Vitamin D Related Genes, CYP24A1 and CYP27B1, and Colon Cancer Risk

Linda M. Dong,1,2 Cornelia M. Ulrich,1,2 Li Hsu,1,2 David J. Duggan,3 Debbie S. Benitez,5 Emily White,1,2 Martha L. Slattery,4 Fred M. Farin,2 Karen W. Makar,1 Christopher S. Carlson,1,2 Bette J. Caan,5 John D. Potter,1,2 and Ulrike Peters1,2

1Fred Hutchinson Cancer Research Center and 2University of Washington, Seattle, Washington; 3Translational Genomics Research Institute, Phoenix, Arizona; 4University of Utah Health Sciences Center, Salt Lake City, Utah; and 5Division of Research, Kaiser Permanente Medical Care Program, Oakland, California

Abstract

Genetic association studies investigating the role of vitamin D in colon cancer have primarily focused on the vitamin D receptor (VDR), with limited data available for other genes in the vitamin D pathway, including vitamin D activating enzyme 1-α-hydroxylase (CYP27B1) and vitamin D deactivating enzyme 24-α-hydroxylase (CYP24A1). We evaluated whether 12 tagging single nucleotide polymorphisms (SNP) in CYP24A1, identified by resequencing the gene in 32 Caucasian samples, and 1 SNP in CYP27B1 were associated with colon cancer risk. In addition, we evaluated whether these two genes modify associations between colon cancer on the one hand and total vitamin D intake and UV-weighted sun exposure on the other, as well as other variants in VDR. Unconditional logistic regression was used to calculate odds ratios (OR) and 95% confidence intervals (95% CI) for the association between polymorphisms and haplotypes in CYP27B1 and CYP24A1 in a multicenter population-based case-control study of 1,600 cases and 1,949 controls. The CYP24A1 polymorphism IVS4-66T > G showed a statistically significant association with risk of colon cancer overall, particularly for proximal colon cancer. When stratified by anatomic site, we also found statistically significant associations for three CYP24A1 polymorphisms with risk of distal colon cancer (IVS4 + 1653C > T; OR for CC versus TT, 0.81; 95% CI, 0.68-0.96; IVS9 + 198T > C: OR for CC versus TT, 1.33; 95% CI, 1.03-1.73; and within whites only: +4125bp 3' of STPC > G: OR for GG versus CC, 1.44; 95% CI, 1.20-1.73). In addition, a possible interaction between CYP27B1 and UV-weighted sun exposure with proximal colon cancer was observed. As this is the first study to evaluate these genes in relation to colon cancer, additional studies are needed to confirm these results. (Cancer Epidemiol Biomarkers Prev 2009;18(9):2540–8)

Introduction

Colorectal cancer is the third most commonly diagnosed cancer among men and women in the United States (1). Despite screening and decreasing incidence rates over the past few decades, colorectal cancer is expected to account for 8% to 9% of all cancer deaths in 2008 (1). Ever since Garland et al. (2) hypothesized that lower risk of colon cancer was related to higher sunshine exposure and vitamin D, a number of experimental and epidemiologic studies investigating the potential chemopreventive effect of vitamin D have been conducted, most of which are consistent with an inverse association. Conclusions from recent review articles of observational studies strongly support the hypothesis that higher levels of vitamin D (both dietary intake and serum concentrations) reduce the risk of colorectal cancer (3, 4). Although results from the calcium and vitamin D supplementation trial of the Women’s Health Initiative (5) failed to show an effect on colorectal cancer risk, high vitamin D supplementation prior to enrollment and the relatively low vitamin D dose chosen for the intervention (400 IU of vitamin D3) complicate its interpretation (6). Further, a nested case-control study within the Women’s Health Initiative did observe a statistically significant inverse trend between baseline concentrations of serum 25-hydroxyvitamin D and risk of colorectal cancer (5).

Aside from its role in bone metabolism and regulation of blood calcium concentrations, vitamin D is postulated to reduce epithelial cell proliferation and to promote differentiation in various cell cultures, including colon-derived cells, and animal studies (7–9). In addition, vitamin D has been shown to induce cell-cycle arrest and apoptosis in colorectal tumor cell lines and premalignant adenoma cell lines (10). In the vitamin D pathway, the focus has been primarily on four genetic variants (Taq1, Bsm1, Fok1, and Apa1) in the vitamin D receptor (VDR) because the cellular effects of vitamin D are mediated primarily through binding, in the biologically active form of 1,25-dihydroxyvitamin D, to VDR, which regulates the transcription of numerous genes (11). However, the discovery that genes encoding for enzymes involved in the activation and inactivation of vitamin D are also expressed in colon cells (12, 13) suggests other genes that...
may contribute towards the association of vitamin D with colorectal cancer risk.

The presence of vitamin D hydroxylase enzymes in colon cells indicates a potential anticarcinogenic effect of 25-hydroxyvitamin D on colon tissue through local production of the biologically active form of vitamin D (1,25-dihydroxyvitamin D) from 25-hydroxyvitamin D (12). The enzyme 1α-hydroxylase, encoded by CYP27B1, is responsible for converting vitamin D into the active VDR-binding form. CYP24A1 encodes for the catabolic enzyme 24-hydroxylase and is responsible for inactivating vitamin D metabolites. Despite their important role in vitamin D metabolism, genetic variants in CYP24A1 and CYP27B1 have not previously been investigated in relation to colon cancer risk.

Detailed information about the genetic variation in CYP27B1 from resequencing has only recently become available,6 and such information is missing for CYP24A1. The purpose of this study was to take advantage of the recent genetic characterization of CYP27B1 and to resequence CYP24A1 for a comprehensive analysis of the relation between genetic variation in these two key vitamin-D-pathway genes and colon cancer risk in a large population-based case-control study of individuals from California, Minnesota, and Utah. We also examined whether the chemopreventive effect of vitamin D is modified by genetic variation in CYP24A1 and CYP27B1.

Materials and Methods

Study Population. This study was based on a multicenter population-based case-control study with 1,993 colon cancer cases and 2,410 controls recruited from the Kaiser Permanente Medical Care Program of Northern California, from an eight-county area in Utah, and from the metropolitan Twin Cities area of Minnesota. Cases were eligible if they were ages 30 to 79 y at the time of diagnosis, with first primary colon cancer (ICD-O Ed.2 codes 18.0 and 18.2-18.9) between October 1991 and September 1994, and with the mental competence to complete the interview. Cases with the following conditions were not eligible: tumors in the rectosigmoid junction or rectum, and cases with pathology report indicating familial adenomatous polyposis, Crohn's disease, or ulcerative colitis. A rapid-reporting system was used to identify all incident cases, with a majority being interviewed within 4 mo of diagnosis. Response and participation proportions have been described previously (14). Of the cases that were asked to participate in the study, 76% cooperated.

Controls were matched to cases by 5-year age groups and sex. Controls from the Kaiser Permanente Medical Care Program were randomly selected from membership lists. In Utah, controls who were <65 y of age were randomly selected from lists generated using random-digit dialing and driver’s license lists, and those ≥65 y were randomly selected from Health Care Financing Administration lists. In Minnesota, control participants were identified from driver's license or state identification lists. Of all controls asked to participate, 64% cooperated (14).

Resequencing and Tagging Single Nucleotide Polymorphism Selection. To identify relevant and novel polymorphisms in CYP24A1 (chr20:203,395-52,223,931; 20537 bp), we sequenced the promoter region (2 kb upstream), the complete coding sequence (12 exons, including intron/exon boundaries), as well as conserved intronic regions identified by phylogenetic footprinting analysis using the ECR Browser (15). We sequenced 32 of the International HapMap Project Centre d’Etude du Polyorphismme Humain (HapMap CEPH) individuals with European ancestry, given that our study population was predominantly Caucasian (95%) and to allow comparability with data from HapMap. Resequencing using standard dye-sequencing (Applied Biosystems) was conducted at the Center for Genomics and Environmental Health at the University of Washington. The subsequent fragments were then aligned and compared with each other and the downloaded GenBank reference sequence using Sequencher (GeneCodes) to determine genetic variation. All alignments were visually inspected to ensure that all variants were comprehensively identified.

Eighty-six variants were detected through our resequencing efforts, of which 30 (6 with minor allelic frequency ≥5%) had not been previously described in dbSNP (The Single Nucleotide Polymorphism database). These plus six additional variants located in regions that we did not resequence but were genotyped in HapMap (Data release, January 20, 2006) were included in the tagging SNP (tagSNP) selection for CYP24A1. For CYP27B1, we used existing data from the National Institute of Environmental Health Sciences (NIEHS) SNPs. Common SNPs with minor allele frequency ≥5% were selected using a three-step approach. To include potentially functionally relevant SNPs, we selected all nonsynonymous polymorphisms. Thereafter, we identified tagSNPs using the htsSNP program developed by Clayton to capture the common variation in the gene (minimum r2 of 0.80, minor allele frequency >5%). Finally, we examined how well the selected tagSNPs were able to identify common haplotypes (>5%) using the software Haploview leading to the inclusion of additional SNPs. A total of 16 SNPs for CYP24A1 and 3 SNPs for CYP27B1 were selected for genotyping.

Genotyping. Of the 4,403 cases and controls with valid study data, 3,680 (83% of cases and 85% of controls) provided a blood sample. Genomic DNA was extracted with a success rate of >95% for both cases and controls from peripheral blood lymphocytes or immortalized cell lines. Staff was blinded to case/control status. Each plate

7 http://ecrbrowser.dcode.org/
6 http://egp.gs.washington.edu/
8 http://www-gene.cimr.cam.ac.uk/clayton/software/stata
contented samples from both cases and controls, as well as duplicate quality-control samples interspersed among the plates (147 duplicates of samples from controls). All genotyping was done by matrix-assisted laser desorption/ionization—time-of-flight mass spectrometry on the Sequenom MassARRAY 7K platform using the iPLEX Gold (low-plex) reaction and conducted at the Translational Genomics Research Institute in March 2007. In total, we received successful genotyping results for 13 of 19 SNPs (CYP24A1, n = 12; CYP27B1, n = 1) for a total of 3,549 subjects (80% of cases and 81% of controls). SNPs were excluded for the following reasons: two SNPs (CYP27B1-1073C > G and IVS8 +113A > C) failed at final assay design and could not be replaced as no other tagSNP was available, one SNP (CYP24A1 EX6 +12C > T) had a call rate <95%, two SNPs (CYP24A1 EX1 +96G > A, IVS7 +204C > T) had <98% concordance rate with sequencing data, and one SNP (CYP24A1 IVS7-1457C > T) had positive calls in water-negative controls. The call rate was >95% for all 13 SNPs. Blinded duplicates displayed >99% concordance for any SNP. Using a goodness-of-fit test, the allele frequencies among Caucasians controls did not deviate from Hardy-Weinberg Equilibrium (P > 0.05) for any SNP. The 12 successfully genotyped tagSNPs covered the genetic variation of all CYP24A1 SNPs with an r² > 80% (mean r², 0.89) except for 7 SNPs that were unsuccessfully included in our assays. The tagSNPs had either low or moderate correlation with these 7 SNPs (r² range, 0.15-0.78).

Data on VDR variants (Bsm1 and Fok1) were previously collected and published on this study population and were made available for this analysis (16).

**Dietary and Lifestyle Data.** Detailed in-person interviews were used to collect demographic, dietary, and lifestyle data from all eligible study participants. Study participants were asked about their lifestyle during the year, 2 y prior to the date of diagnosis or selection. During the in-person interview, information was collected on dietary intake, physical activity, medical history and drug use, demographic factors, smoking, reproductive history (for women), and family history of cancer and colorectal polyps. Quality control methods were used to monitor the interviews (17). Dietary intake was ascertained using a modified version of the diet history questionnaire designed and validated for the Coronary Artery Risk Development in Young Adults study (18). Use of multivitamins, single vitamin, and mineral supplements was also ascertained. We defined supplemental vitamin D as 10 μg (400 IU)/d if a participant indicated either regular multivitamin or vitamin D supplement use (three times a week for ≥1 mo). Total vitamin D combines dietary and supplemental vitamin D intake. Calcium intake was restricted to dietary intake only because supplemental dosage was not provided and doses in multivitamins are often low and highly variable.

The estimates of UV-weighted hours of sun exposure (UV index-hours/week) were based on the average hours per week spent outdoors in the daylight as reported by subjects during each season (spring, summer, fall, winter) of the referent year, multiplied by the UV index for each season in the geographic area of the study center, and divided by four to average over the four seasons (19).

**Statistical Analyses.** Unconditional logistic regression, adjusted for age, sex, race, and study center, was used to estimate odds ratios (OR) and 95% confidence intervals (95% CI) for the association between individual polymorphisms and colon cancer. Trends were tested by including a variable coded 0, 1, and 2 for the number of rare alleles. Statistical significance was defined as P ≤ 0.05. We evaluated associations for the entire study population (all ethnicities), as well as restricted to non-Hispanic whites. Dietary factors included as potential confounders or effect modifiers were energy-adjusted using the residual method. Additional adjustment for education, income, body mass index, cigarette smoking, physical activity, long-term alcohol use, non-steroidal anti-inflammatory drug (NSAID) use, family history of colorectal cancer, dietary fiber, folate, red meat, fat intake, and multivitamin use did not result in meaningful changes of risk estimates for genotypes and were not included in the reported analyses. Global tests of association were conducted by simultaneously including genotypes (heterozygotes and homozygotes rare allele) of all CYP24A1 SNPs in a model and comparing that model to one that included none of the genotypes. Adjustment for multiple testing was achieved through this multilocus global test (degrees of freedom = 2 × number of SNPs in a gene) within the likelihood ratio χ² (20). When SNPs were highly correlated (r² > 0.80), only one of the SNPs was included in the gene-based model.

Haplotype frequencies were estimated and associations evaluated using HaploStats (version 1.3.1) in R (version 2.4.1), assuming an additive model. If the true underlying model is recessive or dominant, associations for haplotypes may not be apparent due to misspecification. Analyses for haplotypes were restricted to non-Hispanic whites, the largest racial group in our study population, and adjusted for age, sex, and study center. A global score statistic was used to evaluate the overall difference in haplotype frequencies between cases and controls.

Polytomous regression was used to estimate associations between genotypes and risk of proximal colon cancer (cecum, ascending colon, hepatic flexure, and transverse colon) and distal colon cancer (splenic flexure, descending colon, and sigmoid colon).

To evaluate whether genetic variants modified the association between measures of vitamin D and colon cancer risk and whether there was evidence of gene-gene interactions among CYP24A1, CYP27B1, and VDR, we investigated interactions through the inclusion of cross-product terms in the regression models. We conducted an omnibus test for multiplicative interaction between each gene and the variable of interest (e.g. total vitamin D) by simultaneously including all cross-product terms for the gene of interest (coded as dummy variables for heterozygotes and variant homozygotes) with the variable of interest in a model (coded as continuous) and comparing that with a model that included only the main effects for the genotypes and variable of interest. For omnibus tests of interaction between two genes, genotypes were reduced to a binary variable combining heterozygotes and variant homozygotes. If the omnibus test for interaction was statistically significant, we further investigated multiplicative interaction between individual SNPs and the variable of interest using the log likelihood ratio test to compare the fit of logistic models with and without interaction variables. Adjustment for multiple testing was also achieved through this multilocus global test.
Several publicly available web-based tools were used to assess the potential functional significance of variants, including the UCSC Genome Browser to obtain conservation scores (21), SIFT (22), PolyPhen (23), Splice Site Prediction tool by Neural Network (24), ESEfinder (25), RESCUE-ESE (26), and UTRScan (27).

Results

The majority of the study population were non-Hispanic white (Table 1). Controls had a higher proportion of college graduates, and higher-income study participants. Use of multivitamin and vitamin D supplements was higher in controls. Nonsignificant inverse associations for the highest quartiles of total vitamin D (proximal OR, 0.83; 95% CI, 0.66-1.05; distal OR, 0.80; 95% CI, 0.63-1.01) and UV-weighted sun exposure (proximal OR, 0.71; 95% CI, 0.55-0.92; distal OR, 0.84; 95% CI, 0.65-1.09) were observed. These results have been previously published (28).

When we investigated the polymorphisms in CYP24A1 individually, two CYP24A1 polymorphisms displayed a statistically significant association with risk of colon cancer. A statistically significant trend was observed with increasing number of rare alleles at loci IVS2 + 523C > T (P trend = 0.05; Table 2) with a statistically nonsignificant inverse association between this variant and colon cancer risk (all ethnicities OR CC + TC versus TT, 0.83; 95% CI, 0.66-1.03; non-Hispanic whites OR, 0.82; 95% CI, 0.65-1.02). A statistically significant inverse association between colon cancer risk and carrying at least one copy of the rare allele at IVS4+667 > G (all ethnicities OR GG + GT versus TT, 0.83; 95% CI, 0.72-0.96) was also observed. The global association P value for SNPs in CYP24A1 and colon cancer was 0.10 (df, 20). A haplotype analysis of CYP24A1 variants (IVS7 + 921C > T, IVS7 + 1307A > T, IVS9 + 198T > C) did not reveal any statistically significant associations with colon cancer risk (P global = 0.27; Table 3).

As the association with colon cancer differs for several risk factors by anatomic site, we examined CYP24A1 associations for distal and proximal colon cancer. Four CYP24A1 variants were statistically significantly associated with site-specific colon cancer, although the global tests were not statistically significant (proximal P global = 0.24; df, 20; distal P global = 0.13; df, 20). The CYP24A1 variant IVS4+667 > G, which showed a statistically significant inverse association with risk of colon cancer, was similarly associated with risk of proximal colon cancer (OR GG + GT versus TT, 0.82; 95% CI, 0.68-0.99) and was suggestive of an association with distal colon cancer (OR GG + GT versus TT, 0.85; 95% CI, 0.70-1.02; Table 2). Furthermore, we observed statistically significant associations for three other CYP24A1 variants with distal colon cancer. Those carrying at least one copy of the rare allele at IVS4 + 1653C > T had a 19% lower risk (OR, 0.81; 95% CI, 0.68-0.96), which was similar among non-Hispanic whites (OR TT + CT versus CC, 0.79; 95% CI, 0.66-0.94). In addition, two CYP24A1 variants had statistically nonsignificant trends, but genotypes with positive associations with distal colon cancer risk: IVS9 + 198T > C was associated with an increased risk of distal colon cancer (all ethnicities OR CC versus TT, 1.33; 95% CI, 1.03-1.73; and non-Hispanic whites OR, 1.46; 95% CI, 1.11-1.91); a borderline association limited to non-Hispanic whites between variant +4125bp 3' of STP C > G and distal colon cancer (OR GG versus CC, 1.44; 95% CI, 1.00-2.05) was also observed. Among the four CYP24A1 variants with a statistically significant association, the variant IVS4+667 > G was in strong linkage disequilibrium (D' = 1.0) with IVS4 + 1653C > T but in weak linkage disequilibrium (D' = 0.39-0.47) with the other two CYP24A1 variants. It was only weakly correlated (r^2 = 0.1) with the other three variants. The variant IVS4 + 1653C > T was in weak linkage disequilibrium (D' = 0.03-0.37) and weakly correlated (r^2 = 0.01-0.03) with IVS9 + 198T > C + +4125bp 3' of STP C > G, respectively. The variant +4125bp 3' of STP C > G was in moderate linkage disequilibrium (D' = 0.79) and moderately correlated (r^2 = 0.40) with IVS9 + 198T > C. Adjustments for the other CYP24A1 genotypes did not change any of the observed associations substantially, suggesting that the remaining four SNPs are independently associated with distal colon cancer risk.

Among SNPs in CYP24A1, three variants (IVS7 + 921C > T, IVS7 + 1307A > T, and IVS9 + 198T > C) were located within a haplotype block spanning intron 7 through intron 9 of CYP24A1. The haplotype containing the rare allele for IVS9 + 198T > C was possibly associated with risk of distal colon cancer (OR, 1.15; 95% CI, 0.99-1.33; Table 3). Although not statistically significant, this result is consistent with the positive association seen for IVS9 + 198T > C. The positive association found for the group composed of rare haplotypes is probably due to the majority of these haplotypes carrying the rare allele for IVS9 + 198T > C.

Table 1. Characteristics of cases and controls

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cases (n = 1,600)</th>
<th>Controls (n = 1,949)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)*</td>
<td>64.9 ± 9.8</td>
<td>65.0 ± 10.1</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>895 (55.9)</td>
<td>1,036 (53.2)</td>
</tr>
<tr>
<td>Female</td>
<td>705 (44.1)</td>
<td>913 (46.8)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White, not Hispanic</td>
<td>1,461 (91.4)</td>
<td>1,814 (93.1)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>62 (3.9)</td>
<td>78 (4.0)</td>
</tr>
<tr>
<td>African American</td>
<td>72 (4.5)</td>
<td>33 (1.7)</td>
</tr>
<tr>
<td>Other</td>
<td>4 (0.2)</td>
<td>3 (0.2)</td>
</tr>
<tr>
<td>Study center</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kaiser</td>
<td>769 (48.1)</td>
<td>804 (41.3)</td>
</tr>
<tr>
<td>Utah</td>
<td>578 (36.1)</td>
<td>765 (39.0)</td>
</tr>
<tr>
<td>Minnesota</td>
<td>253 (15.8)</td>
<td>249 (12.9)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;12 y</td>
<td>262 (16.4)</td>
<td>249 (12.8)</td>
</tr>
<tr>
<td>High school grade</td>
<td>451 (28.2)</td>
<td>546 (28.0)</td>
</tr>
<tr>
<td>Some college or post-high school</td>
<td>530 (33.1)</td>
<td>634 (32.6)</td>
</tr>
<tr>
<td>College graduate or higher</td>
<td>357 (22.3)</td>
<td>520 (26.7)</td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$20,000</td>
<td>401 (25.7)</td>
<td>455 (24.8)</td>
</tr>
<tr>
<td>$20,000-$40,000</td>
<td>533 (35.9)</td>
<td>627 (34.2)</td>
</tr>
<tr>
<td>$40,000-$60,000</td>
<td>348 (23.4)</td>
<td>438 (23.9)</td>
</tr>
<tr>
<td>&gt;$60,000</td>
<td>203 (13.7)</td>
<td>313 (17.1)</td>
</tr>
<tr>
<td>Mean UV-weighted hours of sun exposure (UV index B/wk)†</td>
<td>85.8 ± 20.9</td>
<td>87.4 ± 68.7</td>
</tr>
<tr>
<td>Mean total vitamin D (μg/d)†</td>
<td>10.6 ± 6.0</td>
<td>11.1 ± 6.2</td>
</tr>
<tr>
<td>Multivitamin supplement use‡</td>
<td>520 (32.5)</td>
<td>654 (33.6)</td>
</tr>
<tr>
<td>Vitamin D supplement use‡</td>
<td>25 (1.6)</td>
<td>53 (2.7)</td>
</tr>
<tr>
<td>Tumor site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distal</td>
<td>790 (49.4)</td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td>771 (48.2)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>39 (2.4)</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Continuous variables are displayed as mean values ± SD, and frequencies are displayed as counts (percentage).

*Defined as age at diagnosis for cases and age at recruitment for controls.

†Energy adjusted (residual method) dietary intake.

‡Supplement use defined as regular use (at least three times a week for at least one month) over the referent period.
Table 2. Odds ratios and 95% confidence intervals for the association between polymorphisms in CYP24A1 and CYP27B1 and colon cancer risk

<table>
<thead>
<tr>
<th>Position/Genotype</th>
<th>MAF (%)†</th>
<th>Controls (n)</th>
<th>All colon cancer</th>
<th>Proximal colon cancer</th>
<th>Distal colon cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cases (n)</td>
<td>OR† (95% CI)</td>
<td>P for trend</td>
<td>Cases (n)</td>
</tr>
<tr>
<td>CYP24A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IUV1-105G &gt; C (rs2248137)</td>
<td>0.401</td>
<td>667</td>
<td>546</td>
<td>1.00</td>
<td>269</td>
</tr>
<tr>
<td>IUV2-32C &gt; T (rs2762942)</td>
<td>0.062</td>
<td>1,712</td>
<td>1,430</td>
<td>1.00</td>
<td>691</td>
</tr>
<tr>
<td>CYP27B1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IUV1-105A &gt; G (rs2259735)</td>
<td>0.416</td>
<td>662</td>
<td>510</td>
<td>1.00</td>
<td>249</td>
</tr>
<tr>
<td>IUV4-66T &gt; G (rs4809958)</td>
<td>0.174</td>
<td>1,314</td>
<td>1,144</td>
<td>1.00</td>
<td>555</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*MAF (minor allelic frequency) calculated from results from genotyping restricted to non-Hispanic White controls.
†Adjusted for age, race, sex and study center.
Table 3. Association between CYP24A1 haplotypes and colon cancer among non-Hispanic whites

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Controls %</th>
<th>All colon cancer</th>
<th>Proximal colon cancer</th>
<th>Distal colon cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases %</td>
<td>OR* (95% CI) P</td>
<td>Cases %</td>
<td>OR* (95% CI) P</td>
</tr>
<tr>
<td>Block 1 (IVS7 + 921C &gt; T, IVS7 + 1307A &gt; T, IVS9 + 1981T &gt; C)^1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-A-T</td>
<td>39.4</td>
<td>38.0</td>
<td>1.00</td>
<td>39.2</td>
</tr>
<tr>
<td>T-A-C</td>
<td>31.1</td>
<td>32.3</td>
<td>1.07 (0.95-1.20)</td>
<td>0.29</td>
</tr>
<tr>
<td>T-T-T</td>
<td>17.4</td>
<td>15.8</td>
<td>0.94 (1.01-1.08)</td>
<td>0.39</td>
</tr>
<tr>
<td>T-A-T</td>
<td>8.6</td>
<td>9.4</td>
<td>1.14 (0.95-1.38)</td>
<td>0.16</td>
</tr>
<tr>
<td>rare</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td>0.27</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^1 Among non-Hispanic whites adjusted for age, sex and study center.

**Discussion**

We detected a statistically significant inverse association between CYP24A1 variant IVS4-667T > G and colon cancer overall. The association for IVS4-667T > G was also seen for proximal colon cancer. The most consistent pattern, however, was with risk of distal colon cancer. Specifically, CYP24A1 variants IVS4-667T > G and IVS4 + 1653C > T were associated with a reduced risk, and two CYP24A1 variants, IVS9 + 1981T > C and +4125bp 3' of STP C > G, were associated with increased risk of distal colon cancer. Our findings further suggest evidence of a possible interaction between CYP24A1 and CYP27B1 with distal colon cancer, and an interaction between CYP27B1 and UV-weighted hours of sun exposure with proximal colon cancer.

Based on observations that CYP27B1 is also expressed in the colon (12), it has been suggested that there may be a paracrine/autocrine role for 1,25-dihydroxyvitamin D, reducing cell proliferation and inducing differentiation locally (30). The involvement of CYP27B1 in colon carcinogenesis is further supported by the finding that its expression was much lower or even lost in high-grade, poorly differentiated colon cancer cells compared with normal cells (31). In contrast, the gene encoding for CYP24A1 has been observed to be more highly expressed in undifferentiated colon cancer cells than in normal colon cells (32), potentially preventing tissue accumulation of 1,25-dihydroxyvitamin D and its anticarcinogenic effects. Taken together, results from experimental studies support the importance of CYP27B1 and CYP24A1 in the role of vitamin D on colon cancer.

To our knowledge, this is the first study to examine CYP24A1 or CYP27B1 variants and colon cancer using a comprehensive approach. Of the statistically significantly associated CYP24A1 variants, IVS9 + 1981T > C, although not highly conserved (conservation score <0.1; ref. 21), may be particularly interesting as it could be involved in the regulation of splicing, given that it is located in the intron directly before an alternatively spliced exon. Exon 10 is not present in previously identified human mRNA transcripts listed in GenBank (33), which implies that exon 10 has a weak splicing site, but the functional impact of this transcript is yet to be determined. To investigate whether this intron variant could be linked with abnormal CYP24A1 splicing, we conducted a small experimental study evaluating CYP24A1 expression among EBV-lymphoblastoid cell lines with different genotypes of IVS9 + 1981T > C (see supplementary material). We were unable to show, however, that normalized CYP24A1 expression for all transcripts, or specifically the ratio of exon-10-containing transcripts to other transcripts, differed by genotype. However, a high degree of biological variability between individual cell lines diminished our ability to determine whether IVS9 + 1981T > C is...
involved in the regulation of CYP24A1 expression and will require additional experiments to be conducted.

Of the other independently associated CYP24A1 variants, +4125bp 3′ of STP C > G is located in the 3′ untranslated region, which could influence mRNA stability; but we were not able to detect any regulatory elements in the immediate region surrounding this variant (27). This tagSNP +4125bp 3′ of STP C > G, as well as IVS9 + 198T > C are correlated with synonymous SNP Ex8-33G > A (r2 = 0.51 and r2 = 0.62, respectively) which is highly conserved (conservation score = 0.49; ref. 21). Although it does not result in an amino acid change, its residue is located within the K α-helix of the protein, which may be near or within a substrate recognition site (34, 35). It has recently been proposed that synonymous mutations are not necessarily silent and could result in a different folding configuration and/or function of the protein (36). Taken altogether, there are two strong candidate SNPs (and/or function of the protein) (36) with evidence of linkage disequilibrium (r2 > 0.5) with other identified CYP24A1 variants. Thus, we could not determine what variant or factor, if any, is driving the observed association with IVS4 + 1653C > T.

In this study, CYP24A1 variant IVS4 + 1653C > T was associated with a decreased risk of distal colon cancer. As it resides within an intronic area that is not conserved and investigation of its possible function did not reveal any interesting leads, it is unlikely to be responsible for the observed association. However, this variant is not in strong linkage disequilibrium (r2 > 0.5) with other identified CYP24A1 variants. Thus, we could not determine what variant or factor, if any, is driving the observed association with IVS4 + 1653C > T.

As CYP24A1 and CYP27B1 are involved in the regulation of vitamin D, it is highly plausible that these two genes may modify an association between measures of vitamin D and colon cancer risk. Although there was no evidence of an interaction between CYP24A1 and either vitamin D measure, an inverse association between UV-weighted sun exposure and proximal colon cancer risk may depend on CYP27B1 IVS6-29T > C status: among those carrying the rare allele, increasing sun exposure was associated with lower cancer risk. This variant is located within a poorly conserved area and is in strong linkage disequilibrium with one other CYP27B1 variant, -1073C > G, located in the 5′ promoter region of the gene. The functional significance of both -1073C > G and IVS6-29T > C is unknown, but variants located in the 5′ region of a gene may have an impact on transcriptional or translational control (37). Although we detected a statistically significant interaction with CYP27B1, our global test was based on only one SNP for CYP27B1 (df = 2), and may be due to chance.

In this study, we observed an association between CYP24A1 variants and colon cancer, which seemed to be most consistent for distal colon cancer. There is extensive support from epidemiologic and experimental studies that risk factors have a different association with the development of colon tumors by anatomic site. The proximal and distal colon subsites have been characterized with both morphologic and genetic differences that would signify a potential difference in their etiology (38, 39). Of the epidemiologic studies that evaluated vitamin D by subsite, several suggested a stronger inverse association with distal tumors (40-42), whereas other studies did not find a meaningful difference by subsite (28, 43-46) or found an association with proximal only (47). However, the majority of studies that evaluated serum levels of 25-hydroxyvitamin D by subsite suggested a stronger inverse association with distal adenomas or cancer (48-52).

An explanation for why the associations with vitamin D and CYP24A1 may be specific to distal colon cancer is unclear. Further studies assessing the role of CYP24A1 variants are required to verify whether this association is truly specific to the distal colon.

The selection of tagSNPs based on resequencing data allowed us to conduct a more comprehensive analysis of common genetic variation in CYP24A1 than would have a candidate SNP approach. As CYP24A1 is a large gene (20 kb), we may have missed some genetic variation by not resequencing the entire CYP24A1 intronic region (77% sequenced). However, our resequencing strategy focused on regions of the gene most likely to be functionally relevant. We were not able to genotype all preselected tagSNPs due to incompatibility of some SNPs with the multiplexing assay, but the average coverage for CYP24A1 was high (mean r2, 0.89). TagSNP selection for CYP27B1 was also based on resequencing

---

**Table 4. Association between UV-weighted hours of sun exposure and proximal colon cancer risk stratified by CYP27B1 genotype**

<table>
<thead>
<tr>
<th>Position/Genotype</th>
<th>UV-weighted sun exposure*</th>
<th>∆ UV index-h/wk</th>
<th>OR† (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVS6-29T &gt; C (rs4064536)</td>
<td>TT</td>
<td>342/897</td>
<td>1.02 (0.89-1.18)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>325/783</td>
<td>0.84 (0.70-1.00)</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>80/198</td>
<td>0.94 (0.69-1.28)</td>
</tr>
<tr>
<td><strong>P</strong> interaction</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC + CT</td>
<td>0.86 (0.74-1.00)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*UV-weighted sun exposure based on the average hours per week of daylight reported, weighted by the UV index for each season in the geographic area of the study center.

†Adjusted for age, race, sex, and study center.

---

Cancer Epidemiol Biomarkers Prev 2009;18(9). September 2009
data (98% sequenced); however, there was very little genetic variation (nucleotide diversity among multietnic panel = 2.04 × 10⁻³), perhaps due to its important role in maintaining calcium homeostasis. This study included a large number of cases and collected detailed data on dietary and environmental factors, thus allowing us to control for known confounders. As this was a case-control study, dietary data were collected retrospectively and may be subject to recall bias; however, data were ascertained rapidly after enrollment and collected with established and validated procedures (18). A limitation is that vitamin D status was assessed by vitamin D intake and sun exposure. Serum 25-hydroxyvitamin D in prospectively collected samples provides a more direct measurement. Although some CYP24A1 results were nominally statistically significant, they should be interpreted with caution given that the global test for CYP24A1 was not statistically significant and there remains a potential role of chance in these findings. None of these SNPs reached genome-wide significance. Nonetheless, the probable functional significance of some SNPs provides us with the rationale to present results for individual SNPs and will make replication a worthwhile exercise.

The primary focus of our study was to evaluate the impact of common genetic variants in key genes in the vitamin D pathway other than VDR. Our results suggest that several variants in CYP24A1 are associated with colon cancer. Etiologic studies and vitamin D prevention strategies should take the genetic variation of CYP24A1 and the potential interaction between UV-weighted sun exposure and CYP27B1 into account.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References

Cancer Epidemiology, Biomarkers & Prevention 2009;18(9). September 2009
Cancer Epidemiology, Biomarkers & Prevention

Vitamin D Related Genes, CYP24A1 and CYP27B1, and Colon Cancer Risk

Linda M. Dong, Cornelia M. Ulrich, Li Hsu, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-09-0228

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2009/08/25/1055-9965.EPI-09-0228.DC1

Cited articles
This article cites 48 articles, 25 of which you can access for free at:
http://cebp.aacrjournals.org/content/18/9/2540.full.html#ref-list-1

Citing articles
This article has been cited by 3 HighWire-hosted articles. Access the articles at:
/content/18/9/2540.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.