Linkage disequilibrium for DNA haplotypes near the cystic fibrosis locus in two South European populations


1Unitat de Genètica Molecular, Fundació d’Investigació Santa Creu i Sant Pau, Pare Claret, 167, E-080025 Barcelona, Spain
2Molecular Genetics Department, St. Mary’s Hospital, Medical School, London W2 1PG, UK
3Istituto di Scienze Biologiche, Facoltà di Medicina e Chirurgia, Università di Verona, Verona, Italy
4Cattedra di Genetica, Università degli Studi di Urbino, Urbino, Italy
5Dipartimento di Sanità Pubblica e Biologia Cellulare II, Università di Roma, Rome, Italy

Summary. Three polymorphic DNA marker loci (INT1L1, D7S23 and D7S399) map to a chromosomal region that is very close to the cystic fibrosis (CF) locus in terms of genetic distance. These marker loci have been used to analyse the linkage disequilibrium in 137 CF families from two South European countries (Italy and Spain). The markers can be analysed for differences in linkage disequilibrium more easily in these populations than in North Europeans, in whom the disequilibrium between the allelic systems defined by the probes and CF is much greater and on a “plateau” through the genetic region. The different levels of disequilibrium found in the studied populations suggest that D7S399 and D7S23 are both closer to CF than INT1L1, and provide additional information on the origins and homogeneity of the CF defect.

Introduction

Cystic fibrosis (CF) is the most common severe autosomal recessive disorder affecting Caucasian populations, with incidence in the range from 1:500 to 1:3000 (Thompson 1980). Genetic linkage analysis using DNA markers has assigned the CF locus to the long arm of chromosome 7 (q31) (Knowlton et al. 1985; Tsui et al. 1985; Wainwright et al. 1985; White et al. 1985; Scambler et al. 1986; Estivill et al. 1986). Two of these marker loci, D7S8 (probe pJ3.11) and MET, frame the CF locus with a genetic distance of less than 1 cM each (Beaudet et al. 1986; Lathrop et al. 1988).

We have “jumped” approximately 600 kb from MET towards the CF locus and isolated the gene int-1-like protein 1 (INT1L1) (Estivill et al. 1987a; Wainwright et al. 1988). Two probes (pPT-3 and pXV-2c) at the INT1L1 gene, and two at the D7S23 locus (probes pCS.7 and pKM.19) detect restriction fragment length polymorphisms (RFLPs; Estivill et al. 1987a, 1987b). pPT-3 is a coding sequence at the 3’ end of the gene INT1L1, pXV-2c is a noncoding sequence in the middle of the gene, pCS.7 includes the CpG island and the first exon of INT1L1, and pKM.19 maps 30 kb upstream from INT1L1 (Estivill et al. 1987a, b).

Offprint requests to: X. Estivill, Unitat de Genètica Molecular, Fundació d’Investigació Santa Creu i Sant Pau, Pare Claret, 167, E-080025 Barcelona, Spain

Linkage disequilibrium between alleles at these loci and CF has improved the informativity of CF families for prenatal diagnosis (Beaudet et al. 1988) and carrier exclusion (Farrall et al. 1987; Krawczak et al. 1988), and suggests that approximately 85% of chromosomes carrying the CF mutation are from a single mutational event (Estivill et al. 1987a, b, 1988; Beaudet et al. 1988). Molecular studies of INT1L1 and the adjacent locus D7S23, and the analysis of families in which there is a recombination event demonstrate that CF maps in the interval between D7S23 and D7S8 (Estivill et al. 1987a, b, 1988; Farrall et al. 1988).

We have isolated a new marker called pMP6d-9 (D7S399), which is approximately 160 kb on the D7S8 side of D7S23 (Estivill et al. 1989). The allelic system detected by the new marker is in strong linkage disequilibrium with both CF and D7S23. However, in the primarily North European population studied, the “plateau” of allelic association between D7S23, D7S399 and CF did not provide information on the position of the CF mutation within this interval.

In order to determine more precisely the location of CF relative to these marker loci using linkage disequilibrium, it is essential to study families from populations where variation in the level of disequilibrium is found. Linkage disequilibrium can provide invaluable information about the evolution and population genetics of a disease if close markers are available. Thus, genetic drift, independent mutations generating allelic heterogeneity, and specific variation of recombination rates are mechanisms that can generate different patterns of allelic associations in different populations (Bodmer 1972; Hill and Robertson 1968).

Many North European countries, such as Denmark, show no variation at all, with every CF chromosome having an identical haplotype; these populations are uninformative with respect to analysis of disequilibrium. We present here an analysis of disequilibrium between the new marker and CF in populations from two South European countries (Italy and Spain), in which we previously found that INT1L1 and D7S23 show lower linkage disequilibrium with CF than in other populations (Estivill et al. 1987b, 1988). The differences in the level of disequilibrium found in these populations suggest alternative locations for CF in relation to D7S399 (pMP6d-9) and D7S23 (pKM.19). In addition, the different distribution of haplotypes on CF chromosomes suggests that independent mutations could be present in South European populations.
Materials and methods

Cystic fibrosis families

DNA samples were obtained from 85 Italian [41 Verona (VER) and 44 Urbino (URB)] and 42 Spanish [Barcelona (BCN)] CF families with at least one CF child. Part of the data on the Spanish families has been reported in a previous study (Estivill et al. 1989) and some of the Italian families have been previously analysed for pXV-2c and pKM.19 (Estivill et al. 1988). Diagnosis was based on clinical features and at least one positive sweat test.

Probes, DNA analysis and haplotypes

The following probes were used: pXV-2c (Estivill et al. 1987a), pKM.19 (Estivill et al. 1987b) and pMP6d-9 (Estivill et al. 1989). For each probe, inserts were excised from low-melting-temperature agarose gels after restriction digestion and were labelled with [α-32P]dCTP, using random oligonucleotide primers (Feinberg and Vogelstein 1984), to a specific activity of 1–2 × 10⁸ dpm. Genomic DNA (5 μg) was digested with restriction enzymes, electrophoresed and transferred to nylon membranes. Hybridisation and autoradiography were as described previously (Maniatis et al. 1982).

Haplotypes for both parental chromosomes (normal and CF allele) were determined using the CF child in each family to establish phase. The correlation coefficient and the relative risk or odds ratio were used as measures of allelic association (Estivill et al. 1987b, 1988).

Results

It was possible to establish phase in 112 out of 137 CF families; these 224 CF chromosomes and 224 normal chromosomes were used in this study. Figure 1 shows the segregation pattern of the MspI polymorphism detected with probe pMP6d-9 in a typical CF family.

The haplotype distribution defined by the three RFLPs for normal and CF chromosomes in the studied South European families is shown in Table 1. A single haplotype, − + + (pXV-2c/TaqI, pKM.19/PstI and pMP6d-9/MspI) is found for 65% of CF chromosomes. Two other haplotypes, + − − and ++ +, are present in 12% and 11% of CF chromosomes, respectively, in the overall population. The distribution of haplotypes on the CF chromosomes is different in the three populations studied, but not significantly so. Haplotype − − − accounts for 8% of CF chromosomes in the BCN families, and for only 1% in the Italian families. A different distribution is also found for haplotype + − −, which is present in 9% of CF chromosomes in the VER families and in 3% and 0% of the BCN and URB families, respectively. Haplotype + + + accounts for at least 13% of the VER and BCN CF chromosomes and for only 4% of the URB CF chromosomes.

Markedly significant disequilibrium (P < 0.01) is found between CF and allelic systems defined by the three loci (Table 2). The highest level of disequilibrium with CF was found with D7S23 (pKM.19) in the URB and BCN families, and with D7S399 (pMP6d-9) in the VER sample (P << 0.001) (Table 2 and Fig. 2). Disequilibrium between INT1L1 (pXV-2c) and CF was lower in the VER and BCN families (P < 0.01) than in the URB families. In spite of this variation in the level of association between CF and the three allelic systems in the studied groups, heterogeneity tests only reached the level of significance for INT1L1 (pXV-2c) (P< 0.05) (Table 3). No heterogeneity was found for INT1L1/CF when the URB data was excluded (data not shown).

Marker/marker linkage disequilibrium was determined for the three loci (Table 4). Significant association is found for all pairs of loci on CF chromosomes (P << 0.001). The lowest

---

Table 1. Haplotypes for the D7S23 and D7S399 loci in Italian and Spanish populations

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>pXV-2c</th>
<th>pKM.19</th>
<th>pMP6d-9 MspI</th>
<th>CF</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TaqI</td>
<td>PstI</td>
<td>MspI</td>
<td>URB</td>
<td>VER</td>
</tr>
<tr>
<td>− − −</td>
<td>− −</td>
<td>− −</td>
<td>− − − −</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>+ − −</td>
<td>+ −</td>
<td>+ −</td>
<td>+ − − +</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>− + −</td>
<td>− +</td>
<td>− +</td>
<td>− + + +</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>− − +</td>
<td>+ −</td>
<td>+ −</td>
<td>+ − + +</td>
<td>54</td>
<td>46</td>
</tr>
<tr>
<td>+ − +</td>
<td>+ −</td>
<td>+ −</td>
<td>+ − + +</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>+ + −</td>
<td>+ +</td>
<td>+ +</td>
<td>+ + + +</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>+ + +</td>
<td>+ +</td>
<td>+ +</td>
<td>+ + + +</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>− + − − −</td>
<td>− +</td>
<td>− +</td>
<td>− + − − −</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>74</td>
<td>76</td>
</tr>
</tbody>
</table>

---

Fig. 1. Segregation of the allelic system detected for pMP6d-9 in a family with two affected CF sibs. pMP6d-9 reveals a diallelic polymorphism (13.0 and 8.5 + 4.5 kb) in MspI digests. DNA (5 μg) was digested with MspI, electrophoresed in a 0.8% agarose gel, and hybridised with pMP6d-9.