Smoking and Genetic Risk Variation across Populations of European, Asian, and African-American Ancestry - A Meta-analysis of Chromosome 15q25

Li-Shiun Chen[1], **Nancy L. Saccone**[2], **Robert C. Culverhouse**[3], **Paige M. Bracci**[4], **Chien-Hsiun Chen**[5,6], **Nicole Dueker**[7], **Younghun Han**[8], **Hongyan Huang**[9], **Guangfu Jin**[10], **Takashi Kohno**[11], **Jennie Z. Ma**[12], **Thomas R. Przybeck**[1], **Alan R. Sanders**[13], **Jennifer A. Smith**[14], **Yun Ju Sung**[9], **Angie S. Wenzlaff**[15], **Chen Wu**[16], **Dankyu Yoon**[17,18], **Ying-Ting Chen**[5], **Yu-Ching Cheng**[19], **Yoon Shin Cho**[20], **Sean P. David**[21,22,23], **Jubao Duan**[13], **Charles B. Eaton**[23], **Helena Furberg**[24], **Alison M. Goate**[1], **Dongfeng Gu**[25,26], **Helen M. Hansen**[27], **Sarah Hartz**[1], **Zhibin Hu**[10], **Young Jin Kim**[17,21], **Steven J. Kittner**[28], **Douglas F. Levinson**[29], **Thomas H. Mosley**[14], **Thomas J. Payne**[30], **DC Rao**[9], **John P. Rice**[1], **Treva K. Rice**[9], **Tae-Hwi Schwantes-An**[2], **Sanjay S. Shete**[8], **Jianxin Shi**[31], **Margaret R. Spitz**[8], **Yan V. Sun**[14], **Fuu-Jen Tsai**[32], **Jen C. Wang**[1], **Margaret R. Wrensch**[27], **Hong Xian**[3], **Pablo V. Gejman**[13], **Jiang He**[33], **Steven C. Hunt**[34], **Sharon L. Kardia**[14], **Ming D. Li**[35], **Dongxin Lin**[16], **Braxton D. Mitchell**[20], **Taesung Park**[36], **Ann G. Schwartz**[15], **Hongbing Shen**[10], **John K. Wiencke**[27], **Jer-Yuarn Wu**[5,6], **Jun Yokota**[37], **Christopher I. Amos**[8], and **Laura J. Bierut**[1]

[1]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, United States of America

[2]Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, United States of America

[3]Department of Internal Medicine, Washington University School of Medicine, St. Louis, Missouri, United States of America

[4]Department of Epidemiology and Biostatistics, UCSF, San Francisco, California, United States of America

[5]National Genotyping Center, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

[6]Graduate Institute of Chinese Medical Science, China Medical University, Taichung, Taiwan

[7]Department of Epidemiology and Public Health, University of Maryland, Baltimore, Maryland, United States of America

[8]Department of Epidemiology, University of Texas M.D. Anderson Cancer Center, Houston, Texas, United States of America

[9]Division of Biostatistics, Washington University School of Medicine, St. Louis, Missouri, United States of America

[10]Department of Epidemiology and Biostatistics, Nanjing Medical University, Nanjing, China

[11]Division of Genome Biology, National Cancer Center Research Institute 1-1, Tokyo, Japan

[12]Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia, United States of America

Corresponding Author: Li-Shiun Chen, M.D, M.P.H. Sc.D, Department of Psychiatry (Box 8134), Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110, (314) 362-3932, chenli@psychiatry.wustl.edu.

[13]Department of Psychiatry and Behavioral Sciences, North Shore University Health System Research Institute, and University of Chicago, Chicago, Illinois, United States of America

[14]Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan, United States of America

[15]Karmanos Cancer Institute, Wayne State University, Detroit, Michigan, United States of America

[16]Department of Etiology & Carcinogenesis, Cancer Institute, Chinese Academy of Medical Sciences, Beijing, China

[17]Interdisciplinary Program in Bioinformatics, College of Natural Science, Seoul National University, Seoul, Korea

[18]Center for Immunology and Pathology, National Institute of Health, Seoul, Korea

[19]Department of Medicine, University of Maryland Medical Center, Baltimore, Maryland, United States of America

[20]Center for Genome Science, National Institute of Health, Seoul, Korea

[21]Center for Health Sciences, SRI International, Menlo Park, California, United States of America

[22]Department of Medicine, Stanford University School of Medicine, California, United States of America

[23]Department of Family Medicine, Brown University, Providence, Rhode Island, United States of America

[24]Department of Epidemiology, Memorial Sloan Kettering Cancer Center, New York, New York, United States of America

[25]Fuwai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[26]Chinese National Center for Cardiovascular Disease Control and Research, Beijing, China

[27]Neurological Surgery, Division of Epidemiology, Helen Diller Family Cancer Center, San Francisco, California, United States of America

[28]Department of Neurology, University of Maryland, Baltimore, Maryland, United States of America

[29]Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, California, United States of America

[30]University of Mississippi Medical Center, Jackson, Mississippi, United States of America

[31]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, United States of America

[32]School of Post-Baccalaureate Chinese Medicine, China Medical University, Taiwan

[33]Department of Epidemiology, Tulane School of Public Health and Tropical Medicine, New Orleans, Louisiana, United States of America

[34]Department of Internal Medicine, University of Utah, Salt Lake City, Utah, United States of America

[35]Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, Virginia, United States of America

[36]Department of Statistics, College of Natural Science, Seoul National University, Seoul, Korea

[37]Division of Multistep Carcinogenesis, National Cancer Center Research Institute 1-1, Tokyo, Japan

Abstract

Recent meta-analyses of European ancestry subjects show strong evidence for association between smoking quantity and multiple genetic variants on chromosome 15q25. This meta-analysis extends the examination of association between distinct genes in the *CHRNA5-CHRNA3-CHRNB4* region and smoking quantity to Asian and African American populations to confirm and refine specific reported associations.

Association results for a dichotomized cigarettes smoked per day (CPD) phenotype in 27 datasets (European ancestry (N=14,786), Asian (N=6,889), and African American (N=10,912) for a total of 32,587 smokers) were meta-analyzed by population and results were compared across all three populations.

We demonstrate association between smoking quantity and markers in the chromosome 15q25 region across all three populations, and narrow the region of association. Of the variants tested, only rs16969968 is associated with smoking (p < 0.01) in each of these three populations (OR=1.33, 95%C.I.=1.25–1.42, p=1.1×10^{-17} in meta-analysis across all population samples). Additional variants displayed a consistent signal in both European ancestry and Asian datasets, but not in African Americans.

The observed consistent association of rs16969968 with heavy smoking across multiple populations, combined with its known biological significance, suggests rs16969968 is most likely a functional variant that alters risk for heavy smoking. We interpret additional association results that differ across populations as providing evidence for additional functional variants, but we are unable to further localize the source of this association. Using the cross-population study paradigm provides valuable insights to narrow regions of interest and inform future biological experiments.

Keywords

INTRODUCTION

Recent genetic meta-analyses including tens of thousands of subjects of European ancestry show strong evidence of association between smoking quantity (cigarettes per day, CPD) and multiple genetic markers on chromosome 15q25 [Liu, et al. 2010; Saccone, et al. 2010; TAG 2010; Thorgeirsson, et al. 2010]. Those studies synthesized evidence across many independent datasets to highlight specific variants in the region of the *CHRNA5-CHRNA3-CHRNB4* gene cluster associated with smoking behavior in European ancestry subjects. It is important to determine the biological mechanisms underlying these associations; however, the high linkage disequilibrium (LD) in this region among individuals of European ancestry makes it difficult to differentiate potentially causal variants from the many correlated variants. Because the genetic architecture of chromosome 15q25 varies across populations, comparing associations across diverse populations with differing genetic architecture can help refine the region of association and point to variants more likely to have functional relevance [Rotimi and Jorde 2010; Saccone, et al. 2008; Zaitlen, et al. 2010].

The most robust genetic finding on chromosome 15q25 in subjects of European ancestry is the region tagged by rs16969968, rs1051730, and other correlated variants. This finding has been replicated for smoking-related traits in multiple distinct datasets [Baker, et al. 2009; Berrettini, et al. 2008; Keskitalo, et al. 2009; Saccone, et al. 2009; Saccone, et al. 2007;

Sherva, et al. 2008; Stevens, et al. 2008; Thorgeirsson, et al. 2008; Weiss, et al. 2008], and has now been reported as the most significant genome-wide association in recent meta-analyses of European ancestry subjects (e.g. rs16969968, p = 5.57×10^{-72}, or rs1051730, p = 2.75×10^{-73}) [Liu, et al. 2010; Saccone, et al. 2010; TAG 2010; Thorgeirsson, et al. 2010]. We will use the term "bin" to denote a group of correlated SNPs (r^2 0.7) that may constitute the same association signal in European ancestry samples [Carlson, et al. 2004]. Under this definition and using the 1000 Genomes Pilot 1 CEU as the European ancestry reference sample [Durbin, et al. 2010], the single bin tagged by rs16969968 and rs1051730 includes 52 known variants. This bin, which we will call bin A, groups together and unifies the most significant meta-analysis findings as well as individual dataset reports of SNPs associated with nicotine dependence, heavy smoking, lung cancer, and other smoking related diseases in European ancestry datasets.

There are additional markers of interest in this region that are not strongly correlated with bin A. Because of the clear association between smoking behavior and bin A, each of the large-scale meta-analyses of European ancestry samples carried out association tests conditional on bin A variants for other SNPs to determine whether additional genetic markers in 15q25 are associated after adjusting for effects of bin A [Liu, et al. 2010; Saccone, et al. 2010; TAG 2010; Thorgeirsson, et al. 2010]. After conditioning on bin A, the meta-analyses identified additional SNPs in this region associated with smoking behavior. These SNPs can be grouped into three distinct bins (B, C, D) (Table 1). Bin B, tagged by rs588765 and rs880395, is associated with genome-wide significance among heavy versus light smokers but only in analyses conditioning on bin A (p=1.2×10^{-9}) [Saccone, et al. 2010]. Notably, bin B is also associated with mRNA levels of *CHRNA5* in brain and lung [Falvella, et al. 2010; Smith, et al. 2010; Wang, et al. 2009]. Bin C, tagged by rs6495308 [Liu, et al. 2010], rs2036534 [Thorgeirsson, et al. 2010], rs7163730, rs9788682, rs684513 [TAG 2010], and rs578776 [Saccone, et al. 2010], is associated with heavy smoking after conditioning on bin A (p values from 9.1×10^{-5} to 6.3×10^{-9}). In contrast to bin B, bin C is less significant in conditional analysis compared to single SNP analysis. Bin D is represented by rs2869046, which also displayed residual association after conditioning on bin A (p=4.8×10^{-5}) [Thorgeirsson, et al. 2010]. Markers from these different bins (A, B, C, and D) are only modestly correlated with one another, with r^2 0.52 in the 1000 Genomes Pilot 1 CEU (N=180; Table 2).

Differences in the correlational structure of markers spanning the region 15q25 between populations result in distinct sub-bins of correlated markers among Asian and African American populations which provide an opportunity to refine the source of the previously reported signals. For example, bin A, consisting of 52 variants including rs16969968, separates into 20 sub-bins in Asians (based on 1000 Genomes Pilot 1 JPT/CHB) and 38 sub-bins in African Americans (based on combined information from the 1000 Genomes Pilot 1 YRI and HapMap 3 Release 2 ASW) [Altshuler, et al. 2010]. In particular, rs16969968 and rs1051730 are highly correlated in European ancestry (r^2=1) and Asian populations (r^2=1), but display only moderate correlation (r^2=0.40) in the African American population. These differences in genetic architecture can be used to dissect the association signals.

The purpose of this meta-analysis is to determine if bins A, B, C, and D show consistent association with smoking behavior across populations and, if so, to leverage these differences in genetic correlation across populations to refine the genetic associations in this region previously reported in subjects of European ancestry. We expect a sub-bin showing consistent evidence across all three populations to be more likely to contain a variant altering a biological mechanism. We performed meta-analyses of results from a total of 27 datasets: nine European ancestry samples (used to evaluate consistency with previous results), seven Asian samples, and eleven African American samples. We tested for

association between smoking phenotypes and the four distinct bins (A through D) across all three populations. This cross-population study therefore improves our understanding of genetic risk for smoking by highlighting potentially functional variants.

METHODS

Samples

Results from 27 datasets, containing a total of 32,587 smokers with measures of cigarettes per day, contributed to the meta-analyses. Of these datasets, 9 consisted of European ancestry subjects (N=14,786), 7 consisted of Asians (N=6,889), and 11 consisted of African Americans (N=10,912). Twenty datasets were samples of unrelated individuals. The remaining seven datasets were family-based studies, for which the primary analyses involved an extraction of unrelated individuals. To be included in the analyses, each subject was required to have reported smoking cigarettes in his/her lifetime. Genotyping varied among studies from extensive coverage based on genome wide association genotyping to only a limited number of candidate SNPs genotyped in this 15q25 region. Text S1 provides additional details for each dataset, including recruitment, primary phenotypes, definitions for smokers and cigarettes per day, DNA source, genotyping platforms, and genotyping quality control. Table S1 shows the sample size and demographics for each participating dataset. Four of nine datasets of European ancestry were included in the previous report [Saccone, et al. 2010] (see Table S1 for the overlap, which involves only European-ancestry samples). The informed consent from participants and approval from the appropriate institutional review boards were obtained.

Phenotypes

Smoking quantity was assessed with cigarettes smoked per day (CPD). The primary phenotype for analysis was a dichotomous trait contrasting light smoking controls (CPD 10) to heavy smoking cases (CPD>20). In addition, a 4-level ordered trait (CPD 10; 11 CPD 20; 21 CPD 30; CPD 31 coded as 0, 1, 2, 3 respectively) was developed for confirmatory analysis. The only exception was one study (Women's Health Initiative) that measured smoking amount with different threshold levels (CPD 14, 15 CPD 24, 25 CPD 34, CPD 35), and CPD 14 defined the light smoking controls which was contrasted with CPD 25 as heavy smoking cases.

Variants for Analyses

Multiple SNPs in 15q25 have been identified as associated with smoking behavior in studies of European ancestry subjects. We focused on the results highlighted in the most powerful studies, namely the large meta-analyses [Liu, et al. 2010; Saccone, et al. 2010; TAG 2010; Thorgeirsson, et al. 2010] (Table 1). Table 2 lists the 11 targeted SNPs and illustrates how linkage disequilibrium (LD, measured by r^2) structure varies across different populations. We used SNAP [Johnson, et al. 2008] with 1000 Genomes Pilot 1 reference samples [Altshuler, et al. 2010; Durbin, et al. 2010] and HapMap3 ASW to obtain LD estimates for our three populations: CEU for European ancestry, JPT/CHB for Asians, and ASW/YRI for African Americans.

It is important to examine not just the 11 previously identified SNPs listed in Table 1, but all SNPs correlated with these 11 SNPs in Europeans. We used a two-step process to define distinct groups of correlated SNPs, which we call bins. First, we grouped previously identified SNPs by their correlation in the 1000 Genomes Pilot 1 CEU (i.e. European ancestry) reference sample, using r^2 0.7 as our threshold [Durbin, et al. 2010]. Under this strategy, the 11 previously identified SNPs listed in Table 1 are partitioned into four groups: Group A (rs16969968, rs1051730), Group B (rs588765, rs880395), Group C (rs6495308,

rs2036534, rs7163730, rs9788682, rs684513, rs578776), and Group D (rs2869046) (Table 2). From these four groups, we established the bins by including all SNPs correlated (r^2 0.7) in the European ancestry reference sample with at least one of the SNPs defining the bin. The threshold of 0.7 was chosen to provide an inclusive collection of tested SNPs. Using SNAP to obtain correlated variants in a bin based on 1000 Genomes Pilot 1 CEU, we identified 52 SNPs in bin A, 111 SNPs in bin B, 82 SNPs in bin C, and 15 SNPs in bin D.

Next, we partitioned these SNPs within a bin into "sub-bins" based on r^2 0.8 in the Asian and African American populations. The higher threshold of 0.8 was used for sub-bins in the other populations to refine the focus of the analyses. In Asians, we identified 20 sub-bins for bin A, 39 sub-bins for bin B, 24 sub-bins for bin C, and 7 sub-bins for bin D. In African Americans, we identified 38 sub-bins for bin A, 37 sub-bins for bin B, 26 sub-bins for bin C, and 7 sub-bins for bin D.

Statistical Analyses and Meta-Analyses

We evaluated the genetic associations between heavy smoking and each genotyped SNP in three populations. Standardized scripts were developed centrally by the coordinating site (Washington University) for analyses of all participating datasets at each individual research center. Results were returned to the coordinating site for quality checks and meta-analyses. Individual SNP analyses were performed using SAS (SAS Institute, Cary, NC).

In each dataset, association between heavy versus light smoking based on cigarettes per day and all SNPs was evaluated with logistic regression models as the primary analysis. Genotypes were coded additively as the number of non-reference alleles, where the reference allele was defined as the major allele in the European ancestry population in dbSNP [Sherry, et al. 2001]; consistency of allelic coding was confirmed by comparing allele labels and allele frequencies across all datasets within each population. Age as a continuous variable and gender were included as covariates. Secondary analyses of the 4-level cigarettes per day trait used linear regression models with the same covariates, assuming that the trait has a simple linear relationship with the predictors.

Analyses were stratified by ancestry: European, Asian, and African American. We evaluated the effect of each bin A SNP using single SNP association analyses. For bins B, C, and D, both single SNP association and conditional analyses controlling for bin A were performed. Analyses conditional on bin A (rs16969968) served as our primary analysis model for bins B, C, and D because they were targeted due to previously reported results of analyses conditional on bin A in European ancestry meta-analyses.

For each ancestry group, every dataset with at least one genotyped SNP in a given sub-bin contributed to the meta-analysis of that sub-bin. For each sub-bin, a SNP was selected as the target. In samples where the target SNP was missing, we used the results from the SNP with highest correlation (r^2) with the target SNP in the sub-bin defined by the 1000 Genome Pilot 1 JPT/CHB for Asians, and the 1000 Genome Pilot 1 YRI or HapMap3 ASW project for African Americans.

We used PLINK to perform meta-analyses and generate overall summary odds ratios (OR), standard errors, and p values [Purcell, et al. 2007]. The R package, rmeta, was used to confirm results and generate meta-analysis plots. Meta-analyses results were based on fixed effect models to determine the evidence for association within our collected samples, so we are not making a general inference about what might be observed in other samples.

Multiple Test Correction

Our primary analysis was to determine if any intersecting sub-bins across Asians and African Americans would display evidence of consistent association when comparing heavy versus light smokers, where we defined a consistent association as having the same direction and p value < 0.01 in both populations. Our binning strategy resulted in 100 single sub-bin tests and 67 conditional association tests across the four bins: a total of 167 tests. Because the probability of any particular test resulting in a p-value < 0.01 in both non-European populations by chance would be 0.0001 (=0.01×0.01), results consistently associated in both populations would remain significant after Bonferroni correction (167×0.0001< 0.05).

RESULTS

Genetic associations in bin A

Bin A (tagged by rs16969968 and rs1051730 in Europeans) includes 52 SNPs correlated (r^2 0.7) in the European ancestry reference sample. This bin separates into 20 sub-bins in Asian populations and 38 sub-bins in African American populations. We had adequate coverage to test 9 of these 20 sub-bins in Asian data and 27 of these 38 sub-bins in African American data.

We detected a strong association between the dichotomous phenotype of heavy smoking versus light smoking and bin A in European ancestry data (OR=1.31, 95% C.I.=1.22–1.40, p=1.3×10^{-14}). The only sub-bin showing consistent association with heavy smoking across the other two populations is tagged by rs16969968 (Asian population: OR=1.64, 95%C.I.=1.15–2.32, p=5.8×10^{-3}; African American population: OR=1.62, 95% C.I.=1.21–2.17, p=1.1×10^{-3}). As noted in the Methods, because the probability of any particular test resulting in a p<0.01 in both populations by chance alone would be 0.0001, this result of consistent association in both populations remained significant after Bonferroni correction.

Figure 1 shows all SNPs in bin A, and the only consistently associated sub-bins (p<0.01 in both Asians and African Americans). Bin A variants span six genes in the European ancestry population, the sub-bin tagged by rs16969968 in the Asian population spans three genes, and the sub-bin tagged by rs16969968 in the African American population spans only one gene (CHRNA5). Figure 2 provides a forest plot summary of the stratified meta-analyses for the bin/sub-bin tagged by rs16969968, the only consistent association for bin A, in all three populations. Each plot lists ORs for each contributing sample. The overall cross-population meta-analysis across all datasets gave an OR of 1.33 (95%C.I.=1.25–1.42, p=1.1×10^{-17}).

In European and Asian populations, rs16969968 and rs1051730 are highly correlated. However, due to the different LD structure in African Americans, rs16969968 and rs1051730 represent two different sub-bins (r^2=0.40 in HapMap 3 Release 2 ASW). In our analysis of African Americans, there is stronger evidence of association between the dichotomous phenotype heavy smoking versus light smoking and the sub-bin tagged by rs16969968 (OR=1.62, 95% C.I.=1.21–2.17, p=1.1×10^{-3}), compared to the sub-bin tagged by rs1051730 (OR=1.15, 95%C.I.=1.03–1.28, p=1.1×10^{-2}). This stronger finding is seen despite the lower minor allele frequency (MAF) and much smaller available sample and for rs16969968 (MAF =0.06, 667 cases/1140 controls) compared to that for rs1051730 (MAF=0.12, 1712 cases/5640 controls).

For bin A, no tested sub-bin other than the one tagged by rs16969968 shows consistent association across populations. The meta-analyzed genetic associations between all available constituent sub-bins and heavy smoking are shown in Table S2.

Genetic associations in bin B

Bin B (tagged by rs588765 and rs880395 in Europeans) includes 111 SNPs correlated (r^2 0.7) in the European ancestry reference sample, which was partitioned into 39 sub-bins in Asian and 37 sub-bins in African American ancestry reference samples. We had adequate coverage to test 10 of these 39 sub-bins in Asian samples and 22 of these 37 sub-bins in African American samples. Consistent with the previous report [Saccone, et al. 2010] which used some of these same data (see Table S1 for the overlap, which involves only European-ancestry samples), we find that in European ancestry samples, bin B is associated (OR=1.27, 95%C.I.=1.16–1.38, p=8.7×10^{-8}) with heavy smoking in conditional analyses with rs16969968; bin B is not associated in single SNP analyses (OR=1.0, 95%C.I.=0.94–1.07, p=0.99). In Asian samples, testing for SNP association conditioning on rs16969968 show an association between heavy smoking and bin B, with the strongest result for the sub-bin tagged by rs514743 (OR=1.30, 95% C.I.= 1.07–1.58, p=9.7×10^{-3}), which is similar to the single SNP test (OR=1.28, 95% C.I.=1.05–1.56, p=0.014). In African American subjects, there is a trend of association for the same sub-bin in conditional association (OR=1.16, 95%C.I.=0.99–1.36, p=0.064) (Table 3), compared to the single SNP association (OR=1.05, 95%C.I.=0.96–1.15, p=0.24). Thus, we found evidence of association in the Asian samples consistent with the association observed in the samples of European ancestry, but only a trend toward association in the African American subjects. The meta-analyzed conditional and single SNP associations between these constituent sub-bins and heavy smoking are shown in Tables S3 and S6.

Genetic associations in bin C

Bin C (tagged by rs6495308, rs2036534, rs7163730, rs9788682, rs684513, and rs578776 in Europeans) includes 82 SNPs correlated (r^2 0.7) in the European ancestry reference sample, which was partitioned into 24 sub-bins in Asian and 26 sub-bins in African American reference samples. We had adequate coverage to test 12 of these 24 sub-bins in Asian samples and 19 of these 26 sub-bins in African American samples. Consistent with the previous studies [Liu, et al. 2010; Saccone, et al. 2010; TAG 2010; Thorgeirsson, et al. 2010], in European ancestry samples there is an association between heavy smoking and bin C (OR=0.79, 95% C.I.=0.72–0.86, p= 2.5×10^{-7}) in association tests conditioning on rs16969968, as well as an association in a single SNP analysis (OR=0.77, 95% C.I.=0.71–0.83, p= 4.0×10^{-11}).

Neither the Asian nor African American populations provide strong evidence of association with heavy smoking in any tested sub-bin in bin C under conditional association tests (all p>0.01). In the Asian data, the strongest single SNP signal was the sub-bin tagged by rs6495308 (OR=0.83, 95%C.I.= 0.72–0.96, p=9.8×10^{-3}). In the African American data, there was no evidence of consistent association in either single SNP or conditional analyses (p>0.01) for the sub-bin tagged by rs6495308 or any other sub-bin. The meta-analyzed conditional and single SNP associations between tested sub-bins and heavy smoking are shown in Tables S4 and S7.

Genetic associations in bin D

Bin D (tagged by rs2869046 in Europeans) includes 15 SNPs correlated (r^2 0.7) in the European ancestry reference sample, which was partitioned into 7 sub-bins in Asians and 7 sub-bins in African Americans. We had adequate coverage to test 2 of these 7 sub-bins in Asian samples and 3 of the 7 sub-bins in African American samples. We found no evidence of association between bin D and heavy smoking in European ancestry data, or across populations in single SNP or conditional association analyses (p>0.1). The meta-analyzed genetic associations between available sub-bins and heavy smoking conditional and single

SNP associations are shown in Tables S5 and S8. All bins were tested in secondary analyses using the 4 level phenotype measured by cigarettes per day and results were similar.

DISCUSSION

This collaborative genetic meta-analysis of smoking behavior is the first to show consistent association in the chromosome 15q25 region with heavy smoking, across samples representing three genetically distinct populations – European ancestry, Asian, and African American. Previous meta-analyses examined only European ancestry data to definitively identify associations between chromosome 15q25 and smoking behavior. Smaller individual studies of Asians and African Americans have previously examined this region for association with smoking and related phenotypes. Smoking quantity has been reported as associated with variants correlated with rs16969968 in subjects of Asian and African American descent [Amos, et al. 2010; Li, et al. 2005; Li, et al. 2010; Saccone, et al. 2009; Schwartz, et al. 2010; Shiraishi, et al. 2009; Wu, et al. 2009]. Our meta-analysis synthesizes reported findings of individual SNP associations and compares genetic associations across multi-population samples to take the correlations between genetic variants within each population into account. Our meta-analysis strengthens the evidence of association between the specific SNP rs16969968 in bin A and heavy smoking across these diverse populations.

The strongest association signal seen in this gene cluster in European ancestry populations is represented by a group of 52 correlated variants, including rs16969968, which we call bin A. Due to these high correlations, the ability to statistically refine the association between smoking and these SNPs is very limited when using only European ancestry subjects. However, the LD structure between these 52 variants breaks down into 20 sub-bins in Asians and 38 sub-bins in African Americans.

By requiring consistent genetic effects across the three populations, we can refine a genetic association to variants that are more likely to reflect potential functional variants. Two SNPs in bin A are the most frequently reported from previous meta-analyses of smoking behavior in European ancestry subjects: rs16969968 and rs1051730. They are highly correlated (r^2=1) in European ancestry and Asian populations, but display only modest correlation in African Americans (r^2=0.40) (HapMap 3 Release 2 ASW). We can leverage this difference in LD architecture to differentiate the association of heavy smoking with these two variants.

In our meta-analysis of African-Americans, rs16969968 is more strongly associated with heavy smoking (OR=1.62, 95%C.I.=1.21–2.17, p=0.0011, N=1807) than rs1051730 (OR=1.15, 95%C.I.=1.03–1.28, p=0.011, N=7352). SNP rs16969968 is the most strongly associated polymorphism across all three populations and the only variant meeting the consistent association threshold in our study. In addition, SNP rs16969968 causes an amino acid change in the nicotinic receptor α5 subunit and alters function of its receptor [Bierut, et al. 2008]. The observed consistent associations across diverse populations, combined with the results of biological experiments on rs169669968, provide converging evidence that rs16969968, rather than rs1051730, is most likely one causative variant in this region driving the strongest association signal.

Prior meta-analyses in European ancestry populations have reported additional association signals distinct from bin A, and they cluster into three groups. Bin B, a group of 111 variants highly correlated in Europeans, includes the previously reported associated SNPs rs588765 and rs880395. The association with bin B previously reported in Europeans was seen only in association analyses conditioning on rs16969968 [2]. Bin B consists of 39 sub-bins in Asian subjects and 37 sub-bins in African American subjects. In conditional analyses, we found evidence of association between bin B and heavy smoking in the Asian data (OR=1.30, 95%

C.I.= 1.07–1.58, p=$9.7×10^{-3}$), as well as reproducing the European ancestry finding (OR=1.27, 95%C.I.=1.16–1.38, p=$8.7×10^{-8}$). In the African American data, there was a trend toward association in the same direction (OR=1.16, 95%C.I.=0.99–1.36, p=0.064).

Bin B variants, located upstream of the coding region of *CHRNA5,* are associated with variability in *CHRNA5* mRNA levels in European ancestry samples [Falvella, et al. 2010; Smith, et al. 2010; Wang, et al. 2009]. Low levels of *CHRNA5* mRNA expression are associated with lower risk for nicotine dependence. No data exist on *CHRNA5* mRNA expression in other populations, and further work to examine expression data and smoking behavior in other populations is needed. Because the risk allele of rs16969968 occurs primarily on the low mRNA expression alleles represented by bin B, conditional SNP analysis controlling for bin A (rs16969968) is important to distinguish between these two distinct mechanisms [Saccone, et al. 2010; Wang, et al. 2009]. This is an important example to demonstrate how a genetic effect could be better detected and characterized when additional related variants are taken into account.

Bin C, a group of 82 variants correlated in Europeans, consisted of 24 sub-bins in subjects of Asians and 26 sub-bins in African Americans. Variants reported in previous meta-analyses of European ancestry (rs6495308 [Liu, et al. 2010], rs2036534 [Thorgeirsson, et al. 2010], rs7163730, rs9788682, rs684513 [TAG 2010], rs578776 [Saccone, et al. 2010]) were all examined. However, no tested SNP in bin C was consistently associated with heavy smoking with $p < 0.01$ in both Asians and African Americans. Similarly, we have no consistent associations with bin D which contains 15 SNPs correlated in Europeans and consists of 7 sub-bins in Asians and 7 sub-bins in African Americans.

In undertaking this project, we faced numerous challenges. First, smoking behavior differs substantially across populations. Smoking quantity distributions differ across populations; smokers of European ancestry smoke more heavily than do Asians or African Americans. As a result, we decided to compare heavy (>20 cigarettes per day) versus light smoking (10 cigarettes per day) in our primary association analysis to more closely capture the contrast between nicotine dependent smokers and non-dependent smokers. We then confirmed the consistency of results using the full distribution of smoking quantity in subsequent analyses.

Second, genotyping coverage varied between studies, and several studies in our meta-analysis had only a few variants genotyped. As a result, not all sub-bins were tested and the sample size varied across the tested sub-bins. For example, SNP rs55853698 which was imputed and reported as highly associated with smoking quantity in a previous meta-analysis of European ancestry subjects [Liu, et al. 2010], lies in bin A, but no genotyping data was available for testing this SNP or the sub-bin it represents in Asian and African American populations. Use of imputed data has the potential to mitigate these problems. However, imputation was not possible for our low-coverage studies. Therefore, the concerns about having untested SNPs and unequal subjects in the region would remain even with imputed data. We believe it is important to report our findings based on directly genotyped variants, and the interpretation of the consistent associations is not expected to change with imputation.

In addition, we were not able to perform thorough admixture tests in all datasets due to variable genotyping. Population stratification unaccounted for by our stratified analyses of self identified ancestry - European, Asian, and African American - could be a confounder in our results. Although we are leveraging the admixture to separate the effects of different genetic variants, there may be differential admixture in the cases and controls among African Americans. The impact of varied genetic architecture within given, broadly defined populations as well as within and across populations represented by individual sites (e.g.,

Japanese, Chinese, and Korean) needs to be elucidated in future larger scale studies with sufficient representation of individuals from different population backgrounds and more comprehensive genotyping.

Lastly, an association seen in one population that is not consistent across all three may nonetheless represent a true biological signal. Lack of consistency for the association may simply reflect differences in power for our population samples. Issues that can affect power across our three diverse populations include sample size, minor allele frequency, and even population-specific effect size. The last factor could arise in a variety of ways, including differences in LD structure, background variation, and marker information content. Thus we suggest caution when interpreting the negative or non-consistent association results from this study. Though these results strengthen the evidence for rs16969968 as a likely causal variant, this region remains in need of further interrogation with additional genotyping and standardized imputation across all populations.

Despite the limitations of this study, this meta-analysis refines the association signals with heavy smoking across samples representing European ancestry, Asian, and African American populations. In particular, for bin A, we present evidence showing rs16969968 is a likely causal variant for heavy smoking among the common SNPs in the bin. Our evidence also suggests there are additional distinct genetic variants in the chromosome 15q25 region associated with smoking, but we are unable to clearly identify these other associations across all three populations. For example, we extend the finding of association with bin B in European ancestry samples to an association in Asians, and a trend towards association in African Americans.

This consistent pattern of cross-population association despite many unmeasured genetic and environmental differences has provided important evidence to support true causal variants. It also provides critical information by narrowing a region of interest so laboratory experiments that must follow association studies can focus on a smaller number of variants. Thus, this study represents an important step on the pathway from association to function.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Bonnen PE, de Bakker PI, Deloukas P, Gabriel SB, et al. Integrating common and rare genetic variation in diverse human populations. Nature. 2010; 467(7311):52–8. [PubMed: 20811451]

Amos CI, Gorlov IP, Dong Q, Wu X, Zhang H, Lu EY, Scheet P, Greisinger AJ, Mills GB, Spitz MR. Nicotinic acetylcholine receptor region on chromosome 15q25 and lung cancer risk among African Americans: a case-control study. J Natl Cancer Inst. 2010; 102(15):1199–205. [PubMed: 20554942]

Baker TB, Weiss RB, Bolt D, von Niederhausern A, Fiore MC, Dunn DM, Piper ME, Matsunami N, Smith SS, Coon H, et al. Human neuronal acetylcholine receptor A5-A3-B4 haplotypes are associated with multiple nicotine dependence phenotypes. Nicotine Tob Res. 2009; 11(7):785–96. [PubMed: 19436041]

Berrettini W, Yuan X, Tozzi F, Song K, Francks C, Chilcoat H, Waterworth D, Muglia P, Mooser V. Alpha-5/alpha-3 nicotinic receptor subunit alleles increase risk for heavy smoking. Molecular Psychiatry. 2008; 13(4):368–73. [PubMed: 18227835]

Bierut LJ, Stitzel JA, Wang JC, Hinrichs AL, Grucza RA, Xuei X, Saccone NL, Saccone SF, Bertelsen S, Fox L, et al. Nicotine Dependence and the a5-a3-b4 Nicotinic Receptor gene cluster: Variants in the Nicotinic Receptors Alter the Risk for Nicotine Dependence. American Journal of Psychiatry. 2008; (165):1163–1171. [PubMed: 18519524]

Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet. 2004; 74(1):106–20. [PubMed: 14681826]

Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. Nature. 2010; 467(7319): 1061–73. [PubMed: 20981092]

Falvella FS, Galvan A, Colombo F, Frullanti E, Pastorino U, Dragani TA. Promoter polymorphisms and transcript levels of nicotinic receptor CHRNA5. J Natl Cancer Inst. 2010; 102(17):1366–70. [PubMed: 20733116]

Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008; 24(24): 2938–9. [PubMed: 18974171]

Keskitalo K, Broms U, Heliovaara M, Ripatti S, Surakka I, Perola M, Pitkaniemi J, Peltonen L, Aromaa A, Kaprio J. Association of serum cotinine level with a cluster of three nicotinic acetylcholine receptor genes (CHRNA3/CHRNA5/CHRNB4) on chromosome 15. Hum Mol Genet. 2009; 18(20):4007–12. [PubMed: 19628476]

Li MD, Beuten J, Ma JZ, Payne TJ, Lou XY, Garcia V, Duenes AS, Crews KM, Elston RC. Ethnic- and gender-specific association of the nicotinic acetylcholine receptor alpha4 subunit gene (CHRNA4) with nicotine dependence. Human Molecular Genetics. 2005; 14(9):1211–9. [PubMed: 15790597]

Li MD, Yoon D, Lee JY, Han BG, Niu T, Payne TJ, Ma JZ, Park T. Associations of variants in CHRNA5/A3/B4 gene cluster with smoking behaviors in a Korean population. PLoS One. 2010; 5(8):e12183. [PubMed: 20808433]

Liu JZ, Tozzi F, Waterworth DM, Pillai SG, Muglia P, Middleton L, Berrettini W, Knouff CW, Yuan X, Waeber G, et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. Nat Genet. 2010; 42(5):436–40. [PubMed: 20418889]

Lumley T. rmeta: Meta-analysis, R package version 2.15.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3):559–75. [PubMed: 17701901]

Rotimi CN, Jorde LB. Ancestry and disease in the age of genomic medicine. N Engl J Med. 2010; 363(16):1551–8. [PubMed: 20942671]

Saccone NL, Culverhouse RC, Schwantes-An TH, Cannon DS, Chen X, Cichon S, Giegling I, Han S, Han Y, Keskitalo-Vuokko K, et al. Multiple independent loci at chromosome 15q25.1 affect smoking quantity: a meta-analysis and comparison with lung cancer and COPD. PLoS Genet. 2010; 6(8)

Saccone NL, Saccone SF, Goate AM, Grucza RA, Hinrichs AL, Rice JP, Bierut LJ. In search of causal variants: refining disease association signals using cross-population contrasts. BMC Genet. 2008; 9:58. [PubMed: 18759969]

Saccone NL, Wang JC, Breslau N, Johnson EO, Hatsukami D, Saccone SF, Grucza RA, Sun L, Duan W, Budde J, et al. The CHRNA5-CHRNA3-CHRNB4 nicotinic receptor subunit gene cluster affects risk for nicotine dependence in African-Americans and in European-Americans. Cancer Res. 2009; 69(17):6848–56. [PubMed: 19706762]

Saccone SF, Hinrichs AL, Saccone NL, Chase GA, Konvicka K, Madden PA, Breslau N, Johnson EO, Hatsukami D, Pomerleau O, et al. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Human Molecular Genetics. 2007; 16(1):36–49. [PubMed: 17135278]

Schwartz AG, Wenzlaff AS, Bock CH, Ruterbusch JJ, Chen W, Cote ML, Artis AS, Van Dyke AL, Land SJ, Harris CC, et al. Admixture mapping of lung cancer in 1812 African-Americans. Carcinogenesis. 2010; 32(3):312–7. [PubMed: 21115650]

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001; 29(1):308–11. [PubMed: 11125122]

Sherva R, Wilhelmsen K, Pomerleau CS, Chasse SA, Rice JP, Snedecor SM, Bierut LJ, Neuman RJ, Pomerleau OF. Association of a single nucleotide polymorphism in neuronal acetylcholine receptor subunit alpha 5 (CHRNA5) with smoking status and with 'pleasurable buzz' during early experimentation with smoking. Addiction. 2008; 103(9):1544–52. [PubMed: 18783506]

Shiraishi K, Kohno T, Kunitoh H, Watanabe S, Goto K, Nishiwaki Y, Shimada Y, Hirose H, Saito I, Kuchiba A, et al. Contribution of nicotine acetylcholine receptor polymorphisms to lung cancer risk in a smoking-independent manner in the Japanese. Carcinogenesis. 2009; 30(1):65–70. [PubMed: 19005185]

Smith RM, Alachkar H, Papp AC, Wang D, Mash DC, Wang JC, Bierut LJ, Sadee W. Nicotinic alpha5 receptor subunit mRNA expression is associated with distant 5′ upstream polymorphisms. Eur J Hum Genet. 2010

Stevens VL, Bierut LJ, Talbot JT, Wang JC, Sun J, Hinrichs AL, Thun MJ, Goate A, Calle EE. Nicotinic receptor gene variants influence susceptibility to heavy smoking. Cancer Epidemiol Biomarkers Prev. 2008; 17(12):3517–25. [PubMed: 19029397]

TAG. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet. 2010; 42(5):441–7. [PubMed: 20418890]

Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature. 2008; 452(3):638–642. [PubMed: 18385739]

Thorgeirsson TE, Gudbjartsson DF, Surakka I, Vink JM, Amin N, Geller F, Sulem P, Rafnar T, Esko T, Walter S, et al. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. Nat Genet. 2010; 42(5):448–53. [PubMed: 20418888]

Wang JC, Cruchaga C, Saccone NL, Bertelsen S, Liu P, Budde JP, Duan W, Fox L, Grucza RA, Kern J, et al. Risk for nicotine dependence and lung cancer is conferred by mRNA expression levels and amino acid change in CHRNA5. Hum Mol Genet. 2009

Weiss RB, Baker TB, Cannon DS, von Niederhausern A, Dunn DM, Matsunami N, Singh NA, Baird L, Coon H, McMahon WM, et al. A candidate gene approach identifies the CHRNA5-A3-B4 region as a risk factor for age-dependent nicotine addiction. PLoS Genet. 2008; 4(7):e1000125. [PubMed: 18618000]

Wu C, Hu Z, Yu D, Huang L, Jin G, Liang J, Guo H, Tan W, Zhang M, Qian J, et al. Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. Cancer Res. 2009; 69(12):5065–72. [PubMed: 19491260]

Zaitlen N, Pasaniuc B, Gur T, Ziv E, Halperin E. Leveraging genetic variability across populations for the identification of causal variants. Am J Hum Genet. 2010; 86(1):23–33. [PubMed: 20085711]
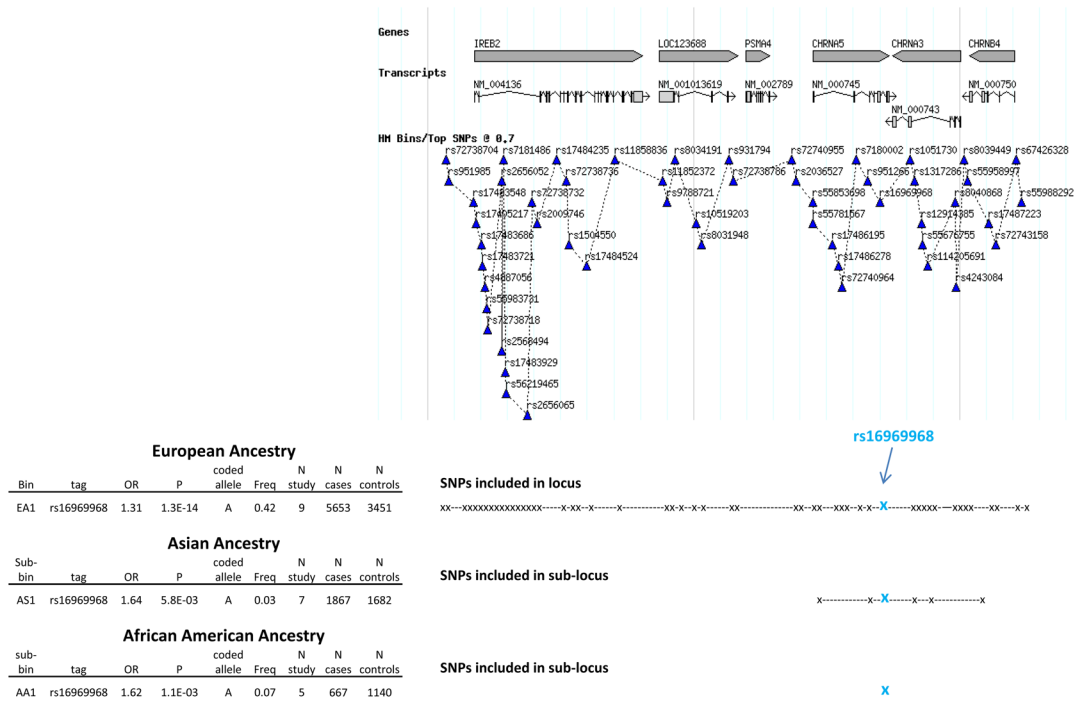
Figure 1. Top associations with heavy vs. light smoking: Bin A across three populations
Figure 1 shows all 52 SNPs in bin A, and also the only consistently associated sub-bins (p<0.01 in both Asians and African Americans). This figure also lists corresponding odds ratios for the association between the sub-bins and heavy smoking. bin A variants span across six genes in the European ancestry population, the sub-bin tagged by rs16969968 in the Asian population spans across two genes, and the sub-bin tagged by rs16969968 in the African American population spans across only one gene (CHRNA5).
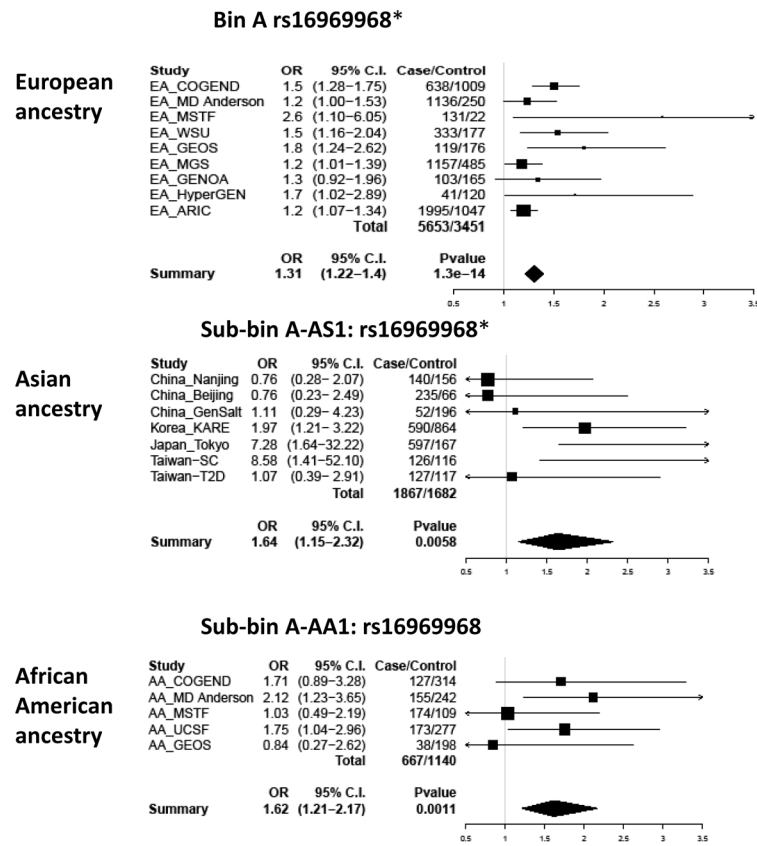
Figure 2. Rs16969968 and heavy smoking in samples of European, Asian, and African American ancestry

The ORs and 95% CIs are for the effect per allele using additive coding in the logistic regression with age and sex as covariates. *When rs16969968 is not available, rs1051730 and rs951266 are used as proxy SNPs in European and Asian ancestry samples.

Table 1

Genetic variants associated with smoking quantity reported in meta-analyses in subjects of European ancestry

Bin	SNP	References	Coded allele	Coded Allele frequency		
				CEU	JPT/CHB	ASW/YRI
A	rs16969968	Saccone et al. 2010	A	0.42	0.03	0.07
A	rs1051730	TAG 2010, Liu et al. 2010, Thorgeirsson et al. 2010	T	0.42	0.03	0.12
B	rs588765/rs880395	Saccone et al. 2010	T	0.39	0.05	0.23
C	rs6495308	Liu et al.2010	C	0.20	0.80	0.31
C	rs2036534	Thorgeirsson et al. 2010	C	0.19	0.50	0.24
C	rs7163730	TAG 2010	G	0.19	0.49	0.25
C	rs9788682	TAG 2010	A	0.19	0.28	0.20
C	rs684513	TAG 2010	G	0.20	0.25	0.21
C	rs578776	Saccone et al. 2010	T	0.24	0.87	0.55
D	rs2869046	Thorgeirsson et al. 2010	C	0.45	0.39	0.17

Allele frequencies based on 1000 Genomes Pilot 1 and HapMap 3 Release 2.

Table 2

Correlation between the examined variants (r^2) in European, Asian, and African American ancestry populations: BinsA, B, C, and D denote four groups of correlated SNPs ($r^2 > 0.7$) in the European ancestry reference sample (1000 Genomes Pilot 1 CEU).

European Ancestry

	rs16969968	rs1051730	rs588765	rs880395	rs6495308	rs2036534	rs7163730	rs9788682	rs684513	rs578776	rs2869046
	Bin A										
rs16969968	1.0										
rs1051730	1.0	1.0	Bin B								
rs588765	0.44	0.44	1.0								
rs880395	0.46	0.46	0.76	1.0	Bin C						
rs6495308	0.18	0.18	0.11	0.05	1.0						
rs2036534	0.17	0.17	0.1	0.15	0.75	1.0					
rs7163730	0.17	0.17	0.1	0.15	0.75	1.0	1.0				
rs9788682	0.17	0.17	0.1	0.15	0.75	1.0	1.0	1.0			
rs684513	0.1	0.1	0.11	0.16	0.7	0.85	0.85	0.85	1.0		
rs578776	0.23	0.23	0.04	0.01	0.78	0.66	0.66	0.66	0.61	1.0	Bin D
rs2869046	0.18	0.18	0.54	0.52	0.07	0.14	0.14	0.14	0.15	0.04	1.0

Asian Ancestry

	rs16969968	rs1051730	rs588765	rs880395	rs6495308	rs2036534	rs7163730	rs9788682	rs684513	rs578776	rs2869046
	Bin A										
rs16969968	1.0										
rs1051730	1.0	1.0	Bin B								
rs588765	0	0	1.0								
rs880395	0	0	0.1	1.0	Bin C						
rs6495308	0.1	0.1	0.21	0.08	1.0						
rs2036534	0.03	0.03	0.05	0.04	0.07	1.0					
rs7163730	0.03	0.03	0.05	0.04	0.06	0.97	1.0				
rs9788682	0.01	0.01	0.02	0	0.1	0.38	0.39	1.0			
rs684513	0.01	0.01	0.02	0.02	0.08	0.28	0.29	0.64	1.0		
rs578776	0.17	0.17	0.11	0.06	0.62	0.02	0.02	0.02	0.02	1.0	Bin D
rs2869046	0.04	0.04	0	0.01	0.01	0.11	0.13	0.13	0.09	0.02	1.0

African American Ancestry

	rs16969968	rs1051730	rs588765	rs880395	rs6495308	rs2036534	rs7163730	rs9788682	rs684513	rs578776	rs2869046
	Bin A										
rs16969968	1.0										
rs1051730	0.4	1.0	Bin B								
rs588765	--	0.04	1.0								
rs880395	--	0.02	0.68	1.0	Bin C						
rs6495308	0.04	0.06	0.14	0.11	1.0						
rs2036534	0.03	0.05	0.06	0.04	0.09	1.0					
rs7163730	0.03	0.05	0.06	0.04	0.09	1.0	1.0				
rs9788682	0.02	0.03	0.06	0.05	0.03	0.6	0.6	1.0			
rs684513	--	0.04	0.08	0.06	0.16	0.62	0.62	0.48	1.0		
rs578776	0.1	0.01	0.55	0.34	0.22	0.08	0.08	0.05	0.13	1.0	Bin D
rs2869046	--	0.03	0.03	0.01	0.13	0	0	0.04	0.02	0.02	1.0

We used SNAP to obtain LD values from HapMap 3 ASW and 1000 Genomes Pilot 1 reference populations. --: data unavailable