Cystic Fibrosis and CFTR Gene

Pascale Fanen, Afia Hasnain

Service de Biochimie-Génétique, Inserm U.654, Hôpital Henri Mondor, 94010 Créteil, France (PF, AH)

Published in Atlas Database: September 2001

Online updated version: http://AtlasGeneticsOncology.org/Educ/CistFibID30032EL.html

DOI: 10.4267/2042/37827

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 2.0 France Licence.

© 2002 Atlas of Genetics and Cytogenetics in Oncology and Haematology

I. Background

The early description of cystic fibrosis (CF) dates back to late 30s. In 1936, Fanconi identified the association between the congenital CF of the pancreas and bronchectasis shortly followed by Andersen who in 1938 gave the complete anatomo-pathologic description of CF.

In 1953, Di Sant’Agnese described an excess of sodium chloride in the sweat of children affected by CF. This discovery shortly leads to the use of sweat chloride test, the only reliable diagnostic test of the disease available to this date.

In early eighties, the abnormality of transport of salts was precisely described by Quinton (Quinton 1983) who explained the defect of permeability of chloride ions (Cl-) in the affected epithelial cells of sweat glands, and later by Knowles (Knowles 1983) who observed the same phenomenon in the respiratory epithelium.
II. Incidence

The CF is the most common lethal autosomal recessive hereditary disorder, worst in the European origin populations, affecting an average one out of 2500 live births (= q2), i.e., one out of 25 (= 2pq) individuals is a carrier of this disease. However, this frequency varies according to geographic and ethnic origin of the patients. In Europe, depending on different regions, one child out of 1800 to 3500 live births is affected. In France alone, there are more than two millions carriers. In France, approximately 250 new cases are reported every year. Or in other words between 4000 and 6000 individuals are affected in our country.

III. Clinical manifestation

The clinical presentation of CF varies among individuals from different families as well as between individuals belonging to the same family. In majority of cases the disease is diagnosed before adolescence, but some remain asymptomatic till the adult age. Clinically and biologically, it is not possible to distinguish the heterozygotes from the individuals not carrying CF mutation.

The clinical presentation varies with age. In 10% of the affected newborns, it presents as meconium ileus (intestinal obstruction due to abnormally thick meconium). Later the symptomatology involves two major organ systems, respiratory and digestive, manifesting as repetitive respiratory infections and signs of malabsorption.

The respiratory involvement predominates. It is related to an obstruction of bronchioles by thick and viscous mucous favoring the growth of microorganisms. This explains the repeated respiratory infections by the opportunistic germs, initially Staphylococcus aureus and Haemophilus influenzae and later Pseudomonas aeruginosa, bacteria naturally resistant to a number of antibiotics that slowly become sole pathogenic agents of respiratory airways.

The digestive involvement with exocrine pancreatic insufficiency (defect in the production of enzymes) is seen in 85% of the patients, leading to the obstruction of pancreatic ducts and thus lipoprotein malabsorption. The state of exocrine pancreas, PI (pancreatic insufficiency) or PS (pancreatic sufficiency), serves as a marker of the gravity of the phenotype, being serious in the former than the latter (Kerem 1990).

The abnormality in the sweat glands leads to an excess of sodium chloride secreted in the sweat, this loss of salt is usually responsible for acute dehydration in case of exposure to heat. This disease involves some other organs particularly the reproductive system and the liver.

Atrophy and absence of vas deferens caused by obstructive azoospermia renders 98% of men sterile, on the other hand 80% of affected females are fertile. The sterility in females is explained by an abnormality in cervical mucus.

The hepatic involvement in 30% of cases starts as a hepatomegaly and in 9% of cases as a hepatic insufficiency. It is due to the obstruction of the intrahepatic or extrahepatic biliary tracts by compression at the level of pancreas. In 2-5% of cases this further leads to biliary cirrhosis.

If not treated, the median survival rate is 3 to 5 years. There is no effective treatment of CF, but the timely diagnosis and symptomatic treatment has made it possible to increase the mean survival to 25 to 30 years. The symptomatic treatment is available in the form of respiratory physiotherapy, antibiotic therapy, nebulisation with bronchodilators and mucolytics, administration of proteases inhibitors for pulmonary symptoms and administration of substitute pancreatic enzymes and vitamins for pancreatic insufficiency.

In advanced stages, the triple (heart-lung-liver) transplant has shown promising results but the major hindrance is the availability of donor organs.

Among the latest therapeutic advances, the gene therapy of CF, with the goal of administering the genetic sequence into the respiratory epithelial cells for the expression of CFTR wild type, has met a number of obstacles:

i) expression of transgene at a very low level than in the in vitro and in vivo models;
ii) transient expression;
iii) impossible readministration;
iv) no improvement in the symptoms related to infection, inflammation and obstruction of airways, or in the respiratory function; however, to date 20 clinical trials are held.

Other promising strategies are under study with the objective of compensating for the defect in the production and/or function of the CFTR protein depending on the type of mutation.

IV. Diagnosis

The diagnosis of CF depends on the sweat chloride test. Even today it is considered to be the most reliable test for CF. The techniques used these days are simple but the quantity of sweat required is a limiting factor especially in neonates raising the risk of error to 30%. The values of this test can vary according to laboratories; a value of less than 40 mmoles/l of chloride is generally considered as normal.

For the range between 40 and 60 mmoles/l, the interpretation is doubtful and the test should be repeated.

For the diagnosis to be established, the sweat chloride concentration must exceed 60 mmoles/l on two or more separate sweat tests. There are other supportive tests used particularly in new borns. The oldest being the BM test (levels of albumin in the meconium) is less reliable. The level of immunoreactive trypsine in the
plasma is of great value in case of a doubtful diagnosis in newborns. The hypertrypsinemia is not however pathognomonic of CF; it is also seen in transient neonatal hypertrypsinemia (associated to hypoxia and birth asphyxia) or persistent (associated with trisomies 13, 18 or 21, renal insufficiency, and pancreatic, hepatic or intestinal pathologies). Also, with positive result, and especially in case of no other pathology detected, only the sweat chloride test is reliable. Another technique consists of measuring the bioelectrical potential difference that exists between the skin and nasal mucosa. This value is significantly raised in case of CF. This technique was described by Knowles (Knowles 1981), and later simplified by Alton (Alton 1987; Alton 1990). It is valuable in following cases:
1. For the early diagnosis in the neonates presenting with a suspected digestive pathology, where the sweat test is very difficult to be carried out,
2. For the confirmation of a doubtful diagnosis with some clinical signs and a borderline or negative sweat chloride test,
3. For the patient follow-up, as there is a correlation between the seriousness of respiratory involvement measured by FEV1 study (forced expiratory volume in first second) and the value of bioelectrical potential difference. The method is simple, less expensive and accessible. The test is well tolerated but the results are difficult to interpret in cases of inflammation of nasal mucosa or in the presence of nasal polyps.

V. CFTR gene and its mutations

V.1. Introduction
The biochemical nature of the basic defect in CF remained obscure for a very long time, which forced the researchers to employ the technique of reverse genetics (or positional cloning) rather than classical genetics.

In 1985, CF locus was localized on the long arm of chromosome 7.

In 1989, the gene implicated in CF was isolated by (Kerem 1989; Riordan 1989; Rommens 1989). The genetic analysis showed that this gene, which is responsible for this disorder, contains 27 exons spanning over 250 kb of chromosome 7 (7q31) and encodes an mRNA of 6.5 kb.

V.2. The DF508 mutation
The most frequent mutation, a deletion of three nucleotides resulting in the deletion of phenylalanine on position 508 (DF508) is responsible for 70% of CF alleles. The description of this mutation reinforces the role of this gene in causing CF, as it was never found in a normal individual in a homozygous state.
A European collaborative study (EWGCFG 1990) comprising of 4871 CF chromosomes and 3539 normal chromosomes showed the great heterogeneity of this anomaly. There is a north-west/south-east gradient, for example 88% of DF508 cases found in Denmark and 50% in Italy. In the French population, the principal mutation (DF508) represents approximately 65-70% of CF chromosomes, with strong regional variations being 64% in Languedoc-Roussillon to 81% in occidental Britannia. This elevated frequency suggests the possible existence of a selective advantage of heterozygotes in the north-European populations. The analysis of genetic markers associated with DF508 suggests that only one mutational event has occurred in the past (Morral 1994). To explain the spread of this mutation in European population, the hypothesis of a selective advantage of heterozygotes was proposed. For example, the heterozygotes would be protected against dehydration due to diarrhea caused by enterotoxins of Escherichia coli and of Vibrio cholerae (Baxter 1988).

V.3. Spectrum of CFTR gene mutations
Since the discovery of CFTR gene, a number of laboratories have characterized the different mutations responsible for CF. They are united under a network of exchanging information, created by the discoverers of this gene, the International Consortium for the Genetic Analysis of CFi (http://www.genet.sickkids.on.ca/cftr/).
It has rapidly established a spectrum of these mutations found in different populations studied all over the world.
The genetic studies carried out before the discovery of this gene showed a very few mutations, in contrast to what was found after the discovery of this gene. More than 1200 mutations are described since the cloning of this gene, out of which 4 (excluding DF508) represent more than 2%. All the other mutations are rare, mostly found in a single family. The frequency of certain mutations can vary among different geographic groups. This is how the nonsense mutation W1282X makes 48% of CF alleles among the Ashkenazi Jews and only 2% of the total CF alleles (Kerem 1995). Similarly, the frequency of the G551D mutation represents approximately 5% of
alleles in the populations of Celt origin (Ireland, Écosse, Britain) (Hamosh 1992).
The majority of molecular defects of CFTR gene are the point mutations out of which 42% are missense mutations, 24% small insertions/deletions with a frame shift, 16% nonsense mutations, 16% mutations of splicing and 2% deletion of an amino acid. Some major deletions are also reported.
One of the particularities of CFTR gene is the existence of deleted transcripts of one or more exons among normal individuals. These transcripts are due to anomalies leading to alternative splicing out of which the most frequent and well studied is the deleted transcript of exon 9 \(9^-\). The presence or absence of this exon is correlated with a "polymorphism" of sequence of the intron 8 situated near the acceptor site of splicing. This polypyrimidic sequence constitutes of 5, 7 or 9 thymidines \(5T, 7T\) or \(9T\). If the presence of \(7T\) or \(9T\) reassures a normal splicing to 90%, the presence of \(5T\) doesnít produce more than 10 to 40% of normal mRNA, the 60 to 90% of mRNA 9- cannot give a functional CFTR (Chu 1991; Chu 1992; Chu 1993).

### V.4. Genotype-phenotype correlations
Approximately half of the patients affected by CF are homozygous for the mutation DF508. DF508 homozygotes present a classical form of the disease with an increase in electrolytes in sweat, pancreatic insufficiency and obstructive pathology of lungs. Comparing the clinical presentation of patients homozygous for DF508 with the patients of different genotypes gives the phenotypic consequences of these other mutations. Having said that the DF508 accounts for 66 % of the mutations, 40 % of these patients are compound heterozygotes with DF508 on one allele and another mutation of CFTR gene on the other chromosome.

Generally, the pulmonary function, the age of onset of the disease, and the amount of chloride salt are all related to a particular genotype. On the other hand, the seriousness and variety of symptoms differing in the same family explain that the genotype, at the level of CFTR gene, cannot be solely held responsible for the phenotype (Zielenski 2000). The only exception being the pancreatic function that is identical in the affected members of the same family (Corey 1989).

The mutation A455E was strongly associated with the state of pulmonary function. In Holland, the mutation A455E is relatively frequent. The analysis of 33 compound heterozygous patients DF508 / A455E revealed the improved pulmonary function tests and a reduced level of colonization by P. aeruginosa as compared to homozygous DF508 individuals from the same population (Gan 1995). The results suggest that the mutation A455E gives rise to a less severe form of lung involvement than the DF508 in the patients in Holland. A similar study was carried out on 9 Canadian compound heterozygotes DF508 / A455E, revealed a less severe form of lung involvement than the 5 homozygotes DF508 from the same population (De Braekeeleer 1997). The studies show that A455E produce a moderate form of lung involvement and the A455E has a dominant effect on the severe alleles than DF508.

Concerning the pancreatic function, a Canadian study showed that the patients with one or two missense mutations R117H, R334W, R347P, A455E or P574H have a conserved exocrine pancreatic function (PS) or pancreatic sufficiency, as compared to those having two alleles of splicing mutations, nonsense or frame shift mutations and some mutations of missense which always lead to pancreatic insufficiency (PI or pancreatic insufficiency) (Kristidis 1992). The mutations associated with a normal pancreatic function are considered as moderate and those associated with pancreatic insufficiency as severe. Similarly, the patients having a PI mutation on one allele and a PS mutation on the other have a phenotype PS. With a PS mutation, the activity of the CFTR protein is sufficient for the pancreatic function. However, multi centric study showed that out of 396 homozygotes DF508, 10 conserve a pancreatic function (Consortium, 1993). This type of analysis is complicated by several phenomena. The effect of one mutation can be modified by a second mutation inherited in cis on the same allele. Two cases have been described; the polymorphism of the polypyrimidic tract of the intron 8, and the association of two nucleotide substitution on the same allele.

The group of Tümmler at Hanover described a patient carrying a complex genotype R553X / DF508-R553Q (Dork 1991). The patient presents a pancreatic sufficiency associated with a typical pulmonary involvement but an abnormally low sweat test compared to 9 patients R553X / DF508, suggesting that the mutation R553Q can modify the effect of DF508. This hypothesis was partly valid in vitro (Teem 1993). This same group also described two unrelated patients carrying this genotype: DF508 / S1251N-F508C. They proposed that the polymorphism F508C could aggravate the effect of S1251N (Kalir 1992). The polymorphism of the polypyrimidic tract of intron 8 modifies the penetration of the missense mutations R117H of the exon 4. This mutation was not only associated with a moderate (PS) phenotype (Kristidis 1992), but also with a congenital bilateral absence of vas deferens (CBAVD) (Rigot 1991).
In 1993, the group of Cutting showed the importance of polymorphism of the intron 8 in the modification of phenotype of a patient (Kiesewetter 1993). In fact, as the mutation R117H is related in cis with allele 5T or 7T, the same genotype R117H / DF508, is associated either to a moderate CF or to a CBAVD and in extreme cases may not have any consequences.

VI. The CFTR protein and its functions

VI.1. Structure of the CFTR protein

The protein sequence of the CFTR is composed of 1480 amino acids. The analysis of its primary sequence allows constituting the possible tertiary structure of CFTR protein with the hydrophilic profile of its amino acids (Riordan 1989). It consists of two repeated motifs each composed of a hydrophobic membrane-spanning domain (MSD) containing six helices and an important hydrophilic region for binding with ATP (NBF or Nucleotide Binding fold). These two motifs are linked by a cytoplasmic (R domain) encoded by exon 13, containing a number of charged residues and the majority of the phosphorylation sites (probable substrates of protein kinases A and/or C) (figure 1). Homologies exist between the primary structure of CFTR protein and members of membrane protein families, the family of ABC transporters (ATP-binding cassette). The proteins belonging to the family of ABC transporters are responsible for the active transport of substrates across the cell membrane, where ATP hydrolysis serves as the source of energy.

VI.2. Cl- channel Function

The early hypotheses regarding the function of CFTR protein revolved around two possibilities. The first postulated that the CFTR protein is a Cl- channel. This hypothesis was compatible with the defect in permeability of Cl- ions at the CF epithelial apical membranes. The other proposed that the CFTR protein is not an ionic channel but it plays a role in the regulation of Cl- channel either by associating with them, or by transporting them, in or out of the cell, as a regulatory factor for Cl- channels. The latter, seemed more realistic in the light of following observations:
- the phenotypic abnormality seen in CF epithelium, particularly the sodium absorption by the respiratory epithelium. It seemed difficult to justify multiple phenotypic abnormalities to a single Cl- channel.
- the primary sequence of the CFTR protein didn’t resemble any other ion channel known so far.

In the preliminary functional studies, the cDNA of the CFTR wild type was expressed in the respiratory (Rich 1990) and pancreatic epithelial cells (Drumm 1990) in the patient’s homozygote for the mutation DF508. The expression of CFTR wild type gene restores the permeability of Cl- regulated by cAMP of the deficient cells contrary to the introduction of cDNA of CFTR DF508, which doesn’t correct the defect (Rich 1990). These results put on evidence the causal relation between the mutations of CFTR gene and the CF phenotype. This came as the first proof that the CFTR gene is responsible for CF. But these results didn’t identify the function of the CFTR protein. In addition, they suggested that the CFTR protein was not itself a Cl-channel, its presence in the cell allows the activation of endogenous Cl- channel by cAMP. The later results determined the function of CFTR protein as a Cl-channel.

The expression of CFTR wild type gene in the heterologous system has shown that indeed CFTR works as a Cl- channel (Anderson 1991; Bear 1991; Berger 1991; Drumm 1991). The CFTR protein was purified from the cells sf9, reconstituted in the proteoliposomes and fused to the lipid bilayer (Bear 1992). In the acellular system, the properties observed were the same as in the secretory epithelial cells. This reconstitution represented the ultimate prove that the CFTR protein is at least a Cl- channel.
Figure 1: proposed structure CFTR protein, by Riordan et al. 1989 - Pascale Fanen.

Figure 2: CFTR, a multi functional protein, by Schwiebert et al. 1999 - Pascale Fanen
VI.3. Correlation of CFTR gene mutations with the Cl- channel function

The molecular anomalies have variable effects on the CFTR protein and its functions. Welsh and Smith have proposed a classification of these anomalies in relation to the Cl- channel function (Welsh & Smith 1993) (figure 3).

VI.3.1. Class 1: mutations altering the production of the protein.

These mutations result in the total or partial absence of the protein. This class includes the nonsense mutations and those that produce a premature stop codon (anomalies of splicing and frameshift mutations). In certain cases (R553X), the mutated mRNA is unstable and doesn't produce the protein (Hamosh 1991). In other cases, the abnormal protein produced will probably be unstable and degrade rapidly. This is what produces the truncated protein or the protein containing the aberrant sequence (anomalies of splicing or the frameshift).

Functionally, these mutants are characterized by a loss of conductance of Cl- channel in the affected epithelia.

VI.3.2. Class 2: mutations altering the cellular maturation of the protein

A number of mutations alter the maturation of the protein and thus the transport of these proteins to the plasma membrane. In this way, the protein is either absent from the plasma membrane or present in a very small quantity. The mutations of this class represent the majority of CF alleles (DF508).

VI.3.3. Class 3: mutations disturbing the regulation of Cl- channel

These mutations are frequently situated in the ATP binding domain (NBF1 and 2). The mutation G551D is an example, which has very low residual function.

VI.3.4. Class 4: mutations altering the conduction of Cl- channel.

Certain segments of membrane spanning domains participate in the formation of an ionic pore. The missense mutations situated in these regions produce a correctly positioned protein that has a cAMP dependant Cl- channel activity. But the characteristic of these channels is different from those of endogenous CFTR channel with a diminution of ion flux and a modified selectivity.

Since the classification by Welsh and Smith, other classes have been proposed to explain the biochemical defects associated with diverse mutations. The class I can be further divided into a class V comprising the mutations altering the stability of mRNA, and a class VI comprising of mutations altering the stability of mature protein (Haardt 1999).

VI.4. The CFTR protein, a multifunctional protein

The discoverers of CFTR gene termed it the transmembrane conductance regulator. In fact, the CFTR protein regulates other channels also, the outwardly rectifying chloride channel (ORCC), epithelial Na+ channel (ENaC) and at least two inwardly rectifying K+channels ROMK1 and ROMK2. Besides being a channel regulator, it also plays a role in transport of ATP, modifying the phenomenon of exocytosis/endocytosis, regulation of pH of intracellular organelles. The figure 3 illustrates this concept of multifunctional CFTR protein.

References


Cystic Fibrosis and CFTR Gene

Fibrosis transmembrane conductance regulator corrects defective chloride channel regulation in cystic fibrosis airway epithelial cells [see comments]. Nature 347, 358-63.


This article should be referenced as such: