Published in final edited form as:
Genes Chromosomes Cancer. 2014 June ; 53(6): 454–466.

The Effect of *UGT1A* and *UGT2B* Polymorphisms on Colorectal Cancer Risk: Haplotype Associations and Gene-Environment Interactions

Andrea Y. Angstadt[1], **Terryl J. Hartman**[2], **Samuel M. Lesko**[3], **Joshua E. Muscat**[4], **Junjia Zhu**[1], **Carla J. Gallagher**[1,4,*], and **Philip Lazarus**[5,*]

[1]Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA

[2]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA

[3]Northeast Regional Cancer Institute, Scranton, PA

[4]Department of Pharmacology, Penn State College of Medicine, Hershey, PA

[5]Department of Pharmaceutical Sciences, College of Pharmacy, Washington State University, Spokane, WA

Abstract

UDP-glucuronosyltransferases (UGTs) play an important role in the phase II metabolism of exogenous and endogenous compounds. As colorectal cancer (CRC) etiology is thought to involve the biotransformation of dietary factors, UGT polymorphisms may affect CRC risk by altering levels of exposure. Genotyping of over 1800 Caucasian subjects was completed to identify the role of genetic variation in nine *UGT1A* and five *UGT2B* genes on CRC risk. Unconditional logistic regression and haplotype analyses were conducted to identify associations with CRC risk and potential gene-environment interactions. *UGT1A* haplotype analysis found that the T-G haplotype in *UGT1A10* exon 1 (block 2: rs17864678, rs10929251) decreased colon cancer risk [proximal (OR = 0.28, 95% CI=0.11–0.69), distal (OR = 0.32, 95% CI=0.12–0.91)] and that the C-T-G haplotype in the 3′ region flanking the *UGT1A* shared exons (block 11: rs7578153, rs10203853, rs6728940) increased CRC risk in males (OR = 2.56, 95% CI=1.10–5.95). A haplotype in *UGT2B15* containing a functional variant (rs4148269, K523T) and an intronic SNP (rs6837575) was found to affect rectal cancer risk overall (OR = 2.57, 95% CI=1.21–5.04) and in females (OR = 3.08, 95% CI=1.08–8.74). An interaction was found between high NSAID use and the A-G-T haplotype (block 10: rs6717546, rs1500482, rs7586006) in the *UGT1A* shared exons that decreased CRC risk. This suggests that UGT genetic variation alters CRC risk differently by anatomical sub-site and gender and that polymorphisms in the *UGT1A* shared exons may have a regulatory effect on gene expression that allows for the protective effect of NSAIDs on CRC risk.

[*]Correspondence to: Philip Lazarus, PhD, Department of Pharmaceutical Sciences, College of Pharmacy, Washington State University, P.O. Box 1495, Spokane, WA 99210, Phone: 509-358-7947, phil.lazarus@wsu.edu or Carla J. Gallagher, PhD, Department of Public Health Sciences, College of Medicine, Penn State University, 500 University Drive, CH69, Hershey, PA 17033, Tel: (717) 531-2973, Fax: (717) 531-0480, cgallagher@hmc.psu.edu.

INTRODUCTION

Phase II enzymes play an important role in the biotransformation of many exogenous and endogenous compounds including drugs, dietary compounds, environmental carcinogens, and hormones (Zheng et al., 2002; Turgeon et al., 2003; Gallagher et al., 2007b). Because these enzymes play such an important role in metabolism, it is thought that polymorphisms in phase II enzymes may affect disease risk. Colorectal cancer (CRC) is one such multifactorial disease with several hypothesized etiologic factors involving the biotransformation of dietary and environmental factors, carcinogens, and hormones (Slattery and Fitzpatrick, 2009). More specifically, CRC has been associated with the consumption of meat that is cooked using high-temperature methods, such as grilling, barbequing, and pan-frying (Larsson and Wolk, 2006; Wu et al., 2006; Miller et al., 2013) which produces the carcinogenic heterocyclic amines (HCAs) (Sinha and Rothman, 1997) and polycyclic aromatic hyodrocarbons (PAHs) (Sinha et al., 2005). Many of the most potent carcinogens within these groups of agents are extensively metabolized by the UDP-glucuronosyltransferase (UGT)-mediated phase II glucuronidation metabolic pathway. UGTs also play an important role in the metabolism of various potentially beneficial agents including antioxidants like flavonoids (Turgeon et al., 2003) and NSAIDs (Kuehl et al., 2005; Kuehl et al., 2006). Therefore, it is possible that CRC risk is affected by dietary components in combination with genetic polymorphisms in the UGT enzyme family.

The UGT superfamily is comprised of two major sub-families (UGT1 and UGT2) (Mackenzie et al., 1997). The single *UGT1A* gene locus is physically located on chromosome band 2q37 and encodes nine functional proteins: UGT1A1, UGT1A3, UGT1A4, UGT1A5, UGT1A6, UGT1A7, UGT1A8, UGT1A9, and UGT1A10 (Mackenzie et al., 2003). The *UGT1A* gene locus is composed of the thirteen first exons on the 5′ end linked, by alternative splicing, to four different common exons on the 3′ end (Gong et al., 2001). The *UGT2B* gene locus is located on chromosome band 4q13 and is comprised of the following functional proteins: UGT2B4, UGT2B7, UGT2B10, UGT2B11, UGT2B15, UGT2B17, and UGT2B28 (Guillemette 2003). The UGT1A and 2B families have been found to be well-expressed in the liver and in tissues of the gastrointestinal and aereodigestive tracts (McDonnell et al., 1996; Giuliani et al., 2001; Gestl et al., 2002; Zheng et al., 2002) with three *UGT1A* genes (*1A7*, *1A8*, and *1A10*) expressed exclusively extra-hepatically (Strassburg et al., 1997; Strassburg et al., 1998), and *UGTs 1A1*, *1A3*, *1A4*, *1A5*, *1A6*, *1A7*, *1A8*, *1A10*, *2B7*, *2B11*, *2B15*, and *2B17* all found to be expressed in the colon (Nakamura et al., 2008).

Sequencing and genotyping data have led to the discovery of over 100 individual variants within the *UGT* genes. Some of these variants in the *UGT1A* gene locus exhibit allele frequencies up to 40–50% in the general population, many of which are found to be in linkage disequilibrium (Maitland et al., 2006). Functional variants have been found in the coding regions and/or promoters of *UGTs 1A1*, *1A3*, *1A4*, *1A7*, *1A8*, *1A9*, *1A10*, *2B4*, *2B7*, *2B10*, *2B15* and *2B17* (Levesque et al., 1997; Levesque et al., 1999; Mackenzie et al., 2000; Miners et al., 2002; Jinno et al., 2003; Villeneuve et al., 2003; Bernard and Guillemette 2004; Duguay et al., 2004; Ehmer et al., 2004; Iwai et al., 2004; Krishnaswamy et al., 2005; Gallagher et al., 2007b; Korprasertthaworn et al., 2012), but as many polymorphisms are

inherited together it is often difficult to identify the actual polymorphism that contributes to the adverse/beneficial effects. Previous case-control studies indicated an increased risk of developing CRC for individuals carrying the *UGT1A1*6* and *UGT1A7*3* variants (Tang et al., 2005). In contrast, it was recently demonstrated that the *UGT2B17* deletion genotype was associated with a decrease in CRC risk (Angstadt et al., 2013).

The goal of the present study was to examine comprehensively the effect of genetic variation in the *UGT1A* and *UGT2B* gene loci on CRC risk. Association studies were conducted by genotyping SNPs (tagSNPs, coding SNPs, and additional SNPs to ensure no gap larger than 10 kb) within and surrounding the nine *UGT1A* genes on 2q37, and the genes encoding UGTs 2B4, 2B7, 2B10, 2B11, and 2B15 on 4q13, in a case-control study of over 1800 Caucasian subjects from central and northeast Pennsylvania, a region deemed at high-risk for CRC (Alpert et al., 2007). Association studies were conducted controlling for known CRC risk factors (Bailey et al., 2009; Wang and Beydoun, 2009) identified using the extensive demographic and dietary data collected from study participants, and interactions between SNPs/haplotypes and HCAs, PAHs, and non-steroidal anti-inflammatory drug (NSAID) use were analyzed to assess whether any gene-environmental relationships were affecting CRC risk in this study population.

MATERIALS AND METHODS

Subjects

Genotyping and association analysis was conducted on individuals from a population-based case-control study conducted from 2006–2011 to investigate CRC risk factors in a contiguous 19-county area in central and northeast Pennsylvania. Potential incident cases in this 19-county region were identified from the Pennsylvania State Cancer Registry and were notified about the study by letter, followed by a telephone call from a study coordinator to explain further the study and answer questions. All cases were newly diagnosed and recruited within 24 months of their CRC diagnosis. The tumors were classified by anatomical site and the histological code of the International Classification of Diseases for Oncology (Organization 2000), including codes C180–C189, C209, and C260. The anatomical sub-site codes are as follows; proximal colon (C180–C184), distal colon (C185–C189), and rectum (C199 and C209). Random digit dialing was used to identify controls residing in the same 19-county region as the cases, described by Waksberg (1978), and they were screened to ensure they had no previous history of cancer. Controls were frequency matched to cases based on sex, age, and race. Of those contacted, 57% of eligible cases and 51% of eligible controls participated in the study and provided a DNA sample.

Written consent was obtained from all participants, a personal interview was scheduled at the home of the participant, and a self-administered food frequency questionnaire was mailed with instructions to complete it before the interview. Data on socio-demographic factors, medical history, alcohol use, lifetime tobacco exposure, physical activity, height, weight, medication use, and other lifestyle-related factors were collected by trained interviewers during the in-person interviews. For health and lifestyle-related factors, such as weight, diet, and physical activity, data prior to diagnosis were collected for cases. Participants completed a modified version of the Diet History Questionnaire (DHQ), a

validated Food Frequency Questionnaire (FFQ) developed by the National Cancer Institute (NCI) (Thompson et al., 2002). The modifications aimed to capture the distinct meat eating patterns of Pennsylvania residents in this catchment area and included the addition of processed meat items commonly consumed in this population, such as specific German sausages, Italian cured meats, and corned beef. The reference period was a year prior to the interview for controls and a year prior to diagnosis for cases. Visual materials were provided to facilitate the correct recall of portion sizes and food preferences. Questions on preferred meat cooking methods, doneness levels for individual meat subtypes, and intake of processed meat items were used to generate estimated exposure to meat-derived mutagens with the NCI's Computerized Heterocyclic Amines Resource for Research in Epidemiology of Disease (CHARRED) software application (Friday and Bowman, 2006). The CHARRED program estimates exposure (nanograms/day) to HCAs, 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP), 2-amino-3,8-dimethylimidazo[4,5,-f]quinoxaline (MeIQx) and 2-amino-3,4,8-trimethylimidazo[4,5,-f]quinoxaline (DiMeIQx), and one PAH, benzo(a)pyrene (BaP). In addition, the CHARRED application generated total mutagenic activity (revertant colonies/day), a measure of overall mutagenic potential that accounts for differences in mutagenic activity between the various compounds.

SNP Selection and genotyping

Genotypes for SNPs in the *UGT1A* and *UGT2B* genes representing people with European Ancestry (CEPH) were downloaded from the International HapMap Project (Altshuler et al., 2005). Linkage disequilibrium (LD) in the *UGT1A* and *UGT2B* genes was determined using Haploview Software (Barrett et al., 2005). LD was estimated between all pairs of SNPs using the D′ statistic, and haplotype block structure was determined using the Solid Spine of LD option, with the block extended if pairwise D′ between SNPs was greater than 0.80. From this analysis, 96 tag, coding, and additional SNPs within and surrounding the nine *UGT1A* genes on 2q37 to fill in gaps greater than 10 kb were used to design an Illumina GoldenGate genotyping assay (Illumina San Diego, CA). A separate GoldenGate genotyping assay was designed containing 16 tagging and coding SNPs and one deletion/insertion polymorphism within five *UGT2B* genes on 4q13; two UGT2B genes, *UGTs 2B17* and *2B28*, were not profiled by the SNPs examined in the present study as these genes had been previously examined for associations with CRC risk using the same sample set (Angstadt et al., 2013).

Oral buccal cell swabs, saliva, and/or blood samples were collected for genomic DNA (gDNA) isolation and gDNA was isolated from oral buccal cell swabs using standard phenol: chloroform isolation. gDNA from saliva was isolated using an Oragene DNA Kit (DNA Genotek Inc, Ontario, Canada) while blood gDNA was isolated using QIAamp DNA Blood mini kit (Qiagen, Valencia, CA). Picogreen analysis was used to quantify the amount of double stranded DNA (dsDNA) for each genomic sample (Life Technologies, Grand Island, NY).

Statistical Analysis

Dietary characteristics between cases and controls were compared using the χ^2-test for categorical variables and non-parametric Wilcoxon Rank-Sum test for continuous variables.

If continuous dietary variables appeared non-normally distributed then the appropriate transformation was performed (for example, log-transformation) to normalize their distributions. T-tests were used on transformed data and then confirmed by the Wilcoxon Rank-Sum test. Likelihood ratio tests were used to evaluate the fit of each model. A total of 854 Caucasian cases and 969 Caucasian controls were genotyped on the *UGT1A* assay, and 897 Caucasian cases and 955 Caucasian controls on the *UGT2B* assay. Although the final number of genotyped samples is similar between datasets, 470 of the samples in the *UGT2B* dataset do not overlap with the *UGT1A* dataset due to limited to quantities of DNA. In order to control for implausible dietary data, individuals who reported < 500 or >5000 kcal/day (n = 53 for *UGT1A*, 50 for *UGT2B*) were excluded from the analysis along with individuals ≤ 35 years of age (n = 13). After this exclusion, a total of 816 cases and 941 controls were analyzed from the *UGT1A* assay and 857 cases and 932 controls were analyzed from the *UGT2B* assay. The *UGT1A* and *UGT2B* data sets were analyzed independently in all types of analyses.

Hardy-Weinberg Equilibrium (HWE), allele frequencies, and identification of haplotype blocks in the study dataset were conducted using the control sample set in the Haploview software, defining blocks by the solid spine of LD. SNPs were excluded if the call rate was <90% and/or a Hardy-Weinberg Equilibrium $p < 1 \times 10^{-3}$. This study is powered to detect associations with SNPs as low as 5% frequency for an OR >2.0 (>95% power) and can additionally detect lower effect sizes (OR>1.5) for common SNPs (>30%) with >80% power. The SAS PROC HAPLOTYPE procedure (Czika and Yu 2004) was used to conduct separate haplotype analysis on the *UGT1A* and *UGT2B* sample sets, using the haplotype block definitions from the Haploview software. The procedure utilizes the Expectation Maximization (EM) algorithm to generate maximum likelihood estimates of haplotype frequencies given a multilocus sample of genetic marker genotypes under the assumption of HWE. The initializing method was INIT=RANDOM, which initializes haplotype frequencies with random values from a Uniform (0,1) distribution. The haplotype frequency threshold was set to 5%, and haplotypes with a lower frequency were excluded from subsequent logistic regression analysis. The standard errors and the confidence intervals for each haplotype were estimated under binomial assumption, by default. The total probability of an individual having a particular haplotype compared with all other haplotype possibilities was determined. These values were used in the following statistical analysis assuming an additive statistical model (comparing the probability of one haplotype with all other haplotypes combined).

Unconditional logistic regression models were used to calculate odds ratios (ORs) and 95% confidence intervals (CIs) for associations between individual SNPs and haplotypes and CRC risk. Three statistical models were tested for individual SNP logistic regression analysis: additive: (BB) > (BA) > (AA), dominant: [(BB + (BA)] vs (AA), and recessive: (BB) vs [(BA) + (AA)], with B being the minor allele (MA). Multivariate models were used adjusting for potential confounding variables that were selected *a priori*: age (continuous), sex (male, female), total energy intake (kcal/d, continuous), body mass index (BMI; kg/m^2, continuous), smoking status (never, current, or former), family history of CRC (yes, no; first degree relative), alcohol (g/d, continuous), physical activity (yes, no; ≥ 1 hr/week of

vigorous activity), education (no college degree, college degree or above), and regular non-steroidal anti-inflammatory drug (NSAID) use (yes, no; regular use defined as at least 3 times a week for 1 year prior to diagnosis for cases and 1 year prior to interview for controls). According to a 10% change-in-estimate criterion, age, education, sex, BMI, family history, NSAID use, and physical activity were included in the final multivariate model. Dietary carcinogen levels (PhIP, MeIQx, DiMeIQx, and BaP) were adjusted for total energy intake (kcal/day) by the nutrient density method (grams per 1000 kcal) (Willett et al., 1997) and separated into quintiles of intake based on the distribution among controls. The lowest level (quintile 1) for each respective dietary carcinogen served as the referent quintile. Since the *UGT* gene family is known to metabolize environmental carcinogens, high intake of each carcinogen was individually tested as a possible cofounding covariate and added into the model if necessary according to a 10% change-in-estimate criterion (Greenland and Rothman 2008). Associations stratified by sex, anatomical sub-site (colon and rectum), and high carcinogen intake (quintile 5 of each dietary carcinogen) were also investigated. Reported *P*-values are 2-sided after correcting for the effects of multiple testing within each genomic region using SAS PROC MULTITEST. The False Discovery Rate (FDR) method described by Benjamini and Hochberg (1995) was used and controls the rate at $\text{level} \leq \frac{m_0}{m}\alpha \leq \alpha$ when you have independent *P*-values that are uniformly distributed under their respective null hypotheses. An adjusted *P*-value < 0.05 was considered significant for all tests.

The PheWAS-View software was used to visually integrate study results, to discover novel relationships between SNPs and phenotypes, and to produce forest plots (Pendergrass et al., 2012). All statistical analyses were performed with SAS version 9.2/9.3 (SAS Institute, Inc., Cary, NC) and JMP Pro 10 (SAS Institute, Inc., Cary, NC).

RESULTS

Study Population

The basic demographic characteristics of study participants analyzed for each assay are summarized in Table 1. Due to limited quantities of DNA there are differences in the sample number for the *UGT2B* and *UGT1A* datasets (Table 1). Overall, a significant difference ($P<0.05$) was found between cases and controls in the UGT1A and UGT2B sample sets for age, BMI, history of CRC in 1st degree family, NSAID use, physical activity, education, and BaP intake (Table 1). Although cases and controls were age-matched during recruitment, a significant difference in age was observed in the study since some subjects were excluded from the genotyping analysis due to low quantities of DNA. While smoking status was significantly ($P = 0.022$) different for cases versus controls for both sample sets, pack-years was only significant ($P = 0.014$) for the UGT1A sample set. While a significant difference was found for MeIQx and DiMeIQx intakes only in the UGT1A sample set, this appeared to be driven by a few outliers (n=5) not observed for the UGT2B sample set (results not shown).

UGT1A SNP Associations with CRC Risk

Unconditional logistic regression analysis using an additive, dominant, and recessive statistical model was conducted for individual SNPs and haplotypes within the *UGT1A* gene family for the effect on CRC risk controlling for age, education, sex, BMI, family history, NSAID use, and physical activity. Three of the 96 SNPs did not amplify and 8 of the SNPs (rs2741028, rs11893247, rs6706988, rs17863773, rs10176426, rs12474980, rs12463641, and rs17862878) failed HWE in the controls, leaving a total of 85 SNPs for analysis. No significant associations were found between individual *UGT1A* SNPs and CRC risk overall or after stratification analysis by sex, cancer sub-site (colon versus rectum), or levels of carcinogen intake when applying the FDR multiple testing correction for all SNPs genotyped.

UGT1A Haplotype Associations with CRC Risk

Haploview software was used on the controls to divide the *UGT1A* gene region into eleven haplotype blocks, and each haplotype (Fig. 1, Supplementary Table 1) was then analyzed for impact on CRC risk. Haplotype block 7 was divided into four blocks (7.1, 7.2, 7.3, and 7.4) due to computational power for the analysis (Fig. 1, Supplementary Table 1). SAS PROC HAPLOTYPE was then used to assign the probability that each individual possesses a particular haplotype compared with all other haplotype possibilities, which was then analyzed in unconditional logistic regression analysis controlling for the same covariates in the individual SNP analysis. Therefore, the analysis, assuming an additive statistical model, reported the risk associated with a specific haplotype when compared with all other haplotypes in the population at a frequency greater than 5%. Several *UGT1A* haplotype blocks were associated with CRC risk in this analysis. In the overall analysis, the T-T-A-G-A haplotype in block 4 (Fig. 1) was found to increase significantly CRC risk (OR= 2.44, 95% CI=1.29–4.6; Table 2). In addition, stratifications by sex and cancer sub-site also yielded FDR-adjusted significance for haplotypes in blocks 2, 5, 6, 7.1, 7.4, 9, and 11 (Table 2). The significant decrease in cancer risk for the T-G haplotype (rs17864678, rs10929251) in block 2 was found in both proximal (OR = 0.29, 95% CI =0.11–0.69) and distal (OR = 0.32, 95% CI =0.12–0.95) colon cancer patients, making it associated with colon cancer risk and not rectal cancer risk. The significance found between decreased proximal colon cancer risk and haplotypes in blocks 7.1 (OR = 0.24, 95% CI = 0.085–0.69) and 7.4 (OR=0.26, 95% CI= 0.091–0.71) is similar, as the two blocks were divided within one block created by Haploview and they contain three overlapping SNPs (rs1604144, rs12988520, and rs7240193). The C-T-G haplotype in block 11 (rs7578153, rs10203853, and rs6728940) was found in males to increase CRC risk overall (OR = 2.56, 95% CI =1.10–5.95) and the risk of proximal colon cancer (OR = 4.06, 95% CI = 1.30–12.6). No association with cancer risk was observed for any *UGT1A* blocks specifically in females even after stratification by sub-site (results not shown).

UGT2B SNP Associations with CRC Risk

As conducted for SNPs in the *UGT1A* gene family, unconditional logistic regression analysis using an additive, dominant, and recessive statistical model was analyzed on individual SNPs and haplotypes within the *UGT2B* gene loci for the effect on CRC risk

controlling for age, education, sex, BMI, family history, NSAID use, and physical activity. Two of the 16 SNPs did not amplify, 1 SNP was not a polymorphism in our population (rs7439366), and 1 SNP was genotyped in only 84% of our patients (rs7668258), leaving a total of 12 SNPs and one deletion/insertion polymorphism (rs35922514) for analysis. All SNPs were consistent with HWE. Overall, no significant associations were found between individual *UGT2B* SNPs and CRC risk even after stratification analysis by sex and high carcinogen intake, when applying the FDR multiple testing correction for all SNPs genotyped. In rectal cancer patients, a few SNPs yielded borderline associations (rs4148269, rs61750900, rs835317, and rs11737566) but only one SNP in *UGT2B15*, rs6837575 (minor allele frequency of 0.386 in controls), was found to decrease risk significantly after multiple testing correction using a dominant statistical model (OR= 0.47, 95% CI = 0.29–0.74, FDR $P = 0.020$; Fig. 2, panel A).

UGT2B Haplotype Associations with CRC Risk

SNPs within the *UGT2B* region were divided into two haplotype blocks using Haploview and SAS PROC Haplotype to calculate the probability that a particular individual possesses a certain haplotype compared with all other haplotype possibilities (Fig. 2, Supplementary Table 2). The rs35922514 polymorphism was excluded from the haplotype analysis because it was a deletion/insertion polymorphism and the insertion was only present in 0.6% (11/1761) of the population with no individuals exhibiting the homozygous rare genotype. While no *UGT2B* haplotypes were significantly associated with overall CRC risk, a significant decreased risk was found for the A-G haplotype (rs4148269; MA = A; rs6837575, MA = A) in block 1 (OR=0.39, 95% CI=0.19–0.77, FDR $P = 0.01$) in patients with rectal cancer in an additive statistical model (Figure 2, panel B). An increased risk for rectal cancer was found for the same haplotype block with a C-A haplotype (OR=2.57, 95% CI=1.21–5.04, FDR $P = 0.01$). After further stratification by gender, the A-G (OR=0.26, 95% CI=0.085–0.77, FDR $P = 0.03$) and C-A (OR=3.08, 95% CI=1.08–8.74, FDR $P = 0.035$) haplotypes in block 1 were found to alter significantly rectal cancer risk in females in the same direction; no significant associations were observed specifically in males (results not shown). The two SNPs that make up block 1 are both located in the *UGT2B15* gene; rs4148269 is a missense polymorphism (c.C1568A, K523T) in exon 1 and rs6837575 is in intron 1.

NSAID Use and *UGT1A* Polymorphisms

Gene x environment (GxE) interactions was tested for all SNPs and haplotypes in both the *UGT1A* and *UGT2B* loci with high carcinogen intake (PhIP, MeIQx, DiMeIQx, and BaP) as well as high NSAID use. In addition, the same GxE interactions were tested for all SNPs and haplotypes in both the *UGT1A* and *UGT2B* loci in all stratification analyses. No significant GxE interactions were found with carcinogen intake after conducting a multiple testing correction. In the *UGT1A* gene cluster, the interaction between high NSAID use and the A-G-T haplotype (rs6717546, rs1500482, rs7586006) combined had a significant ($P = 0.027$) interaction after multiple testing correction, leading to decreased CRC risk. The homozygous recessive allele of rs1500482 was significant prior to multiple testing correction ($P = 0.0007$) but did not remain significant after the FDR correction ($P = 0.051$)

for the interaction between high NSAID use and decreased CRC risk (Table 3). No other gene x NSAID interaction was found between individual SNPs and haplotypes.

DISCUSSION

Select polymorphisms within the *UGT* gene family have been found to play an important role in orolaryngeal, gastrointestinal, colorectal, lung, breast, pancreatic, and prostate cancer risk and pathology (Mackenzie et al., 2000; Strassburg et al., 2002; Butler et al., 2005; Tseng et al., 2005; Gallagher et al., 2007a; Gallagher et al., 2009; Chen et al., 2010; Angstadt et al., 2013; Wang et al., 2013). For CRC, *UGT* polymorphisms have been associated with treatment outcome and risk, but these have all been identified at the single gene and polymorphism level. These include recent studies of whole-gene deletion polymorphisms in *UGTs 2B17* and *2B28*, which demonstrated that the *UGT2B17* deletion genotype was associated with a decrease in CRC risk (Angstadt et al., 2013).

The present study is the first comprehensive study to examine polymorphisms and haplotypes within the genomic regions of the *UGT1A* and *UGT2B* gene families on CRC risk. This study also analyzed the impact of these polymorphisms in combination with environmental cofactors on CRC risk, but although the dietary and demographic questionnaires were extensive in nature, the inability to acquire comprehensively clinical information, such as anatomical sub-site, is a limitation. This study did perform a multiple testing correction on each analysis to account for potential false positives but the stratification analysis is limited in power and will need further validation. Although this study did not find significance when applying a multiple testing correction for individual SNPs in the *UGT1A* gene region with CRC risk, haplotypes within the *UGT1A* loci were found to affect CRC risk overall and after stratification by sex and anatomic sub-site. Previous studies of specific *UGT1A* polymorphisms in CRC risk have found that the *UGT1A1*6* allele (G71R in exon 1) is associated with an increased risk for CRC in a Chinese population (Tang et al., 2005). In addition, the *UGT1A7*3* allele (N129/R131/W208R) increases CRC risk in Caucasians (Strassburg et al., 2002; van der Logt et al., 2004), predicted low UGT1A7 activity and higher intake of DiMeIQx was positively associated with CRC in African Americans (Butler et al., 2005), and UGT1A7 polymorphisms increased CRC risk in a Chinese population (Chen et al., 2006; Lu et al., 2011). The results from the present study are consistent with these previous findings since the increased risk for CRC in Caucasian males was found in a haplotype (block 11) in the 3′ region flanking the *UGT1A* shared exons, which would represent all *UGT1A* genes including *UGTs 1A1* and *1A7*. Interestingly, additional associations within haplotype blocks 2 and 5 yielded two different haplotypes affecting CRC risk in opposite directions, with one haplotype increasing risk and the other decreasing risk. These data are consistent with the fact that UGTs metabolize both potentially harmful as well as beneficial dietary agents, which could either increase or protect against CRC depending on the dietary exposure (Turgeon et al., 2003; Kuehl et al., 2005; Kuehl et al., 2006). The associated haplotype block 2 covers all of *UGT1A10* exon 1, which includes the substrate binding domain, and approximately 2 kb of the *UGT1A10* promoter. The associated haplotype block 5 covers a section of intron 1 of *UGT1A8* and intron 1 of *UGT1A10*. To date, no known associations have been identified between polymorphisms in the *UGT1A8* and *UGT1A10* genes and CRC

risk. Although the SNPs genotyped in this study may not have any known functional roles, the significant association indicates that the haplotypes are tagging a variant that does have functional relevance. Polymorphic expression analysis of *UGT1A* genes in colon cancer did find that *UGT1A8* and *UGT1A10* were up-regulated in colon tumor tissue compared with healthy tissue from the same patients (Wang et al., 2013). Therefore, the specific haplotypes associated with increasing colon cancer risk may be leading to coordinated increased *UGT1A8* and *UGT1A10* expression and activity during the colon tumorigenic process. Further exploration into the in vitro and in vivo effects of these haplotypes is warranted to assess better their effects on UGT1A expression.

While there has been other previous data to suggest that an interaction exists between DiMeIQx intake and *UGT1A7* variants (Butler et al., 2005) as well as between BaP intake and variants in the *UGT1A1* promoter at positions −53 and −3156 (Girard et al., 2008), the findings in the present study are in agreement with a recent publication that did not find any significant interactions between *UGT1A* SNPs and HCA and PAH intake after adjustment for multiple testing (Gilsing et al., 2012).

In addition to examining potential interactions between HCA and PAH exposure and UGT variants, the present study is the first to analyze fully interactions between *UGT* genes and NSAID use. Previous studies have shown that high NSAID use is associated with lower CRC risk (Muscat et al., 1994; Flossmann and Rothwell 2007; Rostom et al., 2007; Cinar et al., 2010; Din et al., 2010). In addition, multiple UGT enzymes, most notably UGTs 1A1, 1A9, 2B4, 2B7, 2B15, and 2B17 have been found to be involved in the hepatic catalysis of NSAID glucuronidation (Kuehl et al., 2005). Unlike the results described above for meat mutagens, an interaction was found between the A-G-T block 10 (rs6717546, rs1500482, and rs7586006) haplotype near the 3′ end of the *UGT1A* gene region and high NSAID use that decreased CRC risk. The homozygous recessive SNP rs1500482, located within that haplotype, yielded a borderline association in single SNP analysis, but it did not pass the FDR correction. This is consistent with the inverse association observed previously between aspirin use and adenoma risk that was restricted to individuals with variant *UGT1A6* genotypes (≥ one T181A+R184S or R184S allele) who have a delayed aspirin metabolism (Chan et al., 2005). In addition, it has recently been shown that, in addition to the alternate splicing of individual *UGT1A* exon 1′s onto the *UGT1A* common region encoded by exon's 2–5, there is an additional splicing event that occurs within exon 5 that is common for all *UGT1A* genes (Bellemare et al., 2010b). The *UGT1A* terminal exon 5 exhibits two splice variants, isoform 1 being enzymatically active and isoform 2 being enzymatically inactive, for each of the UGT1A enzymes. While themselves enzymatically inactive, isoform 2 UGT1A's inhibit the glucuronidation activity of their isoform 1 counterparts (Girard et al., 2007; Levesque et al., 2007; Bellemare et al., 2010a,b). As the A-G-T block 10 haplotype is located within the 3′ *UGT1A* region of both isoforms 1 and 2, this haplotype may have a regulatory effect on production of the inactive UGT1A isoform 2, allowing for a delayed NSAID metabolism, which in turn could provide a protective effect in vivo against the development of CRC.

Results from this study indicate that a single SNP (rs6837575) in the *UGT2B* gene region was significantly associated with rectal cancer risk after correction for multiple testing.

While future studies will be required to test whether this and other SNPs examined in this study are themselves functional or are tagging other functional SNPs, the rs6837575 SNP was also associated with overall and female rectal cancer risk within a haplotype (rs6837575 and rs4148269) that tags the *UGT2B15* exon 1 and intron 1 region, emphasizing its impact on the disease. Recent analysis of the functional coding *UGT2B15* rs4148269, c.1568C>A, polymorphism suggested that this SNP (or SNPs in LD with it) can alter *UGT2B15* expression in the liver and breast (Sun et al., 2011). Although the c1568C>A SNP was not associated with altered enzyme activity against oxazepam (Court et al., 2004), it was found to be in LD with c.1761T>C (rs3100) in the *UGT2B15* 3′ untranslated region. Reporter gene assays showed that the 1761T allele resulted in significantly higher *UGT2B15* expression levels than the 1761C allele in HepG2, MCF-7, LNCaP, and Caco-2 cell lines (Sun et al., 2011). The lower expression of *UGT2B15* associated with the 1761C allele is consistent with the protective effect on rectal cancer risk observed in the present study since UGT2B15 exhibits high glucuronidation activity against ibuprofen (Kuehl et al., 2006), the most commonly taken NSAID in the study population. Lower UGT2B15 expression, resulting in lower ibuprofen metabolism and excretion (Din et al., 2010), could therefore potentially accentuate ibuprofen's protective effect against rectal cancer. The differential effect of the *UGT2B15* c.1568C>A polymorphism in rectal vs. colon cancer may be due to differences in expression of the *UGT2B15* gene in different parts of the colorectum, which needs to be established in future studies. In addition, recent analysis using the same population as the present study concluded that different meat-related compounds were associated with proximal colon, distal colon, and rectal cancer, suggesting that the etiology of each cancer may be different (Miller et al., 2013) and may be the reason why polymorphic associations were observed for specific anatomical sub-sites. It has also been shown that *UGT2B15* is expressed in the liver at a higher level in men than women (Sun et al., 2012). While similar gender-specific analysis has not as yet been performed for colon or rectum, this may help explain the fact that the haplotypes only affected risk for female rectal cancer patients.

In conclusion, the results from the present study indicate that genetic variation in the *UGT1A* and *UGT2B* loci appears to play an important role for CRC risk. Haplotype analysis indicated that *UGT1A* variants located in several genes are responsible for altering colon and rectal cancer risk, suggesting that abnormal expression or activity of more than one *UGT1A* gene may be responsible for different anatomical sub-sites of CRC. In addition, some UGT variants affect CRC risk differently in men and women. These findings agree with the fact that UGT enzymes are known to exhibit different expression and activity patterns in men and women (Gallagher et al., 2007b; Sun et al., 2012). Although no significant interactions between polymorphisms and meat mutagens were found in the present study, the interaction observed between the *UGT1A* A-G-T haplotype and high NSAID use that decreased CRC risk highlights the role of UGT polymorphisms on xenobiotic levels (such as NSAIDs), which can then effect disease development or prevention. Also of interest in this study is the finding that a functional variant in *UGT2B15* affects rectal cancer risk. A more comprehensive study of the functional impact of the associated polymorphisms in this study determining how these SNPs affect UGT expression and activity would aid in a better understanding of the role they play in the pathogenesis of each anatomical sub-site and in men versus women.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Functional Genomics Core Facility at the Penn State University College of Medicine for genotyping services.

Supported by: National Institute of Health grant R00-CA131477 (to Carla Gallagher) and Pennsylvania Department of Health CURE funds #4100038714 (to Philip Lazarus) and #4100050904 (to Carla Gallagher).

References

Alpert HA, Prokup I, Lesko SM. Colorectal cancer incidence and mortality in northeastern pennsylvania. Registry Management. 2007; 34:99.

Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. Nature. 2005; 437:1299–1320. [PubMed: 16255080]

Angstadt AY, Berg A, Zhu J, Miller P, Hartman TJ, Lesko SM, Muscat JE, Lazarus P, Gallagher CJ. The effect of copy number variation in the phase II detoxification genes ugt2b17 and ugt2b28 on colorectal cancer risk. Cancer. 2013; 119:2477–2485. [PubMed: 23575887]

Bailey RL, Miller PE, Mitchell DC, Hartman TJ, Lawrence FR, Sempos CT, Smiciklas-Wright H. Dietary screening tool identifies nutritional risk in older adults. Am J Clin Nutr. 2009; 90:177–183. [PubMed: 19458013]

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21:263–265. [PubMed: 15297300]

Bellemare J, Rouleau M, Harvey M, Guillemette C. Modulation of the human glucuronosyltransferase ugt1a pathway by splice isoform polypeptides is mediated through protein-protein interactions. J Biol Chem. 2010a; 285:3600–3607. [PubMed: 19996319]

Bellemare J, Rouleau M, Harvey M, Tetu B, Guillemette C. Alternative-splicing forms of the major phase ii conjugating ugt1a gene negatively regulate glucuronidation in human carcinoma cell lines. Pharmacogenomics J. 2010b; 10:431–441. [PubMed: 19997083]

Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Royal Statistical Society. 1995; B:289–300.

Bernard O, Guillemette C. The main role of ugt1a9 in the hepatic metabolism of mycophenolic acid and the effects of naturally occurring variants. Drug Metab Dispos. 2004; 32:775–778. [PubMed: 15258099]

Butler LM, Duguay Y, Millikan RC, Sinha R, Gagne JF, Sandler RS, Guillemette C. Joint effects between udp-glucuronosyltransferase 1a7 genotype and dietary carcinogen exposure on risk of colon cancer. Cancer Epidemiol Biomarkers Prev. 2005; 14:1626–1632. [PubMed: 16030093]

Chan AT, Tranah GJ, Giovannucci EL, Hunter DJ, Fuchs CS. Genetic variants in the ugt1a6 enzyme, aspirin use, and the risk of colorectal adenoma. J Natl Cancer Inst. 2005; 97:457–460. [PubMed: 15770010]

Chen G, Giambrone NE Jr, Dluzen DF, Muscat JE, Berg A, Gallagher CJ, Lazarus P. Glucuronidation genotypes and nicotine metabolic phenotypes: Importance of functional ugt2b10 and ugt2b17 polymorphisms. Cancer Res. 2010; 70:7543–7552. [PubMed: 20876810]

Chen K, Jin M, Zhu Y, Jiang Q, Yu W, Ma X, Yao K. Genetic polymorphisms of the uridine diphosphate glucuronosyltransferase 1a7 and colorectal cancer risk in relation to cigarette smoking and alcohol drinking in a chinese population. J Gastroenterol Hepatol. 2006; 21:1036–1041. [PubMed: 16724991]

Cinar M, Dinc A, Simsek I, Erdem H, Koc B, Pay S, Kilic S. Evaluation of the short-term efficacy of nsaids on patients with active ankylosing spondylitis in daily practice: A 3-month, longitudinal, observational study. Rheumatol Int. 2010; 30:331–340. [PubMed: 19466421]

Court MH, Hao Q, Krishnaswamy S, Bekaii-Saab T, Al-Rohaimi A, von Moltke LL, Greenblatt DJ. Udp-glucuronosyltransferase (ugt) 2b15 pharmacogenetics: Ugt2b15 d85y genotype and gender are major determinants of oxazepam glucuronidation by human liver. J Pharmacol Exp Ther. 2004; 310:656–665. [PubMed: 15044558]

Czika, W.; Yu, X. Gene frequencies and linkage disequilibrium. Genetic analysis of complex traits using sas. NC: SAS Publishing Arnold Myron Saxton; 2004.

Din FV, Theodoratou E, Farrington SM, Tenesa A, Barnetson RA, Cetnarskyj R, Stark L, Porteous ME, Campbell H, Dunlop MG. Effect of aspirin and nsaids on risk and survival from colorectal cancer. Gut. 2010; 59:1670–1679. [PubMed: 20844293]

Duguay Y, Baar C, Skorpen F, Guillemette C. A novel functional polymorphism in the uridine diphosphate-glucuronosyltransferase 2b7 promoter with significant impact on promoter activity. Clin Pharmacol Ther. 2004; 75:223–233. [PubMed: 15001974]

Ehmer U, Vogel A, Schutte JK, Krone B, Manns MP, Strassburg CP. Variation of hepatic glucuronidation: Novel functional polymorphisms of the udp-glucuronosyltransferase ugt1a4. Hepatology. 2004; 39:970–977. [PubMed: 15057901]

Flossmann E, Rothwell PM. Effect of aspirin on long-term risk of colorectal cancer: Consistent evidence from randomised and observational studies. Lancet. 2007; 369:1603–1613. [PubMed: 17499602]

Friday, J.; Bowman, S. Mypyramid equivalents database for usda survey food codes, 1994–2002 version 1.0. U.S. Department of Agriculture, Agriculture Research Service; 2006.

Gallagher CJ, Ahn K, Knipe AL, Dyer AM, Richie JP Jr, Lazarus P, Muscat JE. Association between haplotypes of manganese superoxide dismutase (sod2), smoking, and lung cancer risk. Free Radic Biol Med. 2009; 46:20–24. [PubMed: 18930810]

Gallagher CJ, Kadlubar FF, Muscat JE, Ambrosone CB, Lang NP, Lazarus P. The ugt2b17 gene deletion polymorphism and risk of prostate cancer. A case-control study in caucasians. Cancer Detect Prev. 2007a; 31:310–315. [PubMed: 17935910]

Gallagher CJ, Muscat JE, Hicks AN, Zheng Y, Dyer AM, Chase GA, Richie J, Lazarus P. The udp-glucuronosyltransferase 2b17 gene deletion polymorphism: Sex-specific association with urinary 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol glucuronidation phenotype and risk for lung cancer. Cancer Epidemiol Biomarkers Prev. 2007b; 16:823–828. [PubMed: 17416778]

Gestl SA, Green MD, Shearer DA, Frauenhoffer E, Tephly TR, Weisz J. Expression of ugt2b7, a udp-glucuronosyltransferase implicated in the metabolism of 4-hydroxyestrone and all-trans retinoic acid, in normal human breast parenchyma and in invasive and in situ breast cancers. Am J Pathol. 2002; 160:1467–1479. [PubMed: 11943730]

Gilsing AM, Berndt SI, Ruder EH, Graubard BI, Ferrucci LM, Burdett L, Weissfeld JL, Cross AJ, Sinha R. Meat-related mutagen exposure, xenobiotic metabolizing gene polymorphisms and the risk of advanced colorectal adenoma and cancer. Carcinogenesis. 2012; 33:1332–1339. [PubMed: 22552404]

Girard H, Levesque E, Bellemare J, Journault K, Caillier B, Guillemette C. Genetic diversity at the ugt1 locus is amplified by a novel 3′ alternative splicing mechanism leading to nine additional ugt1a proteins that act as regulators of glucuronidation activity. Pharmacogenet Genomics. 2007; 17:1077–1089. [PubMed: 18004212]

Girard H, Butler LM, Villeneuve L, Millikan RC, Sinha R, Sandler RS, Guillemette C. Ugt1a1 and ugt1a9 functional variants, meat intake, and colon cancer, among caucasians and african-americans. Mutat Res. 2008; 644:56–63. [PubMed: 18675828]

Giuliani L, Gazzaniga P, Caporuscio F, Ciotti M, Frati L, Agliano AM. Can down-regulation of udp-glucuronosyltransferases in the urinary bladder tissue impact the risk of chemical carcinogenesis? Int J Cancer. 2001; 91:141–143. [PubMed: 11149414]

Gong QH, Cho JW, Huang T, Potter C, Gholami N, Basu NK, Kubota S, Carvalho S, Pennington MW, Owens IS, Popescu NC. Thirteen udpglucuronosyltransferase genes are encoded at the human ugt1 gene complex locus. Pharmacogenetics. 2001; 11:357–368. [PubMed: 11434514]

Greenland, S.; Rothman, K. Introduction to stratified analysis. In: Rothman, K.; Greenland, S.; Lash, T., editors. Modern epidemiology. 3. Philadelphia, PA: Lippincott Williams and Wilkins; 2008.

Guillemette C. Pharmacogenomics of human udp-glucuronosyltransferase enzymes. Pharmacogenomics J. 2003; 3:136–158. [PubMed: 12815363]

Iwai M, Maruo Y, Ito M, Yamamoto K, Sato H, Takeuchi Y. Six novel udp-glucuronosyltransferase (ugt1a3) polymorphisms with varying activity. J Hum Genet. 2004; 49:123–128. [PubMed: 14986168]

Jinno H, Saeki M, Saito Y, Tanaka-Kagawa T, Hanioka N, Sai K, Kaniwa N, Ando M, Shirao K, Minami H, Ohtsu A, Yoshida T, Saijo N, Ozawa S, Sawada J. Functional characterization of human udp-glucuronosyltransferase 1a9 variant, d256n, found in japanese cancer patients. J Pharmacol Exp Ther. 2003; 306:688–693. [PubMed: 12730278]

Korprasertthaworn P, Rowland A, Lewis BC, Mackenzie PI, Yoovathaworn K, Miners JO. Effects of amino acid substitutions at positions 33 and 37 on udp-glucuronosyltransferase 1a9 (ugt1a9) activity and substrate selectivity. Biochem Pharmacol. 2012; 84:1511–1521. [PubMed: 22981363]

Krishnaswamy S, Hao Q, Al-Rohaimi A, Hesse LM, von Moltke LL, Greenblatt DJ, Court MH. Udp glucuronosyltransferase (ugt) 1a6 pharmacogenetics: I. Identification of polymorphisms in the 5′-regulatory and exon 1 regions, and association with human liver ugt1a6 gene expression and glucuronidation. J Pharmacol Exp Ther. 2005; 313:1331–1339. [PubMed: 15761114]

Kuehl GE, Lampe JW, Potter JD, Bigler J. Glucuronidation of nonsteroidal anti-inflammatory drugs: Identifying the enzymes responsible in human liver microsomes. Drug Metab Dispos. 2005; 33:1027–1035. [PubMed: 15843492]

Kuehl GE, Bigler J, Potter JD, Lampe JW. Glucuronidation of the aspirin metabolite salicylic acid by expressed udp-glucuronosyltransferases and human liver microsomes. Drug Metab Dispos. 2006; 34:199–202. [PubMed: 16258079]

Larsson SC, Wolk A. Meat consumption and risk of colorectal cancer: A meta-analysis of prospective studies. Int J Cancer. 2006; 119:2657–2664. [PubMed: 16991129]

Levesque E, Beaulieu M, Green MD, Tephly TR, Belanger A, Hum DW. Isolation and characterization of ugt2b15(y85): A udp-glucuronosyltransferase encoded by a polymorphic gene. Pharmacogenetics. 1997; 7:317–325. [PubMed: 9295060]

Levesque E, Beaulieu M, Hum DW, Belanger A. Characterization and substrate specificity of ugt2b4 (e458): A udp-glucuronosyltransferase encoded by a polymorphic gene. Pharmacogenetics. 1999; 9:207–216. [PubMed: 10376768]

Levesque E, Girard H, Journault K, Lepine J, Guillemette C. Regulation of the ugt1a1 bilirubin-conjugating pathway: Role of a new splicing event at the ugt1a locus. Hepatology. 2007; 45:128–138. [PubMed: 17187418]

Lu PH, Chen MB, Wu XY, Gu JH, Liu Y, Gu RM. Genetic polymorphisms of ugt1a7 and cancer risk: Evidence from 21 case-control studies. Cancer Invest. 2011; 29:645–654. [PubMed: 22085268]

Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Belanger A, Fournel-Gigleux S, Green M, Hum DW, Iyanagi T, Lancet D, Louisot P, Magdalou J, Chowdhury JR, Ritter JK, Schachter H, Tephly TR, Tipton KF, Nebert DW. The udp glycosyltransferase gene superfamily: Recommended nomenclature update based on evolutionary divergence. Pharmacogenetics. 1997; 7:255–269. [PubMed: 9295054]

Mackenzie PI, Miners JO, McKinnon RA. Polymorphisms in udp glucuronosyltransferase genes: Functional consequences and clinical relevance. Clin Chem Lab Med. 2000; 38:889–892. [PubMed: 11097345]

Mackenzie PI, Gregory PA, Gardner-Stephen DA, Lewinsky RH, Jorgensen BR, Nishiyama T, Xie W, Radominska-Pandya A. Regulation of udp glucuronosyltransferase genes. Curr Drug Metab. 2003; 4:249–257. [PubMed: 12769669]

Maitland ML, Grimsley C, Kuttab-Boulos H, Witonsky D, Kasza KE, Yang L, Roe BA, Di Rienzo A. Comparative genomics analysis of human sequence variation in the ugt1a gene cluster. Pharmacogenomics J. 2006; 6:52–62. [PubMed: 16314881]

McDonnell WM, Hitomi E, Askari FK. Identification of bilirubin udp-gts in the human alimentary tract in accordance with the gut as a putative metabolic organ. Biochem Pharmacol. 1996; 51:483–488. [PubMed: 8619894]

Miller PE, Lazarus P, Lesko SM, Cross AJ, Sinha R, Laio J, Zhu J, Harper G, Muscat JE, Hartman TJ. Meat-related compounds and colorectal cancer risk by anatomical subsite. Nutr Cancer. 2013; 65:202–226. [PubMed: 23441608]

Miners JO, McKinnon RA, Mackenzie PI. Genetic polymorphisms of udp-glucuronosyltransferases and their functional significance. Toxicology. 2002; 181–182:453–456.

Muscat JE, Stellman SD, Wynder EL. Nonsteroidal antiinflammatory drugs and colorectal cancer. Cancer. 1994; 74:1847–1854. [PubMed: 8082089]

Nakamura A, Nakajima M, Yamanaka H, Fujiwara R, Yokoi T. Expression of ugt1a and ugt2b mrna in human normal tissues and various cell lines. Drug Metab Dispos. 2008; 36:1461–1464. [PubMed: 18480185]

Organization WH. World health organization: International classification of diseases for oncology. Geneva, Switzerland: 2000.

Pendergrass SA, Dudek SM, Crawford DC, Ritchie MD. Visually integrating and exploring high throughput phenome-wide association study (phewas) results using phewas-view. BioData Min. 2012; 5:5. [PubMed: 22682510]

Rostom A, Dube C, Lewin G, Tsertsvadze A, Barrowman N, Code C, Sampson M, Moher D. Nonsteroidal anti-inflammatory drugs and cyclooxygenase-2 inhibitors for primary prevention of colorectal cancer: A systematic review prepared for the u.S. Preventive services task force. Ann Intern Med. 2007; 146:376–389. [PubMed: 17339623]

Sinha R, Rothman N. Exposure assessment of heterocyclic amines (hcas) in epidemiologic studies. Mutat Res. 1997; 376:195–202. [PubMed: 9202756]

Sinha R, Peters U, Cross AJ, Kulldorff M, Weissfeld JL, Pinsky PF, Rothman N, Hayes RB. Meat, meat cooking methods and preservation, and risk for colorectal adenoma. Cancer Res. 2005; 65:8034–8041. [PubMed: 16140978]

Slattery ML, Fitzpatrick FA. Convergence of hormones, inflammation, and energy-related factors: A novel pathway of cancer etiology. Cancer Prev Res (Phila). 2009; 2:922–930. [PubMed: 19892662]

Strassburg CP, Oldhafer K, Manns MP, Tukey RH. Differential expression of the ugt1a locus in human liver, biliary, and gastric tissue: Identification of ugt1a7 and ugt1a10 transcripts in extrahepatic tissue. Mol Pharmacol. 1997; 52:212–220. [PubMed: 9271343]

Strassburg CP, Nguyen N, Manns MP, Tukey RH. Polymorphic expression of the udp-glucuronosyltransferase ugt1a gene locus in human gastric epithelium. Mol Pharmacol. 1998; 54:647–654. [PubMed: 9765507]

Strassburg CP, Vogel A, Kneip S, Tukey RH, Manns MP. Polymorphisms of the human udp-glucuronosyltransferase (ugt) 1a7 gene in colorectal cancer. Gut. 2002; 50:851–856. [PubMed: 12010889]

Sun C, Southard C, Olopade OI, Di Rienzo A. Differential allelic expression of c.1568c > a at ugt2b15 is due to variation in a novel cis-regulatory element in the 3′utr. Gene. 2011; 481:24–28. [PubMed: 21513781]

Sun C, Southard C, Huo D, Hernandez RD, Witonsky DB, Olopade OI, Di Rienzo A. Snp discovery, expression and cis-regulatory variation in the ugt2b genes. Pharmacogenomics J. 2012; 12:287–296. [PubMed: 21358749]

Tang KS, Chiu HF, Chen HH, Eng HL, Tsai CJ, Teng HC, Huang CS. Link between colorectal cancer and polymorphisms in the uridine-diphosphoglucuronosyltransferase 1a7 and 1a1 genes. World J Gastroenterol. 2005; 11:3250–3254. [PubMed: 15929176]

Thompson FE, Subar AF, Brown CC, Smith AF, Sharbaugh CO, Jobe JB, Mittl B, Gibson JT, Ziegler RG. Cognitive research enhances accuracy of food frequency questionnaire reports: Results of an experimental validation study. J Am Diet Assoc. 2002; 102:212–225. [PubMed: 11846115]

Tseng CS, Tang KS, Lo HW, Ker CG, Teng HC, Huang CS. Udp-glucuronosyltransferase 1a7 genetic polymorphisms are associated with hepatocellular carcinoma risk and onset age. Am J Gastroenterol. 2005; 100:1758–1763. [PubMed: 16086712]

Turgeon D, Carrier JS, Chouinard S, Belanger A. Glucuronidation activity of the ugt2b17 enzyme toward xenobiotics. Drug Metab Dispos. 2003; 31:670–676. [PubMed: 12695357]

van der Logt EM, Bergevoet SM, Roelofs HM, van Hooijdonk Z, te Morsche RH, Wobbes T, de Kok JB, Nagengast FM, Peters WH. Genetic polymorphisms in udp-glucuronosyltransferases and glutathione s-transferases and colorectal cancer risk. Carcinogenesis. 2004; 25:2407–2415. [PubMed: 15319294]

Villeneuve L, Girard H, Fortier LC, Gagne JF, Guillemette C. Novel functional polymorphisms in the ugt1a7 and ugt1a9 glucuronidating enzymes in caucasian and african-american subjects and their impact on the metabolism of 7-ethyl-10-hydroxycamptothecin and flavopiridol anticancer drugs. J Pharmacol Exp Ther. 2003; 307:117–128. [PubMed: 12944498]

Waksberg J. Sampling methods for random digit dialing. Journal of the American Statistical Association. 1978:40–46.

Wang M, Sun DF, Wang S, Qing Y, Chen S, Wu D, Lin YM, Luo JZ, Li YQ. Polymorphic expression of udp-glucuronosyltransferase ugtla gene in human colorectal cancer. Plos One. 2013; 8:e57045. [PubMed: 23468910]

Wang Y, Beydoun MA. Meat consumption is associated with obesity and central obesity among us adults. Int J Obes (Lond). 2009; 33:621–628. [PubMed: 19308071]

Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997; 65:1220S–1228S. [PubMed: 9094926]

Wu K, Giovannucci E, Byrne C, Platz EA, Fuchs C, Willett WC, Sinha R. Meat mutagens and risk of distal colon adenoma in a cohort of u.S. Men Cancer Epidemiol Biomarkers Prev. 2006; 15:1120–1125.

Zheng Z, Fang JL, Lazarus P. Glucuronidation: An important mechanism for detoxification of benzo[a]pyrene metabolites in aerodigestive tract tissues. Drug Metab Dispos. 2002; 30:397–403. [PubMed: 11901093]

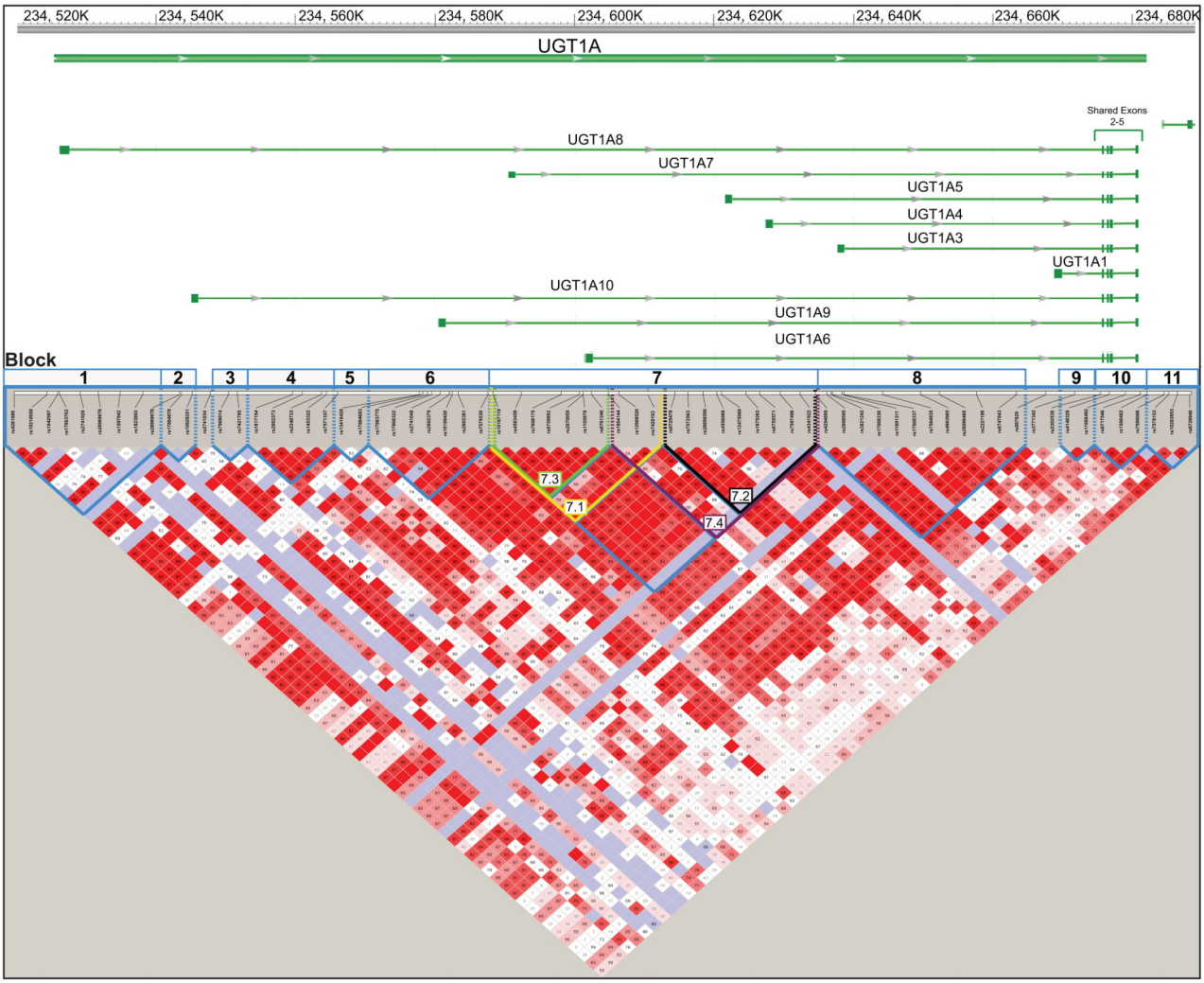

Figure 1.

Linkage disequilibrium (LD) plot of SNPs in the *UGT1A* gene loci on 2q37. Their respective haplotypes were identified by Haploview using controls from the study population. All SNPs shown passed Hardy-Weinberg Equilibrium analysis and were genotyped at a rate >90% in the present study. D′ values are displayed in the squares (empty squares have a pairwise D′=1.00). Red squares show high pairwise LD, gradually coloring down to white squares of low pairwise LD. Blue squares indicate high LD, but low significance. Blue triangles indicate the SNPs in high LD that were grouped into individual haplotype blocks, including the division of haplotype block 7 shown by different colored triangles into 4 blocks because of limited computational power. Lines that extend from the LD plot are dashed as they do not always match the exact chromosomal beginning and end position of each haplotype block.

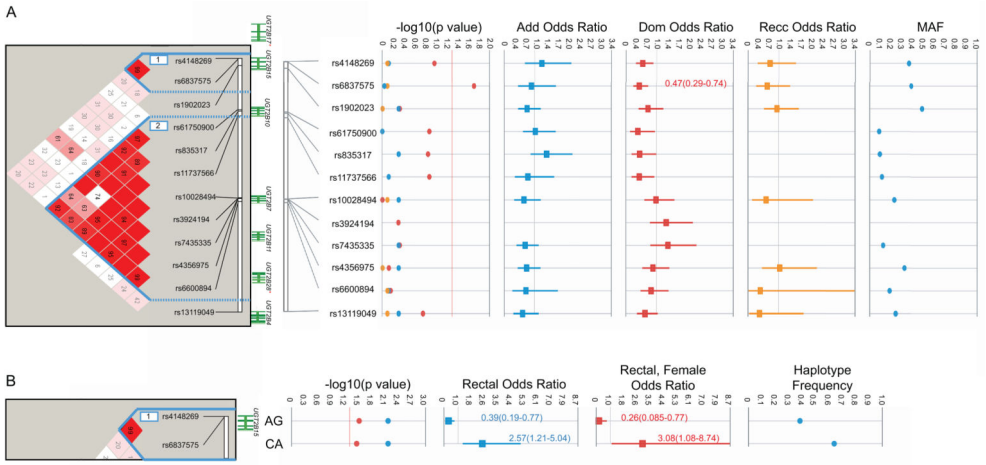

Figure 2.

Schematic representation of the effect the *UGT2B* gene family on rectal cancer risk produced by the PheWAS software (Pendergrass et al., 2012). Forest plot showing the odds ratios and 95% confidence intervals of the effect of individual *UGT2B* SNPs (**A**) and haplotypes in block 1 of *UGT2B15* (**B**) on rectal cancer risk. LD plot demonstrates the haplotype blocks within the region and the position of each *UGT2B* gene [* on genes indicates that this gene has been previously studied for CRC risk associations using the same sample set (Angstadt et al., 2013)]. The *P*-value graphed as the −1og10 (*P*-value) was adjusted for multiple testing by the FDR method and the red line denotes the $P<0.05$ cutoff. Abbreviations are as follows: B = minor allele; Add [Additive Statistical Model, (BB) > (BA) > (AA)]; Dom [(Dominant Statistical Model, (BB + (BA) vs (AA)]; Recc (Recessive Statistical Model, (BB) vs (BA) + (BB)]; MAF (minor allele frequency in population controls). Blank plots in the Additive and Recessive Statistical Model are provided to focus the graphical scale because these SNPs were insignificant and contained a large confidence interval. In addition, rs3924194 and rs7435335 only had two alleles and therefore could not be analyzed in a recessive model.

TABLE 1

Summary of Demographics and Dietary Characteristics for the Study's Caucasian Population

	UGT1A			*UGT2B*		
	Case (n=816)	**Control (n=941)**	***P*-value**	**Case (n=857)**	**Control (n=932)**	***P*-value**
Men	456 (52%)	507 (53%)	0.4	444 (52%)	494 (53%)	0.6129
Age (yrs)[a]	66.6 ± 11.7	62 ± 11.0	<0.0001	66.7 ± 11.6	62.3 ± 11.0	<0.0001
BMI (kg/m^2)[a]	29.9 ± 6.2	29.0 ± 6.0	0.0005	29.6 ± 6.3	28.8 ± 5.9	0.0015
Smoking Status			0.0216			0.0008
Never Smokers	364 (45%)	426 (45%)		375 (44%)	449 (48%)	
Formers Smokers	374 (46%)	399 (41%)		399 (46%)	358 (39%)	
Current Smokers	78 (9%)	126 (14%)		83 (10%)	125 (13%)	
Pack-years[a]	30.6 ± 26.8	26.8 ± 25.2	0.0143	29.5 ± 26.7	26.4 ± 24.9	0.065
History of CRC in 1st Degree Family	114 (17%)	95 (12%)	0.0148	133 (20%)	89 (12%)	0.0001
NSAID use			0.0036			0.0037
High, >3x/week	324 (48%)	428(55%)		326 (47%)	408 (55%)	
Low, ≤ 3x/week	354 (52%)	344 (45%)		365 (53%)	336 (45%)	
Physical Activity			<0.0001			<0.0001
High, >1h/day vigorous exercise	179 (26%)	283 (37%)		188 (27%)	281 (38%)	
Low, <1h/day vigorous exercise	499(74%)	489 (63%)		503 (73%)	463 (62%)	
Education			<0.0001			<0.0001
≤High School	390 (51%)	279 (33%)		397 (50%)	284 (34%)	
>High School	380 (49%)	573 (67%)		404 (50%)	555 (66%)	
Total energy (kcal/day)[a]	1844 ± 843	1824 ± 765	0.767	1885 ± 866	1799 ± 780	0.1393
Alcohol (g) [a]	8.7 ± 28.5	10.2 ± 24.8	0.3496	9.2 ± 29.2	10.8 ± 30.6	0.326
Carcinogens produced from cooking meat (ng/kcal/day)						
PhiP[a]	45.7 ± 71.6	45.2 ± 66.2	0.7286	46.0 ± 72.4	45.2 ± 64.3	0.5629
MeIQx[a]	18.7 ± 22.3	16.3 ± 19.4	0.0267	18.4 ± 21.1	16.7 ± 18.6	0.1612
DiMeIQx[a]	1.6 ± 2.1	1.4 ± 1.8	0.0093	1.6 ± 1.9	1.4 ± 1.8	0.0723
BaP[a]	10.4 ± 18.2	11.2 ± 16.8	0.0202	10.3 ± 18.1	11.3 ± 17.9	0.0276

Genes Chromosomes Cancer. Author manuscript; available in PMC 2015 June 01.

	UGT1A			*UGT2B*		
	Case (n=816)	**Control (n=941)**	***P*-value**	**Case (n=857)**	**Control (n=932)**	***P*-value**
Anatomical Sub-Site						
Proximal	161 (41%)			151 (39%)		
Distal	116 (30%)			119 (31%)		
Rectum	116 (30%)			116 (30%)		

[a]Mean ± SD

TABLE 2

Associations Between *UGT1A* Haplotype Blocks and CRC Risk Using an Additive Statistical Model

Category	Block	SNPs	Haplotype location (bp)[a]	Haplotype	Frequency ± Std Dev	Odds Ratio (95% CI)[b]	FDR *P*-value
Overall	4	rs1817154 rs2602373 rs2248733 rs1453322 rs7608713	234560591–234567051	T-T-A-G-A	0.06 ± 0.004	2.442 (1.291–4.619)	0.030
Male	11	rs7578153 rs10203853 rs6728940	234686157–234689164	C-A-G	0.51 ± 0.009	0.502 (0.286–0.884)	0.044
Male	11	rs7578153 rs10203853 rs6728940	234686157–234689164	C-T-G	0.14 ± 0.006	2.557 (1.099–5.949)	0.044
Proximal	2	rs17864678 rs10929251	234544610–234546229	T-G	0.17 ± 0.006	0.278 (0.112–0.693)	0.012
Proximal	6	rs17863775 rs17868322 rs2741048 rs2602379 rs10189426 rs2602381 rs7579530	234579936–234588585	A-G-A-A-C-C-A	0.50 ± 0.008	0.514 (0.289–0.913)	0.046
Proximal	7.1	rs10167119 rs4583459 rs7608175 rs6759892 rs2070959 rs1105879 rs6761246 rs1604144 rs12988520 rs7420193	234589312–234611523	T-T-C-T-A-T-C-A-C-C	0.12 ± 0.005	0.241 (0.085–0.688)	0.031
Proximal	7.4	rs1604144 rs12988520 rs7420193 rs6725478 rs7572563 rs28898590 rs4556969 rs12475068 rs1875263 rs6755571 rs7597496 rs4341922	234605835–234630443	A-C-C-C-G-G-C-C-C-C-A-C	0.13 ± 0.005	0.255 (0.091–0.71)	0.036
Proximal Male	5	rs13418420 rs17864683	234578762–234579209	C-A	0.27 ± 0.008	3.073 (1.126–7.484)	0.027
Proximal Male	5	rs13418420 rs17864683	234578762–234579209	T-A	0.71 ± 0.008	0.392 (0.17–0.904)	0.028
Proximal Male	11	rs7578153 rs10203853 rs6728940	234686157–234689164	C-T-G	0.14 ± 0.006	4.055 (1.303–12.62)	0.047
Distal	2	rs17864678 rs10929251	234686157–234689164	T-A	0.83 ± 0.006	2.693 (1.13–6.419)	0.032
Distal	2	rs17864678 rs10929251	234686157–234689164	T-G	0.17 ± 0.006	0.324 (0.116–0.907)	0.032
Distal Male	9	rs4148328 rs11888492	234677659–234679974	C-G	0.10 ± 0.005	4.753 (1.462–15.45)	0.029

[a] All SNPs on Chr2q37

[a] Adjusted for age, education, sex, BMI, family history, NSAID use, and physical activity.

TABLE 3

Interaction Effects of High NSAID Use and *UGT1A* Polymorphisms on CRC Risk

Statistical model	Haplotype block	SNP	Haplotype	Interaction expected odds ratio	FDR interaction *P*-value
Recessive		rs1500482		0.103	0.051
Additive	10	rs6717546, rs1500482, rs7586006	A-G-T	0.347	0.027