Cystic fibrosis modifier genes

Jane Davies MRCPCH MD[1] Eric Alton MD FRCP[2] Uta Griesenbach BSc PhD[3]

J R Soc Med 2005;**98**(Suppl. 45):47–54

The cystic fibrosis transmembrane conductance regulator (*CFTR*) gene encodes a chloride channel in the apical membrane of epithelial cells. Either absence or abnormal function of this protein leads to the recessively inherited disease, cystic fibrosis (CF). Of the 1000 or so mutations which have been identified in CFTR,[1] by far the most common is a deletion of phenylalanine at position 508 (*Δ*F508), accounting for approximately 70% of CF chromosomes. Many of the other mutations occur in less than 1% of CF subjects. The mutations fall into five different categories depending on their effect on the CFTR protein:[2] class I mutations lead to defective protein production, class II (including *Δ*F508) to defective protein maturation and processing, class III to defective channel regulation/gating, class IV to altered channel conductance and class V to altered protein stability.

GENOTYPE/PHENOTYPE CORRELATION

CF affects the pancreas, gut, liver, reproductive tract and airways; it is the involvement of the latter which leads to most morbidity and is the most common cause of death. Recently, a large retrospective cohort study of approximately 17 000 patients from the US CF Foundation National Registry, confirmed that *CFTR* genotype affects mortality.[3] However, in individual patients, correlation with CFTR mutation differs with the organ studied. Pancreatic disease, manifest as exocrine deficiency, occurs in over 90% of patients and correlates well with *CFTR* genotype.[4] Pancreatic insufficient (PI) subjects usually possess two mutations in classes I, II or III. Class IV and V mutations, often classified as 'mild', lead to higher levels of residual CFTR function and are generally associated with pancreatic sufficiency (PS).[5] In contrast, genotype/phenotype correlation for lung disease is poor; severity of lung disease can vary greatly even in siblings carrying identical *CFTR* mutations.

If not CFTR itself, which other factors affect phenotype? Clearly, environment plays a role. This encompasses delays in diagnosis, the availability of and adherence to treatments, as well as more conventional 'environmental' factors, such as exposure to tobacco (*in utero* and active or passive smoking) and airborne pollutants. Stochastic factors (those occurring randomly), such as the timing of acquisition of infection, play an effect of unknown, but probably significant, proportions. Finally, the fact that phenotype is more concordant in monozygotic twins than in dizygotic twins or siblings,[6] argues for the influence of additional non-*CFTR* genetic factors, so-called 'modifier genes'. Identification of these genes will not only aid our understanding of pathogenesis, but may also lead to more rational use of conventional treatments and the development of novel drug targets. The evidence for a role for such genetic factors and those that have been identified to date, form the basis of this article.

THE BASIS OF GENETIC VARIATION

The DNA sequence of any two unrelated individuals is approximately 99.9% identical. The remaining 0.1% contains variations (polymorphisms), which may influence the likelihood or severity of disease and affect responses to drugs. The term polymorphism includes single-base nucleotide substitutions (also known as single nucleotide polymorphisms or SNPs), small-scale, multi-base deletions or insertions (also called deletion insertion polymorphisms), and repeat variations (also called short tandem repeats). In contrast to mutations, which occur at low frequencies (less than 1%) in the general population, polymorphisms are common genetic variants, occurring more frequently. Recently, SNPs have received the most attention, mainly due to the development of high throughput analysis.[7] In 2002 the international HapMap project [www.HapMap.org] was founded, which aims to generate a SNP haplotype map of the human genome within the next 3 years. A similar project is being carried out in different strains of mice [www.jax.org], which may provide important supplementary data. It is thought that up to 10 million SNPs exist in the human genome. Within genes, the number of SNPs varies widely, from one or two to several hundred. It is likely, however, that all SNPs are not equally important, and analysis is often restricted to those in: (*a*) exons (the coding region of the gene), particularly if leading to an amino acid substitution; (*b*) regulatory regions, which may affect binding of transcription factors; (*c*) exon/intron boundaries, which may influence splicing; (*d*) conserved regions, as this usually implies a significant function.

[1]Department of Gene Therapy, Faculty of Medicine, Imperial College, Emmanuel Kaye Building, Manresa Road, London SW3 6LR; Departments of [2]Paediatric Respiratory Medicine and [3]Thoracic Medicine, Royal Brompton Hospital, London, UK

Correspondence to: Dr Jane Davies

E-mail: j.c.davies@imperial.ac.uk

Many studies in other diseases have identified significant effects of such polymorphisms: the apolipoprotein E4 allele in early onset Alzheimer disease;[8] the peroxisome proliferator-activated receptor-gamma and type 2 diabetes;[9] the *NOD2* gene and Crohn's disease[10] and the human leucocyte antigen (HLA) gene cluster in sarcoidosis[11] and diabetes.[12] SNPs have also been suggested as prognostic tools for example in breast cancer[13] and cardiovascular disease.[14] Pharmacogenetics,[15] the study of genetic factors influencing variability in drug response and toxicity, has recently become a popular area of research; important advances have been made in several fields, for example oncology.

MODIFIER GENES IN CF

In most patients with CF, respiratory involvement becomes manifest in the first few years of life, although problems may present in adulthood or be apparently absent.[3] Most commonly, chronic infection with bacteria such as *Staphylococcus aureus* and *Pseudomonas aeruginosa* leads to an exaggerated and sustained, neutrophil-dominated inflammatory response, bronchiectasis and death from respiratory failure in over 90% of patients.[16] Based on our understanding of this sequence of events, modifiers might include non-CFTR ion channels or genes involved in host defence, inflammation, epithelial repair, mucin production and airway responsiveness.

ANTIPROTEASES

In the CF airway, chronic inflammation results in an excess of destructive proteases such as neutrophil elastase. These overwhelm their inhibitors, the anti-proteases, of which alpha-$_1$-antitrypsin (α-$_1$-AT) is one of the most abundant. Patients with inherited forms of α-$_1$-AT deficiency are at risk of emphysema,[17] and thus, it was postulated that CF patients with coexisting α-$_1$-AT deficiency would demonstrate a more severe pulmonary phenotype. In this case, the development of recombinant forms of antiproteases could mean a potential new treatment for this subset of patients. In the first of several studies, Doring *et al*. reported on 215 patients in 1994. α-$_1$-AT-deficient CF patients became chronically infected with *P. aeruginosa* at an earlier age, although this did not lead to poorer lung function.[18] Subsequently, Mahadeva *et al*. concluded that α-$_1$-AT-deficient CF subjects (20 of 147) were actually protected from lung disease, having a mean forced expiratory volume in the first second (FEV_1) of 62.5%, significantly greater than the rest of the group at 51.1%.[19] This group has reported remarkably similar findings with α-$_1$-antichymotrypsin, a closely related antiprotease.[20] They later studied 79 'severe' patients, classified on the basis of either death or requirement for transplantation, and found no increase in α-$_1$-AT deficiency in this group.[21] A group from Ireland has reported that 16 (of 134) patients possessing an enhancer polymorphism, leading to increased α-$_1$-AT levels, also had a better prognosis.[22] In contrast, however, in the largest study reported to date (n=716), Frangolias *et al*. have found no relationship between α-$_1$-AT-deficiency and lung function, age at acquisition of *P. aeruginosa*, requirement for transplantation or death.[23]

INFLAMMATION

The HLA region is the most polymorphic in the human genome, encoding hundreds of genes including the major histocompatibility complexes (MHC). MHC class II molecules are critical in antigen presentation and the ensuing inflammatory response. Polymorphisms have been reported in association with a number of diseases including auto-immunity, asthma and allergy, and to date, two reports have linked certain haplotypes with complications in CF. Aron *et al*. reported in 98 adult patients that DR7 was significantly associated with higher IgE levels and risk of *P. aeruginosa* infection[24] (although only 10 patients were non-infected, all were DR7-negative) and an increased frequency of allergic bronchopulmonary aspergillosis (ABPA) in subjects carrying the DR2 allele.[25] In addition, a striking increase in the frequency of DQ2 in patients without ABPA might suggest some protective effect of this allele. Tumour necrosis factor-alpha (TNF-α), a cytokine found in high concentrations within the CF airway, is thought to be pivotal in the promotion of the neutrophil-dominated inflammatory response and has been inversely correlated with lung function.[26] In 1998, Hull *et al*. published results of a study on 53 CF children.[27] Twenty possessed a polymorphism in the promoter region of the TNF-α gene known to be associated with higher (constitutive and induced) levels of the cytokine. These children had a significantly lower FEV_1 and poorer nutritional status than the children without the polymorphism. They also studied glutathione-S transferase (GST), an enzyme involved in oxidative stress, and reported more severe chest radiograph and Schwachman score in children homozygous for the GSTM1 null allele.[27] With regard to TNF-α, in a larger study including both adults and children, Arkwright *et al*.[28] found no such association when studying the same promoter polymorphism, and in 269 adult patients, we have also found no link with any marker of severity.[29]

NITRIC OXIDE SYNTHESIS

The antimicrobial effects of nitric oxide (NO) are being increasingly recognized. High numbers of the AAT trinucleotide repeat sequence in the *NOS-1* gene are associated with low levels of exhaled NO, and were found

in CF patients to confer an increased risk of infection with both *P. aeruginosa* and *Aspergillus fumigatus*, although this did not lead to a more rapid decline in lung function.[30] This group has also recently related a polymorphism in the *NOS-3* gene with risk of infection.[31] This gene, expressed in vascular endothelium, respiratory epithelium and neutrophils, contains a functionally important polymorphism (894G/T), which affects the resistance of NOS-3 to proteolysis. Interestingly, previous work has highlighted gender differences, by demonstrating that circulating oestrogen increases the levels of NOS-3 in the vascular endothelium.[32] Grasemann *et al*. reported higher exhaled NO levels and decreased frequency of *P. aeruginosa* infection in association with the 894T polymorphism in females only.[31] What is not clear from these two related studies is whether *NOS-1* and *NOS-3* are independent modifiers, or whether there is a confounding effect of one upon the other.[33] Further, in a letter on this subject, Mekus and Tummler[34] suggest that these genes may merely be 'hitchhiking' with another gene of relevance, rather than directly involved in disease modification themselves.

GROWTH FACTORS

Polymorphisms leading to high levels of the profibrotic cytokine, tissue growth factor (TGF)-β, have previously been associated with increased pulmonary fibrosis after chemo- and radiotherapy and organ transplantation. Arkwright *et al*. studied TGF-β in a cohort of 261 CF patients (children and adults) from Manchester, UK.[35] Patients with known 'mild' *CFTR* mutations were excluded from analysis. Subjects with at least one high-expressing haplotype had a significantly faster rate of decline in both FEV_1 and forced vital capacity (FVC) than those with low-expressing variants. Interestingly, this would appear to contrast with another study demonstrating that TGF-β expression was highest in the CF patients with the mildest lung disease.[36]

INNATE IMMUNITY

Mannose-binding lectin (MBL), a liver-derived serum protein involved in innate defence, exerts its effects both by direct opsonization of pathogens and by activation of the complement system.[37] Polymorphisms have been identified within both the structural and promoter regions of human *MBL-2*, which result in mutant forms of the protein incapable of forming functional, high order oligomers and low circulating protein levels. Low levels of MBL have been shown to relate to a variety of infective processes including recurrent respiratory infections, and so the *MBL-2* gene was considered a likely candidate as a modifier in the CF lung. In the first study to examine such a link, Garred *et al*. reported that both FEV_1 and FVC were significantly lower in subjects with either one or two structural *MBL-2* mutations, but only following chronic *P. aeruginosa* infection.[38] A trend towards this was visible as early as 8 years of age, although this became significant only at 16 years of age. Although numbers were small, the authors also reported an increased risk of infection with *Burkholderia cepacia*. Another study from France compared patients deficient in MBL with controls closely matched for age and CF genotype (all homozygous ΔF508).[38] In contrast to the earlier study, this group found a significant reduction in lung function only in patients possessing two variant alleles. In support of this, Cutting *et al*. have reported a significant survival disadvantage in American CF patients with 2 mutations.[40] As part of a large study in almost 600 patients, we have recently found that adult patients possessing two structural mutations, but not heterozygotes, have significantly impaired lung function, oxygen saturations and raised inflammatory markers.[41] In contrast to the data from Garred *et al*., this was not seen in our paediatric age group, in whom the majority had well-preserved lung function. This difference highlights an important consideration in modifier gene studies which will be elaborated on later, namely that historical data from current adults may not be applicable to today's children. Treatment regimes have evolved greatly over the last few years and, for example, a polymorphism that was significant in the era before the widespread use of antipseudomonal antibiotics, may be less relevant now. In terms of therapy, MBL is an attractive candidate: several patients worldwide have received purified human MBL, including one young adult with end-stage CF.[42] In response to intravenous treatment, her lung function and inflammatory markers were observed to plateau, although she subsequently died, off treatment, whilst awaiting transplantation. The development of a recombinant form of the protein may prove clinically beneficial. Alternatively, gene-based therapy could be both potentially cheaper and, because of the secreted nature of the protein, somewhat easier than gene therapy for membrane-bound proteins such as CFTR. However, whether such treatments would be useful only for patients with two mutations, or others as well, is uncertain, a factor which will almost certainly impact on the financial viability of product development.

AIRWAY RESPONSIVENESS

Polymorphisms in the β-adrenergic receptor (β-AR) have previously been related to severity of asthma and response to treatment with β_2-agonist drugs.[43] β-AR are important regulators of cAMP in the airway, recent *in vitro* data demonstrating that ion transport via cAMP-dependent CFTR can be activated by β_2-agonists.[44] Buscher *et al*. studied the effects of three polymorphisms in the *β-AR* in 87 young adults and children with CF.[45] Subjects with either

one or two copies of Gly16, an amino acid change leading to down-regulation of the receptor, had significantly reduced lung function and faster rates of decline than those patients homozygous for Arg16. These differences were more marked when only ΔF508 homozygotes were studied. In addition, Gly16 was significantly less common in the CF population than in several groups of healthy controls, possibly implying a survival disadvantage. Unfortunately, the group was not large enough for subgroup analysis on the basis of age, which might have helped support this hypothesis. There were no differences in bronchodilator responsiveness, but in an *in vitro* assay, lymphocytes from these subjects showed a blunted cAMP response to isoproteronol-stimulation suggesting that the clinical findings may relate to differences in level of CFTR function between the two groups.

GASTROINTESTINAL AND LIVER DISEASE

Between 10% and 15% of CF patients are born with meconium ileus (MI), obstruction of the distal small bowel with inspissated meconium.[46] The recognition that strain background was heavily influential in the development of MI in CF mouse models, led to the identification of a genetic locus linked to the disease.[47] The corresponding site on chromosome 19 has now been linked with MI in humans.[48] Of 7 sib pairs where both children had MI, over half possessed two shared haplotypes at this region, whereas in 33 MI-discordant sib pairs, 61% shared neither haplotype. This is a highly significant finding, although the protein encoded at this site remains unidentified. Similarly, familial clustering of portal hypertension suggests that liver disease in CF may be under genetic influence, although few associations have been found with *CFTR* genotype. Duthie *et al*. performed a large multicentre study on 274 unrelated children and adults examining the effect of HLA status.[49] Almost 30% of patients had evidence of chronic liver disease, a higher proportion than reported from most centres. DQ6 was found in 66% of liver disease patients but only 33% of those without. Two other antigens in strong linkage disequilibrium with this locus, DR15 and B7 were also significant risk factors. When portal hypertension was used as a marker of chronicity, these markers were found to be significant for males only, but there was no association with age of onset. MBL-deficiency was shown by one group to be a risk factor for the development of CF liver disease,[50] although we did not find this in our study.[41] Arkwright *et al*. have reported associations between liver disease and both high expressing TGF-β haplotypes[35] and ACE[28] (an enzyme involved in TGF-β activation) polymorphisms. One of the major problems with studies of liver disease is one of definition, which differs widely in clinical practice. Finally, Mekus *et al*. identified loci in a partially imprinted region 3′ of *CFTR* as modifiers of both nutritional and pulmonary phenotype in 34 highly concordant or discordant sib pairs.[51] This region includes both the leptin gene and a candidate for Russell–Silver dwarf syndrome, and is thus likely involved in growth, food intake and energy expenditure.

CFTR POLYMORPHISMS

In addition to these non-CFTR modifiers, over 200 polymorphisms have been identified within CFTR itself[1]. These do not cause CF, but may alter CFTR protein production and/or function. Such alterations may be clinically insignificant in individuals without additional *CFTR* mutations, but have an influence on disease phenotype in patients with co-existing mutations. A good example of this is a polymorphism in intron 8 of the *CFTR* gene. The number of thymidine residues (T) in this locus is either 5, 7 or 9;[52] individuals with the 5T polymorphism, which is present in about 10% of the general population, have reduced amounts of functional CFTR due to inefficient splicing of exon 9. Subjects with 5T who also carry a severe *CFTR* mutation in trans (in their second *CFTR* allele) can either be asymptomatic, infertile due to congenital absence of the vas deferens, or have a subset of CF symptoms (non-classical CF), depending on how much residual CFTR function is retained.[53] Interestingly, the length of a thymidine/guanine (TG) repeat adjacent to 5T also affects the amount of incorrectly spliced CFTR.[54] So-called complex *CFTR* alleles, *CFTR* genes that carry two mutations on the same chromosome, may also influence the phenotype. Albeit rare, there are examples of patients with complex alleles where the second mutation modulated the effect of the first mutation and may have altered the phenotype.[55,56]

ISSUES IN STUDY DESIGN

It will be clear from this, that some published studies exploring the same gene have reached different conclusions. To some extent these discrepant results may reflect true differences in populations, but it seems equally likely, if not more so, that much of this discrepancy relates to study design. There is currently no consensus on many of the issues, such as phenotype definition and statistical analysis, although these could help progress the field. Conversely, it could be argued that studies reaching similar conclusions despite differing methodologies may provide greater support to each other's conclusions than do mere duplicates.

Subjects

The power of any modifier gene study will be determined by population size, the degree of clinical effect of the gene in question and the allele frequency. For investigators

studying the effects of multiple genes, where the magnitude of the second and third factors will differ for each gene, prospective power calculations are difficult. Large groups will increase power, but the clinical data collected may be more limited. For the largest of such studies, the data may be multi-centre, which could lead to lack of consistency. It might be more useful to study smaller but better-defined groups, such as siblings or unrelated subjects matched for age, *CFTR* genotype, gender or CF centre. However, well-designed, large studies may allow subgroup analysis and thus decrease the potential for missing a significant link. As an example, certain genes may modify severity only in patients of one gender, with so-called 'mild' *CFTR* mutations, or at certain ages or stage of disease. Unless driven by pre-defined hypotheses, smaller studies may lack the power to detect significant effects on subgroups. However, smaller studies may generate qualitatively superior data, and thus, may allow a better definition of phenotype. Restricting analysis to twins, siblings or patients within a single centre receiving similar treatment, will reduce the influence of environmental differences, including those related to the level and type of care received. When considering issues of study size, it may be useful to reflect on the experience of the cardiovascular field, in which publication bias of small studies with positive findings has led to conclusions which were refuted once study size increased or meta-analyses were performed.[57] Finally, the fact that CF is a progressive disease makes comparisons within age bands highly desirable. The prognosis of CF has improved dramatically over the last few decades. Patients are now receiving improved treatment, living longer and reaching severe stages of disease at an older age. This means that historical data (e.g. from the childhood of current adults) may not be safely compared or pooled with contemporary data (e.g. that of current children).

Defining phenotype

Modifier gene studies have set out to define genetic factors leading to either a discrete outcome (pancreatic status, liver involvement or infection with specific pathogens) or to continuous variables, which indicate severity (lung function, nutrition). The definition of phenotype is difficult for many of these outcomes. Whereas pancreatic status is usually clear, and can be tested, definitions of liver disease vary, and in the early stages, organ involvement may not be obvious. With regard to infection status and timing, non-expectorating patients, particularly small children, pose a big problem. It is well known from bronchoscopic studies that lower airway infection may go undetected, and that the presence or absence of organisms on upper airway samples may be misleading.[58] Such patients may therefore be easily misclassified. Most modifier gene studies have focused on lung function as the main phenotypic variable. However, the ways in which lung function data have been obtained and analysed differ. Some investigators have used current data or retrospectively acquired longitudinal data, from which they have calculated either averages or rates of decline, and compared these across haplotype groups. Alternative methods have been to compare the ages at which patients have reached defined lung function, e.g. 50% of predicted values, or to group patients into extreme groups (mild and severe) based on lung function, and examine the proportions of haplotypes within each of these. It is unclear from many of these studies whether physiological data have been collected only at times of clinical stability, or whether declines during infective exacerbations may have influenced the data. If the frequency of such exacerbations is itself being analysed, these need to be well defined, which in clinical practice is rarely the case. Finally, age at death or transplantation, although clear cut, is rarely useful, since the clinical status of CF patients has fortunately improved so much in recent years. We and others are attempting to develop sensitive assays related to clinical status, but more closely linked to the basic CFTR defect, such as airway surface liquid height, bacterial adherence and microarray analysis of gene expression. Although these are being developed primarily as end-point assays for future therapeutic trials, they may also aid phenotypic definition for modifier studies. The future gold standard may be carefully designed, prospective studies, recruiting children at diagnosis and following them up for their lifetime with regular monitoring for well-defined outcome measures; such studies are clearly logistically complex and will yield results only in the medium to longer term. The increasingly rigorous methods of analysing polymorphisms need to be matched by a similar level of attention to phenotypic detail.

Statistics and multiple comparisons

A review of statistical methods employed in large modifier studies is outside the scope of this article, although a few concerns should be raised. First, in studies exploring the effects of many genes, some of which may contain many polymorphic sites, multiple comparisons will be required, increasing the likelihood of positive correlations appearing by random chance (α error). This must be taken into account when statistical methods are chosen, when corrections for multiple comparisons are essential. An alternative is to have a second cohort in which hypotheses generated in the first cohort can be tested. With large enough populations, this may be achievable by splitting the group randomly prior to initial data analysis, although detrimental impacts on power caused by lower n numbers must of course be

considered. Secondly, linkage disequilibrium between genes must be considered. If, for example, gene X exerts a major effect on lung function and is linked to gene Y, the latter may appear to be a significant modifier, in the absence of any direct effect. As increasing numbers of genes are identified that appear to affect outcome, statistical techniques need to be adjusted to account for polymorphisms in multiple genes simultaneously.

Ethical issues

Much concern has been raised in the UK in recent years over two issues of relevance here, namely the potential for genetic testing to generate 'prejudice' in areas such as health care or employment and concerns over retention of organs and tissue, without specific consent. With regard to the first of these, we have limited our testing to a list of biologically-plausible CF modifiers; the stored samples will not be tested for unrelated genes without a second application to the ethics committee and specific consent of each patient. Although questions relating to this issue have been raised on occasion by adult patients, after discussion and explanation, no patient, of over 600, has withheld their consent. The successful identification of significant modifiers may pose new ethical challenges for the future. For example, decisions on future pregnancies or terminations could be influenced by a 'mild' or 'severe' modifier within the family. However, unidentified factors which impact on disease progression will always remain and extreme care must be taken by those working in this field that data generated by these studies are not over-interpreted. All institutions carrying out work in these areas should be governed by ethics committees, and informed consent must be sought for inclusion. One requirement of our own committee has been inclusion of specific statements on sample storage and protection of patient anonymity. Finally, children pose a particular ethical challenge.[59] It would probably be considered unethical by most for blood to be taken from a conscious young child purely for research purposes; however, children with CF will require at least annual blood tests, when it is much more acceptable to request that a small portion (if sufficient is obtained) be used for research. In fact, we have adopted this policy even for our adult recruits. Small children are unable to give personal consent, which will be obtained from a carer with parental responsibility, although older children, who can understand the issues may be asked for their assent. Should consent be withdrawn by these children once they are adults (or indeed by any other subject), a system must be in place whereby the samples and data from that individual are excluded from analysis. In our experience, most adults and older children are able to understand the purpose of such studies and indeed the potential benefits.

SUMMARY

Since the recognition that *CFTR* genotype was not a good predictor of pulmonary disease severity in CF, several candidate modifier genes have been identified. It is unlikely that a single modifier gene will be found, but more probable that several haplotypes in combination may contribute, which in itself presents a major methodological challenge. The aims of such studies are to increase our understanding of disease pathogenesis, to aid prognosis and ultimately to lead to the development of novel treatments.

REFERENCES

1 Cystic Fibrosis Mutation Database. [www.genet.sickkids.on.ca/cftr]. Accessed 11 February 2005

2 Welsh M, Ramsey BW, Accurso F, Cutting GR. Cystic fibrosis. In: Scriver AB, Sly WS, Valle D, eds. *The Molecular and Metabolic Basis of Inherited Disease*. New York: McGraw-Hill, 2001:5121–88

3 McKone EF, Emerson SS, Edwards KL, Aitken ML. Effect of genotype on phenotype and mortality in cystic fibrosis: a retrospective cohort study. *Lancet* 2003;**361**:1671–6

4 Santis G, Osborne L, Knight RA, Hodson ME. Independent genetic determinants of pancreatic and pulmonary status in cystic fibrosis. *Lancet* 1990;**336**:1081–4

5 Kerem E, Corey M, Kerem BS, *et al*. The relation between genotype and phenotype in cystic fibrosis—analysis of the most common mutation (delta F508). *N Engl J Med* 1990;**323**:1517–22

6 Bronsveld I, Mekus F, Bijman J, *et al*. Chloride conductance and genetic background modulate the cystic fibrosis phenotype of Delta F508 homozygous twins and siblings. *J Clin Invest* 2001;**108**:1705–15

7 Chen X, Sullivan PF. Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. *Pharmacogen J* 2003;**3**:77–96

8 Rubinsztein DC, Easton DF. Apolipoprotein E genetic variation and Alzheimer's disease: a meta-analysis. *Dement Geriatr Cogn Disord* 1999;**10**:199–209

9 Altshuler D, Hirschhorn JN, Klannemark M, *et al*. The common PPARγ Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 2000;**26**:76–80

10 Hugot JP, Chamaillard M, Zouali H, *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;**411**:599–603

11 Sato H, Grutters JC, Pantelidis P, *et al*. HLA-DQB1*0201: a marker for good prognosis in British and Dutch patients with sarcoidosis. *Am J Respir Cell Mol Biol* 2002;**27**:406–12

12 Kawabata Y, Ikegami H, Kawaguchi Y, *et al*. Age-related association of MHC class I chain-related gene A (MICA) with type 1 (insulin-dependent) diabetes mellitus. *Hum Immunol* 2000;**61**:624–9

13 Iacopetta B, Grieu F, Joseph D. The -174 G/C gene polymorphism in interleukin-6 is associated with an aggressive breast cancer phenotype. *Br J Cancer* 2004;**90**:419–22

14 Boekholdt SM, Peters RJ, de Maat MP, *et al*. Interaction between a genetic variant of the platelet fibrinogen receptor and fibrinogen levels in determining the risk of cardiovascular events. *Am Heart J* 2004;**147**:181–6

15 Goetz MP, Ames MM, Weinshilboum RM. Primer on medical genomics. Part XII: Pharmacogenomics—general principles with cancer as a model. *Mayo Clin Proc* 2004;**79**:376–84

16 Koch C, Hoiby N. Pathogenesis of cystic fibrosis. *Lancet* 1993;**341**:1065–9

17 Doring G, Krogh-Johansen H, Weidinger S, Hoiby N. Allotypes of alpha 1-antitrypsin in patients with cystic fibrosis, homozygous and heterozygous for deltaF508. *Pediatr Pulmonol* 1994;**18**:3–7

18 Mahadeva R, Lomas DA. Genetics and respiratory disease. 2. Alpha 1-antitrypsin deficiency, cirrhosis and emphysema. *Thorax* 1998;**53**:501–5

19 Mahadeva R, Westerbeek RC, Perry DJ, *et al*. Alpha1-antitrypsin deficiency alleles and the Taq-I G–>A allele in cystic fibrosis lung disease. *Eur Respir J* 1998;**11**:873–9

20 Mahadeva R, Sharples L, Ross-Russell RI, Webb AK, Bilton D, Lomas DA. Association of alpha(1)-antichymotrypsin deficiency with milder lung disease in patients with cystic fibrosis. *Thorax* 2001;**56**:53–8

21 Mahadeva R, Stewart S, Bilton D, Lomas DA. Alpha-1 antitrypsin deficiency alleles and severe cystic fibrosis lung disease. *Thorax* 1998;**53**:1022–4

22 Henry MT, Cave S, Rendall J, O'Connor CM, Morgan K, FitzGerald MX, Kalsheker N. An alpha1-antitrypsin enhancer polymorphism is a genetic modifier of pulmonary outcome in cystic fibrosis. *Eur J Hum Genet* 2001;**9**:273–8

23 Frangolias DD, Ruan J, Wilcox PJ, *et al*. Alpha 1-antitrypsin deficiency alleles in cystic fibrosis lung disease. *Am J Respir Cell Mol Biol* 2003;**29**:390–6

24 Aron Y, Polla BS, Bienvenu T, Dall'ava J, Dusser D, Hubert D. HLA class II polymorphism in cystic fibrosis. A possible modifier of pulmonary phenotype. *Am J Respir Crit Care Med* 1999;**159**:1464–8

25 Aron Y, Bienvenu T, Hubert D, Dusser D, Dall'Ava J, Polla BS. HLA-DR polymorphism in allergic bronchopulmonary aspergillosis. *J Allergy Clin Immunol* 1999;**104**:891–2

26 Greally P, Hussein MJ, Cook AJ, Sampson AP, Piper PJ, Price JF. Sputum tumour necrosis factor-alpha and leukotriene concentrations in cystic fibrosis. *Arch Dis Child* 1993;**68**:389–92

27 Hull J, Thomson AH. Contribution of genetic factors other than CFTR to disease severity in cystic fibrosis. *Thorax* 1998;**53**:1018–21

28 Arkwright PD, Pravica V, Geraghty PJ, Super M, Webb AK, Schwarz M, Hutchinson IV. End-organ dysfunction in cystic fibrosis: association with angiotensin I converting enzyme and cytokine gene polymorphisms. *Am J Respir Crit Care Med* 2003;**167**:384–9

29 Low T, Lympany PA, Davies JC, *et al*. The effect of polymorphisms in inflammatory mediators on clinical phenotype in CF. *Pediatr Pulmonol* 2003;**S25**:219

30 Grasemann H, Knauer N, Buscher R, Hubner K, Drazen JM, Ratjen F. Airway nitric oxide levels in cystic fibrosis patients are related to a polymorphism in the neuronal nitric oxide synthase gene. *Am J Respir Crit Care Med* 2000;**162**:2172–6

31 Grasemann H, van's Gravesande KS, Buscher R, *et al*. Endothelial nitric oxide synthase variants in cystic fibrosis lung disease. *Am J Respir Crit Care Med* 2003;**167**:390–4

32 Forstermann U, Boissel JP, Kleinert H. Expressional control of the 'constitutive' isoforms of nitric oxide synthase (NOS I and NOS III). *FASEB J* 1998;**12**:773–90

33 Accurso FJ, Sontag MK. Seeking modifier genes in cystic fibrosis. *Am J Respir Crit Care Med* 2003;**167**:289–90

34 Mekus F, Tummler B. Cystic fibrosis and NOS3. *Am J Respir Crit Care Med* 2004;**169**:319–20

35 Arkwright PD, Laurie S, Super M, Pravica V, Schwarz MJ, Webb AK, Hutchinson IV. TGF-beta(1) genotype and accelerated decline in lung function of patients with cystic fibrosis. *Thorax* 2000;**55**:459–62

36 Wojnarowski C, Frischer T, Hofbauer E, Grabner C, Mosgoeller W, Eichler I, Ziesche R. Cytokine expression in bronchial biopsies of cystic fibrosis patients with and without acute exacerbation. *Eur Respir J* 1999;**14**:1136–44

37 Turner MW. The role of mannose-binding lectin in health and disease. *Mol Immunol* 2003;**40**:423–9

38 Garred P, Pressler T, Madsen HO, *et al*. Association of mannose-binding lectin gene heterogeneity with severity of lung disease and survival in cystic fibrosis. *J Clin Invest* 1999;**104**:431–7

39 Gabolde M, Giulloud-Bataille M, Feingold J, Besmond C. Association of variant alleles of mannose binding lectin with severity of pulmonary disease in cystic fibrosis: cohort study. *BMJ* 1999;**319**:1166–7

40 Buranawuti K, Cheng S, Merlo C, *et al*. Variants in the mannose binding lectin gene modify survival of cystic fibrosis patients. *Pediatr Pulmonol* 2003;**S25**:215

41 Davies JC, Johnson MM, Booth C, *et al*. Age-specific effect of the cystic fibrosis modifier gene, MBL-2. *Pediatr Pulmonol* 2002;**S23**:223

42 Garred P, Pressler T, Lanng S, *et al*. Mannose-binding lectin (MBL) therapy in an MBL-deficient patient with severe cystic fibrosis lung disease. *Pediatr Pulmonol* 2002;**33**:201–7

43 Ramsay CE, Hayden CM, Tiller KJ, Burton PR, Goldblatt J, Lesouef PN. Polymorphisms in the beta2-adrenoreceptor gene are associated with decreased airway responsiveness. *Clin Exp Allergy* 1999;**29**:1195–203

44 Taouil K, Hinnrasky J, Hologne C, Corlieu P, Klossek JM, Puchelle E. Stimulation of beta 2-adrenergic receptor increases cystic fibrosis transmembrane conductance regulator expression in human airway epithelial cells through a cAMP/protein kinase A-independent pathway. *J Biol Chem* 2003;**278**:17320–7

45 Buscher R, Eilmes KJ, Grasemann H, *et al*. beta2 adrenoceptor gene polymorphisms in cystic fibrosis lung disease. *Pharmacogenetics* 2002;**12**:347–53

46 Eggermont E. Gastrointestinal manifestations in cystic fibrosis. *Eur J Gastroenterol Hepatol* 1996;**8**:731–8

47 Rozmahel R, Wilschanski M, Matin A, *et al*. Modulation of disease severity in cystic fibrosis transmembrane conductance regulator deficient mice by a secondary genetic factor. *Nat Genet* 1996;**12**:280–7

48 Zielenski J, Corey M, Rozmahel R, *et al*. Detection of a cystic fibrosis modifier locus for meconium ileus on human chromosome 19q13. *Nat Genet* 1999;**22**:128–9

49 Duthie A, Doherty DG, Donaldson PT, *et al*. The major histocompatibility complex influences the development of chronic liver disease in male children and young adults with cystic fibrosis. *J Hepatol* 1995;**23**:532–7

50 Gabolde M, Hubert D, Guilloud-Bataille M, Lenaerts C, Feingold J, Besmond C. The mannose binding lectin gene influences the severity of chronic liver disease in cystic fibrosis. *J Med Genet* 2001;**38**:310–11

51 Mekus F, Laabs U, Veeze H, Tummler B. Genes in the vicinity of CFTR modulate the cystic fibrosis phenotype in highly concordant or discordant F508del homozygous sib pairs. *Hum Genet* 2003;**112**:1–11

52 Noone PG, Pue CA, Zhou Z, *et al*. Lung disease associated with the IVS8 5T allele of the CFTR gene. *Am J Respir Crit Care Med* 2000;**162**:1919–124

53 Groman JD, Hefferon TW, Casals T, *et al*. Variation in a repeat sequence determines whether a common variant of the cystic fibrosis transmembrane conductance regulator gene is pathogenic or benign. *Am J Hum Genet* 2004;**74**:176–9

54 Cuppens H, Lin W, Jaspers M, *et al*. Polyvariant mutant cystic fibrosis transmembrane conductance regulator genes. The polymorphic (Tg)m locus explains the partial penetrance of the T5 polymorphism as a disease mutation. *J Clin Invest* 1998;**101**:487–96

55 Dork T, Wulbrand U, Richter T, *et al*. Cystic fibrosis with three mutations in the cystic fibrosis transmembrane conductance regulator gene. *Hum Genet* 1991;**87**:441–6

56 Romey MC, Guittard C, Chazalette JP, *et al*. Complex allele [-102T>A+S549R(T>G)] is associated with milder forms of cystic fibrosis than allele S549R(T>G) alone. *Hum Genet* 1999;**105**:145–50

57 Morgan TM, Coffey CS, Krumholz HM. Overestimation of genetic risks owing to small sample sizes in cardiovascular studies. *Clin Genet* 2003;**64**:7–17

58 Armstrong DS, Grimwood K, Carlin JB, Carzino R, Olinsky A, Phelan PD. Bronchoalveolar lavage or oropharyngeal cultures to identify lower respiratory pathogens in infants with cystic fibrosis. *Pediatr Pulmonol* 1996;**21**:267–75

59 Gill D; Ethics Working Group of the Confederation of European Specialists in Paediatrics. Ethical principles and operational guidelines for good clinical practice in paediatric research. Recommendations of the Ethics Working Group of the Confederation of European Specialists in Paediatrics (CESP). *Eur J Pediatr* 2004;**163**:53–7

60 Romieu I, Sienra-Monge JJ, Ramirez-Aguilar M, *et al*. Genetic polymorphism of GSTM1 and antioxidant supplementation influence lung function in relation to ozone exposure in asthmatic children in Mexico City. *Thorax* 2004;**59**:8–10

61 Riordan JR, Rommens JM, Kerem B, *et al*. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 1989;**245**:1066–73

62 Boyle MP. Nonclassic cystic fibrosis and CFTR-related diseases. *Curr Opin Pulm Med* 2003;**9**:498–503