ORIGINAL RESEARCH

Cross-Species Transcriptome Profiling Identifies New Alveolar Epithelial Type I Cell–Specific Genes

Crystal N. Marconett[1,2,3*], Beiyun Zhou[3,4*], Mitsuhiro Sunohara[4], Tiffany M. Pouldar[1,2], Hongjun Wang[4], Yixin Liu[4], Megan E. Rieger[4], Evelyn Tran[1,2,3], Per Flodby[4], Kimberly D. Siegmund[5], Edward D. Crandall[4], Ite A. Laird-Offringa[1,2,3‡], and Zea Borok[2,3,4‡]

Departments of [1]Surgery and [2]Biochemistry and Molecular Medicine, [3]Norris Comprehensive Cancer Center, [4]Department of Medicine, Will Rogers Institute Pulmonary Research Center, Division of Pulmonary, Critical Care and Sleep Medicine, and [5]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California

ORCID IDs: 0000-0002-8463-7126 (C.N.M.); 0000-0002-9084-5933 (B.Z.); 0000-0001-8673-8177 (Z.B.).

Abstract

Diseases involving the distal lung alveolar epithelium include chronic obstructive pulmonary disease, idiopathic pulmonary fibrosis, and lung adenocarcinoma. Accurate labeling of specific cell types is critical for determining the contribution of each to the pathogenesis of these diseases. The distal lung alveolar epithelium is composed of two cell types, alveolar epithelial type 1 (AT1) and type 2 (AT2) cells. Although cell type–specific markers, most prominently surfactant protein C, have allowed detailed lineage tracing studies of AT2 cell differentiation and the cells' roles in disease, studies of AT1 cells have been hampered by a lack of genes with expression unique to AT1 cells. In this study, we performed genome-wide expression profiling of multiple rat organs together with purified rat AT2, AT1, and *in vitro* differentiated AT1-like cells, resulting in the identification of 54 candidate AT1 cell markers. Cross-referencing with genes up-regulated in human *in vitro* differentiated AT1-like cells narrowed the potential list to 18 candidate genes. Testing the top four candidate genes at RNA and protein levels revealed GRAM domain 2 (GRAMD2), a protein of unknown function, as highly specific to AT1 cells. RNA sequencing (RNAseq) confirmed that *GRAMD2* is transcriptionally silent in human AT2 cells. Immunofluorescence verified that GRAMD2 expression is restricted to the plasma membrane of AT1 cells and is not expressed in bronchial epithelial cells, whereas reverse transcription–polymerase chain reaction confirmed that it is not expressed in endothelial cells. Using *GRAMD2* as a new AT1 cell–specific gene will enhance AT1 cell isolation, the investigation of alveolar epithelial cell differentiation potential, and the contribution of AT1 cells to distal lung diseases.

Keywords: alveolar epithelium; type I cell markers; microarray analysis; lung differentiation; transcriptome analysis

Clinical Relevance

This research identifies GRAM domain 2 as a novel marker of alveolar epithelial type I cells. With this knowledge, the field can develop effective tools to understand what role these cells play in diseases of the distal alveolar epithelium.

(Received in original form February 23, 2016; accepted in final form September 27, 2016)

*These authors contributed equally to this work.

‡These authors contributed equally to this work.

This work was supported by the Hastings and Whittier Foundations; National Institutes of Health (NIH) research grants R01HL114094 (I.A.L.-O. and Z.B.), R37HL062569, R01HL112638, and R01HL126877 (Z.B.), R01HL114959 (B.Z.), and U01HL108634 (E.D.C.); Norris Comprehensive Cancer Center NIH core grant P30CA0189; the Cell and Tissue Imaging Core of the University of Southern California Research Center for Liver Diseases (NIH P30 DK048522); and American Cancer Society/Canary postdoctoral fellowship No. PFTED-10-207-01-SIED (C.N.M.). E.D.C. is Hastings Professor and K.T.N. Chair of Medicine. Z.B. is Edgington Chair in Medicine.

Author Contributions: C.N.M., B.Z., P.F., E.D.C., I.A.L.-O., and Z.B. contributed to the design of the research; C.N.M., M.S., T.M.P., H.W., Y.L., M.E.R., and E.T. performed the experiments; C.N.M. and K.D.S. performed or assisted with the data analysis; and C.N.M., B.Z., I.A.L.-O., and Z.B. wrote the manuscript.

Correspondence and requests for reprints should be addressed to Zea Borok, M.D., Will Rogers Institute Pulmonary Research Center, Division of Pulmonary, Critical Care and Sleep Medicine, Keck School of Medicine, University of Southern California, IRD 723 M/C 9520, 2020 Zonal Avenue, Los Angeles, CA 90089-9520. E-mail: zborok@med.usc.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Cell Mol Biol Vol 56, Iss 3, pp 310–321, Mar 2017

Originally Published in Press as DOI: 10.1165/rcmb.2016-0071OC on October 17, 2016
Internet address: www.atsjournals.org

Lung alveolar epithelium is composed of two morphologically distinct cell types, cuboidal surfactant-secreting alveolar epithelial type 2 (AT2) cells and delicate squamoid alveolar epithelial type 1 (AT1) cells with a large surface area that allows efficient gas exchange. AT2 cells uniquely express surfactant proteins, most notably surfactant protein C (SFTPC). Directed expression using the Cre-lox system has been developed for AT2 cells using the *SFTPC* promoter (1–3), allowing studies of differentiation and lineage tracing experiments. These have suggested a role for AT2 cells as the cell of origin for a number of lung diseases, including lung adenocarcinoma (4, 5). However, the contribution of AT1 cells has been studied less precisely because of the lack of uniquely specific AT1 cell markers. Differentiation of AT2 into AT1 and AT1-like cells has been characterized *in vivo* and, using freshly isolated AT2 cells, *in vitro* (3, 6–9). The accepted paradigm has been that although both AT2 and AT1 cells are differentiated cell types that serve specific functions in the alveolar epithelium, only AT1 cells are terminally differentiated (10). However, *in vitro* experiments suggest that AT1 cells can revert back to an AT2 cell-like phenotype (11–13). More recent experiments using the homeodomain-only protein homeobox (*Hopx*) gene as an AT1 cell marker also suggest that AT1 cells can revert to an AT2 cell phenotype *in vivo* (14), although the specificity of *Hopx* as a marker of AT1 cells remains uncertain because of its expression in cells with phenotypic characteristics intermediate between AT2 and AT1 cells (15). Additional AT1 cell–specific markers would facilitate more comprehensive analysis of AT1 cell differentiation potential *in vivo*.

Identification of proteins with highly specific expression in AT1 cells could also enhance purification and isolation strategies for further characterization of AT1 cell biology. Lack of absolute lung specificity of currently available AT1 cell markers limits their usefulness for AT1 cell–specific deletion of functional genes to assess their contribution to alveolar homeostasis. The best characterized of the existing AT1 cell markers, aquaporin 5 (AQP5), is a water channel protein expressed on the apical surface of AT1 cells (11). AQP5 has little to no expression in AT2 cells in rats and in at least some mouse strains and therefore serves as a relatively good marker of AT1 cell differentiation within the distal lung. Mouse models using *Aqp5* regulatory regions driving Cre-IRES-DsRed have been generated to develop AT1 cell–specific mouse genetic tools (16). However, because *Aqp5* is also expressed in salivary and lacrimal glands, as well as in the trachea (17), its utility as a Cre driver for AT1 cell–specific gene deletion can be complex. AQP5 is also expressed in human lymphocytes and dendritic cells (18), complicating *in vivo* analyses of AT1 cell responses to the immune stress and inflammation associated with lung disease and requiring strategies for negative selection to obtain pure AT1 cell populations.

Other known AT1 cell markers, including advanced glycosylation end product–specific receptor (AGER, previously RAGE) (19), podoplanin (PDPN, previously T1α) (20–23), caveolin 1 (CAV1) (24), and HOPX (14, 25), have been used to distinguish AT2 and AT1 cell populations through increased expression in AT1 cells. However, they are also expressed in a wide variety of other cell types and tissues including, in the case of PDPN, lung lymphatic endothelial cells (26). Recently, mouse alveoli have undergone single-cell RNA sequence (RNAseq) analysis revealing potential candidate markers of alveolar epithelial cell (AEC) identity (25), but molecular validation and characterization in human AT1 cells was not included, nor was expression at sites outside the lung addressed. In this study, we harnessed the discovery potential of whole genome transcriptional profiling of numerous purified AT2, AT1, and *in vitro*–derived AT1-like cells and compared that with genome-wide expression levels from 15 other organs. In addition, we harnessed a cross-species comparison between human and rat AT1 cells to identify new AT1 cell–specific genes. Analysis of multiple genome-wide datasets allowed us to refine candidate AT1 cell markers with application in multiple model species as well as in humans. These data led to the identification of GRAM domain 2 (*GRAMD2*) as a highly specific AT1 cell marker and also identified a second gene, *SCNN1G*, as specific to AT1 cells in the lung.

Materials and Methods

Cell Isolation

Rat AT2 and AT1 cells were isolated and differentiated *in vitro* as reported previously (27, 28), with modifications to AT1 cell isolation (*see* online supplement). Endothelial cell isolation is described (*see* online supplement). Human AT2 cells were isolated and differentiated as reported previously (29). Paraffin-embedded normal lung tissue was obtained from remnant human transplant lungs under University of Southern California Institutional Review Board Protocol No. HS-07–00660.

Microarray Analysis

RNA (1 μg) from each rat sample was sent for microarray profiling using Illumina RatRef12 at the Southern California Genotyping Consortium, University of California, Los Angeles. Human microarray data (HT12v4) were analyzed as reported previously (29). Linkage of human and rat gene expression profiles was performed using Entrez identifiers and databases of the Mouse Genome Informatics Web (The Jackson Laboratory, Bar Harbor, ME) (*see* online supplement).

RNAseq Analysis

RNA was isolated from $\sim 10^6$ human AT2 or from *in vitro*–differentiated AT1-like cells using the Illustra Triple Prep Kit (Catalog No. 28-9425-44; GE, Pittsburgh PA). RNA (2 μg) was made into libraries using the Illumina (Madison, WI) Ribo-Zero Gold Magnetic Kit (MRZG 12324) and sequenced using the Illumina HiSeq2000 at the University of Southern California Epigenome Center Core. Base calls were converted into FASTQ files and aligned to the hg19 genome using tophat2-2.0.8b.

Quantitative Polymerase Chain Reactions

RNA was subjected to reverse transcription using random hexamers and Moloney murine leukemia virus reverse transcriptase, followed by quantitative polymerase chain reaction (PCR) using SYBR green (BioRad, Hercules, CA). All reactions were performed using DNA engine Opticon (MJ Research, Waltham, MA) and were normalized to 18S levels as indicated (for primers used, *see* online supplement).

Western Analysis

Western blots were performed as reported previously (30). Blots were incubated with rabbit anti–epithelial sodium channel (ENaC) γ (1:200, sc-21014; Santa Cruz

Biotechnology, Santa Cruz, CA), anti-semaphorin 3B (SEMA3B) (1:500, PAB12040; Abnova, Jhongli, Taiwan), anti-SEMA3E (1:100, AP7976b; Abgent, San Diego, CA), anti-GRAMD2 (1:100, ab84567; Abcam, Cambridge, MA), anti-SFTPC (1:200, AB3786; Millipore, Billerica, MA), and anti-AQP5 (1:200, AQP-005; Alomone Labs, Jerusalem, Israel). Rabbit anti-Lamin A/C (1:1000, sc-20681; Santa Cruz Biotechnology) was used as a loading control. The secondary antibody was goat anti-rabbit IgG horseradish peroxidase (1:5000, sc-2004, Santa Cruz Biotechnology). (For visualization, *see* online supplement).

Immunofluorescence

Four-percent paraformaldehyde-fixed paraffin-embedded sections of mouse and human lungs were deparaffinized, and this was followed by antigen retrieval using Antigen Unmasking Solution (H3301; Vector Labs, Burlingame, CA). (For detailed methods, *see* online supplement.)

Availability of Supporting Data

All microarray and sequencing data have been deposited in Gene Expression Omnibus. GSE38570 = microarray data for rat AT2 and differentiating AT1-like cells. GSE38571 = microarray data for cross-species comparison. GSE59120 = microarray data for rat organs and purified AT1 cells. GSE66627 = RNAseq data of human AT2 and differentiated AT1-like cells.

Results

Transcriptomic Profiling of Rat AEC and Rat Organs

Three independent isolations of rat AT2 cells were differentiated into AT1-like cells. RNA was obtained from freshly isolated AT2 cells on Day (D) 0 and from AT1-like cells on D2, D4, D6, and D8 in culture. AT2 cells differentiated toward an AT1 cell–like morphology, with the shift in cellular phenotype occurring between D2 and D4 *in vitro* (31, 32). Purities of isolated AT2 cells were 92, 94, and 92%, respectively. Three additional preparations of freshly isolated rat AT1 cells (86, 92, and 93% purity, respectively) were harvested, and RNA was extracted. In addition, RNA was extracted from 14 rat tissues: kidney, spleen, ileum, duodenum, colon, stomach, skin, brain, testis, skeletal muscle, trachea, heart, salivary gland, and liver. Analysis of raw expression data across all samples indicated that one pure AT1 cell sample had technical issues with array processing, and it was therefore removed from the downstream analysis. (For samples included in the downstream analysis, *see* Figure E1 in the online supplement.) The relationship between gene expression profiles was examined by unsupervised hierarchical clustering using the top 20% of genes that were most variant across the dataset. Samples clustered into three major groups (Figure 1A). Cluster 1 consisted of pure AT1 cells and differentiated AT1-like cells, cluster 2 consisted of pure AT2 cells, and cluster 3 included all other tissues. Principal component analysis was performed to determine the strongest contribution to variation in the dataset. It revealed that lung cells clustered separately from other tissues and that AT1 cells clustered with AT1-like cells (Figure 1B). Silhouette plotting of cluster designations revealed that the expression profile of pure AT1 cells was highly similar to that of AT1-like cells (Figure 1C). AT1-like cells were therefore combined with pure AT1 cells to increase statistical power for identification of AT1 cell markers.

Aqp5, *Ager*, and *Pdpn*, all known markers that have elevated expression in AT1 and AT1-like cells, showed statistically significantly higher mean expression in AT1 and AT1-like cells in our dataset as compared with pure AT2 cells and other rat tissues (Figure 2). CAV1 and CAV2 are also purported AT1 cell markers (20, 24, 33). *Cav1* was represented by two probes on the array, neither of which showed statistically significantly elevated expression in AT1 and AT1-like cells; however, the first probe showed a substantial difference in expression between primary AT1 cells and *in vitro* transdifferentiated AT1-like cells, which may have confounded the statistical analysis. *Cav2* did not have significantly higher expression in AT1 and AT1-like cells as compared with pure AT2 cells and the other rat tissues analyzed. Overall, the differences between AT1 cell expression of these known markers and other tissues were modest and, although the means were significantly different for most known AT1 cell markers, their expression in AT2 cells or other tissues limits their utility as unique markers of AT1 cell identity (Figure 2).

Discovery of New AT1 Cell Markers

The top 15% of most variant genes across all samples were selected, and the significance of the expression difference between AT1/AT1-like cells and AT2/other tissues was calculated. D2 cells were excluded from the analyses because they are still in the early stages of differentiation. *Green dots* in Figure 3A represent those genes not expressed in AT1 cells but present in other tissues including AT2 cells. One hundred thirty-two probes, representing 131 genes, were specifically underexpressed in AT1 and AT1-like cells (Figure 3A). Conversely, there were 55 probes in the rat, representing 54 genes that were specifically elevated in AT1 and AT1-like cells compared with AT2 cells. We had previously profiled human primary AT2 cells and *in vitro*–derived AT1-like cells (29). We therefore determined the overlap between those genes that are significantly elevated in rat AT1 and AT1-like cells and those that are significantly up-regulated in human AT1-like cells during AEC differentiation (Figures 3B and 3C). We identified 19 overlapping genes (*red dots* in Figure 3B). Therefore, 35% of the candidate AT1 cell markers found in the rat were also elevated in expression during human AEC transdifferentiation.

These 19 AT1 cell–specific genes are listed in Table 1. One of the 19 overlapping genes, *ANAXA8L2*, is an anticoagulant with known expression in blood cells, and it was therefore removed from subsequent analyses because of a lack of specificity for AT1 cells. The remaining 18 genes derived from the microarray analysis encompassed diverse molecular roles. Two of the 18 genes, *AQP5* and *CDKN2B*, were identified previously as AT1 cell markers (16, 34). Of particular interest were the ENaC components, sodium channel nonvoltage gated γ subunit *SCNN1G* (previously ENaCγ) and *SCNN1B* (previously ENaCβ). Two of the three known ENaC channel components were highly expressed in rat AT1 and AT1-like cells as compared with AT2 cells and all other organs tested, including the kidney (Figure 4). *SCNN1A* (previously ENaCα), the third component of the ENaC channel, was also highly expressed in AT1-like cells; however, this did not pass the false-discovery rate cutoff within the microarray data because of variable expression in other tissues. *SCNN1G* RNA was detected at high levels in the kidney, thereby limiting its utility for

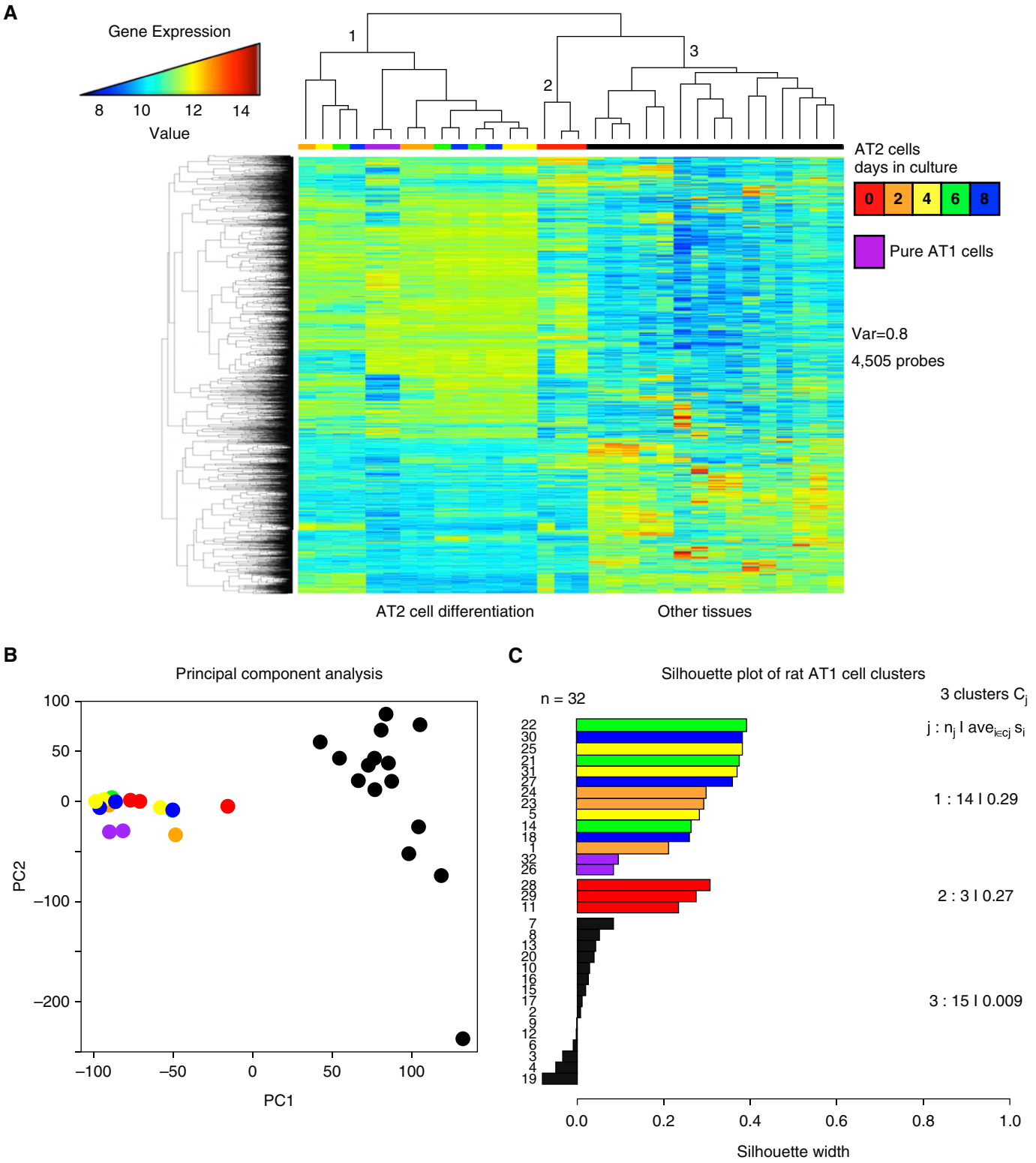
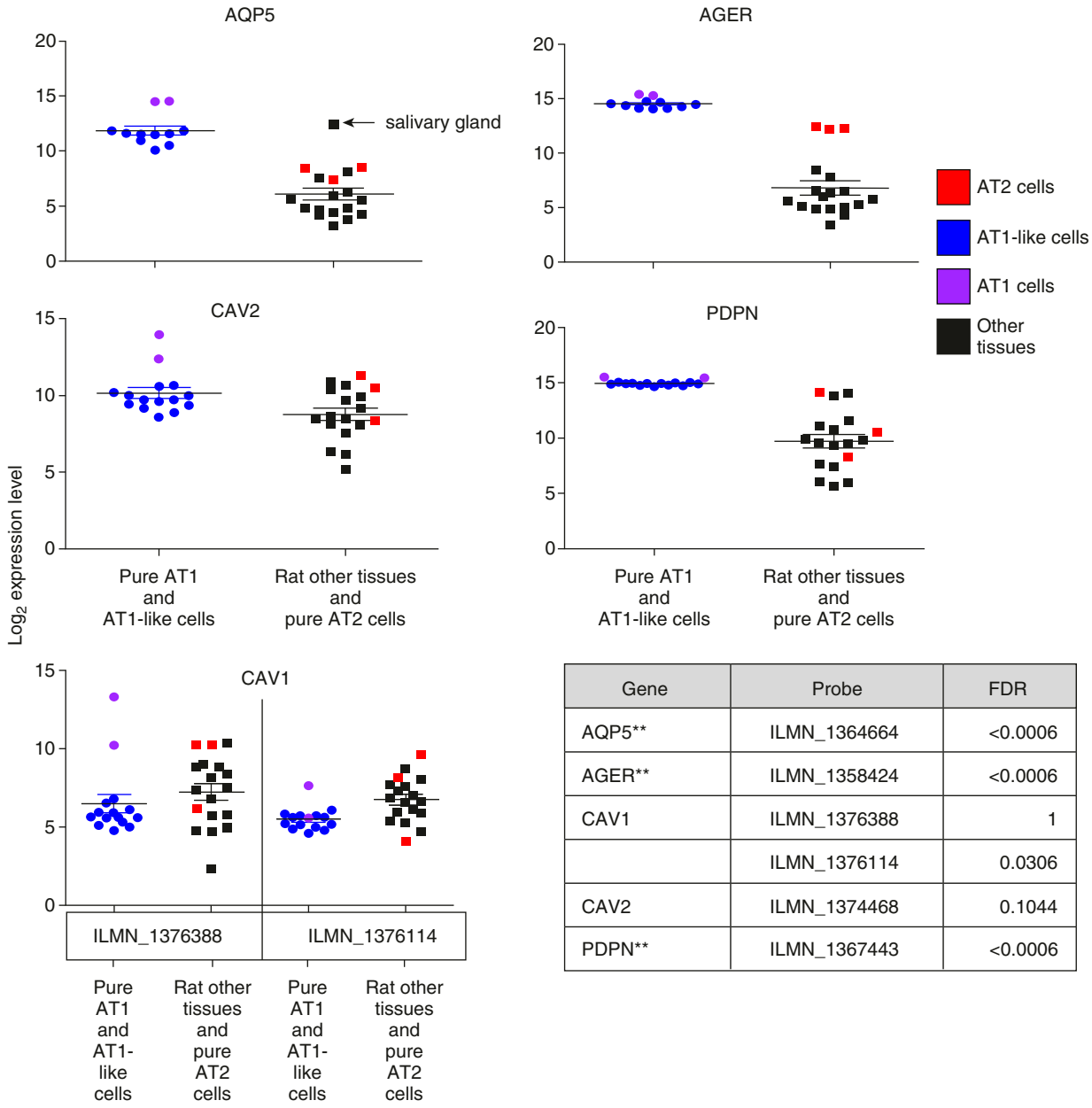


Figure 1. Microarray analysis of gene expression in AT1 and AT2 cells and other tissues. (*A*) Heat map of top 20% of most variant variance stabilization normalization (VSN) normalized gene expression probes, with samples on the *horizontal axis* and gene probes on the *vertical axis*. Low expression (*blue*) and high expression (*red*) are shown. The number of days alveolar epithelial cells were cultured is indicated by the *colored bar* in between the heat map and the dendrogram. (*B*) Principal component analysis of normalized samples. Samples are color-coded by days in culture, as in *A*. (*C*) Silhouette plotting confirms that freshly isolated AT1 cells cluster with differentiating AT1-like cells, as opposed to AT2 cells and other tissues. AT1, alveolar epithelial type 1; AT2, alveolar epithelial type 2; PC, principal component; Var, variance.



Gene	Probe	FDR
AQP5**	ILMN_1364664	<0.0006
AGER**	ILMN_1358424	<0.0006
CAV1	ILMN_1376388	1
	ILMN_1376114	0.0306
CAV2	ILMN_1374468	0.1044
PDPN**	ILMN_1367443	<0.0006

Figure 2. Expression of known AT1 cell markers in rat microarray data. Log2 expression data were generated from microarray experiments: AT1-like cells differentiated in culture (*blue*) (Day 2-6), freshly isolated AT1 cells (*purple*), freshly isolated AT2 cells (*red*), and other tissues (*black*). *Table at bottom right* includes previously described AT1 cell–specific genes, their associated Illumina Probe IDs, and their FDR-corrected *P* values in this study. FDR adjustment is based on the number of tests shown for known genes. AGER, advanced glycosylation end product–specific receptor; AQP5, aquaporin 5; CAV, caveolin; FDR, false-discovery rate; ILMN, Illumina probe number; PDPN, podoplanin. **Indicates significantly greater in rat AT1 and AT1-like cells compared to all others.

modulation of gene expression specifically in AT1 cells. Nevertheless, *SCNN1G* may still be useful within the lung as a marker to distinguish between AT2 and AT1 cells.

In Vitro Validation of Candidate AT1 Cell Markers

The 18 newly identified candidate AT1 cell markers were subjected to quantitative reverse transcription–PCR using the original RNA to verify increased expression in AT1 cells observed in the microarray data. In addition, a second set of rat RNA was isolated from AT1 and AT2 cells and those tissues that, on the basis of the microarray data, might show modest expression of the candidate AT1 cell markers. Samples included three rat AT1 (rAT1) cell samples and duplicates of rat AT2 (rAT2) cells, stomach, skin, brain, kidney, bone, testis, colon, ileum, duodenum, and skeletal muscle. After verification and validation, 4 of the 18 markers were deemed the best candidates for genes specifically expressed in AT1 cells, including *Gramd2*, encoding a protein of unknown function, *Sema3b*, a tumor suppressor gene (35) encoding a secreted protein involved in axonal guidance (36), and *Sema3e*, encoding a semaphorin

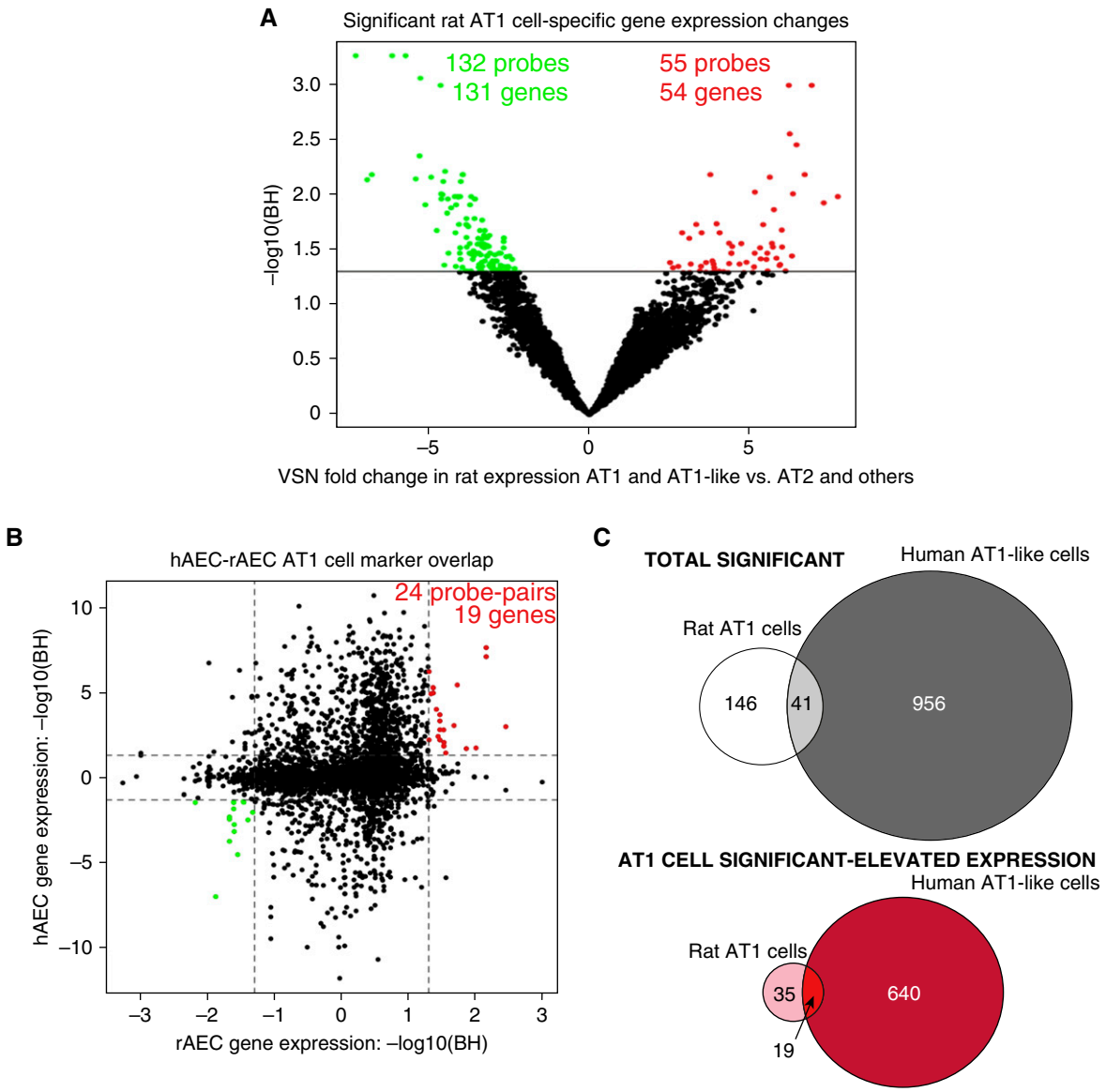


Figure 3. Identification of new rat and human AT1 cell–specific markers. (*A*) Volcano plot of significant differences in gene expression between rat AT1 and AT1-like cells and all other cell types measured: specifically expressed in AT1 and AT1-like cells (*red*) and specifically lacking expression in AT1 and AT1-like cells (*green*). (*B*) Starburst plot of probe pairs between human and rat (*see* online supplement): specifically expressed in rat AT1 and rat AT1-like cells versus AT2 cells/other tissues as well as human AT1-like cells versus human AT2 cells (*red*); genes that lack expression in rat AT1 and AT1-like cells as well as human AT1-like cells but are expressed in all other cells and tissues tested (*green*). (*C*) Venn diagram of overlap between genes. AT1-like cells during alveolar epithelial cell differentiation. *Gray* (*top* Venn) overlap indicates all significantly differentially expressed genes; *red* (*bottom*) overlap indicates only significantly elevated AT1 cell-expressed genes. BH, Benjamini–Hochberg; hAEC, human alveolar epithelial cell; rAEC, rat alveolar epithelial cell.

domain–containing protein involved in embryo vascularization (37) (Figure 5), in addition to the sodium channel subunit *Scnn1g*.

AT1 cell–specific expression of these four candidate AT1 cell markers was verified in rat AT2 cells undergoing differentiation over time, with little to no expression found in freshly isolated AT2 cells at the RNA level (Figure 6A).

Interrogation of protein levels in independent transdifferentiated samples for the four candidate AT1 cell markers revealed a faint signal in rat AT2 cells (Figure 6B), which may be reflective of contaminating AT1 cell material within the cell preparations (which had purities of 92, 92, and 94%, respectively). Levels of the proteins were also examined in human AT2 cells undergoing differentiation, with all showing an absence of SEMA3E, GRAMD2, and AQP5 on D0, with substantially increased expression during differentiation (Figure 6C). Quantitation of AT1 cell marker protein expression showed that GRAMD2 and SCNN1G expression levels increased significantly with time in culture for both human and rat (Figure E2). Although similar trends were observed for SEMA3E and SEMA3B, neither attained

Table 1. Identified Candidate AT1 Cell Markers

Gene Name	rAT1 Fold*	rAT1-BH *P* value†	hAT1 Fold‡	hAT1-BH *P* Value§	Gene Annotation
AQP5	53.81	0.0137	1.61	0.0204	Water channel protein
CDKN2B	49.27	0.0069	6.56	8.01E-08	p15, CDK inhibitor
CLIC5	46.59	0.0342	8.75	0.0004	Chloride intracellular channel
CRLF1	42.66	0.0187	10.23	3.71E-06	Cytokine receptor-like factor
EGFL6	64.22	0.0300	1.65	0.0016	Secreted EGF repeat protein
FSTL3	52.61	0.0300	2.83	0.0143	Inhibits BMP and TGFβ signaling
GPRC5A	61.54	0.0433	2.47	5.24E-06	Retinoic acid inducible G protein
GRAMD2	63.81	0.0210	2.44	0.0008	Unknown
KRT7	79.46	0.0363	2.27	0.0038	Type II cytokeratin
LOC643037	81.19	0.0098	1.38	0.0185	Unknown
MEX3B	6.08	0.0463	3.58	1.21E-05	RNA binding protein
MMP11	21.43	0.0342	2.53	0.0063	Intracellular activated by furin MMP
RADIL	26.61	0.0279	1.77	0.0365	Ras-associated cell adhesion protein
SCNN1B	40.51	0.0385	5.64	9.97E-05	ENaCβ, Na^+ channel component
SCNN1G	21.86	0.0297	2.49	0.0088	ENaCγ, Na^+ channel component
SEMA3B	35.53	0.0342	1.92	0.0015	Tumor suppressor of cell migration
SEMA3E	87.76	0.0035	1.85	0.0010	ligand for PLEXD1
SMARCA1	25.61	0.0429	2.62	5.41E-06	ATPase subunit of NURF complex

Definition of abbreviations: AT1, alveolar epithelial type 1; BH, Benjamini–Hochberg; BMP, bone morphogenetic protein; CDK, cyclin-dependent kinase; EGF, epidermal growth factor; ENaC, epithelial sodium channel; h, human; MMP, matrix metalloproteinase; NURF, nucleosome remodeling factor; PLEXD1, plexin D1; r, rat; TGFβ, transforming growth factor β.

Genes are expressed specifically in both human and rat as potential AT1 cell markers. *ANXA8L2* was excluded from the dataset because it is likely a blood contaminant.

*Fold change in AT1 cells relative to other tissues including AT2 cells in rat.

†BH-corrected *P* value for significance of specific expression in AT1 and AT1-like rat cells.

‡Fold change in AT1 cells relative to AT2 cells in human.

§BH-corrected *P* value for significance of specific expression in AT1-like human cells.

significance in either species. In addition, we used RNAseq to visually inspect transcript levels in human AT2 and AT1-like cells. Expression of *GRAMD2* and *SCNN1G* was absent in human AT2 cells and robustly present in AT1-like cells, whereas SEMA3B showed variable expression in human AT2 cells from different donors, and *SEMA3E* had relatively low levels of expression overall (Figure E3). Accordingly, we selected *GRAMD2* and *SCNN1G* for further study.

Validation of GRAMD2 Expression in Lung Tissue

To assess differential expression in the alveolar epithelium in lung tissue, sections were immunostained for GRAMD2 and known AT1 and AT2 cell markers. Staining of mouse lung sections showed GRAMD2 expression in AT1 cells (Figure E4A). Double staining with the AT2 cell marker pro-SFTPC confirmed that GRAMD2 was not present on the surface of AT2 cells (Figure 7A). In addition, costaining of GRAMD2 with the known AT1 cell marker, AQP5, showed colocalization along the apical surface of AT1 cells (Figure 7B). Furthermore, using human tissue, we found that GRAMD2 does not colocalize with ABCA3, a marker of AT2 cells (Figure E4B).

To address whether *Gramd2* is expressed in other cell types within the lung, we performed RNA isolation and reverse transcription–PCR quantification of rat lung endothelial cells together with *in vitro*–differentiated AT1-like cells. *Gramd2* was not expressed in purified lung endothelial cells (Figure E5A). Immunofluorescence demonstrated that bronchial epithelial tissue adjacent to alveolar tissue was negative for GRAMD2 (Figure E5B). We were thus able to detect GRAMD2 protein and RNA expression specifically in AT1 cells, but not in other cell types examined within human, mouse, and rat lungs.

Discussion

We performed whole transcriptome profiling on enriched populations of rat AT2, AT1, and AT1-like cells, and compared their gene expression profiles with those of other tissues. Selecting for overlap between rat and human resulted in 19 candidate AT1 cell–specific genes, one of which was excluded because of its expression in blood cells. Quantitative reverse transcription–PCR validation indicated that four genes had highly AT1 cell–specific RNA expression, namely *GRAMD2*, *SEMA3B*, *SEMA3E*, and *SCNN1G*. Analysis of *in vitro* AT2 cell differentiation in independent sample sets indicated that both messenger RNA and protein expression of the four top candidate genes was enriched in AT1-like cells in both human and rat, with GRAMD2 and SCNN1G being significantly differentially expressed between AT2 and AT1/AT1-like cells in the lung. Immunofluorescence of mouse and human lung sections confirmed localization of GRAMD2 to AT1, but not AT2, cells.

Public databases list GRAMD2 as a lung-specific protein, with potential expression in the uterus. Indeed, GRAMD2 expression has been reported in uterine tissue on the basis of publicly deposited complementary DNA clones (IMAGp998A1111009) and staining of normal endometrium (The Human Protein Atlas), which may affect female health and reproduction of transgenic

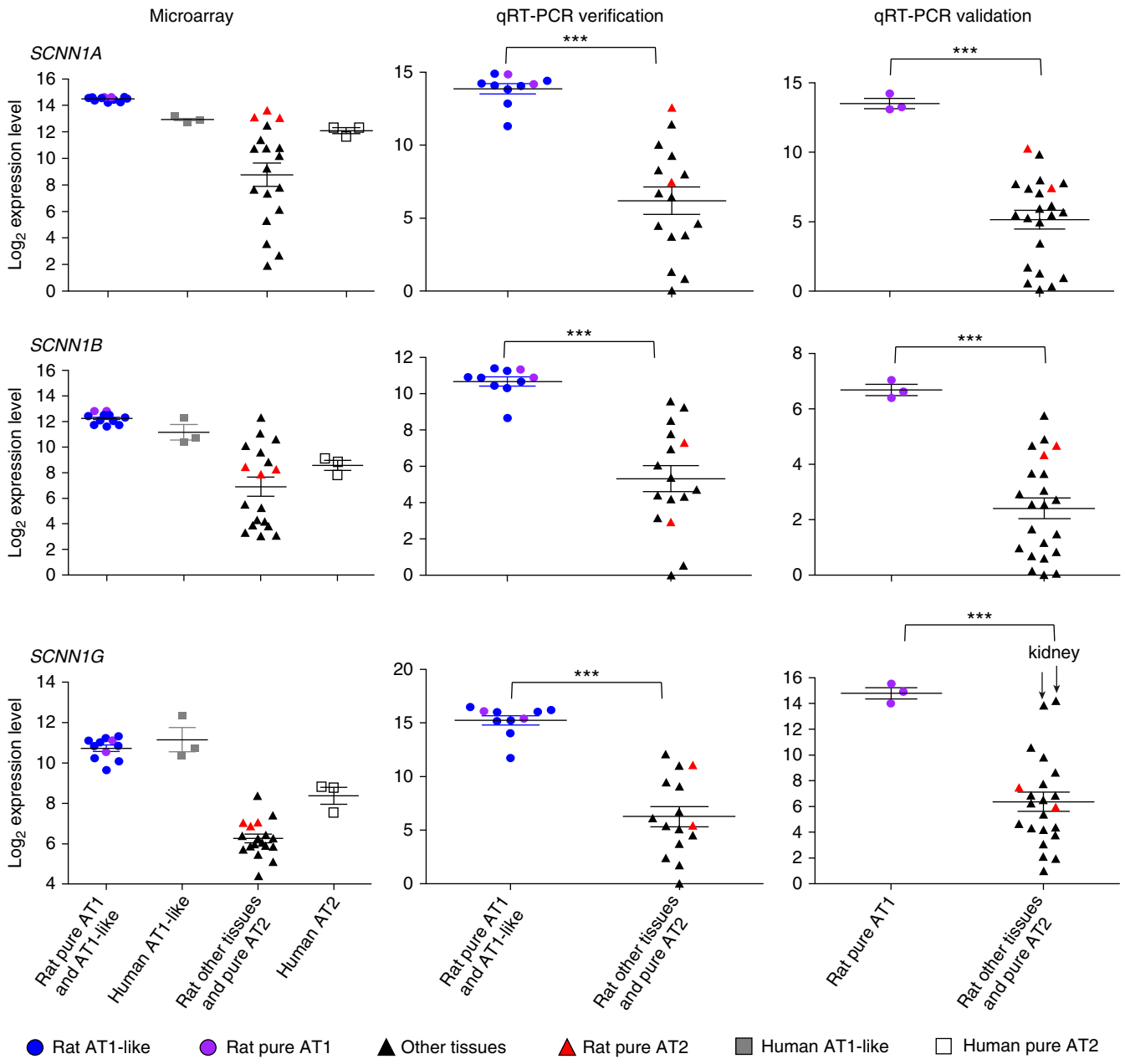


Figure 4. Validation of the epithelial sodium channel (ENaC) messenger RNA in different cell types. (*A*) Expression of *SCNN1A* (ENaCα), *SCNN1B* (ENaCβ), and *SCNN1G* (ENaCγ) is indicated. *Left panel*: data derived from microarray values. *Middle panels*: qRT-PCR verification of microarray results. *Right panels*: qRT-PCR of independent sample isolations for validation purposes. Pure rat AT1 (*purple circles*) and rat AT1-like cells (*blue circles*), human AT1-like cells (*gray squares*), rat AT2 cells (*red triangles*), human AT2 cells (*clear squares*), and rat other tissues (*black triangles*) are indicated. ***$P < 0.0001$. qRT-PCR, quantitative reverse transcription–polymerase chain reaction; SCNN1, sodium channel nonvoltage gated subunit.

mouse models, depending on this gene's function. Fortunately, this would not affect studies in male mammals. On the basis of homology, GRAMD2 is predicted by structural evaluation programs to be a transmembrane protein with a large intracellular domain, a transmembrane region, and an extracellular domain. To our knowledge, no mechanistic work has been published on this protein.

The other potentially new AT1 cell–specific identified gene is *SCNN1G*. The ENaCγ subunit is well described as being expressed in the kidney. Our results are in agreement with *Scnn1g* expression in that organ, albeit at substantially lower levels than in the lung. ENaC mediates sodium transport across the apical membrane to regulate water absorption (38). Our interrogation of ENaC components demonstrated that the messenger RNA of all three primary subunits is expressed at significantly higher levels in AT1 and AT1-like cells than in AT2 cells or other tissues of the body, consistent with previous reports

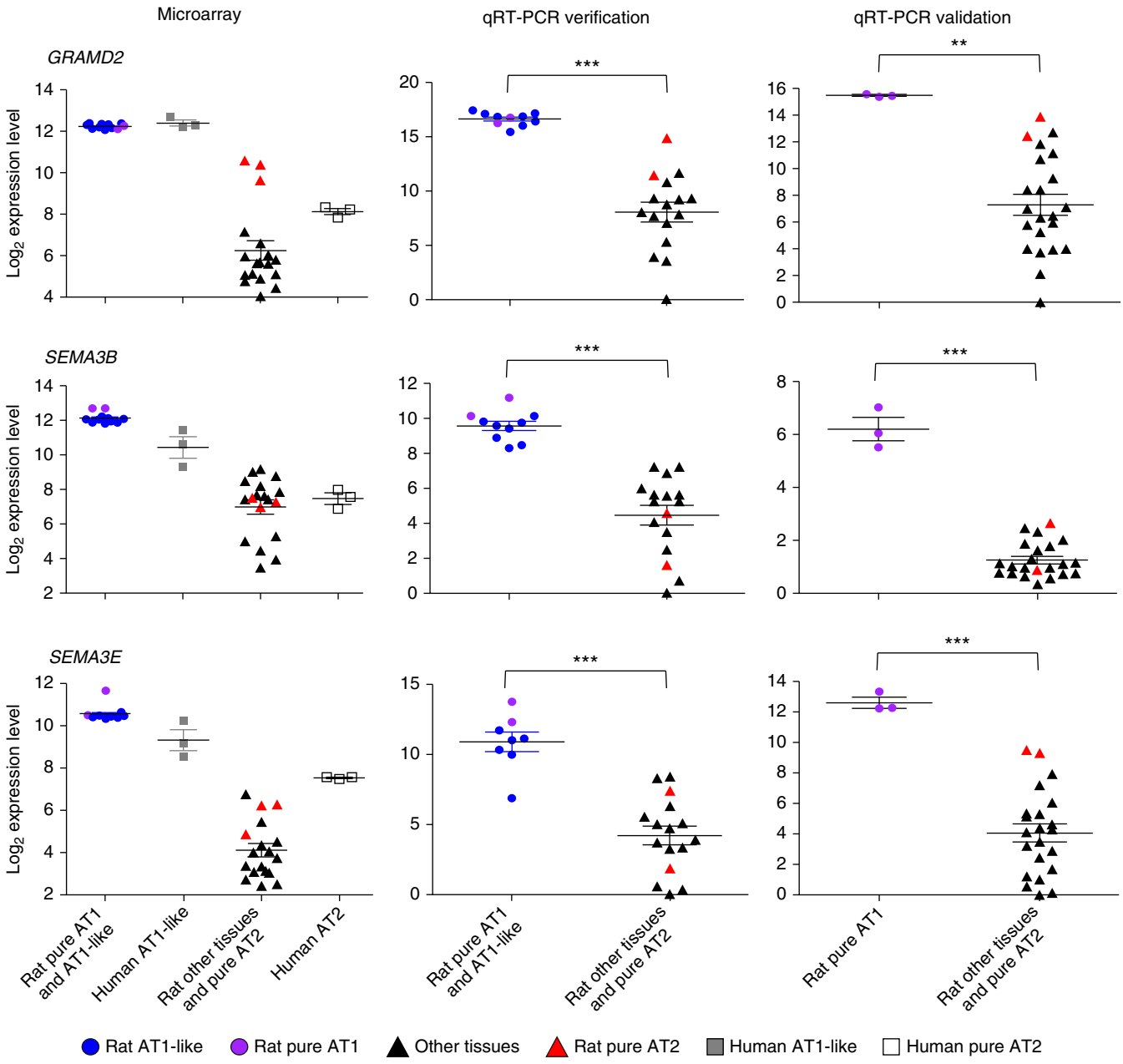


Figure 5. Validation of new AT1 cell marker gene expression. Log2 expression levels of *GRAMD2*, *SEMA3B*, and *SEMA3E* are indicated. *Left panel*: data derived from microarray values. *Middle panels*: qRT-PCR verification of microarray results. *Right panels*: qRT-PCR of independent sample isolations for validation purposes. Samples include pure rat AT1 (*purple circles*) and rat AT1-like cells (*blue circles*), human AT1-like cells (*gray squares*), rat AT2 cells (*red triangles*), human AT2 cells (*clear squares*), and rat other tissues (*black triangles*). ***$P < 0.0001$; **between $P < 0.001$ and $P > 0.0001$. GRAMD2, GRAM domain 2; SEMA3B, semaphorin 3B; SEMA3E, semaphorin 3E.

of ENaC presence in AT1 cells (28). This finding supports the notion that ENaC plays an important role in sodium transport within the lung and that the AT1 cell contributes significantly to overall ENaC activity and sodium transport. Our observations agree with previous work (39) describing relative ENaC levels and activity between AT2 and AT1 cells; however, expression of *Scnn1g* in the kidney limits the utility of this marker to lung-specific applications such as cell isolation and staining techniques.

This study is limited by the tissue samples we selected. Because our profiled tissues were derived from male rats, it remains to be determined if the candidate genes identified herein are expressed in female reproductive organs. Of note is the use of whole organs to investigate gene expression. Ideally, all organs should be separated into their constitutive subpopulations of cells, but it was beyond the scope of this study to perform such separations on >20 organ systems. Alternatively, specificity of expression for the putative AT1 cell genes in tissues

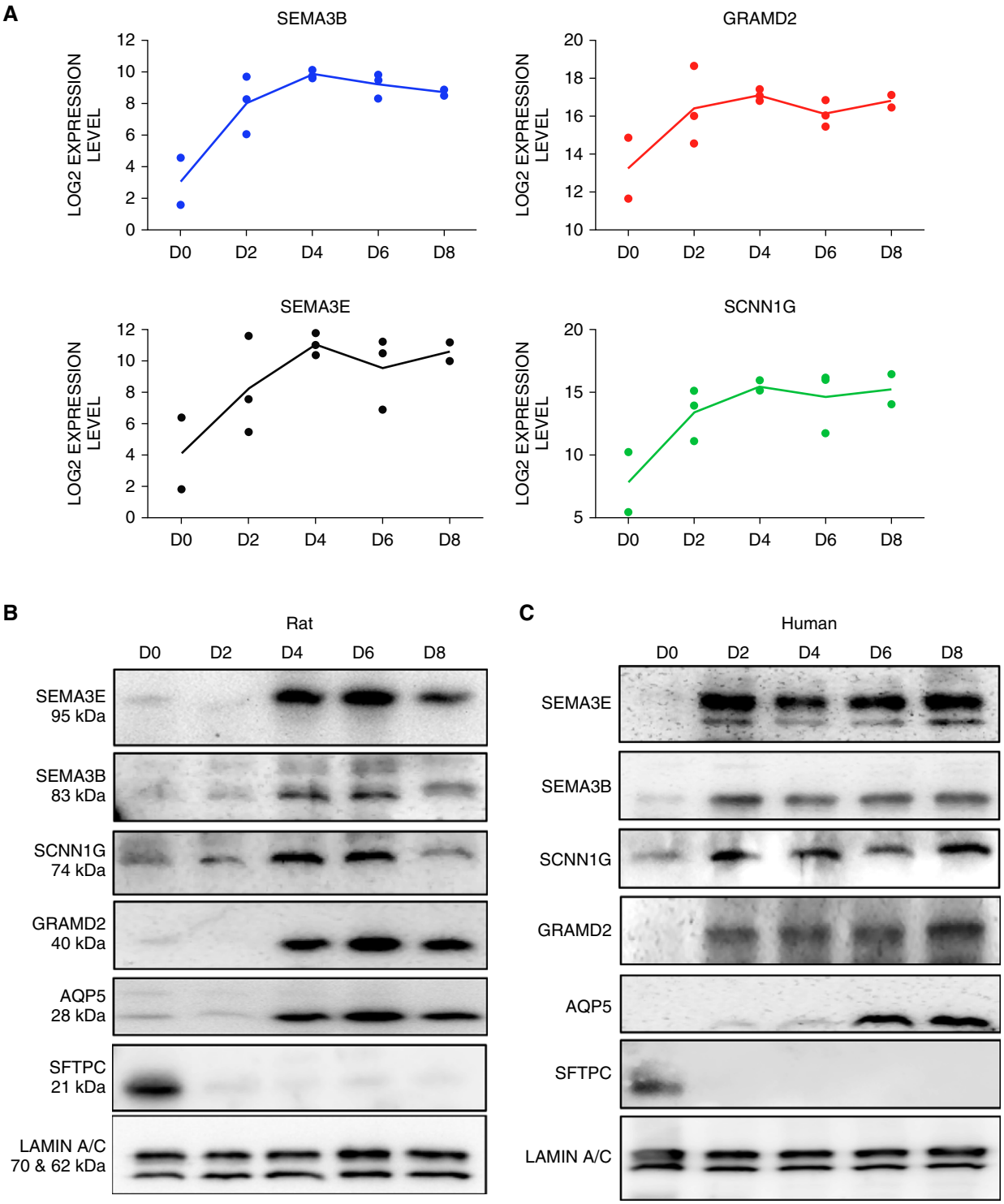


Figure 6. Validation of new AT1 cell–specific marker expression in *in vitro*–differentiated AT1-like cells. (*A*) qRT-PCR of rat AT cell differentiation *in vitro*. Expression levels were normalized to 18S RNA. Three preparations were analyzed. (*B*) Western analysis of rat alveolar epithelial cell differentiation. D0 = freshly isolated AT2 cells and D8 = differentiated AT1-like cells. Time points in between reflect intermediate states of differentiation. Representative blots from four independent experiments are shown. (*C*) Western analysis of human alveolar epithelial cell differentiation. D0 = pure AT2 cells, D8 = differentiated AT1-like cells. Time points in between reflect intermediate states of differentiation. Representative blots from three independent experiments are shown. SFTPC, surfactant protein C.

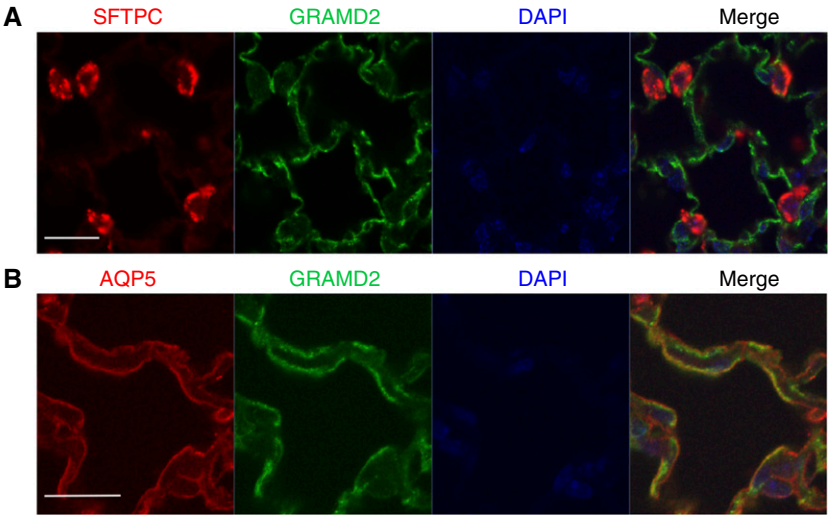


Figure 7. Validation of GRAMD2 expression in AT1 cells in lung tissue. (*A*) Confocal images for GRAMD2/SFTPC double staining in mouse lung sections show that GRAMD2 does not colocalize with SFTPC. DAPI is the nuclear counterstain. *Scale bar*: 20 μm. (*B*) Confocal images for GRAMD2/AQP5 double staining in mouse lung sections shows that GRAMD2 colocalizes with AQP5. DAPI is the nuclear counterstain. *Scale bar*: 20 μm. DAPI, 4′,6-diamidino-2-phenylindole.

can be confirmed using transgenic knockin animal models in which endogenous promoter or enhancer elements could be used to drive expression of Cre to activate Cre-responsive marker genes, such as in the Rosa26 LacZ or Rosa26 Tomato reporter mouse lines (40). Such mice are currently being developed in our laboratory.

Previously described AT1 cell markers such as *Aqp5*, *Pdpn*, and *Hopx* have intrinsic drawbacks for gene deletion specifically in AT1 cells because of their expression in other tissues and/or cell types. RNAseq analysis shows that the recently described candidate AT1 cell marker *HOPX* is expressed in human AT2 cells and also in cells with an intermediate phenotype (15). In this study, we harnessed the discovery potential of whole genome transcriptional profiling of purified AT2, AT1, and *in vitro*–derived AT1-like cells, combined with expression levels genome-wide from 15 other organs, to identify new AT1 cell–specific genes. Integration across species (i.e., between human and rat) allowed the identification of markers relevant for studies of the human lung. Analysis of multiple genome-wide datasets allowed us to refine candidate AT1 cell markers with application to multiple model species including humans and led to the identification of GRAMD2 as the best AT1-cell specific marker. Collectively, these approaches will shed light on the biological roles of the enigmatic AT1 cell in health and disease. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgments: The authors thank J. Alvarez for assistance with AT2 cell preparations and M. Flores for assistance with AT1 cell preparations. Microarray data were generated by the Southern California Genomic Consortium at the University of California, Los Angeles. Histology and microscopy services were provided by the Cell and Tissue Imaging Core of the University of Southern California Research Center for Liver Diseases.

References

1. Perl AK, Tichelaar JW, Whitsett JA. Conditional gene expression in the respiratory epithelium of the mouse. *Transgenic Res* 2002;11:21–29.
2. Rawlins EL, Perl AK. The a"MAZE"ing world of lung-specific transgenic mice. *Am J Respir Cell Mol Biol* 2012;46:269–282.
3. Barkauskas CE, Cronce MJ, Rackley CR, Bowie EJ, Keene DR, Stripp BR, Randell SH, Noble PW, Hogan BL. Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest* 2013;123:3025–3036.
4. Sutherland KD, Proost N, Brouns I, Adriaensen D, Song JY, Berns A. Cell of origin of small cell lung cancer: inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. *Cancer Cell* 2011;19:754–764.
5. Xu X, Rock JR, Lu Y, Futtner C, Schwab B, Guinney J, Hogan BL, Onaitis MW. Evidence for type II cells as cells of origin of K-Ras-induced distal lung adenocarcinoma. *Proc Natl Acad Sci USA* 2012;109: 4910–4915.
6. Adamson IY, Bowden DH. The type 2 cell as progenitor of alveolar epithelial regeneration. A cytodynamic study in mice after exposure to oxygen. *Lab Invest* 1974;30:35–42.
7. Danto SI, Zabski SM, Crandall ED. Reactivity of alveolar epithelial cells in primary culture with type I cell monoclonal antibodies. *Am J Respir Cell Mol Biol* 1992;6:296–306.
8. Cheek JM, Evans MJ, Crandall ED. Type I cell-like morphology in tight alveolar epithelial monolayers. *Exp Cell Res* 1989;184:375–387.
9. Dobbs LG, Williams MC, Gonzalez R. Monoclonal antibodies specific to apical surfaces of rat alveolar type I cells bind to surfaces of cultured, but not freshly isolated, type II cells. *Biochim Biophys Acta* 1988;970: 146–156.
10. Brody JS, Williams MC. Pulmonary alveolar epithelial cell differentiation. *Annu Rev Physiol* 1992;54:351–371.
11. Borok Z, Lubman RL, Danto SI, Zhang XL, Zabski SM, King LS, Lee DM, Agre P, Crandall ED. Keratinocyte growth factor modulates alveolar epithelial cell phenotype *in vitro*: expression of aquaporin 5. *Am J Respir Cell Mol Biol* 1998;18:554–561.
12. Qiao R, Yan W, Clavijo C, Mehrian-Shai R, Zhong Q, Kim KJ, Ann D, Crandall ED, Borok Z. Effects of KGF on alveolar epithelial cell transdifferentiation are mediated by JNK signaling. *Am J Respir Cell Mol Biol* 2008;38:239–246.
13. Danto SI, Shannon JM, Borok Z, Zabski SM, Crandall ED. Reversible transdifferentiation of alveolar epithelial cells. *Am J Respir Cell Mol Biol* 1995;12:497–502.
14. Jain R, Barkauskas CE, Takeda N, Bowie EJ, Aghajanian H, Wang Q, Padmanabhan A, Manderfield LJ, Gupta M, Li D, *et al*. Plasticity of Hopx(+) type I alveolar cells to regenerate type II cells in the lung. *Nat Commun* 2015;6:6727.
15. Liebler JM, Marconett CN, Juul N, Wang H, Liu Y, Flodby P, Laird-Offringa IA, Minoo P, Zhou B. Combinations of differentiation markers distinguish subpopulations of alveolar epithelial cells in adult lung. *Am J Physiol Lung Cell Mol Physiol* 2016;310:L114–L120.
16. Flodby P, Borok Z, Banfalvi A, Zhou B, Gao D, Minoo P, Ann DK, Morrisey EE, Crandall ED. Directed expression of Cre in alveolar epithelial type 1 cells. *Am J Respir Cell Mol Biol* 2010; 43:173–178.
17. Raina S, Preston GM, Guggino WB, Agre P. Molecular cloning and characterization of an aquaporin cDNA from salivary, lacrimal, and respiratory tissues. *J Biol Chem* 1995;270:1908–1912.

18. Moon C, Rousseau R, Soria JC, Hoque MO, Lee J, Jang SJ, Trink B, Sidransky D, Mao L. Aquaporin expression in human lymphocytes and dendritic cells. *Am J Hematol* 2004;75:128–133.
19. Shirasawa M, Fujiwara N, Hirabayashi S, Ohno H, Iida J, Makita K, Hata Y. Receptor for advanced glycation end-products is a marker of type I lung alveolar cells. *Genes Cells* 2004;9:165–174.
20. McElroy MC, Kasper M. The use of alveolar epithelial type I cell-selective markers to investigate lung injury and repair. *Eur Respir J* 2004;24:664–673.
21. Rishi AK, Joyce-Brady M, Fisher J, Dobbs LG, Floros J, VanderSpek J, Brody JS, Williams MC. Cloning, characterization, and development expression of a rat lung alveolar type I cell gene in embryonic endodermal and neural derivatives. *Dev Biol* 1995;167:294–306.
22. Williams MC, Cao Y, Hinds A, Rishi AK, Wetterwald A. T1α protein is developmentally regulated and expressed by alveolar type I cells, choroid plexus, and ciliary epithelia of adult rats. *Am J Respir Cell Mol Biol* 1996;14:577–585.
23. Ramirez MI, Rishi AK, Cao YX, Williams MC. TGT3, thyroid transcription factor I, and Sp1 elements regulate transcriptional activity of the 1.3-kilobase pair promoter of T1α, a lung alveolar type I cell gene. *J Biol Chem* 1997;272:26285–26294.
24. Newman GR, Campbell L, von Ruhland C, Jasani B, Gumbleton M. Caveolin and its cellular and subcellular immunolocalisation in lung alveolar epithelium: implications for alveolar epithelial type I cell function. *Cell Tissue Res* 1999;295:111–120.
25. Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH, Desai TJ, Krasnow MA, Quake SR. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* 2014;509:371–375.
26. Navarro A, Perez RE, Rezaiekhaligh M, Mabry SM, Ekekezie II. T1α/podoplanin is essential for capillary morphogenesis in lymphatic endothelial cells. *Am J Physiol Lung Cell Mol Physiol* 2008;295: L543–L551.
27. Borok Z, Danto SI, Lubman RL, Cao Y, Williams MC, Crandall ED. Modulation of T1α expression with alveolar epithelial cell phenotype *in vitro*. *Am J Physiol* 1998;275:L155–L164.
28. Borok Z, Liebler JM, Lubman RL, Foster MJ, Zhou B, Li X, Zabski SM, Kim KJ, Crandall ED. Na transport proteins are expressed by rat alveolar epithelial type I cells. *Am J Physiol Lung Cell Mol Physiol* 2002;282:L599–L608.
29. Marconett CN, Zhou B, Rieger ME, Selamat SA, Dubourd M, Fang X, Lynch SK, Stueve TR, Siegmund KD, Berman BP, *et al*. Integrated transcriptomic and epigenomic analysis of primary human lung epithelial cell differentiation. *PLoS Genet* 2013;9:e1003513.
30. Zhou B, Francis TA, Yang H, Tseng W, Zhong Q, Frenkel B, Morrisey EE, Ann DK, Minoo P, Crandall ED, *et al*. GATA-6 mediates transcriptional activation of aquaporin-5 through interactions with Sp1. *Am J Physiol Cell Physiol* 2008;295:C1141–C1150.
31. Borok Z, Hami A, Danto SI, Zabski SM, Crandall ED. Rat serum inhibits progression of alveolar epithelial cells toward the type I cell phenotype *in vitro*. *Am J Respir Cell Mol Biol* 1995;12:50–55.
32. Borok Z, Danto SI, Zabski SM, Crandall ED. Defined medium for primary culture *de novo* of adult rat alveolar epithelial cells. *In Vitro Cell Dev Biol Anim* 1994;30A:99–104.
33. Kasper M, Reimann T, Hempel U, Wenzel KW, Bierhaus A, Schuh D, Dimmer V, Haroske G, Müller M. Loss of caveolin expression in type I pneumocytes as an indicator of subcellular alterations during lung fibrogenesis. *Histochem Cell Biol* 1998;109:41–48.
34. Qiao R, Zhou B, Liebler JM, Li X, Crandall ED, Borok Z. Identification of three genes of known function expressed by alveolar epithelial type I cells. *Am J Respir Cell Mol Biol* 2003;29:98–105.
35. Loginov VI, Dmitriev AA, Senchenko VN, Pronina IV, Khodyrev DS, Kudryavtseva AV, Krasnov GS, Gerashchenko GV, Chashchina LI, Kazubskaya TP, *et al*. Tumor suppressor function of the SEMA3B gene in human lung and renal cancers. *PLoS One* 2015;10: e0123369.
36. Chauvet S, Cohen S, Yoshida Y, Fekrane L, Livet J, Gayet O, Segu L, Buhot MC, Jessell TM, Henderson CE, *et al*. Gating of Sema3E/PlexinD1 signaling by neuropilin-1 switches axonal repulsion to attraction during brain development. *Neuron* 2007;56: 807–822.
37. Gu C, Yoshida Y, Livet J, Reimert DV, Mann F, Merte J, Henderson CE, Jessell TM, Kolodkin AL, Ginty DD. Semaphorin 3E and plexin-D1 control vascular pattern independently of neuropilins. *Science* 2005; 307:265–268.
38. Hummler E, Barker P, Gatzy J, Beermann F, Verdumo C, Schmidt A, Boucher R, Rossier BC. Early death due to defective neonatal lung liquid clearance in α-ENaC-deficient mice. *Nat Genet* 1996;12: 325–328.
39. Johnson MD, Widdicombe JH, Allen L, Barbry P, Dobbs LG. Alveolar epithelial type I cells contain transport proteins and transport sodium, supporting an active role for type I cells in regulation of lung liquid homeostasis. *Proc Natl Acad Sci USA* 2002;99: 1966–1971.
40. Feil S, Krauss J, Thunemann M, Feil R. Genetic inducible fate mapping in adult mice using tamoxifen-dependent Cre recombinases. *Methods Mol Biol* 2014;1194:113–139.