What can naturally occurring mutations tell us about the pathogenesis of COPD?

S J Marciniak, D A Lomas

ABSTRACT

The airway and emphysema phenotypes of chronic obstructive pulmonary disease (COPD) cluster in susceptible families. Some of this clustering is likely to be the result of shared genetic factors. There are a number of relatively rare syndromes that predispose to COPD, and a growing number of association and linkage studies that have assessed the genetic factors that predispose smokers to airflow obstruction. A review of the literature was performed to determine what has been learnt about the factors and pathways that predispose some individuals to progressive airflow obstruction while others are relatively resistant. There is very strong evidence that the integration of the extracellular matrix, and in particular the elastin fibres, are important in the pathogenesis of COPD. There is also support for the role of proteinase-antiproteinase imbalance and probably oxidative stress. However, many of the genetic studies have focused on pathways that have already been implicated in disease. It is now essential that future studies take advantage of multiple large cohorts that have been phenotyped for both airways disease and emphysema, and use modern technology to perform unbiased genome-wide analyses of genetic variation in COPD. Only then will we have new insights into the pathways that underlie this common condition.

Human diseases can be deceptively complex. The pathogenesis of some of the commonest conditions remains obscure. An excellent example is chronic obstructive pulmonary disease (COPD); while we know that COPD is caused by inhaled smoke, remarkably few smokers develop airflow obstruction. It is hoped that, by understanding the genetics that underlie this disease, we may be able to tease apart its basic biology and identify new therapeutic avenues. Familial clustering suggests that heritable factors may play a role in the development of COPD. Indeed, a recent study has shown that both the airway and emphysema components cluster independently, suggesting that different genetic factors play a role in the development of these two components of the disease. However, apart from the relatively uncommon $z_1$-antitrypsin mutations which account for 1–2% of cases, the additional genetic factors remain obscure.

In this review we examine the current literature to identify likely candidate genes in order to understand better the pathogenesis of this important disease. We set out by searching the PubMed database with the terms “(COPD OR emphysema) AND (genetic OR mutation OR polymorphism)” on 10 February 2008 and obtained 1071 hits. These titles and abstracts were screened and relevant papers were obtained for detailed assessment. Their bibliographies were also screened for additional sources. Where possible, we have given special emphasis to studies in which COPD was diagnosed using standardised spirometric criteria; cases and control groups were appropriately matched for both smoking and ethnicity and where alleles were shown to be in Hardy-Weinberg equilibrium (HWE) in the control group.

In genetic association studies it is common to consider a statistical significance level of $<0.05$ to be “noteworthy”, but one should be wary. It has been shown that many genetic association studies report false positive findings owing to a failure to appreciate the prior probability of an association and the power of the study to detect a meaningful effect. When the prior probability of an association is low—that is to say, when there is little functional or epidemiological data to support an association—the numbers of subjects required to guard against a false positive result increases. Consequently, the identification of a genetic association in a single study must always be treated with caution and in our review we have concentrated on replicated results (table 1).

EXTRACELLULAR MATRIX

Alveolar tissue consists of epithelial cells, capillaries and extracellular matrix (ECM), the latter comprising a complex network of scaffolding proteins, principally elastin and collagen. The elastin filaments form from tropoelastin monomers that self-assemble into aggregates and then fuse with microfilaments. Multiple covalent cross-links between the lysines in neighbouring filaments provide stability. Cutis laxa is a family of autosomal dominant (OMIM #123700), X-linked (OMIM #304150) and recessive (OMIM #219100, 219200) human diseases characterised by excessively slack connective tissues. Several families with the milder autosomal dominant form show early-onset pulmonary pathology including emphysema, particularly if inherited with the $Z$ allele of $z_1$-antitrypsin. Two groups independently identified separate mutations within the $ELN$ (elastin) gene that cause mild cutis laxa and early onset COPD. The $ELN$ gene maps to 7q11.23 in man but, as chromosome 7 has not been identified in linkage analysis as a site associated with COPD, it is likely that $ELN$ mutations are a rare cause of this disease.

Elastin fibres bind other proteins including fibulins which, in turn, bind multiple ECM components and the basement membrane. The fibulins are a family of six proteins, at least two of which are mutated in severe autosomal recessive...
forms of cutis laxa and whose phenotype often includes early-onset emphysema. A novel mutation in the fibrillin-4 gene (FBLN4; 11q13) was recently identified in autosomal recessive cutis laxa with developmental emphysema. The mutation caused an amino acid substitution in an epidermal growth factor (EGF)-like domain of fibrillin-4, leading to very low levels of extracellular protein. In a consanguineous Turkish family, a mutation in fibrillin, which bears homology to the fibulins, cause Marfan’s syndrome. Moreover, mutations of fibrillin (FBLN5; 14q32.1) was also found to cause cutis laxa and emphysema complicated by recurrent pulmonary infections. Once again, the mutation was located within an EGF-like domain, suggesting these are critical for fibrulins to maintain the integrity of the ECM within the lung. Interestingly, analogous mutations in fibrillin, which bears homology to the fibrulins, cause Marfan’s syndrome. Moreover, mutations of fibrillin (FBN1; 15q21.1) have been described in neonatal Marfan’s syndrome with very early-onset emphysema.

Menkes disease (OMIM #309400), characterised by abnormal hair and dysmorphic features, is caused by mutations in an intracellular copper transporter (ATP7A; Xq13.3). The clinical features are due to defective connective tissue synthesis believed to be the result of dysfunction of lysyl oxidase. This copper-dependent enzyme is required for proper cross-linking of both collagen and elastin fibres. A recent case report described a child with Menkes disease and severe bilateral pan-lobular emphysema who died aged only 14 months. Gene sequencing revealed a splice-site mutation in ATP7A suggesting that proper ECM cross-linking is vital for stability of the lung parenchyma.

In contrast to animal models of COPD, mutations in collagen have not been identified in humans. This does not appear to be due to an incompatibility of mutated collagen with survival as numerous collagen mutations have been described that cause other human diseases. Instead, it may reflect a more important role for elastin integrity in emphysema in humans than in mice. However, aberrant collagen synthesis has been implicated in COPD. The signalling molecule transforming growth factor β1 (TGFB1) enhances collagen synthesis in vivo, and polymorphisms in its gene (TGFB1; 19q13.1) have been associated with COPD, although a recent large study found no association between TGFB1 polymorphisms and the rate of lung function decline in smokers. Intriguingly, the TGFB1 gene maps to a locus on chromosome 19 which has high linkage (LOD 3.3) with forced expiratory volume in 1 s (FEV1) in smokers. However, as is frequently the case with polymorphism studies, the literature is unclear. For example, two polymorphisms (SNPs)—rs1800469 and rs1982073—were found to be independently associated with COPD in two studies but, in another, they were only significant when analysed as part of a haplotype (combination of alleles) while yet another SNP, rs6957, was significant in its own right. Detailed analysis of the Boston Early-Onset COPD Study data revealed further complexity. While some alleles of TGFB1 were associated with FEV1 (rs2241712, rs2241718, rs6957), there was a separate but partially overlapping set of alleles associated with airflow obstruction (rs2241712, rs1800469, rs1982073). TGFB1 protein is inactive when first secreted owing to the presence of an inhibitory N-terminal pro-peptide. It is secreted in association with latent TGFB1 binding proteins (LTBP) which share structural features with fibrulins and are assembled into the ECM. Mice with mutations in LTBP4 develop severe emphysema. Intriguingly, the sole study that has addressed LTBP4 (19q13.1—q13.2) polymorphisms found an association with COPD in man.

Taken together, these studies provide strong evidence in support of a crucial role for the loss of ECM integrity—in
particular the elastic components—in the development of COPD. It is therefore important to consider the enzymes implicated in degradation of the ECM.

PROTEASE-ANTIPROTEASE BALANCE

The protease-antiprotease theory has its roots in the observation that individuals with $\alpha_1$-antitrypsin are particularly susceptible to COPD and in experimental models of emphysema from the 1960s. This theory suggests that the pathogenesis of COPD and emphysema is the result of an imbalance between enzymes that degrade the ECM within the lung and proteins that oppose this proteolytic activity. Many proteases play important roles in remodelling or inflammation within the lung. It is essential that they are controlled by antiproteases to protect against uncontrolled degradation of the ECM.

The best understood example of genetically-induced emphysema results from mutations in the $\alpha_1$-antitrypsin gene (SERPINA1; 14q32.1). These increase the propensity of the protein to form ordered polymers which are incapable of inhibiting its target enzyme, neutrophil elastase. This abnormal behaviour leads to retention of the protein within hepatocytes as Periodic Acid Schiff positive inclusions and results in plasma deficiency of an important protease inhibitor (OMIM #107400). It is now increasingly recognised that mutant $\alpha_1$-antitrypsin can also form polymers within the interstitium and alveolar spaces of the lung. These polymers are chemotactic for neutrophils and so combine with the deficiency of $\alpha_1$-antitrypsin to focus and amplify the inflammatory response within the lung.22 In most Northern European populations the frequency of the most severe Z allele is about 1/2000. Classically, Z $\alpha_1$-antitrypsin homozygotes carry the Glu342Lys mutation and suffer from early-onset emphysema when compared with individuals with normal MM $\alpha_1$-antitrypsin. The onset and progression of emphysema is markedly accelerated by cigarette smoking. Moreover, it appears that even a single allele of Z $\alpha_1$-antitrypsin may increase the risk of COPD. In the longitudinal Copenhagen City Heart Study the ZZ $\alpha_1$-antitrypsin genotype increased the rate of decline of FEV$_1$ by 19% compared with those who were MM homozygotes, causing a 30% increased risk of obstructive lung function and a 50% increased risk of physician-diagnosed COPD.23 The authors found that the frequency of the MZ genotype in their Danish population was as high as 5% and calculated that it would account for 2.4% of cases of COPD. This is in contrast to the ZZ genotype which was causal in only 0.8% of cases. In meta-analysis, heterozygosity for the Z allele, while another increases activity by 25% (His139Arg, the "fast" allele). In 1997 the "slow" variant of EPHX1 was found to segregate with COPD in a large multicentre family-based study and to be associated with COPD in a case-control analysis.27 However, another large study failed to replicate the association with COPD despite having adequate power.28 The latter study included individuals with COPD with and without emphysema, while the studies by DeMeo and colleagues26 included a preponderance of patients with emphysema assessed for lung volume reduction surgery. Nevertheless, while these differences may reflect different COPD phenotypes, they illustrate the need to replicate the findings of genetic association studies in multiple populations before drawing firm conclusions.

Since mutations of $\alpha_1$-antitrypsin so clearly lead to emphysema, one might infer that its target, neutrophil elastase, is central to the pathogenesis of disease. However, mutations in this protease have not been shown to be important, despite being studied extensively in other conditions. Instead, most evidence implicates matrix metalloproteinases (MMPs) in the pathogenesis of COPD. These are zinc-dependent endopeptidases involved in the degradation of many ECM components. An SNP of MMP9 (20q11) was associated with COPD in Japanese and Chinese populations; however, a further Japanese study found an association with emphysema distribution rather than with COPD per se.23 Another large study failed to show an association of MMP9 with COPD, but instead MMP1 (11q22) and MMP12 (11q22) polymorphisms were identified.22 More recently, further data supporting a role for MMP9—but not MMP1—polymorphisms have been published.23 Tissue inhibitors of metalloproteinases (TIMPs) inhibit the MMPs but, thus far, only one polymorphism in TIMP2 (17q25) has been associated with COPD.24

REACTIVE OXYGEN SPECIES

Cigarette smoke contains vast numbers of free radicals that impose an oxidative stress on the lung. Such stress is believed to induce damage through multiple mechanisms, including direct oxidation of cellular lipids and DNA, and through inactivation of key proteins such as $\alpha_1$-antitrypsin. For this reason, much work has gone in to assessing the role of endogenous antioxidant enzymes in protecting against smoke-induced lung damage.

Many toxins in cigarette smoke are subject to first pass metabolism in the liver. Among the many enzymes involved, microsomal epoxide hydrolase (EPHX1; 1q42.1) has been intensely studied in the context of COPD. Several EPHX1 SNPs have been described that affect its activity. One of these leads to a 40% loss of in vitro activity (Tyr113His, the “slow” allele), while another increases activity by 25% (His139Arg, the “fast” allele). In 1997 the “slow” variant of EPHX1 was found to increase the risk of emphysema by a staggering odds ratio of 5.0 and of COPD by an odds ratio of 4.1.29 Since then, numerous studies have attempted to reproduce this effect with varying success.30–45 Contradicting that original study, a recent systematic meta-analysis found homozygosity for the “slow” (Tyr113His) allele to be protective against COPD (odds ratio 0.5).46 This paradoxical result was possible only because five of the eight studies identified by those authors (including the 1997 study by Smith and Harrison)35 36 41 44 45 were excluded on the grounds that alleles in the control group failed to show HWE. HWE refers to the expected distribution of genotypes in a stable population; deviation suggests either a systematic defect in genotyping or an unidentified bias in the selection of subjects. The failure to consider HWE in many studies is an important pitfall in population genetics and is all too common in studies of COPD. It would be fair to say that, despite the considerable effort invested in EPHX1, the role of its alleles in COPD remains...
unclear. Work on this gene continues and a recent analysis of the National Emphysema Treatment Trial (NETT) dataset has suggested a role for EPHX1 polymorphism in both the severity of COPD and the distribution of emphysema. 45

Glutathione S-transferase (GST) comprises a large family of enzymes capable of catalysing the conjugation of reduced glutathione to endogenous and xenobiotic electrophilic compounds. The GSTs are important in the detoxification of many compounds and are highly polymorphic. These polymorphisms have been linked to susceptibility to toxins and carcinogens. SNPs in GSTP1 have been associated with COPD, 39 the distribution of emphysema 40 and more rapid decline in lung function. 30 However, the data should be interpreted with caution as the third of the cohorts 40 has been used in multiple analyses 40 and there was a lack of HWE for GSTP1 in their population. Moreover, no convincing association was found in other studies. 44 51 The null mutation of GSTM1 (1p13.1) has also been associated with COPD, 57 but others have failed to reproduce this finding. 40

Heme oxygenase catalyses the first step in heme degradation. Heme oxygenase 1 (HMOX1; 22q13.1) is the inducible isoform that can be upregulated by a wide range of stresses. Bile pigments generated by heme cleavage are believed to have antioxidant properties, thus HMOX1 induction is protective during cellular oxidant injury and overexpression of HMOX1 in lung tissue protects against hyperoxia. The HMOX1 gene 5'-flanking region contains stretches of GC repeats that are highly polymorphic in length. Yamada and colleagues found a higher proportion of long repeats in patients with COPD and also showed that long repeats were associated with impaired promoter activity. 54 Attempts to reproduce this effect have had varied success. 53 54 In one study an attempt was made to replicate the association in both a case-control study and family-based study. 53 Confusingly, in the case-control study an association was seen between COPD and 30 but not 31 GC repeats, while in the family study the association was seen with the GC31 allele. While HMOX1 GC repeat length has not convincingly been shown to be associated with developing COPD, there are some data to support an association between the long allele and increased severity of disease. 33 55

Superoxide dismutase (SOD) is an important antioxidant enzyme that catalyses the conversion of superoxide to oxygen and hydrogen peroxide. The extracellular isoform (SOD3; 4p15) is abundant in lung parenchyma. In the cross-sectional Copenhagen Heart Study, the R213G allele that results in impaired antioxidant properties, thus SOD3 isoenzyme, but not for other forms of SOD. 57

While biologically very plausible, current genetic evidence fails to provide clear support for the involvement of detoxifying enzymes in the pathogenesis of COPD. Since the potential list of candidates to detoxify cigarette smoke remains long, it would be preferable if future studies were to take an unbiased approach to target identification rather than the current vogue for studying small numbers of candidate genes.

INFLAMMATION

Tumour necrosis factor α (TNF; 6p21) is a multifunctional cytokine whose levels are increased in bronchoalveolar lavage fluid, induced sputum samples and biopsies from patients with COPD. It is a plausible candidate gene for susceptibility to inflammatory disease, especially as well studied promoter polymorphisms clearly alter expression levels. Consequently, considerable effort has been invested into determining whether the promoter polymorphism in TNFs also predisposes smokers to COPD. Much interest was generated when an early study revealed an association (with a staggering odds ratio of over 10) between allele 2 and “bronchitis” in Taiwanese men. 48 This study is difficult to interpret as a third of the men were “never smokers”. Subsequent studies have found little evidence that TNF polymorphisms are associated with or modify the progression of COPD. 47 56 58

Group specific component (GC; 4q12), also known as vitamin D binding globulin, is a multifunctional protein that enhances the neutrophil and monocyte chemotactic activity of complement component 5a. It is a highly polymorphic protein with more than 124 forms, although three (Ga1, Gc1S and Gc2) make up the majority. Kueppers and colleagues found Gc2 homozygotes to be protected from COPD. 62 Others have seen this protective effect, 72 73 while Gc1F homozygosity has been found to be associated with COPD. 72 73 However, a much larger recent study has failed to reproduce these associations. 74

Once again, a biologically plausible pathway has failed to yield unequivocal supportive genetic evidence of association with susceptibility to COPD. The complexity of innate and acquired immunity makes inflammation a far less attractive target for the candidate gene approach in studying COPD than was smoke detoxification. Only an unbiased approach is sufficiently robust to resolve this question.

CONCLUSIONS

Despite enormous efforts invested into understanding the genetics of COPD, most findings support the existing models of ECM integrity, protease/antiprotease balance and probably oxidative stress. It is now essential that future studies take advantage of multiple large cohorts that have been phenotyped with susceptibility to COPD. Much interest was generated when an early study showed that long repeats were associated with impaired promoter activity. 54 Attempts to reproduce this effect have had varied success. 53 54 In one study an attempt was made to replicate the association in both a case-control study and family-based study. 53 Confusingly, in the case-control study an association was seen between COPD and 30 but not 31 GC repeats, while in the family study the association was seen with the GC31 allele. While HMOX1 GC repeat length has not convincingly been shown to be associated with developing COPD, there are some data to support an association between the long allele and increased severity of disease. 33 55

Superoxide dismutase (SOD) is an important antioxidant enzyme that catalyses the conversion of superoxide to oxygen and hydrogen peroxide. The extracellular isoform (SOD3; 4p15) is abundant in lung parenchyma. In the cross-sectional Copenhagen Heart Study, the R213G allele that results in impaired antioxidant properties, thus SOD3 isoenzyme, but not for other forms of SOD. 57

While biologically very plausible, current genetic evidence fails to provide clear support for the involvement of detoxifying enzymes in the pathogenesis of COPD. Since the potential list of candidates to detoxify cigarette smoke remains long, it would be preferable if future studies were to take an unbiased approach to target identification rather than the current vogue for studying small numbers of candidate genes.

INFLAMMATION

Tumour necrosis factor α (TNF; 6p21) is a multifunctional cytokine whose levels are increased in bronchoalveolar lavage fluid, induced sputum samples and biopsies from patients with COPD. It is a plausible candidate gene for susceptibility to inflammatory disease, especially as well studied promoter polymorphisms clearly alter expression levels. Consequently, considerable effort has been invested into determining whether the promoter polymorphism in TNFs also predisposes smokers to COPD. Much interest was generated when an early study revealed an association (with a staggering odds ratio of over 10) between allele 2 and “bronchitis” in Taiwanese men. 48 This study is difficult to interpret as a third of the men were “never smokers”. Subsequent studies have found little evidence that TNF polymorphisms are associated with or modify the progression of COPD. 47 56 58

Group specific component (GC; 4q12), also known as vitamin D binding globulin, is a multifunctional protein that enhances the neutrophil and monocyte chemotactic activity of complement component 5a. It is a highly polymorphic protein with more than 124 forms, although three (Ga1, Gc1S and Gc2) make up the majority. Kueppers and colleagues found Gc2 homozygotes to be protected from COPD. 62 Others have seen this protective effect, 72 73 while Gc1F homozygosity has been found to be associated with COPD. 72 73 However, a much larger recent study has failed to reproduce these associations. 74

Once again, a biologically plausible pathway has failed to yield unequivocal supportive genetic evidence of association with susceptibility to COPD. The complexity of innate and acquired immunity makes inflammation a far less attractive target for the candidate gene approach in studying COPD than was smoke detoxification. Only an unbiased approach is sufficiently robust to resolve this question.

CONCLUSIONS

Despite enormous efforts invested into understanding the genetics of COPD, most findings support the existing models of ECM integrity, protease/antiprotease balance and probably oxidative stress. It is now essential that future studies take advantage of multiple large cohorts that have been phenotyped with susceptibility to COPD. Much interest was generated when an early study showed that long repeats were associated with impaired promoter activity. 54 Attempts to reproduce this effect have had varied success. 53 54 In one study an attempt was made to replicate the association in both a case-control study and family-based study. 53 Confusingly, in the case-control study an association was seen between COPD and 30 but not 31 GC repeats, while in the family study the association was seen with the GC31 allele. While HMOX1 GC repeat length has not convincingly been shown to be associated with developing COPD, there are some data to support an association between the long allele and increased severity of disease. 33 55

Superoxide dismutase (SOD) is an important antioxidant enzyme that catalyses the conversion of superoxide to oxygen and hydrogen peroxide. The extracellular isoform (SOD3; 4p15) is abundant in lung parenchyma. In the cross-sectional Copenhagen Heart Study, the R213G allele that results in impaired antioxidant properties, thus SOD3 isoenzyme, but not for other forms of SOD. 57

While biologically very plausible, current genetic evidence fails to provide clear support for the involvement of detoxifying enzymes in the pathogenesis of COPD. Since the potential list of candidates to detoxify cigarette smoke remains long, it would be preferable if future studies were to take an unbiased approach to target identification rather than the current vogue for studying small numbers of candidate genes.

REFERENCES


55. Ishii T, Matsuse T, Teramoto S, et al. Neither IL-1beta, IL-1 receptor antagonist, nor TNF-alpha polymorphisms are associated with susceptibility to COPD. Respir Med 2000;94:947–51.


Pulmonary puzzle

ANSWER

From the question on page 290

The total white cell count of the BAL fluid was 184/mm³ with 54% eosinophils. Figure 1 shows clusters of granulated and degranulated eosinophils within a background of septal oedema. Comprehensive examination of the BAL fluid and biopsy material yielded no clear infectious cause. These findings are consistent with a diagnosis of acute eosinophilic pneumonia (AEP), which is characterised by progressive respiratory symptoms, diffuse bilateral infiltrates and increased numbers of lung eosinophils in the absence of detectable infection. Biopsy, although not necessary for diagnosing AEP, may be useful for ruling out other causes of pulmonary eosinophilia. Other entities to consider in patients with acute progressive respiratory symptoms and increased numbers of eosinophils in the BAL fluid include drug reactions and fungal and parasitic infections.

The clinical course typically begins with fever averaging 38°C, cough, tachypnoea, hypoxaemia and dyspnoea and often progresses rapidly. Plain radiographs may show an interstitial, alveolar or mixed pattern and pleural effusions are common. Environmental, occupational and recreational exposures are thought to play a significant role in the pathogenesis of AEP. A number of authors have described the development of AEP shortly after the onset of smoking. AEP has also been reported in a patient who became symptomatic shortly after participating in a bicycle race “in dusty conditions”. Our patient had both exposures, as he later admitted to trying cigarette smoking for the first time the evening before his symptoms developed.

At presentation the patient had increased peripheral blood neutrophils, a phenomenon previously reported in cigarette smoke-induced AEP. While peripheral blood eosinophilia characteristically does not occur during the acute phase of AEP, levels may become quite high during convalescence before normalising. The patient’s peripheral blood absolute eosinophil count was 1380/mm³ 1 week after discharge. On rechecking 5 months later, it had returned to the normal range.

While individuals with AEP classically show clinical improvement after the initiation of steroids, most who are treated with only supportive measures also recover completely. Although case reports have described the return of symptoms upon re-exposure to an inciting factor, most patients develop tolerance and do not have recurrence. We advised our patient to avoid cigarette smoke altogether. His symptoms resolved completely with no lasting pulmonary function abnormalities after a course of steroids.

Acknowledgements: The authors wish to thank Dr Rubin Tuder for his assistance with obtaining the histological image.

REFERENCES