Genome-wide association study identifies five loci associated with lung function

Citation for published version:


Digital Object Identifier (DOI):
10.1038/ng.501

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Nature Genetics

Publisher Rights Statement:
Published in final edited form as:

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Genome-wide association study identifies five loci associated with lung function


© 2010 Nature America, Inc. All rights reserved.
Correspondence should be addressed to M.D.T. (mt47@leicester.ac.uk) or I.P.H. (ian.hall@nottingham.ac.uk). *A full list of author affiliations appears at the end of the paper.64 These authors contributed equally to this work.

AUTHOR CONTRIBUTIONS Author contributions are listed in alphabetical order. See Table 1 for definitions of study acronyms.


METHODS Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics. Note: Supplementary information is available on the Nature Genetics website.

COMPETING INTERESTS STATEMENT The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturegenetics./.

Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

URLs. UCSC browser, http://genome.ucsc.edu/.

1Departments of Health Sciences and Genetics, Adrian Building, University of Leicester, Leicester, UK. 2Division of Therapeutics and Molecular Medicine, Nottingham Respiratory Biomedical Research Unit, University Hospital of Nottingham, Nottingham, UK. 3Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary, University of London, London, UK. 4Medical Research Council (MRC) Epidemiology Unit, Institute of Metabolic Science, Cambridge, UK. 5Respiratory Epidemiology and Public Health Group, National Heart and Lung Institute, Imperial College London, London, UK. 6Department of Epidemiology and Public Health, Imperial College London, St. Mary’s Campus, London, UK. 7Department of Twin Research and Genetic Epidemiology, King’s College London, London, UK. 8MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, Scotland, UK. 9Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden. 10Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. 11Wellcome Trust Sanger Institute, Cambridge, UK. 12Department of Community-Based Medicine, University of Bristol, Bristol, UK. 13Department of Social Medicine, University of Bristol, Bristol, UK. 14Avon Longitudinal Study of Parents and Children (ALSPAC) Laboratory, Department of Social Medicine, University of Bristol, Bristol, UK. 15Division of Community Health Sciences, St. George’s University of London, London, UK. 16A full list of members is provided in the supplementary Note. 17Department of Public Health, University of Helsinki, Helsinki, Finland. 18Department of Health Sciences and Finnish Center for Interdisciplinary Gerontology, University of Jyväskylä, Jyväskylä, Finland. 19Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland. 20National Institute for Health and Welfare, Helsinki, Finland. 21Department of Chronic Disease Epidemiology/NICER, Institute of Social and Preventive Medicine, University of Zürich, Zürich, Switzerland. 22Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. 23Klinikum Grosshadern, Munich, Germany. 24Croatian Centre for Global Health, The University of Split Medical School, Split, Croatia. 25Andrija Stampar School of Public Health, Faculty of Medicine, University of Zagreb, Zagreb, Croatia. 26Department of Clinical Sciences, Obstetrics and Gynecology, Institute of Clinical Medicine, University of Oulu, Oulu, Finland. 27Centre for Population Health Sciences, University of Edinburgh, Edinburgh, Scotland, UK. 28Department of Internal Medicine B–Cardiology, Intensive Care, Pulmonary Medicine and Infectious Diseases, University of Greifswald, Greifswald, Germany. 29Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany. 30Institute for Community Medicine, SHIP/Clinical-Epidemiological Research, University of Greifswald, Greifswald, Germany. 31Gen-Info, Zagreb, Croatia. 32Department of LifeCourse and Services, National Institute for Health and Welfare, Oulu, Finland. 33Department of Microbiology and Immunology, Institute of Biomedicine. 34Occupational and Environmental Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. 35Centre for Genetic Epidemiology and Biostatistics, University of Western Australia, Perth, Western Australia, Australia. 36Department of Pulmonary Physiology/West Australian Sleep Disorders Research Institute, Sir Charles Gairdner Hospital, Western Australia, Australia.

Nat Genet. Author manuscript; available in PMC 2010 July 01.
Abstract

Pulmonary function measures are heritable traits that predict morbidity and mortality and define chronic obstructive pulmonary disease (COPD). We tested genome-wide association with forced expiratory volume in 1 s (FEV₁) and the ratio of FEV₁ to forced vital capacity (FVC) in the SpiroMeta consortium (n = 20,288 individuals of European ancestry). We conducted a meta-analysis of top signals with data from direct genotyping (n ≤ 32,184 additional individuals) and in silico summary association data from the CHARGE Consortium (n = 21,209) and the Health 2000 survey (n ≤ 883). We confirmed the reported locus at 4q31 and identified associations with FEV₁ or FEV₁/FVC and common variants at five additional loci: 2q35 in TNS1 (P = 1.11 × 10⁻¹²), 4q24 in GSTCD (2.18 × 10⁻²³), 5q33 in HTR4 (P = 4.29 × 10⁻⁹), 6p21 in AGER (P = 3.07 × 10⁻¹⁵) and 15q23 in THSD4 (P = 7.24 × 10⁻¹⁵). mRNA analyses showed expression of TNS1, GSTCD, AGER, HTR4 and THSD4 in human lung tissue. These associations offer mechanistic insight into pulmonary function regulation and indicate potential targets for interventions to alleviate respiratory disease.
determined. Twin studies in European and US populations give heritability estimates for FEV$_1$ as high as 0.77 (refs. 5,6). Longitudinal studies in families suggest that genetic effects are consistent over time7. Genetic determinants of pulmonary function seem to operate, at least in part, independent of disease status (such as asthma) and smoking status8, suggesting that population-based association studies are a viable way to identify key genetic determinants of lung function.

Adequately powered genome-wide association studies (GWAS) using hundreds of thousands of common SNPs can identify loci associated with common diseases and the quantitative traits that underlie them. Collaborative studies achieving sample sizes in excess of 10,000 have been able to identify associations with common genetic variants with typically modest effect sizes (usually <0.1 s.d.)9. In the past year, GWAS have reported association between an intergenic locus at chromosome 4q31 and FEV$_1$/FVC ratio and COPD, but no large-scale collaborative GWAS have yet been undertaken for lung function10,11.

If common SNPs underlying lung function have modest effects, very large sample sizes will be required to identify them. We therefore established the SpiroMeta consortium to facilitate large-scale meta-analysis of GWAS of lung function. Here we report a meta-analysis of GWAS in the SpiroMeta consortium, comprising 20,288 individuals of European ancestry, that tested association between cross-sectional lung function measures and ~2.5 million genotyped or imputed SNPs (stage 1). We followed up SNPs drawn from the most significantly associated loci in up to 32,184 individuals by direct genotyping (stage 2a) and using in silico summary association data relating to a further 22,092 individuals (stage 2b). These studies confirm the previous reported association at 4q31 and show that five previously unreported loci are robustly associated with lung function.

RESULTS

Genome-wide association with lung function (stage 1)

We included 14 studies of individuals of European ancestry, with sample sizes totaling 20,288 (Table 1). All individuals had measures of FEV$_1$ and FVC and smoking status recorded. FEV$_1$ and (separately) FEV$_1$/FVC measures were adjusted for age, age$^2$, sex, height and ancestry principal components within each study. Genome-wide genotyping was undertaken with a variety of platforms, and standard quality control measures were used (Online Methods and Supplementary Table 1). Genotypes were imputed for ~2.5 million autosomal SNPs from HapMap CEU data and tested for association separately for the inverse-normal transformed residuals of FEV$_1$ and FEV$_1$/FVC under an additive genetic model. We carried out meta-analysis of study-specific test statistics using an inverse variance weighting. We applied genomic control at the study and meta-analysis levels to avoid overinflation of test statistics owing to population structure or relatedness. Test statistic inflation before applying genomic control at the meta-analysis level was modest ($\lambda_{GC} = 1.046$ for FEV$_1$ and 1.035 for FEV$_1$/FVC). The plots of meta-analysis test statistics against expected values under the null hypothesis showed an excess of extreme values even after exclusion of the previously reported11 4q31 locus near HHIP, indicative of additional loci associated with lung function (Supplementary Fig. 1a,b).

We observed independent regions of association at 17 loci with $P < 1 \times 10^{-5}$ for FEV$_1$ and 23 for FEV$_1$/FVC (Figs. 1a,b and 2), including three regions (4q24 in GSTCD, 4q31 near HHIP and 15q23 in THSD4) that reached $P < 5 \times 10^{-8}$ in the stage 1 GWAS data alone, corresponding to a threshold of $P < 0.05$ after adjusting for 1 million independent tests12. SNP rs12504628, which was associated with both FEV$_1$/FVC ($P = 6.48 \times 10^{-13}$; Fig. 2c and Table 2) and FEV$_1$ ($P = 1.50 \times 10^{-10}$; Table 3), lies in an intergenic region upstream of
**HHIP** and spanning ~300 kb at 4q31 that has been associated with lung function11, COPD11 and height9. Our top SNP rs12504628 was in strong linkage disequilibrium (LD; $r^2 = 0.97$) with the previously reported SNP associated with lung function, rs13147758 ($P = 5.30 \times 10^{-10}$ for FEV$_1$ and $P = 1.11 \times 10^{-12}$ 1 for FEV$_1$/FVC in our data), and with SNPs associated previously with height (rs6854783, $r^2 = 0.55$; rs2055059, $r^2 = 0.48$), suggesting a role in skeletal growth and development. The hedgehog gene family, of which **HHIP** is a member, encodes signaling molecules involved in regulating lung morphogenesis, suggesting other mechanisms underlying these associations13. This intergenic region also contains multiple ESTs in human fetal lung (UCSC Browser).

**Follow-up analyses (stage 2)**

To validate potential associations with lung function, we selected 10 SNPs for further genotyping in additional studies of European ancestry (stage 2a, 32,184 individuals; Supplementary Table 2) and 30 SNPs for *in silico* follow-up (stage 2b; Supplementary Table 3). We obtained the *in silico* association results from the Health 2000 study (883 individuals) and from the CHARGE Consortium (21,209 individuals). Meta-analysis of the association results across stages 1, 2a and 2b showed five novel loci reaching genome-wide significance ($P < 5 \times 10^{-8}$): 2q35 in **TNS1**, 4q24 in **GSTCD**, 5q33 in **HTR4**, 6p21 in **AGER** and 15q23 in **THSD4** (Table 2 and Fig. 2). A further locus, 6p21 in **DAAM2**, which was not selected for further genotyping follow-up in stage 2a, fell just below the threshold for genome-wide significance for association with FEV$_1$/FVC after meta-analysis across stages 1 and 2b (rs2395730, $P = 7.98 \times 10^{-8}$; Supplementary Table 3 and Table 2).

The strongest association of FEV$_1$ was at 4q24 in **GSTCD** (rs10516526, $P = 2.18 \times 10^{-23}$; Table 2 and Fig. 2b). Relatively little is known about GSTCD, but the presence of the C-terminal α-helical domain common to the glutathione S-transferase (GST) family of enzymes suggests this protein is involved in cellular detoxification by catalyzing conjugation of glutathione to products of oxidative stress14. GST enzymes also show glutathione peroxidase activity regulating the synthesis of prostaglandins and leukotrienes14. To explore the potential function of GSTCD, we conducted a protein homology search and identified homology with chloride intracellular channels 1, 3, 4, 5 and 6, suggesting a role for GSTCD beyond the GST enzyme family. Genes in the region also include **INTS12** and **NPNT**: **INTS12** associates with RNA polymerase II and mediates 3′-end processing of small nuclear RNA15.

The second locus associated with FEV$_1$ was at 2q35, localized to the **TNS1** gene (nonsynonymous coding SNP rs2571445, $P = 1.11 \times 10^{-12}$; Table 2 and Fig. 2a). The protein this encodes, tensin-1, is an actin-binding protein that contains Src homology 2 domains, suggesting a role in linking cytoskeletal changes with signal transduction16. Tensin-1 may be functionally involved in cell migration17.

Multiple genes potentially underlie the third locus associated with FEV$_1$ at 5q33. The most strongly associated SNPs in this region, rs3995090 and rs689822 (P = 4.29 \times 10^{-9} and $P = 8.17 \times 10^{-9}$; Table 2 and Fig. 2d), are located in an intron in **HTR4** and are part of a cluster of associated SNPs also spanning a **SPINK5**-like gene, **SPINK7**, **SPINK9** and **FBXO38**: **HTR4**, which encodes 5-hydroxytryptamine receptor-4, is expressed in neurons in the respiratory pre-Bötzinger complex. Activation of this G protein–coupled receptor protects spontaneous respiratory activity18. Notably, selective antagonism of HTR4 in human bronchial strips in *vitro* attenuates the facilitation of electric field–stimulated cholinergic contraction by 5-hydroxytryptamine, suggesting a role for HTR4 in mediating airway caliber19. **HTR4** expression has recently been confirmed in airway epithelial type II cells, where receptor stimulation seems to regulate cytokine responses20. The **SPINK** family of serine protease inhibitors may have a role in antimicrobial protection of mucous epithelia21.

*Nat Genet. Author manuscript; available in PMC 2010 July 01.*
F-box protein-38 (encoded by *FBXO38*) is a member of a family of proteins that are believed to mediate protein-protein interactions and protein degradation.

The strongest association with FEV$_1$/FVC was at 6p21, a gene-rich region of the major histocompatibility complex (MHC). The extended LD in this region of the MHC prevented accurate localization of the association signal. However, we observed the peak of association for a nonsynonymous coding SNP in *AGER* (rs2070600, $P = 3.07 \times 10^{-15}$; Table 2 and Fig. 2e), which is a plausible candidate for causal association. *AGER*, also known as RAGE, is a multiligand receptor of the immunoglobulin superfamily. *AGER* is highly expressed in the lung, in particular alveolar epithelial cells, with a potential role in epithelium–extracellular matrix interactions. Reduced *AGER* expression has been identified in individuals with idiopathic pulmonary fibrosis, and *Ager*-/- mice develop age-related pulmonary fibrosis. Another candidate in this region is the nearby gene NOTCH4, a member of the family of transmembrane receptors involved in cell fate decisions. Notch4 is expressed in endothelial cells of the adult mouse lung, where it is believed to regulate angiogenesis.

The second identified association with FEV$_1$/FVC was at 15q23, encompassing the *THSD4* gene (rs12899618, $P = 7.24 \times 10^{-15}$; Table 2 and Fig. 2f). *THSD4* shows homology with members of the thrombospondin family of extracellular calcium-binding proteins that modulate cellular attachment, proliferation and migration and have been implicated in wound healing, inflammation and angiogenesis.

For each of the loci we reported, the estimated effect sizes were broadly consistent across the GWAS (Fig. 3).

### Association of variants with FVC

We tested the top SNP at each of the loci showing genome-wide significant association ($P < 5 \times 10^{-8}$) with FEV$_1$ or FEV$_1$/FVC for association with the other of the two traits, and with FVC in the stage 1 studies (Table 3). In addition to being associated with FEV$_1$, rs10516526 in *GSTCD* was associated with FVC ($P = 2.53 \times 10^{-7}$) but showed no discernible effect on FEV$_1$/FVC.

### Effect of smoking on SNP associations

Adjustment for ever-smoking status in the stage 1 data (Table 3) did not show materially different effect-size estimates for the associations with the sentinel SNPs in *TNS1, GSTCD, HTR4, AGER, THSD4* or *HHIP*. Similarly, adjustments for a quantitative measure of lifetime smoking exposure (pack-years) did not show substantially different effect-size estimates for the identified SNP associations (data not shown). We also tested the associations of the top SNPs in *TNS1, GSTCD, HTR4, AGER* and *THSD4* separately in ever-smokers and never-smokers (Supplementary Table 4); all $P$ values were $>0.05$ for tests of interaction between smoking status and these SNPs on lung function.

### Gene expression

We determined the mRNA expression profiles of *GSTCD, HHIP, THSD4, TNS1, HTR4, AGER* and *NOTCH4* in human lung tissue and a series of primary cells. We detected all transcripts in lung tissue (Supplementary Fig. 2a) and bronchial epithelial cells (Supplementary Fig. 2b); six transcripts (excluding *NOTCH4*) were present in human airway smooth muscle cells. We also detected *GSTCD, TNS1, HTR4, AGER* and *NOTCH4* transcripts in peripheral blood mononuclear cells (Supplementary Fig. 2b). For *AGER*, we noted the presence of two PCR products suggesting an unreported splice variant; we confirmed the presence of the splice variant by sequencing.
DISCUSSION

Our study reports a meta-analysis of GWAS results from 20,288 participants and follow-up analyses in 54,276 participants, identifying five novel genome-wide significant loci for pulmonary function. The regions identified were 4q24 (GSTCD), 2q35 (TNS1) and 5q33 (HTR4) for FEV\textsubscript{1}, and 6p21 (AGER) and 15q23 (THSD4) for FEV\textsubscript{1}/FVC. In addition, we identified a region suggestive of association with FEV\textsubscript{1}/FVC at 6p21 in DAAM2. The companion manuscript from the CHARGE Consortium, which reports a GWAS of lung function in 20,890 participants, also identifies genome-wide significant associations at GSTCD, HTR4 and AGER. Both SpiroMeta and CHARGE confirmed the previously reported association between FEV\textsubscript{1} and FEV\textsubscript{1}/FVC and the 4q31 locus upstream of HHIP.

Our findings have several important implications. First, the loci identified were observed in the whole population studied and were not specific to smokers. The presence of genetic determinants of lung function that do not depend on prior smoking exposure has been suggested by previous studies of heritability. This does not rule out a possible subset of genetic determinants with effects on lung function that are partially or wholly dependent on smoking exposure.

We have also attempted to address the issue of genetic factors that influence smoking behavior. We did not observe any association with the CHRNA3-CHRNA5-CHRNB4 locus previously reported to be associated with cigarette smoke exposure, lung cancer, peripheral arterial disease and COPD (rs1051730, P = 0.23 for FEV\textsubscript{1} and 0.56 for FEV\textsubscript{1}/FVC). The associations we show in GSTCD, TNS1, HTR4, THSD4 and AGER do not seem to be attenuated by adjustment for qualitative or quantitative adjustments for smoking exposure. None of these loci have been implicated in published GWAS of smoking quantity, although a recent report suggested a role for TSHD4 variants in smoking cessation.

SNPs showing association with height could also show association with lung function measures because of incomplete adjustment for height, or because of SNP effects on skeletal growth with consequences for both height and lung function. The 4q31 locus near HHIP has shown convincing association with height. An association was recently reported between height and rs185819 at 6p21 (ref. 34). Although this association signal was broad, reflecting the extended LD across this region of the MHC, rs185819 was in weak LD ($r^2 = 0.069$) with rs2070600 (the sentinel SNP we reported for FEV\textsubscript{1}/FVC in AGER). These findings leave open the possibility of shared genetic determinants of growth of pulmonary function and height, but they do not suggest that our findings are primarily accounted for by inadequate adjustment for height.

The level of FEV\textsubscript{1} at a given time point in an individual depends on two potentially independent processes: the maximum lung function obtained during development, and the rate of decline of lung function with age. Lung function reaches a maximum by age 25–35 years. The populations studied in SpiroMeta cover a wide range of ages except the very elderly; as expected, FEV\textsubscript{1} and FVC values were much lower in children. At least for the loci we identify, there was little evidence for age-specific effects, suggesting that the genetic risk factors identified operate across the age ranges; these findings again are in keeping with those of previous epidemiological studies. Our analyses were based on cross-sectional measures of lung function; additional studies in cohorts with longitudinal data will be needed to identify determinants of the gradients of development and decline in lung function with age.

The magnitude of the estimated effect on untransformed FEV\textsubscript{1} of rs10516526 in GSTCD was 52 ml per copy of the G allele (frequency, 0.06). This is equivalent to about 3 years of
FEV\textsubscript{1} decline in the nonsmoking population\textsuperscript{35}. Allelic effect sizes on FEV\textsubscript{1} of the more common variants (minor allele frequencies \textasciitilde0.4) were 19–23 ml for rs3995090 in \textit{HTR4} and rs2571445 in \textit{TNS1}. Individually, the five loci we describe account for a small proportion (0.07\%–0.14\%) of the variance in FEV\textsubscript{1} and in FEV\textsubscript{1}/FVC (Table 2) in the general population.

After exclusion of the locus near \textit{HHIP} and the five reported regions, meta-analysis test statistics still showed an excess of extreme values compared with expected values under the null, particularly for FEV\textsubscript{1}. Although we cannot rule out the possibility of residual population stratification, this indicates the potential to detect further loci associated with lung function (Supplementary Fig. 1a,b). We have provided a list of the top 2000 associations for FEV\textsubscript{1} and for FEV\textsubscript{1}/FVC (Supplementary Table 5) as a resource to other investigators.

We imputed nongenotyped SNPs using two software implementations\textsuperscript{36,37} that share similar underlying population genetic models\textsuperscript{38}. This methodology facilitates meta-analysis across different marker sets and improves coverage across the genome, and its utility has been empirically shown in several large GWAS. However, the power to detect associations with rare alleles is limited. The loci we report include two relatively infrequent SNPs, \textit{GSTCD} (rs10516526, minor allele frequency 0.06) and \textit{AGER} (rs2070600, minor allele frequency 0.05); these SNPs were directly genotyped in the majority of stage 1 subjects (16,514 and 15,386 individuals, respectively).

The associations we report relate to the general population but were of comparable magnitude after the exclusion of documented cases of asthma or COPD (data not shown). Although pulmonary function is an important predictor of morbidity and mortality \textit{per se}, it will be important to investigate, in appropriately powered studies, whether the risk alleles in the genes identified in this study act as independent susceptibility markers for COPD or influence the development of airway obstruction in other diseases, such as asthma.

Our expression profiling studies identified expression of all of the candidate genes in relevant tissues. Further work is required to elucidate the mechanisms underlying the novel association signals we describe. In broad terms, however, it is notable that the most probable candidate genes in the regions identified seem to be involved either in developmental pathways important for lung growth or in tissue remodeling pathways that might be expected to alter airway architecture.

In conclusion, the results presented here from the SpiroMeta consortium, together with those reported by the CHARGE Consortium\textsuperscript{30}, provide strong evidence for newly identified genetic loci that act as important determinants of pulmonary function.

**ONLINE METHODS**

**Study design**

The study consisted of two stages. In stage 1, a meta-analysis was conducted on directly genotyped and imputed SNPs from 14 studies of individuals of European ancestry, with a total sample size of 20,288. Details of these studies are given in Table 1. This meta-analysis provided loci for further genotyping in up to 32,184 individuals of European origin (stage 2a) and \textit{in silico} comparisons in 22,092 individuals of European origin (stage 2b).

**Stage 1 samples**

The SpiroMeta consortium consists of 14 GWAS studies: ALSPAC, B58C-T1DGC, B58C-WTCCC, EPIC (obese and population-based substudies), the EUROSPAN studies (Korcula,
NSPHS, ORCADES and Vis), FTC (incorporating the FinnTwin16 and Finnish Twin Study on Aging), KORA S3, NFBC1966, SHIP and TwinsUK (see Table 1 for definitions of acronyms). The primary analyses on FEV$_1$ and FEV$_1$/FVC included 20,288 individuals of European descent. The measurements of FEV$_1$ and FVC are described in the Supplementary Note.

**Genome-wide genotyping and quality control**

The platforms used were Affymetrix 500K GeneChip array (four studies), Illumina HumanHap 550 Beadchip (one study), Illumina 317K (four studies), Affymetrix Genome-Wide SNP6.0 (one study), Illumina Hap370Cnv (one study), Illumina Hap300 v1 (one study) and Illumina Hap300 v2 (two studies). Each individual study applied quality-control criteria as described in Supplementary Table 1.

**Imputation**

Imputation of nongenotyped SNPs was undertaken with MACH36 or IMPUTE37 with preimputation filters and parameters as shown in Supplementary Table 1. SNPs were excluded if the imputation information, assessed using r$^2$ (MACH) or .info (IMPUTE), was <0.3. In total, 2,705,257 autosomal SNPs were analyzed.

**Transformation of data and genotype-phenotype association analysis**

Linear regression of age, age$^2$, sex, height and ancestry principal components was undertaken on FEV$_1$ (milliliters) and FEV$_1$/FVC (percentage). The residuals were transformed to ranks and subsequently to normally distributed z scores, and were then used as the phenotype for association testing under an additive genetic model using software specified in Supplementary Table 1. Appropriate tests for association in related individuals were applied where necessary, as described in the Supplementary Note.

**Meta-analysis of stage 1 data**

All stage 1 study effect estimates were corrected using genomic control\textsuperscript{40} and were oriented to the forward strand of the NCBI build 36 reference sequence of the human genome, consistently using the alphabetically higher allele as the coded allele. Study-specific lambda estimates are shown in Supplementary Table 1. The pooled effect-size estimate and s.e.m. were computed using inverse variance weighting, and genomic control was applied to the pooled effect-size estimates. To describe the effect of imperfect imputation on power, we report ‘N effective’, the sum of the study-specific products of the sample size and the imputation quality metric. Meta-analysis statistics and figures were produced using R version 2.7.0.

**Selection of SNPs for stage 2**

Ten leading SNPs were selected for stage 2a genotyping follow-up (Supplementary Table 2). Thirty leading SNPs were selected for stage 2b in silico exchange, according to $P$ value (under the threshold of $5 \times 10^{-5}$), N effective (≥70% of the total sample size) and evidence from supporting SNPs (Supplementary Table 3).

**Stage 2a samples (follow-up genotyped data)**

We genotyped 10 SNPs in up to 32,184 individuals from the ADONIX, BHS, BRHS, BWHHS, Gedling, GS:SFHS, HCS, KORA F4, NFBC1986, Nottingham Smokers and NSHD studies. The characteristics of the studies are summarized in Table 1, and stage 2a study information is provided in the Supplementary Note.
Stage 2b samples (in silico data)

The CHARGE Consortium includes four population-based studies with data on FEV₁ and FEV₁/FVC: the Atherosclerosis Risk in Communities (ARIC) study, the Cardiovascular Health Study (CHS), the Framingham Heart Study (FHS) and the Rotterdam Study (RS). Details are provided in the companion paper in this issue from the CHARGE Consortium. Given differences between the analysis approaches for GWAS adopted by the SpiroMeta and CHARGE consortia, the CHARGE analyses were undertaken using the analysis approach adopted by the SpiroMeta consortium (21,209 individuals; larger than the sample in the companion paper, which excluded subjects with missing or incomplete pack-years data). We also included 883 population-based subjects from the Health 2000 study in the stage 2b analysis.

Combined analysis of stage 1 and 2 samples

Meta-analysis of data from stages 1, 2a and 2b was conducted using inverse variance weighting. We described associations as genome-wide significant if \( P < 5 \times 10^{-8} \).

Secondary analyses

To examine the effect of smoking on the causal pathway between the SNPs and the traits of interest, an adjustment for smoking was applied. The subgroups of ‘ever-smokers’ and ‘never-smokers’ were analyzed separately, and the stratum-specific estimated effects were combined within each individual study using inverse variance weights before meta-analyzing over studies. Additional adjustments were undertaken by adjusting for pack-years among the ever-smokers with these data available, and repeating the analyses.

PCR expression profiling

The mRNA expression profiles of GSTCD, HHIP, THSD4, TNS1, HTR4, AGER and NOTCH4 were determined in human lung tissue and primary cell samples using RT-PCR, including RNA from lung (Ambion/ABI), brain, airway smooth muscle cells and human bronchial epithelial cells (Clonetics). Peripheral blood mononuclear cells were isolated from whole blood using 6% (w/v) dextran and 42%–51% (v/v) Percoll gradients (Sigma). Ethical approval for the use of primary cells was obtained from the local ethics committees. Total RNA was extracted from samples using an RNeasy kit (Qiagen) as directed by the manufacturer. cDNA was generated from 1 μg of RNA template using random hexamers and a SuperScript kit (Invitrogen) as directed by the manufacturer. PCR assays were designed to cross intron-exon boundaries and where splice variation was known, in order to detect all variants. Primer sequences are given in Supplementary Table 6. All PCR was done using Platinum Taq High Fidelity (Invitrogen) with 100 ng of cDNA template in a 25-μl reaction. Cycling conditions were as follows: 94 °C for 3 min, 35 cycles of 94 °C for 45 s, 55 °C for 30 s, and 72 °C for 90 s.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the many colleagues who contributed to collection and phenotypic characterization of the clinical sampling, genotyping and analysis of the GWAS data. We especially thank those who kindly agreed to participate in the studies.

Major funding for this work is from the following sources (in alphabetical order): Academy of Finland (including project grants 104781, 120315 and 1114194) and Center of Excellence in Complex Disease Genetics; Arthritis Research Campaign; Asthma UK; AstraZeneca; Biocenter Oulu, University of Oulu; Biocentrum Helsinki;

Nat Genet. Author manuscript; available in PMC 2010 July 01.
Biotechnology and Biological Sciences Research Council project grant; British Heart Foundation (including project grants PG06154/22043 and PG97012) and Senior Research Fellowship FS05/125; British Lung Foundation; Cancer Research United Kingdom; Chief Scientists Office, part of the Scottish Government Health Directorate (including grant CZD/16/6); Department of Health Air Pollution PRP (ref. no. 0200029); ENGAGE project (HEALTH-F4-2007-201413); European Commission (EURO-BLCS, FP-5/QLG1-CT-2000-01643, FP-7/2007-2013, FP-6 LSBB-CT-2006-018996 (GABRIEL), FP-6 LSBB-CT-2006-01947 (EUROSPAN), HEALTH-F2-2008-201865-GEFOS and FP-5 GenomEUtwin project QLG2-CT-2002-01254); Finnish Ministry of Education; German Federal Ministry of Education and Research (BMBF, including grants 01ZZ06030, 01ZZ0701 and 01GI0883 and German Asthma and COPD Network (COSYCONET) grant 01GI0883; German Ministry for Education, Research and Cultural Affairs; German National Genome Research Network (NGFN-2 and NGFN-plus); Healthway, Western Australia; HEFCE Science Research Investment Fund; Helmholtz Zentrum München; German Research Center for Environmental Health, Neuherberg, Germany; International Osteoporosis Foundation; Juvenile Diabetes Research Foundation International; Leicester Biomedical Research Unit in Cardiovascular Science (NIHR); Medical Research Council UK (including grants G0500539, G0501942, G0000943 and G990146); Medical Research Fund of the Tampere University Hospital; Ministry for Social Affairs of the Federal State of Mecklenburg-West Pomerania; MRC Human Genetics Unit; Munich Center of Health Sciences, as part of LMUinnovativ; National Human Genome Research Institute; National Institute for Health Research comprehensive Biomedical Research Centre award to Guy’s & St. Thomas’ NHS Foundation Trust in partnership with King’s College London; National Institute for Health Research Cambridge Biomedical Research Centre; National Institute of Allergy and Infectious Diseases; National Institute of Child Health and Human Development; National Institute of Diabetes and Digestive and Kidney Diseases; National Heart, Lung, and Blood Institute (grant 5R01HL087679-02 through the STAMPEED program (IRL1MH083268-01)); Oulu University Hospital; PHOEBE (FP6, LHSG-CT-2006-518418); Public Population Project in Genomics (Genome Canada and Genome Quebec), Republic of Croatia Ministry of Science, Education and Sports (research grant 108-1080315-0302); Royal Society; Siemens Health Care Sector; Swedish Heart and Lung Foundation (grant 20050561); Swedish Medical Research Council (project no. K2007-60X-2070-01-3); Swedish Research Council for Working Life and Social Research (FAS, grants 2001-0263 and 2003-0139); the Great Wine Estates of the Margaret River region of Western Australia; UBS Wealth Foundation (grant BA298Q7-DZZ); UK Department of Health Policy Research Programme; University of Nottingham; University of Bristol; US National Institutes of Health (U01 DK062418); US National Institutes of Health–National Institute of Mental Health (U01 MH67306-02); Wellcome Trust (including grants 068545/Z/02, 076113/B/04/Z, 079895, 077016/Z/05/Z, 075883 and 086160/Z/08/A); and Zentren für Innovationskompetenz (BMBF grant 03ZIK012).

References


Figure 1. Manhattan plots of association results for FEV₁ and FEV₁/FVC (analysis stage 1). (a, b) Manhattan plots ordered by chromosome position. SNPs for which $-\log_{10} P > 5$ are indicated in red. The six loci indicated by arrows showed association with FEV₁ (a) or FEV₁/FVC (b, $P < 5 \times 10^{-8}$) in the meta-analysis of data from stages 1, 2a and 2b.
Figure 2.
Regional association plots of six lung function–associated loci. (a–f) Statistical significance of each SNP on the $-\log_{10}$ scale as a function of chromosome position (NCBI build 36) in the meta-analysis of stage 1 data alone. The sentinel SNP at each locus is shown in blue; the correlations ($r^2$) of each of the surrounding SNPs to the sentinel SNP are shown in the indicated colors. The six loci included are those that showed association with FEV$_1$ or FEV$_1$/FVC ($P < 5 \times 10^{-8}$) in the meta-analysis of data from stages 1, 2a and 2b. The combined $P$ values for all stages are indicated by arrows. The relevant trait (FEV$_1$ or FEV$_1$/FVC ratio) is indicated for each plot. For rs12504628, the plot shows only the association of FEV$_1$/FVC; this SNP was associated ($P < 5 \times 10^{-8}$) with both FEV$_1$ and FEV$_1$/FVC. Fine-scale recombination rate is plotted in blue39. Combined $P$ value from stages 1 and 2a only; SNP rs3995090 had low imputation quality in the CHARGE Consortium data and so was not included in stage 2b.
Figure 3.
Forest plots of the stage 1 meta-analysis for the six lung function–associated loci. Each of the SNPs included in the figure showed genome-wide significant association ($P < 5 \times 10^{-8}$) with either FEV$_1$ or FEV$_1$/FVC in the data from stages 1, 2a and 2b. The plots show the meta-analysis of the stage 1 data for each sentinel SNP. The contributing effect (transformed beta) from each study is shown by a square, with confidence intervals indicated by horizontal lines. The contributing weight of each study to the meta-analysis is indicated by the size of the square. The combined meta-analysis estimate in the stage 1 data is shown at the bottom of each graph.
Table 1

study characteristics

<table>
<thead>
<tr>
<th>Study</th>
<th>N total</th>
<th>N male</th>
<th>N female</th>
<th>Age range (y) at FEV1/FVC measurement</th>
<th>Mean age, y (s.d.)</th>
<th>Mean FEV1, l (s.d.)</th>
<th>Mean FVC, l (s.d.)</th>
<th>Mean FEV1/FVC (s.d.)</th>
<th>N neversmokers</th>
<th>N smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1: GWAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALSAC</td>
<td>1,391</td>
<td>676</td>
<td>715</td>
<td>8–9</td>
<td>8.7 (0)</td>
<td>1.70 (0.26)</td>
<td>1.93 (0.31)</td>
<td>0.88 (0.07)</td>
<td>1,391</td>
<td>0</td>
</tr>
<tr>
<td>B3C T1DGC</td>
<td>2,343</td>
<td>1,131</td>
<td>1,212</td>
<td>44–45</td>
<td>44.5 (0)</td>
<td>3.31 (0.78)</td>
<td>4.19 (0.96)</td>
<td>0.79 (0.08)</td>
<td>692</td>
<td>1,651</td>
</tr>
<tr>
<td>B3C WTCCC</td>
<td>1,372</td>
<td>691</td>
<td>681</td>
<td>44–45</td>
<td>44.5 (0)</td>
<td>2.93 (0.75)</td>
<td>4.18 (0.96)</td>
<td>0.79 (0.08)</td>
<td>394</td>
<td>978</td>
</tr>
<tr>
<td>EPIC obese cases</td>
<td>1,104</td>
<td>476</td>
<td>628</td>
<td>39–76</td>
<td>59.1 (8.8)</td>
<td>2.35 (0.69)</td>
<td>2.84 (0.87)</td>
<td>0.82 (0.17)</td>
<td>489</td>
<td>615</td>
</tr>
<tr>
<td>EPIC population based</td>
<td>2,336</td>
<td>1,100</td>
<td>1,236</td>
<td>39–77</td>
<td>59.2 (9.0)</td>
<td>2.50 (0.72)</td>
<td>3.04 (0.90)</td>
<td>0.85 (0.16)</td>
<td>1,061</td>
<td>1,275</td>
</tr>
<tr>
<td>FTC</td>
<td>134</td>
<td>13</td>
<td>121</td>
<td>23–76</td>
<td>57.4 (19.3)</td>
<td>2.69 (0.94)</td>
<td>2.93 (0.61)</td>
<td>0.79 (0.09)</td>
<td>104</td>
<td>30</td>
</tr>
<tr>
<td>KORA S3</td>
<td>555</td>
<td>261</td>
<td>294</td>
<td>29–73</td>
<td>47.6 (9.0)</td>
<td>3.43 (0.78)</td>
<td>4.18 (0.99)</td>
<td>0.83 (0.07)</td>
<td>266</td>
<td>289</td>
</tr>
<tr>
<td>Korcula</td>
<td>825</td>
<td>300</td>
<td>525</td>
<td>18–90</td>
<td>55.5 (13.5)</td>
<td>2.84 (0.81)</td>
<td>3.37 (0.93)</td>
<td>0.84 (0.09)</td>
<td>397</td>
<td>428</td>
</tr>
<tr>
<td>NFBC1966</td>
<td>4,556</td>
<td>2,182</td>
<td>2,374</td>
<td>31–31</td>
<td>31.0 (0)</td>
<td>3.96 (0.79)</td>
<td>4.73 (0.99)</td>
<td>0.84 (0.06)</td>
<td>1,648</td>
<td>2,908</td>
</tr>
<tr>
<td>NSPHS</td>
<td>549</td>
<td>255</td>
<td>294</td>
<td>18–91</td>
<td>50.0 (19.1)</td>
<td>3.02 (0.95)</td>
<td>3.68 (1.12)</td>
<td>0.82 (0.09)</td>
<td>464</td>
<td>85</td>
</tr>
<tr>
<td>ORCARES</td>
<td>692</td>
<td>322</td>
<td>370</td>
<td>19–93</td>
<td>54.9 (15.3)</td>
<td>2.88 (0.84)</td>
<td>3.58 (0.98)</td>
<td>0.80 (0.09)</td>
<td>404</td>
<td>288</td>
</tr>
<tr>
<td>SHIP</td>
<td>1,777</td>
<td>870</td>
<td>907</td>
<td>25–85</td>
<td>52.3 (13.7)</td>
<td>3.28 (0.89)</td>
<td>3.87 (1.03)</td>
<td>0.87 (0.06)</td>
<td>773</td>
<td>1,004</td>
</tr>
<tr>
<td>TwinUK</td>
<td>1,885</td>
<td>0</td>
<td>1,885</td>
<td>18–79</td>
<td>48.4 (12.2)</td>
<td>2.73 (0.56)</td>
<td>3.40 (0.61)</td>
<td>0.80 (0.08)</td>
<td>943</td>
<td>942</td>
</tr>
<tr>
<td>Vis</td>
<td>769</td>
<td>323</td>
<td>446</td>
<td>18–88</td>
<td>56.3 (15.3)</td>
<td>3.39 (1.22)</td>
<td>4.38 (1.43)</td>
<td>0.77 (0.09)</td>
<td>328</td>
<td>441</td>
</tr>
<tr>
<td>Stage 1 sample size</td>
<td>20,288</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 2a: studies with direct genotyping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADONIX</td>
<td>1,338</td>
<td>635</td>
<td>703</td>
<td>25–75</td>
<td>49.2 (13.6)</td>
<td>3.35 (0.86)</td>
<td>4.24 (1.02)</td>
<td>0.79 (0.07)</td>
<td>743</td>
<td>595</td>
</tr>
<tr>
<td>BHS</td>
<td>4,350</td>
<td>1,793</td>
<td>2,557</td>
<td>18–96</td>
<td>50.1 (17.0)</td>
<td>3.02 (0.97)</td>
<td>3.89 (1.16)</td>
<td>0.77 (0.08)</td>
<td>2,459</td>
<td>1,891</td>
</tr>
<tr>
<td>BRHS</td>
<td>3,897</td>
<td>3,897</td>
<td>0</td>
<td>60–79</td>
<td>68.7 (5.5)</td>
<td>2.57 (0.69)</td>
<td>3.37 (0.84)</td>
<td>0.77 (0.11)</td>
<td>1,132</td>
<td>2,765</td>
</tr>
<tr>
<td>BWHHS</td>
<td>3,644</td>
<td>0</td>
<td>3,644</td>
<td>59–80</td>
<td>68.8 (5.5)</td>
<td>1.98 (0.52)</td>
<td>2.82 (0.76)</td>
<td>0.71 (0.09)</td>
<td>2,060</td>
<td>1,584</td>
</tr>
<tr>
<td>Gedling</td>
<td>1,263</td>
<td>632</td>
<td>631</td>
<td>27–80</td>
<td>56.2 (12.3)</td>
<td>2.85 (0.85)</td>
<td>3.68 (1.02)</td>
<td>0.77 (0.07)</td>
<td>633</td>
<td>630</td>
</tr>
<tr>
<td>GS/SHFS</td>
<td>5,474</td>
<td>2,254</td>
<td>3,220</td>
<td>18–89</td>
<td>46.0 (14.3)</td>
<td>3.15 (0.87)</td>
<td>4.11 (1.03)</td>
<td>0.77 (0.10)</td>
<td>3,005</td>
<td>2,469</td>
</tr>
<tr>
<td>HCS</td>
<td>2,850</td>
<td>1,511</td>
<td>1,339</td>
<td>59–73</td>
<td>66.1 (2.8)</td>
<td>2.44 (0.68)</td>
<td>3.42 (0.92)</td>
<td>0.72 (0.09)</td>
<td>1,319</td>
<td>1,531</td>
</tr>
<tr>
<td>KORA F4</td>
<td>1,305</td>
<td>610</td>
<td>695</td>
<td>41–61</td>
<td>51.6 (5.7)</td>
<td>3.32 (0.81)</td>
<td>4.28 (1.00)</td>
<td>0.78 (0.06)</td>
<td>499</td>
<td>806</td>
</tr>
<tr>
<td>NFBC1986</td>
<td>4,946</td>
<td>2,379</td>
<td>2,567</td>
<td>15–77</td>
<td>16.0 (0.38)</td>
<td>3.77 (0.70)</td>
<td>4.30 (0.84)</td>
<td>0.88 (0.08)</td>
<td>3,708</td>
<td>1,238</td>
</tr>
<tr>
<td>Nottingham Smokers</td>
<td>509</td>
<td>280</td>
<td>229</td>
<td>40–89</td>
<td>59.5 (10.4)</td>
<td>2.00 (0.95)</td>
<td>3.01 (1.06)</td>
<td>0.64 (0.16)</td>
<td>0</td>
<td>509</td>
</tr>
<tr>
<td>Study</td>
<td>$N_{total}$</td>
<td>$N_{male}$</td>
<td>$N_{female}$</td>
<td>Age range (y) at FEV$_1$/FVC measurement</td>
<td>Mean age, y (s.d.)</td>
<td>Mean FEV$_1$, l (s.d.)</td>
<td>Mean FVC, l (s.d.)</td>
<td>Mean FEV$_1$/FVC (s.d.)</td>
<td>Neversmokers</td>
<td>$N_{never-smokers}$</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------</td>
<td>------------</td>
<td>--------------</td>
<td>------------------------------------------</td>
<td>-------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>---------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>NSHD</td>
<td>2,608</td>
<td>1,308</td>
<td>1,300</td>
<td>53–53</td>
<td>53 (0)</td>
<td>2.80 (0.70)</td>
<td>3.50 (0.89)</td>
<td>0.81 (0.19)</td>
<td>1,080</td>
<td>1,528</td>
</tr>
<tr>
<td>Stage 2a sample size</td>
<td>32,184</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stage 2b: studies with in silico data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health 2000</td>
<td>883</td>
<td>427</td>
<td>456</td>
<td>30–75</td>
<td>50.2 (11.0)</td>
<td>3.32 (0.91)</td>
<td>4.19 (1.08)</td>
<td>0.79 (0.07)</td>
<td>266</td>
<td>617</td>
</tr>
<tr>
<td>CHARGE$^a$ (ref. 30)</td>
<td>21,209</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 2b sample size</td>
<td>22,092</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total sample size</strong></td>
<td>74,564</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Characteristics are shown for studies analyzed in stage 1 (GWAS meta-analysis), stage 2a (direct genotyping follow-up, 10 SNPs) and stage 2b (in silico follow-up, 30 SNPs). Stage 1 studies: ALSPAC, Avon Longitudinal Study of Parents and Children; B58C-T1DGC, British 1958 Birth Cohort–Type 1 Diabetes Genetics Consortium; B58C-WTCCC, British 1958 Birth Cohort–Wellcome Trust Case Control Consortium; EPIC obese cases, European Prospective Investigation into Cancer and Nutrition, Obese Cases; EPIC population based, European Prospective Investigation into Cancer and Nutrition Cohort; FTC, Finnish Twin Cohort incorporating FinnTwin16 and FITSA; KORA S3, Cooperative Health Research in the Region of Augsburg; the Korcula study; NFBC1966, Northern Finland Birth Cohort of 1966; NSPHS, Northern Sweden Population Health Study; ORCADES, Orkney Complex Disease Study; SHIP, Study of Health in Pomerania; the TwinUK study; the Vis study. Stage 2a studies: ADONIX, Adult-Onset Asthma and Nitric Oxide; BHS, Busselton Heart Study; BRHS, British Regional Heart Study; BWHHS, British Women’s Heart and Health Study; the Gedling study; GS:SFHS, Generation Scotland: Scottish Family Health Study; HCS, Hertfordshire Cohort Study; KORA F4, Cooperative Health Research in the Region of Augsburg; NFBC1986, Northern Finland Birth Cohort of 1986; the Nottingham Smokers study; NSHD, Medical Research Council National Survey of Health and Development (also known as the British 1946 Birth Cohort). Stage 2b studies: Health 2000, Finnish Health 2000 survey; CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology.

$^a$ Characteristics of the constituent studies of the CHARGE Consortium are presented in the companion paper from the CHARGE Consortium30.
### Table 2

Loci associated with lung function

<table>
<thead>
<tr>
<th>Chr.</th>
<th>Measure</th>
<th>SNP ID (NCBIID6 position), function</th>
<th>Coded allele (function)</th>
<th>Stage</th>
<th>Coded allele freq.</th>
<th>N</th>
<th>Beta (s.e.m.)</th>
<th>P</th>
<th>Joint meta-analysis of all stages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Beta (s.e.m.)</td>
<td></td>
<td>P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>r²</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N total</td>
</tr>
<tr>
<td>2</td>
<td>FEV₁</td>
<td>rs2571445 (218,391,399), TNS1 (ns)</td>
<td>G</td>
<td>2a</td>
<td>0.60</td>
<td>18,710</td>
<td>0.057 (0.011)</td>
<td>2.41 × 10⁻⁷</td>
<td>0.035 (0.005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2b</td>
<td>0.61</td>
<td>29,922</td>
<td>0.031 (0.008)</td>
<td>9.03 × 10⁻⁵</td>
<td>0.050 (0.004)</td>
</tr>
<tr>
<td>4</td>
<td>FEV₁</td>
<td>rs10516526 (106,908,353), GSTCD (intron)</td>
<td>G</td>
<td>2a</td>
<td>0.06</td>
<td>20,179</td>
<td>0.135 (0.022)</td>
<td>6.67 × 10⁻¹⁰</td>
<td>0.089 (0.009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2b</td>
<td>0.07</td>
<td>31,217</td>
<td>0.092 (0.013)</td>
<td>8.51 × 10⁻¹²</td>
<td>0.081 (0.009)</td>
</tr>
<tr>
<td></td>
<td>FEV₁/FVC</td>
<td>rs12504628 (145,655,774), HHIP (upstream)</td>
<td>T</td>
<td>2a</td>
<td>0.56</td>
<td>19,400</td>
<td>-0.077 (0.011)</td>
<td>6.48 × 10⁻¹³</td>
<td>-0.077 (0.011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2b</td>
<td>0.56</td>
<td>9,032</td>
<td>-0.077 (0.011)</td>
<td>6.48 × 10⁻¹³</td>
<td>-0.077 (0.011)</td>
</tr>
<tr>
<td>5</td>
<td>FEV₁</td>
<td>rs3998990 (147,826,008), HTR4 (intron)</td>
<td>C</td>
<td>2a</td>
<td>0.41</td>
<td>18,792</td>
<td>0.058 (0.011)</td>
<td>1.36 × 10⁻⁷</td>
<td>0.038 (0.006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2b</td>
<td>0.41</td>
<td>29,630</td>
<td>0.029 (0.008)</td>
<td>3.04 × 10⁻⁴</td>
<td>0.087 (0.004)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.36</td>
<td>883</td>
<td>-0.037 (0.050)</td>
<td>5.76 × 10⁻⁴</td>
<td>0.087 (0.004)</td>
</tr>
<tr>
<td>5</td>
<td>FEV₁</td>
<td>rs6889822 (147,826,900), HTR4 (intron)</td>
<td>G</td>
<td>2a</td>
<td>0.40</td>
<td>19,652</td>
<td>0.057 (0.011)</td>
<td>1.73 × 10⁻⁷</td>
<td>0.034 (0.006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2b</td>
<td>0.38</td>
<td>21,442</td>
<td>0.025 (0.0072)</td>
<td>5.76 × 10⁻⁴</td>
<td>0.034 (0.006)</td>
</tr>
<tr>
<td>6</td>
<td>FEV₁/FVC</td>
<td>rs2070600 (32,259,421), AGER (ms)</td>
<td>T</td>
<td>2a</td>
<td>0.05</td>
<td>19,183</td>
<td>0.115 (0.023)</td>
<td>5.86 × 10⁻⁷</td>
<td>0.088 (0.011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2b</td>
<td>0.04</td>
<td>26,229</td>
<td>0.069 (0.014)</td>
<td>7.90 × 10⁻⁷</td>
<td>0.088 (0.011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
<td>21,998</td>
<td>0.130 (0.031)</td>
<td>2.24 × 10⁻⁵</td>
<td>0.088 (0.011)</td>
</tr>
<tr>
<td>6</td>
<td>FEV₁/FVC</td>
<td>rs2395730 (39,892,343), DAAM2 (intron)</td>
<td>C</td>
<td>2a</td>
<td>0.419</td>
<td>19,887</td>
<td>0.048 (0.011)</td>
<td>5.86 × 10⁻⁶</td>
<td>0.044 (0.008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2b</td>
<td>0.424</td>
<td>21,560</td>
<td>0.038 (0.013)</td>
<td>3.22 × 10⁻³</td>
<td>0.044 (0.008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19,875</td>
<td>0.082 (0.014)</td>
<td>7.17 × 10⁻⁹</td>
<td>0.044 (0.008)</td>
</tr>
<tr>
<td>15</td>
<td>FEV₁/FVC</td>
<td>rs12889618 (69,432,174), THSD4 (intron)</td>
<td>G</td>
<td>2a</td>
<td>0.84</td>
<td>26,124</td>
<td>0.043 (0.011)</td>
<td>7.51 × 10⁻³</td>
<td>0.060 (0.008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2b</td>
<td>0.85</td>
<td>21,050</td>
<td>0.072 (0.018)</td>
<td>4.21 × 10⁻⁵</td>
<td>0.060 (0.008)</td>
</tr>
</tbody>
</table>

Shown are the top SNPs or SNPs for each independent locus associated (P < 5 × 10⁻⁷) with FEV₁ or FEV₁/FVC in a joint analysis of up to 74,564 individuals of European ancestry from the SpiroMeta GWAS (stage 1), follow-up genotyping (stage 2a) and in silico data from the CHARGE Consortium (companion paper in this issue 30) and the Health 2000 study (stage 2b). The six genome-wide significant (P < 5 × 10⁻⁸) loci are indicated in bold. Results are also shown for rs2395730 (DAAM2), as this locus fell just below genome-wide significance after meta-analysis of stage 1 and 2b data (it was not selected for stage 2a genotyped follow-up). For the GWAS (stage 1 and 2b) data, the sample sizes (N) shown are the effective sample sizes. Effective sample size within each study is the product of sample size and imputation quality metric. Total sample size is the sum of the N_effective values across all stage 1 and 2b studies, plus the sample size from the directly genotyped (stage 2a) studies; for each SNP, this is lower than 74,564 individuals owing to missing genotypes and imperfect imputation. Joint meta-analysis includes data from stages 1, 2a and 2b. Beta values reflect effect-size estimates on an inverse-normal transformed scale after adjustments for age, age², sex, height and ancestry principal components. SNP rs12504628 showed association with both FEV₁ and FEV₁/FVC (P < 5 × 10⁻⁸); statistics for the strongest association (FEV₁/FVC) are shown. As this association has been previously published, rs12504628 was not assessed in the follow-up studies, so statistics are not presented for stages 2a and 2b for rs12504628. SNP rs3995090 was genotyped in stage 2a follow-up studies and was available in Health 2000 study data, but owing to low imputation quality for rs3995090 in the CHARGE Consortium data, an alternative SNP in the region, rs6889822, was selected for in silico exchange with the CHARGE Consortium. The estimated proportion of variance explained by each SNP in the joint meta-analysis is shown (r²), ns, nonsynonymous coding SNP.
Table 3

Relation of SNPs at genome-wide significant loci to \( \text{FEV}_1 \), FVC and \( \text{FEV}_1/\text{FVC} \), and impact of adjustment for smoking in stage 1 (spiroMeta GWAs) data

<table>
<thead>
<tr>
<th>SNP ID (NCBI36 position), function</th>
<th>Chr.</th>
<th>Coded allele</th>
<th>Noncoded allele</th>
<th>N</th>
<th>Measure</th>
<th>Unadjusted stage 1 analyses</th>
<th>Smoking-adjusted stage 1 analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Beta</td>
<td>s.e.m.</td>
</tr>
<tr>
<td>rs2571445 (218391399), TNS1 (ns)</td>
<td>2</td>
<td>G</td>
<td>A</td>
<td>18,710</td>
<td>( \text{FEV}_1 )</td>
<td>0.057</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FVC} )</td>
<td>0.041</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FEV}_1/\text{FVC} )</td>
<td>0.026</td>
<td>0.011</td>
</tr>
<tr>
<td>rs10516526 (106908333), GSTCD (intron)</td>
<td>4</td>
<td>G</td>
<td>A</td>
<td>20,179</td>
<td>( \text{FEV}_1 )</td>
<td>0.135</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FVC} )</td>
<td>0.112</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FEV}_1/\text{FVC} )</td>
<td>0.038</td>
<td>0.022</td>
</tr>
<tr>
<td>rs12504628 (145655734), HHIP (upstream)</td>
<td>4</td>
<td>T</td>
<td>C</td>
<td>19,400</td>
<td>( \text{FEV}_1 )</td>
<td>-0.068</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FVC} )</td>
<td>-0.021</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FEV}_1/\text{FVC} )</td>
<td>-0.077</td>
<td>0.011</td>
</tr>
<tr>
<td>rs3995090 (147826008), HTR4 (intron)</td>
<td>5</td>
<td>C</td>
<td>A</td>
<td>18,792</td>
<td>( \text{FEV}_1 )</td>
<td>0.058</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FVC} )</td>
<td>0.023</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FEV}_1/\text{FVC} )</td>
<td>0.045</td>
<td>0.011</td>
</tr>
<tr>
<td>rs6889822 (147826900), HTR4 (intron)</td>
<td>5</td>
<td>G</td>
<td>A</td>
<td>19,652</td>
<td>( \text{FEV}_1 )</td>
<td>0.057</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FVC} )</td>
<td>0.025</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FEV}_1/\text{FVC} )</td>
<td>0.046</td>
<td>0.011</td>
</tr>
<tr>
<td>rs2070600 (32559421), AGER (ns)</td>
<td>6</td>
<td>T</td>
<td>C</td>
<td>19,183</td>
<td>( \text{FEV}_1 )</td>
<td>0.027</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FVC} )</td>
<td>-0.043</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FEV}_1/\text{FVC} )</td>
<td>0.115</td>
<td>0.023</td>
</tr>
<tr>
<td>rs12899618 (69432174), THSD4 (intron)</td>
<td>15</td>
<td>G</td>
<td>A</td>
<td>19,875</td>
<td>( \text{FEV}_1 )</td>
<td>0.028</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FVC} )</td>
<td>-0.024</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \text{FEV}_1/\text{FVC} )</td>
<td>0.082</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Each SNP included in table showed genome-wide significant association (\( P < 5 \times 10^{-8} \)) with either \( \text{FEV}_1 \) or \( \text{FEV}_1/\text{FVC} \) in data from stages 1, 2a and 2b. Beta values shown reflect effect-size estimates on an inverse-normal transformed scale in the stage 1 data. Effective sample size was calculated within each study as the product of sample size and imputation quality metric and were summed across studies to calculate total effective sample size (N). Estimates from analyses unadjusted for smoking status and adjusted for ever-smoking versus never-smoking status are shown.