JARO 03: 140–148 (2001)
DOI: 10.1007/s101620020029

JARO
Journal of the Association for Research in Otolaryngology

An Inner Ear Gene Expression Database

ZHENG-YI CHEN,[1,2] AND DAVID P. COREY[3–5]

[1]*Neurology Service, Massachusetts General Hospital, Boston, MA 02114, USA*
[2]*Department of Neurology, Harvard Medical School, Boston, MA 02115, USA*
[3]*Neurosurgery Service, Massachusetts General Hospital, Boston, MA 02114, USA*
[4]*Department of Neurobiology, Harvard Medical School, Boston, MA 02114, USA*
[5]*Howard Hughes Medical Institute, Chevy Chase, MD 20815-6789, USA*

Received: 30 May 2001; Accepted: 10 August 2001; Online publication: 18 October 2001

ABSTRACT

Microarray technology has provided an unprecedented opportunity to study gene expression profiles at a whole-genome level. As a first step toward a comprehensive understanding of inner ear gene expression, mouse cochleas were examined at two developmental stages (P2 and P32) using GeneChip oligonucleotide arrays. A large number of genes and ESTs (>10,000) were found to be expressed in the cochlea. Expression profiles derived from duplicate samples at the same developmental stages showed general agreement and indicated the reproducibility of the assay. The expression of many known hair-cell genes was detected in the whole-cochlea samples, demonstrating the relatively high sensitivity of the assay. Genes highly expressed only at P2 or P32 were also identified and their expression patterns correlate with their functions in the cochlea. A web-based database with external links was set up for public access, which should facilitate the discovery of genes important in the development and function of the inner ear and should aid the identification of additional deafness genes.

Keywords: cochlea, utricle, microarray, genomic, Wnt

INTRODUCTION

The development and the function of the inner ear are products of the genes expressed, the timing of their expression, and the functional pathways in which they operate. Knowing the pattern of gene expression in the inner ear will help in understanding development and function. Knowledge of genes expressed, in combination with their chromosomal locations, will also assist in identifying genes for hereditary hearing loss. Traditional approaches, such as RT-PCR (including differential display), cDNA library screening, and sequencing of cDNAs from inner ear libraries, have had some success in identifying inner ear genes (Gong et al. 1996; Heller et al. 1998; Skvorak et al. 1999). There are, however, limitations to those approaches. RT-PCR and cDNA library screening can deal with only very limited numbers of genes. Sequencing of cDNA libraries from inner ear has identified thousands of genes, many of them novel. In order to recognize low-copy-number genes with random sequencing of clones, however, an impractically large number of genes have to be sequenced. For instance, about two-thirds of the ~20,000 genes expressed by a typical cell type are expressed at low abundance, i.e., at <5 copies per cell (Zhang et al. 1997). If 50,000 clones from a cDNA library were sequenced at random, we calculate that more than half of these low abundance transcripts (a third of genes expressed) would be missed. The SAGE (serial analysis of gene expression) method to profile expression (Velculescu et al. 1995; 2000) relies on the sequencing of a large number of clones of concatenated, unique sequence tags derived from multiple cDNAs. While SAGE analysis does not depend on previous knowledge of the cDNA sequences (therefore new

Correspondence to: Zheng-Yi Chen, Ph.D. • WEL415 • Neurology Department • Massachusetts General Hospital • Boston, MA 02114. Telephone: (617) 724-8599; fax: (617) 726-5256; email: zhengyi@helix.mgh.harvard.edu

genes can be discovered) and is very quantitative, it is also laborious and requires the preparation, screening, and sequencing of thousands of clones for each sample.

Microarray technology, with which tens of thousands of genes can be monitored simultaneously, offers an alternative and efficient approach to studying gene expression in the inner ear. There are different forms of microarray technology, but all use the same principle: to immobilize targets (cDNAs, oligonucleotides, etc.) on the surface of a solid support, to hybridize a sample to the targets, and to record all the hybridization signals simultaneously. Since many targets (ranging from 10,000 cDNAs to 400,000 oligonucleotides) can be immobilized on one surface, comprehensive expression patterns can be detected in a given sample through several rounds of hybridization (Lander 1999). The success of the human and mouse genome projects has produced an enormous amount of sequence data. Chips commercially available have sequences representing 60,000 human genes and EST clusters, a large fraction of the genome.

We have used oligonucleotide arrays (GeneChip, Affymetrix, Santa Clara, CA) to profile gene expression in the mouse cochlea. We chose the GeneChip system because it is commercially available and covers the highest number of genes. Importantly, the data obtained can be compared with data from other resources using the same platform. We report here gene expression in the mouse cochlea assayed with the GeneChip technology. These data are available as a searchable database that can be downloaded from the web.

METHOD

The oligonucleotide array set Mu30K (comprising Mu11K plus Mu19K) was used for the study. This set covers ~13,000 known genes and ~21,000 EST clusters. Because of the overlap between some of the ESTs and known genes, many genes or ESTs are represented more than once on the chip. The total number of unique genes covered by the set is ~22,000. For simplicity, we will treat each entry on the chip (a total of 33,945 known genes or EST clusters) as one gene in our analysis.

Two developmental stages were used for the study, postnatal day 2 (P2) and postnatal day 32 (P32). Since the inner ear is not fully functional at P2 but is mature at P32, the comparison may reveal genes important in the maturation process. The CBA/CaJ mouse strain was used because it has good hearing (based on ABR testing) even at 2 years old (Zheng et al. 1999). P32 mice were anesthetized with isoflurane; P2 mice were anesthetized with CO_2. Mice were sacrificed by decapitation and the skull was hemisected, allowing a medial approach to the temporal bone that offers access to the cochlea. The cochleas were removed into MEM (Invitrogen, Carlsbad, CA) buffered with 10 mM HEPES to pH 7.4. P32 cochleas were transferred into chilled RNAlater solution (Ambion, Austin, TX) and stored at 4°C. P2 cochleas were frozen in liquid nitrogen and stored at −80°C. A minimum of 1 μg of total RNA was required to produce sufficient cRNA for GeneChip hybridization (Z-Y Chen, unpublished data). For this study, an average of ten mouse cochleas were collected for each of four samples, which produced an average of 1.5 μg total RNA.

The procedures for RNA extraction and cDNA and cRNA synthesis followed the methods recommended by Affymetrix. Total RNA was extracted using the RNeasy RNA isolation kit (Qiagen, Valencia, CA). Tissues were homogenized in 700 μL of RLT solution with β-mercaptoethanol. An equal volume of 70% ethanol was added to the homogenized lysate, mixed well, and spun through an RNeasy mini spin column. The column was washed once with RW1 solution and then transferred to a new collection tube. To avoid DNA contamination, 10 μL DNase I with 70 μL RDD buffer (RNase-free DNase set, Qiagen) was mixed and added to the column for 15 min. The column was washed with RW1, followed by wash with 500 μL of RPE solution. Diethyl pyrocarbonate-treated H_2O (50 μL) was used to elute RNA. Oligonucleotide assays require high-quality RNA for cDNA synthesis. The total RNA was further purified with 3M NaAc and ethanol in the presence of glycogen.

First- and second-strand cDNA synthesis used the T7-(dT)24 (designed by Affymetrix and synthesized and HPLC purified by Genset, La Jolla, CA) and Gibco/BRL SuperScript Choice system (Invitrogen, Carlsbad, CA), following the manufacturer's manual. To improve the efficiency of cDNA synthesis from total RNA, synthesis was done at 42°C. Double-stranded cDNA was then purified with an equal-volume phenol/chloroform/isoamyl alcohol (25:24:1) mix and separated by spinning in a tube containing Phase Lock Gels (PLG; Eppendorf-5 Prime, Inc. Boulder, CO). The aqueous-cDNA-containing phase was purified with 7.5 M NH_4Ac and ethanol in the presence of glycogen.

The Enzo Bioarray High Yield RNA Transcript Labeling Kit (Enzo Diagnostics, Farmingdale, NY), was used for *in vitro* transcription to make cRNA. The T7 promoter linked to the (dT)24 used for cDNA synthesis served as the binding site for T7 RNA polymerase for *in vitro* transcription. Following the kit instructions, we routinely obtained ~11 μg of cRNA from ~1.5 μg total RNA. The cRNA was then purified with Qiagen's RNeasy kit before it was fragmented in 5 × RNA fragmentation solution [40 mM Tris-acetate (pH 8.1), 100 mM KOAc, and 30 mM MgOAc] at 94°C for 35 min. Gel electrophoresis was run with ds cDNA, unfragmented

cRNA, and fragmented cRNA to confirm that no obvious degradation of cRNA occurred during the preparation and that the cRNA was fragmented to the optimal size of 35–200 bases for hybridization. Hybridization and scanning were conducted in a core facility, using the protocol provided by Affymetrix. The final amount of the cRNA from each sample was ~10 μg, which is only enough to hybridize one chip. The murine 30K chip set contains five chips. Sequential hybridizations were therefore carried out with the same sample material in order to hybridize all five chips; presumably each cRNA from the sample bound to only the chip that contained the matching sequence.

The GeneChip Analysis Suite V3.0 (Affymetrix, SantaClara, CA) was used to analyze the data. Each entry on a GeneChip is represented by 40 spots: 20 Perfect Match oligonucleotides each have a unique sequence exactly matching the target gene, and 20 MisMatch oligonucleotides each have the same sequences as the cognate Perfect Match oligonucleotides with the exception of one base in the middle. The signals derived from hybridization to MisMatch oligonucleotides therefore represent the nonspecific background. The analysis software quantifies the results using several different parameters: *Average Difference* (Avg Diff) is an indicator of the expression level of a transcript (ranging from 0 to >50,000). It was calculated by taking the difference between the signals of Perfect Match oligonucleotides and MisMatch oligonucleotides and averaging the differences over the entire probe set. The analysis software scaled different experimental data sets to the same average level (1500 as defined by the software) to allow comparison. When the Avg Diff was less than zero, it was arbitrarily set to zero. *Absolute Call* is determined by a decision matrix in Affymetrix's software that is designed to distinguish signals from specific and nonspecific hybridization to score the presence or absence of each transcript. Three calls were made for each transcript: P–present, M–marginal, or A–absent. Genes with higher Avg Diff were more likely to be called present. However, it is possible that a gene may have a high Avg Diff value yet still be called A (absent) by the software; conversely, genes with low Avg Diff can be called present. *Fold Change* reflects the expression level changes. It was based on the ratio of Avg Diff between the experimental and baseline data and incorporates the noise level into the calculation.

RESULTS

A large number of genes is expressed in the cochlea

From four hybridizations (two P2 and two P32 cochleas), a total of 17,300 genes and ESTs were found to be expressed in at least one of the samples, which represents 51% of the genes and ESTs on the chip. Individually, 13,456, 13,551, 11,785, and 12,871 genes and ESTs were detected in P2a, P2b, P32a, and P32b, respectively.

TABLE 1

Detection of hair-cell genes in the whole-cochlea samples

Genes	*P2a*[a]	*P2b*[a]	*P32a*[b]	*P32b*[b]
Myosin-Iβ	+	+	+	+
Myosin-VI	+	+	−	+
Myosin-VIIa	−	+/−	−	−
Calretinin	+	+	+	+
Parvalbumin	+	+	+	+
S100	+	+	+	+
Oncomodulin	−	−	+	+
Diaphanous	−	+	+	+
p27 Kip2	+	+	−	−
Math1	−	−	−	−
PMCA2	−	+/−	−	−

[a] P2a and P2b are the two hybridizations done with the P2 mouse cochlea.
[b] P32a and P32b are the two P32 mouse cochleas. A + indicates a "present" call, − indicates "absent," and +/− is "marginal."

It is estimated that there are about 3000 hair cells in each mouse cochlea, so each sample used for cRNA preparation and hybridization (from ten cochleas) contained ~30,000 hair cells. However, hair cells constitute only a few percent of the total number of cells in the cochlea. Since some known hair cell genes are represented on the GeneChip, we first asked how many of them were detected. Over half of a representative group of known hair cell genes were detected in the whole cochlea sample (Table 1). A previous study has shown that many genes with 2–5 copies/cell can be detected from samples with 50,000 cultured cells (Mahadevappa and Warrington 1999). However, the sensitivity is directly related to the amount of cRNA used in the hybridization. With ~10 μg of cRNA for each hybridization, compared with 15 μg as suggested by Affymetrix, a decreased sensitivity of detection of ~15% has been observed (Mahadevappa and Warrington 1999). In our experiment using the spiked internal controls at various concentrations (the biotin synthesis pathway genes from the bacteria *E.coli*—bioB, bioC, bioD—supplied by Affymetrix), we could detect most of the transcripts present in 3–5 copies/cell (data not shown). With an average of 11 μg cRNA for each hybridization, and with hair cells constituting a small percentage of total cells, we expect that many hair cell genes with transcripts of more than 4–10 copies/cell were detected but that genes expressed at lower levels may have been missed.

To determine sources contributing to the variability of this method, we examined each developmental stage twice. For P2 mice, 20 cochleas were collected

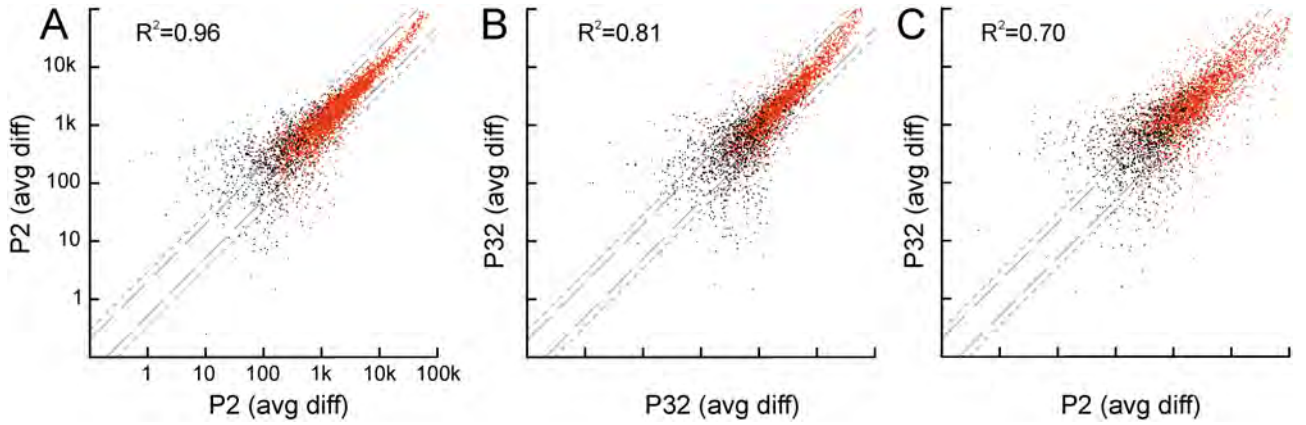

FIG. 1. Scatter plots showing the correlation of expression patterns. **A** Two P2 samples compared with each other. **B** Two P32 samples. **C** A P32 sample compared with a P2 sample. Average difference values (Avg Diff determined by the analysis software) were used to quantify expression and were plotted on a log scale. The correlation coefficient between samples (R^2) is also shown for each comparison. The dashed gray line demarcates less than twofold difference between samples in Avg Diff, whereas the dotted gray line demarcates a less than threefold difference. Red dots represent the genes that were scored as present in at least one of the two samples, and black dots are genes scored as not above background. A higher correlation is evident in **A** and **B,** with **A** showing the highest correlation. The lower correlation in **C** indicates changes in gene expression during development.

and total RNA was extracted. Then the RNA was divided into halves and each half was used to synthesize the cDNA and cRNA. Differences in measured expression between the two samples should reflect the variations inherent in cDNA and cRNA synthesis, variation between chips, and subsequent hybridization and scanning. For P32 mice, two batches of ten cochleas were collected independently. Each batch then went through RNA extraction, cDNA and cRNA synthesis, and hybridization. Differences between these samples should include the additional variation brought about by sample collection and RNA extraction.

In both P2 hybridizations about 13,500 genes and ESTs were classified by the software as being expressed. However, about 20% were scored as expressed in one or the other, but not both. A scatter plot of the expression level, as well as the calculated correlation coefficient, showed that overall the two expression patterns are very similar (Fig. 1). Most genes with high expression have expression levels that differ by less than twofold between the two samples, as indicated by dashed lines.

In the P32 samples, about 12,000 genes and ESTs were scored as expressed. About 9400 genes and ESTs were found in both, but about 24% were found to be expressed in just one or the other. The scatter plot showed that, as in P2 samples, most of the genes and ESTs classified as present have expression levels within twofold between samples. However, more genes in the P32 samples had greater than twofold variation between samples, even among highly expressed genes. The overall correlation of expression between the two P32 samples was less than in the two P2 samples, as indicated by a lower correlation coefficient. Therefore, a larger variation is likely to be introduced by the sample collection procedure than by cRNA synthesis alone.

Much of the variation between two samples from the same developmental stage is associated with the genes with low expression levels. When the 2000 genes with the highest expression levels (represented by Avg Diff $\geq$7221 in the case for P2 and $\geq$7730 for P32) were compared between the two P2 or the two P32 samples, 97% and 94% of them, respectively, showed concordance in being classified as present or absent in both samples. However, when 2000 genes with expression levels close to background were chosen (Avg Diff between 1000 and 1290), only 75% and 70% of genes showed concordance.

Finding candidate genes within a large gene family

Wnt genes encode a group of secreted signaling molecules involved in cell–cell interaction. They are especially important in the development of neural crest, kidney, body axis, distal limbs, and planar polarity of the eye (Mlodzik 1999; Wodarz and Nusse 1998); in addition, the Wnt signaling pathway plays a central role in the etiology of colorectal cancer (Bienz and Clevers 2000). Recently, Wnt-8c was found to act synergistically with Egf-19 in inner ear development to mediate mesodermal and neural signals (Ladher et al. 2000). A Wnt-5a knockout mouse was found to have a defect in the outgrowth of the external ears (Wang et al. 1998a; Yamaguchi et al. 1999). However, a link between most of the Wnt genes and inner ear development has not been described.

There are at least 20 mammalian Wnt genes divided into 16 subgroups (http://www.stanford.edu/

TABLE 2

Wnt gene expression in cochlea[a]

Samples	*Wnt1*	*Wnt3*	*Wnt3A*	*Wnt4*	*Wnt5a*	*Wnt5b*	*Wnt6*	*Wnt7a*	*Wnt7b*	*Wnt8b*	*Wnt8D*	*Wnt10a*	*Wnt10b*	*Wnt11*
P2a	−	−	−	−	+	−	−	−	+	−	−	−	−	−
P2b	−	−	−	+	+	+/−	−	−	+	−	−	+/−	−	−
P32a	−	−	−	+	+	−	−	−	+	−	−	−	−	−
P32b	−	−	−	+	+	+/−	−	−	+	−	−	−	−	−

[a]Wnt genes on the GeneChip were identified by searching the database with the "Find" function and putting "Wnt" in the EntrezDef field. There may be additional ESTs that represent Wnt genes but are not identified as such. Presence or absence (+ or −) indicated is according to the AbsCall field in the database.

~rnusse/Wntgenes/mouseWnt.html), and 14 of them are represented on the GeneChips. By examining the expression of these 14 Wnt genes, we found five Wnt genes to be expressed in the cochlea at P2 or P32 (Table 2): Wnt-4, Wnt-5a, and Wnt-7b were expressed at both stages, while Wnt-5b and Wnt-10 were scored as marginal in one of two samples. Wnt-4 and Wnt-5a had previously been reported to be expressed in the developing inner ear (Kelley 2000; Wang et al. 1998a). Wnt-7a is also reportedly expressed but we failed to detect it with GeneChips, which may result from a lack of Wnt-7a expression at the developmental stages studied or a lack of sensitivity. To understand the function of Wnt genes in inner ear development, this analysis identifies five (Wnt-4, -5a, -5b, -7b, and -10) out of the 14 tested as candidates for further study.

Genes differentially expressed between two stages reflect developmental and functional differences

A comparison between P2 and P32 cochleas should reveal many genes that are up- or down-regulated during development. Because there is some random variation in measured expression level even at the same developmental stage, we compared only genes that were scored as either present or absent in both hybridizations of a given age. More than 700 genes are found to be up-regulated from P2 to P32 (defined as more than doubling in expression level), whereas less than 800 were down-regulated. A few examples illustrate the analysis.

The gene with the greatest down-regulation was otoconin-95. Expression levels of otoconin-95 decreased over 100-fold from P2 to P32, and at P32 expression was scored as absent. Otoconin-95 is expressed by nonsensory epithelial cells of the membranous labyrinth and is a major component of otoconia (Verpy et al. 1999; Wang et al. 1998b). The expression data suggest that, after the initial formation of otoconia, very little additional otoconin-95 may be required for their maintenance.

Like otoconin-95, both α- and β-tectorin are dramatically down-regulated at P32, a clear indication of their initial involvement in tectorial membrane formation but limited turnover once it is formed. This correlates well with a recent study of tectorin expression, which showed robust expression of α- and β-tectorin in the cochlea beginning at embryonic day 14 (E14) but undetectable expression by P22 (Legan et al. 2000; Rau et al. 1999).

Fourteen collagen genes, represented at 27 locations on the GeneChip, were found to be down-regulated from P2 to P32. Only two collagen genes were up-regulated. As collagens are a major component of the cochlea, serving, for instance, as a constituent of the tectorial membrane, the down-regulation of a large number of collagen genes suggests that these are required mostly during early development of cochlea.

A large number of genes with functions in chromosome organization, cell cycle, and protein synthesis are down-regulated at P32, including histones, D-type cyclins, and ribosomal proteins. This indicates that the mouse cochlea is undergoing considerably reduced cell division at P32 compared with P2. D-type cyclins are primarily involved in the G1/S transition during the cell cycle, forming a serine/threonine kinase holoenzyme complex with cdk4 (Hirai et al. 1995). Indeed, cdk4 also shows significant down-regulation from P2 to P32.

Interestingly, a group of negative cell-growth-control genes—including melanoma-activity-inhibition gene, *Kip2, gas1,* and *gas5*—are down-regulated during this period. *Kip2,* a cyclin-dependent kinase inhibitor, is expressed in rat hair cells at P0 but not at P35 (Corwin et al. 2000). *Kip2* has been shown to be upregulated in G(1)/G(0) phase, preventing the cells from entering S-phase, both in retina and myoblasts (Dyer and Cepko 2000; Reynaud et al. 1999). The higher expression of *Kip2* at P2 suggests that Kip2 is required at an early postnatal stage to inhibit the activity of cyclin kinase, for differentiation of hair cells. Later in development the cyclin kinase activity may be completely diminished so that *Kip2* is not required, or alternatively the role of *Kip2* is replaced by similar gene(s). The GeneChip analysis therefore provides us with a candidate gene potentially involved in hair cell

TABLE 3

Genes up-regulated in cochlea from P2 to P32

Channels
- Ion channel homolog RIC
- Sodium channel beta-1
- Brain potassium channel protein-1
- Calcium channel gamma subunit (Cacng6)
- Potassium channel (Kcnk1)
- Mercurial-insensitive water channel 2 (mMIWC2)
- Aquaporin-4 (Aqp4)
- MAT8
- Clcn3
- Connexin 30
- Connexin 43

Transporters
- Sodium-dependent choline transporter 2
- FXYD domain-containing ion transport regulator 6
- Amino acids transporter NAT-1
- Solute carrier family 22 (organic cation transporter) (Slc22a5)
- Solute carrier family 20 (Slc20a1)
- Na+/H+ exchanger (NHE-1) (Slc9a1)
- Calcium-binding transporter
- Mitochondrial carnitine/acylcarnitine carrier protein
- Carrier protein C2
- Na+, K+ ATPase, alpha 1
- Na+, K+ ATPase, alpha 2
- Na+, K+ ATPase, alpha 3
- Na+, K+ ATPase, beta
- Na+, K+ ATPase, beta 1
- Na+, K+ ATPase, beta 2
- Na+, K+ ATPase, beta 3
- Na+, K+ ATPase, gamma
- Ca++ ATPase, sarcoplasmic reticulum type
- Similar to AF159856 N system amino acid transporter NAT-1
- Calcium-binding transporter

Calcium-Binding Proteins
- Parvalbumin
- Intracellular calcium-binding protein (MRP8)
- Calretinin
- S100 A13
- Oncomodulin

differentiation and provides the timing of its expression. *Gas1, gas5,* and melanoma-activity-inhibition gene have not been previously shown to be expressed in the inner ear. Their association with the differentiation process in other tissues (Blesch et al. 1994; Coccia et al. 1992; Del Sal et al. 1992; Evdokiou and Cowled 1998; Stoll et al. 2001) suggests a role in cochlea development.

Among genes up-regulated from P2 to P32 are channels (10) and transporters (12) (Table 3). These include a sodium channel subunit, potassium channels, a chloride channel, and water channels. The ion transporters include Na/K-ATPases and a calcium transporter. (Another 50 channels are expressed at both stages.) The mature cochlea has an elaborate system of ion transport in the organ of Corti and stria vascularis, which serves to maintain the high-potassium endolymph (Steel and Kros 2001). In addition, hair cells and the spiral ganglion neurons on which they synapse use a variety of ion channels to convey the encoded acoustic stimulus. The up-regulation of both channel and transporter genes is consistent with these functions.

Up-regulated as well are genes encoding calcium-binding proteins, including S-100, oncomodulin, calretinin, parvalbumin, and intracellular calcium-binding protein (MRP8). Some of these are highly expressed in hair cells (Pack and Slepecky 1995; Sakaguchi et al. 1998), which must buffer considerable calcium influx through both transduction channels and synaptic voltage-gated calcium channels.

A searchable database

To make these data most accessible to the research community, we have constructed a searchable database using FileMaker Pro (Filemaker, Santa Clara, CA), and it can be downloaded from our laboratory web site (http://www.mgh.harvard.edu/depts/coreylab/index.html). The download file is about 10 Mb and is unzipped to a 30 Mb database. Users may search for genes using any of three attributes (Gene Name, UniGene ID, or Accession Number). In the case of a novel gene that has neither a name nor an accession number, one can use an EST for that gene to identify a UniGene cluster and then search with the UniGene number.

The database can be used to retrieve expression information on genes expressed in the cochlea. For example, one can type in "calretinin" (a calcium-binding protein) in the "EntrezDef" field to learn that calretinin is expressed in both P2 and P32 mouse cochlea. The up-regulation of the gene at P32 vs. P2 is evident by the high Avg Diff value at P32 or by the large "fold change" (>5). In addition, links to various public databases are included to facilitate understanding of protein function. Each gene entry has links, when available, to the Mouse Genome Informatics (MGI), TIGR, and UniGene databases to provide more information.

It is our intention to update the database frequently, to include additional expression data from developing mouse utricle and cochlea and from human inner ear. More extensive links will be added to the database, such as functional annotation of genes, updated UniGene ID, chromosome locations, and MGI entries.

DISCUSSION

We have shown using oligonucleotide array technology that a large number of genes are expressed in the mouse cochlea. Even though we have only four tissue samples in the present study, the expressed genes represent the largest number yet described in the mammalian inner ear.

Since the genes identified to date and arrayed on chips are primarily from tissues other than the inner ear, there could be concern that array technology would miss genes important for hearing. We think the concern is unfounded: First, many genes can play multiple roles in different tissues, so it is not surprising to notice a large number of deafness genes also expressed in other tissues (e.g., myosin VIIa in testis, myosin VI in kidney, Brn-3.1 in the retina, and Math1 in brain). Perhaps relatively few genes are expressed only in the inner ear. Second, the number of known genes is increasing rapidly. The Human Genome Project has produced a draft sequence of the human genome and similar progress is anticipated for the Mouse Genome Project within 2001 (Collins et al. 1998; Lander et al. 2001; Venter et al. 2001). The results of the human genome sequencing project indicate that there are ~30,000 genes in the human genome. More than 30,000 sequences are represented on the mouse GeneChips used here and these most likely represent ~20,000 distinct genes (data from Affymetrix), so it is likely that a significant fraction of the mouse genome is already covered. Chips derived from genomic sequence rather than expressed sequences, and that cover most of the genes in the genome, are expected in the near future. It is likely that very soon all mouse and human genes and many of their splice variants will be represented on GeneChips. These can be used to identify rare inner ear genes without having to sequence inner ear libraries.

Sensitivity of detection is of great importance for expression profiling in the inner ear because of limited material. Using material harvested from the whole cochlea, we detected a number of genes expressed by hair cells (myosin VI, calretinin, parvalbumin). Some genes were not consistently detected in our study, such as myosin VIIa, diaphanous, and PMCA2. There are several possible reasons for this. First, the detection sensitivity is directly related to the number of hair cells present in the sample. In our study we typically used 10 cochleas (with ~30,000 hair cells) per assay. In a separate study of isolated utricular macula (with ~50,000 hair-cells), most known hair-cell genes were detected (Z-Y Chen, data not shown). Therefore, by increasing the amount of tissue used, more genes should be detected. Complexity of tissue used can also affect sensitivity of detection. Since the cochlea consists of many different cell types the abundance of cell-type-specific transcripts was reduced. By using a particular organ with fewer cell types (such as the organ of Corti), we expect to improve the sensitivity of detection.

Very-low-copy-number mRNAs (e.g., 1–2 copies/cell) will not be detected by the GeneChip assay or any other microarray approaches, using the current protocol for amplification. One way around this problem is to use a normalized inner ear cDNA library as the source for cRNA synthesis. The probability of detecting low-copy-number genes would be greatly increased but it would not be quantitatively representative of expression level. It is also possible to perform SAGE analysis using the normalized cDNA libraries. The combination of microarray and SAGE analysis may help to identify most of the rare transcripts.

Results from the GeneChip assay are largely reproducible between samples, at least for genes expressed at levels significantly above background. The higher degree of concordance between the two P2 hybridizations compared with the P32 hybridizations can be explained mainly by the common sample collection and RNA extraction. A recent study using a well-controlled cell line has revealed the importance of duplication in reducing variations in a microarray assay (Lee et al. 2000). Our data confirm the usefulness of running duplicate samples to determine the inherent variability in sample preparation, especially for the limited tissue obtainable from the inner ear.

These data, from these two postnatal developmental stages, cannot reveal genes important in early development of the inner ear. The limited sample numbers also preclude the application of algorithms (such as Genecluster) to group genes involved in different functional pathways. Therefore, future experiments should comprehensively profile expression patterns in the developing inner ear, at a variety of pre- and postnatal stages. Similarly, cochlear samples include a large number of cell types, each with their own function and developmental program. Greater insight will come from expression profiles of individual organs, such as the utricle and the organ of Corti, or even single cell types.

GeneChip studies can produce an overwhelming amount of data, and insight may come only through intelligent sifting procedures. Having a global view of expression patterns of inner ear genes can help to narrow down quickly the genes potentially important in the development and function of this essential organ. The survey of the Wnt gene family identified some members that are more likely than others to control ear development. A similar survey can be performed with members of the Frizzled gene family of seven-transmembrane receptors for Wnt. The time course of Frizzled gene expression correlated with that of Wnt genes could identify the potential Frizzled receptors for various Wnt isoforms, in the context of inner ear development. Finally, with more samples it will be possible to cluster the expression data to explore functional pathways, such as downstream effectors of Wnt signaling, at particular developmental stages.

Annotations and links in our current database are

just the first step in enabling convenient searches. A future goal is to include more links to functional annotations, chromosomal locations, and expanded MGI ID. For instance, the inclusion of chromosomal locations should expedite the search for deafness genes.

ACKNOWLEDGMENTS

We thank Dr. Heidi Rehm and Dr. Duan-Sun Zhang for dissection of the mouse cochleas, Mr. Eric Bloomquist for the database setup, and the MIT Biopolymer Core Facility for performing hybridizations. This work was supported by NIDCD grant RO1-DC02281 and NIDCD contract 263MD008288 (to DPC), and by a grant from the Deafness Research Foundation (to ZYC). DPC is an Investigator of the Howard Hughes Medical Institute.

REFERENCES

BIENZ M, CLEVERS H. Linking colorectal cancer to Wnt signaling. Cell 103:311–320, 2000.

BLESCH A, BOSSERHOFF AK, APFEL R, BEHL C, HESSDOERFER B, SCHMITT A, JACHIMCZAK P, LOTTSPEICH F, BUETTNER R, BOGDAHN U. Cloning of a novel malignant melanoma-derived growth-regulatory protein, MIA. Cancer Res. 54:5695–5701, 1994.

COCCIA EM, CICALA C, CHARLESWORTH A, CICCARELLI C, ROSSI GB, PHILIPSON L, SORRENTINO V. Regulation and expression of a growth arrest-specific gene (*gas5*) during growth, differentiation, and development. Mol. Cell. Biol. 12:3514–3521, 1992.

COLLINS FS, PATRINOS A, JORDAN E, CHAKRAVARTI A, GESTELAND R, WALTERS L. New goals for the U.S. Human Genome Project: 1998–2003. Science 282:682–689, 1998.

Corwin JT, Montcouquiol M, Shi Q, Karaoli T, Karimi K, Whitley M, Byrne M (2000) Molecular changes associated with the loss of proliferation in mammalian hair cell epithelia. Assoc. Res. Otolaryngol. Abstr. 5702.

DEL SAL G, RUARO ME, PHILIPSON L, SCHNEIDER C. The growth arrest-specific gene, *gas1,* is involved in growth suppression. Cell 70:595–607, 1992.

DYER MA, CEPKO CL. p57(Kip2) regulates progenitor cell proliferation and amacrine interneuron development in the mouse retina. Development 127:3593–3605, 2000.

EVDOKIOU A, COWLED PA. Tumor-suppressive activity of the growth arrest-specific gene GAS1 in human tumor cell lines. Int. J. Cancer 75:568–577, 1998.

GONG TW, HEGEMAN AD, SHIN JJ, ADLER HJ, RAPHAEL Y, LOMAX MI. Identification of genes expressed after noise exposure in the chick basilar papilla. Hear. Res. 96:20–32, 1996.

HELLER S, SHEANE CA, JAVED Z, HUDSPETH AJ. Molecular markers for cell types of the inner ear and candidate genes for hearing disorders. Proc. Natl. Acad. Sci. U.S.A. 95:11400–11405, 1998.

HIRAI H, ROUSSEL MF, KATO JY, ASHMUN RA, SHERR CJ. Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. Mol. Cell. Biol. 15:2672–2681, 1995.

KELLEY MW (2000) A potential role for the Wnt signaling pathway in the development of the organ of Corti. Assoc. Res. Otolaryngol. Abstr. 5593.

LADHER RK, ANAKWE KU, GURNEY AL, SCHOENWOLF GC, FRANCIS–WEST PH. Identification of synergistic signals initiating inner ear development. Science 290:1965–1967, 2000.

LANDER ES. Array of hope. Nat. Genet. 21:3–4, 1999.

LANDER ES, LINTON LM, BIRREN B, NUSBAUM C, ZODY MC, BALDWIN J, DEVON K, DEWAR K, DOYLE M, FITZHUGH W, FUNKE R, GAGE D, HARRIS K, HEAFORD A, HOWLAND J, KANN L, LEHOCZKY J, LEVINE R, MCEWAN P, MCKERNAN K, MELDRIM J, MESIROV JP, MIRANDA C, MORRIS W, NAYLOR J, RAYMOND C, ROSETTI M, SANTOS R, SHERIDAN A, SOUGNEZ C, STANGE-THOMANN N, STOJANOVIC N, SUBRAMANIAN A, WYMAN D, ROGERS J, SULSTON J, AINSCOUGH R, BECK S, BENTLEY D, BURTON J, CLEE C, CARTER N, COULSON A, DEADMAN R, DELOUKAS P, DUNHAM A, DUNHAM I, DURBIN R, FRENCH L, GRAFHAM D, GREGORY S, HUBBARD T, HUMPHRAY S, HUNT A, JONES M, LLOYD C, MCMURRAY A, MATTHEWS L, MERCER S, MILNE S, MULLIKIN JC, MUNGALL A, PLUMB R, ROSS M, SHOWNKEEN R, SIMS S, WATERSTON RH, WILSON RK, HILLIER LW, MCPHERSON JD, MARRA MA, MARDIS ER, FULTON LA, CHINWALLA AT, PEPIN KH, GISH WR, CHISSOE SL, WENDL MC, DELEHAUNTY KD, MINER TL, DELEHAUNTY A, KRAMER JB, COOK LL, FULTON RS, JOHNSON DL, MINX PJ, CLIFTON SW, HAWKINS T, BRANSCOMB E, PREDKI P, RICHARDSON P, WENNING S, SLEZAK T, DOGGETT N, CHENG JF, OLSEN A, LUCAS S, ELKIN C, UBERBACHER E, FRAZIER M. Initial sequencing and analysis of the human genome. Internation Human Genome Sequencing Consortium. Nature 409:860–921, 2001.

LEE ML, KUO FC, WHITMORE GA, SKLAR J. Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. Proc. Natl. Acad. Sci. U. S. A. 97:9834–9839, 2000.

LEGAN PK, LUKASHKINA VA, GOODYEAR RJ, KOSSI M, RUSSELL IJ, RICHARDSON GP. A targeted deletion in alpha-tectorin reveals that the tectorial membrane is required for the gain and timing of cochlear feedback. Neuron 28:273–285, 2000.

MAHADEVAPPA M, WARRINGTON JA. A high-density probe array sample preparation method using 10- to 100-fold fewer cells. Nat. Biotechnol. 17:1134–1136, 1999.

MLODZIK M. Planar polarity in the Drosophila eye: a multifaceted view of signaling specificity and cross-talk. EMBO J 18:6873–6879, 1999.

PACK AK, SLEPECKY NB. Cytoskeletal and calcium-binding proteins in the mammalian organ of Corti: cell type-specific proteins displaying longitudinal and radial gradients. Hear. Res. 91:119–135, 1995.

RAU A, RICHARDSON GP. Tectorin mRNA expression is spatially and temporally restricted during mouse inner ear development. J. Comp. Neurol. 405:271–280, 1999.

REYNAUD EG, PELPEL K, GUILLIER M, LEIBOVITCH MP, LEIBOVITCH SA. p57(Kip2) stabilizes the MyoD protein by inhibiting cyclin E-Cdk2 kinase activity in growing myoblasts. Mol. Cell. Biol. 9:7621–7629, 1999.

SAKAGUCHI N, HENZL MT, THALMANN I, THALMANN R, SCHULTE BA. Oncomodulin is expressed exclusively by outer hair cells in the organ of Corti. J. Histochem. Cytochem. 46:29–40, 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.

STEEL KP, KROS CJ. A genetic approach to understanding auditory function. Nat. Genet. 27:143–149, 2001.

STOLL R, RENNER C, ZWECKSTETTER M, BRUGGERT M, AMBROSIUS D, PALME S, ENGH RA, GOLOB M, BREIBACH I, BUETTNER R, VOELTER W, HOLAK TA, BOSSERHOFF AK. The extracellular human melanoma inhibitory activity (MIA) protein adopts an SH3 domain-like fold. EMBO J. 20:340–349, 2001.

VELCULESCU VE, VOGELSTEIN B, KINZLER KW. Analysing uncharted transcriptomes with SAGE. Trends Genet. 16:423–425, 2000.

VELCULESCU VE, ZHANG L, VOGELSTEIN B, KINZLER KW. Serial analysis of gene expression. Science 270:484–487, 1995.

VENTER JC, ADAMS MD, MYERS EW, LI PW, MURAL RJ, SUTTON GG,

SMITH HO, YANDELL M, EVANS CA, HOLT RA, GOCAYNE JD, AMANATIDES P, BALLEW RM, HUSON DH, WORTMAN JR, ZHANG Q, KODIRA CD, ZHENG XH, CHEN L, SKUPSKI M, SUBRAMANIAN G, THOMAS PD, ZHANG J, GABOR MIKLOS GL, NELSON C, BRODER S, CLARK AG, NADEAU J, MCKUSICK VA, ZINDER N, LEVINE AJ, ROBERTS RJ, SIMON M, SLAYMAN C, HUNKAPILLER M, BOLANOS R, DELCHER A, DEW I, FASULO D, FLANIGAN M, FLOREA L, HALPERN A, HANNENHALLI S, KRAVITZ S, LEVY S, MOBARRY C, REINERT K, REMINGTON K, ABU-THREIDEH J, BEASLEY E, BIDDICK K, BONAZZI V, BRANDON R, CARGILL M, CHANDRAMOULISWARAN I, CHARLAB R, CHATURVEDI K, DENG Z, FRANCESCO VD, DUNN P, EILBECK K, EVANGELISTA C, GABRIELIAN AE, GAN W, GE W, GONG F, GU Z, GUAN P, HEIMAN TJ, HIGGINS ME, JI RR, KE Z, KETCHUM KA, LAI Z, LEI Y, LI Z, LI J, LIANG Y, LIN X, LU F, MERKULOV GV, MILSHINA N, MOORE HM, NAIK AK, NARAYAN VA, NEELAM B, NUSSKERN D, RUSCH DB, SALZBERG S, SHAO W, SHUE B, SUN J, WANG ZY, WANG A, WANG X, WANG J, WEI MH, WIDES R, XIAO C, YAN C ET AL. The Sequence of the Human Genome. Science 291:1304–1351, 2001.

VERPY E, LEIBOVICI M, PETIT C. Characterization of otoconin-95, the major protein of murine otoconia, provides insights into the formation of these inner ear biominerals. Proc. Natl. Acad. Sci. U.S.A. 96:529–534, 1999.

WANG W, VAN DE WATER T, LUFKIN T. Inner ear and maternal reproductive defects in mice lacking the *Hmx3* homeobox gene. Development 125:621–634, 1998a.

WANG Y, KOWALSKI PE, THALMANN I, ORNITZ DM, MAGER DL, THALMANN R. Otoconin-90, the mammalian otoconial matrix protein, contains two domains of homology to secretory phospholipase A2. Proc. Natl. Acad. Sci. U.S.A. 95:15345–15350, 1998b.

WODARZ A, NUSSE R. Mechanisms of Wnt signaling in development. Annu. Rev. Cell. Dev. Biol. 14:59–88, 1998.

YAMAGUCHI TP, BRADLEY A, MCMAHON AP, JONES S. A Wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo. Development 126:1211–1223, 1999.

ZHANG L, ZHOU W, VELCULESCU VE, KERN SE, HRUBAN RH, HAMILTON SR, VOGELSTEIN B, KINZLER KW. Gene expression profiles in normal and cancer cells. Science 276:1268–1272, 1997.

ZHENG QY, JOHNSON KR, ERWAY LC. Assessment of hearing in 80 inbred strains of mice by ABR threshold analyses. Hear. Res. 130:94–107, 1999.