




EDITORIAL

The Idiopathic Pulmonary Fibrosis Cell Atlas

 Nir Neumark,^{1,2*}  Carlos Cosme, Jr.,^{1*} Kadi-Ann Rose,¹ and  Naftali Kaminski¹

¹Section of Pulmonary, Critical Care, and Sleep Medicine, Yale University School of Medicine, New Haven, Connecticut; and

²Interdepartmental Program in Computational Biology and Bioinformatics, Yale University School of Medicine, New Haven, Connecticut

Submitted 21 September 2020; accepted in final form 22 September 2020

INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a fatal lung disease that is characterized by replacement of the normal lung anatomy with active remodeling and deposition of extracellular matrix (ECM) accompanied by a shift in lung cellular communities. The most common outcome of this debilitating disease is respiratory failure and subsequent death. The disease predominately afflicts the middle aged and elderly and has a median life expectancy after diagnosis of 3.8 yr. Although two FDA-approved drugs consistently slow down disease progression, lung transplantation is still the only curative treatment.

Historically, much of our understanding of pulmonary fibrosis has been derived from detailed pathological analysis of human lungs that elucidated the unique morphological characteristics of IPF and observations derived from animal models of disease. In the last decade insights derived from genetic studies and from transcriptomic studies of the human lung have given investigators a better appreciation of the complexity and extent of the molecular and cellular changes that determine the lung phenotype in IPF (11). However, a detailed description of all the cell-specific molecular changes in the human fibrotic lung was until recently still lacking.

The advent of single-cell RNA sequencing (scRNAseq) has afforded researchers the opportunity to comprehensively interrogate tissues at cell-level resolution (8). In contrast to the bulk RNA sequencing protocols that analyzed gene expression in a whole piece of tissue, which averaged the signal of multiple cells without distinction of the cellular source of the signal, scRNAseq profiles expression of each individual cell within a sample. This technology allows researchers to closely examine every single individual cell within a sequenced sample to identify their cell types, functions, and ultimately the molecular signatures and cellular interactions guiding their function within the context of that tissue and human disease. With such breadth and depth, scRNAseq is a powerful tool that is in particular appealing to better untangle the intricate multicellular complexity of the human IPF lung.

Several groups embarked independently on exploiting the power of scRNAseq within the context of IPF to create a comprehensive transcriptomic “map” of all human cells in patients with IPF (1–6, 10). The insights from these papers are substantial and range from establishing evidence of profibrotic attributes of distinct cell populations to identification of novel and ectopic cell populations in the human IPF lung. But the “big-

data” nature of scRNAseq poses distinct challenges; in a single experiment of a single sample, investigators can profile 5,000–10,000 cells, with a few thousand transcripts per cell, and usually investigators will run multiple samples. In the lung, if no enrichment techniques are used, we are dealing, based on one estimate, with ~40 known cell types including epithelial, endothelial, mesenchymal, and myelomonocytic populations, and maybe 10–15 new ones (9). Thus, a detailed description of the cell- and disease-specific changes in gene expression is beyond the scope of any single paper or the bandwidth and research interests of a single research group. And although the data are usually deposited in public data repositories, mining the data requires significant expertise in bioinformatics and computational biology, as well as access to powerful computer systems and storage systems. Aiming to make IPF scRNAseq data accessible to the scientific and medical community, we decided to create an easy-to-use data dissemination portal called the IPF Cell Atlas (www.ipfcellatlas.com). With this interactive and easy-to-use web tool, users have access to a variety of embedded visualization tools that allow exploration of several independently procured single-cell data sets on IPF. We have designed this website to be easily accessible and intuitive for any user, regardless of their technical background or computational skill. We believe that improving accessibility to the newest and most advanced data on IPF will have a profound impact on our understanding of IPF and will accelerate development of therapeutic solutions for this devastating disease.

SINGLE-CELL STUDIES INCLUDED IN THE IPF CELL ATLAS

Here, we focus our attention on the main findings of four groups whose data are currently available on the IPF Cell Atlas (Table 1). In 2018, Reyfman et al. (6) applied scRNAseq to eight control lung samples and eight pulmonary fibrosis explants, of which four were IPF, two systemic sclerosis, one polymyositis, and one chronic hypersensitivity pneumonitis. Single-cell sequencing of their 16 samples returned a data set of 76,070 cells that clustered into 16 resident lung cell types based on canonical markers and performed bulk RNA seq and histological validations. The authors identified a novel fibrotic lung macrophage population that was found only in patients with pulmonary fibrosis. These cells expressed profibrotic molecules such as CHI3L1, MARCKS, IL1RN, PLA2G7, MMP9, and SPP1 as well as MAFB1 and SIGLEC1. In addition, they provided evidence that supported previous findings of increased epithelial cell heterogeneity within the IPF lung (10). A subsequent paper by Morse et al. (4) also reported the existence of this profibrotic alveolar macrophage in IPF patients. Their

* N. Neumark and C. Cosme Jr. contributed equally to this work.
Correspondence: N. Kaminski (naftali.kaminski@yale.edu).

Table 1. *Studies included in www.IPFCellAtlas.com*

Publication	Name on IPF Cell Atlas	GEO Accession No.	Cell Number	Subjects by Disease	Highlighted Findings
Reyfman et al., <i>AJRCCM</i> 2019 (6)	Misharin	GSE121611	76,070	5 IPF 2 SSC 1 PM 1 CHP 8 Control	Profibrotic alveolar macrophages; epithelial cell heterogeneity in fibrotic lungs
Morse et al., <i>ERJ</i> 2019 (4)	Lafyatis	GSE128033	47,771	3 IPF Upper and Lower lobe from each 7 Control	Profibrotic alveolar macrophages (SPP1+/MERTK+); epithelial cell heterogeneity in fibrotic lungs
Adams et al., <i>Science Advances</i> 2020 (1)	Kaminski/Rosas	GSE136831	312,928	32 IPF 28 Control 18 COPD	Aberrant basaloid cells; invasive fibroblasts and myofibroblasts; COL15+ endothelial cells in IPF; profibrotic alveolar macrophages; altered regulatory networks
Habermann et al., <i>Science Advances</i> 2020 (3)	Banovich/Kropski	GSE135893	114,396	12 IPF 3 CHP 2 NSIP 1 Sarcoidosis 10 Control	Pathological KRT5-/KRT17+ epithelial cells; multiple fibroblast subtypes

Total: 551,000 cells from 134 individuals [52 idiopathic pulmonary fibrosis (IPF), 4 chronic hypersensitivity pneumonitis (CHP), 2 nonspecific interstitial pneumonitis (NSIP), 2 systemic sclerosis (SSC), 1 polymyositis (PM), 1 sarcoidosis, 18 chronic obstructive pulmonary disease (COPD), 54 control]. *AJRCCM*, *American Journal of Respiratory and Critical Care Medicine*; *ERJ*, *European Respiratory Journal*.

scRNAseq study, which analyzed the transcriptomes of 24,220 cells from the upper and lower lobes of seven control and three IPF samples, identified three distinct pulmonary macrophage subsets in the IPF lung: FABP4^{hi} macrophages, SPP1^{hi}/MERTK macrophages, and FCN1^{hi} macrophages. The SPP1^{hi}/MERTK macrophages displayed characteristics akin to the profibrotic alveolar macrophages identified in Reyfman et al. (6) and were identified by Morse et al. (4) to accumulate predominately in the lower lobes of the IPF lung. This year, the scale of IPF scRNAseq data was substantially augmented when two papers published in the same issue of *Science Advances* (1, 3) reported the sequencing results of over 400,000 cells from 105 human lungs. In the Habermann et al. (3) study, scRNAseq was applied to single-cell suspensions of 10 nonfibrotic and 20 pulmonary fibrosis lungs to produce a data set of 114,396 sequenced cells clustered into 31 distinct cell subsets. Our group (1) profiled 312,928 cells from distal lung parenchyma samples obtained from 32 IPF lungs, 18 chronic obstructive pulmonary disease (COPD) lungs, and 28 control donor lungs. Both groups observed epithelial cell heterogeneity, profibrotic macrophages, and fibroblast and myofibroblast subtypes in the fibrotic lung. Most interesting is the parallel observation by both groups of a novel pathological epithelial cell population coexpressing basal, epithelial, mesenchymal, senescence, and developmental markers (1, 3). Whereas Habermann et al. identified these cells as KRT5-/KRT17+ pathological, ECM-producing epithelial cells, we have named this population “aberrant basaloid cells” and demonstrated that these cells were localized at the edge of the myofibroblastic foci in the IPF lung. In addition, we identified a novel COL15-expressing population of vascular endothelial cells that in IPF are prominent in the distal lung, in areas of fibrosis and remodeling, but in the nonfibrotic lung are restricted to peribronchial regions and are never seen in the lung parenchyma (1). What we also noticed was a substantial agreement between most data sets: the aberrant basaloid cells can be found in at least three of the data sets (1, 3, 6), as are the profibrotic macrophages. Impressively, despite differences in

dissociation techniques and analytic approaches, the results in these four data sets are congruent. The novel macrophage, epithelial, and endothelial cell populations are seen, as well as the overall phenomenon of proximalization of the distal lung with replacement of alveolar epithelial cells and alveolar endothelial capillary populations, with epithelial and endothelial cells that are usually restricted to the airways. Overall, it can be observed from these single-cell works on IPF that scRNAseq is robust and reproducible and holds significant promise for understanding of IPF and, by extension, developing novel therapeutic and diagnostic approaches to the disease. Looking at these data sets, it became very obvious to us that the discoveries reported by the original groups represent only the tip of the iceberg—and that creating an easy-to-use data mining tool is required to realize the full discovery potential of these data sets.

THE IPF CELL ATLAS WEBSITE

Owing to the “big data” nature of scRNAseq, analysis of these data sets is inherently computational and requires a heightened level of skill and comfort with programming, data science, and statistics, in addition to understanding of the biological fields relevant to the data sets being analyzed. For many biological and medical researchers, the need to develop these skills can be a significant barrier to access single-cell data. While some advocate expanding education in quantitative disciplines to address this, we feel that tools that provide the research community easy and intuitive access to the swaths of biological data available are urgently required. Within the context of single-cell work on IPF, we sought to minimize the barriers by creating an easy-to-navigate web interface that provides researchers public access to the four major single-cell IPF data sets mentioned above with several standardized visualization tools for independent differential gene expression analysis. Any scientist, regardless of their computational skill or physical location, can easily access this website for independent data mining of the

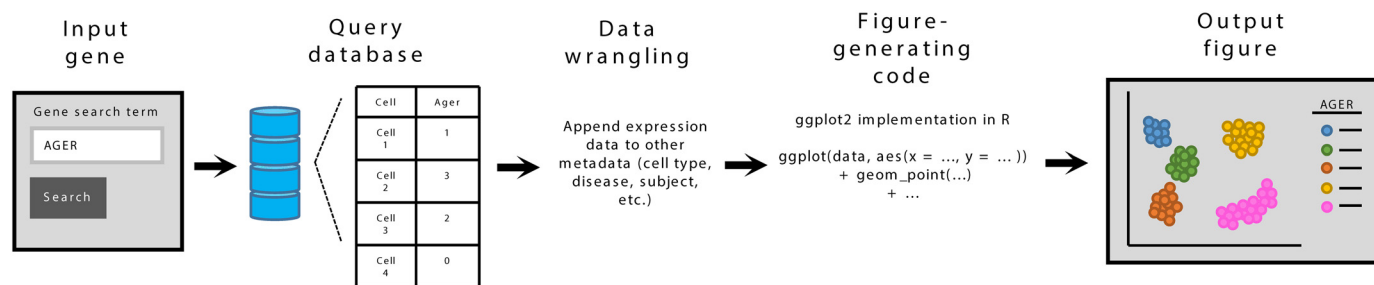


Fig. 1. Flowchart of the IPF Cell Atlas gene search and plot generation pipeline. Briefly, a user will input a gene into the gene search box provided in all analysis tools and will click the search button after toggling the tool-specific plotting options. The website then runs a query of our gene expression-containing database to selectively pull out the data for the input gene. The expression data are appended to the cell metadata for a given data set before being run through R code from the `ggplot2` package to output the figure.

data sets and novel expression pattern exploration. Our website, known as the Idiopathic Pulmonary Fibrosis Cell Atlas (or IPF Cell Atlas for short), can be accessed at www.ipfcellatlas.com.

To code the front-end web interface and back-end database-interacting functions, we utilized the Shiny package (<https://shiny.rstudio.com/>) in the R programming language (9). As most of our single-cell data analysis is performed in the R language, having the ability to code the website in a familiar language both greatly reduced website development time and improved integration of figure-generating code that used the `ggplot2` R package (11). For fast access to gene expression levels, cell-specific gene expression data were stored on a MySQL

database hosted on an Amazon Web Services (AWS) cloud-based server. Database-interacting functions were coded into the back end of the IPF Cell Atlas to facilitate querying of our MySQL database whenever a user inputs a gene into a search box. A diagram of the IPF Cell Atlas' "front-back-front" input processing model is provided in Fig. 1.

One of our main goals when designing the IPF Cell Atlas was to create a web interface that was clean, clear, and easily interpretable for all users regardless of their technical background. For navigation, the portal uses a series of embedded drop-down menus that users can click to move through the various available data sets and visualization tools. The upper panel contains drop-

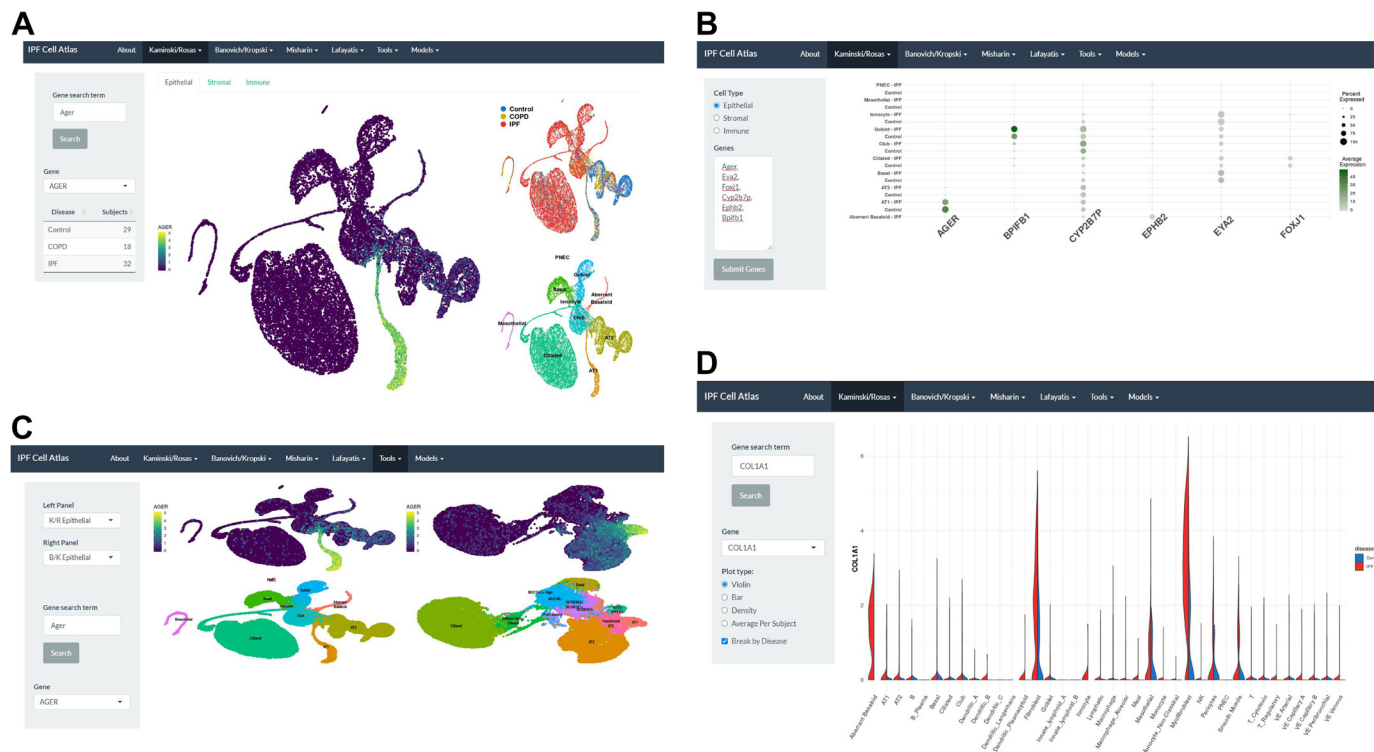


Fig. 2. Screenshots of analysis tools that are available on the IPF Cell Atlas website. *A*: UMAP Explorer feature that allows the user to visualize gene expression of an input gene within a selected cell grouping (i.e., epithelial, immune, stromal, mesenchymal, immune; data set dependent) on a 2-dimensional UMAP projection clustered by cell type. *B*: Batch Explorer features that plot a dot plot of differential gene expression for up to 8 genes within a selected cell grouping. In addition to expression differences between control and IPF samples, the dot plot provides information on the average level of expression of each gene per cell type as well as the percentage of cells of that cell type expressing each gene. *C*: data set comparison tool for comparing gene expression on the different UMAPs generated for each of the data sets. *D*: Gene Explorer feature that allows a user to input a gene and observe cell-specific differential expression patterns in a violin plot, bar graph, density plot, or box plot format. Users can toggle the type of plot and whether to display results that compare control to IPF.

down tabs for each of the four scRNAseq data sets currently available on the website as well as a “Tools” tab for data set comparison and a “Models” tab for a transcriptional regulatory model of IPF. Several visualization tools are available that are meant to facilitate a multifaceted analysis of each of the data sets. These tools include the UMAP (Uniform Manifold Approximation and Projection) Explorer, Gene Explorer, Cell Explorer, and Batch Explorer (Fig. 2). The UMAP Explorer plots the expression of an input gene within each specific cell identified in our data set on a two-dimensional UMAP plot where individual cells are grouped by cell type annotation. This feature is particularly useful for qualitatively assessing cell-specific gene expression patterns in comparison to other cell types identified in each of the data sets. Within this tool, the user can toggle the cell groupings (epithelial, stromal, mesenchymal, immune) that they are most interested in. The Gene Explorer feature provides a more quantitative assessment of differential gene expression patterns and allows the user to plot gene expression between control and IPF samples as a violin plot, bar graph, density plot, or box plot by subject. The Batch Explorer, which is currently limited to the Kaminski/Rosas data set, allows the user to plot differential gene expression between control and IPF samples for up to eight genes of interest in a dot plot format. These types of plots can be particularly informative given that they also provide estimates for the percentage of a given cell type expression of a given gene as well as the average expression of that gene within the given cell type and disease state. Finally, the Cell Explorer function, which is also currently limited to the Kaminski/Rosas data set, allows the user to select a specific cell type for which to plot the average gene expression of all genes within that cell type for the control samples (x-axis) versus the IPF samples (y-axis). Users can input a gene to specifically highlight among the mass of points. In addition to these visualization tools, users who feel their own independent

computational analysis of each data set would be more fruitful can also freely download the data from the Gene Expression Omnibus (GEO), in which links to each GEO page are provided in the drop-down menu of each data set.

A STEP-BY-STEP USAGE GUIDE

Accessing and taking full advantage of the website is straightforward. Upon entering the website at www.ipfcellatlas.com, a user must first decide a few things: 1) the gene(s) whose expression patterns are of interest, 2) the visualization tool that will best display the data, and 3) the specific data set for which to explore and analyze. After these questions are resolved, a user can then begin exploring the data provided on the website easily. As a step-by-step example, suppose a user wants to see a bar graph showing expressional differences of Col4a1 in all identified cell types between control and IPF samples of the Kaminski/Rosas data set. Here, 1) the gene of interest is Col4a1, 2) the visual aid is a bar graph that is found in the Gene Explorer tool, and 3) the data set of interest is “Kaminski/Rosas.” First, the hypothetical user will click the “Kaminski/Rosas” tab on the upper panel to display the drop-down menu of visualization tools. Next, the user will click the Gene Explorer tool from the drop-down menu and will be routed to a new page containing input and toggle boxes on the left of the screen and a display panel in the center. In the case of the Gene Explorer, the user will be greeted by a gene input box, a toggle switch for the type of plot (violin plot, bar graph, density plot, or box plot), and a toggle switch for splitting these plots by disease to compare control and IPF samples. Because the user wants to see a bar graph comparing Col4a1 expression between control and IPF samples, they will now need to input “Col4a1” into the “Gene search term” box, then click the “Bar” option under “Plot type” to output a bar graph and then the “Break by Disease”

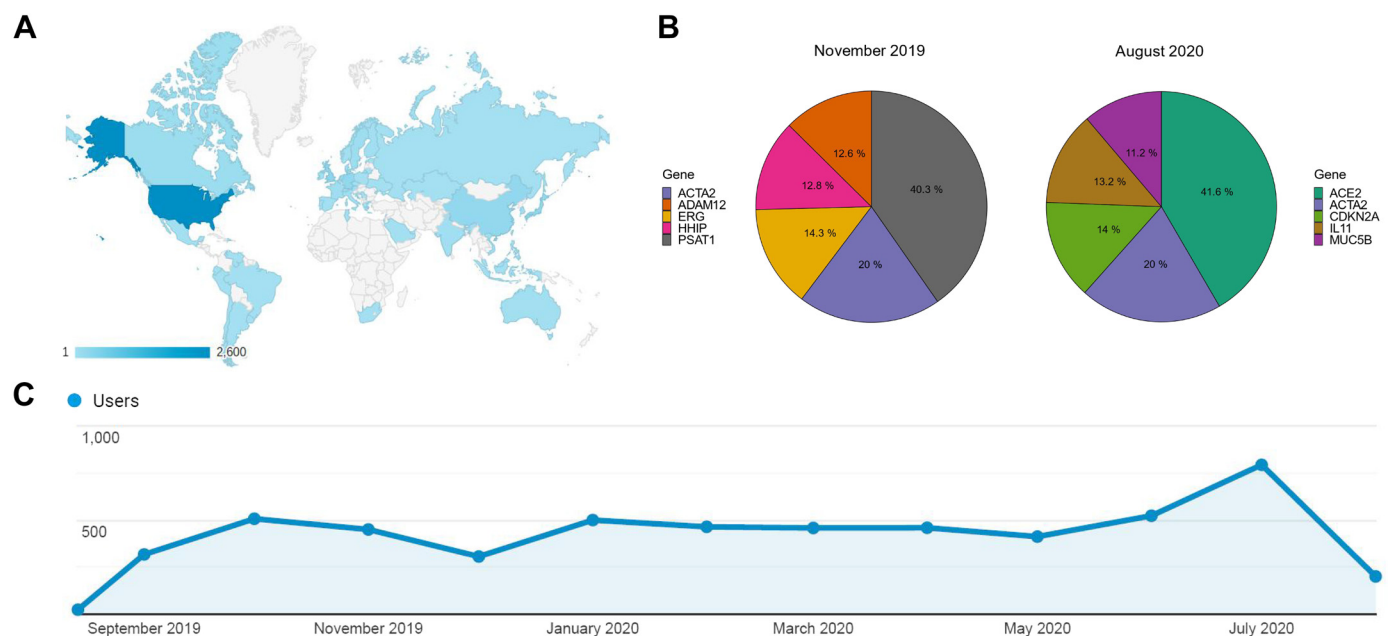


Fig. 3. Website usage statistics provided through Google Analytics. *A*: heatmap of unique users to the portal from across the globe. *B*: pie charts of the top 5 searched genes before the pandemic (November 2019) and during the pandemic (August 2020). Percentages shown represent the percentage of searches of that gene within the 5 genes shown. *C*: line graph displaying the number of unique users accessing the website each month since it was made public.

checkbox at the bottom of the panel to split the bars for each cell type by control and IPF. Finally, the user will click “Search” for the website to run the query on Col4a1 and output a split bar graph of identified cells of the Kaminski/Rosas data set that is split by control and IPF. The same general step-by-step process is applicable, with some minor adjustments, to the other visualization tools available on the portal.

WEBSITE USAGE BY THE NUMBERS

Since being made public in September of 2019, the IPF Cell Atlas portal has seen more than 4,000 unique users from five continents and 56 countries, with ~60% of users coming from the United States per Google Analytics (Fig. 3). There are on average 500 unique visitors to the website each month. July 2020 saw our portal reach a new record for monthly users with 794 unique visitors. Over 12,000 sessions on the portal have been engaged that have led to over 100,000 searches of more than 7,400 unique genes. At the time of writing this editorial, the most frequently searched gene is ACE2, the COVID-19 cellular receptor. This is, of course, due to the COVID-19 pandemic. ACE2 had zero searches in November 2019 but increased to be the top searched gene by August 2020 with 1,966 searches (Fig. 3). Other top searched genes include the myofibroblast marker ACTA2 (942 searches), the senescence marker CDKN2A (662 searches), the cytokine IL11 (623 searches), and MUC5B, the gene whose promoter variant is the most associated gene variant with IPF (530 searches). Although the top genes searched may represent current and emerging interest, we find the fact that 7,400 unique genes have been accessed by over 4,000 individual users most exciting, as it represents the diversity, extent, and reach of interest the IPF Cell Atlas generated.

FUTURE WORK AND CONCLUDING REMARKS

The IPF Cell Atlas in its current form represents the first iteration of what could be a significantly more ambitious project. Based on current use it seems that the IPF Cell Atlas is indeed expanding the access of the academic community to single-cell data from human lung diseases, and, as the ACE2 results suggest, investigators utilize this atlas for scientific objectives that could not have even been predicted when the atlas was designed. We plan to continue to develop the website to include integrating new single-cell data sets and analytical tools as they are made available. New analysis tools may include a heatmap generator, differential cell-cell interaction networks, and improved support for cross-data set gene expression comparison, therapeutic signature exploration, and integrating data from other IPF-specific OMICs approaches, especially spatial transcriptomic as well as single-cell epigenomic, proteomic, and metabolomic data. However, for the lung community at large to benefit, the IPF Cell Atlas should also be expanded in scope; IPF is after all only one of many diffuse parenchymal lung diseases. Thus, data sets from other diffuse parenchymal lung diseases, as well as reference data sets, should be included (7). This would pose significant challenges, as presently there is little agreement about essential information required to harmonize data such as minimal sample attributes (metadata) that need to be provided for each sample, conventions for cell type and state definitions, clear delineation of anatomical section collection, and an effort to develop representative sampling of genders,

ethnicities, races, exposures, disease stages, presentations, and other clinically relevant attributes (8). As more data sets are added and the diversity of the populations and diseases widens, the value of user-friendly, simple to use, single-cell data portals will increase, to allow exploration of the multicellular complexities of human advanced lung disease, facilitate discovery and hypothesis generation, and eventually be a significant milestone in the campaign to cure IPF, as well as other diffuse parenchymal lung diseases.

GRANTS

This work was supported by a generous donation from Three Lake Partners for the construction of the IPF Cell Atlas and by NIH National Heart, Lung, and Blood Institute Grants R01 HL-127349, R01 HL-141852, and U01 HL-145567 to N.K.

DISCLOSURES

N.K. served as a consultant to Biogen Idec, Boehringer Ingelheim, Third Rock, Pliant, Samumed, NuMedii, Theravance, LifeMax, Three Lake Partners, Optikira, and Astra Zeneca over the last 3 years, reports equity in Pliant, a grant from Veracyte, and nonfinancial support from MiRagen and Astra Zeneca, and has intellectual property on novel biomarkers and therapeutics in IPF licensed to Biotech. None of the other authors has any conflicts of interest, financial or otherwise, to disclose.

AUTHOR CONTRIBUTIONS

N.N., C.C. and N.K. prepared figures; N.N., C.C. and N.K. drafted manuscript; N.N., C.C., K.R. and N.K. edited and revised manuscript; C.C. and N.K. approved final version of manuscript.

REFERENCES

- Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, Chu SG, Raby BA, Deluiliis G, Januszyk M, Duan Q, Arnett HA, Siddiqui A, Washko GR, Homer R, Yan X, Rosas IO, Kaminski N. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv* 6: eaba1983, 2020. doi:10.1126/sciadv.aba1983.
- Carraro G, Mulay A, Yao C, Mizuno T, Konda B, Petrov M, Lafkas D, Arron JR, Hogaboam CM, Chen P, Jiang D, Noble PW, Randell SH, McQualter JL, Stripp BR. Single cell reconstruction of human basal cell diversity in normal and IPF lung. *Am J Respir Crit Care Med* In press. doi:10.1164/rccm.201904-0792OC.
- Habermann AC, Gutierrez AJ, Bui LT, Yahn SL, Winters NI, Calvi CL, Peter L, Chung MI, Taylor CJ, Jetter C, Raju L, Roberson J, Ding G, Wood L, Sucre JM, Richmond BW, Serezani AP, McDonnell WJ, Mallal SB, Bacchetta MJ, Loyd JE, Shaver CM, Ware LB, Bremner R, Walia R, Blackwell TS, Banovich NE, Kropski JA. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci Adv* 6: eaba1972, 2020. doi:10.1126/sciadv.aba1972.
- Morse C, Tabib T, Sembrat J, Buschur K, Bittar H, Valenzi E, Jiang Y, Kass D, Gibson K, Chen W, Mora A, Benos P, Rojas M, Lafyatis R. Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur Respir J* 54:1802441, 2019. doi:10.1183/13993003.02441-2018.
- Raredon MS, Adams TS, Suhail Y, Schupp JC, Poli S, Neumark N, Leiby KL, Greaney AM, Yuan Y, Horien C, Linderman G, Engler AJ, Boffa DJ, Kluger Y, Rosas IO, Levchenko A, Kaminski N, Niklason LE. Single-cell connectomic analysis of adult mammalian lungs. *Sci Adv* 5: eaaw3851, 2019. doi:10.1126/sciadv.aaw3851.
- Reyffman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, et al. Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. *Am J Respir Crit Care Med* 199: 1517–1536, 2019. doi:10.1164/rccm.201712-2410OC.
- Schiller HB, Montoro DT, Simon LM, Rawlins EL, Meyer KB, Strunz M, Vieira Braga FA, Timens W, Koppelman GH, Budinger GR, Burgess JK, Waghray A, van den Berge M, Theis FJ, Regev A,

- Kaminski N, Rajagopal J, Teichmann SA, Misharin AV, Nawijn MC. The Human Lung Cell Atlas: a high-resolution reference map of the human lung in health and disease. *Am J Respir Cell Mol Biol* 61: 31–41, 2019. doi:10.1165/rcmb.2018-0416TR.
8. Schupp JC, Yan X, Kaminski N. Towards a cell atlas of the human airway. *Am J Respir Crit Care Med* In press. doi:10.1164/rccm.202007-2977ED.
9. Travaglini KJ, Nabhan AN, Penland L, Sinha R, Gillich A, Sit RV, Chang S, Conley SD, Mori Y, Seita J, Berry GJ, Shrager JB, Metzger RJ, Kuo CS, Neff N, Weissman IL, Quake SR, Krasnow MA. A molecular cell atlas of the human lung from single cell RNA sequencing (Preprint). *bioRxiv*, 2019. doi:10.1101/742320.
10. Xu Y, Mizuno T, Sridharan A, Du Y, Guo M, Tang J, Wikenheiser-Brokamp KA, Perl AT, Funari VA, Gokey JJ, Stripp BR, Whitsett JA. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. *JCI Insight* 1: e90558, 2016. doi:10.1172/jci.insight.90558.
11. Yu G, Ibarra GH, Kaminski N. Fibrosis: lessons from OMICS analyses of the human lung. *Matrix Biol* 68-69: 422–434, 2018. doi:10.1016/j.matbio.2018.03.014.

