

High-Resolution Landmark Framework for the Sequence-Ready Mapping of Xq23–q26.1

Helen E. Steingruber,¹ Andrew Dunham,¹ Alison J. Coffey,¹ Sheila M. Clegg, Gareth R. Howell, Gareth L. Maslen, Carol E. Scott, Rhian Gwilliam, Paul J. Hunt, Elizabeth C. Sotheran, Elizabeth J. Huckle, Sarah E. Hunt, Pawandeep Dhama, Cari Soderlund, Margaret A. Leversha, David R. Bentley, and Mark T. Ross²

The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, UK

We have established a landmark framework map over 20–25 Mb of the long arm of the human X chromosome using yeast artificial chromosome (YAC) clones. The map has approximately one landmark per 45 kb of DNA and stretches from DXS7531 in proximal Xq23 to DXS895 in proximal Xq26, connecting to published framework maps on its proximal and distal sides. There are three gaps in the framework map resulting from the failure to obtain clone coverage from the YAC resources available. Estimates of the maximum sizes of these gaps have been obtained. The four YAC contigs have been positioned and oriented using somatic-cell hybrids and fluorescence in situ hybridization, and the largest is estimated to cover ~15 Mb of DNA. The framework map is being used to assemble a sequence-ready map in large-insert bacterial clones, as part of an international effort to complete the sequence of the X chromosome. PAC and BAC contigs currently cover 18 Mb of the region, and from these, 12 Mb of finished sequence is available.

The complete sequence of the human genome is being determined as an international collaborative effort with the work proceeding on a chromosome-by-chromosome basis. Essential to the co-ordination of the work on individual chromosomes is the availability of an accurate framework map of landmarks (Bentley et al. 1998). Landmarks are used to construct the sequence-ready map in bacterial clones, to integrate confirmatory map data from other sources, and to define boundaries between sequencing groups. The last is particularly relevant to the sequencing of the human X chromosome in view of the large number of participating groups (see <http://webace.sanger.ac.uk/HGP/>).

Whole genome landmark framework maps have been produced by genetic linkage (Murray et al. 1994; Dib et al. 1996), by yeast artificial chromosome (YAC) content mapping (Chumakov et al. 1995; Hudson et al. 1995), and by radiation hybrid (RH) mapping (Hudson et al. 1995; Schuler et al. 1996; Stewart et al. 1997; Deloukas et al. 1998). The STS density of these maps (2–10 markers/Mb) is too low for the assembly of bacterial clone contigs over long distances, even when using large-insert bacterial and P1 artificial chromosome libraries [(BAC) Shizuya et al. 1992; (PAC) Ioannou et al. 1994]. There is a need, therefore, to produce higher density framework maps (15–20 markers per Mb) using RH mapping or YACs. The need is particularly acute for the X chromosome, because neither the whole genome

RH maps nor the genetic maps provide ordered landmarks at a density equivalent to that obtained for the autosomes.

Although there are particular limitations to the X chromosome framework map generated by whole genome approaches, the intense interest in X-linked disease genes has ensured the rapid progress of YAC contig assembly and landmark ordering over most of the chromosome. The editorial committees of successive X chromosome workshops have produced consensus landmark maps based on these individual regional efforts. The latest such map, the product of the Seventh X Chromosome Workshop (Sanger Centre, 1–4 October 1996), showed YAC continuity over most of the chromosome (X Chromosome Editorial Committee, unpubl.). At this time, the region between proximal Xq23 and proximal Xq26 was furthest from completion. In this work we report a high-resolution landmark framework map of this part of the X chromosome and summarize the progress made in the sequence-ready mapping of the region.

RESULTS

A landmark framework map of Xq23–q26.1 was constructed as follows: First, low-resolution order information was obtained for 69 published STS landmarks from Xq22–q26.3 using a panel of 11 rodent–human X chromosome deletion hybrids (Fig. 1). Second, assays developed from the STSs between Xq23 and Xq26.1 (54) were used to screen three YAC libraries with an expected 22-fold coverage of the X chromosome. Con-

¹These authors have contributed equally.

²Corresponding author.

E-MAIL mtr@sanger.ac.uk; FAX 44-(0) 1223-494919.

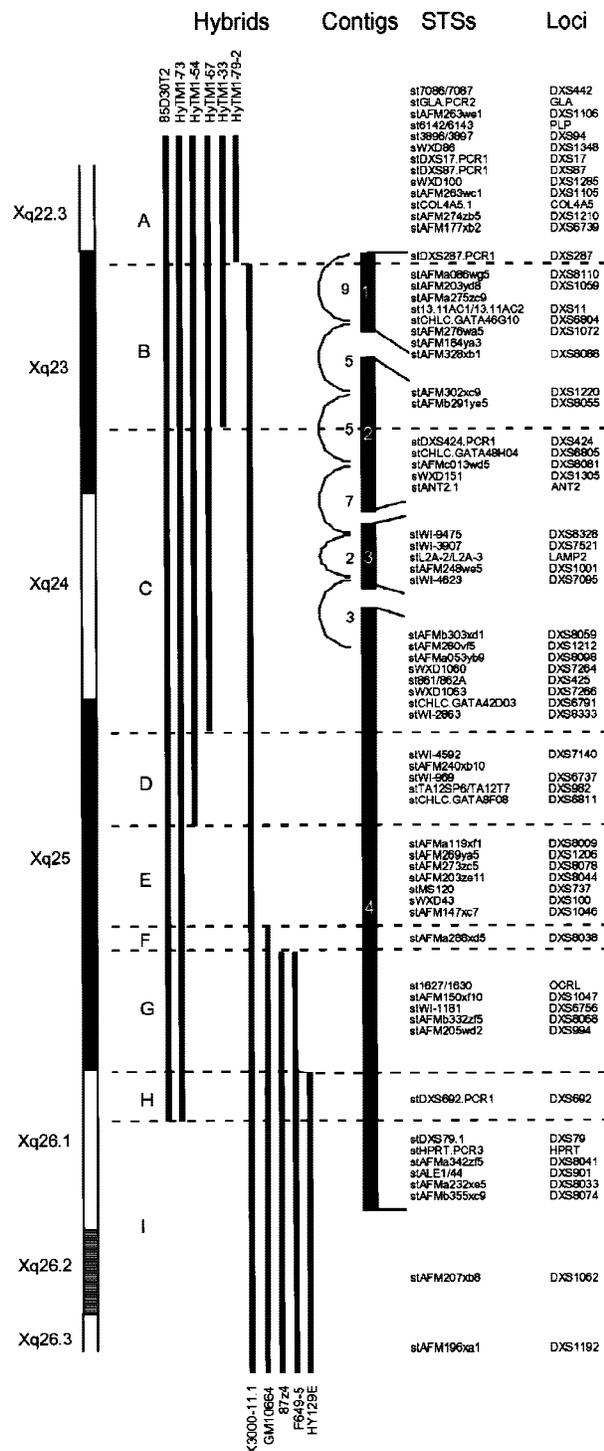


Figure 1 Hybrid interval map and framework map summary. A G-banded representation of Xq22.3–q26.3. The deletion-hybrid breakpoints define nine intervals (A–I) into which the 69 STSs have been placed. Within deletion intervals, STSs and their associated loci are shown in the order in which they appear on the framework map. The approximate positions and extents of the four YAC contigs are indicated by numbered bars. Numbers within the arcs at left of the contigs indicate the number of separate FISH experiments, using different PAC probe combinations, that support the contig order (arc drawn between bars) or orientation (within bars).

firmed positive clones were rearranged into microtitre plates and gridded onto hybridization membranes (the “X polygrid”). Third, further landmarks were generated from YAC clones at the ends of contigs and hybridized to the X polygrid filter to detect overlaps between contigs. Fourth, appropriate YAC end probes were also hybridized to filter arrays of the three libraries to identify new clones for gap closure. Further published landmarks that became available were incorporated into the framework map by hybridization to the X polygrid filter. These included STSs generated from the ends of PAC clones in the growing bacterial clone map (see below).

The extent of the landmark framework map is summarized in Figure 1, and the complete map can be seen in Figure 2. The map contains a total of 503 landmarks, equating to an average landmark separation of ~45 kb. The 294 STSs on the map include 35 polymorphic microsatellite markers and STSs derived from five genes [*ANT2* (Ku et al. 1990; Schiebel et al. 1994), *LAMP2* (Mattei et al. 1990), *XIAP* (Liston et al. 1996), *OCRL* (Silver et al. 1987), and *HPRT* (Pai et al. 1980)] plus four ESTs (loci DXS7028E, DXS7032E, DXS7571E, and DXS8321). The remaining 209 landmarks comprise 195 YAC end probes and 14 other hybridization probes. These landmarks are ordered on four YAC contigs containing 382 clones. The regions covered by the four contigs are as follows (ordered centromere to telomere): DXS7531–DXS8088 (contig 1, ~2–2.5 Mb); DXS7347–DXS7746 (contig 2, ~3.5–4 Mb); DXS7868–DXS7328 (contig 3, ~1–1.5 Mb); and DXS7323–DXS895 (contig 4, ~15 Mb).

The average depth of YAC coverage beneath each landmark is 8.4. At three points, clone coverage is much lower and most of the linking clones are deleted: In contig 1, there is a region that is spanned by only three clones (yM976E1, yM900H7, and yX74A12), the last two of which appear to have large central deletions; in contig 2, a single deleted YAC (yM963A8) joins two regions with deep clone coverage; and at the distal end of contig 4, two deleted YACs (yM941D7 and yX64G2) join the bulk of the contig to the segment containing the *HPRT* gene. In the last of these cases, the weak link is spanned by a bacterial clone contig (Chr_Xctg488 in Fig. 2).

We were unable to close the three gaps using the YAC resources described. However, the four contigs are ordered and oriented by multiple pieces of data. The hybrid interval map is wholly consistent with the landmark content of the contigs and orders the four contigs relative to each other. Additionally, the map orientates contigs 1, 2, and 4 (Fig. 1). Contig 1 also contains 13 YAC clones found at the distal end of the published Xq22 contig of Kendall et al. (1997). At the distal end, contig 4 overlaps the proximal ends of the Xq26 contigs of Cole et al. (1992) and Pilia et al. (1996).

Additional confirmation of contig order and orientation have been provided by metaphase and interphase fluorescence in situ hybridization (FISH). PAC clones isolated using selected STSs from the framework map have been used as probes in a series of two-color hybridization experiments. As summarized in Figure 1, these have confirmed the order of the four contigs using multiple separate probe combinations. These FISH data have also confirmed the orientations of contigs 1 and 2 and provided the orientation of contig 3.

Approximate estimates of the upper limits for the gap sizes have been obtained by interphase FISH. A pair of PAC clones flanking each gap was hybridized simultaneously with a control pair of known separation to interphase chromosome spreads. The estimated sizes obtained (s.d.s in parentheses) were 223 kb (\pm 112 kb), 236 kb (\pm 105 kb), and 292 kb (\pm 133 kb), for gaps 1, 2, and 3, respectively. Although the errors in this experiment are large, the maximum sizes predicted within 95% confidence intervals are 440, 440, and 550 kb for the three gaps. In an alternative approach, probes located on either side of each gap were hybridized to Southern blots of genomic DNA that had been digested with infrequently cutting restriction enzymes and separated by pulsed-field gel electrophoresis. This approach was successful only in the case of gap 2, in which common *Bss*HII and *Eag*I fragments of 360 kb were observed for probes derived from the distal end of contig 2 and the proximal end of contig 3.

DISCUSSION

The Framework Map

By analyzing the landmark content of YAC clones, we have assembled a landmark framework for the sequence-ready mapping of human Xq23-q26.1 with a resolution of ~45 kb. There are three gaps in the landmark framework in which chromosome walking in BAC and PAC libraries will be required for closure. The available evidence suggests that the majority of these gaps in YAC resources will be represented in the bacterial libraries. For example, of eight tested gaps in a chromosome 22 YAC contig (Collins et al. 1995), seven were closed using bacterial clones (I. Dunham, pers. comm.). Furthermore, our results indicate that the three gaps described here are small and should be closed by relatively few steps using current PAC and BAC resources.

The framework map information was applied successfully in the search for the X-linked lymphoproliferative disease gene [*XLP* (MIM 308240); Purtilo et al. 1975]. The critical region for the *XLP* gene, defined by a deletion of ~3–5 Mb of Xq25 between DXS6791 (stCHLC.GATA42D03) and DXS100 (pr45d), is completely encompassed by the framework map (contig 4). Landmarks from the map were used in the construc-

tion of bacterial clone contigs, from which genomic sequence and exon-trapping data were generated for candidate gene identification. This approach led to the identification of a gene mutated in XLP patients and encoding a novel SH2-domain-containing protein (*SH2D1A*; Coffey et al. 1998).

Other disease genes that are potentially within the coverage of the framework map include X-linked arthrogyrosis (Zori et al. 1998), a gene that may be involved in the progression of ovarian carcinoma (Choi et al. 1997), a gene for X-linked nonprogressive congenital cerebellar hypoplasia (Illarioshkin et al. 1996), the congenital generalized hypertrichosis gene (*CGH*; Figueroa et al. 1995), a gene involved in axonal motor-sensory neuropathy with deafness and mental retardation (Priest et al. 1995), and several nonspecific mental retardation loci [*MRX23* (Gregg et al. 1996); *MRX27* (Gedeon et al. 1996); *MRX35* (Gu et al. 1996); *MRX47* (des Portes et al. 1997); *MRX46* (Yntema et al. 1998)].

Comparison with Other Maps

Several published genetic and physical maps encompass the region between DXS7531 and DXS895 and share STSs with the framework map. Comparison of the marker order between these maps is therefore possible.

Twenty-two STSs are shared with the Whitehead Institute radiation hybrid map (Hudson et al. 1995), and overall, there is excellent agreement between the two maps. The only notable discrepancy concerns the Whitehead framework marker stCHLC.GATA48H04, which is found in YAC contig 2, but in the RH map, occupies a position proximal to markers from YAC contig 1. FISH analysis of bacterial clones confirms the framework map order by placing stCHLC.GATA48H04 distal to stAFM302xc9.

Fifteen STSs from the framework map are ordered in bins of the Stanford Human Genome Center RH map (Stewart et al. 1997). There is concordance between the order of the 1000:1 bins on the RH map and the landmark order described here, with landmarks from contigs 1 and 2 occupying bin 20 and landmarks from contig 4 found in bin 21, apart from the most distal landmark (stAFMb355xc9), which is found in bin 24. However, within the bins, some marker reorderings are suggested by the higher resolution YAC content data.

Two maps of particular importance to the X chromosome community are the Généthon genetic map (Dib et al. 1996) and the whole X chromosome YAC-STS map of Nagaraja et al. (1997), which was recently superseded in Xq23-q26 by a regional map (Nagaraja et al. 1998). The genetic map and the framework map share 28 STSs, and Figure 3 shows the differences in marker order between the two. Support for the framework map order is provided by the bacterial clone map

assembly (see below). A total of 163 STS landmarks, all of which are included in Figure 3, are shared with the map described by Nagaraja et al. (1998). In many areas there is good agreement between the two maps. In others there are differences in local marker order between

the maps, which often involve the inversion of small groups of markers. More serious discrepancies are seen for sWXD1100, sWXD1070, and sWXD1822. Evidence for the framework marker order described here is considered below.

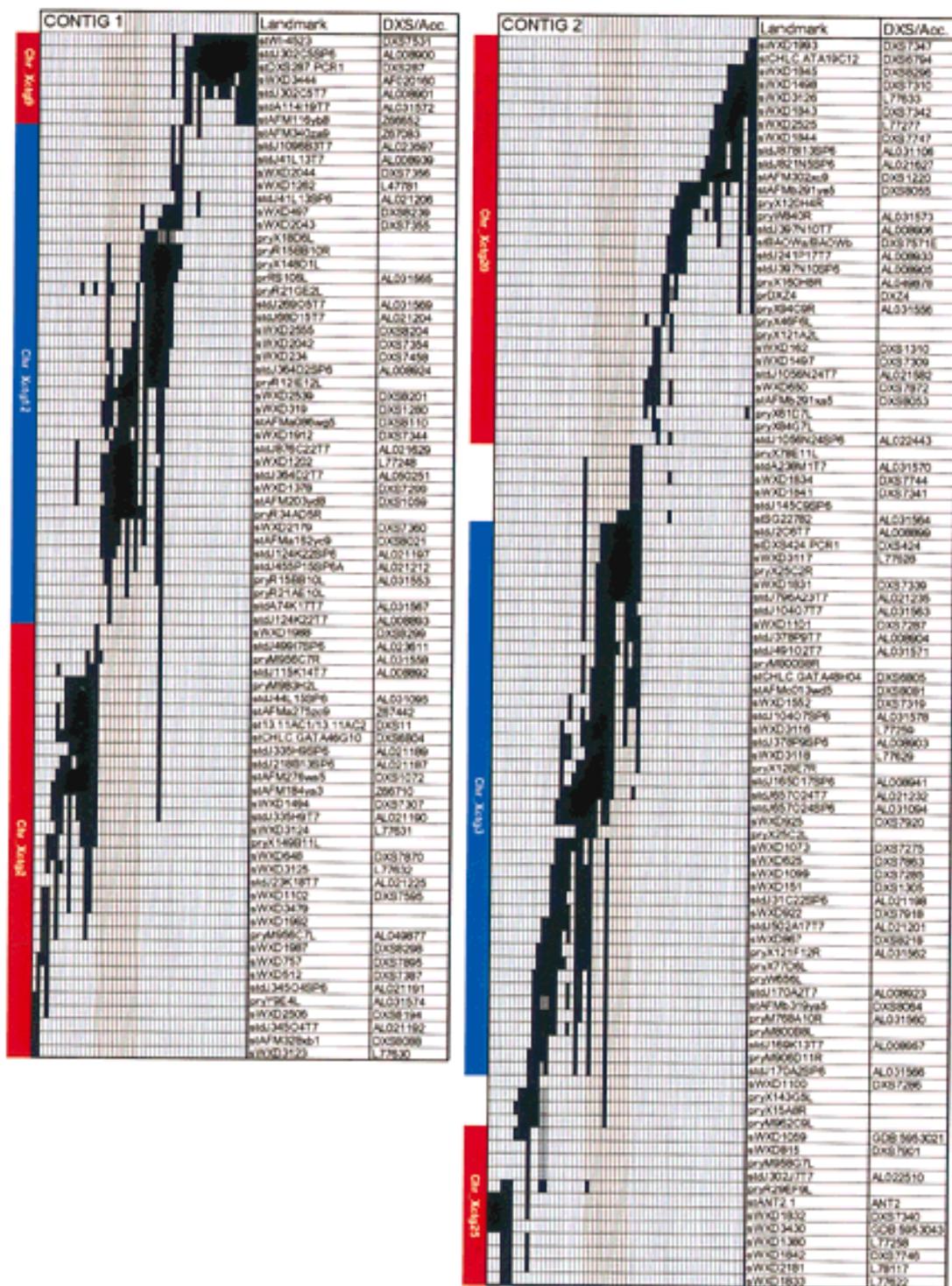


Figure 2 (See p. 757 for legend.)

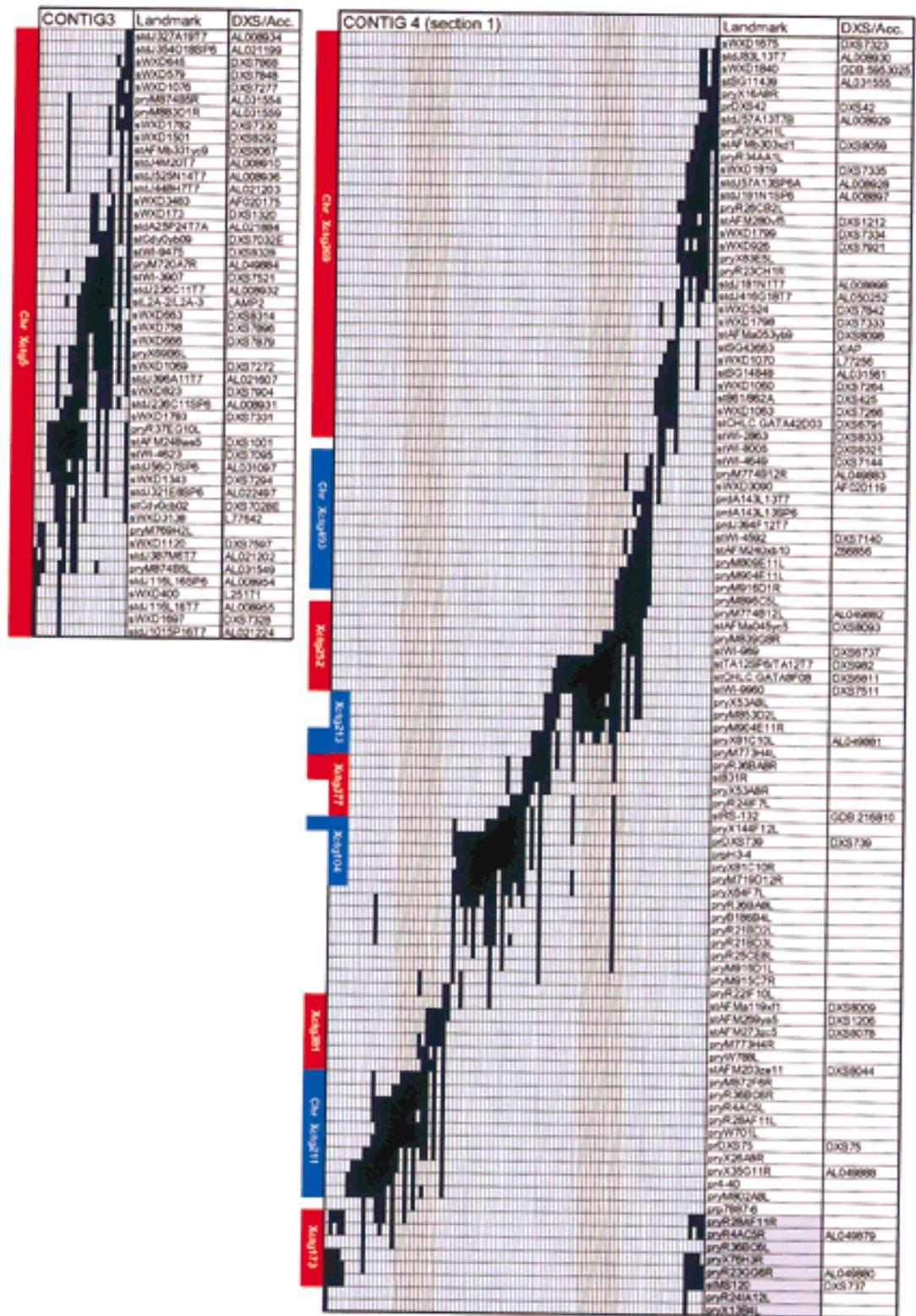


Figure 2 (See p. 757 for legend.)

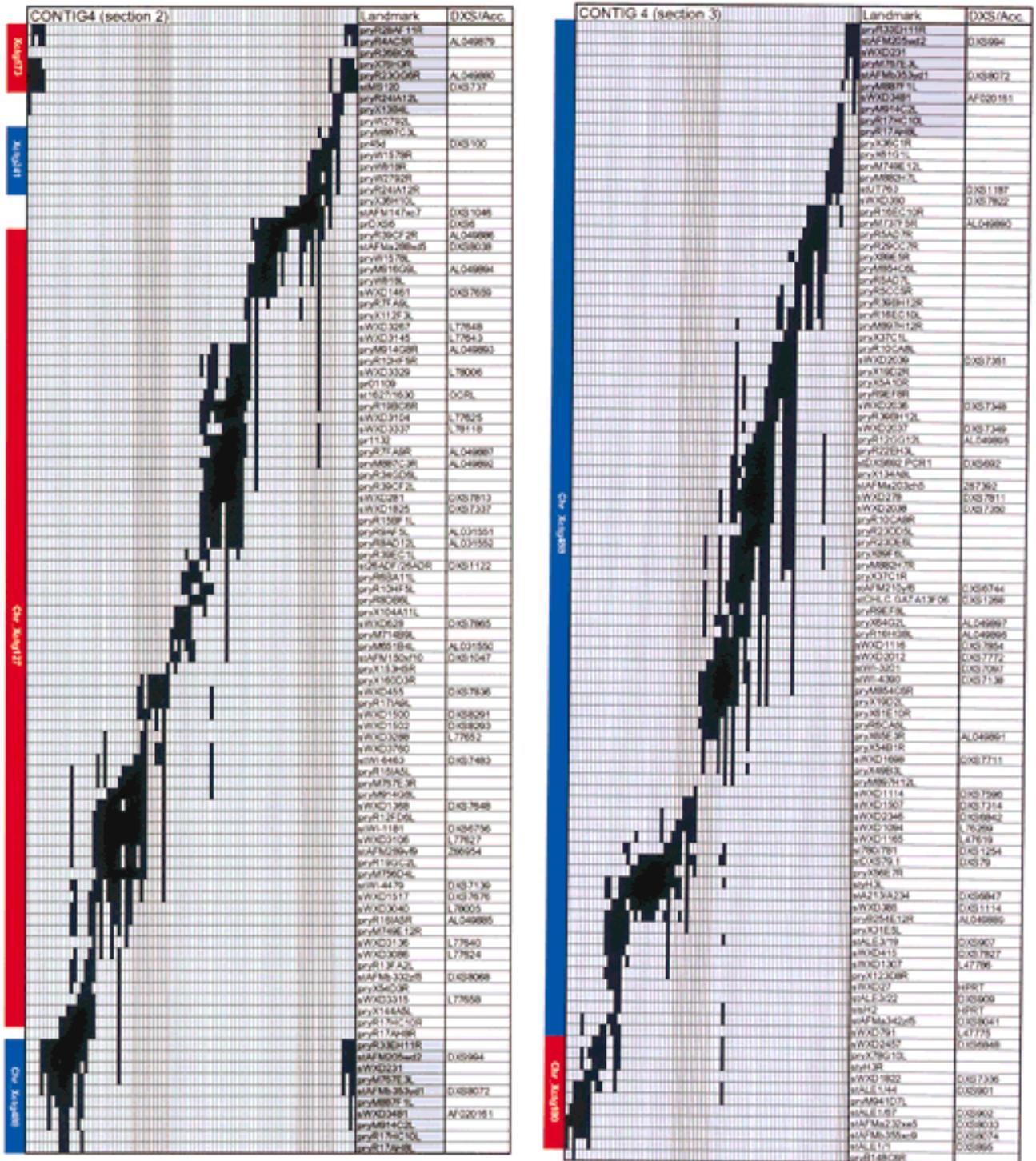


Figure 2 (See facing page for legend.)

It is interesting to note that two of the three gaps in the framework map coincide with two of the five gaps in the map of Nagaraja et al. (1998; Fig. 3), despite the use of a different but overlapping set of YAC clone resources in the two studies. The gap in the framework map between contigs 2 and 3 is closed in the map of Nagaraja et al. (1998). Particularly on the distal side, the join is made using clones from other libraries than those used in this study, and many of these are deleted. The evidence suggests, therefore, that two of the gaps in our map may be unclonable using the YAC system and the third may represent a region of instability in yeast.

Use of the Framework Map in Sequence-Ready Mapping

The availability of a high-resolution landmark framework has allowed the assembly of sequence-ready bacterial clone contigs over ~18 Mb (75%–90%) of Xq23-q26.1 (see <http://www.sanger.ac.uk/HGP/ChrX>). The RPCI1, RPCI3, RPCI4, RPCI5, RPCI6, RPCI11, and RPCI13 PAC and BAC libraries (kindly provided by Pieter de Jong and Joe Catanese; see <http://bacpac.med.buffalo.edu/>) were used as a source of clones, and contigs were assembled by a combination of restriction fingerprinting and landmark content (Mungall et al. 1996). Where necessary, the YAC end probes from the framework map are being converted into STSs for this purpose (see Fig. 2 for accession numbers). The locations and extents of the 20 bacterial clone contigs are illustrated in Figure 2, and all bacterial clone contigs can be viewed in detail via <http://www.sanger.ac.uk/HGP/ChrX>. Nine of the contigs are >1 Mb in size, with the largest (Chr_Xctg488) estimated at 2.7 Mb.

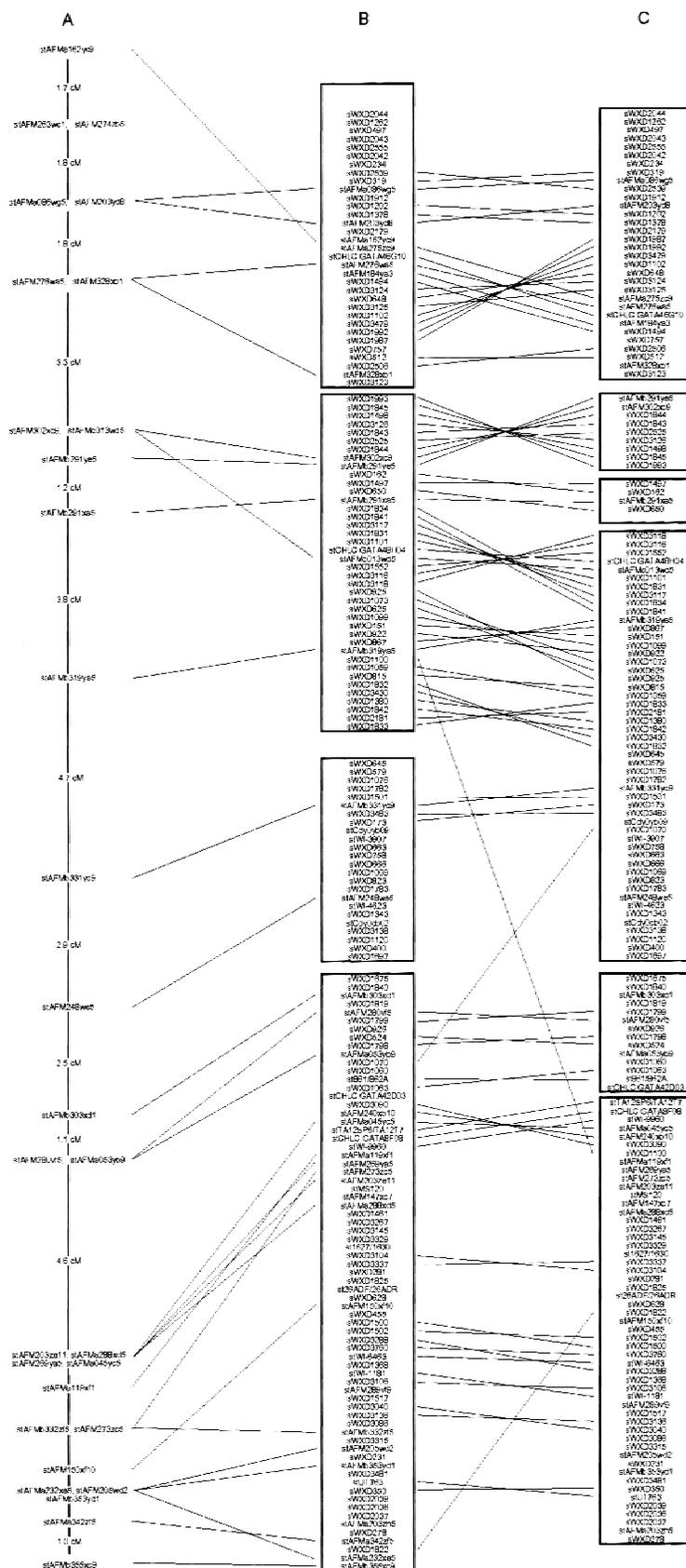
We have observed very close agreement between the framework map and the marker order on the bacterial clone contigs (e.g., Fig. 4), with any observed differences being restricted to minor local marker rear-

rangements suggested by the higher resolution bacterial clone maps. The assembly of the bacterial clone contigs has provided independent confirmation of the accuracy of the framework map in cases of conflict with other maps. Concerning discrepancies with the genetic map, the identification of PAC clones containing both stAFMa162yc9 and stAFM203yd8 supports the more distal location of the former suggested by the framework map; the relocation of stAFMc013wd5 is supported by the assembly of Chr_Xctg3 (Fig. 2) that contains both this marker and stAFMb319ya5; more distally, stAFMa119xf1, stAFM269ya5, stAFM273zc5, and stAFM203ze11 lie within Chr_Xctg381, stAFMa288xd5, stAFM150xf10, and stAFMb332zf5 in Chr_Xctg127, and stAFM205wd2, stAFMb353yd1, and stAFMa342zf5 in Chr_Xctg488 (all Fig. 2), and in each case the framework map order is maintained; and finally, stAFMa232xe5 and stAFMb355xc9 share multiple PAC clones in Chr_Xctg180, thus supporting their proximity on the framework map. Concerning the major clashes with Nagaraja et al. (1998), sWXD1070 and sWXD1822 are at the positions of Chr_Xctg269 and Chr_Xctg180, respectively, predicted by the framework map; PAC clones positive for sWXD1100 are not yet incorporated into Chr_Xctg3 by restriction fingerprinting, though there are some landmark content data supporting this link.

PAC clones from the contigs summarized in Figure 2 covering an estimated 17 Mb of DNA are currently being sequenced, either at the Sanger Centre or at the Genome Sequencing Center (St. Louis, MO). This collaborative effort has yielded a total of 12 Mb of finished sequence to date. The finished sequence, together with unfinished sequence from the region, is available from ftp://ftp.sanger.ac.uk/pub/human/sequences/Chr_X/.

The completion of the sequence-ready map of Xq23-q26.1 is underway. The few remaining markers from the framework map that lie in the gaps between existing bacterial clone contigs are being targeted for BAC screening. Otherwise, chromosome walking from

Figure 2 The landmark framework map of Xq23-q26.1. The landmark order on the four YAC contigs is illustrated. YAC clones are shown as thin vertical bars with positive landmark content indicated in black. Cases in which YAC clones have not been tested with specific markers are indicated by checkered boxes. Contig 4 is shown in three sections. For ease of alignment, some landmarks (solid gray boxes) and YAC clones have been shown duplicated at the end and beginning of appropriate sections. Landmark STSs (st or sWXD prefix) and probes (pr) are shown in the first column, and DXS numbers, database accession numbers, or Genome Database (GDB) identifiers in the second. Published STSs are derived from the Génethon (stAFM), Whitehead Institute (stWI), Co-operative Human Linkage Center (stCHLC), Center for Genetics in Medicine (sWXD) maps and databases, from GDB (various), or from the literature (see Methods). PAC end STSs are named after the PAC clone of origin with the suffix SP6 or T7 (e.g., stdJ364D2SP6). PAC clones are derived from the Roswell Park Cancer Institute libraries RPCI1, RPCI3, RPCI4, RPCI5 (dj), and RPCI6 (dA). YAC end hybridization probes are named after the clones from which they are derived with the "L" or "R" suffix (see Methods). Clones are from the CEPH (yM), ICI (yR), ICRF [yX (ICRFy900), yY (ICRFy901), yZ (ICRFy905)], and Washington University libraries (yA, yB), or were provided by colleagues (see Methods). Where YAC end probes have been sequenced and used to design STSs, EMBL accession numbers are shown. Three of these STSs were used for YAC library screening in preference to the end hybridization probe [stSG11439 (pryR17HB2R), stSG14848 (pryM949F11R), and stSG22782 (pryM963A8R)]. At *left* of each YAC contig are shown the extents of the corresponding bacterial clone contigs as red or blue boxes. Bacterial clone contigs are numbered and have the prefix Chr_Xctg (where space allows). In contig 4 (section 1), contigs X_CTG213, 377, and 104 are labeled in boxes to the *right* of those describing their extents because of space constraints. All of the bacterial clone contigs can be viewed in detail by following instructions at <http://www.sanger.ac.uk/HGP/ChrX/>. To avoid confusion, the lower numbered contig will be maintained on the WWW site as contigs join. A magnified version of this map showing YAC clone names can be viewed via <http://www.sanger.ac.uk/HGP/ChrX/>.



the ends of the bacterial contigs is being applied where such markers do not exist. Full details of all of our X chromosome bacterial clone contigs and sequencing progress can be obtained via <http://www.sanger.ac.uk/HGP/ChrX/>.

METHODS

Landmarks

Full details of all landmarks and their associated positive YAC clones can be obtained by querying our X chromosome database via the internet (see below).

The majority of published STS primer pairs were obtained from the Génethon genetic map and database (Dib et al. 1996), the Co-operative Human Linkage Center (Buetow et al. 1994), the Whitehead Institute/MIT Center for Genome Research map (Hudson et al. 1995), the Center for Genetics in Medicine (CGM, Washington University) X chromosome map (Nagaraja et al. 1997), and from the Genome Database (GDB; <http://gdbwww.gdb.org/>) (Fasman et al. 1997). stALE1/1 (DXS895), stALE1/44 (DXS901), stALE1/67 (DXS902), stALE3/19 (DXS907), and stALE3/22 (DXS909) are described in Cole et al. (1991); and styH3L, styH3R, and stH2 are described in Cole et al. (1992). Primer sequences for stANT2.1 (*ANT2*) are 5'-ACCCTAAAC-CCAGCCTTGAC-3' and 5'-TGAGACGATTGTC-TACAGG-3', and stDXS79.1 (DXS79) are 5'-CTGAACAGGAATAACTTGACA-3' and 5'-AGGATTACTACAGAGGGGAT-3', for stCOL4A5.1 (*COL4A5*) are 5'-TCCTAGGATGCAGTGTCTCATTGTC-3' and 5'-GGATACAGCAGGATTAGTAGCACC-3', and for stSG43663 (*XIAP*) are 5'-ACTTGTGTACCTGCAGACATCA-3' and 5'-GCTTCATAATCTGCCATTGGAT-3'. G. Porta kindly provided st831R, derived from the right end of YAC yWXD831 (CGM map).

YAC clone ends were obtained by the vectorette method of Riley et al. (1990), with the modifications described by Coffey et al. (1992). The only further modification here was that primers pYACL (5'-AATTATCACTACGGAATTC-3'; *TRP1* arm of pYAC4) or pYACR (5'-CCGATCTCAAGATTACGGAATTC-3'; *URA3* arm; Coulson et al. 1991) were used in conjunction with the universal vectorette primer (224) instead of the Sup4-2 and Sup4-3 primers described in Coffey et al. (1992). Many of

Figure 3 Comparison of the framework map with other published maps. A comparison is shown between the landmark order on the framework map (B), the Génethon genetic map (A) (Dib et al. 1996), and the Center for Genetics in Medicine YAC-STS map (C) (Nagaraja et al. 1998). All STSs shared by the framework map and these other maps are shown. Genetic distances are shown for A, except in cases where distance is <1 cM. Boxes around STSs in B and C indicate that markers are from the same YAC contig. In the comparison of B and C, lines are drawn only when landmarks show differences in order between the two maps.

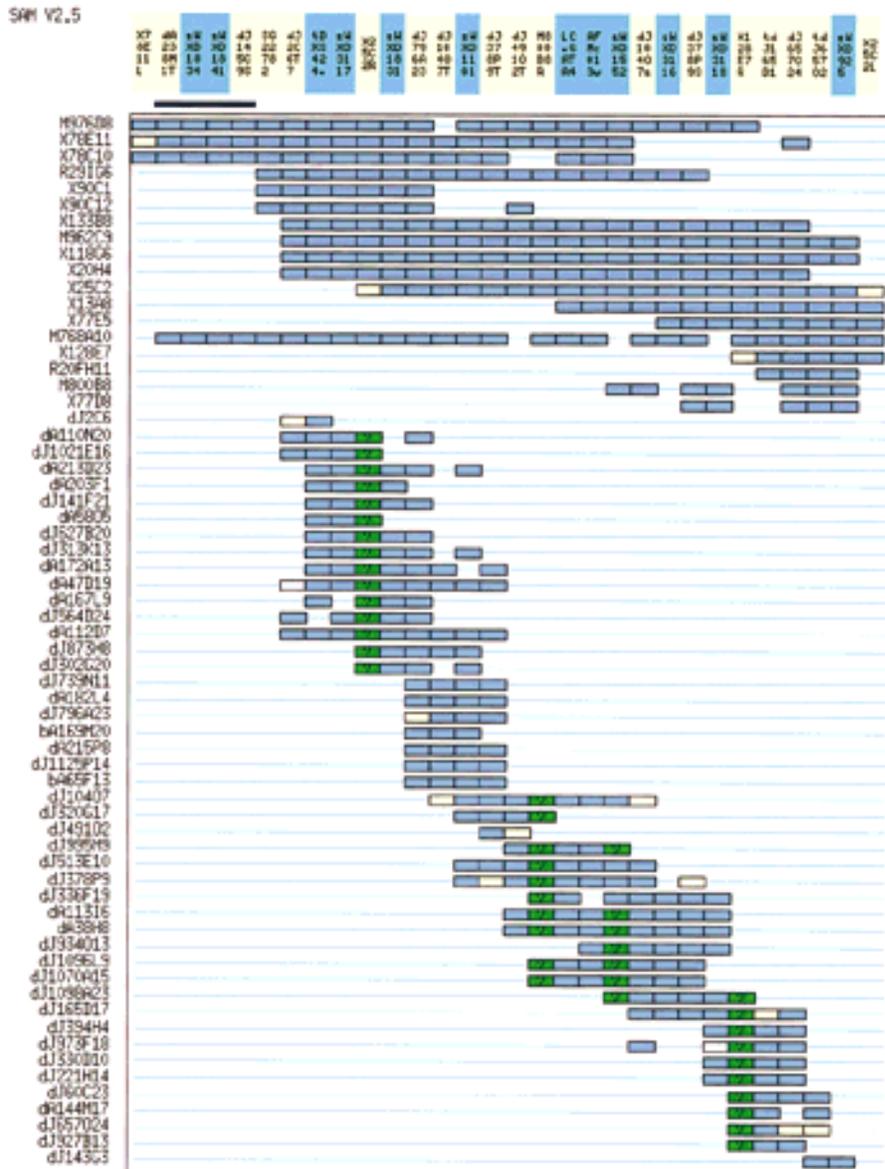


Figure 4 Segments of YAC contig 2 and of the underlying bacterial clone contig (Chr_Xctg3). The image is taken from the graphical display of the contig assembly program SAM. Landmarks are shown at *top*, with clone ends (YAC or PAC) in yellow and other STSs in blue–green. YACs are named according to Fig. 2; PACs are prefixed dj (RPC11, RPC13, RPC14, and RPC15) or dA (RPC16); and BACs, bA (RPC11). Dark blue boxes indicate strong positive landmark–clone associations; and pale gray, weak associations; a yellow box indicates the end of the YAC or PAC clone from which a probe or STS was derived; green boxes show cases in which bacterial clones have not been screened for the relevant markers.

the YAC end probes have been converted to STSs by cycle sequencing of the vectorette PCR products using 224, pYACR, or pYACL primers. Following repeat masking (REPEATMASKER; A.F.A. Smit and P. Green, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>), primer pairs were designed using PRIMER (Lincoln et al. 1991). Where STSs have been defined, database accession numbers are shown in Figure 2.

Some PAC clones in the bacterial map were also used as a source of end STSs for YAC library screening. DNA was prepared from 15-ml cultures using the Qiagen Qiawell-8 kit and

was used for cycle sequencing using primers from the T7 (5'-TAATAC-GACTCACTATAG-3') and SP6 (5'-ATTTAGGTGACACTATAG-3') promoter regions of the pCYPAC2 vector (<http://bacpac.med.buffalo.edu/>). Primer pairs were designed following repeat masking, as described above, and STS accession numbers are shown in Figure 2.

Three PAC clone-end hybridization probes (prdA143L13SP6, prdA143L13T7, and prdJ394F12T7) were obtained using the vectorette method described above for YAC clones. In the first round of PCR, the 224 primer was used in conjunction with PACS2 (5'-CGATCCTCCCGAATTGACTA-3'; 3P6 end) or PACT2 (5'-CTGGGTT-GAAGGCTCTCAAG-3'; T7 end). In the secondary amplification, the respective vector primers were 3P6PAC (5'-ATTTAGGTGACAC-TATAG-3') and T7PAC (5'-TAAT-ACGACTCACTATAGGGAGA-3').

Hybridization probes pr4-40, pr1132, and pr01109 were kindly provided by D. Toniolo (IGBE, CNR, Pavia, Italy), and pr7887-6 and prpH3-4 by B. Sylla (International Agency for Research on Cancer, Lyon, France). Hybridization probes detecting the loci DXS6, DXS42, DXS75, DXS100, and DXS739 were identified from the GDB. Data on the association of YACs with the DXZ4 locus were obtained from the Max Planck Institute for Molecular Genetics database (<http://ixdb.mpimg-berlin-lahlem.mpg.de/>; Leser et al. 1999).

Deletion–Hybrid Interval Mapping

The HyTM1 hamster–human hybrids 79-2, 33, 54, 57, and 73 carry X chromosomes with nested terminal deletions of Xq whose breakpoints lie between the *HPRT* locus in Xq26.1 and *COL4A5* in Xq22.3 (Farr et al. 1992). The other deletion cell lines used were X3000-11.1 (Xq24–qter; Nussbaum et al. 1986), GM10664 [94-3] (Xq25–qter; Ledbetter et al. 1991), F649-5 (Xq25–qter; Reilly et al. 1988), and 85D30T2 (Xpter–Xq26), 87z4 (Xq26–Xqter), and HY129E (Xpter–q11:Xq26–qter) (all Forabosco et al. 1992).

The landmark content of each hybrid was assessed using the PCR, in a reaction containing 50 ng of DNA, 0.5 mM of each dNTP, 67 mM Tris-HCl (pH 8.8), 16.7 mM $(\text{NH}_4)_2\text{SO}_4$, 6.7 mM MgCl_2 , 1.7 $\mu\text{g/ml}$ of BSA, 10 mM 2-mercaptoethanol, 10 ng/ μl of each primer, and 2 units of *Taq* polymerase (Perkin-Elmer). Amplification was carried out for 35 cycles of 93°C for

30 sec, the annealing temperature for 30 sec, and 72°C for 30 sec, with an initial denaturation of 5 min at 94°C and a final extension of 5 min at 72°C. Products were analyzed by electrophoresis on a 2.5% agarose gel. Positive control DNAs were from human male placenta (Sigma) and from the hybrid Clone-2D that contains the entire X chromosome (Goss and Harris 1977). Mouse and hamster DNAs were used as negative controls.

YAC Library Screening

The YAC libraries screened were the ICI library (Anand et al. 1990), the ICRF library (Larin et al. 1991), and the CEPH ("mega") library (Chumakov et al. 1995). The yWXD clones were kindly provided by D. Schlessinger (see <http://www.ibt.wustl.edu/cgm/jcgm.html>). The RS clones were kindly provided by D. Nelson (Nelson et al. 1991), as were YACs yA29B10 and yB186B4, which are from the Washington University YAC libraries (Brownstein et al. 1989). The six other clones from the Washington University libraries were identified in the present study.

Screening of YAC libraries was by the hybridization of landmark probes (STS PCR products, YAC end probes, PAC end probes, etc.) to high-density clone arrays (Bentley et al. 1992). Hybridization membranes were prepared from entire libraries, from selected plates, or from selected clones (the X polygrid; see below). For STS markers, YAC end probes, and PAC end probes, PCR products were excised from 2.5% agarose gels and stored at 4°C in 10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA ($T_{0.1E}$). Five microliters of the $T_{0.1E}$ buffer was taken for radiolabeling by reamplification for 20 cycles in the presence of 0.22 μ M [α - 32 P]dCTP (3000 Ci/mmol). All other probes were labeled using the random hexamer priming method of Feinberg and Vogelstein (1983). Probes were denatured in 5 \times SSC buffer with 2.5 mg/ml sonicated human placental DNA (Sigma) and then were added to the hybridization solution (6 \times SSC, 10 \times Denhardt's, 50 mM Tris-HCl at pH 7.4, 10% dextran sulfate, 1% *N*-lauroyl sarcosine). Hybridization was performed overnight at 65°C, and then filters were rinsed twice in 2 \times SSC and washed twice for 30 min in 0.5 \times SSC, 1% *N*-lauroyl sarcosine at 65°C. Autoradiography was performed using Kodak X-OMAT S film at -70°C using two intensifying screens.

For many STS markers, the scale of the hybridization experiment was minimized by using the PCR to locate microtiter plates containing positive clones. DNA pools containing clones from 16 microtiter plates and from individual plates were prepared as described by Bentley et al. (1992). Positive plates were identified by screening first the superpools then the appropriate single-plate pools using the PCR conditions described above for deletion-hybrid screening. The STS PCR products were then hybridized to colony arrays prepared from the positive microtiter plates to identify individual clones.

Following initial detection, PCR testing of colony material was used to confirm the presence of STSs detected by hybridization, where applicable.

Computer Programs and Contig Assembly

An ACeDB database was established for the storage and visualization of the data described (Durbin and Thierry-Mieg 1991). This database (Xace) contains all of the Sanger Centre X chromosome mapping data (G. Maslen and C. Scott, unpubl.) and can be accessed using the procedures detailed at <http://www.sanger.ac.uk/HGP/ChrX>.

For contig assembly, the landmark-YAC information was

analysed using the program SAM (Soderlund and Dunham 1995). The information in Xace is automatically used to update the SAM marker file daily. SAM uses a rapid stochastic assembly algorithm to generate one or more plausible marker orders, maximizing the number of consecutive markers in each clone. The solution obtained with a specified set of markers is displayed graphically (see Fig. 4), and the program allows for manual manipulation of marker order when there are other data that recommend such alterations.

ACeDB and SAM are available via the Sanger Centre's World Wide Web site at <http://www.sanger.ac.uk/>.

Fluorescence in Situ Hybridization

DNA from PAC clones identified by STS screening (see Mungall et al. 1996) was prepared using a standard alkaline-lysis protocol. DNA was labeled with biotin-16-dUTP or digoxigenin-11-dUTP (Boehringer Mannheim) by nick translation. Labeled probes were hybridized to DAPI-banded metaphase spreads prepared from a normal human female lymphoblastoid cell line and to interphase nuclei prepared from a normal human male fibroblast cell line, as previously described (Leversha 1997). Hybridization signals were detected with successive incubations of 4–8 μ g/ml of avidin-Texas Red (Molecular Probes), followed by 4 μ g/ml of biotinylated anti-avidin (Vector) together with a 1:500 dilution of mouse-anti-digoxigenin FITC (Sigma) and finally, 4–8 μ g/ml of avidin-Texas Red together with 8 μ g/ml of goat-anti-mouse Cy2 (Molecular Probes).

For contig ordering and orientation, multiple metaphase chromosome spreads were scored for probe order relative to the centromere, or multiple interphase nuclei were scored for order relative to a third anchored clone. For the estimation of gap sizes, multiple interphase nuclei were analyzed to compare the relative separations of a control clone pair and a test pair.

Pulsed-Field Gel Electrophoresis and Southern Hybridization

Human female genomic DNA (6 μ g) in agarose blocks was digested using 20 units of *Bss*HII, *Eag*I, *Nae*I, *Not*I, *Sac*II, *Sma*I, *Sfi*I, *Sal*I, *Mlu*I, *Pvu*I, or *Nru*I for 3 hr; then, fragments were separated for 24 hr at 6 V/cm on a 1% agarose gel in 0.5 \times TBE buffer, using a Bio-Rad CHEF-DRIII electrophoresis system with a ramped switching time of 50–90 sec. Gels were treated with 0.25 M HCl for 2 \times 15 min, then denaturing solution (1.5 M NaCl, 0.5 N NaOH) for 2 \times 30 min, and then were blotted onto Hybond N+ (Amersham) using 0.4 N NaOH. Following neutralization of the membrane in 0.2 M Tris-Cl (pH 7.5), 2 \times SSC, blots were hybridized with labeled STS PCR products as described above for YAC colony filters.

ACKNOWLEDGMENTS

We thank Rakesh Anand, Tony Monaco, Hans Lehrach, Denis Le Paslier, and Daniel Cohen for YAC libraries; Pieter de Jong and Joseph Catanese for the RPCI PAC and BAC libraries; Nigel Carter for advice and discussion on FISH; Danita Pearson and Jennie Conquer for library screening reagents; Christine Farr, Antonino Forabosco, and Robert Nussbaum for DNA from deletion-hybrid cell lines; David Schlessinger and David Nelson for YAC clones; Kate Rice for design of primers; and Daniela Toniolo, Giovanni Porta, and Bakary Sylla for landmark information and reagents. This work was supported by an award from the Wellcome Trust.

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

- Anand, R., J.H. Riley, R. Butler, J.C. Smith, and A.F. Markham. 1990. A 3.5 genome equivalent multi access YAC library: Construction, characterisation, screening and storage. *Nucleic Acids Res.* **18**: 1951-1956.
- Bentley, D.R., C. Todd, J. Collins, J. Holland, I. Dunham, S. Hassock, A. Bankier, and F. Gianelli. 1992. The development and application of automated gridding for efficient screening of yeast and bacterial ordered libraries. *Genomics* **12**: 534-541.
- Bentley, D.R., K.D. Pruitt, P. Deloukas, G.D. Schuler, and J. Ostell. 1998. Coordination of human genome sequencing via a consensus framework map. *Trends Genet.* **14**: 381-384.
- Brownstein, B.H., G.A. Silverman, R.D. Little, D.T. Burke, S.J. Korsmeyer, D. Schlessinger, and M.V. Olson. 1989. Isolation of single-copy human genes from a library of yeast artificial chromosome clones. *Science* **244**: 1348-1351.
- Buetow, K.H., J.L. Weber, S. Ludwigsen, T. Scherpbier-Heddema, G.M. Duyk, V.C. Sheffield, Z. Wang, and J.C. Murray. 1994. Integrated human genome-wide maps constructed using the CEPH reference panel. *Nat. Genet.* **6**: 391-393.
- Choi, C., S. Cho, I. Horikawa, A. Berchuk, N. Wang, E. Cedrone, S.W. Jhung, J.B. Lee, J. Kerr, G. Chenevix-Trench et al. 1997. Loss of heterozygosity at chromosome segment Xq25-26.1 in advanced human ovarian carcinomas. *Genes, Chromosomes, Cancer* **20**: 234-242.
- Chumakov, I.M., P. Rigault, I. Le Gall, C. Bellané-Chantelot, A. Billault, S. Guillou, P. Soularue, G. Guasconi, E. Poullier, I. Gros et al. 1995. A YAC contig map of the human genome. *Nature (Suppl.)* **377**: 175-297.
- Coffey, A.J., R.G. Roberts, E.D. Green, C.G. Cole, R. Butler, R. Anand, F. Gianelli, and D.R. Bentley. 1992. Construction of a 2.6-Mb contig in yeast artificial chromosomes spanning the human dystrophin gene using an STS-based approach. *Genomics* **12**: 474-484.
- Coffey, A.J., R.A. Brooksbank, O. Brandau, T. Oohashi, G.R. Howell, J.M. Bye, A.P. Cahn, J. Durham, P. Heath, P. Wray et al. 1998. Host response to EBV infection in X-linked lymphoproliferative disease results from mutations in an SH2-domain encoding gene. *Nat. Genet.* **20**: 129-135.
- Cole, C.G., P.N. Goodfellow, M. Bobrow, and D.R. Bentley. 1991. Generation of novel sequence tagged sites (STSs) from discrete chromosomal regions using *Alu*-PCR. *Genomics* **10**: 816-826.
- Cole, C.G., I. Dunham, A.J. Coffey, M.T. Ross, S. Meier-Ewert, M. Bobrow, and D.R. Bentley. 1992. A random STS strategy for construction of YAC contigs spanning defined chromosomal regions. *Genomics* **14**: 256-262.
- Collins, J.E., C.G. Cole, L.J. Smink, C.L. Garrett, M.A. Leversha, C.A. Soderlund, G.L. Maslen, L.A. Everett, K.M. Rice, A.J. Coffey et al. 1995. A high-density YAC contig map of human chromosome 22. *Nature (Suppl.)* **377**: 367-379.
- Coulson, A., Y. Kozono, R. Shownkeen, and R. Waterston. 1991. The isolation of insert-terminal YAC fragments by genomic sequencing. *Technique* **3**: 17-23.
- Deloukas, P., G.D. Schuler, G. Gyapay, E.M. Beasley, C. Soderlund, P. Rodriguez-Tomé, L. Hui, T.C. Matisse, K.B. McKusick, J.S. Beckmann et al. 1998. A physical map of 30,000 human genes. *Science* **282**: 744-746.
- des Portes, V., N. Soufir, A. Carrie, P. Billuart, T. Bienvenu, M.C. Vinet, C. Beldjord, G. Ponsot, A. Kahn, J. Boue, and J. Chelly. 1997. Gene for nonspecific X-linked mental retardation (*MRX47*) is located in Xq22.3-q24. *Am. J. Med. Genet.* **72**: 324-328.
- Dib, C., S. Fauré, C. Fizames, D. Samson, N. Drouot, A. Vignal, P. Millasseau, S. Marc, J. Hazan, E. Seboun et al. 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**: 152-154.
- Durbin, R. and J. Thierry-Mieg. 1991. A *C. elegans* database; documentation, code and data available from anonymous ftp servers (lirmm.lirmm.fr, cele.mrc-lmb.cam.ac.uk, and ncbi.nlm.nih.gov).
- Farr, C.J., M. Stevanovic, E.J. Thomson, P.N. Goodfellow, and H.J. Cooke. 1992. Telomere-associated chromosome fragmentation: Applications in genome manipulation and analysis. *Nat. Genet.* **2**: 275-282.
- Fasman, K.H., S.I. Letovsky, R.W. Cottingham, and D.T. Kingsbury. 1997. The GDB(TM) Human Genome Database Anno 1997. *Nucleic Acids Res.* **25**: 72-80.
- Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6-13.
- Figuera, L.E., M. Pandolfo, P.W. Dunne, J.M. Cantú, and P.I. Patel. 1995. Mapping of the congenital generalized hypertrichosis locus to chromosome Xq24-q27.1. *Nat. Genet.* **10**: 202-207.
- Forabosco, A., S. Fini, A. Ferlini, and M. Rocchi. 1992. Panel of somatic cell hybrids with breakpoints in the Xq25-Xq26 boundary. In *3rd X chromosome workshop* (ed. IIGB Press), p. 80. International Institute of Genetics and Biophysics, Naples, Italy.
- Gedeon, A.K., I.A. Glass, J.M. Connor, and J.C. Mulley. 1996. Genetic localization of *MRX27* to Xq24-26 defines another discrete gene for non-specific X-linked mental retardation. *Am. J. Med. Genet.* **64**: 121-124.
- Goss, S.J. and H. Harris. 1977. Gene transfer by means of cell fusion. *J. Cell Sci.* **25**: 17-37.
- Gregg, R.G., C. Palmer, S. Kirkpatrick, and A. Simantel. 1996. Localization of a non-specific X-linked mental retardation gene, *MRX23*, to Xq23-q24. *Hum. Mol. Genet.* **5**: 411-414.
- Gu, X.X., R. Decorte, P. Marynen, J.-P. Fryns, J.-J. Cassiman, and P. Raeymaekers. 1996. Localization of a new gene for non-specific mental retardation to Xq22-q26 (*MRX35*). *J. Med. Genet.* **33**: 52-55.
- Hudson, T.J., L.D. Stein, S.S. Gerety, J. Ma, A.B. Castle, J. Silva, D.K. Slonim, R. Baptista, L. Kruglyak, S.-H. Xu et al. 1995. An STS-based map of the human genome. *Science* **270**: 1945-1954.
- Illarioshkin, S.N., H. Tanaka, E.D. Markova, N.N. Nikolskaya, I.A. Ivanova-Smolenskaya, and S. Tsuji. 1996. X-linked nonprogressive congenital cerebellar hypoplasia: Clinical description and mapping to chromosome Xq. *Ann. Neurol.* **40**: 75-83.
- Ioannou, P.A., C.T. Amemiya, J. Garnes, P.M. Kroisel, H. Shizuya, C. Chen, M.A. Batzer, and P.J. de Jong. 1994. A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.* **6**: 84-89.
- Kendall, E., W. Evans, H. Jin, J. Holland, and D. Vetrie. 1997. A complete YAC contig and cosmid interval map covering the entirety of human Xq21.33 to Xq22.3 from DXS3 to DXS287. *Genomics* **43**: 171-182.
- Ku, D.-H., J. Kagan, S.-T. Chen, C.-D. Chang, R. Baserga, and J. Wurzel. 1990. The human fibroblast adenine nucleotide transporter gene. *J. Biol. Chem.* **265**: 16060-16063.
- Larin, Z., A.P. Monaco, and H. Lehrach. 1991. Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc. Natl. Acad. Sci.* **88**: 4123-4127.
- Ledbetter, S.A., C.E. Schwarz, K.E. Davies, and D.H. Ledbetter. 1991. New somatic cell hybrids for physical mapping in distal Xq and the fragile X region. *Am. J. Med. Genet.* **38**: 418-420.
- Leser, U., H. Roest Crolius, H. Lehrach, and R. Sudbrak. 1999. IXDB, an X chromosome integrated database (update). *Nucleic Acids Res.* **27**: 123-127.
- Leversha, M.A. 1997. Fluorescence *in situ* hybridization. In *Genome mapping: A practical approach* (ed. P.H. Dear), pp. 199-225. Oxford University Press, Oxford, UK.
- Lincoln, S.E., M.J. Daly, and E.S. Lander. 1991. PRIMER: A computer program for automatically selecting PCR primers (version 0.5) (<http://www.genome.wi.mit.edu/>).
- Liston, P. N. Roy, K. Tamai, C. Lefebvre, S. Baird, G. Cherton-Horvat, R. Farahani, M. McLean, J.-E. Ikeda, A. MacKenzie, and R. Korneluk. 1996. Suppression of apoptosis in mammalian cells by

- NAIP and a related family of IAP genes. *Nature* **379**: 349–353.
- Mattei, M.-G., J. Matterson, J.W. Chen, M.A. Williams, and M. Fukuda. 1990. Two human lysosomal membrane glycoproteins, h-lamp-1 and h-lamp-2, are encoded by genes localized to chromosome 13q34 and chromosome Xq24-25, respectively. *J. Biol. Chem.* **265**: 7548–7551.
- Mungall, A.J., C.A. Edwards, S.A. Ranby, S.J. Humphray, R.W. Heathcote, C.M. Clee, C.L. East, E. Holloway, A.P. Butler, C.F. Langford et al. 1996. Physical mapping of chromosome 6: A strategy for the rapid generation of sequence-ready contigs. *DNA Sequence* **7**: 47–49.
- Murray, J.C., K.H. Buetow, J.L. Weber, S. Ludwigsen, T. Scherpbier-Heddema, F. Manion, J. Quillen, V.C. Sheffield, S. Sunden, G.M. Duyk et al. 1994. A comprehensive human linkage map with centiMorgan density. *Science* **265**: 2049–2054.
- Nagaraja, R., S. MacMillan, J. Kere, C. Jones, S. Griffin, M. Schmatz, J. Terrell, M. Shomaker, C. Jermak, C. Hott et al. 1997. X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. *Genome Res.* **7**: 210–222.
- Nagaraja, R., S. MacMillan, C. Jones, M. Masisi, G. Pengue, G. Porta, S. Miao, A. Casamassimi, M. D'Urso, B. Brownstein, and D. Schlessinger. 1998. Integrated YAC/STS physical and genetic map of 22.5 Mb of human Xq24-q26 at 56-kb inter-STS resolution. *Genomics* **52**: 247–266.
- Nelson, D.L., A. Ballabio, M.F. Victoria, M. Pieretti, R.D. Bies, R.A. Gibbs, J.A. Maley, A.C. Chinault, T.D. Webster, and C.T. Caskey. 1991. *Alu*-primed polymerase chain reaction for regional assignment of 110 yeast artificial chromosome clones from the human X chromosome: Identification of clones associated with a disease locus. *Proc. Natl. Acad. Sci.* **88**: 6157–6161.
- Nussbaum, R.L., S.D. Airhart, and D.H. Ledbetter. 1986. A rodent-human hybrid containing Xq24-Xqter translocated to hamster chromosome expresses the Xq27 folate-sensitive fragile site. *Am. J. Med. Genet.* **23**: 457–466.
- Pai, G.S., J.A. Sprenkle, T.T. Do, C.E. Mareni, and B.R. Migeon. 1980. Localization of loci for hypoxanthine phosphoribosyltransferase and glucose-6-phosphate dehydrogenase and biochemical evidence of nonrandom X chromosome expression from studies of a human X-autosome translocation. *Proc. Natl. Acad. Sci.* **77**: 2810–2813.
- Pilia, G., S. MacMillan, R. Nagaraja, S. Mumm, J. Weissenbach, and D. Schlessinger. 1996. YAC/STS map of 9 Mb of Xq26 at 100-kb resolution, localizing 6 ESTs, 6 genes, and 32 genetic markers. *Genomics* **34**: 55–62.
- Priest, J.M., K.H. Fishbeck, N. Nouri, and B.J. Keats. 1995. A locus for axonal motor-sensory neuropathy with deafness and mental retardation maps to Xq24-q26. *Genomics* **29**: 409–412.
- Purtillo, D.T., C.K. Cassel, J.P.S. Yang, R. Harper, S.R. Stephenson, B.H. Landing, and G.F. Vewter. 1975. X-linked recessive progressive combined variable immunodeficiency (Duncan's disease). *Lancet* **1**: 935–941.
- Reilly, D.S., R.A. Lewis, D.H. Ledbetter, and R.L. Nussbaum. 1988. Tightly linked flanking markers for the Lowe oculocerebrorenal syndrome with application to carrier assessment. *Am. J. Hum. Genet.* **42**: 748–755.
- Riley, J., R. Butler, D. Ogilvie, R. Finniear, D. Jenner, S. Powell, R. Anand, J.C. Smith, and A.F. Markham. 1990. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* **18**: 2887–2890.
- Schiebel, K., A. Mertz, M. Winkelmann, R. Nagaraja, and G. Rappold. 1994. Localization of the adenine nucleotide transporter gene *ANT2* to chromosome Xq24-q25 with tight linkage to DXS425. *Genomics* **24**: 605–606.
- Schuler, G.D., M.S. Boguski, E.A. Stewart, L. Stein, G. Gyapay, K. Rice, R.E. White, P. Rodriguez-Tomé, A. Aggarwal, E. Bajorek et al. 1996. A gene map of the human genome. *Science* **274**: 540–546.
- Shizuya, H., B. Birren, U.-J. Kim, V. Mancino, T. Slepak, Y. Tachiiri, and M. Simon. 1992. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci.* **89**: 8794–8797.
- Silver, D.N., R.A. Lewis, and R.L. Nussbaum. 1987. Mapping the Lowe oculocerebrorenal syndrome to Xq24-q26 by use of restriction fragment length polymorphisms. *J. Clin. Invest.* **79**: 282–285.
- Soderlund, C. and I. Dunham. 1995. SAM: A system for iteratively building marker maps. *Comp. Appl. Biosci.* **11**: 645–655.
- Stewart, E.A., K.B. McKusick, A. Aggarwal, E. Bajorek, S. Brady, A. Chu, N. Fang, D. Hadley, M. Harris, S. Hussain et al. 1997. An STS-based radiation hybrid map of the human genome. *Genome Res.* **7**: 422–433.
- Yntema, H.G., B.C. Hamel, A.P. Smits, T. van Roosmalen, B. van de Helm, H. Kremer, H.H. Ropers, D.F. Smeets, and H. van Bokhoven. 1998. Localisation of a gene for non-specific X-linked mental retardation (*MRX46*) to Xq25-q26. *J. Med. Genet.* **35**: 801–805.
- Zori, R.T., J.L. Gardner, J. Zhang, M.J. Mullan, R. Shah, A.R. Osborn, H. Houlden, M.R. Wallace, S. Roberts, and T.P. Yang. 1998. Newly described form of X-linked arthrogryposis maps to the long arm of the human X chromosome. *Am. J. Med. Genet.* **78**: 450–454.

Received September 16, 1998; accepted in revised form June 7, 1999.



High-Resolution Landmark Framework for the Sequence-Ready Mapping of Xq23–q26.1

Helen E. Steingruber, Andrew Dunham, Alison J. Coffey, et al.

Genome Res. 1999 9: 751-762

Access the most recent version at doi:[10.1101/gr.9.8.751](https://doi.org/10.1101/gr.9.8.751)

References This article cites 52 articles, 16 of which can be accessed free at:
<http://genome.cshlp.org/content/9/8/751.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genome Research* go to:
<http://genome.cshlp.org/subscriptions>