Genetics of COPD

Hidetoshi Nakamura

ABSTRACT
Previous family studies suggested that genetic variation contributes to COPD susceptibility. The only gene proven to influence COPD susceptibility is SERPINA1, encoding α1-antitrypsin. Most studies on COPD candidate genes except SERPINA1, have not been consistently replicated. However, longitudinal studies of decline in lung function, meta-analyses of candidate gene studies, and family-based linkage analyses suggested that variants in EPHX1, GST, MMP12, TGFβ1, and SERPINE2 were associated with susceptibility to COPD. A genome-wide association (GWA) study has recently demonstrated that CHRNA3/5 in 15q25 was associated with COPD compared with control smokers. It was of interest that the CHRNA3/5 locus was associated with nicotine dependence and lung cancer as well. The associations of HHIP on 4q31 and FAM13A on 4q22 with COPD were also suggested in GWA studies. Another GWA study has shown that BICD1 in 12p11 was associated with the presence or absence of emphysema. Although every genetic study on COPD has some limitations including heterogeneity in smoking behaviors and comorbidities, it has contributed to the progress in elucidating the pathogenesis of COPD. Future studies will make us understand the mechanisms underlying the polygenic disease, leading to the development of a specific treatment for each phenotype.

KEY WORDS
COPD, emphysema, FEV1, genome-wide association study, linkage analysis

INTRODUCTION
Cigarette smoking is the primary risk factor for the development of chronic obstructive pulmonary disease (COPD). However, it is thought that only 15-20% of smokers develop COPD. Previous family studies suggested that genetic variation contributes to COPD susceptibility. The only gene definitely proven to influence COPD susceptibility is SERPINA1, encoding α1-antitrypsin.1 There have been many studies on COPD candidate genes other than SERPINA1, but most of them have not been consistently replicated. This may be attributable to heterogeneity of the study populations and COPD phenotypes. As hypotheses, different genes may contribute to mild, moderate, or severe airflow limitation. Emphysematous changes and airway wall thickness may be influenced by distinct gene polymorphisms. Races and gender2 are also known to effect on associations between genetic variation and COPD phenotypes. Moreover, smoking habit is significantly regulated by genetic factors,3 which can modify the outcomes of genetic studies on COPD. COPD usually develops in elderly patients with other coexistent disorders such as lung cancer and cardiovascular diseases. These disorders are also promoted by smoking and influenced by distinctive genetic factors. All these findings, not always stratified in previous studies, lead to limitations of candidate gene approaches on COPD. In this article, genetic studies of COPD including candidate gene studies, family-based linkage analyses, and genome-wide association studies (GWAS) will be reviewed. The author believes that genetic studies on COPD contribute to understanding the pathogenesis of the polygenic disease with multiple phenotypes despite the difficulties described above.

CANDIDATE GENE STUDIES
α1-antitrypsin is a major circulating inhibitor of serine proteases. Severe α1-antitrypsin deficiency usually results from a PiZZ or PiZnull genotype. This rare recessive trait is commonly seen in individuals of Northern European origin. About 1-2% of COPD patients were reported to be α1-antitrypsin deficiency in the United States while there are only about 20 pedigrees with Siyama genotype in Japan.4 Development of panlobular emphysema and decline in lung function occur in both smokers and nonsmokers with the...
Health Study to examine gene variants associated with different ethnic groups. The genetic factors to COPD susceptibility between different exposures and the differential contribution of these genetic factors to COPD would be illustrated as a result of the volume and heterogeneity of most of these findings have not been consistently replicated in the researches.

Since α1-antitrypsin deficiency was reported to cause COPD, a lot of candidate gene association studies have been published. Candidate genes have been selected in terms of putative pathogenesis of COPD such as the protease-antiprotease imbalance, the oxidative stress and antioxidants, and cytokines and chemokines related to airway inflammation. Over 100 published COPD candidate gene studies suggested each genetic variant’s contribution to one of COPD phenotypes in the reported ethnic groups. However, most of these findings have not been consistently replicated as a result of the volume and heterogeneity of the researches.

Sandford et al. have taken advantage of the Lung Health Study to examine gene variants associated with rapid decline in lung function. They selected 300 continued smokers with the most rapid decline and 300 with the least rapid decline in FEV1 among the approximately 6,000 individuals. They have tested over 50 genes and those with altered risk for decline in FEV1 are presented in Table 1. Other investigators have studied the same (microsomal epoxide hydrolase, glutathione S-transferase, heme oxygenase-1) or additional genes (MMP-12, TAM, 5-HT1A, 5-HT2C, ADAM33, glutamate cysteine ligase, CYP3A5) in different cohorts of smokers in which longitudinal lung function data are available (Table 1). Hunninghake et al. have recently tested for an association between single nucleotide polymorphisms (SNPs) in the MMP-12 gene and lung function in more than 8,300 subjects in seven cohorts. The minor allele of a SNP in MMP-12 was associated with a positive effect on lung function in children with asthma and in adults who smoke, and was associated with a reduced risk of COPD in adult smokers during over 30 years of follow-up. Association studies of genetic variants with decline in lung function have reasonably suggested the roles of several molecules related to proteases-antiproteases imbalance, oxidative stress, and cytokine-mediated inflammation in the development of airflow limitation.

Castaldi et al. have recently conducted a systematic review and meta-analysis of all population-based, case-control candidate gene COPD studies before 16 July 2008. They identified 27 genetic variants with adequate data for quantitative meta-analysis. Of these variants, four (GST, TGFβ1, TNF, SOD3) were significantly associated with COPD susceptibility (Table 2). SOD3 encodes extracellular superoxide dismutase, the most important extracellular scavenger of superoxide. Although only two studies about the SOD3 variant were included in the meta-analysis, Juul et al. included over 9,000 individuals (978 COPD

<table>
<thead>
<tr>
<th>Candidate Gene (Gene Symbol)</th>
<th>Polymorphism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>α1-antitrypsin (SERPINA1)</td>
<td>MZ</td>
<td>6</td>
</tr>
<tr>
<td>MMP-1 (MMP1)</td>
<td>-1607 + G</td>
<td>8</td>
</tr>
<tr>
<td>MMP-12 (MMP12)</td>
<td>-82A/G</td>
<td>9</td>
</tr>
<tr>
<td>ADAM 33 (ADAM33)</td>
<td>F + 1, S_1, etc</td>
<td>10</td>
</tr>
<tr>
<td>microsomal epoxide hydrolase (EPHX1)</td>
<td>Tyr113 &gt; His, Arg139 &gt; His</td>
<td>6, 12</td>
</tr>
<tr>
<td>glutathione S-transferase (GST)</td>
<td>GSTP1 Ile105 &gt; Val</td>
<td>11, 19</td>
</tr>
<tr>
<td>heme oxygenase-1 (HMOX1)</td>
<td>L-allele (≥33 GT repeats)</td>
<td>12, 20</td>
</tr>
<tr>
<td>glutamate cysteine ligase, catalytic subunit (GCLC)</td>
<td>-129C/T</td>
<td>13</td>
</tr>
<tr>
<td>cytochrome P450 3A5 (CYP3A5)</td>
<td>*1/*3 vs *3/*3</td>
<td>14</td>
</tr>
<tr>
<td>interleukin-1β (IL1B), interleukin-1 receptor antagonist (IL1RN)</td>
<td>IL1B/IL1RN haplotypes</td>
<td>15</td>
</tr>
<tr>
<td>interleukin-4 receptor α (IL4RA)</td>
<td>551RR</td>
<td>16</td>
</tr>
<tr>
<td>interleukin-6 (IL6)</td>
<td>-174G/C</td>
<td>17</td>
</tr>
<tr>
<td>β2-adrenergic receptor (ADRB2)</td>
<td>Glu27 &gt; Gln</td>
<td>18</td>
</tr>
</tbody>
</table>

ADAM, a disintegrin and metalloprotease.

Table 1 Significant genetic associations with decline in lung function

<table>
<thead>
<tr>
<th>Candidate Gene (Gene Symbol)</th>
<th>Polymorphism</th>
<th>OR (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>glutathione S-transferase (GST)</td>
<td>GSTM1 null variant</td>
<td>1.45 (1.09-1.92)</td>
</tr>
<tr>
<td>TGF-β1 (TGFβ1)</td>
<td>+297/T/C</td>
<td>0.73 (0.64-0.83)</td>
</tr>
<tr>
<td>TNF-α (TNF)</td>
<td>-308G/A</td>
<td>1.19 (1.01-1.40)</td>
</tr>
<tr>
<td>super oxide dismutase-3 (SOD3)</td>
<td>R213G</td>
<td>1.97 (1.24-3.13)</td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval.

severe deficiency, although smoking increases the risk accordingly. α1-antitrypsin deficiency accounts for only a small part of COPD in the world, but it illustrates the interaction between genes and environmental exposures and the differential contribution of genetic factors to COPD susceptibility between different ethnic groups.

Since α1-antitrypsin deficiency was reported to cause COPD, a lot of candidate gene association studies have been published. Candidate genes have been selected in terms of putative pathogenesis of COPD such as the protease-antiprotease imbalance, the oxidative stress and antioxidants, and cytokines and chemokines related to airway inflammation. Over 100 published COPD candidate gene studies suggested each genetic variant’s contribution to one of COPD phenotypes in the reported ethnic groups. However, most of these findings have not been consistently replicated as a result of the volume and heterogeneity of the researches.

Sandford et al. have taken advantage of the Lung Health Study to examine gene variants associated with COPD susceptibility by a systematic review and meta-analysis of case-control candidate gene studies before 16 July 2008. They identified 27 genetic variants with adequate data for quantitative meta-analysis. Of these variants, four (GST, TGFβ1, TNF, SOD3) were significantly associated with COPD susceptibility. Of these variants, four (GST, TGFβ1, TNF, SOD3) were significantly associated with COPD susceptibility. Of these variants, four (GST, TGFβ1, TNF, SOD3) were significantly associated with COPD susceptibility. Of these variants, four (GST, TGFβ1, TNF, SOD3) were significantly associated with COPD susceptibility. Of these variants, four (GST, TGFβ1, TNF, SOD3) were significantly associated with COPD susceptibility. Of these variants, four (GST, TGFβ1, TNF, SOD3) were significantly associated with COPD susceptibility.

Table 2 Significant genetic associations with COPD susceptibility by a systematic review and meta-analysis of case-control candidate gene studies

<table>
<thead>
<tr>
<th>Candidate Gene (Gene Symbol)</th>
<th>Polymorphism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>glutathione S-transferase (GST)</td>
<td>GSTM1 null variant</td>
<td>1.45 (1.09-1.92)</td>
</tr>
<tr>
<td>TGF-β1 (TGFβ1)</td>
<td>+297/T/C</td>
<td>0.73 (0.64-0.83)</td>
</tr>
<tr>
<td>TNF-α (TNF)</td>
<td>-308G/A</td>
<td>1.19 (1.01-1.40)</td>
</tr>
<tr>
<td>super oxide dismutase-3 (SOD3)</td>
<td>R213G</td>
<td>1.97 (1.24-3.13)</td>
</tr>
</tbody>
</table>
cases and 7604 controls).22

**LINKAGE ANALYSES**

Family-based linkage analyses have brought us a great advance in this field. Although COPD patients are usually aged and linkage studies are basically difficult, the Boston early-onset COPD study recruited relatively young (<52 years) and severe patients (FEV1 ≤40% predicted) and their relatives (72 pedigrees and 560 family members).23 The study has suggested the association of SERPINE2 in the chromosome 2q with FEV1/FVC24 and TGFB1 in the chromosome 19q with pre-bronchodilator FEV125 (Table 3). SERPINE2 encodes a 44-kDa cellular and extracellular matrix-associated serine protease inhibitor, mainly involved in coagulation and fibrinolysis. A mechanism by which SERPINE2 may contribute to the development of COPD has yet to be identified. Framingham study has suggested that FEV1 was linked to SMO2 in the chromosome 6q27 in a general population.26,27 SMO2 encodes secreted modular calcium-binding protein 2, which has homology to the protease inhibitor α1-antitrypsin and contains multiple domains with the potential to act as protease inhibitors. These observations obtained from linkage analyses are thought to be more reliable for the detection of susceptibility genes for COPD than those in case-control candidate gene studies. However, differences in study populations and target phenotypes should be noted in these analyses as well. It was not surprising that the genes identified in the Boston early-onset COPD study were different from those in the Framingham study since the former represented one of the severest subpopulations of COPD while the latter might represent subjects with relative airflow limitations in the general population. In general it is difficult to find disease susceptibility genes by linkage analyses because of relatively small numbers of microsatellite markers (300-400).

**GENOME-WIDE ASSOCIATION STUDIES**

GWAS require hundreds or thousands of participants genotyped in 300,000-500,000 SNP markers. Recent advances in genetical analyses have enabled us to conduct GWAS in COPD. Pillai et al. have demonstrated that 2 SNPs at CHRNA3/5 in 15q25 were associated with COPD (n = 823) when compared with control smokers (n = 810).28 This finding was replicated by the other case-control populations and family-based association analyses. CHRNA3/5 encodes nicotinic cholinergic receptor alpha 3, 5, known to be associated with nicotine addiction and lung cancer.29,30 The association of HHIP locus on 4q31 was also consistently replicated but did not reach genome-wide significance levels. HHIP encodes hedgehog interacting protein, belonging to the hedgehog gene family that play an important role in regulating morphogenesis and lung development.31 Cho et al. have reported that variants in FAM13A are associated with COPD by increasing the numbers of participants (2940 cases and 1380 smoking controls).32 FAM13A encodes family with sequence similarity 13, member A. While little is known about FAM13A function, gene expression analyses in cell lines have demonstrated a consistent increase in response to hypoxia. Differences in respiratory epithelial cell expression of FAM13A have been noted during differentiation into pulmonary type II cells in vitro.33 Kong et al. has reported that a SNP in BICD1 was associated with the presence or absence of emphysema in GWAS. They divided 2380 COPD patients from the Norway, ECLIPSE, and NETT into those with or without emphysema by chest-CT analyses (Table 4).34 BICD1 encodes bicaudal D homolog 1 (Drosophila). BICD protein forms a complex with dynactin and plays a role in mediating dynein function. SNPs in BICD1 gene have been associated with telomere length in leukocytes.34 Association among variants in BICD1, telomere length, and emphysema is expected to be investigated.

More genes including HHIP and FAM13A have been reported to be associated with lung function in GWAS in general populations (Table 5).35,36 PTCH1 is known to encode a membrane receptor of hedgehog protein including HHIP. Thus the variants in HHIP and PTCH1 may influence lung development possibly related to COPD phenotypes.37 In contrast AGER is highly expressed in the lung and its absence contributes to the pathogenesis of idiopathic pulmonary fibrosis. Since AGER signaling is involved in host defense, inflammation, and tissue remodeling,
variants in AGER may be relevant to the development of COPD. Further information on molecular functions of the candidate genes associated with lung function is available in the references. GWAS can theoretically pick up phenotype-specific gene polymorphisms though they still have similar limitations in selecting COPD and control populations to candidate gene case-control studies. In addition the commonly used multistage GWA approach, using very stringent P values in the identification cohort, might result in missing relevant findings that do not show genome-wide significance.

**GENETICS OF SMOKING BEHAVIORS**

A number of linkage analyses have suggested that the specific loci influence smoking behaviors including age at initiation, intensity (number of cigarettes smoked per day), and cessation. GWAS of cigarette smoking behaviors have suggested that CHRNA3/5 in 15q25 was associated with smoking intensity and nicotine dependence. This locus was identified not only as that for COPD but also for lung cancer by GWAS although it is unclear whether the effects of these variants on susceptibility to COPD and lung cancer are due to their influences on smoking behaviors. A genotype-environmental interaction is defined as a non-additive contribution of genetic and environmental factors to the expression of a phenotype. Therefore independent contributions of genetic factors and smoking to COPD are very difficult to be evaluated since the environmental factors can be regulated by the same and/or different genetic factors as/from those for COPD.

Minematsu et al. have reported that the deletion polymorphism of CYP2A6, a major nicotine metabolizing enzyme, was associated with smoking intensity and development of emphysema after normalizing lifelong cigarette consumption (Table 6). Defective alleles of CYP2A6 were also reported to reduce the risk of lung cancer partly independent of their effects on smoking behaviors. These observations are thought to be related to its function to activate procarcinogens including nitrosamines. Ishii et al. have recently reported the association of a SNP in SLC6A4, encoding 5-hydroxytryptamine transporter, related to nicotine dependence with COPD which partly depends on tobacco consumption. They also suggested the association of another SNP with a depression score. The numbers of participants of these Japanese studies are relatively small, but they are of interest that susceptibility to COPD is potentially influenced by genetic factors related to smoking habits other than CHRNA3/5. Although COPD is known to develop only in part of smokers, it is also true that heavy smokers tend to become COPD. Therefore the factors that increase smoking intensity may overlap those for COPD. Further systematic genetic studies will be necessary to elucidate the relationship between genetic variants related to smoking behaviors and COPD.

**FUTURE PERSPECTIVES**

Genetic studies of COPD had started as candidate gene case-control studies. After a couple of family-based linkage analyses were reported, GWAS have recently been applied to identify novel genetic variants associated with COPD. GWAS are hypothesis free, and functional assessment of identified genetic variants is inevitable. Even in GWAS, heterogeneity of study populations is the biggest problem that can result in distinct outcomes. Environmental factors including smoking status and air pollution should carefully be taken into account. As a target phenotype, emphysema has already been used in a GWA study, but criteria to judge the presence or absence of emphysema on chest-CT analysis are not yet established. Since perfect matching of COPD cases and control smokers populations is impossible, other genomic approaches should be considered as well. DeMeo et al. have utilized genomic regions from 56 lung-tissue gene-expression microarrays to select 889 SNPs to be tested for association with COPD, and concluded that IREB2 near the CHRNA3/5 locus was associated with COPD. IREB2 encodes iron-responsive element binding protein 2 that is an RNA binding protein involved in maintaining human cellular iron metabolism. IREB2 mRNA expression was increased in emphysema and risk of COPD partly independent of their effects on smoking behaviors.
increased in lung tissues from COPD patients than those from normal subjects. Whole-genome gene expression studies are useful in understanding the pathogenesis of COPD and selecting candidate genes for genetical studies.

Since most COPD patients are elderly with comorbidities such as asthma, lung cancer, hypertension, cardiovascular diseases, diabetes mellitus, etc., it will be important to stratify these coexistent diseases in future genetical studies of COPD. Regan et al. have stated that they will enroll 10,000 smokers between ages of 45 and 80 years and exclusion criteria are a history of other lung disease except asthma (e.g. pulmonary fibrosis, extensive bronchiectasis, cystic fibrosis), active cancer under treatment, lung cancer, myocardial infarction, other cardiac hospitalization, etc. in the Genetic Epidemiology of COPD (COPDGene) study design. These criteria may be reasonable but it is difficult to find best criteria to analyze genetic factors of COPD patients with comorbidities. Race and gender have been clearly defined in the COPDGene study. To define the phenotypes CT-images will be obtained both on inspiratory and expiratory phases in addition to spirometry in this study.

Recent studies have suggested novel mechanisms underlying the pathogenesis of COPD such as apoptosis and senescence of lung cells, activated autoimmune processes, chronic airway infection, epigenetic regulation such as modifications of histone deacetylases (HDACs), and impaired lung development. All these mechanisms will provide us information about additional candidate gene approaches and functional analyses of novel gene variants. Genetic studies of COPD have confirmed the hypotheses of the disease mechanisms such as protease-antiprotease imbalance and excessive oxidative stress, and will further make sure the novel mechanisms described above. One of the most important purposes of genetical studies of COPD is a clinical application of pharmacogenetics. However, few COPD pharmacogenetics studies have been performed. Most studies have examined the role of variants in the β2-adrenergic receptor gene in bronchodilator response, but the findings were inconclusive. In future, personalized treatment of specific phenotypes of COPD based on pharmacogenetics studies will be expected including existing COPD therapies such as long-acting β-agonists and muscarinic antagonists, and smoking cessation therapy.

Although every genetic study on COPD has some limitations and should be carefully interpreted, we believe that it has greatly contributed to the progress in elucidating the pathogenesis of COPD. Future studies will bring us insight into mechanisms underlying various phenotypes of COPD including emphysema, pulmonary hypertension, and mucus secretion, leading to the development of a specific treatment for each disease process.

REFERENCES

18. Joos L, Weir TD, Connnett JE et al. Polymorphisms in the β2 adrenergic receptor and bronchodilator response, bronchial hyperresponsiveness, and rate of decline in


