The CATCH 22 acronym outlines the main clinical features of 22q11.2 deletions (cardiac defects, abnormal facies, thymic hypoplasia, cleft palate and hypocalcemia), usually found in DiGeorge (DGS) and velo-cardio-facial (VCFS) syndromes. Hemizygosity of this region may also be the cause of over 100 different clinical signs. The CATCH 22 locus maps within a 1.5 Mb region, which encompasses several genes. However, no single defect in 22q11.2 hemizygous patients can be ascribed to any gene so far isolated from the critical region of deletion. We have identified a gene in the CATCH 22 critical region, whose functional features and tissue-specific expression suggest a distinct role in embryogenesis. This gene, \textit{UFD1L}, encodes the human homolog of the yeast ubiquitin fusion degradation 1 protein (UFD1p), involved in the degradation of ubiquitin fusion proteins. Cloning and characterization of the murine homolog (\textit{Ufd1l}) showed it to be expressed during embryogenesis in the eyes and in the inner ear primordia. These data suggest that the proteolytic pathway that recognizes ubiquitin fusion proteins for degradation is conserved in vertebrates and that the \textit{UFD1L} gene hemizygosity is the cause of some of the CATCH 22-associated developmental defects.

INTRODUCTION

The CATCH 22 syndrome includes a wide variety of developmental defects associated with microdeletions of the chromosomal region 22q11.2 (1). Notable examples of the CATCH 22 phenotype are the DiGeorge syndrome (MIM*188400), the velo-cardio-facial syndrome (VCFS; MIM 192430) and the Opitz GBBB syndrome (MIM*145410). The DGS/VCFS characteristic features are cleft palate, conotruncal cardiac abnormalities and typical facies, with prominent nose, broad nasal root, narrow palpebral fissures and retrognathia (2,3). However, the DGS/VCFS is clinically heterogeneous, presenting additional abnormalities which include psychiatric disorders, eye defects, upper-limb malformations, renal and urological tract malformations, cerebellar atrophy, tracheal defects and hearing loss (1–7). These different abnormalities are supposed to have a common etiology, based on neural crest cell differentiation defects (8,9). More than 80% of DGS/VCFS patients have detectable deletions in the 22q11.2 region (8,10). The CATCH 22 locus has been restricted to a 1.5 Mb region and several genes have been positioned in this interval (11). None of these genes can account for the defects occurring in CATCH 22 patients, the only possible exception being that of the clathrin heavy chain (CLTD) (12). Since the expression pattern of all the genes mapping in the CATCH 22 region is similar, we suggested that the region itself behaves as a single functional structure (11). Disruption of any of these genes or some common regulatory element would be sufficient to cause the observed phenotype.

We have now identified a gene (\textit{UFD1L}), deleted in CATCH 22 patients, which encodes the human homolog of the yeast ubiquitin (Ub) fusion degradation 1 protein (UFD1p), an essential component of the Ub-dependent proteolytic pathway which degrades Ub fusion proteins (13,14). We also cloned the mouse homolog cDNA (\textit{Ufd1l}) and examined its developmental expression by whole mount \textit{in situ} hybridization. The available data suggest a relationship between \textit{UFD1L} and some clinical features observed in the CATCH 22 syndrome.

RESULTS

Identification and characterization of the \textit{UFD1L} gene and its murine homolog (\textit{Ufd1l})

We isolated a 110 bp cDNA (B6) by direct selection from a six tissue cDNA mixture, using two yeast artificial chromosomeombs (YACs) from the CEPH library (966a8 and 706b10) as driver DNA. The YAC clones span 300 kb of chromosome 22q11.2, which include the \textit{TUPLE1} gene and extend to telomeric sequences (15). The YAC contig encompasses markers D22S941 and D22S942, which border the most commonly deleted region.
in DGS/VCFS (16). A DNA sequence databank search using B6 cDNA showed complete identity to a human expressed sequence tag (EST) (GenBank accession no. A21736). The original EST cDNA clone sequencing revealed a 1029 bp stretch as the longest open reading frame (ORF). The peptide sequence obtained from this ORF showed significant identities to the ubiquitin fusion degradation 1 protein (UFD1p), a yeast protein whose expression and function have been demonstrated recently in Saccharomyces cerevisiae (14). No similar protein product has been identified in other organisms. However, nucleotide sequences closely related to UFD1 are present in the Caenorhabditis elegans genome (GenBank accession no. Z96935). The C.elegans UFD1-like sequences span a 1350 bp region with five putative exons. The translated protein has 30% identity to the yeast UFD1p and 35% to the human homolog (Fig. 1A). Computer-assisted comparison of the A21736 cDNA-derived peptide with other proteins also detects a sequence encoded by a partial cDNA from rice (Oryza sativa) (GenBank accession nos D23997 and D23991) (Fig. 1B).

In order to follow the UFD1L gene expression during development by whole mount in situ hybridization, we isolated the mouse homolog cDNA (Ufd1l), by screening a mouse brain cDNA library with the human clone. The DNA sequence in the coding region is almost identical to the human cDNA. The putative protein products differ by only eight amino acids (Fig. 1A). However, the human protein (as translated from cDNA A21736) shows a tandem repetition of a 36 amino acid domain, which is present in the other species as a single copy (Fig. 1A). RT-PCR studies on human fetal RNAs revealed amplification fragments corresponding to both RNA forms, with and without the tandem duplication, the second appearing to be the main transcript. In cDNA A21736 (with the duplication), the nucleotide blocks corresponding to the repeated domains are separated by a 9 bp intervening sequence, coding for three amino acids. The beginning of this sequence conforms to the best donor splicing site consensus on the intronic side (GTAAGT), the likely cause of an intra-exonic alternative splicing. The duplicated amino acid stretch represents one of the most conserved regions in the UFD1 proteins. It does not show any homology to known peptides. In a zoo blot assay, UFD1p hybridization signals were also detected in genomic DNA of several other mammalian species (Fig. 2).

Mapping the UFD1L within the CATCH 22 critical region and deletion analysis in patients

A cosmID clone (co23) containing most of the UFD1L gene, was used for fluorescence in situ hybridization (FISH) analysis of metaphase chromosomes from 13 DGS/VCFS patients carrying a 22q11.2 deletion, detected with the scF5 probe (15) and nine unrelated unaffected controls (Fig. 3A). Each member of the chromosome 22 pair showed two distinct signals in unaffected individuals, one for co23 and one for a control probe. However, in all affected DGS/VCFS patients tested, one of the chromosomes 22 had two signals, while the other chromosome 22 showed only the control signal. These results demonstrated that DGS/VCFS patients were hemizygous for the UFD1L gene. The 5'-coding region of the UFD1L cDNA turned out to be identical to an EST obtained by exon amplification from a human chromosome 22-specific cosmID library (GenBank accession no. H55625), further confirming the gene location on chromosome 22.
some 22. Fiber-FISH experiments showed that the co23 position is 30–50 kb telomeric to TUPLE1 (Fig. 3B). The **UFD1L** gene spans ~40 kb of genomic DNA.

**Northern blot analysis of UFD1L and expression of Ufd1l during development**

The A21736 cDNA was used as a probe for Northern blot analysis of poly(A)* RNA from several human adult and fetal tissues. This probe detected a 1.2 kb transcript in all the tissues examined, at different levels (Fig. 4). In the adult tissue blot, heart, skeletal muscle and pancreas displayed the highest expression levels (Fig. 4A). High expression levels were also observed in fetal liver and kidney. In fetal mRNA the **UFD1L** cDNA probe detected additional 1.8 and 6 kb signals (Fig. 4B). These may represent either **UFD1L** splice variants or homologous genes expressed during development. We favor the second hypothesis, as we recently have isolated a new cDNA with a 130 internal nucleotide stretch identical to a portion of cDNA A21736 (data not shown).

The expression pattern of the murine homolog of **UFD1L** was investigated on E8.5–E12.5 embryos by whole mount **in situ** RNA hybridization. For these experiments, the most 3’ 600 bp of the murine cDNA were used to generate an antisense riboprobe as well as a sense control probe. Localized expression was observed at high levels in the otocyst between E9.5 and E11.5 (Fig. 5), fading away after E12. High expression levels were also observed in the developing eye (Fig. 5). Cryostat sections of the same embryos localized the positivity to the anterior structures of the eye (lens) (data not shown). Lower, more variable expression levels were also found at different sites in the embryos, such as the developing brain, the lungs and the cardiac outflow tract (data not shown).

**DISCUSSION**

The **UFD1L** gene is the first described homolog of yeast genes involved in the ubiquitin fusion protein degradation (UFD) pathway (1,4). In eukaryotes, post-translational conjugation to ubiquitin is an obligatory preliminary step for degradation of many proteins (17). An isopeptide bond is formed between the C-terminal Gly of Ub and the ε-amino group of a Lys residue in the accepting protein, resulting in a branched poly-ubiquitinated polypeptide. Unlike the branched Ub conjugates, linear Ub adducts are translationally synthesized and cleaved rapidly at the Ub–protein junction by Ub-specific proteases (13,14). When the Ub C-terminal Gly residue is converted into another residue, the cleavage of ubiquitin fusion proteins is inhibited (14). However, proteins with non-removable N-terminal Ub moieties are short lived and recognized promptly by the UFD pathway as a degradation signal (14). The yeast **UFD1p** is one of five proteins active in the UFD pathway isolated thus far. **UFD1p** acts at a post-ubiquitination step and is essential for yeast viability (14). The presence of homologous proteins in vertebrates suggests that this mechanism has been conserved during the whole evolutionary process, although its physiological substrate(s) and functional significance is unknown at present.

The high level expression during embryogenesis suggests that the **UFD1L** gene may be involved in the development of some ectoderm-derived structures. Evidence to date seems to suggest a role for Ub-mediated mechanisms in differentiation (18–24), including the fate determination of the neural crest cells (22), which are defective in the CATCH22 phenotype. These ectodermal cells, originating from the mid-otic placode level to somite 3, migrate through the circumpharyngeal region to populate the pharyngeal and aortic arches, the cardiac outflow tract and the proximal great vessels (9). When they delaminate from the neural tube and start migrating, they accumulate ubiquitin conjugates, which persist during the whole differentiation period (25). If the pharyngeal arches are deprived of incoming cardiac neural crest cells, they increase the expression of proteosomal mRNA, to replace the missing mesenchyme normally produced by the colonizing cells (26).

Ubiquitination must also be important for the development of the inner ear, as suggested by elevated ubiquitin C-terminal hydrolase levels in non-neuronal cochlea cells both during its early development, and later until after hearing has begun (27). The high expression levels of **UFD1L** in the otic vesicle would support this hypothesis. The otocyst, along with some neural crest derivatives, gives rise to most of the inner ear structures. Although neither conductive nor sensorineural hearing defects are usually listed as cardinal features, they are common in CATCH 22. Temporal bone pathology findings in DGS patients with hypacusia include cochlear hypoplasia and dysplasia, sensory ganglia hypoplasia, absence and defective innervation of the semicircular canals, cristae and saccule hypoplasia and reduced number of ganglion neurons (28–31). The expression pattern of **UFD1L** therefor is suggestive of an active role of this protein in determining the hearing loss and certain ocular features sometimes found in CATCH 22 patients. **UFD1L** is the second gene mapping in the CATCH 22 critical region whose functional characteristics relate to specific clinical manifestations of the syndrome. The **CTDL** gene has been suggested to be related to the occurrence of hypotonia (12).

The CATCH 22 syndrome is a complex disorder with significant phenotypic variability and variable penetrance (1). Given the large size of the deletions, it is likely that a number of contiguous genes in the 22q11.2 region contribute to the phenotype (11). Several transcriptional units sharing very similar expression patterns have been identified in the critical region (11,32). It has been suggested that disrupted expression of any of these genes results in the observed phenotype (11). However, since no correlation has been established between the extension and location of the deletion and phenotype, it is difficult to explain...
Figure 3. (A) FISH analysis with cosmid clone co23 in a DGS patient showing a signal on one chromosome 22 only (red signals). A cosmid clone for the βARK2 gene (41), mapping at 22q11, was used as a control probe and indicates the chromosome 22 pair (yellow signals). (B) Fiber-FISH analysis showing the alignment of three probes in the DGSRO. scF5 (15) (containing TUPLE1; red array) and the more centromeric cosmid clone co29 (44) (containing DVL22; left green array) appear contiguous, with a partial overlap (yellow signals due to combined green and red fluorescences). Clone co23 (right green array), containing part of the UFD1L gene, hybridizes 30–50 kb telomeric to scF5.
the CATCH 22 syndrome as an effect of a single gene. For this reason, different mechanisms, including position effects and disruption of a functional transcriptional architecture, have been proposed (11,33). These hypotheses do not exclude the existence in the critical region of a major gene related to specific phenotypic manifestations. Similar examples have been reported in other autosomal diseases, including the split hand/split foot malformation (34), Williams syndrome (35,36), Langer–Giedion syndrome (37), DEFECT 11 syndrome (38) and spinal muscular atrophy (39,40).

Further studies of UFD1L expression during human embryogenesis and loss-of-function experiments in mice will provide insights into the understanding of the influence of this gene in the pathogenesis of CATCH 22 syndrome.

MATERIALS AND METHODS

cDNA isolation and characterization

YACs were subcloned in cosmids (SuperCos, Stratagene). Total yeast DNA was partially digested with Sau3AI and ligated to the BamHI-digested cosmid. Cosmid clones were hybridized to total human genomic DNA, and ∼50 positive cosmid clones were isolated. Sau3AI fragments from 20 randomly picked cosmids were conjugated to biotin by filling of recessed ends using biotin-dATP and the Klenow fragment of DNA polymerase I. Biotinylated DNA was hybridized to cDNA synthesized from six tissue poly(A)+ RNA (heart, liver, placenta, skeletal muscle, fetal brain, kidney), for 48 h at 70°C. cDNAs previously were ligated to custom synthesized adapters. cDNA-biotinylated DNA hybrids were isolated using streptavidin-conjugated magnetic beads (Dynabeads, Dynal). cDNAs were amplified by PCR using an oligonucleotide designed on the adapter sequence as a polymerization primer and cloned in pBluescript plasmid vector (Stratagene). Single clone cDNA PCR products were blotted onto nylon membranes (Hybond, Amersham) and hybridized to human genomic DNA, yeast genomic DNA and pBR322 DNA, in order to rule out clones containing repetitive sequences, vector DNA contaminants and ribosomal cDNA. The remaining cDNAs were hybridized to grid-arrayed cosmid clones. Those hybridizing one or more cosmids were automatically sequenced.

(ABI 370A DNA Sequencer, Perkin Elmer T3, T7 fluorescent primer sequencing kit).

For the mouse cDNA isolation, 10⁶ cDNA clones from a mouse brain library in λgt10 phage vector (Clontech) were hybridized with the human A21736 cDNA. Data bank searches (GenBank, GenEMBL, Swiss Prot and PIR) were run through the Blastn, Blastp and BlastX network service. Sequence analysis was performed using the GCG package. For pairwise alignment, the gap program was used. Known protein motifs were searched with PROSITE.

Northern blot

Northern blot filters were purchased from Clontech Laboratories and probed according the manufacturer’s instructions. The filters were washed at a final stringency of 0.1× SSC, 0.1% SDS at 65°C. Human β-actin (Clontech) was used as an internal standard.

Zoo blot

Zoo blot filters were purchased from Clontech Laboratories and hybridized according to the manufacturer’s instructions. The filters were washed at a final stringency of 0.1× SSC, 0.1% SDS at 65°C. cDNA A21736 was used as a probe.

Chromosome fluorescent in situ hybridization (FISH)

Metaphase chromosomes were prepared from human peripheral blood lymphocytes obtained from 13 DGS/VCFS patients and nine normal controls, as described elsewhere (41). Purified cosmid DNA was labeled with biotinylated dATP and digoxigenated dUTP by nick translation (Life Technologies) and hybridized to chromosome spreads, according to routine procedures. Chromosomes were counterstained with DAPI and visualized as described (41,42). Fiber-FISH analysis was performed using sodium hydroxide-treated slides (43). Cosmid clones were either biotin labeled and detected with fluorescin isothiocyanate–avidin or were digoxigenin labeled and detected with anti-digoxigenin rhodamine-conjugated antibody. Images were charge-coupled device captured and were merged though a Vidas image analyzer (Zeiss).

Mouse embryo in situ hybridization

Antisense probe was transcribed with T7 polymerase from a pBluescript plasmid clone containing the last 600 bp of the mouse Ufd1l cDNA. A sense control probe was produced with T3 polymerase from the same plasmid. DNase was used to remove the DNA template after probe synthesis. The 8.5–12.5 day post-conception (d.p.c.) embryos were obtained from CD1×CD1 matings, dissected and fixed overnight in 4% paraformaldehyde in phosphate-buffered saline (PBS), dehydrated through increasing methanol concentrations and stored in 100% methanol at −20°C. The hybridization protocol included proteinase K treatment for 15 min (8.5 d.p.c. embryos), 20 min (11.5 d.p.c. embryos) or 25 min (12.5 d.p.c. embryos) at room temperature. Pre-hybridization and hybridization were performed overnight at 65°C. The final color reaction for probe detection on embryos was allowed to proceed for 30 min (8.5 d.p.c. embryos) to 90 min (10.5–12 d.p.c. embryos) at room temperature. Pictures were taken at the dissecting microscope.
Figure 5. Whole mount in situ hybridization of mouse embryos, using an antisense riboprobe transcribed from the 3′-terminal region (600 bp) of Ufd1L cDNA. Sense riboprobe from the same sequence was used as a negative control. (A) In E11.5 embryos, hybridization of the otocyst and the central portion of the eyes is evident (right), compared with a sense RNA probed control (left). (B) Detail of the head region of a hybridized E11.5 embryo.

ACKNOWLEDGEMENTS

We would like to thank Donata Penso for nucleotide sequencing, Michele Bisceglia, Massimo Braga, Angelo Boselli and Simona Livraghi for histological analysis of mouse embryos and Giandomenico Palka, Mauro Magnani and Alex Varshavsky for useful discussions. We are grateful to the UK Genome Project for computer resources and reagents. The work was supported by grants from the Italian Telethon (Grants E.399 and D27), Italian Ministry of Health and MURST.

REFERENCES


