GENOMICS 36, 492–506 (1996)
ARTICLE NO. 0495

Mapping Human Telomere Regions with YAC and P1 Clones: Chromosome-Specific Markers for 27 Telomeres Including 149 STSs and 24 Polymorphisms for 14 Proterminal Regions

A. VOCERO-AKBANI,*,[1] C. HELMS,* J-C. WANG,* F. J. SANJURJO,* J. KORTE-SARFATY,*,[2] R. A. VEILE,* L. LIU,* A. JAUCH,† A. K. BURGESS,* A. V. HING,*,[3] M. S. HOLT,* S. RAMACHANDRA,* A. J. WHELAN,*,[4] R. ANKER,*,[5] L. AHRENT,* M. CHEN,* M. R. GAVIN,* K. IANNANTUONI,* S. M. MORTON,* S. D. PANDIT,* C. M. READ,* T. STEINBRUECK,* C. WARLICK,*,[6] D. A. SMOLLER,‡ AND H. DONIS-KELLER*,§,[7]

*Department of Surgery, Division of Human Molecular Genetics, Box 8109, and §Departments of Genetics and Psychiatry, Washington University School of Medicine, St. Louis, Missouri 63110; †Institut für Humangenetik und Anthropologie, Im Neuenheimer Feld 328, D-69120 Heidelberg, Germany; and ‡Genome Systems, Inc., 8620 Pennell Drive, St. Louis, Missouri 63114

Received October 18, 1995; accepted June 14, 1996

A YAC library enriched for telomere clones was constructed and screened for the human telomere-specific repeat sequence (TTAGGG). Altogether 196 TYAC library clones were studied: 189 new TYAC clones were isolated, 149 STSs were developed for 132 different TYACs, and 39 P1 clones were identified using 19 STSs from 16 of the TYACs. A combination of mapping methods including fluorescence *in situ* hybridization, somatic cell hybrid panels, clamped homogeneous electric fields, meiotic linkage, and BLASTN sequence analysis was utilized to characterize the resource. Forty-five of the TYACs map to 31 specific telomere regions. Twenty-four linkage markers were developed and mapped within 14 proterminal regions (12 telomeres and 2 terminal bands). The polymorphic markers include 12 microsatellites for 10 telomeres (1q, 2p, 6q, 7q, 10p, 10q, 13q, 14q, 18p, 22q) and the terminal bands of 11q and 12p. Twelve RFLP markers were identified and meiotically mapped to the telomeres of 2q, 7q, 8p, and 14q. Chromosome-specific STSs for 27 telomeres were identified from the 196 TYACs. More than 30,000 nucleotides derived from the TYAC vector-insert junction regions or from regions flanking TYAC microsatellites were compared to reported sequences using BLASTN. In addition to identifying homology with previously reported telomere sequences and human repeat elements, gene sequences and a number of ESTs were found to be highly homologous to the TYAC sequences. These genes include human coagulation factor V (F5), Wee1 protein tyrosine kinase (WEE1), neurotropic protein tyrosine kinase type 2 (NTRK2), glutathione *S*-transferase (GST1), and β tubulin (TUBB). The TYAC/P1 resource, derivative STSs, and polymorphisms constitute an enabling resource to further studies of telomere structure and function and a means for physical and genetic map integration and closure.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. U53008–U53019, U57699–U57704, U57849–U57876, U58872–U58879.

[1] Present address: Division of Molecular Oncology, Departments of Pathology and Medicine and Howard Hughes Medical Institute, Washington University School of Medicine, St. Louis, MO 63110.

[2] Present address: Department of Microbiology, University of Washington, Seattle WA 98195.

[3] Present address: Department of Pediatrics, Washington University School of Medicine, St. Louis, MO 63110.

[4] Present address: Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110.

[5] Present address: Department of Pathology, Mayo Clinic, Rochester, MN 55905.

[6] Present address: University of Minneapolis, Institute of Human Genetics, Minneapolis, MN 55455.

[7] To whom correspondence should be addressed at Department of Surgery, Division of Human Molecular Genetics, Box 8109, Washington University School of Medicine, St. Louis, MO 63110. Telephone: (314) 362-8629. Fax: (314) 362-8630. E-mail: helen@hdklab.wustl.edu.

INTRODUCTION

In most eukaryotes the ends of linear chromosomes terminate in tandemly repeated simple G-rich sequences; in humans the terminal repeat consists of $(TTAGGG)_n$ repeats ranging from ~2 to 15 kb in length (Moyzis *et al.*, 1988; Allshire *et al.*, 1989). Human telomere regions constitute complex dynamic structures with important functional features. The presence of repeated sequences within and shared among telomere regions (and in some cases interstitially) and the observed heteropolymorphism of chromosome ends have presented challenges to the delineation of the precise genomic structure and sequence organization for these regions. Nonetheless, the availability of detailed physical and genetic maps for each human chromosome (which may consist of collections of maps of telomeres from different individuals) is essential to achieve clo-

0888-7543/96 $18.00

sure of the maps and to make possible inclusion of the entire human genome in disease gene searches. While a number of polymorphic markers have been mapped to the terminal bands of all of the nonacrocentric chromosomes, their order within these regions is for the most part unresolved. Meiotic linkage markers have been reported to lie near the termini of 4p (NIH/CEPH, 1992), 11p (Browne *et al.,* 1995), 16p, 16q (NIH/CEPH, 1992), 18q (Van Kessel *et al.,* 1994), Xp (Schmitt *et al.,* 1994, Henke *et al.,* 1993), and Xq (Freije *et al.,* 1992). Physical maps of telomere regions have also appeared [for example, for 2q (Macina *et al.,* 1994), 7q (Riethman *et al.,* 1993), 14q (Cook *et al.,* 1994), and Xqter (Kvaloy *et al.,* 1994)]. These maps provide a framework for studies of recombination at terminal regions of the genome (reviewed in Rappold, 1993) and provide the foundation for gaining insight into gene organization and regulation within these regions. For overall reviews of telomere literature, see Kipling (1995) or Blackburn and Greider (1995).

In recent years the relevance of telomere studies to aging, cancer, and inherited disorders has been appreciated. Since Counter *et al.* (1992) described their observations of telomere shortening associated with chromosome instability, numerous other reports of phenomena correlating telomerase activity and telomere stability with cell immortalization have followed (see for example, Kim *et al.,* 1994 and Martin, 1994). Examples of disorders mapping to terminal regions of chromosomes include achondroplasia–hypochondroplasia mapping to 4pter (LeMerrer *et al.,* 1994; Francomano *et al.,* 1994), facioscapulohumeral muscular dystrophy mapping to 4qter (Bengtsson *et al.,* 1994); α thalassaemias mapping to 16pter (Wilke *et al.,* 1990), triphalangeal thumb mapping to 7qter (Heutink *et al.,* 1994), and multiple mental retardation syndromes (see, e.g., Flint *et al.,* 1995). New mapping ventures for other inherited disorders by genomic screening would be facilitated by the availability of markers that provide high resolution in terminal regions where the linkage maps are most often expanded. However, the genomic structure and gene organization of many terminal regions of human chromosomes are scarcely known.

We have previously reported preliminary studies for telomere polymorphisms from four telomere regions [2qter, 8pter, 14qter, 12p (NIH/CEPH 1992)] and detailed linkage mapping studies for 7qter (Hing *et al.,* 1993), 14qter (Pandit *et al.,* 1995), and Xqter (Freije *et al.,* 1992). We report here the construction and characterization of a specialized YAC library enriched for telomere sequence-containing clones including physical and genetic mapping studies. STSs we developed during the course of the study will prove useful for producing unified maps of the human genome and for studying chromosome stability, particularly in disease states and aging. We also report BLASTN sequence comparisons that indicate the presence of coding sequences near telomere regions and highlight the complexity of the sequence organization near the telomere and within other regions of the genome containing sequences homologous to TTAGGG. Now that the Human Genome Project has entered the sequencing phase, the need for understanding genomic organization near telomeres is becoming increasingly important, especially for closure. Reagents developed from the study of proterminal regions ultimately will be used to complete the genome sequencing.

MATERIALS AND METHODS

TYAC Library Construction and Screening for $(TTAGGG)_n$-Containing Clones

The TYAC library was prepared using the method described in Riethman *et al.* (1989) with CGM1 leukocyte DNA as the human genomic component of the library. Briefly, CGM-1 DNA was partially digested with endonuclease *Eco*RI and size-fractionated on sucrose density gradients. Vector DNA and ligation reactions were supplied by H. Riethman and prepared as previously described (Riethman *et al.,* 1989). We performed the transformations using spheroplasts from the yeast host strain AB1380 (Burke *et al.,* 1987) and cells plated on selective regeneration medium lacking uracil. Yeast transformants were picked, gridded onto selective medium plates, and grown on nylon filters (Hybond, Amersham) with controls of yeast pTYAC1 clones with and without human telomere sequences. Colony filter lifts were prepared by standard yeast methods and hybridized with a ^{32}P-labeled DNA probe containing $(TTAGGG)_{40}$ derived from the plasmid clone pHuR93 (Moyzis *et al.,* 1988). Probed filters were washed and exposed to X-ray film for 16–48 h. Five independent isolates from each clone positive for the TTAGGG repeat were used in a secondary round of filter lifts and hybridization to the $(TTAGGG)_{40}$ probe. Clones positive from this screen were numbered and reserved for further work.

TYAC Clone Sizing

TYAC clones were sized on 1% agarose CHEF (clamped homogeneous electric fields) gels in 0.5× TBE buffer using yeast chromosomes and λ ladder as size standards (λ ladder agarose plugs was prepared by ligation of WT λ molecules (NE BioLabs) and then embedding in SeaPlaque agarose (FMC). High-molecular-weight YAC clone DNA was prepared embedded in agarose beads and size-separated using a Bio-Rad CHEF DR II apparatus. Gels were electrophoresed at 200 V for 12–14 h with a 60-s switch time followed by 14–16 h with a 90-s switch time. Clone DNA was transferred to Zetabind (AMF, Cuno Laboratories, Meriden CT) or Magnagraph (MSI, Westboro, MA) nylon membrane in an overnight alkaline (0.4 *N* NaOH) Southern blotting. Blots were probed with ^{32}P-labeled $(TTAGGG)_{40}$ DNA to detect human telomeric sequences on YACs and ^{32}P-labeled λ DNA for the detection of size standards. Probed blots were exposed to X-ray film (Kodak) for 12–48 h and the developed autoradiograms were used to estimate the YAC sizes.

TYAC Vector-Insert Junction Assay Development

PCR assays were developed from TYAC vector-insert junction restriction fragments using the ligation-mediated PCR technique described by Mueller and Wold (1989). Yeast genomic YAC DNA was digested with a variety of blunt-cut restriction enzymes (e.g., *Rsa*I, *Alu*I, *Pvu*II, *Eco*RV), and the fragments were ligated to the Mueller/Wold linker (1989): Linker sequences, GCGGTGACCCGGGAGATC-TGAATTC and GAATTCAGATC. The ligated fragments were then specifically amplified by PCR using primers derived from the flanking vector sequence and from the linker sequence. The vector primer sequence used was TAGGCGTATCACGAGGCCCTT.

The resulting PCR products were size-fractionated on agarose mini-gels, and bands of approximately 200–500 bp were excised for sequencing. Some PCR fragments were sequenced using an ABI

Model 373A DNA sequencer with a vector primer and fluorescent dye termination kit. For most PCR product sequencing, the manual cycle sequencing method of Srivastava *et al.* (1992) was used with one primer end-labeled with ^{32}P. PCR primers were selected from the end-fragment sequences using the PRIMER program (Lincoln and Lander, Whitehead Institute) and were prepared using an ABI Model 394 DNA/RNA oligonucleotide synthesizer. Primers were tested for the ability to PCR amplify the expected product size in two or more human genomic DNA samples as well as the original TYAC DNA embedded in agarose beads.

Five-microliter PCRs with 50 ng template DNA and 50 μM each primer were generally carried out using a PCR cocktail containing 50 mM KCl, 10 mM Tris–HCl, pH 8.3, 1.5 mM $MgCl_2$, 2.5 pmol each dNTP, 0.5 U AmpliTaq (Perkin–Elmer Cetus). Twenty-four to thirty cycles of PCR were carried out on Perkin–Elmer Cetus or Omnigene machines using a 1-min denaturation at 94°C of template DNA, primer annealing for 1–2 min (annealing temperatures shown in the tables), and polymerization at 72°C typically for 2 min. A modification of this protocol was used to sequence the vector-insert junctions of P1 clones, using as vector primer the Sp6 or T7 promoter sequences from the P1 vector.

Somatic Cell Hybrid Mapping

All PCR assays developed from insert DNA were used to screen a human/rodent DNA panel from somatic cell hybrids containing single human chromosomes (many are of Mapping Panel 2 from the NIGMS Human Genetic Mutant Cell Repository (Camden, NJ). Human/rodent somatic cell lines used in the vector-insert junction tests were, with the exception of the line for chromosome 10 (a gift from Dr. Carol Jones of the Eleanor Roosevelt Institute), purchased from NIGMS Human Genetic Mutant Cell Repository. The following list summarizes the human chromosome component of the hybrids, followed by the NIGMS catalog number and the type of background: Chinese hamster (ch) or mouse (mo). Chr1, GM13139 (mo); chr2, GM10826B (ch) or GM11712 (mo); chr3, GM10253 (ch) or GM11713 (mo); chr4, GM10115 (ch) or GM10481 (mo, includes human chr 11); chr5, GM10014 (ch) or GM11714 (mo); chr6, GM11580 (ch) or GM10629 (ch); chr7, GM10791 (ch) or GM10790 (ch); chr8, GM10156B (ch) or GM10897 (ch); chr9, GM10611 (ch); chr10, R342-A4 (ch); chr11, GM10927A (ch) or GM11087A (mo); chr12, GM10868 (ch); chr13, GM10898 (ch); chr14, GM10479 (mo) or GM11535 (ch); chr15, GM 11418 (ch) or GM11715 (mo); chr16, GM10567 (mo); chr17, GM10321 (mo) or GM10498 (mo); chr18, GM11010 (ch); chr19, GM10449 (ch) or GM10612 (ch); chr20, GM13140 (mo); chr21, GM08854 (mo) or GM10323 (mo); chr22, GM10027 (ch) or GM10888 (ch); chrX, GM06318B (ch) or GM10324 (mo) or GM10826 (ch); and chrY, GM06317 (ch). Rodent background (control) lines were GM00345B (mo), GM00347A (mo), GM05862 (mo), GM10658 (ch), and GM10908 (ch). An assay was determined successful if the human controls amplified the correct sized product and somatic hybrid background lines did not. Somatic cell lines were judged positive for the assay if they produced a PCR product of the expected size.

Subclone Libraries for Meiotic Marker Development

Cosmid subclone libraries were prepared from TYAC47, TYAC49, and TYAC90. Partial *Mbo*I-digested fragments were ligated into the *Bam*HI site of the vector sCOS1 (Evans *et al.*, 1989). Bacteriophage λ subclone libraries for clones TYAC73, TYAC89, TYAC98, TYAC123, TYAC145, and TYAC188 were prepared in Lambda Dash II. Transfectants were screened for the presence of human genomic DNA by colony filter hybridization to ^{32}P-labeled human DNA using standard methods. Existing cosmid clones for TYAC190 (HTY146), TYAC192 (HTY255), TYAC193 (HTY275), and TYAC196 (HTY2070) were obtained from H. Riethman and used for microsatellite screening.

Microsatellite Marker Development

A. Screening TYAC subclone libraries prepared for selected clones. Subclone libraries were screened with human genomic DNA, dCA/dGT repeat (Pharmacia), and yeast genomic probes. Clones containing human and microsatellite DNA that did not hybridize with yeast DNA were used to develop microsatellite markers. CA repeats present in the clone were sequenced from the microsatellite using the PCR product and a modification of the ligation-mediated strategy outlined above. After the subclone DNA was digested with a blunt end restriction endonuclease, a PCR product was produced using a GGGG-GT_n primer (Pandolfo, 1992) that primes from the CA repeat and a linker primer. The product was sequenced and another primer selected to sequence back through the CA repeat, using the clone DNA as template. Usually the sequencing primer also could be used, with another primer developed from sequence flanking the repeat, for verification of the chromosome assignment via PCRs on the SCH panel and for the polymorphism detections. In one case, TYAC49, a cosmid clone (c1-7) was subcloned into M13 and screened for CA repeats, which were sequenced using the Sequenase V2 kit from U. S. Biochemical.

Genomic DNAs from four to eight parents of the CEPH (Centre d'Etude du Polymorphism Humain) family panel were tested for polymorphism using the PCR primers from the newly developed marker. Markers displaying more than one PCR product length in this test were used to genotype CEPH families and the genotype data used in linkage analyses. The chromosome assignment of the sequence flanking the CA repeat was also verified by amplification of the somatic cell hybrid (SCH) panel using an STS developed from a sequencing primer and a second oligo identified from sequence flanking the repeat.

B. Direct microsatellite screening of P1 clones. Microsatellite repeats were detected using P1 DNA. Following the ligation-mediated procedure described above, with Mueller/Wold linkers and the G_4GT_n primers, we developed the other microsatellite markers.

Restriction Fragment Length Polymorphism (RFLP) Screening

RFLP analysis was carried out as previously described (Donis-Keller *et al.*, 1987) using cosmid subclones (HTY275C24, HTY146C48, HTY146C3, HTY146C6, HTY2070C14, HTY2070C24) obtained from H. Riethman. Polymorphisms found using some telomere YAC subclones have been described previously (HTY146, Helms *et al.*, 1992; Hing *et al.*, 1993; NIH/CEPH Collaborative Group, 1992; HTY275, NIH/CEPH Collaborative Group, 1992; HTY2070, NIH/CEPH Collaborative Group, 1992, and Pandit *et al.*, 1995).

Genotyping Polymorphisms and Linkage Analyses

The primary 40-family CEPH reference pedigree panel (537 individuals) was used to collect genotypic data for the telomere markers. DNA was obtained from the CEPH or prepared from lymphoblastoid cell lines purchased from NIGMS or provided by Ray White (University of Utah). Microsatellite marker PCRs were optimized using a radioisotopically labeled primer on CEPH panel parents prior to collection of family genotypes. PCR conditions were as described for the SCH panel, using 50 ng of each CEPH individual template DNA and 22–25 PCR cycles including γ-^{32}P-labeled primer for radioisotopic labeling of the products. PCR products were size-fractionated on 0.4-mm-thick sequencing gels, dried to Whatman 3M paper, and exposed to X-ray film for 12–36 h. Initial sizing of alleles for each marker was accomplished using an M13 sequencing ladder electrophoresed in gel lanes adjacent to CEPH parents, and then a selection of CEPH parental DNAs representing the allele size distribution were pooled and used as allele size standards in the genotyping gels. Data were entered into a Hypercard application written for Macintosh computers (Pedigree Stacks, Six Ponds Software) that formats the data for the linkage program analysis.

The computer program CRI-MAP (Version 2.4; P. Green, unpublished) was used for all linkage analyses. Genotype data for new markers were merged with CEPH version 6, 7, or 7.1 genotype data, and our more recently acquired data. Final calculations were performed, and local support for the marker order was determined using the CRI-MAP option FLIPS2.

Fluorescence in Situ Hybridization (FISH)

Labeled probes used throughout the study included cosmid subclone DNA or DNA pools of cosmid subclones, P1 clone DNAs, and *Alu*-PCR products from TYAC clones (Lengauer *et al.*, 1992). Inter-*Alu*-PCR products were prepared from TYAC agarose bead prep DNA using the primers CL1 and CL2, derived from the 5′ and 3′ conserved regions of the *Alu* repeat element [CL1: TCCCAAAGTGCTGGGATTACA; CL2: CTGCACTCCAGCCTGGG (Lengauer *et al.*, 1992)]. *Alu*-PCRs were prepared using an approximately 100-ng equivalent of yeast DNA template (miniprep or agarose bead prep), 0.25 μM each of the CL1 and CL2 primers, 250 μM each of the four dNTPs, 10 mM Tris–HCl, 50 mM KCl, 1.5 mM $MgCl_2$, and 2.5 U AmpliTaq polymerase (Perkin–Elmer Cetus) in a 100-μl total reaction volume. After 30 cycles of PCR (5-min predenaturation at 96°C, 1-min denaturation at 97°C, 30-s annealing at 37°C, and 6-min extension at 72°C), 5-μl fractions of the PCR were electrophoresed on small acrylamide gels to determine the number of differently sized products in each reaction and estimate DNA yields. Approximately 2 μg each DNA probe was biotin-labeled via nick-translation and 100–150 ng was used in hybridization to pro-metaphase chromosome spreads prepared from phytohemagglutinin-stimulated blood lymphocytes as described in Lichter *et al.* (1988). Five to 20 μg human Cot 1 DNA (Gibco BRL) and 5 μg salmon testes DNA (Sigma) per hybridization reaction were used to reduce nonspecific hybridization. Lymphocytes from several normal male donors were used throughout the study. The probed metaphase slides were incubated with fluorescein avidin DCS (Vector Laboratories, CA) and signal amplified via incubation in 5 μg/ml biotinylated anti-avidin D affinity purified from goat (Vector Laboratories), followed by a second layer in fluorescein avidin DCS, and then counterstained in the final wash in 200 ng/ml 4,6-diamino-2-phenylindole (DAPI) and 200 ng/ml propidium iodide (PI). Up to 50 metaphases were examined for each probe. Visualization of fluorescence *in situ* hybridization (FISH) signals was accomplished using an Olympus Vanox microscope and photographed using Kodak Ektachrome ASA 400 color film. Following the FISH, cytogenetic banding Giemsa or DAPI banding patterns often were observed and photographed.

P1 Clone Screening Using the Vector-Insert Junction or Microsatellite PCR Assays

STSs found to derive from a single telomere were screened for the ability to amplify DNA pools prepared from the DuPont human genomic P1 library clones (3 genomic equivalents) (Shepherd *et al.*, 1994). STSs for 17 chromosome ends were used to screen the P1 library (for 1p, sJCW15; 1q, sAVA42; 2p sAVA22, 2q, sJCW5; 3p, sAKB8; 4q, sSR3; 6q sAVA3; 7q, sAVH6; 9q, sAVA19; 10p, sAVA28; 10q, sAVA4; 11q, sJCW13; 13q, sLDA1; 17p, sFJS14; 18q, sJCW2; 21q, sSR8; and for 22q, sAVA39). Three to four rounds of DNA pool PCR screenings were designed to provide the plate, row, and column positions of clones responsible for PCR amplification in each pool. In the final screening, P1 clones from positive wells were restreaked and individual colonies from each position were retested for the ability to act as PCR template. No more than three positives for each screen were tracked to their library well position.

BLASTN DNA Sequence Analysis

Vector-insert junction sequences and sequences flanking microsatellites developed from the TYACs were used in GenInfo Blast searches of over 530,000 nonredundant NCBI database entries (10/95) to find sequence matches having high homology (includes PDB, GenBank, GenBank updates, EMBL, EMBL updates, dbEST, dbSTS, VECTOR, etc.). Local BLASTN was performed on a Sun Sparc2 workstation using the TYAC sequences and known telomere-associated sequences from the NCBI databases.

RESULTS

Telomere YAC (TYAC) Library Construction, Screening, and Clone Sizing

A genomic YAC library was constructed using the cloning system developed by Riethman *et al.* (1989). The vector was designed with only one tetrahymena telomere under the assumption that for the YAC clones to remain stable in yeast, a second telomere would be recruited from a human DNA insert, thereby creating a human telomere sequence-specific library. We used a ligation mixture consisting of *Eco*RI partially digested human genomic DNA (CGM-1) ligated to the *Eco*RI site of the pTYAC vector, transformed the mixture into the host yeast strain AB1380 (Burke *et al.*, 1987), and plated the cells on media selecting for growth of cells containing vector sequences. A total of 11,951 clones were picked and gridded to selective medium plates. Clone DNAs were bound to nylon membranes and screened for human telomere-specific sequences by hybridization with a radioactively labeled $(TTAGGG)_{40}$ insert from the clone pHuR93 (Moyzis *et al.*, 1988). Approximately 1–2% of the transformants were positive and these were plated for single colony isolation. Five colonies for each transformant were gridded and screened again by hybridization to the telomere repeat probe. Altogether 189 independent transformants that hybridized with the probe in the secondary screening were isolated. In addition, 7 TYACs (5 described in Riethman *et al.*, 1989, and 2 additional from H. Riethman) were included, bringing the set of clones to 196.

TYAC clones were size-fractionated on CHEF gels. Southern blots prepared from the gels were probed with radioactively labeled human genomic DNA, and sizes were determined by comparison to bacteriophage λ concatemer molecular size markers. Clone sizes ranged from less than 25 kb to more than one Mb, with the mean near 200 kb. In cases where two YACs were detected in a single yeast colony (eight clones), only the larger TYAC was included in sizing counts.

Chromosome Mapping and Subregional Localization of TYAC Clones

Several approaches were used to characterize the library as a whole and to group the TYAC clones by chromosome origin. CHEF analysis identified a set of clones ranging from 100 to 400 kb that were selected for initial chromosome mapping studies. Eight clones that appeared from CHEF analysis to contain more than one TYAC, suggesting either molecular rearrangements or independent TYACs within the same yeast cell, were not studied further. The principal means used to assign TYACs to chromosomes were FISH analysis and mapping TYAC STSs against a panel of genomic DNAs from rodent/human somatic cell hybrid lines. For the most part, these localization results directed additional experiments to develop meiotic markers and construct regional linkage maps.

Comparison of the DNA sequences used to develop TYAC STSs to previously described DNA sequences using BLASTN provided another means of obtaining chromosomal localization information for homologous sequences and detailed information on TYAC sequence content.

TYAC chromosome assignments from SCH panels and TYAC STSs. A ligation-mediated PCR strategy was implemented to sequence human inserts at the TYAC vector and human insert junction (VIJ) from each of 182 clones. At least one working assay was developed for 130 TYACs. In addition, 12 STS assays that reveal simple sequence repeat polymorphisms (SSRPs) from 12 TYACs were developed and are described separately. Characteristics of these microsatellites are shown in Table 1, which includes 17 microsatellites that map to telomeres. Altogether we developed 149 working STS assays for 132 different TYACs.

Genomic DNA from a panel of 40 SCH human/rodent hybrid cell lines that contain, for the most part, single human chromosomes within either a hamster or a mouse genome background (Dubois and Naylor, 1993) was used to chromosomally localize the TYAC STSs. Most assays were repeated with appropriate control assays with human genomic DNA, with rodent cell genomic DNA, and without genomic DNA substrate. The results for 36 TYACs that map to specific telomere regions using STSs developed from VIJ sequences are shown in Table 2. One TYAC (193) was mapped against a SCH panel using RFLPs. For 94 of the STS assays a single PCR product of the expected size amplified one human chromosome from the panel. In other cases a dominant PCR product and one or more weaker PCR products of the expected size amplified more than one SCH cell line, and in other cases multiple amplifications of equal intensity from more than one SCH cell line were apparent. For 50 STS assays more than 1 chromosome appeared to be positive. Of these 50 assays, 19 detected 2 or 3 chromosomes. In multiple cases all human chromosome-containing SCH cell line DNAs were amplified. In 2 cases (TYAC20 and TYAC46) the rodent background amplified the expected size PCR product, precluding human chromosome assignment, but perhaps indicating a conserved gene sequence. For 10 TYACs more than one STS had been developed from different regions of the clone. The STS assays amplified the same chromosomes for 9 of these TYACs while 1 of the 2 TYAC123 STS assays amplified chromosomes 10 and 20, and the TYAC123 polymorphic STS amplified only chromosome 10.

P1 library telomere clone screening using TYAC derived STSs. Initial FISH mapping experiments assigned a group of TYACs to telomere regions and for some of these we screened the DuPont P1 library (Shephard *et al.*, 1994) to provide a more convenient source of clone DNA for polymorphic marker development and to improve the quality of signal for subsequent FISH mapping. For 16 selected TYACs we identified 39 P1 clones using 19 STS screening assays developed from TYAC sequence. At least 2 clones were recovered for each of these assays as expected from a 3-genome-equivalent library. However, P1s were not found for 3 STS assays for TYACs derived from chromosomes 1q, 3p, and 13q. The P1 library clones averaged 80–85 kb in size and were used directly for FISH hybridizations (Fig. 1, Table 2) and for polymorphic marker development (Tables 1 and 2, Fig. 2). Additional STSs were developed from some the P1 clones (17p, 1q, 3 for 7qter), and chromosomal localization was confirmed by a combination of FISH mapping, amplification of the SCH panel, and linkage analyses.

FISH mapping TYACs and P1s. The FISH subregional chromosomal localization was accomplished using several types of hybridization probes derived from the TYACs. For the most part we used *Alu*-PCR products generated from TYAC clones as probes according to previously published procedures (Lengauer *et al.*, 1992). In addition, when available, cosmid or bacteriophage λ subclones of the TYACs were used as individual clone probes or groups of subclones were pooled and used as a probe source. For example, for TYACs from which polymorphisms were derived from subclones, chromosome localization was verified by FISH mapping with the subclone (Table 2, Fig. 1). P1 clones that were identified using TYAC STSs were used directly as probes and in general provided stronger signal than *Alu*-PCR products from TYACs. Approximately 50 metaphase chromosome spreads from freshly prepared lymphocytes from whole blood samples were examined for each probe. Each of the 181 clones containing a single TYAC from the library and the 7 other TYACs included in the study were tested for chromosomal localization using at least one probe source and each experiment was performed at least twice. Localization results were obtained for most of the TYACs (Table 2, Fig. 1, and data not shown). For 59 clones, hybridization to a single terminal cytogenetic band or multiple telomeres was observed as the principal or minor component. For 41 TYACs, hybridization signal derived entirely from a single nonacrocentric telomere while in 18 other cases multiple nonacrocentric telomeres were detected. However, for 87 clones interstitial signals or hybridization to centromere regions were evident with no detectable signals within telomeric bands. For 13 clones, signals were evident from both telomeric and interstitial regions. In addition, six clones hybridized specifically to single chromosome centromere regions (2, 6, 7, 9, and 16) (data not shown). In a number of cases where multiple localizations were observed, differential intensity of signal was evident.

Meiotic marker development and linkage mapping. TYACs were selected for meiotic marker development in cases where the SCH panel analysis and FISH consistently mapped the clone to a single telomere. In the early phases of the TYAC library characterization studies, bacteriophage λ or cosmid subclone libraries were

TABLE 1

Microsatellite Markers Mapped near Human Telomeres Derived from TYAC/P1 Clones

Chr Tel	Clone(s)	STS name (locus)	Repeat type[a]	GenBank Acc. No.	BLASTN server result	PCR primers (5′ → 3′)	Anneal temp.	No. Alleles	Size range (bp)	% HET	cM to a proximal locus
1q	P1-2428 (TYAC96)	sJCW14	(CA)11	U53008	U	CCAGTATATAATAGAAACGCAGAACC GGGCACATGTTTCTTGTAGATTTCTG	57	2	169–170	<32	ND
2p	47.2p2 (TYAC47)	sRA11	(CA)21	U53009	U	CTCCACTTAAGCTTGGTTTACA CCCTAAAAGCCTCACTACATG	58	9	119–135	82	4.7 cM to D2S323 (SSRP)
6q	145.20 (TYAC145)	sAVA3	(CA)21	U53019	U	GGCTAATAAATGCTTAGAGCC GGGCTTAGTTGTTTTCCATTAG	58	9	100–116	74	0 cM to D6S281 (SSRP)
7q[b]	HTY146c30A (TYAC190) P1-2681, P1-2682	sAVH6 (D7S594)	(GT)33	U53014	G, r	GCCCAGCCTGTGTATGGTTA CTAAAGGAAGTTCTACAGG	58	10	217–235	84	4.1 cM to D7S559 (SSRP) (6.0 cM to D7S559 when RFLPs in map)
10p	73.78 (TYAC73)	sAVA2	Cx	U53011	U	GGCTTAAAAGTAGGTAATTAC CGTCTTGGAACTTTCAGAAT	55	14	151–185	81	2.5 cM to D10S249 (SSRP)
10q	123.69 (TYAC123)	sAVA4	(CA)15	U53013	G, H[i]	AGTGCATGAGCCAATTCCTTA GCTGGTCCCTATCATTAGGTAG	57	17	161–197	59	9.6 cM to D10S212 (SSRP)
11p[c]	yRM2209 (TYAC176)	p194B (D11S2071)	Cx	U12896	nh	AGGGCAATGAGGACATGAAC ATGTGGCTGGTCCACCTG	60	18	168–202	85	<5% recombination with HRAS (Browne *et al.*, 1995)
11q[d]	P1-2388 (TYAC22)	sJCW13	(CA)14	U53017	R, G[i]	TATTGAGTGTCCCTGCTCCC AAAGTTTCCTTGTGGCCCTT	57	8	197–215	83	4.3 cM to D11S912 17.2 cM from distal 11q marker D11S968
12p[d]	c1-7M5 (TYAC49)	sJKS1	Cx	U53010	H, A	GGGTGACAGAGCAAGACTCTG GGTGGTGAATTTGTCACTTTCT	64	14	119–161	74	21 cM from distal 12p marker D12S352
13q	HTY255c26 (TYAC192)	sLDA1 (D13S103)	(GT)13	U53015	G[i]	TACAAAGCACTGTGAGAAATGG ATGCTGGAGCAAGCTAGGAC	56	5	100–150	36	Near 13qter-not uniquely placed
14q[e]	HTY2070c2e (TYAC196)	sCAW1 (D14S826)	(CA)17	U29384	R	TGCTGTTGGACTCAGGTAGCTA TCTCTAAAGCTACTATAACCCAG	61	7	145–161	74	2.9 to D14S543 (SSRP)
18p	89.115 (TYAC89)	sAVA5	Cx	U53012	h	GCCTCATTTTCATGGAGGTGG CTCATTTTTCAGTCCATCTGCAAG	62	8	225–237	80	0 cM to D18S59 (SSRP)
18q[f]	TYAC118	CU18-010 (D18S497; =D18S70)	(CA)17	Z17140	U	ATTGCCATTCAAGGCTGAAC GTTTTGGGAATGTCAAGAAGTACC	55	14	118–144	84	=D18S70 (SSRP locus); 8 cm from D18S461 (Gyapay *et al.*, 1994)
21q[g]	TYAC24	21qter-qt3 (D21S1575)	Cx	X83559	R	GAAACCCATCTCACATGCAG GAAGTGCTCTAAGAACTTGC	55	11	311–343	78	(Blouin *et al.*, 1995)
22q	P1-2391 (TYAC104)	sJCW16	(CA)10	U53018	R	TTGCAGACAGCAGACTACAGG TTCAGTCTGTGGCTGTCCAG	56	5	194–210	76	16.3 from D22S274 (SSRP)
Xq[h]	c250-9k (CGM250)	sDF2 (DXS1108)	Cx	ND	—	ACTAGGCGACTAATACAGTGGTGC GTGAATTCATCATATGTGATTTCC	65	7	163–177	75	1.6 cM to DXS52 (VNTR)
Xq/Yq[h]	c364-10g (CGM364)	sDF1 (DXYS154)	(GT)18	U53016	—	GGCCTGAATTCATTTATTATTCTAATAG GAACAGGCAAAGATGCCCACTCTC	65	10	220–244	70	1.3 cM to DXS1108 (SSRP)

Note. The corresponding TYAC clone is shown in parentheses beneath the name of the clone from which the markers are derived, except the Xq/Yq markers, which were derived from CGM library clones that map near the telomere. The 7qter P1 clones contain the microsatellite sequence. BLASTN analysis was performed using the GenInfo server default parameters.The interpretation of the BLASTN column symbols is as follows: uppercase letters indicate the database match at $\geq e^{-6}$, while lower case indicates $10^{-6} \geq e \geq 0.75$; U, unique sequence; T, subtelomere sequence homology; R, common human repeat homology; NH, nonhuman sequence homology; G, human gene(s), EST or cDNA homology; A, poly(A) in sequence; H, human sequence homology (i.e., with an STS, or no known gene). ND, not determined.

[a] Type of microsatellite repeat and the number of repeats found in the clone sequenced; Cx, complex repeat.

[b] Telomere marker we have reported previously (Hing *et al.*, 1993).

[c] Telomere marker developed by Browne *et al.* (1995) is present within TYAC.

[d] TYAC polymorphism markers are not the most distal markers in the linkage maps; their distance to the physical end is yet to be determined.

[e] Telomere marker we have reported previously (Pandit *et al.*, 1995).

[f] Telomere marker developed from TYAC by R. Straub (described in Van Kessel *et al.*, 1994); in this case the primer sequences flanking the CA repeat also flank the CA repeat from the Généthon marker D18S70.

[g] Telomere marker developed by Blouin *et al.* (1995) is within TYAC sequences.

[h] X and Y polymorphisms we have reported previously (Freije *et al.*, 1992).

[i] Highly significant match, and is discussed in the text.

constructed and screened for RFLPs and microsatellite repeats. Lambda subclone libraries were prepared for 6 TYACs (73, 89, 98, 123, 145, and 188) and cosmid libraries for 3 TYACs (47, 49, 90). In addition a set of existing cosmids (from H. Riethman) for 4 TYACs (190, 192, 193, and 196) were also screened for the ability to detect polymorphism. Later P1 clones were utilized and microsatellite polymorphism markers developed directly from P1 DNA using a modification of the VIJ method that we implemented for STSs.

Microsatellites. A total of 12 dinucleotide repeat polymorphisms were developed from 12 TYAC clones or their P1 derivatives (Table 1, Fig. 2). Three of the CA-positive subclones we sequenced consisted of complex repeats. For example, sJKS1 included a $(GAAA)_5(TA)_4(CA)_{19}$ element that was found to be highly polymorphic in the CEPH pedigrees (14 alleles, Het 74%). The other 9 microsatellites sequenced contained simple repeats that ranged in length from 11 to 33 CA dinucleotides. Ten of the microsatellites were highly informative with heterozygosities of at least 74%, while two markers had relatively low heterozygosities (32 and 36%). In addition, two microsatellites have been developed by oth-

TABLE 2

TYAC Library Clones Mapping to Specific Telomere Regions of Chromosomes: Localization by Methodology, Clone Size, STS Primer Sequences, and BLASTN Results

Chromosome Assignment by Method			YAC	Size Kb	STS †	GenBank Acc. No.	Primer Sequences	PCR Prod. Size	Anneal Temp °C	BLASTN Server Result
SCH panel	FISH sublocal.	Meiotic Linkage								
1	1p36.3		TYAC179	200	sJCW15	U57699	CACGCCACACAAGCTGAC GGAGTTTGCATAGACAGTGAAGA	101	56	U
1	1q44		TYAC110	300	sAVA42	U57701	CTTCTCTTCTGCTTTCCACACC TGGAAGCTCCAAGATAACTTCC	106	56	T*
2	2pter	¥ 2pter (sRA11)	TYAC47	350	sAVA22	U57702	TCTCCATTATGCTTAGGTCTCC GCATCAATCTGTGGAAAATG	133	53	nh
2	2pter	¥ 2pter (sRA11)	TYAC47*	350	sMC1	U57703	ATCTCCATTTGCTTAGGTCTCC AGCATCAATCTTGTGGANGATG	145	50	nh
2,1	2q37		TYAC75	75	sSMM1	U57704	TTCACATTTTAAGATATACTGAGG TGTATNCTGAAAACAAAATGGTACT	108	58	g
2	2qter		TYAC99	250	sJCW5	U57849	TTAGCAGGGGACCCAGTG CATCAGGTTCCTGTGGGTG	113	56	t,nh
2,8	2qter/8pter+	2qter, 8pter (HTY275C24)	TYAC193	150	ND	--	--	--	--	--
3	3pter		TYAC148	300	sAKB8	U57700	ACAGAACCAGAATACCTGAGAAGG TGCAAAAAGAAGTACTACACTGCC	165	58	U
3	3qter		TYAC162	250	sAVA54	U57850	TGTACAGGACCAGCTCCACA AGAGTTCTAATTCTCACTTTGTAAGG	96	56	U
4	4qter		TYAC18	100	sSR3	U57851	AATTCCCCCTCATAGCTGTC GAGAGCTGCCTCTCTTGATG	117	56	U
5,6	5qter		TYAC125	50	rsAKB1	U57852	GGAGTGCTCACCACCATT GGGGTAAATGTAGGTGGAGA	106	60	U
5	1p36.3,1p31.1,1q42.1q44, 3q29, 4q27,5q35,9q34,16p13.3,17p13		TYAC57	275	sAVA25	U57853	TGATTGGTACTATGTCAGGGTCC ACGTAACTGTTAACATCCTTTCCC	104	56	g
6	6qter	¥ 6qter (sAVA3)	TYAC145	250	sFJS21	U57854	TGATAGTGTAGAATTCGGCCG TGTAAATGTGCTGAGTGCGG	103	52	T*
7,18	7p22, 8p23, 16p13.3, 2q27, 6q27		TYAC21	50	rsAVA14	U57855	TTAAGCAGAGACACTACTTTGAAGG TATCACGAGGCCTTCGTCTT	214	55	nh
7	7q36	¥ 7qter (sAVH6)	TYAC6	275	sAVHa	U57856	TGTTAAACCCAGGACTTTTTCC TTGTTTTTACCGACCAAGGC	131	53	r
7	7qter		TYAC87	400	sMO1	ND	ACACTGTTCCANAAGACAG AGGGATATTGGTCTATAG	125	56	U
ND	7q36		TYAC109	275	ND	--	--	--	--	--
ND	7q36		TYAC130	250	ND	--	--	--	--	--
CHO	8p23		TYAC20	150	rsAVA13	U57857	TCAAGAATTCCCAGCGATG TCTCTGGCCAAATCAGCC	152	56	G
ND	8qter		TYAC48	200	2058V1	U11829	--	--	--	--
9	3q29, 9q34.3, 11p15		TYAC43	100	sAVA19	U57858	TGAATCTCCACTCGCATCTG CTGGCTACCCTGAAACTACCC	130	55	g

Note. SCH, rodent/human somatic cell hybrid panel. Where multiple FISH signals were detected, the strongest FISH signals are underlined and all hybridization signals to telomeric bands are listed; however, a "+" indicates that an additional interstitial signal(s) was detected. ND, not determined; CHO, somatic hybrid background line (Chinese hamster ovary) was positive for the assay.

† Marker maps proximal to at least one other microsatellite marker.

^ Marker developed in another laboratory maps to TYAC. BLASTN analysis was performed using the GenInfo server default parameters. The BLASTN result is coded with uppercase letters indicating a database match at $\geq e^{-6}$, and lowercase letters indicating $10^{-6} \geq e \geq 0.75$. U, Unique sequence; T, subtelomere sequence homology; G, g, gene(s) homology; R, r, repeat(s) homology; H, h, nongene human sequence homology; nh, nonhuman sequence homology.

¥ Chromosome end has a corresponding SSRP marker developed from the TYAC sequences; in the case of chromosome 7, the SSRP marker D7S594 is derived from sequences closest to HTY146C6.

* Highly significant match and is discussed in the text.

ers from TYACs that we mapped to chromosomes 11p (TYAC176) (Browne *et al.,* 1995) and 18q (TYAC118) (Van Kessel *et al.,* 1994) (Table 1). We also determined that a previously described microsatellite for 18q (Gyapay *et al.,* 1994) is the same as that developed from the 18q TYAC118 and that one other previously described microsatellite from 21q (Blouin *et al.,* 1995) is contained within TYAC24 (Table 1). Each microsatellite marker amplifies a single chromosome from the SCH panel. Figure 2 shows representative PCR assay results for 7 of these linkage markers.

RFLPs. Cosmid subclone libraries were screened for RFLPs using methods we have previously reported (Schumm *et al.,* 1988). Useful polymorphisms were identified from three of the TYACs that map to chromosomes 2q, 7q, 8p, and 14q (Table 3). The 7qter

TABLE 2—*Continued*

10	10pter	¥ 10pter (sAVA2)	TYAC73	325	sAVA28	U57862	TTCTTGCTTCAGCCGGAG TCCGGAACTCACATCGGT	131	55	G*
10, 20	10qter	¥ 10qter (sAVA4)	TYAC123	275	rsFJS16	U57859	AGATTGTGTTGAGGTGATGGG AGGGTATGCAGATGTCTGTGG	105	60	h
ND	10q26		TYAC93	250	ND	--	--	--	--	--
11	11q25		TYAC59	100	sAJW5	U57860 U57861	GGGTTAATTGAGCTCACTCCC GTAAACAAGCAAGACTCTGGCC	142	59	U,g (2 sequences)
11	11qter	† ¥ 11q25 (sJCW13)	TYAC22	50	sJCW1	U57863	TCAAGAATTCCCAGCGATG TCTCTGGCCAAATCAGCC	152	56	r,g
12	12pter	† ¥ 12p13 (sJKS1)	TYAC49	>950	sJKS3	U57864	AATTCCATCCACTGTGGCTC TGGAAGACATATGAGTTGCTGC	103	60	U
12	12pter		TYAC14	100	sAVH27	U57865	CCTCGGTTTAGCAGAGCAGC GGGAGAATGGGGTCTCTCC	140	60	U
16	16qter; 15q11		TYAC135	100	sAVA47	U57868	CCCAGGCACAGAGAATCATT CGTCCTACAGAAGCTCTTCTACG	110	55	U
17	17pter		TYAC98	100	sFJS14	U57869	TCCAGCTAGAACTTCGTGCA GCACATCCTACAGTAGCTGCC	115	60	U
17	17q25; 6q27; + [repeated 17q11++, 17q25+]		TYAC3	175	sSR5	U57870	GAATTCAGGAGCCTGAGAGG TGTATGGCTAGTATGCTACAGC	150	55	U
18	18pter; 10pter; 9qter; Yq interstitial	¥ 18pter (sAVA5)	TYAC89	225	sAVA33	U57871	TGTAAGACAAGAAGGGGTGG CTTTTACCCCTCAATTTATTTG	117	51	T*
ND	18p		TYAC128	200/250	ND	--	--	--	--	--
18	18q23		TYAC45	300	sAVA20	U57872	CCACGAGAACAGTGAACATCC TCAGAACCATGAGTGAAAGTGC	101	58	T*
18	18qter	^¥ 18qter (D18S70)	TYAC118	150	sJCW2	U57873	CTTGTCTTCAAGACTCATCAGACC GTTCGTAAGATACATCAATGCAGC	158	56	h
ND	18q23		TYAC41	250	sFJS1	U57874	--	--	--	U
ND	20pter		TYAC195	180	ND	--	--	--	--	--
21	21qter	^¥ 21qter (D21S1575)	TYAC24	100	sSR8	U57875	ATGTGATGGGCAGGTATTTG AAGTCTTCTGTGCCCTCAGC	138	55	T*
22	22qter	¥ 22qter (sJCW16)	TYAC104	75	sAVA39	U57876	CTAGCCTGTGTTCTTTTTTGGG AAGGTAAGACCGACAGAAAGACC	130	56	nh

TYAC146 was found to contain five RFLPS, one of which is a highly informative (HET 92%) variable number tandem repeat (VNTR) (Hing *et al.*, 1993). We identified five RFLPs from TYAC193 and found by Southern blot analysis of somatic cell hybrid DNA that three of the RFLPs hybridized to DNA from 2q-containing cell lines while the remaining two RFLPs mapped to chromosome 8-containing cell lines (data not shown). We and others subsequently determined that the TYAC193 sequence derives from the telomere of 2q (Anker *et al.*, in preparation; Macina *et al.*, 1994). In addition, we detected telomere deletions using cosmid subclones from this TYAC that will be described more fully elsewhere (Anker *et al.*, in preparation).

Linkage mapping. All polymorphisms shown in Tables 1 and 3 mapped by linkage analyses to the same chromosomes that had been identified by SCH and FISH mapping. Altogether 19 proterminal regions are marked by these informative polymorphisms. Regional linkage maps for the distal regions of 6 chromosomes (6q, 10p, 10q, 11q, 18p, and 22q) are shown in Fig. 3. In addition, the distance between each of the TYAC derived markers and the next proximal marker in 19 regional telomere maps are indicated in Tables 1 and 3. For the distal regions of chromosomes 2q, 6q, 7q, 8p, 10p, 10q, 11p, 14q, 18p, 18q, and 22q the TYAC-derived markers represent the most distal loci of the linkage maps and in some cases they extend the maps considerably beyond previous maps for these chromosomes. For example, sJCW16, which maps to 22qter, extends the endpoint of the meiotic map by 16.3 cM. With the exception of 2q and 8p, which are terminated with RFLP polymorphisms, 11 telomeres (2p, 6q, 7q, 10p, 10q, 11p, 14q, 18p, 18q, 21q, 22q) now have highly informative microsatellite markers derived from TYACs that serve as the meiotic endpoints for these chromosomes. In collaboration with Heighway and colleagues we have also mapped a 17pter PCR-based RFLP to the most distal region of a 17p linkage map (White *et al.*, 1996).

In four instances TYAC polymorphisms did not map to the most distal positions of the meiotic maps, although all four TYAC clones were physically localized within terminal cytogenetic bands. The 1q polymorphism, sJCW14, was very weakly informative (HET less than 32%), and since the two alleles differ by a single basepair, it was difficult to genotype. Therefore this marker was not fully genotyped and incorporated into a 1q linkage map. While the 11q microsatellite, sJCW13, mapped within the most distal cytogenetic band (Fig. 1 and data not shown), meiotic mapping placed sJCW13 17.2 cM proximal to the most distal

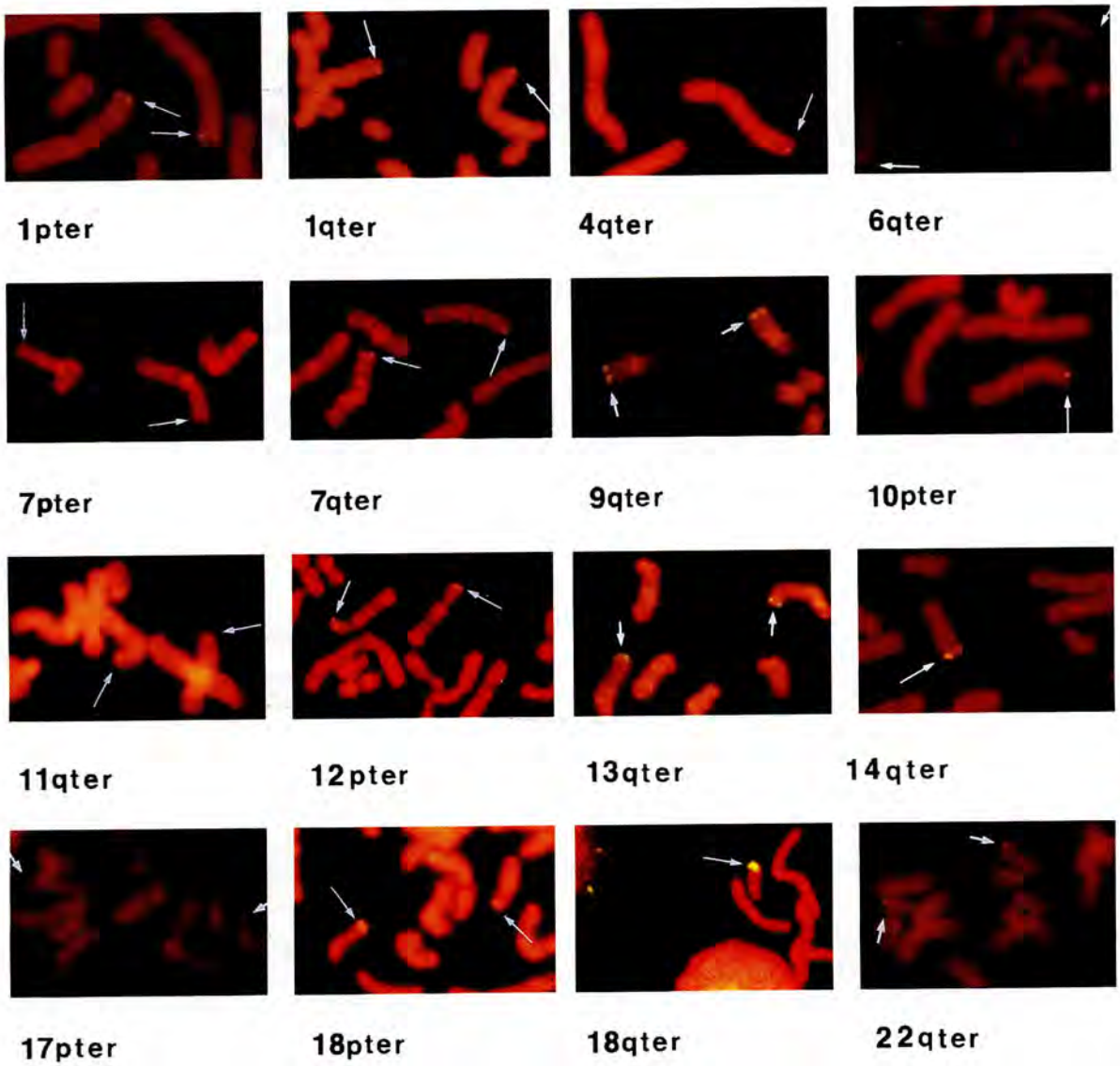

FIG. 1. FISH mapping of 16 TYAC/P1 clones to telomeres. FITC signals at telomeres are indicated with white arrows. The panel results were obtained using biotin-labeled DNA probes prepared from the following clones: for 1pter, the P1 clone W30-H2 (identified using TYAC179 STS); for 1qter, the P1 clone 2428 (identified using TYAC96 STS); for 4qter, the P1 clone 2663 (identified using TYAC18 STS); for 6qter, the cosmid C20 (subclone of TYAC145); for 7pter, the *Alu*-PCR products from TYAC21; for 7qter, cosmid 3 (subclone of TYAC190); for 9qter, the P1 clone 2417 (identified using TYAC43 STS); for 10pter, the *Alu*-PCR products from TYAC73; for 11qter, the P1 clone 2388 (identified using TYAC22 STS); for 12pter, cosmid 7 (subclone of TYAC49); for 13qter, a cosmid pool of clones C22, C26, C29, and C31 (subclones of TYAC192); for 14qter, cosmid 24 (subclone of TYAC196); for 17pter, the P1 clone 2629 (identified using TYAC98 STS); for 18pter, a cosmid pool consisting of clones C15, C30, C35, and C37 (subclones of TYAC89); for 18qter, the *Alu*-PCR products from TYAC45; for 22qter, the P1 clone 2391 (identified using TYAC104 STS). The weak 7pter signals do not photograph well, but all results were verified through the multiplicity of metaphases examined for each hybridization experiment.

marker, D11S968 (Fig. 3). Linkage mapping placed the marker sJKS1 derived from the 950-kb TYAC49 at the terminus of the 12p meiotic map. FISH mapping also placed the cosmid subclone containing the microsatellite sequence at 12pter (Fig. 1); however, additional FISH mapping of subclones from TYAC49 and other independently nonpolymorphic clones showed that six of these clones mapped more telomeric than TYAC49-derived clones (Kobayashi *et al.*, 1994). The microsatellite marker for 13q, sLDA1, was found to be weakly informative (HET 36%) and, although it mapped by FISH to the 13q telomere (Fig. 1), it was not uniquely placed with odds of 1000:1.

DNA Sequence Analysis Using BLASTN

Each of the DNA sequences determined from the vector-insert regions of the TYACs and sequences flanking microsatellite repeats from the TYAC/P1 clones were compared to over 530,000 nonredundant NCBI database entries using the BLASTN sequence analysis program. The 164 independent sequences from 136 TYACs and 4 different TYAC-derived P1s ranged in length from 50 to

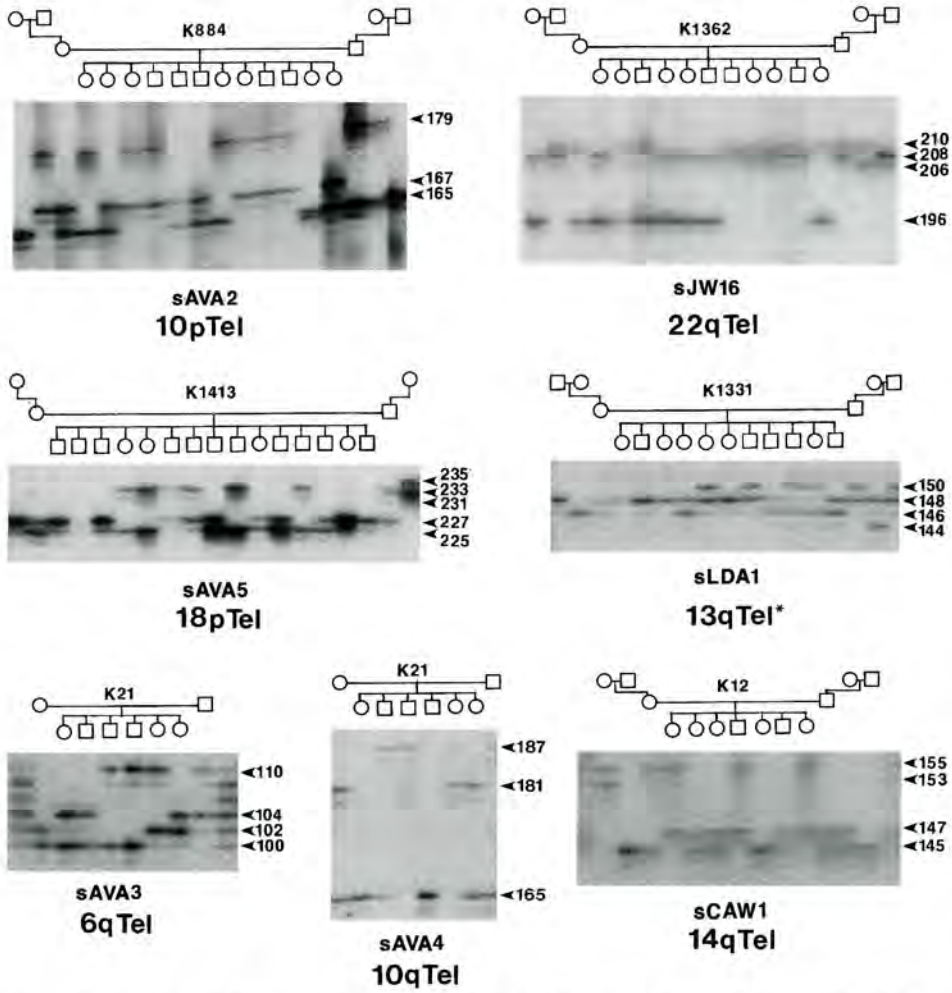

FIG. 2. Representative gels showing allelic segregation through CEPH reference pedigrees for seven TYAC/P1 microsatellite markers. PCR product sizes (in bp) for alleles observed in each pedigree are indicated to the right of the gel image. Linkage maps with five of these markers (sAVA2, sJCW16, sAVA5, sAVA3, and sAVA4) are shown in Fig. 3; sLDA1 was not mapped to a unique telomeric position, and sCAW1 may be found in the 14 map presented in Pandit *et al.* (1995).

560 bp. We also included 3 sequences that we had not developed but that were found in the TYAC clones. Altogether more than 30,000 bp of TYAC sequence were tested for homology to database entries. Matches to the database entries were placed in groups that include reported telomere sequences (T), human repeat sequences (R), human genes (G), nonhuman sequences (NH), and unique sequence (U) (Tables 1 and 2 and data not shown). An "*e*" value of 0.75 was set as the threshold for significance for BLASTN searching. We distinguished between highly significant homologies (designated by a capital letter in Table 2), e^{-6} or greater, and lower homologies (lowercase designation in the tables). In addition, each sequence was reevaluated for consistency of identity by inspection of the BLASTN output and table code letters were modified accordingly.

BLASTN analysis of TYAC sequences mapped specifically to telomere regions. Altogether 49 sequences were analyzed from 38 TYACs representing 29 telomere regions. About one-half (20) of the 38 TYAC sequences did not find a significant match in the databases and were therefore assigned to the unique sequence (U) category. Several previously reported ESTs were identified from TYAC sequences that mapped near 3 telomeres: 8p (TYAC20-rsAVA13), 10p (TYAC73-sAVA28), and 13q (TYAC192-sLDA1), suggesting the presence of genes within telomere regions. TYAC192 identified 4 ESTs with strikingly high homology (80–100% to each sequence). Two of these ESTs were deposited by the WashU-Merck EST Project and the third was deposited by Genexpress. Since the TYAC192 sequence spanned only 325 bp of genomic DNA, it is likely that all 3 ESTs derive from the same gene. In fact, comparison to overlaps with our sequence indicates that two of the ESTs probably represent opposite strands of the same sequence. TYAC73 sequence matched with 95% identity to a WashU-Merck Project EST, providing evidence for a coding sequence very near the telomere of 10p. The TYAC20 sequence (190 bp) demonstrated identity of 161/184 nucleotides (87%) to a designated human "gene" sequence reportedly mapped to chromosome 4p, although the mode of map-

TABLE 3

RFLP Markers Mapped near Human Telomeres Derived from TYAC Clones

Chr.	Locus	RFLP name	Enzyme	RFLP type	No. Alleles	HET	cM to a proximal locus
2q	D2F69S1	HTY275C24 (TYAC193)	*Eco*RI	ND	6	61	3 cM to D2S125 (SSRP)
	D2F69S1	HTY275C24 (TYAC193)	*Hind*III	ND	6	61	3 cM to D2S125 (SSRP)
	D2F69S1	HTY275C24 (TYAC193)	*Msp*I	ND	6	65	3 cM to D2S125 (SSRP)
7q[a]	D7S593	HTY146C48-1 (TYAC190)	*Hind*III	ND	3	50	4.5 cM to D7S559 (SSRP)
	D7S593	HTY146C48-2 (TYAC190)	*Hind*III	ND	4	31	4.5 cM to D7S559 (SSRP)
	D7S593	HTY146C48-3 (TYAC190)	*Hind*III	ND	5	44	4.5 cM to D7S559 (SSRP)
	D7S591	HTY146C3 (TYAC190)	*Msp*I	VNTR	>20	92	5.4 cM to D7S559 (SSRP)
	D7S592	HTY146C6 (TYAC190)	*Msp*I	Mono	1	34	6.0 cM to D7S559 (SSRP)
8p	D8S69S2	HTY275C24 (TYAC193)	*Eco*RI	ND	2	31	3 cM to D8S201 (SSRP)
	D8S69S2	HTY275C24 (TYAC193)	*Hind*III	ND	2	37	3 cM to D8S201 (SSRP)
14q[a]	D14S82	HTY2070C14 (TYAC196)	*Msp*I	ND	6	70	2.9 to D14S543 (SSRP)
	D14S82	HTY2070C24 (TYAC196)	*Eco*RI	ND	2	30	2.9 to D14S543 (SSRP)

Note. VNTR, variable number of tandem repeat polymorphism; RFLP, restriction fragment length polymorphism; SSRP, simple sequence repeat polymorphism; ND, not determined.

[a] Chromosome end has a corresponding SSRP marker developed from the TYAC sequences; in the case of chromosome 7, the SSRP marker D7S594 is derived from sequences closest to HTY146C6.

ping is not evident from this database deposit (GenBank D38378) or whether, in fact, the sequence is known to code for a gene. Six of the TYAC sequences match with strikingly high homology to previously reported sequence for telomere clones from 1q (Negorev *et al.*, 1994), 2q (Macina *et al.*, 1994), 6q, 18p, 18q (Strathdee *et al.*, 1994), 21q (Reston *et al.*, 1995), and are likely to be TYAC clones from the library that we describe in this report.

Of the 15 telomere region sequences flanking microsatellite repeats that were tested (Table 1), 3 appeared to be unique sequence (from TYAC47-sRA11, TYAC73-sAVA2, and TYAC96-sJCW14) since no homologous sequences were found in the databases. However, 5 of the

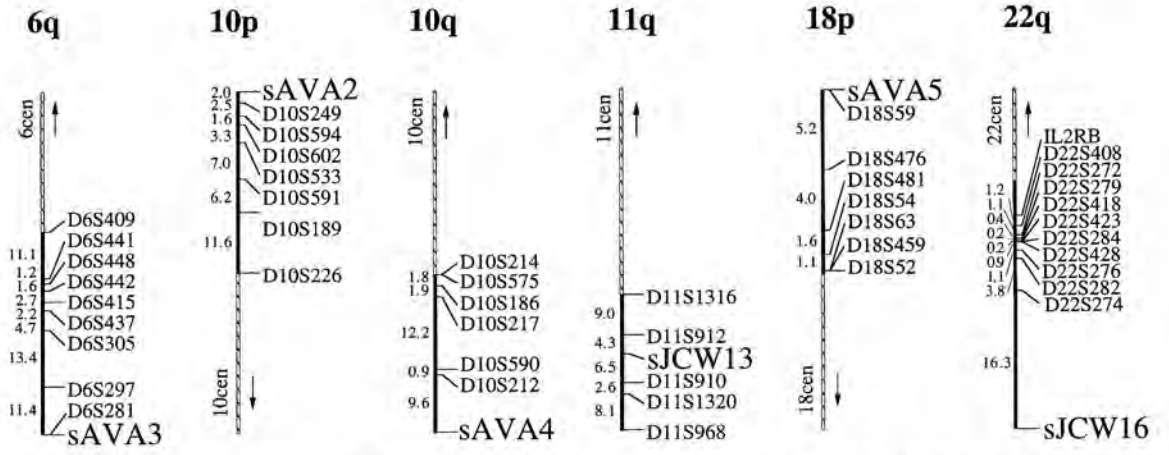

FIG. 3. Regional proterminal multipoint linkage maps incorporating six TYAC SSRPs. The TYAC markers are shown in large font. TYAC SSRPs were mapped with respect to Généthon markers (Gyapay *et al.*, 1994) and in the case of 22q, two additional proximal markers (IL2RB and D22S408). Other markers in the region that were not uniquely placed using the criteria of 1000:1 odds for order are not shown. Map distances are in Kosambi cM.

sequences [from TYAC22-sJCW13, TYAC104-sJCW16, TYAC123-sAVA4, TYAC190-sAVH6, and TYAC192-sLDA1 (described above)] found matches with reported genes, although in some cases the homology was not highly significant. TYAC123-sAVA4 (164 bp) identified two ZN α 2 glycoprotein gene sequences with 89% identity found over the same 56-bp segment. Interestingly, the TYAC 123 sequence that maps to the terminus of chromosome 10q also recognized a large number of microsatellite repeat markers (>50) from many regions of the genome and a number of D segments of unknown function. This finding suggests that there are sequence elements that are repeated in many regions of the genome that contribute to the overall structure of the microsatellite regions yet have the characteristic of local uniqueness in sequence to support sufficient specificity for PCR amplification and enable mapping of the microsatellites to specific locations. The sequences recognized by the microsatellite markers, with homologies of 85–97% over segments of 20–50 bp, often included part of the sequence homologous with the ZN α 2 glycoprotein gene sequence. Similarly, the microsatellite flanking sequence from sJCW16 (TYAC104) also identified a number of cDNAs and microsatellite sequences. TYAC22-sJCW13 identified a WashU-Merck Project EST with homology to cytochrome P450 and a large number of other ESTs with homology over ~50-bp regions. This homology may be due to the presence of PTR5 and MER12 repeat elements that are reportedly present in some of these sequences. The microsatellite flanking sequence from the chromosome 14q terminus marker sCAW1 (TYAC196) identified L1- and KPN-like repeats in addition to identifying sequence from the 4p terminal Huntington disease locus region.

BLASTN analysis of other TYAC sequences. BLASTN analysis was performed for an additional 103 TYAC sequences. Of these sequences surveyed, 49 appeared to be unique, 18 identified repeat elements, and 12 identified reported human gene sequences. Several of the gene sequence matches are noteworthy. For example, TYAC40-sAVA18 (U58875) matched to a human transmembrane receptor mRNA, also known as neurotropic protein tyrosine kinase type 2 (NTRK2), with 95–98% identity for 170 of our 218-bp input sequence. This gene has been recently mapped to 9q22 (Nakagawara *et al.,* 1995), which is consistent with our SCH chr 9 assignment for sAVA18 and with one of the dominant FISH signals that we observed (weaker signals to multiple telomeres were also observed). TYAC160-sAVA53 (138 bp) (U58877) matched with almost complete identity to β tubulin (TUBB) sequence from mouse, human, and TUBB sequences from number of other organisms. There are at least 3 functional copies of TUBB in human and many additional pseudogenes. Our SCH panel mapping and FISH did not identify the chromosomal regions reported for the human TUBB gene and pseudogenes (6p, 8, and 13), and it is possible that our TYAC sequence may identify additional locations for TUBB sequences. TYAC33-sMSH8 (343 bp) (GenBank U58873) demonstrated very high homology to human and rat Wee1 protein tyrosine kinase (~100 bp and 94% identity). Our SCH mapping places sMSH8 on chromosome 3 and FISH results indicate localization of the TYAC33 to 4p16, while human Wee1 protein tyrosine kinase has been mapped to 11p15. The reason for these apparent disparate mapping results is not clear, but it is possible that we have identified a pseudogene and that the TYAC contains other sequences that strongly hybridize to 4p rather than chromosome 3. TYAC35-sMSH10 (U58874) matches with 95% identity to exon 10 of human coagulation factor V (F5), which has been mapped to human chromosome 1, consistent with our SCH panel mapping results for sMSH10. TYAC25-rsAVA15 (U58872) demonstrates close homology to the human glutathione *S*-tranferase π gene. Our FISH mapping localized TYAC25 to multiple human telomeres, and chromosome assignment by SCH could not be accomplished due to homology to rodent genomic DNA.

As expected, a number of TYAC sequences matched to repetitive elements that are widely distributed in cDNAs and noncoding sequences. Highly significant matches were found to *Alu* (many of which are transcribed), *Kpn,* line repeats, and sequences containing MER22 and MSR1 elements. For example, TYAC72-sAVA27 (GenBank U58876) found >90% identity with 9000 matches to *Alu* sequence in the databases. Interestingly, TYAC163-rsAKB7 (GenBank U58878) identified 75 database entries that share homology with α satellite III or Y chromosome-specific repeats, consistent with our FISH results, which included hybridization to the centromere regions of chromosomes 14, 15, 21, and 22 and Yq11.2. In addition, one STS sJCW7 (TYAC86) (GenBank U58879) identified the subtelomeric sequence pHC1103 (Martin-Gallardo *et al.,* 1995).

SUMMARY AND DISCUSSION

45 TYAC Library Clones Mapped to 31 Specific Telomere Regions

A variety of mapping methodologies used to chromosomally position the TYAC sequences provided independent evidence for the telomeric locations of 44 of the TYAC and derivative P1 clones (Tables 1–3, and data not shown). In cases where FISH mapping indicated multiple chromosome assignments, a dominant hybridization was consistent with mapping data from SCH panels, meiotic mapping, and in some cases BLASTN identified matches to previously mapped sequences. We have demonstrated that the TYACs, which range in length from 25 to greater than 950 kb, derive from 31 telomere regions. FISH mapping indicates telomeric locations; however, the level of resolution obtained with our analyses was not definitive and in several cases (e.g., 12p, 11q) we found that the TYAC localized to the terminal band but did not appear to originate from the termi-

nus (Kobayshi *et al.*, 1994). This may be due to the presence of subtelomere repeats that met the criteria for TTAGGG homology during the screening that defined the library. Alternatively, in some cases, it is possible that these sequences are telomeric in the DNA from the individual who was the cloning source, but not telomeric in the individuals who have been studied by the various mapping methods. This can be verified by experiments using DNA from the founder of the TYAC library. Mapping of the P1 clones (which originated from a different individual) identified with STSs from the TYACs should provide some additional information in this regard. Nonetheless, FISH mapping provided a useful guide for the localization of the majority of sequence contained within the probe derived from the TYACs and P1s. STS mapping also gave single chromosome localization for the majority of TYACs mapped to a single chromosome by FISH, except in 5 cases. It is possible that the STS is duplicated on these chromosomes and that the duplicated region is beneath the level of detection obtainable with FISH. Definitive proof of the telomeric origin of the 45 TYACs can be obtained from physical mapping experiments such as Bal 31 and Rec A-assisted (RARE) cleavage (Ferrin and Camerini-Otero, 1991, 1994), which has recently been demonstrated to be feasible for human telomeres. Such experiments have been done demonstrating the telomeric origin of TYACs from 1q, 2q, 6q, 7q, 18p, 18q, and 21q (Negorev *et al.*, 1994; Macina *et al.*, 1994; Strathdee *et al.*, 1994; Reston *et al.*, 1995).

We have identified a number of ESTs and coding sequences from known genes within these TYACs as well as a number of unique sequences that may contain coding regions derived form telomeric locations. The microsatellite markers that are reported here will be useful reagents for map integration and for disease gene hunts since they now allow coverage of the proterminal regions of 14 chromosomes. Others have identified genes and pseudogenes that map near the telomeres, and it will be interesting in future to characterize these regions more fully and determine their biological significance, especially in light of the dynamic nature of telomere length and heteropolymorphism.

It is also apparent that clones for several telomeres appear not to be represented in our library. We have not identified clones for the telomere regions of chromosomes 5p, 6p, 9p, 15q, 19p, 19q, 20q, Xp, and Yp, although we have candidates for several of these termini. In addition, we have not found specific clones for the short arms of chromosomes 13, 14, 15, 21, and 22, although FISH has identified multiple TYACS hybridizing to the acrocentric short arms (13p, 14p, 15p, 21p, and 22p) and several pericentromeric regions (1q11, 3q11, 4q11). Three of these clones showed signal at 4p, consistent with the finding of Youngman *et al.* (1992), who demonstrated that the terminal 60 kb of 4p has sequences that are homologous to 13p, 15p, 21p, and 22p. However, several of these clones did not appear to hybridize to 4p. It may be possible that, at least for some of the other chromosomes for which we have not identified clones, subtelomere repeats may be present on a number of chromosomes and/or at interstitial locations, and since we have assigned such clones low priority for study, these telomeres have been overlooked. However, it should be noted that we have only partially mapped a number of clones. We identified a relatively large number of clones (5) that originated from 7q36. This may be due to the fact that the 7q terminus appears to be relatively free of subtelomere repeats and is therefore easier to map and may be easier to clone.

Different Mapping Methods Give Apparent Differing Chromosomal Localizations

It is also apparent that many TYACs did not localize to the distal termini of human chromosomes, and in a significant number of instances, SCH and FISH mapping produced different chromosome localizations. The TYAC library was assembled from multiple rounds of filter hybridization to the human telomere-specific repeat TTAGGG. The FISH results indicate that similar sequences are present in many regions of the chromosome and probably served as telomere stabilizing elements during the construction of the library. For example, a number of clones appear to derive from centromere regions. In addition, subtelomere repeats are known to reside on different chromosomes, which almost certainly explains the apparent hybridization to multiple telomeres by the same TYAC. The structures of subtelomere repeats are not well known and it is likely that as these TYACs are studied in more detail, additional subtelomere repeats will be uncovered. Another interesting possibility is that fragile sites could contain sequences that were sufficiently homologous to the telomere repeat to support the construction of stable TYACs. The presence of the 9q22 gene sequence for NTRK2 within TYAC40 and amplified by the STS sAVA18 provides strong evidence for the localization of the TYAC, which is consistent with the SCH and FISH mapping. A fragile site is located within 9q22 and it is possible that this region provided the TTAGGG homologous sequences that were cloned in our library.

Heteropolymorphism first reported in detail by Brown *et al.* (1990) may be responsible for some of the differences that we observe between FISH and other mapping results. Since different individuals' chromosomes were used to construct the SCH panel, for FISH experiments, for meiotic mapping, and for BLAST searching, it is possible that some of the sequences could be present or absent in different individuals used for mapping. These issues are currently being addressed, for example, by repeating FISH using chromosomes from the donor used to construct the TYAC library. In other cases, presumably the sequences originating from chromosomes not positive by FISH (but positive on the SCH by PCR) either were not large enough for FISH detection, or were not present on the chromosomes used for FISH. This can be tested directly

by PCR amplification of genomic DNA from the individuals used for FISH.

It is also interesting to note that a large number of ESTs and cDNAs were identified by our TYAC STSs, suggesting the locations of numerous pseudogenes, many of which may be located in distal regions. This finding suggests that STS-based mapping or DNA sequencing alone could give misleading results with respect to the large-scale organization of the genome. The importance of using multiple mapping methods, including meiotic mapping, has been well demonstrated by this study of the TYAC library. Each method has its limitations, in terms of resolution and reproducibility, but the accumulation of multiple independent segments of information can provide a more coherent view of the genomic terrain. The TYACs and related resources described in this report should provide ample reagents to move forward with more detailed characterization of human telomeres and other regions of the genome that contain TTAGGG homologous regions.

Availability of TYAC Data from Public Databases

The information contained within the tables and figures shown in this work and additional information related to the TYAC resource will be available from several public databases with access over the Internet. The most comprehensive listing of TYAC information (much of it contained within a searchable database) can be found in GenLink (http://www.genlink.wustl.edu). In addition, DNA sequence information will be available from GenBank (http://www.ncbi.nlm.nih.gov) and other marker information can be accessed through GDB (http://gdbwww.gdb.org).

ACKNOWLEDGMENTS

We thank Dr. Thomas Cremer in particular for guidance in FISH analyses, for advice over the course of the project, and for reviewing the manuscript prior to submission. We thank M. V. Olson for discussions and advice during the early phases of this work. We also thank M. V. Olson and H. Riethman for providing 7 TYAC clones, a number of TYAC cosmid subclones, and ligation mixtures from which the TYAC library was constructed. The technical assistance of Mary Pat Leicke from the Olson laboratory is gratefully acknowledged as are the contributions from the H.D-K. laboratory research staff, administrative staff, and students, in particular, Mary Akin, Stephanie Amen, Suzanne Cole, Meghan Dierks, William Dilley, Jim Dutchik, Michelle Lacy, Martin Lee, Jing Li, Brian Luketin, Bill Malone, Mark Schaller, Ed Shelton, and Chris Tierney. This work was supported in part by NIH Grants HG00100 and P41-HG01066 (to H.D-K.).

REFERENCES

Allshire, R. C., Dempster, M., and Hastie, N. D. (1989). Human telomeres contain at least three types of G-rich repeat distributed nonrandomly. *Nucleic Acids Res.* **17:** 4611–4627.

Anker, R. *et al.*, In preparation.

Bengtsson, U., Altherr, M. R., Wasmuth, J. J., and Winokur, S. T. (1994). High resolution fluorescence *in situ* hybridization to linearly extended DNA visually maps a tandem repeat associated with facioscapulohumeral muscular dystrophy immediately adjacent to the telomere of 4q. *Hum. Mol. Gen.* **3**(10): 1801–1805.

Blackburn, E. H., and Greider, C. W. (Eds.) (1995). "Telomeres," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Blouin, J. L., Christie, D. H., Gos, A., Lynn, A., Morris, M. A., Ledbetter, D. H., Chakravarti, A., and Antonarakis, S. E. (1995). A new dinucleotide repeat polymorphism at the telomere of chromosome 21q reveals a significant difference between male and female rates of recombination. *Am. J. Hum. Genet.* **57:** 388–394.

Brown, W. R. A., MacKinnon, P. J., Villasante, A., Spurr, N., Buckle, V. J., and Dobson, M. J. (1990). Structure and polymorphism of human telomere-associated DNA. *Cell* **63:** 119–132.

Browne, D. L., Smith, E. A., Dietz-Band, J., Riethmann, H. C., Phromchotikul, T., and Litt, M. (1995). Dinucleotide repeat polymorphism at the human chromosome 11p telomere (D11S2071). *Genomics* **25:** 600–601.

Burke, D. T., Carle, G. F., and Olson, M. V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236:** 806–812.

Cook, G. P., Tomlinson, I. M., Walter, G., Riethman, H., Carter, N. P., Buluiwela, L., Winter, G., and Rabbitts, T. L. H. (1994). A map of the human immunoglobulin VH locus completed by analysis of the telomeric region of chromosome 14. *Nature Genet.* **7:** 162–168.

Counter, C. M., Avilion, A. A., LeFeuvre, C. E., Stewart, N. G., Greider, C. W., Harley, C. B., and Bacchetti, S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* **11**(5): 1921–1929.

Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T. P., Bowden, D. W., Smith, D. R., Lander, E. S., Botstein, D., Akots, G., Rediker, K. S., Gravius, T., Powers, J. A., Watt, D. E., Kauffman, E. R., Bricker, A., Phipps, P., Muller-Kahle, H., Fulton, T. R., Ng, S., Schumm, J. W., Braman, J. C., Knowlton, R. G., Barker, D. F., Crooks, S. M., Lincoln, S. E., Daly, M. J., and Abrahamson, J. (1987). A genetic linkage map of the human genome. *Cell* **51:** 319–337.

Dubois, B. L., and Naylor, S. L. (1993). Characterization of NIGMS human/rodent somatic cell hybrid mapping panel 2 by PCR. *Genomics* **16:** 315–319.

Evans, G. A., Lewis, K., and Rothenberg, B. E. (1989). High efficiency vectors for cosmid microcloning and genomic analysis. *Gene* **79:** 9–20.

Ferrin, L. J., and Camerini-Otero, R. D. (1991). Selective cleavage of human DNA: RecA-assisted restriction endonuclease (RARE) cleavage. *Science* **254:** 1494–1497.

Ferrin, L. J., and Camerini-Otero, R. D. (1994). Long-range mapping of gaps and telomeres with RecA-assisted restriction endonuclease (RARE) cleavage. *Nature Genet.* **6:** 379–383.

Flint, J., Wilkie, A. O. M., Buckle, V. J., Winter, R. M., Holland, A. J., and McDermid, H. E. (1995). The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nature Genet.* **9:** 132–139.

Francomano, C. A., de Luna, R. I. O., Hefferon, T. W., Bellus, G. A., Turner, C. E., Taylor, E., Meyers, D. A., Blanton, S. H., Murray, J. C., McIntosh, I., and Hecht, J. T. (1994). Localization of the achondroplasia gene to the distal 2.5 Mb of human chromosome 4p. *Hum. Mol. Genet.* **6:** 787–792.

Freije, D., Helms, C., Watson, M. S., and Donis-Keller, H. (1992). Identification of a second pseudoautosomal region near the Xq and Yq telomeres. *Science* **258:** 1784–1787.

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.

Helms, C., Mishra, S. K., Riethman, H., Burgess, A. K., Ramachandra, S., Tierney, C., Dorsey, D., and Donis-Keller, H. (1992). Closure of a 2.4 cM genetic linkage map of human chromosome 7q with centromere and telomere polymorphisms. *Genomics* **14:** 1041–1054.

Henke, A., Fischer, C., and Rappold, G. A. (1993). Genetic map of the human pseudoautosomal region reveals a high rate of recombination in female meiosis at the Xp telomere. *Genomics* **18:** 478–485.

Heutink, P., Zguricas, J., van Oosterhout, L., Breedvelt, G. J., Tester, L., Sandkujl, L. A., Snijders, P., Weisenbach, J., Linkhout, D., Hovius, S., and Oostra, B. A. (1994). The gene for triphalangeal thumb maps to the subtelomeric region of chromosome 7q. *Nature Genet.* **6:** 287–291.

Hing, A. V., Helms, C., and Donis-Keller, H. (1993). VNTR and microsatellite polymorphisms within the subtelomeric region of 7q. *Am. J. Hum. Genet.* **53:** 509–517.

Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L. C., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* **266:** 2011–2015.

Kipling, D. (1995). "The Telomere," Oxford Univ. Press, Oxford.

Kobayashi, H., Montgomery, K. S., Bohlander, S. K., Adra, C. N., and Lim, B. L. (1994). Fluorescence *in situ* hybridization mapping of translocations and deletions involving the short arm of human chromosome 12 in malignant hematologic diseases. *Blood* **84:** 3773–3482.

Kvaloy, K., Galvagni, F., and Brown, W. R. A. (1994). The sequence organization of the long arm pseudoautosomal region of the human sex chromosomes. *Hum. Mol. Genet.* **3:** 771–778.

LeMerrer, M., Rousseau, F., Legeai-Mallet, L., Landais, J-C., Pelet, A., Bonaventure, J., Sanak, M., Weissenbach, J., Stoll, C., Munnich, A., and Maroteaux, P. (1994). A gene for achondroplasia-hypochondroplasia maps to chromosome 4p. *Nature Genet.* **6:** 318–321.

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

Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D. C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* suppression hybridization using recombinant DNA libraries. *Hum. Genet.* **80:** 224–234.

Macina, R. A., Negorev, D. G., Spanisc, C., Ruthig, L. A., Hu, S.-L., and Riethman, H. R. (1994). Sequence organization of the human 2q telomere. *Hum. Mol. Genet.* **3:** 1847–1853.

Martin, G. M. (1994). Genetic modulation of telomeric terminal restriction-fragment length: Relevance for clonal aging and late-life disease. *Am. J. Hum. Genet.* **55:** 866–869.

Martin-Gallardo, A., Lamerdin, J., Sopapan, P., Friedman, C., Negorev, A., and Trask, B. J. (1995). Molecular analysis of a novel subtelomeric repeat with polymorphic chromosomal distribution. *Cytogenet. Cell Genet.* **71:** 289–295.

Moyzis, R. K., Buckingham, J. M., Cram, S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L., and Wu, J. R. (1988). A highly conserved repetitive DNA sequence (TTAGGG)n present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. USA* **85:** 6622–6626.

Mueller, P. R., and Wold, B. (1989). In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246:** 780–786.

Nakagawara, A., Liu, X.-G., Ikegaki, N., White, P. S., Yamashiro, D. J., Nycum, L. M., Biegel, J. A., and Brodeur, G. M. (1995). Cloning and chromosomal localization of the human TRK-B tyrosine kinase receptor gene (NTRK2). *Genomics* **25:** 538–546.

Negorev, D. G., Macina, R. A., Spais, C., Ruthig, L. A., Hu, X.-L., and Riethman, H. C. (1994). Physical analysis of the terminal 270 kb of DNA from human chromosome 1q. *Genomics* **22:** 569–578.

NIH/CEPH Collaborators Mapping Group (1992). A comprehensive genetic linkage map of the human genome. *Science* **258:** 67–86 and 148–162.

Overhauser, J., Silverman, G. A., and van Kessel, A. G. (1995). Report of the third international workshop on human chromosome 18 mapping. *Cytogenet. Cell Genet.* **71**(2): 106–107.

Pandit, S. D., Wang, J. C., Veile, R. A., Mishra, S. K., Warlick, C. A., and Donis-Keller, H. (1995). Index, comprehensive microsatellite, and unified linkage maps for human chromosome 14 with cytogenetic tie points and a telomere microsatellite marker. *Genomics* **29:** 653–664.

Pandolfo, M. (1992). A rapid method to isolate $(GT)_n$ repeats from yeast artificial chromosomes. *Nucleic Acids Res.* **20:** 1154.

Rappold, G. A. (1993). The pseudoautosomal regions of the human sex chromosomes. *Hum. Genet.* **92:** 315–324.

Reston, J. T., Hu, X.-L., Mascina, R. A., Spais, C., and Riethman, H. C. (1995). Structure of the terminal 300 kb of DNA from human chromosome 21q. *Genomics* **26:** 31–38.

Riethman, H. C., Moyzis, R. K., Meyne, J., Burke, D. T., and Olson, M. V. (1989). Cloning human telomeric DNA fragments into *Saccharomyces cerevisiae* using a yeast-artificial-chromosome vector. *Proc. Natl. Acad. Sci. USA* **86:** 6240–6244.

Riethman, H. C., Spais, C., Buckingham, J., Grady, D., and Moyzis, R. K. (1993). Physical analysis of the terminal 240 kb of DNA from human chromosome 7q. *Genomics* **17:** 25–32.

Schmitt, H., Blin, N., Zanki, H., and Scherthan, H. (1994). Telomere length variation in normal and malignant human tissues. *Genes Chromosomes Cancer* **11:** 171–177.

Schumm, J. W., Knowlton, R. G., Braman, J. C., Barker, D., Botstein, D., Akots, G., Brown, V., Gravius, T., Helms, C., Hsaio, K., Rediker, K., Thurston, J., and Donis-Keller, H. (1988). Detection of more than 500 single copy RFLPs by random screening. *Am. J. Hum. Genet.* **42:** 143–159.

Shepherd, N. S., Pfrogner, B. D., Coulby, J. N., Ackerman, S. L., Vaidyanathan, G., Sauer, R. H., Balkenhol, T. C., and Sternberg, N. (1994). Preparation and screening of an arrayed human genomic library generated with the P1 cloning system. *Proc. Natl. Acad. Sci. USA* **91:** 2629–2633.

Srivastava, A. K., Montanaro, V., and Kere, J. (1992). Simplified template preparation and improved direct sequencing using Taq polymerase. *PCR Methods Appl.* **1**(4): 255–256.

Strathdee, G., Harrison, W., Riethman, H. C., Goodart, S. A., and Overhauser, J. (1994). Interstitial deletions are not the main mechanism leading to 18q deletions. *Am. J. Hum. Genet.* **54:** 1085–1091.

Van Kessel, A. G., Straub, R. E., Silverman, G. A., Gerken, S., and Overhauser, J. (1994). Report of the second international workshop on human chromosome 18 mapping. *Cytogenet. Cell Genet.* **65**(3): 142–165.

Wilkie, A. O. M., Buckle, V. J., Harris, P. C., Lamb, J., Barton, N. J., Reeders, S. T., Lindenbaum, R. H., Nicholls, R. D., Barrow, M., Bethlenfalvay, N. C., Hutz, M. H., Tolmie, J. L., Weatherall, D. J., and Higgs, D. R. (1990). Clinical features and molecular analysis of the alpha thalassemia/mental retardation syndromes. I. Cases due to deletions involving chromosome band 16p13.3. *Am. J. Hum. Genet.* **46:** 1112–1126.

White, G. R. M., Stack, M., Sautibanez-Koref, M., Liscia, D. S., Venesio, T., Wang, J-C., Helms, C., Donis-Keller, H., Betticher, D. C., Altermatt, H. J., Hoban, P. R., and Heighway, J. (1996). High levels of loss at the 17p telomere suggest the close proximity of a tumour suppressor. *Br. J. Cancer* **74:** in press.

Wilkie, A. O. M., Lamb, J., Harris, P. C., Finney, R. D., and Higgs, D. R. (1990). A truncated human chromosome 16 associated with alpha thalassaemia is stabilized by addition of telomeric repeat $(TTAGGG)_n$. *Nature* **346:** 868–871.

Youngman, S., Bates, G. P., Williams, S., McClatchey, A. I., Baxendale, S., Sedlacek, Z., Altherr, M., Wasmuth, J. J., MacDonald, M. E., Gusella, J. F., Sheer, D., and Lehrach, H. (1992). The telomeric 60kb of chromosome arm 4p is homologous to telomeric regions of 13p, 15p, 21p, and 22p. *Genomics* **14:** 350–356.