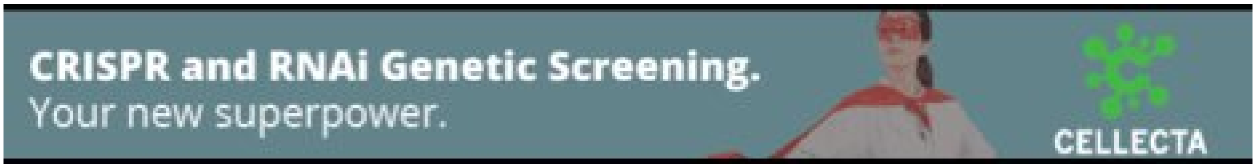
4.5-Mb YAC STS contig at 50-kb resolution, spanning Xq25 deletions in two patients with lymphoproliferative syndrome.

G Porta, S MacMillan, R Nagaraja, et al.

Genome Res. 1997 7: 27-36
Access the most recent version at doi:10.1101/gr.7.1.27



RESEARCH

4.5-Mb YAC STS Contig at 50-kb Resolution, Spanning Xq25 Deletions in Two Patients with Lymphoproliferative Syndrome

Giovanni Porta,[1,2] Sandra MacMillan,[1] Ramaiah Nagaraja,[1] Steven Mumm,[1] Ileana Zucchi,[1,3] Giuseppe Pilia,[1] Shirong Maio,[1] Terence Featherstone,[1] and David Schlessinger[1,4]

[1]Department of Molecular Microbiology and Center for Genetics in Medicine, Washington University School of Medicine, St. Louis, Missouri 63110; [2]Department of Human Genetics, II Facolta di Medicina, Pavia, Italy; [3]Istituto Technologie Biomedicale Avanziata, 20131 Milano, Italy

---

Sequence-tagged site (STS) content mapping in yeast artificial chromosomes (YACs) was used to cover the region deleted in two patients affected with X-linked lymphoproliferative disorder. The order of markers includes, centromere to telomere, DXS8009–DXS1206–DXS8078–DXS8044–DXS982–DXS6811–DXS8093–AFM240xb10–DXS75–DXS737–DXS100–DXS6–DXS1046–DXS8038. The order of six major markers is confirmed by fluorescent in situ hybridization, and all the markers assigned by linkage mapping fall within a 1.6-cM interval. The contig comprises 90 clones containing 89 STSs, yielding a resolution of 50 kb; DNA in a gap just telomeric to DXS8044 has not been found in >20 equivalents of YACs or bacterial clones. The two deletions were found to have centromeric breakpoints that lie close to DXS1206 and may be identical; the telomeric breakpoints are ~150 kb apart, one falling between DXS737 and DXS100, the other between DXS100 and DXS1046. Several STSs near the breakpoints show weak amplification from more than one site; one gives products from three groups of YACs, and lie, respectively, within 50 kb of the centromeric and the two telomeric deletion borders. Such partially duplicated segments of DNA are candidates for involvement in the formation of the deletions.

[A hyperlinked version of Table 1 is available at http://www.cshl.org/gr.]

Among the group of X-linked immunodeficiencies, lymphoproliferative disorder (LYP; MIM308240) is unique in causing a specific defect in the T cell-mediated response to Epstein–Barr virus (Bare et al. 1974; Purtilo et al. 1977). As a result, patients progress to disease that varies from agammaglobulinemia to the development of lymphomas. The pathogenesis of the disease involves massive tissue destruction resulting from infiltration of activated T cells and macrophages. Extranodal lymphoma has been observed in many patients, and has been linked to a complex interaction with the mechanism of apoptosis. In fact, in a mouse model for one form of lymphoproliferative disease, the cell surface antigen *fas*, known to be involved in lymphocyte apoptosis, is defective in both T and B cells infected with Epstein–Barr virus (Ogasawara et al. 1993).

The severe cytotoxic T-cell response in LYP has been ascribed by different groups to polyclonally activated alloreactive cytotoxic T cells or to unregulated autoimmune killer and natural killer cell activity against both EBV-infected and -uninfected cells (Sullivan and Woda 1989). These EBV-provoked cells show cytotoxicity not only to EBV-infected B cells, but also to fibroblast and hepatocyte target cells, so that some host component is also attacked; but recent reviews (Sullivan and Woda 1989; Conley 1992) have pointed out that traditional approaches to pathophysiology have left unclear the nature of the immunological crippling by LYP, the basis of hepatic damage and liver failure, and the reason for variable responses. In contrast, direct study of the gene could reveal the causative factor and the nature and extent of relevant interactions (Seemayer et al. 1995).

[4]**Corresponding author.**
E-MAIL davids@sequencer.wustl.edu; FAX 314-362-3203.

7:27–36 ©1997 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/97 $5.00

Positional cloning to find the disease gene has seemed more feasible since a number of linkage mapping studies and analysis of a large deletion restricted LYP to Xq25 (Wyandt et al. 1989). The candidate region was further refined when two affected boys from different families showed smaller deletions (Skare et al. 1993), detected by the absence of Xq25-specific probes. Wu et al. (1993) ordered several of the probes and suggested from in situ hybridization that the deletions were ~2 Mb in extent.

We have constructed a yeast artificial chromosome (YAC) sequence-tagged site (STS) contig that spans the two deletions reported (Wu et al. 1993). The STS content mapping results (1) reorder DXS6 and DXS100 and provide long-range cloned coverage of the region, which proves to be ~3 Mb; (2) place and order a number of additional markers that detect polymorphism and could narrow the localization of the disease gene; and (3) hint at a possible origin for the deletion formation.

RESULTS

Construction of the Map in Relation to Deletion Borders

The current map of the region is shown in Figure 1. Mapping began with the recovery of cognate YACs for a set of probes (DXS100, DXS739, DXS75, DXS6, and DXS982) that had been shown to lie in or near the LYP locus in previous studies (Bare et al. 1974; Wu et al. 1993; note that other probes that had been shown to be linked to LYP, including DXS42, were excluded from the deletion span and were not specifically studied further to develop this contig). Some overlaps among the initial set of YACs, determined by STS content, provided coverage of about one-third the entire region. Mapping then continued by repeated cycles of end-walking (Kere et al. 1992), with the reiterative synthesis of new STSs from insert ends of YACs at the borders of nascent contigs. The STSs were also supplemented with new polymorphic markers and some additional probes in the region as they became available from the community or from screening efforts in the Washington University Genome Center. In every case the STSs were tested against the two deletion DNAs to verify that mapping was continuing in the target zone, and to permit the recognition of the borders of the deletion regions when they were crossed by YACs.

Figure 1 contains 85 YACs, selected for quality and coverage from a total of 140 recovered from six YAC libraries; five PACs for two loci are also included. The map is formatted with 89 STSs (Table 1) across ~4.5 Mb of DNA. This brings the resolution of ordered, relatively well-spaced markers to ~50 kb. We note that the STSs include seven dinucleotide repeat markers on the Centre d'Etude du Polymorphisme Humain (CEPH) Genethon linkage map and a tetranucleotide repeat marker from the Cooperative Human Linkage Center (CHLC).

The internal consistency of STS content in the

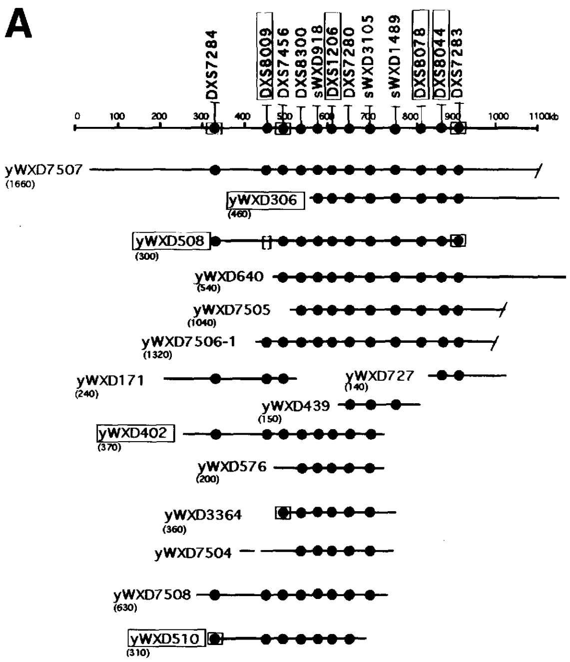


Figure 1 The map was inferred by the SEGMAP program (see Methods). A gap intervenes between the shorter 1-Mb centromeric portion (*A*) and the 4-Mb telomeric portion (*B*), with the two segments drawn in slightly different formats for clarity (see text). In a few instances, an STS was apparently deleted in a YAC, and is indicated by a vertical bracket. YACs are drawn proportional to their DNA content. STS content is indicated by a filled circle, or a square that indicates the origin of an STS from a YAC insert end. The single heavy bar under the distance scale shows a point at which only a single clone bridges two regions of the contig. Dashed YACs are larger in size than their STS content indicates, suggesting possible chimerism. The size of each YAC (and the calculated size for dashed YACs in kb) are given in parentheses at the *left* of each represented clone. STSs are indicated by DXS nos. where available, and by sWXD accession nos. for those still being processed at the Genome Database; YACs are indicated by yWXD nos., and PACs by pWXD nos. (see Methods and Table 1). Boxed YAC names indicated clones placed in the region by FISH. Boxed STSs (*top* scale) are Genethon markers that detect polymorphism.

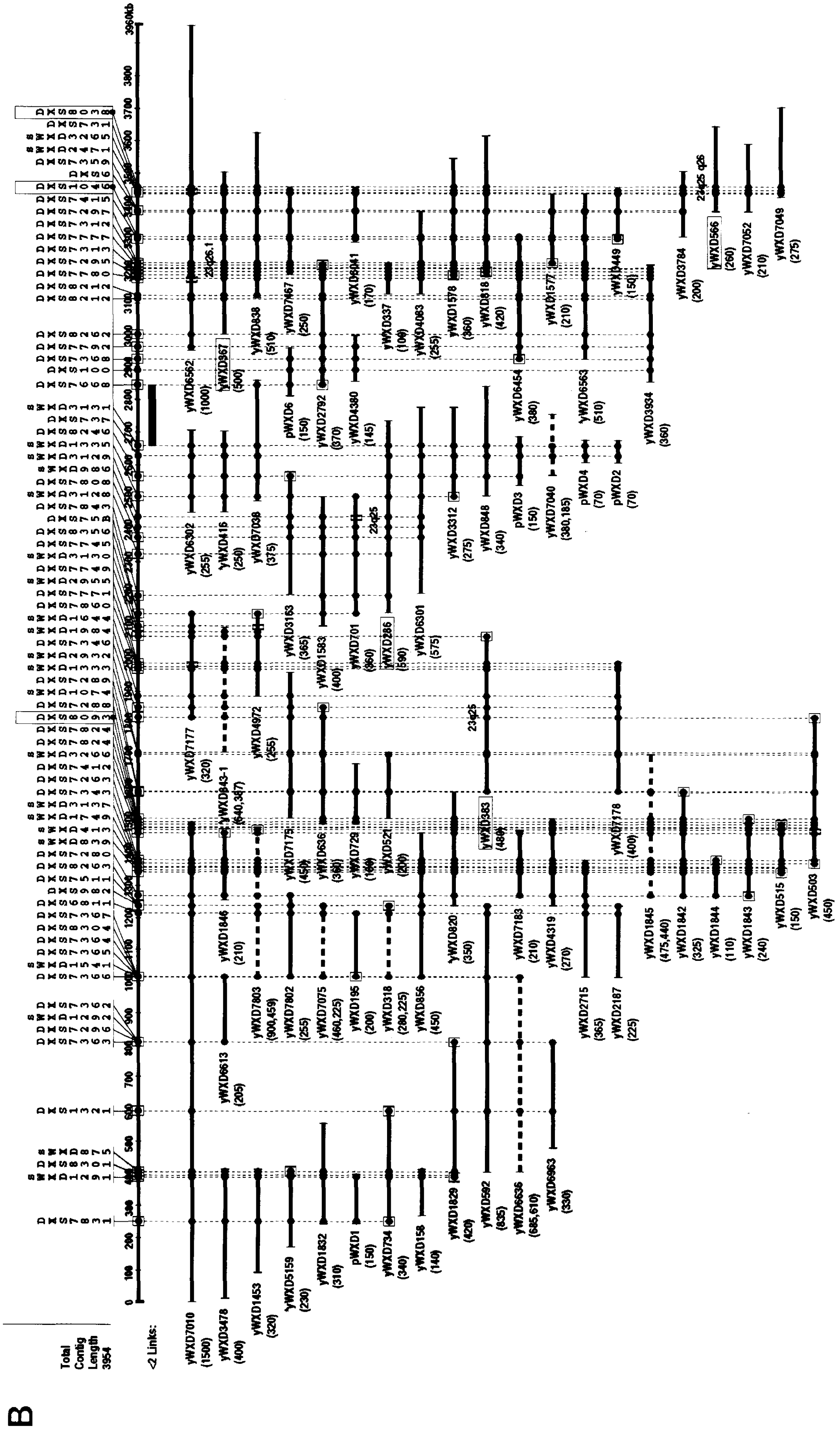


Figure 1 *(See facing page for legend.)*

Table 1. STSs Used to Format the YAC Contig

STS Name	DXS Number	Alias 1	Size (bp)	Buffer	Primer A	Primer B	GenBank Acc	GDB ID
sWXD43	DXS100	DXS100C-168	168	TNK100	GCTTAGAGTGCTAGTTTAGG	AGAGAATGGGGACTTTAAGAGG	L14992	G00-192-511
sWXD174	DXS1321	734L	204	TNK50	CCACTGCCTTTAGAAAAACAC	ATTCACTGAGAACAAGACAAAAAC	L15227	G00-235-865
sWXD201	DXS7456	F5A1L	64	TNK100	CATATCATTCCACTGCCTTC	AAAGAAATTATCAGTCACCTTCTC	L24595	G00-584-187
sWXD269	DXS7808	XY788R	149	TNK100	CTTTGGGAGATGTTGGTTTG	AGATATTATGCCATTAAGCAGGAG	L24576	G00-600-706
sWXD299	DXS737	DXS737	167	TNK100	GAGGAAAAAATCAAGTCCATGT	TGTCTAGTGTATGACTCATCTCCA	M87768	G00-182-659
sWXD397	DXS7824	XSTS3240	189	TNK100	CATAAATCTCAACTTTGGTTTCTC	TGGCTCCAGGAAAATATGC	L24619	G00-600-763
sWXD413		XSTS5118	78	TNK50	CTCACTGTTGCTGTAATACTTC	CATTATTGATAGGGGGAGG	L24635	
sWXD416	DXS8212	F11C4R	104	TNK100	TTTATCTTCAAACTATGTATGCCC	ATTGTATCTGAGATTACTGTTGTG	L24602	G00-637-708
sWXD443	DXS7831	734R	195	TNK100	TTAAGTCAGTTTTTCCAAAAGCAC	ATCTTATTTGATCCTTGCATTTCC	L25167	G00-600-784
sWXD525	DXS7843	F1F4L	97	TNK100	CATGCACTGAGTGTTATGG	CGGGTTCATCCTAATCAAG	L24603	G00-600-820
sWXD585	DXS6	DXS6	151	TNK100	TCAAAGTGATGGCAGATTCAGTG	GGAAATGTTATTAGGATAGGGG		
sWXD596	DXS75B	DXS75B	199	TNK100	AATCATATCAAGTGTCTTTTCTGG	ATAAGTTTTTGGCATGTTCTGTTTC		
sWXD620	DXS7860	F20F11R	146	TNK100	ATTCCCCTCTCCCCCATTG	CCCCAGAACCTTGTAAGAATC	L24593	G00-600-877
sWXD839		515R	82	TNK100	TTTTGAAAGCATTCTATGGCAATC	ATGAAGTATTACAGCAACAGTGAG	L24614	1238759
sWXD862	DXS7412	I97C12R	82	TNK100	AGTGAACTATAACCTAACTTGATG	GTCCTCTATTGGCTAAAACTC	L41322	G00-556-512
sWXD868	DXS7415	449L	80	TNK100	AGAATATCAGCAAAGATAGAGAAG	TATGTATTTTTCTGTTGTTGGGTG	L41210	G00-556-521
sWXD875		F24C3L	108	TNK50	CACAACAAAACACTAGACATTTCC	CTCCCTTAAATTGCCTATATTTCC	L78123	
sWXD914	DXS1206	AFM269ya5	167	TNK100	CACAATCTGCTGTCTGCAAT	AGCATGGGACTTCTCAACC	Z23914	G00-199-640
sWXD918		E143H6R	122	TNK100	CAATATGGCATCAAGTCATGTTTC	TCCTAATGGCTTAGTGGTC	L24876	1296865
sWXD921	DXS8212	367R	61	TNK100	TTTATCTTCAAACTATGTATGCCC	ATGAAACATTCTTGACAAACTCAG	L77994	
sWXD986		S7.043-61	61	TNK100	ATCAGGCAGAAATTTGGGGG	GACTCTGATGCACTCTTCAG	L42649	1238560
sWXD1001	DXS7263	I246G6L	91	TNK100	GTGGAGGAGAGGTTTATATTTTG	GGAAGGCAATGATACCAAC	L41277	G00-556-065
sWXD1074		F20A8R	110	TNK50	GTTCGTGGCTTACAGAATC	CAGAGGCTCAGAGAGAATC	L47776	
sWXD1090	DXS7280	I182G11R	86	TNK100	CAAATAGAGCTTATTCTAGGTTGG	GGAAATGATTCTACAGCTTTTGGG	L41278	G00-556-116
sWXD1095	DXS7282	162L	61	TNK100	GAATTCAATAATTTTACTGTGTTTCTCAC	CACTGAGTTATCAGGCCAGCTG	L41187	G00-556-122
sWXD1096	DXS7283	508L	83	TNK100	ATCAAGGCATAGTTATCAAATGAG	CTATGACTGCTGTTGTAGG	L41188	G00-556-125
sWXD1097	DXS7284	510L	75	TNK100	CGATGCTGCACTTCTTTTGCTGTGTG	GATGTAAATACCTTAGTTCTTCTTAG	L41333	G00-556-128
sWXD1103	DXS7288	I182G11L	130	TNK100	TAAGAGAAGGAGGACTAGG	CTACACAAATGGTAAGTGACTGG	L41280	G00-556-140
sWXD1110	DXS7289	I65H7L	115	TNK100	CCTTGATTTCCTTTCTATTCTTTG	GAAGTGAACTGGATATTGGTG	L41281	G00-556-143
sWXD1291		I317A8L	99	TNK100	TGCTATCCCTCTGCTGATG	GCTTCTTGTTGGTGAAACTG	L77659	1296881
sWXD1292		I317A8R	90	TNK100	AGTGTAGAAAAAGAAGTAGGTAAG	CTTCGTCAAACTCTCACCC		
sWXD1333		XSTS4832	178	TNK100	ATGTTTTGGGGCATTGGGG	AAAGTGAAGGGTAGAGTAAATTTG	L77266	1296827
sWXD1335		F4C3L	120	TNK100	TGATTGTTTGACCCCACCC	ACAGCTCTGTGTATCAGACC	L47789	
sWXD1340	DXS8246	F2F9R	128	TNK100	AAGGGGGTATTGGTTGTAG	CCTATAAAGTCTTCCATAACATTG	L48036	G00-668-407
sWXD1350	DXS7295	E97H7L	136	TNK100	CACCCCTACATCTTCTAATC	GGCATTCACTAACCCATTAAC	L41192	G00-556-161
sWXD1351	DXS7296	831R	102	TNK100	GTTAGGAGGAGGAGAAATG	GTTTTGTGAATGAGAGTATAGAAG	L41193	G00-556-164
sWXD1376	DXS7297	E33H3R	77	TNK100	ATTCGTGTTCACTTCAGCC	AGATTGTAAATTGCACTTTGAACC	L41194	G00-556-167
sWXD1465	DXS8356	F22C4R	92	TNK100	TCTCCTCCTACAATCCATC	AACAACTTGGTGGTCTCTGTC	L48000	G00-681-639
sWXD1489		CF1a	233	TNK100	TCACGGTATTAGTAGATGCTGGTG	AGGTGTTATGTGGATTGTAGAGTG	D16468	1238713
sWXD1503	DXS7668	I216A12L	105	TNK100	GAATTCATTAGTTAAGGACTCAC	CCTAAAAGTCATTACCTTTAGCG	L41195	G00-600-283
sWXD1505	DXS7312	I216A12R	83	TNK100	AAAGGAGGAGAGGATGAGG	TTAGATAGTGTTTGAGCACAATTC	L41196	G00-556-212
sWXD1506	DXS7313	818R	105	TNK100	AAAGGTAAGTGTTTAGGTAATGAG	TAGCAGAAGAGTGGGAAAG	L41197	G00-556-215
sWXD1541		DF83	353	TNK100	AGGAATGCACCTACATATGTGGCA	CAGACTGCTTATTTATGGAGATCC	D16470	1238562
sWXD1610	DXS1046	AFM147xc7	143	TNK100	GGNCTTAGACACACGGTAGT	ATTTGTATTCACCTGGAGGC		G00-189-351
sWXD1611	DXS8009	AFMa119xf1	253	TNK100	AGCCTTCGTCCTCTATATATTTC	GCAATTCAAAAGATTCTGATTAATT		603146
sWXD1612	DXS8044	AFM203ze11	273	TNK100	CAAGGTTGAAAGGGCA	AAATGTAGCAGATGCCAAT		609462
sWXD1613	DXS8078	AFM273zc5	202	TNK100	TGCATCCCCATAGTAATTGGT	CAAATGGCAGGATTTCCC		612060

Table 1. *(Continued)*

STS Name	DXS Number	Alias 1	Size (bp)	Buffer	Primer A	Primer B	GenBank Acc	GDB ID
sWXD1616	DXS8093	AFMa045yc5	156	TNK100	AAGGGCAGGTTTTTTG	AGGGCTCTGAGGACTG		613491
sWXD1644		AFM240XB10	170	TNK50	CATGTAGAGCAAACAAACCCTA	TGCAGTATTTCCCCAGTGTATT	Z66856	1223323
sWXD1749		PF12	205	TNK100	GCAATGCCTTATATGAATGCAGTG	ATGATCAGGGTCACTTTCTGTGAG	D16477	1238715
sWXD1839	DXS7745	F33B3R	93	TNK100	GGGAGTATAATGAGTAGAAGTATC	TTGTTTCCTTTGCTGCACAAAATG	L42707	G00-600-514
sWXD1848	DXS7343	503L	87	TNK100	GTATCGAGGATCACATTTCAAC	CTCATCACAGTATCTGACATATAG	L41158	G00-556-305
sWXD1849	DXS7749	383L	80	TNK100	GGAAAGAAGGAAGTGAAGTTTTAC	CACATTTGCTTCATAGCACTC	L42710	G00-600-526
sWXD1914	DXS8038	AFMa288xd5	132	TNK100	GTGGACTGTCTCCGTAACC	CCAAGATGTGAGCATTTTTC		608337
sWXD1982	DXS7771	F33B3L	68	TNK100	GATGAGCAAATAATAGTCAAGTGG	CACTGGGTTATGTGCTTTTC	L42726	G00-600-592
sWXD1984		636R	360	TNK100	ATTTTCTTTGGGAATAATCACCTC	TCTCTTCTCCTTTTATACACTATC		
sWXD1985	DXS7346	503R	130	TNK100	GAATAAGAACCTTGTTGCTTTTAG	CAAACACTTATGTTGTGCTGG	L41108	G00-556-314
sWXD2011		515L	100	TNK100	TAGATGAATCTTAACTATGTCTGC	CTGGTTGGTGAAAACTAGG	L78122	
sWXD2060	DXS7359	F16A7R	134	TNK100	CAGAACTTCCTTCTTAATACAAAC	CTTTGGCAGCAAATATCCTC	L41200	G00-556-353
sWXD2062	DXS8262	I99F10R	63	TNK100	TCTCCCACTGAAGGATATG	CATAGCATGTTTATTTGTAATAGCC	L47482	G00-668-456
sWXD2063	DXS7777	M847B2L	84	TNK100	TCTTGAGCTGAGATCATCC	CTCAAATCTCTTTTTCTCTCCC	L41803	G00-600-610
sWXD2076	DXS7780	I208G4R	185	TNK100	GTTTATTTGTGCTGAGTTGCC	CTATCAGTAGATTGTAGAGCCC	L42730	G00-600-619
sWXD2145	DXS6811	GATA8F08	214	TNK100	CATTTTTCCAGTGCCAGAT	TGCTGGGAGGCTTTTTATTA		G00-366-394
sWXD2191	DXS7362	25L	107	TNK100	ACCACCCACAGAGAGCAAG	ACAATCCCCCATCTCCTCC	L41203	G00-556-362
sWXD2192	DXS7363	147L	80	TNK100	CCCCAAAACTCTTTGCAGAAC	ACCTAAGCCCCAAGACAAACC	L41204	G00-556-365
sWXD2193	DXS7364	195L	102	TNK100	GCTTAAATTCACGCATTATACTGG	ATTCTTTCCCTTTATCAGTTTTTGAG	L41205	G00-556-368
sWXD2194	DXS7365	25R	178	TNK100	GAGGCAGTGGATAACAGTTGGGG	CTCACAACCCTACACTTTTACATTAC	L41206	G00-556-371
sWXD2195	DXS7366	XSTS0195	64	TNK100	TGTCACAATTTAGCGAACATTGC	GATAAATACGCTAGTGTTATAACC	L41207	G00-556-374
sWXD2196	DXS8263	318R	88	TNK100	GCATTTTCTCTTAGTCTGTGG	CTGGACATTATAAAAATTAGCAATGC	L47527	G00-668-459
sWXD2197	DXS7367	XSTS0147	104	TNK100	AAACTGTTACCTTGCCTG	CTAGATGAAGATTAAAATTTGACTCC	L41208	G00-556-377
sWXD2198	DXS7368	I240E2L	62	TNK100	CAAATCTACACTGCTTTTCATTGG	GGAATACAAAGGGGACAGG	L41209	G00-556-380
sWXD2199	DXS7792	I126H5L	120	TNK100	CACATCAGCATTTCTTCTCAATC	AAAACTACACCAATCAAAACTGTC	L42742	G00-600-655
sWXD2205	DXS8300	XDB-D12	105	TNK100	TCCCCCAGAAGAGATTGAC	GTTCTATCCAGGCACTCAG	L47528	G00-677-336
sWXD2471		XSTS0516	64	TNK100	CAGATTCAGAGGGAATTACTTAAC	CATATTGGGAATAACACTTAACAC		
sWXD2526	DXS8301	64F1L	113	TNK100	GCAAAGTAGAGAACCATTTCC	GGCACAAAAATCCCCCAAG	L48006	G00-677-339
sWXD2703	DXS7511	DXS7511	129	TNK100	GCGAGGTCAGGTCAATGAAA	TCATTATTTAGGATTGACATTTTGG	G05260	G00-588-700
sWXD2777	DXS8304	XSTS0856	102	TNK100	CACCCCAGCAACAATTATTAAC	ACATTCTCCTTGATCTCAATGAAC	L48009	G00-677-348
sWXD2892	DXS982	DXS982	800	TNK50	GGTCATTATGAATGGTCAAAAGGAATTTGACTGC	CCGTCACATATAGAGATAGTATTACTGAACC		G00-216-814
sWXD2932		XSTS1845	86	TNK100	ATTGATGGGTAATTACAG	GATGGGATTATTTTTATTC	L77322	1296839
sWXD2955		I27E3L	132	TNK50	CTAGAAATGAAATAGAAAGGGAAC	GGTAGAAAAAGGAAGAATGTAGAG	L77308	1297109
sWXD3105		M727D3R	102	TNK100	AACTATTCAGAGGAGAGGG	CTGTAGTTTTCTTGTTTTATGTGTC	L77626	1297147
sWXD3122	DXS731	DXS731	100	TNK100	CTCACCATTGGGTCTTCATACA	TATGATAGGCATGAATTGTGTCTG	M87767	G00-182-276
sWXD3128		M802A8L	109	TNK100	CACTAATGCACTACAATTTAACTC	GTAAATAACAGAAAGACTGGTAAAAG	L77635	
sWXD3129		XSTS7038	76	TNK100	GATTTCAATGTGTTGCTGGG	ACAGGAAACACGAAAATGCAAGG	L77262	1296841
sWXD3130		I158C1L	123	TNK100	TCAAACCATTCCCAAAGCC	ATTCACACTTTTATCACCCATTTC	L77754	1297113
sWXD3131		I275H8R	127	TNK100	GGACAGATTATTTGGACAGG	CATAAGACAGATTATAGCAGACAG	L77636	
sWXD3137		yX145F1R	119	TNK100	GACAGTTAATTTCAAGGGTGAG	CATAATGTGTTAGGCATTTGTGAG	L77641	
sWXD3265		I345F2R	116	TNK100	TCAGTATCAGTCAATGAATGAAGG	TGAGAGAAAGTGGGGGTGG	L77646	
sWXD3266		yX34G7R	98	TNK100	TCTTTTAAGTTCTGTTCCCTTTTG	TATGGCTTGGTTCCCTTTG	L77647	

The sWXD accession name (see Methods) is shown, along with DXS numbers, aliases, and sizes of amplified products. Where DXS numbers are not given, STSs have been submitted and numbers requested. All sequences, including those not yet given GenBank accession numbers, are available on the Center for Genetics in Medicine web page (see Methods). A hyperlinked version of these data is also available at http://www.cshl.org/gr.

YACs was one indication of the validity of the map. A second came from the self-consistent patterns of Alu- and L1-containing *Taq*I fragments generated from YACs spanning about half the interval covered by the contig; as in the case of Xq27 (Zucchi et al. 1996), there were no discrepancies observed in the profile obtained from YACs from several different collections (data not shown).

In another approach to verification of marker order, inter-Alu products from YACs containing each of a set of markers were differentially labeled and ordered by fluorescent in situ hybridization (FISH). An example is given in Figure 2. Along a chromosome that is incompletely isolated but oriented by its banding pattern, clones containing the markers DXS1206, DXS982, and DXS100 lie in that order from centromere to telomere. The results are in agreement both with the contig map of Figure 1 and with pairwise hybridizations of the markers (not shown).

Still another set of results that are strongly consistent with the order of markers as shown came when each STS was tested against DNA from each of the two deletion patients 739 (3.5 Mb) and 503 (3 Mb), and its position compared with the inferred order of Figure 1. Sharp borders were seen between the STSs included and excluded from each deletion, with complete consistency as indicated in the summary of Figure 3. Each breakpoint was localized within a distance of ~100 kb. The centromeric breakpoints for both deletions fell between STSs DXS7284 (sWXD1097) and DXS8009 (sWXD1611), and the two telomeric breakpoints fell at two distinct locations ~200 kb apart—between STSs sWXD3131 and DXS7668 (sWXD1503), and DXS8262 (sWXD2062) and DXS8212 (sWXD416), respectively.

For one location, the right end of yWXD734, no DNA extending appreciably further in the centromeric direction was recovered in 10 genomic equivalents of YACs or 10 equivalents of cosmids and bacterial artificial chromosomes (see Discussion). The contigs centromeric and telomeric of the remaining gap (Fig. 1) are aligned and oriented by links to adjoining contigs and by additional cytogenetic and genetic analyses.

The resolution and clone depth of the map was sufficient to resolve several apparent discrepancies in the ensemble of the data. For example, a megaYAC, 802A8, contained a large number of STSs from the region around DXS1206, but an STS made from an end of the YAC insert mapped several megabases further toward the telomere, near DXS75. Because none of the STSs in the intermediate region established by the YAC STS contig were found in the megaYAC, it was inferred to contain a large internal deletion, and has not been shown in the map of Figure 1.

Sequences of YAC insert ends in the region were generally low in GC content. The percent GC for a random selection of 26 is listed in Table 2; most of them were too low in GC content to afford STS development. The overall level of GC (34%) corresponds to the very lowest GC content DNA in the genome, indicative of a low content of genes (see Discussion).

Duplicated STSs

From the start of mapping in this region, some hybridization probes and STSs showed weak signals with clones from apparently different locations. Scoring of STSs could be made less ambiguous by using more stringent conditions for the PCR amplification step, but some ambiguities remained. An extreme example of apparently regionally duplicated sequences was seen with STSs from the ends of YAC yWXD162. Clone yWXD162 yields amplification products for four STSs that fall in a single set of YACs (sWXD918, 914, 1095, and 2205 in Fig. 1). STSs from both insert ends also amplified products only from YACs in Xq25, but at more than one site.

162R, an STS from the right insert end of the YAC, gave amplification signals from a second location more centromeric to the contig shown here (data not shown), suggesting that 162R contains a partially duplicated sequence. More extensive indications of duplication were seen with the STS from the other end of the YAC (sWXD1095, 162L). It amplified a product of the appropriate size from YACs at three different locations. One position coincided with the location of the four internal STSs in yWXD162, near DXS1206. The other two included some YACs just centromeric and another group just telomeric to DXS100. PCR assays with two clones from each region are shown in Figure 4. [Consistently, YACs from two of the locations yielded a second, less intense amplification product of somewhat smaller size. The secondary product was never seen with total human DNA, and may represent a truncated STS (R. Nagaraja and S. MacMillan, unpubl.).] Suggestively, these three positions fall, respectively, near the telomeric breakpoints of the larger and smaller deletions and near the centromeric breakpoint(s) of both deletions.

DISCUSSION

In addition to the self-consistency of STS content in

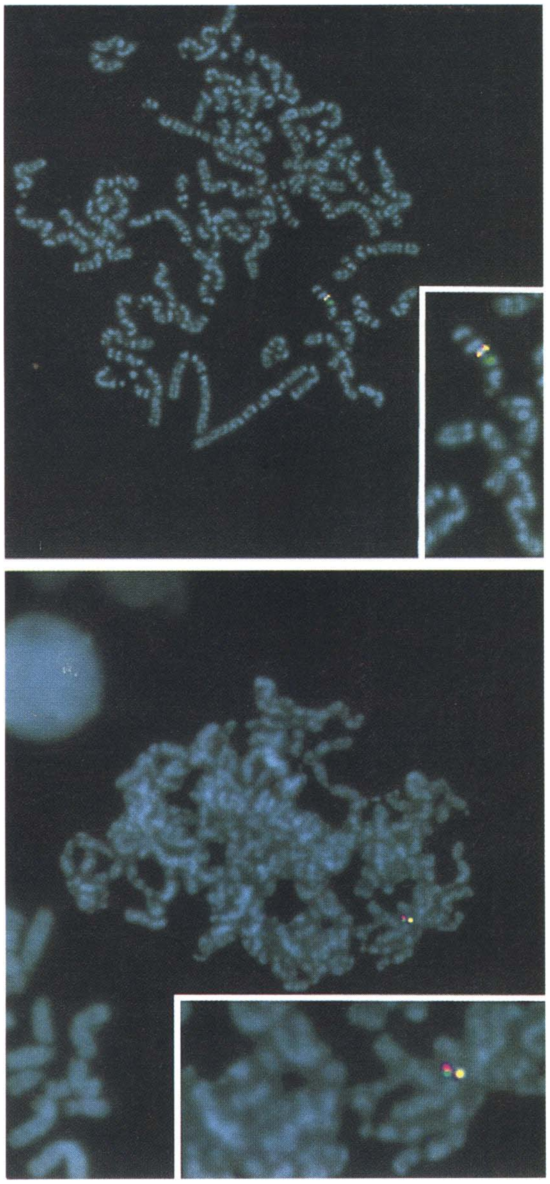


Figure 2 FISH of selected YACs. (*Top*) Hybridization of yWXD640, containing DXS1206 (green); yWXD856, containing DXS982 (red); and cosmid cWXD715 (U234G3), containing DXS100 (green) to a male prometaphase chromosome spread. The order of markers from the centromere along Xq is DXS1206 (green), followed by DXS982 and DXS100, which are overlapping and not resolved (red/green), appearing as yellow. Inset shows an enlargement of the area containing signals. (*Bottom*) Hybridization of yWXD640 (DXS1206), labeled with red and green dyes, appearing as yellow; yWXD856 (DXS982), labeled with digoxigenin and detected with anti-digoxigenin-FITC (green); and cWXD715 (DXS100, red) to a male G1 PCC metaphase plate. The orientation of the chromosome from centromere to telomere is unequivocally determined by the banding pattern, and the order of the markers from the centromere along Xq is DXS1206 (red and green, appearing as yellow), DXS982 (green), DXS100 (red). Inset shows an enlargement of the area containing signals.

the YACs, the contig as shown in Figure 1 is verified by auxiliary fingerprinting assays, and by FISH studies of YACs and cosmids containing six major markers. The order of markers also agrees with the preliminary placement of a subset of probes (Wu et al. 1993), except for DXS6, which is now placed more telomeric to DXS100. This placement of DXS6 has been confirmed both by FISH and by the critical finding that DXS6 maps outside of the intervals spanned by the two deletions.

Recently the content of polymorphic markers in this region has been augmented; all are found in the contig. Because the total recombination distance in this physical interval is only ~1.6 cM, the order of markers on the genetic map is relatively weak, and cannot be stringently compared with the physical order. In ongoing work, however, longer-range linkage and physical orders are in accord for a contig extended several megabases in both the centromeric and telomeric directions (R. Nagaraja and S. MacMillan, in prep.).

Refining the Localization of the LYP Gene

In a few cases, and notably in the case of Duchenne muscular dystrophy (see Mandel et al. 1992), interstitial deletions have localized a hereditary disease lesion to a region small enough for molecular analysis. In the case of LYP, the reduction of the target

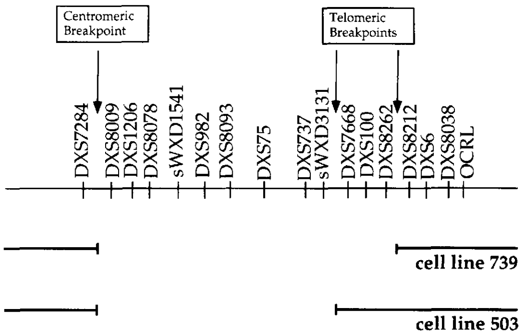


Figure 3 Localization of deletion breakpoints in DNAs from two LYP patients. For clarity, the relative locations of a small number of STSs of Fig. 1 are shown in relation to the positions of the breakpoints in the 739 and 503 deletions. The arrows indicate the location of centromeric and telomeric breakpoints.

Table 2. Fraction of GC in Sequences of YAC Insert Ends

Insert-end	Base pair	% GC	Insert-end	Base pair	% GC
788 L	142	26	162 R	155	45.2
788 R	225	27	592 L	189	34.4
3163 L	221	38	286 L	162	39.6
1356 R	126	20	839 R	169	33.8
1357 R	127	19.7	856 R	198	27.8
763 R	87	33.4	5159 L	164	29.3
606 R	106	44	3478 L	142	33.1
606 L	52	40.4	3478 R	151	43.1
592 R	83	27.8	1829 L	169	43.2

A sample set of YAC insert-sequences across the region; all but five of them (sWXD269, 788R; sWXD1100, 162R; sWXD875; 5159L; sWXD1290, 3478R; and sWXD1291, 1829L) were too high in AT level to yield workable STSs. The overall average % GC was 34.

region to a 3-Mb interval provides an entree to the next stage of the search for the gene, although this is a large target area for a single gene search. On the other hand, there are several indications that this region contains relatively few genes. First, the patients show no symptoms other than LYP, so that there are clearly no other genes in the region whose loss results in significant pathology. Second, rare-cutter restriction analysis of about half the region (G. Porta, unpubl.) shows few if any of the CpG islands that are telltales for an estimated half of all genes in the genome (Antequera and Bird 1993). Third, the GC content of the region, an indicator of gene concentration (Bernardi 1993), is quite low. One indication of this is the fraction of YAC insert ends that showed high AT levels (Table 2). Also, in direct analyses of genomic DNA fractionated by gradient centrifugation and probed with a series of markers from the region, probes for DXS100 and DXS6, for example, indicated an ambient isochore with the low GC level of ~41% GC. The sum of these results suggest that the LYP gene is likely to be one of a small number encoded in the DNA that is absent in the 503 deletion.

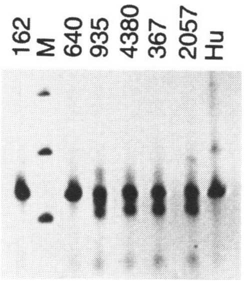


Figure 4 PCR-based amplification of STS sWXD1095 (162L) from groups of YACs at three loci in Xq25. Assays are shown for pairs of YACs from each of three loci (Fig. 1 and text): yWXD162 and 640; 935 and 4380; 367 and 2057. Lane *2* shows a set of marker DNA fragments of 50, 100, and 200 bp. The expected 61-bp product appears in total human DNA (*right* lane) as well as in the YACs.

Deletion Formation and Disease Etiology

In an increasing number of cases, starting with the classic hotspots seen in the DMD gene (Lindlof et al. 1989), recombination between very highly repetitive elements has been shown to underlie the formation of a fraction of deletions involved in human pathophysiology. For example, inter-Alu recombination has been implicated in hypercholesterolemia (Lehrman et al. 1986); hereditary C3 deficiency (Botto et al. 1992), and mucopolysaccharidosis type IVA (Hori et al. 1995). In other cases, however, local low-frequency repetitive elements have been implicated in deletion formation that leads to hereditary pathology. Two notable examples have been analyzed on the X chromosome. CRI-S232 and G1.3, which are interspersed in Xp22.3, have been suggested to be involved in the generation of deletions, including those that lead to Kallmann's syndrome (Ballabio et al. 1990; Li et al. 1992). Another example is the deletion formation in Xq28, which occurs in a number of cases between the multiple copies of the gene that has one of its sites in intron 22 of the factor VIII gene (Freije and Schlessinger 1992). Those deletions result in hemophilia A (Tuddenham et al. 1994).

The deletions analyzed here may provide another comparable case, because duplicated segments including the STS from YAC 162L lie at or near the borders of the deletions. Recombination based on pairing and looping out of the homologous segments may then be a significant source of deletions resulting in LYP.

METHODS

Clones and STS Content Mapping

YACs, P1 artificial chromosomes (PACs), and STSs are referred to either by DXS numbers where available, or by their accession nos. in the Washington University Genome Center database, with the corresponding prefixes yWXD, pWXD, and sWXD (see legend to Fig. 1); details, including YAC library locations and chimera information, are in Genome Database (http://gdbwww.gdb.org) and on the Center Web page at http://genome.wustl.edu/cgm/cgm.html. YACs are available from American Type Culture Collection (ATCC). Cosmid U234G3, containing probe DXS100, was screened from the Livermore library by a PCR assay for the probe. PACs were screened from the library prepared by Dr. P. de Jong (Roswell Park Cancer Institute, Buffalo, NY). YACs yWXD7802 (yX31C9) and 7803 (yX34G7) were kindly provided by Drs. A. Coffey and D. Bentley (Sanger Centre, Hinxton, UK). Cellular DNAs with 3 Mb (503) or 3.5 Mb (739) deletions were as reported (Wu et al. 1993). The most probable order of markers in Figure 1 is inferred by the version 3.41 SEGMAP program of P. Green and C. Magness (University of Washington; obtained from chuck@chimera. biotech.washington.edu).

FISH Method

Cosmids and Alu–PCR amplification products from the YACs (Lengauer et al. 1992) were labeled by incorporation of biotin-11-dUTP (Sigma) or digoxigenin-11-dUTP (Boehinger Mannheim) using a nick translation kit (GIBCO-BRL). Labeled DNAs in the range of 50–800 bp, with a median size of ~200 bp, were used for the in situ hybridization experiments.

In situ hybridization and washing procedures were performed as described (Srivastava et al. 1995). Triple hybridizations were done with 100 ng each of biotin-, digoxigenin-, or biotin–digoxigenin-labeled DNA from two YACs and 50 ng of biotin-, or digoxigenin-labeled DNA from a cosmid, plus 120 µg of human Cot1 DNA (GIBCO-BRL), 5 µg of sheared salmon sperm DNA, and 5 µg of *Escherichia coli* tRNA in 10 µl of 10% dextran sulfate, 50% formamide, and 1× SSC, and hybridized overnight at 37°C. Biotin-labeled probes were detected using Avidin conjugated with Cy-3 (Sigma) resulting in red signal at the location of the probe on the chromosome. Digoxigenin-labeled probes was detected with antidigoxigenin conjugated with fluorescein isothiocyanate (FITC) (B.M.) and amplified with an antisheep antibody conjugated with FITC (Jackson Laboratories). This gives a green signal at the location of the probe on the chromosome.

Slides were examined under a Zeiss Axiovert fluorescent microscope. DAPI replication banding was viewed with an 02 filter set, the FITC signal with an 09 filter set, and the Cy3 signal with a 15 filter set. Images were captured using a cooled CCD (Photometrics, Tucson, AZ) and displayed on a Macintosh computer. Images were processed as described in Featherstone and Huxley (1993). Chromosomes were identified by the 4′,6-diamidino-2-phenylindole (DAPI) replicational banding pattern.

ACKNOWLEDGMENTS

We are grateful to Allison Coffey and David Bentley for many useful discussions and the provision of YACs and STSs. This work was supported by grants HG00201 and HG00247, Giovanni Porta's portion was funded by the Associazione Italiana Ricerca sul Cancro, and the Telethon E 294.

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

Antequera, F. and A.A. Bird. 1993. Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci.* **90:** 11995–11999.

Ballabio, A., B. Bardoni, S. Guioli, E. Basler, and G. Camerino. 1990. Two families of low-copy number repeats are interspersed on Xp22.3: Implications for the high frequency of deletions in this region. *Genomics* **8:** 263–270.

Bare, R.S., C.J. DeLor, K.P. Clauser, P. Hurtubise, W. Henle, J.F. Hewetson. 1974. Fatal infection mononucleosis in a family. *N. Engl. J. Med.* **290:** 363–367.

Bernardi, G. 1993. The isochore organization of the human genome and its evolutionary history — A review. *Gene* **135:** 57–66.

Botto, M., K.Y. Fong, A.K. So, R. Barlow, R. Routier, B.J. Moreley, and M.J. Walport. 1992. Homozygous hereditary C3 deficiency due to a partial gene deletion. *Proc. Natl. Acad. Sci.* **89:** 4957–4961.

Conley, M.E. 1992. Molecular approaches to analysis of X-linked immunodeficiencies. *Annu. Rev. Immunol.* **10:** 215–238.

Featherstone, T. and C. Huxley. 1993. Extrachromosomal maintenance and amplification of yeast artificial chromosome DNA in mouse cells. *Genomics* **17:** 267–278.

Freije, D. and D. Schlessinger. 1992. A 1.6-Mb contig of yeast artificial chromosomes around the human factor VIII gene reveals three regions homologous to probes for the DXS115 locus and two for the DXYS64 locus. *Am. J. Hum. Genet.* **51:** 66–80.

Hori, T., S. Tomatsu, Y. Nakashima, A. Uchiyama, S. Fukuda, K. Sukegawa, N. Shimozawa, Y. Suzuki, N. Kondo, T. Horiuschi, et al. 1995. Mucopolysaccharidosis type IVA: Common double deletion in the N-acetylgalactosamine-6-sulfatase gene (GALNS). *Genomics* **26:** 535–542.

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

Lehrman, M.A., D.W. Russell, J.L. Goldstein, and M.S. Brown. 1986. Exon-Alu recombination deletes 5 kilobases from the low density lipoprotein receptor gene, producing a null phenotype in familial hypercholesterolemia. *Proc. Natl. Acad. Sci.* **83:** 3679–3683.

Lengauer, C., E.D. Green, and T. Cremer. 1992. Fluorescence in situ hybridization of YAC clones after Alu-PCR amplification. *Genomics* **13:** 826–828.

Li, M., P.H. Yen, and L.J. Shapiro. 1992. Characterization of a low copy repetitive element S232 involved in the generation of requent deletions of the distal short arm of the human X chromosome. *Nucleic Acids Res.* **20:** 1117–1122.

Lindlof, M., A. Kiuru, H. Kaariainen, H. Kalimo, H. Lang, H. Pihko, J. Rapola, H.J. Somer, M. Somer, M.L. Savontaus, et al. 1989. Gene deletions in X-linked muscular dystrophy. *Am. J. Hum. Genet.* **44:** 496–503.

Mandel, J.L., A.P. Monaco, D. Nelson, D. Schlessinger, and H.F. Willard. 1992. Genome analysis and the human X chromosome. *Science* **258:** 103–109.

Ogasawara, J.R., R. Watanabe-Fukunaga, M. Adachi, A. Matsuzawa, T. Kasugai, Y. Kitamura, N. Itoh, T. Suda, and S. Nagtata. 1993. Lethal effect of the anti-Fas antibody in mice. *Nature* **364:** 806–809.

Purtilo, D.T., D. DeFlorio, L.M. Hutt, J. Bhawan, J.P.S. Yang, R. Otto, and W. Edwards. 1977. Variable phenotypic expression of an X-linked recessive lymphoproliferative syndrome. *N. Engl. J. Med.* **297:** 1077–1081.

Seemayer, T.A., T.G. Gross, R.M. Egler, S.J. Pirruccello, J.R. Davis, C.M. Kelly, M. Okano, A. Lanyi, and J. Sumegi. 1995. X-linked lymphoproliferative disease- twenty-five years after the discovery. *Pediatr. Res.* **38:** 471–478.

Skare, J., B.-L. Wu, S. Madan, V. Pulijaal, D.T. Purilo, D. Haber, D. Nelson, B. Sylla, H. Grierson, and H. Nitowsky. 1993. Characterization of three overlapping deletions causing X-linked lymphoproliferative disease. *Genomics* **16:** 254–255.

Srivastava, A.K., T. Featherstone, K. Wein, and D. Schlessinger. 1995. YAC contigs mapping the human COL4A5 and COL4A6 genes and DXS118 within Xq21.3-q22. *Genomics* **26:** 502–509.

Sullivan, L.J. and B.A. Woda. 1989. X-linked lymphoproliferative syndrome. *Immunodeficiency Rev.* **1:** 325–247.

Tuddenham, E.G., R. Schwaab, J. Seehafer, D.S. Millar, J. Gitschier, M. Higuchi, S. Bidichandani, J.M. Connor, L.W. Hoyer, A. Yoshioka, et al. 1994. Haemophilia A: Database of nucleotide substitutions, deletions, insertions, and rearrangements of the factor VIII gene, second edition. *Nucleic Acids Res.* **22:** 3511–3533.

Wu, B-L, A. Milunsky, D. Nelson, B. Schmeckpeper, G. Porta, D. Schlessinger, and J. Skare. 1993. High-resolution mapping of probes near the X-linked lymphoproliferative disease (XLP) locus. *Genomics* **17:** 163–170.

Wyandt, H., H. Grierson, W. Sanger, J. Skare, A. Milunsky, and D.T. Putilo. 1989. Chromosome deletion of Xq25 in an individual with X-linked lymphoproliferative disease. *Am. J. Med. Genet.* **33:** 425 –430.

Zucchi, I., S. Mumm, G. Pilia, S. MacMillan, R. Reinbold, L. Susani, J. Weissenbach, and D. Schlessinger. 1996. YAC/STS map across 12 Mb of Xq27 at 25 kb resolution, merging Xq26-qter. *Genomics* **34:** 42–54.

Received August 12, 1996; accepted in revised form November 12, 1996.