Molecular architecture of the chick vestibular hair bundle

Jung-Bum Shin1,2,8, Jocelyn F Krey2,8, Ahmed Hassan3, Zoltan Metlagel3, Andrew N Tauscher3, James M Pagana1,2, Nicholas E Sherman1, Erin D Jeffery1, Kateri J Spinelli2, Hongyu Zhao2, Phillip A Wilmarth5, Dongseok Choi6, Larry L David3, Manfred Auer3 & Peter G Barr-Gillespie2,7

Hair bundles of the inner ear have a specialized structure and protein composition that underlies their sensitivity to mechanical stimulation. Using mass spectrometry, we identified and quantified >1,100 proteins, present from a few to 400,000 copies per stereocilium, from purified chick bundles; 336 of these were significantly enriched in bundles. Bundle proteins that we detected have been shown to regulate cytoskeleton structure and dynamics, energy metabolism, phospholipid synthesis and cell signaling. Three-dimensional imaging using electron tomography allowed us to count the number of actin-actin cross-linkers and actin-membrane connectors; these values compared well to those obtained from mass spectrometry. Network analysis revealed several hub proteins, including RDX (radixin) and SLC9A3R2 (NHERF2), which interact with many bundle proteins and may perform functions essential for bundle structure and function. The quantitative mass spectrometry of bundle proteins reported here establishes a framework for future characterization of dynamic processes that shape bundle structure and function.

An outstanding example of a specialized organelle devoted to a single purpose, the vertebrate hair bundle transduces mechanical signals for the inner ear, converting sound and head movement to electrical signals that propagate to the central nervous system. Protruding from the apical surface of a sensory hair cell, a bundle typically consists of 50–100 actin-filled stereocilia and, at least during development, an axonemal kinocilium1. A bundle enlists ~100 transduction channels, which are mechanically gated by tip links as external forces oscillate the bundle; the opening and closing of the channels in turn modulates the hair cell’s membrane potential, controlling neurotransmitter release.

Because hair bundles have a reduced protein complement and carry out a specialized task, once we know which proteins are present—as well as their concentrations and interactions—understanding bundles’ assembly and operation seems possible. Although genetics studies have identified many proteins essential for bundle function2, others may have escaped detection because they are essential during development or, in some cases, can be compensated for by paralogs. To discover these additional proteins, biochemical strategies are essential. Although bundles are scarce, quantitative mass spectrometry3 has the sensitivity and accuracy to detect and quantify the bundle’s protein complement.

Our previous analysis of hair-bundle proteins using mass spectrometry detected 59 proteins, including several that are critical for bundle function4. Here, using a more sensitive mass spectrometer, we detected over 1,100 proteins from chick vestibular bundles and identified those proteins selectively targeted to bundles. Many bundle-enriched proteins are expressed from deafness-associated genes, confirming their essential function in the inner ear. We also imaged stereocilia using electron tomography and counted actin-actin cross-linkers and actin-membrane connectors; those counts compared favorably to mass-spectrometric estimates for cross-linker and connector proteins. To place the bundle’s proteome into a network of functional and structural interactions, we assembled an interaction map that highlights the central roles in hair-bundle function played by actin, phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P2), Ca2+ and CALM (calmodulin). Moreover, we identified two other key hub proteins: the ezrin-radixin-moesin (ERM) family member RDX (radixin), important in hair-bundle function5, and SLC9A3R2 (NHERF2; solute carrier family 9 member 3 regulator 2), a PDZ-domain adaptor protein that couples RDX to many transmembrane proteins6. The comprehensive view offered by quantitative mass spectrometry reveals functional pathways in hair bundles and, on the basis of the absence of key protein families, also rules out alternative mechanisms.

RESULTS

Mass spectrometry of purified hair bundles

Using liquid-chromatographic tandem mass spectrometry (LC-MS/MS), we identified proteins from hair bundles and epithelia of utricles (Supplementary Fig. 1), vestibular organs that detect linear acceleration, from embryonic day 20–21 chicks; at this age, utricles are functional7. Bundles (BUN) were enriched 40-fold, to ~80% purity.


Figure 1. Quantitative analysis of chick hair-bundle proteins. (a) Top, proteins identified in bundles and epithelium (two or more experiments). Bottom left, representation of bundle proteins as bundle-enriched, unenriched and epithelium-enriched (by protein frequency). Bottom right, same as left except with the mole fractions of proteins in each class summed. (b) Calibration curve relating mole fraction of human protein standards spiked into *E. coli* extract to riBAQ value. The number of identified proteins is indicated for each data point (mean ± s.d.). The points corresponding to mole fractions of 10^{-2}, 10^{-3} and 10^{-4} were fit with a line constrained through the origin (y = 1.02x; R = 0.999). (c) Calibration curve relating mole fraction determined from riBAQ values to mole fraction measured using quantitative immunoblots with purified proteins as standards; data points are mean ± s.e.m. and are fit by y = 0.98x (R = 0.97). Data for CKB and GAPDH were from ref. 4. Dotted lines in a,b correspond to unity. (d) Abundance distribution of bundle and epithelium proteins; single Gaussian fits. (e) Enrichment distribution of proteins detected in bundles and epithelium; single Gaussian fit. (f) Cumulative protein molar abundance, from highest to lowest. The most abundant proteins in bundles and epithelium are indicated. (g) Mole fractions of proteins in epithelium (left) and bundle (right); the slope of the line connecting them represents bundle-to-epithelium enrichment.

Proteins most highly enriched in the epithelium are indicated at left, those most highly enriched in bundles at right. Hue represents relative enrichment (power coefficient of fit connecting points) for each protein. Far right, proteins detected only in bundles.

Quantification using relative iBAQ values

To quantify hair-bundle proteins, we used the iBAQ algorithm, which divides the sum of all precursor-peptide intensities by the number of theoretically observable peptides^11. We normalized each protein’s iBAQ value to the sum of all iBAQ values, generating a relative iBAQ (riBAQ) value for each protein. Although a previous report demonstrated the linearity of the iBAQ approach^11, we sought a more rigorous validation: does riBAQ accurately report the mole fraction of each protein? We mimicked experiments with complex protein mixtures by detecting human proteins diluted in an *Escherichia coli* extract as a protein background. Only the more abundant human proteins were detected, demonstrating the limitations in detecting proteins at low mole fractions. We carried out a linear regression (log_{10} riBAQ = 1.02 ± 0.01 × log_{10} mole fraction) with the 10^{-2}, 10^{-3} and 10^{-4} mole fraction data points (Fig. 1b). Although the 10^{-5} data point did not fall on the regression line, only two of eight proteins were detected. We conclude that the correspondence between riBAQ and mole fraction is nearly exact, at least above to a mole fraction of ~10^{-5} (Fig. 1b).

To independently verify our riBAQ calibration, we measured the concentrations of five hair-bundle proteins using quantitative immunoblotting (Fig. 1c). We only used proteins for which we had purified protein standards, which allowed us to generate accurate standard curves with known amounts of protein. The fit was very close to 1:1 (y = 0.98x; R = 0.97), confirming that riBAQ values reported mole fraction accurately.

Quantification of hair-bundle proteins

With knowledge of total number of molecules per stereocilium, or of molecules per stereocilium of one accurately measured protein, mole fraction values can be used to estimate the number of molecules per stereocilium for any protein. Because actin monomers are present at minimal levels in stereocilia^8 and because each stereocilium has 400,000 filamentous actin molecules (by electron tomography; see below), we used this estimate and actin’s mole fraction value to convert mole fraction values for each hair bundle protein into molecules per stereocilium (Supplementary Table 1). The distribution of protein abundance values was similar for hair-bundle and epithelium proteins (Fig. 1d), indicating that low-abundance proteins were similarly detected in both preparations.
Proteins that are selectively targeted to hair bundles may be particularly important for function. Given that bundles constitute ~2% of the epithelium (Online Methods), targeted proteins could have a bundle-to-epithelium ratio (enrichment) as large as 50-fold. Because stereocilia are not closed compartments, however, diffusible cell-body proteins will also be present in bundles, with an enrichment of ~1. Finally, the bundle preparation will also contain cell-body contaminants. The broad histogram of binned enrichment values reflected the presence of all three types of proteins (Fig. 1e).

We determined the contamination fraction, the average BUN/UTR ratio for proteins known not to be in stereocilia, by measuring relative molar abundances of proteins from mitochondria and nuclei, which are absent from stereocilia12. We chose 81 nuclear and mitochondrial proteins detected in three or more utricle experiments; proteins that were not detected in bundles were assigned an enrichment value of 0. The contamination fraction estimated from these 81 proteins was 0.20 ± 0.25 (mean ± s.d.), which suggests that the BUN preparation contains ~80% hair bundles; bundles were thus purified approximately 40-fold.

To validate the estimated contamination fraction, we used quantitative immunoblotting (Supplementary Figs. 2a,b and 3) to measure the presence in the BUN sample of five proteins known to be absent from hair bundles: ATP1A1 (ATPase, Na+/K+ transporting, alpha 1 polypeptide), found on the basolateral membranes of hair cells and supporting cells13; HSPA5 (heat shock protein 5; GRP78), an endoplasmic reticulum marker; MDH2 (malate dehydrogenase 2), a cytoplasmic marker; PTDJ (protein tyrosine phosphatase receptor J, also known as the supporting cell antigen), present on supporting cell apical surfaces14; and VIM (vimentin), an intermediate filament protein found in cell bodies of hair cells and supporting cells of the vestibular system15. Using immunocytochemistry, we confirmed that these proteins are absent from hair bundles (Supplementary Fig. 2c). As controls, we also examined actin and FSCN2 by immunoblotting (Supplementary Fig. 2a); each is concentrated in hair bundles (Supplementary Fig. 2c). In 14 measurements from 6 sets of BUN and UTR samples, we measured a contamination fraction of 0.30 ± 0.14 (mean ± s.d.), similar to that estimated by mass spectrometry. As further confirmation of the bundle enrichment of proteins detected in our analysis, we observed appropriate distribution of 16 bundle proteins by immunocytochemistry and 14 bundle proteins by immunoblotting (Supplementary Figs. 4 and 5).

We used riBAQ measurement errors, propagated in combination with the error in the calibration slope measurement, to estimate errors in protein enrichment and abundance reported in Supplementary Table 1. To determine the slope error, we plotted log_{10} riBAQ against log_{10} mole fraction (reversed axes from Fig. 1b) for all human proteins detected in the 10^{-2}–10^{-4} mole fraction range. We used a linear mixed-effects model to generate a fit through the data, constrained through the origin; the slope of the calibration curve was 1.00 ± 0.03. Using the contamination fraction and the mole fraction of each protein in hair-bundle and epithelium samples, we corrected the abundance of each protein to reflect its actual concentration in bundles (Supplementary Table 1). Because of the substantial uncertainty in the contamination fraction, these corrected values are much more reliable for bundle-enriched proteins.

The most abundant proteins in hair bundles included ACTG1 (gamma actin, representing all actins), CKB (brain creatine kinase), OCM (parvalbumin CPV3), TUBA5 and TUBB4B (representing all alpha and beta tubulins) and FSCN2 (fascin 2) (Fig. 1f). Glycolytic enzymes were abundant, as were HSPA8 and HSP90AA1 (representing the 70 and 90 kDa heat shock protein families). Only seven proteins (0.6%) accounted for 50% of the total proteome molar abundance in bundles (Fig. 1f). We highlighted proteins with high bundle or epithelium enrichment, plotting mole fraction for all proteins detected in both samples against approximate bundle purity (Fig. 1g). Some proteins were only identified in bundles—for example, PDZD7 (PDZ domain containing 7)—presumably because their epithelium concentration is below the limit of detection. We may have underestimated the hair-bundle riBAQ values for high-molecular-mass proteins, however. Because we carried out strong cation exchange (SCX) purification only for BUN gel slices that were adjacent to sample wells (Online Methods), recovery of proteins from these slices may have been less efficient than from the corresponding UTR slices. Consequently, our estimates of concentrations and enrichment values for CDH23 (cadherin 23), GPR98 (VGLR1; G-protein coupled receptor 98) and USH2A (usherin), all of which are especially large, may be too low.
We used the contamination fraction to determine which proteins were reliably present in hair bundles. Of the 1,095 proteins detected in two or more experiments each of bundles and epithelium, 336 had a bundle-to-epithelium enrichment higher than the contamination fraction with a P value <0.05, adjusted for the false-discovery rate (FDR), the rate of incorrect assignments among enriched proteins. Many

| Identifier or protein group | Description | Protein symbol | Group members | Paralog in BUN? | BUN/UTR ratio | Corrected molecules per SC |
|----------------------------|-------------|---------------|---------------|---------------|---------------|---------------------------|
| ACT                        | Actin gamma 1 | ACTG1         | ACTG1, ACTA1, ACTA2, ACTB, ACTB1L2, ACTC1, ACTG2 | Yes | 7 | 400,000 |
| Fascin 2                   | Fascin 1    | FSCN1         | FSCN1         | Yes | 40 | 40,000 |
| Radixin                    | Radixin     | RDX           | RDX, EZR, MSN | Yes | 9 | 6,800 |
| Plastin 1                  | Plastin 1   | PLS1          | PLS1          | Yes | 12 | 5,500 |
| Xin actin-binding repeat-containing protein 2 | Xin actin-binding repeat-containing protein 2 | XIRP2 | XIRP2 | Yes | 13 | 4,600 |
| Chloride intracellular channel 5 | Chloride intracellular channel 5 | CLIC5 | CLIC5 | Yes | 6 | 2,400 |
| Solute carrier family 9 member 3 regulator 2 | Solute carrier family 9 member 3 regulator 2 | SLC9A3R2 | SLC9A3R2 | No | 17 | 2,000 |
| Spectrin, alpha, non-erythrocytic 1 | Spectrin, alpha, non-erythrocytic 1 | SPTAN1 | SPTAN1 | No | 0.5 | 1,400 |
| Fasclin 1                  | Fasclin 1   | FSCN1         | FSCN1         | Yes | 4 | 1,300 |
| Actin related protein 1 homolog A, centrinactin alpha | Actin related protein 1 homolog A, centrinactin alpha | ACTR1A | ACTR1A | No | 1.7 | 1,300 |
| Destrin (actin depolymerizing factor) | Destrin (actin depolymerizing factor) | DSTN | DSTN | No | 1.3 | 1,300 |
| Spectrin, beta, non-erythrocytic 1 | Spectrin, beta, non-erythrocytic 1 | SPTBN1 | SPTBN1 | No | 0.4 | 1,300 |
| Actin, alpha 4 | Actin, alpha 4 | ACTN4 | ACTN4 | Yes | 1.4 | 1,100 |
| TWF2 | TWF2 | TWF2 | TWF2, TWF2/WDR2 | No | 4 | 950 |
| Tropomyosin 1 | Tropomyosin 1 | TMT1 | TMT1, TMT3 | Yes | 4 | 300 |
| Actin, alpha 1 | Actin, alpha 1 | ACTN1 | ACTN1, ACTN2 | Yes | 0.9 | 790 |
| Capping protein, beta 2 | Capping protein, beta 2 | CAPZB2 | CAPZB2, CAPZB1 | No | 2 | 690 |
| Capping protein, alpha 2 | Capping protein, alpha 2 | CAPZA2 | CAPZA2 | Yes | 1.3 | 540 |
| Actinin related protein 2/3 complex, subunit 1A, 41 kDa | Actinin related protein 2/3 complex, subunit 1A, 41 kDa | ARPC1A | ARPC1A | Yes | 1.2 | 470 |
| Gelsolin | Gelsolin | GSN | GSN | No | 0.4 | 190 |
| Profilin 2 | Profilin 2 | PFN2 | PFN2 | No | 5 | 180 |
| Actinin related protein 2/3 complex, subunit 2, 34 kDa | Actinin related protein 2/3 complex, subunit 2, 34 kDa | ARPC2 | ARPC2 | No | 1.7 | 180 |
| EPS8-like 2 | EPS8-like 2 | EPS8L2 | EPS8L2 | No | 0.97 | 130 |
| Ras-related C3 botulinum toxin substrate 3 | Ras-related C3 botulinum toxin substrate 3 | RAC3 | RAC3, RAC1, RAC2, RHOG | No | 1.3 | 130 |
| Chloride intracellular channel 4 | Chloride intracellular channel 4 | CLIC4 | CLIC4, CLIC6 | Yes | 1.0 | 100 |
| Myosin, heavy chain 10, non-muscle | Myosin, heavy chain 10, non-muscle | MYH10 | MYH10, MYH1 | Yes | 0.2 | 95 |
| Capping protein, alpha 1 | Capping protein, alpha 1 | CAPZA1 | CAPZA1 | No | 0.3 | 250 |
| Actinin related protein 2/3 complex, subunit 4, 20 kDa | Actinin related protein 2/3 complex, subunit 4, 20 kDa | ARPC4 | ARPC4 | Yes | 1.1 | 210 |
| Gelsolin | Gelsolin | GSN | GSN | No | 0.4 | 190 |
| Profilin 2 | Profilin 2 | PFN2 | PFN2 | No | 5 | 180 |
| Actinin related protein 2/3 complex, subunit 2, 34 kDa | Actinin related protein 2/3 complex, subunit 2, 34 kDa | ARPC2 | ARPC2 | No | 1.7 | 180 |
| EPS8-like 2 | EPS8-like 2 | EPS8L2 | EPS8L2 | No | 0.97 | 130 |
| Ras-related C3 botulinum toxin substrate 3 | Ras-related C3 botulinum toxin substrate 3 | RAC3 | RAC3, RAC1, RAC2, RHOG | No | 1.3 | 130 |
| Chloride intracellular channel 4 | Chloride intracellular channel 4 | CLIC4 | CLIC4, CLIC6 | Yes | 1.0 | 100 |
| Myosin, heavy chain 10, non-muscle | Myosin, heavy chain 10, non-muscle | MYH10 | MYH10, MYH1 | Yes | 0.2 | 95 |
| Capping protein, alpha 1 | Capping protein, alpha 1 | CAPZA1 | CAPZA1 | No | 0.3 | 250 |
| Actinin related protein 2/3 complex, subunit 4, 20 kDa | Actinin related protein 2/3 complex, subunit 4, 20 kDa | ARPC4 | ARPC4 | Yes | 1.1 | 210 |

Table 1. Quantification of actin and actin-binding proteins in chick stereocilia.

Identifier or protein group column gives the Ensembl or NCBI identifier; description column gives the common name of the principal entry; the protein with most mass-spectrometric evidence; protein symbol is the official protein symbol (based on human genes); group members lists symbols for all proteins in bundles that are summed together in an entry (alphabetical order after principal entry); paralog in BUN indicates whether a paralog is present in bundles; BUN/UTR ratio indicates the bundle/epithelium enrichment; corrected molecules per SC is the estimated molecules per stereocilium, determined with ribAQ quantification and corrected for enrichment. Actin cross-linking proteins are indicated in bold; the predominant actin-to-membrane connector family is indicated in underlined bold.
actin-associated proteins were present at 100 or more copies per stereocilium (Fig. 2a and Table 1). Moreover, proteins known to be in 1:1 stoichiometric complexes were present at the expected relative abundances (Fig. 2a), which independently corroborated our quantification.

‘Deafness proteins’ are enriched in hair bundles

We ranked proteins by bundle enrichment and labeled those proteins encoded by deafness- or vestibular-dysfunction-associated genes (Fig. 2b). A list of 7,112 Online Mendelian Inheritance in Man (OMIM) terms and their mapping to human genes and MGI marker accession IDs (downloaded October 2012) was used to identify deafness-associated proteins in the list of proteins identified from BUN and UTR samples, including redundant proteins present in protein groups (Supplementary Table 1). We used two terms to search the OMIM data: “deafness” and “Usher syndrome.”

Most deafness-associated proteins detected were enriched in hair bundles; 4% of the 277 proteins enriched >2-fold were associated with deafness in the OMIM database18, compared to only 0.7% of 2,667 proteins enriched <2-fold (P < 10⁻⁴, Fisher’s exact test). The OMIM database has 163 entries annotated with deafness or Usher syndrome, corresponding to 0.7% of the ~23,500 genes in the human Ensembl database. Proteins enriched >2-fold were also significantly (P < 10⁻²) associated with mouse deafness entries in the Mouse Genome Database (MGD)19. The list of proteins enriched in bundles over epithelium is thus a rich reference for proteins with demonstrated significance for auditory and vestibular function.

The OMIM and MGI databases do not include all genes associated with deafness or vestibular dysfunction that are expressed in stereocilia. Adding other known deafness proteins (Supplementary Table 1), including those implicated by means of targeted mutagenesis, our mass spectrometry experiments detected 22 of 27 mouse deafness-associated proteins known to be expressed in stereocilia2. These 22 proteins had an average bundle-to-epithelium enrichment of 29 ± 12 (mean ± s.e.m.), confirming that the enrichment analysis successfully identified functionally important proteins. Only DFNB31 (whirlin), CLRNL1 (clarin 1), LHPL5 (TMHS; lipoma HMGIC fusion partner-like 5), USH1G (Sans) and STRC (stereocilin) were not detected. Two of these proteins we did not expect to detect: USH1G transcripts are undetectable in embryonic day 20–21 utricles20, which generated maximum and minimum estimates of cross-linker staining inhomogeneity. We omitted out-of-plane cross-linkers in model building of our high-threshold (low estimate) rendered maps.

We measured in several independent subvolumes the total actin filament length, actin-actin spacing, number of actin-actin cross-linkers and number of actin-membrane connectors (Fig. 3d,e,l,s). A prototypical chick utricle stereocilium visualized by fluorescence and transmission electron microscopy was ~250 nm in diameter, ~5 μm in length and hence ~0.2 fl in volume. Electron tomography indicated that there were ~210 actin filaments in a stereocilium of that diameter, or ~400,000 actin monomers in filaments (~3 mM).

Table 2 Candidates for mapped but uncloned deafness-associated genes

| Human deafness locus | Chicken protein identifier | Homologous human gene | Chr | Description | Protein symbol | BUN/UTR ratio |
|----------------------|----------------------------|-----------------------|-----|-------------|----------------|---------------|
| AUNA1                | ENSGALP00000027391         | ENSG00000136167       | 13  | Plastin 2   | PLS2           | 21            |
| AUNA1                | ENSGALP00000027417         | ENSG00000102547       | 13  | Calcium binding protein 39-like | CAB39L | Bundle only |
| DFNA16               | ENSGALP0000015632         | ENSG00000071909       | 2   | Myosin IIIb | MYO3B          | 3             |
| DFNA18               | ENSGALP0000004164         | ENSG00000120756       | 3   | Plastin 1   | PLS1           | 4             |
| DFNA32               | ENSGALP00000039755         | ENSG00000177106       | 11  | Epidermal growth factor receptor kinase substrate 8-like protein 2 | EPS8L2 | 1 |
| DFNB5                | ENSGALP00000022269         | ENSG00000128050       | 4   | Multifunctional protein ADE2 | PAICS | 1 |
| DFNB5                | ENSGALP00000022326         | ENSG00000109265       | 4   | Uncharacterized protein KIAA1211 | KIAA1211 | 3 |
| DFNB5                | ENSGALP0000013870         | ENSG00000108039       | 10  | Xaa-Pro aminopeptidase 1 | XPNPEP1 | 3 |
| DFNB5                | ENSGALP000000415053       | ENSG00000187164       | 10  | Shootin-1 | SHOOTIN1 | 5 |
| DFNB5                | ENSGALP00000023253         | ENSG00000108682       | 10  | PDZ domain containing 7 | PDZD7 | Bundle only |
| DFNB85               | ENSGALP0000007825         | ENSG00000175662       | 17  | TOM1-like protein 2 | TOM1L2 | 1 |
| DFNX5                | ENSGALP0000009526         | ENSG00000102024       | X   | Plastin 3  | PLS3           | 3             |
| DFNX5                | ENSGALP0000031748         | ENSG00000165704       | X   | Hypoxanthine-guanine phosphoribosyl-transferase | HPRT | 4 |

The human deafness locus column gives the unmapped human deafness locus identifier; chick protein identifier is the chicken Ensembl identifier for the gene encoding the protein mapping to a human deafness-associated gene; homologous human gene is the Ensembl identifier for the human gene to which a given chick protein maps; Chr is the human chromosome; description gives the descriptive name for the protein; BUN/UTR ratio is the bundle/utricle enrichment in chick.

Table 2 Candidates for mapped but uncloned deafness-associated genes

Actin cytoskeleton structure from electron tomography

To further validate the mass-spectrometric data, we used electron tomography22 to count cytoskeletal structures in stereocilia from chick utricles (Fig. 3). We generated tomograms from stereocilia oriented longitudinally (Fig. 3a–c), transversely (Fig. 3f–h and Supplementary Fig. 6) and obliquely (Fig. 3i–k) to the plane of section. Each data set has distinct advantages for quantification. The longitudinal view allows us to follow individual actin filaments for long distances, but out-of-plane cross-linkers are less reliably detected owing to limited tilt-related data anisotropy. Transverse views allow ready measurement of actin-actin distances in all orientations, but the number of actin-actin links that can be detected is relatively small because of the ultrathin sections. Oblique views allow more reliable detection of out-of-plane cross-linkers, but these views are complicated by the section plane jumping from one actin filament to another. To interpret density maps obtained by electron tomography, we used two density thresholds to build simple ball-and-stick models. The use of two thresholds, which generated maximum and minimum estimates of cross-linker numbers, addressed difficulties in objectively thresholding density maps, which was complicated by reconstruction noise and possible staining inhomogeneity. We omitted out-of-plane cross-linkers in model building of our high-threshold (low estimate) rendered maps.

We measured in several independent subvolumes the total actin filament length, actin-actin spacing, number of actin-actin cross-linkers and number of actin-membrane connectors (Fig. 3d,e,l,s). A prototypical chick utricle stereocilium visualized by fluorescence and transmission electron microscopy was ~250 nm in diameter, ~5 μm in length and hence ~0.2 fl in volume. Electron tomography indicated that there were ~210 actin filaments in a stereocilium of that diameter, or ~400,000 actin monomers in filaments (~3 mM).

The human deafness locus column gives the unmapped human deafness locus identifier; chick protein identifier is the chicken Ensembl identifier for the gene encoding the protein mapping to a human deafness-associated gene; homologous human gene is the Ensembl identifier for the human gene to which a given chick protein maps; Chr is the human chromosome; description gives the descriptive name for the protein; BUN/UTR ratio is the bundle/utricle enrichment in chick.
Using the low threshold (upper limit) and averaging the data from six subvolumes per tomogram, we estimated 62,000 ± 1,000 (longitudinal-orientation tomogram) and 91,000 ± 2,000 (oblique-orientation tomogram) cross-linkers per prototypical stereocilium (mean ± s.e.m.). The conservative lower-limit estimates with a high-density threshold, which also do not consider out-of-plane cross-linkers, were 30,000 ± 1,000 and 36,000 ± 2,000. Assuming three cross-linkers for every 36 nm of actin filament, the theoretical maximum is ~87,000.

We also counted actin-to-membrane connectors with electron tomography. The prototypical stereocilium has ~52 actin filaments adjacent to the plasma membrane; as a binding site should appear every 36 nm along each actin filament, each stereocilium could contain as many as 7,200 actin-membrane connectors. We counted 7,300 ± 1,100 connectors per stereocilium with the low-density threshold and 5,800 ± 900 using the high threshold (Fig. 3g–i).

Stereociliary protein network

Focusing on the actin cytoskeleton, we identified potential interactions between hair-bundle proteins and depicted them using graph theory with spring-electrical modeling, generating a mathematically defined, undirected graph that illustrates these relationships. We chose most of the major cytoskeletal proteins that were significantly enriched in hair bundles (Table 1) to seed the network. Searching the STRING (http://string-db.org/) and BioGRID (http://thebiogrid.org/) databases identified some interactions for these proteins; however, not all known interactions are in these or other protein-protein interaction databases. We therefore manually curated our protein interaction list (Supplementary Table 2) by searching PubMed for each protein in the network, allowing us to both validate interactions and identify additional ones. All interactions identified are given in Supplementary Table 3.

The network (Fig. 4 and Supplementary Fig. 7) highlights important hair-bundle proteins, as well as signaling molecules and ions. In most molecular networks, most nodes have only a few links but others—highly connected hubs—have many, which hold the sparsely linked nodes together. For the 69 nodes with at least two interactions, the protein-interaction distribution data were fit well by a power law, with $P(k) \propto k^{-1.3}$, $R = 0.79$ and $P < 10^{-3}$. The average clustering coefficient $C_i = 2n_i/k_i(k_i - 1)$, where $n_i$ is the number of...
links connecting the $k_i$ neighbors of node $i$ to each other, represents how nodes interconnect. The $C_i$ value of 0.24 measured for our network indicates strong clustering. Nodes with the largest numbers of links were actin (33 interactions), PtdIns(4,5)P$_2$ (20), SLC9A3R2 (12), CALM (9), RDX (8) and Ca$^{2+}$ (8).

**RDX and SLC9A3R2**

RDX and SLC9A3R2, identified as hubs in the hair-bundle protein network, were each detected in chick utricle bundles by immunoblotting (Fig. 5a and Supplementary Fig. 8). We also found that SLC9A3R2 expressed in a human embryo kidney cell line (HEK293T) immunoprecipitated with RDX; SLC9A3R2 binds to activated ERM proteins, and binding was indeed stronger to RDX with a threonine-to-aspartate mutation that mimics the activating phosphorylation (Fig. 5b and Supplementary Fig. 8). In bullfrog stereocilia, RDX is activated by PtdIns(4,5)P$_2$ and phosphorylation is found in a narrow band above basal tapers, at the site of the ankle links.

**Figure 4** Interaction network for hair-bundle proteins. Symbols (nodes) represent bundle proteins or second messengers; only nodes with two or more interactions are plotted, with the exception of OCM and CALB2. Underlined labels indicate deafness-associated proteins. Node colors indicate functional classification (same key as in Fig. 2b); node symbol size represents protein abundance in bundles. Ca$^{2+}$, PtdIns(4,5)P$_2$ and cyclic AMP are indicated by diamond node symbols. Solid links represent interactions validated with literature citations; Supplementary Table 3 lists all interactions and evidence. Dotted links correspond to interactions involving paralogs of bundle proteins; dashed links represent hypothetical interactions (for example, SLC9A3R2 interactions from Table 3). The layout of the plot is controlled by the density of links between nearby nodes. The distribution of nodes and links in the plot is fit well by a power law, which indicates that the plot contains a few highly connected nodes (hubs) and many other less-connected nodes. Supplementary Figure 7 reproduces this figure with each link hyperlinked to a PubMed reference supporting the interaction.
Similarly, SLC9A3R2 and total RDX were concentrated in the bottom half of each stereocilium (Fig. 5c, d), and phosphorylated total ERM (pERM) was only found above stereociliary tapers (Fig. 5c). Likewise, pERM was concentrated in the upper half of supporting-cell microvilli (Fig. 5c, inset). Although MYO7A also appears in a band above basal tapers, its distribution was restricted to those of RDX and SLC9A3R2 (Fig. 5c). Finally, we detected in hair bundles the Rho-family GTPases RHOA, RAC3 and CDC42, which control the actin cytoskeleton (Supplementary Table 1). Actin polymerization in stereocilia, which is dynamic at least through late development, is controlled by the myosin III and V families5. The ARP2/3 complex likely is inactive in stereocilia. Because we did not detect any activating WAVE/WAVE family members, however, the ARP2/3 complex likely is inactive in stereocilia. We did not detect the MYO15A binding partner DFNB31, we found previously31. Mass spectrometry thus confirmed that each stereocilium has ~48,000 cross-linkers, in good agreement with the tomography estimates (33,000–77,000).

The ERM family, which cross-links membrane-protein complexes to actin at ~36-nm intervals, likely contributes most of the actin–membrane connecters. By mass spectrometry, EKR (ezrin), RDX and MSN (moesin) together totaled 6,800 molecules per stereocilium (Table 1), with RDX accounting for the majority; this value is within the range estimated by tomography (5,800–7,300).

Remaining actin-to-membrane connectors may be members of the myosin superfamily. Mass-spectrometric quantification gave estimates of myosin abundance that corresponded well with independent measurements. By quantitative immunoblotting, bullfrog bundles have >700 molecules per stereocilium of MYO6, >400 of MYO7A and 100 of MYO1C29; mass spectrometry estimated 6,600 chick MYO6, 250 MYO7A and ~50 each of MYO1C and the closely related MYO1H.

Actin polymerization in stereocilia, which is dynamic at least through late development, is controlled by the myosin III and myosin XV families. We detected 430 myosin III molecules per stereocilium, nearly all of which was MYO3B. Notably, the concentration of MYO3B was very close to that of its binding partner ESPN and that of PFN2, the profilin paralog that binds to ESPN.

We detected 50 molecules per stereocilium of MYO15A. Although we did not detect the MYO15A binding partner DFNB31, we found 130 molecules of EPS8L2 (EPS8-like 2); because its paralog EPS8 (epidermal growth factor receptor pathway substrate 8) binds the MYO15A-DFNB31 complex, EPS8L2 might partner with MYO15A in the vestibular system. Altogether we counted ~7,500 myosin molecules per stereocilium, which could account for remaining actin-to-membrane connectors.

Several other proteins also control polymerization of actin networks. We detected five of seven subunits of the ARP2/3 complex, which mediates polymerization of branched actin networks; at 3 µM (~340 molecules per stereocilium), the ARP2/3 complex is present at concentration similar to that in human neutrophils. Because we did not detect any activating WASP/WAVE family members, however, the ARP2/3 complex likely is inactive in stereocilia. We did not detect any Ena/VASP family members, and although we detected one formin (DIAPH2) in one bundle experiment, its enrichment value suggested that it was a contaminant. Together these results suggest that control of actin-filament polymerization in late development involves only myosin-mediated mechanisms. Although actin may not treadmill from tip to taper, the stability of stereociliary actin filaments suggests that their barbed ends, data. We compared the tomography estimates to abundances of known actin-actin cross-linkers and actin-membrane connectors measured by mass spectrometry using the rBAQ method (Table 1). We detected three classes of cross-linkers: fascins (40,000 molecules per stereocilium of FSCN2, 1,300 FSCN1), plastins (5,500 PSL1, also known as fimbrin; 460 PSL2; 400 PSL3) and epsins (710 ESPN and 90 ESPNL), similar to what we found previously31. Mass spectrometry thus estimates that each stereocilium has ~48,000 cross-linkers, in good agreement with the tomography estimates (33,000–77,000).
at stereociliary tips, are capped. If one barbed-end capper is present per filament and if filaments run the length of the stereocilium, there should be $\sim 210$ cappers per stereocilium. We detected two main cappers: $\sim 700$ CAPZ (capping protein) heterodimers per stereocilium and 950 TWF2 (twinnilin-2) molecules. The excess of cappers over actin filaments suggests that they compete for filament ends, which could occur differentially in longer and shorter stereocilia.

Pointed ends of stereociliary actin filaments progressively terminate to form stereociliary tapers, suggesting a systematic capping or depolymerization there. We did not detect tropomodulin, the best-characterized pointed-end capper, nor did we detect taperin, a candidate pointed-end capper. The pointed ends of stereociliary actin filaments either terminate on the taper membrane or gather into the central rootlet material; if the former, MYO6 could anchor the pointed ends to the membrane protein PTPRQ (protein tyrosine phosphatase, receptor type, Q), present at 1,500 copies per stereocilium, or if the latter, the rootlet component TRIOBP (TRIO and F-actin binding protein) (detected in one experiment only) could cross-link filament ends to the rootlet.

At 4,600 molecules per stereocilium, XIRP2 (Xin-related protein 2) is the most abundant protein in hair bundles without an obvious role. Although most species’ XIRP2 proteins contain >30 actin-binding Xin repeats, chick and rat bundle XIRP2 do not contain these domains (Supplementary Fig. 4b). However, XIRP2 is a paralog of the actin-binding protein LIMA1, also known as EPLIN, suggesting that XIRP2 may nevertheless bind actin.

Network analysis highlighted the role of several other cytoskeletal proteins. The largest hair-bundle hub was actin, with 33 interactions; PtdIns(4,5)P$_2$ had 20 interactions, which was expected given its prominence in stereocilia and actions in actin polymerization and membrane targeting. As is clear from co-clustering of the two hubs in the network map (Fig. 4), many bundle proteins bind both actin and PtdIns(4,5)P$_2$; moreover, several bundle proteins not known to interact with actin do bind to PtdIns(4,5)P$_2$, suggesting that the phospholipid may concentrate some membrane-associated proteins in stereocilia. Also prominent were CALM and Ca$^{2+}$; with respectively 9 and 8 interactions, they may concentrate some membrane-associated proteins in stereocilia. The estimated CALM concentration (~600 µM) is nearly identical to the 70 µM estimated for bullfrog hair bundles by quantitative immunoblotting. The membrane area of the prototypical stereocilium is ~700 µm$^2$; if the density of the Ca$^{2+}$ pump in chick bundles is the same as in bullfrog (2,000 molecules µm$^{-2}$; ref. 50), each stereocilium would have 8,000 Ca$^{2+}$ pumps. Mass spectrometry estimated significantly fewer, ~1,700 (mostly ATP2B2), perhaps because transmembrane peptides are not well detected in LC-MS/MS experiments.

**Conclusions**

Many proteins enriched in hair bundles are encoded by deafness-associated genes, suggesting that other bundle-enriched proteins may be linked to deafness in the future. Mass spectrometry allows us to identify functionally important bundle proteins that have not yet been identified by genetics, such as proteins that carry out multiple functions in an organism and could have an embryonic lethal phenotype. Indeed, genetic screens for deafness likely miss ubiquitously expressed proteins with developmental roles; by contrast, the mass spectrometric approach can in principle identify any protein that contributes to maintenance and function of the hair bundle.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Accession codes.** Mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with data set identifier PXD000104.

**Note:** Supplementary information is available in the online version of the paper.

**ACKNOWLEDGMENTS**

Mass spectrometry was carried out by the W.M. Keck Biomedical Mass Spectrometry Laboratory and The University of Virginia Biomedical Research Facility. We thank K. McDonald, R. Zalpuri and G. Min of the University of California Berkeley Electron Microscopy Laboratory for assistance with the production of images.

**Other hair bundle proteins**

Mechanotransduction molecules are rare in hair bundles; there is only one tip link and two active transduction channels per stereocilium. Nevertheless, we detected ~20 molecules per stereocilium each of the tip link cadherins, CDH23 and PCDH15 (protocadherin 15), as well as 60 USH1C (harmin) molecules, which cluster to anchor the upper end of a tip link. Although the tip link of a single stereocilium should only contain two each of the cadherins, they are also present in the lateral links of developing bundles and in kinociliary links (Supplementary Fig. 4). Our inability to detect other transduction proteins—such as the elusive transduction channel—likely arises from the limited dynamic range of mass spectrometry, rather than lack of sensitivity. In each mass spectrometry run, we matched ~10$^4$ spectra to chick peptides; because only 1–10 molecules of the transduction channel may be present for every 10$^6$ bundle molecules, substantial enrichment may be required to detect it above the background of actin and other cytoskeletal proteins.

Axonemal kinocilia are present in the hair-bundle preparation; besides tubulin, we detected the axonemal dyneins DNAH5 and DNAH9, the radial spoke head molecules RSHL1, RSPH6A and RSPH9, the intraflagellar transport molecule IFT172 and the axonemal small GTPase ARL13B.

Diffusible Ca$^{2+}$ buffers were prominent; we estimated 63,000 OCM, 8,000 CALB2 (calretinin) and 7,300 CALM (calmodulin) per stereocilium, together corresponding to a total of ~2 mM Ca$^{2+}$ binding sites. The estimated CALM concentration (~60 µM) is nearly identical to the 70 µM estimated for bullfrog hair bundles by quantitative immunoblotting. The membrane area of the prototypical stereocilium is ~700 µm$^2$; if the density of the Ca$^{2+}$ pump in chick bundles is the same as in bullfrog (2,000 molecules µm$^{-2}$; ref. 50), each stereocilium would have 8,000 Ca$^{2+}$ pumps. Mass spectrometry estimated significantly fewer, ~1,700 (mostly ATP2B2), perhaps because transmembrane peptides are not well detected in LC-MS/MS experiments.
high-pressure freezing and imaging. D. Jorgens provided mentoring in high-pressure freezing. We would like to thank A. Cheng, B. Carragher and C. Potter for help with electron microscopy data collection at the National Resource for Automated Molecular Microscopy, supported by US National Institutes of Health (NIH) National Center for Research Resources grant RR017573. For technical assistance, we acknowledge A. Snyder of the Advanced Light Microscopy Core at The Jungers Center (Oregon Health & Science University), supported by shared instrumentation grants S10 RR023432 and S10 RR025440 from the National Center for Research Resources (NIH). Work described here was supported by NIH grants K59/R00 DC009412 (J.B.S.), F32 DC012455 (J.F.K.), R01 DC026368 (P.G.B.-G.), R01 DC011034 (P.G.-B.G.), P30 DC05983 (P.G.B.-G.), R01 EY007755 (L.L.D.), P30 EY10572 (L.L.D.) and P01 GM051487 (M.A.).

AUTHOR CONTRIBUTIONS
J.-B.S. and P.G.B.-G. designed the overall approach and analysis. J.F.K. carried out immunoblotting and immunocytochemistry experiments of Figures 2 and 5, as well as Supplementary Figures 2 and 4; she also carried out quantitative immunoblots of Figure 1. J.-B.S. and J.M.P. prepared hair-bundle and epithelium samples for mass spectrometric analysis. A.H., Z.M., A.N.T. and M.A. carried out electron tomography and analyzed tomography data. N.E.S. and E.D.J. carried out immunoprecipitation experiments of Figure 5. D.C. carried out the statistical analyses. P.G.B.-G. carried out mass spectrometry data analysis, with assistance from P.A.W. and L.L.D.

The manuscript was written by P.G.B.-G.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Published online at http://www.nature.com/doffinder/10.1038/nree.3312. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.
ONLINE METHODS

Animals. Animal experiments reported here were approved by the Oregon Health & Science University Institutional Animal Care and Use Committee; the approval number was A684. All experiments began by killing the animal, which was done using methods approved by American Veterinary Medical Association Panel on Euthanasia.

Preparation of samples for mass spectrometry. Utricle hair bundles were purified from embryonic day 20–21 chicks using the twist-off method. We estimated the fraction of epithelium protein accounted for by bundles using two independent methods. In the first, we divided the amount of bundle protein per utricle (16 ng; ref. 4), measured with a fluorescence protein assay, by the estimated recovery (~40%); this value was then divided by the protein per utricle (estimated here at 2.4 ± 0.2 µg). This approach suggested bundle protein was 1.7% of the utricle's total protein. In the second method, we estimated the areas taken up by bundles and cell bodies in images of plastic sections of fixed, osmium-stained utricles examined by light microscopy. Using Fiji (http://fiji.sc/) to measure regions of interest, this method estimated that bundles make up 1.8 ± 0.6% of the total protein in the utricle. Given the uncertainties in each, the methods suggested that bundles make up ~2% of the total protein in the utricle.

Separation of proteins by a short SDS-PAGE run before reduction, alkylation and trypsin digestion substantially increased sensitivity and reproducibility of detection in comparison to other methods, in part because we could remove interfering polymers from the agarose used for bundle isolation. NuPAGE LDS sample buffer (Invitrogen) with 50 mM dithiothreitol was added to a combined final volume of 80 µl per 100 utricles' worth of bundles; samples were heated to 70 °C for 15 min. Epithelial proteins were also solubilized with NuPAGE LDS sample buffer. Proteins were separated by running –1 cm into a NuPAGE 4–12% Bis-Tris gel (1.5 mm × 10 well; one or two lanes per bundle sample); gels were rinsed with water, then stained with Imperial Protein Stain (Thermo Scientific). The 1 cm of gel with separated proteins was manually sliced into six pieces.

Gel pieces were transferred to siliconized tubes and washed and destained in 200 µl 50% methanol overnight. The gel pieces were dehydrated in acetonitrile, rehydrated in 30 µl of 10 mM dithiothreitol in 0.1 M ammonium bicarbonate and reduced at room temperature (20–25 °C) for 0.5 h. The DTT solution was removed and the sample alkylated at room temperature (20–25 °C) for 0.5 h with 30 µl of 50 mM iodoacetamide in 0.1 M ammonium bicarbonate. The reagent was removed and the gel pieces dehydrated in 100 µl acetonitrile. The acetonitrile was removed and the gel pieces rehydrated in 100 µl of 0.1 M ammonium bicarbonate. The pieces were dehydrated in 100 µl acetonitrile, the acetonitrile removed and the pieces completely dried by vacuum centrifugation. The gel pieces were rehydrated in 20 µg/ml trypsin (Sigma-Aldrich T6567 proteomics grade, from porcine pancreas, dimethylated) in 50 mM ammonium bicarbonate on ice for 10 min. Any excess enzyme solution was removed and 20 µl of 50 mM ammonium bicarbonate added. The sample was digested overnight at 37 °C and the peptides formed were extracted from the polyacrylamide in two 30-µl aliquots of 50% acetonitrile, 5% formic acid. These extracts were combined and evaporated to 15 µl for MS analysis. For the gel slice immediately adjacent to the agarose in the sample well, peptides were first purified away from interfering polymers on a SCX CapTrap from Bruker-Michrom (TR1/25109/35; size 0.5 × 2 mm, bed volume 0.5 µl). The CapTrap was washed with 50 µl of 1% acetic acid (void volume collected) and then eluted with 25 µl 1 M ammonium acetate into a separate microcentrifuge tube. The eluate was vacuum dried, then the sample reconstituted with 15 µl of 3% acetic acid. A single experiment's worth of hair bundles or epithelium was analyzed by six LC-MS/MS runs, corresponding to the six gel pieces.

Mass spectrometry data acquisition and analysis. The LC-MS/MS system consisted of a Thermo Electron Orbitrap Velos ETD mass spectrometer system; the exceptional mass accuracy of the Orbitrap instrument allows for high resolution of peptide peak m/z (mass-to-charge ratio), which leads to high numbers of confident protein assignments. Peptides were introduced into the mass spectrometer with a Protona nanospray ion source, which was interfaced to a reversed-phase capillary column of 8 cm length x 75 µm internal diameter, self-packed with Phenomenex C18 Jupiter of 10 µm particle size. An extract aliquot (7.5 µl) was injected and peptides eluted from the column by an acetonitrile/0.1 M acetic acid gradient at a flow rate of 0.5 µl/min over 1.2 h. The nanospray ion source was operated at 2.5 kV. The digest was analyzed using the data-dependent capability of the instrument, acquiring—in sequential scans—a single full-scan mass spectrum in the Orbitrap detector at 60,000 resolution to determine peptide molecular weights, and 20 product-ion spectra in the ion trap to determine amino acid sequence.

MaxQuant version 1.2.8.5 software was used for protein identification and quantification. The default contaminants file associated with the MaxQuant download was edited to remove entries known to be present in hair bundles (for example, actin) and to add additional impurities that entered the bundle-purification workflow (keratins, hemoglobins, egg white components). Nevertheless, alpha and beta hemoglobins, which appear in the preparation owing to contamination from red blood cells, are expressed in chick utricles, suggesting that they should not be fully dismissed as components of hair cells. Mass spectrometry data were searched against Ensembl version 66 (released February 2012) using Andromeda; the Ensembl FASTA file was edited by replacing several sequences with longer or full-length sequences, including actin gamma 1 (NP_001007825.1), actin beta (NP_990849.1), fascin 1 (NP_001171603), fascin 2 (NP_001171209), ATP synthase beta (NP_001026562.1), peptidylprolyl isomerase A (NP_001159798.1), calbindin 2 (NP_990647), PDZD7 (XP_003641537.1), espin (XP_417532.3) and CACNA2D2 (XP_427707.3). Protein identifications were reported with an FDR of 1%. Proteins identified with a single unique peptide are flagged in Supplementary Table 1. If a set of peptides for a protein was identical to or completely contained within that of another protein, MaxQuant groups those proteins together (redundant groups); the entry with the largest number of peptides was used to identify the redundant group. Redundant groups that shared more than 20% of their identified peptides were further grouped in our analysis (‘shared-peptide groups’); the entry with the greatest intensity associated with it was used to identify the shared-peptide group. All mass spectrometry proteomics data, including raw data, MaxQuant output files and modified Ensembl FASTA database, have been deposited to the ProteomeXchange Consortium (http://protemecentral.proteomeexchange.org/) via the PRIDE partner repository with the data set identifier PXD000104.

Annotation of the chicken genome is incomplete and occasionally wrong. For all proteins identified in the BUN preparation, we manually examined annotations of the chicken Ensembl entry and Ensembl-identified orthologs of other species, particularly mouse and human, to determine an appropriate description and symbol. Whenever possible, we chose the human ortholog’s gene name, as determined by the Human Genome Organization, for a protein’s symbol.

Gene Ontology annotation of the chicken genome is poor and, in many cases, was not useful for annotation of bundle proteins. Accordingly, we chose a simple, consistent set of ontology annotations to apply to all bundle proteins (see Fig. 2b). All proteins in the BUN preparation had one (and only one) of these ontologies assigned to it (Supplementary Table 1).

Paralog identification. Using data downloaded during October 2012 from the Chicken Ensembl database (http://www.ensembl.org/Gallus_gallus/Info/Index/) BioMart tool, we also identified all paralogs in the chicken genome for each protein entry or group, calculating the average sequence identity for all paralogs matching to a protein or protein group. We also determined which paralogs for a protein or protein group were identified in the combined BUN and UTR data sets, as well as which were present in the group of all proteins that were significantly enriched at P < 0.05 over the contamination fraction. These data are reported in Supplementary Table 1.

Intensity-based mass-spectrometric quantification. To quantify proteins, we used a label-free method that relays each detected peptide’s ion-current signal in the mass spectrometer. As peptides elute from the liquid chromatograph, undergo ionization and are delivered to the mass spectrometer, the Orbitrap instrument we used measures their intensities, as well as measuring their mass-to-charge ratio with high resolution. Intensity depends both on the charge and amount of peptide delivered to the detector, although the efficiency of delivery varies widely from peptide to peptide because of variable recovery following liquid chromatography and differing degrees of ionization. Thus the relationship between intensity measured in the mass spectrometer and the amount of a peptide injected into the liquid chromatograph also varies, which limits quantification accuracy when standards for each protein are lacking. For example, hydrophobic peptides derived from transmembrane segments of integral membrane proteins are particularly poorly recovered, leading to reduced detection of this class of proteins.
proteins. Moreover, mass spectrometers like the one we used are tuned to opti-
mally detect peptides in a relatively narrow mass range, typically 6–30 amino
acids, so proteins with an overabundance of short or long tryptic peptides are
quantified less accurately than those with an average peptide. Nevertheless, mea-
surement of protein abundance uses the sum of many peptide measurements, so
with larger proteins—which generate many peptides—inter-peptide variability
is averaged out and quantification accuracy is improved.

In the iBAQ algorithm, the intensities of the precursor peptides that map to
each protein are summed and divided by the number of theoretically observable
peptides, which is considered to be all tryptic peptides between 6 and 30 amino
acids in length. This operation converts a measure that is proportional to mass
(intensity) into one that is proportional to molar amount (iBAQ). The release of
the MaxQuant we used (version 1.2.2.5) reports for each identified protein both
its summed intensity and its iBAQ value.

To determine absolute amounts of each protein in stereocilia, we generated a
normalized measure of molar abundance, relative iBAQ (riBAQ). We first
removed from the analysis all contaminant proteins that entered our sample-
preparation workflow, which include keratins (from human skin), hemoglobin
(from blood), egg white proteins (for example, ovalbumin) and trypsin. We then
divided each remaining protein’s iBAQ value by the sum of all non-contaminant
iBAQ values, generating a riBAQ value for each protein:

\[
 riBAQ = \frac{\sum_{i=1}^{n} iBAQ_i}{\sum_{i=1}^{n} iBAQ_i}
\]

For riBAQ validation experiments, we spiked one-fifth of a vial of UPS2
standard proteins (Universal Proteomics Standard 2; Sigma-Aldrich) into 25 µg
(500 pmol total) of the E. coli extract. Each experiment thus included 48 human
proteins, eight each at 10 pmol, 1 pmol, 100 fmol, 1 fmol, 1 fmol and 100 amol.
We carried out four independent experiments using methods identical to those
for hair bundles and utricular epithelia, including a 1-cm SDS-PAGE separa-
tion that was followed by slicing the gel into six pieces; reduction, alkylation
and trypsin proteolysis; LC-MS/MS using the Orbitrap mass spectrometer; and
MaxQuant analysis with riBAQ determination.

For quantitative immunoblot validation of riBAQ values, BUN and UTR sam-
ple were loaded into 10- or 15-well Novex NuPAGE Bis-Tris 4–12% gels, along
with dilutions of purified protein standards, and separated by SDS-PAGE. Protein
standards were CALM (bovine brain; EMD Millipore), ANXA5 (recombinant,
from chicken; ImmunoTools GmbH), MYO1C (recombinant, from rat; gift of
L. Coluccio, Boston Biomedical Research Institute) and FSCN2 (recombinant,
from human; USC Life Sciences). Five dilutions between 1 and 200 ng were used
for CALM and ANXA5 standard curves, four dilutions between 0.01 and 1 ng for
MYO1C and four dilutions between 1.25 and 10 ng for FSCN2. SDS-PAGE and
immunoblotting were carried out essentially as described previously. Band
intensities were measured using Fiji imaging software and linear regression analy-
sis was carried out to determine the amount of each protein (ng per ear) in the
bundle and epithelium samples. Three experiments were carried out for each
standard protein. To estimate the mole fraction of each protein, the estimate
of ng per ear was converted to mol per ear, then was divided by the number of
mole of total protein per ear. We estimated moles of protein per ear by dividing
the average molecular mass for all proteins in each sample, which was weighted
from chicken saline (155 mM NaCl, 5 mM KCl, 5 mM d-glucose and 10 mM 4-(2-hydroxyethyl)-1-
pyrrolineethanesulfonic acid (HEPES)) containing 0.1 mM CaCl2. The tissue
was fixed for 2 h at room temperature (20–25 °C) in 3% glutaraldehyde (Electron
Microscopy Sciences), 0.2% tannic acid, 100 mM KCl, 5 mM MgCl2 and 20 mM
3-(N-morpholino)propanesulfonic acid (MOPS) at pH 6.8, then washed three
times for 5 min each at room temperature (20–25 °C) in the same solution without
 glutaraldehyde and tannic acid.

Samples were stained with 1% OsO4 in 0.1 M sodium phosphate buffer for
1 h on ice, followed by three rinses with 0.1 M sodium phosphate buffer and
more rinses with deionized water. Subsequent staining was carried out with
2% aqueous uranyl acetate at room temperature (20–25 °C) for 1 h, followed
immediately by three rinses with deionized water. Samples were coated with a
20% glycerol cryoprotectant and frozen using a Bal-Tec HPM 010 high pressure
freeze dryer. Thereafter, freeze substitution was performed using a freeze substitution
mix containing 0.1% uranyl acetate in 100% acetone. Epon-araldite resin embed-
ding and infiltration was carried out as described. Sections were cut at 70 or
120 nm for screening or for tomography, respectively. Thin sections, nominally
120 nm in thickness, were placed on 2 x 1 mm oval-hole copper-rhodium grids
with 0.6% Formvar coating for imaging and decorated with 10 nm or 15 nm
gold fiducial markers for tomography. For greater stability upon beam exposure and
to minimize charging, a thin film of carbon was deposited on grids containing sec-
tions using a Denton Vacuum system (DV-502). Initial screening and tomography
was primarily done with JEOL1200-EX (TEM only), Philips CM200 FEG and FEI
Tecnai F20 electron microscopes, whereas analyzed tilt series were collected using
an FEI Tecnai T12. LaB6 electron microscope operated at 120 kV.

Images were recorded with a Gatan MegaScan Model 794/20 2K CCD
(JEOL1200-EX), a Gatan First Light 4K CCD (CM200), a Tietz F415 4K CCD
(F20) or a Gatan 1000 2.048 x 2.048-pixel CCD camera (T12). Tomograms for
quantification were collected using ~0.8 μm underfocus at a nominal 13,500x
magnification, corresponding to a pixel size of 0.8 nm at the specimen level. Dual-
axis tilt series were recorded at 1° intervals for angles of up to ±65°. Projections
were aligned and reconstructed into a three-dimensional volume with the soft-
ware package IMOD. Resulting maps were processed either with three iter-
ations of a bilateral filter using PRISM or with successive rounds of smoothing,
diffusion, or median filtering using the Clip program in IMOD. Segmentation
and interactive simplified model building was done using UCSF Chimera. The
simplified model showing actin filaments, cross-linkers and connectors
deliberately used cylinders smaller in diameter than their respective structures
to allow adequate display of the cross-linkers and connectors.

**Stereociliary protein network.** The protein interaction data in Supplementary
Table 2 can be represented in the form of a graph G = \{N,L\}, where the set of
nodes N correspond to bundle proteins and links L correspond to specific
protein interactions. Two vertices \(v, m\) form an edge of the graph \(v \rightarrow m\) in L
because in our data \(\{n,m\} \in L\), the graph is undirected and is drawn with line
segments rather than arrows. To visualize the interrelationships between
these interactions, they are displayed in a drawing (a graph, in the
mathematical sense). To represent the interactions aesthetically, link crossings
were minimized and spacing between nodes was kept relatively even. We used
a straight-edge drawing algorithm, spring-electrical embedding, which mini-
izes the energy of a physical model of the graph in two dimensions. Spring-
electrical embedding uses two forces. The ‘attractive force’ \(f_a = K_1/d^2\) between
adjacent nodes is proportional to the Euclidian distance between them \(d_i\); \(K_1\) is a
spring-like constant that maintains the optimal distance between nodes. The
‘electrical force’ (repulsive) is global and inversely proportional to the Euclidian
distance between nodes: \(f_e = K_2/d^3\). The layout of the graph vertices is then
determined by minimizing the energy function described by these two forces.
These modeling rules draw together nodes with similar interactions and disperse
unrelated ones, thus clustering related proteins, which may carry out similar
roles in bundles. The spring-electrical embedding algorithm is implemented in
Mathematica 8 (Wolfram Research; http://www.wolfram.com/mathematica/),
which we used to generate the network figure.

**Antibody methods.** Primary antibodies were as follows: for ANXA5, polyclonal
anti-chick ANXA5 from G. Richardson (University of Sussex); for α-tubulin,
Sigma-Aldrich DMA1; for ATTPA1, Developmental Studies Hybridoma Bank
a5; for ATP2B2, F2a (ref. 50); for ATP5A1, BD Transduction 612516; for CALM,
Millipore 05-173; for CDC42, Cell Signaling 2466 (11A11); for CDH23, C2367 goat anti-CDH23 EC15/16 junction (raised against CATRPAPDREQRq by Genemed Synthesis, Inc.); for CKB, rabbit anti-chICK from T. Wallimann (ETH Zürich); for CTNNB1, BD Transduction Labs 610154; for DSTN, anti-ADP/collin from J. Bamberg (Colorado State University); for ENO1, Santa Cruz H-300; for ESPN, anti-espin from S. Heller (Stanford University) and pan-espin from B. Kachar (NIH); for FSCN2, anti-CYTFLEFKAGKLFKFD (ref. 31); for GAPDH, Chemicon MAB374; for HSPA5 (GRP78), Abcam ab32618; for KIAA1211, rabbit anti-CVSTEPAWLALAKRKAKAWSD (ref. 31); for GAPDH, Chemicon MAB374; for HSPA5 (GRP78), Abcam ab32618; for KIAA1211, rabbit anti-CVSTEPAWLALAKRKAKAWSD

For immunoblotting, proteins were expressed for 24 h in HEK293T cells using Effectene (Qiagen) and CQRHSHSFSSHSSRKDLNGQKE). Antibodies to MYO6 were obtained from J. Kendrick-Jones (MRC Laboratory of Molecular Biology). We also used anti-SLC9A3R2 (anti-NHERF2) 2331B (from M. Donowitz; Johns Hopkins University) for immunoblotting.

For immunoprecipitation, RDX and SLC9A3R2 were cloned from chicken utricle cDNA into expression vectors with respectively Myc and HA epitope tags. Proteins were expressed for 24 h in HEK293T cells using Effectene (Qiagen) transfection. Cells were lysed using a probe sonicator with PBS containing 1% Triton X-100, 0.5% NP40 and protease inhibitors; the extract was centrifuged at 16,000g for 15 min. The supernatant was removed and 300 µl was incubated with anti-HA-agarose (20 µl 50% slurry; Sigma A2095) overnight at 4 °C. Immunoprecipitates were washed and proteins were eluted with SDS sample buffer. SDSPAGE and immunoblotting were carried out essentially as above.

For immunocytochemistry, dissected utricles were fixed for 25 min in 4% formaldehyde (Electron Microscopy Sciences) in chicken saline. Organs were rinsed in PBS, permeabilized for 10 min in 0.5% Triton X-100 in PBS and blocked for 2 h in 2% bovine serum albumin, 5% normal goat serum in PBS. Organs were incubated overnight at 4 °C with primary antibodies diluted in blocking solution (1:250 dilution for all primary antibodies except anti-SLC9A3R2 and antibodies, both at 1:500), then rinsed three times for 10 min each. Organs were then incubated for 3–4 h in blocking solution with 1:1,000 Alexa Fluor secondary antibodies and 0.4 U/ml Alexa Fluor 488 Phalloidin (Molecular Probes, Invitrogen). Organs were then rinsed three times for 20 min each and mounted on slides in Vectashield (Vector Laboratories) using one Secure-Seal spacer (eight wells, 0.12 mm deep, Invitrogen).

Images were acquired on an Olympus Fluoview FV1000 laser scanning confocal microscope system with AF10-ASW 3.0 acquisition software, using a 60× 1.42 NA Plan Apo objective with 3× zoom and 0.2-µm z-steps. Confocal images were deconvoluted with the optical transfer function optimized for that objective using an iterative algorithm of ten iterations. The histogram was adjusted for the most positive image and applied to all the other images for consistency before saving the images as 24-bit merged TIFF files. All z-stacks were processed using Fiji (http://fiji.sc/); the Reslice function was used to generate an x-z slab of the stack.

51. Vizcaíno, J.A. et al. The Proteomics Identifications database: 2010 update. Nucleic Acids Res. 38, D736–D742 (2010).
52. Barrera, N.P. & Robinson, C.V. Advances in the mass spectrometry of membrane proteins: from individual proteins to intact complexes. Annu. Rev. Biochem. 80, 247–271 (2011).
53. Sosinsky, G.E. et al. The combination of chemical fixation procedures with high pressure freezing and freeze substitution preserves highly labile tissue ultrastructure for electron tomography applications. J. Struct. Biol. 161, 359–371 (2008).
54. Kremer, J.R., Mastronarde, D.N. & McIntosh, J.R. Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. 116, 71–76 (1996).
55. Chen, H., Clyborne, W.K., Sedat, J.W. & Agard, D.A. Prisim: an integrated system for display and analysis of 3-D microscope images. Proc. SPIE 1660, 784–790 (1992).
56. Petteerson, E.F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004).
57. Dumont, R.A., Zhao, Y.D., Holt, J.R., Bahler, M. & Gillespie, P.G. Myosin-I isozymes in neonatal rodent auditory and vestibular epithelia. J. Assoc. Res. Otolaryngol. 3, 375–389 (2002).
58. Goodyear, R.J. & Richardson, G.P. A novel antigen sensitive to calcium chelation is associated with the tip links and kinocilial links of sensory hair bundles. J. Neurosci. 23, 4878–4887 (2003).