Genomics **64,** 203–210 (2000)
doi:10.1006/geno.1999.6109, available online at http://www.idealibrary.com on IDEAL®

# Characterization of a Novel Gene, C21orf6, Mapping to a Critical Region of Chromosome 21q22.1 Involved in the Monosomy 21 Phenotype and of Its Murine Ortholog, orf5

R. Orti,* M. Rachidi,† F. Vialard,* K. Toyama,* C. Lopes,* S. Taudien,‡ A. Rosenthal,‡ M. L. Yaspo,§ P.-M. Sinet,* and J. M. Delabar*,[1]

**UMR 8602 CNRS, UFR Necker Enfants-Malades, 156 rue de Vaugirard, 75730 Paris, France; †Department of Molecular Biology, Institute Pasteur, 25 rue du Docteur Roux, 75015 Paris, France; ‡Department of Genome Analysis, Institute of Mol. Biotechnology, Beutenbergstrasse 11, D-07745 Jena, Germany; and §Max Planck Institute for Molecular Genetics, Ihnestrasse 73, D-14195 Berlin, Germany*

Received September 9, 1999; accepted December 22, 1999

**Phenotypic and molecular analyses of patients with partial chromosome 21 monosomy enabled us to define a region, spanning 2.4 Mb between D21S190 and D21S226, associated with arthrogryposis, mental retardation, hypertonia, and several facial anomalies. The markers of the region were used to screen a total human PAC library (Ioannou, RZPD). We isolated 57 PACs, which formed primary contigs. EST clusters (UNIGENE collection) located in a 6-Mb interval, between D21S260 and D21S263, were mapped in individual bacterial clones. We mapped the WI-17843 cluster to the PAC clone J12100, which contains the two anchor markers LB10T and LA329. The open reading frame extends over 960 bp, with three putative start codons. The 1695-bp cDNA containing a polyadenylation signal should correspond to the full-length cDNA. From the genomic sequence, we deduced that the gene contained five exons and that there was a putative promoter sequence upstream from exon 1. *In silico* screening of DNA databases revealed similarity with a murine EST. The corresponding cDNA (1757 bp) sequence was very similar (>85%) to the human cDNA and had an open reading frame of 876 nucleotides. Somatic hybrid mapping localized the cDNA to mouse chromosome 16. EST analyses and RT-PCR indicated that the third exon in the human gene (exon 2 in the mouse) undergoes alternative splicing. Northern blot hybridization showed that the gene was ubiquitously expressed in humans and mice. The longest mouse clone was used to generate riboprobes, which were hybridized to murine embryos at stages E-9.5, E-10.5, E-12.5, E-13.5, and E-14.5–15, to study the pattern of expression during development. Ubiquitous labeling was observed, with strong signals restricted to limited areas of the telencephalon, the mesencephalon, and the interrhombomeric regions in the central nervous system, and other regions of the body such as the limb buds, branchial arches, and somites.** © 2000 Academic Press

Sequence data from this article have been deposited with the GenBank and EBI Data Libraries under Accession Nos. AF177771 and Y19009.

[1] To whom correspondence should be addressed. Telephone: 33 1 40 61 56 95. Fax: 33 1 40 61 56 90. E-mail: delabar@necker.fr.

## INTRODUCTION

Phenotype–genotype correlation in rare patients with partial monosomy of chromosome 21 enabled us to define a critical region on 21q22.1, located in a region flanked by the genes App and SOD1. Deletion of this region leads to arthrogryposis-like symptoms, mental retardation, hypertonia, and several facial anomalies (Chettouh *et al.,* 1995). To define more accurately the landmarks of this region and to generate new markers, a YAC map was established overlapping the SOD1–App region. FISH analysis of patients using YAC and cosmid clones defined the distal boundary of the critical region as lying between D21S213 and D21S226 (Orti *et al.,* 1997). A carrier of a deletion of the proximal part of the 21q arm (Del21JC) has been described (Korenberg *et al.,* 1991) as presenting normal intelligence and minor physical abnormalies. This deletion has been shown to include the polymorphic marker D21S190 (Dutriaux *et al.,* 1994). Markers D21S226 and D21S190 define a region called M21CR-1, for monosomy 21 chromosomal region 1. Cloning and characterization of the genes from this 2.3-Mb region may provide insight into the potential role of these genes in the phenotype associated with the deletion.

All the markers previously located on the YAC map were used to screen a human PAC library and to construct a sequence-ready map. The resulting contig has been entered in the sequencing process (http://genome.imb-jena.de).

We used this contig to map accurately the EST clusters from the UNIGENE collection, which were previously located by radiation hybrid mapping (Whitehead



0888-7543/00 $35.00

Institute) in a 6-Mb interval (D21S260–D21S263) encompassing the critical region.

Here we report the mapping, cDNA sequence, and genomic structure of a new gene called C21orf6, for chromosome 21 open reading frame 6 (as approved by the Human Gene Nomenclature Committee) and the cDNA sequence of its murine ortholog. Both genes are ubiquitously expressed as deduced from multiple-tissue Northern blots and RT-PCR analyses. Murine embryos and tissue sections were hybridized with riboprobes at various stages of development, and strong signals were found to be restricted to specific regions.

## MATERIALS AND METHODS

*Gene mapping and cDNA extension.* We performed a 5′ rapid amplification of the cDNA using the oligonucleotide 5′ ggacagctcaattttcatcagaaggg 3′ and a human Marathon Ready pancreas cDNA kit (Clontech). The sequence extending from the murine cDNA was generated by PCR with the oligonucleotides 175L, 5′ ttgagatggagcaggctgag 3′, and 175R, 5′ gtccaggctcacattgatcg 3′. The oligonucleotides orfut5′, 5′ gcaagaagggttcaaagatcg 3′, and orf5exon2, 5′ atttcgggcaggactgtagg 3′, were used to amplify DNA fragments specific to the 5′ part of the murine cDNA and to the murine gene. The murine gene was mapped by PCR with the primers Aex3L, 5′ aagacaatggaggggcgatc 3′, and Aex3R, 5′ gacaattcctttgcccactc 3′ (murine exon 3), using DNA from a somatic hybrid hamster cell line, 9-6AZ2, containing mouse chromosomes 16 and Y.

*PCR product cloning and sequencing.* The PCR products were cloned with a TA cloning kit (Invitrogen), extracted with the QIAprep Spin miniprep kit, and sequenced with the thermosequenase kit (Amersham) using an ALF DNA sequencer (Pharmacia).

*Tissue expression pattern.* The tissue expression pattern was studied with human and murine fetal and adult tissues, by Northern blots (Clontech) or with panels of cDNAs (Clontech). The sequences of the oligonucleotides used for RT-PCR studies were orf2R, 5′ gcatgctaaacccagacagg 3′, orf2L, 5′ aagacaatggaggggcgatc 3′ (human cDNA), Aex3L and Aex3R with the mouse cDNA, and 175 L and Aex3R (alternative splicing of mouse cDNAs).

*In toto and tissue section hybridizations.* A sample (2 $\mu$g) of the mouse cDNA clone IMAGE (GenBank GI 1436418) was linearized with the appropriate restriction enzyme and transcribed using the DIG RNA labeling kit (Boehringer Mannheim). Embryos were hybridized *in toto* with 2 $\mu$g of the sense and antisense DIG riboprobe as previously described (Conlon and Rossant, 1992).

*In situ* hybridization of C21orf6 mRNAs was performed using sagittal sections of embryos with $\alpha$-$^{35}$S-UTP-labeled riboprobes. All sections were fixed in 4% (w/v) paraformaldehyde and incubated with proteinase K. They were hybridized at 52°C in a humid chamber and washed with increasingly stringent SSC solutions with the final wash in 2× SSC at 65°C. The negative controls underwent the same procedure with the sense riboprobe.

## RESULTS

### *Characterization of a Human cDNA and of Its Murine Ortholog*

Comparison of the location of the monosomy 21 chromosomal region 1 with the radiation hybrid map (Whitehead Institute) indicated that M21CR-1 is located in the 6-Mb region between D21S260 and D21S263. ESTs (from the UniGene collection) located in this interval were further localized on the YAC map (Orti *et al.*, 1997) and on the PAC contig (unpublished results). One of the positive markers, the WI-17843 cluster, was found to be located on the PAC clone J12100. The 2 longest of the 38 clones from this cluster (IMAGE clones 569355 and 925642; 1695 bp) were sequenced; translation and 3′UTR analyses revealed a 960-bp ORF (open reading frame), three potential ATGs (at positions 67, 85, and 153 bp from the transcription start site), and a poly(A) signal at position 1665 (Fig. 1). The first ATG is given as the most probable start codon by the prediction program ATGpr (http://www.hri.co.jp/atgpr); however, the third ATG is the strongest start site for translation according to Kozak's (1996) classification.

Comparison with the size of the band detected by Northern blotting (Fig. 4a) and the absence of extension by 5′ RACE PCR indicate that these two clones probably correspond to the full-length cDNA. BLASTP analysis revealed no significant sequence similarity. This gene was named C21orf6. The corresponding protein is classified as a secreted protein by SOSUI (http://www.tuat.ac.jp/~mitaku/adv_sosui) and as a cytoplasmic protein by PSORT analysis (http://psort.nibb.ac.jp:8800/). A Prosite scan (Bairoch *et al.*, 1997) revealed the presence of five types of potential protein modification, N-glycosylation, cAMP phosphorylation, PKC phosphorylation, casein kinase II phosphorylation, N-myristoylation, and a site for class II aminoacyl-tRNA synthetase was also revealed.

To study the developmental expression of C21orf6, we have searched a murine ortholog by screening the murine ESTs databases. We identified a very similar sequence (E value $8 \times 10^{-14}$). The corresponding mouse clone (IMAGE 401442) was sequenced and found to be 1757 bp long. Comparison of the ORFs of the two species indicated that the murine clone did not contain a translation start site. The single-pass sequence AA172421 of another mouse clone, (IMAGE 554462) available from databases, extends the 5′ part of the contig but does not extend the ORF. Unfortunately this IMAGE clone is unavailable at present (incorrect numbering in two different libraries).

To extend the 5′ part of the sequence and to confirm the validity of the (IMAGE 554462) clone, oligonucleotides were chosen based on the GenBank sequence AA172421, and an RT-PCR experiment was performed with mouse cDNAs (Clontech). The sequences of these PCR products correct the AA172421 sequence, giving an open reading frame of 292 amino acids. The 5′ untranslated part of the mouse cDNA is only 68% identical to the sequence of the human first exon of the human cDNA. The alignment of the putative protein encoded by this new contig with the human protein confirms that the region of sequence similarity starts at human residue 30 and extends to the end of the human polypeptide with 80% identity and 86% similarity (Fig. 2). The murine translation initiation site corresponds to the third ATG of the human sequence. Analysis of the murine sequence of the mouse protein showed that three protein modification sites for glyco-


**FIG. 1.** Intron/exon organization and sequence of the human C21orf6 gene. The putative promoter sequence was analyzed with MatInspector version 2.2 using the Transfac 3.3 matrices. The core- and the matrix-related regions of potential *cis*-regulatory elements on the plus and the minus (−) strands are underlined. The nucleotides are numbered with respect to the first ATG codon. The *Not*I restriction site LB10T maps to nucleotide position −611 (sequence in boldface type). Broken arrow indicates the transcriptional start site. The polyadenylation signal and the start site for the mouse open reading frame are in boldface type. The intron ends are in lowercase with the size of introns, derived from the genomic sequence of clones E0479 and J12100 (GenBank Accession No. AF129075), given the lowercase letters in boldface type indicate the consensus 5′ splice donor and 3′ splice acceptor. An asterisk indicates the position of the TGA stop codon. The sequences of the ORF2L/R oligonucleotides used for PCR expression analysis with mutiple tissue cDNA panels (MTC) are underlined.

sylation, PKC phosphorylation, and N-myristoylation were conserved (Fig. 2). The murine ortholog was named orf5 according to the MGD Nomenclature Committee.

### *Gene Organization and Computer Promoter Analysis*

PCR screening of the sequence-ready contig mapped this transcriptional unit to the J12100 PAC clone (120 kb). Genomic Southern hybridization of *Eco*RI digests of DNA from PAC clone J12100, from human lymphoblasts, and from CHR21 of the human mouse somatic hybrid WA17 indicated a single location (data not shown). This PAC clone contains the two *Not*I linking clones LB10T and LA329 (Ichikawa *et al.*, 1993). The whole sequence of the clone is available at http://genome.imb-jena.de. Comparison of the sequences of the cDNA and PAC clones made it possible to determine the exon/intron organization (Fig. 1). This gene is located 545 bp downstream from the *Not*I site LB10T, which lies within a putative promoter. Computer analysis of the region upstream from the transcription start site predicts no canonical TATA or CAAT boxes, but did identify two GC-rich elements: 70% G + C from −191 to −89 bp and 61% G + C from −627 to −451 bp. We searched for potential *cis*-acting regulatory elements with Matinspector V2.2 (Quandt *et al.*, 1995) using the TransFac database (Wingender *et al.*, 1997) (Fig. 1). The transcriptional unit is organized into five exons, and the gene spans 13.7 kb (Fig. 3). All intron/exon boundaries are consistent with the consensus sequences for the donor and acceptor splice sites (Fig. 1).

In addition, PCR and Southern blotting showed that the mouse cDNA maps to chromosome 16. We used DNA from a mouse–hamster somatic hybrid 9-6AZ2

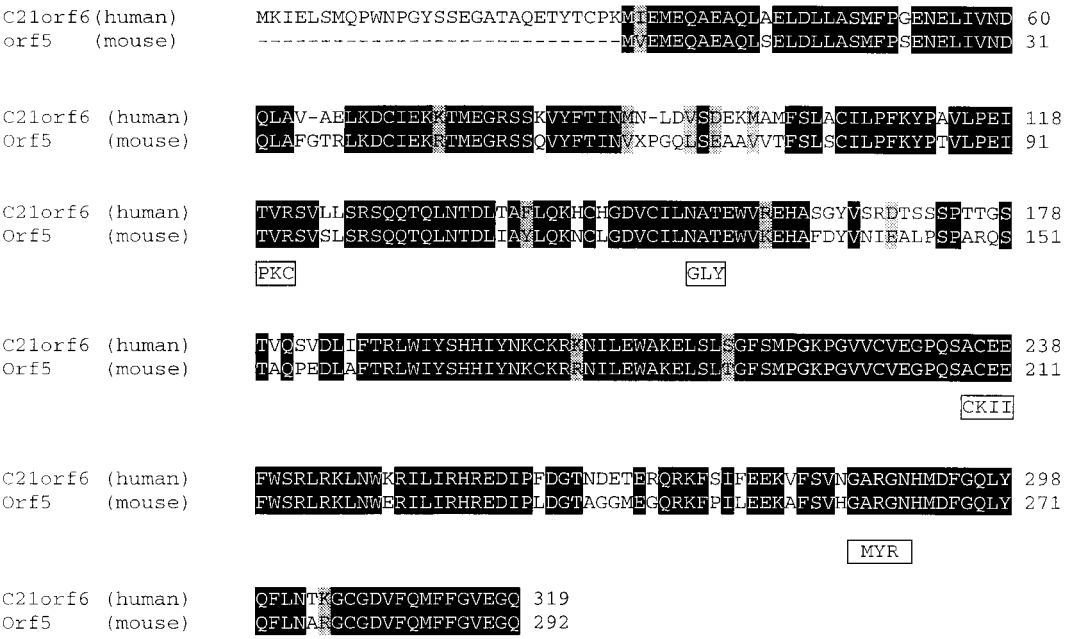


**FIG. 2.** Amino acid sequence comparison of the predicted human C21orf6 protein with its mouse ortholog. Black boxes indicate amino acid identity, and gray boxes indicate amino acid similarity. Boxes below the sequences indicate the protein kinase C phosphorylation site (PKC), N-glycosylation site (GLY), casein kinase II phosphorylation site (CKII), and N-myristoylation site (MYR). The two sequences were aligned using CLUSTALW software.

containing the mouse chromosomes 16 and Y (Hibino *et al.*, 1991) and oligonucleotides derived from the putative exon 3 (data not shown).

*Tissue Expression: Northern Blots and RT-PCR*

Human and mouse C21orf6 genes were found to be widely expressed in various fetal and adult tissues, on Northern blots probed with the human (IMAGE 925642) or mouse clone (IMAGE 401442). A ubiquitous 1.6- to 1.8-kb signal was detected in all tissues tested (Fig. 4a) from fetal and adult human. A similar band (1.6–1.8 kb) was observed with mouse RNA from various tissues from fetuses and adults. RT-PCR on human fetal and adult cDNAs from various tissues (MTC, Clontech) was also performed with oligonucleotides chosen to flank exon 3, which was found to be missing from some of the EST sequences. The expected size of PCR product was 421 bp. Two bands were amplified ubiquitously, but with different relative intensities: one of the expected size and another of 353 bp (Figs. 4c and 4d) resulting from the alternative splicing of exon 3. RT-PCR was also performed using a mouse MTC cDNA panel. RT-PCR with oligonucleotides corresponding to putative murine exon 3 demonstrated the presence of the cDNA in all tissues. The alternative splicing is conserved in the mouse; two bands were generated from brain cDNA with two oligonucleotides 175L and Aex3R flanking exon 2, which corresponds to exon 3 in the human sequence (Fig. 4f). To confirm that there is no murine equivalent to human exon 1, we performed RT-PCR on testis cDNA and genomic DNA with an oligonucleotide in the 5′ untranslated part of orf5 and another in the potential exon 2. The sizes of

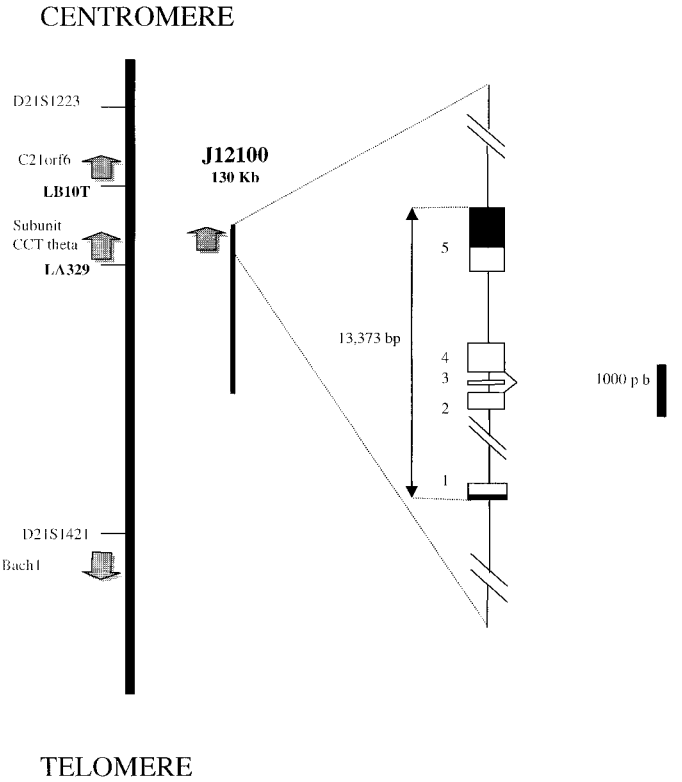


**FIG. 3.** Mapping position of the human gene C21orf6 on chromosome 21. Left vertical line: the gene is oriented from telomere to centromere proximal to the *Not*I linking clone LB10T (545 bp). As regards exon/intron organization, the gene spans 13,373 kb, boxes indicate the exons, and open boxes show the open reading frame. The 3′ and 5′ untranslated regions are shown as black boxes. Alternative splicing between exons 2 and 4 is indicated. Arrows indicate the direction of transcription of the known genes and of C21orf6.

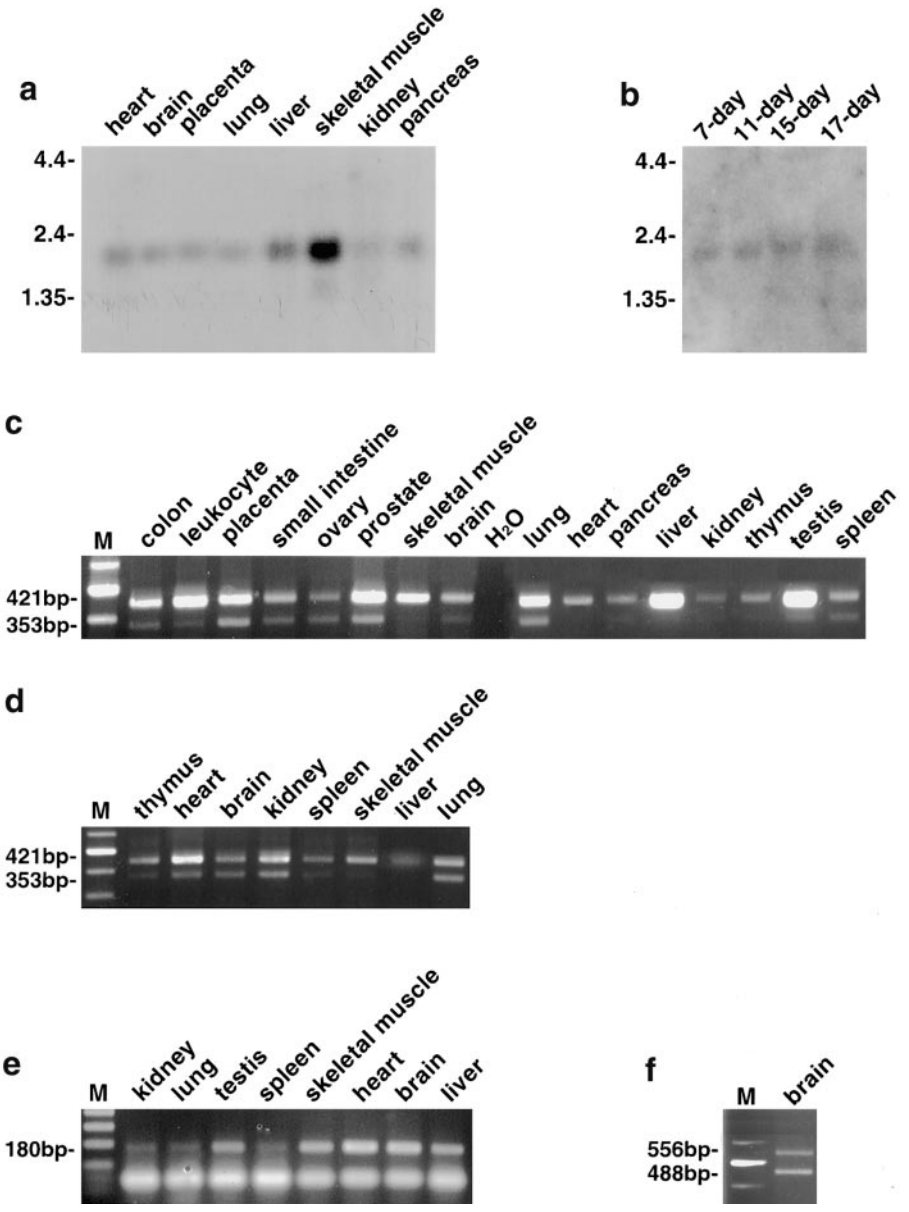


**FIG. 4.** Expression pattern of the human C21orf6 and mouse orf5. (**a**) Northern blot with RNA from various human adult tissues. The marker sizes are indicated on the left in kilobases. (**b**) Northern blot (Clontech) with RNA from mouse embryos at 7, 11, 15, and 17 days. (**c**) RT-PCR with an adult human multiple tissue (cDNA panel) PCR was performed with oligonucleotides located in exons 2 and 4. The PCR products obtained indicate alternative splicing of exon 3. The sizes of the bands are 421 and 353 bp. The marker (M) is a 100-bp ladder. (**d**) RT-PCR with a human fetal tissue cDNA panel. (**e**) RT-PCR with an adult mouse multiple tissue cDNA panel. PCR was performed with oligonucleotides located in the putative exon 3; the PCR product is 180 bp. (**f**) The alternative splicing is conserved in the mouse gene. PCR was performed with brain cDNA and oligonucleotides flanking exon 2; the PCR products are 556 and 486 bp.

the two PCR products differ (392 bp for the cDNA and around 450 bp for the genomic DNA), suggesting that these oligonucleotides amplify a fragment of genomic DNA containing intron 1 (data not shown).

*Expression Pattern of the Mouse orf5 Gene*

We analyzed the spatial expression pattern of this gene, by performing whole-mount *in situ* hybridization with mouse embryos from 9.5 to 12.5 days postcoitum (dpc). Hybridizations were performed with the antisense and the sense riboprobes generated from the mouse clone (IMAGE 401442). The gene was expressed ubiquitously, but its intensity of expression depended on the tissue (Fig. 5a). Strong signals were detected at 9.5 dpc in the somites, branchial arches, and vertebrate anlages (Fig. 5a). At 10.5 dpc, high levels of expression were observed in the branchial arches and in the limb buds (Fig. 5b). In the brain at 12.5 dpc, strong signals were observed in the telencephalon and mesencephalon (Fig. 5c). The dorsal view of the brain showed that the signal was stronger in the area located between rhombomeric segments as if expression were down-regulated in the rhombomeric segments (Fig. 5d). The whisker follicles (wf) and the extremities of

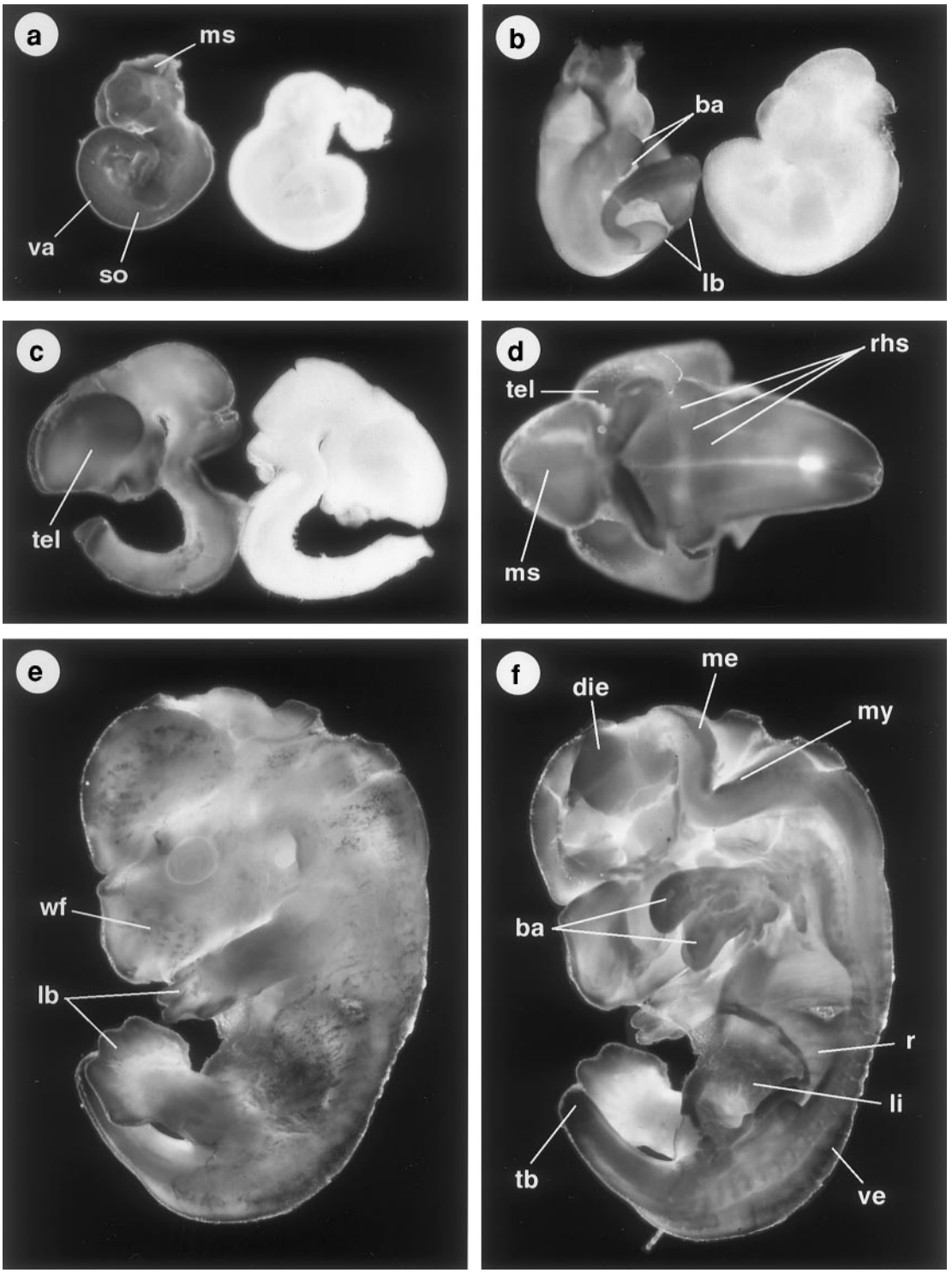


**FIG. 5.** Expression of the mouse orf5 gene. Hybridizations were performed with whole-mount embryos and the antisense (**a, b, c** (**left**) and **d, e, f**) and sense riboprobes (right in **a, b,** and **c**). (**a**) E-9.5 dpc, sagittal view; ms, mesencephalon; va, vertebrate anlages; so, somites. (**b**) E-10.5 dpc, lb, limb buds; ba, branchial arches. (**c**) E-12.5 dpc sagittal view of the brain and part of the spinal cord; tel, telencephalon. (**d**) E-12.5 dpc upper view of the brain; tel, telencephalon; ms, mesensencephalon; rhs, rhombomeric segments (strong expression is seen between the rhombomeric segments). (**e**) A 12.5-dpc embryo outer view. wf, whisker follicles; lb, limb buds. (**f**) Inner view after sagittal section of the embryo shown in (**e**): die, diencephalon; me, metencephalon; my, myelencephalon; ba, branchial arches; r, ribs; li, liver; ve, vertebrae; lb, limb bud; tb, tail bud.

the limb buds were strongly labeled at 12.5 dpc (Fig. 5e). The sagittal half-section of the 12.5-dpc embryo provided an internal view of the embryo (Fig. 5f). Signals were detected in the liver, ribs, limb, and tail buds.

Tissue section hybridization was performed with $^{35}$S-

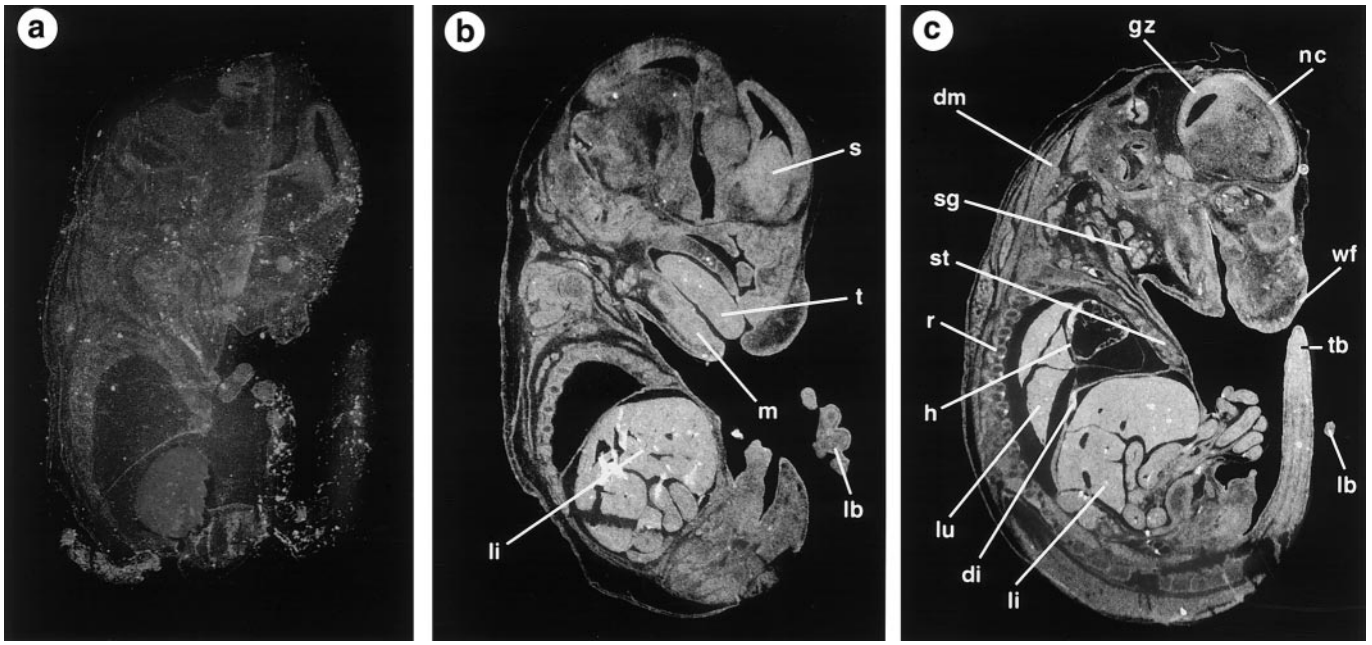


**FIG. 6.** Tissue section hybridization of the mouse orf5 gene. Sagittal sections of an E-13.5 dpc (**b**) and an E-14.5 to E-15 dpc (**c**) embryo hybridized with the antisense riboprobe. Control hybridization of a sagittal section of an E-13.5 dpc embryo with the corresponding sense riboprobe (**a**). (**b**) s, striatum; t, tongue; m, mandibula; li, liver; lb, limb bud. (**c**) nc, neopallial cortex; gz, germinative zone; sg, submandibullar gland; wf, whisker follicle; st, sternum; di, diaphragm; dm, dorsal muscle; r, rib; h, heart; lu, lung; li, liver; tb, tail bud; lb, limb bud.

labeled riboprobes generated from the same clone. These hybridizations also gave a weak ubiquitous signal but with strong expression signals restricted to certain areas: the brain, the neopallial cortex, the germinative zone, and the striatum (Figs. 6a and 6b); the cephalic region, the tongue, the mandibula, and the submandibullar gland were strongly labeled. Strong expression was also detected in the lung, heart, diaphragm, liver, and ribs. Lower levels of expression were detected in dorsal muscle and in bones. The strong signals observed for whisker follicles, limb, and tail buds by whole-mount hybridization were also observed in tissue section hybridization.

## DISCUSSION

We characterized a new gene, C21orf6, located in the monosomy 21 chromosomal region 1 and its mouse ortholog. This gene is located in the J12100 PAC clone proximal to the linking clone LB10T. BLAST searches of databases revealed no significant similarity with known genes. The product of this gene may be a cytoplasmic (PSPORT) or secreted protein (SOSUI). Of the five protein motifs identified by database screening, four are conserved in the murine protein: three phosphorylation sites and a myristoylation site. Exon 3 in humans, which corresponds to the putative exon 2 in the mouse, undergoes alternative splicing, giving rise in both species to two isoforms that may differ in function. Although C21orf6 by Northern blot and RT-PCR was found to be ubiquitously expressed in mouse embryos, whole-mount and tissue section hybridizations showed that strong signals were restricted to specific regions, particularly the branchial arches, some brain tissues, somites, and limb buds. High-level expression was observed in the regions that give rise during developpment to structures involved in the pathologic phenotype. The branchial arches are very much involved in the formation of the head and neck. The phenotype putatively associated with the deletion of the critical region includes several facial anomalies: low hair line, hypertelorism, large nose, large ears, arched palate, and short neck (Chettouh *et al.*, 1995). In the central nervous system, strong expression was detected in the telencephalon, the mesencephalon, and the interrhombomeric regions. Expression seemed to be down-regulated in rhombomeres. Mental retardation is associated with MC21CR-1. The dynamic expression of this gene in the developing limb buds suggests that this gene is involved in distal limb development. Interestingly, it has been reported that arthrogryposis-like symptoms may result from neurogenic, myogenic, and connective tissue disorders (Gordon, 1998). These observations suggest possible roles for C21orf6 in the pathogenesis.

Three other genes have been mapped in this region: GriK1, a kainate subtype of ionotropic glutamate receptor (Gregor *et al.*, 1994), may provide a protective mechanism against hyperexcitability (Cossart *et al.*, 1998). The theta subunit of a chaperonin complex (CCT$\theta$) (Nomura *et al.*, 1994) was located near the linking clone LA329 (Fig. 3) (Yamazaki *et al.*, 1995). This cytosolic chaperonin is involved in the folding of tubulin (Frydman *et al.*, 1992) and actin (Gao *et al.*, 1992) and in neurogenesis (Roobol *et al.*, 1995).

The third gene, Bach1, is a transcriptional regulator

gene expressed ubiquitously (Blouin *et al.*, 1998; Ohira *et al.*, 1998). This gene is described as being involved in hematopoiesis (Igarashi *et al.*, 1998).

To evaluate the involvement of these genes in the monosomy 21 phenotype, further functional studies including immunochemistry and the construction of knockout models are required.

## ACKNOWLEDGMENTS

We thank Dr. C. A. Kozak for the gift of the 9-6AZ2 Chinese hamster mouse hybrids and the E36 hamster cell line. This work was supported by the Centre National de la Recherche Scientifique and by European Union (Contract BMH4-CT98-3039).

## REFERENCES

Bairoch, A., Bucher, P., and Hofmann, K. (1997). The PROSITE database, its status in 1997. *Nucleic Acids Res.* **25**(1): 217–221.

Blouin, J. L., Duriaux Sail, G., Guipponi, M., Rossier, C., Pappasavas, M. P., and Antonarakis, S. E. (1998). Isolation of the human BACH1 transcription regulator gene, which maps to chromosome 21q22.1. *Hum. Genet.* **102**(3): 282–288.

Chettouh, Z., Croquette, M. F., Delobel, B., Gilgenkrants, S., Leonard, C., Maunoury, C., Prieur, M., Rethore, M. O., Sinet, P. M., Chery, M., and Delabar, J. M. (1995). Molecular mapping of 21 features associated with partial monosomy 21: Involvement of the APP-SOD1 region. *Am. J. Hum. Genet.* **57**(1): 62–71.

Conlon, R. A., and Rossant, J. (1992). Exogenous retinoic acid rapidly induces anterior ectopic expression of murine Hox-2 genes in vivo. *Development* **116**(2): 357–368.

Cossart, R., Esclapez, M., Hirsch, J. C., Bernard, C., and Ben-Ari, Y. (1998). GluR5 kainate receptor activation in interneurons increases tonic inhibition of pyramidal cells. *Nat. Neurosci.* **1**(6): 470–478.

Dutriaux, A., Rossier, J., van Hul, W., Nizetic, D., Theophille, D., Delabar, J. M., van Broeckhoven, C., Potier, and M. C. (1994). Cloning and characterization of a 135- to 500-kb region of homology on the long arm of human chromosome 21. *Genomics* **22:** 472–477.

Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J. S., Tempst, P., and Hartl, F. U. (1992). Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits. *EMBO J.* **11**(13): 4767–4778.

Gao, Y., Thomas, J. O., Chow, R. L., Lee, G. H., and Cowan, N. J. (1992). A cytoplasmic chaperonin that catalyzes beta-actin folding. *Cell* **69**(6): 1043–1050.

Gordon, N. (1998). Arthrogryposis multiplex congenita. *Brain Dev.* **20**(7): 507–511.

Gregor, P., Gaston, S. M., Yang, X., O'Regan, J. P., Rosen, D. R., Tanzi, R. E., Patterson, D., Haines, J. L., Horvitz, H. R., Uhl, G. R., and Brown, R. H., Jr. (1994). Genetic and physical mapping of the GLUR5 glutamate receptor gene on human chromosome 21. *Hum. Genet.* **5:** 565–570.

Hibino, Y., Mariano, T. M., Kumar, C. S., Kozak, C. A., and Pestka, S. (1991). Expression and reconstitution of a biologically active mouse interferon-γ receptor in hamster cells. Chromosomal location of an accessory factor. *J. Biol. Chem.* **266**(11): 6948–6951.

Ichikawa, H., Hasada, F., Arai, Y., Shimizu, K., and Ohiva, M. (1993). A *Not*I restriction map of the entire long arm of human chromosome 21. *Nat. Genet.* **4:** 361–366.

Igarashi, K., Hoshino, H., Muto, A., Suwabe, N., Nishikawa, S., Nakauchi, H., and Yamamoto, M. (1998). Multivalent DNA binding complex generated by small Maf and Bach1 as a possible biochemical basis for beta-globin locus control region complex. *J. Biol. Chem.* **273**(19): 11783–11790.

Korenberg, J. R., Kalousek, D. K., Anneren, G., Pulst, S. M., Hall, J. G., Epstein, C. J., and Cox, D. R. (1991). Deletion of chromosome 21 and normal intelligence: Molecular definition of the lesion. *Hum. Genet.* **87**(2): 112–118.

Kozak, M. (1996). Interpreting cDNA sequences: Some insights from studies on translation. *Mamm. Genome* **7**(8): 563–574.

Nomura, N., Miyajima, N., Sazuka, T., Tanaka, A., Kawarabayasi, Y., Sato, S., Nagase, T., Seki, N., Ishikawa, K., and Tabata, S. (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**(1): 27–35.

Ohira, M., Seki, N., Nagase, T., Ishikawa, K., Nomura, N., and Ohara, O. (1998). Characterization of a human homolog (BACH1) of the mouse Bach1 gene encoding a BTB-basic leucine zipper transcription factor and its mapping to chromosome 21q22.1. *Genomics* **47:** 300–306.

Orti, R., Megarbane, A., Maunoury, C., van Broeckhoven, C., Sinet, P. M., and Delabar, J. M. (1997). High-resolution physical mapping of a 6.7-Mb YAC contig spanning a region critical for the monosomy 21 phenotype in 21q21.3–q22.1. *Genomics* **43:** 25–33.

Quandt, K., Fresh, K., Karas, H., Wingender, E., and Werner, T. (1995). MatInd and MatInspector—New fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23:** 4878–4884.

Roobol, A., Holmes, F. E., Hayes, N. V., Baines, A. J., and Carden, M. J. (1995). Cytoplasmic chaperonin complexes enter neurites developing in vitro and differ in subunit composition within single cells. *J. Cell Sci.* **108**(Pt. 4): 1477–1488.

Salamov, A. A., Nishikawa, T., and Swindells, M. B. (1998). Assessing protein coding region integrity in cDNA sequencing projects. *Bioinformatics* **14**(5): 384–390.

Werner, Th. (1999). Models for prediction and recognition of eukaryotic promoters. *Mamm. Genome* **10:** 168–175.

Wingender, E., Kel, A. E., Kel, O. V., Karas, H., Heinemeyer, T., Dietze, P., Knuppel, R., Romaschenko, A. G., and Kolchanov, N. A. (1997). Transfac, TRRD and COMPEL. Towards a federated database system on transcriptional regulation. *Nucleic Acids Res.* **25:** 265–268.

Yamazaki, M., Ono, A., Watanabe, K., Sasaki, K., Tashiro, H., and Nomura (1995). Nucleotide sequence surrounding the locus marker D21S246 on human chromosome 21. *DNA Res.* **2**(4): 187–189.