Expression Profiling of the Developing Mouse Lung

Insights into the Establishment of the Extracellular Matrix

Thomas J. Mariani, Jeremy J. Reed, and Steven D. Shapiro

Departments of Pediatrics, Cell Biology and Physiology, Internal Medicine, and the Program in Lung Development, Washington University School of Medicine and Saint Louis Children's Hospital, St. Louis, Missouri

We have undertaken a comprehensive gene expression profiling of the entire process of murine lung development using oligonucleotide-based microarrays. Our data reveals the expression pattern of ~ 11,000 genes throughout the morphologic stages of lung development. This includes known genes with unappreciated pulmonary expression and novel genes with undefined functions. Traditional gene expression analysis techniques verify a high degree of confidence in the microarray data. Examination of the data confirms previously known patterns of expression for extracellular matrix genes and provides new information regarding relationships in temporal expression among groups of these genes. Large-scale cluster analysis reveals associations in the expression profile of specific genes with defined developmental processes. For instance, we identify groups of genes, which are coordinately expressed with extracellular matrix genes during lung development. These data should serve as a resource for the pulmonary research community and assist in deciphering the molecular mechanisms governing normal lung development as well as those involved in aberrant developmental pathology.

Mammalian lung development is a complex morphogenetic process, which initiates near mid-gestation and continues through early postnatal life. The lung arises as two lateral buds that emerge from the ventral foregut endoderm at ~ 9 days after fertilization (in mouse) and undergo numerous rounds of dichotomous branching to form the bronchial tree. This stage of development is referred to as the pseudoglandular phase, histologically characterized by loose mesenchyme surrounding undifferentiated epithelial tubes. Recent advances have been made in the understanding of the genetic and molecular mechanisms involved in these early processes which, in part, exhibit direct analogies to vertebrate limb development (1, 2). For instance, lung structure initiation and branching morphogenesis involve coordinated regulation of molecular pathways, including Sonic Hedgehog–, Bone Morphogenetic Protein/Transforming Growth Factor-β–, Fibroblast Growth Factor– and Retinoic Acid–related signaling (3). There are at least three subsequent phases of lung development that can be distinguished morphologically; the canalicular, saccular, and alveolar stages. A detailed understanding of the physiologic processes governing these latter stages of lung development, in particular the alveolar stage that results in the development of terminal airways and airsacs capable of functional gas-exchange, remains incompletely defined.

Large-scale gene expression analysis provides a powerful tool that can aid in the understanding of the molecular status of a cell, tissue, or organ (4–7). This technique has been used to discriminate static differences between cells/tissues and to compare changes in gene expression over time. In the former case, gene expression profiling provides data that can improve the description and our comprehension of a particular specimen. For instance, certain morphologically indistinguishable tumors can be discriminated on the basis of gene expression analysis data (6). In the latter case, mathematical clustering techniques have proven useful for identifying functional inter-relationships between genes based upon their expression patterns (8–10).

We have undertaken a large-scale gene expression analysis of mammalian lung development using oligonucleotide-based microarrays. We have chosen to study the mouse as our model system, based upon future applications of our current data set to genetically engineered animal models. Our analysis encompasses all recognized stages of development beginning at embryonic day 12 and continuing to adulthood. These data provide an abundance of information for gene expression during lung development, including cloned genes with previously unappreciated lung expression and expressed sequence tags (ESTs) with undefined function. Additionally, cluster analysis reveals previously unappreciated relationships between the expression patterns of genes that may have functional significance in this complex process.

Materials and Methods

Isolation of Lung RNA

Timed-pregnant Swiss-Webster mice were purchased from Taconic (Charles River, MA). Lungs were isolated by manual dissection with the aid of a dissecting microscope, where necessary (E12–E18). Stage of embryonic lung development was confirmed by visual inspection of the embryos and histologic evaluation of littermate lung sections. RNA was isolated from individual (E18, P1, P4, P7, P10, P14, P21, and adult) or pooled (E12, E14, E16) whole lungs using a modified guanidinium:phenol extraction method (Trizol; Invitrogen, Carlsbad, CA) according to manufacturer's instructions. RNA quality and integrity was confirmed by denaturing gel electrophoresis.

Generation of cRNA "Target" and Chip Hybridization

Ten micrograms of total RNA was used to generate "target" cRNA for hybridization to Affymetrix Mu11K chipset subA and subB oligonucleotide microarrays. A single chip set was used for each time-point. Before target generation, individual RNA samples were pooled so that each "target" was derived from a minimum of three individual lungs. This step was performed to minimize both biologic variability and the number of microarrays necessary to generate informative data. After pooling, RNA was re-purified using Qiagen purification kits (Valencia, CA). Target

cRNA was synthesized according to manufacturer's protocol. Briefly, total RNA was reverse transcribed into a complete cDNA library (SuperScript Choice kit; Invitrogen) using a T7/T$_{24}$ primer (Genset Corp., La Jolla, CA). This cDNA was subsequently *in vitro* transcribed into biotin-labeled cRNA (High Yield RNA Transcript Labeling kit; Enzo Diagnostics, Farmingdale, NY). The resulting cRNA was fragmented and assessed for quality by gel electrophoresis. Twenty-five micrograms of each biotin-labeled cRNA was hybridized to an Affymetrix Mu11K sub A and B chip according to the manufacturer's protocol. Four biotinylated hybridization controls (BioB, BioC, BioD, and Cre) were included in each hybridization reaction to verify consistent hybridization efficiency.

Microarray Data Analysis

Chips were scanned and the intensity of signal was computed using GeneChip Analysis Suite version 3.3 (Affymetrix, Santa Clara, CA). All chip intensities were scaled to an average of 1,500 units for comparison (inter-experiment scaling) purposes. All data values for each probe set were saved as a flat file for further analysis using GeneSpring software from Silicon Genetics (Redwood City, CA). Cluster analysis was performed using normalized average difference values using the standard correlation. The data for each probe set was normalized independently, by assigning the median average difference value for the probe set to 1 and representing all other values for that probe set as a ratio to that median average difference value. Importantly, multiple probe sets interrogating the same gene were normalized independently. For cluster analysis, all probe sets were included without filtering. Normality of distribution for the data set was verified by comparing mean average difference values, number of genes called "present" and number of genes increasing/decreasing at each time-point (*see below*). A complete catalog of this expression data, organized by GenBank accession number, and including expression level values for each probe set at each time-point, is available (http://ilya.wustl.edu/~mga/).

Northern Blot Analysis

Northern blot analysis was performed as previously described (11, 12). Briefly, total RNA from individual samples, the same samples used for microarray analysis before pooling, was used. A quantity of 5 μg RNA was denatured in a formamide/formaldehyde loading buffer, separated by denaturing agarose gel electrophoresis, transferred to a nylon membrane, and hybridized with radiolabeled probes. cDNA clones used as probes were derived from the IMAGE library and purchased from Incyte Genomics (St. Louis, MO).

Results

Validation

A complete database containing all the information related to these experiments will be made available (http://ilya.wustl.edu/~mga/). To assess normality, we analyzed some global parameters of our data set (Table 1). First, the number of genes scored as expressed by the Affymetrix software was within the expected range (maximal at E12 = 48%, minimal at P1 = 40%), did not fluctuate significantly and showed no correlation to changes in the range of relative expression values. Additionally, the number of genes increasing at any given time point were nearly equal, as expected in a random data set (maximal at Adult = 54.7%, minimal at P21 = 46.6%).

We investigated the reliability of the microarray data by candidate gene validation using Northern blot hybridization (Figure 1). Candidate genes chosen for validation included a constitutive house-keeping gene (glyceraldehyde-3-phosphate dehydrogenase), structural protein genes known to be highly expressed in the lung (tropoelastin), and genes with previously undefined lung expression patterns (ADAMTS-1). These analyses confirmed a high degree of concordance between oligonucleotide array results and conventional expression analysis results, providing a high degree of confidence in our overall data set. Additionally, these results indicated a moderate degree of biologic variability, thus reinforcing our experimental strategy.

Due to our interest in the processes that control establishment of lung structure, we focused on the expression of genes known to encode components of the extracellular matrix. Of particular relevance to pulmonary structure and physiology are collagen and elastin. Elastin provides passive recoil to the lung, which is essential for expiration and tissue integrity. Elastogenesis is a highly regulated developmental process that is believed to peak at birth and decline through maturation with unappreciable tropoelastin mRNA expression in the adult human (13) and rodent (14–16). The microarray profile for tropoelastin expression in the developing mouse lung, confirmed by Northern blot analysis, revealed two distinct peaks of maximal expression, at embryonic day 18 and at postnatal days 10–14 (Figure 1C). It is likely that these separate peaks reflect separate compartments of the lung that become elastogenic at different stages, the late embryonic peak representing the vascular compartment and the neonatal peak representing the alveolar interstitial compartment. Although differences in the expression pattern in these two lung structures have been previously appreciated, the temporal distinction in elastogenesis between these two phases (vascular and alveolar intersitium) has not been described. An analysis of smooth muscle α-actin gene expression, a marker for elastogenic cells in the lung (17), showed a biphasic profile similar to but temporally preceding tropoelastin expression. This reflects the establishment of an elastogenic, myofibroblast cell population in the lung before the onset of elastogenesis itself (18, 19).

Collagens provide tensile strength and some are major components of the basement membrane. Simple visualization of the expression patterns of genes for the interstitial collagens (type I and type III) and the basement membrane collagens (type IV) revealed previously unappreciated
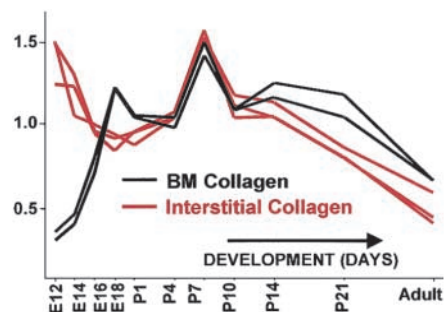
TABLE 1

Microarray data characteristics

Stage	E12	E14	E16	E18	P1	P4	P7	P10	P14	P21	Adult
# Present	5,173	5,407	5,822	5,294	6,327	6,052	6,016	5,797	5,910	5,348	5,802
Mean Avg. Diff.	2,151	1,967	1,900	1,892	1,822	1,830	1,946	1,851	1,842	1,828	1,762
% Increasing		51.1	52.1	48.5	52.9	51.9	48.5	49.1	49.6	46.6	54.7
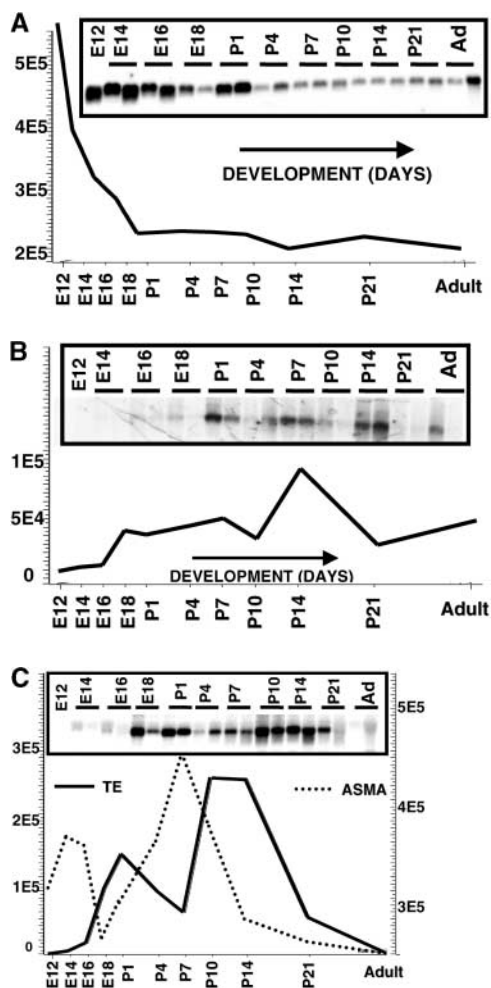
Figure 1. Candidate gene validation of expression. Equal amounts of total RNA were examined for expression by Northern blot analysis (*insets*) of (*A*) glyceraldehyde-3-phosphate dehydrogenase, (*B*) ADAMTS-1, and (*C*) tropoelastin, and compared with results from microarrays (*graphs*). The expression pattern observed for these genes are very comparable to the microarray profile, supporting a high degree of confidence in the microarray data set. Additionally, the expression profile for α-smooth-muscle actin is presented in *C*. This protein is specifically expressed in elastogenic cells and its expression precedes tropoelastin, suggesting that this cell population is established just before elastogenesis itself.

Figure 2. Expression profiles of interstitial versus basement membrane collagens. Shown are the normalized average difference values generated for a subset of interstitial collagen (α1[I], α1[III]) and basement membrane collagen (α1[IV], α2[IV]) genes. Note the similarity in expression patterns within the groups of interstitial collagens and basement membrane collagens. These data support the hypothesis that genes sharing function show similar expression profiles. Importantly, there is a considerable difference in the expression patterns between these groups of genes, supporting the ability to discriminate function based upon expression patterns.

patterns and relationships (Figure 2). Interstitial collagen genes were expressed in early embryonic lung development and were expressed at maximal levels postnatal day 7 (P7), coincident with alveogenesis. Basement membrane collagen gene expression (type IV) was lowest at day 12 of embryonic gestation (E12), and showed two peaks, at E18 and again at P7. These profiles reflect structural requirements of the developing lung, with interstitial collagens likely needed to establish the lung structure in early development and again postnatally for elongation of alveolar septae. Conversely, basement membrane collagens are needed for cellular compartmentalization (endothelial:medial, epithelial:mesenchymal) during later embryonic development and again for vascularization of the gas-exchange units during alveogenesis.

Importantly, the groups of functionally related collagen genes (Types I and III versus Type IV) showed similar patterns of expression within each group.

Hierarchical Clustering

Clustering techniques were applied to determine functional relationships among genes. Hierarchical (agglomerative) clustering was used to generate a dendrogram revealing the relationship between the profiles of each gene with every other gene in the data set (20). Analysis of this dendrogram (Figure 3A) revealed a group of six probe sets with similar profiles (most highly correlated), including two elements interrogating α1(I) procollagen and one for α1(III) procollagen. The adjacent location of the two genetic elements probing expression of α1(I) procollagen serves as an internal control for the technology and validates the utility of the mathematical clustering procedure. Furthermore, the juxtaposition of type I and type III collagen genes supports a functional relationship between proximal members in the dendrogram. In addition to the collagen genes, this group included the transcription factor COUP-TF1 (21) and the P311 gene (22), the developmental expression of which have not previously been characterized in the lung. These data point out a novel relationship between the expression of a transcription factor (COUP-TF1) and interstitial matrix genes in the developing lung.

In the same dendrogram, another group of three probe sets shared the most highly correlated expression profile (Figure 3B). Two of these genes encode primary components of the basement membrane, α1(IV) and α2(IV) procollagen. The additional transcript in this group was microtubule-actin cross-linking factor (MACF), a cytoskeletal component known to be expressed in the lung (23). Again, the close proximity of these two basement membrane-specific collagen genes, which are functionally related, supports the utility of these methods. In fact, in this case the data reveal the expected correlation of expression of these two collagen molecules that physically in-
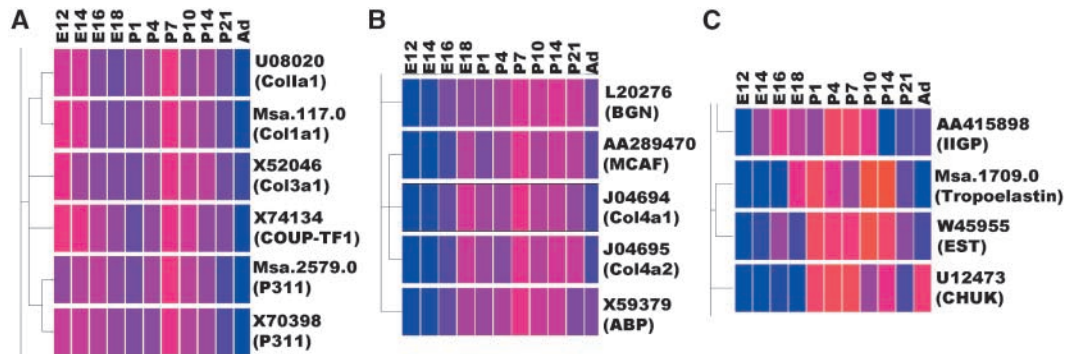
Figure 3. Hierarchical clustering. The entire expression data set was subjected to hierarchical clustering using the standard correlation. In this view, data for each probe set is displayed on the horizontal axis with individual time-points (E12–Adult) represented by *colored boxes*. The color of the box represents the normalized level of expression relative to the median expression for that probe set, with data points above the median shaded *red* and data points below the median shaded *blue*. (*A*) Within the dendrogram generated for the 13,069 probe sets, a group of six probe sets are shown which have the highest correlation of expression. These six probe sets represent four genes, with two genes interrogated by duplicate probe sets. Note that the probe sets showing the highest degree of correlation in the dendrogram interrogate the same gene (α1[I] procollagen or P311). Two functionally related molecules, α1(I) and α1(III) procollagen, cluster together into a node representing 0.045% (6/13,069) of the data set, along with the transcription factor COUP-TF1. (*B*) A separate group of three probe sets are shown to have the highest correlation in expression patterns. Of these three, two interrogate α1(IV) and α2 (IV) procollagen, molecules that physically interact for form type IV collagen heterotrimers. (*C*) The probe set interrogating tropoelastin is most highly correlated with an EST having no similarity to any sequence in GenBank.

teract with one another to form a heterotrimeric type IV collagen. The relationship between the collagen genes and a cytoskeletal component is presently unclear. Additionally, in this dendrogram, tropoelastin shared the highest correlation with a single probe set encoding an EST with no similarity to any sequences in GenBank (Figure 3C).
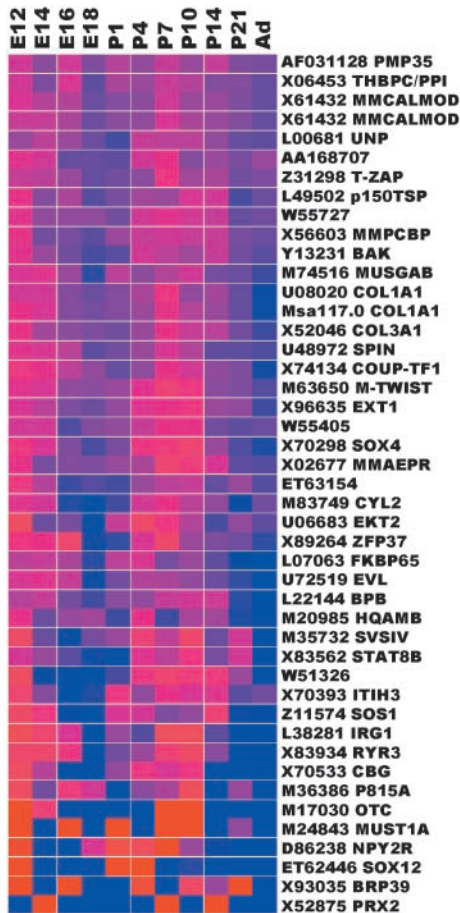
K-Means Clustering

Divisive clustering techniques were applied to determine functional relationships within groups of genes (24) by dividing the entire data set into smaller groups of genes that share a high degree of similarity in expression over time. K-means cluster analysis was used to generate 50 clusters of probe sets with highly similar expression profiles. The clusters contained a median value of 236 probe sets, with a range between 861 and 63. Focusing on the regulation of extracellular matrix corroborated the hierarchical clustering methods and revealed relationships among and between extracellular matrix genes and regulatory molecules.

Again, α1(I) procollagen, α1(III) procollagen and COUP-TF1 co-clustered into a group of 124 probe sets, further supporting a functional relationship between these genes (Figure 4). Additional members of this cluster included: prolyl-4-hydroxylase β subunit, a molecule involved in post-translational collagen modification (25); the immunophillin FKBP65, an ER-localized extracellular matrix–binding protein with peptidyl-prolyl cis-trans isomerase activity (26, 27); ext1, an ER-localized protein responsible for the development of hereditary multiple exotoses, a benign tumor with precocious bone growth (28, 29); Prx2, a homeobox gene previously believed to play a role in vascular extracellular matrix regulation (30, 31); m-Twist, a bHLH transcription factor which may regulate type II collagen gene expression (32); the Sox-4 and -12 transcription factors; CYL2 a D-type cyclin; calmodulin, a calcium-dependent protein kinase; and the RyR3 calcium channel.

In relation to elastogenesis in the developing lung, tropoelastin clustered with 182 probe sets (Figure 5), including TGF-β1 (TGF-1), TGF-β3, and platelet-derived growth factor receptor, molecules known to regulate the expression of tropoelastin (33). This cluster also included LTBP-2 and fibulin, two molecules known to bind to elastic fibers (34–37). Additional genes included the transcription factors Nkx-5.2, lung-specific Kruppel-like factor (LKLF), mesenchyme fork-head (MFH)-1, as well as FGF receptor-1, CRABP, MT2-MMP, and Type XV collagen. Thus, this cluster contains multiple molecules previously shown either to regulate tropoelastin gene expression, to be essential to lung elastogenesis, or to physically interact with elastin/elastic fibers.

Two separate K-means clusters contained distinct subsets of basement membrane collagens. One cluster of 154 probe sets included genes for the basement membrane components fibronectin, laminin α4, α5, and β2, and entactin (Figure 6). Additional genes in this cluster encoded the extracellular matrix molecules α1 and α2 (VI) procollagen and fibrillin-1. Also included were capsulin/cor-1/pod-1, a transcription factor essential for lung development (38), the alveolar epithelial cell–specific product SP-C, β1 integrin, h2-calponin STAT-6, and cyclin G2.

Another cluster of 298 probe sets (Figure 7) contained α1, α2, and α5 (IV) procollagen genes as well as laminin β3 and laminin α2/merosin. Additional genes in this cluster included a number of known extracellular matrix molecules such as perlecan, biglycan, SPARC, Gla/MGP, LTBP-3, lysyl oxidase 2 (a member of the gene family responsible for post-translational crosslinking of collagen molecules) (39), and caspin, a collagen-associated serpin (40). A number of vascular markers, including tie flk-1 and sm22α (a marker of vascular smooth muscle), were present in this group. Also included were osf-2/Cbfa1, a regulator of matrix production (41, 42), FGF receptor-3 and FGF receptor-4, β1 integrin, α-catenin, Mac25 (an activin-binding protein),
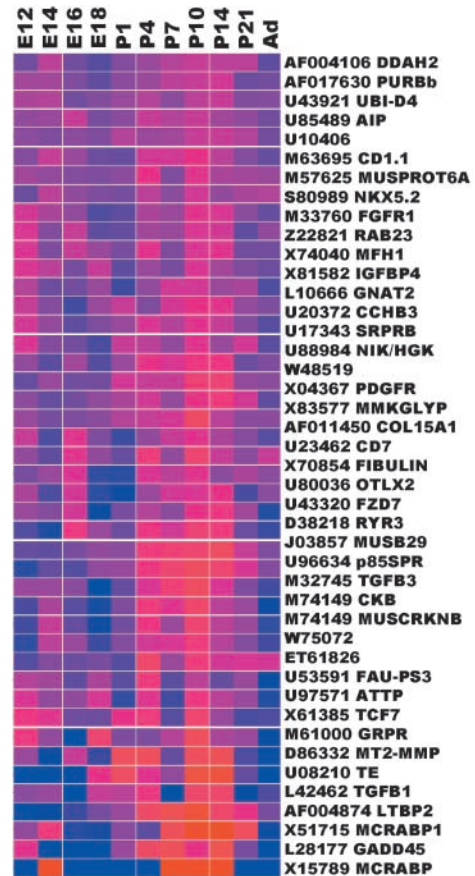
Figure 4. Interstitial collagen K-means cluster. K-means clustering was used to generate 50 groups of probe sets, each group showing a high correlation in expression. Shown is the cluster containing α1(I) and α1(III) procollagen. The ESTs have been removed from the list for visualization. This cluster which represents 0.9% (124/13,069) of the probe sets includes numerous extracellular matrix–related genes.

Figure 5. Tropoelastin K-means cluster. K-means clustering was used to generate 50 groups of probe sets, each group showing a high correlation in expression. Shown is the cluster containing tropoelastin. The ESTs have been removed from the list for visualization. This cluster which represents 1% (182/13,069) of the probe sets includes numerous elastogenesis-related genes.

NRC-1, and the lung epithelial cell products Clara cell secretory protein (CC10) and phosphatidylcholine-phospholipase D2 (PC-PLD2). The biologic explanation for partitioning of these two subsets of basement membrane protein genes is not readily apparent, and is of significant biologic interest. Additionally, the clustering of these structural genes with specific regulatory molecules suggests a functional, physiologic relationship in the developing lung.

Discussion

Experimental Design

We chose to combine samples from several individual animals and run a single array (set) for each experimental time point beginning day E12. The rationale for this decision is multifold. First, technical limitations, such as input mass of RNA and purity of isolated samples, restrained us from investigating the earliest events in lung development immediately subsequent to budding. With advances in target generation, these obstacles will soon be overcome. Second, in support of the single array per time point, quality control

tests of numerous arrays (replicate experiments) and experience with thousands of individual experiments indicate a reasonably low degree of technical variability (M. Watson, personal communication). Recent publications have reinforced this low degree of technical variability, placing the error rate near 0.3% (43) or 2% (44). Third, combining samples controls for the biologic variability, which is certain to be present. Thus, in the massive amount of data which would be generated from multiple arrays at each timepoint, significant changes in gene expression were likely to be masked by the biologic variability. We incorporated into the experimental design candidate gene validation, using conventional technology to verify our results. Our candidate gene validation revealed a high concordance between microarray data and conventional expression data, resulting in a high degree of confidence in our microarray data set. Importantly, duplicate probe sets that interrogate a single gene showed a very high degree of similarity (Figure 3 and data not shown), confirming technical accuracy.

Extracellular Matrix Gene Networks

Analysis of expression microarray data is complex and the data can be interpreted using many strategies. We have taken
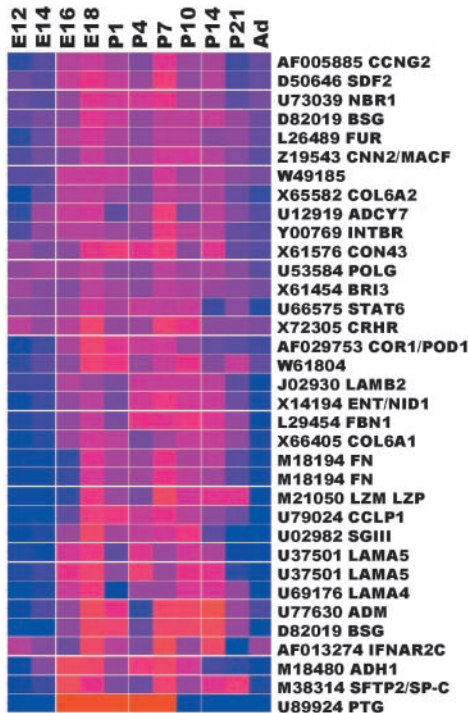
Figure 6. Fibronectin/entactin K-means cluster. K-means clustering was used to generate 50 groups of probe sets, each group showing a high correlation in expression. The ESTs have been removed from the list for visualization. Shown is the cluster representing 1% of the probe sets (154/13,069), which contains distinct basement membrane molecules including fibronectin, entactin, and laminin α4 and α5.

a stepwise approach, from simple examination to correlative mathematical evaluation, to the analysis of our data set. This evaluation strategy provides novel and interesting informa-

tion at each step, and highlights both the fidelity and the utility of the technology and the data set. First, simple assessment of single gene profiles reveals the developmental expression pattern of 11,000 genes and ESTs in the lung. This includes supportive information about genes with previously identified lung expression, like tropoelastin, and novel information about genes with unknown lung expression developmental profiles, like ADAMTS-1 (Figure 1). Second, simple examination of groups of genes with similar functions often shows close associations between their expression patterns. For instance, we find that multiple members of the collagen gene family show highly similar expression patterns, but can be divided by function (interstitial versus basement membrane components), based upon their profiles over time (Figure 2). Finally, agglomerative and divisive cluster analyses can be applied to our data set to distinguish groups of genes with related profiles and, at least in some instances, genes which are functionally (and physically) related. Hierarchical clustering indicated the highest degree of relatedness was found between the expression of 2 α chains of type IV collagen, or alternatively between α1(I) and α1(III) procollagen (Figure 3). K-means clustering mapped some basement membrane-related genes together, yet discriminated divisions between other basement membrane genes (Figures 4–7). The physiologic significance of these groupings is unexpected, but of great interest. Finally, these clustering methods have provided putative functional and/or regulatory relationships between genes. For instance, the closely related profiles between tropoelastin and LTBP-2, as evidenced by cluster analysis, reinforce the proposed physical interaction between these two molecules within developing elastic fibers (34, 37). Similarly, the inclusion in this cluster of the transcription factor essential to lung development, LKLF (45), suggests a potential target for this regulatory molecule and a mechanism by which its necessity for lung development is revealed.
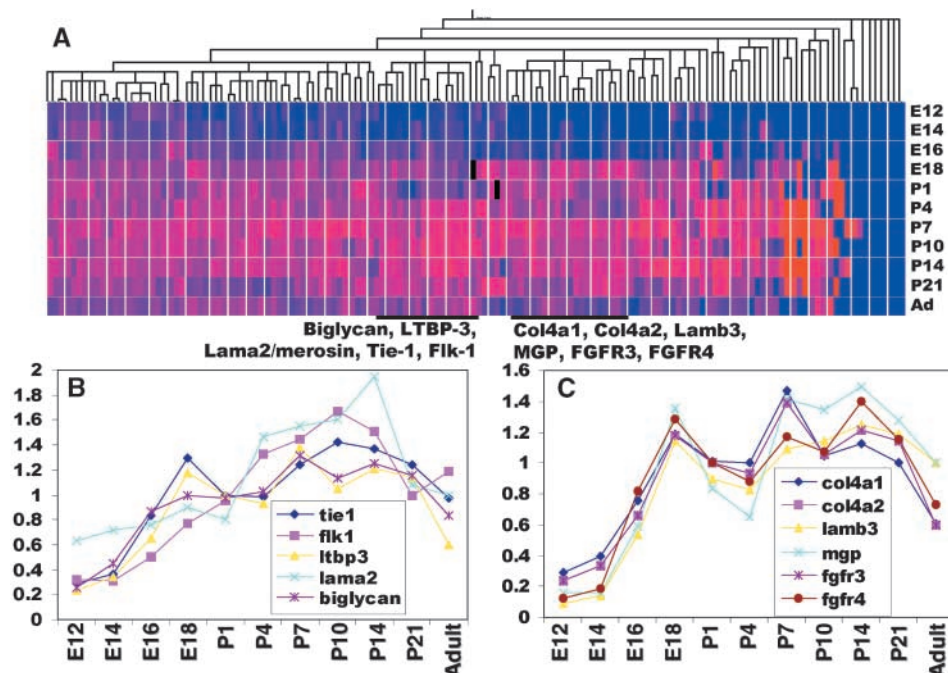

Figure 7. Type IV collagen K-means cluster. K-means clustering was used to generate 50 groups of probe sets, each group showing a high correlation in expression. This cluster which represents 2% (298/13,069) of the probe sets includes numerous basement membrane related genes including α1 (IV) and α2 (IV) procollagen. The ESTs have been removed from the list for visualization. All probe sets in the cluster are displayed in a dendrogram with relative similarity in expression based upon hierarchical clustering (*A*). Individual profiles for two subsets within this cluster, which share a high degree of similarity, are also shown. One group of 16 probe sets contained many vascular-related genes (*B*). Another group of 21 probe sets contained type IV collagen genes and the genes for two FGF receptors implicated in alveogenesis (*C*).

The cluster analysis also reveals a previously unidentified relationship between FGFR gene expression and basement membrane gene expression during lung development. These data suggest that FGF signaling, and specifically FGFR3 and FGFR4, are involved in the regulation of basement membrane formation in the lung. Interestingly, the compound mutation of these two genes leads to the failure of terminal lung development. Our data suggest a potential mechanism for this previously unexplained developmental defect. This hypothesis is supported by recent work showing that FGF signaling is capable of regulating morphogenesis via regulation of basement membrane production (46).

Certainly, the interpretation of expression profiling data from a complex tissue or organ, such as the lung, is subject to changes in the relative cell populations within that sample. This has been previously addressed for microarray data generated from adult lung tissue (5). These authors point out that this fact can provide insight into the nature of the changes in the cell population of one's sample, but also could confound one's interpretation of the data. For instance, it is possible that the number of extracellular matrix genes expressed at or near a maximum at P7 could reflect an increase in the relative number of fibroblastic cells within the lung or an increase in the metabolic activity of these cells at that time point. Although this fact increases the complexity of data interpretation, it does not limit the validity or utility of the data, for individual genes or mathematically related groups of genes. The consistent correlation of functionally related extracellular matrix genes in our data set supports this conclusion. For example, $\alpha 1$ (IV) and $\alpha 2$ (IV) procollagen genes, which encode proteins that form subunits of a single heterotrimeric molecule, showed the most highly correlated expression pattern within a dynamic cell population.

Applications for Database

This data (available at http://ilya.wustl.edu/~mga/) will be useful as a general resource to the pulmonary community for the identification and confirmation of gene expression patterns. This database includes the initial identification of literally hundreds of genes expressed during lung development, including genes with previously defined functions and those without any defined function or identity (ESTs). Using manual and mathematical analyses, the data set can be used to derive putative functions for genes involved in lung development. Such analyses can also be used to construct molecular maps of biochemical pathways essential for the development of this organ. Finally, this data set will serve as the initial phase of a long-term project to identify comprehensive molecular alterations associated with defined lung developmental abnormalities, and potentially, abnormal disease processes in the adult lung.

References

1. Warburton, D., and M. Lee. 1999. Current concepts on lung development. *Curr. Opin. Pediatr.* 11:188–192.
2. Warburton, D., M. Schwarz, D. Tefft, G. Flores-Delgado, K. Anderson, and W. Cardoso. 2000. The molecular basis of lung morphogenesis. *Mech. Dev.* 92:55–81.
3. Cardoso, W. 2000. Lung morphogenesis revisited: old facts, current ideas. *Dev. Dyn.* 219:121–130.
4. Iyer, V., M. Eisen, D. Ross, G. Schuler, T. Moore, J. Lee, J. Trent, L. Staudt, J. Hudson, M. Boguski, D. Lashkari, D. Shalon, D. Botstein, and P. Brown. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science* 283:83–87.
5. Kaminski, N., J. Allard, J. Pittet, F. Zuo, M. Griffiths, D. Morris, X. Huang, D. Sheppard, and R. Heller. 2000. Global analysis of gene expression in pulmonary fibrosis reveals distinct programs regulating lung inflammation and fibrosis. *Proc. Natl. Acad. Sci. USA* 97:1778–1783.
6. Golub, T., D. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. Mesirov, H. Coller, M. Loh, J. Downing, M. Caligiuri, C. Bloomfield, and E. Lander. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531–537.
7. Sherlock, G. 2000. Analysis of large-scale gene expression data. *Curr. Opin. Immunol.* 12:201–205.
8. Hill, A., C. Hunter, B. Tsung, G. Tucker-Kellog, and E. Brown. 2000. Genomic analysis of gene expression in *C. Elegans*. *Science* 290:809–812.
9. Spellman, P., G. Sherlock, M. Zhang, V. Iyer, K. Anders, M. Eisen, P. Brown, D. Botstein, and B. Futcher. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol. Biol. Cell* 9:3273–3297.
10. White, K., S. Rifkin, P. Hurban, and D. Hogness. 1999. Microarray analysis of Drosophila development during metamorphosis. *Science* 286:2179–2184.
11. Mariani, T., S. Sandefur, J. Roby, and R. Pierce. 1998. Collagenase-3 induction in rat lung fibroblasts requires the synergistic effects of TNFa and 12-lipoxygenase metabolites: A model of macrophage-induced fibroblast driven extracellular matrix remodeling during inflammatory lung injury. *Mol. Biol. Cell* 9:1411–1424.
12. Mariani, T., S. Dunsmore, Q. Li, X. Ye, and R. Pierce. 1998. Regulation of lung fibroblast tropoelastin expression by alveolar epithelial cells. *Am. J. Physiol.* 274(1, Pt. 1):L47–L57.
13. Mariani, T., S. Sandefur, and R. Pierce. 1997. Elastin in lung development. *Exp. Lung Res.* 23:131–145.
14. Pierce, R., M. Mariencheck, S. Sandefur, E. Crouch, and W. Parks. 1995. Glucocorticoids upregulate tropoelastin expression during late stages of fetal lung development. *Am. J. Physiol.* 268:L491–L500.
15. Noguchi, A., and H. Samaha. 1991. Developmental changes in tropoelastin gene expression in the rat lung studied by in situ hybridization. *Am. J. Respir. Cell Mol. Biol.* 5:571–578.
16. Wendel, D., D. Taylor, K. Albertine, M. Keating, and D. Li. 2000. Impaired distal airway development in mice lacking elastin. *Am. J. Respir. Cell Mol. Biol.* 23:320–326.
17. Mariani, T., E. Crouch, J. Roby, B. Starcher, and R. Pierce. 1995. Increased elastin production in experimental granulomatous lung disease. *Am. J. Pathol.* 147:988–1000.
18. Bostrom, H., K. Willetts, M. Pekny, P. Leveen, P. Lindahl, H. Hedstrand, M. Pekna, M. Hellstrom, S. Gebre-Medhin, M. Schalling, M. Nilsson, S. Kurland, J. Tornell, J. Heath, and C. Betsholtz. 1996. PDGF-A signaling is a critical event in lung alveolar myofibroblast development and alveogenesis. *Cell* 85:863–873.
19. Lindahl, P., L. Karlsson, M. Hellstrom, S. Gebre-Medhin, K. Willetts, J. Heath, and C. Betsholtz. 1997. Alveogenesis failure in PDGF-A-deficient mice is coupled to lack of distal spreading of alveolar smooth muscle cell progenitors during lung development. *Development* 124:3943–3953.
20. Eisen, M., P. Spellman, P. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95:14863–14868.
21. Zhou, C., S. Tsai, and M. Tsai. 2000. From apoptosis to angiogenesis: new insights into the roles of nuclear orphan receptors, chicken ovalbumin upstream promoter-transcription factors, during development. *Biochim. Biophys. Acta* 1470:M63–M68.
22. Taylor, G., E. Hudson, J. Reseau, and G. V. Woude. 2000. Regulation of P311 expression by Met-hepatocyte growth factor/scatter factor and the ubiquitin/proteasome system. *J. Biol. Chem.* 275:4215–4219.
23. Bernier, G., M. Pool, M. Kilcup, J. Alfoldi, Y. D. Repentigny, and R. Kothary. 2000. Acf7 (MACF) is an actin and microtubule linker protein whose expression predominates in neural, muscle, and lung development. *Dev. Dyn.* 219:216–225.
24. Alon, U., N. Barkai, D. Notterman, K. Gish, S. Ybarra, D. Mack, and A. Levine. 1999. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc. Natl. Acad. Sci. USA* 96:6745–6750.
25. Gong, Q., T. Fukuda, C. Parkinson, and S. Cheng. 1988. Nucleotide sequence of a full-length cDNA clone encoding a mouse cellular thyroid hormone binding protein (p55) that is homologous to protein disulfide isomerase and the beta-subunit of prolyl-4-hydroxylase. *Nucleic Acids Res.* 16:1203.

26. Patterson, C., T. Schaub, E. Coleman, and E. Davis. 2000. Developmental regulation of FKBP65. An ER-localized extracellular matrix binding-protein. *Mol. Biol. Cell* 11:3925–3935.
27. Davis, E., T. Broekelmann, Y. Ozawa, and R. Mecham. 1998. Identification of tropoelastin as a ligand for the 65-kD FK506-binding protein, FKBP65, in the secretory pathway. *J. Cell Biol.* 140:295–303.
28. Ahn, J., H. Ludecke, S. Lindow, W. Horton, B. Lee, M. Wagner, B. Horsthemke, and D. Wells. 1995. Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (EXT1). *Nat. Genet.* 11:137–143.
29. McCormick, C., Y. Leduc, D. Martindale, K. Mattison, L. Esford, A. Dyer, and F. Tufaro. 1998. The putative tumour suppressor EXT1 alters the expression of cell-surface heparan sulfate. *Nat. Genet.* 19:158–161.
30. Bergwerff, M., A. Gittenberger-de Groot, M. DeRuiter, L. van Iperen, F. Meijlink, and R. Poelmann. 1998. Patterns of paired-related homeobox genes PRX1 and PRX2 suggest involvement in matrix modulation in the developing chick vascular system. *Dev. Dyn.* 213:59–70.
31. Bergwerff, M., A. Gittenberger-de Groot, L. Wisse, M. DeRuiter, A. Wessels, J. Martin, E. Olson, and M. Kern. 2000. Loss of function of the Prx1 and Prx2 homeobox genes alters architecture of the great elastic arteries and ductus arteriosus. *Virchows Arch.* 436:12–19.
32. Fang, M., C. Glackin, A. Sadhu, and S. McDougall. 2001. Transcriptional regulation of alpha 2(I) collagen gene expression by fibroblast growth factor-2 in MC3T3–E1 osteoblast-like cells. *J. Cell. Biochem.* 80:550–559.
33. Mariani, T., and R. Pierce. 1999. Development of lung elastic matrix. *In* G. Gaultier, J. Bourbon, and M. Post, editors. Lung Development. Oxford University Press, New York. 28–45.
34. Gibson, M., G. Hatzinikolas, E. Davis, E. Baker, G. Sutherland, and R. Mecham. 1995. Bovine latent transforming growth factor beta 1-binding protein 2: molecular cloning, identification of tissue isoforms, and immunolocalization to elastin-associated microfibrils. *Mol. Cell. Biol.* 15:6932–6942.
35. Roark, E., D. Keene, C. Haudenschild, S. Godyna, C. Little, and W. Argraves. 1995. The association of human fibulin-1 with elastic fibers: an immunohistological, ultrastructural, and RNA study. *J. Histochem. Cytochem.* 43:401–411.
36. Sasaki, T., W. Gohring, N. Miosge, W. Abrams, J. Rosenbloom, and R. Timpl. 1999. Tropoelastin binding to fibulins, nidogen-2 and other extracellular matrix proteins. *FEBS Lett.* 460:280–284.
37. Shipley, J., R. Mecham, E. Maus, J. Bonadio, J. Rosenbloom, R. McCarthy, M. Baumann, C. Frankfater, F. Segade, and S. Shapiro. 2000. Developmental expression of latent transforming growth factor beta binding protein 2 and its requirement early in mouse development. *Mol. Cell. Biol.* 20:4879–4887.
38. Quaggin, S., L. Schwartz, S. Cui, P. Igarashi, J. Deimling, M. Post, and J. Rossant. 1999. The basic-helix-loop-helix protein pod1 is critically important for kidney and lung organogenesis. *Development* 126:5771–5783.
39. Smith-Mungo, L., and H. Kagan. 1998. Lysyl oxidase: properties, regulation and multiple functions in biology. *Matrix Biol.* 16:387–398.
40. Kozaki, K., O. Miyaishi, O. Koiwai, Y. Yasui, A. Kashiwai, Y. Nishikawa, S. Shimizu, and S. Saga. 1998. Isolation, purification, and characterization of a collagen-associated serpin, caspin, produced by murine colon adenocarcinoma cells. *J. Biol. Chem.* 273:15125–15130.
41. Lee, K., H. Kim, Q. Li, X. Chi, C. Ueta, T. Komori, J. Wozney, E. Kim, J. Choi, H. Ryoo, and S. Bae. 2000. Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. *Mol. Cell. Biol.* 20:8783–8792.
42. Kern, B., J. Shen, M. Starbuck, and G. Karsenty. 2001. Cbfa1 contributes to the osteoblast-specific expression of Type I collagen genes. *J. Biol. Chem.* 276:7101–7107.
43. Mody, M., Y. Cao, Z. Cui, K. Tay, A. Shyong, E. Shimizu, K. Pham, P. Schultz, D. Welsh, and J. Tsien. 2001. Genome-wide gene expression profiles of the developing mouse hippocampus. *Proc. Natl. Acad. Sci. USA* 98: 8862–8867.
44. Geraci, M., M. Moore, T. Gesell, M. Yeager, L. Alger, H. Golpon, B. Gao, J. Loyd, R. Tuder, and N. Voelkel. 2001. Gene expression patterns in the lungs of patients with primary pulmonary hypertension: a gene microarray analysis. *Circ. Res.* 88:555–562.
45. Wani, M., S. Wert, and J. Lingrel. 1999. Lung Kruppel-like factor, a zinc finger transcription factor, is essential for normal lung development. *J. Biol. Chem.* 274:21180–21185.
46. Li, X., Y. Chen, S. Scheele, E. Arman, R. Haffner-Krausz, P. Ekblom, and P. Lonai. 2001. Fibroblast growth factor signaling and basement membrane assembly are connected during epithelial morphogenesis of the embryoid body. *J. Cell Biol.* 153:811–822.