Cell Population–resolved Multiomics Atlas of the Developing Lung

Mereena G. Ushakumary[1]*, Song Feng[1]*, Gautam Bandyopadhyay[2], Heather Olson[1], Karl K. Weitz[1], Heidi L. Huyck[2], Cory Poole[2], Jeffrey M. Purkerson[2], Soumyaroop Bhattacharya[2], M. Cecilia Ljungberg[3,4], Thomas J. Mariani[2], Gail H. Deutsch[5], Ravi S. Misra[2], James P. Carson[6], Joshua N. Adkins[1,7], Gloria S. Pryhuber[2], and Geremy Clair[1]

[1]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington; [2]Department of Pediatrics, School of Medicine and Dentistry, University of Rochester, Rochester, New York; [3]Department of Pediatrics, College of Medicine, Baylor University, Houston, Texas; [4]Jan and Dan Duncan Neurological Research Institute at Texas Children's Hospital, Houston, Texas; [5]Department of Laboratory Medicine and Pathology, University of Washington, Seattle, Washington; [6]Texas Advanced Computing Center, University of Texas at Austin, Austin, Texas; and [7]Department of Biomedical Engineering, Oregon Health and Science University, Portland, Oregon

ORCID IDs: 0000-0002-9185-3994 (G.S.P.); 0000-0003-0478-8520 (G.C.).

Abstract

The lung is a vital organ that undergoes extensive morphological and functional changes during postnatal development. To disambiguate how different cell populations contribute to organ development, we performed proteomic and transcriptomic analyses of four sorted cell populations from the lung of human subjects 0–8 years of age with a focus on early life. The cell populations analyzed included epithelial, endothelial, mesenchymal, and immune cells. Our results revealed distinct molecular signatures for each of the sorted cell populations that enable the description of molecular shifts occurring in these populations during postnatal development. We confirmed that the proteome of the different cell populations was distinct regardless of age and identified functions specific to each population. We identified a series of cell population protein markers, including those located at the cell surface, that show differential expression and distribution on RNA *in situ* hybridization and immunofluorescence imaging. We validated the spatial distribution of alveolar type 1 and endothelial cell surface markers. Temporal analyses of the proteomes of the four populations revealed processes modulated during postnatal development and clarified the findings obtained from whole-tissue proteome studies. Finally, the proteome was compared with a transcriptomics survey performed on the same lung samples to evaluate processes under post-transcriptional control.

Keywords: proteomics; transcriptomics; pulmonary development; cell types

The unique architecture of the mammalian lung is a result of evolutionary adaptation to breathing air (1, 2). To fulfill its diverse functions, the human respiratory system comprises cell types derived from embryonic neuroectoderm, mesoderm, and endoderm (1, 3). Many cell types with specific proportions and localization form the diverse pulmonary tissue niches (3, 4), including pleura, bronchi, proximal and distal small airways, submucosal glands, alveolar saccules, and vasculature (1, 5–8). Throughout development, cellular and molecular events drive morphogenesis and achieve functions required for

(Received in original form March 4, 2024; accepted in final form October 24, 2024)

*These authors contributed equally to this work.

Supported by National Heart, Lung, and Blood Institute Molecular Atlas of Lung Development Program Human Tissue Core grants U01HL122700 and U01HL148861 (G.S.P.) and U01HL148860 (J.P.C., J.N.A., G.C.). Parts of this work were performed in the Environmental Molecular Science Laboratory, a U.S. Department of Energy national scientific user facility at the Pacific Northwest National Laboratory. Battelle operates the Pacific Northwest National Laboratory for the Department of Energy under contract DE-AC05-76RLO01830.

Author Contributions: G.C., G.S.P., J.P.C., T.J.M., and J.N.A. conceptualized the study and coordinated the team efforts. H.L.H., R.S.M., and G.S.P. collected the samples. G.H.D. performed hematoxylin and eosin and pathology evaluations. H.O. prepared the samples for mass spectrometry analyses. K.K.W. performed the mass spectrometry experiments. M.G.U. and S.F. performed data analysis and interpretations. G.B. and S.B. performed the analysis of the RNA datasets. C.P. and J.M.P. performed multiplexed immunofluorescence experiments. M.C.L. performed *in situ* hybridization experiments. M.G.U., S.F., and G.C. wrote the draft manuscript.

Correspondence and requests for reprints should be addressed to Geremy Clair, Ph.D., Biological Sciences Division, Pacific Northwest National Laboratory, 902 Battelle Boulevard, P.O. Box 999, MSIN K8-98, Richland, WA 99352. E-mail: geremy.clair@pnnl.gov.

Clinical Relevance

This paper details the change in the proteome and transcriptome occurring in human normal development. The results also enable the identification of protein markers of cell types and populations.

breathing and defense against pathogens. Understanding these mechanisms is critical, as abnormal morphogenesis or function can lead to acute or chronic respiratory diseases in children and adults (9–12). Although traditional experimental approaches have defined diverse aspects of human lung development, significant gaps remain in our understanding of the molecular and signaling pathways that control the formation and functioning of pulmonary tissues. High-throughput omics technologies now enable the precise measurement of thousands of biomolecules from the same tissue. Most omics studies performed during pulmonary mammalian development have focused on the profiling of bulk tissue metabolites (13), lipids (14, 15), proteins (16, 17), RNAs (18–21), and chromatin accessibility (22). Although there are dozens of omics datasets in human tissues (23–25), most were performed from fetal and adult tissue, and only a limited number of unbiased studies were performed on infants tissues because of the scarcity of tissue repositories containing these type of cases (15, 19, 22, 26). Two studies combining proteomics and transcriptomics during development, one from bulk pulmonary tissue (27) and another from laser microdissected alveolar parenchyma (26), suggested a limited correlation between protein and RNA abundances, likely due to post-transcriptional events. Although informative, these studies were performed on heterogeneous whole-lung tissues composed of many cell types. In the past decade, single-cell RNA sequencing (scRNA-seq) has been performed on pulmonary tissues, demonstrating further the complexity of the organ and the importance of studying key cell populations in health and disease (3, 20, 28–31). Although the field of single-cell proteomics is progressing (32–35), the current throughput limits the execution of large studies on individual dissociated cells. Waiting for single-cell proteomics to become amenable to extensive studies, we opted for an intermediate approach using FACS to isolate the four main cell lineages composing the lung (i.e., epithelial, endothelial, nonendothelial mesenchymal, and mixed immune cells). We used this strategy previously to characterize molecular differences among the cell populations (15, 36, 37) at a single time point (e.g., 20 mo of age) for the proteome and the lipidome (15, 36) and between 1 day and 4 years for the transcriptome (37). In this study, we analyze cell population–specific temporal abundance shifts in the proteome and transcriptome in lung cells from donors between 0 and 8 years of age obtained from the Molecular Atlas of Lung Development Program's Biorepository for Investigation of Neonatal Diseases of the Lung (https://brindl.urmc.rochester.edu) (38).

Here, we performed deep proteome profiling to identify biochemical pathways and cell markers characterizing different cell populations. Comparison of the proteomics dataset with RNA sequencing performed on the same sorted cell populations from the same lung sample, defined processes that are potentially post-transcriptionally regulated. The unique proteomics data and analyses provided from a range of human pediatric lung sample ages will serve as a unique resource for the study of molecular events mediating normal postnatal lung development.

Methods

Detailed methods can be found in the data supplement.

Human Donors

Human lung tissue was obtained through the nonprofit United Network for Organ Sharing facilitated by the International Institute for the Advancement of Medicine and the National Disease Research Interchange. The process for procuring lungs can be found at https://www.protocols.io/view/602-2-donor-acceptance-criteria-for-urmc-htc-hubma-4r3l28dnql1y/v1. Lungs were of transplant quality; researchers were offered tissue only if it could not be placed for transplantation. Consent was given for the use of the deidentified tissue of a deceased donor in nonhuman subjects research overseen by the University of Rochester Research Subjects Review Board protocol (RSRB00056775). Donor metadata are provided in the Dataset E1 in the data supplement, and distal tissue hematoxylin and eosin stains are shown in Figure E1.

Tissue Dissociation

The primary human lung cell dissociation protocol is available at https://www.protocols.io/view/702-b-3-urmc-htc-lung-tissue-digestion-sop-workshe-81wgbpbjqvpk/v1 and has been previously applied (34).

FACS

Enrichment of lung cell populations was performed using FACS as described previously (37). Briefly, cells were sorted sequentially to obtain populations of mixed immune cells, endothelial cells, epithelial cells, and mesenchymal cells (Figure 1).

Protein Extraction and Digestion

Proteins and lipids were extracted from the same samples of sorted lung cell populations using a multiomics extraction method called MPLEx (39).

Liquid Chromatography and Mass Spectrometry

Five hundred nanograms of trypsin digested peptides were analyzed using reverse phase separation (C18) using a Waters nanoACQUITY UPLC system interfaced with a Q Exactive Plus Orbitrap (Thermo Scientific) mass spectrometer. Peptides were separated over a 180-minute gradient. Liquid chromatography effluent was analyzed using a top 12 data-dependent acquisition mode. Raw data were deposited at MassIVE (https://massive.ucsd.edu) under accession number MSV000094189.

Transcriptomics

Transcriptomic assessment was performed using bulk RNA sequencing on sorted cells derived from the lung tissues collected from the same 11 pediatric donors as the proteomics survey.

Proteomics Data Analysis

Raw mass spectrometry data were analyzed using MaxQuant (Max Planck Institute of Biochemistry, version 1.6.0.16; https://www.maxquant.org/) against the *Homo sapiens* UniProt database including splicing variants (downloaded September 10, 2020) and standard MaxQuant contaminants. MaxLFQ was used for quantification. Default parameters were used except for a matching window of 1.5 minutes for match between run.

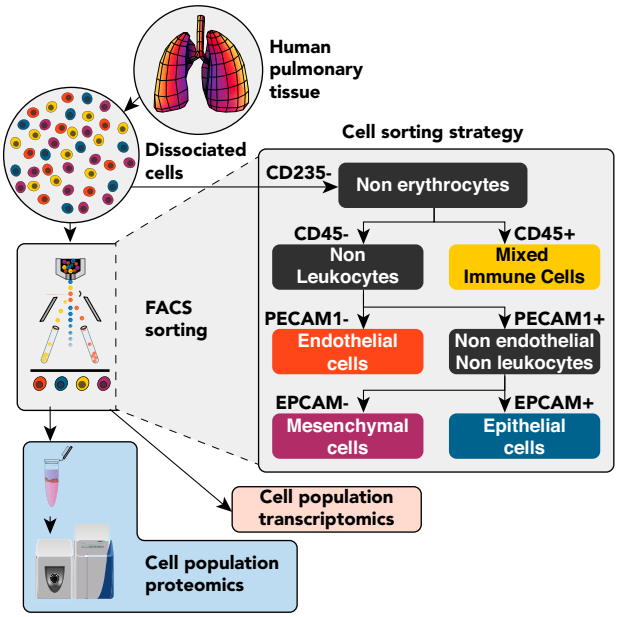

Figure 1. Schematic representation of human lung tissue dissociation, cell-sorting strategy, and cell population proteomics and transcriptomics. Dissociated cells from 11 donors were sequentially sorted for $CD45^+$ mixed immune cells, $PECAM1^+$ endothelial cells, $EpCAM^+$ epithelial cells, and $CD45^-$, $PECAM^-$, and $EPCAM^-$ mesenchymal cells. Sorted populations were divided into two tubes: one for MPLEx multiomic sample preparation and proteomic assay and the other for transcriptomics.

Proteomics and Transcriptomics Quantitative Analyses

Quantitative data were further processed using the homemade package RomicsProcessor version 1.0 (https://zenodo.org/records/10459578). Student's *t* tests and trend analyses were performed using linear and quadratic regression models to identify the analytes that exhibited significant changes in normalized abundances across time (i.e., age), as previously described (27). Enrichment analyses were performed using DAVID (40).

Immunohistochemistry from the Human Protein Atlas

Immunohistochemistry images were reproduced with authorization (Courtesy of Human Protein Atlas version 22) (41) (http://v22.proteinatlas.org/human proteome/tissue).

Immunofluorescence Verification of Alveolar Epithelial Cell Markers

The protocol generating multiplexed immunofluorescence images is adapted from the protocol available at https://www.protocols.io/view/814-1-multiplexed-immunofluorescence-phenocycler-f-c7nrzmd6.html. CD55 was imaged using immunofluorescence on deparaffinized and rehydrated slides.

RNA *In Situ* Hybridization

Using the high-throughput *in situ* hybridization experimental approach previously described (42) and adapted for human tissues, we collected images of gene expression from left upper lobe tissue blocks from the Biorepository for Investigation of Neonatal Diseases of the Lung. Probe sequences details are provided in the data supplement.

Results

Experimental Plan

Human lung samples were collected from 11 donors with ages ranging from 1 day to 8 years old. As depicted in Figure 1, for each donor, after cell dissociation using an enzymatic cocktail, cell types were sequentially sorted as previously described (37) and detailed in the Methods section (Figure 1). The four cell lineages sorted are the following: mixed immune cells ($CD235^-$ and $CD45^+$), endothelial cells ($CD235^-$, $CD45^-$, and $PECAM^+$), epithelial cells ($CD235^-$, $CD45^-$, $PECAM^-$, and $EPCAM^+$), and nonendothelial mesenchymal cells ($CD235^-$, $CD45^-$, $PECAM^-$, and $EPCAM^+$). Each sorted population for each donor was separated in two tubes; one was used for transcriptomics and the other for the MPLEx procedure to extract polar metabolites, proteins, and lipids from the exact same sample (39). Proteomics was performed by liquid chromatography–tandem mass spectrometry.

Cell Populations Have Distinct Proteomes Showing Functional Specializations

The proteomics profiling of the different cell populations can indicate the cellular functions achieved by each population and provide clues on how they contribute collaboratively to tissue functions. This study enabled the quantification of 5,540 proteins across cell populations and postnatal ages. As different cell populations achieve diverse functions, we anticipated that their proteome would diverge regardless of the age at donation. A principal-component analysis (PCA) confirmed that the proteomes mostly clustered on the basis of their population (i.e., endothelial, epithelial, mesenchymal, and mixed immune cells). Although epithelial and immune proteomes separated in a three-component PCA space, the proteomes of the endothelial and mesenchymal populations overlapped slightly (Figure 2A). To establish the function achieved by the different populations, we filtered the proteins that were either uniquely detected or more abundant in each population compared with all others ($P < 0.05$, Student's *t* test). These proteins were considered population markers (Figure 2B and *see* Dataset E1). The epithelial cells had the highest number of markers, with 1,492 uniquely detected or highly abundant markers. Endothelial cells and immune cells had 1,262 and 912 protein markers, respectively. For mesenchymal cells, we identified only 247 makers.

We performed functional enrichment analyses for the markers of each population using the DAVID bioinformatics resources (43) to evaluate the function performed by the different populations. Figure 2C depicts representative Gene Ontology (GO) terms enriched for biological processes. The integral list of GO terms and Kyoto Encyclopedia of Genes and Genomes pathways associated with the population markers is available in Dataset E1. These analyses corroborated anticipated functions of the different cell populations.

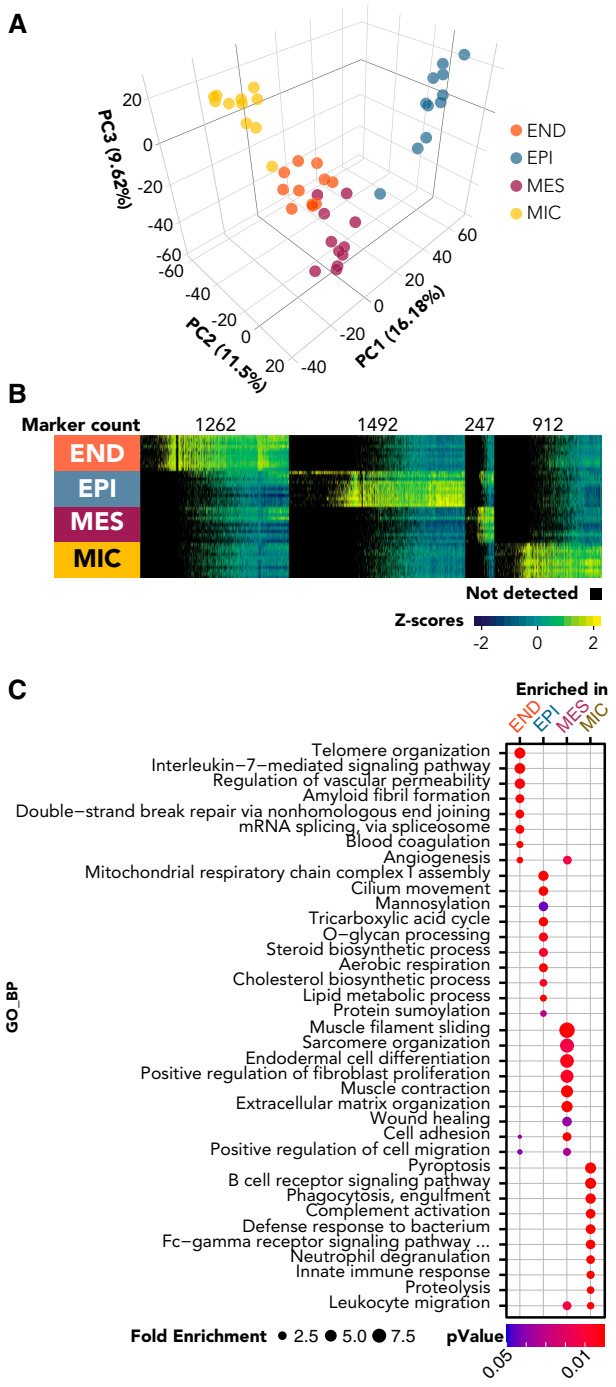

Figure 2. Proteomic profiling of sorted cell populations. (*A*) Principal-component analysis (PCA) shows clustering of proteins on the basis of population. Epithelial and immune proteomes separated in a three-component PCA space; the proteomes of the endothelial and mesenchymal populations overlapped slightly. (*B*) Uniquely abundant proteins of sorted cell population. (*C*) GO terms for BP enriched in each cell population on the basis of their fold changes and *P* values. BP = biological processes; END = endothelial cells; EPI = epithelial cells; GO = Gene Ontology; MES = mesenchymal cells; MIC = mixed immune cells; PC = principal component.

Endothelial cell markers were enriched in GO terms related to angiogenesis (GO:0001525), blood coagulation (GO:0007596) or other vascular function (e.g., GO:0043114, GO:0030948). Epithelial cells were enriched in GO terms related to aerobic respiration (e.g., GO:0032981, GO:0042776, GO:0009060). One of the main functions of alveolar type 2 (AT2) cells is the production of the lipid-rich alveolar surfactant (44), and the enrichment analysis revealed that the epithelial population was enriched in many lipid-related GO terms (e.g., GO:0006695, GO:0006629, GO:0006633). As a ciliated epithelium covers the surface of the airways (7, 8), various cilium-related terms were also enriched in the markers of this population (e.g., GO:0060271 GO:0003341, GO:0003351, GO:0060285). The immune population markers were associated with function related to defense against bacteria. Many immune processes were enriched in this population, including pyroptosis (GO:0070269), defense response to bacterium (GO:0042742), phagocytosis (GO:0006911), leukocyte migration (GO:0050900), neutrophil degranulation (GO:0043312), and inflammatory response (GO:0006954). Finally, the mesenchymal population was enriched in function related to contractile elements or muscle, likely because of the presence of smooth muscle cells, and potentially myofibroblasts, in this population (e.g., GO:0030049 GO:0045214, GO:0006936). Other GO terms enriched for this population are related to fibroblast functions in wound healing (GO:0042060) (45) and extracellular matrix deposition (e.g., GO:0030198, GO:0030199, GO:0007160) (46, 47). These enrichment analyses show that the cell-sorting method was largely effective but also revealed unexpected population-specific functions that will require further validation. For example, terms related to telomere, RNA splicing, and DNA repair were enriched in the endothelial population, which hypothetically may be linked to waves of proliferation expected in this cell population postnatally (48).

Identification of Cell Subpopulation–Distinguishing Cell Surface Protein Markers

Cell types and cell states are commonly resolved through protein markers, often

corresponding with functional properties of the cells. Immunophenotyping and sorting of cells on the basis of surface antigens is one of the most common uses of flow cytometry, to enable the identification and enriching of new or rare cellular populations. Identifying surface markers retained on dissociated cells is a prerequisite to selecting optimal protein targets for antibody-mediated cell sorting. To help identify cell surface markers potentially amenable to such applications, we extracted a list of 739 marker that have been either experimentally evidenced or *in silico* predicted to be localized at the surface of the human cells. To do so, we used experimental data from the subcellular Human Protein Atlas database (49) and predictions from the *in silico* human surfaceome (50) or observed to be localized at the cell surface (Figure 3A). Then, we used scRNA-seq datasets collected by the Molecular Atlas of Lung Development Program consortium (https://www.lungmap.net/cell-cards/) (3, 24) to identify 133 handpicked markers that may present some expression specificity at the cell type level, for those we provide their relative expression in different cell types as well as immunohistochemistry-stained images from the Human Protein Atlas (*see* Dataset E2 and Figure 3B). Some pan-population endothelial makers expressed in most endothelial cell types included TBHD, CDH5, and CD34. In addition, some markers associated with specific cell types were found. PRX and CA4 were shown to be expressed exclusively in capillaries (28), EDNRB was demonstrated to be distinctive in capillary type 2 cells (aerocytes) in a murine model (51), and ARCKR1 (also known as DARC) is expressed in murine venular endothelial cells (52). Epithelial markers suggested included the AT2 markers LAMP3 (53) and ABCA3 (54). CCDC170 was reported to be highly expressed in ciliated cells throughout the body and was also identified as candidate surface marker (55). Two previously known AT1 pneumocyte marker were identified in this population: AGER (also called RAGE) (53) and AQP4 (56). Mesenchymal markers included fibroblasts markers PDGFR-β, LMCD1 (57), and ADH1B (58). Finally, mixed immune cell surface markers were composed of the pan-leukocyte maker integrin αL (ITGAL also known as CD11a),

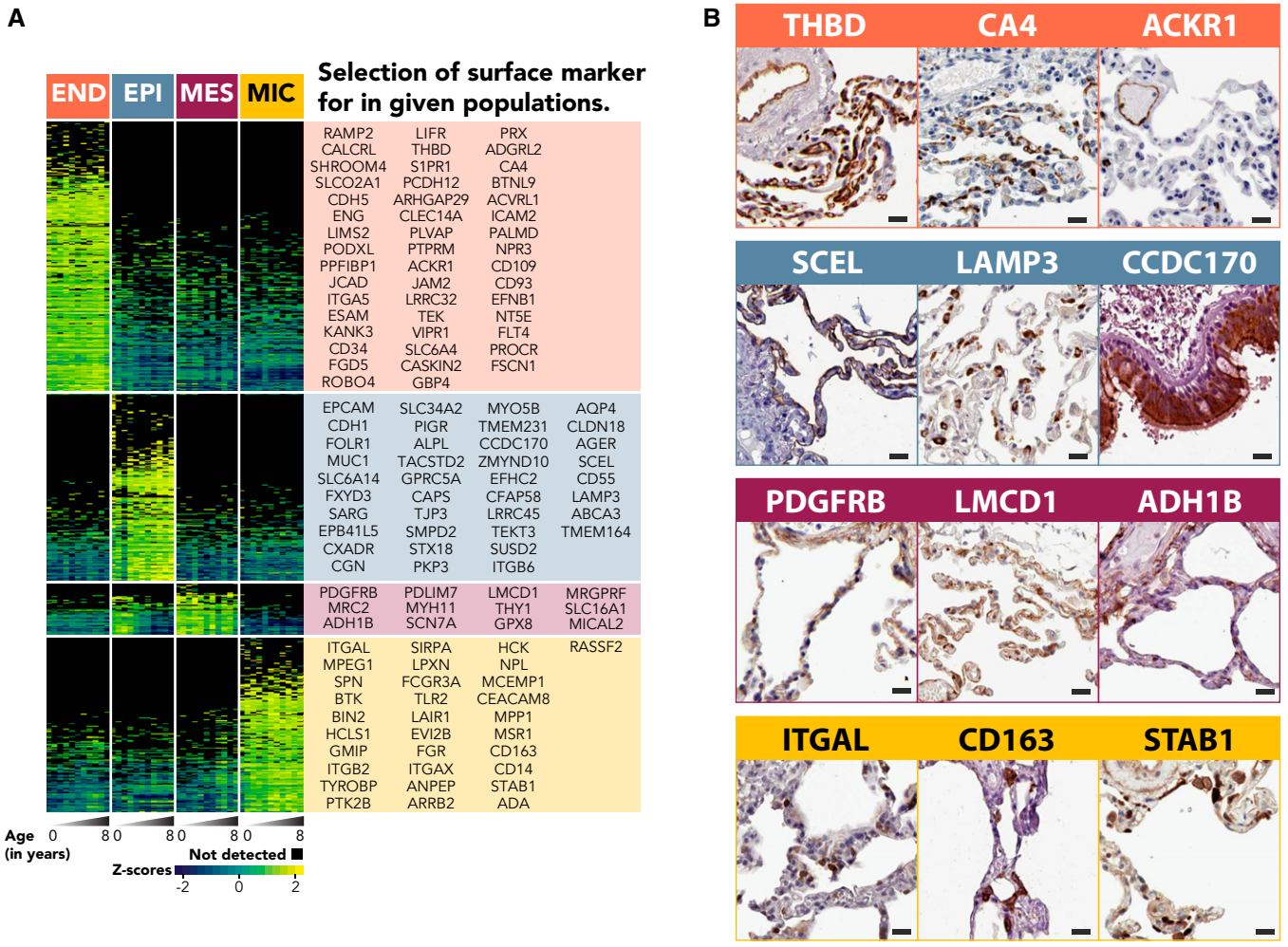

Figure 3. Identification and validation of cell surface markers. (*A*) Seven hundred thirty-nine cell population markers were identified that are localized predominantly at the cell surface. (*B*) Images from the Human Protein Atlas verify the localization of cell population surface proteins identified from the dataset. The first-row images of THBD, CA4, and ACKR1 represent END populations; the second-row images of SCEL, LAMP3, and CCDC170 represent EPI populations; the third-row images of PDGFR, LMCD1, and ADH1B represent MES populations; and the fourth-row images of ITGAL, CD163, and STAB1 represent MIC. Scale bars, 20 μm. Figure 3B used with permission by (41) and Human Protein Atlas; *see* www.proteinatlas.org.

the macrophage markers CD163 and CD14 (59), and STAB1 (stabilin-1), a scavenger receptor involved in leukocyte trafficking (60).

CD55 and SCEL (sciellin) RNA was found to be enriched in AT1 cells in scRNA-seq datasets (3, 24). To verify their tissue protein distribution, we performed immunostaining. CD55 (Figure 4A), also known as DAF (decay-accelerating factor), is a glycoprotein involved in the regulation of the complement pathway (61). It was found covering most of the surface of the alveolar space but not airway regions, consistent with AT1 gas exchanging interface. Some immune cells localized in the alveolar space (likely alveolar macrophages) also showed staining for CD55. SCEL (Figure 4B), a protein described as precursor of cornified envelope (62), was found uniquely in regions corresponding to AT1 cells. Regions immunostained with SCEL generally overlapped with luminal cell regions stained with RAGE, a known marker of AT1. SCEL was not at the apical surface of cells expressing SFTPC, a marker of AT2 cells. Further validations, using antibody-based methodologies, would be required confirm that the protein does not delocalize after being expressed.

We also performed RNA *in situ* hybridization for CD55 and known or newly identified endothelial markers (e.g., PALMD, PRX, ACKR1, THBD, FLT4; Figure 4C). CD55 was found selectively in alveolar regions in a similar fashion as in immunofluorescence. A signal is also seen in some airway-like regions that is so general that it is suspected to be nonspecific RNA detection. PALMD had a ubiquitous expression but with notably high expression in arteries surrounded with smooth muscle. Similar to the CD55 image, a signal was also detected in some airway-like regions (consistent with unspecific binding). As expected, the expression of PRX, a known marker of capillaries (28), was restricted to selective cells localized in the alveolar parenchyma. ACKR was present in vasculature around larger bronchioles, potentially systemic vessels, and "corner

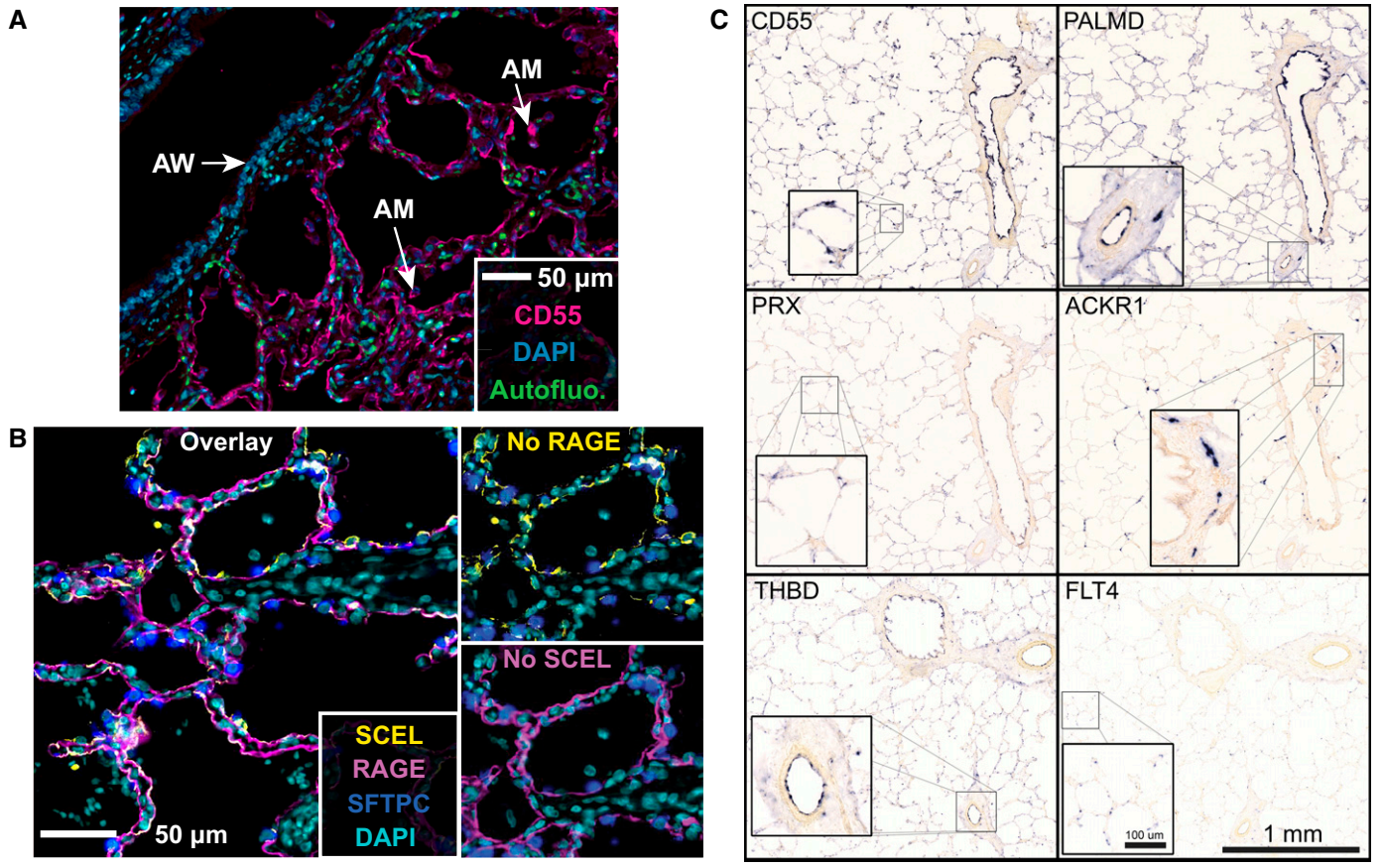

Figure 4. Validation of some of the cell markers by immunofluorescence and *in situ* hybridization. (*A*) Immunofluorescence showing the localization of CD55, arrows indicate AW, AMs performed on Biorepository for Investigation of Neonatal Diseases of the Lung (BRINDL) donor D129 (normal pathology, female, age 3 yr). Scale bars, 50 µm. (*B*) Multiplexed immunofluorescence showing the colocalization of AGER/RAGE and SCEL at the surface of alveoli, SFTPC is used as an AT2 marker performed on BRINDL donor D016 (normal pathology, male, age 9 yr). Scale bars, 50 µm. (*C*) Detection of representative mRNA probes by *in situ* hybridization. Most genes shown here express in endothelial cells, with epithelial CD55 included for comparison. PALMD is a general endothelial marker that is seen lining vessels (inset), in capillaries, and in the systemic venous endothelial cells. PRX is found in capillaries only (inset). Scale bars: 1 mm; 100 µm (inset). ACKR1 is found in the systemic venous endothelial cells in the bronchial vascular plexus (inset) and unexpectedly in the alveolar septa. THBD is seen in vessels (inset) and capillaries. FLT4 is found only in capillaries (inset) and lymphatic cells. Off-target/nonspecific uptake of some mRNA probes is seen in the ciliated lining of bronchioles. *In situ* hybridization was performed on serial sections of a single tissue block from donor D011 (normal pathology, female, age 20 mo). AM = alveolar macrophage; Autofluo. = autofluorescence; AW = airway; RAGE = receptor for advanced glycation endproducts; SCEL = sciellin.

vessels in alveoli" but not in smooth muscle surrounded arteries. TBHD was localized in regions surrounded by smooth muscle surrounded regions consistent with arterial endothelial cells, in the alveolar parenchyma likely in capillaries, and in interstitium surrounding medium arteries and airways consistent with the expected localization of endothelium of veins or lymphatic capillaries. Finally, FLT4 was found in regions consistent with capillaries (e.g., both lymphatic and alveolar capillaries). Overall, these validations demonstrate that protein markers with differential cellular distribution that could be used for cell typing or cell sorting of specific populations. Similar coupled analyses of the proteomic and scRNA-seq datasets with spatial transcriptomic and immunochemical imaging will be used for the discovery of usable, cell-distinguishing markers for appropriate for different applications.

Cell Population Proteomes Are Dynamic during Postnatal Development

We recently established the proteome of the developing lung for donors of the same age range as in the present study (27). That study, performed from tissue blocks comprising diverse cell types, did not allow us to evaluate the contribution of the different cell populations to the modulations observed. Here, we sorted the four broad cell populations composing the pulmonary tissue and confirmed that they had distinct proteomes. We separated the proteome analysis by cell population to evaluate population-specific age-dependent trends. First, a PCA was performed for each cell population (Figure 5A). For most populations, the proteome of the samples with age less than 30 days segregated from the proteome of older donors, demonstrating that the cell population proteomes are temporally dynamic. The exception to this is that the immune proteome of a 4-month-old donor (D024, the youngest of the other donors) was not separated from the proteome of donors less than 1 month old on the first principal component.

Proteome Trend Analysis and Pathway Modulation during Postnatal Development

To explore further the temporal changes occurring in each population, for each quantifiable protein (with abundance value in at least one-third of the samples of a given population), we performed trend analyses fitting linear and quadratic models to the protein abundances on the basis of their age. On the basis of the trend-fitting *P* values, each quantified protein was categorized into one of five protein sets: linear decreasing, linear increasing, quadratic concave, quadratic convex, and unspecific. The abundance of the proteins in the unspecific protein set did not significantly fit either a linear or a quadratic model.

We performed enrichment analysis on the four specific protein sets to understand the functional dynamic for each cell population. The complete list of enriched GO terms and Kyoto Encyclopedia of Genes and Genomes and Reactome pathways is provided in Dataset E2. For the biological process GO terms, we grouped the enriched terms in a table to identify pathways following similar or opposite trends in different cell populations (*see* Dataset E2). The GO terms related to mitochondrial respiration (GO:0042776, GO:0009060, GO:0006120, and GO:0032981) were enriched in the linear decreasing protein set for the four cell populations, suggesting a higher energy demand in early life, when higher cellular proliferation is expected (Figure 5B). These GO terms were also enriched in the set of proteins linearly increasing in epithelial cells. The other terms that were enriched in the protein set decreasing in all four cell populations were GO:0043044 (ATP-dependent chromatin remodeling), GO:0000398 (mRNA splicing, via spliceosome), and GO:0045944 (positive regulation of transcription from RNA polymerase II promoter). The reasons for the higher abundance of spliceosomal subunits postnatally compared with more mature cell populations remains to be explored. Moreover, the proteins related to proliferation (e.g., growth factor signaling, transcription, spliceosome), epigenetics (e.g., chromatin remodeling), respiration (e.g., mitochondrial functions, aerobic respiration), and developmental regulation (e.g., angiogenesis, signal transduction, GTPase activity) decreased with age during postnatal development. The GO term GO:0050821 (protein stabilization) showed a pattern of linear increase across the cell types. The other terms that were enriched in the protein set increasing in one or more cell populations include GO:0022617 (extracellular matrix disassembly), GO:0006879 (cellular iron ion homeostasis), GO:0006508 (proteolysis), and GO:0016241 (regulation of macroautophagy). GO terms for proteins related to cholesterol homeostasis (GO:0042632) and fatty acid β oxidation or metabolic process were linear increasing in all cell types except epithelial cells (GO:0006635 and GO:0001676). These results show the complexity in cellular response at different stages of early life and within each population.

Cell Population–Specific Proteomics Disambiguate Abundance Modulation for Specific Pathways

We previously performed similar trend analyses for the proteome of bulk pulmonary tissues (24). For some pathways, interpretation was difficult because of the convoluted nature of the data generated (e.g., originating from a mixture of cell populations). An example that we highlighted with bulk tissue proteomics (27) is the proteasome. In the whole-lung tissue, most of the proteasomal proteins followed decreasing trends, including the proteolytic subunits β1 (PSB1), β2 (PSB2), and β5 (PSB5). However, the immunoproteasome proteolytic subunits iβ1 (PSB9), iβ2 (PSB10), and iβ5 (PSB8) and the PA28-αβ particle followed an inverse trend (27). In the population-specific proteome presented here, we first compared the abundance of proteasomal subunits regardless of age in the different cell populations (Figure 5C and *see* Dataset E3). We generally observed that proteasome subunits were more abundant in the endothelial and immune cell populations compared with the epithelial and mesenchymal ones. Thirty-four of 40 proteasomal proteins quantified had a lower abundance in the epithelial population compared with the endothelial population ($P < 0.05$, *t* test) and 33 of 40 compared with the immune population. The largest abundance differences were observed for the immunoproteasome-specific proteins iβ1 (PSB9), iβ2 (PSB10), and iβ5 (PSB8) and the PA28-αβ particle proteins PSME1 and PSME2. Similarly, 29 of 40 and 32 of 40 subunits were significantly less abundant in mesenchymal cells compared with endothelial and immune cells, respectively. The proteasome subunits generally follow a decreasing abundance trend with age in the epithelial and mesenchymal populations, with 30 of 40 and 21 of 40 proteins fitted to a decreasing trend (linear model fitting $P < 0.05$). Fewer proteins were modulated in endothelial and immune cells. However, the subunits that were significantly modulated

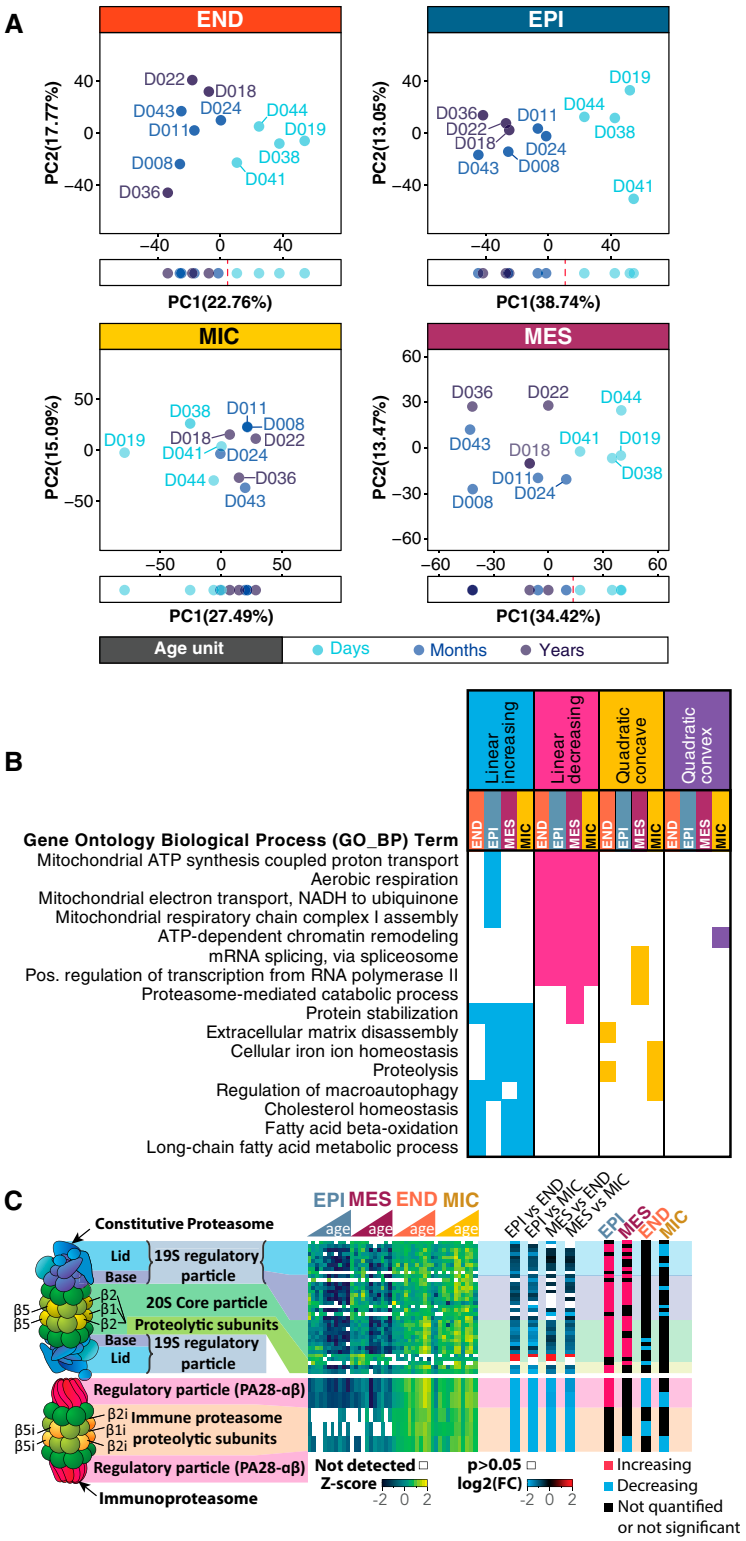

Figure 5. Proteomics changes in individual cell population by age. (*A*) Principal-component analysis showing age-dependent protein clustering within each sorted cell population. The proteome of the samples with age less than 30 days segregated from the proteome of older donors, with an exception for the immune proteome of a 4-month-old donor, D024. (*B*) Age-dependent patterns of pathway enrichment in sorted cell population on the basis of the proteomics results. A complete table and list of proteins contributing to each Gene Ontology biological process term are provided in Dataset E3. (*C*) The scaled abundance of proteasomal subunits has been represented in each cell population regardless of age. Note that proteasome subunits were more abundant in the END and MIC populations compared with the EPI and MES. Log_2(FC) and linear trend direction are displayed only for proteins with $P < 0.05$ (Student's *t* test) or $P < 0.05$ (linear fitting), respectively. FC = fold change.

were mostly immunoproteasome subunits and followed an opposite trend. Six of 40 proteasomal proteins increased with age in endothelial cells and 9 of 40 in the immune population. Taken together, these results suggest that proteasomal activity, in particular immunoproteasome activity, is higher and increases with age in immune and endothelial cells. Conversely, the abundance of the proteasome is faint in epithelial and mesenchymal cells with age.

Protein and Transcript Anticorrelated Functions Highlight Potential Post-transcriptional Regulation

RNA sequencing was performed on the sorted cell populations, with the goal of identifying the proteins and processes regulated at the post-transcriptional level during postnatal development. Overall, a limited number of specific transcript–protein pairs displayed directly opposite abundance trends: 27 in endothelial cells, 33 in epithelial cells, 14 in mesenchymal cells, and 3 in mixed immune cells. Therefore, to discern differences between proteins and mRNAs, we compared the enriched GO biological process terms exhibiting distinct trends in proteomics and transcriptomics data. Endothelial cells exhibited a greater number of enriched biological processes following the opposite trend (whether linear or quadratic), with 47 terms presenting an inverse trend in RNA and protein, suggesting that the highest degree of post-transcriptional regulation may occur in endothelial cells. Epithelial cells had 28 terms with opposite trends, mixed immune cells had 15, and mesenchymal cells had 8 (*see* Dataset E4). For simplicity, only a few processes showing opposite linear trends in association with age are described below.

Processes related to antigen processing and presentation (GO:0019882, GO:0002503) and T-cell activation (GO:0050870) were associated with different trends both in epithelial and endothelial cells. For instance, in endothelial cells, whereas the transcripts associated with these immune processes decreased with age, the proteins increased over time (Figure 6 and *see* Dataset E4). Conversely, in epithelial cells, although the transcripts associated with these immune processes increased with age, the proteins decreased. Interestingly, in mesenchymal cells, both RNA and proteins associated with antigen processing and presentation increased over time. These results suggest a complex and cell population–specific post-transcriptional regulation of antigen processing and presentation. In endothelial cells the processes DNA recombination (GO:0006310), positive regulation of DNA reparation (GO:2000781), and regulation of nucleotide-excision repair (GO:2000819) were associated with RNA increasing in abundance, while the proteins decreased with age. In epithelial cells, hematopoietic stem cell differentiation (GO:0060218) and DNA topological changes (GO:0006265) were associated with increased RNAs and decreased proteins with age. In mesenchymal cells, GO terms associated with response to toxic substances (GO:0009636 and GO:0010033) were associated with linearly increasing RNAs but decreasing proteins with age. Finally, in immune cells, the terms response to LPS (GO:0032496) and cellular calcium ion homeostasis (GO:0006874) followed the opposite temporal trend, with the RNA decreasing and the proteins increasing with age. These processes, playing pivotal roles in immunity and DNA integrity, seemed to be particularly regulated at the post-transcriptional level.

Discussion

Recently, we conducted deep proteome profiling of human lung tissue blocks spanning from birth to 8 years of age (27). In that study, we identified transcription factors modulated during postnatal development and confirmed the existence of substages of pulmonary postnatal development. Moreover, the research provided insights into the intricate remodeling of the extracellular matrix, as the alveolar parenchyma densifies concurrently with alveolar septal thinning. Because it was performed on bulk tissue, that study was unable to unambiguously distinguish the relative contributions of different cell populations. In the present study, the use of cell sorting with proteomics and

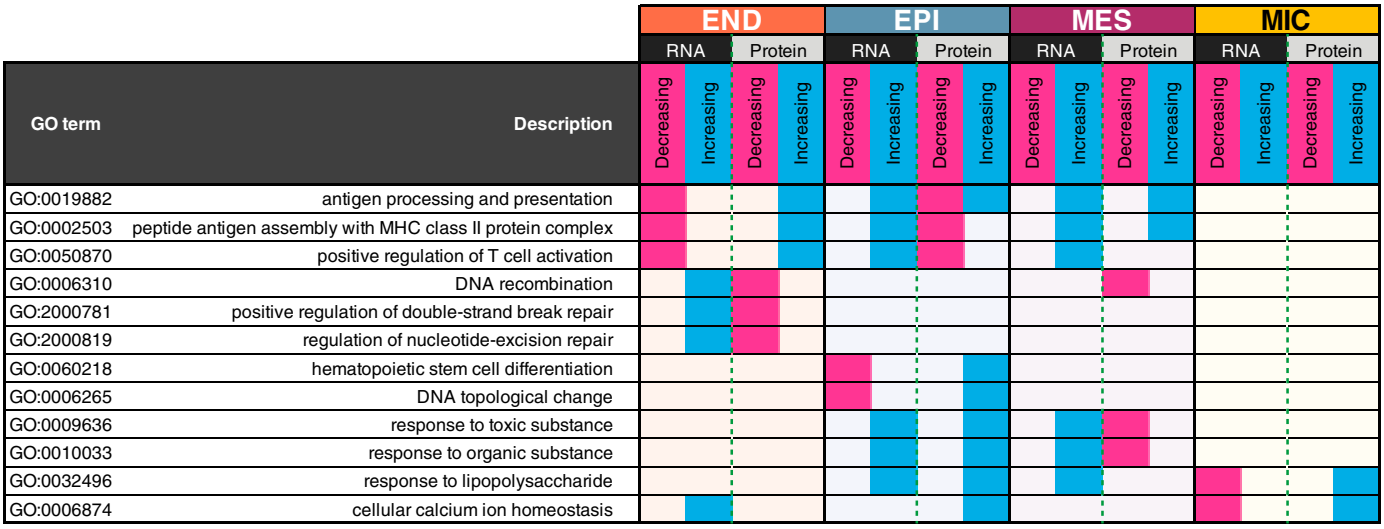

Figure 6. Comparison of proteomics and transcriptomics pathway enrichment. This table shows GO biological process terms associated with RNA and proteins following opposite temporal linear trend (increasing or decreasing with age) for at least one cell population (orange, END; blue, EPI; purple, MES; yellow, MIC). For each specific enriched GO term, pink indicates RNAs or proteins that decrease with age, and blue indicates RNAs or proteins that increase with age.

transcriptomics enabled the investigation of different cell populations to the postnatal pulmonary proteome. Pan-population markers have been identified for different lung cell lineages, including epithelial cells (e.g., EpCAM), endothelial cells (e.g., CDH5, PECAM1), or immune cells (e.g., CD45). The markers available for sorting specific pulmonary cell types are rather limited, except for immune cells, as historically, the immunology community heavily invested in identifying cellular markers (63, 64). Recently, Travaglini and colleagues used pan-population and immune cell sorting to broaden their ability to detect rarer cell types or subtypes (4). We perceived the limited number of available markers as both a constraint and an opportunity. While we opted to use pan-population markers to sort cells into the four main lineages of pulmonary cells (i.e., endothelial, epithelial, immune, and mesenchymal cells), we used the data generated to identify cell type–specific markers (and cell surface markers) on the basis of scRNA-seq surveys. Some of the identified markers were used to develop a highly multiplexed immunofluorescence panel (the CODEX panel; https://www.protocols.io/view/813-1-multiplexed-immunofluorescence-phenocycler-f-6qpvr38dpvmk/v2), enabling the visualization, localization, and counting of specific cell types in pulmonary tissues. In a recent publication by Dylag and colleagues, we used this multiplex panel to quantify cell types in both control subjects and patients with bronchopulmonary dysplasia (65). Knowing protein-level cellular markers allows us to infer shifts in cell-type proportions from the proteome of tissue blocks (bulk proteome). The shifts observed were systematically confirmed by multiplexed immunofluorescence imaging cell counting. More important, we anticipate that the list of protein markers generated here will aid in identifying cell-sorting markers useful for extracting specific cell populations relevant to disease, either to obtain deep omics profiling of these cell types or to generate disease models such as organoids (66, 67) or organs-on-a-chip (68). We believe that the identification of sorting markers is a prerequisite for conducting in-depth omics studies, such as RNA sequencing, proteomics, or lipidomics, to explore the perturbations within cell relationships. Example of relationships that could be explored include the relationship between capillary 1 cells (also known as aerocytes) (3, 51) and AT1 cells that may be perturbed in bronchopulmonary dysplasia (69) or the relationship between fibroblasts and AT2 cells perturbed in pulmonary fibrosis (70, 71). As the cells were sorted using broad markers, it is possible that some cell types not presenting these markers end up being present in an unexpected, sorted cell population. Careful validation of markers to be used for imaging or sorting will be necessary prior to its use for biological interpretations.

The use of sorted cell proteomics helped discern the contributions of different cell populations to specific pathway-level changes, which are observed in bulk omics analyses. The examination of the abundance changes of the different component of proteasome is a perfect illustration of this. In bulk proteomics data, the core proteasome was decreasing through development and the immunoproteasome components increased. Here, the deconvolution of the proteome into lineage allows us a better interpretation. Our observations reveal that proliferative epithelial and mesenchymal cells exhibit elevated proteasome expression compared with their quiescent counterparts, likely reflecting higher protein turnover rates in proliferative cells. Interestingly, the proteasome was constitutively present in endothelial and immune cells regardless of age. Although the enhanced degradation of abnormal or exogenous proteins in immune cells was expected, the similar pattern in endothelial cells was unforeseen. Speculatively, pulmonary endothelial cells, such as Cap1 (also known as aerocytes), may require a functional proteasome to mitigate oxidative stress associated with their oxygen transport function (72). More generally, we hypothesized that the GO terms and pathways associated with proteins decreasing with age might be linked to the reduction of cellular proliferation with age. Therefore, it is plausible that the reactivation of these pathways is crucial during pulmonary tissue repair processes. For instance, glycolysis and oxidative phosphorylation, vital for ATP production, decline with age across all cell populations. Recent evidence highlights the upregulation of these pathways as essential for AT2 differentiation into AT1 cells during alveolar tissue reparation (73). Similarly, spliceosome proteins exhibit high abundance postnatally but gradually decrease across all cell populations. A study carried out in a murine model revealed that proliferative "migratory" AT2 cells express higher amounts of spliceosome components compared with quiescent AT2 cells (74). On the other hand, pathways with protein abundance increasing with age likely reflect the specialization of cell types composing the sorted populations. A current limitation of our study is that we did not perform deep characterization of specific cell type/state, but we truly hope that the markers identified will enable us to perform sorted cell-type studies in the near future.

Finally, we compared the proteome with the transcriptome of the same cell populations. This exercise was challenging, as only a few protein transcript pairs were modulated concurrently in the two datatypes, because of either a difference in abundance variations of proteins and RNAs or real difference between transcriptional and translational regulatory events. We therefore focused on pathway-level modulation and identified a few pathways that seemed to be post-transcriptionally regulated. For example, in both endothelial and epithelial cells, antigen presentation was found to follow opposite trends at the transcript and protein levels. As the lung epithelium directly interacts with the external environment, it is logical for these cells to be critical regulators of immunity. The expression of MHC-II by AT2 cells was shown essential for an optimal response against viral infections (74). Lung endothelial cells have also been shown to express antigens (27, 75). However, how antigen presentation is regulated in epithelial and endothelial cells remains unclear (76), and our data suggest that post-transcriptional regulation may be essential in this process, which would require further investigations. Single-cell proteomics and spatial proteomics are progressing at rapid pace, but the throughput of these technologies remains limiting (34, 77). The results of the work performed here could be compared with spatially resolved proteomics to identify cell population markers uniquely localized in specific structures (e.g., epithelial cells in the alveolar parenchyma). The use of such a strategy would make it possible to refine the list to more focused cell types.

Conclusions

In this study, we conducted a comprehensive investigation into the dynamics of the proteome and transcriptome of the four main cell populations composing human lung during the early postnatal period.

The proteome profiles of endothelial, epithelial, immune, and mesenchymal cell populations were distinctive. We confirmed the relative specificity of the newly identified AT1 markers CD55 and SCEL. We also showed the differential spatial expression of multiple endothelial markers (e.g., PALMD, PRX, ACKR1, THBP, FLT4). Through trend analysis, we gained valuable insights into the temporal patterns of protein expression. We anticipate that this resource will serve researchers and clinicians to generate new hypotheses on postnatal development augmented by identification and use cell type–specific markers.

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

Acknowledgment: Donor tissue was supplied through the U.S. Department of Health and Human Services Donation and Transplantation Network, the National Disease Research Interchange, and the International Institute for Advancement of Medicine, organizations that link donor families to the scientific community. The authors are grateful for the generosity of the donor families and honor their loss. The authors acknowledge Charles Ansong for initiating this work.

References

1. Whitsett JA, Kalin TV, Xu Y, Kalinichenko VV. Building and regenerating the lung cell by cell. *Physiol Rev* 2019;99:513–554.
2. Cupello C, Hirasawa T, Tatsumi N, Yabumoto Y, Gueriau P, Isogai S, *et al.* Lung evolution in vertebrates and the water-to-land transition. *Elife* 2022; 11:e77156.
3. Sun X, Perl AK, Li R, Bell SM, Sajti E, Kalinichenko VV, *et al.*; NHLBI LungMAP Consortium. A census of the lung: CellCards from LungMAP. *Dev Cell* 2022;57:112–145.e2.
4. Travaglini KJ, Nabhan AN, Penland L, Sinha R, Gillich A, Sit RV, *et al.* A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* 2020;587:619–625.
5. Tsuchiya T, Doi R, Obata T, Hatachi G, Nagayasu T. Lung microvascular niche, repair, and engineering. *Front Bioeng Biotechnol* 2020;8:105.
6. Burgstaller G, Sengupta A, Vierkotten S, Preissler G, Lindner M, Behr J, *et al.* Distinct niches within the extracellular matrix dictate fibroblast function in (cell free) 3D lung tissue cultures. *Am J Physiol Lung Cell Mol Physiol* 2018;314:L708–L723.
7. Kadur Lakshminarasimha Murthy P, Sontake V, Tata A, Kobayashi Y, Macadlo L, Okuda K, *et al.* Human distal lung maps and lineage hierarchies reveal a bipotent progenitor. *Nature* 2022;604:111–119.
8. Basil MC, Cardenas-Diaz FL, Kathiriya JJ, Morley MP, Carl J, Brumwell AN, *et al.* Human distal airways contain a multipotent secretory cell that can regenerate alveoli. *Nature* 2022;604:120–126.
9. Clement A, Eber E. Interstitial lung diseases in infants and children. *Eur Respir J* 2008;31:658–666.
10. Gough A, Linden M, Spence D, Patterson CC, Halliday HL, McGarvey LP. Impaired lung function and health status in adult survivors of bronchopulmonary dysplasia. *Eur Respir J* 2014;43:808–816.
11. Glasser SW, Hardie WD, Hagood JS. Pathogenesis of interstitial lung disease in children and adults. *Pediatr Allergy Immunol Pulmonol* 2010; 23:9–14.
12. Barst RJ, Ertel SI, Beghetti M, Ivy DD. Pulmonary arterial hypertension: a comparison between children and adults. *Eur Respir J* 2011;37: 665–677.
13. Lee DD, Park SJ, Zborek KL, Schwarz MA. A shift from glycolytic and fatty acid derivatives toward one-carbon metabolites in the developing lung during transitions of the early postnatal period. *Am J Physiol Lung Cell Mol Physiol* 2021;320:L640–L659.
14. Dautel SE, Kyle JE, Clair G, Sontag RL, Weitz KK, Shukla AK, *et al.* Lipidomics reveals dramatic lipid compositional changes in the maturing postnatal lung. *Sci Rep* 2017;7:40555.
15. Kyle JE, Clair G, Bandyopadhyay G, Misra RS, Zink EM, Bloodsworth KJ, *et al.* Cell type-resolved human lung lipidome reveals cellular cooperation in lung function. *Sci Rep* 2018;8:13455.
16. Moghieb A, Clair G, Mitchell H, Kitzmiller JA, Zink EM, Kim YM, *et al.* Time-resolved proteome profiling of normal lung development. *Am J Physiol Lung Cell Mol Physiol* 2018;315:L11–L24.
17. Clair G, Piehowski PD, Nicola T, Kitzmiller JA, Huang EL, Zink EM, *et al.* Spatially-resolved proteomics: rapid quantitative analysis of laser capture microdissected alveolar tissue samples. *Sci Rep* 2016;6: 39223.
18. Negretti NM, Plosa EJ, Benjamin JT, Schuler BA, Habermann AC, Jetter CS, *et al.* A single-cell atlas of mouse lung development. *Development* 2021;148:dev199512.
19. Kho AT, Bhattacharya S, Tantisira KG, Carey VJ, Gaedigk R, Leeder JS, *et al.* Transcriptomic analysis of human lung development. *Am J Respir Crit Care Med* 2010;181:54–63.
20. Bhattacharya S, Myers JL, Baker C, Guo M, Danopoulos S, Myers JR, *et al.* Single cell transcriptomic profiling identifies molecular phenotypes of newborn human lung cells [preprint]. bioRxiv; 2021 [accessed 2023 Dec 1]. Available from: https://www.biorxiv.org/content/10.1101/2020.06.16.156042v2.
21. Beauchemin KJ, Wells JM, Kho AT, Philip VM, Kamir D, Kohane IS, *et al.* Temporal dynamics of the developing lung transcriptome in three common inbred strains of laboratory mice reveals multiple stages of postnatal alveolar development. *PeerJ* 2016;4:e2318.
22. Duong TE, Wu Y, Sos BC, Dong W, Limaye S, Rivier LH, *et al.* A single-cell regulatory map of postnatal lung alveologenesis in humans and mice. *Cell Genom* 2022;2:100108.
23. Sikkema L, Ramirez-Suastegui C, Strobl DC, Gillett TE, Zappia L, Madissoon E, *et al.*; Lung Biological Network Consortium. An integrated cell atlas of the lung in health and disease. *Nat Med* 2023; 29:1563–1577.
24. Guo M, Morley MP, Jiang C, Wu Y, Li G, Du Y, *et al.*; NHLBI LungMAP Consortium. Guided construction of single cell reference for human and mouse lung. *Nat Commun* 2023;14:4566.
25. Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, *et al.* Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv* 2020;6: eaba1983.
26. Ding J, Ahangari F, Espinoza CR, Chhabra D, Nicola T, Yan X, *et al.* Integrating multi-omics longitudinal data to reconstruct networks underlying lung development. *Am J Physiol Lung Cell Mol Physiol* 2019;317:L556–L568.
27. Clair G, Bramer LM, Misra R, McGraw MD, Bhattacharya S, Kitzmiller JA, *et al.* Proteomic analysis of human lung development. *Am J Respir Crit Care Med* 2022;205:208–218.
28. Schupp JC, Adams TS, Cosme C Jr, Raredon MSB, Yuan Y, Omote N, *et al.* Integrated single-cell atlas of endothelial cells of the human lung. *Circulation* 2021;144:286–302.
29. Raredon MSB, Adams TS, Suhail Y, Schupp JC, Poli S, Neumark N, *et al.* Single-cell connectomic analysis of adult mammalian lungs. *Sci Adv* 2019;5:eaaw3851.
30. Angelidis I, Simon LM, Fernandez IE, Strunz M, Mayr CH, Greiffo FR, *et al.* An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat Commun* 2019;10:963.
31. Guo M, Du Y, Gokey JJ, Ray S, Bell SM, Adam M, *et al.* Single cell RNA analysis identifies cellular heterogeneity and adaptive responses of the lung at birth. *Nat Commun* 2019;10:37.
32. Zhu Y, Clair G, Chrisler W, Shen Y, Zhao R, Shukla A, *et al.* Proteomic analysis of single mammalian cells enabled by microfluidic nanodroplet sample preparation and ultrasensitive nanoLC-MS. *Angew Chem Int Ed Engl* 2018;57:12370–12374.
33. Budnik B, Levy E, Harmange G, Slavov N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol* 2018;19:161.
34. Woo J, Williams SM, Markillie LM, Feng S, Tsai CF, Aguilera-Vazquez V, *et al.* High-throughput and high-efficiency sample preparation for single-cell proteomics using a nested nanowell chip. *Nat Commun* 2021;12:6246.

35. Gatto L, Aebersold R, Cox J, Demichev V, Derks J, Emmott E, *et al.* Initial recommendations for performing, benchmarking and reporting single-cell proteomics experiments. *Nat Methods* 2023;20:375–386.
36. Du Y, Clair GC, Al Alam D, Danopoulos S, Schnell D, Kitzmiller JA, *et al.* Integration of transcriptomic and proteomic data identifies biological functions in cell populations from human infant lung. *Am J Physiol Lung Cell Mol Physiol* 2019;317:L347–L360.
37. Bandyopadhyay G, Huyck HL, Misra RS, Bhattacharya S, Wang Q, Mereness J, *et al.* Dissociation, cellular isolation, and initial molecular characterization of neonatal and pediatric human lung tissues. *Am J Physiol Lung Cell Mol Physiol* 2018;315:L576–L583.
38. Gaddis N, Fortriede J, Guo M, Bardes EE, Kouril M, Tabar S, *et al.* LungMAP portal ecosystem: systems-level exploration of the lung. *Am J Respir Cell Mol Biol* 2024;70:129–139.
39. Nakayasu ES, Nicora CD, Sims AC, Burnum-Johnson KE, Kim YM, Kyle JE, *et al.* MPLEx: a robust and universal protocol for single-sample integrative proteomic, metabolomic, and lipidomic analyses. *mSystems* 2016;1:e00043-16.
40. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44–57.
41. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, *et al.* Proteomics: tissue-based map of the human proteome. *Science* 2015;347:1260419.
42. Ljungberg MC, Sadi M, Wang Y, Aronow BJ, Xu Y, Kao RJ, *et al.* Spatial distribution of marker gene activity in the mouse lung during alveolarization. *Data Brief* 2018;22:365–372.
43. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, *et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* 2022;50:W216–W221.
44. Tang X, Snowball JM, Xu Y, Na C-L, Weaver TE, Clair G, *et al.* EMC3 coordinates surfactant protein and lipid homeostasis required for respiration. *J Clin Invest* 2017;127:4314–4325.
45. Ito Y, Correll K, Schiel JA, Finigan JH, Prekeris R, Mason RJ. Lung fibroblasts accelerate wound closure in human alveolar epithelial cells through hepatocyte growth factor/c-Met signaling. *Am J Physiol Lung Cell Mol Physiol* 2014;307:L94–L105.
46. Eickelberg O, Kohler E, Reichenberger F, Bertschin S, Woodtli T, Erne P, *et al.* Extracellular matrix deposition by primary human lung fibroblasts in response to TGF-beta1 and TGF-beta3. *Am J Physiol* 1999;276: L814–L824.
47. Liu X, Dai K, Zhang X, Huang G, Lynn H, Rabata A, *et al.* Multiple fibroblast subtypes contribute to matrix deposition in pulmonary fibrosis. *Am J Respir Cell Mol Biol* 2023;69:45–56.
48. Zanini F, Che X, Knutsen C, Liu M, Suresh NE, Domingo-Gonzalez R, *et al.* Developmental diversity and unique sensitivity to injury of lung endothelial subtypes during postnatal growth. *iScience* 2023;26: 106097.
49. Thul PJ, Lindskog C. The human protein atlas: a spatial map of the human proteome. *Protein Sci* 2018;27:233–244.
50. Bausch-Fluck D, Goldmann U, Muller S, van Oostrum M, Muller M, Schubert OT, *et al.* The *in silico* human surfaceome. *Proc Natl Acad Sci U S A* 2018;115:E10988–E10997.
51. Gillich A, Zhang F, Farmer CG, Travaglini KJ, Tan SY, Gu M, *et al.* Capillary cell-type specialization in the alveolus. *Nature* 2020;586: 785–789.
52. Thiriot A, Perdomo C, Cheng G, Novitzky-Basso I, McArdle S, Kishimoto JK, *et al.* Differential DARC/ACKR1 expression distinguishes venular from non-venular endothelial cells in murine tissues. *BMC Biol* 2017;15:45.
53. Little DR, Gerner-Mauro KN, Flodby P, Crandall ED, Borok Z, Akiyama H, *et al.* Transcriptional control of lung alveolar type 1 cell development and maintenance by NK homeobox 2-1. *Proc Natl Acad Sci U S A* 2019;116:20545–20555.
54. Rindler TN, Stockman CA, Filuta AL, Brown KM, Snowball JM, Zhou W, *et al.* Alveolar injury and regeneration following deletion of ABCA3. *JCI Insight* 2017;2:e97381.
55. Coan M, Rampioni Vinciguerra GL, Cesaratto L, Gardenal E, Bianchet R, Dassi E, *et al.* Exploring the role of fallopian ciliated cells in the pathogenesis of high-grade serous ovarian cancer. *Int J Mol Sci* 2018; 19:2512.
56. Kreda SM, Gynn MC, Fenstermacher DA, Boucher RC, Gabriel SE. Expression and localization of epithelial aquaporins in the adult human lung. *Am J Respir Cell Mol Biol* 2001;24:224–234.
57. Xie T, Wang Y, Deng N, Huang G, Taghavifar F, Geng Y, *et al.* Single-cell deconvolution of fibroblast heterogeneity in mouse pulmonary fibrosis. *Cell Rep* 2018;22:3625–3640.
58. Danopoulos S, Bhattacharya S, Mariani TJ, Al Alam D. Transcriptional characterisation of human lung cells identifies novel mesenchymal lineage markers. *Eur Respir J* 2020;55:1900746.
59. Dewhurst JA, Lea S, Hardaker E, Dungwa JV, Ravi AK, Singh D. Characterisation of lung macrophage subpopulations in COPD patients and controls. *Sci Rep* 2017;7:7143.
60. Patten DA, Shetty S. More than just a removal service: scavenger receptors in leukocyte trafficking. *Front Immunol* 2018;9:2904.
61. Sun X, Funk CD, Deng C, Sahu A, Lambris JD, Song WC. Role of decay-accelerating factor in regulating complement activation on the erythrocyte surface as revealed by gene targeting. *Proc Natl Acad Sci U S A* 1999;96:628–633.
62. Champliaud MF, Baden HP, Koch M, Jin W, Burgeson RE, Viel A. Gene characterization of sciellin (SCEL) and protein localization in vertebrate epithelia displaying barrier properties. *Genomics* 2000;70:264–268.
63. Hu Y, Liu C, Han W, Wang P. A theoretical framework of immune cell phenotypic classification and discovery. *Front Immunol* 2023;14: 1128423.
64. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, *et al.* CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res* 2019;47:D721–D728.
65. Dylag AM, Misra RS, Bandyopadhyay G, Poole C, Huyck HL, Jehrio MG, *et al.* New insights into the natural history of bronchopulmonary dysplasia from proteomics and multiplexed immunohistochemistry. *Am J Physiol Lung Cell Mol Physiol* 2023;325:L419–L433.
66. Sucre JM, Wilkinson D, Vijayaraj P, Paul M, Dunn B, Alva-Ornelas JA, *et al.* A three-dimensional human model of the fibroblast activation that accompanies bronchopulmonary dysplasia identifies Notch-mediated pathophysiology. *Am J Physiol Lung Cell Mol Physiol* 2016;310: L889–L898.
67. Hein RFC, Conchola AS, Fine AS, Xiao Z, Frum T, Brastrom LK, *et al.* Stable iPSC-derived NKX2-1^{+} lung bud tip progenitor organoids give rise to airway and alveolar cell types. *Development* 2022;149: dev200693.
68. Huh D, Leslie DC, Matthews BD, Fraser JP, Jurek S, Hamilton GA, *et al.* A human disease model of drug toxicity-induced pulmonary edema in a lung-on-a-chip microdevice. *Sci Transl Med* 2012;4:159ra147.
69. Tsikis ST, Hirsch TI, Fligor SC, Quigley M, Puder M. Targeting the lung endothelial niche to promote angiogenesis and regeneration: a review of applications. *Front Mol Biosci* 2022;9:1093369.
70. Moss BJ, Ryter SW, Rosas IO. Pathogenic mechanisms underlying idiopathic pulmonary fibrosis. *Annu Rev Pathol* 2022;17:515–546.
71. Adams TS, Marlier A, Kaminski N. Lung cell atlases in health and disease. *Annu Rev Physiol* 2023;85:47–69.
72. Kotamraju S, Matalon S, Matsunaga T, Shang T, Hickman-Davis JM, Kalyanaraman B. Upregulation of immunoproteasomes by nitric oxide: potential antioxidative mechanism in endothelial cells. *Free Radic Biol Med* 2006;40:1034–1044.
73. Wang Z, Wei D, Bin E, Li J, Jiang K, Lv T, *et al.* Enhanced glycolysis-mediated energy production in alveolar stem cells is required for alveolar regeneration. *Cell Stem Cell* 2023;30:1028–1042.e7.
74. Konkimalla A, Konishi S, Kobayashi Y, Kadur Lakshminarasimha Murthy P, Macadlo L, Mukherjee A, *et al.* Multi-apical polarity of alveolar stem cells and their dynamics during lung development and regeneration. *iScience* 2022;25:105114.
75. Claser C, Nguee SYT, Balachander A, Wu Howland S, Becht E, Gunasegaran B, *et al.* Lung endothelial cell antigen cross-presentation to CD8^{+}T cells drives malaria-associated lung injury. *Nat Commun* 2019;10:4241.
76. Kawasaki T, Ikegawa M, Kawai T. Antigen presentation in the lung. *Front Immunol* 2022;13:860915.
77. Park J, Yu F, Fulcher JM, Williams SM, Engbrecht K, Moore RJ, *et al.* Evaluating linear ion trap for MS3-based multiplexed single-cell proteomics. *Anal Chem* 2023;95:1888–1898.