Three Novel CFTR Polymorphic Repeats Improve Segregation Analysis for Cystic Fibrosis

Ausilia Elce,1,2 Angelo Boccia,1,2 Giuseppe Cardillo,1,2 Sonia Giordano,1,2 Rossella Tomaiuolo,1,2,3 Giovanni Paolella,1,2,3,4* and Giuseppe Castaldo1,2,3,4*

BACKGROUND: Molecular diagnosis for cystic fibrosis (CF) is based on the direct identification of mutations in the CFTR gene [cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7)] (detection rate about 90% with scanning procedures) and on segregation analysis of intragenic polymorphisms for carrier and prenatal diagnosis in about 20% of CF families in which 1 or both causal mutations are unknown.

METHODS: We identified 3 novel intragenic polymorphic repeats (IVS3polyA, IVS4polyA, and IVS10CA repeats) in the CFTR gene and developed and validated a procedure based on the PCR followed by capillary electrophoresis for large-scale analysis of these polymorphisms and the 4 previously identified microsatellites (IVS1CA, IVS8CA, IVS17bTA, and IVS17bCA repeats) in a single run. We validated the procedure for both single- and 2-cell samples (for a possible use in preimplantation diagnosis), and on a large number of CF patients bearing different genotypes and non-CF controls.

RESULTS: The allelic distribution and heterozygosity results suggest that the 3 novel polymorphisms strongly contribute to carrier and prenatal diagnosis of CF in families in which 1 or both causal mutations have not been identified. At least 1 of the 4 previously identified microsatellites was informative in 78 of 100 unrelated CF families; at least 1 of all 7 polymorphisms was informative in 98 of the families. Finally, the analysis of haplotypes for the 7 polymorphisms revealed that most CF mutations are associated with different haplotypes, suggesting multiple slippage events but a single origin for most CFTR mutations.

CONCLUSIONS: The analysis of the 7 polymorphisms is a rapid and efficient tool for routine carrier, prenatal, and preimplantation diagnosis of CF.

Cystic fibrosis (CF) is 1 of the most frequent (1 in 2500) autosomal recessive diseases with a severe outcome among Caucasians. Substantial genetic heterogeneity has been reported among CF patients, and >1600 mutations have been described in CFTR [cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7)], the gene responsible for the disease (1). The mutations show different distributions, depending on the ethnic/geographic origin of the CF patients. The most common mutation, F508del, shows a decreasing prevalence from northwest Europe (up to 80% of CF alleles) to southeast Europe (about 50% of CF alleles), where the disease has a greater molecular heterogeneity. A dozen other mutations are found globally, and although some mutations are peculiar to specific ethnic/geographic groups (2), the large majority of mutations are private (3).

Molecular analysis contributes to CF diagnosis and is the lone approach for identifying CF carriers through cascade screening and for performing prenatal diagnosis in high-risk couples (4). Molecular diagnosis is based on the use of commercial kits for analyzing panels of mutations. In southern Italy, for example, this first-level approach detects 70%–80% of CF alleles (5). Second-level approaches, which are based on the analysis of all the CFTR-coding regions, detect approximately 90% of mutant CF alleles (6). A few CF alleles bear large gene rearrangements (7). The combined use of all these techniques cannot guarantee detection of both mutations in all CF patients, because about 10% of CF alleles remain uncharacterized. Consequently, 1 or both causal mutations are unknown in about 20% of
Identification of 3 Novel Polymorphic Repeats in \textit{CFTR}

CF patients/families, and carrier and prenatal diagnosis is based on segregation analysis of \textit{CFTR} polymorphic markers, i.e., 4 microsatellites within the \textit{CFTR} gene (4, 8–10). Segregation analysis is also useful for preimplantation genetic diagnosis (11, 12).

During a project designed to analyze all \textit{CFTR} intronic conserved sequence tags (13) in CF patients bearing undetected mutations, we identified 3 novel intragenic polymorphisms—2 mononucleotide repeats and 1 dinucleotide repeat.

We developed and validated a novel procedure based on the PCR followed by capillary electrophoresis for large-scale analysis of both the 3 novel intragenic polymorphisms and the 4 previously identified microsatellites (IVS1CA, IVS8CA, IVS17bTA, and IVS17bCA) in a single run.

Materials and Methods

\textbf{STUDY PARTICIPANTS}

We studied 145 unrelated individuals from southern Italy: (a) 20 CF patients homozygous for the F508del mutation; (b) 60 CF patients bearing other known genotypes; (c) 16 CF patients bearing 1 (n = 15) or both (n = 1) undetected mutations; and (d) 49 non-CF individuals. To evaluate the informative value of this novel procedure, we also tested 100 unselected DNA samples from unrelated CF patients and from the 2 parents. The whole project of analyzing \textit{CFTR} conserved sequence tags in CF patients was approved by our institutional ethics committee. All individuals included in the study provided informed consent to use their DNA sample for research purposes. A sample of DNA extracted from a blood sample was obtained from all individuals for diagnostic purposes. In all cases, direct gene sequencing was used to scan the entire coding region of \textit{CFTR}, including exon–intron boundaries (protocols available on request). PCR analysis followed by agarose gel electrophoresis was used to test for the 7 most frequent \textit{CFTR} rearrangements (7).

\textbf{SINGLE- AND 2-CELL ANALYSIS}

Laser microdissection (14) was used to isolate aliquots of 1 and 2 cells from samples of the buccal mucosa collected from laboratory donors. DNA was extracted from these cells with the GenomiPhi V2 DNA Amplification Kit and then amplified according to the manufacturer’s protocols. Reamplification of the entire genome from cells obtained by laser microdissection was checked by electrophoresis in agarose gels (20 g/L) in 1× Tris-borate-EDTA buffer [concentration of 10× solution: 1.0 mol/L Tris, 0.9 mol/L boric acid, 0.01 mol/L EDTA, pH at 25 °C: 8.4 (0.10)].

\textbf{IDENTIFICATION AND LARGE-SCALE ANALYSIS OF THE IVS10CA, IVS3polyA, AND IVS4polyA REPEATS}

The identification of the 3 novel polymorphisms was performed by sequence analysis of intronic conserved sequence tags in the \textit{CFTR} gene (13). The first is an A repeat within intron 3 at nucleotide 3267 and includes 3 alleles; the second, another A repeat, is within \textit{CFTR} intron 4 at nucleotide 52115 and includes 3 alleles; the third is a CA repeat, identified in \textit{CFTR} intron 10 at nucleotide 86287, and includes 9 different alleles. For large-scale analysis of the 3 polymorphisms, we set up a procedure based on the PCR followed by capillary electrophoresis to analyze, under the same conditions, both the 3 novel polymorphisms (i.e., IVS3polyA, IVS4polyA, and IVS10CA repeats) and the 4 previously identified \textit{CFTR} microsatellites (IVS1CA, IVS8CA, IVS17bTA, and IVS17bCA repeats). All polymorphic regions were amplified with the same PCR protocol. DNA (200 ng/reaction) was mixed into 5 μL 5 PRIME HotMasterMix (Eppendorf), the selected primers (0.16 μmol/L), and autoclaved water in a total reaction volume of 50 μL per well. PCR amplification was performed in a GeneAmp PCR System 9700 thermocycler (Applied Biosystems). We used Primer 3 software (version 4.0.0; available at http://primer3.sourceforge.net/) to design primers for PCR reactions. Forward primers were modified at the 5′ end with the addition of the 6′-FAM fluorochrome (IVS3polyA and IVS17bCA) or the Applied Biosystems fluorochromes NED (IVS8CA and IVS10CA), VIC (IVS4polyA and IVS17bTA), and PET (IVS1CA). The thermal profile was 95 °C for 1.5 min for enzyme activation, followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s, and elongation at 65 °C for 1 min. A final elongation of 5 min at 65 °C followed the last PCR cycle. All amplification products were checked by electrophoresis in an agarose gel (20 g/L) with the Tris-borate-EDTA buffer described above. One microliter of amplification products was diluted with 19 μL of autoclaved DNase-free water and mixed with 12.5 μL of Hi-Di™ Formamide (Applied Biosystems) and 1.5 μL of GeneScan™ 500 LIZ™ Size Standard (Applied Biosystems). Capillary electrophoresis was performed on the Applied Biosystems 3130 Genetic Analyzer. GeneMapper™ (version 4.0; Applied Biosystems) was used to collect data, track lanes, and measure sizes of fragments used as controls. The number of repeats obtained for each DNA sample was also estimated by sequencing fragments of different sizes. The sequences of the primers used and the PCR protocols are described in Table 1 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol55/issue7. Detailed capillary electrophoresis protocols are available upon request.
STATISTICAL ANALYSIS

The PowerMarker program (version 3.25) was used to reconstruct the haplotype of each individual from the collected genotypic data (15). The algorithm implemented in the software is the iterative expectation-maximization method (16). The frequency of the alleles, heterozygosity (both expected and observed; heterozygosity is the frequency of heterozygotes in the studied population), the power of discrimination (the probability that 2 individuals chosen at random from a given population have different genotypes), and the polymorphism information content (defined as the probability that the marker genotype of a given offspring will allow deduction, in the absence of crossing over, of which of the 2 marker alleles of the carrier parent it received) were evaluated for each locus by means of a specific function (17) on MATLAB R2008a (The MathWorks) that evaluates Hardy–Weinberg proportions with the Monte Carlo naive method (18).

Results

The analysis can be performed within 5 h of sampling (including DNA extraction). Ten DNA samples bearing different haplotypes were independently tested by 2 operators; the results were concordant and corresponded in all cases to those previously obtained by direct DNA sequencing. Fig. 1 shows 2 examples of electropherograms obtained from the analysis of the 7 polymorphisms. The alleles of all polymorphisms can be easily identified, and the use of different fluorochromes allows resolution of the alleles of 2 of the polymorphic loci (IVS3polyA and IVS4polyA) that show some overlap in the electrophoresis run. For each run, we used a panel of 16 labeled single-stranded fragments of different sizes to help in the interpretation of allele length. These markers, which appear as a different color, are not shown in the electropherograms in Fig. 1.
We applied the procedure to the analysis of DNA extracted from cells obtained by laser capture. In particular, we performed 12 experiments with DNA extracted from 1 cell and clearly resolved the haplotype for the 7 microsatellites in all but 2 cases (see Fig. 1B for an example); in 2 experiments, the DNA was not satisfactory to obtain PCR amplification. Similarly, we performed 10 experiments with DNA extracted from 2 cells and clearly resolved the haplotype for the 7 microsatellites in all but 1 case; in 1 experiment, the DNA was not satisfactory to obtain PCR amplification.

We analyzed DNA extracted from blood cells of CF patients and controls. Table 1 shows the allelic distribution of the 3 novel polymorphisms and that of the 4 previously identified microsatellites for (a) 90 CF chromosomes bearing the F508del mutation, (b) 102 CF chromosomes bearing non-F508del mutations, and (c) 98 non-CF chromosomes.

The IVS3polyA polymorphism has 3 alleles (Table 1). The 18-repeat allele is present in 96.7% of F508del chromosomes, the 19-repeat allele was the most frequent among non-CF chromosomes, and the 17-repeat and 19-repeat alleles were equally distributed among non-F508del CF chromosomes. Finally, the 17-repeat allele was identified in 3 chromosomes bearing the F508del mutation and in 1 chromosome bearing the N1303K mutation.

The IVS4polyA polymorphism also has 3 alleles (Table 1). The 19-repeat allele was the most frequent in all groups of chromosomes. All of the other CF and non-CF chromosomes had the 18-repeat allele, except for 1 non-CF chromosome, which had the 17-repeat allele.

The IVS10CA repeat polymorphism has 9 alleles with 17–25 CA repeats (Table 1). The 23-repeat allele was present in 94.4% of F508del chromosomes. Eight of 9 alleles were present at different frequencies, both in CF non-F508del chromosomes and in non-CF chromosomes. The 25-repeat allele was present in only 2 F508del chromosomes.

For the 4 previously described polymorphisms, the IVS1CA polymorphism has 9 alleles. The 21-repeat allele was present in all but 1 chromosome bearing the F508del mutation and in about 50% of the chromosomes bearing other mutations. The same allele was more rare (about 20%) in non-CF chromosomes (Table 1).

The IVS8CA repeat polymorphism has 10 alleles with 15–24 CA repeats. About 61% of F508del chromosomes had the 23-repeat allele, and about 32% of the chromosomes had the 17-repeat allele. Among non-F508del CF chromosomes, the 16-repeat and 23-repeat alleles were the most frequent, followed by the 17-repeat allele. The 16-repeat allele was detected in 86.7% of non-CF chromosomes.

The IVS17bTA repeat polymorphism features 16 alleles (even if other studies detected other alleles) distributed at different frequencies in the 3 populations (Table 1). Eight non-F508del CF chromosomes (7.8%) had large gene deletions (i.e., dele14b-17b, dele17a-17b, and dele17a-18) and did not show the IVS17bTA allele (which is deleted).

Finally, the IVS17bCA repeat polymorphism has 6 alleles (Table 1), with the 13-repeat allele being present in 97.8% of CF chromosomes with the F508del mutation. The same allele was also the most frequent in non-F508del CF chromosomes and in non-CF chromosomes. Six CF chromosomes had the 17a-18 deletion that included the 17bCA locus.

We calculated the heterozygosity of each polymorphism in the 2 populations (i.e., 96 CF patients and 49 healthy controls; Table 2). The IVS17bTA repeat polymorphism and the novel IVS10CA repeat polymorphism showed the highest heterozygosities in CF patients (0.74 and 0.73, respectively) and in healthy individuals (0.49 and 0.61, respectively). Moreover, these 2 polymorphisms also have the highest values for the power of discrimination and polymorphism information content. The IVS8CA polymorphism has a high heterozygosity in CF patients but a low heterozygosity in healthy individuals, whereas the IVS1CA polymorphism has a high heterozygosity (0.70) in healthy individuals. The other polymorphisms have lower heterozygosities and discriminatory power.

At least 1 of the 4 previously identified microsatellites (IVS1CA, IVS8CA, IVS17bTA, and IVS17bCA repeats) was informative in 78 of 100 CF families; at least 1 of all 7 polymorphisms was informative in 98 of 100 families.

Fig. 2 illustrates a pedigree for a prenatal diagnosis of CF, 1 of the most frequent types of cases observed in clinical practice, i.e., that in which only 1 mutation is known (F508del) in the CF-affected child. Analysis of the 7 polymorphisms permitted a clear definition of the haplotype for each parental chromosome and confirmed the diagnosis.

Finally, we defined the haplotype for the 7 microsatellites associated with all of the chromosomes from CF patients and non-CF individuals (see Table 3 in the online Data Supplement). We identified 143 different haplotypes, and most haplotypes seemed to be specifically associated with 1 of the 3 groups of alleles (CF alleles bearing the F508del mutation, CF alleles bearing non-F508del mutations, and non-CF alleles). The 98 chromosomes from non-CF individuals were associated with 59 different haplotypes. Similarly, most frequent CFTR mutations are associated with more haplotypes (26 different haplotypes for the 90 chromosomes bearing the F508del mutation, 13 alleles for the 20 chromosomes bearing the N1303K mutation, and 7 haplotypes for the 10 chromosomes bearing the G542X mutation). All of the haplotypes observed in chromosomes bearing each mutation seem to be derived from an ancestor haplotype via slip-
Table 1. Allelic frequencies of the 7 CFTR polymorphisms in 90 CF chromosomes bearing the F508del mutation (A), 102 CF chromosomes bearing non-F508del mutations (B), and 98 non-CF chromosomes (C).a

<table>
<thead>
<tr>
<th>Allele</th>
<th>IVS1CA</th>
<th>Allele</th>
<th>IVS3polyA*</th>
<th>Allele</th>
<th>IVS4polyA*</th>
<th>Allele</th>
<th>IVS8CA</th>
<th>Allele</th>
<th>IVS10CA*</th>
<th>Allele</th>
<th>IVS17bTA</th>
<th>Allele</th>
<th>IVS17bCA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>2.0</td>
<td>0</td>
<td>17</td>
<td>3.3</td>
<td>1.0</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>1.0</td>
<td>15</td>
<td>0</td>
<td>2.0</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>1.0</td>
<td>2.1</td>
<td>18</td>
<td>96.7</td>
<td>49.0</td>
<td>23.5</td>
<td>18</td>
<td>14.4</td>
<td>22.5</td>
<td>10.8</td>
<td>16</td>
<td>1.1</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>2.9</td>
<td>6.3</td>
<td>19</td>
<td>050.0</td>
<td>76.5</td>
<td>19</td>
<td>85.6</td>
<td>77.5</td>
<td>89.2</td>
<td>17</td>
<td>32.2</td>
<td>15.7</td>
</tr>
<tr>
<td>21</td>
<td>98.9</td>
<td>49.0</td>
<td>19.1</td>
<td>18</td>
<td>0</td>
<td>2.0</td>
<td>20</td>
<td>0</td>
<td>11.8</td>
<td>12.2</td>
<td>22</td>
<td>0</td>
<td>2.0</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>31.4</td>
<td>43.6</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>18.6</td>
<td>2.0</td>
<td>23</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>4.9</td>
<td>24.4</td>
<td>20</td>
<td>1.1</td>
<td>0</td>
<td>22</td>
<td>1.1</td>
<td>4.9</td>
<td>2.0</td>
<td>25</td>
<td>0</td>
<td>2.0</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>6.9</td>
<td>2.1</td>
<td>21</td>
<td>3.3</td>
<td>1.0</td>
<td>23</td>
<td>94.4</td>
<td>11.8</td>
<td>2.0</td>
<td>28</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>1.1</td>
<td>1.0</td>
<td>0</td>
<td>22</td>
<td>0</td>
<td>1.0</td>
<td>24</td>
<td>2.2</td>
<td>7.8</td>
<td>0</td>
<td>29</td>
<td>0</td>
<td>2.0</td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>1.0</td>
<td>2.1</td>
<td>23</td>
<td>61.1</td>
<td>39.2</td>
<td>4.1</td>
<td>25</td>
<td>2.2</td>
<td>0</td>
<td>30</td>
<td>14.4</td>
<td>19.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td></td>
<td>31</td>
<td>58.9</td>
<td>25.5</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32</td>
<td>23.3</td>
<td>9.8</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33</td>
<td>1.1</td>
<td>4.9</td>
<td>9.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34</td>
<td>1.1</td>
<td>2.9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
<td>0</td>
<td>0</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>36</td>
<td>0</td>
<td>1.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45</td>
<td>0</td>
<td>1.0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

a Asterisks indicate the 3 novel polymorphisms.
page events (see Table 2 in the online Data Supplement). In addition, the 17 CF chromosomes bearing unknown mutations were associated with 16 different haplotypes (see Table 3 in the online Data Supplement).

**Discussion**

We identified 3 novel intragenic polymorphisms within intronic sequences of the *CFTR* gene. We evaluated the allelic distributions of these polymorphisms and their heterozygosity in a CF population similar to that currently observed in a clinical setting and in a group of unrelated non-CF individuals. The novel IVS10CA polymorphism and the previously described IVS17bTA polymorphism had higher heterozygosities and discrimination power, both in CF patients and in healthy control individuals. The other polymorphisms revealed some interesting data. Although the

---

### Table 2. Observed and expected heterozygosity, power of discrimination (PD), and polymorphism information content (PIC) for the 7 polymorphisms in 96 CF patients (A) and 49 healthy control individuals (B).a

<table>
<thead>
<tr>
<th></th>
<th>IVS1CA</th>
<th>IVS3polyA</th>
<th>IVS4polyA</th>
<th>IVS8CA</th>
<th>IVS10CA</th>
<th>IVS17bTA</th>
<th>IVS17bCA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>Heterozygosity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>0.41</td>
<td>0.70</td>
<td>0.40</td>
<td>0.39</td>
<td>0.02</td>
<td>0.22</td>
<td>0.73</td>
</tr>
<tr>
<td>Expected</td>
<td>0.44</td>
<td>0.70</td>
<td>0.41</td>
<td>0.39</td>
<td>0.31</td>
<td>0.20</td>
<td>0.64</td>
</tr>
<tr>
<td>PD</td>
<td>0.64</td>
<td>0.86</td>
<td>0.57</td>
<td>0.54</td>
<td>0.32</td>
<td>0.25</td>
<td>0.81</td>
</tr>
<tr>
<td>PIC</td>
<td>0.41</td>
<td>0.66</td>
<td>0.34</td>
<td>0.31</td>
<td>0.26</td>
<td>0.20</td>
<td>0.59</td>
</tr>
</tbody>
</table>

*a Asterisks indicate the 3 novel polymorphisms.*

---

![Fig. 2. Example of CF prenatal diagnosis based on the analysis of the 7 polymorphisms.](image-url)
IVS3polyA polymorphism has a limited heterozygosity; it has the 18-repeat allele, which is present in about 97% of chromosomes with the F508del mutation, and this allele is present in <25% of wild-type chromosomes (Table 1). Thus, this marker can be strongly informative in families in which one of the parents is a carrier of the F508del mutation and the other parent is a carrier of an unknown mutation, as in the case presented in Fig. 2, which is a frequent clinical occurrence because the F508del is the most frequent CFTR mutation worldwide (19). The same is true for the IVS1CA polymorphism, which has 9 different alleles, of which the 21-repeat allele is frequent among CF chromosomes and more rare in non-CF controls. For this polymorphism, we confirmed the high heterozygosity observed in other ethnic/geographic groups, particularly in non-CF chromosomes (10). The IVS8CA polymorphism has a lower heterozygosity than the IVS10CA and IVS17bTA polymorphisms; however, its 16-repeat allele is in linkage with nonaffected chromosomes, whereas this allele was more rare in CF chromosomes bearing non-F508del mutations and was observed in only a single F508del chromosome. Thus, the analysis of the IVS8CA polymorphism can be informative in most cases of linkage diagnosis, in which the analysis of polymorphic loci must discriminate the nonaffected chromosome from the CF chromosome in the parents of the CF patient. Finally, for the IVS4polyA polymorphism, we observed a low heterozygosity of 0.02 (vs the expected heterozygosity of 0.31) because of the high number of individuals homozygous for this polymorphism. In fact, the marker is not in Hardy–Weinberg equilibrium in any of the 3 populations of our study. The same is true for the IVS17bTA polymorphism, particularly in the population of control individuals.

The combined use of the 7 polymorphisms instead of the 4 short tandem repeat polymorphisms currently used produces a large increase in the diagnostic power of linkage analysis for CF carrier and prenatal diagnosis, which is required in about 20% of CF families in which 1 or both CFTR mutations have not been identified (7). As shown in Table 3 in the online Data Supplement, 23 of 26 haplotypes observed in CF chromosomes bearing the F508del mutation are not present in non-CF chromosomes or in CF chromosomes bearing other mutations. Thus, most of the chromosomes bearing the F508del mutation are associated with a peculiar haplotype. Similarly, 49 of 59 haplotypes observed among non-CF chromosomes have not been observed in chromosomes bearing CF mutations. Finally, 58 of 72 haplotypes observed in CF chromosomes bearing non-F508del mutations are present in neither non-CF chromosomes nor F508del CF chromosomes. In fact, use of the 7 polymorphisms increased the informative value of the analysis to 98% in 100 CF families, whereas the informative value of the 4 previously identified markers was 78%. Furthermore, combined use of the 7 polymorphisms permits study of the CFTR gene from intron 1 to intron 17b, the interval within which most CFTR mutations lie (1), thus reducing the risk of missing recombination events (20) between the CFTR mutation(s) and the polymorphic alleles.

The procedure we have developed to analyze the 7 markers under the same conditions seems to be efficient for analyzing both DNA extracted from blood cells and DNA obtained from samples of 1 or 2 cells. The assay takes <5 h, including DNA extraction. This short assay time is an advantage when prenatal or additional preimplantation diagnosis is required, given that embryos must be implanted within 24 h from the sampling of blastomeres for molecular analysis (12). Indeed, among monogenic disorders, CF is the most commonly requested for preimplantation genetic diagnosis, but the large number of mutations responsible for CF is the major limitation in molecular diagnosis (12). Other limitations to preimplantation genetic diagnosis are the preferential amplification of alleles, allele dropout, and complete amplification failure (21, 22). The isothermal whole-genome amplification of DNA from 1 or 2 cells by multiplex PCR that we used in our protocol reduces preferential amplification and allele dropout (21). In fact, in DNA samples from 1 or 2 cells, we clearly resolved alleles of different sizes (7–45 repeats for the IVS17bTA polymorphism). The use of segregation analysis rather than mutation searching in preimplantation diagnosis has the advantage that a single test is applicable to all families, independent of the CFTR genotype. Furthermore, segregation analysis may be used in addition to direct mutation search for prenatal diagnosis, as suggested by European CF guidelines (23). In this case, the same analysis also provides control of internal amplification, maternal contamination, and paternity.

Finally, the haplotype distributions in chromosomes from the controls and those bearing CF mutations warrant some comment. We identified 143 different haplotypes (59 of which are found in 98 non-CF chromosomes; see Table 3 in the online Data Supplement). The haplotype 22-19-16-18-30-13 (IVS1CA, IVS3polyA, IVS4polyA, IVS8CA, IVS10CA, IVS17bTA, IVS17bCA) was the most frequent in non-CF chromosomes, and a number of other haplotypes seem to have been derived from this ancestor haplotype via slippage phenomena. About 25% of non-CF chromosomes bear the 22-19-16-18-7-17 haplotype or related haplotypes, which presumably have been derived from the ancestor haplotype via a recombination event. Chromosomes bearing the F508del muta-
Identification of 3 Novel Polymorphic Repeats in CFTR

Authors’ Disclosures of Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: Grants from the Italian Cystic Fibrosis Research Foundation, grant 19/07 with the contribution of the Latina delegation; Ministero della Salute and Regione Campania (L.229/99, 2004-05 and L.502/94, 2004, 2005); Regione Campania (DGRG 23/62/07); and MIUR (Ministry of University and Research), PS 35–126/Ind.

Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References


Clinical Chemistry 55:7 (2009) 1379