Global gene annotation analysis and transcriptional profiling identify key biological modules in hypoxic pulmonary hypertension

Sina A. Gharib,1,3 Daniel L. Luchtel,2 David K. Madtes,1,3* and Robb W. Glenny1*

1Division of Pulmonary and Critical Care Medicine and 2Department of Environmental and Occupational Health Sciences, School of Public Health and Community Medicine, University of Washington, Seattle, and 3Section of Pulmonary and Critical Care Medicine, Fred Hutchinson Cancer Research Center, Seattle, Washington

Submitted 8 November 2004; accepted in final form 15 March 2005

Gharib, Sina A., Daniel L. Luchtel, David K. Madtes, and Robb W. Glenny. Global gene annotation analysis and transcriptional profiling identify key biological modules in hypoxic pulmonary hypertension. Physiol Genomics 22: 14–23, 2005. First published March 22, 2005; 10.1152/physiolgenomics.00265.2004.—Chronic hypoxic pulmonary hypertension is an important clinical disorder causing significant morbidity. Despite recent discoveries, many molecular mechanisms involved in its pathogenesis remain unexplored. We have undertaken a systematic and unbiased approach to gain global insights into this complex process. By combining transcriptional profiling with rigorous statistical methods and cluster analysis, we identified the dominant temporal patterns of gene expression during progression and regression of hypoxic pulmonary hypertension. We next integrated these results with global gene annotation analysis to identify key biological themes involved in the development and resolution of hypoxic pulmonary hypertension and vascular remodeling. This novel approach assigned biological roles to thousands of candidate genes based on their temporal expression profiles and membership in specific biological modules. Our procedure confirmed several molecular pathways and gene products known to be important in hypoxic pulmonary hypertension. Furthermore, we discovered several novel candidates and molecular mechanisms, including IQ motif containing GTPase-activating protein-1 (IQGAP1), decorin, insulin-like growth factor binding protein-3 (IGFBP3), and lactotransferrin, that may play crucial roles in hypoxic pulmonary hypertension and vascular remodeling. Our methodology of integrating transcriptional profiling, cluster analysis, and global gene annotation provides new insights into the pathophysiology of pulmonary hypertension and is applicable to other models of human disease.

pulmonary hypertension; hypoxia; microarray; gene annotation; pulmonary vascular remodeling

CHRONIC PULMONARY ARTERIAL HYPERTENSION is a common clinical disorder causing significant morbidity and mortality (31). Chronic pulmonary hypertension has several different etiologies, but hypoxic pulmonary vasoconstriction is the primary mechanism in the development of pulmonary hypertension associated with sleep apnea, chronic obstructive pulmonary disease, and mountain sickness. Chronic elevation of pulmonary arterial pressure due to hypoxia induces profound remodeling of the vasculature (37). The media thickens due to hypertrophy and proliferation of smooth muscle cells and proliferation of fibroblasts in the adventitia. Extracellular matrix (ECM) remodeling with increased deposition of elastin and collagens and activation of matrix metalloproteinases and growth factors are other hallmarks of this process. Although the histopathological changes seen in pulmonary arterioles during chronic hypoxia are less severe than those seen in primary pulmonary hypertension, hypoxic pulmonary hypertension is a much more common clinical entity (31).

In recent years, several important regulators and molecular pathways in the development of pulmonary vascular remodeling due to chronic hypoxia have been identified. They fall into diverse functional groups, including transcription factors (hypoxia-inducible factor-1), vasoactive molecules (prostacyclin, endothelin-1, nitric oxide), growth factors (transforming growth factor-β, vascular endothelial growth factor), and ECM components (matrix metalloproteinases, tenasin C) (21). The literature on genes involved in regression of pulmonary hypertension and vascular remodeling is more limited, although the expression patterns of several ECM-degrading genes during regression of pulmonary vascular remodeling in a rat model of hypoxic pulmonary hypertension have been reported (40). Despite important advances in elucidating the molecular mechanisms of hypoxic pulmonary hypertension and vascular remodeling, we believe that only a subset of the critical genes participating in this process has been identified to date.

The present study was designed to undertake a systematic search for novel genes and mechanisms involved in progression and regression of pulmonary hypertension and vascular remodeling in a murine model of hypoxia-induced chronic pulmonary hypertension. We utilized global expression profiling during induction and resolution of hypoxic pulmonary hypertension to obtain complex temporal patterns of gene expression and systematically narrowed candidate gene lists using rigorous statistical criteria and classification algorithms. We independently confirmed differential gene expression, using quantitative real-time RT-PCR for several candidates and protein level for one candidate gene. By statistically linking thousands of differentially expressed genes to gene annotation databases, we obtained a global view of the key biological processes in hypoxic pulmonary hypertension. Integration of cluster analysis with gene annotation enabled us to identify several novel pathways and candidate genes that may be involved in development and regression of pulmonary hypertension.

MATERIALS AND METHODS

Experiment design. Ninety 8-wk-old male Balb/C mice (Jackson Laboratory) were randomized to experimental and control groups (Fig. 1). All mice were housed in the same specific pathogen-free environment. The experimental group was placed in a hypobaric chamber and exposed to 0.5 atm inspired partial pressure of oxygen...
(P\textsubscript{O\textsubscript{2}}) = 75 Torr) for 21 days and then returned to sea level (P\textsubscript{O\textsubscript{2}} = 150 Torr) for an additional 14 days. Control mice were housed at sea level for the entire duration of experiment (35 days). On days 1, 3, 7, 14, and 21, five mice from each group were anesthetized with an intraperitoneal injection of ketamine (70 mg/kg) and xylazine (10 mg/kg) and killed by exsanguination; this study was approved by the University of Washington’s Institutional Animal Care and Use Committee and was performed in accordance with our institutional animal guidelines. The pulmonary arteries were flushed with 30 ml of RNase-free PBS solution, and whole lungs were then removed, immediately frozen in liquid nitrogen, and stored at −70°C. The remaining mice (n = 20) in the experimental group were returned to normoxia on day 21. On days 22, 24, 28, and 35, another five mice from each group were killed and their lungs removed and processed as described above. At each time point, the right ventricle-to-total heart mass ratio (RV/T) of each mouse was measured to determine the presence of right ventricular hypertrophy (RVH) resulting from chronic hypoxia-induced pulmonary hypertension.

Pulmonary vascular immunohistochemistry and morphometry. Excised lungs were fixed via intratracheal instillation with 4% formalin and maintained at a transmural pressure of 25 cmH\textsubscript{2}O overnight. Lung samples were embedded in paraffin, and 5-μm sections were stained with a monoclonal antibody against α-smooth muscle actin (Dako) with dianemobenidine (DAB) as the chromogen (43). The extent of vascular remodeling was determined by measuring percent wall thickness defined as (area\textsubscript{ext} − area\textsubscript{int})/area\textsubscript{ext} × 100, where area\textsubscript{ext} and area\textsubscript{int} are the areas bounded by the external and internal elastic lamina, respectively (29). Fifty-eight arterioles of comparable size were measured in 13 lungs (n = 5, chronic hypoxia; n = 5, normoxia; n = 3, chronic hypoxia with return to normoxia). Statistical significance was determined using the Wilcoxon-Mann-Whitney rank sum test.

RNA isolation and cDNA microarray hybridization. Whole lungs of experimental and control mice were pooled separately at each of the nine time points. Total RNA was isolated by use of the Qiagen RNaseasy Maxi Kit (Qiagen). RNA integrity was confirmed with the Agilent Bioanalyzer 2100 system (Agilent Technologies). Fifty micrograms of total RNA from experimental and control mice were labeled with Cy5 and Cy3 monoreactive dyes with the use of a modified amino-allyl reverse transcription protocol (Joseph DeRisi, [http://www.microarrays.org](http://www.microarrays.org)). Four replications with dye swapping were performed at each time point for a total of 36 microarray labeling experiments. cDNA microarrays were manufactured at Fred Hutchinson Cancer Research Center and consisted of 15,388 murine genes and expressed sequence tags (ESTs) from a clone set provided by the National Institute on Aging (NIA; [http://lgsun.grc.nia.nih.gov/cDNA/15k.html](http://lgsun.grc.nia.nih.gov/cDNA/15k.html)) (39). All clones had been resequenced at the NIA with an error rate estimation of <2% before the manufacturing of cDNA microarrays (23). We used the latest gene identifiers for this clone set provided by NIA (Feb. 2004). After overnight hybridization, each array was scanned using a GenePix 4000A dual-laser scanner. Image analysis was performed using GenePix Pro 4.0 software (Axon Instruments). For each spot, the median background intensity was subtracted from the median spot intensity to minimize background noise. Complete details of experiment design, hybridization procedures, and raw microarray data in compliance with Minimum Information About a Microarray Experiment (MIAME) have been deposited at the Gene Expression Omnibus website ([http://www.ncbi.nlm.nih.gov/projects/geo](http://www.ncbi.nlm.nih.gov/projects/geo), query GSE19099).

cDNA microarray data analysis. Cy3 and Cy5 dyes have different fluorescence properties that result in systematic dye bias. We addressed this issue by performing four replications at each time point with dye swapping. We then utilized a Lowess global intensity normalization procedure for each microarray experiment ([http://www.r-project.org](http://www.r-project.org), sma package) (45). Differential gene expression at each time point was determined using Significance Analysis of Microarrays (SAM; [http://www-stat.stanford.edu/~tibs/SAM](http://www-stat.stanford.edu/~tibs/SAM)) (41). Missing values were imputed using a K-nearest-neighbor algorithm. The multiple comparisons problem was addressed by choosing a false discovery rate (FDR) of 5%.

Clustering of temporal gene expression data was performed using partitioning around medoids (PAM) algorithm ([http://www.r-project.org](http://www.r-project.org), cluster package) (24). This novel method iteratively assigns data set objects to k representative objects (medoids) such that the total dissimilarity of all objects to their nearest medoid is minimized. PAM is more robust than other partitioning algorithms such as k-means that implicitly assume each cluster has a spherical normal distribution. PAM also uses a graphical display known as the silhouette plot and a corresponding quality index to determine the optimal number of clusters in the data set.

**Quantitative real-time RT-PCR.** Quantitative real-time RT-PCR (qRT-PCR) of selected candidate genes was performed using Taqman Gene Expression Assays (Applied Biosystems). For each gene, cDNA was synthesized from 2 μg of total RNA using SuperScript II reverse transcriptase kit (Invitrogen). qRT-PCR was performed in duplicate using the ABI Prism 7700 Sequence Detection System. After performing validation experiments using serial dilutions and the relative standard curve method, we used the comparative cycle threshold (CT) method (2−ΔΔCT) to calculate relative gene expression under experimental and control conditions normalized to 18S control (26).

**ELISA analysis.** We undertook an independent set of experiments to assess the differential protein expression of one candidate gene, insulin-like growth factor-binding protein-3 (IGFBP3). Forty adult male Balb/C mice were placed in a hypobaric chamber and exposed to 0.5 atm pressure for 3 wk and then returned to sea level pressure for an additional 2 wk. On days 7, 14, 21, 28, and 35, eight mice were killed and their lungs removed and weighed. Lungs were not pooled. Thirteen mice kept at sea level pressure were killed and used as controls. Harvested lungs were homogenized in 1.0 ml of protease inhibitor cocktail (Roche Diagnostics). The homogenate was then vigorously mixed with a buffer containing 0.5% Triton X-100, 150 mM NaCl, 15 mM Tris, 1 mM CaCl\textsubscript{2}, and 1 mM MgCl\textsubscript{2} (pH 7.40); incubated for 30 min at 4°C; and then centrifuged at 10,000 g for 20 min.

The supernatants were stored at −70°C. IGFBP3 concentrations for control and hypoxic mice were measured in duplicate using a
DuoSet ELISA kit (R&D Research) and according to the manufacturer’s instructions using IGFBP3 standard curves. IGFBP3 protein level was normalized to measured lung weight for each animal. Student’s t-test was used to determine statistical difference in IGFBP3 levels between control and hypoxic mice.

**Gene annotation.** Gene annotation of all 15,388 genes/ESTs present on cDNA microarrays was obtained by linking them to the Gene Ontology (GO) database (http://www.geneontology.org) with the use of Gene MicroArray Pathway Profiler software (GenMAPP 2.0, http://www.genmapp.org). Next, we used MAPPFinder (an ancillary program in GenMAPP) to assess whether specific molecular pathways were overrepresented among the differentially expressed genes and within specific gene clusters (10). We confirmed the results of our gene annotation approach using another program, Expression Analysis Systematic Explorer (EASE; http://david.niaid.nih.gov/david/ease.htm) (18). Although EASE, like GenMAPP, links microarray expression data to their GO annotation and identifies abundantly represented molecular pathways, it uses a different statistical approach to test for significance.

**RESULTS**

**Progression and regression of pulmonary hypertension and RVH.** During 21 days of exposure to hypobaric hypoxia, mice developed pulmonary hypertension and RVH as evidenced by a progressive increase in their RV/T compared with control mice (Fig. 2) (2). Pulmonary hypertension regressed after hypoxic mice were returned to normoxia, resulting in the resolution of RVH.

**Development of vascular remodeling due to chronic hypoxic pulmonary hypertension.** Pulmonary vascular morphometry based on medial wall thickness of immunostained arterioles was performed on mice exposed to 21 days of hypoxia (n = 5), mice exposed to 21 days of hypoxia and then returned to normoxia for 14 days (n = 3), and control mice (n = 5). There was a significant increase in medial thickness of mice exposed to hypobaric hypoxia compared with controls (36.3 ± 1.4 vs. 32.5 ± 2.3%, means ± SE, P = 0.05), consistent with vascular remodeling (Fig. 3). There was no difference in the medial wall thickness of pulmonary arterioles between chronically hypoxic mice that were returned to normoxia and control mice, implying regression of pulmonary vascular remodeling after return to normoxia [34.1 ± 2.9 vs. 32.5 ± 2.3%, P = not significant (NS)].

**Statistical analysis of cDNA microarray data.** After completing image analysis and nonlinear normalization of all 36 microarrays, we had data from >15,000 genes with four replicates per gene during each of nine time points, totaling >500,000 gene expression values. Figure 1 outlines our general approach to reduce the size and complexity of the microarray data. Initially, we determined statistically significant gene expression using SAM, choosing an FDR of 5%. We focused our attention on only the genes that met this criterion during at least two time points. This approach reduced our gene list from 15,388 to 1,752 genes. Next, we assessed the robustness of our selection criteria by reanalyzing the data at an FDR of 1%. Over 95% of the genes we had identified remained significant at an FDR of 1% during at least one time point. We then verified our results from SAM by determining significant gene expression using a Bayesian approach outlined by Baldi and Long (4) and implemented as Cyber-T (http://visitor.ics.uci.edu/genex/cybert/). After obtaining Bayesian P values for each gene at each time point, we applied FDR analysis following the method of Benjamini and Hochberg (6). At an FDR of 5%, there was a 98.1% overlap between Bayesian analysis and SAM, independently confirming the robustness of our statistical approach.

We then performed PAM clustering algorithm on the 1,752 significant genes and obtained the dominant temporal expression patterns of this data set by dividing it into seven clusters (Fig. 4). These clusters revealed distinct temporal patterns of gene expression during progression and regression of pulmonary hypertension and vascular remodeling. Although cluster analysis is a purely statistical classification method, we were able to interpret biological meaning from its results based on our experimental design. **Cluster 1** showed genes that are upregulated during development of pulmonary hypertension but not significantly changed after return to normoxia. Expression patterns of several members of this cluster were highly correlated with the pattern of progression and regression of RVH, further suggesting an association between gene expression and development of pulmonary hypertension (20). These genes included a metalloproteinase (MMP2, r = 0.81), a serine protease inhibitor (Wdfc1, r = 0.78), and a growth factor (IGF1, r = 0.71), all of which are examples of processes contributing to vascular remodeling (21). **Cluster 2** identified genes primarily downregulated during progression of pulmo-

![Fig. 2. Development of right ventricular hypertrophy (RVH) due to chronic hypoxia-induced pulmonary hypertension as measured by right ventricle-to-total heart mass ratio (RV/T). Hypoxic mice (gray boxes) developed RVH compared with control mice kept at sea level (white boxes). After return to normoxia, RVH resolved in the hypoxic group. \*P < 0.005 and \**P < 0.05 (hypoxic vs. control mice, Student’s t-test).](image_url)

![Fig. 3. Development of vascular remodeling due to chronic hypoxic pulmonary hypertension. Representative α-smooth muscle actin-immunostained pulmonary arterioles show increased medial wall thickness in a hypoxic mouse compared with a normoxic animal. Unlike several other mammals such as rats, mice do not undergo extensive pulmonary vascular remodeling due to chronic hypoxia.](image_url)
nary hypertension. Clusters 3 and 4 depicted genes up- or downregulated only after a return to normoxia; they may therefore be important in regression of pulmonary vascular remodeling. Clusters 5 and 6 revealed genes that were differentially expressed during both progression and regression of pulmonary hypertension. This finding may imply that both processes share many of the same pathways. Cluster 7 is comprised of only 11 genes (by far the smallest cluster) and is dominated by members of the hemoglobin family. In general, these genes are upregulated during the hypoxic phase of the experiment. We believe this cluster simply represents the presence of nucleated red blood cells within harvested lungs and will therefore vary from sample to sample. Our clustering approach captured the salient patterns of gene expression in our model. Because our perturbation consisted of 21 days of hypoxia followed by 14 days of normoxia, we expected the dominant gene expression patterns to “pivot” around day 22. This is precisely the pattern we observed in clusters 1–6, which represent 99.3% of the differentially expressed genes. A full gene list for each cluster is available as online Supplementary Material (available at the Physiological Genomics web site).1

Gene annotation analysis and identification of significant molecular processes. We utilized two independent methods, GenMAPP/MAPPFinder and EASE, to obtain global insights

1 The Supplemental Material for this article is available online at http://physiolgenomics.physiology.org/cgi/content/full/00265.2004/DC1.
into the pathogenesis of hypoxic pulmonary hypertension by determining significant molecular processes involved in its development and regression. We initially linked our entire microarray data set consisting of 15,388 genes/ESTs to the GO database using GenMAPP. This allowed us to annotate our gene products based on their molecular functions, biological processes, and cellular components. We then applied MAPPFinder to the subset of 1,752 differentially expressed genes to statistically determine the molecular pathways and processes that were most overrepresented during progression and regression of pulmonary hypertension. MAPPFinder calculates a z-score and obtains a permutation-based P value to determine statistical significance. Next, we confirmed our approach by performing a similar but independent analysis using the EASE program. EASE uses a variant of the one-tailed Fisher exact probability and the Gaussian hypergeometric probability distribution to determine an “EASE score” that ranks overrepresented GO processes. Additionally, EASE performs a permutation analysis to calculate the global FDR. FDR addresses the multiple comparisons problem that arises when many processes are evaluated simultaneously. Table 1 lists several of the highest scoring GO processes identified by both EASE and MAPPFinder. Cell adhesion was the most statistically overrepresented biological process in our subset of 1,752 differentially expressed genes.

Although cluster analysis is a statistical classification method, we designed our experiment such that putative biological roles could be inferred for genes within specific clusters. We used MAPPFinder and EASE to assess whether specific biological processes were assigned to individual clusters of differentially expressed genes. Cluster 1 (Fig. 4) showed the temporal profile of genes upregulated during hypoxic pulmonary hypertension. Table 2 lists the dominant processes when MAPPFinder and EASE are applied to this cluster. Both methods reveal very similar results and confirm that extracellular matrix constituents, cell adhesion, and metal ion homeostasis are the most abundantly represented processes in this cluster. Cluster 3 depicted genes upregulated only after return to normoxia, when pulmonary hypertension regresses. Highly significant molecular processes in this cluster identified by both MAPPFinder and EASE were overwhelmingly represented by ribosomal activity and protein synthesis (Table 3). Table 4 lists the dominant processes in cluster 4, whose members are downregulated only after return to normoxia. Interestingly, GO analysis reveals that cell adhesion is an important process in both cluster 1 and cluster 4, suggesting a reciprocal role for cluster members during progression and regression of pulmonary hypertension, respectively. Several other clusters of differentially expressed genes also had dominant biological themes, implying the functional relevance of our clustering approach (see Fig. 7; data for all clusters are available as Supplemental Material).

### Table 1. Highly significant overrepresented processes among all 1,752 differentially expressed genes as identified by MAPPFinder and EASE

<table>
<thead>
<tr>
<th>GO Category</th>
<th>EASE Score</th>
<th>FDR</th>
<th>Z-Score</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell adhesion</td>
<td>2.28 × 10⁻⁶</td>
<td>0</td>
<td>5.042</td>
<td>0</td>
</tr>
<tr>
<td>Heat shock protein activity</td>
<td>7.57 × 10⁻⁴</td>
<td>0</td>
<td>4.218</td>
<td>0</td>
</tr>
<tr>
<td>Extracellular space</td>
<td>1.11 × 10⁻⁴</td>
<td>0</td>
<td>3.428</td>
<td>0</td>
</tr>
<tr>
<td>Cytoskeleton</td>
<td>0.00342</td>
<td>0</td>
<td>3.152</td>
<td>0</td>
</tr>
<tr>
<td>Heparin binding</td>
<td>0.00868</td>
<td>0</td>
<td>2.936</td>
<td>0.005</td>
</tr>
<tr>
<td>Transition metal ion homeostasis</td>
<td>0.00494</td>
<td>0</td>
<td>2.905</td>
<td>0.01</td>
</tr>
<tr>
<td>Enzyme inhibitor activity</td>
<td>0.00525</td>
<td>0</td>
<td>2.832</td>
<td>0.01</td>
</tr>
</tbody>
</table>

EASE, expression analysis systematic explorer; FDR, false discovery rate; GO, Gene Ontology.

### Table 2. Highly significant overrepresented processes among genes in cluster 1 as identified by MAPPFinder and EASE

<table>
<thead>
<tr>
<th>GO Category</th>
<th>EASE Score</th>
<th>FDR</th>
<th>Z-Score</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECM conferring tensile strength</td>
<td>1.15 × 10⁻⁷</td>
<td>0</td>
<td>10.612</td>
<td>0</td>
</tr>
<tr>
<td>Collagen</td>
<td>3.94 × 10⁻⁷</td>
<td>0</td>
<td>10.612</td>
<td>0</td>
</tr>
<tr>
<td>ECM</td>
<td>4.51 × 10⁻⁷</td>
<td>0</td>
<td>8.116</td>
<td>0</td>
</tr>
<tr>
<td>Cell adhesion</td>
<td>2.73 × 10⁻⁷</td>
<td>0</td>
<td>6.608</td>
<td>0</td>
</tr>
<tr>
<td>Basement membrane</td>
<td>0.00284</td>
<td>0</td>
<td>5.616</td>
<td>0.002</td>
</tr>
<tr>
<td>Metal ion homeostasis</td>
<td>0.00216</td>
<td>0</td>
<td>6.121</td>
<td>0.004</td>
</tr>
<tr>
<td>Development</td>
<td>0.000683</td>
<td>0</td>
<td>3.018</td>
<td>0.007</td>
</tr>
</tbody>
</table>

### Table 3. Highly significant overrepresented processes among genes in cluster 3 as identified by MAPPFinder and EASE

<table>
<thead>
<tr>
<th>GO Category</th>
<th>EASE Score</th>
<th>FDR</th>
<th>Z-Score</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytosolic ribosome</td>
<td>9.56 × 10⁻¹⁰</td>
<td>0</td>
<td>11.906</td>
<td>0</td>
</tr>
<tr>
<td>Ribosome</td>
<td>1.18 × 10⁻¹⁰</td>
<td>0</td>
<td>10.209</td>
<td>0</td>
</tr>
<tr>
<td>Structural constituent of ribosome</td>
<td>3.37 × 10⁻¹²</td>
<td>0</td>
<td>10.036</td>
<td>0</td>
</tr>
<tr>
<td>Ribonucleoprotein complex</td>
<td>3.55 × 10⁻⁹</td>
<td>0</td>
<td>8.31</td>
<td>0</td>
</tr>
<tr>
<td>Structural molecular activity</td>
<td>3.76 × 10⁻⁸</td>
<td>0</td>
<td>7.612</td>
<td>0</td>
</tr>
<tr>
<td>Protein biosynthesis</td>
<td>2.40 × 10⁻⁷</td>
<td>0</td>
<td>5.482</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 4. Highly significant overrepresented processes among genes in cluster 4 as identified by MAPPFinder and EASE

<table>
<thead>
<tr>
<th>GO Category</th>
<th>EASE Score</th>
<th>FDR</th>
<th>Z-Score</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat shock protein activity</td>
<td>0.000201</td>
<td>0</td>
<td>7.508</td>
<td>0</td>
</tr>
<tr>
<td>Chaperone activity</td>
<td>0.000139</td>
<td>0</td>
<td>4.886</td>
<td>0</td>
</tr>
<tr>
<td>Response to stress</td>
<td>0.00127</td>
<td>0</td>
<td>4.825</td>
<td>0</td>
</tr>
<tr>
<td>Cell adhesion</td>
<td>0.000796</td>
<td>0</td>
<td>4.209</td>
<td>0</td>
</tr>
<tr>
<td>Protein folding</td>
<td>0.00266</td>
<td>0</td>
<td>3.978</td>
<td>0.001</td>
</tr>
<tr>
<td>Cytoskeleton</td>
<td>0.000891</td>
<td>0</td>
<td>3.556</td>
<td>0.001</td>
</tr>
<tr>
<td>Microtubule cytoskeleton</td>
<td>0.000793</td>
<td>0</td>
<td>3.981</td>
<td>0.002</td>
</tr>
</tbody>
</table>
**BIOLOGICAL MODULES IN HYPOXIC PULMONARY HYPERTENSION**

HYPOXIC PULMONARY HYPERTENSION. Finally, Hoshikawa et al. did not report significant vascular remodeling in mice; this could be due to strain differences between the mice they studied (C57BL/6) and our Balb/C mice, which have been reported to undergo pulmonary vascular remodeling under chronic hypoxia (17).

An innovative feature of our approach was to initially obtain global insights into this disorder by identifying the key biological processes activated during its development and resolution and then to search for novel candidate genes and processes within these biological modules. Utilizing transcriptional profiling, we measured differential gene expression during progression and regression of hypoxic pulmonary hypertension and clustered genes into groups based on their temporal expression patterns. By clustering genes with similar temporal profiles into distinct groups, putative biological roles for candidate genes during progression and regression of pulmonary hypertension were determined.

Next, we used global gene annotation analysis to statistically identify highly represented biological processes in our model of hypoxic pulmonary hypertension. Furthermore, we were able to show that distinct biological themes were present within specific gene clusters, validating our methodology of integrating biological function with statistical classification (Fig. 7).

We would like to stress that the above analyses were performed in a systematic and unbiased fashion, with no a priori interest in a specific molecular pathway or gene product.

A prominent feature of chronic hypoxic pulmonary hypertension is the development of vascular and ECM remodeling. This process likely involves complex interactions among cell adhesion molecules, ECM constituents, and proteases and their inhibitors. These processes are also among the most statistically significant ones we identified through global gene annotation analysis (Table 1). Cluster analysis of temporal patterns of gene expression enabled us to assign several of these biological processes to distinct clusters and thereby integrate functional metal ion homeostasis but has not been reported to be involved in vascular remodeling.

**ELISA confirmation of IGFBP3 expression.** IGFBP3 protein levels in whole lung homogenates were significantly reduced in mice (n = 8 for each time point) exposed to 7, 14, and 21 days of hypoxia compared with normoxic animals (n = 13) by 38% (P < 0.05), 45% (P < 0.02), and 43% (P < 0.05), respectively. Mice returned to normoxia (for 1 or 2 wk) after 3 wk of hypoxia (n = 8 for each time point) had a nonsignificant reduction in IGFBP3 levels (27%, P = NS during both time points) compared with normoxic mice. These results are summarized in Fig. 6.

**DISCUSSION**

Pulmonary hypertension due to chronic hypoxia is a clinically important disorder, yet many of the molecular pathways involved in its pathogenesis remain unexplored. Because hypoxic pulmonary hypertension is a complex and dynamic process, studying its evolution and resolution over time can provide valuable insights into its pathogenesis. This study is the first report of the temporal expression profiles of thousands of genes during both progression and regression of hypoxic pulmonary hypertension in mice. Hoshikawa et al. (19) have used transcriptional profiling in a comparative investigation of gene expression in C57BL/6 mice and Sprague-Dawley rat lungs exposed to hypoxia. However, the primary purpose of that novel study was to compare differential gene expression in mice vs. rats during hypoxia. Additionally, their experiment was limited to two time points (1 and 3 wk of hypoxia) and did not sample animals during the resolution of hypoxic pulmonary hypertension. Finally, Hoshikawa et al. did not report significant vascular remodeling in mice; this could be due to strain differences between the mice they studied (C57BL/6) and our Balb/C mice, which have been reported to undergo pulmonary vascular remodeling under chronic hypoxia (17).

An innovative feature of our approach was to initially obtain global insights into this disorder by identifying the key biological processes activated during its development and resolution and then to search for novel candidate genes and processes within these biological modules. Utilizing transcriptional profiling, we measured differential gene expression during progression and regression of hypoxic pulmonary hypertension and clustered genes into groups based on their temporal expression patterns. By clustering genes with similar temporal profiles into distinct groups, putative biological roles for candidate genes during progression and regression of pulmonary hypertension were determined.

Next, we used global gene annotation analysis to statistically identify highly represented biological processes in our model of hypoxic pulmonary hypertension. Furthermore, we were able to show that distinct biological themes were present within specific gene clusters, validating our methodology of integrating biological function with statistical classification (Fig. 7).

We would like to stress that the above analyses were performed in a systematic and unbiased fashion, with no a priori interest in a specific molecular pathway or gene product.

A prominent feature of chronic hypoxic pulmonary hypertension is the development of vascular and ECM remodeling. This process likely involves complex interactions among cell adhesion molecules, ECM constituents, and proteases and their inhibitors. These processes are also among the most statistically significant ones we identified through global gene annotation analysis (Table 1). Cluster analysis of temporal patterns of gene expression enabled us to assign several of these biological processes to distinct clusters and thereby integrate

---

**Fig. 5. Temporal expression profiles of selected candidate genes from microarray experiments (black boxes) confirmed with quantitative real-time RT-PCR (gray boxes).** IQGAP1, IQ motif containing GTPase-activating protein-1; IGFBP3, insulin-like growth factor binding protein-3.

**Fig. 6. IGFBP3 protein levels by ELISA in whole lung homogenates of mice (n = 8 for each time point) exposed to 0.5 atm pressure for up to 21 days (days 7, 14, 21) and then returned to sea level pressure for 7 or 14 days (days 28 and 35).** The IGFBP3 concentration of each mouse has been normalized to its lung mass (in g) and is shown as mean ± SE. **P < 0.05 and **P < 0.02 (hypoxic vs. control mice, Student’s t-test).**

---

**Physiol Genomics • VOL 22 • www.physiolgenomics.org**
statistical classification of genes with their biological function (Tables 2–4). For example, many ECM components were upregulated during development of pulmonary hypertension, including collagens-1, -3, -4 and -18, matrix metalloproteinase-2, vimentin, and tenascin C; these genes were all grouped into cluster 1 (Fig. 4 and Table 2). Many of these ECM components have been previously reported to be differentially expressed during hypoxia-induced pulmonary hypertension (40), including the report on tenascin C in mice of Hoshikawa et al. (19).

However, not all genes involved in ECM remodeling were grouped into cluster 1. This makes intuitive sense, since vascular remodeling is a dynamic process resulting from both up- and downregulation of many gene products. Cluster 2 depicts genes primarily downregulated during hypoxic pulmonary hypertension. A member of this cluster, decorin, is a small proteoglycan that interacts with a variety of ECM components and regulates cell attachment, migration, and proliferation. Decorin prevents smooth muscle cell proliferation in balloon-injured arteries by inhibiting transforming growth factor (TGF)-β and platelet-derived growth factor signaling, and has been shown to suppress tumor cell-mediated angiogenesis by inhibiting vascular endothelial growth factor (VEGF) (16, 32).

On the basis of these findings, it is being considered as a possible therapeutic target for prevention of vascular injury, angiogenesis-dependent tumor growth, and pulmonary fibrosis (16, 32, 36). Decorin expression is significantly reduced in our model of pulmonary hypertension (Fig. 5), suggesting a state of unhindered smooth muscle proliferation and ECM remodeling. Decorin has been shown to be expressed in pulmonary vessels but has not been previously reported in pulmonary vascular remodeling and may therefore represent a novel target for therapeutic intervention in pulmonary hypertension (13).

Insulin-like growth factors (IGFs) comprise an important family of molecules regulating cell adhesion, growth, and proliferation. Differentially expressed members of this family in our model include IGF-1 and IGFBP3, which were assigned to clusters 1 and 2, respectively. IGF-1 is a potent stimulator of growth, proliferation, and migration in various tissues including smooth muscle cells, but there are only a few in vitro studies investigating its role in pulmonary arterial hypertension (3). IGFBP3 can bind to and inhibit the actions of IGF-1 and has additional anti-proliferative and apoptotic properties mediated through TGF-β and VEGF pathways (9). Although IGFBP3 has not been implicated in pulmonary vascular remodeling, its expression has been shown to

Fig. 7. Simplified graphical representation of significantly overrepresented biological themes in hypoxic pulmonary hypertension based on the Gene Ontology (GO) database. GO assigns gene products to modules (ovals) based on 3 distinct ontologies: molecular function, biological process, and cellular component (boxes). By determining the biological modules that are overabundant within specific gene clusters (shown in different colors), the time-dependent pattern of activation of these pathways can be elucidated. For example, the blue ovals highlight the key processes activated in cluster 1; the gene expression of members of this cluster are upregulated during the development of pulmonary hypertension and vascular remodeling (see also Fig. 4, Tables 1–4). For clarity, only 3 clusters have been displayed.
be significantly decreased in human atherosclerotic plaques, and recent studies have characterized complex interactions between IGFBP3, IGF-1, collagen-1, and matrix metalloproteinases, suggesting a role for these genes in vascular and ECM remodeling (12, 25, 30, 44). In our model, IGF-1 expression is upregulated during development of pulmonary hypertension, whereas IGFBP3 expression is significantly downregulated both at the mRNA and protein levels (Figs. 5 and 6). This finding implies a loss of inhibition of IGF-1-mediated growth and proliferation and suggests a novel mechanism for pulmonary arterial remodeling.

Another member of cluster 1, lactotransferrin, is upregulated during development of pulmonary hypertension and vascular remodeling in our model (Fig. 5). Lactotransferrin is an iron-binding protein with diverse functions, including metal ion homeostasis, iron metabolism, regulation of inflammatory response, and protease activity, but has not been previously described in pulmonary hypertension or pulmonary vascular remodeling (8). Lactotransferrin enhances mobility of fibroblasts and resulted in collagen gel contractile activity in an in vitro model of ECM remodeling (38). Interestingly, lactotransferrin has been shown to bind to and dissociate IGFBP3 from IGF-1, implicating it in the regulation of the IGF signaling system (5). Furthermore, orally administered human lactotransferrin has been recently demonstrated to promote VEGF-induced angiogenesis in rats (34). If the role of lactotransferrin is confirmed in pulmonary hypertension and vascular remodeling, this endogenous protein may be a potential target for intervention in this disease.

Cell adhesion was the most overrepresented biological process in our model of hypoxic pulmonary hypertension (Table 1). Candidate genes involved in cell adhesion were members of several clusters, suggesting complex expression patterns of their gene products. Several of these candidates were downregulated during both progression and regression of pulmonary hypertension and vascular remodeling (cluster 6), including α-catenin and IQGAP1. IQGAP1 is emerging as a key regulator of cell-cell adhesion, cytokinesis, and signal integration in various in vitro cell culture models (7). The expression of IQGAP1 is significantly decreased in our animal model of pulmonary hypertension (Fig. 5). If we postulate that IQGAP1 plays a novel role in pulmonary vascular remodeling, our findings are in contrast to in vitro studies that show overexpression of IQGAP1 enhances cell migration and motility (27). However, more recent data reveal a more complex role for IQGAP1, whereby suppression of IQGAP1 expression using small interfering RNA promoted cell-cell dissociation and scattering, a process likely to be important for vascular and ECM remodeling (33).

Our study has several limitations. By using gene expression profiling, we have only measured transcriptome levels, and our results need to be verified at the protein level. This is an important issue, since mRNA and protein levels are not necessarily tightly correlated, due to, for example, posttranscriptional modification of gene products. We have confirmed the differential expression of IGFBP3 at the protein level and plan to verify other candidate genes in future studies. Another limitation of our approach is the use of whole lungs to isolate RNA, precluding our ability to determine the anatomical location of differential gene expression. Although we observed that chronic hypoxia produces only modest vascular remodeling in Balb/c mice, the histopathological changes found in these lungs were limited to the pulmonary arteries and arterioles, suggesting that the vasculature is likely to be a dominant site for differential gene expression. An alternative method would have been to use laser capture microdissection (LCM) to obtain site-specific tissue (11). However, this method has several limitations including very low RNA yield, RNA degradation during tissue acquisition, and potential sampling bias due to limited tissue availability (28). Although LCM is a promising and improving technique, we believe that the whole organ approach is the most physiologically appropriate model to study pulmonary hypertension and vascular remodeling.

We decided to pool the lungs of hypoxic and control mice separately at each time point before performing microarray replications. Pooling allowed us to assess the total variability of gene expression (i.e., the sum of both technical and biological noise). Pooling has been widely used in microarray studies in mice and humans and may confer advantages for statistical analyses (1, 15, 22, 35). However, because we cannot account for animal-to-animal differences, our results may be biased due to the presence of an outlier animal. We addressed this problem by using relatively large numbers of mice at each time point (5 experiment and 5 control for a total of 90 mice), thereby reducing the potential contribution of an outlier. Furthermore, by obtaining gene expression profiles during nine time points, we were able to focus our attention on genes differentially expressed during at least two time points, further minimizing outlier effects.

A potential drawback of our experiment design is our inability to distinguish genes differentially expressed because of hypoxia from those due solely to pulmonary hypertension. However, in several clinical conditions (e.g., chronic obstructive pulmonary disease, sleep apnea, chronic mountain sickness), hypoxia is an important cause of pulmonary hypertension and both conditions coexist; therefore, hypoxia-induced chronic pulmonary hypertension is a physiologically relevant animal model (31).

We are confident in our statistical approach to determine significant differential gene expression. We performed four replications with dye swapping at each time point, addressed the multiple comparisons problem by limiting the FDR to 5%, and performed sensitivity analyses to confirm the robustness of our approach. Because all our candidate genes met these criteria during at least two time points, the actual statistical confidence in their differential expression is likely to be much higher. We confirmed differential gene expression for four candidates using qRT-PCR and for one candidate at the protein level. We utilized a recently developed clustering method, partitioning around medoids, to classify temporal expression patterns of candidate genes into biologically meaningful clusters. This algorithm has not been previously applied to time-series gene expression data and is a more robust alternative to the k-means clustering technique.

Transcriptional profiling has been used extensively as a “hypothesis generating” tool and as a method to classify different disease entities and even predict survival (42). Important clinical and molecular insights have been obtained, for example, by utilizing microarrays to study primary pulmonary hypertension in humans (14). However, a powerful feature of gene expression analysis is its ability to obtain large-scale “snapshots” of entire biological processes. We have combined
transcriptional profiling with rigorous statistical methods and global gene annotation analysis to gain unique insights into a complex physiological process and to identify the key biological modules involved in the development and regression of hypoxic pulmonary hypertension. By clustering genes based on their temporal expression and annotating individual clusters, we were able to integrate statistical classification of genes with distinct biological functions. This novel methodology designates putative roles for thousands of candidate genes in the context of their temporal expression patterns and memberships in specific biological modules. Our systematic and unbiased approach confirmed several molecular pathways and gene products known to be important in hypoxic pulmonary hypertension. Furthermore, we discovered several novel candidates and molecular mechanisms that may play crucial roles in hypoxia-induced pulmonary hypertension and vascular remodeling. The entire database generated by our study has been made publicly available at our website (http://fmrc.pulmcce.

washington.edu/microarray.html) and at the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/projects/geo/, query GSE 1909) to allow independent corroboration of our results and facilitate further research into pulmonary hypertension.

GRANTS

This work was supported by National Heart, Lung, and Blood Institute Grants K08-HL-74223 (S. A. Gharib), R01-HL-63994 (D. K. Madtes), and R01-HL-56239 (R. W. Glenny; the Royalty Research Fund, University of Washington (R. W. Glenny, S. A. Gharib); and the Charles E. Culpeper Biomedical Pilot Initiative, Rockefeller Brothers Fund (R. W. Glenny, S. A. Gharib).

REFERENCES


