SIGNALING MECHANISMS OF PULMONARY ARTERIAL HYPERTENSION

By

Yajing Ji

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Pharmacology & Toxicology—Doctor of Philosophy

2022

ABSTRACT

SIGNALING MECHANISMS OF PULMONARY ARTERIAL HYPERTENSION

By

Yajing Ji

Pulmonary arterial hypertension (PAH) is a severe and life-threatening disease that is characterized by elevated pulmonary blood pressure. A challenge in treating PAH is that while the current generation of therapeutics alleviate symptoms, they fail to target the underlying causes of the disease. Initially it was thought that PAH is caused by increased pulmonary vasoconstriction; it is now understood that PAH mainly results from remodeling of the pulmonary vasculature. Further characterization of the underlying mechanisms of PAH will identify new pharmacological targets to treat PAH.

In this dissertation I seek to address this challenge from three distinct perspectives. In Chapter 2, I investigated the signaling network downstream of TGFβ and highlighted the MRTF/SRF pathway as potential therapeutical targets for PAH given its pivotal role regulating expression of contractile proteins in PASMCs. In Chapter 3, I aim to test whether TGFβ and the silencing of BMPR2, a member of the TGFβ family of receptors, contribute to the activation of lung fibroblasts in vitro. My results presented do not replicate the role of BMPR2 silencing found in other studies. This could be caused by the relatively short duration of BMPR2 silencing in our system. Finally, in Chapter 4, I perform a combined meta-analysis of several publicly available transcriptomic datasets of lung tissues from PAH patients. Using this approach, I identify PAH-associated signaling pathways, and chemical compounds which reverse a PAH-associated gene expression signature. My findings also suggest that while we bin PAH patients into various

subtypes in the clinic, on a transcriptional level, PAH patients tend to group into distinct gene expression clusters without relying on their clinical subtype. These findings improve our understanding of PAH biology and also highlight several potential drug targets for PAH.

ACKNOWLEDGEMENTS

I would first like to thank my labmates, including Behirda, Benita, Cassie, Clarissa, Erika, Eva, Hoa, Huijie, Indi, Jeff, Kate, Melissa, Tom, Sean and Zhangzhe for all their help with experiments and helpful discussions during lab meeting. I would like to thank all of the undergraduate, rotation, and summer students who have worked with me. Completion of this dissertation would not have been possible without the support provided by the department of pharmacology and toxicology. I would especially like to thank our current program director, Dr. Ann Dorrance for her support in science and graduate school. I would also like to thank the administrative and support staff from both departments, especially Bradley, Beverly, and Wendy, for helping with paperwork and guiding me.

I would like to thank my committee members: Dr. Harkema, Dr. Jackson, Dr. Roth and Dr. Krieger-Burke for years of support and advice. They have written letters of support, provided valuable feedback, and guided my project.

Finally, I would like to acknowledge my mentor Dr. Rick Neubig and thank him for all of the time and effort they put into my training. He has been a great role model for me, as a good person, a great scientist, and a wonderful leader. I appreciate it that he trusts me to give me autonomy and let me learn by making mistakes. He has helped me become the scientist I am today and has prepared me for wherever my career takes me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

αSMA alpha smooth muscle actin

ACVRL1 Activin A receptor type II-like 1

BMPR2 Bone morphogenetic protein receptor type

C2CP Canonical Pathways group of gene signatures in MSigDB

C6 Oncogenic Signatures group of gene signatures in MSigDB

CAF Cancer-associated fibroblasts

CAV1 caveolin-1

CMap Connectivity map

CNN1 Calponin 1

DEG Differentially expressed gene

DEGS Differentially expressed gene set

E2 Estrogen

ECs Endothelial cells

ECM Extracellular matrix

EGFR Epidermal growth factor receptor

EndoMT Endothelial-to-mesenchymal transition

ENG Endoglin

eQTL Expression-quantitative trait loci

ET Endothelin

FDR False discovery rate

GPCR G-coupled protein receptor

GWAS Genome-wide association study

HPASMC Human pulmonary arterial smooth muscle cell

IP Prostacyclin receptor

KCNK3 potassium channel subfamily K, member 3

logFC log Fold Change

miRNA micro RNA

MLC2 Myosin light chain 2

mPAP mean pulmonary arterial pressure

mRNA messenger RNA

MRTF Myocardin-related transcription factor

MSigDB Molecular Signatures Database

NO Nitric oxide

PAH Pulmonary arterial hypertension

PCA Principle components analysis

pMLC2 Phosphorylated myosin light chain 2

PPI Protein-protein interaction

ROCK Rho kinase

ROCK Rho-associated kinase

S1P Sphingosine-1-Phosphate

S1PR Sphingosine-1-Phosphate Receptor

SMAD Mothers against decapentaplegic homolog

SOX18 SRY-Box Transcription Factor 17

SphK1 Sphingosine kinase 1

SRF Serum response factor

ssGSEA Single sample gene set enrichment analysis

TGF-β Transforming growth factor β

CHAPTER 1 : INTRODUCTION

Definition and Classification of Pulmonary Arterial Hypertension

Pulmonary hypertension (PH) is a disease characterized by increased pulmonary arterial pressure (mPAP ≥20 mmHg). The World Health Organization (WHO) classifies patients with PH into five groups based on etiology [1]. Group 1 PH is pulmonary arterial hypertension (PAH), which is caused by abnormalities in the pulmonary vasculature that result in increased pulmonary vascular resistances. PAH is a rare disease with an estimated incidence of 5-15 cases per one million people per year [2]. Compared to other types of PH, PAH tends to affect adults in their midlife and is uncommon in elderly patients [3]. PAH preferentially affects females over males (female:male ratio ranges from 1.7 to 4.8:1.0) [2] and predominantly affects Caucasians (73 percent of registered cases) [4]. PAH is a devastating cardiovascular disease, which causes right heart failure and eventually death [5]. Recent improvements in our understanding of PAH have resulted in improved pharmacotherapeutic methods, leading to decreased population-based death rates from 4.6 to 1.7 per million from 2007-2011 [6]. However, there is still no cure for PAH and the 5-year survival rate is only approximately 40% [7]. One reason for the lack of effective therapeutic strategies is that the molecular mechanisms underlying PAH pathogenesis are still largely unclear.

PAH is a proliferative vasculopathy characterized by vasoconstriction, uncontrolled proliferation, fibrosis and thrombosis. The pathology of pulmonary arteries appears to be monomorphic across PAH patients [8]. However, PAH can be further classified into five subgroups based on their unique features. Idiopathic PAH (IPAH) is the PAH subgroup with neither family history nor known risk factors. The heritable PAH (HPAH) subgroup includes patients with familial germline alterations that convey increased risk of developing PAH. The associated PAH (APAH)

subgroup is associated with diseases known to cause PAH such as Scleroderma and HIV. The final two PAH subgroups are PAH which is induced by either a drug or toxin, PAH with pulmonary veno-occlusive disease and/or pulmonary capillary hemangiomatosis. Schistosomiasis appears to be the most common cause of PAH (APAH) worldwide. Second to that, over half of the remaining PAH cases are classified as IPAH, and approximately 10% of the remaining cases are classified as HPAH [9]. Albeit the differences among subgroups, all PAH patients share pathogenetic mechanisms and respond to the same set of medications.

Pathology of PAH

Initially, PAH was believed to result from vasoconstriction. However only a small portion of PAH patients show high vasoreactivity and respond to vasodilators [10]. PAH pathogenesis has now expanded to pulmonary vascular remodeling, which is characterized by medial hypertrophy/hyperplasia, intimal and adventitial fibrosis, *in situ* thrombotic lesions, and plexiform lesions, as well as peri-vascular infiltration of inflammatory cells. There is an increase in αSMA positive cells in the vasculature, which are responsible for media hypertrophy and occlusion of the vascular lumen. This vascular remodeling occurs mainly in distal muscular-type pulmonary arterial vessels and small pre-capillary arterioles (with diameters of 70–500 μm and 20–70 μm, respectively, in humans) [11]. The vasculopathy in PAH involves all three vascular layers (intima, media, adventitia) and various cell types.

Intima

Endothelial cells (ECs) form the intima layer of the vasculature and are essential for maintaining vascular homeostasis. They are responsible for producing vasoconstrictors, vasodilators, growth factors, and prothrombotic and antithrombotic mediators. Endothelial abnormalities are pivotal for initiating the pathogenesis and for promoting the development of PAH[8]. High shear stress and chronic hypoxia are two key contributors to endothelial injury and can induce endothelial apoptosis and cause DNA damage[12]. Endothelial dysfunction increases vasoconstriction, thrombosis, and proliferation of surrounding cells (predominantly smooth muscle cells and myofibroblasts) [12]. Interestingly, during the early stages of PAH development, there is an elevated level of endothelial cell apoptosis, which disrupts endothelial functions. However, at later stages of PAH, levels of endothelial apoptosis are reduced[13, 14]. This is because endothelial apoptosis that occurs early during PAH progression selects for subclones of endothelial cells that are inherently more resistant to apoptosis and have higher rates of proliferation [13, 14]. ECs in PAH have also switched into an inflammatory type, resulting in elevated release of inflammatory factors and elevated inflammatory responses[15]. These pathological characteristics of ECs are maintained after they are isolated from PAH patients and culture in vitro. This suggests that these pathological characteristics are a result of a genetic or epigenetic alteration or rewiring of signaling pathways in the cells.

Media

Smooth muscle cells (SMCs) in the vascular media layer of PAH patients have increased hypertrophy and hyperplasia[16]. This hyperplasia mainly results from either the inherent

characteristics of SMCs or external environmental triggers. SMCs isolated from PAH patients demonstrate faster growth compared to SMCs from controls with/without stimulation of growth factors[16]. This may result from re-wired signaling pathways in the SMCs that render them more responsive to signals from growth factors[16]. Meanwhile, endothelial cells are one of main sources of pro-proliferative stimulation, likely acting through a paracrine manner[17]. Surrounding inflammatory cells also act on SMCs and promote the vascular remodeling [12].

Adventitia

Cells that reside in vascular adventitia include fibroblasts, immune cells and pericytes[18]. Fibroblasts are responsible for the production and degradation of extracellular matrix (ECM) proteins in the adventitia. Adventitial fibroblasts from PAH patients are prone to be hyperproliferative, apoptosis-resistant and proinflammatory[18]. The over-proliferative fibroblasts are α smooth muscle actin (αSMA) positive and can migrate into the lumen of the vasculature, where the cells can occlude the vessels[19]. The proinflammatory fibroblasts contribute to inflammation in the lungs, in part by producing IL-6 which activates macrophages[20]. Fibroblasts in PAH increase the extracellular matrix (ECM) protein deposition in the adventitia[21], resulting in greater vascular stiffness. This increase in vascular stiffness can also affect the growth of SMCs and ECs [11].

Increased inflammation is an important pathological finding in PAH patients and contributes to disease progression[20]. Perivascular proinflammatory infiltrates include macrophages, monocytes, and lymphocytes and are found in the lungs of PAH patients. Inflammatory markers, such as IL-1β, IL-6[22] and IL-8 are increased in serum and in lung tissues

of PAH patients compared to the control group. These inflammatory mediators can drive the pathogenesis of PAH. For examples, IL-1β can induce endothelial mesenchymal transition (EndoMT)[23]. EndoMT makes ECs more proliferative and migratory. IL-6 is found to directly control the balance between proliferation and apoptosis of SMCs. These inflammatory features of PAH pathogenesis are especially prominent in PAH associated with autoimmune disease such as scleroderma and lupus[24].

Pericytes are resident αSMA positive mesenchymal cells in the adventitia. They are one of the central regulators of vascular development, remodeling and inflammation[11]. Pulmonary arterioles in PAH patients have excessive pericyte coverage, which also contributes to pathological accumulation of αSMA positive cells in the PAH vasculature [11].

Plexiform Lesions

PAH, especially APAH and IPAH, is characterized by complex vascular formations called plexiform lesions (PLs); glomeruloid-like vascular structures that originate from the pulmonary arteries[8]. PLs result from hyperproliferative endothelial cells, apoptosis-resistant SMCs and inflammatory cell infiltration[14]. In PLs, endothelial cells form slit-like “channels” within the occluded vessel lumen that are surrounded by SMCs, myofibroblasts, inflammatory cells and connective tissue matrix [25]

Pseudo-malignant Disease

The PAH pathology described above highlights several outcomes of apoptosis resistance, increased cell growth, inflammation, and angiogenesis in the vasculature of PAH patients[11].

These PAH hallmarks are very similar to characteristics of cancer[26]. PAH can be considered a pseudo-malignant disease, as first suggested in 1998 [26]. Investigations during the last decade have expanded the shared hallmarks between PAH and Cancer: genome instability and mutations, deregulation of cellular energetics, inflammation and avoidance of immune destruction[27].

A large number of proteins that are critical for PAH disease progression also play a central role in cancer. These include transcription factors and transcriptional co-activators including c-Myc, NFAT, and YAP/TAZ, among others[28]. These findings would suggest that the physiological signaling pathways which are co-opted to drive cancer progression may also be co-opted in the vasculature of PAH patients to drive PAH disease progression. This also suggests the exciting prospect that existing cancer drugs have the potential to be repurposed to treat PAH.

One theory of carcinogenesis is that cancer starts with a single cell that gradually accumulates genetic alterations, resulting in expansion of increasingly malignant clones[28]. Cancers have demonstrated genetic variants and genomic instability[28]. A list of genetic variants has also been identified in HPAH. Additionally, in IPAH patients, instability in short DNA microsatellite sequences has been reported in plexiform lesions[29]. Somatic chromosome abnormalities have also been reported in PAH lungs and cells[30].

The Warburg effect is characterized by a switch of cellular metabolism from glucose oxidation to glycolysis despite the presence of oxygen; this phenomenon provides metabolic intermediates for macromolecule synthesis[31]. This effect was originally described in cancer cells and is thought to be mostly caused by high demands for macromolecules in rapidly dividing cancer cells. Recent studies suggest that pulmonary vascular cells also demonstrate a Warburg phenotype[32]. The Warburg effect has also been described in different animal models of PH[33,

34]. The Warburg effect is shared by both PAH and cancer, this suggests that both diseases may share a similar metabolic requirement to sustain a hyperproliferative state.

Infiltration of immune and inflammatory cells in tumors creates an inflammatory microenvironment and promotes tumor growth[26]. Similarly, there is perivascular infiltration of innate and adaptive immune cells in PAH patients[12]. These immune cells produce inflammatory mediators that stimulate proliferation in a paracrine and endocrine manner. Increased numbers of regulatory T cells have recently been reported in peripheral blood from patients with IPAH[35] and regulatory T cells have been associated with cancer[36]. Also, both cancers and PAH can actively suppress immune responses. Cancers use inhibitory checkpoints such as CTLA-4 (cytotoxic T-lymphocyte-associated protein 4) or PD1 (programed cell death) to escape immune defenses[27]. PAH patients have reduced numbers of natural killer (NK) cells, which normally prevent vascular remodeling[27].

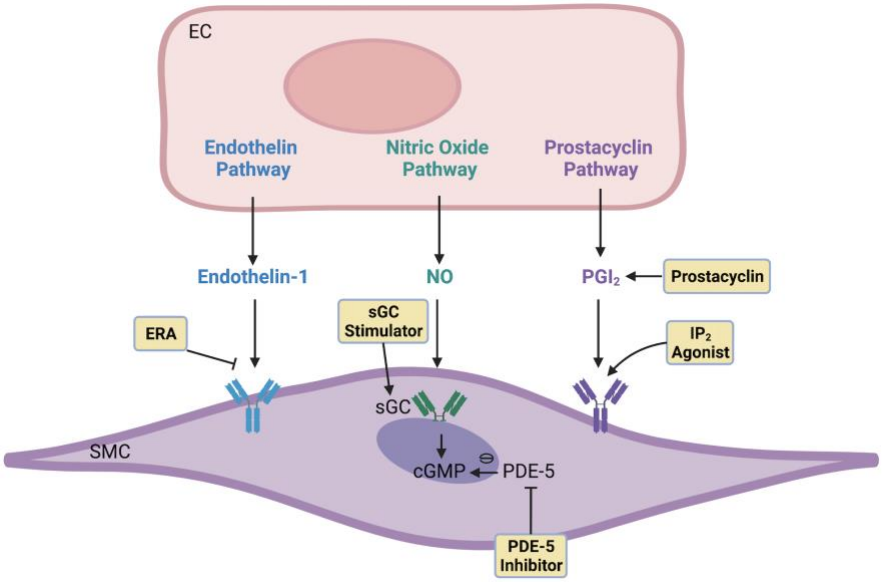
One difference between PAH and cancer is that the cells that contribute to PAH progression do not metastasize to distant organs. However, the cancer paradigm of PAH provides a new perspective to understand its pathobiology. The cancer-like concept has started a new field of investigation regarding the potential use of antiproliferative and/or oncological drugs in PAH. Several medications previously indicated in cancer treatment have been applied in the clinical trials against PAH, such as tyrosine kinase inhibitors (i.e. imatinib) [37]. These medications, such as Imatinib may have potential benefits in PAH. However, their clinical practice is still limited because of their safety profile and severe side-effects [37]. Further optimization or development of these potential medications are necessary.

Current Therapy

Calcium Channel Blockers

About 5-10% of the IPAH patients respond to a vasodilator challenge with significant reduction of pulmonary artery pressure. This includes patients with IPAH, HPAH, and drug/toxin induced PAH. These patients respond well to high doses of calcium channel blockers and have an excellent prognosis [10, 38]. Patients with associated forms of PAH are rarely vasoreactive [39]. Their increased vascular resistance most likely results from vascular remodeling instead of solely vasoconstriction.

Figure 1.1. Pathways targeted in current therapies for pulmonary arterial hypertension [40].



Endothelial dysfunction is one of the initiating events in PAH pathogenesis. ECs release factors that regulate vascular tone and proliferation of multiple cell types. Damage to ECs disrupts vascular homeostasis. Current pharmacological therapies for PAH target endothelial

cells through three different signaling pathways: Endothelin (ET) signaling, Nitric oxide-cyclic guanosine monophosphate (NO-cGMP) signaling and prostacyclin signaling [40].

Endothelin signaling

Endothelin is a 21-amino-acid peptide, predominantly produced by endothelial cells, and also by other cell types, including pulmonary artery smooth muscle cells and lung fibroblasts[41]. There are three peptide isoforms, ET-1, ET-2, and ET-3. ET-1 is the most abundant isoform in the cardiovascular system. There are two endothelin receptor isoforms, the ETAR and ETBR. Both are G protein-coupled receptors. While both ETAR and ETBR are simultaneously expressed in most cell types, only ETBR is expressed in endothelial cells[41]. ET-1 signaling mediates vasoconstriction in smooth muscle cells through activation of phospholipase C (PLC), which signals through inositol trisphosphate (IP3) and diacylglycerol (DAG)[42]. In contrast, ET-1 induces NO prostacyclin (prostaglandin I2) synthesis through ETBR[43]. ET-1 also induces vascular smooth muscle cell proliferation mediated by cytochrome p-450 arachidonic acid metabolites[17]. Plasma and lung ET-1 expression are increased in PH patients and its expression levels correlate with disease severity. ET promotes PASMC and fibroblast proliferation and stimulates fibroblasts to deposit extracellular matrix proteins. ET-1 treatment results in reduction of BMPR2 and BMP4 expression and enhances p38 MAPK activation in PASMCs. This contributes to the susceptibility of PAH in people with BMPR2 mutations[44]. Antagonists that target ETAR or antagonists that target both ETAR and ETBR have been developed to treat PAH. Two of these dual-receptor antagonists, Bosentan and Macitentan, and the ETAR-selective antagonist

Ambrisentan are approved by the US FDA for the treatment of PAH. The main adverse effects of ET antagonists are hepatotoxicity and peripheral edema [45].

NO-cGMP Signaling

The NO signaling is mainly mediated by guanylate cyclase/cyclic guanylate monophosphate (cGMP) pathway [46]. This, in turn, leads to dephosphorylation of myosin light chain which results in the dilation of blood vessels [46]. cGMP also regulates cellular proliferation and inflammation in the vasculature. NO inhibits vascular smooth muscle cell (VSMC) proliferation through cell cycle arrest [47]. In PAH, dysfunctional endothelial cells fail to produce sufficient levels of NO, which affects vascular tone. NO gas inhalation is beneficial to PH patients, especially pediatric patients, and it is mostly used in patients with severe PH[48]. The effects of cGMP are limited by its degradation which is induced by phosphodiesterases (PDEs), especially PDE-5. Thus PDE-5 inhibitors have great therapeutic potential for PAH. PDE-5 inhibitors have also been shown to decrease oxidative stress and demonstrate anti-inflammatory properties. Sildenafil, the first PDE-5 inhibitor, was first approved for treating erectile dysfunction, and then subsequently approved to treat PAH. Later, tadalafil and vardenafil are also used as the therapeutics to treat PAH[49]. Guanylate cyclase is the enzyme synthesizing cGMP. Riociguat, a soluble guanylate cyclase stimulator, was also approved to treat PAH. Riociguat not only acts synergistically with endogenous NO, but also directly stimulates guanylate cyclase activity in a mechanism that is independent of NO[50]. Additionally, Riociguat demonstrates antifibrotic, antiproliferative, and anti-inflammatory effects in multiple preclinical studies [50].

Prostacyclin pathway

Prostacyclin is a prostanoid formed by the cyclooxygenase (COX) pathway from endogenous arachidonic acid. The prostacyclin receptor (IP) is a G-protein coupled receptor (GPCR) that is found on platelets SMCs and ECs. Prostacyclin-induced activation of IP increases intracellular cAMP, activates protein kinase A, and consequently causes inhibition of platelet aggregation, relaxation of smooth muscle, and vasodilation of the pulmonary arteries[51]. The vasoconstriction, thrombosis and over proliferation in PAH is partly due to a lack of endogenous prostacyclin secondary to prostacyclin synthase downregulation[51]. Prostacyclin and its analogues (prostanoids) are widely used in the clinical management of PAH patients because of their potent vasodilating, antithrombotic, antiproliferative and anti-inflammatory properties. However, prostacyclin has a short half-life (only minutes) and this makes clinical use of prostacyclin challenging. Prostacyclin has been delivered by Inhalation, parental and subcutaneous[52]. Adverse events related to prostacyclin analogues are often due to the abrupt changes of the plasma levels of the drug. The side effects include jaw pain, diarrhea, flushing, and arthralgias [52].

One oral alternative to the currently available prostacyclin analogues is Selexipag, a selective IP receptor agonist. Selexipag is rapidly hydrolyzed in the hepatic microsomes to form an active metabolite[53]. Both Selexipag and its active metabolite selectively binds to IP receptor, while the active metobalites have a longer half-life (7.9 h)[54]. Prostacyclin is the most effective treatment for PAH, but to maximize this therapeutic strategy it will be important to generate new analogs that have increased bioavailability and longer half-lives.

For lower-risk profile patients, monotherapy of oral endothelin receptor antagonists or phosphodiesterase inhibitors may be appropriate to halt disease progression of PAH[55]. For patients with severe symptoms, rapid progression or markers of poor clinical prognosis, parenteral, inhaled or oral prostacyclins would be incorporated in the therapeutic regimen[56]. Combination therapy containing a parenteral agent is also commonly used for patients with the most severe conditions or patients who are not responsive to first line treatment. Patients not responding to medical treatment would be considered for atrial septostomy or lung transplantation[55].

Pathogenesis of PAH

There have been multiple germline mutations identified in PAH. However, the low penetrance of the mutations makes "second-hit" hypothesis widely accepted[57]. In addition to the mutations, a probable second genetic variant or environmental trigger is required to drive the PAH pathogenesis[57].

Genetic Contributors

BMPR2

Genetic studies in familial cases of PAH have revealed 20 genes that are associated with risk of developing PAH[58]. Among them, the most common heterozygous germline mutations are found in bone morphogenetic protein receptor 2 (BMPR2)[58]. Mutations in BMPR2 have been reported in up to 80% of HPAH patients[58]. Interestingly, up to 25% of IPAH patients also have abnormal BMPR2 structure or function[59]. The mutations in BMPR2 found in IPAH could

be caused by insufficient family histories, sporadic mutations, or low penetrance in family members[59]. BMPR2 mutations only show about 20% penetrance even in HPAH[60]. BMPR2 mutation carriers who actually developed PAH are found to have lower BMPR2 levels than those who did not develop the disease[60]. This suggests that additional regulatory mechanisms help control BMPR2 levels and activity.

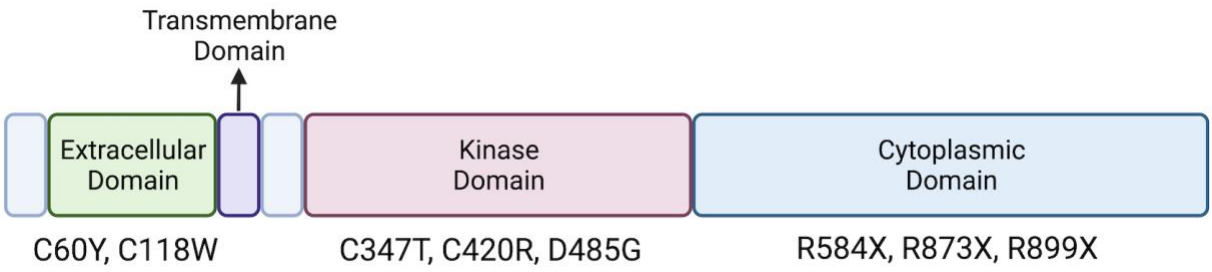
BMPR2 is a receptor in the transforming growth factor-β superfamily. In humans, the genomic structure of the BMPR2 gene covers at least 190 kb[58]. It has 13 exons which encode a 4 kb transcript that encodes a protein of 1,038 amino acids[58]. BMPR2 harbors four discrete functional domains, including an extracellular ligand binding domain encoded by exons 2 and 3, a transmembrane domain encoded by exon 4, and a serine/threonine kinase domain from exon 5 to exon 11[58]. These common domains are highly conserved in all the TGFβ superfamily receptors. Unique to BMPR2 is its extra-long C-terminal cytoplasmic tail which is encoded by exons 12 and 13[61]. While mutations are spread through the whole BMPR2 gene, these variants appear more frequently in the regions which encode the kinase domain[61]. There are over 300 BMPR2 mutations found, including missense, nonsense, and frameshift all of which result in loss of function[61]. BMPR2 dysfunction is mainly caused by haploinsufficiency, since the BMPR2 variants have lower expression level or fail to translocate to the cellular membrane[61]. Some PAH patients with BMPR2 mutations also have reduced levels of wildtype BMPR2. This suggests that certain BMPR2 mutations are dominant negatives[62].

Several pieces of evidence highlight the important role of BMPR2 in PAH pathogenesis. BMPR2 mutation carriers develop PAH 10 years earlier than noncarriers and have a worse prognosis [30]. Very few PAH patients with a BMPR2 mutation respond to acute vasodilators [59].

Overall, BMPR2 mutations demonstrate autosomal dominant inheritance and monogenetically drive the development of PAH but with low penetrance. In the model of "second-hit", additional triggers beyond the genetic background are required for patients to develop PAH[57]. The trigger could be the presence of modifier allele, or perhaps environmental factors that induce epigenetic alterations that interact with BMPR2 variant alleles on a signaling level. People carrying BMPR2 mutations have a higher risk to develop PAH in females (4:1) than males (2:1) [63]. PAH has demonstrated clear sex differences[63]. Estrogen is one of the well-studied triggers[63].

Additionally, BMPR2 variants also display genetic anticipation, common mechanisms such as elongation of repeat sequences due to polymerase slipping are not found in BMPR2, suggesting that this is the result of another yet to be determined mechanism[64].

Figure 1.2. Germline BMPR2 mutations in Pulmonary Hypertension[65].



Other Genetic Mutations

In addition to mutations in BMPR2, several other genes associated with PAH have been identified. Mutations in the activin A receptor type II-like 1 (ACVRL1), endoglin (ENG), and SMAD family member 9 (SMAD9) represent 1% of PAH patients[66]. Like BMPR2, the ACVRL1 is also a member of the TGF-β receptor family. ENG is part of the TGF-β receptor complex[66]. Finally,

SMAD9 is an important intracellular signaling molecule downstream of the TGF-β receptor[66]. Overall, the convergence of these variants in genes that encode on proteins in the TGF-β pathway highlight the critical role for this pathway in PAH. There are also additional variants in other signaling pathways as well. Mutations in caveolin-1 (CAV1) and the potassium channel subfamily K, member 3 (KCNK3) are also found in about 1% PAH patients[66]. CAV1 is a scaffolding plasma membrane-associated protein involved in cell cycle progression and many other signal transduction mechanisms. KCNK3 encodes a potassium channel[63]. In recent genomic studies, variants in the SRY-Box Transcription Factor 17 (SOX17) and T-box (TBX) are identified. SOX17 is a member of the conserved SOX family of transcription factors and is widely expressed during development; this gene regulates vasculogenesis and remodeling [67]. TBX gene encodes transcription factors that are involved in several developmental and cardiovascular diseases [68]. The majority of the mutations in PAH show autosomal dominant inheritance. Mechanisms may include haploinsufficiency or dominant negative effects. However, mutations in the eukaryotic translation initiation factor 2-alpha kinase (EIF2AK4) is autosomal recessive[69]. Mutated EIF2AK4 is predominantly found in patients with PAH that is associated with pulmonary veno-occlusive disease (PVOD) but also rarely found in patients with IPAH/HPAH[69].

These genetic findings facilitate the investigation of signaling pathways for PAH and should help in the development of personalized treatments. Meanwhile, ongoing studies are focused on using more robust patient populations to identify additional variants/genes that are associated with PAH.

Molecular Mechanisms Underlying PAH Pathogenesis

Development of PAH involves complex interactions among various cell types and multiple signaling pathways. Several signaling pathways have been relatively well characterized in PAH and shown great potential as therapeutic targets.

TGFβ Superfamily Signaling in PAH

TGFβ signaling is known to be critical in the maintenance of the biological homeostasis of pulmonary vasculature. TGFβ superfamily receptors are heterotetrametric complexes of type I and type II dual specific kinase receptors [70]. The type I receptors possess a GS domain (SGSGSG) that is critical for the regulation of receptor catalytic activity[71]. Ligand binding promotes interaction between receptor species, inducing phosphorylation of the type I GS domain leading to catalytic activation of the complex by the constitutively kinase-active type II receptor[71]. This then activates members of the SMAD family of signal transducers[71]. SMADs include three groups: receptor-regulated SMADs (R-SMADs), inhibitory SMADs (I-SMADs), and a common-mediator SMAD (i.e. SMAD4)[72]. Upon type I receptor mediated phosphorylation/activation of R-SMADs, they form heterometric complexes with SMAD4 and then translocate into the nucleus where they regulate gene expression[73].

The TGFβ superfamily includes TGFβs, BMPs, growth and differentiation factors (GDFs) and activins[73]. Canonically, TGF-β1 activates phosphorylation of SMAD2/3 and BMPs activate phosphorylation of SMAD1/5/8[74]. In addition to SMAD signaling, TGF-β superfamily receptors also activate non-SMAD signaling including MAP kinase pathways, Rho GTPase signaling pathways, and phosphatidylinositol-3-kinase/AKT pathways[75]. The Rho GTPases, including RhoA, Rac and Cdc42, play important roles in controlling dynamic cytoskeletal organization, cell

motility, and gene expression through a variety of effectors[76]. In different cells, TGFβ and BMP signaling serve different functions, such as embryo development, differentiation and osteogenesis[71].

As described above, several mutations have been found in BMPR2 in PAH patients. BMPR2 is widely expressed in various types of cells and tissues[77]. Homozygous BMPR2(-/-) knockout in mice is embryonic fatal whereas heterozygous BMPR2(+/-) mice are viable but do not develop pulmonary hypertension without a "second hit" such as hypoxia[77]. By contrast, both smooth-muscle-specific and endothelial-specific silencing of BMPR2 in mice cause increased right ventricular systolic pressure (RVSP)[64, 78]. Both BMPs and TGFβ induce differentiation and inhibit proliferation of PASMCs[16]. Interestingly, PASMCs from PAH patients with BMPR2 mutations exhibit an enhanced mitogenic response to TGFβ[64, 78]. The BMPR2 deficiency also induced apoptosis of ECs, which is a key process that initiate PAH[64, 78]. PAECs expressing a mutant BMPR2 also release higher levels of TGFβ into the medium, thereby accelerating SMC growth [79]. PAH patients show high levels of TGFβ1 in plasma and lung tissue[80]. The reduced BMP signaling and overactive TGFβ signaling have been recognized as one of the essential drivers of PAH[81].

How the loss of BMPR2 function impairs pulmonary vascular cell function in PAH is not completely understood. One hypothesis is that BMPR2 deficiency increases upstream ligands expression, which enhances the activation of the SMAD2/3 signaling. Mechanistically, BMP7 has been found to increase SMAD signaling in cells with siRNA knock-down of BMPR2 (e.g., BMP7) [82]. Missense mutations in the cytoplasmic tail of BMPR2 result in disrupted SMAD-independent pathways [32] instead of activation of SMAD signaling [34]. Additionally, BMPR2 mutations could

also alter the non-SMAD signaling pathways. In mouse epithelial cells transfected with BMPR-II mutants (C60Y, R491W, R899X) demonstrated ligand-independent activation of p38MAPK [83].

Some existing treatments that enhance cyclic nucleotide levels alter the balance of TGFβ and BMPR2 signaling. Riociguat and prostanoids can inhibit TGFβ signaling and favor the beneficial BMPR2 pathway[84]. Their beneficial effect on PAH can partly be explained by the inhibition of TGFβ signaling[84]. Several investigations have tried to restore the balance between BMP and TGFβ signaling to halt PAH progress[81]. Upregulation of BMPR2 by gene delivery counteracts the TGFβ induced effects and ameliorates PAH in a BMPR2-mutant mouse model[63].

GPCRs

G-protein-coupled receptors (GPCRs) are seven-transmembrane-domain proteins that regulate diverse intracellular signaling pathways. Activation of GPCRs initiates signaling through heterotrimeric G proteins and G-protein-independent pathways[85]. Signaling through the G-protein independent pathway is mediated by G-protein-coupled receptor kinase (GRK)-induced phosphorylation and arrestin coupling[85]. GPCR activity is regulated by factors that modify their ability to bind to and hydrolyze GTP to GDP[85]. Heterotrimeric G proteins are composed of three subunits, Gα, Gβ and Gγ[86]. Each G protein can couple to a diverse set of receptors and mediate different signaling. Gα effector pathways includes adenylyl cyclase, cGMP phosphodiesterase, phospholipase C and RhoGEFs[87]. Gβγ can recruit GRKs to the membrane and regulate potassium channels, voltage-dependent Ca2+ channels, phosphoinositide 3 kinase and mitogen-activated protein kinases[88]. There are four major Gα isotypes: Gs, Gi/o, Gq/11 and $G_{12/13}$. The vasoconstrictor response is mediated by Gi-, Gq-, or $G_{12/13}$-coupled GPCRs for ET-1 and

angiotensin II (Ang II)[89]. Vasoconstriction is mainly driven by Ca2+-dependent phosphorylation of Ser19 of myosin light chain (MLC), which is opposed by vasodilators. Additionally, activation of $G_{12/13}$ by vasoconstrictor GPCRs stimulates $G_{12/13}$-dependent RhoA GEFs to increase the activity of RhoA[89]. In turn, RhoA activates Rho associated kinase (ROCK) and then increased MLC phosphorylation, which leads to increased vasoconstriction[89]. Continued activation of vasoconstrictor GPCRs leads to a phenotypic switch, which promotes downregulation of contractile proteins, and upregulation of proliferative genes, increased migration capability and increased synthesis of extra cellular matrix components[90]. Activation of the GPCRs downstream signaling pathways also induce cytokine/chemokine productions in immune cells, VSMCs and fibroblasts, which contributes to the vascular inflammation and remodeling[90]. GPCRs are widely expressed in cardiovascular system and play an important role in the development of PH. Because of this, GPCRs are an exciting class of drug targets for PAH[90].

Rho signaling

RhoA/ROCK signaling plays key role in smooth muscle cell contraction, cell migration, and stress fiber formation[91]. RhoA is coupled with $G_{12/13}$ and ROCK, a serine–threonine kinase is the signaling node downstream of ET-1 and Ang II in VSMCs[76]. The first discovered ROCK inhibitor is Fasudil, which has been approved for treatment of cerebral vasospasm after subarachnoid hemorrhage[92]. With implication of ROCK in the pathogenesis of PAH, ROCK inhibitors have been tested in PAH animal models and shown promising protective effect [93]. For example, Fasudil markedly reduced right ventricular systolic pressure in hypoxia and Monocrotaline induced PAH in rats[91, 94]. Fasudil inhalation is as effective as NO at reducing pulmonary arterial

pressure in PAH patients, but only slightly decreased pulmonary vascular resistance[95]. There have not been any clinical trials assessing the long-term efficacy of fasudil in the treatment of PH[95]. Novel ROCK inhibitors are under development[96].

Wnt signaling

The Wnt signaling pathway is broadly separated into the β-catenin-dependent (canonical) and the β-catenin-independent (non-canonical) pathways. Canonical Wnt signaling is required for differentiation of vascular smooth muscle cells[97]. Wnt signaling also promotes proliferation and migration of pulmonary arterial endothelial cells[98]. The Wnt signaling is aberrantly activated in laser-microdissected plexiform lesions of PAH[99]. Wnt signaling has a major role in preserving pulmonary vascular homeostasis and its alteration can lead to vascular remodeling in PAH[98]. Activation of the Wnt is found to be essential for the establishment of human pulmonary endothelium-pericyte interactions[100]. Reduced Wnt signaling could decrease the viability of newly formed vessels and contribute to vascular pruning in PAH[100]. Mice with endothelial specific loss of Wnt5a showed a similar response to chronic hypoxia as wildtype mice but failed to recover after re-exposure to normoxia[100]. Wnt-specific compounds are under development to modulate the pathway activity to target PAH[101]. However, with the large range of cellular functions of Wnt pathway across the body, it is essential to minimize the potential off-target effects and systemic toxicity during the development of Wnt pathway modulators[101].

Estrogen

Women have a higher incidence of IPAH and heritable PAH than men, but female patients generally have better outcomes than male patients. This is called “estrogen paradox”[31]. It is believed that 17β-estradiol (E2) has divergent effects in the pulmonary vasculature versus right ventricle[102]. E2 has many metabolites such as 16α-hydroxy-estrogens. Several of these metabolites are biologically active and important modifiers of disease development[102]. There are two estrogen receptors (ERs): ERα and Erβ, which both are nuclear receptors. In genomic pathway, upon E2 binding to ERs, they translocate to nucleus, bind to an estrogen responsive element (ERE) and serve as a transcription factor[103]. In the nongenomic pathway, a mechanism that is independent of ER’s role as a transcription factor, upon E2 binding ERs directly activate kinases or second messengers and initiate rapid cellular effects. These nongenomic effects include activation of eNOS or prostacyclin synthase and activation of MAP kinases[103]. ERs are expressed in the cardiovascular, respiratory and multiple other organ systems. The function of E2 is dependent on the ER subtypes, their coactivators and corepressors and the cellular contexts[103].

The mechanism underlying the estrogen paradox is not fully understood. E2 is believed to have three-tier effects in PAH. E2 is protective in healthy pulmonary vasculature[104]. However, E2 stimulates angioproliferative increasing the risk and driving the progression of PAH[104]. Meanwhile, E2 inhibits/delays transition to maladaptive RV remodeling[104]. Although with complicated effects in PAH, several drugs targeting the estrogen pathways are under development to treat PAH[102]. One mechanism of the RV protection of E2 is E2/ERα/BMPR2/apelin axis[105]. Both E2 and ERα increased in BMPR2 levels in PH RVs and

isolated RV cardiomyocytes[105]. The protective effect of E2 on RV is mediated through BMP2 and apelin[105]. Aromatase is the enzyme converting androgens to E2. The effect of E2 can be reduced by inhibiting its synthesis enzyme with anastrozole and E2 receptor antagonist, tamoxifen. Both anastrozole and tamoxifen inhibited pulmonary hypertension in different animal models of PH, partly in a sex-dependent way[104].

Epigenetics

Epigenetic modification are involved in the pathological mechanisms of PAH[106]. Two common processes in Epigenetics are DNA methylation and histone modification[106]. The methylation pattern in CpG islands in gene promoter regions contributes to modifying gene transcription through regulation the binding of transcriptional activators/inhibitors[107]. Histone acetylation and deacetylation change the form of chromatin and regulate transcriptional activity[108]. The coordination of DNA methylation and histone modification are essential regulators of gene transcription[109].

Reactive oxygen species (ROS) are O_2 derivatives that are generated in all vascular cell types and are important in vascular biology[110]. SOD is one of the main ROS enzymatic scavengers[109], and its promoters and introns contain CpG islands[111]. The SOD2 gene methylation was shown in pulmonary artery in PAH fawn-hooded rat model[111]. Treatment with a DNMT inhibitor can reverse the decreased level of SOD2 in the rat model[111]. Also, SIRT3, a mitochondrial deacetylase, suppresses mitochondrial oxidative metabolism through inhibiting many mitochondrial enzymes[112]. Mice lacking SIRT3 spontaneously develop pulmonary hypertension; a loss-of-function SIRT3 polymorphism is associated with PAH development in

humans [112]. Through the regulation of mitochondrial metabolism, SIRT3 is involved in cell survival, proliferation and apoptosis, which are critical to PAH development[112]. Epigenetics plays a pivotal role in regulating genes involved in oxidative activities in cells, and thus contributes to the PAH pathogenesis.

Multiple studies focus on develop epigenetic modulation-based for PAH. Histone acetylation is implicated in PAH development[113]. The levels of HDAC1 and HDAC5 are increased in the lung samples from both PAH patients and a PH rat model[113]. HDAC inhibitors can ameliorate the PH phenotype in rat models[114]. However broad-spectrum HDAC inhibitors such as trichostatin A showed worse right ventricular dysfunction[114]. Thus a more specific HDAC inhibitor could be more promising with reduced adverse effects[114].

Bromodomain-containing protein 4 (BRD4) is a transcriptional modulator in the BET family[115]. BRD4 has higher affinity against proteins with multiple acetylated residues[115]. It interacts with hyper-acetylated histone regions and promotes the transcription[115]. BRD4 is involved in inhibiting apoptosis, promoting cell survival and stimulating hyperproliferation. BRD4 can also drive the cells into a proinflammatory phenotype, which show increased transcription of cytokines including interleukin 6 (IL-6) and interleukin 8 (IL-8)[116]. In PAH, BRD4 protein level is increased in the human lung tissues[117]. BRD4 is responsible for triggering the proliferation/apoptosis imbalance of PASMCs, and the elevated inflammation in the lung[117]. One clinically available BET inhibitor successfully reduces pulmonary hypertension in PAH animal models[117, 118] and demonstrates great potential as PAH treatment[118].

Bioinformatic Studies of PAH

PAH pathogenesis is complex and involves multiple cell types and various genetic and environmental factors. Gene expression studies, enabled by microarrays and more recent advances in RNA sequencing, are powerful screening technologies to detect groups of co-regulated genes or pathways that are involved in various aspects of PAH[119]. The rapid development of the sequencing technologies and corresponding analysis tools enabled more proficient systematic analysis of gene expression profiles to reveal the pathogenesis of PAH.

Lung tissues are collected during transplantation or biopsy and control samples are usually sourced from failed donors. Multiple RNA expression studies have been conducted on whole lung tissue samples[120], primary cells isolated from lungs and circulating cells from blood[121]. These studies aim to identify genes and pathways associated with PAH pathogenesis, detect new potential biomarkers and evaluate the effect of potential therapeutics on disease progression[120, 121].

Two studies of the lung tissues have identified downregulation of BMPR2 in PAH samples with or without BMPR2 mutations[122]. This agrees with the observation that BMPR2 expression is reduced at the protein level in lung tissue sections from PAH patients[123]. Another finding from these studies is that increased level of estrogen receptor 1 (ESR1) in female IPAH patients [122]. This finding may suggest why PAH preferentially affects females.

One RNA sequencing study of endothelial cells identified a correlation between BMPR2 dysfunction and reduced expression of collagen IV (COL4) and ephrin A1 (EFNA1), which are essential proteins in ECs migration and adhesion[124]. This suggests that BMPR2 deficiency correlates with endothelial dysfunction, which is a factor that drives PAH pathogenesis early

during disease progression[124]. Analysis of specific cell types, rather than whole tissues, may provide an advantage for identifying expression patterns that are linked to disease status. Expression profiling of PASMCs that were derived from IPAH patients demonstrated that these cells have increased expression of genes involved in cell proliferation, mitosis, and cytokine signaling[125]. Most notably, the adhesion G protein-coupled receptor (GPCR), ADGRG6/GPR126, is elevated in PASMCs derived from IPAH patients[125]. Interestingly, knockdown of ADGRG6/GPR126 with siRNA increases the proliferation of the IPAH PASMCs that were derived from IPAH patients and were subsequently grown in culture[125]. This suggests that the upregulation of ADGRG6 in IPAH PASMCs may be a compensatory mechanism that blunts further proliferation of PASMCs. Analysis of peripheral blood mononuclear cells (PBMCs) that were identified from IPAH patients had alteration in expression of genes related to inflammation and endothelial functions[126]. These genes show potential as biomarkers to assess the severity of PAH or perform personalized treatments[126].

In addition to protein-coding RNAs (mRNA), the noncoding RNAs such as microRNAs (miRNAs) have also been of interest[127]. MicroRNAs are small, non–coding endogenous RNA molecules that are ∼21 to 25 nucleotides in length[127]. One miRNA can regulate the expression of multiple genes; and vice versa multiple miRNAs can regulate one protein together[128]. Many miRNA are evolutionarily conserved and are responsible for modulating cellular differentiation, proliferation, survival, and metabolism[128]. One study found miR-210 expression in PASMCs is induced by hypoxia in PASMCs[129]. Upregulation of miR-210 increases apoptotic resistance in PASMCs, ultimately resulting in hyperplasia[129]. Another study demonstrated that circulating miR-150 is reduced in peripheral blood of PAH patients and correlates with a poor survival

rate[130]. Similarly, circulating miR-26a is found reduced in IPAH patients and positively correlate with the pulmonary function[131].

These large-scale studies of expression profiles on tissues derived from PAH patients has resulted in the identification of several genes and signaling pathways that are essential for PAH disease progression[120, 122, 132, 133]. In addition to identifying new potential therapeutic targets for the treatment of PAH, these studies have also identified several promising biomarkers that could have utility in developing better PAH diagnostic tests[134]. Still, these findings that are derived from high throughput gene expression-based analysis always require further verification and validation using in vitro or in vivo methods[135]. Changes in mRNA levels do not necessarily affect protein levels of the encoded gene[135]. Furthermore, the RNA expression profiles are prone to artifacts that are induced by sample collection and storage methods, RNA quality, and the method used to measure gene expression[135]. One challenge is that since PAH is a rare disease, it is difficult to collect large cohorts of samples, and the cohorts that are collected have significant covariates such as age, sex, severity of PAH, and comorbidities, as well as differences in ongoing therapeutic strategies. The presence of these confounders often makes analysis of these expression studies difficult since they may have a significant contribution to inter-patient heterogeneity in expression and the PAH disease state[136]. Because of these issues, differentially expressed genes identified in one study are sometimes not reflected in other similar studies[120, 122, 132, 133]. This makes it challenging to identify differentially expressed genes that are important in PAH, instead of those that are a result of experimental confounders.

There are multiple advances in the bioinformatic analysis algorithm, for example, through clustering of samples and genes into groups and analysis of signaling pathway activation by measuring expression of gene signatures, rather than individual genes[137]. More “omics” data such as proteomic and metabolomic data are becoming available, sometimes using the same tissue samples that were used for mRNA expression profiling[138]. The integration of all the clinical, genomic, transcriptomic, proteomic, and metabolomic data would provide the best opportunity to illustrate a complete picture of the molecular mechanism that underly PAH pathogenesis.

In total, these data suggest that TGF-β signaling is a clear driver of PAH pathogenesis. Most studies of this signaling pathway in PAH have focused on canonical downstream SMAD signaling. However, there is a growing body of evidence that non-canonical signaling pathways downstream of TGF-β are also relevant in regulating the vascular remodeling process. The impact of decreased BMPR2 signaling on TGF-β pathway activation, especially activation of non-canonical downstream pathways remains unclear. In this dissertation, I focus on better understanding the signaling pathways downstream of TGF-β, specifically on understanding the interaction between the SMAD and MRTF signaling pathways. I also focus on understanding how BMPR2 affects this signaling interaction. PAH pathogenesis is complex, and many signaling pathways contribute to disease progression. Advances in bioinformatics techniques and next generation sequencing technologies have produced a wealth of genome-scale transcriptomic data; in this dissertation I leverage these data to perform a meta-analysis of these data with the goal of identifying signaling pathways that are altered in PAH patients.

CHAPTER 2 : TGF-β1 INCREASES EXPRESSION OF CONTRACTILE GENES IN HUMA PULMONARY ARTERIAL SMOOTH MUSCLE CELLS BY POTENTIATING SPHINGOSINE-1-PHOSPHATE SIGNALING

Abstract

Pulmonary arterial hypertension (PAH) is a disease characterized by elevated pulmonary arterial pressure and carries a very poor prognosis. Understanding PAH pathogenesis is needed to support new therapeutic strategies. TGF-β drives vascular remodeling and increases vascular resistance by regulating differentiation and proliferation of smooth muscle cells (SMCs). Also, sphingosine-1-phosphate (S1P) has been implicated in PAH but the relation between these two signaling mechanisms is not well understood. Here, we define the signaling networks downstream of TGF-β in human pulmonary arterial smooth muscle cells (HPASMC) which involves SMAD signaling as well as Rho GTPases. Activation of Rho GTPases regulates myocardin-related transcription factor (MRTF) and serum response factor (SRF) transcription activity and results in upregulation of contractile genes expression. Our genetic and pharmacologic data show that in HPASMC, upregulation of alpha smooth muscle actin (αSMA) and Calponin (CNN1) by TGF-β is dependent on both SMAD and Rho/MRTF-A/SRF transcriptional mechanisms.

The time course of TGF-β-induced phosphorylated myosin-light chain 2 (pMLC2), a measure of RhoA activation, is slow, as is regulation of the Rho/MRTF/SRF-regulated genes (αSMA and CNN1). These results suggest that TGF-β1 activates Rho/pMLC2 through an indirect mechanism which was confirmed by sensitivity to cycloheximide treatment. As a potential mechanism for this indirect action, TGF-β1 upregulates mRNA for sphingosine kinase (SphK1), the enzyme that produces sphingosine-1-phosphate (S1P), an upstream Rho activator as well as mRNA levels of the S1P Receptor 3 (S1PR3). Both a SphK1 inhibitor and S1PR3 inhibitors (PF543 and TY52156/VPC23019) reduce TGF-β1-induced αSMA upregulation. Overall, this suggests a

model where TGF-β1 activates Rho/MRTF-A/SRF by potentiating an autocrine/paracrine S1P signaling mechanism through SphK1 and S1PR3.

Significance Statement

We suggest a model wherein, downstream of TGF-β1, S1P signaling bridges the interaction between SMAD and Rho/MRTF signaling in regulating αSMA expression in HPASMCs. The Rho/MRTF pathway is a signaling node in the αSMA regulation network and is a potential therapeutic target for the treatment of PAH.

Introduction

Pulmonary arterial hypertension (PAH) is characterized by elevated mean pulmonary arterial blood pressure and right heart failure which often leads to death. In the US, PAH has a prevalence of 12.4 cases/million people[139]. PAH patients are diagnosed at a mean age of 50 ±14 and their three year-survival rate is 63%[140, 141]. The standard of care for PAH is focused on alleviating symptoms, but fails to stop disease progression. To develop therapeutic approaches which halt PAH progression it is critical to better understand the cellular and molecular mechanisms underlying disease progression.

Both vasoconstriction and vascular remodeling contribute to increased vascular resistance in the pulmonary circulation, ultimately leading to elevated blood pressure[142]. Vasoconstriction results from the contraction of smooth muscle cells (SMCs) [142]. Dysregulated proliferation, migration, and hypertrophy of SMCs contribute to vascular remodeling [142]. TGF-

β regulates differentiation and proliferation of PASMCs [16] and is elevated in the serum of PAH patients[80]. TGF-β signaling is one of the main drivers of altered SMC behavior in PAH[81]. Generally, smooth muscle cells are categorized into two mutually exclusive phenotypes. The contractile phenotype has high expression of contractile proteins and a lower proliferation rate and is less migratory[143]. The proliferative phenotype shows low expression of contractile proteins and an elevated rate of proliferation and migration [143]. In PAH, PASMC have both increased proliferation and elevated levels of contractile proteins such as αSMA[144]. Interestingly, TGF-β promotes differentiation of PASMC isolated from non-diseased lungs but paradoxically drives proliferation of PASMC isolated from PAH patients[16]. Further investigation of how TGF-β regulates the proliferation and differentiation of PASMCs is important to understand the mechanism of this dual proliferative and contractile SMC phenotype in PAH pathogenesis.

TGF-β regulates gene expression through SMAD2/3, which is generally considered to be the canonical signaling pathway[145]. TGF-β also regulates gene expression through SMAD-independent mechanisms such as Rho/MRTF/SRF[145]. TGF-β activates the RhoA subfamily of small GTPases which, in turn, induces actin polymerization. Actin polymerization drives the transcription cofactor MRTF to translocate into the nucleus where MRTF binds to serum response factor (SRF) to regulate gene expression[146]. The suite of genes regulated by SMAD signaling and Rho/MRTF/SRF signaling overlaps in TGF-β modulated fibrosis and cellular migration, suggesting that these two transcriptional mechanisms may cooperate to regulate gene transcription[147]. Both SMAD and Rho/MRTF/SRF upregulate the expression of contractile genes and markers of differentiation, such as αSMA[146]. Elevated expression of contractile

genes results in increased contractility of SMCs[148], which could result in excessive vasoconstriction and SMC hypertrophy in PAH. During myofibroblast differentiation, SMAD3 interacts directly with an SRF-associated complex and mediates TGF-β-induced expression of SM22, another contractile protein[149]. It is unknown whether a similar mechanism applies to the transcriptional regulation of αSMA and CNN1 in HPASMC.

It is also unclear how TGF-β activates Rho/MRTF/SRF. In fibroblasts, TGF-β increases the level of the Rho activator S1P by upregulating its synthetic enzyme SphK1[150]. This suggests a model wherein TGF-β activates Rho signaling through S1P. S1P is a bioactive sphingolipid which binds to a GPCR family of S1P receptors, activates RhoA, and stimulates expression of αSMA and CNN1 in SMC[151]. SMAD3 activation is responsible for SphK1 upregulation in C2C12 myoblasts[152]. It is possible that TGF-β upregulates SphK1 levels through a SMAD pathway in HPASMC. This in turn could elevate S1P synthesis activating Rho/MRTF/SRF in an autocrine/paracrine manner. The plasma S1P level is increased in idiopathic PAH patients and in a rodent model of PAH compared to healthy patients and normal rats[153]. Genetic and pharmacologic inhibition of SphK1 activity is protective in PAH animal models[153, 154], highlighting the importance of S1P signaling in PAH. Further clarification of the signaling interaction of between TGF-β and the S1P pathways will increase our understanding of the molecular mechanisms underlying PAH pathogenesis.

In this study we explore mechanisms of crosstalk between SMAD signaling and Rho/MRTF/SRF signaling in order to better understand how TGF-β modulates SMCs. We provide evidence that S1P signaling bridges the SMAD and Rho/MRTF/SRF pathway to co-regulate gene

expression in HPASMCs, and we also identify a critical S1P receptor, S1PR3, involved in these mechanisms.

Materials and Methods

Cell Culture

Human Pulmonary Artery Smooth Muscle Cells (ThermoFisher, Waltham, MA #C0095C) were cultured in Medium 231 (ThermoFisher #M231500) supplemented with SMGS (smooth muscle growth supplement, ThermoFisher #S00725) and 1% Antibiotic-Antimycotic (ThermoFisher #15240062). HPASMCs (passage 6-8) were starved in 0.1% SMGS Medium 231 overnight prior to any experiments.

Compounds and Antibodies

Recombinant human TGF-β1 protein was purchased from Research And Diagnostic Systems, Inc.(Minneapolis, MN). Y-27632 (#S1049) was purchased from Selleckchem, Houston, TX. SIS3 (#15945), JTE-013 (#10009458), TY 52156 (#19119), VPC23019 (#13240), PF-543 (#17034) and SLP7111228 (#23290) were purchased from Cayman Chemical (Ann Arbor, MI). All compounds were dissolved in DMSO and frozen at -20 ^{o}C. Antibodies against MRTF-A (#sc21558) and MRTF-B (#sc98989) were purchased from Santa Cruz (Dallas, TX). Antibodies against MRTF-A (#14760), SMAD2/3 (#8685), pMLC (#3674) and MLC (#3672) were ordered from Cell Signaling (Danvers, MA). αSMA antibody (#7817) and pSMAD3 antibody (#52903) was purchased from Abcam (Cambridge, MA) and CNN1 antibody (13938-1-AP) was purchased from Proteintech (Rosemont, IL). All secondary antibodies [Donkey anti-Mouse680 (#C31216-02), Donkey anti-Mouse800 (#C90507-03), Donkey anti-Goat680 (#C41105-05), Donkey anti-Rabbit680 (#C40130-

02) and Donkey anti-Rabbit800 (#C90129-05)] are all purchased from LI-COR (Lincoln, NE). The MRTF/SRF pathway inhibitor CCG-222740[155] was obtained from the lab of Dr. Scott Larsen at the University of Michigan.

siRNA Transfection

ON-TARGETplus siRNA for MRTF-A (Dharmacon #L-015434-00-0010, Lafayette, CO), MRTF-B siRNA (Dharmacon #L-019279-00-0010) and non-targeting pool control (Dharmacon #D-001810-10-05) were used based on the manufacturer's protocol. siRNAs were diluted in OptiMEM, mixed with DharmaFECT (Dharmacon #T-2001-01) and then mixed with fresh medium 231 with 5% SMGS to a final concentration of 25 nM. Cells were seeded at a density of ~80% confluence and were transfected overnight. The next day the cells were serum starved for 16-20 hours prior to the treatment with TGF-β post-transfection.

RT-qPCR

HPASMCs were re-suspended in complete medium and 180,000 cells were seeded in each well of a 6-well plate. The cells were allowed to reach confluence (approximately four days) before being serum-starved in 0.1% SMGS Medium 231 overnight. Cells were treated as described in the figure legends and total cellular RNA was collected using the RNeasy kit (Qiagen, Hilden, Germany #74104) according to the manufacturer's protocol. The High-Capacity cDNA RT kit (ThermoFisher #4368814) was used to reverse transcribe the RNA into cDNA following the manufacturer's protocol. SYBR Green PCR Master Mix (ThermoFisher #4309155) was used to perform qPCR following the manufacturer's protocol on the Stratagene Mx3000P qPCR machine.

Fold-change of gene expression was normalized to GAPDH and analyzed by the $\Delta\Delta C_T$ method. Primer sequences are listed in Table 2.1.

Immunoblotting

HPASMC were cultured and treated as described in the figure legends. Total cellular protein was collected in 2x Laemmli Sample Buffer (Biorad, Hercules, CA #1610737). After heating the samples at 100 ^{o}C for 10 mins, protein samples were resolved on 10% (MRTF) or 12% (pMLC2/MLC2) polyacrylamide gels and transferred to PVDF Membranes (Millipore Sigma, Burlington, MA #IPFL00010). Blots were blocked in Odyssey Blocking buffer in PBS (LI-COR # 927-40000) at room temperature for 1 h, then incubated with primary antibody at room temperature for 1 h or overnight at 4^{o} C. Blots were washed three times for 5 min each with Tris-buffered saline with 0.1% Tween-20 (TBST) and then incubated with the appropriate secondary antibodies diluted 1:10,000 in blocking buffer at room temperature for 1 h. After 3 washes for 5 mins each with TBST, blots were imaged using a LI-COR Odyssey FC instrument and analyzed using Image Studio Lite software v.4.0.

Cell Proliferation

HPASMCs were cultured and were seeded in 96-well plates at a density of 10,000 cells/well. The next day, the cells were starved in 1% SMGS Medium 231 for 14~18 hours. Then the cells were cultured in 1% SMGS Medium 231 with or without 10 ng/mL TGF-β1 for two days. Cells were then fixed with 3.7% formaldehyde for 10 mins. After three washes with PBS for 5 mins, cells were stained with 500 ng/mL DAPI. Images were captured at the center of each well

using a Cytation 3 automated microscope (Biotek). All images were blinded by an automated R script before quantification. Cell numbers were then quantified using Image J.

Statistical Analysis

Data were analyzed through either Paired t-test or One-way Repeated Measures ANOVA followed by Dunnett's post-test using GraphPad Prism 7. Data are presented as the mean ± SD, and $P<0.05$ is considered statistically significant.

Results

Both SMAD and Rho/MRTF/SRF pathways are necessary for TGF-β-induced contractile gene expression

To identify the mechanism by which TGF-β1 regulates expression of the contractile proteins αSMA and CNN1 in HPASMC, we first inhibited phosphorylation of SMAD3 using SIS3. TGF-β1-induced αSMA and CNN1 expression were reduced to control levels by 10 μM SIS3, a SMAD3 phosphorylation inhibitor (**Figure 2.1A**). This suggests that phosphorylation of SMAD3 is important for TGF-β1-induced contractile gene expression. In order to test the role of Rho/MRTF/SRF in regulating contractile gene expression, we used the ROCK inhibitor Y27632 and the MRTF/SRF pathway inhibitor CCG-222740 [155]. Y27632 and CCG-222740 reduced TGF-β1-induced expression of contractile genes by ~ 60% and ~100% respectively (**Figure 2.1A & 2.1B**). Finally, siRNA-mediated silencing of MRTF-A reduced TGF-β1-induced protein levels of αSMA, however this effect was not observed when MRTF-B was silenced. (**Figure 2.1C).** Taken together,

these data show that both SMAD and the Rho/MRTF-A/SRF pathway are required for TGF-β1-induced contractile gene expression.

TGF-β indirectly activates Rho signaling

In order to better understand how TGF-β1 activates Rho signaling in HPASMC, we first measured the kinetics of TGF-β1-induced phosphorylation of myosin-light chain 2 (MLC2), a commonly used readout of Rho activation[156]. To assess the kinetics of TGF-β1-induced MLC2 phosphorylation, HPASMCs were treated with TGF-β1 for 1 to 9 hours; phosphorylated MLC2 (pMLC2) was increased only after 6-9 hours **(Figure 2.2B)**. In contrast, S1P-induced MLC2 phosphorylation is maximal after 0.5-1 hour (**Figure 2.2A**). S1P is a GPCR agonist that signals through $G_{12/13}$, which can rapidly activate RhoA and induce MLC2 phosphorylation. The delayed kinetics of TGF-β1-induced MLC2 phosphorylation suggests that the action of TGF-β1 is through an indirect, perhaps transcriptional/translational signaling mechanism. Consistent with this hypothesis, cycloheximide blocked TGF-β1-induced MLC2 phosphorylation but did not suppress, and even slightly enhanced, S1P-induced MLC2 phosphorylation (**Figure 2.2**). Thus TGF-β1-induced Rho activation, as detected by pMLC2 levels, requires the translation of new proteins (**Figure 2.2B**). In contrast, TGF-β1-induced SMAD3 phosphorylation peaks at 1 h and is not blocked by cycloheximide, consistent with the expected direct activation of SMAD3 phosphorylation by the TGF-β receptor (**Figure 2.7**).

TGF-β induces mRNA expression of SphK1

Given our hypothesis that TGF-β1 induces Rho activation indirectly by upregulating Rho activators, we next wanted to identify which factors may be mediating this process. We found that TGF-β1 increases the mRNA level of Endothelin 1, Connective tissue growth factor and SphK1. However, in our preliminary data, only S1P signaling inhibitors reduced TGF-β1-induced αSMA expression, so we focused the remainder of our studies on S1P signaling. We compared the kinetics by which TGF-β1 upregulates SphK1 with that for expression of αSMA and CNN1 (**Figure 2.3A**). HPASMCs were treated with TGF-β1 for 0-24 h. αSMA mRNA was upregulated by TGF-β1 at 6h, while CNN1 mRNA was increased at 3h. The TGF-β1-induced αSMA expression, thus, showed a lagging response compared to CNN1 (**Figure 2.3A**). αSMA mRNA was upregulated by TGF-β1 only at 12h while SphK1 mRNA was clearly increased at 3 and 6 hours, similar to results observed in fibroblasts where TGF-β1 activates Rho signaling through S1P [152]. The peak of SphK1 mRNA, 6h, is approximately the same time that αSMA mRNA started to increase. Importantly, the delayed increase in mRNA level of αSMA compared to SphK1 supports the idea that increased S1P could activate αSMA expression. The time course of the increase in CNN1 mRNA (**Figure 2.3A**) was more similar to that for SphK1 than for αSMA but CNN1 mRNA had a delayed peak. It started to rise at 3 hours, before the TGF-β1-induced phosphorylation of MLC2 (**Figure 2.2B**), however its upregulation at 24 hours was reduced by the ROCK inhibitor Y27632 (**Figure 2.1A**). This suggests that TGF-β1 regulation of CNN1 expression may have distinct mechanisms, Rho-dependent mechanism at later times but driven by other mechanisms early.

TGF-β modulates the expression level of an S1P receptor

In HPASMC, S1P elevates αSMA protein levels ~ 1.4 fold (**Figure 2.8**). In addition to the levels of S1P, the amount and composition of its receptors also determine cellular responses. Consequently, we tested whether TGF-β1 also modulates S1P receptor levels in addition to its effects on SphK1. Based on the literature, HPASMC express three subtypes of S1P receptors. Both S1PR2 and S1PR3 are coupled to $G_{12/13}$ which, in-turn, will result in Rho activation (Wamhoff, B. R 2008). We found that S1PR3 mRNA was elevated after 3 h of TGF-β1 treatment and was further increased after 12 h. S1PR1 and S1PR2 mRNA were unaffected by TGF-β1 (**Figure 2.3B**). The TGF-β1 stimulation of S1PR3 mRNA levels was reduced by the SMAD3 phosphorylation inhibitor SIS3 from 2.7- to 1.6- fold over control (95% confidence interval of the difference is -2.5 to 0.3, **Figure 3C**). Although this effect was not statistically significant due to variability, treatment with the SMAD inhibitor SIS3 reduced the stimulation of S1PR3 expression by about 65%.

A S1PR3 antagonist but not a S1PR2 antagonist reduces TGF-β-induced αSMA expression

To determine which S1P receptors might be functionally relevant for TGF-β1-induced αSMA and CNN1 expression, HPASMC were treated with TGF-β1 for 24 hours along with the S1PR2 antagonist JTE013, the S1PR3 antagonist TY52156, or the dual S1PR1/3 antagonist VPC23019. Both the S1PR3 and the S1PR1/3 antagonist reduced TGF-β1-induced stimulation of αSMA and CNN1 mRNA levels at 24 hours (**Figure 2.4A**). There was no effect of the S1PR2 antagonist JTE013. αSMA protein levels were also reduced by the two antagonists (TY52156 and VPC23019) targeting the S1PR3 receptor (**Figure 2.4B**). Inhibition of CNN1 protein levels by these two antagonists was only modest and did not achieve statistical significance. The overlapping

effect of S1PR3 and S1PR1/3 antagonist leads us to conclude that TGF-β1 regulates αSMA expression through S1P signaling via the S1PR3, while CNN1 expression may also be regulated by S1PR3-independent signaling mechanisms.

SphK1 inhibitors reduce TGF-β-induced αSMA expression

To test whether SphK1 regulates TGF-β-induced contractile genes expression, HPASMCs were treated with TGF-β1 for 24 hours with and without a SphK1 inhibitor, PF-543 at 10 μM (**Figure 5**). PF-543 decreased both the αSMA and CNN1 mRNA. PF-543 trended towards decreasing αSMA protein, but not CNN1 protein. This finding is similar to that of the S1PR3 receptor antagonists. Overall, these findings suggest that SphK1 is important for regulating TGF-β-induced αSMA expression and that CNN1 expression is not completely dependent on S1P signaling.

Discussion

TGF-β signaling is enhanced in PAH patients [80]and transgenic mice overexpressing TGF-β1 spontaneously develop PAH[157]. The contributions and interactions of SMAD and Rho/MRTF and other mechanisms downstream of TGF-β that contribute to PASMC activation remain controversial[144, 157, 158]. Also, the RhoA-activating GPCR agonist S1P plays an important role in PAH[145]. Here, in HPASMCs, we investigated the interaction between TGF-β1 and S1P signaling to further investigate the signaling network downstream of TGF-β, especially the SMAD and Rho/MRTF pathways.

SMAD signaling interacts with the Rho/MRTF/SRF pathway in multiple contexts. For example, in Monc-1 neural crest cells, RhoA directly regulates the phosphorylation of SMAD[159]. In cardiac myoblasts, an MRTF-A/pSMAD complex serves as a transcriptional regulatory element controlling the expression of αSMA[160]. In HPASMCs, the Zabini group reported that loss of SMAD3 disinhibits MRTF and drives the αSMA expression in PAH [144]. However, the Hansmann group did not observe reduced expression of SMAD3 in the lungs from SUGEN/Hypoxia rats and pulmonary arteries from PAH patients and claimed that TGF-β1 signaling drives αSMA expression in HPASMC through canonical SMAD3 activation instead of by SMAD3 downregulation[157]. Our results confirm that both SMAD and Rho/MRTF/SRF signaling are involved in the regulation of αSMA and CNN1 in HPASMC but implicate a sequential mechanism rather than a protein-protein interaction in these effects. We also show that MRTF-A, but not MRTF-B is required for regulating αSMA expression in HPASMC.

The interaction between TGF-β and S1P signaling has been characterized in fibrosis and cancer. In those settings, TGF-β increases the expression of SphK1 and in turn the level of S1P in fibroblasts, which contributes to TGF-β-mediated modulation of gene expression. SphK1/S1P has been reported to mediate TGF-β1-induced proliferation in rat PASMCs [161], however, we did not observe a significant proliferative effect of TGF-β1 in human PASMC (**Figure 2.9**). This highlights potential differences between the responses of rat and human PASMC to the same stimuli. In addition to SphK1, TGF-β upregulates S1PR3 through the SMAD3 signaling axis in lung adenocarcinoma cell lines and in those cells the S1PR3-mediated signaling drives the lung carcinoma cells growth [162]. We found a similar interaction between TGF-β and S1P signaling in regulation of αSMA in HPASMC; TGF-β1 elevates the level of αSMA, which is reduced by inhibition

of either SphK1 or S1PR3[150]. Since SphK1 and S1PR3 are two critical components of the S1P signaling pathway, this highlights the importance of the S1P signaling in regulating αSMA expression and further supports their potential as PAH drug targets. Our results showed the S1PR3 antagonists markedly reduced the elevation of αSMA levels, but this was not completely. One possibility is that TGF-β1 regulates αSMA expression through the SMAD pathway in parallel. Alternatively, TGF-β1 may also upregulate other Rho activators such as Endothelin-1 and CTGF, which in turn contribute to the αSMA elevation, even though blocking either Endothelin-1 or CTGF alone did not reduce TGF-β1-induced αSMA expression. In PAH, PASMC from PAH patients showed elevated S1PR2 levels [154]. Silencing of S1PR2 or pharmacological inhibition of S1PR2 has been shown to ablate S1P-stimulated SMC proliferation[154]. S1P regulates proliferation and differentiation of SMCs through different S1P receptors, which are coupled to different G proteins[163]. Thus, TGF-β-induced S1P levels could result in increased differentiation or proliferation of PASMCs based on the composition of S1P receptors of the cells.

Phosphorylation of MLC2 causes contraction of SMC. Additionally, increased expression of the contractile protein αSMA also contributes to the elevated vasoconstriction and SMC hypertrophy[144, 148]. We found both SMAD3 and Rho/MRTF-A/SRF are important mediators of TGF-β1-induced αSMA expression in HPASMC. In this study we demonstrated the essentiality of the S1P synthetic enzyme SphK1 and S1PR3 in the regulation of αSMA expression by TGF-β1, which implicates a model how TGF-β1 activates MRTF-A/SRF and modulates expression of contractile genes in HPASMCs. In total, these findings suggest that MRTF-A/SRF is a potential therapeutic target to reduce vascular contraction and SMC hypertrophy in PAH.

Figure 2.1. TGF-β1-induced αSMA and CNN1 expression is dependent on both the SMAD pathway and the Rho/MRTF/SRF pathway.

HPASMCs were seeded into 6-well plates (for the siMRTF protein experiment, cells were seeded into 60 mm petri dish) and allowed to reach confluence. Then the cells were serum deprived in 0.1% SMGS Medium 231 overnight before being treated the next day. **A.** Cells were pre-treated with 10 µM SMAD3 inhibitor SIS3 for 1h and then along with 10 ng/ml TGF-β1 for another 12h. Cells were also concurrently treated with 10 µM ROCK inhibitors Y-27632 and 10 ng/ml TGF-β1 for 24h. DMSO was used as a vehicle control. After treatment total cellular RNA was extracted and RT-qPCR was used to measure αSMA and CNN1 expression. **B.** HPASMC were co-treated with 10 ng/ml TGF-β1 and 10 µM CCG-222740 or DMSO. Both mRNA and protein were collected to test the expression levels of αSMA and CNN1. **C.** HPASMC cells were transfected with 5 µM siMRTF-A or siMRTF-B as described in materials and methods. The cells were then starved as described above before being treated with 10 ng/ml TGF-β1 for 24 hours. Protein (N=3 Santa Cruz anti-MRTF-A, N=1 Cell Signaling anti-MRTF-A) levels of αSMA, MRTFs and GAPDH were examined. (N>=3. Mean with SD. Paired t-test * P<0.05, ** P<0.01)

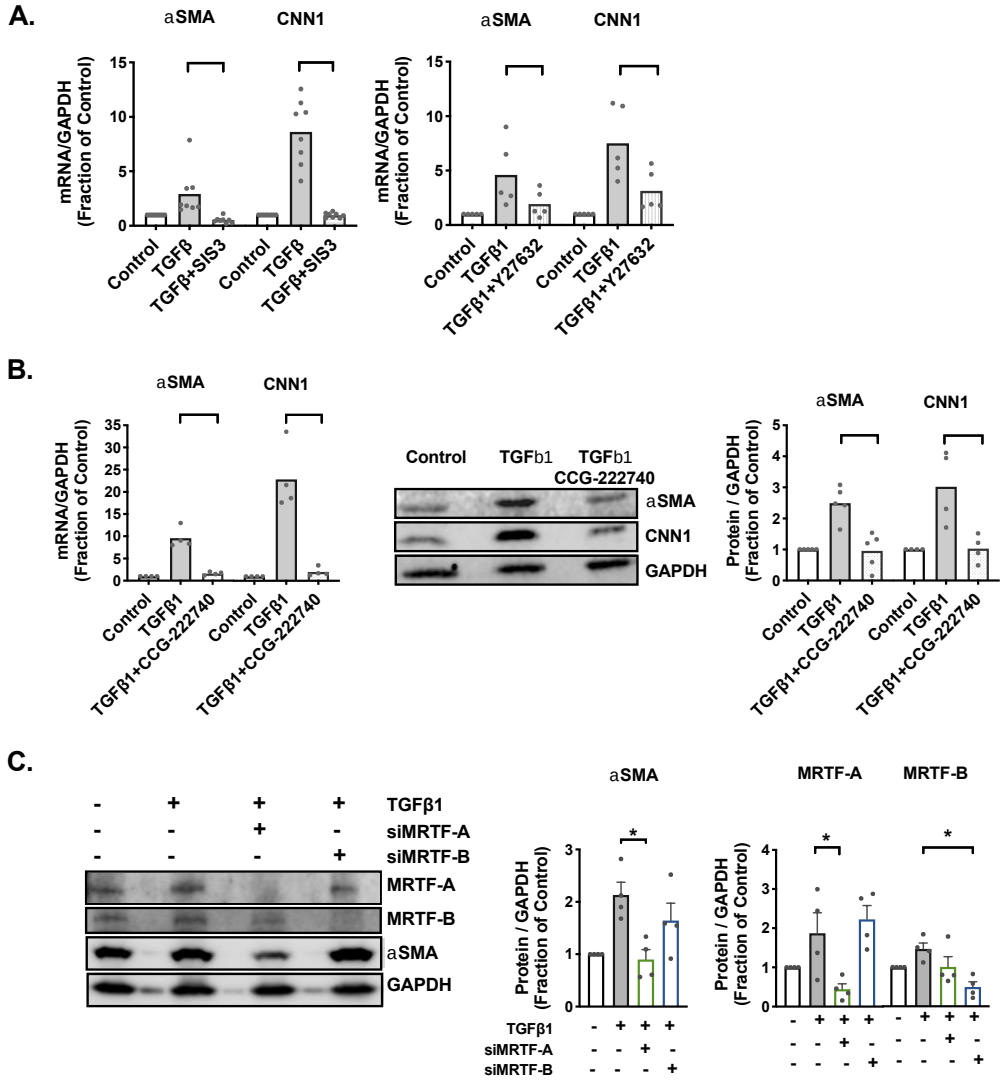


Figure 2.2. TGF-β1 indirectly activates the Rho signaling.
HPASMCs were seeded into 6-well plates and allowed to reach confluence before being serum deprived in 0.1% SMGS Medium 231 overnight. The next day the cells were treated with 10 ng/ml cycloheximide or vehicle and were stimulated with 10 ng/ml TGF-β1 for 1 h, 3 h, 6 h or 9 h. Protein was extracted and the levels of pMLC, total MLC, and GAPDH were measured. The ratio of pMLC/MLC is compared. (N=5. Data are represented as Mean with SD. Paired t test * P<0.05)

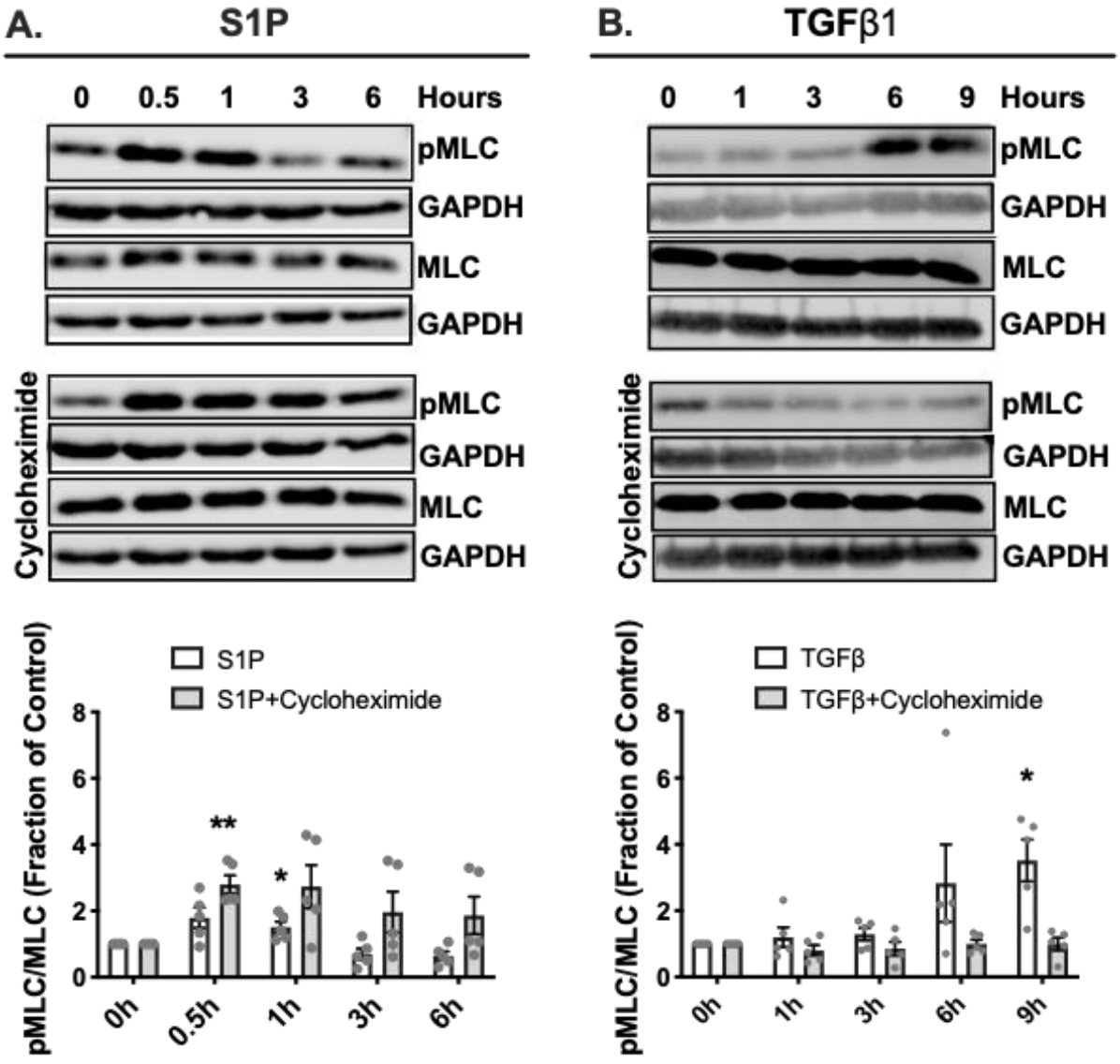


Figure 2.3. Kinetics of TGF-β1-induced contractile genes and TGF-β1 modulation of S1P receptors.

HPASMC were seeded into 6-well plates and allowed to reach confluence before being serum deprived in 0.1% SMGS Medium 231 overnight. **A.** Cells were then treated with 10 ng/ml TGF-β1 for 1, 3, 6, 12, 24 h in low serum medium. Afterwards, mRNA was extracted and the levels of αSMA, CNN1, and SphK1 were measured by RT-qPCR. **B.** The mRNA levels of S1PR1, S1PR2 and S1PR3 were measured by RT-qPCR. **C.** HPASMC were pretreated with 10 μM SIS3 for 1 hour and then treated with 10 ng/ml TGF-β1 for 12 hours in low serum medium. RT-qPCR was used to assess the mRNA level of S1PR3. (N=5. Data are represented as Mean with SD. For Figure 2.3A, One-way Repeated Measures ANOVA followed by Dunnett's post-test was performed. For Figure 3B, Paired t-test was conducted. * P<0.05, ** P<0.01)

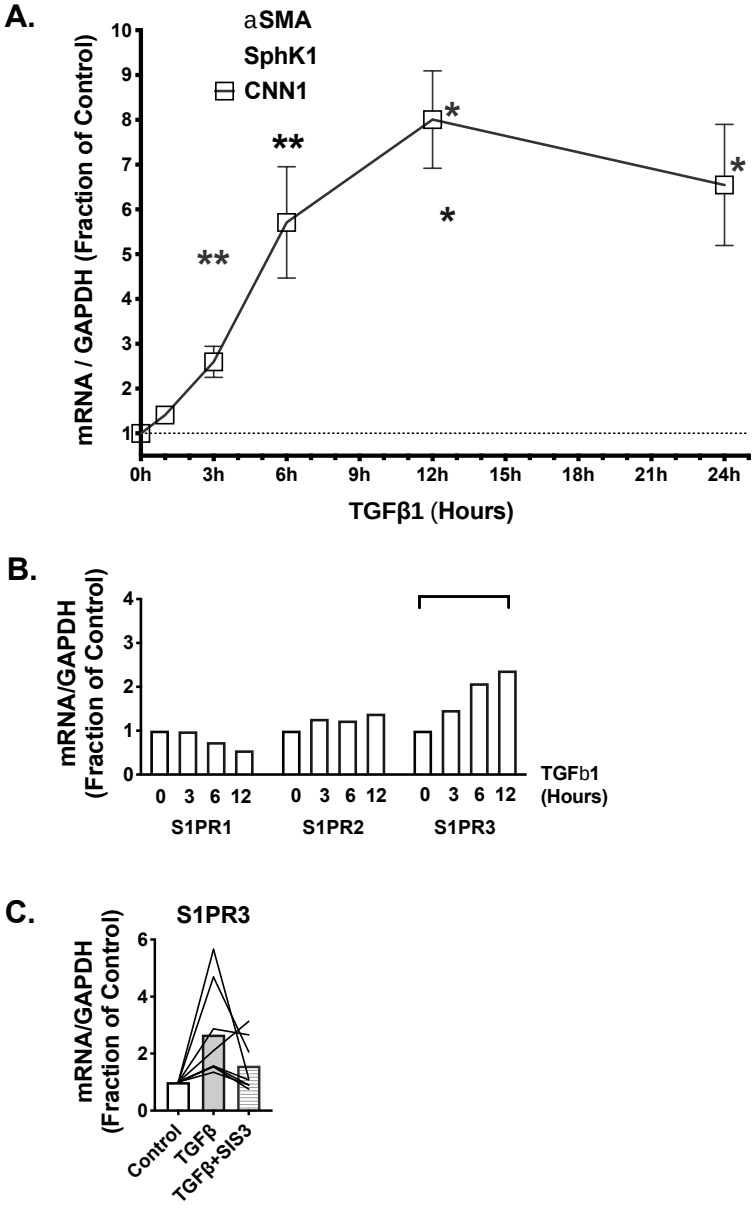


Figure 2.4. TGF-β1-induced αSMA and CNN1 expression is mediated by S1PR3 but not S1PR2. HPASMC were seeded into 6-well plates and were allowed to reach confluence before being serum deprived in 0.1% SMGS Medium 231 overnight. Cells were treated with 10 ng/ml TGF-β1 for 24 h and 5 µM JTE013 (S1PR2 antagonist), 1 µM TY52156 (S1PR3 antagonist), 2 µM VPC23019, or vehicle control. The protein and mRNA levels of αSMA and CNN1 were analyzed by western blot and RT-qPCR respectively. Both blots were probed on the same membrane, but the bands were cropped to include only relevant lanes. αSMA protein levels in TGF-β1 treated groups with or without TY52156 and VPC23019 show a mean of differences of -0.47 with 95% CI of -0.93 to 0.01 and -0.94 with 95% CI of -1.58 to -0.31. (N>=3. Data are represented as Mean with SD. Paired t-test* P<0.05)

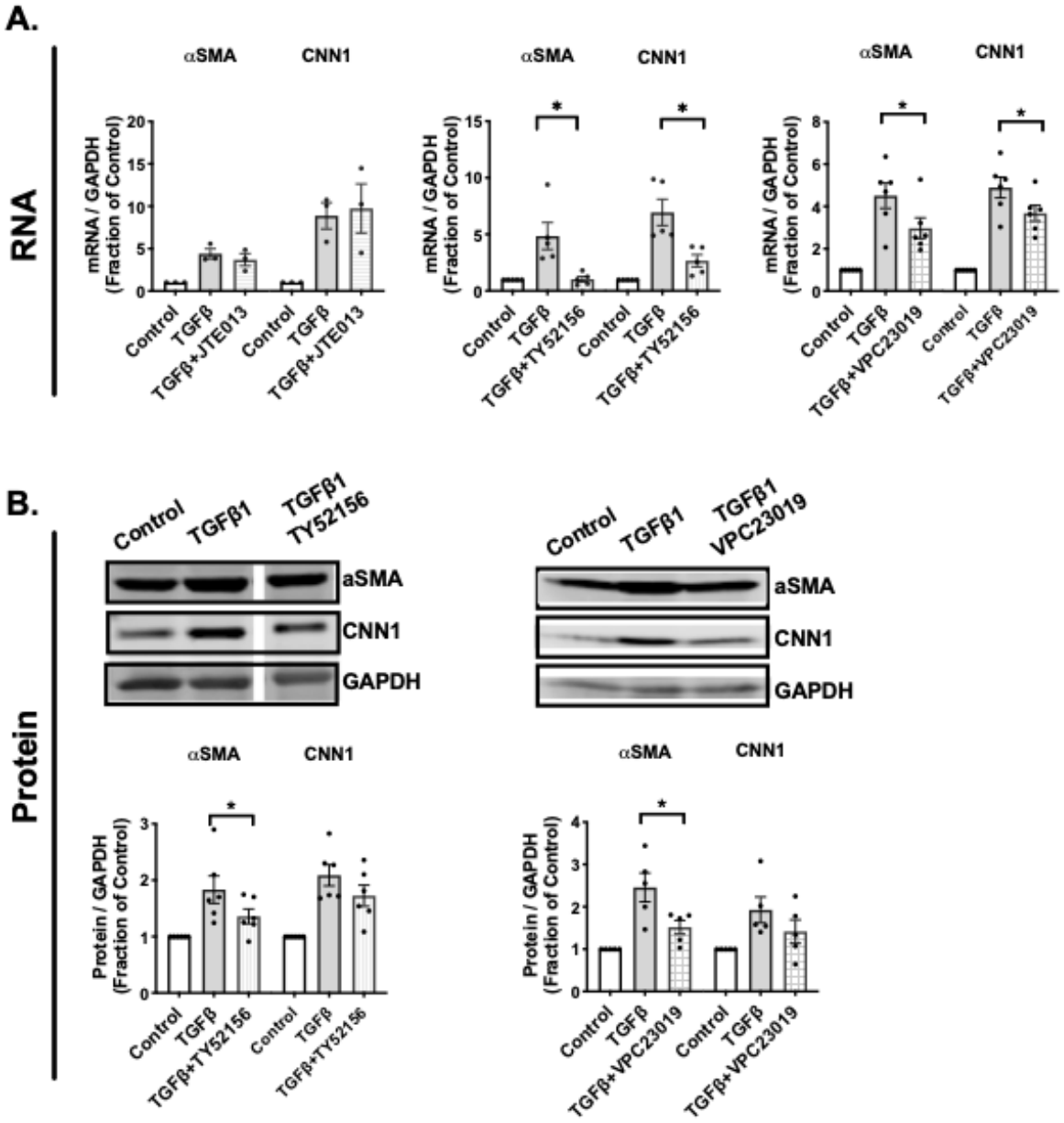


Figure 2.5. TGF-β1-induced αSMA and CNN1 expression is reduced by SphK1 inhibitors.
HPASMC were seeded into 6-well plates and were allowed to reach confluence before being serum deprived in 0.1% SMGS Medium 231 overnight. Cells were treated with 10 ng/ml TGF-β1 for 24 h and 10 μM PF542 or vehicle control. The protein and mRNA levels of αSMA and CNN1 were analyzed by western blot and RT-qPCR respectively. Both blots were probed on the same membrane, but the bands were cropped to include only relevant lanes. The αSMA protein levels in TGF-β1 treated groups with or without PF-543 show a mean of differences of -0.6779 with 95% CI of -1.646 to 0.2898. (N>=3. Data are represented as Mean with SD. Paired t-test* P<0.05)

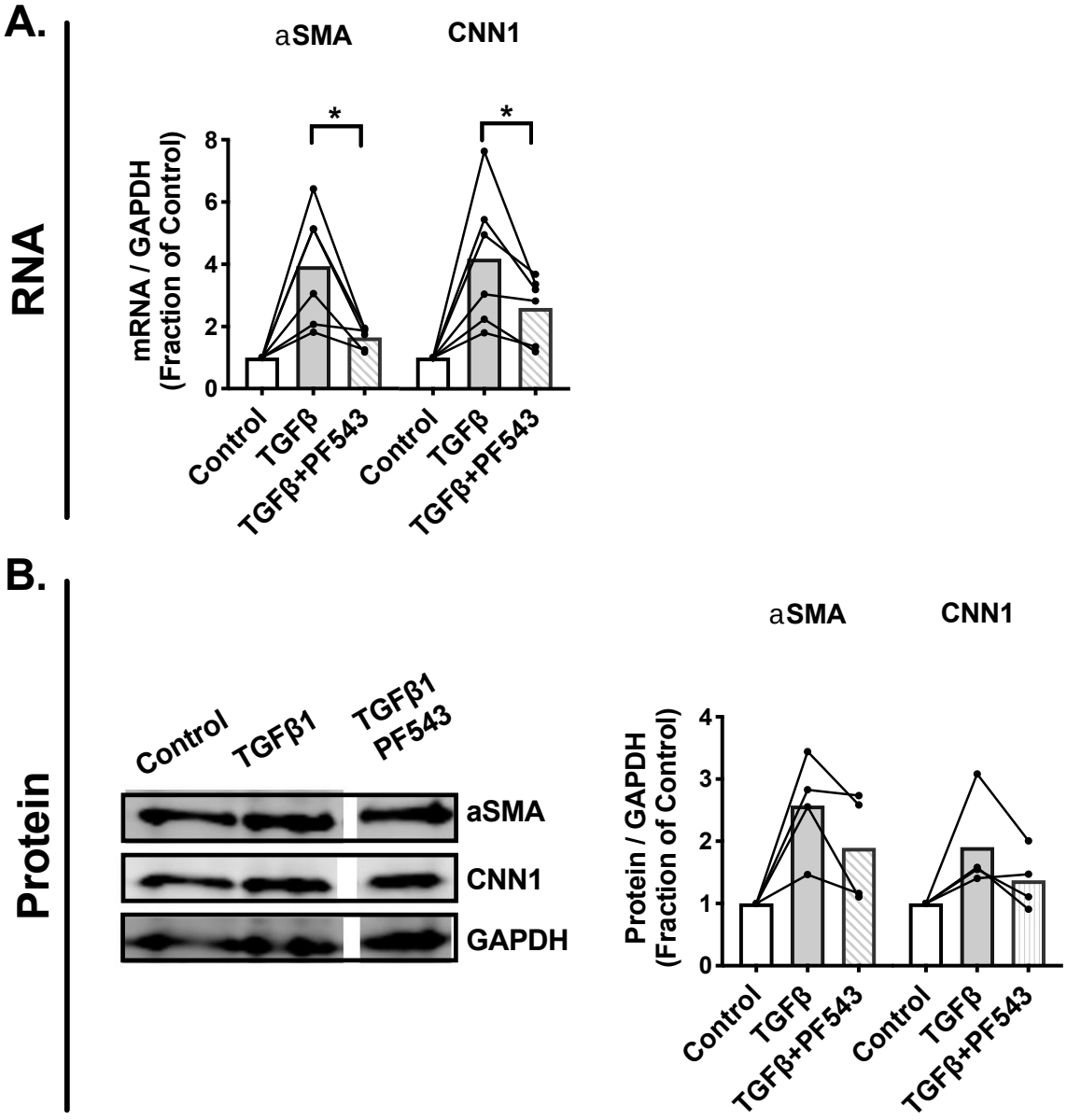


Figure 2.6. Summary overview of the interaction of TGF-β and S1P signaling in regulation of αSMA expression in HPASMC.

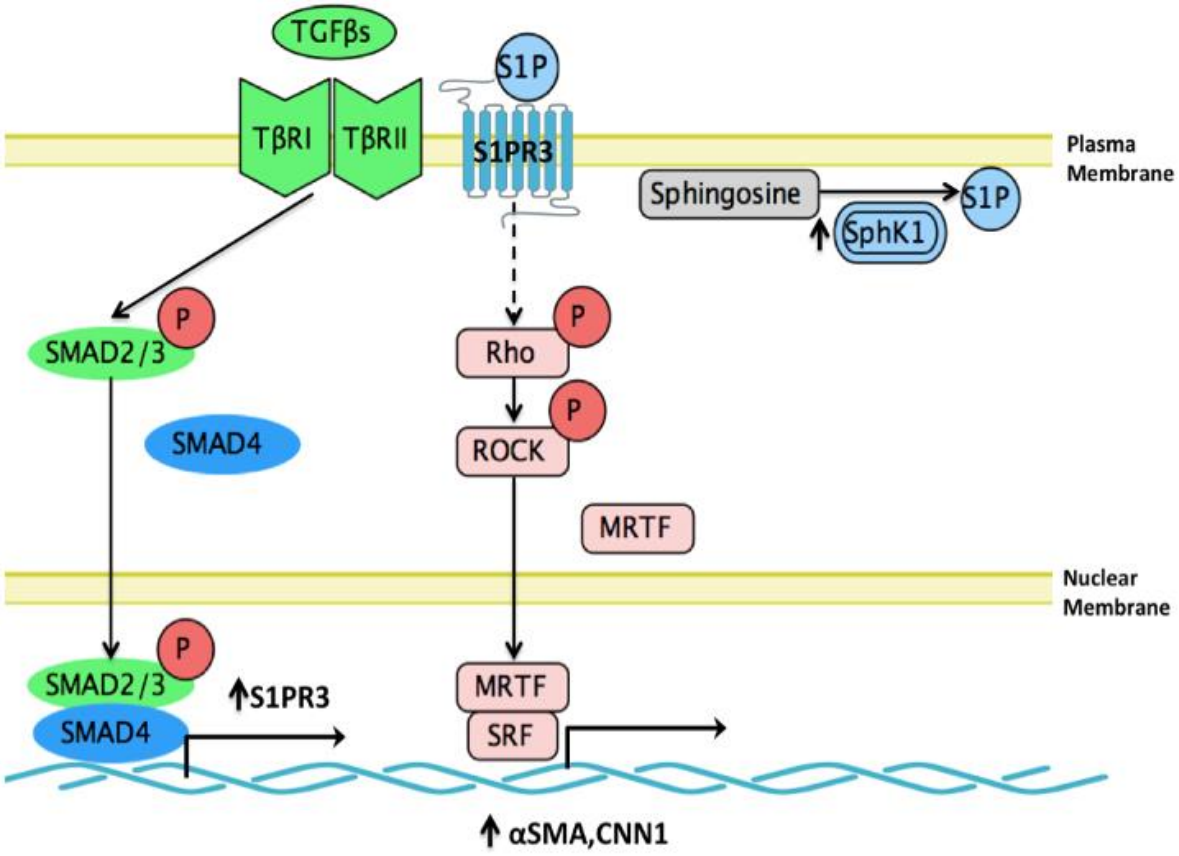


Figure 2.7. TGF-β1 induce SMAD3 phosphorylation with and without treatment of cycloheximide.

HPASMCs were seeded into 6-well plates and allowed to reach confluence before being serum deprived in 0.1% SMGS Medium 231 overnight. HPASMCs were treated with 10-ng/mL cycloheximide or vehicle and were stimulated with 10 ng/mL TGF-β1 for 1, 3, 6 or 9 h. Protein was extracted and the levels of pSMAD3, total SMAD3, and GAPDH were measured. There is an unspecific band right above pSMAD3, which is most likely pSMAD2. Antibody against SMAD2/3 was used and the lower band is SMAD3. The ratio of pSMAD3/SMAD3 is compared. (N=5. Data are represented as mean with SD. Paired t test * P<0.05)

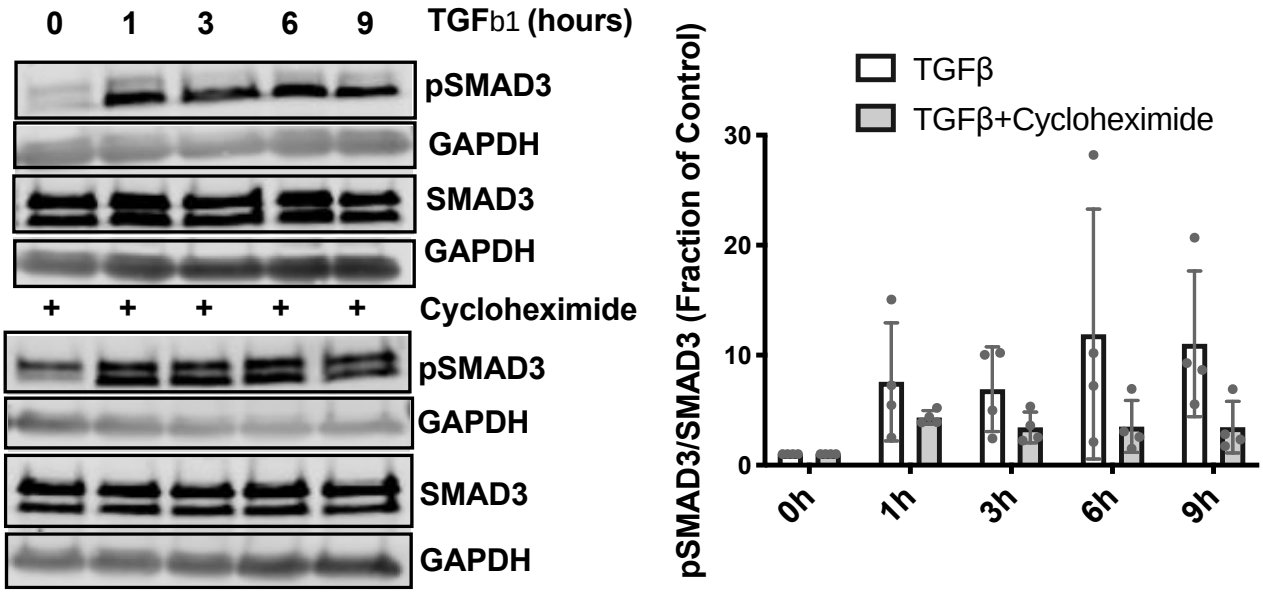


Figure 2.8. S1P increases the αSMA protein level in HPASMC.
HPASMCs were seeded into 6-well plates and allowed to reach confluence before being serum deprived in 0.1% SMGS Medium 231 overnight. HPASMC were then treated with 10 µM S1P or vehicle for 48 h. Protein was extracted and the levels of αSMA and GAPDH were measured. (N=3. Data are represented as mean with SD. Paired t test * P<0.05)

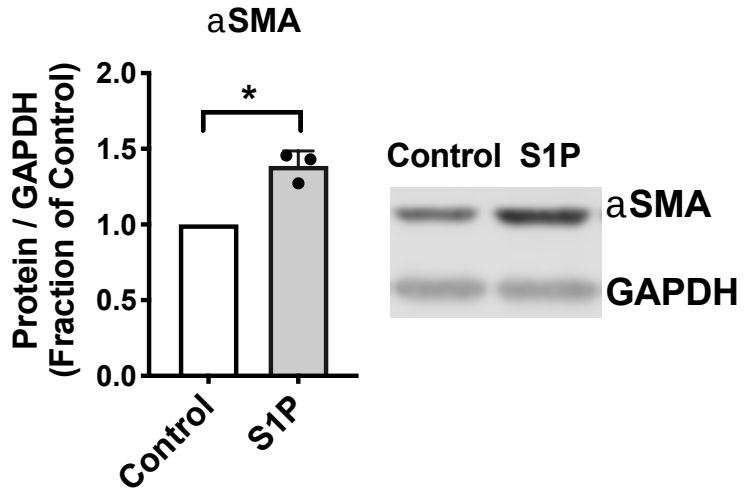


Figure 2.9. TGF-β1 does not induce the proliferation of HPASMC.
HPASMCs were seeded into 96-well plates at a density of 10,000 cells/well. Then cells were serum deprived in 1% SMGS Medium 231 overnight. The cells were then treated with either vehicle or 10 ng/mL TGF-β1 for 2 days. The cells were fixed and stained with DAPI. The Cytation3 was used to image the center of each well and the Image J was used to quantify the number of DAPI-stained nuclei per well. (N=3. Data are represented as Mean with SD. Analyzed by Paired t test.)

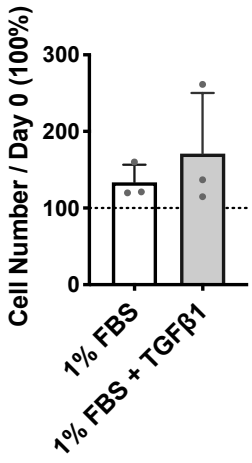


Table 2.1. Primers used for RT-qPCR.

Genes	Forward Sequence (5’ -> 3’)	Reverse Sequence (5’ -> 3’)
GAPDH	AATCCCATCACCATCTTCCAG	AAATGAGCCCCAGCCTTC
αSMA	CTCCCAGGGCTGTTTTCC	CCATGTTCTATCGGGTACTTCAG
CNN1	AACAACTTCATGGACGGCCT	TCTCCAGCTGGTGCCAATTT
SphK1	AGGCTGAAATCTCCTTCACG	CTCCATGAGCCCGTTCAC
S1PR1	CAGCAAATCGGACAATTCCTC	AGCGACCAAGTAAAGAGCG
S1PR2	TTTCTGGAGGGCAACACG	AGTAGCCCCAAGTCTCTATCTG
S1PR3	AGCACTTCAGAATGGGATCTTC	GGTCAAAGTAAGGTAGCTCTCC

CHAPTER 3 : DISSECTING THE INTERACTION BETWEEN TGFβ SIGNALING AND BMPR2 LOSS IN PAH-ASSOCIATED FIBROBLAST

Abstract

Germline mutations in BMPR2, a member of the TGFβ family of receptors, are the most frequent heritable genomic alterations found in PAH patients; all of these mutations are loss of function. In addition to mutations, decreased BMPR2 expression or reduced downstream signaling is also observed. Additionally, the literature suggests that PAH patients have increased TGFβ levels and increased activation of TGFβ downstream signaling. Imbalanced TGFβ and BMP signaling is one of the main drivers of PAH pathogenesis.

Activated fibroblasts, which are myofibroblast-like or proinflammatory, are observed in the adventitia of vessels in PAH patients. These activated fibroblasts showed upregulation of αSMA and increased expression of ECM. They are more proliferative and migratory. They also released chemokines and inflammatory factors, which could recruit macrophages and monocytes, or affect the growth of surroundings cells. Evidence shows that the transition of fibroblasts into myofibroblast-like or proinflammatory cells, could be the initiating factors of the PAH pathogenesis instead of injuries to endothelial cells.

This study aims to test whether TGFβ and the silencing of BMPR2 contribute to the activation of lung fibroblasts. These results show that while BMPR2 was efficiently silenced in human lung fibroblast cells, expression of genes and proteins that are associated with fibroblast activation were not altered. Furthermore, BMPR2 deletion in human lung fibroblasts failed to alter expression of several inflammatory cytokines, even though this was observed in murine fibroblasts. These results could be explained by the differences between cell types or genetic or epigenetic modifications resulting from long term BMPR2 deficiency.

Introduction

Pulmonary arterial hypertension (PAH) is a severe and life-threatening disease characterized by elevated pulmonary blood pressure. PAH can be categorized into different subgroups, the two most prevalent ones are idiopathic PAH (IPAH), which has an unknown cause, and associated pulmonary arterial hypertension (APAH), which is linked to other diseases like scleroderma. While current treatments slow PAH disease progression, they are ultimately not curative. This highlights the need for further investigation of the molecular mechanisms underlying PAH pathogenesis, to ultimately discover new therapeutic targets.

Initially, it was believed that PAH results from vasoconstriction. However only a small portion of PAH patients show high vasoreactivity and responses to vasodilators[10]. PAH pathogenesis is now known to result from pulmonary vascular remodeling, which is characterized by medial hypertrophy/hyperplasia, intimal and adventitial fibrosis, *in situ* thrombosis, and plexiform lesions, as well as peri-vascular infiltration of inflammatory cells. The vasculopathy in PAH involves all three vascular layers (intima, media, adventitia) and a multitude of cell types. The interaction among different cell types is also essential in driving PAH disease progression. For example, injured endothelial cells (ECs) result in an imbalance of vasodilators and vasoconstrictors. This causes increased contraction of smooth muscle cells (SMCs) leading to vascular overconstriction. Extensive studies have been carried out on pulmonary SMCs and ECs. In this study, we focused on the contributions of fibroblasts to PAH pathogenesis, since more and more evidence shows that fibroblasts are not just a bystander in the PAH pathological development.

Traditionally, PAH pathogenesis is characterized as a “inside-out” model, which is initiated with endothelial injuries. However, more recent evidence proposes an “outside in” model for PAH pathogenesis, wherein adventitial fibroblasts serve as injury sensor and the vascular pathology then progresses inward[18, 20]. Following injury and stresses such as hypoxia, resident adventitial fibroblasts are activated and exhibit different functional characteristics that contribute to pulmonary vascular remodeling[21].

In response to various stimuli, activated fibroblasts are believed to transition into a proinflammatory phenotype, which is characterized by high expression of inflammatory cytokines, such as IL-1β, IL-6, and CCL5 (RANTES). These fibroblasts are also responsible for macrophage and monocyte recruitment to the pulmonary vasculature[164]. In PAH patients, inflammatory responses, such as monocyte/macrophage influx, are observed in the adventitia[165]. Additionally, several inflammatory cytokines such as IL-1β and IL-6 are elevated in IPAH and APAH and these cytokine levels are also negatively associated with the survival of PAH patients[166].

Activation of resident fibroblasts results in differentiation into a myofibroblast phenotype[167]. Fibroblast-to-myofibroblast transition is a key process in fibrosis. Myofibroblasts are characterized by increased expression of α smooth muscle actin (αSMA). Early and dramatic increases in the population of αSMA-expressing myofibroblasts in the adventitia is observed in hypoxia-induced pulmonary hypertension[168]. Since myofibroblasts are more proliferative and migratory than un-differentiated fibroblasts, this cell population can migrate into the neointima and contribute to lesion progression[21].

Myofibroblasts are the principal producers of collagen and other ECM proteins including fibronectin and elastin in the remodeled pulmonary adventitia[169]. Increased turnover of ECM proteins and upregulation of matrix metalloproteinase (MMP) activity is observed in the adventitia during the development of pulmonary hypertension[170]. The increased production and altered matrix composition ultimately contributes to further changes in cell growth, behavior, and differentiation. For example, upregulated MMP expression may be necessary for the fibroblast to migrate through the adventitial matrix into the media and intima[171].

In total, these studies demonstrate that activated adventitial fibroblasts play a significant role in influencing the tone and structure of the vascular wall following a variety of injuries or stresses. This occurs both directly, through increasing deposition of ECM, and indirectly by producing chemokines that promote accumulation of leukocytes, or by releasing molecules that serve as mitogenic stimuli for adjacent SMCs or ECs[19].

TGF-β is a cytokine that induces the transition of a fibroblasts into a myofibroblast phenotype by stimulating αSMA expression and collagen production[172]. Bone morphogenetic protein receptor 2 (BMPR2) is a receptor **tyrosine kinase** in the transforming growth factor-β superfamily of receptors. Loss of function mutations in BMPR2 have been reported in up to 80% of HPAH patients. Interestingly, up to 25% of IPAH patients also have abnormal BMPR2 structure or function. Decreased BMPR2 levels in lung tissue are also found in APAH patients. BMPR2 mutation carriers develop PAH 10 years earlier than noncarriers, and these patients tend to have a worse prognosis.

BMPs bind to BMPR2 and induce phosphorylation of SMAD1/5 to regulate gene expression. TGF-βs bind to TGF-β receptors and induce phosphorylation of SMAD2/3, regulating gene expression. Downstream of both BMPs and TGF-βs, there are also non-canonical pathways, such as Rho GTPase. Active Rho signals downstream to Rho-associated kinase (ROCK), which activates and drives myocardin-related transcription factor (MRTF) into the nucleus where it can bind to serum response factor (SRF) and regulate gene transcription. An imbalance in these signaling inputs, wherein TGF-β signaling is overactive and BMP signaling is reduced, is a driver of PAH pathogenesis. It has been reported that loss of BMPR2 protein expression or activity leads to hyperactivated TGF-β signaling and favors endothelial-to-mesenchymal transition (EndoMT) allowing ECs to gain myofibroblast-like characteristics[173]. Similarly, we hypothesized that silencing of BMPR2 in fibroblasts will make them more susceptible to TGF-β-induced fibroblast-to-myofibroblast transition. Another interesting finding is that BMPR2 deletion in cancer-associated fibroblasts upregulates several inflammatory genes[174]. Thus, we propose that BMPR2 silencing could drive fibroblasts into a proinflammatory state in PAH. The goal of this study is to investigate the role of BMPR2 and TGFβ on promoting the fibroblast-to-myofibroblast transition, with a focus on the myofibroblast-associated proinflammatory phenotype. Ultimately, these findings could help explain the pathophysiological role of BMPR2 deficiency during PAH pathogenesis.

Methods

Cell Culture

The human WI 38 Cell Line (MilliporeSigma, Burlington MA, Waltham, #90020107) was cultured in Eagle's Minimum Essential Medium (EMEM) (ATCC, Manassas VA, #30-2003; ThermoFisher, Waltham MA, #670086) supplemented with 10% Tet free fetal bovine serum (FBS) (Thermofisher) and 5% Anti-anti (Thermofisher). HEK-293T cells is cultured in Dulbecco's Modified Eagle Medium (DMEM) (Thermofisher) with 10% fetal bovine serum (FBS) (Thermofisher) and 5% Anti-anti (Thermofisher).

Compounds and Antibodies

Recombinant human TGF-β1 protein was purchased from Research And Diagnostic Systems, Inc.(Minneapolis, MN). Antibodies against SMAD2/3 (#8685), pMLC2 (#3674) and MLC2 (#3672) were ordered from Cell Signaling (Danvers, MA). pSMAD3 antibody (#52903) was purchased from Abcam (Cambridge, MA). BMPR2 antibody (19087-1-AP), αSMA antibody (55135-1-AP) and GAPDH antibody (60004-1-Ig) was purchased from Proteintech (Rosemont, IL). All secondary antibodies [Donkey anti-Mouse800 (#C90507-03), Donkey anti-Rabbit680 (#C40130-02)] are all purchased from LI-COR (Lincoln, NE).

Lentivirus Cloning

All the shRNA sequences were cloned into the Tet-pLKO-puro vector (from Dmitri Wiederschain, Addgene plasmid # 21915) following the protocol listed in the Tet-pLKO Manual on the Addgene website. The presence of the correct insert in the final plasmid was confirmed

by Sanger sequencing. The targeted sequences of shRNAs are as follows: shControl (shLacZ): CGCTAAATACTGGCAGGCGTT, shBMPR2-1: GCCTATGGAGTGAAATTATTT, shBMPR2-2: CCCTCTCTTGATCTAGATAAT.

Virus Preparation and Infection

HEK-293T cells were seeded into 10-cm plates and were allowed to attach overnight. The next day at approximately 60-70% confluence, the cells were transfected with a plasmid cocktail containing 10 µg of the Tet-pLKO-puro plasmid, 3.75 µg of psPAX2 (Addgene plasmid #12260), 1.25 µg of pMD2.G (Addgene plasmid #12259), and 20 µL of Lipofectamine 2000 in 400 µL of OptiMEM. The next morning the medium was changed to 10 mL of fresh culture medium, and the next day each plate was changed with 10 mL of fresh culture medium. After 24 h, the culture medium was harvested and filtered through a 0.45- micron syringe filter. Virus was stored at 4 C and was used within a week.

WI38 fibroblasts were seeded into 10-cm plates and were allowed to attach overnight. The next day at approximately 70% confluence the medium was changed to 10-mL of fresh EMEM and was supplemented with 5 mL of viral supernatant. The next morning, the medium was changed, and the cells were incubated an additional 24 h. The cells were then treated with 2.5 µg/mL puromycin until all of the untransformed cells on the kill control plate died (approximately 72 h). We did not pick individual clones for the cell lines, but instead used a pooled infection approach. Validation of the silencing efficiency was tested by qPCR and immunoblotting to detect BMPR2 mRNA and protein levels, respectively.

For experiments, WI38 were re-suspended in complete medium and 250,000 cells were seeded in each well of a 6-well plate. The cells were allowed to reach confluence before being serum-starved in 1% FBS EMEM with 1 ng/µL Doxycycline overnight. The next day, the medium was changed to fresh 1% FBS EMEM with 1 ng/µL Doxycycline with or without TGF-β treatment, as described in the Figure legends.

RT-qPCR

Cells were treated as described in the figure legends and total cellular RNA was collected using the RNeasy kit (Qiagen, Hilden, Germany #74104) according to the manufacturer's protocol. The High-Capacity cDNA RT kit (ThermoFisher #4368814) was used to reverse transcribe the RNA into cDNA following the manufacturer's protocol. SYBR Green PCR Master Mix (ThermoFisher #4309155) was used to perform qPCR following the manufacturer's protocol on the QuantStudio 7 Flex Real-Time PCR System qPCR machine. Fold-change of gene expression was normalized to GAPDH and analyzed by the $2^{-\Delta\Delta CT}$ method. Primer sequences are listed in **Table 3.1**.

Immunoblotting

Total cellular protein was collected in 2x Laemmli Sample Buffer (Biorad, Hercules, CA #1610737). After heating the samples at 100 ^{o}C for 10 mins, protein samples were resolved on 4-20% MINI-PROTEAN TGX GELs (Bio-Rad, Hercules, CA, #4561094) and transferred to PVDF Membranes (Millipore Sigma, Burlington, MA #IPFL00010). Blots were blocked in Odyssey Blocking buffer in PBS (LI-COR # 927-40000) at room temperature for 1 h, then incubated with primary antibody at room temperature for 1 h or overnight at 4 ^{o}C. Blots were washed three

times for 5 min each with Tris-buffered saline with 0.1% Tween-20 (TBST) and then incubated with the appropriate secondary antibodies diluted 1:10,000 in blocking buffer at room temperature for 1 h. After 3 washes for 5 mins each with TBST, blots were imaged using a LI-COR Odyssey FC instrument and analyzed using Image Studio Lite software v.4.0.

Statistical Analysis

Data were analyzed by a Paired t-test (for two sample comparisons) using GraphPad Prism 7. Data are presented as the mean ± SD, and $P<0.05$ is considered statistically significant.

Results

Confirming Knockdown efficiency of BMPR2

Since our goal was to assess the role of BMPR2 in lung fibroblasts, we first sought to use small hairpin RNA to generate stable BMPR2 deleted cell lines. We were able to generate these cells, but stable deletion of BMPR2 appeared to confer a growth disadvantage to the cells. Thus we were unable to grow the cells long term. As an alternative approach, we opted to use an inducible shRNA system wherein the cells have stable integration of the shRNA, and expression of the shRNA and subsequent gene silencing is under control of a doxycycline-inducible promoter. Expression of two shBMPR2 hairpins and one shControl (shLacZ) hairpin was induced with 1 ng/µL doxycycline for 48 h prior to sample collection and assessment of BMPR2 expression. In some experimental settings, cells were also treated with 10 ng/µL TGF-β for 6 hours or 24 hours concurrent with the final 6 or 24 hours of doxycycline treatment.

To assess knockout efficiency, we measured BMPR2 mRNA and protein levels with RT-qPCR and immunoblotting, respectively. Without the Doxycycline treatment, fibroblasts transduced with shLacZ or shBMPR2 showed similar expression levels of BMPR2 with or without stimulation of TGF-β. When the shRNA was induced by Doxycycline, cells expressing shBMPR2 had about a 75% reduction in BMPR2 expression at the mRNA and protein levels at 48 h. This indicated that our inducible shRNA system efficiently depleted BMPR2 expression.

BMPR2 Silencing Did Not Affect TGF-β Induced Phosphorylation of SMAD3 and MLC2 or mRNA expression of αSMA

TGF-β stimulates the fibroblast-to-myofibroblast transition, which is associated with increased αSMA expression. TGF-β1-induced αSMA upregulation is dependent on downstream SMAD and Rho/MRTF/SRF activation in SMCs (Chapter 2). Activation of these two pathways can be evaluated by measuring the phosphorylation state of SMAD3 or MLC2, respectively. To test the effect of BMPR2 silencing on TGF-β1-induced MLC2 and SMAD3 phosphorylation, or expression of αSMA at the mRNA level, we used our inducible shBMPR2 fibroblasts which were treated with 10 ng/ μL TGF-β1 for 6 hours to measure phosphorylation of SMAD3 and MLC2 and 24 hours to measure mRNA expression of αSMA. We found that while TGF-β1 stimulates the phosphorylation of SMAD3 and MLC2, this did not seem to be affected by BMPR2 silencing. Similarly, we found that BMPR2 silencing does not potentiate TGF-β1-induced αSMA mRNA upregulation. In the absence of TGF-β1, BMPR2 silencing did not affect fibroblast-to-myofibroblast transition marker, αSMA in WI38.

BMPR2 Silencing Did Not Alter Cytokine Expression

Cancer-associated fibroblasts with BMPR2 deletion have increased expression of multiple cytokines, including CCL5 and CXCL10[174]. This led us to hypothesize that reduced BMPR2 activity in lung fibroblasts would have the same effect, and thus could help explain the pro-inflammatory phenotype observed in PAH patients. In our experiment, we measured CCL5 and CXCL10 expression since these two proteins were upregulated in CAFs, and we also measured IL-1β and IL-6 expression since these proteins are elevated in PAH patients. While we found that TGF-β1 treatment induced IL-1β mRNA and decreased CXCL10 mRNA, we did not observe any effect from BMPR2 silencing.

Discussion

While we successfully silenced BMPR2 in WI38 fibroblasts, BMPR2 silencing did not affect TGF-β induced phosphorylation of SMAD3. Increased phosphorylation of SMAD3 is found in PAH patients, but it is possible that this could be driven by increased TGF-β levels instead of increased TGF-β receptor signaling due to BMPR2 dysfunction. TGF-β-induced SMAD3 phosphorylation is increased in BMPR2-mutant mouse primary PASMCs compared to control[175]. Similarly, increased phosphorylation of SMAD2 in endothelial cells with BMPR2 deletion is also observed after TGF-β stimulation[173]. Since my studies were in fibroblasts, not ECs or SMCs, cell type differences may explain these discrepancies. This may also explain why the gatekeeper function of BMPR2 for EndoMT in ECs is not replicated in our studies.

The interaction of BMPR2 silencing and non-SMAD signaling is not well characterized. Studies showed that upregulation of BMPR2 with gene delivery can modulate non-SMAD signaling such as increased phosphoinositide-3 kinase (PI3K) signaling and decreasing p38-mitogen activated protein kinase (p38MAPK) *in vivo*[176]. However, the Rho/MRTF/SRF pathway might not be affected by BMPR2 deficiency in a similar manner.

Surprisingly, we did not observe increased cytokine expression in BMPR2 deficient cells either. Since CAFs are isolated from mice, species differences could play a role. Additionally, the PASMCs which had greater SMAD3 phosphorylation are also isolated from WT or BMPR2-mutant BMPR2 mice. For the ECs where BMPR2 acts as gatekeeper for EndoMT, BMPR2 is deleted or mutated by CRISPR/Cas9. All of these experiments were carried out at passage 7-40 after clonal selection[173], which means that these knockout cells were grown in culture for a prolonged period of time. The suggests that long-term culture may result in the epigenetic and signaling pathway re-wiring. In agreement with this hypothesis, literature evidence suggests that cells isolated from PAH patients or animal models maintain pathological features such as increased proliferation, anti-apoptosis, and altered metabolism in vitro. This could suggest that these abnormal phenotypes do not rely on specific factors within the PAH microenvironment but are rather driven by cell intrinsic alterations on a genetic or epigenetic level.

For our inducible BMPR2 silencing system, BMPR2 is only depleted for 1-2 days prior to performing phenotypic assays, and these timescales are likely too short to permit any significant epigenetic or transcriptomic re-wiring. This could also be a possible explanation as to why we could not replicate the findings found in other cell types. To further validate this hypothesis, it will be necessary to silence BMPR2 over a longer timescale and/or to add TGF-β well after BMPR2

knockdown is established. We could then compare those effects with short-term BMPR2 silencing. A primary goal would be to test whether the long term BMPR2 silencing could replicate the findings observed in other cell types upon BMPR2 silencing or deletion. Second, the comparison between short term and long term BMPR2 silencing may help us to identify signaling pathways which are rewired or possible genetic or epigenetic modifications within the cells.

Figure 3.1. Test the Efficiency of BMPR2 Silencing.
WI-38 cells were seeded into 6-well plates and were culture until they were confluent. Once the cells reached confluence the cells were serum deprived in 1% FBS EMEM and were induced with 1 ng/mL doxycycline for 24 h. The cells were cultured for an additional 24 h in fresh 1% FBS EMEM and 1 ng/mL doxycycline. During this final 6 or 24 h of induction, the cells were concurrently treated with 10 ng/mL TGF-β1 for 6 or 24 h, respectively. Protein and mRNA levels of BMPR2 and GAPDH were analyzed by immunoblotting and RT-qPCR, respectively. Data are represented as mean -/+ SD. N >= 3 biological replicates. Paired t-test where * indicates statistically significant p-values < 0.05.

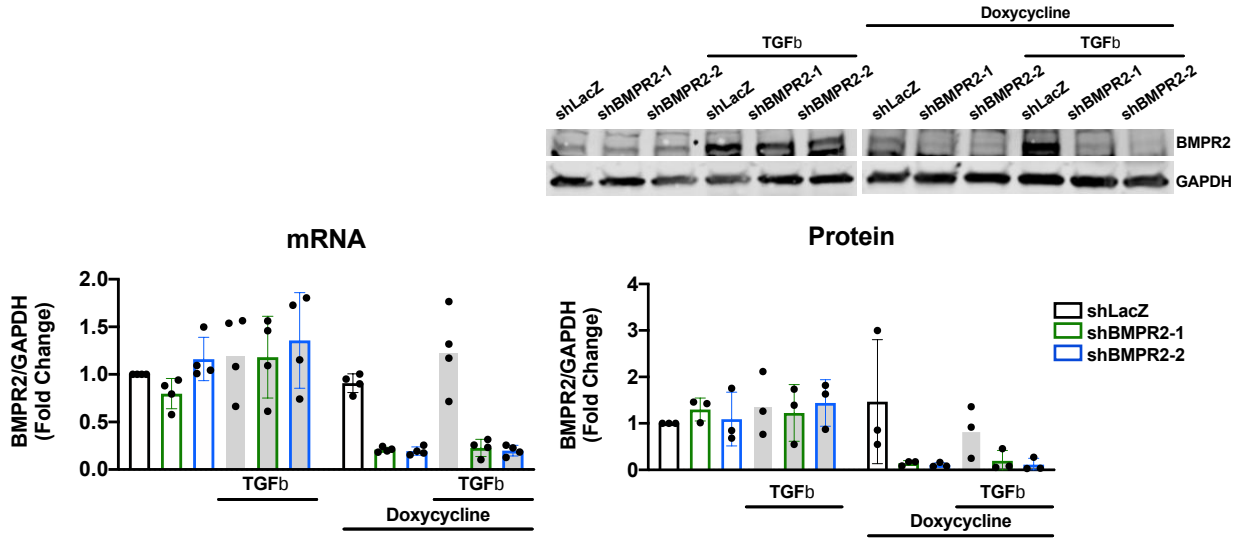


Figure 3.2. BMPR2 silencing did not affect the activity of TGF-β1 downstream signaling.
WI-38 cells were seeded into 6-well plates and were culture until they were confluent. Once the cells reached confluence the cells were serum deprived in 1% FBS EMEM and were induced with 1 ng/mL doxycycline for 24 h. The cells were cultured for an additional 24 h in fresh 1% FBS EMEM and 1 ng/mL doxycycline. During this final 6 or 24 h of induction, the cells were concurrently treated with 10 ng/mL TGF-β1 for 6 or 24 h, respectively. Protein levels of phosphorylated SMAD3, SMAD3, phosphorylated MLC2, MLC2 and GAPDH were analyzed by immunoblotting. mRNA levels of αSMA were measured by RT-qPCR. Data are represented as mean -/+ SD. N >= 1 biological replicates. Paired t-test where * indicates statistically significant p-values < 0.05.

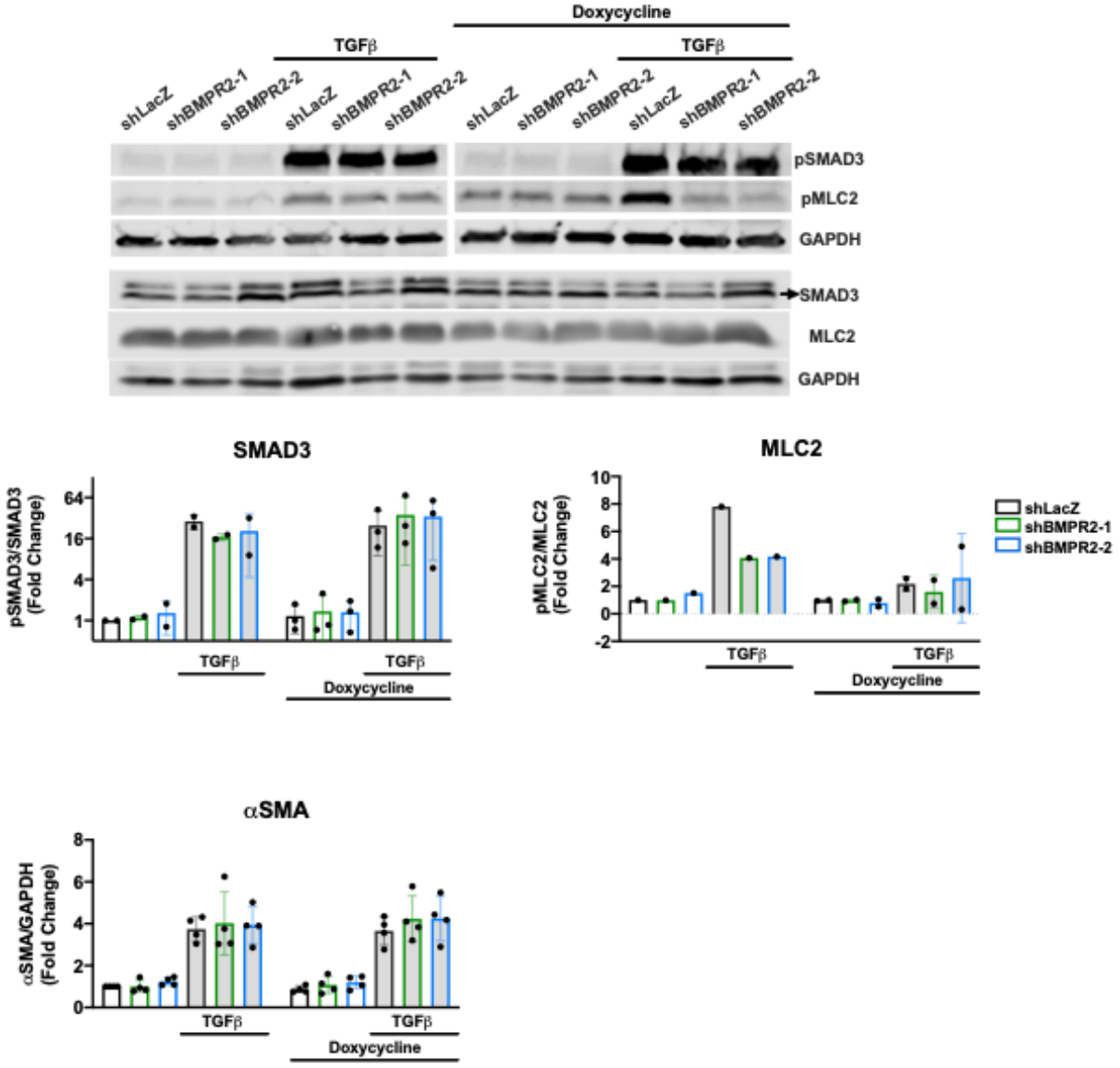


Figure 3.3. BMPR2 silencing did not change the mRNA level of cytokines in fibroblasts.
WI-38 cells were seeded into 6-well plates and were culture until they were confluent. Once the cells reached confluence the cells were serum deprived in 1% FBS EMEM and were induced with 1 ng/mL doxycycline for 24 h. The cells were cultured for an additional 24 h in fresh 1% FBS EMEM and 1 ng/mL doxycycline. During this final 6 or 24 h of induction, the cells were concurrently treated with 10 ng/mL TGF-β1 for 6 or 24 h, respectively. mRNA levels of ILβ -1, IL-6, CCL5 and CXCL10 were measured by RT-qPCR. Data are represented as mean -/+ SD. N = 4 biological replicates. Paired t-test where * indicates statistically significant p-values < 0.05.

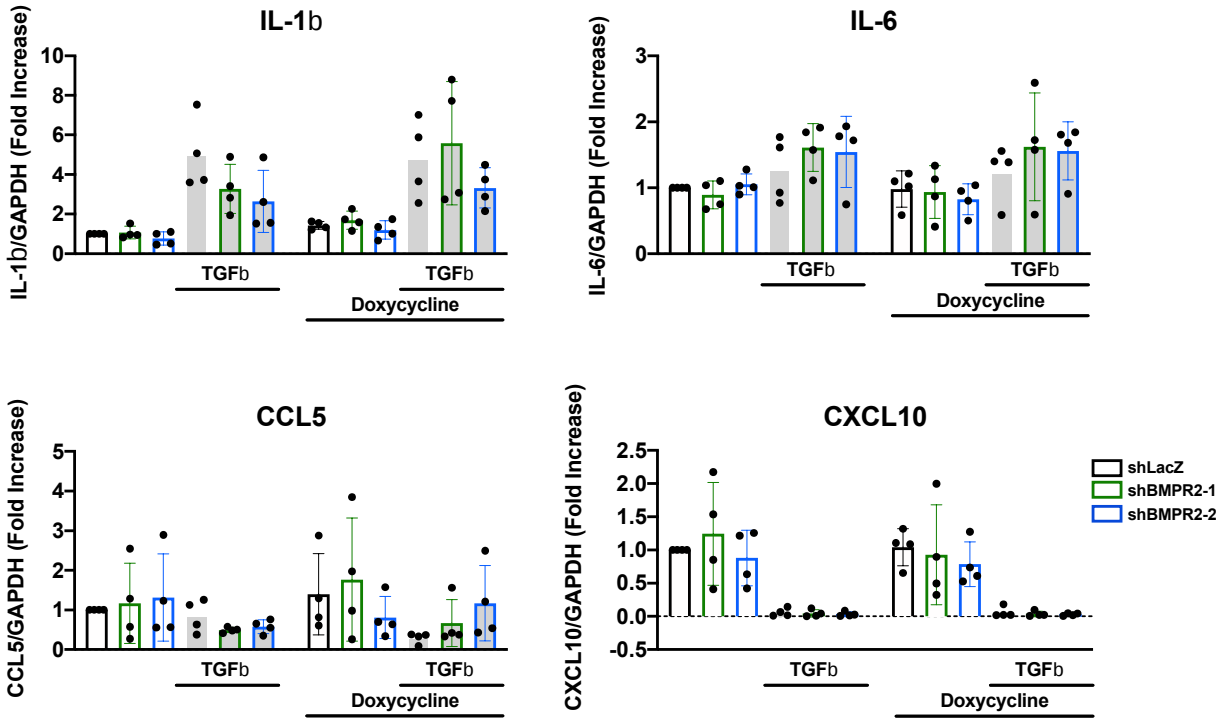


Table 3.1. Primers for qPCR.

Genes	Forward Sequence (5’ -> 3’)	Reverse Sequence (5’ -> 3’)
GAPDH	AATCCCATCACCATCTTCCAG	AAATGAGCCCCAGCCTTC
αSMA	CTCCCAGGGCTGTTTTCC	CCATGTTCTATCGGGTACTTCAG
BMPR2	GGCTGACTGGAAATAGACTGG	GGCTGACTGGAAATAGACTGG
IL-1β	ATGCACCTGTACGATCACTG	ACAAAGGACATGGAGAACACC
IL-6	CCACTCACCTCTTCAGAACG	CATCTTTGGAAGGTTCAGGTTG
CCL5	TGCCCACATCAAGGAGTATTTC	CCATCCTAGCTCATCTCCAAAG

CHAPTER 4 : TRANSCRIPTOMIC CHARACTERIZATION OF PULMONARY ARTERIAL HYPERTENSION AT THE PATHWAY LEVEL

Abstract

Pulmonary arterial hypertension (PAH) is a severe and life-threatening disease characterized by elevated pulmonary blood pressure. PAH can be categorized into different subgroups, which include idiopathic PAH (IPAH) and PAH that is associated with other diseases (APAH). While current treatments slow PAH disease progression, they are ultimately not curative. This highlights the need for further investigation into the molecular mechanisms underlying PAH pathogenesis, to ultimately discover new therapeutic targets.

There have been several studies which have transcriptionally profiled and compared control and PAH samples using either gene expression microarrays or RNA-Seq. One challenge is that each of these independent studies analyzed only a small number of patient samples, which limits the analytical methods that can be performed. To circumvent this problem, we have pooled multiple publicly available PAH gene expression datasets, which enables us to use more sophisticated computational techniques. While the published studies where the datasets were derived, focused on studying differential expression at a single-gene level, in this study we leveraged gene set-based computational techniques to study the transcriptional alterations in PAH on a pathway level.

We have identified differentially expressed genes and gene signatures between control and PAH samples. Many of them were associated with signaling pathways that have an established role in PAH biology. But interestingly, we were also able to identify new genes and pathways that have not been previously associated with PAH, and thus warrant further experimental validation. This study also compared IPAH with APAH samples with the goal of identifying subtype-specific transcriptional signatures. We also performed an unbiased clustering

approach to identify PAH subgroups based on their transcriptomic signatures. Excitingly, the samples did not cluster based upon their clinical subtype, but rather into several distinct expression-based clusters, suggesting that same PAH subtype patients with different transcriptional profiles may have divergent drug responses. These studies have identified several signaling pathways which warrant further experimental validation and investigation.

Introduction

Pulmonary arterial hypertension (PAH) is a disease characterized by elevated mean pulmonary arterial pressure (mPAP ≥20 mmHg)[1]. PAH is rare with an estimated incidence rate of 5-15 cases per one million people[2]. The pathology in the pulmonary arteries of PAH patients is characterized by vasoconstriction, uncontrolled proliferation, fibrosis and thrombosis, which are consistently observed in PAH patients[8]. Despite similar clinical presentations, PAH patients can be subdivided into 5 distinct subgroups based on their genetic backgrounds or other associated diseases that may contribute to causing the disease. Idiopathic PAH (IPAH), the most common subtype, is characterized by lack of family history of PAH or known risk factors for developing PAH[2]. The heritable PAH (HPAH) subgroup includes patients with familial germline alterations, such as alterations in the BMPR2 gene, that convey increased risk of developing PAH[2]. The associated PAH (APAH) subgroup is associated with other diseases such as Scleroderma or HIV[2]. The final two PAH subgroups are PAH induced by drugs or toxins and PAH with pulmonary veno-occlusive disease and/or pulmonary capillary hemangiomatosis[2].

Current PAH medications mainly target the nitric oxide pathway, endothelin signaling, and prostacyclin signaling[31]. Despite these medications, PAH remains a lethal disease without a cure. The 5-year survival of PAH patients is only 50-60%[177]. One reason for this is that current PAH medications work mostly by reducing vasoconstriction but do not alter the vascular remodeling that results from over-proliferation and anti-apoptosis of cells within the pulmonary vasculature[31]. While current treatment strategies are effective in relieving symptoms, in order to further halt the pathologic progression of PAH, drugs are needed that better target the underlying pathological mechanisms of PAH, especially the vascular remodeling.

The pathogenesis of PAH involves multiple cell types and is driven by a complicated molecular signaling network. The development of microarray and RNA sequencing technology and sophisticated data analysis techniques during the last few decades have made it possible to study transcriptional alterations during PAH pathogenesis on a global level[120, 122, 132, 133]. There have been many transcriptomic studies using PAH tissues, compared to other diseases, such as cancer which has been studied in detail. Interestingly, the unique PAH vascular pathology, plexiform lesions, is characterized by hyperproliferative and apoptosis resistant endothelial cells (ECs) and smooth muscle cells (SMCs) [26]. The vascular abnormalities in PAH patients mimic some phenotypes found in tumors, including cellular overgrowth, apoptosis resistance, inflammation, and angiogenesis[26]. These similarities provide us with the unique opportunity to leverage existing bioinformatic resources developed to study cancer to analyze transcriptomic data from PAH patients.

Pulmonary tissue is anatomically difficult to access, which makes it difficult to acquire a large cohort of patient samples for transcriptomic analysis. Another problem, partly due to the

limited availability, is significant heterogeneity among these small sequencing cohorts[136]. This makes it difficult to dissect the positive findings from the noise caused by confounder effects. To circumvent this problem, we compiled and harmonized a meta-dataset from multiple publicly available PAH gene expression datasets. This new combined dataset contains transcriptomic profiles from over 100 PAH patients. Another constraint of the existing bioinformatic studies on these samples is that most of studies focused analysis on a single-gene level. In this study, we leveraged gene-sets based computational techniques[137]. Gene sets are lists of genes with correlated expression found in canonical pathways or in experimental signatures from publications[137]. Using these approaches, we studied transcriptional alterations in PAH on a pathway level, rather than on a single gene level[137]. The goals of this study are to provide new insights into the transcriptional alterations underlying PAH pathogenesis, define transcription-related subgroups, and ultimately to identify new therapeutic targets and biomarkers.

Methods

Datasets description and preprocessing

Gene expression data of PAH samples was obtained from the NCBI-GEO database (http://www.ncbi.nlm.nih.gov/geo/). We used four GEO series: GSE15197, GSE48149, GSE113439 and GSE117261 (**Table 4.4**). For GSE15197, we only analyzed the control and PAH samples, and discarded the pulmonary fibrosis samples prior to any analysis. For GSE48149, we only analyzed the normal, IPAH, and scleroderma-related PAH samples. The microarray data were Robust Multi-array Average (RMA)-normalized and log-transformed using Transcriptome Analysis software (Thermofisher). One confounder with the samples in the GSE117261 dataset is

that the samples in this study were collected and analyzed at different times, which resulted in a large batch effect within this sequencing study. To correct this, we used the "combat" algorithm[178] which uses an Empirical Bayes method to correct for batch effects. When all of the datasets were merged and filtered to include only genes shared among all 4 studies, significant inter-study batch effects which segregated samples were also observed and corrected by "combat". In order to check the variances among samples before and after "combat" processing, we conducted principle component analysis (PCA) to reduce the dimensionality of the datasets and visualize the variability among samples (**Figure 4.6A**). Since GSE15197 did not provide much clinical information, including PAH subgroups, for the included samples, only GSE48149, GSE113439 and GSE117261 were used for the comparison of IPAH and APAH and for the molecular subgroup analysis. For these analyses a separate dataset that contained only the relevant samples from these three studies was constructed, merged, combat-normalized, and PCA-tested (**Figure 4.6B**) prior to downstream analysis.

Gene set score generation

To analyze the transcriptomic data on a pathway activation level, we first collapsed the gene-level data to pathway-level data using single sample Gene Set Enrichment Analysis (ssGSEA) using the pre-computed pathway activation gene sets in MSigDB database v7.2 (https://www.gsea-msigdb.org/gsea/msigdb). We focused our analysis on only the most informative groups of gene sets, which included Hallmark (H), C2-Canonical Pathways (C2CP) and Oncogenic Signatures (C6). In some cases, there is a significant overlap in the genes contained in two or more signatures, which might conceal other essential signatures in differential-expression

analysis and result in false correlation in cluster-based analysis. We used the R package “ReCiPa” (Redundancy Control in Pathway Databases)[179] to eliminate highly redundant gene sets and only include one representative gene set from each cluster of overlapping gene sets for the subsequent analysis. We set the maximum overlap at 80% and minimum overlap at 10% to merge similar datasets that fall within this threshold. As a final step, gene sets were z-score normalized.

Identification of differentially expressed genes and gene sets

We used R package “limma” to conduct differential expression analysis on individual genes and gene signature scores. Genes and gene sets with a fold discovery rate (FDR)-corrected p-value < 0.01 and absolute log fold change (logFC) >= 0.6 during the analysis were considered to be differentially expressed genes (DEGs) and differentially expressed gene sets (DEGSs). Volcano plots of FDR vs logFC were generated in GraphPad Prism (v.7) to highlight the DEGs. We also used the R package “ComplexHeatmap” to illustrate the DEGs and DEGSs across samples with unsupervised hierarchical clustering on the genes or gene sets.

Differential expression analysis was performed on the four combined datasets to compare control and PAH samples, and on the three combined datasets to compare control and IPAH samples, and control and APAH samples separately. For the analysis on PAH subgroups, DEGs that were either shared between these two analyses, or were unique to one of the two analyses were further characterized to investigate the similarities and differences between IPAH and APAH.

Further Analysis of DEGs

In order to interpret the biological functions of the DEGs, we used the String (https://string-db.org/) to generate protein-protein interaction (PPI) networks. With the list of DEGs, we used MSigDB tools to detect possible signaling pathways that they might be involved in with FDR < 0.05. We also used Connectivity Map (CMap) (https://clue.io/cmap) to identify compounds or CMap class that show the opposite expression signatures from the PAH-associated DEGs. This may help uncover novel treatments for PAH by identifying drugs that could reverse pathological gene expression in patients.

Classifing PAH samples based on their transcriptomic profiles

We used the normalized gene signature scores of the PAH samples from the combined three datasets to study the variation among them. We used all gene signatures from the Hallmark, Oncogenic Signatures (C6) and the top 117 ones from C2CP with the greatest relative standard deviation (RSD). We used the R package "ComplexHeatmap" to conduct unsupervised hierarchical clustering of both the samples and the gene sets to identify molecular subgroups within the PAH patients. We also annotated the PAH samples with specific clinical subgroup information to see if they are correlated with the molecular subgroups.

Results

Significant DEGs between Control and PAH

In order to investigate the differential expression between Control and PAH samples I combined the four datasets to have a single normalized and batch corrected meta-dataset with

a total of 58 Control and 109 PAH samples. We first used the R package “limma” to detect the DEGs, which we classified as having an adjusted p-value < 0.01 and a |logFC| >= 0.6 (**Figure 4.1A**). We identified 45 downregulated genes and 34 upregulated genes for subsequent analysis (**Figure 4.1B**). The first question we asked was whether we actually gained any new biological information by merging multiple PAH datasets together. To do this, we re-analyzed each individual dataset with the same protocol, then compared the results with our combined analysis. Interestingly, in multiple cases, the individual datasets had hundreds or thousands of differentially expressed genes. This could be a real biological difference, noise due to small data sets, or possibly experimental confounders. Interestingly, the 6 differentially expressed genes that were identified in the combined analysis were not identified in any of the single-dataset analyses. Of these six genes four of them (RPA4, CALB2, HBD, and TFPI2) were upregulated and two of them (FGFBP1 and KRT4) were downregulated. Interestingly, HBD, a hemoglobin subunit, had previously been identified as a “hub gene node” in a transcriptomic analysis of PBMCs isolated from control and IPAH patients[180]. This transcriptomic dataset was not included in our analysis since it analyzed blood and not lung tissue samples. Our combined analysis is beneficial from a discovery standpoint since it identified PAH-associated genes not uncovered in any single dataset. It is also exciting that one of the novel DEGs is also identified essential in another PAH study.

In order to study the biological function of the core 79 differentially expressed genes we determine the possible associations among these core genes by leveraging biological information of known or predicted protein-protein and gene-gene interactions using String. Most of the 79 genes were associated with the main network (**Figure 4.1C**), suggesting that there is a high degree

of interaction between all of the significant genes identified in our meta-analysis. Some of the singleton genes that were not associated with the network were still interesting and warrant further investigation. For example, SNPs near PDE1A are associated with IPAH in a GWAS analysis[181], raising the possibility that altered PDE1A expression, either through an eQTL or another molecular mechanism, may be important for PAH pathogenesis. Furthermore, the DEGs PDE7B and PDE8B are also members of the phosphodiesterase family, which is one of the main therapeutic targets in PAH. One caveat in this analysis is that it's unclear what the drug treatment status is for the patients included in this study, so we are unable to rule out drug-induced expression change in the PAH samples. WIF1 and SFRP2 are upregulated genes in PAH in my data and are associated with WNT signaling, which has been implicated in PAH pathogenesis[101]. To further analyze the genes that are associated with the main network, we first computed the overlap between this subset of genes and a set of pre-computed gene signatures in MSigDB (**Figure 4.1D**). We focused our analysis on the gene sets in the Hallmark, C2CP and C6 groups of signatures, since these groups contain the most biologically relevant signatures for our analysis. Of the 66 gene sets with an overlap FDR < 0.05, a significant number of the gene sets were involved in the inflammatory responses. For example, "HALLMARK_IL6_JAK_STAT3_SIGNALING" is consistent with the finding that the STAT3 activation has been implicated in PAH pathogenesis and is a potential therapeutic target[182]. "HALLMARK_ EPITHELIAL_ MESENCHYMAL_ TRANSITION" was also identified, and is interesting since EMT in epithelial cells is biologically similar to the fibroblast-to-myofibroblast transition, which ultimately leads to elevated lung fibrosis which is often observed in vascular adventitia of PAH[11]. It is also analogous to endothelial mesenchymal transition (EndoMT), which is implicated in PAH pathogenesis[183].

Overall, our DEG findings successfully validate some of the signaling pathways important for PAH pathogenesis and also provide some new genes for further validation and follow-up.

Pathway-level DEGSs between Control and PAH

Single sample gene set enrichment analysis (ssGSEA) is an analytical approach wherein single-gene level data is collapsed into pathway-level data by grouping genes based on *a priori* defined gene sets that represent specific signaling pathways. This approach is ideal for this study because it allows signatures to be projected onto each individual sample which is not affected by the grouping of samples. This approach also generates a more reliable dataset since it integrates data from all measured genes, rather than just relying on a subset of statistically significant genes, or on any one individual gene of interest. As a first pass analysis we sought to identify differentially activated genesets, herein referred to as differentially expressed gene sets (DEGSs), between control and PAH samples. The cutoff we employed for DEGSs was the same as the one we employed for single-gene analysis with a p-value < 0.01 and a |logFC| >= 0.6. Most of the DEGSs identified were in the C2 collections, which makes sense since they contained the majority of the total gene sets in our analysis. In order to interpret the DEGSs from C2, these gene sets were first clustered with unsupervised hierarchical clustering to identify gene sets which best segregate different sample groups (**Figure 4.2**). The DEGSs from hallmark were just listed in **Table 4.1**. First, we sought to determine whether this analysis was able to re-capture known biological information about PAH. Previous studies have identified key signaling pathways that are upregulated in lung tissues from PAH patients, which include TGFβ signaling and WNT signaling[101, 175]. Several GPCR-related signatures were also upregulated in PAH in this analysis,

including “PID_S1P_S1P3_PATHWAY” and “PID_INTEGRIN5_PATHWAY”. The observation that the “PID_S1P_S1P3_PATHWAY” signature is upregulated in PAH samples agrees with several recent studies demonstrating that the importance of Sphingosine-1-phosphate (S1P) signaling in PAH disease development and its potential as a therapeutic target[153]. Rho signaling is downstream of TGFβ and GPCR signaling and “REACTOME_RHO_GTPASES_ACTIVATE_CIT” is also upregulated in the PAH samples. In line with the importance of the Rho signaling pathway in PAH, the Rho-Associated Kinase (ROCK) inhibitor Fasudil was previously tested clinically in PAH patients, and next-generation ROCK inhibitors are now under development[31]. Other therapeutic targets of clinical interest include various HDAC proteins[31]. Accordingly, we also found that the “BIOCARTA_HDAC_PATHWAY” signature is upregulated in the PAH samples. Finally, YAP/TAZ signaling has been implicated in PAH pathogenesis[184], and a YAP/TAZ gene set was upregulated in the PAH samples.

In total, these findings suggest that this analysis is able to capture known PAH biology on a signaling pathway level. But ultimately, we wanted to determine whether this analysis was able to identify new therapeutic targets for treating PAH. The best example of this was the finding that several metabolism-associated signaling pathways were downregulated in PAH samples. These pathways included“KEGG_ PENTOSE_ PHOSPHATE_ PATHWAY”, “KEGG_ OXIDATIVE_ PHOSPHORYLATION”, “KEGG_ GLYCOLYSIS_ GLUCONEOGENESIS” and “WP_ CHOLESTEROL_ BIOSYNTHESIS_ PATHWA”. This is potentially interesting since, as in tumors, the Warburg effect is also observed in the lungs of PAH patients, suggesting that these differential metabolic properties in PAH patients may be a good therapeutic target. Also, signaling pathways from

oncogenic gene sets also provide novel understanding of PAH and the possibility to repurpose some anti-cancer drugs to PAH treatment.

Differences between IPAH and APAH

DEGSs between IPAH and APAH

While PAH patients are binned into different groups based on genetic background or associated disease, these groups to date had no influence on treatments assigned to each patient. The heterogeneity among PAH patients and the complexity of PAH pathogenesis are obstacles in the development of new PAH treatments. We took advantage of the transcriptomic expression profiles of PAH samples to detect transcriptomic differences among the various subgroups. IPAH and APAH are the two most common clinical subgroups in our datasets, most likely because of their high incidence in the population, and so we chose to focus on finding molecular differences between these two subtypes. Studying IPAH is of particular importance, given the uncertainty of how these patients develop PAH. We first calculated differential gene expression between either control and IPAH samples or control and APAH samples, then calculated the intersection between these two gene lists. Genes which were found on both gene lists were termed as pan-PAH genes, and genes which were specific to one subtype of PAH were considered as IPAH-specific or APAH-specific genes (**Figure 4.3A**). Generally, there was a high degree of overlap between the pan-PAH genes identified in this targeted analysis and the 79 DEGs previously found from the combined four datasets, as is to be expected.

Pathway-level DEGSs between IPAH and APAH

We also used ssGSEA to identify DEGSs between IPAH and APAH to characterize differences between these two subtypes. The DEGSs from C2 and C6 that differ between IPAH and APAH were plotted as heatmap with gene sets unsupervised hierarchical clustered (**Figure 4.3B/C**), while DEGSs form Hallmark was just shown in a table (**Table 4.2**). One interesting finding is IPAH samples show upregulated "BIOCARTA_DNAFRAGMENT_PATHWAY" and "WP_HISTONE_MODIFICATIONS", indicating they might be sensitive to drugs which modify the epigenome, or compounds which are protective against DNA damages. While "WP_APOPTOSISRELATED_NETWORK_DUE_TO_ALTERED_NOTCH3_IN_OVARIAN_CANCER" is upregulated in APAH samples, which is in accordance with the finding that notch3 signaling is essential for vascular remodeling in PAH[185]. Also, "BIOCARTA_ACE2_PATHWAY" is upregulated in APAH. However, decreased levels of serum ACE2 have been found in APAH patients, suggesting that in some cases these transcriptional signatures may contrast with biological observations. The reduction in ACE2 levels could just be caused by the heterogeneity of samples or differences between lung tissues and serum. ACE2 infusion has demonstrated some beneficial effects for PAH patients. However, the mechanisms by which ACE2 infusion exerts beneficial effects is unknown and it might only benefit a subgroup of PAH patients[186].

Use CMap to Analyze DEGs to Identify Compounds targeting PAH or each PAH subgroups

We next used the 79 DEGs and DEGs specific to each PAH subgroup as "PAH signature", "IPAH signature" and "APAH signature", and sought to identify chemical compounds that reverse this signature. If a compound can reverse PAH-associated gene expression, then it may be

effective at treating PAH. To accomplish this, we leveraged the CMap, wherein thousands of compounds were profiled against a panel of cell lines and all the gene expression profiles were collected. We compared all three PAH signatures with the CMap signatures to identify compounds which reversed the PAH signatures. We used a reversal score cutoff of 80 (out of 100) for individual compounds, and a reversal score cutoff of 10 for CMap classes (groups of signatures from similar perturbations). The lists of the CMap classes and descriptions of the compounds from the three analysis are compared in a Venn diagram (**Figure 4.4)** and listed in **Table 4.3**. The "PAH signature" analysis successfully identified compounds targeting signaling pathways that are already clinically targeted in PAH patients, including a nitric oxide production inhibitor (Brazilin), a prostanoid receptor agonist (16,16-dimethylprostaglandin-e2), and a guanylate cyclase activator (Isoliquiritigenin). Additionally, a retinoid receptor agonist (SA-792709) was found to reverse the signatures of both IPAH and APAH. This is in accordance with the attenuation of pulmonary hypertension caused by stimulation of PPARγ[187]. PPARγ may regulate several mediators of pulmonary hypertension, such as nitric oxide, endothelin 1, prostacyclin, and inflammation[187]. This analysis also identified compounds which we would expect might be effective against PAH, based pre-clinical knowledge of PAH pathogenesis. For example, we found that a TGFβ receptor inhibitor and an aromatase inhibitor (Exemestane) both reversed the PAH signature, which is consistent with the idea that TGFβ and estrogen are important mediators of PAH pathogenesis. For IPAH specifically, both an EGFR inhibitor (Tyrphostin-AG-82) and an AKT inhibitor (Triciribine) were identified as specifically reversing the IPAH signature. They have also been found to attenuate pulmonary vascular remodeling in preclinical studies[188]. These results that align with known biology indicate the power of this

analysis to look for potential novel PAH treatments, particularly for IPAH patients. Still, the findings need be scrutinized and validated in studies *in vitro* or *in vivo*. For example, EGFR inhibitors gefitinib, erlotinib did not attenuate PAH, while dacomitinib indeed inhibited hypoxia induced abnormalities in PASMCs[189]. Imatinib, a protein-tyrosine kinase inhibitor, has been tested in clinical trials for PAH (NCT01392469). However, tyrosine kinase inhibitors have demonstrated limiting cardiotoxicity[190].

Molecular subgroups of PAH

While our analysis found differences between IPAH and APAH samples, we next wanted to take an unbiased approach to identify transcriptomic subclusters of PAH samples. This approach would allow us to group patients based on signaling pathway alterations and could help us dissect the pathogenesis mechanism underlying PAH, since patients with similar transcriptomic profiles should share similar underlying mechanisms. Transcriptomic classification of PAH in theory may identify therapeutic targets specific for subsets of PAH patients that may be more effective than the current generalized therapeutic approaches. We used all the Hallmark and C6 gene sets, and the top 10% of the most variable C2 gene sets to perform unsupervised hierarchical clustering.

From the clustering results (**Figure 4.5**), we found that for gene sets clusters, group 3 includes several pathways that are relevant to cell cycle, such as HALLMARK_MYC_TARGETS, HALLMARK_DNA_REPAIR and HALLMARK_G2M_CHECKPOINT, along with HALLMARK_ GLYCOLYSIS. Upregulation of these signaling pathways could indicate increased proliferation in these lung samples. For gene set clusters, group 4 includes multiple inflammatory pathways, such

as HALLMARK_INFLAMMATORY_RESPONSE, HALLMARK_IL6_JAK_STAT3_SIGNALING, IL2_UP.V1, and HALLMARK_TNFA_SIGNALING_VIA_NFKB. These inflammatory pathways are relatively high in both sample clusters 2 and 4. Sample cluster 2 especially is enriched for APAH samples. Additionally, several IPAH samples in this cluster also have high activated inflammation-related signaling pathways compared to other IPAH samples. This shows that a small group of PAH patients indeed share a very similar transcription profiles and would benefit targeted therapies aiming to reverse their transcriptomic signatures.

Interestingly, clustering separated the PAH samples into four main groups that all included both IPAH and APAH patients. As a preliminary quality control step, we confirmed that these samples were not simply segregating based upon the dataset from which the samples were derived. Second, we tried to correlate the molecular subgroups with the clinical defined subgroups. Interestingly, the samples did not cluster based upon IPAH/APAH status, suggesting that binning patients into these subtypes may not be enough to guide targeted therapy. This indicates the clinical classification methods failed to represent the transcriptomic features of the patients.

Discussion

We have successfully identified differentially expressed genes and gene sets between control and PAH samples. With the substantial number of samples provided by aggregating four datasets, our confidence in the identified DEGs is increased. In every analysis we performed, we were able to confirm known PAH biology, which increases our confidence in this methodology. Also, we were also able to identify new genes and pathways which warrant further exploration

and validation. This indicates that we have been able to identify new potential mechanisms from our findings. Ultimately, the genes and signaling pathways identified in this study are potential new therapeutic targets or biomarkers for PAH.

However, with the large numbers of DEGs and DEGSs, it requires careful scrutinization with literature and further validation of the findings with biological experiments. While all of the samples included in this study are from similar tissue samples, the characteristics of the patients in terms of age, ancestry, sex, and other confounders, is highly variable. Unfortunately, we also don't have detailed clinical data for all the patients such as their hemodynamic condition and current therapeutic regime. The lung tissue samples used also include multiple cell types instead of just vascular cells, where the main pathology exists. Even within the vasculature, the pathogenesis mechanism in different cell types might differ. This problem can be solved by single cell RNA sequencing, which current is still limited technically and costly. By profiling these samples on a single cell level, it will almost certainly be possible to identify signaling and transcriptional alterations on a more granular level. This is especially important to investigate underlying mechanism specific to a certain cell type.

While we only compared IPAH and APAH samples in this study, it would also be interesting to compare samples from other PAH subtypes as well. This analysis will require further generation of RNA-Seq data from these samples, which is made difficult by the rarity of patients with these subtypes. These transcriptional analyses may be most powerful for HPAH patients, since known germline variants are involved in PAH pathogenesis. Transcriptomic analysis of samples from these patients will allow us to identify signaling pathways that are differentially

activated in response to a germline variant. This analysis may also shed light on the why there is substantial incomplete penetrance of these variants.

Although we have identified four molecular subgroups of PAH samples, the unsupervised clustering results depend heavily on the samples and gene sets selected for the analysis. In order to demonstrate clinical significance with the molecular subgroups, a much larger number of patient samples numbers is required. However, what this present study provides is a proof of concept to show that a scaled-up dataset, with perhaps an order of magnitude more samples, would be able to segregate samples based on transcriptional profiles, and not based on pathology-based classification. If complete clinical information can be integrated with these transcriptional data, it would be possible to further identify genes/pathways associated with patient survival, and also determine whether specific molecular subgroups of PAH patients have better/worse overall survival.

In addition to transcriptomic profiles, genetic information, epigenetic features, metabonomics and proteomics together would offer a more complete picture of PAH development. Along with the development of bioinformatic analysis tools, collaboration among multiple-centers to standardize the gathering of the samples and establish an omics database would greatly advance the understanding the PAH. Personalized treatment has been introduced to many diseases and would also improve the prognosis of PAH, a lethal cardiovascular disease.

Figure 4.1A. Volcano plots of DEGs between control and PAH.
Differential gene expression between control and PAH samples was calculated with limma, as described in Materials and Methods. Genes with an absolute logFC >= 0.6 and an FDR less than 0.01 were considered statistically significant. All genes which pass these cutoffs are highlighted in red.

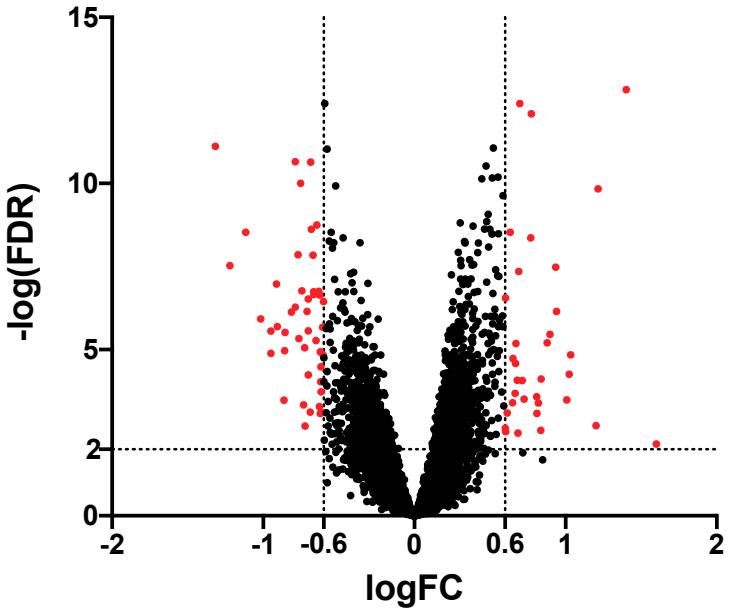


Figure 4.1B. Heatmap of DEGs between control and PAH.
Statistically significant genes from Figure 1A were replotted as a heatmap. Blue boxes are genes which are downregulated in PAH samples compared to control, and red boxes are genes that are upregulated in PAH samples compared to control. Samples were split into Control (Top, magenta), and PAH (Top, blue). Genes were clustered with unsupervised hierarchical clustering.

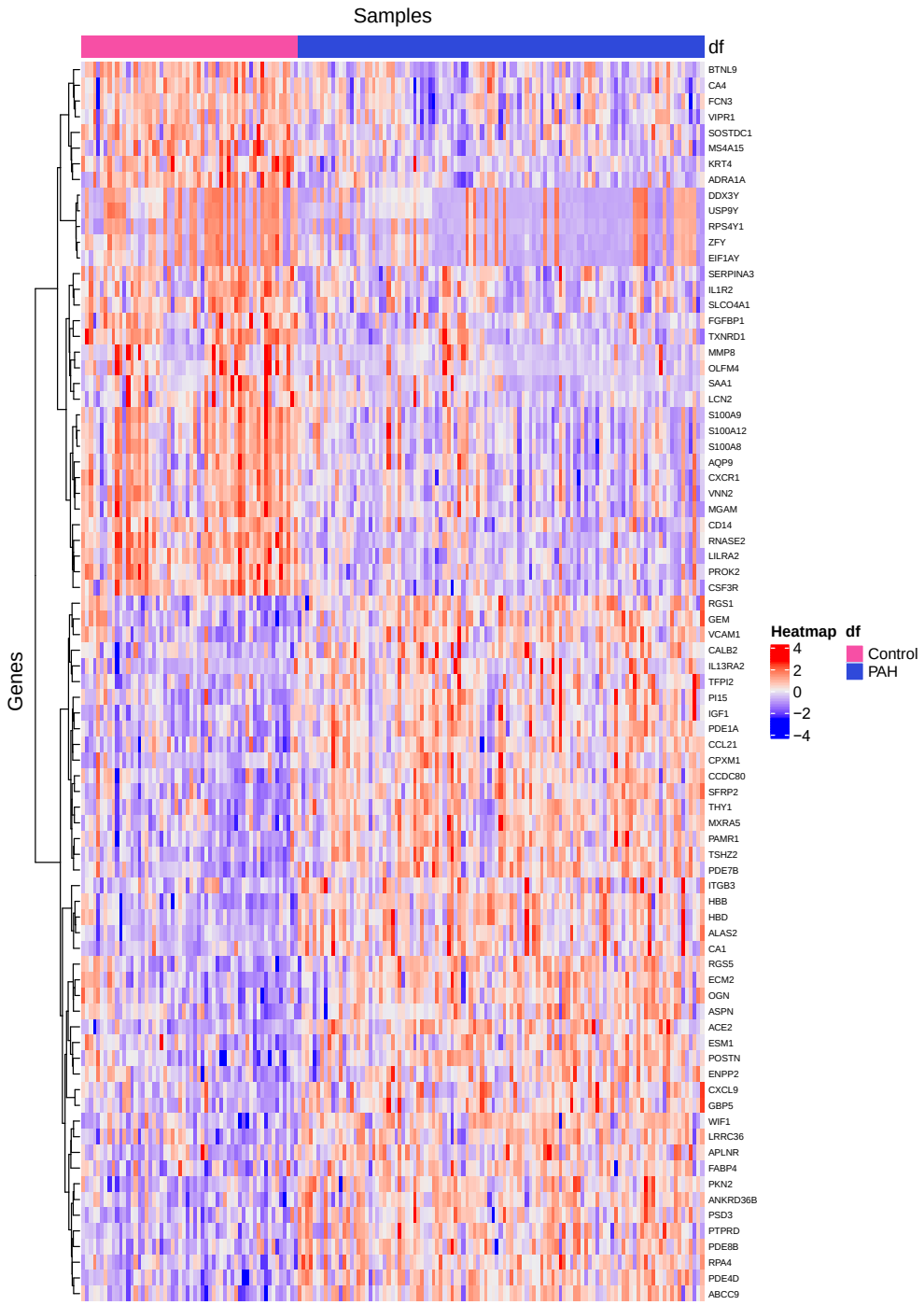


Figure 4.1C. Protein-Protein Interaction Network.
Network analysis on all statistically significant genes was performed with String as described in Methods.

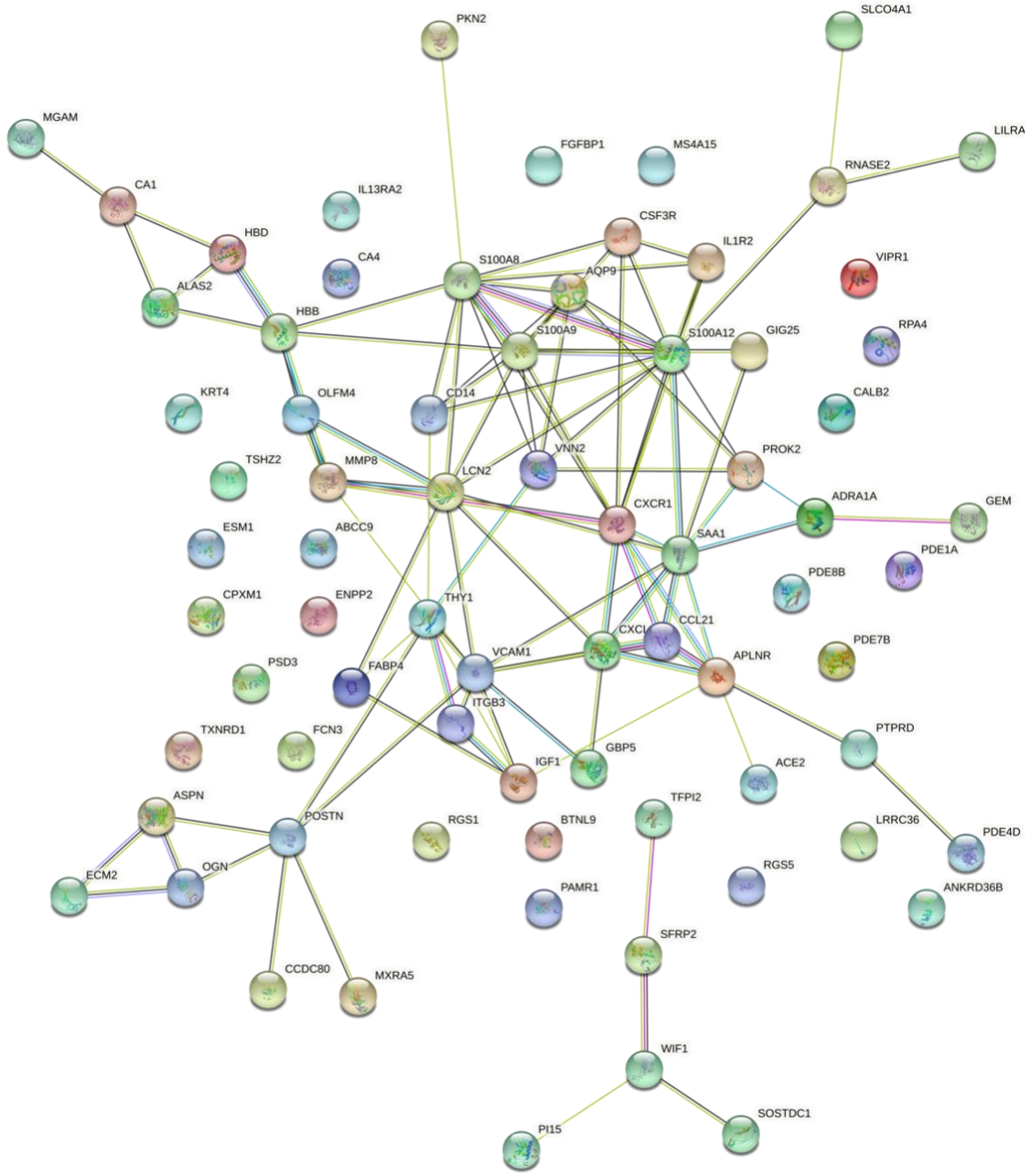


Figure 4.1D. Overlap between the PAH gene signature and MSigDB gene sets.
The overlap between the statistically significant genes identified in Figure 1A and all gene sets in the C2CP, C6, and H groups of MSigDB gene sets was calculated. P-values are derived from calculating the hypergeometric distribution between the PAH gene set and each individual MSigDB gene set. FDR-corrected p-values (x-axis) were plotted for each MSigDB gene signature (y-axis). The bar color indicates the total number of overlapping genes between the PAH gene set and the MSigDB gene set. Gene sets with an FDR < 0.05 were considered statistically significant.

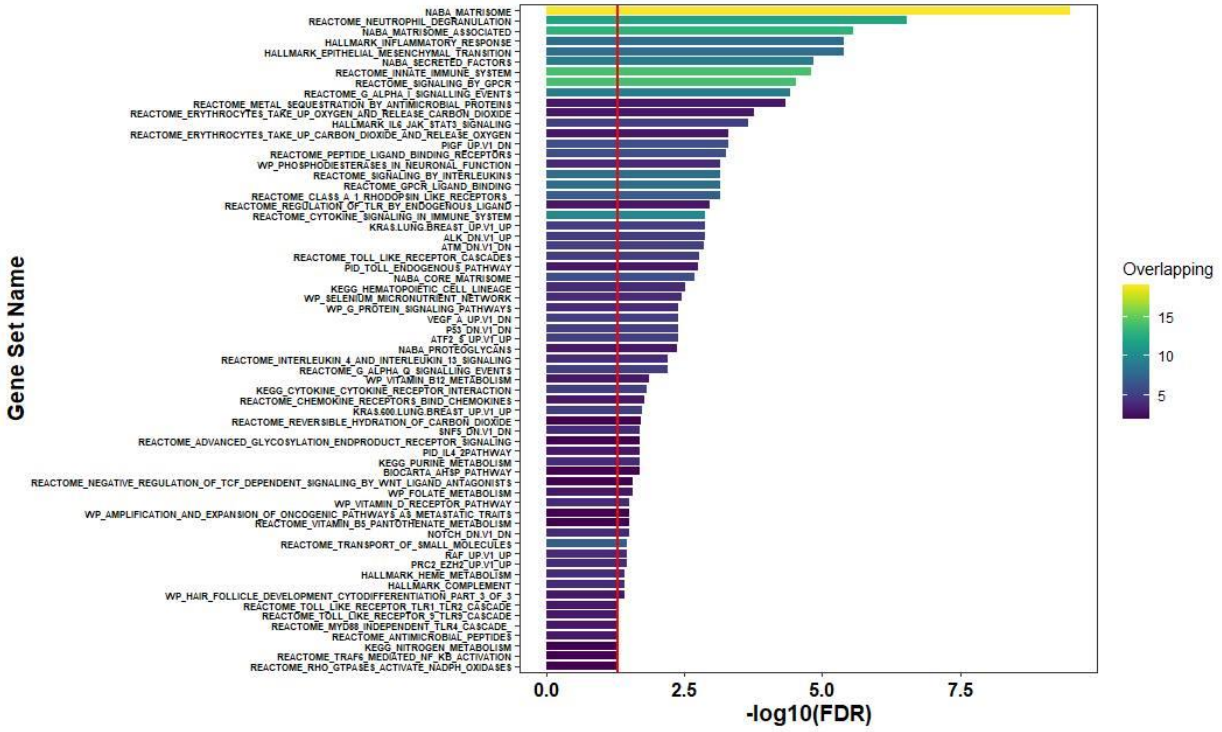


Figure 4.2. Differentially expressed C2 gene sets in control vs PAH samples.
ssGSEA was used to calculate signature scores for all gene sets in the MSigDB C2 collection on every individual Control and PAH sample. Z-score normalized signature scores (heatmap rows), were plotted for each control or PAH sample (heatmap columns). The color of the tiles in the heatmap is not indicative of whether the gene set is up- or down-regulated, rather, the colors express whether there is low or high expression of the gene set in that particular sample vs. all other samples for the same gene set. Gene sets were clustered by unsupervised hierarchal clustering and samples were binned into a control and a PAH group.

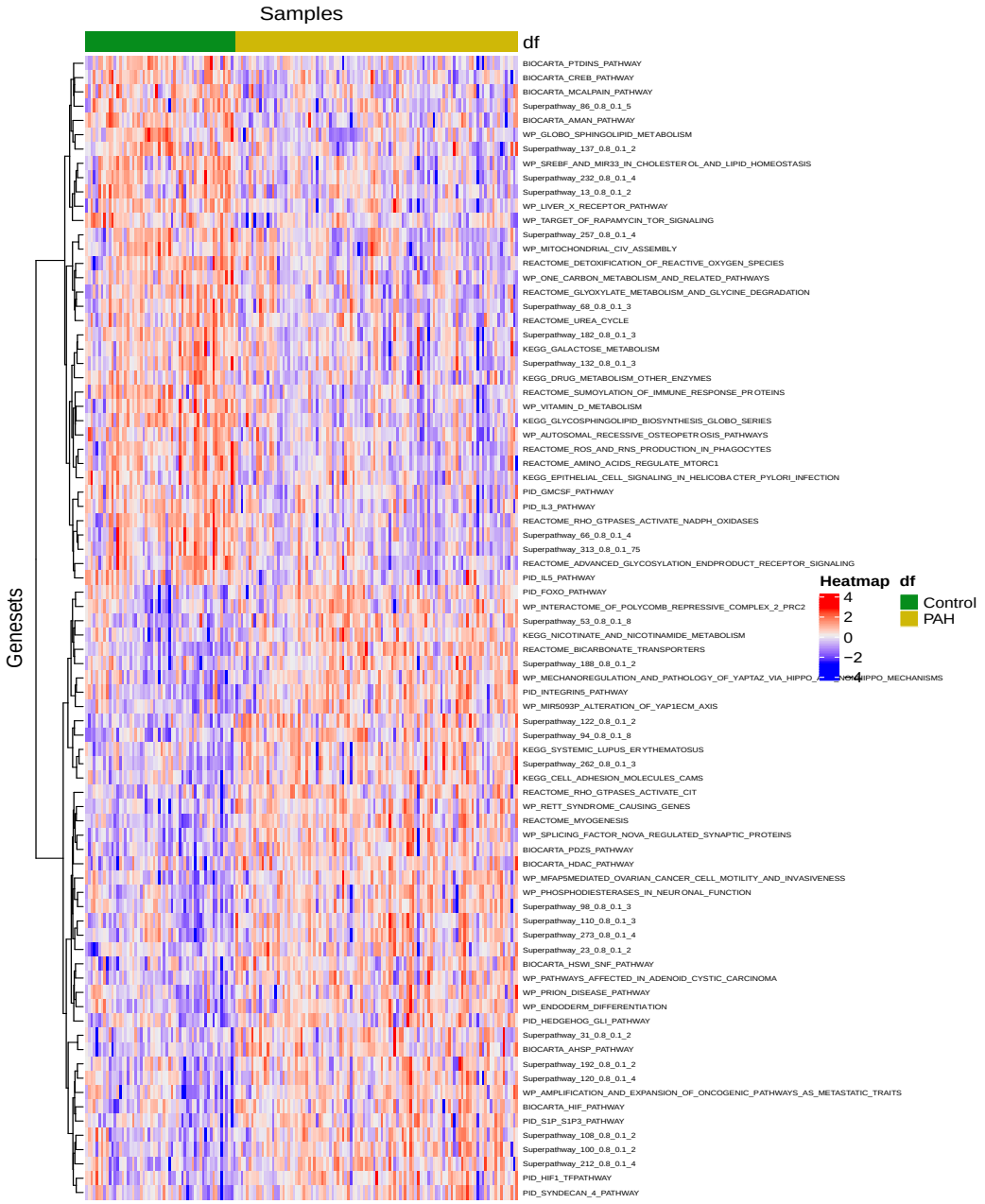


Figure 4.3A. Differentially expressed genes in IPAH and APAH samples.
Differential gene expression between control and IPAH samples or control and APAH samples was calculated and differentially expressed genes that are shared between both sample classes (top section), are unique to APAH (middle section), or are unique to IPAH (bottom section) were plotted as a heatmap. Samples (columns) were split into control, IPAH, and APAH according to the annotation bar above the heatmap. Gene expression values were z-score normalized such that a red box indicates high expression of that gene in that particular sample relative to all other samples, and a blue box indicates low expression of that gene in that particular sample relative to all other samples.

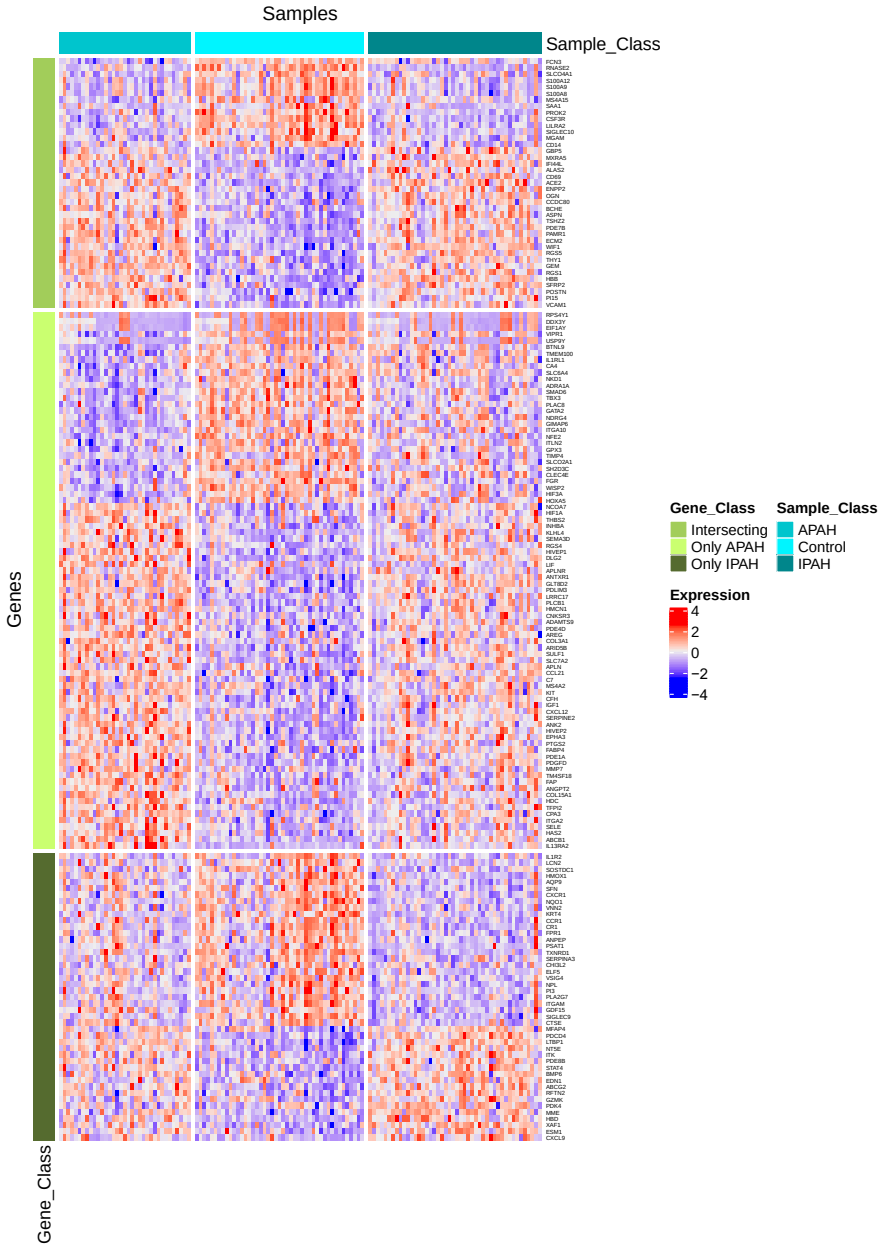


Figure 4.3B. Differential expression of MSigDB C6 gene sets between control, IPAH, and APAH samples.

ssGSEA was used to calculate signature expression for all signatures in the MSigDB C6 group of signatures. Differential signature expression between control and IPAH or control and APAH samples was calculated with limma. Statistically significant gene sets were stratified into gene sets that are differentially expressed in both APAH and IPAH, only APAH, or only IPAH, z-score normalized across samples, and plotted as a heatmap. Red boxes indicate that there is high expression of that gene set in that particular sample, relative to other samples. Blue boxes indicate that there is lower expression of that gene set in that particular sample.

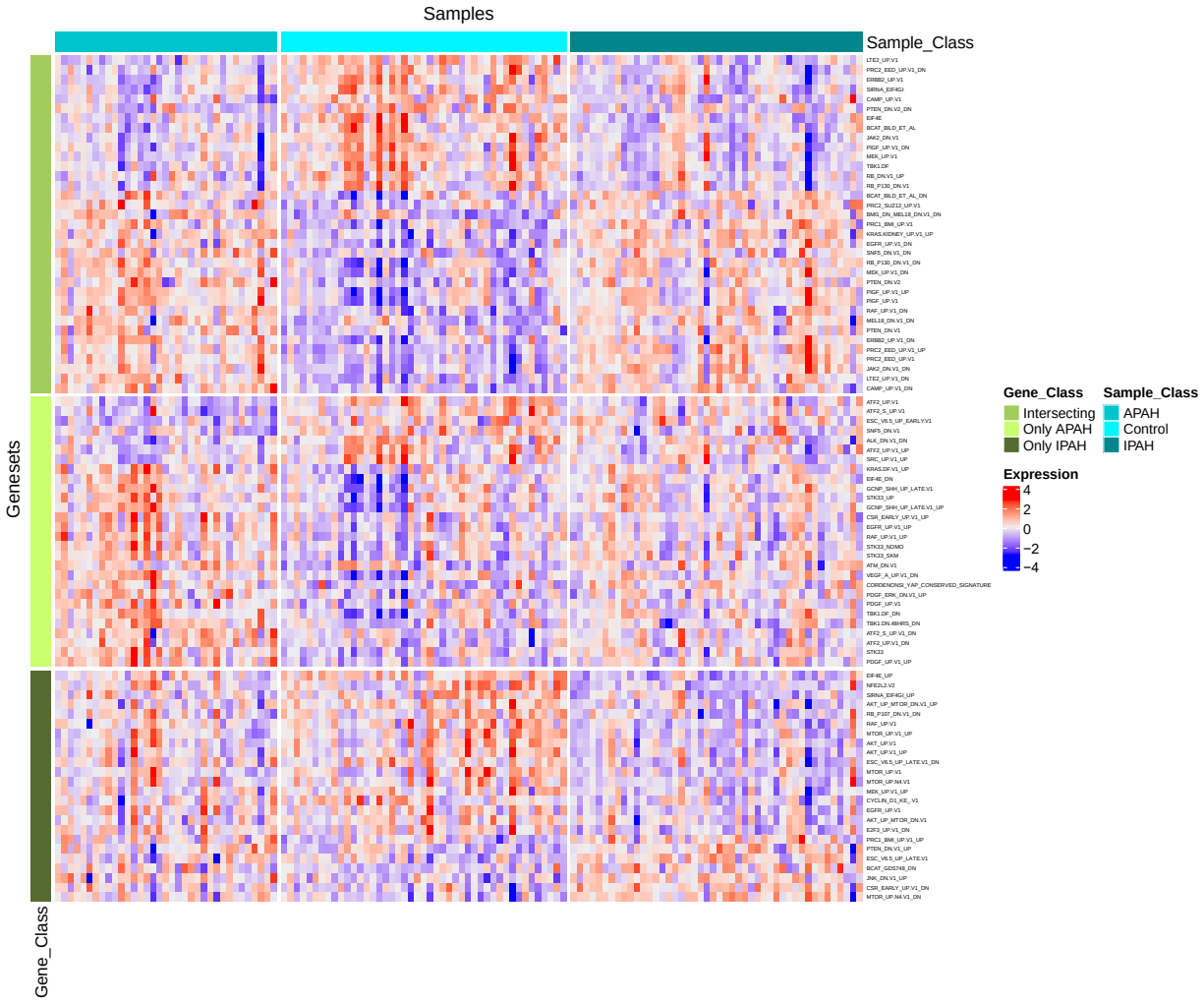


Figure 4.3C. Differential expression of MSigDB C2 gene sets between control, IPAH, and APAH samples.

ssGSEA was used to calculate signature expression for all signatures in the MSigDB C2 group of signatures. Differential signature expression between control and IPAH or control and APAH samples was calculated with limma. Statistically significant gene sets were stratified into gene sets that are differentially expressed in both APAH and IPAH, only APAH, or only IPAH, z-score normalized across samples, and plotted as a heatmap. Red boxes indicate that there is high expression of that gene set in that particular sample, relative to other samples. Blue boxes indicate that there is low expression of that gene set in that particular sample, relative to all other samples.

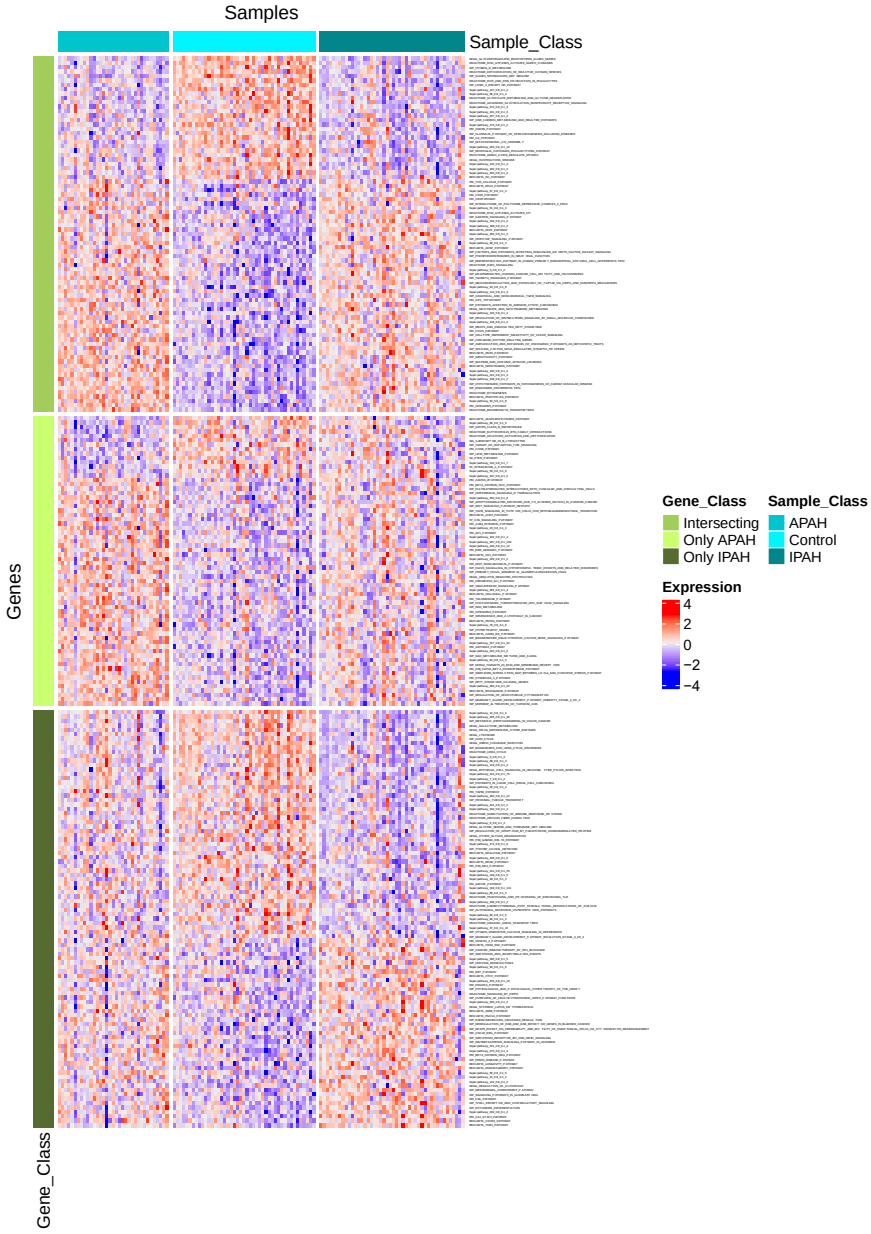


Figure 4.4. Identification of compounds which reverse the PAH gene signature.
Differentially expressed genes were analyzed with Connectivity Map to identify compounds which reverse expression of the PAH gene signature. Compounds or Cmap Classes that reverse the PAH expression signature are listed in green (for the pan-PAH signature), blue (for the control vs APAH signature), or red (for the control vs IPAH signature). The names of the gene sets from each sector of the Venn diagram are listed in the appropriate box.

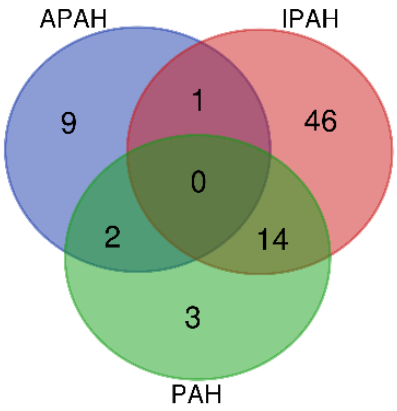


Figure 4.5. Clustering of PAH samples by gene set.
ssGSEA was used to calculate gene set expression score for all MSigDB C2CP, C6, and Hallmark gene sets for all PAH samples in the dataset. All gene sets were z-score normalized across all samples. Samples and genes were clustered with unsupervised hierarchical clustering. The clinical subtype of each sample is indicated above the heatmap. A red box indicates that there is high expression of the gene set in that particular sample, relative to all other samples in the dataset. A blue box indicates that there is low expression of the gene set in that particular sample, relative to all other samples in the dataset.

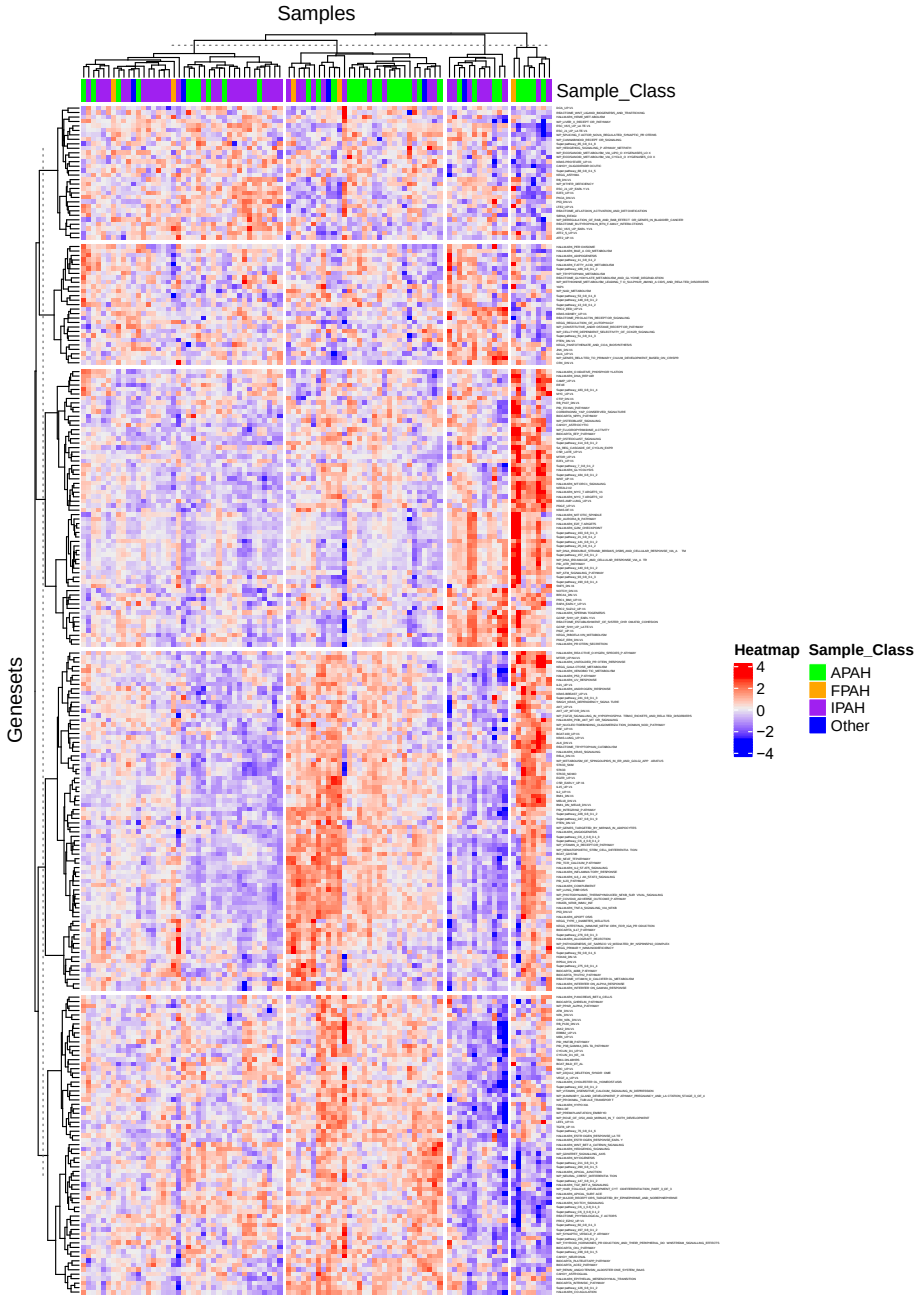


Figure 4.6. Removal of batch effects with combat.
A) All four groups of samples or **B)** All samples except for those in GSE15197 were merged together and PCA analysis was performed before (left) and after (right) combat was performed to remove batch effects. Different colors correspond to different initial datasets.

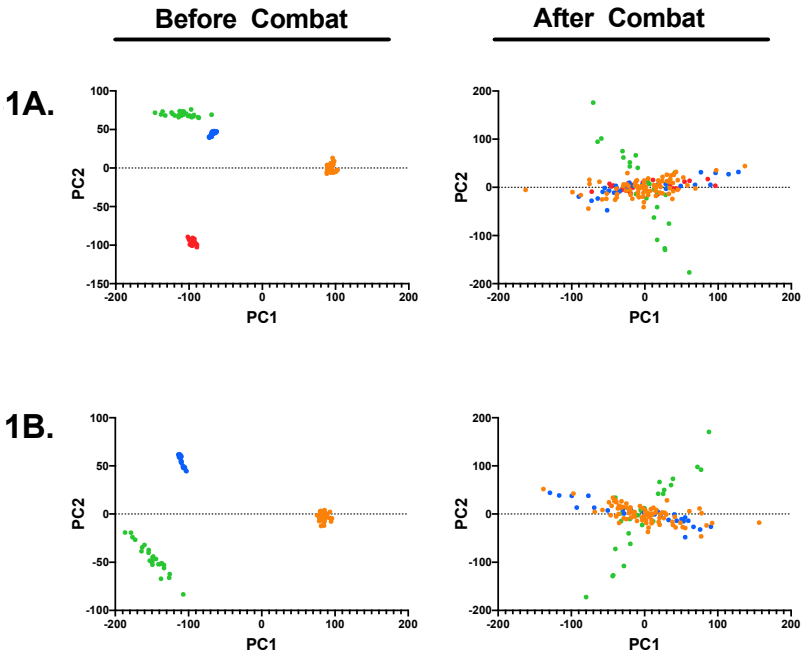


Table 4.1. Differentially expressed gene sets in control vs PAH samples.
ssGSEA was used to calculate signature scores for all gene sets in MSigDB Hallmark and C6 on every individual Control and PAH sample. Statistically significant differences in gene set expression were calculated with limma.

	logFC	FDR
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY	0.826612	1.43E-05
HALLMARK_UV_RESPONSE	0.636968	0.001288
MEL18_DN.V1_DN	-0.79051	0.000126
JAK2_DN.V1_DN	-0.77017	0.000159
JNK_DN.V1_UP	-0.71428	0.000629
PTEN_DN.V1	-0.66731	0.00174
CAMP_UP.V1_DN	-0.64022	0.002731
EGFR_UP.V1_DN	-0.61658	0.003298
LTE2_UP.V1_DN	-0.61518	0.003298
LTE2_UP.V1	0.622886	0.003298
NFE2L2.V2	0.632646	0.002907
EIF4E_UP	0.663083	0.00174
TBK1.DF	0.81352	0.000121

Table 4.2. Differential expression of MSigDB Hallmark gene sets between control, IPAH, and APAH samples.
ssGSEA was used to calculate signature expression for all signatures in the MSigDB Hallmark group of signatures. Differential signature expression between control and IPAH or control and APAH samples was calculated with limma. Gene set that are upregulated in the PAH samples are in red, gene sets that are downregulated in the PAH samples are in blue. Gene sets that are differentially expressed in both APAH and IPAH are highlighted with a yellow box.

Control vs. APAH	Control vs. IPAH
HALLMARK_UV_RESPONSE_DN	HALLMARK_HEDGEHOG_SIGNALING
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	HALLMARK_UNFOLDED_PROTEIN_RESPONSE
HALLMARK_ANGIOGENESIS	HALLMARK_MYC_TARGETS_V2
HALLMARK_COAGULATION	HALLMARK_DNA_REPAIR
HALLMARK_HEDGEHOG_SIGNALING	HALLMARK_P53_PATHWAY
HALLMARK_OXIDATIVE_PHOSPHORYLATION	HALLMARK_GLYCOLYSIS
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY	HALLMARK_UV_RESPONSE
	HALLMARK_XENOBIOTIC_METABOLISM
	HALLMARK_OXIDATIVE_PHOSPHORYLATION
	HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY

Table 4.3. Gene sets from each sector of the Venn diagram in Figure 4.4.

APAH PAH	**IPAH**
Glucocorticoid receptors agonist Neurogenesis of non-pluripotent C2C12 myoblast inducer	PPAR receptor agonist Glucokinase activator GABA receptor antagonist IP1 prostacyclin receptor agonist HSP inhibitor FXR antagonist C2 domain containing LOF HIV protease inhibitor Cathepsin inhibitor Cholesterol inhibitor NKL subclass homeoboxes and pseudogenes LOF PPAR receptor antagonist Aromatase inhibitor Estrogen receptor antagonist TGF β receptor inhibitor Estrogen receptor agonist T-type calcium channel blocker Exportin antagonist Adenosine receptor antagonist Glycogen synthase kinase inhibitor Sodium channel blocker Chymotrypsin inhibitor Inhibitor of translocation of PKCq in T cells Proteasome inhibitor Tyrosinase inhibitor Glucosyltransferase inhibitor SYK inhibitor EGFR inhibitor Histone lysine methyltransferase inhibitor BTK inhibitor Cytotoxic lipid peroxidation product MDM inhibitor DNA synthesis inhibitor Calcineurin inhibitor Immunostimulant BCL inhibitor Leukocyte elastase inhibitor CDC inhibitor vitamin analog Metalloproteinase inhibitor HIF activator AKT inhibitor ATR kinase inhibitor Acetylcholine receptor antagonist Adrenergic receptor antagonist
APAH IPAH	
Retinoid receptor agonist	
IPAH PAH	
Opioid receptor antagonist Nitric oxide production inhibitor Prostanoid receptor agonist Guanylate cyclase activator Protein tyrosine kinase inhibitor TRPV agonist NFkB pathway inhibitor Antineoplastic Sodium/potassium/chloride transporter inhibitor PKC activator Tubulin inhibitor Neuropeptide receptor antagonist Apoptosis stimulant	
APAH	
Akt Signaling GOF NADH ubiquinone oxidoreductase supernumerary subunits LOF G2 M Checkpoint LOF Breast cancer resistance protein inhibitor Serpin peptidase inhibitors LOF MEK inhibitor DDR1 inhibitor Dopamine receptor antagonist Dipeptidyl peptidase inhibitor	
PAH	
Rho GTPase activating proteins LOF Caspase activator GK0582 inhibitor	

Table 4.4. Demographic Characteristics of Patients in the Datasets.
Patient demographics from each of the individual datasets that were analyzed in this study.

GEO	N		Age		Sex		Subgroups
	Control	PAH	Control	PAH	M	F	
GSE 15197	13	18	60 ±11	44±10	12	19	unknown
GSE 48149	9	17	53 ±		10	16	8 IPAH, 10 APAH
GSE 117261	25	58	1-64	7-79	33	50	32 IPAH, 17 APAH
GSE 113439	11	15		41±12	3	12	6 IPAH, 8 APAH

CHAPTER 5 : CONCLUSION AND FUTURE DIRECTIONS

Results

This dissertation focuses on investigating the underlying mechanisms of PAH and aims to identify and characterize potential therapeutic targets. The aim of the first part of this dissertation is to characterize the role of TGFβ signaling in PAH pathogenesis. In particular, I sought to clarify the signaling crosstalk between the canonical SMAD pathway and non-canonical MRTF pathways which signal downstream of TGFβ. A major finding in this dissertation is the identification of a signaling connection between TGFβ and S1P, which offers a new insight into how both pathways are involved in disease pathogenesis. Mechanistically, my data suggests that TGFβ activates the S1P signaling pathway by inducing the autocrine/paracrine signaling through multiple S1PR isoforms. Although cell surface receptors are pharmacologically tractable, redundancy in downstream signaling pathways may make therapies that target these receptors less effective. Instead, it may be more effective to target shared signaling nodes downstream of multiple receptors. I hypothesized that MRTF/SRF acts as a transcriptional node that regulates SMC hypertrophy. Consistent with this hypothesis, MRTF/SRF-mediated transcription is essential for PAH pathogenesis in animal models.

In addition to canonical TGFβ signaling, another member of the TGFβ superfamily, BMPR2, has been implicated in PAH. Activation of TGFβ signaling and reduced BMPR2 signaling is believed to drive PAH pathogenesis[81]. Consistent with this idea, patients with heritable PAH harbor loss of function mutations in BMPR2[60]. A second goal of this dissertation was to understand the signaling crosstalk between canonical TGFβ signaling and BMPR2. Surprisingly, BMPR2 deletion did not alter TGFβ activity in fibroblasts. One explanation for this finding is that acute loss of BMPR2 may result in different signaling alterations than long-term deletion of BMPR2.

Understanding the temporal effects of BMPR2 will be important since short-term deletion of BMPR2 may not reflect the biology of patients with germline BMPR2 alterations.

Finally, I leveraged transcriptomic data to gain a global understanding of which signaling pathways are altered in the vasculature of PAH patients. I found TGFβ, S1P, and Rho signaling are upregulated in tissue samples from PAH patients.

My results also identified several signaling pathways that were differentially activated between control and PAH samples. For example, SARS and GLP-1 related pathways. These findings could help direct further investigation of the PAH pathogenesis. We also tried to find drugs which can reverse a PAH-associated transcriptional signature, some of which are consistent with clinical PAH treatment regimens. This would provide novel thoughts for the repurposing medication for PAH treatment.

Limitations

Many of the experiments described in this dissertation relied on primary smooth muscle cells, since primary cells better reflect the physiological conditions than established cell lines. However, we used SMCs isolated from pulmonary arteries. While these cells may be a good model for arterial smooth muscle cells, it is unclear whether these cells faithfully model arteriolar smooth muscle cells. Since vasoconstriction and vascular remodeling occurring in pulmonary arterioles are also essential contributors to the elevated pressure in the pulmonary vasculature, it is equally important to understand the signaling mechanisms in these cells[8]. While SMCs in the pulmonary arteries and arterioles are highly similar, they are exposed to different oxygen gradients and levels of mechanical pressure[191]. As such, these cells may behave differently in

certain contexts, and may be especially sensitive to differences in oxygen concentrations. While it is likely that my findings here will be applicable to SMCs in both the arteries and arterioles, this is still something that will need to be experimentally validated. Another shortcoming is that many of my experiments were performed with the WI38 fibroblast cell line. While this cell line is a good model to study physiological properties of fibroblasts, it may not fully recapitulate the adventitial fibroblasts in the lung.

In this dissertation I took two complementary approaches, pharmacological inhibition and genetic modification, to perturb several signaling pathways. One challenge with these types of pharmacology experiments is deciding on the correct compound concentration, particularly in *in vitro* experiments. For example, it is difficult to directly translate the Ki of an inhibitor to a concentration for experiments done in a cellular setting. In my experiments, the concentrations that I've used have been based on both the literature and the experimental Ki. However, the possibility remains that the observed effect is off target since some compounds were used at relatively high concentrations. Some of the effects I observed are subtle, which makes dose-dependent response experiments difficult to perform. On the other hand, most genetic modifications are an all or nothing effect since CRISPR generally will result in biallelic loss of a gene. For genes that are essential for cell survival it is often not possible to generate cell lines with stable deletion of that gene. In an ideal situation, verification of results through both pharmacological and genetic inhibition would generate a more solid conclusion.

It was difficult to generate WI38 cells with stable long-term depletion of BMPR2 using both shRNA and CRISPR. The BMPR2 deficient cells have a doubling time of 3 days and are susceptible to death after passaging. This is likely because BMPR2 is essential for the survival of

these cells. As a result, I established an inducible knockdown system and found that this methodology was more conducive to reproducibly perturbing BMPR2. However, the cells were induced to silence BMPR2 immediately before an experiment, and only for at most a few days. It is uncertain whether this short-term knockdown of BMPR2 faithfully recapitulates the condition in PAH patients who may have experienced dysregulation of BMPR2 for years or potentially for their entire lifetime. It will be necessary to perform experiments to understand the temporal consequences of BMPR2 deletion. This gene silencing system used lentivirus and antibiotics for selection, but it is unknown whether lentivirus and antibiotics change the characteristics of the fibroblasts

My computational analysis revealed several genes and signaling pathways that could be important in PAH pathogenesis. One limitation I faced was a lack of high-quality transcriptomic data from PAH patient tissue samples. And the ones that do exist are generally smaller cohorts from different centers, which makes correcting for batch effects a challenge. Another limitation is that we also lack the detailed clinical information of all the samples, including the treatments the patients received or the disease stage at sample collection. The inability to correlate the bioinformatic findings with clinical outcomes makes it difficult to demonstrate the biologically significance of our findings.

Future Directions

The primary focus of my dissertation was to investigate signaling alterations *in vitro.* To expand upon this, it will be necessary to validate these signaling mechanisms in *in vivo* models of PAH. There are several established animal models for PAH including the Monocrotaline (MCT)

model, Hypoxia model, and SUGEN/Hypoxia model[192]. Another research group has already produced evidence that the MRTF pathway is important in PAH by demonstrating that CCG-1423, a compound developed in our laboratory, reduced PAH symptoms in an animal model[144]. PAH animal models may incompletely mimic the pathology observed in humans, such as plexiform lesions[193]. Additionally, the mechanisms underlying the elevated pulmonary hypertension in each animal model is not fully understood[193]. The MCT model is believed to be driven by increased inflammation, and the hypoxia model is more proximal to the physiological condition that most PAH patients experience[193]. Still, the majority of the experimental and preclinical studies in PAH are conducted via these animal models; it will be necessary to characterize which models most faithfully recapitulate the disease pathology observed in humans. Using this approach, I would first identify a set of signaling pathways which are believed to be involved in the PAH pathogenesis by analyzing human lung samples. Then I will evaluate their expression level in the lung tissues collected from PAH animal models in order to assess how well they can represent human disease. It is likely that different animal models will have differing degrees in how faithfully they recapitulate various aspects of PAH biology. Performing this analysis will help guide our interpretation of the scientific findings generated from the PAH animal models.

Another challenge is to identify specific changes in one single cell type *in vivo*. PAH is a progressive disease of lung vasculature with multifactorial etiology, and all three vascular cell types contribute to pathogenesis. Some studies focused on only individual cell types *in vitro* to address their role in this disease[124,[125]. Some microarray and RNAseq studies characterized the lung tissues from patients or PAH animal models [132, 133]. For lung tissues, it not only includes the three vascular cell types, but also other cell types that reside in the vasculature

including respiratory epithelium, pericytes, and immune cells[8]. Whether the genetic features we dissected from the bioinformatic data are shared among all cell types or are predominantly in a single cell type is unknown. Additionally, the interaction between the various cell types is poorly understood. Recently, research groups have started to employ single-cell RNA sequencing (scRNA-Seq) in PAH tissues[194, 195]. This offers a lung-wide perspective of the alterations in each cell type. This is especially true in immune cells, where most studies have focused on the released cytokines rather than the genetic changes of the immune cells [194, 195]. Analysis of cell-specific molecular signatures can reveal potential cellular markers associated with PAH, which would be masked in previous bulk sequencing analysis of the whole lung tissue. One exciting question is whether different cells within one patient represent different disease stages. If so, it may be possible to reconstruct the alterations that occur during disease progression in a single patient, so that we can better understand disease evolution.

One major limitation of sc-RNAseq is the undesired variability or bias in the cell-type composition of single-cell suspensions prepared from inherently complex tissues such as the lung. Problems of diverse patient backgrounds, such as disease stage and clinical treatment prior to sample biopsy, limit the implementation of this technology to study PAH. As an alternative approach, we could leverage spatial profiling methods which combine RNA sequencing and fluorescent staining to gain a spatial understanding of transcriptomic alterations and cellular subpopulations in the development of angio-obliterative vascular lesions[196, 197].

With the combination of sc-RNA seq and other evolving omics methods, we can have a better characterization of pathways and signaling hubs at a cellular level. It will also allow for the correlation of genetic changes of different cell types spatially and temporarily.

Pulmonary Arterial Hypertension is a complex disease, and this complexity has thus far precluded the development of effective therapies. Challenges include the complex genetics underlying heritability, cell-to-cell signaling interactions, and an incomplete knowledge of how this disease progresses. These challenges, in the genomics age, are now within reach. The combination of new sequencing technologies, and the computational methods to analyze that data, with existing PAH animal models will allow us to identify new therapeutic targets, and hopefully new therapies.

REFERENCES

REFERENCES

1. Simonneau, G., et al., *Haemodynamic definitions and updated clinical classification of pulmonary hypertension.* Eur Respir J, 2019. **53**(1).

2. McGoon, M.D., et al., *Pulmonary arterial hypertension: epidemiology and registries.* J Am Coll Cardiol, 2013. **62**(25 Suppl): p. D51-9.

3. Pugh, M.E., et al., *Causes of pulmonary hypertension in the elderly.* Chest, 2014. **146**(1): p. 159-166.

4. Frost, A.E., et al., *The changing picture of patients with pulmonary arterial hypertension in the United States: how REVEAL differs from historic and non-US Contemporary Registries.* Chest, 2011. **139**(1): p. 128-37.

5. Rosenkranz, S., *Pulmonary hypertension 2015: current definitions, terminology, and novel treatment options.* Clin Res Cardiol, 2015. **104**(3): p. 197-207.

6. Stein, P.D., F. Matta, and P.G. Hughes, *Scope of problem of pulmonary arterial hypertension.* Am J Med, 2015. **128**(8): p. 844-51.

7. Farber, H.W., et al., *Five-Year outcomes of patients enrolled in the REVEAL Registry.* Chest, 2015. **148**(4): p. 1043-54.

8. Pietra, G.G., et al., *Pathologic assessment of vasculopathies in pulmonary hypertension.* J Am Coll Cardiol, 2004. **43**(12 Suppl S): p. 25S-32S.

9. Peacock, A.J., et al., *An epidemiological study of pulmonary arterial hypertension.* Eur Respir J, 2007. **30**(1): p. 104-9.

10. Sitbon, O., et al., *Long-term response to calcium channel blockers in idiopathic pulmonary arterial hypertension.* Circulation, 2005. **111**(23): p. 3105-11.

11. Humbert, M., et al., *Pathology and pathobiology of pulmonary hypertension: state of the art and research perspectives.* Eur Respir J, 2019. **53**(1).

12. Guignabert, C. and P. Dorfmuller, *Pathology and pathobiology of pulmonary hypertension.* Semin Respir Crit Care Med, 2013. **34**(5): p. 551-9.

13. Pogoriler, J.E., et al., *Persistence of complex vascular lesions despite prolonged prostacyclin therapy of pulmonary arterial hypertension.* Histopathology, 2012. **61**(4): p. 597-609.

14. Sakao, S., K. Tatsumi, and N.F. Voelkel, *Endothelial cells and pulmonary arterial hypertension: apoptosis, proliferation, interaction and transdifferentiation.* Respir Res, 2009. **10**: p. 95.

15. Cho, J.G., et al., *Endothelial to Mesenchymal Transition Represents a Key Link in the Interaction between Inflammation and Endothelial Dysfunction.* Front Immunol, 2018. **9**: p. 294.

16. Morrell, N.W., et al., *Altered growth responses of pulmonary artery smooth muscle cells from patients with primary pulmonary hypertension to transforming growth factor-beta(1) and bone morphogenetic proteins.* Circulation, 2001. **104**(7): p. 790-5.

17. Ljuca, F. and G. Drevensek, *Endothelin-1 induced vascular smooth muscle cell proliferation is mediated by cytochrome p-450 arachidonic acid metabolites.* Bosn J Basic Med Sci, 2010. **10**(3): p. 223-6.

18. El Kasmi, K.C., et al., *Adventitial fibroblasts induce a distinct proinflammatory/profibrotic macrophage phenotype in pulmonary hypertension.* J Immunol, 2014. **193**(2): p. 597-609.

19. Stenmark, K.R., et al., *Role of the adventitia in pulmonary vascular remodeling.* Physiology (Bethesda), 2006. **21**: p. 134-45.

20. Maiellaro, K. and W.R. Taylor, *The role of the adventitia in vascular inflammation.* Cardiovasc Res, 2007. **75**(4): p. 640-8.

21. Stenmark, K.R., et al., *The adventitia: Essential role in pulmonary vascular remodeling.* Compr Physiol, 2011. **1**(1): p. 141-61.

22. Steiner, M.K., et al., *Interleukin-6 overexpression induces pulmonary hypertension.* Circ Res, 2009. **104**(2): p. 236-44, 28p following 244.

23. Maleszewska, M., et al., *IL-1beta and TGFbeta2 synergistically induce endothelial to mesenchymal transition in an NFkappaB-dependent manner.* Immunobiology, 2013. **218**(4): p. 443-54.

24. Chaisson, N.F. and P.M. Hassoun, *Systemic sclerosis-associated pulmonary arterial hypertension.* Chest, 2013. **144**(4): p. 1346-1356.

25. Ohta-Ogo, K., et al., *CD44 expression in plexiform lesions of idiopathic pulmonary arterial hypertension.* Pathol Int, 2012. **62**(4): p. 219-25.

26. Boucherat, O., et al., *The cancer theory of pulmonary arterial hypertension.* Pulm Circ, 2017. **7**(2): p. 285-299.

27. Rai, P.R., et al., *The cancer paradigm of severe pulmonary arterial hypertension.* Am J Respir Crit Care Med, 2008. **178**(6): p. 558-64.

28. Cool, C.D., et al., *The hallmarks of severe pulmonary arterial hypertension: the cancer hypothesis-ten years later.* Am J Physiol Lung Cell Mol Physiol, 2020. **318**(6): p. L1115-L1130.

29. Yeager, M.E., et al., *Microsatellite instability of endothelial cell growth and apoptosis genes within plexiform lesions in primary pulmonary hypertension.* Circ Res, 2001. **88**(1): p. E2-E11.

30. Aldred, M.A., et al., *Somatic chromosome abnormalities in the lungs of patients with pulmonary arterial hypertension.* Am J Respir Crit Care Med, 2010. **182**(9): p. 1153-60.

31. Sommer, N., et al., *Current and future treatments of pulmonary arterial hypertension.* Br J Pharmacol, 2020.

32. Guignabert, C., et al., *Pathogenesis of pulmonary arterial hypertension: lessons from cancer.* Eur Respir Rev, 2013. **22**(130): p. 543-51.

33. Bonnet, S., et al., *An abnormal mitochondrial-hypoxia inducible factor-1alpha-Kv channel pathway disrupts oxygen sensing and triggers pulmonary arterial hypertension in fawn hooded rats: similarities to human pulmonary arterial hypertension.* Circulation, 2006. **113**(22): p. 2630-41.

34. McMurtry, M.S., et al., *Dichloroacetate prevents and reverses pulmonary hypertension by inducing pulmonary artery smooth muscle cell apoptosis.* Circ Res, 2004. **95**(8): p. 830-40.

35. Ulrich, S., et al., *Increased regulatory and decreased CD8+ cytotoxic T cells in the blood of patients with idiopathic pulmonary arterial hypertension.* Respiration, 2008. **75**(3): p. 272-80.

36. Valzasina, B., et al., *Tumor-induced expansion of regulatory T cells by conversion of CD4+CD25- lymphocytes is thymus and proliferation independent.* Cancer Res, 2006. **66**(8): p. 4488-95.

37. Hoeper, M.M., et al., *Imatinib mesylate as add-on therapy for pulmonary arterial hypertension: results of the randomized IMPRES study.* Circulation, 2013. **127**(10): p. 1128-38.

38. Rich, S., E. Kaufmann, and P.S. Levy, *The effect of high doses of calcium-channel blockers on survival in primary pulmonary hypertension.* N Engl J Med, 1992. **327**(2): p. 76-81.

39. Montani, D., et al., *Long-term response to calcium-channel blockers in non-idiopathic pulmonary arterial hypertension.* Eur Heart J, 2010. **31**(15): p. 1898-907.

40. Tsai, H., Y.K. Sung, and V. de Jesus Perez, *Recent advances in the management of pulmonary arterial hypertension.* F1000Res, 2016. **5**: p. 2755.

41. Chester, A.H. and M.H. Yacoub, *The role of endothelin-1 in pulmonary arterial hypertension.* Glob Cardiol Sci Pract, 2014. **2014**(2): p. 62-78.

42. Kohan, D.E., et al., *Regulation of blood pressure and salt homeostasis by endothelin.* Physiol Rev, 2011. **91**(1): p. 1-77.

43. Houde, M., L. Desbiens, and P. D'Orleans-Juste, *Endothelin-1: Biosynthesis, Signaling and Vasoreactivity.* Adv Pharmacol, 2016. **77**: p. 143-75.

44. Maruyama, H., et al., *Endothelin-Bone morphogenetic protein type 2 receptor interaction induces pulmonary artery smooth muscle cell hyperplasia in pulmonary arterial hypertension.* J Heart Lung Transplant, 2015. **34**(3): p. 468-78.

45. Pulido, T., et al., *Macitentan and morbidity and mortality in pulmonary arterial hypertension.* N Engl J Med, 2013. **369**(9): p. 809-18.

46. Chester, A.H., M.H. Yacoub, and S. Moncada, *Nitric oxide and pulmonary arterial hypertension.* Glob Cardiol Sci Pract, 2017. **2017**(2): p. 14.

47. Tsihlis, N.D., et al., *Nitric oxide inhibits vascular smooth muscle cell proliferation and neointimal hyperplasia by increasing the ubiquitination and degradation of UbcH10.* Cell Biochem Biophys, 2011. **60**(1-2): p. 89-97.

48. Mandras, S., et al., *Combination Therapy in Pulmonary Arterial Hypertension-Targeting the Nitric Oxide and Prostacyclin Pathways.* J Cardiovasc Pharmacol Ther, 2021. **26**(5): p. 453-462.

49. Barnes, H., et al., *Phosphodiesterase 5 inhibitors for pulmonary hypertension.* Cochrane Database Syst Rev, 2019. **1**: p. CD012621.

50. Ghofrani, H.A., et al., *Riociguat for the treatment of pulmonary arterial hypertension.* N Engl J Med, 2013. **369**(4): p. 330-40.

51. Ruan, C.H., et al., *Prostacyclin therapy for pulmonary arterial hypertension.* Tex Heart Inst J, 2010. **37**(4): p. 391-9.

52. Wirth, J.A. and H.I. Palevsky, *Prostacyclin Therapy for Pulmonary Arterial Hypertension Evolves Again With the Development of an Implantable Delivery System.* Chest, 2017. **152**(6): p. 1100-1102.

53. Coghlan, J.G., C. Picken, and L.H. Clapp, *Selexipag in the management of pulmonary arterial hypertension: an update.* Drug Healthc Patient Saf, 2019. **11**: p. 55-64.

54. Simonneau, G., et al., *Selexipag: an oral, selective prostacyclin receptor agonist for the treatment of pulmonary arterial hypertension.* Eur Respir J, 2012. **40**(4): p. 874-80.

55. Galie, N., et al., *Risk stratification and medical therapy of pulmonary arterial hypertension.* Eur Respir J, 2019. **53**(1).

56. Gaine, S. and V. McLaughlin, *Pulmonary arterial hypertension: tailoring treatment to risk in the current era.* Eur Respir Rev, 2017. **26**(146).

57. Yang, P. and P.B. Yu, *In Search of the Second Hit in Pulmonary Arterial Hypertension.* Circ Res, 2019. **124**(1): p. 6-8.

58. Machado, R.D., et al., *Mutations of the TGF-beta type II receptor BMPR2 in pulmonary arterial hypertension.* Hum Mutat, 2006. **27**(2): p. 121-32.

59. Sztrymf, B., et al., *Clinical outcomes of pulmonary arterial hypertension in carriers of BMPR2 mutation.* Am J Respir Crit Care Med, 2008. **177**(12): p. 1377-83.

60. Tojais, N.F., et al., *Codependence of Bone Morphogenetic Protein Receptor 2 and Transforming Growth Factor-beta in Elastic Fiber Assembly and Its Perturbation in Pulmonary Arterial Hypertension.* Arterioscler Thromb Vasc Biol, 2017. **37**(8): p. 1559-1569.

61. Chaikuad, A., et al., *Structural consequences of BMPR2 kinase domain mutations causing pulmonary arterial hypertension.* Sci Rep, 2019. **9**(1): p. 18351.

62. Kim, M.J., et al., *Clinical significance linked to functional defects in bone morphogenetic protein type 2 receptor, BMPR2.* BMB Rep, 2017. **50**(6): p. 308-317.

63. Orriols, M., M.C. Gomez-Puerto, and P. Ten Dijke, *BMP type II receptor as a therapeutic target in pulmonary arterial hypertension.* Cell Mol Life Sci, 2017. **74**(16): p. 2979-2995.

64. Hong, K.H., et al., *Genetic ablation of the BMPR2 gene in pulmonary endothelium is sufficient to predispose to pulmonary arterial hypertension.* Circulation, 2008. **118**(7): p. 722-30.

65. Waite, K.A. and C. Eng, *From developmental disorder to heritable cancer: it's all in the BMP/TGF-beta family.* Nat Rev Genet, 2003. **4**(10): p. 763-73.

66. Ma, L. and W.K. Chung, *The role of genetics in pulmonary arterial hypertension.* J Pathol, 2017. **241**(2): p. 273-280.

67. Francois, M., P. Koopman, and M. Beltrame, *SoxF genes: Key players in the development of the cardio-vascular system.* Int J Biochem Cell Biol, 2010. **42**(3): p. 445-8.

68. Arora, R., R.J. Metzger, and V.E. Papaioannou, *Multiple roles and interactions of Tbx4 and Tbx5 in development of the respiratory system.* PLoS Genet, 2012. **8**(8): p. e1002866.

69. Best, D.H., et al., *EIF2AK4 Mutations in Patients Diagnosed With Pulmonary Arterial Hypertension.* Chest, 2017. **151**(4): p. 821-828.

70. Heldin, C.H. and A. Moustakas, *Signaling Receptors for TGF-beta Family Members.* Cold Spring Harb Perspect Biol, 2016. **8**(8).

71. Massague, J., *TGF-beta signal transduction.* Annu Rev Biochem, 1998. **67**: p. 753-91.

72. Heldin, C.H., K. Miyazono, and P. ten Dijke, *TGF-beta signalling from cell membrane to nucleus through SMAD proteins.* Nature, 1997. **390**(6659): p. 465-71.

73. Cai, J., et al., *BMP signaling in vascular diseases.* FEBS Lett, 2012. **586**(14): p. 1993-2002.

74. Garcia de Vinuesa, A., et al., *BMP signaling in vascular biology and dysfunction.* Cytokine Growth Factor Rev, 2016. **27**: p. 65-79.

75. Hata, A. and Y.G. Chen, *TGF-beta Signaling from Receptors to Smads.* Cold Spring Harb Perspect Biol, 2016. **8**(9).

76. Jaffe, A.B. and A. Hall, *Rho GTPases: biochemistry and biology.* Annu Rev Cell Dev Biol, 2005. **21**: p. 247-69.

77. Maarman, G., et al., *A comprehensive review: the evolution of animal models in pulmonary hypertension research; are we there yet?* Pulm Circ, 2013. **3**(4): p. 739-56.

78. Tada, Y., et al., *Molecular effects of loss of BMPR2 signaling in smooth muscle in a transgenic mouse model of PAH.* Am J Physiol Lung Cell Mol Physiol, 2007. **292**(6): p. L1556-63.

79. Yang, X., et al., *Expression of mutant BMPR-II in pulmonary endothelial cells promotes apoptosis and a release of factors that stimulate proliferation of pulmonary arterial smooth muscle cells.* Pulm Circ, 2011. **1**(1): p. 103-10.

80. Yan, Y., et al., *Elevated levels of plasma transforming growth factor-beta1 in idiopathic and heritable pulmonary arterial hypertension.* Int J Cardiol, 2016. **222**: p. 368-374.

81. Rol, N., et al., *TGF-beta and BMPR2 Signaling in PAH: Two Black Sheep in One Family.* Int J Mol Sci, 2018. **19**(9).

82. Yu, P.B., et al., *Bone morphogenetic protein (BMP) type II receptor deletion reveals BMP ligand-specific gain of signaling in pulmonary artery smooth muscle cells.* J Biol Chem, 2005. **280**(26): p. 24443-50.

83. Rudarakanchana, N., et al., *Functional analysis of bone morphogenetic protein type II receptor mutations underlying primary pulmonary hypertension.* Hum Mol Genet, 2002. **11**(13): p. 1517-25.

84. Tielemans, B., et al., *TGFbeta and BMPRII signalling pathways in the pathogenesis of pulmonary arterial hypertension.* Drug Discov Today, 2019. **24**(3): p. 703-716.

85. Hilger, D., M. Masureel, and B.K. Kobilka, *Structure and dynamics of GPCR signaling complexes.* Nat Struct Mol Biol, 2018. **25**(1): p. 4-12.

86. Downes, G.B. and N. Gautam, *The G protein subunit gene families.* Genomics, 1999. **62**(3): p. 544-52.

87. Kristiansen, K., *Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function.* Pharmacol Ther, 2004. **103**(1): p. 21-80.

88. Smrcka, A.V., *G protein betagamma subunits: central mediators of G protein-coupled receptor signaling.* Cell Mol Life Sci, 2008. **65**(14): p. 2191-214.

89. Kauffenstein, G., et al., *Emerging role of G protein-coupled receptors in microvascular myogenic tone.* Cardiovasc Res, 2012. **95**(2): p. 223-32.

90. Strassheim, D., et al., *A current view of G protein-coupled receptor - mediated signaling in pulmonary hypertension: finding opportunities for therapeutic intervention.* Vessel Plus, 2018. **2**.

91. Oka, M., et al., *Rho kinase-mediated vasoconstriction is important in severe occlusive pulmonary arterial hypertension in rats.* Circ Res, 2007. **100**(6): p. 923-9.

92. Zhang, Y. and S. Wu, *Effects of fasudil on pulmonary hypertension in clinical practice.* Pulm Pharmacol Ther, 2017. **46**: p. 54-63.

93. Antoniu, S.A., *Targeting RhoA/ROCK pathway in pulmonary arterial hypertension.* Expert Opin Ther Targets, 2012. **16**(4): p. 355-63.

94. Abe, K., et al., *Long-term treatment with a Rho-kinase inhibitor improves monocrotaline-induced fatal pulmonary hypertension in rats.* Circ Res, 2004. **94**(3): p. 385-93.

95. Fujita, H., et al., *Acute vasodilator effects of inhaled fasudil, a specific Rho-kinase inhibitor, in patients with pulmonary arterial hypertension.* Heart Vessels, 2010. **25**(2): p. 144-9.

96. Vaidya, B., et al., *Advances in treatment of pulmonary arterial hypertension: patent review.* Expert Opin Ther Pat, 2017. **27**(8): p. 907-918.

97. Zhao, Y., et al., *An essential role for Wnt/beta-catenin signaling in mediating hypertensive heart disease.* Sci Rep, 2018. **8**(1): p. 8996.

98. Boucherat, O. and S. Bonnet, *MicroRNA signature of end-stage idiopathic pulmonary arterial hypertension: clinical correlations and regulation of WNT signaling.* J Mol Med (Berl), 2016. **94**(8): p. 849-51.

99. Awad, K.S., et al., *Novel signaling pathways in pulmonary arterial hypertension (2015 Grover Conference Series).* Pulm Circ, 2016. **6**(3): p. 285-94.

100. Yuan, K., et al., *Loss of Endothelium-Derived Wnt5a Is Associated With Reduced Pericyte Recruitment and Small Vessel Loss in Pulmonary Arterial Hypertension.* Circulation, 2019. **139**(14): p. 1710-1724.

101. de Jesus Perez, V., et al., *Targeting the Wnt signaling pathways in pulmonary arterial hypertension.* Drug Discov Today, 2014. **19**(8): p. 1270-6.

102. Lahm, T., R.M. Tuder, and I. Petrache, *Progress in solving the sex hormone paradox in pulmonary hypertension.* Am J Physiol Lung Cell Mol Physiol, 2014. **307**(1): p. L7-26.

103. Heldring, N., et al., *Estrogen receptors: how do they signal and what are their targets.* Physiol Rev, 2007. **87**(3): p. 905-31.

104. Tofovic, S.P. and E.K. Jackson, *Estradiol Metabolism: Crossroads in Pulmonary Arterial Hypertension.* Int J Mol Sci, 2019. **21**(1).

105. Frump, A.L., et al., *17beta-Estradiol and estrogen receptor alpha protect right ventricular function in pulmonary hypertension via BMPR2 and apelin.* J Clin Invest, 2021. **131**(6).

106. Chelladurai, P., W. Seeger, and S.S. Pullamsetti, *Epigenetic mechanisms in pulmonary arterial hypertension: the need for global perspectives.* Eur Respir Rev, 2016. **25**(140): p. 135-40.

107. Nakao, M., *Epigenetics: interaction of DNA methylation and chromatin.* Gene, 2001. **278**(1-2): p. 25-31.

108. Gilmore, J.M. and M.P. Washburn, *Deciphering the combinatorial histone code.* Nat Methods, 2007. **4**(6): p. 480-1.

109. Xu, X.F., F. Cheng, and L.Z. Du, *Epigenetic regulation of pulmonary arterial hypertension.* Hypertens Res, 2011. **34**(9): p. 981-6.

110. Touyz, R.M. and E.L. Schiffrin, *Reactive oxygen species in vascular biology: implications in hypertension.* Histochem Cell Biol, 2004. **122**(4): p. 339-52.

111. Archer, S.L., et al., *Epigenetic attenuation of mitochondrial superoxide dismutase 2 in pulmonary arterial hypertension: a basis for excessive cell proliferation and a new therapeutic target.* Circulation, 2010. **121**(24): p. 2661-71.

112. Paulin, R., et al., *Sirtuin 3 deficiency is associated with inhibited mitochondrial function and pulmonary arterial hypertension in rodents and humans.* Cell Metab, 2014. **20**(5): p. 827-839.

113. Zhao, L., et al., *Histone deacetylation inhibition in pulmonary hypertension: therapeutic potential of valproic acid and suberoylanilide hydroxamic acid.* Circulation, 2012. **126**(4): p. 455-67.

114. Bogaard, H.J., et al., *Suppression of histone deacetylases worsens right ventricular dysfunction after pulmonary artery banding in rats.* Am J Respir Crit Care Med, 2011. **183**(10): p. 1402-10.

115. Donati, B., E. Lorenzini, and A. Ciarrocchi, *BRD4 and Cancer: going beyond transcriptional regulation.* Mol Cancer, 2018. **17**(1): p. 164.

116. Khan, Y.M., et al., *Brd4 is essential for IL-1beta-induced inflammation in human airway epithelial cells.* PLoS One, 2014. **9**(4): p. e95051.

117. Meloche, J., et al., *Bromodomain-Containing Protein 4: The Epigenetic Origin of Pulmonary Arterial Hypertension.* Circ Res, 2015. **117**(6): p. 525-35.

118. Van der Feen, D.E., et al., *Multicenter Preclinical Validation of BET Inhibition for the Treatment of Pulmonary Arterial Hypertension.* Am J Respir Crit Care Med, 2019. **200**(7): p. 910-920.

119. Zeng, Y., et al., *Screening of Hub Genes Associated with Pulmonary Arterial Hypertension by Integrated Bioinformatic Analysis.* Biomed Res Int, 2021. **2021**: p. 6626094.

120. Mura, M., et al., *Osteopontin lung gene expression is a marker of disease severity in pulmonary arterial hypertension.* Respirology, 2019. **24**(11): p. 1104-1110.

121. Bull, T.M., et al., *Gene microarray analysis of peripheral blood cells in pulmonary arterial hypertension.* Am J Respir Crit Care Med, 2004. **170**(8): p. 911-9.

122. Rajkumar, R., et al., *Genomewide RNA expression profiling in lung identifies distinct signatures in idiopathic pulmonary arterial hypertension and secondary pulmonary hypertension.* Am J Physiol Heart Circ Physiol, 2010. **298**(4): p. H1235-48.

123. Atkinson, C., et al., *Primary pulmonary hypertension is associated with reduced pulmonary vascular expression of type II bone morphogenetic protein receptor.* Circulation, 2002. **105**(14): p. 1672-8.

124. Rhodes, C.J., et al., *RNA Sequencing Analysis Detection of a Novel Pathway of Endothelial Dysfunction in Pulmonary Arterial Hypertension.* Am J Respir Crit Care Med, 2015. **192**(3): p. 356-66.

125. Gorr, M.W., et al., *Transcriptomic analysis of pulmonary artery smooth muscle cells identifies new potential therapeutic targets for idiopathic pulmonary arterial hypertension.* Br J Pharmacol, 2020. **177**(15): p. 3505-3518.

126. Ulrich, S., et al., *Peripheral blood B lymphocytes derived from patients with idiopathic pulmonary arterial hypertension express a different RNA pattern compared with healthy controls: a cross sectional study.* Respir Res, 2008. **9**: p. 20.

127. Rana, T.M., *Illuminating the silence: understanding the structure and function of small RNAs.* Nat Rev Mol Cell Biol, 2007. **8**(1): p. 23-36.

128. Filipowicz, W., S.N. Bhattacharyya, and N. Sonenberg, *Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?* Nat Rev Genet, 2008. **9**(2): p. 102-14.

129. Gou, D., et al., *miR-210 has an antiapoptotic effect in pulmonary artery smooth muscle cells during hypoxia.* Am J Physiol Lung Cell Mol Physiol, 2012. **303**(8): p. L682-91.

130. Rhodes, C.J., et al., *Reduced microRNA-150 is associated with poor survival in pulmonary arterial hypertension.* Am J Respir Crit Care Med, 2013. **187**(3): p. 294-302.

131. Schlosser, K., R.J. White, and D.J. Stewart, *miR-26a linked to pulmonary hypertension by global assessment of circulating extracellular microRNAs.* Am J Respir Crit Care Med, 2013. **188**(12): p. 1472-5.

132. Hsu, E., et al., *Lung tissues in patients with systemic sclerosis have gene expression patterns unique to pulmonary fibrosis and pulmonary hypertension.* Arthritis Rheum, 2011. **63**(3): p. 783-94.

133. Yang, L., et al., *Activation of BK Channels Prevents Hepatic Stellate Cell Activation and Liver Fibrosis Through the Suppression of TGFbeta1/SMAD3 and JAK/STAT3 Profibrotic Signaling Pathways.* Front Pharmacol, 2020. **11**: p. 165.

134. Hemnes, A., et al., *Role of biomarkers in evaluation, treatment and clinical studies of pulmonary arterial hypertension.* Pulm Circ, 2020. **10**(4): p. 2045894020957234.

135. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics.* Nat Rev Genet, 2009. **10**(1): p. 57-63.

136. Kariotis, S., et al., *Biological heterogeneity in idiopathic pulmonary arterial hypertension identified through unsupervised transcriptomic profiling of whole blood.* Nat Commun, 2021. **12**(1): p. 7104.

137. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

138. Hasin, Y., M. Seldin, and A. Lusis, *Multi-omics approaches to disease.* Genome Biol, 2017. **18**(1): p. 83.

139. Burger, C.D., et al., *The evolution of prostacyclins in pulmonary arterial hypertension: from classical treatment to modern management.* Am J Manag Care, 2016. **22**(1 Suppl): p. S3-15.

140. McGoon, M.D. and D.P. Miller, *REVEAL: a contemporary US pulmonary arterial hypertension registry.* Eur Respir Rev, 2012. **21**(123): p. 8-18.

141. Leopold, J.A. and B.A. Maron, *Molecular Mechanisms of Pulmonary Vascular Remodeling in Pulmonary Arterial Hypertension.* Int J Mol Sci, 2016. **17**(5).

142. Vaillancourt, M., et al., *Adaptation and remodelling of the pulmonary circulation in pulmonary hypertension.* Can J Cardiol, 2015. **31**(4): p. 407-15.

143. Gomez, D. and G.K. Owens, *Smooth muscle cell phenotypic switching in atherosclerosis.* Cardiovasc Res, 2012. **95**(2): p. 156-64.

144. Zabini, D., et al., *Loss of SMAD3 Promotes Vascular Remodeling in Pulmonary Arterial Hypertension via MRTF Disinhibition.* Am J Respir Crit Care Med, 2018. **197**(2): p. 244-260.

145. Xing, X.Q., et al., *Sphingosine kinase 1/sphingosine 1-phosphate signalling pathway as a potential therapeutic target of pulmonary hypertension.* Int J Clin Exp Med, 2015. **8**(8): p. 11930-5.

146. Guo, X. and S.Y. Chen, *Transforming growth factor-beta and smooth muscle differentiation.* World J Biol Chem, 2012. **3**(3): p. 41-52.

147. He, W. and C. Dai, *Key Fibrogenic Signaling.* Curr Pathobiol Rep, 2015. **3**(2): p. 183-192.

148. Bai, X., et al., *GRAF3 serves as a blood volume-sensitive rheostat to control smooth muscle contractility and blood pressure.* Small GTPases, 2020. **11**(3): p. 194-203.

149. Qiu, P., X.H. Feng, and L. Li, *Interaction of Smad3 and SRF-associated complex mediates TGF-beta1 signals to regulate SM22 transcription during myofibroblast differentiation.* J Mol Cell Cardiol, 2003. **35**(12): p. 1407-20.

150. Yamanaka, M., et al., *Sphingosine kinase 1 (SPHK1) is induced by transforming growth factor-beta and mediates TIMP-1 up-regulation.* J Biol Chem, 2004. **279**(52): p. 53994-4001.

151. Kluk, M.J. and T. Hla, *Signaling of sphingosine-1-phosphate via the S1P/EDG-family of G-protein-coupled receptors.* Biochim Biophys Acta, 2002. **1582**(1-3): p. 72-80.

152. Cencetti, F., et al., *Transforming growth factor-beta1 induces transdifferentiation of myoblasts into myofibroblasts via up-regulation of sphingosine kinase-1/S1P3 axis.* Mol Biol Cell, 2010. **21**(6): p. 1111-24.

153. Gairhe, S., et al., *Sphingosine-1-phosphate is involved in the occlusive arteriopathy of pulmonary arterial hypertension.* Pulm Circ, 2016. **6**(3): p. 369-80.

154. Chen, J., et al., *The sphingosine kinase 1/sphingosine-1-phosphate pathway in pulmonary arterial hypertension.* Am J Respir Crit Care Med, 2014. **190**(9): p. 1032-43.

155. Hutchings, K.M., et al., *Pharmacokinetic optimitzation of CCG-203971: Novel inhibitors of the Rho/MRTF/SRF transcriptional pathway as potential antifibrotic therapeutics for systemic scleroderma.* Bioorg Med Chem Lett, 2017. **27**(8): p. 1744-1749.

156. Yu, X., et al., *Activation of G protein-coupled estrogen receptor 1 induces coronary artery relaxation via Epac/Rap1-mediated inhibition of RhoA/Rho kinase pathway in parallel with PKA.* PLoS One, 2017. **12**(3): p. e0173085.

157. Calvier, L., et al., *Chronic TGF-beta1 Signaling in Pulmonary Arterial Hypertension Induces Sustained Canonical Smad3 Pathways in Vascular Smooth Muscle Cells.* Am J Respir Cell Mol Biol, 2019. **61**(1): p. 121-123.

158. Tang, Y., et al., *Mechanisms of TGF-beta-induced differentiation in human vascular smooth muscle cells.* J Vasc Res, 2011. **48**(6): p. 485-94.

159. Chen, S., et al., *RhoA modulates Smad signaling during transforming growth factor-beta-induced smooth muscle differentiation.* J Biol Chem, 2006. **281**(3): p. 1765-70.

160. Parmacek, M.S., *Myocardin-related transcription factor-A: mending a broken heart.* Circ Res, 2010. **107**(2): p. 168-70.

161. Wang, J., et al., *SphK1/S1P mediates TGF-beta1-induced proliferation of pulmonary artery smooth muscle cells and its potential mechanisms.* Pulm Circ, 2019. **9**(1): p. 2045894018816977.

162. Zhao, J., et al., *TGF-beta/SMAD3 Pathway Stimulates Sphingosine-1 Phosphate Receptor 3 Expression: IMPLICATION OF SPHINGOSINE-1 PHOSPHATE RECEPTOR 3 IN LUNG ADENOCARCINOMA PROGRESSION.* J Biol Chem, 2016. **291**(53): p. 27343-27353.

163. Wamhoff, B.R., et al., *Sphingosine-1-phosphate receptor subtypes differentially regulate smooth muscle cell phenotype.* Arterioscler Thromb Vasc Biol, 2008. **28**(8): p. 1454-61.

164. Li, M., et al., *Emergence of fibroblasts with a proinflammatory epigenetically altered phenotype in severe hypoxic pulmonary hypertension.* J Immunol, 2011. **187**(5): p. 2711-22.

165. Savai, R., et al., *Immune and inflammatory cell involvement in the pathology of idiopathic pulmonary arterial hypertension.* Am J Respir Crit Care Med, 2012. **186**(9): p. 897-908.

166. Soon, E., et al., *Elevated levels of inflammatory cytokines predict survival in idiopathic and familial pulmonary arterial hypertension.* Circulation, 2010. **122**(9): p. 920-7.

167. Phan, S.H., *Genesis of the myofibroblast in lung injury and fibrosis.* Proc Am Thorac Soc, 2012. **9**(3): p. 148-52.

168. Sobin, S.S., et al., *Changes in arteriole in acute and chronic hypoxic pulmonary hypertension and recovery in rat.* J Appl Physiol Respir Environ Exerc Physiol, 1983. **55**(5): p. 1445-55.

169. Stenmark, K.R., et al., *Hypoxic activation of adventitial fibroblasts: role in vascular remodeling.* Chest, 2002. **122**(6 Suppl): p. 326S-334S.

170. Chelladurai, P., W. Seeger, and S.S. Pullamsetti, *Matrix metalloproteinases and their inhibitors in pulmonary hypertension.* Eur Respir J, 2012. **40**(3): p. 766-82.

171. Shi, Y., et al., *Role of matrix metalloproteinases and their tissue inhibitors in the regulation of coronary cell migration.* Arterioscler Thromb Vasc Biol, 1999. **19**(5): p. 1150-5.

172. Gao, P.J., et al., *Differentiation of vascular myofibroblasts induced by transforming growth factor-beta1 requires the involvement of protein kinase Calpha.* J Mol Cell Cardiol, 2003. **35**(9): p. 1105-12.

173. Hiepen, C., et al., *BMPR2 acts as a gatekeeper to protect endothelial cells from increased TGFbeta responses and altered cell mechanics.* PLoS Biol, 2019. **17**(12): p. e3000557.

174. Pickup, M.W., et al., *BMPR2 loss in fibroblasts promotes mammary carcinoma metastasis via increased inflammation.* Mol Oncol, 2015. **9**(1): p. 179-91.

175. Ogo, T., et al., *Inhibition of overactive transforming growth factor-beta signaling by prostacyclin analogs in pulmonary arterial hypertension.* Am J Respir Cell Mol Biol, 2013. **48**(6): p. 733-41.

176. Harper, R.L., et al., *BMPR2 gene therapy for PAH acts via Smad and non-Smad signalling.* Respirology, 2016. **21**(4): p. 727-33.

177. Gall, H., et al., *The Giessen Pulmonary Hypertension Registry: Survival in pulmonary hypertension subgroups.* J Heart Lung Transplant, 2017. **36**(9): p. 957-967.

178. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods.* Biostatistics, 2007. **8**(1): p. 118-27.

179. Vivar, J.C., et al., *Redundancy control in pathway databases (ReCiPa): an application for improving gene-set enrichment analysis in Omics studies and "Big data" biology.* OMICS, 2013. **17**(8): p. 414-22.

180. Zhao, E., H. Xie, and Y. Zhang, *Identification of Differentially Expressed Genes Associated with Idiopathic Pulmonary Arterial Hypertension by Integrated Bioinformatics Approaches.* J Comput Biol, 2020.

181. Kimura, M., et al., *A genome-wide association analysis identifies PDE1A|DNAJC10 locus on chromosome 2 associated with idiopathic pulmonary arterial hypertension in a Japanese population.* Oncotarget, 2017. **8**(43): p. 74917-74926.

182. Paulin, R., J. Meloche, and S. Bonnet, *STAT3 signaling in pulmonary arterial hypertension.* JAKSTAT, 2012. **1**(4): p. 223-33.

183. Ranchoux, B., et al., *Endothelial dysfunction in pulmonary arterial hypertension: an evolving landscape (2017 Grover Conference Series).* Pulm Circ, 2018. **8**(1): p. 2045893217752912.

184. Bertero, T., et al., *Vascular stiffness mechanoactivates YAP/TAZ-dependent glutaminolysis to drive pulmonary hypertension.* J Clin Invest, 2016. **126**(9): p. 3313-35.

185. Morris, H.E., et al., *Notch3 signalling and vascular remodelling in pulmonary arterial hypertension.* Clin Sci (Lond), 2019. **133**(24): p. 2481-2498.

186. Guignabert, C., F. de Man, and M. Lombes, *ACE2 as therapy for pulmonary arterial hypertension: the good outweighs the bad.* Eur Respir J, 2018. **51**(6).

187. Sutliff, R.L., B.Y. Kang, and C.M. Hart, *PPARgamma as a potential therapeutic target in pulmonary hypertension.* Ther Adv Respir Dis, 2010. **4**(3): p. 143-60.

188. Garat, C.V., et al., *Inhibition of phosphatidylinositol 3-kinase/Akt signaling attenuates hypoxia-induced pulmonary artery remodeling and suppresses CREB depletion in arterial smooth muscle cells.* J Cardiovasc Pharmacol, 2013. **62**(6): p. 539-48.

189. Yu, X., et al., *Dacomitinib, a new pan-EGFR inhibitor, is effective in attenuating pulmonary vascular remodeling and pulmonary hypertension.* Eur J Pharmacol, 2019. **850**: p. 97-108.

190. Hill, R.P., R.S. Bush, and P. Yeung, *The effect of anaemia on the fraction of hypoxic cells in an experimental tumour.* Br J Radiol, 1971. **44**(520): p. 299-304.

191. Hill, M.A. and G.A. Meininger, *Arteriolar vascular smooth muscle cells: mechanotransducers in a complex environment.* Int J Biochem Cell Biol, 2012. **44**(9): p. 1505-10.

192. Sztuka, K. and M. Jasinska-Stroschein, *Animal models of pulmonary arterial hypertension: A systematic review and meta-analysis of data from 6126 animals.* Pharmacol Res, 2017. **125**(Pt B): p. 201-214.

193. Stenmark, K.R., et al., *Animal models of pulmonary arterial hypertension: the hope for etiological discovery and pharmacological cure.* Am J Physiol Lung Cell Mol Physiol, 2009. **297**(6): p. L1013-32.

194. Saygin, D., et al., *Transcriptional profiling of lung cell populations in idiopathic pulmonary arterial hypertension.* Pulm Circ, 2020. **10**(1).

195. Asosingh, K., et al., *Single-cell transcriptomic profile of human pulmonary artery endothelial cells in health and pulmonary arterial hypertension.* Sci Rep, 2021. **11**(1): p. 14714.

196. Amamoto, R., et al., *Probe-Seq enables transcriptional profiling of specific cell types from heterogeneous tissue by RNA-based isolation.* Elife, 2019. **8**.

197. Xiao, L. and J. Guo, *Single-Cell in Situ RNA Analysis With Switchable Fluorescent Oligonucleotides.* Front Cell Dev Biol, 2018. **6**: p. 42.