Localisation of two candidate genes for mental retardation using a YAC physical map of the Xq21.1–21.2 subbands

Laurence Colleaux, Melanie May, Jérome Belougne, Denis Lepaslier, Charles Schwartz, Michel Fontes

Abstract

Genetic studies in families with X linked mental retardation have suggested the localization of several MR genes in the human q21 region. Since the establishment of cloned resources is an essential step towards the cloning of genes involved in inherited diseases, we built a yeast artificial chromosome (YAC) contig and an STS map of this part of the X chromosome. The contig, which extends from PGK1 in Xq13.3 to DXS1002 in Xq21.2, consists of 30 YACs mapped with 21 markers and spans about 6 Mb. The YAC contig was used as a framework to localise several previously known genes and CEPH/Genetion polymorphic markers, as well as to construct a physical map of the region surrounding one of these genes. We recently localised a presumed MR locus to the region flanked by DXS233 (proximal) and CHM (distal). In the present work, the zinc finger gene, ZNF6, has been shown to lie within this region and to be highly expressed in brain, making it a good candidate MR gene. Similarly the VDAC1 gene has been mapped between DXS986 and DXS72 and its candidate gene status for the Allan-Herndon-Dudley syndrome is discussed.

(J Med Genet 1996;33:353–357)

Key words: Xq21; mental retardation; YAC contig.

The dark staining band Xq21 is assumed to be a region in which there are few genes. However, there is strong evidence that this region may contain genes involved in brain function. For example, chromosomal rearrangements such as deletions, microduplications, or inversions of part of this chromosomal region have been reported as being associated with mental retardation. Invariably, in males with interstitial deletions of the Xq21 band, the mental retardation is associated with some other entity, usually choroideremia (CHM), but sometimes both CHM and deafness (DFN3). Other strong evidence for the presence of a gene or genes in Xq21 involved in mental retardation conditions comes from the fact that some X linked mental retardation (XLMR) syndromes, such as the Miles-Carpenter, Goldblatt, Cowchock-Fishbeck, or Allan-Herndon-Dudley syndromes, as well as several non-syndromic MR entities such as MRX7, MRX8, or MRX31, exhibit maximal lod scores with markers in this region. Characterisation of the deletion of Xq21 patients with overlapping phenotypes has recently allowed us to localise a gene involved in MR within the DXS233/CHM interval. Using newly isolated cDNA clones to screen for mutations in patients with some forms of XLMR linked to the 2-2 Mb Xq13.3 region between DXS441 and PGK1, we have identified the XNP (X linked nuclear protein) gene as being responsible for the ATR-X syndrome. Based on the success of these combined analyses, we decided to conduct physical mapping studies on the Xq21 region. Previous analyses have focused mainly on the ordering of probes using a panel of deletion or translocation patients. A report of PFGE experiments, conducted with some probes from the region, has been published. However, no cloned resources, encompassing at least part of the interval between PGK1 and DXS1002, have been described so far, neither has a complete map including physical distances. We therefore decided to cover this region with YACs and to link this new contig to our already existing YAC contig around PGK1. This was accomplished using an STS based approach to isolate MegaYACs from the CEPH library and resulted in the generation of a 6 Mb YAC contig.

This YAC contig was then used to localise precisely several previously known genes and to build a long range restriction map of the DXS995–DXS1002 region. Our results show that the MR critical region between DXS233 and pJ7.6A is no larger than 1.2 Mb and that the zinc finger gene ZNF6 lies within this region. We also identified, within the same MR critical region, a new potential CpG island which might be indicative of at least a second transcription unit within this MR critical region. Moreover, the resources we have produced allowed us to localise another previously known gene, the VDAC1 gene, between DXS986 and DXS72. This result, in combination with the potential function of the VDAC1 protein, makes the VDAC1 gene a good candidate for being involved in the Allan-Herndon-Dudley syndrome.
Table 1  Previously unpublished STSs used to construct the contig

<table>
<thead>
<tr>
<th>Name</th>
<th>Probe</th>
<th>Primer sequences</th>
<th>Product size (bp)</th>
<th>Temp (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS71</td>
<td>pHX79</td>
<td>AATTTGTTGCAAGGCACCTGT</td>
<td>129</td>
<td>55</td>
</tr>
<tr>
<td>DXS447</td>
<td>pX65H7</td>
<td>TCCAAGTTAAGTGGGTCGTC</td>
<td>226</td>
<td>57</td>
</tr>
<tr>
<td>VDAC1</td>
<td>cDNA</td>
<td>TGAGCAAGAAGTGGAAAGCG</td>
<td>208</td>
<td>60</td>
</tr>
<tr>
<td>DXS72</td>
<td>pX65H7</td>
<td>TCTTTGACATTCAGAAAGAC</td>
<td>221</td>
<td>55</td>
</tr>
<tr>
<td>ZNF6</td>
<td>5' cDNA</td>
<td>GCTTTGACCTCGAGCTGCG</td>
<td>228</td>
<td>56</td>
</tr>
<tr>
<td>CHM</td>
<td>cDNA</td>
<td>GACACAAAAGTGGACCCAGG</td>
<td>244</td>
<td>56</td>
</tr>
</tbody>
</table>

The primer sequences are given in a 5' to 3' orientation. The primers derive from the previously published probe, the name of the corresponding probe is indicated. cDNA: primer design according to the cDNA content of the corresponding gene.

Methods

YAC LIBRARY SCREENING

Pools of the CEPH MegaYAC library were screened by PCR for the presence of DXS71, DXS447, and DXS72 using the primers and the annealing conditions shown in table 1. Screening for locus DXS326 was carried out using the primers described in Barker and Fain. The PCR conditions were 96°C for 30 seconds, one minute at the annealing temperature, and 72°C for 40 seconds for 40 cycles.

LONG RANGE RESTRICTION MAPPING

YAC DNA embedded in agarose was digested to completion with either SfiI, NotI, MlaI, or EagI using the buffer supplied by the manufacturer (New England Biolabs). The fragments were then electrophoresed on a 1% agarose (Gibco BRL) gel in a CHEF DRII PFGE apparatus (Biorad) using switch times of 30 to 120 seconds at 200 V for 40 from 4 seconds. The DNA was then transferred to nylon Biodyne membranes (Pall Industries).

PROBE PREPARATION AND HYBRIDISATION CONDITIONS

Probes were labelled using the random priming procedure as previously described. Southern blot hybridisations were carried out in a standard SSC buffer (5 x SSC, 5 x Denhardt's, and 0.1% SDS) for 16 to 18 hours at 65°C in a rotary hybridisation oven. Washes were performed as follows: 20 minutes in 2 x SSC/0.1% SDS, 20 minutes in 0.5 x SSC/0.1% SDS, and five to 20 minutes in 0.1 x SSC/0.1% SDS according to the stringency of the washes. A Northern blot filter was purchased from Clontech and hybridised according to the manufacturer's recommendation. All membranes were exposed to X-OMAT AR film (KODAK) from one hour to five days.

Results

ISOLATION OF YAC CLONES AND CONTIG CONSTRUCTION

As a preliminary step towards the construction of a YAC contig encompassing the Xq21 re-
region, the CEPH MegaYACs library was screened for the loci DXS571, DXS447, DXS72, and DXS326. Three DXS571 positive YACs (813b10, 829b12, 921g10), two DXS447 positive YACs (854e7, 967e12), three DXS72 positive YACs (767e4, 821a10, 321f8), and three DXS326 positive YACs (909g7, 922b5, 949e11) were characterised at this stage. The CEPH-Genethon database was then used to identify YACs that either mapped to the region on the basis of their STS contents (six YACs positive for DXS995 (632b10, 659h3, 666a6, 668e6, 725b3, and 894e7), one additional YAC positive for DXS571 (657e12), one YAC positive for DXS1225 (910h1), and two YACs positive for DXS986 (946e8, 927b5)) or appeared to contig with YACs that had already been placed in the contig (951f11, 792a5, 948a2 that were in contig with 910h1 and 749h4, 955a12, 904f2, 838c5 that were in contig with 821a10). In total, 28 CEPH MegaYACs were further studied with 21 markers including five genes.

The individual STS or probe content of each YAC was then determined by the combined use of PCR and hybridisation. In addition to the previously published original set of STSs from this region, several additional ones were developed. They are listed in table 1 and the STS data are summarised in table 2. This led to the assembly of two contigs spanning from PGK1 to DXS26 and from DXS995 to DXS1002. Since an 850 kb YAC and cosmId contig encompassing the DFN3 locus and linking the two loci DXS26 and DXS995 had previously been described, two additional YACs from the ICRF library, ICRFy901e1023 and ICRFy900B1210 respectively, were also studied and included in the final unique contig.

RESTRICTION MAPPING OF THE MR CRITICAL REGION AND ESTIMATION OF THE CONTIG SIZE

As described in the introduction, we previously localised the proximal and distal borders of the candidate region for one MR gene in Xq21 between DXS233 and CHM respectively. We therefore decided to construct a restriction map of this region to order closely spaced markers and estimate the true distances between loci. Mapping was conducted using the four rare cutter endonucleases SfiI, MluI, NotI, and EagI and the YACs 979a10, 909e7, 922b5, and 949e11. The map was assembled using probes derived from the left and right ends of the pYAC4 vector, as well as internal probes corresponding to STSs or genes. Eventually, membranes were hybridised with total human DNA to ensure that all restriction fragments had been detected. The result is shown in fig 1. Using this map, it was possible to order a number of markers from the region. From the centromere to the telomere, the following order was found: DXS232-DXS121-DXS233-DXS355 (DXS326, DXS1167, ZNF6)-CHM-DXS1002-DXS95. This is in agreement with data previously obtained by other groups.

Analysis of the rare cutter restriction map of the region showed at least one cluster of sites containing a NotI site, which could correspond to a CpG island, approximately 400 kb distal to DXS233, between DXS355 and DXS326. Previous studies have shown that the human ZNF6 gene is localised between DXS233 and CHM. A probe specific for the 5' end of the cDNA, hybridised to a 300 kb SfiI fragment, a 400 kb MluI fragment, and a 200 kb EagI fragment, positioning that end of the gene approximately 800 kb distal to DXS233. A probe corresponding to the 3' end
of the gene hybridised to a 200 kb SfiI fragment and a 350 kb EagI fragment. According to the map (fig 1), this result indicates that the 5' specific probe hybridises to a fragment distal to the one hybridised by the 3' end, thereby establishing that ZNF6 is transcribed telomere to centromere and is localised between DXS355 and DXS326.

Recent studies have mapped the human gene encoding the mitochondrial voltage dependent anion channel gene, VDAC1, into the Xq13-q21 interval. We therefore decided to localise this gene precisely on the YAC contig by PCR amplification to associated STSs. The results show that it lies between DXS986 and DXS72. Similarly to what we did for the ZNF6 gene, the use of specific 5' end or 3' end specific probes and preliminary data on the physical map of the region surrounding the VDAC1 gene allowed us to determine that it is orientated from the telomere to the centromere.

Finally, to approximate the size of the contig we combined the chimerism/size data determined either by ourselves or through the CEPH-Genethon database. The three non-overlapping, non-chimeric YACs, 946e8, 813a2, and 949e11, give a minimum size of approximately 3.8 Mb between DXS447 and DXS1002. According to our long range restriction maps, the minimum distances between DXS995-DXS355 and between VDAC1-DXS169 are about 1 Mb and 400 kb respectively. The size of the most proximal YACs (326b10, 910h1, 657e12) and their chimerism status according to the CEPH databases suggest that the distance between PGK1 and DXS447 should be about 600 kb. Finally, a previous publication reported the distance between DXS26 and DXS995 to be 250 kb. Therefore, considering all these data, we estimate the size of our contig to be about 6 Mb from PGK1 to DXS1002.

**Expression Studies of the ZNF6 Gene**

Since the ZNF6 gene, which maps between DXS233 and CHM, encodes for a putative zinc finger transcriptional factor, its position and putative function led us to speculate that a deficiency in the expression of such a gene could be involved in a form of mental retardation. In order to study its candidate gene status, we performed a northern blot experiment, using RNAs extracted from a variety of fetal tissues (fig 2). A 5.5 kb transcript was identified and showed that the ZNF6 gene is highly expressed in fetal brain with lower expression in heart, kidney, and lung and very low expression in liver. In addition to this 5.5 kb transcript, a larger one, about 7.5 kb, was also identified in kidney and lung. Whether this larger fragment corresponds to an alternatively spliced transcript remains to be determined; however, our finding that the ZNF6 gene is highly expressed in brain makes it a good candidate for being involved in the MR syndrome.

**Discussion**

The present work describes the generation of a 6 Mb YAC contig spanning the genomic region between PGK1 in Xq13.3 and DXS1002 in Xq21.2. When we started this work, a small YAC contig linking DXS26 and DXS995 had been described, and pulse field electrophoresis analysis of genomic DNA had shown the close proximity of loci DXS232, DXS121, and DXS233. However, despite these data, the region running from DXS447 to DXS1002 was split into many contigs and several probes or STSs, such as DXS326, DXS355, or ZNF6, were not included in any contig. The contig we have built contains 30 YACs and 21 STSs have been ordered within this contig. Because the STS order we deduced from our YAC contig is consistent with the ordering of probes resulting from the use of a panel of deletion or translocation patients, we think that these YACs are representative of the corresponding genomic region and should provide useful cloned resources for further analysis.

Since our previous results concerning the molecular analysis of males with mental retardation and deletions of Xq21 placed the putative MR region between DXS233 and CHM, we built a long range restriction map of the DXS232 to DXS1002 interval. It shows...
Localisation of two candidate genes for mental retardation using a YAC physical map of the Xq21.1–2.2 subbands


