Chronic obstructive pulmonary disease (COPD) is defined as airflow obstruction that is not fully reversible. It results from abnormal inflammation following exposure to noxious particles or gases. This is typically exposure to cigarette smoke but may also include exposure to biomass fuels and some industrial dusts. COPD clusters within families, suggesting that heritable factors play a role in the pathogenesis of this disease. The only genetic factor that is widely accepted to be associated with COPD is severe deficiency of Α₂-antitrypsin. This depletes an important member of the antiproteinase screen and causes excessive intrapulmonary inflammation. Indeed, even asymptomatic non-smoking heterozygotes for the Z allele (PiMZ) without airflow obstruction have increased intrapulmonary inflammation.

Severe Α₂-antitrypsin deficiency is found in only 1–2% of all individuals with COPD. There is growing interest in the genetic factors that predispose to COPD in those individuals who do not have Α₂-antitrypsin deficiency. Early papers focused on association studies that looked at genetic variation within individuals with COPD compared with those individuals who were matched for all factors that are associated with COPD (most importantly age, smoking history and ethnic background) but who did not have airflow obstruction. The studies were typically small (100–150 cases and controls) and were often confounded by failure to match cases and controls carefully. There was often the issue of correcting for multiple comparisons as well as the inherent complexity of the COPD phenotype. Larger family-based studies have shown the independent clustering of the airway disease and emphysema component of COPD within families. This suggests that different genetic factors predispose to each of these components of the phenotype. The only way of overcoming the inherent variation in COPD is to focus on groups of well-characterised individuals with components of the COPD phenotype or to undertake studies with large sample sizes and then to replicate any positive findings in other cohorts. This is now the case with candidate gene studies and there is good evidence that heterozygotes for Α₂-antitrypsin deficiency (PiMZ) and polymorphisms in genes involved in oxidative stress (microsomal epoxide hydrolase, glutathione S-transferase, haeme oxygenase and superoxide dismutase 3) are associated with an increased risk of COPD. Furthermore, polymorphisms in transforming growth factor-beta, a protein important in maintaining the extracellular matrix, are also likely to be important in the pathogenesis of this disease. More recently, a minor allele of a single nucleotide polymorphism (SNP) in matrix metalloproteinase-12 (MMP-12) has been shown to protect against COPD in adult smokers.

The limitation of association studies is that genes are assessed in pathways that are already recognised to be associated with COPD—the proteinase–antiproteinase and oxidative stress pathways—and genes that maintain the integrity of the extracellular matrix. In recent years the collection of large cohorts has allowed unbiased genome-wide association studies (GWAS) in individuals with COPD. The largest study was undertaken in a cohort from Bergen, Norway, and then replicated in the International COPD Genetics Network, the National Emphysema Treatment Trial with controls from the Normative Ageing Study and then finally in the Boston Early Onset COPD cohort. Top hits from this analysis were SNPs in the alpha-nicotinic acetylcholine receptor and the hedgehog interacting protein (HHIP). The SNP in the α-nicotinic acetylcholine receptor (rs8034919) was also identified in three GWAS of lung cancer and is thought to be important in peripheral vascular disease and nicotine addiction. It is likely that this SNP functions as a marker for an addiction gene. Individuals who carry this SNP may require more cigarettes to satisfy nicotine addiction, may inhale more deeply and may find it more difficult to withdraw from cigarette smoking. Alternatively, the association may result from linkage disequilibrium with SNPs in the iron responsive element binding protein 2 (IREB2). This was identified by expression analysis in lung tissue from individuals with COPD and then confirmed in three separate COPD cohorts. IREB2 is localised to the human epithelial cell surface and may play a role in protecting against epithelial damage from oxidative stress. Finally, the analysis of even larger numbers of individuals with COPD has identified a SNP in FAM13A as being associated with COPD. The role of this gene in the disease is unclear but expression has been associated with hypoxia. FAM13A has also been associated with lung function in a second independent study. A detailed analysis of these genes in well-characterised cohorts showed that SNPs in the α-nicotinic acetylcholine receptor are associated with smoking intensity, airflow obstruction and emphysema. SNPs in the hedgehog interacting protein are associated with systemic features of COPD (low body mass index) and exacerbations, while SNPs in FAM13A are associated with airflow obstruction.

Taken together, there is clear evidence that genes associated with the protease–antiprotease pathway, oxidative stress and the integrity of the extracellular matrix are involved in the susceptibility of smokers to COPD. GWAS have identified genes that are likely to have a role in addiction and other genes that may protect against oxidative stress. It is clear that the gene mining has only just begun, with good prospects for identifying further novel genes associated with COPD. This is important as they will allow us to stratify individuals with COPD and identify pathways that provide new insights into the mechanisms of disease. Clearly, the long-term aim is to intervene in these pathways and so prevent the relentless progression of the pulmonary and systemic manifestations of COPD. It is clear that further collaboration is required to bring together even larger numbers of well-phenotyped cohorts to identify more genes that are associated with COPD. The potential value of this approach has been demonstrated in two large meta-analyses examining lung function in general population samples. These studies, run by the Spirometa and CHARGE consortia, identified a number of novel genes which predict forced
expiratory volume in 1 s (FEV₁) and/or the ratio of FEV₁ to forced vital capacity. Four genes identified were common to both studies - the previously observed HHIP, GSTCD (glutathione S transferase C-terminal domain), HTR4 (the serotonin 4 receptor subtype) and AGER (the receptor for advanced glycosylation end products). While these studies concentrated on general population samples, it is highly likely that genetic factors predicting FEV₁ will be strong risk factors for COPD and other diseases where airflow obstruction is present; analyses addressing this issue are ongoing. The populations studied in each consortium included primary discovery samples of >20,000 in each. It seems to us therefore that the future for COPD genetics is very bright, but only if we think big!

There have been many more studies addressing the genetics of asthma than have been performed examining COPD, although the overall conclusion that very large studies are required to identify reproducible effects is the same when considering asthma. Initial attempts at hunting an asthma gene used predominantly family-based linkage approaches, which resulted in multiple loci on almost every chromosome being identified in different populations. Candidate gene studies have also been widely performed but, as with COPD, the ability to replicate findings has been very limited. The genes which have attracted the most interest from these studies include ADAM33, PHF11, CPRA and DPP10. While there have been a number of studies which have had replicated initial associations, the effect size of individual SNPs for asthma risk has generally been small and functional mutations remain to be identified. For example, in a study which involved genotyping the complete 1958 birth cohort for risk SNPs in these genes prioritised on the basis of linkage data, the cohort for risk SNPs in these genes genotyping the complete 1958 birth cohort for risk SNPs in these genes was about 20,000 in each. It seemed to us therefore that the future for COPD genetics is very bright, but only if we think big!

We think big!

There have been many more studies addressing the genetics of asthma than have been performed examining COPD, although the overall conclusion that very large studies are required to identify reproducible effects is the same when considering asthma. Initial attempts at hunting an asthma gene used predominantly family-based linkage approaches, which resulted in multiple loci on almost every chromosome being identified in different populations. Candidate gene studies have also been widely performed but, as with COPD, the ability to replicate findings has been very limited. The genes which have attracted the most interest from these studies include ADAM33, PHF11, CPRA and DPP10. While there have been a number of studies which have had replicated initial associations, the effect size of individual SNPs for asthma risk has generally been small and functional mutations remain to be identified. For example, in a study which involved genotyping the complete 1958 birth cohort for risk SNPs in these genes prioritised on the basis of linkage data, the cohort for risk SNPs in these genes was about 20,000 in each. It seemed to us therefore that the future for COPD genetics is very bright, but only if we think big!

We think big!

There have been many more studies addressing the genetics of asthma than have been performed examining COPD, although the overall conclusion that very large studies are required to identify reproducible effects is the same when considering asthma. Initial attempts at hunting an asthma gene used predominantly family-based linkage approaches, which resulted in multiple loci on almost every chromosome being identified in different populations. Candidate gene studies have also been widely performed but, as with COPD, the ability to replicate findings has been very limited. The genes which have attracted the most interest from these studies include ADAM33, PHF11, CPRA and DPP10. While there have been a number of studies which have had replicated initial associations, the effect size of individual SNPs for asthma risk has generally been small and functional mutations remain to be identified. For example, in a study which involved genotyping the complete 1958 birth cohort for risk SNPs in these genes prioritised on the basis of linkage data, the cohort for risk SNPs in these genes was about 20,000 in each. It seemed to us therefore that the future for COPD genetics is very bright, but only if we think big!

We think big!

There have been many more studies addressing the genetics of asthma than have been performed examining COPD, although the overall conclusion that very large studies are required to identify reproducible effects is the same when considering asthma. Initial attempts at hunting an asthma gene used predominantly family-based linkage approaches, which resulted in multiple loci on almost every chromosome being identified in different populations. Candidate gene studies have also been widely performed but, as with COPD, the ability to replicate findings has been very limited. The genes which have attracted the most interest from these studies include ADAM33, PHF11, CPRA and DPP10. While there have been a number of studies which have had replicated initial associations, the effect size of individual SNPs for asthma risk has generally been small and functional mutations remain to be identified. For example, in a study which involved genotyping the complete 1958 birth cohort for risk SNPs in these genes prioritised on the basis of linkage data, the cohort for risk SNPs in these genes was about 20,000 in each. It seemed to us therefore that the future for COPD genetics is very bright, but only if we think big!

We think big!

There have been many more studies addressing the genetics of asthma than have been performed examining COPD, although the overall conclusion that very large studies are required to identify reproducible effects is the same when considering asthma. Initial attempts at hunting an asthma gene used predominantly family-based linkage approaches, which resulted in multiple loci on almost every chromosome being identified in different populations. Candidate gene studies have also been widely performed but, as with COPD, the ability to replicate findings has been very limited. The genes which have attracted the most interest from these studies include ADAM33, PHF11, CPRA and DPP10. While there have been a number of studies which have had replicated initial associations, the effect size of individual SNPs for asthma risk has generally been small and functional mutations remain to be identified. For example, in a study which involved genotyping the complete 1958 birth cohort for risk SNPs in these genes prioritised on the basis of linkage data, the cohort for risk SNPs in these genes was about 20,000 in each. It seemed to us therefore that the future for COPD genetics is very bright, but only if we think big!

We think big!