Nicotine & Tobacco Research, 2016, 196–203
doi:10.1093/ntr/ntv049
Original investigation
Advance Access publication March 5, 2015

Original investigation

CYP2A6 Longitudinal Effects in Young Smokers

Dale S. Cannon PhD[1], Tait R. Medina MA[2,3], Robin J. Mermelstein PhD[4], Donald Hedeker PhD[5], Amanda V. Bakian PhD[1], Hilary Coon PhD[1], Edwin H. Cook MD[6], Cindy Hamil BS[7], Robert B. Weiss PhD[7]

[1]Department of Psychiatry, University of Utah School of Medicine, Salt Lake City, UT; [2]Department of Sociology, Indiana University, Bloomington, IN; [3]Institute for Health Research and Policy, University of Illinois, Chicago, IL; [4]Institute for Health Research and Policy and Psychology Department, University of Illinois, Chicago, IL; [5]Department of Health Sciences, University of Chicago, Chicago, IL; [6]Institute for Juvenile Research, Department of Psychiatry, University of Illinois, Chicago, IL; [7]Department of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT

Corresponding Author: Dale S. Cannon, PhD, Department of Psychiatry, University of Utah School of Medicine, 650 Komas Drive, Suite 206, Salt Lake City, UT 84108-3528, USA. Telephone: 801-587-3776; Fax: 801-585-5723; E-mail: dale.cannon@hsc.utah.edu

Abstract

Introduction: The present study sought to identify time-dependent within-participant effects of *CYP2A6* genotypes on smoking frequency and nicotine dependence in young smokers.
Methods: Predicted nicotine metabolic rate based on *CYP2A6* diplotypes (*CYP2A6* diplotype predicted rate [CDPR]) was partitioned into Normal, Intermediate, and Slow categories using a metabolism metric. Growth-curve models characterized baseline and longitudinal CDPR effects with data from eight longitudinal assessments during a 6-year period (from approximately age 16–22) in young smokers of European descent (N = 296, 57% female) who had smoked less than 100 cigarettes lifetime at baseline and more than that amount by Year 6. Phenotypes were number of days smoked during the previous 30 days and a youth version of the Nicotine Dependence Syndrome Scale (NDSS). A zero-inflated Poisson growth-curve model was used to account for the preponderance of zero days smoked.
Results: At baseline, Intermediate CDPR was a risk factor relative to both Normal and Slow CDPR for smoking frequency and the NDSS. Slow CDPR was associated with the highest probability of smoking discontinuation at baseline. However, due to CDPR time trend differences, by young adulthood these baseline effects had been reordered such that the greatest risks for smoking frequency and the NDSS were associated with Normal CDPR.
Conclusions: Reduced metabolism *CYP2A6* genotypes are associated with both risk and protective effects in novice smokers. However, differences in the time-by-CDPR effects result in a reordering of genotype effects such that normal metabolism becomes the risk variant by young adulthood, as has been reliably reported in older smokers.

Introduction

CYP2A6, the major nicotine metabolism gene,[1,2] may pose different risk liabilities at different points in post-initiation smoking. In nicotine dependent adult smokers of European ancestry, slower metabolizer *CYP2A6* genotypes reliably are protective.[3,4] However, in young smokers of European ancestry, the findings have been inconsistent. Some studies of young smokers have found slower metabolizer genotypes to be protective,[5,6] but others find them to be associated with higher risk.[7–9] Rubinstein et al.,[10] using the nicotine metabolite ratio, found that slower metabolic rate was associated with higher cigarettes per day and greater Nicotine Dependence (ND) in adolescents.

One interpretation of the discrepant results in young smokers is the hypothesis that the effect of slower nicotine metabolism transitions rapidly from risk to protective over early stages of smoking progression.[8,9] This hypothesis would account for the conflicting results of cross-sectional studies in young smokers if those studies observed different points in the fast-changing early course of smoking. It also would account for the discrepancy between the risk effect of slow metabolism in some youth studies versus the reliable finding that it is protective in adults with ND. An appropriate test of this hypothesis would be an analysis of longitudinal within-participant *CYP2A6* effects, which would eliminate cohort differences inherent in the comparison of cross-sectional studies. Further, a longitudinal analysis would explicitly model time effects to estimate the effects of reduced metabolism relative to normal metabolism over the period in which the unexpected risk of slow metabolism is thought to occur. Although results of cross-sectional analyses conducted at different times using youth cohorts followed longitudinally have been reported,[9] only two of the adolescent studies cited above report within-participant temporal effects.[5,8] Those studies reached different conclusions regarding the relation between reduced metabolism alleles and the acquisition of ND. Audrain-McGovern et al.[5] found a protective effect using a latent growth-curve model but O'Loughlin et al.[8] found a risk effect using a survival analysis. Neither study observed a robust reversal in the direction of the effect of reduced metabolism alleles, although O'Loughlin et al.[8] reported a nonsignificant protective trend once participants acquired ND.

The aim of the present study is to use longitudinal analyses to identify time-dependent within-participant *CYP2A6* effects on smoking frequency and ND over the critical period for smoking uptake during adolescence and young adulthood in which smoking progresses from very low smoking quantity/frequency (Q/F) prior to smoking 100 cigarettes lifetime to higher Q/F subsequent to that lifetime cigarette exposure threshold. We sought to characterize the relative effects of normal, intermediate, and slow metabolism alleles over time using growth-curve models with data from eight longitudinal assessments during a 6-year period (from approximately age 16–22). Growth-curve models are informative for the question under consideration in that they provide estimates of initial *CYP2A6* effects and the rate of phenotypic change associated with *CYP2A6* variants over time. Differences in time trends might lead to a reordering of relative genetic effects over time, that is, a transition for a given allele from protective to risk or vice versa relative to another allele. Our prediction from previous studies[7–10] is that, relative to normal metabolic rate, slower metabolism is associated with greater risk of smoking frequency and ND at baseline. Further, based on the risk pattern found later in smoking history,[4,11] we predict a steeper time trend for normal metabolism that eventually reverses the ordering of *CYP2A6* effects on smoking.

Methods

Participants

Participants included in the present analyses (N = 296) were from the genetic component of the Social and Emotional Contexts of Adolescent Smoking Patterns study.[12–14] All were of European ancestry as determined by a cluster analysis of 64 ancestry informative markers (Supplementary methods and Figures S-1 and S-2). Ninety-two percent (92%) self-identified as non-Hispanic white, 3% self-identified as Hispanic white, and 5% self-identified as being of other race/ethnicity ancestries. Fifty-seven percent (57%) were female. Mean age at baseline was 15.6 years (SD = 0.61).

Social and Emotional Contexts of Adolescent Smoking Patterns is a longitudinal study of smoking in 1263 participants of multiple race/ethnicity ancestries recruited as 9th and 10th graders from Chicago area schools. Students who had ever smoked but were not yet daily smokers were oversampled, which is appropriate as we are looking for genetic markers for smoking progression rather than for smoking initiation or dependence severity.[12]

Assessments were made at baseline, 6, 15, 24, and 33 months, and then annually from Year 4 on. The last wave for which data are available at this writing was Year 6. During the Year 5 or 6 assessments, participants were invited into the genetic arm of the project, and 1019 (81% of the original cohort) participated.

To identify participants who progressed from very light, infrequent smoking to heavier, more frequent smoking, two smoking phenotype selection criteria were used: (1) participants had to have smoked at least a puff but fewer than 100 cigarettes lifetime at baseline and (2) they had to have smoked at least 100 cigarettes lifetime by Year 6 (cf. Supplementary methods for further discussion of these two selection criteria). Because smoking Q/F and ND symptoms are positively associated in Social and Emotional Contexts of Adolescent Smoking Patterns with having met the 100 cigarettes lifetime threshold, these criteria selected participants who were very novice smokers at baseline line but who progressed to higher smoking Q/F 6 years later (Table 1). For example, there were no daily smokers at baseline whereas 31% were daily smokers at Year 6.

CYP2A6 Genotyping

Saliva samples were collected with Oragene OG-500 kits (DNA Genotek, Ontario, Canada) under the supervision of the field study team, and genomic DNA was purified as previously described.[12] *CYP2A6* SNPs were genotyped with TaqMan assays (Life Technologies) for rs1801272 (NM_000762.5:c.479T>A, A allele = *2), rs28399433 (NM_000762.5:c.-48T>G, G allele = *9) and rs1137115 (NM_000762.5:c.51A>G, A allele = *1A(51A)). The allelic state of rs28399435 (NM_000762.5: c.86G>A, A allele = *14) was tested only in *1A(51) homozygotes using polymerase chain reaction (PCR) amplification and Sanger dideoxy terminator sequencing (ABI 3730 capillary sequencer) of *CYP2A6* exon 1 (chr19:41356151-41356352). Copy number variants were genotyped using TaqMan Copy Number assays with Hs07545275_cn detecting *CYP2A6* intron 1 loss (deleted in both *CYP2A6*4 and *CYP2A6*12) and Hs07545274_cn detecting *CYP2A6* intron 7 loss (deleted only in *CYP2A6*4). Copy number assays were performed as duplex real-time PCR reactions on the QuantStudio 12 Flex Real-Time PCR system using the RPPH1 gene as the reference assay. *CYP2A6* copy number was calculated using CopyCaller v2.0 software (Life Technologies) from four technical replicates for each sample. Allele counts and frequencies are shown in Supplementary Table S-1 and diplotype counts and frequencies in Supplementary Table S-2. Diplotype counts did not deviate substantially from the null hypothesis of Hardy–Weinberg equilibrium using a multi-allelic exact conditional test (likelihood ratio P value = .01) implemented in the ExactoHW software package.[15]

Phenotypes

Days Smoked

DAYS, the measure of smoking frequency, was the number of days during the past 30 days on which the participant "smoked or tried cigarettes." DAYS was chosen in preference to cigarettes per day

Table 1. Days Smoked, Percent Smoking Zero Days, and NDSS by CDPR and Years Since Baseline. DAYS is the Mean of all Reported Smoking Frequencies During the Previous 30 Days, Including Reports of Zero Days

Years since BL	DAYS MN (*SD*)			% smoking zero days			NDSS MN (*SD*)		
	Slow	Intermediate	Normal	Slow	Intermediate	Normal	Slow	Intermediate	Normal
0	1.31 (2.57)	3.17 (5.40)	2.00 (3.03)	48.50	49.89	50.04	1.10 (0.37)	1.26 (0.52)	1.10 (0.37)
0.5	3.19 (6.81)	3.89 (5.63)	3.95 (6.72)	49.68	49.14	50.11	1.20 (0.55)	1.30 (0.71)	1.18 (0.57)
1.25	5.32 (9.18)	8.60 (11.21)	6.57 (10.02)	50.61	49.49	49.50	1.35 (0.74)	1.39 (0.82)	1.39 (0.76)
2	8.19 (10.57)	10.34 (12.02)	8.73 (11.28)	50.09	47.85	49.16	1.56 (0.89)	1.53 (0.81)	1.53 (0.81)
2.75	10.86 (11.75)	12.08 (12.21)	12.23 (12.54)	46.79	44.79	46.19			
4	11.41 (12.62)	11.20 (11.96)	14.89 (12.53)	43.48	46.86	40.96	2.08 (0.88)	2.05 (0.98)	2.15 (0.84)
5	15.71 (12.27)	14.54 (12.60)	16.57 (12.38)	35.42	41.85	40.88	2.07 (0.83)	1.98 (0.93)	2.15 (0.90)
6	14.95 (12.75)	12.12 (11.76)	16.05 (12.90)	39.74	44.79	40.68	1.93 (0.82)	1.87 (0.82)	2.18 (0.91)

BL = baseline; MN = mean; NDSS = Nicotine Dependence Syndrome Scale.

because, past 30 days smoking is the more commonly used measure of adolescent smoking and may better represent nondaily smoking, which is the norm among adolescent light smokers.

Nicotine Dependence Syndrome Scale

The youth-specific version of the Nicotine Dependence Syndrome Scale (NDSS), a multidimensional measure of ND, was administered.[16] Beyond quantity measures of smoking, the NDSS assesses components of ND (primarily in this version Drive/Tolerance) and is predictive of progression to daily smoking.[13]

CYP2A6 Diplotype Predicted Rate

Estimated nicotine metabolic rate was derived from a predicted metabolism metric based on *CYP2A6* genotype[17,18] (cf., Supplementary methods, Table S-2 and Figure S-3). The metric was partitioned into three levels, which we denominated the *CYP2A6* diplotype predicted rate (CDPR). If the predicted metabolism metric less than 0.79, then CDPR = Slow (N = 42 [14%]); if 0.79 ≤ metric less than 0.87, then CDPR = Intermediate (N = 52 [18%]); and if metric ≥ 0.87, then CDPR = Normal (N = 202, 68%). In this partitioning of the metric, the null *CYP2A6* alleles *2 and *4, the likely null allele *12,[18] as well as the homozygous *9/*9 diplotype were assigned to Slow CDPR. Intermediate CDPR comprised *9 heterozygotes absent a null allele, and the *1A(51A) homozygote. Sequencing of *CYP2A6* exon 1 confirmed that all 14 *1A(51A) homozygotes were rs28399435 G/G homozygotes as well; therefore, none of the *1A(51A) homozygotes were reclassified to *1(51A)/*14 diplotypes and Normal CDPR. Normal CDPR comprised the Normal homozygote, the Normal/*1A(51A) diplotype, and one instance of *1A(51A)/*1X2.

Two other allele groupings were considered. In one, Intermediate and Slow CDPR as defined above were collapsed into a single category to facilitate comparison of our results with a previous study that used a metabolic rate dichotomy.[5] In the other, the *1A(51A) homozygote was moved to normal CDPR to permit determination of whether the *9 heterozygote absent a null allele was sufficient to observe a baseline risk effect for Intermediate CDPR.

Analytic Strategy

A zero-inflated Poisson (ZIP) growth-curve model was used to model change over time in DAYS and a linear growth-curve model was used to model change over time in NDSS. Zero-inflated models[19,20] such as ZIP are appropriate when the outcome is a count variable with many zeros, as is the case with DAYS (Table 1). A ZIP model assumes that zero counts are generated via two processes. The first generates only zero counts, and the second is a Poisson process that generates both zero and positive counts.[21] Here we attribute zero counts in the first process to smoking discontinuation and zero counts in the second process to random variability in days smoked. A ZIP model accounts for these two processes by fitting two regressions simultaneously—a logistic regression of the probability of smoking discontinuation and a Poisson regression of smoking frequency conditional on discontinuation.

Age at baseline (centered on its grand mean of 15.6), gender (1 = male, 0 = female) and self-reported race/ethnicity (1 = non-Hispanic white, 0 = any other self-reported race/ethnicity) were included as controls. Preliminary analyses found no evidence of a gender by CDPR interaction, indicating that it was unnecessary to stratify analyses by gender. CDPR was included as a set of dummy variables (ie, Slow CDPR and Normal CDPR), with Intermediate CDPR as the reference category. Time was coded as years since baseline (ie, 0, 0.5, 1.25, 2, 2.75, 4, 5, and 6). To allow the time trend to vary by CDPR, each CDPR dummy variable was interacted with time.

To illustrate time trends, predicted values were calculated for each category of CDPR at each time using estimates from the growth-curve models. Predicted values were calculated with controls held at their mean and then plotted by time. To test CDPR differences at each time, a series of Wald tests were calculated using estimates from the growth-curve model. Since all possible pairwise contrasts of the CDPR effect were tested, a Bonferonni correction was applied and a P value less than or equal to .017 was considered statistically significant (P value = .05/3 = .017).

Because of the computational intensity of nonlinear growth models, only a random intercept for the ZIP growth curve model was included.[22] For the linear growth model, both a random intercept and trend factor were included, and residual variances were freely estimated over time.[23] The sample size for the DAYS analysis was 296 subjects with 2274 observations over eight waves and the sample size for the NDSS analysis was 212 subjects with 1374 observations over seven waves. The NDSS was not administered at Year 2.75. To be included in analyses participants had to meet the two smoking phenotype selection criteria discussed earlier and have nonmissing values on the outcome at the first and last waves of data collection. Growth-curve models were estimated in Mplus Version 7.11 using Full Information Maximum Likelihood.[24] Parameter estimates, standard errors (*SE*s), and P values for the Poisson and

logistic portions of the ZIP growth-curve model predicting DAYS and the linear growth-curve model predicting NDSS are given in Supplementary Table S-3. Plots of predicted values are given in Figures 1–3. Results from Wald tests of pairwise contrasts of the CDPR effect at each time for DAYS are given in Table 2. Results of additional analyses are provided in Supplementary Tables S-4 and S-5. Here we present only the results of primary interest.

Results

Descriptive statistics for DAYS, smoking discontinuation, and NDSS by CDPR category and by wave are shown in Table 1.

DAYS

Smoking Frequency

Significant CDPR effects were found at baseline for the smoking frequency portion of the ZIP growth-curve model predicting DAYS (Supplementary Table S-3, Poisson). Both Slow and Normal CDPR, as compared to Intermediate CDPR, were associated with smoking fewer days per month on average at baseline, net of controls (B = –.139; SE = .059, P < .05, and B = –.152; SE = .042, P = 2.6E–04 for Slow and Normal, respectively). While an expected positive main effect for time was found (B = .128; SE = .008, P = 2.8E–52), of main importance is the significant time by Normal CDPR interaction (B = .044; SE = .009, P = 3.0E–06), which indicates that the positive time trend was more pronounced for Normal CDPR as compared to Intermediate.

Figure 1 plots the expected number of days smoked over time for each category of CDPR using the estimates from the Poisson portion of the ZIP growth-curve model. While Intermediate CDPR was associated with smoking more days on average at baseline, the steeper time trend associated with Normal CDPR resulted in a reordering of the CDPR categories by the end of the observation period such that the predicted number of days smoked was highest for Normal CDPR, followed by Intermediate, and then Slow. The by-time Wald tests of all pairwise contrasts (Table 2, Poisson) indicate that the risk effect of Intermediate CDPR for smoking frequency was short lived, that is, through Year 1.25 (mean age = 16.9 years). Normal CDPR became a risk factor for smoking frequency relative to Slow CDPR beginning in Year 2.75 (mean age = 18.4 years) and relative to Intermediate CDPR by Year 6 (mean age = 22.5 years). The only significant control in the Poisson portion was gender, with males predicted to smoke more frequently than females (B = .091, SE = .022, P = 3.9E–05).

Smoking Discontinuation

Significant CDPR effects were found at baseline for the smoking discontinuation portion of the ZIP growth-curve model predicting DAYS (Supplementary Table S-3, Logistic). Slow CDPR, as compared to Normal CDPR, significantly increased the probability of discontinuation, net of controls (B = .885; SE = .263, P = 7.6E–04). A significant negative main effect for time (B = –.131; SE = .055, P < .05) and a significant time by CDPR interaction was found with the downward trend in the probability of discontinuation significantly more negative for Slow CDPR (B = –.269; SE = .086, P < .005) and marginally more negative for Normal CDPR (B = –.115; SE = .062, P < .10) as compared to Intermediate CDPR.

Figure 2 plots the predicted probability of discontinuation over time for all levels of CDPR using estimates from the ZIP growth-curve model. At baseline, the predicted probability of discontinuation was highest for Slow CDPR, the steeper decline associated with Slow CDPR, however, resulted in a reordering of the CDPR categories. By-time tests of all pairwise contrasts of the CDPR effect (Table 2, Logistic) indicate that the protective effect of Slow CDPR on smoking discontinuation lasted only through Year 1.25 (mean age = 16.9 years). The only significant control in the logistic portion was age at baseline, indicating that adolescents who were older at baseline had a higher probability of smoking discontinuation (B = .158, SE = .078, P = 4.2E–02).

In an additional DAYS analysis in which only *9 heterozygotes absent null alleles were included in Intermediate CDPR and *1A(51A) homozygotes were coded as Normal CDPR, intercept effects net of controls relative to Intermediate CDPR were retained for both Normal CDPR (B = –.264; SE = .052, P = 3.9E–07) and Slow

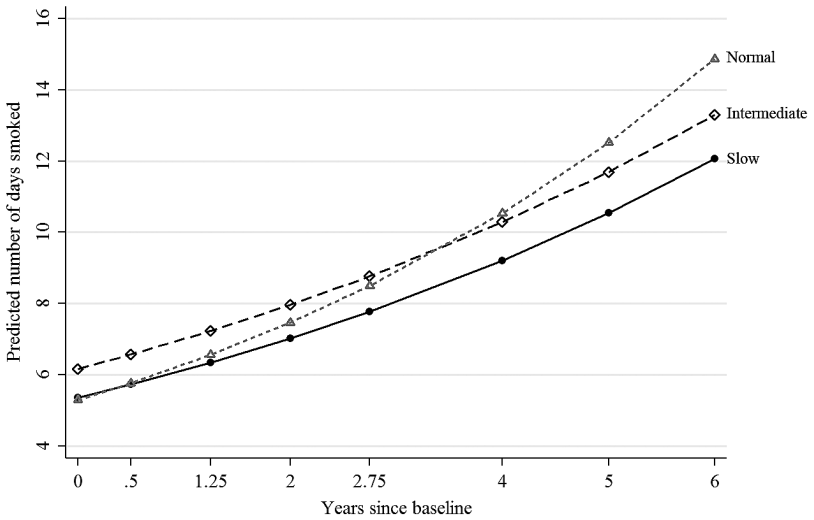

Figure 1. Predicted number of days smoked by *CYP2A6* diplotype predicted rate (CDPR) and time calculated using estimates from Poisson portion of the zero-inflated Poisson growth-curve model predicting DAYS (Supplementary Table S-3, Poisson). Solid line with circle represents Slow CDPR, long-dashed line with diamond represents Intermediate CDPR and short-dashed line with triangle represents Normal CDPR. Predicted number of days smoked was calculated with controls held at their mean value.

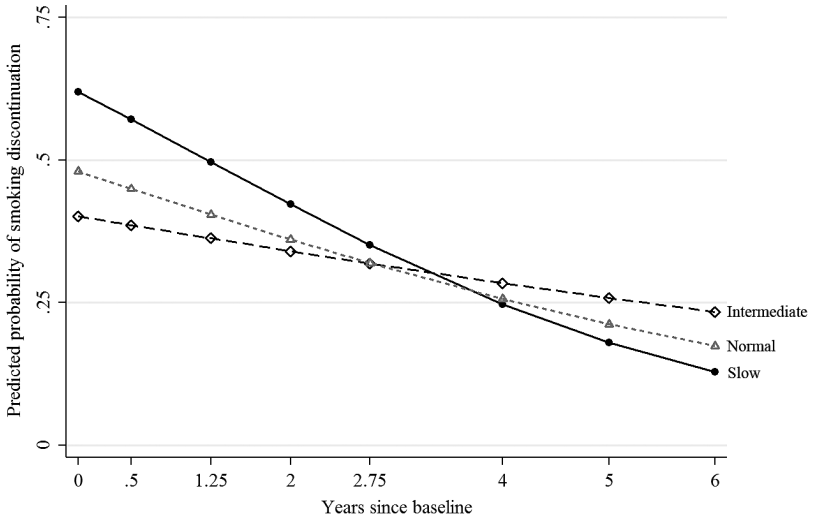

Figure 2. Predicted probability of smoking discontinuation by *CYP2A6* diplotype predicted rate (CDPR) and time calculated using estimates from the logistic portion of the zero-inflated Poisson growth-curve model predicting DAYS (Supplementary Table S-3, Logistic). Solid line with circle represents Slow CDPR, long-dashed line with diamond represents Intermediate CDPR and short-dashed line with triangle represents Normal CDPR. Predicted probability of smoking discontinuation was calculated with controls held at their mean value.

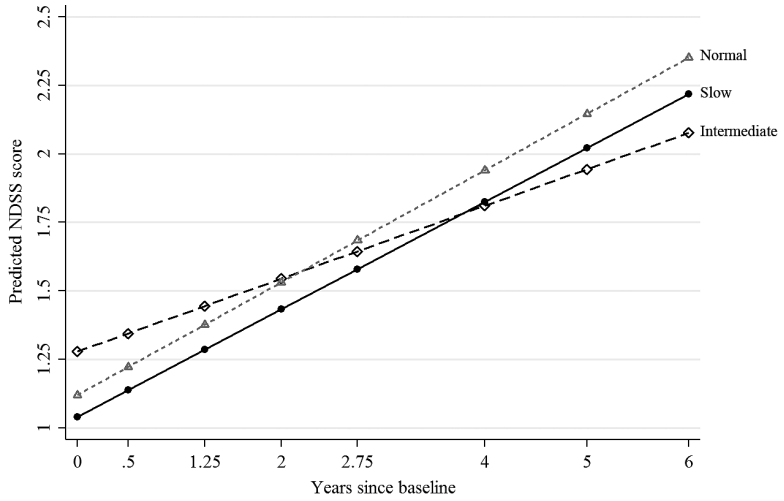

Figure 3. Predicted Nicotine Dependence Syndrome Scale (NDSS) score by *CYP2A6* diplotype predicted rate (CDPR) and time calculated using estimates from the linear growth-curve model predicting NDSS (Supplementary Table S-3, NDSS). Solid line with circle represents Slow CDPR, long-dashed line with diamond represents Intermediate CDPR and short-dashed line with triangle represents Normal CDPR. Predicted NDSS score was calculated with controls held at their mean value.

CDPR (B = –.216; SE = .069, P = 1.7E–03) and there was a steeper time trend for Normal CDPR (B = .058, SE = .01, P = 2.1E–08) (cf., Supplementary Table S-4). When Intermediate and Slow CDPR were collapsed into one category in a DAYS analysis, there was no main effect of CDPR but the time trend for Normal CDPR was steeper relative to that for the combined group, (B = .038, SE = .008, P = 3.2E–06) (cf. Supplementary Table S-5).

NDSS Findings

The substantive NDSS findings largely parallel the smoking frequency portion of the DAYS analyses. Again there were significant CDPR effects at baseline, with Slow and Normal CDPRs associated with less ND than Intermediate CDPR, net of controls (B = .238; SE = .107, P < .05, and B = –.158; SE = .078, P < .05, respectively). A significant positive main effect for time (B = .133; SE = .026, P = 4.5E–07) and a significant time by CDPR interaction for Normal CDPR (B = .072; SE = .029, P < .05) were found. Figure 3 plots the predicted NDSS score over time by CDPR category and illustrates a similar reordering of the categories as was observed with the smoking frequency portion of DAYS.

Discussion

The results of longitudinal analyses suggest a complex set of relations among smoking phenotypes, nicotine metabolic rate as estimated by *CYP2A6* variants, and time. The growth-curve models elegantly characterize CDPR effects over time as baseline effects and rate of change effects. Intermediate CDPR initially was a risk factor relative to both Normal and Slow CDPR for smoking frequency and for the NDSS. For smoking discontinuation, however, the differential effect initially was associated with Slow CDPR, relative to both Normal and Intermediate CDPR. That is, Slow CDPR was associated with

Table 2. Bonferonni Corrected P Values From Wald Tests of CDPR Contrasts at Each Time Using Estimates From ZIP Growth-Curve Model Predicting DAYS

	Smoking frequency (Poisson)			Smoking discontinuation (Logistic)		
Years since BL	Intermediate vs. Normal	Intermediate vs. Slow	Normal vs. Slow	Intermediate vs. Normal	Intermediate vs. Slow	Normal vs. Slow
0	**2.6E–04**	1.8E–02	8.0E–01	1.1E–01	**7.6E–04**	**8.3E–03**
0.5	**6.5E–04**	**1.1E–02**	9.0E–01	1.3E–01	**1.1E–03**	**9.3E–03**
1.25	**3.7E–03**	**5.4E–03**	4.0E–01	2.2E–01	**4.1E–03**	**1.6E–02**
2	3.1E–02	**2.5E–03**	8.5E–02	4.7E–01	3.9E–02	6.2E–02
2.75	2.6E–01	**1.5E–03**	**6.2E–03**	9.7E–01	3.9E–01	3.2E–01
4	3.7E–01	**2.8E–03**	**2.2E–05**	3.7E–01	3.8E–01	7.8E–01
5	2.5E–02	**1.2E–02**	**1.1E–06**	2.0E–01	1.0E–01	3.9E–01
6	**1.8E–03**	4.8E–02	**5.1E–07**	1.4E–01	3.9E–02	2.3E–01

BL = baseline; CDPR = *CYP2A6* diplotype predicted rate; ZIP = zero-inflated Poisson. A bold *P* value indicates a significant difference between CDPR categories at the .017-level. Since all CPDR contrasts were tested, a Bonferonni correction was applied (*P* value = .05/3 = .017). For each test the degrees of freedom equal one. See Supplementary Table S-3 for the parameter estimates, *SE*s, and *P* values for the ZIP growth-curve model predicting DAYS.

the highest probability of discontinuation at baseline, but a steeper negative time trend led to the disappearance of this effect within 2 years. Thus, at the earliest time periods surveyed, Slow CDPR was protective by one estimate while Intermediate CDPR was a risk by two others. However, the Normal CDPR time trends for the NDSS and for smoking frequency were steeper so that by young adulthood it became the risk allele, as it is in studies of older adults.

Although estimates may be specific to our cohort and methods, the more important conclusion we draw from these within-participant findings is that different metabolic rates are associated with different trajectories of smoking progression. *CYP2A6* alleles may be associated with different initial effects on smoking as well as different rates of smoking progression, resulting in a given allele being protective at one point and a risk at another point, relative to other alleles.

The intriguing theoretical questions raised by these results are why smoking enhancement by intermediate metabolism occurs at all initially, and why it does not persist relative to normal metabolism. Among time-dependent variables that might influence the relative effects of different nicotine metabolic rates are amount of lifetime cigarette exposure, smoking Q/F, and the development of ND. These smoking phenotypes all increased over the course of the study in a highly-correlated fashion (data not shown), and no time-dependent variable can be isolated as causative. Another time-dependent variable not assessed here that might interact with nicotine metabolic rate is brain development.[25,26]

Changes in *CYP2A6* effects may be attributable to the relative shift in smoking motivation from pre-ND smoking for nicotine's positive reinforcement effects to post-ND smoking to avoid nicotine withdrawal.[26–29] Perhaps the interpretation of our results is as simple as the possibility that low nicotine dose in nondependent novice smokers is optimally appetitive for smokers with intermediate nicotine metabolic rate whereas, subsequent to ND acquisition, smokers with normal nicotine metabolic rate require higher nicotine intake to avoid withdrawal. If so, the low NDSS scores observed at Year 6 (Table 1), when Normal CDPR was a risk factor for smoking frequency relative to the other two CDPR categories, suggest ND severity does not have to be great in order to observe the pattern of *CYP2A6* effects reported in adult samples with greater ND. Our results shed no light on the role of appetitive nicotine effects in novice smokers. Clearly, further research would be required to support the interpretation suggested here.

If it is generally the case that both protective and risk effects of reduced metabolism occur initially, depending on the phenotype and degree of nicotine metabolism reduction, but that differential time effects eventually lead to the ordering of protective/risk effects found in adults, then clearly the ordering of effects in a cross-sectional study will depend on the phenotype, the point in smoking progression at which the cross-sectional data were obtained, and how alleles are coded. There are several points of agreement with two earlier studies. First, the initial enhancement of smoking frequency and the NDSS by Intermediate CDPR in this study was observed at about the same age and smoking Q/F as that of the sample in the Rubinstein et al. study,[10] which found an enhancement of cigarettes per day and ND by slow metabolism. In the Rubinstein study, age = 16.1 and cigarettes per day *M* = 2.86, *SD* = 3.35, median = 1.78 (compare Year 0, Table 1). Thus, both studies found smoking enhancement effects in young, very novice smokers. Second, our NDSS results are consistent with those of Audrain-McGovern[5] with respect to a steeper time trend for ND development associated with normal metabolic rate. Those investigators did not find any intercept (baseline) effects, but they classified all reduced metabolism variants as slow. We observed baseline NDSS enhancement in the Intermediate but not the Slow category. Thus, it is possible that combining the reduced metabolism variants obscured an enhancement effect in the Audrain-McGovern et al. study.

Our NDSS findings are not consistent with a previous report that slow metabolism increases ND acquisition over time.[8] That study used three metabolic rate categories similar to ours, but it found the slowest rate to be associated with increased risk of ND acquisition. Our ND measure (the NDSS) was more similar to that of Audrain-McGovern (the mFTQ) than that of O'Loughlin (ICD-9 diagnostic criteria dichotomized into ND vs. non-ND), and both Audrain-McGovern and the present study used growth-curve model analysis while O'Louglin used survival analysis. It is unclear whether the discrepancy between the O'Loughlin study and the other two studies is due to one of these methodological differences or to cohort differences. Finally, we found evidence that the *9 heterozygote absent null alleles is at least sufficient, if not necessary, for an enhancement effect. No other reduced metabolism allele was frequent enough to test its effects independently.

Limitations

The number of participants included in our analyses was small due to our selection criteria for initial levels of smoking and progression, as well as the restriction to European ancestry. However, having up to eight observations per participant increased power. The advantage of a longitudinal design such as ours over cross-sectional analyses is that differences in *CYP2A6* effects at different time points cannot be

attributed to between-cohort sampling error.[9] Another limitation is that the results are generalizable only to young smokers of European ancestry who progress from very light, infrequent smoking during adolescence to heavier, more frequent smoking in young adulthood. This was not a population study of all adolescents, but rather captured those at elevated risk of smoking progression. It is not known whether the general *CYP2A6* effects identified here would generalize to other phenotypes, such as smoking cessation.[6]

Conclusions

Based on observed time-dependent within-participant effects, we conclude that reduced activity *CYP2A6* variants have both protective and risk effects in very novice smokers who progress to more frequent smoking by young adulthood. Specifically, variants associated with intermediate metabolic rate (particularly Normal/*9) enhance smoking frequency and ND relative to both normal and slower rate variants, while slower rate is associated with higher probability of smoking discontinuation. By young adulthood, the risk effect of intermediate rate disappears in favor of the pattern of rate differences robustly observed in heavy smoking adults, that is, a positive association between nicotine metabolic rate and risk. Further research is necessary to identify which of a host of time-dependent parameters may account for this shift in intermediate rate from a risk to a protective effect, but we propose that the transition in nicotine metabolism rate effects is in part a function of different smoking motives at different stages of smoking progression.

Supplementary Material

Supplementary methods, Tables S1 to S5, and Figures S1 to S3 can be found online at http://www.ntr.oxfordjournals.org

Funding

This work was supported by the National Cancer Institute at the National Institutes of Health (P01CA98262).

Declaration of Interests

None declared.

Acknowledgments

We gratefully acknowledge the invaluable contributions to this work of Kathleen Diviak and the Participant Interaction Core of the Social and Emotional Contexts of Adolescent Smoking Patterns Study.

References

1. Benowitz NL. Pharmacology of nicotine: addiction, smoking-induced disease, and therapeutics. *Annu Rev Pharmacol Toxicol*. 2009;49:57–71. doi:10.1146/annurev.pharmtox.48.113006.094742.
2. Malaiyandi V, Sellers EM, Tyndale RF. Implications of CYP2A6 genetic variation for smoking behaviors and nicotine dependence. *Clin Pharmacol Ther*. 2005;77(3):145–158. doi:10.1016/j.clpt.2004.10.011.
3. Ray R, Tyndale RF, Lerman C. Nicotine dependence pharmacogenetics: role of genetic variation in nicotine-metabolizing enzymes. *J Neurogenet*. 2009;23(3):252–261. doi:10.1080/01677060802572887.
4. Thorgeirsson TE, Gudbjartsson DF, Surakka I, et al. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet*. 2010;42(5):448–453. doi:10.1038/ng.573.
5. Audrain-McGovern J, Al Koudsi N, Rodriguez D, Wileyto EP, Shields PG, Tyndale RF. The role of CYP2A6 in the emergence of nicotine dependence in adolescents. *Pediatrics*. 2007;119(1):e264–e274. doi:10.1542/peds.2006-1583.
6. Chenoweth MJ, O'Loughlin J, Sylvestre MP, Tyndale RF. CYP2A6 slow nicotine metabolism is associated with increased quitting by adolescent smokers. *Pharmacogenet Genomics*. 2013;23(4):232–235. doi:10.1097/FPC.0b013e32835f834d.
7. Huang S, Cook DG, Hinks LJ, et al. CYP2A6, MAOA, DBH, DRD4, and 5HT2A genotypes, smoking behaviour and cotinine levels in 1518 UK adolescents. *Pharmacogenet Genomics*. 2005;15(12):839–850. doi:01213011-200512000-00002.
8. O'Loughlin J, Paradis G, Kim W, et al. Genetically decreased CYP2A6 and the risk of tobacco dependence: a prospective study of novice smokers. *Tob Control*. 2004;13(4):422–428. doi:10.1136/tc.2003.007070.
9. Rodriguez S, Cook DG, Gaunt TR, Nightingale CM, Whincup PH, Day IN. Combined analysis of CHRNA5, CHRNA3 and CYP2A6 in relation to adolescent smoking behaviour. *J Psychopharmacol*. 2011;25(7):915–923. doi:10.1177/0269881111405352.
10. Rubinstein ML, Shiffman S, Moscicki AB, Rait MA, Sen S, Benowitz NL. Nicotine metabolism and addiction among adolescent smokers. *Addiction*. 2013;108(2):406–412. doi:10.1111/j.1360-0443.2012.04026.x.
11. Kumasaka N, Aoki M, Okada Y, et al. Haplotypes with copy number and single nucleotide polymorphisms in CYP2A6 locus are associated with smoking quantity in a Japanese population. *PLoS One*. 2012;7(9):e44507. doi:10.1371/journal.pone.0044507.
12. Cannon DS, Mermelstein RJ, Hedeker D, et al. Effect of neuronal nicotinic acetylcholine receptor genes (CHRN) on longitudinal cigarettes per day in adolescents and young adults. *Nicotine Tob Res*. 2014;16(2):137–144. doi:10.1093/ntr/ntt125.
13. Dierker L, Mermelstein R. Early emerging nicotine-dependence symptoms: a signal of propensity for chronic smoking behavior in adolescents. *J Pediatr*. 2010;156(5):818–822. doi:10.1016/j.jpeds.2009.11.044.
14. Selya AS, Dierker LC, Rose JS, Hedeker D, Mermelstein RJ. Risk factors for adolescent smoking: parental smoking and the mediating role of nicotine dependence. *Drug Alcohol Depend*. 2012;124(3):311–318. doi:10.1016/j.drugalcdep.2012.02.004.
15. Engels WR. Exact tests for Hardy-Weinberg proportions. *Genetics*. 2009;183(4):1431–1441. doi:10.1534/genetics.109.108977.
16. Sterling KL, Mermelstein R, Turner L, Diviak K, Flay B, Shiffman S. Examining the psychometric properties and predictive validity of a youth-specific version of the Nicotine Dependence Syndrome Scale (NDSS) among teens with varying levels of smoking. *Addict Behav*. 2009;34(6–7):616–619. doi:10.1016/j.addbeh.2009.03.016.
17. Bloom AJ, Harari O, Martinez M, et al. Use of a predictive model derived from in vivo endophenotype measurements to demonstrate associations with a complex locus, CYP2A6. *Hum Mol Genet*. 2012;21(13):3050–3062. doi:10.1093/hmg/dds114.
18. Bloom J, Hinrichs AL, Wang JC, et al. The contribution of common CYP2A6 alleles to variation in nicotine metabolism among European-Americans. *Pharmacogenet Genomics*. 2011;21(7):403–416. doi:10.1097/FPC.0b013e328346e8c0.
19. Greene WH. Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. Working Paper EC-94-10. 1994:1–36. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1293115. Accessed November 3, 2014.
20. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*. 1992;34(1):1–14. http://dl.acm.org/citation.cfm?id=149270. Accessed November 3, 2014.
21. Long JS. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage; 1997.
22. Hall DB. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics*. 2000;56(4):1030–1039. www.ncbi.nlm.nih.gov/pubmed/11129458. Accessed November 3, 2014.

23. Muthén BO, Curran PJ. General longitudinal modeling of individual differences in experimental designs: a latent variable framework for analysis and power estimation. *Psychol Methods*. 1997;24(4):371. doi:10.1037/1082-989X.2.4.371.
24. Muthén LK, Muthén BO. *Mplus User's Guide*. 7th ed. Los Angeles, CA: Muthén & Muthén; 2012.
25. Dwyer JB, McQuown SC, Leslie FM. The dynamic effects of nicotine on the developing brain. *Pharmacol Ther*. 2009;122(2):125–139. doi:10.1016/j.pharmthera.2009.02.003.
26. Lydon DM, Wilson SJ, Child A, Geier CF. Adolescent brain maturation and smoking: what we know and where we're headed. *Neurosci Biobehav Rev*. 2014;45:323–342. doi:10.1016/j.neubiorev.2014.07.003.
27. Baker TB, Piper ME, McCarthy DE, Majeskie MR, Fiore MC. Addiction motivation reformulated: an affective processing model of negative reinforcement. *Psychol Rev*. 2004;111(1):33–51. doi:10.1037/0033-295X.111.1.33.
28. Mathew AR, Wahlquist AE, Garrett-Mayer E, Gray KM, Saladin ME, Carpenter MJ. Affective motives for smoking among early stage smokers. *Nicotine Tob Res*. 2014;16(10):1387–1393. doi:10.1093/ntr/ntu093.
29. Piper ME, Bolt DM, Kim SY, et al. Refining the tobacco dependence phenotype using the Wisconsin Inventory of Smoking Dependence Motives. *J Abnorm Psychol*. 2008;117(4):747–761. doi:10.1037/a0013298.