Exploring regulatory networks of miR-96 in the developing inner ear

Morag A. Lewis1,2, Annalisa Buniello1,2, Jennifer M. Hilton2, Fei Zhu2, William I. Zhang2, Stephanie Evans2, Stijn van Dongen3, Anton J. Enright3 & Karen P. Steel1,2

Mutations in the microRNA Mir96 cause deafness in mice and humans. In the diminuendo mouse, which carries a single base pair change in the seed region of miR-96, the sensory hair cells crucial for hearing fail to develop fully and retain immature characteristics, suggesting that miR-96 is important for coordinating hair cell maturation. Our previous transcriptional analyses show that many genes are misregulated in the diminuendo inner ear and we report here further misregulated genes. We have chosen three complementary approaches to explore potential networks controlled by miR-96 using these transcriptional data. Firstly, we used regulatory interactions manually curated from the literature to construct a regulatory network incorporating our transcriptional data. Secondly, we built a protein-protein interaction network using the InnateDB database. Thirdly, gene set enrichment analysis was used to identify gene sets in which the misregulated genes are enriched. We have identified several candidates for mediating some of the expression changes caused by the diminuendo mutation, including Fos, Myc, Trp53 and Nr3c1, and confirmed our prediction that Fos is downregulated in diminuendo homozygotes. Understanding the pathways regulated by miR-96 could lead to potential therapeutic targets for treating hearing loss due to perturbation of any component of the network.

The diminuendo mouse carries a single base-pair change in the seed region of the microRNA gene Mir96. Heterozygotes exhibit rapidly progressive deafness, hair cell degeneration and stereocilia disorganisation, while homozygotes have a vestibular phenotype in addition to deafness and have lost their outer hair cells and most inner hair cells by four weeks of age. Microarrays carried out on diminuendo homozygotes (Dmdo/Dmdo) and wildtype littermates at four days old (P4), before the hair cell degeneration commences, found 96 transcripts to be significantly affected and, on a broader scale, a transcriptome-wide signature of misregulation of the microRNA1. Mutations in MIR96 have also been reported to cause dominant progressive deafness in humans1,2. miR-96 is one of three microRNAs clustered together on chromosome 6; the others are miR-182 and miR-183. All three are expressed in multiple sensory organs, including the olfactory epithelium, vomeronasal organ, olfactory bulb, pineal body, tongue papilla and dorsal root ganglia1 in addition to sensory hair cells3.

miR-96 appears to be a master regulator of hair cell differentiation. Diminuendo homozygote hair cells fail to develop the potassium currents, synaptic machinery and connections typical of a mature hair cell, retaining juvenile characteristics4. However, the mechanisms by which the diminuendo mutation causes the failure of hair cell differentiation remain unclear. Several of the genes downregulated in Dmdo/Dmdo mice at P4, including Ptprq, Gfi1, Kcna10 and Slc26a5 (prestin), cause deafness when mutated7–10 and some of the unique characteristics of mice null for these genes are replicated in the diminuendo mouse, such as the short outer hair cells of prestin mutants8,9, and the delay in hair cell stereocilia differentiation seen in Ptprq functional nulls11. However, none of these genes can entirely account for the diminuendo phenotype, nor are any of them direct targets of miR-96; rather, the transcriptome-wide effects of the diminuendo mutation suggest it is regulating many genes through multiple pathways.

Regulatory network construction and analysis is a complex problem with no one easy solution. However, knowing both the starting point (the mutant miR-96) and the end points (the misregulated transcripts) means a set of hypothetical pathways can be drawn up, examined for consistency and tested in vivo. Here we present a regulatory network for miR-96 based on our existing data, our new expression data and on reports in the literature, and the results of testing some of its predictions. We have also constructed a protein-protein interaction network based on microarray data together with known and predicted targets of miR-96, and carried out gene

1Wolfson Centre for Age-Related Diseases, King’s College London, Guy’s Campus, London SE1 1UL, UK. 2Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK. 3EMBL EBI, Hinxton, Cambridge CB10 1SD. Correspondence and requests for materials should be addressed to M.L. (email: morag.lewis@kcl.ac.uk) or K.S. (email: Karen.steel@kcl.ac.uk)
set enrichment analyses on the microarray data. Our results demonstrate the differences between the methods, and also identify several genes highlighted as important in more than one analysis, which are candidates for therapeutic intervention.

**Results**

**Microarray on newborn diminuendo homozygotes shows similar broad effects to that at P4.** Because the diminuendo mutation appears to delay development, we carried out a microarray on the organ of Corti of newborn homozygote diminuendo and wildtype mice to examine the transcriptome at a younger age than our previous P4 analysis. We found fewer genes significantly affected at P0, but of the 18 that were, 17 showed the same direction of regulation in the P4 microarray (Supplementary Fig. S1, Supplementary Table S1). We then carried out a Sylamer analysis to examine the direct effects of the mutant microRNA on the transcriptome. Sylamer analyses examine all possible heptamers in the 3′UTRs of genes in the microarray sorted by expression value and detect whether any heptamers are significantly more common in up- or down-expressed genes than would be expected by chance. In this case, the heptamer complementary to the wildtype miR-96 seed was enriched in upregulated genes and the heptamer complementary to the mutant miR-96 seed was enriched in downregulated genes in the mutant sample. This is very similar to the results of the P4 microarray and shows that the mutant microRNA is having a broad effect on many genes at this earlier age (Fig. 1).

**Testing microarray and network genes.** We had previously confirmed the misregulation of 15 genes from the P4 microarray by qRTPCR, but prior to building the network, 10 more genes were tested and the misregulation was confirmed in 7 of them (Fig. 2a; Bhhk40, Ral32, Od3f4, Mom5, Dynlbr2, Meig1, Myo3a). We also tested the potential intermediate regulators Nr3c1, Foxo1, Foxo3 and Mitf, which are all confirmed targets of miR-96 in different tissues, Poud3, a known deafness gene which regulates Gfi1, and Thrb which has been shown to regulate Slc26a5 in the outer hair cells. Zic2 was identified previously as a target by miRanda prediction and preliminary networks created using Ingenuity IPA suggested Htt and Tnc as potential intermediate regulators. Of those nine potential intermediates, only expression of Zic2 and Poud3 were significantly affected when tested by qRTPCR (Fig. 2b, adjusted P < 0.05).

The misregulation of genes by mutant miR-96 can only be relevant if those genes are co-expressed with miR-96 in the hair cells, so we tested 7 of the misregulated genes by immunohistochemistry and confirmed that Ansa4, Meig1, Homer1, Kena10, Kaim, Gif2ec and Tppp were present in wildtype hair cells (Supplementary Fig. S2). Nr3c1, Foxo3, Thrb and Poud3 are known to be expressed in the hair cells and Foxo1 is expressed in the otic vesicle early in development. The remaining 4 genes, Mitf, Htt, Zic2 and Tnc, were tested by immunohistochemistry and all found to be expressed in wildtype hair cells (Supplementary Fig. S2). No difference in staining pattern or intensity was observed in Dmdo/Dmdo homozygote mutant hair cells.

**Identifying direct targets of miR-96 using an Ago2 pulldown assay.** Identifying the targets of miR-96 is critical for exploring the networks it controls. We looked for direct targets of miR-96 by immunoprecipitating Ago2 (one of the RISC complex proteins to which microRNAs bind and direct to their targets) and extracting the RNA bound to it using RNA from the olfactory bulbs and organs of Corti of six Dmdo/Dmdo homozygotes and six sex-matched wildtype littermates at P4, and analysed the RNA by microarray. We used the olfactory bulb as well as the organ of Corti because miR-96 is known to be expressed there and it offers more material than the organ of Corti.
Compiling a list of direct targets of miR-96. We included several targets that have been reported in other studies: Foxo1, Foxo3, Mitf, Adcy6, Clc5, Scarb1, Nr3c1, Scn3a, Spast, Gpc3, Insig2, Irs1, Slc1a1, Gpc1, Alk, Rad51, Rev1, Hbp1 and Reck (see Supplementary Table S3 for references). We also included Aqp5, Celsr2 and Odf2, which were found to be targets by luciferase assay in our previous study\(^1\), Hspa2, Ppop8, Sdc2, Arf2, Casc1, Gadd2 and Sfrs3, which were upregulated in diminuendo homoygote at P4 and P0, and have complementary matches to the miR-96 seed region in their 3’ UTRs; Zic2, which is a predicted target and is upregulated at P4; and

There were no genes reported as significantly misregulated in mutants (adjusted P < 0.05) after p-value adjustment for multiple testing. However, 216 genes from the organ of Corti and 23 from the olfactory bulb were downregulated in mutant samples with an unadjusted P < 0.01, suggesting these were more strongly bound by the wildtype miR-96 seed region. We tested these genes against three microRNA target predictors: miRDB (http://www.mirdb.org/cgi-bin/search.cgi)\(^7\), starBase (http://starbase.sysu.edu.cn/index.php)\(^2\) and Diana-microT (http://diana.cslab.ece.ntua.gr/microT)\(^7\). We also used our target list compiled previously using miRanda and did a text search for the presence of complementary matches to the miR-96 seed region in the 3’UTR (a less stringent prediction method) (Supplementary Table S2). From the resulting lists, 12 targets were selected based on the strength of their fluorescence in the microarray and on their bioinformatics predictions: Abhin1, Unkl, Cacna1c, Akap7, Tmem97, Gnb4, Mnx1, C11orf6, 6820408C15Rik, Sox5, Ankr27 and Zfp251. We checked for the presence of the 12 transcripts in wildtype organ of Corti cDNA by RTPCR and qRTPCR. 6820408C15Rik was not detected, Unkl, Cacna1c and Akap7 qRT-PCR probes did not work reliably, suggesting a very low transcript level, and none of the remaining 8 genes were significantly upregulated in diminuendo mutant organ of Corti (Fig. 3a, adjusted P < 0.05), as would be expected of a target of miR-96. We hoped to identify multiple targets through the Ago2

Figure 2. Testing gene expression in diminuendo homozygotes. qRTPCR was carried out on cDNA from P4 organs of Corti in wildtype (green) and diminuendo homozygote (red) littermates to test gene expression changes. Error bars are standard deviation (adjusted P < 0.05). (a) Confirmation of misregulation of gene expression from the P4 microarray (Chira1: wildtype n = 3, mean = 1.00 ± 0.00 (s.d.); homozygote n = 3, mean = 0.82 ± 0.10 (s.d.) Gif2e2: wildtype n = 5, mean = 1.01 ± 0.01 (s.d.); homozygote n = 5, mean = 1.08 ± 0.13 (s.d.) Rasd2: wildtype n = 5, mean = 1.01 ± 0.01 (s.d.); homozygote n = 5, mean = 2.06 ± 0.54 (s.d.) Odf3b: wildtype n = 5, mean = 1.02 ± 0.01 (s.d.); homozygote n = 5, mean = 2.51 ± 0.63 (s.d.) Dynhrb2: wildtype n = 5, mean = 1.01 ± 0.01 (s.d.); homozygote n = 5, mean = 3.12 ± 1.02 (s.d.) Meig1: wildtype n = 5, mean = 1.01 ± 0.01 (s.d.); homozygote n = 5, mean = 3.41 ± 0.72 (s.d.) Bhlhe40: wildtype n = 6, mean = 1.02 ± 0.01 (s.d.); homozygote n = 6, mean = 1.75 ± 0.15 (s.d.) Homer1: wildtype n = 6, mean = 0.99 ± 0.05 (s.d.); homozygote n = 6, mean = 0.85 ± 0.23 (s.d.) Morn5: wildtype n = 6, mean = 1.02 ± 0.01 (s.d.); homozygote n = 6, mean = 2.72 ± 0.99 (s.d.) Myo3a: wildtype n = 6, mean = 1.01 ± 0.02 (s.d.); homozygote n = 6, mean = 4.33 ± 1.57 (s.d.) Wilcoxon t-test: Chira1: p = 0.28; Gif2e2: p = 0.84; Rasd2: p = 0.03; Odf3b: p = 0.03; Dynhrb2: p = 0.03; Meig1: p = 0.03; Bhlhe40: p = 0.02; Homer1: p = 0.28; Morn5: p = 0.02; Myo3a: p = 0.02. (b) Testing targets and intermediate genes suggested by the literature. Foxo3: wildtype n = 5, mean = 1.09 ± 0.16 (s.d.); homozygote n = 5, mean = 1.03 ± 0.28 (s.d.) Mitf: wildtype n = 6, mean = 1.03 ± 0.04 (s.d.); homozygote n = 6, mean = 1.19 ± 0.45 (s.d.) Htt: wildtype n = 5, mean = 1.03 ± 0.02 (s.d.); homozygote n = 5, mean = 1.20 ± 0.54 (s.d.) Thrb: wildtype n = 5, mean = 1.00 ± 0.00 (s.d.); homozygote n = 5, mean = 1.22 ± 0.44 (s.d.) Foxo1: wildtype n = 5, mean = 1.11 ± 0.16 (s.d.); homozygote n = 5, mean = 1.28 ± 0.63 (s.d.) Nr3c1: wildtype n = 5, mean = 1.04 ± 0.07 (s.d.); homozygote n = 5, mean = 1.35 ± 0.80 (s.d.) Zic2: wildtype n = 5, mean = 1.02 ± 0.02 (s.d.); homozygote n = 5, mean = 1.40 ± 0.48 (s.d.) Tnc: wildtype n = 5, mean = 1.00 ± 0.00 (s.d.); homozygote n = 5, mean = 1.68 ± 0.64 (s.d.) Pou4f3: wildtype n = 5, mean = 1.01 ± 0.01 (s.d.); homozygote n = 5, mean = 1.68 ± 0.43 (s.d.). Wilcoxon t-test: Foxo3: p = 0.28; Mitf: p = 0.63; Htt: p = 0.84; Thrb: p = 0.28; Foxo1: p = 1; Nr3c1: p = 0.95; Zic2: p = 0.03; Tnc: p = 0.28; Pou4f3: p = 0.03.
Oshpl2, a predicted target which is specifically expressed in the outer hair cell stereocilia. We included all these genes in our list of direct targets of miR-96 (Supplementary Table S3).

Compiling the misregulated gene list. We compiled a list of genes and associated expression values, consisting of all the misregulated genes from the P4 and P0 microarrays with adjusted P < 0.1, and those genes with expression changes detected by qRT-PCR with adjusted P < 0.05 (Supplementary Table S4). The less stringent P value cutoff was chosen in order to maximise the genes available for network construction. We had several indicators that genes with adjusted P < 0.1 still offer biological relevance in this case. For example, Dynlnrb2 has an adjusted P = 0.0522 and was shown to be significantly upregulated by qRT-PCR (Fig. 2a, adjusted P = 0.0004). Sdc2 has adjusted P = 0.089 in the P0 microarray, but at P4 it has adjusted P = 0.009 and by qRT-PCR, also at P4, it was shown to be upregulated with adjusted P = 0.026. Pnpla8 has adjusted P = 0.075 in the P4 microarray and adjusted P = 0.036 in the P0 microarray. In the case of the P0 microarray, there were 18 genes with adjusted P < 0.05, and raising the p-value cutoff to 0.1 meant including just 6 more. For the P4 microarray, raising the threshold to p < 0.1 meant 46 genes were added to the initial 85 with p < 0.05. If the microarray gave a significant result but qRT-PCR testing did not, the microarray expression value was used where quantitative data was required. We also added genes identified by qRT-PCR in our previous studies. This misregulated gene list includes a few genes predicted to be direct targets, so there is some overlap between this list (Supplementary Table S4) and the list of direct targets (Supplementary Table S3).

Construction of a regulatory network. To create this network, the miR-96 direct targets (Supplementary Table S3; 31 genes) were treated as one separate group and the rest of the misregulated gene list, excluding the predicted targets, were a second separate group (Supplementary Table S4; 188 genes). The two groups were first connected to each other using Ingenuity IPA. The software has the option to connect input genes using its own manually curated database; for the regulatory network we report here, only this database was used to connect the input genes, and only regulatory interactions were permitted. Not every misregulated gene could be accounted...
for within the network (most notably Ptprq, for which no upstream regulators have been described), but when no further genes could be connected, every link was checked against the original published article and incorrect links were removed (for example, many papers studying Htt carry out experiments using the mutant poly-CAG form, which is not relevant for this network). Each pathway within the network was checked for consistency with respect to the observed gene misregulation, and the intermediate genes along that pathway were annotated with a predicted change of expression/activity (Fig. 4a). Where intermediate genes could be up- or downregulated, and either way could explain a number of observed gene expression changes, a weighting algorithm was used to determine which prediction explained more changes, first taking into account known misregulation, then predicted misregulation, of genes both up- and downstream (Fig. 4b). The resulting network is internally consistent and offers multiple hypotheses for further testing (Fig. 5).

The majority of genes (47) have only one downstream connection. 18 genes have two downstream connections, twelve genes have three, and eight genes have four. The maximum number of downstream connections belongs to miR-96, which has 31 direct targets predicted or confirmed in the literature. In between the two extremes are fourteen genes with between 7 and 17 downstream connections: Agt (10), Cebpa (7), Sdc2 (7), Parp1 (8), Irs1 (9), Foxo1 (10), Foxo3 (9), Htt (8), Mitf (10), Myc (10), Nr3c1 (14), Tgfβ1 (17), Tnf (17) and Trp53 (10). Of these, Sdc2 and Parp1 are known to be upregulated in diminuendo homozygotes, and only Irs1 and Foxo3 lack a direct connection to at least one gene with known misregulation (Fig. 5).

If these genes really do mediate the effects of miR-96, then they offer candidate therapeutic targets for modulating the network as a whole and ultimately the expression of many genes in the hair cells. Trp53, for example, regulates Ppp1r15a, Bhlhe40 and Gfi1 directly, and indirectly regulates Slc26a5, Ocm and Anxa4 via Myc and Fos (Fig. 4a). Trp53 is predicted to be downregulated in diminuendo homozygotes, either by a reduction in mRNA or protein levels, or by decreased activity, and either way could explain a number of observed gene expression changes, a weighting algorithm was used to determine which prediction explained more changes, first taking into account known misregulation, then predicted misregulation, of genes both up- and downstream (Fig. 4b). The resulting network is internally consistent and offers multiple hypotheses for further testing (Fig. 5).

The protein-protein interactome analysis of the genes with multiple connections (hub genes) has to be checked, for example by examining expression in the hair cells and testing for misregulation of mRNA or protein in diminuendo homozygotes.

From the entire network, we chose interesting nodes (genes) based on the number of their downstream connections, their position within the network, and their potential influence on downstream nodes, particularly Ocm and Slc26a5. The genes those nodes represented were tested by qRT-PCR, and Fos was found to be significantly downregulated in diminuendo homozygotes (adjusted P < 0.05), fitting the prediction. None of the other seven tested were significantly affected (Fig. 3b; Ets1, Tgfβ1, Sp1, Myc, Agt, Rhoa, Trp53), but this doesn’t rule them out as genuine interactors in the miR-96 regulatory network, because they may have altered protein levels or activity, neither of which would be detected by qRT-PCR. Myc, Sp1, Ets1 and Fos were also found to be present in wildtype hair cells by immunohistochemistry (Supplementary Fig. S2); no obvious difference in expression was observed in diminuendo homozygote hair cells.

Protein-protein interactome analysis. The regulatory network uses only regulatory interactions, and does not take advantage of the extensive protein-protein interaction (PPI) data publicly available. A different approach to network construction is to build a PPI network using the publicly available results from many experiments. This has the advantage of being less constrained by bias towards well-studied “favourite genes” such as Cdc258 but it does not result in a directional network, or in predictions of up- or downregulation that can be tested. We used InnateDB, a public PPI database combining data from multiple sources, to carry out over-representation and network analyses. In the first analysis we focussed on transcription factor and pathway over-representation, using the misregulated gene list (Supplementary Table S4) plus Fos, but no transcription factors or pathways were over-represented after adjustment for multiple testing (adjusted P < 0.05, Supplementary Table S5). We then carried out a network analysis on both the misregulated gene list (Supplementary Table S4 plus Fos) and the miR-96 direct target list (Supplementary Table S3) together, to obtain a PPI network which was analysed in Cytoscape (Supplementary Fig. S3), using the Network Analyst tool to calculate node degree, betweenness centrality and closeness centrality (Supplementary Figs S3, S4). The node degree is the number of neighbours a node has; nodes with a high degree are hub nodes. Betweenness centrality is a measure of how important a node is for connecting distant parts of the network, and closeness centrality measures how important a node is within its immediate vicinity.

Eighteen nodes had a high degree (>75, Table 1), but only Ywhae also scored highly on both centrality measures (Table 1, Supplementary Fig. S4). Of the high-degree nodes, five are predicted targets of miR-96 (Foxo3, Foxo1, Rad51, Irs1 and Nr3c1), and five are known to be misregulated in diminuendo homozygotes (Gfi1, Nphp1, Parp1, Fos and Fadd). The others are Ywhae, Ywhah, Ywhaz, Pgls, Hsd3b7, IDB-74, Ubc and Cenpc1. Ywhae, Ywhah and Ywhaz are members of the 14-3-3 protein family, which are ubiquitous adapter proteins, and Ubc is part of the ubiquitination system, so they have many known interaction partners. Pgls is a hydrolase involved in the pentose phosphate pathway39, Hsd3b7 is an enzyme required for bile acid synthesis40, and Cenpc1 is a component of the kinetochore plate41. Their high degree in the PPI network may reflect biological relevance, or it may be, like the 14-3-3 proteins, because they have a lot of interaction partners and among those partners are genes misregulated in the diminuendo homozygote organ of Corti. IDB-74 is an InnateDB identifier representing a large protein complex, which is why it has a high degree.

Gene Set Enrichment Analysis. Both the regulatory network and the PPI network were constructed knowing the context of the transcriptome data, that it came from a mouse carrying a mutation in miR-96. For both those networks, we chose to add the predicted miR-96 targets with the aim of connecting as many of the misregulated genes to miR-96 as possible. Gene set enrichment analysis, however, takes as input only the misregulated
genes, and looks for collections of genes which occur more often than would be expected by chance. These gene sets are predefined by the user, and for this analysis we used the following gene sets available at MsigDB (http://msigdb.org).
Figure 5. The Regulatory Network. The regulatory network constructed using regulatory interactions reported in the literature. Up- and downregulation is indicated by red and green colouration respectively. Pale green or pink means the node’s misregulation has been reported in the microarray, and intense green or red means the node’s misregulation has been shown by qRTPCR. miR-96 is shown in yellow. Arrows indicate positive regulation (activation, increased transcription), and bars show negative regulation (repression, reduction in activity); solid lines indicate direct regulation (for example, direct binding of protein to promoter) and dotted lines indirect regulation (where other genes may mediate the observed effect of one gene upon another).

Table 1. Nodes from the protein-protein interaction network with high degree (>75), showing betweenness and closeness scores and their misregulation in diminuendo homozygotes, if known (2 significant figures). mRNA fold changes and adjusted p-values are from P0 and P4 microarray analyses, with the exception of Fos, for which qRTPCR data is shown (transformed to fold change; see Fig. 3 for data).
We compared the most upregulated to the most downregulated genes, and also the most affected to the least affected genes, regardless of the direction of regulation. The input data consisted of the genes with a fold change greater than 10% from the P4 and P0 microarray data and the qRTPCR with adjusted P < 0.05. Where multiple probes were present for the same gene, the fold change with the lowest P-value was included, resulting in 2653 genes (Supplementary Table S6). Significantly enriched gene sets (false discovery rate (FDR) < 0.05) are shown in Table 2, as are all the gene sets with nominal P < 0.05.

One of the regulatory factors identified as having targets enriched in the upregulated genes is miR-96, which is a good positive control. Myc is the other factor with a significantly enriched target set among the upregulated genes, and is of particular interest, being downstream of Gfi1 and upstream of Fos, Slc26a5 and Ocm in the regulatory network (Fig. 4a). Among the factors with a nominal P < 0.05 are Egr4, Pou1f1, Tcf3, Cebpb and Vdr, which have been previously linked to inner ear development and/or deafness, as has Myc itself33–38. These genes are implicated in mediating the pattern of misregulation detected in diminuendo homozygotes and are, therefore, also candidates for further testing, and potential therapeutic targets. Multiple gene sets representing pathways and processes, which can also be targets for manipulation by drugs or small molecules, were identified as enriched in the input data, including the p53 pathway and synaptic transmission, which fits with our previous reports of synaptic defects in diminuendo homozygotes6.

**Discussion**

miR-96 is a master regulator of hair cell differentiation, and controls many genes. Here we have used several different tools to identify genes which play important roles in mediating the action of miR-96.

Our manually created regulatory network demonstrates a novel approach to network creation, and the resulting network has several important features. Firstly, some of the genes are independently known to be involved in hearing, such as Fgf8, Slc26a5 (prestin), Pou4f3 and Gfi19,10,39,40, suggesting that other genes in the network may also underlie deafness. Secondly, several of the genes are known to be involved in cellular stress response

**Table 2.** Gene Set Enrichment Analysis results showing gene sets with a nominal P < 0.05 from the microarray and qRT-PCR data, divided into genes regulated by regulatory factors, genes associated with cellular locations, genes belonging to defined pathways and genes linked to processes. *gene sets with FDR < 0.05.

<table>
<thead>
<tr>
<th>Type of gene set</th>
<th>Gene sets which are enriched in:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upregulated genes</strong></td>
<td>Downregulated genes</td>
</tr>
<tr>
<td>Regulatory factor target list</td>
<td>miR-96*, Myc*, Stat6, Spt1, CebpB</td>
</tr>
<tr>
<td>Genes associated with cellular location</td>
<td>Ribosome*</td>
</tr>
<tr>
<td>Genes involved in specific pathways</td>
<td>p53 pathway*</td>
</tr>
<tr>
<td></td>
<td>Peptide chain elongation*</td>
</tr>
<tr>
<td></td>
<td>3′UTR-mediated translational regulation*</td>
</tr>
<tr>
<td></td>
<td>Nonsense-mediated decay (enhanced by exon junction complex)*</td>
</tr>
<tr>
<td></td>
<td>Srp-dependent cotranslational protein targeting to membrane*</td>
</tr>
<tr>
<td></td>
<td>Translation*</td>
</tr>
<tr>
<td></td>
<td>Protein kinase activity</td>
</tr>
<tr>
<td></td>
<td>Collagen formation</td>
</tr>
<tr>
<td></td>
<td>RNA binding</td>
</tr>
<tr>
<td></td>
<td>Extracellular matrix organisation</td>
</tr>
<tr>
<td></td>
<td>Epithelial-mesenchymal transition*</td>
</tr>
<tr>
<td></td>
<td>Influenza viral RNA transcription and replication*</td>
</tr>
<tr>
<td></td>
<td>Influenza life cycle*</td>
</tr>
<tr>
<td></td>
<td>Metabolism of mRNA*</td>
</tr>
<tr>
<td></td>
<td>Metabolism of proteins</td>
</tr>
<tr>
<td></td>
<td>Cellular biosynthetic process</td>
</tr>
<tr>
<td></td>
<td>Skeletal development</td>
</tr>
<tr>
<td></td>
<td>Positive regulation of cell proliferation</td>
</tr>
<tr>
<td></td>
<td>Inflammatory response</td>
</tr>
</tbody>
</table>

pathways, such as Trp53, Hif1α and Nfe2l2 (reviewed in Simmons et al.41), suggesting that the network may be perturbed during stress triggered by environmental factors. Thirdly, a number of the network molecules are already known to be targeted by drugs, like Trp53, NfkB1 and Nr3c1, providing routes to manipulation of the network as a whole.

While most of the misregulated genes detected by microarray and qRT PCR are present in the protein–protein interaction network created using InNateDB, approximately 31% are not, including Prestin, Chrnα10 and Pou4f3. This may be because there are important targets of miR-96 that we have not discovered, and so are missing from the initial input, or it may be that the targets are correct but there is insufficient data from previous publications or databases on the interactions between the input molecules. Of the nodes identified as important by network analysis in Cytoscape, Parp1, Gfi1, Nphp1, Fos, Fadd, Nr3c1, Irs1, Rad51, Foxo1 and Foxo3 are present in the regulatory network, and Parp1, Gfi1, Nphp1, Fos and Fadd have been shown by microarray or qRT PCR to be misregulated in diminuendo homozygotes.

The gene set enrichment analysis implicates further regulators and regulatory pathways in the diminuendo phenotype, some of which are present in the regulatory network but most of which are not, suggesting that there are more pathways to be discovered. miR-96 targets (defined by MSigDB as genes sharing the motif GTGCCAAA, which matches the wildtype seed region, in their 3’ UTRs) are enriched in the upregulated genes, as would be expected. This correlates well with the Syamler results (Fig. 1). Of the transcription factors whose targets are enriched in the misregulated genes, Egr4, Pou1f1, Tcf3, Myc, Cebpβ and Vdr have been previously implicated in inner ear development and/or deafness33–38. None of the 31 direct targets of miR-96 (Supplementary Table S3) were implicated by GSEA, which could be because gene sets controlled by the direct targets were not available at the time of analysis, or it could mean that none of the 31 direct targets control the cascade of misregulated genes seen in diminuendo homozygote inner ears. Of the genes implicated by GSEA, only Atf2 is predicted by StarBase (http://starbase.sysu.edu.cn/index.php)39 to be a target of miR-96.

We have tried three methods of network analysis, all of which identified genes for further testing, but none of which offered a universal solution. Indeed, there is no common gene identified by all three methods as a mediator of the effects of miR-96 (Fig. 6). This reflects the different strengths, biases and weaknesses of the methods used to create the networks. The regulatory network is very dependent on what is reported in the literature, and emphasises connecting miR-96 to the misregulated genes observed in the diminuendo mutant, while the PPI network relies on large databases of protein–protein interactions, which do not offer directional, regulatory links between interacting partners. Gene Set Enrichment Analysis can be useful because it does not rely on correctly identifying the targets of miR-96, many of which remain unknown, but it does need comprehensive gene sets, which may not exist for all the genes mediating the effects of miR-96 in the inner ear. More data may help in the construction of more accurate and comprehensive networks. Transcriptome data from microarrays and RNA-seq, which is available from public databases such as ArrayExpress, may prove useful because they combine directional regulatory interactions with a genome-wide scope, hopefully avoiding at least some of the bias of well-studied genes. The networks are also all subject to a common limitation in that they cannot identify novel interactions, relying as they do on previous studies in other tissues or organisms. The lack of inner ear-specific data is also a potential problem; interactions observed in one type of tissue may not apply to the organ of Corti. Our analyses implicate Myc, Gfi1 and Fos as important candidates; Fos regulates both Slc26a5 and Ocm42,43, two of the most downregulated genes in the P1 microarray42, and when we tested it we confirmed its predicted downregulation in diminuendo homozygotes. In the regulatory network Fos is controlled by Gfi1 via Myc44,45, and Gfi1 is known to be downregulated in diminuendo homozygotes. Myc also regulates Anxa4, which is upregulated in diminuendo homozygotes and has recently been found to be specifically expressed in hair cells, along with Rassd2, another gene upregulated in diminuendo homozygote organ of Corti46. Both Gfi1 and Fos were also identified as important genes in the protein–protein interaction network, and Myc’s targets were detected by GSEA to be significantly enriched in the upregulated genes. All three have been implicated in hearing and deafness, Gfi1 in particular; the hair cells of mice homozygous for a Gfi1 knockout allele degenerate around P019. Mice homozygous for a null allele of Fos show a reduced startle response at higher frequencies47, and Myc has been implicated in protection against noise exposure-induced damage; guinea-pigs inoculated with adenovirus carrying c-myc showed a smaller threshold shift and less cochlear damage after noise exposure than controls inoculated with vector25. The other candidate targets are those genes identified by multiple networks (Fig. 6). Further testing in vitro would determine whether small molecule agonists or antagonists can modify the expression of these genes, and whether that has an effect further down the line on genes such as Ocm, Ptprq and Slc26a5, whose misregulation in diminuendo homozygotes appears to result in specific aspects of the observed phenotype41,42.

miR-96 controls many genes, and it is likely that many of the connections between the microRNA and the misregulated genes still remain to be discovered. However, the potential benefits of finding genes and pathways which can be modified by small molecules or drugs are significant. Altering the expression of important genes and pathways may prevent or delay hearing loss caused by mutations in miR-96, or indeed any other trigger which affects the miR-96 network, whether genetic or environmental. Discovering ways to counteract the effect of dysfunctional miR-96 could ultimately lead to therapeutic treatments to delay or prevent progressive hearing loss with a range of aetiologies in the human population.

Materials and Methods

Ethics statement. Mouse studies were carried out in accordance with UK Home Office regulations and the UK Animals (Scientific Procedures) Act of 1986 (ASPA) under UK Home Office licences, and the study was approved by both the Wellcome Trust Sanger Institute and the King’s College London Ethical Review Committees. Mice were culled using methods approved under these licences to minimize any possibility of suffering.
RNA extraction. The organs of Corti of newborn (P0) or four-day-old (P4) mice were dissected during a short time window (6 to 7.5 hours after lights on) and stored at −20 °C in RNAlater stabilisation reagent (QIAGen, cat. no. 76106). RNA was extracted using QIAshredder columns (QIAGen, cat. no. 79654) and the RNeasy mini kit (QIAGen, cat. no. 74104), following the manufacturer's instructions. RNA concentration was measured using a Nanodrop spectrophotometer (ND-8000).

Ago2 pulldown. The organs of Corti and olfactory bulbs of mice at P4 were dissected and immediately frozen in liquid nitrogen. Samples were treated as described previously24, with the following modifications; samples were homogenized in 300 ul ice-cold buffer, 30 ul of beads were used per sample, prebound to 8.8 ug anti-mouse monoclonal Ago2 antibody (Wako Pure: 018-22021) according to the Dynabead protocol (Invitrogen: 100.07D immunoprecipitation kit). Samples were incubated with antibody for 90 minutes. After immunoprecipitation, 300 ul TRIzol and 60 ul chloroform was added to extract RNA.

Microarray. For the microarray on P0 mice, organ of Corti samples from six Dmdo/Dmdo homozygotes and six sex-matched wildtype littermates were submitted for microarray analysis. For the Ago2 pulldown microarray, samples were pooled in pairs by genotype, resulting in three Dmdo/Dmdo homozygote pools and three wildtype pools from each tissue type (organ of Corti and olfactory bulb), and the twelve resulting pools were submitted for microarray analysis. RNA was amplified and purified using the Illumina TotalPrep-96 RNA Amplification kit (Ambion, UK), according to the manufacturer's instructions. Biotin-labeled cRNA was then normalized to a concentration of 150 ng/ul and 1500 ng was hybridised to Illumina MouseWG-6 v2 beadarrays (Illumina, CA, USA) for 16 hours (overnight) at 58 °C. Following hybridisation, beadarrays were washed and stained with streptavidin-cy3 (GE Healthcare, UK). Beadarrays were then scanned using the Beadarray reader and image data was then processed using Genome Studio software (Illumina, CA, USA). The data were normalised using a quantile normalisation, assuming that the overall intensity distributions of the arrays should be comparable48,49 and analysed using the LUMI and LIMMA packages from Bioconductor50,51. P-values were adjusted for multiple tests as described in Benjamini and Hochberg52. The microarray data have been submitted to the ArrayExpress database (www.ebi.ac.uk/arrayexpress) and assigned the identifiers E-MTAB-3735 (P0 data) and E-MTAB-3737 (Ago2 pulldown data).

Figure 6. The intersection of genes predicted to be important by the three methods of analysis. 28 genes were detected as having target genes enriched in the microarray data by GSEA, two of which, miR-96 and Myc, are also hub nodes (nodes with more than 7 downstream connections) in the regulatory network. The Cytoscape analysis of the protein-protein interaction network resulted in seventeen genes which were critical to the network, five of which (Parp1, Irs1, Foxo1, Foxo3 and Nr3c1) are hub nodes in the regulatory network.
Immunohistochemistry. P4 and P5 pups were collected, fixed in 10% formalin, embedded in paraffin and cut into 8μm sections. Immunohistochemistry was carried out using a Ventana Discovery machine and reagents according to the manufacturer's instructions: (DAkB84 Kit (cat.no 760-124), OmniMap DAB anti-Rb Detection Kit (cat.no 760-149), Hematoxylin (cat.no 760-2021), Bluing reagent (cat.no 760-2037), CCI (cat.no 950-124), EZPrep (cat.no 950-100), LCS (cat.no 650-010), RiboWash (cat.no 760-105), Reaction Buffer (cat.no 95-300), and Ribonuclease (cat.no 760-107)). For each antibody, at least three wildtype mice were tested, according to the same protocol, and from each animal, at least three sections were used per antibody. Primary antibodies used were: Anxa4 (Abcam, cat.no AB33009, diluted 1:100), Meig1 (Abcam, cat.no AB52012, diluted 1:7.5), Tppp (Abcam, cat.no AB48731, diluted 1:50, Homer1 (Lifespan Biosciences, cat.no LS-C30485, diluted 1:100), KcnA10 (Lifespan Biosciences, cat.no LS-C31214, diluted 1:25), Kalrn (Abcam, cat.no AB2015, diluted 1:5), Gf2e2 (Abcam, cat.no AB28178, diluted 1:125), Mif (Abcam, cat.no AB24875, diluted 1:10), Zic2 (Santa Cruz, cat.no sc-28151, diluted 1:10), Htt (Millipore, cat.no MAB2166, diluted 1:500), Fos (Santa Cruz, cat.no sc-52, diluted 1:75), Mhc (Santa Cruz, cat.no sc-764, diluted 1:25), Sp1 (Santa Cruz, cat.no sc-59, diluted 1:100), Ets1 (Abcam, cat.no AB59217, diluted 1:100) and Tnc (Santa Cruz, cat.no sc-20932, diluted 1:25). The secondary antibodies were anti-goat (Jackson ImmunoResearch, cat.no 705-065-147, diluted 1:100), anti-rabbit (Jackson ImmunoResearch, cat.no 711-065-152, diluted 1:100), and anti-mouse IgG1 (Epitomics, cat.no 3021-1, diluted 1:500). All antibodies were diluted in staining solution (10% foetal calf serum, 0.1% Triton, 2% BSA and 0.5% sodium azide in PBS). Images were taken using a Zeiss Axioskop 2 microscope with the Plan Neofluar 63× 1.4NA objective, a Zeiss AxioCam camera and the associated AxioCam software. Adobe Photoshop was used to process and prepare images; minimal adjustments were made, including rotation and scaling. Where image settings were altered, the adjustment was applied to the entire image.

qRT-PCR. cDNA was made from normalised organ of Corti RNA using Superscript II Reverse Transcriptase (Invitrogen, cat. no. 11904-018) after treatment with DNase 1 (Sigma, cat. no: AMP-D1). Quantitative RT-PCR was carried out using SYBR Green and reagents from Applied Biosystems (Master Mix: 4364340 and Bio-Rad (SsoFast and SsoAdvanced Master mixes, cat. nos 1725232, 1725281), and the 2^(-ΔΔCt) equation was used to calculate relative expression levels53. Hprt was used as an internal control, and the quantity of sensitive tissue sample was checked using Jag1 or Negr, which are expressed in supporting cells4,54–56. Pairs were only used when Jag1/Negr levels did not differ more than 30% between wildtype and Dmdo/Dmdo homozygote littersmates. At least three technical replicates of each sample were carried out for each reaction, and each primer/probe set was tested in at least five wildtype-homozygote littermate pairs, with the exception of Chnn1 and Chsl, which due to problems with the probes had only three technical replicates each. Primer details are in Supplementary Table S7; standard primers were designed by primer357 and qRT PCR primers were purchased from Applied Biosystems (Life Technologies). The Wilcoxon rank sum test was chosen to determine significance as a suitable test for small sample sizes and populations of unknown characteristics58. P-values were adjusted using the method described in Benjamini and Hochberg62 to take multiple testing into account.

Network construction and analysis software. Networks were generated with the aid of IPA (Ingenuity® Systems, www.ingenuity.com). InnateDB59 was used for interactome construction and pathway over-representation analysis with Cytoscape60 for visualisation and further analysis, and the Gene Set Enrichment Analysis (GSEA) software41 for enrichment analysis. InnateDB and GSEA both provide statistical analyses which are corrected for multiple testing.

References


Acknowledgements
We thank Rosalind Lacey, Claire Pearce and Dionne Cooper for animal care, and Chris McGee and Ruben Bautista-Garcia for assistance with the microarrays. This work was supported by the Wellcome Trust (grant numbers 100669 and 098051). J.M.H. was supported by a Wellcome Trust Sanger Institute Clinical PhD Fellowship.

Author Contributions
Dissection and RNA extraction from P0 mice was carried out by J.M.H. Dissection at P4 was done by M.A.L., S.E. and W.Z. and P4 RNA extraction was done by M.A.L., S.E., F.Z. and W.Z. Ago2 pulldown and RNA extraction was carried out by M.A.L. and A.B. S.V.D. and A.I.E. analysed the P0 microarray and performed the Sylamer analysis, and M.A.L. analysed the Ago2 pulldown microarray. cDNA synthesis and qRTPCR was carried out by M.A.L., W.Z., F.Z. and S.E. Immunohistochemistry was done by M.A.L., F.Z. and S.E. M.A.L. created and analysed the networks and carried out the gene set enrichment analysis. The study was designed and the paper was written by M.A.L. and K.P.S and reviewed by all authors.

Additional Information
Accession codes: ArrayExpress: E-MTAB-3735, E-MTAB-3737.
Supplementary information accompanies this paper at http://www.nature.com/srep
Competing financial interests: The authors declare no competing financial interests.
How to cite this article: Lewis, M. A. et al. Exploring regulatory networks of miR-96 in the developing inner ear. Sci. Rep. 6, 23363; doi: 10.1038/srep23363 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/