Genomics **64,** 1–14 (2000)
doi:10.1006/geno.1999.6051, available online at http://www.idealibrary.com on IDEAL®

A 6-Mb High-Resolution Physical and Transcription Map Encompassing the Hereditary Prostate Cancer 1 (HPC1) Region

John D. Carpten,*[,1] Izabella Makalowska,† Christiane M. Robbins,* Nyasha Scott,* Raman Sood,* Tim D. Connors,‡ Tom I. Bonner,§ Jeffrey R. Smith,* Mezbah U. Faruque,* Dietrich A. Stephan,* Heather Pinkett,* Sharon D. Morgenbesser,‡ Kui Su,‡ Chris Graham,* Simon G. Gregory,¶ Hawys Williams,¶ Louise McDonald,¶ Andreas D. Baxevanis,† Kathy W. Klingler,‡ Greg M. Landes,‡ and Jeffrey M. Trent*

**Cancer Genetics Branch and †Gene Technology Branch, National Human Genome Research Institute, and §National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892; ‡Genzyme Genetics Corporation, Framingham, Massachusetts 01701; and ¶The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom*

Received August 31, 1999; accepted November 4, 1999

Several hereditary disease loci have been genetically mapped to the chromosome 1q24–q31 interval, including the hereditary prostate cancer 1 (HPC1) locus. Here, we report the construction of a 20-Mb yeast artificial chromosome contig and a high-resolution 6-Mb sequence-ready bacterial artificial chromosome (BAC)/P1-derived artificial chromosome (PAC) contig of 1q25 by sequence and computational analysis, STS content mapping, and chromosome walking. One hundred thirty-six new STSs, including 10 novel simple sequence repeat polymorphisms that are being used for genetic refinement of multiple disease loci, have been generated from this contig and are shown to map to the 1q25 interval. The integrity of the 6-Mb BAC/PAC contig has been confirmed by restriction fingerprinting, and this contig is being used as a template for human chromosome 1 genome sequencing. A transcription mapping effort has resulted in the precise localization of 18 known genes and 31 ESTs by database searching, exon trapping, direct cDNA hybridization, and sample sequencing of BACs from the 1q25 contig. An additional 11 known genes and ESTs have been placed within the larger 1q24–q31 interval. These transcription units represent candidate genes for multiple hereditary diseases, including HPC1.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AF172072–AF172082, AQ761940–AQ762052, AQ797399–AQ797430, and AF181666–AF181679.

[1] To whom correspondence should be addressed at NHGRI, NIH, Building 36, Room 3D04, 9000 Rockville Pike, Bethesda, MD 20892. Telephone: (301) 435-5626. Fax: (301) 435-5465. E-mail: jdc@nhgri.nih.gov.

INTRODUCTION

Our laboratory performed a genome-wide scan of families at high risk for prostate cancer and revealed evidence of a major prostate cancer susceptibility locus (HPC1; OMIM 601518) on 1q24–q31 (Smith *et al.*, 1996). Prostate cancer linkage to markers from the 1q24–q31 region has been confirmed by additional independent data sets (Cooney *et al.*, 1997; Hsieh *et al.*, 1997). Interestingly there is significant genetic heterogeneity in prostate cancer as several other prostate cancer susceptibility loci, including Xq27 and 1q42, have been reported by us and others (Xu *et al.*, 1998; Berthon *et al.*, 1998). As part of a positional cloning effort toward the isolation of the HPC1 gene, we have constructed a detailed physical map of the 1q24–q31 interval using yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), and P1-derived artificial chromosomes (PACs) (Shizuya *et al.*, 1992; Kim *et al.*, 1996; Ioannou *et al.*, 1994; Dausett *et al.*, 1992). Genomic clones have been identified by database searching and PCR screening of commercially available library DNA pools. These genomic clones represent reagents for mapping and identifying HPC1 candidate genes.

In addition to HPC1, several other hereditary disease loci have been genetically mapped to 1q24–q31. Examples include the camptodactyly arthropathy pericarditis syndrome (CACP; OMIM 208250) locus, which was mapped to this interval by linkage analysis (Bahabri *et al.*, 1998). This autosomal recessive disorder is characterized by congenital flexion contractures, noninflammatory synovial hyperplasia, and pericarditis (Athreya and Schumacher 1978; Ochi *et al.*, 1983; Martinez-Lavin *et al.*, 1983). CACP patients have also been diagnosed with coxa vara, a rare hip joint deformity (Bulutlar *et al.*, 1986). Additionally, the hyperparathyroidism/jaw tumor syndrome (HPRT2; OMIM 145001)

0888-7543/00 $35.00

has been genetically mapped to a 14.3-cM interval encompassing 1q24–q31 (Teh *et al.*, 1996; Hobbs *et al.*, 1999; Williamson *et al.*, 1999). This autosomal dominant disorder, similar to but distinct from MEN1, results in hyperparathyroidism associated with an increased risk of parathyroid adenomas with some affected individuals developing rare "fibro-osseous" tumors of the mandible and maxilla (Jackson *et al.*, 1990; Fujikawa *et al.*, 1998). Renal hamartomas and cystic kidney disease have also been reported in HRPT2 kindreds (Teh *et al.*, 1996). Loss of heterozygosity has been reported for some HRPT2 kindreds, suggesting the possible role of a tumor suppressor gene in the etiology of this syndrome (Teh *et al.*, 1996).

To date we have completed a 20-Mb YAC contig by database searching, STS content mapping, and chromosome walking spanning the interval flanked by D1S212 to D1S412. Also, a 6-Mb BAC/PAC contig of 1q25 has been constructed and is currently being used as a resource for human genome sequencing by the Sanger Genome Sequencing Centre Human Chromosome 1 Sequencing Project. A number of known genes and potential novel transcripts have been precisely mapped by sample sequencing, exon trapping, and cDNA hybridization (Kupfer *et al.*, 1995; Church *et al.*, 1994; Connors *et al.*, 1994). These genes as well as any new transcription units identified from this physical map serve as candidate genes for multiple human diseases.

MATERIALS AND METHODS

YAC database search. A database search was performed against the Whitehead Institute for Biomedical Research/MIT Center for Genome Research (MIT–WICGR) human physical mapping database (Hudson *et al.*, 1995). All published markers within the 1q24–q31 interval were cross-referenced against the MIT–WICGR whole human genome YAC contig database. A subset of YACs was selected from this database based upon known STS content of the YACs. YACs having any unambiguous hits to markers known to map outside of the 1q24–q31 interval were excluded to decrease the number of possible chimeric YACs in the pool.

Polymerase chain reaction (PCR). All PCRs were carried out using 10 ng of template DNA with 2.25 mM Mg^{2+}, 250 nM dNTPs, 333 nM each forward and reverse primer, PCR buffer II (Perkin–Elmer), and 0.6 units of AmpliTaq Gold Polymerase (Perkin–Elmer) in a 15-μl total PCR volume. All PCRs were carried out in either a Model 9600 or a Model 9700 Thermocycler (Perkin–Elmer) using the following cycling protocol: initial denaturation of 94°C for 12 min; 94°C for 15 s, 55°C for 15 s, 72°C for 15 s for 35 cycles; along with a final extension at 72°C for 10 min. All PCRs were separated by agarose gel electrophoresis using 2% agarose gels.

YAC DNA isolation. YAC DNA was isolated by spheroplasting as described previously (Carpten *et al.*, 1994). Briefly, single-colony purified YAC clones were grown in 30 ml of SD medium at 30°C for 48 h. Following centrifugation, pellets were resuspended in sorbitol buffer with zymolase 100T (ICN). Spheroplasts were pelleted and lysed at 68°C in 10% SDS. Following treatment with potassium acetate, DNA was cleaned up with Strataclean resin (Stratagene). DNA was precipitated at room temperature with isopropanol, treated with RNase, and reprecipitated. Purified YAC DNA was resuspended in 100 μl of TE buffer, pH 8.0. This protocol typically yields approximately 10–15 μg of yeast DNA.

YAC end isolation. YAC end sequences were generated by a modification of the protocol described by Riley *et al.* (1990). Briefly, 1 μg of YAC DNA was digested with 20 units of *Rsa*I, *Alu*I, or *Eco*RV according to the manufacturer's recommendations (BRL, Life Technologies). All subsequent steps were as described by Riley *et al.* (1990) to attain end sequences for both the left (T) and the right (U) ends of all YAC clones. YAC end PCR products were sequenced using either the left or the right YAC vector-specific primers 1207 or 1208, respectively, or with the universal vectorette primer. Cycle sequencing reactions were carried out using fluorescent dye terminator chemistry and AmpliTaq FS as described below. Sequencing reactions were separated using an Applied Biosystems 377 XL automated DNA sequencer (Perkin–Elmer).

BAC/PAC library screening. Human BAC and PAC clones were identified by PCR screening DNA pools representing commercially available BAC or PAC libraries. Libraries screened were human (CITB) BAC and human (RPCI1) PAC libraries (Genome Systems) or the human (CITB) BAC library (Research Genetics). Libraries were screened according to the distributor's recommendations. PCR was performed as described above.

BAC/PAC DNA isolation. Forty-microliter cultures of single colony purified BAC/PAC clones were used for DNA isolation. DNA was isolated by alkaline lysis with an AutoGen 850 automated DNA isolation system using the manufacturer's recommended protocol (Autogen). Subsequently, the BAC/PAC DNA was resuspended in 600 μl dH_2O, treated with RNase (Ambion), and purified over a Microcon 100 column (Amicon).

Sizing of BAC/PAC inserts. Approximately 1 μg of BAC/PAC DNA was digested for 2 –3 h with 40 units of *Not*I (Boehringer Mannheim). Restriction fragments were separated by CHEF gel electrophoresis using a CHEF MAPPER gel electrophoresis system (Bio-Rad). The Auto Algorithm feature was programmed for the separation of fragments ranging in size from 50 to 500 kb on a 1% agarose gel in 0.5× TBE at 14°C for 18 h with an included angle of 120°. Lambda concatamers were used as a DNA size standard (Bio-Rad). Gels were stained with ethidium bromide for UV detection of DNA bands.

Restriction fingerprinting of genomic clones. BAC/PAC clones were fingerprinted as previously described (Gregory *et al.*, 1997). Briefly, fingerprinting reactions were carried out by concurrent enzymatic reactions using *Hin*dIII and *Sau*3AI restriction endonucleases for digestion and *Taq* FS (Perkin–Elmer) for fluorescence end labeling of restriction fragments using ddA fluorescently modified with one of three individual dyes: HEX, NED, or TET (Perkin–Elmer). Separate fingerprinting reactions were carried out for the different dyes. Samples were mixed and incubated at 37°C for 1 h. Following ethanol precipitation, all three dye-labeled reactions were pooled for each BAC DNA sample. Samples were mixed with a fluorescently labeled DNA standard, and fragments were separated on an Applied Biosystems 377 automated DNA sequencer (Perkin–Elmer). Following gel data collection and analysis, fragment analysis and contig assembly were carried out using FPC (Soderlund *et al.*, 1997).

Exon trapping. Exon trapping was performed as previously described by Church *et al.* (1994) and Connors *et al.* (1994). Following PCR analysis, putative exon traps were sequenced using Applied Biosystems Big Dye d-rhodamine terminator chemistry (Perkin–Elmer). Sequencing reactions were separated using Applied Biosystems Model 377 automated DNA sequencers (Perkin–Elmer). Data were subsequently tracked and analyzed with Applied Biosystems DNA Analysis Sequencing Software 3.2 (Perkin–Elmer). Sequences were edited and assembled using Sequencher sequence analysis software (GeneCodes). Sequences from the exon traps were compared to the nonredundant nucleotide and dbEST databases using a PowerBlast sequence homology algorithm (Altschul *et al.*, 1997; Zhang & Madden, 1997).

cDNA selection. BAC inserts were removed from vector sequences by digestion with *Not*I and separated by CHEF gel electro-

phoresis as described above. Bands representing BAC inserts were excised from the gel using a sterile razor blade. DNA was purified by treatment with β-agarase (BRL Life Technologies) using the manufacturer's recommendations. One hundred nanograms of purified BAC DNA was labeled with [α-^{32}P]dCTP by random priming using the MegaPrime Kit (Amersham). Probes were suppressed with Cot-1 DNA and the oligonucleotides $(CA)_{10}$, $(GA)_{10}$, and $(GATA)_8$, for 2 h at 65°C. IMAGE Consortium high-density cDNA library filters obtained from Genome Systems, Inc. were screened with suppressed BAC insert probes using the manufacturer's recommendations. Positive clones were obtained from Genome Systems, and ends were sequenced as described above. cDNA inserts were also used as probes on BAC Southern blots to confirm proper mapping of cDNAs to the correct parent BAC clones.

Southern blotting. Southern blotting was performed as described elsewhere (Sambrook *et al.* 1989). Briefly, 2 μg of BAC DNA was digested with *Eco*RI (BRL Life Technologies) and separated on a 1% agarose gel in 1× TBE buffer. DNA was transferred to Genescreen nylon membranes (NEN Life Sciences) by capillary transfer.

BAC/PAC insert terminal end sequencing. BAC/PAC DNA was isolated as described above. BACs were sequenced using M13 forward and reverse primers, and PACs were sequenced using T7 and SP6 primers as per the manufacturer's recommendations. Then 2× sequencing reactions were set up using the Big Dye Terminator Chemistry (Perkin–Elmer) as follows: 500 ng BAC DNA, 375 nM primer, 16 μl Big Dye Terminator Reagent Mix in a 40-μl total reaction. Cycle sequencing was performed in an MJ Tetrad Thermocycler (MJ Research) using the following cycling conditions: 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 50°C for 20 s, 60°C for 4 min. Sequencing reactions were cleaned using Centri-Sep columns according to the manufacturer's recommendations (Princeton Separations), dried in a Speed Vac (Savant), redissolved/denatured in 3 μl of loading buffer (95% formamide/50 mM EDTA) at 90°C for 3 min, and analyzed with an Applied Biosystems 377 XL automated DNA sequencer (Perkin–Elmer). Gel files were tracked and analyzed using Applied Biosystems DNA Analysis Sequencing Software 3.2 (Perkin–Elmer).

Shotgun library construction and single-stranded DNA isolation. BAC DNA was isolated as described above. BAC DNA was sent to SeqWright Corp. (Houston, TX) for shotgun library construction in M13 phage vector. We picked approximately 1400 individual M13 plaques, which were gridded into 96-well microtiter plates and inoculated with *Escherichia coli* strain JM101 in 2X-YT medium for single-stranded DNA isolation and library storage. Single-stranded DNA was isolated in a 96-well format using the high-throughput preparation of M13 DNA (THERMOMAX DNA Preparation Protocol from the Washington University Sequencing Center, St. Louis, MO, http://genome.wustl.edu/gsc/index.shtml).

Sample sequencing. We sequenced single-stranded DNA using the DYEnamic Direct Energy Transfer fluorescently labeled M13 forward sequencing primer (Amersham). Briefly, 200 ng of single-stranded template DNA was used in an 8-μl reaction for ddA/ddC, and 400 ng of template DNA was used in a 16-μl reaction for ddG/ddT with Thermo Sequenase (Amersham Pharmacia). Sequencing reactions were carried out on an Applied Biosystems CATALYST 800 Molecular Biology LabStation (Perkin–Elmer) using the following protocol (95°C for 5 s, 55°C for 10 s, 72°C for 60 s for a total of 15 cycles). The four dye primer reactions were subsequently pooled and precipitated with 132 μl of 95% ethanol and 5 μl glycogen (Boehringer Mannheim), dried by vacuum, and resuspended in 3 μl of loading buffer. Following brief denaturation, sequencing reactions were electrophoresed on an ABI 377 XL Automated DNA Sequencer using the manufacturer's protocols. Gel files were tracked and analyzed using Applied Biosystems DNA Analysis Sequencing Software 3.2 (Perkin–Elmer).

RESULTS

YAC Contig Construction

A database search against the MIT–WICGR human physical mapping database was performed to identify megaYAC clones covering the interval flanked by D1S212 and D1S533. All YACs from MIT–WICGR singly linked contigs WC-1.16 and WC-1.17 were cross-checked against the MIT–WICGR human YAC database to eliminate possible chimeric clones. Any clones that had unambiguous hits to markers not specific to chromosome 1 were considered to be chimeric and were eliminated from consideration unless a chimeric clone was absolutely necessary for linking two markers in the contig. Forty human megaYAC clones were selected for inclusion into the YAC contig (Table 1). Overlaps were confirmed by STS content mapping of these YACs by PCR using markers from WICGR singly linked contigs WC-1.16 and WC-1.17 and novel STSs generated by YAC end sequencing (Table 1). Of the initial set of 48 YAC clones, 47 clones were believed to be nonchimeric and were confirmed as being nonchimeric by YAC end isolation and mapping of end sequences to a chromosome 1-specific somatic cell hybrid (data not shown). Table 2 lists primers derived from YAC ends.

Several doubly linked contigs (WC-1274, WC-929, WC-919, and WC-651) flanked proximally by D1S212 and distally by D1S533 were identified within singly linked contigs WC-1.16 and WC-1.17, suggesting that there were gaps in the WICGR YAC contig representing this area of the genome. We used chromosome walking facilitated by YAC end isolation to close gaps between adjacent doubly linked contigs. One such gap was observed between WC-1274 and WC-929. DNA markers flanking this gap were D1S117 proximally and D1S466 distally. One YAC clone, y954D4, was positive for the marker D1S466 and was a good candidate for anchoring the distal chromosome walk with respect to this gap. The novel STS y20Ta (Table 2), which represents the left end of the y954D4 insert, was used to screen the CEPH megaYAC library DNA pools (data not shown). Two YACs that were not represented in the WICGR database, y615F2 and y662C9, were identified. Upon STS content analysis, these two YACs were shown to be positive for a novel STS derived from the end of y807D7 (Table 2), which is also positive for the marker D1S117, suggesting that this gap was closed. Although one gap remains between markers AFMB330XE9 and D1S533, 20 Mb are complete. We have isolated ends from YACs flanking the gap, and we are in the process of closing this gap in the YAC contig.

1q25 BAC Contig Construction

For the construction of a BAC contig of 1q25, a total of 31 total markers (Table 1) including YAC end-

TABLE 1

YAC STS Content

Clone ID	Size (kb)	D1S212	D1S2791	D1S2078	D1S3424	SOAT	D1S215	D1S2883	D1S2357	D1S3379	D1S3640	D1S117	y22U	**y20Ta**	**y18Ua**	**D1S466**	**D1S2619**	**D1S2818**	**Y20Ur**	**y19Ta**	**y13Tr**	**y14Ur**	**y16Tr**	**y12Ta**	**D1S2472**	**D1S158**	**y11Tr**	**D1S2127**	**y13Ur**	**y10Ur**	**D1S2701**	**D1S240**	**y12Ua**	**y11Ur**	**D1S254**	**D1S444**	**D1S191**	**D1S2848**	**D1S2165**	**D1S1599**	**y10Ta**	**y8Tr**	**D1S202**	**y7Ur**	**y6Ua**	**D1S1642**	D1S518	D1S1604	D1S222	y50U	D1S238	y7Tr	y8Ua	y6Tr	D1S2877	D1S461	D1S428	D1S542	D1S492	y48T	y3Ua	Z39494	y2Ua	D1S2151	y1Tr	y48U	D1S422	D1S384	D1S1593	y2Ta	y35U	y33U	D1S2625	y30Ua	AFMB330XE9	D1S2757	D1S1614	D1S412	D1S533
y744A9	1550	+	+	+																																																																											
790F6	1390			+	+	+																																																																								
y849D3	nd				+	+	+																																																																							
y920B6	1730						+	+	+																																																																					
y756C7	1330								+	+																																																																					
y850BH12	1370								+	+	+	+																																																																	
y713F1	750										+	+																																																																	
y807D10	1790										+	+	+																																																																
y615F2	nd												+	+																																																															
y662C9	nd												+	+																																																															
y954D4	920													+	+	+	+	+	+																																																												
y789H12	930														+	+	+	+	+																																																												
y956F10	1150														+	+	+	+	+	+																																																									
y964H4	1470														+	+	+	+	+	+																																																									
y913C9	1140																			+	+	+	+																																																							
y940F12	1900																					+	+	+	+	+	+	+	+	+																																															
y823C9	1110																										+	+	+	+	+	+	+	+																																											
y910F9	1530																							+	+	+	+	+	+	+	+	+	+	+																																											
y956B9	1710																													+	+	+	+	+	+	+	+	+	+	+	+																																					
y937B10	1430																																		+	+	+	+	+	+	+	+	+	+	+	+																																
y858E2	1750																																				+	+	+	+	+	+	+	+	+	+																																
y947C3	1760																																									+	+	+	+	+	+	+	+	+	+	+	+																									
y744C4	1560																																											+	+	+	+	+	+	+	+	+																										
y817B6	1450																																												+	+	+	+	+	+	+	+	+	+																								
y948F10	1480																																																+	+	+	+	+	+																								
y971C3	1170																																																	+	+	+	+	+	+	+	+	+																				
y709B8	140																																																						+	+																						
y852A8	550																																																								+	+	+	+																		
y750H3	450																																																									+	+	+	+																	
y931D11	1550																																																											+	+	+	+	+	+	+												
y805F3	1340																																																														+	+	+	+	+	+	+	+								
y895F11	1310																																																																+	+	+	+	+									
y792E5	710																																																																				+	+	+	+							
y756A10	1500																																																																						+	+	+	+					
y870F4	790																																																																								+	+					
y845D8	1280																																																																									+	+				
y806H3	980																																																																									+	+				
y770B8	nd																																																																											+	+		
y822B12	950																																																																												+	+	+
y746G3	1040																																																																												+	+	+

HPC1 Interval

HRPT2 Interval

CACP

Note. Clone ID is "y" followed by CEPH mega YAC library address. Clone insert sizes are given in kilobase pairs. DNA markers used for STS content are given across the top of the table. A plus sign depicts the presence of that marker within a YAC clone. The markers in boldface are those used for BAC/PAC library screening.

TABLE 2

Novel STSs

Marker name	GenBank Accession No.	Forward primer sequence	Reverse primer sequence	Product size (bp)
y22U	AQ797414	CACATTTATAGCTGAGGG	TGTGGATCTCAGTCATCC	126
y20Ta	AQ797412	CACTTTGTAACCTCCGCA	CCTGGTGCTCATCAACAT	135
456G14cR	AQ762022	TTTCTCAAGCATCTGGGC	AAAGGTGGGTCTCCTTAG	135
336O2cR	AQ762017	CCATAACCTGAGACAGAC	CTGTGAGGAGTTACATGG	235
55C10cR	AQ762025	TATCTGACCAGTTCACCG	AGAAGAACAGGCTTAGG	187
23P11cR	AQ761988	TCTAAGGCTCGTGAGTTCG	TGATGCAGTCCAGGAATGC	199
336O2cF	AQ762016	TCAGTGCCGAAGAGAAAC	GAAAGGTCTCAGCTCTTC	203
333A6cR	AQ762015	GCTTTGGGTTTATGAGATCC	TCTTGTTAGCAACATGCTGG	202
122H1cR	AQ761952	TACCTCAGTGGGATTCTG	CACTGGAGCATGAATCTC	154
313H7cF	AQ762012	CCTTCCCATCTCACAAATG	GGTTCTTTGGAAATCCCAG	243
333A6cF1	AQ762014	TTGGGATTCAGACAGATCTGG	GGCCACTCAAAAGTTATCAC	156
362E18cF	AQ762019	TGGCTCCTAGAAAGGTAG	TACCCTGGGTATTGCATG	107
285P16cF	AQ762009	ACAGGGAGTGAATGATGAGG	GTCTTAGGTTTCTCGCTTGG	168
122H1cF	AQ761951	CAGATATACAGGTCACAGG	AAACATGGTGAGAAGCAG	171
278C23cF	AQ762002	CAATGCAGAGAAGTTGGTC	GGGTGTTAAAGTGCTGAAC	155
223H12cR	AQ761981	ATCATCAGGACCTCACTC	AGCCATGACTGAGAGAAG	144
285P16cR	AQ762010	TAATGGCTCTCCGCCATTGTTG	ACTTCAGGATGATCCTCTTGGG	226
278C23cR	AQ762003	ATGTCCAAAAGGCTTTCCTATG	GTAAGCTACAATCTCATCTTGC	225
180A9cF	AQ761964	TAAAGAACGGTCCTTGGTGC	GCTATTCTCCACTACTGAGG	134
y20Ur	AQ797413	CATGAGGACAGGAATGCA	TCAGTCAAGTGGGAGAAG	135
24J24cR	AQ761994	TTCCAGGCATGACGTATG	GTTGACTCCTTGAGACTG	233
223H12cF	AQ761980	AAAATGGTGTTCCCAAGTCTC	GAGTTTAGAGCATGGATATCC	148
270B19cR	AQ762001	TATCTGGTCTTCAGACCC	AACAGAGAGCAAGGTCAC	180
24J24cF	AQ761993	GAGAAGCCAGTACCATAAG	AACCTGGAGAATAGCAAGG	106
105G1cF	AQ762043	TGGTGGTAGTGTAGTGTG	TAAGCCTAACCCTAACCC	262
245D23cR	AQ761989	AACCCTCCAAGATGGTACAG	ACTAGAAGCCATGGCATACC	222
53H18cR	AQ762052	AGATCATGAGTGGGACTG	TTGAGGGAAGCTATCCAC	175
112B20cF	AQ762044	TAGAGTCCACCTGAACAG	TCAACTAGTTAGTGGCCC	120
y19Ta	AQ797410	CCTGCAGCAAAACCCACAG	GTCCTCTGTTCTCACCTCA	98
y13Tr	AQ797405	CAACAGGACTGAGTTCAAGTC	TGTATCCACCACTTCGAG	129
187H3cR	AQ761966	TTAACTCAGGCTCAGCTTCC	ATATAACAGGGCAAGTGGGC	128
211P21cR	AQ761974	TATCCACATGATGGACTGGC	TAGGAAGAGACAGTGTACCC	116
536E3cF	AQ762023	AGGCCAGAAAGGACTTTGAC	GAAGGCTTTGATCTGCAGAG	128
338D11cF	AQ762018	CTTTAAGGTGACCCAGAGTTGC	GAGTTTTATCTGGGCATGTGCC	118
187H3cF	AQ761965	CTAAGATTGGGAAGGTGG	CATCTGAAGTGACCTGAG	151
107A4cR	AQ761940	CAGCAGATGTGGAACTAGAG	GATGACATTGGCATCTCAGG	144
229A1cF	AQ761984	AAGACCCAGATCAACAGCCAAC	TCTTCACATCAGCCAGAATCCC	126
536E3cR	AQ762024	CAGAATTACAGAGACGTGCC	CCACACTTGGTATTGCTGAC	131
324G22cF	AQ762013	GAAACAAACGCGCGATAAG	GGTGGATGCTGAGTGAATAG	273
57P8R	AQ762029	GCTCGATACAGAAACCAG	TTGACAGGTCCACTCACC	89
127C7cR	AQ762047	ATCTTCGACTTTGTGGCTTG	CTGCAAAGGGACATTACAAG	89
y14Ur	AQ797407	AGGTGTGCTTCTAGTCAGTG	AGGTGTGCTTCTAGTCAGTG	207
127C7cF	AQ762046	CACACTAAGCTTAAGAATCTTTG	TGACCACTTGATTTAGAGTAATG	153
y16Tr	AQ797408	GAAACTGAGTTCCCAGAG	GAAACTGAGTTCCCAGAG	129
110O11cR	AQ761945	GAGCCTTTTGTGATTGGC	CCATGTTGAGAAGCACAG	129
y12Ta	AQ797403	TGCTTGCTTATGGTGCAGAG	CTGATTGAATGTTTCTGAGC	169
110M15cF	AQ761943	AGTCCACTATGCTGTGTG	CTTGGAGTAAATGCTCGC	129
96P8cF	AQ762040	GGAAAGAGATTACCAACCACAG	GGAGAATCTACCAGTTCTTGAG	127
8J24cF	AQ762038	TGAATTGGTTGGAGGAGCAC	TGCAGTCTCCTAAGGAAGTC	113
110O11cF	AQ761944	AAAGGCCAGCATGATTGG	CTTACCATGGCCAGAATG	105
215N23cF	AQ761975	ATGGCAAAAGCAGGAGTGAGAGAG	GGGTAAATCCTTCATGGCTGAGTG	119
8J24cR	AQ762039	GCAAAGGAGCAATTTGGTATATGG	ATACACTCAGCCATCTTATCTGAG	121
80M18cF	AQ762036	TGCCAATGACCGAGATTG	TGTAACTTGGCCAAGGTG	238
178P9cF	AQ761962	TGAGATACGTGCAACTCC	CTCTTCCTCACCTCAGGC	116
215N23cR	AQ761976	ATGGCAAAAGCAGGAGTGAGAGAG	GGGTAAATCCTTCATGGCTGAGTG	179
233G14cR	AQ761986	TCTTCTCCAAGCTGCACACC	ATTAGCTGCTACTGGTGGCC	115
178P9cR	AQ761963	ATGCCCAATGGAATTCCG	GGATTTCACTCTTCTGGG	136
233G14cF	AQ761985	AAGAGGCATGTATCACCAGG	AGAATTGCAGCTGGAGGATG	156
35013cF	AQ762048	TCCATATCTGTCACCTGAAG	CAGCACACAGAATTATGACC	96
y11Tr	AQ797401	GATTCCAGTGTCACCACT	CCATGCTAATTGTGCACTG	171
35O13cR	AQ762049	CCTCTGAGTTTGCATCTC	GGTCAATTAGAGCTGGAC	111
238N13cF	AQ761987	ACCCTCTTCTTAGGGTAGAG	TGTTCTTCTGAGGACAGCTG	136
193G4cF	AQ761967	TTTGTATGGCTCCTGTGCAC	TAAAGGGTTACAGCGGACAC	109

TABLE 2—*Continued*

Marker name	GenBank Accession No.	Forward primer sequence	Reverse primer sequence	Product size (bp)
56G5cF	AQ762028	TGCTAACTCAACTTTACAAGC	AATTTGCTCATCATATGCTGC	118
193G4cR1	AQ761968	TTCAGATGGTTGGGATGC	ATTACCTCTGCAGGAGAC	127
98H19cR	AQ762042	AATCCCTCTTGGGCATTC	CGTAGGGAAATGGATTGG	108
y13Ur	AQ797406	GAAGGAGAAAGGAAGGAAGC	TTAGGCAGAGACAACTGC	135
120K12cF	AQ761949	ATTACTCCCTTTGAGAAGCC	TAAAACCAGGCAATCCAGGG	145
250F22cR	AQ761995	TCTGCACAGAGCCTTTCAC	GCTGATTGCTGTCTTTGTG	153
98H19cF	AQ762041	TTTCCATCTCAGACCCTTCG	AGACTCTCATAGGGTTGCTG	104
y10Ur	AQ797400	GTATGCAGGGCATATGTTC	TTTCCACCTGCAGTCTGT	118
253N17R	AQ761997	AGGTACTGGGTGAACAGTGGT	AGTTGGCATCTCTGCCTGCCTG	275
135A22R	AQ761954	AAAGGGTGTGGTTAGCACTG	ATTTGTGCTCTCTGCCTGTG	200
55H19cR	AQ762027	GACAACAGAAGTACCAGG	TGGAGACATGTTCATGGC	114
y12Ua	AQ797404	CAATGAAGTACTTCCTGC	TTCTTATTTCCCCAGGTC	83
135A22F	AQ761953	CACAGTGAAGGAAACAAC	CCATAAGAAAGAATGAGATC	219
253N17F	AQ761996	TGAGAAATAGCAGGGGAAGC	ATGTTTGAGACTAGCCTGGG	190
284A1F	AQ762008	CTGTGTCAGGAAACCAGGTT	ACGGACTAATGACAGTCG	119
279B7cR	AQ762005	GCAACATCTATACCCACC	CTGAATACGGTGGAAGTC	171
62L3cF	AQ762032	ACTTGTTGGTGAGCTCTGTG	AGACTGGAGTCCAATTCAGG	173
55H19cF	AQ762026	GAAAGACACTGGAACTGG	TTACTTCCCAAGGCTCAG	170
y11Ur	AQ797402	GCCAAGTATGTGTCAATG	TTCACTGAAGGACAGTGG	178
62L3cR	AQ762033	AGAGATGTGAAAACCCTGAC	CCTCTCATAGTTTTGATGGC	167
279B7cF	AQ762004	CGATCCTGAGATGGAAAG	CTTTCATGCTCACACTGC	125
37P10cR	AQ762051	TGCCCTCATGCAGTTTACC	TGCATGTAATCTGTTCCCTG	121
204I12cR	AQ761971	TAATACCCAAGGTGTGGTC	GTGCCTGTTACTCTATG	319
279M17cF	AQ762006	TTATGTGTCAGGGAGTGG	TTAGAGGTACAGGTACCAGC	124
37P10cF	AQ762050	CATACTGTTAGGATGGCTGC	TACTTGAGGGAGAAGTAGGC	180
59O14cR	AQ762030	TTTTCCCTGAGCACCTAG	CATGTCCCATGATGGAAG	164
206I2cR	AQ761973	CCAGCTGTGAACTCTTTG	TTAGAATAGGGCCAGCAC	274
204I12cF	AQ761970	AGTGACGGAAAATCCCTG	TTGGTGACTCCACCTTAC	137
110H19cF	AQ761941	GGACTTTGGACAAAAGTC	GAAGCCGATTTGAGTTTC	141
206I2cF	AQ761972	GCGAACATATTCCAGTTGTC	CTCCAGGTATTTAAGGTAGC	94
59O14tF	AQ762031	GGTGGAATCAGTCTCTTG	TTGCCTCTTGATAGGAGG	145
110H19tR	AQ761942	GCCAAATAAAGAGCTCCC	CATGTTCTCTCTGCATGC	103
279M17cR	AQ762007	GTTAAGTCTGGCTTCTCTCC	TGCTGTTGAGAGACCATGTG	122
y37T	AQ797420	CTAAAGCTTGAAGTGAGAGG	GGTCACTGGAAAAAGAGCACAG	183
118G19cR	AQ761948	CCAGTGTAAGGTTAGGCAAG	CACCTGACAGATTCTAGTGG	141
366A8F	AQ762020	GGTATTGGCAGTGATATGC	GTTGGCATTGTAATCTC	166
153O3cR	AQ761957	TGCCAGTTGAATGGTGTC	GCTAGCCTAATTAGAGTGG	130
164L12tF	AQ761959	TCACATTAGCCAGCCTTCTC	CAAACTTGGCCTTTGCCTTC	125
268N14cR	AQ762000	ATACTCAGCATGCTGTGGTG	TCCATCTCTAGAGGTGAAGC	165
366A8R	AQ762021	GCCAACATGAACTCCAAG	GTTGCCATTTGAAGCCAC	171
153O3cF	AQ761956	ATGGTTTCTTGGAGGGTTCC	TATCCCTAGTCTTGAGCAGG	112
248A21cF	AQ761990	GTCTAGCTTTTCTAGCCATC	AAGAGAGGAGTCATAGCTAG	112
268N14cF	AQ761999	CAGTGCATCTTAGTTTGGAC	TGGTCTAGTGTTACACTGTC	93
77J13cF	AQ762034	ATTCCCAGTGCTTAGCATG	AGTAACTAGTTCGGTTC	85
174L6qF	AQ761961	AAAATATTCCAGTGCCTGGGCTGC	GTCTTGTCCAATTGAGTGCTGAGG	101
77J13cR	AQ762035	TAGTGAAGATCCGTGCAG	TACCTTGCAGGTGCTAAG	167
216P5cR	AQ761979	CTTCCTTGGACTATTGG	ATATCACGCCACTGCATTCC	225
116D22F	AQ761946	TGCTTAATAAAGCCAGCC	TTTAACGGAAGTCCTGAGGAC	196
304P7cR	AQ762011	ATGACAGAAGGCAAAGGG	TGTTTAGGTCATGGGAGCAG	163
116D22R	AQ761947	GACACCATCAACATGGTGAGT	TTGCTCAAGATCCCAGAGCCA	269
y10Ta	AQ797399	TTGAGTCATGGCAGCAGATC	TTGAGTCATGGCAGCAGATC	157
168F1cR	AQ761960	AAATATGCATCGCCTTCCGC	TTGCTGTGACTCTGAACTGG	150
25B1cF	AQ761998	GATGCCACAACCAGATAAGG	TATACTGGTGCAGACTGGTG	139
202K21cF	AQ761969	CTATTCAGTATGGGTGTC	GTGTCCTTGCATGCCCTTGAG	238
y8Tr	AQ797427	GCTGCTTTAAGATAGCCATC	TATCCTGATCACACAGAGCC	225
24B12sF	AQ761991	AGACTAGACTTCGAGTGCTG	GGACTTGTTTACTGACAGG	130
216P11F	AQ761977	ATGCTGCCTATGATCTCTTTCT	CTCATAGGCTAGTCTTGAGTG	187
156O9sR	AQ761958	TGCACAGCTTGTTCATGG	AAACACACATTGGGTGGC	235
87E12sF	AQ762037	GTGCCACATGTATGTTTAC	AACTGTCTCCTAACATAGG	155
y7Ur	AQ797426	GGCAGATTTGTAGGTGATG	TAGCAATCATGTGGACAGC	185
24B12sR	AQ761992	GAACTGCAGAAGTGCAAAC	TTCCTAGTGTGTGGTTTCC	124
216P11R	AQ761978	TTTATGTCCCTGGAGAGGCA	AGCTTCTGACACACCTCCTCT	242
13C1cF	AQ761955	TCAGCCTAATGCTATGAAGTG	GAGATTAGACTCCCAACTCTC	131
y6Ua	AQ797424	CATTTGGAGAGACATCAACA	CTATCTTGGTTCTATGGGTA	78
y50U	AQ797422	TGCACTTTCCATTCTGC	TTGAGCACTTTAAGGAG	164

TABLE 2—*Continued*

Marker name	GenBank Accession No.	Forward primer sequence	Reverse primer sequence	Product size (bp)
y8U	AQ797428	ATAGTCCTAGTCCCTTGC	TTTCACCGTGTTAGCCAG	122
y6T	AQ797423	CCTCATACCATAGCTAC	GCCATTCACCTCACTG	97
y48T	AQ797430	CGCAATAATAGCCTTAGGC	CATGAAATAAAGGGAAGGG	101
y3U	AQ797421	CTGTTCACATGCAGCTAC	AGCAGGCATATGACCA	107
y2U	AQ797416	CTTCCCAAGGTGCTAAG	GATCCTCAATAAGTCCTGAG	151
y1T	AQ797411	CAGCATTCTATAGCTC	TTGAAGGGCAGGACAGA	94
y48U	AQ797429	CACAAGGTTCATTCTAGTCATC	TGTTTTCTTGCGGGTGTATGTG	165
y2T	AQ797415	GGTCCCTGTTACCTTC	CAGACTGTGAAGACTAC	126
y35U	AQ797419	CATCATACCAAGTCACATG	AACTTCCATTGGCAGAATG	108
y33U	AQ797418	CAGAAAACCTACATCTTCCC	AAGCCTGAGTGAGATACTGG	141
y30U	AQ797417	ACAAATATACAAATGCA	TTGAGATAGTGCCATGC	80

derived STSs and published STS markers confirmed within the YAC contig between and inclusive of D1S466 and D1S1642 were used to screen commercially available BAC and PAC library DNA pools by PCR. This starting collection of 31 markers in an approximately 6-Mb interval suggests an average distance of 210 kb between adjacent markers. The initial screen with 31 markers resulted in the identification of 78 BAC clones. STS content analysis was performed to generate BAC contigs; however; due to low marker density, only small contigs were generated that were spaced by distances of unknown length with clustering of markers in certain areas (Fig. 1). D1S466 and D1S2619 were both contained within three overlapping BACs (b278C23, b7F19, and b285P16), and D1S2818 was also linked to this cluster, as it was present within BACs b278C23 and b7F19 (Fig. 1). A total of 8 BAC clones were identified by the marker D1S202, with 4 of these clones also being positive for y8Tr and 1 of the 8 clones being positive for y7Ur, suggesting clustering of these markers.

All BAC and PAC clones identified by markers known to map to this interval were subjected to direct end sequencing to generate new STSs primarily to facilitate chromosome walking (Table 2). End sequences were compared against the NCBI nonredundant nucleotide database using PowerBlast for the identification of repetitive elements and also to identify homologous sequences including gene and EST homologies (Ferlanti *et al.*, 1999). STS 245D23cR had 100% homology to the ATP-dependent RNA helicase (*DEAD*) gene sequence, confirming the exact location of this gene within the 1q25 contig (Fig. 1). A total of 108 new STSs were generated from BAC end sequencing and were subsequently mapped to the D1S466–D1S1642 interval (Table 2). These new STSs were used to screen BAC and PAC library DNA pools further by PCR and resulted in the identification of 133 unmapped BAC/PAC clones (Fig. 1). STS content mapping with these new markers resulted in the completion of a 6-Mb sequence-ready contig of the 1q25 region as indicated in Fig. 1.

Restriction Fingerprinting

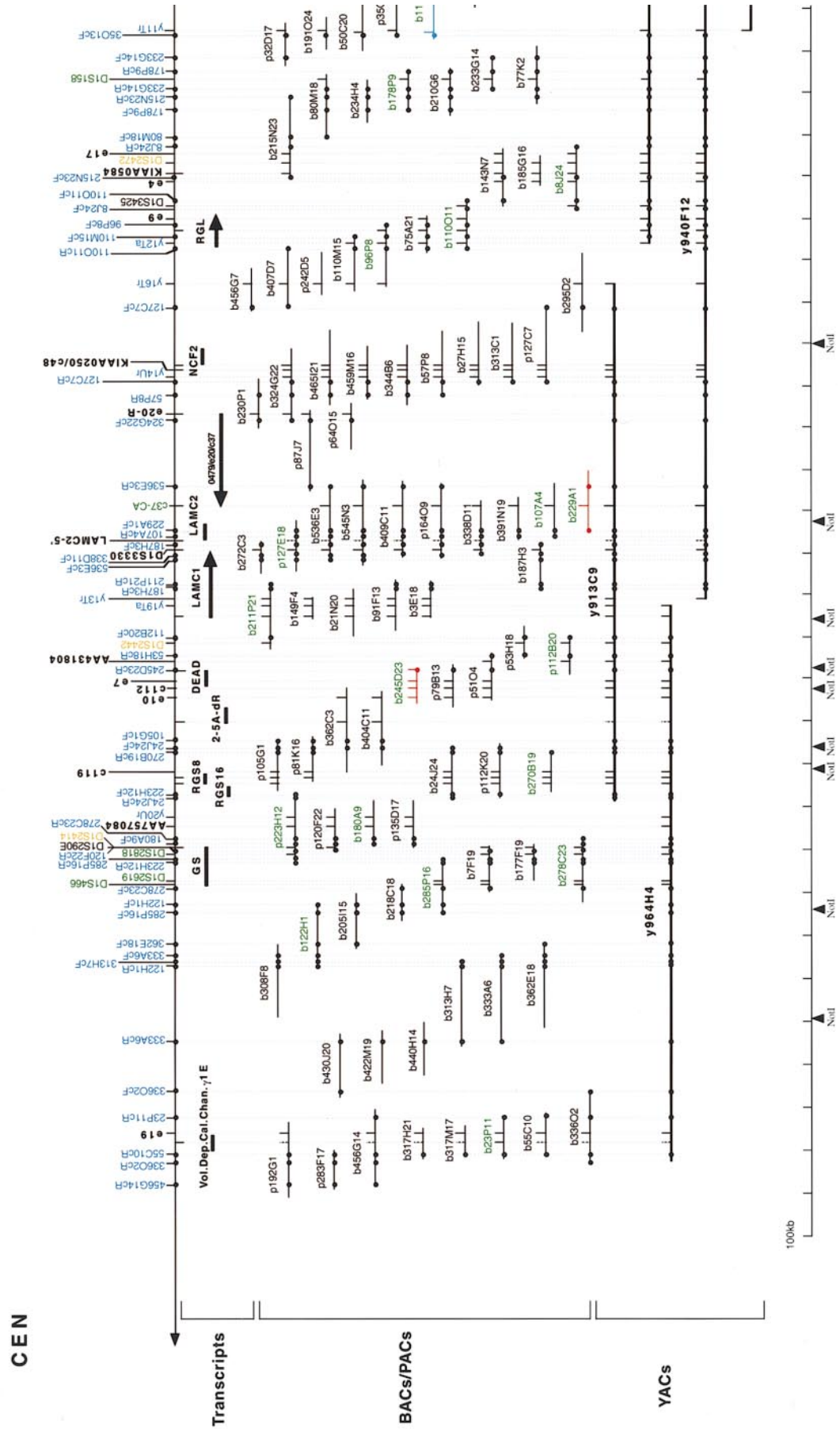
A subset of BAC and PAC clones from this contig have been subjected to restriction fingerprinting as a prerequisite for genome sequencing by the Sanger Genome Sequencing Centre. Sanger chromosome 1 sequence contig 70 initially contained a small group of overlapping PAC clones identified by the Sanger Centre using markers known to map to the 1q25 region, and finished sequence for several of these clones has been archived. Overlaps determined by STS content mapping between adjacent BAC/PAC clones identified as part of the HPC1 positional cloning effort have been confirmed by fingerprinting experiments at the Sanger Genome Sequencing Centre (data not shown). To date, clones representing the region flanked by D1S2127 and D1S202 represent Sanger chromosome 1 contig 70. These clones are currently in the queue for shotgun library construction and sequencing. Clones representing the D1S466 to D1S2127 region are currently being fingerprinted to determine the integrity of the 6-Mb BAC/PAC contig constructed as part of the HPC1 positional cloning effort.

Identification of Novel Simple Sequence Repeat Polymorphisms

For the genetic refinement of disease loci, we attempted to isolate novel simple sequence repeat polymorphisms (SSRPs) from the 1q25 contig. Five BAC/PAC clones (b238N13, b277P6, b204I12, b59O14, and b164L12) were used as templates to generate small insert libraries. Libraries were probed with a pool of radiolabeled oligos (CA_{10}, AG_{10}, and $GATA_8$) for the identification of novel SSRPs within the interval for use in genetic refinement of disease gene loci. A total of seven novel SSRPs were generated by this method (Table 3) and mapped back to the contig (Fig. 1). Sample sequencing of BAC b173P17 and b77J13 also led to the identification of seven novel SSRPs (Table 3).

Transcript Mapping and Identification

For the identification and mapping of transcription units within the 1q25 contig, we performed database


1q25 Contig
CEN
Transcripts
BACs/PACs
YACs
Vol.Dep.Cal.Chan.γ1E
GS
RGS16
RGS8
2-5A-dR
DEAD
LAMC1
LAMC2
NCF2
RGL
y964H4
y913C9
y940F12
100kb
NotI

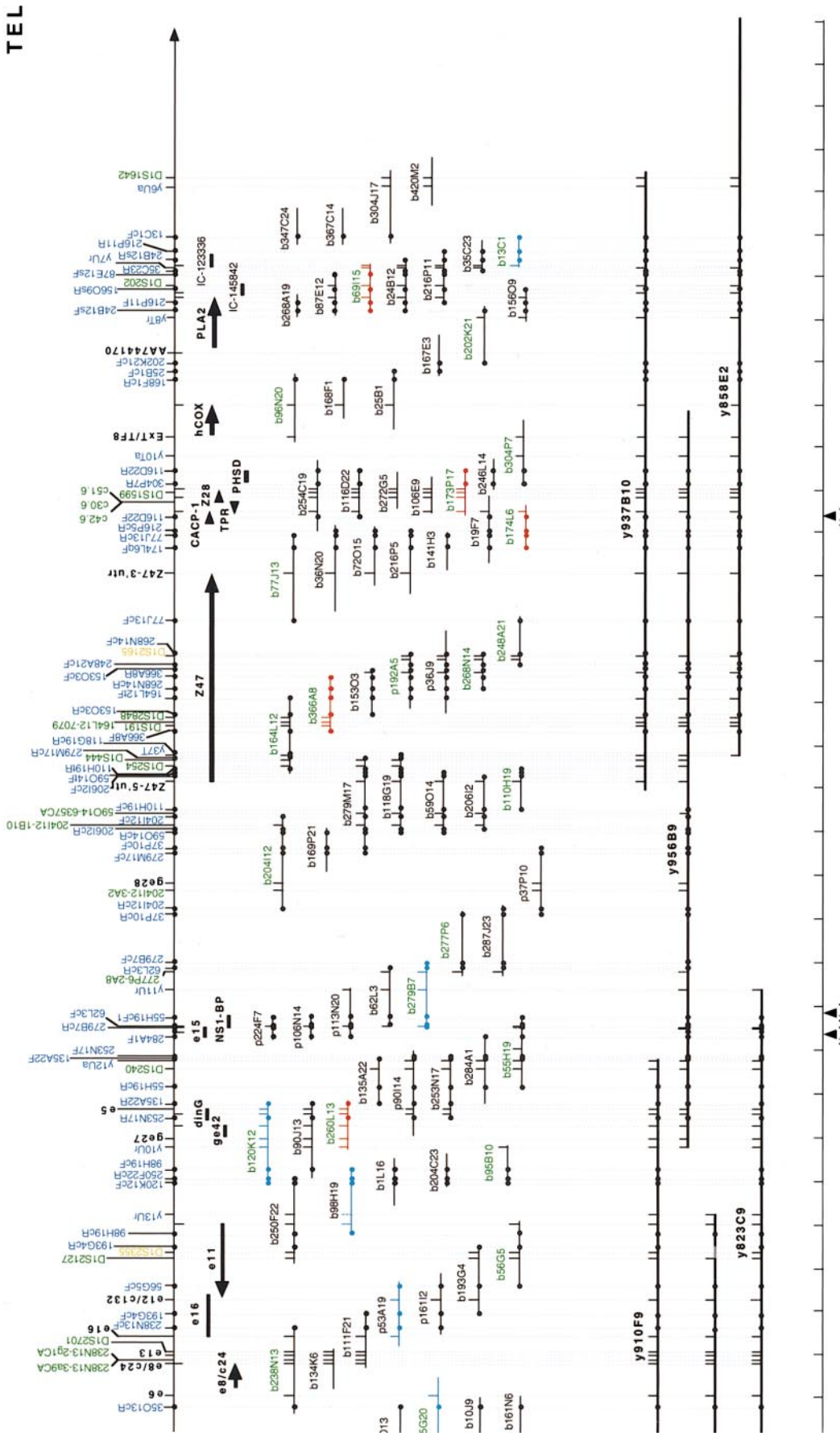


FIG. 1. Diagram of the 1q25 contig. DNA markers labeled in blue are novel nonpolymorphic unique STSs. DNA markers labeled in gold are public access nonpolymorphic STSs. DNA markers labeled in green are either public or novel simple sequence repeat polymorphisms (public SSRPs are labeled by D1S designations). Transcript-related DNA markers are labeled in black and represent either public access ESTs or exon trap sequences that have been precisely mapped by our group. Closed circles represent novel STSs generated from BAC/PAC terminal sequences. All confirmed genomic clones (YAC, BAC, PAC) in this contig are drawn to scale. Red lines representing BAC/PAC clones have been used as templates for sample sequencing. BAC/PAC clones depicted as cyan lines represent clones in the sequence assembly phase at Sanger Genome Sequencing Centre. BAC/PAC clone names labeled in green have been used as templates for exon-trapping experiments. Arrows representing genes are pointing in the 5′ to 3′ direction. *Not*I sites are based on BAC/PAC digests.

TABLE 3

Novel SSRPs

Marker name	GenBank Accession No.	Forward primer sequence	Reverse primer sequence	Size (bp)
238N13-3A9	AF181673	AGAAAGGGTGATGCCAGTAC	ACCACAACCTTGCATGAAGG	137
238N13-2G1	AF181674	TTATTGGAGAACTAGGGGC	GCCACATAAATGGCAGTAG	137
277P6-2A8	AF181675	GATCCTTCACTAGTTACTCCC	TTATCCCTCGAGTTAACAGCC	118
204I12-3A2	AF181676	ATCCAGGTTGCTGCAAATGC	CTATGGAACCAGCCTGAATG	180
204I12-1B10	AF181677	CACTTGCAAGCTTAGGTCAC	ACTGGCTATTGGTCAGAGAC	154
59O14-6357	AF181678	GTAATCCCAGCACTTTGGAAG	CAGGTCTATCAAATGTTGTGGC	244
164L12-7079	AF181679	TCCCGTTGATGATTTGAGAG	GGGATCAAGAATCAAGGTAC	254
c84CA	AF181669	AATTGAGCCACTCTCAGG	GAAGTAGAATGGTGACCC	133
c114GA	AF181670	CACAGAGGAGATAAACAGTG	TTTCAGTGTCACTGCAATGG	178
c124CA	AF181672	AGGCAGGAGAATTCTCTCTC	AAGCAGATTGTCCTCCATGG	129
c117TAA	AF181671	ATGCCACTATCTGAGTCC	ACATCTTCCTGGAACTCC	193
c42.6	AF172081	CCAAACACCTTCCATTAGGC	TTAGGAGATGGAGGTGAGAG	315
c30.6	AF172079	GTTTGCATCATGAAAATGAGTGCAG	ATATGAGGTGATTCTAAGTAGCAGG	175
c51.6	AF172081	TAATTGTCTGCCACGAGTGC	AGTTCGTCAGAAGTGTGTGC	162

searching, direct hybridization of cDNA libraries, exon trapping, and sample sequencing. A thorough search of the NCBI Human GeneMap 98 (Deloukas *et al.*, 1998) was performed for an expanded region (D1S212–D1S412) due to the relatively low resolution of the GeneBridge 4 radiation hybrid panel used for EST mapping. Primers were derived from EST or gene sequence and were used for mapping against the contig. All ESTs that mapped to the contig were searched against UniGene (http://www.ncbi.nlm.nih.gov/UniGene/index.html) to identify EST clusters to collect the greatest amount of known sequence for each potential transcript. In addition to mapping ESTs and known genes from the database, we used several molecular techniques to identify novel transcripts within the 1q25 contig. Several BAC inserts were purified, radiolabeled, and used as probes to screen the IMAGE consortium high-density cDNA filters. Positive clones were mapped back to the contig by either PCR or Southern blotting. Also, a subset of BACs and PACs were selected for exon-trapping experiments (Fig. 1).

A number of BACs (b260L13, b174L6, b173P17, and b69I15) were subjected to shotgun library construction in M13 vector, and the corresponding libraries were sample sequenced (Fig. 1). We used the PHRED/PHRAP/CONSED suite (Ewing *et al.*, 1998; Gordon *et al.*, 1998) for contig assembly and sequence editing. Sequence contigs were searched against the NCBI nonredundant nucleotide database using the PowerBlast algorithm (Altschul *et al.*, 1997; Zhang & Madden, 1997). Prior to gene analysis, repeat sequences were removed using Repeat Masker (A. F. A. Smit and P. Green, unpublished data). Sequence data were organized and initially analyzed using WebBlast (Ferlanti *et al.*, 1999). Sequence contigs were subsequently analyzed with several gene prediction programs including GRAIL, GENSCAN, MZEF, and FGENES (Guan *et al.*, 1992; Burge and Karlin, 1997; Zhang, 1997; Solovyev and Salamov, 1997) using an analysis workbench called GeneMachine (http://genome.nhgri.nih.gov/genemachine). All BLAST information and exon prediction results were viewed and annotated using Sequin (http://www.ncbi.nlm.nih.gov/Sequin).

Several known genes and novel transcripts were identified by sample sequencing, as were ESTs that were also identified by other methods (Table 4). Information derived from sample sequencing is shown in Table 5. Sample sequencing of b173P17 resulted in the precise mapping of two known human genes (TPR and Phosducin) and one EST Z28464 (Table 5). The previously unmapped megakaryocyte stimulating factor gene (MSF) was also placed within the 1q25 contig by sample sequencing (Fig. 1). MSF and the human TPR gene share sequence at their respective 3′ untranslated regions and are oriented in opposite directions. Sample sequencing also revealed that the respective 5′ ends of TPR and the Z28464 transcript are very close to each other, and these genes are oriented in opposite directions. This is evident as the 5′ ends of these two transcripts are within the same b173P17 sample sequence contig (Table 5). Several exons from the phospholipase A2 gene along with genomic sequence for this gene were identified from sample sequence derived from the BAC clone b69I15. We have also mapped the human homologue of the *Caenorhabditis elegans* hemicentin gene him-4 (AF074901) by sample sequencing. This gene, which we call Z47, extends over approximately 600 kb of human genomic sequence. Table 4 shows corresponding accession numbers for those transcripts mapped to the 1q24–q31 region by our group and the methods by which those transcripts were identified. A detailed analysis of transcripts from this contig including expression profiling, tissue distribution, and extension of partial transcripts into full-length transcripts is under way (manuscript in preparation).

DISCUSSION

Several hereditary disease loci have been shown to map to the 1q24–q31 interval including HPC1, CACP

TABLE 4

Transcripts Mapped to the 1q24–q31 Region

Unigene	Accession No.	Name	Source	Known information
Hs. 14553	L21934	SOAT	DB	Sterol-*O*-acetyl transferase; formation of cholesterol ester from cholesterol
Hs. 7175	A006H48	LHX4	DB	Similar to mouse LIM homeobox gene
Hs. 1355	J05036	CTSE	DB	Cathepsin E gene; aspartic protease found in the gastrointestinal tract
None	U97276	QSCN6	DB	Quiescin gene; may play a role in transition into cellular quiescence
Hs. 65441	L29384	Vol.Dep.Cal.Chan-g1E	DB/ET	Voltage-dependent calcium channel α-1E subunit
Hs. 126918	stsg40094	e19	DB	
Hs. 170171	Y00387	D1S290E/GLUL	DB	Glutamine synthetase; control of normal body pH
None	AA757084	AA757084	ET	
Hs. 183601	U70426	A28RGS14/RGS16	DB	Regulation of G protein signaling
None	None	RGS8	ET	Regulation of G protein signaling/human homologue of rat RGS8
Hs. 20982	H17666	c119	DB	
None	L10381	2-5A-dR	DB	2-5A-dependent RNase; regulator of interferon action
Hs. 18033	stsg16431	e10	DB	
Hs. 23756	AI042017	c112	DB	Similar to *E. coli N*-acetylneuraminate lyase
Hs. 36300	A009P09	e7	DB	
Hs. 74578	L13848	DEAD	DB	ATP-dependent RNA helicase
None	AA431804	AA431804	ET	
Hs. 87428	J03202	LAMC1	DB	Laminin C1/LAMB2; cellular adhesion molecule
Hs. 54451	WI-7043	LAMC2	DB	Laminin C2/LAMB2t; cellular adhesion molecule
Hs. 158244	KIAA0479	KIAA0479/e20/c47	DB	
Hs. 15087	KIAA0250	KIAA0250	DB	
Hs. 949	M32011	NCF2	DB	Oxidative phosphorylation through the NADPH enzyme system
Hs. 146957	KIAA0959	KIAA0959	DB/ET	Homology to mouse RGL; member of the Ras pathway
None	T79647	e9	DB	
Hs. 26835	Z38864	D1S3425	DB	
Hs. 192208	R48958	e4	DB	
Hs. 106794	KIAA0584	KIAA0584	DB	
Hs. 128769	stsg31677	e17	DB	
H2.200305	stsg30862	e16	DB	Manosidase gene
None	A009C04	e6	DB	
None	WI-22797	e13	DB	
Hs. 16769	A009Q28	e8/c24	DB	
Hs. 48778	STSG22079	e11	DB	
Hs. 117788	STSG22920	c132/e12	DB	
Hs. 52438	AA657408	ge27	ET/SS	
None	AA469731	ge42	ET/SS	Possible zinc finger protein
H2s.28441	Y10571	dinG	ET/SS	Possible transcription factor; has RING domain
None	stgdb191000	e5	DB/SS	
Hs. 113928	stsg27101	e15	DB	
Hs. 197298	AJ012449	NS1-BP	DB/ET	Influenza virus binding protein; *Drosophila* ring canal protein
None	None	Z47	SS/ET	Homology to *C. elegans* hemicentrin 4 protein (Z47068)
Hs. 154083	U70136	MSF	SS/BH	Synovium lubrication and growth stimulation
Hs. 169750	X66397	TPR	SS/BH	Promoter translocation to TRK to form met oncogene
Hs. 69388	Z28464	Z28	SS	
Hs. 550	M33478	PDC/PHSD	SS	Human homologue to rat phototranducin; retinal protein
Hs. 92309	D28235	D1S3433/PTGS2/COX2	DB/ET	Involved in the inflammatory process
None	AA744170	AA744170	Sanger	
Hs. 3278	M72393	PLA2G4A	DB/ ET/SS	Catalyzes release of arachodonic acid from membrane phosopholipids
Hs. 144508	R79368	IC-145842	BH	
None	R00375	IC-123336	BH	
Hs. 65765	Cda1be02	Z39494	DB	
None	stSG23258	M78285	DB	
None	sts-M78354	M78354	DB	
Hs. 75256	S59049	RGS1/BL34	DB	Mediate programmed changes in response to extracellular signals
Hs. 78944	L13463	RGS2/G0S8	DB	Control of proliferation and differentiation in normal cells
None	AF147717	UCH37/c17	DB	Thiolproteases that catalyze proteolytic processing of ubiquitin
Hs. 5722	A005A18	c190	DB	No homologies
Hs. 181353	Y15014	c78	DB	β3 galactosyl transferase/brainiac?

Note. DB, database searching; ET, exon trapping; BH, BAC insert hybridization to cDNA. All known information was derived from Online Mendelian Inheritance in Man or from homology to other known proteins.

TABLE 5

Sample Sequence Contigs Containing Gene Sequences

Sequence contig	Accession No.	Clone source	Transcripts (GenBank Accession No.)
173P17-C19.8	AF172075	b173P17	Phosducin (M33478)
173P17-C20.8	AF172076	b173P17	Z28 (Z28464)
173P17-C23.8	AF172078	b173P17	Z28 (Z28464)
173P17-C24.8	AF172079	b173P17	Phosducin (M33478)
173P17-C25.8	AF172080	b173P17	Phosducin (M33478)
173P17-C26.8	AF172081	b173P17	TPR (X66397) and Z28 (Z28464)
260L13-C32	AQ839801	b260L13	ge42 (AA469731)
260L13-C19	AQ839802	b260L13	ge42 (AA469731)
260L13-C39	AQ839803	b260L13	ge42 (AA469731)
260L13-C53	AQ839804	b260L13	ge42 (AA469731)
260L13-C56	AQ839852	b260L13	ge42 (AA469731)
260L13-C55	AQ839853	b260L13	ge42 (AA469731)
260L13-C49	AQ839811	b260L13	dinG (Y10571)
260L13-C35	AQ839812	b260L13	dinG (Y10571)
260L13-C33	AQ839813	b260L13	dinG (Y10571)
260L13-C50	AQ839814	b260L13	dinG (Y10571) and e5 (stgdb191000)
260L13-C54	AQ839854	b260L13	ge27 (AA657408)
69I15-C77	AQ839850	b69I15	PLA2G4A (M72393)
69I15-C78	AQ839855	b69I15	PLA2G4A (M72393)
69I15-C79	AQ839851	b69I15	PLA2G4A (M72393)
174L6-C69	AQ839805	b174L6	MSF (U70136)
174L6-C80	AQ839806	b174L6	TPR (X66397)
174L6-C82	AQ839807	b174L6	TPR (X66397)
174L6-C84	AQ839808	b174L6	MSF (U70136)
174L6-C93	AQ839809	b174L6	TPR (X66397)
174L6-C105	AQ839810	b174L6	TPR (X66397)

syndrome, and the HRPT2 syndrome. A 20-Mb region flanked proximally by D1S212 and distally by D1S533 is covered in 40 megaYACs and serves primarily as a resource for the localization of ESTs previously mapped at radiation hybrid resolution. YAC contig confirmation and assembly consisted of a search of the MIT–WICGR human physical map for clones believed to contain a set of polymorphic markers within the HPC1 interval along with novel STSs generated from YAC insert terminal sequences. Furthermore, we have constructed a complete 6-Mb sequence-ready contig of 1q25 flanked proximally by D1S466 and distally by D1S1642 using BACs and PACs. Genomic clones were identified by PCR screening of commercially available library DNA pools. This physical map consists of a total of 9 nonchimeric CEPH megaYACs, 160 BAC clones, and 34 PAC clones.

YAC clones were selected from the database based on the notion of being nonchimeric. However, two new clones identified by our group to close a gap in the contig were found to be chimeric. Interestingly, 3/160 BACs (1.9%) from the 1q25 contig were found to be chimeric (as derived from mapping end sequences to somatic cell hybrid DNA; data not shown). This may in fact represent an underestimate of the chimerism rate for BAC libraries screened here, as all BAC end sequences were not mapped from every clone identified due to the presence of repetitive elements in some BAC end-derived sequence.

Physical maps serve as a resource for identifying new polymorphic markers for genetic refinement of disease loci. A total of 10 new SSRPs were identified by probing BAC insert subclones with primers representing polymorphic stretches and also by the analysis of sample sequence contigs. Several public databases suggest a genetic distance of 3 cM for the region flanked by D1S466 and D1S1642. The GeneBridge 4 radiation hybrid data in the MIT–WICGR human physical map database suggest a distance of 6 cR or 2 Mb based on an average of 334 kb/cR for chromosome 1. However, we estimate an actual physical distance of 6 Mb based upon our contig. Three nonoverlapping megaYAC clones from the contig, y913C9 (1140 kb), y910F9 (1530 kb), and y937B10 (1430 kb), suggest a minimal distance of 4 Mb for this interval. Although the correlation between genetic distance and actual physical distance is not absolute, the lack of recombination here may suggest that this region is a cold spot for recombination.

Physical clones also serve as a means for mapping and identifying candidate genes. All transcripts identified by exon trapping and sample sequencing turned out to have associated ESTs in the database. These findings suggest that EST coverage may be close to actual gene representation for this region of human chromosome 1. Sizing of BAC inserts using *Not*I gives an estimate of CpG content for this region of the genome (Kusuda *et al.*, 1990). CpG islands have been shown to be associated with 60% of the known genes throughout the genome (Larsen *et al.*, 1992). *Not*I digests of BAC clones within the 1q25 contig revealed a paucity of *Not*I sites with one identified at the distal 2

Mb of the contig flanked by D1S254 and D1S1642. Only 10 transcripts mapped to this area of the contig. Similarly, the central 2 Mb of the contig flanked by D1S158 and D1S254 contain two *Not*I sites with 13 possible transcripts mapping to this area of the contig. The most proximal area of the contig flanked by D1S466–D1S158 contains 8 *Not*I sites and contains 22 mapped transcription units. This finding correlates well with reports that suggest that, in general, clustering of CpG islands is positively related to transcriptionally active regions of the genome (Cross and Bird, 1995). A more detailed restriction map using multiple rare-cutter enzymes is needed to support these findings more definitively. It is also possible that many more transcription units may be identified within this interval.

All transcription units mapped to this interval represent candidate genes for several human hereditary diseases. Recently, we have identified mutations in a gene formerly known as MSF in patients with CACP (Marcelino *et al.*, 1999). We are now attempting to identify mutations for HPC1 and HRPT2 in genes that we have mapped to the 1q24–q31 interval. Moreover, a 6-Mb sequence-ready BAC/PAC contig has been submitted to the Sanger Human Genome Sequencing Centre for human genome sequencing as well as confirmation of contig integrity. This high-resolution 6-Mb BAC/PAC contig should provide a major resource for human genetics, disease gene research, and the ultimate physical map; the human genome project.

ACKNOWLEDGMENTS

The authors thank G. Olayinka, H. Tinkle, J. Turner, S. McGee, M. Phipps, A. Duda, and L. Michalowsky for outstanding technical support. The authors recognize M. Warman and J. Marcelino for the CACP collaboration. This work was supported through the National Human Genome Research Institute Intramural Program.

REFERENCES

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25:** 3389–3402.

Athreya, B. H., and Schumacher, H. R. (1978). Pathologic features of a familial arthopathy associated with congenital flexion contractures of fingers. *Arthritis Rheum.* **21:** 429–437.

Bahabri, S. A., Suwairi, W. M., Laxer, R. M., Polinkovsky, A., Dalaan, A. A., and Warman, M. L. (1998). The camptodactyly-arthropathy-coxa vara-pericarditis syndrome: Clinical features and genetic mapping to human chromosome 1. *Arthritis Rheum.* **41:** 730–735.

Berthon, P., Valeri, A., Cohen-Akenine, A., Drelon, E., Paiss, T., Wohr, G., Latil, A., Millasseau, P., Mellah, I., Cohen, N., Blanche, H., Bellane-Chantelot, C., Demenais, F., Teillac, P., Le Duc, A., de Petriconi, R., Hautmann, R., Chumakov, I., Bachner, L., Maitland, N. J., Lidereau, R, Vogel, W.,Fournier, G., Mangin P., Cohen, D., and Cussenot, O. (1998). Predisposing gene for early-onset prostate cancer, localized on chromosome 1q42.2–43. *Am. J. Hum. Genet.* **62:** 1416–1424.

Bulutlar, G., Yazici, H., Ozdogan, H., and Schreuder, I. (1986). A familial syndrome of pericarditis, arthritis, camptodactyly, and coxa vara. *Arthritis Rheum.* **29:** 436–438.

Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268:** 78–94.

Carpten, J. D., DiDonato, C. J., Ingraham, S. E., Wagner-McPherson, C., Nieuwenhuijsen, B. W., Wasmuth, J. J., and Burghes, A. H. (1994). A YAC contig of the region containing the spinal muscular atrophy gene (SMA): Identification of an unstable region. *Genomics* **24:** 351–356.

Church, D. M., Stotler, C. J., Rutter, J. L., Murrell, J. R., Trofatter, J. A., and Buckler, A. J. (1994). Isolation of genes from complex sources of mammalian genomic DNA using exon trapping. *Nat. Genet.* **6:** 98–105.

Connors, T., Burn, T., and Landes, G. (1994). Exon trapping vector pSPL3-CAM: Improved shotgun subcloning of cosmid-derived fragments. *Focus* **16:** 111–112.

Cooney, K. A., McCarthy, J. D., Lange, E., Huang, L., Miesfeldt, S., Montie, J. E., Oesterling, J. E., Sandler, H. M., and Lange, K. (1997). Prostate cancer susceptibility locus on chromosome 1q: A confirmatory study. *J. Natl. Cancer Inst.* **89:** 955–999.

Cross, S. H., and Bird, A. P. (1995). CpG islands and genes. *Curr. Opin. Genet. Dev.* **5:** 309–314.

Dausett, J., Ougen, P., Adberrahim, H., Billault, A., Sambucy, J. L., Cohen, D., and Le Paslier, D. (1992). The CEPH YAC library. *Behring. Inst. Mitt.* **91:** 13–20.

Deloukas, P., Schuler, G. D., Gyapay, G., Beasley, E. M., Soderlund, C., Rodriguez-Tome, P., Hui, L., Matise, T. C., McKusick, K. B., Beckmann, J. S., Bentolila, S., Bihoreau, M., Birren, B. B., Browne, J., Butler, A., Castle, A. B., Chiannilkulchai, N., Clee, C., Day, P. J., Dehejia, A., Dibling, T., Drouot, N., Duprat, S., Fizames, Fox, S., Gelling, S., Green, L., Harrison, P., Hocking, R., Holloway, E., Hunt, S., Keil, S., Lijnzaad, P., Louis-Dit-Sully, C., Ma, J., Mendis, A., Miller, J., Morissette, J., Muselet, D., Nussbaum, H. C., Peck, A., Rozen, S., Simon, D., Slonim, D. K., Staples, R., Stein, L. D., Stewart, E. A., Suchard, M. A., Thangarajah, T., Vega-Czarny, N., Webber, C., Wu, X., Auffray, J. C., Nomura, N., Sikela, J. M., Polymeropoulos, M. H., James, M. R., Lander, E. S., Hudson, T. J., Myers, R. M., Cox, D. R., Weissenbach, J., Boguski, M. S., and Bentley, D. R. (1998). A physical map of 30,000 human genes. *Science* **282:** 744–746.

Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8:** 175–185.

Ferlanti, E. S., Ryan, J. F., Makalowska, I., and Baxevanis, A. D. (1999). WebBLAST 2.0: An integrated solution for organizing and analyzing sequence data. *Bioinformatics* **15:** 422–423.

Fujikawa, M., Okamura, K., Sato, K., Mizokami, T., Tamaki, K., Yanagida, T., and Fujishima, M. (1998). Familial isolated hyperparathyroidism due to multiple adenomas associated with ossifying jaw fibroma and multiple uterine adenomyomatous polyps. *Eur. J. Endocrinol.* **138:** 557–561.

Gordon, D., Abajian, C., and Green. P. (1998). Consed: A graphical tool for sequence finishing. *Genome Res.* **8:** 195–202.

Gregory, S. G., Howell, G. R., and Bentley, D. R. (1997). Genome mapping by fluorescent fingerprinting. *Genome Res.* **7:** 1162–1168.

Guan, X., Mural, R. J., Einstein, J. R., Mann, R. C., and Uberbacher, E. C. (1992). GRAIL: An integrated artificial intelligence system for gene recognition and interpretation. *In* Proceedings of the Eighth IEEE Conference on AI Applications, pp. 9–13.

Hobbs, M. R., Pole, A. R., Pidwirny, G. N., Rosen, I. B., Zarbo, R. J., Coon, H., Heath, H., III, Leppert, M., and Jackson, C. E. (1999). Hyperparathyroidism-jaw tumor syndrome: The HRPT2 locus is within a 0.7-cM region on chromosome 1q. *Am. J. Hum. Genet.* **64:** 518–525.

Hsieh, C-L., Oakley-Girvan, I., Gallagher, R. P., Wu, A. H., Kolonel, L. N., The, C. Z., Halpern, J., West, D. W., Paffenbarger, R. S., Jr,, and Whittemore, A. S. (1997). Prostate cancer susceptibility locus

on chromosome 1q: A confirmatory study. *J. Natl. Cancer Inst.* **89:** 1893–1894.

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. H., Hu, S., 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., Faure, 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.

Ioannou, P. A., Amemiya, C. T., Garnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A., and de Jong, P. J. (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.* **6:** 84–89.

Jackson, C. E., Norum, R. A., Boyd, S. B., Talpos, G. B., Wilson, S. D., Taggert, R. T., and Mallette, L. E. (1990). Hereditary hyperparathyroidism and multiple ossifying jaw fibromas: A clinically and genetically distinct syndrome. *Surgery* **108:** 1006–1013.

Kim, U. J., Birren, B. W., Slepak, T., Mancino, V., Boysen, C., Kang, H. L., Simon, M. I., and Shizuya, H. (1996). Construction and characterization of a human bacterial artificial chromosome library. *Genomics* **34:** 213–218.

Kupfer, K., Smith, M. W., Quackenbush, J., and Evans, G. A. (1995). Physical mapping of complex genomes by sample sequencing: a theoretical approach. *Genomics* **27:** 90–100.

Kusuda, J., Hirata, M., Yoshizaki, N., Kameoka, Y., Takahashi, I., and Hashimoto, K. (1990). Over 80% of NotI sites are associated with CpG rich islands in the sequenced human DNA. *Jinrui Idengaku Zasshi* **35:** 277–282.

Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics* **13:** 1095–1107.

Marcelino, J., Carpten, J. D., Suwairi, W. M., Gutierrez, O. M., Schwartz, S., Robbins, C., Sood, R., Makalowska, I., Baxevanis, A., Johnstone, B., Laxer, R. M., Zemel, L., Kim, C. A., Herd, J. K., Ihle, J., Williams, C., Johnson, M., Raman, V., Bahabri, S. A., Trent, J. M., and Warman, M. L. Mutations in a secreted proteoglycan, encoded by the gene *CACP*, cause camptodactyly-arthropathy-coxa vara-pericarditis syndrome indicating its essential role in tendons and joints. *Nat. Genet.* **23:** 319–322.

Martinez-Lavin, M., Buendia, A., Delgado, E., Reyes, P., Amigo, M.-C., Sabanes, J., Zghaib, A., Attie, F., and Salinas, L. (1983). A familial syndrome of pericarditis, arthritis, and camptodactyly. *N. Engl. J. Med.* **309:** 224–225.

Ochi, T., Iwase, R., Okabe, N., Fink, C. W., and Ono, K. (1983). The pathology of the involved tendons in patients with familial arthropathy and congenital camptodactyly. *Arthritis Rheum.* **26:** 896–900.

Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C., and Markham, A. F. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucl. Acids Res.* **18:** 2887–2890.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992).Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. USA* **89:** 8794–8797.

Smith, J. R., Freije, D., Carpten, J. D., Gronberg, H., Xu, J., Isaacs, S. D., Brownstein, M. J., Bova, G. S., Guo, H., Bujnovszky, P., Nusskern, D. R., Damber, J.-E.; Bergh, A., Emanuelsson, M., Kallioniemi, O. P., Walker-Daniels, J., Bailey-Wilson, J. E., Beaty, T. H., Meyers, D. A., Walsh, P. C., Collins, F. S., Trent, J. M., and Isaacs, W. B. (1996). Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. *Science* **274:** 1371–1374.

Soderlund, C., Longden, I., and Mott, R. (1997). FPC: A system for building contigs from restriction fingerprinted clones. *CABIOS* **13:** 523–535.

Solovyev, V. V., and Salamov, A. A. (1997). The Gene-Finder computer tools for analysis of human and model organisms genome sequences. *In* "Proceedings of the Fifth International Conference on Intelligent Systems for Molecular Biology, Halkidiki, Greece," pp. 294–302, AAAI Press, Menlo Park, CA.

Teh, B. T., Farnebo, F., Kristoffersson, U., Sundelin, B., Cardinal, J., Axelson, R., Yap, A., Epstein, M., Heath, H., III, Cameron, D., and Larsson, C. (1996). Autosomal dominant primary hyperparathyroidism and jaw tumor syndrome associated with renal hamartomas and cystic kidney disease: Linkage to 1q21–q32 and loss of the wild type allele in renal hamartomas. *J. Clin. Endocr. Metab.* **81:** 4204–4211.

Williamson, C., Cavaco, B. M., Jausch, A., Dixon, P. H., Forbes, S., Harding, B., Holtgreve-Grez, H., Schoell, B., Pereira, M. C., Font, A. P., Loureiro, M. M., Sobrinho, L. G., Santos, M. A., and Thakker, R. V. (1999). Mapping the gene causing hereditary primary hyperparathyroidism in a Portuguese kindred to chromosome 1q22–q31. *J. Bone Miner. Res.* **14:** 230–239.

Xu, J., Meyers, D., Freije, D., Isaacs, S., Wiley, K., Nusskern, D., Ewing, C., Wilkens, E., Piroska, B., Bova, G. S., Walsh, P., Isaacs, W., Schleutker, J., Matikainen, M., Tammela, T., Visakorpi, T., Kallioniemi, O. P., Ewing, C., Berry, R., Schaid, D., French, A., McDonnell, S., Schroeder, J., Blute, M., Thibodeau, S., Gronberg, H., Emanuelsson, M., Damber, J.-D., Anders, B., Jonsson, B.-A., Smith, J., Bailey-Wilson, J., Carpten, J., Stephan, D., Gillanders, E., Asmundson, I., Kainu, T., Freas-Lutz, D., Baffoe-Bonnie, A., Van Aucken, A., Sood, R., Collins, F., Brownstein, M., and Trent, J. M. (1998). Evidence for a prostate cancer susceptibility locus on the X chromosome. *Nat. Genet.* **20:** 175–179.

Zhang, J., and Madden, T. L. (1997). PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation. *Genome Res.* **7:** 649–656.

Zhang, M. Q. (1997). Identification of protein coding regions in the human genome based on quadratic discriminant analysis. *Proc. Natl. Acad. Sci. USA* **94:** 565–568.