Published in final edited form as:

Pharmacogenet Genomics. 2012 June ; 22(6): 429–440. doi:10.1097/FPC.0b013e3283527c1c.

CYP2A6 and CYP2B6 genetic variation and its association with nicotine metabolism in South Western Alaska Native people

Matthew J. Binnington¹, Andy Z.X. Zhu¹, Caroline C. Renner², Anne P. Lanier², Dorothy K. Hatsukami³, Neal L Benowitz⁴, and Rachel F. Tyndale¹

¹Centre for Addiction & Mental Health, Departments of Psychiatry, Pharmacology and Toxicology, University of Toronto, Toronto, Ontario.

²Alaska Native Tribal Health Consortium (ANTHC), Clinical & Research Services, Anchorage, Alaska.

³Department of Psychiatry, University of Minnesota, Minneapolis, Minnesota.

⁴Departments of Medicine, Bioengineering & Therapeutic Sciences, and Clinical Pharmacology, University of California San Francisco, San Francisco, California

Abstract

Objectives—Alaska Native people (AN) have a high prevalence of tobacco use and associated morbidity and mortality when compared to the general U.S. population. Variation in the *CYP2A6* and *CYP2B6* genes, encoding enzymes responsible for nicotine metabolic inactivation and procarcinogen activation, has not been characterized in AN and may contribute to the increased risk.

Methods—AN people ($n = 400$) residing in the Bristol Bay region of South Western Alaska were recruited for a cross-sectional study on tobacco use. They were genotyped for *CYP2A6**1X2A, *1X2B, *1B, *2, *4, *7, *8, *9, *10, *12, *17, *35 and *CYP2B6**4, *6, *9 and provided plasma and urine samples for measurement of nicotine and metabolites.

Results—*CYP2A6* and *CYP2B6* variant frequencies among the AN Yupik people ($n=361$) were significantly different from other ethnicities. Nicotine metabolism (as measured by the plasma and urinary ratio of metabolites *trans*-3’hydroxycotinine to cotinine [(3HC/COT)] was significantly associated with *CYP2A6* ($P < 0.001$) but not *CYP2B6* genotype ($P = 0.95$) when controlling for known covariates. Of note, plasma 3HC/COT ratios were high in the entire Yupik people, and among the Yupik *CYP2A6* wild-type participants they were substantially higher than previously characterized racial/ethnic groups ($P < 0.001$ vs. Caucasians and African Americans).

Correspondence: Dr. Rachel F. Tyndale, Room 4326, Medical Sciences Building, 1 King’s College Circle, University of Toronto, Toronto, Ontario M5S 1A8, Canada, Telephone: (+1) 416 978 6374, Fax: (+1) 416 978 6395.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest

R.F.T. holds shares in Nicogen Research, a company focused on novel approaches to smoking cessation treatments. No Nicogen funds were used in this work, others affiliated with Nicogen did not review the manuscript, and none of the data in the manuscript benefits any commercial aspect of Nicogen. R.F.T. has also been a paid consultant for Novartis and McNeil. N.L.B. serves as a paid expert witness in litigation against tobacco companies and a paid advisor to several pharmaceutical companies developing or marketing medications for smoking cessation.

Conclusions—Yupik AN people have a unique *CYP2A6* genetic profile which associated strongly with *in vivo* nicotine metabolism. More rapid *CYP2A6*-mediated nicotine and nitrosamine metabolism in the Yupik people may modulate tobacco-related disease risk.

Keywords

CYP2A6; CYP2B6; nicotine; tobacco; smoking; genetic variation; Alaska Native people

Introduction

Tobacco use is the most prominent cause of preventable death in the world [1]. Its consumption results in greater incidence of cardiovascular disease, pulmonary disease, and many cancers [2]. Alaska Native (AN) peoples have the highest prevalence of tobacco use among ethnic minorities in the U.S., where 41% are current smokers compared to 20% of the general U.S. population [3]. Further, 11% of AN peoples residing in Alaska currently use smokeless tobacco (ST) vs. 4% of non-native Alaska residents [3]. Additionally, the AN population is characterized by use of a homemade ST mixture known as iqmik, a combination of leaf tobacco and ash from a local tree fungus, *Phellinus igniarius* [4, 5]. The fungus ash raises product pH, which is expected to result in greater oral bioavailability and absorption of nicotine [4].

Although tobacco use is highly prevalent among the AN population, average daily individual tobacco consumption remains low. AN people are typically light cigarette users – defined as consumption of 10 cigarettes per day (CPD) [6]. Despite low levels of tobacco consumption AN smokers still report difficulty quitting [7], which is similar to reports in African-American light smokers [8]. Notably, AN peoples experience high rates of many tobacco-related diseases. Cancer is the leading cause of AN population mortality [9], with lung cancer the primary cause of cancer death and elevated rates of oral and esophageal cancer incidence compared to the U.S. overall [10, 11].

Nicotine is the compound in tobacco primarily responsible for the addictiveness of its use [12, 13]. *In vivo* the majority of nicotine is inactivated to cotinine (COT), mediated principally by the hepatic enzyme cytochrome P450 2A6 (*CYP2A6*) [14]. Cotinine is then further converted to its proximate metabolite *trans*-3'-hydroxycotinine (3HC), a reaction exclusively conducted by *CYP2A6* [15, 16]. The gene encoding *CYP2A6* is highly polymorphic (<http://www.cypalleles.ki.se/cyp2a6.htm>), and substantial interethnic and interindividual variation in *CYP2A6* allele frequencies and resulting activity exists [17–19].

Variation in *CYP2A6* activity modifies tobacco use behaviors as dependent smokers regulate intake to maintain levels of nicotine in the brain and plasma [20, 21]. Thus, lower rates of nicotine metabolism, due to *CYP2A6* decreased- and loss-of-function alleles, result in decreased nicotine intake in moderate to heavy smokers [22–24]. In order to accurately assess the impact of *CYP2A6* decreased- and loss-of-function alleles on enzyme activity a reliable phenotypic marker is necessary. The ratio of *trans*-3'-hydroxycotinine to cotinine (3HC/COT) is a validated phenotype for *CYP2A6* activity due to exclusive mediation of the conversion of COT to 3HC by *CYP2A6* [16], and the 3HC/COT ratios in plasma and saliva are highly correlated with the rate of *in vivo* nicotine clearance ($r = 0.70$ – 0.95) [16]. The concordance between *CYP2A6* genotype and plasma 3HC/COT has also been demonstrated in a population of light smoking African-Americans, who exhibit a similar pattern of low and irregular tobacco use as AN people [8]. Further, the ratio of free 3HC/free COT in urine strongly correlates with the ratio in plasma ($r = 0.70$) and may also be employed as a biomarker of *CYP2A6* activity [25]. The current study represents the first study of plasma

and urine 3HC/COT as a measure of CYP2A6 activity in ST users, compared to cigarette smokers.

In addition to CYP2A6, CYP2B6 may also metabolize nicotine. CYP2B6 possesses a lower affinity than CYP2A6 for nicotine but may contribute to its metabolism in individuals with low CYP2A6 activity [26]. The *CYP2B6* gene is also highly polymorphic (<http://www.cypalleles.ki.se/cyp2b6.htm>). One prevalent variant allele, *CYP2B6**6, has an allele frequency ranging from 14–62% among different ethnic groups [27]. Studies examining the influence of CYP2B6 on *in vivo* nicotine metabolism have provided inconsistent results. Associations have been observed between the *CYP2B6**6 variant and a faster rate of nicotine clearance [28], as well as the *CYP2B6**4 allele and an elevated plasma 3HC/COT ratio [29]. However, *CYP2B6* genotype had no impact on *in vivo* nicotine plasma level from nicotine patch in treatment-seeking abstinent smokers [30], and *in vitro* analysis demonstrated that correlation of CYP2B6 expression with nicotine C-oxidation was dependent on amount of CYP2A6 [31]. Additionally, *CYP2A6* and *CYP2B6* are both located within the *CYP2ABFGST* gene cluster on chromosome 19q13.2, separated by only 150 kb [32], and significant linkage disequilibrium has been shown between the two genes [33]. It is possible that linkage disequilibrium of *CYP2B6* alleles with *CYP2A6* is responsible for the associations observed between *CYP2B6* and *in vivo* nicotine metabolism phenotypes in some studies.

We investigated *CYP2A6* and *CYP2B6* gene variation in a population of Yupik AN tobacco users and non-users. Distinct increased, decreased, and loss-of-function alleles characteristic of previously studied ethnic populations [19, 34–37] were chosen for comparison of variant allele frequencies. We then determined if *CYP2A6* genotype associated with the plasma and urine 3HC/COT phenotypes derived from *ad libitum* tobacco use; an analysis performed in current users of cigarettes or ST only. We also examined whether *CYP2B6* variation influenced the rates of nicotine metabolism, and sought to clarify whether this effect was through an association with *CYP2A6*. This is the first study examining the impact of *CYP2A6* and *CYP2B6* genetic variation on *in vivo* nicotine metabolism in an AN population.

Methods

Study design

Briefly, individuals ($n = 400$) self-identified as “Alaska Native” were recruited for a cross-sectional study in the Bristol Bay region of southwest Alaska, an area characterized by a high prevalence of cigarette and ST use. Participants were recruited through a variety of means: public radio announcements, written communication to villages from the Bristol Bay Area Health Corporation (BBAHC), posted flyers, and direct VHF radio messages to communities. During recruitment individuals were provided with a brief study description, inclusion criteria, and an estimation of their time required in participating. Subjects were recruited into the following mutually exclusive categories: cigarette smoker – regular user only manufactured cigarettes, commercial chew tobacco user – regular user only commercial chew tobacco, iqmik user – regular user of iqmik for previous 30 days, dual tobacco user – regular user of both cigarettes and commercial chew and/or iqmik, and tobacco non-user – never user (< 100 times) and former user (> 100 times but no use in the past year). Current tobacco users must have used tobacco in the 24 hours prior to recruitment and regularly for the past 30 days. For greater detail on study recruitment refer to Renner et al. (2011 – submitted). AN tribal self-reported participants included Yupik ($n = 361$), Aleut ($n = 19$), Athabascan ($n = 11$), Inupiaq ($n = 4$), Cupik ($n = 1$), Tlingit ($n = 1$), and “Unknown” ($n = 3$). Analysis was restricted to Yupik people only to reduce the impact of population stratification. Recruitment occurred during 41 visits to 16 distinct villages with

an average of 7 people recruited per village visit, minimizing the relatedness of the participants. Thus, we expect that these findings are representative of the total Yupik population although some unidentified intra-population substructure may be a study limitation.

Measurements

Demographic information, health status, dietary practices, tobacco use history, level of nicotine dependence, and attitude towards tobacco products were recorded via a structured interview at enrollment. There were no significant differences in baseline participant characteristics among Yupik tobacco users ($n = 265$) between the subgroups of smokers ($n = 143$), commercial chew users ($n = 74$), iqmik users ($n = 20$), and dual product users ($n = 28$), with the exception of age where commercial chew and iqmik users were significantly older than dual users. Plasma and urine samples were provided at enrolment, and the levels of COT and 3HC in plasma and nicotine and its metabolites in urine were quantified using methods described previously [16, 38]. Using these same methods the intraday coefficients of variation for COT and 3HC at 150 ng/mL were 1.8% and 3.3% respectively, based on six replicates on three different days [39]. While the populations being compared were recruited at different times, the analysis of plasma/urine COT and 3HC were all performed in the same laboratory with appropriate internal and external controls.

Previously in the literature urinary 3HC/COT has been reported as the ratio of free 3HC to free COT [not deconjugated, [25] or the total 3HC to total COT [deconjugated, [40]], however current analyses were performed primarily on the ratio of total (free plus glucuronide-conjugated) 3HC to free COT as CYP2A6 produces all 3HC available, from free COT, some of which has been further glucuronidated. The ratio of total 3HC to free COT is used throughout this paper unless otherwise indicated, as dividing the sum of the total CYP2A6 product (both the free 3HC primary metabolite and the 3HC-glucuronide secondary metabolite) by free COT (the enzymatic substrate) provides a better reflection of enzyme activity than the ratio of free 3HC/free COT or total 3HC/total COT.

CYP2A6 genotyping assays

DNA was extracted from whole blood through use of the G1N70 GenElute miniprep kit purchased from Sigma-Aldrich (St. Louis, MO) according to manufacturer instructions. DNA was successfully genotyped for all 361 Yupik study participants. Established alleles for *CYP2A6* (*CYP2A6**1X2A, *1X2B, *1B, *2, *4, *7, *8, *9, *10, *12, *17, *35) and *CYP2B6* (*CYP2B6**4, *6, *9) were genotyped according to two-step allele-specific PCR and haplotyping assays as previously described [34, 36, 43].

Statistical analyses

Log-transformation of the plasma and urinary 3HC/COT ratios was performed prior to statistical analyses as the ratios were non-normally distributed, and also log-transformed 3HC/COT ratios correlates more strongly with nicotine clearance [44]. The χ^2 test was used to determine Hardy-Weinberg equilibrium and examine the differences in allele frequencies in this sample compared with those of previously reported ethnicities. Univariate analysis of variance, corrected for multiple testing, was used to assess differences in 3HC/COT ratio within non-genetic factors (gender, age, BMI, etc.) between individual *CYP2A6* genotypes, by *CYP2A6* genotype grouping (NM, IM, SM), and among compared non-Yupik ethnicities. Bonferroni's correction was used for *post hoc* analyses. Haplotyping results of *CYP2A6* and *CYP2B6* were produced, and linkage disequilibrium quantified, by the Haploview© family of software [45]. Linear regression models were used to measure the association of *CYP2A6* genotype with the plasma and urinary 3HC/COT ratios while controlling for variables known to also affect this measure.

Results

Comparisons of population *CYP2A6* and *CYP2B6* allele frequencies

Yupik genotype frequencies did not deviate significantly from Hardy-Weinberg equilibrium; in all cases $P > 0.05$. Yupik *CYP2A6* and *CYP2B6* variant allele frequencies were compared to a selected panel of representative values for differing ethnicities including Japanese, Korean, Chinese, African-American, and Caucasian populations. Frequency comparisons for all alleles are depicted in Table 1, and select variants are graphed to illustrate differences in Figure 1 [8, 19, 24, 27, 34, 36, 46]. Yupik allele frequencies were distinct from Japanese values for *CYP2A6**1B, *CYP2A6**7, *CYP2A6**9, and *CYP2B6**6, from Koreans for *CYP2A6**7, *CYP2A6**9, and *CYP2B6**6, and from Chinese for *CYP2A6**4, *CYP2A6**7, *CYP2A6**9, and *CYP2B6**6. Differences in variant allele frequencies were also found when contrasted to African-American values (*CYP2A6**4, *CYP2A6**10, *CYP2A6**17, *CYP2A6**35, *CYP2B6**6) and those of Caucasians (*CYP2A6**2, *CYP2A6**4, *CYP2A6**10, *CYP2A6**12, *CYP2B6**6). Notably, the directions of difference from Yupik participants (i.e. sometimes higher sometimes lower) were not consistent within comparisons to each of the ethnicities, and across the range of measured alleles, indicating the unique allelic pattern of Yupik *CYP2A6* and *CYP2B6* genetic variation.

CYP2A6 genotype association with the 3HC/COT ratio

Strong concordance was found between *CYP2A6* genotype and plasma and urinary 3HC/COT ($n = 265$) (Figure 2, Table 2); Yupik former tobacco users and nonusers ($n = 96$) did not have this smoking-related ratio, and thus were not included in this and further phenotype analysis. As expected, a strong gene-dose effect was observed with prevalent decreased function *CYP2A6* variants (*CYP2A6**9) and loss of function variants (*CYP2A6**4). Further, two copies of the *CYP2A6**4 gene deletion allele caused the characteristic complete loss of activity as indicated by a lack of 3HC production. The *CYP2A6**1B variant has previously been shown to associate with enhanced *CYP2A6* enzyme activity in European American and African American populations [8, 43]. However, Yupik *CYP2A6**1B individuals in the current study did not have significantly higher plasma 3HC/COT ($F(2, 147) = 1.62$, $P = 0.24$) or urinary 3HC/COT ($F(2, 147) = 2.89$, $P = 0.10$), likely due to a low numbers of study participants without one or more copies of this variant.

Association of plasma and urinary 3HC/COT with *CYP2A6* genotype grouping

Participants were categorized according to the following established *CYP2A6* genotype grouping strategy [8, 14, 47]: individuals without *CYP2A6* decreased or loss-of-function variants were defined as normal metabolizers (NM) ($n = 147$), individuals with one copy of a decreased function allele (*CYP2A6**9 and *CYP2A6**12) were defined as intermediate metabolizers (IM) ($n = 31$), and individuals with two decrease-of-function allele copies, or one or more copies of a loss-of-function allele (*CYP2A6**2, *CYP2A6**4, *CYP2A6**10) were grouped as slow metabolizers (SM) ($n = 87$). While in some genetic analyses those who lack activity (i.e. *CYP2A6**4/*4, Table 2) are grouped as poor metabolizers, in this study only 6 poor metabolizers were found. Thus they were grouped together with the slow metabolizers to both increase statistical power and enable comparisons to previous publications where the *CYP2A6* NM, IM, and SM groupings have been used [8, 14, 47]. The *CYP2A6* genotype groupings were found to significantly associate with the plasma 3HC/COT ratios in all Yupik tobacco users ($F(2, 265) = 55.5$ $P < 0.001$), and the subgroups of Yupik smokers ($F(2, 143) = 26.1$ $P < 0.001$) and commercial chew users ($F(2, 74) = 19.6$ $P < 0.001$) (Figure 3a, b, c). Moreover, this relationship in all tobacco users remained significant when controlling for age, gender and BMI ($F(2, 247) = 57.0$ $P < 0.001$).

Association of *CYP2A6* genotype groupings with the urinary 3HC/COT ratio among Yupik people was also determined. A significant relationship was observed between *CYP2A6* genotype and the urinary total 3HC/free COT ratio ($F(2, 265) = 36.4$ $P < 0.001$), which remained significant when controlling for age, gender, and BMI ($F(2, 247) = 12.7$ $P < 0.001$). A significant relationship was also seen for the alternative ratios, the urinary free 3HC/free COT ratio ($F(2, 265) = 26.2$ $P < 0.001$) and the urinary total 3HC/total COT ratio ($F(2, 265) = 34.1$ $P < 0.001$) (Figure 4a, b, c). While the absolute values of the urinary 3HC/COT ratio for each genotype differed when calculated from the different forms of urinary ratios, the impact of *CYP2A6* IM or SM genotype relative to NM remained the same.

CYP2A6-independent variables that impact the 3HC/COT ratio

A strong correlation was observed between the individual 3HC/COT plasma and urine ratios assessed in Yupik tobacco users (Pearson $R = 0.72$, $P < 0.05$). Extensive interindividual variation was found in plasma and urinary 3HC/COT phenotypes, with the proportions varying greater than 650-fold and 730-fold among all tobacco users (range = 0.003–2.059 in plasma, 0.045 – 33.01 in urine), and 14-fold and 26 fold among *CYP2A6* wild-type individuals (range = 0.150 – 2.059 in plasma, 1.25 – 33.01 in urine). Thus, a univariate analysis to discern factors influencing the plasma and urinary 3HC/COT ratios that are independent of variation in the *CYP2A6* gene (age, gender, BMI) was conducted [8, 25]. In order to eliminate confounding by individuals possessing reduced- or loss-of-function *CYP2A6* genotypes this analysis was limited to participants possessing the *CYP2A6**1/*1 genotype ($n = 147$). The plasma 3HC/COT ratio was found to be higher in females ($P < 0.05$), older individuals ($P < 0.001$), and those with lower BMI ($P < 0.05$) (Table 3) as seen previously [8]. The urine 3HC/COT ratio was not impacted by age, gender, BMI or *CYP2B6* genotype.

In the total Yupik tobacco user group ($n = 265$) a significant association with *CYP2B6* genotype (*CYP2B6**1/*1, *CYP2B6**1/*6, *CYP2B6**6/*6 groups) was observed in both plasma ($F(2, 265) = 10.0$ $P < 0.001$) (Figure 5a) and urine ($F(2, 265) = 8.2$ $P < 0.001$). However this effect was lost following control for *CYP2A6* genotype in either measure ($P > 0.05$), suggesting the association of *CYP2B6* with plasma and urine 3HC/COT was predicated on linkage disequilibrium with *CYP2A6* (Figure 5b).

Gene-gene interaction of CYP2A6 and CYP2B6

To better characterize possible interaction between *CYP2A6* and *CYP2B6* genotypes, linkage disequilibrium was assessed through haplotyping analysis, shown in Figure 5c and 5d. The *CYP2B6**6 allele was significantly associated with *CYP2A6**1B; thus given the high percentage of *CYP2A6**1B among Yupik (65.4% allele frequency - Table 1) it is likely that linkage disequilibrium between the wild type *CYP2A6**1B allele and *CYP2B6**6 is responsible for the high plasma 3HC/COT ratio observed in *CYP2B6**6 individuals. In addition, *CYP2A6**7 and *CYP2A6**8 (together *CYP2A6**10) and *CYP2B6**4 and *CYP2A6**9 (together *CYP2B6**6), were found to be in full linkage disequilibrium as these single variants were not found alone in any subject. Associations with *CYP2B6**6 were also observed for *CYP2A6**4 and *CYP2A6**10. Linkage disequilibrium between *CYP2A6**4 and *CYP2B6**6 did not result in the association of elevated plasma 3HC/COT ratio with *CYP2B6**6, as *CYP2A6**4 leads to complete loss of CYP2A6 activity. Further, the low number of Yupik individuals possessing the *CYP2A6**10 variant (2.8% allele frequency – Table 1) indicated it was also unlikely to be responsible for the interaction with *CYP2B6* and 3HC/COT.

Modeling of the plasma and urinary 3HC/COT ratios

To test the impact of individual variables on the plasma and urinary 3HC/COT ratios independently, while controlling for the influence of other factors, a pair of linear regression models including gender, age, BMI, *CYP2B6* genotype, and *CYP2A6* genotype grouping were developed (Table 4). Consistent with univariate analysis female sex ($P < 0.01$) and increasing age ($P < 0.001$) were associated with a higher plasma 3HC/COT ratio, while increasing BMI ($P < 0.001$) and *CYP2A6* genotype subgroup (IM, SM vs. NM) ($P < 0.001$) were associated with a lower 3HC/COT value. BMI ($P < 0.05$) and *CYP2A6* ($P < 0.001$) genotype significantly associated with the urine 3HC/COT ratio as well, while gender and age had no effect. *CYP2B6* genotype had no impact on the 3HC/COT ratio in plasma or urine. The final models accounted for 38% of the total variability in plasma 3HC/COT and 20% of total variability in urine 3HC/COT.

Comparisons of population 3HC/COT ratios

The mean plasma 3HC/COT ratio of the total Yupik tobacco use population ($n = 265$) was higher than the mean population ratios previously assessed in African-Americans and Caucasians [8, 48], reaching significance when compared to the African-American group ($F(2, 1097) = 27.17$ $P < 0.001$) (Figure 6a). To determine whether the elevated Yupik mean population 3HC/COT ratio was due mainly to differences in prevalence of decreased or loss-of-function *CYP2A6* gene variants, a second comparison was conducted in only *CYP2A6**1/*1 individuals (Figure 6b). Among people with this genotype, Yupik people had a significantly higher mean ratio than both Caucasians and African-Americans ($F(2, 640) = 26.63$ $P < 0.001$) with the latter groups having near-equal values among the wild-type individuals.

Discussion

The AN Yupik people occupy a unique position in *CYP2A6* and *CYP2B6* genetics with allele frequencies that differed substantially from other racial/ethnic groups. There was a strong relationship between plasma and urine 3HC/COT ratios derived from *ad libitum* tobacco and *CYP2A6* genotypes among both smokers and ST users. Notably, Yupik possessed a higher mean plasma 3HC/COT ratio than other racial/ethnic groups - suggesting elevated *CYP2A6* activity - which may contribute to their higher prevalence of tobacco use and tobacco-related cancers.

The finding that Yupik *CYP2A6* and *CYP2B6* allele frequencies were distinct from those of previously examined ethnicities indicates the unique profile of the Yupik population for these genes. It was hypothesized that the pattern of allele variation would resemble that of Asian populations based on migration of the ancestral AN population from central Asia through Siberia and across the Bering Strait during the last glacial maximum [49–52]. One possibility for a difference from Asian *CYP2A6* allele frequencies could be admixture of Yupik and Caucasian populations in Alaska, as admixture has been demonstrated in a recent study by examining Y-chromosome variation in Alaska's Aleutian islands [53]. However, the frequencies of Yupik alleles that significantly differ from Chinese do not consistently occupy an intermediary position between Chinese and Caucasians, arguing against this form of admixture. Several other possible explanations for the unique variant allele frequencies among AN also exist: a founder effect of the initial migrants randomly sampled from the larger ancestral group [54, 55], severe genetic bottleneck events (e.g. illnesses which have dramatically decreased the population) during or after migration [56], the greater influence of genetic drift on a small itinerant population [57], and/or novel environmental forces selecting for increased or decreased prevalence of variants [58, 59].

In assessing the influence of *CYP2A6* genetic variants on nicotine metabolism, the 3HC/COT ratio in plasma or urine is a useful phenotype measurement due to its inherent specificity to *CYP2A6* [16] and due to the long half-lives of cotinine (~16h) and *trans*-3'-hydroxycotinine (formation-dependent) [60, 61], consistency across sampling points [39], a low variance over time [62], and regular association with *CYP2A6* genotype irrespective of tobacco consumption level [8, 14, 42]. Results of the current study indicate there is a strong relationship between *CYP2A6* genotype and the 3HC/COT measure among Yupik tobacco users. Variants previously established as decreased activity (ex. *CYP2A6**9) or inactive alleles (ex. *CYP2A6**4) maintain their degree of impact [8, 47], a robust gene-dose effect is observed among prevalent alleles, and these findings are consistent across all tobacco users despite marked differences in the main route of nicotine absorption between tobacco products. Specifically, a significant association of *CYP2A6* genotype with plasma 3HC/COT was observed in ST users for the first time. In examining the different types of ratios which could be utilized in urine, the biologically rational one containing product over substrate (total 3HC/COT) was shown to be highly related to the *CYP2A6* genotype, however both the free/free and total/total were also significantly related, although absolute levels differed. It will be important to confirm this finding (the utility of all three forms of this ratio) in populations with vastly different 3HC or COT rates of glucuronidation.

Consistent with previous data a higher plasma 3HC/COT ratio was observed in females ($P=0.07$) [63] and an increasing mean plasma ratio was observed with increasing age and decreasing BMI [8]. However, only BMI was found to significantly impact the 3HC/COT ratio in urine, differing from variable impact in plasma 3HC/COT as well as previous studies of the urinary ratio [25]. Plasma and urinary 3HC/COT ratios are higher in females likely due to the influence of estrogenic hormones on *CYP2A6* gene transcription [63]. The influence of additional environmental factors, such as glucuronidation rate, urine flow rate, and urine pH may contribute as well [25]. The regression model of urinary 3HC/COT was able to predict 20% of ratio variability, compared to the 38% predictive capability seen in plasma. Additional parameters have a substantial effect on urinary 3HC/COT and identifying them will assist in the utility of this *CYP2A6* activity measurement in future studies.

As the plasma and urinary 3HC/COT ratios are specific phenotype markers for *CYP2A6* activity [14, 16], the discovery in the Yupik cohort of an association between the *CYP2B6**6 allele and these measures was not expected. While *CYP2B6* may play a role in *in vivo* nicotine metabolism, especially in individuals with low *CYP2A6* activity [26], the conversion of 3HC to COT is exclusively mediated by *CYP2A6* [15]. The association of *CYP2B6* genotype with plasma and urine 3HC/COT was removed when controlling for *CYP2A6* genotype, by limiting testing to *CYP2A6* wild-type individuals only, and through regression analysis of 3HC/COT. The association of *CYP2B6**6 with 3HC/COT appears to be explained by its occurrence at a disproportionately high frequency in the *CYP2A6* normal metabolizer subgroup due to linkage disequilibrium with *CYP2A6**1B. This is consistent with prior *in vitro* work demonstrating that correlation of *CYP2B6* protein levels and nicotine C-oxidation activity was due to a correlation of *CYP2B6* and *CYP2A6* protein expression [31].

Although weak linkage disequilibrium between the *CYP2A6* and *CYP2B6* genes has been documented previously in some studies [29, 33], including our own [30]. Data from International Hapmap Project suggests a relatively low level linkage disequilibrium between *CYP2B6**6 and *CYP2A6* variants for Caucasians ($r^2 = 0$), Japanese ($r^2 = 0$), and African Americans ($r^2 = 0$). This demonstrates the uniqueness of the *CYP2A6*-*CYP2B6* association strength in Yupik people. The current findings establish linkage disequilibrium between *CYP2A6* and *CYP2B6* as responsible for a spurious association of *CYP2B6* genotype with

increased *in vivo* nicotine metabolism, and suggest this spurious association may occur in other populations. Thus, we conclude the increased plasma and urine 3HC/COT ratios in Yupik by *CYP2B6**6 genotype is due to the association of this allele with *CYP2A6* wild-type genotype.

The average rate of CYP2A6 activity in the Yupik cohort was higher than those observed in previously measured race/ethnicity. The higher mean plasma 3HC/COT ratio likely does not result from overexpression of known *CYP2A6* gene duplication alleles, as the frequencies of *CYP2A6**1X2A and *CYP2A6**1X2B are negligible in the study group. A portion of the elevated CYP2A6 activity level found in Yupik tobacco users may be explained by a high frequency of the *CYP2A6**1B allele, which has been previously shown to associate with a greater mean 3HC/COT ratio in individuals with at least one copy [8, 43]. Additional factors may underlie the elevated Yupik 3HC/COT ratio such as inheritance of distinctive population gain-of-function *CYP2A6* allele(s), altered rates of COT or 3HC glucuronidation, or the existence of dietary or environmental inducers in the population. AN people consume a unique diet with great seasonal variation in food intake and high levels of smoked meat and fish [64]. Contact with potential CYP inducers may also occur through time spent in fish smokehouses or wood-burning steam baths, routine seasonal sources of environmental smoke exposure [65]. Unique inducers may also exist in the different tobacco products used by this population of Yupik, such as iqmik or hand-rolled cigarettes [4].

High CYP2A6 activity, as suggested by the 3HC/COT ratio, may contribute to the disproportionately high levels of tobacco-related disease in the AN population due to potentially greater tobacco carcinogen activation [66, 67]. Increased CYP2A6 activity may also be a factor in the high prevalence of tobacco use in the population [24].

One of the strengths of this study was our ability to associate *CYP2A6* genotype with CYP2A6 activity phenotype in a large, previously unstudied population. A limitation of this study is that its findings may not be representative of the total AN population, as there are many distinct AN ethnic groups, given the overrepresentation of Yupik and female subjects. In conclusion, this study establishes the unique *CYP2A6* and *CYP2B6* genetic profile of the Yupik and provides further evidence for the relationship between *CYP2A6* genotype and plasma and urine 3HC/COT phenotype markers of CYP2A6 activity in light smoking populations and smokeless tobacco users. Additionally, this work suggests that Yupik people possess an unusually high rate of CYP2A6 activity and a high frequency of *CYP2B6* slow metabolizers. This phenomenon may contribute to the high prevalence of tobacco use in the AN group and disproportionate rates of tobacco-related disease and/or alter response to clinically used drugs such as efavirenz and bupropion which are metabolized by CYP2B6.

Acknowledgments

The scientific team would like to express their gratitude for the leadership and direction from the members of the Board of Directors of the Bristol Bay Area Health Corporation, the members of the Ethics Committee of that organization and the Community Advisory Board for this study, and the BBAHC Director of Community Health Services, Ms. Rose Loera, all who contributed their time and expertise to making this study possible. We would also like to acknowledge contributions of Ms. Kim Hatt, Ms. Helen Peters and Ms. Ana Chartier who were study assistants to the project. In addition, we would like to acknowledge Drs. David Ashley and Tom Bernert for advise on study design and Qian Zhou and Ewa Hoffmann for their assistance with the DNA extraction and genotyping assays. This study was supported NIDA/NCI NARCH III (Indian Health Service Grant) U26IHS30001/01, NCI grant CA114609 and CIHR MOP86471 and was approved by the Health Sciences Research Ethics Board of the University of Toronto - #22398.

Funding Sources:

We also thank CAMH and the CAMH foundation (RFT), NIH grants DA020830 (RFT, NB) and DA012353 (NB), the Canada Foundation for Innovation (#20289 and #16014), and the Ontario Ministry of Research and Innovation (RFT).

References

1. WHO. World Health Organization Report on the Global Tobacco Epidemic, 2008. 2008.
2. CDCP. 2004 Surgeon General's Report - The Health Consequences of Smoking. Department of Health and Human Services: Centers for Disease Control and Prevention. 2004.
3. Hagan, K.; Provost, D. Alaska Native Health Status Report. Alaska Native Epidemiology Center. Anchorage AK: Alaska Native Tribal Health Consortium; 2009.
4. Renner CC, Enoch C, Patten CA, Ebbert JO, Hurt RD, Moyer TP, et al. Iqmik: a form of smokeless tobacco used among Alaska natives. *Am J Health Behav.* 2005; 29:588–594. [PubMed: 16336113]
5. Hurt RD, Renner CC, Patten CA, Ebbert JO, Offord KP, Schroeder DR, et al. Iqmik--a form of smokeless tobacco used by pregnant Alaska natives: nicotine exposure in their neonates. *J Matern Fetal Neonatal Med.* 2005; 17:281–289. [PubMed: 16147838]
6. Smith JJ, Ferucci ED, Dillard DA, Lanier AP. Tobacco use among Alaska Native people in the EARTH study. *Nicotine Tob Res.* 2010; 12:839–844. [PubMed: 20547558]
7. Boles M, Rohde K, He H, Maher JE, Stark MJ, Fenaughty A, et al. Effectiveness of a tobacco quitline in an indigenous population: a comparison between Alaska Native people and other first-time quitline callers who set a quit date. *Int J Circumpolar Health.* 2009; 68:170–181. [PubMed: 19517876]
8. Ho MK, Mwenifumbo JC, Al Koudsi N, Okuyemi KS, Ahluwalia JS, Benowitz NL, et al. Association of nicotine metabolite ratio and CYP2A6 genotype with smoking cessation treatment in African-American light smokers. *Clin Pharmacol Ther.* 2009; 85:635–643. [PubMed: 19279561]
9. Day GE, Lanier AP. Alaska native mortality 1979–1998. *Public Health Rep.* 2003; 118:518–530. [PubMed: 14563909]
10. Lanier AP, Kelly JJ, Maxwell J, McEvoy T, Homan C. Cancer in Alaska Native people 1969–2003. *Alaska Med.* 2006; 48:30–59. [PubMed: 17140152]
11. Lanier AP, Day GE, Kelly JJ, Provost E. Disparities in cancer mortality among Alaska Native people, 1994–2003. *Alaska Med.* 2008; 49:120–125. [PubMed: 18491804]
12. Benowitz NL. Nicotine addiction. *N Engl J Med.* 2010; 362:2295–2303. [PubMed: 20554984]
13. Messina ES, Tyndale RF, Sellers EM. A major role for CYP2A6 in nicotine C-oxidation by human liver microsomes. *J Pharmacol Exp Ther.* 1997; 282:1608–1614. [PubMed: 9316878]
14. Benowitz NL, Swan GE, Jacob P 3rd, Lessov-Schlaggar CN, Tyndale RF. CYP2A6 genotype and the metabolism and disposition kinetics of nicotine. *Clin Pharmacol Ther.* 2006; 80:457–467. [PubMed: 17112802]
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; 277:1010–1015. [PubMed: 8627511]
16. Dempsey D, Tutka P, Jacob P 3rd, Allen F, Schoedel K, Tyndale RF, et al. Nicotine metabolite ratio as an index of cytochrome P450 2A6 metabolic activity. *Clin Pharmacol Ther.* 2004; 76:64–72. [PubMed: 15229465]
17. Benowitz NL, Perez-Stable EJ, Herrera B, Jacob P 3rd. Slower metabolism and reduced intake of nicotine from cigarette smoking in Chinese-Americans. *J Natl Cancer Inst.* 2002; 94:108–115. [PubMed: 11792749]
18. Mwenifumbo JC, Tyndale RF. Genetic variability in CYP2A6 and the pharmacokinetics of nicotine. *Pharmacogenomics.* 2007; 8:1385–1402. [PubMed: 17979512]
19. Nakajima M, Fukami T, Yamanaka H, Higashi E, Sakai H, Yoshida R, et al. Comprehensive evaluation of variability in nicotine metabolism and CYP2A6 polymorphic alleles in four ethnic populations. *Clin Pharmacol Ther.* 2006; 80:282–297. [PubMed: 16952495]
20. Ray R, Schnoll RA, Lerman C. Nicotine dependence: biology, behavior, and treatment. *Annu Rev Med.* 2009; 60:247–260. [PubMed: 19630572]

21. Rao Y, Hoffmann E, Zia M, Bodin L, Zeman M, Sellers EM, et al. Duplications and defects in the CYP2A6 gene: identification, genotyping, and in vivo effects on smoking. *Mol Pharmacol*. 2000; 58:747–755. [PubMed: 10999944]
22. Malaiyandi V, Lerman C, Benowitz NL, Jepson C, Patterson F, Tyndale RF. Impact of CYP2A6 genotype on pretreatment smoking behaviour and nicotine levels from and usage of nicotine replacement therapy. *Mol Psychiatry*. 2006; 11:400–409. [PubMed: 16402128]
23. Liu T, David SP, Tyndale RF, Wang H, Zhou Q, Ding P, et al. Associations of CYP2A6 genotype with smoking behaviors in southern China. *Addiction*. 2011; 106:985–994. [PubMed: 21205058]
24. Schoedel KA, Hoffmann EB, Rao Y, Sellers EM, Tyndale RF. Ethnic variation in CYP2A6 and association of genetically slow nicotine metabolism and smoking in adult Caucasians. *Pharmacogenetics*. 2004; 14:615–626. [PubMed: 15475735]
25. Swan GE, Lessov-Schlaggar CN, Bergen AW, He Y, Tyndale RF, Benowitz NL. Genetic and environmental influences on the ratio of 3'hydroxycotinine to cotinine in plasma and urine. *Pharmacogenet Genomics*. 2009; 19:388–398. [PubMed: 19300303]
26. 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; 22:419–425. [PubMed: 15265511]
27. Zanger UM, Klein K, Saussele T, Blievernicht J, Hofmann MH, Schwab M. Polymorphic CYP2B6: molecular mechanisms and emerging clinical significance. *Pharmacogenomics*. 2007; 8:743–759. [PubMed: 17638512]
28. Ring HZ, Valdes AM, Nishita DM, Prasad S, Jacob P 3rd, Tyndale RF, et al. Gene-gene interactions between CYP2B6 and CYP2A6 in nicotine metabolism. *Pharmacogenet Genomics*. 2007; 17:1007–1015. [PubMed: 18004205]
29. Johnstone E, Benowitz N, Cargill A, Jacob R, Hinks L, Day I, et al. Determinants of the rate of nicotine metabolism and effects on smoking behavior. *Clin Pharmacol Ther*. 2006; 80:319–330. [PubMed: 17015050]
30. Lee AM, Jepson C, Shields PG, Benowitz N, Lerman C, Tyndale RF. CYP2B6 genotype does not alter nicotine metabolism, plasma levels, or abstinence with nicotine replacement therapy. *Cancer Epidemiol Biomarkers Prev*. 2007; 16:1312–1314. [PubMed: 17548706]
31. Al Koudsi N, Tyndale RF. Hepatic CYP2B6 is altered by genetic, physiologic, and environmental factors but plays little role in nicotine metabolism. *Xenobiotica*. 2010; 40:381–392. [PubMed: 20307138]
32. Hoffman SM, Nelson DR, Keeney DS. Organization, structure and evolution of the CYP2 gene cluster on human chromosome 19. *Pharmacogenetics*. 2001; 11:687–698. [PubMed: 11692077]
33. Haberl M, Anwald B, Klein K, Weil R, Fuss C, Gepdiremen A, et al. Three haplotypes associated with CYP2A6 phenotypes in Caucasians. *Pharmacogenet Genomics*. 2005; 15:609–624. [PubMed: 16041240]
34. Al Koudsi N, Ahluwalia JS, Lin SK, Sellers EM, Tyndale RF. A novel CYP2A6 allele (CYP2A6*35) resulting in an amino-acid substitution (Asn438Tyr) is associated with lower CYP2A6 activity in vivo. *Pharmacogenomics J*. 2009; 9:274–282. [PubMed: 19365400]
35. Fukami T, Nakajima M, Yoshida R, Tsuchiya Y, Fujiki Y, Katoh M, et al. A novel polymorphism of human CYP2A6 gene CYP2A6*17 has an amino acid substitution (V365M) that decreases enzymatic activity in vitro and in vivo. *Clin Pharmacol Ther*. 2004; 76:519–527. [PubMed: 15592323]
36. Mwenifumbo JC, Myers MG, Wall TL, Lin SK, Sellers EM, Tyndale RF. Ethnic variation in CYP2A6*7, CYP2A6*8 and CYP2A6*10 as assessed with a novel haplotyping method. *Pharmacogenet Genomics*. 2005; 15:189–192. [PubMed: 15861044]
37. Oscarson M, McLellan RA, Asp V, Ledesma M, Bernal Ruiz ML, Sinues B, et al. Characterization of a novel CYP2A7/CYP2A6 hybrid allele (CYP2A6*12) that causes reduced CYP2A6 activity. *Hum Mutat*. 2002; 20:275–283. [PubMed: 12325023]
38. Xia, Y.; McGuffey, JE.; Wang, L.; Bernert, JT. Analysis of urinary nicotine metabolite profiles by LC atmospheric pressure ionization tandem mass spectrometry; 49th ASMS Conference on Mass Spectrometry and Allied Topics; 2001; Chicago, IL.

39. Lea RA, Dickson S, Benowitz NL. Within-subject variation of the salivary 3HC/COT ratio in regular daily smokers: prospects for estimating CYP2A6 enzyme activity in large-scale surveys of nicotine metabolic rate. *J Anal Toxicol.* 2006; 30:386–389. [PubMed: 16872570]
40. 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; 17:3526–3535. [PubMed: 19029401]
41. Kandel DB, Hu MC, Schaffran C, Udry JR, Benowitz NL. Urine nicotine metabolites and smoking behavior in a multiracial/multiethnic national sample of young adults. *Am J Epidemiol.* 2007; 165:901–910. [PubMed: 17322544]
42. Benowitz NL, Pomerleau OF, Pomerleau CS, Jacob P 3rd. Nicotine metabolite ratio as a predictor of cigarette consumption. *Nicotine Tob Res.* 2003; 5:621–624. [PubMed: 14577978]
43. Mwenifumbo JC, Lessov-Schlaggar CN, Zhou Q, Krasnow RE, Swan GE, Benowitz NL, et al. Identification of novel CYP2A6*1B variants: the CYP2A6*1B allele is associated with faster in vivo nicotine metabolism. *Clin Pharmacol Ther.* 2008; 83:115–121. [PubMed: 17522595]
44. Levi M, Dempsey DA, Benowitz NL, Sheiner LB. Prediction methods for nicotine clearance using cotinine and 3-hydroxy-cotinine spot saliva samples II. Model application. *J Pharmacokinet Pharmacodyn.* 2007; 34:23–34. [PubMed: 17206525]
45. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005; 21:263–265. [PubMed: 15297300]
46. Mwenifumbo JC, Zhou Q, Benowitz NL, Sellers EM, Tyndale RF. New CYP2A6 gene deletion and conversion variants in a population of Black African descent. *Pharmacogenomics.* 2010; 11:189–198. [PubMed: 20136358]
47. Lerman C, Jepsen C, Wileyto EP, Patterson F, Schnoll R, Mroziewicz M, et al. Genetic variation in nicotine metabolism predicts the efficacy of extended-duration transdermal nicotine therapy. *Clin Pharmacol Ther.* 2010; 87:553–557. [PubMed: 20336063]
48. Lerman C, Tyndale R, Patterson F, Wileyto EP, Shields PG, Pinto A, et al. Nicotine metabolite ratio predicts efficacy of transdermal nicotine for smoking cessation. *Clin Pharmacol Ther.* 2006; 79:600–608. [PubMed: 16765148]
49. Goebel T, Waters MR, O'Rourke DH. The late Pleistocene dispersal of modern humans in the Americas. *Science.* 2008; 319:1497–1502. [PubMed: 18339930]
50. Schurr TG. The peopling of the New World: perspectives from molecular anthropology. *Ann Rev Anthropol.* 2004:551–583.
51. Kitchen A, Miyamoto MM, Mulligan CJ. A three-stage colonization model for the peopling of the Americas. *PLoS One.* 2008; 3:e1596. [PubMed: 18270583]
52. Fagundes NJ, Kanitz R, Eckert R, Valls AC, Bogó MR, Salzano FM, et al. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am J Hum Genet.* 2008; 82:583–592. [PubMed: 18313026]
53. Zlojutro M, Rubicz R, Crawford MH. Mitochondrial DNA and Y-chromosome variation in five eastern Aleut communities: evidence for genetic substructure in the Aleut population. *Ann Hum Biol.* 2009; 36:511–526. [PubMed: 19588274]
54. Anderson DG, Gillam JC. Paleoindian colonization of the Americas: Implications from an examination of phisiography, demography, and artifact distribution. *American Antiquity.* 2000; 65:43–66.
55. Templeton AR. The reality and importance of founder speciation in evolution. *Bioessays.* 2008; 30:470–479. [PubMed: 18404703]
56. Crawford MH. Genetic structure of circumpolar populations: a synthesis. *Am J Hum Biol.* 2007; 19:203–217. [PubMed: 17286261]
57. Miller, FP.; Vandome, AF.; McBrewster, J., editors. Genetic drift: Genetic drift. Population bottleneck, founder effect, allopatric speciation, antigenic drift, small population size, hardy-weinberg principle, neutral theory of molecular evolution. Saarbrücken, Germany: Alphascript Publishing; 2009.
58. Bamshad M, Wooding SP. Signatures of natural selection in the human genome. *Nat Rev Genet.* 2003; 4:99–111. [PubMed: 12560807]

59. Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, et al. Detecting recent positive selection in the human genome from haplotype structure. *Nature*. 2002; 419:832–837. [PubMed: 12397357]
60. Benowitz NL, Jacob P 3rd. Metabolism of nicotine to cotinine studied by a dual stable isotope method. *Clin Pharmacol Ther*. 1994; 56:483–493. [PubMed: 7955812]
61. Benowitz NL, Jacob P 3rd. Trans-3'-hydroxycotinine: disposition kinetics, effects and plasma levels during cigarette smoking. *Br J Clin Pharmacol*. 2001; 51:53–59. [PubMed: 11167665]
62. Mooney ME, Li ZZ, Murphy SE, Pentel PR, Le C, Hatsukami DK. Stability of the nicotine metabolite ratio in ad libitum and reducing smokers. *Cancer Epidemiol Biomarkers Prev*. 2008; 17:1396–1400. [PubMed: 18559554]
63. Benowitz NL, Lessov-Schlaggar CN, Swan GE, Jacob P 3rd. Female sex and oral contraceptive use accelerate nicotine metabolism. *Clin Pharmacol Ther*. 2006; 79:480–488. [PubMed: 16678549]
64. Nobmann ED, Ponce R, Mattil C, Devereux R, Dyke B, Ebbesson SO, et al. Dietary intakes vary with age among Eskimo adults of Northwest Alaska in the GOCADAN study 2000–2003. *J Nutr*. 2005; 135:856–862. [PubMed: 15795447]
65. Knudson KJ, Frink L, Hoffman BW, Price TD. Chemical characterization of Arctic soils: activity area analysis in contemporary Yup'ik fish camps using ICP-AES. *Journal of Archaeological Science*. 2004; 31:443–456.
66. Wassenaar CA, Dong Q, Wei Q, Amos CI, Spitz MR, Tyndale RF. Relationship Between CYP2A6 and CHRNA5-CHRNA3-CHRNA4 Variation and Smoking Behaviors and Lung Cancer Risk. *J Natl Cancer Inst*. 2011
67. Rodriguez-Antona C, Gomez A, Karlgren M, Sim SC, Ingelman-Sundberg M. Molecular genetics and epigenetics of the cytochrome P450 gene family and its relevance for cancer risk and treatment. *Hum Genet*. 2010; 127:1–17. [PubMed: 19823875]
68. Yoshida R, Nakajima M, Watanabe Y, Kwon JT, Yokoi T. Genetic polymorphisms in human CYP2A6 gene causing impaired nicotine metabolism. *Br J Clin Pharmacol*. 2002; 54:511–517. [PubMed: 12445030]
69. Fukami T, Nakajima M, Yamanaka H, Fukushima Y, McLeod HL, Yokoi T. A novel duplication type of CYP2A6 gene in African-American population. *Drug Metab Dispos*. 2007; 35:515–520. [PubMed: 17267622]
70. Wall JD, Pritchard JK. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet*. 2003; 4:587–597. [PubMed: 12897771]

Figure 1. Illustration of unique *CYP2A6* and *CYP2B6* allele frequencies among Yupik ($n = 361$)
 $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ when compared to Yupik in individual χ^2 tests.

Figure 2. Association of *CYP2A6* genotype with the a) plasma and b) urine 3HC/COT ratios among Yupik tobacco users ($n = 265$)

CYP2A6 genotypes and their associated plasma and urine ratios; each dot represents the value for one individual and the line represents the mean 3HC/COT ratio of each genotype group. See Table 2 for statistical comparisons between genotype groups.

Figure 3. Association of *CYP2A6* genotype groupings with the plasma 3HC/COT ratio among a) Yupik tobacco users ($n = 265$), and the subgroups of b) Yupik smokers ($n = 143$), and Yupik commercial chew ST users ($n = 74$)

The plasma ratios were significantly associated with *CYP2A6* genotype groups. Mean plasma 3HC/COT values for Yupik tobacco users are (mean ± 95% CI): SM = 0.26 ± 0.03 , IM = 0.44 ± 0.06 , and NM = 0.61 ± 0.05 . Mean plasma 3HC/COT values for Yupik smokers are (mean ± 95% CI): SM = 0.27 ± 0.11 , IM = 0.42 ± 0.09 , and NM = 0.60 ± 0.07 . Mean plasma 3HC/COT values for Yupik commercial chew users are (mean ± 95% CI): SM = 0.27 ± 0.05 , IM = 0.39 ± 0.07 , and NM = 0.63 ± 0.10 . Values were log-transformed prior to statistical analysis as they were not normally distributed. ** $P < 0.01$, *** $P < 0.001$.

Figure 4. Association of *CYP2A6* genotype groupings with the urinary a) Total 3HC/Free COT ratio, b) Free 3HC/Free COT, and c) Total 3HC/Total COT ratio among Yupik tobacco users ($n = 265$)

The urinary ratios were significantly associated with *CYP2A6* genotype groups. Mean urinary Total 3HC/Free COT values are (mean $\pm$ 95% CI): SM = 3.24 ± 0.70 , IM = 4.91 ± 0.89 , and NM = 6.08 ± 0.66 . Mean urinary Free 3HC/Free COT values are (mean $\pm$ 95% CI): SM = 2.75 ± 0.56 , IM = 4.40 ± 0.82 , and NM = 5.20 ± 0.54 . Mean urinary Total 3HC/Total COT values are (mean $\pm$ 95% CI): SM = 1.28 ± 0.19 , IM = 2.02 ± 0.33 , and NM = 2.49 ± 0.22 . Values were log-transformed prior to statistical analysis as they were not normally distributed. *** $P < 0.001$.

Figure 5. The association of *CYP2B6* genotype with plasma 3HC/COT ratio among Yupik tobacco users ($n = 265$)

a) The 3HC/COT ratio significantly associated with *CYP2B6* genotype. Mean 3HC/COT values for each *CYP2B6* genotype group are (mean $\pm$ 95% CI): *CYP2B6**1/*1 = 0.38 ± 0.06 , *CYP2B6**1/*6 = 0.46 ± 0.05 , and *CYP2B6**6/*6 = 0.60 ± 0.08 . Prior to statistical analysis values were log-transformed as they were not normally distributed. The number of individuals is listed on the x -axis. * $P < 0.05$, *** $P < 0.001$. **b)** The majority of variability in the 3HC/COT ratio results from *CYP2A6*, and not *CYP2B6*, genotype. Mean plasma 3HC/COT ratios were stratified by *CYP2A6* and *CYP2B6* genotype. A greater effect on mean plasma 3HC/COT is observed when comparing column size by *CYP2A6* metabolizer grouping along the z -axis than *CYP2B6* genotype along the x -axis. Note no subjects possessed the combination of *CYP2A6*1M and *CYP2B6**6/*6 genotypes. Haplotype analysis of genotyped variants within *CYP2A6* and *CYP2B6* displayed by **c)** D' score and **d)** R^2 value. D' and R^2 values are provided for each pair of SNPs analyzed, ranging from 0

(no linkage disequilibrium) to 100 (Full linkage disequilibrium) for both measures. Haplotype map cells are also shaded based on the strength of corresponding disequilibrium relationships. Note a D' score of 100 indicates that only two or three haplotypes were observed for a pair of SNPs, and a score <100 signifies the occurrence of four haplotypes, indicating historical recombination has occurred. R^2 describes the statistical correlation of SNPs at two loci and a score of 100 indicates just two haplotypes are observed [70]. Tag SNPs were utilized for hybrid alleles *CYP2A6**1*B* ($rs = 41349651$), *4 ($rs = 41349978$), and *12 ($rs = 41356316$).

Figure 6. Yupik possess an elevated rate of CYP2A6 metabolism compared to characterized ethnicities

(a) Plasma 3HC/COT values for each ethnicity were: (mean ± 95% CI): Yupik = 0.48 ± 0.04 , Caucasians = 0.43 ± 0.03 , and African-Americans = 0.35 ± 0.02 . (b) When analysis was limited to *CYP2A6**1/*1 individuals, Yupik possessed a significantly higher plasma 3HC/COT ratio than both Caucasians and African Americans. Plasma 3HC/COT values for the *CYP2A6**1/*1 subgroups within each ethnicity are (mean ± 95% CI): Yupik = 0.61 ± 0.05 , Caucasians = 0.44 ± 0.03 , and African-Americans = 0.43 ± 0.04 . Prior to statistical analysis values were log-transformed as they were not normally distributed. Number of individuals is listed on the x-axis. *** $P < 0.001$.

References: [8, 48]

Total Yupik (n = 361) CYP2A6 and CYP2B6 allele frequencies compared to previously genotyped ethnicities

Table 1

Shading indicates comparator ethnicity allele frequency was significantly different ($P < 0.05$) from Yupik in individual chi-square test.

Genotype	Allele Frequencies (%)					
	Yupik	Japanese	Korean	Chinese	African-American	Caucasian
CYP2A6						
*1X2A	0.0	0.0	0.2	0.4	0.0	0.7
*1X2B	0.0	0.0	0.0	-	3.4	0.0
*1B	65.3	48.4	57.0	51.3	18.2	33.5
*2	0.4	0.0	0.2	0.0	0.9	2.2
*4 ^a	14.5	19.0	10.8	8.5	2.7	1.2
*7	0.0	12.5	9.4	5.7	0.0	0.0
*8	0.0	0.0	0.0	0.0	0.0	0.0
*9	8.9	19.0	19.6	15.7	9.6	7.1
*10	1.9	3.2	4.1	4.3	0.0	0.0
*12	0.4	1.6	0.0	0.0	0.4	2.0
*17	0.0	0.0	0.0	-	8.0	0.0
*35	0.0	0.8	-	0.5	2.9	0.0
CYP2B6						
*4	0.0	6.0	4.5	9.1	0.0	3.9
*6	51.7	18.0	15.9	18.4	32.8	25.6
*9	0.0	0.0	0.0	-	0.0	0.0

^aThe current study employed the CYP2A6*4H assay, which detect the CYP2A6*4A/C, *4D-H variants. Allele frequencies in Japanese, Koreans, Chinese, and Caucasians were found using the CYP2A6*4D assay, which detects *4A/C and *4D only. References: [8, 18, 19, 22, 24, 27, 34, 36, 43, 46, 68, 69]

Table 2
Frequency of CYP2A6 genotypes and their associated plasma and urine 3HC/COT ratios among Yupik tobacco users (*n* = 265)

The mean plasma and urine total 3HC/free COT ratios for each genotype group are listed along with the percentage relative to the CYP2A6 wild-type reference group. As the CYP2A6*1A/*1B and CYP2A6*1B/*1B groups did not significantly differ from the CYP2A6*1A/*1A group, all CYP2A6*1/*1 individuals were included in the CYP2A6 wild-type reference group as done previously [8]. Plasma 3HC/COT was calculated using free 3HC and free COT, and urinary 3HC/COT was calculated using free plus glucuronidated 3HC and free COT. Prior to statistical analysis values were log-transformed as they were not normally distributed. As the CYP2A6*1A/*1B and CYP2A6*1B/*1B groups did not significantly differ from the CYP2A6*1A/*1A group, all CYP2A6*1/*1 individuals composed the reference group as done previously (Ho et al. 2009).

CYP2A6 Allele	Genotype	Observed frequency (<i>n</i>)	Mean Plasma 3HC/COT	SD	% of Wild-type Plasma 3HC/COT	<i>P</i>	Mean Urine 3HC/COT	SD	% of Wild-type Urine 3HC/COT	<i>P</i>
*1B	*1/*1	9	0.53	0.29	86	0.24	4.63	3.03	76	0.10
	*1/*1B	36	0.55	0.32	90		5.41	3.62	89	
	*1B/*1B	102	0.64	0.33	104		6.49	4.20	107	
Wild-type Reference Group	*1/*1	147	0.61	0.32	100	-	6.08	4.03	100	-
*2	*1/*2	1	0.43	-	70	-	6.46	-	105	-
*4	*1/*4	54	0.31	0.13	51	<0.001	3.76	3.50	62	<0.001
	*4/*4	6	0.01	0.00	2		0.08	0.03	<1	
*9	*1/*9	30	0.44	0.17	72	<0.001	4.85	2.45	80	0.02
	*9/*9	2	0.12	0.05	20		1.79	1.72	29	
*10	*1/*10	15	0.36	0.12	58	<0.001	3.87	3.06	64	<0.001
*12	*1/*12	1	0.58	-	95	-	6.58	-	108	-
Compound Heterozygotes	*4/*9	9	0.12	0.08	20	<0.001	1.20	0.97	20	<0.001

Table 3
Factors that influence the plasma and urine 3HC/COT ratios among CYP2A6*1/*1 Yupik tobacco users (n = 147)

	n	Mean Plasma 3HC/COT	SD	P value	Mean Urine 3HC/COT	SD	P value
Gender							
Males	67	0.55	0.24	0.05	5.44	3.20	0.09
Females	80	0.66	0.37		6.62	4.57	
Age							
Younger (< 33)	73	0.52	0.28	< 0.001	5.90	4.54	0.39
Older (≥ 33)	74	0.70	0.34		6.27	3.48	
r		0.31		< 0.001	0.11		0.15
BMI^a							
Low (< 27)	62	0.67	0.35	0.03	6.68	4.76	0.14
High (≥ 27)	76	0.55	0.29		5.58	3.25	
r		-0.17		0.04	-0.17		0.04
CYP2B6 Genotype							
CYP2B6*1/*1	20	0.50	0.27	0.23	5.00	3.18	0.29
CYP2B6*1/*6	65	0.62	0.35		6.31	4.86	
CYP2B6*6/*6	62	0.63	0.32		6.20	3.25	

^a Age was categorized by median split at 33.

^b Body mass index (BMI) was categorized by median split at 27.0; BMI was not available for 10 individuals. r = Spearman's correlation coefficient.

Table 4
Factors that influence the plasma and urine 3HC/COT ratios in Yupik tobacco users ($n = 246$)

a) For total model of plasma 3HC/COT in Yupik tobacco users ($R^2 = 0.38$. b) For total model of urine 3HC/COT in Yupik tobacco users ($R^2 = 0.20$). Analyses were also conducted using the urinary free 3HC/free COT and total 3HC/total COT ratios, in which no changes were seen in the significance level of predictors tested with the ratio above. Regression analysis was performed on log transformed plasma and urine 3HC/COT values as they were not normally distributed. Age and BMI were included as continuous variables. BMI data was not available for 19 individuals.

a) Plasma 3HC/COT – Yupik Tobacco Users			
	Standardized β	95% CI	P value
Gender (Female vs. Male)	0.117	0.003–0.035	0.02
Age (Increasing)	0.248	0.001–0.002	< 0.001
BMI (Increasing)	–0.177	–0.004–0.001	< 0.001
<i>CYP2B6</i> Genotype (<i>CYP2B6</i> *1/*1 $\Rightarrow$ *1/*6 $\Rightarrow$ *6/*6)	–0.007	–0.014–0.012	0.91
<i>CYP2A6</i> Genotype Grouping (NM $\Rightarrow$ IM $\Rightarrow$ SM)	–0.537	–0.062–0.043	< 0.001
b) Urinary 3HC/COT – Yupik Tobacco Users			
	Standardized β	95% CI	P value
Gender (Female vs. Male)	0.057	–0.030–0.89	0.32
Age (Increasing)	0.007	–0.002–0.002	0.90
BMI (Increasing)	–0.129	–0.010–0.001	0.03
<i>CYP2B6</i> Genotype (<i>CYP2B6</i> *1/*1 $\Rightarrow$ *1/*6 $\Rightarrow$ *6/*6)	0.050	–0.025–0.068	0.36
<i>CYP2A6</i> Genotype Grouping (NM $\Rightarrow$ IM $\Rightarrow$ SM)	–0.404	–0.152–0.080	< 0.001