Perfect Conserved Linkage Across the Entire Mouse Chromosome 10 Region Homologous to Human Chromosome 21

Tim Wiltshire,[1,2,5] Mathew Pletcher,[1,5] Susan E. Cole,[1,3] Melissa Villanueva,[1] Bruce Birren,[4] Jessica Lehoczky,[4] Ken Dewar,[4] and Roger H. Reeves[1,6]

[1]Department of Physiology, Johns Hopkins School of Medicine, Baltimore, Maryland 21205 USA; [4]Whitehead Institute/MIT Center for Genome Research, Cambridge, Massachusetts 02141 USA

The distal end of human Chromosome (HSA) 21 from *PDXK* to the telomere shows perfect conserved linkage with mouse Chromosome (MMU) 10. This region is bounded on the proximal side by a segment of homology to HSA22q11.2, and on the distal side by a region of homology with HSA19p13.1. A high-resolution PAC-based physical map is described that spans 2.8 Mb, including the entire 2.1 Mb from *Pdxk* to *Prmt2* corresponding to HSA21. Thirty-four expressed sequences are mapped, three of which were not mapped previously in any species and nine more that are mapped in mouse for the first time. These genes confirm and extend the conserved linkage between MMU10 and HSA21. The ordered PACs and dense STS map provide a clone resource for biological experiments, for rapid and accurate mapping, and for genomic sequencing. The new genes identified here may be involved in Down syndrome (DS) or in several genetic diseases that map to this conserved region of HSA21.

Trisomy 21 is the most frequent human aneuploidy at birth, occurring in 1 out of 700 live births (Hassold et al. 1996). For this reason, and because it is the smallest human autosome, analysis of HSA21 has played a leading role in the Human Genome Project (Antonarakis 1998). The telomeric region of HSA21q22.3 is associated with several phenotypes of Down syndrome (DS), including congenital cardiac defects (Korenberg et al. 1994). Several monogenic human disorders have been localized to this region, including the gene for Cystatin B, mutation of which is responsible for progressive myoclonus epilepsy of Unverricht and Lundborg type (Pennacchio et al. 1996; Pennacchio and Myers 1996); autoimmune polyendocrinopathy-candidasis ectodermal dystrophy type I (APECED), arising from mutations in the *AIRE* gene (Nagamine et al. 1997; The Finnish-German APECED Consortium 1997); and Bethlem benign congenital myopathy, which is correlated with mutations in the *COL6A1* and *COL6A2* genes (Jobsis et al. 1996). Several human disorders for which no causative genes have been found also map to this region, including bipolar affective disorder (Straub et al. 1994), Knobloch syndrome (Sertie et al. 1996), holoprosencephaly (Muenke et al. 1995), and two forms of nonsyndromic deafness (Bonne-Tamir et al. 1996; Veske et al. 1996). Despite the many diseases mapping to

HSA21, several lines of evidence suggest that this chromosome may have a lower gene density than other chromosomes, for example, the relative paucity of the 30,000 mapped UniGene sequences found on HSA21 (Antonarakis 1998).

HSA21 includes regions of conserved linkage with mouse Chromosomes 10, 16, and 17 (Mouse Genome Informatics, http://www.informatics.jax.org). A number of mouse models for DS, both multigenic and single genes have been developed to study the mechanisms that underlie one or more of the anomalies that occur in DS (Kola and Herzog 1998). Several models have been developed with dosage imbalance for extended regions of MMU16 that correspond to HSA21 (Reeves et al. 1995; Sago et al. 1998; Smith et al. 1995) but not for DS traits that arise from dosage imbalance in this region of MMU10.

We constructed a high-resolution physical map comprised of a contiguous set of PAC clones that span the region of MMU10 with homology to HSA19, HSA22, and the entire region of conserved linkage with HSA21 (Fig. 1). Thirty-four transcripts are mapped, including twelve that are localized for the first time on the MMU10 physical map. Twenty-four transcripts in the region of conserved synteny with HSA21 are found in the corresponding positions on the human and mouse chromosomes and are likely to represent a substantial portion of genes from this region. The PACs are an important resource for constructing both single gene and multi-gene mouse models of diseases in the MMU10/HSA21 conserved synteny region. A minimal

Present addresses: [2]Center for Neurobiology and Behavior, University of Pennsylvania, Philadelphia, Pennsylvania 19104 USA; [3]Lewis Thomas Laboratory, Princeton University, Princeton, New Jersey 08544 USA.
[5]These authors contributed equally to this work.
[6]Corresponding author.
E-MAIL rreeves@welchlink.welch.jhu.edu; FAX (410) 955-0461.
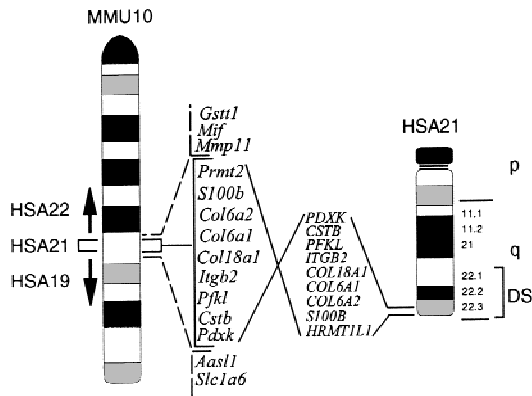
Figure 1 An interstitial segment of MMU10 shows perfect conserved linkage with HSA21. The HSA21-homologous segment is inverted with respect to the HSA21 telomere and centromere, but all genes mapped in both species are in the same relative positions. This conserved segment is flanked by regions of conserved synteny with HSA22 and HSA19. A subset of the 24 markers mapped to MMU10 and HSA21 is shown.

tiling path of these clones provides sequencing template for the entire region of conserved linkage.

RESULTS

Isolation of PAC Clones Forming a 2.8-Mb Physical Map

A YAC-based physical map for the HSA21/MMU10 region of conserved synteny was described previously (Cole et al. 1999). Six YACs that span a large part of the 2.1 Mb region were used as hybridization probes to screen the entire RPCI-21 PAC library (Fig. 2). Of these, YAC I2H8 was known to be chimeric and YAC 60D3 proved to be unstable. Assembly of the PAC/BAC contig demonstrated that two other YACs, 114H5 and 113G5, had substantial internal deletions. In both cases these deletions correspond to regions of instability in human YACs for the region around *Pfkl* and *Adarb1* (Lapenta et al. 1998).

Hybridizations with the six minimal-tiling path YACs identified a core set of 137 PACs. End sequences from selected clones and additional MMU10/HSA21 markers were used as probes on the PAC library to extend coverage. PACs were arrayed on nylon filters for screening by hybridization with the original YACs and additional markers and DNAs were prepared for PCR-based STS content analysis. Nine D10Mit markers mapped to this region were also mapped on the contig. PAC clones were underrepresented in the mouse RPCI-21 library for the region just distal to *Pfkl*. The mouse RPCI-22 BAC library was screened to provide additional depth of coverage for this region. The final contig spans 2.8 Mb and includes the entire HSA21 region of conserved linkage, anchored in flanking HSA19- and HSA22-homologous sequences (Fig. 2).

The sizes of all PAC inserts (average 135 kb) were determined on pulsed field gels following *Not*I digestion, and PAC overlaps and integrity of individual clones were confirmed by assessment of restriction endonuclease fingerprints and by PAC-to-PAC hybridizations on Southern blots. The final contig contains 156 clones, providing an average 7 × coverage across 2.8 Mb, that is, each marker hit 7 PACs on average. In all, 91 markers were mapped on the 2.8 Mb contig (1/31 kb). Thirty-four of these are genes or ESTs and 40 additional markers are derived from PAC end sequences (Table 1). The positions of all expressed sequences and 47 of the PAC and YAC ends (*D10Jhu* markers) are shown on 52 clones describing the minimal tiling path in Figure 2. The total STS content for all clones with all markers is available from http://oscar.med.jhu.edu/roger/roger.html.

Thirty-one percent (42) of the 137 PAC clones recovered from the initial screening with YAC probes could not be shown to be part of this region by STS content mapping and were considered false positives. This estimate may be high since some clones may have been real positives for YAC I2H8, which has a substantial chimeric region representing ~11% (~300 kb) of the aggregate probe length in the YAC minimal tiling path. Positive clones for this chimeric segment would not map to the MMU10 region.

The entire clone set was independently ordered by semi-automated restriction fingerprint analysis (Marra et al. 1997). Separate restriction fragment maps were built using *Hin*dIII and *Eco*RI digests. Clone order was consistent throughout the contig for maps made with either enzyme, and with the STS content map. However, the minimal tiling path assembled by fingerprinting included two apparent gaps in coverage, at *Adarb1* and *Mit242*, when compared to the STS content map (Fig. 2). Sequencing of clone 644M8 (AC007641) demonstrated that it overlaps with 364C8, as predicted by fingerprinting but not STS content analysis. If this should prove incorrect for the second apparent gap, multiple bridging PACs are available to complete the tiling path. Because the fingerprint based map revealed overlap between clones that did not necessarily share an STS, the minimal tiling path based on fingerprint analysis contained only 13 clones as opposed to 19 when STS content mapping was used alone. The 2.1 Mb aggregate insert size of these 13 clones is close to estimates of the size of the HSA21-homologous region based on YAC fragmentation mapping (Cole et al. 1999).

Identification of Chromosomal Boundary Regions

The 2.8 Mb contig extends 550 kb into the HSA22 conserved synteny region. The most distal known gene on HSA21 is *HRMT1L1* (mouse *Prmt2*) and this gene is present on the MMU10 minimal tiling path ~200 kb from *Mmp11*, whose human ortholog is on HSA22.
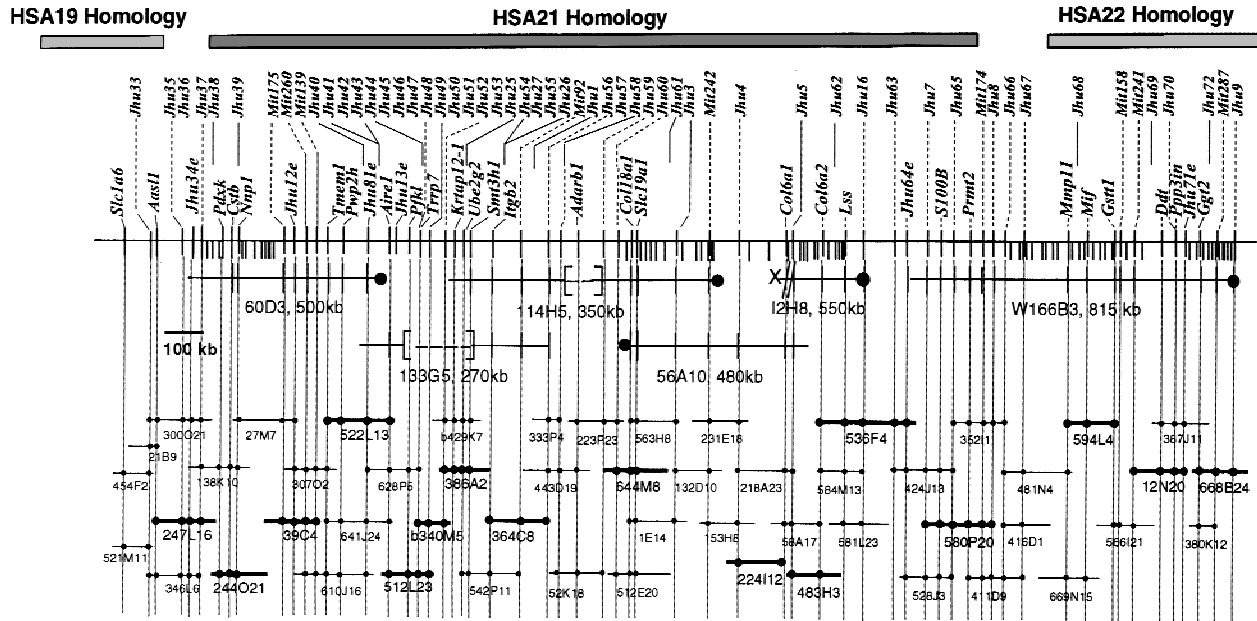
Figure 2 PAC-based STS content map of the region of MMU10 that shows conserved linkage with HSA19, HSA21, and HSA22. The contig spans 2.8 Mb and includes the entire HSA21 homologous region. Fifty-two of 156 clones are shown with addresses in the RPCI-21 or RPCI-22 libraries (BACs from the RPCI-22 library have the prefix „"b"). STS markers are shown above and transcripts adjacent to the chromosome. The prefix "D10" has been omitted for Jhu and Mit STS markers. Six YACs that were used as hybridization probes are shown: [—-] indicates a deleted region, **X**-// a chimeric end, and filled circles the YAC centromeric ends. YAC fragmentation events are indicated as short lines extending beneath the main marker line (Cole et al. 1999). Clones are drawn to scale. Presence of a marker is indicated with a filled circle (a subset of the STS markers are shown here). Clones of the minimal tiling path that are currently being sequenced are indicated by bold lines. Sequence is available for clones 340M5 (GenBank accession no. AC007433), 644M8 (accession no. AC007641), 536F4 (accession no. AC007937), and 580P20 (accession no. AC006507).

This narrows the proximal boundary of HSA21/22 conserved synteny which was localized previously to three overlapping PACs (Cole et al. 1998) that are being sequenced currently at the University of Oklahoma Genome Center (see accession nos. AC006507, AC005302, and AC005818). This contig also establishes *Pdxk* distal to *Cstb* on MMU10 (Fig. 2) and locates the HSA21/19 boundary region between pyridoxal kinase, *Pdxk* (*PDXK* is on HSA21), and *Aasl1*, a mammalian homolog of an acetolactate synthase gene from *Escherichia coli* found on MMU10 and HSA19. Mouse EST 851439 (*D10Jhu34e*) is located between these genes but does not have a mapped human ortholog. The contig extends approximately 250 kb into the HSA19 region. Minimal tiling path clones selected for sequencing include the entire HSA21/MMU10 region and extend across the borders of MMU10 homology with HSA19 and HSA22. Sequence for a boundary of chromosome fusion in mammalian evolution has not been reported.

Gene Content

In the course of contig construction, 57 kb of MMU10 sequence was derived from 74 PAC ends. These sequences were masked and used to search EST and nonredundant databases by BLAST search. Thirty percent of the sequence was repetitive based on Repeat-Masker. The unique end sequence recognized 11 genes or ESTs, 8 of which were not localized previously in mouse (Table 2). Four additional genes, *Pwp2h*, *Pdxk*, *Slc1a6*, and *Ube2g2*, were mapped for the first time in mouse using cognate cDNAs to the human genes whose positions on HSA21 or HSA19 were known.

A total of 34 transcripts were mapped on the mouse physical map (Table 2). Thirty-one of these that were known or found to have human orthologs were used to align the comparative mouse/human map. All transcripts map to corresponding positions on HSA21 or HSA22. Three mouse ESTs identified by BLAST search using PAC end sequences do not recognize human ESTs in GenBank. The 31 transcribed sequences with human orthologs provide one expressed sequence per 90 kb on average.

Five HSA21 orthologs were mapped to high resolution in this interval for the first time. Periodic tryptophan protein 2 yeast homolog (*Pwp2h*) (Lafreniere et al. 1996) and *Ube2g2*, ubiquitin-conjugating enzyme e2g 2 (Katsanis and Fisher 1998) are homologs of *S. cerevisiae* genes mapped to the MMU10 recombinational map in the XREFdb project as *D10Xrf337* and *D10Xrf369/453*, respectively (Bassett et al. 1997). Pyridoxal kinase, *Pdxk*, (Hanna et al. 1997); transient re-

Table 1. PCR Primers and Reaction Conditions

Symbol	Primer name	Product size (bp)	Primer sequence		Reaction conditions[a]	Accession no.
D10Jhu25	100K2.T7	104	F-5'TACTTCAGGGTGTGTGGCAG	R-5'ACAGGGTGAGGAGGGACC	Prog 45	AF167489
D10Jhu26	100K2.SP6	480	F-5'CCTAGGTGAAGCGTCCAGAG	R-5'GGGGACTTAAAAAGCTTGCC	Prog 30	AF167488
D10Jhu27	386A2.T7	223	F-5'ACTCAACCTGGGCCTCAAC	R-5'TCACCTAGCGTGAGTGCAAG	Prog 30	AF167518
D10Jhu33	599G22.SP6	125	F-5'TCAGCGGGAAATGACATACA	R-5'TGGGGATCTTCAGTGACACA	Prog 42	AF167508
D10Jhu34e	300O21.T7	186	F-5'TGAGTGCCTCAGAAGAGGGT	R-5'AGAACTCTGGTGCGGTCAGT	Prog 45	AF167501
D10Jhu35	247L16.T7	256	F-5'CCAAGACTGTGTGCCAGAAC	R-5'GCAGTTAGAACTTCCGTCCG	Prog 42	AF167498
D10Jhu36	331J21.SP6	419	F-5'GGGAGTTGGGGAGATACTGG	R-5'CATGCTTGCTTCTGCTTGAC	Prog 30	AF164436
D10Jhu37	235F4.SP6	146	F-5'CGCTAAAGGTGGATGCTTTC	R-5'CTGTCCGGGAACACATGAC	Prog 10 at 57	AF167496
D10Jhu38	247L16.SP6	287	F-5'TTCCTGTCCCATCTCCAGAC	R-5'TGATTCCAGTCAAGCCTTCC	Prog 42	AF167497
D10Jhu39	331J21.T7	314	F-5'ACCCATATGCAGAGACCCAG	R-5'CCTCATCTGCCAACAATCAG	Prog 30	AF167502
D10Jhu40	628P5.SP6	260	F-5'CCGCCAAGTGCTAGATTTTC	R-5'GTAGCTAAGCGGCTCCTGTG	Prog 1 at 60	AF167509
D10Jhu41	64013.SP6	391	F-5'CAGCTAACTAAAGCCACGGG	R-5'TCACTGCTACCAATGCCAAG	Prog 30	AF167528
D10Jhu42	166J15.T7	238	F-5'ACATTGCCTGGTCCTCTCTG	R-5'AGCCACATGGCTCCTATCAG	Prog 10 at 57	AF167495
D10Jhu43	578K13.T7	198	F-5'TTTGCTGGGGGTACAAGTTC	R-5'CTAAACAGATGCCCCTGGAC	Prog 10	AF167507
D10Jhu44	64012.T7	252	F-5'CTCGACCTTTTCTGTCCTCG	R-5'CTTTCCGTATACGCGATTCC	Prog 10	AF167529
D10Jhu45	628P5.T7	140	F-5'CCAGGAGGAGGAAGCTGTAG	R-5'AATACCCCCTTCCACCACTC	Prog 10 at 60	AF167510
D10Jhu46	166J15.SP6	139	F-5'CTCTACCGCAACTCTCAGCC	R-5'TTGGGGTTTAAGTCTGCCAC	Prog 10 at 57	AF167494
D10Jhu47	340M5.SP6	348	F-5'ACTGGGACCACCAACTGAAG	R-5'CCAGAAAGAGCCTCACGAAC	Prog 45	AF167513
D10Jhu48	578K13.SP6	281	F-5'AGATGCATACTGCATGGGTG	R-5'AGTCCCTGCCGATGACATAC	Prog 10	AF167506
D10Jhu49	112I15.T7	266	F-5'AATGCCAATCAAACCTGAGC	R-5'TTTGCCCCTTCAATTACAGC	Prog 1	AF167490
D10Jhu50	340M5.T7	331	F-5'GTAGGGGTGCCAGTGTCATT	R-5'CCCTTTCCCCCACAATCTAT	Prog 10	AF167514
D10Jhu51	335G5.SP6	266	F-5'CAGAGAGAATTCCCCCATCA	R-5'AACTCCCAACCAACACCAAG	Prog 45	AF167512
D10Jhu52	429K7.SP6	185	F-5'AGTCCTGGTCACCAACAAGG	R-5'AGAAATGGGATGCTGTGACC	Prog 45	AF167523
D10Jhu53	118N22.SP6	209	F-5'GTAGGGGAGACCTTTGGGAC	R-5'AGACGATGTCCGGTGGTAAG	Prog 45	AF167491
D10Jhu54	118N22.T7	258	F-5'ACGTAGGTGCCGAGTTCATC	R-5'GAAACCAGGCTGACTGAAGGC	Prog 45	AF167492
D10Jhu55	364C8.SP6	103	F-5'CCAGTTCTCCAAACCTCCAG	R-5'CAGTCAGGCTCACAGAGCAG	Prog 30	AF167516
D10Jhu56	364C8.T7	230	F-5'CCGTGCTAAGAACTACCATGC	R-5'CCCAAACACACACCTCACAC	Prog 1 at 60	AF167517
D10Jhu57	333P4.T7	258	F-5'TTGTCATCTGCCTCTACCCC	R-5'AAGTGCGTAAGCCAGAAAGC	Prog 45	AF167511
D10Jhu58	52K18.SP6	187	F-5'TGCCTATACCACTGTGCCTG	R-5'AGCCGTGTTTTTCCTGTGTC	Prog 45	AF167503
D10Jhu59	52K18.T7	429	F-5'AGCCTGCTTCTTGGTGTCTC	R-5'CATGCTATGTCTGTCCGTGG	Prog 45	AF167504
D10Jhu60	644M8.T7	275	F-5'CTAAGCCTGCATCTTGGGTC	R-5'TCATCAGACCTGGAGCAGTG	Prog 30	AF167531
D10Jhu61	644M8.SP6	213	F-5'TAATGCTCTTGAACCCCTCG	R-5'CTCACAGGGGTCCTGAAAAG	Prog 30	AF167530
D10Jhu63	424J13.T7	381	F-5'AGAACAGAGAGCCAACAGCC	R-5'CTCGAAGACAGTCGTGATGC	Prog 45	AF167521
D10Jhu64e	536F4.T7	222	F-5'CGGCATACAATGTCAAGTGG	R-5'TTGTTCCCTTGATCCTCCTG	Prog 10	AF167515
D10Jhu65	424J13.SP6	392	F-5'ATCGGTGGCCGTATAGTCGT	R-5'CACTCTTGGGAAGGAAGCTG	Prog 30	AF167522
D10Jhu66	566H1.T7	288	F-5'CATGATGGTTGCTTGTCCTG	R-5'TGGGGTCTGACTGAAGAACC	Prog 30	AF167505
D10Jhu67	481N4.SP6	441	F-5'AGGAAGGGAAGGAGAAGCAG	R-5'CATCCTGTAGAAAGCGAGCC	Prog 10	AF167525
D10Jhu68	481N4.T7	207	F-5'GTTGCTGACTCTCCTCCCTG	R-5'TAGAGAAGGTTCTGCCTGCC	Prog 45	AF167526
D10Jhu69	12N23.SP6	154	F-5'GGTTCTGCCTGGGAATCTC	R-5'TCCAACATCCTTTTTCTCGG	Prog 10 at 57	AF167493
D10Jhu70	387J11.SP6	297	F-5'GGAAATCTGGAGAGACGCAG	R-5'TGAATGTGGCTGAGTCGAAG	Prog 10 at 57	AF167519
D10Jhu71e	12N23.T7	233	F-5'CTTTTCCAGCCCTCACTCAG	R-5'AAAGGCAACATCCCACTCTG	Prog 10 at 57	AF167537
D10Jhu72	387J11.T7	228	F-5'GGGAGGAAGGAGGCAGATAG	R-5'CGGGACTTGCTGTTTTAAGC	Prog 10 at 57	AF167520
D10Jhu78	455H18.T7	183	F-5'CCTCAGGCTGCTCTGACTCT	R-5'TGTCTGGCATCAACAGCTTC	Prog 45	AF167535
D10Jhu79	58L3.T7	107	F-5'TCCATCTTTGAGACATGGGAG	R-5'CATAGACTGGGCCTTAGGGA	Prog 42	AF167536

[a]Prog. 1 (94°C × 30"/55°C × 30"/72°C × 1') 35 cycles. Prog. 10: As Prog. 1 but 45 cycles. Prog. 30: As Prog. 1 but 57°C annealing. Prog. 42: As Prog. 1 but 58°C annealing and 45 cycles. Prog. 45: Touchdown 72°C to 58°C in 2°C steps–5 cycles each.

ceptor potential-related channel, *Trrp 7* (human *TRPC7*) (Kudoh et al. 1997); and the mouse ortholog of *D21S2048E* (*D10Jhu81e*) (Scott et al. 1997) are also localized on the physical map. The human *D21S2048E* gene has a number of aliases including *HES1,* homolog of (zebrafish protein) ES1 (Table 2). The symbol *Hes1* was assigned previously to a different mouse gene, a homolog of *Drosophila* hairy enhancer of split, which

Table 2. Transcripts Defining MMU10 Conserved Linkages with HSA19, 21, and 22

Mouse locus name	Human locus name and aliases	Genbank accession number	Gene name	Human location
Slc1a6[c]	SLC1A6, EAAT4	D83262[a]	Solute carrier family 1, member 6	19p13.1
Aasl1[c]	AASSL1, 209L8	AI156588[b]	Human homolog of acetolactate synthetase like gene	19p13.1
D10Jhu34e[c]	No known human orthologs	AI613583[b]	Mouse EST 851439	(19 or 21)[d]
Pdxk[c]	PDXK	AI529482[a]	Pyridoxal kinase	21q22.3
Cstb	CSTB, STFB		Cystatin B	21q22.3
Nnp1	NNP-1, D21S2056E		Novel nuclear protein 1	21q22.3
D10Jhu12e	Homology to human genomic sequence from 21q22.3 region		Mouse EST 515094	21q22.3
Tmem1	TMEM1		Transmembrane protein 1	21q22.3
Pwp2h[c]	PWP2H	H50886[a]	Yeast periodic tryptophan protein 2 homolog	21q22.3
D10Jhu81e[c]	D21S2048E, GT335, HES1, hES1, KNP-Ia, KNP-Ib	AA510875[a]	Homolog of ES1	21q22.3
Aire1	AIRE, APECED, APS1, PGA I		Autoimmune polyendocrinopathy syndrome, type I	21q22.3
D10Jhu13e	C21ORF2		mRNA mapping to APECED disease region (HSA 21q22.3)	21q22.3
Pfkl	PFKL, PFK		Phosphofructokinase, liver type	21q22.3
Trrp7[c]	TRPC7	NM003307[b]	Transient receptor potential-related channel 7	21q22.3
Krtap12-1	KRTAP12-1		High-cysteine keratin associated protein	21q22.3
Ube2g2[c]	UBE2G2, UBC7	AA759396[a]	Ubiquitin-conjugating enzyme E2G 2	21q22.3
Smt3h1	SMT3H1, SMT3A, SMT3		Supressor of mif two, 3 homolog 1	21q22.3
Itgb2	ITGB2, CD18		Integrin, beta-2	21q22.3
Adarb1	ADARB1, RED1		Adenosine deaminase, RNA-specific, B1	21q22.3
Col18a1	COL18A1		Collagen, type XVIII, alpha-1	21q22.3
Slc19a1	SLC19A1, FOLT, RFC1, IFC1		Solute carrier family 19, member 1	21q22.3
Col6a1	COL6A1		Collagen, type VI, alpha-1	21q22.3
Col6a2	COL6A2		Collagen, type VI, alpha-2	21q22.3
Lss	LSS, OSC		Lanosterol synthase	21q22.3
D10Jhu64e[c]	No known human orthologs	AA727480[b]	Mouse EST 1210378	(21q22.3)[d]
S100b	S100B		S100 calcium-binding protein, beta	21q22.3
Prmt2	PRMT2, HRMT1L1		hnRNP methyltransferase-1-like 1	21q22.3
Mmp11	MMP11, STMY3		matrix metalloproteinase 11	22q11.2
Mif	MIF		Macrophage migration inhibitory factor	22q11.2
Gstt1	GSTT1		Glutathione S-transferase, theta 1	22q11.2
Ddt	DDT		D-dopachrome tautomerase	22q11.2
Ppp3in[c]	PPP3IN, KIAA0330	AB002328[b]	Protein phosphatase 3 inhibitor	22q11–12
D10Jhu71e[c]	No known human orthologs	AI615712[b]	Mouse EST 974977	(22q12)[d]
Ggt2[c]	GGT2	U30509[a]	Gamma-glutamyl transpeptidase	22q12

Human map positions are from OMIM, GDB, and this analysis. Genes/ESTs are listed in the order that they appear on MMU10, telomeric to centromeric.
[a]Accession number of sequence that was used to map the gene.
[b]accession number of the EST/gene that was discovered from PAC end-sequence homology searches.
[c]transcripts mapped for the first time in mouse.
[d]predicted location in human.

maps to a region of MMU16 that shows conserved synteny with HSA3 (Feder et al. 1994). Mouse *Hes1* is not related to human *D21S2048E/HES1*. *D10Jhu64e* hits EST 1210378 which shows very weak similarity (p = 5×10^{-4} using BLASTx) to the putative mouse CAMP protein (AF119384).

Two genes were mapped in the HSA19 conserved synteny region. *Aasl1* (human *AASL1*) was identified by PAC-end sequence 21B9.SP6 which perfectly matches mouse EST 1493949 (AI156588). This EST in turn shows strong conservation (p = 1.1×10^{-55})

with a human gene related to the B isozyme of the large catalytic subunit of the *E. coli* acetolactate synthase gene (AHAS-I) (Joutel et al. 1996), and to two cosmids from HSA19p13.1 (accession nos. AC003956 and AC004794). *Slc1a6*, solute carrier family 1, member 6, was mapped distal to *Aasl1* on MMU10. Its human ortholog (*SLC1A6*) is also found in genomic sequence from HSA19p13.1 (accession no. AC004659).

Two genes are also newly mapped in the regions of conserved synteny with HSA22. These are gamma-glutamyl transpeptidase 2, *Ggt2* (Figlewicz et al. 1993)

and calcineurin inhibitor, *Ppp3in* (human *PPP3IN*) (Lai et al. 1998). In all, six genes and an orphan EST were ordered in this conserved segment (Fig. 2). Across the entire contig, the positions of all genes that had identifiable human orthologs mapped to corresponding positions in both species. Eighteen of the 34 transcribed sequences were mapped previously on the YAC contig that provided the basis for PAC contig assembly (Cole et al. 1999).

DISCUSSION

A 2.8 Mb physical map was prepared using a combination of strategies. YAC-based hybridization screening of arrayed PACs and BACs identified a set of clones whose relative positions and overlaps were determined by STS-content mapping. End-sequence analysis of 38 clones provided additional STS markers to refine the contig. Fingerprinting methods were used to confirm clone order and to identify a tiling path of minimally overlapping clones. Thirteen clones spanning the 2.1-Mb MMU10/HSA21 conserved linkage region are currently being sequenced (Fig. 2).

The use of both yeast and bacterial-based clones generally overcomes difficulties in cloning regions that are unstable in one or the other of these disparate hosts. However, a region just distal to *Pfkl* appears to be unstable in both. We observed previously that YACs derived from this region of mouse and human are unstable (Cole et al. 1999). The only verified chimeric clone was a BAC that came from this region. Repeated probing of the complete RPCI-21 PAC library and subsequent probing of segment 2 of the RPCI-22 BAC library with several markers from this region, identified only five clones even though the combined libraries provide a total of 17.5 × genomic coverage. To date there is little information to indicate whether bacterial-based clones will provide complete genome coverage. When sequence data is available, it will be of interest to see if there is a common explanation for the scarcity of PAC clones and the instability of human and mouse derived YAC clones in this region.

Chromosomes are rearranged when compared among species, and rodents have undergone an unusually high number of genomic rearrangements per unit of evolutionary time (Graves 1996). Not surprisingly, the degree of conservation between mouse and human chromosomes depends on the scale and resolution of the analysis, and varies for different genome regions. Oakey and colleagues (Oakey et al. 1991) constructed long-range restriction maps spanning more than 11 Mb of homologous regions of MMU1 and HSA1, showing that gene content, order, and spacing were nearly identical in human and mouse throughout the length of this region. Others have shown a high degree of conservation for 2 Mb segments of HSA1/MMU3 (Kingsmore et al. 1989), and for human and mouse X

chromosomes (Cooper et al. 1993). A high-resolution YAC-based fragmentation map of distal MMU16 shows perfect conservation of gene order across 4.5 Mb of conserved linkage with HSA21, with conserved gene spacing for about two thirds of this length (Cabin et al. 1998). This MMU10 contig also shows perfect conservation of gene content and order when compared to the orthologous region of HSA21, with average conserved marker spacing of 90 kb. The availability of detailed comparative maps has initiated considerable interchange of genetic, genomic, and functional information between the mouse and human genomes (Carver and Stubbs 1997).

This region of MMU10 is not well-resolved on recombinational maps. Recombinational maps correlate with physical maps in order, but not relative spacing of genes because the frequency of recombination is a biological phenomenon only partly based on physical distance. The length of these maps is affected by gender and by the strain and species of mice involved in crosses (Reeves et al. 1991; Reeves et al. 1990). A set of intersubspecific crosses locates several genes from *Pfkl* to *S100b* in a 1.5 cM interval (Cole et al. 1999), whereas these genes are nonrecombinant on the interspecific JAX BSS cross (Rowe et al. 1994). Recombination is suppressed across the HSA19, HSA21, and HSA22 region of MMU10 in the BSS cross as evidenced by more than 30 nonrecombinant genes and ESTs at one locus, which we estimate to represent more than 8 Mb (~4–5 cM) based on mouse and human physical maps. The majority of these are orthologs of genes on HSA19. All six genes and ESTs mapped on BSS that have known HSA21 orthologs are included and ordered in the contig presented here. Three of these recognize two genes and were mapped initially in the XREFdb project based on their conservation with yeast genes or ORFs (Bassett et al. 1997).

Availability of both mouse and human sequence will provide a powerful tool for identifying genes, regulatory sequences, and conserved regions whose functions are not currently understood. For the MMU10/HSA21 conserved linkage, sequencing efforts for both species are ongoing, and mouse will likely provide the first mammalian genomic sequence of portions of this region. Mouse sequence provides some specific advantages in gene discovery. More than 15,000 mouse EST clusters representing genes expressed at developmental stages from conception to gastrulation have been developed, and nearly half have no human or mouse counterparts (Ko et al. 1998). Mouse sequence will provide a more conserved query set than human genomic sequence for these and other ESTs representing tissues and developmental stages that are inaccessible in human beings.

One of the major goals of research in DS is to correlate the dosage imbalance of specific genes on HSA21

with specific DS phenotypes. The perfect conservation of markers on these high resolution human and mouse physical maps indicates that mice at dosage imbalance for this region would provide an accurate model for the corresponding genetic insult in DS. The robust clone set and imminent comparative sequence for this region, provide the resources necessary to advance studies of this complex genetic disorder.

METHODS

PAC and BAC Library Screening

High-density plugs of selected YACs were prepared according to standard methods (Schwartz and Cantor 1984), and plugs were loaded onto an 8 cm wide preparative well of a 1% low-melt agarose CHEF gel. The conditions of electrophoresis were varied according to the size of the YAC being isolated using the CHEF-Mapper (Bio-Rad Laboratories, Hercules, CA) auto-algorithm. A segment of the gel was cut off and Southern blotted onto nylon, and hybridized with [^{32}P]dCTP labeled mouse Cot1 DNA to accurately identify the position of the YAC band. The YAC was excised from the remaining section of the gel and ~400 ng of isolated YAC DNA was labeled in agarose with [^{32}P]dCTP by random primer extension (Feinberg and Vogelstein 1983). Unincorporated nucleotides were removed from the labeling reaction using Sephadex NICK spin columns (Pharmacia, Piscataway, NJ.). Prior to the hybridization of labeled YAC DNA to library filters, all YAC probes were prehybridized with Mouse Cot1 DNA and $(CA)_{10}/(GT)_{10}$ oligonucleotides as suppressive blocking for repetitive DNA. The full protocol for labeling and hybridizations using YAC probes is available at http://physiology.med.jhu.edu/roger/protocols/prot1.html.

The mouse PAC (RPCI-21) and BAC (RPCI-22) libraries (Roswell Park Cancer Institute, Buffalo, NY) were screened by hybridization. Initial screenings used YACs as probes. In subsequent screenings the probes used were primarily PCR products generated from PAC end-sequence data (Table 1), cDNAs, and ESTs, also labeled by random priming methods. For smaller PCR products, priming was accomplished using one or both of the specific PCR primers instead of random hexamers.

Clone Analysis and Contig Construction

PAC DNA was prepared from 5 ml overnight cultures using a modified alkaline lysis procedure. Three milliliters of culture and 300 µl of each of P1, P2, and P3 (Qiagen "midi-prep" solutions) were used for alkaline lysis stages and 800 µl of isopropanol was used to precipitate DNA from the supernatant. After centrifugation and a 70% ethanol wash, the DNA pellet was resuspended in 40 µl of H_2O. PAC DNA for end-sequencing was prepared from 200 ml overnight cultures according to the modified protocol for BACs using P100 midi-prep columns (Qiagen Inc., Valencia, CA). The yield of DNA using this method ranged from 20–80 µg.

*Not*I digests of mini-prep PAC DNA were used to determine clone insert size. Twenty-microliter reactions containing 5 µl of DNA were digested with 5 units of *Not*I enzyme for 2 hr and subsequently run on a pulsed field gel (16 hr run time, switch time 5–15 sec, 6.0 V/cm, 14°C, 1% agarose in 0.5× TBE buffer). *Eco*RI and *Hin*dIII digests of mini-prep DNA were used to fingerprint clones (Marra et al. 1997). Clone ordering was performed manually using FPC (http://www.sanger.ac.uk).

PAC end-sequencing reactions included 2–3 µg of PAC DNA, 16 µl of Big Dye Terminator Mix (PE-ABI, Foster City, CA), 15–30 pmoles of T7 or SP6 primers, and 1 µl of Thermofidelase I (Fidelity Systems Inc., Gaithersburg, MD) in a 40-µl reaction volume. Cycling conditions were: 95°C for 5 min, (95°C for 30 sec, 55°C for 20 sec, 60°C for 4 min) for 45 cycles. Reaction products were purified through G-50 spin columns and run on ABI 377 sequencers. PAC end-sequences were masked using RepeatMasker (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker) and analyzed for sequence similarities using the BLAST network (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast?Jform = 0). Primer 3.0 (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi) was used to select PCR primers (Table 1). PCR was performed with diluted mini-prep PAC DNA in 15-µl reactions consisting of 1× buffer (20 mM Tris-HCl at pH 8.3, 50 mM KCl, and 2.5 mM $MgCl_2$), 0.2 mM each dNTP, 1 µM each of STS primer, and 0.5 units of *Taq* polymerase (Roche, Indianapolis, IN.) with the following conditions: 94°C for 4 min, 94°C for 30 sec, annealing for 30 sec (temperatures listed in Table 1), 72°C for 30 sec, for 35 cycles.

REFERENCES

Antonarakis, S.E. 1998. 10 years of Genomics, Chromosome 21, and Down syndrome. *Genomics* **51:** 1–16.
Bassett, D.E., Jr., M.S. Boguski, F. Spencer, R.H. Reeves, S. Kim, T. Weaver, and P. Hieter. 1997. Genome cross-referencing and

XREFdb: Implications for the identification and analysis of genes mutated in human disease. *Nat. Genet.* **15:** 339–344.

Bonne-Tamir, B., A.L. DeStefano, C.E. Briggs, R. Adair, B. Franklyn, S. Weiss, M. Korostishevsky, M. Frydman, C.T. Baldwin, and L.A. Farrer. 1996. Linkage of congenital recessive deafness (gene DFNB10) to chromosome 21q22.3. *Am. J. Hum. Genet.* **58:** 1245–1249.

Cabin, D.E., J.W. McKee-Johnson, L.E. Matesic, T. Wiltshire, E.E. Rue, A.E. Mjaatvedt, Y.K. Huo, J.R. Korenberg, and R.H. Reeves. 1998. Physical and comparative mapping of distal mouse chromosome 16. *Genome Res.* **8:** 940–950.

Carver, E.A. and L. Stubbs. 1997. Zooming in on the human-mouse comparative map: Genome conservation re-examined on a high-resolution scale. *Genome Res.* **7:** 1123–1137.

Cole, S.E., T. Wiltshire, and R.H. Reeves. 1998. Physical mapping of the evolutionary boundary between human Chromosomes 21 and 22 on mouse Chromosome 10. *Genomics* **50:** 109–111.

Cole, S.E., T. Wiltshire, E.E. Rue, D. Morrow, P. Hieter, C. Brahe, E.M. Fisher, N. Katsanis, and R.H. Reeves. 1999. High-resolution comparative physical mapping of mouse Chromosome 10 in the region of homology with human Chromosome 21. *Mamm. Genome* **10:** 229–234.

Cooper, P., J.T. Keer, V.M. McCabe, R.M. Hamvas, S.D. Brown, S. Rastan, and N. Brockdorff. 1993. High-resolution comparative mapping of the proximal region of the mouse X chromosome. *Genomics* **15:** 570–575.

Feder, J.N., L. Li, L.Y. Jan, and Y.N. Jan. 1994. Genomic cloning and chromosomal localization of HRY, the human homolog to the *Drosophila* segmentation gene, hairy. *Genomics* **20:** 56–61.

Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* **132:** 6–13.

Figlewicz, D.A., O. Delattre, G. Guellaen, A. Krizus, G. Thomas, J. Zucman, and G.A. Rouleau. 1993. Mapping of human gamma-glutamyl transpeptidase genes on Chromosome 22 and other human autosomes. *Genomics* **17:** 299–305.

Graves, J.A. 1996. Mammals that break the rules: Genetics of marsupials and monotremes. *Annu. Rev. Genet.* **30:** 233–260.

Hanna, M.C., A.J. Turner, and E.F. Kirkness. 1997. Human pyridoxal kinase. cDNA cloning, expression, and modulation by ligands of the benzodiazepine receptor. *J. Biol. Chem.* **272:** 10756–10760.

Hassold, T., M. Abruzzo, K. Adkins, D. Griffin, M. Merrill, E. Millie, D. Saker, J. Shen, and M. Zaragoza. 1996. Human aneuploidy: Incidence, origin, and etiology. *Environ. Mol. Mutagen.* **28:** 167–175.

Jobsis, G.J., H. Keizers, J.P. Vreijling, M. de Visser, M.C. Speer, R.A. Wolterman, F. Baas, and P.A. Bolhuis. 1996. Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. *Nat. Genet.* **14:** 113–115.

Joutel, A., A. Ducros, S. Alamowitch, C. Cruaud, V. Domenga, E. Marechal, K. Vahedi, H. Chabriat, M.G. Bousser, and E. Tournier-Lasserve. 1996. A human homolog of bacterial acetolactate synthase genes maps within the CADASIL critical region. *Genomics* **38:** 192–198.

Katsanis, N. and E.M. Fisher. 1998. Identification, expression, and chromosomal localization of ubiquitin conjugating enzyme 7 (UBE2G2), a human homologue of the *Saccharomyces cerevisiae ubc7* gene. *Genomics* **51:** 128–131.

Kingsmore, S.F., J. Snoddy, D. Choubey, P. Lengyel, and M.F. Seldin. 1989. Physical mapping of a family of interferon-activated genes, serum amyloid P-component, and alpha-spectrin on mouse chromosome 1. *Immunogenetics* **30:** 169–174.

Ko, M.S. , T.A. Threat, X. Wang, J.H. Horton, Y. Cui, X. Wang, E. Pryor, J. Paris, J. Wells-Smith, J.R. Kitchen et al. 1998. Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome. *Hum. Mol. Genet.* **7:** 1967–1978.

Kola, I. and P.J. Herzog. 1998. Down syndrome and mouse models. *Curr. Opin. Gen. Dev.* **8:** 316–321.

Korenberg, J.R., X.N. Chen, R. Schipper, Z. Sun, R. Gonsky, S. Gerwehr, N. Carpenter, C. Daumer, P. Dignan, C. Disteche et al. 1994. Down syndrome phenotypes: The consequences of chromosomal imbalance. *Proc. Natl. Acad. Sci.* **91:** 4997–5001.

Kudoh, J., K. Nagamine, S. Asakawa, I. Abe, K. Kawasaki, H. Maeda, S. Tsujimoto, S. Minoshima, F. Ito, and N. Shimizu. 1997. Localization of 16 exons to a 450-kb region involved in the autoimmune polyglandular disease type I (APECED) on human chromosome 21q22.3. *DNA Res.* **4:** 45–52.

Lafreniere, R.G., D.L. Rochefort, N. Chretien, C.E. Neville, R.G. Korneluk, L. Zuo, Y. Wei, J. Lichter, and G.A. Rouleau. 1996. Isolation and genomic structure of a human homolog of the yeast periodic tryptophan protein 2 (PWP2) gene mapping to 21q22.3. *Genome Res.* **6:** 1216–1226.

Lai, M.M., P.E. Burnett, H. Wolosker, S. Blackshaw, and S.H. Snyder. 1998. Cain, a novel physiologic protein inhibitor of calcineurin. *J. Biol. Chem.* **273:** 18325–18331.

Lapenta, V., V. Sossi, P. Gosset, C. Vayssettes, T. Vitali, N. Rabatel, F. Tassone, J.L. Blouin, H.S. Scott, S.E. Antonarakis et al. 1998. Construction of a 2.5-Mb integrated physical and gene map of distal 21q22.3. *Genomics* **49:** 1–13.

Marra, M.A., T.A. Kucaba, N.L. Dietrich, E.D. Green, B. Brownstein, R.K. Wilson, K.M. McDonald, L.W. Hillier, J.D. McPherson, and R.H. Waterston. 1997. High throughput fingerprint analysis of large-insert clones. *Genome Res.* **7:** 1072–1084.

Muenke, M., L.J. Bone, H.F. Mitchell, I. Hart, K. Walton, K. Hall-Johnson, E.F. Ippel, J. Dietz-Band, K. Kvaloy, and C.M. Fan. 1995. Physical mapping of the holoprosencephaly critical region in 21q22.3, exclusion of SIM2 as a candidate gene for holoprosencephaly, and mapping of SIM2 to a region of Chromosome 21 important for Down syndrome. *Am. J. Hum. Genet.* **57:** 1074–1079.

Nagamine, K., P. Peterson, H.S. Scott, J. Kudoh, S. Minoshima, M. Heino, K.J. Krohn, M.D. Lalioti, P.E. Mullis, S.E. Antonarakis et al. 1997. Positional cloning of the APECED gene. *Nat. Genet.* **17:** 393–398.

Oakey, R.J., M.G. Caron, R.J. Lefkowitz, and M.F. Seldin. 1991. Genomic organization of adrenergic and serotonin receptors in the mouse: Linkage mapping of sequence-related genes provides a method for examining mammalian chromosome evolution. *Genomics* **10:** 338–344.

Pennacchio, L.A., A.E. Lehesjoki, N.E. Stone, V.L. Willour, K. Virtaneva, J. Miao, E. D'Amato, L. Ramirez, M. Faham, M. Koskiniemi et al. 1996. Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy. *Science* **271:** 1731–1734.

Pennacchio, L.A. and R.M. Myers. 1996. Isolation and characterization of the mouse cystatin B gene. *Genome Res.* **6:** 1103–1109.

Reeves, R.H., M.R. Crowley, W.S. Moseley, and M.F. Seldin. 1991. Comparison of interspecific to intersubspecific backcrosses demonstrates species and sex differences in recombination frequency on mouse chromosome 16. *Mamm. Genome* **1:** 158–164.

Reeves, R.H., M.R. Crowley, B.F. O'Hara, and J.D. Gearhart. 1990. Sex, strain, and species differences affect recombination across an evolutionarily conserved segment of mouse Chromosome 16. *Genomics* **8:** 141–148.

Reeves, R.H., N.G. Irving, T.H. Moran, A. Wohn, C. Kitt, S.S. Sisodia, C. Schmidt, R.T. Bronson, and M.T. Davisson. 1995. A mouse model for Down syndrome exhibits learning and behaviour deficits. *Nat. Genet.* **11:** 177–184.

Rowe, L.B., J.H. Nadeau, R. Turner, W.N. Frankel, V.A. Letts, J.T. Eppig, M.S.H. Ko, S.J. Thurston, and E.H. Birkenmeier. 1994. Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. *Mamm. Genome* **5:** 253–274.

Sago, H., E.J. Carlson, D.J. Smith, J. Kilbridge, E.M. Rubin, W.C. Mobley, C.J. Epstein, and T.T. Huang. 1998. Ts1Cje, a partial trisomy 16 mouse model for Down syndrome exhibits learning and behavioral abnormalities. *Proc. Natl. Acad. Sci.* **95:** 6256–6261.

Schwartz, D.C. and C.R. Cantor. 1984. Separation of yeast

chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37:** 67–75.

Scott, H.S., H. Chen, C. Rossier, M.D. Lalioti, and S.E. Antonarakis. 1997. Isolation of a human gene (HES1) with homology to an *Escherichia coli* and a zebrafish protein that maps to chromosome 21q22.3. *Hum. Genet.* **99:** 616–623.

Sertie, A.L., M. Quimby, E.S. Moreira, J. Murray, M. Zatz, S.E. Antonarakis, and M.R. Passos-Bueno. 1996. A gene which causes severe ocular alterations and occipital encephalocele (Knobloch syndrome) is mapped to 21q22.3. *Hum. Mol. Genet.* **5:** 843–847.

Smith, D.J., Y. Zhu, J. Zhang, J.-F. Cheng, and E.M. Rubin. 1995. Construction of a panel of transgenic mice containing a contiguous 2-Mb set of YAC/P1 clones from human chromosome 21q22.2. *Genomics* **27:** 425–434.

Straub, R.E., T. Lehner, Y. Luo, J.E. Loth, W. Shao, L. Sharpe, J.R. Alexander, K. Das, R. Simon, and R.R. Fieve. 1994. A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. *Nat. Genet.* **8:** 291–296.

The Finnish-German APECED Consortium, Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. 1997. An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. *Nat. Genet.* **17:** 399–403.

Veske, A., R. Oehlmann, F. Younus, A. Mohyuddin, B. Muller-Myhsok, S.Q. Mehdi, and A. Gal. 1996. Autosomal recessive non-syndromic deafness locus (DFNB8) maps on chromosome 21q22 in a large consanguineous kindred from Pakistan. *Hum. Mol. Genet.* **5:** 165–168.