Published in final edited form as:
Nat Genet. 2010 January ; 42(1): 45–52. doi:10.1038/ng.500.

Meta-analyses of genome-wide association studies identify multiple novel loci associated with pulmonary function

Dana B. Hancock[1,29], **Mark Eijgelsheim**[2,29], **Jemma B. Wilk**[3,29], **Sina A. Gharib**[4,5,29], **Laura R. Loehr**[1,7], **Kristin D. Marciante**[5,6], **Nora Franceschini**[7], **Yannick M.T.A. van Durme**[2,8], **Ting-hsu Chen**[9,10], **R. Graham Barr**[11,12,13], **Matthew B. Schabath**[14], **David J. Couper**[15], **Guy G. Brusselle**[2,8], **Bruce M. Psaty**[5,6,16,17,18], **Cornelia M. van Duijn**[2], **Jerome I. Rotter**[19], **André G. Uitterlinden**[2,20], **Albert Hofman**[2], **Naresh M. Punjabi**[21], **Fernando Rivadeneira**[2,20], **Alanna C. Morrison**[22], **Paul L. Enright**[23], **Kari E. North**[7,24], **Susan R. Heckbert**[6,16,17], **Thomas Lumley**[25,30], **Bruno H.Ch. Stricker**[2,20,26,27,30], **George T. O'Connor**[9,10,30], and **Stephanie J. London**[1,28,30]

[1]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina [2]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands [3]Departments of Neurology and Medicine, Boston University School of Medicine, Boston, Massachusetts [4]Center for Lung Biology, University of Washington, Seattle, Washington [5]Department of Medicine, University of Washington, Seattle, Washington [6]Cardiovascular Health Research Unit, University of Washington, Seattle, Washington [7]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina [8]Department of Respiratory Medicine, Ghent University and Ghent University Hospital, Ghent, Belgium [9]Pulmonary Center, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts [10]The National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts [11]Division of General Medicine, Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, New York [12]Division of Pulmonary, Allergy and Critical Care, Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, New York [13]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York [14]Risk Assessment, Detection, and Intervention, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida [15]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina [16]Department of Epidemiology, University of Washington, Seattle, Washington [17]Center for Health Studies, Group Health, Seattle, Washington [18]Department of Health Services, University of Washington, Seattle, Washington [19]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California [20]Department of Internal Medicine, Eramus Medical Center, Rotterdam, The

Correspondence should be addressed to S.J.L. (london2@niehs.nih.gov).
[29]These authors contributed equally to this work
[30]These authors jointly directed this work.

Author Contributions

ARIC: D.B.H., L.R.L., N.F., M.B.S., D.J.C., N.M.P., A.C.M., K.E.N., S.J.L.
CHS: S.A.G., K.D.M., R.G.B., B.M.P., J.I.R., P.L.E., S.R.H., T.L.
FHS: J.B.W., T.C., G.T.O.
RS: M.E., Y.M.T.A.van.D., G.G.B., C.M.van.D., A.G.U., A.H., F.R., B.H.Ch.S.
Study design: T.L., B.H.Ch.S., G.T.O., S.J.L.; Data analysis: D.B.H., M.E., J.B.W., L.R.L., K.D.M., N.F., T.C.; Drafting of manuscript: D.B.H., M.E., J.B.W., S.A.G. Critical revision of manuscript: D.B.H., M.E., J.B.W., S.A.G., L.R.L., K.D.M., N.F., Y.M.T.A.van.D., T.C., R.G.B., M.B.S., D.J.C., G.G.B., B.M.P., C.M.van.D., J.I.R., A.G.U., A.H., N.M.P., F.R., A.C.M. P.L.E., K.E.N., S.R.H., T.L., B.H.Ch.S., G.T.O., S.J.L.
The authors declare no competing financial interests.

Netherlands [21]Departments of Medicine and Epidemiology, Johns Hopkins University, Baltimore, Maryland [22]Human Genetics Center, Division of Epidemiology and Disease Control, University of Texas Health Science Center at Houston, Houston, Texas [23]College of Public Health, University of Arizona, Tucson, Arizona [24]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC [25]Department of Biostatistics, University of Washington, Seattle, Washington [26]Netherlands Genomics Initiative-sponsored Netherlands Consortium for Healthy Aging, Rotterdam, The Netherlands [27]Inspectorate of Health Care, The Hague, The Netherlands [28]Laboratory of Respiratory Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina

Abstract

Measurements of lung function by spirometry are heritable traits that reflect respiratory health and predict morbidity and mortality. We meta-analyzed genome-wide association studies for two clinically important measures, forced expiratory volume in the first second (FEV_1) and its ratio to forced vital capacity (FEV_1/FVC), an indicator of airflow obstruction. This meta-analysis included 20,890 participants of European ancestry from four CHARGE consortium studies: Atherosclerosis Risk in Communities (ARIC), Cardiovascular Health Study (CHS), Framingham Heart Study (FHS), and Rotterdam Study (RS). We identified eight loci associated with FEV_1/FVC (*HHIP*, *GPR126*, *ADAM19*, *AGER-PPT2*, *FAM13A*, *PTCH1*, *PID1*, and *HTR4*) and one locus associated with FEV_1 (*INTS12-GSTCD-NPNT*) at or near genome-wide significance ($P<5\times10^{-8}$) in CHARGE; all but 3 loci (*FAM13A*, *PTCH1*, and *PID1*) replicated with the SpiroMeta consortium. Our findings of novel loci influencing pulmonary function may offer insights into chronic lung disease pathogenesis.

Introduction

Pulmonary function is an easily measurable and reliable index of the physiological state of the lungs and airways[1]. Pulmonary function also predicts mortality in the general population, even among never smokers with only modestly reduced pulmonary function and without respiratory symptoms2,3. The peak level of pulmonary function attained in early adulthood and its subsequent decline with age are likely influenced by genetic and environmental factors. Tobacco smoking is a major environmental cause of accelerated decline in pulmonary function with age. Other inhaled pollutants also appear to contribute. Familial aggregation studies suggest a genetic contribution to lung function with heritability estimates exceeding 40%4,5, but little is known about specific genetic factors involved. A relatively uncommon deficiency of α1-antitrypsin is the only established genetic risk factor for accelerated decline in pulmonary function and development of chronic obstructive pulmonary disease (COPD), especially in smokers[4,6]. However, α1-antitrypsin accounts for little of the population variability in pulmonary function4. Candidate gene studies suggest that other genetic variants may influence the time course of pulmonary function and its decline in relation to smoking, but these putative genetic risk factors remain unknown4.

Forced expiratory volume in the first second (FEV_1) and its ratio to forced vital capacity (FEV_1/FVC) are two clinically relevant pulmonary function measures. While both FEV_1 and FVC are influenced by lung size and can be reduced by restrictive lung diseases, obstructive lung disease leads to proportionately greater reduction in FEV_1 than FVC. Therefore, a reduced FEV_1/FVC, an indicator of airflow obstruction that is independent of lung size, is the primary criterion for defining an obstructive ventilatory defect[1]. Whereas low FEV_1/

FVC indicates the presence of airflow obstruction, FEV_1 is used to classify severity and follow the progression of obstructive lung disease over time[5,7,8].

The first genome-wide association study (GWAS) for pulmonary function evaluating 70,987 single nucleotide polymorphisms (SNPs) in about 1,220 Framingham Heart Study (FHS) participants revealed no genome-wide significant loci[9]. Recently, a GWAS of FEV_1/FVC using 2,540,223 SNPs in 7,691 FHS participants identified several chromosome 4q31 SNPs near *HHIP* with genome-wide significance[10]. A GWAS of COPD[11] also implicated the *HHIP* region along with *CHRNA3/5* on chromosome 15, previously associated with nicotine dependence[12,13].

We conducted meta-analyses of GWAS results for a cross-sectional analysis of pulmonary function (FEV_1/FVC and FEV_1) in 20,890 individuals of European ancestry from four Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium[14] studies: Atherosclerosis Risk in Communities (ARIC), Cardiovascular Health Study (CHS), FHS, and Rotterdam Study (RS-I and RS-II). Given that cigarette smoking is a major risk factor for pulmonary function decline, we conducted meta-analyses with adjustment for smoking status and quantity, and in subgroups of ever and never smokers. Significant findings and other selected high-signal hits were evaluated for replication with the SpiroMeta consortium, an independent consortium having a combined sample size of 20,228 participants of European ancestry as described in the accompanying manuscript.

Results

Meta-analyses of CHARGE genome-wide association results

Meta-analyses for FEV_1/FVC and FEV_1 were conducted using approximately 2,534,500 SNPs in 20,890 CHARGE participants of European ancestry (N=7,980 from ARIC, N=3,140 from CHS, N=7,694 from FHS, N=1,224 from RS-I, and N=852 from RS-II) and in subgroups of ever (N=11,963) and never smokers (N=8,927). Characteristics of the cohort participants are presented in Table 1. We applied genomic control, although cohort-specific genomic inflation factors (λ_{gc}) were low (for FEV_1/FVC ranging from 1.00 (RS-I and RS-II) to 1.05 (ARIC) and for FEV_1 ranging from 1.01 (RS-II) to 1.05 (FHS)) suggesting minimal population stratification. The meta-analysis λ_{gc} was 1.04 for FEV_1/FVC and 1.03 for FEV_1 in all participants. Quantile-quantile (Q-Q) plots show large deviations between observed and expected *P* values for high-signal SNPs in analyses of FEV_1/FVC and FEV_1 in all participants (Supplementary Fig. 1a,b), FEV_1/FVC in never smokers (Supplementary Fig. 2a), and FEV_1 in ever smokers (Supplementary Fig. 3c). Genome-wide significant associations ($P<5\times10^{-8}$) were found for multiple SNPs in each of these analyses (Fig. 1a,b for overall analyses and Supplementary Fig. 2b,d and Supplementary Fig. 3b,d for analyses stratified by ever/never smoking). The top 2,000 SNPs associated with each measure, FEV_1/FVC and FEV_1, beyond genome-wide significance ($P>5\times10^{-8}$) are presented in Supplementary Table 1.

For FEV_1/FVC, genome-wide significant associations were seen for 119 SNPs at seven loci (Supplementary Table 2). The SNP with the smallest *P* value, rs1980057 ($P=4.90\times10^{-11}$), is located on chromosome 4q31.22, 81 kb away from the 5'-end of *HHIP*. There were 27 other genome-wide significant SNPs in the *HHIP* region (Fig. 2a). Additionally, 69 genome-wide significant SNPs were located in or near the 3'-end of *GPR126* on chromosome 6q24.1, with the top SNP (rs3817928) having $P=2.60\times10^{-10}$ (Fig. 2b). Fifty-nine of these 69 *GPR126* SNPs were associated with FEV_1/FVC at genome-wide significance among never smokers (Supplementary Table 2). Seven chromosome 5q33.3 SNPs located in *ADAM19* (Fig. 2c), two correlated chromosome 6p21.32 SNPs ($r^2=0.66$, Fig. 2d) located in two genes (*AGER* and *PPT2*), four chromosome 4q22.1 SNPs near the 5'-end of *FAM13A* (Fig. 2e), two

chromosome 9q22.32 SNPs in *PTCH1* (Fig. 2f), and six chromosome 2q36.3 SNPs near the 3'-end of *PID1* (Fig. 2g) were also significantly associated with FEV_1/FVC in all participants. SNPs in *AGER*, *PPT2*, *PTCH1*, and *PID1* had minor allele frequencies (MAFs) between 4 and 10%, while all other significantly associated SNPs had MAFs exceeding 10%. Absolute β values (per-allele change in FEV_1/FVC) ranged from 0.44 to 1.14%. The β directions were consistent across the CHARGE cohorts for all genome-wide significant SNPs except for the *GPR126* SNPs noted in Supplementary Table 2. A borderline significant association ($P=5.37\times10^{-8}$, MAF=0.42, β=−0.43) with FEV_1/FVC was noted for the chromosome 5q33.1 SNP rs11168048 in *HTR4* (Fig. 2h). Cohort-specific association results for SNPs with the smallest *P* value from each locus implicated at or near genome-wide significance are shown in Supplementary Table 3.

For FEV_1, genome-wide significant associations were observed for 46 chromosome 4q24 SNPs in or near four adjacent genes (Supplementary Table 4). The SNP with the smallest *P* value, rs17331332 ($P=4.00\times10^{-10}$), is located near *NPNT*. The 45 other significantly associated SNPs include four SNPs located near the 5'-end of *NPNT*, five SNPs located in *INTS12* or near its 3'-end, seven SNPs located in *FLJ20184* or near its 3'-end, and 29 SNPs located in *GSTCD*. *FLJ20184* encodes a hypothetical protein according to several genome browsers including the UCSC genome browser[15], but there is no approved HUGO gene name for this locus[16]. The SNP rs17331332 is correlated at $r^2>0.5$ with most other significantly associated SNPs in this region (Fig. 3), suggesting that the associations in the four adjacent genes represent one independent finding. The significantly associated SNPs had MAFs between 6 and 8%. The absolute β values (per-allele change in FEV_1) ranged from 55.92 to 71.43 mL (Supplementary Table 4), and the β directions were consistent across the CHARGE cohorts for all 46 genome-wide significant SNPs (Supplementary Table 3 for rs17331332). Among these 46 SNPs, 39 were associated with FEV_1 at genome-wide significance among ever smokers (Supplementary Table 4).

To evaluate whether other loci may also influence pulmonary function, we created Q-Q plots for FEV_1/FVC and FEV_1 among all participants after removing SNPs (1,862 for FEV_1/FVC and 284 for FEV_1) at or close to genome-wide significance and nearby SNPs correlated at $r^2>0.2$ with the top SNP for each locus. The resulting Q-Q plots show some excess of small *P* values for FEV_1/FVC (Supplementary Fig. 4a) and FEV_1 (Supplementary Fig. 4b).

Putative functional polymorphisms

Three SNPs among the 119 genome-wide significant SNPs for FEV_1/FVC are non-synonymous (missense) polymorphisms: rs11155242 (Lys to Gln) in *GPR126*, rs1422795 (Ser to Gly) in *ADAM19*, and rs2070600 (Gly to Ser) in *AGER*. The Polymorphism Phenotyping (PolyPhen) program[17] predicts that the amino acid substitutions resulting from rs11155242 and rs1422795 cause benign changes but predicts that rs2070600 has a possibly damaging impact on the structure and function of AGER.

All other SNPs implicated for FEV_1/FVC or FEV_1 are intergenic, intronic, or located in 3' untranslated regions. Of these, three intronic *GPR126* SNPs (rs9496346, rs1040525, and rs6929442) and one intergenic SNP near *NPNT* (rs10516529) are located in transcription factor binding sites, according to the UCSC genome browser[15].

Replication with the SpiroMeta consortium

Thirty high-signal SNPs associated with FEV_1/FVC (18 SNPs from eight loci) or FEV_1 (12 SNPs from three loci) at or close to genome-wide significance were tested in the SpiroMeta consortium. We evaluated these SNPs in 16,178 SpiroMeta participants of European

ancestry with complete quantitative smoking data using the CHARGE analytic method, which included adjustment for smoking status and pack-years, and performed joint meta-analyses of CHARGE GWAS and SpiroMeta replication results (Table 2 and Table 3). *P* values that exceeded the significance threshold in SpiroMeta ($P<8.33\times10^{-4}$ based on 60 tests) or the genome-wide significance threshold in joint meta-analyses ($P<5\times10^{-8}$) were considered significant evidence for replication.

For FEV_1/FVC, among 18 SNPs tested for replication, six SNPs in three loci were significantly associated with this measure in SpiroMeta: rs1980057 and rs1032295 near *HHIP* ($r^2=0.72$), rs2070600 in *AGER* and rs10947233 in *PPT2* ($r^2=0.66$), and rs11168048 and rs7735184 in *HTR4* ($r^2=0.93$) (Table 2). Their joint meta-analysis *P* values ranged from 3.21×10^{-20} to 6.23×10^{-11} (Table 2). Five additional SNPs in *GPR126* (rs3817928, rs7776375, and rs6937121) and *ADAM19* (rs2277027 and rs1422795) were not significantly associated with FEV_1/FVC at the stringent threshold in SpiroMeta, but these SNPs were associated at genome-wide significance in the joint meta-analysis with *P* values ranging from 9.93×10^{-11} to 1.25×10^{-8} (Table 2). For replicated SNPs, the allele frequencies and the direction and magnitude of the associations with FEV_1/FVC were similar between consortia (Table 2). Further, the *HHIP*, *ADAM19*, and *HTR4* SNPs were significantly associated with FEV_1 in SpiroMeta (Supplementary Table 5). The *HHIP* SNP rs1980057 and *HTR4* SNPs rs11168048 and rs7735184 were also associated with FEV_1 at genome-wide significance in the joint meta-analysis (*P* ranging from 5.86×10^{-9} to 1.58×10^{-8}, Supplementary Table 5). SNPs in *FAM13A*, *PTCH1*, and *PID1* that gave genome-wide significance in CHARGE were not confirmed in analyses with SpiroMeta.

For FEV_1, among the 12 SNPs tested for replication, eight SNPs from one locus with four adjacent genes were significantly associated with this measure in SpiroMeta, including rs17331332 and rs17036341 near *NPNT*, rs11727189 and rs17036090 in or near *INTS12*, rs17036052 and rs17035960 in or near *FLJ20184*, and rs11097901 and rs11728716 in *GSTCD* (Table 3). For replicated SNPs, the allele frequencies and the direction and magnitude of the associations with FEV_1 were similar between consortia, and *P* values from joint meta-analysis ranged from 4.66×10^{-17} to 9.42×10^{-14} (Table 2). None of these SNPs were significantly associated with FEV_1/FVC in CHARGE or SpiroMeta (Supplementary Table 5).

Associations in individuals with normal pulmonary function

To address whether the genetic associations hold even among people with normal pulmonary function, we repeated the meta-analyses after excluding individuals with asthma or COPD, leaving 17,855 individuals (N=6,912 from ARIC, N=2,634 from CHS, N=6,371 from FHS, N=1,126 from RS-I, and N=812 from RS-II). Asthma was defined by self-report of ever having asthma or self-report of ever having physician-diagnosed asthma. COPD was defined spirometrically as having both FEV_1/FVC and FEV_1 less than the lower limit of normal values using NHANES III prediction equations[18,19]. Comparing the original meta-analyses to the meta-analyses with exclusions for asthma and COPD, β estimates were highly correlated for the high-signal SNPs tested for replication (Pearson's r>0.99 for 18 FEV_1/FVC SNPs and 12 FEV_1 SNPs). β estimates remained highly correlated for SNPs with *P* values as high as 0.01 in the original meta-analyses (r=0.92 for FEV_1/FVC and r=0.96 for FEV_1). As expected, there was some attenuation in *P* values for many of the SNPs in our implicated loci given the substantial power loss due to both reduced sample size and the truncation of the FEV_1/FVC and FEV_1 distributions, but there was substantial overlap in the top-ranking SNPs between the two meta-analyses (results not shown). The *P* values for some top-ranking SNPs became smaller, including several *ADAM19*, *FAM13A*, and *HTR4* SNPs associated with FEV_1/FVC. Of note, 12 SNPs in *HTR4*, a locus with one SNP rs11168048 showing borderline genome-wide significance in the original meta-analysis,

gave genome-wide significance in the subset of individuals without asthma or COPD (P=6.93×10^{-9} for rs11168048).

Discussion

In meta-analyses of GWAS results in 20,890 CHARGE participants of European ancestry, we identified genome-wide significant associations with FEV_1/FVC for SNPs in seven novel independent loci (*GPR126*, *ADAM19*, *AGER-PPT2*, *FAM13A*, *PTCH1*, *PID1*, and *HTR4*) and with FEV_1 for one novel independent locus annotated by at least three genes (*INTS12-GSTCD-NPNT*). The SpiroMeta consortium independently reported genome-wide significant associations of *GSTCD*, *HTR4*, *AGER*, *TNS1*, and *THSD4* with FEV_1/FVC and FEV_1 in an independent sample of 20,228 individuals of European ancestry (accompanying manuscript). Both consortia confirm previous GWAS findings implicating the *HHIP* region for FEV_1/FVC[10].

Several SNPs near the hedgehog interacting protein (*HHIP*) gene were associated with FEV_1/FVC at genome-wide significance in CHARGE and SpiroMeta, confirming earlier GWAS findings in FHS[10]. The hedgehog (Hh)-signaling pathway is crucial in several embryonic development processes, including the branching morphogenesis of the lung[20,21]. Furthermore, several polymorphisms in three genes of the Hh-signaling pathway (*IHH*, *HHIP*, and *PTCH1*) were significantly associated in a GWAS of adult height[22]. Several *PTCH1* SNPs were also significantly associated with FEV_1/FVC in CHARGE, but these associations were not confirmed in SpiroMeta. Epithelial cells produce Hh protein, which binds to its membrane receptor (encoded by *PTCH1*) on mesenchymal cells and orchestrates tissue and organ patterning. Hh pathway dysfunction during fetal life in humans is responsible for severe lung malformations[23,24]. In adults, the Hh-signaling pathway may participate in the response of the airway epithelium to injury, such as smoking and hyperoxia[25,26].

A non-synonymous *AGER* SNP (rs2070600) was associated with FEV_1/FVC at genome-wide significance in our study and independently confirmed in SpiroMeta. The AGER protein, a membrane-bound or soluble pattern recognition receptor, belongs to the immunoglobulin superfamily of cell surface receptors. The SNP rs2070600 has functional significance, e.g., higher ligand affinity and production of proinflammatory proteins upon activation[27]. In healthy adult mice and humans, AGER is highly expressed in the lung[28], and its absence contributes to the pathogenesis of idiopathic pulmonary fibrosis[29,30]. AGER signaling is involved in host defense, inflammation, and tissue remodeling, which are relevant processes for accelerated decline in pulmonary function with age.

Polymorphisms in *HTR4* were associated with FEV_1/FVC at genome-wide significance in the joint meta-analysis of CHARGE and SpiroMeta results. *HTR4* encodes a G-coupled transmembrane receptor that regulates cAMP production in response to 5-hydroxytryptamine (serotonin). Elevated levels of free serotonin have been found in the plasma of symptomatic asthmatics[31], and serotonin signaling pathways involving *HTR4* have been implicated in cholinergic and immune-mediated airway reactivity[32,33]. Upon activation by serotonin, *HTR4* in human airway epithelial cells regulates the release of a pro-inflammatory cytokine, a signature characteristic of asthma[34].

ADAM19 SNPs were associated with FEV_1/FVC at genome-wide significance in CHARGE and in the joint meta-analysis with SpiroMeta. *ADAM19* is a member of “a disintegrin and metalloprotease” (ADAM) family of membrane-anchored glycoproteins that control cell-matrix interactions and help regulate growth and morphogenesis. Polymorphisms in another ADAM family member, *ADAM33*, have been associated with bronchial hyperresponsiveness

and accelerated lung function decline in asthmatics and the general population[35–37]. *ADAM19* has not been previously implicated in human pulmonary disorders, but it is abundantly expressed in alveolar epithelial cells and bronchial smooth muscle tissue[38].

GPR126 polymorphisms were associated with FEV_1/FVC at genome-wide significance in CHARGE and in the joint meta-analysis with SpiroMeta. *GPR126* belongs to a superfamily of G protein-coupled receptors involved in cell adhesion and signaling[39]. While its precise function has not been elucidated, its expression in mice is temporally increased during embryonic organ development and is highest in the adult lung[40]. In humans, recent GWA studies have linked *GPR126* variants with adult height, and more specifically, with trunk height[41–43]. We adjusted all analyses for standing height. Therefore, we repeated analyses for *GPR126* SNPs adjusting for sitting height (a more reliable indicator of trunk height) in ARIC, where both height variables were measured, and associations with FEV_1/FVC remained significant. Thus, these associations are not likely due to residual confounding by trunk height.

Genome-wide significant associations with FEV_1 were observed in CHARGE for numerous SNPs spanning at least three genes on chromosome 4q24, and these associations were significant for all eight SNPs tested for replication in SpiroMeta. There is moderate to strong linkage disequilibrium among the chromosome 4q24 SNPs, and the specific genes influencing FEV_1 remain speculative. The genes are ordered *INTS12-GSTCD-NPNT* along chromosome 4q24, and joint meta-analysis with SpiroMeta showed that SNPs from the genes *INTS12* and *GSTCD* had the most significant associations with FEV_1. The product of *INTS12* is a subunit of the Integrator complex that associates with the C-terminal domain of RNA polymerase II and mediates 3'-end processing of small nuclear RNAs[44]. GSTCD (glutathione S-transferase, C-terminal domain) could influence lung function via mechanisms involving the detoxification by glutathione S-transferases of xenobiotics that might damage the lungs.

The most distal gene in the chromosome 4q24 region, *NPNT*, encodes nephronectin, which is expressed in fetal and adult lung[45,46]. The *NPNT* SNP rs10516529 is located in a binding site for the transcription factor POU6F1 (also known as mPOU homeobox protein), which is known to be expressed in adult lung and hypothesized to play a role in lung development[47–49]. A fourth predicted gene in the region, *FLJ20184*, is located proximal to the other three genes. Although *FLJ20184* encodes a hypothetical protein of unknown function, *FLJ20184* contains allelic variants associated with successful smoking cessation in a GWAS of patients in smoking cessation trials[50].

The identified genetic factors gave estimated effect sizes consistent with those for well-established risk factors for pulmonary function decline. Carrying one copy of an implicated reference allele resulted in a FEV_1 difference ranging from 50 to 70 mL. These effect sizes correspond to approximately 2.8–3.9 years of age-related decline in pulmonary function based on a mean decline of about 18 mL/year and to approximately 1.7–2.3 years of active smoking-related decline based on a mean decline of about 30 mL/year[51]. Second-hand smoke exposure has also been associated with decline in FEV_1 (15 mL decline for a 10-year exposure in the home and 41 mL decline for a 10-year workplace exposure)[52]. For FEV_1/FVC, carrying one copy of an implicated reference allele resulted in a difference ranging from 0.30 to 1%. The lower effect size estimates are comparable with the mean FEV_1/FVC decline related to second-hand smoking (0.35 for a 10-year exposure in the home and 0.14 for a 10-year workplace exposure)[52]. These comparisons demonstrate that the identified genetic factors have a moderate impact on pulmonary function. Individuals carrying these polymorphisms will have lower pulmonary function than predicted at a given age, thus placing them at greater risk for developing COPD and greater risk of mortality2,3.

A GWAS of COPD identified *CHRNA3/5* on chromosome 15 as a susceptibility locus[11]. *CHRNA3/5* has also been associated with nicotine dependence12,[13]. In CHARGE, one identified SNP in this locus (rs1051730) was associated with FEV_1/FVC (*P*=0.00070) and FEV_1 (*P*=0.016), while the other identified SNP in this locus (rs8034191) was not associated with FEV_1/FVC (*P*=0.11) or FEV_1 (*P*=0.36). The nominal evidence for replication may reflect differences in study design and a potential gene-environment interaction involving smoking.

Our study has several important strengths. The CHARGE cohorts are well-phenotyped with pulmonary function measures passing stringent quality control criteria, thus minimizing measurement error. Our large sample size of 20,890 participants offers a powerful resource to examine associations of common SNPs with modest to large effects[14]. However, we likely have insufficient power to detect associations of polymorphisms with small effect sizes or low frequencies. Replication in an independent consortium with similar power offered the opportunity to confirm true genetic associations.

Population-based cohorts are subject to population stratification, and analytic steps were taken to minimize this potential bias. Cohort-specific λ_{gc} values were low (1.00 to 1.05), and a genomic control adjustment was made in the meta-analyses to reduce inflation in the test statistics. The two largest cohorts, with the largest (albeit modest) λ_{gc} values (ARIC and FHS), incorporated principal components as potential confounders in their cohort-specific association tests. Although we cannot eliminate the possibility that some findings are subject to residual confounding by population stratification, the Q-Q plots showing deviations between observed and expected *P* values for many high- to moderate-signal SNPs and the replication of association for multiple top loci in SpiroMeta suggest a multifactorial influence on pulmonary function.

Our study identified several novel loci related to two clinically important pulmonary function measures with evidence for replication, including *GPR126*, *ADAM19*, *AGER-PPT2*, and *HTR4* for FEV_1/FVC and *INTS12-GSTCD-NPNT* for FEV_1 and confirmed previous reports of association with FEV_1/FVC in the *HHIP* region. These loci include genes with biologically plausible functions, and their identification here warrants future investigations to elucidate the mechanisms underlying their influence on pulmonary function. A few of the associated polymorphisms are potentially functional, but most of the associated polymorphisms likely tag for yet unidentified functional variants. Fine mapping in these regions might identify and characterize such variants. Understanding the genetic determinants of pulmonary function is paramount in identifying the biological mechanisms that lead to its decline and ultimately lessening the mortality burden associated with reduced pulmonary function.

Methods

Pulmonary function measurements

Study design details of the participating CHARGE cohorts are described elsewhere[14,53–58]. Study protocols were approved by the relevant institutional review boards, and all participants provided written informed consent.

Pulmonary function testing was conducted by trained spirometry technicians at a single visit for RS and at more than one visit for ARIC, CHS, and FHS. FEV_1/FVC and FEV_1 measures meeting American Thoracic Society/European Respiratory Society criteria for acceptability were tested for association with SNPs in participants of European ancestry who were successfully genotyped and provided informed consent for genetic testing.

In ARIC and CHS, pulmonary function measures and questionnaire data from the baseline visit were analyzed. ARIC measurements were made with a Collins Survey II water-seal spirometer (Collins Medical, Inc.) and Pulmo-Screen II software (PDS Healthcare Products, Inc.)[59]. CHS measurements were made with a Collins Survey I water-seal spirometer (Collins Medical, Inc.) and software from S&M Instruments60,61.

In three generations of families participating in FHS, data from the most recent examination were analyzed. Eligible examinations providing spirometry and questionnaire data included examinations 13, 16, 17, and 19 in the original cohort (in approximate two-year intervals); examinations three, five, six, and seven in the offspring generation (in approximate four-year intervals); and the one examination completed to date for the third generation. Equipment used in the standard protocol evolved as technology improved over the decades of study[62]. A Collins Survey water-filled spirometer (Collins Medical, Inc.) was used for most examinations, with measurements made by Eagle II microprocessor (Collins Medical, Inc.) or by software from the S&M Instruments. In more recent examinations, a Collins Comprehensive Pulmonary Laboratory dry rolling-seal spirometer and Collins 2000 Plus/SQL Software (Collins Medical, Inc.) were used.

In RS, pulmonary function was measured at the fourth center visit of participants from the original cohort (RS-I) and the second center visit of participants from the first extension cohort (RS-II). Spirometry was performed using a SpiroPro® portable spirometer (Erich Jaeger GmbH)[63,64].

Genotyping, imputation, and quality control

Different genotyping platforms were used across the cohorts (Table 1)14. Imputation was conducted using either MACH65 or BIMBAM[66] to generate approximately 2.5 million autosomal SNP genotype dosages for meta-analysis. The imputation methods perform similarly, although MACH generally produces higher accuracy rates than the imputation process used in BIMBAM (fastPHASE)[67]. Differing imputation methods across cohorts is not a source of bias for meta-analysis since all comparisons using the imputed data are within-cohort comparisons.

ARIC—Among 8,861 self-identified white ARIC participants genotyped, 8,127 participants remained after exclusions for call rate<95%, genotypic and phenotypic sex mismatch, discordances with previous genotype data, suspected first-degree relative of an included individual based on genotype data, more than eight standard deviations for any of the first 10 principal components using EIGENSTRAT[68], or outlying average identity-by-state estimates using PLINK[69]. Of these, 7,980 participants had pulmonary function measures and complete covariate information.

A total of 704,588 autosomal genotyped SNPs remained after exclusions for call rate<95%, MAF<1%, Hardy-Weinberg equilibrium (HWE) $P<10^{-5}$, or lacking strand annotation. MACH (version 1.00.16)[65] was used to impute all autosomal SNPs with reference to HapMap CEU (release 21, build 35)70 from these 704,588 SNPs. Imputed SNPs failing additional quality control criteria (monomorphism, HWE $P<10^{-6}$, or genotype frequencies between two genotyping phases differed by $P<10^{-6}$) were excluded, leaving 2,515,866 genotyped or imputed SNPs for analysis.

CHS—CHS genotyped 3,980 participants free of cardiovascular disease at baseline with available DNA and consent to genetic testing. After exclusions for call rate <95%, sex mismatch, or discordance with prior genotyping, 3,291 white participants remained. Of these, 3,140 had pulmonary function measures and complete covariate information.

A set of 306,655 autosomal genotyped SNPs remained after exclusions for call rate<97%, HWE $P<10^{-5}$, more than two duplicate errors or Mendelian inconsistency (for reference HapMap CEU trios)70, heterozygote frequency>0, or no mapping in dbSNP. Imputation of autosomal SNPs was based on these 306,655 SNPs using BIMBAM (version 0.99)66 with reference to HapMap CEU (release 22, build 36)70. The analysis data set included 2,543,887 genotyped or imputed SNPs.

FHS—A total of 8,481 participants remained after exclusions for call rate<97%, heterozygosity more than five standard deviations from the mean, or excessive non-inheritance. The analysis data set included 7,694 participants with complete spirometry and covariate data.

MACH (version 1.00.15)[65] was used for imputation based on 378,163 autosomal SNPs remaining after exclusions for HWE $P<10^{-6}$, call rate<97%, differential missingness related to genotype (mishap procedure in PLINK69) with $P<10^{-9}$, Mendelian errors>100, MAF<1%, or those not present in HapMap. Two hundred unrelated individuals with high call rate were used to infer model parameters, which were subsequently applied to all 8,481 individuals. Imputation, using HapMap CEU (release 22, build 36),[70] produced genotype dosages on 2,543,887 genotyped or imputed SNPs.

RS—All RS participants with available DNA were genotyped; 5,974 RS-I participants and 2,157 RS-II participants remained after exclusion for call rate<97.5%, excess autosomal heterozygosity, sex mismatch, or outlying identity-by-state clustering estimates. Of these, 1,224 RS-I participants and 852 RS-II participants had pulmonary function measures and complete covariate information.

After exclusions for call rate<98%, HWE $P<10^{-6}$, and MAF<1%, 512,349 autosomal SNPs in RS-I and 466,389 autosomal SNPs in RS-II were used for imputation in MACH (version 1.00.15 for RS-I and 1.00.16 for RS-II)[65] with reference to the 2,543,887 SNPs of the HapMap CEU (release 22, build 36)70.

Statistical analysis

In cross-sectional analyses, FEV_1/FVC and FEV_1 were tested for association with SNP genotypes using a one degree-of-freedom additive model of the dosage value (estimated reference allele count with a fractional value ranging from 0 to 2.0) as a predictor in linear regression models. Associations were examined overall and stratified into ever and never smokers. Overall models were adjusted for age, sex, standing height, smoking status (current/past/never), and pack-years of smoking. Current, past, or never smoking was based on questionnaire responses, and pack-years were calculated for current and past smokers by multiplying smoking dose (packs/day) and duration (years). Stratified models used the same covariates as the overall models, except that the ever-smoker stratum included adjustment for smoking status as current/past and the never-smoker stratum included no smoking-related covariates. Additional study-specific covariates included recruitment cohort (FHS), recruitment center (ARIC and CHS), and principal component eigenvalues for population stratification adjustments (10 components for ARIC and statistically significant components for FHS). Models were implemented using ProbABEL[71] in ARIC, R72 in CHS, linear mixed effects models with fixed effects for SNPs and random effects for individuals correlated within families73 in FHS, and MACH2QTL65 in RS as implemented in GRIMP[74]. In FHS, the kinship package in R generated a covariance matrix for each family based on the kinship coefficient for each relative pair. The kinship matrix, which includes the full set of family-specific covariance matrices, specified the covariance matrix for the random effects.

GWAS results from the four cohorts were combined using inverse variance weighted meta-analysis in METAL (http://www.sph.umich.edu/csg/abecasis/metal/). Meta-analysis was performed on approximately 2,534,500 SNPs after applying genomic control for each study and filtering SNPs with extremely low imputation quality ratios (<0.01) and MAF (<1%). The genome-wide significance threshold was defined *a priori* as $P<5\times10^{-8}$, the Bonferroni adjustment for one million independent tests75. Information on SNP function and position relative to genes, microRNA, and transcription factor binding sites was obtained using a Perl script (J.B.W.) that queries tables of the UCSC genome browser[15] (hg18, March 2006 genome build). Functional effects of non-synonymous SNPs on protein structure and function were predicted using PolyPhen17.

Replication in the SpiroMeta consortium

We exchanged 30 SNPs for replication testing with the SpiroMeta consortium (accompanying manuscript). No additional genotyping was required, as these SNPs were available from the SpiroMeta GWAS. We aimed to select two SNPs from each of the top genes implicated for FEV_1/FVC or FEV_1, nearly all exceeding genome-wide significance. The SNP with the lowest *P* value in or near each gene was selected. A second SNP, genotyped (instead of imputed) in at least one cohort, was selected with preference for non-synonymous SNPs and SNPs not in strong linkage disequilibrium with the first selected SNP. Only one SNP was available for *AGER*, *PPT2*, *TSPYL4*, and *NT5DC1*. Four SNPs were selected from two linkage disequilibrium blocks for the largest gene, *GPR126*. In total, 18 SNPs from nine genes (eight independent loci) implicated for FEV_1/FVC and 12 SNPs from seven genes (three independent loci) implicated for FEV_1 were tested for replication.

Unlike CHARGE, SpiroMeta used normalized residuals as phenotypes, adjusted for age[2] rather than age, and did not adjust for smoking. For better comparison, SpiroMeta conducted modified analyses following the CHARGE analytic method described above in 16,178 participants from adult cohorts with complete quantitative smoking data available. Results from the CHARGE GWAS and SpiroMeta replication were combined in a joint meta-analysis using inverse variance weighting with METAL. SpiroMeta results with $P<8.33\times10^{-4}$, based on an overly conservative Bonferroni correction for 60 tests (30 SNPs tested for association with two traits, FEV_1/FVC and FEV_1), or joint meta-analysis results with $P<5\times10^{-8}$ (genome-wide significance threshold) were considered statistically significant.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01ES043012). The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions, along with Dr. Grace Chiu, Dick Howard, and Miguel Quibrera for their analytic contributions.

The Cardiovascular Health Study research reported in this article was supported by contract numbers N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, grant numbers U01 HL080295 and R01 HL087652 from the National Heart, Lung, and Blood Institute, with additional contribution from the National Institute of Neurological Disorders and Stroke. A full list of principal CHS investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center for Research Resources grant M01-RR00425 to the Cedars-

Sinai General Clinical Research Center Genotyping core and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

Research was conducted in part using data and resources from the Framingham Heart Study of the National Heart, Lung, and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the FHS investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung, and Blood Institute's FHS (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc. for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. J.B.W. is supported by a Young Clinical Scientist Award from the Flight Attendant Medical Research Institute (FAMRI).

The Rotterdam Study was supported from grants from the Netherlands Organisation of Scientific Research Netherlands Organisation for Scientific Research (NOW) Investments (175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/ NWO (050-060-810), Erasmus Medical Center, Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors thank Pascal Arp, Mila Jhamai, Dr. Michael Moorhouse, Marijn Verkerk, and Sander Bervoets for their help in creating the Rotterdam GWAS database; Dr. Tobias A. Knoch, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions, the Erasmus Computing Grid, Rotterdam, The Netherlands, and the national German MediGRID and Services@MediGRID part of the German D-Grid (German Bundesministerium fur Forschung und Technology (#01 AK 803 A–H and # 01 IG 07015 G), for access to grid resources.

References

1. Wilk JB, et al. Evidence for major genes influencing pulmonary function in the NHLBI family heart study. Genet Epidemiol 2000;19:81–94. [PubMed: 10861898]
2. Hole DJ, et al. Impaired lung function and mortality risk in men and women: findings from the Renfrew and Paisley prospective population study. BMJ 1996;313:711–716. [PubMed: 8819439]
3. Schunemann HJ, Dorn J, Grant BJ, Winkelstein W Jr, Trevisan M. Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study. Chest 2000;118:656–664. [PubMed: 10988186]
4. DeMeo DL, Silverman EK. Genetics of chronic obstructive pulmonary disease. Semin Respir Crit Care Med 2003;24:151–160. [PubMed: 16088534]
5. Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. Am J Respir Crit Care Med 2001;163:1256–1276. [PubMed: 11316667]
6. Silverman EK, Sandhaus RA. Clinical practice. Alpha1-antitrypsin deficiency. N Engl J Med 2009;360:2749–2757. [PubMed: 19553648]
7. Crapo RO. Pulmonary-function testing. N Engl J Med 1994;331:25–30. [PubMed: 8202099]
8. Pellegrino R, et al. Interpretative strategies for lung function tests. Eur Respir J 2005;26:948–968. [PubMed: 16264058]
9. Wilk JB, Walter RE, Laramie JM, Gottlieb DJ, O'Connor GT. Framingham Heart Study genome-wide association: results for pulmonary function measures. BMC Med Genet 2007;8 Suppl 1:S8. [PubMed: 17903307]
10. Wilk JB, et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. PLoS Genet 2009;5:e1000429. [PubMed: 19300500]
11. Pillai SG, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. PLoS Genet 2009;5:e1000421. [PubMed: 19300482]
12. Berrettini W, et al. Alpha-5/alpha-3 nicotinic receptor subunit alleles increase risk for heavy smoking. Mol Psychiatry 2008;13:368–373. [PubMed: 18227835]

13. Saccone SF, et al. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Hum Mol Genet 2007;16:36–49. [PubMed: 17135278]
14. Psaty BM, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. Circ Cardiovasc Genet 2009;2:73–80. [PubMed: 20031568]
15. Kent WJ, et al. The human genome browser at UCSC. Genome Res 2002;12:996–1006. [PubMed: 12045153]
16. Eyre TA, et al. The HUGO Gene Nomenclature Database, 2006 updates. Nucleic Acids Res 2006;34:D319–D321. [PubMed: 16381876]
17. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res 2002;30:3894–3900. [PubMed: 12202775]
18. Ferguson GT, Enright PL, Buist AS, Higgins MW. Office spirometry for lung health assessment in adults: A consensus statement from the National Lung Health Education Program. Chest 2000;117:1146–1161. [PubMed: 10767253]
19. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S population. Am J Respir Crit Care Med 1999;159:179–187. [PubMed: 9872837]
20. Chen MH, Wilson CW, Chuang PT. SnapShot: hedgehog signaling pathway. Cell 2007;130:386. [PubMed: 17662951]
21. Warburton D, et al. Molecular mechanisms of early lung specification and branching morphogenesis. Pediatr Res 2005;57 26R–37R.
22. Weedon MN, et al. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 2008;40:575–583. [PubMed: 18391952]
23. Villavicencio EH, Walterhouse DO, Iannaccone PM. The sonic hedgehog-patched-gli pathway in human development and disease. Am J Hum Genet 2000;67:1047–1054. [PubMed: 11001584]
24. Whitsett JA, Wert SE, Trapnell BC. Genetic disorders influencing lung formation and function at birth. Hum Mol Genet 13 Spec No 2004;2:R207–R215.
25. Lemjabbar-Alaoui H, et al. Wnt and Hedgehog are critical mediators of cigarette smoke-induced lung cancer. PLoS One 2006;1:e93. [PubMed: 17183725]
26. Pogach MS, Cao Y, Millien G, Ramirez MI, Williams MC. Key developmental regulators change during hyperoxia-induced injury and recovery in adult mouse lung. J Cell Biochem 2007;100:1415–1429. [PubMed: 17167788]
27. Hofmann MA, et al. RAGE and arthritis: the G82S polymorphism amplifies the inflammatory response. Genes Immun 2002;3:123–135. [PubMed: 12070776]
28. Chavakis T, et al. The pattern recognition receptor (RAGE) is a counterreceptor for leukocyte integrins: a novel pathway for inflammatory cell recruitment. J Exp Med 2003;198:1507–1515. [PubMed: 14623906]
29. Englert JM, et al. A role for the receptor for advanced glycation end products in idiopathic pulmonary fibrosis. Am J Pathol 2008;172:583–591. [PubMed: 18245812]
30. Queisser MA, et al. Loss of RAGE in pulmonary fibrosis: molecular relations to functional changes in pulmonary cell types. Am J Respir Cell Mol Biol 2008;39:337–345. [PubMed: 18421017]
31. Lechin F, van der Dijs B, Orozco B, Lechin M, Lechin AE. Increased levels of free serotonin in plasma of symptomatic asthmatic patients. Ann Allergy Asthma Immunol 1996;77:245–253. [PubMed: 8814052]
32. Dupont LJ, et al. The effects of 5-HT on cholinergic contraction in human airways in vitro. Eur Respir J 1999;14:642–649. [PubMed: 10543288]
33. Idzko M, et al. The serotoninergic receptors of human dendritic cells: identification and coupling to cytokine release. J Immunol 2004;172:6011–6019. [PubMed: 15128784]
34. Bayer H, et al. Serotoninergic receptors on human airway epithelial cells. Am J Respir Cell Mol Biol 2007;36:85–93. [PubMed: 16873768]
35. Van Eerdewegh P, et al. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. Nature 2002;418:426–430. [PubMed: 12110844]

36. Jongepier H, et al. Polymorphisms of the ADAM33 gene are associated with accelerated lung function decline in asthma. Clin Exp Allergy 2004;34:757–760. [PubMed: 15144468]
37. van Diemen CC, et al. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. Am J Respir Crit Care Med 2005;172:329–333. [PubMed: 15879414]
38. Dijkstra A, et al. Expression of ADAMs ("a disintegrin and metalloprotease") in the human lung. Virchows Arch 2009;454:441–449. [PubMed: 19255780]
39. Bjarnadottir TK, Fredriksson R, Schioth HB. The adhesion GPCRs: a unique family of G protein-coupled receptors with important roles in both central and peripheral tissues. Cell Mol Life Sci 2007;64:2104–2119. [PubMed: 17502995]
40. Moriguchi T, et al. DREG, a developmentally regulated G protein-coupled receptor containing two conserved proteolytic cleavage sites. Genes Cells 2004;9:549–560. [PubMed: 15189448]
41. Gudbjartsson DF, et al. Many sequence variants affecting diversity of adult human height. Nat Genet 2008;40:609–615. [PubMed: 18391951]
42. Lettre G, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet 2008;40:584–591. [PubMed: 18391950]
43. Soranzo N, et al. Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. PLoS Genet 2009;5:e1000445. [PubMed: 19343178]
44. Baillat D, et al. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell 2005;123:265–276. [PubMed: 16239144]
45. Brandenberger R, et al. Identification and characterization of a novel extracellular matrix protein nephronectin that is associated with integrin alpha8beta1 in the embryonic kidney. J Cell Biol 2001;154:447–458. [PubMed: 11470831]
46. Huang JT, Lee V. Identification and characterization of a novel human nephronectin gene in silico. Int J Mol Med 2005;15:719–724. [PubMed: 15754038]
47. Wey E, Lyons GE, Schafer BW. A human POU domain gene, mPOU, is expressed in developing brain and specific adult tissues. Eur J Biochem 1994;220:753–762. [PubMed: 7908264]
48. Cardoso WV. Transcription factors and pattern formation in the developing lung. Am J Physiol 1995;269:L429–L442. [PubMed: 7485515]
49. Warburton D, et al. The molecular basis of lung morphogenesis. Mech Dev 2000;92:55–81. [PubMed: 10704888]
50. Uhl GR, et al. Genome-wide association for methamphetamine dependence: convergent results from 2 samples. Arch Gen Psychiatry 2008;65:345–355. [PubMed: 18316681]
51. Kohansal R, et al. The natural history of chronic airflow obstruction revisited: an analysis of the Framingham offspring cohort. Am J Respir Crit Care Med 2009;180:3–10. [PubMed: 19342411]
52. Eisner MD, et al. Secondhand smoke exposure, pulmonary function, and cardiovascular mortality. Ann Epidemiol 2007;17:364–373. [PubMed: 17300955]
53. The Atherosclerosis Risk in Communities (ARIC) Study: design objectives. The ARIC investigators. Am J Epidemiol 1989;129:687–702. [PubMed: 2646917]
54. Fried LP, et al. The Cardiovascular Health Study: design and rationale. Ann Epidemiol 1991;1:263–276. [PubMed: 1669507]
55. Dawber TR, Kannel WB. The Framingham study. An epidemiological approach to coronary heart disease. Circulation 1966;34:553–555. [PubMed: 5921755]
56. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. Prev Med 1975;4:518–525. [PubMed: 1208363]
57. Hofman A, et al. The Rotterdam Study: objectives and design update. Eur J Epidemiol 2007;22:819–829. [PubMed: 17955331]
58. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991;7:403–422. [PubMed: 1833235]

59. The National Heart, L., and Blood Institute. Chapel Hill, NC: ARIC Coordinating Center, School of Public Health, University of North Carolina; 1989. Atherosclerosis Risk in Communities (ARIC) Study. Quality assurance and quality control, version 1.0.
60. Enright PL, Kronmal RA, Higgins M, Schenker M, Haponik EF. Spirometry reference values for women and men 65 to 85 years of age. Cardiovascular health study. Am Rev Respir Dis 1993;147:125–133. [PubMed: 8420405]
61. Enright PL, Kronmal RA, Higgins MW, Schenker MB, Haponik EF. Prevalence and correlates of respiratory symptoms and disease in the elderly. Cardiovascular Health Study. Chest 1994;106:827–834. [PubMed: 8082366]
62. Givelber RJ, et al. Segregation analysis of pulmonary function among families in the Framingham Study. Am J Respir Crit Care Med 1998;157:1445–1451. [PubMed: 9603122]
63. van Durme YM, et al. Prevalence, incidence, and lifetime risk for the development of COPD in the elderly: the Rotterdam study. Chest 2009;135:368–377. [PubMed: 19201711]
64. Miller MR, et al. Standardisation of spirometry. Eur Respir J 2005;26:319–338. [PubMed: 16055882]
65. Li Y, Abecasis GR. Mach 1.0: Rapid Haplotype Reconstruction and Missing Genotype Inference. Am J Hum Genet 2006;S79:2290.
66. Guan Y, Stephens M. Practical issues in imputation-based association mapping. PLoS Genet 2008;4:e1000279. [PubMed: 19057666]
67. Pei YF, Li J, Zhang L, Papasian CJ, Deng HW. Analyses and comparison of accuracy of different genotype imputation methods. PLoS One 2008;3:e3551. [PubMed: 18958166]
68. Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–909. [PubMed: 16862161]
69. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575. [PubMed: 17701901]
70. The International HapMap Project. Nature 2003;426:789–796. [PubMed: 14685227]
71. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics 2007;23:1294–1296. [PubMed: 17384015]
71. Team, RDC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2007.
72. Abecasis GR, Cardon LR, Cookson WO, Sham PC, Cherny SS. Association analysis in a variance components framework. Genet Epidemiol 21 Suppl 2001;1:S341–S346.
73. Estrada K, et al. GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. Bioinformatics in press. 2009
74. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 2008;32:381–385. [PubMed: 18348202]

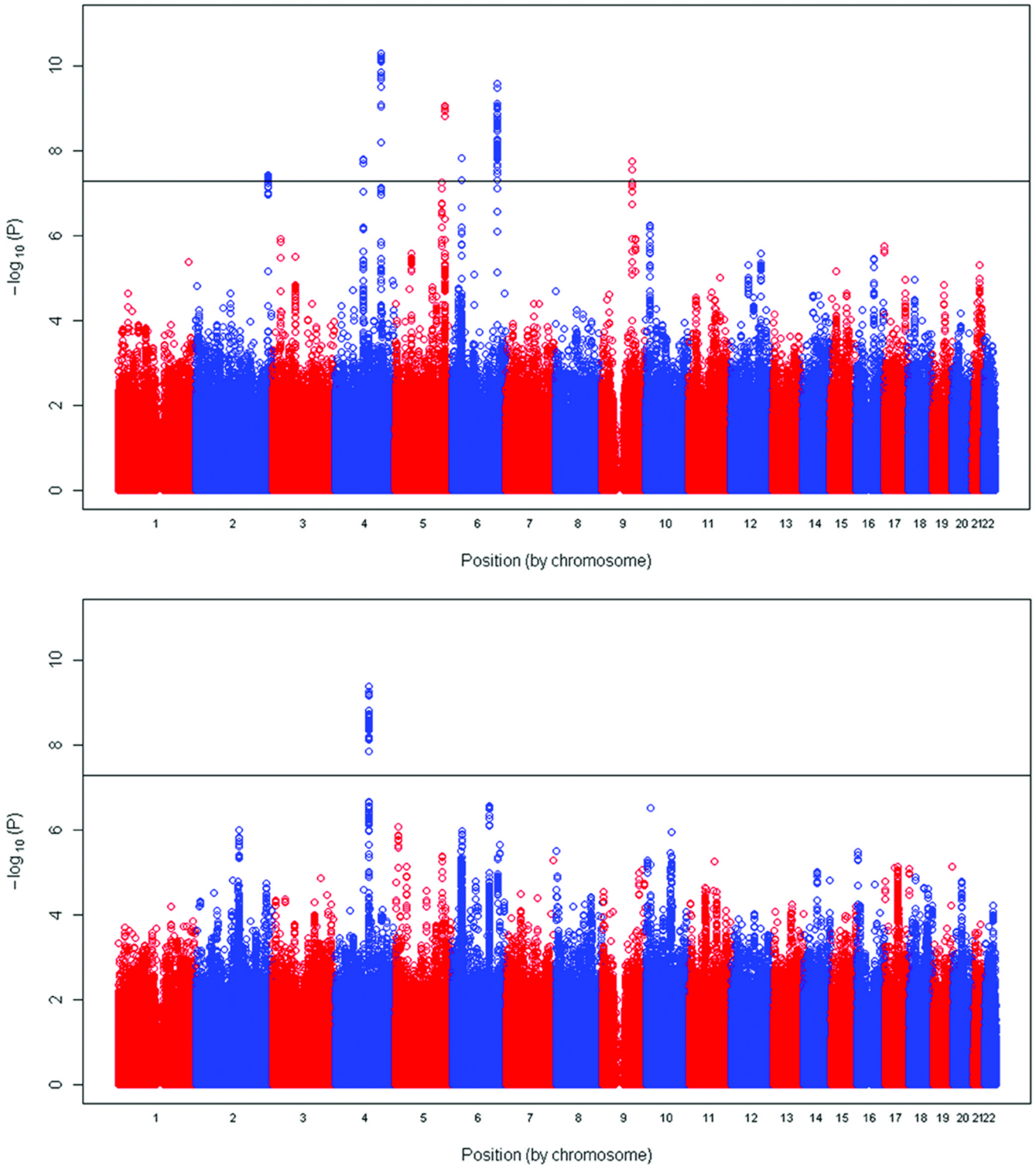
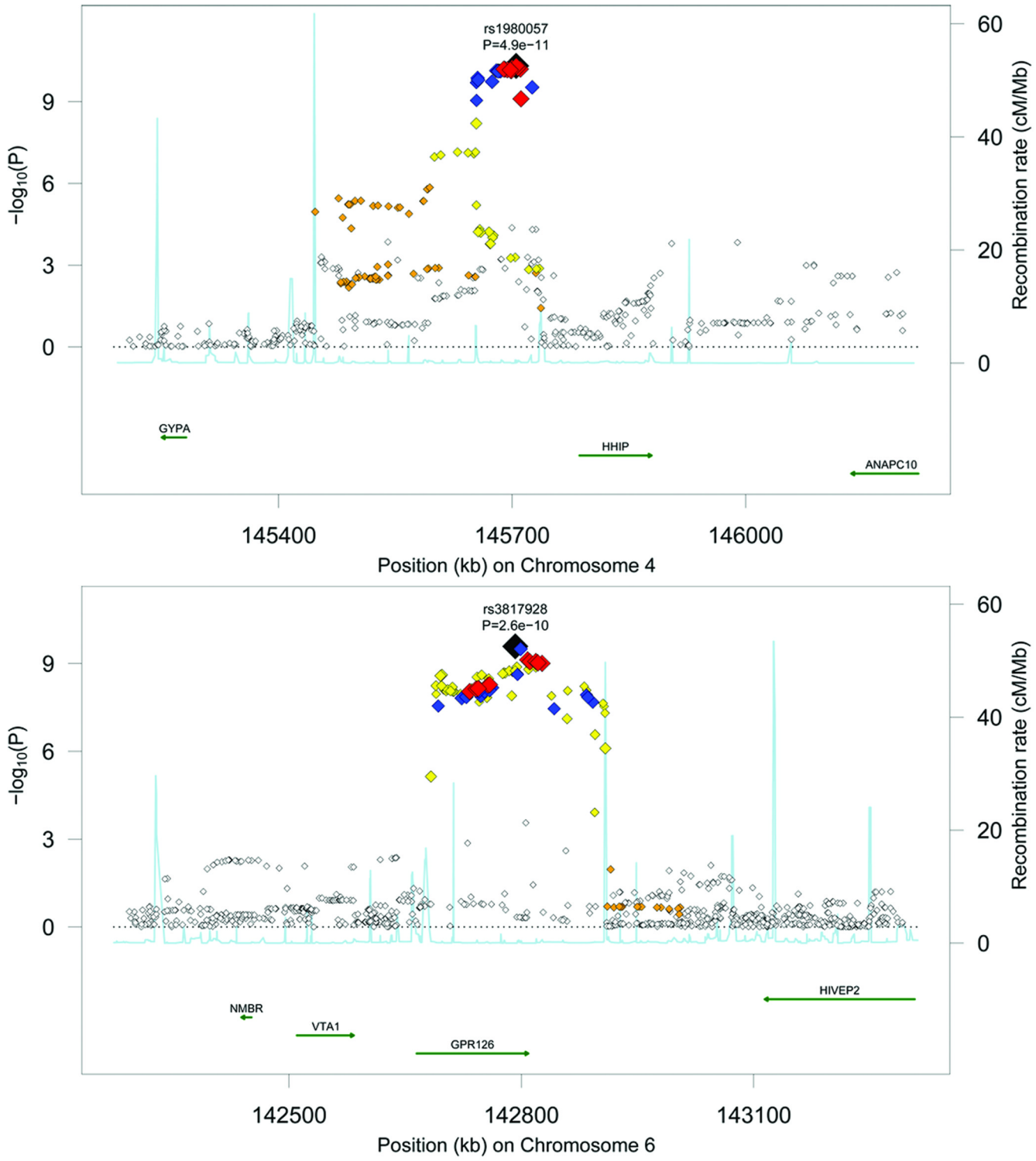
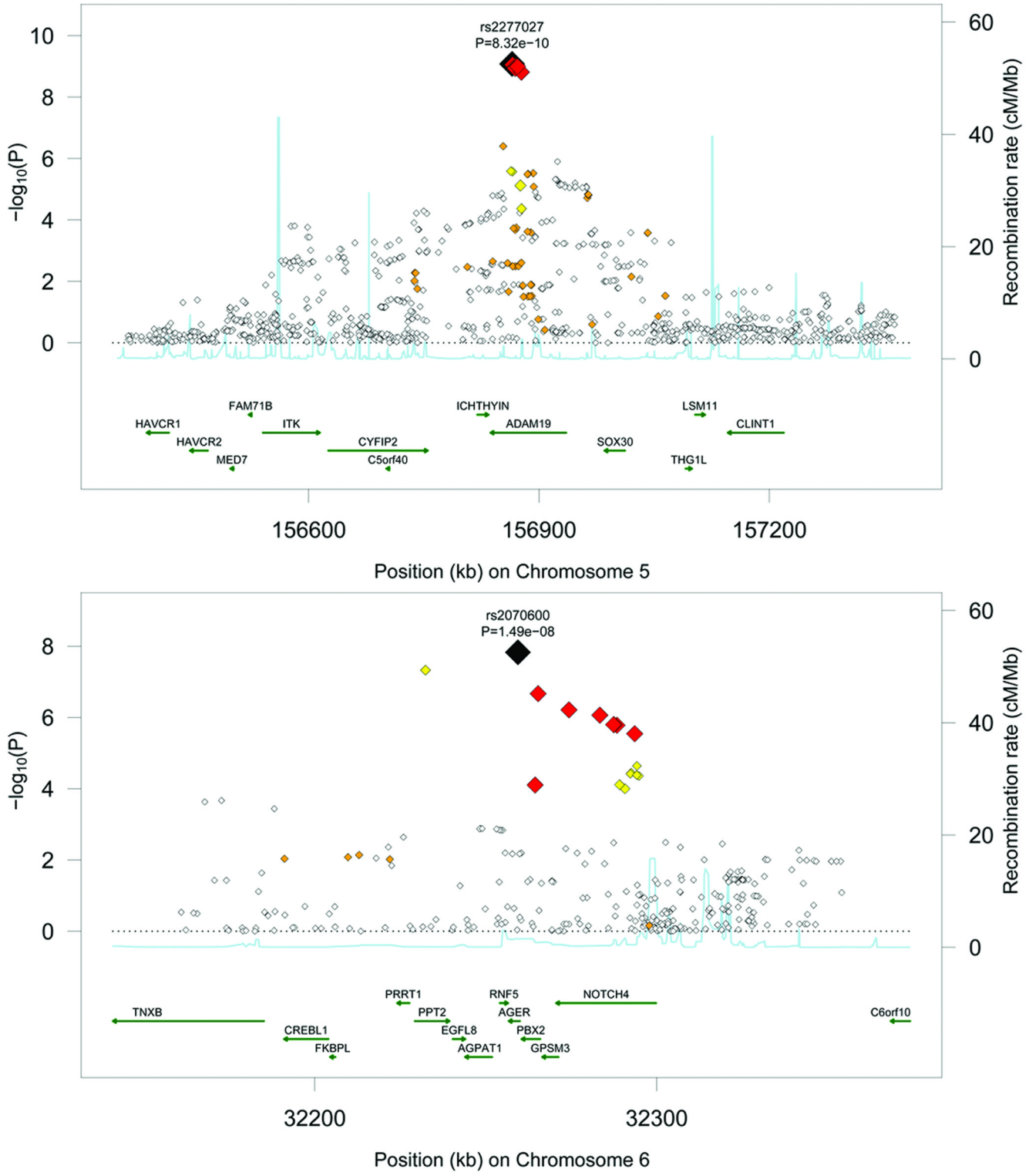


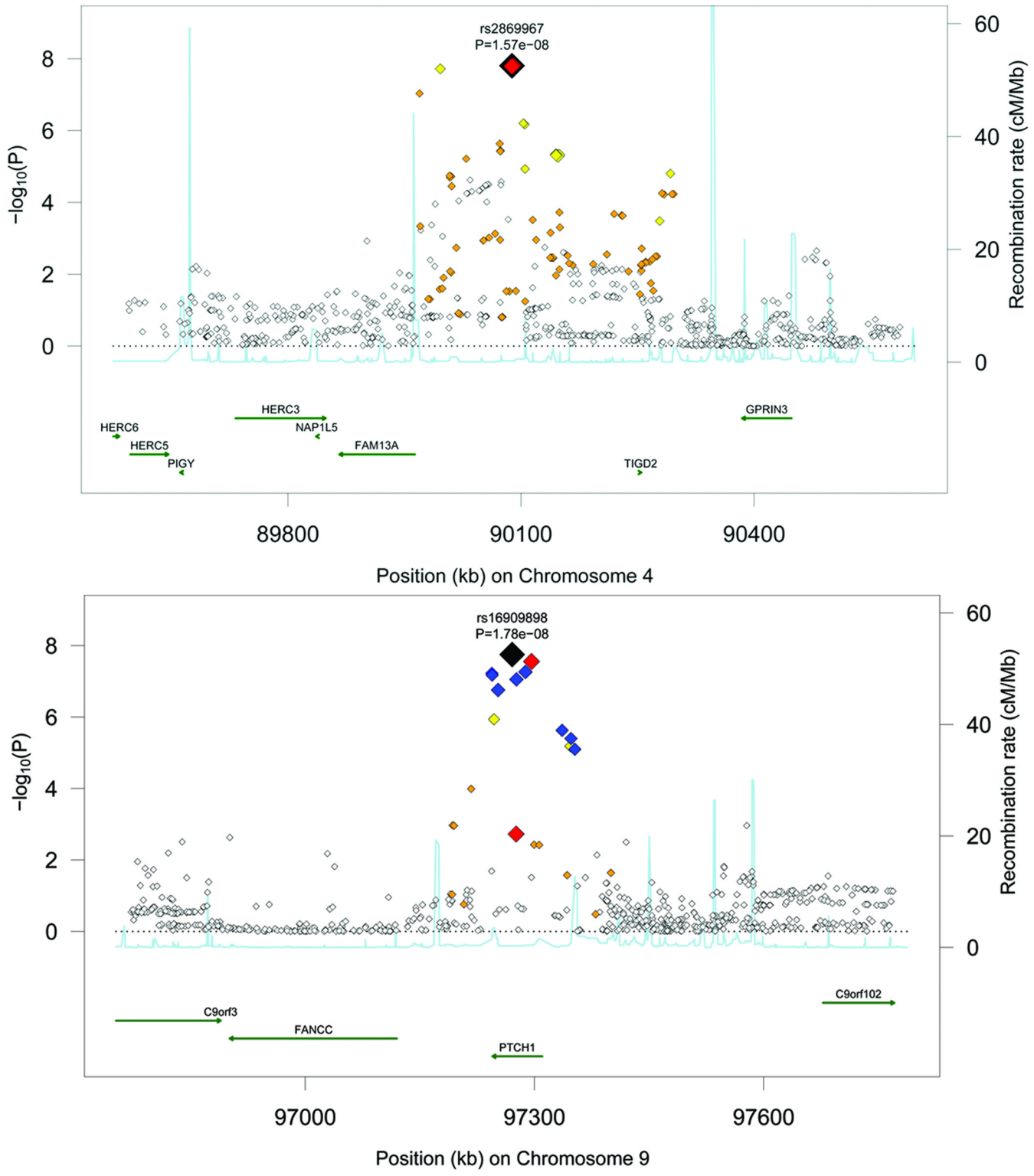
Figure 1.

Meta-analyses of approximately 2,534,500 SNPs tested for association with (a) FEV_1/FVC and (b) FEV_1 in all participants from the CHARGE consortium. The Manhattan plots (also

known as $-\log_{10}(P)$ association plots) show the chromosomal position of SNPs exceeding the genome-wide significance threshold ($P<5\times10^{-8}$ as indicated by the solid black line).







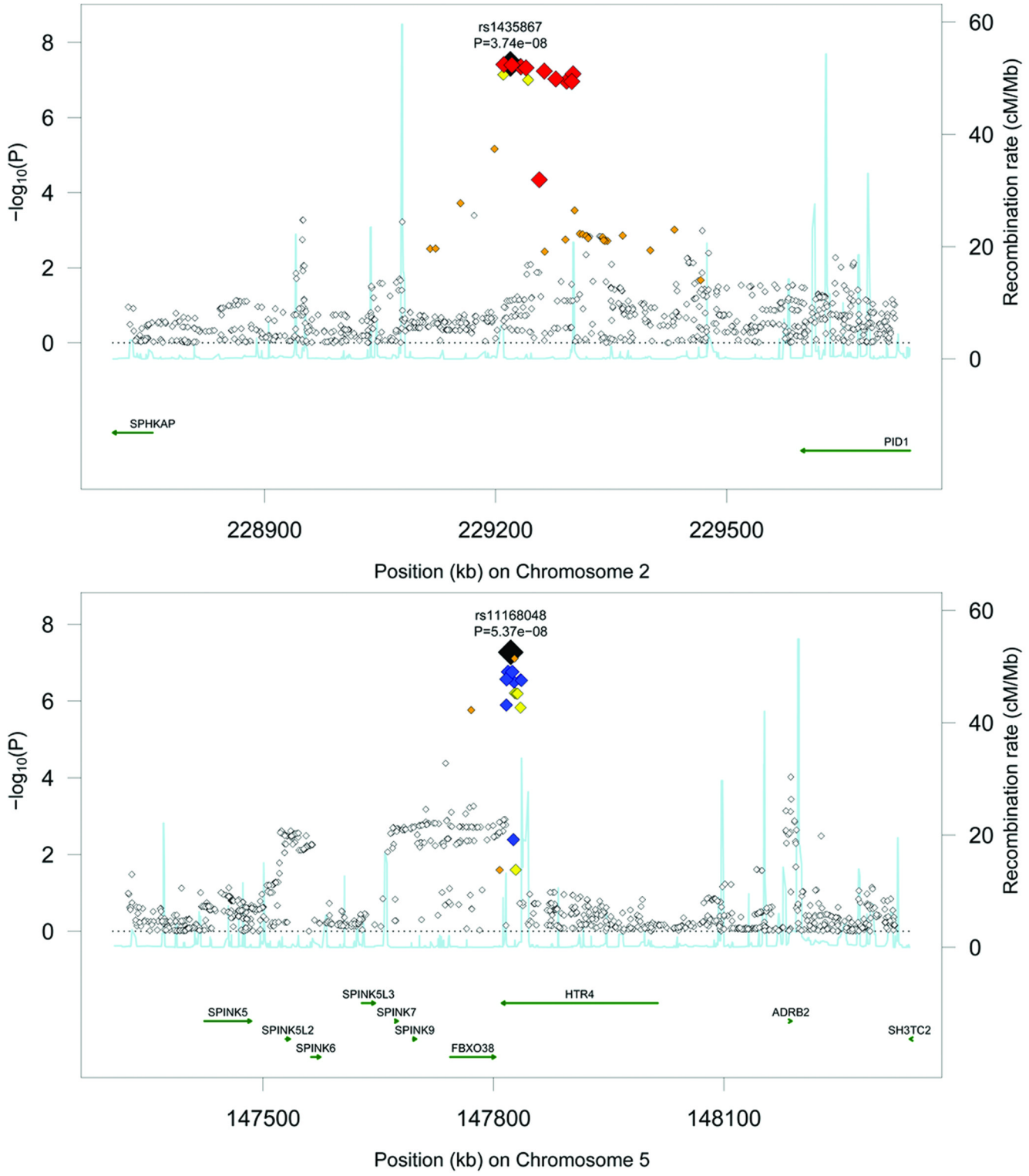


Figure 2.

Regional association plots for loci associated with FEV_1/FVC in the CHARGE consortium at or near genome-wide significance, including (a) *HHIP* on chromosome 4q31.22, (b) *GPR126* on chromosome 6q24.1, (c) *ADAM19* on chromosome 5q33.3, (d) *AGER-PPT2* on

chromosome 6p21.32, (e) *FAM13A* on chromosome 4q22.1, (f) *PTCH1* on chromosome 9q22.32, (g) *PID1* on chromosome 2q36.3, and (h) *HTR4* on chromosome 5q33.1. For each locus, correlations between the target SNP (the SNP with the lowest *P* value depicted in black) and other SNPs in the region are depicted in red when $r^2=1$, blue when $0.8\leq r^2<1$, yellow when $0.5\leq r^2<0.8$, orange when $0.2 \leq r^2<0.5$, and white when $r^2<0.2$. The r^2 values were based on the HapMap CEU population. Gene annotations are shown in green, and estimated recombination rates from HapMap are shown in light blue.

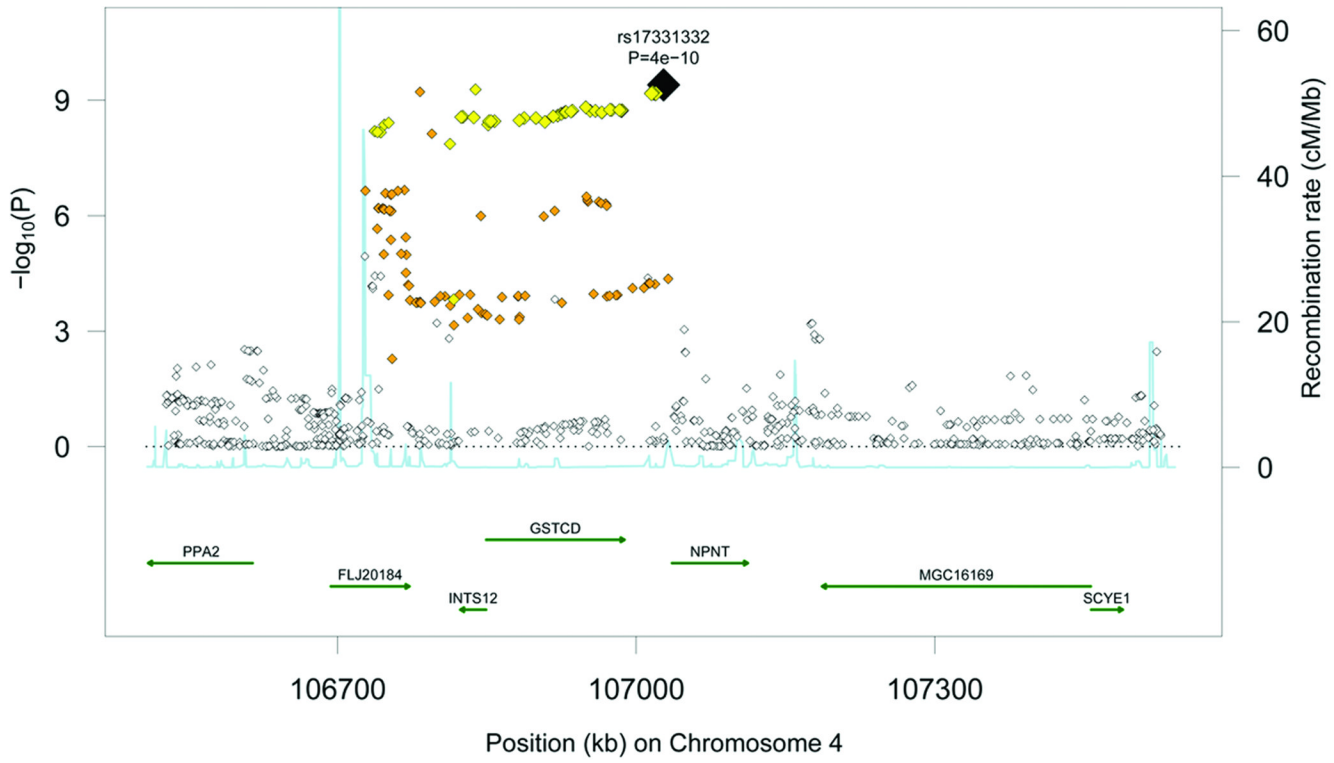


Figure 3.

Regional association plot for the chromosome 4q24 locus associated with FEV_1 in the CHARGE consortium at genome-wide significance, which includes *FLJ20184, INTS12, GSTCD,* and *NPNT*. Correlations between the target SNP (the SNP with the lowest *P* value depicted in black) and other SNPs in the region are depicted in red when $r^2=1$, blue when $0.8\leq r^2<1$, yellow when $0.5\leq r^2<0.8$, orange when $0.2 \leq r^2<0.5$, and white when $r^2<0.2$. The r^2 values were based on the HapMap CEU population. Gene annotations are shown in green, and estimated recombination rates from HapMap are shown in light blue.

Table 1

Characteristics of cohort participants in the CHARGE consortium at the time of pulmonary function assessment

Cohort, number	% male	Age (y)		Height (m)		BMI (kg/m^2)		Current smoking			Former smoking			Pulmonary function			
								Pack-years			Pack-years			FEV_1 (ml)		FEV_1/FVC (%)	
		Mean	s.d	Mean	s.d	Mean	s.d	%	Mean	s.d	%	Mean	s.d	Mean	s.d	Mean	s.d
ARIC 7,980	47.2	54.3	5.7	1.69	0.09	27.0	4.9	25.0	35.8	20.3	35.2	24.1	21.4	2,941	777	73.6	8.0
CHS 3,140	39.0	72.3	5.4	1.65	0.09	26.3	4.4	10.8	45.1	25.4	40.0	29.9	26.4	2,116	659	70.5	10.5
FHS 7,694	46.1	51.9	14.6	1.69	0.10	27.3	5.3	15.3	32.1	23.8	38.5	19.0	19.3	3,038	944	75.1	8.0
RS-I 1,224	45.4	74.5	5.6	1.67	0.09	27.4	4.0	11.7	40.1	23.4	58.8	24.7	22.5	2,320	728	73.1	8.3
RS-II 852	44.7	67.2	6.3	1.68	0.09	27.7	4.1	14.0	37.4	21.4	52.3	21.5	22.0	2,716	782	75.8	9.1

Genotyping platforms used: ARIC, Affymetrix GeneChip SNP Array 6.0; CHS, Illumina Human 370CNV BeadChip; FHS, Affymetrix GeneChip Human Mapping 500K Array and 50K Human Gene Focused Panel; RS-I, Illumina Infinium II HumanHap 550K Single and Duo Bead Chips; RS-II, Illumina Infinium II HumanHap 500K Dup and 610K Quad Bead Chips. FEV_1, forced expiratory volume in one second; FEV_1/FVC, ratio of forced expiratory volume in one second to forced vital capacity.

Table 2

Joint meta-analysis of SNPs selected from the top 8 loci implicated for FEV_1/FVC in the CHARGE GWAS and tested for replication with FEV_1/FVC in the SpiroMeta consortium. SNPs are grouped together by the nearest gene and ordered by the CHARGE GWAS *P* value. *P* values highlighted in bold exceeded the threshold for significance ($P<5\times10^{-8}$ for the GWAS and joint meta-analysis, $P<8.33\times10^{-4}$ for replication).

				CHARGE GWAS			SpiroMeta replication			Joint meta-analysis	
SNP	Chr	Gene/ Nearest gene	Reference allele	Allele frequency[a]	β, per-allele change in FEV_1/FVC (%)	*P*	Allele frequency[a]	β, per-allele change in FEV_1/FVC (%)	*P*	β, per-allele change in FEV_1/FVC (%)	*P*
rs1980057	4	*HHIP*	T	0.40	0.51	**4.90×10^{-11}**	0.45	0.54	**1.09×10^{-10}**	0.52	**3.21×10^{-20}**
rs1032295	4	*HHIP*	T	0.58	−0.47	**6.28×10^{-9}**	0.55	−0.46	**1.15×10^{-7}**	−0.47	**4.37×10^{-15}**
rs3817928	6	*GPR126*	A	0.78	−0.59	**2.60×10^{-10}**	0.79	−0.21	3.99×10^{-2}	−0.42	**1.17×10^{-9}**
rs7776375	6	*GPR126*	A	0.71	−0.51	**1.33×10^{-9}**	0.73	−0.19	4.87×10^{-2}	−0.37	**6.71×10^{-9}**
rs6937121	6	*GPR126*	T	0.71	−0.49	**2.46×10^{-9}**	0.72	−0.18	5.62×10^{-2}	−0.35	**1.25×10^{-8}**
rs11155242	6	*GPR126*	A	0.80	−0.54	**9.13×10^{-9}**	0.80	−0.16	1.26×10^{-1}	−0.37	1.45×10^{-7}
rs2277027	5	*ADAM19*	A	0.71	0.49	**8.32×10^{-10}**	0.66	0.24	5.04×10^{-3}	0.38	**9.93×10^{-11}**
rs1422795	5	*ADAM19*	T	0.66	0.48	**1.16×10^{-9}**	0.66	0.24	6.65×10^{-3}	0.37	**2.62×10^{-10}**
rs2070600	6	*AGER*[b]	T	0.04	1.06	**1.49×10^{-8}**	0.06	0.94	**4.40×10^{-7}**	1.00	**3.15×10^{-14}**
rs10947233	6	*PPT2*[b]	T	0.04	1.14	**4.71×10^{-8}**	0.04	1.05	**3.43×10^{-5}**	1.10	**6.66×10^{-12}**
rs2869967	4	*FAM13A*	T	0.61	0.44	**1.57×10^{-8}**	0.59	0.13	1.28×10^{-1}	0.30	1.91×10^{-7}
rs6830970	4	*FAM13A*	A	0.65	0.46	**1.92×10^{-8}**	0.64	0.11	2.00×10^{-1}	0.30	6.63×10^{-7}
rs16909898	9	*PTCH1*	A	0.90	0.77	**1.78×10^{-8}**	0.89	0.20	1.70×10^{-1}	0.50	5.34×10^{-7}
rs10512249	9	*PTCH1*	A	0.10	−0.73	**2.79×10^{-8}**	0.10	−0.19	1.67×10^{-1}	−0.48	5.75×10^{-7}
rs1435867	2	*PID1*	T	0.93	−0.84	**3.74×10^{-8}**	0.93	−0.075	6.51×10^{-1}	−0.50	1.53×10^{-5}
rs10498230	2	*PID1*	T	0.07	0.86	**3.87×10^{-8}**	0.07	0.075	6.56×10^{-1}	0.50	1.46×10^{-5}
rs11168048	5	*HTR4*	T	0.58	−0.43	5.37×10^{-8}	0.58	−0.36	**3.97×10^{-5}**	−0.40	**1.08×10^{-11}**
rs7735184	5	*HTR4*	T	0.40	0.40	1.74×10^{-7}	0.41	0.34	**5.62×10^{-5}**	0.37	**6.23×10^{-11}**

[a] Weighted average reference allele frequency for combined cohorts.

[b] *AGER* and *PPT2* SNPs are considered to represent one locus given their correlations.

FEV_1/FVC, forced expiratory volume in one second to forced vital capacity ratio

Table 3

Joint meta-analysis of SNPs selected from the top 3 loci implicated for FEV_1 in the CHARGE GWAS and tested for replication with FEV_1 in the SpiroMeta consortium. SNPs are grouped together by the nearest gene and ordered by the CHARGE GWAS *P* value. *P* values highlighted in bold exceeded the threshold for significance ($P<5\times10^{-8}$ for the GWAS and joint meta-analysis, $P<8.33\times10^{-4}$ for replication).

				CHARGE GWAS			SpiroMeta replication			Joint meta-analysis	
SNP	Chr	Gene/ Nearest gene	Reference allele	Allele frequency[a]	β, per-allele change in FEV_1 (mL)	*P*	Allele frequency[a]	β, per-allele change in FEV_1 (mL)	*P*	β, per-allele change in FEV_1 (mL)	*P*
rs17331332	4	*NPNT* [b]	A	0.08	60.35	**4.00×10^{-10}**	0.07	52.11	**2.45×10^{-6}**	56.79	**5.69×10^{-15}**
rs17036341	4	*NPNT* [b]	C	0.93	−58.85	**6.29×10^{-10}**	0.94	−53.82	**6.38×10^{-7}**	−56.65	**2.18×10^{-15}**
rs11727189	4	*INTS12* [b]	T	0.06	63.43	**5.28×10^{-10}**	0.05	66.38	**1.60×10^{-8}**	64.70	**4.66×10^{-17}**
rs17036090	4	*INTS12* [b]	T	0.93	−56.6	**1.37×10^{-8}**	0.94	−59.51	**7.85×10^{-8}**	−57.90	**5.61×10^{-15}**
rs17036052	4	*FLJ20184* [b]	T	0.06	71.43	**6.07×10^{-10}**	0.04	68.19	**5.78×10^{-7}**	70.08	**1.83×10^{-15}**
rs17035960	4	*FLJ20184* [b]	T	0.07	56.9	**3.87×10^{-9}**	0.06	49.96	**4.52×10^{-6}**	53.85	**9.42×10^{-14}**
rs11097901	4	*GSTCD* [b]	T	0.07	58.85	**1.52×10^{-9}**	0.06	59.49	**4.03×10^{-8}**	59.14	**3.26×10^{-16}**
rs11728716	4	*GSTCD* [b]	A	0.07	57.52	**1.80×10^{-9}**	0.06	57.41	**7.61×10^{-8}**	57.47	**7.20×10^{-16}**
rs3749893	6	*TSPYL4* [c,d]	A	0.37	−26.55	5.35×10^{-7}	0.37	4.01	4.62×10^{-1}	−11.71	2.05×10^{-3}
rs1052443	6	*NT5DC1* [c,d]	A	0.62	26.12	7.50×10^{-7}	0.62	−3.75	4.90×10^{-1}	11.61	2.17×10^{-3}
rs6555465	5	*ADCY2* [d]	A	0.18	−31.23	1.32×10^{-6}	0.19	0.12	9.86×10^{-1}	−16.24	4.96×10^{-4}
rs7710510	5	*ADCY2* [d]	T	0.19	−30.34	1.72×10^{-6}	0.21	−2.77	6.77×10^{-1}	−17.20	1.77×10^{-4}

[a] Weighted average reference allele frequency for combined cohorts.

[b] *NPNT*, *INTS12*, *FLJ20184*, and *GSTCD* SNPs are considered to represent one locus given their correlations.

[c] *TSPYL4* and *NT5DC1* SNPs are considered to represent one locus given their correlations.

[d] *TSPYL4*, *NT5DC1*, and *ADCY2* SNPs were not associated with FEV_1 at genome-wide significance, but these loci had the next smallest *P* values for association.

FEV_1, forced expiratory volume in one second