Multiplex Allele-Specific Fluorescent PCR for Haplotyping the IVS8 (TG)$_m$(T)$_n$ Locus in the CFTR Gene

Catherine Costa,1* Jean-Marc Costa,2 Josiane Martin,1 Brigitte Boissier,1 Michel Goossens,1 and Emmanuelle Giordon1

1 Groupe hospitalier Henri Mondor-Albert Chenevier, AP-HP, Laboratoire de Génétique Moléculaire et Unité Inserm U841, Créteil, France; 2 Hôpital Américain de Paris, Laboratoire de Biologie Moléculaire, Neuilly-sur-Seine, France; * address correspondence to this author at: Groupe hospitalier Henri Mondor-Albert Chenevier, AP-HP, Laboratoire de Génétique Moléculaire et Unité Inserm U841, équipe 11, 94010 Créteil, France. Fax: 33 149 812 842; e-mail: catherine.costa@hmn.aphp.fr.

BACKGROUND: Precise genotyping of the intron 8 poly(TG) and poly(T) tracts of the cystic fibrosis transmembrane conductance regulator (CFTR) gene is of clinical relevance in CFTR pathology. The (TG)$_m$ locus influences the penetrance of the (T)$_5$ allele, which may be associated with male infertility by congenital bilateral absence of the vas deferens (CBAVD) or other CFTR-related disorders (CFTR-RD), in particular in CBAVD patients to identify those carrying the (TG)$_{12}$ or, more importantly, the (TG)$_{13}$ allele in trans with a severe mutation, since these patients might develop other symptoms related to mild CF and thus need clinical evaluation and long-term follow-up. Moreover, a new (TG)$_{15}$(T)$_{5}$ allele has been described recently in CF twins, in cis with a CF-causing mutation (11).

Several methods have been developed for (TG)$_m$(T)$_n$ genotyping, which can be multistep and time consuming, not always designed to determine phase, and difficult to interpret (1, 10, 12–14). We describe a new method for simultaneous poly(TG) and poly(T) genotyping in a single assay that relies on a multiplex allele-specific fluorescent PCR and allows direct phase detection.

We first set up the multiplex allele-specific fluorescent assay using 52 known DNA samples representing various IVS8 (TG)$_m$(T)$_n$ genotypes that had been previously determined using a validated strategy relying on a combination of denaturing gradient gel electrophoresis (DGGE) and DNA sequencing (12) (Table 1). The test was performed in parallel in 2 laboratories (Créteil and Neuilly) after the samples had been anonymized.

RESULTS: We observed a 100% match in both validation steps. Results found in CBAVD and CFTR-RD patients are in keeping with data in the literature.

CONCLUSIONS: The assay proved to be simple, rapid, and accurate for single-test (TG)$_m$(T)$_n$ genotyping and suited for analysis in clinical laboratories.

Cystic fibrosis transmembrane conductance regulator gene (CFTR) mutations have been associated with a wide spectrum of phenotypes, ranging from classic cystic fibrosis (CF) to atypical presentations and CFTR-related disorders (CFTR-RDs), including male infertility by congenital bilateral absence of the vas deferens (CBAVD) (1–3), disseminated bronchiectasis (4, 5), and chronic pancreatitis (6, 7). More than 1500 gene defects have been described (www.genet.sickkids.on.ca/cftr/), with effects ranging from mild to severe. The (T)$_5$ variant of the polyphenylalanine tract of intron 8 (IVS8) acceptor splice site—contrary to the other common, neutral (T)$_7$ and (T)$_9$ alleles—affects the normal splicing of exon 9 and is considered a CFTR-RD–associated defect with incomplete penetrance. It has a frequency of 19% to 37% in CBAVD patients (1–3, 8).

The disease penetrance of the (T)$_5$ variant is influenced by the adjacent (TG)$_m$ locus, the shorter (TG)$_m$ associated with shorter (T)$_n$ repeats being less favorable for exon 9 splicing efficiency (9). Penetration of the (TG)$_{11}$(T)$_5$, (TG)$_{12}$(T)$_5$, and (TG)$_{13}$(T)$_5$ haplotypes in trans with a CF-causing allele have been assessed as having 11%, 78%, and 100% disease penetrance, respectively (10), making discrimination between these variants of clinical relevance. It is especially important in CBAVD patients to identify those carrying the (TG)$_{12}$ or, more importantly, the (TG)$_{13}$ allele in trans with a severe mutation, since these patients might develop other symptoms related to mild CF and thus need clinical evaluation and long-term follow-up. Moreover, a new (TG)$_{15}$(T)$_{5}$ allele has been described recently in CF twins, in cis with a CF-causing mutation (11).

Several methods have been developed for (TG)$_m$(T)$_n$ genotyping, which can be multistep and time consuming, not always designed to determine phase, and difficult to interpret (1, 10, 12–14). We describe a new method for simultaneous poly(TG) and poly(T) genotyping in a single assay that relies on a multiplex allele-specific fluorescent PCR and allows direct phase detection.

We first set up the multiplex allele-specific fluorescent assay using 52 known DNA samples representing various IVS8 (TG)$_m$(T)$_n$ genotypes that had been previously determined using a validated strategy relying on a combination of denaturing gradient gel electrophoresis (DGGE) and DNA sequencing (12) (Table 1). The test was performed in parallel in 2 laboratories (Créteil and Neuilly) after the samples had been anonymized.

3 Nonstandard abbreviations: CF, cystic fibrosis; CFTR, cystic fibrosis transmembrane conductance regulator; CFTR-RD, CFTR-related disorder; CBAVD, congenital bilateral absence of the vas deferens; DGGE, denaturing gradient gel electrophoresis.
The assay was then applied in Créteil for genotyping 170 new samples: 62 healthy individuals from the general population (normal controls) and 108 patients referred for CFTR gene studies in a diagnostic setting because of CBAVD (n/H1005 68) or pulmonary disease (n/H11005 40), mainly disseminated bronchiectasis. Written informed consent to perform CFTR studies was obtained from the patients at the time of referral to the laboratory.

We extracted genomic DNAs from EDTA anticoagulated whole blood samples using varied protocols, mostly a phenol-chloroform method or commercial kits (Nucleon, BACC3, Amersham Biosciences; DNA High Pure precipitation kit, Roche Diagnostics; FlexiGene, Qiagen). The multiplex allele-specific fluorescent PCR is based on the use of 3 different primers, specific for (T)5, (T)7, or (T)9 repeats and labeled with 3 different fluorophores, and a common primer allowing simultaneous amplification in a single assay of the poly(TG) and poly(T) tracts. The multiplex PCR reactions were performed on a GeneAmp PCR system 9700 (Applied Biosystems) in 25-μL reactions using the Qia- gen Multiplex PCR kit, following the manufacturer’s instructions, with 50 ng DNA and a mix of 0.25 μmol/L of each primer: forward primer, 5’-ACATAAAACAAGCATCTATTGAAA; reverse primers, (T)5: 5’-Hex-CCAAATCCCTGTTAAAAACA, (T)7: 5’-Fam-CCAAATCCCTGTTAAAAACA, (T)9: 5’-Ned-CCAAATCCCTGTTAAAAACA. The reaction started with an initial denaturation of 15 min at 95 °C, followed by 35 cycles at 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 45 s and a final extension step of 10 min at 72 °C. We added 2 μL of the purified PCR products to 9.8 μL formamide and 0.2 μL Genescan-500 Rox or Tamra size standard (Applied Biosystems). The fluorescent PCR products were sized on a monocapillary sequencer (Neuilly) or a 16-capillary sequencer (Créteil) (ABI Prism 310 or 3100 Genetic Analyzer, Applied Biosystems). The assay can be applied to both sequence analyzer instruments by adapting the fluorophores and Genescan size standard to the system used: Fam and Hex for both; Tet or Ned for ABI 310 or ABI 3100, respectively. The results were processed using Genescan 3.7 software (Applied Biosystems). Each product was identified by its fluorescence color [poly(T)] and size [poly(TG)] (Fig. 1). We first used different known combinations of haplotypes to establish a calibration of the assay before testing the 108 new samples. Then 2 control DNAs were included in each experiment for blind genotyping of the 170 new samples: a (TG)11(T)7 homozygote and a [(TG)10(T)9][(TG)12(T)5] compound heterozygote.

Analysis of the 52 samples previously genotyped showed a 100% match between the 2 laboratories and compared with the reference methods. Similarly, we observed a 100% match when testing the 170 samples in parallel with the new assay and sequencing. Non-specific amplification was not detected, but replication slippage patterns occurred, as commonly observed for microsatellites. Some representative haplotypes are shown in Fig. 1, and the frequency of the different haplotypes identified is shown in Table 1.

In this study, we found the (T)5 variant in 24% of CBAVD alleles, with (TG)13(T)5 and (TG)12(T)5 in 24% and 60% of (T)5 alleles, respectively, a result in keeping with our recent data (3). In patients with

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Validation study, n*</th>
<th>Normal controls</th>
<th>CBAVD</th>
<th>Other CFTR-RD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(TG)9(T)9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(TG)9(T)9</td>
<td>18</td>
<td>9 (7.3)</td>
<td>46 (33.8)</td>
<td>20 (25.0)</td>
</tr>
<tr>
<td>(TG)10(T)7</td>
<td>17</td>
<td>32 (25.8)</td>
<td>23 (16.9)</td>
<td>13 (16.2)</td>
</tr>
<tr>
<td>(TG)11(T)9</td>
<td>6</td>
<td>4 (3.2)</td>
<td>3 (2.2)</td>
<td>3 (3.7)</td>
</tr>
<tr>
<td>(TG)11(T)7</td>
<td>22</td>
<td>58 (46.8)</td>
<td>27 (19.9)</td>
<td>27 (33.8)</td>
</tr>
<tr>
<td>(TG)13(T)9</td>
<td>11</td>
<td>15 (12.1)</td>
<td>4 (2.9)</td>
<td>8 (10.0)</td>
</tr>
<tr>
<td>(TG)13(T)5</td>
<td>12</td>
<td>5 (4.0)</td>
<td>5 (3.7)</td>
<td>4 (5.0)</td>
</tr>
<tr>
<td>(TG)12(T)9</td>
<td>12</td>
<td>1 (0.8)</td>
<td>20 (14.7)</td>
<td>5 (6.3)</td>
</tr>
<tr>
<td>(TG)12(T)5</td>
<td>5</td>
<td>0</td>
<td>8 (5.9)</td>
<td>0</td>
</tr>
<tr>
<td>Total (T)5</td>
<td>29</td>
<td>6 (4.8)</td>
<td>33 (24.3)</td>
<td>9 (11.3)</td>
</tr>
<tr>
<td>Total alleles</td>
<td>104</td>
<td>124</td>
<td>136</td>
<td>80</td>
</tr>
</tbody>
</table>

* As the 52 samples used for the first validation step were selected on their genotypes, percentages are not relevant.
a pulmonary disease suggesting CF, the (T)5 variant accounted for 11.2% of alleles, with (TG)12(T)5 observed in 55.5% of (T)5 alleles. The absence of (TG)13(T)5 may be due to the limited sample size. It would be interesting to investigate further the distribution of (T)5 haplotypes in CFTR-related pulmonary disorders, as little data is available (15). Frequencies of the (T)5 allele and of (TG)m(T)n combinations found in healthy individuals were similar to those previously reported in our laboratory in non-CF chromosomes of CF carriers (12).

Among the methods for detection of all (TG)m(T)n alleles, sequencing is the most frequently used. Interpretation may be greatly helped by prior determination of the poly(T) tract, which implies 2 steps, but some cases remain difficult to interpret or require family analysis. DGGE analysis has the advantage of assessing both haplotypes, but it is time consuming and the interpretation is tedious (12). More recently, melting curve analysis of hybridization probes has appeared as an attractive method for (TG)m(T)n genotyping through a single-step assay (16); however, in our experience, some haplotypes cannot be determined accurately (17). Recently, poly(TG) tract genotyping in patients already known to carry the (T)5 variant has been described using single allele-specific fluorescent PCR (14). It is thus a second-step method, and non-specific PCR products may be obtained for non-(T)5 alleles, probably because the test is restricted to the detection of (T)5 alleles. The assay described here follows the principle of allele-specific fluorescent PCR, but has been designed to amplify specifically the common (T)5, (T)7, and (T)9 alleles, thus preventing non-specific amplification, and is able to detect all linked (TG) repeats. It cannot truly detect other poly(T) variants such as (T)3 (18), (T)6 (19), or (T)11 (20). These

Fig. 1. (A), Electropherograms obtained in different combinations of haplotypes. Each haplotype is defined by its color, specific for (T)n, and size, indicative of (TG) length. (B), Expected values for the haplotypes observed in our series. No value is indicated for those not observed in our laboratory or described in the literature.
variants are indeed very rare, as none has been identified in the 18 years’ experience of CFTR gene studies in our laboratory. Of these, only the (T)₃ variant may be of diagnostic value and could easily be included by adding a (T)₃ allele–specific primer. Direct (TG)ₘ(T)ₙ genotyping prevents the need for linkage family studies, which is really an advantage, especially in CBAVD and other CFTR-RDs where family members may not be available. Moreover, single-test analysis makes our assay a method of choice when screening for the (T)₃ variant, as needed in diagnostic settings. Not all tests designed for detection of frequent CF mutations have included it in their panel. Such tests are to be used as a first step in dealing with different situations for CFTR testing because, as the (T)₃ variant is not a classic CF mutation, it should not be tested for carrier testing but considered for diagnostic purposes only.

In conclusion, the assay we designed for single test (TG)ₘ(T)ₙ genotyping proved to be simple, rapid, and inexpensive to use and suited to analysis of a large number of samples, in specialized and routine laboratories for CFTR gene studies. It has proved to be of accurate and reliable interpretation compared with the other methods used or tested in our laboratory, thereby superseding them.

Grant/Funding Support: None declared.
Financial Disclosures: None declared.

Acknowledgment: David Kerridge is acknowledged for rereading the manuscript.

References


DOI: 10.1373/clinchem.2008.103259