ORIGINAL ARTICLE

Exome Array Analysis Identifies a Common Variant in *IL27* Associated with Chronic Obstructive Pulmonary Disease

Brian D. Hobbs[1,2], Margaret M. Parker[1], Han Chen[3], Taotao Lao[1], Megan Hardin[1,2], Dandi Qiao[1], Iwona Hawrylkiewicz[4], Pawel Sliwinski[4], Jae-Joon Yim[5], Woo Jin Kim[6], Deog Kyeom Kim[7], Peter J. Castaldi[1,8], Craig P. Hersh[1,2], Jarrett Morrow[1], Bartolome R. Celli[2], Victor M. Pinto-Plata[9], Gerald J. Criner[10], Nathaniel Marchetti[11], Raphael Bueno[12], Alvar Agustí[13], Barry J. Make[14], James D. Crapo[14], Peter M. Calverley[15], Claudio F. Donner[16], David A. Lomas[17], Emiel F. M. Wouters[18], Jorgen Vestbo[19], Peter D. Paré[20], Robert D. Levy[20], Stephen I. Rennard[21,22], Xiaobo Zhou[1], Nan M. Laird[5], Xihong Lin[3], Terri H. Beaty[23], Edwin K. Silverman[1,2], and Michael H. Cho[1,2]; on behalf of the NETT Genetics Investigators, ECLIPSE Investigators, COPDGene Investigators*, and International COPD Genetics Network Investigators*

[1]Channing Division of Network Medicine, [2]Division of Pulmonary and Critical Care Medicine, [8]Division of General Medicine and Primary Care, and [12]Division of Thoracic Surgery, Brigham and Women's Hospital, Boston, Massachusetts; [3]Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, Massachusetts; [4]National Tuberculosis and Lung Disease Research Institute, Warsaw, Poland; [5]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul National University College of Medicine, Seoul, South Korea; [6]Kangwon National University, Chuncheon, South Korea; [7]Seoul National University Boramae Medical Center, Seoul, South Korea; [9]Department of Critical Care Medicine and Pulmonary Disease, Baystate Medical Center, Springfield, Massachusetts; [10]Division of Pulmonary and Critical Care Medicine and [11]Department of Thoracic Medicine and Surgery, Temple University School of Medicine, Philadelphia, Pennsylvania; [13]Thorax Institute, Hospital Clinic, IDIBAPS, University of Barcelona, CIBERES, Barcelona, Spain; [14]National Jewish Health, Denver, Colorado; [15]University of Liverpool, Liverpool, United Kingdom; [16]Mondo Medico di I.F.I.M. srl, Multidisciplinary and Rehabilitation Outpatient Clinic, Borgomanero (NO), Italy; [17]University College London, London, United Kingdom; [18]University Hospital Maastricht, Maastricht, the Netherlands; [19]University of Manchester, Manchester, United Kingdom; [20]University of British Columbia, Vancouver, British Columbia, Canada; [21]Pulmonary and Critical Care Medicine, University of Nebraska Medical Center, Omaha, Nebraska; [22]Clinical Discovery Unit, AstraZeneca, Cambridge, United Kingdom; and [23]Department of Epidemiology, Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, Maryland

ORCID IDs: 0000-0001-9564-0745 (B.D.H.); 0000-0002-4907-1657 (M.H.C.).

Abstract

Rationale: Chronic obstructive pulmonary disease (COPD) susceptibility is in part related to genetic variants. Most genetic studies have been focused on genome-wide common variants without a specific focus on coding variants, but common and rare coding variants may also affect COPD susceptibility.

Objectives: To identify coding variants associated with COPD.

Methods: We tested nonsynonymous, splice, and stop variants derived from the Illumina HumanExome array for association with COPD in five study populations enriched for COPD. We evaluated single variants with a minor allele frequency greater than 0.5% using logistic regression. Results were combined using a fixed effects meta-analysis. We replicated novel single-variant associations in three additional COPD cohorts.

Measurements and Main Results: We included 6,004 control subjects and 6,161 COPD cases across five cohorts for analysis. Our top result was rs16969968 ($P = 1.7 \times 10^{-14}$) in *CHRNA5*, a locus previously associated with COPD susceptibility and nicotine dependence. Additional top results were found in *AGER*, *MMP3*, and *SERPINA1*. A nonsynonymous variant, rs181206, in *IL27* ($P = 4.7 \times 10^{-6}$) was just below the level of exome-wide significance but attained exome-wide significance ($P = 5.7 \times 10^{-8}$) when combined with results from other cohorts. Gene expression datasets revealed an association of rs181206 and the surrounding locus with expression of multiple genes; several were differentially expressed in COPD lung tissue, including *TUFM*.

Conclusions: In an exome array analysis of COPD, we identified nonsynonymous variants at previously described loci and a novel exome-wide significant variant in *IL27*. This variant is at a locus previously described in genome-wide associations with diabetes, inflammatory bowel disease, and obesity and appears to affect genes potentially related to COPD pathogenesis.

Keywords: chronic obstructive pulmonary disease; genetics; exome; IL-27

(Received in original form October 21, 2015; accepted in final form January 15, 2016)

*A complete list of members may be found before the beginning of the References.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Crit Care Med Vol 194, Iss 1, pp 48–57, Jul 1, 2016

Originally Published in Press as 10.1164/rccm.201510-2053OC on January 18, 2016
Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the Subject: Chronic obstructive pulmonary disease (COPD) is a complex disease, with susceptibility modified in part by genetic variants. Genome-wide association studies have identified genetic variants that impact COPD susceptibility; however, the majority of COPD heritability remains unexplained. Genome-wide association studies have not been optimized to capture coding genetic variants that affect COPD susceptibility.

What This Study Adds to the Field: In our analysis of the protein-coding region of the genome (the exome), we identified an *IL27* variant associated with COPD susceptibility. This variant is highly correlated with genetic variants associated with diabetes, inflammatory bowel disease, and obesity, and it also regulates expression of several genes in lung tissue and in blood, including *TUFM*.

Chronic obstructive pulmonary disease (COPD) is a highly prevalent condition that is projected to be the third leading cause of death worldwide by 2020 (1). COPD is a complex disease whose development depends on both environmental and genetic risk factors. The genetic component of COPD is demonstrated by studies showing strong familial aggregation. Furthermore, several genome-wide association studies (GWASs) have implicated COPD risk loci at chromosomal regions 15q25 and 19q13, and near genes *HHIP*, *FAM13A*, *RIN3*, *MMP12*, and *TGFB2* (2–6). Although multiple COPD susceptibility loci have been identified through GWASs, known COPD risk loci explain only a small fraction of the observed variability in risk (7). Therefore, a large portion of the heritability of COPD has yet to be explained.

Analyzing coding variation may reveal novel pathobiology contributing to the development of COPD. While the majority of genome-wide association variants in complex diseases are likely regulatory, coding variants, accounting for only 1% of the genome by size, are overrepresented among these associations (8). In addition, uncommon (minor allele frequency [MAF], 1–5%) and rare (MAF $<$ 1%) genetic variants, not well captured by GWAS arrays, are one of several possible causes of "missing heritability" (9). Rare coding variants are important in COPD susceptibility, as illustrated by alpha-1 antitrypsin deficiency (AATD), a genetic disorder in which rare variants in a serine protease inhibitor (*SERPINA1*) greatly impact COPD susceptibility (10). Traditional GWAS genotyping arrays have a large portion of genetic markers outside the coding genome, so many GWAS associations in complex disease have yet to be functionally classified. Restricting analysis to the coding regions of the genome (the exome) could enable more direct biological and functional interpretation of association study results. Exome genotyping arrays have been developed to specifically query genetic variation in the coding genome. These exome arrays contain approximately 250,000 putatively functional (nonsynonymous, stop, or splice) single-nucleotide polymorphism (SNP) probes (11). In complex disease phenotypes such as insulin secretion, type 2 diabetes risk, blood lipid levels, and fasting glucose levels, this exome array technology has already been employed to add to working knowledge of these phenotypes (12–15). We hypothesized that an exome genotyping analysis would identify coding variants showing an association with COPD susceptibility.

Some of the results of this study were previously reported in the form of an abstract (16).

Methods

Study Cohorts

The primary analysis included individuals from two family-based cohorts identified through a proband with COPD (the BECOPD [Boston Early-Onset COPD] study and the ICGN [International COPD Genetics Network] study), as well as three case–control studies (the TCGS [Transcontinental COPD Genetics Study] in Poland and in Korea, and the COPDGene [Genetic Epidemiology of COPD] study, which included non-Hispanic white [NHW] and African American [AA] individuals). Institutional review board approval and written informed consent were obtained for all of these cohorts. Persons were labeled unaffected with COPD if they had an FEV_1/FVC ratio greater than or equal to 0.7 and an FEV_1 greater than or equal to 80% of the predicted value. Moderate to severe COPD was defined by an FEV_1/FVC ratio less than 0.7 and an FEV_1 less than 80% of the predicted value, and severe COPD was defined by an FEV_1/FVC ratio less than 0.7 and an FEV_1 less than 50% of the predicted value.

Single-Variant Testing for Association with COPD

All individuals were genotyped using the Illumina HumanExome arrays (v1.1 and v1.2; Illumina, San Diego, CA).

The project was supported by National Institutes of Health grants T32 HL007427, K08 HL097029 (M.H.C.), R01 HL113264 (M.H.C.), R01 HL089856 (E.K.S.), R01 HL089897 (J.D.C.), R01 HL084323 (E.K.S.), R01 HL127200 (X.Z.), R21 HL120794 (X.Z.), P01 HL083069 (E.K.S.), P01 HL105339 (E.K.S.), and P01 HL114501 (Augustine M. K. Choi). The project was also supported by the Alpha-1 Foundation. The International COPD Genetics Network was funded by GlaxoSmithKline. The COPDGene project (clinicaltrials.gov identifier NCT00608764) is also supported by the COPD Foundation through contributions made to an industry advisory board comprised of AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Novartis, Pfizer, Siemens, and Sunovion Pharmaceuticals.

Author Contributions: B.D.H. and M.H.C.: participated in study conception and design, data acquisition, analysis of data, interpretation of analysis results, manuscript drafting, and manuscript editing for intellectual content; M.M.P.: participated in data acquisition, analysis of data, interpretation of analysis results, and manuscript editing for intellectual content; H.C., X.L., N.M.L., T.H.B., and E.K.S.: participated in study design, interpretation of analysis results, and manuscript editing for intellectual content; T.L., M.H., D.Q., I.H., P.S., J.-J.Y., W.J.K., D.K.K., P.J.C., C.P.H., J.M., B.R.C., V.M.P.-P., G.J.C., N.M., R.B., A.A., B.J.M., J.D.C., P.M.C., C.F.D., D.A.L., E.F.M.W., J.V., P.D.P., R.D.L., S.I.R., and X.Z.: participated in data acquisition and manuscript editing for intellectual content.

Correspondence and requests for reprints should be addressed to Michael H. Cho, M.D., M.P.H., Channing Division of Network Medicine, Brigham and Women's Hospital, 181 Longwood Avenue, Boston, MA 02115. E-mail: michael.cho@channing.harvard.edu

Single-variant testing was performed on nonsynonymous, stop, and splice-site exome array variants with cohort-based MAF greater than 0.5%. Analyses were adjusted for age, sex, pack-years of smoking, and ancestry-based principal components. Association analysis for the TCGS-Korea, TCGS-Poland, COPDGene AA, and COPDGene NHW cohorts were performed using an additive genetic model through logistic regression in PLINK v1.9 (17, 18). To match the analytical methods for gene-based association testing (*see* later), the BEOCOPD and ICGN family-based cohorts were analyzed using logistic regression via generalized estimating equations implemented in GWAF version 2.2 (19, 20). Results were combined in a fixed effects meta-analysis using METAL (version 2011-03-25) (21, 22). Top coding variants ($P < 1 \times 10^{-5}$) not previously implicated in COPD susceptibility were further assessed with previously reported genotyped and imputed GWAS data derived from the ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints), GenKOLS (Genetics of Chronic Obstructive Pulmonary Disease, Norway), and NETT/NAS (National Emphysema Treatment Trial and Normative Aging Study) studies using the same analysis and covariates (with the exception of NETT/NAS, in which all control subjects are male) (2, 4).

Gene-based Testing for Association with COPD

Statistical tests that combine multiple variants in a gene may be more powerful than tests of individual variants, particularly when the variants are individually rare. Thus, in each analysis cohort (BEOCOPD and ICGN, TCGS-Korea, TCGS-Poland, COPDGene AA, and COPDGene NHW) and for each gene, we combined nonsynonymous, splice, and stop variants with a cohort-specific MAF less than 5%. Our analysis was limited to genes with a minimum minor allele count of 5. We used SNP-set (Sequence) Kernel Association Test (SKAT)-O (23), for all gene-based COPD association testing. The BEOCOPD and ICGN family-based cohorts were combined, using a covariate to denote study, to meet minimum sample size requirements of the statistical model, and then analyzed using a logistic mixed model in GMMAT version 0.5 (24). We performed a meta-analysis of SKAT-O results using SKAT version 1.0.1 (25, 26) and MetaSKAT version 0.40 (27), assuming homogeneous effects of SNPs across study cohorts. For both single-variant and gene-based meta-analyses, we adjusted for multiple testing using the Bonferroni correction, requiring that a single variant or gene be present in at least three of the five meta-analysis groups.

Secondary Analyses

We performed meta-analysis of white cohorts (European ancestry) to assess possible differences in results due to racial heterogeneity among cohorts in the primary analysis. We additionally performed an analysis with cases defined by severe COPD (Global Initiative for Chronic Obstructive Lung Disease spirometry stages 3–4) to search for enrichment of any association signals, as has been done in prior COPD GWASs (6). Finally, we evaluated the association statistics for a list of 119 genes identified from previous GWASs of COPD and lung function as well as genes implicated in Mendelian diseases leading to syndromes that included COPD (subsequently referred to as candidate genes) (*see* Table E2 in the online supplement) (6, 28, 29).

Additional cohort information and details of the methods used in this study are available in the online supplement.

Results

Participant Details

After subject quality control measures, a total of 6,004 control subjects and 6,161 cases were available for analysis. Baseline characteristics are shown in Table 1. As the two family-based cohorts, BEOCOPD and ICGN, were combined (*see* Methods section), we considered five analysis groups: (*1*) BEOCOPD and ICGN, (*2*) TCGS-Korea, (*3*) TCGS-Poland, (*4*) COPDGene AA, and (*5*) COPDGene NHW.

Marker Filtering for Analysis and Calculation of Statistical Significance Level

The number of nonsynonymous, splice, and stop variants present in at least three of the five analysis groups in the meta-analysis was 21,947, for a Bonferroni-adjusted, exome-wide, single-variant significance threshold of 2.3×10^{-6}. For gene-based SKAT-O meta-analysis, 12,133 genes in at least three cohorts required a significance threshold of 4.1×10^{-6}. Table E3 illustrates the number of variants and gene sets considered in both single-variant testing and gene-based testing with SKAT-O.

Single-Variant Meta-analysis for Association with COPD Affection Status

In our COPD affection status single-variant meta-analysis, we identified a single exome-wide significant variant, rs16969968 in *CHRNA5*, yielding $P = 1.7 \times 10^{-14}$. The *CHRNA5* locus is known to be associated with COPD susceptibility and nicotine addiction (2, 4, 6). No novel variants reached exome-wide significance in our meta-analysis; however, rs181206 in the interleukin 27 gene (*IL27*) on chromosome 16 was just below the exome-wide significance threshold with meta-analysis ($P = 4.7 \times 10^{-6}$). Top results from the meta-analysis across the five analysis cohorts are shown in Table 2. Of the top 10 results, the majority were common variants with overall effect allele frequency greater than 5%; however, variants in *AGER*, *FAM208B*, *CRAMP1L*, and *SERPINA1* were uncommon, with effect allele frequency between 1% and 5%. Note that the 10th most significant variant association ($P = 1.4 \times 10^{-4}$) was rs28929474, the *SERPINA1* Z allele, which causes AATD (10). Although persons with known AATD were excluded from all study cohorts, three individuals in the TCGS-Poland study cohort were found to be homozygous for the Z allele. Removal of these three subjects diminished, but did not eliminate, the association for this variant in the TCGS-Poland study population (*P* value increased from 0.009 to 0.04) and in the overall meta-analysis results (*P* value increased from 1.4×10^{-4} to 2.4×10^{-4}).

The severe COPD association analysis (Table E4) showed stronger association (lower *P* values and higher effect sizes) at several known COPD loci, including the regions around *CHRNA5* and *ADAMTS7*, *MMP3*, and *SERPINA1*. No additional exome-wide significant single-variant associations were discovered in this analysis. To assess nonsynonymous, splice, and stop variant associations with COPD in a more racially homogeneous set of individuals, we performed a meta-analysis limited to the white (European ancestry) cohorts (BEOCOPD and ICGN combined,

Table 1. Cohort Compositions and Demographics, Separated by COPD Status (Control Subjects, Moderate to Severe COPD, and Severe COPD)

	BEOCOPD (*n = 201 Pedigrees*)	ICGN (*n = 1,103 Pedigrees*)	TCGS-Korea	TCGS-Poland	COPDGene African Americans	COPDGene Non-Hispanic Whites
Control subjects						
Number of subjects	560	696	219	307	1,715	2,507
Males, %	41.6	48.3	96.8	67.4	58.1	49.4
Age, yr	39.3 (26.3–51.5)	54.9 (48.1–60.7)	53 (46–59)	58.3 (54.3–62.9)	51.8 (48.2–55.6)	59.3 (52.4–65.9)
Pack-years of smoking	1.66 (0–15.9)	25.1 (15.7–38.8)	25.5 (16.5–34)	32.2 (22.9–41.1)	32.7 (21.2–43.8)	35 (23.4–47)
FEV_1, % predicted	94.5 (86.9–103)	97.5 (88.6–108)	93.6 (88.2–100)	102 (92.5–111)	96.6 (88.9–106)	95.5 (88–104)
Moderate to severe cases						
Number of subjects	366	1,769	149	304	796	2,777
Males, %	39.9	58.6	99.3	70.1	55.5	55.7
Age, yr	50.8 (46.4–59)	59.4 (54.9–63.6)	69 (65–74)	62.2 (57.8–67.7)	58.2 (52.4–64.8)	65.2 (59–71)
Pack-years of smoking	37.5 (25.4–54)	45 (32.1–64.5)	40 (31–52)	40.3 (30–52.8)	37.8 (25.2–51.6)	49.8 (38–70.5)
FEV_1, % predicted	29.2 (18.5–51.8)	39.2 (26.8–52.7)	33.2 (27.4–40.1)	28.7 (22.3–35.4)	54 (39.3–67.1)	50 (34.6–65.5)
Severe cases						
Number of subjects	251	1,099	149	304	343	1,385
Males, %	33.1	60.9	99.3	70.1	58.0	57.7
Age, yr	50 (46.3–52.4)	59.5 (55.4–63.4)	69 (65–74)	62.2 (57.8–67.7)	60.1 (54.2–66.2)	65.8 (59.9–71.1)
Pack-years of smoking	39.1 (26.4–54)	46.4 (34–68)	40 (31–52)	40.3 (30–52.8)	38.5 (26–54)	52 (40–73.5)
FEV_1, % predicted	22.6 (15.6–29.8)	29.9 (22.1–37.5)	33.2 (27.4–40.1)	28.7 (22.3–35.4)	36.5 (26.6–43.3)	34.6 (26.2–42.2)

Definition of abbreviations: BEOCOPD = Boston Early-Onset COPD; COPD = chronic obstructive pulmonary disease; COPDGene = Genetic Epidemiology of COPD; ICGN = International COPD Genetics Network; TCGS = Transcontinental COPD Genetics Study.
Values represent number, percent, or the median and interquartile range (lower quartile–upper quartile).

TCGS-Poland, and COPDGene NHW). The results were similar overall (Table E5).

Gene-based Meta-analysis for Gene Association with COPD Affection Status

In gene-based association analysis, 16,006 genes were evaluated; 12,133 genes were present in at least three of the five analysis groups. Our top association was in the gene *KIAA0020*, but it did not reach exome-wide significance ($P = 1.2 \times 10^{-4}$). Our second most significant association was in *AGER*. When the variant rs2070600 (a top association from single-variant analysis) was removed from the *AGER* gene set, the association signal was markedly attenuated ($P = 0.46$). The top 10 most significant P values from SKAT-O meta-analysis of moderate to severe COPD are provided in Table 3. Analyses of severe COPD and limiting the sample to white subjects did not result in any significant associations (Tables E6 and E7).

Candidate Gene Evaluation

In the meta-analysis of single nonsynonymous, splice, and stop variants, 14 variants in candidate genes were nominally significant ($P < 0.05$). Because the majority of our candidate genes were selected on the basis of significance in prior GWASs of COPD and lung function, we examined linkage disequilibrium (LD) between our variants and the previously reported GWAS SNPs. Most of these variants had either a high correlation coefficient (r^2) or normalized coefficient of LD (D′) with previously reported SNPs. The results of these candidate gene evaluations are reported in Tables E8 and E9.

Further Assessment of *IL27* Variant rs181206 in Additional Populations

The nonsynonymous variant rs181206 in *IL27* was our only novel finding close to exome-wide significance; therefore, we sought to determine whether this variant could be tested in additional available cohorts with genome-wide SNP genotyping data. As this variant was common, genotyped or imputed data for this variant were available (*see* Table E10) in all three additional COPD cohorts (NETT/NAS, ECLIPSE, and GenKOLS). When these results for rs181206 were combined with our exome array data in a meta-analysis, the association signal for rs181206 was $P = 5.7 \times 10^{-8}$, exceeding the threshold for exome-wide significance.

We sought evidence for a functional effect of rs181206 in computational prediction models. This variant is predicted to be "probably damaging" by PolyPhen2 (30) and has a combined annotation-dependent depletion (31) scaled score of 23.5, which suggests that this variant has a deleterious effect on the protein. However, we also considered whether this variant (or variants in LD with it) could have effects on gene expression. Our

Table 2. Top 10 Results from Meta-analysis Across All Cohorts of Nonsynonymous, Stop, and Splice Variant Associations with Moderate to Severe COPD

Chr	SNP	Gene	Effect Allele	Effect Allele Overall Frequency	β Value	SE	Test for Heterogeneity		Direction*†	*P* Value
							I^2 Statistic	Q Statistic		
15	rs16969968	*CHRNA5*	A	0.36	0.27	0.035	68	13	+++++	1.7×10^{-14}
16	rs181206	*IL27*	A	0.69	−0.16	0.036	1.2	4.0	−−−−−	4.7×10^{-6}
11	rs8177374	*TIRAP*	T	0.16	0.19	0.045	31	5.8	+−−++	2.6×10^{-5}
6	rs2070600	*AGER*	T	0.042	−0.33	0.079	28	5.6	−−++−	2.7×10^{-5}
11	rs679620	*MMP3*	T	0.48	0.12	0.030	0	2.5	+++++	3.8×10^{-5}
6	rs10499052	*AKD1*	A	0.27	0.15	0.036	0	1.8	+++++	4.0×10^{-5}
6	rs59056467	*MICAL1*	T	0.32	0.13	0.033	0	1.5	+−+++	4.7×10^{-5}
10	rs41290259	*FAM208B*	A	0.98	0.58	0.15	0	1.2	+??++	8.8×10^{-5}
16	rs61746451	*CRAMP1L*	T	0.011	−0.57	0.14	0	2.4	−?+−−	8.9×10^{-5}
14	rs28929474	*SERPINA1*	T	0.021	0.45	0.12	50	4.0	+?+?+	1.4×10^{-4}

Definition of abbreviation: Chr = chromosome; COPD = chronic obstructive pulmonary disease; SNP = single-nucleotide polymorphism.
Only variants observed in at least three of the five analysis cohorts are reported.
*The direction of effect is reported for each analysis cohort. The symbols from left to right represent (*1*) Boston Early-Onset COPD study and International COPD Genetics Network study, (*2*) Transcontinental COPD Genetics Study (TCGS) Korea cohort, (*3*) TCGS Poland cohort, (*4*) Genetic Epidemiology of COPD Study (COPDGene) African American cohort, and (*5*) COPDGene non-Hispanic white cohort.
†The symbol for each cohort is "+" for effect allele association with increased COPD susceptibility, "−" for association with decreased COPD susceptibility, and "?" if the variant was not present at a sufficient allele count in the analysis cohort.

nonsynonymous *IL27* variant, rs181206, is in LD ($r^2 > 0.5$ in HapMap3 Utah residents of northern and western European ancestry [CEU]) with previously reported noncoding GWAS variants of body mass index (32), early-onset inflammatory bowel disease (33), Crohn's disease (34), type 1 diabetes mellitus (35), and type 1 diabetes autoantibodies (36). Table E11 shows the *P* values from publicly available GWAS results for association of 16p11.2 variants with body mass index, type 1 diabetes, and inflammatory bowel disease. We evaluated the overall COPD association signal (primary analysis plus replication cohort meta-analysis) for rs181206 and the previously reported noncoding GWAS variants, and we report these results in Figure E1. These variants lie in the chromosome 16p11.2 region, which is complex, with 26 genes in a ±500-kb window around our lead association (*see* Table E12). Gene expression effects of rs181206 and the related 16p11.2 noncoding GWAS variants were evaluated in public or previously published studies of expression quantitative trait loci (eQTL) in blood, sputum, and lung. Sixteen different eQTL genes (of which 11 are in the ±500-kb window around variant rs181206) were identified and are shown in Table E13. None of these eQTL sources specifically pointed to *IL27*, but at least one source identified an effect on Tu, translation elongation factor, mitochondrial (*TUFM*) in both the blood and the lung for each variant examined at the 16p11.2 locus.

In an attempt to further explore the activity at the 16p11.2 locus related to COPD and lung function, we assessed lung tissue gene expression obtained from a separate study of former smokers with severe COPD and with normal spirometry (37). Details of this analysis are provided in the online supplement. Several genes in the region, including *CD19*, *LAT*, *SULT1A1*, *SULT1A2*, and *TUFM*, had nominal statistical significance (unadjusted *P* value < 0.05), but did not meet the false discovery rate threshold of

Table 3. Top 10 Results from Meta-analysis Across All Cohorts of SKAT-O Gene-based Association with Moderate to Severe COPD

Gene Set	BEOCOPD and ICGN	TCGS-Korea	TCGS-Poland	COPDGene AA	COPDGene NHW	Meta-analysis
KIAA0020	0.054	0.34	0.51	0.47	1.3×10^{-3}	1.2×10^{-4}
AGER	0.34	0.35	0.48	0.42	1.7×10^{-5}	1.3×10^{-4}
RRP1B	0.33	0.69	0.49	4.9×10^{-5}	0.25	1.6×10^{-4}
SQSTM1	0.045	—	0.74	0.35	2.2×10^{-3}	1.6×10^{-4}
RUFY4	0.060	0.11	0.63	0.14	0.040	2.3×10^{-4}
TYRO3	0.36	—	—	7.8×10^{-5}	0.18	2.8×10^{-4}
IKZF2	0.078	—	0.58	0.12	0.011	4.2×10^{-4}
TAF12	6.6×10^{-3}	—	0.43	—	0.025	4.4×10^{-4}
GON4L	1	0.86	0.40	4.9×10^{-4}	0.24	7.3×10^{-4}
MYEOV2	0.023	—	0.39	9.1×10^{-3}	0.13	9.1×10^{-4}

Definition of abbreviations: AA = African Americans; BEOCOPD = Boston Early-Onset COPD; COPD = chronic obstructive pulmonary disease; COPDGene = Genetic Epidemiology of COPD; ICGN = International COPD Genetics Network; NHW = non-Hispanic whites; SKAT-O = SNP-set (Sequence) Kernel Association Test; TCGS = Transcontinental COPD Genetics Study.
Results are given as *P* values. Gene sets present in at least three of the five cohort-level analyses are reported.

less than 5%, for differential expression between COPD cases and control subjects. *CD19*, *SULT1A1*, and *TUFM* were significantly associated with FEV_1 and FEV_1/FVC ratio with a false discovery rate less than 5%.

Discussion

Despite substantial progress in identifying genetic risk factors for COPD from GWASs, much of the genetic risk for COPD remains undefined. Coding variants are known to predispose individuals to COPD, but most studies to date have been focused on the entire genome, the majority of which is noncoding. We performed a comprehensive assessment of coding variation on the exome array, analyzing both single-variant and gene-based associations with COPD affection status as the primary outcome phenotype, and included five study populations with 6,004 unaffected and 6,161 affected persons. Our study did not identify a large number of coding variants responsible for COPD susceptibility. These results are consistent with studies of other complex diseases showing that while coding variation, accounting for less than 1% of the genome, has a disproportionate contribution to disease susceptibility, the majority of heritability for most complex diseases is likely due to regulatory variation.

Gene-based analyses may have more power to detect an association than single-variant associations. However, our gene-based SKAT-O meta-analysis did not reveal any exome-wide significant genes associated with COPD affection status. We were able to recover evidence for the previously described association of COPD with *AGER*, though this association was driven by a single coding variant (rs2070600) that has been reported to be functional (38, 39). Our top gene-based COPD association was the pumilio domain–containing protein KIAA0020 (*KIAA0020*) ($P = 1.2 \times 10^{-4}$). The *KIAA0020* protein product is also known as the minor histocompatibility antigen HA-8, where minor histocompatibility antigens are T-cell epitopes that are presented by major histocompatibility complexes to function in immune regulation (40). While *KIAA0020* has not previously been studied in COPD, it is expressed in lung tissue (41, 42) and is present at moderate to high levels in bronchial respiratory epithelial cells, pulmonary macrophages, and pneumocytes (43, 44). Future studies with increased sample size are needed to corroborate an association of *KIAA0020* with COPD affection status.

Our top single nonsynonymous variant was rs16969968 in *CHRNA5*, a gene known to be associated with COPD susceptibility. The SNP in *CHRNA5* is in strong LD ($r^2 = 0.89$; $D' = 0.95$) with rs8034191, one of the first-identified COPD associations through GWAS (2). The 15q25 locus encompassing *CHRNA5* is known to be complex, with a prior association with smoking behavior (45), and multiple independent variants explaining the association signals for COPD and smoking behavior (46, 47). A previous mediation analysis found that some but not all of the COPD risk at this locus is attributable to smoking behavior (48). Supporting this data is evidence of association of the 15q25 locus with airflow obstruction, even in never-smokers (49). While we adjusted for pack-years of smoking, we cannot rule out that the observed increase in strength of association of rs16969968 in severe COPD is due to increased smoke exposure. Another top association was with rs679620 in *MMP3*, which is in the same region recently associated with severe COPD in a GWAS meta-analysis (6). The variant rs2070600 in *AGER* is associated with COPD and quantitative computed tomography–identified emphysema, and was first identified as genome-wide significant in GWASs of FEV_1/FVC ratio (50–53).

We also identified a new exome-wide significant association of COPD with the nonsynonymous variant rs181206 in *IL27*. The variant rs181206 is a missense variant in which the minor allele (and COPD risk allele) G results in a leucine-to-proline amino acid change in the IL-27 subunit α-protein (54). Of interest, variants at 16p11.2 have previously been identified in GWASs of early-onset inflammatory bowel disease (33), Crohn's disease (34), type 1 diabetes mellitus (35), and type 1 diabetes autoantibodies (36), adding to existing evidence of pleiotropy in GWASs (55) and to a potential biological link between autoimmune disease and COPD (56–58). Interestingly, effects at this variant appear to be discordant between diabetes and inflammatory bowel disease; the direction of effect that we found at this locus in COPD seems to be concordant with inflammatory bowel disease but discordant with risk to diabetes (59). The former may be of particular interest, given that some studies suggest an increased risk of COPD in patients with inflammatory bowel disease (60, 61).

IL27 (along with *EBI3*) encodes a heterodimeric cytokine complex that has a complex role in immune regulation and has been found to exhibit both pro- and antiinflammatory properties (62). Regarding COPD, Huang and colleagues showed that two *IL27* polymorphisms, c.-964A/G (rs153109, LD with rs181206, $r^2 = 0.49$ in HapMap3 CEU; $r^2 = 0.32$ in HapMap3 Japanese individuals from Tokyo, Japan, and Han Chinese individuals from Beijing, China [JPT+CHB]) and g.2905T/G (rs17855750, LD with rs181206, $r^2 = 0.03$ in HapMap3 CEU, $r^2 = 0.016$ in HapMap3 JPT+CHB), are associated with COPD susceptibility in a Chinese population (63). Cao and colleagues reported higher levels of IL27 in the sputum and plasma of patients with COPD than in those of healthy control subjects, with a negative correlation between IL27 levels and FEV_1 in patients with COPD (64). Angata and colleagues recently suggested serum IL27 levels are a promising biomarker for COPD exacerbation (65). Our *in silico* analysis suggests the amino acid change of rs181206 could impact *IL27* function. Alternatively, this variant could be in LD with a regulatory variant. In a study of lymphoblastoid cell lines, *IL27* expression was decreased in healthy individuals with two copies of the rs1968752 (also in LD with rs181206) risk allele (33).

In fact, despite our hypothesis that using the exome array would identify functional coding variants, our discovered variant may be a marker for broader regulatory effects at the complex 16p11.2 region. An integrative analysis of expression data from the ECLIPSE study with previously published GWAS results identified the 16p11.2 locus as harboring evidence for association with COPD and gene expression and also showed that *CCDC101* in the 16p11.2 locus had evidence of colocalization (46). In previous work, *CCDC101* was also identified in a consensus module in a network-guided approach to COPD (66). Our survey of

eQTL data indicates that *CDC37P1*, *SULT1A1*, and *TUFM* are seen as blood and lung and/or sputum eQTL signals in each of the five examined SNPs at the 16p11.2 locus. Further, *SULT1A1* and *TUFM* are nominally significant for differential expression in COPD (46). Adding to the complexity of the 16p11.2 region, a common, approximately 0.45-Mb 16p11.2 inversion has been associated with the joint susceptibility to asthma and obesity (67). This 16p11.2 inversion encompasses the nonsynonymous variant rs181206. Although the LD structure between rs181206 and the inversion haplotype is not reported, some of the SNPs in LD with rs181206 also appear to be in LD with the inversion haplotype. The inversion genotype was strongly associated with expression levels of *TUFM* ($P = 3.0 \times 10^{-40}$) and *SULT1A4*, *SPNS1*, *SULT1A1*, and *CCDC101* (67). Gonzalez and colleagues (67) proposed that the expression variation of genes in the 16p11.2 region was "most likely" driven by the inversion itself. In the same study, *IL27* levels were somewhat correlated ($P = 2 \times 10^{-2}$) with the inversion genotype, though not to the same degree as *TUFM*, *SULT1A4*, *SPNS1*, *SULT1A1*, and *CCDC101*.

The *TUFM* gene is intriguing in relation to COPD pathogenesis. *TUFM* has been described as interacting with the mitochondria-localized, nucleotide-binding, leucine-rich repeats (*NLRX1*) gene to reduce cytokine response and augment virus-induced autophagy (68, 69). In a recent study, three independent cohorts of patients with COPD were demonstrated to have decreased levels of *NLRX1* mRNA, which was associated with increased disease severity, decreased pulmonary function, poorer quality of life, and worse prognosis (70). The *TUFM* and *NLRX1* relationship, together with the association between variants in the 16p11.2 region and *TUFM* expression levels, and evidence for differential *TUFM* expression in COPD tissue provide support for a potential role of *TUFM* and more generally lend biologic plausibility to the association between COPD susceptibility and the 16p11.2 locus containing rs181206.

Our study has some potential limitations that deserve comment. First, despite our analysis of over 12,000 subjects, lack of statistical power may be a limitation. The number of subjects required to identify association with rare variants likely requires large sample sizes, and our effective sample size was decreased because of imbalances in cases and control subjects within individual cohorts. Second, COPD is a heterogeneous disease (71–73). Even when degree of airflow obstruction is similar, the clinical presentation of individuals can vary markedly, including the severity of emphysema and the frequency of exacerbations. Refining and studying phenotypes related to COPD pathogenesis may give additional power to detect novel genetic associations in COPD. In particular, efforts to define biologically similar subtypes of COPD may lead to improved ability to detect genetic contributions to COPD. The racial and/or ethnic heterogeneity of the COPD cohorts analyzed may also limit our observations. This heterogeneity was most notable in the gene-based analysis, where only 4,519 of 16,006 gene sets were common across all cohorts (where 9,023 gene sets were common across the white cohorts). We performed an analysis across all cohorts with the hope of maximizing power, but we also attempted to assess the impact of racial heterogeneity on our results by analysis only among the white cohorts. The results were not significantly different from when we included all cohorts in the meta-analysis, suggesting that racial heterogeneity across cohorts did not significantly impair our ability to discover genetic associations. However, the lack of a difference in findings may have been affected by the reduced sample size with the white cohort–only meta-analysis. Third, by using an exome genotyping array (as opposed to exome or whole-genome sequencing), we were unable to detect very low–frequency and private variants, which were explicitly excluded from the exome arrays. Fourth, our functional investigation of this locus is limited, and additional studies are needed to confirm the gene(s) involved, identify the mechanisms for increased susceptibility, and elucidate their role in the context of other genes related to COPD.

In summary, in a large study of coding variation in COPD, we identified a novel, nonsynonymous variant, rs181206, in *IL27* that is associated with COPD susceptibility. This variant is in LD with variants identified by GWASs in body mass index, type 1 diabetes, and inflammatory bowel disease. Our work and the work of others suggest that this locus is complex and that susceptibility may be due to an effect on *IL27* or other genes at this locus, particularly the mitochondria- and autophagy-related gene *TUFM*. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors acknowledge and thank Augustine M. K. Choi, M.D., for his support of this project.

COPDGene Investigators

Administrative Core: James Crapo, M.D. (principal investigator); Edwin Silverman, M.D., Ph.D. (principal investigator); Barry Make, M.D.; and Elizabeth Regan, M.D., Ph.D.

Genetic Analysis Core: Terri Beaty, Ph.D.; Nan Laird, Ph.D.; Christoph Lange, Ph.D.; Michael Cho, M.D.; Stephanie Santorico, Ph.D.; John Hokanson, M.P.H., Ph.D.; Dawn DeMeo, M.D., M.P.H.; Nadia Hansel, M.D.; M.P.H., Craig Hersh, M.D., M.P.H.; Peter Castaldi, M.D., M.Sc.; Merry-Lynn McDonald, Ph.D.; Emily Wan, M.D.; Megan Hardin, M.D.; Jacqueline Hetmanski, M.S.; Margaret Parker, Ph.D.; Marilyn Foreman, M.D.; Brian Hobbs, M.D., M.M.Sc.; Robert Busch, M.D.; Adel El-Bouiez, M.D.; Dandi Qiao, Ph.D.; Elizabeth Regan, M.D.; Eitan Halper-Stromberg; Ferdouse Begum; Sungho Won; and Sharon Lutz, Ph.D.

Imaging Core: David A. Lynch, M.B.; Harvey O. Coxson, Ph.D.; MeiLan K. Han, M.D., M.S.; Eric A. Hoffman, Ph.D.; Stephen Humphries M.S.; Francine L. Jacobson, M.D.; Philip F. Judy, Ph.D.; Ella A. Kazerooni, M.D.; John D. Newell, Jr., M.D.; Elizabeth Regan, M.D.; James C. Ross, Ph.D.; Raul San Jose Estepar, Ph.D.; Berend C. Stoel, Ph.D.; Juerg Tschirren, Ph.D.; Eva van Rikxoort, Ph.D.; Bram van Ginneken, Ph.D.; George Washko, M.D.; Carla G. Wilson, M.S.; Mustafa Al Qaisi, M.D.; Teresa Gray, Alex Kluiber, Tanya Mann, Jered Sieren, Douglas Stinson, Joyce Schroeder, M.D.; and Edwin Van Beek, M.D., Ph.D.

PFT QA Core, Salt Lake City, Utah: Robert Jensen, Ph.D.

Data Coordinating Center and Biostatistics, National Jewish Health, Denver, Colorado: Douglas Everett, Ph.D.; Anna Faino, M.S.; Matt Strand, Ph.D.; and Carla Wilson, M.S.

Epidemiology Core, University of Colorado Anschutz Medical Campus, Aurora, Colorado: John E. Hokanson, M.P.H., Ph.D.; Gregory Kinney, M.P.H., Ph.D.; Sharon Lutz, Ph.D.; Kendra Young Ph.D.; Katherine Pratte, MSPH; and Lindsey Duca, M.S.

VA Ann Arbor Healthcare System, Ann Arbor, Michigan: Jeffrey L. Curtis, M.D.; Carlos H. Martinez, M.D., M.P.H.; and Perry G. Pernicano, M.D.

Baylor College of Medicine, Houston, Texas: Nicola Hanania, M.D., MS; Philip Alapat, M.D.;

Venkata Bandi, M.D.; Mustafa Atik, M.D.; Aladin Boriek, Ph.D.; Kalpatha Guntupalli, M.D.; Elizabeth Guy, M.D.; Amit Parulekar, M.D.; and Arun Nachiappan, M.D.

Brigham and Women's Hospital, Boston, Massachusetts: Dawn DeMeo, M.D., M.P.H.; Craig Hersh, M.D., M.P.H.; George Washko, M.D.; and Francine Jacobson, M.D., M.P.H.

Columbia University, New York, New York: R. Graham Barr, M.D., DrPH, Byron Thomashow, M.D.; John Austin, M.D.; Belinda D'Souza, M.D.; Gregory D. N. Pearson, M.D.; and Anna Rozenshtein, M.D., M.P.H.

Duke University Medical Center, Durham, North Carolina: Neil MacIntyre, Jr., M.D.; Lacey Washington, M.D.; and H. Page McAdams, M.D.

Health Partners Research Foundation, Minneapolis, Minnesota: Charlene McEvoy, M.D., M.P.H.; and Joseph Tashjian, M.D.

The Johns Hopkins University, Baltimore, Maryland: Robert Wise, M.D.; Nadia Hansel, M.D., M.P.H.; Robert Brown, M.D.; Karen Horton, M.D.; and Nirupama Putcha, M.D., M.H.S.

Los Angeles Biomedical Research Institute at Harbor UCLA Medical Center, Torrance, California: Richard Casaburi, Ph.D., M.D.; Alessandra Adami, Ph.D.; Janos Porszasz, M.D., Ph.D.; Hans Fischer, M.D., Ph.D.; Matthew Budoff, M.D.; and Harry Rossiter, Ph.D.

Michael E. DeBakey VA Medical Center, Houston, Texas: Amir Sharafkhaneh, M.D., Ph.D.; and Charlie Lan, D.O.

Minneapolis VA Hospital, Minneapolis, Minnesota: Christine Wendt, M.D.; and Brian Bell, M.D.

Morehouse School of Medicine, Atlanta, Georgia: Marilyn Foreman, M.D., M.S.; Gloria Westney, M.D., M.S.; and Eugene Berkowitz, M.D., Ph.D.

National Jewish Health, Denver, Colorado: Russell Bowler, M.D., Ph.D.; and David Lynch, M.D.

Reliant Medical Group, Worcester, Massachusetts: Richard Rosiello, M.D.; and David Pace, M.D.

Temple University, Philadelphia, Pennsylvania: Gerard Criner, M.D.; David Ciccolella, M.D.; Francis Cordova, M.D.; Chandra Dass, M.D.; Gilbert D'Alonzo, D.O.; Parag Desai, M.D.; Michael Jacobs, PharmD.; Steven Kelsen, M.D., Ph.D.; Victor Kim, M.D.; A. James Mamary, M.D.; Nathaniel Marchetti, D.O.; Aditi Satti, M.D.; Kartik Shenoy, M.D.; Robert M. Steiner, M.D.; Alex Swift, M.D.; Irene Swift, M.D.; and Maria Elena Vega-Sanchez, M.D.

University of Alabama, Birmingham, Alabama: Mark Dransfield, M.D.; William Bailey, M.D.; J. Michael Wells, M.D.; Surya Bhatt, M.D.; and Hrudaya Nath, M.D.

University of California, San Diego, California: Joe Ramsdell, M.D.; Paul Friedman, M.D.; Xavier Soler, M.D., Ph.D.; and Andrew Yen, M.D.

University of Iowa, Iowa City, Iowa: Alejandro Comellas, M.D.; John Newell, Jr., M.D.; and Brad Thompson, M.D.

University of Michigan, Ann Arbor, Michigan: MeiLan Han, M.D.; Ella Kazerooni, M.D.; and Carlos Martinez, M.D.

University of Minnesota, Minneapolis, Minnesota: Joanne Billings, M.D.; and Tadashi Allen, M.D.

University of Pittsburgh, Pittsburgh, Pennsylvania: Frank Sciurba, M.D.; Divay Chandra, M.D., M.Sc.; Joel Weissfeld, M.D., M.P.H.; Carl Fuhrman, M.D.; and Jessica Bon, M.D.

University of Texas Health Science Center at San Antonio, San Antonio, Texas: Antonio Anzueto, M.D., Sandra Adams, M.D., Diego Maselli-Caceres, M.D.; and Mario E. Ruiz, M.D.

International COPD Genetics Network Investigators

Edwin K. Silverman, Brigham & Women's Hospital, Boston, Massachusetts; David A. Lomas, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom; Barry J. Make, National Jewish Medical and Research Center, Denver, Colorado; Alvar Agustí and Jaume Sauleda, Hospital Universitari Son Dureta, Fundación Caubet-Cimera and CIBER Enfermedades Respiratorias, Spain; Peter M. A. Calverley, University of Liverpool, Liverpool, United Kingdom; Claudio F. Donner, Division of Pulmonary Disease, S. Maugeri Foundation, Veruno (NO), Italy; Robert D. Levy, University of British Columbia, Vancouver, British Columbia, Canada; Peter D. Paré, University of British Columbia, Vancouver, British Columbia, Canada; Stephen Rennard, Section of Pulmonary & Critical Care, University of Nebraska Medical Center, Omaha, Nebraska; Jørgen Vestbo, Department of Cardiology and Respiratory Medicine, Hvidovre Hospital, Copenhagen, Denmark; and Emiel F.M. Wouters, University Hospital Maastricht, Maastricht, the Netherlands.

Additional acknowledgments can be found in the online supplement.

References

1. Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, Barnes PJ, Fabbri LM, Martinez FJ, Nishimura M, *et al*. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013;187:347–365.
2. Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, Need AC, Feng S, Hersh CP, Bakke P, Gulsvik A, *et al*.; ICGN Investigators. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet* 2009;5:e1000421.
3. Wilk JB, Chen TH, Gottlieb DJ, Walter RE, Nagle MW, Brandler BJ, Myers RH, Borecki IB, Silverman EK, Weiss ST, *et al*. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. *PLoS Genet* 2009;5:e1000429.
4. Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Hersh CP, DeMeo DL, Hunninghake GM, Litonjua AA, Sparrow D, *et al*. Variants in *FAM13A* are associated with chronic obstructive pulmonary disease. *Nat Genet* 2010;42:200–202.
5. Cho MH, Castaldi PJ, Wan ES, Siedlinski M, Hersh CP, Demeo DL, Himes BE, Sylvia JS, Klanderman BJ, Ziniti JP, *et al*.; ICGN Investigators; ECLIPSE Investigators; COPDGene Investigators. A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. *Hum Mol Genet* 2012;21:947–957.
6. Cho MH, McDonald ML, Zhou X, Mattheisen M, Castaldi PJ, Hersh CP, Demeo DL, Sylvia JS, Ziniti J, Laird NM, *et al*.; NETT Genetics, ICGN, ECLIPSE and COPDGene Investigators. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med* 2014;2:214–225.
7. Zhou JJ, Cho MH, Castaldi PJ, Hersh CP, Silverman EK, Laird NM. Heritability of chronic obstructive pulmonary disease and related phenotypes in smokers. *Am J Respir Crit Care Med* 2013;188: 941–947.
8. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009;106:9362–9367.
9. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, *et al*. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–753.
10. Silverman EK, Sandhaus RA. Alpha$_1$-antitrypsin deficiency. *N Engl J Med* 2009;360:2749–2757.
11. Abecasis Lab at the University of Michigan Center for Statistical Genetics. Exome chip design [updated 2013 Aug 6; accessed 2016 Jan 27]. Available from: http://genome.sph.umich.edu/wiki/Exome_Chip_Design
12. Huyghe JR, Jackson AU, Fogarty MP, Buchkovich ML, Stančáková A, Stringham HM, Sim X, Yang L, Fuchsberger C, Cederberg H, *et al*. Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. *Nat Genet* 2013;45: 197–201.
13. Flannick J, Thorleifsson G, Beer NL, Jacobs SB, Grarup N, Burtt NP, Mahajan A, Fuchsberger C, Atzmon G, Benediktsson R, *et al*.; Go-T2D Consortium; T2D-GENES Consortium. Loss-of-function

mutations in *SLC30A8* protect against type 2 diabetes. *Nat Genet* 2014;46:357–363.
14. Peloso GM, Auer PL, Bis JC, Voorman A, Morrison AC, Stitziel NO, Brody JA, Khetarpal SA, Crosby JR, Fornage M, *et al*.; NHLBI GO Exome Sequencing Project. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am J Hum Genet* 2014;94: 223–232.
15. Wessel J, Chu AY, Willems SM, Wang S, Yaghootkar H, Brody JA, Dauriz M, Hivert MF, Raghavan S, Lipovich L, *et al*.; EPIC-InterAct Consortium. Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. *Nat Commun* 2015;6:5897.
16. Hobbs BD, Hardin M, Hawrylkiewicz I, Sliwinski P, Yim JJ, Kim WJ, Kim DK, Agusti A, Make BJ, Calverley PM, *et al*. Coding variant associations with lung function and COPD using an exome array [abstract]. *Am J Respir Crit Care Med* 2015;191:A2912.
17. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015;4:7.
18. Purcell S, Chang C. PLINK version 1.9 [accessed 2016 Jan 27]. Available from: https://www.cog-genomics.org/plink2
19. Chen MH, Yang Q. GWAF: genome-wide association/interaction analysis and rare variant analysis with family data. R package version 2.2 [accessed 2016 Jan 27]. Available from: http://CRAN.R-project.org/package=GWAF
20. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing [accessed 2016 Jan 27]. Available from: http://www.R-project.org/
21. Abecasis G, Li Y, Willer C; Abecasis Lab at the University of Michigan Center for Statistical Genetics. METAL. Version release 2011-03-25 [accessed 2016 Jan 27]. Available from: http://genome.sph.umich.edu/wiki/METAL_Program
22. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26:2190–2191.
23. Lee S, Miropolsky L, Wu M. SKAT: SNP-set (Sequence) Kernel Association Test. R package version 1.0.1 [accessed 2016 Jan 27]. Available from: http://CRAN.R-project.org/package=SKAT
24. Chen H, Wang C, Conomos MP, Stilp AM, Li Z, Sofer T, Szpiro AA, Chen W, Brehm JM, Celedon JC, *et al*. Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. *Am J Hum Genet* 2016;98: 653–666.
25. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet* 2011;89:82–93.
26. Lee S. MetaSKAT: Meta-Analysis for SNP-set (Sequence) Kernel Association Test. R package version 0.40 [accessed 2016 Jan 27]. Available from: http://CRAN.R-project.org/package=MetaSKAT
27. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, Christiani DC, Wurfel MM, Lin X; NHLBI GO Exome Sequencing Project—ESP Lung Project Team. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet* 2012;91: 224–237.
28. Soler Artigas M, Loth DW, Wain LV, Gharib SA, Obeidat M, Tang W, Zhai G, Zhao JH, Smith AV, Huffman JE, *et al*.; International Lung Cancer Consortium; GIANT consortium. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat Genet* 2011;43:1082–1090.
29. Hersh CP, DeMeo DL, Silverman EK. Chronic obstructive pulmonary disease. In: Silverman EK, Shapiro SD, Lomas DA, Weiss ST, editors. Respiratory genetics. Boca Raton, FL: CRC Press; 2005. pp. 253–296.
30. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–249.
31. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–315.
32. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, Powell C, Vedantam S, Buchkovich ML, Yang J, *et al*.; LifeLines Cohort Study; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; 518:197–206.
33. Imielinski M, Baldassano RN, Griffiths A, Russell RK, Annese V, Dubinsky M, Kugathasan S, Bradfield JP, Walters TD, Sleiman P, *et al*.; Western Regional Alliance for Pediatric IBD; International IBD Genetics Consortium; NIDDK IBD Genetics Consortium; Belgian-French IBD Consortium; Wellcome Trust Case Control Consortium. Common variants at five new loci associated with early-onset inflammatory bowel disease. *Nat Genet* 2009;41: 1335–1340.
34. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R, *et al*. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010; 42:1118–1125.
35. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, *et al*.; Type 1 Diabetes Genetics Consortium. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 2009;41:703–707.
36. Plagnol V, Howson JM, Smyth DJ, Walker N, Hafler JP, Wallace C, Stevens H, Jackson L, Simmonds MJ, Bingley PJ, *et al*.; Type 1 Diabetes Genetics Consortium. Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases. *PLoS Genet* 2011;7:e1002216.
37. Morrow J, Qiu W, DeMeo DL, Houston I, Pinto-Plata VM, Celli BR, Marchetti N, Criner GJ, Bueno R, Washko GR, *et al*. Network analysis of gene expression in severe COPD lung tissue samples. *Am J Respir Crit Care Med* 2015;191:A1253.
38. Park SJ, Kleffmann T, Hessian PA. The G82S polymorphism promotes glycosylation of the receptor for advanced glycation end products (RAGE) at asparagine 81: comparison of wild-type rage with the G82S polymorphic variant. *J Biol Chem* 2011;286: 21384–21392.
39. Xie J, Reverdatto S, Frolov A, Hoffmann R, Burz DS, Shekhtman A. Structural basis for pattern recognition by the receptor for advanced glycation end products (RAGE). *J Biol Chem* 2008;283: 27255–27269.
40. Spierings E. Minor histocompatibility antigens: past, present, and future. *Tissue Antigens* 2014;84:374–360.
41. The Broad Institute of MIT and Harvard. GTEx Portal. Version: 4. Build 200 [accessed 2016 Jan 27]. Available from: http://www.gtexportal.org/home/
42. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580–585.
43. The human protein atlas [accessed 2016 Jan 27]. Available from: http://www.proteinatlas.org/
44. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, *et al*. Tissue-based map of the human proteome. *Science* 2015;347: 1260419.
45. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010;42:441–447.
46. Castaldi PJ, Cho MH, Zhou X, Qiu W, Mcgeachie M, Celli B, Bakke P, Gulsvik A, Lomas DA, Crapo JD, *et al*. Genetic control of gene expression at novel and established chronic obstructive pulmonary disease loci. *Hum Mol Genet* 2015;24:1200–1210.
47. Saccone NL, Culverhouse RC, Schwantes-An TH, Cannon DS, Chen X, Cichon S, Giegling I, Han S, Han Y, Keskitalo-Vuokko K, *et al*. Multiple independent loci at chromosome 15q25.1 affect smoking quantity: a meta-analysis and comparison with lung cancer and COPD. *PLoS Genet* 2010;6: e1001053.

48. Siedlinski M, Tingley D, Lipman PJ, Cho MH, Litonjua AA, Sparrow D, Bakke P, Gulsvik A, Lomas DA, Anderson W, *et al*.; COPDGene and ECLIPSE Investigators. Dissecting direct and indirect genetic effects on chronic obstructive pulmonary disease (COPD) susceptibility. *Hum Genet* 2013;132:431–441.
49. Wilk JB, Shrine NR, Loehr LR, Zhao JH, Manichaikul A, Lopez LM, Smith AV, Heckbert SR, Smolonska J, Tang W, *et al*. Genome-wide association studies identify *CHRNA5/3* and *HTR4* in the development of airflow obstruction. *Am J Respir Crit Care Med* 2012; 186:622–632.
50. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, Zhao JH, Ramasamy A, Zhai G, Vitart V, *et al*.; Wellcome Trust Case Control Consortium; NSHD Respiratory Study Team. Genome-wide association study identifies five loci associated with lung function. *Nat Genet* 2010;42:36–44.
51. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, Franceschini N, van Durme YM, Chen TH, Barr RG, *et al*. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet* 2010; 42:45–52.
52. Burkart KM, Manichaikul A, Wilk JB, Ahmed FS, Burke GL, Enright P, Hansel NN, Haynes D, Heckbert SR, Hoffman EA, *et al*. *APOM* and high-density lipoprotein cholesterol are associated with lung function and per cent emphysema. *Eur Respir J* 2014;43:1003–1017.
53. Cho MH, Castaldi PJ, Hersh CP, Hobbs BD, Barr RG, Tal-Singer R, Bakke P, Gulsvik A, San José Estépar R, Van Beek EJ, *et al*.; NETT Genetics, ECLIPSE, and COPDGene Investigators. A genome-wide association study of emphysema and airway quantitative imaging phenotypes. *Am J Respir Crit Care Med* 2015;192:559–569.
54. National Center for Biotechnology Information. dbSNP short genetic variations (dbSNP Build ID: 144). Bethesda, MD: National Library of Medicine [accessed 2016 Jan 27]. Available from: http://www.ncbi.nlm.nih.gov/projects/SNP/
55. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* 2013;14: 483–495.
56. Núñez B, Sauleda J, Antó JM, Julià MR, Orozco M, Monsó E, Noguera A, Gómez FP, Garcia-Aymerich J, Agustí A; PAC-COPD Investigators. Anti-tissue antibodies are related to lung function in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2011;183:1025–1031.
57. Cosio MG, Saetta M, Agustí A. Immunologic aspects of chronic obstructive pulmonary disease. *N Engl J Med* 2009;360: 2445–2454.
58. Hancock DB, Artigas MS, Gharib SA, Henry A, Manichaikul A, Ramasamy A, Loth DW, Imboden M, Koch B, McArdle WL, *et al*. Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. *PLoS Genet* 2012;8:e1003098.
59. Wang K, Baldassano R, Zhang H, Qu HQ, Imielinski M, Kugathasan S, Annese V, Dubinsky M, Rotter JI, Russell RK, *et al*. Comparative genetic analysis of inflammatory bowel disease and type 1 diabetes implicates multiple loci with opposite effects. *Hum Mol Genet* 2010; 19:2059–2067.
60. Jess T, Loftus EV Jr, Harmsen WS, Zinsmeister AR, Tremaine WJ, Melton LJ III, Munkholm P, Sandborn WJ. Survival and cause specific mortality in patients with inflammatory bowel disease: a long term outcome study in Olmsted County, Minnesota, 1940–2004. *Gut* 2006;55:1248–1254.
61. Ekbom A, Brandt L, Granath F, Löfdahl CG, Egesten A. Increased risk of both ulcerative colitis and Crohn's disease in a population suffering from COPD. *Lung* 2008;186:167–172.
62. Yoshida H, Hunter CA. The immunobiology of interleukin-27. *Annu Rev Immunol* 2015;33:417–443.
63. Huang N, Liu L, Wang XZ, Liu D, Yin SY, Yang XD. Association of interleukin (*IL*)-*12* and *IL-27* gene polymorphisms with chronic obstructive pulmonary disease in a Chinese population. *DNA Cell Biol* 2008;27:527–531.
64. Cao J, Zhang L, Li D, Xu F, Huang S, Xiang Y, Yin Y, Ren G. IL-27 is elevated in patients with COPD and patients with pulmonary TB and induces human bronchial epithelial cells to produce CXCL10. *Chest* 2012;141:121–130.
65. Angata T, Ishii T, Gao C, Ohtsubo K, Kitazume S, Gemma A, Kida K, Taniguchi N. Association of serum interleukin-27 with the exacerbation of chronic obstructive pulmonary disease. *Physiol Rep* 2014;2:e12069.
66. McDonald ML, Mattheisen M, Cho MH, Liu YY, Harshfield B, Hersh CP, Bakke P, Gulsvik A, Lange C, Beaty TH, *et al*.; GenKOLS, COPDGene and ECLIPSE study investigators. Beyond GWAS in COPD: probing the landscape between gene-set associations, genome-wide associations and protein-protein interaction networks. *Hum Hered* 2014;78:131–139.
67. González JR, Cáceres A, Esko T, Cuscó I, Puig M, Esnaola M, Reina J, Siroux V, Bouzigon E, Nadif R, *et al*. A common 16p11.2 inversion underlies the joint susceptibility to asthma and obesity. *Am J Hum Genet* 2014;94:361–372.
68. Lei Y, Wen H, Yu Y, Taxman DJ, Zhang L, Widman DG, Swanson KV, Wen KW, Damania B, Moore CB, *et al*. The mitochondrial proteins NLRX1 and TUFM form a complex that regulates type I interferon and autophagy. *Immunity* 2012;36:933–946.
69. Lei Y, Wen H, Ting JP. The NLR protein, NLRX1, and its partner, TUFM, reduce type I interferon, and enhance autophagy. *Autophagy* 2013;9: 432–433.
70. Kang MJ, Yoon CM, Kim BH, Lee CM, Zhou Y, Sauler M, Homer R, Dhamija A, Boffa D, West AP, *et al*. Suppression of NLRX1 in chronic obstructive pulmonary disease. *J Clin Invest* 2015;125:2458–2462.
71. Han MK, Criner GJ. Update in chronic obstructive pulmonary disease 2012. *Am J Respir Crit Care Med* 2013;188:29–34.
72. Rennard SI, Vestbo J. The many "small COPDs": COPD should be an orphan disease. *Chest* 2008;134:623–627.
73. Hersh CP, Make BJ, Lynch DA, Barr RG, Bowler RP, Calverley PM, Castaldi PJ, Cho MH, Coxson HO, DeMeo DL, *et al*.; COPDGene and ECLIPSE Investigators. Non-emphysematous chronic obstructive pulmonary disease is associated with diabetes mellitus. *BMC Pulm Med* 2014;14:164.