Sex-Based Genetic Association Study Identifies CELSR1 as a Possible Chronic Obstructive Pulmonary Disease Risk Locus among Women

Permalink
https://escholarship.org/uc/item/5mw2151m

Journal
American Journal of Respiratory Cell and Molecular Biology, 56(3)

ISSN
1044-1549

Authors
Hardin, Megan
Cho, Michael H
Sharma, Sunita
et al.

Publication Date
2017-03-01

DOI
10.1165/rcmb.2016-0172oc

Peer reviewed
Sex-Based Genetic Association Study Identifies CELSR1 as a Possible Chronic Obstructive Pulmonary Disease Risk Locus among Women

Megan Hardin1,2, Michael H. Cho1,2, Sunita Sharma3, Kimberly Glass1, Peter J. Castaldi1, Merry-Lynn McDonald1, Hugues Aschard4, Jody Senter-Sylvia1, Kelan Tantisira1,2, Scott T. Weiss1,2, Craig P. Hersh1,2, Jarrett D. Morrow1,2, David Lomas5, Alvar Agusti6, Per Bakke7, Amund Gulsvik8, George T. O'Connor9, Josée Dupuis10,11, John Hokanson12, James D. Crapo13, Terri H. Beaty14, Nan Laird4, Edwin K. Silverman1,2, Dawn L. DeMeo1,2, and the COPDGene and Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-Points Investigators

1Channing Division of Network Medicine and 2Division of Pulmonary and Critical Care Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts; 3Division of Pulmonary Sciences and Critical Care Medicine, Department of Medicine, University of Colorado School of Medicine, Aurora, Colorado; 4Harvard School of Public Health, Boston, Massachusetts; 5Department of Medicine, University College London, London, United Kingdom; 6Thoracic Institute, Hospital Clinic, Barcelona, Spain; 7Department of Clinical Science, University of Bergen, Bergen, Norway; 8Department of Geriatric Medicine Ullevaal, Institute of Clinical Medicine, Oslo University Hospital University of Oslo, Oslo, Norway; 9Boston University School of Medicine, Boston, Massachusetts; 10Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts; 11National Heart, Lung, and Blood Institute’s Framingham Heart Study, Framingham, Massachusetts; 12Department of Epidemiology, Colorado School of Public Health, University of Colorado Denver, Denver, Colorado; 13Division of Pulmonary Sciences and Critical Care Medicine, National Jewish Health, Denver, Colorado; and 14Johns Hopkins School of Public Health, Baltimore, Maryland

Abstract

Chronic obstructive pulmonary disease (COPD) is a complex disease with strong environmental and genetic influences and sexually dimorphic features. Although genetic risk factors for COPD have been identified, much of the heritability remains unexplained. Sex-based genetic association studies may uncover additional COPD genetic risk factors. We studied current and former smokers from COPD case–control cohorts (COPDGene non-Hispanic whites and African Americans, Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-Points, and Genetics of Chronic Obstructive Lung Disease). COPD was defined as post-bronchodilator forced expiratory volume in 1 second/forced vital capacity less than 0.70 and forced expiratory volume in 1 second percent predicted less than 80. Testing was performed across all cohorts and combined in a meta-analysis adjusted for age, pack-years, and genetic ancestry. We first performed genome-wide single-nucleotide polymorphism (SNP)-by-sex interaction testing on the outcome of COPD affection status.

Keywords: chronic obstructive pulmonary disease; genetics; growth and development; sex; genome-wide association study

We performed sex-stratified association testing for SNPs with interaction \( P \) less than 10\(^{-6}\). We examined over 8 million SNPs in four populations, including 6,260 subjects with COPD (40.6% female) and 5,269 smoking control subjects (47.3% female). The SNP rs9615358 in the cadherin gene CELSR1 approached genome-wide significance for an interaction with sex \((P = 1.24 \times 10^{-7})\). In the sex-stratified meta-analysis, this SNP was associated with COPD among females (odds ratio, 1.37 [95% confidence interval, 1.25–1.49]; \( P = 3.32 \times 10^{-7} \)) but not males (odds ratio, 0.90 [95% confidence interval, 0.79–1.01]; \( P = 0.06 \)). CELSR1 is involved in fetal lung development. In a human fetal lung tissue dataset, we observed greater CELSR1 expression in female compared with male samples. This SNP-by-sex genome-wide association analysis identified the fetal lung development gene, CELSR1, as a potential sex-specific risk factor for COPD. Identifying sex-specific genetic risk factors may reveal new insights into sexually dimorphic features of COPD.

Acknowledgments: This work was supported by National Institutes of Health/National Heart, Lung, and Blood Institute (NHLBI) grants K12 HL120004 (M.H.), R01HL113264 (M.H.C.), R01 HL089856 (E.K.S.), R01 HL089897 (E.K.S. and J.D.C.), P01HL105339 (M.H.C. and E.K.S.), HL089438 (D.L.D.), and NHLBI contract HHSN268201500001I (G.T.O'C. and J.D.). This project was also supported by the 2013 Sheila J. Goodnight, M.D., F.C.C.P., Clinical Research Grant in Women’s Lung Health (M.H.); the COPDGene project (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. The content of this article is solely the responsibility of the authors, and does not necessarily represent the official views of the NHLBI or the National Institutes of Health.
Clinical Relevance

There is accumulating evidence that the susceptibility to and severity of chronic obstructive pulmonary disease (COPD) is sexually dimorphic, but the biologic drivers of COPD-related dimorphism are poorly understood. Identifying sex-specific genetic associations for COPD may highlight new genes and pathways to consider for sex-specific diagnostic and therapeutic approaches to COPD.

Chronic obstructive pulmonary disease (COPD) is a complex disease characterized by fixed airflow obstruction that results from a combination of genetic risk factors and environmental exposures. COPD was the third leading cause of death in the US in 2011 (1). Although traditionally considered a disease of men, the prevalence of COPD is increasing among female smokers (2). Furthermore, there has been a growing recognition that men and women may have differing susceptibility and presentation of disease. On average, compared with males, females with COPD are more likely to report dyspnea and cough (3), and have less radiographic emphysema, but more airway remodeling (4). We have recently demonstrated that, in the COPDGene population, on average, women with COPD have less radiographic emphysema than men for the same degree of lung function, but, among subgroups with severe or early-onset COPD, women have equivalent emphysema despite less pack-years smoking (5). Women are more prevalent in populations with severe, early-onset COPD (6, 7), and women who smoke demonstrate greater lung function decline than men (8).

Cigarette smoking is the strongest environmental risk factor for COPD; however, the variability in clinical response to smoke exposure and the familial aggregation of airflow obstruction support a genetic component as well. Although several genetic risk factors for COPD have been identified through genome-wide analysis (9), to our knowledge, no study has investigated sex-specific associations for COPD. Our aim was to identify sexually dimorphic genetic risk factors for COPD, with the goal of revealing new insights into COPD susceptibility and severity.

Genotype-by-sex interaction is likely a major component of many human traits and diseases (10, 11), including waist:height ratio (12), type II diabetes (13), and asthma (14). The genetic composition of the autosomes is shared between men and women, and, therefore, these differences likely result from variation at the level of the regulatory genome (10).

There is evidence to suggest that, similar to the complex traits mentioned previously here, COPD may also have sexually dimorphic genetic features. Lung function (the ratio of forced expiratory volume in 1 second [FEV1] to forced vital capacity [FVC]) demonstrated sexually dimorphic heritability estimates in a population of nonsmokers (15). There is a female predominance of early-onset COPD, and female, smoking, first-degree relatives of subjects with early-onset COPD demonstrate greater lung function impairment compared with male first-degree relatives (6). In a cohort of subjects with alpha-1 antitrypsin deficiency, variants in the COPD risk gene CHRNA3/IREB2 demonstrated a significant single-nucleotide polymorphism (SNP)-by-sex interaction for lung function outcomes, with a significant association among males, but not females (16). Collectively, these sex-based differences in COPD susceptibility and presentation support a role for genetic influences on sexually dimorphic features.

We hypothesized that genome-wide association testing that specifically accounts for sex-specific genetic effects would identify new variants associated with COPD, as well as demonstrate sexually dimorphic association in known COPD risk alleles. We tested this hypothesis using four large case–control populations of COPD to detect a SNP-by-sex interaction for the outcome, COPD affection status.

Materials and Methods

Study Subjects

Study subjects were current or former smokers with at least 10 pack-years smoking participating in previously described case–control COPD studies: COPDGene non-Hispanic white and African American (17), ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-Points) (18), and Genes of Chronic Obstructive Lung Disease (GenKOLS) subjects (19). All subjects completed standardized spirometry (20). Cases had moderate to severe airflow limitation (Global Initiative for Chronic Obstructive Lung Disease spirometry grades II–IV) with post-bronchodilator FEV1/FVC less than 0.70 and FEV1 less than 80% predicted. Control subjects were current or former smokers with normal lung function. Subjects with known alpha-1 antitrypsin deficiency were excluded. COPDGene subjects included non-Hispanic white or African American subjects, and these subgroups were analyzed independently of each other. ECLIPSE and GenKOLS subjects were white of European descent. All subjects signed informed consent before participation in any study activity, and Institutional Review Board approval was obtained from all study sites (see Table E5 in the online supplement; COPDGene, 2007P000554; ECLIPSE, 2005P002467; GenKOLS, 2009P001700).
Genotyping

Genotyping using Illumina genome-wide SNP platforms, quality control, and imputation have been previously described for all populations (19, 21, 22). Imputation was performed using 1,000 genomes phase I v3 European (EUR) for non-Hispanic white ancestry, and cosmopolitan reference panels for African Americans.

Statistical Analysis

We first modeled the interaction between SNP and sex for COPD case–control status. We performed a separate logistic regression analysis in each population using probABEL v0.4.5 (23), testing for a significant SNP-by-sex interaction adjusting for age, sex, pack-years smoking and principal components. We combined these results in a meta-analysis using METAL (metal-2010-02-08-patch2). We then performed sex-stratified genome-wide association study (GWAS) for SNPs that demonstrated SNP-by-sex interactions. We chose an inclusion threshold of P less than 10^{-8} for carrying forward interaction SNPs, as standard significance thresholds may miss SNPs of interest.

We performed sex-stratified GWAS in all cohorts, adjusting for age, pack-years, and principal components using PLINK v1.9 (https://www.cog-genomics.org/plink2) (24), and combined these results in stratified meta-analyses using METAL (version 2010-08-01) (25).

We examined variants previously identified through GWAS as COPD risk factors for a SNP-by-sex interaction (CHRNA3, rs4416442; FAM13A, rs12914385; HHIP, rs13141641; RIN3, rs754388; MMP3/12, rs626750; TGFBR2, rs4846480) (9), as well as sex-specific association with COPD affection status. We created regional association plots using LocusZoom (v1.1) (26) and measured linkage disequilibrium using SNAP (27).

Human Fetal Lung Gene Expression Profiling

Postconception age- and sex-matched human fetal lung tissue samples from 366 samples were acquired through the tissue retrieval program sponsored by the National Institute of Child Health and Development (Rockville, MD), the University of Maryland Brain and Tissue Bank for Developmental Disorders (Baltimore, MD), and the Center for Birth Defects Research (University of Washington, Seattle, WA), as previously described (28). Total mRNA was isolated using the Illustra RNAspin mini RNA isolation kit (GE Healthcare, Piscataway, NJ), and genome-wide gene expression profiles were measured using HumanRef8 v2 BeadChips (Illumina, San Diego, CA). Placental cotinine was measured, and only fetal lung tissues without intrauterine smoke exposure were included (29).

Sex-Specific Differences in CELSR1 and FAR2 Expression during Human Lung Development

We assessed sex-specific differential expression of CELSR1 and FAR2 genes using linear models with adjustment for postconception age.

Expression Quantitative Trait Loci Analysis

To identify a functional role for top SNPs from the SNP-by-sex interaction, we tested our top SNPs as expression quantitative trait loci (eQTL) using a previously reported analysis identifying whole-blood eQTLs from 121 subjects (40 female, 81 male) with COPD (30).

Sex-Specific Differences in CELSR1 Expression in Adult Human Lung Tissue

We performed a sex-specific gene expression analysis for CELSR1 using data from adult human lung tissue samples from 111 subjects with COPD and 40 control smokers with normal lung function (extended methods available in the online supplement). We assessed sex-specific differential expression of the CELSR1 gene among subjects with COPD and control subjects using linear models with adjustment for age, race, and pack-years of smoking.

Sex-Specific Association of CELSR1 in Lifetime Never-Smokers

We also investigated the association of CELSR1 with obstructive lung disease in 2,229 lifetime never-smokers from the Framingham Heart Study population.

The Framingham Heart Study includes spirometry measures of three generations of family members. For the replication analysis of CELSR1 associations, offspring cohort members and the third generation were included. A total of 2,229 individuals were lifetime never-smokers, had genotyping and spirometry data, and met the definition of COPD (n = 100, 44 females, FEV1/FVC < 0.7 and FEV1 < 70% predicted) or control (n = 2,129, 1,156 females, FEV1/FVC ratio ≥ 0.7 and FEV1 > 80% predicted). Framingham participants were genotyped using the Gene Chip Human Mapping 500K Array Set (Affymetrix, Santa Clara, CA). All four SNPs selected for replication (rs9615358, rs9615981, rs9615982, and rs7286446) in the Framingham Heart Study nonsmokers were imputed using the HaploType Reference Consortium panel. Imputation quality was good (r² = 0.92 for all four SNPs). Logistic regression models were fitted using generalized estimating equations to account for familial correlation, with COPD affection status as the outcome, and included age, sex, SNP, and a SNP-by-sex interaction term.

Results

Demographics

There were 6,260 cases with COPD (40.6% female) and 5,269 smoking control subjects (47.3% female) included in this GWAS analysis (Table 1). Males tended to have greater pack-years smoking history. Lung function was similar among male and female cases, except for lower lung function in males from ECLIPSE. Lung function was similar between male and female control subjects (Table 1).

SNP-by-Sex Interaction

Over 4.7 million genotyped and imputed SNPs across all four cohorts were included in the interaction GWAS and meta-analysis (as shown in Figure E1 in the online supplement). We identified several variants that approached the prespecified significance threshold of 5 × 10^{-8} for an interaction with sex (Table 2). The SNP, rs9615358, in the gene CELSR1 had the lowest P value (P = 1.24 × 10^{-7}) for the SNP-by-sex interaction.

The SNP, rs7294481, located upstream from the gene FAR2 approached significance for the SNP-by-sex interaction (P = 3.00 × 10^{-7}). A SNP near the gene, C9orf123 approached significance for an interaction (P = 2.95 × 10^{-7}); however, the direction of effect was not the same in all studies, and this SNP was not followed up.
In a combined interaction model evaluating individual SNP-by-sex interactions, there was no evidence for association with COPD or SNP-by-sex interaction (P = 0.33) for CELSR1 in lifetime never-smokers in the Framingham Heart Study data.

**Sex-Stratified Analysis of Top Interaction SNPs**

We performed genetic association analysis stratified by sex for the top five interaction SNPs. The top SNP, rs9615358, in the gene, CELSR1, approached genome-wide significance for COPD affection status in females (odds ratio [OR], 1.37; 95% confidence interval [CI], 1.25–1.49; P = 3.32 × 10⁻⁴), but not in males (OR, 0.90; 95% CI, 0.79–1.01; P = 0.06) (Figure 1, Table 3). This genotyped SNP is located within an intronic region of the gene CELSR1. The SNP, rs9615358, demonstrated evidence as a blood eQTL for CELSR1 (30) (P = 0.0028). This SNP is in linkage disequilibrium with the SNP, rs6008795 (r² = 0.76), which is predicted to cause a leucine-to-proline missense mutation (31). rs6008795 was nominally associated with COPD in the meta-analysis (females, P = 3.3 × 10⁻⁴; males, P = 0.01). The SNP, rs7294481, upstream from FAR2, was associated with a protective effect among male subjects (OR, 0.83 [95% CI, 0.74–0.92]; P = 2.49 × 10⁻³), and an effect in the opposite direction among females (OR, 1.15 [95% CI, 1.06–1.25]; P = 0.003) (Figure 2, Table 3).

**Candidate Gene Analysis**

We examined top SNPs from several genes/regions that have recently been identified as associated with COPD affection status (RIN3, FAM13A, HHIP, CHRNA3/IREB2) and severe COPD (MMP3/12, TGFβ2) (9). The SNP, rs754388, in the gene, RIN3 locus demonstrated a nominally significant SNP-by-sex interaction (P = 0.03) with a smaller P value and greater OR in females (OR, 1.41 [95% CI, 1.28–1.54]; P = 1.84 × 10⁻⁷) compared with males (OR, 1.17 [95% CI, 1.05–1.28]; P = 0.01) (Table 4). The FAM13A, HHIP, CHRNA3/IREB2, MMP3/12, and TGFβ2 SNPs did not demonstrate significant SNP-by-sex interactions for the COPD case–control outcome.

**Sex-Stratified GWAS**

We additionally performed a sex-stratified GWAS testing all SNPs in all cohorts, and
then combined the results within each sex via meta-analysis (Tables E1 and E2). Among the female samples, the top sex-stratified GWAS results included several variants in the *FAM13A* locus that were associated with risk of COPD at the genome-wide level (rs2869966: OR, 1.35 [95% CI, 1.26–1.45]; P = 6.63 × 10⁻⁸). The P value for the association between these variants and COPD among the male samples was not as low (rs2869966: OR, 1.19 [95% CI, 1.10–1.28]; P = 8.12 × 10⁻⁵). The P value for the SNP-by-sex interaction for the top SNP, rs2869966, was 0.04. The top SNP among the males was rs5607733 in the gene, *CHRNA3* (OR, 1.35 [95% CI, 1.26–1.45]; P = 1.61 × 10⁻⁹). The P value for this SNP was not as low among female subjects (OR, 1.22 [95% CI, 1.12–1.33]; P = 1.48 × 10⁻⁴).

The SNP-by-sex interaction was not significant at a P value of 0.12.

### Tissue Expression Data
Sexually dimorphic features in the lung may begin during lung development, and the gene, *CELSR1*, is involved in fetal lung maturity. As such, we examined the sex-specific expression of our top associated genes in a set of fetal lung tissue samples. In a linear regression model comparing tissue log expression levels and adjusting for age, the gene, *CELSR1*, was differentially expressed by sex, with significantly higher levels of expression in fetal lung tissue from females compared with male fetal lung tissue from males (female log expression, 8.58; male log expression, 8.45; P = 0.0007; Figure E2). *FAR2* additionally demonstrated greater expression in female fetal lung tissue compared with male fetal lung tissue (female log expression, 6.97; male log expression, 6.89; P = 0.03; Figure E3).

### Discussion
This is the first GWAS, to our knowledge, to examine sexually dimorphic genetic risk factors for COPD. Our main findings show that the *CELSR1* gene is a candidate for COPD in female smokers. Variants in this gene approached genome-wide significance for a SNP-by-sex interaction for the outcome, COPD affection status. In a sex-stratified analysis, these SNPs were associated with COPD affection status among female subjects, but not among male subjects. *CELSR1* is a known early lung development gene; our top SNP appears to play a role in controlling *CELSR1* gene expression, and *CELSR1* gene expression was higher in female lung tissue from pseudoglandular and cannillaric phases of lung development compared with male lung tissue. Collectively, these observations suggest that *CELSR1* may play a role in sexually dimorphic COPD susceptibility, and that this susceptibility may have developmental origins.

Several challenges exist to identifying sex-specific risk factors for complex traits. Early sex-specific association studies were

### Table 2. Single-Nucleotide Polymorphisms with P less than 10⁻⁶ for the Single-Nucleotide Polymorphism-by-Sex Interaction on COPD Case–Control Outcome

<table>
<thead>
<tr>
<th>Chr</th>
<th>Gene</th>
<th>SNP</th>
<th>Reference Allele</th>
<th>MAF</th>
<th>Interaction P Value</th>
<th>Main Effect P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>CELSR1</td>
<td>rs9615358</td>
<td>A</td>
<td>0.14</td>
<td>1.24 × 10⁻⁷</td>
<td>0.04</td>
</tr>
<tr>
<td>9</td>
<td>C9orf123</td>
<td>rs7034632</td>
<td>T</td>
<td>0.33</td>
<td>2.95 × 10⁻⁷</td>
<td>0.08</td>
</tr>
<tr>
<td>12</td>
<td>FAR2</td>
<td>rs7294481</td>
<td>T</td>
<td>0.42</td>
<td>3.00 × 10⁻⁷</td>
<td>0.27</td>
</tr>
<tr>
<td>22</td>
<td>CELSR1</td>
<td>rs9615982</td>
<td>T</td>
<td>0.14</td>
<td>4.00 × 10⁻⁷</td>
<td>0.27</td>
</tr>
<tr>
<td>22</td>
<td>CELSR1</td>
<td>rs7286446</td>
<td>T</td>
<td>0.14</td>
<td>5.49 × 10⁻⁷</td>
<td>0.22</td>
</tr>
<tr>
<td>22</td>
<td>CELSR1</td>
<td>rs9615981</td>
<td>T</td>
<td>0.14</td>
<td>5.51 × 10⁻⁷</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Definition of abbreviations: Chr, chromosome; COPD, chronic obstructive pulmonary disease; MAF, minor allele frequency; SNP, single-nucleotide polymorphism.

P values from logistic regression for COPD case–control status, modeling the interaction between SNP and sex, and adjusted for age, pack-years, and principal components. Results from a meta-analysis, including COPDGene non-Hispanic white, COPDGene African American, Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-Points, and Genetics of Chronic Obstructive Lung Disease subjects. Main effect P value represents result from logistic regression in the absence of the SNP × sex interaction term.
plagued with false-positive associations; one group demonstrated that, before 2007, most claims of a sex-difference in genetic effect were either spurious or insufficiently documented (32). Identification of a significant SNP-by-sex interaction requires more power than standard GWASs, and most cohorts assembled for the study of complex diseases lack sufficient sample size to overcome this. In contrast, sex-stratified genetic association tests that are performed in the absence of interaction testing are at increased risk for type I error. Lack of findings in one subgroup could be a result of chance or insufficient power to detect an association. Despite these limitations, several recent studies have demonstrated sex-specific risk factors associated with complex traits using a combination of interaction and sex-stratified testing (11, 14).

COPD is a disease with sexually dimorphic differences in susceptibility and presentation, and we therefore hypothesized that, similar to other complex diseases, there would be sexual dimorphism in the genetic architecture of this disease. To date, we are not aware of any other study that has investigated sexually dimorphic genetic risk factors for COPD at the genome-wide level. Genetic association studies of asthma, another respiratory disease with dimorphic features, have demonstrated sex-specific effects. For example, in a recent analysis, Myers and colleagues (14) identified sex-specific associations with asthma in several genes, including IRF-1 (5q31.1), using a combination of SNP-by-sex interaction testing and stratified association tests. For our analysis, we combined four large COPD case–control cohorts with sufficient female representation to perform two-stage testing, first testing for the SNP-by-sex interaction, then examining our top interaction SNPs in a sex-stratified analysis to identify risk alleles. Our findings identify variants that have biological plausibility for playing a role in sexually dimorphic features of COPD susceptibility.

The CELSR1 gene encodes the cadherin EGF LAG seven-pass G-type receptor 1. CELSR1 is involved in the planar cell polarity pathway of lung branching morphogenesis, and mice with CELSR1 mutations demonstrate abnormal lung development, with smaller lungs and reduced branching compared with healthy lungs (33). Because this gene appears to play a role in lung development, we examined expression levels in human fetal lung tissue samples. We observed that this gene is dimorphically expressed in the human fetal lung. In our genome-wide SNP-by-sex testing, we found variants in the CELSR1 gene that approached significance for an interaction by sex. Although these variants did not meet the prespecified threshold for genome-wide significance, the sex-stratified GWAS combining all four cohorts provides support that CELSR1 may be involved in COPD risk for women, but not for men. This is further supported by our findings that CELSR1 demonstrates differential expression by sex in the fetal lung. Although our most significant SNP is located within an intronic region in the gene, it is in close linkage disequilibrium with a missense SNP. The lung epithelial tube formation is coordinated by the planar cell polarity pathway, of which CELSR1 plays an integral part. Therefore, it is possible that variants within this gene could lead to disruption of the actin–myosin lung cytoskeleton, as well as dysregulated branching morphogenesis. As seen in mouse models, this results in smaller lungs, with less epithelial branches and narrow or absent epithelia.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Male Subjects</th>
<th>Female Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio (95% CI)</td>
<td>P Value</td>
</tr>
<tr>
<td><strong>CELSR1 SNPs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7294481</td>
<td>0.83 (0.74–0.92)</td>
<td>2.49 × 10⁻⁵</td>
</tr>
<tr>
<td>rs9615358</td>
<td>0.90 (0.79–1.01)</td>
<td>0.06</td>
</tr>
<tr>
<td>rs9615982</td>
<td>0.86 (0.71–0.98)</td>
<td>0.01</td>
</tr>
<tr>
<td>rs7286446</td>
<td>0.87 (0.76–0.98)</td>
<td>0.02</td>
</tr>
<tr>
<td>rs9615981</td>
<td>0.87 (0.76–0.98)</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>FAR2 SNPs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7294481</td>
<td>0.83 (0.74–0.92)</td>
<td>2.49 × 10⁻⁵</td>
</tr>
</tbody>
</table>

**Table 3.** Top Sex-Stratified Single-Nucleotide Polymorphisms in the Male and Female Stratified GWAS Study

**Definition of abbreviations:** CI, confidence interval; GWAS, genome-wide association; SNP, single-nucleotide polymorphism.

Presenting sex-stratified analysis of the top SNPs from the SNP × sex interaction test. SNP rs7034632 is excluded from this analysis, as the direction of the interaction effect was not uniform among all cohorts. This meta-analysis includes COPDGene non-Hispanic white and African American, Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-Points, and Genetics of Chronic Obstructive Lung Disease subjects, and is adjusted for age, pack-years, and principal components.
Chr Gene SNP Published – for sexually dimorphic response to the risk of COPD. It is possible that variants in to cigarettes (8), suggesting that lung more likely to develop COPD at an

Recent findings suggest that maximal lung function achieved in early adulthood may influence later development of COPD (36). Several studies suggest that women are more likely to develop COPD at an earlier age (6), and with less exposure to cigarettes (8), suggesting that lung development may play a role in women’s risk of COPD. It is possible that variants in the CELSR1 gene could be responsible for sexually dimorphic response to early-life exposures through altering the size, morphology, and actin–myosin cytoskeleton of female lungs. We only observe statistically significant gene expression differences of CELSR1 by sex in fetal lung tissue. In an adult lung tissue expression dataset of cases with COPD and control subjects, we do not observe differential expression of CELSR1 or FAR2. In a prior sex-stratified analysis of gene expression in subjects with COPD, we broadly observed that there was limited differential gene expression by sex, but extensive differential gene regulation by sex (37). From those data, we observe that CELSR1 is strongly differentially targeted in a female-only COPD gene expression network compared with a male-only COPD gene expression network (false discovery rate adjusted \( P = 2.3 \times 10^{-8} \)), suggesting a potential sexually dimorphic contribution by CELSR1 to COPD through sex-specific gene regulation, not necessarily gene expression.

We additionally identified several variants upstream from the FAR2 gene that also approached significance for an association with COPD affection status. The effect estimate for these SNPs was in the opposite direction for males and females, but the overall \( P \) value in the sex-stratified analysis did not meet the prespecified genome-wide significance threshold. There was a modest increased expression of this gene among female subjects in the fetal lung expression dataset. The role of the FAR2 gene product has not been well described. Modest expression has been identified in pneumocytes and lung endothelial cells (38). The pattern on differential gene targeting by sex observed for CELSR1 was not observed for FAR2.

The candidate gene analysis of known COPD genes and sex-stratified GWASs did not identify significant sex-based differences in known COPD genes; however, there were some interesting differences in how these variants were associated within each sex. We found that previously identified variants upstream from RIN3 did demonstrate a significant SNP-by-sex interaction in our candidate gene approach. The \( P \) value for this variant was much lower among female subjects than male subjects; however, the effect size and direction of effect were similar between both groups. The sex-stratified analysis identified known COPD SNPs within the top genome-wide significant SNPs for each sex; however, different SNPs rose to the top for each sex. Among the females, SNPs in the gene, FAM13A, were significant for COPD. The effect size was in the same direction in both groups, but of greater magnitude among females. The SNP-by-sex interaction was 0.04 for the top SNP, which did not reach genome-wide significance. GWASs have previously demonstrated that this gene is associated with COPD (9, 19) as well as with lung function (39). The sex-stratified approach is essentially a subset analysis; our findings could be related to differences in power to detect an association between the two groups. Alternatively, they could suggest sexually dimorphic associations between these variants that would be better demonstrated through more functional approaches.

Table 4: Interaction and Sex-Stratified Results Published Top COPD GWAS Affection Alleles

<table>
<thead>
<tr>
<th>Chr</th>
<th>Gene</th>
<th>SNP</th>
<th>Published ( P ) value</th>
<th>All Subjects Interaction ( P ) Value</th>
<th>Males ( P ) Value</th>
<th>Females ( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>CHRNA3</td>
<td>rs4416442</td>
<td>( 1.12 \times 10^{-14} )</td>
<td>0.08</td>
<td>( 1.68 \times 10^{-5} )</td>
<td>( 1.78 \times 10^{-9} )</td>
</tr>
<tr>
<td>4</td>
<td>FAM13A</td>
<td>rs12914385</td>
<td>( 6.38 \times 10^{-14} )</td>
<td>0.25</td>
<td>( 1.70 \times 10^{-5} )</td>
<td>( 3.30 \times 10^{-5} )</td>
</tr>
<tr>
<td>4</td>
<td>HHIP</td>
<td>rs13141641</td>
<td>( 1.57 \times 10^{-12} )</td>
<td>0.59</td>
<td>( 1.70 \times 10^{-5} )</td>
<td>( 3.72 \times 10^{-6} )</td>
</tr>
<tr>
<td>14</td>
<td>RIN3</td>
<td>rs754388</td>
<td>( 5.25 \times 10^{-9} )</td>
<td>0.03</td>
<td>( 1.04 \times 10^{-2} )</td>
<td>( 1.84 \times 10^{-7} )</td>
</tr>
<tr>
<td>11</td>
<td>MMP3/12</td>
<td>rs826750</td>
<td>( 5.35 \times 10^{-9} )</td>
<td>0.38</td>
<td>( 7.81 \times 10^{-5} )</td>
<td>( 1.82 \times 10^{-2} )</td>
</tr>
<tr>
<td>1</td>
<td>TGF2B</td>
<td>rs4846480</td>
<td>( 1.25 \times 10^{-7} )</td>
<td>0.86</td>
<td>( 2.48 \times 10^{-3} )</td>
<td>( 1.54 \times 10^{-3} )</td>
</tr>
</tbody>
</table>

Definition of abbreviations: COPD, chronic obstructive pulmonary disease; Chr, chromosome; GWAS, genome-wide association; SNP, single-nucleotide polymorphism.

Published \( P \) values are results from prior COPDGene affection status analysis (19); interaction \( P \) values are results from SNP × sex interaction test for the outcome, COPD case–control status, including COPDGene non-Hispanic white and African American, ECLIPSE (Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-Points), and GenKOLS (Genetics of Chronic Obstructive Lung Disease) subjects. Male and female \( P \) values indicate results from male- and female-stratified meta-analysis of COPDGene non-Hispanic white, African American, ECLIPSE, and GenKOLS subjects. Results are from stratified logistic regression, adjusting for age, pack-years, and principal components.
Our study has several limitations. First, interaction testing requires very large sample sizes to achieve adequate power. We were limited by the availability of adequate COPD case–control datasets with sufficient female subjects, and therefore did not identify any variants that met the prespecified genome-wide threshold for a significant interaction. However, the gene, CELSR1, had many SNPs that approached this threshold and demonstrated differences in association between males and females, and we found supportive evidence of sexually dimorphic expression in the developing lung. This analysis highlights the need for greater inclusion of women in COPD cohorts, and future studies with larger sample sizes may better identify SNP–by-sex interactions. Functional follow-up studies will better identify the mechanisms behind these identified differences. Second, female subjects with COPD in our analysis were younger and had less cumulative smoking exposure (pack-years), although we did adjust for this by including age and pack-years as covariates in our model. Third, we used all available cohorts that had suitable populations of males and females for this analysis. Therefore, we do not have an additional population to replicate our top finding. As additional cohorts are developed that include greater participation of female subjects, future studies can include replication of our findings here.

Finally, racial differences have been demonstrated in COPD susceptibility and presentation (7). Other analyses of COPD-related phenotypes in this population have identified risk variants in the African American population that were not present among the non-Hispanic white population, although these risk alleles were rare (40). We included the African American subjects in our meta-analysis of SNP–by-sex interaction and case–control affection status, as has been done previously (9, 40). However, our study size is underpowered to identify SNP–by-sex interactions among the African American population alone in this study. There is a need to develop more data from African American genotyped populations to better study race-specific sexually dimorphic risk factors for COPD.

In conclusion, this is the first GWAS to identify sex-specific genetic risk factors for COPD. We identified variants in the lung development gene, CELSR1, that are associated with COPD in women, but not men. These findings could identify potential pathways for developmental origins of chronic airflow obstruction among women.

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

**COPDGene Investigators—Core Units**

**Administrative Core:** James D. Crapo, M.D. (Principal Investigator [PI]), Edwin Silverman, M.D., Ph.D. (PI), Barry Make, M.D., Elizabeth Regan, M.D., Ph.D.

**Genetic Analysis Core:** Terri Beatty, Ph.D., Nan Laird, Ph.D., Christophe 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.P.H., Craig Hersch, M.D., M.P.H., Peter Castaldi, M.D., M.S., Merryl Lynn McDonald, Ph.D., Erny Walt, M.D., Megan Hardin, M.D., Jacqueline Hetmanski, M.S., Margaret Parker, M.S., Marilyn Foreman, M.D., Brian Holbs, M.D., Robert Busch, M.D., Adel El-Boueiz, M.D., Peter Castaldi, M.D., Megan Hardin, M.D., Dandi Qiao, Ph.D., Elizabeth Regan, M.D., Elan Halper-Stromberg, Ferdhouse Begum, Sung Ho Won, Shannon Lutz, Ph.D.

**Imaging Core:** David A. Lynch, M.D., 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., Ph.D., Raul San Jose Estepar, Ph.D., Berend C. Stoel, Ph.D., Juerg Tschirren, Ph.D., Eva van Rijkooort, Ph.D., Bram van Ginneken, Ph.D., George Washko, M.D., Carla G. Wilson, M.S., Mustafa Al Qaisi, M.D., Teresa Gray, Alex Kluber, Tanya Mann, Jereed Sieren, Douglas Stinson, Joyce Schroeder, M.D., 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., 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, M.S.P.H., Lindsey Duca, M.S.

**COPDGene Investigators—Clinical Centers**

Ann Arbor Veterans Administration, Ann Arbor, Michigan: Jeffrey L. Curtis, M.D., Carlos H. Martinez, M.D., M.P.H., Perry G. Pernicano, M.D.

Baylor College of Medicine, Houston, Texas: Nicola Hanania, M.S., M.S., Philip Alapat, M.D., Verketa Bandi, M.D., Mustafa Atik, M.D., Ahadin Boriek, Ph.D., Kaipatha Guntupalli, M.D., Elizabeth Guy, M.D., Amit Parulekar, M.D., Arun Nachiappan, M.D.

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

Columbia University, New York, New York: R. Graham Barr, M.D., Dr.P.H., Byron Thomsow, M.D., John Austin, M.D., Belinda D’Souza, M.D., Gregory D. N. Pearson, M.D., Anna Rozansplein, M.D., M.P.H.

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

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

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

Los Angeles Biomedical Research Institute at Harbor University of California Los Angeles 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 Budoft, M.D., Harry Rossitter, Ph.D.

Michael E. DeBakey VAMC, Houston, TX: Amir Sharafkhaneh, M.D., Ph.D., Charlie Lan, D.O.

Minneapolis Veterans Affairs, Minneapolis, Minnesota: Christine Wendt, M.D., Brian Bell, M.D.

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

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

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

Temple University, Philadelphia, Pennsylvania: Gerard Criner, M.D., David Ciccone, M.D., Francis Cordova, M.D., Chandra Dass, M.D., Gilbert D’Abronzo, D.O., Parag Desai, M.D., Michael Jacobs, Ph.D., Steven Kelsen, M.D., Ph.D., Victor Kim, M.D., A. James Mamyry, M.D., Nathaniel Marchetti, D.D., Aditi Satt, M.D., Kartik Shency, M.D., Robert M. Steiner, M.D., Alex Swift, M.D., Irene Swift, M.D., 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., Hirudaya Nith, M.D.

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

University of Iowa, Iowa City, Iowa: Alejandro Cornellas, M.D., John Newell Jr., M.D., Brad Thompson, M.D.

University of Michigan, Ann Arbor, Michigan: Meilan Han, M.D., Ella Kazerooni, M.D., Carlos Martinez, M.D.

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

University of Pittsburgh, Pittsburgh, Pennsylvania: Frank Sciurba, M.D., Divay Chandra, M.D., M.S., Joel Weissfeld, M.D., M.P.H., Carl Fuhrman, M.D., 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., Mario E. Ruiz, M.D.
References

34. Weiss ST, Ware JH. Overview of issues in the longitudinal analysis of respiratory data. Am J Respir Crit Care Med 1996;154: S208–S211.


