AWARD NUMBER: W81XWH-19-1-0415

TITLE: Integrated Molecular Pathogenesis of Pulmonary Fibrosis

PRINCIPAL INVESTIGATOR: Dr. Jonathan Kropski, MD

CONTRACTING ORGANIZATION: Vanderbilt University Medical Center
Nashville, TN

REPORT DATE: August 2021

TYPE OF REPORT: Annual

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

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE
August 2021

2. REPORT TYPE
Annual

3. DATES COVERED
01Aug2020-31Jul2021

4. TITLE AND SUBTITLE

Integrated Molecular Pathogenesis of Pulmonary Fibrosis

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-19-1-0415

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 "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.

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 USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE	Unclassified	10	19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

	Page
1. Introduction	4
2. Keywords	4
3. Accomplishments	4
4. Impact	8
5. Changes/Problems	8
6. Products	8
7. Participants & Other Collaborating Organizations	9
8. Special Reporting Requirements	10
9. Appendices	10

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 by the end of the second year was to enroll, collect and generate scRNA-seq data from 2/3 of the proposed 75 patients. The ongoing COVID-19 pandemic continued to cause disruptions to our workflow and sample acquisition, related largely to reduced numbers of transplants due to intermittent hospital capacity issues. However, we did continue to make progress, and in spite of the disruptions due to COVID-19, during the current reporting period we collected lung tissue from 22 new lungs with PF – for a total of 42 non-IPF ILD lungs. These include a diversity of diagnoses including unclassifiable ILD, CTD-ILD, chronic hypersensitivity pneumonitis, coal worker pneumoconiosis (CWP), IPAF, NSIP, sarcoidosis, and post-ARDS pulmonary fibrosis. From the majority of these samples were able to collect samples from both a highly fibrotic and a less affected region of the lung; in the remaining cases, disease was homogenous and samples from multiple regions were pooled for sequencing. We have completed scRNA-seq on 8 of the 22 new individuals collected this year, and the remaining 14 are in our sequencing queue. The data quality has remained

excellent, nearly all samples passing our quality control thresholds for cells captured, genes per-cell, and mitochondrial content. 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 40 IPF lungs and 52 control lungs. Prior to COVID-19 related shutdowns, our sample acquisition was proceeding on or ahead of schedule, however given the circumstances of this past year and conditions anticipated for the next several months, we expect some reduction in transplant numbers sample availability this year. Consequently, we expect to approach our target of 75 total samples during this funding year, but recognize it is possible that conditions may require us to request a no-cost extension to complete our sample accrual.

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

Aim 2: Our data analysis is performed iteratively, as batches of library sequencing are completed. 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 52 controls (declined donors) and 40 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. In last year's report we described an ongoing study using these data to identify how gene expression programs associated with entry, infection and response to SARS-CoV-2 are altered in patients with PF. This study (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 now published in Nature Communications.

In addition, among the 42 PF samples collected to date are 6 from individuals with PF and a history of coal mining with suspected coal-worker's pneumoconiosis (Figure 1). We have jointly analyzed 5 of these samples, and library sequencing of the 6th will be completed in the next few weeks. 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 in the next several months.

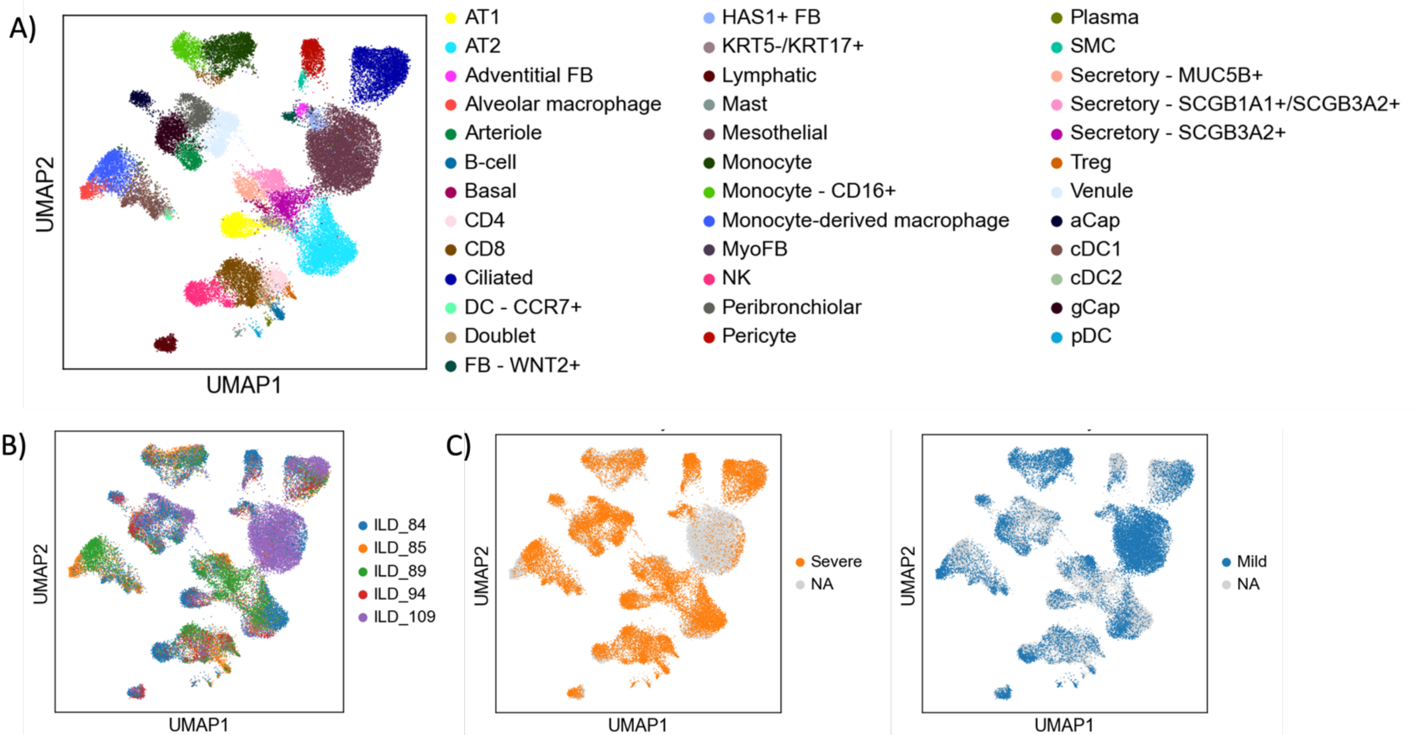

Figure 1. scRNA-seq of Coal-worker's-pneumoconiosis (CWP)-associated pulmonary fibrosis. UMAP embedding depiciting integrated analysis and cell-type annotation of 40,134 cells from 5 unique CWP lungs annotated by A) cell-type, B) donor, and C) severity.

Additionally, three of the samples we sequenced this past year were from patients with post-ARDS pulmonary fibrosis as a sequellae of SARS-CoV-2 infection. We have performed initial analysis on this post-COVID ARDS form of pulmonary fibrosis. Our initial analyses suggest emergence of KRT5-/KRT17+ matrix-producing epithelial cells in this form of pulmonary fibrosis, but with less activated-states compared to those in IPF (Figure 2). This manuscript is currently being finalized and we anticipate submitting it for publication in the fall of 2021.

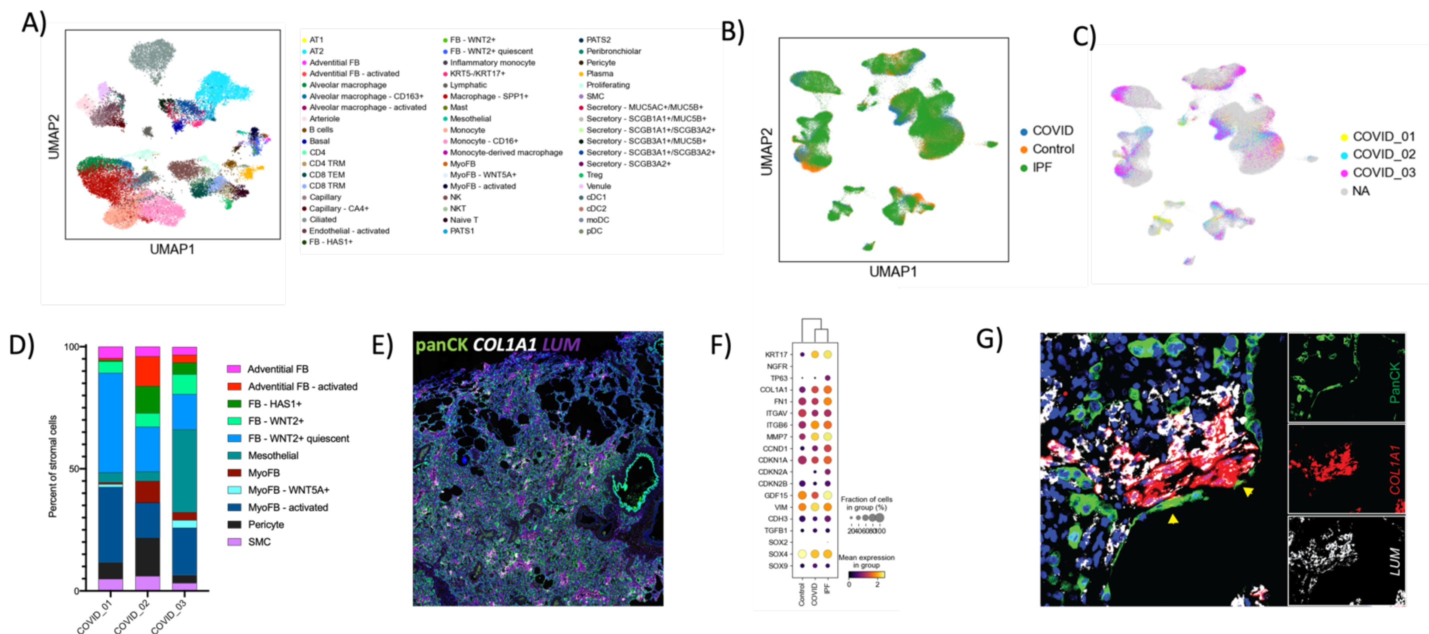

Figure 2. Post-ARDS pulmonary fibrosis following SARS-CoV-2 infection. UMAP embedding depicting 160,012 cells from 3 post-COVID PF, 10 control, and 12 IPF lungs annotated by A) cell-type, B) diagnosis, and C) donor. D) Comparison of fibroblast subtypes across post-COVID PF donors. E) Immunofluorescence image from post-COVID PF depicting fibroblast location. F) Dotplot depicting expression of KRT5-/KRT17+ molecular

programs across diagnoses. G) Immunofluorescence image from post-COVID PF depicting KRT5-/KRT17+ cell localization.

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
Post-ARDS fibrosis 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 begun structured analyses of the imaging and pathology. We have continued optimization of organoid culture models, with an emphasis on population identification and cell-sorting strategy as we have found there are considerable changes in cell-surface markers in diseased lungs. As proof of concept, we isolated AT2 cells from a PF lung using a CD45-/CD31-/CD326+/Lysotracker positive strategy by FACS. Analysis of cytospins revealed >95% concordance of Lysotracker+ status with pro-SP-C expression, revealing a high-degree of purity. Coculture of these AT2 cells with primary human lung fibroblasts in small-airway growth media (SAGM) x 21 days led to the establishment and expansion of AT2 organoids which was regulatable with small-molecule based-interventions (Figure 3).

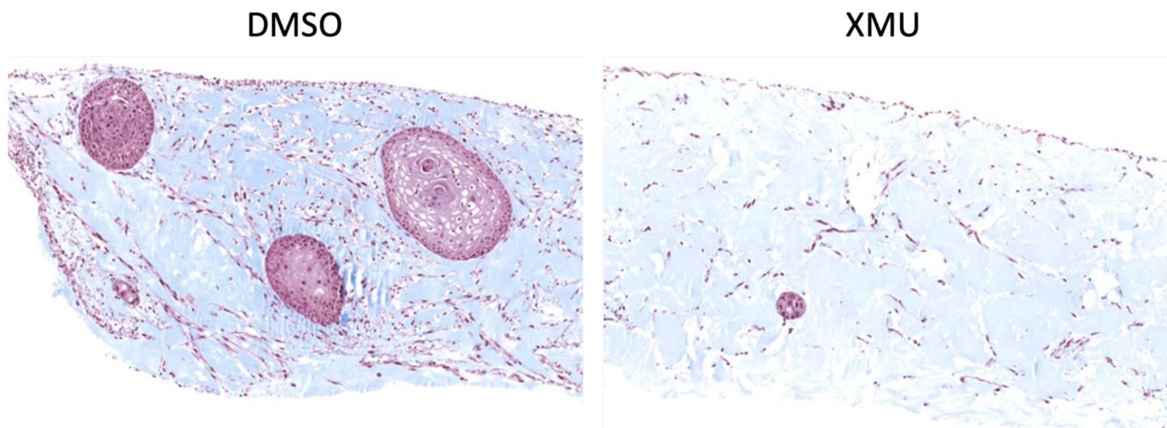

Figure 3. AT2-derived organoids from PF lungs. CD45-/CD31-/CD326+/Lysotracker+ AT2 cells were isolated from a PF lung and cocultured with primary lung fibroblasts (5,000:50,000 ratio) in SAGM x21 days in the presence or absence of Wnt/catenin-pathway inhibitor XMU.

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. Dr. Bui also presented the COVID-19 analyses described above – now published in Nature Communications – at the American Society of Human Genetics conference in October of 2020. She also presented at an American Thoracic Society Webinar in October of 2020.

How were the results disseminated to communities of interest? In addition to the Nature Communications manuscript noted above, this work was presented at the American Society of Human Genetics conference in October of 2020, at an American Thoracic Society Webinar in October of 2020, at the Human Cell Atlas Lung Biological Network webinar in February of 2021, at the American Thoracic Society International Virtual Conference in May 2021, at several university

seminars (Northwestern University, University of Washington, University of Southern California), and at the Vermont Stem Cell Conference in July 2021.

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.

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.

Muus C et al. Integrated analyses of single-cell atlases reveal age, gender, and smoking status associations with cell type-specific expression of mediators of SARS-CoV-2 viral entry and highlights inflammatory programs in putative target cells. Nature Medicine. March 2, 2021. https://doi.org/10.1038/s41591-020-01227-z. PMID: 33654293

Bui LT and Winters NI et al. Single-cell RNA-sequencing reveals dysregulation of molecular programs associated with SARS-CoV-2 severity and outcomes in patients with chronic lung disease. Nat Communications. 2021 July 14. doi: https://doi.org/10.1038/s41467-021-24467-0. PMID: 34262047 PMCID: PMC8280215

Books or other non-periodical, one-time publications.
Nothing to Report.
Other publications, conference papers, and presentations.
Poster American Society of Human Genetics – October 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: Jonathan A. Kropski
Project Role: PI
Nearest person month worked: 3
Research Identifier (e.g. ORCID ID): 0000-0002-8923-1344
Contribution to project: Dr. Kropski oversaw sample collection, processing, and analysis of scRNA-seq data, and performance of ex-vivo experiments.

Name: Carla L. Calvi
Project Role: Senior Research Specialist
Nearest person month worked: 1
Research Identifier (e.g. ORCID ID): N/A
Contribution to Project: Ms. Calvi assisted with biospecimen processing and related cell-culture studies.

Name: Molly Dixon
Project Role: Research Assistant
Nearest person month worked: 6
Research Identifier (e.g. ORCID ID): N/A
Contribution to Project: Ms. Dixon assisted with biospecimen processing and IHC studies.

Name: Jason Gokey
Project Role: Co-investigator
Nearest person month worked: 4
Research Identifier (e.g. ORCID ID): 0000-0001-8244-7831
Contribution to Project: Dr. Gokey performed organoid studies and IHC studies.

Name: Chase J. Taylor
Project Role: Research Assistant
Nearest person month worked: 7
Research Identifier (e.g. ORCID ID): N/A
Contribution to project: Mr. Taylor processed biospecimens for scRNA-seq library preparation and assisted with data analysis. He coordinated transfer of sequencing libraries to Tgen, and performed RNA-ISH preliminary validation studies.

Name: Peter Ghattas
Project Role: Research Assistant
Nearest person month worked: 3
Research Identifier (e.g. ORCID ID): N/A
Contribution to project: Mr. Ghattas assisted Mr. Taylor with biospecimen processing, storage, and organization.

Name: Natalie Geis
Project Role: Research Assistant
Nearest person month worked: 3
Research Identifier (e.g. ORCID ID): N/A
Contribution to project: Ms. Geis assisted with biospecimen processing and IHC studies.

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

Jonathan Kropski MD (PI)
New: VUMC89846(1U01HL152976-01 (Kotton)); Unspecified-Three Lakes Foundation (Kropski)
Ended: 2018099 Doris Duke Charitable Foundation (Kropski); 5K08HL130595-06 (Kropski)

What other organizations were involved as partners?

None

8. Special Reporting Requirements

Collaborative award

9. Appendices

Nothing to report