THE EFFECT OF MIR-29 AND TGFB ON FIBROSIS

by

Tyler J. Creamer

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

June, 2018

© 2018 Tyler J. Creamer
ABSTRACT

Fibrosis is defined as excess accumulation of extracellular matrix (ECM) that impedes normal function of the parenchyma and is the cause of up to 45% of deaths in the developed world. With no favorable therapeutic strategy, development of new therapeutic interventions is necessary. Transforming growth factor beta (TGF-β), a potent pro-fibrotic agonist, promotes deposition of ECM in part through upregulation of collagens, the major structural components of fibrotic scars. Loss of miR-29, a potent anti-fibrotic miRNA family that is regulated by TGF-β and that targets multiple collagen species, is often associated with fibrosis. In previous work, our laboratory showed that maintenance of miR-29a via self-complementary adeno-associated virus (scAAV) transduced to the liver was sufficient to ameliorate liver fibrosis in a carbon tetrachloride (CCl4)-induced liver fibrosis murine model. Conversely, targeting of miR-29b to hepatocytes was unable to recapitulate the scAAV result with miR-29a. To systematically interrogate the efficacy of miR-29 as a therapeutic agent in different cell populations, we generated a tetracycline-regulated miR-29a and green fluorescent protein (GFP) reporter transgenic mouse to temporally and spatially control miR-29a expression. Using 2 different constitutive reverse tetracycline-controlled transactivators (rtTA), transgene expression can be induced in many tissues and cell types with doxycycline. Unexpectedly, the induction of miR-29a to more than 3-fold normal levels across the liver was not sufficient to prevent CCl4-induced liver fibrosis. Upon investigation of the spatial expression of the transgene, we saw exclusion of GFP from the Acta2⁺, collagen-producing cells. Additional evidence suggested that the miR-29 transgene is being
epigenetically regulated in high collagen-producing cells, which would be the desired therapeutic target of this miR-29a-based therapy.

Using high throughput RNA sequencing and a physical interaction pulldown assay, we were able to infer function of miR-29a in the different liver cell populations in vitro in the presence or absence of synthetic injury. We showed that miR-29a interacted with ECM components and epigenetic regulators and was able to normalize TGF-β induction of paracrine signals as well as non-coding RNAs. Collectively, my data indicate that miR-29a is a potent therapeutic when expressed in specific cells during fibrotic injury by regulating more than collagens.

Advisor: Daniel Warren, Ph.D.

Reader: Elena Gallo MacFarlane, Ph.D.
PREFACE

There are many people to thank for their support and guidance during graduate school. First, I want to thank my mentor, Dan Warren. Dan has been a great mentor these past few years both, professionally and personally. I appreciate all of the guidance in understanding and interpreting the biology, as well as how research and a lab is run. I also appreciate all the discussions over the years about politics, raising kids, and how to fix any and everything around the house and with my cars. Dan has been a good friend and I hope to continue that friendship in the years to come. I’d like to thank previous members of the lab Matt Knabel and Kalyani Ramachandran for their mentorship and guidance.

I would like to thank Dave Valle and Sandy Muscelli from the Human Genetics program. Being able to participate in this program has been incredible and I thank them very much for the experience. To my thesis committee, Hal Dietz, Sarah Wheelan, Jeff Corden, Kirby Smith, and Anthony Leung, I want to thank you for your guidance and support to complete my PhD. To my Human Genetics classmates, I want to thank you for going through the experience with me. I especially want to thank Jevon Cutler. You were indispensable as a good friend and someone to talk to throughout the whole process. I always looked forward to our lunch time talks about science, the lab, and life. I wish you and your family all the best in your next chapter in Boston.

The Dietz lab has been a source of collaborators, resources, friends, and good times. Over the years, I have worked with many people that have come through the Dietz lab and everyone has always been helpful and fun. I want to thank Juan Calderon, Shira Ziegler, Joe Shin, James Beckett, Manny Seman-Senderos, Nicole Wilson, Cailtin
Bowen, Varun Nagpal, Ben Kang, and Sara Cooke, for being good friends and lab mates. I also want to thank the MacFarlane lab, Elena MacFarlane and Rusty Bagirzadeh. I look forward to joining lab for the next stage of my career.

I also want to thank Jeff Corden, who has been an extraordinary scientific mentor over the years. Jeff gave me my first taste of research when I worked in his lab during high school and has been helping me build the necessary skills for my career ever since.

Last but not least by any means, I would like to thank my family. My parents, Bob and Suzi, have always been supportive of my choices and an inspiration for to achieve more. My brother, Nick, and sister, Morgan, have also been there as those I can call on talk to. My kids, Henry and John, are the wonderful reasons to come home in the evening. But most of all, my wife, Georgie, has been my rock. Thank you so much for everything. I couldn’t have done it without you.
TABLE OF CONTENTS

ABSTRACT ................................................................................................................................................ II

PREFACE ................................................................................................................................................. IV

TABLE OF CONTENTS .............................................................................................................................. VI

LIST OF FIGURES ..................................................................................................................................... XI

CHAPTER 1 - FIBROSIS AND MIRNAS: THE BIG ROLES OF SMALL RNAs IN FIBROPROLIFERATIVE DISEASE 1

I. WHAT IS FIBROSIS? ................................................................................................................................ 2

II. COLLAGEN AND THE EXTRACELLULAR MATRIX ...................................................................................... 3

III. TREATMENTS FOR A DISEASE OF EVERY ORGAN ........................................................................................... 4

IV. MICRORNAs IN NORMAL AND PATHOPHYSIOLOGY ...................................................................................... 5

V. THERAPEUTIC DELIVERY OF miR-29A AMELIORATES MURINE LIVER FIBROSIS ...................................................... 6

CHAPTER 2 – THERAPEUTIC ASSESSMENT OF HEPATOCYTE-TARGETED MIR-29B OLIGONUCLEOTIDE IN A
MURINE MODEL OF LIVER FIBROSIS .......................................................................................................12

I. ABSTRACT ........................................................................................................................................13

II. INTRODUCTION ..........................................................................................................................................14

III. RESULTS ..............................................................................................................................................17

MiR-29a and miR-29b are equivalent regulators of collagens and other pro-fibrotic factors ..............17
Hepatocyte-directed miR-29b oligonucleotide does not alleviate CCl4-induced liver fibrosis...........18
Qualitative difference in collagen composition in miR-29b oligo treated mice...............................19

IV. DISCUSSION .......................................................................................................................................27

Therapeutic restoration of miR-29 may not be sufficient to ameliorate hepatic fibrosis...............27
Qualitative difference in collagen fibers in miR-29b treated animals..............................................28
Which cell type with increased miR-29 levels is sufficient to produce antifibrotic effects? ..........28

V. METHODS .......................................................................................................................................30

Small RNA transfection and TGFβ stimulation of HuH7 and LX2 cells ...........................................30
CCl₄-induced liver fibrosis in mice ....................................................................................................30
Histology preparation, staining, and analysis ..................................................................................31
RNA isolation and RT-qPCR analysis ................................................................................................31
Western Blot for protein analysis ....................................................................................................32
Serum Analysis ..................................................................................................................................33

CHAPTER 3 - EPGENETIC REGULATION OF COL1A1 LOCUS SILENCES CIS-TRANSGENE INSERTIONS ....34

I. ABSTRACT ........................................................................................................................................35

II. INTRODUCTION ...............................................................................................................................36

III. RESULTS .........................................................................................................................................39

TRE_eGFP-miR29a Construct ..........................................................................................................39
Col1A1tm1[TRE_EGFP-miR29a] expression during CCl₄-induced fibrosis model ....................................40
TGFβ stimulation in MEFs Col1A1tm1[TRE_EGFP-miR29a] +/-; M2-rtTA+/− mice ........................................42
Metanalysis of TGFβ-stimulated NMuMg cells ................................................................................43
Epigenetic regulation of Col1a1 locus conserved in human cells and produces novel eRNA ..............44

IV. DISCUSSION ...................................................................................................................................55

Expression of Tg29, in all but Acta2+ cells, during injury is not protective against hepatic fibrosis.....55
Epigenetic regulation of the Col1a1 locus responsible for proximal gene expression dysregulation ..56
Epigenetic regulation of Col1a1 in other mouse models ..................................................................56

V. METHODS .......................................................................................................................................59

Generation of Col1A1tm1[TRE_EGFP-miR29a] mice ............................................................................59
CCl₄-induced liver fibrosis in mice ....................................................................................................59
Histology preparation and staining ..................................................................................................60
Immunofluorescence of FFPE tissue sections ..................................................................................60
CHAPTER 4 - IDENTIFICATION OF TGFß AND MIR-29A TARGETS USING TANDEM HIGH-THROUGHPUT RNA ANALYSIS

I. ABSTRACT ........................................................................................................................................66

II. INTRODUCTION ...............................................................................................................................67

III. RESULTS .........................................................................................................................................69

Differential effect of TGF-ß on contraction and migration in epithelial cells and fibroblasts ..........69

Total RNA-Seq in HuH7 and LX2 reveals the potency of TGFß and miR-29a on the transcriptome....71

TGFß shows common regulation of ECM components and cell-type specific regulation of different pathways in HuH7 and LX2 cells......................................................................................................72

Expression data, combined with biotinylated miR-29a pull-down, identifies novel regulatory mechanisms of collagen synthesis by miR-29a ........................................................................................74

MIR-29a differentially regulates TGF-ß-mediated injury in cells of different lineages ...............77

IV. DISCUSSION ...................................................................................................................................96

In vitro versus in vivo: TGFß stimulation is a good surrogate for in vivo fibrotic injury ..........96

MiR-29a targets more than just fibrosis: miR-29a in cancer ............................................................96

ncRNAs: Functional or Consequential? ............................................................................................98

V. METHODS .....................................................................................................................................100

Small RNA transfection and TGFß stimulation of HuH7 and LX2 cells .................................100

RT-qPCR .......................................................................................................................................100

Collagen gel contraction assay .....................................................................................................101

Scratch Mobility Assay ..................................................................................................................101
Western Blot for protein analysis ................................................................. 102
RNA-sequencing preparation and analysis .................................................... 102
Biotinylated miRNA pulldown assay and gene expression array ....................... 103

VI. ACKNOWLEDGEMENTS ............................................................................. 106

CONCLUSION AND FUTURE DIRECTIONS .................................................... 107

APPENDIX - CAPN9<sup><sup>-/-</sup></sup> MICE ARE PROTECTED AGAINST EMT-MEDIATED FIBROSIS IN MULTIPLE MURINE MODELS OF INJURY ................................................................. 110

I. EXPLANATION OF PROJECT PARTICIPATION ............................................. 111

II. ABSTRACT FROM MANUSCRIPT ............................................................. 113

III. INTRODUCTION FROM MANUSCRIPT ................................................. 114

IV. ABBREVIATED RESULTS FROM MANUSCRIPT ...................................... 117

Capn9<sup><sup>-/-</sup></sup> mice are resistant to bleomycin induced lung fibrosis ......................... 117
Capn9<sup><sup>-/-</sup></sup> mice are protected from carbon tetrachloride (CCl<sub>4</sub>) induced liver fibrosis ................................ 118
Capn9<sup><sup>-/-</sup></sup> mice are protected from angiotensin-2 (Ang II) induced heart fibrosis ................... 119

V. DISCUSSION FROM MANUSCRIPT ......................................................... 123

VI. ABBREVIATED METHODS FROM MANUSCRIPT .................................. 127

Small molecule inhibition of calpains .............................................................. 127
RNA in situ hybridization .............................................................................. 127
Histological analysis ..................................................................................... 127
RNA analysis .............................................................................................. 128
Bleomycin-induced pulmonary fibrosis .......................................................... 129
Angiotensin II-induced cardiac fibrosis .......................................................... 130
CCl<sub>4</sub>-induced hepatic fibrosis ..................................................................... 130
Immunofluorescence .................................................................................... 131
Collagen quantification with hydroxyproline assay ........................................ 132
REFERENCES ......................................................................................................................................................................................... 133

CURRICULUM VITAE ................................................................................................................................................................................ 141
LIST OF FIGURES

FIGURE 1.1 – INFLUENCE OF TGFβ ON FIBROBLAST ACTIVATION ................................................................. 9

FIGURE 1.2 – SCHEMATIC OF CONTROL AND MIR-29A CONTAINING SCAAV AND TROPISM OF AAV IN MURINE LIVER SECTIONS ........................................................................................................................ 10

FIGURE 1.3 – PREVENTION AND REVERSAL MODELS OF THERAPEUTIC INTERVENTION .................. 11

FIGURE 2.1 – EXPRESSION CHANGES OF COL1A2 AND FBN1 ± TGFβ ± MIR-29A OR MIR-29B ............. 21

FIGURE 2.2 – DESIGN AND ASSESSMENT OF MIR-29B OLIGO IN LIVER FIBROSIS MODEL .................... 22

FIGURE 2.3 – MASSON TRICHROME STAIN OF LIVERS AFTER 12 WEEKS OF CCL4 ................................. 23

FIGURE 2.4 – EXAMPLE OF APERIO IMAGESCOPE MARKUP OF COLLAGEN DEPOSITION IN LIVER TISSUE ............................................................................................................................................................... 24

FIGURE 2.5 – IN SILICO QUANTIFICATION OF COLLAGEN CONTENT FROM LIVER SECTIONS ............. 25

FIGURE 2.6 – MOLECULAR ANALYSIS OF FIBROSIS-ASSOCIATED GENES IN CCL4-TREATED LIVERS ...... 26

FIGURE 3.1 – EXPRESSION PATTERN OF COL1A1TM1(TRE_EGFP-MIR29A) WITH M2-RTTA .................... 45

FIGURE 3.2 – EXPRESSION PATTERN OF COL1A1TM1(TRE_EGFP-MIR29A) WITH RTTA3 ......................... 46

FIGURE 3.3 – MATURE MIR-29A LEVELS AFTER DOXYCYCLINE INDUCTION ............................................ 47

FIGURE 3.4 – HISTOLOGICAL AND MOLECULAR ANALYSIS OF COL1A1TM1[TRE_EGFP-MIR29A] ± 6WK CCL4 ........................................................ .................................................. 48

FIGURE 3.5 – REGULATION OF PRO-FIBROTIC, MIR-29 TARGETS IN COL1A1TM1[TRE_EGFP-MIR29A] MICE ........... 49

FIGURE 3.6 – DISTRIBUTION OF ACTA2+ CELLS IN COL1A1TM1[TRE_EGFP-MIR29A] MICE ......................... 50

FIGURE 3.7 – MEFS STIMULATED WITH TGFβ – 10X MAGNIFICATION ................................................ 51

FIGURE 3.8 – MEFS STIMULATED WITH TGFβ – 20X MAGNIFICATION ................................................ 52

FIGURE 3.9 – EPIGENETIC CHANGES TO COL1A1 LOCUS WITH TGFβ STIMULATION .......................... 53
CHAPTER 1 - Fibrosis and miRNAs: The Big Roles of Small RNAs in Fibroproliferative Disease
I. What is Fibrosis?

Fibrosis, which can be considered the result of a normal wound healing process gone awry, is characterized by excess accumulation of extracellular matrix (ECM) that disrupts the normal function of the parenchyma. It can be caused by many different insults, including repeated chemical or physical injury, exposure to chemotherapeutic or toxic chemicals, viral or parasitic infection, or underlying genetic predisposition. Some of the pathological mechanisms by which organ or tissue function is disrupted are increased tissue stiffness, diverted blood flow, or alteration of the cytokine milieu in injured tissue. Fibrosis can affect all solid organs and is estimated to be the cause of 45% of death in the developed world. With no compelling therapeutic options, fibrosis is a major physical and financial burden on society; understanding the balance between wound healing and fibrosis is essential to finding an effective therapeutic strategy.

Cardiac, pulmonary, renal, dermatologic, and hepatic fibrosis have been extensively studied and share many common histopathological features including excess collagen deposition, activation and migration of alpha smooth muscle actin (ACTA2) expressing myofibroblasts to site of injury, and increased secretion of active cytokines such as transforming growth factor beta (TGF-β) and epidermal growth factor (EGF). Other molecular and histological characteristics appear to be organ and tissue specific. For example, cardiac fibrosis after a myocardial infarct has substantial perivascular fibrosis, with milder interstitial fibrosis, resulting from collagen type I and III deposition, while hepatic fibrosis from carbon tetrachloride (CCl4) exposure or chronic HCV infection develops with collagen tracts that encompass hepatic lobules and connect between vasculature, known as bridging fibrosis. Pulmonary fibrosis models are more...
varied in their pathology, which depends on the type and mode of administration of the injury-inducing agent. For example, bleomycin given by intratracheal instillation will yield lots of inflammation and immune cell infiltrate that results in interstitial collagen deposition whereas bleomycin administered through systemic diffusion via osmotic pump yields markedly less immune reaction and predominately subpleural collagen deposition. Notwithstanding differences in pathologic features of ECM content and distribution, the cellular composition in most models of fibrosis are, at the most basal level, similar and generally involve an injured epithelium and quiescent fibroblasts that become activated. Understanding the differences and similarities between fibrosis models in their cellular, molecular, and histopathological composition will be critical to the development of new therapies that will be useful and effective against all modes of fibrosis.

II. Collagen and the Extracellular Matrix

The ECM consists of multiple structural molecules including proteoglycans, elastic fibers, fibronectin, laminins, and collagens. The collagens are a diverse group of structural and regulatory proteins that makes up to 30% of total protein in a given individual and a large proportion of the fibrotic lesions in fibroproliferative diseases. There are 28 species of collagen, named in order of their discovery. Collagen type I is a hetero- or homotrimer, fibrillar collagen that becomes crosslinked and interacts with other ECM to form rigid fibrils. Collagens and the ECM are necessary for many normal physiologic functions including but not limited to providing structure and rigidity of a tissue, acting as a repository of latent growth factors, sensing of mechanical stimuli, serving as a scaffold for cellular infiltrate during normal wound healing, and regulating
cell-matrix interactions\textsuperscript{11,12}. Collagen is regulated both at the transcriptional level, via Mothers against dpp family members 2 and 3 (SMAD2 and SMAD3) and other transcription factors\textsuperscript{13}, and thorough regulation of its maturation, hydroxylation, propeptide cleavage, folding\textsuperscript{14,15}, intracellular trafficking, which is regulated by SEC31A and COPII vesicles\textsuperscript{16}, and by extracellular modification like crosslinking of the fibrils by LOX and LOXL2. Collagen and ECM can form intricate complexes with laminins (LAMB1 or LAMC1), fibrillin (FBN1), and integrins, which mediate cell-matrix interactions and thus mediate perception of mechanical changes in its environment\textsuperscript{17,18}.

Collagens and ECM are very important for the structure and function of healthy and disease tissues. To better interrogate and ameliorate diseases associated with aberrant collagen and ECM regulation, composition, or deposition, elucidating the function of each component and how it relates to different cell types is imperative.

III. Treatments for a Disease of Every Organ

Current fibrosis therapies, including organ transplantation and medications such as Pirfenidone, offer organ-specific treatment but are not comprehensively anti-fibrotic. Pan anti-fibrotic therapies that exploit the commonalities between the different modes of fibrosis are currently not available. As previously stated, the development of fibrosis in solid organs follows similar molecular pathways that can be targeted by therapies that exploit common molecular characteristics. Pro-fibrotic cytokines are upstream modulators of the final fibrotic effectors molecules that control many critical transcriptional and protein changes across the cell and have been considered attractive therapeutic targets. For example, Pirfenidone, a clinically available medication for some
cases of idiopathic pulmonary fibrosis, targets proliferation and activation of fibroblasts via TGF-β signaling\textsuperscript{19}. TGF-β is a potent growth factor that activates fibroblasts by signaling through cell surface receptors, creating a signaling cascade that upregulates expression, maturation, and secretion of collagen, as well as migration via increases in the stress fibers ACTA2 (Figure 1.1a,b). Another pro-fibrotic effect of TGF-β is to reduce transcription of the miR-29 locus via SMAD3\textsuperscript{20}, and thus result in decreased levels of mature miR-29, an anti-fibrotic microRNA that targets a variety of ECM components including collagen, elastin, and fibrillin.

IV. MicroRNAs in Normal and Pathophysiology

MicroRNAs (miRNAs) are small, 21-24 nucleotide RNA species that can functionally repress a gene of interest through transcript degradation and/or translational inhibition\textsuperscript{21}. The basal RNA induced silencing complex (RISC), which consists of different combinations of argonaute proteins (AGO1-4), a miRNA, and the mRNA target, facilitates miRNA-mediated repression\textsuperscript{22,23}. Canonical targeting of mRNA transcripts is mediated through complementary binding of the seed sequence (nucleotides 2-8 of the miRNA) to the 3’ untranslated region of the mRNA\textsuperscript{24}. Any given miRNA can regulate multiple transcripts in a cell\textsuperscript{24}, thus allowing for post-transcriptional regulation of multiple targets within a given pathway, first noted in the context of important regulators of development \textsuperscript{25}. Many miRNAs are also evolutionarily conserved across metazoans and have been functionally validated in plants and animals\textsuperscript{26,27}. As of July 2018, there are 1984 miRNA producing loci described in the human genome that encode 2693 unique miRNAs\textsuperscript{28,29}. Being potent regulators of multiple mRNAs, miRNAs can also have a
significant role in disease pathophysiology including Mendelian disorders (e.g. Choroideraemia or DFNA50)\textsuperscript{30}, cancers as tumor suppressor- or oncogenes (e.g. miR-17-92 cluster or miR-29, respectively)\textsuperscript{31,32}, viral infection\textsuperscript{33}, wound healing, and fibrosis\textsuperscript{34,35}. The roles that these small RNAs play are diverse and important both for disease progression but also prevention or amelioration. MiR-29 has been shown to be a positive effector of cancer and fibrosis progression through its regulation of DNA methyltransferases (DNMT3A/B), RAS-related nuclear protein (RAN) and collagens. In multiple models of fibrosis, mature miR-29 levels are significantly reduced during and after injury and prevention and amelioration of fibrosis can occur if miR-29 levels are sustained or increased during or after injury\textsuperscript{35,36}.

V. Therapeutic Delivery of miR-29a Ameliorates Murine Liver Fibrosis\textsuperscript{*}

Early in my time in the lab, I worked with Matt Knabel and Kalyani Ramachandran to demonstrate that a clinically relevant miRNA delivery system can be used to sustain miR-29a levels during fibrotic injury in a murine model of hepatic fibrosis\textsuperscript{35}. Self-complementary adeno-associated virus (scAAV) was chosen as a gene delivery system because it had been shown in clinical trials to be an effective and well tolerated method to deliver a small gene to a cell population of interest\textsuperscript{37,38}. The serotype of the viral capsid in which the DNA is packaged displays a cell type specific pattern of infection that is different for each serotype\textsuperscript{39}. The scAAV construct included a GFP

\textsuperscript{*} Text excepts and figures reproduced or adapted from Knabel MK and Ramachandran K et al. (2015) with permission.
reporter gene, to track active transduction of the virus in the animals, and ~200bp of the human miR-29a locus in the 3’UTR of GFP (Figure 1.2a). Our virus had an AAVrh74 serotype* that has high tropism and specificity for the liver (Figure 1.2b).

Immunofluorescent co-staining for GFP and a hepatocyte specific cell marker in liver sections showed overlap of GFP derived from the virus and albumin, the hepatocyte specific marker. We also saw exclusion of GFP from the quiescent hepatic stellate cell (HSC) population when co-staining for GFP and either vimentin or desmin, which are imperfect markers for inactivated HSCs. We hypothesized that using the non-integrating scAAV to replace miR-29a during fibrotic injury would be sufficient to reduce or ameliorate hepatic fibrosis.

To test whether miR-29a could prevent or reverse hepatic fibrosis, we administered the scAAV at different time points relative to injury (Figure 1.3a,b). In the prevention model, mice were inoculated with $>10^{11}$ viral genomes per mouse 1 week prior to a 4-week course of carbon tetrachloride (CCl$_4$), a potent hepatotoxin; in the reversal model, fibrosis was established in mice with 4 weeks of CCl$_4$ injury before inoculation with the scAAV, upon which the course of injury was continued for an additional 8 weeks. The miR-29a containing scAAV was able to prevent injury-induced loss of mature miR-29a and significantly reduced and prevented pathologic deposition of collagen and fibrosis (Figure 1.3a). Pathologist blinded to the study scored Masson trichrome (MTC) stained liver sections from the scAAV.miR29a.GFP treated animals significantly lower than

---

* In Knabel MK and Ramachandran K et al. (2015) the serotype is described as scAAV8 but this was the previous, now-obsolete nomenclature and, more accurately, it is AAVrh74 serotype.
those from the control treated animals on a scale of 0 – 4, where 0 is no fibrosis and 4 is significant fibrosis. Similar effect was observed in the reversal model, with scAAV.miR-29a.GFP-treated animals showed less fibrosis than controls and even reversed existing fibrosis when compared to the 4-week baseline animals. This data indicated that preservation of miR-29a to non-injured levels during fibrotic injury using the scAAV was able to prevent and reverse CCl₄-induced murine hepatic fibrosis. Somewhat surprisingly, this outcome was achieved with miR-29a containing virus transducing predominately hepatocytes, the non-collagen producing fraction of the liver, suggesting that miR-29a targets, other than collagen itself, were mediating this effect. The subsequent chapters of my thesis aim to identify cell-type specific miR-29a targets that are driving the therapeutic benefit of the scAAV.miR29a.GFP, and to explore its effects on epithelial to mesenchymal transition (EMT) in different modes of fibrosis.
Figure 1.1 – Influence of TGFβ on fibroblast activation

(A) Quiescent fibroblast that expresses miR-29a and COL1A1 at baseline
(B) Activated fibroblast, stimulated with TGF-β. Increased signaling from TGF-β receptors and phosphorylated states of SMAD proteins, increased COL1A1 and pro-fibrotic genes, increased maturation, secretion, and crosslinking of collagen, increased ACTA2 stress fibers, and loss of mature miR-29a.
Figure 1.2 – Schematic of control and miR-29a containing scAAV and tropism of AAV in murine liver sections

(A) Schematic of the scAAV. Control and miR-29a-containing virus with GFP reporter within the inverted terminal repeat sequences (B) Immunofluorescent staining of GFP, Albumin, Desmin, or Vimentin in liver tissue transduced with scAAV.GFP.
Figure 1.3 – Prevention and reversal models of therapeutic intervention

(A) Schematic of the scAAV administration regimen. Control and miR-29a-containing virus with GFP reporter within the inverted terminal repeat sequences (B) Immunofluorescent staining of GFP, Albumin, Desmin, or Vimentin in liver tissue transduced with scAAV.GFP.
CHAPTER 2 – Therapeutic Assessment of Hepatocyte-targeted miR-29b Oligonucleotide in a Murine Model of Liver Fibrosis
I. ABSTRACT

MicroRNAs play key roles in regulation of diverse groups of pathways in normal and disease pathology, including fibrosis. The miR-29 family has significant influence over the regeneration of murine hepatic tissue after damage caused by administration of fibrosis-inducing agents such as bleomycin and carbon tetrachloride (CCl4). Previously our group demonstrated that scAAV-mediated preservation of miR-29a levels ameliorates murine hepatic fibrosis (Knabel and Ramachandran et al., 2015). In order to overcome the limitations of scAAV tropism to the liver, we set out to test the efficacy of a small RNA oligonucleotide that could be biochemically modified to target other organ or tissues of interest. In this study, we tested the therapeutic potential of subcutaneous delivery of a modified miR-29b oligo, developed by miRagen Therapeutics, in a mouse model of hepatic fibrosis. Unfortunately, this treatment regime was unable to normalize miR-29b levels, and had only marginal therapeutic benefit, did not affect the RNA or protein level of other known fibrotic markers, and caused only minor qualitative differences in the collagen fibers morphology in the miR-29b treated mice. These data indicated that, although we were unable to deliver therapeutic doses of miR-29b oligo, this treatment still had measurable effects on the collagen fiber quality.
II. INTRODUCTION

Effective therapeutic strategies require two characteristics to be successful: (1) be able to target the cell population necessary for therapeutic benefit and (2) have the ability to ameliorate a pathologic process. Multiple publications have shown that miR-29 can do the latter in that it can ameliorate various models of fibrosis such as renal, pulmonary, and hepatic fibrosis, as well as some forms of cancer by targeting pro-fibrotic and anti-proliferative processes both in vitro and in vivo\(^5,32,35,36,40-42\). The next step necessary to create a therapy is to develop a strategy for directing miR-29 to the cell type of interest in each of the disease states. Understanding what effects miR-29 has in a given cell type and which cells have the most robust therapeutic benefits when targeted is the key to developing an effective anti-fibrotic treatment.

In the CCl\(_4\)-induced hepatic fibrosis model, injury occurs in the hepatocytes, which make up \(~80\%\) of the liver by volume\(^43\), in consequence of a toxic CCl\(_4\) breakdown product generated by the cytochrome P450 family member CYP2E1\(^44,45\). Given that CYP2E1 is not expressed in hepatic stellate cells (HSC), the collagen producing and effector cells of fibrosis, these cells are not predicted to be directly affected by CCl\(_4\). Thus, it seems likely that paracrine factors from the injured hepatocytes and/or immune infiltrate responding to hepatocellular disturbance can activate the normally quiescent HSC population\(^46,47\). Previous studies aimed at increasing miR-29a or miR-29b in a given tissue during fibrotic injury have used delivery methods specific to the tissue of interest. For example, nearly all tissues were affected when using a constitutive over expressing transgenic mouse to look at renal fibrogenesis; the whole lung was affected when using a Sleeping beauty-mediated transposon therapy in a
pulmonary fibrosis study; and most specific in expression pattern, certain cells of the liver were affected when using an scAAV\textsuperscript{35,48}. The scAAV study from Knabel and Ramachandran et al. (2015) showed a significant number of hepatocytes that expressed virally transduced GFP protein, while quiescent hepatic stellate cells did not. Additionally, in all of the studies described, miR-29a and miR-29b were protective against fibrotic injury both molecularly and histologically.

MiRNAs negatively regulate stability or translation of multiple RNA targets in any given cell. Unlike certain drugs that need to be metabolically activated prior to beneficial action, the machinery required for miRNA regulation is present in all cells, so addition of a given miRNA should elicit regulation of RNA targets that are expressed and accessible to the RNA-induced silencing complex (RISC). The miR-29 family consists of 3 species, transcribed from 2 loci in the human and mouse genome (miR-29a/b/c from miR-29a/b1 and miR-29c/b2 loci)\textsuperscript{49}. MiR-29a/b/c all share a common ‘seed’ sequence (nucleotides 2-8), which is most significant factor for the miRNA target specificity. Interactions at the 3’ end of the miRNA also play a role but the rules that govern this interaction more complex than just complementation\textsuperscript{50,51}. Some studies have shown miRNAs of the same family to have highly similar lists of targets with divergent targets specific to 3’ end sequence differences of the miRNAs\textsuperscript{52}. Despite the fact that they diverge at the sequence level at the 3’ end by 5 nucleotides, both miR-29a and miR-29b, display therapeutic benefit when administered during fibrotic injury.

Use of therapeutic small RNAs is currently being tested in many pre-clinical trials and in at least one human trial\textsuperscript{53,54}. Currently under development is a use of a modified miR-29b (MiRagen Therapeutics) to treat skin fibrosis (Clinical trial ID: NCT02603224),
which, given ease of accessibility, circumvents many complex delivery issues. In order to test whether this oligonucleotide could be targeted to harder-to-reach cell populations, we worked with MiRagen Therapeutics to assess whether an oligonucleotide of miR-29b could be modified to selectively target liver populations in a CCl₄-induced liver fibrosis model.
III. RESULTS

*MiR-29a and miR-29b are equivalent regulators of collagens and other pro-fibrotic factors*

All three miR-29 family members are known regulators of pro-fibrotic genes including collagens, elastin, and fibrillin. MiR-29a and miR-29b vary in sequence by 5 nucleotides, all of which are located outside of the ‘seed’ sequence. In view of the results obtained using ssAAV-miR-29a, we wanted to determine if there were differences between miR-29a and miR-29b targets within our cells of interest, human hepatocytes and hepatic stellate cell (HSC) lines (HuH7 and LX2, respectively). The cells were grown in culture and transfected with exogenous miR-29a or miR-29b in the presence or absence of the profibrotic cytokine transforming growth factor beta (TGF-ß). In both HuH7 and LX2 cells, known miR-29 targets were similarly regulated by exogenous miR-29a or miR-29b (Figure 2.1). When stimulated with TGF-ß, miR-29a and miR-29b were significantly reduced and expression of profibrotic genes such as *COL1A2* and *FBN1* was significantly increased. In the presence of both TGF-ß and exogenous miR29a or miR-29b, expression of *COL1A2* and *FBN1* was normalized or significantly reduced compared to TGF-ß treatment alone. Collagen, type I is a heterotrimeric collagen fiber that consists of 2 *COL1A1* molecules and one *COL1A2* molecule. *COL1A2* mRNA was used as the metric for collagen expression because baseline measurements of *COL1A1* transcript in HuH7 cells were below the threshold of detection. Both members of the miR-29 family were able to regulate ECM components that TGFß had significantly increased in both cell types.
Hepatocyte-directed miR-29b oligonucleotide does not alleviate CCl4-induced liver fibrosis

As miR-29a has been shown to be therapeutically beneficial when given prior to or during fibrotic injury and miR-29a and miR-29b have been shown to both down-regulate expression of pro-fibrotic ECM components in both hepatocyte and HSC populations, a cohort of WT C57BL/6 mice were used to test the efficiency of a modified miR-29b oligonucleotides, designed to target hepatocytes, in alleviating the pathology of fibrosis. The concentrations and exact modifications of the tested miR-29b oligos, M12723, M12768, and M12770, are proprietary information of MiRagen Therapeutics. M12723, M12768, and M12770 were tested in a fibrosis reversal model whereby mild fibrosis is established for 4 weeks by intraperitoneal injections of CCl4, prior to therapeutic intervention and then continued for 8 weeks concurrently with miRNA treatment. Intraperitoneal injections of CCl4 and subcutaneous injections of miR-29b oligos were given twice and trice a week, respectively (Figure 2.2a). The animals were sacrificed 3 days after last CCl4 and miR-29b injections. Levels of miR-29a and miR-29b, as measured by qPCR, were significantly reduced in all injured livers, including those from miR-29b oligo treated animals (Figure 2.2b). This result was not entirely unexpected as work conducted by MiRagen Therapeutics had shown that the modified oligos have a short half-life in vivo and it was conceivable that no detectable increase in miR-29b in the treated animals would be observable 3 days after the last injection prior to sacrifice. Alternatively, miR-29b oligo concentration or modification was not sufficient to raise miR-29b levels or achieve efficient uptake by hepatocytes. Future study designs could take this failure into account in order to gain a more accurate measure of peak
modified oligo levels. Blood serum analysis of liver enzymes showed that animals were equally injured by CCl₄, independently of miR-29b oligo injections (Figure 2.2c). This implied that administration of miR-29 oligos does not rescue the CCl₄ from injuring the liver.

*Qualitative difference in collagen composition in miR-29b oligo treated mice*

To assess the histological differences between conditions, liver sections were stained with Masson Trichrome stain to visualize collagen fibers blue and aid in tracking the differences in collagen accumulation, distribution, and intensity (Figure 2.3). No overt differences of collagen deposition across multiple sections were observed, and quantification of the total collagen content and quality of collagen fibers with Aperio ImageScope software revealed that, as expected, total deposited collagen in liver sections was significantly higher in CCl₄-treated animals compared to corn oil-treated controls, while no significant differences between oligo- and saline-injected mice were observed (Figure 2.4 and Figure 2.5). Interestingly, the same analysis revealed differences in the quality of collagen fibers deposited, as categorized in ‘weak’, ‘positive’, and ‘strong’ staining intensities categories. The miR-29b oligo treated animals had significantly less weak collagen staining, a trend toward less positive staining, and equivalent levels of strong staining, suggesting that miR-29b given to the hepatocytes may have had an effect on the less structured, diffuse collagen but not the more intensely stained collagen fibers. Molecular markers associated with fibrosis, measured at the RNA and protein levels, showed no altered regulation by miR-29b oligo treatment (Figure 2.6), including
predicted and validated targets of miR-2955 *Col1a1, Col3a1*, and *Serpinh1*, which are associated collagen protein synthesis and maturation during injury.
RNA levels for pro-fibrotic genes COL1A2 and FBN1 were measured by qPCR in HuH7 and LX2 cells stimulated with 10ng/ml TGFβ ± control, miR29a, or miR-29b oligonucleotide. Each cell line in each primer set was normalized saline-treated, control transfected cells.
Figure 2.2 – Design and assessment of miR-29b oligo in liver fibrosis model

(A) Experimental design for reversal study. (B) qPCR for mature miR-29a or miR-29b from murine livers. RNA was normalized using 18S rRNA. (C) Liver enzyme levels measured from serum taken at time of sacrifice.
Figure 2.3 – Masson Trichrome stain of livers after 12 weeks of CCl₄

Masson Trichrome stain of murine liver sections ± 12 weeks of CCl₄ ± 8 weeks of miR-29b oligo
An example of the Aperio ImageScope software discerning different intensity of staining in the collagen fibers of the Masson trichrome stained livers. Left three panels are Masson trichrome stained sections after 12 weeks of CCl₄ injury at 10x magnification; Middle three panels are same magnification, same fields of the same sections above using the positive pixel count (PPC) markup to show stained fibers in shades of red; Right three panels are 20x magnification inset images of middle panels.
Figure 2.5 – In silico quantification of collagen content from liver sections

Total collagen quantification (composed of ‘weak’, ‘positive’, and ‘strong’ staining intensities) and the individual intensities across whole slide sections of the miR-29b oligo cohort.
Figure 2.6 – Molecular analysis of fibrosis-associated genes in CCL4-treated livers

Molecular analysis of WT B6 mice given CCl₄ for 12 weeks. Therapeutic intervention began at week 4 with either saline or 3 different versions of modified miR-29b oligo. (A) qPCR analysis quantifying Col1A1 and Col3A1 mRNA levels. 18S rRNA used as normalizing RNA species. (B) Western blot for Serpinh1 protein, also known as Hsp47, and normalizer Actin (Actb). (C) Quantification of Serpinh1 mRNA by qPCR and protein by Western blot, pictured above. All p-values were calculated with Two-Tailed Student T-tests.
IV. DISCUSSION

*Therapeutic restoration of miR-29 may not be sufficient to ameliorate hepatic fibrosis*

Targeting of modified miR-29b to the epithelial fraction (hepatocytes) of the liver during fibrotic injury was not sufficient restore miR-29b to pre-injury levels and therefore not able to ameliorate excess collagen deposition and fibrosis. Although data suggests that miR-29b was trafficked to the hepatocytes, as evidenced by quantitative differences in collagen composition of the livers treated with miR-29b, compared to untreated, miR-29b levels not being sustained throughout injury and recovery periods may have resulted in inefficient regeneration of tissue. We cannot definitively say that miR-29b was not restored to the hepatocyte fraction during injury or after injury without confirming spatially where the modified oligo was located. Although hepatocytes compose 80% of the liver by volume, they contribute ~50% by cell number. It’s possible that the loss of miR-29a/b across the entire liver, as evidenced by qPCR, may be explained by loss of miR-29 from the non-hepatocyte fraction. In order to know if modified miR-29b levels is actually increased in the hepatocytes, we would need a method of visualizing the modified oligo through some quantitative *in situ* assay coupled with a lineage tracing method to specifically identify hepatocytes from the other cellular components of the liver. Regardless of specific mechanism, our data demonstrate that the modified miR-29b did not recapitulate the therapeutic benefit that we previously observed using with the scAAV-miR29a. Three possibilities may explain this discordance: (1) the modified miR-29b oligos were not able to restore miR-29b to therapeutic levels in the hepatocytes, (2) miR-29b is not sufficient for clinically relevant therapeutic benefit where miR-29a is.
Qualitative difference in collagen fibers in miR-29b treated animals

Although administration of the miR-29b oligos were not able to fully rescue CCl₄-treated mice from overall increased collagen deposition and liver fibrosis, we did detect a significant difference between the saline and miR-29b oligo-treated groups. The distribution of collagen fiber intensities of Masson trichrome stained sections was different and showed significantly less diffuse, lightly stained collagen fibers in the miR-29b-treated animals. This may suggest that even suboptimal levels of miR-29b may be able to partially regulate collagen deposition, turnover, maturation, or crosslinking, either directly or by modulation of paracrine signals upstream of collagen deposition. MiR-29-mediated regulation of paracrine signaling via the cytokine TGF-β2 and one of its receptors (TβRII) is likely, given that members of this signaling pathway are validated miR-29 targets with canonical seed sequence binding sites in their 3’UTRs. Another possibility is that modified miR-29b inhibited epithelial to mesenchymal transition (EMT) of targeted hepatocytes to bile-duct progenitors and possibly other cells types, and thereby reduced the portion of the collagen content that would be deposited from hepatocyte-derived fibroblasts.

Which cell type with increased miR-29 levels is sufficient to produce antifibrotic effects?

As has been proposed earlier, a possible reason for the outcome of these experiments is that delivery of miR-29 to the hepatocyte fraction is not sufficient for fibrosis prevention or reversal. Our previous studies showed strong overlap between scAAV-mediated GFP expression and the hepatocyte specific marker albumin but a lack of overlap with the hepatic stellate cell (HSC) markers, vimentin or desmin. This had led to the conclusion
that hepatocyte transduction of miR-29a was able to ameliorate murine hepatic fibrosis. However, it is possible that scAAV-infected cells other than hepatocytes or quiescent HSCs that may play a significant role in fibrosis development. Some examples may include Acta2+ myofibroblasts, or CD45+ Kupffer cells, or a range of other non-resident cells that infiltrate liver tissue post injury. Our immunofluorescence images from the previous study did not show a large population of unknown GFP+ cells in the liver at baseline or with injury suggesting that the key cellular components may be small in number but potent in its regulation of the phenotype. Collectively our data suggest that identifying miR-29 targets and the cell-type in which its action is most important to prevention of fibrosis is important to the design of viable therapeutics for treatment of fibrotic diseases.
V. METHODS

*Small RNA transfection and TGFβ stimulation of HuH7 and LX2 cells*

HuH7 or LX2 cells were grown in a 37°C incubator in 5% CO₂ in full serum media [FS Media; DMEM (Corning Cellgro; Cat# 10-013-CV) + 10% FBS (ThermoFisher Scientific; Cat# 26140079) + 1x Antimycotic-Antibiotic (ThermoFisher Scientific; Cat# 15240062) + GlutaMAX (ThermoFisher Scientific; Cat# 35050061)]. At ~70% confluency, media was replaced with low serum media [LS Media; DMEM + 0.5% FBS + AntiAnti + GlutaMAX] for 24hr prior to transfection. Cells were transfected with hsa-miR-29a, hsa-miR-29b, or cel-miR-67 control oligo (Dharmacon; Cat# C-300504-07-0005, C-300521-05-0005, and CN-001000-01-05) using Lipofectamine RNAiMAX (ThermoFisher Scientific; Cat# 13778075) as per manufacturer’s protocol. After 3hr with transfection reagent in media, the media was replaced with LS media ± 10ng/ml of recombinant human TGFβ1 (BioLegend; Cat# 580702) and incubated for 24hr.

*CCL4-induced liver fibrosis in mice*

Wild-type C57BL/6J mice, aged 8 weeks, received biweekly intraperitoneal injections of 1ml/kg of carbon tetrachloride, diluted 1:7 in corn oil (Sigma Aldrich; Cat# 289116 and C8267). Injections of the miR-29b oligonucleotide or saline control were on the right, hind leg, 3 times per week. After 12 weeks of injury, mice were euthanized and serum and organs were harvested for analysis. All housing and procedures were conducted according the protocol approved by the Animal Care and Use Committee of Johns Hopkins University, School of Medicine.
Histology preparation, staining, and analysis

Tissue was fixed with 10% Formalin for 24hr and then placed in 70% ethanol. The Johns Hopkins Reference Histology Lab embedded tissues in paraffin, cut sections for slides, and performed the Masson Trichrome staining. Picrosirius Red staining was done on 10µm sections affixed to Fisherbrand™ Superfrost™ Plus Microscope Slides (Fisher Scientific; Cat# 12-550-15). Protocol adapted from Histological and histochemical methods: theory and practice59. In brief, sections were deparaffinized, rehydrated, incubated with a picrosirius red solution (1% Direct Red 80 in saturated picric acid solution) for 1hr, washed with acidified water (0.5% glacial acetic acid diluted in water) twice for 15min, and then dehydrate in 3 changes of 100% ethanol before mounting with coverslip. Stained slides were sent to Johns Hopkins Oncology Tissue Services for whole slide imaging. Quantitative and qualitative analysis of trichrome stain was conducted using the Aperio ImageScope software.

RNA isolation and RT-qPCR analysis

Total RNA was isolated from cultured cells or tissue homogenate using Trizol (Invitrogen; Cat# 15596026) as per manufacturer’s protocol. RNA was treated with DNaseI (New England Biolabs; Cat# M0303S) before cDNA synthesis using the High Capacity cDNA Reverse Transcription Kit (Invitrogen, Cat# 4368814). Quantitative PCR using SYBR Green PCR Master Mix (ThermoFisher Scientific; Cat# 4309155) or Taqman Gene Expression Master Mix (ThermoFisher Scientific; Cat# 4369016) were used to amplify and quantify cDNA of interest. Housekeeping genes ACTB or 18S rRNA were used for relative quantification. GraphPad Prism was used for plot generation and
statistical analysis where p-values were generated with two-tailed Student T-tests when comparing two groups and a two-way ANOVA and Krustal multiple hypothesis testing when comparing more than two groups.

**Western Blot for protein analysis**

Approximately 20mg of tissue was lysed in M-PER™ Mammalian Protein Extraction Reagent (ThermoFisher Scientific; Cat# 78501) using a FastPrep24 tissue homogenizer (MP Biomedicals; Cat# 116004500) and quantified using a BCA protein assay (ThermoFisher Scientific; Cat# 23225). Using ~20μg of protein per lane, samples were loaded into 10% Criterion™ XT Bis-Tris acrylamide gels (Bio-Rad; Cat# 3450112) and run at constant voltage for ~1.5hr. The protein was then transferred to PVDF membrane using Bio-Rad Trans Blot Turbo transfer system, as per manufacturer’s protocol. The membrane was then blocked for 1hr in Odyssey® Blocking Buffer (LI-COR; Cat# 927-50000), incubated with 1° antibody overnight at 4°C, washed 3 times with PBS + 0.1% Tween20 (PBST), incubated with 2° antibody (IRDye® 800CW or 680LT from LI-COR) for 1hr at room temperature, washed 3 times with PBST, and rinsed with PBS twice before imaging on an Odyssey imaging system. Primary antibodies include: Actb [1:5000] (Sigma; Cat# ), Acta2 [1:1000] (Abcam; Cat# ), and Serpinh1 [1:1000] (ThermoFisher Scientific; Cat# ). Proteins of interest were quantified using ImageStudio Lite software from LI-COR and normalized to Actb levels.
Serum Analysis

After sacrifice, blood was retrieved from the right atrium using a 21 gauge needle.

Approximately ~300-400 µl of whole blood was allowed to clot at room temperature for 15 minutes before serum was isolated by spinning the blood for 5 minutes at 5,000xg in Microtainer serum separator spin tubes (BD; Cat# BD365967). Serum was flash frozen in liquid nitrogen and stored at -80°C until serum analysis was conducted by the Phenotyping Core in the Department of Molecular and Comparative Pathobiology at Johns Hopkins School of Medicine.
CHAPTER 3 - Epigenetic Regulation of Col1a1 Locus Silences

Cis-Transgene Insertions
I. ABSTRACT

In this study, we derived transgenic mice from embryonic stem cells that utilize a recombinase mediated cassette exchange to insert a temporally and spatially controlled, tetracycline-regulated transgene expressing miR-29a and a reporter GFP, downstream of the Col1A1 locus in, what has been previously described as, a ‘transcriptionally open’ region. Upon inducing liver fibrosis in these mice by carbon tetrachloride (CCl4) injections, we noted dysregulation of the transgene expression in alpha smooth muscle actin (Acta2) positive, collagen producing cells of the liver. This result was recapitulated in vitro with mouse embryonic fibroblasts stimulated with the pro-fibrotic cytokine TGF-β where there was an increase in Acta2⁺; GFP⁻ cells. Due to the reduced number of GFP⁺ cells in culture with TGF-β, we hypothesized a change in the epigenetic landscape at the locus downstream of Col1a1 may be silencing transgene expression. A meta-analysis of a ChIP-Seq dataset mapping H3K27ac-bound chromatin in the presence or absence of TGFβ in the murine mammary gland cell line NMuMg showed significant changes in the locus proximal to transgene insertion site with TGF-β stimulation. Other mice created using an insertion site within 1kb of our transgene locus, also showed aberrant expression from their transgene in high collagen expressing cell types. In summary, we created a transgenic mouse that can temporally and spatially express miR-29a and GFP; additionally, expression of this transgene can be ablated in the presence of pro-fibrotic stimulus both in vivo and in vitro, with evidence for epigenetic changes being the cause. Beyond our mouse, many other mouse models with the same or similar insertion loci may be affected by this newly described regulation.
II. INTRODUCTION

Biomedical research relies on model organisms to allow the study of molecular and physiologic processes in a living organism. Direct patient cells or tissues are difficult to access and are limited in quantity making it difficult to rigorously test multiple conditions. Model organisms, including yeast, nematodes, mice, and rats, have been used successfully to investigate the normal or disease processes that affect human biology. They allow the experimenter to attain greater sample sizes through shorter life cycles, the ability to obtain vital tissues or samples, and opportunity to use genetic modifications to test functions of targets of interest or mimic rare disease mutations.

Some of the greatest contributions from model organisms have come from the ability to modify their genome to interrogate the function of individual genes as well as reproduce specific disease associated mutations. In mice, a random insertion of DNA of interest, commonly known as the transgenic mouse model, has been used since the 1990s with the first somatic transgenic mouse being created in 1974\textsuperscript{60}. The technique of homologous recombination was developed in 1987, making targeted insertion possible. With these ever-improving technologies many transgenic, knock-out, and knock-in mouse models have been created and utilized for biomedical research. According to the International Mouse Strain Resource (IMSR)\textsuperscript{61} as of June 2018, there are 32,620 mouse strains created and, if modified sperm and ova are included, that number jumps to 260,103 strains available from around the world. These mouse models encompass an incredible number of genes and with many different mutations that are used for virtually all fields of study.
MicroRNAs (miRNAs) are an important, evolutionarily conserved family of regulatory molecules that influence the expression pattern of multiple genes with a single RNA species. Many miRNAs have been implicated in the regulation of disease phenotypes including cancer and fibrosis through their ability to modulate expression of critical effectors of pathogenesis. For example, increased expression of the Mirc1/MiR17-92 cluster in cystic fibrosis patients has been negatively correlated with lung function\textsuperscript{62}. The miR-29 family has been shown to be down-regulated during fibrotic injury to the heart, liver, lung, and kidneys\textsuperscript{5,20,35,63}, and its forced expression in tissues under fibrotic stress has been shown to be therapeutically beneficial\textsuperscript{35,36}. In order to generate a genetic tool that would allow controlled expression of this locus and aid our investigation of miR-29a as a therapeutic for fibrosis, we used mouse embryonic stem (ES) cells from the NCI repository to create a transgenic mouse carrying an inducible miR-29a transgenic allele.

The transgenic mouse created in our study was derived from an ES cell library generated by Dr. Scott Lowe’s group\textsuperscript{64}, in which different miRNAs in a miR-30 backbone are targeted to a transgene insertion site originally described and modified to allow FRT-based recombination by Dr. Rudolf Jaenish’s group\textsuperscript{65}. This library collection of ES cells contains the miRNA-inducing transgenes under temporal and spatial control through the use of tetracycline regulating element (TRE) as a promoter, and a reverse tetracycline-controlled transactivator (rtTA), which is transcribed from a different locus. The spatial control comes from the promoter elements regulating expression of the rtTA, while the temporal control comes from the need for doxycycline to transcriptionally
activate the TRE in the transgene locus. This system, commonly referred as TET/ON\textsuperscript{66}, allows tight control of the expression of transgenes of interest.

In creating this mouse model and testing its therapeutic potential in a hepatic fibrosis model, we uncovered a novel epigenetic phenomenon that regulates the transcriptional activity of this commonly used transgene insertion site. During the course of our study, we determined that our transgene was not expressed in a subset of cells that co-localize with markers of the collagen-producing myofibroblast population. This epigenetic phenomenon likely affects all transgenes inserted into this locus, which is located proximal to the 3’ untranslated region (UTR) of \textit{Col1A1} and has been used to generate 18 published mouse strains\textsuperscript{67} and an unknown number of unpublished mouse strains.
III. RESULTS

*TRE_eGFP-miR29a Construct*

To create a system with the ability to temporally and spatially control miR-29a expression, we derived mice from ES cells that contained a GFP reporter with a mature miR-29a construct in the 3’UTR, all under the control of multiple tetracycline regulated elements (TRE), and inserted downstream of the Col1A1 locus (Col1A1^{tm1(TRE_EGFP-miR29a)})\(^{64,65}\) (Figure 3.1a). The ES cells also contained M2-rtTA at the Rosa26 locus. The M2-rtTA was described in the literature as widely expressed in multiple tissues with strong expression in the small intestine and liver\(^{68}\). Mice heterozygous for the Col1A1^{tm1(TRE_EGFP-miR29a)} (Tg29) and heterozygous or homozygous for the M2-rtTA were given doxycycline containing chow for 10 days and expression of the GFP reporter was assessed via immunofluorescence (Figure 3.1b). The expression pattern of GFP in the homozygous M2-rtTA mice matched what was previously described in the literature\(^{65}\), with strong GFP signal in both the small intestine and liver and weak fluorescence in pericardial and interstitial fibroblasts in the heart. A few organs, including the lung, showed little to no GFP expression (Figure 3.1b). In an effort to get more ubiquitous expression of the transgene, we crossed the Tg29 mouse to a B6N.FVB(Cg)-Tg(CAG-rtTA3)4288Slowe/J (rtTA3) mouse, which expresses an enhanced rtTA (rtTA3), randomly inserted as a transgene and under the control of a strong synthetic (CAG) promoter. This line, which had been shown to display stronger and wider expression pattern than the M2-rtTA\(^{64}\), induced expression of the Tg29 transgene in a greater range of tissues (Figure 3.2a,b). Expression in the liver was strong but was not complete after 10 days of doxycycline induction. To test if this could be due to insufficient exposure to
doxycycline, we administered doxycycline-containing chow for 50 day. Under these conditions, we saw near complete induction of the transgene across the liver (Figure 3.2c). To confirm miR-29a was also increased, we used qPCR to access mature miR-29a transcript levels across various tissues (Figure 3.3a). In the presence of doxycycline, both \textit{M2-rtTA} and \textit{rtTA3} significantly increased miR-29a in the liver. Other organs showed increases in mature miR-29a that correlated with the amount of GFP seen by microscopy (data not shown).

\textit{Col1A1^{tm1(TRE_EGFP-miR29a)}} expression during \textit{CCL4}-induced fibrosis model

With robust GFP- and miR-29a expression in the liver after doxycycline induction using the \textit{rtTA3}, we used a hepatic fibrosis model to test the efficacy the Tg29 mouse in restoring miR-29a expression during fibrotic injury. Carbon tetrachloride (\textit{CCL4}) is a potent hepatotoxin that, administered over a course of 4-, 6-, or 8-weeks can induce varying amounts of liver damage, collagen deposition, and fibrosis. Mice that were heterozygous for the \textit{Tg29} transgene (\textit{Tg29}^{+/−}) and either negative for rtTA (\textit{rtTA}^{−/−}) or heterozygous for \textit{rtTA3} (\textit{rtTA3}^{+/−}) were injured with a 6-week course of \textit{CCL4} diluted in corn oil to 1 ml/kg, given twice a week through intraperitoneal injections. Doxycycline chow was given to both \textit{rtTA3}^{+/−} and \textit{rtTA3}^{+/+} mice, beginning 1 week prior to the first \textit{CCL4} injection. Mice were sacrificed 1 day after the last \textit{CCL4} injection and assessed for histological and molecular markers of fibrosis. Quantitative PCR showed significant increase in miR-29a levels in the \textit{rtTA3}^{+/−} mice compared to their \textit{rtTA3}^{−/−} controls even in the presence of injury (Figure 3.4a). Surprisingly, given that miR-29a expression is suppressed in the context of fibrosis, we did not observe a significant decrease of miR-
29a levels in the liver of CCl4-treated animals, even in the absence of doxycycline, possibly due to low level leakiness of the transgene locus. Masson trichrome (MTC) stain, which stains collagen fibrils blue, showed significant fibrosis with CCl4-treated animals compared to corn oil-treated animals, and a qualitative decrease with rtTA3+/− (Figure 3.4b). MTC visualizes collagen content, the effector molecule of fibrosis, but the pathology relies on more than just collagen deposition. The regulation of collagen consists of chaperone proteins for proper folding, crosslinking enzymes to increase rigidity of ECM, and matrix metalloproteases to regulate ECM turnover, including Sparc and Mmp2 genes. The injured mice exhibited significant increase all of these fibrotic markers and only partially corrected mRNA levels with the addition of miR-29a. Even though miR-29a levels were sustained during fibrotic injury, histological and molecular markers of fibrosis persisted. This is contrary to findings in the literature and our own studies showing that miR-29a expression was sufficient to ameliorate fibrosis.

Continuing our characterization of the Tg29 mouse during fibrotic injury, we examined the distribution of Acta2, a marker for activated fibroblasts or myofibroblasts that have high collagen expression, and GFP proteins by immunofluorescence (IF) (Figure 3.6a). With CCl4 injections, mice significantly increased the amount of Acta2+ cells in the liver, which correlated spatially with the collagen deposition and fibrotic tracts. Interestingly, there was no significant overlap of Acta2 and GFP immunofluorescence. This observation was consistent across all other organs surveyed including kidney and lungs (Figure 3.6b). Additionally, when we crossed Tg29 transgenic mice to an Acta2-rtTA, in place of rtTA3, we observed little to no GFP fluorescence (Figure 3.6c). The fact that miR-29a expression is not being sustained in the collagen producing myofibroblasts may
explain the otherwise discordant observations of increased miR-29a levels by qPCR without significant protection from fibrosis. Possible explanations for transgene not being expressed in these cells are regulation of the \textit{rtTA} locus in these cells, regulation of the \textit{Tg29} locus in these cells, or doxycycline availability in these cells. The \textit{rtTA3} locus is randomly inserted in the mouse genome so there could be some cell type specific regulation of the \textit{rtTA} transgene locus in the Acta2\textsuperscript{+} cells; we believe this is not likely given that we have also used the M2-\textit{rtTA} knocked-in to the Rosa26 locus in both a 4- and 6-week CCl\textsubscript{4} injury experiments and saw the same pattern of exclusion of GFP from the Acta2\textsuperscript{+} cells (data not shown). The 2 different rtTAs, under different promoters from different insertion sites in the genome, would need to be regulated by the same event, which seems highly unlikely.

\textit{TGFβ stimulation in MEFs Col1A1\textsuperscript{1tm1(TRE_EGFP-miR29a)} +/-; M2-rtTA +/- mice}

To determine why the Acta2\textsuperscript{+} cells do not express the \textit{Tg29} transgene, we utilized an \textit{in vitro} system to examine the effects of TGFβ and doxycycline exposure in a homogenous cell population like mouse embryonic fibroblasts (MEFs). \textit{Tg29\textsuperscript{+/-}; M2-rtTA \textsuperscript{+/-}} MEFs were isolated and grown in 37°C + 5% CO\textsubscript{2} before being passaged onto glass coverslips. Cells on the glass coverslips were grown in the presence or absence of doxycycline (1mg/ml) and/or TGF-β (10ng/ml) to mimic \textit{in vivo} conditions of transgene activation, fibrotic injury, or both. Immunofluorescence for Acta2 and GFP was visualized in the cells 24hr after TGF-β stimulation and 48hr after doxycycline addition (Figure 3.7-8). MEFs showed Acta2 expression at baseline that was increased with TGF-β stimulation and reduced with transgene expression. When co-stimulated with TGF-β and doxycycline,
there was an increase in Acta2 signal that showed no overlap with GFP\(^+\) cells. In other words, the cells expressing highest levels of Acta2 were GFP\(^-\).

**Metanalysis of TGFβ-stimulated NMuMg cells**

We hypothesized that mechanism by which the Tg29 transgene expression is repressed in Acta2\(^+\) cells could be epigenetic in nature. In order to interrogate the influence of TGF-β on the epigenome, we conducted a meta-analysis, looking at transcriptomic and epigenetic changes in Namru murine mammary glands (NMuMg) cells treated with TGF-β for 0, 4, or 24hr\(^69\), and monitor whether the epigenetic status of the transgene insertion site were altered by TGF-β stimulation. In Navendar et al. (2017), investigators stimulated NMuMg cells with 10ng/ml TGFβ1 for 0, 4, or 24hr and then isolated RNA for RNA-Seq and, in parallel, conducted chromatin immune-precipitation and sequencing (ChIP-Seq) using an antibody for histone 3 lysine 27 acetylation (H3K27ac).

Transcriptomic analysis showed increased expression of markers of epithelial to mesenchymal transition (EMT) as well as increased expression of fibrotic markers like Acta2, Colla1, and Pdgfrb. The Colla1 transcript was significantly increased in both the 4 and 24hr timepoints and there were differential peaks of H3K27ac-bound DNA in both 4 and 24hr with TGF-β compared to 0hr (Figure 3.9). Mapping the transgene location, relative to the ChIP-Seq peaks, we noticed that the Colla1 locus is being epigenetically modified proximal to the transgene insertion site when given TGF-β. This confirms our hypothesis that there are epigenetic landscape changes proximal to our transgene under conditions that lend to increased Colla1 expression.
Using the ENCODE database\textsuperscript{70}, we were able to verify that the H3K27ac marks we saw in NMuMg cells were conserved in some human cell lines, although we could not further investigate the effect of TGF-β from those datasets. Others have noted that genomic loci associated with H3K27ac histone modifications produce long non-coding RNAs, known as eRNAs\textsuperscript{71,72}, which are derived from these enhancer regions. To determine if exposure to TGF-β influences the human locus in a similar manner, we tested if an eRNA homologous region proximal to the \textit{COL1A1} locus, was induced by TGF-β in human HuH7, epithelial-derived hepatocytes that do not produce collagen, and LX2 cells, which are derived from collagen-producing human hepatic stellate cells (HSCs) (Figure 3.10). After exposure to TGF-β, HuH7 cells show a modest but significant increase in the eRNA, whereas LX2 cells, show a large increase in the novel eRNA. This data suggests that the epigenetic modification of the Col1a1 locus, driven by either injury or TGFβ stimulation, appears to be conserved between mouse and human, is associated with increased collagen expression but likely interferes with the proximal transgene expression.
Figure 3.1 – Expression pattern of Col1A1\(^{tm1(TRE_{EGFP-miR29a})}\) with M2-rtTA

(A) Schematic of the Col1A1\(^{tm1(TRE_{EGFP-miR29a})}\) and Rosa26_M2-rtTA genomic constructs. (B) Immunofluorescence of GFP signal in Col1A1\(^{tm1(TRE_{EGFP-miR29a})}\) heterozygous mice with 0, 1, or 2 copies of the M2-rtTA after 10 days of doxycycline induction. Top panels are lung sections and bottom panels are liver sections of the given genotypes at 10x magnificaiton.
Figure 3.2 – Expression pattern of Col1A1<sup>tm1(TRE_EGFP-miR29a)</sup> with rtTA3

(A) Schematic of the Tg(CAG-rtTA3) genomic sequence. (B) Immunofluorescence of GFP in liver, kidney, pancreas, and lung of mice heterozygous for Col1A1<sup>tm1(TRE_EGFP-miR29a)</sup> and heterozygous for Tg(CAG-rtTA3) after 10 days of doxycycline induction. (C) Comparison of 10 days vs 50 days of doxycycline induction of liver sections.
Figure 3.3 – Mature miR-29a levels after doxycycline induction

RNA levels as quantified by qPCR using Taqman probe designed for the mature miR-29a-3p RNA normalized to 18S rRNA. Doxycycline induction occurred for 50 or 35 days in the rtTA3 and M2-rtTA mice, respectively. Mice were heterozygous for Col1A1tm1(TRE_EGFP-miR29a) and heterozygous for Tg(rtTA3) or homozygous for Rosa26_M2-rtTA.
Figure 3.4 – Histological and molecular analysis of Col1A1\textsuperscript{tm1(TRE_EGFP-miR29a)} ± 6wk CCl\textsubscript{4}

(A) Mature miR-29a levels measured by qPCR with Taqman probe for mature-miR-29a-3p in corn oil treated ± rtTA3 and CCl\textsubscript{4}-treated ± rtTA3 animals. (B) Masson trichrome staining of 4 individuals from each condition after 6 weeks of control or injury with 50 days of doxycycline chow.
Figure 3.5 – Regulation of pro-fibrotic, miR-29 targets in Col1A1<sup>tm1(TRE_EGFP-miR2a9)</sup> mice

Quantitative RT-qPCR for Col1a1, Col1a2, Mmp2, and Sparc transcripts in the 6-week CCl<sub>4</sub> fibrosis model.
Figure 3.6 – Distribution of Acta2+ cells in Col1A1^{tm1(TRE_EGFP-miR29a)} mice

Immunofluorescent staining of Acta2 and GFP in Col1A1^{tm1(TRE_EGFP-miR29a)} mice. (A) IF in mice after 6 weeks of corn oil and CCl4 treated livers of Col1A1^{tm1(TRE_EGFP-miR29a)} mice Tg29^{+/-}; rtTA3^{+/-} [Dox chow = 50 days] (B) CAG-rtTA3 mice: IF in liver, kidney, and lung sections [Dox chow = 50 days] (C) Acta2-rtTA mice: IF in liver, kidney and lung sections [Dox chow = 14 days]. All images were taken at 20x magnification.
Figure 3.7 – MEFs stimulated with TGFβ – 10X Magnification

Mouse embryonic fibroblasts imaged at 10x magnification. Left panels are without doxycycline in the media; Right panels have had 1mg/ml doxycycline added for 48hrs. Top panels are without TGFβ; Bottom panels are with 10ng/ml TGFβ1 added for 24hr. White arrows highlight Acta2⁺; GFP⁺ cells.
Mouse embryonic fibroblasts imaged at 20x magnification. Left panels are without doxycycline in the media; Right panels have had 1mg/ml doxycycline added for 48hrs. Top panels are without TGFβ; Bottom panels are with 10ng/ml TGFβ1 added for 24hr. White arrows highlight Acta2⁺; GFP⁺ cells.
Figure 3.9 – Epigenetic changes to Col1a1 locus with TGFβ stimulation

Genomic map of Col1a1 locus in the mouse genome with reads from H3K27ac ChIP-Seq analysis visualized. Samples in green are 0hr TGFβ, samples in red are 4hr 10ng/ml TGFβ, and blue samples are 24hr TGFβ. All samples are done in biological triplicates. On the annotation track at the bottom, there is an annotation for the Col1A1<sup>tm1(TRE<sub>E</sub>_EGFP-miR29a)</sup> transgene insertion site, ~1800bp downstream of the Col1a1 3’UTR.
Figure 3.10 – COL1A1 eRNA in human cells treated with TGFβ

Quantitative RT-PCR for novel eRNA in human HuH7 and LX2 cell lines treated with 10ng/ml TGFβ for 24hr.
IV. DISCUSSION

*Expression of Tg29, in all but Acta2+ cells, during injury is not protective against hepatic fibrosis*

The *Col1A1*\textsuperscript{tm1(TRE_EGFP-miR29a)} mouse harbors a miR-29a containing cassette with a GFP reporter, located downstream of the *Col1a1* gene. The temporal and spatial control of this transgene depend on the presence of doxycycline and of the rtTA, which is expressed from a different locus and under a promoter that can be ubiquitous or specific for the tissue of interest. Using two different ubiquitously expressed rtTAs, *Rosa26_M2-rtTA* and *Tg(CAG_rtTA3)*, we documented similar patterns of expression in many tissues. With either rtTA, the *Col1A1*\textsuperscript{tm1(TRE_EGFP-miR29a)} transgene was able to sustain miR-29a levels during fibrotic injury. This was associated with a qualitative reduction in staining with MTC, a slight reduction of some collagens and fibrosis-associated miR-29a targets at the RNA level. However, this fell short of normalization, and we did not observe a dramatic improvement of fibrosis. This seemingly discordant result of sustained mature miR-29a levels with persisting fibrosis may be explained by spatial expression of the miR-29a transgene. Specifically, the GFP reporter was not expressed in Acta2+ cells in any of the tissues surveyed. The importance of the lack of overlap between Acta2 and GFP is paramount as it has been shown that Acta2 is a marker of activated fibroblasts, also known as myofibroblasts, the major collagen producers in different models of fibrosis\textsuperscript{3,73,74}. Thus, the failure to increase miR-29a levels in the cell most responsible for the production of matrix proteins in fibrosis may account for the minimal protein observed in *Col1A1*\textsuperscript{tm1(TRE_EGFP-miR29a)} mice.
Epigenetic regulation of the Col1a1 locus responsible for proximal gene expression dysregulation

The expression pattern of the Col1A1\textsuperscript{tm1(TRE_EGFP-miR29a)} transgene was distinct and non-random, with absence of transgene expression in specific cells of the liver, and independently of which ubiquitously expressed rtTA was used. The pattern was also not the result of limited doxycycline availability, as evidenced by our \textit{in vitro} studies with MEFs, showing an absence of GFP expression in Acta2\textsuperscript{+} cells. The observed expansion of the GFP\textsuperscript{-} population of cells in the fibrotic tracts \textit{in vivo} also implies these cells have a role in fibrosis. The fact that the transgene insertion site was located so close to the Col1a1 gene, paired with the fact that Col1a1 expression is increased in these cells, prompted us to examine how exposure to TGF-β modified the epigenetic architecture of this locus in association with upregulation of Col1a1 expression. We determined increased Col1A1 transcript expression upon exposure to TGF-β correlated with increased H3K27ac marks proximal to the Col1a1 and transgene insertion loci. These changes and the existence of a novel eRNA derived from this locus was predicted and verified by qPCR in human cell lines treated with TGF-β. All of these data describe a novel phenomenon of epigenetic regulation of transgenes inserted into the commonly used transgene insertion site downstream of the Col1a1 3’UTR.

Epigenetic regulation of Col1a1 in other mouse models

Beyond the scope of our miR-29a expressing mouse, other strains of mice have been created with the same insertion site, either through use of the KH2 ES cells or by other means with targeted insertion downstream of the Col1a1 3’UTR. We were “unlucky”
with our study in that this transgene regulation event occurred in a cell population that is important to the phenotypic outcome of our disease of interest, but others that use this locus for targeted insertion may not initially notice a difference in their model system or disease. In our search of the literature, there are 18 known mouse models using the \textit{Col1a1} locus for targeted insertions. Not included in the list of 18 are \textit{Col1a1}^{FRT-Cre-ER-T2-}\textit{FRT} and \textit{Col1a1}^{FRT-STOP-FRT-Cre-ER-T2} mice created by the Kirsch group at Duke University\textsuperscript{75} as these mice were part of a publication that was later retracted due to an recognized issue of inefficient recombination of their alleles, and which was assumed to be due to reduced antibiotic availability \textit{in vivo}. In the paper, they hypothesize that antibiotic availability \textit{in vivo} did not allow efficient recombination in a subset of cells. Later in the retraction, the group describes \textit{in vitro} experiments on sarcoma cells and MEFs, both of which are high-collagen producing cells, and conclude that inefficient recombination cannot be attributed to lack of drug availability, given the \textit{in vitro} setting. Given that the insertion site described by the Kirsch group maps within 300bp of the \textit{Tg29} transgene locus, we believe that the inefficient recombination they observed might be explained by the same epigenetic silencing we have described.

In conclusion, we created a mouse that offers temporally and spatially controlled miR-29a expression from a transgene downstream of \textit{Col1a1}. As a consequence of the genomic location of the insertion, the transgene is silenced in a high-collagen producing state such as that associated with fibrotic injury or TGF-\(\beta\) stimulation. Being a popular site for targeted transgene insertion, this study informs the scientific community about expression pattern of previously created and putative mouse models using this genomic
locus for insertion. This study may also validate a previously retracted study by providing a potential mechanism to explain their unexpected observations.
V. METHODS

Generation of Col1A1\textsuperscript{tm1(TRE_EGFP-miR29a)} mice

Mouse ES cells (M001315) were obtained from the Fredrick National Laboratory, sponsored by the National Cancer Institute, containing the TRE-GFP\textsubscript{miR29a} construct. The ES cells were injected into blastocysts and implanted into a pseudo-pregnant C57/BL6 mouse by the Johns Hopkins Transgenic Core Facility (https://www.hopkinsmedicine.org/core/home.htm). Chimeric mice were selected for breeding to the C57/BL6 background. Col1A1\textsuperscript{TRE-GFP\_miR29a} mice were genotyped using a 3 primer PCR strategy for the positional insertion and copy number of the transgene downstream of the Col1A1 locus.

<table>
<thead>
<tr>
<th>Primer Type</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT Reverse primer</td>
<td>5’ – ACCTCCCAGGCCTCAGGTCT – 3’</td>
</tr>
<tr>
<td>WT Forward primer</td>
<td>5’ – ATGACATCATCAAGGAAACCCTGG – 3’</td>
</tr>
<tr>
<td>Tg29 Forward primer</td>
<td>5’ – GAGCACCACCTTTCCAAAGCAC – 3’</td>
</tr>
</tbody>
</table>

CCL\textsubscript{4}-induced liver fibrosis in mice

Mice heterozygous for Col1A1\textsuperscript{TRE-GFP\_miR29a} with or without a single copy of B6N.FVB(Cg)-Tg\textsuperscript{(CAG-rtTA3)4288Slowc/J}, aged ~8 weeks, received biweekly intraperitoneal injections of 1ml/kg of carbon tetrachloride, diluted 1:7 in corn oil (Sigma Aldrich; Cat# 289116 and C8267). All mice were placed on doxycycline containing chow (Envigo; Cat# TD.01306), starting 1 week prior to first injury. After 6 weeks (or 12 injections) mice were euthanized and organs were harvested for analysis. All housing and procedures were conducted according the protocol approved by the Animal Care and Use Committee of Johns Hopkins University, School of Medicine.
**Histology preparation and staining**

Tissue was fixed with 10% Formalin for 24hr and then placed in 70% ethanol. The Johns Hopkins Reference Histology Lab embedded tissues in paraffin blocks and performed the Masson Trichrome staining on tissue sections. (Sections were cut [5µm thickness] and fixed to glass slides for histological and immunofluorescent analyses.)

**Immunofluorescence of FFPE tissue sections**

Immunofluorescence for tissue sections used 5µm sections affixed to Fisherbrand™ Superfrost™ Plus Microscope Slides (Fisher Scientific; Cat# 12-550-15). Sections were deparaffinized and rehydrated in sequential incubations in: xylene, xylene, 100% EtOH, 90% EtOH, 70% EtOH, and PBS. Citrate buffer antigen retrieval was conducted to unmask antigens fixed with formalin. Slides were incubated in 90°C citrate buffer for 15 minutes in a 90°C water bath. The Coplin jar containing the buffer and slides were removed from the water bath and allowed to equilibrate to room temperature for 30 minutes. Slides were then dunked in PBS and excess liquid was wiped away with a Kimwipe. Without allowing the samples to dry out, sections were outlined with a hydrophobic pen and blocking buffer was placed on the samples for ~1hr [blocking buffer: 5% FBS in PBS]. Blocking buffer was then aspirated off and primary antibody in blocking buffer was placed on the sections for 1hr. The excess antibody was aspirated off and the sections were washed 3 times with PBS. Secondary antibody, that was raised against the host of the primary antibody, diluted in blocking buffer, was incubated on the sections for 30 minutes. Excess antibody was aspirated off and sections were then rinsed
with 3 washes of PBS before mounted using VECTASHEILD hard-set mounting media containing DAPI (VECTOR Industries; Cat# H-1500). Antibody used: Acta2 [1:200] (Abcam; Cat# ), GFP [1:200] (Abcam; Cat# ).

**Immunofluorescence of formaldehyde-fixed cells**

Immunofluorescence protocol was adapted from Warren DS et al. In brief, cells were grown to appropriate level of confluency, washed with PBS, and then fixed with 3% formaldehyde for 20 minutes at room temperature. Cells were washed 3 times with PBS and then permeabilized with 1% Triton in PBS for 5 minutes. Excess detergent was removed with 3 more washes in PBS. Cell were incubated with 5% BSA in PBS solution for 15 minutes to block nonspecific antigen interaction. The cells were then incubated with primary antibodies at appropriate dilutions in 5% BSA in PBS for 30 minutes. Cells were then washed 3 times in PBS before incubating with secondary antibody raised against the host of the primary (30 minutes in 5% BSA in PBS). Excess antibody was washed away with 5 washes of PBS before mounting in VECTASHEILD hard-set mount containing DAPI (VECTOR Industries; Cat# H-1500). Antibody used: Acta2 [1:200] (Abcam; Cat# ), GFP [1:200] (Abcam; Cat# ).

**Mouse embryonic fibroblast isolation and culture**

Mouse embryonic fibroblasts (MEFs) were isolated from embryos of a Col1a1^{TRE-GFP\_miR29a} heterozygous; M2-rtTA heterozygous breeding. The pregnant female was sacrificed at day e13.5 using halothane (the day the mucosal plug is first observed is considered day e0.5). Removal of the uterine horn, containing embryos, and sequential
washing in 70% EtOH in a 50mL conical tube then 3x PBS in 50mL conical tubes.

Transfer uterine horn into 10cm dish filled with PBS and cut through uterus between the embryos, removing the placenta and yolk sac. Transfer each embryo to a new 6cm dish (move embryos not currently being dissected onto ice). Working on one embryo at a time, remove the head and transfer to a labeled 1.5ml tube for DNA extraction and genotyping. Remove fetal liver (red organ) and other organs (darker colored tissue) from abdomen. Place remaining tissue in lid of 6cm plate and using a sterilized blade, triturate tissue. Transfer disrupted tissue to 15mL conical tube and add 10mL PBS before centrifuging at 500g for 5 minutes at room temperature. Move to sterile hood and aspirate PBS. Disrupt the cell pellet by adding 2mL of trypsin and pipetting up and down, gently. Incubate at 37°C in water bath for 10 minutes. Add 10ml full serum media [DMEM + 10% FBS + AntiAnti + GlutaMAX] to quench trypsin and gently pipette up and down, against side of tube, to break up any remaining clumps. Transfer suspension to a 10cm culture dish (Passage 0). Allow cells to grow to confluency at 37°C in the presence of 5% CO₂ (~2-4 days, changing the media every 24-28hr). Some clumps of debris may be visible on the early passage MEFs—this is normal and will go away with later passages. Passage cells into two 15cm dishes using full serum media supplemented with 2-mercaptoethanol [120µM final concentration] (Passage 1) and allow cells to grow to confluency at 37°C (2-4 days). Trypsinize cells and resuspend at 3x10⁶ cells/mL of MEF Freezing Media [70% DMEM, 20% FBS, 10% DMSO], made fresh. Aliquot 1mL of cells suspension in a 2ml cryopreservation tube with a screw top. Allow to freeze slowly in a -80°C freezer and transfer to liquid nitrogen within a week for long-term storage.
TGFβ stimulation of HuH7 and LX2 cells

HuH7 or LX2 cells were grown in a 37°C incubator in 5% CO₂ in full serum media [FS Media; DMEM (Corning Cellgro; Cat# 10-013-CV) + 10% FBS (ThermoFisher Scientific; Cat# 26140079) + 1x Antimycotic-Antibiotic (ThermoFisher Scientific; Cat# 15240062) + GlutaMAX (ThermoFisher Scientific; Cat# 35050061)]. At ~70% confluency, media was replaced with low serum media [LS Media; DMEM + 0.5% FBS + AntiAnti + GlutaMAX] for 24hr prior to stimulation. Cells were stimulated with LS media ± 10ng/ml of recombinant human TGFβ1 (BioLegend; Cat# 580702) and incubated for 24hr.

RT-qPCR

Total RNA isolated with Trizol reagent (Invitrogen; Cat# 15596026) as per manufacturer’s protocol, and then treated with DNaseI (New England Biolabs; Cat# M0303S) before cDNA synthesis using the High Capacity cDNA Reverse Transcription Kit (Invitrogen, Cat# 4368814). Quantitative PCR using SYBR Green PCR Master Mix (ThermoFisher Scientific; Cat# 4309155) or Taqman Gene Expression Master Mix (ThermoFisher Scientific; Cat# 4369016) were used to amplify and quantify cDNA of interest.
VI. ACKNOWLEDGMENTS

I would like to thank Dr. Greg Hannon and Dr. Scott Lowe and the NCI Mouse Models of Human Cancer Consortium (MMHCC) for creating and making available the mouse ES cells used to generate the transgenic mouse. I would also like to thank Chip Hawkins and Holly Wellington of the JHU-SOM Transgenic Mouse Core for their help with ES microinjections and expertise in mouse generation.
CHAPTER 4 - Identification of TGFβ and miR-29a Targets

Using Tandem High-throughput RNA Analysis
I. ABSTRACT

In normal wound healing, the initial epithelial damage leads to activation of fibroblasts, extracellular matrix (ECM) deposition, repair and subsequent removal of excess ECM. In fibrosis, repeated injury cycles overcome the normal remodeling process and tip the balance toward net accumulation of ECM. In this study, we investigate how TGF-β and miR-29a modulate hepatic fibrosis by analyzing the transcriptome of cells being injured, human hepatocytes, and ECM-producing cells, hepatic stellate cells, separately. By sequencing total RNA from hepatocyte and hepatic stellate cell lines exposed to TGF-β in the presence or absence of a miR-29a oligonucleotide, we tracked the changes caused by fibrotic injury that are reversed with a therapeutic intervention. To determine the likelihood of physical (or direct) miRNA-mRNA interaction in either cell-type, complexes of biotinylated miR-29a, argonaute (AGO) proteins and target mRNAs were isolated and quantified. Combined analyses identified both known and novel pathways and genes regulated by TGFβ and miR-29a in a cell-type specific manner, including a novel eRNA downstream of the COL1A1 locus. We show that, in epithelial cells, miR-29a regulates pro-fibrotic and cancer-related genes including EGFR, RELA, BRAF, MAPK1, MAP4K4, INSR, IRS1, CALM3, and ELK1 and YY1 transcription factors and that it modulates the TGF-β response in epithelial cells through MMP15, paracrine signals, and inhibition of the switch from E-cadherin to N-cadherin. In fibroblasts, miR-29a regulates collagen synthesis and maturation at nearly every level, including modulation of transcripts coding for TGF-β, integrins, platelet derived growth factors (PDGFs), collagens, collagen modifying enzymes P3H4 and SERPINH1, collagen secreting factor SEC31A, and ECM crosslinking enzymes LOX, and SPARC.
II. INTRODUCTION

MicroRNAs (miRNAs or miRs) are small 21-24 nucleotide, non-coding RNAs that modulate expression of target genes through physical interaction via the RNA induced silencing complex (RISC)\textsuperscript{24,50,51,77}. The basal RISC is comprised of an argonaute protein (AGO), miRNA, and target mRNA transcript. MicroRNAs are generally highly conserved in metazoans at the sequence and function levels and are also found in other phyla\textsuperscript{78,79}. In plants, miRNA have more sequence complementarity to their RNA targets making target prediction easier through in silico analysis; in metazoans canonical binding between target mRNA and miRNA occurs in the miRNA seed sequence (nucleotides 2-8 of the mature sequence) and the mRNA 3’UTR. Target prediction using seed sequence binding and evolutionary conservation create large putative target lists that do not always validate experimentally. Other prediction methods include AGO precipitations and high throughput sequencing of bound miRNAs and mRNAs, combined with \textit{in silico} analyses. Some of these studies suggested that 3’ complementary binding as well as GC content can influence miR-mRNA interactions\textsuperscript{52,80}. The gold standard method for miRNA target validation has been luciferase assays in which the effect of a given miRNA is assessed by measuring luciferase activity in cells transfected with a construct in which the 3’UTR of the putative mRNA target is placed downstream of a luciferase expression cassette\textsuperscript{81}. This quantifies the interaction through luciferase activity but requires plasmid cloning, transfection optimization for your cell line of interest, as well as equipment for reading the luciferase assay, all for a single target validation. In this study, target identification is achieved by whole transcriptomic analysis of cells treated with the miRNA of interest, miR-29a, coupled with a pull-down assay that quantifies physical transcriptome-wide
interactions between miRNA and mRNA targets. With an unbiased, high throughput design, we were able to predict and validate miR-29a targets in our cells of interest.

MiR-29a is a potent anti-fibrotic molecule that is significantly reduced after fibrotic injury in multiple organ fibrosis models including heart, lung, liver, and kidney. The fibrotic insult can come in the form of chemical injury, viral infection, cancer, or genetic predisposition. The use of miR-29a as a therapeutic is being researched and implemented by multiple pharmaceutical companies for both cancer and multiple modes of fibrosis. Our previous studies have shown that the therapeutic benefit of forced miR-29 expression in a model of hepatic fibrosis is dependent on the targeted cell-type (Chapter 1,2,3), and highlight the importance of determining the specific targets that miR-29 regulates in each of cellular compartment.

In this study, an unbiased analysis of epithelial cell and collagen producing cells reveals the transcriptomic differences associated with miR-29a in the presence and absence of transforming growth factor beta (TGFβ) stimulation and determines direct miR-29a, thus creating a short list of molecular effectors for putative fibrosis therapeutic targets.
III. RESULTS

* Differential effect of TGF-β on contraction and migration in epithelial cells and fibroblasts *

TGF-β is a potent cytokine that is responsible for activation of fibroblasts *in vivo* and is used *in vitro* to mimic fibrotic injury of the liver. TGF-β has been shown to downregulate miR-29a both *in vitro* and *in vivo*. By studying the effect of TGF-β on cell lines derived from the different cell populations involved in the fibrotic process *in vivo*, we can partially mimic *in vitro* some of the signaling that occurs in injured livers and test the effects of potential therapeutics. HuH7, a hepatocellular carcinoma line, was used as a surrogate for human hepatocytes; LX2 cells are spontaneously immortalized hepatic stellate cells. To measure the effect of exogenous TGF-β and miR-29a on our cells of interest, we transfected HuH7 and LX2 cells with control or miR-29a oligonucleotide and then stimulated these cells with vehicle alone or TGF-β. RNA expression changes were measured 24 hours after TGF-β stimulation (Figure 4.1a). Transfection with miR-29a oligonucleotide was able significantly increase miR-29a expression even in the absence and in the presence of TGF-β, which, as expected, decreased miR-29a levels in un-transfected cells. In order to test the efficacy of the miR-29a mimic, we assessed its effect on known TGFβ-induced, miR-29a targets in LX2 cells. TGF-β significantly increased collagen and SPARC mRNA and protein levels, while transfection with miR-29a was able to reduce expression of these targets to levels comparable to those of non-TGF-β treated cells (Figure 4.2a-b). Interestingly, while these trends were recapitulated in HuH7 at the mRNA level, this cell line expressed overall lower levels of *COL1A1* and *SPARC* (data not shown; higher cycle threshold values in qPCR) and showed no SPARC protein
present by Western Blots, suggesting that additional regulatory mechanism might be at play.

In order to test the physical changes\textsuperscript{87,88} induced by TGF-β on HuH7 and LX2 cells, we used the collagen contraction and scratch migration assays to assess the physical exertion on the cellular environment and cellular migration, respectively\textsuperscript{89,90}. The collagen contraction assay uses the change in the area of a collagen disk, embedded with cells, as a metric for the contractile forces the cells exert on the matrix either at baseline or in response to a given treatment. This metric can be used as a surrogate for a measure of cell activation and migration. The cells are first transfected with small RNAs then embedded into an acellular collagen matrix. Once the matrix has solidified, the disks are floated in a medium with or without 10ng/ml of human recombinant TGFβ1 and the disk area is measured at baseline and every 24hr following (Figure 4.3a,b). HuH7 cells did not show any significant contraction of the collagen disks and were not influenced by exposure to TGF-β or miR-29a transfection. In contrast, LX2 cells showed baseline contraction, which was exacerbated by TGF-β and relaxed by transfection with miR-29a. This suggests that epithelial HuH7 cell line does not physically act upon their environment in the same way as the LX2 fibroblast cell line. Within the LX2 cells’ range of contraction, miR-29a has a greater influence on cell contraction than TGF-β.

The scratch mobility assay is another method to measure the physiological effect of TGFβ and miR-29a on cells \textit{in vitro}. With the same stimulation and transfection conditions as the collagen contraction assay, a confluent monolayer of HuH7 or LX2
cells was scratched with a P200 pipet tip and the area of the scratch was then measured at baseline and 24hr later (Figure 4.4a,b). Similar to what we observed in the collagen contraction assay, HuH7 cells were not responsive to TGF-β or miR-29a (data not shown), while LX2 cells were increasingly migratory when exposed to TGF-β, while miR-29a alleviated this effect. These data suggested the existence of consistent differences in molecular and physiological responses to both TGF-β and miR-29a between the epithelial and mesenchymal derived cell populations that could be examined by a systematical and unbiased approach, with the ultimate goal of furthering our understanding of the events associated with fibrotic injury in the presence or absence of therapeutic miRNAs.

**Total RNA-Seq in HuH7 and LX2 reveals the potency of TGFβ and miR-29a on the transcriptome**

In order to (a) describe the effects of TGF-β on HuH7 and LX2 cells, (b) describe miR-29a targets in HuH7 and LX2 cells, and (c) determine what miR-29a regulates in different cell populations in a setting of injury to ameliorate fibrosis, we sequenced RNA from all of the 8 previously described conditions (HuH7 and LX2 in the presence and absence TGF-β and/or miR-29a) (Figure 4.5a) on a HiSeq2500. Illumina TruSeq Stranded Total RNA Sample kit with Ribo-Zero was used to obtain non-ribosomal and non-coding RNAs as well as mRNAs. We analyzed both individual comparisons to determine the TGF-β effect or miR-29a targets and in combination to assess the impact of both miR-29a and TGF-β. For the latter analysis, we describe the transcripts that are increased with TGF-β as well as those normalized or reduced with miR-29a transfection.
in the presence of TGF-β. First, we used a transcriptome-wide analysis across all conditions [principal component analysis (PCA)] to display the transcriptomic variation across the samples relative to each other (Figure 4.5b). Principal component 1 (PC1 or the x-axis of the PCA plot), which explains the most variation across the samples, segregates the samples by TGF-β stimulation status and PC2 (or the y-axis), which explains the second most variation, segregates the samples according to small RNA transfection conditions. This trend is common between both HuH7 and LX2 cells with TGF-β having a larger effect on the HuH7 transcriptome and miR29a having a larger effect on the LX2 transcriptome [89% vs 75% of variance explained], while miR-29a affects only a subset of these transcripts, despite having a profound effect on fibrosis in vivo.

* TGFβ shows common regulation of ECM components and cell-type specific regulation of different pathways in HuH7 and LX2 cells

Another way to visualize the effect of TGF-β is by an MA plot, in brief an MA plot visualizes the mean normalized read count by the log fold change of each transcript. It can be used to visualize the magnitude and direction of fold change of all transcripts across 2 conditions (Figure 4.6a). The top 75 genes increased with TGFβ- compared to baseline, ranked by fold-change, and with a false discovery rate (FDR) ≤ 0.05, were used to generate a heatmap with hierarchical clustering, which visualizes the relative changes between replicates and the conditions of those individual genes (Figure 4.6b). The heatmaps of the most changed genes show both common and cell type specific differences. Common gene patterns include ECM components, Wnt signaling
components, and transcription factors known to be associated with TGF-β signaling like SMAD3 and JUNB. Cell type specific differences include induction of collagen chaperones in LX2 cells and induction of Nodal signaling agonists, LEFTY1/2, in HuH7 cells by TGF-β. We used WebGestalt\(^\text{\textregistered}\) to find gene ontology (GO) terms significantly overrepresented in the list of transcripts regulated by TGF-β (4,936 increased and 5,049 decreased transcripts across both cell types) (Figure 4.7). The terms focal adhesion, regulation of the cytoskeleton, and integrin signaling were significantly overrepresented in both HuH7 and LX2. Cell type specific patterns such as NFκB signaling and protein synthesis associated pathways were enriched in either HuH7 or LX2 cells, respectively. The LX2 fibroblasts have much greater number of collagen transcripts upregulated as well as collagen maturation components such as SERPINH1 and LOX. Protein and ribosome biosynthesis associated targets like WDR75 and WDR43, emphasized a putative role of fibroblasts as protein factories to synthesize more ECM and other components associated with the wound healing process. The epithelial-derived HuH7 cells didn’t induce collagen maturation associated genes but increased genes associated with adherens junctions, NFκB signaling, as well as endocytosis. To test the hypothesis that TGF-β can be used in vitro to mimic fibrotic injury, a subset of TGF-β-increased transcripts were quantified in 2 different models of fibrosis (Figure 4.8a,b). In both of hepatic and cardiac fibrosis models in mice, the TGF-β-increased transcripts in the experiments with human cell lines were also increased in fibrotic organs. Collagens and integrins, which are known to be increased in fibrosis, as well as newly described TGF-β-induced targets SKP2, S100A16, FYTTD1, and GOLGA7 were significantly increased in
at least one of the fibrosis models. The upregulation of these targets in both mice and humans also suggest the relevance of these transcripts in the context of fibrosis.

In addition to significant mRNA changes, TGF-β influenced non-coding RNA species in both HuH7 and LX2 cells. Loci associated with TGF-β increased transcription showed proximal ncRNA transcription initiation with TGF-β stimulation. Regions downstream of COL1A1, previously described in Chapter 3 of this thesis, and which was confirmed by qPCR, also showed increase in LX2 cells but not HuH7 cells (Figure 4.9a). Other ncRNAs not immediately proximal to the COL1A1 gene also displayed increased levels when stimulated with TGF-β (Figure 4.9b). Of the ncRNAs that were differentially expressed with TGF-β there were both common and cell line specific RNAs (e.g. LINC01969 and a locus downstream of TGFB2 in LX2 cells, respectively). Some of these ncRNAs have had function attributed to them in the literature, such as linc-YY1 in muscle cell differentiation92,93, and others have not been previously described, such as the eRNA derived from the locus downstream of COL1A1.

*Expression data, combined with biotinylated miR-29a pull-down, identifies novel regulatory mechanisms of collagen synthesis by miR-29a*

Our first analysis of the RNA-Seq data showed TGFβ to be a potent regulator of many different transcripts in each cell population and, from the PCA, we would expect TGFβ to have a greater influence on the transcriptome than miR-29a. Continuing our analysis, we compared HuH7 or LX2 cells with or without exogenous miR-29a to assess the miR-29a regulome. These individual comparisons in each cell type will highlight RNA species that
are regulated at the level of transcript abundance, either directly through transcript
degradation or turnover or indirectly through regulation of positive regulator of the
transcript. Of note, miRNAs can also translationally inhibit their targets without effecting
mRNA transcript levels and these targets will be missed in our screen. We know that
miR-29a had less of an impact on the transcriptome than TGF-β, in terms of total number
of significantly changed transcripts compared to unstimulated conditions, and control
transfected cells, but there were still many transcripts regulated by miR-29a, including
2,316 significantly increased and 3,953 significantly decreased transcripts across both
cell lines. The MA plots to visualize the distribution of transcripts regulated by miR-29a
showed a distinct pattern of miR-29a negatively regulating more than positively
regulating transcripts, which aligns with the canonical function of miRNAs as repressive
elements (Figure 4.10a). We also see more significantly decreased transcripts in the LX2
cells compared to HuH7 cells (Figure 4.11). A heatmap with hierarchical clustering of the
top 75 genes, decreased by transfection with miR-29a compared to baseline, ranked on
fold-change with an FDR ≤ 0.05, replicates (Figure 4.10b) includes genes previously
described as miR-29a targets, TargetScan predicted targets of miR-29, and novel miR-
29a regulated RNAs. Predicted targets include ECM components like collagens and
fibrillin and collagen crosslinking enzyme lysyl oxidase, whereas the novel putative miR-
29a regulated transcripts include lysine demethylases and other epigenetic regulators,
protein transport factor SEC16A, as well as ncRNAs. An analysis of GO terms
enrichments in the significantly increased or decreased transcript lists, highlights ECM
organization, focal adhesion, and chromatin organization associated targets in both cell
lines (Figure 4.11).
Having created a list of transcripts that are negatively regulated in the presence of excess miR-29a, we can infer both direct and indirect targets of miR-29a. To further divide these putative targets into direct and indirect targets, we conducted a biotinylated miRNA pulldown assay, previously described by Lal et al.\(^{82}\), to isolate and quantify mRNAs that physically interact with miR-29a. The transcripts can be measured in a high throughput manner using a gene expression bead array. The calculated enrichment ratio (ER) score is used as a quantitative metric for physical interaction between miRNA and mRNA target (i.e., direct target as opposed to indirect target) (Figure 4.12a). HuH7 and LX2 cells were transfected cells with biotinylated control or miR-29a oligo and, after 24 hours, streptavidin was used to isolate biotinylated miRNA in complex with AGO2 and its mRNA targets. Pull-down enriched and input RNAs, collected from cell lysis homogenate prior to pull-down, were measured by sequencing using the Illumina HumanHT-12v4 Gene Expression BeadChip. An ER score is considered significant if it is greater than 2 standard deviations above the mean of all ER scores within a cell type. In HuH7 cells, a significant ER score was $\geq 2.03$ and in LX2 it was $\geq 2.66$. There were 926 and 858 genes with significant ER scores in HuH7 and LX2, respectively. Of the genes with significant ER scores in HuH7 and LX2 cells, 32% and 33% were TargetScan predicted targets of miR-29 (Figure 4.12b). Using the analysis package Sylamer\(^{94}\), we also showed that the miR-29a seed sequences were over-represented in the 3’UTRs of transcripts ranked by ER score in both HuH7 and LX2 cells (Figure 4.13). Of note, the seed sequence corresponding to the control oligo, cel-miR-67, is over-represented at the
bottom of our ranked list suggesting that there are targets of the *C. elgans* miRNA in the human transcriptome.

Combining the analyses from both high-throughput assays, the RNA-Seq and streptavidin pulldown, we were able to compile a list of direct miR-29a targets in both HuH7 and LX2 cells. Transcripts that were significantly reduced with miR-29a as compared to control oligo by RNA-Seq, had a significant ER score, and were expressed with a transcripts per million reads (TPM) above 1 were compiled into a list for each cell line. HuH7 had transcripts of 233 genes, 123 of which were TargetScan predicted targets of miR-29, and LX2 had 288 genes with 142 being TargetScan predicted. A heatmap of the top 75 genes from each cell line revealed targets from a variety of molecular pathways (Figure 4.14). The lists included ECM components and collagen chaperones, genes associated with multiple cancers including pancreatic cancer and renal cell carcinoma, as well as the insulin signaling pathway (e.g., previous predicted targets: *EGFR, RELA, BRAF, MAPK1, MAP4K4, FLOT1/2, PMP22* and the *ELK1* and *YY1* transcription factors; novel putative targets: *INSR, IRS1, CALM3, and MMP15*).

*MiR-29a differentially regulates TGF-β-mediated injury in cells of different lineages*

Our analysis of the individual influences of TGFβ or miR-29a in our cell lines and the comparison of TGF-β-treated cells to TGF-β- and miR-29a-treated cells (i.e., ‘injured’ cells ± therapeutic intervention) can inform the study of physiologically relevant TGF-β-induced, miR-29a targets. MiR-29 is sufficient to ameliorate the fibrotic phenotype and, therefore, it must regulate an important molecular target in the context of injury. The
criteria to create a gene list of physiologically relevant, miR-29a targets consists of: (a) a gene that is increased with TGF-β stimulation, (b) regulated by miR-29a, (c) reduced or normalized when co-stimulated TGF-β and miR-29a. When we visualize the comparison between the TGF-β-treated and the TGF-β- and miR-29a-treated cells, we see that very few transcripts are significantly changed (Figure 4.15a). If we further subset the list of significantly changed transcripts by transcripts that were significantly increased with TGF-β, reduced with miR-29a, reduced or normalized in the presence of TGF-β and miR-29a, and had a significant ER score, we obtain a list of 77 and 114 directly targeted genes in HuH7 and LX2 cells, respectively. These genes were used to generate a heatmap visualizing the relative expression changes of the given genes across all conditions and replicates (Figure 4.15b). Of the 77 and 114 genes identified, nearly half (~53%) were TargetScan predicted targets of miR-29, meaning that the molecular screen assays with specific set of filters identified 36 and 52 novel miR-29a targets with respect to fibrotic injury across the 2 cell lines. In the HuH7 cells, novel miR-29a targets included MMP15, PDGFA, TTYH3, ITGA5, and the ncRNA linc-YY1. Indirectly in HuH7 cells, miR-29a inhibited the loss of epithelial markers and increase of mesenchymal markers E-caderhin (CDH1) and N-cadherin (CDH2), respectively (Figure 4.16a). Interestingly, expression of cytokines PDGFA and PDGFC are directly regulated in HuH7 cells, where there is little to no expression of PDGF receptors, providing evidence for miR-29a in the hepatocyte fraction being at least partially protective by modulating the paracrine signaling caused by fibrotic injury to the epithelium. MMP15 has also been implicated in progression of EMT in carcinomas and provides some more evidence for restoration of miR-29a hindering the progression of fibrosis, in this case through depletion of hepatocyte-derived
myofibroblasts. Novel putative miR-29a targets that are increased with injury in LX2 cells include collagen maturation proteins \textit{P3H4} and \textit{COLGALT1} (Figure 4.16b). All levels of collagen synthesis, maturation, secretion, and crosslinking are modulated by miR-29a, ranging from cytokines, their respective receptors, secondary messengers, collagen transcripts, collagen chaperones and other protein maturation machinery, secretory vesicles, and the extracellular modifying enzymes. MiR-29a acts upon nearly every aspect of collagen production expressed in these fibroblasts. In summary, miR-29a modulates the effect of TGFß through different pathways in different cell types. From regulating EMT and paracrine signaling in the epithelium to regulating all aspects of collagen production, secretion, and maturation, intra- and extracellularly, miR-29a is protective against fibrotic injury to varying degrees and in varying ways in multiple cell types.
Figure 4.1 – Mature miR-29a levels in HuH7 and LX2 cells ± TGFB ± miR29a mimic

Mature miR-29a levels in HuH7 and LX2 cells measured by Taqman qPCR. RNA was collected 24hr after 10ng/ml TGFB stimulation began. Samples within each cell line were normalized to control transfected, saline treated cells.
Figure 4.2 – Expression of COL1A1 and SPARC in TGFβ stimulated cells ± miR-29a mimic

(A) RNA levels of COL1A1 measured by qPCR and normalized to βActin expression. (A’) RNA levels of SPARC measured by qPCR and normalized to βActin expression. (B) Western blot analysis of protein from HuH7 and LX2 cells 24hr after 10ng/ml TGFβ stimulation or saline control. 10µg of protein lysate was loaded into each lane and proteins were quantified and normalized to βActin levels. Sample within cell lines were normalized to control transfected, saline treated cells.
Figure 4.3 – Collagen contraction assay in HuH7 and LX2 cells

(A) Solidified collagen disks embedded with cells transfected with miR-29a or control oligo and stimulated with 10ng/ml final concentration of TGFβ after 48hr. HuH7 cells on top row; LX2 cells on bottom row. (B) Quantification of area of collagen disks at given time points after TGFβ stimulation. HuH7 cell in shades or tints of red; LX2 cells in shades of blue.
Figure 4.4 – Scratch mobility assay in LX2 cells

(A) Transfected with control or miR-29a, confluent LX2 cells were scratched with P200 pipet tip and then stimulated with TGFβ or saline. Baseline images were taken immediately after scratch was made and 24hr after. (B) The area of scratch was measured in pixels from images of the same resolution of the same area of the scratch. Measurements were taken at 3 areas per well for technical triplicates and 3 wells were used for biological triplicates at each time point.
Figure 4.5 – RNA-Seq experimental design and PCA

A

HuH7  or  LX2

Unstimulated  TGFβ

Control  miR-29a  Control  miR-29a

B

(A) Experimental design of cell types and conditions that were used in the RNA-Seq experiments.
(B) Principal component analysis on all transcripts in all samples. Samples segregated first by TGFβ simulation vs vehicle control in PC1 and then segregated by miRNA oligo transfection in PC2.
Figure 4.6 – Analysis of TGFβ-increased transcripts in HuH7 and LX2

(A) MA Plots of the comparison between baseline and TGFβ treated cells. (B) Heatmap showing the expression difference of genes that are increased in the TGFβ fraction with an FDR ≤ 0.05 and are expressed [TPM ≥ 1.0]. Each column represents a replicate of RNA-Seq and color corresponds to relative expression (within a row) for a given gene.
The Venn diagram displays the number of transcripts either significantly increased or decreased with TGFβ [red = HuH7 specific; blue = LX2 specific; purple = dysregulated in both cell types]. Terms on the right are GO terms enriched in lists of transcripts of a given condition (i.e., increased or decreased with TGFβ stimulation when compared to vehicle treated cells).
Figure 4.8 – In vivo fibrotic injury recapitulates in vitro TGFβ stimulation in multiple models of fibrosis

RNA expression of a subset of TGFβ increased transcripts from RNA-Seq analysis in multiple fibrosis models. (A) RNA isolated and quantified from control and fibrotic livers from 4 weeks CCl4 treated mice. [N = 5 per group] (B) RNA isolated and quantified from control and fibrotic hearts from 7-day saline and angiotensin II-treated mice. [N = 5 per group]
Figure 4.9 – Differential ncRNA expression proximal to TGFβ-increased genes

(A) Integrated Genome Viewer (IGV) track showing read pileups at the COL1A1 locus. Inset is the region downstream of the 3’UTR of COL1A1 showing increased number of reads in the TGFβ stimulated LX2 cells. (B) Integrated Genome Viewer (IGV) track showing read pileups at different loci in the human genome. Top: ~30kb downstream of the 3’UTR of TGFB2. Bottom: the LINC01969 genomic locus.
Figure 4.10 – Analysis of miR-29a regulated transcripts in HuH7 and LX2

(A) MA Plots of the comparison between control and miR-29a oligo transfected cells. (B) Heatmap showing the expression difference of genes that are decreased in the miR-29a fraction with an FDR ≤ 0.05 and are expressed [TPM ≥ 1.0]. Each column represents a replicate of RNA-Seq and color corresponds to relative expression (within a row) for a given gene.
Figure 4.11 – GO terms analysis in transcripts that are changed with exogenous miR-29a

The Venn diagram displays the number of transcripts either significantly increased or decreased with miR-29a transfection [red = HuH7 specific; blue = LX2 specific; purple = dysregulated in both cell types] Terms on the left are GO terms enriched in lists of transcripts of a given condition (i.e., increased or decreased with miR-29a transfection when compared to control oligo transfected cells).
Figure 4.12 – Biotinylated miRNA pulldown assay

(A) Schematic of the biotinylated miRNA pulldown assay, modified from Lal A et al. (2011).
Cells transfected with biotinylated control or miR-29a oligo had input and pulldown RNA fractions quantified on gene expression array and enrichment ratio (ER) score calculated. (B) Plot of the individual ratios that go into the ER score [x-axis = enrichment of transcripts in miR-29a pulldown fraction vs control oligo; y-axis = ratio of depletion of transcript expression from miR-29 input fraction vs control oligo]. Highlighted in red are TargetScan predicted miR-29a targets.
Figure 4.13 – Sylamer analysis of the 3’UTRs of ranked gene lists

Sylamer analysis looking at enrichment of miRNA seed sequences in the 3’UTRs of a ranked list of transcripts. The transcript lists were ranked on ER scores, largest to smallest, of all probes assayed on the Illumina beadchip. MiR-29 seed had the largest peak of enrichment while the seed sequence corresponding to the control oligo, cel-miR-67, had the lowest peak.
Figure 4.14 – Heatmaps of putative miR-29a targets in HuH7 and LX2 cell lines

Heatmap showing the expression difference of genes that are decreased in the miR-29a fraction with an FDR ≤ 0.05 and are expressed [TPM ≥ 1.0]. Each column represents a replicate of RNA-Seq and color corresponds to relative expression (within a row) for a given gene.
Figure 4.15 – Analysis of TGFβ-regulated, miR-29a target transcripts in HuH7 and LX2

(A) MA Plots of the comparison between TGFβ stimulated cells with control or miR-29a oligo.
(B) Heatmap showing the expression difference of genes across all conditions. Genes in the
heatmaps are expressed [TPM ≥1.0], significantly increased with TGFβ and decreased with miR-
29a relative to WT and reduced with both TGFβ and miR-29a. [Significance is an FDR ≤ 0.05].
Each column represents a replicate of RNA-Seq and color corresponds to relative expression
(within a row) for a given gene.
Figure 4.16 – Regulation of TGFβ response by miR-29a in HuH7 and LX2 cells

MiR-29a-mediated TGFβ signaling in HuH7 cells (A) and LX2 cells (B). Highlighted in red are miR-29a targets that are increased with TGFβ and reduced or normalized in the presence of miR-29a.
IV. DISCUSSION

In vitro versus in vivo: TGFβ stimulation is a good surrogate for in vivo fibrotic injury

In vivo fibrosis models in mice have been a great resource to dissect pathology of molecular events and the histological changes that occur during and after injury. When tissue is sampled for molecular studies, a biopsy of heterogeneous cellular makeup is taken. Then an averaging technique, such as Western blot or qPCR, is used to measure protein or RNAs of interest. However, due to sampling of multiple cell populations in different ratios, varying from biopsy to biopsy, significant but rare differences are easily missed. Immunofluorescence and other single target visualization methods are very helpful when the list of targets to interrogate is small but are difficult and expensive for even medium throughput analyses. Digestion and sorting of cells from wild-type and injured organs is also technically difficult due to the fibrotic nature of the disease. To understand fibrosis and how fibrotic injury can influence different cell population in an organ, we sought to develop a system in which to assess the response of individual cell populations to equal injury and therapeutic perturbations. In this study, hepatocyte and hepatic stellate cell lines (HuH7 and LX2 cells, respectively) were used as a surrogate for epithelial and collagen producing cell populations within the injured liver, and injury and treatment regime was applied uniformly to cell monolayers. The validation of some of the molecular changes observed in vitro in multiple in vivo models of fibrosis highlights the potential of our approach to the identification of physiologically relevant, fibrosis-associated targets.

MiR-29a targets more than just fibrosis: miR-29a in cancer
Conducting a transcriptome wide survey of changes in RNAs in the presence or absence of miR-29a, in conjunction with a physical interaction assay, gave an unbiased approach for miR-29a target determination in multiple cell types. Not only did we identify novel fibrosis associated targets, we also showed that the regulome of miR-29a encompasses genes associated with renal and pancreatic cancer, epigenetic regulation, as well as ECM production and maturation. The function of miR-29a goes beyond just the collagens and ECM effectors of fibrosis. The roles of miR-29a in cancer are being actively researched and some studies have shown miR-29 to be dysregulated with fibrotic cancers such as pancreatic cancer or hepatocellular carcinoma and correlate with aberrant epigenetic machinery and increased metastasis\textsuperscript{40,96}. Understanding what miR-29 regulates in different transcriptomic contexts (i.e., cells of different lineages with their different expression patterns like in HuH7 versus LX2 cells) helps inform the scientific and medical community about the global effects when miR-29 is dysregulated in a multiple cell populations or tumors. Showing that miR-29a, in epithelial-derived HuH7 cells, regulates paracrine signaling using the platelet derived growth factor (PDGF) signaling pathway hints at how an epithelial derived cancer can create fibrotic tumors like pancreatic carcinoma. Additionally, regulation of collagen protein, maturation, and secretion by miR-29a in the fibroblast LX2 cell lines, explains how a significant increase in collagen and ECM productions can come from a modest (~30-50%) reduction of miR-29a levels during fibrotic injury. It also highlights the potency of miR-29a and explains how restoration to non-injury levels in the AAV study was sufficient to prevent and reverse hepatic fibrosis. Novel targets of miR-29a such as \textit{FZD5}, \textit{MMP15}, \textit{COLGALT1}, or \textit{ITGA5} are all associated with profibrotic pathways and are consistent with the
antifibrotic effect of miR-29a. Other interesting, novel miR-29a targets include ncRNAs such as linc-YY1 which has been shown to influence the specificity of the transcription factor Ying Yang 1 (YY1), including at the miR-29 locus\textsuperscript{92,93}.

\textit{ncRNAs: Functional or Consequential?}

The non-coding RNAs were quantified by RNA-Seq but, unfortunately, were not included on the Illumina gene expression array and, therefore, the physical interaction between miRNA and ncRNA could not be quantified. Multiple ncRNAs regulated by either TGF\(\beta\), miR-29a, or both were discovered and quantified through our RNA-Seq experiments and some of these ncRNAs were then validated in replicated experiments by qPCR. The function of these ncRNA is not well characterized and, in some cases, completely unknown. Of note, we noticed overlap of some of these ncRNAs with epigenetic marks that correlate with open chromatin. In the literature some of these ncRNAs have been characterized as enhancer RNAs (eRNAs); their exact role in epigenetic regulation has not been fully elucidated. The putative eRNA, downstream of the \textit{COL1A1} locus, that we have shown to be increased with TGF\(\beta\) stimulation, has not been seen in the FANTOM CAT5 ncRNA database\textsuperscript{97}. In other studies, it has been shown that there is a change in the epigenetic landscape of the analogous locus in mice\textsuperscript{69}. We propose that in wild-type mouse cells, treated with TGF\(\beta\) or injured with a profibrotic stimulus, we would see a significant increase in the eRNA from this locus. It would also be an interesting experiment to knock down these eRNAs with an siRNA in the presence of TGF\(\beta\) to see if the effect on Col1a1 transcript levels and determine whether this eRNAs has direct role in local gene expression.
The aim of this study was to determine the effect of the profibrotic cytokine TGFβ and the antifibrotic miRNA miR-29a on cell populations of different lineages. We also aimed to determine if TGFβ is not only a good surrogate for molecular characterization of fibrotic injury, but also one of cellular physiology. With our careful interrogation with cellular migration assays and high throughput transcriptomic approaches, we were able to show that TGF-β causes different cellular migration and contraction phenotypes in these epithelial and mesenchymal derived cells that can be characterized and quantified. We also showed that TGF-β induces both common and cell type specific responses (e.g., differential contraction and migration, ECM production as well as regulation through different cellular signaling pathways). The regulome of miR-29a was shown to influence genes involved in fibrosis, cancer, and epigenetic modifications. Novel miR-29a targets like MMP15, FZD5, and COLGALT1 were shown to physically interact with miR-29a and TGF-β induction is normalized by the addition of exogenous miR-29a. Through our studies we have identified putative targets for that may form the basis for future antifibrotic therapies.
V. METHODS

Small RNA transfection and TGFβ stimulation of HuH7 and LX2 cells

HuH7 or LX2 cells were grown in a 37°C incubator in 5% CO₂ in full serum media [FS Media; DMEM (Corning Cellgro; Cat# 10-013-CV) + 10% FBS (ThermoFisher Scientific; Cat# 26140079) + 1x Antimycotic-Antibiotic (ThermoFisher Scientific; Cat# 15240062) + GlutaMAX (ThermoFisher Scientific; Cat# 35050061)]. At ~70% confluency, media was replaced with low serum media [LS Media; DMEM + 0.5% FBS + AntiAnti + GlutaMAX] for 24hr prior to transfection. Cells were transfected with hsa-miR-29a or cel-miR-67 control oligo (Dharmacon; Cat# C-300504-07-0005 and CN-001000-01-05) using Lipofectamine RNAiMAX (ThermoFisher Scientific; Cat# 13778075) per the manufacturer’s protocol. After 3hr with transfection reagent, the media was replaced with LS media ± 10ng/ml of recombinant human TGFβ1 (BioLegend; Cat# 580702) and incubated for 24hr.

RT-qPCR

Total RNA was isolated with Trizol reagent (Invitrogen; Cat# 15596026) per the manufacturer’s protocol, and treated with DNaseI (New England Biolabs; Cat# M0303S) before cDNA synthesis using the High Capacity cDNA Reverse Transcription Kit (Invitrogen, Cat# 4368814). Quantitative PCR using SYBR Green PCR Master Mix (ThermoFisher Scientific; Cat# 4309155) or Taqman Gene Expression Master Mix (ThermoFisher Scientific; Cat# 4369016) was used to amplify and quantify cDNA of interest. GraphPad Prism was used for plot generation and statistical analysis where p-values were generated with two-tailed Student T-tests when comparing two groups and a
two-way ANOVA and Krustal multiple hypothesis testing when comparing more than two groups. All groups were analyzed with biological triplicates in each condition and each cell line expression levels are normalized to the control oligo, saline treated condition. Each experiment was conducted at least twice.

**Collagen gel contraction assay**

A modified protocol from Ngo P et al. (2006)\(^89\) was used for the collagen gel contraction assay. In brief, acellular collagen solution was embedded with HuH7 or LX2 cells transfected with small RNAs and placed in wells of a 24-well plate. Once collagen disks solidified, they were floated in equivalent disk volume of low serum media ± 20ng/ml TGFβ [10ng/ml final concentration] and allowed to contract for 72hr, taking pictures every 24hr. ImageJ software was used to measure disk size at varying timepoints.

**Scratch Mobility Assay**

The scratch mobility assay was modified from Liang CC et al. (2006)\(^90\). In brief, HuH7 and LX2 cells were grown to ~75% confluency and then serum starved for 24hr before being transfected with miR-29a or control. After 3hr with transfection reagents in media, a scratch was made with a P200 pipet tip in a straight line across the cells. Cells were washed once in warm PBS and fresh LS media ± 10ng/ml TGFβ1 was added. At the same magnification and resolution, pictures were taken at time of scratch and 24hr post cell monolayer injury. Each scratch was measured and quantified with ImageJ software at 3 different places along the scratch.
Western Blot for protein analysis

Cells were lysed in M-PER™ Mammalian Protein Extraction Reagent (ThermoFisher Scientific; Cat# 78501) and quantified using a BCA protein assay (ThermoFisher Scientific; Cat# 23225). Approximately 10µg of protein was loaded into each lane of 10% Criterion™ XT Bis-Tris acrylamide gel (Bio-Rad; Cat# 3450112) and run at constant voltage for ~1.5hr. The protein was then transferred to PVDF membrane using Bio-Rad Trans Blot Turbo transfer system, as per manufacturer’s protocol. The membrane was then blocked for 1hr in Odyssey® Blocking Buffer (LI-COR; Cat# 927-50000), incubated with 1° antibody overnight at 4°C, washed 3 times with PBS + 0.1% Tween20 (PBST), incubated with 2° antibody (IRDye® 800CW or 680LT from LI-COR) for 1hr at room temperature, washed 3 times with PBST, and rinsed with PBS twice before imaging on an Odyssey imaging system. Primary antibodies include: Actb [1:5000] (Sigma; Cat# ), Acta2 [1:1000] (Abcam; Cat# ), and Serpinh1 [1:1000] (ThermoFisher Scientific; Cat# ). Proteins of interest were quantified using ImageStudio Lite software from LI-COR and normalized to Actb levels. All groups were analyzed with biological duplicates in each condition and the expression levels of each cell line are normalized to the control oligo, saline treated condition. Each experiment was conducted at least twice.

RNA-Sequencing preparation and analysis

Total RNA was isolated from HuH7 and LX2 cells using Trizol reagent (Invitrogen; Cat# 15596026) as per manufacturer’s protocol. The Illumina TruSeq Stranded Total RNA Sample Prep kit (Illumina; Cat# 20020596) was used for library preparation and the
samples were sequenced on the Illumina HiSeq2500, Rapid Run 100bp x 100bp paired end run. Reads were aligned to the genome using RSEM, originally described by Li B and Dewey CN (2011)\(^98\), and transcripts were quantified and compared using DESeq2\(^99\).

**Biotinylated miRNA pulldown assay and gene expression array**

The miRNA-mRNA capture protocol was adapted and modified from Lal A et al. (2011)\(^82\). Using 6-well plates, 3 wells per replicate were transfected (~1 x 10\(^6\) cells per well) with biotinylated hsa-miR-29a and 3 wells per replicate with control cel-miR-67, both at 20 nM. Cells were incubated at 37\(^\circ\)C in 5% CO\(_2\) for 24hrs before harvesting.

Activate the Streptavidin-Dynabeads (Invitrogen; Cat# 112.05D) were activated by transferring 100\(\mu\)l of beads to new tube and wash 3 times with 500\(\mu\)l Binding Buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1M NaCl) using a magnetic separator to remove excess buffer between washes. Next, wash beads twice with 500\(\mu\)l Buffer A (0.1 M NaOH and 0.05 M NaCl) and once with 500\(\mu\)l Buffer B (0.1 M NaCl). Block the beads with 500\(\mu\)l Lysis Buffer [20 mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM MgCl\(_2\), 0.3% IGEPAL CA-630 (Nonidet P40 substitute, Fluka)] supplemented 10\(\mu\)l RNase-free BSA (ThermoFisher Scientific; Cat# AM2616) and 10\(\mu\)l RNase-free yeast tRNA (ThermoFisher Scientific; Cat# AM7119) for 30 minutes at 4\(^\circ\)C with rotation. Spin briefly to remove liquid from the cap before washing twice with 500\(\mu\)l Lysis Buffer. The activated beads can be stored short term at 4\(^\circ\)C. Wash the cells with PBS and then trypsinize them with 500\(\mu\)l TrypLE Express (Gibco; Cat# 12605-010). Add 500\(\mu\)l FS media to each well, combine triplicates, and transfer cells to 15ml conical tube. Centrifuge at 1600 RPM for 4 minutes at room temperature, remove supernatant, and
wash cell pellet with 3mL PBS. Spin again, remove PBS, and resuspend pellet in 700µl Lysis Buffer + 3.75µl RNaseOUT (Invitrogen; Cat# 10777019) + 7.5µl 100X Halt protease inhibitor cocktail (ThermoFisher Scientific; Cat# 87786). Transfer lysate to a 1.5ml Eppendorf tube and pipet up and down 20 times with P1000. Incubate on ice for 20 min. Centrifuge lysate at 10,000xg for 15 minutes at 4°C. Take 50ul of lysate (to be used as input) and add 1 ml Trizol (incubate for 5 minutes to allow the complete dissociation of nucleoprotein complexes). Mix by inversion and store at -80°C to be processed at same time as pulldown fractions. Add remaining lysate to the pre-blocked Dynabeads and incubate with rotation at 4°C for 4hr. Briefly spin to remove beads that adhered to caps and remove unbound material using magnetic separators. Wash 5 times with 500µL Lysis Buffer. Add 100µl of Lysis Buffer + 10ul of DNasel (1U/ul) and incubate at 37°C for 7 mins. After incubation, add 500µl Lysis Buffer and remove supernatant using magnetic separator. Next, add 100ul of Lysis Buffer + 3ul of 20 mg/ml Proteinase K (Promega; Cat# PRMC5005) and 1ul of 10% SDS to the beads. Incubate at 55°C for 15-20min. Using magnetic separator, collect supernatant (~100ul). To the beads, add 200ul of Lysis Buffer and using the magnetic separator, collect the supernatant (~200ul) to combine with previous supernatant. To isolate RNA, add 300ul from the lower layer of acid phenol:CHCl3:isoamyl alcohol (Invitrogen; Cat# AM9720) and mix well for 1 minute room temperature. Centrifuge for 15 minutes at max speed (>16,000xg) in table top centrifuge and collect ~250ul of upper aqueous layer. Add 1:10 volume of 3M NaOAc and 5ul of Glycoblu (ThermoFisher Scientific; Cat# AM9515) and mix well. Then add 2.5 volumes of cold 100% EtOH and incubate overnight at -20°C. [Note: Add NaOAC and glycoblu BEFORE EtOH]. Pellet RNA by centrifuging at max speed for 30 minutes
at 4°C. Discard supernatant, taking care not to disturb pellet, and add 350µL 70% cold EtOH. Centrifuge at max speed for another 5 minutes and remove supernatant, again taking care not to disturb the RNA pellet. Spin down tubes for a quick spin and with a P20 remove remaining EtOH. Let samples dry for 5 minutes with caps open and resuspend RNA in 25µl nuclease-free H₂O. RNA samples were then submitted to the Sidney Kimmel Cancer Center Microarray Core at Johns Hopkins University School of Medicine for further processing. The samples were then run on Illumina HumanHT-12v4 Gene Expression BeadChip for quantification of ~48k transcripts.
VI. ACKNOWLEDGEMENTS

I would like to thank our collaborator Ashish Lal for the guidance and supplies to complete the biotinylated miRNA pulldown assay. I would also like to thank Sarah Wheelan, Paul Schaughney, Luigi Marchioni, and the Next Generation Sequencing Core at Sydney Kimmel Cancer Center at Johns Hopkins University School of Medicine for their help and insight of the bioinformatic analysis of the high throughput datasets.
CONCLUSION AND FUTURE DIRECTIONS
Fibrosis is a disease of many etiologies and affects all major solid organs. The medical and financial burden of fibrosis on society is great and is compounded by the lack of good therapeutic interventions. MiR-29 is a potent antifibrotic molecule that has been shown in multiple studies, performed by us and others, to prevent or reverse established fibrosis. It has thus a great potential as an antifibrotic treatment.

Using a self-complementary adeno associated virus (scAAV) as a vector for miR-29a transduction, we were able to prevent and reverse existing fibrosis in a murine hepatic fibrosis model. Using a miR-29b oligonucleotide with a modification for hepatocyte-specific uptake, we tested the efficacy of subcutaneous injections of miRNAs for reversal of the same murine hepatic fibrosis model. Although there were qualitative differences in collagen content, miR-29 oligos did not prevent the development of fibrosis. To further investigate the potential therapeutic benefit of increased miR-29 levels, we used embryonic stem cells with a tetracycline regulated (TRE) GFP and miR-29a cassette inserted downstream of the Col1a1 locus to derive a new mouse line with temporal and spatial control of miR-29a overexpression (Col1A1tm1(TRE_EGFP-miR29a)).

Interestingly, the Col1A1tm1(TRE_EGFP-miR29a) mouse was not protective against hepatic fibrosis. Upon further interrogation of the expression pattern of the transgene, we saw that the Acta2+ population of the liver did not express GFP and, therefore, transgene derived miR-29a. With further ex vivo analyses of mouse embryonic fibroblasts and a meta-analysis of the published data we determined that the transgene is likely epigenetically regulated by changes to the Col1a1 locus chromatin landscape triggered by fibrotic injury. Together our data suggest that elevated miR-29a in the myofibroblast population is key to preventing fibrosis. To further tease out the role of miR-29a as an
antifibrotic agent, we utilized an *in vitro* system with homogenous cellular populations and environments. We were able to characterize HuH7 and LX2 cell lines using high throughput sequencing techniques to show that TGFβ influences many ECM production pathways, as well as upregulates protein production machinery and different signaling pathways in the respective cell lines. MiR-29a counters some of the TGFβ induced changes by normalizing of collagen production, paracrine signaling, and novel ncRNAs as well as the switch of cadherins from E-cadherin to N-cadherin.

To further understand the role of miR-29a on ncRNAs and the regulation on the epigenetic landscape, we are planning to conduct another high throughput assay to interrogate the changes in chromatin accessibility with TGFβ and/or miR-29a. The assay for transposase-accessible chromatin using sequencing (ATAC-Seq) is a powerful tool when paired with RNA-Seq to show genome wide chromatin changes that correlate with gene expression changes. This approach will facilitate a better understanding of the epigenomic changes that are induced by fibrotic injury and how miR-29a influences those changes.

In conclusion, I hope my data further informs the scientific community as to the molecular mechanism of miR-29a and hope that this knowledge ultimately leads to the development of an effective therapy for fibrotic disorders.
Appendix - Capn9\textsuperscript{\textminus/-} Mice are Protected Against EMT-mediated Fibrosis in Multiple Murine Models of Injury\textsuperscript{*}

\textsuperscript{*} Excerpts and figures in this appendix have been reproduced or adapted with permission from Kim D and Beckett JD et al. Calpain 9 as a Novel Therapeutic Target in TGF\textbeta{}-Induced Mesenchymal Transition and Fibrosis. (2018).
I. EXPLANATION OF PROJECT PARTICIPATION

The ‘Capn9 project’, to be described further in this appendix, investigated the role of calpains in epithelial to mesenchymal transition (EMT) during fibrotic injury in multiple organ systems and in vitro. This work was done separately from, but in parallel to, my main thesis work and was similar to some of my thesis aims in that we investigated the mechanisms of fibrotic injury and possible therapeutic treatments (e.g., EMT and calpain inhibition). Work conducted on the ‘Capn9 project’ spanned multiple research groups including ours (Warren lab), other groups at Johns Hopkins University (Dietz and Mitzner labs), and a private business venture group, Blade Therapeutics. The culmination of experimental results from multiple participants resulted in a publication, currently under final stages of the review process at Science Translational Medicine, and the excerpts and figures from the manuscript will be described in this appendix as it was submitted to the journal for publication. The contributions to the ultimate publication were great by multiple authors, including myself who participated in the husbandry and colony maintenance of Capn9−/− mice, in vivo experiments of fibrosis models in the lung, heart, and liver, aiding in the preparation and writing of the manuscript, and experiments in response to reviewers’ comments. As I am not first author and have not written the bulk of the manuscript submitted, I wish to give credit to James Beckett and Dave Kim as first authors and Hal Dietz as senior author of the manuscript. My contribution to this project was a significant portion of my time and effort during my graduate education and I wish to describe this work as part my thesis. Proceeding are excerpts, either in part or in full, from the manuscript that encompass the aspects of the project to which I contributed. Many more experiments were conducted under the umbrella of the ‘Capn9’ project, by
our and other groups, which were not included in the final manuscript including other
organ fibrosis models and different time points of the liver model, but they will not be
further described.
II. ABSTRACT FROM MANUSCRIPT

Fibrosis is a common pathologic outcome of chronic disease resulting in the replacement of normal tissue parenchyma with a collagen-rich extracellular matrix by myofibroblasts. While the progenitor cell types and cellular programs giving rise to myofibroblasts through mesenchymal transition can vary between tissues and diseases, their contribution to fibrosis initiation, maintenance, and progression is thought to be pervasive. In this study we showed that the ability of transforming growth factor-β (TGFβ) to efficiently induce myofibroblast differentiation of cultured epithelial cells, endothelial cells, or quiescent fibroblasts is dependent upon the induced expression and activity of dimeric calpains, a family of non-lysosomal cysteine proteases that regulate varied cellular events through post-translational modification of diverse substrates. siRNA-based gene silencing demonstrated that TGFβ-induced mesenchymal transition of a murine breast epithelial cell line was dependent upon induction of expression of calpain 9 (CAPN9), an isoform previously thought to be restricted to the gastrointestinal tract. Mice lacking functional CAPN9 due to biallelic targeting of Capn9 were viable and fertile, but showed overt protection from bleomycin-induced lung fibrosis, carbon tetrachloride-induced liver fibrosis, and angiotensin II-induced cardiac fibrosis and dysfunction. A predicted loss-of-function allele of CAPN9 is common in Southeast Asia, with the frequency of homozygosity matching the prediction of Hardy-Weinberg equilibrium; together with the highly spatially-restricted pattern of CAPN9 expression under physiologic circumstances and the heartiness of the murine knockout, these data provide a strong signature for tolerance of therapeutic strategies aimed at CAPN9 antagonism.
III. INTRODUCTION FROM MANUSCRIPT

The replacement and distortion of tissue parenchyma with fibrillar collagens and other extracellular matrix (ECM) proteins – thereby compromising organ function – is a common feature of chronic disease and contributes to nearly 45% of deaths in the industrialized world\textsuperscript{1,2,100}. While collagen deposition is an indispensable component of tissue homeostasis, chronic injury or dysregulation of wound-healing can lead to pathologic scarring, a condition termed fibrosis\textsuperscript{101}. In some instances, provocations that induce tissue fibrosis have been identified, for example: genetic conditions (e.g., stiff skin syndrome\textsuperscript{102}, SP-C deficient interstitial lung disease\textsuperscript{103}, short telomere pulmonary fibrosis\textsuperscript{104}); persistent infections (e.g., hepatitis C-induced liver cirrhosis); chemical exposures (e.g., bleomycin-induced lung fibrosis, gadolinium associated nephrogenic systemic fibrosis, alcoholic liver cirrhosis); and chronic inflammation secondary to autoimmune disorders (e.g., systemic sclerosis). In other cases, such as the majority of idiopathic pulmonary fibrosis presentations, the driver of fibrosis is unknown.

Regardless of the initiating events, all fibrotic disorders show accumulation of activated fibroblasts that are invasive, synthetic, contractile, proliferative, and long-lived\textsuperscript{105}. The profibrotic cytokine transforming growth factor beta (TGF\textbeta) can induce differentiation of these so-called myofibroblasts from epithelial cells, endothelial cells, or resident fibroblasts (in a process known as mesenchymal transition), and all have been implicated in fibrotic diseases. While the specific source of myofibroblasts in fibrotic diseases remains controversial and is likely varied, the prevailing view is that mesenchymal transition plays a prominent role in most if not all contexts\textsuperscript{106}. Typical alterations in cellular phenotype that attend TGF\textbeta-mediated epithelial- or endothelial-to-
mesenchymal transition (EpMT or EnMT, respectively; EMT collectively) include down-regulation of markers of a mature polarized cell state (e.g., E-cadherin) and induction of mesenchymal markers – e.g., expression of α-smooth muscle actin (αSMA), vimentin, fibrillar collagens, and matrix metalloproteases (MMPs)\textsuperscript{107}. Efforts to fate-map cells in fibrotic models of lung, liver, and heart fibrosis provide ample evidence for and against a role of EMT in the accumulation of myofibroblasts \textit{in vivo}\textsuperscript{108}. Nevertheless, TGFβ-SMAD signaling induces canonical EMT transcription factors\textsuperscript{107}, and genetic deletion of EMT transcription factors in lung alveolar cells or in hepatocytes blunts experimentally induced organ fibrosis\textsuperscript{109,110}. Given the clear role of TGFβ in fibrosis, we reasoned that a distal molecular event that is critical for TGFβ-induced mesenchymal transition would be an attractive therapeutic target for multiple etiologies of fibrosis.

We were intrigued by the association in the literature between multiple EMT-related disease processes and the increased expression or activity of calpains – a family of calcium-dependent non-lysosomal cysteine proteases that cleave diverse substrates to regulate cell activities including differentiation, adhesion, invasion, migration, synthetic repertoire, and survival\textsuperscript{111}. For example, calpain activity has been mechanistically linked to the invasive behavior of epithelial tumors\textsuperscript{112-115}, normal wound healing\textsuperscript{116}, cardiac fibrosis after tissue injury\textsuperscript{117-120}, and lung fibrosis in response to bleomycin\textsuperscript{121}. These observations led to our hypothesis that specific calpain cleavage products are required for mesenchymal transition and that inhibition of calpain activity may have therapeutic value in fibrotic disorders.

Of the 15 calpain isoforms expressed by humans, CAPN1 and CAPN2 are the best characterized, and are termed the conventional classical calpains\textsuperscript{122}. Active CAPN1
and 2 enzymes consist of a heterodimer formed with small regulatory subunit CAPNS1 or alternatively CAPNS2, and the activity of these conventional dimeric calpains is tightly regulated by the endogenous calpain inhibitor calpastatin (CAST). CAST is thought to specifically inhibit all dimeric calpains. CAST binds near the active site cleft of dimeric calpains in the presence of calcium and prevents engagement of substrates but is protected from hydrolysis by not binding the active site itself\textsuperscript{123-125}. CAPN1, 2, S1, and CAST are ubiquitously expressed; however, other calpain isoforms are expressed primarily in specific tissues or organs. For example, large subunit CAPN9 is reportedly chiefly expressed in the gastrointestinal tract\textsuperscript{126}, while CAPNS2 reportedly shows predominant expression in the skin and esophagus\textsuperscript{127}.

In this study, we used complementary methods to implicate calpain 9 (\textit{Capn9}) and calpain small subunit 2 (\textit{Capns2}), in TGFβ-induced myofibroblast differentiation \textit{in vitro} and in multiple experimentally induced models of fibrosis \textit{in vivo}. \textit{Capn9} showed highly restricted physiologic expression but can be potently induced in naïve cell types by TGFβ, suggesting the potential of high tolerance for therapeutic strategies aimed at antagonism.
IV. ABBREVIATED RESULTS FROM MANUSCRIPT

Capn9−/− mice are resistant to bleomycin induced lung fibrosis.

We next sought to define the therapeutic potential of calpain inhibition in animal models of fibrosis. We utilized mice lacking CAPN9 function that were generated by deleting Capn9 exon 3, which contains the catalytic cysteine in the peptidase domain. Deletion of the 119 nucleotides corresponding to exon 3 is predicted to create a frameshift and hence a premature termination codon (PTC) a short distance (6 codons) into exon 4 in mature mRNA. We confirmed these predictions using RT-PCR and sequencing of cDNA (data not shown). Contrary to prediction, the PTC did not initiate degradation of the mutant transcript via nonsense-mediated mRNA decay (data not shown). Loss of expression of CAPN9 protein was confirmed using a commercially available antibody from Abnova, but not using a reagent from Sigma, which detected a protein of the predicted mass for all dimeric calpains in both stomach and lung of wildtype and Capn9−/− animals indicating lack of specificity for CAPN9. In contrast, the Abnova antibody appears specific for CAPN9 but cross-reacts with a slightly smaller protein in the lung highlighting the limitations of the reagents for certain applications (data not shown).

Experimentally induced lung fibrosis is commonly achieved via exposure to the chemotherapeutic agent bleomycin. Systemic delivery of bleomycin by a subcutaneously implanted osmotic pump results in penetrant subpleural fibrosis with minimal inflammation, mirroring the histologic findings in scleroderma or idiopathic pulmonary fibrosis-related interstitial lung disease. Compared to wildtype mice, Capn9−/− mice
exposed to bleomycin are protected from fibrosis, as assessed histologically by Masson’s trichrome staining (Figure A.1a), blinded histological observation (Figure A.4b), and total lung collagen content (Figure A.1c). Bleomycin lung injury in wildtype mice corresponded with a dramatic mortality increase; however, Capn9<sup>−/−</sup> mice showed no mortality with systemic delivery of bleomycin (Figure A.1d) and exhibited improved lung function tests, as illustrated by a restored respiratory resistance and moderated total lung capacity (TLC) and lung compliance (Figure A.1e-g). Fibrosis in wildtype animals exposed to bleomycin was accompanied by a profound increase in the number of αSMA-positive myofibroblasts in the lung interstitium; whereas αSMA immunoreactivity was restricted to airways and blood vessels in vehicle-treated wildtype or bleomycin-treated Capn9<sup>−/−</sup> animals (Figure A.1h). Robust upregulation of Capn9 message accompanied myofibroblast accumulation in wildtype animals was detected by RT-PCR and in situ RNA hybridization (RNA-ISH) (Figure 1.4i).

Capn9<sup>−/−</sup> mice are protected from carbon tetrachloride (CCl<sub>4</sub>) induced liver fibrosis

Chronic CCl<sub>4</sub>-induced liver injury drives periportal inflammation and fibrosis in mice<sup>35</sup>. Wildtype mice treated with CCl<sub>4</sub> for 4 weeks developed bridging fibrosis while Capn9<sup>−/−</sup> mice were protected from fibrosis as shown by low-power bright-field picrosirius red (PSR) staining (Figure A.2a), birefringence of mature collagen fibrils under polarized light (Figure A.2b), blinded histologic scoring (Figure A.5c), and total liver collagen content (Figure A.2d). Robust upregulation of Capn9 message was detectable adjacent to fibrotic tracks in CCl<sub>4</sub>-treated wildtype animals by RNA-ISH (Figure A.2e). These histologic changes were accompanied by markers of liver damage such as elevated serum
levels of the liver enzymes alanine transaminase (ALT) and aspartate transaminase (AST). Wildtype and Capn9−/− mice receiving injections of CCl₄ had equal injury as indicated by elevated liver function tests (Figure A.2f); however, unlike wildtype controls, Capn9−/− animals developed minimal fibrosis.

Capn9−/− mice are protected from angiotensin-2 (Ang II) induced heart fibrosis

Ang II is capable of inducing a fibrotic synthetic repertoire in cardiac fibroblasts in vitro and in vivo¹²⁸-¹³¹. In wildtype animals, Ang II infusion results in cardiac arterial intima proliferation and perivascular collagen accumulation with collagen fibers extending into the cardiac interstitial tissue¹³². Here we showed that Capn9−/− mice with Ang II treatment are protected from fibrosis in intimal regions and developed minimal perivascular fibrosis, as shown by Masson’s trichrome stained slides (Figure A.6a), quantification of blue staining on trichrome (Figure A.3b), total heart collagen content (Figure A.6c), and αSMA immunoreactivity (Figure A.3d). Despite systolic blood pressure equivalent to wildtype animals receiving AngII (Figure A.3e), Capn9 deficient animals showed normalization of Ang II-induced left ventricular cardiac function as monitored by echocardiography (Figure A.3f).
Figure A.1 – Bleomycin induced pulmonary fibrosis
Figure A.2 – CCl₄ hepatic fibrosis
Figure A.3 – AngII induced cardiac fibrosis
V. DISCUSSION FROM MANUSCRIPT

Despite the significant contribution of fibrosis to global disease burden, few therapies target the accumulation or function of the cell type primarily responsible for pathologic ECM production, the myofibroblast\textsuperscript{105}. The lack of effective therapies reflects the tremendous difficulty in targeting fibrosis, a pathology that develops from the co-option of complex biological processes involved in tissue development and injury response, including EMT\textsuperscript{133}. Moreover, fibrosis is a common – and indolent – pathologic outcome of a group of highly heterogeneous disorders, compounding the challenge to achieve a comprehensive understanding of disease pathogenesis and vulnerabilities. Nevertheless, the recent observation that curative hepatitis treatment culminates in reversal of fibrosis establishes that a fibrotic extracellular matrix is a modifiable consequence of disease\textsuperscript{134-136}.

In this paper, we employed \textit{in vitro} assays of TGFβ-induced mesenchymal transition to investigate the role of calpains in the differentiation of myofibroblasts from various cell types. Our data show that TGFβ induces calpain activity, while inhibition of calpains preserves early TGFβ signaling events but prevents the induction of a mesenchymal phenotype. We were initially surprised to observe that TGFβ can cause myofibroblast differentiation utilizing calpain isoforms that are not normally expressed in the progenitor cell (i.e. NMuMG cells), but are rapidly and potently induced by TGFβ. In retrospect, this appears to be an elegant strategy to tightly regulate a process that is necessary and productive in development and wound healing, but can prove deleterious if left unchecked. Restriction of this myofibroblast-promoting activity to calpains that are recruited following a fibrogenic stimulus, such as TGFβ, may prevent initiation of this
pathway in response to physiological stimuli that normally result in activation of the constitutively expressed calpains. Indeed, prior work has demonstrated TGFβ-induced translational upregulation of other EMT-inducing factors in NMuMG cells including DAB2 and ILE1, albeit by a different mechanism than observed for CAPN9\textsuperscript{137}. Our data do not preclude the involvement or sufficiency of other calpains or calpain-independent mechanisms in mesenchymal transition in other cell types and contexts; indeed, this appears to be required given the observation of normal development and tissue homeostasis in $Capn9$ deficient animals\textsuperscript{126}. Moreover, a predicted loss-of-function mutation in humans is relatively common, found in homozygosity, and appears to be in Hardy-Weinberg equilibrium suggesting tolerance for CAPN9 deficiency. Although congenital loss of a gene is an imperfect model of therapeutic antagonism, these observations suggest the potential for a broad therapeutic window.

Our data build upon previous studies that show protection from bleomycin-induced lung fibrosis\textsuperscript{121,138} and Ang II-induced cardiac fibrosis\textsuperscript{117} upon broad spectrum calpain inhibition with calpeptin treatment or calpastatin overexpression, respectively. Importantly, both calpeptin and calpastatin are antagonists of the ubiquitous dimeric calpains, as well as CAPN9, and calpeptin inhibits other cysteine proteases including cathepsins\textsuperscript{139,140}. To our knowledge, this is the first description of a role for $Capn9$ or $Capns2$ in fibrosis and in TGFβ-induced myofibroblast differentiation. Indeed, existing literature describes CAPN9 function as restricted to the gastrointestinal tract, where co-expression of CAPN9 with CAPN8 in gastric pit cells forms a complex, termed G-calpain, that is involved in gastric mucosal injury response\textsuperscript{126,140}. The tissue restricted alternative small subunit, CAPNS2, appears to be largely redundant with CAPNS1 in
vitro. However, CAPNS2 binds to the large catalytic subunits with lower affinity, does not undergo activation-mediated autolysis, and conveys some substrate specificity differences in vivo.\textsuperscript{141,142}

The development of clinically useful calpain small molecule inhibitors has been hampered by the challenge of developing agents with specificity for calpains over other cysteine proteases, such as some cathepsins. Despite these challenges, the existence of CAST, a highly specific inhibitor of dimeric calpains, indicates that the design of calpain specific inhibitors is possible. Crystal structures of CAPN2-CAST complexes reveal that while CAST is an unstructured protein in solution, in the presence of calcium it binds to calpain and wraps over the active site.\textsuperscript{125,143} Calpain amino acid residues that interact with CAST near the active site are highly conserved in each of the CAST-inhibited isoforms; whereas CAPN3, which escapes CAST inhibition, is divergent, suggesting that calpain specificity may be achieved by inhibitors that display affinity for amino acids adjacent to the active site.\textsuperscript{144} Furthermore, crystal structures of the protease cores of CAPN1 and CAPN9 point to a number of differences that may be exploited for the development of CAPN9 antagonists.\textsuperscript{143}

Despite this progress, a number of limitations to this study should be considered. First, the observation that CAPN9-deficient NMuMG cells fail to support efficient TGFβ-induced EMT despite abundant expression of CAPN1, CAPN2, and CAPNS1 suggests the importance of a CAPN9 cleavage event that remains to be identified. While many calpain substrates have been recognized, the identification of calpain isoform-specific substrates has proved challenging. Knowledge regarding the precise function of CAPN9 that influences fibrosis in model systems will inform the clinical utility of strategies.
aimed at its specific antagonism. As stated above, the lack of small molecule antagonists that display high bioavailability, potency, and specificity for CAPN9 – or indeed all dimeric calpains – limits the immediate clinical application of these findings. Finally, our observations are based on prophylactic modulation of chemically induced models of fibrosis, which have a limited ability to predict efficacy in human presentations of disease or in disease reversal. Such issues will be best addressed using pharmacologic manipulations in both preclinical models and human clinical trials.
VI. ABBREVIATED METHODS FROM MANUSCRIPT

Small molecule inhibition of caplains

MDL-28170 (Enzo Lifesciences; Cat# BML-PI130), calpeptin (Tocris; Cat# 0448), 2-
APB (Sigma; Cat# D9754), CA-074-OMe (Sigma; Cat# C5857) were dissolved in
DMSO for 100 mM stock solutions. SB431542 (Sigma; Cat# S4317) was dissolved in
DMSO for a 10 mM stock solution.

RNA in situ hybridization

Formalin fixed, paraffin embedded histology sections were pretreated with the
RNAscope Target Retrieval kit (ACD; Cat# 322000) then hybridized with the Mm-
Capn9-O1 probe (ACD; Cat# 487221). Immunofluorescent detection of hybridized lung
samples was performed with the RNAscope Fluorescent Multiplex Detection kit (ACD;
Cat# 320851); chromogenic detection of hybridized liver samples was performed with
the RNAscope 2.5 HD Red kit (ACD; Cat# 322360) according to the manufacturer’s
instructions. Fluorescent maximum intensity projection images were captured on a Zeiss
LSM780 using a 40x objective. Chromogenic slides were captured on a Nikon 80i
microscope.

Histological analysis

Tissue was fixed with 10% formalin for 24 h prior to embedding in paraffin and
sectioning for histological staining. Masson’s trichrome staining was conducted by AML
Laboratories (St. Petersburg, FL) for the lung and liver histology and the Reference
Histology Laboratory at Johns Hopkins University (Baltimore, MD) for the heart
histology. Picrosirius red (PSR) staining was performed using standard techniques. Lung and liver histology images were captured by whole-slide scan performed by HistoTox Labs (Boulder, CO). Histological scoring of lung and liver histology was performed by a blinded histopathologist (LNH ToxPath Consulting, LLC).

**RNA analysis**

RNA was harvested from cells using the RNeasy Kit (Qiagen; Cat# 74106) with DNase digestion (Qiagen; Cat# 79254). RNA was harvested from dissected mouse organs following homogenization in an automatic bead homogenizer, FastPrep24 (MP Biomedicals; Cat# 116004500), in Trizol (Invitrogen; Cat# 15596026) according to the manufacturer’s instructions. cDNA was generated using a high capacity RNA to cDNA kit (Applied Biosystems; Cat# 4387406). Quantitative PCR was done using ABI TaqMan probes for mouse *Col1A1* (Mm00801555_g1), *Mmp2* (Mm00439498_m1), *Mmp9* (Mm00442991_m1), *Gapdh* (Mm99999915_g1), porcine *ACTA2* (Ss04245588_m1), *CDH1* (Ss03377287_u1), *VIM* (Ss04330801_gH), *MMP2* (Ss03394318_m1), *MMP9* (Ss03392092_g1) and *GAPDH* (Ss03375629_u1). Canine qPCR was performed using Sybr Green (ThermoFisher Scientific; Cat# 4309155) and canine *CAPN9* primers were: sense (5’-GCAGAGACCTTCGCAACTAA-3’) and antisense (5’-GCTGCATTTCCTGGATCAATGG-3’); canine *GAPDH* primers were: sense (5’-AACATCATCCTGCTTCCAC-3’) and antisense (5’-GACCACCTGCTCCTCAGTG-3’). Samples were run on a QuantStudioTM 7 Flex Real-Time PCR system (ThermoFisher Scientific). Ct values were corrected for loading and calculated using the 2^−ΔCt method. The average numerical value of the TGFβ1 only
treated sample was normalized to 1 for each gene unless indicated otherwise. cDNA gel electrophoresis of *Capn9* amplicons was performed by RNA isolation and total cDNA synthesis (as above). Amplification of mouse *Capn9* cDNA was performed with Flash Phusion (ThermoFisher Scientific; Cat# F548L) for 40 cycles with the following primers: exon 1 sense (5’-CTTTGTGTGGAAACGGCCAG-3’); exon 3 sense (5’-AGAAAGCACTGACCAGGGTG-3’); exon 5 sense (5’-TGGAAGACTTCACCTGGGGGT-3’); exon 9 sense (5’-TGCAACCTCACACCTGATGC-3’); exon 5 antisense (5’-TCCATGGCTTCAATGGCCTC-3’); exon 7 antisense (5’-CCCAAGGGTTACGGACTCTG-3’); exon 9 antisense (5’-AGCTTCCTTGGGATGGTGC-3’); exon 11 antisense (5’-TCCCACTCAGGTGTCGTC-3’); exon 13 antisense (5’-TGTGGAGTCGGCTTTTGAAG-3’). Mouse *Gadph* cDNA amplicons were generated with primers sense (5’-CAGGAGAGTGTTTCCTCGTCC-3’) and antisense (5’-TTCCCATTCTGCGCTTGAG-3’). Band intensities were measured with ImageJ (NIH) and normalized to *Gapdh* loading control.

**Bleomycin-induced pulmonary fibrosis**

Bleomycin was delivered by osmotic pump as previously described. Briefly, osmotic minipumps (Alzet; Cat# 1007D) containing either 100 µL saline vehicle or (100 U/kg) pharmaceutical grade bleomycin (Teva Generics; NDC# 00703-3155-01) designed to deliver their contents at 0.5 µL/h for 7 days were implanted under isoflurane anesthesia under the loose skin on the back of the mice, slightly posterior to the scapulae. Pumps
were removed on day 10 as recommended by the manufacturer and mice were sacrificed on day 35. Pulmonary function testing was performed on a randomly selected subset of animals using a flexi-VentTM ventilator (SCIREQ) as previously described \textsuperscript{145}. Tidal volume was set to 0.2 mL of 100% oxygen at a rate of 150 Hz with a positive end expiratory pressure (PEEP) of 3 cmH\textsubscript{2}O.

\textit{Angiotensin II-induced cardiac fibrosis}

Wildtype mice 6 weeks of age were anesthetized using isoflurane and an osmotic pump (Alzet; Cat# 2004) delivering Ang II (Sigma; Cat# A9525) at 1.2 \(\mu\)g/kg/min or saline was implanted beneath the mid-scapular loose skin. Pumps were left for 28 days and mice were sacrificed for tissue collection. Blood pressures were measured by tail cuff plethysmography in the week prior to sacrifice. Mice were habituated to the system, and, on following day, at least 3 blood pressure readings were obtained per mouse and averaged. Left ventricular heart function was determined as previously described\textsuperscript{146}. In brief, ventral hair was removed with Nair the day prior to echocardiograms. All echocardiograms were performed on unsedated mice using a Vevo 2100 (VisualSonics) equipped with a 40 MHz linear transducer. M-mode echocardiography was acquired from the parasternal short axis view of the left ventricle.

\textit{CCl\textsubscript{4}-induced hepatic fibrosis}

Male C57BL/6 and \textit{Capn9}\textsuperscript{-/-} mice, aged 10-12 weeks, received intraperitoneal injections twice a week with 1 mL/kg of CCl\textsubscript{4} diluted 1:7 in corn oil for 4 weeks. Mice were sacrificed 3 days after final injection and subject to whole body perfusion with PBS via
cardiac puncture. Serum was isolated from whole blood using blood collection tubes (BD Biosciences; Cat# 365967) as per manufacturer’s protocol. Liver enzyme serum concentrations were determined by the Johns Hopkins Department of Comparative & Molecular Pathology Phenotyping core facility.

**Immunofluorescence**

NMuMG cells were plated and grown on sterilized glass coverslips and submitted to EMT assays as described above. Coverslips were fixed with 4% paraformaldehyde (PFA) in PBS for 10 minutes at room temperature and blocked in immunofluorescence blocking buffer (5% donkey serum, 0.1% Triton X-100, in PBS) for 1 hour at room temperature. Slides were incubated with mouse anti-E cadherin (BD Biosciences; Cat# 610405) 1:500 in blocking buffer overnight at 4°C, then probed with donkey anti-mouse Alexa 488 (Molecular Probes; Cat A-21202) 1:500, rhodamine-phalloidin (Molecular Probes; Cat# R415) 1:500, and DAPI (Molecular Probes; Cat# D1306) 1:25000 for 1 hour at room temperature in the dark. After washing with PBS, they were mounted on slides (ThermoFisher Scientific; Cat# 12-550-15) with Prolong Gold mounting media (Molecular Probes; Cat# P10144). Images were obtained with a Zeiss LSM710. Formalin fixed paraffin embedded tissue was sectioned at 5 μm, deparaffinized in xylene, rehydrated, and blocked for 1 hour at room temperature (PBS, 0.1% Tween 20, 5% normal goat serum). Blocked samples were incubated with mouse anti-αSMA-Cy3 (Sigma; Cat# C6198) 1:1000 overnight at 4C. Slides were counter stained with DAPI and mounted. Images were acquired on a Zeiss LSM710 (lung) or a Leica DMi8 (heart).
Collagen quantification with hydroxyproline assay

Total collagen content was determined in freshly harvested lung tissue using the hydroxyproline assay kit (Sigma; Cat# MAK008) and normalized to total collagen content (Pierce; Cat# 23225). Liver and heart total collagen content was determined from formalin fixed paraffin embedded tissue samples (QuickZyme Biosciences; Cat# QZBtotcol1) and normalized to protein content (QuickZyme Biosciences; Cat# QZBtotprot1) according to the manufacturer’s protocols.
REFERENCES


Mohit Navandar1, Angela Garding1,*, Sanjeeb Kumar Sahu1, Abhijeet Pataskar1, Sandra Schick1, Vijay K. Tiwari1. ERK signalling modulates epigenome to drive epithelial to mesenchymal transition. *Oncotarget* (2017).


138


Sadoshima, J. & Izumo, S. Molecular characterization of angiotensin II--induced hypertrophy of cardiac myocytes and hyperplasia of cardiac fibroblasts. Critical role of the AT1 receptor subtype. *Circulation Research* **73**, 413-423, doi:10.1161/01.res.73.3.413 (1993).


Bachofner, J. A. et al. Direct antiviral agent treatment of chronic hepatitis C results in rapid regression of transient elastography and fibrosis markers fibrosis-4 score and


Curriculum Vitae
Tyler Jones Creamer
1019 W 38th Street
Baltimore, MD 21209
Phone: (410) 913-1069
Email: tcreamer@jhmi.edu

EDUCATION
Johns Hopkins University, School of Medicine, Baltimore, MD   2011-2018
PhD in Human Genetics and Molecular Biology, Institute of Genetic Medicine
Thesis: The Effect of miR-29 and TGFβ on Fibrosis
Mentors: Dr. Daniel Warren

University of East Anglia, Norwich, Norfolk, UK      2006-2009
Bachelor of Science, Hons – Biological Sciences

Whitman College, Walla Walla, WA      2004-2006

WORK EXPERIENCE
Johns Hopkins University, School of Medicine, Baltimore, MD   2012-2018
PhD Candidate, Laboratory of Dr. Daniel Warren, Department of Surgery
The lab of Dr. Warren is focused on understanding and preventing fibroproliferative diseases. For my PhD thesis, I focused on the influence of miR-29a as a potential therapeutic for liver fibrosis. Both in vivo and in vitro approaches allowed us to study liver fibrosis progression in tissue as well as individual cell populations contributions to the pathophysiology. We utilized high-throughput sequencing methods to understand the influence of injury and therapeutic on the transcriptomes of different liver cell populations.

Johns Hopkins University, School of Medicine, Baltimore, MD   Spring 2012
Rotation Student, Laboratory of Dr. Haig Kazazian
The Kazazian lab and Dr. Haig Kazazian is a world renown expert in retrotransposon biology. Under the supervision of Dr. Kazazian and lab members, I utilized an in vitro retrotransposon assay in U2OS cells to determine transposition efficiency of variants of the L1 retrotransposon family found in primates on the human genome. With the immense amount of DNA sequencing being done both in the lab and the clinic, increasing our understanding of the variation in L1 transposons can aide in determining active versus quiescent transposable elements and, therefore, putative causes of detrimental transposon insertions.

Johns Hopkins University, School of Medicine, Baltimore, MD   Fall 2011
Rotation Student, Laboratory of Dr. Sarah Wheelan
Under the supervision of Dr. Sarah Wheelan and members of her lab, I began using the programming languages R, Python, and Perl to quantify repetitive sequences in the human genome and their spatial distribution across the chromosomes. Parsing the sequences of the genome into words of a given size, I was able to map the distribution of k-mers. Understanding nucleotide sequences are important for understanding, not only coding effects, but also structure and folding of the DNA.

Johns Hopkins University, School of Medicine, Baltimore, MD   2009-2011
Research Specialist, Laboratory of Dr. Jeffry Corden, Department of Molecular Biology and Genetics
The Corden lab studied the effect of the carboxyl terminal domain (CTD) of RNA polymerase II and associated genes on transcription in baker’s yeast, Saccharomyces cerevisiae. During my tenure in the Corden lab, we developed a tandem affinity purification method of CTD-associated proteins and the nascent RNAs that were in close proximity. Using high-throughput sequencing methods, we showed that CTD-binding and associated proteins, Nrd1, Nab3, and Sen1, were responsible for regulating snRNA and snoRNA during stress responses.
Thesis Research conducted in Laboratory of Dr. Tamas Dalmay, University of East Anglia, School of Biological Sciences, Norwich, Norfolk, UK 2008-2009
Dr. Tamas Dalmay and the Dalmay Lab were interested in understanding miRNAs in the development of plants and animals. My work consisted of validating the lab’s work that miR-140 regulated a number of targets. Specifically, I was tasked with cloning 3’UTR constructs from genes discovered to be regulated through an over-expression and repression assay into Luciferase containing plasmids for the miRNA-target validation process.

John Innes Centre, Sainsbury Laboratory, Norwich, Norfolk, UK Summer 2008
Research Technician in Laboratory of Sir Dr. David Baulcombe
The Baulcombe lab are pioneers in the field of miRNA biology and have discovered multiple enzymes used by plants to regulate gene expression and translation via miRNAs. While working in the Baulcombe lab, I used ammonium persulphate precipitation and HPLC to try and devise a purification method for these enzymes, RNA polymerase IV and RNA polymerase V, from cauliflower.

Johns Hopkins University, School of Medicine, Baltimore, MD Summer 2006
Research Technician, Laboratory of Dr. Jeffry Corden, Department of Molecular Biology and Genetics
The Corden lab was interested in the carboxyl terminal domain (CTD) of RNA polymerase II and the evolutionary conservation across metazoa. During the summer I spent in the Corden lab, we developed methods to isolate and purify RNA pol II from Plasmodium falciparum, the protozoa responsible of Malaria. I learned in depth knowledge of the life cycle of P. falciparum and how to isolate protozoan proteins from infected human blood.

Johns Hopkins University, School of Medicine, Baltimore, MD 2003-2004
Research Practicum, Laboratory of Dr. Jeffry Corden, Department of Molecular Biology and Genetics
The Research Practicum class offered through my high school, Baltimore Polytechnic Institute, allowed students to work in an academic research setting for half of the school day for the duration of the last year. I conducted my work in the lab of Dr. Jeffry Corden where I investigated the CTD for RNA polymerase II in yeast. We used a truncated form of RNA pol II CTD and grew the yeast over 20 successive generations and measured the length of the CTD over that time. We showed, that due to growth advantages of a full length CTD, the yeast had successively larger CTD lengths as the passages continued. It was an incredibly insightful year that began my journey to becoming a research scientist.

MENTORSHIP/TEACHING EXPERIENCE
Johns Hopkins University, Center for Computational Genomics, Baltimore, MD 2015-2017
Teaching Assistant, Practical Genomics Workshop: From Biology to Biostatistics
Tasks included help with course organization, writing questions for small group discussions, and aiding registrants with computer and analysis questions during the 4-day workshop.

Johns Hopkins University, School of Medicine, Baltimore, MD Spring 2013
Graduate Teaching Assistant, Advanced Topics in Human Genetics
Tasks included help with course organization, writing and grading the mid-term exam, and directing the discussion during a weekly seminar portion of the course.

PRESENTATIONS


PUBLICATIONS


*Authors contributed equally