AWARD NUMBER: W81XWH-19-1-0416

TITLE: Integrated molecular pathogenesis of pulmonary fibrosis

PRINCIPAL INVESTIGATOR: Dr. Nicholas Banovich, PhD

CONTRACTING ORGANIZATION: Translational Genomics Research Institute and Vanderbilt University Medical Center

REPORT DATE: August 2020

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
Aug 2020	Annual	08/01/2019 - 07/31/2020

4. TITLE AND SUBTITLE

Integrated Molecular Pathogenesis of Pulmonary Fibrosis

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-19-1-0416

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dr. Nicholas Banovich, PhD
Dr. Jonathan Kropski, MD

E-Mail: nbanovich@tgen.org, jon.kropski@vumc.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Translational Genomics Research Institute, 445 N 5th St, Phoenix AZ, 85004

Vanderbilt University Medical Center,
1211 Medical Center Dr., Nashville, TN
37232

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Pulmonary fibrosis (PF) is a heterogeneous clinical syndrome that represents the end-stage of chronic interstitial lung diseases. Dozens of different occupational, environmental, immune and genetic risk factors have been associated with PF, and through the past several decades, risk factor exposures have been the driving force in the diagnostic classification of PF, thus in the current paradigm, there are dozens of different â€œdiagnosesâ€ of pulmonary fibrosis.

15. SUBJECT TERMS
Single cell RNA-seq, pulmonary fibrosis

16. SECURITY CLASSIFICATION OF: U			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified	Unclassified		USAMRMC

TABLE OF CONTENTS

1. Introduction

Pulmonary fibrosis (PF) is a heterogeneous clinical syndrome that represents the end-stage of chronic interstitial lung diseases. Dozens of different occupational, environmental, immune and genetic risk factors have been associated with PF, and through the past several decades, risk factor exposures have been the driving force in the diagnostic classification of PF, thus in the current paradigm, there are dozens of different “diagnoses” of pulmonary fibrosis. This emphasis on distinction has focused much attention on the most “common” form of this syndrome (Idiopathic Pulmonary Fibrosis, IPF), which comprises only 20% of PF patients. Today there are 2 modestly effective FDA-approved treatments for IPF; however, for the 80% of PF patients with other diagnoses, there are no known effective treatments. The current paradigm emphasizing diagnostic distinction has limited progress in understanding how different risk factors lead to a common end-stage lung pathology. In order to rapidly accelerate progress towards better treatments for all PF patients, a radical departure from this approach is needed. We believe any subdividing of PF should be driven by demonstrated relevant differences in disease biology; to this end, it has become clear that a more nuanced understanding of “upstream” disease mechanism of disease initiation and propagation, as well as the convergent “downstream” mechanisms of lung fibrosis is critical. By leveraging the inherent heterogeneity of disease state in the PF lung, we will employ innovative single-cell genomic approaches – in particular single cell RNA-seq (scRNA-seq) and culture models to recreate the molecular natural history of disease, determine the convergent mediators and pathways that drive PF pathogenesis and identify mechanistically-driven disease endotypes.

2. Keywords

Pulmonary fibrosis
Interstitial lung disease
Genomics
Single cell RNA sequencing (scRNA-seq)

3. Accomplishments

What were the major goals of the project?

This project consists of three specific aims:
Aim 1. Profile the clinical, cellular and molecular landscape of PF lungs.

Aim 2: Determine the conserved cell-type specific gene expression programs driving PF pathogenesis

Aim 3. Determine the mechanisms underlying molecular endotypes of pulmonary fibrosis.

What was accomplished under these goals?

Aim 1: The major objective of the first year was to enroll, collect and generate scRNA-seq data from 1/3 of the proposed 75 patients. After we received final IRB approvals in late October 2020, the laboratory and programmatic shut-downs due to the COVID-19 pandemic starting in March 2020 (4 months after our IRB approvals) were a substantial impediment to progress during this year. However, in spite of the disruptions due to COVID-19, during the current reporting period we collected lung tissue and performed scRNA-seq from 20 patients with PF: unclassifiable ILD (7), CTD-ILD (1), IPFA (3), NSIP (4), cHP (2), and sarcoidosis (3). From the majority of these samples we successfully collected and sequenced samples from both a highly fibrotic (20) and a less affected (14) region of the lung; in the remaining cases, disease was homogenous and samples from multiple regions were pooled for sequencing. In addition, while not directly supported by DOD funds but with planned utilization of the data for integrated analyses as a part of these studies, we have now also generated scRNA-seq libraries and/or performed sequencing from 29 IPF lungs and 34 control lungs. Prior to

COVID-19 related shutdowns, our sample acquisition was proceeding on or ahead of schedule, and barring unforeseen developments, we anticipate reaching our target of 75 PF lungs during year 3 of the award.

Tasks by PI:
Sample acquisition – Banovich and Kropski
Single cell library generation – Banovich and Kropski
Sequencing – Banovich

Aim 2: We have begun analyzing the scRNA-seq data generated to identify cell-type-specific gene expression changes associated with PF. To this end, we have integrated this dataset with a second dataset, generated by us under the support of the NIH/NHLBI, which consists of lung tissue from 34 controls (declined donors) and 29 IPF patients. Together, this is the most comprehensive dataset of scRNA-seq data from PF. Using a negative binomial regression framework, we have identified hundred of differentially expressed (DE) genes in the majority of cell-types. We have made comparisons between control lung and all disease, as well as between control lung and individual diagnosis – identifying both conserved and specific DE programs. Additionally, as chronic lung disease has proven to be a significant co-morbidity with regards to the recent COVID19 pandemic, we have undertaken a number of analyses to identify how gene expression programs associated with entry, infection and response to SARS-CoV-2 are altered in patients with PF. Our findings suggest that there are significant gene expression changes associated with PF that create a molecular environment more favorable to SARS-CoV-2 entry and infection (**Figure 1**). This manuscript (led by Dr.Linh Bui, a postdoctoral fellow from Dr. Banovich's lab and Dr. Nichelle Winters, a pulmonary/critical care fellow from Dr. Kropski's laboratory), is being finalized for submission within the next month.

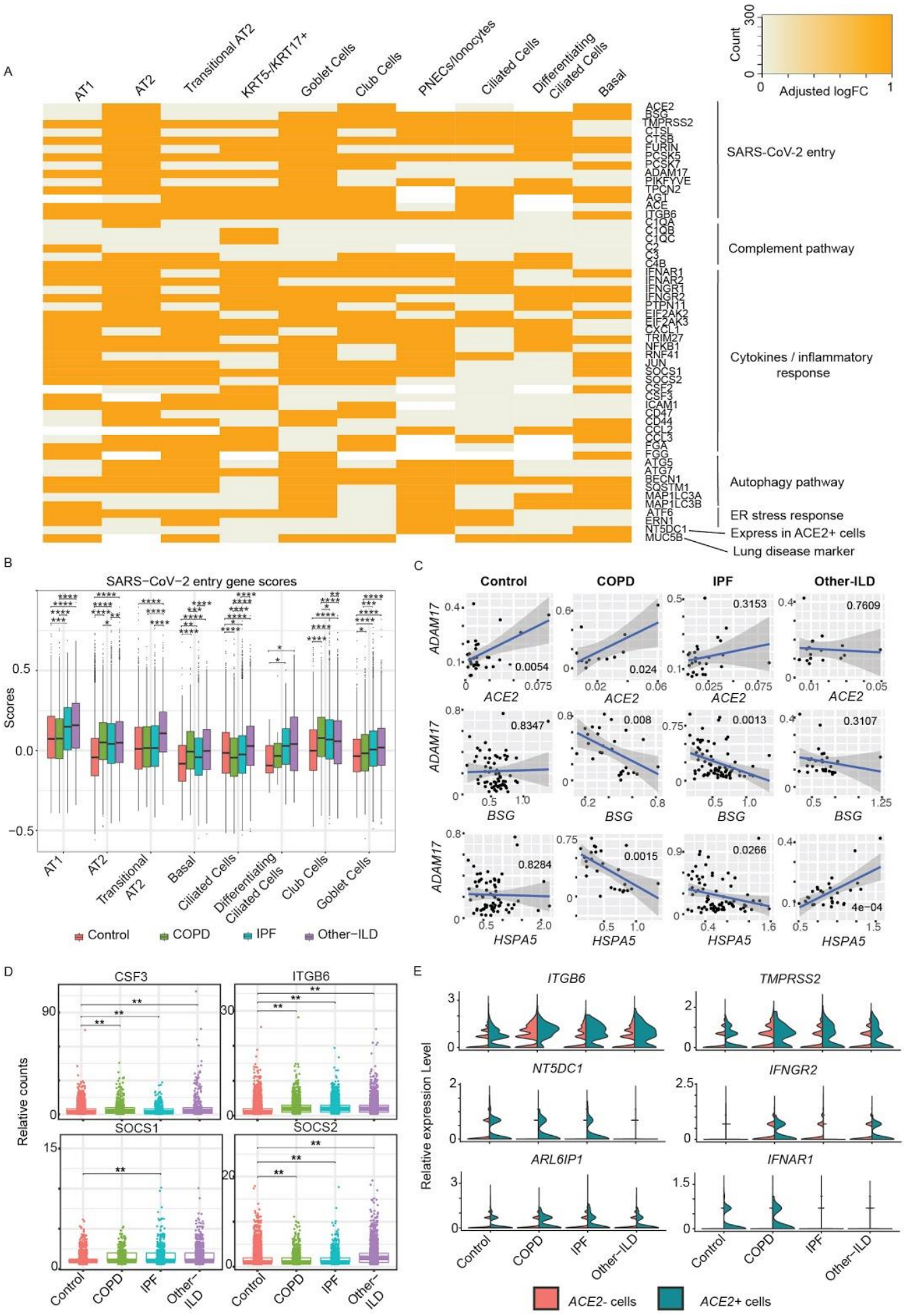
A
AT1
AT2
Transitional AT2
KRT5-/KRT17+
Goblet Cells
Club Cells
PNECs/Ionocytes
Ciliated Cells
Differentiating Ciliated Cells
Basal
Count
300
0
0
Adjusted logFC
1
ACE2
BSG
TMPRSS2
CTSL
CTSB
FURIN
PCSK5
PCSK7
ADAM17
PIKFYVE
TPCN2
AGT
ACE
ITGB6
C1QA
C1QB
C1QC
C2
C3
C4B
IFNAR1
IFNAR2
IFNGR1
IFNGR2
PTPN11
EIF2AK2
EIF2AK3
CXCL1
TRIM27
NFKB1
RNF41
JUN
SOCS1
SOCS2
CSF2
CSF3
ICAM1
CD47
CD44
CCL2
CCL3
FGA
FGG
ATG5
ATG7
BECN1
SQSTM1
MAP1LC3A
MAP1LC3B
ATF6
ERN1
NT5DC1
MUC5B
SARS-CoV-2 entry
Complement pathway
Cytokines / inflammatory response
Autophagy pathway
ER stress response
Express in ACE2+ cells
Lung disease marker
B
SARS−CoV−2 entry gene scores
Scores
0.5
0.0
−0.5
AT1
AT2
Transitional AT2
Basal
Ciliated Cells
Differentiating Ciliated Cells
Club Cells
Goblet Cells
Control
COPD
IPF
Other−ILD
C
Control
COPD
IPF
Other-ILD
ADAM17
ACE2
BSG
HSPA5
0.0054
0.024
0.3153
0.7609
0.8347
0.008
0.0013
0.3107
0.8284
0.0015
0.0266
4e−04
D
CSF3
ITGB6
SOCS1
SOCS2
Relative counts
Control
COPD
IPF
Other−ILD
E
ITGB6
TMPRSS2
NT5DC1
IFNGR2
ARL6IP1
IFNAR1
Relative expression Level
Control
COPD
IPF
Other−ILD
ACE2- cells
ACE2+ cells

Figure 1. Expression of SARS-CoV2 entry-related factors in chronic lung disease. 2A. Heatmap shows the majority of the Covid-19 response genes is upregulated in the Disease samples compared to Control, binom test p_value = 9.909e-07. 2B. SARS-CoV-2 entry module score in different CT, Tukey_HSD post-hoc statistical test, p-value < 0.05: *, p-value < 0.01: **, p-value < 0.001: ***, p-value < 0.0001: ****. 2C. Gene expression correlation between ADAM17 vs. ACE2, BSG (CD147) and HSPA5 (GRP78): significant correlation in COPD and IPF 2D. Boxplot shows differences in gene expression of some Covid-19 reponse genes in AT2 cell types, **: p_value_adj < 0.05 (negative binomial test, 4 variables: Age, Ethnicity, Smoking_status and Dataset) 2E. High expression of some interesting genes in ACE2+ (blue) vs. ACE2- (red) AT2 cells

In addition, among the 20 PF samples sequenced during the past year were 3 from individuals with PF and a history of coal mining with suspected coal-worker's pneumoconiosis. Working in collaboration with Dr. Frank McCormack at the University of Cincinnati, we identified a subset of macrophages that exhibit osteoclast-like features in these lungs, and collaborative work in Dr. McCormack's lab has demonstrated these osteoclast like macrophages are both profibrotic and their polarization is driven at least in part through the RANK/RANKL pathway. This manuscript is being finalized and we anticipate will be submitted within the next month.

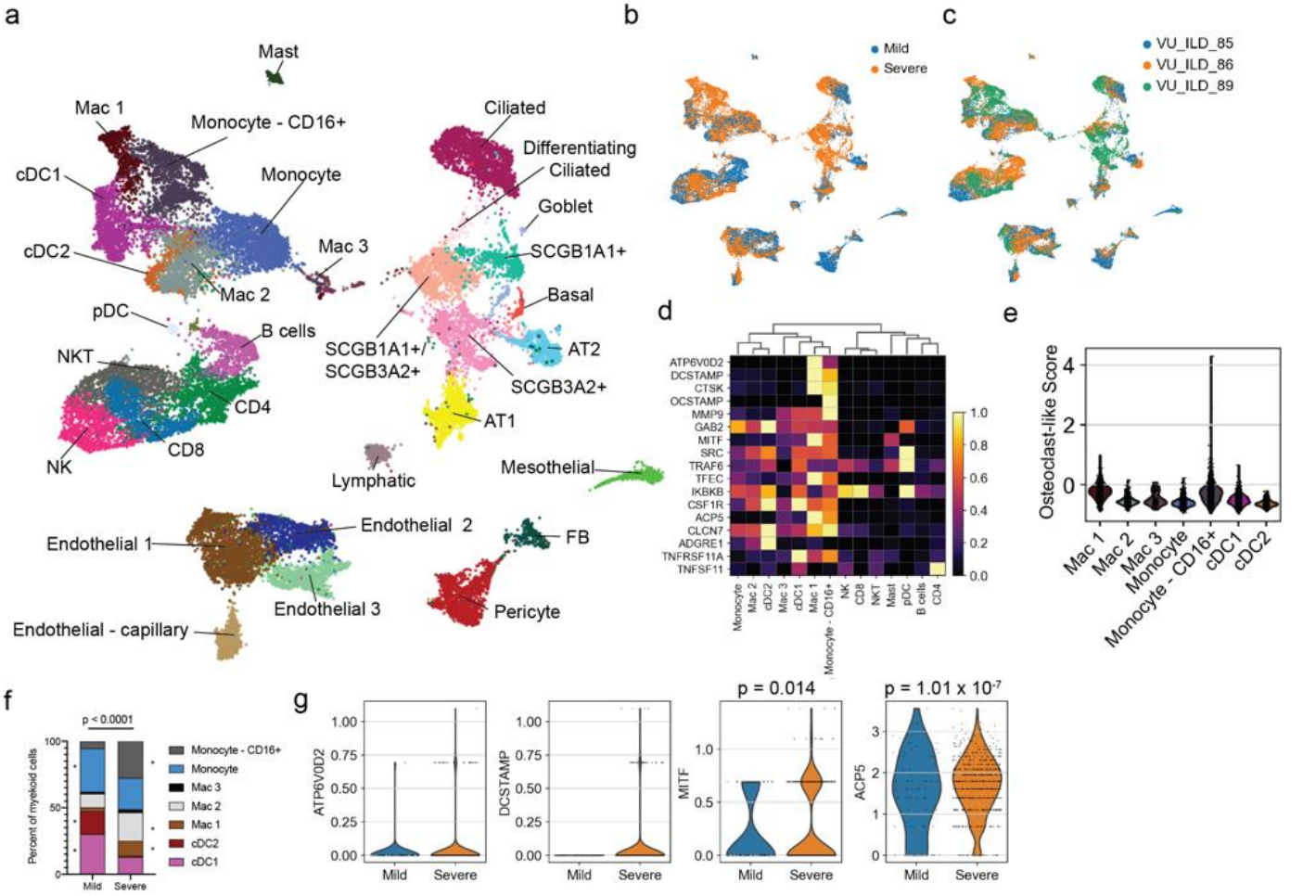

Figure 2. scRNA-seq of coal-worker's pneumoconiosis. Single-cell suspensions were generated from severely diseased (Severe) and relatively preserved (Mild) regions of 3 silicosis lungs removed at the time of lung transplantation, and were processed in parallel for scRNA-seq using the 10X Genomics 5' assay. Uniform manifold approximation and projection (UMAP) embedding of 29,334 cells from explanted lung tissue from 3 patients with silicosis who underwent lung transplantation grouped by a) cell-type following recursive clustering, b) Mild vs. Severe status, and c) lung donor. d) Heatmap depicting relative expression of selected osteoclast-related genes in immune cells. e) An "osteoclast-like" gene score (score_genes function in Scanpy v1.46) was generated for myeloid cells using genes from (d) as input. f) Comparison of myeloid cell subtypes between Mild and Severely affected regions. Distribution was compared by Chi-square. g) Relative expression of selected osteoclast-like genes within myeloid cells comparing Mild and Severely affected regions. P-values reflect FDR-adjusted negative binomial test results.

Tasks by PI:
Data integration: Banovich and Kropski
Cell-type annotation: Kropski
Differential expression analysis: Banovich
COVID19 analyses: Banovich and Kropski
CWP analyses: Kropski/Banovich

Aim 3: At present, statistical power for endotype analyses is still limited, and we plan to begin these analyses once larger samples numbers are available in the coming year. We have continued optimization of organoid culture models for ex-vivo validation studies.

What opportunities for training and professional development has the project provided? Dr. Linh Bui received new training around novel computational analyses to perform the work carried out here. She was scheduled to present this work at a number of conferences which were unfortunately cancelled. This experience was invaluable her Dr. Bui's successful competition for an F32 award. Dr. Nichelle Winters (a postdoctoral fellow supported by T32 funding) also was scheduled to present an abstract at the ATS 2020 International Conference which was unfortunately cancelled due to the COVID-19 pandemic.

How were the results disseminated to communities of interest? This work was presented as a poster at the Biology of Genomes meeting in May of 2020 and the American Thoracic Society meeting in August of 2020.

What do you plan to do during the next reporting period to accomplish the goals? We will continue to work towards the goals as outlined in the SOW. We anticipating remaining on schedule with this work.

4. Impact
What was the impact on the development of the principal discipline(s) of the project?
Nothing to Report.
What was the impact on other disciplines?
Nothing to Report.
What was the impact on technology transfer?
Nothing to Report.
What was the impact on society beyond science and technology?
Nothing to Report.

5. Changes/Problems
Nothing to Report.

6. Products
Publications, conference papers, and presentations:
Journal publications.
Sungnak, W et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. Nat. Med. April 23, 2020. doi: https://doi.org/10.1038/s41591-020-0868-6. PMID: 32327758
Books or other non-periodical, one-time publications.
Nothing to Report.
Other publications, conference papers, and presentations.
Poster CSHL Biology of Genomes – May 2020 (virtual)

Poster American Thoracic Society – Aug. 2020 (virtual)

Website(s) or other Internet site(s)

Nothing to Report.

Technologies or techniques

Nothing to Report.

Inventions, patent applications, and/or licenses

Nothing to Report.

Other Products

Nothing to Report.

7. Participants & Other Collaborating Organizations

What individuals have worked on the project?

Name: Nicholas Banovich, PhD
Project Role: P.I.
Nearest person month worked: 4
Contribution to the project: Dr. Banovich, along with co-PI Dr. Kropski, oversees the overall aims of this project, in particular the genomic and computational analyses. He participates in data interpretation, presentation, and publication.

Name: Linh Bui, PhD
Project Role: Postdoctoral Researcher
Nearest person month worked: 12
Contribution to the project: Dr. Bui is responsible for curation, QC, and analysis of genomic data generated in this project.

Name: Lance Peter, MSc
Project Role: Research Associate
Nearest person month worked: 12
Contribution to the project: Mr. Peter is responsible for processing primary lung tissue from patients with ILD, including 10X genomics library preparation and next generation sequencing. Mr. Peter is also responsible for sample management and organization.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nicholas Banovich, PhD
New: 1R01CA254271 (Brown)
Ended: None

What other organizations were involved as partners?

Organization name: Norton Thoracic Institute at Dignity Health
Location of Organization: Phoenix Arizona
Partner's contribution to the project: Collaboration

- The Norton Thoracic Institute collects biopsies from explant lungs to be used in this study. Rajat Walia, MD, serves as the primary collaborator.

8. Special Reporting Requirements

Collaborative award

9. Appendices

Nothing to report