JARO 7: 160–172 (2006)
DOI: 10.1007/s10162-006-0032-0

JARO
Journal of the Association for Research in Otolaryngology

Characterization of an Abundant COL9A1 Transcript in the Cochlea with a Novel 3′ UTR: Expression Studies and Detection of miRNA Target Sequence

THERU A. SIVAKUMARAN,[1,2] BARBARA L. RESENDES,[1,2] NAHID G. ROBERTSON,[1] ANNE B. S. GIERSCH,[2,3] AND CYNTHIA C. MORTON[1,2,3,4]

[1]*Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Boston, MA 02115, USA*

[2]*Harvard Medical School, Boston, MA 02115, USA*

[3]*Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA*

[4]*Brigham and Women's Hospital, New Research Building, Room 160, 77 Avenue Louis Pasteur, Boston, MA 02115, USA*

Received: 11 November 2005; Accepted: 7 March 2006; Online Publication: 19 April 2006

ABSTRACT

EST N66408 represents one of several large unique clusters expressed in the Morton human fetal cochlear cDNA library. N66408 is 575 bp in size and initial BLAST analysis of this sequence showed no homology to any known genes or expressed sequence tags (ESTs) from other organs or tissues. Sequence of the original cochlear clone from which N66408 was derived revealed that the corresponding cDNA was about 700 bp in size, including 125 bp at its 5′ end with homology to the 3′ end of *COL9A1* in addition to 575 bp of novel sequence. RT-PCR analysis using primers specific to *COL9A1* isoforms 1 and 2 detected expression of both isoforms in human fetal cochlea. Tissue *in situ* hybridization using the novel 3′ UTR sequence as probe showed abundant expression in spiral limbus and spiral ligament, and a moderate level of expression in the organ of Corti. dbEST analysis of ESTs specific to the 3′ UTR of *COL9A1* showed 19 ESTs derived from various tissues; three polyadenylation sites were identified and the majority of these ESTs were derived from overlapping polyadenylation signals at the second site (position 749–758). Comparison of the 3′ UTR of human *COL9A1* with its orthologs as well as with dbEST uncovered a highly conserved region around the overlapping polyadenylation signals at position 749–758 in mammals. A search of the microRNA database revealed a highly conserved target sequence for miR-9 immediately preceding the overlapping polyadenylation signals in the novel 3′ UTR of *COL9A1*, suggesting its role in posttranscriptional regulation of *COL9A1*.

Keywords: *COL9A1*, long form or isoform 1, short form or isoform 2, novel 3′ UTR, ESTs, cochlea

INTRODUCTION

Hearing loss is the most common sensory disorder in humans. Approximately one in 1,000 children in the United States has congenital deafness (Morton 1991), and at least half of these cases has a genetic etiology. About 70% of cases attributed to genetic causes represent nonsyndromic deafness, and the remaining 30% are syndromic and associated with other clinical findings. Among nonsyndromic deafness, autosomal recessive inheritance accounts for about 80% of cases, autosomal dominant inheritance is observed in about 20%, and X-linked (about 1%) and mitochondrial (less than 1%) forms are comparatively infrequently observed. The unparalled heterogeneity seen in heritable deafness suggests the involvement of many genes in the hearing process. More than 95 loci have been mapped for nonsyndromic hearing loss (i.e., 40 autosomal dominant, 51 autosomal recessive, 4 X-linked, and several mito-

Correspondence to: Cynthia C. Morton · Brigham and Women's Hospital · New Research Building, Room 160, 77 Avenue Louis Pasteur, Boston, MA 02115, USA. Telephone: +1-617-5254535; fax: +1-617-5254533; email: cmorton@partners.org

chondrial loci), and more than 60 genes have been identified (as of February 2006) as responsible for syndromic and nonsyndromic sensorineural hearing loss (http://webhost.ua.ac.be/hhh).

The impressive genetic heterogeneity in nonsyndromic deafness confounds traditional genetic linkage analysis. An alternative method for identification of disease-causing genes is one that utilizes tissue-specific cDNA libraries (Hedrick et al. 1984; Jones and Reed 1989; Gurish et al. 1992). To identify genes responsible for hearing loss and to gain a more fundamental understanding of human hearing at a molecular level, we constructed a human fetal cochlear cDNA library (Robertson et al. 1994; Skvorak et al. 1999). Various methods were undertaken to analyze transcripts in this library, namely, the combined approaches of subtractive hybridization and differential screening and expressed sequence tag (EST) production and analysis (Robertson et al. 1994; Skvorak et al. 1999; Resendes et al. 2002). More than 15,000 ESTs were generated. About 8,500 ESTs corresponded to more than 2,300 known genes, more than 1,770 EST clusters were expressed in other libraries, and approximately 340 EST clusters were unique to the cochlear library (Skvorak et al. 1999; Resendes et al. 2002). Analysis of the cochlear ESTs revealed 788 discrete genetic loci, some of which reside within genomic regions for mapped deafness disorders and represent potential positional candidate genes for hearing impairment (http://hearing.bwh.harvard.edu). With these genomic approaches, several novel auditory genes such as *COCH*, *OTOR*, and *PFET1* (now designated *KCTD12*) have been identified (Robertson et al. 1997, 2000; Resendes et al. 2004). Missense mutations in *COCH* were subsequently shown to be etiologic in the autosomal dominant sensorineural hearing loss and vestibular disorder, DFNA9 (Robertson et al. 1998). *OTOR* is an inner ear preferentially expressed gene as determined by UniGene dbEST and NCBI BLAST searches (Beisel et al. 2004). The recently described *KCTD12* is an intronless gene abundantly expressed in the fetal inner ear and a variety of other fetal organs, but at barely detectable levels in adult organs (Resendes et al. 2004).

In this work, we present characterization of an EST from one of the unique EST clusters identified from the sequence analysis of the human fetal cochlear cDNA library. This EST was subsequently determined to represent a novel 3′ UTR of the short form of a collagen gene, *COL9A1*, and was selected for further investigation based on the fundamental role of collagens in the development and maintenance of hearing. Collagens expressed in the inner ear and essential for hearing are illustrated by the association of defects in several syndromic disorders including hearing loss, such as mutations in *COL1A1* and *COL1A2* in osteogenesis imperfecta (Kuivaniemi et al. 1991, 1997); *COL2A1* and *COL11A1* in classic Stickler syndrome (Ahmad et al. 1991; Winterpacht et al. 1993; Richards et al. 1996); *COL11A2* in nonocular Stickler syndrome (Sirko-Osadsa et al. 1998) and otospondylomegaepiphyseal dysplasia (OSMED) syndrome (van Steensel et al. 1997); *COL11A1* in Marshall syndrome (Griffith et al. 2000); and *COL4A3*, *COL4A4*, and *COL4A5* in Alport syndrome (Barker et al. 1990; Lemmink et al. 1994; Mochizuki et al. 1994). Mutations in *COL11A2* have also been found in DFNA13, an autosomal dominant nonsyndromic sensorineural hearing loss (McGuirt et al. 1999).

MATERIALS AND METHODS

EST sequence analysis

A unique EST cluster representing the 3′ UTR of *COL9A1* was derived from the unsubtracted, nonnormalized Morton fetal cochlear cDNA library. Briefly, mass *in vivo* excision of the human fetal (16–22 weeks gestational age) cochlear cDNA library, constructed in the UniZap vector, was performed according to the manufacturer's protocol (Robertson et al. 1994) (Stratagene, La Jolla, CA, USA). The cDNA library was then contributed to the IMAGE Consortium to generate human cochlear ESTs, and sequences subsequently deposited in GenBank (Skvorak et al. 1999).

Sequence analysis

Nucleotide sequence of cDNA clones was determined by using an ABI PRISM dye-terminator cycle-sequencing system (PE Applied Biosystems, Foster City, CA, USA). Sequence analysis was performed using the University of Wisconsin Genetics Computer Group software (Devereux et al. 1984). The nucleotide sequence of N66408 was compared to sequences contained in various databases including the GenBank primate database and sequences in the EST, STS, and nonhuman mammalian databases (Skvorak et al. 1999) using the BLAST network service of the National Center for Biotechnology Information (Altschul et al. 1997).

Northern blot analysis

Total cellular RNAs were extracted (Chirgwin et al. 1979) from second trimester human fetal cochlea, brain, skeletal muscle, testis, eye, placenta, thymus, spleen, tongue, liver, kidney and bone marrow. All human tissues were obtained according to guidelines

established by the Partners Human Research Committee. Ten micrograms of each of the RNAs was electrophoresed in 1% agarose–formaldehyde gels and transferred to GeneScreen (DuPont, Wilmington, DE, USA) filters (Thomas 1980). Filters were prehybridized for 2–4 h and hybridized overnight at 42°C with a ^{32}P-labeled probe derived from the original human EST N66408 clone. After incubation, filters were washed in 0.1× SSC in 0.1% SDS at 42–55°C prior to autoradiography overnight using XAR-5 film (Kodak, Rochester, NY, USA) with intensifying screens at −80°C.

Reverse transcription-polymerase chain reaction

Reverse transcription-polymerase chain reaction (RT-PCR) was performed with 2–3 μg of total RNAs extracted from all human tissues using the SuperScript™ III first-strand synthesis system (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. Full-length *COL9A1* transcripts of both long and short forms (also referred to as isoforms 1 and 2, respectively) were amplified by using six sets of overlapping gene-specific primers with 5 min initial denaturation at 96°C, 30 cycles of 97°C for 30 s, 58°C for 30 s, and 72°C for 30 s, and a final extension of 7 min at 72°C. Primer sequences are shown in Table 1. Expression of the two different isoforms of *COL9A1* in various tissues was studied by RT-PCR using 5′ end isoform-specific primers and 3′ UTR sequence-specific primers.

Tissue *in situ* hybridization

Riboprobe preparation

Radioactive *in situ* hybridization was carried out with a 450-bp PCR-generated probe synthesized from the novel 3′ UTR of *COL9A1*, amplified by using gene-specific primers with T7 and SP6 RNA polymerase-specific promoter sequences at the 5′ end of forward and reverse primers, respectively. Primers and conditions used for amplification were as follows: forward primer (5′-GAA TTG TAA TAC GAC TCA CTA TAG GGA GTT GCA GTA GTT ATT TTC ATT TAG-3′) and reverse primer (5′-CAA GCT ATT TAG GTG ACA CTA TAG AAC ATG CAA GAG ATT AAG AAA CCA-3′); initial denaturation at 96°C for 5 min; 30 cycles of 97°C for 30 s, 50°C for 30 s, and 72°C for 30 s, and a final extension at 72°C for 7 min. The gel-purified PCR product with a final concentration of 1 μg/μL was used to prepare ^{35}S-labeled riboprobe using an *in vitro* transcription kit (Roche Molecular Biochemicals, Mannheim, Germany). The efficiency of transcription and incorporation of ^{35}S-UTP into riboprobe were evaluated by 5% acrylamide-urea gel electrophoresis and quantification of counts by using a scintillation counter, respectively.

Tissue

Cochlear tissue from a 20-gestational-week-old human fetus was fixed in 4% paraformaldehyde/1× PBS for 3 days, decalcified in 0.1 M EDTA/1× PBS, pH 7.3 for 7 days, and dehydrated before being embedded in paraffin by standard histologic procedures. Sections (17 μm) were cut with a cryomicrotome (Leica, Malven, PA, USA), collected on silyated slides (PGC Scientific, Gaithersburg, MD, USA), dried at 37°C for 45 min, and stored frozen at −20°C until use.

In situ hybridization

Tissue sections were deparaffinized, fixed in 4% paraformaldehyde in PBS, and treated with proteinase K. After washing in 0.5× SSC, sections were prehybridized in 50% deionized formamide, 0.3 M NaCl, 20 mM Tris (pH 8.0), 5 mM EDTA, 1×

TABLE 1

Oligonucleotide primers used in *COL9A1* RT-PCR

Fragments	*Forward Primer*	*Reverse Primer*
Isoform 1 specific	COL9A1F1-F: CCA GAA CAC ATA GTC CTA GGG TAA	COL9A1F1-R: ATG ATC TTA TGC CAC TGG GAA T
Isoform 1	COL9A1F2-F: AAC AGC AGC CTT TTC GAA TTT	COL9A1F2-R: AGG GTC ACC AAC AGG TCC TAC
Isoform 2 specific	Iso2-F: ACG CTT TGG CAA CCG CTA CTC	Iso2-R: ACA AGC AGA GCC CCA ACA
COL9A1	COL9A1F3-F: GTA TTC CTG GAC CCC CTG GT	COL9A1F3-R: CTA CCC TTT TCA CCA GCA ACA C
COL9A1	COL9A1F4-F: ATT GCA GGG GTT ACC AGG TCT A	COL9A1F4-R: GAG TCT GGA CGC TTA AGA CTG G
COL9A1	COL9A1F5-F: ATT AAG CAG GTT TGC ATG AGA GT	COL9A1F5-R: GCA CCG TTC TTC TAT ATG TGA TTG
COL9A1	COL9A1F6-F: GAA CCA CGC CTG GTG AAG	COL9A1F6-R: CCA AAC AGT TGT TTT TGT TTA TTG G
Novel 3′ UTR	COL9A1novel -F: AGT TGC AGT AGT TAT TTT CAT TTA G	COL9A1novel-R: AAT AAG CTC AGT AAC TTC TTT TCA TTG

Denhardt's solution, 10% dextran sulfate, and 10 mM dithiothreitol for 2 h at 60°C. ^{35}S-labeled antisense and sense RNA probes (3×10^6 cpm/slide) were then added separately to the hybridization solution and incubated for 12–18 h at 60°C. After hybridization, sections were washed for 20 min in 2× SSC, 10 mM β-mercaptoethanol, and 1 mM EDTA, treated with RNase A (20 μg/mL) for 30 min at room temperature, and washed at high stringency (0.1× SSC, 10 mM β-mercaptoethanol, and 1 mM EDTA) for 2 h at 60°C. Sections were dehydrated, dipped in photographic emulsion NTB-3 (Kodak), and maintained at 4°C for 2–4 weeks of exposure. Sections were then developed and counterstained with hematoxylin and eosin (H&E), and coverslips were applied with 50% Canada balsam in methyl salicylate (Sigma, St. Louis, MO, USA).

Identification of polyadenylation sites in the 3′ UTR of *COL9A1*

Fifteen hundred bases of sequence of the 3′ UTR starting from the first base after the stop codon of *COL9A1* were retrieved from the Ensembl genome browser (Gene ID no. ENSG00000112280) and analyzed to identify the number and location of polyadenylation sites using the ESTparser program http://tagc.univ-mrs.fr/bioinfo/ESTparser/ (Beaudoing and Gautheret 2001). ESTparser predicts potential polyadenylation signals in the sequence, performs BLAST searches against the EST databases, and filters the output to produce a general picture of alternative polyadenylated forms and the tissues in which they occur. The sequence was also compared to the human EST database to retrieve *COL9A1* 3′ UTR-specific ESTs expressed in various tissues. Because the actual 3′ ESTs are not consistently annotated in the database, we used criteria proposed by Beaudoing et al. (2000) to select ESTs for further analysis: (1) ESTs of at least 40 nucleotides (nt) matching the *COL9A1* 3′ UTR with at least 95% similarity over the entire length of the EST sequence, except for the allowed 25 and 5 nt mismatches at the EST 5′ and 3′ ends; (2) internal priming assessed by adenine stretches in the *COL9A1* 3′ UTR flanking the 3′ extremity of the EST [i.e., internal priming as a result of cDNA oligo d(T) primers hybridizing to internal poly (A) stretches instead of the actual poly (A) tail]; and (3) ESTs with poly (A) or poly (T) termini (in antisense strand) of length 10 nt or more at the end of the transcript. ESTs that fulfilled the first two criteria were aligned to their corresponding positions in the 3′ UTR of *COL9A1* to identify tissue-specific polyadenylation sites.

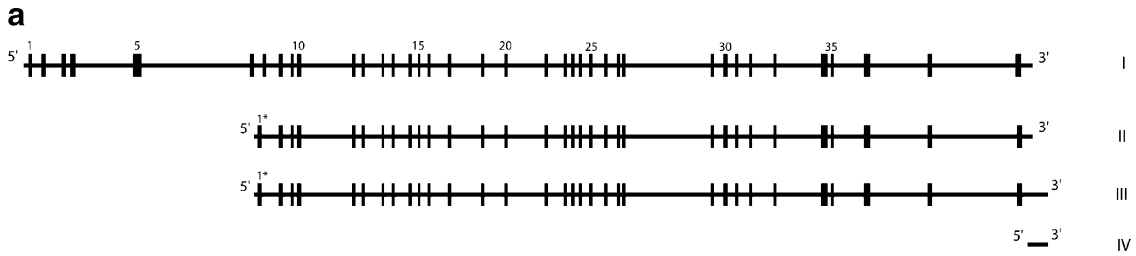
A large fraction of human polyadenylation sites are flanked by upstream and downstream sequence elements, located around positions 0 to −50 and +20 to +60, respectively, relative to the polyadenylation signal and distinguishing true polyadenylation sites from randomly occurring AAUAAA hexamers. The vicinity of the polyadenylation sites in the 1.5-kb 3′ UTR of *COL9A1* was also analyzed with the ERPIN program (http://tagc.univ-mrs.fr/erpin/) to identify true polyadenylation sites within this sequence based on sequence present upstream of the polyadenylation signal and downstream of the cleavage site (Legendre and Gautheret 2003).

Identification of regulatory motif in the novel 3′ UTR

About 1.5 kb of sequence from the 3′ UTR of *Col9a1* was retrieved from the Ensembl genomic sequences of other species, including mouse (ENSMUSG00000026147), rat (ENSRNOG00000012920), dog (ENSCAFG00000002589), chicken (ENSGAGG00000015970), fugufish (SINFRUG00000146962), pufferfish (GSTENG0030712001), and western clawed frog (ENSEXTG00000017926). To identify conserved regulatory elements, the 3′ UTR of *COL9A1* was compared with that of its orthologs retrieved from Ensembl as well as with EST databases of mouse and other species using BLAST (Altschul et al. 1997). The microRNA database (http://www.microrna.org) was searched for the presence of a microRNA target in the 3′ UTR of *COL9A1* and in its rat and mouse orthologs. The 3′ UTR of the canine ortholog was also evaluated for any conserved microRNA target sequences by using the miRranda computer program (John et al. 2004).

RESULTS

DNA sequencing and subsequent cluster analysis of the human fetal cochlear ESTs identified groups of overlapping clones comprising over 2,200 clusters. One of these clusters contained 21 ESTs derived from 17 overlapping clones. EST N66408, representing this cluster, is 575 bp in size and initial BLAST analysis of this sequence showed no homology to any known gene in the human and nonhuman mammalian databases. DNA sequencing of the original clone from which this EST was derived yielded an ~0.7-kb insert. BLAST analysis of the entire sequence revealed homology of 125 bp of the 5′ end to the 3′ UTR of exon 38 of *COL9A1* and 575 bp of novel sequence (Fig. 1b).



b

```
2041 tgagcgaggc cccccagggc tggcaggaat tcctggagtg cctggacccc cgggacctcc
2101 tgggcttccc ggtttctgtg agccagcctc ctgcaccatg caggctggtc agcgagcatt
2161 taacaaaggg cctgaccctt gaaaggctta ctgctgcatg gctgtctgca tgaaccacgc
2221 ctggtgaagg agcctgggtg agaaacacca tccaaagctg gggcaaagat gattaccttc
2281 agcatgatta caatgtatta ccttcagtat gattacagaa gtcctacttg acaatcacat
2341 atagaagaac ggtgctattc agtaagttct ctttcctttc ccttggaggg aagacagcag
2401 agtcatcagt taaaaaaaaa aaaaaagaaa accaaacacc tcccttgaac aaatttatac
2461 tcctgttccc aggatcttga gctttagtgt gctataccta tgtgtcttat cgtgggccac
2521 tgtgccAATA AAcaaaaaca actgtttggt ttacctcagt tgcagtagtt attttcattt
2581 agaagttgtt ctcagattat tgtttcagtt atatagagga ttactagact agttatgaag
2641 aaccccacta cattcaatgg aattggtgct taaaatctca tcgatgtgct gtctctggag
2701 tgataagaaa gggctacatc tcccgaaatg atttctttac gtcatgtatt ggtttccttc
2761 ttcaccttga acttttgttg aactgtatgt actttacccc aaacctgtta atattttgag
2821 cgcttctatg tgaaagcaaa gaaataattt taatactcag gcattcataa attttattga
2881 tgagattatt tattttaaag gtttgaggta acatctctgg ttgtaccaaa gAAGAAATAA
2941 Atatggtttc ttaatctctt gcatgttttc ttataaataa tcatgttcaa tgaaaagaag
3001 ttactgagct tatttagatc ATTAAAcatt actt
```

FIG. 1. (**a**) Schematic representation of *COL9A1*. I. Long form of *COL9A1*. Vertical boxes represent coding exons. The 5′ region of the long form encoded by exons 1–7 spliced to exon 8 encodes an amino-terminal nontriple-helical domain of 268 residues, including an NC4 domain of 245 amino acids and a putative signal peptide of 23 amino acids. II. Short form of *COL9A1*. The 5′ region is encoded by an alternative exon 1* present in intron 6 and noncoding exon 7 of the long form spliced to exon 8; the resulting protein lacks the NC4 domain and contains only 25 amino acid residues of nontriple-helical sequence at the amino terminal end. III. Short form of *COL9A1* with novel 3′ UTR expressed in human fetal cochlea. IV. Alignment of EST N66408 to the isoform of *COL9A1* detected in human fetal cochlea (III). (**b**) Nucleotide sequence of partial cDNA of *COL9A1*. Sequence in bold lower case type represents the novel 3′ UTR discovered in the human fetal cochlear cDNA library. ***tga*** (bold and italized) denotes the stop codon. Capitalized and italicized letters indicate the three putative polyadenylation signals.

Northern blot analysis

To determine the size as well as the relative level of expression of the full-length transcript from which cochlear EST N66408 was derived, a Northern blot analysis of a panel of human fetal tissues was performed. A transcript of about 3.5 kb in length was observed to be highly expressed in cochlea following overnight exposure of the filter (lane 3, Fig. 2A). N66408 was either expressed at a much lower level or was undetectable in other tissues analyzed. An extended exposure of the blot revealed very low level expression in testis and kidney and confirmed the absence of expression in other human fetal tissues including brain, thymus, spleen, lung, tongue, liver, bone marrow, and skeletal muscle (Fig. 2A).

RT-PCR analysis

Five overlapping sets of short form *COL9A1* transcript-specific primers (Table 1, Fig. 1a) used for RT-PCR of cDNA synthesized from total RNA extracted from 21-gestational-week-old human fetal cochlea revealed expression of the short form (data not shown). Furthermore, sequence analysis showed that this transcript contained the novel 3′ UTR with no recognized sequence similarity to *COL9A1* transcripts. RT-PCR analysis performed to assess expression of the novel 3′ UTR sequence in human fetal tissues showed an intense amplification product of this sequence in cochlea and in patella, a notably less intense product in testis, brain, eye, trachea, and skeletal muscle, and no detectable expression in placenta, epiglottis, and skull (Fig. 2B). RT-PCR was also carried out to evaluate expression of both long and short forms of *COL9A1* in various tissues using 5′ end isoform-specific primers. Expression of the long form of *COL9A1* was seen in trachea, patella, cochlea, and skull (Fig. 2C). The short form of *COL9A1* was found to be expressed in human fetal cochlea, eye, patella, testis, brain, skeletal muscle, epiglottis, trachea, and skull, with no detectable expression in placenta

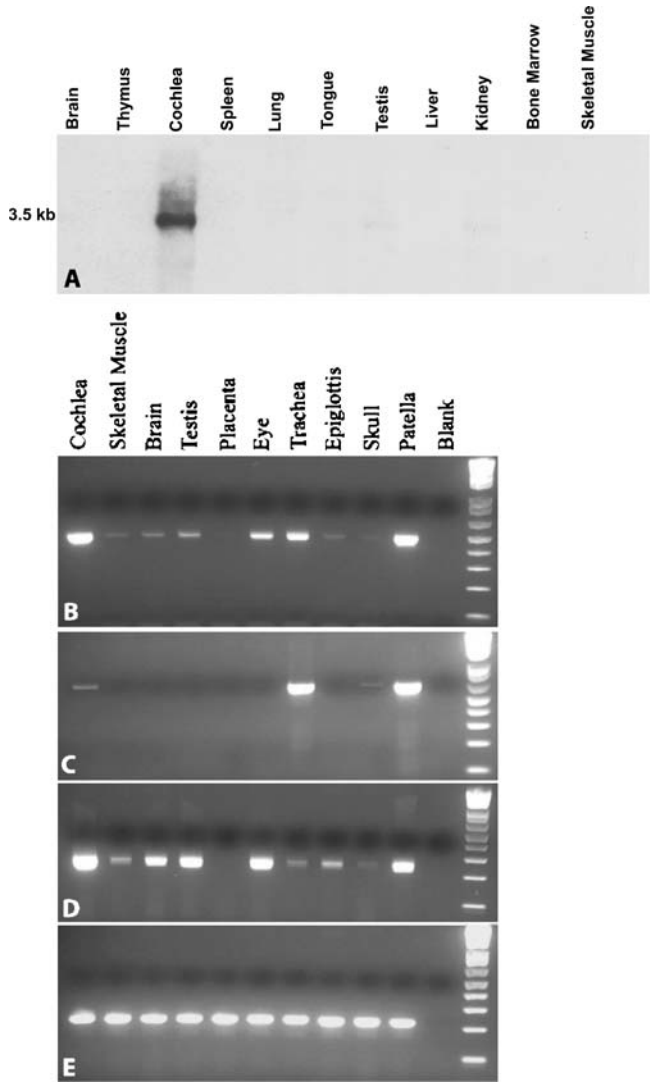


FIG. 2. (**A**) Autoradiogram of Northern blot analysis of human fetal RNA samples hybridized with EST N66408. High level expression was detected in cochlea; low level expression was observed in testis and kidney after an extended exposure of the blot (not shown). (**B**) RT-PCR analysis of novel 3′ UTR of *COL9A1* shows expression in all tissues except placenta. (**C**) RT-PCR analysis of the long form of *COL9A1* with isoform-specific primers detects expression in cochlea, trachea, skull, and patella. (**D**) RT-PCR analysis of the short form of *COL9A1* using isoform-specific primers reveals expression in all tissues except placenta. (**E**) RT-PCR analysis of *GAPDH* demonstrates bands of equal intensity in all tissues.

(Fig. 2D). For a positive control, RT-PCR analysis using *GAPDH*-specific primers showed products of equal intensities in all tissues, confirming the quality of cDNAs (Fig. 2E).

In situ hybridization

Tissue *in situ* hybridization was performed by using an ^{35}S-labeled riboprobe prepared from the novel 3′ UTR sequence of the short form of *COL9A1* (Fig. 2A) to cochlear sections from 20-gestational-week-old human fetal inner ear, and a strong hybridization signal was observed in the spiral limbus and spiral ligament region (Fig. 3). Lower level expression was seen in the organ of Corti and stria vascularis. The *in situ* hybridization studies with antisense probe did not reveal any signal in human fetal cochlear sections (data not shown).

Multiple polyadenylation signals in the 3′ UTR of *COL9A1*

Analysis of a 1.5-kb region from the 3′ UTR of *COL9A1* with the ESTparser program detected 17 potential polyadenylation signals: an AATAAA hexamer at two sites and seven of its single-base variants in 15 additional sites. A single nucleotide polymorphism at position 266 was found to result in an additional primary polyadenylation signal at position 264–269 ($AA^{C}/_{T}AAA$). The ESTparser program

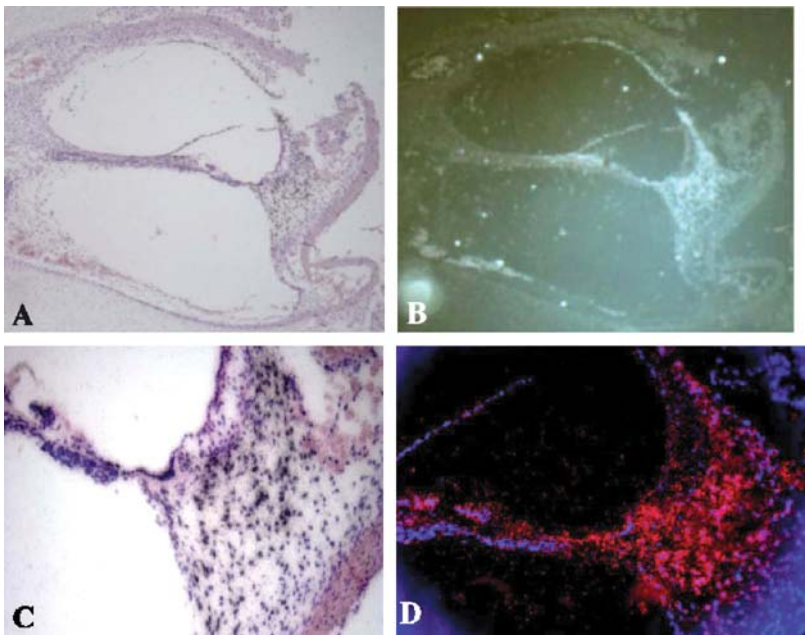


FIG. 3. Tissue *in situ* hybridization of novel 3′ UTR sequence of the short form of *COL9A1* on human fetal cochlear sections. (A) and (B) show cross sections (4.5×) of the human cochlear duct at 20 weeks gestational age. Images A and B, captured in bright and dark field, respectively, show abundant *COL9A1* expression in the spiral limbus and spiral ligament. (C) Magnified image (20×), in bright field, showing expression of the novel 3′ UTR transcript in the spiral ligament and organ of Corti. (D) Expression of the 3′ UTR of *COL9A1* transcripts in the spiral ligament, organ of Corti and spiral limbus captured in dark and bright field (magnified 20×). The red color represents the signals captured in dark field and pseudocolored, and the blue color indicates the tissue morphology, stained with hematoxylin and eosin, captured in bright field and pseudocolored.

detected only 19 ESTs from dbEST with sequence identity/similarity to the 1.5-kb region of the 3′ UTR of *COL9A1*; the majority of ESTs were found to have putative polyadenylation sites either at positions 774 or 853.

Our search of dbEST as well as the Unigene database showed 63 ESTs corresponding to the 1.5-kb sequence from the 3′ UTR of *COL9A1*. Of these, 23 ESTs were derived from the cochlear library, 13 were chondrosarcoma-specific, eight were from pooled tumors, six were eye-specific, and the remaining 13 were identified from inner and middle ear, pooled skin, placenta, colon cancer, malignant melanoma, pooled fetal tissue, and fetal heart libraries. Alignment of these ESTs to the 3′ UTR of *COL9A1* demonstrated no tissue-specific 3′ end other than for cochlea; four of the cochlear ESTs had a tissue-specific 3′ end. Twenty-nine of these ESTs were probably derived from a poly (A) stretch in genomic DNA, and these ESTs were not included for further analysis. Six of the remaining 34 ESTs were duplicates derived from three different clones and three of these duplicates were excluded. Analysis of the 3′ end of the remaining 31 ESTs (Fig. 4a) showed no poly (A)/poly (T) stretches in 12; these ESTs were not considered for prediction of polyadenylation sites.

A total of 19 ESTs with poly (A)/poly (T) tails were evaluated to infer the cleavage sites (Table 2). Of these 19 ESTs, 14 (70%) were found to be derived from one of the two overlapping polyadenylation signals (AAGAAA/AATAAA) located at position 749–758, three ESTs (including N66408) were from the ATTAAA hexameric variant at position 839–844, and two were from the first AATAAA polyadenylation signal located at position 343–348. Transcripts derived from the latter putative polyadenylation site represent the long or short forms of *COL9A1*. Ten of the 19 ESTs were from human fetal cochlea, seven of which were derived from the overlapping signals present at position 749–758. Analysis of mouse ESTs-specific to the 1.5-kb 3′ UTR of *Col9a1* showed only five ESTs fulfilling all criteria (data not shown) and derived from three different polyadenylation signals, one of which corresponds to the AATAAA hexamer at overlapping polyadenylation signals at position 749–758.

Analysis of 3′ UTR sequences revealed two AAUAAA sites at positions 343–348 and 753–758 as positive sites with scores of 12.14 (E value of 4.42×10^{-4}) and 13.86 ($E = 4.01 \times 10^{-5}$), respectively. The score is computed directly from the lodscore profile and E is the number of solutions expected by chance

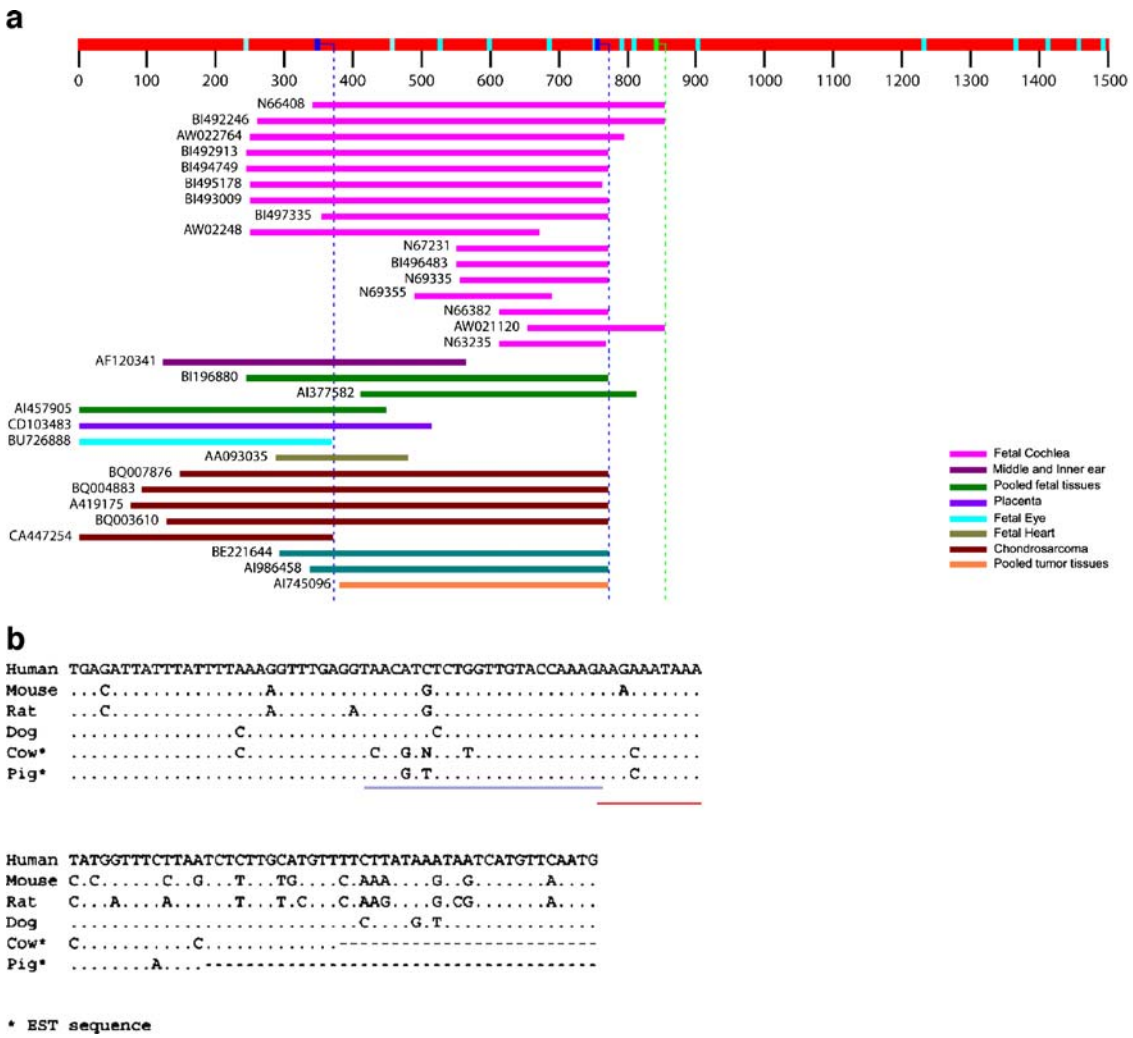


FIG. 4. (**a**) Alignment of ESTs specific to the 3′ UTR of *COL9A1*. The top line in red represents the query 3′ UTR sequence of *COL9A1* obtained from the genomic sequence and the boxes within the red line show the position of the potential polyadenylation signals. The dark blue box represents the AATAAA hexamer, the light blue boxes indicate the positions of hexameric variants, and the green box shows the ATTAAA signal at position 839–844. Overlapping light and dark blue boxes represent the overlapping polyadenylation signals AAGAAATAAA at position 749–758. The vertical broken line denotes the putative polyadenylation sites and indicates the first "A" nucleotide in the poly (A) stretch. (**b**) Highly conserved 3′ UTR of *COL9A1* in mammals. The blue underlined region represents the highly conserved miR-9 target sequence in the novel 3′ UTR and the red underlined region represents the overlapping polyadenylation signals AAGAAATAAA at position 749–758 starting from the first nucleotide immediately following the stop codon. The overlapping polyadenylation signals are not found in the *Col9a1*-specific EST sequences obtained from cow and pig as AAGCAATAAA is present at this position. Dots denote nucleotide sequence identity, and dashes in the cow and pig sequence represent the end of the available sequence from the corresponding ESTs.

TABLE 2

Distribution of polyadenylation sites in *COL9A1* ESTs

Putative polyadenylation site	*Supporting ESTs*
Site 1 at position 371	Eye: BU726888, Chondrosarcoma: CA447254
Site 2 at position 774	Cochlea: BI492913, BI494749, BI495178, BI493009, BI497335, BI496483, N69335, BQ007876 Chondrosarcoma: BQ004883, CA419175, BQ003610 Malignant melanoma: AI986458 Pooled tumors: AI745096 Pooled fetal tissues: BF196880
Site 3 at position 853	Cochlea: N66408, BI492246, AW021120

at this score level. These *E* values indicate that these two sites are very likely to be true sites (Gautheret, personal communication).

miR-9 target sequence in the 3′ UTR

The 3′ UTR of *COL9A1* was compared with that of its orthologs from mouse, rat, dog, chicken, fugufish, zebrafish, and pufferfish. This analysis revealed ~80% to 90% homology at two regions within the novel 3′ UTR in dog, mouse, and rat. One of these sites contained the highly conserved overlapping polyadenylation signals, where most of the transcripts were found to be cleaved during polyadenylation. This region was also conserved in pig and cow (Fig. 4b). We did not observe any such similarity between the 3′ UTR of human *COL9A1* and its orthologs in chicken, fugufish, zebrafish, and pufferfish. A search for regulatory motifs in the 3′ UTR of *COL9A1*, especially within the highly conserved region, revealed a highly conserved target sequence for one of the microRNAs, miR-9, immediately preceding the overlapping polyadenylation signals at position 749–758 in the novel 3′ UTR of *COL9A1* and conserved orthologs in mouse and rat. Analysis of the 3′ UTR of the dog ortholog with the miRranda program demonstrated a highly conserved potential miR-9 target sequence, which is also highly conserved in cow and pig ESTs homologous to the 3′ UTR of *COL9A1*.

DISCUSSION

Analysis of EST N66408 revealed homology to *COL9A1* at the 5′ end and showed no homology to any known gene transcript in the remainder of the sequence. An independent cDNA library prepared with human fetal cochlear RNA obtained from our cochlear RNA collections, using suppression subtractive hybridization against a mixture of cDNAs from various fetal tissues to increase the proportion of tissue-specific genes in the cDNA library, also revealed abundant expression of *COL9A1*-specific clones (Luijendijk et al. 2003).

Type IX collagen is one of the important components in the tectorial membrane of the organ of Corti, along with three other collagens (collagen types II, V, and XI) and noncollagenous proteins (alpha-tectorin, beta-tectorin, otogelin, and glycoproteins) (Goodyear and Richardson 2002). Type IX collagen is a heterotrimer of polypeptide chains encoded by three genes, *COL9A1*, *COL9A2*, and *COL9A3*, and these chains fold into the triple-helix structure, an important characteristic feature of members of the collagen family of extracellular matrix proteins (Olsen 1997). The triple helix of *COL9A1* consists of COL1, COL2, and COL3 domains, which are separated and flanked by noncollagenous segments, domains NC1–NC4. *COL9A1* is synthesized in two forms (Nishimura et al. 1989): a short form of about 2.6 kb and a long form of approximately 3.3 kb. The *COL9A1* long form (also known as v1 or *COL9A1*_v1), found in hyaline cartilage, is transcribed from an upstream promoter and generates a product that contains an additional N-terminal globular domain of 245 amino acids, called NC4 (Wu et al. 1992). In contrast, the *COL9A1* short form (v2 or *COL9A1*_v2), expressed in vitreous and several other tissues (Svoboda et al. 1988; Fitch et al. 1988; Wright and Mayne 1988; Perris et al. 1991; Swiderski and Solursh 1992), uses an alternative downstream promoter and an alternative exon; the globular NC4 domain is replaced with a short alternative amino acid sequence at the N-terminal end (Nishimura et al. 1989). The two other polypeptide chains, transcribed from *COL9A2* and *COL9A3*, also lack this NC4 domain (Mayne and Brewton 1993).

Expression of the novel short form of *COL9A1* in cochlea

Northern blot analysis of RNAs from various human tissues showed a strong hybridizing band corresponding to an ~3.5-kb transcript in fetal cochlea. Further RT-PCR analysis confirmed expression of the full-length transcript of the short form of *COL9A1* with the novel 3′ UTR, representing a new alternative transcript of *COL9A1*, designated as v3 or *COL9A1*_v3. Sequence analysis of 22/26 *COL9A1*-specific clones obtained from a suppression subtracted human fetal cochlear cDNA library (Luijendijk et al. 2003) also showed the novel 3′ UTR in 15 clones. In addition to human fetal cochlea, several other tissues showed expression of this novel 3′ UTR by RT-PCR analysis. Cochlea, trachea, skull, and patella were found to express both short and long forms; it remains to be determined whether the novel 3′ UTR is specific to the short form or present in both isoforms. This new variant with the novel 3′ UTR might have been derived from alternative polyadenylation sites present in the 3′ UTR. Northern blot analyses of RNA from chicken sternal cartilage was previously reported to show two different transcripts of *Col9a1* produced by usage of alternative polyadenylation sites (Svoboda et al. 1988).

Expression of the *COL9A1* short form with the novel 3′ UTR (v3) has not been described previously, although expression of the short isoform (v2) has been detected in other tissues including chicken cornea (Svoboda et al. 1988; Fitch et al. 1988; Nishimura et al. 1989), vitreous humor (Wright and Mayne 1988; Yada et al. 1990), notochord, axial skeleton, chondrocranium, Meckel's cartilage (Swiderski

and Solursh 1992), and perinotochordal matrix (Perris et al. 1991). Swiderski and Solursh (1992), while comparing the spatial and temporal expression patterns of transcripts of avian type II collagen and the long and short forms of the alpha 1 type IX collagen by *in situ* hybridization, observed expression of the short form of *Col9a1* in the developing chicken (stages 25–28) nonchondrogenic notochord along with type II collagen mRNA. They also found expression of the short form of *Col9a1* to be more restricted in the developing chondrogenic vertebrae of the axial skeleton, chondrocranium, and Meckel's cartilage, although expression of both long and short forms of *Col9a1* was seen along with type II collagen transcripts (Swiderski and Solursh1 1992). Using Northern blot and RNase protection assays, Savontaus et al. (1998) detected expression of both isoforms and only the short form of *Col9a1* in the developing mouse embryonic limbs and eyes, respectively. Expression of the short form of *Col9a1* was also observed in bovine nucleus pulposus, the gel-like central zone of the young intervertebral disk that has a similar collagen phenotype to that of hyaline cartilage (Wu and Eyre 2003).

Expression of novel 3′ UTR of *COL9A1* in human fetal cochlear sections

Our tissue *in situ* hybridization studies detected abundant expression of the novel 3′ UTR of *COL9A1* in the spiral ligament and limbus. The high level and widespread expression of the novel 3′ UTR of *COL9A1* in the cochlea is in agreement with a strong hybridizing band in human fetal cochlea by Northern blot analysis. Immunolocalization studies of type IX collagen have demonstrated its expression in the tectorial membrane, an acellular structure that overlies the cochlear sensory hair cells and plays an important role in the transduction process, along with type II collagen in the mouse inner ear (Slepecky et al. 1992; Goodyear and Richardson 2002; Suzuki et al. 2005). The tectorial membrane is composed of two different components: the radial collagen fibril and striated sheet matrix (Goodyear and Richardson 2002). Immunoelectron microscopic findings showed that types II and IX collagens colocalize in collagen fibril bundles, indicating the contribution of type IX collagen to the three-dimensional integrated structure of type II collagen molecules (Asamura et al. 2005).

Multiple polyadenylation sites in the 3′ UTR of *COL9A1*

About 29% of human mRNAs have multiple polyadenylation sites (Beaudoing et al. 2000) and the choice of poly (A) site usage may influence stability, translational efficiency, or localization of an mRNA in a tissue- or disease-specific manner (Edwalds-Gilbert et al. 1997). Polyadenylation sites are primarily defined by a polyadenylation signal of an AATAAA hexamer or its single-base variants, located between 10 and 35 nt upstream of the actual cleavage and polyadenylation site. Analysis of ESTs specific to the *COL9A1* 3′ UTR revealed three putative polyadenylation sites present within the novel 3′ sequence expressed in cochlea at positions 371, 774, and 853 following the stop codon. Our analysis demonstrated that position 774 was the site used most commonly and that this site comprises two overlapping polyadenylation signals, AAGAAA/AATAAA. It is unknown which of these two overlapping signals was involved in cleavage and polyadenylation. However, given that the AATAAA hexamer is the most commonly observed polyadenylation signal (58.2%) (Colgan and Manley 1997; Beaudoing et al. 2000) and AAGAAA is the least commonly observed variant, it seems likely that AATAAA is involved in the cleavage and polyadenylation of the transcripts at position 774. Three cochlear ESTs were derived from the putative polyadenylation site at 853. It is unknown whether expression of two alternative transcripts in cochlea is cellular-specific or whether expression of both isoforms occurs in all cell types within the cochlea. Because the difference between the two transcripts of *COL9A1* is about 80 bases in size, it is understandable as to why two distinct bands were not detectable by Northern blot analysis of human fetal cochlear RNA.

Highly conserved miR-9 target in the novel 3′ UTR

A high degree of sequence conservation seen in mammals in the 3′ UTR of *COL9A1*, particularly around one of the polyadenylation signals, suggests the presence of important key regulatory elements necessary for proper polyadenylation. A search of the microRNA (miRNA) database (http://www.microrna.org), as well as an analysis using miRanda software (John et al. 2004) to identify potential regulatory motifs in the 3′ UTR of *COL9A1*, revealed a putative target sequence for one of the microRNAs, miR-9, in the immediate upstream vicinity of the AATAAA hexameric signal at position 753–758. MicroRNAs are about 22 nt single-stranded RNAs that regulate expression of target genes by interacting with complementary sites in the 3′ UTR of the target mRNAs (Bartel 2004). When target sites are not perfectly matched to the miRNA, which is seen in the majority of 3′ UTRs, the miRNA suppresses translation of the mRNA by an unknown mechanism (Zeng et al. 2002;

Zeng and Cullen 2003). In cases with perfect or near-perfect sequence matches, miRNAs mediate cleavage of the target mRNA (Yekta et al. 2004). These miRNAs have critical roles in a variety of biological processes. In *Caenorhabditis elegans*, the *let-7* RNA acts to promote the transition from the late-larval to adult cell fates in the same manner that *lin-4* RNA acts earlier in development to promote progression from the first larval stage to the second stage (Reinhart et al. 2000; Slack et al. 2000). In mammals, miRNAs were shown to be involved in developmental regulation of subsets of lymphoid lineages (Chen et al. 2004). As of February 2006, 326 human miRNA sequences giving rise to 319 unique mature miRNAs are found in the microRNA database (miRBase) http://www.sanger.ac.uk/Software/Rfam/mirna/index.shtml. miR-9 is encoded by three genes, *miR-9-1*, *miR-9-2*, and *miR-9-3*, located on human chromosomes 1, 5, and 15, respectively (Lagos-Quintana et al. 2002). Recently, miR-9 was shown to play an important functional role during neural cell differentiation (Krichevsky et al. 2003, 2005; Smirnova et al. 2005).

As many miRNA target sequences are conserved across large evolutionary distances, they must be subjected to strong functional constraints (John et al. 2004). The target sequence, which we found in the 3′ UTR of the short form of *COL9A1,* was highly conserved in dog, cow, pig, mouse, and rat, and this region was shown to have a potential candidate target for miR-9 (John et al. 2004). The expression pattern of miR-9 in inner ear and its functional importance remain to be elucidated.

Importance of *COL9A1* in hearing

To date, mutations and polymorphisms in *COL9A1* have not been reported with any form of hearing loss. However, a small insertional mutation in intron 8 has been shown in affected members of a family with autosomal dominant multiple epiphysial dysplasia (Czarny-Ratajczak et al. 2001). Recently, a knockout mouse model lacking both isoforms of *Col9a1* showed progressive hearing loss as well as an abnormality in the shape of the tectorial membrane (Suzuki et al. 2005). Electron microscopic analysis of the tectorial membrane revealed disorganized, aggregated, and fused fiber-like structures in the 1-month-old homozygous knockout mice (Suzuki et al. 2005). Deficiency of alpha 1 type IX collagen was also found to lead to a functional knockout of the other two *Col9* polypeptide chains, alpha 2 and alpha 3, which were transcribed normally (Hagg et al. 1997). Type II collagen, which is colocalized with type IX collagen, was likewise not immunohistochemically detected in the tectorial membrane of the homozygous *Col9* knockout mice, suggesting that lack of type IX collagen may also affect the three-dimensional structure of type II collagen (Asamura et al. 2005). Humans and mice with mutations in *COL2A1* have hearing loss associated with several other phenotypic abnormalities (Richards et al. 2000; Donahue et al. 2003; Lopponen et al. 2004). Given the expression pattern of *COL9A1* and the electrophysiological and morphological abnormalities of the inner ear shown in the homozygous knockout mice for type IX collagen, it is likely that type IX collagen plays an important functional role along with type II collagen in the development and function of the inner ear, and that mutations in *COL9A1* may be expected to result in hearing loss in humans.

ACKNOWLEDGMENTS

The authors wish to recognize the support of this research by the National Institutes on Deafness and Other Communication Disorders, DC03402 (to C.C.M) and F32DC00405 (to B.L.R). Also, the authors thank Dr. Hannie Kremer for communication of the sequence analysis of the *COL9A1*-specific clones from the suppression subtracted human fetal cochlear cDNA library, Drs. Massimo Loda and Yu Yang for technical assistance with the *in situ* hybridization provided through the Dana Farber–Harvard Cancer Center *in situ* hybridization core, Dr. Charles Lee for his help in photography of the cochlear *in situ* hybridization, and Drs. Bjorn Olsen, Kirk Beisel and Garret Soukup for critical review of the manuscript.

REFERENCES

AHMAD NN, ALA-KOKKO L, KNOWLTON RG, JIMENEZ SA, WEAVER EJ, MAGUIRE JI, TASMAN W, PROCKOP DJ. Stop codon in the procollagen II gene (COL2A1) in a family with the Stickler syndrome (arthro-ophthalmopathy). Proc. Natl. Acad. Sci. U. S. A. 88:6624–6627, 1991.

ALTSCHUL SF, MADDEN TL, SCHAFFER AA, ZHANG J, ZHANG Z, MILLER W, LIPMAN DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402, 1997.

ASAMURA K, ABE S, IMAMURA Y, ASZODI A, SUZUKI N, HASHIMOTO S, TAKUMI Y, HAYASHI T, FASSLER R, NAKAMURA Y, USAMI S. Type IX collagen is crucial for normal hearing. Neuroscience 132:493–500, 2005.

BARKER DF, HOSTIKKA SL, ZHOU J, CHOW LT, OLIPHANT AR, GERKEN SC, GREGORY MC, SKOLNICK MH, ATKIN CL, TRYGGVASON K. Identification of mutations in the COL4A5 collagen gene in Alport syndrome. Science 248:1224–1227, 1990.

BARTEL DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297, 2004.

BEAUDOING E, GAUTHERET D. Identification of alternate polyadenylation sites and analysis of their tissue distribution using EST data. Genome Res. 11:1520–1526, 2001.

BEAUDOING E, FREIER S, WYATT JR, CLAVERIE JM, GAUTHERET D. Patterns of variant polyadenylation signal usage in human genes. Genome Res. 10:1001–1010, 2000.

BEISEL KW, SHIRAKI T, MORRIS KA, POMPEIA C, KACHAR B, ARAKAWA T, BONO H, KAWAI J, HAYASHIZAKI Y, CARNINCI P. Identification of unique transcripts from a mouse full-length, subtracted inner ear cDNA library. Genomics 83:1012–1023, 2004.

CHEN CZ, LI L, LODISH HF, BARTEL DP. MicroRNAs modulate hematopoietic lineage differentiation. Science 303:83–86, 2004.

CHIRGWIN JM, PRZYBYLA AE, MACDONALD RJ, RUTTER WJ. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299, 1979.

COLGAN DF, MANLEY JL. Mechanism and regulation of mRNA polyadenylation. Genes Dev. 11:2755–2766, 1997.

CZARNY-RATAJCZAK M, LOHINIVA J, ROGALA P, KOZLOWSKI K, PERALA M, CARTER L, SPECTOR TD, KOLODZIEJ L, SEPPANEN U, GLAZAR R, KROLEWSKI J, LATOS-BIELENSKA A, ALA-KOKKO L. A mutation in COL9A1 causes multiple epiphyseal dysplasia: further evidence for locus heterogeneity. Am. J. Hum. Genet. 69:969–980, 2001.

DEVEREUX J, HAEBERLI P, SMITHIES O. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12:387–395, 1984.

DONAHUE LR, CHANG B, MOHAN S, MIYAKOSHI N, WERGEDAL JE, BAYLINK DJ, HAWES NL, ROSEN CJ, WARD-BAILEY P, ZHENG QY, BRONSON RT, JOHNSON KR, DAVISSON MT. A missense mutation in the mouse Col2a1 gene causes spondyloepiphyseal dysplasia congenita, hearing loss, and retinoschisis. J. Bone Miner. Res. 18:1612–1621, 2003.

EDWALDS-GILBERT G, VERALDI KL, MILCAREK C. Alternative poly(A) site selection in complex transcription units: means to an end? Nucleic Acids Res. 25:2547–2561, 1997.

FITCH JM, MENTZER A, MAYNE R, LINSENMAYER TF. Acquisition of type IX collagen by the developing avian primary corneal stroma and vitreous. Dev. Biol. 128:396–405, 1998.

GOODYEAR RJ, RICHARDSON GP. Extracellular matrices associated with the apical surfaces of sensory epithelia in the inner ear: molecular and structural diversity. J. Neurobiol. 53:212–227, 2002.

GRIFFITH AJ, GEBARSKI SS, SHEPARD NT, KILENY PR. Audiovestibular phenotype associated with a COL11A1 mutation in Marshall syndrome. Arch. Otolaryngol. Head Neck Surg. 126:891–894, 2000.

GURISH MF, BELL AF, SMITH TJ, DUCHARME LA, WANG RK, WEIS JH. Expression of murine beta 7, alpha 4, and beta 1 integrin genes by rodent mast cells. J. Immunol. 149:1964–1972, 1992.

HAGG R, HEDBOM E, MOLLERS U, ASZODI A, FASSLER R, BRUCKNER P. Absence of the alpha1(IX) chain leads to a functional knockout of the entire collagen IX protein in mice. J. Biol. Chem. 272:20650–20654, 1997.

HEDRICK SM, COHEN DI, NIELSEN EA, DAVIS MM. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature 308:149–153, 1984.

JOHN B, ENRIGHT AJ, ARAVIN A, TUSCHL T, SANDER C, MARKS DS. Human MicroRNA targets. PLoS Biol. 2:e363, 2004.

JONES DT, REED RR. Gold: an olfactory neuron specific G-protein involved in ororant signal transduction. Science 244:790–795, 1989.

KRICHEVSKY AM, KING KS, DONAHUE CP, KHRAPKO K, KOSIK KS. A microRNA array reveals extensive regulation of microRNAs during brain development. RNA 9:1274–1281, 2003.

KRICHEVSKY AM, SONNTAG KC, ISACSON O, KOSIK KS. Specific microRNAs modulate ES cell-derived neurogenesis. Stem Cells, 2005.

KUIVANIEMI H, TROMP G, PROCKOP DJ. Mutations in collagen genes: causes of rare and some common diseases in humans. FASEB J. 5:2052–2060, 1991.

KUIVANIEMI H, TROMP G, PROCKOP DJ. Mutations in fibrillar collagens (types I, II, III, and XI), fibril-associated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels. Hum. Mutat. 9:300–315, 1997.

LAGOS-QUINTANA M, RAUHUT R, YALCIN A, MEYER J, LENDECKEL W, TUSCHL T. Identification of tissue-specific microRNAs from mouse. Curr. Biol. 12:735–739, 2002.

LEGENDRE M, GAUTHERET D. Sequence determinants in human polyadenylation site selection. BMC Genomics 4:7, 2003.

LEMMINK HH, MOCHIZUKI T, VAN DEN HEUVEL LP, SCHRODER CH, BARRIENTOS A, MONNENS LA, VAN OOST BA, BRUNNER HG, REEDERS ST, SMEETS HJ. Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal recessive Alport syndrome. Hum. Mol. Genet. 3:1269–1273, 1994.

LOPPONEN T, KORKKO J, LUNDAN T, SEPPANEN U, IGNATIUS J, KAARIAINEN H. Childhood-onset osteoarthritis, tall stature, and sensorineural hearing loss associated with Arg75-Cys mutation in procollagen type II gene (COL2A1). Arthritis Rheum. 51:925–932, 2004.

LUIJENDIJK MW, VAN DE POL TJ, VAN DUIJNHOVEN G, DEN HOLLANDER AI, TEN CAAT J, VAN LIMPT V, BRUNNER HG, KREMER H, CREMERS FP. Cloning, characterization, and mRNA expression analysis of novel human fetal cochlear cDNAs. Genomics 82:480–490, 2003.

MAYNE R, BREWTON RG. New members of the collagen superfamily. Curr. Opin. Cell Biol. 5:883–890, 1993.

MCGUIRT WT, PRASAD SD, GRIFFITH AJ, KUNST HP, GREEN GE, SHPARGEL KB, RUNGE C, HUYBRECHTS C, MUELLER RF, LYNCH E, KING MC, BRUNNER HG, CREMERS CW, TAKANOSU M, LI SW, ARITA M, MAYNE R, PROCKOP DJ, VAN CAMP G, SMITH RJ. Mutations in COL11A2 cause non-syndromic hearing loss (DFNA13). Nat. Genet. 23:413–419, 1999.

MOCHIZUKI T, LEMMINK HH, MARIYAMA M, ANTIGNAC C, GUBLER MC, PIRSON Y, VERELLEN-DUMOULIN C, CHAN B, SCHRODER CH, SMEETS HJ ET AL. Identification of mutations in the alpha 3(IV) and alpha 4(IV) collagen genes in autosomal recessive Alport syndrome. Nat. Genet. 8:77–81, 1994.

MORTON NE. Genetic epidemiology of hearing impairment. Ann. N. Y. Acad. Sci. 630:16–31, 1991.

NISHIMURA I, MURAGAKI Y, OLSEN BR. Tissue-specific forms of type IX collagen-proteoglycan arise from the use of two widely separated promoters. J. Biol. Chem. 264:20033–20041, 1989.

OLSEN BR (1997) COLLAGEN IX. Int. J. Biochem. Cell Biol. 29:555–558.

PERRIS R, KROTOSKI D, BRONNER-FRASER M. Collagens in avian neural crest development: distribution *in vivo* and migration-promoting ability *in vitro*. Development 113:969–984, 1991.

REINHART BJ, SLACK FJ, BASSON M, PASQUINELLI AE, BETTINGER JC, ROUGVIE AE, HORVITZ HR, RUVKUN G. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 403:901–906, 2000.

RESENDES BL, ROBERTSON NG, SZUSTAKOWSKI JD, RESENDES RJ, WENG Z, MORTON CC. Gene discovery in the auditory system: characterization of additional cochlear-expressed sequences. J. Assoc. Res. Otolaryngol. 3:45–53, 2002.

RESENDES BL, KUO SF, ROBERTSON NG, GIERSCH AB, HONRUBIA D, OHARA O, ADAMS JC, MORTON CC. Isolation from cochlea of a novel human intronless gene with predominant fetal expression. J. Assoc. Res. Otolaryngol. 5:185–202, 2004.

RICHARDS AJ, YATES JR, WILLIAMS R, PAYNE SJ, POPE FM, SCOTT JD, SNEAD MP. A family with Stickler syndrome type 2 has a mutation in the COL11A1 gene resulting in the substitution of glycine 97 by valine in alpha 1 (XI) collagen. Hum. Mol. Genet. 5:1339–1343, 1996.

RICHARDS AJ, BAGULEY DM, YATES JR, LANE C, NICOL M, HARPER PS, SCOTT JD, SNEAD MP. Variation in the vitreous phenotype of Stickler syndrome can be caused by different amino acid substitutions in the X position of the type II collagen Gly–X–Y triple helix. Am. J. Hum. Genet. 67:1083–1094, 2000.

ROBERTSON NG, KHETARPAL U, GUTIERREZ-ESPELETA GA, BIEBER FR, MORTON CC. Isolation of novel and known genes from a human fetal cochlear cDNA library using subtractive hybridization and differential screening. Genomics 23:42–50, 1994.

ROBERTSON NG, SKVORAK AB, YIN Y, WEREMOWICZ S, JOHNSON KR, KOVATCH KA, BATTEY JF, BIEBER FR, MORTON CC. Mapping and characterization of a novel cochlear gene in human and in mouse: a positional candidate gene for a deafness disorder, DFNA9. Genomics 46:345–354, 1997.

ROBERTSON NG, LU L, HELLER S, MERCHANT SN, EAVEY RD, MCKENNA M, NADOL JB JR., MIYAMOTO RT, LINTHICUM FH JR., LUBIANCA NETO JF, HUDSPETH AJ, SEIDMAN CE, MORTON CC, SEIDMAN JG. Mutations in a novel cochlear gene cause DFNA9, a human nonsyndromic deafness with vestibular dysfunction. Nat. Genet. 20:299–303, 1998.

ROBERTSON NG, HELLER S, LIN JS, RESENDES BL, WEREMOWICZ S, DENIS CS, BELL AM, HUDSPETH AJ, MORTON CC. A novel conserved cochlear gene, OTOR: identification, expression analysis, and chromosomal mapping. Genomics 66:242–248, 2000.

SAVONTAUS M, IHANAMAKI T, PERALA M, METSARANTA M, SANDBERG-LALL M, VUORIO E. Expression of type II and IX collagen isoforms during normal and pathological cartilage and eye development. Histochem. Cell Biol. 110:149–159, 1998.

SIRKO-OSADSA DA, MURRAY MA, SCOTT JA, LAVERY MA, WARMAN ML, ROBIN NH. Stickler syndrome without eye involvement is caused by mutations in COL11A2, the gene encoding the alpha2(XI) chain of type XI collagen. J. Pediatr. 132:368–371, 1998.

SKVORAK AB, WENG Z, YEE AJ, ROBERTSON NG, MORTON CC. Human cochlear expressed sequence tags provide insight into cochlear gene expression and identify candidate genes for deafness. Hum. Mol. Genet. 8:439–452, 1999.

SLACK FJ, BASSON M, LIU Z, AMBROS V, HORVITZ HR, RUVKUN G. The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. Mol. Cell 5:659–669, 2000.

SLEPECKY NB, CEFARATTI LK, YOO TJ. Type II and type IX collagen form heterotypic fibers in the tectorial membrane of the inner ear. Matrix 12:80–86, 1992.

SMIRNOVA L, GRAFE A, SEILER A, SCHUMACHER S, NITSCH R, WULCZYN FG. Regulation of miRNA expression during neural cell specification. Eur. J. Neurosci. 21:1469–1477, 2005.

SUZUKI N, ASAMURA K, KIKUCHI Y, TAKUMI Y, ABE S, IMAMURA Y, HAYASHI T, ASZODI A, FASSLER R, USAMI S. Type IX collagen knock-out mouse shows progressive hearing loss. Neurosci. Res. 51:293–298, 2005.

SVOBODA KK, NISHIMURA I, SUGRUE SP, NINOMIYA Y, OLSEN BR. Embryonic chicken cornea and cartilage synthesize type IX collagen molecules with different amino-terminal domains. Proc. Natl. Acad. Sci. U. S. A. 85:7496–7500, 1998.

SWIDERSKI RE, SOLURSH M. Localization of type II collagen, long form alpha 1(IX) collagen, and short form alpha 1(IX) collagen transcripts in the developing chick notochord and axial skeleton. Dev. Dyn. 194:118–127, 1992.

THOMAS PS. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U. S. A. 77:5201–5205, 1980.

VAN STEENSEL MA, BUMA P, DE WAAL MALEFIJT MC, VAN DEN HOOGEN FH, BRUNNER HG. Oto-spondylo-megaepiphyseal dysplasia (OSMED): clinical description of three patients homozygous for a missense mutation in the COL11A2 gene. Am. J. Med. Genet. 70:315–323, 1997.

WINTERPACHT A, HILBERT M, SCHWARZE U, MUNDLOS S, SPRANGER J, ZABEL BU. Kniest and Stickler dysplasia phenotypes caused by collagen type II gene (COL2A1) defect. Nat. Genet. 3:323–326, 1993.

WRIGHT DW, MAYNE R. Vitreous humor of chicken contains two fibrillar systems: an analysis of their structure. J. Ultrastruct. Mol. Struct. Res. 100:224–234, 1988.

WU JJ, EYRE DR. Intervertebral disc collagen. Usage of the short form of the alpha1(IX) chain in bovine nucleus pulposus. J. Biol. Chem. 278:24521–24525, 2003.

WU JJ, WOODS PE, EYRE DR. Identification of cross-linking sites in bovine cartilage type IX collagen reveals an antiparallel type II-type IX molecular relationship and type IX to type IX bonding. J. Biol. Chem. 267:23007–23014, 1992.

YADA T, SUZUKI S, KOBAYASHI K, KOBAYASHI M, HOSHINO T, HORIE K, KIMATA K. Occurrence in chick embryo vitreous humor of a type IX collagen proteoglycan with an extraordinarily large chondroitin sulfate chain and short alpha 1 polypeptide. J. Biol. Chem. 265:6992–6999, 1990.

YEKTA S, SHIH IH, BARTEL DP. MicroRNA-directed cleavage of HOXB8 mRNA. Science 304:594–596, 2004.

ZENG Y, CULLEN BR. Sequence requirements for micro RNA processing and function in human cells. RNA 9:112–123, 2003.

ZENG Y, WAGNER EJ, CULLEN BR. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol. Cell 9:1327–1333, 2002.