CYP2A6 metabolism in the development of smoking behaviors in young adults

Emily Olsson1, Joseph Bloom2,3, Sarah Bertelsen4, John P Budde2, Naomi Breslau5, Andrew Brooks6, Robert Culverhouse7, Grace Chan8, Li-Shiun Chen2, David Chorlian9, Danielle M Dick10, Howard J Edenberg11, Sarah Hartz2, Dorothy Hatzukami12, Victor M Hesselbrock8, Eric O Johnson13, John R Kramer14, Samuel Kuperman14, Jacquelyn L Meyers9, John Nurnberger15, Bernice Porjesz16, Nancy L Saccone16, Marc A Schuckit17, Jerry Stitzer18, Jay A Tischfield6, John P Rice2, Alison Goate4 & Laura J Bierut2

ABSTRACT

Cytochrome P450 2A6 (CYP2A6) encodes the enzyme responsible for the majority of nicotine metabolism. Previous studies support that slow metabolizers smoke fewer cigarettes once nicotine dependent but provide conflicting results on the role of CYP2A6 in the development of dependence. By focusing on the critical period of young adulthood, this study examines the relationship of CYP2A6 variation and smoking milestones. A total of 1209 European American young adults enrolled in the Collaborative Study on the Genetics of Alcoholism were genotyped for CYP2A6 variants to calculate a previously well-validated metric that estimates nicotine metabolism. This metric was not associated with the transition from never smoking to smoking initiation nor with the transition from initiation to daily smoking (P > 0.4). But among young adults who had become daily smokers (n = 506), decreased metabolism was associated with increased risk of nicotine dependence (P = 0.03) (defined as Fagerström Test for Nicotine Dependence score ≥4). This finding was replicated in the Collaborative Genetic Study of Nicotine Dependence with 335 young adult daily smokers (P = 0.02). Secondary meta-analysis indicated that slow metabolizers had a 53 percent increased odds (OR = 1.53, 95 percent CI 1.11–2.11, P = 0.009) of developing nicotine dependence compared with normal metabolizers. Furthermore, secondary analyses examining four-level response of time to first cigarette after waking (>60, 31–60, 6–30, ≤5 minutes) demonstrated a robust effect of the metabolism metric in Collaborative Study on the Genetics of Alcoholism (P = 0.03) and Collaborative Genetic Study of Nicotine Dependence (P = 0.004), illustrating the important role of this measure of dependence. These findings highlight the complex role of CYP2A6 variation across different developmental stages of smoking behaviors.

Keywords CYP2A6, genetics, nicotine dependence, smoking, young adults.

INTRODUCTION

The development of nicotine dependence requires smoking initiation, conversion from experimental to daily use and finally the development of advanced smoking behaviors (Belsky et al., 2013; Bierut, 2011). Although the majority of adult smokers initiate smoking during adolescence, rates of daily smoking substantially increase...
during young adulthood (1 percent at ages 12–17, 12 percent at ages 18–25, 14 percent at ages 26 or more) (NSDUH, 2015). Furthermore, among those who report current daily smoking, the proportion of individuals who smoke a pack or more a day also dramatically increases with age (12 percent at ages 12–17, 23 percent at ages 18–25 and 33 percent at ages 26 or more) (NSDUH, 2015). Increasing our knowledge of what factors drive some young adults and not others to transition from initiation to daily smoking and then to advanced smoking behaviors is important for effectively preventing the progression toward nicotine dependence.

One genetic factor that plays an important role in the development of smoking behaviors is variation in the gene Cytochrome P450 2A6 (CYP2A6), which encodes the cytochrome P450 enzyme responsible for the majority of oxidation of nicotine to cotinine; this is the primary pathway of nicotine metabolism in humans (Hukkanen et al., 2005). The CYP2A6 locus is highly polymorphic, and alleles with reduced function have been associated with slower rates of nicotine metabolism. Common variants define multiple CYP2A6 haplotypes in individuals of European ancestry (Haberl et al., 2005), and the majority of inter-individual variation in the metabolism of nicotine to cotinine is explained by CYP2A6 genotypes in European Americans (Bloom et al., 2011).

The region on chromosome 19 encompassing CYP2A6 is genome-wide significantly associated with cigarettes per day in large meta-analyses of European ancestry adults (TAG, 2010; Thorgeirsson et al., 2010). Among nicotine dependent adults, several studies demonstrate that slower metabolizers smoke fewer cigarettes per day (Benowitz, 2008; Malaiyandi et al., 2005). This observation is thought to reflect that smokers naturally titrate cigarette consumption to maintain steady nicotine levels.

Studies in youth present conflicting results regarding the effect of nicotine metabolism on the development of nicotine dependence and other smoking behaviors (Audrain-McGovern et al., 2007; Cannon et al., 2016; Chenoweth et al., 2016; Huang et al., 2005; Moolchan et al., 2009; O’Loughlin et al., 2004; Rubinstein et al., 2008; Rubinstein et al., 2013). Some studies suggest that slow nicotine metabolism is associated with an increased risk of nicotine dependence (Chenoweth et al., 2016; O’Loughlin et al., 2004; Rubinstein et al., 2013), possibly reflecting an increased sensitivity to initial nicotine exposure among youth who metabolize nicotine more slowly. In contrast, other studies suggest that slower metabolizers have a decreased risk for dependence and related symptoms (Audrain-McGovern et al., 2007; Rubinstein et al., 2008), paralleling findings in adults regarding reduced heaviness of smoking among slow metabolizers.

Our goal was to investigate the ways in which variation in CYP2A6 relates to the development of smoking behaviors during the critical period of young adulthood in a sample of European Americans. A better understanding of how variation in nicotine metabolism contributes to the acquisition of smoking milestones will add to our fundamental knowledge of the developmental processes that lead to nicotine dependence and has the potential to identify individuals at increased susceptibility during this critical period.

**MATERIALS AND METHODS**

**Primary sample description**

The Collaborative Study on the Genetics of Alcoholism (COGA) is a United States multi-center, family study that aims to identify genes that contribute to alcohol use disorders and related phenotypes (Begleiter et al., 1995). Since 2005, the adolescent and young adult study in COGA has used a longitudinal design to examine the development of substance use disorders in youth from high-risk (defined as recruited through alcohol dependent probands with multiple affected family members) and community comparison families. Members aged 12–22 years were recruited from six sites across the United States and interviewed approximately every 2 years with ongoing data collection. Detailed descriptions of the COGA prospective adolescent and young adult sample have been previously published (Chorlian et al., 2015; Dick et al., 2013).

**Smoking behaviors in COGA**

Assessments were performed using the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA), which gathers detailed information on substance use with high reliability and validity (Bucholz et al., 1994; Bucholz et al., 1995; Hesselbrock et al., 1999). Smoking initiation was evaluated with the question ‘Have you ever smoked a full cigarette?’ Daily smoking was defined as smoking at least 4 days per week for at least a month as performed in previous analyses (Kapoor et al., 2012). This was assessed among individuals who had initiated smoking using the questions ‘When you were smoking regularly, how many days per week did you usually smoke cigarettes?’ and ‘For how long did you smoke this many cigarettes at that rate?’

Among individuals who reported daily smoking (n = 506), several measures of more advanced smoking behaviors were examined in this analysis that focused on the period of heaviest smoking. Time to first cigarette after waking was derived from the question ‘During this period when you were smoking the most, about how many minutes after you woke up did you smoke your first
cigarette? and the four response options are the following: more than 1 hour, 31–60 minutes, 6–30 minutes and within 5 minutes. For the analyses, time to first cigarette was dichotomized into >5 minutes (n = 338, 67 percent) and ≤5 minutes after waking (n = 168, 33 percent). Cigarettes per day was evaluated with the question ‘During the period of time when you were smoking the most, about how many cigarettes did you usually have per day?’ and the four response options are as follows: 10 or fewer, 11–20, 21–30 and 31 or more cigarettes. Cigarettes per day was dichotomized into ≤20 (n = 367, 74 percent) and >20 cigarettes (n = 131, 26 percent) in the analyses as performed in previous studies (Belsky et al., 2013; Broms et al., 2006). A total Fagerström Test for Nicotine Dependence (FTND) score during the heaviest period of smoking was calculated at each interview using responses to these two questions as well as responses to questions assessing the four remaining criteria (Heatherton et al., 1991). For the analyses, nicotine dependence was defined as a FTND score of four or more (n = 306, 61 percent), which is a sensitive and specific cut-off for smoking biomarkers (Huang et al., 2008) and has been used in previous genetic studies (Bierut et al., 2007; Saccone et al., 2009).

Given the longitudinal design of this study, an endorsement of smoking initiation or daily smoking at any interview at age 30 years or younger was used to capture these behaviors during young adulthood. The highest FTND score across available interviews at age 30 years or younger was chosen to capture the lifetime maximum, and time to first cigarette as well as cigarettes per day were set at these interviews.

Genotyping

Bloom et al. (2011) developed a metric based on several genetic variants in CYP2A6 to estimate nicotine metabolism. Cross-validation estimates that this metric predicts approximately 70 percent of the variance in metabolism of orally administered nicotine to cotinine in European Americans (Bloom et al., 2012). Our goal was to use this CYP2A6 metabolism metric to test whether CYP2A6 variation predicts cigarette smoking behaviors in young adulthood.

Five CYP2A6 single nucleotide polymorphisms (SNPs) (rs1801272, rs28399442, rs28399433, rs11371115 and rs28399435) were genotyped using the LGC (Teddington, Middlesex, UK) genomics competitive allele-specific polymerase chain reaction (KASP) (lgcgenomics.com). The CYP2A6 copy number variant (CNV) was genotyped with TaqMan 5’ Nuclease Assays (Life Technologies, Thermo Fisher Scientific, Waltham, MA, USA). The CNV assay was run in duplicate, and genotype calls were made using CopyCaller (Thermo Fisher Scientific, Waltham, MA, USA) software. The program PEDCHECK (O’Connell and Weeks, 1998) was used to examine Mendelian inheritance, and only individuals with no Mendelian inconsistencies were included in the analyzed sample. The metabolism metric was calculated based on the genotypes of the five CYP2A6 SNPs and the CNV using an algorithm described in Table S1 (adapted from Bloom et al., 2012).

A set of 64 ancestry informative markers was genotyped as part of a 96 SNP Biorepository Panel by the Rutgers University Cell and DNA Repository. These markers were used in SNPrelate, a function in R, to assign ancestry groups. HapMap populations were included as reference groups. There was high concordance (98 percent) between self-reported and genetically determined ancestry among European Americans. Only individuals with a genetic ancestry of European American were included in the analysis because the metric was optimized for this population (Bloom et al., 2011).

Primary sample selection

The analysis was restricted to individuals who had reached young adulthood (19 years or older) because we were interested in transitions to daily smoking and advanced smoking behaviors, outcomes that often occur during this time. In the COGA adolescent and young adult study, 1209 European ancestry individuals with a last interview age of 19 years or older were genotyped for the CYP2A6 variants, and participants for the analyses were drawn from this group (Fig. 1). The sample used to analyze daily versus non-daily smokers consisted of 776 individuals who had initiated smoking (64 percent of all subjects). For transitions to advanced smoking behaviors, we focused on the sample of 506 daily smokers (65 percent of initiators, 42 percent of all subjects, described in Table 1).
**Replication COGEND sample**

The Collaborative Genetic Study of Nicotine Dependence (COGEND) is a multi-center case-control study designed to identify genes that contribute to nicotine dependence (Saccone et al., 2007). Community based recruitment enrolled participants ages 25–45 years old. Cases were required to be current smokers and have an FTND score of four or more. Controls were required to have smoked at least 100 cigarettes and have a lifetime maximum FTND score of zero or one. For this analysis, only subjects who self-reported as being of European ancestry were examined (previous analyses using EIGENSTRAT have shown a high correspondence with genetic ancestry groups; Saccone et al., 2009). Genotyping of variants to calculate the metabolism metric in COGEND has been previously described (Bloom et al., 2012). We focused on the subsample of 377 COGEND young adults ages 25–30 that overlapped with the ages of the primary COGA sample. From this group, 335 (89 percent) reported smoking every day or nearly every day for at least 2 months and were considered daily smokers. Replication sample characteristics of these daily smokers are described in Table 1.

**Primary data analysis**

Data were analyzed using the Statistical Analysis System. Logistic regression was used to model dichotomous outcomes of smoking initiation, daily smoking, nicotine

---

**Table 1** Characteristics of primary and replication samples of European American young adults.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>COGA young adult European American daily smokers (n = 506)</th>
<th>Replication: COGEND young adult European American daily smokers (n = 335)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, n (percent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>288 (57 percent)</td>
<td>129 (39 percent)</td>
</tr>
<tr>
<td>Females</td>
<td>218 (43 percent)</td>
<td>206 (61 percent)</td>
</tr>
<tr>
<td>Age at last interview examined, years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± sd</td>
<td>24.4 ± 3.3</td>
<td>27.8 ± 1.7</td>
</tr>
<tr>
<td>Range</td>
<td>19–30</td>
<td>25–30</td>
</tr>
<tr>
<td>No. of interviews examined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± sd</td>
<td>4.0 ± 1.4</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>1–6</td>
<td>1</td>
</tr>
<tr>
<td>Family status, n (percent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>From high-risk families</td>
<td>464 (92 percent)</td>
<td></td>
</tr>
<tr>
<td>From comparison families</td>
<td>42 (8 percent)</td>
<td></td>
</tr>
<tr>
<td>Lifetime DSM-IV alcohol dependence, n (percent)</td>
<td>185 (37 percent)</td>
<td>35 (11 percent)</td>
</tr>
<tr>
<td>No. of extended families</td>
<td>310</td>
<td></td>
</tr>
<tr>
<td>No. of nuclear families (full siblings)</td>
<td>431</td>
<td></td>
</tr>
<tr>
<td>FTND score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± sd</td>
<td>4.2 ± 2.6</td>
<td>3.0 ± 3.3</td>
</tr>
<tr>
<td>Range</td>
<td>0–10</td>
<td>0–10</td>
</tr>
<tr>
<td>Nicotine dependence (FTND ≥ 4), n (percent)</td>
<td>306 (60 percent)</td>
<td>166 (50 percent)</td>
</tr>
<tr>
<td>Time to first cigarette after waking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More than 1 hour</td>
<td>80 (16 percent)</td>
<td>168 (50 percent)</td>
</tr>
<tr>
<td>31–60 minutes</td>
<td>73 (14 percent)</td>
<td>19 (6 percent)</td>
</tr>
<tr>
<td>6–30 minutes</td>
<td>185 (37 percent)</td>
<td>67 (20 percent)</td>
</tr>
<tr>
<td>Within 5 minutes</td>
<td>168 (33 percent)</td>
<td>81 (24 percent)</td>
</tr>
<tr>
<td>Cigarettes per day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 or fewer</td>
<td>182 (36 percent)</td>
<td>171 (51 percent)</td>
</tr>
<tr>
<td>11–20</td>
<td>185 (37 percent)</td>
<td>78 (23 percent)</td>
</tr>
<tr>
<td>21–30</td>
<td>81 (16 percent)</td>
<td>45 (13 percent)</td>
</tr>
<tr>
<td>31 or more</td>
<td>50 (10 percent)</td>
<td>41 (12 percent)</td>
</tr>
<tr>
<td>Metabolism metric*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± sd</td>
<td>0.86 ± 0.07</td>
<td>0.86 ± 0.07</td>
</tr>
<tr>
<td>Range</td>
<td>0.44–0.90</td>
<td>0.44–0.90</td>
</tr>
<tr>
<td>Metabolism status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low (Metric ≤ .85)</td>
<td>134 (26 percent)</td>
<td>103 (31 percent)</td>
</tr>
<tr>
<td>Normal (Metric &gt; .85)</td>
<td>372 (74 percent)</td>
<td>232 (69 percent)</td>
</tr>
</tbody>
</table>

*Distribution of metabolism metric in COGA and COGEND young adult daily smokers provided in Fig. S1. COGA, Collaborative Study on the Genetics of Alcoholism; COGEND, Collaborative Genetic Study of Nicotine Dependence; FTND, Fagerström Test for Nicotine Dependence.
dependence, time to first cigarette and cigarettes per day. In the primary analyses in COGA and COGEND, the continuous metabolism metric, sex, study site and last interview age were included as predictor variables. In COGA, family structure was accounted for using generalized estimating equations via PROC GENMOD. Results from the COGEND replication sample were meta-analyzed with the primary COGA results (Table 2) using a publically available Statistical Analysis System macro (>http://www.hsph.harvard.edu/donna-spiegelman/software/metaanal/). Meta-analyses results were based on fixed effect models to determine the evidence for association within the collected samples. In these analyses, we did not observe heterogeneity between the two studies based on the Q statistic ($P > 0.1$).

Secondary data analyses

Secondary analyses were performed to further explore our primary findings. First, individuals were divided into slow and normal metabolizers using a cut-off of $\leq 0.85$ on the metabolism metric as previously described (Chen et al., 2014). This cut-off captures the lowest quartile of metabolizers, and this dichotomous variable was used in logistic regression models of smoking behaviors. Second, because the majority of the COGA sample was recruited from families at high-risk for alcoholism, the primary analyses examining the continuous metabolism metric and smoking milestones were repeated with the covariate of lifetime DSM-IV alcoholism dependence. Third, after observing an association between the metabolism metric and the time to first cigarette dichotomous variable ($> 5$ and $\leq 5$ minutes), the four-level variable of time to first cigarette after waking ($> 60$, $31–60$, $6–30$, $\leq 5$ minutes) was also investigated in cumulative logistic regression models. These analyses were performed to assess whether the continuous metabolism metric predicted response across the four ordinal categories.

Ethics statement

Institutional review boards at all COGA and COGEND sites approved the study design and the studies were carried out in accordance with the Declaration of Helsinki. Written consent was received from all study participants.

RESULTS

Participant characteristics

Demographic, behavioral and metabolism metric characteristics of the COGA and COGEND samples are presented in Table 1. The primary COGA sample of young adult daily smokers consisted of 506 European American individuals from 431 nuclear families from 310 extended families. The mean age at last interview was 24 years, 43 percent were female, and 92 percent came from families at high-risk for alcoholism, with 37 percent meeting the criteria for lifetime DSM-IV alcohol dependence. Among these daily smokers, 61 percent were nicotine dependent, 33 percent smoked within 5 minutes after waking and 26 percent smoked greater than 20 cigarettes per day (Fig. 1 and Table 1). A total of 26 percent of the young adults were slow metabolizers, and the distribution of the metabolism metric (Fig. S1) was similar to that seen in other samples (Bloom et al., 2012; Chen et al., 2014).

The COGEND replication sample of young adult daily smokers consisted of 335 European Americans with an

| Table 2 | Logistic regression models examining the association of a continuous CYP2A6 metabolism metric and smoking milestones in young adults. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | Metabolism Metric in COGA young adults | Replication: metabolism metric in COGEND young adults | Meta-analysis of results |
| Among all young adults | Beta | SE | $P$-value | Beta | SE | $P$-value | Beta | SE | $P$-value |
| (COGA $n = 1209$) |                   |                   |                   |                   |                   |                   |                   |                   |
| Smoking initiation | 0.46 | 0.97 | 0.63 | – | – | – | – | – |
| Among young adult ever-smokers | – | – | – | – | – | – | – | – |
| (COGA $n = 776$) |                   |                   |                   |                   |                   |                   |                   |                   |
| Daily smoking | –0.92 | 1.16 | 0.42 | – | – | – | – | – |
| Among young adult daily smokers | – | – | – | – | – | – | – | – |
| (COGA $n = 506$; COGEND $n = 335$) |                   |                   |                   |                   |                   |                   |                   |                   |
| Nicotine dependence | 3.49 | 1.62 | 0.03 | 4.36 | 1.86 | 0.02 | 3.86 | 1.21 | 0.002 |
| Smoked five or fewer minutes after waking | 2.44 | 1.34 | 0.07 | 4.63 | 1.82 | 0.01 | 3.21 | 1.07 | 0.003 |
| Smoked more than 20 cigarettes per day | –1.10 | 1.59 | 0.49 | 1.53 | 1.85 | 0.41 | 0.01 | 1.17 | 0.99 |

All models include sex, study site and age of last interview as covariates. Analyses with Collaborative Study on the Genetics of Alcoholism (COGA) were also adjusted for familial clustering.
average age at interview of 28 years and the majority were female (61 percent). Among COGEND young adult daily smokers, 50 percent were nicotine dependent, 24 percent smoked within 5 minutes after waking, 25 percent smoked greater than 20 cigarettes per day and 31 percent were slow metabolizers (distribution in Fig. S1).

CYP2A6 metabolism metric and early smoking behaviors

The continuous CYP2A6 metabolism metric was not associated with smoking initiation (P = 0.63) nor with the development of daily smoking (P = 0.42) in the COGA young adults (Table 2). Of the 270 young adults who initiated smoking but did not transition to daily smoking, essentially all of them (98 percent) failed to develop any of the more advanced smoking behaviors, including nicotine dependence and smoking within 5 minutes after waking. This supports the notion that daily smoking is a prerequisite for the development of advanced smoking behaviors. Therefore, subsequent analyses of advanced smoking milestones focused on the 506 daily smokers.

CYP2A6 metabolism metric and advanced smoking behaviors in daily smokers

The CYP2A6 haplotypes predictive of slower metabolism were associated with an increased risk of nicotine dependence in both the primary COGA and replication COGEND samples of young adult daily smokers (Table 2). In multivariate models adjusting for age, sex and study site, the continuous CYP2A6 metabolism metric had a significant effect in COGA (P = 0.03) and COGEND (P = 0.02), where a slow predicted metabolism was associated with an increased risk of nicotine dependence defined by an FTND score ≥4 (Table 2). Secondary analyses showed that slow metabolizers (defined by a metric of ≤0.85) had a 53 percent increased odds (OR = 1.53, 95 percent CI 1.11–2.11, P = 0.009) of developing nicotine dependence as compared with normal metabolizers (metric > 0.85) in meta-analyses of COGA and COGEND studies (Table S2). Figure 2 illustrates this association by showing that a larger proportion of slow metabolizers in both COGA and COGEND developed nicotine dependence as compared with normal metabolizers.

Consistent with the nicotine dependence results, a lower metabolism metric was associated with an increased risk of smoking within 5 minutes after waking (Table 2, Fig. 2). The continuous CYP2A6 metabolism metric had a trending effect in COGA (P = 0.07) and a significant effect in COGEND (P = 0.01). In secondary meta-analysis, slow metabolizers had a 57 percent increased odds (OR = 1.57, 95 percent CI 1.13–2.18, P = 0.007) of smoking within 5 minutes after waking compared with normal metabolizers (Table S2). The CYP2A6 metabolism metric was not associated with smoking more than 20 cigarettes per day in either sample or meta-analysis (Table 2, Fig. 2). Secondary analyses examining the effect of the metabolism metric on smoking behaviors after controlling for DSM-IV alcohol dependence illustrates similar results (Table S3), supporting that the associations are not dependent on alcoholism status.

Robustness of effect of CYP2A6 metabolism metric on time to first cigarette after waking

Secondary analyses using all four responses of time to first cigarette after waking (>60, 31–60, 6–30, ≤5 minutes) demonstrated a more robust effect of the metabolism metric in both COGA (P = 0.03) and COGEND (P = 0.004) as compared with the dichotomous time to first cigarette (>5 and ≤5 minutes) used in our primary analysis (Table S4 and Table 2, respectively). Figure S2 illustrates that across the four categories, there was an increased proportion of slow metabolizers at shorter times to first cigarette after waking among COGA daily smokers. In COGEND daily smokers, we observed a similar trend, except in the category of 31–60 minutes that only had 19 individuals (6 percent of sample, Table 1).
Taken together, these results support a possible ‘dosage effect’ where predicted slower metabolism was correlated with smoking sooner after waking.

**DISCUSSION**

Young adulthood is a critical developmental period for the progression from initiation to more advanced smoking milestones (NSDUH, 2015). This study links variation in a genome-wide significant gene, CYP2A6, with the development of smoking behaviors in two independent samples of European American young adults. Using specific CYP2A6 polymorphisms, we calculated a nicotine metabolism metric, which has been previously shown to account for approximately 70 percent of the variance in metabolism of orally administered nicotine to cotinine in European Americans (Bloom et al., 2012; Bloom et al., 2011). Our primary finding is that slower nicotine metabolism is associated with an elevated risk of developing nicotine dependence among young adult daily smokers, adding important insight into the role of variation in CYP2A6 across stages of smoking development, as illustrated in Fig. 3.

Despite having an important role in the development of nicotine dependence among daily smokers, variation in CYP2A6 was not associated with smoking initiation nor the progression to daily smoking (step 1 in Fig. 3). Previous twin studies support that environmental influences primarily drive early adolescent nicotine use, and that the role of heritable factors on smoking behaviors increases throughout young adulthood (Kendler et al., 2008; Koopmans et al., 1999). Our results are consistent with this model by providing evidence of a specific gene that impacts the transition from daily smoking to nicotine dependence, without influencing initiation and daily smoking.

The observation that decreased predicted nicotine metabolism is associated with increased risk of nicotine dependence in young adult daily smokers also builds on previous studies conducted in adolescents (step 2 in Fig. 3). O’Loughlin et al. (2004) followed 228 non-dependent smokers in grade seven over approximately 30 months and found that individuals with less active genetic variants in CYP2A6 were more likely to develop nicotine dependence but smoked fewer cigarettes per day once dependent. In a follow-up study examining 421 adolescents who had ever smoked a cigarette, Chenoweth et al. (2016) similarly found that slow metabolism conferred by CYP2A6 variation was associated with increased risk of nicotine dependence in adolescence. Huang et al. (2005) examined variation in CYP2A6 in 1518 adolescents enrolled in a longitudinal study in the United Kingdom and similarly found that individuals with variants associated with slower metabolism were more likely to be current versus former smokers at age 18 years compared with normal metabolizers. Rubinstein et al. (2013) assessed a biomarker of the rate of nicotine metabolism (the nicotine metabolite ratio) in 164 adolescent smokers and found that slower metabolizers showed greater symptoms of

**Figure 3** A theoretical framework of the development of smoking behaviors in relation to CYP2A6, Cytochrome P450 2A6 (CYP2A6) variation. Steps 1 and 2 are supported by this paper and previous studies. Steps 3 and 4 are supported by previous studies (reviewed in Benowitz, 2008; Malayandi et al., 2005)
dependence. Our findings in two independent samples expand on these earlier results by demonstrating that during early young adulthood, when many advanced smoking behaviors develop, slow metabolizers who smoke daily continue to have a greater risk of lifetime dependence.

The increased susceptibility to developing nicotine dependence encountered by youth who are slow metabolizers compared with normal metabolizers has been hypothesized to reflect prolonged exposure to nicotine because of its longer half-life during initial smoking experiences (Chenoweth et al., 2013; Malaiyandi et al., 2005; Rubinstein et al., 2013). Although accumulating evidence supports this role, it is important to note that a few studies show the opposite effect where slow metabolism is associated with decreased risk of smoking behaviors in youth (Audrain-McGovern et al., 2007; Moolchan et al., 2009; Rubinstein et al., 2008). For example, Audrain-McGovern (2007) examined 222 adolescent ever-smokers of European ancestry and found that normal CYP2A6 metabolizers developed symptoms of dependence at a faster rate than slower CYP2A6 metabolizers. Other studies suggest that the increased risk of slower metabolizers for developing nicotine dependence in adolescence disappears by young adulthood (Chenoweth et al., 2016). Cannon et al. (2016) followed 296 participants across ages 16–24 years and found that using a CYP2A6 diplotype predictive metric, intermediate metabolism compared with slow and normal was a risk factor for smoking frequency and nicotine dependence. By the end of the study at age 24 years, however, the individuals with predicted normal metabolism were at greatest risk for these smoking behaviors. Many possible explanations exist for these discrepant results. One hypothesis is that the effect of slower nicotine metabolism transitions rapidly from increasing risk to being protective (Cannon et al., 2016 O’Loughlin et al., 2004), and previous studies may have observed different developmental periods in the fast-changing early course of smoking behaviors. The ascertainment of subjects and baseline smoking behaviors also varies across studies, which may influence the power to detect associations. Furthermore, previous studies use different measures of smoking behaviors and nicotine dependence, and it is possible that they capture different components of dependence that are differentially influenced by CYP2A6 metabolism.

Our results suggest that time to first cigarette after waking is a critical contributor to the association between the CYP2A6 metabolism metric and development of nicotine dependence assessed by the FTND criteria among daily smokers. Little consensus exists on the best measure of nicotine dependence, but research supports that two items from the FTND score, time to first cigarette after waking and cigarettes smoked per day, are strong, valid, reliable predictors of quitting behaviors, which are key indicators of dependence (Baker et al., 2007; Borland et al., 2010; Hyland et al., 2006). Studies also suggest that these two measures are distinct predictors of addiction (Borland et al., 2010; Lessov et al., 2004), chronic obstructive pulmonary disease (Guertin et al., 2015) and lung cancer (Gu et al., 2014), suggesting the possibility that different genetic factors may contribute to urgency to smoke and levels of cigarette consumption. In a sample of over 1000 young adults, Haberstick et al. (2007) found that time to first cigarette was the most informative measure of heritable factors from the FTND score. Our results complement these findings by illustrating that necessity to smoke measured by time to first cigarette after waking at least partly drives the association of the CYP2A6 metabolism metric and nicotine dependence in young adult daily smokers. Although the physiologic mechanism underlying this association remains unknown, slow metabolizer daily smokers likely have more consistent nicotine levels throughout the day compared with fast metabolizer daily smokers, which may contribute to a feeling of greater necessity to smoke in the morning when nicotine levels are low.

These findings in young adults should be considered in the context of the literature about adult smoking. Previous studies of adults demonstrate that once dependent, slower metabolizers smoke fewer cigarettes to reach target blood nicotine levels (Benowitz, 2008) (step 3 in Fig. 3). Although we did not observe an effect of slow metabolism on risk of smoking more than 20 cigarettes per day among daily smokers (Table 2, Fig. 2), only 26 percent of these young adults were heavy smokers, and heaviness of smoking continues to increase throughout adulthood (NSDUH, 2015). In the entire COGEND sample ages 25–45 years, a previous analysis demonstrated that among nicotine dependent smokers, slower metabolism is associated with decreased cigarette consumption (Bloom et al., 2012). It is possible that slow metabolism is primarily protective at high levels of cigarette consumption, which is most evident in older populations of adults. Overall, these findings underscore that variation in CYP2A6 has a variety of effects on smoking behaviors across stages of development: slow metabolism leads to increased risk for developing nicotine dependence in young adult daily smokers through time to first cigarette, but once dependent, slow metabolism is protective against heaviness of smoking.

Another important consideration is that the fraction of slow metabolizers in the population of smokers has been observed to decrease with age, suggesting that slow metabolizers are more likely to quit smoking (Benowitz, 2008) (step 4 in Fig. 3). In the COGEND dataset, among current nicotine dependent smokers ages 25–30 years,
we found that 36 percent (60/166) were slow metabolizers. However, among current nicotine dependent smokers over 30 years old, only 28 percent (250/883) were slow metabolizers (Chi-square, \( P = 0.04 \)), supporting that proportionally more slow metabolizers have quit by this time. Furthermore, other studies directly support that slow nicotine metabolism, measured by CYP2A6 genotypes or the nicotine metabolite ratio, is associated with increased cessation rates in both youth (Chenoweth et al., 2013) and adults enrolled in clinical trials (Chen et al., 2014; Ray et al., 2009). Taken together, these findings suggest that across development, slow metabolizers may quit smoking more easily. Therefore, the observation that slow metabolism is associated with increased risk of nicotine dependence may be most pronounced in samples of youth when symptoms of dependence are first developing and before cessation attempts occur.

The findings reported here have limitations. First, this study focused on individuals of European Ancestry because the metabolism metric was optimized for this population (Bloom et al., 2011). Second, the precise timing of length of transitions between smoking behaviors could not be examined in these analyses because the smoking questions did not assess age of onset. Third, the majority of the COGEND participants were from families at high risk for alcoholism and rates of DSM-IV alcohol dependence are high in this sample, which may affect the generalizability of the findings. Secondary analyses that include DSM-IV alcohol dependence as a covariate and replication of the primary findings in a community-based recruitment sample (COGEND), however, support the conclusion that the findings are not specific to a high-risk population.

In summary, using a validated CYP2A6 metabolism metric, this study demonstrates that slower nicotine metabolism is associated with an increased risk of nicotine dependence in two independent samples of young adult daily smokers. These findings add important knowledge about the complex role of CYP2A6 variation across different developmental stages of smoking behaviors.

Acknowledgements

COGEND was supported by U10AA008401 (NIAAA) and COGEND was supported by P01CA089392 (NCT). EO was supported by T32GM07200 (NIGMS), UL1TR000448 (NCATS), TL1TR000449 (NCATS) and F30AA023685 (NIAAA). SJH was supported by K08DA032680 (NIDA), LJB was supported by R01DA032680 (NIDA) and P30CA091842 (NCI).

Disclosure/Conflict of Interest

LJB, AG, and the spouse of NLS are listed as inventors on Issued U.S. Patent 8,080,371, ‘Markers for Addiction’ covering the use of certain SNPs in determining the diagnosis, prognosis, and treatment of addiction. JN is an investigator for Assurex and an investigator and consultant for Janssen. The other authors declare no conflict of interest.

Authors Contributions

LJB, AG, JPR, DMD, HJE, VMH, JRK, SK, JN, BP, MAS and JAT contributed to the conception and design of COGA. LJB, AG, JPR, EOJ, NLS, NB, DH and JS contributed to the conception and design of COGEND. AB and JAT managed the DNA biorepository. JPB performed the genotyping. SB cleaned the genetic data. EO performed the statistical analyses. All authors assisted with analysis design and interpretation of findings. EO and LJB drafted the manuscript. All authors critically reviewed the manuscript, provided important intellectual feedback and approved the manuscript.

References


SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

Table S1. Predicted metabolism metric and CYP2A6 diplotypes based on copy number and 5