



Published in final edited form as:

Mol Carcinog. 2010 November ; 49(11): 974–980. doi:10.1002/mc.20682.

Interactions between cigarette smoking and selected polymorphisms in xenobiotic metabolizing enzymes in risk for colorectal cancer: a case-only analysis

Mala Pande¹, Chris I. Amos^{1,2,3}, Cathy Eng⁴, and Marsha L. Frazier^{1,3}

¹ Department of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas

² Department of Bioinformatics and Computational Biology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas

³ Program in Human and Molecular Genetics, The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, Texas

⁴ Department of Gastrointestinal Medical Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas

Abstract

Objective—The metabolism of xenobiotics is complex and involves multiple steps and multiple enzymes. Genetic variation in the genes encoding these enzymes as well as the level of exposure to the substrates of these enzymes could alter metabolism and clearance of potential carcinogens and thus alter cancer susceptibility. This study examined interaction effect between smoking and two single nucleotide polymorphisms (SNPs)—*CYP1A1* c.1384A>G (p.Ile462Val) and *EPHX1* c.337T>C (p.Tyr113His)—in modulating colorectal cancer (CRC) risk. The SNPs were selected a priori based on functional significance.

Methods—In a case-only analysis, unconditional logistic regression was used to examine the associations between smoking and each SNP and between the two SNPs in 786 patients with nonfamilial CRC.

Results—There was significant multiplicative interaction for CRC risk between smoking and *EPHX1* c.337T>C (odds ratio [OR] = 1.37, 95% confidence interval [CI] = 1.03–1.81, $P = 0.03$), particularly among smokers with a history of greater than 20 pack-years of smoking (OR = 1.52, 95% CI = 1.07–2.16, $P = 0.02$). In addition, there was gene-gene interaction between *EPHX1* c.337T>C and *CYP1A1* c.1384A>G (OR = 1.61, 95% CI = 1.02–2.55, $P = 0.04$).

Conclusion—Smokers with any variant allele of *EPHX1* were at increased risk for CRC, as were individuals with any variant allele of *CYP1A1* together with any variant allele of *EPHX1*. Thus, the study of gene-environment and gene-gene interactions may help to identify high-risk subgroups that can be targeted for intensive smoking cessation and CRC screening interventions.

Keywords

Genetic epistasis; Microsomal Epoxide Hydrolase; Cytochrome P-450 CYP1A1; gene-environment interaction

Introduction

Colorectal cancer (CRC) is the third most common cancer in both men and women, representing 10% of all cancers in men and women [1]. It is also the second most frequent cause of cancer mortality, accounting for 9% of all deaths from cancer in men and women. An estimated 146,970 incident cases of CRC and 49,920 deaths due to CRC were expected in the United States in 2009 [1]. Risk for CRC is influenced by both environmental and genetic factors. Smoking is a putative risk factor for many cancers, and mounting evidence indicates that smoking is a risk factor for CRC [2]. Genetic risk factors for CRC include the highly penetrant but rare mutations seen in familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer syndrome (HNPCC or Lynch syndrome), which demonstrate a Mendelian inheritance and account for ~5% of CRC cases [3]. There is evidence that other, more common, low-penetrance gene mutations may also be associated with risk for CRC (reviewed in [4]).

Cigarette smoking is a widespread source of exposure to carcinogens like polycyclic aromatic hydrocarbons (PAHs) and nitrosamines. These potentially carcinogenic substances are metabolized by many different metabolic enzymes that vary in their expression and activity levels due to polymorphisms in the genes encoding them. These enzymatic variations can result in differences in the metabolism and clearance of carcinogens and therefore modify cancer risk.

The enzyme that principally metabolizes PAHs is encoded by the *CYP1A1* gene. Two single nucleotide polymorphisms (SNPs) in *CYP1A1* have frequently been examined for their relationship to cancer susceptibility: a T>C polymorphism in the 3'-untranslated region, which creates an *MspI* restriction site, and a c.1384A>G polymorphism that results in an amino acid change from isoleucine to valine (p.Ile462Val). The process by which PAHs are metabolized involves a phase I activation reaction in which highly reactive intermediates with mutagenic potential are generated [5,6]. The *CYP1A1* c.1384A>G SNP is associated with increased activity of the enzyme [7] and thus increased activation and carcinogenic potential of PAHs and the associated activated metabolites.

PAHs are also substrates for the *EPHX1*-encoded enzyme microsomal epoxide hydrolase, which hydrolyzes epoxides; the metabolic process can result in both detoxified or more toxic product [8]. Variations in *EPHX1*, such as non-synonymous SNPs c.337T>C (p.Tyr113His) and c.416A>G (p.His139Arg), influence enzyme activity and may thus modulate the risk of PAH-mediated CRC, especially in smokers. Specifically, the *EPHX1* c.337T>C variant allele is associated with reduced enzyme activity [8].

Earlier, in a retrospective study on a cohort of individuals with Lynch syndrome (individuals with a genetic predisposition for CRC due to an inherited pathogenic mutation in one of the DNA mismatch repair genes), we reported on the effect of polymorphisms in certain candidate xenobiotic metabolizing genes on risk for CRC [9]. We found that the *CYP1A1* c.1384A>G and *MspI* variant alleles were associated with an earlier age at onset of CRC, and the data were suggestive of an interaction between *CYP1A1* c.1384A>G and *EPHX1* c.337T>C SNPs in modulating age-associated CRC risk (P for interaction term = 0.036; Wald $\chi^2 P = 0.044$).

Since the PAHs present in cigarette smoke are metabolized by polymorphic enzymes it is reasonable to assume that the effect of genetic background on CRC risk could be modified by smoking. However, although gene-gene interaction was explored in our previous study on Lynch syndrome, the sample size was too small (smoking and genotype information were available for only 167 individuals) for examination of possible gene-environment interaction between smoking and the metabolic gene SNPs that were found to be associated with CRC

risk. Therefore, in the current study, we examined a series of 786 nonfamilial CRC cases to test the hypothesis that smoking interacts with the polymorphic *CYP1A1* c.1384A>G or *EPHX1* c.337T>C in influencing risk for CRC. We also used this sample of sporadic CRC cases to validate the gene-gene interactions observed between *CYP1A1* c.1384A>G and *EPHX1* c.337T>C in our previous study and to test the hypothesis that individuals with variant alleles of both SNPs are at significantly higher risk for CRC.

Methods

Study population

The study population consisted of a series of 794 patients with histopathologically confirmed CRC enrolled as participants in TexGen at The University of Texas M. D. Anderson Cancer Center from April 2002 to July 2007. The TexGen research project is an ongoing project that collects and stores biological material (blood and tissue) for future genetic and medical research at the Texas Medical Center. TexGen enrolled all new patients in the Gastrointestinal, Genitourinary, and Urology Centers at M. D. Anderson Cancer Center who were age 18 years or older and consented to the study. We studied all patients from the TexGen series with non-syndromic CRC (any patients with known hereditary or familial CRC were excluded). Blood samples were collected in EDTA tubes at the same time that the patient underwent venipuncture for other tests. The samples were processed on the same day or the next day and stored at -80°C . The process for DNA extraction and purification involved spinning the blood samples at 2800 rpm for 10 min, then 1–4 ml of plasma was separated, leaving the buffy coat for DNA extraction. DNA was extracted on an Autopure LS automated DNA purification instrument (Gentra Systems Inc., Minneapolis, MN) according to the manufacturer's instructions. The steps for purification of DNA included lysis of red blood cells and nucleated cells, protein precipitation to pellet cellular proteins, DNA recovery using 100% isopropanol, washing of the DNA pellet with 70% isopropanol, and DNA hydration before storage.

The demographic and epidemiological data for the patients in this study were obtained from the Patient History Database questionnaire. This questionnaire is a medical intake form that all newly registered patients are required to complete at presentation to M. D. Anderson Cancer Center. The form has been in use since December 1999, and 93% of all newly registered patients complete the questionnaire. The form is a template that the clinician uses to guide the primary medical evaluation, and besides eliciting smoking history, it asks for core risk factor, family history, demographic, and quality-of-life information that the patient self-reports. Certified clinical coding specialists abstract the data and enter them into a web-based Oracle database. Quality control checks are performed at regular intervals. The Institutional Review Board approved the use of TexGen samples and data for this study.

Participants were characterized at study enrollment as “never-smokers” if they had smoked fewer than 100 cigarettes in their lifetime and as “ever-smokers” if they had smoked at least 100 cigarettes. The ever-smokers were further divided into “former smokers” and “current smokers”. Former smokers were defined as subjects who had quit smoking at least 1 year before their cancer diagnosis, and current smokers were defined as those who were smoking within 1 year before their cancer diagnosis. Smoking dose was calculated in pack-years by multiplying the number of packs of cigarettes smoked per day by the number of years of smoking. On merging the TexGen and Patient History Database data files, we excluded 6 individuals who had incomplete smoking data, so that our final study sample consisted of 794 individuals.

Genotyping

Genotyping for *EPHX1* c.337T>C polymorphism was performed with the Taqman assay using protocols described by the manufacturer (Applied Biosystems, Foster City, CA). Primer and probe sequences were obtained from the National Cancer Institute's SNP500Cancer database (<http://snp500cancer.nci.nih.gov/>). The dual-384-well GeneAmp PCR system 9700 (Applied Biosystems) was used for polymerase chain reactions (5 μ l) that included 5 ng of sample DNA, 2.5 μ l of 2 X Taqman genotyping master mix (Applied Biosystems), and 900 nM for each primer and 200 nM for each probe. The polymerase chain reaction conditions were 95°C for 10 min, followed by 50 cycles of 92°C for 30 s and 60°C for 1 min. The reacted plates were then read using the ABI Prism 7900HT sequence detection system, and genotypes were automatically called by the built-in software. Positive and negative controls were used in each genotyping assay, and more than 5% of samples were randomly selected and run in duplicates with 100% concordant results.

The *CYP1A1* c.1384A>G polymorphism was not amenable to genotyping on the Taqman platform, so genotyping was performed on a pyrosequencer (PSQ 96 System; Biotage, Inc.) using the method described previously [9,10]. For quality control, 5% of the samples were run as duplicates in the pyrosequencing reactions with no conflicting results.

Statistical Analysis

We estimated the genotype frequencies for *CYP1A1* c.1384A>G and *EPHX1* C.337T>C and tested each SNP for Hardy-Weinberg equilibrium. A case-only design was used to examine the interactions between smoking and each of these SNPs and also gene-gene interaction between these two SNPs.

In a case-only analysis, it has been shown that the odds ratio (OR) for the association between the environmental exposure and the gene is a function of the OR for the environmental factor alone, the genotype alone, and their joint effects in a regular case-control study [11]. Khoury and Flanders [12] have elaborated that the case-only OR can be interpreted as the multiplicative interaction between gene and environment under the assumption that the gene and environment are independent of each other in risk for disease. To address the independence assumption between the genes and smoking in our study, it can be stated that individuals do not routinely know their genetic status for *CYP1A1* or *EPHX1* SNPs and that neither of these genes is known to be associated with smoking behavior, so it is unlikely that an individual's genotype influences whether or not he or she smokes. Another method to determine independence between exposures is to examine their relationship in a sample of controls from the same base population as the cases, as suggested by Piegorsch et al. [11]. A lack of association between the exposures among the controls would indicate that the exposures are independent. Therefore, to determine whether smoking and genotypes for each of the SNPs were independent, we examined their relationship in a sample of 722 controls from the same base population as the cases (these were controls for a lung cancer study at M. D. Anderson Cancer Center for whom genotyping data for the two SNPs were available [13]), and we found no differences in genotype frequencies between smokers and nonsmokers for either SNP ($\chi^2 P > 0.05$). Therefore, we made the assumption that the gene and environmental exposures were independent and our interaction estimates were valid. Similarly, in testing for gene-gene interaction, linkage disequilibrium may cause nonindependence between the genes, which can invalidate a case-only design to measure gene-gene interaction [14]. Since genes on different chromosomes are unlikely to be correlated and the genes we tested are located on separate chromosomes (*CYP1A1* on chromosome 15 and *EPHX1* on chromosome 1), the independence assumption for these two SNPs was not violated.

Unconditional logistic regression was used to estimate the case-only ORs and 95% confidence intervals (CIs) to assess the interaction effect of smoking and each of the SNPs and to determine the influence of interaction between the two SNPs on CRC risk. Data for the *CYP1A1* SNP were dichotomized by combining the heterozygous and homozygous variant genotypes (AG+GG) because of the low frequency of the GG genotype. The *EPHX1* SNP c.337T>C was examined in both an additive and a codominant model. The wild-type genotypes formed the reference group. Gender, age and ethnicity were included as covariates in the regression model as potential confounders. The interaction effects were also examined by cancer stage (localized vs. distant) and site (colon vs. rectal). Lastly, in stratified analyses, the gene-gene interaction was analyzed within smoker/non-smoker and gender subgroups. All statistical analyses were performed using STATA 8.0 statistical software (StataCorp LP, College Station, TX).

Results

Of the 794 participants, we had complete genotyping data for 786. Genotyping of the remaining 8 samples failed for one or both SNPs, and these samples were therefore dropped from the analysis.

Participant characteristics and genotype frequencies are presented in Table 1. The study population was mostly non-Hispanic White (80%), and there were more men (59%) than women. Smokers constituted 47.2% of the population, and among the smokers, about half (50.4%) had a history of greater than 20 pack-years of smoking. Both of the SNPs were in Hardy-Weinberg equilibrium (*CYP1A1* c.1384A>G, $\chi^2 P = 0.73$; *EPHX1* c.337T>C, $\chi^2 P = 0.36$).

Information about gene-environment and gene-gene interactions and CRC risk is presented in Table 2. The case-only OR comparing the CRC risk between ever-smokers with the variant *EPHX1* c.337T>C genotypes and never-smokers with the homozygous wild-type genotype was statistically significant (OR = 1.37, 95% CI = 1.03–1.81, $P = 0.03$), suggesting the presence of an interaction between smoking and *EPHX1* (Table 2). Further, the interaction effect between *EPHX1* c.337T>C and greater than 20 pack-years of smoking was also significant (OR = 1.52, 95% CI = 1.07–2.16, $P = 0.02$), suggesting that the *EPHX1* c.337T>C variant genotypes influence susceptibility for CRC particularly among heavy smokers (>20 pack years). There was no evidence of a similar interaction between smoking and *CYP1A1* c.1384A>G (OR = 0.73, 95% CI = 0.46–1.14, $P = 0.17$).

The OR for gene-gene interaction between the two SNPs was significant for individuals with any variant allele of *CYP1A1* c.1384A>G and any variant allele of *EPHX1* c.337T>C (OR = 1.61, 95% CI = 1.02–2.55, $P = 0.04$) and more so for individuals with the *EPHX1* homozygous variant genotype (OR = 2.96, 95% CI = 1.57–5.57, $P = 0.001$) compared to individuals with the homozygous wild-type genotypes of the two SNPs (referent), suggesting the presence of an interaction between these two SNPs. Adjusting for gender, age and ethnicity did not significantly alter any of the main effect estimates (results not shown); therefore, the unadjusted estimates are presented (Table 2).

Although the power was limited for subgroup analyses, on stratification by smoking status, the OR for interaction effect was statistically significant for the *CYP1A1* any variant and *EPHX1* homozygous variant genotype (CC) in both never-smokers (OR = 2.67, 95% CI = 1.16–6.15, $P = 0.02$) and ever-smokers (OR = 3.59, 95% CI = 1.33–9.65, $P = 0.01$), but the OR was much higher in ever-smokers (Table 3). The OR for interaction effect did not differ by tumor stage or site. Similarly, there was no interaction effect between gender and either

of the SNPs. However, in subgroup analysis by gender, a gene-gene interaction effect was seen in males but not in females (Table 4).

Discussion

In this case-only analysis, there was evidence of departure from multiplicativity indicating a gene-environment interaction between *EPHX1* c.337T>C and smoking, particularly in smokers with a history of greater than 20 pack-years of smoking. In addition, there was a gene-gene interaction between *CYP1A1* c.1384A>G and *EPHX1* c.337T>C. Therefore, our results suggest that (1) the impact of cigarette smoking on CRC risk is synergistically increased among individuals who carry a variant allele of *EPHX1* c.337T>C and (2) individuals who carry variant alleles of both *CYP1A1* c.1384A>G and *EPHX1* c.337T>C are more susceptible to CRC. Furthermore, in an analysis stratified by smoking status, we found that the OR for gene-gene interaction effect was statistically significant for the *EPHX1* homozygous variant genotype (CC) in both never-smokers and ever-smokers but that the OR was higher in ever-smokers. Similarly, the OR for gene-gene interaction was significant in males but not in females, although overall, there was no significant interaction between gender and either of the SNPs. These results were based on a limited number of patients, but they suggest that compared with nonsmokers with wild-type alleles for both SNPs, smokers with variant alleles of *CYP1A1* c.1384A>G and smokers with the homozygous variant genotype for *EPHX1* c.337T>C are at significantly increased risk for CRC and the gene-gene interaction effect may be gender specific, as it was evident only in males.

In our previous study on individuals with Lynch syndrome, we found evidence for multiplicative interaction between *CYP1A1* c.1384A>G and *EPHX1* c.337T>C (P for interaction term = 0.036; Wald χ^2 P = 0.044) with a greater than multiplicative hazard ratio for the combined effect of having a variant allele of each of these SNPs (hazard ratio, 3.09, 95% CI, 1.58–6.04; P = 0.001). The purpose of the present study was to see if these findings were replicable in cases of nonsyndromic CRC. Even though the study designs used to obtain the interaction estimates were different (retrospective cohort versus case-only design), the results from both studies were statistically significant.

Most of the association studies for *EPHX1* and smoking in colorectal carcinogenesis have examined colorectal adenoma as the outcome [15–18]. Of three studies with CRC as the outcome one reported an increased frequency of the c.337C variant allele in CRC cases compared to controls [19], one found a reduced CRC risk associated with the c.337C allele [20] and one large case-control study found no association [21]. Similarly, there have been conflicting results for the interaction effect between smoking and *EPHX1*. Ulrich et al. [18] found that the variant *EPHX1* c.337C allele increased adenoma risk among smokers and the risk was highest among those with greater than 25 pack-years of smoking (similar to our findings for CRC risk), whereas other studies reported reduced risk associated with the low activity c.337C allele in the presence of smoking [15–17]. Further validation of the influence of *EPHX1* on risk of CRC in the presence of smoking may therefore be warranted.

Few studies have evaluated the association between *CYP1A1* and CRC, though *CYP1A1* has been extensively evaluated in other smoking-related cancers (reviewed in [22]). Slattery et al. examined the association between *CYP1A1* c.1384A>G and smoking in risk of CRC in 1026 cases and 1185 controls and found that the individuals at highest risk for CRC were men who were currently smoking and had any *CYP1A1* variant allele. The authors concluded that the impact of smoking on CRC risk may depend on *CYP1A1* genotype [23]. Fan et al. used a case-only study similar to our own study to determine the interactions between certain polymorphisms in metabolic enzymes and smoking in 207 Chinese patients with CRC and found a significant gene-gene interaction between *CYP1B1* 1294G and

SULT1A1 638A alleles (OR 2.68, 95% CI = 1.16–6.26) and gene-environment interaction between *CYP1B1* 1294G and smoking (OR 2.62, 95% CI = 1.01–6.72) [24]. However, the results of their study and our study cannot be compared since the polymorphisms examined in the two studies were different.

PAHs in cigarette smoke are substrates for both *CYP1A1* and *EPHX1*, and these two enzymes act sequentially to metabolize PAHs. Therefore, a biological interaction effect may exist between *CYP1A1* and *EPHX1*. First, *CYP1A1* converts benzo(a)pyrene to the active benzo(a)pyrene 7,8 epoxide. This is then hydrated by *EPHX1* to a transhydrodiol derivative, benzo(a)pyrene 7,8 diol, a product that is less toxic [16]. However, the diol derivative is also a primary substrate for CYP enzymes that oxidize it further to benzo(a)pyrene 7,8 dihydrodiol 9,10 epoxide (BPDE), which is highly reactive and capable of forming DNA adducts. Therefore, these genes may interact to play a more complex role in cancer susceptibility.

The case-only approach was appropriate for our study since it was used to validate a priori findings of an interaction effect between two SNPs. However, a case-only study does have the disadvantage of not allowing evaluation of the independent effect of either of the exposures, smoking alone or the *CYP1A1* and *EPHX1* genotypes alone, but only allowing evaluation of their interactions. It also does not allow assessment for departures from joint additive effects (can only test departures from joint multiplicative effects) of the exposure and genotype or the genotypes with each other. However, the case-only design is efficient (smaller sample size required to assess interaction than in a case-control design) [14] and offers less potential for misclassification of exposures.

Though both *CYP1A1* c.1384A>G and *EPHX1* c.337T>C are nonsynonymous SNPs and therefore likely have functional consequences, we queried two programs (PolyPhen [<http://genetics.bwh.harvard.edu/pph/>] and SIFT [<http://blocks.fhcrc.org/sift/SIFT.html>]) that predict the impact of an amino acid substitution on protein function to assess whether these SNPs are potentially deleterious. Polyphen classified *CYP1A1* c.1384A>G as “benign,” and SIFT classified it as “tolerated.” *EPHX1* c.337T>C, on the other hand was classified as “possibly damaging” by Polyphen and “intolerant” by SIFT. The *EPHX1* c.337T>C SNP may therefore be important to follow up on in functional studies in CRC.

Our study was underpowered for subgroup analysis by race/ethnicity due to the low frequency of the *CYP1A1* variant genotype. The OR for interaction between *CYP1A1* c.1384A>G and *EPHX1* c.337T>C (homozygous variant) lost significance when we analyzed the non-Hispanic Whites alone ($OR_{(EPHX1:CC \text{ genotype})} = 2.2$, 95% CI = 0.87–5.56, $P = 0.09$). However, evidence for interaction between *EPHX1* c.337T>C and smoking remained significant when the analysis was limited to non-Hispanic whites—that is, among whites, ever-smokers with one or two copies of the *EPHX1* variant c.337C allele had OR = 1.4 (95% CI = 1.05–1.97, $P = 0.02$) compared to never-smokers with the TT genotype. Overall, including ethnicity as a covariate while testing for interaction, did not alter the main effect estimates. Analyzing interaction effects in a larger sample of each ethnic subgroup may be required to validate these findings more globally.

Our finding of evidence for a gene-gene interaction between *CYP1A1* c.1384A>G and *EPHX1* c.337T>C in risk for sporadic CRC is especially meaningful, since to our knowledge this interaction has not previously been described and since this finding validates a similar interaction seen in our previously reported study [9] in a different study population. Future plans for these findings would include evaluating the gene-gene and gene-environment relationship as predictors of CRC recurrence and survival. In conclusion, while low penetrance genes like *CYP1A1* and *EPHX1* may raise the cancer risk only slightly

independently, in combination they may greatly increase cancer susceptibility. Therefore, individuals who have multiple genetic susceptibility alleles and are smokers may be a subgroup that could be targeted for more intensive interventions than is recommended for the general population.

Acknowledgments

Grant support: This study was supported in part by NIH Cancer Center Support grant CA16672, and National Cancer Institute grants CA57730, CA070759 and CA133996.

We thank Dr. Kelly Merriman and Elizabeth Thompson for their assistance in providing the TexGen data and samples.

Abbreviations

CRC	colorectal cancer
SNP	single nucleotide polymorphism
PAH	polycyclic aromatic hydrocarbon

Reference List

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin.* 2009; 59:225–249. [PubMed: 19474385]
2. Giovannucci E. An updated review of the epidemiological evidence that cigarette smoking increases risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev.* 2001; 10:725–731. [PubMed: 11440957]
3. Rowley PT. Inherited susceptibility to colorectal cancer. *Annu Rev Med.* 2005; 56:539–554. [PubMed: 15660526]
4. de Jong MM, Nolte IM, te Meerman GJ, van der Graaf WT, de Vries EG, Sijmons RH, et al. Low-penetrance genes and their involvement in colorectal cancer susceptibility. *Cancer Epidemiol Biomarkers Prev.* 2002; 11:1332–1352. [PubMed: 12433710]
5. Gonzalez FJ, Gelboin HV. Role of human cytochromes P450 in the metabolic activation of chemical carcinogens and toxins. *Drug Metab Rev.* 1994; 26:165–183. [PubMed: 8082563]
6. Parke DV. The cytochromes P450 and mechanisms of chemical carcinogenesis. *Environ Health Perspect.* 1994; 102:852–853. [PubMed: 9644192]
7. Crofts F, Taioli E, Trachman J, Cosma GN, Currie D, Toniolo P, et al. Functional significance of different human CYP1A1 genotypes. *Carcinogenesis.* 1994; 15:2961–2963. [PubMed: 8001264]
8. Omiecinski CJ, Hassett C, Hosagrahara V. Epoxide hydrolase--polymorphism and role in toxicology. *Toxicol Lett.* 2000; 112–113:365–370.
9. Pande M, Amos CI, Osterwisch DR, Chen J, Lynch PM, Broaddus R, et al. Genetic variation in genes for the xenobiotic-metabolizing enzymes CYP1A1, EPHX1, GSTM1, GSTT1, and GSTP1 and susceptibility to colorectal cancer in Lynch syndrome. *Cancer Epidemiol Biomarkers Prev.* 2008; 17:2393–2401. [PubMed: 18768509]
10. Pande M, Chen J, Amos CI, Lynch PM, Broaddus R, Frazier ML. Influence of methylenetetrahydrofolate reductase gene polymorphisms C677T and A1298C on age-associated risk for colorectal cancer in a caucasian lynch syndrome population. *Cancer Epidemiol Biomarkers Prev.* 2007; 16:1753–1759. [PubMed: 17855693]
11. Piegorsch WW, Weinberg CR, Taylor JA. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. *Stat Med.* 1994; 13:153–162. [PubMed: 8122051]
12. Khoury MJ, Flanders WD. Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! *Am J Epidemiol.* 1996; 144:207–213. [PubMed: 8686689]

13. Zhao H, Spitz MR, Gwyn KM, Wu X. Microsomal epoxide hydrolase polymorphisms and lung cancer risk in non-Hispanic whites. *Mol Carcinog.* 2002; 33:99–104. [PubMed: 11813302]
14. Yang Q, Khoury MJ, Sun F, Flanders WD. Case-only design to measure gene-gene interaction. *Epidemiology.* 1999; 10:167–170. [PubMed: 10069253]
15. Cortessis V, Siegmund K, Chen Q, Zhou N, Diep A, Frankl H, et al. A case-control study of microsomal epoxide hydrolase, smoking, meat consumption, glutathione S-transferase M3, and risk of colorectal adenomas. *Cancer Res.* 2001; 61:2381–2385. [PubMed: 11289100]
16. Huang WY, Chatterjee N, Chanock S, Dean M, Yeager M, Schoen RE, et al. Microsomal epoxide hydrolase polymorphisms and risk for advanced colorectal adenoma. *Cancer Epidemiol Biomarkers Prev.* 2005; 14:152–157. [PubMed: 15668489]
17. Tranah GJ, Giovannucci E, Ma J, Fuchs C, Hankinson SE, Hunter DJ. Epoxide hydrolase polymorphisms, cigarette smoking and risk of colorectal adenoma in the Nurses' Health Study and the Health Professionals Follow-up Study. *Carcinogenesis.* 2004; 25:1211–1218. [PubMed: 14988221]
18. Ulrich CM, Bigler J, Whitton JA, Bostick R, Fosdick L, Potter JD. Epoxide hydrolase Tyr113His polymorphism is associated with elevated risk of colorectal polyps in the presence of smoking and high meat intake. *Cancer Epidemiol Biomarkers Prev.* 2001; 10:875–882. [PubMed: 11489754]
19. Harrison DJ, Hubbard AL, MacMillan J, Wyllie AH, Smith CA. Microsomal epoxide hydrolase gene polymorphism and susceptibility to colon cancer. *Br J Cancer.* 1999; 79:168–171. [PubMed: 10408710]
20. Sachse C, Smith G, Wilkie MJ, Barrett JH, Waxman R, Sullivan F, et al. A pharmacogenetic study to investigate the role of dietary carcinogens in the etiology of colorectal cancer. *Carcinogenesis.* 2002; 23:1839–1849. [PubMed: 12419832]
21. Robien K, Curtin K, Ulrich CM, Bigler J, Samowitz W, Caan B, et al. Microsomal epoxide hydrolase polymorphisms are not associated with colon cancer risk. *Cancer Epidemiol Biomarkers Prev.* 2005; 14:1350–1352. [PubMed: 15894702]
22. Bartsch H, Nair U, Risch A, Rojas M, Wikman H, Alexandrov K. Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol Biomarkers Prev.* 2000; 9:3–28. [PubMed: 10667460]
23. Slattery ML, Samowitz W, Ma K, Murtaugh M, Sweeney C, Levin TR, et al. CYP1A1, cigarette smoking, and colon and rectal cancer. *Am J Epidemiol.* 2004; 160:842–852. [PubMed: 15496536]
24. Fan C, Jin M, Chen K, Zhang Y, Zhang S, Liu B. Case-only study of interactions between metabolic enzymes and smoking in colorectal cancer. *BMC Cancer.* 2007; 7:115. [PubMed: 17603900]

Table 1

Case Characteristics and Genotype Frequencies

Characteristic	No. (%) of Patients
Gender	
Male	466 (59.3)
Female	320 (40.7)
Ethnicity	
Non-Hispanic White	633 (80.5)
Hispanic	68 (8.7)
Black	59 (7.5)
Asian/Pacific Islander	15 (1.9)
Other	11 (1.4)
Age, years	
Median	59
Mean \pm SD	58.4 \pm 12.7
Smoking status	
Never-smoker	415 (52.8)
Ever-smoker	371 (47.2)
Former smoker	313 (39.8)
Current smoker	58 (7.4)
Pack-years of smoking	
\leq 20	184 (49.6)
$>$ 20	187 (50.4)
CYP1A1 c.1384A>G	
AA	699 (88.9)
AG	84 (10.7)
GG	3 (0.38)
EPHX1 c.337T>C	
TT	380 (48.4)
TC	324 (41.2)
CC	82 (10.4)

SD, standard deviation.

Table 2
Gene-environment and Gene-gene Interaction between Smoking and Metabolic Gene SNPs *CYP1A1* c.1384A>G and *EPHX1* c.337T>C

	No. (%) of Patients	OR (95% CI)	P	OR (95% CI)*	P
Gene-Gene interaction					
<i>CYP1A1</i> c.1384A>G (AA as referent; n=699)					
<i>EPHX1</i> c.337T>C (TT as referent; n=380)					
TC + AG/GG	36	1.31 (0.78–2.16)	0.28	1.3 (0.77–2.2)	0.32
CC + AG/GG	18	2.96 (1.57–5.57)	0.001	2.96 (1.49–5.89)	0.002
TC/CC + AG/GG	54	1.61 (1.02–2.55)	0.04	1.6 (0.98–2.6)	0.06
Gene-Smoking interaction					
<i>CYP1A1</i> c.1384A>G (AA as referent; n=699)					
Smoking status (never-smoker as referent; n=415)					
Ever + AG/GG	35	0.73 (0.46–1.14)	0.17	0.77 (0.47 – 1.27)	0.31
Former + AG/GG	27	0.66 (0.4–1.08)	0.09	0.7 (0.41 – 1.2)	0.2
Current + AG/GG	8	1.11 (0.5–2.49)	0.79	1.15 (0.49 – 2.71)	0.75
Pack-years of smoking (never-smoker as referent; n=415)					
≤20 + AG/GG	21	0.9 (0.52–1.54)	0.7	0.83 (0.47 – 1.49)	0.54
>20 + AG/GG	14	0.56 (0.30–1.05)	0.07	0.7 (0.36 – 1.36)	0.29
<i>EPHX1</i> c.337T>C (TT as referent; n=380)					
Smoking status (never-smoker as referent; n=415)					
Ever + TC/CC	207	1.37 (1.03–1.81)	0.03	1.35 (1.01 – 1.8)	0.04
Former + TC/CC	176	1.39 (1.04–1.87)	0.03	1.37 (1.01 – 1.85)	0.04
Current + TC/CC	31	1.25 (0.72–2.16)	0.43	1.24 (0.71 – 2.16)	0.44
Pack-years of smoking (never-smoker as referent; n=415)					
≤20 + TC/CC	99	1.21 (0.85–1.71)	0.27	1.19 (0.84 – 1.69)	0.32
>20 + TC/CC	107	1.55 (1.09–2.2)	0.01	1.54 (1.07 – 2.21)	0.02

OR, odds ratio; CI, confidence interval.

* OR adjusted for age, gender and race (White vs. other)

Table 3

Gene-gene Interaction between Metabolic Gene SNPs *CYP1A1* c.1384A>G and *EPHX1* c.337T>C Stratified by Smoking and Gender

	No. (%) of Patients	OR (95% CI)	P	OR (95% CI)*	P
Never-smoker					
<i>CYP1A1</i> c.1384A>G (AA as referent; n=363)					
<i>EPHX1</i> c.337T>C (TT as referent; n=216)					
TC + AG/GG	20	1.29 (0.68–2.47)	0.43	1.14 (0.57 – 2.3)	0.7
CC + AG/GG	10	2.67 (1.16–6.15)	0.02	3.07 (1.2 – 7.81)	0.02
TC/CC + AG/GG	30	1.57 (0.87–2.82)	0.13	1.46 (0.78 – 2.77)	0.24
Ever-smoker					
<i>CYP1A1</i> c.1384A>G (AA as referent; n=336)					
<i>EPHX1</i> c.337T>C (TT as referent; n=164)					
TC + AG/GG	16	1.46 (0.66–3.26)	0.35	1.66 (0.72 – 3.83)	0.23
CC + AG/GG	8	3.59 (1.33–9.65)	0.01	3.33 (1.16 – 9.54)	0.02
TC/CC + AG/GG	24	1.82 (0.87–3.84)	0.11	1.99 (0.91 – 4.34)	0.08
Males					
<i>CYP1A1</i> c.1384A>G (AA as referent; n=415)					
<i>EPHX1</i> c.337T>C (TT as referent; n=221)					
TC + AG/GG	24	2.13 (1.07 – 4.26)	0.03	2.21 (1.05 – 4.65)	0.04
CC + AG/GG	13	4.58 (2.0 – 10.44)	<0.001	4.53 (1.82 – 11.27)	0.001
TC/CC + AG/GG	37	2.63 (1.38 – 5.0)	0.003	2.7 (1.35 – 5.4)	0.005
Females					
<i>CYP1A1</i> c.1384A>G (AA as referent; n=284)					
<i>EPHX1</i> c.337T>C (TT as referent; n=159)					
TC + AG/GG	12	0.72 (0.34 – 1.55)	0.41	0.75 (0.34 – 1.65)	0.47
CC + AG/GG	5	1.67 (0.57 – 4.94)	0.35	1.76 (0.56 – 5.56)	0.33
TC/CC + AG/GG	17	0.87 (0.43 – 1.74)	0.69	0.9 (0.43 – 1.85)	0.77

* OR adjusted for age, gender and race (White vs. other) for analysis stratified by smoking and age, race and smoking for analysis by gender