PII: S0736-5748(96)00113-X

NOVEL GENES EXPRESSED IN THE CHICK OTOCYST DURING DEVELOPMENT: IDENTIFICATION USING DIFFERENTIAL DISPLAY OF RNA

TZY-WEN L. GONG,† ADRIAN D. HEGEMAN,† JOOYOUNG J. SHIN,†
KENDRA H. LINDBERG,‡ KATE F. BARALD‡ and MARGARET I. LOMAX†‡§

†Kresge Hearing Research Institute, Department of Otolaryngology/Head–Neck Surgery, University of Michigan Medical School, Ann Arbor, MI 48109-0648, U.S.A.; ‡Department of Anatomy and Cell Biology, University of Michigan Medical School, Ann Arbor, MI 48109-0616, U.S.A.

Abstract—Differential display of mRNA is a technique that enables the researcher to compare genes expressed in two or more different tissues or in the same tissue or cell under different conditions. The method is based on polymerase chain reaction amplification and comparison of specific subsets of mRNA. We have used this method to clone partial complementary DNAs (cDNAs; amplicons) for genes expressed in the otocyst in order to identify genes that may be involved in development of the inner ear. A full length cDNA was isolated from an embryonic quail head library with an amplicon (KH121) obtained from the otocyst. This avian cDNA encoded a novel, 172-amino acid acidic protein and detected a major transcript of *ca* 0.8 kb in RNA from chick embryos and several neonatal chick tissues. The full length avian cDNA had high sequence identity to several human cDNAs (expressed sequence tags) from human fetal tissues, including cochlea, brain, liver/spleen and lung, and from placenta. The human homologue of the avian gene encoded a protein that was 183 amino acids long and had 75.6% amino acid sequence identity to the avian protein. These results identified both the avian and human homologues of an evolutionarily conserved gene encoding a small acidic protein of unknown function; however, expression of this gene was not restricted to otocysts. © 1997 ISDN

Key words: differential display; human EST; acidic domain.

Birds provide an ideal model system for two types of studies in the auditory system: development of the ear,[18] and repair and regeneration of the auditory epithelium after noise trauma.[10,43,44,49,50] Development of the ear involves a genetic cascade, that is, a programmed series of sequential gene activation and inactivation.[18] Genetic analysis of embryonic lethal mutations in *Drosophila*,[41,53] identified a large class of genes involved in the early events of vertebrate development, for example, axis specification, segmentation, and pattern formation. Many of these genes are transiently expressed during embryonic development and encode transcription factors containing either homeodomains (HOX genes) or paired domains (PAX genes[8]). These transcription factors regulate the expression of many as yet unidentified target genes. The DNA-binding domain of *Hox* and *Pax* genes has been conserved throughout evolution, permitting the isolation of the vertebrate homologues by using the *Drosophila* gene as a hybridization probe. One example of a transcription factor involved in patterning the early events in ear development is Pax2. The *Pax2* gene is expressed in the otocyst[6,40] and presumably activates the expression of additional target genes required for correct ear development. This genetic approach cannot be used, however, to identify genes involved in the later stages of development of vertebrate embryonic structures such as the ear, since no comparable structures exist in invertebrates.

The second type of study that can be performed in birds involves repair and regeneration of the auditory epithelium after acoustic overstimulation. Regeneration of the auditory epithelium after damage by noise or ototoxic drugs occurs in birds but not mammals. This process has been examined both histologically and physiologically. It has been clearly demonstrated in birds that hair cells in the auditory system can regenerate if lost due to trauma.[10,43,44] Such regeneration not only repopulates the damaged auditory epithelium with immature hair cells, but also restores its functions.[35,39,49,50]

Thus, although the process of regeneration has been studied at the morphological and electrophysiological levels, very little is known about the molecular signals that trigger hair cell regeneration, or the physiological processes involved during regeneration. We have recently demonstrated that several genes encoding proteins involved in signal transduction are expressed at higher levels in the chicken basilar papilla after acoustic trauma.[22]

We are also interested in the molecular events involved in early ear development and in repair and regeneration of the auditory epithelium after acoustic overstimulation. In several organ systems, notably skeletal muscle[54] and liver,[7] genes that are expressed during development are re-expressed during regeneration. Our hypothesis is that some genes that are expressed in the developing ear (otocyst and associated neuronal structures) may be re-expressed during regeneration of the auditory epithelium and thus provide important markers for molecular analysis of both ear development and regeneration. A corollary of this hypothesis is that genes activated during regeneration may represent previously unidentified genes important in development. To test this hypothesis, it will be necessary to identify genes expressed in the otocyst during ear development. Very little is known about the process at the molecular level, and what is known is summarized in our recent review.[18] The molecular mechanisms involved in mesodermal and neuronal induction of otic placode and otocyst are not well understood. Neither are the genetic repertoires of placode cells or otocyst epithelia well characterized. However, in recent years, gene expression studies have revealed that both the placode and the developing otocyst have unique patterns of gene expression that restrict some genes and their protein products to specific cells at critical periods in the developmental process. These genes include: *dix3*,[13] *dix4*,[3] *otx1*,[30,51] *msxD*,[13] *GH6*,[52] *Nkx5.1*,[45] *Pax2*,[6,33,40] *SOHo-1*,[12] *goosecoid*,[20] and secreted growth factors such as *FGF3*,[38] *wnt3*,[26] *Xwnt4*,[37] and *BMP4*,[57] receptor tyrosine kinases, including *ret*,[46] and *sek1*,[45] and *Delta*, which is a membrane-bound ligand in the Notch signaling system (reviewed in Refs 5,17).

As a first step in identifying molecular events involved in later stages of development of the inner ear, we have applied the method of differential display of messenger RNA (mRNA)[31,32,36] to isolate genes expressed in the otocyst during development, since the otocyst is the embryonic precursor of the vertebrate ear.[18] In this paper we describe the results of differential display analysis of genes expressed in the otocyst. One of these complementary DNAs (cDNAs) and its mammalian homologue have been completely sequenced and the expression pattern examined.

1. EXPERIMENTAL PROCEDURES

1.1. *Isolation of RNA*

Total RNA was isolated by the acid–phenol method.[9] For differential display experiments, chick embryonic RNA was isolated from Hamburger and Hamilton (HH)[24] stage 23–24 (3.5d–4d) chick embryos. Embryos were aseptically removed from the eggs and rinsed in sterile phosphate buffered saline (PBS)+Ca^{2+} (Dulbecco's formulation; Life Technologies). Otocysts were dissected from embryos and placed immediately into guanidinium isothiocyanate (GTC) solution. From *ca* 250 otocysts we obtained 30 μg of total RNA. After the otocysts were dissected, everything in the head remaining rostral to visceral arch III and caudal to the mandibular arch, referred to hereafter as *hindbrain region*, was removed and placed in GTC solution for RNA isolation. From *ca* 40 embryonic hindbrain regions, we obtained 40 μg of total RNA, which was treated with RNase-free DNase I to reduce DNA contamination, extracted with acid–phenol, and ethanol precipitated.

For construction of an embryonic quail head cDNA library, heads were dissected from 42 embryos at stage 19 and from 75 embryos at stage 21–22 for RNA isolation. Total RNA from both stages was combined. Poly(A)$^+$ RNA was isolated from 572 μg RNA by the magnetic bead method, with a Poly A-Tract mRNA isolation kit (Promega, Madison, WI). The yield was 2.8 μg. For Northern blot analysis, total RNA was isolated from 1-week-old white Leghorn hatchling chickens or from 2-month-old mice. Tissues were dissected, placed immediately into GTC solution and disrupted with a Polytron (Brinkman) prior to total RNA isolation. Poly(A)$^+$ RNA was isolated as described above.

1.2. DNA probes and primers

DNA fragments used as hybridization probes were isolated in low melting point agarose and were labeled by the random primer method[15,16] with [α-^{32}P]dCTP (>3000 Ci/mmol; DuPont-New England Nuclear). All oligonucleotide primers were synthesized in the University of Michigan DNA Synthesis Facility. The cDNA for human glyceraldehyde-3-phosphate dehydrogenase (GAPDH)[14] was obtained from Maria Alexander-Bridges, Massachusetts General Hospital.

1.3. Differential display of mRNA

Differential display reactions were performed as described previously[22] with the RNAmap kit (GenHunter, Inc., Brookline, MA). MMLV reverse transcriptase and an anchored 3' primer (dT)$_{12}$MdG (M = a mixture of dA, dC, and dG) were used to generate first strand cDNA from total RNA. Aliquots of the resulting single-stranded cDNA population (first strand cDNA) were used as template DNAs for polymerase chain reaction (PCR) amplification, as described by Liang et al.[31] with the same anchored 3' primer (dT)$_{12}$MdG and five different 10-mer arbitrary primers (AP-1 to AP-5; RNAmap kit) on RNA from both otocyst and hindbrain region. [α-^{35}S]dATP was included in the PCR reaction mixture to label the resulting cDNA fragments (amplicons). Bands representing amplicons that were present in duplicate cDNA samples from otocysts, but not from the hindbrain region, were isolated, re-amplified by identical PCR reactions containing only unlabeled dNTPs, and subcloned into plasmid vectors as described previously.[22] Plasmid DNA was prepared using a modification of the alkaline lysis method.[48] These amplicons were presumed to be derived from genes that were differentially expressed in otocysts compared to hindbrain region. However, as indicated in Section 2, that was not always the case.

1.4. Construction and screening of a cDNA library from embryonic quail head mRNA

A cDNA library from embryonic quail head mRNA was generated to provide a source of full length cDNAs enriched in genes expressed in the otocyst. A unidirectional cDNA library was constructed in λUniZAP-XR (Stratagene, La Jolla, CA) from 1.5 μg poly(A)$^+$ RNA, according to the manufacturer's instructions. Briefly, first strand cDNA synthesis was primed with an oligo(dT) primer containing a XhoI adapter. The second strand cDNA was synthesized by using RNaseH to nick the RNA strand of the RNA–DNA hybrid, followed by nick translation with Escherichia coli DNA polymerase holoenzyme to generate the DNA duplex. Double-stranded cDNA was treated with DNA polymerase to create blunt ends, ligated to EcoR1 adapters, and digested with XhoI. The resulting cDNAs were passed over a Sephacryl S-400 column to isolate cDNAs longer than 500 bp, which were ligated to λUniZAP-XR vector arms at the EcoRI (5') and XhoI (3') sites within the pBlueScript II plasmid vector. The resulting recombinant DNAs were packaged with Gigapack II (Stratagene) and used to infect E. coli strain LE392. The primary library contained 1.5×10^6 plaque forming units (p.f.u.). One round of amplification generated a secondary library with an estimated titer of 1×10^{10} p.f.u./ml. The fraction of nonrecombinant plaques (3%) was estimated from the number of blue LacZ$^+$ plaques on plates containing X-gal, a chromogenic β-galactoside. An aliquot of the library was plated and screened by plaque hybridization[48] with a chick β-actin cDNA probe.[27] The library contained 1.5% β-actin cDNAs, an appropriate percentage of cDNAs for this highly abundant mRNA. Therefore, we could conclude that this library has a high probability of containing moderately abundant mRNAs.

The secondary (amplified) embryonic quail head cDNA library was screened for full length cDNAs by plaque hybridization.[48] A total of 5×10^5 p.f.u. was plated on E. coli and transferred to nitrocellulose membranes (BA85; Schleicher & Schuell, Keene, NH). Plaque hybridization was performed at 60°C overnight in 5 × standard saline citrate (SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0), 5 × Denhardt's solution (1 × = 0.2% each Ficoll, polyvinylpyrrolidone, and bovine serum albumin; BSA), 0.1% sodium dodecyl sulfate (SDS), and sheared, denatured fish sperm DNA (100 μg/ml). Hybridization solutions contained 5×10^5 c.p.m./ml of denatured cDNA probe. The most stringent post-hybridization washes were performed in 0.2 × SSC, 0.1% SDS at 60°C. Positive plaques were purified by successive rounds of hybridization. The plasmids contained within the λUniZAP-XR DNA were excised in vivo, as described previously[25] by co-infecting cells with both lambda phage and helper M13 phage and plating the infected cells on ampicillin plates to select cells containing the excised plasmid.

1.5. *DNA sequence analysis*

Purified plasmid DNA was sequenced either manually by the dideoxy-chain termination method (Sequenase 2, Amersham) or on an Applied Biosystems, Inc. model 373A automated DNA sequencer in the University of Michigan DNA Sequencing Core. DNA sequences were aligned with PC/GENE sequence analysis programs (Intelligenetics, Campbell, CA). Nucleotide database searches were performed with the FASTA program of the University of Wisconsin Genetics Computer Group (GCG) and with the BLAST[4] search program at the National Center for Biotechnology Information (NCBI) against GenBank, EMBL, dbEST [the database of expressed sequence tag (EST) sequences], and nonredundant PDB databases. Protein sequences were analyzed for secondary structure predictions either with PC/GENE or GCG programs, including Prosite, Motif, and Blocks searches. Protein searches were performed with FASTA via the GenQuest server at the Johns Hopkins University against SwissProt (release 32), Genome Sequence Database (GSDB-LANL), PDB, dbEST, and Genpept (release of Oct. 1995) databases and by BEAUTY searches.

1.6. *Northern blot analysis*

Chick hatchling poly(A)$^+$ RNA (1 μg) was subjected to electrophoresis through denaturing 1% agarose–2.2 M formaldehyde gels in a buffer containing 0.02 M 3-(*N*-morpholino)propane sulfonic acid (pH 7.0), 8 mM sodium acetate, and 1 mM ethylenediaminetetra-acetic acid (EDTA). RNA was transferred to Nytran membranes (Schleicher & Schuell, Keene, NH) by capillary transfer in 10 × SSC and immobilized by baking at 80°C for 2 hr *in vacuo*. The membrane was hybridized overnight at 65°C in a buffer containing 5 × SSPE (1 × SSPE = 0.15 M sodium chloride, 0.05 M sodium phosphate, 5 mM EDTA, pH 7.4), 5 × Denhardt's, 1% SDS, and denatured fish sperm DNA (100 μg/ml). DNA probes labeled by the random priming method were added to a final concentration of 4 × 10^6 c.p.m./ml. The post-hybridization washes were performed 65°C in 0.2 × SSC–0.5% SDS.

2. RESULTS

2.1. *Differential display analysis of genes expressed in the otocyst*

To identify genes involved in the development of the vertebrate inner ear, our strategy was to use differential display[31,32,36] to compare genes expressed in the otocyst with genes expressed in the surrounding tissues (*hindbrain region*) (Fig. 1). We isolated and sequenced three PCR products (amplicons) that, based on differential display gels, appeared to be present in RNA isolated from otocysts, but were not present, or were present at very low levels, in RNA from the hindbrain region. The DNA sequence of each of the three amplicons was compared with all sequences in the GenBank database. One partial cDNA clone (KH119, 359 bp) had 97.8% sequence identity to a known sequence in the database, the chick ND5 gene. The ND5 gene is located on mitochondrial DNA and encodes a subunit of the mitochondrial enzyme NADH-dehydrogenase (ND). The two remaining partial cDNAs, designated KH120 (268 bp) and KH121 (367 bp), respectively, had no significant sequence identity with any known DNA sequence in the database and were therefore presumed to be derived from previously unidentified genes.

2.2. *KH121 identified an embryonic avian cDNA encoding a previously unidentified acidic protein*

To determine the type of protein encoded by a novel gene, it is necessary to isolate and sequence the full length cDNA, then translate the cDNA sequence to deduce the amino acid sequence. Therefore, we used both the 268-bp cDNA insert from KH120 and the 367-bp cDNA insert from KH121 as hybridization probes to screen an embryonic quail head cDNA library constructed in λUniZAP-XR. No full length cDNA was obtained with the KH120 partial cDNA probe; therefore, the nature of this gene is still unknown. We identified two plaques from the embryonic quail head cDNA library that hybridized with the KH121 cDNA probe, excised the plasmids from the purified phage *in vivo*, and determined the complete DNA sequence of one cDNA (KH210, 777 bp). This cDNA encoded a novel protein: a small, 172 amino acid, acidic polypeptide (Fig. 2) with a predicted molecular weight of 19 kDa. The deduced protein sequence contained many charged amino acid residues and was particularly rich in acidic amino acids, yielding a calculated isoelectric point (pI) of 5.13. There is no obvious hydrophobic, transmembrane domain. The theoretical protein sequence
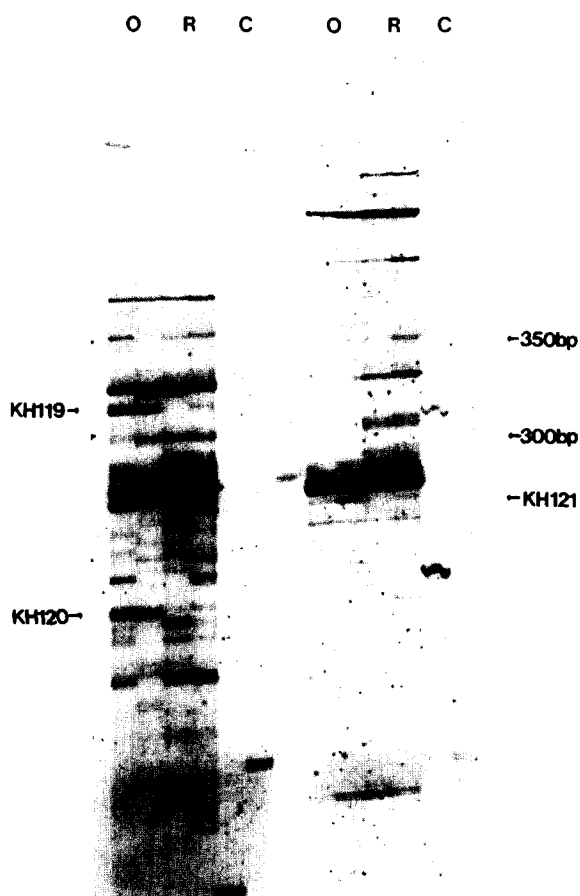
O   R   C   O   R   C

Fig. 1. Identification of partial cDNAs (amplicons) KH119, KH120 and KH121 by differential display of mRNA. Total RNA from chick otocysts (O) and hindbrain region (R) was treated with RNase-free DNase I to remove DNA and was subjected to the differential display procedure. First strand cDNA was generated from RNA with MMLV reverse transcriptase and the $(dT)_{12}MdG$ anchored primer (see Section 1.3). PCR amplification was performed using $(dT)_{12}MdG$ and either the AP-4 (5'-GGTACTCCAC-3') or AP-5 (5'-GTTGCGATCC-3') amplification primer (GenHunter RNAmap kit). PCR cycling conditions were 94°C for 30 sec, 40°C for 2 min, 72°C for 30 sec, for a total of 40 cycles, followed by 72°C for 5 min. PCR products were analyzed by electrophoresis at 60 W for 3.5 hr on a denaturing 6% acrylamide–7 M urea gel. The dried gel was exposed to X-Omat film overnight. Left panel, PCR products generated with the $(dT)_{12}MdG$ and AP-4 primers. Right panel, PCR products generated with the $(dT)_{12}MdG$ and AP-5 primers. Lanes: O, duplicate PCR reactions performed on RNA from otocysts; R, duplicate PCR reactions performed on RNA from hindbrain region; C, control reactions containing no reverse transcriptase to test for DNA contamination. The left control lane contained RNA from the O samples; the right control lane contained RNA from the R samples. The arrows indicate three unique PCR products (amplicons), designated KH119, KH120, KH121, that were observed in the duplicate O lanes, but not in the duplicate R lanes.

deduced from the embryonic quail cDNA had no significant sequence identity to any known protein using the BLASTP search algorithm, nor could we identify any known protein motif in the embryonic quail protein sequence. We concluded, therefore, that the cDNA we isolated and sequenced (KH210) was the transcript of a novel, previously unidentified gene encoding a small, highly acidic protein of unknown function.

2.3. Identification of a human homologue of the embryonic quail gene

It was important to determine whether or not the human genome contained a homologue of the novel avian gene we identified by differential display, since the presence of a mammalian homologue might indicate that the protein was conserved throughout vertebrate evolution. We used the BLAST algorithm to search the GenBank and dbEST nucleic acid databases for mammalian cDNAs with significant sequence identity to the full length quail cDNA (KH210). This search identified several uncharacterized human cDNAs, ESTs,[1,2] with high sequence identity to the embryonic quail cDNA.

```
HUMAN    - MSAARESHP-HGVKRSASPDDDLGSSNWEAADLGNEERKQKFLRLMGAGK -49
           ||·||·|    ||·||·||||    ||  | |||||||||||||||||||||||
QUAIL    - MSSARDSQAQHGLKRAASPD---GSGSWQAADLGNEERKQKFLRLMGAGK -47


HUMAN    - KEHTGRLVIGDHKSTSHFRTGEEDKKINEELESQYQQSMDSKLSGRYRRH -99
           |||||||||||·||||||||||||·|||||||||||| ·||| |||
QUAIL    - KEHTGRLVIGDHRSTSHFRTGEEDKKMNEELESQYQQSMDSTMSGRNRRH -97


HUMAN    - CGLGFSEVEDHDGEGDVAGDDDDDDDDSPDPESPDDSESDSESEKEESAE -149
           |||||||  ·· |  ||       | | || ·|||| |·|| ·||||
QUAIL    - CGLGFSEFQEVEEEA--AGH-------SSDHESSEDSESGSDSEQDESAE -138


HUMAN    - ELQAAEHPDEVEDPKNKKDAKSNYKMMFVKSSGS -183
           ||||||  ||·  | |||·||||||||||||·|||
QUAIL    - ELQAAEKHDEAAVPENKKEAKSNYKMMFVKASGS -172
```

Fig. 2. Comparison of the deduced amino acid sequences of the quail head and human proteins. The deduced amino acid sequence of the embryonic quail protein was obtained by translating a full length cDNA sequence (KH210, GenBank accession No. U37722). The first ATG in the sequence was surrounded by a sequence with nine out of 11 matches to the Kozak consensus sequence for translation initiation.[28] Similarly, the human protein sequence was deduced from a full length cDNA (KH251: GenBank accession No. U51678) from a human placental cDNA library. Protein sequences were aligned using the CLUSTAL program of PC/GENE. Symbols: |, identical amino acids; ·, conservative substitutions; -, gap inserted to optimize the amino acid alignment.

The DNA sequences of these ESTs were identical, indicating that all were derived from the same human gene. The cDNAs that gave rise to the EST sequences were present in libraries from different tissues, indicating that the gene was expressed in a wide range of cell types. We obtained and completely sequenced one full length human cDNA (KH251; I. M.A.G.E. CloneID 145052, Research Genetics) identified by the I.M.A.G.E. Consortium[29] at Washington University. The deduced human protein sequence was slightly longer than the avian protein (183 vs 172 amino acids long and 20 kDa vs 19 kDa predicted molecular weight) (Fig. 2). The human protein had a high proportion of charged amino acids and was particularly rich in acidic amino acids such as aspartate (D = Asp) and glutamate (E = Glu), including a central region of the protein containing eight contiguous Asp residues, resulting in a predicted pI of 4.57. Database searches with the deduced human protein identified no known protein with significant sequence identity nor any known protein motif. This lack of amino acid sequence identity supported the conclusion that these proteins and the genes encoding them have not been described previously.

2.4. The novel quail and human proteins are highly conserved

Alignment of the quail and human protein sequences with the FSTPSCAN program of PC/GENE (Fig. 2) demonstrated 75.6% overall amino acid sequence identity between the quail and the human protein sequences. The sequences are co-linear except for the single amino acid deletion in the human protein at residue 10, the three amino acids inserted between codons 20–21, and eight contiguous Asp residues in the central region of the protein, which are present only in the human protein. Notably, there were two regions of sequence conservation: a region of high sequence identity (73/78 amino acids) in the N-terminal half of the protein and a well conserved region at the C-terminus. The amino acid sequence surrounding the eight contiguous Asp residues in the center of the protein was not as well conserved and represents the region of greatest sequence divergence between these two proteins. However, there are approximately the same number of negatively charged amino acid residues (Asp or Glu) flanking this stretch of eight Asp residues. Thus, the most important feature of this domain may be a high concentration of negatively charged amino acids, not the exact amino acid sequence.

2.5. The gene encoding the embryonic quail cDNA is expressed in both embryonic and neonatal avian tissues

Northern blot analysis of total RNA from whole chick embryos (4 days) or from various tissues from hatchling chicks indicated that the gene encoding the novel acidic protein is widely expressed. The KH210 cDNA probe detected a major transcript of *ca* 0.8 kb in total RNA from brain, heart,

eye, liver, kidney, and skeletal muscle of chick hatchlings, as well as in 4-day embryos (Fig. 3). The same transcript is also present in 2- and 6-day chick embryos (data not shown). In addition, two minor transcripts of ca 1.8 and 1.5 kb were detected in all tissues examined. These larger transcripts appeared to be less abundant than the smaller 800 bp transcript. It is not clear whether these larger transcripts represented an alternatively spliced mRNA or cross-hybridization to a transcript from a related member of a gene family. Preliminary *in situ* hybridization experiments with the quail cDNA in whole chick embryos indicated that the gene was expressed in many cell types and was not restricted to the otocyst (Lewis and Barald, unpublished data).

The pattern of expression of the human homologue of the avian gene can also be deduced to some extent from the distribution of the human EST sequences in cDNA libraries. ESTs with high sequence identity to the quail head cDNA (KH210) were derived from cDNAs isolated from several fetal tissue libraries (cochlea, brain, lung, placenta) and from neonatal or adult tissues (brain, lung, white blood cells, melanocytes, and prostate gland). Expression in a wide variety of tissues is consistent with the tissue distribution observed by Northern analysis in the chick. It also suggests that expression may be higher in fetal tissues, since the cDNA was found primarily in fetal and neonatal cDNA libraries. It should be noted that this novel gene is expressed in human fetal cochleas, at least during the stage of fetal development at which the cochleas were isolated for construction of the cDNA library.[47] The role of this protein during development of the inner ear remains to be determined.

An additional question that we hope to address with these novel cDNAs is whether their genes represent candidate genes for hereditary syndromic or nonsyndromic deafness. To answer this question, we attempted to map the gene for the mammalian homologue of KH121. Southern hybridization analysis of a human chromosome mapping panel with the human cDNA identified two genes, one on human chromosome 11, the other on chromosome 14 (data not shown). These mapping results suggest that there are either two related expressed genes, or an expressed gene and a pseudogene, in the human genome. Four nonsyndromic deafness genes have been mapped to these two chromosomes: DFNA11[55] to 11q12.3-q21; DFNB2[23] to 11q13.5; DFNA9[34] to 14q12-q13; and DFNB5[19] to 14q12. Isolation of the expressed gene will be required to distinguish between these
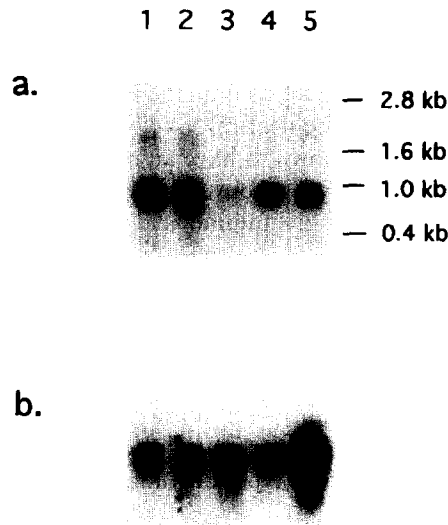

Fig. 3. Expression of the embryonic quail gene (KH210) in chicken tissues. Northern blot analysis was performed on poly(A)$^+$ RNA (1 μg per lane) from 4-day chick embryos (lane 1) or tissues from 1-week-old hatchling chicks, including brain (lane 2), eye (lane 3), heart (lane 4), and skeletal muscle (lane 5). RNA was subjected to electrophoresis on denaturing 1% agarose-2.2 M formaldehyde gels and transferred to Nytran membranes. (A) Northern blot hybridized with the KH210 cDNA probe. The most stringent post-hybridization wash was carried out at 0.2 × SSC–0.5% SDS at 65°C. The sizes of the mRNA transcripts, estimated from the migration of RNA molecular weight standards, are ca 0.8 kb and 2.3 kb. (B) Northern blot hybridized with the human GAPDH cDNA probe. The same membrane was stripped and reprobed with a cDNA for GAPDH as a control for RNA loading. Final wash conditions were: 0.2 × SSC–0.5% SDS at 65°C.

two genomic regions and to determine whether the expressed gene maps near, and is thus a candidate for, a deafness gene on one of these two chromosomes.

3. DISCUSSION

Little is known at the molecular level about the processes involved in development of the inner ear, or in regeneration of the auditory epithelium after noise trauma. To begin to fill this gap in our knowledge, we have designed experiments to identify genes expressed in the otocyst during development of the ear. In several organ systems, notably skeletal muscle[54] and liver,[7] genes that are expressed during development are re-expressed during regeneration. Genes that meet the criteria of being differentially expressed in the otocyst will eventually be used to test our hypothesis that such genes might be re-expressed in the avian basilar papilla during regeneration following acoustic trauma; however, none of the genes isolated in this study meet the criteria of being differentially expressed in the otocyst. Isolation of the ND5 gene encoding a subunit of the mitochondrial NADH-dehydrogenase complex may indicate increased levels of mitochondrial biogenesis in the otocyst at this stage of development. In view of the known effect of mitochondrial DNA mutations on hearing,[56] this observation deserves further examination.

We were able to use the avian cDNA (KH121) identified by differential display analysis of chick otocyst RNA to obtain a full length avian (quail) cDNA and subsequently to identify and sequence the human homologue of this gene. These analyses indicated that the gene we had identified encoded a novel protein that is highly conserved between mammals (human) and birds. FASTA and BLASTP searches of the entire SwissProt database with either the entire protein sequence or the highly conserved N-terminal region of the human and quail protein sequences did not reveal any protein motif or any region of significant sequence homology to known proteins. Although the central, acidic region of both the quail and human proteins (Fig. 2) showed some sequence similarity to acidic activation domains found in several proteins with nucleotide binding properties, the sequence identity could not be extended beyond this region. We concluded, therefore, that this gene encoded a highly conserved novel protein of unknown function.

These results also illustrate the importance of the Human Genome Project, and in particular, the EST project,[1,2,29] for gene discovery studies in other organisms such as birds. Database searches, particularly with sequences from the 5' end of a cDNA, can demonstrate quickly that the human genome contains a homologous gene. Furthermore, the pattern of expression of the human gene can be deduced from the cDNA libraries that contain the EST sequence. The presence of the human homologue of KH121 in a human fetal cochlear cDNA library suggested that even though the gene was not otocyst-specific, it may yet be shown to have an important function in development of the ear. The challenge for studies of these novel genes will be to define the function of the protein, based only on information on the primary structure (amino acid sequence). This is an area of biology that lags far behind our ability to clone and sequence genes.[42]

We were unable to identify full length cDNAs for KH120 in the embryonic quail head cDNA library we constructed. This could be due to low levels of gene expression at the embryonic stages used for constructing the library. Alternatively, there may have been high level expression in the otocyst but not in other cell types, leading to dilution of the otocyst mRNA in RNA from total head, which was used in the construction of the library. Clearly, the availability of an otocyst-specific cDNA library, which has recently been constructed (D. Fekete, personal communication) and made available to us, will enhance these types of studies by optimizing our ability to isolate full length otocyst-specific cDNAs.

To test the corollary of our hypothesis, that is, that genes expressed during regeneration may represent genes important for development, we have begun to identify genes expressed after acoustic trauma and during regeneration of the basilar papilla in chickens.[22] In addition to several cDNAs encoding proteins involved in signal transduction, one cDNA, KH129, was shown to be present at higher concentrations after noise,[22] but could not be identified because the cDNA sequence did not have significant sequence identity to known genes in the GenBank database. This cDNA, and others isolated during regeneration after acoustic trauma, will be tested for expression in the otocyst or in later embryonic structures derived from the otocyst.

We achieved only limited success in mapping the human homologue of the gene encoding the KH121 cDNA (amplicon). We identified genomic regions on human chromosomes 11 and 14, respectively; however, we do not know which chromosome contains the expressed gene. We have successfully mapped the mouse homologue of KH129[21] to a region of mouse chromosome 4 that is syntenic with human chromosome 1p36. Although the hereditary deafness gene DFNA2 has been mapped to the short arm of human chromosome 1,[11] these mapping studies on the mouse and human genes for KKH129[21] excluded this gene as a candidate for DFNA2. Such mapping studies will be an important aspect of analyzing these novel cDNAs and determining whether or not the gene is a candidate gene for a hereditary deafness.

REFERENCES

1. Adams M., Kelley J., Gocayne J. D., Dubnick M., Polymeropoulos M. H., Xiao H., Merri C. R., Wu A., Olde B., Moreno R. F., Kerlavage A. R., McCombie W. R. and Venter J. C. (1991) Complementary DNA sequencing: expressed sequence tags and the human genome project. *Science* **252**, 1651–1656.
2. Adams M. *et al.* (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* **377**, 3–17.
3. Akimenko M. A., Ekker M., Wegner J., Lin W. and Westerfield M. (1994) Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head. *J. Neurosci.* **14**, 3475–3486.
4. Altschul S. F., Stephen F., Gish W., Miller W., Myers E. W. and Lipman D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
5. Artavanis-Tsakonas S., Matsuno K. and Fortini M. E. (1995) Notch signalling. *Science* **267**, 225–232.
6. Barald K. F., Lindberg K. H., Hardiman K., Kavka A. I., Lewis J. E., Victor J. C., Gardner C. A. and Poniatowski A. (1997) Immortalized cell lines from embryonic avian and murine otocysts: Tools for molecular studies of the developing inner ear. *Int. J. Devl. Neurosci.* **15**, 523–540.
7. Belayew A., Tilghman S. M. (1982) Genetic analysis of α-fetoprotein synthesis in mice. *Mol. Cell Biol.* **2**, 1427–1435.
8. Chalepakis G., Stoykova A., Wihnholds J., Tremblay P. and Gruss P. (1993) Pax: gene regulators in the developing nervous system. *J. Neurobiol.* **24**, 1367–1384.
9. Chomcszynski P. and Sacchi N. (1987) Single step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162**, 156–159.
10. Cotanche D. A., Lee K. H., Stone J. S. and Picard D. A. (1994) Hair cell regeneration in the bird cochlea following noise damage or ototoxic drug damage. *Anat. Embryol.* **189**, 1–18.
11. Couche P., Van Camp G., Djoyodiharjo B., Smith S. D., Frants R. R., Padberg G. W., Darby J. K., Huizing E. H., Cremers W. R. J., Kimberling W. J., Oostra B. A., Van de Heyning P. H. and Willems P. J. (1994) Linkage of autosomal dominant hearing loss to the short arm of chromosome 1 in two families. *N. Engl. J. Med.* **331**, 425–431.
12. Deitcher D. L., Fekete D. M. and Cepko C. L. (1994) Asymmetric expression of a novel homeobox gene in vertebrate sensory organs. *J. Neurosci.* **14**, 486–498.
13. Ekker M., Akimenko M. A., Bremiller R. and Westerfield M. (1992) Regional expression of three homeobox transcripts in the inner ear of zebrafish embryos. *Neuron* **9**, 27–35.
14. Ercolani L., Florence B., Denaro M. and Alexander M. (1988) Isolation and complete sequence of a funcational human glyceraldehyde-3-phosphate dehydrogenase gene. *J. Biol. Chem.* **263**, 15335–15341.
15. Feinberg A. P. and Vogelstein B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6–13.
16. Feinberg A. P. and Vogelstein B. (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Addendum. *Anal. Biochem.* **137**, 266–267.
17. Fekete, D. (1996) Cell fate specification in the inner ear. *Curr. Opin. Neurobiol.* **6**, 533–541.
18. Fritzsch B. F., Barald K. F. and Lomax M. I. Early embryology of the vertebrate ear. In *Development of the Auditory System* (eds Rubel E. W., Popper A. N. and Fay R. R.) Springer-Verlag, Berlin, in press.
19. Fukushima K., Ramesh A., Srikumari Srisailapathy C. R., Ni L., Chen A., O'Neill M., Van Camp G., Coucke P., Smith S. D., Kenyon J. B., Jain P., Wilcox E. R., Zbar R. I. S. and Smith R. J. H. (1995) Consanguineous nuclear families used to identify a new locus for recessive non-syndromic hearing loss on 14q. *Human Molec. Genetics* **4**, 1643–1648.
20. Gaunt S. J., Blum M. and DeRobertis E. M. (1993) Expression of the mosue goosecoid gene during mid-embryogenesis may mark mesenchymal cell lineages in the developing head, limbs and body wall. *Development* **117**, 769–778.
21. Gong T.-W. L., Burmeister M. and Lomax M. I. (1996) The novel gene *D4Mille* maps to distal mouse chromosome 4. *Mammalian Genome* **7**, 790–791.
22. Gong T. W. L., Hegemann A. D., Shin J. J., Adler H. J., Raphael Y. and Lomax M. I. (1996) Identification of genes expressed after noise exposure in the chick basilar papilla. *Hearing Res.* **96**, 20–32.
23. Guilford P., Ayadi H., Blanchard S., Chaib H., Le Pasilier D., Weissenbach J., Drira M. and Petit C. (1994) A human gene responsible for neurosensory, non-syndromic recessive deafness is a candidate homologue of the mouse sh-1 gene. *Human Molec. Genetics* **3**, 989–993.

24. Hamburger V. and Hamilton V. L. (1951) A series of normal stages in the development of the chick embryo. *J. Morphol.* **88**, 48–92.
25. Hegeman A. D., Brown J. S. and Lomax M. I. (1995) Isolation and sequence of the mouse cDNA for the sarcomeric isoform of cytochrome *c* oxidase subunit VIII. *Biochim. Biophys. Acta* **1261**, 311–314.
26. Hollyday M., McMahon T. A. and McMahon A. P. (1995) *Wnt* expression pattern in chick embryo nervous system. *Mech. Dev.* **52**, 9–25.
27. Kost T. A., Theodorakis N. and Hughes S. H. (1983) The nucleotide sequence of the chick cytoplasmic beta-actin gene. *Nucleic Acids Res.* **11**, 8287–8301.
28. Kozak M. (1987) An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **20**, 8125–8148.
29. Lennon G., Auffray C., Polymeropoulos M. and Soares M. B. (1996) The 1. M.A.G.E. consorium: an integrated molecular analysis of genomes and their expression. *Genomics* **33**, 151–152.
30. Li Y., Allende M. L., Finkelstein R. and Weinberg E. S. (1994) Expression of two zebrafish orthodenticle-related genes in the embryonic brain. *Mech. Dev.* **48**, 229–244.
31. Liang P., Averboukh L. and Pardee A. B. (1993) Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimization. *Nucleic Acids Res.* **21**, 3269–3275.
32. Liang P. and Pardee A. B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**, 967–971.
33. Lindberg K. H., Lomax M. I., Hegeman A. D. and Barald K. F. PAX-2: An early marker of otic placode induction that defines otocyst development in the chick. *Abstracts of the Midwinter Meeting of the Association for Research in Otolaryngology*, 5–9 February 1995, p. 105.
34. Manolis E. N., Yandavi N., Nadol J. B. Jr, Eavey R. D., McKenna M., Rosenbaum S., Khetarpal U., Halpin C., Merchant S. N., Duyk G. M., MacRae C., Seidman C. E. and Seidman J. G. (1996) A gene for non-syndromic autosomal dominant progressive postlingual sensorineural hearing loss maps to chromosome 14q12-13. *Human Molec. Genetics* **5**, 1047–1050.
35. Marean G. C., Burt J. M., Beecher M. D. and Rubel E. W. (1993) Hair cell regeneration in the European starling (*Sturnus vulgaris*): recovery of pure-tone detection thresholds. *Hearing Res.* **71**, 125–136.
36. McClelland M., Methieu-Daude F. and Welsh J. (1995) RNA fingerprinting and differential display using arbitrarily primed PCR. *Trends Genet.* **11**, 242–246.
37. McGrew L. L., Otte A. and Moon R. (1992) Analysis of Xwnt-4 in embryos of *Xenopus laevis*: a Wnt family member expressed in the brain and floor plates. *Development* **115**, 463–473.
38. McKay I. J., Lewis J. and Lumsden A. (1996) The role of FGF-3 in early inner ear development: an analysis in normal and *kreisler* mutant mice. *Dev. Biol.* **174**, 370–378.
39. Niemiec A. J., Raphael Y. and Moody D. B. (1994) Return of auditory function following structural regeneration after acoustic trauma: behavioral measures from quail. *Hearing Res.* **75**, 209–224.
40. Nornes H. O., Dressler G. R., Knapik E. W., Deutsch U. and Gruss P. (1990) Spatially and temporally restricted expression of Pax-2 during murine neurogenesis. *Development* **109**, 797–809.
41. Nusslein-Vollhard C. (1991) Determination of the embryonic axes of *Drosophila*. *Development* **1**, 1–10.
42. Oliver S. G. (1996) From DNA sequence to biological function. *Nature* **379**, 597–600.
43. Raphael Y. (1992) Evidence for supporting cell mitosis in response to acoustic trauma in the avian inner ear. *J. Neurocytol.* **21**, 663–671.
44. Raphael Y. (1993) Reorganization of the chick basilar papilla following acoustic trauma. *J. Comp. Neurol.* **330**, 521–532.
45. Rinkwitz-Brandt S., Justus M., Oldenettel I., Arnold H. H. and Bober E. (1995) Distinct temporal expression of mouse *Nkx*-5.1 and *Nkx*-5.2 homeobox genes during brain and ear development. *Mech. Dev.* **52**, 371–381.
46. Robertson K. and Mason I. (1995) Expression of *ret* in the chicken embryo suggests roles in regionalization of the vagal neural tube and somites and in development of multiple neural crest and placodal lineages. *Mech. Dev.* **53**, 329–344.
47. Robertson N. G., Khetarpal U., Gutierrez-Espeleta G. A., Bieber F. R. and Morton C. C. (1994) Isolation of novel and known genes from a human fetal cochlear cDNA library using subtractive hybridization and differential screening. *Genomics* **23**, 42–50.
48. Sambrook J., Fritsch E. F. and Maniatis T. *Molecular Cloning: A Laboratory Manual*, Vols 1–3. Cold Spring Harbor Press, New York.
49. Saunders J. C. and Adler H. J. (1992) The structural and functional aspects of hair cell regeneration in the chick as a result of exposure to intense sound. *Exp. Neurol.* **115**, 13–17.
50. Saunders J. C., Doan D. E., Cohen Y. E., Adler H. J. and Poje C. P. (1996) Recent observations on the recovery of structure and function in the sound-damaged chick ear. In *Auditory System Plasticity and Regeneration* (eds Salvi R. J., Henderson D., Fiorino F. and Colleti V.), pp. 62–83. Thieme Medical Publishers, New York.
51. Simeone A., Acampora D., Mallamaci A., Storaiuolo A., D'Apice M. R., Nigro V. and Boncinelli E. (1993) A vertebrate gene related to orthodenticle contains a homeodomain of the bicoid class and demarcates anterior neuroectoderm in the gastrulating mouse embryo. *EMBO J.* **12**, 2735–2747.
52. Stadler H. S. and Solursh M. (1994) Characterization of the homeobox-containing gene GH6 identifies novel regions of homeobox gene expression in the developing chick embryo. *Dev. Biol.* **161**, 251–262.
53. St Johnston D., Nusslein-Vollhard C. (1992) The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**, 201–219.
54. Sutherland C. J., Esser K. A., Elsom V. L., Gordon M. L. and Hardeman E. C. (1993) Identification of a program of contractile protein gene expression initiated upon muscle differentiation. *Dev. Dyn.* **196**, 25–36.
55. Tamagawa Y., Kitamura K., Ishida T., Ishikawa K., Tanaka H., Tsuji S. and Hishizawa M. (1996) A gene for a dominant form of non-syndromic sensorineural deafness (DFNA11) maps within the region containing the DFNB2 recessive deafness gene. *Human Molec. Genetics* **5**, 849–852.
56. Wallace D. C. (1994) Mitochondrial DNA mutations in diseases of energy metabolism. *J. Bioenerg. Biomembranes* **26**, 241–250.
57. Wu D. K. and Oh S. H. (1996) Differential expression of bone morphogenetic proteins in the developing vestibular and auditory sensory organs. *J. Neurosci.* **16**, 6463–6475.