Published in final edited form as:

Addiction. 2012 September ; 107(9): 1696–1704. doi:10.1111/j.1360-0443.2012.03871.x.

Genetic Influences on Developmental Smoking Trajectories

Christina N. Lessov-Schlaggar, Sean D. Kristjansson, Kathleen K. Bucholz, Andrew C. Heath, and Pamela A. F. Madden

Department of Psychiatry, Washington University School of Medicine, St. Louis, MO

Abstract

Aims—To investigate the relative contribution of genetic and environmental factors on smoking trajectory membership and to test whether individual smoking trajectories represent phenotypic thresholds of increasing genetic risk along a common genetic liability dimension.

Design—Prospective study of a birth cohort of female like-sex twin pairs.

Setting—Participants completed diagnostic interview surveys 4 times from adolescence (average age 16) through young adulthood (average age 25).

Participants—Female twins who had smoked 100 cigarettes lifetime (n=1466 regular smokers).

Measurements—Number of cigarettes smoked per day during the heaviest period of smoking (2 waves) or during the past 12 months (2 waves).

Findings—A 4-trajectory class solution provided the best fit to cigarette consumption data and was characterized by Low (n=564, 38.47%), Moderate (n=366, 24.97%), and High level smokers (n=197, 13.44%), and smokers who increased their smoking from adolescence to young adulthood (n=339, 23.12%). The best genetic model fit was a 3-category model that comprised the Low, a combined Increasing + Moderate, and High trajectories. This trajectory categorization was heritable (72.7%) with no evidence for significant contribution from shared environmental factors.

Conclusions—The way that smoking patterns develop in adolescence has a high level of heritability.

Keywords

smoking trajectories; twins; heritability

Introduction

Empirical identification of developmental smoking trajectories has advanced understanding of how the smoking habit unfolds from adolescence through young adulthood and has helped define subgroups of smokers who differ in the level and rate of progression of cigarette smoking. Prospective investigations have consistently shown the existence of three to five smoking subgroups (excluding never smokers), comprised of persistent low level smokers; early smoking initiators with rapid progression to moderate or heavy persistent smoking; later smoking initiators; persistent heavy smokers; and early initiators who quit or

Corresponding Author: Christina N. Lessov-Schlaggar, PhD, 660 S. Euclid, Campus Box 8134, Saint Louis, MO 63110, office (314) 747-1450, fax (314) 286-2213, schlaggc@psychiatry.wustl.edu.

Declaration of Interest: The study was supported by National Institutes of Health grants AA009022, AA017688, AA017915, AA011998, AA12640, HD049024, DA012854, DA14363, and DA027046. None of the authors have any connection to the tobacco, alcohol, pharmaceutical or gaming industries.

reduce their smoking by later adolescence/early adulthood [1–6]. Correlates of developmental smoking trajectories include parent demographics and characteristics of the home environment (e.g., parental attitude toward smoking, smoking rules inside the home), as well as race/ethnicity, smoking expectancies, academic performance, internalizing (depression) and externalizing (e.g., deviance) behaviors, peer smoking, educational attainment, alcohol and drug use, and tobacco dependence [4,6–16]. There is substantial overlap in associations between trajectories [4,9,14,17] with differences most consistently seen between non-smokers and any other smoking trajectory [1,3,4,9].

Overlapping profiles suggest shared liability for smoking trajectory classification despite distinct smoking trajectories. This study investigated whether this shared liability is influenced by latent genetic factors and whether individual smoking trajectories represent phenotypes marking increasing genetic risk along a common genetic liability dimension.

Methods

Sample

The Missouri Adolescent Female Twin Study (MOAFTS) is a prospective study of a birth cohort of female like-sex twin pairs, born 1975–1985. Twins were ascertained from Missouri birth records in a cohort-sequential design with recruitment, over two years, of a cohort of 13.5, 15.5, 17.5, and 19.5 year-olds with continuous recruitment of 13-year old twins in subsequent years. Twins were first assessed at mean age 16 (baseline), with one-year brief follow-up interview with most of the sample (wave 2), and with additional follow-up in young adulthood (wave 4: mean age 22; wave 5: mean age 25) [18]. Wave 3 comprised only retest interviews for a subsample of the target cohort. For wave 4, twins were all aged at minimum 18 and all target sample twins who had not previously refused or been withdrawn from the study by their parents were re-invited to participate. Wave 5 assessments focused on past two-year behavior. The study was approved by the Institutional Review Board at Washington University School of Medicine. All participants provided written informed consent.

Measures

Defining regular smokers—All assessment waves included questions on history of cigarette use. Individuals who reported having smoked 100 cigarettes at waves 1 (total $n=3446$), 2 ($n=2927$), 4 ($n=3787$) or 5 ($n=3427$) were considered lifetime regular smokers ($n=1609$). The 100 cigarette threshold defines an “ever smoker” according to the CDC (CDC, 2002).

Diagnostic interview assessments—The twins and at least one parent informant (typically mother) were assessed over the telephone using a modified version of the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA, [19]) and the Diagnostic Interview for Children and Adolescents (DICA); the smoking section was modified from the Composite International Diagnostic Interview (CIDI). The interview included detailed assessments of history of alcohol and other substance use disorders (DSM-IV diagnoses of alcohol, nicotine dependence, and illicit drug use/dependence), psychopathology (e.g. depression, conduct disorder), and general health, school, and home environment.

Smoking quantity—Many smoking trajectory studies [3,4,8,11,12,17,20,21] including our own work [5], have used daily or weekly quantity and/or frequency measures to construct smoking trajectories. Here, we used the number of cigarettes smoked *per day* (CPD) as the dependent variable. For waves 1 (baseline) and 4, CPD was defined based on reports of daily quantity smoked during the heaviest period of smoking in order to capture

maximal vulnerability to smoking exposure. At waves 2 and 5, quantity smoked during the period of heaviest smoking was not assessed. Instead, CPD questions focused on usual daily quantity smoked in the past 12 months. CPD was categorized as 0–5, 6–10, 11–19, and 20+ across all waves.

Age as a covariate—There was a wide age range at baseline (ages 12–23) necessitating statistical adjustment for age to control for age-related differences in quantity smoked in order to reduce potential bias of assigning individuals to a heavy smoking trajectory due to their older age at baseline. For smokers who were first assessed at wave 4 ($n=303$, 20.67% of smokers included in trajectory analyses) and therefore had missing wave 1 and 2 data, we used their age at wave 4 as their ‘baseline’ age. In order for age to be scaled the same, it was dichotomized at the median, older than age 15 (or age 22) at wave 1 (or wave 4).

Data Analysis

Developmental smoking trajectories—We applied general growth mixture modeling (GMM), where trajectory classes are captured by a latent categorical variable, while class intercepts and trend (linear, quadratic) are captured by latent continuous growth factors [22,23]. GMMs were fit to the 4-category CPD data using maximum-likelihood estimation implemented in MPlus software [24]. Model building proceeded by testing the fit of an “unconditional” (no covariates) single group growth curve model and building to multiple trajectories. The same process was undertaken with inclusion of age as a covariate, regressing latent trajectory class membership on age using multinomial logistic regression. The Bayesian Information Criterion (BIC) [25] was the primary model selection criterion where the model with the lowest absolute BIC value was the most parsimonious. The general recommendation is to add trajectory groups as long as the BIC continues to decrease and the model is substantively meaningful [26]. We also considered the Lo-Mendell-Rubin Likelihood Ratio Test (LMR LRT), which tests the significance of the -2 times log-likelihood difference of a model with k and a model with $k-1$ trajectory classes where a non-significant test ($p>0.05$) favors the $k-1$ solution. Continuous growth factor means were estimated for each trajectory, with the exception of the intercept for the last trajectory that was modeled (e.g., trajectory 3 in the 3-class model or trajectory 4 in the 4-class model) for model identification. Growth factor variances were set to zero. Twins were treated as independent observations at this point of the analysis.

For each twin, posterior probabilities of trajectory group membership were estimated for as many trajectories as were being modeled. Individuals were assigned to the trajectory for which they had the highest posterior probability value. A value of 0.70 indicated reliable individual assignment to a given trajectory [27]. Ideally, individuals would have high posterior probabilities for the trajectory to which they are assigned and low posterior probabilities for the other trajectories.

Ordering of smoking trajectories for genetic threshold models—We used correlates from the literature [8,9,14,28] and from descriptive analyses to rank-order the smoking trajectories into a single categorical variable. Using multinomial logistic regression in Stata [29] and controlling for data non-independence using a robust estimator for standard errors, we examined the association of smoking trajectories (dependent variable) and lifetime history of each of the following correlates (independent variables): (1) regular alcohol use, defined as drinking at least one full drink/month for 6 months in a row; (2) first consumed one full drink of alcohol at age <15 ; (3) DSM-IV alcohol dependence; (4) ever having used marijuana; (5) first marijuana use at age <16 ; (6) endorsement of any DSM-IV marijuana abuse symptom (marijuana dependence was only assessed at wave 4 and had low prevalence (4.44%) among regular smokers); (7) DSM-IV childhood conduct disorder; and

several cigarette use behaviors including (8) ages of onset of first, regular, and daily cigarette use; (9) DSM-IV nicotine dependence symptom count; and (10) DSM-IV nicotine dependence.

Genetic liability threshold models—When modeling the genetic and environmental contribution to individual differences in a categorical phenotype, we assume a latent, normally distributed genetic liability for the behavior. When this liability exceeds a theoretical threshold, the behavior is expressed. Multiple threshold models assume a progressive increase in liability from the lowest to the highest phenotypic category. Assuming, for the sake of argument, that we have a 4-category variable, the multiple threshold model assumes that the liability for membership in the middle smoking trajectories is intermediate to that for membership in the low and high categories. A series of models were fit to the 4-category CPD variable and to alternative categorizations that combined across categories. Threshold models were fit to monozygotic (MZ) and dizygotic (DZ) twin pair contingency tables using Mx software [30]. Since contingency tables exclude singleton twins whose co-twin has missing data, singleton MZ and DZ twins were included in the models as separate groups and thresholds were equated to the MZ and DZ contingency tables from complete twin pairs. The best liability threshold model fit was the one with lowest (negative is better) Akaike Information Criterion (AIC; [31]), an indicator of model parsimony and goodness of fit to the data [32].

Additionally, for each liability threshold model, we tested the significance of the relative contribution of genetic (A), shared environmental (C), and non-shared environmental (E) influences on individual differences in trajectory group membership. This was accomplished by testing the relative fit of nested genetic (“AE”) or environmental (“CE”) models compared to the fit of the full “ACE” model using the chi-square difference test which compares chi-squares of the full and each nested model at the degrees of freedom difference between the two models. A significant chi-square difference ($p < 0.05$) suggested significant deterioration of the nested compared to the full model fit favoring the full model fit. For each liability threshold model, we also tested whether thresholds could be equated across zygosity groups by comparing models with freely estimated thresholds (which controls for prevalence differences in trajectory membership between MZ and DZ twins) and models with thresholds equated across MZ and DZ twins, using the chi-square difference test.

Results

Developmental smoking trajectories

There were 1609 lifetime regular smokers. We selected 1466 individuals with data on at least 2 of 4 waves, of which 533 (36.4%), 425 (29.0%), and 508 (34.7%) twins had data for two, three, or four waves, respectively. Mean sample age at waves 1 ($n=1162$), 2 ($n=1056$), 4 ($n=1367$), and 5 ($n=1268$) was 16.0 ($SD=2.38$, range 12–23), 17.3 ($SD=2.25$, range 14–23), 22.1 ($SD=2.68$, range 18–28), and 25.2 ($SD=2.57$, range 21–32). A total of 111 individuals (7.57%) were African American. Table 1 shows that for both the unconditional and age-adjusted models, the 4-class solutions had both the lowest BIC values and significantly differed from the 3-class models ($LMR\ LRT < 0.05$), providing the best fit to the data. In addition, the BICs for the age-adjusted models were lower than the BICs for the respective unconditional models suggesting better fit. Analysis was repeated excluding 67 regular smokers who were older than 19 years of age (20–23 years old) at wave 1 with identical results. Further, twins who were first assessed at wave 4 ($n=303$) were equally distributed across the 4 trajectories (20.0 – 21.5%), suggesting that they did not contribute to bias in the trajectory solution.

The 4-class age-adjusted model is shown in the Figure. The x-axis shows average age at each assessment and the y-axis shows the probability of heavy smoking (20+ CPD) given trajectory group membership. The smoking trajectories were characterized by smokers at Low (n=564, 38.47%), Increasing (n=339, 23.12%), Moderate (n=366, 24.97%), and High (n=197, 13.44%) risk for heavy smoking. The pattern of daily cigarette consumption across waves of assessment and trajectory classes is shown in Table S1.

Mean posterior probabilities of membership in each of the low, increasing, moderate, and high risk trajectories were, respectively, 0.80 (SD=0.18, range 0.42–0.99), 0.78 (SD=0.16, range 0.35–0.995), 0.75 (SD=0.17, range 0.38–0.99), and 0.84 (SD=0.15, range 0.42–0.999). The proportion of individuals with posterior probabilities ≥ 0.70 in each respective trajectory was 71.45%, 67.55%, 62.57%, and 87.82%, suggesting reliable trajectory group assignment.

The 4-class solution can be graphically depicted separately for each of the 0–5, 6–10, and 11–19 CPD categories (Figures S1–S3). Good trajectory separation is seen for the 0–5 CPD category (Figure S1). Less well-defined trajectory separation is seen for the intermediate CPD categories (Figures S2, S3). These results suggest that the heaviest and lightest smoking levels were most informative for trajectory estimation.

Genetic liability threshold models

Trajectory group data were available from 277 MZ and 204 DZ twin pairs and from 196 MZ and 291 DZ singleton twins. A total of 17 individuals were excluded from genetic analysis because 14 were from pairs with inconsistent within-pair zygosity assignments, and 1 twin pair (2 twins) and 1 additional twin had missing zygosity.

The pattern of correlate associations (Table 2) shows that the Low trajectory had significantly lower prevalence across correlates compared to the High trajectory, and for most correlates compared to the Increasing and Moderate trajectories. The High trajectory also had significantly elevated risk for substance use and conduct disorder relative to both Increasing and Moderate trajectories. Clearly, the Low and High trajectories could be appropriately ordered as the lowest and highest risk categories, respectively. The Increasing and Moderate trajectories had similar rates of substance use involvement and conduct disorder, making it difficult to determine their relative rank-order. However, the Increasing trajectory had significantly fewer nicotine dependence symptoms, and thus was ranked as less severe than the “moderate” trajectory. Interestingly, there was a trend for lower prevalence of alcohol dependence and earlier age of experimentation with cigarettes in the Increasing relative to the Moderate group. The direction of these trends was at odds with each other in terms of rank-order decisions. In the literature, trajectories similar to the Increasing trajectory have milder profiles compared to trajectories similar to the Moderate trajectory [8,9,14].

Altogether, correlate associations with smoking trajectories in the literature and in our data supported the following rank order of smoking trajectories: 1=Low, 2=Increasing, 3=Moderate, and 4=High.

Table 3 shows the results of liability threshold model fitting. For each trajectory categorization, the best fitting model is indicated in *italics*. The model that fit the data best overall based on lowest AIC (***bold italics***) was the 3-category model that comprised the Low, combined Increasing + Moderate, and High risk trajectories. This trajectory categorization was heritable (72.7%) with no evidence for significant contribution from shared environmental factors since the shared environmental component could be equated to

zero without significant deterioration of model fit ($p=0.097$). However, equating the genetic component to zero significantly worsened model fit ($p<0.05$).

Discussion

One aim of this study was to investigate the relative contribution of genetic and environmental influences on membership in developmental smoking trajectories among female lifetime regular smokers. Four smoking trajectories represented the best characterization of smoking behavior and included low, moderate, and high-level smokers as well as smokers who increased their smoking over time, consistent with previous studies [1,8,28]. The majority of individuals assigned to a given trajectory had high probability of membership in that trajectory suggesting reliable classification.

This study extends previous findings by showing that smoking trajectory group membership among regular smokers is heritable, with up to 72.7% of the variance attributable to latent genetic factors. However, since genetic analyses excluded never smokers and individuals who smoked <100 cigarettes lifetime, heritability estimates may be elevated. One study showed greater trajectory group concordance in MZ compared to DZ pairs suggesting genetic influences; however, estimates of heritable and environmental factors were not quantified [33]. Parent smoking, a proxy of familial smoking risk, has been associated with smoking trajectory membership compared to abstainers or low-level smokers [3,6,11,13,34]. With some exceptions [17], parent smoking does not distinguish between different smoking trajectories [3–5,8,11,13,21]. Lack of relationship between parent smoking and offspring trajectory group membership is inconsistent with evidence presented here that smoking trajectory membership is heritable. However, parent smoking was often defined based on adolescent report of their parents' current or recent (past-year) smoking status [3–5,11,13,17], providing a broad and non-specific measure of parent smoking behavior. One study showed that while adolescent report of parent smoking status was not related to adolescent smoking trajectories, detailed assessment of family smoking history conducted when the adolescents were young adults was associated with smoking trajectory membership in a dose-dependent manner [5], suggesting that genetic factors could be important contributors to trajectory classification.

A second aim of the study was to examine whether the phenotypically distinct trajectory groups represent different severity thresholds along an underlying normally-distributed latent liability of genetic and environmental risk. Model fitting showed that the Increasing and Moderate risk trajectories could be combined into a single category, suggesting that they were not distinguishable in terms of severity of genetic vulnerability and they were both intermediate in risk to the low and high risk trajectories. Perhaps these results could have been predicted considering that the Increasing and Moderate trajectories had very similar patterns of lifetime substance use involvement and rates of conduct disorder, but were different from both the Low and High trajectories. The fact that the Increasing and Moderate trajectories were phenotypically distinguishable despite high degree of shared risk suggests that there are other, perhaps time-variant risk factors that affect differences in smoking between these two trajectories. In other studies, individuals who increase their smoking in mid to late adolescence similar to twins in the Increasing trajectory, display lower levels of risk on factors such as parent smoking, peer smoking, alcohol and drug use, school grades, and deviance in adolescence that are similar to those of low-level smoking trajectories, but in young adulthood have negative outcomes similar to individuals in high-level smoking trajectories, such as lower life satisfaction, greater negative affect, problems with alcohol and drug use, greater deviance, lower educational attainment, and more positive beliefs about health and psychological consequences of smoking [2,4,5,8,9,14]. These results suggest late emergence of risk for heavy smoking that may be associated with life changes

in the transition from adolescence to young adulthood. From a genetic perspective, change in smoking behavior across adolescence could mean that inherited risk for smoking is moderated by environmental factors. It could be that individuals who increase their smoking in later adolescence are at overall moderate to high risk for smoking but they grow up in protective environments, such as those with more highly educated parents, non-smoking home rules, two-parent homes, and high parent involvement and monitoring [1,9,13]. Twin studies have shown that inherited liability to substance involvement is more pronounced in higher risk environments such as urban compared to rural settings [35], at lower levels of socioeconomic status [36], at lower religiosity [37,38], with low parental monitoring [39] and low parental closeness [40], and at higher levels of substance using peers [41].

There are several limitations of this study. First, the sample was mostly White and results might not generalize to other racial/ethnic groups. One study identified three smoking trajectories in both White and African American boys (non-smokers, light smokers, and heavy smokers); however, the African American sample smoked significantly less than the White sample [12]. Other studies have found that African Americans, Hispanics, and Asians, relative to Whites, are more likely to belong to low level smoking trajectories such as Experimenters and Quitters and less likely to belong to increasing or heavy smoking trajectories [2,9]. Second, the sample was all female, and the results may not generalize to men. Some studies have reported no sex differences [7,9,14,20,21,34], others have reported higher proportion of females in heavier smoking trajectories [1,3,5,6,13,17,28,42], and still others have concluded that there was no clear pattern of sex differences [8,43]. Where trajectories have been identified separately by sex, the same number and pattern of smoking trajectories were found for both sexes [28,42]. Overall, there appears to be consistency in trajectory group solutions across ethnicity and sex, though it is possible that our mostly White and all-female sample could be biased toward heavier smoking trajectories. Third, due to the nature of the existing data, the quantity smoked variable was defined based on either the quantity reported during the heaviest periods of smoking or usual smoking quantity in the past 12 months. It is possible that using data on quantity smoked during the heaviest period of smoking contributed to bias toward assigning smokers to heavier smoking trajectories. Finally, the results of this study have to be interpreted within the study constraints where the best fitting model was chosen based on a limited number of alternative models that were tested.

In conclusion, this study identified four patterns of smoking behavior across adolescence and young adulthood, similar to a large body of literature on smoking trajectories, in a sample of female twins who had smoked 100+ cigarettes in their life. What this study adds is evidence that trajectory group membership is significantly heritable, and that genetically, there is no distinction between individuals who have consistent moderate levels of smoking and individuals who increase their smoking in later adolescence; that is, individuals from these two trajectories could be combined into one category for genetic analysis. These results suggest that phenotypic distinctions between developmental smoking trajectories do not necessarily indicate genetic distinctions between the trajectories. The discrepancy could be due to hidden, or unmodeled, interactions between genetic risk and environmental factors [33]. Such analyses are beyond the scope of the current investigation but will be taken into account in future investigations on the genetics of smoking trajectories.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to the twins for their long term commitment to the study. This research was made possible by the Missouri Family Registry.

References

1. Abroms L, Simons-Morton B, Haynie DL, Chen R. Psychosocial predictors of smoking trajectories during middle and high school. *Addiction* (Abingdon, England). 2005; 100:852–61.
2. Audrain-McGovern J, Rodriguez D, Tercyak KP, Cuevas J, Rodgers K, Patterson F. Identifying and characterizing adolescent smoking trajectories. *Cancer Epidemiol Biomarkers Prev.* 2004; 13:2023–34. [PubMed: 15598757]
3. Brook JS, Pahl K, Ning Y. Peer and parental influences on longitudinal trajectories of smoking among African Americans and Puerto Ricans. *Nicotine Tob Res.* 2006; 8:639–51. [PubMed: 17008191]
4. Chassin L, Presson CC, Pitts SC, Sherman SJ. The natural history of cigarette smoking from adolescents to adulthood in a Midwestern community sample: multiples trajectories and their psychosocial correlates. *Health Psychology.* 2000; 19:223–31. [PubMed: 10868766]
5. Lessov-Schlaggar CN, Hops H, Brigham J, Hudmon KS, Andrews JA, Tildesley E, et al. Adolescent smoking trajectories and nicotine dependence. *Nicotine Tob Res.* 2008; 10:341–51. [PubMed: 18236299]
6. Weden MM, Miles JN. Intergenerational Relationships Between the Smoking Patterns of a Population-Representative Sample of US Mothers and the Smoking Trajectories of Their Children. *American journal of public health.* 2011
7. Brook DW, Brook JS, Zhang C, Whiteman M, Cohen P, Finch SJ. Developmental trajectories of cigarette smoking from adolescence to the early thirties: personality and behavioral risk factors. *Nicotine Tob Res.* 2008; 10:1283–91. [PubMed: 18686175]
8. Karp I, O'Loughlin J, Paradis G, Hanley J, Difranza J. Smoking trajectories of adolescent novice smokers in a longitudinal study of tobacco use. *Annals of epidemiology.* 2005; 15:445–52. [PubMed: 15967392]
9. Orlando M, Tucker JS, Ellickson PL, Klein DJ. Developmental trajectories of cigarette smoking and their correlates from early adolescence to young adulthood. *Journal of consulting and clinical psychology.* 2004; 72:400–10. [PubMed: 15279524]
10. Riggs NR, Chou CP, Li C, Pentz MA. Adolescent to emerging adulthood smoking trajectories: when do smoking trajectories diverge, and do they predict early adulthood nicotine dependence? *Nicotine Tob Res.* 2007; 9:1147–54. [PubMed: 17978988]
11. Vitaro F, Wanner B, Brendgen M, Gosselin C, Gendreau PL. Differential contribution of parents and friends to smoking trajectories during adolescence. *Addictive behaviors.* 2004; 29:831–5. [PubMed: 15135568]
12. White HR, Nagin D, Replogle E, Stouthamer-Loeber M. Racial differences in trajectories of cigarette use. *Drug Alcohol Depend.* 2004; 76:219–27. [PubMed: 15561473]
13. Bernat DH, Erickson DJ, Widome R, Perry CL, Forster JL. Adolescent smoking trajectories: results from a population-based cohort study. *J Adolesc Health.* 2008; 43:334–40. [PubMed: 18809130]
14. Costello DM, Dierker LC, Jones BL, Rose JS. Trajectories of smoking from adolescence to early adulthood and their psychosocial risk factors. *Health Psychol.* 2008; 27:811–8. [PubMed: 19025277]
15. Otten R, Wanner B, Vitaro F, Engels RC. Disruptiveness, peer experiences and adolescent smoking: a long-term longitudinal approach. *Addiction* (Abingdon, England). 2009; 104:641–50.
16. Pollard MS, Tucker JS, Green HD, Kennedy D, Go MH. Friendship networks and trajectories of adolescent tobacco use. *Addictive behaviors.* 2010; 35:678–85. [PubMed: 20332061]
17. White HR, Pandina RJ, Chen PH. Developmental trajectories of cigarette use from early adolescence into young adulthood. *Drug Alcohol Depend.* 2002; 65:167–78. [PubMed: 11772478]

18. Heath AC, Howells W, Bucholz KK, Glowinski AL, Nelson EC, Madden PA. Ascertainment of a mid-western US female adolescent twin cohort for alcohol studies: assessment of sample representativeness using birth record data. *Twin Res.* 2002; 5:107–12. [PubMed: 11931688]
19. Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr, et al. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *Journal of studies on alcohol.* 1994; 55:149–58. [PubMed: 8189735]
20. Chassin L, Presson CC, Sherman SJ, Edwards DA. The natural history of cigarette smoking: predicting young-adult smoking outcomes from adolescent smoking patterns. *Health Psychology.* 1990; 9:701–16. [PubMed: 2286181]
21. Stanton WR, Flay BR, Colder CR, Mehta P. Identifying and predicting adolescent smokers' developmental trajectories. *Nicotine Tob Res.* 2004; 6:843–52. [PubMed: 15700920]
22. Muthen B, Muthen LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcoholism, clinical and experimental research.* 2000; 24:882–91.
23. Muthen B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics.* 1999; 55:463–9. [PubMed: 11318201]
24. Muthén, LK.; Muthén, BO. *MPlus User's Guide.* 4. Los Angeles, CA: Muthén & Muthén; 1998–2006.
25. Schwarz G. Estimating the dimension of a model. *Ann Statistics.* 1978; 6:461–4.
26. Nagin DS, Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychol Methods.* 2001; 6:18–34. [PubMed: 11285809]
27. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. *Annu Rev Clin Psychol.* 2010; 6:109–38. [PubMed: 20192788]
28. Rosendahl KI, Galanti MR, Gilljam H. Trajectories of smokeless tobacco use and of cigarette smoking in a cohort of Swedish adolescents: differences and implications. *Nicotine Tob Res.* 2008; 10:1021–7. [PubMed: 18584465]
29. StataCorp. *Stata Statistical Software Release 9.* Release 9 ed. College Station, TX: Stata Corporation; 2005.
30. Neale, MC.; MBS; Xie, G.; Maes, HH. 6. VCU Box 900126, Richmond, VA 23298: Department of Psychiatry; 2003. *Mx: Statistical Modeling.*
31. Akaike H. Factor analysis and AIC. *Psychometrika.* 1987; 52:317–32.
32. Singer, JD.; Willett, JB. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence.* New York: Oxford University Press, Inc; 2003.
33. Jackson, KM.; Sher, KJ.; Rose, RJ.; Kaprio, J. Institute NC. Phenotypes and Endophenotypes: Foundations for Genetic Studies of Nicotine use and Dependence. Vol. 20. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute; 2009. Trajectories of tobacco use from adolescence to adulthood: Are the most informative genotypes tobacco specific?. *NCI Tobacco Control Monograph* NIH Publication No. 09-6366
34. de Leeuw R, Scholte R, Vermulst A, Engels R. The relation between smoking-specific parenting and smoking trajectories of adolescents: how are changes in parenting related to changes in smoking? *Psychol Health.* 2010; 25:999–1021. [PubMed: 20437293]
35. Rose RJ, Dick DM, Viken RJ, Kaprio J. Gene-environment interaction in patterns of adolescent drinking: regional residency moderates longitudinal influences on alcohol use. *Alcoholism, clinical and experimental research.* 2001; 25:637–43.
36. Dick DM, Rose RJ, Viken RJ, Kaprio J, Koskenvuo M. Exploring gene-environment interactions: socioregional moderation of alcohol use. *Journal of abnormal psychology.* 2001; 110:625–32. [PubMed: 11727951]
37. Koopmans JR, Slutske WS, van Baal GC, Boomsma DI. The influence of religion on alcohol use initiation: evidence for genotype X environment interaction. *Behavior genetics.* 1999; 29:445–53. [PubMed: 10857249]
38. Timberlake DS, Rhee SH, Haberstick BC, Hopfer C, Ehringer M, Lessem JM, et al. The moderating effects of religiosity on the genetic and environmental determinants of smoking initiation. *Nicotine Tob Res.* 2006; 8:123–33. [PubMed: 16497606]

39. Dick DM, Viken R, Purcell S, Kaprio J, Pulkkinen L, Rose RJ. Parental monitoring moderates the importance of genetic and environmental influences on adolescent smoking. *Journal of abnormal psychology*. 2007; 116:213–8. [PubMed: 17324032]
40. Miles DR, Silberg JL, Pickens RW, Eaves LJ. Familial influences on alcohol use in adolescent female twins: testing for genetic and environmental interactions. *Journal of studies on alcohol*. 2005; 66:445–51. [PubMed: 16240551]
41. Agrawal A, Balasubramanian S, Smith EK, Madden PA, Bucholz KK, Heath AC, et al. Peer substance involvement modifies genetic influences on regular substance involvement in young women. *Addiction (Abingdon, England)*. 2010; 105:1844–53.
42. White HR, Johnson V, Buyske S. Parental modeling, parenting behavior effects on offspring alcohol, cigarette use, A growth curve analysis. *Journal of substance abuse*. 2000; 12:287–310. [PubMed: 11367605]
43. Soldz S, Cui X. Pathways through adolescent smoking: a 7-year longitudinal grouping analysis. *Health Psychol*. 2002; 21:495–504. [PubMed: 12211517]

Figure.
Developmental smoking trajectories in female lifetime regular smokers (≥ 100 cigarettes)
graphed as a function of heavy smoking

Table 1

Model statistics for the unconditional (no covariates) and age-adjusted models

Model	Unconditional		Age-Adjusted	
	BIC	LMR LRT	BIC	LMR LRT
1 class	11453.183		13594.463	
2 classes	10845.905	<.001	10729.093	<.001
3 classes	10725.690	<.001	10527.279	<.001
4 classes	10641.435	<.001	10431.938	<.001
5 classes	10643.089	0.014	10436.086	0.011

BIC=Bayesian Information Criterion; LMR LRT=Lo-Mendell-Rubin Likelihood Ratio Test; Models were fit to daily quantity smoked data categorized as 0–5, 6–10, 11–19, and 20+ cigarettes per day.

Table 2

Substance use and psychopathology across four trajectory classes

Substance Use and Psychopathology	Low	Increasing	Moderate	High
Regular alcohol use, %	62.4	65.2	66.9	70.6 ^L
First full drink < age 15, %	33.3	44.3	36.9	52.8 ^{LM}
Alcohol dependence, %	14.9	15.6	21.6 ^{L*}	33.0 ^{LM}
Ever tried marijuana, %	66.3	74.6 ^L	75.4 ^L	83.8 ^{LM}
First tried marijuana < age 16, %	19.3	33.6 ^L	27.3 ^L	42.6 ^{L*} M
Any marijuana abuse symptom, %	3.9	8.0 ^L	7.7 ^L	13.2 ^{LM}
Childhood conduct disorder, %	5.5	12.1 ^L	10.7 ^L	21.8 ^{LM}
Age first tried cigarette smoking, mean (SD)	13.9 (2.72)	12.5 ^L (2.54)	13.0 ^{L*} (3.10)	12.2 ^{LM} (3.06)
Age of onset of regular smoking, mean (SD)	16.5 (2.43)	15.4 ^L (2.47)	15.1 ^L (2.66)	14.3 ^{LM} (2.47)
Age of onset of daily smoking, mean (SD)	16.9 (2.29)	16.0 ^L (2.34)	15.8 ^L (2.41)	15.0 ^{LM} (2.34)
Nicotine dependence symptom count, mean (SD)	1.88 (1.63)	3.21 ^L (1.84)	3.67 ^{L*} (1.62)	4.59 ^{LM} (1.37)
Nicotine dependence, %	27.7	58.9 ^L	57.0 ^L	76.7 ^{LM}

CPD=Cigarettes Per Day;

L, I, M prevalence or mean significantly different ($p < 0.05$) from that of the Low, Increasing, or Moderate trajectories, respectively;

* 0.05 < $p < 0.06$ for the comparison with the Increasing trajectory

Table 3

Genetic liability threshold models showing the relative contribution of genetic (A), shared environmental (C), and non-shared environmental (E) factors on developmental smoking trajectory group membership adjusted for age. The best fitting model for each categorization is indicated in italics. The best fitting model overall is indicated in bold italics.

Models	Variance Components				Model Fit		Relative Model Fit		
	A (95% CI)	C (95% CI)	E (95% CI)	χ^2	df	p	AIC	χ^2 diff	p diff
<i>4 categories (Low, Increasing, Moderate, High) [MZ, DZ thresholds equated]</i>									
ACE	43.8 (18.9, 71.3)	31.2 (5.6, 52.8)	25.0 (19.1, 32.4)	112.500	31	<.001	50.500		
AE	76.5 (69.9, 81.8)	--	23.5 (18.2, 30.1)	118.065	32	<.001	54.065	5.565	1 0.018
CE	--	65.4 (58.5, 71.4)	34.6 (28.6, 41.5)	124.737	32	<.001	60.737	12.237	1 <.001
<i>3 categories (Low + Increasing, Moderate, High) [MZ, DZ thresholds equated]</i>									
ACE	31.4 (7.0, 58.9)	49.0 (23.3, 69.7)	19.6 (13.9, 27.3)	33.590	16	0.006	1.59		
AE	82.2 (75.7, 87.2)	--	17.8 (12.8, 24.3)	45.669	17	<.001	11.669	12.079	1 0.001
CE	--	73.6 (66.7, 79.4)	26.4 (20.6, 33.3)	40.035	17	0.001	6.035	6.445	1 0.011
<i>3 categories (Low, Increasing+ Moderate, High) [MZ, DZ thresholds equated]</i>									
ACE	46.0 (15.6, 76.4)	25.0 (0, 50.8)	29.0 (21.5, 38.6)	29.128	16	0.023	-2.872		
AE	72.7 (64.2, 79.5)	--	27.4 (20.5, 35.8)	31.89	17	0.016	-2.11	2.761	1 0.097
CE	--	60.8 (52.5, 68.0)	39.2 (32.0, 47.5)	38.048	17	0.002	4.048	8.92	1 0.003
<i>3 categories (Low, Increasing, Moderate + High) [MZ, DZ thresholds free]</i>									
ACE	33.7 (4.4, 66.2)	42.6 (12.3, 67.6)	23.8 (17.1, 32.5)	89.797	14	<.001	61.797		
AE	78.3 (70.7, 84.3)	--	21.7 (15.8, 29.3)	96.976	15	<.001	66.976	7.179	1 0.007
CE	--	69.8 (62.2, 76.2)	30.2 (23.8, 37.8)	94.958	15	<.001	64.958	5.161	1 0.023
<i>2 categories (Low + Increasing+ Moderate, High) [MZ, DZ thresholds equated]</i>									
ACE	61.7 (9.4, 87.4)	14.5 (0, 58.2)	23.8 (12.5, 40.6)	13.111	5	0.022	3.111		
AE	77.4 (62.2, 87.9)	--	22.6 (12.1, 37.8)	13.458	6	0.036	1.458	0.346	1 0.556
CE	--	62.5 (48.0, 74.3)	37.5 (25.7, 52.0)	18.449	6	0.005	6.449	5.338	1 0.021
<i>2 categories (Low, Increasing+ Moderate + High) [MZ, DZ thresholds equated]</i>									
ACE	44.0 (0, 79.0)	25.3 (0, 62.4)	30.7 (20.1, 44.2)	19.567	5	0.002	9.567		
AE	71.2 (58.9, 80.9)	--	28.8 (19.1, 41.1)	20.965	6	0.002	8.965	1.398	1 0.237
CE	--	60.8 (49.4, 70.6)	39.2 (29.4, 50.6)	23.442	6	0.001	11.442	3.876	1 0.049
<i>2 categories (Low + Increasing, Moderate + High) [MZ, DZ thresholds free]</i>									
ACE	23.4 (0, 54.0)	63.2 (34.5, 84.7)	13.4 (7.6, 22.1)	9.72	4	0.045	1.720		

Models	Variance Components			Model Fit			Relative Model Fit		
	A (95% CI)	C (95% CI)	E (95% CI)	χ^2	df	p	χ^2 diff	df diff	p diff
AE	88.7 (81.7, 93.5)	--	11.3 (6.5, 18.3)	24.705	5	<.001	14.705	1	<.001
CE	--	81.8 (74.4, 87.5)	18.2 (12.5, 25.6)	12.657	5	0.027	2.937	1	0.087

χ^2 =chi-square goodness of fit statistic; df=degrees of freedom; p=p value of model fit; AIC=Akaike Information Criterion; diff= difference