Construction of a YAC Contig Covering Human Chromosome 6p22

Patrizia Malaspina, Antonella Roetto,* Flavia Trettel, Carla Jodice, Paola Blasi, Marina Frontali,† Massimo Carella,‡ Brunella Franco,§ Clara Camaschella,*,¶ and Andrea Novelletto[1]

*Dipartimento di Biologia, Università di Roma Tor Vergata, 00133 Rome, Italy; *Dipartimento di Scienze Biomediche ed Oncologia Umana, Ospedale San Luigi, Orbassano, 10100 Turin, Italy; †Istituto di Medicina Sperimentale, C.N.R., 00137 Rome, Italy; ‡IRCCS-Ospedale CSS, 71013 San Giovanni Rotondo, Italy; §TIGEM, San Raffaele Biomedical Science Park, 20132 Milan, Italy; and ¶C.I.O.S., C.N.R., 10100 Turin, Italy*

A contig covering human chromosome 6p22 that consists of 134 YAC clones aligned based on the presence/absence of 52 DNA markers is presented. This contig overlaps with the 6p23 contig at its telomeric end and with the 6p21.3 contig at its centromeric end. The order of loci within the contig resolves the relative positions of several genetically mapped markers. Among the additional markers used here, there are eight novel PCR assays. The 12 known genes and anonymous ESTs located within the contig establish a first step toward a transcriptional map of this region. The instability of YAC clones observed during this work is also discussed.   © 1996 Academic Press, Inc.

INTRODUCTION

The cytogenetic band 6p22 is a Giemsa positive band (GTG⁺; Holmquist, 1992) spanning approximately from genetic positions 34 to 46, or 5–6% of the entire genetic length of chromosome 6 (Gyapay *et al.,* 1994). Cytogenetically it has been measured to span 21% of the length of the chromosome 6 short arm (Francke, 1994). This band is delimited centromerically by 6p21, a very gene-rich region containing the MHC, and telomerically by 6p23. Both regions have been subjected to extensive physical analysis (Campbell *et al.,* 1993; Nemani *et al.,* 1994; Olavesen *et al.,* 1995; Burt *et al.,* 1996; Totaro *et al.,* 1996). The intensive effort to generate new genetic markers (Gyapay, 1994) has identified several polymorphic and nonpolymorphic microsatellites that have been later anchored to 6p22 by FISH. However, in the first-generation physical map of the human genome (Cohen *et al.,* 1993) this band was poorly covered, and in the recent report by Chumakov

et al. (1995) it is still represented by two separate paths of levels 3 and 1, respectively. In the STS-based map of the human genome (Hudson *et al.,* 1995) two "doubly linked" contigs overlap this region.

Several genes have been assigned to 6p22. By using panels of somatic cell hybrids, Albig *et al.* (1993) assigned to this band the cluster of H1 histone genes; Abbott and Povey (1991) assigned the prolactin gene; Durkin *et al.* (1994) assigned brain-specific ESTs by PCR; and Pappas *et al.* (1995) found six ESTs mapping to 6p21.3–p23, one of which was identified as a zinc finger protein. A number of disease causing genes or cDNAs linked to the MHC are reported in OMIM with a possible location overlapping 6p22. These include the methionine-tRNA-i gene 1 (TRMI1), the transcription factor ID4, the COL11A2 gene involved in Stickler syndrome and otospondylomegaephyseal dysplasia, the branched-chain alpha-keto acid dehydrogenase E1-beta gene (BCKDHB), and one of the processed pseudogenes of the dihydrofolate reductase (DHFRP2). The location for the Sry-related gene SOX4 (Farr *et al.,* 1993) has been shown to overlap 6p22 by somatic cell hybrid analysis. Recently, the gene for hereditary idiopathic hemochromatosis (HFE) has been shown to be closely linked to markers in 6p22 (Raha-Chowdhury *et al.,* 1995). Moreover, evidence from studies in independent family sets has accumulated (reviewed by Peltonen, 1995) in favor of a major locus predisposing to schizophrenia in a region overlapping 6p22. Finally, Foulkes *et al.* (1993), based on the finding of loss of heterozygosity, suggested the presence of a tumor supressor gene relevant to ovarian carcinoma whose putative position may overlap 6p22.

To provide complete physical coverage of 6p22, we constructed a single YAC contig spanning from D6S306 centromerically to D6S109 telomerically, juxtaposing with the contiguous resource of YACs extended up to 6p24 (Olavesen *et al.,* 1995). In this work, the content of the same set of YAC clones was checked for an integrated assemblage of polymorphic genetic markers, anonymous STSs and ESTs, known genes, and anony-

mous subcloned PCR products. Both PCR screening and hybridization probes were used to determine the relationship between the YAC clones.

MATERIALS AND METHODS

Isolation and preparation of YAC DNA. YAC clones from the CEPH (Albertsen *et al.,* 1990; Chumakov *et al.,* 1995) and ICI (Anand *et al.,* 1991) libraries were obtained from the CEPH (Paris), ICRF (London), and DIBIT-HSR (Milan) resource centers. The original cultures and cultures obtained from single colonies were grown in AHC medium and resuspended in 80 μl LMP agarose plugs containing 7 $\times$ 10^7 cells. Plugs were treated according to standard techniques. Two microliters of molten plug was used in PCR assays, whereas an entire plug was used for restriction enzyme digestion, agarose electrophoresis, and Southern transfer. Probe labeling and hybridization were performed according to standard techniques. Filter washing was performed at high stringency.

Development of new markers. We generated new STSs and probes by using three different methods: (1) *Alu*-PCR products were generated from selected YACs. PCR with 250 ng of one of the primers described by Nelson *et al.* (1989) was carried out for 35 cycles with 1.5 mM MgCl$_2$ and an annealing temperature of 65°C. After agarose electrophoresis, bands distinctive for recombinant yeast were eluted and reamplified under the same conditions to reduce contamination by nonspecific products. Some of these bands were subcloned (BTVC1.5, BTVD1.4, BTVG2.9), while others were eluted from 1% LMP agarose gels and used as hybridization probes to filters containing either *Alu*-PCR products or digested DNA of individual YACs in the region. Hybridization was carried out as described by Wapenaar *et al.* (1994) in the presence of total human DNA to suppress signals from repetitive sequences (Sealey *et al.,* 1985). (2) PCR products were generated with a pool of primers designed on splicing donor and acceptor site consensus sequences (Pritchard *et al.,* 1994; Senapathy *et al.,* 1990) (BTV9F, SSADH). PCR was performed with 25 pmol each of 3 primers with 1.5 mM MgCl$_2$ for 35 cycles with an annealing temperature of 45°C. After agarose electrophoresis, bands not observed in wildtype yeast were eluted and reamplified under the same conditions. (3) The third method relied on the observation that nonspecific PCR products were generated by some markers on certain sets of YACs. In some cases the presence/absence pattern of such bands was consistent with a location of the corresponding target sequences within the contig. The bands were then excised, subcloned, and sequenced, and specific primers were developed (BTVOL506).

Products obtained with the above methods were subcloned in pUC18 as blunt-ended inserts (Sure Clone kit, Pharmacia) and sequenced with universal primers with the dideoxy chain termination method by *Taq* polymerase (AmpliCycle kit, Perkin–Elmer). The sequences were then checked for their repetitive content by comparison with human repetitive elements (Jurka *et al.,* 1992). Primers were designed to obtain PCR products in the 100- to 500-bp range from the nonrepetitive regions. The genomic origin of the STSs was checked on a somatic cell hybrid retaining the sole human chromosome 6 on hamster background (GM10629A; NIGMS Human Genetic Mutant Cell Repository).

A specific assay for human butyrophilin (BT) was developed: the PCR product obtained using bovine primers (Amadou *et al.,* 1995) on human genomic DNA was directly sequenced to design human-specific primers. PCR conditions of the above-mentioned markers are reported in Table 1.

Alu-vector probes were generated from individual YACs by using *Alu* and vector primers as previously described (Nelson *et al.,* 1989; Zoghbi and Chinault, 1995). The bands were eluted from agarose gels and used as hybridization probes to digested DNA of individual YACs using the protocol for suppression of signals from repetitive sequences (Sealey *et al.,* 1985).

Probes, PCR primers, and conditions. All STS markers used the buffer recommended by the manufacturer (Promega), 0.2 mM dNTPs, 10 pmol of each primer, 2 μl of molten yeast plug as template in a final volume of 20 μl, and the same PCR thermal profile with the exception of the annealing conditions and number of cycles. The standard conditions were 4 min at 94°C, then the appropriate number of cycles of 45 s at 94°C, annealing, and 45 s at 72°C, followed by a final extension of 5 min at 72°C. In all experiments, distilled water and yeast strain AB1380 DNA were used as negative controls, and genomic DNA was used as a positive control. The markers taken from the literature and the corresponding PCR annealing and cycling conditions are listed in Table 2.

RESULTS

Initial Assembly of Contigs

An initial set of 154 YACs identified from the first release of Généthon data (Cohen *et al.,* 1993) to reside potentially in 6p22 was assembled. Clonal cultures were grown and DNA prepared to retest YACs individually with the markers used to construct the genetic map, D6S422, D6S299, D6S461, D6S464, D6S306, D6S258, D6S276, D6S265, and D6S273. Among the YACs detected by tridimensional screening at Généthon to contain these microsatellite markers, 24% were not confirmed by individual testing . On the other hand, we found 33 YACs positive with at least one of these markers and not reported as positive in Généthon files. We then added individual analyses for markers D6S109, SOX4, D6S506, D6S507, D6S242, D6S105, H1.5, and TRMI1.

The 76 YACs positive for at least one of the markers listed above could be arranged in four separate contigs: the first one contained D6S109, D6S422, SOX4, D6S507, and D6S506; the second one contained D6S276, D6S461, D6S299, and D6S242; the third one contained D6S464, D6S306, H1.5, and D6S105; and the fourth one contained D6S273, D6S258, and D6S265. These results showed that the STS content of the YACs could provide a finer resolution than the genetic map. In fact, Chumakov *et al.* (1995) reported that the loci D6S464, D6S306, D6S258, D6S276, and D6S265 represent the single position 0.46 together with the HLA cluster. On the other hand, our analysis grouped D6S276 in a cluster mapped 4 cM telomerically and placed D6S258 and D6S265 in the proximity of D6S273. When our pool of YACs was assayed by hybridization with a probe specific for the telomeric MHC gene HLA-F (Amadou *et al.,* 1995), only YACs of the fourth contig produced a positive signal. This placed the latter subgroup of YACs centromeric to the others. This region was not considered for further analysis.

Filling Gaps

To fill the gaps between the contigs and to extend the coverage to the 6p22–p23 boundary (Olavesen *et al.,* 1995), we used different methods to generate additional STS markers in the form of PCR assays. This search was directed toward single-copy loci to avoid the problems associated with contig assembly in the presence of repeated or highly homologous loci (Foote *et al.,* 1992). Two methods were used to select low-copy

TABLE 1

New Markers Developed

Name (original YAC)	Size (bp)	Primers	Annealing conditions	MgCl$_2$ (mM)	Cycles
BTVC1.5 (738_b_5)	105	CAGATGCTATGCTTCAC CCAACAGATGGGTTTAG	52°C, 30 s	1.5	30
BTVG2.9 (927_e_10)	160	ACTCAGTATAATTCCATAG AACTGTATTGCTTGTACAC	52°C, 30 s	1.5	30
BTVD1.4 (934_d_11)	450	CTTTAATTCAAGAAGAGAAC GTGAAGCATAGGTGATTG	49°C, 30 s	1.5	30
BTV9F (830_b_5)	264	CCTGATTTAAAGCACATGGC GCATTTACTGTCTTCTTGCC	53°C, 30 s	1.5	35
BTVOL506	230	CCAACATGCAGACACACTAC GACTATAAGGGGGAAAGGC	55°C, 30 s	1.5	35
SSADH (775_g_10)	203	GTGTCTTCGGCAGGCTTC GGTTTGTCAATCAGTTGTGC	55°C, 30 s	1.5	35
L34820	94	TATGTGTGTTACGGGGGC AATAATGGATGGCATGTACC	54°C, 30 s	1.5	35
BT	198	CAGAGAGATTTGACTCCTGG CAGTACCCATTTCCATACAAC	65°C, 45 s	2.5	30

YAC fragments: preliminary amplification of inter-*Alu* DNA or amplification with a pool of primers potentially specific for coding sequences (see Materials and Methods). Of 67 bands identified in YAC-bearing but not wildtype yeast, 29 (43%) were successfully picked and reamplified in a second PCR run under the same conditions. The 16 products cloned and fully sequenced were compared with a compilation of human repetitive elements with FASTA. Six of them still contained repetitive sequences that prevented a specific assay, 1 was an artifact resulting from primer multimerization during the repeated amplifications, and 9 did not show any relevant homology with repetitive elements (optimal alignment scores 49–99) and were considered for PCR development. The check for repetitive content by computer analysis alone turned out to be highly effective. When PCR was performed under stringent conditions (Table 1), in 9 of 9 cases a sharp band of the expected molecular weight was obtained. When tested as probes against digested human genomic DNA, all of these products hybridized to a single band, as expected for single-copy sequences. By performing specific PCR assays on a hamster hybrid containing only human chromosome 6, four of nine STSs were excluded from this chromosome, thus confirming that material from chromosomes other than chromosome 6 is abundant in the YACs used. By comparison with sequences in GenBank, one of the clones contained a 244-bp insert that shared a stretch of identity with the human SSADH gene (Chambliss *et al.,* 1995). We then designed an additional PCR assay (L34820) to amplify the 3′-end of coding and part of the 3′-untranslated regions of human SSADH cDNA specifically.

We observed a nonspecific PCR product generated by marker D6S506 on genomic DNA. A group of adjacent YACs replicated this spurious band, suggesting the presence of a second locus homologous to D6S506 in 6p22. Subcloning and sequencing of this band led to another STS (BTVOL506, Table 1).

The description of the butyrophilin locus (Amadou *et al.,* 1995; Vernet *et al.,* 1993) in a region of the mouse genome showing homology with the human 6p22 led us to develop a specific PCR assay for the human BT.

These novel markers were used to rescreen the Mega-CEPH and ICI libraries. Additional YACs were also included in our analysis based on the results of screening reported by the MIT Genome Center with CHLC and newly developed STSs (10 markers).

The set of YACs, now totaling 243, was analyzed with all of the above-mentioned markers as well as a series of ESTs and known genes mapped or possibly mapping to 6p22. One hundred thirty-four YAC clones turned out to contain at least one marker and could be now arranged in a single contig reaching at its telomeric end a region already covered in 6p23 (Olavesen *et al.,* 1995).

Figure 1 shows a simplified version of the contig comprising 65 YACs. The remaining YACs were omitted for simplicity of presentation but the overall number of YACs positive for each marker is reported. A variable redundancy is apparent, but it is not possible to determine whether this is the result of an unequal representation in the original libraries or is due to a still incomplete detection during the screening procedure.

The order of markers was established by the presence of two or more markers within a single YAC (Fig. 1, last row) and by building up the number of overlapping clones. In addition, inter-*Alu* band probes were used as hybridization probes to reinforce the link between adjacent loci further.

Resolving Ambiguities

The order of markers presented in Fig. 1 generated the minimum number of internal deletions irrespective of the number of markers missing in each case. The relative positions of D6S242 vs D6S461 and AF-M342xe5 vs SOX4 could not be resolved because the two markers in each of these pairs showed a perfect

TABLE 2

Previously Described Markers Used in 6p22 Physical Mapping

| Name | PCR conditions | | | References |
	Annealing	MgCl$_2$ (mM)	Cycles	
Histone gene H1.1	50°C, 60 s	1.5	30	Albig *et al.,* 1993
Histone gene H1.2	56°C, 45 s	1.5	30	Albig *et al.,* 1993
Histone gene H1.3	63°C, 30 s	1.5	30	Albig *et al.,* 1993
Histone gene H1.4	63°C, 30 s	1.5	30	Albig *et al.,* 1993
Histone gene H1.5	58°C, 45 s	1.5	35	Albig *et al.,* 1993
D6S422	52°C, 45 s	1.5	35	Gyapay *et al.,* 1994
AFM342xe5	47°C, 45 s	1.5	30	Gyapay *et al.,* 1994
AFM082tel	59°C, 45 s	1.5	30	Gyapay *et al.,* 1994
D6S461	55°C, 30 s	1.5	35	Gyapay *et al.,* 1994
D6S299	52°C, 45 s	1.5	35	Gyapay *et al.,* 1994
D6S276	55°C, 45 s	1.5	30	Gyapay *et al.,* 1994
D6S464	52°C, 45 s	1.5	35	Gyapay *et al.,* 1994
D6S265	55°C, 30 s	1.5	35	Gyapay *et al.,* 1994
D6S258	55°C, 30 s	1.5	35	Gyapay *et al.,* 1994
D6S273	53°C, 45 s	1.5	30	Gyapay *et al.,* 1994
D6S306	52°C, 45 s	1.5	35	Gyapay *et al.,* 1994
EST449	52°C, 45 s	1.5	35	Pappas *et al.,* 1995
EST2534	57°C, 30 s	1.5	30	Pappas *et al.,* 1995
EST798	54°C, 30 s	1.5	35	Pappas *et al.,* 1995
SOX4	65°C, 30 s	1.5	35	Farr *et al.,* 1993
WI9755	58°C, 45 s	1.5	35	Hudson *et al.,* 1995
D6S1078	62°C, 45 s	1.5	30	Hudson *et al.,* 1995
D6S1074	60°C, 60 s	1.5	30	Hudson *et al.,* 1995
WI7513	62°C, 45 s	1.5	30	Hudson *et al.,* 1995
D6S1328	52°C, 45 s	1.5	30	Hudson *et al.,* 1995
WI3998	56°C, 45 s	1.5	30	Hudson *et al.,* 1995
D6S1921	52°C, 45 s	1.5	30	Hudson *et al.,* 1995
D6S1468	56°C, 45 s	1.5	30	Hudson *et al.,* 1995
D6S1918	58°C, 90 s	1.5	30	Hudson *et al.,* 1995
D6S1422	58°C, 45 s	1.5	30	Hudson *et al.,* 1995
125 5	52°C, 45 s	1.5	30	Hudson *et al.,* 1995
150 1	52°C, 45 s	1.5	30	Hudson *et al.,* 1995
100 1	58°C, 30 s	1.5	30	Hudson *et al.,* 1995
GATAP19326	56°C, 45 s	1.5	30	Hudson *et al.,* 1995
STS_G9945	58°C, 90 s	2.5	30	Hudson *et al.,* 1995
D6S1558	65°C, 45 s	1.5	30	Hudson *et al.,* 1995
D6S1016	60°C, 45 s	1.5	30	Buetow *et al.,* 1994
D6S1050	60°C, 45 s	1.5	30	Buetow *et al.,* 1994
D6S1002	60°C, 60 s	1.5	30	Stone *et al.,* 1994
D6S109	65°C, 30 s	1.5	35	Ranum *et al.,* 1991
D6S506	60°C, 60 s	1.5	30	GDB
D6S507	67°C, 60 s	1.5	30	GDB
D6S986	67°C, 120 s	1.5	30	GDB
D6S242	58°C, 45 s	1.5	35	GDB
PRL	66°C, 60 s	1.5	30	Abbott and Povey, 1991
				Amadou *et al.,* 1995
TRMI1	50°C, 60 s	1.5	30	Polymeropoulos *et al.,* 1991
PGC	55°C, 45 s	1.5	30	Hayano *et al.,* 1988
DSP	57°C, 45 s	1.5	30	Arnemann *et al.,* 1991
DHFRP2	55°C, 45 s	1.5	30	Anagnou *et al.,* 1985
D6S346E	62°C, 60 s	1.5	30	Durkin *et al.,* 1994
D6S105	54°C, 60 s	1.5	30	Weber *et al.,* 1992
729LK	65°C, 30 s	2.5	28	H. Zoghbi, *pers. comm.,* 1994, Houston, TX

correlation for their presence/absence in YACs (Fig. 1, last two rows). This was also true for SSADH vs L34820. However, we were able to determine their relative order by hybridizing the flanking YACs with a cDNA probe overlapping the 3′-end of the gene. This result further confirmed the assignment of the SSADH gene to this chromosomal region.

On the other hand, YAC deletions in the H1 histone cluster generated ambiguities to order the corresponding markers. Among several best-fitting orders, Fig. 1 shows that obtained by Volz *et al.* (1994) on the basis of long-range restriction mapping.

The entire set of markers is "doubly linked" to adjacent markers within the contig with three exceptions

LIBRARY	CLONE	ASSAY SIZE (Kb)
CEPH	19_a_5	
ICI	11HH2	250
CEPH	10_g_4	
ICI	20BF3	400
CEPH	57_g_4	
MegaCEPH	689_g_1	140–580
MegaCEPH	878_e_3	1380
CEPH	423_b_1	
CEPH	387_d_2	
MegaCEPH	964_f_1	850
MegaCEPH	733_e_8	860
MegaCEPH	753_h_12	1770
MegaCEPH	727_b_2	1380
MegaCEPH	908_d_11	1170
MegaCEPH	928_c_4	1540
MegaCEPH	950_h_11	1760
MegaCEPH	912_c_11	570–740
MegaCEPH	905_g_1	1250
MegaCEPH	732_d_11	660
MegaCEPH	764_h_9	1160
MegaCEPH	974_a_2	
MegaCEPH	712_f_12	500
MegaCEPH	667_a_8	740
MegaCEPH	899_g_1	850
MegaCEPH	767_g_12	1740
MegaCEPH	790_g_7	1040
MegaCEPH	792_g_12	610
MegaCEPH	947_f_6	840
MegaCEPH	771_g_5	1360–1680
MegaCEPH	901_a_10	1130
MegaCEPH	755_g_10	1510
MegaCEPH	720_c_11	1250
MegaCEPH	872_h_10	1630
MegaCEPH	759_h_10	450
MegaCEPH	944_h_10	800
MegaCEPH	755_g_2	1270
MegaCEPH	970_b_7	1190
MegaCEPH	952_d_3	830
MegaCEPH	874_b_3	1430
MegaCEPH	965_d_7	1050
MegaCEPH	935_a_8	1160
MegaCEPH	947_f_9	1260
CEPH	364_d_1	
MegaCEPH	732_b_8	650
MegaCEPH	819_d_5	1060
MegaCEPH	989_g_6	
MegaCEPH	756_b_6	1210
MegaCEPH	933_h_3	1200
MegaCEPH	890_a_9	750
MegaCEPH	815_b_12	
MegaCEPH	927_e_10	890
MegaCEPH	837_a_8	790
MegaCEPH	884_e_9	1670
MegaCEPH	780_b_6	1340–1500
MegaCEPH	805_d_6	1260
MegaCEPH	922_b_2	1700
MegaCEPH	958_h_9	920–1780
MegaCEPH	760_c_9	1370
MegaCEPH	966_e_10	870
MegaCEPH	927_f_7	1020
MegaCEPH	848_d_5	800
MegaCEPH	738_b_5	
MegaCEPH	958_a_10	1690
MegaCEPH	830_b_5	770
MegaCEPH	936_c_1	1070

FIG. 1. Results of YAC alignment based on marker content of clones. A subset of 65 YACs is shown. Presence (+) or absence (−) of DNA markers in YAC clones is indicated. The second row reports the method of marker ascertainment: P, PCR; S, Southern blotting. The penultimate row reports the overall number of YACs containing the corresponding marker. The last row reports the number of YACs containing simultaneously each marker and the marker immediately to its right.

(Fig. 1, last row). The first "singly linked" interval is between markers WI7513 and BTVD1.4, which are both contained only in YAC 966_e_10; the second is between D6S276 and SSADH (755_g_10), and the third is between BT and D6S1016 (905_g_1).

An estimate of the size of the YAC contig can be made by summing the sizes of the smallest overlapping and nonoverlapping YACs with no known deletion across the contig. By these means the region covered is 8.0–9.0 Mb.

The order of markers in the contig refines the relative positions of loci reported in the CEPH genetic map (Chumakov *et al.,* 1995). In particular, D6S285 and D6S422 (3 cM) are separated by 9 markers, as determined by our results and those by Olavesen *et al.* (1995); D6S422 and D6S461 (5 cM) are separated by 17 markers. D6S461 and D6S299, not genetically resolved, are now oriented based on the discordant results of YACs 720_c_11, 944_h_10, 970_b_7, and 947_f_9 (Fig. 1). In agreement with the MIT data (contig WC140), D6S276, genetically mapped 4 cM from [D6S461–D6S299], turns out to be adjacent to D6S299, with its position being defined by discordant results in several YACs. D6S299 and D6S464 (4 cM) are separated by 22 markers and 2 additional markers separate the latter from D6S306. Our contig also supports the consensus map reported by Volz *et al.* (1994) with odds of only 14:1. YAC 423_b_1 places D6S105 telomerically although very close to H1.5. The finding that the HLA-F probe does not hybridize with YACs containing D6S464, D6S306, and D6S276 is in agreement with the location of HLA further centromeric to the H1 histone cluster and, by extension, to the above-mentioned loci (Amadou *et al.,* 1995; Volz *et al.,* 1994). This is in line with the observation that among the 42 YACs reported by Chumakov *et al.* (1995) to contain genetic position 0.46, not one scored simultaneously positive for any of the three microsatellite markers and the MHC cluster markers. The order of markers here reported confirms that of the doubly linked MIT contig WC140 and establishes a link between WC140 and WC1316. In the integrated map (Hudson *et al.,* 1995), however, D6S1422 is placed telomerically to D6S1918, and AFM342xe5 is placed telomerically to D6S1328, an order not supported in our contig. In addition, the same authors report AFM342xe5 and WI7513 to comap in the same radiation hybrid, whereas we resolve their order with four intervening markers.

Inter-*Alu* bands and anonymous probes from the insert-vector boundary of selected YACs were generated to reinforce the link between adjacent loci further. 792G12_2 is a specific *Alu* band obtained by *Alu*-PCR of YAC 792_g_12 (Fig. 1). LE952D3, LE912C11, and LE20BF3, obtained by the corresponding YAC clones by using *Alu*-vector protocols, when used as hybridization probes on a subset of YACs demonstrated that the corresponding inserts are oriented as reported in Table 3.

Instability of YACs

Thirty cultures of the Mega-YAC library were examined in double or triple replicates by two laboratories that obtained them through different resource centers, i.e., after several rounds of mitotic propagation. Of the YACs in this group, 15 gave identical positive results for both members of each pair, 1 gave identical negative results, while 5, 4, 2, and 3 exhibited a loss of one, two, three or four markers, respectively (47% of YACs with losses). The finding of such a high rate of marker loss raises questions as to whether it may affect the ordering of loci. Forty-one markers were examined in at least a pair of subcultures of the same original YAC clone, and 155 markers × pair assays were performed. Overall 31 markers showed loss in one culture among the examined pairs (10% rate of loss). Thus there is a 90% chance for a given marker to be retained in a YAC, ensuring that, with the number of clones examined in this work and found positive (Fig. 1, last two rows), the order of loci can be reliably inferred. Indeed, visual inspection of the contig showed that the ordering of loci involved in losses in each pair is supported by several other YAC clones. An exception is represented by the H1 cluster-containing region, where the order of loci was drawn from long-range restriction mapping (Volz *et al.,* 1994). Recurrent marker loss and the low efficiency in obtaining YACs from this region (Fig. 1, penultimate row) might be the result of enhanced mitotic instability of YACs containing portions of this region of the genome.

DISCUSSION

We have constructed an integrated map of chromosome 6p22 consisting of YACs, polymorphic genetic markers, anonymous STS and ESTs, anonymous subcloned *Alu*-PCR products, and known genes. We primarily used PCR screening to determine the relationships between the YAC clones, but hybridization probing was also used to confirm them. We developed novel STS markers and integrated their analysis with previously described markers in the characterization of the same set of YAC clones.

We resolved the relative positions of nine markers used to construct the Généthon genetic map (Gyapay *et al.,* 1994), confirmed the presence of six of them in this region, and assigned the remaining three to the more proximal MHC-containing region. In fact, in their recent report, Burt *et al.* (1996) physically mapped D6S265 and D6S258 in the proximity of HLA-A and HLA-F genes. D6S276, one of the markers used in the recent radiation hybrid framework map (Gyapay *et al.,* 1996), is retained in many of the YACs harboring D6S299 and in none of those harboring either D6S464 or D6S306. Its position is thus considered more telomeric than previously reported, perhaps as a result of a limited number of recombination events with respect to neighboring markers. The repositioning of

TABLE 3

Hybridization Probes

Name (original YAC)	Probe type	Position	Positive clones
LE952D3 (952_d_3)	*Alu*-vector (left end)	Telomeric end	874_b_3, 965_d_7, 970_b_7, 935_a_8, 947_f_9
LE912C11 (912_c_11)	*Alu*-vector (left end)	Centromeric end	753_h_12, 964_f_1
LE20BF3 (20BF3)	*Alu*-vector (left end)	Centromeric end	753_h_12, 11HH2

D6S276 would imply the breakdown of the level 1 path between genetic positions 0.42 and 0.46 (Chumakov *et al.,* 1995).

We superimposed on this map 46 additional markers. These include the SOX4, prolactin, butyrophilin, and methionine-t-RNA genes as well as the H1 histone cluster containing several genes (Albig *et al.,* 1993). We could physically locate this cluster for the first time between D6S1468 and D6S1016. In addition, we refined the location of two of six ESTs with a 6p22–p23 assignment (Pappas *et al.,* 1995) by positioning them within the contig. We assigned for the first time the gene for NAD$^+$-dependent succinic semialdehyde dehydrogenase (SSADH) to this portion of the genome and found it to be oriented from telomere to centromere.

The number of markers per genetic interval correlates fairly well with the genetic distances reported by Chumakov *et al.* (1995), providing no evidence for a biased distribution. By virtue of the new markers we were able to bring all paths between adjacent loci to level 1.

When comparing our results with those of the MIT Genome Center, the STS WI9755 was not found in any of our YACs. In particular, we examined individually YACs 787_d_1, 847_c_8, 927_f_10, and 844_h_10, reported with three "definite" YAC/STS hits and one "ambiguous" YAC/STS hit in the contig WC140. The first two were indeed confirmed to share markers at the telomeric end of 6p22. With this exception, we observed a better agreement with the individual doubly linked contigs rather than the integrated map, since in some instances it reports an alternative order of markers. In this work we examined YACs individually, thus excluding the imprecision introduced by the tridimensional screening scheme (Hudson *et al.,* 1995). Recurrent loss of markers in independent YACs or complex rearrangements in radiation hybrids may contribute to the observed discrepancies with the integrated map.

We excluded from this region the DHFRP2 and DSP genes and the EST D6S346E, reported in the literature with 6pter-qter, 6pter-p21, and 6p22.1–p21.3 assignments, respectively. Moreover, we excluded the EST798 (Pappas *et al.,* 1995), an expressed sequence assigned to 6p22–p23. Olavesen *et al.* (1995) excluded its presence from 6p23 as well, thus leaving the precise position of this sequence still to be determined.

The anchoring of our contig is established by the assignment of YACs 878_e_3 and 753_h_12 by FISH (Bray-Ward *et al.,* 1996). These clones are found on the proximal side of our contig (Fig. 1) and produced hybridization signals on proximal 6p22 and the 6p21.3–p22 boundary, respectively.

Centromerically, our contig reaches the contig reported by Burt *et al.* (1996), sharing with it four YACs and three markers, and ensuring that 6p22 is physically covered up to its proximal end. In addition, we could orient and order several markers reported to lie in the proximity of the 6p21.3–p22 boundary (Amadou *et al.,* 1995; Raha-Chowdhury *et al.,* 1995; Vernet *et al.,* 1993). Based on the volume of metaphase chromosome (Heslop-Harrison *et al.,* 1989) and the relative length of bands (Franke, 1994), the 6p22 band can be estimated to span 13 Mb. However, this measurement is hardly comparable with the length of our contig, due to variations in the extent of chromosome condensation. To its telomeric end, our contig reaches that published by Olavesen *et al.* (1995), sharing with it five YACs and three markers. The continuous resource of aligned YAC clones now spans from 6p21.3 to distal 6p24.3 (Davies *et al.,* 1995).

Our work enabled us to evaluate the instability of YAC clones upon several runs of mitotic propagations and revealed a remarkable proportion of YACs with marker losses. This might considerably affect the usefulness of YACs in constructing contigs based on STS content. We used a large number of independent clones to increase internal consistency and found reproducible links between adjacent loci in most cases.

The expressed genes localized within the contig and the reagents utilized in this work provide the first core of a transcriptional map of this chromosomal region.

REFERENCES

Abbott, C., and Povey, S. (1991). Development of human chromosome-specific PCR primers for characterization of somatic cell hybrids. *Genomics* 9: 73–77.

Albertsen, H. M., Adberrahim, H., Cann, H. M., Dausset, J., Le Paslier, D., and Cohen, D. (1990). Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. *Proc. Natl. Acad. Sci. USA* 87: 4256–4260.

Albig, W., Drabent, B., Kunz, J., Kalff-Suske, M., Grzeschik, K., and Doenecke, D. (1993). All known human H1 histone genes except the H1° gene are clustered on chromosome 6. *Genomics* 16: 949–654.

Amadou, C., Ribouchon, M. T., Mattei, M. G., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., Avoustin, P., and Pontarotti, P. (1995). Localization of new genes and markers to the distal part of the human major histocompatibility complex (MHC) region and comparison with the mouse: New insights into the evolution of mammalian genomes. *Genomics* 26: 9–20.

Anagnou, N. P., Antonarakis, S. E., O'Brien, S. J., and Nienhuis, A. W. (1985). Chromosomal localization and racial distribution of the polymorphic hDHFR-psi-1 pseudogene. *Clin. Res.* 33: 328A.

Anand, R., Ogilvie, D. J., Butler, R., Riley, J. H., Finniear, R. S., Powell, S. J., Smith, J. C., and Markham, A. F. (1991). A yeast artificial chromosome contig encompassing the cystic fibrosis locus. *Genomics* 9: 124–130.

Arnemann, J., Spurr, N. K., Wheeler, G. N., Parker, A. E., and Buxton, R. S. (1991). Chromosomal assignment of the human genes coding for the major proteins of the desmosome junction, Desmoglein DGI (DSG), desmocollins DGII/III (DSC), desmoplakins DPI/II (DSP), and plakoglobin DPIII (JUP). *Genomics* 10: 640–645.

Bray-Ward, P., Menninger, J., Liemann, J., Desai, T., Mokady, N., Banks, A., and Ward, D. C. (1996). Integration of the cytogenetic, genetic, and physical maps of the human genome by FISH mapping of CEPH YAC clones. *Genomics* 32: 1–14.

Buetow, K. H., Weber, J. L., Ludwigsen, S., Scherpbier-Heddema, T., Duyk, G. M., Sheffield, V. C., Wuang, Z., and Murray, J. C. (1994). Integrated human genome-wide maps constructed using the CEPH reference panel. *Nature Genet.* 6: 391–394.

Burt, M. J., Smit, D. J., Pyper, W. R., Powell, L. W., and Jazwinska, E. C. (1996). A 4.5-megabase YAC contig and physical map over the hemochromatosis gene region. *Genomics* 33: 153–155.

Campbell, R. D., and Trowsdale, J. (1993). Map of the human MHC. *Immunol. Today* 14: 349–352.

Chambliss, K. L., Caudle, D. L., Hinson, D. D., Moomaw, C. R., Slaughter, C. A., Jakobs, C., and Gibson, K. M. (1995). Molecular cloning of the mature NAD$^+$-dependent succinic semialdehyde dehydrogenase from rat and human. *J. Biol. Chem.* 270: 461–467.

Chumakov, I. M., Rigault, P., Le Gall, I., Bellannè-Chantelot, C., Billault, A., Guillou, S., Soularue, P., Guasconi, G., Poullier, E., Gros, I., Belova, M., Sambucy, J. L., Susini, L., Gervy, P., Glibert, F., Beaufils, S., Bui, H., Massart, C., De Tand, M. F., Dukasz, F., Lecoulant, S., Ougen, P., Perrot, V., Saumier, M., Soravito, C., Bahouayila, R., Cohen-Akenine, A., Barillot, E., Bertrand, S., Codani, J. J., Caterina, D., Georges, I., Lacroix, B., Lucotte, G., Sahbatou, M., Schmit, G., Sangouard, M., Tubacher, E., Dib, C., Fauré, S., Fizames, C., Gyapay, G., Millasseau, P., NGuyen, S., Muselet, D., Vignal, A., Morissette, J., Menninger, J., Lieman, J., Desai, T., Banks, A., Bray-Ward, P., Ward, D., Hudson, T., Gerety, T., Foote, S., Stein, L., Page, D. C., Lander, E. S., Weissenbach, J., Le Paslier, D., and Cohen, D. (1995). A YAC contig map of the human genome. *Nature* 377: 175–183.

Cohen, D., Chumakov, I., and Weissenbach, J. (1993). A first-generation physical map of the human genome. *Nature* 366: 698–701.

Davies, A. F., Stephens, R. J., Olavesen, M. G., Heather, L., Dixon, M. J., Magee, A., Flinter, F., and Ragoussis, J. (1995). Evidence of a locus for orofacial clefting on human chromosome 6p24 and STS content map of the region. *Hum. Mol. Genet.* 4: 121–128.

Durkin, A. S., Nierman, W. C., Zoghbi, H. Y., Jones, C., Kozak, C. A., and Maglott, D. R. (1994). Chromosome assignment of human brain expressed sequence tags (ESTs) by analysing fluorescently labeled PCR products from hybrid cell panels. *Cytogenet. Cell Genet.* 65: 86–91.

Farr, C. J., Easty, D. J., Ragoussis, J., Collignon, J., Lovell-Badge, R., and Goodfellow, P. N. (1993). Characterization and mapping of human SOX4 gene. *Mamm. Genome* 4: 577–584.

Foote, S., Vollrath, D., Hilton, A., and Page, D. C. (1992). The human Y chromosome: Overlapping DNA clones spanning the euchromatic region. *Science* 258: 60–66.

Foulkes, W. D., Ragoussis, J., Stamp, G. W. H., Allan, G. J., and Trowsdale, J. (1993). Frequent loss of heterozygosity on chromosome 6 in human ovarian carcinoma. *Br. J. Cancer* 67: 551–559.

Francke, U. (1994). Digitized and differentially shared human chromosome ideograms for genomic application. *Cytogenet. Cell Genet.* 65: 206–219.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., and Weissenbach, J. (1994). The 1993–94 Généthon human genetic linkage map. *Nature Genet.* 7: 246–339.

Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-Czarny, N., Spillett, D., Muselet, D., Prud'Homme, J., Dib, C., Auffray, C., Morissette, J., Weissenbach, J., and Goodfellow, P. N. (1996). A radiation hybrid map of the human genome. *Hum. Mol. Genet.* 5: 339–346.

Hayano, T., Sogowa, K., Ikihara, Y., Fuji-Quriyama, Y., and Takahashi, K. (1988). Primary structure of human pepsinogen C gene. *J. Biol. Chem.* 263: 1382–1385.

Heslop-Harrison, J. S., Leitch, A. R., Schwarzacher, T., Smith, J. B., Atkinson, M. D., and Bennett, M. D. (1989). The volumes and morphology of human chromosomes in mitotic reconstructions. *Hum. Genet.* 84: 27–34.

Holmquist, G. P. (1992). Chromosome bands, their chromatine flavors, and their functional features. *Am. J. Hum. Genet.* 51: 17–37.

Hudson, T. J., Stein, L. D., Gerety, S. S., Ma, J., Castle, A. B., Silva, J., Slonim, D. K., Baptista, R., Kruglyak, L., Xu, S., Hu, X., Colbert, A. M. E., Rosenberg, C., Reeve-Daly, M. P., Rozen, S., Hui, L., Wu, X., Vestergaard, C., Wilson, K. M., Bae, J. S., Maitra, S., Ganiatsas, S., Evans, C. A., DeAngelis, M. M., Ingalls, K. A., Nahf, R. W., Horton, L. T., Jr., Anderson, M. O., Collymore, A. J., Ye, W., Kouyoumjian, V., Zemsteva, I. S., Tam, J., Devine, R., Courtney, D. F., Renaud, M. T., Nguyen, H., O'Connor, T. J., Fizames, C., Fauré, S., Gyapay, G., Dib, C., Morissette, J., Orlin, J. B., Birren, B. W., Goodman, N., Weissenbach, J., Hawkins, T. L., Foote, S., Page, D. C., and Lander, E. S. (1995). An STS-based map of the human genome. *Science* 270: 1945–1954.

Jurka, J., Walichiewicz, J., and Milosavljevic, A. (1992). Prototypic sequences for human repetitive DNA. *J. Mol. Evol.* 35: 286–291.

Nelson, D. L., Ledbetter, S. A., Corbo, L., Maureen, F. W., Ramirez-Solis, R., Webster, T. D., Ledbetter, D. H., and Caskey, T. (1989) Alu polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. Natl. Acad. Sci. USA* 86: 6686–6690.

Nemani, M., Cherif, D., Chesne, H., Pelandakis, M., Ougen, P., Berger, R., Weissenbach, J., Le Paslier, D., Choen, D., and Cann, H. M. (1994). A YAC contig in 6p23 based on sequence tagged sites. *Genomics* 22: 388–396.

Olavesen, M. G., Davies, A. F., Broxholme, S. J., Wixson, J. L., Senger, G., Nizetic, D., Campbell, R. D., and Ragoussis, J. (1995). An integrated map of human chromosome 6p23. *Genome Res.* 5: 342–358.

Pappas, G., Polymeropoulos, M. H., Boyle, J. M., and Trent, J. M. (1995). Regional assignment by hybrid mapping of 36 expressed sequence tags (ESTs) on chromosome 6. *Genomics* 25: 124–129.

Peltonen, L. (1995). All out for chromosome six. *Nature* 378: 665–666.

Polymeropoulos, M. H., Xiao, H., Rath, D. S., and Merril, C. R. (1991). Trinucleotide repeat polymorphism at the human Met-tRNA-i gene 1 (TRMI). *Nucleic Acids Res.* 19: 4306.

Pritchard, M., Fuentes, J. J., Bosch, A., and Estivill, X. (1994). Towards a transcriptional map of human chromosome 21q22.2–22.3. Third EUROGEM meeting, Montpellier, France. [Abstract.]

Raha-Chowdhury, R., Bowen, D. J., Stone, C., Pointon, J. J., Terwilliger, J. D., Shearman, J. D., Robson, K. J. H., Bomford, A., and Worwood, M. (1995). New polymorphic microsatellite markers place the hemochromatosis gene telomeric to D6S105. *Hum. Mol. Genet.* 4: 1869–1874.

Ranum, L. P., Chung, M. Y., Duvick, L. A., Zoghbi, H. Y., and Orr, H. T. (1991). Dinucleotide repeat polymorphism at the D6S109 locus. *Nucleic Acids Res.* 19: 1171.

Sealey, P. G., Whittaker, P. A., and Southern, E. M. (1985). Removal of repeated sequences from hybridization probes. *Nucleic Acids Res.* 13: 1905–1922.

Senapathy, P., Shapiro, M. B., and Harris, N. L. (1990). Splice junctions, branchpoint size and exons: Sequence statistics, identification and applications to genome project. *Methods Enzymol.* 183: 252–278.

Stone, C., Pointon, J. J., Jazwinska, E. C., Halliday, J. W., Powell, L. W., Robson, K. J. H., Monaco, A. P., and Weatherall, D. J. (1994). Isolation of the CA dinucleotide repeats close to D6S105; linkage disequilibrium with hemocromatosis. *Hum. Mol. Genet.* 3: 2043–2046.

Totaro, A., Rommens, J. M., Grifa, A., Lunardi, C., Carella, M., Huizenga, J. J., Roetto, A., Camachella, C., De Sandre, G., and Gasparini, P. (1996). Hereditary hemochromatosis: Generation of a transcription map within a refined and extended map of the HLA class I region. *Genomics* 31: 319–326.

Vernet, C., Boretto, J., Mattei, M., Takahashi, M., Jack, L. J. W., Mather, I. H., Rouquier, S., and Pontarotti, P. (1993). Evolutionary study of multigenic family mapping close to the human MHC class I region. *J. Mol. Evol.* 37: 600–612.

Volz, A., Boyle, J. M., Cann, H. M., Cottingham, R. W., Orr, H. T., and Ziegler, A. (1994). Report of the Second International Workshop on human chromosome 6. *Genomics* 21: 464–472.

Wapenaar, M. C., Schiaffino, M. V., Bassi, M. T., Scaefer, L., Chinault, A. C., Zoghbi, H. Y., and Ballabio, A. (1994). A YAC based binning strategy facilitating assembly of cosmid contig: 1.6 megabases of overlapping cosmid in Xp22. *Hum. Mol. Genet.* 3: 1155–1161.

Weber, J. L., Kwitkok, A. E., May, P. E., and Zoghbi, H. Y. (1992). Dinucleotide repeat polymorphism at the D6S105 locus. *Nucleic Acids Res.* 19: 968.

Zoghbi, H. Y., and Chinault, A. C. (1995). Generation of YAC contigs. *In* "YACs: A User Guide" (B. H. Brownstein and D. L. Nelson, Eds.), Freeman, New York.