RESEARCH

Gene Identification in 1.6-Mb Region of the Down Syndrome Region on Chromosome 21

Miki Ohira,[1,3,7] Naohiko Seki,[1] Takahiro Nagase,[1] Emiko Suzuki,[1,3] Nobuo Nomura,[1] Osamu Ohara,[2] Masahira Hattori,[4] Yoshiyuki Sakaki,[4] Toshihiko Eki,[5] Yasufumi Murakami,[5] Toshiyuki Saito,[6] Hitoshi Ichikawa,[3] and Misao Ohki[3]

[1]Laboratory of Gene Structure 1 and [2]Laboratory of DNA Technology, Kazusa DNA Research Institute, Chiba 292, Japan; [3]Radiobiology Division, National Cancer Center Research Institute, Tokyo 104, Japan; [4]Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo 108, Japan; [5]Laboratory of Cellular Physiology, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), Tsukuba 305, Japan; [6]Genome Research Group, National Institute of Radiological Sciences, Chiba 292, Japan

The Down syndrome (DS) region has been defined by analyses of partial trisomy 21. The 2.5-Mb region between *D21S17* and *ERG* is reportedly responsible for the main features of DS. Within this 2.5-Mb region, we focused previously on a distal 1.6-Mb region from an analysis of Japanese DS patients with partial trisomy 21. Previously we also performed exon-trapping and direct cDNA library screening of a fetal brain cDNA library and identified a novel gene *TPRD*. Further screening of a fetal heart cDNA library was performed and a total of 44 possible exons and 97 cDNA clones were obtained and mapped on a *Bam*HI map. By rescreening other cDNA libraries and a RACE reaction, we isolated nearly full-length cDNAs of three additional genes [holocarboxylase synthetase (*HCS*), G protein-coupled inward rectifier potassium channel 2 (*GIRK2*), and a human homolog of *Drosophila minibrain* gene (*MNB*)] and a coding sequence of a novel inward rectifier potassium channel-like gene (*IRKK*). The gene distribution and direction of transcription were determined by mapping both ends of the cDNA sequences. We found that these genes, except *IRKK*, are expressed ubiquitously and are relatively large, extending from 100 kb to 300 kb on the genome. These nearly full-length cDNA sequences should facilitate understanding of the detailed genome structure of the DS region and help to elucidate their role in the etiology of DS.

[The sequence data described in this paper have been submitted to EMBL/GenBank/DDBJ under accession nos. D86550, D86865–D86908, D87291, and D87327–D87328.]

Down syndrome (DS) is the most frequent birth defect (1 in 1000 newborns) and is caused by trisomy 21. Patients exhibit certain clinical features, such as mental retardation, congenital heart defect, and distinct facial and physical appearances (Epstein 1986). Recent studies of DS patients who showed a triplication of only part of chromosome 21 (partial trisomy 21) suggested the existence of a region that was essential for the pathogenesis of DS (Rahmani et al. 1989; Korenberg et al. 1990; Delabar et al. 1993). This region is called the DS region. The 2.5-Mb region from *D21S17* to *ERG* is thought to be essential for the main phenotypic characteristics of DS (Delabar et al. 1993), although an association with other regions cannot be excluded (Korenberg et al. 1994). The cloning and characterization of genes in this region are necessary for understanding the pathogenesis of DS. The physical map for this region is nearly complete (Patil et al. 1994; Ohira et al. 1996a; Osoegawa et al. 1996). Various efforts toward the construction of transcription maps, including an expressed-sequence tag (EST) collection (Chiang et al. 1995), cDNA selection (Cheng et al. 1994; Peterson et al. 1994; Tassone et al. 1995), and exon-trapping (Lucente et al. 1995; Yaspo et al. 1995; H. Chen et al. 1996), are currently under way (for review, see Shimizu et al. 1995).

Within this 2.5-Mb region, we focused on a 1.6-Mb region between a *Not*I site LA68 (*D21S396*,

[7]Corresponding author.
E-MAIL oohira@kazusa.or.jp; FAX + 81-438-52-3931.

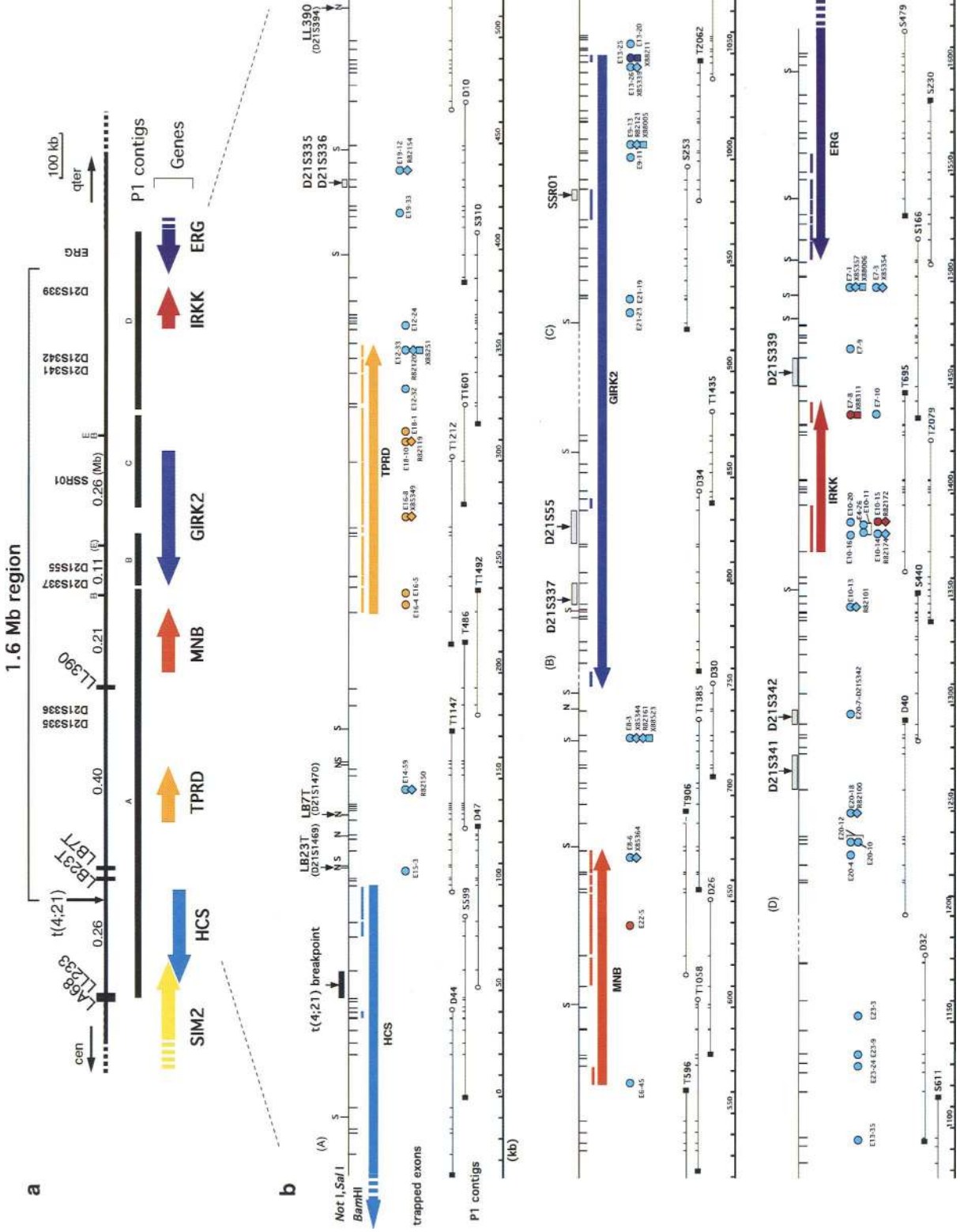

Figure 1 *(See facing page for legend.)*

which is mapped distal to *D21S17*) and *ERG* (see Fig. 1) from an analysis of a Japanese DS family with partial trisomy 21. The patients of the family had a der(4)t(4;21)(q35;q22.2) chromosome in addition to two normal chromosome 21 (Korenberg et al. 1990), and the proximal border of their triplicated region was revealed to be distal to LA68 (Ohira et al. 1996a). This 1.6-Mb region has been triplicated in most of the patients with partial trisomy 21 reported so far, and therefore it may be associated with the phenotypic characteristics commonly exhibited by these patients (such as mental retardation, open mouth, and a flat nasal bridge) (Korenberg et al. 1990, 1994; Delabar et al. 1993). To make a transcription map of this 1.6-Mb region, we constructed a P1 contig map (Ohira et al. 1996a) and performed exon-trapping and cDNA library screening using the P1 clones. Thus far, by combining some of the cDNA clones and exons, we have identified *TPRD*, a novel 9-kb gene with unknown function (Ohira et al. 1996b). Although many genes have been identified in chromosome 21, only two genes other than *TPRD* have been localized in this 1.6-Mb region. One is the G protein-coupled inward rectifier potassium channel 2 gene (*GIRK2*), which has been reported to be mutated in the *weaver* mouse and to be located on the *D21S55* locus in the 1.6-Mb region by Patil et al. (1995). They also reported a cDNA fragment homologous to the *Drosophila minibrain* (*mnb*) gene located 175 kb proximal from *D21S55*. In this study we identified five genes, including these genes, in the 1.6-Mb region, isolated nearly full-length cDNA sequences, and determined their distribution and direction of transcription.

RESULTS

Gene Identification by Direct cDNA Library Screening

Most of the 1.6-Mb region could be covered with 28 P1 clones. By direct cDNA library screening using these 28 P1 clone DNAs, we obtained a total of 97 cDNA clones. Sixty-seven clones were isolated previously from a fetal brain cDNA library (Ohira et al. 1996b) and an additional 30 clones were obtained from a fetal heart cDNA library. These 97 cDNAs were ascertained to be derived from the P1 clones by hybridization to blots of *Bam*HI-digested P1 DNAs. Single-run sequencing of these cDNAs was carried out and PCR primers were designed from both ends of the inserts. PCR analysis using genomic DNAs from human lymphocytes and a chromosome 21q-specific human–hamster hybrid cell line was performed to detect chimeric cDNA clones (see Methods). Although some of the primers were not appropriate for PCR, probably because of splicing junctions, repeats, or conserved sequences, four chimeric cDNAs (two from the fetal brain library and two from the fetal heart library) were identified and excluded.

Comparison of the end sequences allowed us to divide the 93 cDNA clones into four groups and 12 solitary clones. Northern analysis using various human tissue RNA blots showed that representative cDNA clones of the four cDNA groups (FB29-16, FB18-5, FH5-17, and FH4-3) detected four different types of transcripts of 6.5 + 9.0, 8.0 + 9.0, 6.5, and 2.8 + 4.6 kb, respectively (data not shown). The remaining 12 solitary clones could not detect any transcripts in this analysis. Characterization of these clones is described below.

Figure 1 Gene distribution and location of trapped exons in the 1.6-Mb region. (*a*) Schematic representation of the 1.6-Mb region, four P1 contigs, and the identified genes. The *Not*I restriction map of the human chromosome 21q22.2 region (Ichikawa et al. 1993) and DNA markers are shown at the *top*. Vertical bars indicate the *Not*I sites. In addition to *Not*I sites, *Bss*HII and *Eag*I sites, identified in the region distal to LL390 (Ohira et al. 1996a), are also shown as B and E, respectively. Four P1 contigs (A–D) are illustrated *below* the *Not*I map. Genes are shown by thick arrows. *SIM2* (Chen et al. 1995) and *ERG* (Reddy et al. 1987) are partly shown. (*b*) Location of trapped exons and gene distribution on the *Bam*HI restriction map of the 1.6-Mb region. The restriction map is shown at the *top*. *Bam*HI sites are represented by bars without a symbol, and *Not*I and *Sal*I sites are indicated by bars with N and S, respectively. Identified genes and trapped exons are shown under the *Bam*HI map. Colored horizontal bars indicate *Bam*HI fragments hybridized by nearly full-length cDNA sequences and colored thick arrows represent the range of the identified genes. Circles indicate our trapped exons. Diamonds and squares indicate exons that were reported by Lucente et al. (1995) and by Chen et al. (1996), respectively, and identical to ours. Exon color shows the correspondence with the identified genes. P1 contigs (A–D) are shown under the restriction map. Under the P1 contigs, the physical distance from the proximal end of a P1 clone S599 is represented by a thick line with 50-kb intervals. The proximal end of the *ERG* was determined. The *ERG* sequence was published previously by Reddy et al. (1987). PCR with 3′ end primers of *ERG* indicated that these primers were contained in S166, T1526, and S230, but not in T1063. The 3′ end of *ERG* was also mapped by hybridization to a 10-kb *Bam*HI–*Sal*I fragment of S230 P1 DNA, as shown.

To isolate longer cDNA sequences, we used an oligo(dT) primed and size-fractionated cDNA library of the human immature myeloid cell line KG-1 (Nomura et al. 1994). With respect to the gene of 8.0 + 9.0-kb transcripts, previously we reported the 9-kb cDNA sequence that we named *TPRD* (a gene containing tetratricopeptide repeat motifs in the Down syndrome region) (Ohira et al. 1996b). By screening the KG-1 cDNA library with FB29-16 (6.5 + 9.0-kb transcript gene) and FH5-17 (6.5-kb transcript gene), we obtained kg-24 (6.5-kb insert) and kg-68 (4.5-kb insert), respectively. Considering the size of the cDNA insert, kg-24 was expected to cover most of the 6.5-kb transcript of the 6.5 + 9.0-kb gene. A cDNA that extends to 9.0-kb has not been obtained. To isolate the remaining 5′ cDNA sequence (~2 kb) of the 6.5-kb gene, we performed 5′ rapid amplification of cDNA ends (RACE) PCR (Frohman et al. 1988) using fetal brain cDNA with gene-specific primers generated from the 5′ end of the kg-68 sequence, and obtained a 2.0-kb RACE product. Sequence analysis of these clones showed that the 6.5 + 9.0-kb transcripts and a 6.5-kb transcript corresponded to the holocarboxylase synthetase gene (*HCS*) and a human homolog of the *Drosophila minibrain* gene (*MNB*), respectively. Regarding FH4-3 (2.8 + 4.6-kb gene), subsequent analysis revealed that it was derived from chromosome 17 and that the sequence on chromosome 21 might be a pseudogene (data not shown). Therefore, we identified three genes (*HCS*, *TPRD*, and *MNB*) by cDNA library screening (Fig. 1).

Gene Identification by Exon-Trapping

We also carried out exon-trapping using 24 P1 clones from T1147 to S166 (see Fig. 1). The 100 clones that were initially isolated were sequenced, and 52 independent clones were obtained. After excluding false clones that contained *Alu* repeats, a vector sequence, or an *Escherichia coli* sequence, 44 possible exons remained and were mapped on the *Bam*HI fragments of the original P1 clones (see Fig. 1; Table 1). By comparison to the cDNA sequences described above, we found that five exons (E16-4, E16-5, E16-8, E18-1, and E18-10) were derived from the *TPRD* gene and one exon (E22-5) was from the *MNB* gene. In addition, a BLASTN homology search against the GenBank/EMBL database indicated that one exon (E13-25) was part of the known gene human G protein-coupled inward rectifier potassium channel 2 (*GIRK2*). To isolate full-length cDNA for *GIRK2*, E13-25 insert DNA was used for screening of an oligo(dT) primed and size-fractionated cDNA library of whole brain. A cDNA clone TB-2, whose insert size was 2.5 kb, was obtained. Because the E13-25 detected two transcripts of 2.5 + 4.5 kb by Northern analysis (Fig. 2), TB-2 cDNA might have a nearly full-length sequence of 2.5-kb *GIRK2* transcript.

In the homology search, E7-8 was highly homologous to rat inward rectifier potassium channel proteins (Table 1). This result suggests that E7-8 might be part of a novel potassium channel gene. By Northern analysis of multiple tissues, four sizes of transcripts (2.0 + 3.0 + 4.5 + 8.5 kb) were detected, especially in the lung, kidney, and pancreas, by E7-8 (see Fig. 2). We performed a RACE reaction using kidney cDNA and obtained both the 5′ and 3′ RACE products. The 5′ RACE product contained the E10-15 sequence, indicating that E10-15 was also part of the novel potassium channel gene. We designated this gene *IRKK* (inward rectifier K^+ channel from kidney) and performed a sequence analysis. Both RACE products were combined to an ~1.5-kb sequence (RACE7-8). Thus far, no other exons have shown any remarkable homology. Therefore, we identified two potassium channel genes, *GIRK2* and *IRKK*, by exon-trapping.

Gene Characterization

The sequences of four nearly full-length cDNAs [kg-24 (*HCS*), *TPRD*, kg-68 + 2.0 kb RACE product (*MNB*), and TB-2 (*GIRK2*)] and 5′ and 3′ RACE products of E7-8 (*IRKK*) were determined. To elucidate the gene distribution in the 1.6-Mb region, we carried out hybridization of the cDNAs to *Bam*HI-digested P1 fragments. In addition, to determine the directions of transcription, the 5′ and 3′ ends of the cDNA sequences were also mapped on P1 clones by PCR using end-specific primers. Figure 1 shows the results of these analyses.

Known Genes

The holocarboxylase synthetase gene (*HCS*), whose product catalyzes the biotinylation of four biotin-dependent carboxylases, was reported previously by Suzuki et al. (1994) and León-Del-Rio et al. (1995). The cDNA kg-24 (EMBL/GenBank/DDBJ accession no. D87328) had a 6465-bp insert and a poly(A) sequence and, based on a comparison to published data, had an additional 1.1-kb sequence in the 5′ noncoding region and a 3.0-kb sequence in the 3′ noncoding region. The 3′ end was on the P1 clones T1003 and T1335, which were just adjacent to the LL233 *Not*I site (Ohira et al. 1996a), and the 5′ end

Table 1. Trapped Exons and Results of Database Searches

exon name	accession#	length	P1 name	BamHI frag.	BLASTN		BLASTX	
		(bp)		(kb)	putative identification	P(N)	putative identification	P(N)
E15-3	D86879	338	T1147,D47	BN5.0			Q04503;DP87 protein	4.10E-06
E14-59	D86878	149	T486,T1147	B6.8	R82150;human genomic 5A2	identical		
E16-4	D86880	54	T1212	B11.5	F13530;human partial cDNA	7.10E-14		
E16-5	D86908	86	T1212	B11.5	D83077;human TPRD gene	5.00E-28		
E16-8	D86881	86	T1212	B30	X85349;trapped exon SNSB0302	identical		
E18-10	D86882	73	T1601	B25	R82119;human genomic 17G3	identical		
E18-1	D86883	104	T1601	B25	D83077;human TPRD gene	5.00E-28		
E12-32	D86873	71	S310	B15				
E12-33	D86871	103	S310	B7.0	R82120,X88251;human genomic 18C3	identical		
E12-24	D86872	89	S310	B3.4	Z28811;HSB98C051 partial cDNA	5.50E-04	P09949;neurotoxin III homolog(PA-TX)	5.10E-04
E19-33	D86885	103	D10	B3.0				
E19-12	D86884	91	D10	B9.0	R82154,X88251;human genomic 5E5	identical		
E6-45	D86898	95	T1058	B25				
E22-5	D86893	138	T906	B29	X70794;D.melanogaster mnb gene	3.10E-12	P14680;yeast protein kinase YAK1	0.01
E8-6	D86905	85	T1385,T906	B8.5	X85364;trapped exon snsb32	identical		
E8-3	D86904	132	T1385	B7.5	X85344,R82161,X88523;trapped exon SNS03D17	identical		
E21-23	D86892	173	S253	BS11				
E21-19	D86891	67	S253	BS11,1.2				
E9-11	D86906	103	T2062	B3.5				
E9-13	D86907	51	T2062	B6.0	R82121,X88005;human genomic 18G8	identical		
E13-26	D86875	115	S611	B5.0	X85339;trapped exon SNS03B01	identical		
E13-25	D86877	72	S611	B3.2	U24660;human GIRK2, X88211	3.20E-22		
E13-20	D86874	129	S611	B4.0				
E13-35	D86876	165	S611	B4.0				
E23-24	D86894	72	D32	B4.5				
E23-9	D86896	99	D32	B0.9				
E23-3	D86895	167	D32	B15				
E20-4	D86889	84	D40	B9.0				
E20-12	D86887	123	D40	B2.2				
E20-10	D86886	46	D40	B2.2				
E20-18	D86888	78	D40	B22	R82100;human genomic 10E5	identical		
E20-7	D86890	138	S440	B14	Z18605;human STS D21S342	identical		
E10-13	D86866	123	T2079,S440	B3.9	R82101;human genomic 11A12	identical		
E10-16	D86869	182	T2079,T695	B21				
E10-14	D86867	124	T2079,T695	B21	R82174;human genomic 9E4	identical		
E10-20	D86870	188	T2079,T695	B21			P30263;cata-hanpo catalase	3.60E-04
E10-15	D86868	140	T2079,T695	B21	R82172;human genomic 9E1	identical		
E10-11	D86865	158	T2079,T695	B21				
E4-26	D86897	100	T2079,T695	B21				
E7-8	D86902	377	T695	B10	X83585,X88311;rat inward rectifying K-channel	1.30E-17	P35560;rat IRKA	4.40E-12
E7-10	D86899	53	T695	B10			P10244;myb-related protein B	0.075
E7-9	D86903	184	S166	B7.0	X87344;human EST	5.10E-11		
E7-1	D86900	41	S166	BS8.0	X85357,X88006;human trapped exon SNSD0318	identical		
E7-3	D86901	90	S166	BS8.0	X85354;trapped exon SNSC0317	identical		

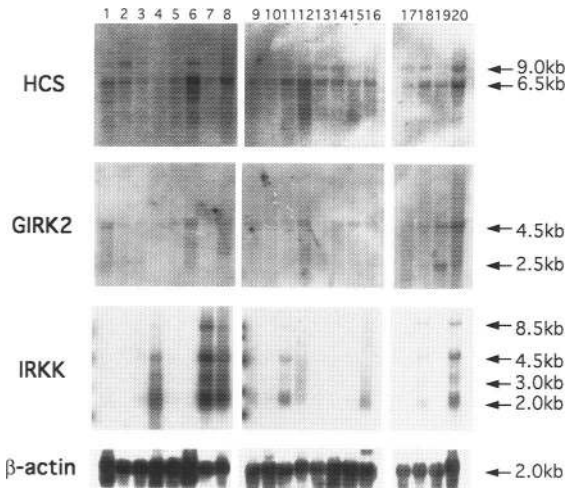

Figure 2 Multiple-tissue Northern blot analysis of *HCS*, *GIRK2*, and *IRKK*. Inserts of the kg-24, TB-2, and 3′ RACE product of *IRKK* were used as probes. The blots used are MTN made from multiple human tissue poly(A) RNAs (1, heart; 2, brain; 3, placenta; 4, lung; 5, liver; 6, skeletal muscle; 7, kidney; 8, pancreas; 9, spleen; 10, thymus; 11, prostate; 12, testis; 13, ovary; 14, small intestine; 15, colon; 16, peripheral blood leukocyte; 17, fetal brain; 18, fetal lung; 19, fetal liver; 20, fetal kidney). The same blots were rehybridized with β-*actin* cDNA.

was on the P1 clone D47, and not on S599 or T1147 (Fig. 1). This suggests that this gene starts distal of the t(4;21) translocation breakpoint in Japanese DS patients with partial trisomy 21, covers a region of ~160–200 kb, and is transcribed in a telomere-to-centromere direction. This result was in agreement with the mapping data recently reported by Blouin et al. (1996).

The 9009-bp nucleotide sequence and physical mapping of the *TPRD* gene have been reported previously (Ohira et al. 1996b).

The human homolog of the *Drosophila mnb* serine/threonine protein kinase gene (*MNB*) has been suggested to reside in the 1.6-Mb region by Patil et al. (1995), and the nucleotide sequence and mapping of this gene have been published recently by Guimerá et al. (1996) and Shindoh et al. (1996). In our laboratory, a 6381-bp region of the *MNB* cDNA sequence (EMBL/GenBank/DDBJ accession no. D86550) was obtained by combining the 2.0-kb 5′ RACE product (1–2028) and 4.5 kb of kg-68 cDNA (1933–6381). The cDNA length was in good agreement with the results of Northern analysis (~6.5 kb; data not shown). The E22-5 exon sequence was found at nucleotides 1963–2110. Our MNB cDNA sequence had an additional 1.4 kb (in the 5′ region) and 2.2–2.5 kb (in the 3′ region) sequences in comparison with the two published sequences. Coding sequence of our cDNA was identical to that of Shindoh et al. (1996), whereas the sequence of Guimerá et al. (1996) contained 27 bp of insertion in the 5′ region (nine residues between amino acids 68 and 69), suggesting the existence of alternatively spliced species. *MNB* extends ~100 kb, and is transcribed in a centromere-to-telomere direction. This result was in good agreement with those of the published papers, and provided improved mapping data.

The *GIRK2* gene was reported previously by Ferrer et al. (1995) and Tsaur et al. (1995), and was also revealed to be mutated in the *weaver* mouse by Patil et al. (1995). Moreover, nearly full-length cDNA sequence of this gene (*Kir3.2*, accession no. U52153) has been submitted recently to GenBank/EMBL databases. Our cDNA clone TB-2 contains a 2447-nucleotide sequence (accession no. D87327). The exon E13-25 sequence corresponds to nucleotides 405–476. We found 2.5- and 4.5-kb transcripts by Northern analysis (Fig. 2). A 4.5-kb transcript was seen in most of the tissues examined, but a 2.5-kb transcript was detected only in some tissues, such as the fetal liver and fetal kidney. No cDNAs extending to 4.5 kb have been obtained yet. Using hybridization, the cDNA insert detected three *Bam*HI fragments in D34/T1435, S253/T2062, and S611, as represented in Figure 1. PCR using 5′ end primers detected T2062 and S611, but 3′ end primers did not detect any positive P1 clones. This suggests that the 3′ end of *GIRK2* might fall between D30 and D34. PCR with 3′ end primers using YAC DNAs that cover this gap supported this speculation (data not shown). Therefore, this gene might extend ~300 kb, and is transcribed in a telomere-to-centromere direction.

A Novel Gene IRKK

IRKK is a novel member of the potassium channel protein family. We obtained a 1489-bp cDNA sequence containing a 1125-bp open reading frame (ORF) that encoded a 375-amino acid protein (accession no. D87291) (Fig. 3a). The nucleotide sequence surrounding the predicted initiation codon at nucleotide 355 contained an in-frame stop codon located 51 bp upstream, which agrees with the Kozak (1987) consensus sequence. The combined RACE product (RACE7-8) has no poly(A) sequence nor any polyadenylation signal sequences. The sequences of exons E7–8 and E10–15 correspond to nucleotides 239–615 and 99–238, respectively. Homology searching of the 375-amino-acid protein

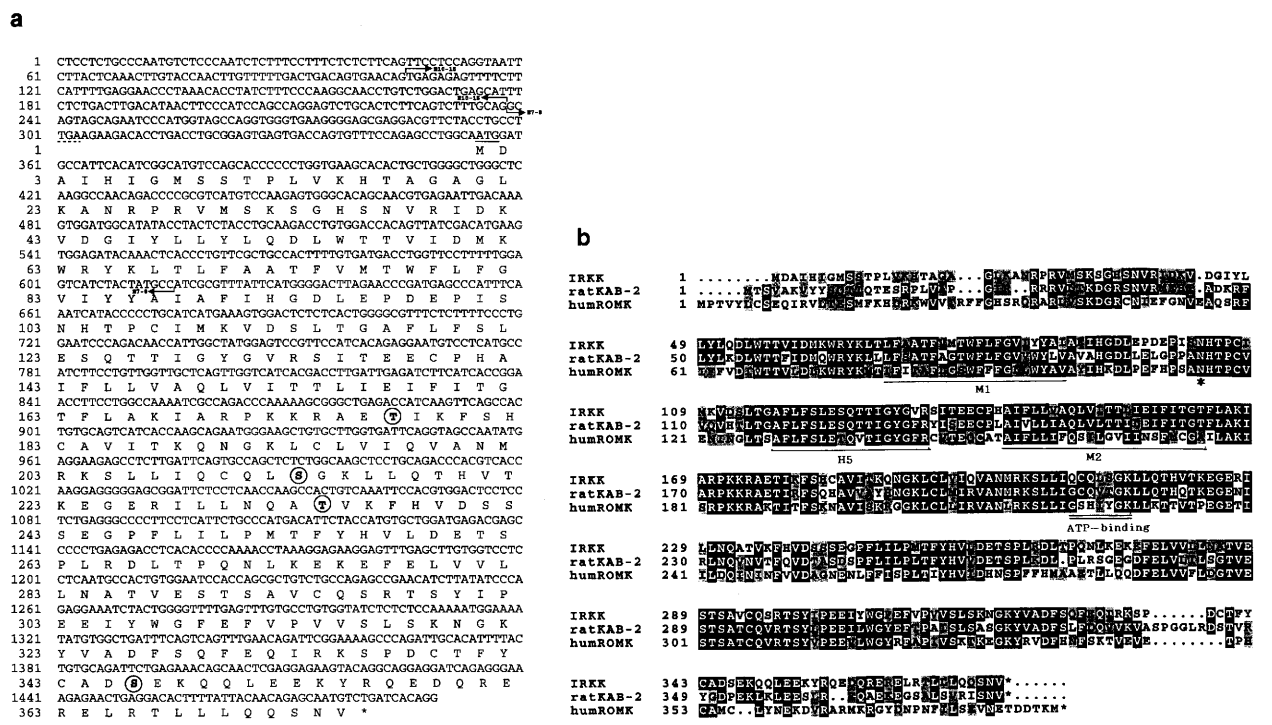

Figure 3 (*a*) Nucleotide sequence and predicted amino acid sequence of the *IRKK* gene. The sequence was constructed with 5′ RACE product (1–572) and 3′ RACE product (457–1489). The stop codon is denoted by an asterisk. The possible translation start site (355–357) is underlined, and an in-frame stop codon located 51 bp upstream is also underlined with a dotted line. The exons E7–8 (239–615) and E10–15 (106–238) sequences are also represented. Four potential protein kinase C phosphorylation sites are circled. (*b*) Comparison of the deduced amino acid sequence of *IRKK* with those of rat K_{AB}-*2* and human *ROMK1*. Identical or similar residues are printed in reverse types. Two putative membrane-spanning hydrophobic segments (M1 and M2) and a pore-forming region (H5) are underlined. A potential N-glycosylation site (asparagine 103 in IRKK) is marked by an asterisk. A Walker type-A motif (see text) representing an ATP-binding loop is underlined.

showed strong homology with ATP-regulated rat inward rectifier potassium channel K_{AB}-*2* gene product (Takumi et al. 1995) (X83585, 64.4% identity in a 343-amino acid overlap) and human ROMK1 potassium channel (Ho et al. 1993; Shuck et al. 1994) (U03884, 50.6% in a 332-amino-acid overlap) (Fig. 3b). Hydropathicity profile analysis (Kyte and Doolittle 1982) indicated two putative membrane-spanning hydrophobic segments (M1 and M2) with a pore-forming region (H5), which were also found in K_{AB}-*2* and ROMK1 (Fig. 3b). A potential N-glycosylation site (Pless and Lennarz 1977) was also seen at asparagine in the predicted extracellular domain of the M1-H5 linker. A Walker type-A motif [GX_4GKX_7 (I/V)] representing a phosphate-binding loop (Saraste et al. 1990) was found in K_{AB}-2 and ROMK1, whereas IRKK contained a similar sequence but had glutamine instead of the first glycine. IRKK also contains four potential protein kinase C phosphorylation sites (positions 177, 213, 234, and 346) (Kishimoto et al. 1985). Transcripts of ~2.0, 3.0, 4.5, and 8.5 kb were detected by multiple-tissue Northern blot analysis with adult and fetal kidney to give the most intense hybridization signals (Fig. 2). It is unclear at present whether all of these transcripts are strictly identical to *IRKK*. Mapping of the 1.5-kb *IRKK* sequence to the *Bam*HI map showed that it extends at least 40–70 kb in a centromere-to-telomere direction.

Other cDNAs and Exons

The remaining 12 solitary cDNA clones (insert size 1–5 kb) with no expression by Northern analysis were also sequenced. Seven of these contained *Alu* or *MER* repetitive sequences. We examined all 12 sequences by FRAMES, which detects ORFs, but none of them had a sufficient ORF (all <500 bp). In addition, these cDNAs seemed to contain no exon–intron junctions, because the length of PCR products amplified from genomic DNAs between 5′- and 3′-end primers of the cDNA sequences seemed to be the same as the length of the cDNA inserts. Al-

though the possibility remained that they could have been derived from the 5′- or 3′-noncoding region, we judged that these clones might be attributable to contamination of genomic clones or immature species of transcripts in cDNA libraries, and excluded them from the transcription map.

Forty-four exons were localized on the *Bam*HI map by hybridization (Fig. 1). The results of homology searching are shown in Table 1. Fifteen exons were identical to the exon sequences reported by Lucente et al. (1995) and five exons were identical to the exon sequences by H. Chen et al. (1996). Nine exons corresponded to *TPRD*, *MNB*, *GIRK2*, and *IRKK*, as described above. The remaining 35 exons without any homology to known genes were analyzed further by Northern analysis using a blot of poly(A) RNAs from fetal brain, fetal heart, and KG-1 cells. These exons displayed no positive bands in these tissues. We then synthesized PCR primers from the exon sequences for subsequent RACE reactions using fetal brain and fetal heart cDNAs. Some of the exon sequences were amplified in these cDNAs, but no RACE product containing ORFs were obtained.

DISCUSSION

Chromosome 21 is one of the most widely examined chromosomes, partly because it is the smallest and a good model for genome analysis, and partly because one additional copy of this chromosome results in the complex phenotypes specific to DS. The molecular basis of the pathogenesis of DS is still unknown. Therefore, the construction of a detailed transcription map is particularly important for understanding the pathogenesis of DS, and for understanding the genome structure in detail. In this paper, we applied exon-trapping and direct cDNA library screening to a 1.6-Mb region of the DS region. We focused on this 1.6-Mb region from an analysis of a Japanese DS family with partial trisomy 21. To identify the genes that are associated with DS, our first goal was to identify all of the genes expressed in the 1.6-Mb region.

How many genes should be identified to complete the transcription map of this region? With regard to other chromosome regions that have been reported recently, on average every 20- to 50-kb region contains one transcription unit (Olsen et al. 1994; Brody et al. 1995; Ansari-Lari et al. 1996; E.Y. Chen et al. 1996; Gong et al. 1996; Kioschis et al. 1996). Although all of these data were obtained from a so-called gene-rich region, most of the 1.6-Mb region may be in a G-band region (band 21q22.2) that tends to contain fewer genes than an R-band region (Saccone et al. 1993). Although genes may remain unidentified, the extent of the five genes (160–200, 100, 100, 300, and 40–70 kb) in this paper suggests that the 1.6-Mb region may have rather large transcription units and may contain smaller number of genes (Fig. 1).

Direct cDNA library screening with P1 DNAs provided many cDNA clones that corresponded to three highly expressed genes (32 clones to *HCS*, 23 to *TPRD*, and 8 to *MNB*). However, false clones containing genome-like sequences with no sufficient ORFs (12%, 12/97) or chimeric sequences (4%, 4/97) were also obtained. We excluded these clones by sequence analysis, by searching the ORFs, and/or by PCR using primers generated from both ends of the inserts.

To isolate genes expressed at a low level or in other tissues, exon-trapping was also performed using individual P1 DNAs (24 P1 clones from T1147 to S166). By use of this method, the *GIRK2* gene and a novel potassium channel gene (*IRKK*) could be obtained. Comparison of 44 exon sequences and the identified genes revealed that five, one, one, and two exons (in total, nine exons) were contained in *TPRD*, *MNB*, *GIRK2*, and *IRKK*, respectively (Fig. 1). Because exon-trapping was not applied to P1 clones D44 and S599, no exon corresponded to *HCS*. Lucente et al. (1995) have reported the localization of 102 exons to the 2.5-Mb region between *D21S17* and *ERG*. By comparing our exon map to their map, we found that 15 exons were identical to those reported by Lucente et al., and three of these were contained in *TPRD* or *IRKK* (two in *TPRD*, and one in *IRKK*; Table 1 and Fig. 1). However, except for these three exons, no other previously reported exons corresponded to the five genes. This may suggest that our exons detect genes in the 1.6-Mb region more efficiently, although the density of exons is lower than Lucente et al. (one exon per ~36 kb vs. ~25 kb). Recently H. Chen et al. (1996) have reported 599 potential exons isolated from chromosome 21. Homology searching showed that five exons were identical to ours, and that seven exons corresponded to the four genes in the 1.6-Mb region (two to *TPRD*, three to *MNB*, one to *GIRK2*, and one to *IRKK*). This efficiency was almost identical to ours (nine exons to the four genes). They estimated that their exons detected ~40% of chromosome 21 genes, so if the number of their exons mapped in the 1.6 Mb was determined, the gene number of this region can be estimated. Combination of exons by these three groups will provide more information of genes in this region. We also mapped 44 trapped

exons to *Bam*HI fragments by hybridization. Two, four, four, 16, and four putative exons were located in the intervals *HCS–TPRD* (~130 kb), *TPRD–MNB* (~200 kb), *MNB–GIRK2* (~100 kb), *GIRK2–IRKK* (~450 kb), and *IRKK–ERG* (~70 kb), respectively (Fig. 1). These exons, especially those in long intervals such as *TPRD–MNB* and *GIRK2–IRKK*, might contain unidentified gene fragments. Further analysis of these exons by RT–PCR with various tissue RNAs is in progress. Once the expressing tissues are identified, RACE experiments will be a powerful approach for isolating full-length cDNAs.

We still have no biological evidence for a correlation between these five genes and the pathogenesis of DS. Further studies, including an elucidation of each protein function, dose effects, and production in transgenic mice, will be necessary. However, we found that the t(4;21) translocation breakpoint in the Japanese DS family was located in the 5′ region of the *HCS* gene (Fig. 1). This suggests that the third copy of HCS protein may be truncated and that the trisomy of *HCS* might not be associated with DS features exhibited by these Japanese DS patients with partial trisomy 21, although the possibility remains that the truncated product of *HCS* also causes their DS phenotypes. The level of transcript should be examined. With respect to the neighboring two genes, *TPRD* and *MNB*, their possible associations to DS phenotypes have been discussed by us (Ohira et al. 1996b) and by Guimerá et al. (1996) and Shindoh et al. (1996). We also identified two inward rectifier potassium channel genes, *GIRK2* and *IRKK*, in the 1.6-Mb region. IRKK is highly homologous to ATP-sensitive rat K_{AB}-2 and human ROMK1, and therefore may have a similar function, although IRKK itself had only an incomplete ATP-binding motif. Homology between IRKK and G-protein-coupled GIRK2 was much lower; 38.8% identity in a 320-amino-acid overlap. The expression patterns of two genes were distinctly different: *GIRK2* was expressed rather ubiquitously, whereas *IRKK* was expressed in some limited tissues. These features suggest that two potassium channels might have distinct functions in different tissues. Concerning the association to the pathogenesis of DS, *GIRK2* could play some role in central nervous abnormality because of the fact that it was mutated in the *weaver* mouse (Patil et al. 1995); whereas *IRKK*, from its specific expression in kidney and pancreas, could associate to the urinary malformation of DS.

In summary, we identified five genes in the 1.6 Mb of the Down syndrome region and isolated nearly full-length cDNA sequences of four genes. These genes, which are positional candidates for DS, should be analyzed further to determine whether they are associated with the phenotypes of DS. The nearly full-length cDNAs we isolated should be good materials for constructing transgenic mice, gene targeting, and functional analyses. In addition, our transcription map should facilitate the understanding of the detailed structure of the human genome.

METHODS

Gene Fragment Isolation

Exon-trapping and cDNA library screening using P1 clones were described previously (Ohira et al. 1996b). To minimize the bias, P1 DNAs were not mixed, but rather were used individually for each trapping/screening experiment. Exon-trapping was carried out with a pSPL3 vector according to Buckler et al. (1991). For direct cDNA library screening, human fetal brain and fetal heart cDNA libraries were purchased from Stratagene. P1 clone DNA used as a probe was prepared by a standard alkaline lysis method (Sambrook et al. 1989) and then purified by ethidium bromide–CsCl centrifugation to reduce the background hybridization by *E. coli* chromosome DNA contamination. Plaques (5×10^5) of each library were screened by each ^{32}P-labeled P1 clone DNA with human placenta DNA to suppress repetitive sequences, under the conditions described previously (Ohira et al. 1996b).

Possible exons and positive cDNA clones were sequenced by single-run sequencing and used for hybridization to blots of the *Bam*HI-digested P1 DNAs to confirm whether these clones were derived from the source P1 clones. In addition, PCR primers were designed from the exons and both ends of the cDNAs, and used for PCR with genomic DNAs from human lymphocytes, a human-hamster hybrid cell line 2Fur1 that contains human chromosome 21q as a sole human component, and its background hamster cell line GlyB, to confirm their origin and to exclude chimeric cDNA sequences.

For further screening to isolate longer cDNAs, we used an oligo(dT)-primed and size-fractionated cDNA library prepared from a human immature myeloid cell line KG-1 or human whole brain. These libraries was constructed as described previously (Nomura et al. 1994; Ohira et al. 1996b).

RACE–PCR

RACE (Frohman et al. 1988) was carried out using a Marathon cDNA amplification kit (Clontech). Marathon-ready (Clontech) cDNAs prepared from human fetal brain or human kidney were used in PCR with an anchor primer provided by the manufacturer and a gene-specific primer designed from the exon and cDNA sequences. PCR was performed in 20 µl of 1× LA PCR buffer II [25 mM TAPS buffer (pH 9.3), 50 mM KCl, 2 mM $MgCl_2$, 1 mM 2-mercaptoethanol] (Takara Shuzo), using 0.5 unit Ex *Taq* DNA polymerase (Takara Shuzo) and 0.55 mg/ml *Taq* Start antibody (Clontech). The conditions for PCR were a 1-min denaturation step at 94°C, followed by 30 cycles of (denaturation at 94°C for 30 sec, annealing and extension at 68°C for 3 min) and finally a 5-min extension at 68°C. PCR reactions were performed on Perkin Elmer PJ2000 thermal

cyclers. PCR products were analyzed by gel electrophoresis using 0.8% agarose, purified from the gel, and cloned into a sequencing vector pGEM-T (Promega).

DNA Sequencing

cDNA clones and RACE–PCR products were sequenced by a shotgun method as follows. The cDNA inserts were prepared by restriction endonuclease digestion, purified by agarose gel electrophoresis, and sonicated. Fragmented DNAs were blunted with mung bean endonuclease and T4 DNA polymerase. After purification of 700- to 1100-bp fragments by agarose gel electrophoresis, these fragments were subcloned into a *Sma*I-digested M13mp18 sequencing vector. Single-stranded DNAs were prepared by an automatic DNA isolation robot (PI100, Kurabo) and used as templates for sequencing. The sequencing reaction and subsequent analysis were performed as described previously (Nomura et al. 1994). Single-run sequencing of exons and cDNA clones was carried out using double-stranded DNAs as described by Ohira et al. (1996b).

Northern Blot Analysis

Human multiple-tissue Northern (MTN) blots were purchased from Clontech. For hybridization, insert DNA of the cDNA clone was labeled with [^{32}P]dCTP by random priming. The hybridization conditions were described previously (Ohira et al. 1996b).

ACKNOWLEDGMENTS

We thank Dr. A. Ootsuyama (National Cancer Research Institute) for help in constructing the system of cDNA library screening. We thank T. Kato, K. Sato, and S. Takahashi for the sequencing analysis. We are also grateful to K. Ishikawa for help with the computer analysis, and to K. Ozawa, T. Tajino, K. Yamada, N. Suzuki, M. Yoshida, and K. Ishii for their technical assistance. We thank Dr. N. Miyajima for valuable help with the computer analysis. This work was supported in part by Grants-in-Aid for Creative Basic Research (Human Genome Program) and Scientific Research on Priority Areas from the Ministry of Education, Science and Culture; by a grant from the Special Coordination Funds for the Promotion of Science and Technology from the Science and Technology Agency; by a Grant-in-Aid for the Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health and Welfare of Japan, and by grants from the Kazusa DNA Research Institute.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Ansari-Lari, M.A., D.M. Muzny, J. Lu, F. Lu, C.E. Lilley, S. Spanos, T. Malley, and R.A. Gibbs. 1996. A gene-rich cluster between the CD4 and triosephosphate isomerase genes at human chromosome 12p13. *Genome Res.* **6:** 314–326.

Blouin, J.L., G.D. Sail, and S.E. Antonarakis. 1996. Mapping of the human holocarboxylase synthetase (HCS) to the Down syndrome critical region of chromosome 21q22. *Ann. Genet.* **39:** 185–188.

Brody, L.C., K.J. Abel, L.H. Castilla, F.J. Couch, D.R. McKinley, G.Y. Yin, P.P. Ho, S. Merajver, S.C. Chandrasekharappa, J. Xu, J.L. Cole, J.P. Struewing, J.M. Valdes, F.S. Collins, and B.L. Weber. 1995. Construction of a transcription map surrounding the BRCA1 locus of human chromosome 17. *Genomics* **25:** 238–247.

Buckler, A.J., D.D. Chang, S.L. Graw, J.D. Brook, D.A. Haber, P.A. Sharp, and D.E. Housman. 1991. Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci.* **88:** 4005–4009.

Chen, E.Y., M. Zollo, R. Mazzarella, A. Ciccodicola, C. Chen, L. Zuo, C. Heiner, F. Burough, M. Repetto, D. Schlessinger, and M. D'Urso. 1996. Long-range sequence analysis in Xq28: Thirteen known and six candidate genes in 219.4 kb of high GC DNA between the RCP/GCP and G6PD loci. *Hum. Mol. Genet.* **5:** 659–668.

Chen, H., R. Chrast, C. Rossier, A. Gos, S.E. Antonarakis, J. Kudoh, A. Yamaki, N. Shindo, H. Maeda, S. Minoshima, and N. Shimizu. 1995. *Single-minded* and Down syndrome? *Nature Genet.* **10:** 9–10.

Chen, H., R. Chrast, C. Rossier, M. Morris, M. Lalioti, and S. Antonarakis. 1996. Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping. *Genome Res.* **6:**747–760.

Cheng, J.F., V. Boyartchuk, and Y. Zhu. 1994. Isolation and mapping of human chromosome 21 cDNA: Progress in constructing a chromosome 21 expression map. *Genomics* **23:** 75–84.

Chiang, P.W., G. Dzida, J. Grumet, J.-F. Cheng, W.-J. Song, E. Crombez, M.L. Van Keuren, and D.M. Kurnit. 1995. Expressed sequence tags from long arm of human chromosome 21. *Genomics* **29:** 383–389.

Delabar, J.M., D. Theophile, Z. Rahmani, Z. Chettouh, J.-L. Blouin, M. Prieur, B. Noel, and P.-M. Sinet. 1993. Molecular mapping of twenty-four features of Down syndrome on chromosome 21. *Eur. J. Hum. Genet.* **1:** 114–124.

Epstein, C.J. 1986. *The consequences of chromosome imbalance: Principles, mechanisms, and models.* Cambridge University Press, Cambridge, UK.

Ferrer, J., C.G. Nichols, E.N. Makhina, L. Salkoff, J. Bernstein, D. Gerhard, J. Wasson, S. Ramanadham, and A. Permutt. 1995. Pancreatic islet cells express a family of inwardly rectifying K^+ channel subunits which interact to form G-protein-activated channels. *J. Biol. Chem.* **270:** 26086–26091.

Frohman, M.A., M.K. Dush, and G.R. Martin. 1988. Rapid production of full-length cDNA from rare transcripts: Amplification using a single gene-specific oligo-nucleotide primer. *Proc. Natl. Acad. Sci.* **85:** 8998–9002.

Gong, W., B.S. Emanuel, J. Collins, D.H. Kim, Z. Wang, F.

Chen, G. Zhang, B. Roe, and M.L. Budarf. 1996. A transcription map of the DiGeorge and velo-cardio-facial syndrome minimal critical region on 22q11. *Hum. Mol. Genet.* **5:** 789–800.

Guimerá, J., C. Casas, C. Pucharcos, A. Solans, A. Domenech, A. Planas, J. Ashley, M. Lovett, X. Estivill, and M. Pritchard. 1996. A human homologue of *Drosophila minibrain* (*MNB*) is expressed in the neuronal regions affected in Down syndrome and maps to the critical region. *Hum. Mol. Genet.* **5:**1305–1310.

Ho, K., C.G. Nichols, W.J. Lederer, J. Lytton, P.M. Vassilev, M.V. Kanazirska, and S.C. Hebert. 1993. Cloning and expression of an inwardly rectifying ATP-regulated potassium channel. *Nature* **362:** 31–38.

Ichikawa, H., F. Hosoda, Y. Arai, K. Shimizu, M. Ohira, and M. Ohki. 1993. A *Not*I restriction map of the entire long arm of human chromosome 21. *Nature Genet.* **4:** 361–366.

Kioschis, P., U.C. Rogner, E. Pick, S.M. Klauck, N. Heiss, R. Siebenhaar, B. Korn, J.F. Coy, J. Laporte, S. Liechti-Gallati, and A. Poustka. 1996. A 900-kb cosmid contig and 10 new transcripts within the candidate region for myotubular myopathy (MTM1). *Genomics* **33:** 365–373.

Kishimoto, A., K. Nishiyama, H. Nakanishi, Y. Uratsuji, H. Nomura, Y. Takeyama, and Y. Nishizawa. 1985. Studies on the phosphorylation of myelin basic protein by protein kinase C and adenosine 3′:5′-monophosphate-dependent protein kinase. *J. Biol. Chem.* **260:** 12492–12499.

Korenberg, J.R., H. Kawashima, S-M. Pulst, T. Ikeuchi, N. Ogasawara, K. Yamamoto, S.A. Schonberg, R. West, L. Allen, E. Magenis, K. Ikawa, N. Taniguchi, and C.J. Epstein. 1990. Molecular definition of a region of chromosome 21 that causes features of the Down syndrome phenotype. *Am. J. Hum. Genet.* **47:** 236–246.

Korenberg, J.R., X.N. Chen, R. Schipper, Z. Sun, R. Gonsky, S. Gerwehr, N. Carpenter, C. Daumer, P. Dignan, C. Disteche, J.M. Graham, Jr., L. Hugdins, B. McGillivray, K. Miyazaki, N. Ogasawara, J.P. Park, R. Pagon, S. Pueshel, G. Sack, B. Say, S. Schuffenhauer, S. Soukup, and T. Yamanaka. 1994. Down syndrome phenotypes: The consequences of chromosomal imbalance. *Proc. Natl. Acad. Sci.* **91:** 4997–5001.

Kozak, M. 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* **196:** 947–950.

Kyte, J. and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157:** 105–132.

León-Del-Rio, A., D. Leclerc, B. Akerman, N. Wakamatsu, and R.A. Gravel. 1995. Isolation of a cDNA encoding human holocarboxylase synthetase by functional complementation of a biotin auxotroph of *Escherichia coli*. *Proc. Natl. Acad. Sci.* **92:** 4626–4630.

Lucente, D., H.M. Chen, D. Shea, S.N. Samec, M. Rutter, R. Chrast, C. Rossier, A. Buckler, S.E. Antonarakis, and M.K. McCormick. 1995. Localization of 102 exons to a 2.5 Mb region involved in Down syndrome. *Hum. Mol. Genet.* **4:** 1305–1311.

Nomura, N., N. Miyajima, T. Sazuka, A. Tanaka, Y. Kawarabayashi, S. Sato, T. Nagase, N. Seki, K. Ishikawa, and S. Tabata. 1994. Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001–KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1. *DNA Res.* **1:** 27–35.

Ohira, M., H. Ichikawa, E. Suzuki, M. Iwaki, K. Suzuki, F. Saito-Ohara, T. Ikeuchi, I. Chumakov, H. Tanahashi, K. Tashiro, Y. Sakaki, and M. Ohki. 1996a. A 1.6-Mb P1-based physical map of the Down syndrome region on chromosome 21. *Genomics* **33:** 65–74.

Ohira, M., A. Ootsuyama, E. Suzuki, H. Ichikawa, N. Seki, T. Nagase, N. Nomura, and M. Ohki. 1996b. Identification of a novel gene containing the tetratricopeptide repeat motifs of the Down syndrome region on chromosome 21. *DNA Res.* **3:** 9–16.

Olsen, A., S. Teglund, D. Nelson, L. Gordon, A. Copeland, A. Georgescu, A. Carrano, and S. Hammarström. 1994. Gene organization of the pregnancy-specific glycoprotein region on human chromosome 19: Assembly and analysis of a 700-kb cosmid contig spanning the region. *Genomics* **23:** 659–668.

Osoegawa, K., R. Susukida, S. Okano, J. Kudoh, S. Minoshima, N. Shimizu, P.J. de Jong, J. Groet, J. Ives, H. Lehrach, D. Nizetic, and E. Soeda. 1996. An integrated map with cosmid/PAC contigs of a 4-Mb Down syndrome critical region. *Genomics* **32:** 375–387.

Patil, N., A. Peterson, A. Rothman, P.J. de Jong, R.M. Myers, and D.R. Cox. 1994. A high resolution physical map of 2.5 Mbp of the Down syndrome region on chromosome 21. *Hum. Mol. Genet.* **3:** 1811–1817.

Patil, N., D.R. Cox, D. Bhat, M. Faham, R.M. Myers, and A.S. Peterson. 1995. A potassium channel mutation in weaver mice implicates membrane excitability in granule cell differentiation. *Nature Genet.* **11:**126–129.

Peterson, A., N. Patil, C. Robbins, L. Wang, D.R. Cox, and R.M. Myers. 1994. A transcript map of the Down syndrome critical region on chromosome 21. *Hum. Mol. Genet.* **3:** 1735–1742.

Pless, D.D. and W.J. Lennarz. 1977. Enzymatic conversion of proteins to glycoproteins. *Proc. Natl. Acad. Sci.* **74:** 134–138.

Rahmani, Z., J.L. Blouin, N. Creau-Goldberg, P.C. Watkins, J.-F. Mattei, M. Poissonnier, M. Prieur, Z. Chettouh, A. Nicole, A. Aurias, P.-M. Sinet, and J.-M. Delabar. 1989. Critical role of the D21S55 region on chromosome 21 in the pathogenesis of Down syndrome. *Proc. Natl. Acad. Sci.* **86:** 5958–5962.

Reddy, E.S.P., V.N. Rao, and T.S. Papas. 1987. The *erg* gene: A human gene related to the *ets* oncogene. *Proc. Natl. Acad. Sci.* **84:** 6131–6135.

Saccone, S., A. De Sario, J. Wiegant, A.K. Raap, G. Della Valle, and G. Bernardi. 1993. Correlations between isochores and chromosomal bands in the human genome. *Proc. Natl. Acad. Sci.* **90:**11929–11933.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Saraste, M., P.S. Sibbald, and A. Wittinghofer. 1990. The P-loop—A common motif in ATP- and GTP-binding proteins. *Trends Biol. Sci.* **15:** 430–434.

Shimizu, N., S.E. Antonarakis, C. Van Brockhoven, D. Patterson, K. Gardiner, D. Nizetic, N. Créau, J.-M. Delabar, J. Korenberg, R. Reeves, J. Doering, A. Chakravati, S. Minoshima, O. Ritter, and J. Cuticchia. 1995. Report of the fifth international workshop on human chromosome 21 mapping 1994. *Cytogenet. Cell Genet.* **70:** 147–182.

Shindoh, N., J. Kudoh, H. Maeda, A. Yamaki, S. Minoshima, Y. Shimizu, and N. Shimizu. 1996. Cloning of a human homolog of the *Drosophila minibrain*/rat Dyrk gene from "the Down syndrome critical region" of chromosome 21. *Biochem. Biophys. Res. Comm.* **225:**92–99.

Shuck, M.E., J.H. Bock, C.W. Benjamin, T.-D. Tsai, K.S. Lee, J.L. Slightom, and M.J. Bienkowski. 1994. Cloning and characterization of multiple forms of the human kidney ROM-K potassium channel. *J. Biol. Chem.* **269:** 24261–24270.

Suzuki, Y., Y. Aoki, Y. Ishida, Y. Chiba, A. Iwamaysu, T. Kishino, N. Niikawa, Y. Matsubara, and K. Narisawa. 1994. Isolation and characterization of mutations in the human holocarboxylase synthetase cDNA. *Nature Genet.* **8:** 122–128.

Takumi, T., T. Ishii, Y. Horio, K. Morishige, N. Takahashi, M. Yamada, T. Yamashita, H. Kiyama, K. Sohmiya, S. Nakanishi, and Y. Kurachi. 1995. A novel ATP-dependent inward rectifier potassium channel expressed predominantly in glial cells. *J. Biol. Chem.* **270:** 16339–16346.

Tassone, F., H. Xu, H. Burkin, S. Weissman, and K. Gardiner. 1995. cDNA selection from 10 Mb of chromosome 21 DNA: Efficiency in transcriptional mapping and reflections of genome organization. *Hum. Mol. Genet.* **4:** 1509–1518.

Tsaur, M.L., T. Menzel, F.-P. Lai, R. Espinosa III, P. Concannon, R.S. Spielman, C.L. Hanis, N.J. Cox, M.M. Le Beau, M.S. German, L.Y. Jan, G.I. Bell, and M. Stoffel. 1995. Isolation of a cDNA clone encoding a KATP channel-like protein expressed in insulin-secreting cells, localization of the human gene to chromosome band 21q22.1, and linkage studies with NIDDM. *Diabetes* **44:** 592–596.

Yaspo, M.L., L. Gellen, R. Mott, B. Korn, D. Nizetic, A. Poustka, and H. Lehrach. 1995. Model for a transcript map of human chromosome 21: Isolation of new coding sequences from exon and enriched cDNA libraries. *Hum. Mol. Genet.* **4:** 1291–1304.

Received September 23, 1996; accepted in revised form December 3, 1996.

Gene identification in 1.6-Mb region of the Down syndrome region on chromosome 21.

M Ohira, N Seki, T Nagase, et al.

Genome Res. 1997 7: 47-58
Access the most recent version at doi:10.1101/gr.7.1.47