Genome-wide association and large scale follow-up identifies 16 new loci influencing lung function

María Soler Artigas, Daan W Loth, Louise V Wain, Sina A Gharib, Ma'en Obeidat, Wenbo Tang, Guangju Zhai, Jing Hua Zhao, Albert Vernon Smith, Jennifer E Huffman, Eva Albrecht, Catherine M Jackson, David M Evans, Gemma Cadby, Myriam Fornage, Ani Manichaikul, Lorna M Lopez, Toby Johnson, Melinda C Aldrich, Thor Aspelund, Inês Barroso, Harry Campbell, Patricia A Cassano, David J Couper, Gudny Eiriksdottir, Nora Franceschini, Melissa Garcia, Christian Gieger, Gauti Kjartan Gislason, Ivica Grkovic, Christopher J Hammond, Dana B Hancock, Tamara B Harris, Adaikalavan Ramasamy, Susan R Heckbert, Markku Heliövaara, Georg Homuth, Pirro G Hysi, Alan L James, Stipan Jankovic, Bonnie R Joubert, Stefan Karrasch, Norman Klopp, Beate Koch, Stephen B Kritchevsky, Lenore J Launer, Yongmei Liu, Laura R Loehr, Kurt Lohman, Ruth JF Loos, Thomas Lumley, Khalid A Al Balushi, Wei Q Ang, R Graham Barr, John Beilby, John D Blakey, Mladen Boban, Vesna Boraska, Jonas Brisman, John R Britton, Guy G Brusselle, Cyrus Cooper, Ivan Curjuric, Santosh Dahgam, Ian J Deary, Shah Ebrahim, Mark Eijgelsheim, Clyde Francks, Darya Gaysina, Raquel Granell, Xiangjun Gu, John L Hankinson, Rebecca Hardy, Sarah E Harris, John Henderson, Amanda Henry, Aroon D Hingorani, Albert Hofman, Patrick G Holt, Jennie Hui, Michael L Hunter, Medea Imboden, Karen A Jameson, Shona M Kerr, Ivana Kolcic, Florian Kronenberg, Jason Z Liu, Jonathan Marchini, Tricia McKeever, Andrew D Morris, Anna-Carin Olin, David Porteous, Dirkje S Postma, Stephen S Rich, Susan M Ring, Fernando Rivadeneira, Thierry Rochat, Avan Aihie Sayer, Ian Sayers, Peter D Sly, George Davey Smith, Akshay Sood, John M Starr, André G Uitterlinden, Judith M Vonk, S Goya Wannamethee, Peter H Whincup, Cisca Wijmenga, O Dale Williams, Andrew Wong, Massimo Mangino, Kristin D Marciante, Wendy L McArdle, Bernd Meibohm, Alanna C Morrison, Kari E North, Ernst Omenaas, Lyle J Palmer, Kirsi H Pietiläinen, Isabelle Pin, Ozren Polašek, Anneli Pouta, Bruce M Psaty, Anna-Liisa Hartikainen, Taina Rantanen, Samuli Ripatti, Jerome I Rotter, Igor Rudan, Alicja R Rudnicka, Holger Schulz, So-Youn Shin, Tim D Spector, Ida Surakka, Veronique Vitart, Henry Völzke, Nicholas J Wareham, Nicole M Warrington, H-Erich Wichmann, Sarah H Wild, Jemma B Wilk, Matthias Wjst, Alan F Wright, Lina Zgaga, Tatijana Zemunik, Craig E Pennell, Fredrik Nyberg, Diana Kuh, John W Holloway, H Marike Boezen, Debbie A Lawlor, Richard W Morris, Nicole Probst-Hensch, The International Lung Cancer Consortium, GIANT consortium, Jaakko Kaprio, James F Wilson, Caroline Hayward, Mika Kähönen, Joachim Heinrich, Arthur W Musk, Deborah L Jarvis, Sven Gläser, Marjo-Riitta Järvelin, Bruno H Ch Stricker, Paul Elliott, George T O'Connor, David P Strachan, Stephanie J London,Ian P Hall, Vilmundur Gudnason & Martin D Tobin

Nature Genetics: doi:10.1038/ng.941

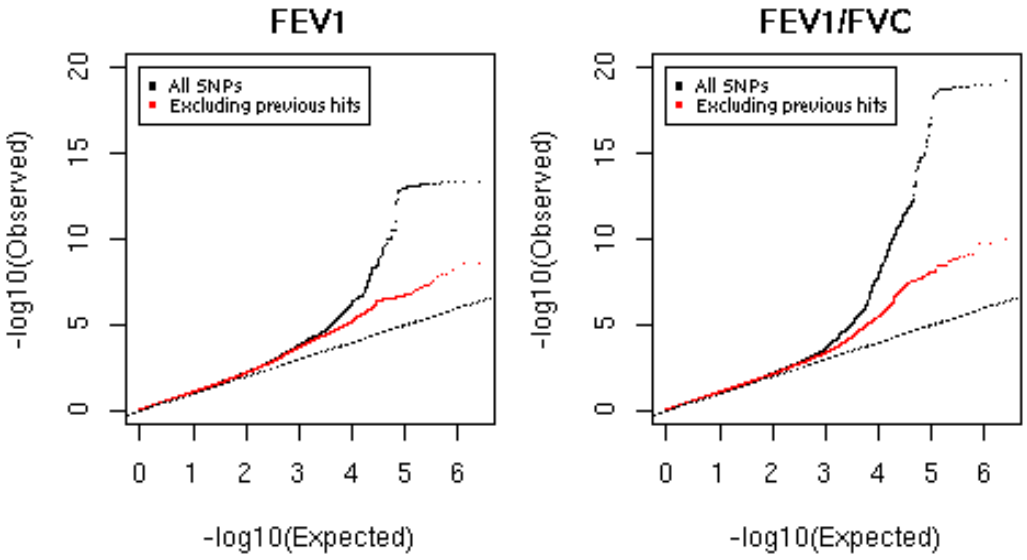
Table of Contents

Nature Genetics: doi:10.1038/ng.941

Supplementary Figures

Supplementary Figure 1 Association test statistics for FEV_1 and FEV_1/FVC.

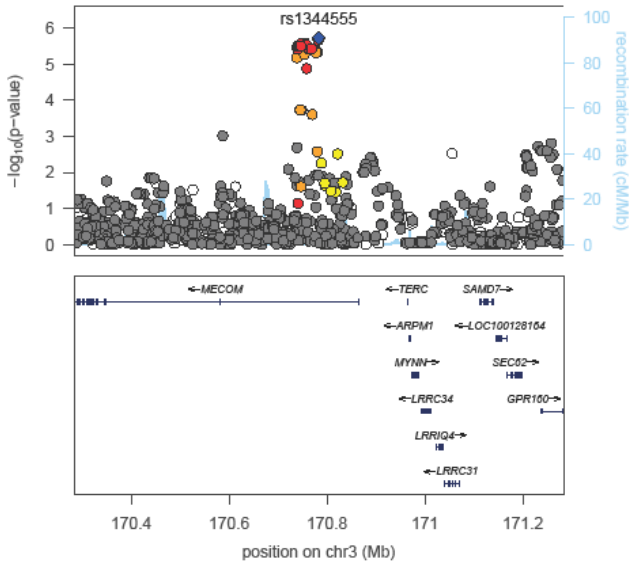
A) Quantile-Quantile (QQ) plots show $-\log_{10}(P)$ of observed genome-wide association results against expected association results for FEV_1 and FEV_1/FVC. λ_{GC} before genomic control was 1.12 for FEV_1 and 1.09 for FEV_1/FVC. The QQ plot for all SNPs is shown in black. The QQ plot for remaining SNPs after exclusion of SNPs from loci previously reported by the SpiroMeta and CHARGE consortia[1,2] is shown in red; SNPs within 500kb of the lead reported SNP were excluded. The previously reported loci were *TNS1*, *FAM13A*, *GSTCD*/*NPNT*, *HHIP*, *HTR4*, *ADAM19*, *AGER*, *GPR126*, *PTCH1*, and *TSHD4.*

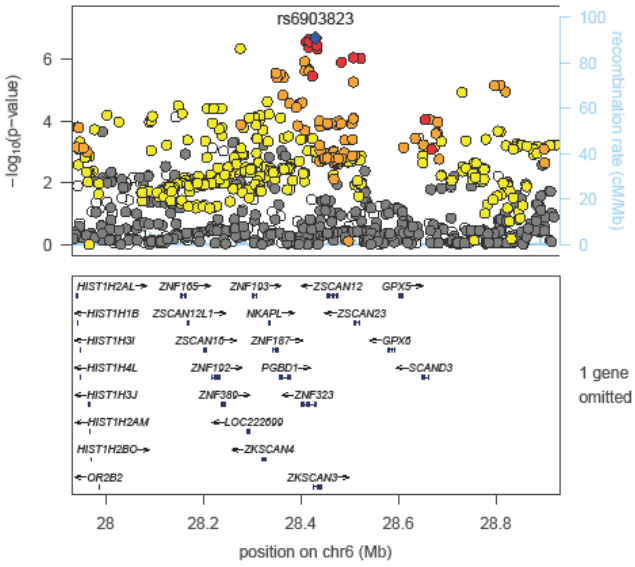


Nature Genetics: doi:10.1038/ng.941

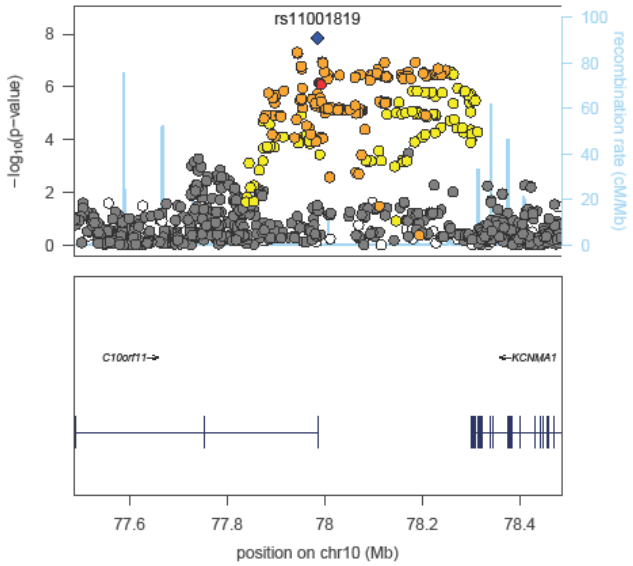
B) Regional association plots of 16 novel lung function-associated loci. Statistical significance of each SNP on the $-\log_{10}(P)$ scale as a function of chromosome position (NCBI Build 36) in the meta-analysis of Stage 1 data alone. The sentinel SNP at each locus is shown in blue with the correlations (r^2) of surrounding SNPs to the sentinel indicated by colour (red: r^2>0.8, orange: r^2>0.5, yellow: r^2>0.2, grey: r^2<0.2, white: r^2 unknown). The fine scale recombination rate is shown in blue.

FEV_1 only



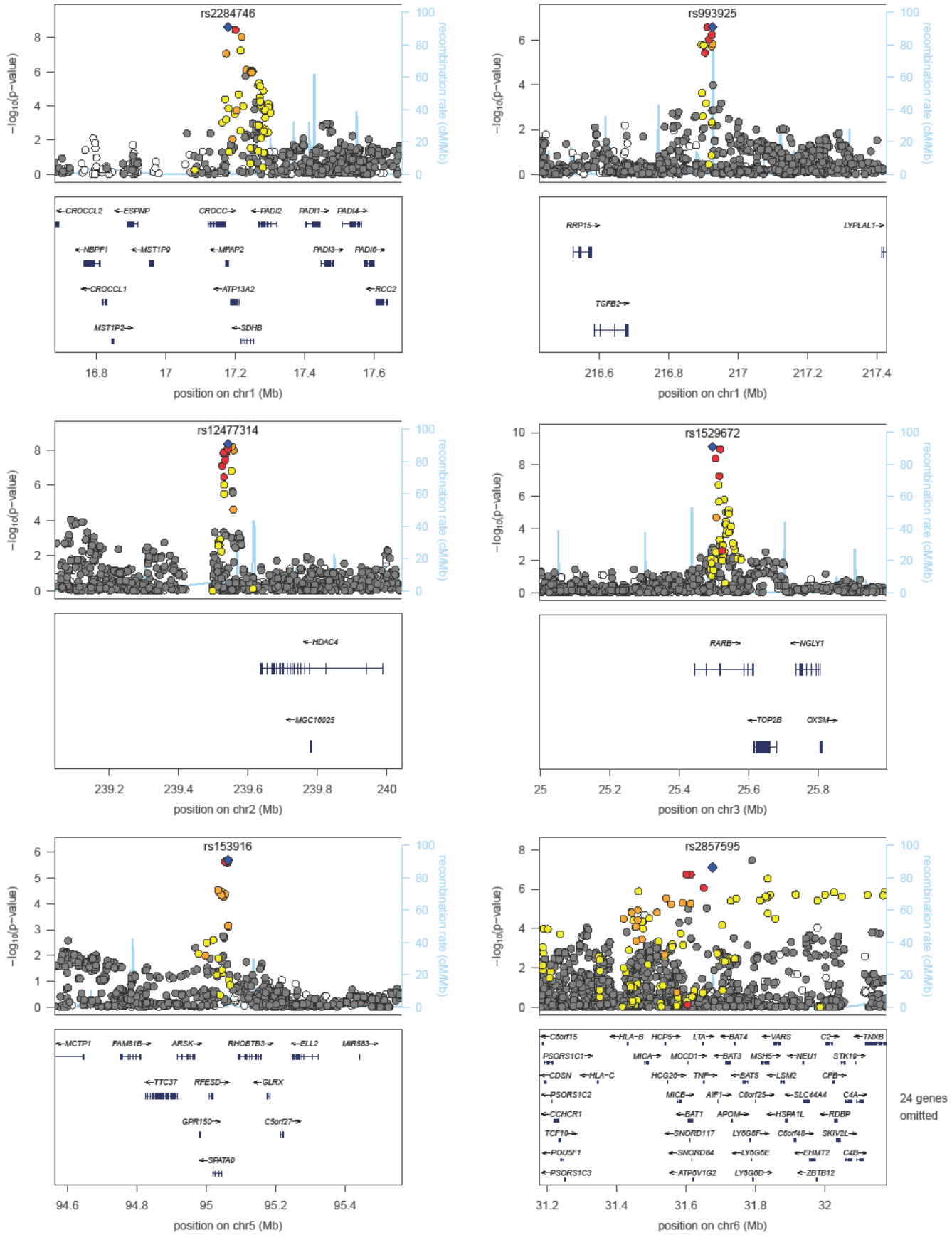


F

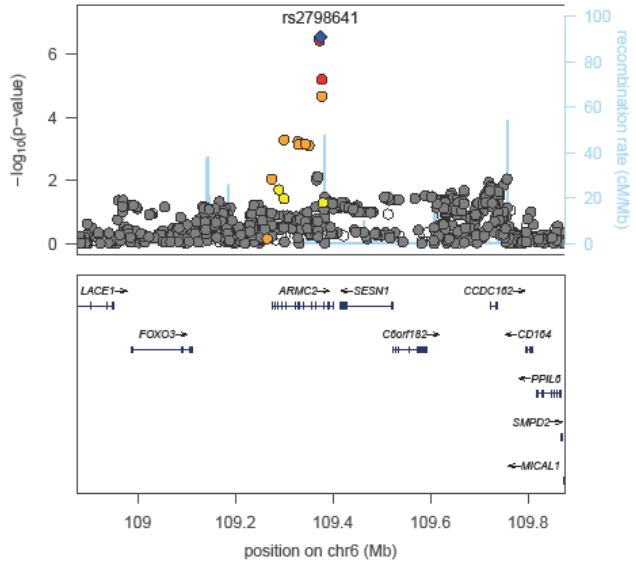


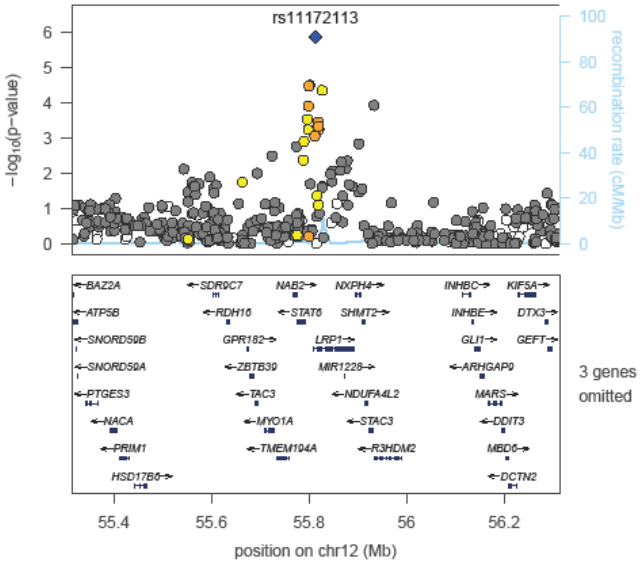
Nature Genetics: doi:10.1038/ng.941

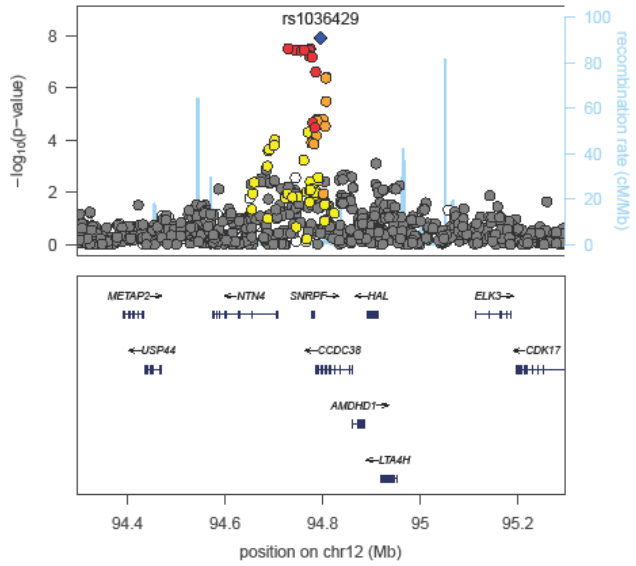
FEV$_1$/FVC only

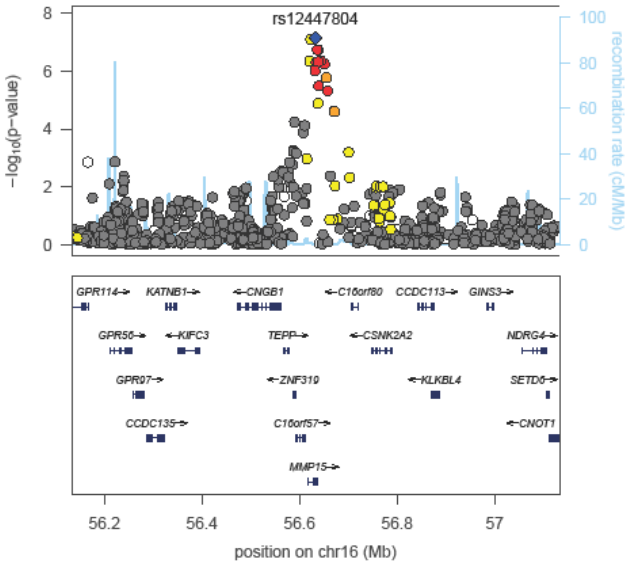


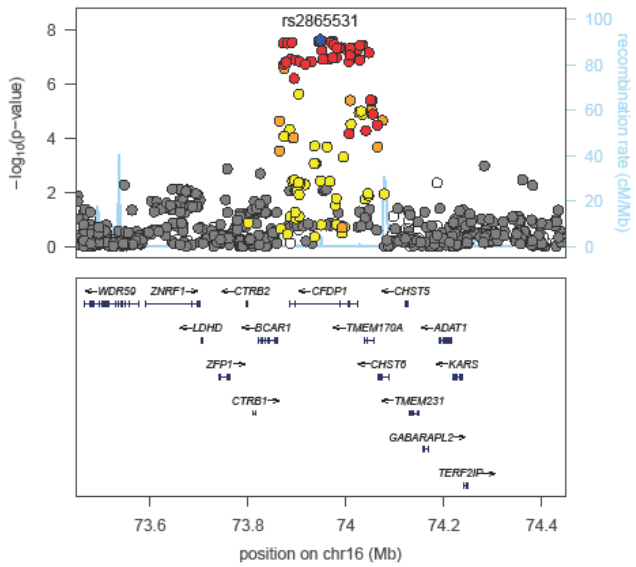
Nature Genetics: doi:10.1038/ng.941

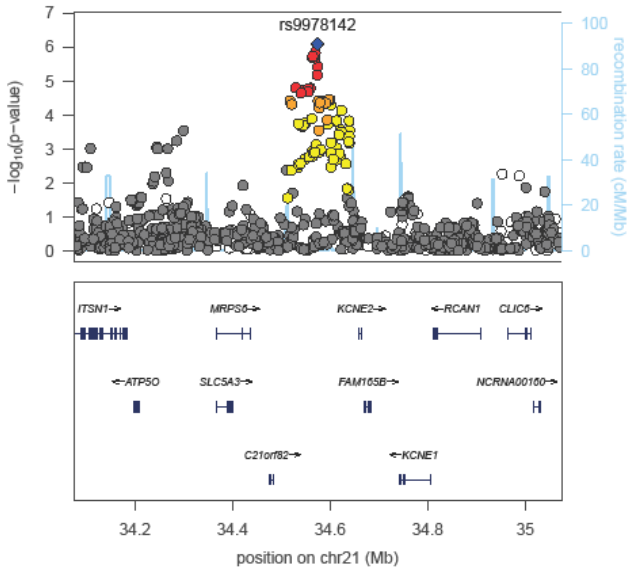








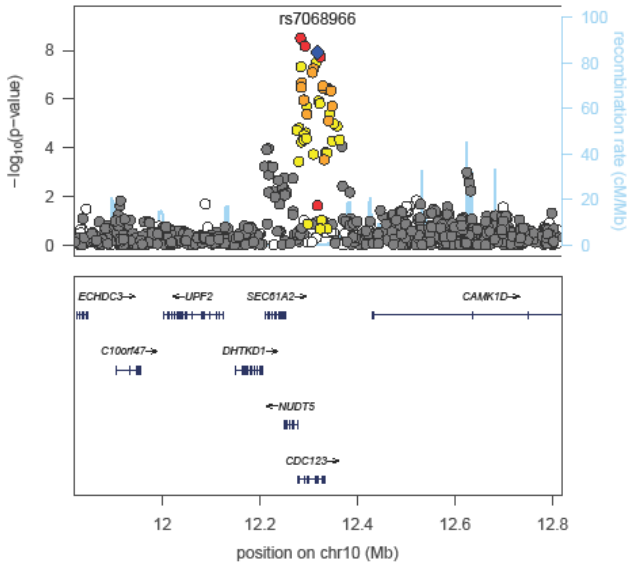




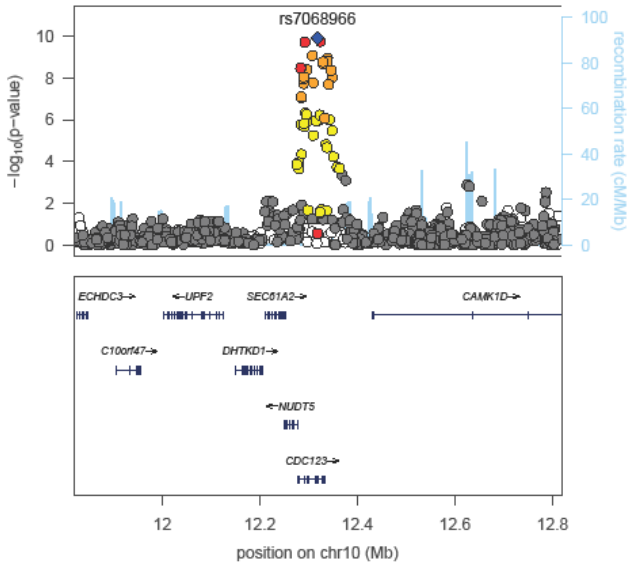
Nature Genetics: doi:10.1038/ng.941

Both FEV_1 and FEV_1/FVC

FEV_1



FEV_1/FVC



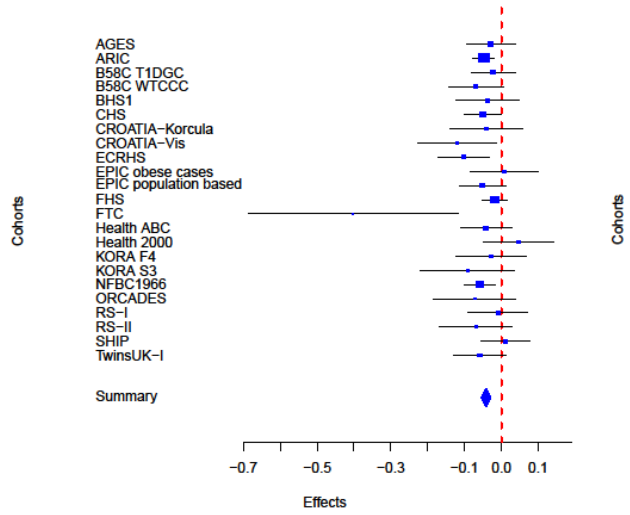
Nature Genetics: doi:10.1038/ng.941

C) Forest plots for the 16 loci associated with lung function for Stage 1 and Stage 2 separately. Each of the SNPs included in the figure showed genome-wide significant association ($P<5\times10^{-8}$) with either FEV_1 or FEV_1/FVC in the data from Stages 1 and 2. For each SNP there is a plot for the meta-analysis of the Stage 1 data and another for the meta-analysis of the Stage 2 data. 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 is shown at the bottom of each graph.

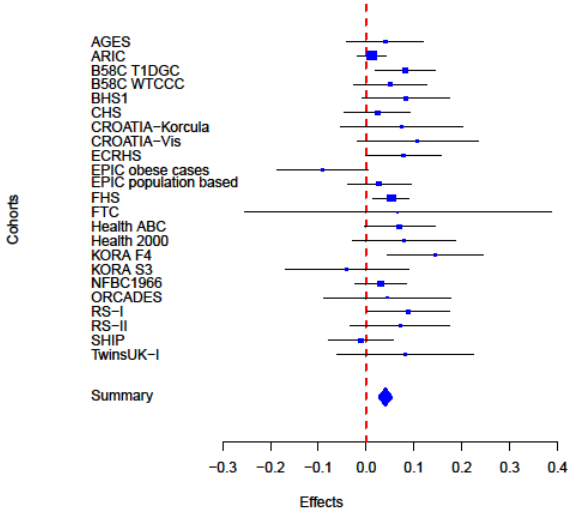
Nature Genetics: doi:10.1038/ng.941

Stage 1

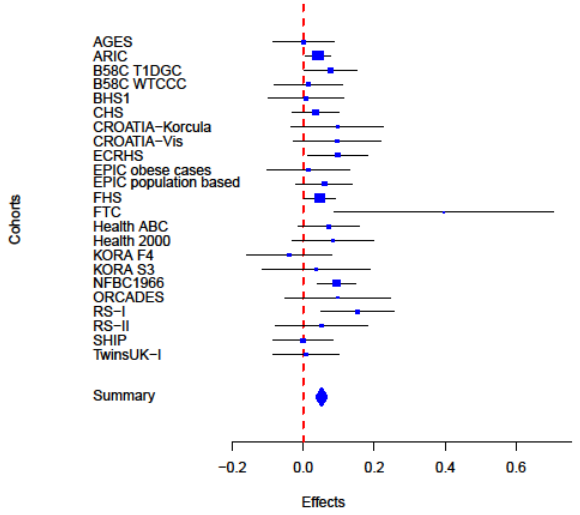
FEV1/FVC rs2284746



FEV1/FVC rs993925

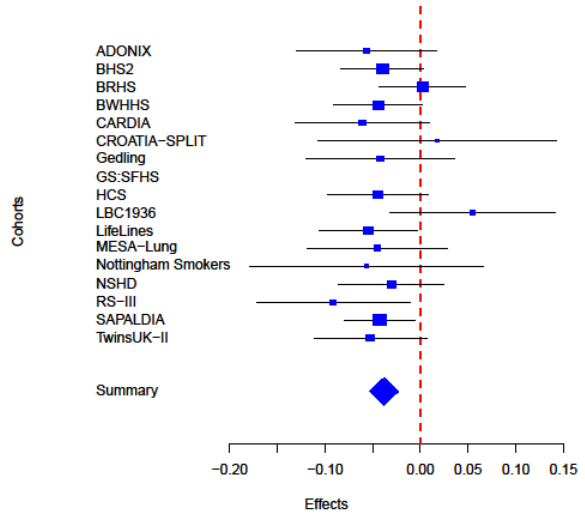


FEV1/FVC rs12477314

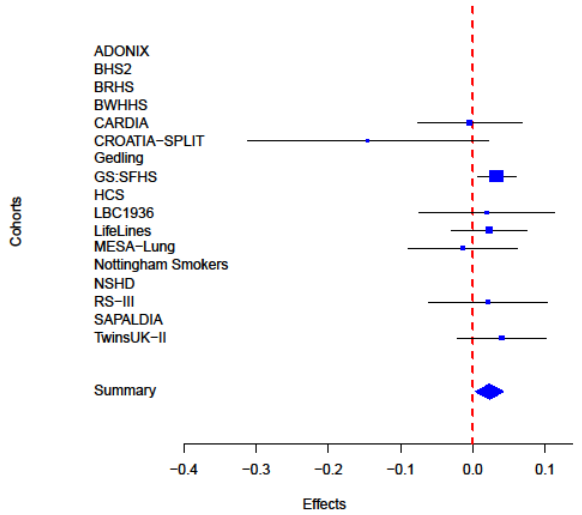


Stage 2

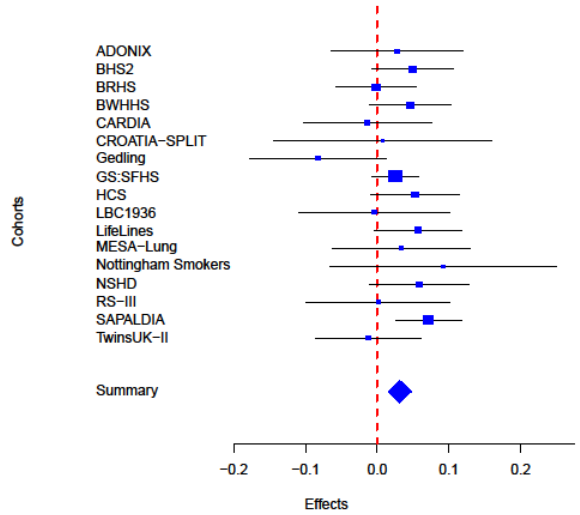
FEV1/FVC rs2284746



FEV1/FVC rs993925

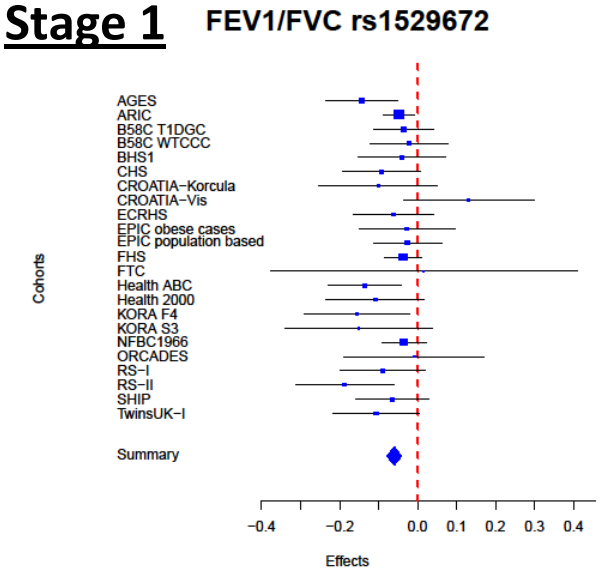


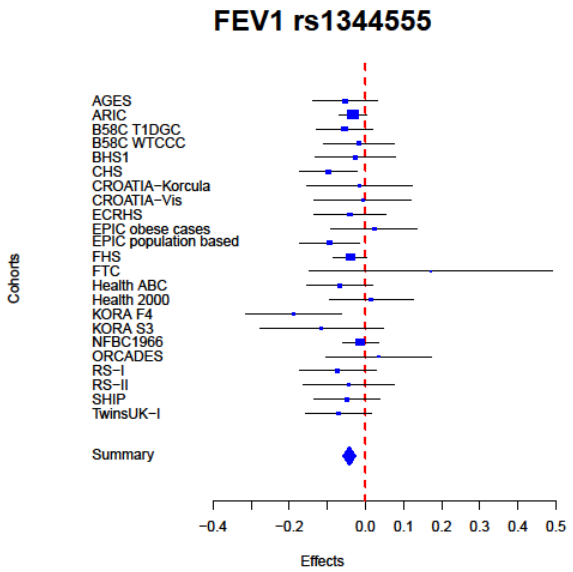
FEV1/FVC rs12477314

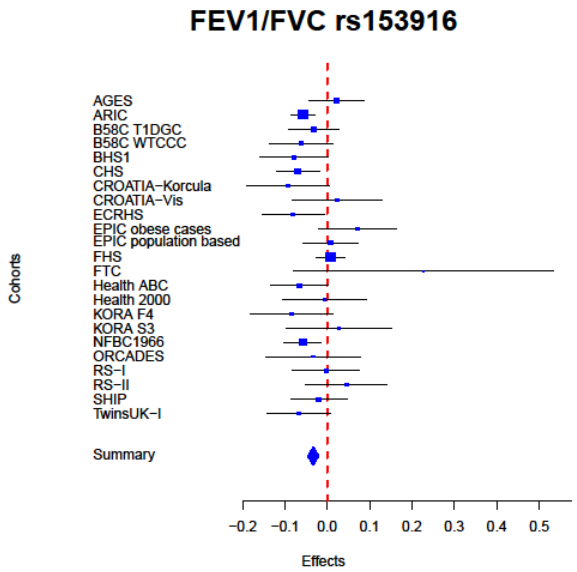


Nature Genetics: doi:10.1038/ng.941

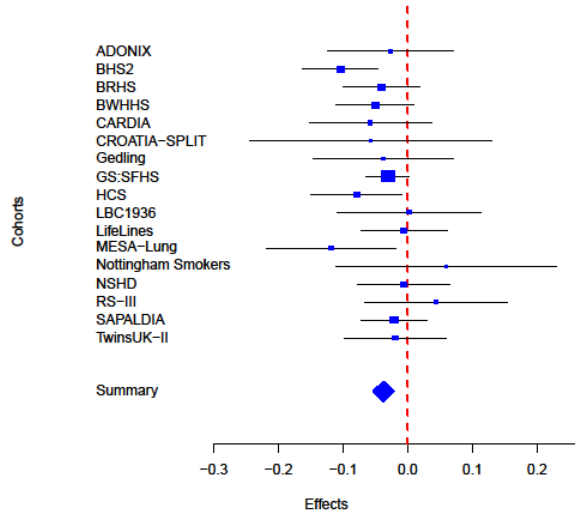
Stage 1

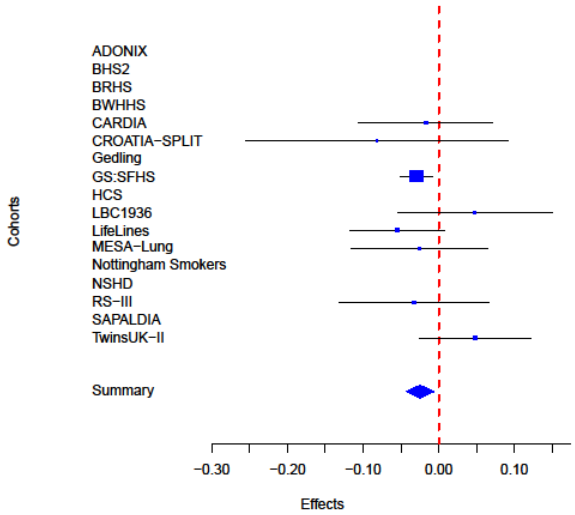


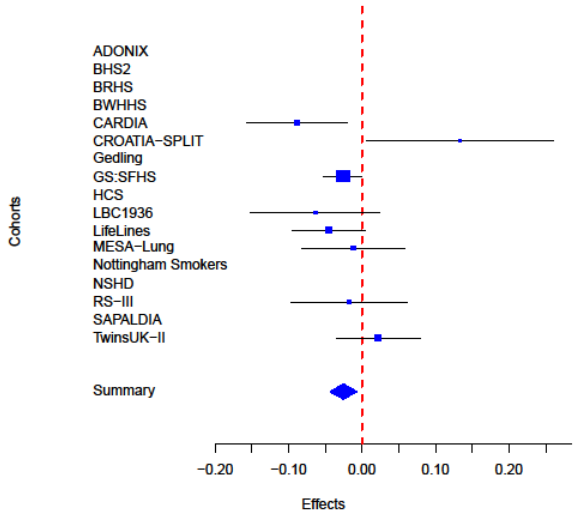




Stage 2



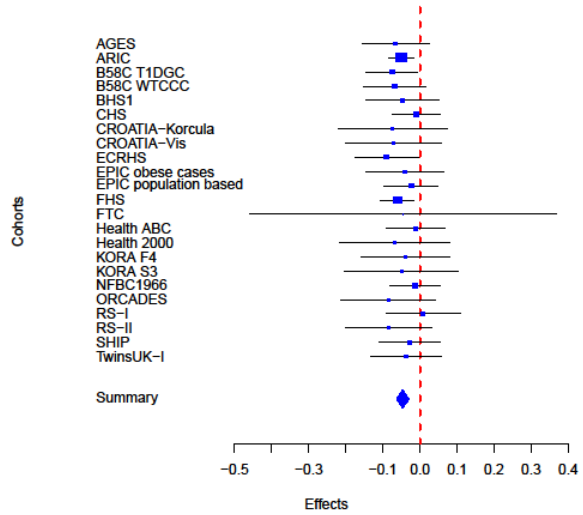




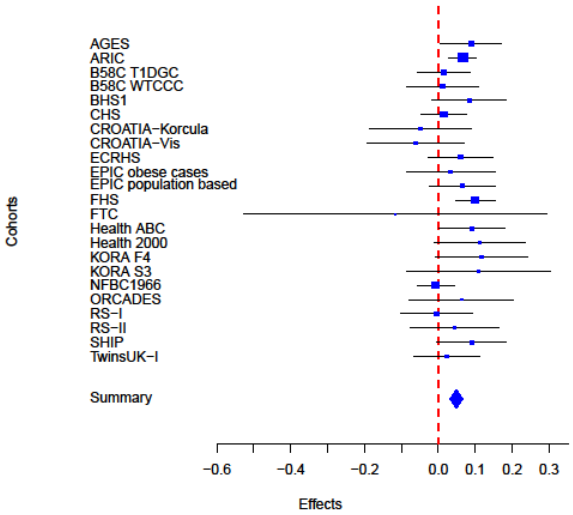
Nature Genetics: doi:10.1038/ng.941

Stage 1

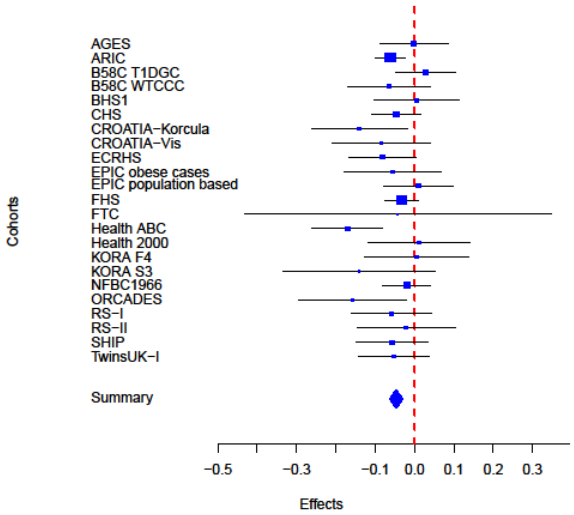
FEV1 rs6903823



FEV1/FVC rs2857595

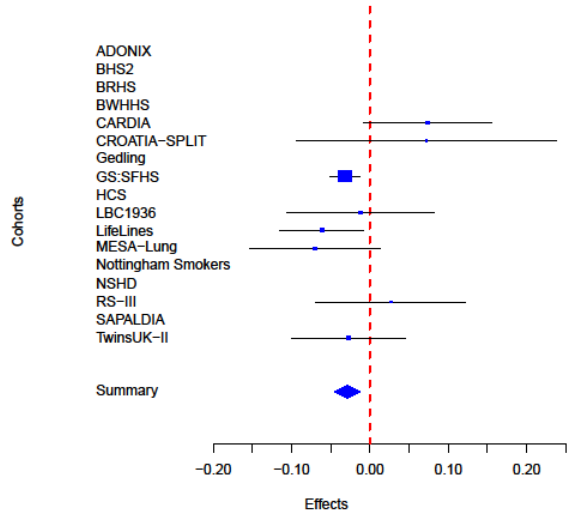


FEV1/FVC rs2798641

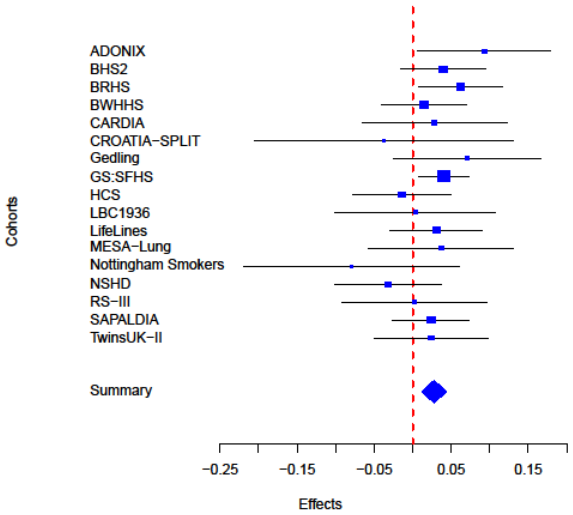


Stage 2

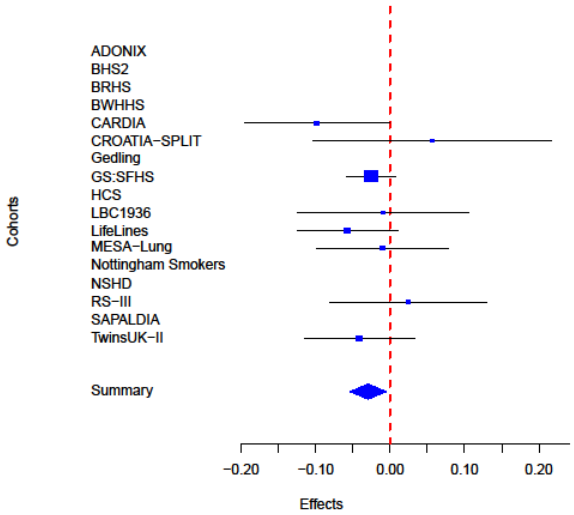
FEV1 rs6903823



FEV1/FVC rs2857595

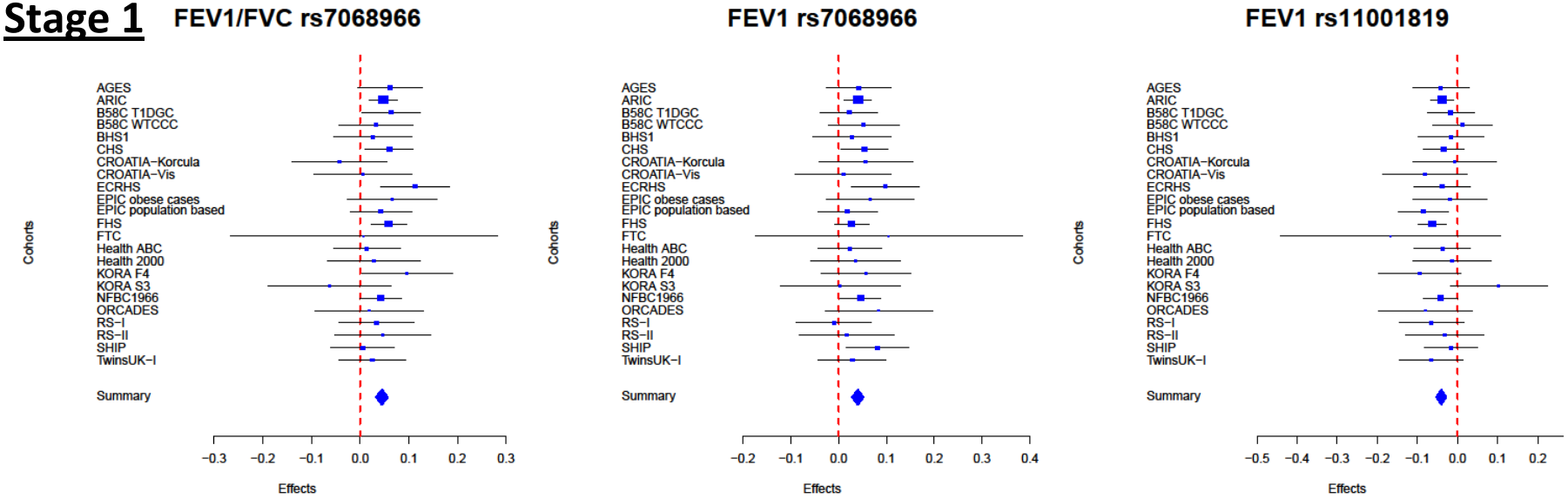


FEV1/FVC rs2798641

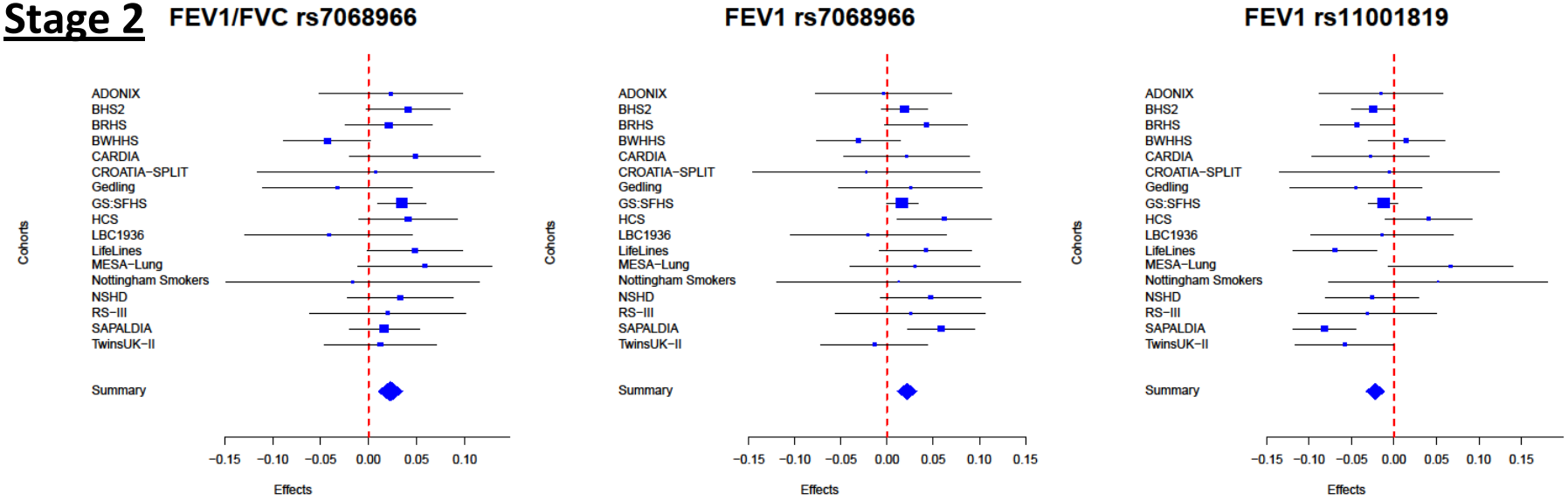


Nature Genetics: doi:10.1038/ng.941

Stage 1



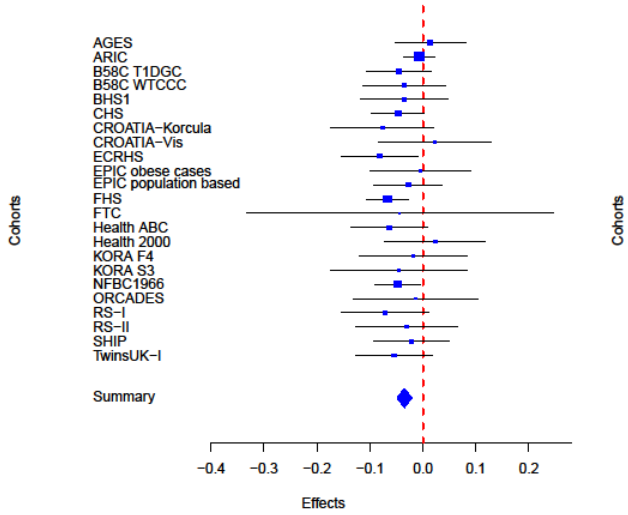
Stage 2



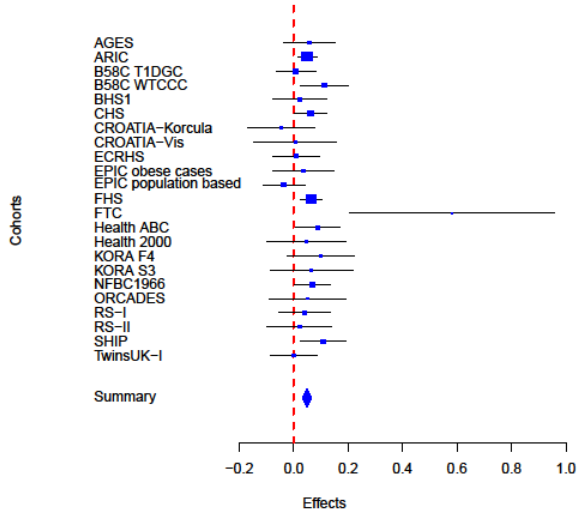
Nature Genetics: doi:10.1038/ng.941

Stage 1

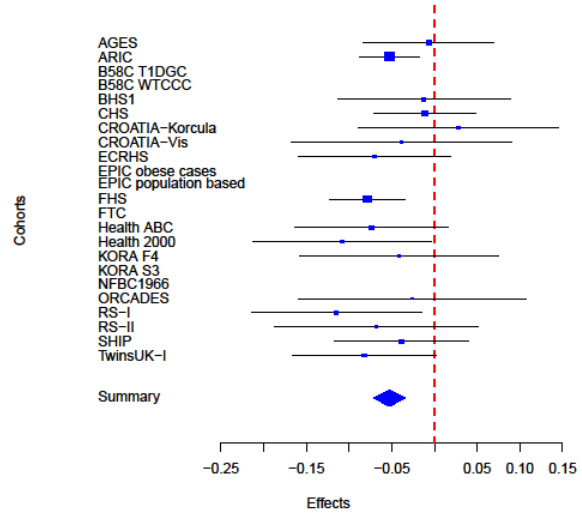
FEV1/FVC rs11172113



FEV1/FVC rs1036429

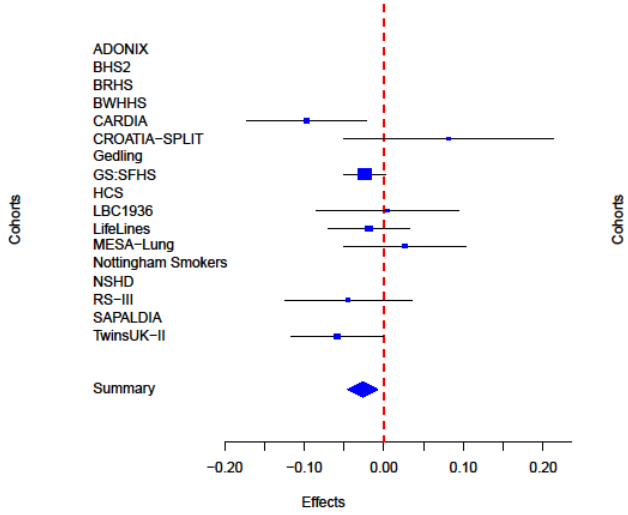


FEV1/FVC rs12447804

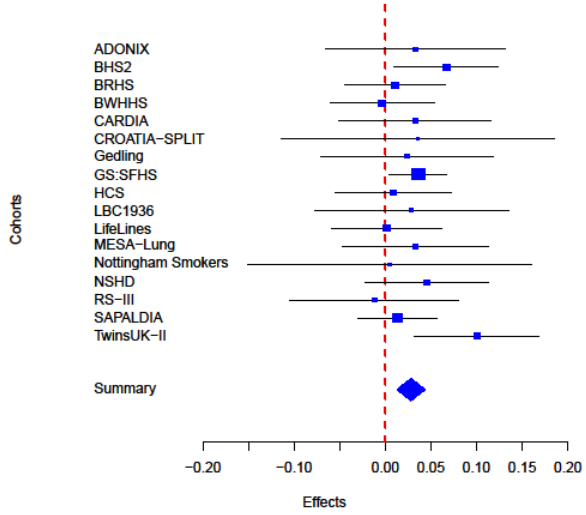


Stage 2

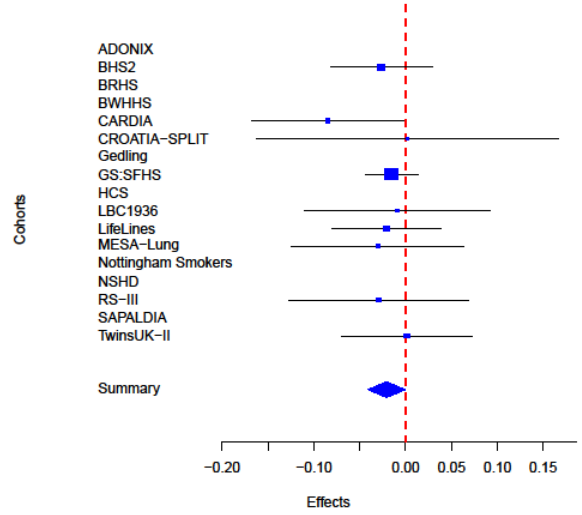
FEV1/FVC rs11172113



FEV1/FVC rs1036429

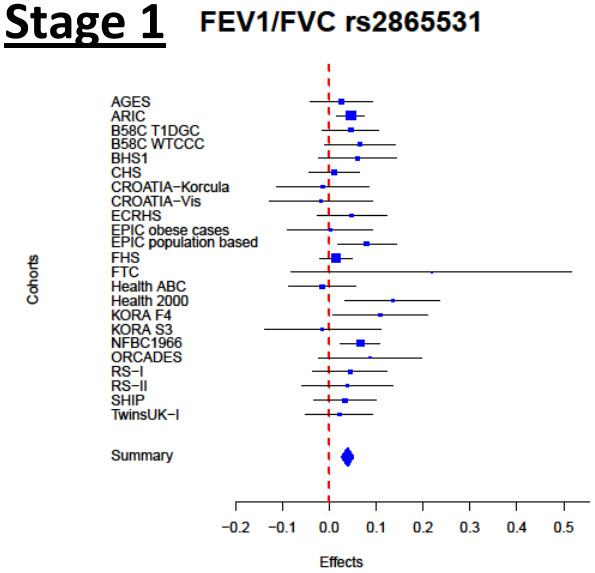


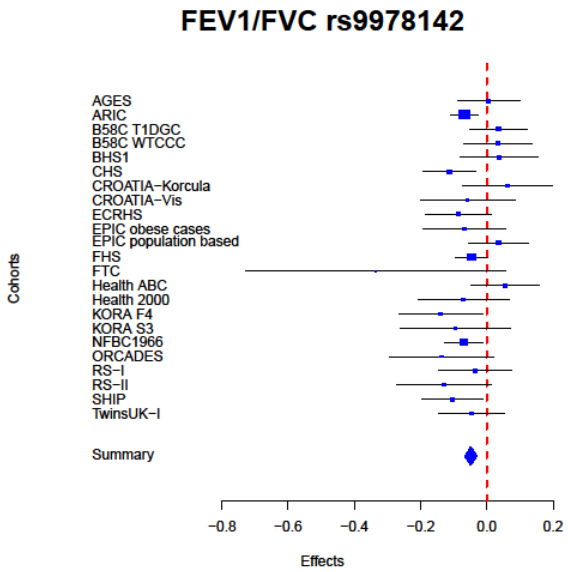
FEV1/FVC rs12447804



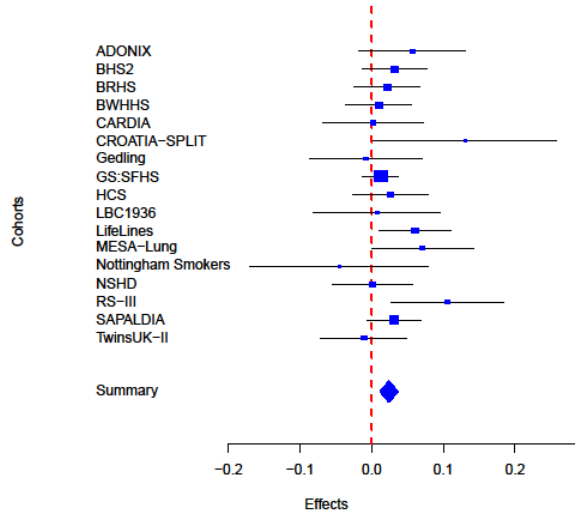
Nature Genetics: doi:10.1038/ng.941

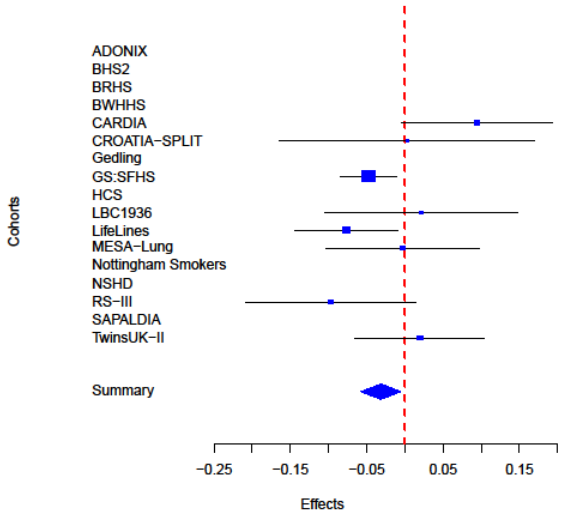
Stage 1





Stage 2





Nature Genetics: doi:10.1038/ng.941

Supplementary Tables

Supplementary Table 1 Sample population characteristics and genotyping platform details for each study.

A) Sample population characteristics for each study. Characteristics are shown for studies analyzed in Stage 1 (GWAS meta-analysis), Stage 2 (follow-up). Stage 1 studies: AGES, Age, Gene/Environment Susceptibility; ARIC, Atherosclerosis Risk in Communities; B58C-T1DGC, British 1958 Birth Cohort–Type 1 Diabetes Genetics Consortium; B58C-WTCCC, British 1958 Birth Cohort–Wellcome Trust Case Control Consortium; BHS1, Busselton Health Study 1; CHS, Cardiovascular Health Study; the CROATIA- Korcula study; the CROATIA-Vis study; ECRHS, the European Community Respiratory Health Survey; EPIC obese cases, European Prospective Investigation into Cancer and Nutrition, Obese Cases; EPIC population based, European Prospective Investigation into Cancer and Nutrition Cohort; FHS, Framingham Heart Study; FTC, Finnish Twin Cohort incorporating FinnTwin16 and FITSA; H2000, Finnish Health 2000 survey; Health ABC, Health, Aging, and Body Composition; KORA F4, Cooperative Health Research in the Region of Augsburg; KORA S3, Cooperative Health Research in the Region of Augsburg; NFBC1966, Northern Finland Birth Cohort of 1966; ORCADES, Orkney Complex Disease Study; RS-I and RS-II, Rotterdam Study; SHIP, Study of Health in Pomerania; the TwinsUK-I study. Stage 2 studies: ADONIX, Adult-Onset Asthma and Nitric Oxide; BHS2, Busselton Health Study 2; BRHS, British Regional Heart Study; BWHHS, British Women's Heart and Health Study; CARDIA, Coronary Artery Risk Development in Young Adults; the CROATIA-Split study; the Gedling study; GS:SFHS, Generation Scotland: Scottish Family Health Study; LBC1936, Lothian Birth Cohort 1936; the LifeLines study; MESA-Lung, Multi-Ethnic Study of Atherosclerosis; HCS, Hertfordshire Cohort Study; the Nottingham Smokers study; NSHD, Medical Research Council National Survey of Health and Development (also known as the British 1946 Birth Cohort); RS-III, Rotterdam Study; SAPALDIA, Swiss study on Air Pollution and Lung Disease in adults; the TwinsUK-II study. * Total number of individuals with both genotype and phenotype data and which were included in the analyses described in this paper.

Nature Genetics: doi:10.1038/ng.941

Study	*N* Total	*N* male	*N* female	Age range (y) at FEV_1/FVC measurement	Mean age, y (s.d.)	Mean FEV_1, L (s.d.)	Mean FVC, L (s.d.)	Mean FEV_1/FVC (s.d.)	*N* never-smokers	*N* ever-smokers	Genomic inflation factor (λ) FEV_1		Genomic inflation factor (λ) FEV_1/FVC	
Stage 1: GWAS											Ever-smokers	Never-smokers	Ever-smokers	Never-smokers
AGES	1689	686	1003	66-95	76.19 (5.63)	2.13(0.69)	2.86(0.85)	0.74(0.11)	803	886	1.009	1.012	1.003	1.003
ARIC	9078	4279	4799	44-66	54.27 (5.70)	2.94 (0.78)	3.99 (0.98)	0.74 (0.08)	3620	5458	1.034	1.007	1.019	1.019
B58C T1DGC	2343	1131	1212	44–45	44.5 (0)	3.31 (0.78)	4.19 (0.96)	0.79 (0.08)	692	1651	1.009	0.999	1.023	1.009
B58C WTCCC	1372	691	681	44–45	44.5 (0)	2.93 (0.75)	4.18 (0.96)	0.79 (0.08)	394	978	0.999	0.996	1.007	0.99
BHS1	1168	455	713	17-91	52.98 (17.07)	2.81 (0.97)	3.68 (0.11)	0.76 (0.09)	653	515	1.02	1.034	1.02	1.015
CHS	3140	1226	1914	65-95	72.3 (5.4)	2.12 (0.66)	3.00 (0.87)	0.71 (0.11)	1543	1597	1.035	1.021	1.02	1.033
CROATIA-Korcula	825	300	525	18–90	55.5 (13.5)	2.84 (0.81)	3.37 (0.93)	0.84 (0.09)	397	428	1.039	1.014	0.999	1.041
CROATIA-Vis	769	323	446	18–88	56.3 (15.3)	3.39 (1.22)	4.38 (1.43)	0.77 (0.09)	328	441	1.019	1.002	1.066	1.027
ECRHS	1594	784	810	19-48	33.90 (7.17)	3.78(0.82)	4.59 (1.03)	0.83 (0.07)	699	895	1.018	1.014	1.005	1.024
EPIC obese cases	1104	476	628	39–76	59.1 (8.8)	2.35 (0.69)	2.84 (0.87)	0.82 (0.17)	489	615	1.005	1.02	1.02	1.014
EPIC population based	2336	1100	1236	39–77	59.2 (9.0)	2.50 (0.72)	3.04 (0.90)	0.85 (0.16)	1061	1275	1.013	1.008	1.002	1.018
FHS	7911	3650	4261	19-92	52.2 (14.6)	3.03 (0.94)	4.02 (1.14)	0.75 (0.08)	3556	4355	1.02	1.032	1.007	1.026
FTC	134	13	121	23–76	57.4 (19.3)	2.69 (0.94)	2.93 (0.61)	0.79 (0.09)	104	30	1.054	1.011	1.005	1.003
Health ABC	1472	786	686	70-79	73.7 (2.8)	2.31 (0.66)	3.11 (0.81)	0.74 (0.08)	641	831	0.997	1.001	0.998	1.012
Health 2000	821	394	427	30-75	50.47(10.91)	3.29 (0.90)	4.16 (1.07)	0.79 (0.07)	249	572	1.001	1.023	1.007	1
KORA F4	904	426	478	42-61	53.82(4.39)	3.25 (0.79)	4.20 (0.97)	0.77 (0.06)	344	560	1.051	1.013	1.032	1.017
KORA S3	555	261	294	29–73	47.6 (9.0)	3.43 (0.78)	4.18 (0.99)	0.83 (0.07)	266	289	1.014	1.019	1.012	1.029

Nature Genetics: doi:10.1038/ng.941

NFBC1966	4556	2182	2374	31–31	31.0 (0)	3.96 (0.79)	4.73 (0.99)	0.84 (0.06)	1648	2908	1.022	1.011	1.023	1.004
ORCADES	692	322	370	19–93	54.9 (15.3)	2.88 (0.84)	3.58 (0.98)	0.80 (0.09)	404	288	1.014	1.046	1.015	1.051
RS-I	1224	556	668	65-97	74.5 (5.6)	2.31 (0.73)	3.16 (0.92)	0.73 (0.08)	361	863	1.029	1.023	1.026	1.017
RS-II	852	381	471	58-88	67.2 (6.3)	2.71 (0.78)	3.61 (1.08)	0.76 (0.09)	287	565	1.027	1.006	1.009	1.016
SHIP	1777	870	907	25–85	52.3 (13.7)	3.28 (0.89)	3.87 (1.03)	0.87 (0.06)	773	1004	1	0.996	1.016	0.991
TwinsUK-I	1885	0	1,885	18–79	48.4 (12.2)	2.73 (0.56)	3.40 (0.61)	0.80 (0.08)	943	942	0.998	1.012	1.009	1.005
Stage 1 sample size	48201													
Stage 2: follow-up														
ADONIX	1410	660	750	25-75	49.08 (13.54)	3.34 (0.86)	4.24(1.02)	0.79 (0.07)	792	618				
BHS2	3038	1368	1670	18-97	50.0 (16.7)	3.05 (0.951)	3.97 (1.15)	0.78 (0.08)	1633	1405				
BRHS	3862	3862	0	58-80	68.72 (5.48)	2.58 (0.69)	3.38 (0.84)	0.77 (0.17)	1121	2741				
BWHHS	3635	0	3635	59-80	68.83 (5.49)	1.98 (0.52)	2.82 (0.76)	0.71 (0.09)	2055	1580				
CARDIA	1626	768	858	17-32	25.6 (3.33)	3.68 (0.81)	4.70 (1.00)	0.82 (0.06)	932	694				
CROATIA-SPLIT	491	209	282	18-85	49.07 (14.60)	3.19 (0.91)	3.80 (1.06)	0.84 (0.08)	239	252				
Gedling	1266	633	633	27-80	56.14 (12.29)	2.85 (0.85)	3.68 (1.01)	0.77 (0.07)	634	632				
GS:SFHS	10399	4304	6095	18-93	46.37 (14.61)	3.11 (0.87)	4.05 (1.02)	0.77 (0.09)	5674	4725				
HCS	2848	1509	1339	59-73	66.14 (2.84)	2.44 (0.68)	3.42 (0.92)	0.72 (0.09)	1318	1530				
LBC1936	991	501	490	67-71	69.55 (0.84)	2.38 (0.67)	3.04 (0.87)	0.79 (0.10)	437	554				
LifeLines	3078	1232	1846	21-88	54.94 (9.75)	3.15 (0.81)	4.21 (1,01)	0.75 (0.08)	1075	2003				
MESA-Lung	1469	737	732	48-90	66.1 (9.7)	2.57 (0.76)	3.44 (0.99)	0.73 (0.09)	636	833				

Nature Genetics: doi:10.1038/ng.941

Nottingham Smokers	521	236	285	36-89	59.60 (10.48)	1.10 (0.95)	3.02 (1.05)	0.64 (0.16)	0	521				
NSHD	2511	1258	1253	53	53	2.79 (0.70)	3.50 (0.90)	0.80 (0.09)	1045	1466				
RS-III	1247	549	698	46-89	56.59 (5.58)	3.15 (0.85)	4.06 (1.14)	0.78 (0.09)	425	822				
SAPALDIA	5646	2753	2893	18-62	42.0 (11.4)	3.58 (0.84)	4.53 (1.04)	0.79 (0.07)	2653	2993				
TwinsUK-II	2373	0	2373	17-85	53.5(14.3)	2.62(0.61)	3.27(0.65)	0.80 (0.08)	1230	1143				
Stage 2 sample size	46411													
Total sample size	94612													

Nature Genetics: doi:10.1038/ng.941

B) Stage1 study, genotyping, imputation and genotype-phenotype data. Genotyping platforms, filters applied to SNPs and individuals (if any) before imputation, imputation software and genotype-phenotype association software are given. Studies: Age, Gene/Environment Susceptibility (AGES), Atherosclerosis Risk in Communities (ARIC), British 1958 Birth Cohort–Type 1 Diabetes Genetics Consortium (B58C T1DGC), British 1958 70Birth Cohort–Wellcome Trust Case Control Consortium (B58C WTCCC), Busselton Health Study 1(BHS1), Cardiovascular Health Study (CHS), CROATIA- Korcula, CROATIA-Vis, European Community Respiratory Health Survey (ECRHS), European Prospective Investigation into Cancer and Nutrition-obese Cases (EPIC obese cases), European Prospective Investigation into Cancer and Nutrition-cohort (EPIC population based), Framingham Heart Study (FHS), Finnish Twin Cohort incorporating FinnTwin16 and FITSA (FTC), Finnish Health 2000 survey (Health 2000), Health, Aging, and Body Composition (Health ABC), Cooperative Health Research in the Region of Augsburg (KORA F4), Cooperative Health Research in the Region of Augsburg (KORA S3), Northern Finland Birth Cohort of 1966 (NFBC1966), Orkney Complex Disease Study (ORCADES), Rotterdam Studies (RS-I and RS-II), Study of Health in Pomerania (SHIP) and TwinsUK-I. Abbreviations: GWAS= Genome-Wide Association Study, imp'n=imputation, HWE= Hardy Weinberg Equilibrium, MAF= minor allele frequency.
MACH and IMPUTE are two software implementations that share similar underlying population genetic models[3], and BIMBAM has been shown to perform similarly to MACH and IMPUTE in contrast with other imputation methods[4,5].

Nature Genetics: doi:10.1038/ng.941

Study name	GWAS platform	Calling algorithm	Individual call rate filter applied (before imp'n)	SNP call rate filter applied before imp'n	SNP HWE filter applied (before imp'n)	SNP MAF filter applied (before imp'n)	Other filter	No of SNPs after filtering (before imp'n)	Imp'n software and version	NCBI; HapMap CEU version for imp'n	Genotype-phenotype association software and version
AGES	Illumina Hu370CNV	BeadStudio	0.97	0.90	1e10-6	0.01	remove AT/GC SNPs	208340	MACH 1.0.16	36;21a	ProbABEL 0.1
ARIC	Affymetrix 6.0	Birdseed	0.95	0.95	1e10-6	0.01	no chromosomal location	669,450	MACH 1.0.16	36;22	ProbABEL 0.1-3
B58C T1DGC	Illumina 550K	ILLUMINUS	0.98	No	No	No	No	520010	MACH 1.0.13	35;21	ProbABEL 0.0-5b
B58C WTCCC	Affymetrix 500K	CHIAMO	0.98	No	No	No	No	490033	IMPUTE 0.2.0	35;21	SNPTEST 1.1.3
BHS1	Illumina 610-Quad	BeadStudio	0.97	0.99 for SNPS with MAF<1%, 0.95 for all other SNPs	5.7e10-7	0.01	No	549294	MACH 1.0.16	36	Mach2Qtl 1.0.8
CHS	Illumina 370 CNV	BeadStudio	0.95	0.97	1e10-5	heterozygote frequency >0	reproducibility errors<2	306,655	BimBam 0.99	36	R
CROATIA-Korcula	Illumina HumanHap 370cnv	Beadstudio	0.98 (for SNP of call rate >=0.98,MAF>=0.02,HWE>=E-10)	0.98	1e10-6	0.01	No	307728	MACH 1.0.15	36;22	GenABEL 1.4.2 , ProbABEL
CROATIA-Vis	Illumina HumanHap 300 v1	Beadstudio	0.97 (for SNP of call rate >=0.98,MAF>=0.02,HWE>=E-10)	0.98	1e10-6	0.01	No	305068	MACH 1.0.15	36;22	GenABEL 1.4.2, ProbABEL

Nature Genetics: doi:10.1038/ng.941

ECRHS (population based sample from first survey)	Illumina Quad 610k	GenCall	None	None	None	None	None	582892	MACH 1.0	36;22	ProbABEL 0.0-9
EPIC obese cases	Affymetrix 500K	BRLMM	0.94	0.90	1e10-6	0.01	No	397438	IMPUTE 0.3.1	35;21	SNPTEST 1.1.5
EPIC population-based	Affymetrix 500K	BRLMM	0.94	0.90	1e10-6	0.01	No	397438	IMPUTE 0.3.1	35;21	SNPTEST 1.1.5
FHS	Affy 500K + 50K Gene focused	Bayesian robust linear modeling using Mahalanobis distance (BRLMM)	0.97	0.97	1e10-6	0.01	MISHAP p<10-9, mendelian errors>100	378163	MACH 1.0.15	36;22	GWAF
FTC	Illumina 317K	BeadStudio	0.95	0.90	1e10-5	0.01	No	315987	MACH 1.0.16	36;22	PLINK 1.06
Health 2000	Illumina 610K	Illuminus	0.95	0.95	1e10-6	0.01	MDS-plot outliers removed (non-European ancestry)	555388	MACH 1.0	36;22	ProbABEL
Health ABC	Illumina Human 1M-Duo	BeadStudio 3.3.7	0.97	0.95	1e10-6	0.01	No sex mismatch, and cryptic relatedness	914,263	MACH 1.0.16.a	36;22	R version 2.9.2
KORA F4	Affymetrix 6.0	Birdseed2	0.93	No	No	No	No	909622	IMPUTE 0.4.2	36;22	SNPTEST 1.1.5
KORA S3	Affymetrix 500K	BRLMM	0.93	No	No	No	No	490033	MACH 1.0.9	35;21	MACH2QTL 1.0.4
NFBC1966	Illumina HumanCNV370-	Beadstudio	none	0.95	1e10-4	0.01	No	328007	IMPUTE v1.0	35; 21	SNPTEST 1.1.5

Nature Genetics: doi:10.1038/ng.941

	Duo										
ORCADES	Illumina HumanHap 300 v2	Beadstudio	0.98 (for SNP of call rate >=0.98,MAF>=0.02,HWE>=E-10)	0.98	1e10-6	0.01	No	306207	MACH 1.0.15	36;22	GenABEL 1.4.2, ProbABEL
RS-I	Illumina HapMap 550K	BeadStudio	0.98	0.98	1e10-6	0.01	excess autosomal heterozygosity, sex mismatch or outlying identity-by-state clustering estimates	512,349	MACH 1.0.15	36;22	MACH2QTL as implemented in GRIMP
RS-II	Illumina 550K + 610 Quad	GenomeStudio	0.98	0.98	1e10-6	0.01	excess autosomal heterozygosity, sex mismatch or outlying identity-by-state clustering estimates	537,405	MACH 1.0.16	36;22	MACH2QTL as implemented in GRIMP
SHIP	Affymetrix 6.0	BirdseedV2	0.92	No	No	No	QC callrate > 0.86 each Chip	869224	IMPUTE 0.5.0	36;22	SNPTEST 1.1.5
TwinsUK-I	Illumina 317K	Beadstudio	0.95	0.95 if MAF>0.05; <0.99 if 0.01<=MAF<0.05	5.7e10-7	0.01	unexpected relatedness based on pi_hat	296293	IMPUTE 0.5.0	36;22	GenABEL 1.4.2

Nature Genetics: doi:10.1038/ng.941

Supplementary Table 2 Literature search results for genes overlapping or proximal to novel lung function lead SNP and genes which contain a SNP with r^2>0.3 with lead SNP.

We systematically searched the scientific literature as catalogued by PubMed and refined our results using the PubMatrix text mining interface[6]. Additionally, we complemented our association findings by searching the Genetic Association Database—a comprehensive repository of published human genetic association studies[7]. Primers used for mRNA expression profiling are also included for each locus.

Lead SNP, Chr: position, chromosomal band	**Genes containing SNP with r^2>0.3 with sentinel SNP (HapMap release 22 CEU, functional annotation obtained using GeneCruiser).**	**Description of genes in region and biological plausibility**	**Animal model with lung function or pulmonary phenotype**	**Lookup in NHGRI GWAS table by sentinel SNP +/- 500kb and r^2>0.3. (P<5x10^{-8} in reported GWAS study)**
rs2284746 chr1: 17179262 1p36.13	**rs2284746**; ***MFAP2*** intron, sentinel. *rs6657613*; ***ATP13A2*** intron, r^2=0.84. *rs2871775*; ***SDHB*** intron, r^2=0.66. *rs2016693*; ***PADI2*** intron, r^2=0.43. *rs7513616*; ***CROCC*** intron, r^2=0.42.	***MFAP2***: Microfibrillar-associated protein 2 is a major antigen of elastin-associated microfibrils[8] and a candidate for involvement in the etiology of inherited connective tissue diseases. ***ATP13A2*** encodes a member of the P5 subfamily of ATPases which transports inorganic cations as well as other substrates. Mutations in ATP13A2 are associated with a form of Parkinson's Disease (PARK9, Kufor-Rakeb syndrome)[9]. ***SDHB*** is a component of the succinate dehydrogenase complex (II) of the respiratory chain and sporadic and familial mutations in this gene result in paragangliomas and pheochromocytoma[10]. ***PADI2*** encodes a peptidyl arginine deiminase which catalyses the post-translational deimination of proteins by converting arginine residues into citrullines in the presence of calcium ions. PADI2 protein levels are increased in bronchial mucosa and bronchoalveolar lavage cells of smokers[11]. ***CROCC*** encodes rootletin, a component of the rootlet at the proximal end of a cilium[12]. **mRNA expression profiling primers** (*MFAP2*) MFAP2-E8F(TCTGTGTTCGTACAGTGTGTGC), MFAP2-E9R(ACAGAGAGGCCTGGGATACTC) Expected product size: 398bp	*Crocc*-null mice develop photoreceptor degeneration and a lymphocytic lung infiltration consistent with impaired mucociliary clearance[13].	Sentinel SNP rs2284746 also associated with human height[14] (P=4x10^{-29}).
rs993925 chr1: 216926691 1q41	Sentinel SNP is 242.1kb downstream of *TGFB2* and 487.1kb upstream of *LYPLAL1* (no SNPs with r^2>0.3 with sentinel SNP in either gene).	This association signal is distal to *TGFB2*, transforming growth factor beta 2. TGFB2 is a member of a family of cytokines TGFB1-3 with a diverse repertoire of functions including regulating; cell proliferation, adhesion, migration and differentiation mediated via TGFB receptor 1 and 2 activation and smad signalling. TGFB2 has been suggested as a pro-fibrotic cytokine modulating epithelial repair mechanisms and extracellular matrix homeostasis including collagen deposition[15]**.** TGFB2 expression levels have been shown to be elevated in respiratory disease *e.g.* in bronchial epithelium in asthma[16]. **mRNA expression profiling primers** (*TGFB2*)	*Tgfb2-/-* mice have multiple cardiac defects, and dilated conducting airways with collapsed terminal bronchioles postnatally[17]. Tgfb2 contributes to airway remodeling and eosinophil recruitment in	none

Nature Genetics: doi:10.1038/ng.941

		TGFB2F(CGATCCGCGGGCAGATCCTG), TGFB2R(TGGAGGTGCCATCAATACCTGCA) Expected product size: 741bp	OVA-induced asthma model[18]	
rs12477314 chr2: 239542085 2q37.3	rs12477314 is 92.7kb downstream of *HDAC4* and 29.2kb upstream of *FLJ43879*.	***HDAC4*** (histone deacetylase 4) is a member of a family of proteins with enzymatic activity that removes acetylation from histone surrounding DNA thus influencing transcription factor access to the DNA. This deacetylase activity is thought to repress gene transcription. Reduced HDAC activity has been reported in lung tissue from subjects with COPD and this activity correlated with disease severity and inflammatory cytokine expression although HDAC2, 5 and HDAC8 differential expression levels were thought to mediate this observation[19]. Ongoing approaches to recover HDAC activity show potential for therapeutic intervention in COPD[20] **mRNA expression profiling primers** (*HDAC4*) HDAC4-E23F(CATGTGTTTCTGCCTTGCTG), HDAC4-E26R(GTCTGTCCCGTGTTCCCTGT) Expected product size: 523bp	*Hdac4*-null mice die soon after birth, displaying significant skeletal defects and growth retardation that impair mobility and breathing[21].	none
rs1529672 chr3: 25495586 3p24.2	**rs1529672**; ***RARB*** intron, sentinel.	***RARB*** codes for retinoic acid receptor beta, a nuclear receptor responsive to retinoic acid, a vitamin A derivative. This receptor also controls cell proliferation and differentiation. Retinoic acid has been implicated in embryonic lung branching morphogenesis[22]. Epigenetic regulation of the RARB gene promoter has been linked to various cancers including non-small cell lung cancer[23]. **mRNA expression profiling primers** (*RARB*) RARB-E3F(GATGCCAATACTGTCGACTCC), RARB-E6R(CAGAGGACCAAATCCAGCAT) Expected product size: 500bp	*Rarb*-null mice exhibit premature alveolar septation[24].	none
rs1344555 chr3: 170782913 3q26.2	**rs1344555**; ***EVI1*** intron, sentinel	***EVI1***, ecotropic viral integration site 1 codes for a zinc finger transcription factor and evidence suggests a proto-oncogenic role in proliferation in acute myeloid leukaemia[25,26]. The role of EVI1 in the lung is unclear at this time although increases in copy number have been reported in human squamous cell lung cancer[27]. Component of MECOM (*MDS1* and *EVI1* complex locus). **mRNA expression profiling primers** (*EVI1*) EVI1F(TGACCTGGAAACAACCTCGGGC), EVI1R(GCATTGGGAGGCGCCCTGAA) Expected product size: 951bp	*Evi1* is abundantly expressed in the bronchial epithelium of the developing mouse embryo[28].	none
rs153916 chr5: 95062456 5q15	rs153916 is 17.9kb upstream of ***SPATA9*** and 30.1kb upstream of **RHOBTB3**. *rs34895*; ***SPATA9*** intron, r^2=0.65.	Little is known regarding ***SPATA9***, spermatogenesis associated 9. However, this family of proteins were initially identified as mediators of spermatogenesis. Additional functions have now been identified for these proteins *e.g.* SPATA2 has been suggested to have a role in pancreatic development and β-cell proliferation[29]. ***RHOBTB3*** encodes a Rho GTPase which is required for protein transport from endosomes to the trans Golgi network. **mRNA expression profiling primers** (*SPATA9*) SPATAF2(CGAGCAACATTAATTCGTGGAT), SPATAR2(CACCTTCAGCAAACATAGGCT) Expected product size: 406bp	none	none
rs6903823 chr6:	**rs6903823**; ***ZKSCAN3*** intron, sentinel. *rs6912584*; ***ZNF323*** intron, r^2=0.95.	***ZKSCAN3***, zinc finger protein with KRAB and SCAN domains 3 is a transcription factor that has recently been identified as a regulator of many downstream targets including genes involved	none	SNP rs4324798 (r^2=0.42 with

Nature Genetics: doi:10.1038/ng.941

28430275 6p22.1	*rs7766356*; ***ZSCAN23*** 3' UTR, r^2=0.80. *rs7774981*; ***ZSCAN12*** 3'UTR, r^2=0.66. *rs13215054*; ***GPX6*** intron, r^2=0.60. *rs6901575*; ***PGBD1*** intron, r^2=0.51. *rs13205911*; ***ZNF192*** 3'UTR, r^2=0.50. *rs17763089*; ***HIST1H1B*** exon (synonymous), r^2=0.47. *rs200981;* ***HIST1H2AL*** exon (synonymous), r^2=0.42. *rs1150726*; ***ZNF187*** intron, r^2=0.40. *rs203878*; ***ZNF165*** 5'UTR, r^2=0.37. *rs1654774*; ***ZSCAN16*** intron, r^2=0.34. rs7206; ***ZNF193*** 3'UTR, r^2=0.33. *rs17720293*; ***ZKSCAN4*** intron, r^2=0.31. <u>Non-synonymous SNPs (gene, r^2 >0.3 with sentinel):</u> rs13201753 (*ZKSCAN3*, r^2=0.51). rs1997660 (*PGBD1*, r^2=0.49). rs2232423 (*ZSCAN12*, r^2=0.38). rs6456811 and rs3800325 (*PGBD1*, r^2=0.34). rs853684 (*ZNF323*, r^2=0.33).	in; cell growth/cell cycle/signal transduction *e.g.* VEGF, FGF23, cyclin D1, extracellular matrix and proteolysis *e.g.* MMP26, integrin beta 4 and transcription regulators *e.g.* ELF3[30]. Importantly, these functions have been linked to tumour growth[30] and proliferation in multiple myeloma[31]. ***ZNF323*** (zinc finger 323) is a member of the subfamily of C2H2 Kruppel-like zinc finger transcription factors. Preliminary mRNA expression data for *ZNF323* suggests a potential role in multiple organ embryogenesis including the lung[32]. ***ZNF192* (*ZKSCAN8*), *ZNF187* (*ZSCAN26*), *ZNF165* (*ZSCAN7*), *ZNF193*(*ZSCAN9*), *ZSCAN23*, *ZSCAN12*** and ***ZSCAN16*** encode zinc finger and SCAN domain-containing proteins. ***GPX6*** encodes a glutathione peroxidase. ***PGBD1*** piggyBac transposable element derived 1 is within part of a haplotype linked to high CD8+ T-lymphocyte numbers and Hereditary Hemochromatosis[33]. ***HIST1H1B*** and ***HIST1H2AL*** encode members of the H1 histone protein family, basic nuclear proteins which form the nucleosome which wraps and compacts the DNA into chromatin and may play a role in transcription regulation, DNA repair, DNA replication and chromosomal stability. ***ZKSCAN4*** encodes a zinc finger protein with KRAB and SCAN domains which may inhibit glucocorticoid receptor-mediated transactivation[34]. **eQTL result:** ***ZFP57*** encodes a zinc finger protein with a KRAB domain with a role in methylation associated with imprinting[35]. Mutations in this gene have been implicated in transient neonatal diabetes[36]. ***ZSCAN12* (*ZNF96*)** is described above. **mRNA expression profiling primers** (*ZKSCAN3*) ZKSCAN3F(GCGCCGCAGGTTTCAGGTGTT), ZKSCAN3R(CTGCTGTGTCCATTCAGGGGTGA) Expected product size: 300bp		rs6903823) associated with lung cancer[37] (P=2x10^{-8}).
<u>rs2857595</u> Chr6: 31676448 6p21.33	Sentinel SNP is 7.7kb upstream of ***NCR3*** (no SNP in NCR3 with r^2>0.3 with sentinel SNP), 14.6kb upstream of ***AIF1***. *rs3093975*; ***BAT1*** intron, r^2=0.95. *rs1800629*; ***TNFA*** promoter region (upstream), r^2=0.86. *rs3130614*; ***MICB*** intron, r^2=0.67. *rs7750641*; ***TCF19*** exon (non-synonymous), r^2=0.54. *rs2853975*; ***MICA*** intron, r^2=0.52. *rs3130557*; ***PSORS1C1(SEEK1)*** intron, r^2=0.50. *rs2233980*; ***C6orf15*** exon (synonymous), r^2=0.50. *rs1041981*; ***LTA*** exon (non-synonymous), r^2=0.47. *rs2071592*; ***NFKBIL1*** intron, r^2=0.47. *rs2071594*; ***ATP6V1G2*** 3'UTR, r^2=0.46.	***NCR3***, natural cytotoxity triggering receptor 3, blocking antibodies *in vitro* have demonstrated that NCR3 is required for efficient cytotoxicity responses by natural killer cells against normal cells and tumours[38]. AIF1, allograft inflammatory factor 1, is induced by chemokines and interferon and is involved in the regulation of vascular smooth muscle growth. AIF1 has been implicated in the pulmonary vasculopathy that characterizes systemic sclerosis[39,40]. ***TNFA***: rs1800629 (-308G/A) in the promoter region of TNFA (r^2 0.86) has previously been associated with asthma susceptibility in European ancestry subjects[41]. rs1800629 (-308G/A) been associated with lung function in COPD cases[42], but candidate gene studies reporting its association to COPD susceptibility have not provided consistent and conclusive evidence[43] and suggest publication bias. ***BAT1-BAT5*** are HLA-B associated transcripts. ***MICA*** and ***MICB*** encode the MHC class I chain-related proteins A and B[44]. These proteins are stress-induced and ligands for the activation of the cytolytic response of NK cell receptors. They are similar to MHC class I molecules but do not associate with beta-2-microglobulin or bind peptides. Polymorphisms in MICB have been associated with airflow obstruction in asthmatics[45]. ***TCF19*** encodes transcription factor 19. ***PSORS1C1*(*SEEK1*)** is a psoriasis susceptibility candidate gene[46] and overlaps *CDSN***.** The	*Tnfa* overexpression in mice results in increased lung volume, loss of small airspaces and thickened pleural septa[60]. *Lta* knockout mice have increased airway inflammation and remodelling, and elicit a severe Th1 response when challenged by allergens[61] *Cfb*-null mice infected with *P. aeruginosa* pneumonia exhibit increased mortality[62]	SNP rs3117582 (r^2=0.38 with rs2857595) associated with lung cancer[37,57,63] (strongest P=5x10^{-12}). SNP rs3099844 (r^2=0.67 with rs2857595) associated with neonatal lupus[64] (P=5x10^{-10}). SNP rs3131379 (r^2=0.34 with rs2857595)

Nature Genetics: doi:10.1038/ng.941

	rs3132450; ***BAT2(PRRC2A)*** intron, r^2=0.38. *rs3117582*; ***APOM*** intron, r^2=0.38. *rs9267531*; ***CSNK2B*** intron, r^2=0.38. *rs3131383*; ***CLIC1*** 5'UTR, r^2=0.34. *rs3131379*; **MSH5** intron, r^2=0.34. *rs915652*; ***VARS*** intron, r^2=0.34. *rs3130679*; ***C6orf48*** 3'UTR, r^2=0.34. *rs3132550*; ***CDSN*** intron, r^2=0.31. *rs3130628*; ***BAT3(BAG6)*** intron, r^2=0.30. *rs3130618*; ***BAT4(GPANK1)*** exon (non-synonymous), r^2=0.30. *rs1270942*; **CFB** intron, r^2=0.30. *rs389884*; **STK19** intron, r^2=0.30. <u>Non-synonymous SNPs (gene, r^2 >0.3 with sentinel):</u> rs7750641 (*TCF19*, r^2=0.54). rs1041981 (*LTA*, r^2=0.47). rs2233974 (*C6orf15*, r^2=0.38). rs3130618 (*BAT4*, r^2=0.30).	function of PSORS1C1 is unknown. ***CDSN*** encodes a protein found in corneodesmosomes, structures found within human epidermis and other cornified squamous epithelia. ***C6orf15*** **(*STG*)** No function is known for this gene in humans but the strongest signal of association in this region for a study of follicular lymphoma overlapped this gene[47]. ***LTA*** encodes lymphotoxin alpha, a cytokine produced by lymphocytes and a member of the tumour necrosis factor family. This gene has been associated with leprosy[48] and psoriatic arthritis[49]. ***NFKBIL1*** encodes a putative inhibitor of the NFκB complex which has a role in DNA transcription regulation and has shown association with rheumatoid arthritis[50] and may confer susceptibility to chronic thromboembolic-induced pulmonary hyptertension[51]. ***ATP6V1G2*** encodes a component of vacuolar ATPase (V-ATPase), a multisubunit enzyme that mediates acidification of intracellular compartments of eukaryotic cells. ***APOM*** encodes apolipoprotein M which may play a role in high density lipid metabolism[52]. ***CSNK2B*** encodes the beta subunit of casein kinase II which is believed to play a role in a wide range of cellular processes[53] including Wnt/beta-catenin signalling[54]. ***CLIC1*** encodes chloride intracellular channel 1 which may play a role in the maintenance of intracellular membranes[55]. ***MSH5*** encode a MutS-homologous protein with roles in meiotic and mitotic DNA recombination[56]. Common variants in the MSH5 locus have been associated with risk of lung cancer[57]. ***VARS*** encodes valyl tRNA synthetase which binds the tRNA with valine for translation. Other aminoacyl tRNA synthetases have been associated with disease including neurodegeneration[58]. ***C6orf48*** Function unknown. ***CFB*** encodes complement factor B, a component of the alternative pathway of complement activation. Complement factor B is cleaved into Ba and Bb with Bb being a serine protease that associates with C3b to form the alternative pathway C3 convertase. Polymorphisms in this gene have been associated with a reduced risk of age-related macular degeneration[59]. ***STK19*** encodes serine/threonine kinase 19 and may be involved in transcription regulation. **eQTL results:** ***HLA-DRB1*** and ***HLA-DQA1*** encode HLA beta and alpha chain class II paralogues, respectively, that play a key role in the immune system by presenting peptides derived from extracellular proteins. **mRNA expression profiling primers** (*NCR3*) NCR3-E2F(ATGCCAGCCAAGGGAGACT), NCR3-E4R(ACAGCCAGAAGAGGGTATGTGT) Expected product size: 577bp		associated with SLE[65] (P=2x10^{-52}).
<u>rs2798641</u> chr6: 109374743 6q21	**rs2798641*; ARMC2*** intron, sentinel.	***ARMC2***, armadillo repeat containing 2 is a member of a protein family containing a repeating motif of ~42 amino acids consisting of three alpha helices. Protein family members have been identified as having roles in many diverse protein functions including; cell signalling, protein degradation and cytoskeleton functions[66]. At this time little is known regarding the function of ARMC2. **mRNA expression profiling primers** (*ARMC2*) ARMC2F(GTGCTGCTTGCACACAACTTCAT), ARMC2R(TTCCATTGCCGTGCAGAGCTGG)	none	none

Nature Genetics: doi:10.1038/ng.941

		Expected product size: 578bp		
rs7068966 chr10: 12317998 10p13	**rs7068966**; ***CDC123*** intron, sentinel. *rs7072888*; ***NUDT5*** intron, r^2=0.37.	***CDC123***, cell division cycle 123 homolog shows high homology across species and has been shown to be a critical control protein modulating Eukaryotic initiation factor 2 in times of cell stress *e.g.* nutrition and growth in *Saccharomyces cerevisiae*[67]. ***NUDT5*** Human NUDT5 (hNUDT5) is an ADP-ribose (ADPR) pyrophosphatase (ADPRase) that plays important roles in controlling the intracellular levels of ADPR and preventing non-enzymatic ADP-ribosylation of proteins by hydrolyzing ADPR to AMP and ribose 5'-phosphate[68]. **eQTL results** *OSBPL1A* (located on chromosome 18) encodes an oxysterol-binding protein. **mRNA expression profiling primers** (*CDC123*) CDC123-E1F(CGAGGCGTTACCATCAAGA), CDC123-E8R(ATTCACACCATTTTCGGAGAAC) Expected product size: 460bp	none	none
rs11001819 chr10: 77985230 10q22.3	**rs11001819; *c10orf11*** intron, sentinel *rs7078013*; ***KCNMA1*** intron, r^2=0.54.	***C10orf11*** Individuals with a rare deletion that disrupts C10orf11 exhibit cognitive defects[69]. ***KCNMA1*** encodes the potassium large conductance calcium-activated channel, subfamily M, alpha member 1 (also known as BKCa), is present on airway smooth muscle cells and pulmonary endothelium and its activation may lead to bronchial relaxation[70] and pulmonary arterial vasodilation[71]. **mRNA expression profiling primers** (*C10orf11*): C10orf11-E2F(GGAGGAACTCATCTTGGACAAC), C10orf11-E6RB(TCCCATAGTAAACGTAGCGACA) Expected product size: 473bp	none	none
rs11172113 chr12: 55813550 12q13.3	**rs11172113**; ***LRP1*** intron, sentinel *rs167769*; ***STAT6*** intron, r^2=0.36	***LRP1***, low density lipoprotein receptor-related protein 1 is a member of a family of receptors involved in cholesterol homeostasis[72]. However, this receptor interacts with a diverse number of ligands *e.g.* ApoE, MMP-2, MMP-9, uPA-uPAR pathway components, amyloid precursor protein, and mediates the cellular processes related to these proteins including cell signalling and migration[73]. Several of these ligands have been implicated via differential expression in disease progression in COPD including MMP-9, uPA and uPAR[74]. ***STAT6*** encodes a signal transducer and activator of transcription 6, interleukin-4 induced protein and may play a role in allergic inflammation[75] and pathways leading to pulmonary fibrosis[76]. **mRNA expression profiling primers** (*LRP1*) LRP1F(AACCGGTACTACCTGCGCAAGC), LRP1R(TGGCTAGCCGTGCAGTTGGACA) Expected product size: 911bp	*Stat6*-null mice are protected from allergen-induced airway inflammation and eosinophilia[77]. Pharmacologic inhibition of *Stat6* ameliorates ovalbumin-induced bronchial constriction in mice[78].	none
rs1036429 chr12: 94795559 12q22	**rs1036429**; ***CCDC38*** intron, sentinel. *rs4762633*; ***SNRPF*** intron, r^2=0.96.	***CCDC38***, coiled-coil domain containing 38 is a member of a diverse protein family that contain two or more alpha helical domains that twist around each other forming a supercoil. These proteins are involved in a diverse array of functions and includes; skeletal proteins *e.g.* vimentin and motor proteins *e.g.* myosin[79]. At this time little is known regarding the function of CCDC38. ***SNRPF*** Small nuclear ribonucleoprotein F. The presence of anti-RNP antibodies has been associated with increased progression to pulmonary damage in individuals with Systemic Lupus Erythematosus[80].	none	none

Nature Genetics: doi:10.1038/ng.941

		mRNA expression profiling primers (*CCDC38*): CCDC38-E12F(AGAGAGCTGGAAGAGCAGAATC), CCDC38-E15R(TAGCCTTTCCTGTTGGTGTCTT) Expected product size: 498bp **mRNA expression profiling primers** (*SNRPF*): SNRPFF(CGGTACCTGCTGTAGTCACGA),SNRPFR(CTTCACCCAGATGTCCAGACA) Expected product size: 262bp		
rs12447804 chr16: 56632783 16q13	**rs12447804; *MMP15*** intron, sentinel. Non-synonymous SNPs (gene, r^2 >0.3 with sentinel): rs3743563 (*MMP15*, r^2=1).	***MMP15***, matrix metalloproteinase 15 (also called membrane type matrix metalloproteinase 2, MT2-MMP) is a member of a large protease family with diverse functional roles via protease activity and specificity including; tissue remodelling, wound healing, angiogenesis, and tumor invasion. MMP15 expression has been shown to be elevated in primary lung adenocarcinoma[81] and MMP15 has demonstrated anti- apoptosis properties[82]. MMP15 substrates include Pro-MMP-2, gelatin, fibronectin and laminin[83]. **mRNA expression profiling primers** (*MMP15*) MMP15-E1F(CTGCTCCTGGTGCTTCTGG), MMP15-E4R(AGGTCAGTGCTGGAGAAGGTC) Expected product size: 650bp	*Mmp15* expression is increased in the lungs of mice exposed to particulate matter[84].	none
rs2865531 chr16: 73947817 16q23.1	**rs2865531; *CFDP1*** intron, sentinel *rs12933281*; ***TMEM170A*** intron, r^2=0.97. *rs3115960*; ***CHST6*** intron, r^2=0.76.	***CFDP1*** encodes Craniofacial Development Protein 1. ***TMEM170A*** unknown. ***CHST6*** encodes carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 6. Mutations in this gene have been associated with macular corneal dystrophy. **mRNA expression profiling primers** (*CFDP1*) CFDP1E1F(TGCGGTCTTGTGAGTTTGAC), CFDP1E4R(GCTCTTCTGCTTTTACCAACAA) Expected product size: 500bp	*Cfdp1* is highly expressed in developing mouse lung[85]	none
rs9978142 chr21: 34574109 21q22.11	rs9978142 is 84.1kb upstream of ***KCNE2*** and 90kb downstream of ***C21orf82***.	***KCNE2***, potassium channel, voltage-gated, ISK-related subfamily, member 2 has been associated with thyroid dysfunction using KCNE2-/- mice which exhibited; hypothyroidism, dwarfism, alopecia, goitre and cardiac abnormalities[86]. The authors concluded that the KCNQ1-KCNE2 TSH-stimulated thyrocyte K+ channel was crucial for normal thyroid I– accumulation[86]. KCNQ1-KCNE2 K+ channels may modulate transepithelial anion secretion in Calu3 airway epithelial cells[87]. ***C21orf82*** unknown **mRNA expression profiling primers** (*KCNE2*) KCNE2F(ATTGGCGCCAGAACACAACAGCT), KCNE2R(GACATTTTGAACCCAGCCGCACC) Expected product size: 292bp	none	SNP rs9982601 (r^2=0.68 with rs9978142) associated with myocardial infarction[88] (P=6x10^{-11}).

Nature Genetics: doi:10.1038/ng.941

Supplementary Table 3 Association of the 16 novel regions, and previously reported regions, with A) FEV_1 and FEV_1/FVC in children, B) tobacco addiction, C) height and D) lung cancer. The effect of SNPs previously associated with height on lung function are also shown (E).

A) Association of the 16 novel regions, and previously reported regions, with FEV_1 and FEV_1/FVC in children. Effects of the 16 novel SNPs, and effects of SNPs in regions previously reported as associated with lung function, on FEV_1 and FEV_1/FVC in children were looked up in ALSPAC (Avon Longitudinal Study of Parents and Children, N=5062, 50% males, age mean (s.d.)=8.6(0.3)) and in the Raine Study (N=1219, 52% males, age mean (s.d.)=8.1(0.3)). Effect sizes estimates for the children's cohorts were meta-analysed using inverse variance weighting. *P* values still significant after applying a Bonferroni correction for 16 tests ($P<0.003125$) are shown in bold. To enable a comparison of effect sizes between children and adults, effect sizes in the SpiroMeta-CHARGE Stage 2 dataset only (to avoid potential winners' curse bias) are given for the novel loci. Effect sizes in the SpiroMeta-CHARGE GWAS Stage 1 are provided for the previously reported regions.

Nature Genetics: doi:10.1038/ng.941

Chr.	Measure	SNP_ID(NCBI36 position), function	Coded allele	ALSPAC N=5062			Raine N=1219			ALSPAC+Raine meta-analysis			SpiroMeta-CHARGE Stage 2		
				Beta	Se	*P*	Beta	Se	*P*	Beta	Se	*P*	Beta	Se	*P*
Novel loci															
1	FEV_1/FVC	rs2284746 (17179262), *MFAP2*(intron)	G	0.005	0.028	8.58E-01	0.002	0.043	9.61E-01	0.004	0.023	8.60E-01	-0.038	0.007	2.640E-07
1	FEV_1	rs2284746 (17179262), *MFAP2*(intron)	G	-0.064	0.029	2.70E-02	0.096	0.042	2.30 E-02	-0.013	0.024	5.93E-01	0.006	0.007	3.701E-01
1	FEV_1/FVC	rs993925 (216926691), *TGFB2*(downstream)	T	0.052	0.029	7.30E-02	0.015	0.05	7.67E-01	0.043	0.025	8.90E-02	0.023	0.01	1.758E-02
1	FEV_1	rs993925 (216926691), *TGFB2*(downstream)	T	0.058	0.03	5.30E-02	-0.011	0.049	8.20E-01	0.039	0.026	1.26E-01	0.003	0.007	7.292E-01
2	FEV_1/FVC	rs12477314 (239542085), *HDAC4*(downstream)	T	0.093	0.035	8.00E-03	0.063	0.048	1.93E-01	0.083	0.028	4.00E-03	0.031	0.008	8.410E-05
2	FEV_1	rs12477314 (239542085), *HDAC4*(downstream)	T	0.025	0.036	4.87E-01	0.058	0.048	2.31E-01	0.037	0.029	2.03E-01	0.025	0.007	1.816E-04
3	FEV_1/FVC	rs1529672 (25495586), *RARB*(intron)	C	-0.078	0.036	3.00E-01	-0.034	0.053	5.24E-01	-0.064	0.03	3.10E-02	-0.038	0.009	1.160E-05
3	FEV_1	rs1529672 (25495586), *RARB*(intron)	C	0.044	0.037	2.34E-01	0.011	0.053	8.41E-01	0.033	0.03	2.76E-01	-0.011	0.007	9.328E-02
3	FEV_1/FVC	rs1344555 (170782913), *MECOM*(intron)	T	0.016	0.035	6.48E-01	-0.068	0.053	2.00E-01	-0.01	0.029	7.43E-01	-0.017	0.012	1.551E-01
3	FEV_1	rs1344555 (170782913), *MECOM*(intron)	T	-0.013	0.036	7.18E-01	-0.064	0.052	2.19E-01	-0.03	0.03	3.20E-01	-0.025	0.009	6.436E-03
5	FEV_1/FVC	rs153916 (95062456), *SPATA9*(upstream)	T	0.037	0.027	1.71E-01	0.019	0.04	6.31E-01	0.032	0.022	1.60E-01	-0.025	0.009	6.671E-03
5	FEV_1	rs153916 (95062456), *SPATA9*(upstream)	T	0.001	0.028	9.72E-01	0.05	0.04	2.19E-01	0.017	0.023	4.64E-01	0.004	0.007	6.220E-01
6	FEV_1/FVC	rs6903823 (28430275), *ZKSCAN3*(intron)/*ZNF323*(intron)	G	-0.037	0.032	2.48E-01	-0.131	0.049	7.00E-03	-0.065	0.027	1.40E-02	-0.013	0.011	2.337E-01
6	FEV_1	rs6903823 (28430275), *ZKSCAN3*(intron)/*ZNF323*(intron)	G	0.02	0.034	5.56E-01	-0.036	0.049	4.64E-01	0.002	0.028	9.51E-01	-0.029	0.008	4.750E-04
6	FEV_1/FVC	rs2857595 (31676448), *NCR3*(upstream)	G	0.044	0.034	1.96E-01	0.08	0.05	1.09E-01	0.055	0.028	4.90E-02	0.028	0.008	5.364E-04
6	FEV_1	rs2857595 (31676448), *NCR3*(upstream)	G	-0.013	0.035	7.10E-01	-0.076	0.05	1.30E-01	-0.034	0.029	2.41E-01	0.017	0.007	9.411E-03
6	FEV_1/FVC	rs2798641 (109374743), *ARMC2*(intron)	T	-0.102	0.036	5.00E-03	0.049	0.052	3.49E-01	-0.053	0.03	7.30E-02	-0.03	0.012	1.565E-02
6	FEV_1	rs2798641 (109374743), *ARMC2*(intron)	T	-0.115	0.037	2.00E-03	-0.062	0.052	2.33E-01	-0.097	0.03	**1.00E-03**	-0.009	0.01	3.351E-01
10	FEV_1/FVC	rs7068966 (12317998), *CDC123*(intron)	T	0.051	0.027	5.90E-02	-0.03	0.04	4.64E-01	0.026	0.022	2.45E-01	0.023	0.006	3.861E-04
10	FEV_1	rs7068966 (12317998), *CDC123*(intron)	T	0.041	0.029	1.57E-01	0.044	0.04	2.74E-01	0.042	0.024	7.40E-02	0.022	0.005	3.560E-05
10	FEV_1/FVC	rs11001819 (77985230), *C10orf11*(intron)	G	-0.048	0.028	8.60E-02	-0.018	0.041	6.62E-01	-0.038	0.023	9.70E-02	-0.006	0.006	3.174E-01
10	FEV_1	rs11001819 (77985230), *C10orf11(*intron)	G	-0.04	0.029	1.68E-01	0.003	0.041	9.45E-01	-0.026	0.024	2.78E-01	-0.022	0.005	3.100E-05
12	FEV_1/FVC	rs11172113 (55813550), *LRP1*(intron)	T	-0.011	0.028	6.94E-01	-0.054	0.041	1.87E-01	-0.025	0.023	2.85E-01	-0.026	0.01	5.829E-03
12	FEV_1	rs11172113 (55813550), *LRP1*(intron)	T	-0.039	0.029	1.79E-01	-0.032	0.041	4.34E-01	-0.037	0.024	1.21E-01	-0.003	0.007	6.940E-01
12	FEV_1/FVC	rs1036429 (94795559), *CCDC38*(intron)	T	0.066	0.034	5.20E-02	0.014	0.052	7.82E-01	0.05	0.028	7.60E-02	0.028	0.008	3.348E-04
12	FEV_1	rs1036429 (94795559), *CCDC38*(intron)	T	-0.015	0.036	6.77E-01	0.001	0.052	9.79E-01	-0.01	0.03	7.43E-01	0.004	0.006	5.381E-01
16	FEV_1/FVC	rs12447804 (56632783), *MMP15*(intron)	T	-0.06	0.033	6.90E-02	0.089	0.051	8.40E-02	-0.017	0.028	5.52E-01	-0.021	0.01	4.204E-02
16	FEV_1	rs12447804 (56632783), *MMP15*(intron)	T	-0.074	0.034	3.00E-02	-0.006	0.051	9.12E-01	-0.053	0.028	6.10E-02	0.004	0.007	5.705E-01

Contd.

Nature Genetics: doi:10.1038/ng.941

				ALSPAC N=5062			Raine N=1219			ALSPAC+Raine meta-analysis			SpiroMeta-CHARGE Stage 2		
Chr.	**Measure**	**SNP_ID(NCBI36 position), function**	**Coded allele**	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***
16	FEV_1/FVC	rs2865531 (73947817), *CFDP1*(intron)	T	-0.016	0.028	5.68E-01	0.023	0.041	5.85E-01	-0.004	0.023	8.67E-01	0.024	0.006	1.940E-04
16	FEV_1	rs2865531 (73947817), *CFDP1*(intron)	T	-0.037	0.029	2.02E-01	-0.066	0.041	1.12E-01	-0.046	0.024	5.00E-02	0.011	0.005	3.888E-02
21	FEV_1/FVC	rs9978142 (34574109), *KCNE2*(upstream)	T	-0.031	0.04	4.38E-01	0.125	0.059	3.50E-02	0.018	0.033	5.90E-01	-0.031	0.013	1.749E-02
21	FEV_1	rs9978142 (34574109), *KCNE2*(upstream)	T	-0.035	0.042	4.05E-01	0.095	0.059	1.12E-01	0.008	0.034	8.13E-01	-0.015	0.01	1.353E-01
Previously reported regions															
2	FEV_1/FVC	rs2571445(218391399), *TNS* (ns)	G	0.033	0.028	2.39E-01	-0.037	0.042	3.74E-01	0.011	0.023	6.31E-01	0.033	0.007	4.460E-06
2	FEV_1	rs2571445(218391399), *TNS1*(ns)	G	0.05	0.029	8.47E-02	0.045	0.042	2.83E-01	0.048	0.024	4.25E-02	0.047	0.007	9.830E-11
2	FEV_1/FVC	rs10498230(229210747), *PID1*(downstream)	T	0.072	0.052	1.66E-01	-0.073	0.076	3.39E-01	0.026	0.043	5.44E-01	0.068	0.014	1.130E-06
2	FEV_1	rs10498230(229210747), *PID1*(downstream)	T	0.078	0.054	1.49E-01	0.054	0.076	4.81E-01	0.07	0.044	1.13E-01	0.03	0.014	3.601E-02
4	FEV_1/FVC	rs2045517(90089987), *FAM13A*(intron)	T	0.004	0.028	8.86E-01	-0.03	0.042	4.72E-01	-0.007	0.023	7.78E-01	-0.047	0.007	2.000E-11
4	FEV_1	rs2045517(90089987), *FAM13A*(intron)	T	-0.035	0.029	2.27E-01	-0.013	0.042	7.57E-01	-0.028	0.024	2.43E-01	-0.012	0.007	8.933E-02
4	FEV_1/FVC	rs7671167(90103002), *FAM13A*(intron)	T	0.02	0.027	4.59E-01	-0.016	0.041	6.96E-01	0.009	0.023	6.85E-01	-0.042	0.007	1.270E-09
4	FEV_1	rs7671167(90103002), *FAM13A*(intron)	T	-0.024	0.028	3.91E-01	-0.017	0.041	6.74E-01	-0.022	0.023	3.44E-01	-0.017	0.007	1.635E-02
4	FEV_1/FVC	rs10516526(106908353), *GSTCD*(intron)	G	0.158	0.055	4.07E-03	-0.007	0.081	9.35E-01	0.106	0.045	1.98E-02	0.039	0.014	6.167E-03
4	FEV_1	rs10516526(106908353), *GSTCD*(intron)	G	0.069	0.057	2.26E-01	0.169	0.081	3.66E-02	0.102	0.047	2.81E-02	0.108	0.014	4.750E-14
4	FEV_1/FVC	rs17331332(107027556), *NPNT*(upstream)	G	-0.132	0.054	1.45E-02	0.031	0.08	6.95E-01	-0.081	0.045	7.14E-02	-0.057	0.014	5.300E-05
4	FEV_1	rs17331332(107027556), *NPNT*(upstream)	G	-0.074	0.056	1.86E-01	-0.177	0.079	2.57E-02	-0.108	0.046	1.79E-02	-0.102	0.014	1.110E-12
4	FEV_1/FVC	rs6823809(107048244), *NPNT*(intron)	T	0.113	0.039	3.76E-03	0.101	0.101	3.16E-01	0.112	0.036	2.18E-03	0.056	0.011	2.200E-07
4	FEV_1	rs6823809(107048244), *NPNT*(intron)	T	0.043	0.041	2.94E-01	0.103	0.095	2.80E-01	0.052	0.038	1.64E-01	0.05	0.011	4.820E-06
4	FEV_1/FVC	rs1032296(145654138), *HHIP*(upstream)	T	-0.012	0.029	6.79E-01	0.013	0.043	7.56E-01	-0.004	0.024	8.68E-01	-0.05	0.007	3.420E-12
4	FEV_1	rs1032296(145654138), *HHIP*(upstream)	T	0.02	0.03	5.05E-01	-0.053	0.042	2.12E-01	-0.004	0.024	8.57E-01	-0.047	0.007	8.740E-11
4	FEV_1/FVC	rs11100860(145698589), *HHIP*(upstream)	G	0.016	0.029	5.81E-01	-0.02	0.041	6.25E-01	0.004	0.024	8.68E-01	0.064	0.007	6.810E-20
4	FEV_1	rs11100860(145698589), *HHIP*(upstream)	G	-0.015	0.03	6.17E-01	0.056	0.041	1.70E-01	0.01	0.024	6.82E-01	0.041	0.007	4.270E-09
5	FEV_1/FVC	rs11168048(147822546), *HTR4*(intron)	T	-0.072	0.029	1.30E-02	0.011	0.041	7.86E-01	-0.044	0.024	6.08E-02	-0.047	0.007	5.970E-11
5	FEV_1	rs11168048(147822546), *HTR4*(intron)	T	-0.003	0.03	9.20E-01	0.057	0.041	1.71E-01	0.018	0.024	4.70E-01	-0.046	0.007	2.430E-10
5	FEV_1/FVC	rs3995090(147826008), *HTR4*(intron)	C	0.074	0.028	8.22E-03	-0.001	0.04	9.88E-01	0.049	0.023	3.12E-02	0.046	0.007	1.040E-10
5	FEV_1	rs3995090(147826008), *HTR4*(intron)	C	0	0.029	1.00E+00	-0.042	0.04	2.98E-01	-0.014	0.023	5.41E-01	0.045	0.007	3.330E-10
5	FEV_1/FVC	rs1985524(147827981), *HTR4*(intron)	G	-0.083	0.028	3.03E-03	0.04	0.04	3.21E-01	-0.043	0.023	6.10E-02	-0.045	0.007	2.900E-10
5	FEV_1	rs1985524(147827981), *HTR4*(intron)	G	-0.004	0.029	8.90E-01	0.052	0.041	2.05E-01	0.015	0.024	5.35E-01	-0.048	0.007	3.060E-11

Nature Genetics: doi:10.1038/ng.941

				ALSPAC N=5062			Raine N=1219			ALSPAC+Raine meta-analysis			SpiroMeta-CHARGE Stage 2		
Chr.	**Measure**	**SNP_ID(NCBI36 position), function**	**Coded allele**	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***
5	FEV_1/FVC	rs11134779(156869344), *ADAM19*(intron)	G	0.013	0.029	6.54E-01	-0.036	0.042	3.90E-01	-0.003	0.024	9.06E-01	-0.042	0.007	6.010E-09
5	FEV_1	rs11134779(156869344), *ADAM19*(intron)	G	-0.051	0.03	8.91E-02	0.039	0.042	3.60E-01	-0.021	0.024	3.93E-01	-0.027	0.007	2.396E-04
6	FEV_1/FVC	rs2070600(32259421), *AGER*(ns)	T	0.126	0.054	1.96E-02	0.192	0.081	1.83E-02	0.146	0.045	1.15E-03	0.126	0.016	9.070E-15
6	FEV_1	rs2070600(32259421), *AGER*(ns)	T	0.05	0.056	3.72E-01	0.089	0.081	2.73E-01	0.063	0.046	1.75E-01	0.025	0.016	1.271E-01
6	FEV_1/FVC	rs3817928(142792209), *GPR126*(intron)	G	0.106	0.035	2.46E-03	0.045	0.05	3.68E-01	0.086	0.029	2.75E-03	0.059	0.008	2.270E-12
6	FEV_1	rs3817928(142792209), *GPR126*(intron)	G	-0.033	0.036	3.59E-01	0.032	0.05	5.26E-01	-0.011	0.029	7.11E-01	0.023	0.009	8.625E-03
6	FEV_1/FVC	rs262129(142894837), *LOC153910*(unknown)	G	0.118	0.03	8.38E-05	0.054	0.044	2.23E-01	0.098	0.025	8.19E-05	0.056	0.008	2.910E-13
6	FEV_1	rs262129(142894837), *LOC153910*(unknown)	G	-0.021	0.032	5.12E-01	0.017	0.044	6.98E-01	-0.008	0.026	7.61E-01	0.031	0.008	5.440E-05
9	FEV_1/FVC	rs16909859(97244613), *PTCH1*(downstream)	G	0.057	0.05	2.54E-01	0.022	0.09	8.12E-01	0.049	0.044	2.66E-01	0.08	0.013	7.450E-10
9	FEV_1	rs16909859(97244613), *PTCH1*(downstream)	G	0.058	0.052	2.65E-01	-0.128	0.092	1.62E-01	0.013	0.045	7.80E-01	-0.014	0.013	2.933E-01
9	FEV_1/FVC	rs16909898(97270829), *PTCH1*(intron)	G	-0.089	0.048	6.37E-02	0.002	0.088	9.77E-01	-0.068	0.042	1.06E-01	-0.072	0.012	3.940E-09
9	FEV_1	rs16909898(97270829), *PTCH1*(intron)	G	-0.022	0.049	6.53E-01	0.098	0.089	2.67E-01	0.006	0.043	8.85E-01	0.015	0.012	2.211E-01
15	FEV_1/FVC	rs12899618(69432174), *THSD4*(intron)	G	0.045	0.039	2.49E-01	0.092	0.057	1.02E-01	0.06	0.032	6.03E-02	0.076	0.01	1.860E-15
15	FEV_1	rs12899618(69432174), *THSD4*(intron)	G	-0.024	0.041	5.58E-01	-0.029	0.057	6.10E-01	-0.026	0.033	4.39E-01	0.036	0.01	1.567E-04
15	FEV_1/FVC	rs8033889(69467134), *THSD4*(intron)	T	-0.051	0.034	1.34E-01	-0.104	0.051	4.09E-02	-0.067	0.028	1.70E-02	-0.072	0.008	2.030E-17
15	FEV_1	rs8033889(69467134), *THSD4*(intron)	T	0.031	0.035	3.76E-01	0.094	0.051	6.30E-02	0.051	0.029	7.41E-02	-0.044	0.009	3.010E-07
15	FEV_1/FVC	rs2568494(76528019), *IREB2*(intron)	G	0.014	0.029	6.29E-01	0.028	0.043	5.22E-01	0.018	0.024	4.48E-01	0.029	0.007	5.250E-05
15	FEV_1	rs2568494(76528019), *IREB2*(intron)	G	-0.023	0.03	4.43E-01	0.015	0.043	7.27E-01	-0.011	0.025	6.68E-01	0.023	0.007	1.635E-03
15	FEV_1/FVC	rs8034191(76593078), *CHRNA3/5*(intron)	T	0.026	0.029	3.70E-01	-0.01	0.042	8.13E-01	0.014	0.024	5.47E-01	0.032	0.007	9.650E-06
15	FEV_1	rs8034191(76593078), *CHRNA3/5*(intron)	T	-0.03	0.03	3.17E-01	0.007	0.042	8.67E-01	-0.017	0.024	4.75E-01	0.031	0.007	2.070E-05
15	FEV_1/FVC	rs2036527(76638670), *CHRNA5*(uptream)	G	0.038	0.03	2.05E-01	-0.015	0.047	7.57E-01	0.023	0.025	3.66E-01	0.032	0.007	1.190E-05
15	FEV_1	rs2036527(76638670), *CHRNA5*(uptream)	G	-0.015	0.031	6.28E-01	0.04	0.047	3.97E-01	0.002	0.026	9.51E-01	0.036	0.008	2.400E-06
15	FEV_1/FVC	rs8040868(76698236), *CHRNA3*(s)	T	0.039	0.028	1.64E-01	-0.027	0.042	5.19E-01	0.019	0.023	4.25E-01	0.04	0.008	1.140E-06
15	FEV_1	rs8040868(76698236), *CHRNA3*(s)	T	0.003	0.03	9.20E-01	0.041	0.042	3.21E-01	0.016	0.024	5.09E-01	0.039	0.008	2.980E-06

Nature Genetics: doi:10.1038/ng.941

B) Relation of 16 novel lung function loci, and previously reported loci, with tobacco addiction. Effects of the 16 novel SNPs associated with lung function on two smoking phenotypes (cigarettes per day and ever-smokers vs. never-smokers) were looked up in the Oxford-GlaxoSmithKline (Ox-GSK) study, a collaborative effort to investigate the genetic basis of smoking-related behavioral traits [89]. Results for the SpiroMeta-CHARGE joint meta-analysis of Stage 1 and Stage 2 for the lung function measure that showed stronger association are reported for the novel loci. Results for the SpiroMeta-CHARGE GWAS stage (stage 1) for the lung function measure that showed stronger association are reported for the previously reported loci.

Nature Genetics: doi:10.1038/ng.941

				Lung function (Stage 1 and Stage 2 meta-analysis)			Cigarettes per day			Ever vs. Never smoking		
Chr.	**SNP ID(NCBI36 position), function**	**Coded allele**	**Measure**	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***
Novel loci												
1	rs2284746 (17179262), *MFAP2*(intron)	G	FEV_1/FVC	-0.04	0.005	7.50E-16	0.004	0.01	6.85E-01	-0.015	0.018	3.99E-01
1	rs993925 (216926691), *TGFB2*(downstream)	T	FEV_1/FVC	0.034	0.006	1.16E-08	-0.007	0.011	5.12E-01	-0.027	0.019	1.53E-01
2	rs12477314 (239542085), *HDAC4*(downstream)	T	FEV_1/FVC	0.041	0.006	1.68E-12	-0.021	0.012	8.83E-02	0.014	0.022	5.20E-01
3	rs1529672 (25495586), *RARB*(intron)	C	FEV_1/FVC	-0.048	0.006	3.97E-14	-0.013	0.014	3.65E-01	0.024	0.025	3.38E-01
3	rs1344555 (170782913), *MECOM*(intron)	T	FEV_1	-0.034	0.006	2.65E-08	0.011	0.013	3.86E-01	0.01	0.022	6.51E-01
5	rs153916 (95062456), *SPATA9*(upstream)	T	FEV_1/FVC	-0.031	0.005	2.12E-08	-0.021	0.01	3.70E-02	0.02	0.018	2.72E-01
6	rs6903823 (28430275), *ZKSCAN3*(intron)/*ZNF323*(intron)	G	FEV_1	-0.037	0.006	2.18E-10	0.02	0.012	8.86E-02	-0.026	0.021	2.16E-01
6	rs2857595 (31676448), *NCR3*(upstream)	G	FEV_1/FVC	0.037	0.006	2.28E-10	-0.002	0.014	8.55E-01	-0.009	0.025	7.12E-01
6	rs2798641 (109374743), *ARMC2*(intron)	T	FEV_1/FVC	-0.041	0.007	8.35E-09	0.007	0.013	5.69E-01	-0.036	0.025	1.39E-01
10	rs7068966 (12317998), *CDC123*(intron)	T	FEV_1/FVC	0.033	0.005	6.13E-13	0.001	0.01	9.22E-01	0.02	0.018	2.74E-01
10	rs11001819 (77985230), *C10orf11*(intron)	G	FEV_1	-0.029	0.004	2.98E-12	-0.011	0.01	2.70E-01	0.015	0.018	4.05E-01
12	rs11172113 (55813550), *LRP1*(intron)	T	FEV_1/FVC	-0.032	0.006	1.24E-08	-0.006	0.01	5.41E-01	0.022	0.019	2.44E-01
12	rs1036429 (94795559), *CCDC38*(intron)	T	FEV_1/FVC	0.038	0.006	2.30E-11	0.006	0.012	6.35E-01	0.028	0.021	1.85E-01
16	rs12447804 (56632783), *MMP15*(intron)	T	FEV_1/FVC	-0.038	0.007	3.59E-08	-0.005	0.013	6.69E-01	-0.005	0.022	8.06E-01
16	rs2865531 (73947817), *CFDP1*(intron)	T	FEV_1/FVC	0.031	0.005	1.77E-11	0.019	0.01	5.30E-02	0.002	0.018	9.13E-01
21	rs9978142 (34574109), *KCNE2*(upstream)	T	FEV_1/FVC	-0.043	0.008	2.65E-08	0.007	0.013	5.84E-01	0.007	0.024	7.59E-01
Previously reported loci												
2	rs2571445(218391399), *TNS1*(ns)	G	FEV_1	0.047	0.007	9.83E-11	-0.011	0.011	3.09E-01	0.023	0.02	2.47E-01
2	rs10498230(229210747), *PID1*(downstream)	T	FEV_1/FVC	0.068	0.014	1.13E-06	-0.061	0.019	**1.62E-03**	-0.05	0.033	1.38E-01
4	rs2045517(90089987), *FAM13A* (intron)	T	FEV_1/FVC	-0.047	0.007	2.00E-11	-0.01	0.01	3.23E-01	0.004	0.018	8.40E-01
4	rs7671167(90103002), *FAM13A* (intron)	T	FEV_1/FVC	-0.042	0.007	1.27E-09	-0.013	0.01	1.91E-01	0.014	0.018	4.35E-01
4	rs10516526(106908353), *GSTCD*(intron)	G	FEV_1	0.108	0.014	4.75E-14	0.03	0.02	1.23E-01	0.018	0.034	6.08E-01
4	rs17331332(107027556), *NPNT* (upstream)	G	FEV_1	-0.102	0.014	1.11E-12	-0.037	0.019	5.25E-02	-0.031	0.033	3.51E-01
4	rs6823809(107048244), *NPNT*(intron)	T	FEV_1/FVC	0.056	0.011	2.20E-07	0.011	0.026	6.77E-01	-0.045	0.066	5.02E-01
4	rs1032296(145654138), *HHIP*(upstream)	T	FEV_1/FVC	-0.05	0.007	3.42E-12	0.007	0.01	4.97E-01	-0.002	0.018	9.01E-01
4	rs11100860(145698589), *HHIP*(upstream)	G	FEV_1/FVC	0.064	0.007	6.81E-20	-0.002	0.01	8.58E-01	0.001	0.017	9.32E-01

Nature Genetics: doi:10.1038/ng.941

				Lung function (Stage 1 and Stage 2 meta-analysis)			Cigarettes per day			Ever vs. Never smoking		
Chr.	**SNP ID(NCBI36 position), function**	**Coded allele**	**Measure**	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***
5	rs11168048(147822546), *HTR4*(intron)	T	FEV_1/FVC	-0.047	0.007	5.97E-11	0.003	0.01	7.34E-01	-0.012	0.018	5.04E-01
5	rs3995090(147826008), *HTR4*(intron)	C	FEV_1/FVC	0.046	0.007	1.04E-10	-0.002	0.01	8.69E-01	0.014	0.018	4.25E-01
5	rs1985524(147827981), *HTR4*(intron)	G	FEV_1	-0.048	0.007	3.06E-11	0	0.01	9.97E-01	-0.027	0.018	1.32E-01
5	rs11134779(156869344), *ADAM19*(intron)	G	FEV_1/FVC	-0.042	0.007	6.01E-09	-0.015	0.01	1.43E-01	0.012	0.019	5.27E-01
6	rs2070600(32259421), *AGER*(ns)	T	FEV_1/FVC	0.126	0.016	9.07E-15	0.033	0.026	1.98E-01	0.056	0.044	2.03E-01
6	rs3817928(142792209), *GPR126*(intron)	G	FEV_1/FVC	0.059	0.008	2.27E-12	-0.004	0.012	7.16E-01	0.005	0.021	8.10E-01
6	rs262129(142894837), *LOC153910*(unknown)	G	FEV_1/FVC	0.056	0.008	2.91E-13	0.005	0.011	6.56E-01	0.019	0.02	3.39E-01
9	rs16909859(97244613), *PTCH1*(downstream)	G	FEV_1/FVC	0.08	0.013	7.45E-10	-0.005	0.02	7.81E-01	-0.012	0.033	7.09E-01
9	rs16909898(97270829), *PTCH1*(intron)	G	FEV_1/FVC	-0.072	0.012	3.94E-09	0.024	0.018	1.91E-01	0.019	0.031	5.45E-01
15	rs12899618(69432174), *THSD4*(intron)	G	FEV_1/FVC	0.076	0.01	1.86E-15	-0.013	0.014	3.36E-01	-0.012	0.024	6.31E-01
15	rs8033889(69467134), *THSD4*(intron)	T	FEV_1/FVC	-0.072	0.008	2.03E-17	-0.004	0.012	7.57E-01	-0.02	0.022	3.59E-01
15	rs2568494(76528019), *IREB2*(intron)	G	FEV_1/FVC	0.029	0.007	5.25E-05	-0.082	0.01	**2.15E-15**	-0.016	0.018	3.62E-01
15	rs8034191(76593078), *CHRNA3/5*(intron)	T	FEV_1/FVC	0.032	0.007	9.65E-06	-0.09	0.011	**1.59E-17**	-0.02	0.018	2.75E-01
15	rs2036527(76638670), *CHRNA5*(uptream)	G	FEV_1	0.036	0.008	2.40E-06	-0.091	0.011	**6.34E-18**	-0.016	0.018	3.86E-01
15	rs8040868(76698236), *CHRNA3*(coding-synon)	T	FEV_1/FVC	0.04	0.008	1.14E-06	-0.09	0.011	**2.53E-17**	-0.023	0.019	2.26E-01

Nature Genetics: doi:10.1038/ng.941

C) Relation of16 novel lung function loci, and previously reported loci, with height. Effects of the 16 novel SNPs, and previously reported SNPs, associated with lung function on height were looked up in the GIANT dataset[90]. Directions of effect for height have been flipped to refer to the same coded allele as lung function. Both effect sizes for lung and for height can be interpreted as proportion of a standard deviation. *P* values still significant after applying a Bonferroni correction for 16 tests (*P*<0.003125) are shown in bold. Results for the SpiroMeta-CHARGE Stage 1 and Stage 2 meta-analysis for the lung function measure that showed stronger association are reported for the novel loci. Results from the GWAS Stage 1 for the lung function measure that showed stronger association are reported for the previously reported loci. A column indicating whether the effect of each SNP on height is on the same direction as the effect on lung is included. Abbreviations: ns=nonsynonymous, s= synonymous.

Nature Genetics: doi:10.1038/ng.941

Chr.	SNP ID(NCBI36 position), function	Coded allele	Measure	Lung function (Stage 1+ Stage 2 meta-analysis)			Height (GIANT consortium)			Relationship of direction of effect on height to direction of effect on lung function
				Beta	Se	*P*	Beta	Se	*P*	
Novel loci										
1	rs2284746 (17179262), *MFAP2*(intron)	G	FEV_1/FVC	-0.04	0.005	7.50E-16	0.0354	0.0045	**5.638E-15**	different
1	rs993925 (216926691), *TGFB2*(downstream)	T	FEV_1/FVC	0.034	0.006	1.16E-08	0.0105	0.005	3.605E-02	same
2	rs12477314 (239542085), *HDAC4*(downstream)	T	FEV_1/FVC	0.041	0.006	1.68E-12	-0.0029	0.0057	6.124E-01	different
3	rs1529672 (25495586), *RARB*(intron)	C	FEV_1/FVC	-0.048	0.006	3.97E-14	0.0012	0.0063	8.494E-01	different
3	rs1344555 (170782913), *MECOM*(intron)	T	FEV_1	-0.034	0.006	2.65E-08	-0.0145	0.0056	9.681E-03	same
5	rs153916 (95062456), *SPATA9*(upstream)	T	FEV_1/FVC	-0.031	0.005	2.12E-08	0.0027	0.0045	5.513E-01	different
6	rs6903823 (28430275), *ZKSCAN3*(intron)/*ZNF323*(intron)	G	FEV_1	-0.037	0.006	2.18E-10	-0.0017	0.0056	7.617E-01	same
6	rs2857595 (31676448), *NCR3*(upstream)	G	FEV_1/FVC	0.037	0.006	2.28E-10	-0.0148	0.006	1.306E-02	different
6	rs2798641 (109374743), *ARMC2*(intron)	T	FEV_1/FVC	-0.041	0.007	8.35E-09	-0.0042	0.0058	4.723E-01	same
10	rs7068966 (12317998), *CDC123*(intron)	T	FEV_1/FVC	0.033	0.005	6.13E-13	0.0078	0.0045	8.521E-02	same
10	rs11001819 (77985230), *C10orf11*(intron)	G	FEV_1	-0.029	0.004	2.98E-12	0.0024	0.0045	5.964E-01	different
12	rs11172113 (55813550), *LRP1*(intron)	T	FEV_1/FVC	-0.032	0.006	1.24E-08	0.003	0.0047	5.189E-01	different
12	rs1036429 (94795559), *CCDC38*(intron)	T	FEV_1/FVC	0.038	0.006	2.30E-11	-0.0053	0.0056	3.444E-01	different
16	rs12447804 (56632783), *MMP15*(intron)	T	FEV_1/FVC	-0.038	0.007	3.59E-08	0.0077	0.0075	3.054E-01	different
16	rs2865531 (73947817), *CFDP1*(intron)	T	FEV_1/FVC	0.031	0.005	1.77E-11	-0.0129	0.0045	4.418E-03	different
21	rs9978142 (34574109), *KCNE2*(upstream)	T	FEV_1/FVC	-0.043	0.008	2.65E-08	-0.0122	0.0062	4.914E-02	same
Previously reported loci										
2	rs2571445 (218391399), *TNS1*(ns)	G	FEV_1	0.047	0.007	9.83E-11	-0.0032	0.0047	4.914E-01	different
2	rs10498230 (229210747), *PID1*(downstream)	T	FEV_1/FVC	0.068	0.014	1.13E-06	-0.0111	0.0087	2.023E-01	different
4	rs2045517 (90089987), *FAM13A*(intron)	T	FEV_1/FVC	-0.047	0.007	2.00E-11	0.0058	0.0045	2.006E-01	different
4	rs7671167 (90103002), *FAM13A*(intron)	T	FEV_1/FVC	-0.042	0.007	1.27E-09	0.0072	0.0045	1.121E-01	different
4	rs10516526 (106908353), *GSTCD*(intron)	G	FEV_1	0.108	0.014	4.75E-14	0.0032	0.0092	7.275E-01	same
4	rs17331332 (107027556), *NPNT*(upstream)	G	FEV_1	-0.102	0.014	1.11E-12	-0.002	0.0093	8.298E-01	same
4	rs6823809 (107048244), *NPNT*(intron)	T	FEV_1/FVC	0.056	0.011	2.20E-07	-0.0001	0.0072	9.888E-01	different
4	rs1032296 (145654138), *HHIP*(upstream)	T	FEV_1/FVC	-0.05	0.007	3.42E-12	-0.0152	0.0047	1.082E-03	same
4	rs11100860 (145698589), *HHIP*(upstream)	G	FEV_1/FVC	0.064	0.007	6.81E-20	0.0151	0.0045	8.618E-04	same

Nature Genetics: doi:10.1038/ng.941

				Lung function (Stage 1+ Stage 2 meta-analysis)			Height (GIANT consortium)			Relationship of direction of effect on height to direction of effect on lung function
Chr.	**SNP ID(NCBI36 position), function**	**Coded allele**	**Measure**	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	
5	rs11168048 (147822546), *HTR4*(intron)	T	FEV_1/FVC	-0.047	0.007	5.97E-11	-0.0133	0.0049	6.524E-03	same
5	rs3995090 (147826008), *HTR4*(intron)	C	FEV_1/FVC	0.046	0.007	1.04E-10	0.0143	0.0049	3.448E-03	same
5	rs1985524 (147827981), *HTR4*(intron)	G	FEV_1	-0.048	0.007	3.06E-11	-0.015	0.0049	2.156E-03	same
5	rs11134779 (156869344), *ADAM19*(intron)	G	FEV_1/FVC	-0.042	0.007	6.01E-09	-0.0092	0.0047	4.791E-02	same
6	rs2070600 (32259421), *AGER*(ns)	T	FEV_1/FVC	0.126	0.016	9.07E-15	0.0094	0.0114	4.116E-01	same
6	rs3817928 (142792209), *GPR126*(intron)	G	FEV_1/FVC	0.059	0.008	2.27E-12	-0.0368	0.0055	1.967E-11	different
6	rs262129 (142894837), *LOC153910*(unknown)	G	FEV_1/FVC	0.056	0.008	2.91E-13	-0.0443	0.005	9.173E-19	different
9	rs16909859 (97244613), *PTCH*1(downstream)	G	FEV_1/FVC	0.08	0.013	7.45E-10	-0.0266	0.0082	1.226E-03	different
9	rs16909898 (97270829), *PTCH1*(intron)	G	FEV_1/FVC	-0.072	0.012	3.94E-09	0.0313	0.0078	5.392E-05	different
15	rs12899618 (69432174), *THSD4*(intron)	G	FEV_1/FVC	0.076	0.01	1.86E-15	-0.0075	0.0061	2.175E-01	different
15	rs8033889 (69467134), *THSD4*(intron)	T	FEV_1/FVC	-0.072	0.008	2.03E-17	0.0015	0.0054	7.798E-01	different
15	rs2568494 (76528019), *IREB2*(intron)	G	FEV_1/FVC	0.029	0.007	5.25E-05	0.0021	0.0047	6.516E-01	same
15	rs8034191 (76593078), *CHRNA3/5*(intron)	T	FEV_1/FVC	0.032	0.007	9.65E-06	-0.0002	0.0047	9.657E-01	different
15	rs2036527 (76638670), *CHRNA5*(uptream)	G	FEV_1	0.036	0.008	2.40E-06	-0.0005	0.0048	9.165E-01	different
15	rs8040868 (76698236), *CHRNA3*(s)	T	FEV_1/FVC	0.04	0.008	1.14E-06	-0.006	0.0064	3.515E-01	different

Nature Genetics: doi:10.1038/ng.941

D) Relation of 16 novel lung function loci, and previously reported loci, with lung cancer. Effects of the 16 novel SNPs, and previously reported SNPs, associated with lung function on lung cancer were assessed in the International Lung Cancer Consortium (ILCCO) GWAS meta-analysis[37]. Directions of effect for lung cancer have been flipped to refer to the same coded allele as lung function. *P* values still significant after applying a Bonferroni correction for 16 tests (*P*<0.003125) are shown in bold. The ILCCO GWAS meta-analysis only had genotyped data, for this reason proxy SNPs were given when the top SNP was not included in their data. Leading SNP, region name and r^2 between leading SNP and proxy SNP are also provided. Results for the SpiroMeta-CHARGE Stage 1 and Stage 2 meta-analysis are reported for the SNPs that were followed up in Stage 2 and were included in the lung cancer dataset (rs1529672, rs2857595, rs2798641, rs11001819, rs11172113, rs1036429) and results from the GWAS Stage 1 only are provided for the other loci, for the lung function measure that showed stronger association.

Nature Genetics: doi:10.1038/ng.941

Chr.	Proxy SNP ID (NCBI36 position), function	Leading SNP (region name), r^2 with proxy	Coded allele	Lung function				Lung cancer		
				Measure	Beta	Se	*P*	Beta	Se	*P*
Novel loci										
1	rs761423 (17174259), *MFAP2* (intron)	rs2284746 (*MFAP2*), 0.63	T	FEV_1/FVC	0.038	0.007	8.72E-08	0.033	0.017	6.24E-02
1	rs2871775 (17218492), *SDHB* (intron)	rs2284746 (*MFAP2*), 0.66	G	FEV_1/FVC	-0.040	0.007	9.08E-09	-0.015	0.017	3.81E-01
2	rs4591362 (239542675), *HDAC4*(downstream)	rs12477314 (*HDAC4*), 0.94	G	FEV_1/FVC	-0.049	0.009	8.29E-09	-0.020	0.021	3.49E-01
3	rs1529672 (25495586), *RARB* (intron)	rs1529672 (*RARB*), 1	C	FEV_1/FVC	-0.048	0.006	3.97E-14	-0.013	0.027	6.33E-01
3	rs2056777 (25515395), *RARB* (intron)	rs1529672 (*RARB*), 0.77	T	FEV_1/FVC	-0.050	0.009	5.31E-08	-0.011	0.027	6.80E-01
3	rs1362772 (170739927), *MECOM*(intron)	rs1344555 (*MECOM*), 1	T	FEV_1	-0.040	0.009	3.24E-06	-0.010	0.022	6.63E-01
3	rs7642776 (170753972), *MECOM*(intron)	rs1344555 (*MECOM*), 0.94	G	FEV_1	0.038	0.008	5.38E-06	0.013	0.021	5.37E-01
5	rs2548125 (95037182), *SPATA9* (intron)	rs153916 (*SPATA9*), 0.61	G	FEV_1/FVC	-0.029	0.007	3.50E-05	-0.012	0.018	5.08E-01
6	rs209181 (28900456), *LOC401242* (downstream)	rs6903823 (*ZKSCAN3/ZNF323*), 0.69	G	FEV_1	0.035	0.012	2.25E-03	-0.106	0.026	**3.41E-05**
6	rs3099844 (31556955), *HCG26* (downstream)	rs2857595 (*NCR3*), 0.67	C	FEV_1	0.058	0.011	1.92E-07	-0.141	0.027	**2.21E-07**
6	rs2857595 (31676448), *NCR3*(upstream)	rs2857595 (*NCR3*), 1	G	FEV_1/FVC	0.037	0.006	2.28E-10	-0.051	0.022	1.91E-02
6	rs1475055 (109350925), *ARMC2* (intron)	rs2798641 (*ARMC2*), 0.73	T	FEV_1/FVC	0.027	0.008	7.67E-04	-0.011	0.020	5.87E-01
6	rs2798641 (109374743), *ARMC2* (intron)	rs2798641 (*ARMC2*), 1	T	FEV_1/FVC	-0.041	0.007	8.35E-09	0.006	0.022	7.72E-01
10	rs1317549 (12285320), *CDC123* (intron)	rs7068966 (*CDC123*), 0.68	T	FEV_1/FVC	-0.038	0.007	8.95E-08	0.013	0.018	4.78E-01
10	rs4478891 (12307660), *CDC123* (intron)	rs7068966 (*CDC123*), 0.85	G	FEV_1/FVC	0.043	0.007	8.71E-10	-0.005	0.018	7.61E-01
10	rs2130800 (77944824), *C10orf11* (intron)	rs11001819 (*C10orf11*), 0.73	T	FEV_1	0.038	0.007	5.45E-08	0.031	0.017	7.80E-02
10	rs11001819 (77985230), *C10orf11* (intron)	rs11001819 (*C10orf11*), 1	G	FEV_1	-0.029	0.004	2.98E-12	-0.051	0.020	1.21E-02
10	rs2637260 (77990352), *C10orf11* (downstream)	rs11001819 (*C10orf11*), 0.72	T	FEV_1	0.035	0.007	7.38E-07	0.024	0.017	1.72E-01
12	rs11172113 (55813550), *LRP1* (intron)	rs11172113 (*LRP1*), 1	T	FEV_1/FVC	-0.032	0.006	1.24E-08	-0.010	0.018	5.84E-01
12	rs1466535 (55820737), *LRP1* (intron)	rs11172113 (*LRP1*), 0.72	G	FEV_1/FVC	-0.025	0.007	5.76E-04	0.004	0.018	8.45E-01
12	rs7307510 (94761701), *SNRPF* (upstream)	rs1036429 (*CCDC38*), 0.96	T	FEV_1/FVC	0.049	0.009	3.49E-08	0.009	0.023	6.82E-01
12	rs1036429 (94795559), *CCDC38* (intron)	rs1036429 (*CCDC38*), 1	T	FEV_1/FVC	0.038	0.006	2.30E-11	0.011	0.022	6.06E-01
16	rs2304488 (56631711), *MMP15* (intron)	rs12447804 (*MMP15*), 0.88	G	FEV_1/FVC	-0.040	0.008	9.45E-07	0.019	0.021	3.62E-01
16	rs12597233 (56657709), *MMP15* (downstream)	rs12447804 (*MMP15*), 0.87	G	FEV_1/FVC	0.038	0.008	4.80E-06	-0.012	0.022	5.85E-01
16	rs4243111 (73878328), *CFDP1*(downstream)	rs2865531 (*CFDP1*), 0.93	T	FEV_1/FVC	-0.037	0.007	1.52E-07	-0.036	0.018	4.22E-02
16	rs1424013 (74053487), *TMEM170* (intron)	rs2865531 (*CFDP1*), 0.82	T	FEV_1/FVC	0.033	0.007	3.92E-06	0.046	0.018	1.00E-02
21	rs973754 (34555400), *C21orf82* (downstream)	rs9978142 (*KCNE2*), 0.81	G	FEV_1/FVC	-0.043	0.010	2.09E-05	-0.010	0.025	6.78E-01
Previously reported regions										
2	rs1035672 (218383444), *TNS1* (intron)	rs2571445 (*TNS1*), 0.96	G	FEV_1	-0.046	0.007	3.03E-10	-0.020	0.018	2.66E-01

Nature Genetics: doi:10.1038/ng.941

Chr.	Proxy SNP ID (NCBI36 position), function	Leading SNP (region name), r^2 with proxy	Coded allele	Lung function				Lung cancer		
				Measure	Beta	Se	*P*	Beta	Se	*P*
2	rs2571445 (218391399), *TNS1* (non synonymous)	rs2571445 (*TNS1*), 1	G	FEV_1	0.047	0.007	9.83E-11	0.016	0.018	3.72E-01
4	rs2869967 (90088355), *FAM13A* (intron)	rs2045517 (*FAM13A*), 1	T	FEV_1/FVC	0.047	0.007	2.08E-11	0.026	0.018	1.36E-01
4	rs6849143 (90147512), *FAM13A* (intron)	rs2045517 (*FAM13A*), 0.75	T	FEV_1/FVC	-0.038	0.007	5.99E-08	-0.020	0.018	2.51E-01
4	rs11727735 (106851319), *GSTCD* (intron)	rs10516526 (*GSTCD*), 1	G	FEV_1	0.105	0.014	1.65E-13	-0.029	0.036	4.19E-01
4	rs10516526 (106908353), *GSTCD* (intron)	rs10516526 (*GSTCD*), 1	G	FEV_1	0.108	0.014	4.75E-14	-0.029	0.036	4.31E-01
4	rs1828591 (145700230), *HHIP* (upstream)	rs11100860 (*HHIP*), 1	G	FEV_1/FVC	0.063	0.007	1.44E-19	-0.005	0.018	7.82E-01
4	rs1512288 (145710731), *HHIP* (upstream)	rs11100860 (*HHIP*), 1	G	FEV_1/FVC	-0.062	0.007	3.46E-19	0.000	0.018	9.89E-01
5	rs2277027 (156864954), *ADAM19* (intron)	rs11134779 (*ADAM19*), 1	C	FEV_1/FVC	-0.042	0.007	6.65E-09	0.007	0.018	6.93E-01
5	rs1422795 (156868942), *ADAM19* (non synonymous)	rs11134779 (*ADAM19*), 1	T	FEV_1/FVC	0.041	0.007	1.05E-08	-0.007	0.018	7.11E-01
6	rs2070600 (32259421), *AGER* (non synonymous)	rs2070600 (*AGER*), 1	T	FEV_1/FVC	0.126	0.016	9.07E-15	-0.004	0.044	9.27E-01
6	rs2854050 (32293583), *NOTCH4* (intron)	rs2070600 (*AGER*), 1	G	FEV_1/FVC	-0.083	0.016	9.34E-08	0.022	0.040	5.80E-01
6	rs6570507 (142721265), *GPR126* (intron)	rs262129 (*GPR126*), 0.72	G	FEV_1/FVC	-0.051	0.008	2.25E-11	0.006	0.019	7.49E-01
6	rs11155242 (142733242), *GPR126* (non synonymous)	rs3817928 (*GPR126*), 1	C	FEV_1/FVC	0.055	0.009	1.88E-10	0.003	0.022	8.99E-01
6	rs3748069 (142809326), *GPR126* (downstream)	rs262129 (*GPR126*), 0.84	G	FEV_1/FVC	0.053	0.008	2.02E-12	-0.007	0.019	7.01E-01
6	rs7776356 (142818722), *GPR126* (downstream)	rs3817928 (*GPR126*), 1	G	FEV_1/FVC	0.059	0.008	4.16E-12	-0.001	0.021	9.69E-01
9	rs10512249 (97296130), *PTCH1* (intron)	rs16909859 (*PTCH1*), 0.84	G	FEV_1/FVC	0.066	0.012	1.54E-08	-0.067	0.029	1.90E-02
15	rs1913768 (69436598), *THSD4* (intron)	rs8033889 (*THSD4*), 0.673	G	FEV_1/FVC	0.075	0.009	2.77E-15	-0.014	0.024	5.49E-01
15	rs8033889 (69467134), *THSD4* (intron)	rs8033889 (*THSD4*), 1	T	FEV_1/FVC	-0.072	0.008	2.03E-17	0.014	0.025	5.85E-01
15	rs8034191 (76593078), *AGPHD1* (intron)	rs8040868 (*CHRNA3*), 0.70	T	FEV_1/FVC	0.032	0.007	9.65E-06	-0.258	0.018	2.19E-46
15	rs1051730 (76681394), *CHRNA3* (synonymous)	rs8040868 (*CHRNA3*), 0.76	G	FEV_1/FVC	0.032	0.007	1.46E-05	-0.273	0.018	1.91E-51

Nature Genetics: doi:10.1038/ng.941

E) Association of published height SNPs with FEV_1 and FEV_1/FVC. Effects sizes and *P* values for the 180 loci that reached genome-wide significance ($P < 5x10^{-8}$) in the combined Stage1 and 2 analysis of height reported by the GIANT consortium[90] are shown, together with the effect sizes, standard errors and *P* values for FEV_1 and FEV_1/FVC in the SpiroMeta-CHARGE GWAS dataset. Directions of effect for height have been flipped to refer to the same coded allele as lung function. Both effect sizes for lung and for height can be interpreted as proportion of a standard deviation. *P* values still significant after applying a Bonferroni correction for 180 tests ($P < 2.8x10^{-4}$) are shown in bold. Columns indicating whether the effect of each SNP on height is on the same direction as the effect on lung are included.

Nature Genetics: doi:10.1038/ng.941

Chr.	SNP ID(NCBI36 position), Nearest_OMIM_height_gene	Coded allele	Height		FEV_1				FEV_1/FVC			
			Beta	*P*	Beta	Se	*P*	Direction of effect for FEV_1 relative to height	Beta	Se	*P*	Direction of effect for FEV_1/FVC relative to height
1	rs425277(2059032), *PRKCZ*	T	0.022	2.10E-08	0.004	0.008	5.88E-01	same	-0.007	0.008	3.68E-01	different
1	rs2284746(17179262), *MFAP2*	G	0.04	3.90E-29	0.008	0.007	2.78E-01	same	-0.042	0.007	**2.47E-09**	different
1	rs1738475(23409478), *HTR1D*	G	-0.025	3.00E-12	-0.002	0.007	7.73E-01	same	-0.004	0.007	5.76E-01	same
1	rs4601530(24916698), *CLIC4*	T	-0.028	2.20E-12	0.03	0.008	**1.47E-04**	different	0.019	0.008	1.32E-02	different
1	rs7532866(26614131), *LIN28*	G	-0.021	3.40E-08	-0.017	0.008	2.19E-02	same	-0.001	0.007	8.94E-01	same
1	rs2154319(41518357), *SCMH1*	T	-0.03	1.80E-12	-0.008	0.009	3.57E-01	same	-0.003	0.008	6.80E-01	same
1	rs17391694(78396214), *GIPC2*	T	0.042	1.70E-11	0.013	0.013	3.16E-01	same	0.007	0.012	5.90E-01	same
1	rs6699417(88896031), *PKN2*	T	0.021	5.00E-09	0.001	0.007	8.44E-01	same	-0.011	0.007	1.04E-01	different
1	rs10874746(93096559), *RPL5*	T	-0.024	6.70E-11	0.005	0.007	4.69E-01	different	0.015	0.007	3.10E-02	different
1	rs9428104(118657110), *SPAG17*	G	0.041	5.60E-23	-0.013	0.008	1.26E-01	different	0.015	0.008	6.91E-02	same
1	rs11205277(148159496), *SF3B4*	G	0.046	4.80E-32	-0.021	0.008	9.16E-03	different	-0.011	0.008	1.76E-01	different
1	rs17346452(170319910), *DNM3*	T	-0.04	1.40E-23	-0.014	0.008	7.77E-02	same	0.002	0.008	8.09E-01	different
1	rs1325598(175058872), *PAPPA2*	G	0.022	1.10E-09	-0.008	0.007	2.76E-01	different	-0.003	0.007	6.59E-01	different
1	rs1046934(182290152), *TSEN15*	C	0.044	2.10E-31	-0.001	0.007	8.97E-01	different	0.01	0.007	1.75E-01	same
1	rs10863936(210304421), *DTL*	G	0.021	1.90E-09	-0.003	0.007	6.47E-01	different	-0.011	0.007	1.11E-01	different
1	rs6684205(216676325), *TGFB2*	G	0.028	1.50E-12	0.002	0.008	7.94E-01	same	-0.007	0.008	3.75E-01	different
1	rs11118346(217810342), *LYPLAL1*	T	-0.025	1.90E-12	0.008	0.007	2.61E-01	different	0.004	0.007	5.67E-01	different
1	rs10799445(225978506), *JMJD4*	C	-0.032	2.40E-13	0.001	0.008	9.45E-01	different	-0.005	0.008	5.67E-01	same
2	rs4665736(25041103), *DNAJC27*	T	0.029	7.30E-16	-0.008	0.007	2.67E-01	different	-0.012	0.007	9.28E-02	different
2	rs6714546(33214929), *LTBP1*	G	0.026	1.60E-09	0.004	0.008	6.48E-01	same	0.002	0.008	8.24E-01	same
2	rs17511102(37814117), *CDC42EP3*	T	0.06	1.60E-18	0.004	0.014	7.88E-01	same	0.015	0.013	2.69E-01	same
2	rs2341459(44621706), *C2orf34*	T	0.025	7.90E-10	-0.006	0.008	4.53E-01	different	-0.003	0.008	7.13E-01	different
2	rs12474201(46774789), *SOCS5*	G	-0.028	2.60E-13	0.001	0.008	9.37E-01	different	0.003	0.007	6.68E-01	different
2	rs3791675(55964813), *EFEMP1*	T	-0.053	2.50E-35	-0.02	0.008	1.28E-02	same	0.025	0.008	1.88E-03	different
2	rs11684404(88705737), *EIF2AK3*	T	-0.028	9.90E-14	0.02	0.007	7.65E-03	different	0.002	0.007	8.34E-01	different
2	rs7567288(134151294), *NCKAP5*	T	-0.032	2.10E-12	0.008	0.009	3.75E-01	different	0.016	0.009	7.32E-02	different
2	rs7567851(178392966), *PDE11A*	G	-0.037	3.30E-08	-0.009	0.013	4.86E-01	same	0.019	0.013	1.47E-01	different

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	SNP ID(NCBI36 position), Nearest_OMIM_height_gene	Coded allele	Height		FEV_1				FEV_1/FVC			
			Beta	*P*	Beta	Se	*P*	Direction of effect for FEV_1 relative to height	Beta	Se	*P*	Direction of effect for FEV_1/FVC relative to height
2	rs1351164(217980143), *TNS1*	T	0.034	2.10E-14	-0.017	0.009	5.42E-02	different	-0.015	0.009	8.06E-02	different
2	rs12470505(219616613), *CCDC108/IHH*	T	0.041	8.90E-12	0.007	0.012	5.81E-01	same	-0.009	0.012	4.59E-01	different
2	rs2629046(224755988), *SERPINE2*	T	0.024	7.90E-12	-0.009	0.007	1.96E-01	different	-0.006	0.007	4.08E-01	different
2	rs2580816(232506210), *NPPC*	T	-0.045	5.80E-22	0.007	0.009	4.49E-01	different	-0.007	0.009	4.75E-01	same
2	rs12694997(241911659), *SEPT2*	G	0.024	1.20E-08	0.004	0.008	5.94E-01	same	-0.001	0.008	8.59E-01	different
3	rs2597513(13530836), *HDAC11*	T	-0.036	7.40E-10	0.014	0.012	2.45E-01	different	0.025	0.011	2.66E-02	different
3	rs13088462(51046753), *DOCK3*	T	-0.052	3.80E-10	-0.011	0.016	5.18E-01	same	-0.026	0.016	1.07E-01	same
3	rs2336725(53093779), *RTF1*	T	-0.027	9.70E-13	0.002	0.007	7.52E-01	different	0.009	0.007	2.41E-01	different
3	rs9835332(56642722), *C3orf63*	G	0.026	5.30E-13	0.005	0.007	4.41E-01	same	0.006	0.007	3.99E-01	same
3	rs17806888(67499012), *SUCLG2*	T	0.036	2.10E-09	-0.018	0.011	1.16E-01	different	-0.02	0.011	8.55E-02	different
3	rs9863706(72520103), *RYBP*	T	-0.031	4.10E-13	-0.011	0.009	2.00E-01	same	0.002	0.008	8.02E-01	different
3	rs6439167(130533446), *C3orf47*	T	-0.034	8.90E-15	-0.014	0.009	9.99E-02	same	0	0.008	9.68E-01	different
3	rs9844666(137456906), *PCCB*	G	0.024	3.50E-09	-0.01	0.008	2.04E-01	different	-0.001	0.008	9.08E-01	different
3	rs724016(142588260), *ZBTB38*	G	0.07	3.10E-86	-0.003	0.007	6.26E-01	different	0.002	0.007	7.52E-01	same
3	rs572169(173648421), *GHSR*	T	0.033	2.80E-18	0.011	0.008	1.64E-01	same	0.008	0.007	2.73E-01	same
3	rs720390(187031377), *IGF2BP2*	G	-0.029	1.90E-14	0.004	0.007	5.75E-01	different	0.013	0.007	7.13E-02	different
4	rs2247341(1671115), *SLBP/FGFR3*	G	-0.025	1.50E-11	0.017	0.007	1.99E-02	different	0.01	0.007	1.88E-01	different
4	rs6449353(17642586), *LCORL*	T	0.075	7.10E-46	0.013	0.01	2.07E-01	same	0.003	0.01	7.57E-01	same
4	rs17081935(57518233), *POLR2B*	T	0.03	3.70E-11	0.007	0.009	4.07E-01	same	0.015	0.009	7.98E-02	same
4	rs7697556(73734177), *ADAMTS3*	T	0.028	2.00E-14	-0.007	0.007	3.36E-01	different	0.007	0.007	3.32E-01	same
4	rs788867(82369030), *PRKG2/BMP3*	T	-0.043	8.90E-28	0.005	0.008	4.86E-01	different	0.002	0.008	8.03E-01	different
4	rs10010325(106325802), *TET2*	C	-0.024	3.90E-11	-0.009	0.007	2.16E-01	same	-0.014	0.007	4.05E-02	same
4	rs7689420(145787802), *HHIP*	T	-0.073	6.20E-51	-0.008	0.009	4.21E-01	same	-0.008	0.009	3.84E-01	same
4	rs955748(184452669), *WWC2*	G	0.023	4.40E-08	-0.003	0.008	7.19E-01	different	-0.011	0.008	1.54E-01	different
5	rs1173727(32866278), *NPR3*	T	0.034	1.60E-21	-0.009	0.007	1.82E-01	different	-0.002	0.007	7.23E-01	different
5	rs11958779(55037656), *SLC38A9*	G	0.027	1.80E-12	0.005	0.008	4.73E-01	same	-0.008	0.007	2.65E-01	different
5	rs10037512(88390431), *MEF2C*	T	0.032	2.00E-18	0.002	0.007	7.77E-01	same	0.003	0.007	7.19E-01	same

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	SNP ID(NCBI36 position), Nearest_OMIM_height_gene	Coded allele	Height		FEV_1				FEV_1/FVC			
			Beta	*P*	Beta	Se	*P*	Direction of effect for FEV_1 relative to height	Beta	Se	*P*	Direction of effect for FEV_1/FVC relative to height
5	rs13177718(108141243), *FER*	T	-0.04	3.00E-08	0.002	0.014	8.90E-01	different	0.021	0.014	1.33E-01	different
5	rs1582931(122685098), *CEP120*	G	0.023	1.50E-10	0.005	0.007	4.54E-01	same	0.002	0.007	7.95E-01	same
5	rs274546(131727766), *SLC22A5*	G	0.029	7.30E-16	-0.007	0.007	3.13E-01	different	0.001	0.007	8.87E-01	same
5	rs526896(134384604), *PITX1*	T	0.03	2.30E-13	0	0.008	9.51E-01	different	0.007	0.008	3.83E-01	same
5	rs4282339(168188818), *SLIT3*	G	0.036	6.60E-16	0.003	0.009	7.55E-01	same	0.006	0.009	4.97E-01	same
5	rs12153391(171136043), *FBXW11*	C	0.03	3.60E-12	-0.001	0.008	8.96E-01	different	-0.007	0.008	3.99E-01	different
5	rs889014(172916720), *BOD1*	T	-0.03	9.40E-16	0.008	0.007	2.83E-01	different	0.004	0.007	5.72E-01	different
5	rs422421(176449932), *FGFR4/NSD1*	T	-0.031	1.10E-12	0.005	0.009	5.76E-01	different	0.018	0.008	3.90E-02	different
5	rs6879260(179663620), *GFPT2*	T	-0.022	1.60E-09	0.003	0.007	7.20E-01	different	-0.014	0.007	5.38E-02	same
6	rs3812163(7670759), *BMP6*	T	0.036	1.20E-23	-0.026	0.007	**2.14E-04**	different	-0.001	0.007	9.33E-01	different
6	rs1047014(19949472), *ID4*	T	-0.032	1.80E-13	0.005	0.009	5.44E-01	different	-0.011	0.008	1.78E-01	same
6	rs806794(26308656), *Histone cluster*	G	-0.052	1.20E-39	0.006	0.008	4.67E-01	different	-0.013	0.008	9.67E-02	same
6	rs3129109(29192211), *OR2J3*	T	-0.032	2.40E-17	-0.01	0.007	1.61E-01	same	0.001	0.007	8.49E-01	different
6	rs2256183(31488508), *MICA*	G	-0.04	7.80E-29	0.007	0.007	3.11E-01	different	0.01	0.007	1.46E-01	different
6	rs6457620(32771977), *HLA* locus	G	0.029	2.10E-16	-0.002	0.007	7.58E-01	different	-0.025	0.007	2.92E-04	different
6	rs2780226(34307070), *HMGA1*	T	-0.076	8.10E-28	-0.015	0.014	2.63E-01	same	0.023	0.014	8.63E-02	different
6	rs6457821(35510783), *PPARD/FANCE*	C	0.104	2.10E-12	-0.007	0.029	7.96E-01	different	0.058	0.028	3.97E-02	same
6	rs9472414(45054484), *SUPT3H/RUNX2*	T	0.026	1.80E-09	-0.002	0.008	7.97E-01	different	0.003	0.008	7.01E-01	same
6	rs9360921(76322362), *SENP6*	T	-0.042	2.60E-13	0.015	0.011	1.81E-01	different	0.011	0.011	3.36E-01	different
6	rs310405(81857081), *FAM46A*	G	-0.026	2.20E-13	-0.007	0.007	3.14E-01	same	-0.012	0.007	9.00E-02	same
6	rs7759938(105485647), *LIN28B*	T	-0.045	8.30E-31	0.016	0.008	3.93E-02	different	-0.006	0.008	4.26E-01	same
6	rs1046943(109890634), *ZBTB24*	G	-0.02	2.50E-08	-0.012	0.007	8.88E-02	same	0	0.007	9.46E-01	different
6	rs961764(117628849), *VGLL2*	G	0.024	1.30E-11	-0.003	0.007	6.42E-01	different	0.009	0.007	2.23E-01	same
6	rs1490384(126892853), *C6orf173*	T	0.034	3.90E-21	0.022	0.007	2.07E-03	same	-0.002	0.007	7.46E-01	different
6	rs6569648(130390812), *L3MBTL3*	T	-0.04	1.10E-21	-0.002	0.008	8.41E-01	same	0.013	0.008	1.13E-01	different
6	rs7763064(142838982), *GPR126*	G	0.048	1.10E-33	-0.03	0.008	**1.33E-04**	different	-0.052	0.008	**6.80E-12**	different
6	rs543650(152152636), *ESR1*	T	-0.034	1.20E-17	0.002	0.008	7.58E-01	different	0.002	0.008	8.44E-01	different

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	SNP ID(NCBI36 position), Nearest_OMIM_height_gene	Coded allele	Height		FEV_1				FEV_1/FVC			
			Beta	*P*	Beta	Se	*P*	Direction of effect for FEV_1 relative to height	Beta	Se	*P*	Direction of effect for FEV_1/FVC relative to height
6	rs9456307(158849430), *TULP4*	T	0.048	2.20E-09	0.001	0.016	9.38E-01	same	-0.007	0.016	6.76E-01	different
7	rs798489(2768329), *GNA12*	T	-0.048	1.90E-33	0.006	0.008	4.72E-01	different	-0.002	0.008	7.93E-01	same
7	rs4470914(19583047), *TWISTNB*	T	0.029	9.20E-10	-0.008	0.01	4.25E-01	different	-0.014	0.009	1.39E-01	different
7	rs12534093(23469499), *IGF2BP3*	T	0.034	2.00E-14	0.008	0.008	3.46E-01	same	-0.003	0.008	7.43E-01	different
7	rs1708299(28156471), *JAZF1*	G	-0.04	5.80E-25	0.014	0.008	6.07E-02	different	0.002	0.008	8.31E-01	different
7	rs6959212(38094851), *STARD3NL*	T	-0.024	1.60E-09	0.003	0.008	7.21E-01	different	0.002	0.008	7.68E-01	different
7	rs42235(92086012), *CDK6*	T	0.057	7.70E-47	0	0.008	9.50E-01	different	-0.007	0.008	3.52E-01	different
7	rs822552(148281567), *PDIA4*	G	0.025	2.60E-08	-0.023	0.009	8.01E-03	different	-0.003	0.009	7.28E-01	different
7	rs2110001(150147955), *TMEM176A*	G	0.031	3.30E-13	0.019	0.008	1.86E-02	same	0.007	0.008	4.03E-01	same
8	rs1013209(24172249), *ADAM28*	T	-0.025	1.60E-09	0.021	0.008	1.03E-02	different	-0.002	0.008	8.37E-01	same
8	rs7460090(57356717), *SDR16C5*	T	0.058	8.20E-27	0.012	0.011	2.52E-01	same	-0.003	0.01	8.07E-01	different
8	rs6473015(78341040), *PEX2*	C	0.029	6.90E-13	0.011	0.008	1.72E-01	same	-0.002	0.008	8.42E-01	different
8	rs6470764(130794847), *GSDMC*	T	-0.05	1.70E-28	-0.004	0.009	6.39E-01	same	-0.002	0.009	8.34E-01	same
8	rs12680655(135706519), *ZFAT*	G	-0.028	1.60E-14	0.016	0.007	2.15E-02	different	0.014	0.007	4.40E-02	different
9	rs7864648(16358732), *BNC2*	T	0.022	2.10E-08	0	0.008	9.87E-01	different	0.004	0.008	6.30E-01	same
9	rs11144688(77732106), *PCSK5*	G	0.049	9.60E-12	0.001	0.015	9.29E-01	same	0.009	0.015	5.31E-01	same
9	rs7853377(85742025), *C9orf64*	G	0.024	4.50E-08	0.007	0.009	4.29E-01	same	0.008	0.009	3.38E-01	same
9	rs8181166(88306448), *ZCCHC6*	G	-0.026	2.70E-12	0.014	0.007	5.36E-02	different	0.003	0.007	6.68E-01	different
9	rs2778031(90025546), *SPIN1*	T	0.031	9.00E-13	0.005	0.008	5.62E-01	same	0.01	0.008	2.16E-01	same
9	rs9969804(94468941), *IPPK*	C	-0.03	7.70E-17	0.003	0.007	6.58E-01	different	0.015	0.007	3.15E-02	different
9	rs1257763(95933766), *PTPDC1*	G	-0.069	9.90E-10	-0.003	0.022	8.98E-01	same	-0.005	0.022	8.13E-01	same
9	rs473902(97296056), *PTCH1/FANCC*	T	0.065	2.30E-17	-0.012	0.016	4.26E-01	different	-0.024	0.015	1.16E-01	different
9	rs7027110(108638867), *ZNF462*	G	-0.031	2.30E-13	0.024	0.008	3.80E-03	different	0.01	0.008	2.35E-01	different
9	rs1468758(112846903), *LPAR1*	T	-0.026	1.40E-09	-0.004	0.008	6.21E-01	same	-0.006	0.008	4.26E-01	same
9	rs751543(118162163), *PAPPA*	T	0.026	6.50E-10	-0.005	0.008	5.47E-01	different	-0.003	0.008	7.14E-01	different
9	rs7466269(132453905), *FUBP3*	G	-0.032	2.60E-17	-0.012	0.007	1.11E-01	same	-0.02	0.007	6.62E-03	same
9	rs7849585(138251691), *QSOX2*	T	0.029	4.70E-14	0.014	0.008	7.84E-02	same	-0.002	0.008	8.08E-01	different

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	SNP ID(NCBI36 position), Nearest_OMIM_height_gene	Coded allele	Height		FEV_1				FEV_1/FVC			
			Beta	*P*	Beta	Se	*P*	Direction of effect for FEV_1 relative to height	Beta	Se	*P*	Direction of effect for FEV_1/FVC relative to height
10	rs7909670(12958770), *CCDC3*	T	-0.021	3.20E-09	0.002	0.007	7.78E-01	different	-0.001	0.007	8.76E-01	same
10	rs2145998(80791702), *PPIF*	T	0.026	3.60E-13	0	0.007	9.72E-01	different	0.006	0.007	3.59E-01	same
10	rs11599750(101795432), *CPN1*	T	-0.028	1.60E-13	-0.01	0.007	1.55E-01	same	-0.012	0.007	9.58E-02	same
11	rs2237886(2767307), *KCNQ1*	T	0.046	2.20E-13	-0.005	0.012	7.13E-01	different	-0.008	0.012	5.31E-01	different
11	rs7926971(12654616), *TEAD1*	G	0.023	4.40E-10	-0.017	0.007	1.41E-02	different	0.006	0.007	3.59E-01	same
11	rs1330(17272605), *NUCB2*	T	0.022	4.90E-09	-0.007	0.008	3.32E-01	different	0.003	0.007	7.11E-01	same
11	rs10838801(48054856), *PTPRJ/SLC39A13*	G	0.027	3.50E-12	-0.01	0.008	1.94E-01	different	-0.008	0.008	2.87E-01	different
11	rs1814175(49515748), *FOLH1*	T	0.022	1.60E-08	-0.006	0.008	4.80E-01	different	0.002	0.008	8.07E-01	same
11	rs5017948(51270794), *OR4A5*	T	-0.027	3.10E-08	-0.002	0.01	8.63E-01	same	0	0.009	9.70E-01	different
11	rs3782089(65093395), *SSSCA1*	T	-0.058	3.60E-13	-0.019	0.016	2.45E-01	same	0.009	0.016	5.74E-01	different
11	rs7112925(66582736), *RHOD*	T	-0.023	9.00E-10	0.003	0.007	7.14E-01	different	-0.001	0.007	8.98E-01	same
11	rs634552(74959700), *SERPINH1*	T	0.039	3.50E-13	-0.024	0.01	2.15E-02	different	-0.007	0.01	4.72E-01	different
11	rs494459(118079885), *TREH*	T	0.02	1.70E-08	0.003	0.007	6.23E-01	same	-0.003	0.007	6.39E-01	different
11	rs654723(128091365), *FLI1*	C	-0.025	3.60E-11	0.003	0.007	6.98E-01	different	0.002	0.007	7.78E-01	different
12	rs2856321(11747040), *ETV6*	G	0.029	4.50E-15	-0.008	0.007	2.50E-01	different	-0.013	0.007	6.49E-02	different
12	rs10770705(20748734), *SLCO1C1*	C	-0.033	8.00E-18	0.014	0.008	6.45E-02	different	0.004	0.007	6.38E-01	different
12	rs2638953(28425682), *CCDC91*	G	-0.032	6.70E-17	0.018	0.007	1.70E-02	different	-0.003	0.007	6.80E-01	same
12	rs2066807(55026949), *STAT2*	G	0.054	1.00E-13	0.018	0.014	2.07E-01	same	-0.001	0.014	9.16E-01	different
12	rs1351394(64638093), *HMGA2*	T	0.06	1.70E-65	0.012	0.007	8.47E-02	same	0.016	0.007	2.12E-02	same
12	rs10748128(68113925), *FRS2*	T	0.038	2.10E-20	-0.009	0.008	2.63E-01	different	-0.005	0.008	5.37E-01	different
12	rs11107116(92502635), *SOCS2*	T	0.052	1.40E-34	0.008	0.008	3.14E-01	same	-0.006	0.008	4.67E-01	different
12	rs7971536(100897919), *CCDC53/GNPTAB*	T	0.028	8.20E-14	0.012	0.007	9.22E-02	same	0.003	0.007	7.20E-01	same
12	rs11830103(122389499), *SBNO1*	G	0.035	3.90E-15	-0.01	0.009	2.67E-01	different	-0.015	0.009	7.90E-02	different
13	rs7332115(32045548), *PDS5B/BRCA2*	T	-0.023	5.50E-10	-0.006	0.007	4.16E-01	same	0.007	0.007	3.49E-01	different
13	rs3118905(50003335), *DLEU7*	G	0.056	1.10E-45	-0.024	0.008	2.40E-03	different	-0.004	0.008	6.44E-01	different
13	rs7319045(90822575), *GPC5*	G	-0.025	1.20E-11	-0.01	0.007	1.70E-01	same	-0.012	0.007	1.09E-01	same
14	rs1950500(23900690), *NFATC4*	T	0.034	2.20E-18	-0.011	0.008	1.62E-01	different	-0.004	0.008	6.24E-01	different

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	SNP ID(NCBI36 position), Nearest_OMIM_height_gene	Coded allele	Height		FEV_1				FEV_1/FVC			
			Beta	*P*	Beta	Se	*P*	Direction of effect for FEV_1 relative to height	Beta	Se	*P*	Direction of effect for FEV_1/FVC relative to height
14	rs2093210(60027032), *SIX6*	T	-0.032	6.20E-17	-0.007	0.008	3.77E-01	same	0.011	0.008	1.30E-01	different
14	rs1570106(67882868), *RAD51L1*	T	-0.026	8.10E-09	-0.018	0.009	4.56E-02	same	-0.021	0.009	1.46E-02	same
14	rs862034(74060499), *LTBP2*	G	0.028	7.30E-14	-0.013	0.007	8.64E-02	different	-0.003	0.007	6.97E-01	different
14	rs7155279(91555634), *TRIP11*	T	-0.024	1.40E-10	0.019	0.007	1.01E-02	different	0.008	0.007	2.81E-01	different
15	rs16964211(49317787), *CYP19A1*	G	0.05	1.70E-09	0.012	0.016	4.66E-01	same	-0.009	0.016	5.85E-01	different
15	rs7178424(60167551), *C2CD4A*	T	-0.021	5.60E-09	-0.012	0.007	8.08E-02	same	-0.007	0.007	2.93E-01	same
15	rs10152591(67835211), *TLE3*	C	-0.041	2.70E-10	-0.002	0.013	8.45E-01	same	-0.002	0.012	8.65E-01	same
15	rs12902421(69948457), *MYO9A*	T	-0.062	2.90E-08	0.01	0.023	6.75E-01	different	0.026	0.023	2.51E-01	different
15	rs5742915(72123686), *PML*	T	-0.031	1.00E-15	0.002	0.008	8.21E-01	different	0.017	0.008	2.85E-02	different
15	rs11259936(82371586), *ADAMTSL3*	C	0.044	1.70E-35	0.009	0.007	1.90E-01	same	0.021	0.007	2.77E-03	same
15	rs16942341(87189909), *ACAN*	T	-0.13	3.80E-27	0.004	0.024	8.80E-01	different	-0.002	0.023	9.38E-01	same
15	rs2871865(97012419), *IGF1R*	G	-0.057	2.90E-21	-0.004	0.012	7.61E-01	same	0.017	0.012	1.35E-01	different
15	rs4965598(98577137), *ADAMTS17*	T	-0.028	4.30E-13	-0.003	0.008	7.15E-01	same	0.015	0.007	4.02E-02	different
16	rs11648796(732191), *NARFL*	G	0.034	1.20E-13	-0.01	0.01	3.22E-01	different	0.002	0.01	8.69E-01	same
16	rs26868(2189377), *CASKIN1*	T	-0.034	9.00E-17	0.004	0.008	6.11E-01	different	0	0.008	9.87E-01	different
16	rs1659127(14295806), *MKL2*	G	-0.027	1.10E-11	-0.008	0.008	3.08E-01	same	0.005	0.008	5.47E-01	different
16	rs8052560(87304743), *CTU2/GALNS*	C	-0.029	3.30E-08	0.002	0.011	8.43E-01	different	-0.001	0.011	9.52E-01	same
17	rs4640244(21224816), *KCNJ12*	G	-0.024	2.30E-08	-0.004	0.008	6.03E-01	same	0.003	0.008	7.20E-01	different
17	rs3110496(24941897), *ANKRD13B*	G	0.022	7.30E-09	-0.006	0.008	4.54E-01	different	-0.012	0.007	1.05E-01	different
17	rs3764419(26188149), *ATAD5/RNF135*	C	0.035	1.80E-21	-0.016	0.007	2.56E-02	different	0.002	0.007	8.21E-01	same
17	rs17780086(27367395), *LRRC37B*	G	-0.028	2.60E-08	-0.002	0.01	8.22E-01	same	-0.025	0.01	1.31E-02	same
17	rs1043515(34175722), *PIP4K2B*	G	0.023	2.90E-10	-0.005	0.007	4.46E-01	different	-0.027	0.007	**1.22E-04**	different
17	rs4986172(40571807), *ACBD4*	T	-0.032	2.30E-16	-0.009	0.008	2.55E-01	same	-0.008	0.007	2.76E-01	same
17	rs2072153(44745013), *ZNF652*	G	-0.021	3.50E-08	-0.002	0.008	7.84E-01	same	-0.007	0.007	3.54E-01	same
17	rs4605213(46599746), *NME2*	G	-0.021	2.70E-08	0.013	0.008	9.69E-02	different	-0.002	0.007	8.19E-01	same
17	rs227724(52133816), *NOG*	T	0.03	7.40E-15	-0.001	0.007	8.41E-01	different	-0.001	0.007	8.74E-01	different
17	rs2079795(56851431), *TBX2*	T	0.04	2.10E-24	-0.023	0.008	2.78E-03	different	-0.006	0.008	4.24E-01	different

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	SNP ID(NCBI36 position), Nearest_OMIM_height_gene	Coded allele	Height		FEV_1				FEV_1/FVC			
			Beta	*P*	Beta	Se	*P*	Direction of effect for FEV_1 relative to height	Beta	Se	*P*	Direction of effect for FEV_1/FVC relative to height
17	rs2665838(59320197), *CSH1/GH1*	G	0.042	5.10E-25	0.006	0.008	4.31E-01	same	0	0.008	9.82E-01	different
17	rs11867479(65601802), *KCNJ16/KCNJ2*	T	0.025	1.50E-10	-0.007	0.008	3.65E-01	different	0.006	0.008	4.25E-01	same
18	rs4800452(18981609), *CABLES1*	T	0.051	4.20E-30	-0.023	0.009	7.84E-03	different	0	0.009	9.98E-01	different
18	rs9967417(45213498), *DYM*	G	0.038	9.30E-25	-0.004	0.007	6.04E-01	different	-0.002	0.007	7.65E-01	different
18	rs17782313(56002077), *MC4R*	T	-0.028	3.80E-11	-0.016	0.008	5.22E-02	same	0	0.008	9.59E-01	different
19	rs12982744(2128193), *DOT1L*	G	0.03	3.40E-16	-0.01	0.007	1.80E-01	different	-0.002	0.007	7.38E-01	different
19	rs7507204(3379834), *NFIC*	G	-0.036	4.30E-16	-0.006	0.009	5.22E-01	same	-0.007	0.009	3.90E-01	same
19	rs891088(7135762), *INSR*	G	0.029	2.40E-12	0	0.008	9.59E-01	different	0.002	0.008	8.30E-01	same
19	rs4072910(8550031), *ADAMTS10*	G	0.031	3.60E-13	0.002	0.009	8.23E-01	same	-0.003	0.009	6.93E-01	different
19	rs2279008(17144303), *MYO9B*	T	0.025	2.50E-08	-0.004	0.009	6.89E-01	different	0.004	0.009	6.61E-01	same
19	rs17318596(46628935), *ATP5SL*	G	-0.032	5.00E-16	0.006	0.008	4.39E-01	different	0.003	0.008	6.73E-01	different
20	rs1741344(4049800), *SMOX*	T	-0.023	3.30E-09	0.012	0.007	1.09E-01	different	0	0.007	9.82E-01	different
20	rs2145272(6574218), *BMP2*	G	0.039	2.10E-24	-0.018	0.007	1.84E-02	different	0.016	0.007	2.60E-02	same
20	rs7274811(31796842), *ZNF341*	T	-0.041	5.90E-22	-0.003	0.008	7.03E-01	same	-0.004	0.008	5.87E-01	same
20	rs143384(33489170), *GDF5*	G	0.063	1.00E-58	0.004	0.008	1.54E-01	same	0.018	0.008	1.45E-02	same
20	rs237743(47336426), *ZNFX1*	G	-0.041	1.30E-20	0.008	0.007	3.65E-01	different	0.008	0.008	3.41E-01	different
21	rs2834442(34612656), *KCNE2*	T	-0.026	5.10E-12	-0.002	0.007	1.82E-01	same	-0.024	0.007	9.36E-04	same
22	rs4821083(31386341), *SYN3*	T	0.031	3.10E-10	0.03	0.008	1.93E-01	same	-0.014	0.01	1.54E-01	different

Nature Genetics: doi:10.1038/ng.941

Supplementary Table 4 Associations in never-smokers and ever-smokers in the Stage 1 and 2 data, and tests for interaction with smoking.

Each of the SNPs included in the table showed genome-wide significant association ($P<5x10^{-8}$) with either FEV_1 or FEV_1/FVC in the Stage 1 and 2 data. The stratum-specific estimates are shown for Stage 1 (GWAS), Stage 2 and Stage 1 + 2 data for the trait with the strongest association. *P* values for the interaction between ever-smokers and never-smokers in the combined Stage 1 + 2 meta-analysis are shown.

Nature Genetics: doi:10.1038/ng.941

Chr.	SNP_ID (NCBI36 position), function	Measure	Stage 1						Stage 2						Stage 1 + Stage 2 meta-analysis				Interaction
			Ever-smokers			Never-smokers			Ever-smokers			Never-smokers			Ever-smokers		Never-smokers		
			Beta	Se	*P*	Beta	Se	*P*	Beta	Se	*P*	Beta	Se	*P*	Beta	Se	Beta	Se	*P*
1	rs2284746 (17179262), *MFAP2* (intron)	FEV_1/FVC	-0.046	0.009	5.82E-07	-0.036	0.010	3.43E-04	-0.039	0.010	1.04E-04	-0.037	0.011	7.02E-04	-0.043	0.007	-0.036	0.007	5.12E-01
1	rs993925 (216926691), *TGFB2* (downstream)	FEV_1/FVC	0.047	0.010	4.85E-06	0.032	0.011	4.59E-03	0.031	0.014	2.29E-02	0.016	0.014	2.43E-01	0.041	0.008	0.026	0.009	1.91E-01
2	rs12477314 (239542085), *HDAC4* (downstream)	FEV_1/FVC	0.054	0.012	3.58E-06	0.049	0.013	1.14E-04	0.042	0.011	1.90E-04	0.019	0.011	9.56E-02	0.048	0.008	0.032	0.008	1.88E-01
3	rs1529672 (25495586), *RARB* (intron)	FEV_1/FVC	-0.072	0.013	1.64E-08	-0.044	0.014	1.97E-03	-0.047	0.012	7.44E-05	-0.025	0.012	3.80E-02	-0.059	0.009	-0.033	0.009	4.29E-02
3	rs1344555 (170782913), *MECOM* (intron)	FEV_1	-0.053	0.012	5.13E-06	-0.029	0.012	2.05E-02	-0.022	0.013	8.79E-02	-0.029	0.012	1.57E-02	-0.040	0.009	-0.029	0.009	3.81E-01
5	rs153916 (95062456), *SPATA9* (upstream)	FEV_1/FVC	-0.041	0.009	6.39E-06	-0.022	0.010	2.81E-02	-0.022	0.013	9.55E-02	-0.028	0.013	3.31E-02	-0.035	0.008	-0.024	0.008	3.28E-01
6	rs6903823 (28430275), *ZKSCAN3* (intron)/ *ZNF323* (intron)	FEV_1	-0.043	0.012	3.30E-04	-0.051	0.012	4.20E-05	-0.033	0.012	5.79E-03	-0.027	0.011	1.63E-02	-0.038	0.008	-0.037	0.008	9.64E-01
6	rs2857595 (31676448), *NCR3* (upstream)	FEV_1/FVC	0.053	0.012	9.43E-06	0.045	0.013	7.01E-04	0.035	0.011	1.91E-03	0.020	0.012	7.59E-02	0.043	0.008	0.031	0.009	3.11E-01
6	rs2798641 (109374743), *ARMC2* (intron)	FEV_1/FVC	-0.056	0.012	2.97E-06	-0.036	0.013	6.55E-03	-0.038	0.017	2.70E-02	-0.020	0.017	2.45E-01	-0.050	0.010	-0.030	0.010	1.67E-01

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	**SNP_ID (NCBI36 position), function**	**Measure**	**Stage 1**						**Stage 2**						**Stage 1 + Stage 2 meta-analysis**				**Interaction**
			Ever-smokers			**Never-smokers**			**Ever-smokers**			**Never-smokers**			**Ever-smokers**		**Never-smokers**		
			Beta	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	***P***	**Beta**	**Se**	**Beta**	**Se**	***P***
10	rs7068966 (12317998), *CDC123* (intron)	FEV_1/FVC	0.053	0.009	4.55E-09	0.034	0.010	7.48E-04	0.029	0.009	1.31E-03	0.016	0.009	7.72E-02	0.041	0.006	0.024	0.007	7.15E-02
10	rs11001819 (77985230), *C10orf11* (intron)	FEV_1	-0.036	0.010	1.61E-04	-0.047	0.010	2.99E-06	-0.020	0.008	9.85E-03	-0.023	0.007	1.23E-03	-0.026	0.006	-0.031	0.006	5.56E-01
12	rs11172113 (55813550), *LRP1* (intron)	FEV_1/FVC	-0.036	0.010	1.66E-04	-0.034	0.010	1.01E-03	-0.033	0.013	1.31E-02	-0.020	0.013	1.40E-01	-0.035	0.008	-0.029	0.008	5.97E-01
12	rs1036429 (94795559), *CCDC38* (intron)	FEV_1/FVC	0.049	0.011	1.76E-05	0.050	0.012	6.28E-05	0.039	0.011	3.91E-04	0.019	0.011	8.77E-02	0.044	0.008	0.033	0.008	3.45E-01
16	rs12447804 (56632783), *MMP15* (intron)	FEV_1/FVC	-0.058	0.013	7.47E-06	-0.047	0.014	6.79E-04	-0.029	0.014	4.77E-02	-0.011	0.014	4.31E-01	-0.045	0.010	-0.030	0.010	2.71E-01
16	rs2865531 (73947817), *CFDP1* (intron)	FEV_1/FVC	0.034	0.009	2.23E-04	0.045	0.010	9.48E-06	0.033	0.009	2.19E-04	0.014	0.009	1.20E-01	0.034	0.006	0.028	0.007	5.42E-01
21	rs9978142 (34574109), *KCNE2* (upstream)	FEV_1/FVC	-0.055	0.013	1.61E-05	-0.040	0.014	5.06E-03	-0.046	0.019	1.34E-02	-0.018	0.019	3.25E-01	-0.052	0.011	-0.032	0.011	1.94E-01

Nature Genetics: doi:10.1038/ng.941

Supplementary Table 5 Lung function associations (FEV_1 and FEV_1/FVC) in Stage 1 for all previously reported loci (A) and all SNPs taken forward to follow-up in Stage 2 (B).

A) Effects in the SpiroMeta-CHARGE dataset of loci previously reported to be associated with FEV_1, FEV_1/FVC, or COPD[1,2,91-94] providing that they also showed association with lung function, are shown for both lung function measures. Details of the selection of these SNPs are shown in the **Online Methods**. *P* values below $5x10^{-8}$ for FEV_1 and FEV_1/FVC are shown in bold. 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. For each locus, the most significant SNP for either FEV_1 or FEV_1/FVC (as indicated in the third column) in the SpiroMeta-CHARGE dataset (SpiroMeta-CHARGE in 2nd column) is also included. Abbreviations: ns=nonsynonymous, s= synonymous.
Previous studies which report these loci are Cho *et al.* (2010)[91], DeMeo *et al.* (2009)[94], Hancock *et al.* (2010)[1], Pillai *et al.* (2009)[92], Repapi *et al.* (2010)[2] and Wilk *et al.* (2009)[93].

Nature Genetics: doi:10.1038/ng.941

Chr.	Paper reported	Measure	SNP ID (NCBI36 position), function	Region name	Coded allele	FEV_1		FEV_1/FVC		N
						Beta (Se)	*P*	Beta (Se)	*P*	
2	Repapi *et al.*	FEV_1	rs2571445(218391399), *TNS1*(ns)	*TNS1*	G	0.047 (0.007)	**9.83E-11**	0.033 (0.007)	4.46E-06	45839
2	Hancock *et al.*	FEV_1/FVC	rs10498230(229210747), *PID1*(downstream)	*PID1*	T	0.03 (0.014)	3.60E-02	0.068 (0.014)	1.13E-06	44957
4	Hancock *et al.*	FEV_1/FVC	rs2869967(90088355), *FAM13A*	*FAM13A*	T	0.012 (0.007)	9.38E-02	0.047 (0.007)	**2.08E-11**	47710
4	SpiroMeta-CHARGE	FEV_1/FVC	rs2045517(90089987), *FAM13A* (intron)	*FAM13A*	T	-0.012 (0.007)	8.93E-02	-0.047 (0.007)	**2.00E-11**	47675
4	Cho *et al.*	COPD	rs7671167(90103002), *FAM13A*	*FAM13A*	T	-0.017 (0.007)	1.64E-02	-0.042 (0.007)	**1.27E-09**	47723
4	Repapi *et al.*	FEV_1	rs10516526(106908353), *GSTCD*(intron)	*GSTCD-NPNT*	G	0.108 (0.014)	**4.75E-14**	0.039 (0.014)	6.17E-03	47970
4	Hancock *et al.*	FEV_1	rs17331332(107027556), *NPNT* (upstream)	*GSTCD-NPNT*	G	-0.102 (0.014)	1.11E-12	-0.057 (0.014)	5.30E-05	39503
4	SpiroMeta-CHARGE	FEV_1/FVC	rs6823809(107048244), *NPNT*(intron)	*GSTCD-NPNT*	T	0.050 (0.011)	4.82E-06	0.056 (0.011)	2.20E-07	23656
4	SpiroMeta-CHARGE	FEV_1	rs1032296(145654138), *HHIP*(upstream)	*HHIP*	T	-0.047 (0.007)	8.74E-11	-0.050 (0.007)	**3.42E-12**	45318
4	Repapi *et al.*	FEV_1/FVC	rs12504628(145655774), *HHIP*	*HHIP*	T	-0.044 (0.007)	**1.03E-09**	-0.063 (0.007)	**5.54E-19**	46204
4	Wilk *et al.*	FEV_1/FVC	rs11100860(145698589), *HHIP*(upstream)	*HHIP*	G	0.041 (0.007)	**4.27E-09**	0.064 (0.007)	**6.81E-20**	47876
4	Hancock *et al.*	FEV_1/FVC	rs1980057(145705188), *HHIP*	*HHIP*	T	0.042 (0.007)	**4.07E-09**	0.063 (0.007)	**1.06E-19**	47865
5	Hancock *et al.*	FEV_1/FVC	rs11168048(147822546), *HTR4*(intron)	*HTR4*	T	-0.046 (0.007)	**2.43E-10**	-0.047 (0.007)	**5.97E-11**	44976
5	Repapi *et al.*	FEV_1	rs3995090(147826008), *HTR4*	*HTR4*	C	0.045 (0.007)	**3.33E-10**	0.046 (0.007)	**1.04E-10**	47607
5	SpiroMeta-CHARGE	FEV_1	rs1985524(147827981), *HTR4*	*HTR4*	G	-0.048 (0.007)	**3.06E-11**	-0.045 (0.007)	**2.90E-10**	45623
5	Hancock *et al.*	FEV_1/FVC	rs2277027(156864954), *ADAM19*	*ADAM19*	C	-0.026 (0.007)	3.10E-04	-0.042 (0.007)	**6.65E-09**	48023
5	SpiroMeta-CHARGE	FEV_1/FVC	rs11134779(156869344), *ADAM19*(intron)	*ADAM19*	G	-0.027 (0.007)	2.40E-04	-0.042 (0.007)	**6.01E-09**	48075
6	Hancock *et al.* and Repapi *et al.*	FEV_1/FVC	rs2070600(32259421), *AGER*(ns)	*AGER*	T	0.025 (0.016)	1.27E-01	0.126 (0.016)	**9.07E-15**	46314
6	Repapi *et al.*	FEV_1/FVC	rs2395730(39892343), *DAAM2* (intron)	*DAAM2*	C	-0.004 (0.007)	5.95E-01	0.022 (0.007)	1.39E-03	47256
6	SpiroMeta-CHARGE	FEV_1/FVC	rs11756622(39898021), *DAAM2* (intron)	*DAAM2*	T	0.047 (0.019)	1.23E-02	0.064 (0.019)	5.48E-04	28276

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	Paper reported	Measure	SNP ID (NCBI36 position), function	Region name	Coded allele	FEV_1		FEV_1/FVC		N
						Beta (Se)	*P*	Beta (Se)	*P*	
6	Hancock *et al.*	FEV_1/FVC	rs3817928(142792209), *GPR126*	*GPR126-LOC153910*	G	0.023 (0.009)	8.63E-03	0.059 (0.008)	**2.27E-12**	46730
6	SpiroMeta-CHARGE	FEV_1/FVC	rs262129(142894837), *LOC153910*(unknown)	*GPR126-LOC153910*	G	0.031 (0.008)	5.44E-05	0.056 (0.008)	**2.91E-13**	47014
9	SpiroMeta-CHARGE	FEV_1/FVC	rs16909859(97244613), *PTCH1*(downstream)	*PTCH1*	G	-0.014 (0.013)	2.93E-01	0.08 (0.013)	**7.45E-10**	43353
9	Hancock *et al.*	FEV_1/FVC	rs16909898(97270829), *PTCH1*	*PTCH1*	G	0.015 (0.012)	2.21E-01	-0.072 (0.012)	**3.94E-09**	42486
15	Repapi *et al.*	FEV_1/FVC	rs12899618(69432174), *THSD4*	*THSD4*	G	0.036 (0.01)	1.57E-04	0.076 (0.01)	**1.86E-15**	46657
15	SpiroMeta-CHARGE	FEV_1/FVC	rs8033889(69467134), *THSD4* (intron)	*THSD4*	T	-0.044 (0.009)	3.01E-07	-0.072 (0.008)	**2.03E-17**	46995
15	DeMeo *et al* (2009)	COPD	rs2568494(76528019), *IREB2*(intron)	*CHRNA3-CHRNA5-IREB2-LOC123688*	G	0.023 (0.007)	1.64E-03	0.029 (0.007)	5.25E-05	47919
15	Pillai *et al.* (2009)	COPD	rs8034191(76593078), *LOC123688*	*CHRNA3-CHRNA5-IREB2-LOC123688*	T	0.031 (0.007)	2.07E-05	0.032 (0.007)	9.65E-06	47954
15	SpiroMeta-CHARGE	FEV_1	rs2036527(76638670), *CHRNA5*(uptream)	*CHRNA3-CHRNA5-IREB2-LOC123688*	G	0.036 (0.008)	2.40E-06	0.032 (0.007)	1.19E-05	45038
15	SpiroMeta-CHARGE	FEV_1/FVC	rs8040868(76698236), *CHRNA3*(s)	*CHRNA3-CHRNA5-IREB2-LOC123688*	T	0.039 (0.008)	2.98E-06	0.04 (0.008)	1.14E-06	35121

Nature Genetics: doi:10.1038/ng.941

B) Tests for association with lung function for all SNPs followed up in Stage 2. 34 SNPs which showed novel evidence of association ($P < 3x10^{-6}$) were followed up for both lung function measures in Stage 2. The sample sizes (*N*) shown are the effective sample sizes for Stage 1, Stage 2 and the joint analysis of Stage 1 and 2. Effective sample size within each study is the product of sample size and imputation quality metric. $P< 5x10^{-8}$ for Stage 1 and 2 are shown in bold. Abbreviations: ns=nonsynonymous, s= synonymous. For the *CHRNA3/5* and *MMP15* regions, the lead SNP had an effective sample size ≥70 but <80% and so a proxy SNP was also included. For the *CHRNA3/5*, *MICB/NCR3* and *CDC123* regions, there were different top SNPs for FEV_1 and FEV_1/FVC and so both lead SNPs were taken forward.

Nature Genetics: doi:10.1038/ng.941

Chr.	Measure	SNP_ID (NCBI36 position), function	Coded allele	Stage 1			Stage 2			Stage 1 + Stage 2 meta-analysis		
				Beta (Se)	*P*	N	Beta (Se)	*P*	N	Beta (Se)	*P*	N
1	FEV_1/FVC	rs2284746 (17179262), *MFAP2*(intron)	G	-0.042 (0.007)	2.47E-09	45944	-0.038 (0.007)	2.64E-07	35310	-0.04 (0.005)	**7.50E-16**	81254
1	FEV_1	rs2284746 (17179262), *MFAP2*(intron)	G	0.008 (0.007)	2.78E-01	45944	0.006 (0.007)	3.70E-01	35310	0.007 (0.005)	1.48E-01	81254
1	FEV_1/FVC	rs993925 (216926691), *TGFB2*(downstream)	T	0.04 (0.007)	2.54E-07	42402	0.023 (0.01)	1.76E-02	21162	0.034 (0.006)	**1.16E-08**	63564
1	FEV_1	rs993925 (216926691), *TGFB2*(downstream)	T	0.025 (0.007)	1.51E-03	42402	0.003 (0.007)	7.29E-01	21162	0.014 (0.005)	8.71E-03	63564
2	FEV_1/FVC	rs2544527 (15843619), *DDX1*(downstream)	T	-0.04 (0.007)	1.08E-07	45352	0 (0.01)	9.75E-01	21115	-0.026 (0.006)	8.73E-06	66467
2	FEV_1	rs2544527 (15843619), *DDX1*(downstream)	T	-0.024 (0.007)	1.55E-03	45352	-0.017 (0.007)	1.95E-02	21115	-0.021 (0.005)	5.53E-05	66467
2	FEV_1/FVC	rs3769124 (239014101), *ASB1*(intron)	G	-0.038 (0.01)	1.95E-04	44924	-0.032 (0.02)	1.11E-01	10579	-0.036 (0.009)	2.83E-05	55503
2	FEV_1	rs3769124 (239014101), *ASB1*(intron)	G	-0.053 (0.01)	2.76E-07	44924	-0.023 (0.02)	2.44E-01	10579	-0.047 (0.009)	6.50E-08	55503
2	FEV_1/FVC	rs12477314 (239542085), *HDAC4*(downstream)	T	0.052 (0.008)	4.48E-09	45585	0.031 (0.008)	8.41E-05	45704	0.041 (0.006)	**1.68E-12**	91289
2	FEV_1	rs12477314 (239542085), *HDAC4*(downstream)	T	0.032 (0.008)	2.77E-04	45585	0.025 (0.007)	1.82E-04	45704	0.028 (0.005)	1.02E-07	91289
3	FEV_1/FVC	rs1529672 (25495586), *RARB*(intron)	C	-0.06 (0.009)	7.75E-10	40624	-0.038 (0.009)	1.16E-05	45386	-0.048 (0.006)	**3.97E-14**	86010
3	FEV_1	rs1529672 (25495586), *RARB*(intron)	C	-0.037 (0.009)	1.78E-04	40624	-0.011 (0.007)	9.33E-02	45386	-0.02 (0.006)	2.16E-04	86010
3	FEV_1/FVC	rs9310995 (32904119), *TRIM71*(intron)	T	0.017 (0.007)	1.70E-02	44835	-0.013 (0.009)	1.60E-01	21070	0.007 (0.006)	2.36E-01	65905
3	FEV_1	rs9310995 (32904119), *TRIM71*(intron)	T	0.035 (0.007)	1.28E-06	44835	0.009 (0.007)	2.00E-01	21070	0.023 (0.005)	3.60E-06	65905
3	FEV_1/FVC	rs1344555 (170782913), *MECOM*(intron)	T	-0.019 (0.008)	2.61E-02	46067	-0.017 (0.012)	1.55E-01	21104	-0.018 (0.007)	6.65E-03	67171
3	FEV_1	rs1344555 (170782913), *MECOM*(intron)	T	-0.042 (0.008)	1.91E-06	46067	-0.025 (0.009)	6.44E-03	21104	-0.034 (0.006)	**2.65E-08**	67171
4	FEV_1/FVC	rs1541374 (106267809), *TET2*(upstream)	T	-0.026 (0.007)	5.56E-04	45221	-0.014 (0.01)	1.72E-01	20516	-0.022 (0.006)	2.05E-04	65737
4	FEV_1	rs1541374 (106267809), *TET2*(upstream)	T	-0.036 (0.007)	2.43E-06	45221	-0.015 (0.007)	4.36E-02	20516	-0.026 (0.005)	5.80E-07	65737
5	FEV_1/FVC	rs1551943 (52230790), *ITGA1*(intron)	G	0.048 (0.008)	1.20E-08	43787	0.007 (0.008)	3.71E-01	45914	0.026 (0.006)	2.43E-06	89701
5	FEV_1	rs1551943 (52230790), *ITGA1*(intron)	G	0.022 (0.008)	9.93E-03	43787	-0.006 (0.006)	3.53E-01	45914	0.004 (0.005)	3.61E-01	89701
5	FEV_1/FVC	rs153916 (95062456), *SPATA9*(upstream)	T	-0.033 (0.007)	2.06E-06	47530	-0.025 (0.009)	6.67E-03	21428	-0.031 (0.005)	**2.12E-08**	68958
5	FEV_1	rs153916 (95062456), *SPATA9*(upstream)	T	-0.001 (0.007)	8.91E-01	47530	0.004 (0.007)	6.22E-01	21428	0.001 (0.005)	8.20E-01	68958
5	FEV_1/FVC	rs10067603 (131831767), *C5orf56*(downstream)	G	-0.04 (0.008)	1.60E-06	44134	-0.006 (0.011)	6.03E-01	21167	-0.028 (0.006)	1.46E-05	65301
5	FEV_1	rs10067603 (131831767), *C5orf56*(downstream)	G	-0.007 (0.008)	3.83E-01	44134	0.013 (0.008)	1.14E-01	21167	0.002 (0.006)	6.74E-01	65301
6	FEV_1/FVC	rs1928168 (22125717), *AK026189*(intron)	T	0.037 (0.007)	8.99E-08	47936	0.011 (0.009)	2.40E-01	21323	0.028 (0.005)	1.69E-07	69259
6	FEV_1	rs1928168 (22125717), *AK026189*(intron)	T	0.025 (0.007)	2.61E-04	47936	0.002 (0.007)	7.69E-01	21323	0.015 (0.005)	2.25E-03	69259
6	FEV_1/FVC	rs6903823 (28430275), *ZKSCAN3*(intron)/*ZNF323*(intron)	G	-0.027 (0.008)	2.28E-03	47057	-0.013 (0.011)	2.34E-01	21428	-0.021 (0.007)	1.19E-03	68485
6	FEV_1	rs6903823 (28430275), *ZKSCAN3*(intron)/*ZNF323*(intron)	G	-0.046 (0.008)	2.00E-07	47057	-0.029 (0.008)	4.75E-04	21428	-0.037 (0.006)	**2.18E-10**	68485

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	Measure	SNP_ID(NCBI36 position), function	Coded allele	Stage 1			Stage 2			Stage 1 + Stage 2 meta-analysis		
				Beta (Se)	*P*	N	Beta (Se)	*P*	N	Beta (Se)	*P*	N
6	FEV_1/FVC	rs3094548 (29463181), *OR12D2*(upstream)	G	-0.027 (0.008)	1.15E-03	42516	-0.015 (0.01)	1.37E-01	20733	-0.022 (0.006)	3.39E-04	63249
6	FEV_1	rs3094548 (29463181), *OR12D2*(upstream)	G	-0.042 (0.008)	4.11E-07	42516	-0.016 (0.008)	3.60E-02	20733	-0.029 (0.005)	1.45E-07	63249
6	FEV_1/FVC	rs2855812 (31580699), *MICB*(intron)	T	-0.034 (0.008)	5.11E-05	46921	-0.015 (0.011)	1.59E-01	21190	-0.027 (0.006)	2.45E-05	68111
6	FEV_1	rs2855812 (31580699), *MICB*(intron)	T	-0.045 (0.008)	8.57E-08	46921	-0.013 (0.008)	1.06E-01	21190	-0.03 (0.006)	1.80E-07	68111
6	FEV_1/FVC	rs2857595 (31676448), *NCR3*(upstream)	G	0.049 (0.009)	7.86E-08	45540	0.028 (0.008)	5.36E-04	45657	0.037 (0.006)	**2.28E-10**	91197
6	FEV_1	rs2857595 (31676448), *NCR3*(upstream)	G	0.04 (0.009)	1.46E-05	45540	0.017 (0.007)	9.41E-03	45657	0.025 (0.005)	1.30E-06	91197
6	FEV_1/FVC	rs2647044 (32775888), *HLA-DQB1*(upstream)	G	0.053 (0.011)	2.71E-06	44610	0.007 (0.022)	7.63E-01	8381	0.044 (0.01)	5.95E-06	52991
6	FEV_1	rs2647044 (32775888), *HLA-DQB1*(upstream)	G	0.031 (0.011)	6.71E-03	44610	0.009 (0.022)	6.71E-01	8381	0.027 (0.01)	5.62E-03	52991
6	FEV_1/FVC	rs2798641 (109374743), *ARMC2*(intron)	T	-0.047 (0.009)	2.81E-07	46369	-0.03 (0.012)	1.57E-02	20999	-0.041 (0.007)	**8.35E-09**	67368
6	FEV_1	rs2798641 (109374743), *ARMC2*(intron)	T	-0.046 (0.009)	5.39E-07	46369	-0.009 (0.01)	3.35E-01	20999	-0.03 (0.006)	4.69E-06	67368
6	FEV_1/FVC	rs3734729 (150612560), *PPP1R14C*(untranslated-3)	G	-0.045 (0.017)	8.71E-03	43680	-0.058 (0.023)	1.00E-02	20998	-0.05 (0.013)	1.93E-04	64678
6	FEV_1	rs3734729 (150612560), *PPP1R14C*(untranslated-3)	G	-0.085 (0.016)	1.08E-06	43680	-0.021 (0.017)	2.24E-01	20998	-0.055 (0.012)	4.48E-06	64678
10	FEV_1/FVC	rs1878798 (12283489), *CDC123*(intron)	G	0.042 (0.007)	3.48E-09	46164	0.024 (0.009)	1.15E-02	21086	0.035 (0.005)	**9.56E-11**	67250
10	FEV_1	rs1878798 (12283489), *CDC123*(intron)	G	0.042 (0.007)	3.11E-09	46164	0.015 (0.007)	3.65E-02	21086	0.029 (0.005)	**1.84E-09**	67250
10	FEV_1/FVC	rs7068966 (12317998), *CDC123*(intron)	T	0.045 (0.007)	1.28E-10	47085	0.023 (0.006)	3.86E-04	45892	0.033 (0.005)	**6.13E-13**	92977
10	FEV_1	rs7068966 (12317998), *CDC123*(intron)	T	0.04 (0.007)	1.19E-08	47085	0.022 (0.005)	3.56E-05	45892	0.029 (0.004)	**2.82E-12**	92977
10	FEV_1/FVC	rs11001819 (77985230), *C10orf11*(intron)	G	-0.019 (0.007)	6.50E-03	45546	-0.006 (0.006)	3.17E-01	45677	-0.012 (0.005)	7.58E-03	91223
10	FEV_1	rs11001819 (77985230), *C10orf11*(intron)	G	-0.041 (0.007)	1.42E-08	45546	-0.022 (0.005)	3.10E-05	45677	-0.029 (0.004)	**2.98E-12**	91223
12	FEV_1/FVC	rs4762767 (19757396), *AEBP2*(downstream)	G	-0.036 (0.007)	2.42E-06	48016	-0.008 (0.011)	4.47E-01	21324	-0.027 (0.006)	8.15E-06	69340
12	FEV_1	rs4762767 (19757396), *AEBP2*(downstream)	G	-0.028 (0.007)	3.85E-04	48016	-0.012 (0.008)	1.34E-01	21324	-0.021 (0.005)	1.52E-04	69340
12	FEV_1/FVC	rs11172113 (55813550), *LRP1*(intron)	T	-0.035 (0.007)	1.36E-06	45387	-0.026 (0.01)	5.83E-03	20256	-0.032 (0.006)	**1.24E-08**	65643
12	FEV_1	rs11172113 (55813550), *LRP1*(intron)	T	-0.021 (0.007)	3.55E-03	45387	-0.003 (0.007)	6.94E-01	20256	-0.013 (0.005)	1.19E-02	65643
12	FEV_1/FVC	rs1036429 (94795559), *CCDC38*(intron)	T	0.049 (0.008)	1.24E-08	47814	0.028 (0.008)	3.35E-04	46183	0.038 (0.006)	**2.30E-11**	93997
12	FEV_1	rs1036429 (94795559), *CCDC38*(intron)	T	0.01 (0.008)	2.67E-01	47814	0.004 (0.006)	5.38E-01	46183	0.006 (0.005)	2.26E-01	93997
15	FEV_1/FVC	rs2036527 (76638670), *CHRNA5*(uptream)	G	0.032 (0.007)	1.19E-05	45038	0 (0.01)	9.82E-01	20874	0.022 (0.006)	1.81E-04	65912
15	FEV_1	rs2036527 (76638670), *CHRNA5*(uptream)	G	0.036 (0.007)	2.40E-06	45038	0.015 (0.008)	5.44E-02	20874	0.026 (0.005)	6.90E-07	65912
15	FEV_1/FVC	rs12914385 (76685778), *CHRNA3*(intron)	T	-0.03 (0.007)	2.28E-05	47226	0.002 (0.01)	8.08E-01	21327	-0.019 (0.006)	5.17E-04	68553
15	FEV_1	rs12914385 (76685778), *CHRNA3*(intron)	T	-0.034 (0.007)	2.95E-06	47226	-0.015 (0.007)	4.10E-02	21327	-0.025 (0.005)	4.72E-07	68553
15	FEV_1/FVC	rs8040868 (76698236), *CHRNA3*(s)	T	0.04 (0.008)	1.14E-06	35121	-0.005 (0.01)	6.10E-01	21131	0.022 (0.006)	3.01E-04	56252
15	FEV_1	rs8040868 (76698236), *CHRNA3*(s)	T	0.039 (0.008)	2.98E-06	35121	0.012 (0.007)	9.86E-02	21131	0.025 (0.005)	4.06E-06	56252

Contd.

Nature Genetics: doi:10.1038/ng.941

Contd.

Chr.	Measure	SNP_ID(NCBI36 position), function	Coded allele	Stage 1			Stage 2			Stage 1 + Stage 2 meta-analysis		
				Beta (Se)	*P*	N	Beta (Se)	*P*	N	Beta (Se)	*P*	N
16	FEV_1/FVC	rs12447804 (56632783), *MMP15*(intron)	T	-0.053 (0.009)	7.12E-08	35123	-0.021 (0.01)	4.20E-02	23693	-0.038 (0.007)	**3.59E-08**	58816
16	FEV_1	rs12447804 (56632783), *MMP15*(intron)	T	-0.017 (0.009)	8.02E-02	35123	0.004 (0.007)	5.71E-01	23693	-0.004 (0.006)	4.73E-01	58816
16	FEV_1/FVC	rs3743563 (56636666), *MMP15*(missense)	G	0.043 (0.008)	1.80E-07	47179	0.013 (0.008)	1.22E-01	43190	0.028 (0.006)	6.76E-07	90369
16	FEV_1	rs3743563 (56636666), *MMP15*(missense)	G	0.015 (0.008)	8.52E-02	47179	-0.001 (0.007)	8.74E-01	43190	0.006 (0.005)	2.79E-01	90369
16	FEV_1/FVC	rs2865531 (73947817), *CFDP1*(intron)	T	0.039 (0.007)	2.30E-08	47594	0.024 (0.006)	1.94E-04	46286	0.031 (0.005)	**1.77E-11**	93880
16	FEV_1	rs2865531 (73947817), *CFDP1*(intron)	T	0.024 (0.007)	6.30E-04	47594	0.011 (0.005)	3.89E-02	46286	0.016 (0.004)	1.09E-04	93880
16	FEV_1/FVC	rs12716852 (76746239), *WWOX*(intron)	G	0.011 (0.007)	1.24E-01	47510	-0.004 (0.009)	6.85E-01	21228	0.006 (0.005)	2.81E-01	68738
16	FEV_1	rs12716852 (76746239), *WWOX*(intron)	G	0.036 (0.007)	3.45E-07	47510	0.013 (0.007)	7.11E-02	21228	0.025 (0.005)	1.92E-07	68738
21	FEV_1/FVC	rs9978142 (34574109), *KCNE2*(upstream)	T	-0.048 (0.009)	8.23E-07	44577	-0.031 (0.013)	1.75E-02	20693	-0.043 (0.008)	**2.65E-08**	65270
21	FEV_1	rs9978142 (34574109), *KCNE2*(upstream)	T	-0.012 (0.009)	2.47E-01	44577	-0.015 (0.01)	1.35E-01	20693	-0.013 (0.007)	5.57E-02	65270

Nature Genetics: doi:10.1038/ng.941

Supplementary Table 6 Estimated number of undiscovered variants and proportion of variance explained.

Effect sizes and standard errors estimated using non-discovery data are shown for genome-wide significant loci in SpiroMeta-CHARGE Stage 1 or Stage1 + 2 data. Sample sizes (N) are effective sample sizes for the non-discovery data. Estimates of discovery power for the 26 loci associated with lung function, estimates of the number of variants with similar effect sizes, and the proportion of the variance for FEV_1 and FEV_1/FVC that each individual discovered variant explains are shown. The total r^2 accounted for by discovered and putative undiscovered variants in combination is calculated as the sum of the products of r^2 multiplied by the number of estimated variants.

Chr.	SNP ID (NCBI36 position), function	FEV_1 excluding winners' curse bias		FEV_1/FVC excluding winners' curse bias		N	Power	Estimated variants	R2 (%) FEV_1	R2 (%) FEV_1/FVC
		Beta	Se	Beta	Se					
1	rs2284746(17179262), *MFAP2*(intron)	0.006	0.007	-0.038	0.007	35371[n2]	0.707	1.4	0.002	0.072
1	rs993925(216926691), *TGFB2*(downstream)	0.003	0.007	0.023	0.01	21414[n2]	0.214	4.7	0	0.024
2	rs2571445(218391399), *TNS1*(ns)	0.041	0.009	0.034	0.009	29130[s]	0.863	1.2	0.082	0.055
2	rs12477314(239542085), *HDAC4*(downstream)	0.025	0.007	0.031	0.008	45821[n2]	0.341	2.9	0.02	0.031
3	rs1529672(25495586), RARB(intron)	-0.011	0.007	-0.038	0.009	45466[n2]	0.376	2.7	0.003	0.041
3	rs1344555(170782913), *MECOM*(intron)	-0.025	0.009	-0.017	0.012	21313[n2]	0.207	4.8	0.021	0.01
4	rs2045517(90089987), *FAM13A* (intron)	-0.006	0.009	-0.037	0.009	25736[c]	0.654	1.5	0.002	0.067
4	rs10516526(106908353), *GSTCD*(intron)	0.07	0.034	0.035	0.033	7587[sc]	0.627	1.6	0.062	0.016
4	rs11100860(145698589), *HHIP*(upstream)	0.042	0.008	0.058	0.007	40202[f]	0.996	1	0.085	0.163
5	rs153916(95062456), *SPATA9*(upstream)	0.004	0.007	-0.025	0.009	21647[n2]	0.252	4	0.001	0.031
5	rs1985524(147827981), *HTR4*	-0.047	0.017	-0.052	0.017	7204[sc]	0.961	1	0.107	0.134
5	rs11134779(156869344), *ADAM19*(intron)	-0.03	0.01	-0.023	0.01	25917[c]	0.495	2	0.04	0.024
6	rs6903823(28430275), *ZKSCAN3*(intron)/*ZNF323*(intron)	-0.029	0.008	-0.013	0.011	21489[n2]	0.248	4	0.031	0.006
6	rs2857595(31676448), *NCR3*(upstream)	0.017	0.007	0.028	0.008	46107[n2]	0.129	7.8	0.009	0.025
6	rs2070600(32259421), *AGER*(ns)	0.012	0.04	0.093	0.04	7226[sc]	0.695	1.4	0.001	0.084
6	rs2798641(109374743), *ARMC2*(intron)	-0.009	0.01	-0.03	0.012	21173[n2]	0.214	4.7	0.002	0.026
6	rs262129(142894837), *LOC153910*(unknown)	0.014	0.01	0.045	0.01	25317[c]	0.319	3.1	0.008	0.081
9	rs16909859(97244613), *PTCH1*(downstream)	-0.021	0.018	0.062	0.017	22923[c]	0.539	1.9	0.006	0.058
10	rs7068966(12317998), *CDC123*(intron)	0.022	0.005	0.023	0.006	46067[n2]	0.209	4.8	0.024	0.026
10	rs11001819(77985230), *C10orf11*(intron)	-0.022	0.005	-0.006	0.006	45932[n2]	0.108	9.3	0.024	0.002
12	rs11172113(55813550), *LRP1*(intron)	-0.003	0.007	-0.026	0.01	20509[n2]	0.292	3.4	0	0.033
12	rs1036429(94795559), *CCDC38*(intron)	0.004	0.006	0.028	0.008	46311[n2]	0.204	4.9	0.001	0.026
15	rs8033889(69467134), *THSD4* (intron)	-0.039	0.011	-0.072	0.011	28974[s]	0.996	1	0.05	0.174
16	rs12447804(56632783), *MMP15*(intron)	0.004	0.007	-0.021	0.01	24398[n2]	0.059	16.8	0.001	0.015
16	rs2865531(73947817), *CFDP1*(intron)	0.011	0.005	0.024	0.006	46304[n2]	0.175	5.7	0.006	0.028
21	rs9978142(34574109), *KCNE2*(upstream)	-0.015	0.01	-0.031	0.013	20944[n2]	0.253	3.9	0.006	0.024

Nature Genetics: doi:10.1038/ng.941

Total variants		101.5		
Total % variance explained by estimated variants			1.355	3.016
Total % polygenic variance explained by estimated variants			3.388	7.538

[n2] no exclusions in SpiroMeta-CHARGE Stage 2

[s] excluding SpiroMeta consortium discovery GWAS of Repapi et al. (2010)

[c] excluding CHARGE consortium discovery GWAS of Hancock et al. (2010)

[sc] excluding SpiroMeta consortium discovery GWAS of Repapi et al. (2010) and CHARGE consortium discovery GWAS of Hancock et al. (2010)

[f] excluding FHS from SpiroMeta-CHARGE Stage 1

Nature Genetics: doi:10.1038/ng.941

Supplementary Note

PCR expression profiling

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 2.** 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. *GAPDH* gene was used as a positive control for the cDNA quality, and water (H_2O) as a negative control.

Linkage disequilibrium at 6p21.32-22.1 region of association with FEV1 and FEV1/FVC.

The novel signal of association with FEV_1/FVC at SNP rs2857595 (upstream of *NCR3*) is independent (r^2<0.01) of the previously reported association of rs2070600 (non-synonymous SNP in *AGER*).
The novel signal of association with FEV_1 at SNP rs6903823 (intron of *ZKSCAN3*/*ZNF323*) is also independent (r^2=0.01) of SNP rs2070600 (non-synonymous SNP in *AGER*) and is in low LD (r^2=0.31) with rs2857595 (novel signal upstream of *NCR3*) which show association with FEV_1/FVC.

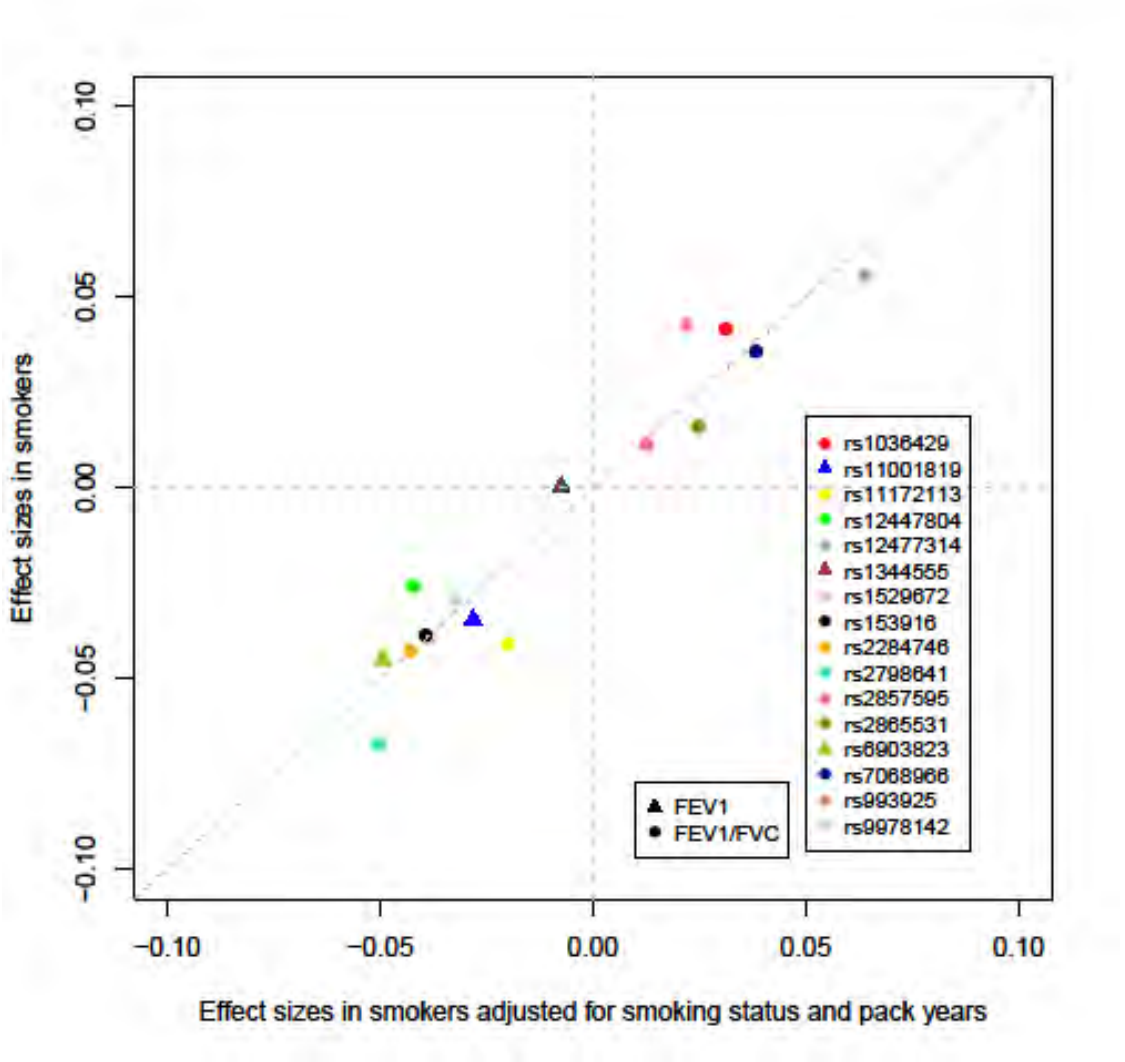
Interaction analysis

The analysis carried out to test for Gene x Smoking interaction (**Supplementary Table 4**) was a Z-test comparing the effect of a given SNP in ever-smokers and in never-smokers.

Nature Genetics: doi:10.1038/ng.941

Effect of pack-years adjustment

To assess the sensitivity of the results to the amount of cigarettes smoked, an analysis was undertaken on ever-smokers adjusting for smoking status (current or ex smoker) and pack-years. The cohorts that were included in this analysis (BHS2, CARDIA, Gedling, HCS, LBC1936, LifeLines, MESA-Lung, Nottingham Smokers, NSHD, SAPALDIA and TwinsUK-II) were a subset of the follow-up studies that were chosen because of the availability of pack-years data. As represented below, effect size estimates for ever-smokers adjusted for smoking status and pack years and effect sizes for ever-smokers without any smoking adjustment were very similar across the novel loci.



Nature Genetics: doi:10.1038/ng.941

Estimation of the number of undiscovered variants and calculation of the proportion of variance explained

Our approach to estimate the number of independent variants associated with lung function measures that have similar effect sizes to the 26 variants reported here is based on the approach developed by Park *et al.*[95]. We estimated winners' curse bias free effect sizes, and obtained the number of undiscovered variants using the discovery power to detect the unbiased effect sizes.

To calculate the unbiased effect sizes for the 26 genome-wide significant ($P<5x10^{-8}$) variants we excluded discovery data for each variant. Effect size estimates for the 16 novel loci reported here were obtained using SpiroMeta-CHARGE Stage 2 data. The discovery data for *HHIP*[92,93] included the data from FHS, therefore we excluded this study in our estimation of effect size estimates. For the remaining 9 loci previously discovered, effect sizes were calculated excluding studies involved in discovery GWAS of Repapi *et al*. 2010, or discovery GWAS of Hancock *et al*. 2010, or discovery GWAS of any of them, as indicated in **Supplementary Table 6.**

To estimate the power for discovered associations we followed the approach used by the ICBP consortium for systolic and diastolic blood pressure[96].This approach takes into account that two phenotypes were analysed in parallel and it also takes into account uncertainty about true effect sizes of discovered variants. The discovery power is expressed as a function of the true effect sizes and it is then integrated with respect to a joint probability distribution for true effect sizes on FEV_1 and FEV_1/FVC. The unbiased estimates of effect sizes are used as the means for the true effect size distributions.

We followed Park *et al*. 2010 to obtain the proportion of variance explained by the inferred number of variants. First, the number of variants of a given effect size was obtained as the inverse of the power to detect them. Then, the proportion of variance explained by a variant of effect size β^2_i and allele frequency p_i was calculated as

$$\frac{2\, p_i(i-p_i)\, \beta^2_i}{V}$$

where V is the phenotypic variance. Finally, the proportion of variance explained by the inferred number of variants is obtained summing the product of the number of variants of a given effect size by the proportion of variance explained by one of them, over each of the 26 genome-wide significant ($P<5x10^{-8}$) variants. We assumed a heritability of 40%[97-100] to estimate the proportion of the additive polygenic variance of each trait.

The confidence interval for the total number of variants was obtained using bootstrapping, as in Park *et al*. 2010.

MAGENTA analysis

We tested whether the Stage 1 GWAS results were enriched for known biological pathways using MAGENTA v2[101]. Briefly, MAGENTA defines a *P* value for each gene that is the lowest SNP *P* value within 110kb upstream and 40kb downstream of the gene and is corrected for gene size, number of SNPs per gene and linkage disequilibrium within the region. For each gene set, the null hypothesis that there is a random distribution of gene association score ranks within the gene set is tested against the alternative hypothesis that there are more gene association score ranks above a given rank cut-off (75^{th} percentile cut-off is recommended for polygenic traits) compared to random sampling of 10,000 gene sets of identical size. For each gene set, a false discovery rate (FDR) is calculated as the fraction of all randomly sampled gene sets (10,000 x number of gene sets tested) that have more genes with *P* value below the cut off (75^{th} percentile) than in the gene set being tested, divided by the fraction of real gene sets that have more genes with *P* value below the cut off (75^{th} percentile) than in the gene set being tested.

Six databases of biological pathways were tested: including Ingenuity Pathway (June 2008, number of pathways n=92), KEGG (2010, n=186), PANTHER Molecular Function (January 2010, n=276), PANTHER Biological Processes (January 2010, n=254), PANTHER Pathways (January 2010, n=141) and Gene Ontology (April 2010, n=9542). Significance thresholds were Bonferroni corrected for each database.

Nature Genetics: doi:10.1038/ng.941

Genes which contain a SNP with r^2>0.3 with any of our sentinel SNPs were flagged in the analysis. Genes in the HLA region on chromosome 6 were included.

No gene sets were significant for FEV_1/FVC. Two gene sets were significant for FEV_1 after multiple testing correction for the number of gene sets tested within the database.

From the KEGG database, a Systemic Lupus Erythematosus (SLE) was significant (P=1x10^{-3}, FDR=0.03). This gene set contains 140 genes with an effective number of genes of 74 in this analysis. Flagged genes for this gene set were *TNF* and *HIST1H2AL*.

From the PANTHER Biological Processes database, the MHC-I mediated immunity pathway was significant (P=2x10^{-3}, FDR=0.019). This gene set contains 22 genes with an effective number of genes of 16 in this analysis. Flagged genes for this gene set were *MICA* and *MICB*.

GRAIL analysis

The previously described GRAIL[102] algorithm was used to mine PubMed abstracts (prior to 2006), the Gene Ontology database and the Novartis Gene Expression Analysis for evidence of functional connections between genes in the 27 regions associated with lung function. For each lung function locus, boundaries were defined for testing in GRAIL such that all genes containing a SNP with r^2>0.3 (see **Supplementary Table 2**) with the leading SNP in the region were included.

No genes were identified that were connected by the literature, Gene Ontology or Expression Atlas to genes in the other associated loci (P<0.01).

CNV analysis

All lead 16 SNPs plus all SNPs with r^2>0.8 with any lead SNP were checked for correlation with known CNVs. 192 SNP genotypes from HapMap release 24 were compared to 5238 GSV CNV genotypes[103] and 1319 McCarroll *et al*[104] CNV genotypes for the same CEU individuals. No SNP had r^2>0.3 with any of the CNVs tested.

eQTL analysis

Within each of the 16 novel regions, the lead SNPs and any SNPs with r^2>0.8 with the lead SNP were searched against a database of expression SNPs (eSNPs) from lymphoblastoid cell lines[105]. We report eSNP associations exceeding a threshold of P<1.2x10^{-7}.

SNP rs3130805, a proxy for rs6903823 in *ZKSCAN3/ZNF323* (r^2=0.69), was significantly associated (P=1.3x10^{-15}) with transcript levels of *ZFP57*. The two top eSNPs in this region showing association with *ZFP57* were rs2535238 (P=2.8x10^{-34}) and rs2747457 (P=3.3x10^{-34}) which were ~350kb away from, and not independent of, rs3130805 (r^2=0.43 and 0.38, respectively).

SNP rs3799499, another proxy for rs6903823 (r^2=0.66), was significantly associated with transcript levels of ZSCAN12 (P=7.8x10^{-10}). This SNP is 7.3Kb away from, and in strong linkage disequilibrium (r^2=1) with, the top eSNP for transcript levels of ZSCAN12 (rs7774981, P=4.1x10^{-10}).

SNP rs2857595, upstream of *NCR3*, showed significant association with transcript levels of *HLA-DRB1* (3 probes, strongest P=1x10^{-10}) and *HLA-DQA1* (P=3.9x10^{-9}). The top eSNP associated with transcript levels of both of these

Nature Genetics: doi:10.1038/ng.941

genes was rs3129763 (P=2x10^{-36} for *HLA-DRB1* and P=2.8x10^{-39} for *HLA-DQA1*) which is 1.02Mb away from, and in weak LD (r^2= 0.30) with, rs2857595.

SNP rs1424013, a proxy for rs2865531 (r^2=0.82), is one of two top eSNPs for transcript levels of *CFDP1* (P=3.3x10^{-10}). The other top eSNP is rs999675 (P=2.9x10^{-10}) is 190Kb from rs1424013, downstream of *CFDP1*, and in linkage disequilibrium with rs1424013 (r^2=0.78).

SNP rs4478891, a proxy for rs7068966 in *CDC123* on chromosome 10, showed significant association with transcript levels of *OSBPL1A* on chromosome 18 (P=1x10^{-7}) suggesting a possible *trans* regulatory effect of rs4478891. No other SNP in the eSNP database showed significant association with transcript levels of *OSBPL1A*.

Nature Genetics: doi:10.1038/ng.941

Individual study descriptions

Stage 1 GWAS samples and phenotype measurement

This section describes study-specific characteristics that are not presented in the tables. All participants provided written informed consent and studies were approved by local Research Ethics Committees and/or Institutional Review boards.

The Reykjavik Study cohort originally comprised a random sample of 30,795 men and women born in 1907-1935 and living in Reykjavik in 1967. A total of 19,381 attended, resulting in 71% recruitment rate. The study sample was divided into six groups by birth year and birth date within month. One group was designated for longitudinal follow-up and was examined in all stages. One group was designated a control group and was not included in examinations until 1991. Other groups were invited to participate in specific stages of the study. Between 2002 and 2006, the **AGES**-Reykjavik study re-examined 5,764 survivors of the original cohort who had participated before in the Reykjavik Study. Pulmonary function was obtained by spirometry (Vitalograph). Testing was conducted with the participant in a standing position and a disposable mouthpiece. Participants were shown how to perform the maneuver by the technician before testing. A successful test session was defined as at least three acceptable maneuvers.

Atherosclerosis Risk in Communities (**ARIC**), is a population based study of risk factors for atherosclerosis and its sequelae[106] in adults from four U.S. field centers aged 45-64 at recruitment in 1987-1989. ARIC spirometry measurements were made with a Collins Survey II water-seal spirometer (Collins Medical, Inc.) and Pulmo-Screen II software (PDS Healthcare Products, Inc.). Genotyping was done using the Affymetrix GeneChip SNP Array 6.0 and imputation was performed using MACH. Quality control steps for genotyping data included exclusions for call rate <95%, minor allele frequency <1%, HWE P<10-5, no chromosomal location, suspected first-degree relative of an included individual based on genotype data, or more than 8 standard deviations for any of the first ten principal components. The current analysis includes 9,078 Caucasian subjects with genotyping data, pulmonary function measures and complete covariate information.

Details of the **British 1958 Birth Cohort** biomedical follow-up have been previously reported[107] and a full technical report is available online (http://www.b58cgene.sgul.ac.uk/report.php). Spirometry at age 44–45 years was done in the standing position without nose clips, using a Vitalograph handheld spirometer as previously described[108,109]. In the analysis, all readings with a best-test variation greater than 10% were excluded.

The **Busselton Health Study** (BHS) is a longitudinal survey of the town of Busselton in the south-western region of Western Australia that began in 1966. In 1994/1995 a cross-sectional community follow-up study was undertaken where blood was taken for DNA extraction. A sample of 1,168 European-ancestry individuals were genotyped using the Illumina 610-Quad BeadChip (BHS1), and subsequent genotyping was carried out on an independent group of 3,038 European-ancestry individuals (BHS2). Spirometric measures of forced expired volume in one second (FEV1) and forced vital capacity (FVC) were assessed as described previously[110,111].

The **CHS** is a population-based cohort study of risk factors for CHD and stroke in adults > 65 years conducted across four U.S. field centers[112]. CHS spirometry measurements were made with a Collins Survey I water-seal spirometer (Collins Medical, Inc.) and software from S&M Instruments. 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 self-identified white participants remained. Of these, 3,140 had pulmonary function measures and complete covariate information.

Nature Genetics: doi:10.1038/ng.941

The **CROATIA-Korcula** study is a family-based, cross-sectional study in the isolated island of Korcula that included about 1,000 examinees aged 18-95. Spirometry was performed in the sitting upright position without nose clips, using a Jaegger Toennisen spirometer. Three readings were taken for each examinee, at least 15 seconds apart[113]. Population stratification was partially taken into account after adjusting the measures for the 3 first principal components drawn from the population genomic kinship matrix computed using the ibs function of the GenABEL[114]. Additional correction of the standard error of effect estimates was undertaken by genomic control[115]. Effect estimates were obtained using the palinear function of the ProbABEL package.

The **CROATIA-Vis** study is a family-based, cross-sectional study in the isolated island of Vis that included about 1,000 examinees aged 18-93. Spirometry was performed in the sitting upright position without nose clips, using a Jaegger Toennisen spirometer. Three readings were taken for each examinee, at least 15 seconds apart[113]. Population stratification was partially taken into account after adjusting the measures for the 3 first principal components drawn from the population genomic kinship matrix computed using the ibs function of the GenABEL[114]. Additional correction of the standard error of effect estimates was undertaken by genomic control[115]. Effect estimates were obtained using the palinear function of the ProbABEL package.

Details of the methods of **ECRHS I** and **ECRHS II**, a multicentre international cohort study, have been published elsewhere[116,117]. Participants within the ECRHS were eligible for inclusion in this analysis if they were identified by random sampling of those who fulfilled the following criteria 1) lived in centres that took part in genome-wide genotyping initiative under the auspices of GABRIEL[118] AND 2) were initially selected to take part in the ECRHS clinical measurements as part of the random sample (ie not specifically selected for inclusion because of any pre-existing disease). Participants were included in this analyses if they also provided a technically satisfactory forced expiratory manoeuvre, compliant with ATS spirometry criteria, at the time of the first survey (aged 20-48). Most centres used the BIOMEDIN water-sealed spirometer for lung function measures and all centres conducted manoeuvres in the sitting position with nose clips on. Further details are available in[119].

The **EPIC Norfolk** GWA cohort includes 2,566 participants randomly selected from the EPIC-Norfolk Study, a population-based cohort study of 25,663 men and women of European descent aged 39-79 years recruited in Norfolk, UK between 1993 and 1997[109]. Respiratory function was assessed by spirometry[120,121]. Forced expiratory volume in 1 second (FEV1) and forced vital capacity (FVC) were measured twice using a portable spirometer (Micro Medical, Rochester, United Kingdom). The higher values of the two readings for FEV1 and FVC were used for the analyses.

The **FHS** is a longitudinal community-based family study that originated in 1948 with the recruitment of adults from the town of Framingham, MA. The offspring of the original cohort were recruited to participate in 1971, and the third generation (Gen3) was recruited starting in 2002. Spirometry has been measured on all three generations of the participating families as part of the clinical examinations. For the original and offspring cohorts, spirometry was performed using a 6-L water-filled Collins survey spirometer connected to an Eagle II microprocessor (Collins Medical, Braintree, MA) or in later offspring examinations to a personal computer running software developed by S&M Instruments, Doylestown, PA. For the Gen3 cohort, spirometry was performed using the CPL System (Collins Medical, Braintree, MA) with a dry rolling-seal spirometer. All of these systems provided an automatic correction for body temperature and pressure, saturated, and provided quality assurance information. Each Framingham participant's latest examination with acceptable pulmonary function data was used in this analysis. Eligible examinations included original cohort examinations 19, 17, 16 and 13; offspring cohort examinations 7, 6, 5, and 3; and Gen3 cohort examination 1. A total of 9,274 subjects were genotyped and exclusions were based on a call rate < 97%, heterozygosity more than 5 standard deviations from the mean, or excessive non-inheritance. A total of 8,481 participants passed the genotyping quality control criteria. Of these, 7,914 participants with complete spirometry and covariate data were used in the final analysis data set.

Nature Genetics: doi:10.1038/ng.941

The subjects for the **Finnish Twin Cohort** (FTC) came from two clinical substudies of twins in Finland: 1. The FinnTwin16 study is a nationwide longitudinal cohort study of health behaviors in twins and their families[122]. The participants were 34 MZ twin pairs, who were a subset of 347 twins recruited from the FinnTwin16 study of five consecutive birth cohorts (1975–1979) of twins (n = 4929 individuals), and who participated in a clinical study of body composition and cardiorespiratory fitness. The study protocol included a bicycle ergometer exercise test with gas exchange analysis. A mass flow sensor of the gas exchange device (Sensormedics Co) was used to measure forced expiratory volume in one second and forced vital capacity before exercise. 2. The Finnish Twin Study on Ageing (FITSA). Participants were recruited from the older Finnish Twin Cohort for a clinical study of functional limitations in older women. Clinical assessment was conducted in 2000-2001 at the University of Jyväskylä. The final sample consisted of 103 monozygotic (MZ) and 114 dizygotic (DZ) twin pairs. Lung function was measured in the standing position using an electronic spirometer (Medikro 202, Kuopio, Finland). The subject was asked to inhale maximally and to exhale as fast as possible into a mouthpiece connected to a flow transducer and a flow-volume curve was created. At least two tests were performed and the best result taken for the analyses. Spirometer was calibrated daily with a three-litre pump and was accurate to within 1%. Both substudies were mainly genotyped as part of the GenomEUtwin project of female MZ pairs [123]. To avoid taking into account the statistical dependence of two related individuals in a family, we selected one twin at random for inclusion in the analysis if both of the MZ twins in a pair had the phenotype.

The Health Aging and Body Composition (**Health ABC**) study is a prospective observational cohort of well functioning individuals aged 70–79 years. The Health ABC study recruited 3,075 community-dwelling African and European Americans, men and women, at two field centers at the University of Pittsburgh, Pennsylvania and the University of Tennessee, Memphis. Spirometry was performed with a horizontal dry rolling seal spirometer (SensorMedics Corporation, Yorba Linda, CA) connected to a computer. Pulmonary function testing followed ATS guidelines for the standardization of spirometry, and is described in detail elsewhere[124]. Health ABC genotyped 1,794 self-described white participants at baseline with available DNA and consent to genetics testing; of these 1,661 passed quality control benchmarks (call rate > 97%, no sex mismatch, and cryptic relatedness), and 1,472 had pulmonary function measurements and complete data on covariates.

The DNA archive established from the **Health 2000** Survey Cohort was used. Details of this study population and phenotyping procedures have been previously reported[125]. Spirometry was done in the standing position without nose clips, using a Vitalograph 2150 spirometer. In the analysis, the maximum permissible difference between the two highest FEV1 and FVC values was 10%.

The **KORA F4** study (Cooperative Health Research in the Region of Augsburg) is a follow-up study to the KORA-Survey 2000 (S4, 10/1999 – 7/2001)[126]. Lung function tests were performed in a subsample of the KORA F4 cohort corresponding to a random population sample of subjects born between 1946 and 1965 (age range 41 – 63 years). A total of 1321 individuals (618 men, 703 women) were examined. Age was 51.6 ± 5.7 years (mean ± SD), weight was 79.2 ± 16.8 kg, and height was 1.70 ± 0.09 m. Spirometry was performed in line with the ATS recommendations [127] using a pneumotachograph-type spirometer (Masterscreen PC, CardinalHealth, Würzburg, Germany) before and after inhalation of 200 µg salbutamol. The spirometer was calibrated daily using a 1L-calibration pump (CardinalHealth, Würzburg, Germany), and additionally, an internal control was used to ensure constant instrumental conditions. The participants performed at least three forced expiratory lung function manoeuvres in order to obtain a minimum of two acceptable and reproducible values. The final values of FVC and FEV1 were determined based on the best maneuver as defined by the highest sum of FVC and FEV1. For participants who did not manage to perform at least two acceptable and reproducible values out of nine trials no lung function measurements were recorded. The presence of acute or chronic respiratory diseases as well as medication was assessed by a standardized questionnaire.

Nature Genetics: doi:10.1038/ng.941

The **KORA S3** study (Cooperative Health Research in the Region of Augsburg) is also known as the third MONICA survey, which was performed 1994-95 in Augsburg, Germany. The objective and protocol of this survey have been published earlier[126]. The third MONICA survey consisted of a random sample of all registered residents of the city of Augsburg aged 25–74 years. For all participants younger than 60, who did not smoke or use inhalers one hour before the test, FEV1 and FVC were determined by spirometry in 1997-98. All spirometric tests were performed strictly adhering to the ECRHS protocol [116]. Tests were accounted valid if at least two technically satisfactory manoeuvres could be obtained throughout a maximum of nine trials. Based on the best manoeuvre, which was defined as the trail with the highest sum of FVC and FEV1, the values of FVC and FEV1 were determined.

The **Northern Finland Birth Cohort** (NFBC) study programme was initiated in the 1960s. The cohort of women and their offspring was established in the provinces of Oulu and Lapland and had an expected date of birth in 1966 comprising 12231 children (NFBC1966, 15). The NFBC1966 had spirometry and other measurements done at the age of 31 years. In NFBC1966, we used a Vitalograph P-model spirometer (Vitalograph Ltd., Buckingham, UK), with a volumetric accuracy of ±2% or ±50 mL whichever was greater. The spirometer was calibrated regularly using a 1-Litre precision syringe. The spirometric manoeuvre was performed three times but was repeated if the coefficient of variation between two maximal readings was >4%.

The **Orkney Complex Disease Study** (ORCADES) is an ongoing family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. Spirometry was performed in the sitting position without nose clips, using a Spida handheld spirometer. Measurements were repeated once and the better reading was used for analysis. Population stratification was partially taken into account after adjusting the measures for the 3 first principal components drawn from the population genomic kinship matrix computed using the ibs function of the GenABEL[114]. Additional correction of the standard error of effect estimates was undertaken by genomic control[115]. Effect estimates were obtained using the palinear function of the ProbABEL package.

The **Rotterdam Study** is a prospective population-based cohort study founded in 1990 in a suburb of Rotterdam, the Netherlands. The first cohort (RS I) consists of 7,983 participants, aged 55 years and over. The second cohort (RS II) was recruited in 2000 with the same inclusion criteria. Performing of spirometry was introduced in 2004. Spirometry was performed by trained paramedical personnel using a SpiroPro® portable spirometer (Erich Jaeger, Hoechberg, Germany), according to American Thoracic Society (ATS)/European Respiratory Society (ERS) guidelines. FEV1, FVC and FEV1/FVC ratio were measured. A total of 6,240 subjects were genotyped in RS I and a total of 2,516 in RS II. Exclusions were based on a call rate < 98%, Hardy-Weinberg p-value < 10^{-6} and MAF < 0.01%. A total of 5,974 for RS I and 2,157 for RS II passed genotyping quality control. Since spirometry was introduced in only 2004, full data with regard to spirometry and covariate data was available in 1,224 participants and 852 for RS II and were used in the final analysis dataset. Further details can be found in Hofman A et al., 2009[128] and Estrada K et al., 2009[129].

The **Study of Health in West Pomerania** (SHIP) is a cross-sectional, population based survey in a region in the Northeast of Germany. Study details are given elsewhere[130,131]. The examinations were conducted using a bodyplethysmograph equipped with a pneumotachograph (VIASYS Healthcare, JAEGER, Hoechberg, Germany) which meets the American Thoracic Society (ATS) criteria[132]. The volume signal of the equipment was calibrated with a 3.0 litre syringe connected to the pneumotachograph in accordance with the manufacturer´s recommendations and at least once on each day´s testing. Barometric pressure, temperature and relative humidity were registered every morning. Calibration of reference gas and volume was examined under ATS-conditions (Ambient Temperature Pressure) and the integrated volumes were BTPS (Body Temperature Pressure Saturated) corrected[132,133]. Lung function variables were measured continuously throughout the baseline breathing and the forced manoeuvres using a VIASYS HEALTHCARE system (MasterScreen Body/Diff.). Spirometry flow volume loops were conducted in accordance with ATS recommendations[133] in a sitting position and with wearing noseclips. The participants performed at least three forced expiratory lung function manoeuvres in order to obtain a minimum of two acceptable and reproducible values[134]. Immediate on-screen error codes indicating the major acceptability (including

Nature Genetics: doi:10.1038/ng.941

start, duration and end of test) and reproducibility criteria supported the attempt for standardised procedures. The procedure was continuously monitored by a physician. The best results for forced vital capacity (FVC), forced expiratory volume in one second (FEV 1), peak expiratory flow (PEF) and expiratory flow at 75%, 50%, 25% of FVC (MEF 75, MEF 50, MEF 25) were taken. The ratio of FEV 1 to FVC was calculated from the largest FEV 1 and FVC.

The **TwinsUK** cohort consisted of a group of twins ascertained to study the heritability and genetics of age-related diseases (www.twinsUK.ac.uk). These unselected twins were recruited from the general population through national media campaigns in the UK and shown to be comparable to age-matched population singletons in terms of disease-related and lifestyle characteristics[135]. Spirometry (Vitalograph model 2150, Buckingham, England) was conducted at the clinical centre during a visit. Twins were instructed before the test and forced vital capacity (FVC) manoeuvres were performed in a standing position, without the use of nose clips. Three manoeuvres were performed and the maximum obtained values for FEV1 were obtained. Because of the relatedness in the TwinsUK cohort, we utilized the GenABEL software package[114] which incorporates a pair-wise kinship matrix calculated using genotyping data in the polygenic model to correct for relatedness and hidden population stratification. The score test implemented in the software was used to test the association between a given SNP and the lung function phenotypes. Two sets of the samples (TwinsUK-I and TwinsUK-II) derived from the TwinsUK were genotyped. The first set – TwinsUK-I, was genotyped using the Infinium assay (Illumina, San Diego, USA) with four fully compatible SNP arrays as previously described[136]. Normalised intensity data were pooled and genotypes called on the basis of the Illluminus algorithm[137]. No calls were assigned if the most likely call was less than a posterior probability of 0.95. Validation of pooling was done by visual inspection of 100 random, shared SNPs for overt batch effects; none were observed. Quality checks similar to those for the WTCCC Study [138] were applied. The second set – TwinsUK-II, were typed with the Infinium 610k assay (Illumina, San Diego, USA) at two different centres, namely the Centre for Inherited Diseases Research (USA) and the Wellcome Trust Sanger Institute. The same pooling procedure used for the TwinsUK-I was applied to the TwinsUK-II. Further, we excluded SNPs that had a low call rate (<95%), Hardy-Weinberg p values < 10−4 and minor allele frequencies < 1%. We also removed subjects if the sample call rate was less than 95%, autosomal heterozygosity was outside the expected range, genotype concordance was over 97% with another sample and the sample was of lesser call rate, non-Caucasian ancestry either self-identified or identified by cluster analysis in STRUCTURE [139] or multidimensional scaling by comparison to the three HapMap phase 2 reference populations (CEU, YRI, CHB+JPT), or unexplained relatedness (estimated proportion of allele shared identical by descent >0.05) to >120 other samples where genotyping failed for >2 % of SNPs [140]. The overall genotyping efficiency of the GWA was 98.7 %. Imputation of genotypes was carried out using the software IMPUTE [141] for both TwinsUK-I and TwinsUK-II.

Stage 2 follow-up samples and phenotype measurement

The **ADONIX** study was established in 2001, and for the period 2001-2004 the study base was established in cooperation with the INTERGENE study. Details of the ADONIX study[142,143] and INTERGENE study[144,145] have been previously reported. Spirometry was done in the standing position with nose clips, using a Vitalograph dry-wedge spirometer as previously described[143]. All readings fulfilled the ATS requirements for standardization of spirometry.

Details of the **British Regional Heart Study** (BRHS) design have been previously reported [146]. Spirometry was done at age 60-79 years. A minimum of 3 measurements of FVC and FEV1 were obtained standing and without noseclips, using a Vitalograph Compact II instrument (Vitalograph Ltd, Buckingham, United Kingdom) as previously described[147]. FEV1 and FVC were recorded for the best test, which was defined in accordance with American Thoracic Society recommendations[148].

Full details of the selection of participants and measurements used in the British **Women's Heart and Health Study** (BWHHS) have been previously reported[149,150]. Lung function tests were carried out using a digital meter Vitalograph with an attached printout of forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1), peak expiratory flow (PEF) and forced mid expiratory flow rate (FEF25–75). The Vitalograph was calibrated each day using a 1 litre syringe and automated so that results were adjusted for ambient temperature. The women then performed some

Nature Genetics: doi:10.1038/ng.941

practice efforts. They were then required to perform a minimum of three reproducible FVC measures (within 5% of maximum FVC produced). The output that produced the highest sum of FVC and FEV1 were used in the analyses. Women who could not perform three reproducible measures or who were unable to attempt the lung function assessment were excluded.

Cross-sectional analyses of data from year 0 examination of the **Coronary Artery Risk Development in Young Adults** (CARDIA) cohort was performed. During 1985-1986, CARDIA randomly recruited 5,115 black and white men and women, aged 18 to 30 years, from the general population at Birmingham, Alabama; Chicago, Illinois; and Minneapolis, Minnesota; and from the membership of the Oakland Kaiser-Permanente Health Plan in Oakland, California. Detailed methods, instruments, and quality control procedures are described at the CARDIA website (http://www.cardia.dopm.uab.edu/ex_mt.htm) and in other published reports[151,152]. Spirometric pulmonary function testing were performed using the Collins survey 8-liter water-sealed spirometer and the Eagle II microprocessor (Warren E. Collins, Inc., Braintree, MA) in a sitting position with noseclips, as per the 1979 American Thoracic Society criteria[153]. Specifically, each subject performed a minimum of three trials with expirations recorded to the FVC plateau, which occurs after six seconds of expiration in adult males and was maintained for at least one second before terminating the forced expiratory maneuver. If, at the end of the three trials, there were at least three acceptable tracings, and with the maximum FVC and FEV1 reproduced to within 5% or 100 mL, whichever is greater, no more trials were performed.

The **CROATIA-Split** study,is an ongoing population-based, cross-sectional study in the Dalmatian City of Split for which 499 of the examinees, aged 18-95, have genotype data available. Spirometry was performed in the sitting upright position without nose clips, using a Jaegger Toennisen spirometer. Three readings were taken for each examinee, at least 15 seconds apart 12. Population stratification was partially taken into account after adjusting themeasures for the3 first principal components drawn from the population genomic kinship matrix computed using the ibs function of the GenABEL package. Additional correction of the standard error of effect estimates due to related individuals was undertaken using the mmscore function in GenABEL Effect estimates were obtained using the palinear function of the ProbABEL package.

The **Gedling** cohort is a general population sample recruited in Nottingham in 1991 (18 to 70 yr of age, n=2,633)[154] and was then followed-up in 2000 (n=1346) when blood samples were taken for DNA extraction and FEV1 and FVC were measured using a calibrated dry bellows spirometer (Vitalograph, Buckingham, UK), recording the best of three satisfactory attempts[155].

Details of the **Generation Scotland: Scottish Family Health Study** (GS:SFHS) population and phenotyping procedures have been previously reported[156]. Spirometry was performed without nose clips using the Ndd Easy One Spirometer (Model 2001). Spirometry was performed a maximum of 4 times until satisfactory readings were obtained, all readings with a best test variation of greater than 10% were excluded. Generalised Estimating Equations (GEE) with robust standard errors were used to adjust for familial correlations.

The **Hertfordshire Cohort Study** (HCS) comprises a nationally unique study of 3,000 men and women born during the period 1931-39 in the English county of Hertfordshire and still resident there[157]. Lung function was measured using a Micro Spirometer (Micro Medical Ltd) in the seated position without noseclips. After at least one practice blow, three FEV1 and FVC readings were recorded. The highest FEV1 and FVC values from satisfactory manoeuvres were used in the analyses - these did not necessarily come from the same blow. DNA on all participants has been collected and is stored in the MRC Epidemiology Resource Centre, University of Southampton.

Details of the **Lothian Birth Cohort 1936** have been previously reported[158]. Lung function assessing peak expiratory flow rate, forced expiratory volume in 1 s, and forced vital capacity (each the best of three), using a Micro Medical Spirometer was assessed, sitting down without nose clips, at age 70 years. The accuracy of the spirometer is ±3% (to ATS recommendations Standardisation of Spirometry 1994 update for flows and volumes).

Nature Genetics: doi:10.1038/ng.941

LifeLines is an observational follow-up study in a large representative sample of the population of the northern provinces of the Netherlands covering three generations[159]. Firstly, a random sample of persons aged between 25 and 50 years are invited to participate. Subsequently their family members if present are invited to also take part (parents, partner, parents in law, children), resulting in a three-generation study. The core of the LifeLines project consists of dedicated data collection and biological sample storage, including genetic samples (''biobank''). All participants receive a number of questionnaires and a basic medical examination and are followed for many years with extensive standardized measurements. Spirometry was performed following ATS guidelines using a Welch Allyn Version 1.6.0.489, PC-based SpiroPerfect with Ca Workstation software. Genotyping was performed on 3078 unrelated individuals with IlluminaCytoSNP-12 arrays[160].

Details of the **Multi-Ethnic Study of Atherosclerosis** (MESA) and MESA Lung Study have been previously reported [161,162] and a full technical report of MESA procedures is available online (www.mesa-nhlbi.org). Spirometry was conducted in 2004-2006 in accordance with the 2005 ATS/ERS recommended guidelines[127] using a dry-rolling-sealed spirometer with automated quality control software (Occupational Marketing, Inc. Houston, TX). All spirometry exams were centrally reviewed by at least one (JLH) of the authors[163].

The **Nottingham Smokers** cohort is an ongoing collection in Nottingham using the following criteria; Caucasian, > 40 years of age and smoking history of > 10 pack years (currently n=538). Lung function measurements (FEV1 and FVC) were recorded at enrolment using a MicroLab ML3500 spirometer (Micro Medical Ltd, UK) recording the best of three satisfactory attempts.

The DNA archive established from the MRC National Survey of Health and Development (**NSHD**, also known as the British 1946 Birth Cohort) was used. Details of this study population have been previously reported[164]. In brief, NSHD is an ongoing prospective birth cohort study consisting of a stratified sample of all births in England, Scotland and Wales in one week in March 1946. The cohort is ethnically homogeneous and initially consisted of 2,547 females and 2,815 males who have been followed up over 20 times since their birth, mostly by means of home interviews. During 1999 when cohort members were all aged 53 years, all those still in contact with the study team and living in the UK were invited to be examined by a trained research nurse at home: 3,035 were successfully contacted[164] and adequate blood samples for DNA extraction obtained from 2,756. Spirometry at age 53 years was done in the standing position without nose clips, using a Micromedical turbine electronic spirometer (Cardinal Health UK 232 Ltd, Basingstoke, UK) as previously described[165,166]. Two measurements were taken and the maximum reading with recorded satisfactory technique (as assessed by the nurse conducting the test) was used for these analyses.

The **Rotterdam Study** is a prospective population-based cohort study founded in 1990 in a suburb of Rotterdam, the Netherlands. The third cohort (RS IIII) consists of 3,932 participants, aged 45 years and over and was recruited in 2006. Spirometry was performed by trained paramedical personnel using a SpiroPro® portable spirometer (Erich Jaeger, Hoechberg, Germany), according to American Thoracic Society (ATS)/European Respiratory Society (ERS) guidelines. FEV1, FVC and FEV1/FVC ratio were measured. A total of 2,420 subjects were genotyped in RS II. Exclusions were based on a call rate < 98%, Hardy-Weinberg p-value < 10^{-6} and MAF < 0.01%. A total of 2,082 passed genotyping quality control. Of these, 1,247 participants with complete spirometry and covariate data were used in the final analysis data set. Further details can be found in Hofman A et al., 2009[128] and Estrada K et al., 2009[129].

Details of the **SAPALDIA** Cohort have been previously reported[118,167-169] and technical details are available online (http://www.sapaldia.net/en/content/view/88/305/). SAPALDIA is a cohort study initiated in 1991 recruiting adults aged 18-60 years from the general population in eight areas of Switzerland. A first follow-up examination was done in 2002 and included the establishment of a DNA- and blood biobank, a second follow-up is currently under way. All participants were administered a detailed health questionnaire. Spirometry was done using a Sensormedics 2200

Nature Genetics: doi:10.1038/ng.941

spirometer and complying to the ATS quality criteria as previously described [167,170]. High quality genotype data was available on up to 5646 participants with complete spirometry and smoking information at baseline.

Look up samples and phenotype measurement

The **Avon Longitudinal Study of Parents and their Children** (ALSPAC) is a population-based birth cohort study consisting initially of over 13 000 women and their children recruited in the county of Avon, UK in the early 1990s[171]. At 8–9 years of age, children's lung function was measured by spirometry (Vitalograph 2120, Maids Moreton, UK) according to American Thoracic Society criteria. Flow-volume curves were reviewed by a respiratory physician to ensure adherence to standards. GWA data and lung function measurements were available on 5062 of the children from this study.

The **GIANT consortium** study description can be found in Lango Allen, H. et al[90].

The **International Lung Cancer Consortium** (ILCC) GWAS meta-analysis was conducted based on 14 lung cancer genomewide association studies with a total of 13,300 lung cancer patients and 19,666 controls[37]. The GWAS studies included in this meta-analysis are the Environment and Genetics in Lung Cancer Etiology (EAGLE), the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC), the Prostate, Lung, Colon, Ovary Screening Trial (PLCO), the Cancer Prevention Study II Nutrition Cohort (CPS-II), the UK study from the Institute for Cancer Research, the International Agency for Research on Cancer (IARC) study in central Europe, the MD Anderson study in Texas, the DeCODE study in Iceland, the Helmholtz-Gemeinschaft Deutscher Forschungszentren (HGF) study in Germany, the Carotene and Retinol Efficacy Trial (CARET) cohort, the HUNG2/Tromso study, the Canadian study from the Samuel Lunenfeld Research Institute, the lung cancer study in France and lung cancer study in Estonia. The detailed study methods and descriptions have been published previously[37].

The **Oxford-GlaxoSmithKline study** (OX-GSK) is a collaborative effort to investigate the genetic basis of smoking-related behavioral traits[89]. The study combines genome-wide chip and imputed genotype from 20 disease, population and control cohorts totaling 41,150 samples. Three smoking related behavioural traits were measured on these samples (a) a smoking quantity phenotype, defined as a semi-quantitative trait based on the self-reported variable of cigarettes smoked per day (N = 15,574) , (b) a smoking initiation binary phenotype, defined as Ever smoker (N=18,598) versus Never smoker (N=15,041), and (c) a smoking cessation phenotype, defined as Current smoker (N=10,123) versus Non-current smoker (N=19,903). Association analyses separately within each cohort under an additive model using covariate effects for age, sex, disease case or control status where applicable, and other cohort-specific covariates. A meta-analysis was then carried out by combining study-specific regression coefficient estimates using a fixed effects model.

Recruitment of the Western Australian Pregnancy (Raine) cohort has previously been described in detail[172-174]. In brief, between 1989 and 1991 2,900 pregnant women were recruited prior to 18-weeks gestation into a randomised controlled trial to evaluate the effects of repeated ultrasound in pregnancy. Children have been comprehensively phenotyped from birth to 21 years of age (average ages of one, two, three, six, eight, ten, fourteen, seventeen and twenty-one) by trained members the Raine research team. Most of the children are of Caucasian ethnicity. Data collection included questionnaires completed by the child's primary care and by the adolescent from age 14, physical assessments by trained assessors at all follow up years, DNA collection from year 14 follow-up. The study was conducted with appropriate institutional ethics approval, and written informed consent was obtained from all mothers. Study individual genotype data was extracted from the genome-wide Illumina 660 Quad Array.

Nature Genetics: doi:10.1038/ng.941

Investigators, contributions, funding and acknowledgements per study

SpiroMeta consortium

Analysis group: María Soler Artigas, Louise V Wain, Martin D Tobin (co-chair), Ian P Hall (co-chair), Toby Johnson.

Expression profiling and bioinformatics group: Ma'en Obeidat, Ian Sayers, Ian P Hall, Louise V Wain, María Soler Artigas, Martin D Tobin

Writing group: María Soler Artigas, Louise V Wain, Martin D Tobin, Ian P Hall, Ma'en Obeidat, David P Strachan, Paul Elliott.

Funding & Acknowledgements:
M.D.T. has been supported by MRC fellowships G0501942 and G0902313. The meta-analysis was undertaken using the University of Leicester's **ALICE**; a High Performance Computing (HPC) cluster, purchased through a £2million HEFCE CIF award. The Universities of Leicester and Nottingham acknowledge receipt of a Collaborative Research and Development grant from the Healthcare and Bioscience iNet, a project funded by the East Midlands Development Agency (*emda*), part-financed by the European Regional Development Fund and delivered by Medilink East Midlands. The SpiroMeta consortium would also like to acknowledge support from Medisearch, The Leicester Medical Research Foundation, via a grant to the University of Leicester. P.E. is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at Imperial College Healthcare NHS Trust. P.E. is an NIHR Senior Investigator.

CHARGE consortium

Analysis group: Daan W Loth, Nora Franceschini , Jemma B Wilk, Kristin D Marciante, Wenbo Tang, Albert V Smith, Stephanie J. London

Expression profiling and bioinformatics group: Sina Gharib, Daan W Loth

Writing group: Daan W Loth, Sina A Gharib, Wenbo Tang, Bruno H Ch Stricker, George T O'Connor, Vilmundur Gudnason, Stephanie J London

Funding & Acknowledgements: The CHARGE Pulmonary Working Group would like to acknowledge funding from the NHLBI (HL105756) and the CHARGE consortium for its organizational support.

Stage 1 GWAS studies

AGES

Investigators: Thor Aspelund, Gudny Eiriksdottir, Melissa Garcia, Gauti Kjartan Gislason, Vilmundur Gudnason, Tamara B Harris, Lenore J Launer, Albert Vernon Smith.

Contributions: project conception, design and management: G.E., M.G., V.G., T.B.H., L.J.L., phenotype collection and data management: T.A., data analysis: G.K.G., A.V.S. **Funding & Acknowledgements**: This study has been funded by NIH contract N01-AG-1-2100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063. The researchers are indebted to the participants for their willingness to participate in the study.

ARIC

Investigators: David J Couper, Nora Franceschini, Dana B Hancock, Bonnie R Joubert , Laura R Loehr, Stephanie J London, Alanna C Morrison, Kari E North.

Contributions: project conception, design and management: S.J.L., N.F., L.R.L., D.J.C., D.B.H., B.R.J., A.C.M., K.E.N., phenotype collection and data management: D.J.C., N.F., L.R.L., A.C.M., K.E.N., data analysis: N.F., D.B.H., L.R.L.

Nature Genetics: doi:10.1038/ng.941

Funding & Acknowledgements: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), 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. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. NIEHS: Supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (ZO1 ES49019).

B58C – T1DGC

Investigators: David P Strachan, Alicja R Rudnicka, Wendy L McArdle

Contributions: project conception, design and management: D.P.S., phenotype collection and data management: A.R.R., D.P.S., genotyping: W.L.M., data analysis: A.R.R., D.P.S. **Funding & Acknowledgements:** This research utilizes resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge (John Todd, Helen Stevens and Neil Walker), which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). We acknowledge use of the DNA from the British 1958 Birth Cohort collection, funded by the Medical Research Council and Wellcome Trust. We thank the Avon Longitudinal Study of Parents and Children laboratory in Bristol and the British 1958 Birth Cohort team, including S. Ring, R. Jones, M. Pembrey, W. McArdle, D.P.Strachan and P. Burton for preparing and providing the control DNA samples.

B58C -WTCCC

Investigators: Wendy L McArdle, Alicja R Rudnicka, David P Strachan.

Contributions project conception, design and management: D.P.S., phenotype collection and data management: A.R.R., D.P.S., genotyping: W.L.M., data analysis: A.R.R., D.P.S. **Funding & Acknowledgements:** The British 1958 Birth Cohort DNA collection was funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the Wellcome Trust Case Control Consortium was funded by the Wellcome Trust grant 076113/B/04/Z.

BHS1

Investigators: Alan L James, Arthur W Musk, Lyle J Palmer, Nicole M Warrington.

Contributions: project conception, design and management: A.L.J., A.W.M., L.J.P., phenotype collection and data management: A.L.J., A.W.M., L.J.P., genotyping: A.L.J., A.W.M., L.J.P., data analysis: N.M.W. **Funding & Acknowledgements:** The Busselton Health Study acknowledges the generous support for the 1994/5 follow-up study from Healthway, Western Australia. The Busselton Health Study is supported by The Great Wine Estates of the Margaret River region of Western Australia. The BHS gratefully acknowledges the assistance of the Western Australian DNA Bank (NHMRC Enabling Facility) with DNA samples and the support provided by The Ark for this study.

CHS

Investigators: Sina A Gharib, Susan R Heckbert, Thomas Lumley, Kristin D Marciante, Bruce M Psaty, Jerome I Rotter.

Contributions: phenotype collection and data management: S.A.G., S.R.H., T.L., B.M.P., genotyping: S.R.H., B.M.P, J.I.R., data analysis: K.D.M., J.I.R. **Funding & Acknowledgements:** The CHS research reported in this article was supported by NHLBI contracts N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, and grants HL080295, HL087652, HL105756 and HL075336 with additional contribution from the NINDS. Additional support from the NIA was provided through AG-023269, AG-15928, AG-20098, and AG-027058. 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 theCedars-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.

CROATIA-Korcula

Investigators: Harry Campbell, Ivica Grkovic, Caroline Hayward, Jennifer E Huffman, Stipan Jankovic, Ozren Polašek, Igor Rudan, Veronique Vitart , Alan F Wright, Lina Zgaga.

Contributions: project conception, design and management: H.C., I.G., S.J., I.R., A.F.W., L.Z. phenotype collection and data management: I.G., S.J., O.P., I.R., L.Z., data analysis: C.H., J.E.H, V.V. **Funding & Acknowledgements:** The CROATIA-Korcula study was funded by grants from the Medical Research Council (UK) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (Vis) (108-1080315-0302).

CROATIA-Vis

Investigators: Harry Campbell, Caroline Hayward, Ozren Polašek, Igor Rudan, Veronique Vitart, Alan F Wright.

Contributions: project conception, design and management: H.C., C.H., O.P., I.R., A.F.W., phenotype collection and data management: H.C., C.H., O.P., I.R., A.F.W., genotyping: C.H., I.R., A.F.W., data analysis: C.H, V.V. **Funding & Acknowledgements:** The CROATIA-Vis study was funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). The Vis study would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health.

ECRHS

Investigators: Deborah L Jarvis, Ernst Omenaas, Isabelle Pin, Matthias Wjst.

Contributions: phenotype collection and data management: D.L.J., E.O., I.P, M.W., genotyping: M.W., data analysis: D.L.J., A.R.. **Funding & Acknowledgements:** EU funding (GABRIEL GRANT Number: 018996, ECRHS II Coordination Number: QLK4-CT-1999-01237).

EPIC

Investigators: Inês Barroso, Ruth JF Loos, Nicholas J Wareham, Jing Hua Zhao.

Contributions: project conception, design and management: N.J.W., phenotype collection and data management: N.J.W., genotyping: I.B., R.J.F.L., J.H.Z., data analysis: J.H.Z. **Funding & Acknowledgements:** The EPIC Norfolk Study is funded by Cancer Research United Kingdom and the Medical Research Council. EPIC would like to thank Manjinder S. Sandhu (Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge,

Nature Genetics: doi:10.1038/ng.941

Cambridge CB2 0SR, UK) for contributions to project conception, design and management and genotyping. I.B. acknowledges funding from the Wellcome Trust (077016/Z/05/Z).

FHS

Investigators: Jemma B Wilk, George T O'Connor.

Contributions: phenotype collection and data management: J.B.W., G.T.O., data analysis: J.B.W. **Funding & Acknowledgements:** G.T.O. acknowledges NIH/NHLBI (N01 HC 25195). J.B.W. received salary support from the Flight Attendant Medical Research Institute (FAMRI).

FTC

Investigators: Jaakko Kaprio, Kirsi H Pietiläinen, Taina Rantanen, Ida Surakka,

Contributions: project conception, design and management: J.K., K.H.P., T. Rantanen., phenotype collection and data management: J.K., K.H.P., T. Rantanen., genotyping: J.K., data analysis: I. Surakka. **Funding & Acknowledgements:** K. H. P. acknowledges Novo Nordisk, Jalmari and Rauha Ahokas, Gyllenberg Foundations and Finnish Foundation for Cardiovascular Research. T. R. acknowledges Academy of Finland (75507). J. K. acknowledges The European Community´s Seventh Framework Programme (FP7/2007-2013), ENGAGE Consortium (HEALTH-F4-2007- 201413), The European Community´s Fifth Framework Programme, GenomEUtwin (QLG2-CT-2002-01254) and Academy of Finland, Center of Excellence in Complexe Disease Genetics (213506 and 129680).

Health ABC

Investigators: Melinda C Aldrich, Patricia A Cassano , Tamara B Harris, Stephen B Kritchevsky, Yongmei Liu, Kurt Lohman, Bernd Meibohm, Wenbo Tang.

Contributions: project conception, design and management: M.C.A., P.A.C., T.B.H., S.B.K., Y.L., B.M., phenotype collection and data management: P.A.C., B.M., W.T., genotyping: Y.L., K.L, data analysis: P.A.C., Y.L., K.L., W.T. **Funding & Acknowledgements:** B.M. acknowledges NIH/NIA (N01-AG-6-2101, N01-AG-6-2103, N01-AG-6-2106), and NIH/NHLBI (R01-HL-074104) for spirometry funding. M.C.A. acknowledges NIH/NHLBI (HL088133), NIH/NIEHS (ES015794) and FAMRI (Flight Attendants Medical Research Institute). S.B.K. acknowledges NIH/NIA (RC1 AG035835) for analysis time funding. Y.L. and S.B.K. acknowledge NIH/NIA (R01 AG032098) for GWAS funding. P.A.C. acknowledges NIH/NHLBI (R01 HL071022). This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106 and was supported in part by the Intramural Research Program of the NIH, National Institute on Aging. The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C.

Health 2000

Investigators: Markku Heliövaara, Mika Kähönen, Samuli Ripatti, Ida Surakka.

Contributions: project conception, design and management: M.H., M.K., phenotype collection and data management: M.H., M.K., genotyping: S.R., I. Surakka., data analysis: M.K., S.R., I. Surakka. **Funding & Acknowledgements:** This study was financially supported by the Medical Research Fund of the Tampere University Hospital.

KORA F4

Investigators: Joachim Heinrich, Stefan Karrasch, Norman Klopp, Holger Schulz.

Nature Genetics: doi:10.1038/ng.941

Contributions: project conception, design and management: J. Heinrich., phenotype collection and data management: S.K, H.S., genotyping: N.K. **Funding & Acknowledgements:** The KORA Augsburg studies were financed by the Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany and supported by grants from the German Federal Ministry of Education and Research (BMBF) in the context of the German National Genome Research Network (NGFN-2 and NGFN-plus). Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ.

KORA S3

Investigators: Eva Albrecht, Christian Gieger, Nicole Probst-Hensch, H Erich Wichmann.

Contributions: project conception, design and management: C.G., H.E.W., phenotype collection and data management: N.P-H, genotyping: C.G., data analysis: E.A. **Funding & Acknowledgements:** The KORA Augsburg studies were financed by the Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany and supported by grants from the German Federal Ministry of Education and Research (BMBF) in the context of the German National Genome Research Network (NGFN-2 and NGFN-plus). Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ. Genotyping for the KORA F3/500K Study which included the lung function characterized KORA C sample was partly funded by UBS Wealth Foundation Grant BA29s8Q7-DZZ.

NFBC1966

Investigators: Paul Elliott, Anna-Liisa Hartikainen, Marjo-Riitta Järvelin, Anneli Pouta, Adaikalavan Ramasamy

Contributors: project conception, design and management: P.E., A-L.H., M-R.J., A.P., phenotype collection and data management: P.E., A-L.H., M-R.J., A.P., genotyping: P.E., A-L.H., M-R.J., A.P., A.R., data analysis: A.R. **Funding & Acknowledgements:** NFBC1966 received financial support from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), the European Commission (EURO-BLCS, Framework 5 award QLG1-CT-2000-01643), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), NIH/NIMH (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, the Medical Research Council, UK (G0500539, G0600705, PrevMetSyn/SALVE) and the Wellcome Trust (project grant GR069224), UK. A.R. acknowledges European Commission (through project GABRIEL (contract number 018996 under the Integrated Program LSH-2004-1.2.5-1)) and Department of Health, UK.

ORCADES

Investigators: Harry Campbell, Caroline Hayward, Veronique Vitart, Sarah H Wild, James F Wilson, Alan F Wright.

Contributions: project conception, design and management: H.C., S.H.W., J.F.W., A.F.W., phenotype collection and data management: H.C., S.H.W., J.F.W., genotyping: H.C., J.F.W., data analysis: C.H., V.V. **Funding & Acknowledgements:** The ORCADES study was funded by the Chief Scientist Office of the Scottish Government, the Royal Society and the MRC Human Genetics Unit. DNA extraction was performed at the Wellcome Trust Clinical Research Facility in Edinburgh. Genotyping was funded by the European Union Framework Programme 6 EUROSPAN project. As a EUROSPAN partner, we are indebted to Yurii Aulchenko, Department of Epidemiology, Erasmus University Medical Center and Anatoly V. Kirichenko, Institute of Cytology and Genetics, Siberian Division of Russian Academy of Sciences, Novosibirsk, Russia for respectively performing imputation of the genotypic data and providing IT facilities. The ORCADES study would like to acknowledge the invaluable contributions of Lorraine Anderson (lorraine.anderson@ed.ac.uk) and the research nurses in Orkney.

RS

Nature Genetics: doi:10.1038/ng.941

Investigators: Guy G Brusselle, Mark Eijgelsheim, Albert Hofman, Daan W Loth, Fernando Rivadeneira, Bruno H Ch Stricker, André G Uitterlinden.

Contributions: project conception, design and management: A. Hofman., phenotype collection and data management: G.G.B, M.E., D.W.L., B.H.Ch.S. genotyping: F.R., A.G.U., data analysis: M.E., D.W.L. **Funding & Acknowledgements:** The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and 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 are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We would like to thank Karol Estrada, Dr. Fernando Rivadeneira, Dr. Tobias A. Knoch, Anis Abuseiris, Luc V. de Zeeuw, and Rob de Graaf (Erasmus MC Rotterdam, The Netherlands), for their help in creating GRIMP, and BigGRID, MediGRID, and Services@MediGRID/D-Grid, (funded by the German Bundesministerium fuer Forschung und Technology; grants 01 AK 803 A-H, 01 IG 07015 G) for access to their grid computing resources.

SHIP

Investigators: Sven Gläser, Georg Homuth, Beate Koch, Henry Völzke.

Contributions: project conception, design and management: S.G., G.H., B.K., H.V., phenotype collection and data management: S.G., B.K., H.V., genotyping: G.H., data analysis: S.G., G.H., B.K, H.V. **Funding & Acknowledgements:** SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the German Asthma and COPD Network (COSYCONET; BMBF grant 01GI0883), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG**.**

TwinsUK

Investigators: Pirro G Hysi, Massimo Mangino, So-Youn Shin, Nicole Soranzo, Tim D Spector, Guangju Zhai.

Contributions: project conception, design and management: T.D.S., G.Z., phenotype collection and data management: C.J.H., P.G. Hysi., M.M., T.D.S., G.Z., genotyping: C.J.H., S-Y.S. data analysis: M.M., G.Z. **Funding & Acknowledgements:** The study was funded by the Wellcome Trust, Arthritis Research Campaign, European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F2-2008-201865-GEFOS and Seventh Framework Programme grant 200800 Treat OA/(FP7/2007-2013), ENGAGE project grant agreement HEALTH-F4-2007-201413, and the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from the UK Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. TDS is an NIHR senior investigator. The project also received support from a Biotechnology and Biological Sciences Research Council project grant (G20234). The authors acknowledge the funding and support of the National Eye Institute (NEI) via an NIH/Center for Inherited Disease Research (CIDR) genotyping project. We thank the staff

Nature Genetics: doi:10.1038/ng.941

from the Genotyping Facilities at the Wellcome Trust Sanger Institute for sample preparation, quality control, and genotyping; Le Centre National de Génotypage, France, for genotyping; Duke University, NC, USA, for genotyping; and the Finnish Institute of Molecular Medicine, Finnish Genome Center, University of Helsinki. Genotyping was also done by CIDR as part of an NEI/NIH project grant.

Stage 2 follow-up studies

ADONIX

Investigators: Jonas Brisman, Santosh Dahgam, Fredrik Nyberg, Anna-Carin Olin.

Contributions: project conception, design and management: J. Brisman., A-C.O., phenotype collection and data management: J. Brisman., A-C.O., genotyping: S.D., F.N., A-C.O., data analysis: S.D., F.N. **Funding & Acknowledgements:** The ADONIX study was funded by the Swedish Research Council for Worklife and Social research (FAS), grants 2001-0263, 2003-0139, Swedish Heart and Lung Foundation grant 20050561 and a collaborative research grant from AstraZeneca.

BHS2

Investigators: John Beilby, Gemma Cadby, Michael L Hunter, Jennie Hui.

Contributions: project conception, design and management: J. Beilby., phenotype collection and data management: J. Beilby., M.L.H., genotyping: J. Beilby., G.C., J. Hui., data analysis: G.C. **Funding & Acknowledgements:** The Busselton Health Study acknowledges the generous support for the 1994/5 follow-up study from Healthway, Western Australia. The Busselton Health Study is supported by The Great Wine Estates of the Margaret River region of Western Australia. The BHS gratefully acknowledges the assistance of the Western Australian DNA Bank (NHMRC Enabling Facility) with DNA samples and the support provided by The Ark for this study.

BRHS

Investigators: Aroon D Hingorani, Richard W Morris, S Goya Wannamethee, Peter H Whincup.

Contributions: phenotype collection and data management: R.W.M., S.G.W., P.H.W., genotyping: A.D.H., R.W.M., data analysis: R.W.M. **Funding & Acknowledgements:** Blood collection and spirometric measures were carried out in 1998-2000 as part of a field study funded by the British Heart Foundation (grant PG/97012). DNA was extracted by Geneservice Limited. A.D.H has a British Heart Foundation Senior Research Fellowship (FS05/125).

BWHHS

Investigators: George Davey Smith, Shah Ebrahim, Debbie A Lawlor, Peter H Whincup.

Contributions: project conception, design and management: G.D.S., S.E., D.A.L., P.H.W., phenotype collection and data management: G.D.S., S.E., D.A.L., P.H.W., genotyping: S.E., D.A.L., data analysis: D.A.L. **Funding & Acknowledgements:** The British Women's Heart & Health Study (BWHHS) is funded by the UK Department of Health Policy Research Programme and the British Heart Foundation. A separate British Heart Foundation project grant (PG/06/154/22043) funded cotinine assays.

CARDIA

Investigators: Myriam Fornage, Xiangjun Gu, Akshay Sood, O Dale Williams.

Nature Genetics: doi:10.1038/ng.941

Contributions: project conception, design and management: A.S., phenotype collection and data management: O.D.W., genotyping: M.F., X.G., data analysis: M.F., X.G. **Funding & Acknowledgements:** The Coronary Artery Risk Development in Young Adults (CARDIA) study is funded by contracts N01-HC-95095, N01-HC-48047, N01-HC-48048, N01-HC-48049, N01-HC-48050, N01-HC-45134, N01-HC-05187, N01-HC-45205 and N01-HC-45204 from the National Heart, Lung and Blood Institute to the CARDIA investigators. A full list of principal CARDIA investigators and institutions can be found at http://www.cardia.dopm.uab.edu/o_pain.htm. Genotyping of the CARDIA participants and statistical data analysis was supported by grants U01-HG-004729 from the National Human Genome Research Institute and R01-HL-084099 from the National Heart, Lung and Blood Institute to MF and K23HL094531-01 from the National Heart, Lung and Blood Institute and M01 RR00997 from the National Center for Research Resources to AS. The authors thank the investigators and staff of the GENEVA coordinating center and genotyping center, as well as the staff and participants of the CARDIA study for their important contributions.

CROATIA-Split

Investigators: Mladen Boban, Vesna Boraska, Ivana Kolcic, Tatijana Zemunik.

Contributions: phenotype collection and data management: M.B., I.K., T.Z., genotyping: V.B., T.Z. **Funding & Acknowledgements:** CROATIA-Split acknowledges Medical Research Council UK and Ministry of Science, Education and Sport of the Republic of Croatia (108-1080315-0302).

Gedling

Investigators: John R Britton, Tricia McKeever.

Contributions: genotyping: J.R.B., T.M. **Funding & Acknowledgements:** The Nottingham Gedling cohort collection was funded by Asthma UK and the British Lung Foundation.

GS:SFHS

Investigators: Catherine M Jackson, Shona M Kerr, Andrew D Morris, David Porteous.

Contributions: project conception, design and management: C.M.J., S.M.K., A.D.M, D.J.P., phenotype collection and data management: C.M.J., A.D.M., genotyping: C.M.J., S.M.K., D.J.P. **Funding & Acknowledgements:** Generation Scotland : Scottish Family Health Study is funded by the Chief Scientists Office , part of the Scottish Government Health Directorate (http://www.sehd.scot.nhs.uk/cso), grant reference number CZD/16/6. Genotyping was conducted by the Genetics Core at the Wellcome Trust Clinical Research Facility, University of Edinburgh, Western General Hospital, Edinburgh, UK.

HCS

Investigators: Cyrus Cooper, John W Holloway, Karen A Jameson, Avan Aihie Sayer.

Contributions: project conception, design and management: C.C., J.W.H., A.A.S., phenotype collection and data management: C.C., K.A.J., A.A.S., genotyping: J.W.H., data analysis: J.W.H., K.A.J. **Funding & Acknowledgements:** The Hertfordshire Cohort Study DNA collection was funded by the Medical Research Council, Arthritis Research Campaign, and International Osteoporosis Foundation. Genotyping was funded by The Asthma, Allergy and Inflammation Research Trust.

LBC1936

Investigators: Ian J Deary, Sarah E Harris, Lorna M Lopez, John M Starr.

Nature Genetics: doi:10.1038/ng.941

Contributions: project conception, design and management: I.J.D., S.E.H., J.M.S., phenotype collection and data management: I.J.D., L.M.L., J.M.S., genotyping: I.J.D., S.E.H., L.M.L., J.M.S., data analysis: L.M.L. **Funding & Acknowledgements:** J.M.S. and I.J.D. acknowledge Biotechnology and Biological Sciences Research Council (BBSRC) (Ref. BB/F019394/1), Research Into Ageing (Ref. 251) and Lifelong Health and Wellbeing Initiative (Ref. G0700704/84698). Lorna M Lopez is the beneficiary of a post-doctoral grant from the AXA Research Fund.

LifeLines

Investigators: H Marike Boezen, Dirkje S Postma, Judith M Vonk, Cisca Wijmenga.

Contributions: project conception, design and management: H.M.B., D.S.P., J.M.V., C. W., phenotype collection and data management: D.S.P., J.M.V., genotyping: C.W., data analysis: H.M.B. **Funding & Acknowledgements:** The LifeLines Cohort Study is supported by the Netherlands Organization of Scientific Research NOW (1750102007006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science the Ministry for Health, Welfare and Sports , the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, the city of Groningen, the University Medical Center Groningen, the University of Groningen and the Dutch Kidney Foundation.

MESA-Lung

Investigators: R Graham Barr, John L Hankinson, Ani Manichaikul, Stephen S Rich.

Contributions: phenotype collection and data management: R.G.B., J.L.H., genotyping: S.S.R., data analysis: A.M., S.S.R. **Funding & Acknowledgements:** MESA and MESA SHARe acknowledges NIH/NHLBI (N01-HC-95159 through N01-HC-95169 and RR-024156). MESA SHARe acknowledges NIH/NHLBI (N02-HL-6-4278). MESA Family acknowledges NIH (R01HL071051, R01HL071205, R01HL071250, R01HL071251, R01HL071252, R01HL071258, R01HL071259). MESA Lung acknowledges NIH/NHLBI (R01 HL077612). MESA Lung SHARe acknowledges NIH/NHLBI (RC1 HL100543).

Nottingham smokers

Investigators: Khalid A Al Balushi, John D Blakey, Ian P Hall, Amanda Henry, Ma'en Obeidat, Ian Sayers.

Contributions: project conception, design and management: I.P.H., phenotype collection and data management:K.A. A-B, J.D.B., I.P.H., A. Henry., M.O., I. Sayers., data analysis: I. Sayers., A. Henry. **Funding & Acknowledgements:** The Nottingham Smokers cohort collection was funded by The University of Nottingham.

NSHD

Investigators: Darya Gaysina, Rebecca Hardy, Diana Kuh, Andrew Wong.

Contributions: phenotype collection and data management: R.H., D.K., genotyping: D.K., A.W., data analysis: D.G., R.H. **Funding & Acknowledgements:** The Medical Research Council National Survey of Health and Development (NSHD, also known as the British 1946 Birth Cohort) is funded by the Medical Research Council. The MRC Unit for Lifelong Health and Ageing is responsible for NSHD. D.G., R.H., D.K., A.W. are all affiliated to the MRC National Survey of Health and Development, MRC Unit for Lifelong Health and Ageing, London, UK.

SAPALDIA

Investigators: Ivan Curjuric, Medea Imboden, Florian Kronenberg, Nicole Probst-Hensch, Thierry Rochat.

Contributions: project conception, design and management: N.P-H., T. Rochat., phenotype collection and data management: N.P-H., genotyping: M.I., F.K., data analysis: I.C., M.I. **Funding & Acknowledgements:** N.P-H. and T.R

Nature Genetics: doi:10.1038/ng.941

acknowledge Swiss National Science Foundation (grants no 4026-28099,3347CO-108796, 3247BO-104283, 3247BO-104288, 3247BO-104284, 32-65896.01,32-59302.99, 32-52720.97, 32-4253.94); the Federal Office for Forest, Environment and Landscape; the Federal Office of Public Health; the Federal Office of Roads and Transport; the canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Zurich; the Swiss Lung League; the canton's Lung League of Basel Stadt/ Basel Landschaft, Geneva, Ticino and Zurich; Freie Akademische Gesellschaft (FAG); UBS Wealth Foundation. F.K acknowledges "Integriertes Forschungs- und Therapiezentrum" (IFTZ) of Innsbruck Medical University to Florian Kronenberg. The study could not have been done without the help of the study participants, technical and administrative support and the medical teams and field workers at the local study sites. Local fieldworkers : Aarau: M Broglie, M Bünter, D Gashi, Basel: R Armbruster, T Damm, U Egermann, M Gut, L Maier, A Vögelin, L Walter, Davos: D Jud, N Lutz, Geneva: M Ares, M Bennour, B Galobardes, E Namer, Lugano: B Baumberger, S Boccia Soldati, E Gehrig-Van Essen, S Ronchetto, Montana: C Bonvin, C Burrus, Payerne: S Blanc, AV Ebinger, ML Fragnière, J Jordan, Wald: R Gimmi, N Kourkoulos, U Schafroth. Administrative staff: N Bauer, D Baehler, C Gabriel, R Gutknecht. SAPALDIA Team: Study directorate: T Rochat, , JM Gaspoz, N Künzli, LJS Liu, NM Probst Hensch, C Schindler. Scientific team: JC Barthélémy, W Berger, R Bettschart, A Bircher, G Bolognini, O Brändli, C Brombach, M Brutsche, L Burdet, M Frey, U Frey, MW Gerbase, D Gold, E de Groot, W Karrer, R Keller, B Knöpfli, B Martin, D Miedinger, U Neu, L Nicod, M Pons, F Roche, T Rothe, E Russi, P Schmid-Grendelmeyer, A Schmidt-Trucksäss, A Turk, J Schwartz, D. Stolz, P Straehl, JM Tschopp, A von Eckardstein, E Zemp Stutz. Scientific team at coordinating centers: M Adam, E Boes, PO Bridevaux, D Carballo, E Corradi, I Curjuric, J Dratva, A Di Pasquale, L Grize, D Keidel, S Kriemler, A Kumar, M Imboden, N Maire, A Mehta, F Meier, H Phuleria, E Schaffner, GA Thun, A Ineichen, M Ragettli, M Ritter, T Schikowski, G Stern, M Tarantino, M Tsai, M Wanner.

Look up studies

ALSPAC

Investigators: David M Evans, Raquel Granell, John Henderson, Wendy L McArdle, Susan M Ring.

Contributions: project conception, design and management: R.G., J. Henderson., phenotype collection and data management: R.G., J. Henderson., genotyping: S.M.R., W.L.M., data analysis: D.M.E. **Funding & Acknowledgements:** D.M.E. is supported by a Medical Research Council New Investigator Award (MRC G0800582 to D.M.E). We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council (G9901462), the Wellcome Trust and the University of Bristol provide core support for ALSPAC. We thank the Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and also 23andME for generating the ALSPAC GWA data. This publication is the work of the authors and they will serve as guarantors for the contents of this paper.

GIANT

Details of the GIANT consortium can be found in Lango Allen, H. et al [90].

ILCCO

Investigators: International Lung Cancer Consortium.

Contributions: phenotype collection and data management: ILCCO, genotyping: ILCCO, data analysis: ILCCO. **Funding & Acknowledgements:** Chu Chen acknowledges NIH/NCI (RO1CA111703). Gary Goodman acknowledges NIH (UO1CA63673). Maria Teresa Landi, Neil Caporaso and Demetrius Albanes acknowledge NIH/NCI/DCEG Intramural funding. Michael Thun and Susan Gapstur acknowledge American Cancer Society. Rayjean Hung acknowledges the funding from Ontario Institute of Cancer Research and Canadian Cancer Society Research Institute (CCSRI 020214). Andres Metspalu acknowledges Estonian Genome Center, University of Tartu, Estonia (SF0180142s08 and by the Centre of Excellence in Genomics via EBC). Lars Vatten acknowledges Norwegian University of Science and

Technology. Inger Njølstad acknowledges Tromsø University, Norway. Chris Amos acknowledges NIH (1P50 CA70907, RO1 CA121197 and U19 CA148127). Margaret Spitz acknowledges NIH (CA55769 and CA127219) NCI GWAS acknowledges Maria Teresa Landi, Neil Caporaso and Demetrius Albanes from the National Cancer Institute, and Michael Thun and Susan Gapstur from the American Cancer Society, for the lung cancer GWAS data. UK GWAS acknowledges Richard Houlston, Peter Broderick, Yufei Wang and Athena Matikidou from the Institute of Cancer Research, and Timothy Eisen from the University of Cambridge for the lung cancer GWAS data. Central Europe GWAS acknowledges Paul Brennan and James McKay from the International Agency for Research on Cancer, David Zaridze from the Cancer Research Centre in Moscow, Neonilia Szeszenia-Dabrowska from the Institute of Occupational Medicine in Lodz, Jolanta Lissowska from M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology in Warsaw, Peter Rudnai from National Institute of Environmental Health in Budapest, Eleonora Fabianova from Specialized Institute of Hygiene and Epidemiology in Banska Bystrica, Dana Mates from National Institute of Public Health in Bucharest, Vladimir Bencko from The First Faculty of Medicine, Charles University in Prague, Czech Republic, Lenka Foretova from Masaryk Memorial Cancer Institute in Brno and Vladimir Janout from Palacky University in Olomouc, for the lung cancer GWAS data. MDACC GWAS acknowledges Chris Amos, Margaret Spitz and Carol Etzel from MD Anderson Cancer Center for lung cancer GWAS data. DeCODE GWAS acknowledges Kari Stefansson and Thorunn Rafnar from DECODE Genetics for lung cancer GWAS data. Germany GWAS acknowledges H-Erich Wichmann and Joachim Heinrich from Helmholtz Center in Munich, Angela Risch from DKFZ and Heike Bickeboller from the University of Göttingen for lung cancer GWAS data. CARET GWAS acknowledges Chu Chen and Gary Goodman from FHCRC for lung cancer GWAS data. CARET would like to acknowledge the support of J. A. Doherty, M. Thornquist, M. Neuhouser, M. Barnett, M. M. Loomis and the CARET study participants. HUNT2/Tromso GWAS acknowledges Lars Vatten from Norwegian University of Science and Techonology and Inger Njølstad from University of Tromsø in Norway for lung cancer GWAS data. Canada GWAS acknowledges Rayjean Hung from Samuel Lunenfeld Research Institute of Mount Sinai Hospital and John McLaughlin from Cancer Care Ontario for lung cancer GWAS data. France GWAS acknowledges Simone Benhamou from INSERM and Christine Bouchardy from Geneva Cancer Registry for lung cancer GWAS data. Estonia GWAS acknowledges Tõnu Vooder from Clinic of Cardiovascular and Thoracic Surgery and Estonian Genome Center, University of Tartu in Estonia, Kristjan Välk from IMCB, Universityu of Tartu and CCRMB Inc, Tartu in Estonia and Andres Metspalu from Estonian Genome Center of University of Tartu in Estonia for lung cancer GWAS data.

Ox-GSK

Investigators: Clyde Francks , Jason Z Liu, Jonathan Marchini.

Contributions: project conception, design and management: C.F., J.M. data analysis: J.Z.L. **Funding & Acknowledgements:** GlaxoSmithKline (GSK), a pharmaceuticals company that is interested in developing new cessation therapies for smoking, funded a postdoctoral fellowship for J.Z.L. at Oxford University. GSK also funded the collection, characterization, and, in some cases, the genotyping and genotype data preparation for several of the cohorts used in this study. A. Roses and P. Matthews played crucial roles in establishing and funding the Medical Genetics activities at GSK.

Raine

Investigators: Wei Q Ang, Patrick G Holt, Craig E Pennell, Peter D Sly.

Contributions: phenotype collection and data management: W.Q.A., P.G. Holt., C.E.P., P.D.S., genotyping: W.Q.A., C.E.P., data analysis: W.Q.A. **Funding & Acknowledgements**: The authors are grateful to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the NH&MRC for their long term contribution to funding the study over the last 20 years and also the following Institutions for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA), Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health

Nature Genetics: doi:10.1038/ng.941

Sciences, The Telethon Institute for Child Health Research and Women and Infants Research Foundation. The authors gratefully acknowledge the assistance of the Western Australian DNA Bank (National Health and Medical Research Council of Australia National Enabling Facility). The authors also acknowledge the support of the National Health and Medical Research Council of Australia (Grant ID 403981 and ID 003209) and the Canadian Institutes of Health Research (Grant ID MOP-82893).

Nature Genetics: doi:10.1038/ng.941

Supplementary references

1. Hancock, D.B. et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet* **42**, 45-52 (2010).
2. Repapi, E. et al. Genome-wide association study identifies five loci associated with lung function. *Nat Genet* **42**, 36-44 (2010).
3. Guan, Y. & Stephens, M. Practical issues in imputation-based association mapping. *PLoS Genet* **4**, e1000279 (2008).
4. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
5. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. *Annu Rev Genomics Hum Genet* **10**, 387-406 (2009).
6. Becker, K.G. et al. PubMatrix: a tool for multiplex literature mining. *BMC Bioinformatics* **4**, 61 (2003).
7. Zhang, Y. et al. Systematic analysis, comparison, and integration of disease based human genetic association data and mouse genetic phenotypic information. *BMC Med Genomics* **3**, 1 (2010).
8. Gibson, M.A., Hughes, J.L., Fanning, J.C. & Cleary, E.G. The major antigen of elastin-associated microfibrils is a 31-kDa glycoprotein. *The Journal of biological chemistry* **261**, 11429-36 (1986).
9. Ramirez, A. et al. Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. *Nat Genet* **38**, 1184-91 (2006).
10. Schimke, R.N., Collins, D.L. & Stolle, C.A. Paraganglioma, neuroblastoma, and a SDHB mutation: Resolution of a 30-year-old mystery. *Am J Med Genet A* **152A**, 1531-5 (2010).
11. Makrygiannakis, D. et al. Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells. *Ann Rheum Dis* **67**, 1488-92 (2008).
12. Yang, J. et al. Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet. *J Cell Biol* **159**, 431-40 (2002).
13. Yang, J. et al. The ciliary rootlet maintains long-term stability of sensory cilia. *Mol Cell Biol* **25**, 4129-37 (2005).
14. Lango Allen, H. et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832-8.
15. Thompson, H.G.R., Mih, J.D., Krasieva, T.B., Tromberg, B.J. & George, S.C. Epithelial-derived TGF-beta2 modulates basal and wound-healing subepithelial matrix homeostasis. *American journal of physiology Lung cellular and molecular physiology* **291**, L1277-85 (2006).
16. Chu, H.W. et al. Transforming growth factor-beta2 induces bronchial epithelial mucin expression in asthma. *The American journal of pathology* **165**, 1097-106 (2004).
17. Sanford, L.P. et al. TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes. *Development* **124**, 2659-70 (1997).
18. Bottoms, S.E., Howell, J.E., Reinhardt, A.K., Evans, I.C. & McAnulty, R.J. Tgf-Beta isoform specific regulation of airway inflammation and remodelling in a murine model of asthma. *PLoS One* **5**, e9674 (2010).
19. Ito, K. et al. Decreased histone deacetylase activity in chronic obstructive pulmonary disease. *The New England journal of medicine* **352**, 1967-76 (2005).
20. Bourbeau, J. & Johnson, M. New and controversial therapies for chronic obstructive pulmonary disease. *Proceedings of the American Thoracic Society* **6**, 553-4 (2009).
21. Vega, R.B. et al. Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell* **119**, 555-66 (2004).
22. Chazaud, C., Dolle, P., Rossant, J. & Mollard, R. Retinoic acid signaling regulates murine bronchial tubule formation. *Mechanisms of development* **120**, 691-700 (2003).
23. Feng, Q. et al. DNA methylation in tumor and matched normal tissues from non-small cell lung cancer patients. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **17**, 645-54 (2008).
24. Massaro, G.D. et al. Retinoic acid receptor-beta: an endogenous inhibitor of the perinatal formation of pulmonary alveoli. *Physiol Genomics* **4**, 51-7 (2000).
25. Bartholomew, C., Kilbey, A., Clark, A.M. & Walker, M. The Evi-1 proto-oncogene encodes a transcriptional repressor activity associated with transformation. *Oncogene* **14**, 569-77 (1997).
26. Gomez-Benito, M. et al. EVI1 controls proliferation in acute myeloid leukaemia through modulation of miR-1-2. *British journal of cancer* **103**, 1292-6 (2010).

Nature Genetics: doi:10.1038/ng.941

27. Kang, J.U., Koo, S.H., Kwon, K.C., Park, J.W. & Kim, J.M. Identification of novel candidate target genes, including EPHB3, MASP1 and SST at 3q26.2-q29 in squamous cell carcinoma of the lung. *BMC Cancer* **9**, 237 (2009).
28. Perkins, A.S., Mercer, J.A., Jenkins, N.A. & Copeland, N.G. Patterns of Evi-1 expression in embryonic and adult tissues suggest that Evi-1 plays an important regulatory role in mouse development. *Development* **111**, 479-87 (1991).
29. Maran, C., Tassone, E., Masola, V. & Onisto, M. The Story of SPATA2 (Spermatogenesis-Associated Protein 2): From Sertoli Cells to Pancreatic Beta-Cells. *Current genomics* **10**, 361-3 (2009).
30. Yang, L., Zhang, L., Wu, Q. & Boyd, D.D. Unbiased screening for transcriptional targets of ZKSCAN3 identifies integrin beta 4 and vascular endothelial growth factor as downstream targets. *The Journal of biological chemistry* **283**, 35295-304 (2008).
31. Yang, L. et al. Evidence of a role for the novel zinc-finger transcription factor ZKSCAN3 in modulating Cyclin D2 expression in multiple myeloma. *Oncogene* **Epub ahead of print**(2010).
32. Pi, H. et al. A novel human SCAN/(Cys)2(His)2 zinc-finger transcription factor ZNF323 in early human embryonic development. *Biochemical and biophysical research communications* **296**, 206-13 (2002).
33. Cruz, E. et al. A new 500 kb haplotype associated with high CD8+ T-lymphocyte numbers predicts a less severe expression of hereditary hemochromatosis. *BMC Med Genet* **9**, 97 (2008).
34. Ecker, K. et al. A RAS recruitment screen identifies ZKSCAN4 as a glucocorticoid receptor-interacting protein. *J Mol Endocrinol* **42**, 105-17 (2009).
35. Li, X. et al. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Developmental cell* **15**, 547-57 (2008).
36. Mackay, D.J. et al. Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. *Nat Genet* **40**, 949-51 (2008).
37. Landi, M.T. et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am J Hum Genet* **85**, 679-91 (2009).
38. Pende, D. et al. Identification and molecular characterization of NKp30, a novel triggering receptor involved in natural cytotoxicity mediated by human natural killer cells. *The Journal of experimental medicine* **190**, 1505-16 (1999).
39. Pendergrass, S.A. et al. Limited systemic sclerosis patients with pulmonary arterial hypertension show biomarkers of inflammation and vascular injury. *PLoS One* **5**, e12106 (2010).
40. Del Galdo, F., Artlett, C.M. & Jimenez, S.A. The role of allograft inflammatory factor 1 in systemic sclerosis. *Curr Opin Rheumatol* **18**, 588-93 (2006).
41. Castro-Giner, F. et al. TNFA -308G>A in two international population-based cohorts and risk of asthma. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology* **32**, 350-61 (2008).
42. Ruse, C.E. et al. Tumour necrosis factor gene complex polymorphisms in chronic obstructive pulmonary disease. *Respir Med* **101**, 340-4 (2007).
43. Smolonska, J., Wijmenga, C., Postma, D.S. & Boezen, H.M. Meta-analyses on suspected chronic obstructive pulmonary disease genes: a summary of 20 years' research. *American journal of respiratory and critical care medicine* **180**, 618-31 (2009).
44. Bahram, S., Bresnahan, M., Geraghty, D.E. & Spies, T. A second lineage of mammalian major histocompatibility complex class I genes. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 6259-63 (1994).
45. Hui, J., Palmer, L.J., James, A.L., Musk, A.W. & Beilby, J.P. AluyMICB dimorphism within the class I region of the major histocompatibility complex is associated with asthma and airflow obstruction in the Busselton population. *Clin Exp Allergy* **36**, 728-34 (2006).
46. Nair, R.P. et al. Sequence and haplotype analysis supports HLA-C as the psoriasis susceptibility 1 gene. *Am J Hum Genet* **78**, 827-51 (2006).
47. Skibola, C.F. et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. *Nat Genet* **41**, 873-5 (2009).
48. Alcais, A. et al. Stepwise replication identifies a low-producing lymphotoxin-alpha allele as a major risk factor for early-onset leprosy. *Nat Genet* **39**, 517-22 (2007).
49. Reich, K. et al. TNF polymorphisms in psoriasis: association of psoriatic arthritis with the promoter polymorphism TNF*-857 independent of the PSORS1 risk allele. *Arthritis Rheum* **56**, 2056-64 (2007).

Nature Genetics: doi:10.1038/ng.941

50. Okamoto, K. et al. Identification of I kappa BL as the second major histocompatibility complex-linked susceptibility locus for rheumatoid arthritis. *Am J Hum Genet* **72**, 303-12 (2003).
51. Kominami, S. et al. HLA-DPB1 and NFKBIL1 may confer the susceptibility to chronic thromboembolic pulmonary hypertension in the absence of deep vein thrombosis. *J Hum Genet* **54**, 108-14 (2009).
52. Dahlback, B. & Nielsen, L.B. Apolipoprotein M--a novel player in high-density lipoprotein metabolism and atherosclerosis. *Curr Opin Lipidol* **17**, 291-5 (2006).
53. Meggio, F. & Pinna, L.A. One-thousand-and-one substrates of protein kinase CK2? *Faseb J* **17**, 349-68 (2003).
54. Gao, Y. & Wang, H.Y. Casein kinase 2 Is activated and essential for Wnt/beta-catenin signaling. *The Journal of biological chemistry* **281**, 18394-400 (2006).
55. Littler, D.R. et al. The enigma of the CLIC proteins: Ion channels, redox proteins, enzymes, scaffolding proteins? *FEBS Lett* **584**, 2093-101 (2010).
56. Lahaye, F. et al. hMSH5 is a nucleocytoplasmic shuttling protein whose stability depends on its subcellular localization. *Nucleic Acids Res* **38**, 3655-71 (2010).
57. Wang, Y. et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat Genet* **40**, 1407-9 (2008).
58. Antonellis, A. & Green, E.D. The role of aminoacyl-tRNA synthetases in genetic diseases. *Annu Rev Genomics Hum Genet* **9**, 87-107 (2008).
59. Francis, P.J., Hamon, S.C., Ott, J., Weleber, R.G. & Klein, M.L. Polymorphisms in C2, CFB and C3 are associated with progression to advanced age related macular degeneration associated with visual loss. *J Med Genet* **46**, 300-7 (2009).
60. Lundblad, L.K. et al. Tumor necrosis factor-alpha overexpression in lung disease: a single cause behind a complex phenotype. *Am J Respir Crit Care Med* **171**, 1363-70 (2005).
61. Kang, H.S. et al. Lymphotoxin is required for maintaining physiological levels of serum IgE that minimizes Th1-mediated airway inflammation. *J Exp Med* **198**, 1643-52 (2003).
62. Mueller-Ortiz, S.L., Drouin, S.M. & Wetsel, R.A. The alternative activation pathway and complement component C3 are critical for a protective immune response against Pseudomonas aeruginosa in a murine model of pneumonia. *Infect Immun* **72**, 2899-906 (2004).
63. Broderick, P. et al. Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res* **69**, 6633-41 (2009).
64. Clancy, R.M. et al. Identification of candidate loci at 6p21 and 21q22 in a genome-wide association study of cardiac manifestations of neonatal lupus. *Arthritis and rheumatism* **62**, 3415-24 (2010).
65. Harley, J.B. et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. *Nat Genet* **40**, 204-10 (2008).
66. Tewari, R., Bailes, E., Bunting, K.A. & Coates, J.C. Armadillo-repeat protein functions: questions for little creatures. *Trends in cell biology* **20**, 470-81 (2010).
67. Bieganowski, P., Shilinski, K., Tsichlis, P.N. & Brenner, C. Cdc123 and checkpoint forkhead associated with RING proteins control the cell cycle by controlling eIF2gamma abundance. *The Journal of biological chemistry* **279**, 44656-66 (2004).
68. Zha, M. et al. Molecular mechanism of ADP-ribose hydrolysis by human NUDT5 from structural and kinetic studies. *Journal of molecular biology* **379**, 568-78 (2008).
69. Tzschach, A. et al. Chromosome aberrations involving 10q22: report of three overlapping interstitial deletions and a balanced translocation disrupting C10orf11. *Eur J Hum Genet* **18**, 291-5 (2009).
70. Morin, C., Sirois, M., Echave, V., Gomes, M.M. & Rousseau, E. Functional effects of 20-HETE on human bronchi: hyperpolarization and relaxation due to BKCa channel activation. *Am J Physiol Lung Cell Mol Physiol* **293**, L1037-44 (2007).
71. Vang, A., Mazer, J., Casserly, B. & Choudhary, G. Activation of endothelial BKCa channels causes pulmonary vasodilation. *Vascul Pharmacol* **53**, 122-9 (2010).
72. Terrand, J. et al. LRP1 controls intracellular cholesterol storage and fatty acid synthesis through modulation of Wnt signaling. *The Journal of biological chemistry* **284**, 381-8 (2009).
73. Lillis, A.P., Mikhailenko, I. & Strickland, D.K. Beyond endocytosis: LRP function in cell migration, proliferation and vascular permeability. *Journal of thrombosis and haemostasis : JTH* **3**, 1884-93 (2005).
74. Gosselink, J.V. et al. Differential expression of tissue repair genes in the pathogenesis of chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* **181**, 1329-35 (2010).
75. Chapoval, S.P. et al. STAT6 Expression in Multiple Cell Types Mediates the Cooperative Development of Allergic Airway Disease. *J Immunol* **186**, 2571-83 (2011).

Nature Genetics: doi:10.1038/ng.941

76. Yamaji-Kegan, K. et al. Hypoxia-induced mitogenic factor (HIMF/FIZZ1/RELMalpha) increases lung inflammation and activates pulmonary microvascular endothelial cells via an IL-4-dependent mechanism. *J Immunol* **185**, 5539-48 (2010).
77. Miyata, S. et al. STAT6 deficiency in a mouse model of allergen-induced airways inflammation abolishes eosinophilia but induces infiltration of CD8+ T cells. *Clin Exp Allergy* **29**, 114-23 (1999).
78. Chiba, Y., Todoroki, M., Nishida, Y., Tanabe, M. & Misawa, M. A novel STAT6 inhibitor AS1517499 ameliorates antigen-induced bronchial hypercontractility in mice. *Am J Respir Cell Mol Biol* **41**, 516-24 (2009).
79. Burkhard, P., Stetefeld, J. & Strelkov, S.V. Coiled coils: a highly versatile protein folding motif. *Trends in cell biology* **11**, 82-8 (2001).
80. Bertoli, A.M. et al. Systemic lupus erythematosus in a multiethnic US Cohort LUMINA XLVIII: factors predictive of pulmonary damage. *Lupus* **16**, 410-7 (2007).
81. Kobayashi, K. et al. Identification of genes whose expression is upregulated in lung adenocarcinoma cells in comparison with type II alveolar cells and bronchiolar epithelial cells in vivo. *Oncogene* **23**, 3089-96 (2004).
82. Abraham, R. et al. Identification of MMP-15 as an anti-apoptotic factor in cancer cells. *The Journal of biological chemistry* **280**, 34123-32 (2005).
83. Gueders, M.M., Foidart, J.-M., Noel, A. & Cataldo, D.D. Matrix metalloproteinases (MMPs) and tissue inhibitors of MMPs in the respiratory tract: potential implications in asthma and other lung diseases. *European journal of pharmacology* **533**, 133-44 (2006).
84. Park, E.J. *et al.* PM 2.5 collected in a residential area induced Th1-type inflammatory responses with oxidative stress in mice. *Environ Res* **111**, 348-55 (2011).
85. Diekwisch, T.G., Marches, F., Williams, A. & Luan, X. Cloning, gene expression, and characterization of CP27, a novel gene in mouse embryogenesis. *Gene* **235**, 19-30 (1999).
86. Roepke, T.K. et al. Kcne2 deletion uncovers its crucial role in thyroid hormone biosynthesis. *Nature medicine* **15**, 1186-94 (2009).
87. Cowley, E.A. & Linsdell, P. Characterization of basolateral K+ channels underlying anion secretion in the human airway cell line Calu-3. *The Journal of physiology* **538**, 747-57 (2002).
88. Kathiresan, S. et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet* **41**, 334-41 (2009).
89. Liu, J.Z. et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet* **42**, 436-40 (2010).
90. Lango Allen, H. et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832-8 (2010).
91. Cho, M.H. et al. Variants in FAM13A are associated with chronic obstructive pulmonary disease. *Nat Genet* **42**, 200-2 (2010).
92. Pillai, S.G. et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet* **5**, e1000421 (2009).
93. Wilk, J.B. et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. *PLoS genetics* **5**, e1000429 (2009).
94. DeMeo, D.L. et al. Integration of genomic and genetic approaches implicates IREB2 as a COPD susceptibility gene. *Am J Hum Genet* **85**, 493-502 (2009).
95. Park, J.H. et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet* **42**, 570-5 (2010).
96. Ehret, G.B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* (2011).
97. Gottlieb, D.J. et al. Heritability of longitudinal change in lung function. The Framingham study. *Am J Respir Crit Care Med* **164**, 1655-9 (2001).
98. Lewitter, F.I., Tager, I.B., McGue, M., Tishler, P.V. & Speizer, F.E. Genetic and environmental determinants of level of pulmonary function. *Am J Epidemiol* **120**, 518-30 (1984).
99. Palmer, L.J. et al. Familial aggregation and heritability of adult lung function: results from the Busselton Health Study. *Eur Respir J* **17**, 696-702 (2001).
100. Wilk, J.B. et al. Evidence for major genes influencing pulmonary function in the NHLBI family heart study. *Genet Epidemiol* **19**, 81-94 (2000).
101. Segre, A.V., Groop, L., Mootha, V.K., Daly, M.J. & Altshuler, D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet* **6**(2010).

Nature Genetics: doi:10.1038/ng.941

102. Raychaudhuri, S. et al. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet* **5**, e1000534 (2009).
103. Conrad, D.F. et al. Origins and functional impact of copy number variation in the human genome. *Nature* **464**, 704-12 (2009).
104. McCarroll, S.A. et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* **40**, 1166-74 (2008).
105. Dixon, A.L. et al. A genome-wide association study of global gene expression. *Nat Genet* **39**, 1202-7 (2007).
106. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am J Epidemiol* **129**, 687-702 (1989).
107. Strachan, D.P. et al. Lifecourse influences on health among British adults: effects of region of residence in childhood and adulthood. *Int J Epidemiol* **36**, 522-31 (2007).
108. Marossy, A.E., Strachan, D.P., Rudnicka, A.R. & Anderson, H.R. Childhood chest illness and the rate of decline of adult lung function between ages 35 and 45 years. *Am J Respir Crit Care Med* **175**, 355-9 (2007).
109. Orfei, L., Strachan, D.P., Rudnicka, A.R. & Wadsworth, M.E. Early influences on adult lung function in two national British cohorts. *Arch Dis Child* **93**, 570-4 (2008).
110. James, A.L. et al. Decline in lung function in the Busselton Health Study: the effects of asthma and cigarette smoking. *Am J Respir Crit Care Med* **171**, 109-14 (2005).
111. Ryan, G. et al. Decline in lung function and mortality: the Busselton Health Study. *J Epidemiol Community Health* **53**, 230-4 (1999).
112. Fried, L.P. et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* **1**, 263-76 (1991).
113. Buist, A.S. Standardization of spirometry. *Am Rev Respir Dis* **136**, 1073-4 (1987).
114. Aulchenko, Y.S., Ripke, S., Isaacs, A. & van Duijn, C.M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294-6 (2007).
115. Amin, N., van Duijn, C.M. & Aulchenko, Y.S. A genomic background based method for association analysis in related individuals. *PLoS One* **2**, e1274 (2007).
116. Burney, P.G., Luczynska, C., Chinn, S. & Jarvis, D. The European Community Respiratory Health Survey. *Eur Respir J* **7**, 954-60 (1994).
117. European Community Respiratory Health Survey II Steering Committee (written by D.Jarvis et al). The European Community Respiratory Health Survey II. *Eur Respir J* **20**, 1071-9 (2002).
118. Moffatt, M.F. et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* **363**, 1211-21 (2010).
119. Chinn, S. et al. Smoking cessation, lung function, and weight gain: a follow-up study. *Lancet* **365**, 1629-35; discussion 1600-1 (2005).
120. Canoy, D. et al. Abdominal obesity and respiratory function in men and women in the EPIC-Norfolk Study, United Kingdom. *Am J Epidemiol* **159**, 1140-9 (2004).
121. Cox, B., Huppert, F. & Whichelow, M. *The Health and Lifestyle Survey: seven years on*, (Dartmouth Publishing Company, Aldershot, United Kingdom, 1993).
122. Kaprio, J., Pulkkinen, L. & Rose, R.J. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. *Twin Res* **5**, 366-71 (2002).
123. Aulchenko, Y.S. et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* **41**, 47-55 (2009).
124. Yende, S. et al. Inflammatory markers are associated with ventilatory limitation and muscle dysfunction in obstructive lung disease in well functioning elderly subjects. *Thorax* **61**, 10-6 (2006).
125. Heistaro, S. Methodology report. Health 2000 survey. in *Publications of National Public Health Institute* (ed. Heistaro, S.) (2000).
126. Wichmann, H.E., Gieger, C. & Illig, T. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1**, S26-30 (2005).
127. Miller, M.R. et al. Standardisation of spirometry. *Eur Respir J* **26**, 319-38 (2005).
128. Hofman, A. et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol* **24**, 553-72 (2009).
129. 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* **25**, 2750-2 (2009).
130. John, U. et al. Study of Health In Pomerania (SHIP): a health examination survey in an east German region: objectives and design. *Soz Praventivmed* **46**, 186-94 (2001).

Nature Genetics: doi:10.1038/ng.941

131. Volzke, H. et al. Reference intervals of serum thyroid function tests in a previously iodine-deficient area. *Thyroid* **15**, 279-85 (2005).
132. Nelson, S.B., Gardner, R.M., Crapo, R.O. & Jensen, R.L. Performance evaluation of contemporary spirometers. *Chest* **97**, 288-97 (1990).
133. Standardization of spirometry--1987 update. Statement of the American Thoracic Society. *Am Rev Respir Dis* **136**, 1285-98 (1987).
134. Quanjer, P.H. et al. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl* **16**, 5-40 (1993).
135. Andrew, T. et al. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Res* **4**, 464-77 (2001).
136. Richards, J.B. et al. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet* **371**, 1505-12 (2008).
137. Teo, Y.Y. et al. A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics* **23**, 2741-6 (2007).
138. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).
139. Pritchard, J.K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-59 (2000).
140. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
141. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).
142. Olin, A.C., Bake, B. & Toren, K. Fraction of exhaled nitric oxide at 50 mL/s: reference values for adult lifelong never-smokers. *Chest* **131**, 1852-6 (2007).
143. Olin, A.C. et al. Height, age, and atopy are associated with fraction of exhaled nitric oxide in a large adult general population sample. *Chest* **130**, 1319-25 (2006).
144. Berg, C. et al. Trends in overweight and obesity from 1985 to 2002 in Goteborg, West Sweden. *Int J Obes (Lond)* **29**, 916-24 (2005).
145. Berg, C.M. et al. Trends in blood lipid levels, blood pressure, alcohol and smoking habits from 1985 to 2002: results from INTERGENE and GOT-MONICA. *Eur J Cardiovasc Prev Rehabil* **12**, 115-25 (2005).
146. Walker, M., Whincup, P.H. & Shaper, A.G. The British Regional Heart Study 1975-2004. *Int J Epidemiol* **33**, 1185-92 (2004).
147. Wannamethee, S.G., Shaper, A.G. & Whincup, P.H. Body fat distribution, body composition, and respiratory function in elderly men. *Am J Clin Nutr* **82**, 996-1003 (2005).
148. Standardization of Spirometry, 1994 Update. American Thoracic Society. *Am J Respir Crit Care Med* **152**, 1107-36 (1995).
149. Lawlor, D.A., Bedford, C., Taylor, M. & Ebrahim, S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J Epidemiol Community Health* **57**, 134-40 (2003).
150. Lawlor, D.A., Ebrahim, S. & Davey Smith, G. Association of birth weight with adult lung function: findings from the British Women's Heart and Health Study and a meta-analysis. *Thorax* **60**, 851-8 (2005).
151. Friedman, G.D. et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol* **41**, 1105-16 (1988).
152. Hughes, G.H. et al. Recruitment in the Coronary Artery Disease Risk Development in Young Adults (Cardia) Study. *Control Clin Trials* **8**, 68S-73S (1987).
153. ATS statement--Snowbird workshop on standardization of spirometry. *Am Rev Respir Dis* **119**, 831-8 (1979).
154. Britton, J.R. et al. Dietary antioxidant vitamin intake and lung function in the general population. *Am J Respir Crit Care Med* **151**, 1383-7 (1995).
155. McKeever, T.M. et al. Prospective study of diet and decline in lung function in a general population. *Am J Respir Crit Care Med* **165**, 1299-303 (2002).
156. Smith, B.H. et al. Generation Scotland: the Scottish Family Health Study; a new resource for researching genes and heritability. *BMC Med Genet* **7**, 74 (2006).
157. Syddall, H.E. et al. Cohort profile: the Hertfordshire cohort study. *Int J Epidemiol* **34**, 1234-42 (2005).

Nature Genetics: doi:10.1038/ng.941

158. Deary, I.J. et al. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr* **7**, 28 (2007).
159. Stolk, R.P. et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *Eur J Epidemiol* **23**, 67-74 (2008).
160. Dolmans, G.H. et al. Wnt Signaling and Dupuytren's Disease. *N Engl J Med* **2011**, 6 (2011).
161. Bild, D.E. et al. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* **156**, 871-81 (2002).
162. Rodriguez, J. et al. The association of pipe and cigar use with cotinine levels, lung function, and airflow obstruction: a cross-sectional study. *Ann Intern Med* **152**, 201-10.
163. Hankinson, J.L. et al. Performance of American Thoracic Society-recommended spirometry reference values in a multiethnic sample of adults: the multi-ethnic study of atherosclerosis (MESA) lung study. *Chest* **137**, 138-45.
164. Wadsworth, M., Kuh, D., Richards, M. & Hardy, R. Cohort Profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). *Int J Epidemiol* **35**, 49-54 (2006).
165. Black, S. et al. Contribution of functional variation in the IL13 gene to allergy, hay fever and asthma in the NSHD longitudinal 1946 birth cohort. *Allergy* **25**, 25 (2009).
166. Wadsworth, M.E. et al. Alpha1-antitrypsin as a risk for infant and adult respiratory outcomes in a national birth cohort. *Am J Respir Cell Mol Biol* **31**, 559-64 (2004).
167. Ackermann-Liebrich, U. et al. Follow-up of the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA 2) 1991-2003: methods and characterization of participants. *Soz Praventivmed* **50**, 245-63 (2005).
168. Downs, S.H. et al. Reduced exposure to PM10 and attenuated age-related decline in lung function. *N Engl J Med* **357**, 2338-47 (2007).
169. Imboden, M. et al. Glutathione S-transferase genotypes modify lung function decline in the general population: SAPALDIA cohort study. *Respir Res* **8**, 2 (2007).
170. Kunzli, N. et al. Longitudinal validity of spirometers--a challenge in longitudinal studies. *Swiss Med Wkly* **135**, 503-8 (2005).
171. Golding, J., Pembrey, M. & Jones, R. ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol* **15**, 74-87 (2001).
172. Newnham, J.P., Evans, S.F., Michael, C.A., Stanley, F.J. & Landau, L.I. Effects of frequent ultrasound during pregnancy: a randomised controlled trial. *Lancet* **342**, 887-91 (1993).
173. Williams, L.A., Evans, S.F. & Newnham, J.P. Prospective cohort study of factors influencing the relative weights of the placenta and the newborn infant. *BMJ* **314**, 1864-8 (1997).
174. Evans, S., Newnham, J., MacDonald, W. & Hall, C. Characterisation of the possible effect on birthweight following frequent prenatal ultrasound examinations. *Early Hum Dev* **45**, 203-14 (1996).

Nature Genetics: doi:10.1038/ng.941