Causes of Variation in the Cystic Fibrosis Phenotype

Garry R. Cutting
Institute of Genetic Medicine, Johns Hopkins Hospital, Baltimore, Md., USA

Introduction

Cystic fibrosis (CF), a single gene disorder that affects 60,000 individuals worldwide, is caused by defects in the CF transmembrane conductance regulator (CFTR) gene. It is observed primarily in individuals of European Caucasian descent with an incidence ranging from 1 in 2,000 to 1 in 4,000 [1]. CF is much less common in native African and native Asian populations. Some Caucasian groups have a higher frequency of CF than others likely due to a founder effect [2, 3]. The cause of the high frequency of carriers of CF mutations in Caucasians (1 in 30) is not known, although heterozygote advantage is a favored explanation [1]. Any implicated mechanism needs to take into account that CF in Caucasians is due to the presence of a common CF-causing mutation (ΔF508; see below) will likely also explain the commonness of CF in Caucasians. Now that we know the etiology of CF, the current challenge is to determine why patients have variable outcomes. This review provides a summary of our current understanding of the role of CFTR variation and presents mounting evidence on the importance of factors independent of CFTR in determining disease severity in CF patients.
Variation in the CF Phenotype

When CF was first described as a distinct entity in the 1940s, most patients died from this disease before their first birthday due to malnutrition caused by a dysfunction of the exocrine pancreas. Treatment of the exocrine pancreatic insufficiency, in particular the development of enzyme replacement therapy, enabled patients to survive beyond their first decade. Median survival has progressively increased since the 1940s and CF patients currently have a median life expectancy of 36.2 years in the United States. Despite impressive increases in median survival, there continue to be wide variations in the age at death of CF patients. Some CF patients die before their 10th birthday, while others live into their 50s and 60s. Lung disease is currently the major cause of morbidity and mortality in CF and is currently the cause of 90% of all deaths due to CF [4]. The severity of pulmonary disease varies considerably among CF patients. For example, some patients develop an obstructive pattern in the first few years of life complicated by an infection with Pseudomonas aeruginosa, fungal organisms and, later in life, atypical mycobacteria. Other patients experience relatively few symptoms of lung disease until their 20s and then present with an episode of hemoptysis followed by a rapid downhill course leading to lung transplantation within a few years. Most patients with CF also suffer from chronic rhinosinusitis and a subset of patients develop nasal polyps.

Pancreatic abnormalities are present in almost all patients with CF. Most patients have severe exocrine deficiency requiring enzyme supplementation to attain sufficient fat and carbohydrate absorption and avoidance of vitamin deficiencies. A small fraction of CF patients (5–10%) have adequate fat and carbohydrate absorption without enzyme supplementation and are designated as pancreatic sufficient. The latter group of patients generally has a milder course with less severe lung disease, milder abnormalities in the sweat chloride concentration and fewer gastrointestinal complications [5]. However, some manifest severe pulmonary disease requiring lung transplantation. The endocrine pancreas is also affected in CF. As more CF patients survive into adulthood, it has been recognized that a large fraction develop glucose intolerance and 30% of CF patients are diagnosed with diabetes by 35 years of age [6]. The development of diabetes has been associated with more severe pulmonary disease and reduced longevity [7].

Other gastrointestinal manifestations of CF include liver disease and intestinal obstruction. Most CF patients have some degree of a hepatobiliary disease at autopsy, and a subset of patients (approximately 5%) develop severe liver disease including cirrhosis, portal hypertension, and esophageal varices [8]. The development of these features is usually associated with reduced survival; about 2% of patients die due to complications associated with hepatic disease [4]. Intestinal obstruction in the neonatal period, termed meconium ileus (MI), is a well-recognized feature that affects approximately 15% of CF patients [4]. Modern pediatric surgical techniques have made this complication fatal only in the rarest of cases. In some situations, the obstruction can be relieved by a gastrograffin enema. Distal intestinal obstruction syndrome (DIOS) affects older children and adults with CF. It has been estimated that between 7 and 40% of CF patients experience an episode of DIOS during their lifetime [9]. MI and DIOS have similar clinical features, and MI patients have been reported to have a higher rate of DIOS leading to the suggestion that they share a common etiology.

An elevation in the concentration of sodium and chloride in sweat is a distinct biochemical feature of CF. An abnormality in the chloride concentration has greater specificity for CF and has been used as a diagnostic tool for over 4 decades. The range of sweat chloride concentrations varies amongst normals and CF patients. Some healthy adults can have sweat chloride concentrations in the range of 40–60 mmol/l. This concentration of chloride in the sweat of an infant or young child, however, is consistent with the diagnosis of CF. Most CF patients have sweat chlorides of >60 mmol/l with an average of about 105 mmol/l. Abnormalities in the ionic concentration of sweat can lead to excessive salt loss and hyponatremic dehydration in a small fraction of CF patients.

The vast majority of men with CF (98%) are infertile due to abnormalities in the development of structures derived from the Wolffian duct. This includes malformation or the absence of the vas deferens and abnormalities in the seminal vesicles. Sperm motility and morphology has also been reported to be abnormal, although sperm acquired by testicular aspiration have produced viable conceptions.

Contribution of CFTR Genotype

The CFTR gene encodes a protein of 1,480 amino acids that functions as a cAMP-regulated chloride channel and as a regulator of other ion channels. CFTR is responsible for regulating ion and fluid transport across the apical
membranes of epithelial cells and plays a central role in the absorption and secretion in the epithelial lining of the pulmonary airways and the pancreatic ducts. Dysfunction of CFTR causes changes in the fluid and salt content of airway surface liquid and pancreatic secretions leading to compromised defense mechanisms and obstruction that eventually destroys both organs. Over 1,400 mutations in the CFTR gene have been reported to the CF Genetic Analysis Consortium (www.genet.sickkids.on.ca/cftr); 1,200 of these mutations have been associated with disease while the remaining 200 appear to be benign variants. Despite this extreme allelic diversity, one mutation, a deletion of phenylalanine at position 508 (termed ΔF508), accounts for 70% of CF alleles worldwide [10]. The ΔF508 mutation is found almost exclusively in European Caucasians and occurs in increasing frequency from the Mediterranean basin to the northern portions of Europe [10]. Approximately 5–20 other CFTR mutations account for 10–20% of CF alleles [10]. The distribution of the ‘less common’ mutations varies depending on the population. The remainder of the CFTR mutations are rare with the exception of a handful that have achieved a high frequency in isolated populations due to a founder effect. The high frequency of the ΔF508 mutation results in half of all CF patients having the same CFTR genotype (i.e. ΔF508/ΔF508). The ΔF508 homozygotes have been studied extensively and this genotype is almost invariably associated with pancreatic insufficiency [11]. These patients have high sweat chloride concentrations and all males with this genotype are infertile. However, the severity of lung disease in ΔF508 homozygotes varies considerably [11, 12].

The large pool of CF patients with an identical CFTR genotype (ΔF508 homozygosity) provides a reference population to assess the effect of other CFTR mutations upon the CF phenotype. Genotype/phenotype studies have revealed that the CFTR genotype is highly correlated with the severity of pancreatic disease [11]. Most mutations are associated with pancreatic insufficiency, however, there are about 2 dozen mutations associated with pancreatic sufficiency. These mutations impair but do not abolish CFTR function [13]. Thus, residual pancreatic function appears to be the result of partial CFTR function. However, this association is not absolute; some patients bearing mutations that permit some CFTR function have pancreatic insufficiency [13]. There are also mutations that produce specific abnormalities of pancreatic function such as the L997F mutation that is strongly associated with pancreatitis [14]. Complications in the gastrointestinal tract, such as MI, generally occur in patients with mutations that cause pancreatic insufficiency. However, patients bearing the G551D mutation, a mutation that causes pancreatic insufficiency, are associated with a lower rate of MI than patients who are ΔF508 homozygotes [15, 16]. While pancreatic sufficiency has been associated with less severe lung disease, it has been difficult to find specific mutations that confer both pancreatic sufficiency and mild lung disease. One example of the latter is the A455E mutation. Two studies have shown that CF patients carrying A455E have slower progression of their lung disease than patients who are homozygous for the ΔF508 mutation [17, 18].

The CFTR genotype moderately correlates with the degree of sweat chloride abnormality and this generally occurs in the context of pancreatic sufficiency. In other words, patients bearing mutations with partial function that confers pancreatic sufficiency also tend to have less abnormal elevations in the sweat chloride concentration [19]. Finally, abnormalities in the reproductive tract of CF males are closely associated with the CFTR genotype. The splice mutation 3849+10kbC→T is associated with a high rate of fertility in CF males despite causing pulmonary features consistent with CF [20].

Two CFTR mutations demonstrate how subtle the association between genotype and phenotype can be. One mutation that causes CF (R117H) is also a common mutation in men with isolated congenital bilateral absence of the vas deferens (CBAVD). The latter condition has identical features to those found in men with CF, but men with CBAVD do not develop life-limiting lung disease even though they carry mutations in each CFTR gene. The difference between men with CF and men with CBAVD who both carry the R117H mutation is that the former have a second mutation in the gene. [21]. Thus, the presence of a second alteration (called 5T) that reduces the amount of CFTR protein produced causes men with CBAVD to also develop life-limiting lung disease. Intriguingly, 5T causes CF disease by itself when it is associated with an abnormality in length of an adjacent sequence variant called the TG tract. Males who have a longer TG tract associated with the 5T variant and another mutation in the CFTR gene (such as ΔF508) generally have CBAVD. Those with the same combination of mutations (e.g. ΔF508 and 5T) who have the more common short TG tract are generally free of abnormalities [22]. This association between CFTR genotype and CF phenotype requires that the multiple alterations in CFTR be considered as contributors to outcome.
Contribution of Genetic Modifiers

Although there is some evidence of an association between CF mutations and lung disease (e.g. A455E and R117H described above), for the most part CFTR genotype does not predict lung disease severity. This is primarily due to the wide variability of lung function observed in CF patients with identical CFTR genotypes [12]. This observation indicates that factors independent of CFTR must contribute substantially to variation in pulmonary disease.

One of the most effective methods to determine the role of genetic factors upon disease variability in humans is to study families with multiple affected children. Siblings share 50% of their genes. Thus, the recurrence of complications among siblings at a higher rate than seen in unrelated patients can implicate genetic factors. One example is MI. If the first sibling in a family has CF and MI, then each subsequent sibling with CF has a 25% risk of also having MI. The recurrence risk for siblings is substantially higher than in unrelated patients (about 15%). Other examples of complications with high recurrence risks and disease features with high similarity among siblings include pancreatic function, liver disease and infections with Pseudomonas [23–26]. However, siblings also share their environment (household, school, clinic, etc.), thus higher recurrence rates among siblings do not necessarily implicate genetic modifiers. A formal method to test for the effect of genetic versus non-genetic modifiers is to study affected twins. Comparison of monozygous (100% gene sharing) and dizygous twins (50% gene sharing) permits estimates of genetic effects (i.e. heritability). Using this approach, investigators have shown that a measure of body mass (height for weight) is primarily determined by genes [27]. Genetic factors independent of CFTR also appear to influence intestinal function in CF patients by altering chloride secretion [28, 29] and may modify growth in CF patients [30].

One method to identify specific genes that modify the CF phenotype is to select candidate genes based upon knowledge of CF pathophysiology. These candidate genes encode proteins that play a role in disease progression (e.g. inflammatory mediators). DNA variants in these candidate genes are identified and the distribution of the DNA variants is determined in CF patients stratified according to disease severity. Finding that variants of a candidate gene are associated with more severe disease implicates the candidate as a modifier. While this approach is powerful, it is also fraught with the problem inherent in dividing patients into different groups. Genetic variants can differ by chance or due to unrecognized differences among the two groups, affecting the distribution of variants that are not correlated with disease severity. Thus, case-control association studies tend to produce a number of false-positive candidate modifier genes. One approach to reduce false positives is to test candidate genes in different patient populations. Using this approach, Drumm et al. [31] tested 10 candidate genes that had previously been shown to modify disease severity in CF patients in a multicenter study of 808 CF patients recruited on the basis of lung disease severity. Of these 10 genetic variants, in only one ‘modifier’ gene (transforming growth factor-β1, TGFβ1) was associated with more severe lung disease. The authors confirmed their finding by testing the TGFβ1 gene in a second group of 498 CF patients [31]. Variants in the TGFβ1 gene have been associated with a susceptibility to asthma and protection from the development of chronic obstructive pulmonary disease in smokers, supporting the concept that CF modifier genes may also play a pathologic role in common diseases [32–35]. Variants of two other genes (tumor necrosis factor-α and mannose-binding lectin) have been associated with the severity of CF lung disease independent of CFTR genotype in more than one group of CF patients [36–40]. These two genes were included in the replication study of Drumm et al. [31] and did not show an association with lung disease severity. However, modifier genes that have a real but small effect can be difficult to detect in every study. Subtle modifier effects upon the phenotype can be masked by variation due to non-genetic factors such as a difference in treatment among centers [41]. Therefore, it is not surprising that a number of candidate modifier genes have shown an association in only one study or have not shown a consistent association with disease severity [42].

Over the next few years, it is likely that new studies will confirm or refute the role of a number of genes in modifying the CF phenotype. In addition, it is expected that other approaches to identifying modifier genes such as positional cloning will reveal genes that would not have been selected based on the candidate gene approach. These genes might have different functions than those envisioned for modifiers but, by virtue of their genetic correlation with disease, they can be proven to be modifiers of CF. This approach was used to identify CFTR as the gene responsible for CF. Family-based studies are ideally suited to the latter approach. Siblings that are affected with CF can be grouped according to the degree of similarity for a measure of disease severity (e.g. lung function). Siblings are then tested for correlations be-
tween the degree of sharing a genetic region and the degree of similarity in disease severity. Thus, siblings who both have severe lung disease would be tested for sharing modifier genes that cause severe lung disease. Siblings who are different; one with severe disease and the other with mild disease, would not be expected to have a modifier gene in common. This method, called concordance, has been applied to CF siblings to demonstrate a role for a region on chromosome 19 that may contain a modifier for MI [43]. A powerful extension of family-based methods is to test parents and their CF-affected offspring for transmission of modifier genes. In this situation, one determines if a variant of a modifier gene was transmitted more often than predicted by chance to offspring grouped according to disease severity. This method, called transmission disequilibrium testing (TDT) is not subject to errors due to population stratification that plague association studies are. However, both parents have to be tested for TDT to be maximally effective. Obtaining DNA from all family members can be difficult especially for older patients. Nevertheless, the family-based studies in Europe, Canada, and in the US have patient collections suitable for TDT.

**Role of Environmental Modifiers**

The improvement in the survival of CF patients over the past 5–6 decades cannot be attributed to modifier genes. Changes in treatment, especially the development of pancreatic enzyme replacement therapy and the creation of multidisciplinary CF care centers has produced substantial changes in outcome for CF patients (see chapter by B. Strandvik ‘Care of Patients with CF’ in this issue). These ‘modifying factors’ are environmental in nature. Identification of specific environmental exposures that affect outcome provides a powerful method to find factors that might easily be manipulated. Assessing the environmental contribution to phenotype has been challenging due to the variety of components involved and the lack of objective measurements for many components. Despite these limitations, passive cigarette smoking has been linked with reduced pulmonary function [44–46]. Elimination of passive smoking for CF patients is a relatively straightforward measure to improve outcome. Aggressive nutritional intervention has also been shown to affect outcome [47–50]. Interestingly, patients vary in their response to nutritional treatment suggesting that other environmental and/or genetic factors may play a role [51]. Pulmonary colonization with *P. aeruginosa* and *Burkholderia cepacia* is an environmentally mediated event [52, 53] that is associated with reduced longevity [54–57]. Again, changes in environmental factors, such as infection control, can prevent the spread of virulent subtypes of these bacteria.

Improving outcome by reducing variation in treatment among care centers so that all patients receive optimized care is a mainstay of current CF treatment. Patient advocacy foundations have utilized their resources and organizational abilities to develop multidisciplinary care

**Fig. 1.** Causes of variation in the major manifestations of CF. The relative contribution of CFTR genotype, modifier genes and environmental factors are estimated for each of the organ systems affected in CF patients. CFTR genotype is the primary determinant of the severity of pancreatic disease and malformation of the male reproductive structures. Modifier genes and environmental factors play important roles in the severity of lung disease.
centers that enable delivery of optimal and uniform levels of care. Recent efforts in the US have been directed at improving the nutritional status of CF patients with the understanding that poor nutritional status is associated with worse lung disease and worse outcome. Finally, compliance is clearly a major issue for patients with CF as with patients with any chronic disorder. Even if a therapy that altered the function of the CFTR protein were developed, successful treatment could hinge upon the willingness of the patient to adhere to a lifelong therapeutic regimen. Thus, emphasis on health behaviors and adherence to routine care and anticipatory management are likely to yield continued improvements in outcome for CF patients.

**Summary**

There has been amazing success in the treatment of CF, a condition that was once frequently fatal in the first year of life. Recognition of the key factors that determine outcome in the individual patient and the development of treatments to address these factors are improving longevity for all CF patients. Identification of the CFTR gene was a key step in understanding pathophysiology at a molecular level and establishing the degree to which variation in CFTR function influences the outcome of CF. These studies revealed that other modifiers, both genetic and non-genetic, play an important role in determining disease severity (fig. 1). The challenge for the next decade will be to identify and characterize those factors so that outcome is optimized for all patients with CF.

**References**

Causes of CF Variation


30 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.


VAKAT
Abstract
Sorry, there is no abstract. Read the first few lines of the text instead!