IRF8 (INTERFERON REGULATORY FACTOR 8)	H-ICSBP, ICSBP, ICSBP1, IRF-8	16q24.1
IRF9 (INTERFERON REGULATORY FACTOR 9)	IRF-9, ISGF3, ISGF3G, p48	14q11.2
IRGM (IMMUNITY-RELATED GTPASE FAMILY, M)	IFI1, IRGM1, LRG-47, LRG47	5q33.1
IRS1 (INSULIN RECEPTOR SUBSTRATE 1)		2q36.3
IRS2 (INSULIN RECEPTOR SUBSTRATE 2)		13q34
JAK1 (JANUS KINASE 1 (A PROTEIN TYROSINE KINASE))	JAK1A, JAK1B	1p32.3-p31.3
JAK2 (JANUS KINASE 2 (A PROTEIN TYROSINE KINASE))		9p24
JUN (JUN ONCOGENE)	AP1, c-Jun	1p32-p31
JUNB (JUN B PROTO-ONCOGENE)		19p13.2
KDR (KINASE INSERT DOMAIN RECEPTOR)	CD309, FLK1, VEGFR, VEGFR2	4q11-q12
MAP2K1 (MITOGEN-ACTIVATED PROTEIN KINASE KINASE 1)	MAPKK1, MEK1, MKK1, PRKMK1	15q22.1-q22.33
MAP3K1 (MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 1)	MAPKKK1, MEKK, MEKK1	5q11.2
MAP3K10 (MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 10)	MLK2, MST	19q13.2
MAP3K11 (MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 11)	MGC17114, MLK-3, MLK3, PTK1, SPRK	11q13.1-q13.3
MAP3K2 (MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 2)	MEKK2, MEKK2B	2q14.3
MAP3K3 (MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 3)	MAPKKK3, MEKK3	17q23.3
MAP3K7 (MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 7)	TAK1, TGF1a	6q16.1-q16.3
MAP3K9 (MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 9)	MLK1, PRKE1	14q24.3-q31
MAPK1 (MITOGEN-ACTIVATED PROTEIN KINASE 1)	ERK, ERK2, ERT1, MAPK2, P42MAPK, PRKM1, PRKM2, p38p40, p41, p41mapk	22q11.21
MAPK12 (MITOGEN-ACTIVATED PROTEIN KINASE 12)	ERK3, ERK6, P38GAMMA, PRKM12, SAPK-3, SAPK3	22q13.33
MAPK14 (MITOGEN-ACTIVATED PROTEIN KINASE 14)	CSBP1, CSBP2, CSPB1, PRKM14, PRKM15RK, SAPK2A, p38, p38ALPHA	6p21.3-p21.2
MAPK3 (MITOGEN-ACTIVATED PROTEIN KINASE 3)	ERK1, P44ERK1, P44MAPK, PRKM3	16p11.2
MAPK8 (MITOGEN-ACTIVATED PROTEIN KINASE 8)	JNK, JNK1, JNK1A2, JNK21B1/2, PRKM8, SAPK1	10q11.22
MMP1 (MATRIX METALLOPEPTIDASE 1)	CLG, CLGN	11q22.3
MMP3 (MATRIX METALLOPEPTIDASE 3)	CHDS6, MMP-3, STMY, STMY1STR1	11q22.3
MMP7 (MATRIX METALLOPEPTIDASE 7)	MMP-7, MPSL1, PUMP-1	11q21-q22
MMP9 (MATRIX METALLOPEPTIDASE 9)	CLG4B, GELB, MANDP2, MMP-9	20q11.2-q13.1
MPO (MYELOPEROXIDASE)		17q23.1

MSTN (MYOSTATIN)	GDF8	2q32.2
MTOR (MECHANISTIC TARGET OF RAPAMYCIN)	FRAP, FRAP1, FRAP2, RAFT1, RAPT1	1p36.22
NFAM1 (NFAT ACTIVATING PROTEIN WITH ITAM MOTIF 1)	CNAIP, FLJ40652	22q13.2
NFAT5 (NUCLEAR FACTOR OF ACTIVATED T-CELLS 5)	NF-AT5, NFATL1, NFATZ	16q22.1
NFKB1 (NUCLEAR FACTOR OF KAPPA LIGHT POLYPEPTIDE GENE ENHANCER IN B-CELLS 1 (P105))	EBP-1, KBF1, MGC54151, NF-kappa-B, NFKB-p105, NFKB-p50	4q24
NFKB2 (NUCLEAR FACTOR OF KAPPA LIGHT POLYPEPTIDE GENE ENHANCER IN B-CELLS 2 (P49/P100))	LYT-10, LYT10	10q24
NFKBIA (NUCLEAR FACTOR OF KAPPA LIGHT POLYPEPTIDE GENE ENHANCER IN B-CELLS INHIBITOR, ALPHA)	IKBA, MAD-3, NFKBI	14q13
NOS2A (NITRIC OXIDE SYNTHASE 2A (INDUCIBLE, HEPATOCYTES))	HEP-NOS, INOS, NOS, NOS2	17q11.2-q12
PDGFB (PLATELET-DERIVED GROWTH FACTOR BETA POLYPEPTIDE)	PDGF2, SIS	22q13.1
PDK1 (PYRUVATE DEHYDROGENASE KINASE, ISOZYME 1)		2q31.1
PDK2 (PYRUVATE DEHYDROGENASE KINASE, ISOZYME 2)		17q21.33
PIK3CA (PHOSPHOINOSITIDE-3-KINASE, CATALYTIC, ALPHA POLYPEPTIDE)	PI3K, p110-alpha	3q26.3
PIK3CB (PHOSPHOINOSITIDE-3-KINASE, CATALYTIC, BETA POLYPEPTIDE)	PI3K, PI3Kbeta, PIK3C1, p110-BETA	3q22.3
PIK3CG (PHOSPHOINOSITIDE-3-KINASE, CATALYTIC, GAMMA POLYPEPTIDE)	PI3CG, PI3K, PI3Kgamma, PIK3	7q22.3
PPARG (PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR-GAMMA)		3p25.2
PRKAA1 (PROTEIN KINASE, AMP-ACTIVATED, ALPHA 1 CATALYTIC SUBUNIT)	AMPK, AMPKa1	5p12
PRKAA2 (PROTEIN KINASE, AMP-ACTIVATED, ALPHA 2 CATALYTIC SUBUNIT)	AMPK, AMPK2, PRKAA	1p31
PRKAB1 (PROTEIN KINASE, AMP-ACTIVATED, BETA 1 NON-CATALYTIC SUBUNIT)	AMPK, HAMPKb	12q24.1
PRKAB2 (PROTEIN KINASE, AMP-ACTIVATED, BETA 2 NON-CATALYTIC SUBUNIT)	MGC61468	1q21.1
PRKAG2 (PROTEIN KINASE, AMP-ACTIVATED, GAMMA 2 NON-CATALYTIC SUBUNIT)	AAKG, AAKG2, CMH6, WPWS	7q36.1
PTEN (PHOSPHATASE AND TENSIN HOMOLOG)	BZS, MMAC1, PTEN1, TEP1	10q23.3
RAF1 (V-RAF-1 MURINE LEUKEMIA VIRAL ONCOGENE HOMOLOG 1)	CRAF, Raf-1, c-Raf	3p25
RPS6KA1 (RIBOSOMAL PROTEIN S6 KINASE, 90KDA, POLYPEPTIDE 1)	HU-1, MAPKAPK1A, RSK, RSK1, S6K-alpha1	1p
RPS6KA2 (RIBOSOMAL PROTEIN S6 KINASE, 90KDA, POLYPEPTIDE 2)	MAPKAPK1C, RSK, RSK3, S6K-alpha, S6K-alpha2, p90-RSK3	6q27
RPS6KB1 (RIBOSOMAL PROTEIN S6 KINASE, 70KDA, POLYPEPTIDE 1)	PS6K, S6K, S6K1, STK14A, p70(S6K)-alpha, p70-S6K	17q23.1
RPS6KB2 (RIBOSOMAL PROTEIN S6 KINASE, 70KDA, POLYPEPTIDE 2)	P70-beta, S6K-beta2, S6K2, SRK, STK14Bp70(S6K)-beta, p70S6Kb	11q13.2
RUNX1 (RUNT-RELATED TRANSCRIPTION FACTOR 1)	AML1, AML1-EVI-1, AMLCR1, PEBP2aB	21q22.3
RUNX2 (RUNT-RELATED TRANSCRIPTION FACTOR 2)	AML3, CBFA1, CCD, CCD1, PEBP2A1, PEBP2A2, PEBP2aA, PEBP2aA1	6p21
RUNX3 (RUNT-RELATED TRANSCRIPTION FACTOR 3)	AML2, CBFA3,PEBP2aC	1p36
SELS (SELENOPROTEIN S)	AD015, SBBI8, SEPS1, VIMP	15q26.3
SEP15 (15 KDA SELENOPROTEIN)		1p31
SEPN1 (SELENOPROTEIN N, 1)	MDRS1, RSMD1, RSS, SELN	1p36.13
SEPP1 (SELENOPROTEIN P, PLASMA, 1)	SELP, SeP	5q31
SEPW1 (SELENOPROTEIN W, 1)	selW	19q13.3
SEPX1 (SELENOPROTEIN X, 1)	MSRB1, SELR, SELX	16p13.3
SLC2A4 (SOLUTE CARRIER FAMILY 2, MEMBER 4)	GLUT4	17p13
SMAD1 (SMAD FAMILY MEMBER 1)	MADH1, MADR1	4q31
SMAD2 (SMAD FAMILY MEMBER 2)	MADH2, MADR2, hSMAD2	18q21.1
SMAD3 (SMAD FAMILY MEMBER 3)	MADH3, MGC60396, Smad3	15q22.33
SMAD4 (SMAD FAMILY MEMBER 4)	DPC4, JIP, MADH4	18q21.1
SMAD7 (SMAD FAMILY MEMBER 7)	FLJ16482, MADH7, MADH8	18q21.1
SOCS1 (SUPPRESSOR OF CYTOKINE SIGNALING 1)	CIS1, CISH1, JAB, SOCS-1, SSI-1, SSI1, TIP3	16p13.13
SOCS2 (SUPPRESSOR OF CYTOKINE SIGNALING 2)	CIS2, Cish2, SOCS-2, SSI-2, SSI2, STATI2	12q

SOD1 (SUPEROXIDE DISMUTASE 1)	IPOA	21q22.11
SOD2 (SUPEROXIDE DISMUTASE 2)	IPO-B, MNSOD	6q25.11
STAT1 (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1, 91KDA)	ISGF-3, STAT91	2q32.2
STAT2 (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2, 113KDA)	ISGF-3, P113, STAT113	12q13.2
STAT3 (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3)	APRF	17q21.31
STAT4 (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 4)		2q32.2-q32.3
STAT5A (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5A)	MGF, STAT5	17q11.2
STAT5B (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5B)	STAT5	17q11.2
STAT6 (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 6 (INTERLEUKIN-4 INDUCED))	IL-4-STAT, STAT6B, STAT6C	12q13
STK11 (SERINE/THREONINE KINASE 11)	LKB1, PJS	19p13.3
TCF7L2 (TRANSCRIPTION FACTOR 7-LIKE 2)	TCF4	10q25.2-q25.3
TERT (TELOMERASE REVERSE TRANSCRIPTASE)	EST2, TCS1, TP2, TRT, hEST2	5p15.33
TGFB1 (TRANSFORMING GROWTH FACTOR, BETA 1)	CED, DPD1, LAP, TGFB, TGFbeta	19q13.1
TGFBR1 (TRANSFORMING GROWTH FACTOR, BETA RECEPTOR 1)	AAT5, ACVRLK4, ALK-5, ALK5, LDS1A, LDS2A, SKR4, TGFR-1	9q22
TLR2 (TOLL-LIKE RECEPTOR 2)	CD282, TIL4	4q32
TLR3 (TOLL-LIKE RECEPTOR 3)	CD283	4q35
TLR4 (TOLL-LIKE RECEPTOR 4)	ARMD10, CD284, TOLL, hToll	9q32-q33
TNF (TUMOR NECROSIS FACTOR (TNF SUPERFAMILY, MEMBER 2))	DASS-280D8.2, DIF, TNF-alpha, TNFA, TNFSF2	6p21.3
TNFRSF1A (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY, MEMBER 1A)	CD120a, TNF-R, TNF-R-I, TNFR55, TNFR60, p55, p55-R, p60	12p13.2
TRAF2 (TNF RECEPTOR-ASSOCIATED FACTOR 2)	MGC	9q34
TSC1 (TUBEROUS SCLEROSIS 1)	KIAA0243, LAM, MGC86987, TSC	9q34
TSC2 (TUBEROUS SCLEROSIS 2)	FLJ43106, LAM, TSC4	16p13.3
TXNRD1 (THIOREDOXIN REDUCTASE 1)	GRIM-12, MGC9145, TR, TR1, TRXR1, TXNR	12q23-q24.1
TXNRD2 (THIOREDOXIN REDUCTASE 2)	SELZ, TR, TR-BETA, TR3, TRXR2	22q11.21
TXNRD3 (THIOREDOXIN REDUCTASE 3)	TGR, TR2, TRXR3	3q21.3
TYK2 (TYROSINE KINASE 2)	JTK1	19p13.2
VDR (VITAMIN D RECEPTOR)		12q13.11
VEGFA (VASCULAR ENDOTHELIAL GROWTH FACTOR A)	MGC70609, VEGF, VEGF-A, VPF	6p12

Supplemental Table 2. Significant interactions previously identified with colon cancer after adjustment for multiple comparisons that were tested in this study

NSAIDs		Smoking		BMI		OBS	
Gene	SNP	Gene	SNP	Gene	SNP	Gene	SNP
FLT1	rs2296188	*FLT1*	rs17086609	*FLT1*	rs1408245	*FLT1*	rs12429309
KDR	rs12502008		rs9319425		rs3936415		rs9554314
	rs12505758		rs9513095	*KDR*	rs17085326	*KDR*	rs1531290
	rs2071559	*KDR*	rs2305948		rs2219471		rs2305948
IRF2	rs3756093		rs2305949*		rs6828477	*EPX*	rs12602498
	rs3822118	*VEGFA*	rs3025033	*VEGFA*	rs2010963		rs2240815
	rs6856910	*IRF2*	rs6827018		rs25648*		rs9892223
	rs9684244	*IRF4*	rs11242865	*IL10*	rs1800890	*NOS2A*	rs16949
IRF4	rs1050975		rs12211228	*DUSP4*	rs2341674		rs4795067
	rs3778607		rs872071	*MAP3K2*	rs3732209		rs8072199
IRF5	rs1874328	*IRF6*	rs861020	*RPS6KA1*	rs12025634		
IRF6	rs2013162	*IL6R*	rs1386821*	*TXNRD1*	rs4964779		
	rs2013196	*JAK2*	rs10815160	*TXNRD2*	rs1044732		
IL15	rs17461269		rs1887429		rs7410379		
TYK2	rs280521		rs7043371	*TXNRD3*	rs777238		
MAP3K10	rs892117	*STAT2*	rs2229363*	*SEPN1*	rs11247735		
EPX	rs10853004	*STAT4*	rs4853540		rs718391		
TXNRD1	rs4523760	*STAT5A*	rs7217728	*TERT*	rs10069690		
	rs4964778	*STAT5B*	rs7218563		rs2242652		
TXNRD2	rs17745445	*STAT6*	rs3024974		rs2736118		
	rs3788314	*MAP3K11*	rs1784223		rs4246742		
	rs5992493		rs7116712				
	rs756661	*NOS2A*	rs2274894				
TERT-CLPTM1L	rs2853668		rs3729508				
TLR2	rs7656411		rs7406657				
TLR3	rs11721827		rs944725				
MAP3K7	rs13208824	*TXNRD2*	rs17745445				
TGFBR1	rs6478974		rs5992493				
Smad3	rs3743343	*TLR2*	rs3804099				
	rs7173811	*NFAM1*	rs13055337				
		NFAT5	rs8049728				
		TGFB1	rs4803455				
		TGFBR1	rs10733710				
		TGFBR1	rs1571590				

*Excluded from analysis due to low MAF in subset up population with RNAseq data.