CHARACTERIZING MONOCYTES/MACROPHAGES IN PULMONARY FIBROSIS

TARGET IDENTIFICATION THROUGH THE TRANSCRIPTOMIC CHARACTERIZATION OF PROFIBROTIC MONOCYTES/MACROPHAGES IN IDIOPATHIC PULMONARY FIBROSIS

By MEGAN VIERHOUT, B.Sc. (Honours)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

MASTER OF SCIENCE (2020) McMaster University
(Medical Sciences) Hamilton, Ontario

TITLE: Target Identification Through the Transcriptomic Characterization of Profibrotic Monocytes/Macrophages in Idiopathic Pulmonary Fibrosis

AUTHOR: Megan Vierhout, B.Sc. (Honours)

SUPERVISOR: Dr. Kjetil Ask, Ph.D

NUMBER OF PAGES: lxxxiii, 83

LAY ABSTRACT

Idiopathic pulmonary fibrosis (IPF) is a disease of unknown cause that results in excessive scarring of the lungs and progressive impairment in lung function. We believe that white blood cells called monocytes and macrophages play a key role in the development and progression of this disease. Overall, it is thought that monocytes, which circulate in the blood, enter the lung tissue and become macrophages. These macrophages then lead to the formation of scar tissue, which is characteristic to IPF. In order to better understand how these cells contribute to IPF, we studied their properties in blood and lung biopsies from IPF patients. We found significant differences between monocytes/macrophages in IPF than those in healthy controls, that may help explain disease progression. We hope that these findings will provide insight into causes of the IPF, and potential avenues for therapeutic intervention.

ABSTRACT

Idiopathic pulmonary fibrosis (IPF) is a disease of unknown pathogenesis characterized by scarring of the lung and declining respiratory function. Originating from bone marrow, circulating monocytes can be recruited into the lung tissue and polarized toward the alternatively activated (M2) profibrotic macrophage phenotype. Recent literature has shown that cluster of differentiation 14 positive (CD14+) monocytes are more abundant in IPF patient blood and are associated with disease outcome and acute exacerbation. Additionally, a 52-gene risk profile from peripheral blood mononuclear cells for outcome prediction in IPF was recently published. Here, we began by characterizing macrophages in human IPF lung tissue. We then assembled a biobank and examined transcriptomic characteristics of blood-derived circulating monocytes from IPF patients.

Various histological assessments were completed on a tissue microarray including lung biopsies from 24 IPF patients and 17 controls, to characterize M2 macrophage expression in human tissue. Whole blood samples were collected from 50 IPF patients and 12 control subjects. CD14+ monocytes were isolated and mRNA was extracted for bulk RNA sequencing. Data were analyzed for differential expression (DE), and Gene Set Enrichment Analysis (GSEA) was performed to examine enrichment of the previously published 52-gene risk profile in our dataset.

We found that M2 macrophage expression was increased in IPF lung tissue compared to controls. CD14+ monocyte levels were significantly elevated in IPF patients

in our cohort compared to control participants, and was negatively correlated with forced vital capacity (FVC). DE analysis comparing IPF and control monocytes yielded a 35-gene signature, with 16 up-regulated genes and 19 down-regulated genes. When comparing the signature related to long transplant-free survival from the published dataset to our data, GSEA demonstrated that this signature is enriched in donors from our dataset, supporting concurrence between the meanings of the two datasets. Overall, these results provide insight to identify targets to modulate monocyte/macrophage function in IPF and potentially affect progressive disease.

ACKNOWLEDGEMENTS

I am immensely grateful for all of the support that I have received throughout the journey of completing my Masters degree. The past two years have been an exciting expedition, and I would like to thank everyone who helped shape this experience. First and foremost, thank you to my supervisor, Dr. Kjetil Ask. Words cannot thank you enough for your continuous positivity, creativity, and guidance throughout the past two years. Your mentorship has shaped me into the scientist I am today. Thank you for all the opportunities and open doors you have provided for me. Thank you for your kindness.

Secondly, I would like to extend my utmost thanks to my supervisory committee members: Dr. Jeremy Hirota, Dr. Nathan Hambly, and Dr. Asghar Naqvi. I have been extremely fortunate to work with you. Thank you for all of the ideas, guidance, and support you have provided. Your innovation and expertise are unmatched.

My fellow graduate students: Olivia Mekhael and Aaron Hayat! I could not have asked for better grad colleagues to share this experience with. I am continuously amazed by your ambition and talent. Thank you for making each and every day enjoyable and memorable.

To the Ask Lab members, past and present- Spencer Revill, Anmar Ayoub, Sonia Padwal, Alex Ognjanovic, Safaa Naiel, Jane Ann Smith, Soumeya Abed, Eric Chu, Victor Tat, Tran Zhou, Limin Liu, Pari Yazdanshenas, and Nicole Bodnariuc. Thank you for your multifaceted support and friendship throughout this process. I am so grateful to have been part of this team.

To our amazing bioinformatician, Dr. Anna Dvorkin-Gheva! I have learned so much for you and am so thankful for all of your help throughout my masters. Thank you for your direction and patience.

To my family, thank you for being a constant in my life throughout this time of great change. You are my biggest support and I cannot thank you enough.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

ATF6	Activating transcription factor 6
BiP	Binding immunoglobulin protein
CCL18	Chemokine ligand 18
CD	Cluster of differentiation
DLCO	Diffusing capacity for carbon monoxide
ER	Endoplasmic reticulum
FEV1	Forced expiratory volume in 1 second
FFPE	Formalin-fixed paraffin-embedded
FVC	Forced vital capacity
GEO	Gene Expression Omnibus
GSEA	Gene set enrichment analysis
H&E	Hematoxylin and eosin
IHC	Immunohistochemistry
IL	Interleukin
ILD	Interstitial lung disease
IPF	Idiopathic pulmonary fibrosis
IRE1α	Inositol requiring enzyme 1 alpha
M1	Classically activated macrophage phenotype
M2	Alternatively activated macrophage phenotype
PBMC	Peripheral blood mononuclear cells
PCA	Principal component analysis
PCR	Polymerase chain reaction
PERK	PKR-like endoplasmic reticulum kinase
RIN	RNA integrity number
sXBP1	Spliced X-box binding protein 1
TMA	Tissue microarray
UPR	Unfolded protein response
TGF-β	Transforming growth factor beta
TMA	Tissue microarray
TFS	Transplant-free survival
XBP1	X-box binding protein 1

DECLARATION OF ACADEMIC ACHIEVEMENT

This is a declaration that the work included in this thesis was completed by Megan Vierhout, under the supervision of Dr. Kjetil Ask. Bioinformatic analysis was completed by Dr. Anna Dvorkin-Gheva.

CHAPTER 1: INTRODUCTION

1.1 Fibrotic Lung Disease

1.1.1 Overview of Fibrotic Lung Disease

In the developed world, 45% of all mortalities are associated with chronic fibroproliferative disease (Wynn, 2008). Fibrosis, defined by the overabundant deposition of extracellular matrix components such as collagen and fibronectin in regions of damaged and inflamed tissue, can affect several organs in the body (Wynn, 2008; Wynn & Ramalingam, 2012). The pathologic deposition of these materials can lead to extensive tissue remodeling, hence having the potential to cause organ failure and death (Wynn, 2008). Interstitial lung disease (ILD) is an umbrella term for diseases characterized by inflammation and fibrosis of the parenchyma. Patients with fibrotic ILD experience progressive fibrosis of the lung, which involves declining lung function, dyspnea, and overall reduced quality of life. (Wong et al., 2020). Overall, fibrotic diseases are poorly understood with limited treatment options, warranting the exploration of etiology and treatment avenues in these conditions.

1.1.2 Idiopathic Pulmonary Fibrosis

Idiopathic pulmonary fibrosis (IPF) is the most common and severe subtype of idiopathic interstitial pneumonia (Wong et al., 2020). IPF is a disease of poorly understood pathogenesis that is characterized by scarring of the lung and declining respiratory function. Due to the progressive nature of the disease, many patients die of respiratory failure, with a median prognosis of a median 2.8 to 4.2 years of survival after

diagnosis Boon et al., 2009; Ley et al., 2011; Maher et al., 2007). Despite the aggressive natural history, treatment options are limited and not curative (Moore et al., 2014). Common symptoms include dyspnea, and dry cough, with characteristic signs including inspiratory crackles, clubbing of the digits, weight loss, and pedal edema (Meltzer & Noble, 2008; Gogali & Wells, 2012). Possible risk factors for IPF include smoking, gastroesophageal reflux disease, genetic predisposition, and exposure to metal dust environmental (Zaman & Lee, 2018). Currently, there are approximately 14,000 to 15,000 patients living with this disease in Canada, with older adults being the most affected subpopulation (Hopkins et al., 2016).

1.1.3 Treatments for IPF

As of 2014, there are two current FDA-approved treatments for IPF: Nintedanib and Pirfenidone. Both of these drugs are antifibrotic medications and are recommended for the treatment of IPF in accordance with the American Thoracic Society/European Respiratory Society Clinical Practice Guidelines. Nintedanib functions as a small molecule inhibitor of tyrosine kinase receptors, such as fibroblast growth factor receptor, platelet-derived growth factor receptor, and vascular-endothelial growth factor receptor. Overall, in clinical trials, Nintedanib was seen to reduce the rate of decline of lung function (Richeldi et al., 2016). Pirfenidone functions as an antifibrotic and anti-inflammatory agent. This medication reduces proliferation of fibroblasts and the accumulation of collagen, and has been seen to target the mitogen-activated protein kinase pathway (Hall et al., 2018). In clinical trials, Pirfenidone was also seen to reduce

the rate of lung function decline (King et al., 2014; Saito et al., 2019). Recent trials have also demonstrated that combined therapy with these two anti-fibrotic medications is safe and tolerable for patients, however evaluation of effectiveness still requires further elucidation (Flaherty et al., 2018; Richeldi et al., 2019). Although both of these drugs decline the rate of progression of IPF, they do have considerable side effects, which may limit quality of life in patients taking the medications. Therefore, improve treatment options are still required in this disease.

1.1.4 Pulmonary Function Measures

Pulmonary function tests are used as one of the techniques to assess the status and progression of IPF, using spirometry, body plethysmography, and measurement of gas transfer to quantify lung function. Spirometry is completed performing a forced expiratory maneuver, whereby the volume expired and airflow are measured. Spirometry is a favourable testing method due to its sensitivity, reproducibility, and limited invasiveness (Moore, 2012). Common measures obtained include forced vital capacity (FVC), which is the maximum volume of air that a patient can exhale when blowing out as fast as they are able, and forced expiratory volume in 1 second (FEV1). Diffusing capacity for carbon monoxide (DLCO) is a valuable measure evaluating how efficient gas transfer is from the alveoli to the bloodstream (Johnson, 2000). In combination, these tests can provide a great deal of insight into the lung health and function of IPF patients.

1.2 Circulating Monocytes

1.2.1 Overview of Circulating Monocytes

Monocytes are white blood cells derived from progenitor cells in the bone marrow. After leaving the bone marrow, these cells circulate in the peripheral blood and play a large role in the body's innate and adaptive immune systems. These monocytes extravasate from the circulatory compartment and enter various tissues in the body where they are acted upon by various stimuli, such as cytokines and growth factors, leading them to differentiate into macrophages. (Zhang et al., 2018). In the setting of tissue injury, monocytes are recruited from the bone marrow to the site of injury, where they often exceed the quantity of tissue-resident macrophages (Wynn & Vannella, 2016). Monocyte subpopulation is determined by levels of expression of Cluster of Differentiation (CD) 14 and CD16. Classical monocytes comprise 85% of the monocyte population, and have high expression of CD14 (100 times higher than the isotype control) and no CD16 expression (CD14++CD16-). Intermediate monocytes make up 5% of total monocytes, and have high CD14 expression and low (10 fold higher than the isotype control) expression of CD16 (CD14++CD16+). Non-classical monocytes comprise 10% of all monocytes and have low expression of CD14 and high expression of CD16 (CD14+CD16++) (Ziegler-Heitbrock et al., 2010; Marimithu et al., 2018).

1.2.2 Cells Expressing CD14

CD14 is a protein on the cell surface that operates as an endotoxin receptor and binds lipopolysaccharide. It is strongly expressed on monocytes and tissue macrophages.

B-cells and neutrophils may also express CD14 to a lesser extent (Naeim et al., 2018). Dendritic cell subsets may also express CD14, however it has been speculated that these cells may be monocyte-derived (Collin et al., 2013), and are present in tissue only.

1.2.3 Circulating Monocytes and Progression of Fibrotic Lung Disease

Monocytes can be recruited from the peripheral circulation into the lungs where they become polarized into alternatively activated, or M2, profibrotic macrophages and act as potential key drivers of progressive fibrotic disease (Zhang et al., 2018). These macrophages can then enter both alveolar and interstitial spaces (Zhou & Moore, 2018). Overall, there is limited information on the prediction of mortality in IPF patients (Ley et al., 2011). However, recent work has raised monocytes to be a predictor of disease outcome. Scott et al. (2019) exhibited that CD14+ monocytes are more abundant in the blood of IPF patients and are associated with poor outcomes and disease progression, which suggests a potential causal role between these cells and fibrogenesis. A recent study has also shown that increased monocyte number is associated with acute exacerbations in IPF patients (Kawamura et al., 2020). It was found that patients with an increased monocyte count were more likely to develop an acute exacerbation, and patients with a higher absolute monocyte count had a shorter time to developing their first acute exacerbation (Kawamura et al., 2020).

Circulating monocytes have been demonstrated to play a role in the progression of IPF (Moore et al., 2014). It has been shown, through multiple murine studies, that the depletion of circulating monocytes significantly reduces the severity of fibrosis. Moore et

al. (2001) showed that the deletion of the main receptor for monocyte-chemoattractant protein-1 known as C-C Motif Chemokine Receptor 2, a receptor that is important in the recruitment of monocytes, protected mice from developing lung fibrosis. Gibbons et al. (2011) deleted the Ly6C^{hi} population of circulating monocytes, and found that fibrosis and the quantity of alternatively activated macrophages in the lung was reduced. Several studies have also exhibited that deletion of monocyte-derived alveolar macrophages, but not tissue-resident alveolar macrophages, ameliorated fibrosis (Misharin et al., 2017; Joshi et al., 2019). This supports that it is the monocytes from the circulatory department that contribute to fibrogenesis. Additionally, in a study examining lung-derived macrophages, it was discovered that circulating monocytes from IPF patients appeared to be predisposed to promote fibroproliferative activity prior to entering the lung (Zhou et al., 2014).

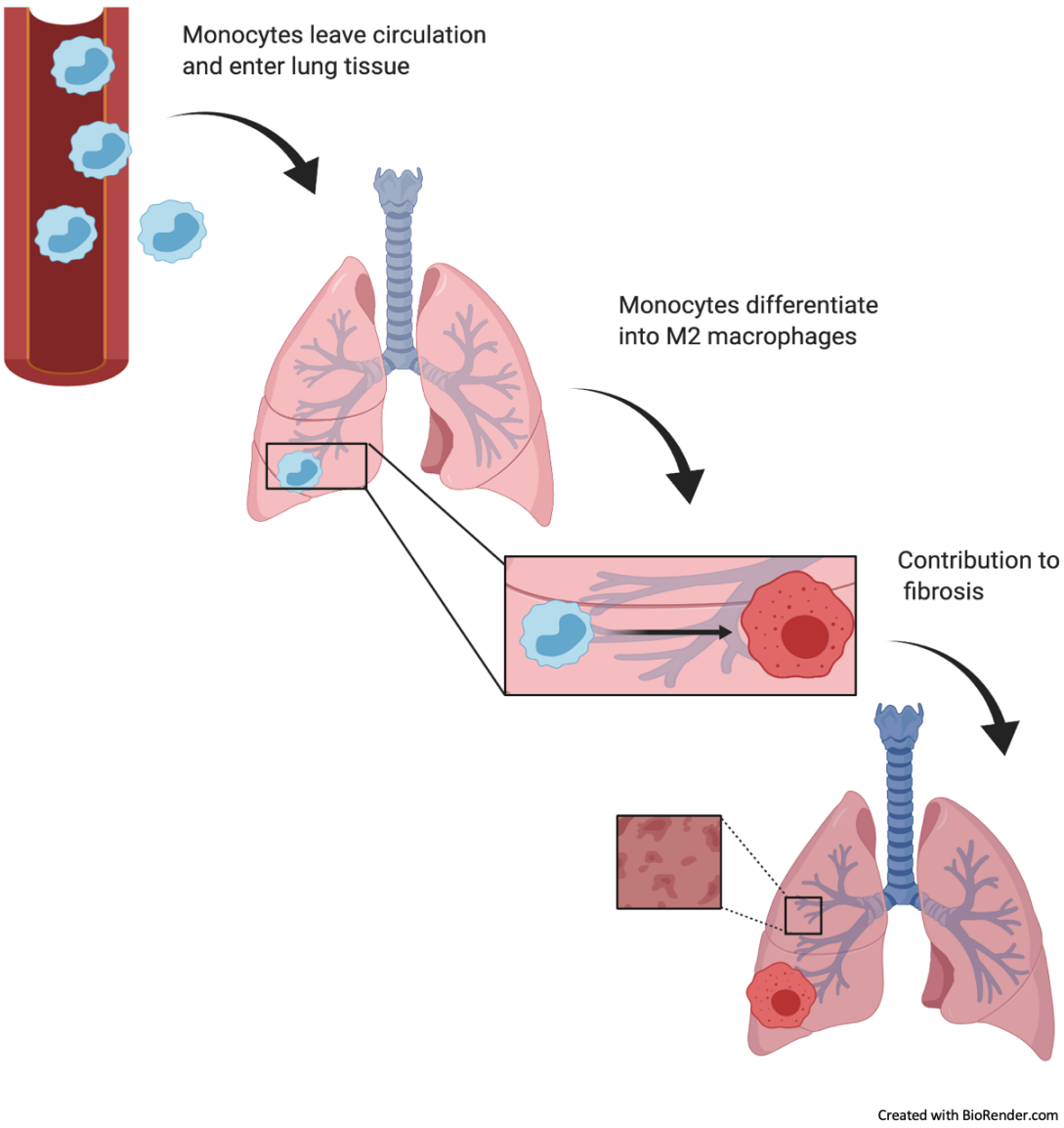


Figure 1: Contribution of Circulating Monocytes to Lung Fibrosis

Monocytes leave the circulation and enter the lung tissue, where they differentiate into alternatively activated macrophages, and contribute to fibrogenesis.

1.3 Alternatively Activated Macrophages

1.3.1 Alternatively Activated Macrophages and Fibrotic Lung Disease

Macrophages, as regulator and effector cells, hold the potential to modulate immune function and are involved in inflammatory defense functions (Sica & Mantovani, 2012). Similar to the TH1-TH2 polarization of T-cells, macrophages may become activated to the M1 (classically activated) phenotype, or M2 (alternatively activated) phenotype (Stifano & Christmann, 2016; Sica & Mantovani, 2012). The M1 phenotype is stimulated by interferon gamma, lipopolysaccharide, and toll-like receptor ligands, and is proinflammatory with microbicidal and tumoricidal properties (Stifano & Christmann, 2016; Sica & Mantovani, 2012; Braga et al., 2015). M2 macrophages are anti-inflammatory, and are known to stimulate tumour growth and aid in organ regeneration. They can also be classified as profibrotic, and release profibrotic factors including transforming growth factor beta (TGF-β) and Galactin-13 (Braga et al., 2015). Repeated microinjuries to the alveolar epithelium induces the production of TGF-β in macrophages, which leads to differentiation of fibroblasts into active myofibroblasts and thus deposition of extracellular matrix components (Fernandez & Eickelberg, 2012). M2 macrophages are polarized in the presence of interleukin (IL)-4 and IL-13, and previous work by our group has also demonstrated that IL-6 mediates hyperpolarization (Ayaub et al., 2019). Chemokine ligand 18 (CCL18), which is produced by alveolar macrophages, is a marker that has been positively correlated with the decline of lung function, and elevated levels in the serum of IPF patients were correlated with mortality (Schupp et al., 2014). In

addition, the expression of mannose receptor CD206 is a feature that characterizes M2 macrophages (Braga et al., 2015). CD206 is found on the surface of alternatively activated macrophages (Kazuo et al., 2019). CD163 is another M2 macrophage marker, while CD68 is a pan-macrophage marker (Minami et al., 2018). CD163 is a scavenger receptor responsible for the uptake of hemoglobin-haptoglobin complexes, and is a membrane-bound protein found on tissue macrophages (Kristiansen et al., 2001; Fabriek et al., 2005). CD68 is a glycoprotein highly expressed on macrophages, and is found in the endosomal and lysosomal compartments of cells, as well as the cell surface (Chistiakov et al., 2017). Throughout this thesis, CD206 and CD163 will be referred to as "M2" macrophage markers, and CD68 as a "pan" macrophage marker. Due to their involvement in matrix remodelling and fibrogenesis, as well as their association with the properties of progression of IPF, M2 macrophages constitute a valid target for investigation in fibrotic lung disease.

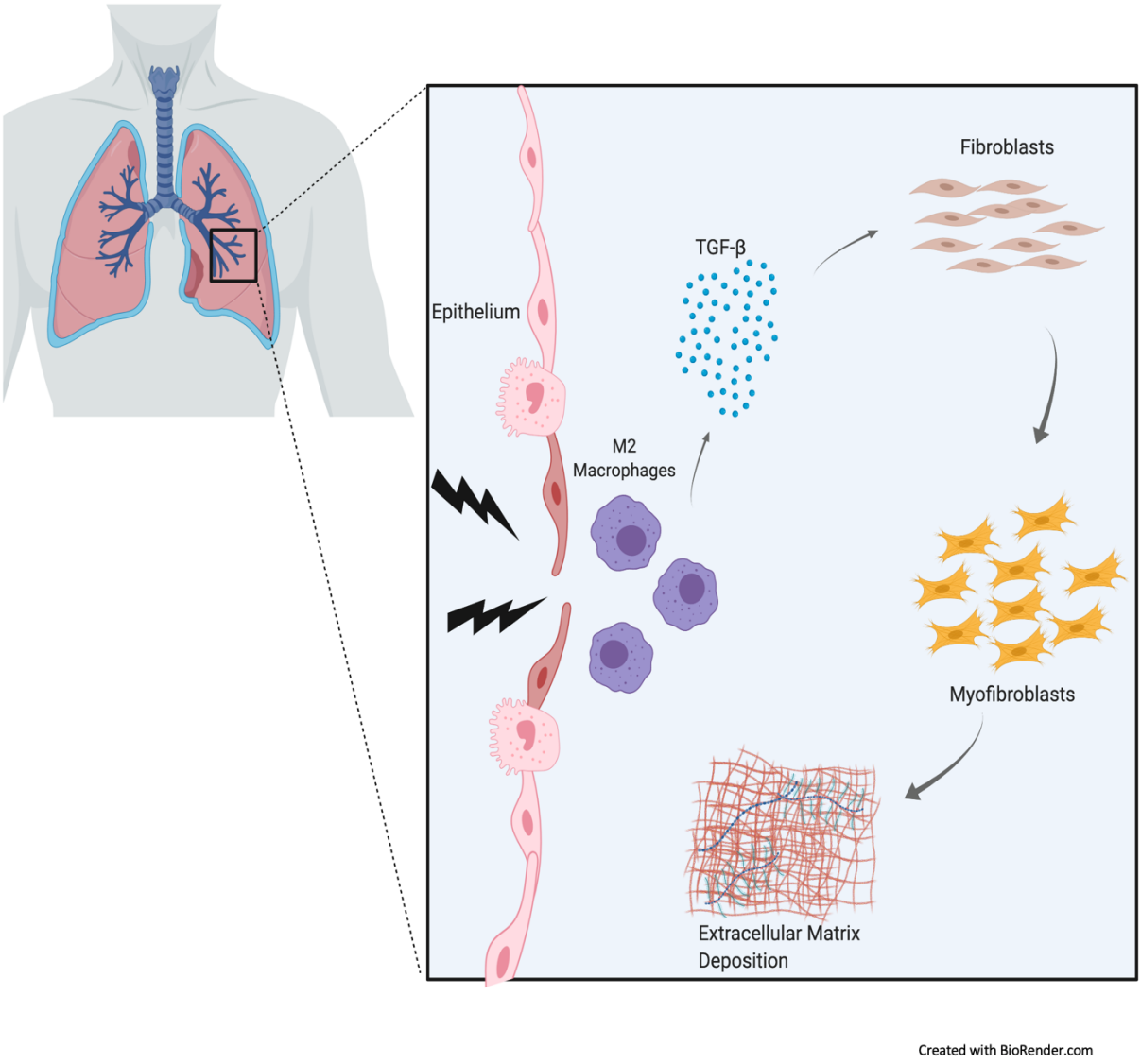


Figure 2: Contribution of M2 Macrophages to Fibrogenesis

In the case of epithelial injury, M2 macrophages are recruited to the site. TGF-β is secreted, which causes differentiation of fibroblasts into myofibroblasts, and subsequent extracellular matrix deposition.

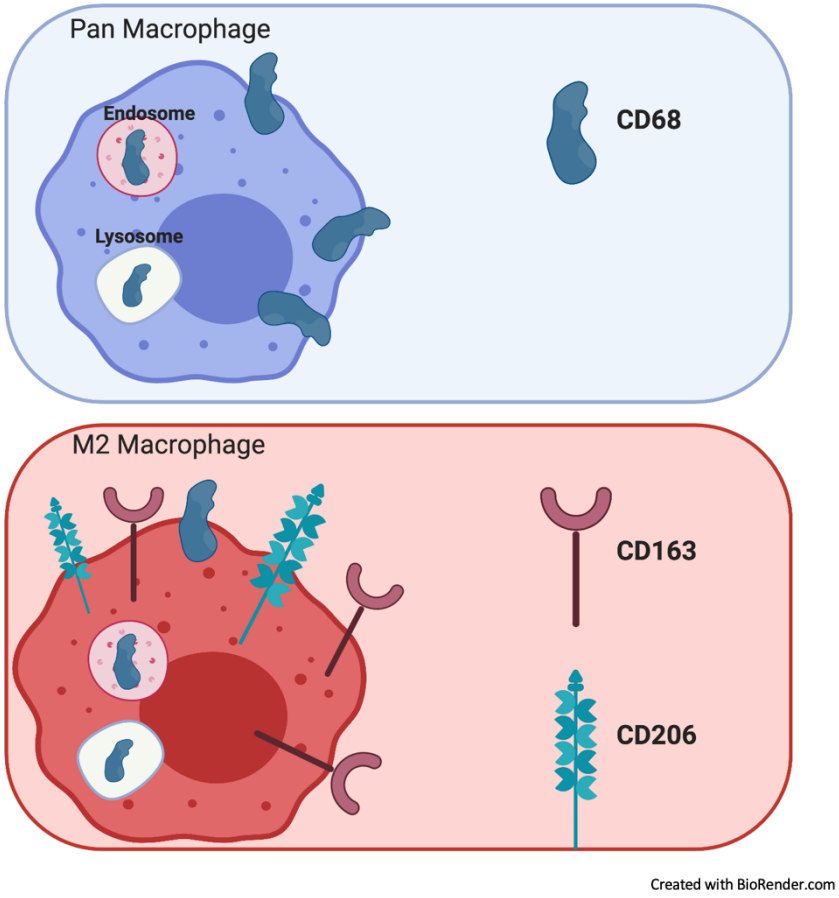


Figure 3: Macrophage Markers

As depicted here, CD68 is a pan macrophage marker (expressed on all macrophages), while CD163 and CD206 are specific to the M2 phenotype. CD68 is found in endosomes and lysosomes as well as the cell surface, while CD163 and CD206 are found on the cell surface.

1.3.2. Endoplasmic Reticulum Stress/The Unfolded Protein Response and M2 Macrophages

Endoplasmic reticulum (ER) stress is involved in fibrotic disease, and plays a role in disease progression. As a key organelle, the ER is the moderator of protein homeostasis and protein folding quality control. The role of the ER includes protein folding, assembly,

and trafficking, and degrading proteins that are misfolded or defective (Burman et al., 2018). Protein synthesis occurs on the ribosomes, after which proteins are translocated to the lumen where they are folded and modified. Most proteins are transported to other locations in vesicles, but those that are misfolded remain at the site of the ER (Marcinak & Ron, 2010). Conditions that disrupt protein processing, including increased protein load, heightened metabolic demands, and decreased chaperone efficiency, cause misfolded proteins to accumulate and effectuate ER stress. Following the onset of ER stress, the unfolded protein response (UPR) is activated, with the goal of restoring protein homeostasis. The UPR pathway has three arms, which are each controlled by an ER transmembrane protein. These include PKR-like endoplasmic reticulum kinase (PERK), activating transcription factor 6 (ATF6), and inositol requiring enzyme 1 alpha (IRE1α). A chaperone called binding immunoglobulin protein (BiP) remains bound to these three proteins when there is no ER stress. However, when there is protein accumulation and cell stress occurs, BiP becomes sequestered from these proteins and UPR is initiated (Burman et al., 2018). Additionally, the ER can expand to have an increased surface area and volume, thus enhancing its ability to fold proteins (Wei et al., 2013). In fibrotic lung disease, chronic ER stress and the UPR play a disease-causing role, such as through modulation of processes of fibrosis, inflammation, and epithelial injury (Burman et al., 2018; Wei et al., 2013). ER stress can contribute to the process of differentiation of fibroblasts to myofibroblasts, which are responsible for the synthesis collagen and extracellular matrix materials (Burman et al., 2018). Thus, it is believed that if we further

evaluate ER stress and the UPR in the setting of fibrotic lung disease, we can gain insight to potentially slow or halt the progression of fibrosis.

The IRE1 pathway is the most highly conserved of all arms of the UPR, being present in both yeast and higher eukaryotes. IRE1α is classified as a serine/threonine protein kinase and endoribonuclease (Hetz et al., 2011). When BiP is sequestered from IRE1α, an RNase domain becomes activated through oligomerization and autophosphorylation, that cleaves X-box binding protein 1 (XBP1) into its active form, spliced X-box binding protein 1 (sXBP1) (Burman et al., 2018). This occurs by the splicing 26 nucleotides from XBP1, which leads to a frameshift and presence of a C-terminal activation domain (Wu et al., 2015). sXBP1 is then translocated to the nucleus and activates genes of the UPR, such as ER degradation enhancing α-mannosidase like protein, which allows for more degradation of misfolded proteins through the ER-associated degradation pathway (Wu et al., 2015; Burman et al., 2018). IRE1 and ATF6 activation cause an increase in the quantity of chaperones in order to increase the protein folding capacity of the cell (Wei et al., 2018). It has also been speculated that the IRE1/XBP1 pathway plays a role in monocyte to macrophage differentiation (Dickhout et al., 2010). When activated in macrophages the pathway also plays a role in the expression of inflammatory cytokines, including IL-6 and tumour necrosis factor alpha (Martinon et al., 2010). Other results supporting the involvement of the IRE1 arm in fibrotic disease include its requirement for myofibroblast activation, the link of myofibroblast ER expansion to XBP1 splicing, the attenuation of liver fibrosis with IRE1α inhibition, and

the reduction of skin fibrosis with IRE1α inhibition (Heindryckx et al., 2016). In another study, Western-blot analysis of IPF lung tissues as compared to chronic obstructive pulmonary disease and human donor lung tissues exhibited that sXBP1 protein was only present in the IPF lungs. In the same study, when comparing sXBP1 transcripts levels between IPF and healthy donors, it was found that the transcripts were only present in the IPF lungs as well (Korfei et al., 2008). Previous work done by our group has drawn the parallel between M2 macrophages and the IRE1/XBP1 pathway. It was shown that XBP1 mRNA is increased in alternatively activated macrophages in vitro, and that treatment with an IRE1 inhibitor led to a decrease in CCL18, supporting its role in polarization of macrophages to the M2 phenotype and thus involvement in pro-fibrotic processes (Ayaub et al., 2019). Therefore, it is justifiable to explore this arm of the UPR as a target in fibrotic lung disease, and attempt to translate previous findings to the patient setting.

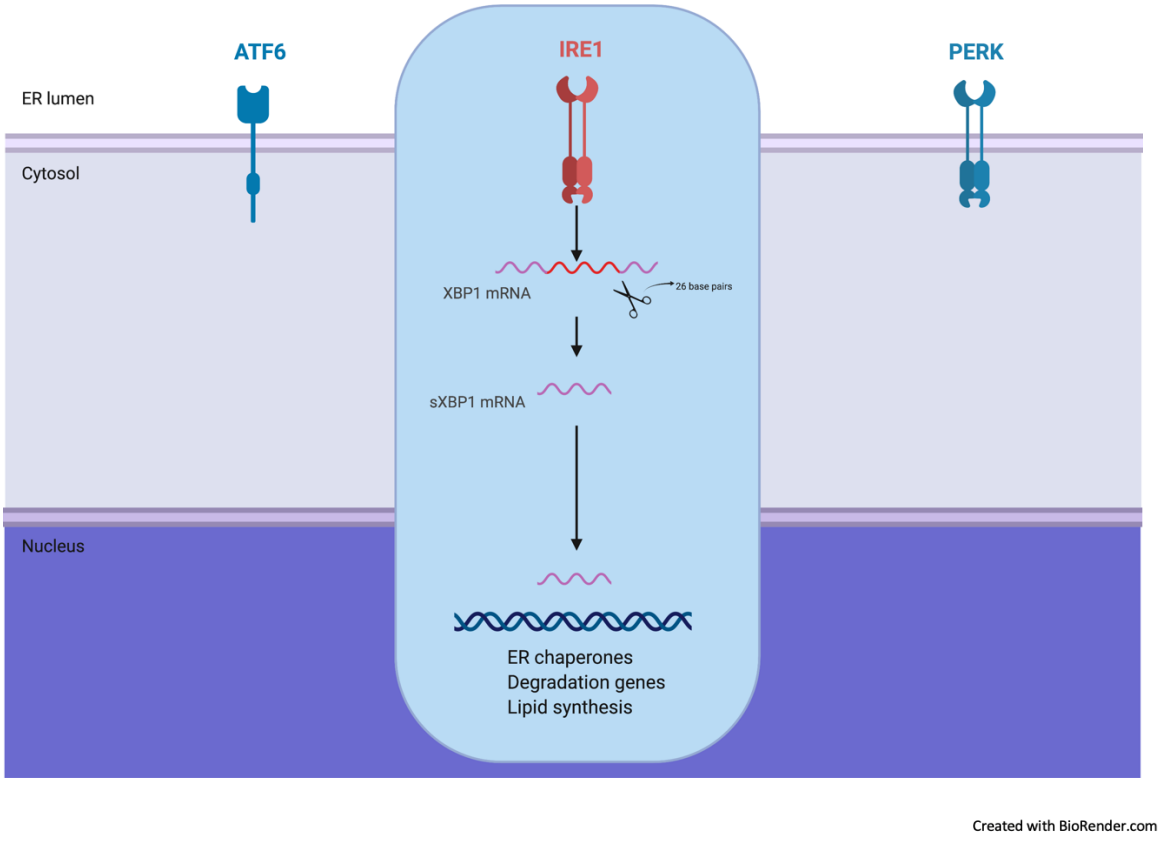


Figure 4: Unfolded Protein Response

The IRE1 pathway is one of three arms of the UPR, along with ATF6 and PERK. When the IRE1 pathway is activated, XBP1 is spliced by 26 base pairs to form sXBP1. sXBP1 is then translocated to the nucleus to activate various UPR genes.

1.4 Transcriptomic Signatures in Fibrotic Lung Disease

In exploration of transcriptomic characteristics of fibrotic lung disease, a study by Boon et al. (2009) has compared the gene profiles obtained from surgical lung biopsies of IPF patients. Since the rate of progression is variable, they aimed to elucidate a gene expression profile that would distinguish between stable and rapidly progressing disease, based on changes in FVC and DLCO over a one year period. Overall, this study defined a molecular expression signature of 134 transcripts that distinguishes stable and progressive IPF (Boon et al., 2009). Another group has explored transcriptomics in the blood compartment of IPF patients. Herazo-Maya et al. (2017) discovered a 52-gene signature from peripheral blood mononuclear cells (PBMC) from IPF patients. This signature was associated with outcome prediction in IPF.

Although human genetic signatures for IPF, both in the lung and blood compartments, have been investigated, this has not yet been explored in monocytes. Monocytes may constitute a valid target for therapeutic intervention in IPF, as they are critical for the fibrotic process, as demonstrated through murine studies, and are the precursors for macrophages in the lung. Therefore, we aimed to explore the characteristics of profibrotic monocytes/macrophages in IPF, and how they may contribute to progression of the disease.

1.5 Hypothesis

Overall Hypothesis

I hypothesize that profibrotic monocytes and macrophages contribute to the fibrotic process in IPF. Circulating profibrotic monocytes are believed to enter the lung and then become polarized to profibrotic macrophages, which contribute to the process of excessive scarring. These cells are characteristically different in IPF compared to healthy controls, thus constituting them as a valid target in this disease.

Specific Hypothesis

I hypothesize that circulating profibrotic monocytes from IPF patients possess a gene signature that differentiates them from healthy controls. I further hypothesize that the transcriptomic characteristics of these monocytes are associated with the progressive nature of IPF, as these cells enter the lung tissue and differentiate into profibrotic macrophages, which are present and increased in the lungs of IPF patients.

1.6 Objectives

1.6.1 Objective 1: Assessment of expression of macrophages in human IPF lung tissue

Since it is believed that monocytes differentiate into profibrotic macrophages in the lung which are responsible for fibrogenesis, we began by examining macrophage presence and expression in IPF. Although monocytes are the precursors for these macrophages, macrophages are the cells that are thought to play an active role in the fibrotic processes in the lung. Therefore, it is reasonable to study monocytes if profibrotic macrophages are characteristic of IPF lung pathology.

Prior to exploring the blood compartment in the following aims, we sought to study formalin-fixed paraffin-embedded (FFPE) lung tissue samples from IPF patients. Samples from 24 IPF patients (from both fibrotic and non-fibrotic areas) were assessed using histological staining for various macrophage markers, including CD206, CD163, and CD68. In-situ hybridization techniques were also utilized to examine potential colocalization of CCL18 and CD68 in IPF lung tissue, to evaluate whether the macrophages in these samples are of the M2 phenotype and are able to produce CCL18. Samples from the FFPE tissue were also cored to look at RNA levels of CCL18 with NanoString® Technology. Lastly, the markers for the activation of the IRE1/XBP1 UPR pathway were examined in these tissues via in-situ hybridization, as our lab has previously linked this pathway to M2 macrophages, and determined that its activation is critical for polarization to the M2 phenotype. Overall, these characterizations confirm the presence and involvement of alternatively activated macrophages in IPF, identify them as a potential target, and provide justification and a groundwork to study monocytes.

1.6.2 Objective 2: Collection of patient blood samples and generation of biobank

In order to study monocytes in IPF, the collection of whole blood samples from patients is required. This was done in collaboration with physicians Dr. Nathan Hambly, Dr. Gerard Cox, and Dr. Martin Kolb in the clinics of the Firestone Institute for Respiratory Health. Samples were collected from IPF patients and healthy controls, from which monocytes were isolated from. Where possible, serial samples were also collected from IPF patients at six month intervals, in order to study disease progression. For

potential future evaluation, PBMC, T-cells, and plasma were collected and added to the biobank.

In order to collect numerous samples suitable for transcriptomic evaluation and RNA sequencing, it is necessary to conduct evaluations to determine the quality of samples. Average blood volume obtained from participants, RNA quantity per 100,000 monocytes, and RNA integrity number (RIN) were calculated. This provides insight into overall suitability of the samples for transcriptomic evaluation, as well as quality control of the samples collected in the biobank. This lays the framework for further analysis of these cells.

1.6.3 Objective 3: Investigate transcriptomic characteristics of monocytes from IPF patients with fibrotic disease and correlate these characteristics with disease progression

As previously explained, there are limited predictors of mortality in IPF (Ley et al., 2011). A recent study conducted by Scott et al. (2019) showed that the quantity of monocytes can act as a predictor of poor outcomes in IPF, however little is known about the characteristics of these monocytes. First, we aimed to confirm the expression of CD14 in monocytes in IPF with analysis of publicly available Gene Expression Omnibus (GEO) single cell RNA sequencing datasets. Different cell populations were analyzed for their expression of CD14, and it was determined which cell population most highly expresses this gene. Then, we aimed to confirm our findings with published literature, where monocyte quantity was compared between IPF and control patients to determine if they

are more abundant in disease. Assessing if monocyte quantity is correlated with pulmonary function methods also provides insight into the involvement of these cells in progression of disease. We then compared the transcriptomic characteristics of these monocytes between IPF and control, by performing bulk RNA sequencing and differential expression analysis. This elucidated any differences that exist in gene expression in the disease. With regards to progression, principal component analysis (PCA) was performed to observe any transcriptomic changes that may occur between samples collected six months apart from the same patient. A previous multicentre study has been conducted on PBMC in IPF, which identified a 52-gene signature associated with poor outcomes (Herazo-Maya et al., 2017). Although this signature is also blood based and PBMC include monocytes, it is not specific to this cell type. Therefore, we aimed to investigate the similarity between this dataset and ours using gene set enrichment analysis (GSEA). Overall, understanding the transcriptomic differences and characteristics in IPF will provide insight into explicating monocytes as viable target in IPF. Understanding the nature, activation status, and characteristics of these cells could allow us to target them prior to their entry into the lung, and potentially reprogram them from entering the profibrotic state.

CHAPTER 2: METHODS

2.1 Tissue Microarray Creation

The IPF TMA was created by previous Ask Lab Masters student, Karun Tandon, with the guidance of molecular pathologists (Dr. Asghar Naqvi and Dr. J.C. Cutz). Regions of interest were selected from hematoxylin and eosin (H&E) slides of the FFPE parent blocks from the surgical lung biopsy cohort, which were then punched and inserted into a host paraffin block using the TMA Master II (3D Histech Ltd) for creation of TMAs. This TMA contains 316 0.6mm diameter tissue cores from 24 IPF patients and 17 controls resected from non-involved regions of lung cancer biopsies. IPF cores were taken from both fibrotic and non-fibrotic regions. The collection and utilization of human tissues was approved by the Hamilton Integrated Research Ethics Board (HiREB# 11-3559 and 13-523-C).

2.2 Core Punching for RNA Extraction

Four 1.0mm diameter cores (with a tissue depth of approximately 4.0mm) were punched from fibrotic areas of IPF lung blocks, using the TMA Master II (3D Histech Ltd). Cores were stored in RNase-free tubes for future RNA extraction.

2.3 Immunohistochemistry (IHC)

Immunohistochemical staining of the TMAs was conducted at the John Mayberry Histology Facility at the McMaster Immunology Research Centre, using the Bond RX immunostainer (Leica). Slides were placed on the Bond RX and dewaxed and hydrated. The slides were then treated with Bond Epitope Retrieval 1 buffer for 20 minutes for

CD68 and CD206, and Bond Epitope Retrieval 2 for CD163. The IHC protocol was then performed using the Bond Polymer Refine Detection Kit (DS9800). The primary dilutions of the antibodies were as follows: CD68 (Dako M0876) 1:50, CD206 (Abcam ab64693) 1:8000, CD163 (Abcam ab182422) 1:1000.

2.4 In-situ Hybridization (RNAscope® and Basescope™ Technology)

RNAscope® (ACD Bio) in-situ hybridization for CCL18 (486618) and CD68 (560598-C2) duplex staining, and Basescope™ single-plex staining for XBP1 (715178) and sXBP1 (715188) were completed using commercially available assays. This was also completed at the John Mayberry Histology Facility at the McMaster Immunology Research Centre, using a program devised by ACD Bio for the Leica Bond RX immunostainer.

2.5 Slide Digitalization

Slides were digitalized using the Olympus VS 120 automated slide scanner. Images of the entire slides were acquired. For IHC stained slides, 20x magnification was utilized, while 40x magnification was utilized for in-situ hybridization stained slides.

2.6 HALO® Histological Quantification

HALO® histological image analysis software (Indica Labs) was used to quantify the scanned images of slides. Custom algorithms were created for each stain based on colour, intensity, and optical density. For quantification of IHC staining, the Multiplex IHC module was used to detect percent cell positivity. For slides stained in-situ hybridization, the ISH module was used to determine percent positive cells (that had at

least one probe puncta present) and average probe number per cell. All algorithms were used in conjunction with the TMA module in the software, that allows for proper segmenting of the TMA slides.

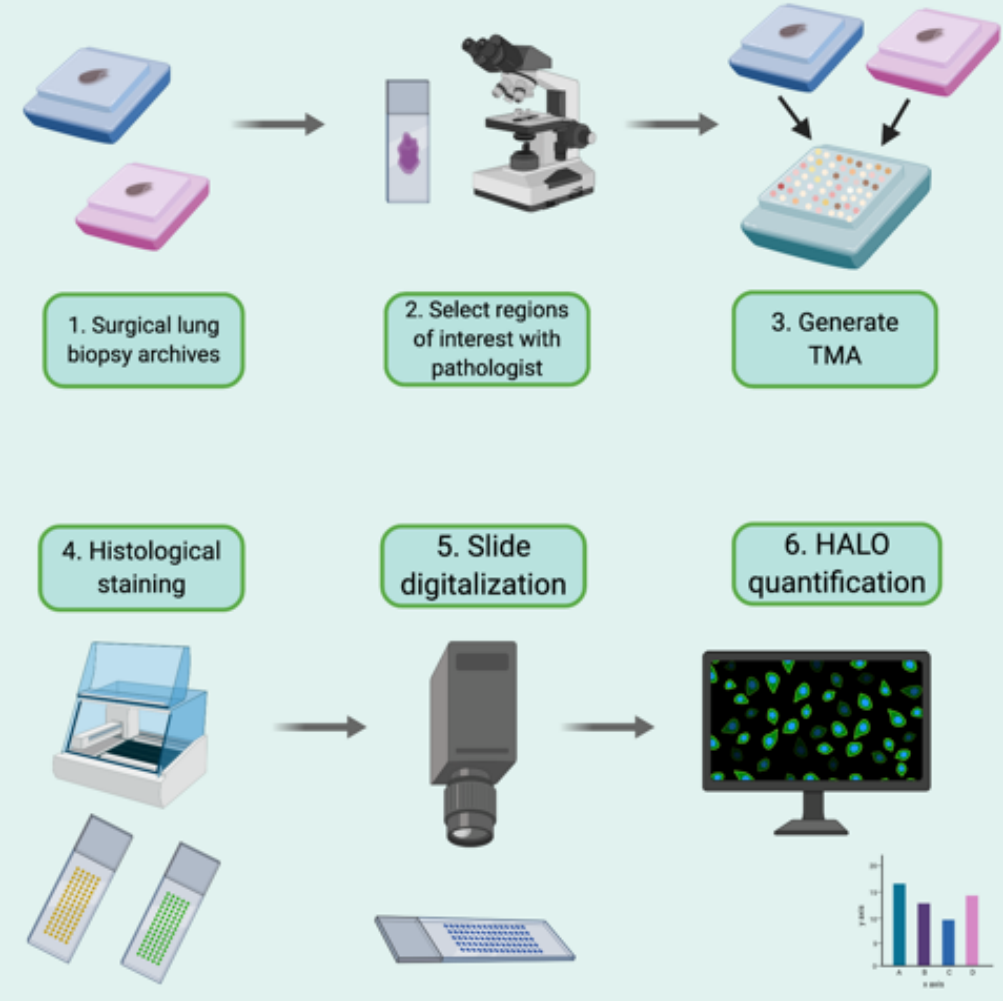


Figure 5: Overall workflow for molecular phenotyping and imaging analysis of FFPE tissues

As shown here, quantitative data from histological analyses of human tissues are obtained through multiple steps, including specimen acquisition, selection of regions of interest, TMA creation, histological assessments, and scanning and quantification.

2.7 Nanostring® Gene Expression Quantification

RNA samples were analyzed using the nCounter® Analysis System at the McMaster Farncombe Metagenomics Facility. Data were preprocessed using the total counts normalization method. The resulting gene expression data was processed using the nSolver® Analysis Software v. 2.6, and analyzed using R. Data are expressed as Log2 values.

2.8 Whole Blood Collection

Collection and utilization of human blood was approved by the Hamilton Integrated Research Ethics Board (HiREB# 2017). Informed consent was obtained from all participants. Consecutive whole blood samples (six months apart) were collected from IPF patients at the Firestone Institute for Respiratory Health, in K2EDTA blood tubes (BD 366643). Several components of the blood were stored for biobanking, including plasma, PBMC (in freezing media, in RNA later, or formalin-fixed), CD4+ T-cells (in RNA later), and CD14+ monocytes (in RNA later or RNA lysis buffer). A total of 66 IPF samples from 50 patients (36 patients with a single visit, 12 patients with two visits, and 2 patients with three visits) were collected for monocyte isolation. Samples were also collected from 12 healthy controls.

2.9 CD14+ Monocyte Isolation

CD14+CD16- monocytes were directly from whole blood isolated using immunomagnetic negative selection (Stemcell Technologies 19669). “The Big Easy” magnet (Stemcell Technologies 18002) was used with 3mL of blood. Isolation techniques

were completed according to the manufacturer's instructions. Monocytes were centrifuged at 1300 rpm for 10 minutes, and then resuspended in RNA lysis buffer (Macherey-Nagel 740906). Cells were stored at -80°C until RNA extraction.

2.10 RNA Extraction

For FFPE cores, RNA was extracted using the Nucleospin total RNA FFPE extraction kit (Macherey-Nagel 740982.50), according to the manufacturer's protocol. For frozen monocytes, mRNA was extracted from using the NucleoSpin® RNA Plus extraction kit (Macherey-Nagel 740984.250), according to the manufacturer's protocol. Sample concentration was obtained with a NanoDrop spectrophotometer. RNA quality was assessed at McMaster Farncombe Metagenomics Facility. For FFPE samples, RNA quality was checked with the Bioanalyzer 2100 (Agilent Technologies) and for monocytes with the High Sensitivity RNA ScreenTape® Device.

2.11 Bulk RNA Sequencing

RNA sequencing was completed at the McMaster Farncombe Metagenomics Facility, using the Illumina HiSeq 1500. The samples were sequenced with single end reads of 75 base pairs, at an average depth of 9.1 million clusters (ranging from 6 million to 14 million). Library preparation was done using the NEBNext Ultra Directional RNA LP kit with poly-A mRNA enrichment beads, which produces strand-specific data.

2.12 RNA Sequencing Data Preprocessing and Normalization

The mapping of the processed reads was performed by using HISAT2 (Kim et al., 2015) with hg38 (UCSC) reference genome and the reads were counted using HTSeq

(Anders et al., 2015). Genes showing low levels of expression across samples were removed, resulting in genes. The remaining values were normalized with *TMM* normalization method (Robinson & Oshlack, 2010) and then transformed with *voom* transformation (Law et al., 2014).

2.13 Analysis of Single Cell RNA Sequencing GEO Datasets

Data were obtained from publicly available datasets containing samples from donors and IPF patients (GEO: GSE122960 [termed 1st dataset] and GSE135893 [termed 2nd dataset]). Since both datasets contained additional samples from patients with other fibrotic diseases, only samples from healthy donors and IPF patients were selected for processing and analysis. Both processing and analysis were performed using *Seurat* package (Butler et al., 2018) in R. Differential expression analyses between cell populations and between IPF patients and donors within each of the cell populations were performed using *Seurat* package. Violin plots were obtained using *Seurat* package.

2.14 Differential Expression Assessment of CD14+ Monocyte RNA Sequencing Data

Differential expression analysis was performed using *limma* package (Ritchie et al., 2015), comparing samples obtained from IPF patients during their first visit to samples obtained from healthy donors with adjustment for monocyte counts. P-values were corrected with BH correction for multiple testing (Benjamini & Hochberg, 1995), and corrected values <0.05 were considered to be significant. Difference in monocyte counts between the samples obtained from donors and the samples obtained from IPF patients was assessed using *limma* as well.

2.15 Principal Component Analysis of CD14+ Monocyte RNA Sequencing Data

IPF samples from 1st and 2nd visits were examined using PCA, with samples collected from the 1st visit used as a reference to the matching samples from the 2nd visit. Clusters of different trends were obtained by using mclust package in R (Scrucca et al., 2002; Fraley & Raftery, 2002). For visualization, PCA analysis and plots were obtained using *rgl* package in R.

2.16 Enrichment Analysis of 52-gene Signature

Enrichment of the published 52-gene signature was examined using GSEA (Subramanian et al., 2005).

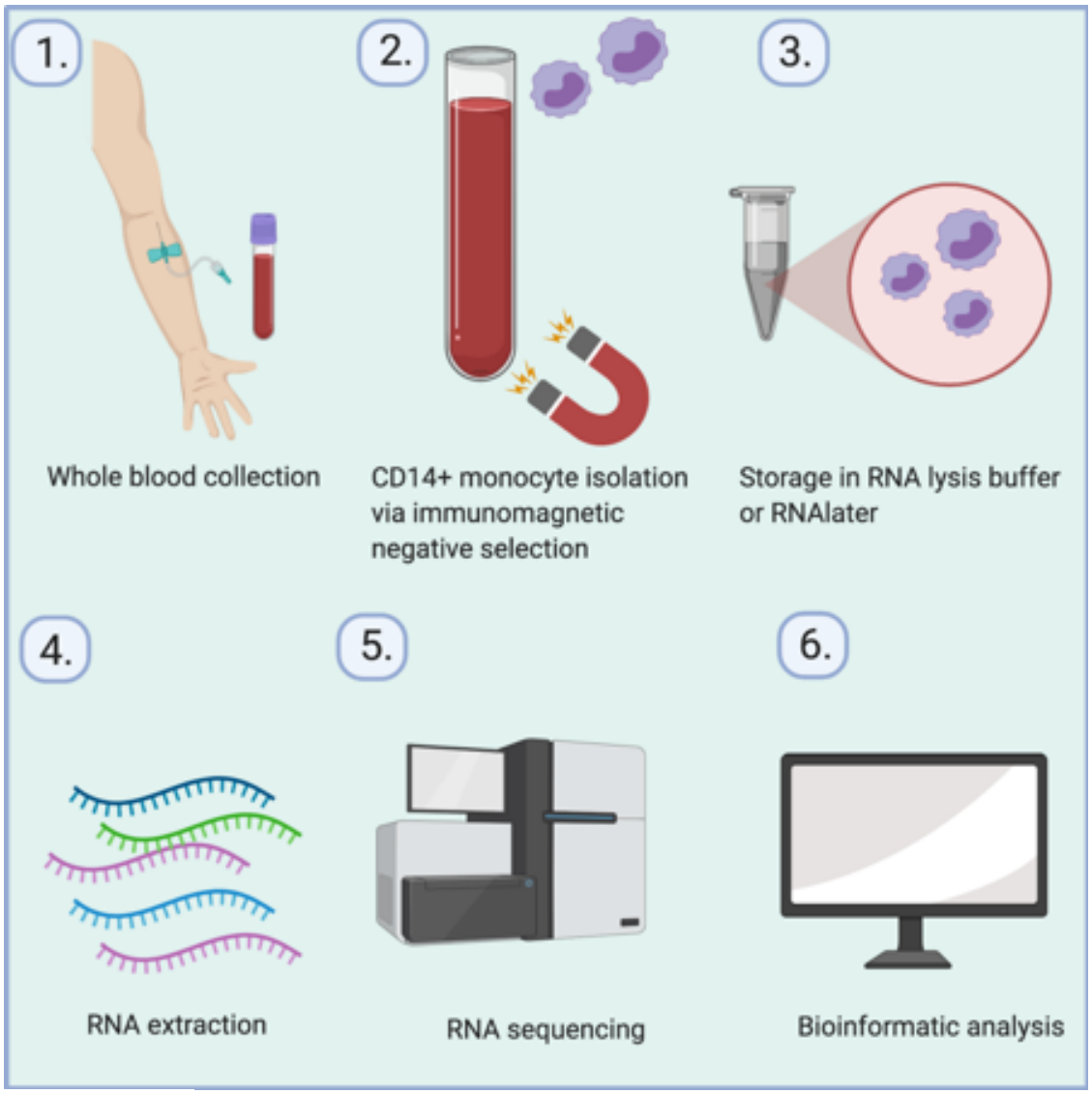


Figure 6: Overall workflow for the collection and transcriptomic analysis of CD14+ monocytes

Depicted above is the pipeline for blood sample collection and processing, to obtain transcriptomic data on CD14+ monocytes.

2.17 Statistical Analysis

All results are conveyed as mean ± SEM, unless otherwise noted. Statistical analyses were performed using GraphPad Prism (Version 8). When comparing groups, a two-tailed t-test was used to determine significance. Paired t-tests were using when comparing parameters from the same patient (ie. from non-fibrotic and fibrotic areas). A p-value of less than 0.05 was considered significant. Cohen's d was used to express effect size, using the formula: Effect size = $\frac{[Mean\ of\ experimental\ group]-[Mean\ of\ control\ group]}{Pooled\ Standard\ Deviation}$

Statistical analyses for Nanostring®, bulk RNA sequencing, and single-cell RNA sequencing data were performed in R.

CHAPTER 3: RESULTS

3.1 Objective 1: Assessment of expression of macrophages in human IPF lung tissue

3.1.1 There is no difference in expression of pan macrophage marker, CD68, in lung tissue from IPF patients compared to controls

In order to determine the overall expression of macrophages in human IPF lung samples, the IPF human lung tissue TMA was immunohistochemically stained for CD68. CD68 is a pan-macrophage marker that is highly expressed on macrophages, and found in the endosomal and lysosomal compartments, in addition to the cell surface (Chistiakov et al., 2017). Clinical characteristics for the IPF patient cohort included on the TMA can be found in **Table 1.** A scanned image of the entire slide was analyzed for CD68 positivity using the Multiplex IHC module in the HALO® Image Analysis Platform software. This was done to confirm previous findings in the lab, utilizing novel and updated algorithms and modules in this software. There were no significant differences seen in pan macrophage marker CD68 expression among fibrotic, non-fibrotic and control cores **(Figure 7 A,B)**.

Table 1: Clinical Characteristics of IPF Patients on TMA

Characteristic		Value
Gender	Male	17
	Female	7
Age (years)		59.0 ± 10.1
Disease Duration (years)		2.6 ± 2.8
FVC (% predicted)		62.9 ± 23.2
FEV1 (% predicted)		68.1 ± 16.5
DLCO (% predicted)		42.8 ± 10.9

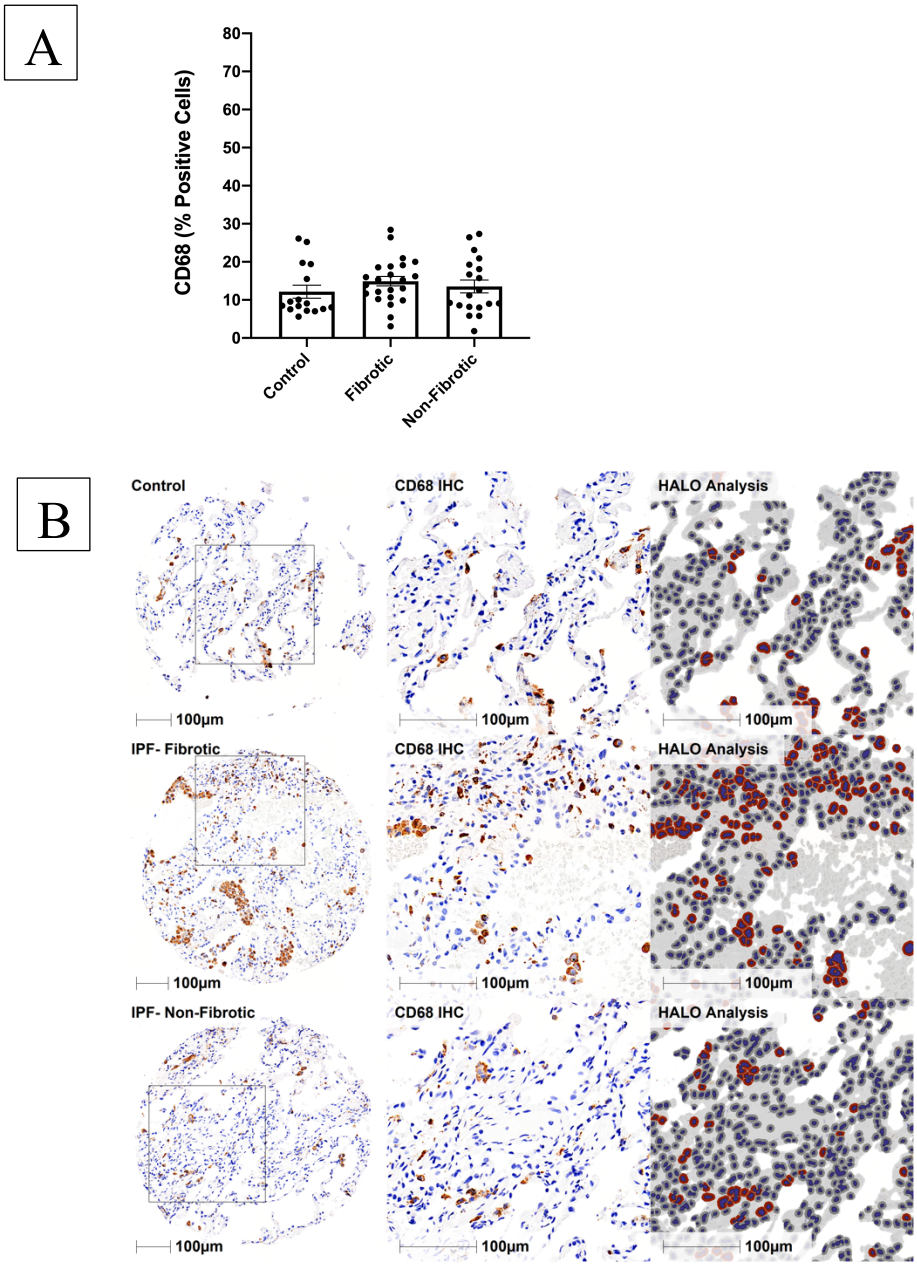


Figure 7: CD68 Immunohistochemical staining on human IPF lung TMA

A. There is no difference in CD68 percent cell positivity in IPF (fibrotic and non-fibrotic regions) compared to control tissues.
B. Representative fibrotic, non-fibrotic, and control tissue core images. HALO® analysis markup is shown, with nuclei in blue and positive CD68 staining in red.

3.1.2 Expression of M2 macrophage marker, CD206, is increased in lung tissue from IPF patients

In order to determine the expression of M2 macrophages in human IPF lung tissue, the IPF human lung tissue TMA was immunohistochemically stained for CD206. CD206 is a feature that characterizes M2 macrophages, and is found on their surface (Braga et al., 2015; Kazuo et al., 2019). A digitalized image of the slide was analyzed for percent positivity of CD206 cells using the Multiplex IHC module in the HALO® Image Analysis Platform software. This was done to confirm previous findings in the lab, utilizing novel and updated algorithms and modules in this software. As expected, CD206 expression was significantly higher in fibrotic regions of IPF lung tissue compared to control lung tissue ($p<0.0005$, Cohen's $d=1.229$). This was also found for non-fibrotic regions of IPF lung tissue compared to controls ($p<0.0005$, Cohen's $d=1.424$). No significant differences were seen between in CD206 expression between fibrotic and non-fibrotic regions **(Figure 8 A,B).**

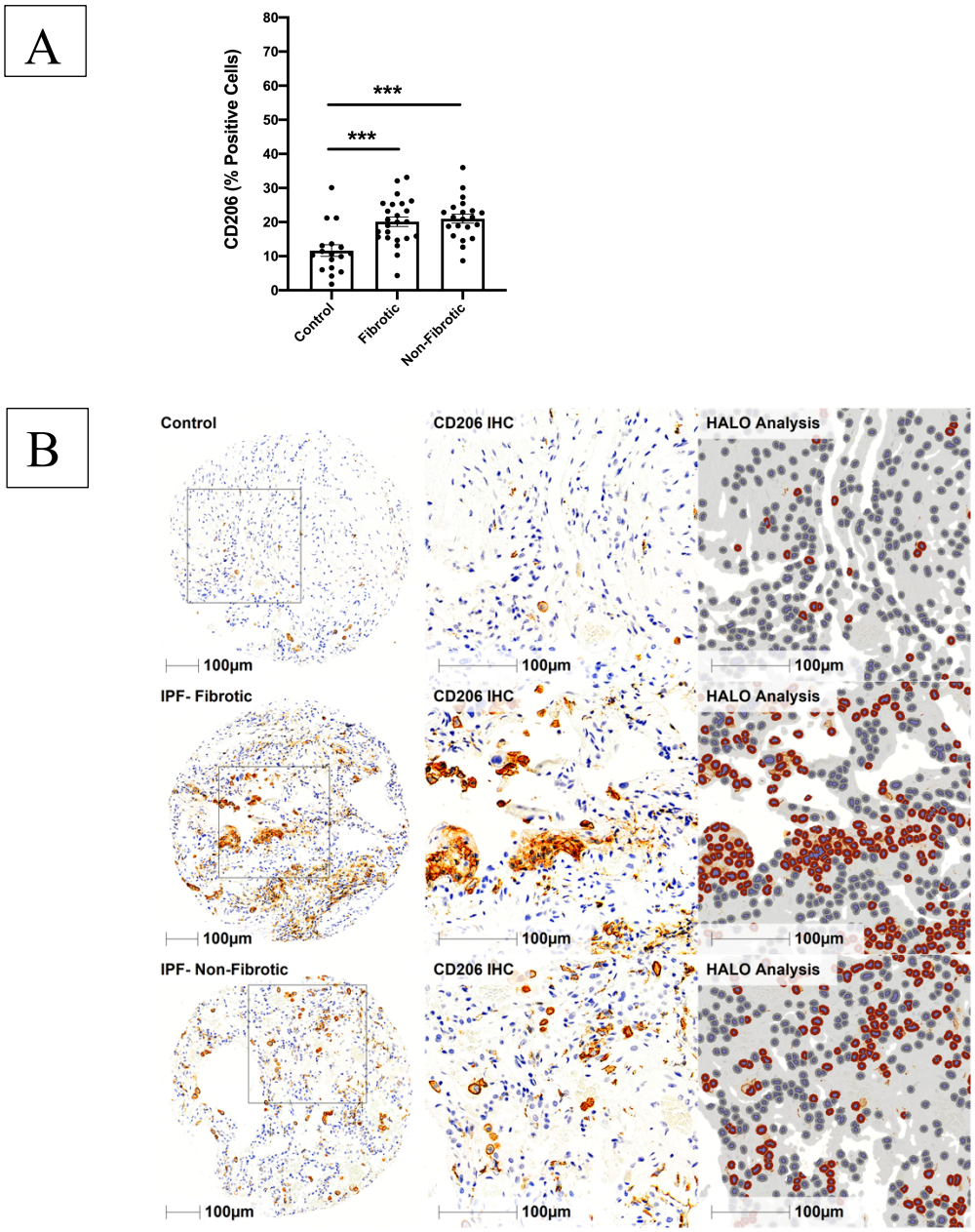


Figure 8: CD206 Immunohistochemical staining on human IPF lung TMA

A. CD206 percent cell positivity is significantly greater in IPF (fibrotic and non-fibrotic regions) compared to control tissues ($p<0.0005$, Cohen's $d=1.229, 1.424$).
B. Representative fibrotic, non-fibrotic, and control tissue core images. HALO® analysis markup is shown, with nuclei in blue and positive CD206 staining in red.

3.1.3 Expression of M2 macrophage marker, CD163, is increased in lung tissue from IPF patients

To further assess M2 macrophage expression, we also evaluated the expression of CD163 on the IPF human lung tissue TMA. CD163 is another M2 macrophage marker, which is a membrane-bound protein (Kristiansen et al., 2001; Fabriek et al., 2005). The TMA was immunohistochemically stained for CD163. A scanned image of the entire slide was analyzed for CD163 positivity using the Multiplex IHC module in the HALO® Image Analysis Platform software. As expected, CD163 expression was significantly higher in fibrotic regions of IPF lung tissue compared to control lung tissue ($p<0.05$, Cohen's $d=1.141$). This was also found for non-fibrotic regions of IPF lung tissue compared to controls ($p<0.05$, Cohen's $d=1.046$). No significant differences were seen between in CD163 expression between fibrotic and non-fibrotic regions **(Figure 9 A,B).**

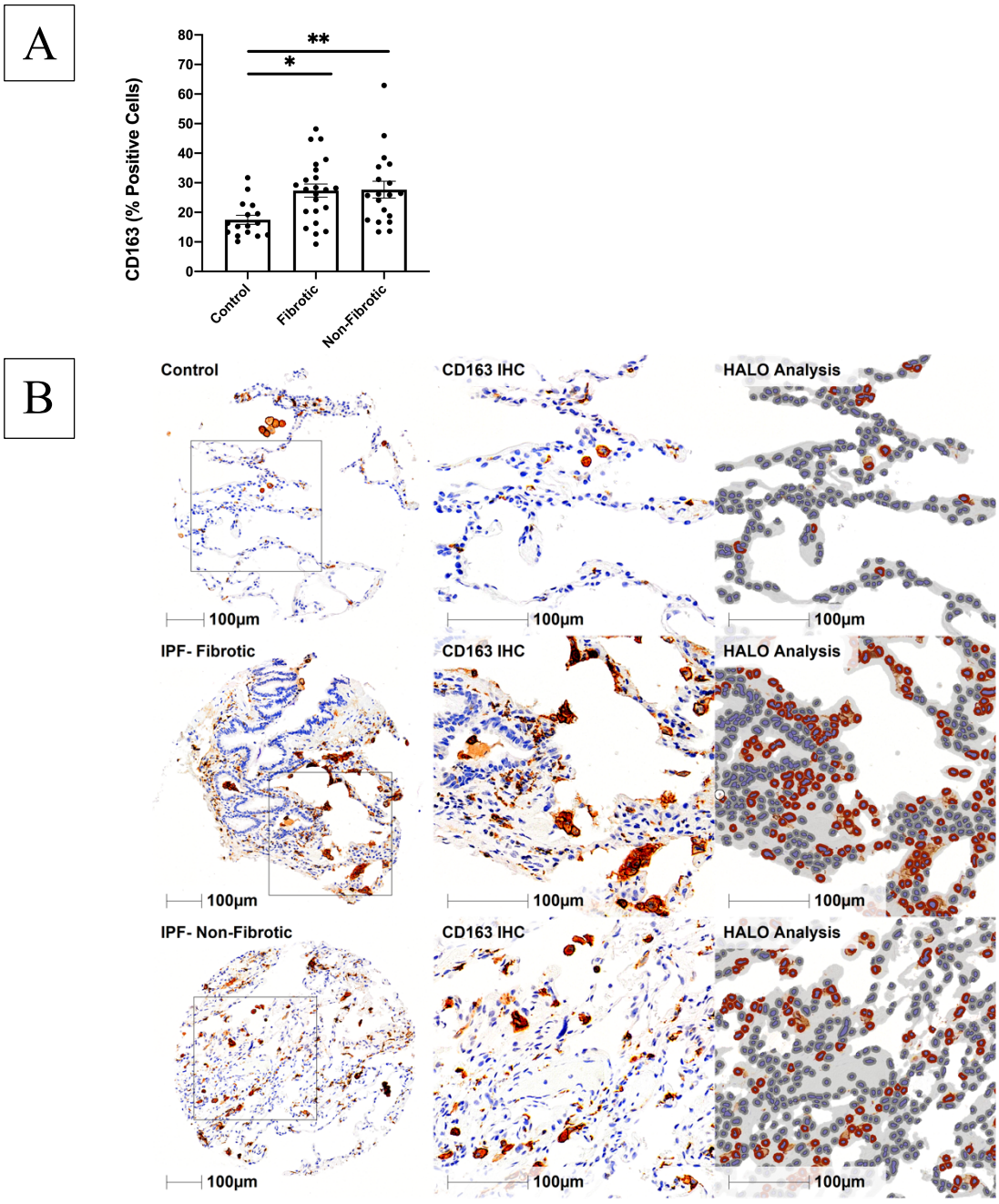


Figure 9: CD163 Immunohistochemical staining on human IPF lung TMA

A. CD163 percent cell positivity is significantly greater in IPF (fibrotic and non-fibrotic regions) compared to control tissues ($p<0.05$, Cohen's $d=1.141, 1.046$).
B. Representative fibrotic, non-fibrotic, and control tissue core images. HALO® analysis markup is shown, with nuclei in blue and positive CD163 staining in red.

3.1.4 CCL18 and CD68 RNA are colocalized in IPF lung tissue, confirming the ability of macrophages to produce M2 marker CCL18 in IPF

Dual RNAscope® fluorescent in-situ hybridization assay was used to stain CCL18 (Cy 5-channel 1) and CD68 (FITC-channel 2) on the human IPF TMA. The slide was digitalized under fluorescent light and the HALO® fluorescent in-situ hybridization module was used to detect colocalization of Cy5 and FITC stained puncta. CCL18 and CD68 puncta were colocalized to the same cells (**Figure 10**). This confirms the ability of macrophages, as detected by pan macrophage marker CD68, to produce CCL18 and thus be of the M2 phenotype in IPF, as exhibited at the RNA level. **Supplementary Figure A1** shows positive and negative probe staining.

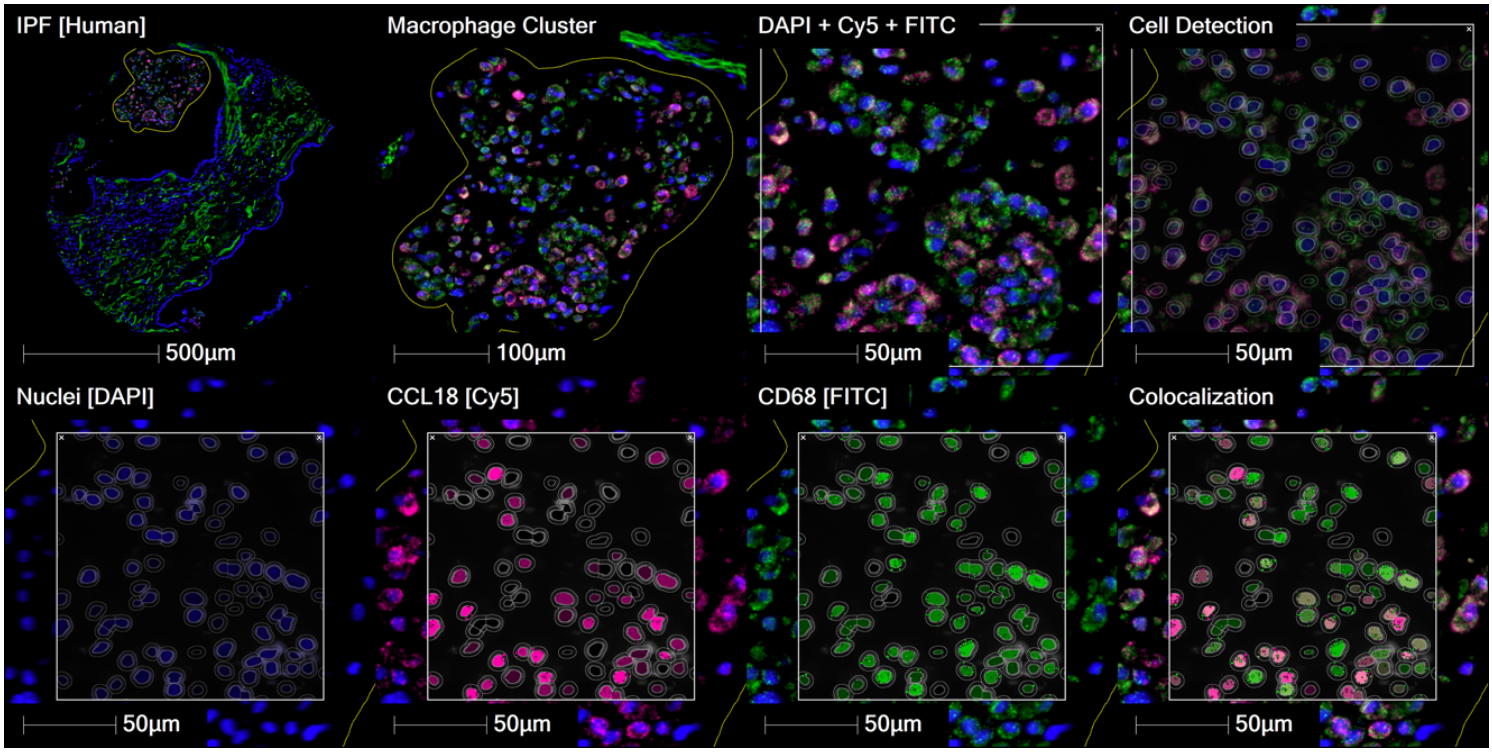


Figure 10: CCL18 and CD68 RNA are colocalized in human IPF lung tissue

RNAscope® fluorescent in-situ hybridization demonstrates that in an IPF TMA core, CCL18 (stained by Cy5, pink) is colocalized with CD68 (FITC, green). This colocalization, in addition to cell and nuclei (DAPI) detection, was detected by HALO® histology analysis software.

3.1.5 CCL18 mRNA levels are increased in IPF FFPE lung tissues compared to control tissues

Approximately 4 FFPE cores (1.0mm in diameter with a tissue depth of approximately 4.0mm) were punched from each patient (n=12) FFPE block and control (n=12) FFPE block for assessment. mRNA was extracted from samples and analyzed through Nanostring® technology for expression of CCL18. Data were preprocessed using the total counts normalization method, and two samples, one IPF and one control, were excluded as outliers. Data are expressed as Log2 values. Consistent with our previous findings, we confirmed that CCL18 gene expression was significantly upregulated in IPF FFPE lung tissues compared to control ($p<0.05$) (**Figure 11**), which supports the presence of M2 macrophages in IPF at the mRNA level. Overall, this is a suitable method to query RNA expression in FFPE tissues, and can be combined with colocalization data from in-situ hybridization (as seen in 3.1.4)

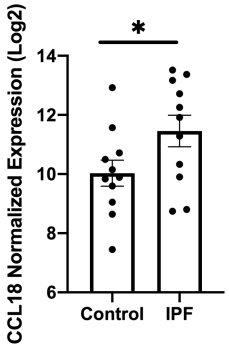
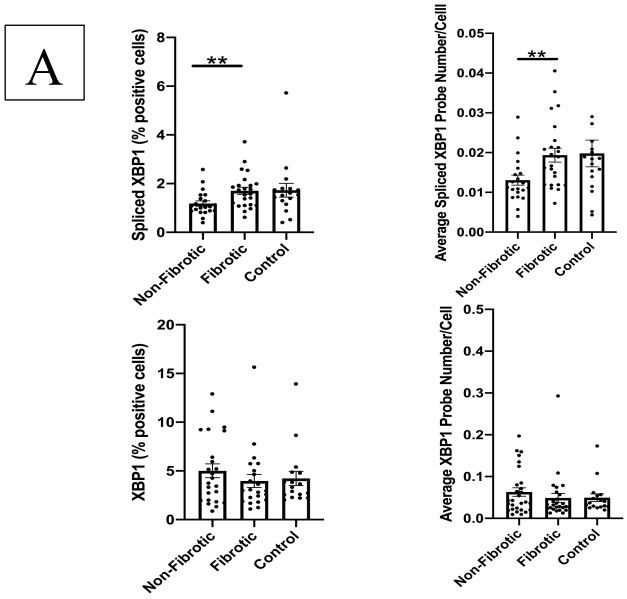


Figure 11: CCL18 gene expression is significantly increased in IPF FFPE lung tissues

Through Nanostring® technology, it was confirmed that CCL18 gene expression was significantly increased in IPF lung tissue (n=11) compared to control (n=11).

3.1.6 Spliced XBP1 is expressed in higher levels in fibrotic cores from IPF patients

In-situ hybridization using Basescope™ technology on the human IPF TMA and quantification with HALO® showed that sXBP1 is increased (both % positive cells and average probe number per cell) in fibrotic cores from IPF patients compared to non-fibrotic cores ($p<0.005$), suggesting increased UPR activation of the IRE1/XBP1 arm in areas of fibrosis. No significant difference was seen between IPF (both fibrotic and non-fibrotic) and control cores, suggesting that UPR activation may be present in cancerous lung, regardless of the region. No significant differences were seen for XBP1, as this form does not signify the activation of the IRE1/XBP1 pathway **(Figure 12 A,B).** Proof of principle of Basescope™ staining can be seen in **Supplementary Figure A2** (positive and negative probe staining) and **Supplementary Figure A3** (increase in sXBP1 staining in vitro with tunicamycin treatment).



B

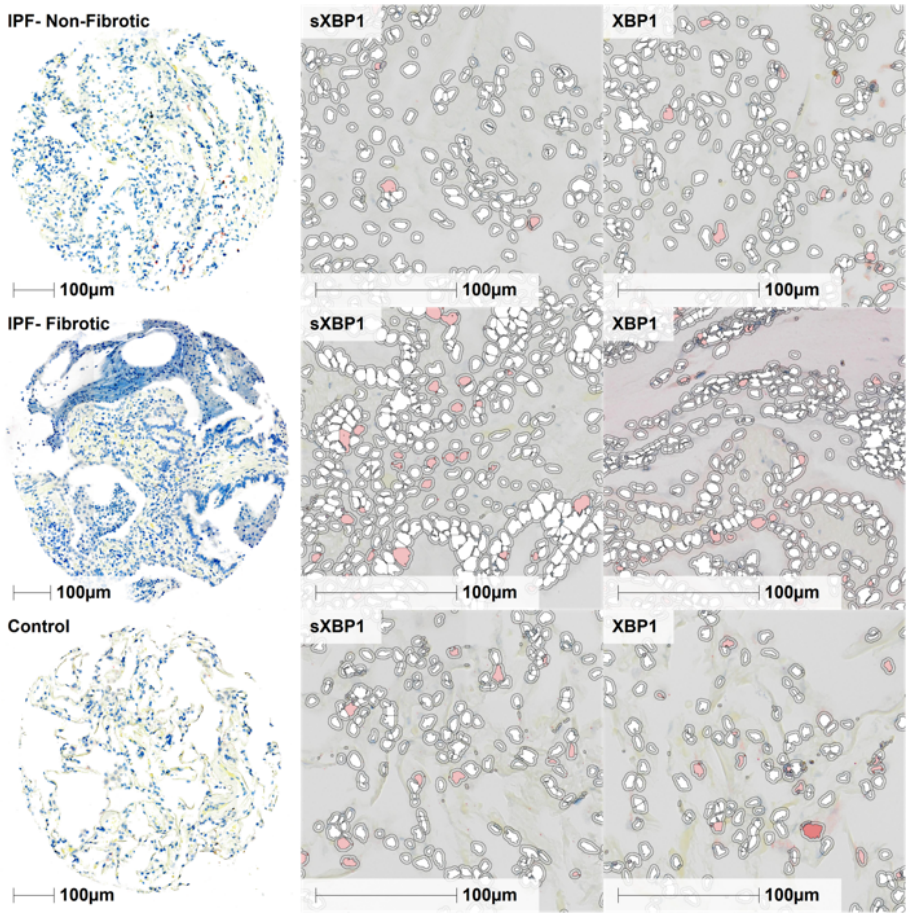


Figure 12: Basescope™ in-situ hybridization for staining of full-length and spliced XBP1 gene variants on human IPF tissue TMA

A. Quantification of percent positive cells and average probe number per cell of sXBP1 and XBP1 in fibrotic versus non-fibrotic cores, and control cores. sXBP1 expression is significantly higher in fibrotic cores ($p<0.005$) compared to non-fibrotic cores. No significant difference was found for full-length XBP1.
B. Representative images of fibrotic, non-fibrotic cores, and control stained for sXBP1 and XBP1, using the "real time tuning" analysis function in HALO® to demonstrate algorithm markup.

3.2 Objective 2: Collection of patient blood samples and generation of biobank

3.2.1 Collection of whole blood samples from patients and controls

Samples were collected from IPF patients and healthy controls at the Firestone Institute of Respiratory Health, in collaboration with respirologists (Dr. Hambly, Dr. Cox, and Dr. Kolb). Samples were collected from January 2018 to May 2019. Up to 6 samples were collected and processed per week. As shown in **Table 2**, samples were collected from 50 patients and 12 healthy controls. Serial blood samples were also collected at 6 month follow up timepoints from the same patients, where 12 patients had 2 samples collected and 2 patients had 3 samples collected. In addition to monocytes, other samples isolated and collected from the whole blood included PBMC and T-cells, and plasma.

Table 2: Sample Collection Information

Single Sample Collected	Two Samples Collected	Three Samples Collected
36	12	2
Total patients: 50		
Total controls (not included in table): 12		
Total samples: 78		

3.2.2 Demographic and Clinical Information

Overall, 50 patients and 12 controls were recruited from the Firestone Institute for Respiratory Health Clinic (as seen in Table 2). **Table 3** and **Table 4** include the demographic and clinical information of the patients and healthy controls from whom the blood samples were collected. For both patients and controls, the majority of participants were males. For IPF patients, the average age was 74.2, while for healthy controls it was 44.8. The majority of patients were ex-smokers, while the majority of healthy controls were never-smokers. In the IPF population, the majority of patients were taking Nintedanib as an antifibrotic medication. Overall, FVC and FEV1 values were lower patients compared to healthy controls, while monocyte quantities (obtained with cell counter after sample isolation) were higher.

Table 3: Demographic and Clinical Characteristics from IPF Patients

Characteristic		Value
Gender	Male	44
	Female	5
Age (years)		74.2 ± 1.0
Smoking Status	Never	9
	Ex	37
	Current	2
Antifibrotic Medication	Nintedanib	24
	Pirfenidone	17
	None	4
FVC (%)		72.0 ± 2.5
FEV1 (%)		80.7 ± 2.6
DLCO (%)		45.7 ± 2.6
Monocytes per mL		330,608 ± 27,354

***some information missing from 5 participants**
****values from most recent patient visit for those with multiple visits**

Table 4: Demographic and Clinical Information from Healthy Controls

Characteristic		Value
Gender	Male	8
	Female	2
Age (years)		44.8 ± 4.5
Smoking Status	Never	7
	Ex	2
	Current	0
FVC (%)		106.0 ± 5.0
FEV1 (%)		106.2 ± 6.0
Monocytes per mL		205,456 ± 29,811

***some information missing from 7 participants**

3.2.3 Assessing and understanding quality of samples in biobank

In order to assess the quality and suitability of our biobank samples for RNA sequencing, the average blood volume obtained from participants, RNA extracted from 100,000 monocytes and RIN were calculated. On average, RNA sequencing completed on cells requires a RIN of 7, and an input volume of ideally 100 ng of RNA. As shown in **Figure 13A**, we obtained an average of 45 mL of blood from healthy controls, and 25 mL of blood from IPF patients. 3mL of whole blood was used for monocyte isolation. As shown above in Table 2 and Table 3, we obtained an average of 330,608 monocytes per mL of blood from IPF patients, and 205,456 monocytes per mL of blood from healthy controls (when calculated using the most recent values we have per participant). Therefore, obtaining approximately 72 ng per 100,000 monocytes **(Figure 13B)** was sufficient to meet the input volume of 100 ng of RNA. Additionally, when determining the quality of RNA, we achieved an average RIN of 9.3 **(Figure 13C),** which exceeds the minimum required RIN of 7.

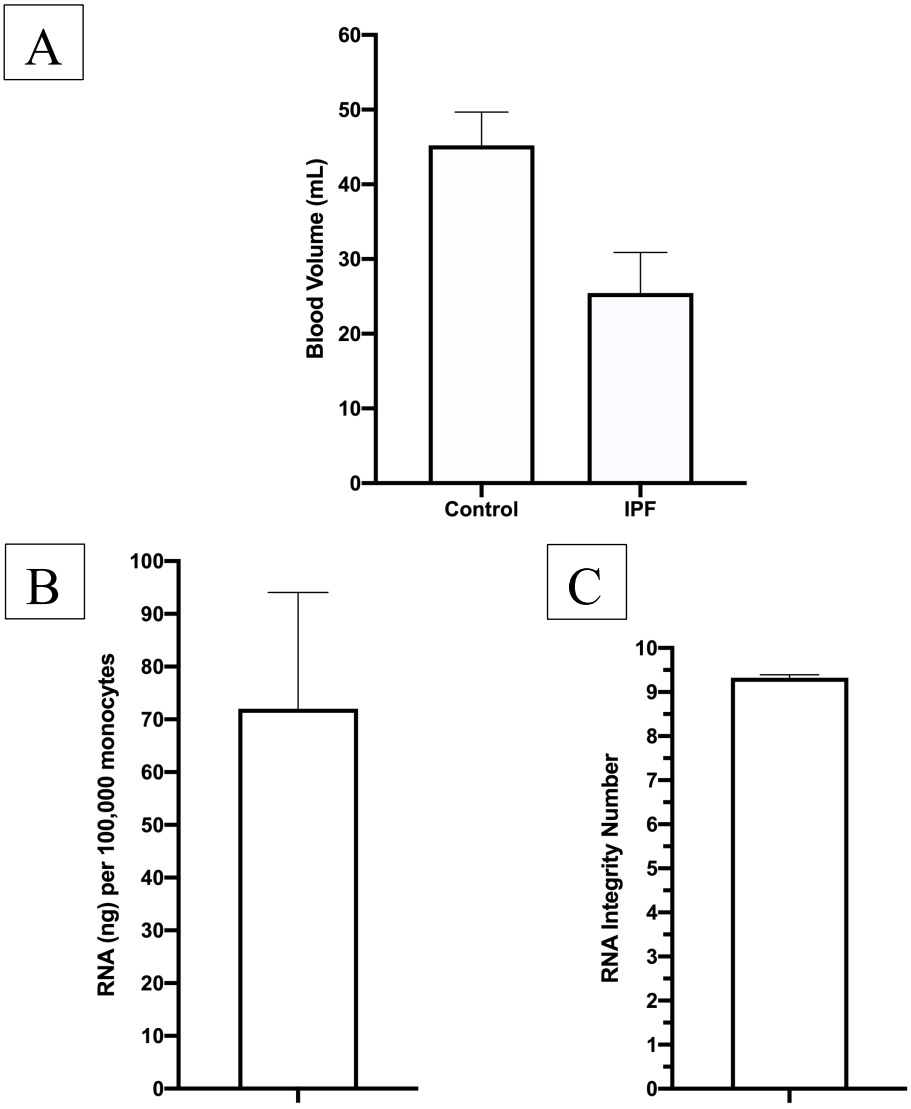


Figure 13: Assessment of sample quantities obtained from participants

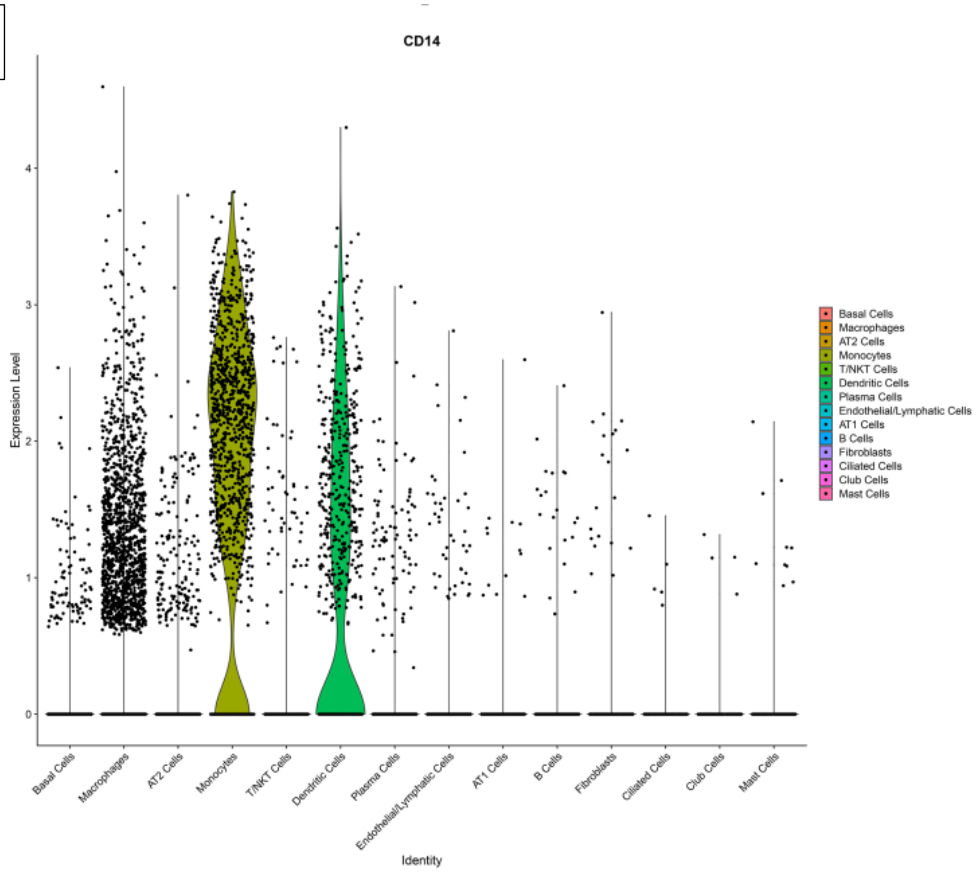
A. The average whole blood volume obtained when collecting from control and IPF subjects.
B. Overall average quantity of RNA extracted per 100,000 monocytes.
C. RIN from samples analyzed on the Agilent Technologies Bioanalyzer 2100.

3.3 Objective 3: Investigate transcriptomic characteristics of monocytes from IPF patients with fibrotic disease and correlate these characteristics with disease progression

3.3.1 Confirming the expression of CD14 in monocytes in IPF

In order to confirm that CD14 is indeed expressed on monocytes in IPF, analysis of human pulmonary fibrosis single cell RNA sequencing data from two GEO deposited datasets (GSE122960 and GSE135893) was performed. GSE122960 includes 14 different cell types in the lung, and GSE135893 includes 27. In both of these datasets, it was found that CD14 was most highly expressed on monocytes (**Figure 14 A,B and Figure 15 A,B**), confirming CD14 to be a marker for monocytes. As expected, CD14 expression was also present in dendritic cells and macrophages.

A



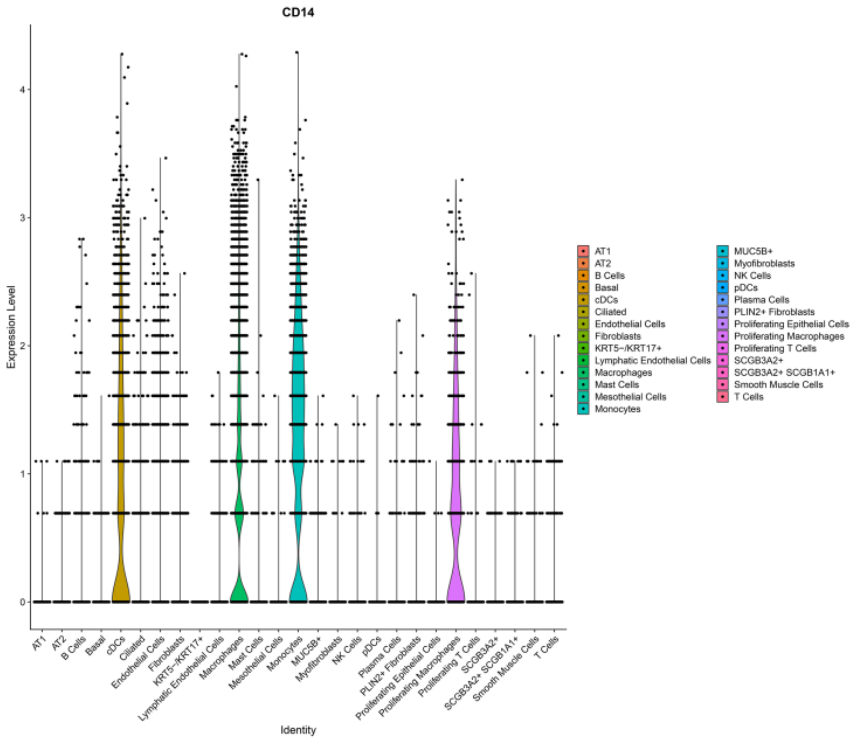
B

CD14 Expression		
Cell Population	**FC**	**P Value**
Monocytes	6.49	$< 10^{-6}$
Dendritic Cells	2.83	$< 10^{-6}$

Figure 14: Examining CD14 expression in GSE122960 pulmonary fibrosis single cell RNA sequencing dataset

A. Violin plot for CD14 gene expression level in various cell types.
B. CD14 is most highly expressed in the monocyte and dendritic cell populations.

A



B

CD14 Expression		
Cell Population	**FC**	**P value**
Monocytes	2.48	$< 10^{-6}$
Macrophages	2.01	$< 10^{-6}$
Conventional Dendritic Cells	1.92	$< 10^{-6}$
Proliferating Macrophages	1.39	$< 10^{-6}$

Figure 15: Examining CD14 expression in GSE135893 pulmonary fibrosis single cell RNA sequencing dataset

A. Violin plot for CD14 gene expression level in various cell types.
B. CD14 is most highly expressed in the monocyte, dendritic cell, and macrophage populations.

3.3.2 Monocyte quantity is increased in IPF patients, as compared to control subjects

The cell count obtained after isolation showed that CD14+ monocyte count is significantly increased in IPF patients compared to controls **(Figure 13**). This is consistent with previous results, which reported that increased monocyte count is associated with poor outcomes in fibrotic disease (Scott et al., 2019).

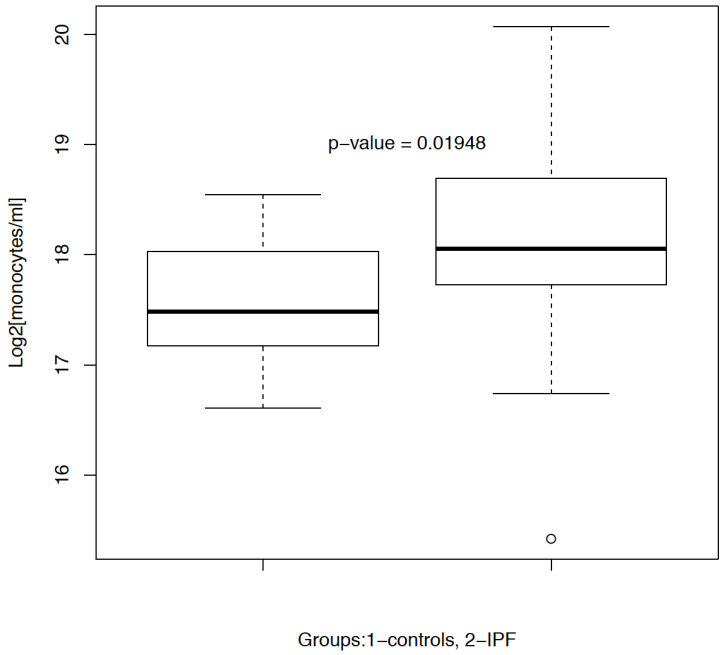


Figure 16: Monocyte number is increased in IPF subjects compared to healthy controls

Box plot (median shown) of log-transformed data displaying increased CD14+ monocyte count in IPF patients compared to controls ($p<0.05$).

3.3.3 Monocyte number is negatively correlated with percent FVC in this cohort

Monocyte cell count and lung function values, including FVC, FEV1, and DLCO, were analyzed via correlation. Matched lung function test values and monocyte counts measured during the same study visit were compared. Monocyte count was determined with an automated cell counter after isolation of from whole blood, while lung function measures were determined with spirometry. It was found that there was a significant negative correlation ($p<0.05$) between monocyte value and FVC in our cohort (**Figure 17A**). There was no significant correlation observed for monocyte quantity and FEV1, and monocyte quantity and DLCO (**Figure 17 B,C**).

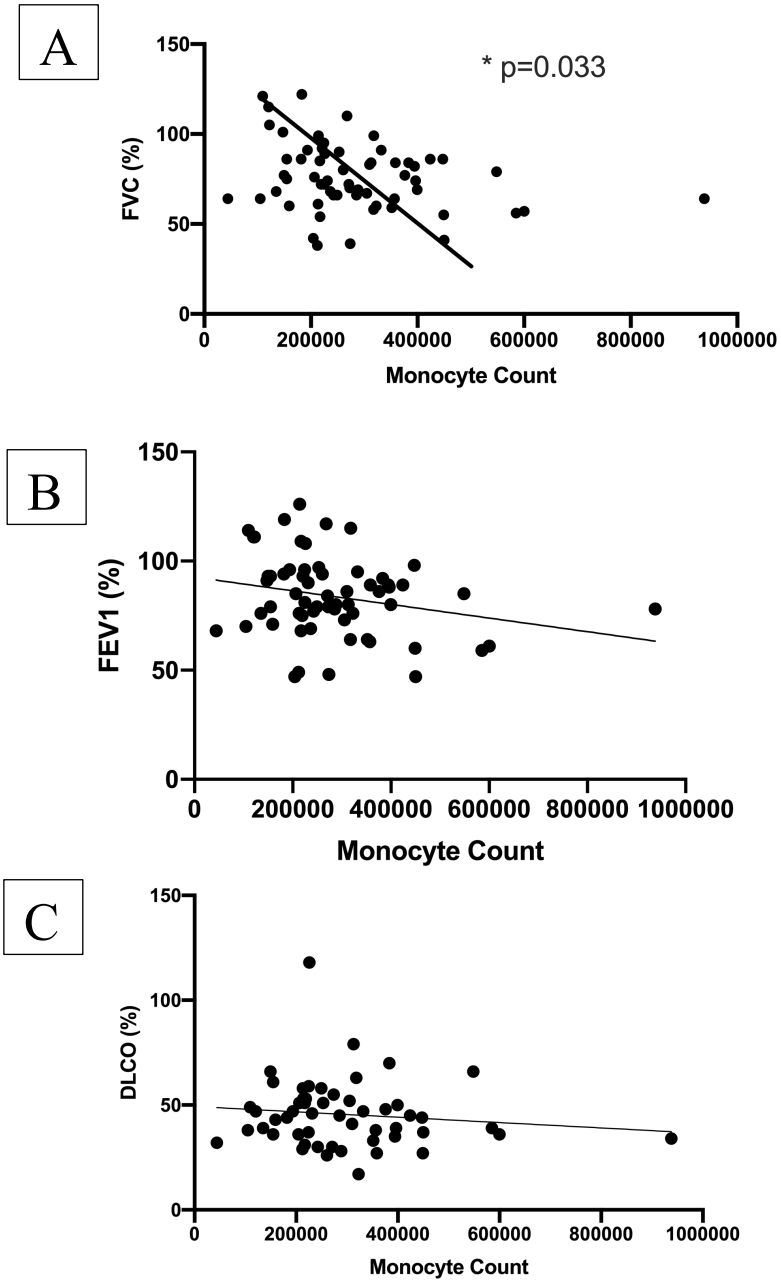


Figure 17: Correlation of monocyte count and lung function measures

A. A significant negative correlation ($p<0.05$) was found between monocyte count and FVC
B. No significant correlation was found between monocyte count and FEV1.
C. No significant correlation was found between monocyte count and DLCO.

3.3.4 There is a 35-gene signature associated with disease determined by differential expression between IPF and control subjects

Differential expression analysis yielded a 35-gene signature, where 16 genes were significantly upregulated in IPF patients (**Table 5**), and 19 genes were significantly downregulated (**Table 6**). All genes displayed are significantly differentially expressed ($p<0.05$). Initially, over 600 genes were seen to be significantly differentially expressed. However, after factoring in the difference in monocyte quantity between IPF and controls, as seen in 3.3.2, a 35-gene signature was determined.

Table 5: 16 upregulated genes in IPF compared to control

	FC	Gene Name
NIPAL2	4.852189	NIPA Like Domain Containing 2
C17orf80	2.573723	Chromosome 17 Open Reading Frame 80
PPP1R32	2.362287	Protein Phosphatase 1 Regulatory Subunit 32
ZNF57	2.247155	Zinc Finger Protein 57
DDX20	2.124149	DEAD-Box Helicase 20
SBF2	1.970128	SET Binding Factor 2
C7orf60	1.927868	Base methyltransferase of 25S rRNA 2 homolog
AGFG1	1.851571	ArfGAP With FG Repeats 1
TBC1D14	1.79721	TBC1 Domain Family Member 14
GOLGA5	1.762521	Golgin A5
TMEM62	1.659073	Transmembrane Protein 62
TBC1D2	1.657096	TBC1 Domain Family Member 2
VWA5A	1.582539	Von Willebrand Factor A Domain Containing 5A
FLVCR2	1.574252	FLVCR Heme Transporter 2
PXK	1.542189	PX Domain Containing Serine/Threonine Kinase Like
IL4R	1.450062	Interleukin 4 Receptor

Table 6: 19 downregulated genes in IPF compared to control

	FC	Gene Name
MSC	-14.3656	Musculin
GATA3	-6.23865	GATA Binding Protein 3
BEX2	-5.10632	Brain Expressed X-Linked 2
FXYD7	-4.24973	FXYD Domain Containing Ion Transport Regulator 7
NFKBIA	-2.97053	NFKB Inhibitor Alpha
MARCKS	-2.54177	Myristoylated Alanine Rich Protein Kinase C Substrate
C19orf48	-2.26709	Chromosome 19 Open Reading Frame 48
C17orf89	-2.21413	NADH:Ubiquinone Oxidoreductase Complex Assembly Factor 8
FAM195A	-2.07853	Family With Sequence Similarity 195, Member A
LTBP3	-1.79622	Latent Transforming Growth Factor Beta Binding Protein 3
POLD2	-1.68778	DNA Polymerase Delta 2, Accessory Subunit
MRPL54	-1.6632	Mitochondrial Ribosomal Protein L54
TRMT61A	-1.62339	TRNA Methyltransferase 61A
IMPDH2	-1.61829	Inosine Monophosphate Dehydrogenase 2
PPIB	-1.56429	Peptidylprolyl Isomerase B
RANBP1	-1.56153	RAN Binding Protein 1
FXN	-1.55647	Frataxin
HSP90B1	-1.48277	Heat Shock Protein 90 Beta Family Member 1
HSP90AB1	-1.46919	Heat Shock Protein 90 Alpha Family Class B Member 1

3.3.5 Principal component analysis of bulk RNA sequencing data on monocytes mRNA conveyed three distinct clusters of samples corresponding to sequential six month follow up visits

Overall, there are 14 patients in our dataset from which we have at least two serial samples collected from. In PCA, when all initial visits were placed at the origin of the plot, three distinct clusters formed, which may be indicative of different directions of progression in disease (**Figure 18**). It was confirmed that these clusters were not as a result of difference in monocyte quantity, RNA quantity, or person performing the isolation.

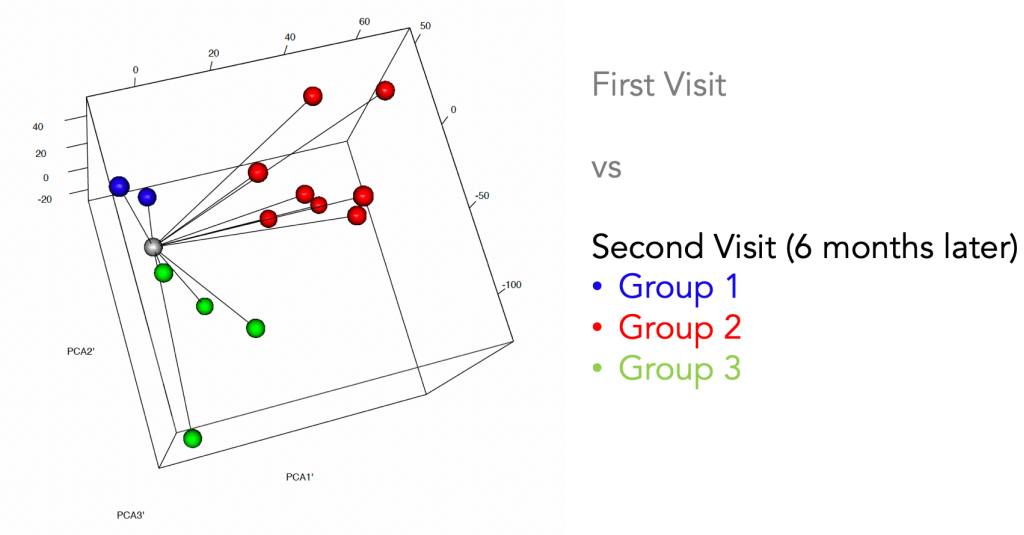


Figure 18: Principal component analysis plot of 14 samples from serial visits, compared to initial visit

All samples from initial visits are placed at the origin of the plot (grey). From there, samples from six months later are spread across the plot, which formed three distinct clusters (blue, red, green).

3.3.6 Part of the published outcome-predicting 52-gene signature is enriched in our dataset

The published outcome-predicting 52-gene signature from PBMC in IPF consisted of 7 genes associated with short TFS, and 45 genes associated with long TFS (Herazo-Maya et al., 2017). When comparing the signature related to long TFS from the published dataset to our data, GSEA demonstrated that this signature is enriched in donors from our dataset ($p<0.005$) **(Figure 19)**, supporting concurrence between the meanings of the two datasets.

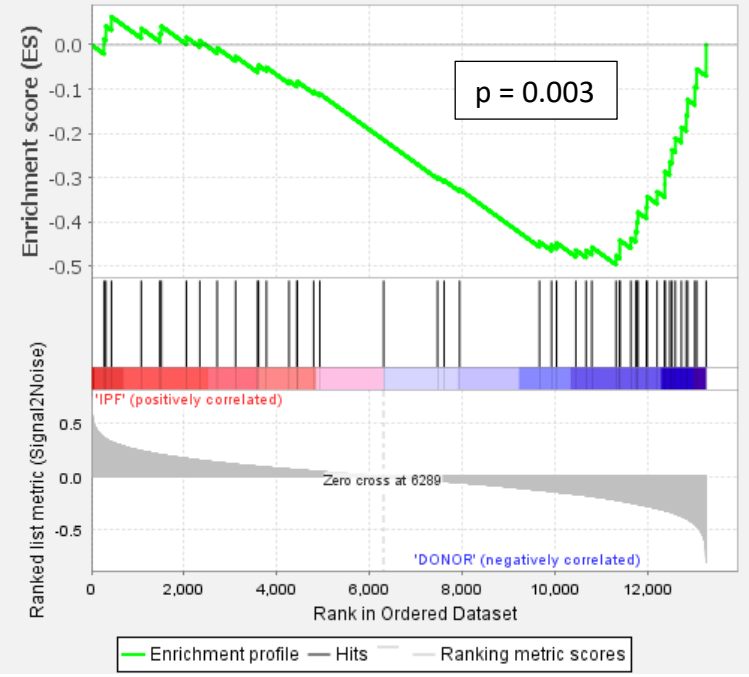


Figure 19: GSEA demonstrating enrichment of the published progression-related 52-gene signature in the dataset obtained from monocytes at the Firestone Institute for Respiratory Health

Through comparing the two datasets through GSEA, it was found that the published 52-gene signature was enriched in controls in our dataset.

CHAPTER 4: DISCUSSION

There is increasing evidence to suggest the role of macrophages in fibrosis, and these cells are considered to have critical involvement in fibrogenesis (Hou et al., 2018). One of such ways is by participating in crosstalk with fibroblasts, and leading them to differentiate into myofibroblasts through the release of TGF-β (Lodyga et al., 2019). These myofibroblasts then contribute to the deposition of extracellular matrix components (Fernandez & Eickelberg, 2012). Due to the involvement in these processes, macrophages constitute a probable target in IPF. Monocytes, as precursor cells for macrophages, are also a reasonable target of interest. These cells can be recruited from the circulation into the lung tissue, where they differentiate into macrophages and likely drive fibrogenesis (Zhang et al., 2018). However, very little is known about these monocytes in IPF, especially in the human system. Previous work has identified circulating fibrocytes as a potential clinical marker for disease progression, as they were seen to be a predictor of early mortality (Moeller et al., 2009). In systemic sclerosis interstitial lung disease, it has also been shown that circulating monocytes have a profibrotic phenotype (Mathai et al., 2010). This supports the investigation of circulating cells as contributors to the progression of IPF. Recently, a study has exhibited that the sheer quantity of monocytes in the blood of patients is predictive of poor outcomes in IPF (Scott et al., 2019). Another recent study found that increased monocyte numbers are a risk factor for acute exacerbation in fibrosing ILD (Kawamura et al., 2020). Despite these associations, the

characteristics of these cells have not been elucidated. Therefore, the examination and characterization of monocytes is critical to understand their role in the progression of IPF.

The first objective of this project was to assess the expression of macrophages in IPF human lung samples. Using TMA technology and a novel pipeline for target identification that our lab has recently optimized including characterization of tissues at the pathology level, histological staining, slide digitalization, and quantitative analysis, we examined the expression of macrophage markers CD68, CD206, and CD163 in human lung tissue. CD68, a pan macrophage marker, was not seen to be increased in IPF lung compared to control. M2 macrophage markers CD206 and CD163 were seen to have increased expression in IPF, both in fibrotic and non-fibrotic regions. The absence of increase of CD68 was unexpected, especially taken together with our results expressing an increase in monocytes in IPF. However, as CD68 is a pan macrophage marker, it cannot be distinguished as to which macrophages are of the M1 phenotype and which are of the M2 phenotype. As we currently do not have staining for the M1 phenotype, we cannot conclude the level of M1 expression in the lung samples. This may suggest that while the overall level of macrophage expression does not increase, more macrophages are polarized to the M2 phenotype. Macrophage polarization has been demonstrated to be a plastic and flexible process (Italiani & Boraschi, 2014). Therefore, CD68 expression would not necessarily be increased in IPF since it is non-specific for the profibrotic phenotype. This could explain why CD68 expression is consistent in IPF and control lung. Additionally, expression of profibrotic markers CD206 and CD163 were elevated in

IPF. CD206 and CD163 has previously been seen to be increased in alveolar macrophages in IPF (Pechkovsky et al., 2010; Vasarmidi et al., 2019). However, it is interesting that these markers are elevated in both fibrotic and non-fibrotic regions. This may be suggestive that M2 macrophages are not present in only regions of active fibrosis, but in regions that may soon become fibrotic as well.

To examine the expression of UPR pathway IRE1 activation in human lung tissue, we utilized Basescope™ in-situ hybridization. This technology has allowed us to identify sXBP1 in FFPE tissues for the first time. As sXBP1 only varies from XBP1 by 26 base pairs, there is not another existing technology, other than polymerase chain reaction (PCR), which allows the detection of this gene. PCR is also only limited to fresh cells, as there is considerable RNA degradation in FFPE tissues. However, Basescope™ probes are sensitive enough to target short regions of base pairs, and thus gave us the ability detect sXBP1, signifying the activation of the IRE1/XBP1 pathway, in lung tissues from human IPF patients. Previous work done in our lab has linked activation of this pathway and the presence of sXBP1 to polarization of macrophages to the M2 phenotype (Ayaub et al., 2019). Although we have detected it in human IPF tissue and found it to have higher expression in fibrotic versus non-fibrotic regions, we are limited in our ability to definitively draw the link to macrophages as this assay was single-plex. Future prospective experiments to confirm this include multiplex assays, where we can also detect and quantify the colocalization of sXBP1 and macrophage markers, to verify our lab's in vitro findings in human lung tissue.

In concurrence with the literature (Scott et al., 2019), our findings confirmed that monocyte number is increased in the blood of IPF patients compared to control. The study by Scott et al. (2019) demonstrated that patients with higher monocyte counts had an increased risk for shorter TFS, and suggests that monocyte count could possibly be incorporated into the clinical assessment of IPF patients. In our cohort, we also compared monocyte count to matched lung function values, which were obtained on the same day as the blood collection for each patient. Interestingly, it was found that monocyte count and FVC had a negative correlation. Not only is FVC an indicator of disease progression in IPF, but lowered FVC is also a predictor of mortality (Snyder et al., 2019). Therefore, this negative correlation further supports monocyte count to be related to poor outcomes and mortality in fibrotic lung disease, and warrants the investigation and characterization of these cells in IPF.

Through bulk RNA sequencing and transcriptomic analysis, we were able to determine a 35 gene signature for monocytes that contains 16 genes that are significantly upregulated and 19 genes that are significantly downregulated in IPF. To our knowledge, this is the first time that the transcriptomic characteristics of monocytes have been explored in IPF in humans on a moderate scale. This signature was determined through the analysis of 50 patient samples and 12 control samples. In order to confirm this signature and strengthen confidence in the findings, it is necessary to include more samples in the assessment. However, this study has established the framework and optimized protocols for further sample collection. To determine the minimum sample size

required to be confident in the signature, we performed a power calculation. In order to detect genes differentially expressed with a fold change of at least 2 in expression level between IPF patients and controls with a two-sided 0.05 significance level test with 90% power, we will need 27 samples in each group (IPF and controls). The cut off for the variation was taken as the 75th percentile of standard deviation obtained from our previous experiment. Therefore, we would require at least 15 additional control samples to confirm this signature. It would also be crucial to collect control samples from age-matched participants, as the current mean age for control participants (44.8) is drastically lower than for IPF participants (74.2). As there are known age-dependent changes in monocytes (Seidler et al., 2010) it would be imperative to obtain consistency in this factor to minimize influence on results.

Through GSEA, we found that the published 52 gene signature in PBMC (Herazo-Maya et al., 2017) to be enriched in our dataset. Specifically, we found that the 45 genes in this signature that are associated with long TFS were enriched in the controls in our cohort. This is as expected, since long TFS is associated with better outcomes. Overall, these findings support that there is agreeance between the meanings of the two datasets. However, only four genes that were differentially expressed in this dataset were also differentially expressed in ours. This may be because it was determined in PBMC rather than monocytes. In order to validate results and strengthen findings, it is important to obtain validation cohorts that also look at the transcriptomic characteristics of CD14+ monocytes. To solidify these signatures, we have established validation cohorts with the

Guangzhou Institute of Respiratory Health in Guangzhou, China, as well as the Women and Brigham's Hospital in Boston. We hope to sequence approximately 100 CD14+ monocyte samples from each of these sites, which will play a role in substantiating the signature that we have determined.

Although it is overall ineffective to consider the individual genes in the signature, one of the upregulated genes that stands out is IL4R. This gene encodes the receptor for IL-4, which is a cytokine that plays a significant role in the polarization of macrophages to the M2, or profibrotic, phenotype. In human IPF, IL-4 levels have been seen to be increased in bronchoalveolar lavage fluid (Park et al., 2009). It has also been discovered that IL-4 is a powerful inducer towards the M2 macrophage phenotype in monocytes from IPF patients (Pechkovsky et al., 2010). Our lab also uses this cytokine in in vitro experiments to skew macrophages to the M2 phenotype. Increased expression of IL4R could suggest an increased sensitivity of monocytes to IL-4, and thus a greater propensity to become M2 macrophages. If more of the recruited monocytes are differentiating into M2 macrophages in the lung, this may provide partial insight as to why there may be increased levels of these cells in IPF. This may also suggest potential treatment strategies with an anti-IL4R antibody.

Understanding and controlling acute exacerbations in IPF presents a long-standing challenge. (Juarez et al., 2015). As these events may result in death for patients, it is crucial to gain a better understanding of the risk factors and pathogenesis for exacerbations (Juarez et al., 2015). As absolute quantity of monocytes has been shown to

be a risk factor for acute exacerbation (Kawamura et al., 2020), further studies to examine the characteristics of these cells in this setting may be beneficial. Similar methodologies of characterization and transcriptomic analysis as used here can also be used to uncover differences in monocytes at the time of acute exacerbation, compared to when the patient is progressing normally. If transcriptomic changes are indeed occurring in the cells, this could provide insight into potential treatment options for the management of acute exacerbations, and what can potentially be done to lessen the severity of these events. However, logistical challenges exist with this as it may be difficult to obtain a blood sample from patients at their regular treatment institution, due to urgency for treatment during acute exacerbation.

Future Directions

Currently, our lab is working on the collection and characterization of additional FFPE lung biopsies for the creation of further TMAs with additional IPF cases. With the creation of these tools, along with the current human TMA currently being used, it would be beneficial to study the overall lung composition of different macrophages in human IPF tissue. As macrophages are the most abundant immune cells in the lung (Byrne et al., 2015), it would be interesting to examine how many of these macrophages are of the profibrotic phenotype in human IPF lung. With our established platform, this would be fairly feasible to execute. Serial sections of the TMA would be cut and stained with CD68, CD206, CD163, and binding immunoglobulin protein and H&E. The overall

quantity of macrophages for each subject on the TMA could be determined through quantification of CD68. To confirm accuracy, the H&E slide could also be analyzed alongside a pathologist to determine the quantity of macrophages in some samples, and then compared with the CD68 quantification. Quantification of CD206 and CD163 would then be performed, and the number of profibrotic macrophages compared to the total number of macrophages could be calculated.

Since our overall hypothesis is that monocytes leave the circulation, enter the lung tissue, and differentiate into profibrotic macrophages, it would be critical to also study the 35 gene signature in the lung compartment. To do this, in-situ hybridization using RNAscope® technology could be performed. From this, we could quantify the expression of these genes in IPF compared to control tissues, and also examine colocalization of different genes using multiplex technology. Using a macrophage marker, such as CD68 or Mannose receptor C-type 1, which is the gene that encodes CD206, we could also study colocalization of genes of interest in the macrophage department, as we have previously done with CCL18 and CD68. This would suggest that the monocytes expressing this gene signature are entering the lung in IPF, and may provide insight into their mechanism in this disease. Additionally, this may also shed light on potential ways to target these macrophages in the lung tissue, based on their gene expression profile.

We discovered the formation of three distinct clusters when comparing patient samples from the first visit and samples collected six months later through PCA, which we believe may be related to disease progression. We have confirmed that these

differences are not due to procedural or experimental variations, such as the person performing the isolation, the quantity of monocytes isolated, or the quantity of RNA obtained. Since we have collected various demographic and clinical characteristics from patients, such as smoking status, antifibrotic medication, and lung function scores, future analysis would include comparing the differences among these three clusters, and if any of the changes can be attributed to these characteristics. IPF is also a heterogenous disease, and so it would also be crucial to compare differences seen here with the clinical course of patients. This would provide meaning for the changes occurring at the transcriptomic level in IPF, and which genes may be related to disease progression. In order to concretely observe changes, more samples would need to be studied, since we presently have only studied 14.

It is also crucial to remember the importance of longitudinal sampling when studying human disease. In collaboration with respirologists at the Firestone Institute for Respiratory Health, we have developed a system for collecting serial samples from the same patients at six month intervals. This is critical to understanding the changes that are occurring, and the overall progression of the disease. In our biobank, we have now collected up to five serial samples from the same patients. This would allow us to study the progression of the patient over a 2.5 year timespan, which may provide significant insight into the transcriptomic changes that are occurring as the disease progresses. Cross-sectional sampling would be less effective in this setting, as it only provides a brief snapshot of this complex disease. We are continuing with longitudinal collection from the

same patients, and hope that as numbers and power are increased, we can gain a better understanding of the changes occurring throughout disease progression.

Although CD14 is a characteristic marker of monocytes in the blood, other cells may also express CD14, such as B-cells and neutrophils (Naeim et al., 2018). This is a current limitation to our study, since we cannot definitively conclude that all cells in the isolated population are monocytes. In order to overcome this limitation, it would be beneficial to store an aliquot of the isolated cells from each patient for flow cytometry, in order to verify purity. Other monocyte markers would need to be incorporated into the assessment to confirm that the cell population is truly monocytes. CD64 is a more sensitive monocyte marker than CD14. This marker is also absent in B-cells. The CD64/CD14 combination is also commonly used in flow cytometry when studying leukemia to identify cells with myelomonocytic differentiation (Naeim et al., 2018). Therefore, incorporating this into future flow cytometry studies would be beneficial in verifying the purity of the isolated cell population.

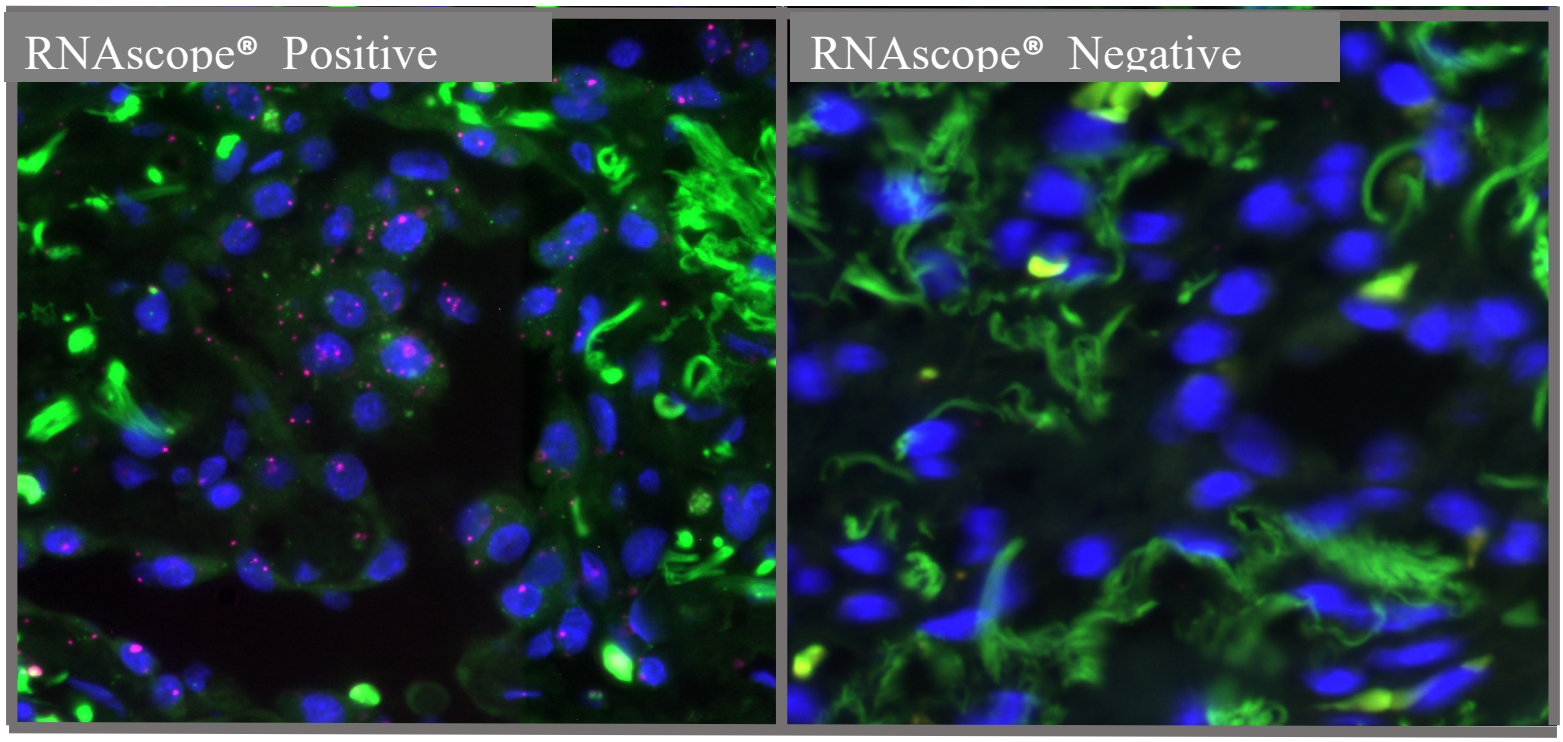
CHAPTER 5: CONCLUSION

Overall, the current study has demonstrated a novel approach for target identification in IPF in the human setting. As our overall hypothesis is that circulating monocytes leave the blood and enter the lung tissue where they differentiate into alternatively activated macrophages and contribute to the scarring process, we aimed to characterize both macrophages and monocytes in human IPF. Through an optimized pipeline for molecular phenotyping and imaging, it was found that profibrotic macrophages are increased in IPF human lung tissue, in both fibrotic and non-fibrotic regions. This was also confirmed at the gene expression level through Nanostring® technology, where M2 marker CCL18 RNA levels were increased in IPF compared to control. These results helped us to confirm M2 macrophages to constitute a valid target in IPF, and provided rationale for us to further study monocytes and build a biobank.

In order to explore the transcriptomic characteristics of monocytes in IPF, we collected 78 whole blood samples in total. These were from 50 patients (some with follow up visits, as seen in Table 4) and 12 controls. After optimizing protocols and performing quality assurance, through bulk RNA sequencing we determined a 35 gene signature associated with monocytes in IPF. In our cohort we also found a negative correlation between monocyte count in the blood and FVC, which supports that these cells may be associated with disease progression. Overall, these findings confirm our hypothesis that these cells express different characteristics in IPF, and provides insight on the function of these cells in this diseases. These differences also support monocytes as a potential

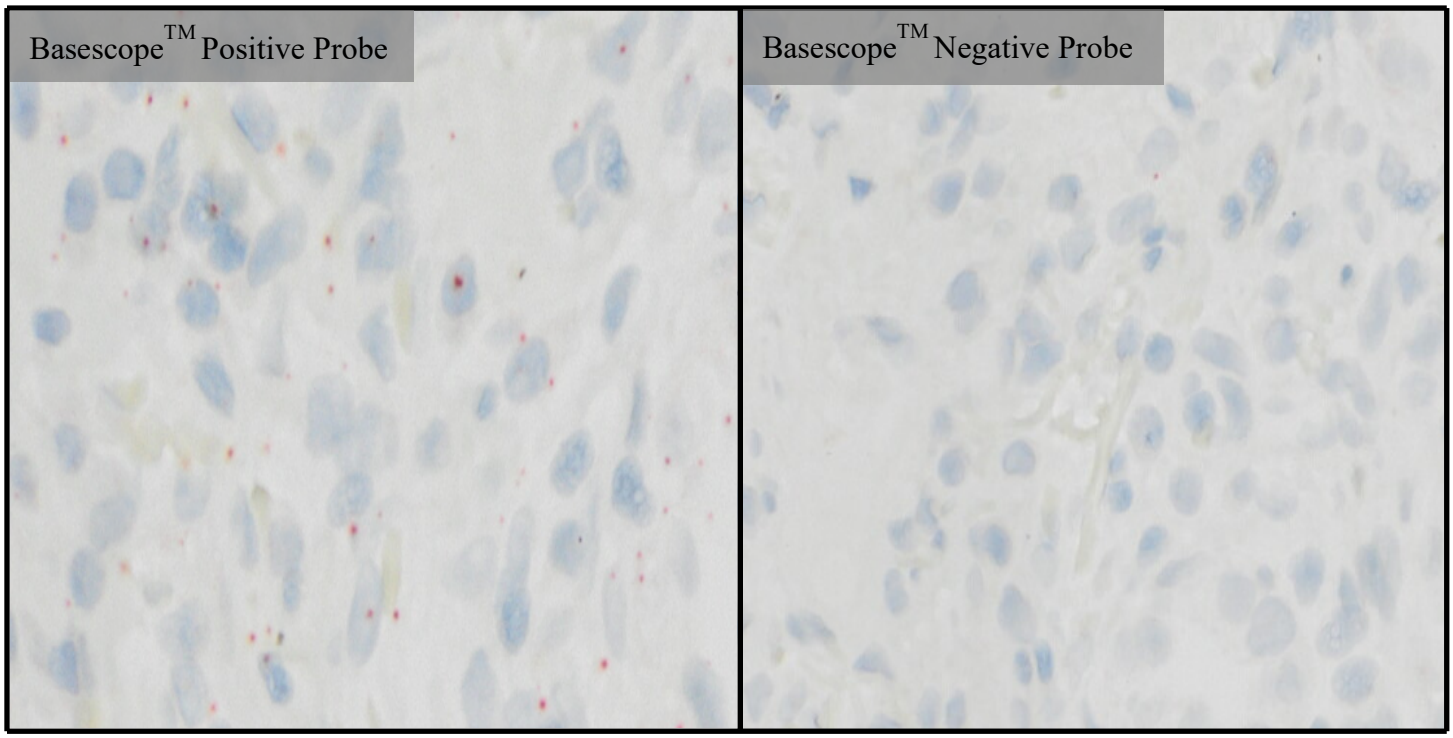
therapeutic target in IPF, and may allow for future identification of methods to target these cells to slow or stop disease progression.

CHAPTER 6: SUPPLEMENTARY FIGURES



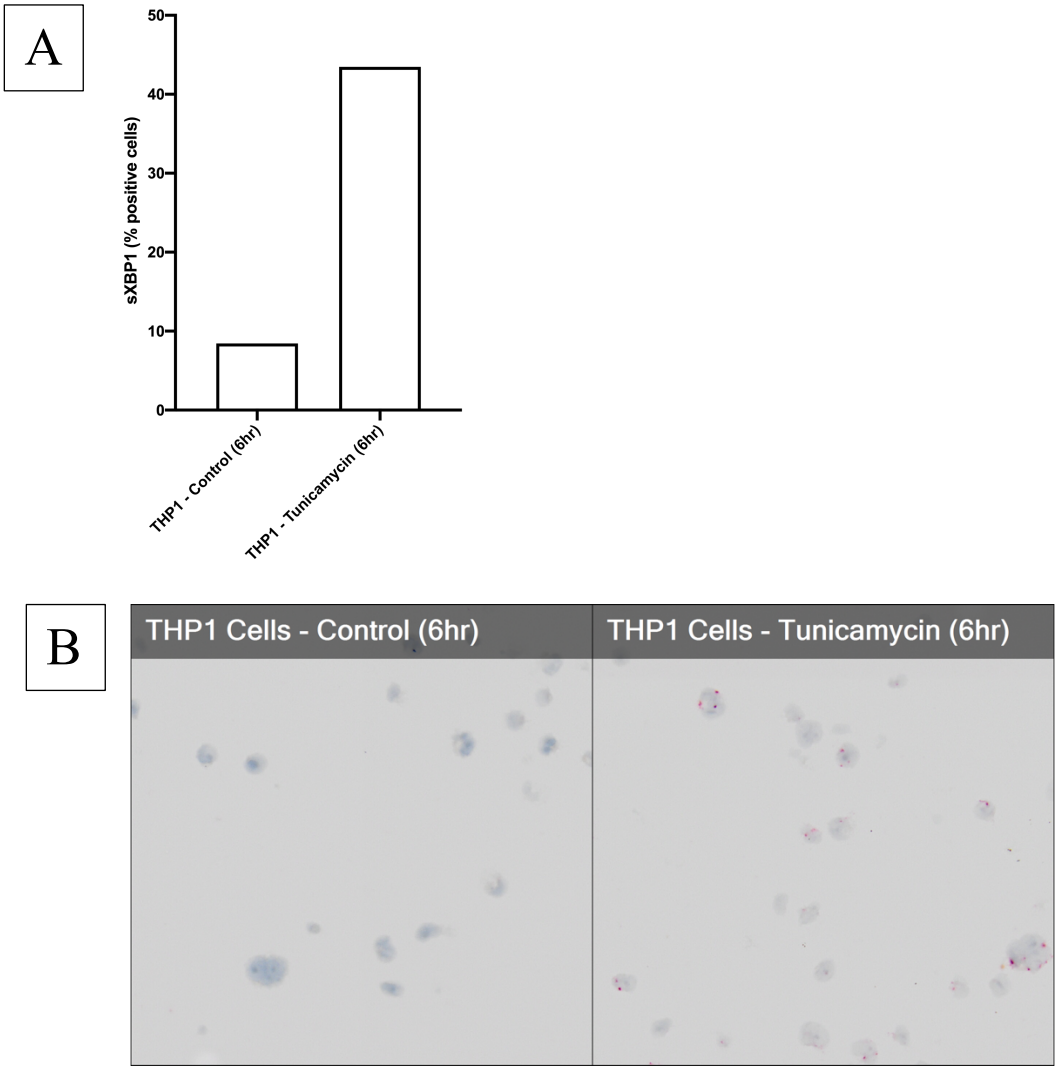
Supplementary Figure A1: RNAscope® Positive and Negative Probe Stains on Human Lung Tissue

Human IPF FFPE lung tissue stained with RNAscope® technology. The positive probe is a house keeping gene found in human cells and the negative probe is a bacterial gene.



Supplementary Figure A2: Basescope™ Positive and Negative Probe Stains on Human Lung Tissue

Human IPF FFPE lung tissue stained with Basescope™ technology. The positive probe is a house keeping gene found in human cells and the negative probe is a bacterial gene.



Supplementary Figure A3: Proof of Principle sXBP1 Staining in THP1 cells

THP1 cells were cultured in vitro. After a 6 hour incubation (control and tunicamycin treated), cells were lifted and fixed in formalin for 24 hours. Cells were then embedded in histogel, embedded in paraffin wax, and stained histologically for sXBP1 using Basescope™ technology.

A. sXBP1 levels are increased in THP1 cells treated with tunicamycin, a known inducer of ER stress.
B. 40X images showing increase in red puncta, which signify sXBP1 mRNA transcripts.

CHAPTER 7: REFERENCES

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)*, *31*(2), 166–169.

Ayaub, E. A., Tandon, K., Padwal, M., Imani, J., Patel, H., Dubey, A., Mekhael, O., Upagupta, C., Ayoub, A., Dvorkin-Gheva, A., Murphy, J., Kolb, P. S., Lhotak, S., Dickhout, J. G., Austin, R. C., Kolb, M. R. J., Richards, C. D., & Ask, K. (2019). IL-6 mediates ER expansion during hyperpolarization of alternatively activated macrophages. *Immunology & Cell Biology*, *97*(2), 203–217.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological), 57*(1), 289-300.

Boon, K., Bailey, N.W., Yang, J., Steel, M.P., Groshong, S., Kervitsky, D., Brown, K.K., Schwarz, M.I., & Schwartz, D.A. (2009). Molecular Phenotypes Distinguish Patients with Relatively Stable from Progressive Idiopathic Pulmonary Fibrosis (IPF). *PLoS ONE 4*, e5134.

Braga, T.T., Agudelo, J.S.H., and Camara, N.O.S. (2015). Macrophages During the Fibrotic Process: M2 as Friend and Foe. *Frontiers in Immunology 6*, 602.

Burman, A., Tanjore, H., and Blackwell, T.S. (2018). Endoplasmic reticulum stress in pulmonary fibrosis. *Matrix Biology 68–69*, 355–365.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature biotechnology, 36*(5), 411–420.

Byrne, A. J., Mathie, S. A., Gregory, L. G., & Lloyd, C. M. (2015). Pulmonary macrophages: key players in the innate defence of the airways. *Thorax, 70*(12), 1189-1196.

Chistiakov, D. A., Killingsworth, M. C., Myasoedova, V. A., Orekhov, A. N., & Bobryshev, Y. V. (2017). CD68/macrosialin: not just a histochemical marker. *Laboratory Investigation, 97*(1), 4-13.

Collin, M., McGovern, N., & Haniffa, M. (2013). Human dendritic cell subsets. *Immunology, 140*(1), 22–30.

Dickhout, J.G., Lhoták, Š., Hilditch, B.A., Basseri, S., Colgan, S.M., Lynn, E.G., Carlisle, R.E., Zhou, J., Sood, S.K., Ingram, A.J., et al. (2010). Induction of the unfolded protein response after monocyte to macrophage differentiation augments cell survival in early atherosclerotic lesions. *The FASEB Journal 25*, 576–589.

Fabriek, B. O., Dijkstra, C. D., & van den Berg, T. K. (2005). The macrophage scavenger receptor CD163. *Immunobiology, 210*(2-4), 153-160.

Fernandez, I., & Eickelberg, O. (2012). The impact of TGF-β on lung fibrosis: From targeting to biomarkers. *Proceedings of the American Thoracic Society, 9*, 111–116.

Flaherty, K. R., Fell, C. D., Huggins, J. T., Nunes, H., Sussman, R., Valenzuela, C., Petzinger, U., Stauffer, J. L., Gilberg, F., Bengus, M., & Wijsenbeek, M. (2018). Safety of nintedanib added to pirfenidone treatment for idiopathic pulmonary fibrosis. *The European respiratory journal*, *52*(2), 1800230.

Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association, 97*(458), 611-631.

Gibbons, M. A., MacKinnon, A. C., Ramachandran, P., Dhaliwal, K., Duffin, R., Phythian-Adams, A. T., van Rooijen, N., Haslett, C., Howie, S. E., Simpson, A. J., Hirani, N., Gauldie, J., Iredale, J. P., Sethi, T., & Forbes, S. J. (2011). Ly6Chi Monocytes Direct Alternatively Activated Profibrotic Macrophage Regulation of Lung Fibrosis. *American Journal of Respiratory and Critical Care Medicine*, *184*(5), 569–581.

Gogali, A., & Wells, A. U. (2012). Diagnostic approach to interstitial lung disease. *Current Respiratory Care Reports, 1*(4), 199-207.

Hall, C. L., Wells, A. R., & Leung, K. P. (2018). Pirfenidone reduces profibrotic responses in human dermal myofibroblasts, in vitro. *Laboratory Investigation, 98*(5), 640-655.

Heindryckx, F., Binet, F., Ponticos, M., Rombouts, K., Lau, J., Kreuger, J., & Gerwins, P. (2016). Endoplasmic reticulum stress enhances fibrosis through IRE1α-mediated

degradation of miR-150 and XBP-1 splicing. *EMBO molecular medicine, 8*(7), 729–744.

Herazo-Maya, J. D., Sun, J., Molyneaux, P. L., Li, Q., Villalba, J. A., Tzouvelekis, A., Lynn, H., Juan-Guardela, B. M., Risquez, C., Osorio, J. C., Yan, X., Michel, G., Aurelien, N., Lindell, K. O., Klesen, M. J., Moffatt, M. F., Cookson, W. O., Zhang, Y., Garcia, J., Noth, I., … Kaminski, N. (2017). Validation of a 52-gene risk profile for outcome prediction in patients with idiopathic pulmonary fibrosis: an international, multicentre, cohort study. *The Lancet. Respiratory medicine*, *5*(11), 857–868.

Hetz, C., Martinon, F., Rodriguez, D., & Glimcher, L. H. (2011). The Unfolded Protein Response: Integrating Stress Signals Through the Stress Sensor IRE1α. *Physiological Reviews*, *91*(4), 1219–1243.

Hopkins, R. B., Burke, N., Fell, C., Dion, G., & Kolb, M. (2016). Epidemiology and survival of idiopathic pulmonary fibrosis from national data in Canada. *European Respiratory Journal, 48*(1), 187-195.

Hou, J., Shi, J., Chen, L., Lv, Z., Chen, X., Cao, H., Xiang, Z., & Han, X. (2018). M2 macrophages promote myofibroblast differentiation of LR-MSCs and are associated with pulmonary fibrogenesis. *Cell Communication and Signaling*, *16*(1), 89.

Hughes, J. M. B., & Pride, N. B. (2012). Examination of the carbon monoxide diffusing capacity (DLCO) in relation to its KCO and VA components. *American journal of respiratory and critical care medicine*, *186*(2), 132-139.

Italiani, P., & Boraschi, D. (2014). From monocytes to M1/M2 macrophages: phenotypical vs. functional differentiation. *Frontiers in immunology, 5,* 514.

Johnson, D. C. (2000). Importance of adjusting carbon monoxide diffusing capacity (DLCO) and carbon monoxide transfer coefficient (K CO) for alveolar volume. *Respiratory medicine, 94*(1), 28-37.

Joshi, N., Watanabe, S., Verma, R., Jablonski, R. P., Chen, C. I., Cheresh, P., Markov, N. S., Reyfman, P. A., McQuattie-Pimentel, A. C., Sichizya, L., Lu, Z., Piseaux-Aillon, R., Kirchenbuechler, D., Flozak, A. S., Gottardi, C. J., Cuda, C. M., Perlman, H., Jain, M., Kamp, D. W., Budinger, G., … Misharin, A. V. (2020). A spatially restricted fibrotic niche in pulmonary fibrosis is sustained by M-CSF/M-CSFR signalling in monocyte-derived alveolar macrophages. *The European respiratory journal*, *55*(1), 1900646

Juarez, M. M., Chan, A. L., Norris, A. G., Morrissey, B. M., & Albertson, T. E. (2015). Acute exacerbation of idiopathic pulmonary fibrosis-a review of current and novel pharmacotherapies. *Journal of thoracic disease, 7*(3), 499–519.

Kawamura, K., Ichikado, K., Anan, K., Yasuda, Y., Sekido, Y., Suga, M., Ichiyasu, H., & Sakagami, T. (2020). Monocyte count and the risk for acute exacerbation of

fibrosing interstitial lung disease: A retrospective cohort study. *Chronic Respiratory Disease*, *17*, 1479973120909840.

Tsuchiya, K., Suzuki, Y., Yoshimura, K., Yasui, H., Karayama, M., Hozumi, H., Furuhashi, K., Enomoto, N., Fujisawa, T., Nakamura, Y., Inui, N., Yokomura, K., & Suda, T. (2019). Macrophage Mannose Receptor CD206 Predicts Prognosis in Community-acquired Pneumonia. *Scientific Reports*, *9*(1), 18750.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature methods, 12*(4), 357–360.

King, T. E., Bradford, W. Z., Castro-Bernardini, S., Fagan, E. A., Glaspole, I., Glassberg, M. K., Gorina, E., Hopkins, P. M., Kardatzke, D., Lancaster, L., Lederer, D. J., Nathan, S. D., Pereira, C. A., Sahn, S. A., Sussman, R., Swigris, J. J., & Noble, P. W. (2014). A Phase 3 Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis. *New England Journal of Medicine*, *370*(22), 2083–2092.

Korfei, M., Ruppert, C., Mahavadi, P., Henneke, I., Markart, P., Koch, M., Lang, G., Fink, L., Bohle, R. M., Seeger, W., Weaver, T. E., & Guenther, A. (2008). Epithelial endoplasmic reticulum stress and apoptosis in sporadic idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, *178*(8), 838–846.

Kristiansen, M., Graversen, J. H., Jacobsen, C., Sonne, O., Hoffman, H. J., Law, S. A., & Moestrup, S. K. (2001). Identification of the haemoglobin scavenger receptor. *Nature, 409*(6817), 198-201.

Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome biology, 15*(2), R29.

Ley, B., Collard, H.R., & King, T.E. (2011). Clinical Course and Prediction of Survival in Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine 183*, 431–440.

Lodyga, M., Cambridge, E., Karvonen, H. M., Pakshir, P., Wu, B., Boo, S., Kiebalo, M., Kaarteenaho, R., Glogauer, M., Kapoor, M., Ask, K., & Hinz, B. (2019). Cadherin-11-mediated adhesion of macrophages to myofibroblasts establishes a profibrotic niche of active TGF-β. *Science Signaling*, *12*(564).

Maher, T. M., A. U. Wells, & Laurent, G.J. (2007). Idiopathic pulmonary fibrosis: multiple causes and multiple mechanisms? *European Respiratory Journal 30*, 835– 839.

Marcinak, S.J., and Ron, D. (2010). The Unfolded Protein Response in Lung Disease. *Proceedings of the American Thoracic Society 7*, 356–362.

Marimuthu, R., Francis, H., Dervish, S., Li, S., Medbury, H., & Williams, H. (2018). Characterization of Human Monocyte Subsets by Whole Blood Flow Cytometry Analysis. *Journal of visualized experiments : JoVE*, (140), 57941.

Martinon, F., Chen, X., Lee, A.-H., and Glimcher, L.H. (2010). Toll-like receptor activation of XBP1 regulates innate immune responses in macrophages. *Nature Immunology 11*, 411– 418.

Mathai, S. K., Gulati, M., Peng, X., Russell, T. R., Shaw, A. C., Rubinowitz, A. N., Murray, L. A., Siner, J. M., Antin-Ozerkis, D. E., Montgomery, R. R., Reilkoff, R. A., Bucala, R. J., & Herzog, E. L. (2010). Circulating monocytes from systemic sclerosis patients with interstitial lung disease show an enhanced profibrotic phenotype. *Laboratory investigation; a journal of technical methods and pathology, 90*(6), 812–823.

Meltzer, E. B., & Noble, P. W. (2008). Idiopathic pulmonary fibrosis. *Orphanet journal of rare diseases, 3,* 8.

Minami, K., Hiwatashi, K., Ueno, S., Sakoda, M., Iino, S., Okumura, H., Hashiguchi, M., Kawasaki, Y., Kurahara, H., Mataki, Y., Maemura, K., Shinchi, H., & Natsugoe, S. (2018). Prognostic significance of CD68, CD163 and Folate receptor-β positive macrophages in hepatocellular carcinoma. *Experimental and therapeutic medicine*, *15*(5), 4465–4476.

Misharin, A. V., Morales-Nebreda, L., Reyfman, P. A., Cuda, C. M., Walter, J. M., McQuattie-Pimentel, A. C., Chen, C.-I., Anekalla, K. R., Joshi, N., Williams, K. J. N., Abdala-Valencia, H., Yacoub, T. J., Chi, M., Chiu, S., Gonzalez-Gonzalez, F. J., Gates, K., Lam, A. P., Nicholson, T. T., Homan, P. J., … Perlman, H. (2017). Monocyte-derived alveolar macrophages drive lung fibrosis and persist in the lung over the life span. *Journal of Experimental Medicine*, *214*(8), 2387–2404.

Moeller, A., Gilpin, S. E., Ask, K., Cox, G., Cook, D., Gauldie, J., Margetts, P. J., Farkas, L., Dobranowski, J., Boylan, C., O'Byrne, P. M., Strieter, R. M., & Kolb, M.

(2009). Circulating Fibrocytes Are an Indicator of Poor Prognosis in Idiopathic Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine, 179*(7), 588–594.

Moore, V. C. (2012). Spirometry: Step by step. *Breathe, 8*(3), 232.

Moore, B.B., Fry, C., Zhou, Y., Murray, S., Han, M.K., Martinez, F.J., Flaherty, K.R. & COMET Investigators (2014). Inflammatory leukocyte phenotypes correlate with disease progression in idiopathic pulmonary fibrosis. *Frontiers in Medicine, 1*, 56.

Moore, B. B., Paine, R., Christensen, P. J., Moore, T. A., Sitterding, S., Ngan, R., Wilke, C. A., Kuziel, W. A., & Toews, G. B. (2001). Protection from Pulmonary Fibrosis in the Absence of CCR2 Signaling. *The Journal of Immunology*, *167*(8), 4368.

Naeim, F., Nagesh Rao, P., Song, S. X., & Phan, R. T. (2018). Chapter 2—Principles of Immunophenotyping. In F. Naeim, P. Nagesh Rao, S. X. Song, & R. T. Phan (Eds.), Atlas of Hematopathology (Second Edition) (pp. 29–56). Academic Press.

Park, S. W., Ahn, M. H., Jang, H. K., Jang, A. S., Kim, D. J., Koh, E. S., Park, J. S., Uh, S. T., Kim, Y. H., Park, J. S., Paik, S. H., Shin, H. K., Youm, W., & Park, C. S. (2009). Interleukin-13 and its receptors in idiopathic interstitial pneumonia: clinical implications for lung function. *Journal of Korean medical science*, *24*(4), 614–620.

Pechkovsky, D. V., Prasse, A., Kollert, F., Engel, K. M. Y., Dentler, J., Luttmann, W., Friedrich, K., Müller-Quernheim, J., & Zissel, G. (2010). Alternatively activated

alveolar macrophages in pulmonary fibrosis—Mediator production and intracellular signal transduction. *Clinical Immunology*, *137*(1), 89–101.

Richeldi, L., Cottin, V., du Bois, R. M., Selman, M., Kimura, T., Bailes, Z., Schlenker-Herceg, R., Stowasser, S., & Brown, K. K. (2016). Nintedanib in patients with idiopathic pulmonary fibrosis: Combined evidence from the TOMORROW and INPULSIS(®) trials. *Respiratory Medicine*, *113*, 74–79.

Richeldi, L., Fletcher, S., Adamali, H., Chaudhuri, N., Wiebe, S., Wind, S., Hohl, K., Baker, A., Schlenker-Herceg, R., Stowasser, S., & Maher, T. M. (2018). No Relevant Pharmacokinetic Interaction Between Nintedanib and Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis (IPF): Results from a Drug-Drug Interaction Study. In *D103. TRANSPLANT AND ILD POTPOURRI* (Vol. 1–308, pp. A7470–A7470). American Thoracic Society.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research, 43*(7), e47.

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology, 11*(3), R25.

Saito, S., Alkhatib, A., Kolls, J. K., Kondoh, Y., & Lasky, J. A. (2019). Pharmacotherapy and adjunctive treatment for idiopathic pulmonary fibrosis (IPF). *Journal of thoracic disease*, *11*(Suppl 14), S1740–S1754.

Schupp, J., Becker, M., Günther, J., Müller-Quernheim, J., Riemekasten, G., & Prasse, A. (2014). Serum CCL18 is predictive for lung disease progression and mortality in systemic sclerosis. *European Respiratory Journal, 43*(5), 1530-1532.

Scott, M.K., Quinn, K., Li, Q., Carroll, R., Warsinske, H., Vallania, F., Chen, S., Carns, M.A., Aren, K., Sun, J. and Koloms, K. (2019). Increased monocyte count as a cellular biomarker for poor outcomes in fibrotic diseases: a retrospective, multicentre cohort study. *The Lancet. Respiratory Medicine,* 7, 497-508.

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *The R journal, 8*(1), 289–317.

Seidler, S., Zimmermann, H. W., Bartneck, M., Trautwein, C., & Tacke, F. (2010). Age-dependent alterations of monocyte subsets and monocyte-related chemokine pathways in healthy adults. *BMC immunology, 11*, 30.

Sica, A., and Mantovani, A. (2012). Macrophage plasticity and polarization: in vivo veritas. *Journal of Clinical Investigation 122*, 787–795.

Snyder, L., Neely, M. L., Hellkamp, A. S., O'Brien, E., de Andrade, J., Conoscenti, C. S., Leonard, T., Bender, S., Gulati, M., Culver, D. A., Kaner, R. J., Palmer, S., Kim, H. J., Asi, W., Baker, A., Beegle, S., Belperio, J. A., Condos, R., Cordova, F., … on behalf of the IPF-PRO^{TM} Registry investigators. (2019). Predictors of death or lung transplant after a diagnosis of idiopathic pulmonary fibrosis: Insights from the IPF-PRO Registry. *Respiratory Research*, *20*(1), 105.

Stifano, G. and Christmann, R.B. (2016). Macrophage involvement in systemic sclerosis: do we need more evidence? *Current rheumatology reports, 18*, 2.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America, 102*(43), 15545–15550.

Vasarmidi, E., Bibaki, E., Koutoulaki, C., Mastrodemou, S., Trachalaki, A., Tzanakis, N., Tsitoura, E., & Antoniou, K. (2019). Evaluation of CD163 expression on alveolar macrophages from BAL of patients with Fibrotic Lung Diseases. *European Respiratory Journal*, *54*(suppl 63), PA4694.

Wei, J., Rahman, S., Ayaub, E., Dickhout, J., and Ask, K. (2013). Protein Misfolding and Endoplasmic Reticulum Stress in Chronic Lung Disease. *CHEST 143*, 1098-1105

Wong, A. W., Ryerson, C. J., & Guler, S. A. (2020). Progression of fibrosing interstitial lung disease. *Respiratory research, 21*(1), 32.

Wu, R., Zhang, Q.-H., Lu, Y.-J., Ren, K., and Yi, G.-H. (2015). Involvement of the IRE1α-XBP1 Pathway and XBP1s-Dependent Transcriptional Reprogramming in Metabolic Diseases. *DNA and Cell Biology 34*, 6–18.

Wynn, T. (2008). Cellular and molecular mechanisms of fibrosis. *The Journal of Pathology 214*, 199–210.

Wynn, T.A., & Ramalingam, T.R. (2012). Mechanisms of fibrosis: therapeutic translation for fibrotic disease. *Nature Medicine 18*, 1028–1040.

Wynn, T. A., & Vannella, K. M. (2016). Macrophages in Tissue Repair, Regeneration, and Fibrosis. *Immunity, 44*(3), 450–462.

Zaman, T., & Lee, J. S. (2018). Risk factors for the development of idiopathic pulmonary fibrosis: A review. *Current pulmonology reports, 7*(4), 118–125.

Zhang, L., Wang, Y., Wu, G., Xiong, W., Gu, W., and Wang, C. Y. (2018). Macrophages: friend or foe in idiopathic pulmonary fibrosis?. *Respiratory Research, 19*. 170.

Zhou, X., Moore, B.B. (2018). Location or origin? What is critical for macrophage propagation of lung fibrosis? *European Respiratory Journal 51*: 1800103

Zhou, Y., Peng, H., Sun, H., Peng, X., Tang, C., Gan, Y., Chen, X., Mathur, A., Hu, B., Slade, M.D. and Montgomery, R.R. (2014). Chitinase 3–like 1 suppresses injury and promotes fibroproliferative responses in mammalian lung fibrosis. *Science Translational Medicine, 6,* 240ra76-240ra76

Ziegler-Heitbrock, L., Ancuta, P., Crowe, S., Dalod, M., Grau, V., Hart, D. N., ... & Lutz, M.B. (2010). Nomenclature of monocytes and dendritic cells in blood. *Blood, 116*(16), e74-e80.