XV-2c and KM.19 haplotype analysis in Chilean patients with cystic fibrosis and unknown CFTR gene mutations

GABRIELA M REPETTO1, ALONSO R PUGA1 and IRIS DELGADO2

1 Programa de Genética Humana and 2 Instituto de Epidemiología y Políticas Públicas, Facultad de Medicina, Clínica Alemana-Universidad del Desarrollo, Santiago, Chile

ABSTRACT

Cystic fibrosis (CF) is caused by mutations in the CFTR gene. More than 1600 mutations have been described, with frequencies that differ worldwide according to the ethnic origin of patients. A small group of mutations are recurrent on several populations. It has been shown that they each tend occur on specific chromosome 7 haplotypes, supporting the notion of a single origin for them. Less than 50% of mutations in Chilean patients have been identified to date. To indirectly assess the possible presence of a predominant founder mutation in the remaining unknown alleles, we evaluated 2 polymorphic markers, XV-2c and KM.19, tightly linked to the CFTR locus. The study was done in Chilean CF patients with unknown or deltaF508 (ΔF508) CFTR mutations and their haplotypes were compared to affected family-based controls. ΔF508 showed marked linkage disequilibrium with XV-2c/KM.19 haplotype B, with 90% of alleles on that haplotype. There was no difference in haplotype distribution between unknown mutations and normal controls. These results support a European origin for ΔF508 alleles in Chilean patients, and make unlikely the presence of a predominant founder mutation in the so-far unknown alleles.

Key terms: CFTR gene, cystic fibrosis, ΔF508, haplotype analysis.

INTRODUCTION

Cystic fibrosis (CF) is one of the most common life-threatening autosomal recessive diseases. It is a complex multisystem disorder, caused by mutations of the gene encoding for the cystic fibrosis transmembrane conductance regulator (CFTR), located on chromosome region 7q31. More than 1600 disease-causing mutations have been described worldwide, but only a small number, approximately 30 of them, have individual frequencies higher than 0.1% in CF patients and the rest have been observed in a few families each (CFGAC, 1994).

The CFTR gene was identified by positional cloning (Beaudet et al., 1989; Rommens et al., 1989). The gene was found to be tightly linked to the D7S23 locus. Polymorphisms in this locus were initially detected by restriction fragment length polymorphisms (RFLP) using probes XV-2c and KM.19, located 175 and 125 kb upstream of the 5' end of the gene. Prior to the discovery of the gene, it was shown that close to 90% of European CF patients shared a common XV-2c/KM.19 (XK) haplotype. This haplotype, designated “B”, was only found in 16% of normal individuals (Beaudet et al., 1989; Kerem et al., 1989). Subsequent research confirmed that this observation reflected the presence of strong linkage disequilibrium, that is, non-random association of alleles at two or more loci on a chromosome, between the most common European mutation, deltaF508 (ΔF508), and haplotype B. This finding supported the theory of a single origin of this mutation (Dork et al., 1992; Kerem et al., 1989; Highsmith et al., 1990; Morral, Bertranpetit et al., 1994). Several
other common mutations are predominantly present on one haplotype. For example, G542X, a mutation of high prevalence in Mediterranean countries and having the third highest frequency in Chilean patients, is also associated with XK haplotype B (Loirat et al., 1997). Mutations N1303K, G551D, W1282X, 394delTT and 3120+1 G>A also appear to have a single origin each (Cashman et al., 1995; Osborne et al., 1992; Dork et al., 2000; Kerem et al., 1995; Shoshani et al., 1992; Dork et al., 1998). In contrast, other mutations have been found on several haplotypes, suggesting multiple origins for them (Morral, Llevadot et al., 1994).

The incidence of the disease and the distribution of the mutations differ among populations according to their ethnic origin. A group of 32 common mutations accounts for 83% of mutant alleles in Caucasian population in the USA, percentage largely attributable to the large proportion of ΔF508 alleles (Heim et al., 2001). In Chilean CF patients, the same panel of 32 mutations has a detection rate of only 50% (Puga et al 2007, manuscript under review). The remainder of local mutations has not been identified, mainly because the ones evaluated so far correspond to mutations that are common in European and North American populations (Rios et al., 1994; Repetto et al., 2001; Molina et al., 2002; Puga et al., 2007). This hinders diagnostic confirmation, particularly in atypical cases, and carrier identification.

The mutations not yet identified in our population could correspond to several of low individual frequency or alternatively, there could be a single or more common founder mutation originated on a specific haplotype. To assess these possibilities, we evaluated XV-2c/KM.19 haplotype distribution in Chilean CF patients with unknown CFTR mutations and compared them to ΔF508 and normal CFTR alleles.

METHODS

Peripheral blood samples were collected from patients registered in the National Cystic Fibrosis Program of Ministerio de Salud de Chile who fulfilled diagnostic criteria for CF (Rosenstein et al., 1998), consisting in typical involvement of at least one organ system and sweat chloride concentration equal or greater than 60 mEq/L. We only included patients that had had molecular testing in our previous study of 32 common mutations (Puga et al., 2007, manuscript under review) and evaluated 3 groups of CF patients: disease alleles not identified, ΔF508/unknown mutation heterozygotes, and ΔF508/ΔF508 homozygotes.

To construct the XK haplotypes, blood samples from both parents were also obtained and analyzed. The protocol was approved by Comité de Ética para la Investigación, Facultad de Medicina, Clínica Alemana-Universidad del Desarrollo. Written informed consent was obtained from all participating individuals and/or their parents.

Genomic DNA was extracted from samples using Easy-DNA kit (Invitrogen, Carlsbad, CA, USA). The polymorphisms were evaluated by restriction enzyme analysis of PCR products, which allows to discriminate between two alleles at each marker based on the presence or absence of a restriction site. To analyze marker XV-2c, genomic DNA was amplified using primers XV-2C (f): 5'-GTT GAA GTG AAT TGA ATG-3' and XV-2C(r): 5'-GTT AAA CTA TGT CAA AG-3'. The products were digested with restriction enzyme Taq I (Rosenbloom et al., 1989). For KM.19, the primers used were KM-19(f): 5'-GCT GCA TCA TAT AAG TTG CC-3' and KM-19(r): 5'-GTT AAA CTA TGT CAA AG-3'. The products were digested with restriction enzyme Pst I (Anwar et al., 1990). PCR and restriction products were separated by electrophoresis on 2.5% agarose gels. The order of the markers on chromosome 7 and their observed recombination fractions is centromere/ XV-2c (0.001) KM.19 (0.006) CFTR/telomere (Lathrop et al., 1988; Beaudet et al., 1989). The polymorphism was analyzed with restriction enzyme Pst I (Anwar et al., 1990). PCR and restriction products were separated by electrophoresis on 2.5% agarose gels. The order of the markers on chromosome 7 and their observed recombination fractions is centromere/ XV-2c (0.001) KM.19 (0.006) CFTR/telomere (Lathrop et al., 1988; Beaudet et al., 1989). Haplotypes for each individual were constructed by segregation analysis of the results of the members of each family. XK haplotypes were labeled as described by Beaudet et al (1989) (see table I).
Data were compiled and analyzed using SPSS version 12 (SPSS Inc., Chicago, IL). To avoid population stratification bias, we used the affected family-based controls (AFBAC) method of analysis (Thomson, 1995), in which haplotypes not transmitted by a healthy carrier parent to an affected child constitute the control or normal population. The haplotype distributions of the mutation-unknown, ΔF508 and normal chromosomes were compared by χ². Two-tailed p-values < 0.05 were considered to be statistically significant.

RESULTS

Samples from 52 unrelated patients with CF were analyzed, distributed according to genotypes as follows: 16 had both mutations unknown, 7 were ΔF508/unknown mutation compound heterozygotes and 29 were ΔF508 homozygotes. Regarding family structure, 21 trios were complete, that is, both non-consanguineous parents were analyzed. Five families had no paternal sample available. Parental samples were unavailable for twenty-six ΔF508 homozygotes. Three of the complete trios and one child/mother duo were uninformative, i.e., the parental and/or child’s haplotypes could not be clearly established and were not included in the analysis. This resulted in 37 alleles with unknown CFTR mutations, 63 bearing ΔF508 and 40 control haplotypes available for analysis.

Results for XV-2c and KM.19 RFLP analysis from a representative family and an isolated ΔF508 homozygous individual are shown in Figure 1. The father on family #005 (Fig. 1A) was homozygous for absence of the restriction site at markers XV-2c and KM.19, therefore labeled “1” for both alleles at both markers, this corresponds to haplotype A. The mother was homozygous for the presence of the Taq I restriction site in marker XV-2c (therefore “2”) and homozygous for the absence of the Pst I site at KM.19 (“1”). This combination corresponds to haplotype C. Their affected child inherited an A haplotype from his father and a C haplotype from the mother, as expected. The parental non-transmitted A and C haplotypes were counted as controls. The RFLP results of a ΔF508 homozygote are shown in Figure 2B. The individual was homozygous for absence of Taq1 restriction site at XV-2c (“1”) and presence of the site at KM.19 (“2”) and therefore was homozygous for haplotype B.

The distribution of XK haplotypes among the unknown mutations and normal chromosomes showed no significant differences, as shown in Table II. Most alleles in both groups were on haplotypes A and C, followed in frequency by B. Only 5% of normal alleles or those bearing unknown mutations were on haplotype D. In contrast, there was marked predominance of haplotype B in ΔF508 chromosomes, with only one allele in each of A and C haplotypes and 4 alleles on haplotype D. A statistically significant difference was observed between the distribution of these haplotypes and that of normal chromosomes ($\chi^2 = 66.615; p < 0.000001$, df 3).

DISCUSSION

This study evaluated XK haplotypes in Chilean patients with CF but unknown mutations, and compared the results with normal, family-based control haplotypes, as well as ΔF508-bearing chromosomes.

No difference was observed in haplotype distribution between normal and unknown-mutant chromosomes. This implies that we
found no evidence of association of mutant CF chromosomes with a specific haplotype in this group of patients with the disease and unidentified causative mutations. This lack of association contrasts with what has been observed for several of the common mutations (Cuppens et al., 1994). The opposite was seen in Chilean ΔF508 chromosomes in our study, with over 90% of them present on XK haplotype B. This finding evidences marked linkage disequilibrium between them, as it has been shown in other populations (Kerem et al., 1989; Highsmith et al., 1990; Dork et al., 1992; Morral, Bertranpetit et al., 1994). A European origin for the ΔF508 mutation observed in Chilean patients is therefore suggested.

**Figure 1:** Restriction enzyme analysis in a family (A) and an isolated case (B). Electrophoresis of restriction products run in 2.5% agarose and stained with SYBR Green. L, molecular weight ladder. Chromosome 7 regions studied for each individual are represented as vertical lines. The number in each allele represents absence (1) or presence (2) of the restriction site for Taq I or Pst I in markers XV-2c (top panel) or KM.19 (bottom panel), respectively. XK haplotypes were labeled as described on Table 1. Panel A: Family #005, child with unknown CFTR mutations (“f”: father, “x”: affected and “m”: mother), with their constructed haplotypes at the bottom. Panel B: Patient #299, a ΔF508 homozygote, with the constructed haplotypes shown on the right.

**TABLE II**

Haplotype distribution among unknown, ΔF508 and normal CF alleles, shown as number of alleles with percentages in parentheses

<table>
<thead>
<tr>
<th>Mutation</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown *</td>
<td>12 (32.4)</td>
<td>8 (21.6)</td>
<td>15 (40.5)</td>
<td>2 (5.4)</td>
<td>37 (100)</td>
</tr>
<tr>
<td>DF508 **</td>
<td>1 (1.6)</td>
<td>57 (90.5)</td>
<td>1 (1.6)</td>
<td>4 (6.3)</td>
<td>63 (100)</td>
</tr>
<tr>
<td>Normal</td>
<td>18 (45.0)</td>
<td>6 (15.0)</td>
<td>14 (35.0)</td>
<td>2 (5.0)</td>
<td>40 (100)</td>
</tr>
</tbody>
</table>

* $\chi^2 = 1.405; p=0.704, df 3$ vs. normal.
** $\chi^2 = 66.615; p < 0.000001, df 3$ vs. normal.
Haplotype heterogeneity for unknown mutations has been observed in other mixed Latin American populations. In Mexican patients, Orozco et al. (2001) showed that 97% of ΔF508 chromosomes are on haplotype B, but no difference in haplotype distribution between unknown mutations and normal chromosomes was found. In Brazilian patients, Raskin et al. (1997) showed that 88% of ΔF508 and only 41% non-ΔF508 were on B haplotypes, though they did not report the distribution in normal population nor describe what constituted the non-ΔF508 group.

Our findings suggest that it is unlikely that a single common mutation could account for the so-far undetected alleles in the Chilean population. Alternatively, if a single or markedly predominant mutation exists, the lack of linkage disequilibrium could be due to recombination between the markers and the mutation, a remote possibility due to the known low recombination fraction between them. Alternatively, a putative common mutation could have multiple origins, but the available worldwide data supporting tight linkage disequilibrium between most common mutations and close haplotypic frames make this improbable (Estivill et al., 1987). These possibilities can be evaluated in further studies of intragenic microsatellite and single nucleotide polymorphism (SNP) markers.

Mis-attributed paternity or uniparental disomy were not evaluated in this study. Nevertheless, the inheritance of the alleles showed expected Mendelian transmission in all trios and no child had alleles that were absent in the parents, making these confounders unlikely.

Our results have several implications. First, in terms of individual families, the feasibility of the analysis highlights its potential use to provide relevant information regarding carrier status or presymptomatic testing for unaffected relatives of individuals with unknown genotypes, who cannot currently benefit from direct mutation testing or can only do so at expensive costs (Beaudet et al., 1989; Lemna et al., 1990; Riveros et al., 2005).

Regarding the molecular genetics of the disease in our population, the deduction that there could be several additional alleles evidences the challenges in improvement of sensitivity of molecular testing for Chilean patients. To solve this problem, full sequencing of the gene and re-analysis of the haplotypic frameworks of newly found alleles are required.

Other investigators in Latin American countries that have performed extensive analysis of the CFTR gene in their patients have shown the presence of several private mutations of low individual frequency. Visich et al. (2002) found 62 mutations in 220 Argentinean patients, including 6 novel mutations and reaching a detection rate of 84%. Only 6 of these, all of them worldwide common mutations, showed frequencies higher than 1%. In Colombia, Keyeux et al. (2003) found 6 common and 13 rare mutations, including 6 novel ones, in 92 patients, reaching a detection rate of 64% after denaturing gradient gel eletrophoresis analysis.

No evidence has been found so far of a common Amerindian CF founder mutation. Thus, the disease in our countries probably originated by a group of worldwide common mutations that were introduced from immigrants, as evidenced by their presence on the same haplotypes in European counterparts. In addition, there is an apparently large additional number of mutations, maybe different between countries, and whose origins remain undetermined.

ACKNOWLEDGEMENTS

This work was funded by Dirección de Investigación, Vice-Rectoría Académica, Universidad del Desarrollo, Santiago, Chile. The authors wish to thank the clinicians involved in contacting patients and families for the study, and Mrs. María Luisa Guzmán, RN and Sonia Margarit, MSc, for assistance in blood sample collection.

REFERENCES

asociados fuertemente al gen CFTR. Rev Med Chil 133: 648-654 (In Spanish)
ROSENBLOOM CL, KEREM BS, ROMMENS JM, TSUI LC, WAINWRIGHT B, WILLIAMSON R, O’BRIEN WE, BEAUDET AL (1989) DNA amplification for detection of the XV-2c polymorphism linked to cystic fibrosis. Nucleic Acids Res 17: 7117