Supplementary Information

TABLE OF CONTENTS

Supplementary methods 2

Study populations 2
Genotyping, quality control and imputation 6
Statistical analysis 8
The lung eQTL study 8
Network analysis 10

Supplementary methods

Study populations

For the identification phase (meta-analysis1), subjects were recruited as participants of the following cohorts:

1. The Dutch Asthma GWAS (DAG) Study: a cohort screened for genetic studies, characterized by the presence of a doctor diagnosis of asthma and hyperresponsiveness and extensive phenotyping.
2. The NELSON cohort study [27]: a population-based cohort screening for lung cancer, including current or ex-smokers with at least 20 pack-years. Blood bank controls, collected in Amsterdam and Utrecht, were added to increase power of the COPD set, and except for the age (range 18-65) no clinical data were available.

For the 1st replication phase (meta-analysis2, 2,048 SNPs) subjects were recruited as participants of the LifeLines cohort study.

LifeLines [S1] is a biobanking initiative collecting random individuals from the general population and their relatives in order to perform intense phenotyping which will be used to better define disease type and/or subtype and relate it to genetic and environmental factors. Amongst others spirometry is performed (following ATS guidelines) and data on respiratory symptoms are collected (using standardized questionnaires) enabling a diagnosis of asthma and/or COPD. For purpose of current study all relatives were excluded.

The 2nd replication phase (meta-analyses3-9) evaluated the top 20 SNPs; subjects were recruited as participants of the LifeLines cohort study (LifeLines2), an independent sample of Lifelines, SAPALDIA, RS-I, RS-II, RS-III, MESA, and ARIC cohorts.

The **SAPALDIA** cohort: a multi-center study in eight geographic areas representing the range of environmental, meteorological and socio-demographic conditions in Switzerland. It was initiated in 1991 (SAPALDIA1) with a follow-up assessment in 2002 (SAPALDIA2) and 2011 (SAPALDIA3). This study has specifically been designed to investigate longitudinally lung function, respiratory and cardiovascular health, and to study and identify the associations of these health indicators with individual long term exposure to air pollution, other toxic inhalants, life style and molecular factors. SAPALDIA2 was used in current study.

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. The third cohort (RS III) consists of 3,932 participants, aged 45 years and over and was recruited in 2006. 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. FEV_1, FVC and FEV_1/FVC ratio were measured. Asthma cases were collected by chart review. Cases were ascertained by 1) Doctor's diagnosis of Asthma and 2) no conflicting respiratory diagnosis. Additionally, prescriptions for respiratory medicines were used for case finding.

The Multi-Ethnic Study of Atherosclerosis (MESA) is a longitudinal study of subclinical cardiovascular disease and risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease.[S2] Between 2000 and 2002, MESA recruited 6,814 men and women 45 to 84 years of age from Forsyth County,

North Carolina; New York City; Baltimore; St. Paul, Minnesota; Chicago; and Los Angeles. Exclusion criteria were clinical cardiovascular disease, weight exceeding 136 kg (300 lb.), pregnancy, and impediment to long-term participation.The MESA Family Study recruited 1,595 African American and Hispanic participants, generally siblings of MESA participants, using the same inclusion and exclusion criteria as MESA except that clinical cardiovascular disease was permitted.[S3] The MESA Air Pollution Study recruited an additional 257 participants from Los Angeles and Riverside County, CA, and Rockland County, NY, using the same criteria as MESA except that participants were ages 50 to 89 who lived in the area ≥ 50% of the year and had no plans to move in the next five years.[S4]

Atherosclerosis Risk in Communities Study (ARIC) study is a community-based prospective cohort of cardiovascular disease and its risk factors.[S5] ARIC recruited 15,792 men and women, aged 45-64 years, from 4 communities in the United States (Forsyth Co, NC; Jackson, MS; Minneapolis suburbs, MN; Washington Co, MD) in 1987-1989. Participants were mostly white in the Minnesota and Washington Co field center, white and African-American in Forsyth Co, and exclusively African-American in the Jackson field center. At baseline, participants answered questionnaires on respiratory and other illnesses. White subjects were classified as having asthma diagnosis if they responded yes to the question "Has a doctor ever said that you had asthma? Among the whites with GWAS data there were 453 cases with asthma and 9203 controls. COPD was defined based on pre-bronchodilator spirometry according to the presence of airflow obstruction was defined as an FEV_1 and FEV_1/FVC both less than the lower limit of normal based on prediction equations that include age, age^2, and height^2 calculated separately by sex.[S6] Unaffected participants were defined by FEV_1, FVC, and FEV_1/FVC all above the lower limit of normal. Individuals below the lower limit of normal for FEV_1 or FEV_1/FVC but not both were excluded from these analyses.[S7] There were 915 COPD, 6,610 controls in this analysis.

Ethics statement

LifeLines cohort and DAG study were approved by the Ethical Committee of the University Medical Center Groningen.

NELSON trial was approved by the institutional review board of the University Medical Center Utrecht.

SAPALDIA was approved by the Overall Regional Ethics Commission for Clinical Medicine (Swiss Academy of Medical Sciences, Basel, Switzerland) and the responsible cantonal ethics committees of the study centres (Ethics Commissions of the cantons Aarau, Basel, Geneva, Grisons, Ticino, Valais, Vaud, and Zürich)

The Rotterdam Study has been approved by the institutional review board (Medical Ethics Committee) of the Erasmus Medical Center and by the review board of the Netherlands Ministry of Health, Welfare and Sports

The MESA Lung SHARe project is approved by the institutional review boards of the participating study centers at Wake Forest University, Columbia University, Johns Hopkins University, University of Minnesota, Northwestern University and University of California - Los Angeles

Institutional Review Board (IRB) approvals were obtained by each participating ARIC study center (the Universities of North Carolina, Mississippi, Minnesota, and John Hopkins University) and the coordinating center (University of North Carolina), and the research was conducted in accordance with the principles described in the Declaration of Helsinki.

Human lung tissue samples for the Lung eQTL Consortium were obtained in accordance with Institutional Review Board guidelines at three sites: Laval University (Quebec, Canada), University of British-Columbia (Vancouver, Canada) and Groningen University (Groningen,

The Netherlands). The study was approved by the ethics committees of the Institut universitaire de cardiologie et de pneumologie de Québec (IUCPQ) and the UBC-Providence Health Care Research Institute Ethics Board for Laval and UBC, respectively. The study protocol was consistent with the Research Code of the University Medical Center Groningen and Dutch national ethical and professional guidelines ("Code of conduct; Dutch federation of biomedical scientific societies"; http://www.federa.org).

Genotyping, quality control and imputation

The Dutch Asthma GWAS cohort was genotyped on Illumina Hap300 and Hap370 platforms. The NELSON COPD cohort was genotyped using Illumina 610 Quad BeadChip containing 620,901 probes. The blood bank controls for the NELSON COPD cohort were genotyped with Illumina 670 Quad BeadChip containing custom and tagging SNPs selected by WTCCC2.

The LifeLines cohort was genotyped on CytoChip containing a selection of 299,140 SNPs to tag the whole genome and capture regions most frequently used in cytogenetic screening. Genotypes were called with an algorithm provided by Illumina and implemented in Genome Studio. Quality control was performed for each dataset to ensure the best quality. SNPs were removed if call rates were <95%, the minor allele frequency was below 5% and/or Hardy-Weinberg equilibrium was not met (HWE; $p<0.0001$). Samples were removed when more than 5% of genotypes were missing, samples were duplicated (PI_HAT value ~ 1) or from individuals related to another individual in the dataset (PI_HAT>0.5) or ethnic outliers (based on components from multidimensional scaling C1 and C2). To enable uniform analyses across all datasets imputation was performed using BEAGLE 3.0 [S8] and HapMap CEU as reference panel (HapMap 2, release 24, genome build 36). Genotype dosages were

converted to regular genotypes with a best-guess method. SNPs imputed with less confidence were removed (correlation r2<0.5).

In the framework of the European GABRIEL study on asthma genetics, genotyping in the SAPALDIA cohort was done on the Illumina Human610 Quad platform at the Centre National de Génotypage in Evry, France on 663 asthmatics and a random sample of 997 non-asthmatic participants. 567589 successfully genotyped autosomal SNPs were imputed to 2.5 Mio by MACH v 1.0 software [S9] using the HapMap v22 CEPH panel of Utah residents with ancestry from northern and Western Europe as reference. Samples with <97% genotyping success rate, non-European origin, cryptic relatedness or sex inconsistencies were excluded from analysis. After quality control procedures, genetic and covariate data from 461 cases of doctor diagnosed asthma, 118 cases of COPD and 656 controls was available.

Details regarding the Rotterdam Study, genotyping, imputation and quality control have been described previously.[36,S10]

Participants in the original MESA cohort, the MESA Family Study and the MESA Air Pollution Study who consented to genetic analyses were genotyped in 2009 using the Affymetrix Human SNP array 6.0. Genotype quality control for these data have been described previously.[S11] Briefly, we filtered on SNP level call rate < 95%, individual level call rate < 95%, heterozygosity > 53%, and removed all monomorphic SNPs. The cleaned genotypic data was deposited with MESA phenotypic data into dbGaP as the MESA SHARe project (study accession phs000209, http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000209.v7.p2); 8,224 consenting individuals (2,685 Caucasian, 2,588 non-Hispanic African-American, 2,174 Hispanic, 777 Chinese) were included, with 897,981 SNPs passing study specific quality control (QC). For MESA Caucasians, IMPUTE version 2.1.0 was used to perform imputation (chromosomes 1-22) using HapMap Phase I and II - CEU as the reference panel (release #22 - NCBI Build 36 (dbSNP b126)).

Statistical analysis

Case-control analysis was performed using X^2 test for both asthma and COPD. Association results were combined in a directional meta-analysis. Same approach was used for later replication stages. All above mentioned analyses were performed using PLINK [S12] in case of NELSON, DAG and LifeLines cohorts. Remaining cohorts performed the analysis using ProbABEL software version 0.1.3.[S13] Logistic regression models were applied to adjust for certain covariates (age, gender, pack-years, center, and principal components).

Directions of the signals were compared between the meta-analyses and if they were the same results were combined with Fishers' method using R.[S14] We used Fishers' method to combine p values resulting from meta-analyses, as meta-analysis of meta-analysis is statically inappropriate.

The lung eQTL study

Non-tumor lung tissues were collected from patients who underwent lung resection surgery at three participating sites: Laval University (Quebec City, Canada), University of Groningen (Groningen, The Netherlands), and University of British Columbia (Vancouver, Canada). Whole-genome gene expression and genotyping data were obtained from these specimens. Gene expression profiling was performed using an Affymetrix custom array testing 51,627 non-control probe sets and normalized using RMA.[30] Genotyping was performed using the Illumina Human1M-Duo BeadChip array. At Laval, lung specimens were collected from patients undergoing lung cancer surgery and stored at the "Institut universitaire de cardiologie et de pneumologie de Québec" (IUCPQ) site of the Respiratory Health Network Tissue Bank of the "Fonds de recherche du Québec – Santé" (www.tissuebank.ca). Written informed

consent was obtained from all subjects and the study was approved by the IUCPQ ethics committee. At Groningen, lung specimens were provided by the local tissue bank of the Department of Pathology and the study protocol was consistent with the Research Code of the University Medical Center Groningen and Dutch national ethical and professional guidelines (“Code of conduct; Dutch federation of biomedical scientific societies”; http://www.federa.org). At Vancouver, the lung specimens were provided by the James Hogg Research Center Biobank at St Paul's Hospital and subjects provided written informed consent. The study was approved by the ethics committees at the UBC-Providence Health Care Research Institute Ethics Board.

The lung eQTL analysis was performed as described before by Fehrman and Hao.[28,29].

Network analysis

To perform gene enrichment base on networks we used publically available tool: GeneMANIA (http://www.genemania.org/).[32] Because GeneMANIA does not support pseudogenes, we queried two genes resulting from our analysis: *DDX1* and *COMMD10*. The settings were the following: we chose to show up to 20 genes in the network, we used automatic weighting for the network. For the network creation we used only physical interactions, predicted interaction, pathways and coexpression. GeneMANIA was accessed on 28th September 2011.

In order to identify the overrepresented pathways in the above mentioned network we used GATHER.[33] GATHER is web-based tool (http://gather.genome.duke.edu/), which performs annotations, among others, to pathways, based on the query genes. The resulting annotations are given Bayes factor and p value to indicate the strength of the annotation with the gene list provided. We queried 22 genes in total: *DDX1*, *COMMD10* and 20 genes resulting from gene enrichment in GeneMANIA: *SBDS, NAGLU, ROGDI, SAP30L, SAP30, ATP2A2, MTHFSD, COMMD1, DAB2IP, RELB, CSTF2, SETD3, CEP250, NBN, RAD50, CDK5, MRE11A, NFKB2, RELA* and *HNRNPK*. GATHER was last accessed on August 3rd 2012.

Association of SNPs with asthma and COPD

SNPs in the DDX1 and COMMD10 locus were associated with both asthma and COPD (Table S4). The meta-analysis results of the GNG5P5 locus were driven by the association with the COPD phenotype, since non of the GNG5P5 SNPs were significantly associated with the asthma phenotype.

CHR	BP	SNP	P_orig_COPD	P_orig_asthma	locus
2	15820130	rs2112101	0.005356	0.02213	*DDX1*
2	15822156	rs6728667	0.003193	0.03643	*DDX1*
2	15822185	rs6728750	0.00247	0.02818	*DDX1*
2	15823917	rs2544534	0.001117	0.0243	*DDX1*
2	15827908	rs1477253	0.0004962	0.01191	*DDX1*
2	15830470	rs2693008	0.0007489	0.01857	*DDX1*
2	15837774	rs2544523	0.0006778	0.01818	*DDX1*
2	15839739	rs2693019	0.0005924	0.02281	*DDX1*
5	115623770	rs10036292	0.004218	0.0623	*COMMD10*
5	115624947	rs10043228	0.00341	0.0623	*COMMD10*
5	115633819	rs254149	0.01456	0.007756	*COMMD10*
13	46738025	rs17069787	2.49E-06	0.4495	*GNG5P5*
13	46739001	rs7994542	0.0001131	0.07985	*GNG5P5*
13	46741378	rs9534578	1.50E-06	0.2737	*GNG5P5*

Supplementary Figure and Table legends

Figure S1. Quantile-quantile plots for asthma and COPD GWAS

Table S1. Characteristics of participants

a – blood bank controls – no data available

b – NELSON controls, numbers given in the table based on this group

* calculated in ever smokers

† forced expiratory volume in one second

‡ forced vital capacity

§FEV1 percent predicted

¶ FEV1/FVC ratio, used for COPD diagnosis

						Smoking habits				*Steroid use in last 12 months (%)*				*Lung function variables***			
	Phenotype	N	Age, yrs, mean (SD)	Height. meters. mean (SD)	Gender male n (%)	Current smoker, n (%)	Never smoker, n (%)	Ex smoker, n (%)	Pack-years, median (p25 – p75)*	Oral steroids, n (%)	Inhaled steroids, n (%)	Oral and inhaled steroids, n (%)	No steroids, n (%)	FEV1 pre med, L, mean (SD)†	FVC pre med, L, mean (SD)‡	FEV1 percent pred, mean (SD)§	FEV1/FVC, mean (SD)¶
	asthma	920	34 (16)	1.69 (0.17)	430 (47)	147 (16.0)	544 (59.1)	226 (24.6)	7.9 (2.1 -17.3)	51 (5.5)	506 (55.0)	23 (2.5)	325 (35.3)	2.81 (0.95)	3.35 (1.44)	86 (21)	79.72 (10.6)
DAG	controls	2.777	55.4 (9.9)	1.73 (0.09)	991 (36)	396 (14)	1,305 (47)	1076 (39)	1.95 (0-11.6)	1 (0)	49 (1.8)	0 (0)	2727 (98.2)	3.25 (0.8)	4.21 (1.0)	98.2 (13.0)	77.2 (5.4)
	COPD	1.030	63.3 (5.6)	1.78 (0.06)	1,030 (100)	410 (39.8)	0 (0)	620 (60.2)	38.7 (29.7-49.5)	n.a.	n.a.	n.a.	n.a.	2.76 (0.70)	3.31 (1.61)	82.3 (18.3)	60.28 (8.5)
NELSON	controls	844[a]+ 964[b]	59.1 (5)	178.2 (6.6)	964 (100)	621 (64,4)	0 (0)	343 (35.6)	34.2 (27.9-46.2)	n.a.	n.a.	n.a.	n.a.	3.10 (0.73)	3.52 (1.44)	91.6 (18.6)	76.5 (4.9)
LifeLines1	asthma	534	44.8 (9.7)	1.74 (0.1)	214 (40)	106 (19.9)	293 (54.9)	135 (25.3)	10.8 (4.9 - 20.5)	3 (0.6)	254 (47.6)	22 (4.1)	255 (47.8)	3.27 (0.78)	4.33 (1.01)	89.3 (12.9)	75.7 (7.4)
	controls	2.568	43 (9.4)	1.75 (0.09)	1,102 (42.9)	266 (10.4)	2,010 (78.8)	276 (10.8)	12.75 (5.5-20.4)	1 (0)	123 (4.8)	6 (0.2)	2438 (94.9)	3.5 (0.83)	4.5 (1.0)	92.2 (13.0)	78.2 (7.3)
	COPD	711	54 (10.6)	1.76 (0.09)	369 (52)	363 (51.1)	0 (0)	348 (48.9)	16.8 (8.5 – 26.7)	0(0)	30 (4.2)	2 (0.3)	679 (95.5)	2.89 (0.77)	4.5 (1.07)	81.6 (13.9)	63.9 (5.7)
	controls	1.854	43.2 (8.6)	1.75 (0.09)	807 (43.5)	805 (43.4)	0 (0)	1049 (56.6)	9 (4-15)	0(0)	19 (1)	0(0)	1835 (99)	3.72 (0.76)	4.7 (0.97)	98.1 (9.14)	79.3 (4.9)
LifeLines2	asthma	317	46.7 (11.2)	1.73 (0.09)	120 (37.9)	41(12.9)	171(53.9)	105 (33.1)	7.4 (3 – 15.5)	0 (0)	136 (42.9)	21 (6.6)	160 (50.5)	3.29 (0.86)	4.34 (1.06)	91.8 (13.8)	75.9 (7.2)
	controls	2.363	48.5 (11.6)	1.74 (0.09)	885 (37.5)	165 (7.2)	1922 (83.3)	220 (9.5)	12 (5-20.5)	0 (0)	40	1	2322	3.4 (0.84)	4.3 (1.05)	94.4 (13.0)	77.5 (5.9)
	COPD	601	56.7 (10.8)	1.74 (0.09)	282 (46.9)	231(38.4)	0 (0)	370 (61.6)	15.2(7 – 25.2)	0 (0)	34 (5.7)	0 (0)	567(94.3)	2.84 (0.72)	4.4 (1.04)	84.3 (13.0)	64.4 (5.4)
	controls	1.868	49.6 (10.9)	1.74 (0.09)	784 (42.0)	601 (32.2)	0 (0)	1267 (67.8)	8.6 (4-16)	0 (0)	28 (1.5)	1 (0)	1839 (98.5)	3.5 (0.78)	4.5 (0.98)	99.7 (10.0)	77.9 (4.5)
SAPALDIA 2	asthma	461	49.0 (11.8)	1.69 (0.09)	212 (46.0)	95 (20.6)	215 (46.6)	151 (32.8)	16.3 (4.9-32.9)	n.a.	n.a.	n.a.	n.a.	3.02 (0.95)	4.17 (1.11)	90.0 (17.2)	72.1 (9.7)
	controls	522	51.4 (11.1)	1.69 (0.10)	244 (46.7)	95 (18.2)	252 (48.3)	175 (33.5)	13.1 (5.1-25.5)	n.a.	n.a.	n.a.	n.a.	3.40 (0.77)	4.35 (0.99)	103.5 (10.9)	77.8 (4.6)
	COPD	118	58.3 (10.0)	1.69 (0.09)	67 (56.8)	44 (37.3)	49 (41.5)	25 (21.2)	37.0 (15.4-52.7)	n.a.	n.a.	n.a.	n.a.	2.83 (0.79)	4.37 (1.17)	89.3 (14.7)	64.6 (4.9)

	controls	134	51.4 (10.4)	1.69 (0.10)	60 (44.8)	30 (22.4)	68 (50.8)	36 (26.9)	14.8 (3.9-27.0)	n.a.	n.a.	n.a.	n.a.	3.39 (0.81)	4.37 (1.17)	103.7 (10.9)	77.0 (4.7)
	asthma	126	65.8 (7.8)	1.64 (0.09)	33 (26.2)	24 (19)	50 (40)	51 (41)	15.4 (4.5-37.4)	n.a.	n.a.	n.a	n.a.	n.a.	n.a.	n.a.	n.a.
	controls	4.241	69.8 (9.2)	1.66 (0.09)	1,499 (35.3	782 (18)	1,854 (44)	1,605 (38)	20 (7.5-37.5)	n.a.	n.a.	n.a	n.a.	n.a.	n.a.	n.a	n.a.
	COPD	229	79.8 (4.9)	1.67 (0.09)	126 (55)	51 (22)	36 (16)	142 (62)	26 (9.8-45)	n.a.	n.a.	n.a.	n.a.	1.79 (0.55)	2.82 (0.79)	80.2 (19.5)	63.1 (6.7)
RS-I	controls	781	79.1 (4.5)	1.65 (0.09)	306 (39)	49 (6)	299 (38)	433 (55)	16.8 (5.7-36.0)	n.a.	n.a.	n.a.	n.a.	2.39 (0.62)	3.06 (0.81)	115 (17)	78.6 (4.9)
	asthma	58	62.9 (6.8)	1.66 (0.06)	15 (26)	7 (12)	23 (40)	28 (48)	21.6 (6-43.8)	n.a.	n.a.	n.a	n.a.	n.a.	n.a.	n.a.	n.a.
	controls	1.584	64.7 (8.0)	1.68 (0.09)	712 (45)	249 (16)	526 (33)	809 (51)	14 (3.6-31)	n.a.	n.a.	n.a	n.a.	n.a.	n.a.	n.a	n.a.
	COPD	186	72.8 (5.1)	1.71 (0.10)	108 (58)	48 (26)	28 (15)	110 (59)	31.7 (16.4-46.0)	n.a.	n.a.	n.a	n.a.	2.04 (0.74)	3.23 (1.04	77.60 (18.93)	62.7 (7.2)
RS-II	controls	783	72.1 (4.9)	1.67 (0.08)	327 (42)	52 (7)	317 (41)	415 (53)	13.9 (3.7 -28.0)	n.a.	n.a.	n.a	n.a.	2.67 (0.62)	3.39 (0.82)	111.06 (14)	79.0 (4.6)
	asthma	71	54.7 (4.5)	1.70 (0.09)	20 (28)	6 (9)	27 (38)	38 (54)	15.5 (1.2-25.7)	n.a.	n.a.	n.a	n.a.	n.a.	n.a.	n.a.	n.a.
	controls	1.714	55.8 (5.6)	1.71 (0.09)	764 (45)	356 (21)	574 (34)	784 (46)	13.8 (4.0-29.0)	n.a.	n.a.	n.a	n.a.	n.a.	n.a.	n.a	n.a.
	COPD	79	56.9 (5.0)	1.72 (0.10)	40 (51)	32 (41)	19 (24)	28 (35)	28.9 (16.2-44.7)	n.a.	n.a.	n.a	n.a.	2.51 (0.90)	3.95 (1.31)	81.1 (19.21)	63.0 (5.1)
RS-III	controls	824	56.5 (5.5)	1.71 (0.09)	353 (43)	137 (17)	288 (35)	399 (48)	12.5 (3.8-26.6)	n.a.	n.a.	n.a	n.a.	3.35 (0.75)	4.17 (0.97)	112.85 (13.32)	80.6 (5.1)
	asthma	453	54.3 (5.8)	1.69 (0.1)	226 (50)	107 (23.62)	181 (39.96)	165 (36.42)	29.6 (14.1-45.0)	n.a.	n.a.	n.a	n.a.	2.59 (0.87)	3.83 (1.07)	80.9 (20.9)	67.2 (10.8)
	controls	9.203	54.8 (5.7)	1.69 (0.1)	4,318 (47)	2,268 (24.64)	3,691 (40.11)	3,239 (35.20)	26.0 (12-40)	n.a.	n.a.	n.a	n.a.	2.95 (0.77)	3.98 (0.98)	93.31 (16.19)	74.03 (7.53)
	COPD	915	55.6 (5.57)	1.70 (0.09)	506 (55)	522 (57.1)	93 (10.2)	300 (32.8)	39 (29-54)	n.a.	n.a.	n.a.	n.a.	2.01 (0.60)	3.47 (0.93)	62.2 (13.4)	58.1 (8.6)
ARIC	controls	6.610	54.1 (5.67)	1.68 (0.09)	3,042 (46)	1,120 (16.9)	3,096 (46.8)	2394 (36.2)	20.3 (9-34)	n.a.	n.a.	n.a.	n.a.	3.14 (0.70)	4.11 (0.93)	99.5 (11.1)	76.6 (4.4)
	asthma	267	61.1 (9.6)	1.68 (0.10)	119 (45)	29 (11)	112 (58)	124 (47)	20 (6-41.3)	n.a.	n.a.	n.a.	n.a.	2.29 (0.73)	3.35 (0.93)	80.1 (17.6)	68.4 (10.5)
	controls	2.381	63.0 (10.2)	1.69 (0.10)	1149 (48)	263 (11)	1,061 (55)	1,053 (44)	19 (6.6-37.8)	n.a.	n.a.	n.a.	n.a.	2.60 (0.76)	3.53 (1.00)	92.9 (16.3)	73.9 (8.1)
	COPD	104	67.1 (8.9)	1.70 (0.09)	51 (49)	19 (18)	15 (14)	70 (67)	37 (22-64)	n.a.	n.a.	n.a.	n.a.	1.72 (0.55)	3.06 (0.87)	61.0 (12.7)	56.6 (9.2)
MESA	controls	979	66.0 (10.0)	1.68 (0.10)	467 (48)	55 (6)	446 (46)	478 (49)	17.3 (7-36)	n.a.	n.a.	n.a.	n.a.	2.72 (0.73)	3.61 (0.96)	98.2 (12.0)	75.5 (5.8)

a - blood bank controls, no demographic data available

b - NELSON controls, all demographic data shown for 946 individuals

Table S2. Linkage disequilibrium between top 20 SNPs in asthma-COPD meta-analysis with highest significance level

SNP	Proxy	Distance	RSquared	DPrime	Chromosome	Coordinate_HG18
rs2112101	rs6728750	2055	0,867	1	chr2	15822185
rs2112101	rs2544534	3787	0,863	0,963	chr2	15823917
rs2112101	rs1477253	7778	0,863	0,963	chr2	15827908
rs2112101	rs6728667	2026	0,838	1	chr2	15822156
rs2112101	rs2693008	10340	0,828	0,927	chr2	15830470
rs2112101	rs2544523	17644	0,609	0,84	chr2	15837774
rs2112101	rs2693019	19609	0,609	0,84	chr2	15839739
rs6728667	rs6728750	29	0,966	1	chr2	15822185
rs6728667	rs2544534	1761	0,9	1	chr2	15823917
rs6728667	rs1477253	5752	0,9	1	chr2	15827908
rs6728667	rs2112101	2026	0,838	1	chr2	15820130
rs6728667	rs2693008	8314	0,804	0,962	chr2	15830470
rs6728667	rs2544523	15618	0,604	0,915	chr2	15837774
rs6728667	rs2693019	17583	0,604	0,915	chr2	15839739
rs6728750	rs6728667	29	0,966	1	chr2	15822156
rs6728750	rs2544534	1732	0,932	1	chr2	15823917
rs6728750	rs1477253	5723	0,932	1	chr2	15827908
rs6728750	rs2112101	2055	0,867	1	chr2	15820130
rs6728750	rs2693008	8285	0,833	0,963	chr2	15830470
rs6728750	rs2544523	15589	0,627	0,916	chr2	15837774
rs6728750	rs2693019	17554	0,627	0,916	chr2	15839739
rs2544534	rs1477253	3991	1	1	chr2	15827908
rs2544534	rs6728750	1732	0,932	1	chr2	15822185
rs2544534	rs6728667	1761	0,9	1	chr2	15822156
rs2544534	rs2112101	3787	0,863	0,963	chr2	15820130
rs2544534	rs2693008	6553	0,83	0,927	chr2	15830470
rs2544534	rs2544523	13857	0,676	0,918	chr2	15837774
rs2544534	rs2693019	15822	0,676	0,918	chr2	15839739
rs1477253	rs2544534	3991	1	1	chr2	15823917
rs1477253	rs6728750	5723	0,932	1	chr2	15822185
rs1477253	rs6728667	5752	0,9	1	chr2	15822156
rs1477253	rs2112101	7778	0,863	0,963	chr2	15820130
rs1477253	rs2693008	2562	0,83	0,927	chr2	15830470
rs1477253	rs2544523	9866	0,676	0,918	chr2	15837774
rs1477253	rs2693019	11831	0,676	0,918	chr2	15839739
rs2693008	rs6728750	8285	0,833	0,963	chr2	15822185
rs2693008	rs1477253	2562	0,83	0,927	chr2	15827908
rs2693008	rs2544534	6553	0,83	0,927	chr2	15823917
rs2693008	rs2112101	10340	0,828	0,927	chr2	15820130
rs2693008	rs6728667	8314	0,804	0,962	chr2	15822156
rs2693008	rs2544523	7304	0,766	0,96	chr2	15837774

rs2693008	rs2693019	9269	0,766	0,96	chr2	15839739
rs2544523	rs2693019	1965	1	1	chr2	15839739
rs2544523	rs2693008	7304	0,766	0,96	chr2	15830470
rs2544523	rs1477253	9866	0,676	0,918	chr2	15827908
rs2544523	rs2544534	13857	0,676	0,918	chr2	15823917
rs2544523	rs6728750	15589	0,627	0,916	chr2	15822185
rs2544523	rs2112101	17644	0,609	0,84	chr2	15820130
rs2544523	rs6728667	15618	0,604	0,915	chr2	15822156
rs2693019	rs2544523	1965	1	1	chr2	15837774
rs2693019	rs2693008	9269	0,766	0,96	chr2	15830470
rs2693019	rs1477253	11831	0,676	0,918	chr2	15827908
rs2693019	rs2544534	15822	0,676	0,918	chr2	15823917
rs2693019	rs6728750	17554	0,627	0,916	chr2	15822185
rs2693019	rs2112101	19609	0,609	0,84	chr2	15820130
rs2693019	rs6728667	17583	0,604	0,915	chr2	15822156
rs10036292	rs10043228	1177	1	1	chr5	1,16E+08
rs10036292	rs7727882	18636	1	1	chr5	1,16E+08
rs10036292	rs254149	10049	0,292	1	chr5	1,16E+08
rs10043228	rs10036292	1177	1	1	chr5	1,16E+08
rs10043228	rs7727882	17459	1	1	chr5	1,16E+08
rs10043228	rs254149	8872	0,292	1	chr5	1,16E+08
rs254149	rs7727882	8587	0,292	1	chr5	1,16E+08
rs254149	rs10043228	8872	0,292	1	chr5	1,16E+08
rs254149	rs10036292	10049	0,292	1	chr5	1,16E+08
rs7727882	rs10043228	17459	1	1	chr5	1,16E+08
rs7727882	rs10036292	18636	1	1	chr5	1,16E+08
rs7727882	rs254149	8587	0,292	1	chr5	1,16E+08
rs17069787	rs9534578	3353	0,901	1	chr13	46741378
rs17069787	rs7989394	139445	0,49	0,7	chr13	46877470
rs17069787	rs7994542	976	0,489	0,894	chr13	46739001
rs7994542	rs7989394	138469	0,611	1	chr13	46877470
rs7994542	rs9534578	2377	0,551	1	chr13	46741378
rs7994542	rs17069787	976	0,489	0,894	chr13	46738025
rs9534578	rs17069787	3353	0,901	1	chr13	46738025
rs9534578	rs7994542	2377	0,551	1	chr13	46739001
rs9534578	rs7989394	136092	0,548	0,78	chr13	46877470
rs7989394	rs7994542	138469	0,611	1	chr13	46739001
rs7989394	rs9534578	136092	0,548	0,78	chr13	46741378
rs7989394	rs17069787	139445	0,49	0,7	chr13	46738025

Table S3. Heterogeneity between asthma and COPD cohorts.

			Meta1 (DAG and NELSON)		Meta2 (LifeLines 1)		Meta3 (LifeLines 2)		Meta4 (SAPALDIA)		Meta5 (RS-I)		Meta6 (RS - II)		Meta7 (RS-III)		Meta8 (MESA)		Meta9 (ARIC)	
CHR	BP	SNP	Q	I	Q	I	Q	I	Q	I	Q	I	Q	I	Q	I	Q	I	Q	I
2	15820130	rs2112101	0,81	0	0,25	25,08	0,08	68,46	0,85	0	0,94	0	0,75	0	0,96	0	0,06	70,94	0,31	1,83
2	15822156	rs6728667	0,6	0	0,3	6,63	0,06	71,84	0,86	0	0,83	0	0,85	0	0,91	0	0,09	64,21	0,17	46,59
2	15822185	rs6728750	0,61	0	0,31	4,67	0,05	73,63	0,95	0	0,78	0	0,81	0	0,92	0	0,95	0	0,16	48,17
2	15823917	rs2544534	0,58	0	0,13	56,84	0,03	78,03	0,97	0	0,69	0	0,94	0	0,85	0	0,1	62,22	0,16	49,26
2	15827908	rs1477253	0,58	0	0,11	60,65	0,05	74,04	0,97	0	0,65	0	0,97	0	0,94	0	0,09	64,99	0,15	51,95
2	15830470	rs2693008	0,57	0	0,38	0	0,03	79,96	1	0	0,7	0	0,7	0	0,91	0	0,07	69,91	0,17	47,7
2	15837774	rs2544523	0,5	0	0,47	0	0,05	74,58	0,81	0	0,65	0	0,64	0	1	0	0,05	74,92	0,19	41,64
2	15839739	rs2693019	0,44	0	0,49	0	0,04	76,24	0,78	0	0,64	0	0,63	0	1	0	0,05	73,87	0,19	41,83
2	15840892	rs1363058	0,84	0	0,33	0	0,81	0	0,81	0	0,78	0	0,68	0	0,65	0	0,04	77,24	0,13	56,14
2	15843619	rs2544527	0,33	0	0,76	0	0,1	63,74	0,94	0	0,42	0	0,49	0	0,74	0	0,26	22,74	0,19	42,46
5	115623770	rs10036292	0,52	0	0,85	0	0,07	68,61	0,06	72,49	0,35	0	0,6	0	0,99	0	0,33	0	0,29	12,2
5	115624947	rs10043228	0,49	0	0,92	0	0,08	67,8	0,06	72,04	0,34	0	0,6	0	0,99	0	0,33	0	0,29	11,34
5	115633819	rs254149	0,73	0	0,3	6,66	0,15	52,3	0,03	77,59	0,15	52,76	0,73	0	0,84	0	0,57	0	0,88	0
5	116557808	rs7718941	0,17	47,1	0,28	14,6	0,12	57,71	0,46	0	0,39	0	0,04	76,07	0,61	0	0,53	0	0,64	0
13	46725490	rs7985155	0,82	0	0,47	0	0,94	0	0,64	0	0,82	0	0,12	58,46	0,94	0	0,33	0	0,94	0
13	46728991	rs4391953	0,94	0	0,71	0	0,63	0	0,62	0	0,82	0	0,12	58,19	0,9	0	0,31	2,72	0,98	0
13	46737339	rs17069785	0,08	67,32	0,76	0	0,91	0	0,31	3,14	0,93	0	0,19	41,77	0,97	0	0,11	61,26	0,88	0
13	46738025	rs17069787	0,92	0	0,9	0	0,9	0	0,52	0	0,91	0	0,12	59,64	0,87	0	0,3	7,18	0,98	0
13	46739001	rs7994542	0,45	0	0,3	5,22	0,47	0	0,65	0	0,89	0	0,31	4,51	0,59	0	0,21	36,93	0,64	0
13	46741378	rs9534578	0,61	0	0,93	0	0,85	0	0,6	0	0,74	0	0,15	52,39	0,84	0	0,32	0	0,95	0

References

28. Fehrmann RS, Jansen RC, Veldink JH, Westra HJ, Arends D et al. (2011) Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. PLoS Genet 7: e1002197.

36. Soler AM, Loth DW, Wain LV, Gharib SA, Obeidat M et al. (2011) Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nat Genet 43: 1082-1090.

27. van Iersel CA, de Koning HJ, Draisma G, Mali WP, Scholten ET et al. (2007) Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). Int J Cancer 120: 868-874.

29. Hao K, Bosse Y, Nickle DC, Pare PD, Postma DS et al. (2012) Lung eQTLs to help reveal the molecular underpinnings of asthma. PLoS Genet 8: e1003029.

30. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4: 249-264.

32. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q (2008) GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome Biol 9 Suppl 1: S4.

33. Chang JT, Nevins JR (2006) GATHER: a systems approach to interpreting genomic signatures. Bioinformatics 22: 2926-2933.

Supplementary References

S1. Stolk RP, Rosmalen JG, Postma DS, de Boer RA, Navis G et al. (2008) Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. Eur J Epidemiol 23: 67-74.

S2. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV et al. (2002) Multi-ethnic study of atherosclerosis: objectives and design. Am J Epidemiol 156: 871-881.

S3. Manichaikul A, Chen WM, Williams K, Wong Q, Sale MM et al. (2012) Analysis of family- and population-based samples in cohort genome-wide association studies. Hum Genet 131: 275-287.

S4. Kaufman JD, Adar SD, Allen RW, Barr RG, Budoff MJ et al. (2012) Prospective study of particulate air pollution exposures, subclinical atherosclerosis,

and clinical cardiovascular disease: The Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Am J Epidemiol 176: 825-837.

S5. 1989) The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am J Epidemiol 129: 687-702.

S6. Hankinson JL, Odencrantz JR, Fedan KB (1999) Spirometric reference values from a sample of the general U.S. population. Am J Respir Crit Care Med 159: 179-187.

S7. Wilk JB, Shrine NR, Loehr LR, Zhao JH, Manichaikul A et al. (2012) Genome-wide association studies identify CHRNA5/3 and HTR4 in the development of airflow obstruction. Am J Respir Crit Care Med 186: 622-632.

S8. Browning BL, Browning SR (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet 84: 210-223.

S9. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol 34: 816-834.

S10. Hofman A, van Duijn CM, Franco OH, Ikram MA, Janssen HL et al. (2011) The Rotterdam Study: 2012 objectives and design update. Eur J Epidemiol 26: 657-686.

S11. Manichaikul A, Naj AC, Herrington D, Post W, Rich SS et al. (2012) Association of SCARB1 variants with subclinical atherosclerosis and incident cardiovascular disease: the multi-ethnic study of atherosclerosis. Arterioscler Thromb Vasc Biol 32: 1991-1999.

S12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559-575.

S13. Aulchenko YS, Struchalin MV, van Duijn CM (2010) ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 11: 134.

S14. R Development Core Team (2011) R: A Language and Environment for Statistical Computing, version .