

Dietary intake alters gene expression in colon tissue: possible underlying mechanism for the influence of diet on disease

Andrew J. Pellatt, Martha L. Slattery, Lila E. Mullany, Roger K. Wolff and Daniel F. Pellatt

Background Although the association between diet and disease is well documented, the biologic mechanisms involved have not been entirely elucidated. In this study, we evaluate how dietary intake influences gene expression to better understand the underlying mechanisms through which diet operates.

Methods We used data from 144 individuals who had comprehensive dietary intake and gene expression data from RNAseq using normal colonic mucosa. Using the DESeq2 statistical package, we identified genes that showed statistically significant differences in expression between individuals in high-intake and low-intake categories for several dietary variables of interest adjusting for age and sex. We examined total calories, total fats, vegetable protein, animal protein, carbohydrates, trans-fatty acids, mutagen index, red meat, processed meat, whole grains, vegetables, fruits, fiber, folate, dairy products, calcium, and prudent and western dietary patterns.

Results Using a false discovery rate of less than 0.1, meat-related foods were statistically associated with 68 dysregulated genes, calcium with three dysregulated genes, folate with four dysregulated genes, and nonmeat-related foods with 65 dysregulated genes. With a more stringent false discovery rate of less than 0.05, there were nine meat-related dysregulated genes and 23 nonmeat-related genes. Ingenuity pathway analysis identified three major networks among genes identified as dysregulated with respect to meat-related dietary variables and three networks among

genes identified as dysregulated with respect to nonmeat-related variables. The top networks (Ingenuity Pathway Analysis network score > 30) associated with meat-related genes were (i) cancer, organismal injury, and abnormalities, tumor morphology, and (ii) cellular function and maintenance, cellular movement, cell death, and survival. Among genes related to nonmeat consumption variables, the top networks were (i) hematological system development and function, nervous system development and function, tissue morphology and (ii) connective tissue disorders, organismal injury, and abnormalities.

Conclusion Several dietary factors were associated with gene expression in our data. These findings provide insight into the possible mechanisms by which diet may influence disease processes. *Pharmacogenetics and Genomics* 26:294–306 Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

Pharmacogenetics and Genomics 2016, 26:294–306

Keywords: calcium, diet, fiber, folate, gene expression, meat, vegetables, whole grains

Department of Internal Medicine, University of Utah, Salt Lake City, Utah, USA

Correspondence to Martha L. Slattery, PhD, MPH, Department of Internal Medicine, University of Utah, 383 Colorow, Salt Lake City, UT 84108, USA
Tel: +1 801 585 6955; fax: +1 801 581 3623;
e-mail: marty.slattery@hsc.utah.edu

Received 4 November 2015 Accepted 17 February 2016

Introduction

Diet has been associated with several diseases including atherosclerosis [1,2], inflammatory bowel disease [3], steatosis and nonalcoholic fatty liver disease [4,5], seizures [6], and various cancers including colon and breast carcinoma [7–9]. Moreover, diet has also been identified as a risk factor for type 2 diabetes mellitus, obesity, and metabolic syndrome [10]. Beyond its role in disease status, diet has been shown to play an important role in a number of physiologic and metabolic processes including

bone remodeling [11] and aspects of the immune response [12]. The biologic mechanisms involved in these processes are less clear.

Studies have evaluated the role of diet in the regulation of specific physiologic and pathologic processes, showing that single dietary metabolites and nutrients play a role in the regulation of certain signaling pathways [11,13–15] or cytokine profiles [16]. Other studies have focused on dietary effects on oxidative stress [17] or the role of potentially chemoprotective nutrients [18]. Some of these studies also have looked at localized gene expression within specific hypothesized pathways and processes, suggesting that diet plays a role in the regulation of gene expression [7,8,16,19]. Recently, the effects of dietary patterns on the entire gene expression profile were investigated, suggesting a differential expression of genes in individuals with a prudent dietary pattern versus

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (www.pharmacogeneticsandgenomics.com).

This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

a western dietary pattern and identified specific canonical pathways associated with both dietary patterns [20].

In this study, we hypothesize that levels of gene expression in normal colonic mucosa will be associated with dietary factors. Dietary factors that we consider include sources of calories, including animal protein and vegetable protein, carbohydrates, and total fat. We also consider trans-fatty acids, total red meat, processed meat, mutagen index, dairy products, fruit, vegetables, refined grains, whole grains, sucrose, dietary fiber, calcium, and folate with gene expression profiles. Finally, we examined the association between prudent dietary pattern and western dietary pattern with gene expression.

Methods

Total RNA was available from normal colonic mucosa for 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) [21]. Cases were identified using a rapid-reporting system, had tumor registry verification of a first primary adenocarcinoma of the colon, and were diagnosed between October 1991 and September 1994; tumor tissue blocks (used for RNA) were obtained for 97% of all Utah cases and for 85% of all KPMRP cases [22]. Individuals with known adenomatous polyposis coli, 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; all study participants signed informed consent forms before participation.

Dietary data

Data were collected by trained and certified interviewers using laptop computers shortly after diagnosis. All interviews were audio-taped as described previously and reviewed for quality control (QC) purposes [23]. Any interview deemed questionable by the interviewer was reviewed centrally; this enabled us to evaluate the data by distinguishing between an interview that was difficult for the interviewer versus one that was difficult for the participant. In addition, we reviewed audio tapes for all individuals whose nutrient levels were considered outliers (over the 95% tile or under the 5% tile of intake); this enabled us to correct any coding errors and determine the quality and face validity of the interview. Dietary information was obtained for the year, 2 years before diagnosis using an extensive diet history questionnaire adapted from the validated CARDIA diet history; most individuals were asked to recall usual dietary intake from 2 to 3 years ago to obtain the prediagnosis diet [24]. Foods were converted into nutrients using the Nutrition Coding Center Nutrient Data System (University of Minnesota, Nutrition Coding Center, Minneapolis, Minnesota, USA), version 19 and were also grouped into categories of similar foods. We assessed

both foods and nutrients with gene expression. Foods units were standard servings per day, which was half cup of fruit, vegetable, or dairy product, meat servings were 2–3 oz, and grain products were half cup of rice-type grains or one slice of bread. Prudent and western dietary patterns were developed on the basis of the principal component program [25]. Our prudent dietary pattern was heavily loaded toward diets high in fruits, vegetables, whole grains, fish, and chicken, whereas the western dietary pattern was highly loaded toward red meat, processed meats, and refined grains and high-sugar, high-fat foods. Additional questions were asked about meat consumption, doneness, and preparation methods that were combined and used to create a mutagen index score [26].

RNA processing

RNA was extracted from formalin-fixed paraffin-embedded colonic tissues. Formalin-fixed paraffin-embedded tissue has been shown to be a reliable source of RNA for use in conjunction with RNA-seq [27]. The study pathologist reviewed slides to delineate carcinoma and normal colonic mucosa. Total RNA was extracted, isolated, and purified using the RecoverAll Total Nucleic Acid isolation kit (Ambion, Carlsbad, California, USA), RNA yields were determined using a NanoDrop spectrophotometer (Wilmington, Delaware, USA).

Sequencing library preparation

Library construction was performed using the TruSeq stranded total RNA sample preparation kit (Illumina, San Diego, California, USA) with Ribo-Zero (Illumina) as described previously [28]. Briefly, ribosomal RNA was removed from 100 ng total RNA using biotinylated Ribo-Zero oligos attached to magnetic beads. Following purification, the rRNA-depleted sample was fragmented and primed with random hexamers. First-strand reverse transcription was performed using superscript II reverse transcriptase (Invitrogen, Carlsbad, California, USA). Second-strand cDNA synthesis was performed 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. The PCR-amplified library was purified using Agencourt AMPure XP beads (Beckman Coulter Genomics, Danvers, Massachusetts, USA).

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 HiSeq2500 were aligned to a sequence database containing the human genome (build GRCh37/hg19, February 2009, from <http://www.genome.ucsc.edu>) plus all splice junctions generated using the USeq MakeTranscriptome application (version 8.8.1, available at: <http://www.useq.sourceforge.net>). Alignment was performed using NovoAlign, version 2.08.01 available from <http://www.novocraft.com>, which also trimmed any adapter sequence. Genome alignments to splice junctions were translated back to genomic coordinates using the USeq SamTranscriptomeParser application. Alignments were then sorted and indexed using the Picard SortSam application (version 1.100, available at: <http://broadinstitute.github.io/picard/>). Aligned read counts for each gene were calculated using pysam (<https://code.google.com/p/pysam/>) and samtools (<http://www.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://www.genome.ucsc.edu>. Over 17 000 genes that were expressed in colon mucosa were analyzed with dietary factors. Rigid QC procedures were carried out to ensure high-quality data; participants failing QC were excluded from the analysis.

Statistical methods

Of the 197 initial tumor/nontumor tissue pairs, 22 patients failed QC on the basis of a low number of sequences, leaving 175 patients with high-quality expression data. Of these, 144 had dietary questionnaire data. For each dietary factor, our analysis focused on contrasting individuals with lower intake to those of individuals with higher intake. Dietary data were evaluated using nutrients per 1000 calories and standard servings of food; sample tertiles were computed for the consumption levels of each dietary factor. For each such dietary variable, we determined which genes showed statistically significant differential expression between the lowest and the highest intake tertile categories after adjusting for age and sex using the Bioconductor package DESeq2 (Stanford, California, USA), written for the R statistical programming environment; these genes were considered dysregulated. DESeq2 assumes that the RNA-Seq counts are distributed according to negative binomial distributions. It utilizes generalized linear modeling and variance-reduction techniques for estimated coefficients to test individual null hypotheses of zero log₂-fold changes between high and low categories (i.e. no differential expression) for each gene. It uses both an independent-filtering method and the Benjamini and Hochberg [29] procedure to improve power and control

the false discovery rate (FDR). The default DESeq2 options were used, including replacement of outliers, as defined by Cook's distance, and the Wald test. For further details on DESeq2, see Love *et al.* [30]. In identifying genes with differential expression, an FDR of 0.10 was used. The fold change calculations for differentially expressed genes was determined by DESeq2 and represents the log₂ change in expression level (i.e. counts) for the high versus low dietary categories.

Bioinformatics analysis was carried out on the list of Ensemble IDs associated with genes identified as differentially expressed between high-consumption and low-consumption categories for the dietary variables of interest using QIAGEN's Ingenuity Pathway Analysis (IPA) [31]. We used genes from ingenuity knowledge base and considered both indirect and direct relationships. Causal and interaction networks were generated. Interaction networks were limited to 35 molecules per network and 25 networks per analysis, and excluded endogenous chemicals. We focused on algorithmically derived interaction networks, which are assigned a score on the basis of their relevance to the genes in the input dataset, the number of focus genes (i.e. dysregulated genes in our data that are in that network), and their connectivity [32]. The score is calculated as $-\log_{10}P$, where P is generated using a Fisher's exact test [33]. Studies have found scores greater than 3 to be significant, with a score of 3 indicating a 1/1000 chance that the focus genes are in a network because of random chance [34–36]. Other studies have opted to utilize more stringent criteria and higher scores to ensure that their discovered networks are highly significant [37,38]; we utilized highly stringent criteria, only including networks with scores over 20. We applied the Benjamini and Hochberg multiple testing correction to assess pathways in IPA.

Availability of data is restricted to that authorized in the patient consent form and in accordance with data transfer agreements and institutional review board requirements.

Results

The majority of our study population included men and the median age was 65 years; 13.2% were current smokers and 37.3% currently took aspirin or nonsteroidal anti-inflammatory drugs on a regular basis (Table 1). The median BMI was 29.6. Within this population, few individuals consumed processed meat, with the highest level of intake being less than one serving per day. The highest red meat consumption was less than two servings per day. Over half of the study participants had over two servings of vegetables per day and one serving of whole grains per day.

Seven genes were differentially expressed for dietary calcium and folate (Table 2). Sixty-five genes were differentially expressed between high-intake and low-intake categories among nonmeat variables (i.e. prudent

Table 1 Description of the study population

Sex [n (%)]	
Male	80 (55.6)
Female	64 (44.4)
Smoking status [n (%)]	
Never	60 (41.7)
Former	65 (45.1)
Current	19 (13.2)
Recent NSAID use [n (%)]	
No	89 (62.7)
Yes	53 (37.3)
Alcohol consumption, long term and current (g/day) [n (%)]	
Low	53 (36.8)
Mod	68 (47.2)
High	23 (16.0)
Race/ethnicity [n (%)]	
Nonhispanic white	124 (86.1)
Black	8 (5.6)
Hispanic	9 (6.3)
Other/NA	3 (2.1)
Education (years) [n (%)]	
< High school	24 (16.7)
High school	42 (29.2)
Some college	50 (34.7)
College grad/higher	28 (19.4)
Age (years) [median (interquartile range)]	65.0 (56–71)
BMI (kg/m ²) [median (interquartile range)]	29.6 (26.6–33.2)
Hours moderate physical activity/week [median (interquartile range)]	1.5 (0.0–3.9)
Meat consumption [median (interquartile range)]	
Processed meat (2–3 oz servings/day)	
Low ^a	0.05 (0.01–0.07)
Intermediate	0.19 (0.14–0.23)
High	0.45 (0.35–0.59)
Red meat (2–3 oz servings/day)	
Low	0.43 (0.29–0.61)
Intermediate	0.95 (0.84–1.07)
High	1.58 (1.41–1.93)
Nonmeat consumption [median (interquartile range)]	
Calcium (mg/1000 kcal)	
Low	300.71 (261.72–326.90)
Intermediate	406.47 (380.31–445.51)
High	585.69 (541.01–697.95)
Fiber (g/1000 kcal)	
Low	7.32 (6.75–8.09)
Intermediate	10.03 (9.52–10.43)
High	13.06 (12.15–15.62)
Folate (mcg/1000 kcal)	
Low	118.35 (103.41–126.46)
Intermediate	151.61 (142.85–161.84)
High	203.63 (189.98–247.94)
Vegetables (1/2 cup servings/day)	
Low	1.41 (1.03–1.90)
Intermediate	2.85 (2.55–3.25)
High	4.59 (4.13–5.74)
Whole grains (servings/day)	
Low	0.43 (0.33–0.63)
Intermediate	1.12 (1.02–1.29)
High	2.14 (1.82–2.52)

NA, not applicable.

^aAssociations are adjusted for age and sex.

dietary pattern, fruits, vegetables, and whole grains). This can be further broken down into one gene with prudent dietary pattern, one gene with fruit intake, 26 genes with vegetable intake, and 37 genes with whole grain intake (Table 3, FDR < 0.1). Several genes were identified as differentially expressed between consumption categories for multiple dietary variables. *TXNDC17* was upregulated for both prudent dietary pattern and vegetable intake; *MUC5AC* was downregulated for vegetable

intake, whole grain intake, and dietary folate. *FOXJ2*, *NECAP1*, and *C3AR1* were upregulated for both calcium intake and vegetable intake. Using a more stringent FDR of less than 0.05, we identify one gene with differential expression between calcium intake categories, one with prudent dietary pattern, 13 with vegetable intake, and eight with whole grain intake (Tables 2 and 3). Among the genes with an FDR of less than 0.05, *FOXJ2* was upregulated with high calcium intake and *TXNDC17* was unregulated with a high prudent dietary pattern. In contrast, five of the eight genes associated with whole grains were downregulated and four of the 13 genes associated with vegetables were downregulated.

Our IPA analysis found three interaction networks to be significantly associated with nonmeat consumption (IPA network score ≥ 20). These IPA networks were identified as functioning with cancer, organismal injury and abnormalities, and tumor morphology (score = 31, focus molecules = 11) (Fig. 1a), cellular function, maintenance, cellular movement, cell death and survival (score = 31, focus molecules = 15) (Fig. 1b), and drug metabolism, molecular transport, and small molecule biochemistry (score = 26, focus molecules = 13) (Fig. 1c).

Among genes identified as significantly differentially expressed between high and low meat-related intake categories, 64 genes were differentially expressed with red meat, two genes were differentially expressed with processed meat, and three genes were differentially expressed with a western dietary pattern (Table 4). Of these, eight genes were differentially expressed by the level of red meat and *LSAMP* was differentially expressed by the level of western dietary pattern when using a more stringent selection criterion of a maximal FDR of less than 0.05. The majority of these genes were downregulated by red meat consumption; of all of the genes identified, only *IL22RA1*, *SLC17A4*, *SLC25A25*, *MUC17*, *ZFY*, *FRK*, *PEX11G*, *NEDD9*, *NR3C2*, and *RPS4Y1* were upregulated by red meat consumption. The two genes identified as being differentially expressed for western dietary pattern were also differentially expressed for red meat; *C3* was downregulated by red meat and processed meat.

In IPA analysis, three interaction networks were significantly associated (score ≥ 20) with meat-related variables (i.e. red meat, processed meat, and western dietary pattern). These IPA networks were identified as hematological system developmental function, nervous system development and function, tissue morphology (IPA network score = 31, focus molecules 22) (Fig. 2a); connective tissue disorders, organismal injury and abnormalities, cardiovascular system development and function (score = 29, focus molecules = 14) (Fig. 2b); and cellular compromise, cellular function and maintenance, and protein degradation (score = 21, focus molecules 11) (Fig. 2c). Supplement Table 1, Supplemental digital content 1, <http://links.lww.com/FPC/B6> shows genes

Table 2 Associations between nutrients and gene expression (FDR < 0.1)^a

	Gene	Average adjusted count		DESeq2 estimated log ₂ -fold change	P value	Multiple comparison adjusted P value
		Low tertile	High tertile			
Calcium	<i>FOXJ2</i>	32.84	135.64	0.87	4.05E-10	7.07E-06
	<i>NECAP1</i>	23.38	68.92	0.71	3.87E-07	2.38E-03
	<i>C3AR1</i>	3.51	21.28	0.63	4.09E-07	2.38E-03
Folate	<i>A1BG</i>	3.84	8.45	0.63	2.12E-05	9.22E-02
	<i>SLC37A2</i>	23.04	79.47	0.61	1.83E-05	9.22E-02
	<i>CLDN4</i>	129.49	87.70	-0.47	1.44E-05	9.22E-02
	<i>MUC5AC</i>	58.73	5.49	-0.50	1.52E-05	9.22E-02

FDR, false discovery rate.

^aAssociations are adjusted for age and sex.

upregulated and downregulated in each of these networks. Some genes were differentially expressed between high-consumption and low-consumption categories for both meat and non-meat dietary variables. *COL1A1* was downregulated for vegetables and meat intake; *MUC5AC* was downregulated for folate, vegetables, whole grains, and meat. Foods and nutrients analyzed that were not associated with any differential gene expression (FDR < 0.1) were mutagen index, total dairy products, carbohydrates, trans-fatty acids, total dietary fat, dietary fiber, refined grain products, and sucrose. In addition, we observed no differences in gene expression by level of vigorous physical activity.

Discussion

Our study provides support for associations between level of dietary intake and gene expression in normal colonic mucosa utilizing RNA-Seq, which enabled the assessment of over 17 000 genes. Some genes were differentially expressed by intake level across multiple dietary variables, generally in the same direction, although there was limited overlap between meat and nonmeat variables. We observed three IPA interaction networks to be highly associated with meat consumption and three IPA interaction networks to be associated with nonmeat consumption. These data suggest that dietary influence on gene expression plays an important role in a number of physiologic and pathologic processes.

The majority of genes that were differentially expressed with high red meat consumption were downregulated. These genes predominantly play a role in extracellular matrix (ECM) assembly and structure, calcium binding, and the regulation of signal transduction through the Wnt signaling pathway. The canonical Wnt signaling pathway triggers β -catenin-dependent gene expression, allowing cell cycle progression at the G1/S transition [39]. In particular, we found that high intake of meat correlated with downregulation of *FZD3*, *TMEM132B*, *CCDC150*, *EGFLAM*, and *NKD1*. Both *TMEM132B* and *NKD1* are known to downregulate Wnt signaling [40]. Thus, meat consumption can contribute toward the dysregulation of a proto-oncogenic pathway. At the same time, we found that high red meat consumption was associated with a

downregulation of *TP53*, encoding for the tumor suppressor p53. Both of these findings are consistent with previous findings that red meat intake can induce pro-carcinogenic gene expression changes through the p53 and Wnt signaling pathways [19]. This is possibly because of the fact that red meat is high in dietary heme, a luminal irritant associated with hyperproliferation and hyperplasia through the downregulation of feedback inhibitors, including Wnt inhibitors [17]. These findings add to the literature that suggest that red meat increases the risk of several chronic diseases including cancer.

High red meat consumption also downregulated the expression of genes associated with ECM. Dysregulation of genes involved in the regulation of the ECM and cell matrix adhesion, such as matrix metalloproteinases (MMPs), collagen-related genes such as *COL4A2*, *COL1A2*, and *COL1A1*, tissue inhibitor of metalloproteinases (TIMPs), and one cut homeobox 2 (*ONECUT2*) have been implicated in a variety of diseases [41]. Moreover, red meat is associated with both acute oxidative stress and delayed cytotoxic stress [17]. Several genes dysregulated by meat, including *THY1*, *NCL*, *COL4A2*, and *KRT1*, contribute to metastasis and angiogenesis. We found that diet downregulated a variety of genes associated with ECM assembly and maintenance, including *TIMP1*, providing an additional mechanistic link between red meat and disease processes.

It is also worth noting that red meat upregulated some genes, namely, genes with products involved in phosphate transport (*SLC17A4* and *SLC25A25*) and cytokine signaling (*MUC17* and *IL22RA1*). Diet is known to play an important role in bone health, with meat protein increasing acid load and regulating calcium balance [42]. The upregulation of genes involved in phosphate transport as well as our observed downregulation of calcium-binding genes (*SPARC*, *CEMIP*, *HSPA5*, *HSP901B*, *VWDE*, *SLC25A25*, *C1R*, *CLDN1*, and *EGFLAM*) may help explain the mechanism by which meat consumption modulates bone metabolism, beyond the effects of dietary protein on acid load. Moreover, the upregulation of *MUC17* and *IL22RA1* may play a role in meat-induced inflammation and inflammation-related diseases. The exact function of *MUC17* is unknown; however, other

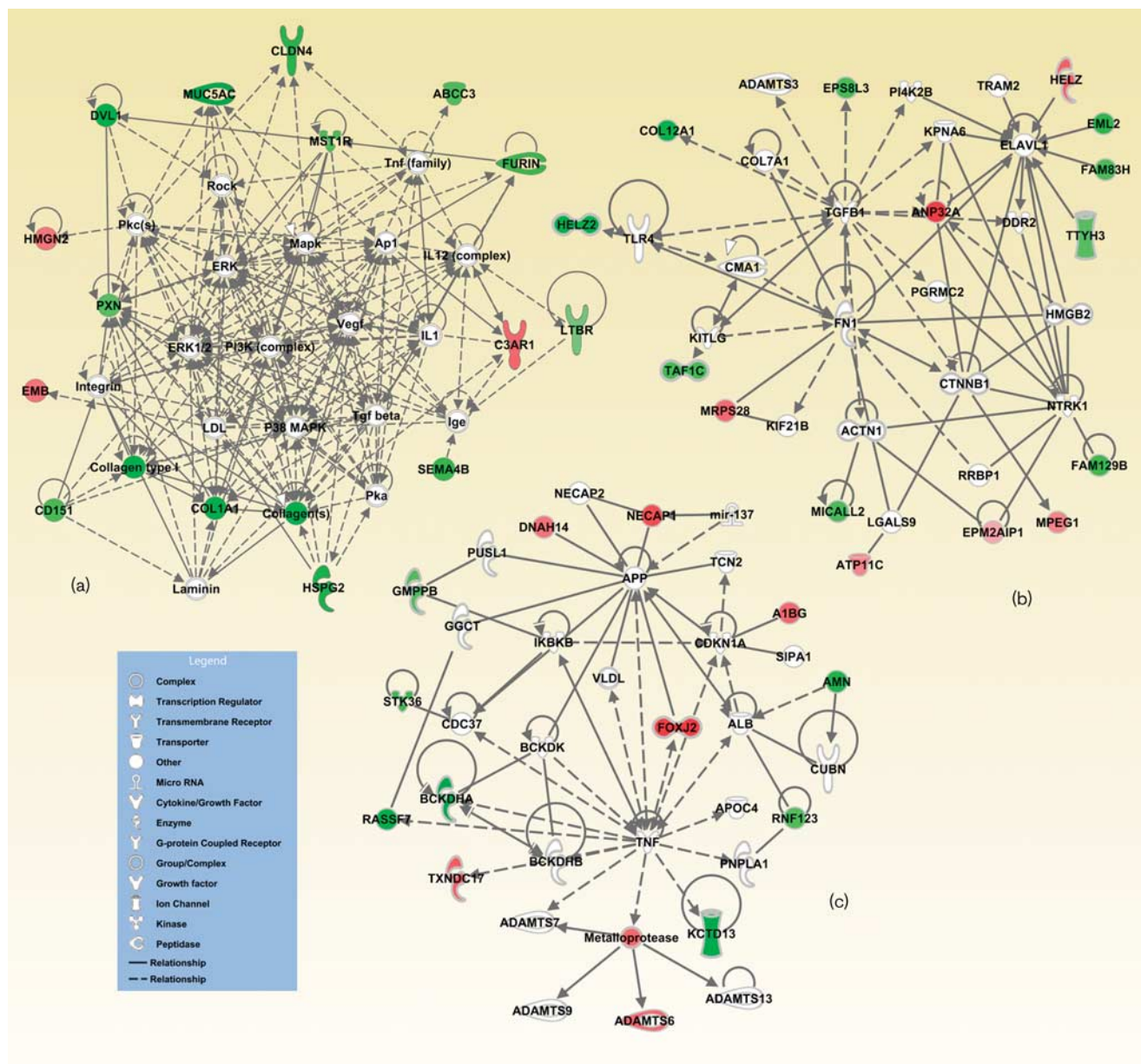
Table 3 Associations between nonmeat foods and gene expression (FDR < 0.1)^a

	Gene	Average adjusted count		DESeq2 estimated log ₂ -fold change	P value	Multiple comparison adjusted P value
		Low tertile	High tertile			
Prudent diet	<i>TXNDC17</i>	6.18	16.74	0.68	4.00E-07	6.99E-03
Fruit	<i>HMG2</i>	8.45	16.66	0.57	4.31E-06	7.50E-02
Vegetables	<i>ANP32A</i>	64.67	622.05	0.86	3.22E-10	2.90E-06
	<i>NECAP1</i>	30.12	160.72	0.76	5.98E-08	1.94E-04
	<i>HELZ</i>	639.22	10029.11	0.66	6.49E-08	1.94E-04
	<i>SLC2A3</i>	50.41	263.66	0.75	9.46E-08	1.95E-04
	<i>FOXJ2</i>	53.51	337.01	0.72	1.09E-07	1.95E-04
	<i>ADAMTS6</i>	9.00	40.07	0.63	5.57E-06	6.26E-03
	<i>HSPG2</i>	285.80	190.14	-0.48	4.47E-06	6.26E-03
	<i>C3AR1</i>	7.36	55.37	0.57	4.98E-06	6.26E-03
	<i>TXNDC17</i>	6.23	14.92	0.62	9.03E-06	9.03E-03
	<i>PXN</i>	124.57	91.47	-0.39	1.54E-05	1.26E-02
	<i>MUC5AC</i>	49.01	3.56	-0.48	1.51E-05	1.26E-02
	<i>FURIN</i>	55.18	36.93	-0.45	2.09E-05	1.56E-02
	<i>DNAH14</i>	21.82	62.62	0.58	6.69E-05	4.63E-02
	<i>RNF123</i>	45.86	30.57	-0.40	7.22E-05	4.64E-02
	<i>MICALL2</i>	81.80	57.71	-0.41	9.91E-05	5.94E-02
	<i>COL1A1</i>	478.40	278.35	-0.53	1.09E-04	6.15E-02
	<i>COL12A1</i>	94.94	58.32	-0.49	1.22E-04	6.27E-02
	<i>STK36</i>	23.04	14.94	-0.43	1.25E-04	6.27E-02
	<i>REG4</i>	75.69	30.88	-0.54	1.68E-04	7.96E-02
	<i>AMN</i>	65.79	41.60	-0.47	1.92E-04	8.62E-02
	<i>PTER</i>	10.51	16.89	0.44	2.02E-04	8.64E-02
	<i>ACAP3</i>	56.17	37.98	-0.41	2.24E-04	9.17E-02
	<i>TAF1C</i>	45.25	31.54	-0.42	2.70E-04	9.83E-02
	<i>TTYH3</i>	60.34	44.08	-0.39	2.58E-04	9.83E-02
	<i>EPM2AIP1</i>	44.29	60.95	0.34	2.73E-04	9.83E-02
	<i>ATP11C</i>	22.62	41.25	0.49	2.87E-04	9.94E-02
Whole grains	<i>MUC5AC</i>	55.79	3.17	-0.62	2.52E-06	2.47E-02
	<i>BCKDHA</i>	16.28	7.89	-0.69	3.59E-06	2.47E-02
	<i>BTG1</i>	74.46	100.65	0.42	6.99E-06	3.20E-02
	<i>ANKRD30BL</i>	118.66	53.70	-0.70	1.05E-05	3.61E-02
	<i>ABCC3</i>	341.13	253.51	-0.41	1.54E-05	3.90E-02
	<i>MPEG1</i>	33.40	49.38	0.50	2.25E-05	3.90E-02
	<i>EPS8L3</i>	143.09	103.38	-0.42	2.27E-05	3.90E-02
	<i>JMJD7</i>	11.64	6.22	-0.63	1.97E-05	3.90E-02
	<i>DVL1</i>	39.48	25.55	-0.53	5.97E-05	5.46E-02
	<i>HELZ2</i>	91.24	58.30	-0.52	5.32E-05	5.46E-02
	<i>GSTM3</i>	4.19	9.13	0.64	5.94E-05	5.46E-02
	<i>NR1I2</i>	56.26	39.67	-0.46	4.25E-05	5.46E-02
	<i>GMPPB</i>	54.06	40.48	-0.40	4.89E-05	5.46E-02
	<i>AMIGO3</i>	55.35	41.00	-0.42	4.93E-05	5.46E-02
	<i>C9orf1172</i>	10.88	6.23	-0.60	5.86E-05	5.46E-02
	<i>TPRN</i>	54.68	35.53	-0.53	6.45E-05	5.54E-02
	<i>EMB</i>	14.29	24.80	0.58	6.99E-05	5.65E-02
	<i>RASSF7</i>	37.58	23.88	-0.51	8.00E-05	5.98E-02
	<i>KCTD13</i>	10.46	5.71	-0.58	8.28E-05	5.98E-02
	<i>RTKL1-TNFRSF6B</i>	89.14	62.54	-0.43	9.18E-05	6.30E-02
	<i>C9orf50</i>	11.32	2.44	-0.61	9.77E-05	6.39E-02
	<i>PFKL</i>	144.14	108.76	-0.38	1.09E-04	6.76E-02
	<i>MRPS28</i>	3.04	7.82	0.62	1.13E-04	6.76E-02
	<i>LTBR</i>	69.00	55.07	-0.33	1.27E-04	6.96E-02
	<i>MST1R</i>	97.85	71.06	-0.41	1.26E-04	6.96E-02
	<i>TERC</i>	27.56	14.83	-0.58	1.37E-04	7.21E-02
	<i>FAM129B</i>	113.83	79.80	-0.44	1.51E-04	7.68E-02
	<i>CTC-435M10.3</i>	11.71	6.39	-0.57	1.57E-04	7.70E-02
	<i>FAM83H</i>	73.50	53.33	-0.41	1.68E-04	7.86E-02
	<i>CHKB-CPT1B</i>	70.38	48.67	-0.45	1.72E-04	7.86E-02
	<i>SEMA3B</i>	54.89	35.79	-0.50	1.95E-04	8.37E-02
	<i>DPM2</i>	18.77	11.53	-0.51	1.97E-04	8.37E-02
	<i>CD151</i>	79.03	57.71	-0.42	2.01E-04	8.37E-02
	<i>RNF146</i>	13.31	19.88	0.46	2.45E-04	9.39E-02
	<i>SEMA4B</i>	52.73	35.60	-0.45	2.39E-04	9.39E-02
	<i>NOC2L</i>	35.57	24.81	-0.44	2.46E-04	9.39E-02
	<i>EML2</i>	21.68	13.83	-0.46	2.68E-04	9.96E-02

FDR, false discovery rate.

^aAssociations are adjusted for age and sex.

Fig. 1



IPA networks associated with dysregulated genes on the basis of the level of nonmeat-related dietary intake. (a) Hematological system development and function, nervous system development and function, tissue morphology (score 31). (b) Connective tissue disorders, organismal injury and abnormalities, cardiovascular system development, and function (score 29). (c) Cellular compromise, cellular function and maintenance, and protein degradation (score 21). Red indicates upregulation and green indicates downregulation.

mucins are reportedly involved in the regulation of inflammation [43] and *MUC17* expression is related to certain cancers, including pancreatic cancer [44]. *IL22* is produced by T lymphocytes and mediates cellular inflammatory responses through the activation of intracellular kinases and has been implicated in a variety of inflammatory diseases, including colon cancer and ulcerative colitis [45], both of which have been associated with red meat consumption [17,46].

Several genes associated with immune response were downregulated by high meat consumption, including *TRIL* and *IGF2*. We found that high red meat consumption decreased the expression of *C1R* and *C1S*. *C1S* encodes for a component of C1 in the classic complement pathway; C1R is the R subcomponent of C1. The downregulation of the components of the complement cascade by meat consumption seems paradoxical; however, evidence suggests that the proximal classical

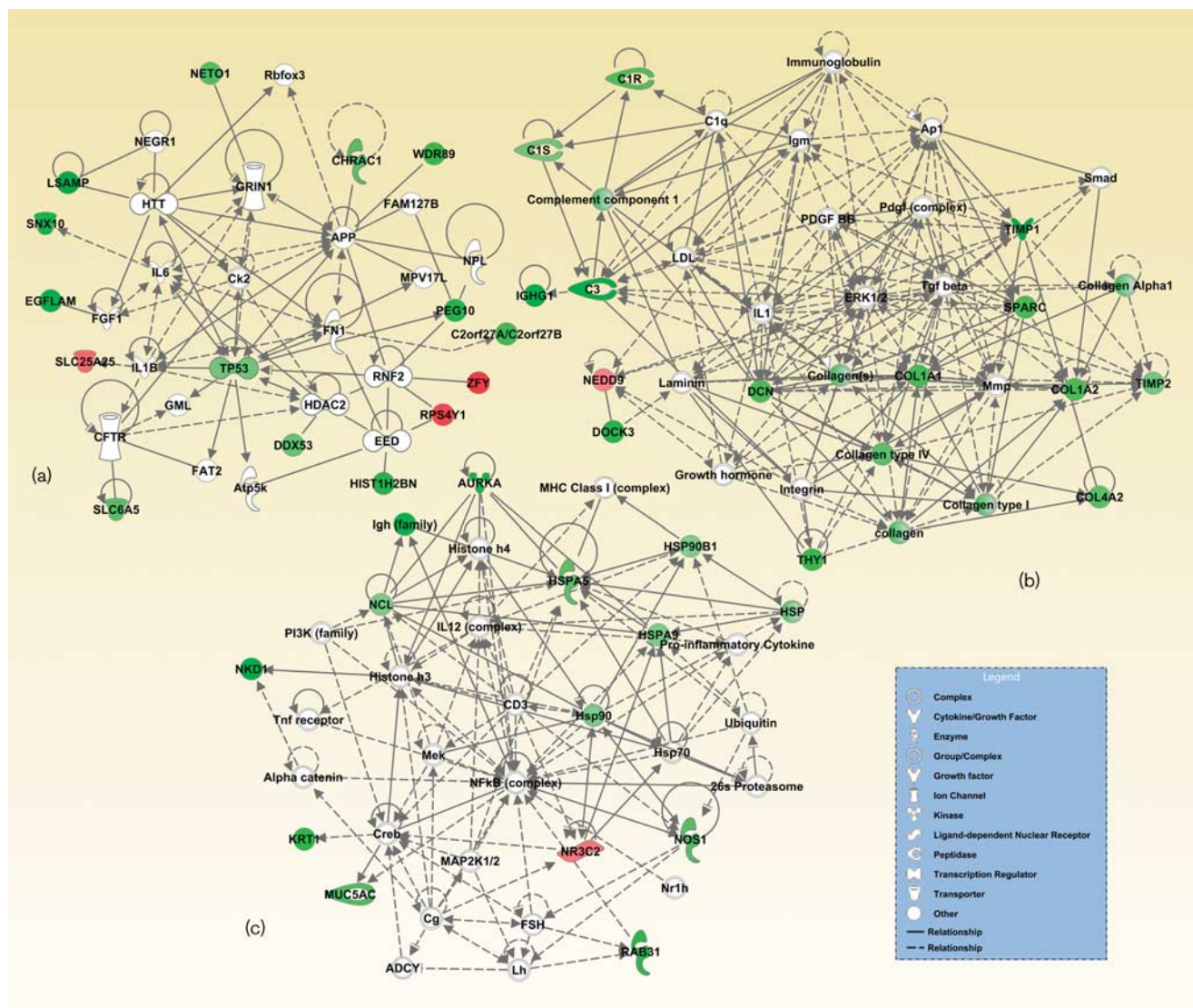
Table 4 Association between meat-related foods and gene expression (FDR < 0.1)^a

	Gene	Average adjusted count		DESeq2 estimated log2-fold change	P value	Multiple comparison adjusted P value
		Low tertile	High tertile			
Western diet	<i>LSAMP</i>	7.21	1.51	-0.70	2.20E-06	3.83E-02
	<i>TDRD15</i>	2.21	0.21	-0.57	6.21E-06	5.42E-02
Meat	<i>ONECUT2</i>	12.45	2.07	-0.92	2.40E-08	4.18E-04
	<i>TIMP1</i>	26.41	11.51	-0.76	1.39E-06	1.22E-02
	<i>CLDN1</i>	12.53	2.57	-0.77	3.78E-06	1.32E-02
	<i>CMSS1</i>	11.46	4.35	-0.77	3.34E-06	1.32E-02
	<i>LSAMP</i>	7.47	1.56	-0.79	3.58E-06	1.32E-02
	<i>SPARC</i>	118.63	74.35	-0.55	8.19E-06	2.38E-02
	<i>TDRD15</i>	1.98	0.14	-0.64	1.54E-05	3.84E-02
	<i>COL1A2</i>	314.64	214.85	-0.54	1.97E-05	4.29E-02
	<i>CEMIP</i>	27.24	16.30	-0.56	3.66E-05	6.73E-02
	<i>NKD1</i>	16.63	6.86	-0.69	3.86E-05	6.73E-02
	<i>ZFY</i>	4.11	9.63	0.65	5.84E-05	7.01E-02
	<i>FZD3</i>	17.37	9.06	-0.63	6.32E-05	7.01E-02
	<i>COL4A2</i>	155.42	108.31	-0.50	6.44E-05	7.01E-02
	<i>TP53</i>	33.67	23.40	-0.44	4.87E-05	7.01E-02
	<i>SLC17A4</i>	46.78	78.07	0.58	6.44E-05	7.01E-02
	<i>CCT2</i>	27.99	18.04	-0.54	5.51E-05	7.01E-02
	<i>HSPA5</i>	50.38	34.01	-0.50	8.23E-05	8.24E-02
	<i>TPX2</i>	18.50	8.63	-0.61	1.47E-04	8.24E-02
	<i>FRK</i>	15.94	23.66	0.50	1.44E-04	8.24E-02
	<i>PTK7</i>	18.21	10.62	-0.57	8.63E-05	8.24E-02
	<i>NCL</i>	109.20	80.61	-0.40	1.03E-04	8.24E-02
	<i>IL22RA1</i>	11.53	22.37	0.61	1.17E-04	8.24E-02
	<i>GPR26</i>	5.57	0.71	-0.55	1.46E-04	8.24E-02
	<i>EGFLAM</i>	4.92	1.58	-0.65	1.30E-04	8.24E-02
	<i>SLC6A5</i>	2.59	0.18	-0.51	1.10E-04	8.24E-02
	<i>RAB31</i>	19.12	10.49	-0.60	1.06E-04	8.24E-02
	<i>TMEM107</i>	37.26	23.81	-0.56	1.24E-04	8.24E-02
	<i>C1S</i>	63.12	47.50	-0.43	1.37E-04	8.24E-02
	<i>C2orf27A</i>	2.37	0.50	-0.57	1.23E-04	8.24E-02
	<i>IGHG1</i>	171.81	74.48	-0.66	1.00E-04	8.24E-02
	<i>MUC5AC</i>	50.46	5.63	-0.54	1.30E-04	8.24E-02
	<i>HSPA9</i>	71.14	53.20	-0.40	1.52E-04	8.26E-02
	<i>TMEM132B</i>	6.82	2.55	-0.65	1.56E-04	8.26E-02
	<i>DCN</i>	39.88	23.71	-0.55	1.98E-04	8.40E-02
	<i>SNX10</i>	7.08	2.27	-0.64	1.99E-04	8.40E-02
	<i>AURKA</i>	10.06	3.66	-0.63	2.06E-04	8.40E-02
	<i>PEX11G</i>	1.32	3.29	0.64	1.76E-04	8.40E-02
<i>COL1A1</i>	445.08	285.52	-0.55	1.87E-04	8.40E-02	
<i>NEDD9</i>	84.09	112.98	0.38	1.96E-04	8.40E-02	
<i>WDR89</i>	11.37	5.46	-0.59	1.77E-04	8.40E-02	
<i>SLC25A25</i>	35.34	54.48	0.46	1.73E-04	8.40E-02	
<i>NR3C2</i>	113.10	154.48	0.41	1.95E-04	8.40E-02	
<i>C1R</i>	44.16	25.70	-0.53	2.09E-04	8.40E-02	
<i>MUC17</i>	107.73	269.60	0.63	2.12E-04	8.40E-02	
<i>RPS4Y1</i>	8.70	19.26	0.59	2.28E-04	8.82E-02	
<i>DOCK3</i>	6.29	2.66	-0.62	2.86E-04	9.55E-02	
<i>CHRAC1</i>	14.62	8.35	-0.51	2.90E-04	9.55E-02	
<i>SPATA17</i>	2.21	0.55	-0.58	2.82E-04	9.55E-02	
<i>NETO1</i>	3.40	0.46	-0.52	2.86E-04	9.55E-02	
<i>TRIL</i>	7.68	4.13	-0.60	2.83E-04	9.55E-02	
<i>DDX53</i>	1.70	0.06	-0.47	2.71E-04	9.55E-02	
<i>HIST1H2BN</i>	5.55	1.23	-0.61	2.74E-04	9.55E-02	
<i>PEG10</i>	6.38	2.44	-0.62	2.74E-04	9.55E-02	
<i>THY1</i>	19.83	10.36	-0.58	3.03E-04	9.65E-02	
<i>GRM2</i>	3.19	1.31	-0.59	3.07E-04	9.65E-02	
<i>TTC40</i>	4.30	0.77	-0.56	3.10E-04	9.65E-02	
<i>TIMP2</i>	94.56	68.12	-0.42	3.30E-04	9.71E-02	
<i>NOS1</i>	6.65	0.87	-0.51	3.48E-04	9.71E-02	
<i>CLSPN</i>	12.16	4.89	-0.61	3.54E-04	9.71E-02	
<i>C3</i>	102.46	57.73	-0.56	3.62E-04	9.71E-02	
<i>CCDC150</i>	7.05	3.32	-0.61	3.22E-04	9.71E-02	
<i>VWDE</i>	4.91	1.77	-0.61	3.50E-04	9.71E-02	
<i>CCT3</i>	36.04	24.11	-0.44	3.37E-04	9.71E-02	
<i>HSP90B1</i>	142.80	105.92	-0.38	3.61E-04	9.71E-02	
Processed meat	<i>C3</i>	101.05	50.36	-0.70	5.56E-06	6.44E-02
	<i>KRT1</i>	2.97	0.08	-0.60	7.40E-06	6.44E-02

FDR, false discovery rate.

^aAssociations are adjusted for age and sex.

Fig. 2



IPA networks associated with dysregulated genes on the basis of the level of meat-related dietary intake. (a) Cancer, organismal injury and abnormalities, and tumor morphology (score 31). (b) Cellular function and maintenance, cellular movement, cell death, and survival (score 31). (c) Drug metabolism, molecular transport, and small-molecule biochemistry (score 26). Red indicates upregulation and green indicates downregulation.

complement components may provide a protective effect against atherosclerosis [47,48]. Although atherosclerosis is an inflammatory disease, Hovland *et al.* [48] suggest that the classic complement pathway may be important for tissue homeostasis by the clearance of cell debris and immune complexes. Thus, it is possible that downregulation of proximal components of the classic pathway, such as components of C1, could contribute paradoxically toward the proinflammatory effects of red meat. However, upregulation of *MUC17* by meat could help explain its association with a number of factors, including inflammation [20,46].

Further evaluation of the genes associated with nonmeat-related variables showed that genes associated with

calcium were upregulated, two of the four genes associated with folate were upregulated, and the *TXNDC17* gene associated with prudent dietary pattern was upregulated. The genes associated with calcium consumption were involved in transmembrane transport, regulation of transcription and translation, cell survival, and the complement system. Calcium plays a key role in muscle contraction and relaxation, blood coagulation, nerve transmission, and keratinocyte differentiation [49]. The genes upregulated by high calcium consumption include *FOXJ2*, encoding for a transcription activator; *NECAP1*, whose product is involved in receptor-mediated endocytosis; and *C3AR1*, encoding for a C3A receptor. A prudent dietary pattern was associated with roughly three

servings of milk per day in one study [20] and was associated with a decreased risk of cancer and the B cell receptor pathways [20]. Our observed calcium-associated upregulation of tumor suppressors and a C3A receptor is consistent with these earlier findings.

Folate-rich foods include green vegetables and citrus fruits, and folate plays a key role as a coenzyme involved in DNA synthesis, amino acid metabolism, and methylations [50]. High versus low folate intake was associated with the upregulation of *SLC37A2* and *CLDN4* and the downregulation of *MUC5AC*. The gene products for *SLC37A2* and *CLDN4* are involved in carbohydrate transport and tight junction organization, respectively. Tight junction proteins, such as claudins, are related to NF- κ B, mTOR, and Nrf2 [51,52], whereas the expression of tight junction proteins is associated with amino acids, such as arginine and tryptophan [51,52]. As folic acid regulates amino acid metabolism, this could explain its role in modulating the expression of tight junction proteins. Dietary folate also downregulates the expression of *MUC5AC* as some mucins have been associated with inflammation [43]; this could help explain the reported association between folate and certain cancers and cardiovascular diseases [50]. We also found that a prudent dietary pattern upregulated *TXNDC17*, which modulates NF- κ B. NF- κ B signaling, plays a role in inflammation, apoptosis, and a variety of cellular stress responses [15]. *TXNDC17* overexpression has been shown to inhibit TNF- α -induced activation of NF- κ B [53]; thus, its upregulation could contribute toward the decrease in inflammation and cancer noted previously with a prudent dietary pattern [20].

We found that the majority of genes differentially expressed between intake levels of whole grains and vegetables were downregulated, especially for vegetable intake. The genes that are differentially expressed with level of whole grain and vegetable consumption are associated with NF- κ B signaling, regulation of apoptosis, cytoskeleton dynamics, and carbohydrate metabolism. Carbohydrates are important modulators of insulin action on glucose metabolism [54], which is consistent with our findings of *PFKL*, *ABCC3*, and *GMPPB* downregulated expression with whole grain consumption, and *HSPG2* and *REG4*, which were downregulated by high vegetable consumption. Both fiber and polyphenols attenuate inflammation [46]. Phenolic compounds are major bioactive compounds in whole grains and polyphenols are known to interfere with the NF- κ B signaling pathway [55], which is consistent with our data showing a downregulation of genes involved in the NF- κ B signaling. Our data suggest that the previously described polyphenolic perturbation of nuclear signaling and anti-inflammatory effects [56] may be explained by the differential gene expression observed with both whole grain and vegetable levels of intake.

Some genes were upregulated by high vegetable intake, including *ANP32A*, *NECAP1*, and *HELZ*. *ANP32A*, which

encodes for a tumor suppressor [57]; and *HELZ*, which encodes a RNA helicase, which has been shown in conjunction with other genes to regulate downstream genes as a transcription factor containing histone methyltransferase activity [58]. In association with SMYD3, *HELZ* plays a role in the proliferation of cancer cells [58]. It has also been suggested that *HELZ* plays a wider role in global translation activation and the phosphorylation (activation) of ribosomal protein S6, which in turn activates eIF4b [59]. This indicates that the upregulation of *HELZ* could help explain the association of vegetables with several factors, including inflammation [20,46], through the translational activation of numerous genes within the inflammatory process.

IPA found three interaction networks with meat consumption and three interaction networks associated with nonmeat consumption. This correlates with earlier IPA studies showing that the western dietary pattern is associated with five canonical pathways related to cancer, six pathways related to immune and/or inflammatory responses, and three pathways related to cardiovascular signaling and that the prudent dietary pattern is associated with nine IPA canonical pathways associated with immune/inflammatory response and six pathways associated with cancer [20]. It has been noted previously that western dietary patterns are associated with a proinflammatory gene expression profile and that prudent dietary patterns are associated with an anti-inflammatory gene expression profile [60], which would be consistent with our described networks. Moreover, our networks suggest that although both meat-related and nonmeat-related dietary variables are associated with the cellular functions of intracellular transport and ECM regulation, the signal transduction pathways differ between meat and nonmeat networks. We found meat consumption to be associated with TP53, NF κ B1, and MAPK signaling through the ERK1/2 signaling pathway. The Ras-MAPK system is involved in cell proliferation [61]; moreover, our findings in this network are consistent with earlier findings that high-fat diets, the cooked meat carcinogen 2-amino-1-methyl-6-phenylimidazo [4,5-b]-pyridine (PhiP), and sodium nitrate can activate the MAPK pathways [9,16, 62]. Nonmeat consumption appears to play a greater role in the regulation of the PI3K complex, TNF-signaling, and TGF β 1-signaling pathways. All of these pathways have been identified as major signaling pathways for colon cancer [63–66].

To the best of our knowledge, this is one of the few studies to examine the role of diet in gene expression using data from a population-based study and comprehensive gene expression data. Our study has a number of strengths including using RNA-Seq, which produces global gene expression data for each RNA sample and is an ideal method to carry out a discovery study such as ours [67,68]. However, it is important to keep in mind that the gene expression profile is most relevant to

current diet exposure. The time between tissue ascertainment and the referent period for diet in our study could be from several months to 3 or possibly 4 years. Although the lack of findings could indicate a disparity in time between exposure and tissue sample acquirement, finding associations would imply that the exposure is recent enough, in that dietary patterns are consistent over time, to alter the expression. In addition, we have utilized normal colonic mucosa; thus, genes would have to be expressed in colon tissue for detection. It should be recognized that different platforms carry different technical strengths and weaknesses that can influence results. Thus, it is essential to validate these findings in other populations using the methods that we used here as well as other platforms to better understand associations between diet and gene expression. We utilized DESeq2 to assess gene expression data adjusting for age and sex. We have previously shown associations between cigarette smoking and alcohol and gene expression in these samples [69]. Similarly, we report statistically significant differential expression as an indicator of dysregulated genes; however, the level of dysregulation that is necessary to result in functional significance is not clear.

Although this study was carried out in a rigorous manner, there are limitations. Our dietary data were collected shortly after diagnosis. Although our dietary questionnaire allowed for reporting of over 800 food items and we implemented extremely rigid QC procedures, it is possible that recall could have been influenced by the disease status. Although this is a possibility, it should be noted that our risk estimates for diet and colon cancer are almost identical to those reported by large cohort studies [70]. In addition, these data reflect foods consumed in the early 1990s. Although we believe that individuals still eat the foods that we report here, different dietary patterns may be more important today than the western and prudent diet that we report.

Our data support the hypothesis that diet is associated with gene expression. Our data showed that diet was involved in deregulation of genes involved in numerous pathways and functions such as Wnt-signaling, MAPK, NF- κ B, TP53, ECM maintenance, cytoskeletal structure, and cell cycle regulation. These data suggest that differential gene expression associated with dietary intake is a possible mechanism by which diet can influence a variety of biological processes and functions. Focused functionality studies evaluating dietary influence on gene expression are needed.

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. The authors would like to acknowledge the contributions of Dr Bette Caan, Judy Morse, and Donna Schaffer, the Kaiser Permanente Medical Research Program, 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.

Author contributions: A.P. compiled and interpreted the data and wrote the manuscript; D.P. carried out statistical analysis, edited the manuscript, and approved the final version of manuscript; L.M. carried out bioinformatics analysis and approved the final manuscript; and R.W. oversaw RNAseq data collection and approved the final manuscript. M.S. obtained funding, collected data for study, assisted in data interpretation and analysis, edited, and approved the final manuscript.

Conflicts of interest

There are no conflicts of interest.

References

- McDaniel J, Askew W, Bennett D, Mihalopoulos J, Anantharaman S, Fjeldstad AS, *et al.* Bison meat has a lower atherogenic risk than beef in healthy men. *Nutr Res* 2013; **33**:293–302.
- Torres N, Guevara-Cruz M, Velazquez-Villegas LA, Tovar AR. Nutrition and atherosclerosis. *Arch Med Res* 2015; **46**:408–426.
- Sarbagili-Shabat C, Sigall-Boneh R, Levine A. Nutritional therapy in inflammatory bowel disease. *Curr Opin Gastroenterol* 2015; **31**:303–308.
- Yang CQ, Shu L, Wang S, Wang JJ, Zhou Y, Xuan YJ, Wang SF. Dietary patterns modulate the risk of non-alcoholic fatty liver disease in Chinese adults. *Nutrients* 2015; **7**:4778–4791.
- Garcia Caraballo SC, Comhair TM, Houten SM, Dejong CH, Lamers WH, Koehler SE. High-protein diets prevent steatosis and induce hepatic accumulation of monomethyl branched-chain fatty acids. *J Nutr Biochem* 2014; **25**:1263–1274.
- Jeong EA, Jeon BT, Shin HJ, Kim N, Lee DH, Kim HJ, *et al.* Ketogenic diet-induced peroxisome proliferator-activated receptor- γ activation decreases neuroinflammation in the mouse hippocampus after kainic acid-induced seizures. *Exp Neurol* 2011; **232**:195–202.
- Zhu QC, Gao RY, Wu W, Guo BM, Peng JY, Qin HL. Effect of a high-fat diet in development of colonic adenoma in an animal model. *World J Gastroenterol* 2014; **20**:8119–8129.
- Lauber SN, Ali S, Gooderham NJ. The cooked food derived carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine is a potent oestrogen: a mechanistic basis for its tissue-specific carcinogenicity. *Carcinogenesis* 2004; **25**:2509–2517.
- Wang Z, Pei H, Kaec M, Lu J. Mammary cancer promotion and MAPK activation associated with consumption of a corn oil-based high-fat diet. *Nutr Cancer* 1999; **34**:140–146.
- Marangoni F, Corsello G, Cricelli C, Ferrara N, Ghiselli A, Lucchin L, Poli A. Role of poultry meat in a balanced diet aimed at maintaining health and wellbeing: an Italian consensus document. *Food Nutr Res* 2015; **59**:27606.
- Kim HJ, Yoon HJ, Kim SY, Yoon YR. A medium-chain fatty acid, capric acid, inhibits RANKL-induced osteoclast differentiation via the suppression of NF- κ B signaling and blocks cytoskeletal organization and survival in mature osteoclasts. *Mol Cells* 2014; **37**:598–604.
- Rafi MM, Shafaie Y. Dietary lutein modulates inducible nitric oxide synthase (iNOS) gene and protein expression in mouse macrophage cells (RAW 264.7). *Mol Nutr Food Res* 2007; **51**:333–340.
- Cao HH, Cheng CY, Su T, Fu XQ, Guo H, Li T, *et al.* Quercetin inhibits HGF/c-Met signaling and HGF-stimulated melanoma cell migration and invasion. *Mol Cancer* 2015; **14**:103.
- Lee WJ, Chen WK, Wang CJ, Lin WL, Tseng TH. Apigenin inhibits HGF-promoted invasive growth and metastasis involving blocking PI3K/Akt pathway and beta 4 integrin function in MDA-MB-231 breast cancer cells. *Toxicol Appl Pharmacol* 2008; **226**:178–191.
- Haar L, Ren X, Liu Y, Koch SE, Goines J, Tranter M, *et al.* Acute consumption of a high-fat diet prior to ischemia-reperfusion results in cardioprotection through NF- κ B-dependent regulation of autophagic pathways. *Am J Physiol Heart Circ Physiol* 2014; **307**:H1705–H1713.

- 16 Al-Gayyar MM, Al Youssef A, Sherif IO, Shams ME, Abbas A. Protective effects of arjunolic acid against cardiac toxicity induced by oral sodium nitrite: effects on cytokine balance and apoptosis. *Life Sci* 2014; **111** (1–2):18–26.
- 17 Ijssennagger N, Rijnsma A, de Wit NJ, Boekschoten MV, Dekker J, Schonewille A, *et al.* Dietary heme induces acute oxidative stress, but delayed cytotoxicity and compensatory hyperproliferation in mouse colon. *Carcinogenesis* 2013; **34**:1628–1635.
- 18 Guo X, Yin S, Dong Y, Fan L, Ye M, Lu J, Hu H. Enhanced apoptotic effects by the combination of curcumin and methylseleninic acid: potential role of Mcl-1 and FAK. *Mol Carcinog* 2013; **52**:879–889.
- 19 Hebels DG, Svejce KM, de Kok MC, van Herwijnen MH, Kuhnle GG, Engels LG, *et al.* Red meat intake-induced increases in fecal water genotoxicity correlate with pro-carcinogenic gene expression changes in the human colon. *Food Chem Toxicol* 2012; **50**:95–103.
- 20 Bouchard-Mercier A, Paradis AM, Rudkowska I, Lemieux S, Couture P, Vohl MC. Associations between dietary patterns and gene expression profiles of healthy men and women: a cross-sectional study. *Nutr J* 2013; **12**:24.
- 21 Slattery ML, Potter J, Caan B, Edwards S, Coates A, Ma KN, Berry TD. Energy balance and colon cancer – beyond physical activity. *Cancer Res* 1997; **57**:75–80.
- 22 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** (1–2):7–14.
- 23 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**:1020–1028.
- 24 Liu K, Slattery M, Jacobs D Jr, Cutter G, McDonald A, Van Horn L, *et al.* A study of the reliability and comparative validity of the cardia dietary history. *Ethn Dis* 1994; **4**:15–27.
- 25 Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. *Am J Epidemiol* 1998; **148**:4–16.
- 26 Kampman E, Slattery ML, Bigler J, Leppert M, Samowitz W, Caan BJ, Potter JD. Meat consumption, genetic susceptibility, and colon cancer risk: a United States multicenter case-control study. *Cancer Epidemiol Biomarkers Prev* 1999; **8**:15–24.
- 27 Li P, Conley A, Zhang H, Kim HL. Whole-Transcriptome profiling of formalin-fixed, paraffin-embedded renal cell carcinoma by RNA-seq. *BMC Genomics* 2014; **15**:1087.
- 28 Slattery ML, Herrick JS, Mullany LE, Gertz J, Wolff RK. Improved survival among colon cancer patients with increased differentially expressed pathways. *BMC Med* 2015; **13**:75.
- 29 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc* 1995; **57**:289–300.
- 30 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**:550.
- 31 QIAGEN's Ingenuity Pathway Analysis; 2014, Valencia, California, USA. Available at: <http://www.ingenuity.com/products/ipa>. [Accessed March 2016].
- 32 Savli H, Szendrői A, Romics I, Nagy B. Gene network and canonical pathway analysis in prostate cancer: a microarray study. *Exp Mol Med* 2008; **40**:176–185.
- 33 Li Y, Carrillo JA, Ding Y, He Y, Zhao C, Zan L, Song J. Ruminant transcriptomic analysis of grass-fed and grain-fed angus beef cattle. *PLoS One* 2015; **10**: e0116437.
- 34 Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, *et al.* Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet* 2009; **18**:2078–2090.
- 35 Yan-Fang T, Dong W, Li P, Wen-Li Z, Jun L, Na W, *et al.* Analyzing the gene expression profile of pediatric acute myeloid leukemia with real-time PCR arrays. *Cancer Cell Int* 2012; **12**:40.
- 36 Naito Y, Kuroda M, Mizushima K, Takagi T, Handa O, Kokura S, *et al.* Transcriptome analysis for cytoprotective actions of rebamipide against indomethacin-induced gastric mucosal injury in rats. *J Clin Biochem Nutr* 2007; **41**:202–210.
- 37 Reyes-Gibby CC, Yuan C, Wang J, Yeung SC, Shete S. Gene network analysis shows immune-signaling and ERK1/2 as novel genetic markers for multiple addiction phenotypes: alcohol, smoking and opioid addiction. *BMC Syst Biol* 2015; **9**:25.
- 38 Jia P, Kao CF, Kuo PH, Zhao Z. A comprehensive network and pathway analysis of candidate genes in major depressive disorder. *BMC Syst Biol* 2011; **5** (Suppl 3):S12.
- 39 Stolz A, Bastians H. Fresh WNT into the regulation of mitosis. *Cell Cycle* 2015; **14**:2566–2570.
- 40 Lv ZD, Zhang L, Liu XP, Jin LY, Dong Q, Li FN, *et al.* NKD1 down-regulation is associated with poor prognosis in breast invasive ductal carcinoma. *Int J Clin Exp Pathol* 2015; **8**:4015–4021.
- 41 Piccardi B, Palumbo V, Nesi M, Nencini P, Gori AM, Giusti B, *et al.* Unbalanced metalloproteinase-9 and tissue inhibitors of metalloproteinases ratios predict hemorrhagic transformation of lesion in ischemic stroke patients treated with thrombolysis: results from the MAGIC study. *Front Neurol* 2015; **6**:121.
- 42 Hannan MT, Tucker KL, Dawson-Hughes B, Cupples LA, Felson DT, Kiel DP. Effect of dietary protein on bone loss in elderly men and women: the Framingham Osteoporosis Study. *J Bone Miner Res* 2000; **15**:2504–2512.
- 43 Ng GZ, Menheniott TR, Every AL, Stent A, Judd LM, Chionh YT, *et al.* The MUC1 mucin protects against *Helicobacter pylori* pathogenesis in mice by regulation of the NLRP3 inflammasome. *Gut* 2015. [Epub ahead of print].
- 44 Shibahara H, Higashi M, Yokoyama S, Rousseau K, Kitazono I, Osako M, *et al.* A comprehensive expression analysis of mucins in appendiceal carcinoma in a multicenter study: MUC3 is a novel prognostic factor. *PLoS One* 2014; **9**:e115613.
- 45 Nardinocchi L, Sonogo G, Passarelli F, Avitabile S, Scarponi C, Failla CM, *et al.* Interleukin-17 and interleukin-22 promote tumor progression in human non-melanoma skin cancer. *Eur J Immunol* 2015; **45**:922–931.
- 46 Andersen V, Olsen A, Carbone L, Tjønneland A, Vogel U. Diet and risk of inflammatory bowel disease. *Dig Liver Dis* 2012; **44**:185–194.
- 47 Vlaicu SI, Tatomir A, Rus V, Mekala AP, Mircea PA, Niculescu F, Rus H. The role of complement activation in atherogenesis: the first 40 years. *Immunol Res* 2016; **64**:1–13.
- 48 Hovland A, Jonasson L, Garred P, Yndestad A, Aukrust P, Lappegård KT, *et al.* The complement system and toll-like receptors as integrated players in the pathophysiology of atherosclerosis. *Atherosclerosis* 2015; **241**:480–494.
- 49 Jiang Y, Liao L, Shrestha C, Li D, Li M, Mu Y, *et al.* Inhibition of 4-nitroquinoline-1-oxide-induced oral carcinogenesis by dietary calcium. *Int J Clin Exp Pathol* 2015; **8**:3529–3542.
- 50 Mönch S, Netzel M, Netzel G, Ott U, Frank T, Rychlik M. Folate bioavailability from foods rich in folates assessed in a short term human study using stable isotope dilution assays. *Food Funct* 2015; **6**:242–248.
- 51 Wang B, Feng L, Jiang WD, Wu P, Kuang SY, Jiang J, *et al.* Copper-induced tight junction mRNA expression changes, apoptosis and antioxidant responses via NF- κ B, TOR and Nrf2 signaling molecules in the gills of fish: preventive role of arginine. *Aquat Toxicol* 2015; **158**:125–137.
- 52 Jiang WD, Wen HL, Liu Y, Jiang J, Kuang SY, Wu P, *et al.* The tight junction protein abundance changes and oxidative damage by tryptophan deficiency or excess are related to the modulation of the signalling molecules, NF- κ B p65, TOR, caspase-(3,8,9) and Nrf2 mRNA levels, in the gill of young grass carp (*Ctenopharyngodon idellus*). *Fish Shellfish Immunol* 2015; **46**:168–180.
- 53 Jeong W, Chang TS, Boja ES, Fales HM, Rhee SG. Roles of TRP14, a thioredoxin-related protein in tumor necrosis factor- α signaling pathways. *J Biol Chem* 2004; **279**:3151–3159.
- 54 Tremblay F, Lavigne C, Jacques H, Marette A. Dietary cod protein restores insulin-induced activation of phosphatidylinositol 3-kinase/Akt and GLUT4 translocation to the T-tubules in skeletal muscle of high-fat-fed obese rats. *Diabetes* 2003; **52**:29–37.
- 55 Kim YH, Kim JL, Lee EJ, Park SH, Han SY, Kang SA, Kang YH. Fisetin antagonizes cell fusion, cytoskeletal organization and bone resorption in RANKL-differentiated murine macrophages. *J Nutr Biochem* 2014; **25**:295–303.
- 56 Farzaei MH, Abdollahi M, Rahimi R. Role of dietary polyphenols in the management of peptic ulcer. *World J Gastroenterol* 2015; **21**:6499–6517.
- 57 Buddaseth S, Göttmann W, Blaszczyk R, Huyton T. Dysregulation of cell cycle control caused by overexpression of the oncogene pp32r1 (ANP32C) and the Tyr > His mutant pp32r1Y140H. *Biochim Biophys Acta* 2013; **1833**:1212–1221.
- 58 Hamamoto R, Furukawa Y, Morita M, Imura Y, Silva FP, Li M, *et al.* SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol* 2004; **6**:731–740.
- 59 Hasgall PA, Hoogewijs D, Faza MB, Panse VG, Wenger RH, Camenisch G. The putative RNA helicase HELZ promotes cell proliferation, translation initiation and ribosomal protein S6 phosphorylation. *PLoS One* 2011; **6**: e22107.
- 60 van Dijk SJ, Feskens EJ, Bos MB, Hoelen DW, Heijligenberg R, Bromhaar MG, *et al.* A saturated fatty acid-rich diet induces an obesity-linked proinflammatory gene expression profile in adipose tissue of subjects at risk of metabolic syndrome. *Am J Clin Nutr* 2009; **90**:1656–1664.

- 61 Walczak K, Turski WA, Rajtar G. Kynurenic acid inhibits colon cancer proliferation in vitro: effects on signaling pathways. *Amino Acids* 2014; **46**:2393–2401.
- 62 Creton SK, Zhu H, Gooderham NJ. The cooked meat carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine activates the extracellular signal regulated kinase mitogen-activated protein kinase pathway. *Cancer Res* 2007; **67**:11455–11462.
- 63 Slattery ML, Lundgreen A, Herrick JS, Caan BJ, Potter JD, Wolff RK. Associations between genetic variation in RUNX1, RUNX2, RUNX3, MAPK1 and eIF4E and risk of colon and rectal cancer: additional support for a TGF- β -signaling pathway. *Carcinogenesis* 2011; **32**:318–326.
- 64 Slattery ML, Lundgreen A, Herrick JS, Wolff RK. Genetic variation in RPS6KA1, RPS6KA2, RPS6KB1, RPS6KB2, and PDK1 and risk of colon or rectal cancer. *Mutat Res* 2011; **706**:13–20.
- 65 Slattery ML, Herrick JS, 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.
- 66 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**:328–338.
- 67 Dorr C, Wu B, Guan W, Muthusamy A, Sanghavi K, Schladt DP, *et al.* Differentially expressed gene transcripts using RNA sequencing from the blood of immunosuppressed kidney allograft recipients. *PLoS One* 2015; **10**: e0125045.
- 68 Guo Y, Sheng Q, Li J, Ye F, Samuels DC, Shyr Y. Large scale comparison of gene expression levels by microarrays and RNAseq using TCGA data. *PLoS One* 2013; **8**:e71462.
- 69 Slattery ML, Pellatt DF, Mullany LE, Wolff RK. Differential gene expression in colon tissue associated with diet, lifestyle, and related oxidative stress. *PLoS One* 2015; **10**:e0134406.
- 70 Figueiredo JC, Hsu L, Hutter CM, Lin Y, Campbell PT, Baron JA, *et al.* Genome-wide diet-gene interaction analyses for risk of colorectal cancer. *PLoS Genet* 2014; **10**:e1004228.