Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

Research Article

Cancer Epidemiology, Biomarkers & Prevention

The Contribution of Common Genetic Variation to Nicotine and Cotinine Glucuronidation in Multiple Ethnic/Racial Populations

Yesha M. Patel[1], Daniel O. Stram[1], Lynne R. Wilkens[2], Sung-Shim L. Park[1], Brian E. Henderson[1], Loic Le Marchand[2], Christopher A. Haiman[1], and Sharon E. Murphy[3]

Abstract

Background: The lung cancer risk of smokers varies by race/ethnicity even after adjustment for smoking. Evaluating the role of genetics in nicotine metabolism is likely important in understanding these differences, as disparities in risk may be related to differences in nicotine dose and metabolism.

Methods: We conducted a genome-wide association study in search of common genetic variants that predict nicotine and cotinine glucuronidation in a sample of 2,239 smokers (437 European Americans, 364 African Americans, 453 Latinos, 674 Japanese Americans, and 311 Native Hawaiians) in the Multiethnic Cohort Study. Urinary concentration of nicotine and its metabolites were determined.

Results: Among 11,892,802 variants analyzed, 1,241 were strongly associated with cotinine glucuronidation, 490 of which were also associated with nicotine glucuronidation ($P < 5 \times 10^{-8}$). The vast majority were within chromosomal region 4q13, near *UGT2B10*. Fifteen independent and globally significant SNPs explained 33.2% of the variation in cotinine glucuronidation, ranging from 55% for African Americans to 19% for Japanese Americans. The strongest single SNP association was for rs115765562 ($P = 1.60 \times 10^{-155}$). This SNP is highly correlated with a *UGT2B10* splice site variant, rs116294140, which together with rs6175900 (Asp67Tyr) explains 24.3% of the variation. The top SNP for nicotine glucuronidation (rs116224959, $P = 2.56 \times 10^{-43}$) was in high LD ($r^2 = 0.99$) with rs115765562.

Conclusions: Genetic variation in *UGT2B10* contributes significantly to nicotine and cotinine glucuronidation but not to nicotine dose.

Impact: The contribution of genetic variation to nicotine and cotinine glucuronidation varies significantly by racial/ethnic group, but is unlikely to contribute directly to lung cancer risk. *Cancer Epidemiol Biomarkers Prev; 24(1); 119–27. ©2014 AACR.*

Introduction

Cigarette smoking is the leading cause of lung cancer-related deaths and nicotine is the agent responsible for tobacco addiction (1, 2). Much research has been directed toward understanding the pharmacology of nicotine and its influence on smoking behavior (3, 4). Smoking history, in the form of the number of cigarettes smoked per day (CPD), gathered through validated questionnaires, possibly in conjunction with plasma levels of nicotine metabolites, aid in evaluating tobacco smoke constituents uptake, individual differences in metabolism, and lung cancer risk (5–7). Surprisingly, notable racial/ethnic differences in lung cancer risk occur among smokers. Moreover these differences persist even after adjustment for smoking rates (i.e., cigarettes/day) and smoking duration (8, 9). For example, in comparison with European Americans, African American and Native Hawaiian smokers have higher overall risks of lung cancer at relatively low rates of consumption (e.g., 10 and 20 CPD), whereas Japanese Americans and Latinos tend to have lower risks than European Americans at this same level of smoking (8). These noted disparities in lung cancer risk among ethnic groups may be related to differences in internal dose and metabolism and may result from common genetic variation. Because nicotine is the known addictive component of cigarette smoke, understanding individual variation in nicotine metabolism is likely to be important in understanding both interindividual and racial/ethnic differences in smoking behavior, the resulting exposure to tobacco carcinogens, and lung cancer susceptibility (10, 11).

The primary pathway of nicotine metabolism is conversion to cotinine. Typically 80% of nicotine is metabolized to cotinine via cytochrome P450 2A6 (CYP2A6)-catalyzed C-oxidation (3, 12–14). CYP2A6 also catalyzes the oxidation of cotinine to *trans*-3′-hydroxycotinine (3-HCOT; ref. 15). The other pathways of nicotine metabolism, *N*-oxidation, and *N*-glucuronidation, each typically contribute <10% to total metabolism, although, in some individuals, *N*-glucuronidation may account for >40% of the excreted nicotine metabolites (16, 17). UGT2B10 and UGT1A4 both catalyze nicotine and cotinine *N*-glucuronidation; however, UGT2B10 is a significantly more efficient catalyst and seems to be the enzyme responsible for nicotine and cotinine glucuronidation in smokers

[1]Department of Preventive Medicine and Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, California. [2]Cancer Research Center of Hawaii, University of Hawaii, Honolulu, Hawaii. [3]Department of Biochemistry, Molecular Biology, and Biophysics and Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Sharon E. Murphy, Masonic Cancer Center, University of Minnesota, 2231 6th Street SE-2-127 CCRB, Minneapolis, MN 55455. Phone: 612-624-7633; Fax: 612-624-3869; E-mail: murph062@umn.edu

doi: 10.1158/1055-9965.EPI-14-0815

Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

(18–23). 3-HCOT is O-glucuronidated, a reaction catalyzed, at least in part, by UGT2B17 (3, 18, 22, 24), an enzyme that does not catalyze *N*-glucuronidation (25). In urine, the sum of nicotine, cotinine, 3-HCOT, and their respective glucuronide conjugates, referred to as "nicotine equivalents" account for 85% to 90% of total nicotine uptake (3). Therefore, nicotine equivalents can be used as a biomarker of nicotine uptake and tobacco exposure (26, 27).

There is noted interindividual variation in metabolism—different people metabolize nicotine and cotinine at different rates (28). Smokers self-modulate their tobacco consumption to maintain the desired effects and optimal concentrations of nicotine in the brain (2). A smoker with a slow rate of metabolism would likely smoke less or extract a lower nicotine dose per cigarette to achieve the same plasma level of nicotine as someone who metabolizes nicotine more quickly. Both CYP2A6 activity and genotype are associated with CPD in Japanese and smokers of European ancestry (4, 29–31). Nicotine glucuronidation is of interest as another possible modulator of smoking behavior, and we previously reported that smokers who carry the *UGT2B10* Asp67Tyr variant, which is associated with reduced nicotine and cotinine *N*-glucuronidation, excrete lower levels of nicotine equivalents (20, 22). Cotinine and nicotine glucuronidation levels, as represented in the urine, are significantly correlated and, due to the longer half-life of cotinine, cotinine glucuronide is a more stable phenotypic measure of variation in glucuronidation (21).

Our prior study was relatively small, analyzed a single variant, and was carried out in smokers with predominantly European American ancestry. The genome-wide association study (GWAS) study described here was carried out in a large multiethnic cohort, in which the urinary concentrations of nicotine and six metabolites were quantified. The significant variation in metabolism across the ethnic groups within this cohort was recently reported (32). As reported previously, nicotine C-oxidation was lower in Japanese Americans and Native Hawaiians compared with European Americans, whereas nicotine and cotinine N-glucuronidation were lower in African Americans (20, 29, 33). The large number of subjects and their varied nicotine metabolism in this cohort allowed us to comprehensively assess the relationship of nicotine glucuronidation to smoking intensity. Because the *N*-glucuronidation of cotinine and nicotine is catalyzed by the same enzymes (18, 34), we have used both nicotine and cotinine glucuronidation phenotypes to identify genetic variation in glucuronidation activity, then used the genetic model developed to test the relationship of glucuronidation to nicotine equivalents. Nicotine and cotinine glucuronidation levels in smokers' urine are correlated; however, the correlation will depend on the other pathways of nicotine and cotinine metabolism, primarily CYP2A6-catalyzed oxidation. Because of the greater catalytic efficiency of CYP2A6-catalyzed nicotine oxidation relative to cotinine oxidation (35, 36), the extent of cotinine glucuronidation will be less influenced by variation in CYP2A6 activity then will nicotine glucuronidation. Therefore, cotinine glucuronidation is a more stable measure of *N*-glucuronidation and an excellent surrogate for nicotine glucuronidation.

There has been great interest in evaluating the role of genetics in understanding the metabolism of nicotine and in predicting cancer risk among smokers (4, 37). Differences in the prevalence of genetic factors may assist in understanding the striking differences in lung cancer risk that have been noted between ethnic groups, especially at low and moderate levels of tobacco exposure. In the present study, we conducted a GWAS in search of common genetic variants that may be associated with nicotine and cotinine glucuronidation in a sample of 2,239 current smokers representing five racial/ethnic populations in the Multiethnic Cohort (MEC) Study.

Materials and Methods

Study population

The MEC consists of more than 215,000 men and women in California and Hawaii ages 45 to 75 years at recruitment, and comprises mainly five self-reported racial/ethnic populations: African Americans, Japanese, Latinos, Native Hawaiians, and European Americans (38, 39). Between 1993 and 1996, adults enrolled in the study by completing a mailed questionnaire asking detailed information about demographic factors, personal behaviors, and prior medical conditions. Potential participants were identified through driver's license files, voter registration lists, and Health Care Financing Administration data files. Between 1995 and 2006, blood specimens and either first morning or overnight urine were collected prospectively from approximately 67,000 participants for genetic and biomarker analyses. The Institutional Review Boards at the Universities of Southern California (Los Angeles, CA) and Hawaii (Honolulu, HI) approved the study protocol. A total of 2,393 current smokers at time of blood draw with no cancer diagnosis were assessed for inclusion.

Phenotypes

Nicotine, cotinine, and 3-HCOT in urine were analyzed by LC/MS-MS in a 96-well plate format using essentially the methods described previously (40, 41). The glucuronide conjugates were determined by analyzing the urine after treatment with β-glucuronidase, quantifying the total nicotine (nicotine plus nicotine *N*-glucuronide), total cotinine (cotinine plus cotinine-*N*-glucuronide), and total 3-HCOT (3-HCOT plus 3-HCOT-O-glucuronide), then calculating glucuronide concentrations as the difference between the free and total analyte. The coefficients of variation were (16.7 for nicotine, 10.1 for cotinine, and 11.4 for 3-HCOT). The main phenotypes analyzed were cotinine and nicotine glucuronidation, the ratio of cotinine glucuronide to total cotinine, and nicotine *N*-glucuronide to total nicotine, respectively. CYP2A6 phenotype was described by the ratio of total 3-HCOT to cotinine. To account for cigarette smoke exposure, nicotine equivalents, the sum of total nicotine, total cotinine, and 3-HCOT total (nmol/mg creatinine) were used for adjustment in analyses (27).

Genotyping and quality control

A total of 2,418 current smokers were genotyped using the Illumina Human1M-Duo BeadChip (1,199,187 SNPs). The genotyping quality control consisted of (i) removing individual samples with $\geq$2% of genotypes not called ($n = 8$), (ii) removing SNPs $\leq$98% call rate ($n = 67{,}761$), (iii) removing known duplicate samples ($n = 25$), and (iv) excluding samples with close relatives (as determined by estimated identity by descent status in pair wise comparisons, $n = 59$), and samples with conflicting or indeterminate sex ($n = 7$). The analysis included 1,131,426 SNPs and 2,239 samples.

Twenty-five replicate samples were included and the concordance was > 99.99%.The missense SNP in *UGT2B10* (rs61750900 Asp67Tyr) was not included on the BeadChip and TaqMan genotyping was not successful, clustering was relatively poor. Two other missense variants (rs147368959 IIe409Thr and

Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

rs111772923 Met>Ile) on chromosome 4 that were identified on the basis of the ESP project (42) and only found in African Americans were successfully genotyped by TaqMan in the majority (2240) of participants.

Genotype imputation

We used SHAPEIT (43) and IMPUTE2 (44) to extend our genotype analysis by imputing all SNPs appearing in the thousand genomes project (45) as of the March 2012 release. This extended our SNP association testing to a total of 11,892,802 genome-wide variants post quality control checks (1,131,426 genotyped and 10,761,376 imputed SNPs/indels). To remove poorly imputed SNPs from analysis, we filtered the data to include SNPs with an IMPUTE2 info score cutoff of $\geq$0.30 and minor allele frequency >1% by ethnic group. The *UGT2B10* missense SNP, rs61750900, was successfully imputed (with imputation scores from 0.94 to 1.0 among all ethnic groups) and our examination of this association was based on the imputed alleles. A *UGT2B10* splice variant, rs116294140, common in African Americans (46) was successfully imputed (imputation scores $\geq$0.93 among all groups).

Statistical analysis

We used a random sample of 19,059 autosomal SNPs with frequency $\geq$ 2% over the five racial/ethnic group samples to estimate principal components of ancestry. We used the program GCTA to compute a genetic relatedness matrix using these 19,059 SNPs and to output the top 10 leading eigenvectors from this matrix to adjust for population stratification in the analyses described below (47, 48).

Single SNP association testing

Individuals with low smoking levels (nicotine equivalents <1.4 nmol/mL, $n = 77$), and low genotype quality measures (as mentioned above) were excluded leaving a total of 2,239 smokers for analysis. For every SNP individually, linear regression models were applied to each phenotype, with adjustment for age, sex, reported ethnicity, nicotine equivalents, and the first 10 principal components described above. For a given SNP, the number of copies of the minor allele carried by each subject was used as the explanatory variable of most interest in the analysis and an additive model was fitted. Estimates, confidence intervals, and *P* values were computed as usual for linear regression, with a *P* value $> 5 \times 10^{-8}$ to establish global significance.

Multiple SNP regression

To determine the relative importance of multiple SNPs in a region or genome-wide, we used multiple regression methods. All SNPs showing globally significant associations were allowed to compete in forward selection regression models and all variables that entered with a significance level of $P < 0.001$ were retained. This *P* value allows for multiple testing of approximately 50 independent tagging SNPs in a given region, this is approximately the number of independent tagging SNPs in regions of similar size examined when fine mapping breast cancer associations in an African American sample (49). This allowed us to estimate the number of independent signals that may be involved in each region associated with each phenotype of interest. We expect some signals to be stronger, weaker, or absent in certain ethnic groups due to LD differences, or allele frequency differences between ethnic/racial groups, thus we also ran ethnic-specific analyses and tested for heterogeneity between ethnic groups in the impact of each SNP on each phenotype.

Results

A total of 2,239 smokers were included in the analysis, 53% were female (Table 1). On average, African American and Latino

Table 1. The descriptive characteristics of the multiethnic sample

		European Americans	African Americans	Latinos	Japanese Americans	Native Hawaiians
N (%)		437 (20%)	364 (16%)	453 (20%)	674 (30%)	311 (14%)
Sex, *N* (%)	Male	190 (44%)	111 (31%)	237 (52%)	388 (58%)	114 (37%)
	Female	247 (56%)	253 (69%)	216 (48%)	286 (42%)	197 (63%)
Age, mean (SD)	Male	63.3 (6.8)	63.5 (6.6)	66.2 (6.3)	63.7 (7.0)	63.0 (7.4)
	Female	64.0 (7.8)	65.3 (7.8)	64.7 (6.4)	63.8 (7.4)	60.4 (6.7]
	All	63.7 (7.4)	64.7 (7.5)	65.5 (6.3)	63.7 (7.1)	61.3 (7.1)
CPD: Mean (SD)	Male	20.8 (12.4)	12.1 (7.6)***	10.5 (8.0)***	15.4 (8.7)***	17.1 (10.8)**
	Female	15.2 (10.2)	10.8 (7.1)***	8.0 (6.3)***	12.1 (7.8)***	14.1 (8.9)
	All	17.6 (11.5)	11.2 (7.3)***	9.3 (7.3)***	14.0 (8.5)***	15.2 (9.7)***
	P^d	0.0001	0.1667	0.0003	0.0002	0.0056
NE[a]: Mean (SD)	Male	67.2 (35.5)	49.8 (30.0)***	45.4 (30.8)***	47.3 (48.9)***	48.1 (26.9)***
	Female	76.4 (48.1)	58.5 (35.2)***	54.8 (36.2)***	57.0 (40.7)***	61.2 (35.5)**
	All	72.4 (43.3)	55.9 (33.9)***	49.9 (33.8)***	51.4 (45.9)***	56.4 (33.1)***
	P^d	0.029	0.018	0.0021	0.0056	0.00050
Cot Gluc[b]: Mean (SD)	Male	58.4 (15.2)	44.7 (22.5)***	55.9 (16.3)	50.6 (14.7)***	54.5 (14.4)*
	Female	56.9 (15.5)	45.8 (23.0)***	60.5 (14.6)*	48.9 (14.5)***	53.5 (14.1)*
	All	57.6 (15.4)	44.0 (22.8)***	58.1 (15.7)	49.9 (14.6)***	53.8 (14.2)**
Nic Gluc[c]: Mean (SD)	Male	35.1 (19.3)	27.7 (19.7)**	39.0 (23.7)	33.4 (17.4)	30.8 (17.6)
	Female	35.7 (19.1)	29.1 (21.1)***	40.7 (21.4)*	33.5 (16.9)	31.2 (18.2)*
	All	35.4 (19.2)	28.7 (20.7)***	39.8 (22.6)**	33.5 (17.2)	31.0 (18.0)*

NOTE: *P* values adjusted for age (and gender) across ethnic groups (with European Americans as the reference) were indicated where significant as: *, $P < 0.05$; **, $P < 0.005$; and ***, $P < 0.0005$.

[a]NE (nicotine equivalents) is the sum of total nicotine, total cotinine, and total 3-hydroxycotinine expressed as nmol/mg creatinine. In an independent analysis of this cohort, NEs were expressed as nmol/mL (32).

[b]Cot Gluc (cotinine glucuronidation) is the ratio of the difference between total cotinine and free cotinine and total cotinine, expressed as percent.

[c]Nic Gluc (nicotine glucuronidation) is the ratio of the difference between total nicotine and free nicotine and total nicotine, expressed as percent.

[d]For CPD and NE, age-adjusted *P* values across ethnic groups were included (with males as the reference).

Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

Table 2. Ethnic differences in cotinine and nicotine glucuronidation[a] per nicotine equivalents

	Change in cotinine glucuronidation per NE[b]				Change in nicotine glucuronidation per NE[b]			
	N	β[c]	SE[c]	*P*[d]	*N*	β[c]	SE[c]	*P*[d]
European Americans	436	−0.007	0.017	0.675	436	−0.006	0.021	0.794
African Americans	364	0.028	0.036	0.428	364	−0.014	0.032	0.669
Latino Americans	453	−0.037	0.022	0.087	453	−0.064	0.032	0.044
Japanese Americans	674	−0.011	0.012	0.385	674	−0.047	0.014	1.30×10^{-3}
Native Hawaiians	311	−0.038	0.024	0.125	311	−0.100	0.031	1.40×10^{-3}
Overall	2239	−0.012	0.009	0.164	2239	−0.042	0.010	6.44×10^{-5}
$P_{heterogeneity}$				1.26×10^{-25}				7.69×10^{-11}

[a]Cotinine glucuronidation is the ratio of the difference of total cotinine and free cotinine over total cotinine.
[b]NE (nicotine equivalents) is the sum of total nicotine, total cotinine, and total 3′-hydroxycotinine expressed as nmol/mg creatinine.
[c]β values and SEs have been adjusted for age, gender, and race.
[d]*P* values adjusted for age, gender, and race.

smokers had lower tobacco smoke exposure compared with European Americans. They smoked significantly fewer CPD (11.2 and 9.3 vs. 17.6), and had significantly lower mean values of nicotine equivalents (55.9 and 49.9 vs. 72.4). The reported CPD for Japanese Americans was higher than for African and Latino Americans, but the level of nicotine equivalents was intermediate. However, if nicotine equivalents is expressed per urine volume, the concentration in African Americans is higher than in European Americans (32). African Americans and Native Hawaiians were found to have significantly lower nicotine and cotinine glucuronidation values than European Americans, both overall and among males and females (Table 1). Cotinine glucuronidation was lower in Japanese Americans, relative to European Americans and glucuronide levels among Latino Americans were similar or slightly higher than for European Americans.

The change in cotinine and nicotine glucuronidation per value of nicotine equivalents is presented in Table 2. For all ethnic groups, other than African Americans, there is a nonsignificant inverse relationship between cotinine glucuronidation and nicotine equivalents (β-ranged from 0.028 to −0.038). A similar inverse relationship between nicotine glucuronidation and nicotine equivalents was statistically significant among Latino Americans, Japanese Americans, and Native Hawaiians (Table 2). The *P* value for heterogeneity is significant for both cotinine and nicotine glucuronidation, indicating that there is a difference in slopes among the ethnic groups.

GWAS of cotinine glucuronidation

The GWAS analysis included 11,892,802 variants in 2,239 smokers. A total of 1,241 variants on 15 chromosomes were found to be strongly associated with cotinine glucuronidation ($P < 5 \times 10^{-8}$). The vast majority (1,076) of these associations was within a mega base of each other within chromosomal region 4q13 (between chr4:58148386 and chr4:79607027). Additional associations were found with variants in regions 1q32, 2q36, 4q12, 4q21, 5p13, 7p22, 7q11, 9q21, 9q31, 10p13, 11p15, 11q24, 12p13, 13q12, 14q21, 14q31, 15q14, 15q26, 16q13, 16q24, 19q13, and 20q13 (Fig. 1A–D and Supplementary Table S1). Through forward selection regression analysis of the 1,241 globally significant variants, we identified 15 independent signals comprising nine different chromosomes (Table 3), with four of the variants located within 190 kb of *UGT2B10* on 4q13. Of the 15 signals, 11 are intergenic and four are intronic variants. By far, the strongest association came from our top SNP in 4q13 (rs115765562, $P = 1.60 \times 10^{-155}$) near the gene *UGT2B10*. This SNP is in high LD ($r^2 = 0.97$) with the top SNP associated with total cotinine levels (rs835317, $P = 7.7 \times 10^{-43}$, data not shown).

Variability explained by SNPs and other variables

We fit forward linear regression models to evaluate the variation of cotinine glucuronidation explained by the most significant SNPs, and other baseline covariates (age, sex, nicotine equivalents, race, and principal components). Of the baseline variables, nicotine equivalents and sex were not important predictors for cotinine glucuronidation (with a combined R^2 of 0.05%, Table 4A). Race was a highly significant ($P < 0.0001$) predictor, explaining 8.5% of variability observed. Principal components were also significant predictors and captured 10.4% of phenotypic variation, and 2.27% when added to the model in addition to race ($P < 0.001$). The principal components correct for population stratification by accounting for a marker's variation in frequency across ancestral populations. They are most likely capturing the effects of admixture percentage as well as race, since three (Native Hawaiians, Latino Americans, and African Americans) of the five ethnic groups considered are admixed (48).

No pairwise interactions were found among the 15 variants deemed independently significant at *P* value < 0.005 (after correcting for the number of pairwise interactions tested). Therefore, we based our analysis on the main effects of the 15 variants; when added to the model, the fraction of variance explained by the model increased dramatically from 11.1% to 44.3%, that is, the variants alone explain 33.2% of variation. It is also important to note that variants on 4q13 near the gene *UGT2B10*, contribute a majority (27.4%) of the explained variability observed in cotinine glucuronidation. Our top most significant SNP, rs115765562, accounts for 24.2% of variability in cotinine glucuronidation.

Genetic score

We further considered the performance of a simple genetic score; a weighted sum of alleles associated with the phenotype using the (univariate) regression coefficient estimates as weights. The weighted genetic score explained a very similar amount of variation (31.3%) as did the total of the main effects of the 15 variants constituting the score (Table 4A).

LD with other variants

Of the 1,241 genome-wide significant associations, we found three missense SNPs, a synonymous variant and one splice variant (Supplementary Table S1). However, none of these protein-altering SNPs are among the 15 variants that are in our final model. We checked to see whether any of the 15 variants are in high LD with these coding variants. The highest correlations between a protein-coding variant and any of the 15 SNPs that entered were between the nonsynonymous SNP rs9530 (gene: *GUSB*, β-glucuronidase), and the intergenic SNP rs6952407 both on chromosome 7q11

Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

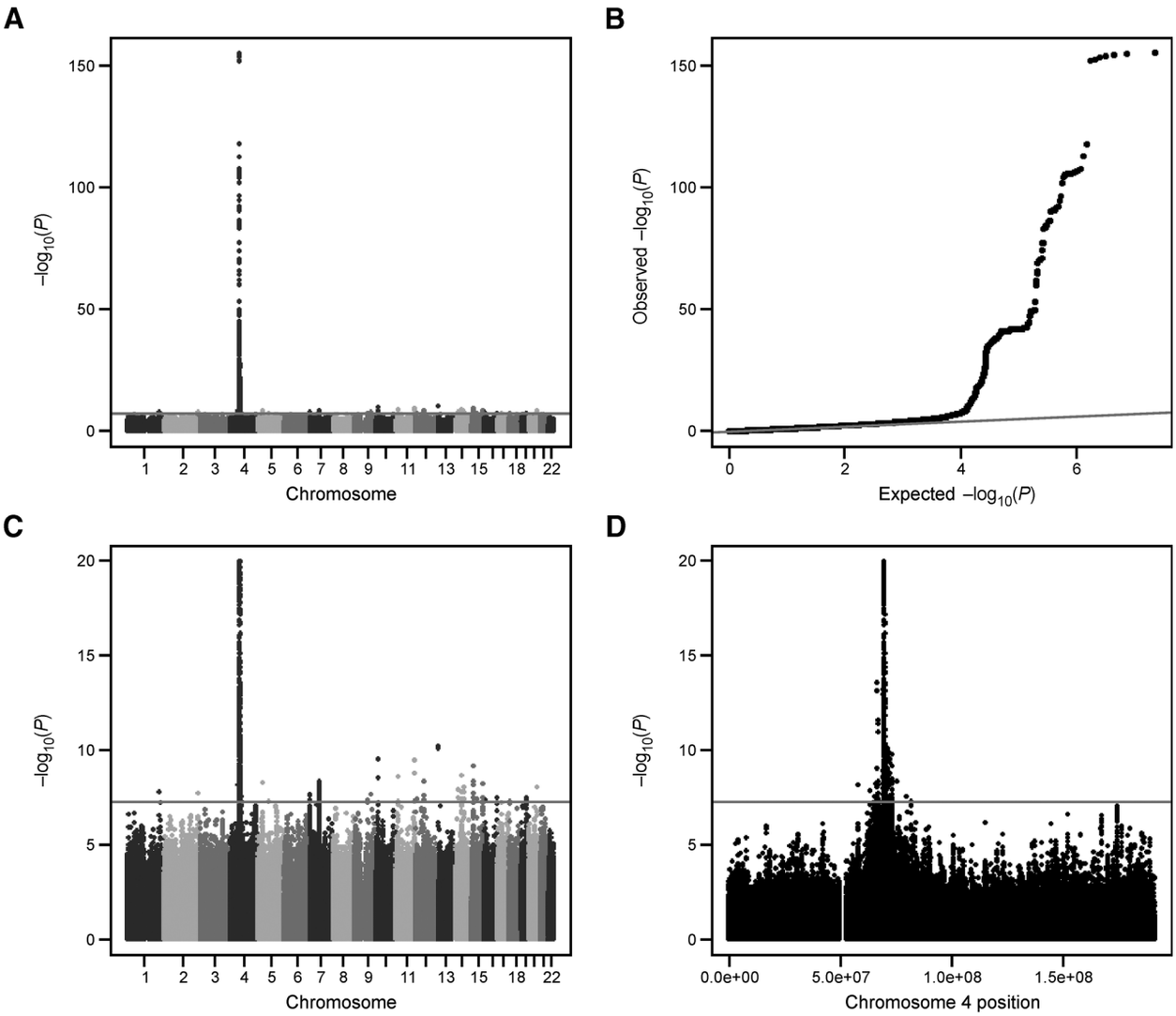

Figure 1.
A, Manhattan plot of the $-\log_{10}$ (*P* values) from the test of association for cotinine glucuronidation plotted as a function of the chromosomal position. Genome-wide significance is defined as the Bonferroni-corrected 5% significance threshold ($P < 5.0 \times 10^{-8}$) and is indicated as a horizontal line. B, quantile–quantile plot of the GWAS results for cotinine glucuronidation. C, Manhattan plot with the scale of the *y*-axis [$-\log_{10}$ (*P*) values] reduced to 1.0×10^{-20} for visual acuity of all significant associations. D, Manhattan plot of chromosome 4–specific $-\log_{10}$ (*P* values) from the test of association for cotinine glucuronidation.

(overall $R^2 = 0.84$). Our top most significant hit on 4q13, rs115765562, was strongly correlated with the splice variant, rs116294140 ($R^2 = 0.60$). Another one of our most significant SNPs on chromosome 4q13, rs141360540, was correlated with the known *UGT2B10* missense SNP rs61750900 as well as synonymous SNP, rs61749966 with R^2 values of 0.34. All other overall correlations between protein-altering SNPs and SNPs in the model were <0.20.

Ethnic-specific analyses

Because the vast majority of the signal is restricted to regions on 4q13, we focused our ethnic-specific analysis on this chromosome. When examining SNPs on 4q13, a total of 99 SNPs were globally significant in one or more of the ethnic-specific analyses but were not found to be significant at *P* value < 5×10^{-8} in the overall results. Among the 404 globally significant associations for African Americans, there were 14 SNPs that were not found in the overall analysis for cotinine glucuronidation, for European Americans there were two new significant associations out of 328, for Japanese Americans there were 72 out of 412, for Latinos 11 out of 497, and there were no new associations in the ethnic-specific analysis for Native Hawaiians (Supplementary Tables S3–S7).

When significant SNPs from the ethnic-specific analyses were allowed to compete with the six independent signals observed from the overall analyses for cotinine glucuronidation on 4q13, only one SNP, rs10029577 a *UGT2B28* variant, additionally entered the model for African Americans. When added to the model with the six 4q13 variants, rs10029577 only explains an additional 0.9% of variation in cotinine glucuronidation in African Americans. No additional SNPs entered the model at $P < 1 \times 10^{-3}$ among any of the other ethnic groups, indicating

Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

Table 3. SNPs that enter stepwise regression

SNP	Alt Rs#	Chr	BP	A1	A2	Info	β	P[a]	Type	Nearest gene
Cotinine glucuronidation										
rs115765562	rs34100980	4	69673553	C	A	1.1377	0.1833	1.60E−155	Intergenic	*UGT2B10*
rs141360540	rs10028938	4	69669216	G	A	0.9535	−0.1587	5.51E−91	Intergenic	*UGT2B10*
rs115219551	rs9997650	4	69669388	A	G	0.8903	−0.0941	1.70E−50	Intergenic	*UGT2B10*
rs294777		4	69682471	A	G	1.2244	0.1905	5.53E−48	Intronic	*UGT2B10*
rs6832720		4	58148386	A	G	1.2635	−0.0585	7.36E−09	Intergenic	*OC255130*
rs1115363		4	67073143	A	T	0.9697	0.0974	7.38E−09	Intergenic	*LOC100144602*
rs6952407		7	66045512	A	G	1.1719	0.033	4.51E−09	Intergenic	*LOC493754/KCTD7*
rs60634637		9	110873780	T	C	1.0462	0.0942	2.14E−08	Intergenic	*KLF4/ACTL7B*
rs4750535		10	14633024	C	T	1.0947	0.0999	2.99E−10	Intronic	*FAM107B*
rs4287304		11	16559630	C	T	0.8419	−0.1028	2.53E−09	Intronic	*SOX6*
chr12:7996130:D	rs202090541	12	7996130	AC	A	0.9222	0.061	3.88E−08	Intronic	*NECAP1/SLC2A14*
rs76513344		14	63047683	C	A	1.0516	−0.0995	2.16E−09	Intergenic	*KCNH5*
rs80332023		15	101073444	C	T	1.0749	−0.12	3.35E−08	Intronic	*CERS3/LASS3*
rs60283548		15	97454880	C	T	1.062	−0.1327	3.57E−08	Intergenic	*7SK/SPATA8*
rs34705275		19	43933725	A	C	1.158	0.0337	3.91E−08	Intergenic	*TEX101/LYPD3*
Nicotine glucuronidation										
rs116224959	rs835315	4	69685772	G	A	1.1326	0.1193	2.56E−43	Intronic	*UGT2B10*
rs4132568		7	96134981	A	C	1.0853	−0.0343	2.69E−08	Intronic	*SHFM1*

[a]Overall *P* values were analyzed via linear regression in PLINK with adjustment for age, sex, race, principal components 1 to 10, and nicotine equivalents.

that the six independent signals sufficiently capture the variability noted in ethnic-specific analyses.

We further examined the ethnic variations explained by the full weighted genetic score (Table 5). The addition of the weighted genetic score to the model for African Americans explains 55% of variability. Amongst Latinos, the genetic score explains 30% variability, and similar variations were noted for Native Hawaiians and European Americans (25.6%, 21%, respectively), with the least variability explained for the Japanese Americans at approximately 19%. This high predictive value of the genetic score in African Americans may be due to the high frequency of the most influential SNP, rs294777, among African Americans (22%), compared with 2% in Latino Americans and null among Native Hawaiians, European Americans, and Japanese Americans.

Nicotine glucuronidation

There were 492 globally significant SNPs for nicotine glucuronidation, most of which were in 4q13 near *UGT2B10* (between positions 69592725 and 7013816); 490 of these top hits were also globally significant for cotinine glucuronidation (Supplementary Table S2 and Supplementary Fig. S1A-S1D). These findings included the original nonsynonymous SNP of interest, rs61750900, and the *UGT2B10* splice variant, rs116294140, found here to be associated at 3.34×10^{-17} and 4.61×10^{-23}, respectively. Two intronic SNPs on chromosome 7 near gene *SHFM1* were also found to be globally significant for nicotine glucuronidation.

In a forward selection analysis as described above, two SNPs, our top most association, a *UGT2B10* intronic SNP rs116224959, and an intronic variant on chromosome 7 near *SHFM1*,

Table 4. Determinants of cotinine and nicotine glucuronidation

Model	R^2	N	Percentage variation explained	
A. Cotinine glucuronidation				
Nicotine equivalents	0.0001	2,239	NA	
+Sex	0.0005	2,239	0.04	
+Age	0.0037	2,239	0.32	
+Race	0.0887	2,239	8.50	
+Principal components 1–10	0.1114	2,239	2.27	Base
Base model: Cotinine glucuronidation = nicotine equivalents + sex + age + race + principal components 1–10				
Data				
Base model + 15 SNPs from stepwise	0.4433	2,239	33.19	Compared with base
Base model + weighted GS with 15 SNPs from stepwise[a]	0.4246	2,239	31.32	Compared with base
B. Nicotine glucuronidation				
Nicotine equivalents	0.0068	2,239	NA	
+Sex	0.0069	2,239	0.01	
+Age	0.0121	2,239	0.52	
+Race	0.0447	2,239	3.26	
+Principal components 1–10	0.0513	2,239	0.66	Base
Base model: Nicotine glucuronidation = nicotine equivalents + sex + age + race + principal components 1–10				
Data				
Base model + 2 SNPs from stepwise (rs116224959 and rs4132568)	0.1412	2,239	8.99	Compared with base
Base model + CotGluc weighted GS with 15 SNPs from stepwise (from model in "A. Cotinine glucuronidation")	0.1365	2,239	8.52	Compared with base

[a]The weighted GS was weighted with the betas from the overall GWAS results.

Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

Table 5. Ethnic-specific percentage of variation explained by GS for cotinine glucuronidation

Base model - cotinine glucuronidation: nicotine equivalents + sex + age + principal components 1–10 + weighted GS					
Ethnic-specific weighted GS	*N*	Base R^2	Weighted GS R^2	Percentage variation explained by weighted GS	*P*
European Americans	437	0.0292	0.2395	21.03	3.47×10^{-24}
African Americans	364	0.1017	0.6508	54.91	1.34×10^{-73}
Latinos	453	0.0629	0.3629	30.00	1.35×10^{-38}
Japanese Americans	674	0.0252	0.2148	18.96	7.89×10^{-33}
Native Hawaiians	311	0.0971	0.3531	25.60	3.17×10^{-23}

rs4132568, entered the model, indicating that there are two independent signals driving the overall association (Table 3). The *UGT2B10* variant, rs116224959, was also among the very top SNPs for cotinine glucuronidation with $P = 8.71 \times 10^{-153}$, and is in high LD with ($R^2 = 0.99$) rs115765562, the top most SNP that remains in the forward selection for cotinine glucuronidation. No new markers were observed when comparing ethnic-specific results with the overall associations for nicotine glucuronidation.

The weighted genetic score comprising the 15 cotinine glucuronidation SNPs explains approximately 8.5% of the variation for nicotine glucuronidation (Table 4B). On its own, rs116224959 explains a majority (7.80%) of variance noted in nicotine glucuronidation, though this is substantially smaller than the 23.8% observed for cotinine glucuronidation for this SNP alone (not shown). The splice variant, rs116294140, explains an overall variation of 4.09%, and 9.08% among African Americans, lower than what is noted for rs116224959 at 7.80% overall and 11.0%.

Additional analyses

We determined the possible effects of the weighted cotinine glucuronidation genetic score on smoking behavior, either as CPD or as nicotine equivalents. We did not find any association between the genetic score and nicotine equivalents ($P = 0.41$). Neither did we find an association between the genetic score and CPD ($P = 0.54$). We also analyzed two *UGT2B10* missense variants, rs147368959 and rs111772923, found only among African Americans with frequencies of 4% and 7% in our dataset. Neither of these variants were significantly associated with cotinine glucuronidation in African Americans ($P = 0.61$ and 0.96, respectively).

One aspect of the results of the multiple regression for cotinine glucuronidation that is puzzling is that in single SNP analyses, 15 chromosomes showed globally significant associations ($P < 5 \times 10^{-8}$), whereas only nine chromosomes are represented among the SNPs chosen in the forward regression analysis using an entry criteria of ($P < 1 \times 10^{-3}$); this was seen despite no LD existing between different chromosomes after correction for principal components. A possible explanation for this is the presence of interactions between those SNPs on the chromosomes not represented in the score and those included in the score. Indeed when we looked between the 15 SNPs in the genetic score and the 10 SNPs which were globally significant but on one of the missing 6 chromosomes, we found significant pairwise interactions ($P = 0.01$) for several of them. However, the amount of variance that these interactions accounted for was very small compared with the large amount explained by the main effects and we did not consider these SNPs further. We also tested for SNP by race interactions for the 15 variants that remained in the forward selection for cotinine glucuronidation. Three significant interactions ($P < 0.01$) were found for race by SNPs (rs115765562, chr12_7996130_D and rs80332023), though when added to the model, these interactions only explained 0.24% of additional variability in cotinine glucuronidation. No such significant SNP by race interactions were found for nicotine glucuronidation.

Discussion

Assessing the genetic contribution in the metabolism of nicotine may be important in assessing the racial/ethnic differences in lung cancer risk among smokers (4, 37, 50) as individuals with a genetic basis for fast metabolism of nicotine may smoke more CPD than those with slower metabolism (2). Prior studies have focused on the catalyst of nicotine C-oxidation, *CYP2A6* and variants in this gene have been reported to be associated with smoking and lung cancer risk (51, 52). However, nicotine glucuronidation may account for up to 40% of the nicotine equivalents excreted by smokers (16, 17). *UGT2B10* catalyzes both nicotine and cotinine glucuronidation (21–23), and our analysis has determined that a high fraction of individual variation of cotinine glucuronidation is explained by genetic differences, which can be parsimoniously characterized using a genetic score of 15 SNPs from nine chromosomes with SNPs near *UGT2B10* showing the strongest associations. The fraction of variance explained by this genetic score is estimated to be 33% overall ethnic groups considered.

We based our analysis primarily on cotinine glucuronidation rather than nicotine glucuronidation because the same enzyme is responsible for their formation and nicotine is temporally more variable than cotinine. SNPs predictive of nicotine glucuronidation were also predictive of cotinine glucuronidation. Of the six variants at 4q13 maintained in the model for cotinine glucuronidation (Table 3), the four SNPs near *UGT2B10* were found to be significantly associated with nicotine glucuronidation at $P < 5 \times 10^{-8}$.

A very small fraction of the globally significant associations involved missense SNPs or other protein-coding SNPs (e.g., splice site variants). Of the SNPs maintained in the forward regression model only one, on 7q11, a far less predictive region than chromosome 4, was in high LD with a missense SNP ($r^2 = 0.84$). The SNP is in the coding region of the enzyme β-glucuronidase, which cleaves glucuronide conjugates. Variation in this enzyme could impact the levels of nicotine and cotinine glucuronide excreted. However, β-glucuronidase is a lysosomal enzyme, only a small amount is present in the plasma, so the influence of this enzyme on circulating nicotine levels would likely be small (53).

Although all of the remaining SNPs selected in the forward selection regression model were either intronic or intergenic, this does not in itself negate the possibility that common missense variation may still be playing an important role in the associations seen here. Focusing on 4q13, we found our most

Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

significant association, rs115765562, to be highly correlated with the splice variant, rs116294140. When forced into the regression model, the missense SNP (the Asp67Try, rs61750900) and the splice site variant, alone explain 24.3% of the variation in cotinine glucuronidation. This compares with the 28.1% including all six chromosome 4 SNPs (a small but strongly significant improvement in R^2). Clearly much of the variation in cotinine glucuronidation could be due to the Asp67Tyr and splice site variants; however, many nearly equivalent alternative choices of best predictors can be found in the chromosome 4 region, reflecting a complex pattern of linkage disequilibrium there, so that genetic regulation, rather than the effect of direct coding changes, cannot be ruled out as a primary mechanism affecting glucuronidation.

We have previously reported lower levels of nicotine equivalents in individuals heterozygous for the *UGT2B10* Asp67Tyr genotype (rs61750900) compared with those without the allele (19, 21). In the present study, unlike the previous report, which had a smaller sample size and included fewer ethnic groups, we do not find that the Asp67Tyr variant is related to nicotine equivalents ($P = 0.62$). In addition, the genetic score is not significantly related to nicotine equivalents. Our overall conclusion is that *UGT2B10* variants are less of a factor in determining nicotine dose than initially suspected.

To date most published GWAS for smoking behavior have been conducted in European populations, motivating exploration in other ancestry groups to help understand the differences in genetic diversity in smoking behavior and tobacco dependence (37, 54). Although our single SNP analysis by ethnic groups did not show notable differences by ethnicity, the weighted genetic score is more predictive in some groups than in others. The fraction of variability in cotinine glucuronidation explained by the genetic score ranges from 55% for African Americans to 19% for Japanese Americans. The high predictive value in African Americans reflects that the most influential of the SNPs (rs294777) included is only common (22%) in African Americans and is not present in European Americans, Japanese Americans, or Native Hawaiians. The predictive value in African Americans also may arguably be driven by a similar pattern of association with the *UGT2B10* splice variant, which has a frequency of 35% in African Americans and from 0.1% to 8.0% in the other groups. The much lower predictive value of the genetic score in Japanese Americans may be due to the higher prevalence of *CYP2A6* null variants in this group (27). We previously reported that *CYP2A6* alleles contribute to variation in plasma nicotine glucuronide levels among European Americans (23), and in the subjects of this study, the CYP2A6 metabolic ratio for Japanese Americans was half the ratio for African Americans (32). In both Japanese Americans and Native Hawaiians, low CYP2A6 activity is associated with decreased nicotine equivalents (29). This relationship may explain the inverse relationship between nicotine equivalents and nicotine glucuronidation that we see with these two groups because decreased CYP2A6-catalyzed nicotine C-oxidation results in increased nicotine *N*-glucuronidation (17, 32). Genotyping of common *CYP2A6* alleles in the current study is on-going (*CYP2A6* is not well covered by GWAS arrays) and these data, with additional nicotine metabolism phenotypes and the GWAS data will be used to gain a more complete understanding of the genetics of nicotine metabolism and tobacco use.

The ethnic differences in nicotine and cotinine glucuronidation are interesting, but do not seem to be directly related to the differences in cancer risk seen between the five racial/ethnic groups. African Americans have the lowest levels of nicotine glucuronidation among the groups, but their lung cancer risk is the highest. There was no significant association between nicotine equivalents and nicotine glucuronidation among African Americans. Therefore, the relatively high prevalence of *UGT2B10* variants in African Americans does not seem to influence smoking levels; however, it may result in decreased detoxification of tobacco carcinogens (55).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: D.O. Stram, L.R. Wilkens, L. Le Marchand, S.E. Murphy
Development of methodology: L. Le Marchand, S.E. Murphy
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): L.R. Wilkens, L. Le Marchand, C.A. Haiman, S.E. Murphy
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y. Patel, D.O. Stram, C.A. Haiman, S.E. Murphy
Writing, review, and/or revision of the manuscript: Y. Patel, D.O. Stram, L.R. Wilkens, S.L. Park, B.E. Henderson, L. Le Marchand, S.E. Murphy
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S.L. Park, L. Le Marchand, S.E. Murphy
Study supervision: D.O. Stram, L. Le Marchand, S.E. Murphy

Grant Support

This work was supported by the NIH (Program Project Grant CA-138338 to D.O. Stram, L. Le Marchand, C.A. Haiman, and S.E. Murphy) and the Ethnic Differences in Cancer: The MEC Study (grant 5UM1CA164973-02 to L. Le Marchand and D.O. Stram).

Received July 14, 2014; revised September 24, 2014; accepted September 30, 2014; published OnlineFirst October 7, 2014.

References

1. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer 2003;3:733–44.
2. Benowitz NL. Nicotine addiction. N Engl J Med 2010;362:2295–303.
3. Hukkanen J, Jacob P III, Benowitz NL. Metabolism and disposition kinetics of nicotine. Pharmacol Rev 2005;57:79–115.
4. Wassenaar CA, Dong Q, Wei Q, Amos CI, Spitz MR, Tyndale RF. Relationship between CYP2A6 and CHRNA5-CHRNA3-CHRNB4 variation and smoking behaviors and lung cancer risk. J Natl Cancer Inst 2011;103:1342–6.
5. Vartiainen E, Seppala T, Lillsunde P, Puska P. Validation of self reported smoking by serum cotinine measurement in a community-based study. J Epidemiol Community Health 2002;56:167–70.
6. Perez-Stable EJ, Benowitz NL, Marin G. Is serum cotinine a better measure of cigarette smoking than self-report? Prev Med 1995;24:171–9.
7. Boffetta P, Clark S, Shen M, Gislefoss R, Peto R, Andersen A. Serum cotinine level as predictor of lung cancer risk. Cancer Epidemiol Biomarkers Prev 2006;15:1184–8.
8. Haiman CA, Stram DO, Wilkens LR, Pike MC, Kolonel LN, Henderson BE, et al. Ethnic and racial differences in the smoking-related risk of lung cancer. N Engl J Med 2006;354:333–42.
9. Le Marchand L, Wilkens LR, Kolonel LN. Ethnic differences in the lung cancer risk associated with smoking. Cancer Epidemiol Biomarkers Prev 1992;1:103–7.

10. Le Marchand L, Derby KS, Murphy SE, Hecht SS, Hatsukami D, Carmella SG, et al. Smokers with the CHRNA lung cancer-associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. Cancer Res 2008;68:9137–40.
11. Nakajima M, Yokoi T. Interindividual variability in nicotine metabolism: C-oxidation and glucuronidation. Drug Metab Pharmacokinet 2005;20:227–35.
12. Benowitz NL, Jacob P3rd. Metabolism of nicotine to cotinine studied by a dual stable isotope method. Clin Pharmacol Ther 1994;56:483–93.
13. Peterson LA, Trevor A, Castagnoli NJr. Stereochemical studies on the cytochrome P-450 catalyzed oxidation of (S)-nicotine to the (S)-nicotine delta 1′(5′)-iminium species. J Med Chem 1987;30:249–54.
14. Messina ES, Tyndale RF, Sellers EM. A major role for CYP2A6 in nicotine C-oxidation by human liver microsomes. J Pharmacol Exp Ther 1997;282:1608–14.
15. Nakajima M, Yamamoto T, Nunoya K, Yokoi T, Nagashima K, Inoue K, et al. Characterization of CYP2A6 involved in 3′-hydroxylation of cotinine in human liver microsomes. J Pharmacol Exp Ther 1996;277:1010–5.
16. Murphy SE, Link CA, Jensen J, Le C, Puumala SS, Hecht SS, et al. A comparison of urinary biomarkers of tobacco and carcinogen exposure in smokers. Cancer Epidemiol Biomarkers Prev 2004;13:1617–23.
17. Yamanaka H, Nakajima M, Nishimura K, Yoshida R, Fukami T, Katoh M, et al. Metabolic profile of nicotine in subjects whose CYP2A6 gene is deleted. Eur J Pharm Sci 2004;22:419–25.
18. Kaivosaari S, Toivonen P, Hesse LM, Koskinen M, Court MH, Finel M. Nicotine glucuronidation and the human UDP-glucuronosyltransferase UGT2B10. Mol Pharmacol 2007;72:761–8.
19. Chen G, Blevins-Primeau AS, Dellinger RW, Muscat JE, Lazarus P. Glucuronidation of nicotine and cotinine by UGT2B10: loss of function by the UGT2B10 Codon 67 (Asp>Tyr) polymorphism. Cancer Res 2007;67:9024–9.
20. Berg JZ, Mason J, Boettcher AJ, Hatsukami DK, Murphy SE. Nicotine metabolism in African Americans and European Americans: variation in glucuronidation by ethnicity and UGT2B10 haplotype. J Pharmacol Exp Ther 2010;332:202–9.
21. Berg JZ, von Weymarn LB, Thompson EA, Wickham KM, Weisensel NA, Hatsukami DK, et al. UGT2B10 genotype influences nicotine glucuronidation, oxidation, and consumption. Cancer Epidemiol Biomarkers Prev 2010;19:1423–31.
22. Chen G, Giambrone NE Jr, Dluzen DF, Muscat JE, Berg A, Gallagher CJ, et al. Glucuronidation genotypes and nicotine metabolic phenotypes: importance of functional UGT2B10 and UGT2B17 polymorphisms. Cancer Res 2010;70:7543–52.
23. Bloom AJ, von Weymarn LB, Martinez M, Bierut LJ, Goate A, Murphy SE. The contribution of common UGT2B10 and CYP2A6 alleles to variation in nicotine glucuronidation among European Americans. Pharmacogenet Genomics 2013;23:706–16.
24. Byrd GD, Chang KM, Greene JM, deBethizy JD. Evidence for urinary excretion of glucuronide conjugates of nicotine, cotinine, and trans-3′-hydroxycotinine in smokers. Drug Metab Dispos 1992;20:192–7.
25. Chen G, Giambrone NE, Lazarus P. Glucuronidation of trans-3′-hydroxycotinine by UGT2B17 and UGT2B10. Pharmacogenet Genomics 2012;22:183–90.
26. Scherer G, Engl J, Urban M, Gilch G, Janket D, Riedel K. Relationship between machine-derived smoke yields and biomarkers in cigarette smokers in Germany. Regul Toxicol Pharmacol 2007;47:171–83.
27. Wang J, Liang Q, Mendes P, Sarkar M. Is 24h nicotine equivalents a surrogate for smoke exposure based on its relationship with other biomarkers of exposure? Biomarkers 2011;16:144–54.
28. Benowitz NL, Jacob P3rd. Individual differences in nicotine kinetics and metabolism in humans. NIDA Res Monogr 1997;173:48–64.
29. Derby KS, Cuthrell K, Caberto C, Carmella SG, Franke AA, Hecht SS, et al. Nicotine metabolism in three ethnic/racial groups with different risks of lung cancer. Cancer Epidemiol Biomarkers Prev 2008;17:3526–35.
30. Fujieda M, Yamazaki H, Saito T, Kiyotani K, Gyamfi MA, Sakurai M, et al. Evaluation of CYP2A6 genetic polymorphisms as determinants of smoking behavior and tobacco-related lung cancer risk in male Japanese smokers. Carcinogenesis 2004;25:2451–8.
31. Kumasaka N, Aoki M, Okada Y, Takahashi A, Ozaki K, Mushiroda T, et al. Haplotypes with copy number and single nucleotide polymorphisms in CYP2A6 locus are associated with smoking quantity in a Japanese population. PLoS ONE 2012;7:e44507.
32. Murphy SE, Park S-SL, Thompson EF, Wilkens LR, Patel YM, Stram DO, et al. Nicotine *N*-glucuronidation relative to *N*-oxidation and *C*-oxidation and UGT2B10 genotype in five ethnic/racial groups. Carcinogenesis 2014;35:2526–33.
33. Benowitz NL, Perez-Stable EJ, Fong I, Modin G, Herrera B, Jacob P III. Ethnic differences in N-glucuronidation of nicotine and cotinine. J Pharmacol Exp Ther 1999;291:1196–203.
34. Kuehl GE, Murphy SE. N-glucuronidation of nicotine and cotinine by human liver microsomes and heterologously expressed UDP-glucuronosyltransferases. Drug Metab Dispos 2003;31:1361–8.
35. Brown KM, von Weymarn LB, Murphy SE. Identification of *N*-(hydroxymethyl) norcotinine as a major product of cytochrome P450 2A6, but not cytochrome P450 2A13-catalyzed cotinine metabolism. Chem Res Toxicol 2005;18:1792–8.
36. Murphy SE, Raulinaitis V, Brown KM. Nicotine 5′-oxidation and methyl oxidation by P450 2A enzymes. Drug Metab Dispos 2005;33:1166–73.
37. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet 2010;42:441–7.
38. Stram DO, Hankin JH, Wilkens LR, Pike MC, Monroe KR, Park S, et al. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. Am J Epidemiol 2000;151:358–70.
39. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 2000;151:346–57.
40. Bloom J, Hinrichs AL, Wang JC, von Weymarn LB, Kharasch ED, Bierut LJ, et al. The contribution of common CYP2A6 alleles to variation in nicotine metabolism among European-Americans. Pharmacogenet Genomics 2011;21:403–16.
41. Murphy SE, Villalta P, Ho SW, von Weymarn LB. Analysis of [3′,3′-d_2]-nicotine and [3′,3′-d_2]-cotinine by capillary liquid chromatography-electrospray tandem mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 2007;857:1–8.
42. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 2012;337:64–9.
43. Delaneau O, Marchini J, Zagury J-F. A linear complexity phasing method for thousands of genomes. Nature Methods 2011;9:179–81.
44. Howie BN, Donnelly P, Marchini J. Impute2: A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5:e1000529.
45. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061–73.
46. (ESP) NESP. Exome Variant Server. [Accessed July 8, 2014]. Available from: http://evs.gs.washington.edu/EVS
47. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82.
48. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–9.
49. Chen F, Chen GK, Millikan RC, John EM, Ambrosone CB, Bernstein L, et al. Fine-mapping of breast cancer susceptibility loci characterizes genetic risk in African Americans. Hum Mol Genet 2011;20:4491–503.
50. Tyndale RF. Genetics of alcohol and tobacco use in humans. Ann Med 2003;35:94–121.
51. Wang L, Zang W, Liu J, Xie D, Ji W, Pan Y, et al. Association of CYP2A6*4 with susceptibility of lung cancer: a meta-analysis. PLoS ONE 2013;8:e59556.
52. Thorgeirsson TE, Gudbjartsson DF, Surakka I, Vink JM, Amin N, Geller F, et al. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. Nat Genet 2010;42:448–53.
53. Gratz M, Kunert-Keil C, John U, Cascorbi I, Kroemer HK. Identification and functional analysis of genetic variants of the human beta-glucuronidase in a German population sample. Pharmacogenet Genomics 2005;15:875–81.
54. Caporaso N, Gu F, Chatterjee N, Sheng-Chih J, Yu K, Yeager M, et al. Genome-wide and candidate gene association study of cigarette smoking behaviors. PLoS ONE 2009;4:e4653.
55. Chen G, Dellinger RW, Sun D, Spratt TE, Lazarus P. Glucuronidation of tobacco-specific nitrosamines by UGT2B10. Drug Metab Dispos 2008;36:824–30.

Published OnlineFirst October 7, 2014; DOI: 10.1158/1055-9965.EPI-14-0815

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

The Contribution of Common Genetic Variation to Nicotine and Cotinine Glucuronidation in Multiple Ethnic/Racial Populations

Yesha M. Patel, Daniel O. Stram, Lynne R. Wilkens, et al.

Cancer Epidemiol Biomarkers Prev 2015;24:119-127. Published OnlineFirst October 7, 2014.