

# Expression and phylogeny of claudins in vertebrate primordia

Richard Kollmar, Shashi Karia Nakamura, James A. Kappler, and A. J. Hudspeth\*

Laboratory of Sensory Neuroscience and Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10021-6399

Contributed by A. J. Hudspeth, June 27, 2001

**Claudins, the major transmembrane proteins of tight junctions, are members of the tetraspanin superfamily of proteins that mediate cellular adhesion and migration. Their functional importance is demonstrated by mutations in claudin genes that eliminate tight junctions in myelin and the testis, abolish  $Mg^{2+}$  resorption in the kidney, and cause autosomal recessive deafness. Here we report that two paralogs among 15 claudin genes in the zebrafish, *Danio rerio*, are expressed in the otic and lateral-line placodes at their earliest stages of development. Related claudins in amphibians and mammals are expressed in a similar manner in vertebrate primordia such as sensory placodes, branchial arches, and limb buds. We also show that the claudin gene family may have expanded along the chordate stem lineage from urochordates to gnathostomes, in parallel with the elaboration of vertebrate characters. We propose that tight junctions not only form barriers in mature epithelia, but also participate in vertebrate morphogenesis.**

Compartments bounded by epithelia are a fundamental feature of any metazoan. Compartmentalization overcomes the limitations imposed by diffusion and facilitates the control of cellular environments, the division of labor among differentiated cell types, and the efficient transport of nutrients and information over long distances. To form effective diffusional barriers, epithelia use two types of intercellular seals, tight junctions and septate junctions (1, 2). Both contain matching strands of intramembrane particles in the apposed cell membranes. Both are selective and dynamic rather than impermeable and static barriers. Both are linked by intracellular membrane-associated proteins to signaling pathways and the cytoskeleton. The two types of junctions are distinguished by their ultrastructure and their distribution among metazoan phyla. At tight junctions the intercellular space is wholly obliterated, whereas at septate junctions the apposed cell membranes are separated by distinct septa and remain 12–18 nm apart. Tight junctions are in general characteristic of chordates, whereas septate junctions are found in nonchordates (1). However, tight junctions also have been found in the blood–brain and blood–testis barriers of adult arthropods (3) and septate junctions occur at the nodes of Ranvier in the vertebrate nervous system (4).

The intercellular barrier in mammalian tight junctions is formed by claudin proteins (2), members of the tetraspanin superfamily of integral membrane proteins that mediate cellular adhesion and migration (5). The mammalian claudin family comprises at least 20 members (6). Mutations in claudin genes eliminate tight junctions in myelin and the testis (7), abolish  $Mg^{2+}$  resorption in the kidney (8), and cause autosomal recessive deafness (9). Claudins by themselves can form the strands of membrane particles that are characteristic of the vertebrate tight junction (10). For septate junctions and arthropod tight junctions, in contrast, the molecular nature of the barrier remains unknown (4).

Compartmentalization is not only a hallmark of metazoans, but also may have facilitated the increase in diversity and complexity during their evolution (11, 12). Compartments in a broad sense—of body plans, genomes, and developmental programs—allowed variation of individual modules while protecting others from the potentially deleterious effects of mutations.

During the evolution of vertebrates, for example, the chordate body plan was retained while novel structures such as the cranium, paired sensory organs, and paired appendages were added (13, 14). This increase in anatomical complexity likely reflected an increase in genetic complexity as many gene families expanded by retrotransposition, tandem duplication, and duplication of whole genomes (15, 16). The redundancy of the duplicated genes presumably allowed them to evolve independently and to acquire novel functions or perform old functions in novel combinations (14, 17). During the embryonic development of chordates, members of all subphyla have been postulated to pass after gastrulation through a common phylotypic period (11, 18). Only then do the primordia of characteristic vertebrate structures become apparent and only thereafter do features specific to the various orders, classes, and families develop.

Vertebrate tight junctions had been until recently investigated in detail only in adults or during early cleavage of the embryo (19). During a cDNA subtraction screen in the zebrafish, *Danio rerio*, we identified two claudin genes that are expressed after gastrulation with exquisite specificity in sensory placodes and the pronephros. This expression during the phylotypic period in characteristic vertebrate structures and the importance of tight junctions for compartmentalization suggested a connection between claudins and vertebrate evolution and ontogeny. We therefore investigated the phylogeny of claudins and their expression in vertebrate primordia.

## Materials and Methods

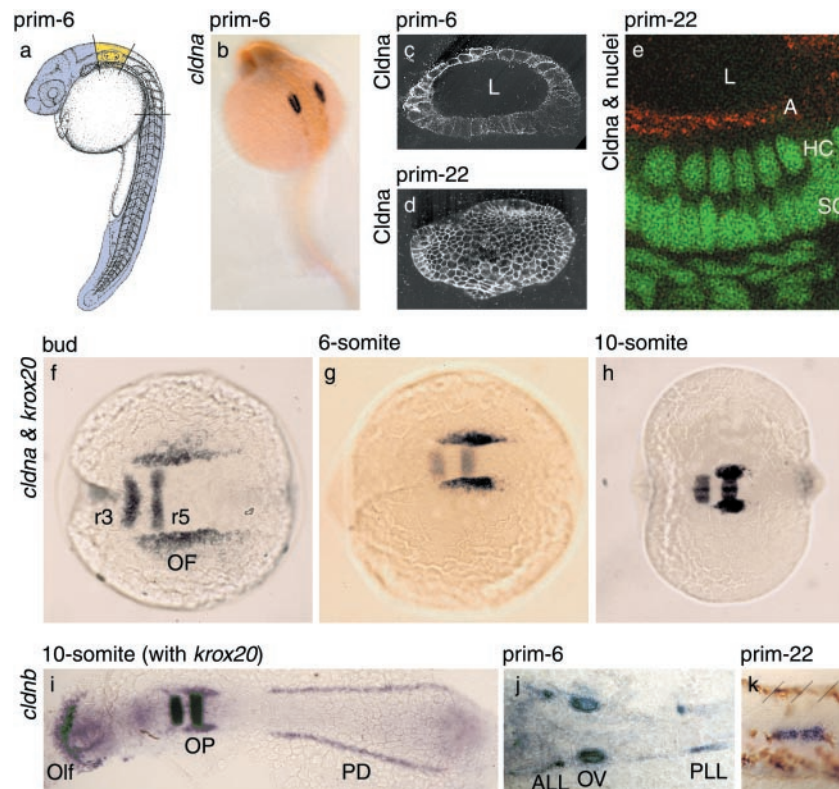
**Subtractive cDNA Hybridization.** Zebrafish embryos at about prim-6 stage were cut transversely with ultra-fine scissors (Fine Science Tools, Foster City, CA) anterior and posterior to the otic vesicle and at the seventh somite (Fig. 1*a*). The hindbrain region containing the otic vesicles was removed (“ear” sample), as were the anterior part of the head, the trunk, and the tail (“head-and-tail” sample). From 150 embryos, 2  $\mu$ g of total ear RNA and 70  $\mu$ g of total head-and-tail RNA were obtained by purification with Trizol (Life Technologies, Rockville, MD), digestion with DNaseI, and centrifugation through a cesium trifluoroacetate cushion. One microgram from each sample was used for cDNA synthesis and limited amplification with a Smart PCR cDNA synthesis kit followed by subtractions with a PCR-Select cDNA subtraction kit (CLONTECH).

**Differential Filter Hybridization.** A plasmid library was prepared with the cDNAs that remained after the forward subtraction with ear cDNA as tester. The PCR-amplified inserts of 376 clones were spotted onto duplicate filter sets and hybridized with radioactively labeled probes from either the forward or the

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession numbers AF359423–AF359436).

\*To whom reprint requests should be addressed at: Laboratory of Sensory Neuroscience and Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, Box 314, New York, NY 10021-6399. E-mail: hudspaj@rockvax.rockefeller.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.



**Fig. 1.** Zebrafish *cldna* and *cldnb* mark the earliest developmental stages of the acoustico-lateralis system. (a) Material dissected for the subtractive cDNA hybridization: ear region in yellow; head and tail regions in mauve. [Reprinted with permission from ref. 48 (Copyright 1995, Wiley-Liss, a subsidiary of John Wiley & Sons, Inc.).] (b and f–h) Whole-mount *in situ* hybridizations with antisense probes for *cldna*, which labels the developing ear, and *krox20*, which marks hindbrain rhombomeres 3 and 5. (c–e) Immunofluorescence detection of Cldna protein throughout the cell membranes of the otic vesicle after acetone permeabilization (c and d), but only at the apical surface of the sensory epithelium, the location of tight junctions, after detergent extraction (e; Cldna in red, nuclei in green). Each panel shows the combination of several adjacent confocal sections; in d, the vesicle was sectioned tangentially. (i–k) Whole-mount *in situ* hybridizations with antisense probes for *cldnb*, which labels placodal structures, including those of the ear and lateral-line organ, and for *krox20*. (k) Migrating placode of the posterior lateral line, with brown pigment cells and boundaries between somites 19 and 22 marked by slanted lines. Views are lateral (a, c–e, and k), dorsolateral (b), or dorsal (f–i), all with anterior to the left, and of whole embryos (a, b, and f–i) or details (c–e, j, and k); in i and j, the specimens have been flattened. L, lumen of the otic vesicle; A, apical surface of the otic vesicle's epithelium; HC, hair-cell nuclei; SC, supporting-cell nuclei; r3 and r5, hindbrain rhombomeres 3 and 5; OF, otic field; Olf, olfactory placode; OP, otic placode; PD, pronephric duct; ALL, anterior lateral-line placode; OV, otic vesicle; PLL, posterior lateral-line placode.

reverse subtraction. For 32 clones, the hybridization signal was more than one standard deviation above average with the forward-subtracted probe, but within one standard deviation of the average with the reverse-subtracted probe.

**In Situ Hybridization.** Whole-mount *in situ* hybridizations were conducted as described (20–22). For double labeling, a digoxigenin-labeled probe for zebrafish *cldna* or *cldnb* mRNA and a fluorescein-labeled probe for *krox20* mRNA were detected sequentially with the same blue chromogen. No labeling was observed in control hybridizations with sense probes.

**Cloning and Sequencing.** The partial *cldna* cDNA from the subtractive hybridization was used to screen a bacteriophage library from 24-h-old zebrafish embryos (Stratagene) and to isolate a full-length cDNA, as confirmed by Northern blotting. A partial cDNA for *Xenopus laevis cldna* was amplified with the degenerate primers 5'-TGGGARGGNHTNTGGATG-3' and 5'-RYNADNGGRTTRAARTCNYG-3' and used to isolate a full-length cDNA from an embryonic stage 30 library (Stratagene). Zebrafish and mouse cDNAs identified by BLAST searches (23) of GenBank were purchased from Incyte Genomics (St. Louis). The zebrafish and frog cDNAs were sequenced completely on both strands.

**Nomenclature.** We numbered each nonmammalian claudin gene like the human ortholog with the lowest suffix. If an ortholog

could not be identified unambiguously, a character suffix was assigned provisionally. If a subsequent study identifies the human ortholog, a character suffix may be changed to the appropriate number.

**Immunofluorescence.** The coding sequence of zebrafish *cldna* was amplified with the PCR primers 5'-CGGGATCCAATTGTTCCACCATGGTATCA-3' and 5'-GCAAGCTTGCTCTAGATCAGACATACCCCTTGGTTCC-3' and inserted between the *Bam*HI and *Hind*III sites of the vector pET-28a(+) (Novagen). The resulting fusion protein was expressed in *Escherichia coli* BL21(DE3), purified by affinity chromatography under denaturing conditions on Ni<sup>2+</sup>-Sepharose (Novagen), and used to immunize rabbits. Antibodies specific for the carboxyl terminus of Cldna were isolated by affinity chromatography (24) with the peptide NH<sub>2</sub>-CANSPPQDQYKATYTARSGGKGYV-COOH linked to Sulfolink gel (Pierce). On a Western blot, the affinity-purified antiserum recognized the recombinant full-length Cldna protein, but not a deletion mutant without the carboxyl terminus. After fixation, zebrafish embryos were permeabilized by extraction with acetone at –20°C for 7 min or with 1% (vol/vol) Triton X-100 in PBS at room temperature for 1 h, and labeled as described (ref. 25, protocol 6.12) with a 1:10,000 dilution of the anti-Cldna serum, 3 μg/ml Alexa-488-conjugated goat anti-rabbit antiserum, and 1 μg/ml of the nuclear stain

To-Pro-3 (Molecular Probes). Fluorescence was detected by confocal microscopy.

**Phylogenetic Analysis.** Protein sequences were aligned with CLUSTALX 1.81 (26), and transmembrane segments were predicted with TMHMM 2.0 (27). The aligned claudin sequences in FASTA format are available as supplemental text, which is published on the PNAS web site, [www.pnas.org](http://www.pnas.org). Pseudoreplicate data sets were generated with SEQBOOT from the PHYLIP 3.573c package (28); columns of the alignment that contain an initial methionine or a gap were disregarded. Maximum-likelihood distances and tree probabilities were calculated with TREE PUZZLE 5.0 (29) by using the WAG substitution matrix and a discrete  $\Gamma$  distribution with four bins. PAUP 4.0 $\beta$ 4a (30) was used in heuristic searches for optimal tree topologies with minimum evolution as the objective function and to draw a consensus tree.

**Locating Introns.** For the human claudin genes, we deduced exon boundaries in the coding regions from the literature (31) or published mRNA and genomic sequences. For mouse *Cldn13* and the zebrafish claudins, we used PRIMER3 ([www.genome.wi.mit.edu/genome\\_software/other/primer3.html](http://www.genome.wi.mit.edu/genome_software/other/primer3.html)) to design primer pairs that flanked the intron positions observed in the human genes (Table 2, which is published as supplemental material). To amplify products shorter than 1 kb, we conducted PCRs containing 0.035 units/ $\mu$ l Amplitaq Gold enzyme (Applied Biosystems), 1 $\times$  Amplitaq Gold buffer, 2.5 mM MgCl<sub>2</sub>, 0.2 mM each dNTP, 0.8  $\mu$ M each primer, and 2.5 ng/ $\mu$ l genomic DNA (Research Genetics, Huntsville, AL) with the following protocol: 95°C for 10 min; 30 cycles of 94°C for 30 s, 68°C for 30 s, 72°C for 2 min; and 72°C for 3.5 min. To amplify products longer than 1 kb, we conducted PCRs containing 0.15  $\mu$ l Expand long-template enzyme mixture (Roche Diagnostics) per 10  $\mu$ l, 1 $\times$  Expand buffer 2, 0.5 mM each dNTP, 0.3  $\mu$ M each primer, and 5 ng/ $\mu$ l genomic DNA with the following protocol: 92°C for 2 min; 10 cycles of 92°C for 10 s, 68°C for 30 s, 68°C for 15 min; 20 cycles of 92°C for 10 s, 68°C for 30 s, 68°C for 15 min plus a 20-s extension per cycle; and 68°C for 7 min. The PCR products were sized by gel electrophoresis and sequenced across the exon-intron boundaries. In all cases, splice donor (GT) and acceptor (AG) sequences bounded the introns.

**Physical Mapping.** The chromosomal locations of the human claudin genes were obtained from the LOCUSLINK database (32) and BLAST searches of the human genome sequence. Zebrafish genes were mapped with the T51 radiation-hybrid panel (Research Genetics), the primers shown in Table 2, and Amplitaq Gold under the conditions described above, except for an annealing temperature of 60°C. The scoring data were analyzed by use of the mapping server at the Max-Planck-Institut für Entwicklungsbiologie in Tübingen, Germany (33). Zebrafish orthologs of genes within 2 Mb of human *CLDN3* and *CLDN4* were found by iterative searches of GenBank. Orthology was assumed if the predicted zebrafish protein in a reciprocal BLASTP search matched the human protein most closely, with at least 60% pairwise sequence identity, and clustered with the human protein on a phylogenetic tree of similar sequences from any organism.

## Results and Discussion

To study the formation of the vertebrate ear from its primordium, the otic placode (34), we conducted a screen for molecular markers in the zebrafish that comprised a cDNA subtraction (35), a differential filter hybridization, and *in situ* hybridizations at prim-6 stage (Fig. 1*a*). At this stage, when the first sensory hair cells have formed (36), one of the genes we identified was expressed exclusively in the otic vesicle (Fig. 1*b*). Because the protein encoded by a full-length cDNA from this gene is most

similar to mammalian claudins, we refer to the zebrafish gene as *claudin a* (*cldna*; see *Materials and Methods* for nomenclature). Consistent with the presence of the protein Cldna in tight junctions, an antiserum against its unique carboxyl terminus detected the protein after detergent extraction solely at the apical surface of the otic vesicle's epithelium (Fig. 1*c–e*). Expression in the ear has previously been described only for mouse *Cldn14* and only after birth (9). Immature tight junctions, however, have been observed in the chicken's ear as early as embryonic stage 15, before the otic cup closes (37).

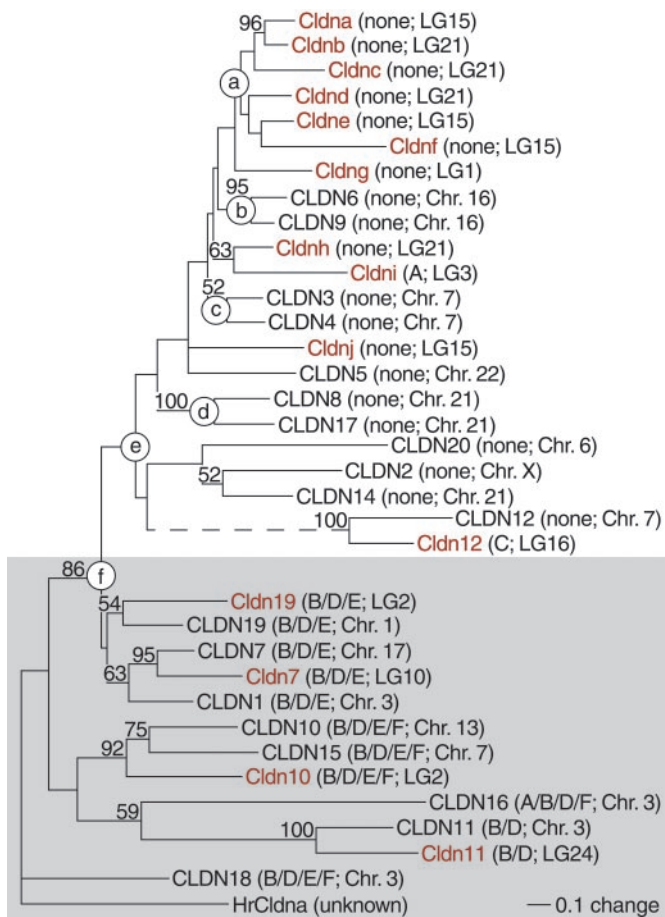
Because the mammalian claudin family comprises at least 20 members (6), we speculated that zebrafish also might harbor multiple claudins and that more than one gene might be expressed in the developing ear. BLAST searches of GenBank with all known human claudin sequences revealed 14 additional zebrafish genes that encode claudin-like proteins; reciprocal database searches with the zebrafish protein sequences always identified a mammalian claudin as the closest match, with at least 43% sequence identity (Fig. 5, which is published as supplemental material). Although our set of 15 zebrafish claudins may be incomplete, it is comparable in size to the mammalian gene family. In whole-mount *in situ* hybridizations with probes for the 14 additional zebrafish claudins, we found that one of them, *cldnb*, also is expressed in the ear at prim-6 stage (Fig. 1*j*).

During the first 36 h of development, the expression patterns of *cldna* and *cldnb* are similar, but not identical. We first detected the mRNAs of both genes at the end of gastrulation (Fig. 1*f* and data not shown), shortly after the time of otic induction (38). Their initial expression domains, diffuse stripes along both sides of the hindbrain primordium, coincide with the presumptive otic fields. At the 10-somite stage, concomitant with the invagination of the otic placode, these domains have shrunk and sharpened (Fig. 1*g* and *h*). Unlike *cldna*, which is expressed exclusively in the ear at least until prim-22 stage, *cldnb* is expressed also in the olfactory placode and the pronephric duct (Fig. 1*i*). In addition, we detected *cldnb* mRNA after prim-6 stage in patches rostral and caudal to the otic vesicles (Fig. 1*j*) that represent placodes of the anterior and posterior lateral lines (39). At later stages, the caudal boundary of the posterior patch coincides with that of the migrating primordium of the posterior lateral line (Fig. 1*k*). Zebrafish *cldna* and *cldnb* thus mark the initial stages of the acoustico-lateralis system; *cldna* is one of the earliest molecular markers exclusive to the ear (34).

Sensory placodes and the pronephros, the domains of *cldna* and *cldnb* expression, are characters of vertebrate embryos that arose along the chordate stem lineage (13). Tight junctions, the cellular structures in which claudins occur, are likewise characteristic of chordates (1). Suspecting a link between the evolution of vertebrates and claudins, we therefore investigated the phylogeny of the claudin gene family, with an emphasis on zebrafish *cldna* and *cldnb*.

First, we asked whether any nonvertebrate claudins were known. Outside the Phylum Chordata we and others (40) found no sequences in GenBank with significant similarity to claudins, even in the sequenced genomes of *Drosophila melanogaster* and *Caenorhabditis elegans*. The closest match we found, with a BLASTP expectation value of 0.0002, was the hypothetical protein F53B3.5 from *C. elegans*, which lacks several of the characteristic residues of claudin proteins (Fig. 5) and is equally similar to other members of the tetraspanin superfamily (5). Among nonvertebrate chordates, we found a single claudin gene from the urochordate *Halocynthia roretzi* (Fig. 5). These findings are consistent with the presence of claudins only in chordates.

We next asked when the claudin gene family had diversified. For a representative set of 36 claudins, we analyzed the aligned protein sequences, intron positions, and chromosomal locations. From the sequence alignment (Fig. 5), we reconstructed a 50%-majority consensus tree with compatible groupings by using



**Fig. 2.** Zebrafish paralogs *cldna* and *cl دنب* are orthologs of mammalian claudins 3–6 and 9. This 50% majority consensus tree with compatible groupings was reconstructed from the alignment in Fig. 5 by using maximum-likelihood distances and a heuristic tree-search algorithm (41) and rooted for display purposes with HrCldna from *Halocynthia roretzi* as the outgroup. Numbers indicate the percentage of 1,000 bootstrap replicates that support the adjacent node. Zebrafish proteins are shown in red, human and urochordate in black; introns in the coding region (Table 1) and the linkage group or chromosome of the corresponding gene (Table 2) are listed in parentheses. Claudins whose genes contain multiple introns in their coding regions are shaded in gray. CLDN12 and Cldn12 appear to be tetraspanins, but only distant relatives of the other claudins; their common branch (dashed line) was shortened to two-fifths to conserve space. Mouse Cldn13 was omitted from this tree because its placement, either with CLDN3 and CLDN4 or with CLDN12 and Cldn12, was the least consistent among bootstrap replicates and its omission markedly increased the tree's likelihood.

maximum-likelihood distances and a heuristic tree-search algorithm (41). The resulting phylogram (Fig. 2) is supported by three independent lines of evidence. First, we obtained essentially the same topology after we replaced the human sequences with their mouse orthologs (data not shown). Next, all claudin genes with multiple introns in the coding region cluster together (Table 1 and Fig. 2, gray shading). The remaining genes lack introns in the coding region save for zebrafish *cldni* and *clدنب12*, each of which contains a single intron in an unusual position (Table 1). The bipartite configuration of the phylogram strongly suggests a concerted loss of introns by retrotransposition between nodes e and f. Finally, sets of genes whose chromosomal locations coincide, presumably because of recent tandem gene duplications, also cluster on the tree (Fig. 2, nodes a to d). Because the zebrafish and human claudins are mostly interspersed in this phylogram, we conclude that by the time ray-

**Table 1.** Conservation of intron positions in the coding regions of human and zebrafish claudins

Intron*	Gene†	Exon boundaries‡
A	<i>CLDN16</i>	Leu-Glu <sub>108</sub> —Val <sub>109</sub> -Ser
	<i>clدنب1</i>	Ala-Gln <sub>44</sub> —Val <sub>45</sub> -Ser
B	<i>CLDN1</i>	Leu-Ser-S—er <sub>75</sub> -Thr-Leu
	<i>CLDN7</i>	Leu-Ser-A—la <sub>75</sub> -Ala-Leu
	<i>clدنب7</i>	Leu-Asp-S—er <sub>75</sub> -Ala-Leu
	<i>CLDN10</i>	Leu-Asp-G—ly <sub>74</sub> -Tyr-Ile
	<i>clدنب10</i>	Leu-Pro-V—a <sub>73</sub> -His-Ile
	<i>CLDN11</i>	Leu-Pro-G—ly <sub>76</sub> -Tyr-Val
	<i>clدنب11</i>	Ile-Pro-A—la <sub>76</sub> -Tyr-Ile
	<i>CLDN15</i>	Leu-Ser-G—ly <sub>74</sub> -Tyr-Ile
	<i>CLDN16</i>	His-Pro-L—eu <sub>143</sub> -Lys-Leu
	<i>CLDN18</i>	Leu-Pro-A—la <sub>74</sub> -Met-Leu
C	<i>CLDN19</i>	Leu-Asp-G—ly <sub>75</sub> -His-Ile
	<i>clدنب19</i>	Leu-Pro-G—ly <sub>75</sub> -Glu-Ile
	<i>clدنب12</i>	Pro-Thr-G—ly <sub>93</sub> -Leu-Leu
D	<i>CLDN1</i>	Leu-Ala-G—ly <sub>130</sub> -Leu-Ala
	<i>CLDN7</i>	Val-Ala-G—ly <sub>130</sub> -Leu-Ala
	<i>clدنب7</i>	Val-Gly-A—la <sub>130</sub> -Leu-Cys
	<i>CLDN10</i>	Leu-Ser-G—ly <sub>128</sub> -Leu-Cys
	<i>clدنب10</i>	Ile-Gly-G—ly <sub>128</sub> -Leu-Cys
	<i>CLDN11</i>	Leu-Leu-A—la <sub>131</sub> -Leu-Cys
	<i>clدنب11</i>	Val-Ile-S—er <sub>131</sub> -Leu-Cys
	<i>CLDN15</i>	Leu-Ala-G—ly <sub>128</sub> -Ile-Cys
	<i>CLDN16</i>	Ile-Ala-G—ly <sub>198</sub> -Thr-Pro
	<i>CLDN18</i>	Val-Ser-G—ly <sub>129</sub> -Leu-Cys
E	<i>CLDN19</i>	Leu-Ala-G—ly <sub>130</sub> -Leu-Cys
	<i>clدنب19</i>	Thr-Gly-G—ly <sub>130</sub> -Leu-Phe
	<i>CLDN1</i>	Asn-Ala-Ar—g <sub>158</sub> -Tyr-Glu
	<i>CLDN7</i>	Asn-Ile-Ly—s <sub>158</sub> -Tyr-Glu
	<i>clدنب7</i>	Asn-Thr-Ly—s <sub>158</sub> -Tyr-Glu
F	<i>CLDN10</i>	Glu-Gln-Ly—s <sub>155</sub> -Tyr-Glu
	<i>clدنب10</i>	Gly-Val-Ar—g <sub>155</sub> -Phe-Glu
	<i>CLDN15</i>	Gly-Thr-Ly—s <sub>155</sub> -Tyr-Glu
	<i>CLDN18</i>	Gln-Thr-Ar—g <sub>168</sub> -Tyr-Thr
	<i>CLDN19</i>	Asn-Ala-Ar—g <sub>158</sub> -Tyr-Glu
	<i>clدنب19</i>	Asn-Ala-Ar—g <sub>158</sub> -Tyr-Glu
	<i>CLDN10</i>	Thr-Pro-Ar—g <sub>191</sub> -Tyr-Thr
	<i>clدنب10</i>	Glu-Lys-Gl—y <sub>193</sub> -Ala-Tyr
<i>CLDN15</i>	Ala-Ala-Se—r <sub>194</sub> -Ala-Arg	
<i>CLDN16</i>	Phe-Lys-A—sp <sub>262</sub> -Val-Gly	
<i>CLDN18</i>	Glu-Thr-As—n <sub>205</sub> -Tyr-Lys	

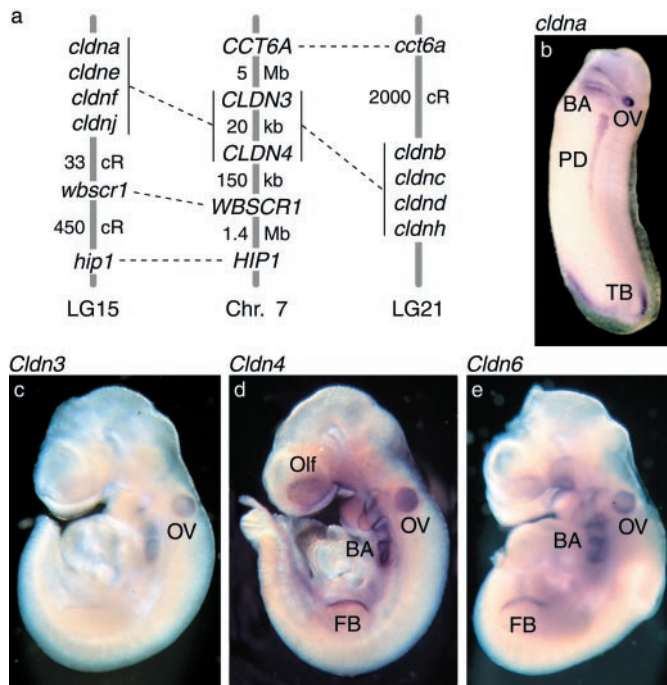
\*See Fig. 5 for the location of the introns in the sequence alignment.

†Symbols capitalized for human and lowercase for zebrafish.

‡Numbers denote the position of the amino acids whose codons border or straddle the intron (—).

finned fishes diverged from the tetrapod lineage, vertebrates already possessed an abundance of claudin genes.

We then examined the homologs of *clدنبا* and *clدنبب* in the zebrafish and in mammals. The phylogram provides strong bootstrap support for *clدنبا* and *clدنبب* being closest paralogs (Fig. 2) that arose from a recent duplication of a gene cluster now located on both zebrafish linkage groups LG15 and LG21 (Fig. 3a). The overlapping but distinct expression patterns of these two genes may reflect degenerative mutations of tissue-specific regulatory elements (17). The phylogram also suggests that mammalian claudins 6 and 9, 3 and 4, and 5 are orthologs of zebrafish *clدنبا* and *clدنبب*. In support of this proposition, we detected conserved synteny between zebrafish LG15 and LG21, where *clدنبا* and *clدنبب* occur, and human chromosome 7, where *CLDN3* and *CLDN4* reside (Fig. 3a). These findings

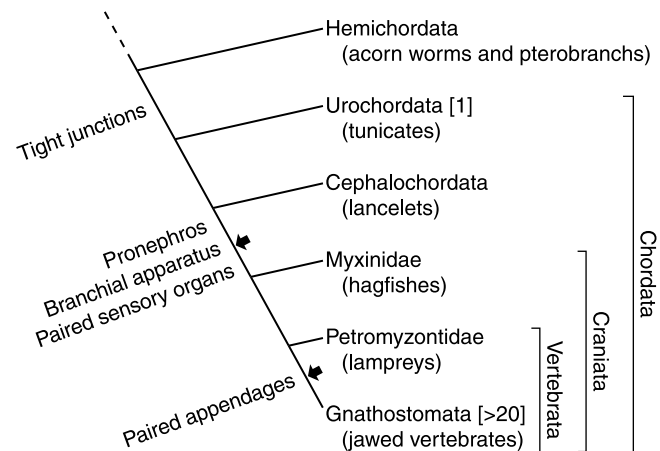


**Fig. 3.** Orthologs of zebrafish *cldna* and *cldnb* are expressed in vertebrate primordia. (a) Conserved syntenies between zebrafish linkage groups 15 and 21, which harbor paralogous clusters of claudin genes that include *cldna* and *cldnb*, and human chromosome 7, which contains *CLDN3* and *CLDN4*. We mapped putative zebrafish orthologs of six genes from the vicinity of *CLDN3* and *CLDN4* (45); two of them, *hip1* and *wbscr1*, occur in LG15 near *cldna* (Table 2). In addition, zebrafish LG21 harbors the ortholog of the human *CCT6A* gene on 7p11 (46). (b) *In situ* hybridization of a stage 30 *X. laevis* embryo with an antisense probe for frog *cldna*. (c–e) *In situ* hybridizations of mouse embryos between embryonic days 9.0 and 9.5 with antisense probes for *Cldn3*, *Cldn4*, and *Cldn6*. All views are lateral. BA, branchial arches; OV, otic vesicle; PD, pronephric duct; TB, tail bud; Olf, olfactory placode; FB, forelimb bud.

indicate that the claudin family continued to grow through gene or genome duplications after the divergence of ray-finned fishes and tetrapods.

Finally, we asked whether the orthologs of zebrafish *cldna* and *cldnb* exhibit similar expression patterns early in organogenesis. In whole-mount *in situ* hybridizations with mice between embryonic days 9.0 and 9.5, we detected *Cldn3*, *Cldn4*, and *Cldn6* mRNAs in the otic vesicle as well as in the olfactory placode, branchial arches, and forelimb bud (Fig. 3 c–e). At a comparable stage of development in *Xenopus* we detected expression of its *cldna* gene, which is most similar to mammalian claudins 4 and 6, in the otic vesicle, branchial arches, pronephric duct, and tailbud (Fig. 3b; see also ref. 42). These results demonstrate that claudin genes closely related to zebrafish *cldna* and *cldnb* are expressed specifically in primordia characteristic of the vertebrate embryo.

We conclude that claudins emerged around the same time as chordates (Fig. 4), when tight junctions replaced septate junctions (1). Like many other gene families (16), the claudins then expanded as the vertebrate body plan evolved from the chordate stem configuration through the addition of novel structures such as the cranium, paired sensory organs, and paired appendages



**Fig. 4.** Concurrent evolution of the claudin family and the vertebrate body plan. The number of distinct claudin genes found to date in each clade is given in brackets. The appearance of those vertebrate characters that we found to be associated with claudin expression is indicated beside the stem lineage (13). Arrows indicate two presumptive genome duplications (16); a third genome duplication may have taken place in ray-finned fishes alone (15, 47).

(13, 14). A more precise specification of this model will require exhaustive screens for claudin genes in the nongnathostome clades shown in Fig. 4. What were the claudins' precursors in nonchordates and were they associated with septate junctions? How many claudin genes did the common chordate ancestor harbor? Did the expansion of the claudin gene family result from whole-genome duplications in the vertebrate lineage? The strong and specific expression of claudins in vertebrate primordia reported here and the randomization of the embryo's left-right axis upon overexpression of frog *cldna* (42) suggest that tight junctions play an active role in morphogenesis. Such a role may explain some of the morphological and neurological abnormalities of individuals with Williams–Beuren syndrome, in which the *CLDN3* and *CLDN4* genes are excised as part of a large deletion (43). The diversification of the claudin gene family may have facilitated the formation of tight junctions during organogenesis at different times and at different sites. Alternatively, the various claudins may mediate selective cell adhesion (44) in vertebrate primordia.

**Note Added in Proof.** An additional zebrafish claudin, *Cldn8*, is encoded by expressed sequence tags fl13b04 and fl24f07. The *cldn8* gene contains no introns within the coding region and resides in LG15.

We thank Dr. Robert Geisler for the analysis of radiation-hybrid mapping data, Dr. Ali Hemmati-Brivanlou for access to his *Xenopus* facility, and Dr. Kathleen Millen for suggestions about mouse *in situ* hybridizations. Ms. Jennifer Nuñez, Ms. Carrie Rosenberger, and Mr. Jak Fak contributed at the beginning of this project. Mr. Adam Kuhl, Ms. Andrea Seymour, Mr. David Chosid, and Ms. Jennifer Voisine maintained the zebrafish colony. Dr. Hemmati-Brivanlou, Dr. Axel Meyer and members of our laboratory group provided comments on the manuscript. This research was supported by a grant from the American Hearing Research Foundation to R.K. and by National Institutes of Health Grant DC00317 to A.J.H., who is an Investigator of the Howard Hughes Medical Institute.

- Green, C. R. & Bergquist, P. R. (1982) *J. Cell Sci.* **53**, 279–305.
- Tsukita, S., Furuse, M. & Itoh, M. (2001) *Nat. Rev. Mol. Cell Biol.* **2**, 285–293.
- Lane, N. J. (1992) in *Tight Junctions*, ed. Cereijido, M. (CRC, Boca Raton, FL), pp. 23–48.
- Bellen, H. J., Lu, Y., Beckstead, R. & Bhat, M. A. (1998) *Trends Neurosci.* **21**, 444–449.
- Bronstein, J. M. (2000) *Curr. Opin. Neurobiol.* **10**, 552–557.

- Tsukita, S. & Furuse, M. (1999) *Trends Cell Biol.* **9**, 268–273.
- Gow, A., Southwood, C. M., Li, J. S., Pariali, M., Riordan, G. P., Brodie, S. E., Danias, J., Bronstein, J. M., Kachar, B. & Lazzarini, R. A. (1999) *Cell* **99**, 649–659.
- Simon, D. B., Lu, Y., Choate, K. A., Velazquez, H., Al-Sabban, E., Praga, M., Casari, G., Bettinelli, A., Colussi, G., Rodriguez-Soriano, J., et al. (1999) *Science* **285**, 103–106.

9. Wilcox, E. R., Burton, Q. L., Naz, S., Riazuddin, S., Smith, T. N., Ploplis, B., Belyantseva, I., Ben-Yosef, T., Liburd, N. A., Morell, R. J., *et al.* (2001) *Cell* **104**, 165–172.
10. Furuse, M., Sasaki, H., Fujimoto, K. & Tsukita, S. (1998) *J. Cell Biol.* **143**, 391–401.
11. Kirschner, M. & Gerhart, J. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 8420–8427.
12. Carroll, S. B. (2001) *Nature (London)* **409**, 1102–1109.
13. Maisey, J. G. (1986) *Cladistics* **2**, 201–256.
14. Shimeld, S. M. & Holland, P. W. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 4449–4452.
15. Meyer, A. & Schartl, M. (1999) *Curr. Opin. Cell Biol.* **11**, 699–704.
16. Holland, P. W. H. (1999) *Semin. Cell Dev. Biol.* **10**, 541–547.
17. Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. & Postlethwait, J. (1999) *Genetics* **151**, 1531–1545.
18. Richardson, M. K. (1995) *Dev. Biol.* **172**, 412–421.
19. Fleming, T. P., Papenbrock, T., Fesenko, I., Hausen, P. & Sheth, B. (2000) *Semin. Cell Dev. Biol.* **11**, 291–299.
20. Harland, R. M. (1991) *Methods Cell Biol.* **36**, 685–695.
21. Parr, B. A., Shea, M. J., Vassileva, G. & McMahon, A. P. (1993) *Development (Cambridge, U.K.)* **119**, 247–261.
22. Thisse, C. & Thisse, B. (1998) in *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Danio rerio)*, ed. Westerfield, M. (Univ. of Oregon Press, Eugene), p. 9.82.
23. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402.
24. Harlow, E. & Lane, D. (1999) *Using Antibodies: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
25. Jowett, T. (1999) *Methods Cell Biol.* **59**, 63–85.
26. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucleic Acids Res.* **25**, 4876–4882.
27. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. (2001) *J. Mol. Biol.* **305**, 567–580.
28. Felsenstein, J. (1989) *Cladistics* **5**, 164–166.
29. Strimmer, K. & von Haeseler, A. (1996) *Mol. Biol. Evol.* **13**, 964–969.
30. Swofford, D. L. (2000) *PAUP\*: Phylogenetic Analysis Using Parsimony (\* and Other Methods)* (Sinauer, Sunderland, MA).
31. Paperna, T., Peoples, R., Wang, Y.-K., Kaplan, P. & Francke, U. (1998) *Genomics* **54**, 453–459.
32. Wheeler, D. L., Church, D. M., Lash, A. E., Leipe, D. D., Madden, T. L., Pontius, J. U., Schuler, G. D., Schriml, L. M., Tatusova, T. A., Wagner, L., *et al.* (2001) *Nucleic Acids Res.* **29**, 11–16.
33. Geisler, R., Rauch, G. J., Baier, H., van Bebber, F., Brobeta, L., Dekens, M. P., Finger, K., Fricke, C., Gates, M. A., Geiger, H., *et al.* (1999) *Nat. Genet.* **23**, 86–89.
34. Torres, M. & Giraldez, F. (1998) *Mech. Dev.* **71**, 5–21.
35. Diatchenko, L., Lau, Y.-F., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D., *et al.* (1996) *Proc. Natl. Acad. Sci. USA* **93**, 6025–6030.
36. Haddon, C. & Lewis, J. (1996) *J. Comp. Neurol.* **365**, 113–128.
37. Ginzberg, R. D. & Gilula, N. B. (1979) *Dev. Biol.* **68**, 110–129.
38. Mendonsa, E. S. & Riley, B. B. (1999) *Dev. Biol.* **206**, 100–112.
39. Raible, D. W. & Kruse, G. J. (2000) *J. Comp. Neurol.* **421**, 189–198.
40. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., *et al.* (2001) *Science* **291**, 1304–1351.
41. Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. (1996) in *Molecular Systematics*, eds. Hillis, D. M., Moritz, C. & Mable, B. K. (Sinauer, Sunderland, MA), pp. 407–514.
42. Brizuela, B. J., Wessely, O. & De Robertis, E. M. (2001) *Dev. Biol.* **230**, 217–229.
43. Morris, C. A. & Mervis, C. B. (2000) *Annu. Rev. Genomics Hum. Genet.* **1**, 461–484.
44. Furuse, M., Sasaki, H. & Tsukita, S. (1999) *J. Cell Biol.* **147**, 891–903.
45. Valero, M. C., de Luis, O., Cruces, J. & Perez J urado, L. A. (2000) *Genomics* **69**, 1–13.
46. Barbazuk, W. B., Korf, I., Kadavi, C., Heyen, J., Tate, S., Wun, E., Bedell, J. A., McPherson, J. D. & Johnson, S. L. (2000) *Genome Res.* **10**, 1351–1358.
47. Wittbrodt, J., Meyer, A. & Schartl, M. (1998) *BioEssays* **20**, 511–515.
48. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. (1995) *Dev. Dyn.* **203**, 253–310.