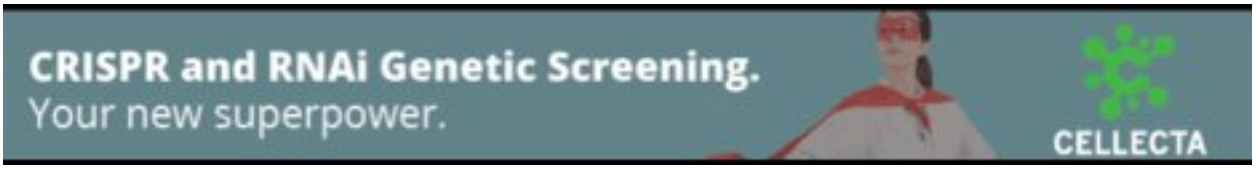
Long-range map of a 3.5-Mb region in Xp11.23-22 with a sequence-ready map from a 1.1-Mb gene-rich interval.

D Schindelhauer, H Hellebrand, L Grimm, et al.

Genome Res. 1996 6: 1056-1069
Access the most recent version at doi:10.1101/gr.6.11.1056



RESEARCH

Long-range Map of a 3.5-Mb Region in Xp11.23–22 with a Sequence-ready Map from a 1.1-Mb Gene-rich Interval

Dirk Schindelhauer,[1,4,5] Heide Hellebrand,[1,4] Lena Grimm,[1] Ingrid Bader,[1] Thomas Meitinger,[1] Manfred Wehnert,[2] Mark Ross,[3,5] and Alfons Meindl[1,6]

[1]Abteilung für Pädiatrische Genetik, Kinderpoliklinik der Universität München, 80336 München, Germany; [2]Institut für Humangenetik, Ernst Moritz-Arndt Universität, Greifswald, Germany; [3]Genome Analysis Laboratory, Imperial Cancer Research Fund, London, UK

Most of the yeast artificial chromosomes (YACs) isolated from the Xp11.23–22 region have shown instability and chimerism and are not a reliable resource for determining physical distances. We therefore constructed a long-range pulsed-field gel electrophoresis map that encompasses ~3.5 Mb of genomic DNA between the loci TIMP and DXS146 including a CpG-rich region around the WASP and TFE-3 gene loci. A combined YAC–cosmid contig was constructed along the genomic map and was used for fine-mapping of 15 polymorphic microsatellites and 30 expressed sequence tags (ESTs) or sequence transcribed sites (STSs), revealing the following order: tel–(SYN–TIMP)–(DXS426–ELK1)–ZNF(CA)$_n$–L1–DXS1367–ZNF81–ZNF21–DXS6616–(HB3–OATL1pseudogenes–DXS6950)–DXS6949–DXS6941–DXS7464E(MG61)–GW1E(EBP)–DXS7927E(MG81)–RBM–DXS722–DXS7467E(MG21)–DXS1011E–WASP–DXS6940–DXS7466E(MG44)–GF1–DXS226–DXS1126–DXS1240–HB1–DXS7469E–(DXS6665–DXS1470)–TFE3–DXS7468E–SYP–DXS1208–HB2E–DXS573–DXS1331–DXS6666–DXS1039–DXS1426–DXS1416–DXS7647–DXS8222–DXS6850–DXS255–CIC-5–DXS146–cen. A sequence-ready map was constructed for an 1100-kb gene-rich interval flanked by the markers HB3 and DXS1039, from which six novel ESTs/STSs were isolated, thus increasing the number of markers used in this interval to thirty. This precise ordering is a prerequisite for the construction of a transcription map of this region that contains numerous disease loci, including those for several forms of retinal degeneration and mental retardation. In addition, the map provides the base to delineate the corresponding syntenic region in the mouse, where the mutants *scurfy* and *tattered* are localized.

[The sequence data described in this paper have been submitted to GenBank under accession no. U66359.]

Multiple genes and disease loci have been mapped to the Xp11.23–22 region. A gene mutated in Wiskott-Aldrich syndrome (WAS) patients has been identified recently by positional cloning (Derry et al. 1994), whereas the genes for a number of disease loci including a type of X-linked congenital stationary night blindness (CSNB) (Bech-Hansen and Pearce 1993), retinitis pigmentosa (xlRP) (Meitinger et al. 1989; Ott et al. 1990) and Aland Island eye disease (Glass et al. 1993) remain to be isolated. In addition, at least four X-linked mental retardation loci (MRX1, MRX8, MRX12, MRX26; Lubs et al. 1996) and X-linked infantile spinal muscular atrophy (Kobayashi et al. 1995) have been linked to this region.

Several yeast artificial chromosome (YAC) contigs from the Xp11.23–22 region have been reported; however, none of these maps provides a full coverage. A YAC contig encompassing the distal part of the Xp11.23 region between the loci TIMP and OATL1 has been published (Coleman et al. 1994; Knight et al. 1994). A zinc-finger gene cluster has been detected on the distal OATL1 YAC (ICRFy900C0874; Knight et al. 1994), which also contains breakpoints for synovial sarcomas (De Leeuw et al. 1993; Chand et al. 1995) and the retinal-expressed sequences MG61, MG81, MG21, and MG44 on the proximal OATL1 YAC (ICRFy900F0501; Geraghty et al. 1993). Other

[4]These authors contributed equally to this work.
[5]Present addresses: D.S.: MRC Human Genetics Unit, Western General Hospital, Edinburgh, UK; M.R.: Sanger Centre, Hinxton Hall, Cambridge, UK.
[6]Corresponding author.
E-MAIL alfons@pedgen.med.uni-muenchen.de; FAX 0049-89-5160-4780.

6:1056–1069 ©1996 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/96 $5.00

YAC contigs encompassing the region between TIMP and the WASP gene locus also have been reported (Derry et al. 1994; Fisher et al. 1995; Kwan et al. 1995a). However, deleted or chimeric YACs have been used in these cases to connect the OATL1 pseudogene cluster with the GF-1 locus (Zon et al. 1990), which led to an incorrect positioning of the WASP locus (Derry et al. 1994; Meindl et al. 1995). Gaps also remained for the region between GF-1 and TFE-3 as well as between the TFE-3 and synaptophysin (SYP) loci (Derry et al. 1994). One YAC that has been reported to contain TFE-3 (Beckmann et al. 1990) and SYP (Özcelik et al. 1990) is chimeric and unstable (Hagemann et al. 1994; Fisher et al. 1995). A more proximally situated YAC contig for the region around DXS255 has been published, but could not be connected with the SYP locus (Fisher et al. 1995). Thus, the physical distances between the OATL1 pseudogene cluster and the GF-1 locus as well as the distances between the SYP locus and the VNTR locus DXS255 remained uncertain.

As most of the YACs characterized for this project also proved to be chimeric, deleted, unstable, or small, an integrative mapping approach was performed. A long-range map of the Xp11.23-22 region was established concomitantly with a YAC–cosmid contig for the entire region flanked by the loci TIMP and DXS146. The pulsed-field gel electrophoresis (PFGE) map revealed a CpG cluster in a 1100-kb region from which a sequence-ready map with a high density of expressed sequence tags (ESTs) and polymorphic markers was constructed.

RESULTS

Long-range Map of the Xp11.23–22 Region

Most of the probes obtained from YACs and other sources were first tested against the human/hamster radiation hybrid cell lines A19C5 and A10F7. X-chromosomal sequences retained in the first hybrid include DXS7 at one end and the DXS226 locus at the other, making it a useful mapping tool for Xp11.23 probes (Berger et al. 1992), whereas the second hybrid contains DXS255 and DXS146, making it a useful tool for Xp11.22 probes. Hybridization of genomic DNA with selected probes mapping between the loci TIMP and DXS146 (Table 1) revealed three *Mlu*I fragments. End-to-end localization of these fragments was shown by the restriction analysis of genomic DNA digested with other rare cutters as well as cosmids spanning the *Mlu*I restriction sites. The PFGE map encompasses a region of ~3.5 Mb. Long-range mapping results are summarized in Table 1 and depicted in Figure 1.

The distal *Mlu*I-fragment with a length of ~1450 kb starts at the TIMP locus and extends proximally to the DXS226 locus. The gene loci ZNF81, OATL1, and MG61 (DXS746) were mapped to the identical *Mlu*I fragment and *Not*I fragments, but to three different *Nru*I fragments. The most distal *Nru*I fragment contains the loci ZNF81 and TIMP, the intermediate (450 kb) contains the OATL1 pseudogene cluster including HB3 (Table 2), and the most proximal contains MG61. The loci DXS226, WASP, and HB-1 (Table 2) were localized to a single, partially digested *Not*I fragment (630 kb). After complete digestion, DXS226 and the WASP gene were mapped together with MG44 to the distal 280-kb *Not*I fragment (Fig. 2), whereas HB1 was mapped to a proximal 350-kb *Not*I fragment, which also contains the expressed sequences Xp664 (DXS7469E; Lee et al. 1995), TFE-3, and T54 (Fig. 2). The HB1 locus was placed distally to Xp664E and TFE-3, because it remains on the 1500-kb *Mlu*I fragment, whereas T54 maps together with TFE-3 and SYP to the second 380-kb *Mlu*I fragment (Fig. 2). The third and most proximal *Mlu*I fragment (~1500 kb) was identified by the novel single-copy probe HB2 (Table 2) that maps proximal to the *Mlu*I restriction site localized adjacent to the SYP gene. The identical *Mlu*I fragment also contains the loci DXS255 and DXS146 (Fig. 3a). The genomic clones HB2 and HB4 identify the same *Nru*I fragment (800 kb) as SYP but different and partially digested *Bss*HII fragments as well as a different *Nru/Mlu*I fragment (Fig. 3b; Table 1). An *Nru*I fragment of ~900 kb is shared by DXS255, ClC-5 (Fisher et al. 1995), and DXS146 (Fig. 3a). Although DXS255 like ClC-5 maps to a 350-kb *Bss*HII fragment, DXS146 hybridizes with a 400-kb *Bss*HII fragment (Table 1; Fig. 1).

Further Characterization of a YAC Contig Around a Zinc Finger Gene Cluster

YACs encompassing the region between the loci TIMP and OATL1 containing a zinc finger gene cluster were isolated from the Imperial Cancer Research Fund (ICRF) YAC library (Lehrach et al. 1990) and colinearity with the genomic map was tested. In contrast to two previous reports (Coleman et al. 1994; Hagemann et al. 1994), we found the YAC ICRFy900A1220 to be chimeric and the YAC ICRFy900A0120 to be deleted at its

Table 1. Fragment Sizes and Probes

Designation	*Not*I	*Bss*HII	*Nru*I	*Mlu*I	Marker type
TIMP	360	120	540	1450	Genomic probe
ELK1	360	420	540	1450	PCR product
L1	360	420	540	1450	cDNA probe
ZNF 81	720	420	540	1450	PCR product
OATL-1	720	450	450	1450	cDNA probe
HB3	720	450	450	1450	Genomic probe
MG61	720	<50	210	1450	PCR product
MG21	280*	<50	210	1450	Genomic clone
DXS1011	280*	<50	210	1450	PCR product
WASP	280*	<50	210	1450	cDNA fragment
MG44	280*	<50	95	1450	DNA fragment
GF-1	280*	110	95	1450	PCR product
DXS226	280*	110	80	1450	DNA fragment
HB1	350*	<50	80	1450	Alu-PCR product
Xp664	350*	<50	110	380	cDNA fragment
TFE-3	350*	100	110	380	cDNA fragment
T54	350*	<50	800	380	cDNA fragment
SYP	>3000	<50	800	380	cDNA fragment
HB2	>3000	600*	800	1500	Genomic fragment
HB4	>3000	600*	800	1500	Genomic fragment
DXS255	>3000	350	900	1500	Genomic fragment
CIC-5	>3000	350	900	1500	cDNA fragment
DXS146	>3000	400	900	1500	Genomic fragment

*Partially digested fragments. Double digests were performed with *Not*I/*Bss*HII, *Not*/*Nru*I, *Not*I/*Mlu*I, *Bss*HII/*Nru*I, *Bss*HI/*Mlu*I, *Nru*/*Mlu*I.

proximal end (see Methods). Two novel genes were isolated from this region. First, a clone designated L1 was isolated from a T-cell cDNA library by YAC hybridization (16H12) and positioned between the loci ELK1 (Rao et al. 1989) and DXS1367 (Schindelhauer et al. 1995). Sequence analysis revealed partial sequence identity with EST04093, which has been isolated from a brain cDNA library (Adams et al. 1993). Hybridization results indicate an additional autosomal gene related to L1 (data not shown). Second, a novel zinc finger gene [ZNF(CA)$_n$] was isolated from a cosmid (ICRFc104A0330) identified with YAC 16H12 (see Methods). The truncated zinc finger gene was shown to contain a polymorphic (CA)$_n$ repeat (Table 2, DXS9820). Fine mapping of the YACs and further cosmids isolated from that region localized ZNF(CA)$_n$ proximally and adjacent to the L1 locus (Fig. 1).

A Sequence-ready Map of a Gene-rich Interval in Xp11.23

The genomic region of ~1100 kb between the OATL1 pseudogene cluster (Chand et al. 1995) and the polymorphic repeat DXS1039 shows a high density of CpG islands on genomic DNA (Fig. 1). We concentrated our efforts on the construction of a cosmid contig for this region. Cosmids from the Lawrence Livermore (LL) X cosmid library distributed under the Reference Library System (Zehetner and Lehrach 1994) and from the ICRF cosmid library (Nizetic et al. 1991) were isolated with probes given in Table 1. Their order and map positions were determined by restriction mapping and hybridization. Gaps between individual cosmids were bridged using cosmid end sequences to isolate overlapping clones (Fig. 4). Only a single YAC located to this region was used for the hybridization of X chromosome-specific cosmid libraries. This small YAC (5H12, Lee et al. 1992) with a size of 120 kb was identified with the probe TFE-3. In all other cases, YACs were unsuited for this purpose owing to their chimerism or instability.

Nearly 20 ESTs or short tandem repeat (STR) markers have been located to this gene-rich interval (Nelson et al. 1995), and during this study,

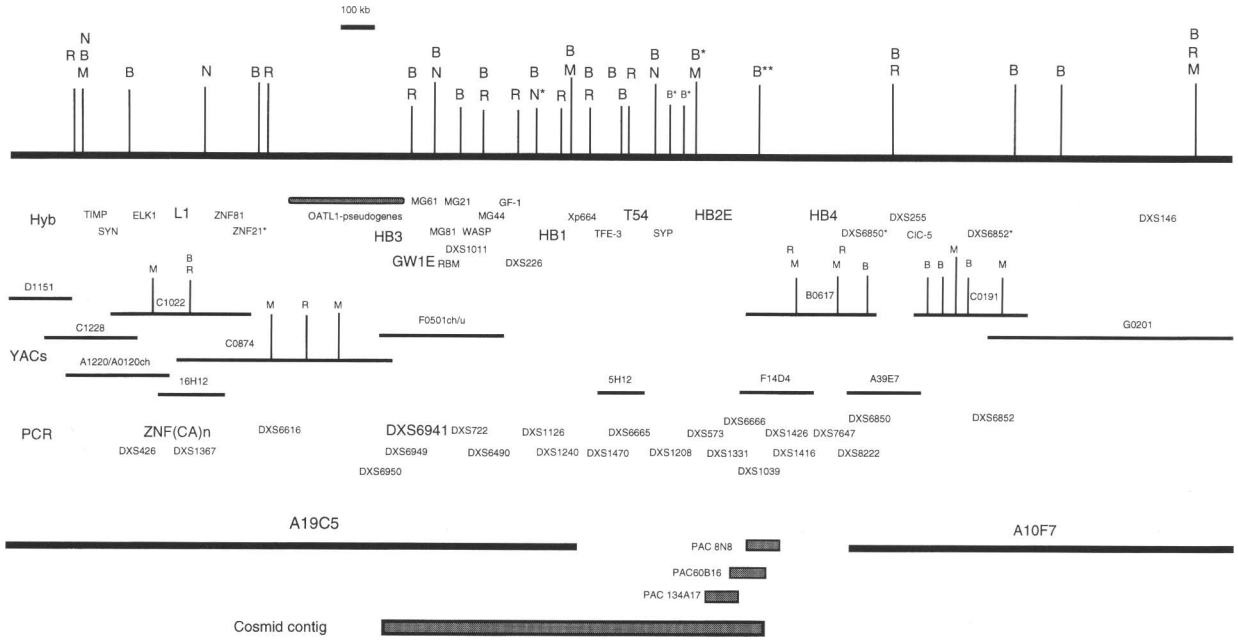


Figure 1 Long-range map of the Xp11.23–22 region. STSs or ESTs that were mapped by hybridization on genomic DNA and to YACs are shown above (probes with * only against YACs), markers mapped by a PCR-based approach on YACs or cosmids below. (R) *Nru*I; (B) *Bss*HII; (M) *Mlu*I; (N) *Not*I. YACs shown in the center line were obtained from either the ICRF library (Lehrach et al. 1990) or from an X chromosome-specific library (Lee et al. 1992). YACs yWXDF14D4/A39E7 were isolated from the St. Louis library (Nagaraja et al. 1994). Additional restriction sites on YACs not seen in genomic DNA are given. The 1100-kb sequence-ready cosmid/PAC contig is shown in detail in Fig. 4. YACs ICRFy900A0120/A1220 are chimeric (ch); YAC ICRFy900F051 is chimeric and unstable (u).

three additional ESTs, one polymorphic repeat (DXS6941), and two STSs were isolated (Table 2). A 1000-kb *Eco*RI map of cloned DNA was generated and this allowed the resolution and mapping of biologically relevant Xp11.23/22 markers, including 14 expressed genes and nine polymorphic markers (the former are marked by an asterisk, the latter are underlined): (HB3–OATL1–pseudogenes)–50 kb–DXS6949–15 kb–DXS6941–25 kb–MG61*(DXS7465E)–10 kb–GW1E*(EBP)–30 kb–MG81*(DXS7927E)–35 kb–RBM*–20 kb–DXS722–10 kb–MG21*(DXS7467E)–35 kb–DXS1011E*–35 kb–WASP*–25 kb–MG44*(DXS7466E)–5 kb–DXS6940–70 kb–GF1*–20 kb–DXS226–45 kb–DXS1126–25 kb–DXS1240–15 kb–HB1(DXS9821)–55 kb–Xp664*(DXS7469E)–30 kb–(DXS6665–DXS1470)–90 kb–TFE3*–75 kb–T54*(DXS7468E)–80 kb–SYP*–35 kb–DXS1208–45 kb–HB2E*(DXS9823)–15 kb–DXS573–30 kb–DXS1331–30 kb–DXS6666–40 kb–DXS1039.

A clone named GW1E was isolated from a placental cDNA library with the cosmid LLNLc110A0842 (Fig. 1) by reciprocal probing. Sequence analysis revealed complete identity with the EBP gene (GenBank accession no. Z37986), which was found by functional cloning as a receptor for the ligand emopamil (Hanner et al. 1995).

A novel expressed gene was cloned from the region by isolating transcripts from a T-cell cDNA library with the YAC 5H12. A complete cDNA sequence was obtained for the corresponding gene and named T54 (GenBank accession no. U66359; Fig. 5). This gene is expressed ubiquitously and shows no significant homology to other genes, although data-base searches revealed an identical EST (GenBank accession no. R37743) isolated from infant brain (Hillier et al. 1995).

Two published ESTs were mapped to the gene-rich interval. RBM has been isolated from a chimeric YAC that contains sequences of the Xp11.23 region (Derry et al. 1995a) and here was finely mapped between the loci GW1E (EBP) and MG81. The placental cDNA clone Xp664 (Lee et al. 1995) was finely mapped between the loci HB1 and DXS1470.

The cosmid contig was extended beyond the

Table 2. Novel ESTs and STSs from the Xp11.23 Region

$(CA)_n$ repeat inside an truncated zinc finger gene (DXS9820)	
ZNF-CA	5′-GTCCTGAGTGTGAGAAGGCCTTCA-3′
ZNF-GT	5′-GTTTAGCTGGAGAAAGAGTAATCTGG-3′
Primer sequences for DXS6941	
WADL-1	5′-GGTGTCTGTGTACAGGTACCTCAG-3′
WADL-2	5′-GCCCAGACCCAGGTCCTTGG-3′
Primer sequences for HB3 (DXS9821)	
HB3-F	5′-TATGAAGGTTAGTTTGGCTGG-3′
HB3-R	5′-AGGAAGCTTGAACTGGGTGG-3′
Primer sequences for GW1E-EBP (DXS9824E)	
PAA3-F	5′-AGCACGCTGGATGCCAAGG-3′
PAA3-R	5′-TCGTCACCATCATAGACCTCC-3′
Primer sequences for HB1 (DXS9822)	
AM3-F1	5′-GGGTTGGAGTTTGTGTTAGG-3′
AM3-R1	5′-GCGTGGCATCAGCAGGCT-3′
Primer sequences for T54 (DXS7468E)	
54-F6	5′-GTTTCTAGCTTCCCTACTGGT-3′
54-R6	5′-CGGACAGATGAAGGCCGAGT-3′
Primer sequences for HB2E (DXS9823E)	
HB2-F	5′-CGTGGTGCGCTATGAGACCC-3′
HB2-R	5′-CATGGCAAAGGTGTTATCATCC-3′

genomic *Mlu*I site (Fig. 1). One isolated genomic clone termed HB2 maps adjacent to a CpG island and was shown to be conserved evolutionarily (data not shown). HB2 was sequenced and GRAIL analysis led to the prediction of the expressed sequence HB2E (DXS9823E, Table 2) that is identical with an EST from the data base (GenBank accession no. H21088). One of the HB2-containing cosmids (LL0XNC01–U86F9) includes the polymorphic repeat DXS573 (Thiselton et al. 1995) and the STS marker DXS1331 (Kere et al. 1992). Extension of the cosmid contig was done to the polymorphic marker DXS1039 (Dib et al. 1996) that could be placed proximal to the STS marker DXS666. In addition, three P1 artificial chromosomes (PAC) clones were isolated encompassing the region between HB2E and DXS1039 (see Figs. 1 and 4).

Toward a Cosmid Contig Connecting the Loci DXS1039 and DXS255

A YAC contig around the DXS255 locus and flanked by the markers DXS6666 and DXS146 was also characterized in this study. Colinearity with the genomic map was demonstrated for all but one YAC. The distal part of the YAC ICRFy900E0250 was shown to be chimeric after subcloning into cosmids (data not shown) and could therefore not connect DXS6666 with the SYP locus, which are only ~160 kb apart.

The constructed cosmid contig was overlapped proximally with two YACs. The YAC ICRFy900B0617 (400 kb) contains the markers DXS6666, DXS1426, DXS1416, DXS7647, DXS6850, as well as the probe HB4, but is deleted for the marker DXS1039. However, this YAC identifies the cosmid (LLNLc110C1459) that contains DXS1039, indicating only a small deletion within it. Another YAC that overlaps with the cosmid contig (yWXDF14D4, Boycott et al. 1996) contains the markers DXS6666, DXS1039, DXS1426, and DXS1416. In addition, comparative hybridization of the LL cosmid library with the YACs ICRFy900B0617 and yWXDF14D4 revealed >10 overlapping clones. The genomic clone HB4 (see above) was isolated from a cosmid identified only by B0617 and was shown to contain the marker DXS7647.

Overlap of the YAC ICRFy900B0617 with the YAC yWXDA39E7 (Boycott et al. 1996) was first

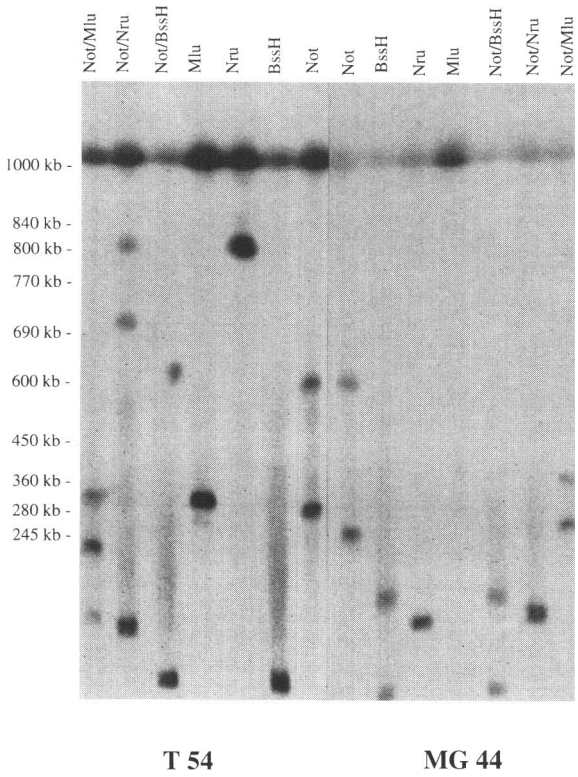


Figure 2 PFGE analysis connecting MG44 with T54. Both probes identify an identical partially digested *NotI* fragment, but different *NotI* end fragments and *MluI* fragments. Hybridization was performed under stringent conditions. Sizes are indicated on the *left*.

demonstrated by PCR reactions with the markers DXS8222 (Boycott et al. 1996) and DXS6850 (Fisher et al. 1995) and second by hybridization of the LL cosmid library revealing six overlapping clones. In addition, three of the cosmid clones found by the YAC yWXDA39E7 are identical with the cosmid clones identified by hybridization with the probe M27β (data not shown).

DISCUSSION

YAC Instability in a Gene-rich Region

A genomic map is an important tool especially for the cloning of gene-rich regions. Colinearity between the constructed YAC contigs and the genomic restriction map has not been attained in previously published reports, resulting in mislocalization of some ESTs like WASP or EST04093 (Derry et al. 1994). In addition, most YACs isolated from various libraries subsequently have proven to be either chimeric, unstable, or very small (Fisher et al. 1995; Kwan et al. 1995a,b; Boycott et al. 1996), making it difficult to construct an accurate physical map. The YACs previously used to connect OATL1 with GF-1 were not reliable as the YAC ICRFy900F0501 (Coleman et al. 1994; Fisher et al. 1995) and YAC ICI-27GF2 (Derry et al. 1994, 1995a; Fisher et al. 1995), both unstable and chimeric. Instability has also been shown for the YAC ICRFy900E021, which has been used previously to connect TFE-3 and SYP (Hagemann 1994). The central and distal parts of the region between ZNF21 and DXS255 have been especially difficult to clone in YACs in one study, with overall rates of 30% chimerism and 40–80% internal deletions (Boycott et al. 1996). The only stable and reliable YACs for this region thus far are quite small (<200 kb) and were isolated either from an X chromosome-specific YAC library (Lee et al. 1992) or from the St. Louis YAC library (Boycott et al. 1996). In all, >100 different YACs were isolated from various YAC libraries without achieving overlapping clones across the entire region (Fisher et al. 1995; Nelson et al. 1995; Boycott et al. 1996; this report). The importance of an integrated mapping approach using a long-range restriction map and a combined YAC-cosmid contig, especially in gene-rich regions, has been emphasized by other groups (Fisher et al. 1995; Boycott et al. 1996).

Of note is the finding that suitable YACs for Xq28, an X region with an even higher gene density than Xp11.23, was similarly difficult because of a plethora of highly chimeric and unstable YACs (Palmieri et al. 1994). Likewise, small unstable clones were found in the region of highest GC content, that is, the area of highest gene density, and it was speculated that this might be attributed to the following mechanisms: (1) some unknown factor intrinsic to the Xq28 region itself; (2) frequent tandem repeats (not actually shown in Xq28 or Xp11.23), which tend to self-delete from YACs; or (3) the high GC content not being as compatible with YACs, which tend to have a high AT content.

To circumvent these problems, we developed a restriction site-based frame for the entire Xp11.23–22 region encompassing ~3.5 Mb. The existing gap between the OATL1 pseudogene cluster and the locus DXS7465E (MG61), which are separated by ~70 kb, could be bridged by three cosmids of which two overlap with the YAC ICRFy900C0874 (Figs. 1 and 4). In addition to HB3 (DXS9821), the marker DXS6950 (Kwan et al. 1995a) that also had been derived from this

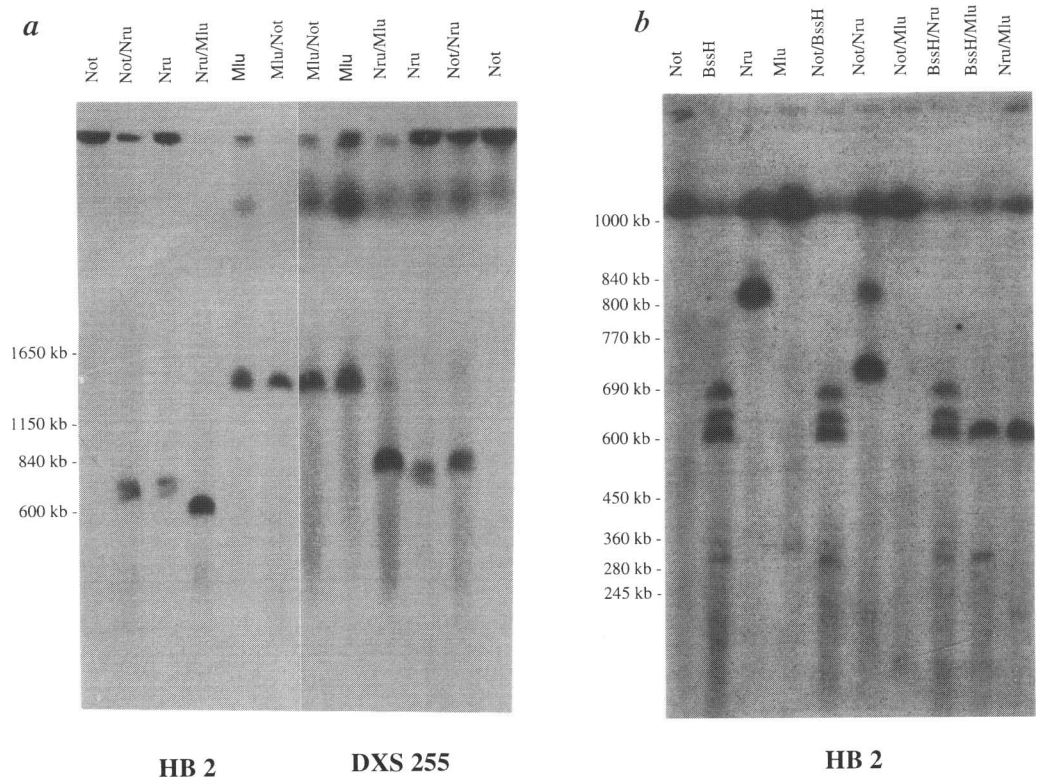


Figure 3 PFGE analysis in Xp11.22. (*a*) High-resolution PFGE connecting HB-2 with DXS255. The probe HB2 and the VNTR DXS255 hybridize with identical *Mlu*I and *Not*I fragments (>3 Mb), but with two different *Nru*I fragments. (*b*) Identification of two putative CpG islands by hybridization with HB2. Partially digested *Bss*HII fragments were not observed with other probes except for HB4.

YAC was mapped into the cosmid contig (e.g., LLNLc110P0252/K0616).

Another remaining gap in the contig between the loci DXS573 and DXSDXS666/1039 could be closed by four overlapping cosmids. Extension of the cosmid contig revealed the localization of DXS1039 proximal to DXS6666. Surprisingly, YAC ICRFy900B0617, which contains DXS6666 at its distal end, is not positive for DXS1039. In contrast, both the DXS6666 containing cosmid LLNLc110A0631 and the DXS1039 containing cosmid LLNLc110C1549 were found by hybridization of the LL cosmid library with this YAC. In addition, the St. Louis YAC yWXDF14D4 as well as the PAC 8N8 (Fig. 4) are positive for both markers, confirming tight physical linkage between them (Figs. 1 and 4) and making likely a microdeletion in YAC ICRFy900B0617.

A Sequence-ready Map Constructed Around Multiple CpG Islands

A sequence-ready map was constructed from a 1100-kb gene-rich region, and further markers to be identified can be mapped easily back to this cosmid contig. In addition, mapping of the cosmids using rare-cutter enzymes revealed more CpG islands not seen in genomic DNA with the rare cutters used. If every CpG island is associated

Figure 4 A cosmid/PAC contig between the loci HB3 and DXS1039 with six novel ESTs/STSs and complete resolution of all polymorphic or EST markers. The cosmids given were obtained either from the ICRF cosmid library, which are indicated by an asterisk (Nizetic et al. 1991; provided by RLDB, Berlin, ICRFc104), or from the LL X-cosmid library (provided by either RLDB, Berlin, LLNLc110; or from Baylor College, Houston, LL0XNC01, starting with U). The cosmid–PAC contig covers a region of ~1100 kb; the established *Eco*RI map encompasses ~1000 kb.

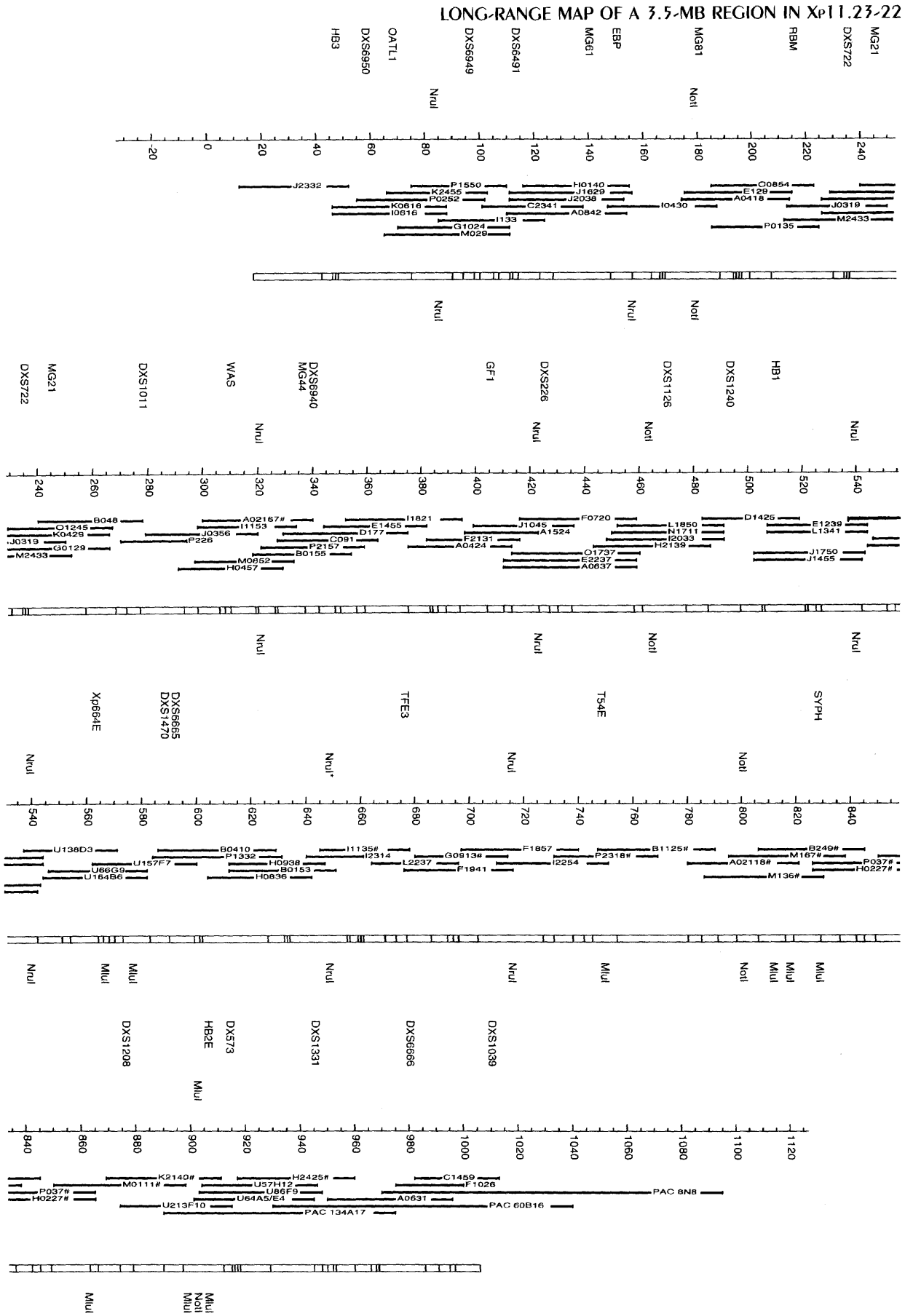


Figure 4 (*See facing page for legend.*)

a

```
        GATCTGAACCCAAACTAAATTTCCCAGCAAGCAGCGCGCCGGCCTGGGAAAAGGAGCAAG ATG GCT GAC TCC AAA GAG GGT GTT   84
                                                                      M   A   D   S   K   E   G   V
    TTG CCG CTG ACG CTG CTT CCA CTG CCC CAA TTT CAT TCG GCT TCA CTC GCA CGT CCG CAC GGA GGC GCT GGC CGA  159
  9  L   P   L   T   L   L   P   L   P   Q   F   H   S   A   S   L   A   R   P   H   G   G   A   G   R
    CTC GAG AGA CGG CGC GGG GCC ATC TCC GGA GGA AAA GGA TTT CTT GAA AAC CGT GGA AGG GAG GGA GCT GCA GAG  234
 34  L   E   R   R   R   G   A   I   S   G   G   K   G   F   L   E   N   R   G   R   E   G   A   A   E
    TGT GAA GCC CCA GGA GGC CCC CAA GGA ACT CGT CAT CCC TTT GAT CCA GAA TTG GCC ATC GCA GGC AGC CAC CAG  309
 59  C   E   A   P   G   G   P   Q   G   T   R   H   P   F   D   P   E   L   A   I   A   G   S   H   Q
    CCC GGC CCC CTG CAG GTC CAC AGA TAC TGG GGC CTT GCG GAT GGG GTG GTG TCC CAG GCT GTG AAG GAG CTC ATT  384
 84  P   G   P   L   Q   V   H   R   Y   W   G   L   A   D   G   V   V   S   Q   A   V   K   E   L   I
    GCG GAA TCC AAG AAG TCT CTG GAA GAA GGA AAG AAT GCG GGT GTC GAC CCC ACG CTC GCT ATC CCC ATG ATC CAG  459
109  A   E   S   K   K   S   L   E   E   G   K   N   A   G   V   D   P   T   L   A   I   P   M   I   Q
    AAA GGA TGC ACC CCC AGC GGG GAA GGG GCA GAC AGC GAA CCC CGG GCA GAG ACA GTG CCA GAG GAG GCT AAT TAT  534
134  K   G   C   T   P   S   G   E   G   A   D   S   E   P   R   A   E   T   V   P   E   E   A   N   Y
    GAG GCG GTC CCC GTG GAG GCC TAT GGG TGG CCA TGC TGC GGG GCA TGG GCT GGA AAC CTA GCG AGG CAT CGG CCG  609
159  E   A   V   P   V   E   A   Y   G   W   P   C   C   G   A   W   A   G   N   L   A   R   H   R   P
    CAC CTT CAA TCA AGT AGT GAA GCC CCG TGT CAA CTC ACT GAG GCC AAG GGG TTA GGG CTG GGG TGC CAA CCT GAC  684
184  H   L   Q   S   S   S   E   A   P   C   Q   L   T   E   A   K   G   L   G   L   G   C   Q   P   D
    CGA GCC CAG GCC TTG ACC CCC ACT GGC CCC TCC CGC ATG CCA AGA CCA GAT GAG GAG CAA GAG AAA GAT AAG GAA  759
209  R   A   Q   A   L   T   P   T   G   P   S   R   M   P   R   P   D   E   E   Q   E   K   D   K   E
    GAT CAG CCT CAA GGG CTG GTG CCT GGA GGA GCT GTG GTG GTT CTT TCT GGC CCT CAC CGA GGC CTC TAT GGG AAG  834
234  D   Q   P   Q   G   L   V   P   G   G   A   V   V   V   L   S   G   P   H   R   G   L   Y   G   K
    GTG GAA GGC CTT GAT CCT GAC AAT GTT CGG GCC ATG GTT CGT CTG GCT GTG TGG AGC CGG GTG GTG ACT GTT AGT  909
259  V   E   G   L   D   P   D   N   V   R   A   M   V   R   L   A   V   W   S   R   V   V   T   V   S
    GAG TAC TAC TGC GGC CTG TCT CCC AGC AGG AGT TTG ACA AGA ACA CCT TGG ATC TCA GGC AAC AGA ACG GAA CTG  984
284  E   Y   Y   C   G   L   S   P   S   R   S   L   T   R   T   P   W   I   S   G   N   R   T   E   L
    CCT CAT GCA CGG AAG ACC CTC TGG AAT CAA GAA CTC TAC ATC CAG CAG GAC AAC TCA GAG AGG AAG CGG AAA CAC 1059
309  P   H   A   R   K   T   L   W   N   Q   E   L   Y   I   Q   Q   D   N   S   E   R   K   R   K   H
    CTT CCA GAC CGA CAG GAT GGC CTG CAG CCA AGA GTG AGA AAG CAG CCC CCA GAA GTC AGC ACT GGT TGC ACA GGG 1134
334  L   P   D   R   Q   D   G   L   Q   P   R   V   R   K   Q   P   P   E   V   S   T   G   C   T   G
    ACC TGC GTG TGC GGT TTG TGG ACA ACA TGT ACA AAG GAG GCC AAT ATT ACA ACA CCA AGA TGA TAA TTGAAGATGTC 1211
359  T   C   V   C   G   L   W   T   T   C   T   K   E   A   N   I   T   T   P   R   *
    CTAAGCCCAGATACCTGTGTATGTCGGACAGATGAAGGCCGAGTCCTGGAAGGCCTGAGGGAAGACATGCTGGAGACCCTGGTTCCCAAGGCAGAGGGT 1310
    GACCGTGTGATGGTGGTGCTGGGCCCACAGACTGGAAGGGTGGGACATTTGCTGAGCCGGGACAGAGCACGGTGACCCGGGATTTGGTGCAACTGCCAA 1409
    GAGAAAATCAGGTGGTGGTGACTTCACTACGATGCCATCTGCCAGTACATGGGCCCTAGTGACACAGATGATGACTGACCCATGGGACTCCTCCCATCC 1508
    CCCAGGCTGGTACCAGTTCTGTACCATATGAGAAAGTTGCCTTCAGAAGGTGGGAAGATCATTGTTCCATCCTCTACTTCTGGTGCAGTCCTGGGACAA 1607
    GGACAAGGGAAAGGGATGGGTGAACCAGTAGGGAAGCTAGAAACAAACCCAATATTTACCAAAATTTAAGGGTATAATAAAAACCATTTCAAGTACTTA 1706
    ATAAAAAAAA
```

b

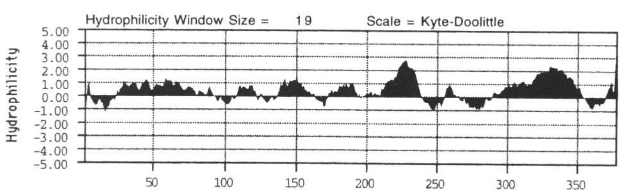


Figure 5 Entire cDNA sequence of T54 (GenBank accession no. U66359) encoding for a predicted novel protein of 378 amino acids. (*a*) The initiation codon is underlined and the stop codon is indicated by an asterisk. The poly(A) site is in bold letters, a putative nuclear location signal is doubly underlined. (*b*) Hydrophilicity profile of T54 using a window of 19 residues (Kyte and Doolittle 1982).

with a gene, then 20–30 additional genes can be expected to map in this region, making a total of at least 40 genes in 1000 kb at an average distance of at most 25 kb. Sequencing of a 75-kb region around HB2E showed two additional Gene Recognition and Analysis Internet Link (GRAIL)-predicted exon clusters.

A similar gene density is unlikely for the adjacent regions between TIMP and the pseudogene cluster OATL1 as well as between the loci DXS1039 and DXS255, for two reasons. First, significantly fewer CpG islands were detected in genomic DNA compared with the above gene-rich interval. Second, both regions were subcloned in colinear and longer (>300 kb) YACs (Coleman et al. 1994; Boycott et al. 1996), which correspond typically to regions of normal gene density. In contrast, preliminary analysis of cosmids and small YACs located between the genes ELK1 (Rao et al. 1989) and ZNF81 (Knight et al. 1994) as well as between DXS1039 and DXS255 indicate a relatively gene-rich region with an average distance of 50 kb, approximately twice as dense as the mean for the entire X chromosome.

ESTs Associated with Polymorphic Repeats

All of the ESTs localized in the Xp11.23 region are associated with one or more polymorphic $(CA)_n$-repeats. Refined mapping of these repeats provides markers for linkage analysis in WAS, which may now be performed using DXS6941 (this report), DXS722 (Thiselton et al. 1995), DXS6940 (Kwan et al. 1995a), and DXS1126 (Donelly et al. 1994). DXS722 was mapped between MG81 and MG21. The gene Xp664 is associated with the polymorphic marker DXS1470 and the adjacent gene TFE-3 with the novel marker DXS8221 (Boycott et al. 1996). DXS573 (Roustan et al. 1992) lies adjacent to HB2E, along with DXS1208 and DXS1039 forming a second group of linked markers proximal to the WAS group. Distally, the ubiquitously expressed gene L1 (EST04093; data not shown) could be mapped

between the polymorphic repeat DXS1367 (Schindelhauer et al. 1994) and the novel ZNF-$(CA)_n$-repeat (DXS9820), thereby constituting a fourth linkage cluster. The precise ordering and grouping of these markers will allow linkage analyses for a variety of diseases localized to this region.

A linkage map for the human genome published recently (Dib et al. 1996) contains a few large voids, one of which corresponds to the region discussed in this report. Thus, the novel and known microsatellites mapped in this study fill in one of the few remaining relatively marker-poor (>3 cM) regions of Xp.

Candidate Genes in the Gene-rich Interval

The characterized ESTs located in the region are either sequences encoding for proteins with unknown functions or transcription factors involved in hematopoesis like WASP (Derry et al. 1994), GF-1 (Zon et al. 1990), or TFE-3 (Beckmann and Kadesch. 1990). A further gene potentially involved in hematopoesis and located in that region will be the human homolog of the mouse *scurfy* gene (see below).

Four of the genes with unknown function (MG61, MG81, MG21, MG44) recently have been isolated from retinal tissues, but eye-restricted expression has not yet been demonstrated (Geraghty et al. 1993). All these cDNAs were excluded as candidate genes for RP2 in a single family (Hardcastle et al. 1995) but remain candidate genes for CSNB1 (Bech-Hansen and Pearce 1993) and the Aland Island eye disease (Glass et al. 1993). The EPB gene (Hanner et al. 1995) lies adjacent to MG61, and the RBM gene (Derry et al. 1995a) next to MG81, and both are, because of their functional role, not expected to be associated with eye diseases.

A complete cDNA sequence was established for the novel gene we named T54. It is expressed ubiquitously, but no homologies to known proteins emerged after a FASTA screen. A Kyte–Doolittle plot and screening for known protein motifs point to a hydrophilic protein with a nuclear location signal (Dingwall and Laskey 1991; Fig. 5).

HB2E is a novel gene that is part of an evolutionarily conserved sequence (data not shown). This sequence starts at a CpG island that contains a genomic *Mlu*I site and a *Not*I site identified on the cosmid level. Searches in the data base revealed an identical cDNA clone (GenBank accession no. H21088; Hillier et al. 1995) that shows distant homologies to a protein phosphatase 1. A search with exon prediction programs GRAIL2 and XPOUND revealed two further clusters of predicted exons that are located to either side of HB2E in the cosmid. Further characterization of these separate sequences is in progress.

Syntenic Regions in the Mouse

Comparative mapping of the mouse and human X chromosomes previously reported in this region demonstrated a partial inversion between these species with breakpoints distally near the human XK locus and proximally between PFC and OATL1 (Blair et al. 1994b). Thus, the loci TIMP or PFC and GF-1 or TFE-3 are still separated in the mouse genome but are adjacent to one another in the human genome (Blair et al. 1995). In mouse, the *scurfy* locus cosegregates with Gf-1 and Tfe-3, and *scurfy* was proposed as a mouse model for the human disease WAS (Blair et al. 1994a). This did not prove to be the case, however, as the mouse WASP gene was not mutated in the *scurfy* mouse (Derry et al. 1995b). Other candidate genes mutated in *scurfy* or *tattered* mice (Merrell et al. 1995) will become evident as this interval is sequenced.

In conclusion, we have constructed a genomic map of Xp11.23 using restriction-site based techniques instead of the unstable and chimeric YACs used previously. This was crucial to provide a more reliable genomic map in this difficult-to-clone region. This map can be used as a base for sequencing this region and characterizing precisely the disease loci in this region.

METHODS

PFGE Analysis

Agarose plugs of high molecular weight DNA were prepared either from blood lymphocytes or from yeast cells, digested with rare cutters acccording to the manufacturer's instructions (Boehringer, Stratagene), and fractionated in 1.5% agarose gels using a "Waltzer" apparatus (Anand et al. 1989). Genomic DNA was separated within a range of 50 to 950 kb (see Figs. 2 and 3b) or alternatively in the range of 300 to 2500 kb (Fig. 3a), whereas digested YAC DNA was separated within a range of 50 to 750 kb. Separated fragments were blotted onto nylon membranes (PALL) and hybridized according to standard protocols. Partially digested fragments are indicated by an asterisk in Figure 1.

Screening of YAC Libraries

YACs were obtained by probe screening from the ICRF reference library (Lehrach et al. 1990), the CEPH-A-YAC library (Albertson et al. 1990), or an X chromosome-specific library (Lee et al. 1992). Positive clones were confirmed by hybridization with the corresponding probes. Most of the YACs isolated with probes from the gene-rich interval were positive in a PCR-based approach but not after hybridization and therefore were not further characterized. Some of the YACs published by other groups were found to be chimeric after fine analysis: YAC ICRFy900A1220 was isolated with the probe TIMP but contains the polymorphic repeat DXS1368, which could be mapped distally to NDP in Xp11.4 (Schindelhauer et al. 1994); YAC ICRF y900A0120 demonstrated a nearly complete colinearity with the genomic map but the proximal end could be excluded from the region as the YAC is negative for the L1 gene; YAC ICRFy900F0501, which was used to connect OATL1 with GF-1 (Hagemann et al. 1994; Kwan et al. 1995b), contains a chimeric part between MG81 and MG21 and is unstable; and YAC ICRFy900E0250 has been subcloned in cosmids revealing colinearity with the region only with a short proximal part but not the large distal one. PAC clones were isolated from a PAC library (Pieter de Jong, unpubl., distributed by Sanger Centre, Hinxton) with the probe HB2 (134A17) and DXS1039 (60B16, 8N8). The YACs yWXDF14D4 and yWXDA39E7 isolated from the St. Louis YAC library (Nagajara et al. 1994) were kindly provided by Kym Boycott (University of Calgary, Canada).

YAC-DNA and YAC-plugs were prepared according to standard methods (Anand et al. 1989). The plugs were digested for 120 min using 10–20 units of rare cutter enzymes per plug. Partially and fully digested products were separated on agarose gels, blotted onto nylon membranes, and hybridized with YAC left- and right-arm probes as well as internal probes where available.

Identification and Characterization of Cosmids

The ICRF cosmid library (Nizetic et al. 1991), distributed by RLDB, Max Planck Institute, Berlin (ICRFc104), and the LL X chromosome-specific library (LL0XNC01, Biomedical Sciences Division, Lawrence Livermore National Library, Livermore, CA), distributed by either Baylor College, Houston (LL0XNC01-U) or by RLDB, Max Planck Institute, Berlin; Zehetner and Lehrach 1994, LLNLc110), were used. Cosmids were isolated with the probes 16H12 (eluted YAC fragment, ICRF A0330), MG61, WASP, DXS226, HB1, SYP, Xp664, TFE-3, 5H12 (eluted YAC fragment), SYP, HB2E, DXS6666, B0617 and A39E7 (both eluted YAC fragments), and DXS255, as well as by hybridization-based cosmid end clones. All cosmid clones were kindly provided by Nadja Pohl (RLDB, Max Planck Institute, Berlin, Group Hans Lehrach).

Isolation of Novel and Published Markers

Polymorphic Repeats

The corresponding cosmids were hybridized with a $(CA)_{18}$ oligonucleotide. Hybridizing fragments were eluted and sequenced. Polymorphisms were determined by the analysis of 100 unrelated females from Southern Germany. DXS6941: GDB-ID:435208; DXS9820 [ZNF$(CA)_n$]: GDB-ID: 1316857.

STSs

HB1 (DXS9822) was generated by inter-*Alu*-PCR (Ledbetter et al. 1990). Template for HB1 was the radiation hybrid A19C5 (Berger et al. 1992). STS HB3 (DXS9821) was derived from a 2.4-kb *Eco*RI fragment from the distal site of the cosmid LLNLc110K0616. The *Eco*RI fragment was subcloned and sequenced subsequently on an ABI sequencer. STS HB4 was derived from cosmid LLNLc110D0919 from which a 3.3-kb *Eco*RI fragment was isolated.

cDNA clones or ESTs

HB2E (DXS9823E): A 8.5-kb *Eco*RI/*Mlu*I fragment was isolated from LL0XNC01–U213F10 (Fig. 4) and sequenced at IMB, Jena (G. Nyakatura and A. Rosenthal). GRAIL analysis revealed two excellent exons that show identity to the EST HS088173 (GenBank accession no. H21088). Both predicted that exons, which are separated by 10 kb in genomic DNA, could be connected using lymphocyte cDNA; YAC 16H12 was used to hybridize a T-cell cDNA library (Clontech). One of the clones (L1) isolated is partially identical to EST 04093 (Adams et al. 1993); YAC 5H12 was used to hybridize the same T-cell library. One clone was shown to be identical with TFE-3; cDNA clone T54 was used to genenerate a complete cDNA sequence (see Fig. 5); GW1E = (EBP): Cosmid LLNLc110A0842 (Fig. 4) was used to screen a placental cDNA library (Lee et al. 1995). Sequence analysis of the picked clone GW1E revealed identity to the EBP cDNA (Hanner et al. 1995, GenBank accession no. Z37986).

ACKNOWLEDGMENTS

We are grateful to Kerry Baldwin-Jedele for help with the manuscript, Tim Strom for continuous support, and Kym Boycott for providing YACs. This work was funded by a grant from the Wilhelm Sander Stiftung (grant no. 93.057.1).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Adams, M.D., A.R. Kerlavage, C. Fields, and J.C. Venter. 1993. 3400 Expressed Sequence Tags Identify Diversity of Transcripts from Human Brain. *Nature Genet.* **4:** 256–267.

Albertson, H.M., H. Abderrahim, H.M. Cann, J. Dausset, D. LePaslier, and D. Cohen. 1990. Construction and characterization of yeast artificial chromosome library containing seven haploid genome equivalents. *Proc. Natl. Acad. Sci.* **87:** 4256–4260.

Anand, R., A. Villasante, and C. Tyler-Smith. 1989.

Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field gel electrophoresis. *Nucleic Acids Res.* **9:** 3425–3433.

Bech-Hansen, N.T. and W.G. Pearce. 1993. Manifestations of X-linked congenital stationary night blindness in three daughters of an affected male: Demonstration of homozygosity. *Am. J. Hum. Genet.* **52:** 71–77.

Beckmann, H., L.K. Su, and T. Kadesch. 1990. TFE3: A helix-loop-helix protein that activates transcription through the immunoglobulin enhancer uE3 motif. *Genes & Dev.* **4:** 167–179.

Berger, W., A. Meindl, B. de Leeuw, A. de Roos, T.J.R. van de Pol, T. Meitinger, S.D. van der Velde-Visser, H. Achatz, A. Geurts van Kessel, F.P.M. Cremers, and H.-H. Ropers. 1992. Generation and characterization of radiation reduced cell hybrids and isolation of probes from the proximal short arm of the human X chromosome. *Hum. Genet.* **90:** 243–246.

Blair, H.J., D.A. Carpenter, V.L. Godfrey, L.B. Russell, J.E. Wilkinson, and E.M. Rinchik. 1994a. The mouse scurfy (*sf*) mutation is tightly linked to *Gata*1 and *Tfe*3 on the proximal X chromosome. *Mamm. Genome* **5:** 652–654.

Blair, H.J., V. Reed, S.H. Laval, and Y. Boyd. 1994b. New insights into the man-mouse comparative map of the X chromosome. *Genomics* **19:** 215–220.

Blair, H.J., M. Ho, A.P. Monaco, S. Fisher, I.W. Craig, and Y. Boyd. 1995. High-resolution comparative mapping of the proximal region of the mouse X chromosome. *Genomics* **28:** 305–310.

Boycott, K.M., G.R. Halley, D. Schlessinger, and N.T. Bech-Hansen. 1996. A 2-megabase physical contig incorporating 43 DNA markers on the human X chromosome at p11.23-p11.22 from ZNF21 to DXS255. *Genomics* **33:** 488–497.

Chand, A., J. Clark, C.S. Cooper, and I.W. Craig. 1995. Long range organization of reiterated sequences, including the OATL1 cluster in Xp11.23. *Genomics* **30:** 545–552.

Coleman, M.P., A.H. Nemeth, L. Campbell, C.P. Raut, J. Weissenbach, and K.E. Davies. 1994. A 1.8 Mb YAC contig in Xp11.23: Identification of CpG islands and physical mapping of CA repeats in a region of high gene density. *Genomics* **21:** 337–343.

De Leeuw, B., W. Berger, R.J. Sinke, R.F. Suijkerbuijk, S. Gilgenkrantz, M.T. Geraghty, D. Valle, A.P. Monaco, H. Lehrach, H.H. Ropers, and A. Geurts van Kessel. 1993. Identification of a yeast artificial chromosome (YAC) spanning the synovial sarcoma specific t(X;18)(p11.2;q11.2) breakpoint. *Genes Chrom. Cancer* **6:** 182–189.

Derry, J.M.J., H.D. Ochs, and U. Francke. 1994. Isolation of a novel gene mutated in Wiskott-Aldrich syndrome. *Cell* **78:** 635–644.

Derry, J.M., J.A. Kerns, and U. Francke. 1995a. RBM3, a novel human gene in Xp11.23 with a putative RNA-binding domain. *Hum. Mol. Genet.* **4:** 2307–2311.

Derry, J.M.J., P. Wiedemann, P. Blair, Y. Wang, J.A. Kerns, V. Lemahieu, V.L. Godfrey, J.E. Wilkinson, and U. Francke. 1995b. The mouse homolog of the Wiskott-Aldrich syndrome protein (WASP) gene is highly conserved and maps near the scurfy (*sf*) mutation on the X chromosome. *Genomics* **29:** 471–477.

Dib, C., S. Faure, C. Fizames, D. Samson, N. Drouot, A. Vignal, P. Millasseau, S. Marc, J. Hazen, E. Seboun, et al. 1996. A comprehensive genetic map of the human genome based on 5.264 microsatellites. *Nature* **380:** 152–154.

Dingwall, C. and R.A. Laskey. 1991. Nuclear targeting—A consensus? *Trends Biochem. Sci.* **16:** 478–481.

Donelly, A., H. Kozman, A.K. Gedeon, S. Webb, M. Lynch, G.R. Sutherland, R.I. Richards, and J.C. Mulley. 1994. A linkage map of microsatellite markers on the X chromosome. *Genomics* **20:** 363–370.

Fisher, S.E., E. Hatchwell, A. Chand, N. Ockendon, A.P. Monaco and I.W. Craig. 1995. Construction of two YAC contigs in human Xp11.23-p11.22, one encompassing the loci OATL1, GATA, TFE3, and SYP, the other linking DXS255 to DXS146. *Genomics* **29:** 496–502.

Geraghty M.T., L.C. Brody, L.S. Martin, M. Marble, W. Kearns, P. Pearson, A.P. Monaco, H. Lehrach, and D. Valle. 1993. The isolation of cDNAs from OATL1 at Xp11.2 using a 480 kb YAC. *Genomics* **16:** 440–446.

Glass, I.A., P. Good, M.P. Coleman, P. Fullwood, M.G. Giles, S. Lindsay, A. H. Nemeth, K.E. Davies, H.A. Willshaw, A. Fielder, M.W. Kilpatrick, and P. Farndon. 1993. Genetic mapping of a cone and rod dysfunction (Aland Island eye disease) to the proximal short arm of the human X chromosome. *J. Med. Genet.* **30:** 1044–1050.

Hanner, M., F.F. Moebius, F. Weber, M. Grabner, J. Striessnig, and H. Glossmann. 1995. Phenylalkylamine Ca^{2+} antagonist binding protein. *J. Biochem.* **270:** 7551–7557.

Hageman, T., R. Surosky, A.P. Monaco, H. Lehrach, F.S. Rosen, and S.-P. Kwan. 1994. Physical mapping in a YAC contig of 11 markers on the human X chromosome in Xp11.23. *Genomics* **21:** 262–265.

Hardcastle, A.J, R.M. Hampson, D.L. Thiselton, M. Najadu, S.E. Jones, and S.S. Bhattacharya. 1995. Refinement of the locus for X-linked retinitis pigmentosa and exclusion of candidate genes. *Am. J. Hum. Genet. (Suppl.)* **57:** A1232.

Hillier, L., N. Clark, T. Dubuque, K. Elliston, M. Hawkins,

M. Holman, et al. 1995. The WashU-Merck EST project. http: //www2.ncbi.nlm.gov/cgi-bin/birx_doc?dbest+254126

Kere, J., R. Nagaraja, S. Mumm, A. Ciccodicola, M. D'Urso, and D. Schlessinger. 1992. Mapping human chromosomes by walking with sequence-tagged sites from end fragments of yeast artificial chromosome inserts. *Genomics* **14:** 241–248.

Knight, J.C., G. Grimaldi, H.J. Thiesen, N.T. Bech-Hansen, C.D.M. Fletcher, and M.P. Coleman. 1994. Clustered organization of Krüppel zinc-finger genes at Xp11.23, flanking a translocation breakpoint at OATL1: A physical map with locus assignments for znf21, znf41, znf81, and elk1. *Genomics* **21:** 180–187.

Kobayashi, H., L. Baumbach, T.C. Matise, A. Schiavi, F. Greenberg, and E.P. Hoffman. 1995. A gene for a severe lethal form of X-linked arthrogryposis (X-linked spinal muscular atrophy) maps to human chromosome Xp11.3-q11.2. *Hum. Mol. Genet.* **4:** 1213–1216.

Kwan, S.-P., T.L. Hagemann, R.M. Blaese, and F.S. Rosen. 1995a. A high-resolution map of genes, microsatellite markers, and new dinucleotide repeats from UBE1 to the GATA locus in the region Xp11.23. *Genomics* **29:** 247–252.

Kwan, S.-P., T.L. Hagemann, B.E. Radtke, R.M. Blaese, and F.S. Rosen. 1995b. Identification of mutations in the Wiskott-Aldrich syndrome gene and characterization of a polymorphic dinucleotide repeat at DXS6940 adjacent to the disease gene. *Proc. Natl. Acad. Sci.* **92:** 4706–4710.

Kyte, J. and R.D.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157:** 105–132.

Ledbetter, S.A., D.L. Nelson, S.T. Warren, and D.H. Ledbetter. 1990. Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequences polymerase chain reaction. *Genomics* **6:** 475–481.

Lee, J.T., A. Murgia, D.M. Sosnoki, I.M. Olivos, and R.L. Nussbaum. 1992. Construction and characterization of a yeast chromosome library for Xpter-Xq27.3: A systematic determination of cocloning rate and X-chromosome representation. *Genomics* **12:** 526–533.

Lee, C.C., A. Yazdani, M. Wehnert, Z.Y. Zhao, E.A. Lindsay, J. Bailey, M.I. Coolbaugh, L. Couch, M. Xiong, A. C. Chinault, A. Baldini, and C.T. Caskey. 1995. Isolation of chromosome-specific genes by reciprocal probing of arrayed cDNA and cosmid libraries. *Hum. Mol. Genet.* **4:** 1373–1380.

Lehrach, H., R. Ormanac, J. Hoheisel, Z. Larin, G. Lennon, A.P. Monaco, D. Nizetic, G. Zehetner, and A. Poustka. 1990. Genome analysis. In *Genetic and physical mapping* (ed. K.E. Davis and S.M. Tilghman), Vol. 1, pp. 39–81, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Lubs, H.A., P. Chiurazzi, J.F. Arena, C. Schwartz, L. Traeneberg, and G. Neri. 1996. XLMR genes: Update 1996. *Am. J. Med. Genet.* **64:** 147–157.

Meindl, A., M.R.S. deCarvalho, D. Schindelhauer, K. Herrmann, L. Grimm, M. Wehnert, M. Ross, and T. Meitinger. 1995. Two novel genes mapped to cosmid contigs in Xp21.1 and Xp11.23. *Cytogenet. Cell Genet.* **71:** 340.

Meitinger, T., N.A. Fraser, B. Lorenz, E. Zrenner, J. Murken, and I.W. Craig. 1989. Linkage of X-linked retinitis pigmentosa to the hypervariable DNA marker M27β (DXS255). *Hum. Genet.* **81:** 283–286.

Merrell, K., J.C. Gonzales, S. Wells, K. Calame, and G.E. Herman. 1995. Genetic analysis of Tattered, an X -linked dominant, developmental mouse mutation. *Mamm. Genome* **6:** 291–294.

Nagaraja, R., J. Kere, S. MacMillan, M.W.J. Masisi, D. Johnson, B.J. Molini, G.R. Halley, K. Wein, M. Trusgnich, B. Eble, et al. 1994. Characterization of four human YAC libraries for clone size, chimerism and X chromosome sequence representation. *Nucleic Acids Res.* **22:** 3406–3411.

Nelson, D.L., A. Ballabio, F. Cremers, A.P. Monaco, and D. Schlessinger. 1995. Report of the sixth international workshop on X chromosome mapping. *Cytogenet. Cell Genet.* **71:** 307–342.

Nizetic, D., G. Zehetner, A.P. Monaco, L. Gellen, B.D. Young, and H. Lehrach. 1991. Construction, arraying and high-density screening of large insert libraries of human chromosomes X and 21: their potential use as reference libraries. *Proc. Natl. Acad. Sci.* **88:** 3233–3237.

Ott, J., S. Bhattacharya, J.D. Chen, M.J. Denton, J. Donald, C. Dubay, G.J. Farrar, G.A. Fishman, D. Frey, A. Gal, et al. 1990. Localizing multiple X chromosome-linked retinitis pigmentosa loci using multilocus homogeneity tests. *Proc. Natl. Acad. Sci.* **87:** 701–704.

Özcelik, T., R.G. Lafreniere, B.T. Archer, P.A. Johnston, H.F. Willard, U. Francke, and T.C. Südhof. 1990. Synaptophysin: Structure of the human gene assignment to the X chromosome in man and mouse. *Am. J. Hum. Genet.* **47:** 551–561.

Palmieri, G., G. Romano, A. Ciccodiola, A. Casamassimi, C. Campanile, T. Esposito, V. Cappa, A. Lania, S. Johnson, R. Reinbold, et al. 1994. YAC contig organization and CpG island analysis in Xq28. *Genomics* **24:** 149–158.

Rao, V.N., K. Huebner, M. Isobe, A. ar-Rushdi, C.M. Croce, and E.S.P. Reddy. 1989. elk, tissue-specific ets-related genes on chromosomes X and near 14 translocation breakpoints. *Science* **244:** 66–70.

Roustan, P., A.R.J. Curtis, S. Kamakari, D.L. Thiselton, S. Lindsay, and S.S. Bhattacharya. 1992. Dinucleotide

repeat polymorphism at the DXS573 locus. *Hum. Mol. Genet.* **2:** 92.

Schindelhauer, D., H. Achatz, T.M. Strom, M. Ross, M.R.S. deCarvalho, and A. Meindl. 1994. Isolation and fine mapping of $(CA)_n$ repeats from the Xp11.23 and Xp11.4 region. *Hum. Mol. Genet.* **6:** 1027.

Thiselton, D.L., S. Lindsay, S. Kamakari, A.J. Hardcastle, P. Roustan, and S.S. Bhattacharya. 1995. Genetic and physical mapping of five novel microsatellite markers on human Xp21.1-p11.22. *Genomics* **25:** 279–281.

Zehetner, G. and H. Lehrach. The Reference Library System—Sharing biological material and experimental data. *Nature* **367:** 489–491.

Zon, L.I., S.F. Tsai, S. Burgess, P. Matsudaira, G.A. Bruns, and S.H. Orkin. 1990. The major human erythroid DNA-binding protein (GF-1): Primary sequence and localization of the gene to the X chromosome. *Proc. Natl. Acad. Sci.* **87:** 668–672.

Received June 25, 1996; accepted in revised form August 15, 1996.