Supplemental Information

Supplemental Methods and Materials

Data collection

Data on pregnancy and perinatal period were collected prospectively using hospital records, communal maternity clinic cards, and a questionnaire administered by midwives to mothers during the 6th-7th month of pregnancy. Information included mothers' smoking habits before and during pregnancy, weight and height before pregnancy as well as indices of family socio-economic status (i.e. parents' occupations, mothers' marital status, education). Participants were re-contacted in 1980 at 14 years of age and were asked to fill in a questionnaire mailed directly to participants, or to their parents (in case participants did not respond), or to school health nurses (in case neither participants nor their parents did respond). Eleven thousand and ten participants (93.6%) returned questionnaires. Altogether with parental and nurses returns, data were available for 97% of the adolescents. Information collected at this time point included participants' smoking habits as well as parents' occupations. Participants were re-contacted at age 31 in 1997-8, and were asked to fill in a questionnaire including questions on participants' smoking habits, occupation, education and employment status. Those who were still living in the provinces of Oulu and Lapland or who had moved to the capital (Helsinki) area were invited to a clinical examination and blood sampling. Questionnaires were returned by 8,767 (75%) participants and 6,033 (71.3% of eligible) subjects attended a clinical examination (1). During the clinical examination participants received the Temperament and Character Inventory (TCI) questionnaire (2, 3) to be filled in at home, and 5,115 (84.4%) returned it.

Maternal smoking. Participants' mothers' smoking habits were assessed by a questionnaire during pregnancy. Mothers reported if they were smoking during the 12 months preceding pregnancy. In case of an affirmative reply, numbers of cigarettes/day were recorded. Mothers were also asked whether they had changed their smoking habits during pregnancy. In case of an affirmative answer, quality of change ('stop', 'cut down', 'increase'), month of pregnancy when change occurred, and number of cigarettes smoked daily after the change were collected. In the present study, we used a variable related to the number of cigarettes the mothers were smoking during the 2nd month of pregnancy and classified mothers as: non-smokers; light-smokers (1-10 cigarettes/day); and heavy-smokers (> 10 cigarettes/day).

Alcohol use was assessed by postal questionnaires sent to cohort members at age 14 and 31 years.

Age 14 years. Participants' alcohol use was inquired with the following questions and response alternatives: "Alcohol use (beer or other alcoholic beverages): 1) I have never drunk any; 2) I have tasted once; 3) I have drunk a few times; 4) I use alcohol monthly; 5) I use alcohol weekly". We classified subjects into 1) non-users (never, merely tasted or had consumed alcohol occasionally; categories 1-3) and 2) regular users (monthly or more often; categories 4-5) [see also Rantakallio *et al.* (4)].

Age 31 years. The postal questionnaire sent at age 31 years included several questions on type of consumed alcohol, amount and frequency and the answers were validated against 7-day food diaries for a subset of the cohort. The data were transformed into daily intake (grams/day) (see below for a detailed description and also Järvelin *et al.* (1); Poikolainen and Vartiainen (5)) and we further classified subjects into abstainers and consumers divided according to tertiles of daily intake.

In more detail, participants were asked the following questions:

1) Do you use any alcoholic drinks, even occasionally?: I have never used; No, because I quit using alcohol __ years ago; Yes, less than once a month; Yes, at least once a month

2) How often do you usually drink beer, cider or long-drinks? (Alcohol content 4.8%): never; once a year or more seldom; twice a year; 3-4 times per year; about once in two months; about once a month; two times a month; about once a week; a few times a week; daily

3) How much do you usually drink beer, cider or long drinks at a time?: less than one bottle (167 ml); 1 bottle (333 ml); 2 bottles (666 ml); 3 bottles (1000 ml); 4-5 bottles (1500 ml); 6-9 bottles (2500 ml); 10-14 bottles (4000 ml); 15 bottles or more (5000 ml); I do not drink these drinks

4) How often do you usually drink light wine (alcohol content 5%)?: never; once a year or more seldom; twice a year; 3-4 times a year; about once in two months; about once a month; two times a month; about once a week; a few times a week; daily

5) How much do you usually drink light wine at a time?: half a glass (80 ml); one glass (160 ml); two glasses (320 ml); about half a bottle (375 ml); a little less than one bottle (575 ml); about one bottle (750 ml); from one to two bottles (1125 ml); more than two bottles (1875 ml); I do not drink wine

6) How often do you usually drink wine (alcohol content 14.5%)?: never; once a year or more seldom; twice a year; 3-4 times a year; about once in two months; about once a month; two times a month; about once a week; a few times a week; daily

7) How much do you usually drink wine at a time?: half a glass (80 ml); one glass (160 ml); two glasses (320 ml); about half a bottle (375 ml); a little less than one bottle (575 ml); about one bottle (750 ml); from one to two bottles (1125 ml); more than two bottles (1875 ml); I do not drink wine

8) How often do you usually drink spirits (alcohol content 37%)?: never; once a year or more seldom; twice a year; 3-4 times a year; about once in two months; about once a month; two times a month; about once a week; a few times a week; daily

9) How much do you usually drink spirits at a time?: less than one restaurant measure (20 ml); one restaurant measure (40 ml); two restaurant measures (80 ml); 3-4 restaurant measures (140 ml); 5-6 restaurant measures (220 ml); 7-10 restaurant measures (340 ml); about half a liter bottle (500 ml); more than a half liter bottle (750 ml); I do not drink spirits

Frequencies of alcohol use were proportioned to 365 days: never = 0; once a year or more seldom = 1/365; twice a year = 2/365; 3-4 times a year = 3.5/365; about once in two months = 6/365; about once a month = 12/365; a few times a month = 24/365; about once a week = 52/365; a few times a week = 182/365; daily = 1.

The following alcohol contents (Vol-%) were applied: for beer 4.8, for light wine 5.0, for wine 14.5, and for spirits 37 (see also Poikolainen and Vartiainen (5)). The alcohol portions were turned into measures of volume as follows: one bottle of beer, cider or long drink 333 ml (1/3 l), one glass of wine 160 ml, and one measure of spirits 40 ml.

The amount of pure alcohol [g] consumed at a time was calculated as: Volume [ml] x Vol-% x 0.816 [g/ml] (please note that one ml of alcohol weighs 0.816 g). The average daily amount of pure alcohol [g/day] was calculated to be: Pure alcohol [g] at a time x Frequency of alcohol use [1/day]. Finally, the alcohol amounts calculated for each type of drink were added together to obtain the total amount of pure alcohol consumed per day.

Co-occurrence of alcohol use and nicotine use and association between alcohol use and *TTC12*-rs10502172/ *CHRNA3*-rs1051730

We found a strong co-occurrence between smoking behavior and alcohol use both at age 14 years (see Table S4) and 31 years (see Table S5).

Association between *CHRNA3*-rs1051730/ *TTC12*-rs10502172 and alcohol use was tested using logistic regression (binary logistic at age 14 and multinomial logistic at age 31 years). Alcohol use was entered in the regression model as dependent variable. At age 14 years alcohol use was binary coded as: non-users; regular users. At age 31 years alcohol use was coded in four categories: non-users, alcohol-users following within the 1st tertile of the distribution of daily alcohol intake; alcohol-users in the 2nd tertile; and alcohol-users in the 3rd tertile. Allelic effect at each locus was modeled as additive. The following covariates were also entered in the models: socioeconomic status, gender, principal components significantly associated with smoking. No significant association was found between *CHRNA3*-rs1051730 and *TTC12*-rs10502172 and alcohol use at age 14/31 years (see Table S6).

Haplotype analyses

Haplotype structures of the 15q25 and 11q23 regions were computed with Haploview v3 (6). Association between haplotype and smoking behavior was performed using χ^2 test run in PLINK version 1.06. Results of the global as well as haplotype-specific (i.e. comparing frequencies of one haplotype vs. all other haplotypes combined) association tests are reported for each haplotype-block in the main text in Tables 2 and 3. Degree of freedoms (df) for these analyses are equal to: N of haplotypes within a block minus 1 for the global tests; 1 for all the haplotype-specific tests. As done for single markers analyses, smoking has been coded differently for *CHRNA5-CHRNA3-CHRNB4* and *TTC12-ANKK1-DRD2.* For *CHRNA5-CHRNA3-CHRNB4*, smoking behavior was dichotomized as: 1) Smoking at 14: regular smokers vs. occasional + non-smokers; 2) Smoking at 31 years: subjects smoking > 10 cigarettes/day vs. subjects smoking 1-10 cigarettes/day + non-smokers (i.e. in both time points we compared heavier smokers with the rest of the sample). In contrast, for the *TTC12-ANKK1-DRD2* gene-cluster, smoking behavior was dichotomized as smokers vs. non-smokers regardless of severity (see statistical analysis section of the main text for explanation).

Correction for population substructure

Principal components (PCs) identify the major axes of variation in multivariate data. It has been shown that PCs provide a good summary of population genetic structure (7) and their inclusion in tests for genotype phenotype association has been shown to correct for population structure (8).

Principal components were calculated from SNPs selected uniformly from the 22 autosomes such that all SNPs had minor allele frequency > 0.01, HWE p-value > 0.005, call rate > 0.995 and no two SNPs were in r2 > 0.2 with any other selected SNP. This gave 61,917 SNPs genotyped in 4,793 individuals. Following (9), SNP data was normalized to have mean zero and variance one. If we denote the resultant genotype matrix by X, where individuals are represented in rows and SNPs in columns, principal components are calculated by an eigenvalue decomposition of XX^t, where X^t is the transpose of X. Eigenvalue decomposition was conducted in the statistical software R.

We adjusted all our analyses for PCs that were significantly associated with smoking at 14 years, smoking at 31 years, novelty seeking, or maternal smoking in pregnancy.

A detailed description of the genetic diversity in Northern Finland, from which the samples in Northern Finland Birth Cohort (NFBC) are drawn, can be found in Sabatti *et al.* (10). Of note, population substructure analyses of NFBC performed by Sabatti *et al.* did not identify clearly separated clusters as expected in a population with strong subpopulation structure, however, the identified fine structure did correlate with geographic location within Northern Finland (10).

Supplemental Results

Interaction between maternal smoking and genotype at *TTC12*-rs10502172 and *CHRNA3*-rs1051730 analyses

Interaction analyses (i.e. multiplicative interaction) were performed between maternal smoking during pregnancy and genotypes at *TTC12*-rs10502172 and *CHRNA3*-rs1051730, at both age 14 and 31 years. Multinomial logistic regression run in SAS v9.2 was used to perform these analyses. Smoking at 14 and 31 years were entered in the regression model as dependent variables and were coded using three categories of severity (age 14: regular/occasional/non-smokers; age 31: heavy/light/non-smokers). Allelic effect at each locus was modeled as additive. The model included also *TTC12*-rs10502172, *CHRNA3*-rs1051730, and maternal smoking main-effects as well as additional covariates (principal components, socioeconomic status, and gender). Statistics presented below refer to the gene*maternal smoking interactive effect on each smoking category (non-smoking is the reference category). Results of these analyses should be interpreted with caution given the low number of participants following in some groups and the elevated number of tests performed.

At age 14 years, a significant interactive effect on smoking behavior was found between *CHRNA3*-rs1051730[A] and maternal smoking in pregnancy on risk of smoking occasionally (Wald chi-square = 8.84, df = 1, *p* = 0.003). Among offspring of non-smoking mothers, *CHRNA3*-rs1051730[A] was associated with increased risk of smoking occasionally at age 14 years [OR (95% CI): 1.15 (0.99-1.35)], although this difference only approached statistical significance (df = 1, Wald chi-square = 3.16, *p* = 0.07). In contrast, among offspring of smoking mothers, *CHRNA3*-rs1051730[A] was associated with significantly reduced risk of being occasional smokers than non-smokers [OR (95% CI): 0.59 (0.39-0.89), df = 1, Wald chi-square = 6.3, *p* = 0.01) at age 14 years. No significant interactive effects were found between *CHRNA3*-rs1051730[A] and maternal smoking on risk of smoking regularly (Wald chi-square = 1.43, df = 1, *p* = 0.23) at age 14. No significant interactive effects were found between *CHRNA3*-rs1051730[A] and maternal smoking on risk of smoking at age 31 years [heavy-smoking: Wald chi-square = 0.29, df = 1, *p* = 0.59]; light-smoking: Wald chi-square = 0.50, df = 1, *p* = 0.48].

No significant interactions were found between genotype at *TTC12*-rs10502172 and maternal smoking:

- *TTC12*-rs10502172 [G]*maternal smoking on smoking regularly 14 years: Wald chi-square = 0.27, df = 1, *p* = 0.61
- *TTC12*-rs10502172 [G]*maternal smoking on smoking occasionally at 14 years: Wald chi-square = 0.50, df = 1, *p* = 0.48
- *TTC12*-rs10502172 [G]*maternal smoking on smoking heavily at 31 years: Wald chi-square = 1.68, df = 1, *p* = 0.20
- *TTC12*-rs10502172 [G]*maternal smoking on light-smoking at 31 years: Wald chi-square = 0.06, df = 1, *p* = 0.81.

Correlations between maternal smoking and genotype *TTC12*-rs10502172 and *CHRNA3*-rs1051730

Association between genotype at *TTC12*-rs10502172 and *CHRNA3*-rs1051730 and maternal smoking were tested by comparing allele frequencies at the two loci in offspring of mothers who were smoking in pregnancy vs. offspring of non-smoking mothers. A significant correlation was found for *CHRNA3*-rs1051730. Offspring of mothers smoking heavily during pregnancy more likely carried *CHRNA3*-rs1051730[A] than offspring of non-smoking mothers [1.39 (1.04-1.87), *P* = 0.03]. No differences in frequency of *CHRNA3*-rs1051730[A] were found when comparing offspring of mothers smoking 1-10 cigarettes/day with offspring of non-smoking mothers [0.96 (0.83-1.10), *P* = 0.53]. No significant correlations were found for *TTC12*-rs10502172[G] [1.12 (0.83-1.502), *p* = 0.45 for offspring of heavy-smoking mothers and 1.06 (0.93-1.21), *p* = 0.36, for offspring of light-smoking mothers].

Supplemental Tables

Table S1. Phenotypic and genotype characterization of the Northern Finland Birth Cohort 1966 study sample (N = 4,762). The subset of the cohort that is included in this study is compared with the participants who were not included because DNA was not collected.

Variable	**Study sample (N = 4,762) N (%)**	**Rest of the cohort (N = 7,469) N (%)**	**Entire cohort (N = 12,231) N (%)**	**Study sample vs. rest of the cohort χ^2 (df), *P* Value**
Maternal smoking at 2nd month of pregnancy				15.95 (2), **<0.001**
Non-smokers	4,000 (86.0)	6,017 (83.5)	10,017 (84.5)	
Light-smokers (1-10 cigarettes/day)	554 (11.9)	977 (13.6)	1,531 (12.9)	
Heavy-smokers (> 10 cigarettes/day)	96 (2.1)	210 (2.9)	306 (2.6)	
Mother's marital status during pregnancy				27.16 (1), **<0.001**
Married	4,605 (96.8)	6,986 (94.8)	11,591 (95.6)	
Unmarried	151 (3.2)	380 (5.2)	531 (4.4)	
Family socio-economic status during pregnancy				53.23 (3), **<0.001**
Professionals	1,107 (23.4)	1,692 (23.1)	2,799 (23.2)	
Skilled workers	1,616 (34.2)	2,415 (33.0)	4,031 (33.5)	
Unskilled workers	1,004 (21.2)	1,931 (26.4)	2,935 (24.4)	
Farmers	1,000 (21.2)	1,278 (17.5)	2,278 (18.9)	
Gender				32.08 (1), **<0.001**
Male	2,286 (48.0)	3,928 (53.3)	6,214 (51.2)	
Female	2,476 (52.0)	3,446 (46.7)	5,922 (48.8)	
Smoking at 14 years				1.01 (2), 0.60
Non-smokers (smoked once/twice in life)	3,787 (83.1)	5,221 (82.9)	9,008 (83.0)	
Occasional-smokers (smoking occasionally or about twice per week)	491 (10.8)	663 (10.5)	1,154 (10.6)	
Regular-smokers (> twice a week)	279 (6.1)	414 (6.6)	693 (6.4)	
Family socio-economic status at 14 years				65.65 (3), **<0.001**
Professionals	1,341 (29.3)	1,847 (29.1)	3,188 (29.2)	
Skilled workers	1,573 (34.4)	2,289 (36.1)	3,862 (35.4)	
Unskilled workers	987 (21.6)	1,573 (24.8)	2,560 (23.5)	
Farmers	674 (14.7)	629 (9.9)	1,303 (11.9)	

Smoking at 31 years				0.23 (2), 0.89
Non-smoker	2,589 (58.0)	681 (58.4)	3,270 (58.0)	
Light smokers (1-10 cigarettes/day)	947 (21.2)	250 (21.4)	1,197 (21.2)	
Heavy smokers (> 10 cigarettes/day)	932 (20.9)	236 (20.2)	1,168 (20.7)	
Socio-economic status at 31 years				7.08 (4), 0.13
Professionals	1,098 (23.3)	311 (25.4)	1,409 (23.8)	
Skilled workers	1,481 (31.5)	350 (28.6)	1,831 (30.9)	
Unskilled workers	1,215 (25.8)	302 (24.7)	1,517 (25.6)	
Farmers	164 (3.5)	50 (4.1)	214 (3.6)	
Others	747 (15.9)	212 (17.3)	959 (16.2)	
***TTC12*-rs10502172**				
AA	984 (20.8)	NA	NA	
TG	2,391 (50.4)	NA	NA	
GG	1,368 (28.8)	NA	NA	
***CHRNA3*-rs1051730**				
GG	2,201 (46.2)	NA	NA	
AG	2,070 (43.5)	NA	NA	
AA	489 (10.3)	NA	NA	
	Mean (SD)			
Novelty seeking (N = 4,057 in the study sample; N = 1,041 in the rest of the cohort; N = 5,098 in the entire cohort)	20.3 (5.9)	20.1 (6.2)	20.3 (6.0)	t = -1.17, df = 5096, *P* = 0.24

Bolded values indicate significance.

Table S2. Association between smoking behavior during adolescence and adulthood and each genetic marker in the 15q25 candidate chromosome regions. High-quantity smokers are compared to low-quantity/non-smokers at both time points (see text for more details). Odds ratios (OR) and 95% confidence intervals (CI) are computed using logistic regression and adjusted for principal components, gender, maternal smoking during pregnancy, and socio-economic status. Allelic effect at each locus is modelled as additive (df = 1). In bold, P < 0.05, thick lines indicate boundaries between haplotype-blocks.

							Smoking at 14 years [Regular (N = 279) vs. occasional+non-smokers (N = 4,278)]			**Smoking at 31 years** [Heavy (N = 932) vs. light+non-smokers (N = 3,536)]		
Chr	**rs number**	**Gene**	**Position**	**Location**	**Risk/other allele**[a]	**MAF**	**N**	**OR (95% CI)**	***P* Value**	**N**	**OR (95% CI)**	***P* Value**
15	rs8034191	*LOC123688*	76593078	Intron	**G**/A	0.33	4446	1.21 (1.01-1.45)	**0.04**	4346	1.29 (1.15-1.45)	**1.18x10^{-5}**
15	rs3885951	*LOC123688*	76612972	flanking 3’UTR	**G**/A	0.06	4443	1.39 (1.01-1.92)	0.05	4343	1.22 (0.98-1.52)	0.08
15	rs2036534	*LOC123688*	76614003	flanking 3’UTR	A/**G**	0.28	4435	1.05 (0.87-1.28)	0.59	4334	1.15 (1.01-1.30)	**0.03**
15	rs6495306	*CHRNA5*	76652948	Intron	A/**G**	0.38	4437	1.23 (1.02-1.48)	**0.03**	4338	1.13 (1.01-1.27)	**0.03**
15	rs680244	*CHRNA5*	76658343	Intron	G/**A**	0.38	4444	1.23 (1.02-1.48)	**0.03**	4344	1.13 (1.01-1.27)	**0.03**
15	rs621849	*CHRNA5*	76659916	Intron	A/**G**	0.38	4430	1.23 (1.02-1.48)	**0.03**	4331	1.13 (1.01-1.27)	**0.03**
15	rs1051730	*CHRNA3*	76681394	Coding	**A**/G	0.32	4444	1.25 (1.04-1.50)	**0.02**	4344	1.29 (1.15-1.45)	**1.05x10^{-5}**
15	rs6495309	*CHRNB4*	76702300	flanking 3’UTR	G/**A**	0.27	4443	1.02 (0.84-1.24)	0.83	4343	1.17 (1.03-1.33)	**0.01**
15	rs1948	*CHRNB4*	76704454	3’UTR	G/**A**	0.34	4441	1.33 (1.10-1.62)	**0.004**	4339	1.14 (1.02-1.28)	**0.03**
15	rs950776	*CHRNB4*	76713073	Intron	A/**G**	0.33	4442	1.37 (1.13-1.67)	**0.002**	4343	1.17 (1.04-1.31)	**0.01**
15	rs12594247	*CHRNB4*	76733688	flanking 5’UTR	**A**/G	0.21	4444	1.03 (0.83-1.28)	0.79	4344	1.00 (0.87-1.14)	0.96
15	rs12900519	*CHRNB4*	76736182	flanking 5’UTR	A/**G**	0.14	4446	1.14 (0.88-1.49)	0.33	4346	1.03 (0.88-1.20)	0.71
15	rs1996371	*CHRNB4*	76743861	flanking 5’UTR	**G**/A	0.35	4439	1.24 (1.04-1.49)	**0.02**	4339	1.21 (1.08-1.36)	**0.001**
15	rs6495314	*CHRNB4*	76747584	flanking 5’UTR	**C**/A	0.35	4437	1.25 (1.04-1.49)	**0.02**	4337	1.21 (1.08-1.36)	**0.001**
15	rs8032156	*CHRNB4*	76751553	flanking 5’UTR	G/**A**	0.30	4377	1.08 (0.89-1.31)	0.42	4283	1.03 (0.91-1.16)	0.66
15	rs8038920	*CHRNB4*	76761600	flanking 5’UTR	G/**A**	0.27	4441	1.10 (0.90-1.35)	0.33	4341	1.15 (1.01-1.30)	0.03
15	rs4887077	*CHRNB4*	76765419	flanking 5’UTR	**A**/G	0.33	4442	1.24 (1.04-1.49)	**0.02**	4343	1.25 (1.12-1.40)	**0.0001**
15	rs11638372	*CHRNB4*	76770614	flanking 5’UTR	**A**/G	0.33	4444	1.24 (1.04-1.49)	**0.02**	4344	1.26 (1.12-1.41)	**0.0001**

CI, confidence interval; MAF, minor allele frequency; OR, odds ratio.

[a] Minor allele is indicated in bold.

Table S3. Association between smoking behavior during adolescence and adulthood and each genetic marker in the 11q23 chromosome regions. Smokers are compared to non-smokers at both time points (see text for more details). Odds ratios (OR) and 95% confidence intervals (CI) are computed using logistic regression and adjusted for principal components, gender, maternal smoking during pregnancy, and socio-economic status. Allelic effect at each locus is modelled as additive (df = 1). In bold, P < 0.05; thick lines indicate boundaries between haplotype-blocks.

							Smoking at 14 years [Regular+occasional smokers (N = 770) vs. non-smokers (N = 3,787)]			**Smoking at 31 years** [Heavy+light smokers (N = 1,879) vs. non-smokers (N = 2,589)]		
Chr	**rs number**	**Gene**	**Position**	**Location**	**Risk/other allele[a]**	**MAF**	**N**	**OR (95% CI)**	***P* Value**	**N**	**OR (95% CI)**	***P* Value**
11	rs4517559	*TTC12*	112685412	flanking 5'UTR	**G**/A	0.49	4443	1.09 (0.98-1.22)	0.12	4343	1.00 (0.92-1.10)	0.94
11	rs2236709	*TTC12*	112691986	Intron	**G**/A	0.30	4444	1.15 (1.02-1.30)	**0.02**	4344	0.99 (0.90-1.09)	0.80
11	rs7927508	*TTC12*	112693407	Intron	**G**/A	0.49	4443	1.09 (0.98-1.22)	0.12	4343	1.00 (0.92-1.09)	0.97
11	rs2156486	*TTC12*	112694133	Intron	A/**C**	0.19	4445	1.18 (1.02-1.37)	**0.03**	4345	0.98 (0.87-1.09)	0.67
11	rs723077	*TTC12*	112699378	Coding	**C**/A	0.43	4437	0.94 (0.84-1.05)	0.27	4337	1.03 (0.94-1.13)	0.47
11	rs10502172	*TTC12*	112704356	Intron	G/**A**	0.46	4427	1.30 (1.16-1.46)	**9.05x10^{-6}**	4329	1.13 (1.03-1.23)	**0.01**
11	rs2303380	*TTC12*	112705919	Intron	**G**/A	0.30	4442	1.23 (1.09-1.38)	**0.001**	4342	1.11 (1.01-1.22)	**0.03**
11	rs4987094	*TTC12*	112721133	Intron	**A**/G	0.20	4445	1.11 (0.97-1.28)	0.13	4345	1.04 (0.93-1.16)	0.48
11	rs2276070	*TTC12*	112735810	Intron	**A**/G	0.24	4443	1.11 (0.97-1.26)	0.13	4343	1.04 (0.94-1.15)	0.48
11	rs719802	*TTC12*	112739889	Intron	**A**/G	0.32	4442	1.19 (1.06-1.34)	**0.004**	4342	1.08 (0.99-1.19)	0.10
11	rs719804	*TTC12*	112739985	Intron	**G**/A	0.17	4429	1.17 (1.02-1.35)	**0.03**	4327	1.09 (0.97-1.23)	0.13
11	rs754672	*TTC12*	112754346	flanking 3'UTR	G/**A**	0.46	4441	1.28 (1.14-1.44)	**2.12x10^{-5}**	4341	1.12 (1.02-1.22)	**0.02**
11	rs877138	*ANKK1*	112761718	flanking 5'UTR	**G**/A	0.30	4445	1.23 (1.09-1.39)	**0.001**	4346	1.10 (1.00-1.21)	0.05
11	rs4590907	*ANKK1*	112768580	Intron	**C**/A	0.24	4440	1.08 (0.95-1.24)	0.23	4340	1.02 (0.92-1.13)	0.70
11	rs7118900	*ANKK1*	112772031	Coding	**A**/G	0.16	4440	1.07 (0.92-1.25)	0.37	4340	1.05 (0.92-1.18)	0.47
11	rs2734849	*ANKK1*	112775370	Coding	A/**G**	0.46	4419	1.25 (1.11-1.40)	**0.0002**	4320	1.11 (1.02-1.22)	**0.02**
11	rs1800497	*ANKK1*	112776038	Coding	**A**/G	0.17	4394	1.10 (0.95-1.27)	0.23	4293	1.04 (0.93-1.18)	0.49
11	rs2242592	*DRD2*	112784640	flanking 3'UTR	**G**/A	0.37	4438	1.18 (1.05-1.32)	0.005	4336	1.08 (0.98-1.18)	0.12
11	rs1076563	*DRD2*	112801119	Intron	A/**C**	0.50	4445	1.23 (1.10-1.37)	**0.0004**	4345	1.08 (0.99-1.18)	0.09
11	rs2471857	*DRD2*	112803549	Intron	**A**/G	0.17	4443	1.10 (0.94-1.27)	0.23	4343	1.03 (0.91-1.16)	0.62
11	rs4620755	*DRD2*	112814829	Intron	**A**/G	0.23	4444	1.08 (0.94-1.23)	0.29	4344	1.01 (0.91-1.13)	0.80
11	rs7125415	*DRD2*	112815891	Intron	**A**/G	0.19	4444	1.08 (0.94-1.25)	0.27	4344	1.01 (0.91-1.13)	0.81
11	rs4648318	*DRD2*	112818599	Intron	**G**/A	0.34	4445	1.21 (1.08-1.36)	**0.001**	4344	1.05 (0.96-1.16)	0.28
11	rs4274224	*DRD2*	112824662	Intron	G/A	0.24	4442	0.98 (0.86-1.12)	0.74	4342	1.02 (0.92-1.13)	0.78
11	rs4581480	*DRD2*	112829684	Intron	**G**/A	0.07	4440	1.19 (0.97-1.46)	0.10	4341	1.03 (0.86-1.22)	0.78
11	rs7131056	*DRD2*	112834984	Intron	**C**/A	0.49	4424	0.99 (0.89-1.11)	0.89	4323	1.02 (0.93-1.11)	0.69
11	rs4938019	*DRD2*	112846601	Intron	**G**/A	0.23	4441	1.06 (0.93-1.21)	0.39	4341	1.02 (0.92-1.14)	0.65
11	rs12364283	*DRD2*	112852165	flanking 5'UTR	**G**/A	0.08	4410	1.06 (0.86-1.31)	0.57	4316	1.05 (0.89-1.24)	0.54

11	rs10891556	*DRD2*	112857971	flanking 5’UTR	**A**/C	0.24	4430	1.08 (0.94-1.22)	0.27	4331	1.02 (0.92-1.13)	0.68
11	rs6589377	*DRD2*	112860946	flanking 5’UTR	A/**G**	0.17	4440	1.16 (0.99-1.35)	0.07	4340	0.99 (0.88-1.11)	0.87
11	rs10736466	*DRD2*	112872777	flanking 5’UTR	**A**/G	0.39	4417	1.14 (1.01-1.27)	**0.03**	4323	1.02 (0.93-1.12)	0.62

CI, confidence interval; MAF, minor allele frequency; OR, odds ratio.

[a] Minor allele is indicated in bold.

Table S4. Co-occurrence of nicotine and alcohol use at age 14. Participants have been classified at age 14 based on alcohol use as 1) non-users (never, merely tasted or had consumed alcohol occasionally) and 2) regular users (monthly or more often).

	Alcohol non-users	**Regular users**	
	N (%)	**N (%)**	**df / χ^2/ *P* Value**
Non-smokers	3,766 (99.6)	15 (0.4)	
Occasional smokers	451 (92.6)	36 (7.4)	2 / 749.7 / < **0.001**
Regular smokers	200 (72.2)	77 (27.8)	

Bolded value indicates significance.

Table S5. Co-occurrence of nicotine and alcohol use at age 31. Alcohol users at age 31 years are categorized based on alcohol daily intake (g/day).

	Alcohol non-users (0 g/day) N (%)	**Alcohol-users, 1st tertile (0-2.4 g/day) N (%)**	**Alcohol-users, 2nd tertile (2.5-8.3 g/day) N (%)**	**Alcohol-users, 3rd tertile (≥8.4 g/day) N (%)**	**df / χ^2/ *P* Value**
Non-smokers	340 (13.4)	898 (35.5)	788 (31.1)	507 (20.0)	
Light smokers	54 (5.8)	236 (25.6)	301 (32.6)	332 (36.0)	6 / 453.9 / <**0.001**
Heavy smokers	30 (3.3)	152 (16.8)	230 (25.4)	494 (54.5)	

Bolded value indicates significance.

Table S6. Association between alcohol use at age 14/31 years and genotype at *TTC12*-rs10502172 and *CHRNA3*-rs1051730. Analyses are performed by using logistic regression, allelic effect is modeled as additive (df = 1), covariates included in the regression models are: gender, socioeconomic status, maternal smoking, principal components to correct for population stratification.

	Alcohol Use at 14 years (Users vs. non-users)	
	OR (95% CI)	**χ^2/ *P* Value**
TTC12-rs10502172[G]	1.08 (0.83-1.40)	0.33 / 0.56
CHRNA3-rs1051730[A]	1.21 (0.92-1.58)	1.89 / 0.17

	Alcohol Use at 31 years					
	1st tertile vs. non-users		**2nd tertile vs. non-users**		**3rd tertile vs. non-users**	
	OR (95% CI)	**χ^2/ *P* Value**	**OR (95% CI)**	**χ^2/ *P* Value**	**OR (95% CI)**	**χ^2/ *P* Value**
TTC12-rs10502172[G]	1.01 (0.86-1.18)	0.02 / 0.89	1.05 (0.89-1.22)	0.31 / 0.58	1.12 (0.95-1.31)	1.85 / 0.17
CHRNA3-rs1051730[A]	1.10 (0.93-1.30)	1.13 / 0.29	1.12 (0.95-1.33)	1.79 / 0.18	1.10 (0.93-1.31)	1.29 / 0.26

Table S7. Association between smoking behavior in adulthood and each genetic marker in the 15q25 candidate chromosome region. Heavy smokers (N = 932) are compared to light/non-smokers (N = 2,637) by using logistic regression. Participants who were not smoking at age 31 or one year before, but who used to smoke in the past, have been excluded from these analyses. Allelic effect at each locus is modelled as additive (df = 1). Covariates entered in the regression models were: gender, socioeconomic status, maternal smoking, and principal components to correct for population stratification. In bold, $P < 0.05$; thick lines indicate boundaries between haplotype-blocks.

							Smoking at 31 years		
Chr	**rs number**	**Gene**	**Position**	**Location**	**Risk/other allele**[a]	**MAF**	**N**	**OR (95% CI)**	***P* Value**
15	rs8034191	*LOC123688*	76593078	Intron	**G**/A	0.33	3458	0.77 (0.68-0.86)	**1.33x10^{-5}**
15	rs3885951	*LOC123688*	76612972	flanking 3'UTR	**G**/A	0.06	3455	0.81 (0.64-1.01)	0.07
15	rs2036534	*LOC123688*	76614003	flanking 3'UTR	A/**G**	0.28	3450	0.85 (0.74-0.96)	**0.01**
15	rs6495306	*CHRNA5*	76652948	Intron	A/**G**	0.38	3454	0.90 (0.80-1.02)	0.09
15	rs680244	*CHRNA5*	76658343	Intron	G/**A**	0.38	3456	0.90 (0.80-1.02)	0.09
15	rs621849	*CHRNA5*	76659916	Intron	A/**G**	0.38	3449	1.11 (0.98-1.25)	0.09
15	rs1051730	*CHRNA3*	76681394	Coding	**A**/G	0.32	3456	1.31 (1.16-1.48)	**1.03x10^{-5}**
15	rs6495309	*CHRNB4*	76702300	flanking 3'UTR	G/**A**	0.27	3456	1.21 (1.07-1.38)	**0.003**
15	rs1948	*CHRNB4*	76704454	3'UTR	G/**A**	0.34	3452	1.11 (0.98-1.25)	0.10
15	rs950776	*CHRNB4*	76713073	Intron	A/**G**	0.33	3455	0.87 (0.77-0.98)	0.02
15	rs12594247	*CHRNB4*	76733688	flanking 5'UTR	**A**/G	0.21	3457	1.02 (0.89-1.17)	0.78
15	rs12900519	*CHRNB4*	76736182	flanking 5'UTR	A/**G**	0.14	3458	1.03 (0.88-1.22)	0.70
15	rs1996371	*CHRNB4*	76743861	flanking 5'UTR	**G**/A	0.35	3453	0.83 (0.74-0.94)	**0.002**
15	rs6495314	*CHRNB4*	76747584	flanking 5'UTR	**C**/A	0.35	3453	1.20 (1.07-1.35)	**0.002**
15	rs8032156	*CHRNB4*	76751553	flanking 5'UTR	G/**A**	0.30	3410	1.02 (0.90-1.15)	0.80
15	rs8038920	*CHRNB4*	76761600	flanking 5'UTR	G/**A**	0.27	3454	0.88 (0.77-1.00)	0.06
15	rs4887077	*CHRNB4*	76765419	flanking 5'UTR	**A**/G	0.33	3455	0.81 (0.72-0.91)	**0.001**
15	rs11638372	*CHRNB4*	76770614	flanking 5'UTR	**A**/G	0.33	3456	1.24 (1.10-1.40)	**0.0005**

CI, confidence interval; MAF, minor allele frequency; OR, odds ratio.

[a] Minor allele is indicated in bold.

Table S8. Association between smoking behavior in adulthood and each genetic marker in the 11q23 chromosome region. Smokers (light + heavy smokers, N = 1,879) are compared to non-smokers (N = 1,690) by using logistic regression. Participants who were not smoking at age 31 or one year before, but who used to smoke in the past, have been excluded from these analyses. Allelic effect at each locus is modelled as additive (df = 1). Covariates entered in the regression models were: gender, socioeconomic status, maternal smoking, and principal components to correct for population stratification. In bold, $P < 0.05$; thick lines indicate boundaries between haplotype-blocks.

							Smoking at 31 years		
Chr	**rs number**	**Gene**	**Position**	**Location**	**Risk/other allele[a]**	**MAF**	**N**	**OR (95% CI)**	***P* Value**
11	rs4517559	*TTC12*	112685412	flanking 5'UTR	**G**/A	0.49	3455	1.03 (0.93-1.13)	0.61
11	rs2236709	*TTC12*	112691986	Intron	**G**/A	0.30	3456	1.00 (0.90-1.12)	0.98
11	rs7927508	*TTC12*	112693407	Intron	**G**/A	0.49	3455	1.02 (0.93-1.13)	0.63
11	rs2156486	*TTC12*	112694133	Intron	A/**C**	0.19	3457	1.01 (0.89-1.14)	0.91
11	rs723077	*TTC12*	112699378	Coding	**C**/A	0.43	3449	1.03 (0.93-1.14)	0.55
11	rs10502172	*TTC12*	112704356	Intron	G/**A**	0.46	3443	1.18 (1.07-1.30)	**0.001**
11	rs2303380	*TTC12*	112705919	Intron	**G**/A	0.30	3454	1.17 (1.05-1.31)	**0.004**
11	rs4987094	*TTC12*	112721133	Intron	**A**/G	0.20	3457	1.05 (0.93-1.19)	0.40
11	rs2276070	*TTC12*	112735810	Intron	**A**/G	0.24	3455	1.03 (0.92-1.16)	0.56
11	rs719802	*TTC12*	112739889	Intron	**A**/G	0.32	3454	1.12 (1.01-1.25)	**0.03**
11	rs719804	*TTC12*	112739985	Intron	**G**/A	0.17	3444	1.16 (1.02-1.33)	**0.02**
11	rs754672	*TTC12*	112754346	flanking 3'UTR	G/**A**	0.46	3453	1.15 (1.05-1.27)	**0.004**
11	rs877138	*ANKK1*	112761718	flanking 5'UTR	**G**/A	0.30	3458	1.15 (1.04-1.28)	**0.01**
11	rs4590907	*ANKK1*	112768580	Intron	**C**/A	0.24	3453	1.01 (0.90-1.14)	0.83
11	rs7118900	*ANKK1*	112772031	Coding	**A**/G	0.16	3452	1.07 (0.93-1.23)	0.33
11	rs2734849	*ANKK1*	112775370	Coding	A/**G**	0.46	3435	1.15 (1.04-1.27)	0.01
11	rs1800497	*ANKK1*	112776038	Coding	**A**/G	0.17	3414	1.09 (0.95-1.24)	0.23
11	rs2242592	*DRD2*	112784640	flanking 3'UTR	**G**/A	0.37	3449	1.09 (0.98-1.20)	0.12
11	rs1076563	*DRD2*	112801119	Intron	A/**C**	0.50	3457	1.11 (1.00-1.22)	0.05
11	rs2471857	*DRD2*	112803549	Intron	**A**/G	0.17	3455	1.07 (0.94-1.22)	0.32
11	rs4620755	*DRD2*	112814829	Intron	**A**/G	0.23	3456	1.00 (0.89-1.12)	0.97
11	rs7125415	*DRD2*	112815891	Intron	**A**/G	0.19	3456	1.00 (0.88-1.13)	0.99
11	rs4648318	*DRD2*	112818599	Intron	**G**/A	0.34	3456	1.05 (0.95-1.17)	0.32
11	rs4274224	*DRD2*	112824662	Intron	G/A	0.24	3455	1.03 (0.92-1.15)	0.64
11	rs4581480	*DRD2*	112829684	Intron	**G**/A	0.07	3454	1.09 (0.90-1.33)	0.37
11	rs7131056	*DRD2*	112834984	Intron	**C**/A	0.49	3436	0.97 (0.88-1.07)	0.55
11	rs4938019	*DRD2*	112846601	Intron	**G**/A	0.23	3453	1.00 (0.89-1.12)	0.99

11	rs12364283	*DRD2*	112852165	flanking 5'UTR	**G**/A	0.08	4316	1.03 (0.86-1.24)	0.73
11	rs10891556	*DRD2*	112857971	flanking 5'UTR	**A**/C	0.24	4331	1.00 (0.89-1.12)	0.97
11	rs6589377	*DRD2*	112860946	flanking 5'UTR	A/**G**	0.17	4340	1.00 (0.87-1.13)	0.94
11	rs10736466	*DRD2*	112872777	flanking 5'UTR	**A**/G	0.39	4323	0.99 (0.89-1.09)	0.83

CI, confidence interval; MAF, minor allele frequency; OR, odds ratio.

[a] Minor allele is indicated in bold.

Table S9. Association between *TTC12*-rs10502172, *CHRNA3*-rs1051730, maternal smoking, socio-economic status, and gender with smoking behavior at 31 years. Participants who were not smoking at age 31 or one year before, but who used to smoke in the past, have been excluded from this analysis. Analyses are computed by using multinomial logistic regression. Non-smokers are the reference category. Effect sizes for each variable are adjusted for the effect of all other variables listed in the table and for principal components to correct for population stratification. Genetic effect is estimated assuming an additive model of number of risk alleles for each locus.

	Smoking at 31 yrs					
	Light-smokers (N = 947) vs. non-smokers (N = 1,690)			**Heavy-smokers (N = 932) vs. non-smokers (N = 1,690)**		
	OR	**(95% CI)**	***P* Value**	**OR**	**(95% CI)**	***P* Value**
***TTC12*-rs10502172[G]**	1.17	(1.04-1.31)	**0.01**	1.20	(1.06-1.36)	**0.004**
***CHRNA3*-rs1051730[A]**	0.95	(0.84-1.08)	0.43	1.29	(1.13-1.47)	**<0.001**
Sex						
Female (N = 2,476)	1			1		
Male (N = 2,286)	0.95	(0.80-1.14)	0.59	2.99	(2.47-3.61)	**<0.001**
Maternal smoking						
Non-smokers (N = 4,000)	1			1		
Light-smokers (N = 554)	1.24	(0.97-1.61)	0.09	1.73	(1.34-2.22)	**<0.001**
Heavy-smokers (N = 96)	1.08	(0.59-2.01)	0.80	1.78	(0.99-3.20)	**0.05**
Socio-economic status						
Professionals (N = 1,098)	1			1		
Skilled workers (N = 1,481)	1.68	(1.34-2.12)	**<0.001**	1.74	(1.34-2.25)	**<0.001**
Unskilled workers (N = 1,215)	2.33	(1.83-2.98)	**<0.001**	3.80	(2.96-4.86)	**<0.001**
Farmers (N = 164)	0.86	(0.52-1.41)	0.54	0.72	(0.41-1.25)	0.24
Others (N = 747)	1.82	(1.39-2.39)	**<0.001**	2.50	(1.88-3.23)	**<0.001**

Bolded values indicate significance.

Table S10. Direct effects of *TTC12*-rs10502172, *CHRNA3*-rs1051730, socio-economic status, maternal smoking during pregnancy, and gender on smoking at 14 and 31 years, and novelty seeking.

Variable	**β (SE)**[a]	**Standardized β**[a]	***P* Value**
Maternal smoking during pregnancy			
Family socio-economic status during pregnancy			
Professionals	ref.		
Skilled workers	0.24 (0.07)	0.24	**<0.001**
Unskilled workers	0.23 (0.09)	0.22	**0.01**
Farmers	-0.15 (0.10)	-0.15	0.14
Mother's marital status during pregnancy			
Married	ref.		
Unmarried	0.50 (0.13)	0.47	**<0.001**
Smoking at 14 years			
TTC12-rs10502172[G]	0.15 (0.03)	0.10	**<0.001**
CHRNA3-rs1051730[A]	0.06 (0.03)	0.04	0.07
Maternal smoking during pregnancy			
Non-smoker	ref.		
Smoker	0.12 (0.04)	0.36	**<0.001**
Gender			
Male	ref.		
Female	0.07 (0.05)	0.07	0.16
Family socio-economic status at 14 years			
Professionals	ref.		
Skilled workers	0.15 (0.07)	0.14	**0.02**
Unskilled workers	0.24 (0.07)	0.23	**0.001**
Farmers	0.00 (0.10)	0.00	0.97
Smoking at 31 years			
TTC12-rs10502172[G]	0.02 (0.04)	0.01	0.63
CHRNA3-rs1051730[A]	0.13 (0.04)	0.07	**<0.001**
Maternal smoking during pregnancy			
Non-smoker	ref.		
Smoker	0.10 (0.04)	0.23	**0.01**
Smoking at 14 years			
Non-smoker	ref.		
Smoker	0.44 (0.04)	0.97	<**0.001**
Gender			
Male	ref.		
Female	-0.79 (0.06)	-0.64	**<0.001**
Socio-economic status at 31			
Professionals	ref.		
Skilled workers	0.20 (0.08)	0.17	**0.01**
Unskilled workers	0.59 (0.08)	0.48	**<0.001**
Farmers	-0.10 (0.17)	-0.08	0.56
Others	0.41 (0.09)	0.33	**<0.001**
Novelty seeking	0.03 (0.01)	0.12	**<0.001**
Novelty seeking			
TTC12- rs10502172[G]	0.28 (0.14)	0.03	**0.04**
CHRNA3-rs1051730[A]	-0.04 (0.14)	0.00	0.78
Gender			
Male	ref.		
Female	1.29 (0.20)	0.22	**<0.001**
Novelty seeking and smoking at 14 (correlation)	0.95 (0.14)	0.16	**<0.001**

[a] The estimates are probit regression estimates from the SEM model fitted to the data.

Supplemental Figures

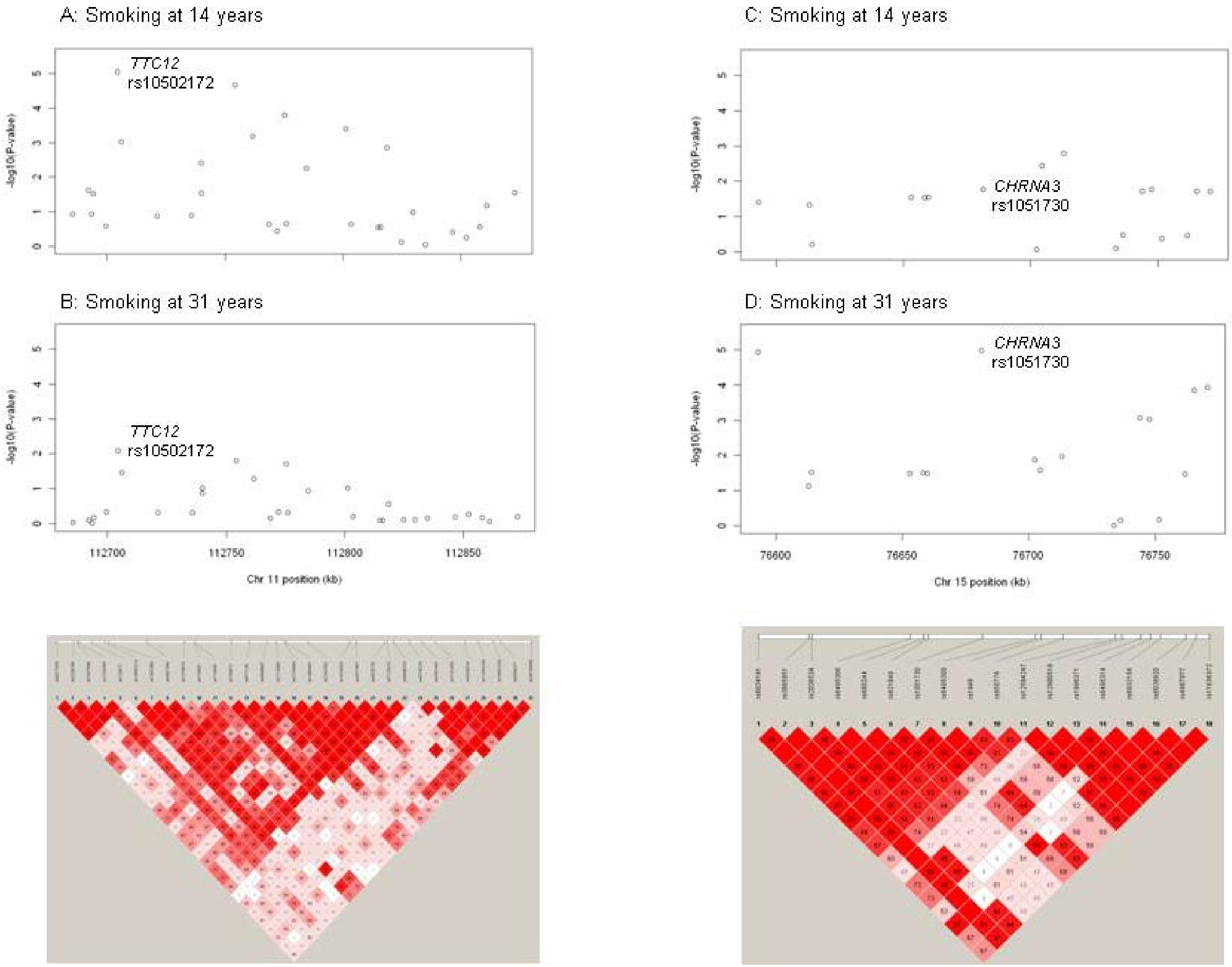

Figure S1. Association between smoking behavior during adolescence and adulthood and genetic variation in the two candidate chromosome regions on 11q23 and 15q25. Linkage disequilibrium structure of the region is also shown.

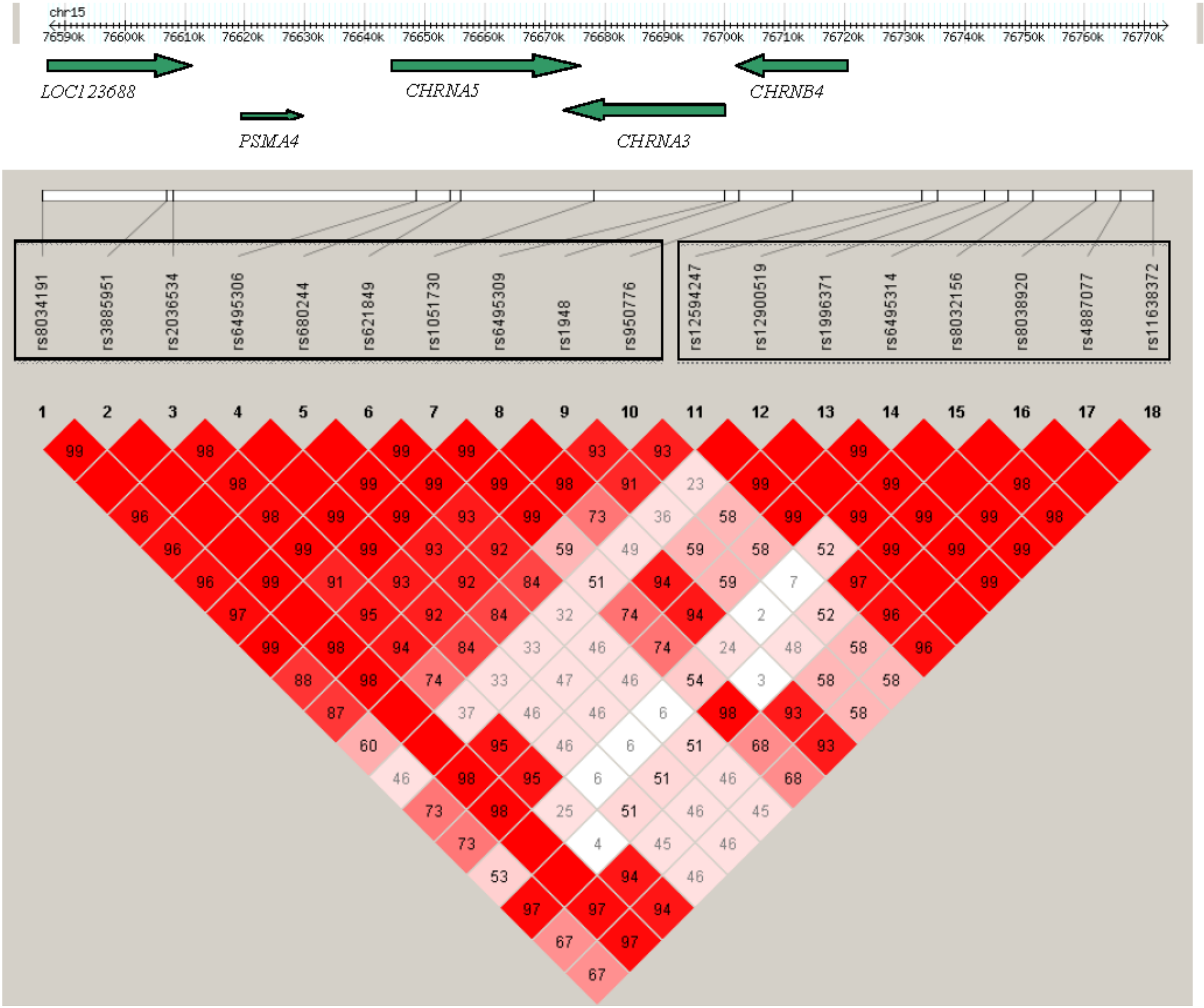

Figure S2. Haplotype-block structure of the chromosome 15q25 region. Linkage disequilibrium pattern coded according to the following color scheme: LOD < 2 and D’ < 1: white; LOD < 2 and D’ = 1: blue; LOD ≥ 2 and D’ < 1: shades of red; LOD ≥ 2 and D’ = 1: bright red.

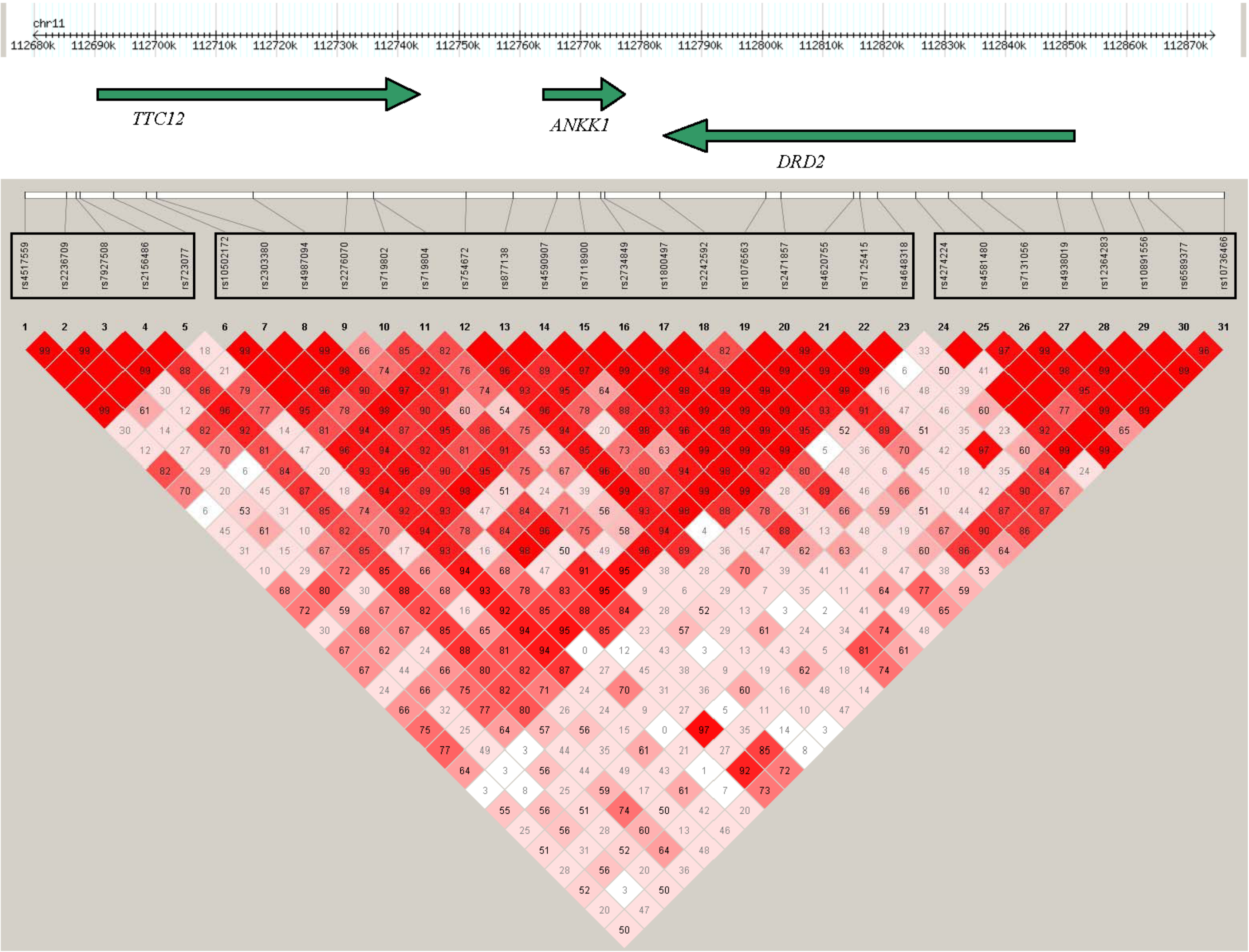

Figure S3. Haplotype-block structure of the chromosome 11q23 region. Linkage disequilibrium pattern coded according to the following color scheme: LOD < 2 and D’ < 1: white; LOD < 2 and D’ = 1: blue; LOD ≥ 2 and D’ < 1: shades of red; LOD ≥ 2 and D’ = 1: bright red.

Supplemental References

1. Jarvelin MR, Sovio U, King V, Lauren L, Xu B, McCarthy MI, *et al.* (2004): Early life factors and blood pressure at age 31 years in the 1966 northern Finland birth cohort. *Hypertension*. 44:838-846.
2. Cloninger CR, Svrakic DM, Przybeck TR (1993): A psychobiological model of temperament and character. *Arch Gen Psychiatry*. 50:975-990.
3. Cloninger CR (1987): A systematic method for clinical description and classification of personality variants. A proposal. *Arch Gen Psychiatry*. 44:573-588.
4. Rantakallio P (1983): Family background to and personal characteristics underlying teenage smoking. Background to teenage smoking. *Scand J Soc Med*. 11:17-22.
5. Poikolainen K, Vartiainen E (1997): Determinants of gamma-glutamyltransferase: positive interaction with alcohol and body mass index, negative association with coffee. *Am J Epidemiol*. 146:1019-1024.
6. Barrett JC, Fry B, Maller J, Daly MJ (2005): Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 21:263-265.
7. Menozzi P, Piazza A, Cavalli-Sforza L (1978): Synthetic maps of human gene frequencies in Europeans. *Science*. 201:786-792.
8. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006): Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 38:904-909.
9. Patterson N, Price AL, Reich D (2006): Population structure and eigenanalysis. *PLoS Genet*. 2:e190.
10. Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, *et al.* (2009): Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet*. 41:35-46.