Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 8447–8451, November 1990
Medical Sciences

Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene

(genetic disease/missense mutation/nonsense mutation/mutational hot spot/pancreatic function)

BAT-SHEVA KEREM*†, JULIAN ZIELENSKI*, DANUTA MARKIEWICZ*, DOMINIQUE BOZON*, EPHRAIM GAZIT‡, JACOB YAHAV§, DARA KENNEDY*, JOHN R. RIORDAN¶, FRANCIS S. COLLINS‖, JOHANNA M. ROMMENS*, AND LAP-CHEE TSUI*,**

*Department of Genetics, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada; ‡Department of Human Genetics, Tel Aviv University, Tel Aviv 69978, Israel; §Chaim Sheba Medical Center, Tel Hashomer 52621, Israel; ¶Department of Biochemistry, The Hospital for Sick Children, and Departments of Biochemistry and Clinical Biochemistry, University of Toronto, Toronto, ON M5S 1A8, Canada; ‖Howard Hughes Medical Institute and Departments of Internal Medicine and Human Genetics, University of Michigan, Ann Arbor, MI 48109; and **Departments of Molecular and Medical Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada

Communicated by Y. W. Kan, August 3, 1990

ABSTRACT **Additional mutations in the cystic fibrosis (CF) gene were identified in the regions corresponding to the two putative nucleotide (ATP)-binding folds (NBFs) of the predicted polypeptide. The patient cohort included 46 Canadian CF families with well-characterized DNA marker haplotypes spanning the disease locus and several other families from Israel. Eleven mutations were found in the first NBF, 2 were found in the second NBF, but none was found in the R-domain. Seven of the mutations were of the missense type affecting some of the highly conserved amino acid residues in the first NBF; 3 were nonsense mutations; 2 would probably affect mRNA splicing; 2 corresponded to small deletions, including another 3-base-pair deletion different from the major mutation (ΔF508), which could account for 70% of the CF chromosomes in the population. Nine of these mutations accounted for 12 of the 31 non-ΔF508 CF chromosomes in the Canadian families. The highly heterogeneous nature of the remaining CF mutations provides important insights into the structure and function of the protein, but it also suggests that DNA-based genetic screening for CF carrier status will not be straightforward.**

The gene responsible for cystic fibrosis (CF) has recently been identified and the major mutation has been defined at the DNA level (1–3), providing a direct means for elucidating the basic defect in this disease and for giving an accurate genetic diagnosis. Sequence analysis of overlapping cDNA clones predicts a protein consisting of 1480 amino acids and containing membrane-associated regions and nucleotide (ATP)-binding folds (NBFs); this putative protein product has been named the cystic fibrosis transmembrane conductance regulator (CFTR) (2). The major mutation corresponds to a 3-base-pair (bp) deletion, which results in the loss of a phenylalanine residue at amino acid position 508 (ΔF508), within the first NBF (NBF1) of CFTR (2, 3). This mutation accounts for ≈70% of all CF chromosomes, and haplotype analysis suggests that all ΔF508 chromosomes are derived from the same origin (3–5).

On the basis of DNA marker haplotype analysis, the remaining 30% of CF chromosomes are expected to be heterogeneous with respect to the nature of the mutations (3). Since each mutation marks a functionally important region of CFTR, it is expected that identification of additional mutations in the CF gene will provide useful information about the structure and function of the protein, in understanding the pathophysiology of the disease, and in development of pharmaceutical agents for better treatment of individuals with CF. Furthermore, identification of CF mutations at the DNA level will provide the required information basis for genetic testing, especially when population screening for CF carriers is considered.

Since the major CF mutation occurs within the first NBF of CFTR, it is reasonable to assume that additional mutations might also be found within the two NBFs for the remaining 30% of CF chromosomes. In this communication, we report the result of a survey of the CF gene, in regions corresponding to the two NBFs and the R-domain of the protein, for the presence of additional mutations.

MATERIALS AND METHODS

DNA Samples. The majority of CF families used in this study have been described (6). Extensive DNA marker haplotype information was available for the CF chromosomes in 46 Canadian families (3). A small number of additional CF families with unique geographic or ethnic backgrounds were included in the mutational screening; the ones reported here were recruited in the Chaim Sheba Medical Center in Israel.

Polymerase Chain Reactions (PCRs). DNA sequences spanning individual exons were amplified by PCR with oligonucleotide primers located in the respective flanking introns. The general procedure for PCR has been described (7, 8). The specific oligonucleotide primers used for the amplification of individual exon sequences are listed in Table 1 (details of the intron/exon boundary sequences will be published elsewhere). All oligonucleotides were purchased from the Hospital for Sick Children DNA Biotechnology Service Center. Approximately 500 ng of genomic DNA from cultured lymphoblastoid cell lines or peripheral blood of each individual was used in each reaction. Typically, DNA samples were heated at 94°C for 6 min and then subjected to 30 cycles of denaturation (94°C for 30 sec), primer annealing (55°C for 30 sec), and extension (72°C for 60 sec) with 2.5 units of *Thermus aquaticus* DNA polymerase (*Taq* polymerase) (Perkin–Elmer/Cetus), followed by a final extension period of 7 min in a Perkin–Elmer/Cetus DNA thermal cycler.

DNA Sequence Determination. Sequencing of cloned DNA was performed according to the dideoxynucleotide chain-termination method essentially as described (2, 9) with the Sequenase kit (United States Biochemical). The method for direct sequencing of PCR products was modified from published procedures (10) with either one of the PCR primers or

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: CF, cystic fibrosis; NBF, nucleotide (ATP)-binding fold; CFTR, CF transmembrane conductance regulator; PCR, polymerase chain reaction; ASO, allele-specific oligonucleotide; PI, pancreatic insufficient; PS, pancreatic sufficient.

†Present address: Department of Genetics, Hebrew University, Jerusalem 91904, Israel.

Table 1. Oligonucleotides for PCR-directed genomic DNA amplification and DNA sequencing

Exon	Primer (5′ → 3′)	Amplified product, bp
	PCR primer	
9	TAATGGATCATGGGCCATGT (9i-5)	560
	ACAGTGTTGAATGTGGTGCA (9i-3)	
10	GCAGAGTACCTGAAACAGGA (10i-5)	491
	CATTCACAGTAGCTTACCCA (10i-3)*	
11	CAACTGTGGTTAAAGCAATAGTGT (11i-5)*	425
	GCACAGATTCTGAGTAACCATAAT (11i-3)	
12	GTGAATCGATGTGGTGACCA (12i-5)*	426
	CTGGTTTAGCATGAGGCGGT (12i-3)	
13	(a) TGCTAAAATACGAGACATATTGCA (13i-5)*	528
	ATCTGGTACTAAGGACAG (C1-1M)*	
	(b) TCAATCCAATCAACTCTATACGAA (X13B-5)*	497
	TACACCTTATCCTAATCCTATGAT (13i-3)*	
19	GCCCGACAAATAACCAAGTGA (19i-5)	454
	GCTAACACATTGCTTCAGGCT (19i-3)*	
20	GGTCAGGATTGAAAGTGTGCA (20i-5)*	473
	CTATGAGAAAACTGCACTGGA (20i-3)	
21	AATGTTCACAAGGGACTCCA (21i-5)	477
	CAAAAGTACCTGTTGCTCCA (21i-3)*	
22	AAACGCTGAGCCTCACAAGA (22i-5)	562
	TGTCACCATGAAGCAGGCAT (22i-3)*	
23	AGCTGATTGTGCGTAACGCT (23i-5)	400
	TAAAGCTGGATGGCTGTATG (23i-3)*	
	Internal sequencing primer	
9	GTACTCCTGTCCTGAAAGA (X9B)	
	CCAGCAACCGCCAACAAC (T16L)	

PCR amplification of each exon was accomplished by one pair of oppositely oriented PCR primers (the primers within the 5′ intron are indicated by i-5 and those within the 3′ intron are indicated by i-3) except exon 13, for which two sets of PCR primers were required. Direct DNA sequencing of the amplified products was performed with the oligonucleotides marked by asterisks; two internal primers were necessary for exon 9.

internal oligonucleotides used as sequencing primer (see Table 1).

Allele-Specific Oligonucleotide (ASO) Hybridization. The oligonucleotide hybridization condition has been described (3). The normal and mutant specific oligonucleotides are listed in Table 2. Hybridization was performed at 37°C and the washings were done twice with 5× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate) for 10 min each at room temperature followed by two washings with 2× SSC for 30 min each at 45°C–57°C; the exact temperature was determined empirically for each ASO as indicated.

Cloning of PCR Products. In one case (ΔI507), in which direct sequencing of PCR-amplified genomic DNA did not allow unequivocal determination of nucleotide sequence alteration in the patient, cloning of the mutant allele was necessary. The two homoduplexes (94 and 97 bp) generated by PCR amplification of the genomic DNA from the father, who was found to contribute the mutant allele in question on the basis of hybridization analysis (data not shown), were purified from a 5% polyacrylamide gel and cloned into Bluescript KS vector (Stratagene). One clone, 5-3-15, was isolated by its lack of hybridization with the oligonucleotide N (3). The sequence obtained from this clone was in perfect agreement with that derived from direct DNA sequencing of the PCR products.

RESULTS AND DISCUSSION

Initially, we had selected a small number of individuals representative of different haplotype groups (3) for detection of unusual CF mutations, as we assumed that there were only a limited number of mutations and that CF chromosomes in the same DNA marker haplotype group might carry identical mutations. Since, as the study continued, many more CF mutations had been discovered than previously anticipated, we included all the CF families (patients) with non-ΔF508 mutations in our study population. We also included several CF individuals whose DNA marker haplotypes were not determined.

Table 2. Methods for detecting mutations

Mutation	Method
A455E	ASO hybridization; washing at 52°C
	N: GTTGTTGG**C**GGTTGCT;
	A455E: GTTGTTGG**A**GGTTGCT
Q493X	Allele-specific PCR; annealing at 57°C
	Common primer: 10i-5
	N: GGCATAATCCAGGAAAACT**G**
	Q493X: GGCATAATCCAGGAAAACT**A**
ΔI507	Heteroduplex DNA formation (see Fig. 2)
ΔF508	ASO hybridization (3)
	Heteroduplex DNA formation (11)
1717-1G → A	ASO hybridization; washing at 47°C
	N: TTTGGTAATA**G**GACATCTCC
	1717-1G → A: TTTGGTAATA**A**GACATCTCC
G542X	ASO hybridization; washing at 45°C
	N: ACCTTCTC**C**AAGAACT
	Q542X: ACCTTCTC**A**AAGAACT
S549I	*Dde* I digestion
	N: 238 + 179 + 13 bp;
	S549I: 417 + 13 bp (elimination of site)
S549R	ASO hybridization; washing at 56°C
	N: ACACTGAG**T**GGAGGTC
	S549R: ACACTGAG**G**GGAGGTC
	Allele-specific PCR; annealing at 55°C
	Common primer: 11i-5
	N: TTCTTGCTCGTTGACCTCC**A**
	S549R: TTCTTGCTCGTTGACCTCC**C**
G551D	ASO hybridization; washing at 45°C
	N: GAGTGGAG**G**TCAACGAG
	G551D: GAGTGGAG**A**TCAACGAG
	*Sau*3A digestion (12)
R560T	*Mae* II digestion
	N: 425 bp;
	R560T: 215 + 210 bp (creation of site)
Y563N	ASO hybridization; washing at 54°C
	N: AGCAGTA**T**ACAAAGATGC
	Y563N: AGCAGTA**A**ACAAAGATGC
P574H	ASO hybridization; washing at 48°C
	N: GACTCTC**C**TTTTGGA
	P574H: GACTCTC**A**TTTTGGA
3659delC	Allele-specific PCR; annealing 60°C
	N: GTATGGTTTGGTTGACTTG**G**
	3659delC: GTATGGTTTGGTTGACTTG**T**
W1282X	Direct DNA sequencing of PCR product

Detection of Mutations. Direct DNA sequencing for the genomic DNA regions spanning the two NBFs (exons 9–12 and 19–23) and the R-domain (exon 13) was performed for 29 individuals carrying 40 CF chromosomes with mutations other than ΔF508. The results are shown in Table 3. A total of 13 mutations were identified: 11 of them were located in the first NBF and 2 were in the second NBF, but none were in the R-domain. Two of these mutations, G551D (12) and W1282X (14), have been reported previously; 2 others, ΔI507 (M. Schwarz, C. Summers, L. Heptinstall, C. Newton, A. Markham, R. Cain, and M. Super, unpublished data) and 1717-1G → A (16), have also been described by other investigators independently.

Different lines of evidence indicate that the sequence alterations detected in this study were disease causing: (*i*)

three of the changes (Q493X, G542X, W1282X) were nonsense mutations that predict production of truncated polypeptides; (*ii*) five of the missense mutations (A455E, S549I, S549R, G551D, R560T) occurred at highly conserved amino acid residues in the first NBF; (*iii*) two mutations altered the sequence at splice junctions—one (1717-1G → A) switched the highly conserved AG (17) to AA, and the other (R560T), which was also a missense mutation, substituted AC for AG immediately upstream of the splice acceptor site, creating an alteration known to reduce efficiency of splicing (18); (*iv*) one mutation (3659delC) caused a single base-pair deletion that would result in a frameshift and premature termination; (*v*) one mutation (ΔI507) corresponded to a 3-bp deletion immediately upstream of the position for the phenylalanine codon deleted in ΔF508; (*vi*) two missense mutations (Y563N and P574H) would result in a drastic change of the encoded amino acids. Although the chromosomes carrying the various mutations described above were not fully sequenced to exclude the possibility of other coexisting alterations, as has been done for ΔF508, none of the changes were found among the normal chromosomes, some of which shared extensive DNA marker haplotype as the mutant chromosomes (data not shown). The absence of the same alterations in haplotype-matched chromosomes for most mutations suggested that the sequence alterations were not polymorphisms.

While the majority of the mutations described in this report (Table 3) appeared to be rare in the population, as only single cases were detected, two of the mutations, A455E and G551D, occurred more than once among the present Canadian chromosome population—two were detected for A455E and three were detected for G551D. In an earlier study, Cutting *et al.* (12) reported that 6 of 35 (17%) non-ΔF508 CF chromosomes screened carried the G551D mutation. Preliminary data from studies of other populations indicated that the G542X and S549I mutations were also relatively abundant; for example, 5 of 59 (8.4%) non-ΔF508 CF chromosomes were found to carry G542X in the Jewish population in Israel and 1 of the CF Arabic patients was homozygous S549I. Together, these four mutations may account for 10–20% of the non-ΔF508 CF chromosomes.

In addition to the mutations described above, we have detected three DNA sequence variations that occur on normal chromosomes: two of them, at nucleotide positions 1716 (G or A) and 2694 (T or G), do not appear to affect the encoded amino acids, but the third one, at position 1540 (A or G), results in either methionine or valine.

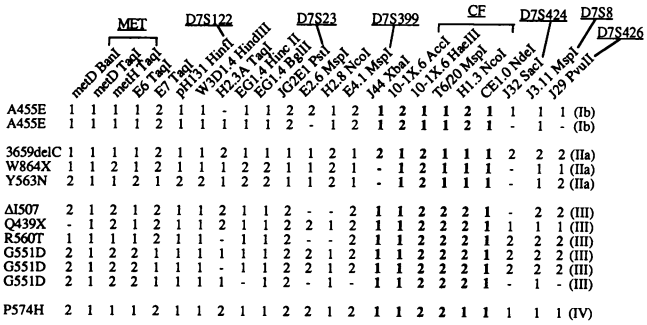

FIG. 1. DNA marker haplotypes associated with the described CF mutations. The DNA markers (with the locus names underlined), and the enzyme revealing the polymorphism, haplotypes, and group names (in parentheses) are defined as described (3).

DNA Marker Haplotype. Determination of the DNA marker haplotype associated with each of the mutations is important, not only in providing a reference for mutant classification but also in tracing the origins of mutations, thereby gaining insights into the possible cause for the high frequency of CF mutations in the Caucasian population.

Our previous study showed that chromosomes carrying the ΔF508 mutation were derived from the same origin (3). To determine whether ancestral relationships existed among CF chromosomes carrying the other mutations, DNA marker haplotypes were analyzed. As shown in Fig. 1, the chromosomes containing A455E and G551D share the same haplotypes for DNA markers spanning the CF gene, groups Ib and III, respectively. The G551D chromosomes detected by Cutting *et al.* (12) also appear to have haplotype III. The ΔI507 mutation described here carries a group III haplotype (Fig. 1) and it is in good agreement with the partial haplotype data for another example of this mutation independently identified (M. Schwarz, C. Summers, L. Heptinstall, C. Newton, A. Markham, R. Cain, and M. Super, unpublished data). Therefore, these haplotype data suggest that CF chromosomes harboring the same mutation were probably derived from a common origin.

Chromosomes within the same haplotype group, however, may harbor different mutations (e.g., those in groups IIa and

Table 3. Mutations in the two NBF regions of the CF gene

Mutation	Amino acid change	Nucleotide change	Exon	No. of CF chromosomes screened: With mutation	No. of CF chromosomes screened: Without mutation	N
A455E	Ala-455 → Glu	C-1496 → A	9	2	84(62)	40(11)
Q493X	Gln-493 → stop	C-1609 → T	10	1	33(5)	28(15)
ΔI507	Deletion of Ile-506 or -507	3-bp deletion	10	1	465	198
ΔF508	Deletion of Phe-508	3-bp deletion	10	330*	136*	198
1717-1G→A	Splice mutation	G-1717-1 → A	Intron 10	1	67	60
G542X	Gly-542 → stop	G-1756 → T	11	1	63	60
S549I	Ser-549 → Ile	G-1778 → T	11	1	66	58
S549R	Ser-549 → Arg	T-1779 → G	11	1	63	60
G551D	Gly-551 → Asp	G-1784 → A	11	3	35(5)	19(5)
R560T	Arg-560 → Thr	G-1811 → C	11	1	36(7)	44(16)
Y563N	Tyr-563 → Asn	T-1819 → A	12	1	59(8)	54(39)
P574H	Pro-574 → His	C-1853 → A	12	1	111	78(4)
3659delC	Frameshift	1-bp deletion	19	1	57(7)	50(43)
W1282X	Trp-1282 → stop	G-3978 → A	20	1	32	8

Numbers in parentheses indicate the number of chromosomes within the same haplotype group (see Table 2 and ref. 3). Except for ΔF508 and ΔI507, most of the CF chromosomes screened were non-ΔF508 chromosomes. N represents the number of parental normal (non-CF) chromosomes examined.

*These numbers are included for comparison (see refs. 3 and 13). G551D and W1282X have been reported previously (12, 14).

III; Fig. 1). It is of interest to note that four different mutations have already been identified on CF chromosomes with haplotype III, which is infrequent among the N chromosomes (3). The number of mutations in the CF gene would therefore greatly exceed that predicted on the basis of haplotype analysis (3), as has been observed thus far.

The extended DNA marker haplotypes were not determined for five of the CF mutations—namely, G542X, W1282X, S549R, S549I, and 1717-1G → A—because the patients in whom these mutations were found were all recently recruited from Israel. The G542X mutation was first found in an Ashkenazic Jewish patient whose other CF chromosome was W1282X; S549R was discovered in a Moroccan–Jewish individual who also carried ΔF508; S549I and 1717-1G → A were both first detected in Arabic patients.

On the basis of ASO-hybridization and haplotype data, we previously suspected that one of the CF chromosomes with haplotype group III might carry a recurring ΔF508 mutation (3). Direct DNA sequence analysis, however, revealed that the mutation carried by this chromosome was ΔI507 (Table 3 and Fig. 1), differing from ΔF508 by a single nucleotide, A instead of T. The marked difference between haplotype III and haplotype Ia, which is associated with ΔF508, suggests that the two mutations are derived from different origins.

Hot Spots for Mutations. There appeared to be a difference in distribution of mutations between the two NBFs—the number of mutations detected in the first NBF (exons 9–12) and 6 times that in the second NBF (exons 19–23). In addition, although the sample size is small, exon 11 seems to contain the most mutations (this study; ref. 12). It is possible that the first NBF, particularly the exon 11 region, is more critical for CFTR function. Alternatively, the DNA sequences encoding the first NBF may be relatively more unstable than other regions. The latter explanation is supported by two lines of observations. First, only two 3-bp deletions have so far been detected in the CF gene; both occur in the same region within exon 10. Second, several sequence alterations have been detected for the AGT codon for serine at position 549 in exon 11; while Cutting *et al.* (12) reported a change of this codon to AAT for asparagine (S549N), this study shows two additional mutations at this position—T to G transversion for S549R and G to T transversion for S549I.

Implications in Protein Structure. Since no major deletion or rearrangement has been detected in CF, the discovery of nonsense, frameshift, and splicing mutations in the gene is surprising. It is also somewhat surprising that no mutation was found in the R-domain, a region that was thought to have a regulatory function (2). The only mutation reported for the region so far was a 2-bp insertion toward the end of this domain (19). The paucity of mutations in this region suggests that either the R-domain is not critical for function or mutations in this part of the protein cannot be tolerated. On the other hand, patients homozygous for nonsense (ref. 20; H. Cuppens, P. Marynen, C. De Boeck, F. Baets, E. Eggermont, H. Van den Berghe, and J. J. Cassiman, unpublished data) or splicing (data not shown) mutations have been detected, suggesting that total loss of CFTR function is not incompatible with life.

The occurrence of ΔI507 and ΔF508 in the same region of the presumptive ATP-binding domain of CFTR is noteworthy. First, it is unclear which 3 bp were deleted in both cases (Fig. 2). In ΔI507, the deleted amino acid may be the isoleucine residue at position 506. Second, the amino acid residues involved, Phe-508 and Ile-507 (or Ile-506), are located in a polypeptide segment linking the two highly conserved regions believed to be the nucleotide contacting points (2, 22). Deletion of either one of the two amino acids is sufficient to disrupt the gene function. Furthermore, amino acid substitutions in this region have been found in two rare variants of normal chromosomes (with Phe-508 replaced by a cysteine and Ile-506 replaced by a valine) without apparent loss of function of the protein (23). Therefore, the length of the polypeptide may be more important than the actual amino acid residues in this region for CFTR. Since the length of the linking segment varies among different ATP-binding proteins and the amino acid sequences are less conserved, the functional importance of Ile-507 and Phe-508 is probably unique to CFTR.

Five of the missense mutations (A455E, S549I, S549R, G551D, R560T) occurred at amino acid residues highly conserved among the various NBFs from different proteins, suggesting that these residues are critical in ATP binding (2). The role of the two other missense mutations (Y563N, P574H) within the NBFs is not immediately apparent, although both mutations would drastically change the amino acid types. It is anticipated, however, that knowledge about these and other (12, 14, 19, 24) naturally occurring mutations will be an essential part in understanding the structure and function of the protein. In conjunction with development of three-dimensional structural analysis and tests for the biological activities of CFTR, it should be possible to define the role of each of these residues at the molecular level and, eventually, the function of the protein.

Clinical Phenotype. Our previous studies (3, 25) showed that the pancreatic function of CF patients was at least in part predisposed by their genotypes at the CF locus. We hypothesized that CF mutations could be divided into two classes—namely, severe and mild (with respect to pancreatic involvement). Patients with two severe alleles are expected to be pancreatic insufficient (PI) and patients with one or two copies of mild alleles are pancreatic sufficient (PS). The ΔF508 mutation has been classified as a severe allele and, as expected, patients homozygous for ΔF508 are almost certainly PI (3, 13).

To examine the clinical consequence of the mutations described in this report, the clinical status of each individual carrying these mutations was reviewed. Ten of the mutations

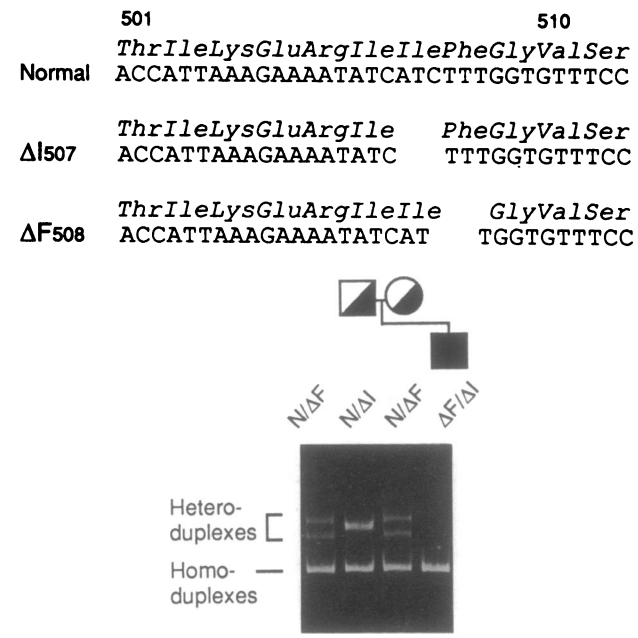

FIG. 2. Identification of the ΔI507 and ΔF508 alleles. (*Upper*) The nucleotide sequence around the mutation sites. (*Lower*) Detecting the mutations by PAGE. The PCR products were prepared from the three family members and separated on a 5% polyacrylamide gel. A DNA sample from a known heterozygous ΔF508 carrier is included for comparison. The normal and mutant homoduplexes are poorly separated in this assay system because the PCR products are 491 and 488 bp long, respectively; however, the two sequences may be easily distinguished with shorter PCR products (11).

detected here (Q493X, ΔI507, 1717-1G → A, G542X, S549I, S549R, G551D, R560T, 3659delC, W1282X) were clearly associated with the PI phenotype and should, therefore, be classified as severe (as for ΔF508) according to our previous classification (3). It is of interest to note that several patients with both CF chromosomes carrying nonsense mutations (e.g., G542X) were reported to have relatively mild lung disease (although all PI as predicted) and that it was argued that truncation or total absence of CFTR might be associated with near normal lung function (ref. 20; H. Cuppens, P. Marynen, C. De Boeck, F. Baets, E. Eggermont, H. Van den Berghe, and J. J. Cassiman, unpublished data). Further studies are required to strengthen the latter argument, however, because good pulmonary condition has also been detected in a small number of patients homozygous for ΔF508 (13).

Two of the mutations, A455E and P574H, were detected in PS patients. Since two PS patients with these mutations in our study population are heterozygous for ΔF508, it is reasonable to assume that A455E and P574H are both mild alleles. A455E was also detected in a PS patient with the other chromosome carrying an unknown mutation. The physiological consequence of Y563N could not be determined because it was found in a PS patient with a yet unclassified mutation in exon 14a (W864X; ref. 14). Obviously, additional case studies will be necessary to clarify the clinical consequences of carrying these potentially mild mutations.

Moreover, we have found a substitution of cysteine for Phe-508 (F508C) in one patient who has typical symptoms of CF with PI, whereas the amino acid substitution has been found in at least two phenotypically normal individuals carrying the ΔF508 on their other chromosome 7 (23). It is possible that the F508C chromosome detected here carries an additional mutation causing the disease. Alternatively, there may be other genetic or environmental factors that contribute to disease severity.

DNA Diagnosis. An immediate application for the information of the defined mutation sites is in CF diagnosis. For each identified mutation, however, it will be essential to devise an easy yet reliable screening procedure to confirm the observation and to estimate its frequency in various populations. We previously suggested that the unique mobility of the heteroduplexes formed between DNA fragments containing the normal and ΔF508 alleles is a rapid, nonradioactive method to detect carrier individuals (11). This method may also be applied to ΔI507; as shown in Fig. 2, the banding pattern of the PCR-amplified genomic DNA from the father, who is the carrier with ΔI507, is clearly distinguishable from that of the mother, who is typical of carriers with ΔF508.

It has become clear that, although ΔF508 accounts for ≈70% of all CF chromosomes, the total number of disease-causing mutations in the CF gene is likely to be large (refs. 12, 14, 16, 19, 24; this study and M. Schwarz, C. Summers, L. Heptinstall, C. Newton, A. Markham, R. Cain, and M. Super, unpublished data). In this study, while we have fully analyzed >46% (2055/4440) of the coding region of the 31 non-ΔF508-carrying chromosomes in the Canadian family collection for which the DNA marker haplotypes have been determined (3), the mutations identified only account for 12 of these chromosomes (38%). Therefore, DNA-based population screening for CF carriers will be complicated by the existence of a large number of different mutant alleles, especially since some of them are expected to have markedly different geographical (ethnic) distributions (4, 5, 12).

It is perhaps also important to point out the pitfalls of some of the common means of DNA diagnosis. For example, both S549N (12) and S549I (this study) destroy the same *Dde* I site; although both of them are disease-causing mutations in this instance, the destruction of a restriction enzyme recognition site is obviously not a reliable method to define a particular mutation. ASO hybridization may also give false results; ΔI507 could be mistaken as ΔF508 under low hybridization stringency (ref. 3; this study). Allele-specific PCR has been developed for ΔF508 (15) and several mutations described here (Q493X, S549R, and 3659delC); this method usually works well but it requires very careful temperature control to distinguish different sequences. These examples illustrate that extreme caution should be used with any DNA testing procedure; testing of a large number of normal chromosomes, particularly those of the same haplotypes, and quality control should be mandatory before any broad-scale screening is implemented. Furthermore, it has been suggested that broad-scale population screening for CF carriers should not begin until at least 95% of the mutations have been identified (21).

The authors wish to thank Richard Rozmahel and Natasa Plavsic for expert technical assistance and Garry Cutting, Arthur Beaudet, Michel Goossens, Maurice Super, and J. J. Cassiman for communicating data prior to publication. The availability of the CF patient data base developed by Mary Corey, Peter Durie, and others at The Hospital for Sick Children in Toronto is also gratefully acknowledged. This research was supported by grants from the National Institutes of Health (DK-34944-5), the Cystic Fibrosis Foundation (USA), and the Canadian Cystic Fibrosis Foundation. J.M.R. and L.-C.T. are recipients of a Postdoctoral Fellowship and a Scientist Award from the Medical Research Council of Canada, respectively.

1. Rommens, J. M., Iannuzzi, M. C., Kerem, B., Drumm, M. L., Melmer, G., Dean, M., Rozmahel, R., Cole, J. L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J. R., Tsui, L.-C. & Collins, F. S. (1989) *Science* **245**, 1059–1065.
2. Riordan, J. R., Rommens, J. M., Kerem, B., Alon, N., Rozmahel, R., Grzelchak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.-L., Drumm, M. L., Iannuzzi, M. C., Collins, F. S. & Tsui, L.-C. (1989) *Science* **245**, 1066–1073.
3. Kerem, B., Rommens, J. M., Buchanan, J. A., Markiewicz, D., Cox, T. K., Chakravarti, A., Buchwald, M. & Tsui, L.-C. (1989) *Science* **245**, 1073–1080.
4. Lemna, W. K., Feldman, G. L., Kerem, B., Fernbach, S. D., Zevkovich, E. P., O'Brien, W. E., Collins, F. S., Tsui, L.-C. & Beaudet, A. L. (1990) *N. Engl. J. Med.* **322**, 291–296.
5. The Cystic Fibrosis Genetic Analysis Consortium (1990) *Am. J. Hum. Genet.* **47**, 354–359.
6. Tsui, L.-C., Zengerling, S., Willard, H. F. & Buchwald, M. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51**, 325–335.
7. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230**, 1350–1354.
8. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
9. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
10. Winship, P. R. (1989) *Nucleic Acids Res.* **17**, 1266.
11. Rommens, J., Kerem, B., Greer, W., Chang, P., Tsui, L.-C. & Ray, P. (1990) *Am. J. Hum. Genet.* **45**, 395–396.
12. Cutting, G. R., Kasch, L. M., Rosenstein, B. J., Zielensky, J., Tsui, L.-C., Antonarakis, S. E. & Kazazian, H. H., Jr. (1990) *Nature (London)* **346**, 366–368.
13. Kerem, E., Corey, M., Kerem, B., Rommens, J., Markiewicz, D., Levison, H., Tsui, L.-C. & Durie, P. (1990) *N. Engl. J. Med.*, in press.
14. Vidaud, M., Fanen, P., Martin, J., Ghanem, N., Nicolas, S. & Goossens, M. (1990) *Hum. Genet.* **85**, 446–459.
15. Balabio, A., Gibbs, R. A. & Caskey, C. T. (1990) *Nature (London)* **343**, 220.
16. Guillermit, H., Fanem, P. & Ferec, C. (1990) *Hum. Genet.* **85**, 450–453.
17. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4853–4857.
18. Vidaud, M., Gattoni, R., Stevenin, J., Vidaud, D., Amselem, S., Chibani, J., Rosa, J. & Goossens, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1041–1045.
19. White, M. B., Amos, J., Hsu, J. M. C., Gerrard., B., Finn, P. & Dean, M. (1990) *Nature (London)* **344**, 665–667.
20. Cutting, G. R., Kasch, L. M., Rosenstein, B. J., Tsui, L.-C., Kazazian, H. H., Jr., & Antonarakis, S. E. (1990) *N. Engl. J. Med.*, in press.
21. The NIH Workshop on Population Screening for the Cystic Fibrosis Gene (1990) *N. Engl. J. Med.* **323**, 70–71.
22. Hyde, S. C., Emsley, P., Hartshorn, M. J., Mimmack, M. M., Gileadi, U., Pearce, S. R., Gallagher, M. P., Gill, D. R., Hubbard, R. E. & Higgins, C. F. (1990) *Nature (London)* **346**, 362–365.
23. Kobayashi, K., Knowles, M., O'Brien, W. E. & Beaudet, A. L. (1990) *Am. J. Hum. Genet.* **47**, in press.
24. Dean, M., White, M., Amos, J., Gerrard, B., Stewart, C., Khaw, K.-T. & Leppert, M. (1990) *Cell* **61**, 863–870.
25. Kerem, B., Buchanan, J. A., Durie, P., Corey, M., Levison, H., Buchwald, M. & Tsui, L.-C. (1989) *Am. J. Hum. Genet.* **44**, 827–834.