A highly polymorphic dinucleotide repeat on the proximal short arm of the human X chromosome: linkage mapping of the synapsin I/A-raf-1 genes

Cordula U. Kirchgessner
University of Massachusetts Medical School

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/oapubs

Part of the Cell Biology Commons, and the Neuroscience and Neurobiology Commons

Repository Citation

This material is brought to you by eScholarship@UMassChan. It has been accepted for inclusion in Open Access Publications by UMass Chan Authors by an authorized administrator of eScholarship@UMassChan. For more information, please contact Lisa.Palmer@umassmed.edu.
A Highly Polymorphic Dinucleotide Repeat on the Proximal Short Arm of the Human X Chromosome: Linkage Mapping of the Synapsin I/A-raf-1 Genes

Cordula U. Kirchgessner,* James A. Trofatter,† Melanie M. Mahtani,‡ Huntington F. Willard,‡ and Louis J. DeGennaro*

*Departments of Neurology and Cell Biology, University of Massachusetts Medical School, Worcester; †Neurogenetics Laboratory, Massachusetts General Hospital, Charlestown; and ‡Department of Genetics, Stanford University Medical Center, Stanford, CA

Summary
A compound (AC)n repeat located 1,000 bp downstream from the human synapsin I gene and within the last intron of the A-raf-1 gene has been identified. DNA data-base comparisons of the sequences surrounding the repeat indicate that the synapsin I gene and the A-raf-1 gene lie immediately adjacent to each other, in opposite orientation. PCR amplification of this synapsin I/A-raf-1-associated repeat by using total genomic DNA from members of the 40 reference pedigree families of the Centre d'Etude du Polymorphisme Humaine showed it to be highly polymorphic, with a PIC value of .84 and a minimum of eight alleles. Because the synapsin I gene has been mapped previously to the short arm of the human X chromosome at Xp11.2, linkage analysis was performed with markers on the proximal short arm of the X chromosome. The most likely gene order is DXS7357/ARAFT1MPDXS7255DXS146, with a relative probability of $5 \times 10^3$ as compared with the next most likely order. This highly informative repeat should serve as a valuable marker for disease loci mapped to the Xp11 region.

Introduction
Synapsin I is the best-characterized member of a family of neuron-specific phosphoproteins (reviewed in Bähler et al. 1990). It is found in neurons, regardless of transmitter type, in both the central and peripheral nervous system and is concentrated at the presynaptic terminal. Evidence suggests that it plays a fundamental role in the regulation of neurotransmitter release.

Using both in situ hybridization to somatic cell hybrids and Southern blot analysis, we have previously mapped the location of the synapsin I gene to the X chromosome of man (at band Xp11) and of the mouse (at XA1-A4) (Yang-Feng et al. 1986). This localization suggested the possibility that X-linked neurological disorders might have as their cause a defect in the synapsin I gene. One prerequisite for studying this prospect is the ability to perform linkage mapping with the human synapsin I gene locus. To this end we have explored the use of highly repetitive DNA sequences as polymorphic markers for linkage analysis. Others have previously shown that dinucleotide-repeats within genes are often highly polymorphic. Through the use of PCR technology these polymorphisms can be used to assess allelic variation and inheritance in a variety of genes (Litt and Luty 1989; Tautz 1989; Weber and May 1989). Such dinucleotide repeats are therefore excellent markers for linkage mapping.

We describe here the identification of a highly informative (AC)n dinucleotide repeat located near the 3' end of the human synapsin I gene. DNA sequence analysis of the region around this repeat revealed that it actually lies within the last intron of the human A-raf-1 gene, known to encode the cellular homologue of a viral oncogene and to be located at Xp11 (Huebner et al. 1986; Beck et al. 1987), and suggests that these two genes lie back-to-back on the X chromo-
some. Using this polymorphic repeat, which we define as the synapsin I/A-raf-1 repeat (Syn1/Araf), we have carried out detailed linkage mapping with other markers on the proximal short arm of the human X chromosome.

Material and Methods

Identification of (AC), Repeats

Previously isolated human genomic clones for synapsin I (C. U. Kirchgesner and L. J. DeGennaro, unpublished data) were digested with various restriction enzymes. Resulting DNA fragments were separated by agarose-gel electrophoresis and transferred to a Zeta probe filter membrane according to established procedures (Sambrook et al. 1989). The filter was prehybridized for 4 h at 55°C in 1% BSA, 1 mM EDTA, 0.5 M Na₂HPO₄ pH 7.2, and 7% SDS and was then hybridized overnight in the same solution containing a radiolabeled probe produced by nick-translation of a synthetic poly(dA-dC)-poly(dG-dT) oligomer (Pharmacia). After hybridization, the filter was washed at 55°C in 1 mM EDTA, 40 mM Na₂HPO₄ pH 7.2, and 1% SDS for 4 × 20 min and was blotted dry and placed on film for autoradiography. Hybridizing DNA fragments were identified, subcloned into M13mp18 sequencing vectors (Messing et al. 1981), and sequenced by the dideoxy-chain termination method (Sanger et al. 1977).

PCR Amplification

Ten nanograms of total genomic DNA were initially denatured at 95°C for 3 min. Amplification of the DNA samples was then performed according to the conditions described in the Gene Amp Kit (Perkin Elmer-Cetus, Norwalk, CT), with the following modifications: gelatin was added to a final concentration of 0.01%, 1 μCi of α-(³²P)dCTP (800 Ci/mmol) was included in each reaction, and the concentration of unlabeled dCTP was reduced to 20 μM. Reactions contained 10 pmol each of two 19-mer oligonucleotide primers (see fig. 1), in a reaction volume of 20–25 μl. Thirty cycles of amplification were performed automatically with an MJ Programmable Thermal Controller (MJ Research, Cambridge, MA). Denaturation was for 1 min at 94°C, annealing was for 2 min at 55°C, and primer extension was for 2 min at 72°C. A final extension step of 7 min at 72°C was also performed. Aliquots of the amplified products were mixed with 2 vol formamide sample buffer, and 3–5 μl aliquots were applied to a 5% acrylamide gel containing 8 M urea. After electrophoresis the gels were fixed, dried, and autoradiographed for several hours. Amplified product sizes were determined relative to a DNA sequencing ladder.

Linkage Analysis

Linkage analysis was performed using the programs LINKMAP and ILINK of the LINKAGE program package (version 4.5) (Lathrop et al. 1984) on a VAX8700. Order and distances for the established markers used in the analysis were obtained from published reports (Drayna and White 1985; Goodfellow et al. 1985; Mahtani and Willard 1988; Mahtani et al. 1989). Linkage analysis in the reference pedigrees of the Centre d’Etude du Polymorphisme Humain (CEPH), Paris, were performed with version 3 of the CEPH data base as well as with an updated version of the data base containing the additional markers SYP and OATL1.

Results

Identification of (AC), Repeats

A Southern blot containing restriction endonuclease-cleaved DNA fragments of human genomic clones encoding a portion of the synapsin I gene were hybridized with a radiolabeled poly(dA-dC)-poly(dG-dT) oligomer. A 1,400-bp BamHI fragment gave a positive hybridization signal (data not shown) and was subsequently cloned into M13mp18 and sequenced. Figure

Figure 1  Sequence of compound (AC), repeat of SYN/Araf1 locus. Individual repeats are boxed. Sequences of the two 19-base primers used in the PCR amplification are indicated by arrows.
1 shows a portion of the DNA sequence of this fragment. Inspection of the sequence revealed a compound dinucleotide repeat as defined by Weber (1990). This compound repeat contains two perfect repeats, with lengths of 18 and 12 (AC) dinucleotides, and an imperfect repeat 9 dinucleotides in length. This repeat is clearly distinct from another dinucleotide repeat, DXS426, located at Xp11 (Coleman et al. 1990; Luty et al. 1990). Also shown in figure 1 are the sequence and position of the two oligonucleotide primers used for PCR amplification of the repeat region from total genomic DNA.

**PIC of the Synl/Araf Repeat Revealed by PCR**

Figure 2 shows the results of polymerase chain reaction amplification of the Synl/Araf repeat region in total DNA from members of a CEPH reference pedigree (family 1418) and from the original synapsin I genomic clone. Amplification of the Synl/Araf repeat in the original phage lambda genomic clone resulted in the synthesis of the expected product of 202 nucleotides in length (fig. 2, lanes M). We defined this product as allele zero (0). Amplification of the repeat sequence in the DNA of CEPH family members revealed that the sequence is highly polymorphic and informative. Table 1 lists the most frequent alleles so far observed and their respective frequencies in the 40 pairs of parents in the CEPH reference pedigrees. An eighth allele of 190 bases in length has been observed at a very low frequency, and we anticipate that alleles of 198 and 200 bases in length may also exist at a very rare frequency in the population. The PIC value calculated for the two perfect repeats taken together, according to the formula of Weber (1990), is .84.

**Linkage Analysis**

Linkage analysis was performed with the CEPH reference pedigrees. Thirty-three of the 40 families proved to be informative for the Synl/Araf repeat and were analyzed further. Detailed linkage analysis was performed with six loci located on the proximal short arm of the X chromosome (listed in table 2). Table 3 gives the lod scores (Z) at varying recombination fractions (θ), as well as the maximum (Z) between the Synl/Araf repeat and the six chosen makers in a two-point linkage analysis. There is at least one recombination event between each of these markers and SYN/ARAF1.

In an initial multipoint analysis we used only four markers (DXS7, TIMP, DXS255, and DXS146) for which order and distances have been well established (Drayna and White 1985; Goodfellow et al. 1985; Mahtani and Willard 1988; Mahtani et al. 1989). A graphic representation of the results of the multipoint linkage analysis is given in figure 3. A gene order that places the synapsin I and A-raf-1 genes between DXS7 and TIMP results in the highest relative probability (5 × 10⁸). The likelihood for alternative gene orders were also examined and are listed in table 4. To evaluate intergene distances maximization of the likelihoods was performed with the ILINK program and is as follows: DXS7—9 cM—SYN/ARAF1—4.2 cM—TIMP—2.8 cM—DXS255—3.4 cM—DXS146. Our data are consistent with physical mapping data that place the A-raf-1 gene distal to DXS146 and DXS255 (Lafreniere et al. 1989).

Once the Synl/Araf repeat was placed on the map

| Table 1 |
|-----------------|-----------------|
| **Synl/Araf Repeat-Allele Sizes and Frequencies in 80 Parents of CEPH Pedigrees** |
| **Size** | **Frequency** | **Nomenclature** |
| 202 | 23.3% | 0 |
| 204 | 38.3% | +2 |
| 206 | 19.2% | +4 |
| 208 | 2.5% | +6 |
| 210 | 2.5% | +8 |
| 196 | 2.5% | −6 |
| 194 | 11.7% | −8 |
Table 2

X-Chromosome Markers Used for Linkage Analysis

<table>
<thead>
<tr>
<th>Locus</th>
<th>Probe</th>
<th>Region</th>
<th>Enzyme</th>
<th>Alleles (kb)</th>
<th>Frequency</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS7 .....</td>
<td>L1.28</td>
<td>Xp11.3</td>
<td>TaqI</td>
<td>12.0</td>
<td>.68</td>
<td>Wieacker et al. 1984</td>
</tr>
<tr>
<td>TIMP ......</td>
<td>TIMP</td>
<td>Xp11</td>
<td>BglII</td>
<td>9.0</td>
<td>.32</td>
<td>Durfy et al. 1986</td>
</tr>
<tr>
<td>DXS255 .....</td>
<td>M27β</td>
<td>Xp11.3-cen</td>
<td>EcoRI and PstI</td>
<td>Multiallelic hypervariable system</td>
<td>Fraser et al. 1987</td>
<td></td>
</tr>
<tr>
<td>DXS146 .....</td>
<td>pTAK8</td>
<td>Xp11.3-Xp11.1</td>
<td>XbaI</td>
<td>3.3</td>
<td>.61</td>
<td>Kruse et al. 1986</td>
</tr>
<tr>
<td>SYP .........</td>
<td>498</td>
<td>Xp11</td>
<td>Eco109</td>
<td>0.8</td>
<td>.38</td>
<td>Ozcelik et al. 1990</td>
</tr>
<tr>
<td>OATL1 ......</td>
<td>HuOAT</td>
<td>Xp21.1-Xp11.2</td>
<td>Scal</td>
<td>4.3</td>
<td>.53</td>
<td>Lafreniere et al. 1989</td>
</tr>
</tbody>
</table>

Table 3

Two-Point Linkage Analysis of SYN/ARAF1

<table>
<thead>
<tr>
<th>Marker Locus</th>
<th>No. of Informative Families</th>
<th>Z at Θ of</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>Θ</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS7 .......</td>
<td>10</td>
<td>8.27</td>
<td>11.91</td>
<td>12.37</td>
<td>11.87</td>
<td>10.9</td>
<td>8.09</td>
<td>4.39</td>
</tr>
<tr>
<td>TIMP .......</td>
<td>22</td>
<td>35.08</td>
<td>35.82</td>
<td>33.71</td>
<td>30.77</td>
<td>27.33</td>
<td>19.32</td>
<td>9.88</td>
</tr>
<tr>
<td>DXS255 ...</td>
<td>28</td>
<td>29.44</td>
<td>37.02</td>
<td>37.12</td>
<td>35.5</td>
<td>32.4</td>
<td>22.92</td>
<td>17.64</td>
</tr>
<tr>
<td>DXS146 ...</td>
<td>11</td>
<td>4.13</td>
<td>9.74</td>
<td>10.95</td>
<td>10.8</td>
<td>10.03</td>
<td>7.36</td>
<td>3.7</td>
</tr>
<tr>
<td>SYP .........</td>
<td>16</td>
<td>24.8</td>
<td>23.85</td>
<td>21.97</td>
<td>20.29</td>
<td>17.52</td>
<td>12.29</td>
<td>6.29</td>
</tr>
<tr>
<td>OATL1 .......</td>
<td>22</td>
<td>22.0</td>
<td>27.66</td>
<td>27.71</td>
<td>26.05</td>
<td>23.55</td>
<td>16.9</td>
<td>8.57</td>
</tr>
</tbody>
</table>

with respect to these markers, we carried out further analyses to include the remaining two loci, synapto-physin (SYP) and ornithine aminotransferase (OATL1), for which the available data were limited. Previous linkage analysis had placed SYP near TIMP, with physical data placing it just proximal to the TIMP locus (Ozcelik et al. 1990). The results of our linkage analysis are given in table 5. They confirm and more firmly support the previously established location, with a 66-fold higher likelihood compared with the next most likely order. Results of a multipoint linkage analysis for the OATL1 locus (SacI polymorphism) remained inconclusive, placing this locus proximal to SYN/ARAF1 and distal to DXS255, but a finer mapping was not possible with the available data. Figure 4 depicts the short arm of the X chromosome, with the newly established order of markers in the Xp11 region.

Discussion

We report the identification of a highly polymorphic marker located on the proximal short arm of the human X chromosome, in association with both the sy-
Table 4
Location of Synapsin I Gene and A-raf-1 Gene, Relative to Test Marker Loci

<table>
<thead>
<tr>
<th>Order</th>
<th>Relative Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS7-[SYN,ARAF1]-TIMP-DXS255-DXS146</td>
<td>4.68 x 10^21</td>
</tr>
<tr>
<td>DXS7-TIMP-DXS255-DXS146-[SYN,ARAF1]</td>
<td>1.00 x 10^14</td>
</tr>
<tr>
<td>[SYN,ARAF1]-DXS7-TIMP-DXS255-DXS146</td>
<td>2.95 x 10^13</td>
</tr>
<tr>
<td>DXS7-TIMP-[SYN,ARAF1]-DXS255-DXS146</td>
<td>1.15 x 10^9</td>
</tr>
<tr>
<td>DXS7-TIMP-DXS255-[SYN,ARAF1]-DXS146</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 5
Location of Synaptophysin Gene, Relative to Test Marker Loci

<table>
<thead>
<tr>
<th>Order</th>
<th>Relative Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS7-[SYN,ARAF1]-TIMP-SYP-DXS255-DXS146</td>
<td>3.8 x 10^4</td>
</tr>
<tr>
<td>DXS7-SYP-[SYN,ARAF1]-TIMP-DXS255-DXS146</td>
<td>5.8 x 10^3</td>
</tr>
<tr>
<td>DXS7-[SYN,ARAF1]-SYP-TIMP-DXS255-DXS146</td>
<td>11.17</td>
</tr>
<tr>
<td>DXS7-[SYN,ARAF1]-TIMP-DXS255-SYP-DXS146</td>
<td>1.26</td>
</tr>
<tr>
<td>SYP-DXS7-[SYN,ARAF1]-TIMP-DXS255-DXS146</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Synapsin I and A-raf-1 genes. This marker is defined by three (AC)_n repeats, which are themselves separated by a repetitive sequence. Because of this unusual sequence structure, we carried out PCR amplification of the entire region, rather than across the (AC)_n repeat regions individually. We were initially concerned that the PIC of the multiple repeat might be reduced if an increase in the number of dinucleotides in one of the repeat stretches was canceled by an equal reduction in one of the other stretches. However, the observed .825 heterozygosity among females conforms well with the calculated PIC value of .84 for a single (AC) stretch of 30 repeats (Weber 1990) and leads us to believe that information is rarely lost in this system.

Additional DNA sequencing and exon-intron border mapping of our human synapsin I gene clones revealed that this (AC)_n repeat lies about 1,000 bp 3' of the synapsin I gene polyadenylation signal. Further, a comparison of our sequence with sequences in the Genbank Database revealed that this repeat lies within the last intron of the gene encoding the cellular homologue of a viral oncogene, A-raf-1 (Huebner et al. 1986; Beck et al. 1987). This result indicates that the synapsin I gene and the A-raf-1 gene lie immediately adjacent to each other on the human X chromosome, 3' and to 3' end, and are transcribed from opposite strands (details of this localization will be published elsewhere). The functional importance of this association remains obscure; however, we have observed a similar juxtaposition of these two genes in the rat genome, which suggests that this arrangement has been conserved in evolution. In eukaryotes no clear functional relationship has been described for other structurally unrelated but overlapping genes transcribed.

Figure 4 Linkage map of proximal short arm of X chromosome. Order of loci is shown in the center; genetic distances, expressed in centiMorgans, are listed to the right and are based on 8 estimates made by ILINK analysis.
from opposing strands (Spencer et al. 1986; Williams and Fried 1986; Adelman et al. 1987).

The fortuitous location of the Synl/Araf repeat has allowed us to establish the order of these loci in relation to other markers on the proximal short arm of the X chromosome. We analyzed our data for both pairwise and multipoint linkage. There are several recombination events between the Synl/Araf repeat and most of the other markers studied, making this analysis very informative. The exception is SYP, where we observed only a single recombination event. Given only these data, a two-point linkage analysis places SYP within 1 cM of SYN/ARAF1. However, we noted that the same recombination event that occurred between SYN/ARAF1 and SYP also occurred between TIMP and SYP (CEPH family 13293, individual 6). This information, taken together with our multipoint linkage analysis, places the synaptophysin gene 2 cM proximal to TIMP and therefore approximately 6 cM from the SYN/ARAF1 locus. The OATL1 sequence (Scal RFLP) has been reported to map distal to DXS255 (Lafreniere et al. 1989), which we confirm here; and we also place this sequence proximal to the Synl/Araf repeat.

Being able, with high confidence, to place the synapsin I gene and the Araf-1 gene on the X chromosome linkage map allows us to analyze the potential role of these genes in X-linked diseases and to use this repeat to study linkage with disease loci which are known to occur on the proximal short arm of the X chromosome. We have previously speculated about the possible involvement of the synapsin I gene in the X-linked disease Rett syndrome (DeGennaro et al. 1987). Rett syndrome is a degenerative neurological disorder affecting both the central and peripheral nervous systems and occurring exclusively in females (Rett 1966; Hagberg et al. 1983). It is believed to be an X-linked dominant, male-lethal gene defect which occurs spontaneously in most cases, although a few familial cases with inheritance through the maternal line are known (Zoghbi 1988). It is therefore of particular interest to study the inheritance of the different alleles of the Synl/Araf repeat in these family cases, in order to look for concordance or discordance in inheritance with the disease. We have undertaken such a study and will report in detail elsewhere that, in at least one of the previously identified families, we have been able to exclude synapsin I and Araf-1 as candidate genes for Rett syndrome.

A number of disease loci have been mapped to the Xp11 region. These include properdin deficiency, Norrie disease, congenital stationary night blindness (CSNB1), X-linked retinitis pigmentosa (RP2), Wiskott-Aldrich syndrome, and incontinentia pigmenti. Properdin deficiency has been mapped distal to DXS7 (Goonewardena et al. 1988). Norrie disease, which is characterized by congenital blindness often accompanied by mental retardation, maps close to DXS7 (reviewed in Ngo et al. 1988). The Wiskott-Aldrich syndrome has been mapped to Xp11 and has been shown to be flanked by TIMP and DXS255 (Greer et al. 1990). These three diseases, therefore, are best analyzed by markers distal or proximal to the Synl/Araf repeat. However, in certain families, especially those that are uninformative for the DXS7 marker, the Synl/Araf repeat may prove useful as an alternative linkage marker.

Retinitis pigmentosa, a progressive degenerative disorder of the retina, is inherited in autosomal or X-linked fashion, with its severest manifestations occurring in X-linked cases. At least two loci for X-linked RP exist: RP3 maps to Xp21, and RP2 maps to Xp11 (reviewed in Humphries et al. 1990). Reported θ values between RP2 and DXS7 are 0–10. It would therefore fall in the vicinity of the Synl/Araf repeat. However, a recent multipoint study combining information obtained from 62 families (Ott et al. 1990) places RP2 close to the centromere, proximal to DXS14. Linkage analysis with additional markers between DXS7 and DXS14, such as the Synl/Araf repeat, may lead to a more precise localization of RP2. This approach has recently been used, for the dinucleotide marker DXS426, in two families with RP2 (Coleman et al. 1990).

Congenital stationary night blindness is an ocular disorder inherited in autosomal and X-linked modes. This disease is characterized by an impairment in neural transmission from photoreceptors to bipolar cells (Miyake et al. 1986). The X-linked form, CSNB1, has been linked to markers at Xp11 (Gal et al. 1989; Musarella et al. 1989; Bech-Hansen et al. 1990) and has also been placed distal to TIMP (Musarella et al. 1989). The Synl/Araf repeat appears to be close to this locus and should provide an excellent diagnostic marker. Its use in family linkage studies may also help to pinpoint the chromosomal location of the responsible gene(s). The linkage relationship between Synl/Araf and CSNB1 is currently under investigation.

In summary, the Synl/Araf repeat, localized in a region of the short arm of the human X chromosome.
in which informative markers are sparse, will provide a powerful tool for linkage mapping, prenatal diagnosis, and carrier detection.

Acknowledgments

We wish to thank Dr. Xandra Breakefield for suggesting the use of dinucleotide repeats and PCR amplification; Paul Mange for technical assistance; Drs. D. Howland, L. Hemmendinger, R. Franco, and J. Haines for valuable discussion; and Dr. David Drachman for his support. We gratefully acknowledge Dr. R. Lafreniere for providing genotypic data. This work was supported by National Institutes of Health Research Grant NS25050 (to L.J.D.), by a grant (to L.J.D.) from the Retinitis Pigmentosa Foundation Fighting Blindness, and by a fellowship (to C.U.K.) from the Boehringer Ingelheim Fonds.

References


Bech-Hansen NT, Field LL, Schramm AM, Reedyk M, Craig IW, Fraser NJ, Pearce WG (1990) A locus for X-linked congenital stationary night blindness is located on the proximal portion of the short arm of the X chromosome. Hum Genet 84:406–408


Musarella MA, Weleber RG, Murphey WH, Young RSL, Anson-Cartwright L, Mets M, Kraft SP, et al (1989) Assignment of the gene for complete X-linked congenital stationary night blindness (CSNB1) to Xp11.3. Genomics 5:727–737
Weber JL (1990) Informativeness of human (dC-dA)n (dG-dT)n polymorphisms. Genomics 7:524–530
Williams T, Fried M (1986) A mouse locus at which transcription from both DNA strands produces mRNAs complementary at their 3' ends. Nature 322:275–279