Genome-wide linkage analysis of severe, early-onset chronic obstructive pulmonary disease: airflow obstruction and chronic bronchitis phenotypes

Edwin K. Silverman1,2,4,*, Jonathan D. Mosley1, Lyle J. Palmer1,4, Matthew Barth1, Jody M. Senter1, Alison Brown1, Jeffrey M. Drazen2,4, David J. Kwiatkowski4, Harold A. Chapman6, Edward J. Campbell5, Michael A. Province7, D. C. Rao7, John J. Reilly2,4, Leo C. Ginn3,4, Frank E. Speizer1,4 and Scott T. Weiss1,4

1Channing Laboratory and 2Division of Pulmonary and Critical Medicine, Department of Medicine, Brigham and Women’s Hospital, 3Pulmonary and Critical Care Unit, Department of Medicine, Massachusetts General Hospital, 4Harvard Medical School, Boston, MA, USA, 5Department of Internal Medicine, University of Utah Health Sciences Center, Salt Lake City, UT, USA, 6Division of Pulmonary and Critical Care Medicine, University of California at San Francisco, San Francisco, CA, USA and 7Division of Biostatistics, Washington University School of Medicine, St Louis, MO, USA

Received October 8, 2001; Revised and Accepted January 20, 2002

Familial aggregation of chronic obstructive pulmonary disease (COPD) has been demonstrated, but linkage analysis of COPD-related phenotypes has not been reported previously. An autosomal 10 cM genome-wide scan of short tandem repeat (STR) polymorphic markers was analyzed for linkage to COPD-related phenotypes in 585 members of 72 pedigrees ascertained through severe, early-onset COPD probands without severe α1-antitrypsin deficiency. Multipoint non-parametric linkage analysis (using the ALLEGRO program) was performed for qualitative phenotypes including moderate airflow obstruction [forced expiratory volume at one second (FEV1) < 60% predicted, FEV1/FVC < 90% predicted], mild airflow obstruction (FEV1 < 80% predicted, FEV1/FVC < 90% predicted) and chronic bronchitis. The strongest evidence for linkage in all subjects was observed at chromosomes 12 (LOD = 1.70) and 19 (LOD = 1.54) for moderate airflow obstruction, chromosomes 8 (LOD = 1.36) and 19 (LOD = 1.09) for mild airflow obstruction and chromosomes 19 (LOD = 1.21) and 22 (LOD = 1.37) for chronic bronchitis. Restricting analysis to cigarette smokers only provided increased evidence for linkage of mild airflow obstruction and chronic bronchitis to several genomic regions; for mild airflow obstruction in smokers only, the maximum LOD was 1.64 at chromosome 19, whereas for chronic bronchitis in smokers only, the maximum LOD was 2.08 at chromosome 22. On chromosome 12p, 12 additional STR markers were genotyped, which provided additional support for an airflow obstruction locus in that region with a non-parametric multipoint approach for moderate airflow obstruction (LOD = 2.13) and mild airflow obstruction (LOD = 1.43). Using a dominant model with the STR markers on 12p, two point parametric linkage analysis of all subjects demonstrated a maximum LOD score of 2.09 for moderate airflow obstruction and 2.61 for mild airflow obstruction. In smokers only, the maximum two point LOD score for mild airflow obstruction was 3.14. These observations provide suggestive evidence that there is a locus on chromosome 12p which contributes to susceptibility to early-onset COPD.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD), which includes chronic bronchitis, peripheral airways disease and emphysema, is the fourth leading cause of death in the United States (1,2). Cigarette smoking is the major environmental risk factor for the development of COPD, but the development of chronic airflow obstruction—the defining characteristic of COPD—is quite variable among smokers (3).

One genetic determinant of COPD, an inherited severe deficiency of the serine protease inhibitor α1-antitrypsin, has been known for more than 35 years (4). General population studies of families and twins have demonstrated familial aggregation of pulmonary function measurements, and previous studies in relatives of COPD patients have demonstrated familial aggregation of reduced lung function (5–7). However, genetic determinants of COPD other than severe α1-antitrypsin deficiency, which
accounts for only 1–2% of COPD cases, have not been proven. A variety of candidate genes have been assessed using case-control association studies, but the results have not been consistent across studies (8). Linkage studies of COPD have not been reported previously.

To identify regions of the genome linked to COPD, we have assembled pedigrees ascertained through single probands with severe, early-onset COPD who did not have severe α1-antitrypsin deficiency. Analogous to other complex disorders including breast cancer and glaucoma, we hypothesized that subjects with severe, early-onset COPD might be more likely to have a genetic etiology for their disease (9,10).

Phenotype definition in studies of COPD genetics is problematic, because many individuals with COPD are undiagnosed. Therefore, assessment of COPD-related phenotypes, including airflow obstruction and chronic bronchitis, is essential for the study of COPD genetics. Airflow obstruction is assessed by pulmonary function tests that measure forced expiratory flow; airflow obstruction is defined by an absolute reduction in the forced expiratory volume at 1 s (FEV₁) coincident with a reduction in the ratio of the FEV₁ to the forced vital capacity (FVC). Among subjects with COPD, greater disease severity is associated with greater reductions in measured FEV₁ compared to predicted values based on the age, height, gender and race of the subject. Chronic bronchitis, defined as a chronic cough productive of phlegm, is present in many but not all COPD patients; individuals with chronic bronchitis but without airflow obstruction do not have COPD (11).

In previous studies of early-onset COPD pedigrees, we have demonstrated an increased risk to current or ex-smoking first-degree relatives of severe, early-onset COPD probands for reduced FEV₁ and chronic bronchitis, with an estimated risk to smoking first-degree relatives for these phenotypes of approximately three times the risk to smokers from the general population (12). We found no increased risk for lifelong non-smoking first-degree relatives for reduced FEV₁ or chronic bronchitis. Thus, a genotype-by-smoking interaction could account for the increased risk to relatives for reduced FEV₁ and chronic bronchitis in early-onset COPD pedigrees.

To maximize information from these rare severe, early-onset COPD probands, we have included first-degree relatives and older second-degree relatives (half-sibs, aunts, uncles and grandparents) of each proband and we have expanded the pedigrees to include first-degree relatives of any pedigree members with moderate airflow obstruction. We now report the results of the first whole genome scan for COPD using autosomal markers at ∼10 cM intervals with phenotypes including airflow obstruction and chronic bronchitis. Based on the linkage results from the genome scan, additional short tandem repeat (STR) markers have been analyzed on chromosome 12p.

## RESULTS

### Demographics and affection status in early-onset COPD families

The age, smoking history and affection status for the 585 members of 72 pedigrees included in the genome scan linkage analyses are presented in Table 1. For the moderate airflow obstruction phenotype (FEV₁ < 60% predicted with FEV₁/FVC < 90% predicted), a high percentage of parents (38%) were affected, but only 48 parents were included in the study. Among siblings, 15% had moderate airflow obstruction; however, only one child of a proband met the criteria for moderate airflow obstruction. When the sample was limited to cigarette smokers with at least 10 pack-years of smoking, an increased percentage of siblings, parents and more distant relatives had moderate airflow obstruction.

<table>
<thead>
<tr>
<th>Relative group/smoking status</th>
<th>n</th>
<th>Age (mean ± SD)</th>
<th>Pack-years (mean ± SD)</th>
<th>Mild airflow obstruction n (%)</th>
<th>Moderate airflow obstruction n (%)</th>
<th>Chronic bronchitis n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probands: all subjects</td>
<td>72</td>
<td>47.7 ± 5.3</td>
<td>38.9 ± 21.6</td>
<td>72 (100)</td>
<td>72 (100)</td>
<td>34 (47)</td>
</tr>
<tr>
<td>Probands: smokers only</td>
<td>66</td>
<td>48.3 ± 4.4</td>
<td>42.3 ± 19.3</td>
<td>66 (100)</td>
<td>66 (100)</td>
<td>31 (47)</td>
</tr>
<tr>
<td>Siblings: all subjects</td>
<td>148</td>
<td>47.0 ± 9.0</td>
<td>22.2 ± 20.9</td>
<td>46 (31)</td>
<td>22 (15)</td>
<td>31 (21)</td>
</tr>
<tr>
<td>Siblings: smokers only</td>
<td>94</td>
<td>48.1 ± 7.4</td>
<td>34.2 ± 16.9</td>
<td>41 (44)</td>
<td>21 (22)</td>
<td>29 (31)</td>
</tr>
<tr>
<td>Parents: all subjects</td>
<td>48</td>
<td>72.5 ± 8.1</td>
<td>38.0 ± 39.4</td>
<td>29 (60)</td>
<td>18 (38)</td>
<td>11 (23)</td>
</tr>
<tr>
<td>Parents: smokers only</td>
<td>30</td>
<td>73.4 ± 6.5</td>
<td>60.2 ± 34.2</td>
<td>27 (90)</td>
<td>18 (60)</td>
<td>11 (37)</td>
</tr>
<tr>
<td>Children: all subjects</td>
<td>125</td>
<td>24.7 ± 5.7</td>
<td>4.3 ± 6.3</td>
<td>11 (9)</td>
<td>1 (1)</td>
<td>20 (16)</td>
</tr>
<tr>
<td>Children: smokers only</td>
<td>25</td>
<td>28.5 ± 3.4</td>
<td>15.2 ± 3.7</td>
<td>5 (20)</td>
<td>0 (0)</td>
<td>9 (36)</td>
</tr>
<tr>
<td>Other relatives: all subjects</td>
<td>161</td>
<td>54.4 ± 18.0</td>
<td>23.1 ± 25.6</td>
<td>52 (32)</td>
<td>21 (13)</td>
<td>31 (19)</td>
</tr>
<tr>
<td>Other relatives: smokers only</td>
<td>100</td>
<td>55.8 ± 15.8</td>
<td>36.3 ± 24.3</td>
<td>43 (43)</td>
<td>16 (16)</td>
<td>22 (22)</td>
</tr>
<tr>
<td>Spouses: all subjects</td>
<td>31</td>
<td>51.7 ± 6.0</td>
<td>37.1 ± 34.5</td>
<td>7 (23)</td>
<td>3 (10)</td>
<td>5 (16)</td>
</tr>
<tr>
<td>Spouses: smokers only</td>
<td>21</td>
<td>52.2 ± 7.1</td>
<td>54.5 ± 28.3</td>
<td>5 (24)</td>
<td>3 (14)</td>
<td>5 (24)</td>
</tr>
</tbody>
</table>

*Smokers includes subjects with ≥10 pack-years.

*All probands were defined as affected for airflow obstruction, but two subjects had FEV₁/FVC > 90% predicted, probably related to insufficient exhalation times.

*Spirometry was not available for one child, one other relative and one spouse. Mild airflow obstruction was defined as FEV₁ < 80% predicted and FEV₁/FVC < 90% predicted. Moderate airflow obstruction was defined as FEV₁ < 60% predicted and FEV₁/FVC < 90% predicted.
Using the more liberal criteria for mild airflow obstruction (FEV₁ < 80% predicted with FEV₁/FVC < 90% predicted), more affected relatives were noted. A high percentage of parents (60%), siblings (31%) and more distant relatives (32%) met this threshold; however, only 9% of children were affected with mild airflow obstruction. Among smokers with at least 10 pack-years of smoking, increased rates of affection with mild airflow obstruction were noted, including 90% of parents.

Chronic bronchitis was present in only 47% of early-onset COPD probands. Moderate rates of chronic bronchitis in siblings (21%), parents (23%), children (16%) and more distant relatives (19%) were noted; these rates were increased among current or ex-smokers in each of these groups of relatives.

### Genome-wide multipoint non-parametric linkage analysis

Genome-wide multipoint non-parametric linkage results using ALLEGRO for moderate airflow obstruction, mild airflow obstruction and chronic bronchitis are presented in Table 2. For moderate airflow obstruction, allele-sharing LOD scores above 1.0 were observed using all subjects on chromosomes 4, 8, 12 and 19 (LOD = 1.24, P = 0.008), 12 (LOD = 1.70, P = 0.003), 15 (LOD = 1.07, P = 0.01), 18 (LOD = 1.01, P = 0.02) and 19 (LOD = 1.54, P = 0.004). Among smokers with at least 10 pack-years of smoking, multipoint non-parametric linkage analysis demonstrated some reduction in the evidence for linkage on chromosomes 4, 12 and 18, while no reduction in linkage evidence was observed on chromosomes 15 and 19, despite the inclusion of phenotypic data from fewer affected individuals.

Multipoint linkage results for mild airflow obstruction in all subjects demonstrated only two regions with allele-sharing LOD scores above 1.0, which were located on chromosomes 4 (LOD = 1.24, P = 0.008) and 18 (LOD = 1.01, P = 0.02) and 19 (LOD = 1.54, P = 0.004). Among smokers with at least 10 pack-years of smoking, multipoint non-parametric linkage analysis demonstrated some reduction in the evidence for linkage on chromosomes 4, 12 and 18, while no reduction in linkage evidence was observed on chromosomes 15 and 19, despite the inclusion of phenotypic data from fewer affected individuals.

Multipoint linkage results for mild airflow obstruction in all subjects demonstrated only two regions with allele-sharing LOD scores above 1.0, which were located on chromosomes 4 (LOD = 1.24, P = 0.008) and 18 (LOD = 1.01, P = 0.02) and 19 (LOD = 1.54, P = 0.004). Among smokers with at least 10 pack-years of smoking, multipoint non-parametric linkage analysis demonstrated some reduction in the evidence for linkage on chromosomes 4, 12 and 18, while no reduction in linkage evidence was observed on chromosomes 15 and 19, despite the inclusion of phenotypic data from fewer affected individuals.

Multipoint linkage results for chronic bronchitis in all subjects demonstrated two regions with allele-sharing LOD scores above 1.0, on chromosomes 19 (LOD = 1.21, P = 0.009) and 22 (LOD = 1.37, P = 0.006). Of note, in smokers only, the LOD scores on chromosomes 19 and 22 increased, and additional regions with LOD scores above 1.0 were detected on chromosomes 6 (LOD = 1.51, P = 0.004) and 12 (LOD = 1.14, P = 0.01).

In the genome scan linkage analysis of moderate airflow obstruction, mild airflow obstruction and chronic bronchitis, two chromosomal regions (12p and 19) were found to have a LOD score above 1.0 in smokers for all three phenotypes. Chromosome 22q, which demonstrated the strongest evidence for linkage to chronic bronchitis, also had some evidence for linkage to moderate airflow obstruction (LOD = 0.63, P = 0.04) and mild airflow obstruction (LOD = 0.73, P = 0.03) in all subjects; however, the strongest evidence for linkage to mild airflow obstruction on chromosome 22 was located ~20 cM from the linkage to chronic bronchitis. Chromosomes 3, 4 and 6 had some evidence for linkage to airflow obstruction and chronic bronchitis as well, but the locations of the linkage peaks differed substantially between the airflow obstruction and chronic bronchitis phenotypes for these chromosomes.

### Additional markers on chromosome 12p: two point linkage analysis

Because our most stringent definition of airflow obstruction showed the strongest evidence for linkage to chromosome 12p in the initial genome scan, fine mapping was performed using four markers from the original genome scan (GATA49D12, Mfd259, GATA6C01 and ATA27A06) and 12 additional STR markers located between 18 and 49 cM on chromosome 12. Since multipoint linkage analysis is sensitive to errors in marker order and intermarker distances, two point linkage analysis was also performed in this region. Two point linkage analysis was performed with one dominant and one recessive model, assuming penetrance of 0.5, disease allele frequency of 0.01 and no phenocopies. Subjects were only classified as unaffected for the mild and moderate airflow obstruction phenotypes if they were at least 30 years old, with at least 10 pack-years of smoking and FEV₁ > 80% predicted. The maximum LOD score and the model that provided this score

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Moderate airflow obstruction</th>
<th>Mild airflow obstruction</th>
<th>Chronic bronchitis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All subjects max LOD</td>
<td>Smokers max LOD</td>
<td>All subjects max LOD</td>
</tr>
<tr>
<td>3</td>
<td>0.68</td>
<td>26</td>
<td>0.76</td>
</tr>
<tr>
<td>4</td>
<td>1.24</td>
<td>70</td>
<td>0.83</td>
</tr>
<tr>
<td>6</td>
<td>0.20</td>
<td>191</td>
<td>0.28</td>
</tr>
<tr>
<td>8</td>
<td>0.99</td>
<td>79</td>
<td>0.89</td>
</tr>
<tr>
<td>12</td>
<td>1.70</td>
<td>36</td>
<td>1.28</td>
</tr>
<tr>
<td>15</td>
<td>1.07</td>
<td>38</td>
<td>1.18</td>
</tr>
<tr>
<td>18</td>
<td>1.01</td>
<td>13</td>
<td>0.50</td>
</tr>
<tr>
<td>19</td>
<td>1.54</td>
<td>42</td>
<td>1.65</td>
</tr>
<tr>
<td>22</td>
<td>0.63</td>
<td>46</td>
<td>0.52</td>
</tr>
</tbody>
</table>

The highest allele-sharing LOD score on each chromosome for each phenotype is presented.
are shown in Table 3 for moderate airflow obstruction, mild airflow obstruction and chronic bronchitis.

For moderate airflow obstruction, the strongest evidence for linkage in all subjects was observed at GATA6C01 (36.1 cM) with a maximum LOD score of 2.09 in a dominant model. Other markers in that region provided some evidence for linkage, with LOD scores above 1.0 in dominant models noted at D12S1715 (35.5 cM), D12S1630 (36.1 cM) and D12S1682 (38.5 cM). D12S308 (31.2 cM) had a maximum LOD score of 1.1 in a recessive model. Assessment of linkage in smokers only did not markedly change the evidence for linkage.

For mild airflow obstruction, the maximum observed LOD score in all subjects was observed at D12S1715 (LOD = 2.61 in a dominant model) at 35.5 cM. Several other markers in the region, including Mfd259 (30.6 cM), D12S1650 (38.5 cM), D12S1682 (38.5 cM) also had maximum LOD scores above 1.0 in dominant models. In smokers only, some of the markers had substantially increased evidence for linkage. For D12S1715, the maximum LOD score in smokers increased to 3.14. For Mfd259, the maximum LOD score increased from 1.16 in all subjects to 2.30 in smokers only; similar increases in LOD scores in smokers only were observed at D12S1630 and GATA6C01.

For chronic bronchitis, no LOD scores above 1.0 were noted in all subjects or smokers only. The maximum observed LOD score in all subjects was 0.56 at D12S363 in a recessive model, while the maximum observed LOD score in smokers only was 0.95 at Mfd259 in a recessive model.

To assess the statistical significance of the mild airflow obstruction linkage to D12S1715 in smokers, which was the most notable linkage result on 12p, a series of simulations were performed. Using the SIMULATE program, 5000 replicates of the 72 pedigrees were created assuming no linkage of D12S1715 to COPD-related phenotypes. These replicates were analyzed for linkage of mild airflow obstruction to D12S1715 with the MSIM program of SLINK, using both a dominant and a recessive model, in all subjects and smokers only (13–15). None of these four analyses revealed any LOD scores above 3.13 in the 5000 replicates. To adjust for the multiple comparisons involved with three phenotypes (moderate airflow obstruction, mild airflow obstruction and chronic bronchitis), these 5000 replicates were analyzed for linkage to all three phenotypes, with a dominant and recessive model, in all subjects and smokers only; thus, 12 analyses were performed in total. Four replicates had a LOD score above 3.13 for at least one of the 12 analyses, yielding an estimated P-value of 0.0008 for the D12S1715 linkage to mild airflow obstruction.

### Additional markers on chromosome 12p: multipoint linkage analysis

The multipoint non-parametric linkage analysis results for chromosome 12, including the 12 additional flanking markers on 12p, are depicted in Figure 1 for all subjects and smokers with at least 10 pack-years of smoking. In addition to the non-parametric allele-sharing LOD scores, the parametric heterogeneity LOD scores are presented for a dominant model with penetrance 0.5, allele frequency 0.01 and no phenocopies.

### Table 3. Two point linkage analysis on chromosome 12p: MMLS approach

<table>
<thead>
<tr>
<th>Marker</th>
<th>cM</th>
<th>Moderate airflow obstruction</th>
<th>Mild airflow obstruction</th>
<th>Chronic bronchitis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All subjects</td>
<td>Smokers</td>
<td>All subjects</td>
</tr>
<tr>
<td>GATA49D12</td>
<td>18</td>
<td>0.13 (D)</td>
<td>0.28 (D)</td>
<td>-0.07 (D)</td>
</tr>
<tr>
<td>D12S1695</td>
<td>19.7</td>
<td>0.08 (R)</td>
<td>0.10 (R)</td>
<td>-0.16 (R)</td>
</tr>
<tr>
<td>D12S1690</td>
<td>20.3</td>
<td>0.04 (D)</td>
<td>0.15 (R)</td>
<td>-0.14 (D)</td>
</tr>
<tr>
<td>D12S89</td>
<td>23.4</td>
<td>0.04 (D)</td>
<td>0.08 (D)</td>
<td>-0.21 (R)</td>
</tr>
<tr>
<td>D12S1581</td>
<td>29.7</td>
<td>0.32 (D)</td>
<td>0.19 (R)</td>
<td>-0.05 (D)</td>
</tr>
<tr>
<td>Mfd259</td>
<td>30.6</td>
<td>0.76 (D)</td>
<td>0.89 (R)</td>
<td>1.16 (D)</td>
</tr>
<tr>
<td>D12S308</td>
<td>31.2</td>
<td>1.10 (R)</td>
<td>0.94 (R)</td>
<td>0.92 (D)</td>
</tr>
<tr>
<td>D12S1303</td>
<td>32.5</td>
<td>-0.07 (D)</td>
<td>-0.03 (D)</td>
<td>0.02 (R)</td>
</tr>
<tr>
<td>D12S1715</td>
<td>35.5</td>
<td>1.44 (D)</td>
<td>1.50 (D)</td>
<td>2.61 (D)</td>
</tr>
<tr>
<td>D12S1630</td>
<td>36.1</td>
<td>1.43 (D)</td>
<td>1.67 (D)</td>
<td>0.95 (D)</td>
</tr>
<tr>
<td>GATA6C01</td>
<td>36.1</td>
<td>2.09 (D)</td>
<td>1.65 (D)</td>
<td>0.54 (D)</td>
</tr>
<tr>
<td>D12S363</td>
<td>36.1</td>
<td>0.25 (R)</td>
<td>0.39 (R)</td>
<td>-0.07 (D)</td>
</tr>
<tr>
<td>D12S1650</td>
<td>38.5</td>
<td>0.67 (D)</td>
<td>0.80 (D)</td>
<td>2.13 (D)</td>
</tr>
<tr>
<td>D12S1682</td>
<td>38.5</td>
<td>1.49 (D)</td>
<td>1.05 (D)</td>
<td>1.54 (D)</td>
</tr>
<tr>
<td>D12S1066</td>
<td>41.6</td>
<td>0.74 (R)</td>
<td>0.35 (D)</td>
<td>0.18 (D)</td>
</tr>
<tr>
<td>ATA27A06</td>
<td>48.7</td>
<td>0.09 (R)</td>
<td>0.05 (R)</td>
<td>0.30 (D)</td>
</tr>
</tbody>
</table>

*For LINKAGE results, maximum LOD of one dominant and one recessive model is presented.

*For markers at identical locations on Marshfield Map, order was determined from Golden Path and Celera databases.
A dominant model was selected, because the most significant evidence for linkage in two point analysis was seen with a dominant model; in addition, we previously found that >50% of early-onset COPD probands reported that at least one parent had COPD, which would be consistent with a dominant mode of inheritance (12).

For moderate airflow obstruction, the maximum allele-sharing LOD score in all subjects was 2.13 ($P = 0.0009$) at 36.1 cM.
DISCUSSION
COPD includes chronic bronchitis, which is associated with increased mucus production within the bronchi, peripheral airways disease, which is related to inflammation and fibrosis of the bronchioles, and emphysema, which involves destruction of the pulmonary alveoli. Like most complex diseases, COPD is relatively unique among complex diseases in that the key genetic risk factor for COPD, α1-antitrypsin deficiency, has been excluded. COPD is probably a heterogeneous syndrome influenced by multiple genetic and environmental factors, as well as genotype-by-environment interactions. To limit the complexity and increase the likelihood of success in novel gene identification in COPD, we have focused on pedigrees of individuals with severe, early-onset COPD in which severe α1-antitrypsin deficiency, a proven genetic risk factor for COPD, had been excluded. COPD is relatively unique among complex diseases in that the key environmental risk factor, cigarette smoking, is both known and readily quantified. Therefore, we have assessed the impact of smoking-related susceptibility by performing linkage analysis in all subjects and in smokers only.

We have focused on phenotypes related to airflow obstruction and chronic bronchitis, and we have performed multipoint non-parametric linkage analysis using a 10 cM genome scan of STR markers. We employed two thresholds to define affection for airflow obstruction—the defining characteristic of COPD. Despite a relatively modest number of affected relatives that met the stringent criteria for moderate airflow obstruction (n = 62), some evidence for linkage was identified at several chromosomal regions, with a maximum LOD score of 1.70 observed on chromosome 12p. Since all subjects with moderate airflow obstruction also had mild airflow obstruction, we would expect substantial overlap in the evidence for linkage to mild and moderate airflow obstruction. When all subjects were included in linkage analysis, the mild airflow obstruction phenotype demonstrated the most suggestive evidence of linkage to chromosomes 8 and 19. Of note, the evidence for linkage to mild airflow obstruction increased substantially in multiple chromosomal regions when affected subjects were limited to cigarette smokers with at least 10 pack-years of smoking; with this stratification, evidence for linkage to chromosomes 12p (maximum LOD = 1.54) and 19 (maximum LOD = 1.64) increased. Thus, stratification by smoking status provided stronger evidence for linkage to the more common, and potentially more heterogeneous, phenotype of mild airflow obstruction in several regions that demonstrated some evidence for linkage to the more stringent phenotype of moderate airflow obstruction. These regions may be sites of genetic determinants that confer a genotype-by-smoking interaction. Alternatively, limiting affected individuals to smokers may have increased the homogeneity of subjects with airflow obstruction and/or chronic bronchitis. For example, individuals with airflow obstruction related to asthma rather than COPD may have been removed from the analysis, increasing the power to detect linkage to smoking-related lung disease.

Chronic bronchitis is absent in many cases of COPD; more than half of the severe, early-onset COPD probands did not have chronic bronchitis at the time they were enrolled into the study. The most compelling evidence for linkage to chronic bronchitis in multipoint non-parametric linkage analysis of all subjects was found on chromosome 22; the multipoint LOD score increased from 1.37 (P = 0.006) to 2.08 (P = 0.001) when the analysis was limited to smokers only. Some evidence for linkage to chromosomes 12 and 19, in the same regions where linkage evidence was found for the airflow obstruction phenotypes, was found for chronic bronchitis in the initial genome scan analysis; however, the evidence for linkage of chronic bronchitis to chromosome 12 diminished with the inclusion of additional markers (Fig. 1C). Thus, it is possible that some shared genetic determinants influence the development of airflow obstruction and chronic bronchitis, and chromosomes 12 and 19 appear to be the most likely candidate regions.

Within the region of linkage to airflow obstruction on chromosome 12p, additional STR markers were genotyped to increase the information content available for linkage analysis. Within this region, both two point maximized maximum LOD score (MMLS) and multipoint linkage analysis were performed. These methods provided complementary approaches to assess linkage. The two point parametric MMLS approach may offer increased power compared to non-parametric methods, but it does require stipulation of parameter values, such as allele frequency and phenocopy rate, which are difficult to estimate. The multipoint approach incorporates information from adjacent markers, but some reduction in pedigree size was required. Multipoint non-parametric linkage analysis using all of the chromosomes 12 markers provided suggestive evidence for linkage to moderate airflow obstruction, with a maximum LOD score of 2.13 in all subjects. Two point parametric linkage analysis demonstrated that many of the flanking markers provided some evidence for linkage to airflow obstruction on chromosome 12p. The maximum two point LOD score for mild airflow obstruction in smokers was 3.14 in a dominant model. Simulations were performed to assess the impact of analyzing multiple phenotypes with a dominant and recessive model in all subjects and smokers only; these simulations indicated that this maximum two point LOD score provided suggestive evidence for linkage with an estimated P-value of 0.0008. Multipoint parametric linkage analysis under non-homogeneity provided evidence for linkage of airflow obstruction to chromosome 12p with holds of 2.58 for moderate airflow obstruction in all subjects and 3.19 for mild...
airflow obstruction in smokers. The stronger evidence for linkage to 12p with parametric rather than non-parametric methods could relate to additional information provided by unaffected subjects in parametric methods and the increased power of parametric methods that include the appropriate mode of inheritance (16). Taken together, these results provide suggestive evidence that a subset of early-onset COPD pedigrees is influenced by a susceptibility locus for airflow obstruction on chromosome 12p.

Although genome scan linkage analysis has not been reported previously in COPD, multiple genome scans have been published for asthma (17–23). Whether asthma, which also includes airflow obstruction as a central feature, and COPD are distinct or related conditions has been an ongoing controversy in pulmonary medicine for decades. Some investigators have argued that asthma and COPD are part of the same basic pathophysiologic process (the Dutch Hypothesis), while others have argued that COPD and asthma are unrelated conditions (3,24). Within our early-onset COPD pedigrees, we cannot determine with certainty whether airflow obstruction and chronic bronchitis in relatives were caused by COPD and/or asthma. For the three chromosomal regions that provided the most compelling evidence for linkage to COPD-related phenotypes (12p, 19 and 22q), chromosomes 12p and 19 have been linked to asthma-related phenotypes in previous linkage studies of asthma. The maximal reported evidence for linkage on chromosome 19 to asthma and airway hyper-responsiveness has been located at 67–88 cM; in the Collaborative Study on the Genetics of Asthma Caucasian group, LOD scores above 1.0 were also observed at ∼42 cM on chromosome 19 (17,19,21). In our COPD pedigrees, the maximal evidence for linkage to airflow obstruction and chronic bronchitis is at ∼42 cM; however, the evidence for linkage to moderate airflow obstruction in all subjects in our COPD pedigrees remains at $P < 0.05$ from 23 to 78 cM. Therefore, it is possible that the asthma and COPD linkage ages on chromosome 19 do overlap. On chromosome 12p, several previous asthma studies have found nominal evidence for linkage to asthma and atopy-related phenotypes ($P < 0.05$) in the same general region demonstrating linkage to airflow obstruction in our study (21,25). Further work will be required to determine whether shared genetic determinants for asthma and COPD are located on chromosomes 12, 19 or other genomic regions.

Our study has several important limitations. To enrich our sample for genetically influenced causes of COPD, we have focused on probands with severe, early-onset disease. However, our strict enrolment criteria limited the available sample size and may have compromised our power to detect linkage. Moreover, the generalizability of this severe, early-onset COPD population to COPD at later ages is uncertain; this population may represent an interesting but rare cause of COPD, similar to severe α1-antitrypsin deficiency. In addition, we have stratified by smoking status, but we were unable to include genotype-by-environment interaction. We have selected clinically reasonable discrete phenotypes that provided more power to detect linkage in analyses of α1-antitrypsin deficiency (26), but quantitative phenotypes may provide more information to detect novel COPD genes in our current study. Future work will include quantitative phenotype linkage analysis. Finally, linkage analysis of three COPD-related phenotypes in all subjects and smokers only does introduce multiple comparisons; however, these phenotypes are correlated, and we assessed the impact of multiple comparisons by simulation of our two point linkage analysis results.

Clearly, assessment of linkage in additional early-onset COPD pedigrees will be required before firm inferences can be made regarding the genetics of COPD susceptibility. However, based on our initial results, several insights can be drawn. First, we have not found evidence for conventional significant linkage in this initial genome scan; LOD score thresholds of 3.3 for significant linkage and 1.9 for suggestive linkage have been suggested for pedigree-based LOD score analysis (27). However, we have identified several regions of suggestive linkage that require further investigation as potential sites of COPD genetic determinants. Our initial genome scan results provided suggestive evidence for linkage of chronic bronchitis in smokers to chromosome 22q. Genotyping flanking STR markers on chromosome 12p provided suggestive evidence for linkage of airflow obstruction to this region. Stratification by the key environmental risk factor, cigarette smoking, provided substantially increased evidence for linkage in some chromosomal regions, despite inclusion of fewer affected subjects. Finally, shared genetic determinants may be present for asthma and COPD. The genetic determinants of COPD are likely to be multiple, and genotype-by-environment interactions are likely critical factors in the development of COPD. Future work in severe, early-onset COPD pedigrees may lead to the identification of novel genetic influences on COPD development, which may result in improved understanding of COPD pathophysiology and new opportunities for prevention and treatment.

**MATERIALS AND METHODS**

**Study participants**

Enrolment of probands with severe early-onset COPD has been described previously (28). In brief, enrolment criteria for probands with severe early-onset COPD included: FEV1 < 40% of the predicted value, age <53 years, and absence of severe α1-antitrypsin deficiency (e.g. PI*Z, PI null-null). The arbitrary age threshold of 53 years was chosen to balance our goal of identifying very young probands with our need to ascertain an adequate number of probands. Subjects who had undergone lung transplantation were not eligible to participate and subjects who had undergone lung volume reduction surgery were included only if preoperative pulmonary function tests were available (which were used in the analysis).

All available first-degree relatives and older second-degree relatives (half-sibs, aunts, uncles and grandparents) of the enrolled COPD probands were invited to participate. In addition to this fixed ascertainment scheme, 49 additional relatives were enrolled, including: (i) 37 first-degree relatives of pedigree members with moderate airflow obstruction (FEV1 < 60% predicted, FEV1/FVC < 90% predicted); (ii) three other relatives who had been diagnosed with COPD; (iii) two first-degree relatives of unavailable, reportedly affected pedigree members; and (iv) seven other relatives. Pedigree sizes ranged from two to 18 genotyped individuals.

Participants gave written informed consent and completed a protocol which included a questionnaire, spirometry and a blood sample. The protocol was approved by the Human Research Committees of Partners HealthCare (Brigham and
Women’s Hospital and Massachusetts General Hospital) and the Brockton/West Roxbury VA Hospital.

**Questionnaire**

Each participant completed a modified version of the 1978 ATS-DLD Epidemiology Questionnaire (12,29). Pack-years of cigarette smoking was calculated as the product of the duration of smoking (in years) and the average number of cigarettes smoked per day, which was divided by 20 to convert to packs. Chronic bronchitis was defined as affirmative responses to questionnaire items for both chronic cough and chronic phlegm production for at least 3 months per year for at least 2 years.

**Pulmonary function tests**

Spirometry was performed with a Survey Tach Spirometer (Warren E. Collins, Braintree, MA). The manoeuvres were performed in a standardized manner with subject seated and wearing a nose clip. Spirometry was performed in accordance with ATS guidelines (30). Subjects were instructed to withhold bronchodilator medications for at least 4 h before spirometry if possible. Pulmonary function test results are expressed as percent of predicted results using predicted equations from Crapo et al. (31) for adult Caucasian participants. For Caucasian participants under 18 years of age, predicted values for FEV\(_1\) were determined from Hsu et al. (32) and predicted values for FEV\(_1\)/FVC were determined from Knudson et al. (33). For African-American participants, predicted spirometric values were determined from Hankinson et al. (34). For seven subjects, previous hospital pulmonary function test results were used; three subjects had undergone lung volume reduction surgery and four other subjects were located at great geographic distance. Spirometry values were missing for three other subjects.

**α\(_1\)**-Antitrypsin studies**

The PI type of each early-onset COPD proband was determined by isoelectric focusing of dithioerythritol-treated serum at pH 4.2–4.9 in polyacrylamide gels, and immunoreactive and functional α\(_1\)-antitrypsin levels were measured as described previously (28). For probands who were found to be carriers of the deficient Z allele, all of their first-degree relatives underwent α\(_1\)-antitrypsin typing by isoelectric focusing; no subjects with severe α\(_1\)-antitrypsin deficiency were identified.

**Genotyping and data cleaning**

A total of 607 individuals were genotyped by the NHLBI Mammalian Genotyping Service, using DNA that was extracted from blood samples with Puregene Kits (Gentra Systems). We analyzed 378 autosomal STR markers with an average spacing of 9.1 cm, ranging from 7.4 cm on chromosome 11 to 10.6 cm on chromosome 14. Marker locations were determined from version 10 of the Marshfield Map at the Marshfield website (http://research.marshfieldclinic.org/genetics). For markers with exactly the same location on the Marshfield Map, a distance of 0.1 cm was placed between these markers after determining relative locations on the Human Genome Working Draft (Golden Path at http://genome.cse.ucsc.edu) and the Celera Genomics Human Genome Sequence (http://cds.celera.com). The average genotype completeness per marker in the initial genome scan was >98%.

On chromosome 12, 12 additional STR markers were genotyped at Brigham and Women’s Hospital. Fluorescent-labeled and unlabeled primers were obtained from Research Genetics and Applied Biosystems. PCR was performed with Taq Gold Polymerase (Applied Biosystems) in MJ Research PCR machines. Product sizes were assessed on an ABI 3100; Genescan and Genotyper Version 3.7 software were used to assist with genotype determination.

To assess for Mendelian inconsistencies in pedigree data, the RELCHECK program was used to determine pedigree relationships based on the marker data (35). Twenty-two subjects failed to match reported familial relationships and were excluded. Using the remaining 585 genotyped individuals, inconsistencies at individual markers were resolved using the PEDCHECK program (36). PEDCHECK revealed an average of 0.95 pedigree inconsistencies per STR marker in the genome scan data.

Marker allele frequencies were estimated using the SIB-PAIR program, with a weighted average based on the number of typed founders in the pedigree (37).

**Linkage analysis**

Phenotypes selected for linkage analysis included chronic bronchitis (defined by questionnaire responses to chronic cough and chronic phlegm production) and airflow obstruction. Because reduced FEV\(_1\) could be caused by restrictive pulmonary processes (e.g. interstitial lung disease, neuromuscular weakness) as well as obstructive pulmonary processes (e.g. COPD, asthma), we included reduction in the ratio of FEV\(_1\)/FVC in definitions of airflow obstruction phenotypes. We selected a moderate airflow obstruction phenotype, defined as FEV\(_1\) < 60% predicted and FEV\(_1\)/FVC < 90% predicted, and a mild airflow obstruction phenotype, defined as FEV\(_1\) < 80% predicted and FEV\(_1\)/FVC < 90% predicted. Thus, all subjects with moderate airflow obstruction were also included within the mild airflow obstruction group. Multipoint linkage analysis of the genome scan genotyping data was performed with the non-parametric affected relative method as implemented in the ALLEGRO program (version 1.1-b) (38). Some trimming of our pedigrees was necessary for the ALLEGRO analyses; 14 unaffected individuals were removed due to computational constraints. For multipoint linkage analysis, the exponential model, the Score = All option (to examine all affected relatives within a pedigree simultaneously) (39), and five interval steps between each marker location were employed. The multipoint allele-sharing LOD scores and the corresponding P-values are reported from the ALLEGRO program output.

In the region of suggestive linkage to airflow obstruction on chromosome 12p, two point parametric linkage analysis with the LINKAGE program as implemented in FASTLINK (Version 4.1 P) was performed (13,40). Because the use of parametric linkage analysis with one dominant and one recessive model has been suggested to have increased power to detect linkage compared to non-parametric methods, we performed two point parametric linkage analysis with one dominant model with penetrance 0.5 and one recessive model with penetrance 0.5, as described in the MMLS approach (16,41). Linkage was evaluated at recombination fractions of 0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35 and 0.4, and the maximum LOD score is reported. To correct for multiple
comparisons, the LOD score threshold for the parametric models should be increased by 0.3, since we have not subtracted this value from the two point LOD scores presented (16). For subjects to be defined as unaffected in the parametric linkage analysis of airflow obstruction phenotypes, we required that they be >30 years of age with FEV₁ > 80% predicted and at least 10 pack-years of smoking. Multipoint linkage analysis using flanking markers on chromosome 12p was performed with ALLEGRO; both parametric linkage analysis under heterogeneity and non-parametric linkage analysis were used.

To assess smoking-related effects, linkage analyses were performed using all subjects and smokers only for both two point and multipoint linkage analysis. In these analyses, phenotype data from subjects with less than 10 pack-years of smoking were changed to missing, but genotype data were included from all subjects. To assess the statistical significance of the two point linkage analysis results, 3000 replicates of the 72 pedigrees were made with the SIMULATE program, and the replicates were analyzed for linkage using the MSIM program of SLINK (13–15).

The complete results of the linkage analyses performed in this study can be found at the University of Arizona/Brigham and Women’s Hospital Program in Genomic Applications website (http://innateimmunity.net).

ACKNOWLEDGEMENTS
We appreciate assistance from a variety of physicians in recruiting subjects for this study. We are grateful for the high quality genotyping performed by the NHLBI Mammalian Genotyping Service. We are especially thankful for the enthusiastic quality genotyping performed by the NHLBI Mammalian Genotyping Service. We are grateful for the high quality genotyping performed by the NHLBI Mammalian Genotyping Service. We are especially thankful for the enthusiastic quality genotyping performed by the NHLBI Mammalian Genotyping Service. We are grateful for the high quality genotyping performed by the NHLBI Mammalian Genotyping Service. We are especially thankful for the enthusiastic quality genotyping performed by the NHLBI Mammalian Genotyping Service. We are grateful for the high quality genotyping performed by the NHLBI Mammalian Genotyping Service. We are especially thankful for the enthusiastic quality genotyping performed by the NHLBI Mammalian Genotyping Service.

REFERENCES


