Tissue expression patterns identify mouse cilia genes

Timothy S. McClintock, Chad E. Glasser, Soma C. Bose, and Daniel A. Bergman

Department of Physiology, University of Kentucky, Lexington, Kentucky

Submitted 15 June 2007; accepted in final form 29 October 2007

McClintock TS, Glasser CE, Bose SC, Bergman DA. Tissue expression patterns identify mouse cilia genes. Physiol Genomics 32: 198–206, 2008. First published October 30, 2007; doi:10.1152/physiolgenomics.00128.2007.—In mammals, cilia are critical for development, sensation, cell signaling, sperm motility, and fluid movement. Defects in cilia are causes of several congenital syndromes, providing additional reasons to identify cilia-related genes. We hypothesized that mRNAs selectively abundant in tissues rich in highly ciliated cells encode cilia proteins. Selective abundance in olfactory epithelium, testes, vomeronasal organ, trachea, and lung proved to be an expression pattern uniquely effective in identifying documented cilia-related genes. Known and suspected cilia-related genes were statistically overrepresented among the 99 genes identified, but the majority encoded proteins of unknown function, thereby predicting new cilia-related proteins. Evidence of expression in a highly ciliated cell, the olfactory sensory neuron, exists for 73 of the genes. In situ hybridization for 17 mRNAs confirmed expression of all 17 in olfactory sensory neurons. Most were also detected in vomeronasal sensory neurons and in neighboring tissues rich in ciliated cells such as respiratory epithelium. Immunoreactivity for one of the proteins identified, Spa17, colocalized with acetylated tubulin in the cilia layer of the olfactory epithelium. In contrast, the ciliary rootlet protein, Crocc, was located in discrete structures whose position was consistent with the dendritic knobs of the olfactory sensory neurons. A compilation of >2,000 mouse genes predicted to encode cilia-related proteins revealed a strong correlation (R = 0.99) between the number of studies predicting a gene’s involvement in cilia and documented evidence of such involvement, a fact that simplifies the selection of genes for further study of the physiology of cilia.

Nearly all vertebrate cells have a single primary cilium. While its function in many cell types is a mystery, available evidence suggests the broad conclusion that the primary cilium is important for signaling and cell polarity (27, 30). A few cell types have elaborated and expanded on this single cilium, generating specialized structures that perform critical functions of several broad types: sensation, development, fluid movement, sperm motility, and cell signaling. The functional significance of cilia in tissues is reflected in the severity and diversity of pathologies caused by defects in cilia. These include anosmia, retinitis pigmentosa and retinal degeneration, polycystic kidney disease, diabetes, neural tube defects and neural patterning defects, chronic sinusitus and bronchiectasis, obesity, heterotaxias, polydactyly, and infertility (3). This list includes defects in highly specialized cilia, exemplified by retinitis pigmentosa and anosmia, as well as defects in primary cilia, exemplified by polycystic kidney disease (23, 24). Defects in cilia are therefore underlying causes of several congenital disease syndromes with pleiotropic symptoms. These include Alstrom syndrome, Bardet-Biedl syndrome, Kartagener syndrome, Meckel-Gruber syndrome, Senior-Loken syndrome, orofacial digital syndrome I, Joubert syndrome, and primary cilia dyskinesia (3). Investigation of mammalian cilia genes will enhance not only our knowledge of ciliary function but also of dysfunctions that contribute to significant human health problems.

The function of cilia in cell signaling is evident in the specialization of cilia for sensation. In the mammalian olfactory epithelium, for example, olfactory cilia are elaborated during the maturation of olfactory sensory neurons (OSNs). This event is a defining feature that distinguishes a mature OSN from its immature antecedent (8, 28). These nonmotile cilia are the sites where odorants interact with odorant receptors, triggering G protein-dependent generation of cAMP and subsequent electrical activity in the OSN. Not surprisingly, the bioinformatics of expression profiling of OSN mRNAs consistently identifies cilia/flagella and ciliogenesis/spermatogenesis as overrepresented biological processes in OSNs (26, 29). This link between OSNs and sperm, and our observation that mRNAs linked to cilia were detected in both OSNs and in cells of the neighboring respiratory epithelium, another tissue enriched in highly ciliated cells, led to the hypothesis that tissue expression pattern could be used as an ab initio approach to identify genes whose products are important to cilia.

The ability of bioinformatics and expression profiling to identify cilia-related genes has been documented several times. Comparison of the genomes of ciliated and nonciliated organisms predicts hundreds of cilia-related genes in Chlamydomonas reinhardtii and Drosophila melanogaster (2, 18). The presence of X box promoter elements, likely targets of the cilia-related transcription factors of the Rfx family, are also predictors of cilia-related genes (10, 34). Proteomics on cilia-enriched samples from C. reinhardtii, Tetrahymena aurelia, and cultured human respiratory epithelial cells predict cilia proteins (21, 22, 31). Expression profiling to compare ciliated vs. nonciliated cells in C. elegans or mRNAs that increased during replacement of flagella in C. reinhardtii predict additional cilia-related genes (5, 32). Although previous efforts have been substantial, we reasoned that additional cilia-related genes remain to be identified. This is especially true in mammals, which were the focus of only one of these earlier studies. We therefore compared our data on gene expression in mouse OSNs with gene expression patterns across >60 mouse tissues to predict mouse cilia-related genes (26, 33). We identified 99 putative cilia-related genes, many of which had not previously been linked to cilia and most of which encode proteins of unknown function.

MATERIALS AND METHODS

Bioinformatics. Mouse tissue expression patterns in which mRNAs were most abundant in olfactory epithelium, testes, vomeronasal...
organ, trachea, and lung were identified via SymAtlas, a repository of microarray data on mRNA abundance in many mouse tissues (33). Seven such genes were identified. The correlation function resident in SymAtlas was applied to these seven mRNAs to identify additional mRNAs sharing their expression patterns. Each of the seven lists of correlated mRNAs was ordered by the strength of correlation and then truncated before the first mRNA whose abundance pattern failed to have two of the five target tissues as the tissues with the largest amounts of that mRNA. This threshold corresponded to correlation coefficients that ranged from 0.90 to 0.99 for the seven mRNAs tested.

We then combined the seven correlations, tracking how many correlation sets included each mRNA and excluding genes whose mRNAs were found in only one correlation set. As a filter for evidence of expression in highly ciliated cells, we then used expression profiling data generated from separating the olfactory epithelium into purified expression in highly ciliated cells, we then used expression profiling data generated from separating the olfactory epithelium into purified expression in tissues rich in highly ciliated cells. Data obtained from SymAtlas (33). Microarray signal intensities reported as gcRMA processed data.

![Diagram of mRNA expression patterns](image)

**Fig. 1.** Tissue expression patterns identify mRNAs most abundant in tissues rich in highly ciliated cells. Shown are the 3 most divergent patterns among the 7 mRNAs used for the correlation analysis that identified putative cilia-related genes. The 20 tissues with the highest signal intensities for each mRNA are listed. Each of the seven lists of correlated mRNAs was ordered by the strength of correlation and then truncated before the first mRNA whose abundance pattern failed to have two of the five target tissues as the tissues with the largest amounts of that mRNA. This threshold corresponded to correlation coefficients that ranged from 0.90 to 0.99 for the seven mRNAs tested. We then combined the seven correlations, tracking how many correlation sets included each mRNA and excluding genes whose mRNAs were found in only one correlation set. As a filter for evidence of expression in highly ciliated cells, we then used expression profiling data generated from separating the olfactory epithelium into purified expression in highly ciliated cells, we then used expression profiling data generated from separating the olfactory epithelium into purified expression in tissues rich in highly ciliated cells. Data obtained from SymAtlas (33). Microarray signal intensities reported as gcRMA processed data.

**Immunohistochemistry.** Cryostat sections at 10-μm thickness on slides (Fisher Superfrost Plus) were prepared from wild-type C57Bl/6J mice of both sexes, aged P21–P28, and fixed in 4% paraformaldehyde as described for in situ hybridization above. Slides were dried with a hair dryer for 20 s and washed in PBS for 5 min, PBS containing 1% Triton X-100 for 30 min, and PBS for 5 min. Blocking was then done with 2% BSA and 0.4% Triton X-100 in PBS for 1 h in a humidated chamber at room temperature. Primary antibodies were used at monoclonal against acetylated tubulin (Sigma no. T6793) at 1/1,000 dilution and rabbit polyclonal antiserum against Sp17 (16) at 1/200 and Croc (Root6) (38) at 1/5,000. The specificity of these antibodies and their utility for immunohistochemistry are well documented (16, 37, 38). Slides were incubated in primary antibodies overnight in the blocking solution at 4°C in a humidated chamber. Washing was in PBS for 10 min, PBS containing 0.5% Tween 20 for 10 min, and PBS for 10 min. Slides were incubated in cyanine-3(Cy3)-conjugated anti-rabbit secondary antibody in PBS for 1 h in a humidated chamber at room temperature in the dark. Washing was then done with PBS for 10 min, PBS containing 0.5% Tween 20 for 10 min, and PBS for 10 min. Slides were then incubated in FITC-conjugated anti-rabbit secondary antibody in PBS for 1 h in a humidated chamber at room temperature in the dark. Washing was in PBS for 10 min twice in the...
Table 1. Predicted mouse cilia-related genes, identified by tissue expression patterns

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Symbol</th>
<th>Frequency</th>
<th>Ciliome Database?</th>
<th>PubMed ID</th>
<th>ISH</th>
<th>OSN+/OSN− Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>69329</td>
<td>1700003M02Rik</td>
<td>7</td>
<td>yes</td>
<td></td>
<td>OSN, RE</td>
<td>6.2</td>
</tr>
<tr>
<td>66353</td>
<td>2310007A19Rik</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>77048</td>
<td>1700007C10Rik</td>
<td>7</td>
<td>yes</td>
<td></td>
<td>OSN, VSN, RE</td>
<td>5.3</td>
</tr>
<tr>
<td>71845</td>
<td>4922501L14Rik</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>75040</td>
<td>4930504H06Rik</td>
<td>7</td>
<td>yes</td>
<td></td>
<td>OSN, VSN, RE</td>
<td>8</td>
</tr>
<tr>
<td>68283</td>
<td>9530077C05Rik</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>223690</td>
<td>Ankrd54</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>77609</td>
<td>C330001K17Rik</td>
<td>7</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>70664</td>
<td>Dnajb14</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>75563</td>
<td>Dnali1</td>
<td>7</td>
<td>yes</td>
<td>16496424</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>330830</td>
<td>Gm7770</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>gnf1m07624_a_at</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>75050</td>
<td>Kif27</td>
<td>7</td>
<td>*17089004</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>329164</td>
<td>LOC329164</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>70144</td>
<td>Lrch3</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>76688</td>
<td>Mdh1b</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>75333</td>
<td>Nme5</td>
<td>7</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>69310</td>
<td>Pacrg</td>
<td>7</td>
<td>yes</td>
<td>16278296</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>21689</td>
<td>Tek1</td>
<td>7</td>
<td>yes</td>
<td>11089920</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>74407</td>
<td>Ttc25</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>78802</td>
<td>Ttc30a1</td>
<td>7</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>237930</td>
<td>Tll6</td>
<td>7</td>
<td>*15890843</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>68625</td>
<td>1110020C03Rik</td>
<td>6</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>69315</td>
<td>1700001L19Rik</td>
<td>6</td>
<td>yes</td>
<td>15242845</td>
<td>OSN, VSN, RE</td>
<td>6.3</td>
</tr>
<tr>
<td>75472</td>
<td>1700009P17Rik</td>
<td>6</td>
<td>yes</td>
<td>15242845</td>
<td>OSN, VSN, RE</td>
<td>5.9</td>
</tr>
<tr>
<td>227736</td>
<td>1700019L03Rik</td>
<td>6</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>216516</td>
<td>4930562D19Rik</td>
<td>6</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>228778</td>
<td>6820408C15Rik</td>
<td>6</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>244653</td>
<td>Hydin</td>
<td>6</td>
<td></td>
<td>12719380</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>67661</td>
<td>Itf172</td>
<td>6</td>
<td></td>
<td>16061793</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>50525</td>
<td>Spag6</td>
<td>6</td>
<td></td>
<td>12167721</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>68777</td>
<td>Tmem53</td>
<td>6</td>
<td>yes</td>
<td>15385169</td>
<td>OSN, VSN, RE</td>
<td>12.5</td>
</tr>
<tr>
<td>22092</td>
<td>Tsga2</td>
<td>6</td>
<td></td>
<td>16354795</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>71860</td>
<td>Wdr16</td>
<td>6</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>66758</td>
<td>Zipf47</td>
<td>6</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>73721</td>
<td>1110017D15Rik</td>
<td>5</td>
<td></td>
<td>15242845</td>
<td>OSN, VSN, RE</td>
<td>6.3</td>
</tr>
<tr>
<td>67085</td>
<td>1700024G13Rik</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>76420</td>
<td>1700027A23Rik</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>73287</td>
<td>1700040L02Rik</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>74895</td>
<td>4930455F23Rik</td>
<td>5</td>
<td>yes</td>
<td>15242845</td>
<td>OSN, VSN, RE</td>
<td>10.6</td>
</tr>
<tr>
<td>75172</td>
<td>4930528G09Rik</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>70977</td>
<td>4931407G18Rik</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>71233</td>
<td>4933434H06Rik</td>
<td>5</td>
<td>yes</td>
<td>15385169</td>
<td>OSN, VSN, RE</td>
<td>12.5</td>
</tr>
<tr>
<td>74934</td>
<td>Armc4</td>
<td>5</td>
<td></td>
<td>16809635</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>67331</td>
<td>Atupb83</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>218232</td>
<td>AW456874</td>
<td>5</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>244608</td>
<td>Ccdc113</td>
<td>5</td>
<td>yes</td>
<td>15257753</td>
<td>OSN, VSN, RE</td>
<td>6.4</td>
</tr>
<tr>
<td>51938</td>
<td>Ccdc39</td>
<td>5</td>
<td>yes</td>
<td>15257753</td>
<td>OSN, VSN, RE</td>
<td>5.7</td>
</tr>
<tr>
<td>105833</td>
<td>Ccdc65</td>
<td>5</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>207175</td>
<td>Cetn4</td>
<td>5</td>
<td></td>
<td>12802058</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>71877</td>
<td>Ehbc1</td>
<td>5</td>
<td>yes</td>
<td>15670853</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>238726</td>
<td>Gm276</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>16977</td>
<td>Lrc23 (Lrp7)</td>
<td>5</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>20686</td>
<td>Spa17</td>
<td>5</td>
<td>yes</td>
<td>15257753</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>72236</td>
<td>Tsnaip1</td>
<td>5</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>76670</td>
<td>Ttc18</td>
<td>5</td>
<td>yes</td>
<td>12802058</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>74052</td>
<td>Ttc21a</td>
<td>5</td>
<td>yes</td>
<td>15670853</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>269701</td>
<td>Wdr66</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>69234</td>
<td>Zipf688</td>
<td>5</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>67483</td>
<td>1700028P14Rik</td>
<td>4</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>66706</td>
<td>4733401H18Rik</td>
<td>4</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>69312</td>
<td>4930418G15Rik</td>
<td>4</td>
<td>yes</td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>70882</td>
<td>Armc3</td>
<td>4</td>
<td></td>
<td>16809635</td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>330323</td>
<td>C330043M08Rik</td>
<td>4</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>382137</td>
<td>D630064A14Rik</td>
<td>4</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>331595</td>
<td>Gm777</td>
<td>4</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
<tr>
<td>67694</td>
<td>Ift74 (Cdc2)</td>
<td>4</td>
<td>yes</td>
<td>15955805</td>
<td>OSN, VSN, RE</td>
<td>2.3</td>
</tr>
<tr>
<td>27356</td>
<td>Ins6b</td>
<td>4</td>
<td></td>
<td></td>
<td>nd</td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table 1.—Continued

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Symbol</th>
<th>Frequency</th>
<th>Ciliome Database?</th>
<th>PubMed ID</th>
<th>ISH</th>
<th>OSN+/OSN− Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>53885</td>
<td>Nphp1</td>
<td>4</td>
<td>yes</td>
<td>12872123</td>
<td>OSN, VSN</td>
<td>2.4</td>
</tr>
<tr>
<td>114402</td>
<td>Zmynd10</td>
<td>4</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>34.5</td>
</tr>
<tr>
<td>66061</td>
<td>0610012D17Rik</td>
<td>3</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>2.3</td>
</tr>
<tr>
<td>71837</td>
<td>1700003E16Rik</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>nd</td>
</tr>
<tr>
<td>75556</td>
<td>1700026D08Rik</td>
<td>3</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>nd</td>
</tr>
<tr>
<td>71864</td>
<td>1700026J04Rik</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>24.7</td>
</tr>
<tr>
<td>73376</td>
<td>1700061J05Rik</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>8</td>
</tr>
<tr>
<td>235345</td>
<td>4833427G06Rik</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>nd</td>
</tr>
<tr>
<td>70966</td>
<td>4931415C17Rik</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>16.3</td>
</tr>
<tr>
<td>207686</td>
<td>A330021E22Rik</td>
<td>3</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>nd</td>
</tr>
<tr>
<td>68644</td>
<td>Abbhl14a</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>4.1</td>
</tr>
<tr>
<td>232566</td>
<td>C730024G19Rik</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>1.1</td>
</tr>
<tr>
<td>241116</td>
<td>Cdc108</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>0.9</td>
</tr>
<tr>
<td>102124</td>
<td>E130303B06Rik</td>
<td>3</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>13.5</td>
</tr>
<tr>
<td>13516</td>
<td>Epyc(Dspg3)</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>nd</td>
</tr>
<tr>
<td>381738</td>
<td>Gm1060</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>3.7</td>
</tr>
<tr>
<td>210762</td>
<td>Gm70</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>1.1</td>
</tr>
<tr>
<td>none</td>
<td>gnf1m11188_a_at</td>
<td>3</td>
<td>nd</td>
<td></td>
<td>OSN, VSN</td>
<td>nd</td>
</tr>
<tr>
<td>69325</td>
<td>1700012B09Rik</td>
<td>2</td>
<td>OSN, VSN, RE</td>
<td></td>
<td></td>
<td>43.2</td>
</tr>
<tr>
<td>654809</td>
<td>4930428O21Rik</td>
<td>2</td>
<td>OSN, VSN, RE</td>
<td></td>
<td></td>
<td>3.4</td>
</tr>
<tr>
<td>217887</td>
<td>BC022687</td>
<td>2</td>
<td>OSN, VSN, RE</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>327951</td>
<td>Cby5d1</td>
<td>2</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>nd</td>
</tr>
<tr>
<td>71200</td>
<td>Dyk2</td>
<td>2</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>nd</td>
</tr>
<tr>
<td>15223</td>
<td>Fovj1</td>
<td>2</td>
<td>yes</td>
<td>9739041</td>
<td>OSN, VSN</td>
<td>5.5</td>
</tr>
<tr>
<td>18415</td>
<td>Hspa4l</td>
<td>2</td>
<td>yes</td>
<td>16923965</td>
<td>OSN, VSN</td>
<td>1.1</td>
</tr>
<tr>
<td>71827</td>
<td>Lrc34</td>
<td>2</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>5.2</td>
</tr>
<tr>
<td>54562</td>
<td>Lrc6</td>
<td>2</td>
<td>yes</td>
<td></td>
<td>OSN, VSN</td>
<td>4</td>
</tr>
<tr>
<td>104362</td>
<td>Meig1</td>
<td>2</td>
<td>osn, re</td>
<td></td>
<td></td>
<td>6.6</td>
</tr>
<tr>
<td>73075</td>
<td>Ppl6</td>
<td>2</td>
<td>osn, re</td>
<td></td>
<td></td>
<td>4.2</td>
</tr>
<tr>
<td>19725</td>
<td>Rfx2</td>
<td>2</td>
<td>osn, re</td>
<td></td>
<td></td>
<td>nd</td>
</tr>
<tr>
<td>66832</td>
<td>Rsh2</td>
<td>2</td>
<td>osn, re</td>
<td></td>
<td></td>
<td>1.1</td>
</tr>
</tbody>
</table>

Frequency, how often the mRNA appeared in the 7 tissue expression pattern correlation gene lists; Ciliome database, whether the gene is listed at http://www.ciliome.com; PubMed ID, a representative publication linking the encoded protein to cilia (*published evidence is indirect or circumstantial); ISH, in situ hybridization results; OSN, olfactory sensory neuron; RE, respiratory epithelium; VSN, vomeronasal sensory neurons; OSN+/OSN− ratio, predicts likelihood of expression in OSNs at 99% for ratios ≥1.3 and at 83% for ratios between 1.2 and 0.5 (26); nd, not determined.

RESULTS AND DISCUSSION

Identification of mouse cilia-related genes by tissue expression pattern. We reasoned that the core elements of cilia and cilia support structures (such as the axoneme, basal body, and intraflagellar transport particles) are enriched in highly ciliated cell types, loosely defined as cells that elaborate large or multiple cilia. We identified seven mRNAs that shared expression in mature OSNs and a subset of cells in neighboring respiratory epithelium and found that they shared a tissue expression pattern of high abundance in tissues with numerous highly ciliated cells: olfactory epithelium, testes, vomeronasal organ, trachea, and lung (Fig. 1). We used these mRNAs to identify 99 mouse genes whose mRNAs had highly correlated tissue expression patterns (Table 1). Of these, 39 are the subject of papers archived in PubMed, and of these 39, 18 are directly linked to cilia or flagella by these publications. This proportion is significant even if we make the cautious prediction that 25% of all genes are linked to cilia (binomial test; H0: cilia gene fraction = 0.25; B* = 3.05, P < 0.001). The frequency with which mRNAs were detected in the seven tissue expression pattern correlations was not correlated with the proportion with proven links to cilia as determined by articles archived in PubMed. This finding argues that even those mRNAs whose abundance patterns were highly similar to only two of the seven test mRNAs are not less likely to be important to cilia.

Confirmation of expression in OSN, a highly ciliated cell. If the mRNAs we identified do indeed encode proteins important to cilia, then these mRNAs must be expressed in the ciliated cells of these tissues. We tested this by in situ hybridization in the olfactory epithelium where the only highly ciliated cells, mature OSNs, can be unambiguously identified in tissue sections because of the position of their cell bodies (Fig. 2A). Published evidence already indicates that 73 of the mRNAs are expressed in OSNs (Table 1; OSN+/OSN− ratios >0.8) (26).
We tested 17 mRNAs, including three that have proven links to cilia (4933434I06Rik, Ift74, and 1110017D15Rik). All 17 mRNAs were detected in the mature OSN layer of the epithelium. Examples are shown in Fig. 2, B–O, and the data are summarized in Table 1 [column titled ISH (in situ hybridization)]. Thirteen of the mRNAs were also detected in the neurons of the vomeronasal organ (not shown, but summarized in Table 1). These tissue sections also contain neighboring regions with other highly ciliated cells, such as respiratory epithelium and nasal gland linings. Of the 17 mRNAs, only A330021E33Rik failed to unequivocally label subsets of cells in these epithelia. Examples of labeling of respiratory epithelium and nasal glands can be seen in the low-magnification images of Fig. 2, B′–O′. Even though the ciliated cells in respiratory epithelium and nasal glands cannot be unambiguously identified by in situ hybridization because their cell bodies are intermingled with other cell types, the presence of the mRNAs in these epithelia is additional evidence supporting the conclusion drawn directly from the localization of the mRNAs to OSNs and vomeronasal sensory neurons. We conclude that our tissue expression pattern approach identified mRNAs expressed primarily by highly ciliated cells.

The cell and tissue expression patterns of the 99 mRNAs we identified were consistent with the expectation that many of the encoded proteins reside in cilia or cilia support structures. However, some may instead have functions in other cellular compartments, such as nuclei or vesicular trafficking compartments, yet still be important to cilia. None of the novel proteins we linked to cilia is as yet a target of useful antibodies. However, our analyses predicted that Spa17 is present in olfactory cilia. Spa17 has previously been localized to sperm flagella and cilia of the respiratory tract and reproductive organs and is known to be expressed by OSNs, but it has not been localized to the cilia of OSNs (6, 12, 16). Not surprisingly, we observed that Spa17 immunoreactivity in the olfactory epithelium overlaps completely with immunoreactivity for acetylated tubulin, a marker for cilia (Fig. 2P). In contrast, immunoreactivity for the ciliary rootlet protein, Crocc, did not (Fig. 2Q). Crocc is the major component of ciliary rootlets that extend basally from the basal bodies of cilia (37, 38). We observed Crocc immunoreactivity in round structures just below immunoreactivity for acetylated tubulin, consistent with its expected location in the dendritic knobs from which olfactory cilia arise. Immunoreactivity for Spa17 and Crocc in neighboring respiratory epithelium gave similar patterns. Spa17 overlapped completely with acetylated tubulin labeling of cilia, and Crocc immunoreactivity was found directly below the cilia layer (data not shown). These data provide additional reasons to expect that many of the 99 genes we identified will prove to be important to cilia.

**Tissue expression pattern predictions overlap with other predictive approaches.** As described above, several other methods have been used to predict cilia-related genes with some degree of success. We hypothesized that our method would overlap significantly with these previous efforts. Indeed, 39 of our 99 genes are probable orthologs of genes previously predicted to be related to cilia (15). This proportion is greater than a random sampling of genes should yield (binomial test; $H_0$: cilia gene fraction $= 0.25$; $B^*$ $= 6.37$, $P < 0.0001$). Of these 39, 27 as yet have no direct evidence of importance to cilia. The mouse genes we identified that are probable orthologs of genes previously linked to cilia were not limited to orthology with mammalian genes. Just six were probable orthologs of genes previously linked to cilia only in humans. Instead, most were orthologs of nonmammalian genes, or of both human and nonmammalian genes.

The basic logic of our approach, an expression pattern largely restricted to tissues with large numbers of highly ciliated cells, appeared to be uniquely consistent with cilia genes. We were unable to find other patterns involving other tissues where cilia have important functions, such as pituitary, retina, kidney, and cartilage/bone (11, 14, 36), that had even a few mRNAs whose abundance was so highly correlated (co-efficiencies $>0.9$). This tight correlation corresponded with rough estimates of the proportion of highly ciliated cells in the tissues, with testes, olfactory epithelium, vomeronasal organ, trachea, and lung having substantial proportions, whereas these other tissues have relatively few highly ciliated cells, or in the case of the retina and its photoreceptors, have cells whose cilia have diverged into a very specialized structure. The frequency with which mRNAs were detected in the seven tissue expression pattern correlations was not correlated with the proportion of them listed in http://www.ciliome.com (15). This finding is consistent with the expectation that our tissue expression pattern approach does not share biases inherent in previous approaches.

The reliability of our search strategy was better than that of previous methods, judging by the frequency of genes shared with other predictive studies. Genes identified by multiple high-throughput approaches would seem more likely to be true cilia-related genes than those detected only once. The frequency of identification of predicted cilia genes is listed at http://www.ciliome.com, which summarizes previous work in this area (Table 1) (15). Only 22% of the genes listed were identified by multiple studies. For the genes we identified, the proportion was 48%. Overall, the available evidence suggests that the proportion of true cilia-related genes in our list will prove to be substantial.

*Many candidates are genes of unknown function.* Like previous bioinformatics approaches to identify cilia-related genes,
our tissue expression pattern approach was not biased by gene type. We detected genes involved in signaling, transcriptional regulation, protein chaperone function, cytoskeletal structure, and cytoskeletal dynamics, all of which are functions needed by cilia. However, what was remarkable about the genes as a group was how little we know about them. For 60 of them, we found no evidence of published work on them or their products. Because our approach was unbiased, these genes could encode products involved in virtually any aspect of cilia function or support of cilia, such as the axoneme, basal body, protein trafficking, and transcription of other cilia-related genes. Expression patterns limited to tissues enriched in motile cilia (testes, trachea, lung) were unsuccessful in identifying a different group of highly correlated mRNAs that could be associated specifically with motility. Our search pattern included tissues with both motile and nonmotile cilia, so we conclude that the genes we identified are likely to encode components of processes shared by both motile and nonmotile cilia. These processes are not necessarily restricted to cilia or the basal body but could instead be processes necessary to support cilia that happen at other locations such as the nucleus or the cell body.

Several of the most interesting proteins encoded by the 99 genes we identified have never been studied. Proteins with coiled-coil domains appear to be important to cilia (22). This domain mediates protein interactions, especially with other coiled-coil domain proteins, predicting that additional coiled-coil domain proteins might be important to cilia (7). Ccdc39, Ccdc65, Ccdc113, and 1700003M02Rik are all predicted to possess coiled-coil domains, and all have putative orthologs in other species that have been linked to cilia by at least one other high-throughput approach (15). Similarly, tetratricopeptide repeat domain (Ttc) proteins and WD (tryptophan aspartic acid) repeat domain proteins have been linked to cilia (1, 4, 9, 13, 20, 25). We identified five such genes, three of which have putative orthologs that have been tentatively linked to cilia in other species: Ttc18, Ttc21a, and Wdr66. Last, 0610012D17Rik encodes a novel protein that is predicted to contain a Tcex-1 domain. Tcex-1 is a dynein light chain that interacts with rhodopsin, implicating it in the trafficking of rhodopsin to the photoreceptor outer segment (35). 0610012D17Rik may therefore be a component of a more general mechanism of trafficking proteins to cilia.

A mouse ciliome resource. Our prediction that the identification of mammalian cilia genes is incomplete was supported not only by our prediction of additional mouse cilia genes but also by published work on genes that were not detected by high-throughput approaches. We detected several proteins with published links to cilia, such as Foxj1 (19), whose genes were absent from http://www.ciliome.com (Table 1) (15). A literature search for mouse and human cilia-related genes revealed 87 genes that have documented importance to cilia but were not detected by high-throughput approaches. Surprisingly, 62% of the mouse genes with documented importance to cilia were not predicted by previous high-throughput methods, nor were probable orthologs of these genes. Combining the genes identified herein (Table 1) with those identified in our literature search and those listed in http://www.ciliome.com identifies 2,128 mouse genes that have been linked to cilia in some fashion (Supplemental File 1).

We expect this mouse ciliome resource to be neither complete, as discussed above, nor wholly accurate. For example, a few genes in our results may be selectively abundant in olfactory epithelium, vomeronasal organ, testes, trachea, and lung for reasons unrelated to cilia. More broadly based bioinformatics and expression profiling approaches are even more likely to make this type of false inclusion error. Similarly, the proteomics approach suffers from limited sensitivity, may misidentify some proteins, and is prone to include proteins from contaminating subcellular fractions. In fact, >75% of the genes predicted as important to cilia by high-throughput approaches were identified by only one of the eight previous studies (Fig. 3), a statistic that argues that many cilia gene predictions are erroneous. Users should therefore keep in mind that the mouse ciliome resource suffers from omissions of cilia-related genes and from inclusion of genes not truly important to cilia and may be complicated by the identification of genes important to cilia in some organisms and tissues but not others.

Even with these imperfections, the compilation of documented and predicted cilia-related genes provides a valuable resource. As the pace of prediction of cilia-related genes has far outstripped the ability to test these predictions, significant opportunities to investigate proteins critical for cilia function and dysfunction have arisen. Supplemental File 1 provides a rational basis for identifying which proteins are the best candidates for as yet undocumented roles in cilia or the support of cilia. We hypothesized that the number of studies linking a gene to cilia would correlate with the frequency of genes independently validated as encoding cilia proteins. Indeed, we found a significant correlation within the exponential relationship between these factors (Fig. 3). Genes identified by multiple approaches are therefore the most promising candidates...
for future studies to test whether the encoded proteins are sufficient and necessary for cilia function.

Conclusions

The tissue expression pattern we discovered proved to be significantly better than chance at identifying known and suspected cilia-related genes. It identified 99 mRNAs, of which 48 have some link to cilia, ranging from documented involvement in cilia to prediction by previous high-throughput studies. Most of the 99 genes encode proteins that have never been studied, supporting the idea that much remains to be learned about cilia. The correlation of documented importance to cilia with the frequency of prediction of links to cilia indicates that our approach was more reliable than previous approaches, but this came at the expense of limiting the number of genes identified. The mouse ciliome resource we compiled from nine different sources contains 2,128 mouse genes with documented or predicted importance to cilia. Users of this resource should understand that additional mammalian cilia-related genes remain to be discovered, and that some proportion of the predictions will prove to be erroneous. Genes identified in multiple studies are the most promising for future study of cilia function and dysfunction.

ACKNOWLEDGMENTS

We thank Drs. Michael O’Rand and Tiansen Li for the gift of antisera against Spai17 and Crocc, respectively.

Present address of D. A. Bergman: Dept. of Biomedical Sciences, Grand Valley State Univ., 130 Padnos Hall, Allendale, MI 49401.

GRANTS

This work was supported by National Institute of Deafness and Other Communications Disorders Grants R01-DC-002736 and R01-DC-007194 and Boyarsky Professorship funds to T. S. McClintock.

REFERENCES


