Human Molecular Genetics, 1997, Vol. 6, No. 2 247–258

The murine homologue of HIRA, a DiGeorge syndrome candidate gene, is expressed in embryonic structures affected in human CATCH22 patients

Laurens G. Wilming[1], C. A. Sylvia Snoeren[1], Angelique van Rijswijk[1,2], Frank Grosveld[2] and Carel Meijers[1,2,*]

[1]Institute of Pediatric Surgery and [2]Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, the Netherlands

Received August 21, 1996; Revised and Accepted November 1, 1996

A wide spectrum of birth defects is caused by deletions of the DiGeorge syndrome chromosomal region at 22q11. Characteristic features include cranio-facial, cardiac and thymic malformations, which are thought to arise from disturbances in the interactions between hindbrain neural crest cells and the endoderm of the pharyngeal pouches. Several genes have been identified in the shortest region of deletion overlap at 22q11, but nothing is known about the expression of these genes in mammalian embryos. We report here the isolation of several murine embryonic cDNAs of the DiGeorge syndrome candidate gene *HIRA*. We identified several alternatively spliced transcripts. Sequence analysis reveals that Hira bears homology to the p60 subunit of the human Chromatin Assembly Factor I and yeast Hir1p and Hir2p, suggesting that Hira might have some role in chromatin assembly and/or histone regulation. Whole mount *in situ* hybridization of mouse embryos at various stages of development show that *Hira* is ubiquitously expressed. However, higher levels of transcripts are detected in the cranial neural folds, frontonasal mass, first two pharyngeal arches, circumpharyngeal neural crest and the limb buds. Since many of the structures affected in DiGeorge syndrome derive from these *Hira* expressing cell populations we propose that haploinsufficiency of *HIRA* contributes to at least some of the features of the DiGeorge phenotype.

INTRODUCTION

Hemizygosity of a region of chromosome 22q11 has been associated with a wide variety of congenital malformations that receive several diagnostic labels, the main features being covered by the acronym CATCH22 [Cardiac defect, Abnormal facies, Thymic hypoplasia, Cleft palate, Hypocalcemia, (interstitial) 22q11 deletions] (1). The congenital malformation syndromes associated with these deletions include DiGeorge syndrome (DGS), velo-cardio-facial syndrome (VCFS), conotruncal anomaly face syndrome and Opitz syndrome (2–8). For the latter syndrome Robin *et al.* recently reported linkage with loci on chromosome 22q11 and Xp22 (9).

Several well characterized DNA probes can detect CATCH22 deletions (3,4,10–15). The interstitial deletions are large, spanning 2–3 Mb of DNA. A limited number of CATCH22 patients have been reported with an interstitial deletion smaller than the commonly deleted region (16–18). The current shortest region of deletion overlap (SRDO1) is defined between the centromeric breakpoint in DGS patient G (16) and the unbalanced translocation breakpoint in cell line GM00980 [from a VCFS patient with a t(11;22)(q25;q11) translocation (19)]. Recently, Kurahashi *et al.* provided evidence for the existence of a second critical region (SRDO2) located in the distal part of the commonly deleted region (17). The two SRDOs are mutually exclusive.

At present it is not clear whether the CATCH22 phenotype is the result of haploinsufficiency of one or more genes. The balanced t(2;22)(q14.1;q11.1) translocation breakpoint in DGS patient ADU (20) was the major target of positional cloning in various laboratories. These attempts led to the isolation of *DGCR3* and *DGCR5*. *DGCR3* is an open reading frame (ORF) of 260 amino acids (aa) that is interrupted by the translocation breakpoint (21). However, no cDNAs were isolated and expression studies were only informative under low stringency conditions. This prompted Sutherland *et al.* (22) to screen for adjacent coding sequences in the region. They identified a series of alternatively spliced transcripts from the *DGCR5* gene. The *DGCR3* sequence is embedded in an intron of *DGCR5* and in the same transcriptional orientation. The *DGCR5* transcripts do not have an obvious protein encoding potential. They may be functional RNAs. The recent refinement of the SRDO places the ADU breakpoint ~100 kb centromeric to the proximal deletion boundary of SRDO1 (16), so the relevance of *DGCR3* and *DGCR5* for CATCH22 etiology is unclear. It could very well be that the CATCH22 phenotype results from haploinsufficiency of several relevant genes in SRDO1 and that the ADU breakpoint exerts a position effect on such genes.

*To whom correspondence should be addressed

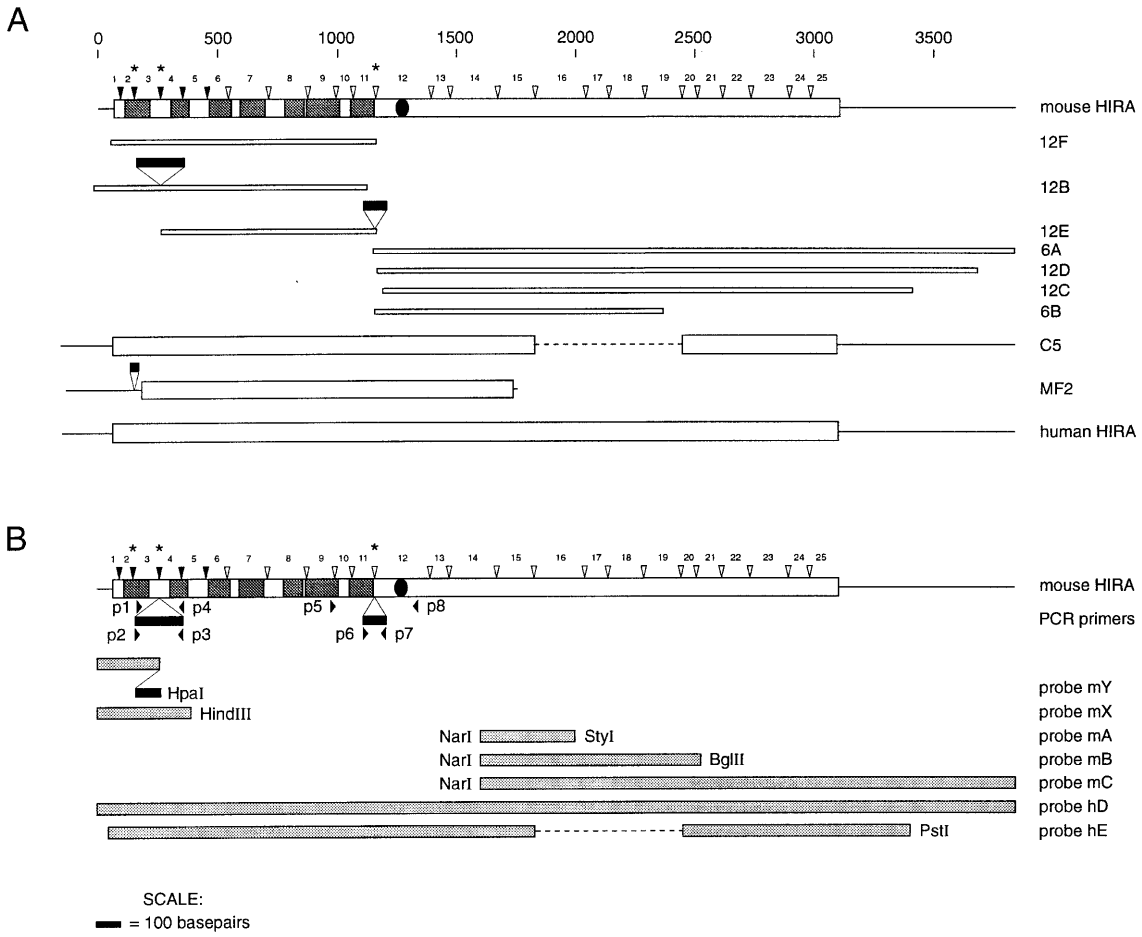


Figure 1. (**A**) A scale schematic drawing of the assembled mouse *Hira* cDNA and the individual cDNA clones isolated from an E11.5 mouse embryo cDNA library (12F, 12B, 12E, 6A, 12D, 12C and 6B). As a reference the human and murine *Tuple1* C5 and MF2 cDNAs and human *HIRA* are shown. Nucleotide positions are indicated on top. Open boxes depict open reading frames; dark gray boxes depict WD40 domain coding regions; black oval indicates the penta-Q stretch coding region; closed arrowheads mark exon–exon boundaries determined from mouse genomic sequences; open arrowheads mark human exon–exon boundaries (exon numbers are between the arrowheads); asterixes above arrowheads show the positions of additional exons 2A, 3A and 11A; black bars indicate the additional exons. (**B**) Drawing of a putative mouse *Hira* transcript with additional exons and the positions of the various primers and probes used. Probes preceded by m are mouse probes, the ones preceded by h are human probes. Restriction sites on mouse or human *HIRA* are given as a reference.

The saturation cloning and sequencing of the 22q11 region are at an advanced stage (23,24). Two transcription maps of SRDO1 have recently been reported (24,25). Currently at least six transcription units have been identified: *HIRA* (*DGCR1*/*TUPLE1*) (26,27), *CLH-22* (clathrin heavy chain like, *CLTCL*/*ES3*/*DGS-K*) (24,25,28,29), *CTP* (mitochondrial citrate transporter protein, *DGS-J*) (24,30), *DGS-I*/*ES2* (24,25), *DGS-H* (24) and *DGS-G* (a serine/threonine kinase) (24). Sequences homologous to the 3′ UTR of the human *Dishevelled 1* gene were mapped just centromeric to the *HIRA* gene (31). It is unclear whether these sequences belong to a gene. Little is known about the expression of the CATCH22 candidate genes in adult and embryonic tissues. It will be important to establish the embryonic expression patterns and functions of the candidate genes.

As an initial step to determine the expression pattern of genes from the CATCH22 SRDO1 we chose to analyze the expression pattern of *HIRA*. The *HIRA* gene (*DGCR1*) was originally isolated as *TUPLE1* by Halford *et al.* (26). Lamour *et al.* isolated human cDNAs containing the complete *TUPLE1* sequence with 621 additional nucleotides in the ORF (27). They renamed the gene *HIRA* because the most significant peptide homology found at the time was with Hir1p and Hir2p, two histone gene repressor proteins from the yeast *Saccharomyces cerevisiae* (32). The predicted HIRA protein is characterized by the presence of seven N-terminal WD40 repeats, two bipartite nuclear localization signals and a penta-Q stretch (frequently seen in transcription factors). The gene product has been implicated in transcriptional regulation based on its homology to gene regulators like yeast Hir1p, Hir2p and Tup1p (32,33).

We report here the identification of cDNA clones of the murine homologue of *HIRA* and studies on the expression of *Hira* during mouse embryogenesis. The sequence of the predicted protein is homologous with the p60 subunit of human Chromatin Assembly Factor 1 (CAF1A) (34). On the basis of sequence homologies with Tup1p, CAF1A, Hir1p and Hir2p, we propose that HIRA is involved in the assembly of chromatin, either by interacting with histones or by regulating their genes. Our results demonstrate that *Hira* mRNA is ubiquitously present from early developmental stages through adulthood. Raised levels of mRNA are detected in the neural folds, pharyngeal arches, circumpharyngeal neural

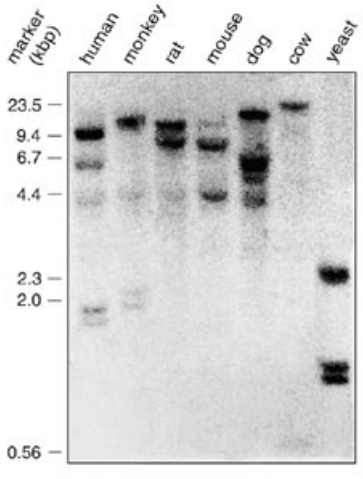


Figure 2. Conservation of the *HIRA* gene. A human *HIRA* probe, hD, detects conserved sequences in *Eco*RI digests of genomic DNA from human, monkey, rat, mouse, dog, cow and yeast.

crest and limb buds. We think that HIRA is normally involved in several aspects of neurogenesis, pharyngeal arch development and limb development by regulating genes that control these processes.

RESULTS

Murine *Hira* cDNA clones coding for a protein with homology with human Chromatin Assembly Factor 1A

We examined the evolutionary conservation of *HIRA* by hybridization of radiolabeled probe hD (from the 4 kb insert of one of our human *HIRA* cDNA clones) (Fig. 1B) to a blot containing *Eco*RI digested genomic DNA from various species. The probe hybridizes to homologous sequences in the DNA of all mentioned organisms (Fig. 2). This high level of conservation from man to yeast indicates an essential function for HIRA.

As an initial step to study whether haploinsufficiency of the *HIRA* gene contributes to the CATCH22 phenotype, we hybridized probe hE (Fig. 1B) to an E11.5 mouse embryo cDNA library under stringent conditions. This screening yielded seven positive cDNA clones of the mouse homologue of the human *HIRA* gene (Fig. 1A). Complete sequencing and assembly of selected clones allowed us to reconstitute a cDNA of 3852 bp (accession number EMBL X99712), containing an ORF of 3045 bp. The 1015 aa predicted protein has seven WD40 repeats (at aa positions 10–44, 68–98, 129–159, 172–202, 234–259, 266–313 and 326–356), two bipartite nuclear localization signals (at aa 267–286 and 626–643), a penta-Q stretch (aa 408–412) and a molecular mass of 111.72 kDa.

Sequence comparison between the human and mouse proteins reveals an overall identity of 95.8% and a similarity of 97.2%. The N-terminal third of the protein, containing the WD40 repeats, is 99.4% conserved (identity and similarity). The difference is in one amino acid in WD40 repeat 3 and one in repeat 6. The C-terminal two-thirds have a relatively higher divergence (identity 93.7%; similarity 96%). Compared to the 1017 aa human HIRA translated sequence, the Hira sequence lacks amino acids 510 and 607 or 608. Our *Hira* sequence is 99% identical to the recently reported mouse *Hira* sequence (35) and 100% identical to the mouse *Tuple1* cDNA clone MF2 (26).

In addition to homology with yeast Hir1p and Hir2p, BlastP and BEAUTY database searches with our mouse Hira sequence reveal high homology to the p60 subunit of human Chromatin Assembly Factor I (CAF1A) (34) and a *Caenorhabditis elegans* protein (36) which, in the absence of a name, we will henceforward refer to as *C.elegans* HIRA-like protein (ceHILP). The overall homology with ceHILP is higher than with Hir1p ($P = 6.1 \times 10^{-77}$ versus 7.6×10^{-45}). We aligned the murine Hira protein sequence with ceHILP, Hir1p, Hir2p and CAF1A (Fig. 3). The overall similarities in this alignment are 39% with ceHILP, 38% with CAF1A, 35% with Hir1p and 31% with Hir2p. It can be readily seen that there are at least three different conserved (functional?) regions: the highest homologies are found in the N-terminal part, region A (aa 1–371) (average similarity 47%); in the middle part, region B (aa 473–658), homology drops (average similarity 18%); C-terminal region C (aa 744–1015) shows again an increased homology (average similarity 35%). Sequence stretches with many gaps in the alignment were excluded from these regions [comparable with the hinge region suggested by Scamps *et al.* (35)]. In the absence of experimental evidence for boundaries of functional domains exon–exon boundaries were chosen to provisionally delimit the regions. The C-terminal boundary of region A corresponds with the insertion point for exon 11A. In our alignment the C-termini of Hira, ceHILP and Hir2p almost coincide. CAF1A terminates at the end of region B and Hir1p terminates in region C.

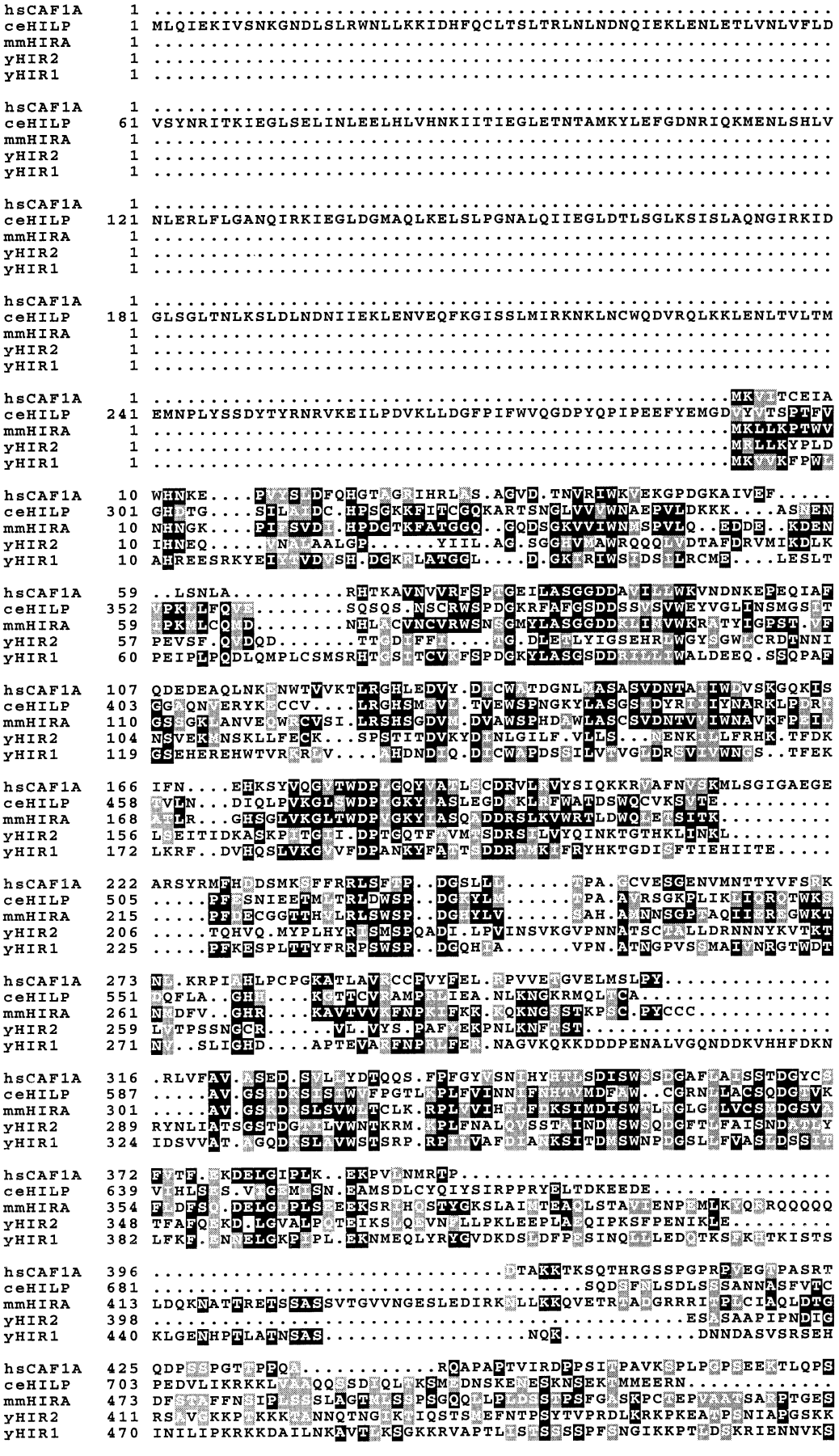
Two Hira transcripts are present from gastrulation stage onward

The presence of *Hira* mRNA in adult mice and embryos was examined by Northern blot analysis of poly(A)$^+$ RNA extracted from embryos at different stages of development and from various adult organs. Probes 12E and mX (Fig. 1A and B) detect transcripts of 4.4 and 2.3 kb (Fig. 4A). The 4.4 kb transcript is more abundant in most organs while the 2.3 kb transcript is more abundant in the adult heart.

In mouse embryos the 4.4 and 2.3 kb *Hira* transcripts are already present during gastrulation stages (E7) (Fig. 4A). At all stages examined the 4.4 kb transcript is most abundant. In the course of development the ratio between the two transcripts does not seem to vary.

Given the length of the human and mouse *HIRA* cDNAs of ~4 kb (26,35), the 4.4 kb *Hira* transcript most likely represents full-length *Hira.* Probes mB and mC (Fig. 1B) only detect the 4.4 kb transcript on the organ and embryo Northern blots, respectively (Fig. 4B). This suggests that the 2.3 kb transcript contains sequences from the 5′ part of *Hira* only. Thus, the 2.3 kb transcript might represent an MF2-like mRNA (26).

To compare *Hira* expression in adult mouse organs with expression in adult human organs, we hybridized a human multiple tissue Northern blot with probe hE (Fig. 1B). Different from what we observed on the Northern blot of adult mouse organs, probe hE detects two transcripts of ~4.5 and 5 kb (Fig. 5A). The ratio between the expression levels of the two transcripts varies from organ to organ. In heart, lung and pancreas the 5.0 kb transcript is most abundant, while in brain and liver the 4.5 kb transcript prevails. Mouse probe mB on the same Northern


hsCAF1A
ceHILP
mmHIRA
yHIR2
yHIR1

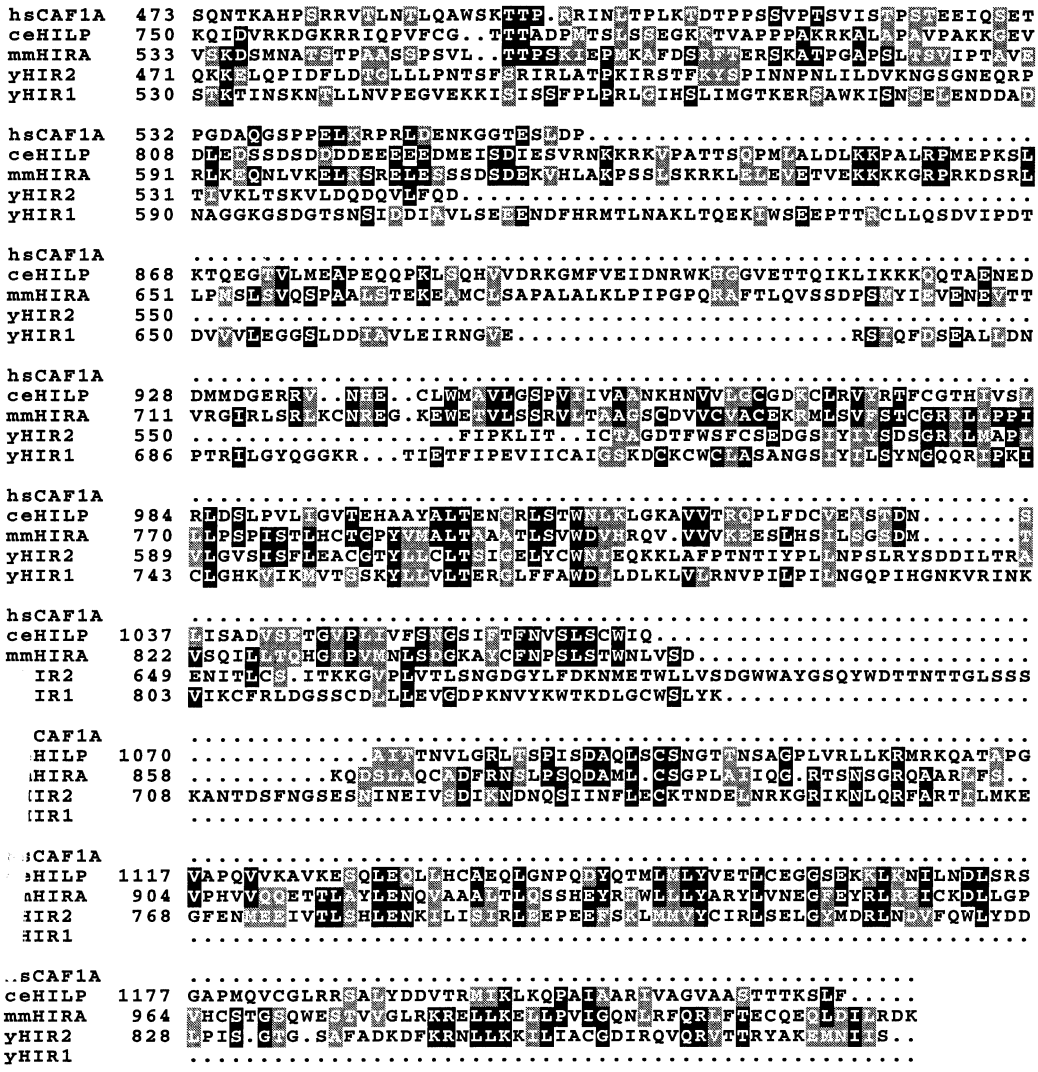


Figure 3. Alignment of mouse Hira sequence with hsCAF1A, ceHILP, yeast Hir2p and Hir1p. Identical amino acids are indicated by black boxes, similar residues by grey boxes. The homologies shown are generated by comparing CAF1A, ceHILP, Hir1p and Hir2p to Hira. Three regions can be distinguished with high, low and intermediate homology. Region A (aa 1–371) spanning the WD repeats; region B (aa 473–658); region C (aa 744–1015). ceHILP has 291 additional N-terminal residues.

blot fails to detect the 5.0 kb transcript (Fig. 5B). Only the 4.5 kb transcript is detected. However, an additional weak signal at 2.4 kb is seen.

Hira is ubiquitously expressed, with enhanced expression in pharyngeal arches, neural folds, circumpharyngeal neural crest and limb buds

The expression of *Hira* in the brain and the heart and its expression during organogenesis, raised the possibility that Hira plays a role in the development of the brain and the pharyngeal region of the head. Therefore we determined the spatial distribution of *Hira* mRNA during mammalian embryogenesis. Whole mount *in situ* hybridization was performed on mouse embryos of 8.5, 9.5 and 10.5 days of gestation, a time-window during which the most critical events in organogenesis are taking place. For these experiments we used single stranded RNA probe mA (Fig. 1B), that only detects the 4.4 kb transcript.

At E8.5 base levels of *Hira* transcripts are present throughout the embryo, including the heart (Fig. 6A and B). Higher levels of transcripts are detected in the cranial neural folds (Fig. 6A). At E9.5 *Hira* transcripts are also widely present (Fig. 6D). Higher levels of expression are seen in pharyngeal arches 1 and 2. We do not detect *Hira* transcripts in the heart anymore at this stage. Sections of gelatin embedded whole mount embryos reveal *Hira* transcripts in the mesenchyme of the pharyngeal arches (not shown). At E10.5 the expression pattern of *Hira* is similar to that at E9.5 (Fig. 6E). The highest levels of transcripts are in the anterior part (including the maxillary component) of arch 1 and the posterior part of arch 2, the frontonasal mass and the limb buds. Just posterior to arches 1 and 2 a circumscribed spot with a high level of *Hira* transcripts is present. This region corresponds well to the region where hindbrain neural crest cells reside before entering the developing gut and the third and fourth pharyngeal arch. In chicken embryos this region is known as the circumpharyngeal neural crest (37).

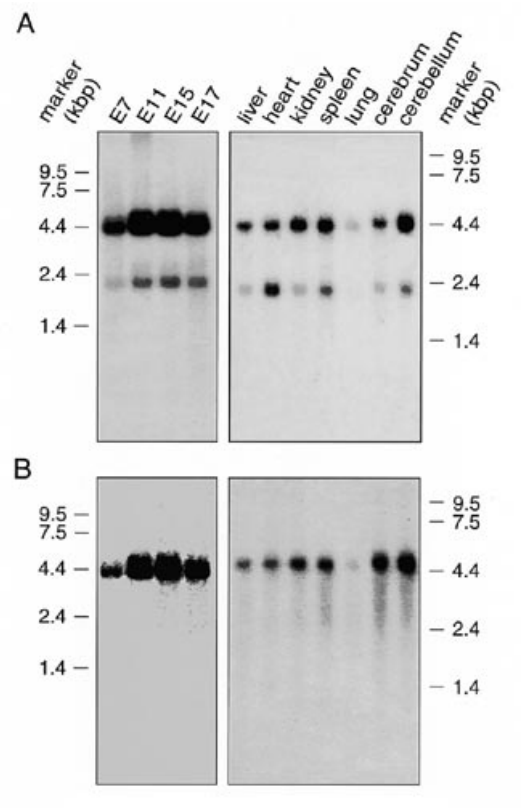


Figure 4. (**A**) Northern blot of poly(A)$^+$ RNA of total mouse embryos (E7, 11, 15 and 17) and adult organs. Two *Hira* transcripts (4.4 and 2.3 kb) are detected by probe mX in all samples. (**B**) Probes mB and mC, which exclude 5′ sequences, only detect the 4.4 kb transcript, indicating that the 2.3 kb transcript represents a 3′ truncated mRNA.

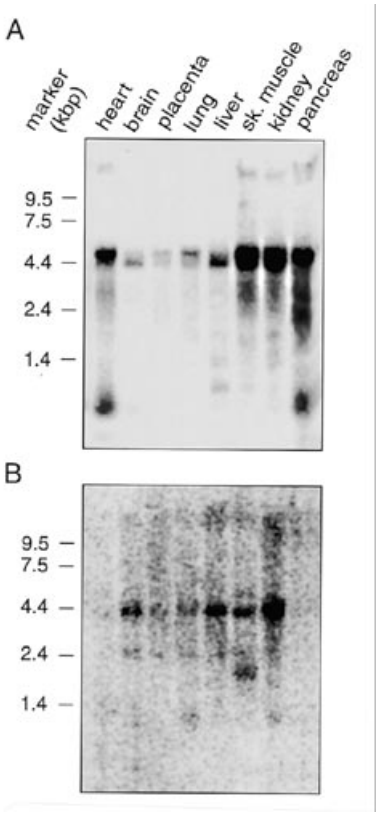


Figure 5. (**A**) Northern blot of poly(A)$^+$ RNA from adult human organs. Probe hE detects two transcripts (4.5 and 5.0 kb). Relative abundances vary from organ to organ. The 4.5 kb transcript is most abundant in heart and pancreas. (**B**) Mouse probe mC only detects the 4.5 kb transcript and an additional faint band of ~2.4 kb.

Control hybridizations with non-complementary (sense) probes are negative (Fig. 6C and F).

cDNA clones 12B and 12E represent splice variants with exons 3A and 11A

Two out of the seven *Hira* cDNAs (clones 12B and 12E) contain non-canonical sequences (Fig. 1A). To investigate whether these cDNAs represent valid transcripts, we partially characterized the genomic organization of the murine *Hira* gene. The intron–exon boundaries we determined coincide with those in the human as recently published by Lorain *et al.* (38). Below we will use their numbering of the exons.

Clone 12B contains a 203 bp insert between exons 3 and 4, which encode part of the first and second WD40 domain respectively (sequence in Fig. 7A). Genomically, the insert is continuous with exon 3 but is separated from exon 4 by an intron. Therefore, clone 12B must have arisen from the use of an alternative splice donor site. The ORF of clone 12B is interrupted by a frameshift, both in cDNA and genomic clones. The next potential translation initiation site is the methionine codon at nucleotide position 510, with a Kozak score of 0.67, or the methionine codon at nucleotide position 546 with a Kozak score of 0.78.

The 3′ terminus of clone 12E has an insert of 107 nucleotides between exons 11 and 12 (sequence in Fig. 7B). Exon 11 encodes the last WD40 domain. The 12E insert causes a frameshift in the ORF, with stopcodons as a result. PCR amplifications with primer sets p5–p8 and p6–p8 (Fig. 1B) on mouse genomic DNA, show that the 12E insert is located within an intron. With primers p5 and p8 we can amplify a 1.3 kb fragment, while with primers p6 and p8 we can amplify a 0.6 kb fragment. The insert is therefore separated from exon 11 and 12 by 0.5 and 0.33 kb introns, respectively.

Genomic sequences show that the 25 bp MF2 insert is positioned between exons 2 and 3. The extra sequences are designated by us as exons 2A (MF2), 3A (12B) and 11A (12E). BlastN and BlastX database searches with exons 3A and 11A do not yield any significant homologies.

Through RNase protection assays, we examined the presence of cellular mRNA species containing exon 3A. We detect full length protected fragments with probe mX (Figs 1B and 8B) in all RNA samples analyzed (Fig. 8A). This indicates the presence of canonical *Hira* transcripts. In addition, two smaller protected fragments are detected with this probe showing the existence of mRNA species that contain exon 3A. Full length protection was also found with probe mY (Fig. 1B) that contains exon 3A sequence (not shown). This also indicates the presence of

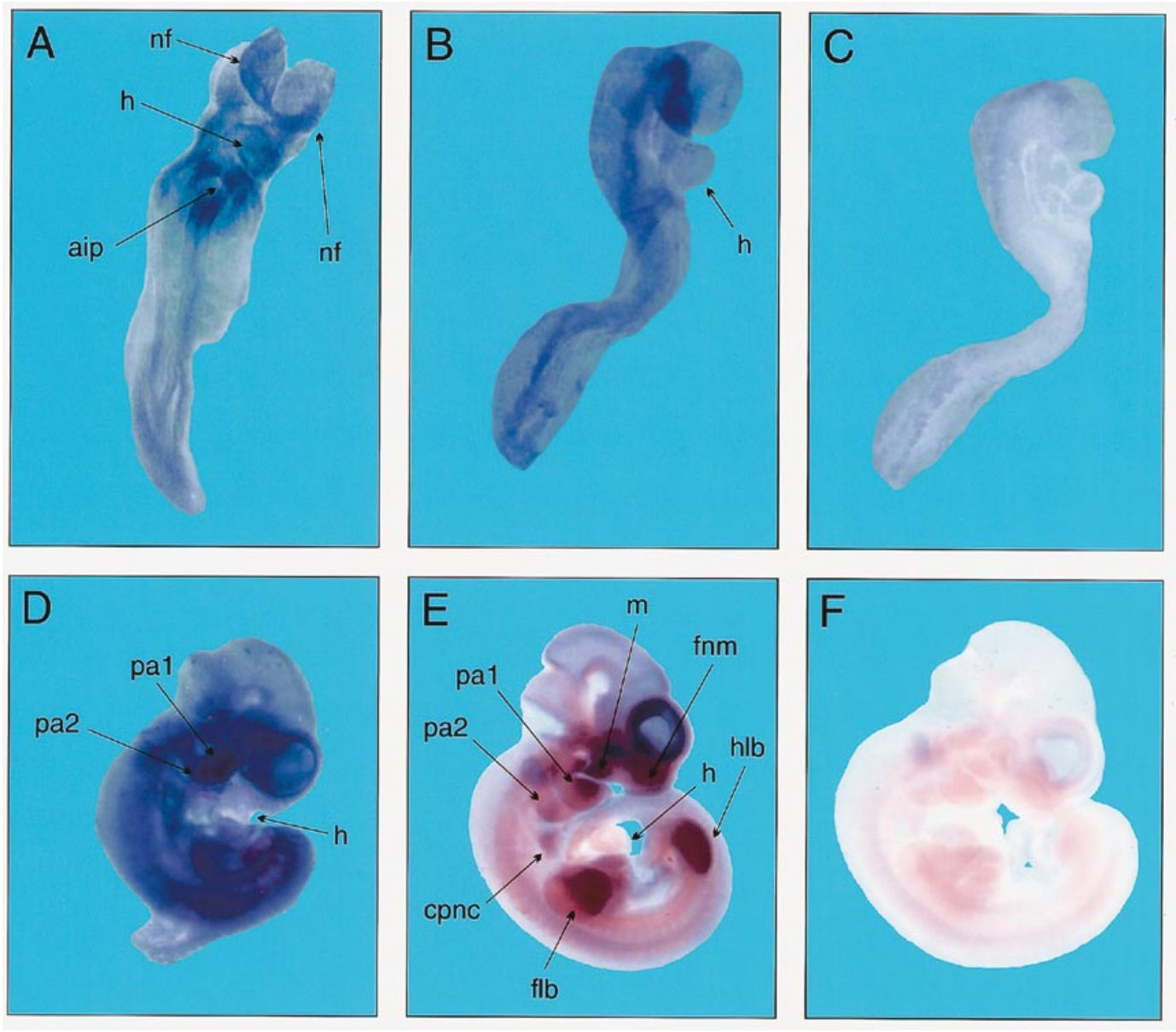


Figure 6. Whole mount *in situ* hybridization of mouse embryos with DIG labelled probe mA. (**A**) Ventral view of an E8.5 embryo. *Hira* transcripts are present ubiquitously, but higher levels are present in the neural folds. (**B**) Lateral view of an E8.5 embryo. The heart is positive for *Hira*. (**C**) E8.5 embryo hybridized with a non-complementary (sense) probe as a negative control. (**D**) Lateral view of an E9.5 embryo shows enhanced expression in the first two branchial arches. The heart is notably negative. (**E**) Lateral view of E10.5 embryos shows additional expression in limb buds, the circumpharyngeal neural crest, the frontonasal mass and the maxillary component of arch 1. (**F**) E10.5 embryo as a negative control. Abbreviations: aip, anterior intestinal portal; cpnc, circumpharyngeal neural crest; flb, forelimb bud; hlb, hindlimb bud; h, heart; m, maxillary component of pharyngeal arch 1; nf, neural folds; pa1, pharyngeal arch 1; pa2, pharyngeal arch 2. Lateral views are shown with the dorsal side of the embryo left and all views show the anterior end up.

transcripts containing exon 3A. None of the negative controls with sense probe or tRNA target yield any protected fragment.

To show the presence of mRNAs containing exon 11A we performed RT-PCR experiments with various primer sets on mouse embryo and adult organ RNAs. We used primer sets p5–p7, p5–p8, p6–p7 and p6–p8 (Fig. 1B) on poly-dT and random primed cDNA from total or poly(A)$^+$ RNA samples. With all four primer sets we can amplify fragments containing exon 11A sequences, both in adult organs and embryos. The presence of transcripts containing exon 3A was likewise confirmed by RT-PCR with primer sets p1–p3, p1–p4, p2–p3 and p2–p4 (Fig. 1B) on the same samples. Exon 3A is present in all samples (not shown).

DISCUSSION

Multiple alignment suggests the presence of three functional regions in Hira

We identified three regions in Hira through alignment with ceHILP, Hir1p, Hir2p and CAF1A. In addition, mouse/human HIRA alignments and the intron/exon structure of the *ceHILP*

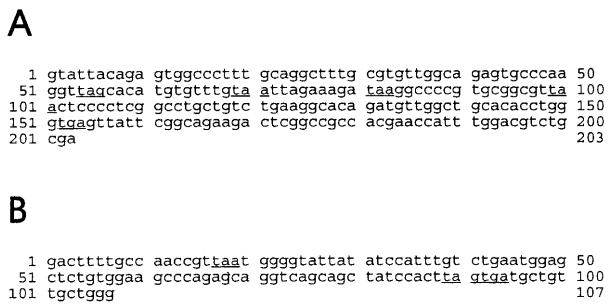


Figure 7. (**A**) Nucleotide sequence of exon 3A (from clone 12B). (**B**) Nucleotide sequence of exon 11A (from clone 12E). Stop codons in the *Hira* reading frame are underlined. Accession numbers: EMBL X99713 and X99722, respectively.

gene further support the presence of these distinct regions within Hira: mouse/human HIRA comparison showed higher homology in the N-terminal third (region A) than in the remainder of the protein; ceHILP exon–exon boundaries are spatially remarkably close to those of Hira. The major part of regions B and C are coded for by large single exons in ceHILP (39,40), suggesting that these regions represent single functional domains.

Region A contains WD40 domains that are present in a wide range of different proteins (41) and are involved in protein–protein interactions (42). In the case of yeast Tup1p, after which Hira was originally named, WD40 domains 1 and 2 interact directly with the homeodomain protein α2p (43).

The homology in regions B and C points to conserved domains although the functions have yet to be demonstrated. A clue to their functions might again come from Tup1p. Recently, Edmonson *et al.* reported that Tup1p binds histones H3 and H4 through a domain outside the WD40 region (44). Strikingly, the CAF complex also binds histones H3 and (acetylated) H4 and is thought to chaperone these core histones to the newly replicated DNA (34). In view of the homology of Hira with CAF1A it might well be that Hira performs an analogous function. Moreover, CAF1A function might be sensitive to gene dosage. *CAF1A* has been mapped to the Down syndrome critical region at 21q22 (45). The presence of three copies of CAF1A in Down syndrome individuals might contribute to the phenotype. CAF1A works in a complex with at least one other subunit (p150). Changes in gene-dosage might perturb the stoichiometry of the complex, hampering its formation. Hira, like CAF1A and Hir1p and Hir2p, is expected to be functional in a complex. Decreased availability of Hira due to haploinsufficiency of its gene may well have an effect on the formation of such a complex.

Hira expression in embryos

In this report we have shown that the *Hira* gene is expressed from as early as day E7 (gastrulation stage) through adulthood. Whole mount *in situ* hybridization shows that expression is ubiquitous, but that higher levels of transcripts are present in the cranial neural folds, subregions of pharyngeal arches 1 and 2, the circumpharyngeal neural crest and the limb buds. It is of interest that *Hira* transcripts are present in neural folds, where neural crest cells will be generated. In chicken embryos the zinc finger gene *Slug* is also expressed in the cranial folds. Treatment of embryos with *Slug* antisense oligos showed that Slug is necessary for neural crest cell migration (46). In mice, the homeobox gene *Msx-3* shows highly restricted expression that partially overlaps with *Hira* expression (47).

Most of the mesenchymal cells in the pharyngeal arches derive from hindbrain neural crest cells. It thus seems likely that the *Hira* expressing cells in pharyngeal arches 1 and 2 represent hindbrain neural crest cells. Special attention needs to be drawn to the regional differences within the arch mesenchyme. Certain subpopulations of cells either in the anterior part (arch 1) or in the posterior part (arch 2) seem to exist. In a recent study it was shown that *Bmp-7* has an overlapping distribution pattern (48). It will be interesting to study whether the various gene products (Slug, Msx-3, Bmp-7 and Hira) function in the same developmental pathway. Another group of *Hira* expressing cells that is evident consists of neural crest cells in the circumpharyngeal region. In birds, it has been demonstrated that these neural crest cells are crucial for proper development of the outflow tract of the heart. In mice, these cells express the proto-oncogene *Ret*, and part of this population of hindbrain neural crest cells is on its way to the developing gut or to the superior cervical ganglion (49).

In general, the elevated levels of *Hira* expression overlap with the expression domains of the cellular retinoic acid binding protein gene *Crabp-1*. *Crabp-1* is also expressed in the fronto-nasal mass, branchial arches 1 and 2 and the neural folds. Exposure of vertebrate embryos to retinoic acid leads to phenocopy of the CATCH22 phenotype (50). The overall *Hira* expression pattern seems to be conserved across species (61) .

Full-size *Hira* transcripts are present in the developing heart at E8.5, where they are no longer detectable at E9.5 and E10.5. The probe we used for the *in situ* experiments does not detect the short transcript, which is predominant in the adult heart. It is possible that between E8.5 and E9.5 the embryonic heart switches from the full-size to the 2.3 kb transcript. The developing heart arises through a complex series of morphogenetic interactions. In the mouse embryo the heart field is induced at E7.5. The straight heart tube is subsequently specified in an antero-posterior sequence (E9.5) to form the various regions and chambers of the looped (E8.5) and mature heart (E15) (51). The expression in the heart at E8.5 suggests that Hira might play a role in cardiac looping. This is of interest in the context of CATCH22, as the heart is one of the affected organs.

Splice variants of *Hira* mRNA

We find that two *Hira* mRNA species (2.3 and 4.4 kb) are present in embryos and adult organs. Both embryonically and at the adult stage (averaged over the individual organs) the 4.4 kb transcript is expressed at higher levels than the 2.3 kb transcript, without dramatic variations in the ratio between the two. Between adult organs the ratio of the expression of the 2.3 and 4.4 kb transcripts varies. The smaller 2.3 kb transcript is the predominant mRNA in the adult heart. Northern blot analysis reveals that this transcript lacks the sequences of the 3′ end of canonical *Hira*. This mRNA thus encodes a protein that consists of mainly WD40 domains. The 2.3 kb transcript could represent the transcript coding for the truncated 12E or MF2 coded polypeptides. The 4.4 kb transcript must represent the full length *Hira* transcript. Scamps *et al.* mention only one full length transcript in fetal mouse RNA (36). Most likely, they used 3′ probes.

Until now, two *HIRA* transcripts have only been reported for human fetal liver (26,27). Though these transcripts were labelled 3.4 and 3.2 kb, and 4.5 and 4.2 kb we assume that these are

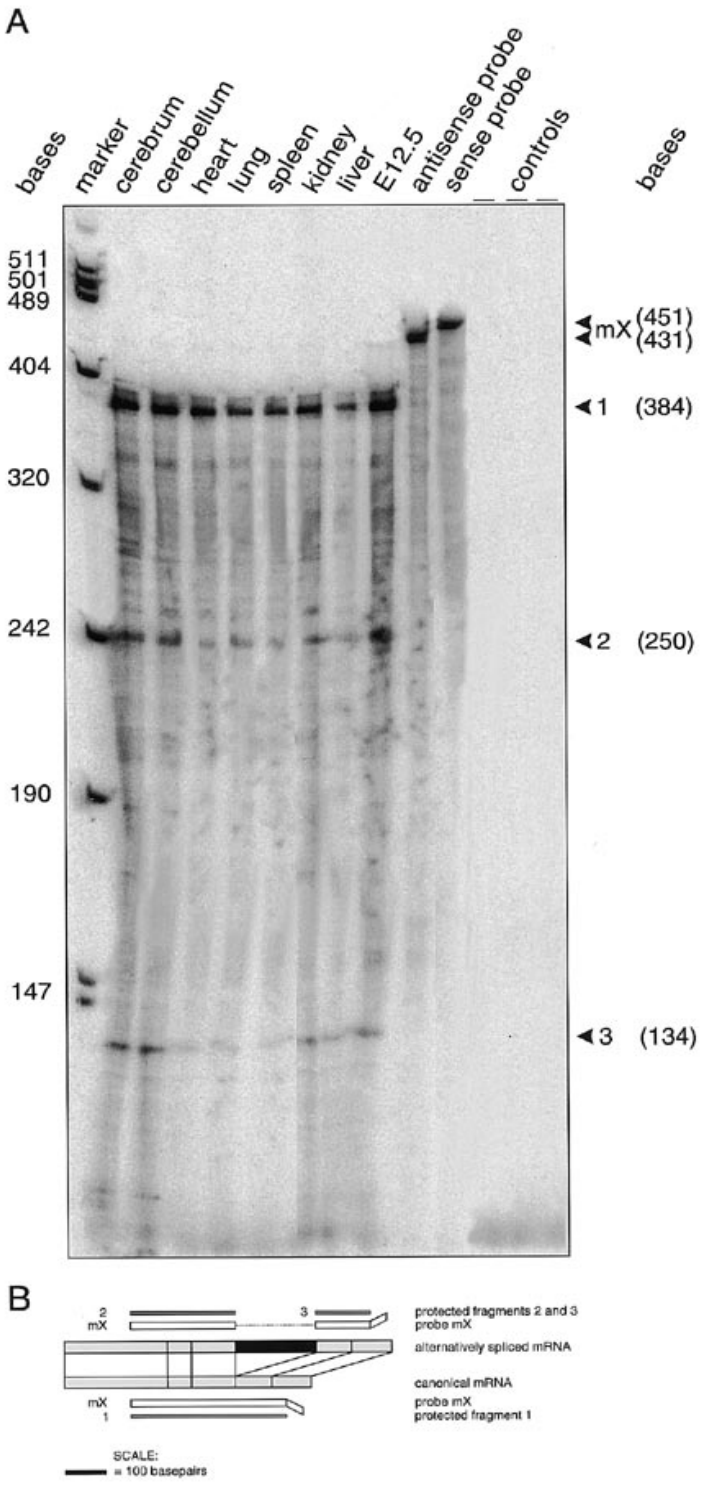


Figure 8. (**A**) Protected RNA fragments resulting from an RNase protection assay with probe mX and the target RNAs shown. Controls were RNAs hybridized to non-complementary (sense) RNA probes or tRNA as a target for antisense probes. Arrowheads point to probe mX (sense and antisense) and the protected fragments: (1) full length protected fragment representing the canonical transcripts; (2) and (3) indicate the presence of transcripts with exon 3A. (**B**) Schematic drawing of probe mX and the relevant fragments of two possible target transcripts. Gray boxes indicate canonical exons, the black box indicates exon 3A. The upper and lower lines indicate the expected size of protected fragments and their corresponding numbers in (A).

identical to our 5.0 and 4.5 kb adult transcripts. In the adult heart and pancreas the 5.0 kb transcript is predominant, whereas the 4.5 kb transcript is predominant in brain. Like the 2.3 kb transcript in the mouse, the 5.0 kb is not detected by a *Hira* probe that excludes WD40 repeat coding regions. One explanation could be that the 5.0 kb transcript represents an RNA that encodes a HIRA like protein with C-terminal WD40 domains. This protein would have an N-terminal extension, which could be similar to the N-terminal

extension of ceHILP. This part of ceHILP, which has no equivalent in HIRA, CAF1A, Hir1p or Hir2p, bears very high homology to other *C.elegans* proteins that are similar to *Schizosaccharomyces pombe* SDS22 and *L.monocytogenes* Internalin (52,53). The yeast SDS22 positively modulates protein phosphatase 1 (54). Internalin is required for the internalization of *L.monocytogenes* by eukaryotic cells (55). Alternatively, the 5.0 kb transcript encodes a protein with N-terminal WD40 repeats followed by hitherto unknown sequences not present in canonical HIRA.

The insertion of exon 3A into the *Hira* mRNA disrupts the canonical *Hira* open reading frame. An N-terminally truncated protein starting from either of the next two in-frame translation start sites lacks the first WD40 domain. This is comparable with the protein coded for by the MF2 *Tuple1* (*Hira*) cDNA, which lacks half of the first WD40 domain due to a 25 bp insert in the mRNA. In Tup1p the array of WD40 domains appears to influence the binding specificity of the individual domains. In the context of other WD40 domains, a single domain has different binding specificities than in isolation (43). The lack of a WD40 domain is therefore very likely to alter the affinity of Hira for other proteins. The insertion exon 11A introduces stop codons, yielding a polypeptide consisting only of WD40 domains. This truncated polypeptide could compete with full length Hira for binding to other proteins. This putative truncated form of Hira resembles CAF1A which also consists mainly of WD40 repeats.

In conclusion, at least four different splice forms of *HIRA* can be identified in humans and mice. This complex expression pattern of the mammalian *HIRA* gene is regulated in an organ- and developmental-specific manner and arises from alternative splicing and exon skipping.

HIRA is being analyzed for point mutations and microdeletions by several groups that are looking for a cause for the CATCH22 phenotype in apparently non-deleted CATCH22 patients. The new exons 2A, 3A and 11A should be included in these screens. Further analyses of the human 5.0 kb transcript should help in identifying more *HIRA* sequences to check for mutations.

Molecular genetics of CATCH22

In the absence of direct evidence that the CATCH22 phenotype is the result of a single gene defect, it is most likely that the phenotype is the result of haploinsufficiency of several genes with major effect. The spatio-temporal expression pattern of *HIRA* renders the *HIRA* gene one of the genes of which haploinsufficiency could lead to part of the CATCH22 phenotype. Since deletion screens in CATCH22 patients have so far not reduced the SRDO to less than 300 kb, it will be necessary to analyze the function of genes with major effect in mice. Therefore, we are analyzing the effect on development of a null mutation of the murine *Hira* gene.

One has to bear in mind, that it will also be important to identify genes with major effect in SRDO2. Furthermore, position effects cannot be excluded. In campomelic dysplasia and aniridia, *SOX9* and *PAX6*, respectively, can be influenced by translocation breakpoints up to 250 kb distant from the disease causing gene (56,57). A similar effect could play a role in CATCH22. It is also possible that as a result of the CATCH22 deletion a heterochromatic region (e.g. close to the centromere) can exert its influence onto former distal regions, thereby silencing the genes in these regions. In this model it is even possible that the genes of major effect are located outside the commonly deleted region. Since there appear to be at least two SRDOs, one cannot exclude that both position effects and gene deletions can lead to the CATCH22 phenotype. Considering the fact that there are many phenocopies of the CATCH22 phenotype, deletions or silencing of different genes on 22q11 could lead to similar phenotypes.

MATERIALS AND METHODS

Embryos

Embryos were derived from crosses between FVB/N mice. The morning of vaginal plug was considered 0.5 days *post coitum* (dpc). The embryos were stages according to the number of developmental day (E8.5–E12.5). The E8.5 embryos were staged under the dissecting microscope by counting somites.

cDNA cloning, sequencing and database searches

A mouse embryo 11.5 dpc cDNA library was purchased from Clontech. Plaques were lifted onto nitrocellulose membranes (Millipore HATF filters) and hybridized as recommended by the manufacturer. The cDNA clones were sequenced by primer walking using the Pharmacia T7 sequencing kit. Sequences were assembled, aligned and compared using the GCG package at the CAOS/CAMM server in Nijmegen. Database searches were performed using Blast programs (58). Multiple sequence alignments were edited manually with the GeneDoc program and printed with BoxShade.

Whole mount *in situ* hybridization

Complementary (antisense) or non-complementary (sense) RNA probes were synthesized from linearized DNA templates using the DIG-UTP labelling kit (Boehringer Mannheim). *In situ* hybridization on whole mouse embryos (E8.5–E11.5) was performed according to the protocol of Wilkinson *et al.* (59), with posthybridization washes modified as follows: 50 and 70 min in solution 1 at 70°C; 20 min in solution 1 and 2 (1:1) at room temperature; 2× 20 min in solution 2 at room temperature; 2× 30 min in solution 2 with RNase A at 37°C; 10 min in solution 2 at room temperature; 2× 30 min in solution 3 at 65°C; 2× 30 min in TBST at room temperature. Adsorption of anti-digoxigenin antibodies to embryo powder was omitted. Embryos were either captured on film and subsequently scanned, or captured electronically with an electronic camera. Images were processed using Adobe Photoshop software.

RNA isolation

Adult organs or embryos were homogenized in GIT or LiCl urea with a tissue homogenizer. Total RNA was purified by ultra centrifugation on a cesium-chloride cushion or by LiCl precipitation and organic extractions. Poly(A)$^+$ RNA was isolated from total RNA with the PolyATtract mRNA isolation system III (Promega).

Southern and Northern analyses

A mouse embryo and a human multiple tissue Northern blot, together with a Zoo blot were purchased from Clontech (Palo Alto, CA). Blots were hybridized as per manufacturer's protocol. The mouse multiple tissue Northern blot was produced by running 5 μg poly(A)$^+$ RNA per lane on a denaturing formalde-

hyde agarose gel, blotting in 20× SSC onto Qiabrene^{+} nylon membrane (Qiagen), and UV cross-linking in a Stratalinker (Stratagene).

RNase protection assay

Complementary (antisense) or non-complementary (sense) [α-^{32}P]UTP labelled RNA probes were synthesized from linearized *Hira* cDNA templates, gel-eluted, hybridized to total RNA at 50°C and subjected to RNase protection assay according to ref. 60.

RT-PCR detection of alternative *Hira* mRNA

RT-PCR was performed on 5 µg total RNA or 1 µg poly(A)$^{+}$ RNA (from E8.5, E9.5, E10.5 and E14.5 mouse embryos and adult mouse brain, spleen, liver and lung) with a mix of 300 ng each of oligo-dT and random hexamer primers according to established protocols. The RT reaction product was used for the following PCR amplification: one cycle of 2 min at 94°C, 2 min at 50°C, 2 min at 72°C; 30 cycles of 1 min at 94°C, 1 min at 50°C, 1.5 min at 72°C and one cycle of 10 min at 72°C. Primer sequences: p1, ctgggaaggttgtgatctgg; p2, ggtattacagagtggccctt; p3, tcgcagacgtccaaatggtt; p4, ccaatgtacgtagcccgctt; p5, atactgctgctgtgctgttg; p6, tttgccaaccgttaatgggg; p7, cccagcaacagcatcactaa; p8, atgctgagcttgtctcccta. Positive controls were 1 pg of cDNAs 12B, E and F. Negative controls were omission of cDNA and a sample from an RT reaction without RNA.

ACKNOWLEDGEMENTS

We thank Drs Frans Lohman, Vincent Lui and Maarten Mulder for comments on the manuscript. We thank Peter Scambler for his gift of the human C5 cDNA and sharing unpublished results. Prof. Jan C. Molenaar is acknowledged for his continuous support of our work. This study was financially supported by the Erasmus University, the Sophia Foundation for Medical Research, the European Community and the Dutch Heart Foundation.

ABBREVIATIONS

E, embryonic age in days *post coitum*; SRDO, smallest region of deletion overlap.

REFERENCES

1. Wilson,D.I., Burn,J., Scambler,P. and Goodship,J. (1993) DiGeorge syndrome: part of CATCH22. *J. Med. Genet.*, **30**, 852–856.
2. de la Chapelle,A., Herva,R., Koivisto,M. and Aula,P. (1981) A deletion in chromosome 22 can cause DiGeorge syndrome. *Hum. Genet.*, **57**, 253–256.
3. Carey,A.H., Roach,S., Williamson,R., Dumanski,J.P., Nordenskjold,M., Collins,V.P., Rouleau,G., Blin,N., Jalbert,P. and Scambler,P.J. (1990) Localization of 27 DNA markers to the region of human chromosome 22q11–pter deleted in patients with the DiGeorge syndrome and duplicated in the der22 syndrome. *Genomics*, **7**, 299–306.
4. Driscoll,D.A., Budarf,M.L. and Emanuel,B.S. (1992) A genetic etiology for DiGeorge syndrome: consistent deletions and microdeletions of 22q11. *Am. J. Hum. Genet.*, **50**, 924–933.
5. Scambler,P.J., Kelly,D., Lindsay,E., Williamson,R., Goldberg,R., Shprintzen,R., Wilson,D.I., Goodship,J.A., Cross,I.E. and Burn,J. (1992) Velo-cardio-facial syndrome associated with chromosome 22 deletions encompassing the DiGeorge locus. *Lancet*, **339**, 1138–1139.
6. Driscoll,D.A., Spinner,N.B., Budarf,M.L., McDonald-McGinn,D.M., Zackai,E.H., Goldberg,R.B., Shprintzen,R.J., Saal,H.M., Zonana,J., Jones,M.C., Mascarello,J.T. and Emanuel,B.S. (1992) Deletions and microdeletions of 22q11.2 in velo-cardio-facial syndrome. *Am. J. Med. Genet.*, **44**, 261–268.
7. Burn,J., Takao,A., Wilson,D., Cross,I., Momma,K., Wadey,R., Scambler,P. and Goodship,J. (1993) Conotruncal anomaly face syndrome is associated with a deletion within chromosome 22q11. *J. Med. Genet.*, **30**, 822–824.
8. McDonald-McGinn,D.M., Driscoll,D.A., Bason,L., Christensen,K., Lynch,D., Sullivan,K., Canning,D., Zavod,W., Quinn,N., Rome,J., Paris,Y., Weinberg,P., Clark III,B.J., Emanuel,B.S. and Zackai,E.H. (1995) Autosomal dominant 'Opitz' GBBB syndrome due to a 22q11.2 deletion. *Am. J. Med. Genet.*, **59**, 103–113.
9. Robin,N.H., Feldman,G.J., Aronson,A.L., Mitchell,H.F., Weksberg,R., Leonard,C.O., Burton,B.K., Josephson,K.D., Laxova,R., Aleck,K.A., Allanson,J.E., Guionalmeida,M.L., Martin,R.A., Leichtman,L.G., Price,R.A., Opitz,J.M. and Muenke.M. (1995) Opitz syndrome is genetically heterogenous with one locus on Xp22, and a second locus on 22q11.2. *Nature Genet.*, **11**, 459–461.
10. Fibison,W.J., Budarf,M.L., McDermid,H., Greenberg,F. and Emanuel,B.S. (1990) Molecular studies of DiGeorge syndrome. *Am. J. Hum. Genet.*, **46**, 888–895.
11. Lindsay,E.A., Halford,S., Wadey,R., Scambler,P.J. and Baldini,A. (1993) Molecular cytogenetic characterization of the DiGeorge syndrome region using fluorescence *in situ* hybridization. *Genomics*, **17**, 403–407.
12. Aubry,M., Demczuk,S., Desmaze,C., Aikem,M., Aurias,A., Julien,J.P. and Rouleau,G. (1993) Isolation of a zinc finger gene consistently deleted in DiGeorge syndrome. *Hum. Mol. Genet.*, **2**, 1583–1587.
13. Kurahashi,H., Akagi,K., Karakawa,K., Nakamura,T., Dumanski,J.P., Sano,T., Okada,S., Takai,S.I. and Nishisho,I. (1994) Isolation and mapping of cosmid markers on human chromosome 22, including one within the submicroscopically deleted region of DiGeorge syndrome. *Hum. Genet.*, **93**, 248–254.
14. Mulder,M.P., Wilke,M., Langeveld,A., Wilming,L.G., Hagemeijer,A., van Drunen,E., Zwarthoff,E.C., Riegman,P.J., Deelen,W.H., van den Ouweland,A.M.W., Halley,D.J.J. and Meijers,C. (1995) Positional mapping of loci in the DiGeorge critical region at chromosome 22q11 using a new marker (D22S183). *Hum. Genet.*, **96**, 133–141.
15. Morrow,B., Goldberg,R., Carlson,C., Gupta,R.D., Sirotkin,H., Collins,J., Dunham,I., O.Donnell,H., Scambler,P., Shprintzen,R. and Kucherlapati,R. (1995) Molecular definition of the 22q11 deletions in velo-cardio-facial syndrome. *Am. J. Hum. Genet.*, **56**, 1391–1403.
16. Levy,A., Demczuk,S., Aurias,A., Depetris,D., Mattei,M.G. and Philip,N. (1995) Interstitial 22q11 microdeletion excluding the ADU breakpoint in a patient with the DiGeorge syndrome. *Hum. Mol. Genet.*, **4**, 2417–2419.
17. Kurahashi,H., Nakayama,T., Osugi,Y., Tsuda,E., Masuno,M., Imaizumi,K., Kamiya,T., Sano,T., Okada,S. and Nishisho,I. (1996) Deletion mapping of 22q11 in CATCH22 syndrome: identification of a second critical region. *Am. J. Hum. Genet.*, **58**, 1377–1381.
18. Dallapiccola,B., Pizzuti,A. and Novelli,G. (1996) How many breaks do we need to CATCH on 22q11? *Am. J. Hum. Genet.*, **59**, 7–11.
19. Fu,W.N., Borgaonkar,D.S., Ledewig,P.P., Weaver,J. and Pomerance,H.H. (1976) Structural aberrations of the long arm of chromosome no. 22. *Clin. Genet.*, **10**, 329–336.
20. Augusseau,S., Jouk,S., Jalbert,P. and Prieur,M. (1986) DiGeorge syndrome and 22q11 rearrangements. *Hum. Genet.*, **74**, 206.
21. Budarf,M.L., Collins,J., Gong,W., Roe,B., Wang,Z., Bailey,L.C., Sellinger,B., Michaud,D., Driscoll,D.A. and Emanuel,B.S. (1995) Cloning a balanced translocation associated with DiGeorge syndrome and identification of a disrupted candidate gene. *Nature Genet.*, **10**, 269–277.
22. Sutherland,H.F., Wadey,R., McKie,J.M., Taylor,C., Atif,U., Johnstone,K.A., Halford,S., Kim,U.J., Goodship,J., Baldini,A. and Scambler,J. (1996) Identification of a novel transcript disrupted by a balanced translocation associated with DiGeorge syndrome. *Am. J. Hum. Genet.*, **59**, 23–31.
23. Collins,J.E., Cole,C.G., Smink,L.J., Garett,C.L., Leversha,M.A., Soderlund,C.A., Maslen,G.L., Everett,L.A., Rice,K.M., Coffey,A.J., Gregory,S.G., William,R., Dunham,A., Davies,A.F., Hassock,S., Todd,C.M., Lehrach,H., Hulsebos,T.J.M., Weissenbach,J., Morrow,B., Kucherlapati,R.S., Wadey,R., Scambler,P.J., Kim,U.J., Simon,M.I., Peyrard,M., Xie,Y.G., Carter,N.P., Durbin,R., Dumanski,J.P., Bentley,D.R. and Dunham,I. (1995) A high-density YAC contig map of human chromosome 22. *Nature*, **377** Suppl., 367–379.
24. Gong,W., Emanuel,B.S., Collins,J., Kim,D.H., Wang,Z., Chen,F., Zhang,G., Roe,B. and Budarf,M.L. (1996) A transcriptional map of the DiGeorge and velo-cardio-facial syndrome minimal critical region on 22q11. *Hum. Mol. Genet.*, **5**, 789–800.
25. Lindsay,E.A., Rizzu,P., Antonacci,R., Juecic,V., Delmas-Mata,J., Lee,C.C., Kim,U.J., Scambler,P.S. and Baldini,A. (1996) A transcription map in the CATCH22 critical region: identification, mapping, and ordering of four novel transcripts expressed in heart. *Genomics*, **32**, 104–112.

26. Halford,S., Wadey,R., Roberts,C., Daw,S.C.M., Whiting,J.A., O'Donnell,H., Dunham,I., Bentley,D., Lindsay,E., Baldini,A., Francis,F., Lehrach,H., Williamson,R., Wilson,D.I., Goodship,J., Cross,I., Burn,J. and Scambler,P.J. (1993) Isolation of a putative transcriptional regulator from the region of 22q11 deleted in DiGeorge syndrome, Sphrintzen syndrome and familial congenital heart disease. *Hum. Mol. Genet.*, **2**, 2099–2107
27. Lamour,V., Lécluse,Y., Desmaze,C., Spector,M., Bodescot,M., Aurias,A., Osley,M.A. and Lipinski,M. (1995) A human homolog of the *S. cerevisiae* HIR1 and HIR2 transcriptional repressors cloned from the DiGeorge syndrome critical region. *Hum. Mol. Genet.*, **4**, 791–799.
28. Sirotkin,H., Morrow,B., Dasgupta,R., Goldberg,R., Patanjali,S.R., Shi,G.P., Cannizzaro,L., Shprintzen,R., Weissman,S.M. and Kucherlapati,R. (1996) Isolation of a new clathrin heavy chain gene with muscle-specific expression from the region commonly deleted in velo-cardio-facial syndrome. *Hum. Mol. Genet.*, **5**, 617–624.
29. Kedra,D., Peyrard,M., Fransson,I., Collins,J.E., Dunham,I., Roe,B.A. and Dumanski,P.J. (1996) Characterization of a second human clathrin heavy chain polypeptide gene (CLH-22) from chromosome 22q11. *Hum. Mol. Genet.*, **5**, 625–631.
30. Heisterkamp,N., Mulder,M.P., Langeveld,A., Ten Hoeve,J., Wang,Z., Roe,B.A. and Groffen,J. (1995) Localization of the human mitochondrial citrate transporter protein gene to chromosome 22q11 in the DiGeorge critical region. *Genomics*, **29**, 451–456.
31. Pizzuti,A., Novelli,G., Mari,A., Ratti,A., Colosimo,A., Amati,F., Penso,D., Sangiuolo,F., Calabrese,G., Palka,G., Salani,V., Gennarelli,M., Mingarelli,R., Scarlatto,G., Scambler,P. and Dallapiccola,B. (1996) Human homologue sequences to the *Drosophila dishevelled* segment polarity gene are deleted in the DiGeorge syndrome. *Am. J. Hum. Genet.*, **58**, 722–729.
32. Sherwood,P.W., Tang,S.V. and Osley,M.A. (1993) Characterization of *HIR1* and *HIR2*, two genes required for regulation of histone gene transcription in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **13**, 28–38.
33. Promisel Cooper,J., Roth,S. and Simpson,R.T. (1994) The global transcriptional regulators SSN6 and TUP1, play distinct roles in the establishment of a repressive chromatin structure. *Genes Dev.*, **8**, 1400–1410.
34. Kaufman,P.D., Kobayashi,R., Kessler,N. and Stillman,B. (1995) The p150 and p60 subunits of chromatin assembly factor I: a molecular link between newly synthesized histones and DNA replication. *Cell*, **81**, 1105–1114.
35. Scamps,C., Lorain,S., Lamour,V. and Lipinski,M. (1996) The HIR protein family: isolation and characterization of a complete murine cDNA. *Biochim. Biophys. Acta*, **1306**, 5–8.
36. Miller,N. and Waterston,R. (PID g687860)
37. Kuratani,S.C. and Kirby,M.L. (1991) Initial migration and distribution of the cardiac neural crest in the avian embryo: an introduction to the concept of the circumpharyngeal crest. *Am. J. Anat.*, **191**, 215–227.
38. Lorain,S., Demczuk,S., Lamour,V., Toth,S., Aurias,A., Roe,B.A. and Lipinski,M. (1996) Structural organization of the WD40 repeat protein-encoding gene HIRA in the DiGeorge syndrome critical region of human chromosome 22. *Genome Res.*, **6**, 43–50.
39. Miller,N. and Waterston,R. (U21322).
40. Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fraser,A., Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M., Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laisster,N., Latreille,P., Lightning,J., Lloyd,C., Mortimore,B., O'Callaghan,M., Parson,J., Percy,C., Rifken,L., Roopra,A., Saunders,D., Shownkeen,R., Sims,M., Smaldon,N., Smith,A., Smith,M., Sonnhammer,E., Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M., Vaughan,K., Waterston,R., Watson,A., Weinstock,L., Wilkinson-Sproat,J. and Wohldman,P. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans. *Nature*, **368**, 32–38.
41. Neer,E.J., Schmidt,C.J., Nambudripad,R. and Smith,T.F. (1994) The ancient regulatory-protein family of WD-repeat proteins. *Nature*, **371**, 297–300.
42. Sondek,J., Bohm,A., Lambright,D.G., Hamm,H.E. and Sigler,P.B. (1996) Crystal structure of a G_A protein ßγ dimer at 2.1 Å resolution. *Nature*, **379**, 369–374.
43. Komachi,K., Redd,M.J. and Johnson,A.D. (1994) The WD repeats of Tup1 interact with the homeo domain protein α2. *Genes Dev.*, **8**, 2857–2867.
44. Edmondson,D.G., Smith,M.M. and Roth,S.Y. (1996) Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes Dev.*, **10**, 1247–1259.
45. Blouin,J.L., Duriaux-Sail,G., Chen,H., Gos,A., Morris,M.A., Rossier,C. and Antonorakis,S.E. (1996) Mapping of the gene for the p60 subunit of the human chromatin assembly factor (CAF1A) to the Down syndrome region of chromosome 21. *Genomics*, **33**, 309–312.
46. Nieto,M.A., Sargent,M.G., Wilkinson,D.G. and Cooke,J. (1994) Control of cell behavior during vertebrate development by *Slug*, a zinc finger gene. *Science*, **264**, 835–839.
47. Shimeld,S.M., McKay,I.J. and Sharpe,P.T. (1996) The murine homeobox gene *Msx-3* shows highly restricted expression in the developing neural tube. *Mech. Dev.*, **55**, 201–210.
48. Wall,N.A. and Hogan,B.L.M. (1995) Expression of bone morphogenetic protein-4 (BMP-4), bone morphogenetic protein-7 (BMP-7), fibroblast growth factor-8 (FGF-8) and sonic hedgehog (SHH) during branchial arch development in the chick. *Mech. Dev.*, **53**, 383–392.
49. Durbec,P.L., Larsson-Blomberg,L.B., Schuchard,A., Constantini,F. and Pachnis,V. (1996) Common origin and developmental dependance on *c-ret* of subsets of enteric and sympathetic neuroblasts. *Development*, **122**, 349–358.
50. Vaessen,M.J., Meijers,J.H.C., Bootsma,D. and Geurts van Kessel,A. (1990) The cellular retinoic acid-binding protein is expressed in tissues associated with retinoic-acid-induced malformations. *Development*, **110**, 371–378.
51. Kaufman,M.H. (1992) *The atlas of mouse development*, Academic Press, London
52. Leimbach,D. and Waterston,R. (Genbank U39849)
53. Lightning,J. (P45969)
54. Ohkura,H. and Yanagida,M. (1991) *S. pombe* gene SDS22+ essential for a midmitotic transition encodes a leucine-rich repeat protein that positively modulates protein phosphatase-1. *Cell*, **64**, 149–157.
55. Mengaud,J., Ohayon,H., Gounon,P., Mege,R.M. and Cossart,P. (1996) E-cadherin is the receptor for internalin, a surface protein required for entry of L. monocytogenes into epithelial cells. *Cell*, **84**, 923–932.
56. Wirth,J., Wagner,T., Meyer,J., Pfeiffer,R.A., Tietze,H.U., Schempp,W. and Scherer,G. (1996) Translocation breakpoints in three patients with campomelic dysplasia and autosomal sex reversal map more than 130 kb from Sox9. *Hum. Genet.*, **97**, 186–193.
57. Fantes,J., Redeker,B., Breen,M., Boyle,S., Brown,J., Fletcher,J., Jones,S., Bickmore,W., Fukushima,Y., Mannens,M., Danes,S., van Heyningen,V. and Hanson,I. (1995) Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum. Mol. Genet.*, **4**, 415–422.
58. Altschul,S.F., Gish,W., Miller,W., Meyers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
59. Wilkinson,D.G. (1992) in Wilkinson,D.G. (ed.), *In Situ Hybridization: A Practical Approach*. IRL Press, Oxford, pp. 75–83.
60. Gilman,M. (1987) in Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds), *Current Protocols in Molecular Biology*. Wiley Interscience, New York, Vol. I, pp.4.7.1.
61. Roberts, C., Daw, S.C.M., Halford, S. and Scambler, P.J. (1997) Cloning and developmental expression analysis of chick *Hira (Chira)*, a candidate gene for DiGeorge syndrome. *Hum. Mol. Genet.* **6**, 235–243.

NOTE ADDED IN PROOF

One of the reviewers pointed to the fact that patients and their families often find the use of the term CATCH22 offensive. We propose to use the acronym in scientific communications only. For clinical purposes the various syndromes could be referred to collectively as the 'chromosome 22q11 deletion syndrome'.