Mutation analysis in 600 French cystic fibrosis patients

F Chevalier-Porst, A M Bonardot, R Gilly, J P Chazalette, M Mathieu, D Bozon

Abstract

The cystic fibrosis transmembrane conductance regulator (CFTR) gene of 600 unrelated cystic fibrosis (CF) patients living in France (excluding Brittany) was screened for 105 different mutations. This analysis resulted in the identification of 86% of the CF alleles and complete genotyping of 76% of the patients. The most frequent mutations in this population after ΔF508 (69% of the CF chromosomes) are G542X (3.3%), N1303K (1.8%), W1282X (1.5%), 1717→G→A (1.5%), 2184delA → 2183 A→G (0.9%), and R553X (0.8%).

METHODS

Fourteen exons of the CFTR gene and their flanking intron-exon junctions were amplified according to Zielenksi et al.10 from genomic DNA. Methods used to detect mutations are: (1) heteroduplex formation followed by polyacrylamide gel electrophoresis (PAGE),17 (2) allele specific oligonucleotide hybridisation (ASO), (3) restriction enzyme digestion showing created or abolished sites. Mutations detected by the abolition of a restriction enzyme site were confirmed by ASO. (4) PCR with modified primers followed by restriction enzyme digestion.18

Haplotype analysis with marker/enzyme pairs XV2C/TaqI and KM19/PulI were performed and are named A, B, C, and D according to Estivill et al19 when the phase in the family is known.

The gene involved in cystic fibrosis (CF), the most common genetic disease among white populations, has been isolated and sequenced.20,21 The protein encoded by this gene, named cystic fibrosis transmembrane conductance regulator (CFTR), contains 1480 amino acids, has two membrane associated domains, two ATP binding folds (NBF), and a large highly charged domain (R) containing phosphorylation sites for protein kinases A and C. The most common mutation responsible for CF is a deletion, in exon 10 (first NBF domain), of a triplet coding for a phenylalanine at position 508 (ΔF508) of the CFTR protein. This mutation has been found with an overall frequency of 68% but with marked variation in populations of different geographical origin.6 Further genetic studies have identified more than 300 different mutations in this gene (Cystic Fibrosis Genetic Analysis Consortium, CFFGC).5 In this study a panel of 105 mutations5-15 has been tested on a sample of 600 unrelated CF patients living mostly in the eastern part of France, where ethnic heterogeneity is an important factor. Identification of the most frequent mutations in a given population is essential for genetic diagnosis and carrier risk assessment.

Material and methods

SAMPLE COMPOSITION

Mutation analysis was performed on 600 unrelated CF patients and their families. The diagnosis of CF was made on the basis of at least two positive sweat tests and on clinical findings. This CF population is composed mostly of patients of French origin (about 87%) but also of mixed European origin and North African origin (Algeria, Tunisia, Morocco).
least one CF chromosome (table): 21 of them are very rare as they were found on only one CF chromosome in our population. Screening of the 47 positive mutations allowed the characterisation of 1031 (85.9%) CF alleles and complete genotyping of 456 (76%) CF patients.

The most frequent mutations after ΔF508 are G542X (3.3%), N1303K (1.8%), W1282X (1.5%), 17171G→A (1.3%), 2184delA→2183A→G (0.9%), and R553X (0.8%). The search for ΔF508 and these six mutations allowed the identification of 944 (79%) CF alleles and the determination of the complete genotype of 386 (64.5%) CF patients in our study.

**Identification of New Sequence Modifications**

**I506M**

In exon 10, a C to G substitution at nucleotide position 1650 changes an isoleucine at position 506 into a methionine (I506M). This substitution gives, in association with a ΔF508 CF chromosome, a particular pattern when exon 10 PCR products (primed with C16B and C16D) are analysed by heteroduplex formation17 (fig 1). I506M is probably a polymorphism as there is a methionine at this position in two other related proteins: the maltose and the ribose transporters of E coli.

**2184delA**

When screening for the frameshift mutation 2184delA→2183A→G, one CF chromosome was found to have only the deletion of an A (2184delA) without the A→G substitution at position 2183 of the deleted sequence. The same mutation has also been identified in a German CF patient (Dörk et al, personal communication). The 11 other positive chromosomes had both changes.

**R1283K**

When screening for W1282X by MseI digestion, a patient showed a pattern similar to this mutation, except that it was on a CF chromosome bearing haplotype C (all the CF chromo-
somes with W1282X, apart from one on haplotype D, are from haplotype B. Direct sequencing of the PCR product of exon 20 from this patient showed a G to A substitution at nucleotide 3980 changing an arginine at position 1283 into a lysine (R1283K) (fig. 2). It has not been established if R1283K is a disease causing mutation as the arginine at this position is not conserved in the related proteins.\textsuperscript{3} R1283K can be detected by two different restriction enzyme digestions: abolition of the Mbol site or creation of an Mbol site giving a pattern similar to G1244E on agarose gel electrophoresis.\textsuperscript{21} The presence of this new mutation highlights the necessity of verification by ASO hybridisation for mutations detected by abolition of a restriction enzyme site.

We also identified four other mutations which have already been reported, a nonsense mutation in exon 4 (Y122X)\textsuperscript{22} and three frameshift mutations in exon 13, 1918delGC, 2118del4, and 2372del8.\textsuperscript{23} This study underlines some other factors. The identification of R334W in two affected sibs and the study of the segregation of this mutation through the family showed that the mother, without a history of CF, was homozygous for R334W. Further clinical investigations showed that she had a positive sweat test (80 mmol/l) and was pancreatic sufficient. This mutation has previously been described as a "mild" mutation.\textsuperscript{24}

As reported by Kälin et al,\textsuperscript{25} the two CF chromosomes (from unrelated patients) positive for S1251N also carry the polymorphism F508C,\textsuperscript{26} but two other CF chromosomes bearing F508C are negative for S1251N and have unknown mutations.

The splice mutation 711 + 1G→T, first described with a high frequency in CF families living in Quebec,\textsuperscript{27} was found in three homozygous and two compound heterozygous patients originating from North Africa. This mutation, also associated with haplotype A, accounts for 5% of CF chromosomes in this population and most of the other North African CF chromosomes carry unidentified mutations.

In our study, W1282X has a higher frequency than in the rest of France (CFGAC). This mutation has been reported with a high frequency (60% of CF chromosomes) in the Ashkenazi Jewish population.\textsuperscript{28} The cause of this frequency in our sample is unknown as the origin of these 18 chromosomes is variable (from the north east to south east of France).

The mutation 3905 insT, described with a frequency of about 9% in Swiss CF chromosomes\textsuperscript{29} has been found in our study in a large gypsy population where it is associated with ΔF508.

Discussion

The present study gives a representative view of the frequency of 105 CF mutations in an important sample of the CF population (600 patients).

From our data, the molecular pathology of the CFTR gene is exceedingly heterogeneous in our population: only 86% of the CF alleles identified with 47 different mutations (out of 105 tested). Most of the unidentified mutations are probably very rare and almost "private", so that systematic sequencing of these chromosomes would be necessary, but this procedure represents an unrewarding amount of work.

Most of the identified CF alleles belong to haplotype B (96%), while only 62% of groups A and D and 46% of groups C were identified. Consequently, for the patients with non-B chromosomes, polymorphism study (RFLP and microsatellites) will be necessary to identify the heterozygotes and allow prenatal diagnosis for the family. At the end of this study, mutation and haplotype analysis made 98.5% of the families informative.

Besides informativeness, the identification of the most frequent mutations in a given population is necessary for genetic counselling for at risk couples when one of the partners is a proven heterozygote. For the other partner, who has an initial risk of being a carrier of 1 in 25, the screening of the seven mutations ΔF508, G542X, N1303K, W1282X, 1717→G→A, 2184delA + 2183A→G, and R553X allows a better estimation of this risk; it drops to 1 in 120 if this screening is negative.

Our data also indicate that the goal of population screening for CF mutations (identification of 90 to 95% of the CF alleles) will be very difficult in this population of mixed ancestry because many different mutations will have to be tested. It seems very unlikely that a frequent but still unknown mutation will be identified, as many European CF chromosomes have already been sequenced.

We are grateful to all the patients and their families. We thank all the clinicians and geneticists for their cooperation and Lap-Chee Tsui for information. This manuscript was typed by Brigitte Biron. This work was supported by APLM (Association Française de Lutte contre la Mucoviscidose et HCL (Hospices Civils de Lyon).


