Published in final edited form as:
Cancer Epidemiol Biomarkers Prev. 2008 September ; 17(9): 2393–2401. doi: 10.1158/1055-9965.EPI-08-0326.

Genetic variation in genes for the xenobiotic-metabolizing enzymes *CYP1A1*, *EPHX1*, *GSTM1*, *GSTT1* and *GSTP1* and susceptibility to colorectal cancer in Lynch syndrome

Mala Pande[1], **Christopher I. Amos**[1,4], **Daniel R. Osterwisch**[1], **Jinyun Chen**[1], **Patrick M. Lynch**[2], **Russell Broaddus**[3], and **Marsha L. Frazier**[1,4,*]

[1]Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA

[2]Department of Gastrointestinal Medicine and Nutrition, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA

[3]Department of Pathology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA

[4]Program in Human and Molecular Genetics, The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX

Abstract

Individuals with Lynch syndrome are predisposed to cancer due to an inherited DNA mismatch repair gene mutation. However, there is significant variability observed in disease expression, likely due to the influence of other environmental, lifestyle, or genetic factors. Polymorphisms in genes encoding xenobiotic-metabolizing enzymes may modify cancer risk by influencing the metabolism and clearance of potential carcinogens from the body. In this retrospective analysis, we examined key candidate gene polymorphisms in *CYP1A1*, *EPHX1*, *GSTT1*, *GSTM1*, and *GSTP1* as modifiers of age at onset of colorectal cancer among 257 individuals with Lynch syndrome. We found that subjects heterozygous for *CYP1A1* I462V (c.1384A>G) developed colorectal cancer 4 years earlier than those with the homozygous wild-type genotype (median ages 39 and 43 years, respectively; log-rank test $P = 0.018$). Furthermore, being heterozygous for the *CYP1A1* polymorphisms, I462V and Msp1 (g.6235T>C), was associated with an increased risk for developing colorectal cancer [adjusted hazard ratio for AG relative to AA = 1.78, 95% CI = 1.16–2.74, $P = 0.008$; and hazard ratio for TC relative to TT = 1.53, 95% CI = 1.06–2.22, $P = 0.02$]. Since homozygous variants for both *CYP1A1* polymorphisms were rare, risk estimates were imprecise. None of the other gene polymorphisms examined were associated with an earlier onset age for colorectal cancer. Our results suggest that the I462V and Msp1 polymorphisms in *CYP1A1* may be an additional susceptibility factor for disease expression in Lynch syndrome since they modify the age of colorectal cancer onset by up to 4 years.

Keywords

xenobiotic metabolism; polymorphisms; Lynch syndrome; colorectal cancer; age at onset

[*]Request for reprints: Department of Epidemiology, Unit 1365, The University of Texas M. D. Anderson Cancer Center, 1155 Hermann P. Pressler Boulevard, Houston, TX 77030. Phone: 713-792-3393; Fax: 713-745-1163; mlfrazier@mail.mdanderson.org. .

Introduction

Lynch syndrome, commonly known as hereditary non-polyposis colorectal cancer (HNPCC), is caused by an inherited pathogenic germline mutation in one of the DNA mismatch repair (MMR) genes, *MLH1*, *MSH2*, *MSH6*, or *PMS2* (1). The mutations have an autosomal dominant pattern of inheritance and result in deficient mismatch repair, predisposing individuals with these mutations to early onset of cancers of the colon, the endometrium, and, less frequently, the stomach, ovaries, small intestine, biliary and uroepithelial tracts, skin, and brain (2).

Wide variation in disease expression both within and between families, particularly in age at onset (3), has been observed among individuals with Lynch syndrome, suggesting that other genetic and environmental factors may modify the effect of the inherited single-gene mutations. Evidence for the role of other genetic factors may also lie in the fact that 25–30% of colorectal cancer (CRC) cases that occur annually are considered familial (4) . Of these cases, about 5–14% are attributable to known inherited deleterious mutations (5). The remaining familial component of incident CRCs is likely due to the effect of variation in common low-penetrance genes or “modifier” genes (4). Therefore, it is of interest to examine low-penetrance genes as potential modifiers of risk for cancer onset in individuals with Lynch syndrome.

Cancer risk resulting from human exposure to exogenous chemicals (xenobiotics), like polycyclic aromatic hydrocarbons (PAH), that are ubiquitous environmental, dietary, and tobacco carcinogens may vary according to the ability to clear the xenobiotics from the body. Polymorphisms in the genes that encode enzymes involved in the metabolism of PAHs (that is, xenobiotic-metabolizing enzymes) such as the cytochrome P450 group (CYPs), the microsomal epoxide hydrolase group (mEH or EPHX) and the glutathione-S-transferase group (GSTs), result in varying activity levels of these enzymes, which can then influence xenobiotic clearance. The metabolism of PAHs involves both activation (phase I) and detoxification (phase II) reactions by these enzymes. During activation, reactive intermediates are formed that can bind to DNA and result in adducts that cause mutations if not repaired, thereby initiating carcinogenesis (6). It is likely that the expression and activity levels of the xenobiotic-metabolizing enzymes determine the relative level of activation and detoxification of carcinogens. These levels are important because increased levels of activation, decreased detoxification, or both, may increase cancer risk. Therefore, we hypothesize that the variation in enzyme activity due to polymorphisms in metabolic genes may explain some of the differences seen in disease risk for CRC in individuals with Lynch syndrome. We identified several candidate metabolic genes that display variation in activity levels of their expressed enzymes and selected common polymorphisms in these genes (i.e., with a minor allele frequency > 5%), to test our hypothesis.

The CYP group of enzymes is involved in the oxidation of many xenobiotics. *CYP1A1* encodes the principal enzyme that metabolizes PAHs. In the phase I reaction, PAHs are metabolically activated, which may generate highly reactive mutagenic metabolites. Two common single nucleotide polymorphisms (SNPs) in *CYP1A1*, a non-coding Msp1 polymorphism in the 3′ untranslated region (nucleotide T to C) and an exon 7 polymorphism Ile462Val, (nucleotide A to G), which are in linkage disequilibrium (LD) (7), have been commonly examined for cancer risk. The functional significance of the Msp1 polymorphism is uncertain but the I462V polymorphism results in higher enzyme activity compared to the homozygous wild-type genotype and the enzyme is expressed in the colon (8). De Jong *et al*. (9) found little evidence that polymorphic variants of *CYP1A1* influence the risk for sporadic CRC; however, Slattery *et al*. (10) have suggested that these variant alleles increase the risk

for sporadic CRC among smokers. Similarly, Talseth *et al*. (11) found an association of the *CYP1A1* Msp1 variant with an increased CRC risk in Lynch syndrome.

EPHX1 plays an important role in both the activation and detoxification of PAHs. Two *EPHX1* polymorphisms, Tyr113His and His139Arg, result in reduced and increased enzyme activity, respectively. Gsur *et al*. (12) demonstrated that genetically reduced mEH activity may be protective against lung cancer, but these same *EPHX1* polymorphisms have not been implicated in risk for sporadic CRC (13;14). We were unable to find any prior studies evaluating their role in the risk for CRC in individuals with Lynch syndrome.

GSTs are a superfamily of proteins that perform the phase II detoxification reactions of PAHs and other xenobiotics. GSTs catalyze the conjugation of reduced glutathione to a variety of potentially carcinogenic electrophilic and hydrophobic compounds. This detoxification reaction inactivates the compounds and renders them water-soluble so they can be readily excreted through urine or bile (15). GSTs include glutathione S-transferase theta1, mu1, and pi1 (*GSTT1*, *GSTM1*, and *GSTP1*). Null alleles exist as a common polymorphism for the *GSTT1* and *GSTM1* genes. Total or partial deletion of these genes results in no enzyme being produced (16;17). An overall increased risk for sporadic CRC has been described for the GSTT1 null allele, in a report by de Jong *et al*. based on 11 studies (9) but their report was inconclusive for the association, or lack thereof, between CRC and *GSTM1* or *GSTP1*. In a recent study examining Lynch syndrome individuals, Felix *et al*. (18) reported that males with the *GSTM1* and *GSTT1* null genotype were at a 3-fold increased risk for an earlier age at onset of CRC compared to those with no deletion of the genes. Another study reported a 6-year shift in the median age at onset among 150 *MLH1* mutation carriers with an earlier onset age associated with the null alleles of both *GSTM1* and *GSTT1* (19). However, two earlier studies found no influence on age at CRC onset for either *GSTM1* (20;21) or *GSTT1* polymorphisms (21). In our study, we report on some of the participants previously examined by Jones *et al*. (20), but our study includes substantially more subjects.

Although many of the genes encoding xenobiotic-metabolizing enzymes have been examined for their influence on the age at onset for CRC in Lynch syndrome, the results have not been consistent. This inconsistency might be due to the limited sample size or heterogeneous populations used in previous studies. A major strength of this study is the larger, predominantly white study population, consisting of a cohort of known carriers of pathogenic MMR gene mutations from many different families. Therefore, we are reporting on a population that is a unique resource for examining possible variation in disease expression associated with metabolic genes in this high-risk group. Specifically, we assessed the association between common polymorphisms in genes involved in xenobiotic-metabolism—*CYP1A1* (rs4646903:g.6235T>C and rs1048903:c.1384A>G p.I462V), *EPHX1* (rs1051740:c.339T>C p.Y113H and rs2234922:c.418A>G p.H139R), *GSTM1* (deletion), *GSTT1* (deletion), and *GSTP1* (rs1695:c.330A>G p.I105V and rs1138272:c. 343C>T p.A114V)—and age at onset of CRC to determine if genetic variation in any of these candidate metabolic genes modifies the age at onset of CRC in individuals with Lynch syndrome.

Materials and methods

Study population

The study population, which consisted of a cohort of individuals with Lynch syndrome has been described previously (22;23). Briefly, most Lynch syndrome probands were recruited from The University of Texas M. D. Anderson Cancer Center's gastrointestinal and other clinics from September 1994 to July 2007. Through the probands (n=119), other first-degree

relatives (n=74) and more distantly related family members (n=64) carrying the inherited mutation were recruited to the study. Study participants were from 130 families. The size of the families varied between 1 and 10 members of which 80 (31%) were singletons, 150 (58%) were between 2-7 members per family and the remaining were single families with 8, 9 and 10 members each. Criteria for mutation testing for Lynch syndrome were young age at onset of CRC (< 45 years), presence of a strong family history (met Amsterdam or relaxed Amsterdam criteria (24)), and suggestive tumor characteristics (defined as presence of microsatellite instability and loss of staining for an MMR protein). A total of 260 confirmed carriers of a MMR gene mutation were recruited. We excluded 3 subjects with mutations in *MSH6* because *MSH6* is associated with a variant form of Lynch syndrome with later age at CRC onset (25). We also did not include any individuals who tested positive for variants of unknown significance in the MMR genes since the pathogenicity of these mutations is not known. Among the remaining 257 study participants, there were 81 different MMR mutations which included deletions (25.7%), insertions (4.3%), nonsense (20.6%), splice site (27.2%) and missense mutations (22.2%). We confirmed that all of these mutations were pathogenic, particularly the missense mutations. They were either reported to be pathogenic by a Clinical Laboratory Improvement Act certified laboratory or confirmed to be pathogenic from the International Collborative Group-HNPCC InSight database (http://www.insight-group.org/) or from the published literature (26-29). All participants provided informed consent, and this retrospective study was approved by the Institutional Review Board.

For most participants (65%), information about demographic characteristics and medical history was obtained from a self-administered health, habits, and history questionnaire. Where questionnaire data were not available (i.e., when individuals were deceased or lost to follow-up), information was abstracted from the subjects' medical records. CRC date of diagnosis was confirmed for the cases by review of pathology reports. Dates of diagnosis for adenoma or other cancers were similarly confirmed from medical records, pathology reports, or both.

Genotyping

All participants provided 10-ml samples of blood from which DNA was extracted using the AUTOPURE LS Automated DNA Purification Instrument (Gentra Systems Inc., Minneapolis, MN) according to the manufacturer's instructions. Genotyping of SNPs, with the exception of *GSTM1* and *GSTT1* deletions, was done using the polymerase chain reaction (PCR) and pyrosequencing methods described by Chen *et al*. (22). The primers used in these procedures are listed in supplementary Table 1 available at the journal website online. The PCR reaction mixture was initially incubated at 95°C for 6 min, followed by 45 cycles at 95°C for 15 s, 59°C for 30 s for both *CYP1A1* SNPs, 59°C for 30 s for both *EPHX1* SNPs, 59°C for 30 s for the *GSTP1* A114V SNP, 61.5°C for 30 s for the *GSTP1* I105V SNP and 56°C for 30 s for *GSTM1/GSTT1*, followed by 72°C for 15 s, and then an extension of 72°C for 5 min. A multiplex PCR method was used for the *GSTM1* and *GSTT1* deletions, and the genotypes were ascertained on a 1.2% agarose gel by examining the gels for bands of the appropriate sizes, as described by Abdel-Rahman *et al*. (30). Positive and negative controls were included and 5% of the samples were run in duplicate for each genotyping assay with 100% concordance.

Statistical analysis

We estimated the genotype frequencies for each of the candidate metabolic gene polymorphisms and tested them for Hardy-Weinberg equilibrium. We used the Kaplan-Meier product limit method to assess the probability that subjects with a particular genotype would remain free of CRC. Thus, we defined CRC as the failure event; data for all other

subjects were censored on the date of last contact for those remaining free of CRC, on the date of diagnosis of adenoma or cancer other than CRC, or on the date of death due to other causes. Total analysis time at risk for all participants consisted of the period from the date of birth to the time of the first CRC event for cases, or the date of censoring for the rest. The disease-free survival curves by genotype were compared using the log-rank (LR) test. The median age at onset was defined as the age at which 50% of the participants remained cancer-free. We performed Cox proportional hazards regression analysis to estimate hazard ratios (HR) and 95% confidence intervals (95% CI) for risk of CRC at any age, comparing the polymorphic to the wild-type genotype. Gender, ethnicity (white or other), and MMR gene mutated (*MLH1* or *MSH2*) were included in the Cox model as potential confounding factors. We obtained unadjusted and adjusted HRs for the main effect of each of the candidate metabolic genes. We further performed stratified analysis (using Kaplan-Meier plots and Cox regression) for the SNPs that had a significant main effect by levels of the other variables, such as gender, ethnicity, gene mutated, to examine whether the effect of the SNP varied within these variables. We also generated interaction terms for SNPs significantly associated with CRC risk with each of the other gene polymorphisms (dichotomized as homozygous wild-type = 0; heterozygous and homozygous variant = 1; or non-deleted = 0 and deleted = 1 for *GSTM1* and *GSTT1*) and tested for multiplicative interaction in the Cox model by including each of the main effect terms and the term for interaction. We used the Huber-White robust variance correction as applied in STATA 8.0 (StataCorp LP, College Station, TX) to correct for any correlations in time to onset of CRC among family members (31;32). The robust variance estimator adjusts for within-cluster correlation (data not independent within groups, but independent across groups) appropriately correcting for within-family correlation of the age of onset. We tested whether the hazard ratios were proportional using the method described by Grambsch *et al*. (33).

Results

Of the 257 participants, 120 (46.7%) developed CRC as the first cancer. We found no differences in the CRC-free survival time or the median age at onset of CRC by gender, ethnicity, or MMR gene mutated (LR test $P > 0.05$ for each of the variables; Table 1). All of the SNPs analyzed were in Hardy-Weinberg equilibrium (exact $P > 0.05$; Table 2). The allele frequencies were close to those described in the literature (we compared allele frequencies in our patients with those reported for white populations since our study participants were predominantly non-Hispanic whites). We compared the genotypic frequencies for each of the gene SNPs between subjects with and without CRC, using Pearson's χ^2 tests and found no difference in the frequencies of *EPHX1* and GST genotypes (data not shown). For *CYP1A1*, the genotype frequencies for both the Msp1 ($\chi^2_{(2)}$ 6.23; $P = 0.04$) and I462V ($\chi^2_{(2)}$ 7.63; $P = 0.02$) SNPs were significantly different between the patients with and without CRC (however, the endpoint of our study was the time to onset of CRC by genotype).

Kaplan-Meier survival curves, comparing CRC-free survival by genotype, were significantly different for the *CYP1A1* I462V SNP (comparing AA, AG, and GG, LR test $P = 0.018$; for AG and GG genotypes combined, LR test $P = 0.036$) and marginally so for the *CYP1A1* T>C SNP (comparing genotypes TT, TC, and CC, LR test $P = 0.059$) (Figure 1). CRC-free survival did not differ by genotype for any of the other polymorphisms or gene deletions ($P > 0.05$). The median age at onset for CRC was 39 years for patients with the *CYP1A1* I462V AG genotype compared to 43 years for those with the wild-type AA genotype. Using Cox proportional hazards regression analysis, we also found an increased risk of early onset CRC associated with the AG genotype compared to the AA genotype (HR = 1.81, 95% CI = 1.19–2.75, $P = 0.005$). Because of the rarity of the *CYP1A1* homozygous variant genotypes for both Msp1 and I462V, we did not analyze them separately (risk estimates were unstable

and have been omitted) but did combine them with the heterozygous genotypes to calculate the risk estimates (Table 2). On stratified analysis for the effects of various *CYP1A1* genotypes by gender and gene mutated on risk of early disease, we found that female gender and presence of the *MSH2* mutation were significantly associated with the risk of early onset CRC (Table 3). Although the direction of association was the same among males and females, and among *MLH1* and *MSH2* mutation carriers, the magnitude of effect was smaller among males and *MLH1* mutation carriers and the results did not achieve statistical significance, perhaps due to smaller numbers in those categories. Furthermore, since our study population was predominantly white, we analyzed the non-Hispanic whites versus others and found that a significant association persisted with an increased risk related to the AG genotype compared to the AA (HR = 1.73, 95% CI = 1.08–2.77, $P = 0.021$) among the non-Hispanic whites (Table 3). Even though gender, ethnicity, or gene mutated did not have significant effects at the $\alpha = 0.05$ level in the Cox model, we retained them as covariates as per convention for their role as potential confounders of the modifier gene-CRC association. The crude and adjusted hazard ratios are listed in Table 2. We tested for and found no evidence of multiplicative interaction by gender, ethnicity, or gene mutated in the Cox model (Wald χ^2 test $P > 0.05$). All the estimates were obtained by applying robust correction using the cluster function in STATA.

In analyzing the *CYP1A1* Msp1 SNP, which was in significant LD with the *CYP1A1* I462V SNP (D' = 0.97; $R^2 = 0.44$; $P < 0.0001$), we found that the TC genotype was associated with an increased hazard for earlier onset of CRC compared to the TT genotype (adjusted HR = 1.62, 95% CI = 1.06–2.45, $P = 0.02$) and that this was more evident among *MSH2* mutation carriers, as in the case of the *CYP1A1* I462V. The proportional hazards assumption was tested using the Schoenfeld residuals and was not violated for either of the *CYP1A1* SNPs ($P > 0.05$ for the global test).

None of the other SNPs analyzed was associated with an earlier age at onset or a difference in risk by genotype. However, we found evidence for multiplicative interaction between *CYP1A1* I462V and *EPHX1* Y113H (Wald χ^2 $P = 0.036$; likelihood ratio test $P = 0.044$) with a greater than multiplicative hazard ratio for the combined effect of having a variant allele of both these SNPs (HR = 3.09, 95% CI = 1.58–6.04, $P = 0.001$). Further, on stratified analysis (Figure 2) we found that in the presence of any polymorphic allele of the *EPHX1* Y113H SNP (genotypes TC and CC), having any polymorphic allele of *CYP1A1* I462V (genotypes AG and GG) was associated with a significantly earlier median age of CRC onset at 37 years compared to 42 years (LR test $P = 0.002$) for having the *CYP1A1* I462V homozygous wild-type allele (AA genotype) . We did not detect interaction of *CYP1A1* I462V SNP with any of the other polymorphisms (P for interaction term > 0.05).

Discussion

This study investigated eight polymorphisms in five candidate genes involved in the metabolism of xenobiotics to determine if they had any effect on the age at onset of CRC as an indicator of phenotypic variation in individuals with Lynch syndrome. The modifying effect of these genes was examined among individuals with a common background for increased susceptibility to CRC due to an inherited deleterious mutation resulting in deficient mismatch repair. Our most prominent finding was an observed shift in the median age at onset of CRC to 4 years earlier among *CYP1A1* I462V heterozygotes compared to those with the homozygous wild-type genotype. In addition, subjects with the heterozygous genotype had an ~80% higher risk for CRC each year than subjects with the homozygous wild-type genotype. A similar trend in risk was observed for the *CYP1A1* Msp1 variant, although the risk estimates were lower. Since the CYP group of enzymes is predominantly involved in activation reactions and the *CYP1A1* I462V polymorphism results in higher

enzyme activity, it is reasonable to assume that the increased CYP1A1 activity may be associated with an increase in the level of activated metabolites that have the potential to cause DNA damage and initiate carcinogenesis.

Comparing our results to evidence in the literature, we found that in a recent study, Talseth *et al*. (11) examined four of the same polymorphisms we did—*CYP1A1* Msp1, *GSTP1* I105V, and *GSTM1* and *GSTT1* deletions—in a mixed Australian (n = 86) and Polish (n = 134) population of MMR gene mutation carriers. The researchers did not find a statistically significant difference between the age at diagnosis of CRC by genotype for any of the genes tested, although they did detect a nonsignificant difference of up to 8 years in the median age at onset of CRC for three of the SNPs (authors did not specify which SNPs). Talseth *et al*. also found that only the *CYP1A1* Msp1 mutant genotype was present significantly more in the CRC-affected subjects than in those without CRC (P = 0.03). This finding was confirmed by our results which showed a significant excess of Msp1 heterozygotes among the CRC affected (P = 0.04). Similarly, for the I462V SNP there were significantly more subjects with the heterozygous genotype among the CRC cases (P = 0.02) than among the cancer free. However, Talseth *et al*. did not examine the *CYP1A1* I462V, which is the SNP that we also found to be most strongly associated with an earlier onset of CRC.

The *GSTM1* and *GSTT1* deletion polymorphisms have been extensively examined in individuals with Lynch syndrome (18-21). In a Finnish population of 150 *MLH1* mutation carriers, the null alleles of *GSTM1* and *GSTT1* each shifted the median age of onset of CRC to 6 years earlier; the authors, however, did not report whether these differences were statistically significant (19). Another recent study examining a homogeneous cohort of 129 South African individuals carrying a single predisposing mutation in *MLH1* reported that men who were carriers of both the *GSTM1* and *GSTT1* null alleles had a 3-fold increase in risk of developing CRC compared to men who had neither null allele (18). There are likely to be inherent genetic differences and differences in environment and lifestyle factors between the two study populations that could explain why we did not find similar results. Nevertheless, our null results for *GSTM1* and *GSTT1* do corroborate the findings of two previous studies on MMR gene mutation carriers, neither one of which found an association between *GSTM1* and *GSTT1* deletions and early onset of CRC (18;19). Of these, the study by Jones *et al*. was on a smaller subset of our population (i.e. 104 MMR gene mutation carriers from 59 families) and a similar analysis detected no association of *GSTM1* with age at onset of colorectal cancer. (20). Though *GSTM1* deletion is common, being present in almost 50% of the general population, the enzyme is not expressed in the colonic mucosal cells (34); hence it may play a very limited role in the metabolism of xenobiotics in the colon. This could perhaps explain why most studies have not found an association of *GSTM1* with risk for sporadic or hereditary cases of CRC. The putative influence of *GSTT1*, however, is biologically plausible because this gene is abundantly expressed in the colon (34) and variation in GSTT1 enzyme levels could influence xenobiotic detoxification in the colon. Although our study did not find a modification of risk for CRC by *GSTT1* in individuals with Lynch syndrome, compiled evidence appears to indicate that those with the gene deletion, that is, those lacking the GSTT1 enzyme, are likely to be at an increased risk for sporadic CRC (9;35;36). Our study could have detected a hazards ratio of ~1.6 for this polymorphism with 80% power at a significance level of 5%, but a larger sample size would be needed to detect less penetrant effects of this (or other) variants on risk.

One of the limitations of our study was a lack of complete data on different sources of exposure to PAHs, which include smoking and consumption of well-done meat. It is likely that the influence of the genetic polymorphisms is dependent on the level of xenobiotic exposure (i.e., the risk associated with the metabolic genes could be further modified by the level of the substrates upon which they act). As demonstrated by Slattery *et al*. (10), the

impact of smoking on sporadic CRC risk is modified by the *CYP1A1* genotype. Further, we had limited power to test the joint effects of the various genes or gene-gene interaction in our Cox model. Although we found that the adverse effect of the *CYP1A1* I462V variant allele was further increased among individuals with the *EPHX1* Y113H variant allele, showing a 5-year earlier shift in median age at onset and a 3-fold increased risk by year compared to having the *CYP1A1* homozygous wild-type allele, these results are based on small numbers and would require validation in larger studies. It is pertinent to note that the interaction effect seen between EPHX1 and CYP1A1 is biologically plausible since both these enzymes are sequentially involved in the biotransformation of PAHs. In the initial phase I reaction, CYP1A1 converts benzo(a)pyrene, to the active benzo(a)pyrene 7,8 epoxide. This is hydrated by EPHX1 to a transhydrodiol derivative benzo(a)pyrene 7,8 diol that is less toxic (37), but the diol derivative is a primary substrate for CYP enzymes that further oxidize it to the highly reactive benzo(a)pyrene 7,8 dihydrodiol 9,10 epoxide (BPDE) – a potent mutagen capable of reacting with DNA to form adducts. Thus these and other genes may interact to play a more complex role in carcinogenesis. Future studies on larger numbers of mutation carriers using statistical tools such as Random Forest and CART analysis may be useful in determining the presence of gene-gene interactions and allow risk predictions to be made based on a combination of genotypes.

It is important to note that although we found an earlier age at CRC onset associated with two SNPs in the *CYP1A1* gene that are in LD, it is possible that neither one of these is the causative SNP. The increased risk may be due to genetic variation at another locus that happens to be in further LD with the *CYP1A1* SNPs examined in our study. Within *CYP1A1*, a reduced risk for sporadic CRC has been described for two other *CYP1A1* SNPs, T461N and −1738A>C (13). The T461N SNP has a very low minor allele frequency (< 5%) and the functional relevance of the −1738A>C SNP is still unknown, but it would be of interest to examine these two *CYP1A1* SNPs in future studies on individuals with Lynch syndrome.

Although we had limited power to examine the association of the SNPs with age at onset by different ethnic subgroups, our results were consistent whether we analyzed the overall population or restricted the analysis to non-Hispanic whites alone. Further, although we examined genetic variation at eight different loci, we do not believe that our results could be due to type I error as a consequence of multiple testing since the polymorphisms examined were in candidate metabolic genes that were selected a priori based on biological or epidemiological evidence for their role in affecting CRC risk. Besides, the univariate *P* value for *CYP1A1* I462V (*P*=0.005) meets the stringent threshold *P*- value of <0.00625 (0.05/n=8 tests) on applying the Bonferroni correction for multiple testing.

There is a possibility that some degree of selection bias may have influenced the results of our study because hospital identified probands are likely to be younger than Lynch syndrome cases identified from the general population. This may be mitigated to some extent in our study by the inclusion of family members with Lynch syndrome, both CRC-affected and non-affected, to the analysis which perhaps restores the age at onset distribution closer to that in the population. Additionally, since the participants were not enrolled on the basis of their genotype, selection bias would not falsely indicate an association even though there may be an earlier shift in the age at onset.

Finally, the results of our study indicate that among MMR gene mutation carriers, the age at onset of CRC may be modified by the *CYP1A1* I462V and Msp1 polymorphisms; the I462V polymorphism increasing risk for an earlier onset age by 4 years on average. Therefore, combining the knowledge of an individual's *CYP1A1* genotype along with other genetic markers and environmental factors that influence CRC risk in MMR gene mutation carriers may improve risk estimates and help identify genetically susceptible high-risk subgroups

with Lynch syndrome that could benefit most from intensive screening and chemopreventive or surgical interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Haidee Chancoco and Domitila Patenia for their technical assistance.

Grant support: National Cancer Institute grant CA 70759, National Institutes of Health Cancer Center Support grant CA 16672; and National Cancer Institute grant CA 57730 (R25 Cancer Prevention Predoctoral Fellowship).

Reference List

1. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. N Engl J Med 2003;348:919–32. [PubMed: 12621137]
2. Watson P, Riley B. The tumor spectrum in the Lynch syndrome. Fam Cancer 2005;4:245–8. [PubMed: 16136385]
3. Peltomaki P, Gao X, Mecklin JP. Genotype and phenotype in hereditary nonpolyposis colon cancer: a study of families with different vs. shared predisposing mutations. Fam Cancer 2001;1:9–15. [PubMed: 14574010]
4. de la Chapelle A. Genetic predisposition to colorectal cancer. Nat Rev Cancer 2004;4:769–80. [PubMed: 15510158]
5. Lynch HT, Lynch JF. Lynch syndrome: history and current status. Dis Markers 2004;20:181–98. [PubMed: 15528784]
6. Nebert DW, Dalton TP. The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. Nat Rev Cancer 2006;6:947–60. [PubMed: 17128211]
7. Smith G, Stanley LA, Sim E, Strange RC, Wolf CR. Metabolic polymorphisms and cancer susceptibility. Cancer Surv 1995;25:27–65. [PubMed: 8718512]
8. Mercurio MG, Shiff SJ, Galbraith RA, Sassa S. Expression of cytochrome P450 mRNAs in the colon and the rectum in normal human subjects. Biochem Biophys Res Commun 1995;210:350–5. [PubMed: 7755610]
9. de Jong MM, Nolte IM, te Meerman GJ, et al. Low-penetrance genes and their involvement in colorectal cancer susceptibility. Cancer Epidemiol Biomarkers Prev 2002;11:1332–52. [PubMed: 12433710]
10. Slattery ML, Samowtiz W, Ma K, Murtaugh M, et al. CYP1A1, cigarette smoking, and colon and rectal cancer. Am J Epidemiol 2004;160:842–52. [PubMed: 15496536]
11. Talseth BA, Meldrum C, Suchy J, Kurzawski G, Lubinski J, Scott RJ. Genetic polymorphisms in xenobiotic clearance genes and their influence on disease expression in hereditary nonpolyposis colorectal cancer patients. Cancer Epidemiol Biomarkers Prev 2006;15:2307–10. [PubMed: 17119063]
12. Gsur A, Zidek T, Schnattinger K, et al. Association of microsomal epoxide hydrolase polymorphisms and lung cancer risk. Br J Cancer 2003;89:702–6. [PubMed: 12915882]
13. Landi S, Gemignani F, Moreno V, et al. A comprehensive analysis of phase I and phase II metabolism gene polymorphisms and risk of colorectal cancer. Pharmacogenet Genomics 2005;15:535–46. [PubMed: 16006997]
14. Robien K, Curtin K, Ulrich CM, et al. Microsomal epoxide hydrolase polymorphisms are not associated with colon cancer risk. Cancer Epidemiol Biomarkers Prev 2005;14:1350–2. [PubMed: 15894702]
15. Whalen R, Boyer TD. Human glutathione S-transferases. Semin Liver Dis 1998;18:345–58. [PubMed: 9875553]

16. Pemble S, Schroeder KR, Spencer SR, et al. Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. Biochem J 1994;300(Pt 1): 271–6. [PubMed: 8198545]
17. Rebbeck TR. Molecular epidemiology of the human glutathione S-transferase genotypes GSTM1 and GSTT1 in cancer susceptibility. Cancer Epidemiol Biomarkers Prev 1997;6:733–43. [PubMed: 9298582]
18. Felix R, Bodmer W, Fearnhead NS, van der ML, Goldberg P, Ramesar RS. GSTM1 and GSTT1 polymorphisms as modifiers of age at diagnosis of hereditary nonpolyposis colorectal cancer (HNPCC) in a homogeneous cohort of individuals carrying a single predisposing mutation. Mutat Res 2006;602:175–81. [PubMed: 17087981]
19. Moisio AL, Sistonen P, Mecklin JP, Jarvinen H, Peltomaki P. Genetic polymorphisms in carcinogen metabolism and their association to hereditary nonpolyposis colon cancer. Gastroenterology 1998;115:1387–94. [PubMed: 9834266]
20. Jones JS, Gu X, Campos IM, Lynch PM, Amos CI, Frazier ML. GSTM1 polymorphism does not affect hereditary nonpolyposis colorectal cancer age of onset. Cancer Epidemiol Biomarkers Prev 2004;13:676–8. [PubMed: 15066938]
21. Heinimann K, Scott RJ, Chappuis P, et al. N-acetyltransferase 2 influences cancer prevalence in hMLH1/hMSH2 mutation carriers. Cancer Res 1999;59:3038–40. [PubMed: 10397239]
22. Chen J, Sen S, Amos CI, et al. Association between Aurora-A kinase polymorphisms and age of onset of hereditary nonpolyposis colorectal cancer in a Caucasian population. Mol Carcinog 2007;46:249–56. [PubMed: 17219423]
23. Frazier ML, O'Donnell FT, Kong S, et al. Age-associated risk of cancer among individuals with N-acetyltransferase 2 (NAT2) mutations and mutations in DNA mismatch repair genes. Cancer Res 2001;61:1269–71. [PubMed: 11245417]
24. Rodriguez-Bigas MA, Boland CR, Hamilton SR, et al. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. J Natl Cancer Inst 1997;89:1758–62. [PubMed: 9392616]
25. Chung DC, Rustgi AK. The hereditary nonpolyposis colorectal cancer syndrome: genetics and clinical implications. Ann Intern Med 2003;138:560–70. [PubMed: 12667026]
26. Cravo M, Afonso AJ, Lage P, Albuquerque C, Maia L, Lacerda C, et al. Pathogenicity of missense and splice site mutations in hMSH2 and hMLH1 mismatch repair genes: implications for genetic testing. Gut 2002;50:405–12. [PubMed: 11839723]
27. Raevaara TE, Korhonen MK, Lohi H, et al. Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. Gastroenterology 2005;129:537–49. [PubMed: 16083711]
28. Barnetson RA, Tenesa A, Farrington SM, Nicholl ID, Cetnarskyj R, Porteous ME, et al. Identification and survival of carriers of mutations in DNA mismatch-repair genes in colon cancer. N Engl J Med 2006;354:2751–63. [PubMed: 16807412]
29. Lagerstedt RK, Liu T, Vandrovcova J, et al. Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics. J Natl Cancer Inst 2007;99:291–9. [PubMed: 17312306]
30. Abdel-Rahman SZ, el Zein RA, Anwar WA, Au WW. A multiplex PCR procedure for polymorphic analysis of GSTM1 and GSTT1 genes in population studies. Cancer Lett 1996;107:229–33. [PubMed: 8947518]
31. Rogers WH. Regression standard errors in clustered samples. Stata Technical Bulletin 1993;13:19–23.
32. Williams RL. A note on robust variance estimation for cluster-correlated data. Biometrics 2000;56:645–6. [PubMed: 10877330]
33. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994;81:515–26.
34. Nishimura M, Naito S. Tissue-specific mRNA expression profiles of human phase I metabolizing enzymes except for cytochrome P450 and phase II metabolizing enzymes. Drug Metab Pharmacokinet 2006;21:357–74. [PubMed: 17072089]
35. Chen K, Jiang QT, He HQ. Relationship between metabolic enzyme polymorphism and colorectal cancer. World J Gastroenterol 2005;11:331–5. [PubMed: 15637738]

36. Rajagopal R, Deakin M, Fawole AS, et al. Glutathione S-transferase T1 polymorphisms are associated with outcome in colorectal cancer. Carcinogenesis 2005;26:2157–63. [PubMed: 16051638]

37. Huan WY, Chatterjee N, Chanock S, et al. Microsomal epoxide hydrolase polymorphisms and risk for advanced colorectal adenoma. Cancer Epidemiol Biomarkers Prev 2005;14:152–7. [PubMed: 15668489]

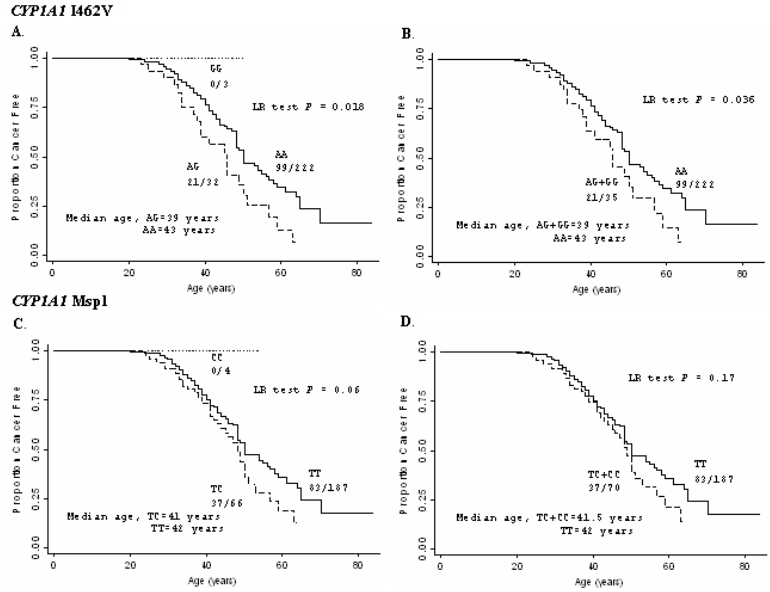

Fig. 1.

Kaplan-Meier plots comparing the age at onset of CRC according to *CYP1A1* I462V genotypes (**A**) AA, AG, and GG (**B**) AA and AG+GG; and *CYP1A1* Msp1 T>C genotypes (**C**) TT, TC, and CC (**D**) TT and TC+CC.

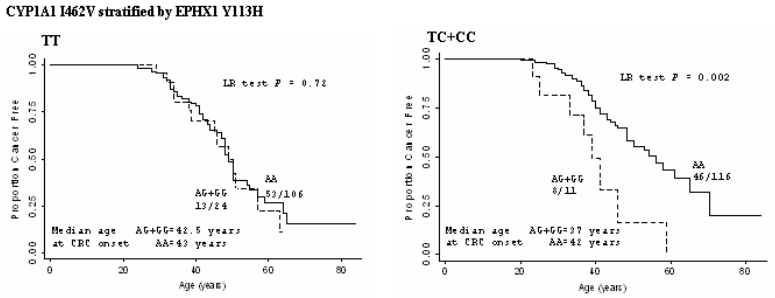

Fig. 2.
Kaplan-Meier plots for *CYP1A1* I462V SNP comparing genotype AA with AG+GG, stratified by *EPHX1* Y113H

Table I

Participant characteristics, frequency of CRC events, and comparison of disease-free survival

	***n* (%)**	**No. of CRC events (failures)**	
	***n* = 257**	**Total = 120**	**Log-rank test[*] *P***
Gender			
Male	112	60	0.19
Female	145	60	
Ethnicity			
White	215 (83.7)	100	0.9
Black	25 (9.7)	13	
Hispanic	16 (6.2)	7	
Asian	1 (0.4)	0	
Non-White (Black, Hispanic, and Asian combined)	42 (16.3)	20	0.83
MMR gene mutated			
MLH1	108 (42)	55	0.24
MSH2	149 (58)	65	

Age[†] (years)		**Mean (SD) Min/Max**	**Student's t-test *P***
CRC +	120 (46.7)	41.9 (10.4) 20/70	0.34
CRC −	137 (53.3)	43.3 (12.4) 18/84	

[*] Test for equality of the Kaplan-Meier estimates of CRC-free survival by participant characteristics.

[†] Age at diagnosis for the CRC-affected and age at censoring for the unaffected (censored at date of last contact or date of diagnosis of adenoma/other cancer/death).

Table 2

Genotype frequencies, comparison of disease-free survival by genotype, and Cox proportional hazards analysis for the age-associated risk of CRC by genotype.

	HWE Exact *P*	*n* (%)	CRC events	Log-rank test *P*	Univariate HR (95% CI)	*P*	Adjusted* HR (95% CI)	*P*
		Total *n* = 257	120					
CYP1A1 I462V								
AA	0.15	222 (86.4)	99	0.018	Ref	-	Ref	-
AG		32 (12.4)	21		**1.81 (1.19-2.75)**	**0.005**	**1.78 (1.16 - 2.74)**	0.008
GG†		3 (1.2)	0		-	-	-	-
Any G (AG+GG)		35 (13.6)	21	0.036	**1.63 (1.08 2.44)**	**0.02**	**1.62(1.06-2.45)**	**0.02**
CYP1A1 Msp1								
TT	0.62	187 (72.8)	83	0.06	Ref	-	Ref	-
TC		66 (25.7)	37		**1.41 (1.01-1.97)**	0.04	**1.53(1.06 - 2.22)**	**0.02**
CC†		4 (1.5)	0		-	-	-	-
Any C(TC+CC)		70 (27.2)	37	0.17	1.31 (0.93-1.84)	0.13	1.44 (0.99 - 2.09)	0.06
EPHX1 Y113H								
TT	0.65	130 (50.6)	66	0.46	Ref	-	Ref	-
TC		108 (42.0)	46		0.9 (0.63-1.27)	0.55	0.88 (0.62-1.26)	0.49
CC		19 (7.4)	8		0.64 (0.27-1.51)	0.31	0.62 (0.25-1.51)	0.29
Any C (TC+CC)		127 (49.4)	54	0.37	0.85 (0.61-1.19)	0.34	0.83 (0.59-1.17)	0.29
EPHX1 H139R								
AA	0.58	155 (60.3)	77	0.34	Ref	-	Ref	-
AG		92 (35.8)	37		0.77 (0.55-1.08)	0.13	0.74 (0.52-1.04)	0.09
GG		10 (3.9)	6		1.18 (0.40-3.49)	0.76	1.17 (0.4-3.4)	0.77
Any G (AG+GG)		102 (39.7)	43	0.26	0.81 (0.58-1.14)	0.22	0.78 (0.55-1.1)	0.16
GSTP1 I105V								
AA	1	120 (46.7)	55	0.75	Ref	-	Ref	-
AG		112 (43.6)	54		1.13 (0.81-1.57)	0.48	1.11 (0.79-1.54)	0.54
GG		25 (9.7)	11		0.93 (0.48-1.79)	0.83	0.85 (0.44-1.65)	0.64
Any G (AG+GG)		137 (53.3)	65	0.64	1.09 (0.79-1.5)	0.78	1.06 (0.76-1.46)	0.73
GSTP1 A114V								

	HWE Exact *P*	*n* (%)	CRC events	Log-rank test *P*	Univariate HR (95% CI)	*P*	Adjusted* HR (95% CI)	*P*
CC	1	227 (88.3)	105	0.49	Ref	-	Ref	-
CT		30 (11.7)	15		0.83 (0.5-1.34)	0.44	0.75 (0.44-1.28)	0.3
GSTM1								
+ (Gene present)	-	144 (56.0)	68	0.81	Ref	-	Ref	-
− (Null)		113 (44.0)	52		0.96 (0.66-1.38)	0.81	0.98 (0.67-1.43)	0.92
GSTT1								
+ (Gene present)	-	198 (77.0)	95	0.79	Ref	-	Ref	-
− (Null)		59 (23.0)	25		1.06 (0.75-1.49)	0.73	1.05 (0.74-1.49)	0.78

*Adjusted for gender, ethnicity (white or non-white), and gene mutated (*MLH1* or *MSH2*).

†Hazard ratios omitted as estimates are unreliable with small numbers. HWE = Hardy-Weinberg equilibrium.

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2011 January 27.

Table 3

Stratified analysis for *CYP1A1* I462V and Msp1 by gender, gene mutated, and ethnicity.

	HR (95% CI)	*P*	HR (95% CI)	*P*
	Male		**Female**	
CYP1A1 I462V				
AA	Ref		Ref	
AG	1.65 (0.88 - 3.10)	0.12	**1.99 (1.12 - 3.53)**	**0.019**
AG+GG	1.65 (0.88 - 3.10)		1.61 (0.91 - 2.86)	0.09
CYP1A1 Msp1				
TT	Ref		Ref	
TC	1.34 (0.75 - 2.4)	0.33	1.58 (0.99 - 2.5)	0.052
TC+CC	1.34 (0.74 - 2.41)	0.33	1.39 (0.87 – 2.24)	0.17
	MLH1		***MSH2***	
CYP1A1 I462V				
AA	Ref		Ref	
AG	1.3 (0.82 - 2.03)	0.24	**2.24 (1.22 - 4.12)**	**0.009**
AG+GG	1.15 (0.79 - 1.81)	0.52	**2.06 (1.14 - 3.72)**	**0.017**
CYP1A1 Msp1				
TT	Ref		Ref	
TC	1.3 (0.82 - 2.07)	0.27	**1.74 (1.12 - 2.69)**	**0.013**
TC+CC	1.19 (0.74 - 1.91)	0.46	**1.62 (1.04 - 2.54)**	**0.034**
	White		**Other**	
CYP1A1 I462V				
AA	Ref		Ref	
AG	**1.73 (1.09 - 2.77)**	**0.021**	2.19 (0.82 - 5.88)	0.12
AG+GG	1.53 (0.97 - 2.43)	0.066	2.11 (0.8 - 5.63)	0.13
CYP1A1 Msp1				
TT	Ref		Ref	
TC	1.5 (0.99 - 2.26)	0.052	1.04 (0.52 - 2.08)	0.91
TC+CC	1.41 (0.93 - 2.12)	0.1	0.93 (0.44 - 1.96)	0.85