Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

M.Sc. Nadine Tüchler

born in: Zwettl, Austria

Oral examination: July 22nd, 2024

Dynamic multi-omics and mechanistic modeling of kidney fibrosis progression

Referees: Prof. Dr. Michael Knop

Dr. Justin Michael Crocker

Dedicated to my family.

Summary

Chronic kidney disease (CKD) affects more than 10% of the world's population and causes millions of deaths annually. Organ fibrosis is the main driver of the pathology underlying CKD, myofibroblasts are the cellular correlate of disease progression and TGF-β is considered as a master regulator of the underlying molecular processes. Despite progress in understanding the disease, there is no specific treatment available, and current diagnostic indicators neither facilitate early detection of CKD nor correlate with actual renal damage. Therefore, there is an urgent need for more time-sensitive biomarkers and a better understanding of diseases progression to provide the basis for new treatment options.

The aim of this thesis was to provide detailed mechanistic insights into CKD progression. To study fibrosis progression, a cellular model was used, consisting of human kidney derived PDGFRβ^{+} mesenchymal cells stimulated with TGF-β. Initial characterization of the model aligned with existing knowledge in the literature, indicating morphological and transcriptional changes along with increased ECM accumulation.

Following this, time-resolved multi-omics data, including transcriptomics, proteomics, phosphoproteomics and secretomics were acquired. The findings revealed previously proposed biomarkers and drug targets for kidney fibrosis, shown at all omics levels. These include SERPINE1, CCN2, CDH11, and integrins, alongside novel factors like LTBP2 and ADAM12, that have not been studied in the context of kidney fibrosis yet. Proteins like carboxypeptidase CPA4, with unknown implications in fibrosis, require further investigation to demonstrate their potential as therapeutic target or marker.

In addition, an integrative analysis approach was employed, using mechanistic modeling combined with footprint methods to estimate transcription factor (TF), kinase and phosphatases activities. TFs implicated in fibrosis were identified and experimentally validated using siRNA knock-down. This underscored the role of E2F1, FLI1 and NR4A1 in modulating fibrosis and ECM deposition. Nevertheless, further exploration is needed to elucidate the role of these TFs in collagen gene regulation and ECM accumulation during fibrosis.

Furthermore, mechanistic modeling generated new hypotheses regarding pathway dynamics over time in the fibrotic context. Overall, this integrative approach contributes to a deeper understanding of the molecular mechanisms driving kidney fibrosis,

offering potential biomarkers and therapeutic targets for clinical translation. Additionally, the dataset generated serves as a valuable resource for further research in the field of CKD.

Zusammenfassung

Chronische Nierenerkrankungen (CKD) betreffen mehr als 10% der Weltbevölkerung und verursachen jährlich Millionen von Todesfällen. Die Organfibrose ist die Hauptursache für die der CKD zugrunde liegende Pathologie, Myofibroblasten sind das zelluläre Korrelat der Krankheitsprogression und TGF-β gilt als Hauptregulator der zugrunde liegenden molekularen Prozesse. Trotz neuer Erkenntnisse im Bereich der Nierenfibrose gibt es keine spezifische Behandlung, und die derzeitigen diagnostischen Indikatoren erleichtern weder die Früherkennung von CKD noch korrelieren sie mit der tatsächlichen Nierenschädigung. Daher besteht ein großer Bedarf an zeitnahen Biomarkern und einem besseren Verständnis des Krankheitsverlaufs, um die Basis für neue Behandlungsmöglichkeiten für Patienten zu schaffen.

Ziel dieser Arbeit war es, detaillierte mechanistische Einblicke in das Fortschreiten von CKD zu gewinnen. Zur Untersuchung des Fortschreitens von Fibrose wurde ein zelluläres Modell verwendet, das aus PDGFRβ+ mesenchymalen Zellen der menschlichen Niere besteht, die mit TGF-β stimuliert wurden. Die anfängliche Charakterisierung des Modells stimmte mit dem vorhandenen Wissen in der Literatur überein und zeigte morphologische und transkriptionelle Veränderungen zusammen mit einer erhöhten ECM-Akkumulation.

Anschließend wurden zeitaufgelöste Multi-omics-Daten, einschließlich Transkriptomics, Proteomics, Phosphoproteomics und Sekretomics, erfasst. Die Ergebnisse enthüllten zuvor vorgeschlagene Biomarker und „Drug Targets“ für Nierenfibrose, die auf allen Omics-Ebenen nachgewiesen wurden. Dazu gehören SERPINE1, CCN2, CDH11 und Integrine sowie neue Faktoren wie LTBP2 und ADAM12, die im Zusammenhang mit Nierenfibrose noch nicht untersucht wurden. Proteine wie die Carboxypeptidase CPA4, deren Bedeutung für die Fibrose unbekannt ist, müssen weiter untersucht werden, um ihr Potenzial als therapeutisches Ziel oder als Marker zu demonstrieren.

Darüber hinaus wurde ein integrativer Datenanalyseansatz verwendet, bei dem mechanistische Modellierung mit Footprint-Methoden kombiniert wurde, um die Aktivitäten von Transkriptionsfaktoren (TF), Kinasen und Phosphatasen abzuschätzen. Dabei wurden TFs, die an der Fibrose beteiligt sind, identifiziert und anschließend durch siRNA-Knockdown Experimente validiert. Darunter wurde die

Rolle von E2F1, FLI1 und NR4A1 bei der Modulation von Fibrose und ECM-Ablagerungen unterstrichen. Dennoch sind weitere Experimente erforderlich, um die genaue Rolle dieser TFs bei der Regulierung von Kollagenen und der ECM-Akkumulation während der Fibrose zu klären.

Darüber hinaus hat die mechanistische Modellierung neue Hypothesen zur Dynamik der Signalwege im Laufe der Zeit im fibrotischen Kontext hervorgebracht. Insgesamt trägt dieser integrative Ansatz zu einem besseren Verständnis der molekularen Mechanismen bei, die der Nierenfibrose zugrunde liegen, und bietet potenzielle Biomarker und therapeutische Ziele für die klinische Umsetzung. Außerdem dient der erzeugte Datensatz als wertvolle Ressource für die weitere Forschung auf dem Gebiet der CKD.

Acknowledgements

First, I want to express my deepest gratitude to my supervisor and mentor, Dr. Rainer Pepperkok, for his unwavering support, invaluable guidance, constructive feedback, and encouragement throughout this journey. His challenging questions and dedication to pushing my boundaries have been instrumental in shaping my research endeavors, and I am truly grateful for the freedom he provided me to explore my own interests.

Special thanks are extended to Prof. Dr. Julio Saez-Rodriguez, my co-supervisor, for fostering a collaborative and supportive environment where individuals can thrive. His caring nature and concern for our well-being and interactions with each other have been greatly appreciated. I am also thankful to Prof. Dr. Christoph A. Merten for welcoming me to EMBL and providing an inspiring environment in which I learnt more than just microfluidics but got embarked by his entrepreneurial drive and motivation.

I am particularly grateful to Prof. Dr. Rafael Kramann and his lab for invaluable biological insights and the cells provided for my research, which have significantly contributed to the success of my project.

My sincere appreciation goes to my examiners Prof. Dr. Michael Knop, Dr. Justin Crocker, Dr. Mikhail Savitski, and Prof. Dr. Benedikt Brors, as well as other members of my thesis advisory committee, Dr. Sara Cuylen-Häring and Prof. Dr. Martina Muckenthaler, for their valuable input and guidance.

I am grateful for the support of all past and present members of the Pepperkok, Saez, and Merten groups, whose support, collaboration and friendship have enriched my PhD experience. These include Karim Karimli, Giriram Kumar Mohana Sundara Shanmugam, Muzamil M. Khan, Joanna Barbara Zukowska, Juan Jung, Sanjana Singh, Magdalena Zimon, George Galea, Karolina Kuodyte, Christopher Randolph Rhodes, Marco Raffaele Cosenza, Ann-Engelke Timm, Erika Schulz, Denes Turei, Martín Garrido Rodríguez-Córdoba, Aurélien Dugourd, Ricardo O. Raminez-Flores, Pau Badia I Mopel, Robin Falleger, Sophia Müller-Dott, Charlotte Boys, Rebecca Terrall Levinson, Olga Ivanova, Christina Schmidt, Ece Kartal, Remi Trimbour, Chiara Schiller, Francesco Ceccarelli, Martine Ballinger, Samantha Seah, Wenwei Ma, Anna Grab, Ramesh Utharala, Nicolas Peschke, Hongxing Hu, Alexis Autour, Jatin Panwar, Vida Vafaizadeh, and Miguel Ibarra.

A special mention goes to Mira L. Burtscher for her indispensable support. Thanks to her and her patience and time in guiding me through the analysis.

I would further like to acknowledge the invaluable assistance and expertise provided by the EMBL Core Facilities and Centers, including the Genome Core Facility, Proteomics

Core Facility, Advanced Light Microscopy Facility, Chemical Biology Core Facility, Protein Expression and Purification Core Facility, Centre for Statistical Data Analysis and Centre for BioImage Analysis. Members of these facilities has played a crucial role in advancing my research, not just by direct contribution or guidance, but also by providing workshops and resources that helped me advance my skillset. I extend my gratitude to Patricia Cabezas, career development specialist, for guiding me through the identification of my interests and her support on my job applications.

I cannot express how grateful I am for conducting my PhD at EMBL and the University of Heidelberg. I greatly enjoyed the collaborative and welcoming atmosphere and am thankful for the many connections I made. Special thanks go to Andreas Brunner, Wolfram Heinrich Seifert Davila, Felix Mikus, Laura Rustarazo-Calvo, and Marco Payr. Thank you for the memorable nights, coffee breaks, mental support, writing sessions, mock interviews, sports buddying, endless number of hugs and simply the amazing time. My time in Heidelberg was greatly enriched by the presence of many more friends who offered invaluable support and guidance, were travel buddies, and made my stay enjoyable through social activities and by engaging in deep discussions. I am immensely grateful for Simona Gioè, Constantin Ahlmann-Eltze, Florian Huber, Ioana Goganau, Alina Maria Ritz, Soham Basu, Laura Sabou, Camilo Ignacio Andino Cappagli, Tim Dullweber, Nedal-Djamil Darif, Anna Jungblut, Fabian Ruberti, Martin Larralde, Felix Langer, Simonne Anne Griffith-Jones, Mathias C. Eymery, Paulina Chojnacka, Donnacha Fitzgerald, Arif Ul Maula Khan, and Nima Vakili for their friendship. To everyone mentioned above and those whose names I could not include, thank you for being part of this incredible journey. Your impact on my academic and personal growth is immeasurable.

It is hard to find words to express my gratitude for the friends that have been supporting me throughout my PhD and beyond, including Adel Boudi, Lukas Zeinzinger, Emmanuel Nii Dauda Dixon und Kai-Chun Li.

Zu guter Letzt möchte ich meiner Familie danken, meinen Eltern Josef und Ingrid Tüchler, meinen Geschwistern Florian Tüchler und Sandra Bauer, sowie meinem Schwager Christoph Bauer, nicht zu vergessen Sarah und Jakob, für ihre fortwährende Unterstützung, Aufmunterung und die regelmäßigen Essenspakete. Schließlich möchte ich mich bei meinem Partner Lukas Koppensteiner und seiner Familie für ihre kontinuierliche Unterstützung bedanken.

Table of Contents

Contributions

- All experiments summarized in this thesis were designed, performed, analyzed and interpreted by me, if not otherwise stated.
- The Cell Profiler pipeline used in this thesis was designed with the input of Dr. Christian Tischer.
- Library preparation and sequencing of the transcriptomics data were performed by Ferris Jung, and the mapping of the transcripts to the human genome was done by Dr. Jonathan Landry, who also guided me through the initial analysis of the transcriptomics data.
- Mandy Rettel performed the sample preparation and MS runs for the secretomics data, while Dr. Jennifer J. Schwarz prepared and measured the samples obtained via cell lysis including proteomics and phosphoproteomics. Initial bioinformatic analysis was performed by Dr. Frank Stein.
- More in depth analysis of the multi-omics data was performed by me and Mira L. Burtscher, M.Sc., who was guiding me through this process.
- The mechanistic network was set up by Mira L. Burtscher, M.Sc.
- Statistical analysis of imaging and RT-qPCR data was supported by Dr. Sarah Kaspar.

Publications

- **Tuechler, N.**, Burtscher, M.L., Khan, M.M., Garrido-Rodriguez, M., Turei, D., Kramann, R., Savitski, M., Merten, C.A., Saez-Rodriguez, J., Pepperkok, R. *Dynamic multi-omics and mechanistic modeling approach provides deeper insight into kidney fibrosis progression.* (in preparation)
- Khan, M.M., Zukowska, J., Jung, J., Galea, G., **Tuechler, N.**, Halavatyi, A., Tischer, C., Haberkant, P., Stein, F., Jung, F., Landry, J., Khan, A.M., Oorschot, V., Becher, I., Neumann, B., Muley, T., Winter, H., Duerr, J., Mall, M.A., Savitski, M., Pepperkok, R. **2023**. *Dextromethorphan inhibits collagen transport in the endoplasmic reticulum eliciting an anti-fibrotic response in ex-vivo and in vitro models of pulmonary fibrosis*. **BioRxiv**, DOI
- Utharala, R., Grab, A., Vafaizadeh, V., Peschke, N., Ballinger, M., Turei, D., **Tuechler, N.**, Ma, W., Ivanova, O., Ortiz, A.G., Saez-rodriguez, R., Merten, A.M. **2022.** *A microfluidic Braille valve platform for on-demand production, combinatorial screening and sorting of chemically distinct droplets.* **Nature Protocols**, 17, 2920-2965. DOI
- Code availability: github.com/NadineTuechler/PhD_thesis_kidney_fibrosis_multiomics_PDGFRb_cells

List of Abbreviations

µl	Microliter
µM	Micromolar
α-SMA	Alpha-smooth muscle actin
ACTA2	Alpha-smooth muscle actin gene
ADAMTS	A Disintegrin And Metalloproteinase with Thrombospondin motifs
adj.P.Val	Adjusted p-value
ADPKD	Autosomal dominant polycystic kidney disease
AKT	Protein kinase B
ATAC-seq	Assay of Transposase Accessible Chromatin Sequencing
B2M	Beta 2-microglobulin
BGN	Biglycan
BHLHE40	Class E basic helix-loop-helix protein 40
BMP	Bone Morphogenic Protein
BSA	Bovine serum albumin
BTP	Bis-tris propane
ChIP seq	Chromatin immunoprecipitation sequencing
CKD	Chronic kidney disease
CO_2	Carbon dioxide
Col I	Collagen I
COL1A1	Collagen type 1a1
COL1A2	Collagen type 1a2
COL3	Collagen type 3
COSMOS	Causal Oriented Search of Multi-Omics Space
CT	Cycle threshold
DCN	Decorin
ddH_20	Deionized water
DEPC	Diethyl pyrocarbonate
Dex	Dextromethorphan
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol

e.g.	Exempli gratia, for example
E2F1	E2F Transcription Factor 1
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
eGFR	Estimated glomerular filtration rate
EMT	Epithelial-to-mesenchymal transition
Endo-MT	Endothelial-to-mesenchymal transition
EPO	Erythropoietin
ERK	Extracellular signal-regulated kinase
et al.	Et alia, and others
FBS	Fetal bovine serum
FC	Fold-change
FDR	False discovery rate
FGF	Fibroblast growth factor
FLI1	Friend leukaemia integration 1 transcription factor
FN1	Fibronectin 1
FSGS	Focal segmental glomerulosclerosis
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GFP	Green fluorescent protein
GLP-1	Glucagon-like peptide-1
GM130	Cis-Golgi matrix protein 1 (GOLGA2)
GOLGA2	Golgin subfamily A member 2
GPCR	G protein-coupled receptors
GRN	Gene regulatory network
GTPases	Guanosine-5'-triphosphate (GTP)-binding proteins
h	Hour
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HNF4G	Hepatocyte nuclear factor 4-gamma
i.e.	Id est, that is
ID2	Inhibitor Of DNA Binding 2
ID3	Inhibitor Of DNA Binding 3
IF	Immunofluorescence
IGFBP3	Insulin-like growth factor-binding protein 3
IgG	Immunoglobulin G
IL	Interleukin

IMAC	Immobilized metal affinity chromatography
IRI	Ischemia-reperfusion injury
JNK	C-Jun N-terminal kinases
k	Hhousand
KCl	Potassium Chloride
KD	Knock-down
kDa	Kilodalton
KH_2PO_4	Potassium dihydrogen phosphate
KIM-1	Kidney injury molecule-1
KPMP	Kidney Precision Medicine Project
KRIS	Kidney risk inflammatory signature
LAP	Latency associated peptide
LC	Liquid chromatography
lcRNA	Long non-coding RNA
Ly	LY2109761
MAPK	Mitogen Activated Protein Kinase
MC	Macromolecular crowdinh
MEK	MAP kinase kinases
min	Minute
miRNAs	Micro RNA
ml	Milliliter
mM	Millimolar
MMP	Matrix metalloproteinase
mRNA	Messenger RNA
MS	Mass spectrometry
MS/MS	Tandem MS
MsigDB	Molecular Signatures Database
NaCl	Sodium chloride
NAG	N-acetyl-β-D-glucosaminidase
$NaHPO_4$	Sodium phosphate
NEPTUNE	Nephrotic Syndrome Study Network
ng	Nanogram
NGAL	Neutrophil gelatinase-associated lipocalin
Nin	Nintedanib
NR4A1	Nuclear Receptor Subfamily 4 Group A Member 1

NRF2	Nuclear factor erythroid 2-related factor 2
NTRK	Neurotrophic tyrosine kinase receptor
PAI-1	Plasminogen activator inhibitor type 1
PBS	Phosphate-buffered saline
PC	Principal component
PCA	Principal component analysis
PDGF	Platelet-derived growth factor
PDGFRβ	PDGF receptor beta
PFA	Paraformaldehyde
PFD	Pirfenidone
PI3K	Phosphoinositide-3-kinase
POSTN	Periostin
PPI	Protein-protein interaction
p-SMAD2	Phosphorylated SMAD2
RAAS	Renin-angiotensin-aldosterone system
RAS	Renin-angiotensin system
RNA	Ribonucleic acid
RT-qPCR	Quantitative reverse transcription polymerase chain reaction
RUNX	Runt-related transcription factor
SDS	Sodium dodecyl sulfate
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	Second
SERPINE1	Serpin family E member 1, PAI-1
SGTL-2	Sodium-glucose cotransporter 2
siRNA	Small interfering RNA
SMAD	Mothers against decapentaplegic homolog
SPHK1	Sphingosine kinase 1
sqrt	Square root
SRGN	Serglycan
sTNFR	Soluble TNF receptor
TAZ	Transcriptional co-activator with PDZ-binding motif
TBS	Tris-buffered saline
TBS-T	TBS-Tween
TCA	Trichloroacetic acid
TEAD4	Transcriptional enhancer factor TEF-3

TF	Transcription factor
TGF-β	Transforming growth factor beta
TMT	Tandem mass tag
TNC	Tenascin-C
TNF-α	Tumor necrosis factor alpha
TRAF	TNF receptor (TNFR) associated factor
UMOD	Uromodulin
UPJ	Uretero-pelvic junction
UUO	Unilateral ureteral obstruction
v/v	Volume per volume
VEGF	Vascular endothelial growth factor
VIM	Vimentin
w/v	Weight/volume
WB	Western blot
YAP	Yes-associated protein

List of Figures

1. Introduction

1.1. Chronic Kidney Disease (CKD)

Chronic kidney disease (CKD) has a prevalence of approximately 10% worldwide (Edeling et al., 2016; Kim et al., 2022; Sziksz et al., 2015; Yamashita & Kramann, 2024). Especially in developed countries, CKD and other fibroproliferative diseases have arisen as a major public health problem (Nitta, 2015; Rockey et al., 2015). CKD is characterized by the excessive accumulation of extracellular matrix (ECM) and its subsequent replacement of functional tissue. The main producers of this ECM are fibroblasts and myofibroblasts which are activated during the course of the disease (Kim et al., 2022).

Most commonly diabetes mellitus and hypertension are the underlying causes of CKD (Mihai et al., 2018). The onset of CKD can also be caused by other factors including autoimmune disorders (like Lupus), long-term infections and chronic inflammation (Edeling et al., 2016; Kim et al., 2022). In general, persistent injury or stress of endothelial and epithelial cells can lead to the onset of the disease. The trans-differentiation of different cell types to myofibroblasts and their proliferation is essential in the progression of CKD as myofibroblasts are the major source of ECM (Kendall & Feghali-bostwick, 2014). Depending on the cell type from which the myofibroblasts originate, the phenotypes and signaling pathways may also differ (Prakash & Pinzani, 2017).

As the kidney consists of a variety of structures and cell types, each with distinct functions, there are categories of renal fibrosis based on the affected structures (J. Park et al., 2018; Rockey et al., 2015). Therefore, glomerulosclerosis describes fibrosis in the glomerulus, tubulointerstitial fibrosis is used for fibrosis of the proximal and distal tubules, while perivascular fibrosis describes the fibrosis around the vasculature (Kim et al., 2022).

While the filtering capacity of the kidneys (defined by estimated glomerular filtration rate, eGFR) decreases with age even in healthy individuals (Xu et al., 2020), the disease is defined by a reduced eGFR or albuminuria (abnormal leaking of plasma albumin into the filtrate) (Edeling et al., 2016; Kim et al., 2022). Based on these parameters, disease progression can be tracked and classified into five stages (Table

1), with kidney failure defined as a eGFR < 15 mL/min/1.73m^2 and an increased albuminuria (Himmelfarb & Ikizler, 2019; Kim et al., 2022; Rasmussen et al., 2019; Yan et al., 2021).

Table 1 CKD classification based on estimated glomerular filtration rate (eGFR). Adjusted after (Cañadas-Garre et al., 2018, 2019)

CKD Stage	eGFR [mL/min/1.73m^2]	Terms
1	≥ 90	Normal or high
2	60-89	Mildly decreased
3a	45-59	Mildly to moderately decreased
3b	30-44	moderately to severely decreased
4	15 - 20	Severely decreased
5	< 15	Kidney failure

As clinical symptoms only appear after irreversible damage to kidney function, CKD is rarely diagnosed in its early stages and often, by the time of diagnosis, the progression to kidney failure is inevitable (Mihai et al., 2018). This is aggravated by the fact that there is no specific treatment for the disease. To compensate for reduced kidney function, patients have to undergo lifelong dialysis or kidney transplantation, resulting in life-altering consequences (Edeling et al., 2016).

Even though several inhibitors of fibrosis, inflammation and oxidative stress-related pathways have shown to be effective in animal models, there is no clinically approved treatment for CKD in humans yet. There are several reasons for this. On the one hand, simple cellular and animal models rarely represent physiological complexity of fibrotic diseases in humans. To study kidney fibrosis in animal models, researchers often focus on a single pathway using genetically modified organisms or induce kidney injury aggressively by mechanical stress or chemical compounds (e.g. Ischemia-reperfusion injury (IRI) and Unilateral ureteral obstruction (UUO)). Moreover, results obtained from most animal models are irrelevant as studies are too short and fibrosis is often reversed in these models when the stimulus is gone, those they do not sufficiently replicate the normal disease progression, nor its defining characteristic of being chronic. On the other hand, promising anti-fibrotic therapies are not tested or have failed due to long, time-consuming and expensive clinical trials that are lacking adequate tools to assess fibrosis progression and regression and therefore resulting in equivocal end-points (Prakash & Pinzani, 2017).

Patients only present with clinical symptoms after irreversible damage to kidney function, by which point disease reversal is almost impossible and life-altering consequences are inevitable. Therefore, CKD is rarely diagnosed in early stages and

ultimately patients will suffer from kidney failure (Mihai et al., 2018; Walraven & Hinz, 2018). However, it has been shown that early diagnosis could have a major impact in slowing down disease progression. The gold standard for diagnosing CKD is based on eGFR and kidney biopsies. Each of them has several drawbacks. eGFR, derived by measuring serum creatine, is influenced by sex, age, ethnicity, muscle mass as well as diet and is only informative after 50% of the kidney function is already lost (Cañadas-Garre et al., 2018; Devarajan, 2010; Rodríguez-Ortiz et al., 2018). Biopsies are invasive, painful and can be biased by sampling variability (Devarajan, 2010).
Other biomarkers which could give information about kidney injury, can be influenced by co-existing chronic diseases and are understudied in patients with CKD. Studies that have investigated biomarkers for acute kidney injury (AKI, characterized by kidney injury up to three months, thereafter CKD), defined AKI based on elevated serum creatine levels, which is largely dependent on the above mentioned variables (Alpern et al., 2013; Himmelfarb & Ikizler, 2019). The same is true for CKD biomarkers. Furthermore, some markers are not tissue specific or occur during the physiological wound healing response (Walraven & Hinz, 2018).

1.1.1. Cellular Mechanisms Underlying CKD

In the following section, cellular components of the kidney and their implications in renal fibrosis will be summarized. Additionally, the complex communication and interplay between different cell types and the changing ECM will be described (Adler et al., 2019; Yamashita & Kramann, 2024).

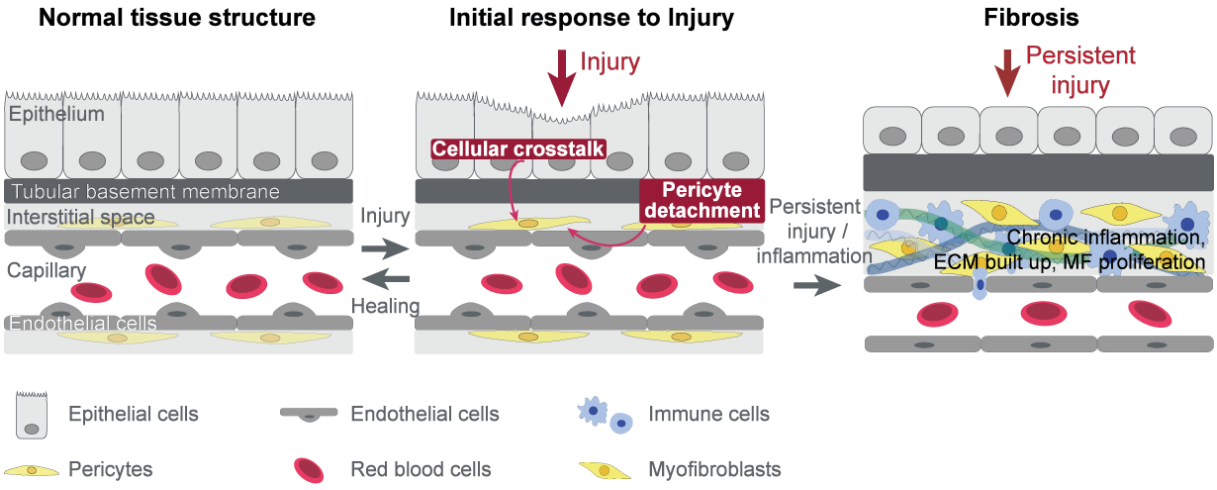

Figure 1 Fibrotic processes in the kidney. In response to tissue injury, wound healing is initiated, facilitated by cellular crosstalk and supported by tissue resident mesenchymal cells. Continuous injury fosters a pro-inflammatory environment, causing the transition of epithelial cells to a mesenchymal-like phenotype driven by EMT and thereby loss if the epithelial function. This change in metabolism and

phenotype leads to activation of latent TGF-β and thereby activation of mesenchymal cells (like pericytes and fibroblasts) that differentiate into myofibroblasts. Activated myofibroblasts in turn, due to the persistent stimulus, produce excess amounts of ECM components. Endo-MT leads to loss of endothelial integrity, in parts driven by detachment of pericytes. In general, it is a viscous cycle driven by cell-cell interactions and change in microenvironment that ultimately result in loss of physiological kidney architecture and function. Modified after (Friedman et al., 2013) and (Mullins et al., 2016).

Tissue Resident Pericytes, Fibroblasts and Myofibroblasts

Resident mesenchymal cells expressing PDGFRβ, including pericytes and fibroblasts, are essential in maintaining the kindey's basement membrane and its overal architecture, ensuring tissue homeostasis (Kim et al., 2022; Kramann & Humphreys, 2014). Pericytes and fibroblasts also regulate oxygen transport and produce renin and erythropoietin (EPO), renal hormons that become deregulated in CKD patients leading to complications such as anemia and hypertension (Kim et al., 2022; Zeisberg & Kalluri, 2015). Stimulation with growth factors, either secreted by other cells or via autocrine signaling, results in differentiation of pericytes and firboblasts, which are the main progenitors of myofibroblasts (proven by genetic lineage tracing and omics approaches) (Kramann & Humphreys, 2014; Kuppe et al., 2021; Yamashita & Kramann, 2024).

Many studies focus on activated myofibroblasts (showing an increased expression of e.g. α-smooth muscle actin (α-SMA), collagens type I, III, V, periostin (POSTN), tenascin-C (TNC), fibronectin (FN)), as their abundance can be indicative for the progression of the functional decline (Schnieder et al., 2016; Zeisberg & Kalluri, 2015). Since myofibroblast numbers seem to correlate with epithelial injury and epithelial cell loss, it has long been argued that epithelial cells transdifferentiate into myofibroblasts (Kendall & Feghali-bostwick, 2014; Kim et al., 2022). However, there are other studies that challenge this view (Kramann et al., 2018; Kuppe et al., 2021). Myofibroblasts are the main effector cells that are secreting ECM components including collagens, fibronectin and glycosaminoglycans (Kim et al., 2022). The initial role of this process is to support tissue repair, after which apoptosis of myofibroblasts is induced (**Figure 1**) (Kim et al., 2022). However, in the context of fibrosis, this leads to ECM accumulation as there is persistent injury, inflammation, and an inhibition of myofibroblast apoptosis. Pericyte-to-myofibroblast differentiation, and the therefore decreasing number of vasculature-associated pericytes drives one of the hallmarks of acute kidney injury to CKD progression which is loss of integrity of the mircovasculature (capilary rarefaction). Therefore, the inhibition of pericyte-to-myofibroblast differntiation provides

a potential target for treatment of renal fibrosis (F. C. Chang et al., 2012; Friedman et al., 2013; Kramann & Humphreys, 2014; Mullins et al., 2016).

Epithelial and Endothelial Cells

The kidney contains specialized epithelial and endothelial cells. In general, endothelial cells line the microvasculature in the kidney, forming the filtration barrier in the glomerulus. There, they merge from afferent to efferent arterioles (Kim et al., 2022; Yamashita & Kramann, 2024). Epithelial cells are arranged in a single layer in the Bowman's Capsule, to the Loop of Henle and the collecting duct controlling reabsorption (Eaton, 2012; Kim et al., 2022). Upon injury, both endothelial and epithelial cells contribute to to the wound healing response and fibrosis by secreting factors stimulating mesenchymal cells and recruiting immune cells. Additionally, they acquire a mesenchymal-like phenotype via Endo-MT and EMT, respectively (**Figure 1**) (Friedman et al., 2013).

Tissue damage can be caused by a variety of factures such as high levels of glucose or free fatty acids like in diabetes, toxins, oxidative stress or physical stretch (Edeling et al., 2016). As a result of ongoing injury, epithelial cells undergo apoptosis, necrosis, autophagy, mitotic catastrophe (e.g. G2/M cell cycle arrest) or oncosis which leads to the recruitment of immune cells (Gewin et al., 2017; Ke et al., 2017). The immune response is initiated to clear the dead cells and help in wound healing. Another result of epithelial cell apoptosis is peritubular capillary rarefaction, increased hypoxia and oxidative stress (Gewin et al., 2017). This can lead to the release of hypoxia-inducible factor-1alpha, thereby initiating dedifferentiation of proximal tubule cells and ECM build up resulting in a vicious cycle as a result of which endothelial and epithelial cell loss lead to functional decline (Edeling et al., 2016).

Immune Cells

Different immune cells and their interactions with renal resident cells are known to influence kidney fibrosis. Neutrophils are thought to be among the first cells that activate other innate immune cells like macrophages (Sziksz et al., 2015; Van Linthout et al., 2014). Further immune cells involved in fibrosis e.g. by TGF-β secretion are eosinophils and macrophages (Van Linthout et al., 2014). One of the proposed targets for therapy of renal fibrosis is to inhibit the differentiation of monocytes into macrophages (Chana et al., 2003; Yamashita & Kramann, 2024). Studies have shown

that there is an increase of pro-inflammatory macrophages (M1) that drives fibrosis by secreting ROS, inflammatory cytokines and pro-fibrotic MMPs and thereby regulating differentiation of resident cells in the kidney (e.g. via TGF-β, PDGF, galnectin3, IL-1, IL-4, IL-10, IL-13, TNF-α) (Kim et al., 2022; Van Linthout et al., 2014; Yamashita & Kramann, 2024). In turn, renal cells like myofibroblasts also secrete factors that control macrophage activity such as macrophage colony stimulating factor (M-CSF). Moreover, it has been shown that depletion of macrophages leads to reduced or enhanced fibrosis. This goes to show that CKD is more complex than we currently understand and further research is needed (Adler et al., 2019).

1.1.2. Molecular Mechanisms Underlining CKD

TGF-β as Master Regulator of CKD

The TGF-β superfamily does not only contain three isoforms of TGF-β, but also seven bone morphogenic protein (BMP) isoforms and activins. Serine and threonine kinase receptors that bind TGF-β and BMP are the starting points for signal transduction (Figure 2). The canonical signaling leads to phosphorylation of SMADs which are forming protein complexes that are translocated into the nucleus for regulation of transcription. While TGF-β leads, via SMAD2/3/4 activation, to transcription of fibrotic target genes (such as α-SMA, type I collagen, and tissue inhibitor of matrix metalloproteinases), BMP binding induces phosphorylation of SMAD1/5/8 and, e.g. induced by BMP7 counteracting the pro-fibrotic effects of TGF-β (Kim et al., 2022; C. H. Park & Yoo, 2022). During normal tissue homeostasis, the effects of TGF-β and BMP balance each other out. Cells developed additional control mechanisms such as the transcription of SMAD7, induced by TGF-β, that counteracts SMAD3 activity. Unfortunately, these control mechanisms are lost in fibrosis further driving disease progression (Kim et al., 2022).

TGF-β originates from cells that either secrete it or release it from the ECM. For instance, activated myofibroblasts secrete TGF-β in an autocrine fashion, thereby forming a positive feedback loop (C. H. Park & Yoo, 2022).

A non-canonical response initiated by TGF-β is the MAP kinase pathway by activation of Ras and signaling via MEK1/2 and subsequently ERK1/2. This leads to SMAD2 phosphorylation and expression of target genes as described above (C. H. Park & Yoo, 2022).

Ubiquitinoylation of TRAF4/6, resulting from interaction with the TGF-β receptor, also leads to activation of SMAD2/3 via the p38/JNK pathway. In addition, c-Jun and AP-1, transcription factors implicated in fibrosis, are phosphorylated by p38/JNK (C. H. Park & Yoo, 2022).

Inhibition of fibroblast apoptosis and enhanced proliferation are the result of activation of PI3K/AKT signaling that can also be targeted by TGF-β signaling.

A crucial factor in cell motility is the remodeling of the cytoskeleton which is regulated by RhoA GTPases. This pathway can also be resulting from TGF-β stimulation, leading to myofibroblast differentiation (C. H. Park & Yoo, 2022).

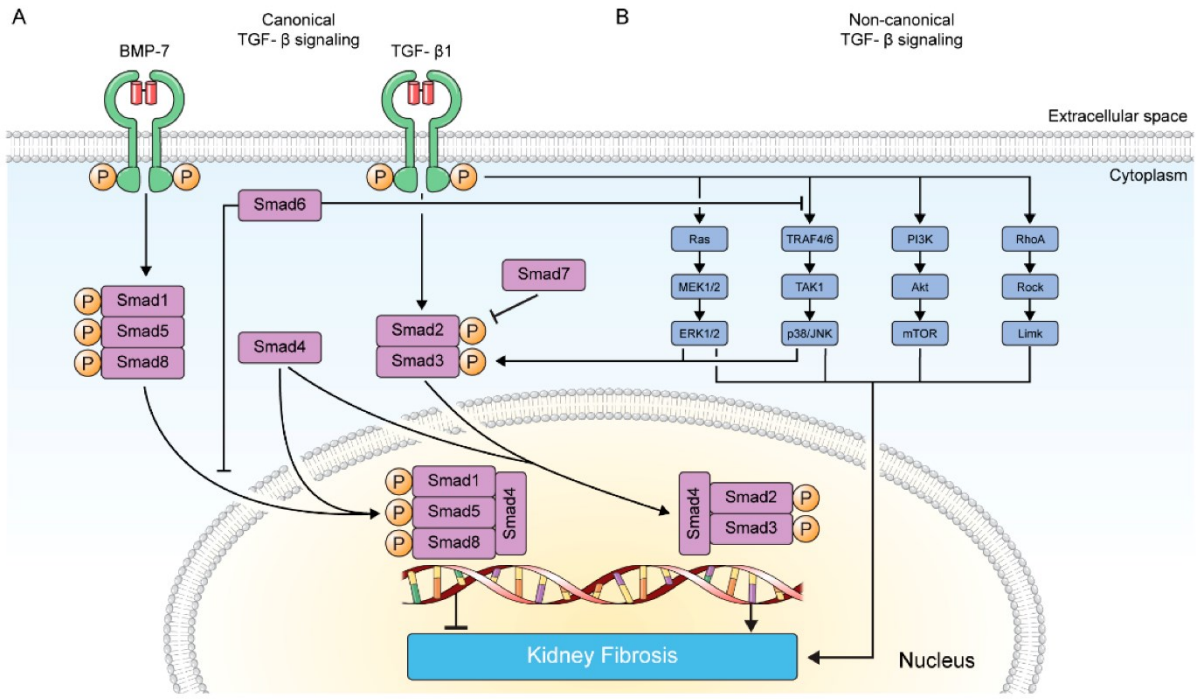

Figure 2 TGF-β and BMP induced signaling mechanisms. Signal transduction is initiated by TGF-β/BMP binding and drives different pathways. **A)** The canonical pathway is based on SMAD proteins controlling transcription of pro-fibrotic target genes. **B)** Signaling via MAPK, p38/JNK, PI3K/Akt and RhoA GTPase are considered non-canonical response to TGF-β stimulation. Taken from (C. H. Park & Yoo, 2022).

As a central regulator of fibrosis, it seems apparent to target TGF-β in order to treat fibrotic diseases, and hence development in this direction is ongoing. However, drug treatments based on this approach were not successful so far due to undesired side effects (Henderson et al., 2020; Meng et al., 2016; C. H. Park & Yoo, 2022). This might be contributed to the roles of TGF-β in other biological contexts such as inflammation and cancer (Zhao et al., 2020). For example, treatment with LY2382770 or fresolimumab, antibodies blocking TGF-β did not lead to an improved kidney function

in the case of diabetic nephropathy or focal segmental glomerulosclerosis (FSGS) (Kim et al., 2022; Yamashita & Kramann, 2024).

So far, only two drugs are approved for the treatment of pulmonary fibrosis: Pirfenidone (PFD) and nintedanib (Nin). Therefore, there was a lot of interest to test them for renal fibrosis, however with mixed outcomes. The specific function of PFD is not entirely understood but is associated with inhibition of production as well as activity of TGF-β (Cho & Kopp, 2010). PFD is currently also tested in clinical trials for the treatment of kidney fibrosis where it has been shown to slow down kidney function decline in patients with diabetic kidney disease or FSGS (Henderson et al., 2020; Kim et al., 2022; Yamashita & Kramann, 2024). Affected pathways by Nin are VEGF, FGF, PDGF as well as lymphocyte-specific protein tyrosine kinase and Src nonreceptor kinase signaling. The inhibited pathways are of therapeutic interest as all of them are involved in fibrosis (Bigaeva, Stribos, et al., 2020). So far, this compound has not been investigated in clinical trials for kidney fibrosis (Yamashita & Kramann, 2024). Even though there are clinical trials ongoing, no specific treatment for kidney fibrosis has been approved to this date (Yamashita & Kramann, 2024).

Integrin Mediated Signaling

In humans, there are 24 known integrin family members forming α–β heterodimers of transmembrane protein complexes (Kim et al., 2022; Nolte et al., 2020). Integrin signaling plays a central role in processes including cell adhesion, migration, proliferation, differentiation as well as apoptosis (Henderson et al., 2020; Kim et al., 2022). As TGF-β is deposited in the ECM in a latent form via binding to latency associated peptide (LAP), the conformational change excreted by mechanical force applied by integrin-expressing cells, triggers the further release of TGF-β. Integrins involved in this process include αvβ1/β3/β5 (myofibroblasts/fibroblasts), αvβ6 (epithelial cells), and αvβ8 (by recruiting MMPs) (Henderson et al., 2020; Kim et al., 2022). This integrin-mediated TGF-β accumulation further drives the vicious cycle of fibrosis and therefore provides an additional therapeutic target (Henderson et al., 2020).

Furthermore, integrins are implicated in assembly of extracellular fibrous structures (Kim et al., 2022). The expression of integrin subunits is cell-type dependent, hence the cellular responses can differ as well as the processes driven by integrin signaling. For instance, activation of myofibroblasts is reduced by decreased cell-ECM signaling

as mediated by integrins. Studies have also shown that inhibition of integrins associated with binding of basement-membrane components aggravated renal fibrosis, while deletion of integrins involved in binding of interstitial matrix components improved kidney fibrosis (Henderson et al., 2020; Kim et al., 2022). While there are ongoing clinical trials to test inhibitors of different integrins, they are mostly focused on treating pulmonary fibrosis (Henderson et al., 2020). Targeting integrin-mediated signaling provides valuable treatment options for fibrosis. However, further investigation, including improved and more targeted therapeutics delivery are needed.

Mechanosensitive Mechanisms

In response to changes of the microenvironment, there are cytoskeletal adaptation that are implicated in yes-associated protein 1 (YAP)/transcriptional coactivator with PDZ-binding motif (TAZ) signaling in fibroblasts. In the nucleus, YAP/TAZ bind to e.g. Runt-related transcription factor (RUNX) and TEA domain family member (TEAD) and control expression of fibrotic target genes associated with myofibroblast differentiation and TGF-β signaling. Due to the importance of these co-activators, therapeutically targeting them could provide promising results (Kim et al., 2022; Piersma et al., 2015).

Reactivation of Developmental Mechanisms

Key developmental pathways, such as Wnt, Notch and Hedgehog, are activated initially as a wound healing response but are also implicated in fibrosis (Edeling et al., 2016). For instance, the Wnt/β-catenin pathway is involved in EMT, differentiation of myofibroblasts and immune cell invasion by upregulation of MMP7. This metalloproteinase promotes a positive feedforward loop by further upregulating β-catenin and MMP7 levels (Kim et al., 2022).

Angiotensin II / RAAS Pathway

Especially important for the kidney is the renin-angiotensin-aldosterone system (RAAS) as it is central in controlling the water as well as sodium concentrations and the expression of Vasopressin. Thereby, the permeability and constriction of vasculature in the kidney are controlled by this mechanism. The deregulation of RAAS was shown to be implicated in myofibroblast differentiation and immune cell invasion. A common treatment for renal diseases (to inhibit RAAS signaling) is the use of

angiotensin II receptor blockers as increased angiotensin II promotes profibrotic pathways including TGF-β, NFκB and Wnt. NFκB/TNF-α signaling is implicated in immune response and inflammation (Kim et al., 2022).

1.1.3. CKD Diagnosis

The "Golden" Standard

CKD diagnosis and prognosis in clinics mainly rely on eGFR and albuminuria, as readouts for renal (filtration) function and damage, and biopsies. Since there is no targeted therapy for CKD available, fibrosis cannot be reversed and lost kidney tissue cannot be fully regenerated. It is therefore crucial to detect CKD early on in order to prevent or slow down disease progression. However, eGFR, albuminuria and biopsy have limitations as diagnostic tools which I will summarize in the following paragraphs (R. Huang et al., 2023; Yan et al., 2021). Additionally, I will list a few of the many proposed biomarkers, especially suggeted for early disease diagnosis.

As a result of hyperfiltration, the ability of the kidney to take over the function of lost nephrons (filtering units), eGFR gains predictive power only when 50% of renal function is lost (Cañadas-Garre et al., 2018; Devarajan, 2010; Rodríguez-Ortiz et al., 2018). This compensatory hypertrophy takes over the clearance of plasma factors, and maintains normal levels of plasma creatinine and urea (used to calculate eGFR). As already mentioned in the initial part of the introduction, these factors are heavily biased by several variables such as age, gender, ethnicity, muscle mass and other medications (Devarajan, 2010).
Albuminuria, proteinuria and microalbuminuria are all terms that are used to describe the abnormal presence of albumin in the urine which can be a measure for kidney damage. Similar to eGFR, albuminuria is present in the urine once significant damage has already occurred. Moreover, the predictive power of this marker has been questioned (Devarajan, 2010; Rodríguez-Ortiz et al., 2018).

A more invasive approach for diagnosis and determination of CKD stage involves biopsy. Besides pain and discomfort experienced by the patients, biopsies are also susceptible to sampling variability and not sensitive to early disease progression (Prakash & Pinzani, 2017; Yan et al., 2021).

Potential New Biomarkers

Newer markers for renal function include beta trace protein (BTP), and β2-microglobulin (B2M). BTP catalyses prostaglandin H2 to D2 conversion and is also used as a marker for cerebrospinal fluid leakage. B2M, as part of the major histocompatibility class I molecule, is found in nearly all nucleated cells. Since both are proteins with low molecular weight, they pass the endothelial barrier in the glomeruli but are reabsorbed in the proximal tubules. While they can also be measured in the serum, increased levels in the urine are indicative for the tubular damage. Compared to serum creatinine (used to calculate eGFR), both proteins are less affected by factors like age, gender and ethnicity and therefore could be used in combination with eGFR or used to develop newer eGFR equations (Lousa et al., 2021).

A kidney tubule specific protein, uromodulin (UMOD), is one of the most abundant glycoproteins in the urine under physiological conditions. Genetic studies have shown its association with hereditary autosomal-dominant tubulointerstitial kidney disease. It has further been highlighted as a biomarker for tubular injury, where its levels decreased in the urine and serum of patients (Lousa et al., 2021). Interestingly, it could be used as an early biomarkers to differentiate healthy form early stage CKD patients (Lousa et al., 2021; Yan et al., 2021). Other examples of markers that could be used in combination are neutrophil gelatinase-associated lipocalin (NGAL) and kidney injury molecule-1 (KIM-1) as well as N-acetyl-β-D-glucosaminidase (NAG). However, the use of NGAL, KIM-1 and NAG as biomarkers showed conflicting results that could be attributed to different causes of CKD, requiring more research to demonstrate their importance and suitability (Lousa et al., 2021; Yan et al., 2021).

Information about the inflammatory state of renal tissue can gained with markers such as interleukin-6 (IL-6), tumor necrosis factor (TNF-α) and soluble TNF receptors (sTNFR1 and 2). Their increase has been linked to renal function decline and poor outcomes (Lousa et al., 2021; Mariani et al., 2018, 2023; Yan et al., 2021).

Matrix metalloproteinases (MMPs) are enzymes that reshape the ECM, but also cleave other molecules like adhesion molecules and growth factors. A well-known MMP in regards to fibrosis is MMP-9, which is expressed in the kidney in low concentrations but increases upon injury. As a result, also urine levels of MMP-9 raise. Similarly,

enhanced MMP-2 levels are present in the urine of patients with renal fibrosis. Even though both show potential as biomarkers for CKD progression, further studies are needed to evaluate their role in the disease (Yan et al., 2021).

ECM remodeling, collagen turnover and increased deposition are a hallmark of fibrosis. Therefore, it is interesting to monitor their changes with disease progression or in response to treatment. Collagen maturation and crosslinking, for deposition, is achieved by peptidase cleavage. So in the case of Col I formation, a PRO-C1 peptide is cleaved off. When Col I fibers are degraded by MMP-2, 9 or 13, the C1M peptide is released. Such peptides get released from formation/degradation of other collagens as well and can be measured in the blood or urine. In a study led by Rasmussen et al., they investigated the change of blood and urine fragments indicative for collagen formation and degradation. Their findings could be used to identify subgroups of patients with differences in collagen turnover. Nevertheless, further studies, especially with a higher number of patients, are necessary to confirm their findings and help understand the information such measurements could provide (Rasmussen et al., 2019). A limitation of using collagens or collagen turnover as a readout of disease progression, is that collagens are not specific to the kidney and results could be affected by other comorbidities.
Expression of periostin (POSTN), another ECM protein and mesenchymal marker, is increased in fibrosis. The positive feed-back loop of POSTN and TGF-β further drives a pro-fibrotic environment. In patients with type 2 diabetes, it is proposed as a promising early biomarker to predict diabetic nephropathy, as increased levels of POSTN are secreted via the urine (Yan et al., 2021).

I have not mentioned microRNAs (miRNAs) in the context of fibrosis yet, as there are contradicting data especially when it comes to the role of miRNAs as biomarkers. miRNAs are generally involved in a variety of cellular processes like proliferation and differentiation. This is achieved by binding to complementary regions of mRNA targets and thereby suppressing translation. Some miRNAs, like mir-16 and mir-451-5p, could be involved in downregulation of IL-6 and MMP-9 levels and therefore play a role in fibrosis. However, to show the potential of miRNAs as biomarkers, more evidence is needed (Yan et al., 2021).

Omics studies in the recent years have led to a panel of peptides that can be detected in the urine. The panel is known under CKD273, as it contains 273 peptides, among which are many collagens, that have been validated in several studies to allow diagnosis of CKD. This provides a valuable tool for diagnosis and even has been shown to allow earlier detection as well as patient stratification of CKD e.g. in the case of diabetic nephropathy (Lousa et al., 2021).
Additionally, there are non-invasive imaging methods available, which however, have their own limitations. Examples are further summarized e.g. in Klinkhammer and Boor (2023) (Klinkhammer & Boor, 2023). To conclude, there are several biomarkers, kits and peptide panels available that show potential in being applied in the clinics in the next decades, but more research is needed to get them to this point. Most likely not a single biomarker but rather a combination of biomarkers will be used to address currently unmet needs.

1.1.4. CKD Treatments

As already mentioned in earlier sections, currently, there are no effective drug treatments available for CKD. However, patients are commonly given the following medications that have been shown to delay the progression of the illness at different rates and with varying efficiency (Himmelfarb & Ikizler, 2019; R. Huang et al., 2023). Well-known examples are glucagon-like peptide-1 (GLP-1) receptor agonists and sodium-glucose cotransporter 2 (SGLT-2) inhibitors. GLP-1 receptor agonists have been shown to reduce risk of albuminuria as well as CKD progression with their antioxidant, anti-inflammatory and antifibrotic actions. SGTL-2 inhibitors show glucose-lowering effects and, in the case of kidney disease, may work to reduce hyperfiltration and hypertension (Himmelfarb & Ikizler, 2019). Other treatments, including Renin-angiotensin system (RAS) blockers, endothelin-1 blockers, vasopressin receptor 2 antagonists and non-steroidal anti-mineralocorticoids help in controlling blood pressure, suppress inflammation and oxidative stress (Himmelfarb & Ikizler, 2019; R. Huang et al., 2023).

In previous sections, I highlighted potential targeted approaches to provide treatment for kidney disease patients, like inhibition of monocyte-to-macrophage transition, differentiation of kidney resident mesenchymal cells to myofibroblasts, and TGF-β signaling, as central regulator of fibrosis. Other approaches include targeting of CTGF

using antibodies, activation of NRF2, inhibition of micro RNAs or galectin-3 antagonists (R. Huang et al., 2023), however with varying outcomes further emphasizing the complexity of CKD treatment.

CKD is driven by a sophisticated network of pathways, including intricate cell-cell and cell-matrix interactions. Additionally, it often coexists with different underlying conditions that create a profibrotic microenvironment. The problem of translating research outcomes of in vitro and in vivo studies into clinical practice and bringing treatment options to patients has been persistent so far (Yamashita & Kramann, 2024). This might only be overcome by improving our understanding of underlying disease mechanisms, dissecting the role of various cell types and the changing ECM as well as their interactions. Therefore, it is crucial to investigate the responses of the renal and immune cells to a variety of stimuli and understand their role in the context of CKD. Moreover, investigations of cellular cross talks are needed to understand the interplay of different cells types, also with respect to the changes in the microenvironment. It is essential to bringing different experimental models together, including cell culture based, organoids, ex vivo studies, and animal models, and exploiting each of their strengths. Correlating findings generated by these studies with clinical data and insights from human studies like the Joslin Kidney Study (Dom et al., 2021; Dubin & Rhee, 2020; Kobayashi et al., 2022; Krolewski et al., 2017; Niewczas et al., 2019; Satake et al., 2021), Kidney Precision Medicine Project (KPMP) (de Boer et al., 2021; El-Achkar et al., 2024; Lake et al., 2021; Mariani et al., 2023), and Nephrotic Syndrome Study Network (NEPTUNE) (Lake et al., 2021; Lassé et al., 2023; Mariani et al., 2018; Pandey & Loscalzo, 2023; Reznichenko et al., 2021), we can hope to bridge the gaps in knowledge and pave the way for more effective treatment options for patients suffering from CKD. Enhanced understanding of kidney disease mechanisms can aid in dissecting patient heterogeneity, facilitating targeted therapy for (at least) a subset of patients. Moreover, it can help to identify a biomarker signature for reliable readouts of disease progression and treatment response, as well as surrogate end points (Henderson et al., 2020).

1.2. Extracellular Matrix (ECM)

In multicellular organisms, the extracellular matrix (ECM) provides a fundamental meshwork of proteins that serves several functions. While it is essential for tissue architecture and cell polarity, the ECM also provides bio-mechanical and bio-chemical cues that are involved in how the cell senses its environment (Naba et al., 2016; Ricard-Blum, 2011). Thereby, the ECM is implicated in controlling processes such as cell migration, proliferation and differentiation (Naba et al., 2016; Ricard-Blum, 2011).
To advance our understanding of ECM components, mass spectrometry (MS) studies have been essential. This has led to identification of roughly 270 core matrisome proteins that can be divided into proteoglycans (like biglycan (BGN), decorin (DCN), and serglycan (SRGN)), collagens and glycoproteins (such as fibronectin (FN1), elastin, laminin and fibrillins) (Naba et al., 2015, 2016).
In the kidney, the basement membrane consists mainly of laminin and Col IV and provides structural support for endothelial and epithelial cells, while the interstitial matrix provides a scaffold for myofibroblasts and fibroblasts (Kim et al., 2022). Migration of immune cells, fibroblasts and other resident renal cells is essential during the process of wound healing, which is supported by the ECM (Henderson et al., 2020). In fibrosis, the normal functional tissue is replaced by extracellular scar tissue. In CKD, this ultimately leads to loss of renal function and end stage renal failure (Kim et al., 2022).
Stimulated e.g. by growth factors like TGF-β, either secreted by other cell types or made available through remodeling of the ECM, differentiation of cells into myofibroblasts and their activation results in remodeling of the ECM (Kim et al., 2022). Not solely the abundance of ECM changes over the progress of renal fibrosis, but also its molecular composition, stiffness, pH and cellular components (Kim et al., 2022). Excessive deposition of extracellular matrix constituents, such as collagen and fibronectin, leads to scarring of the tissue and increased stiffness of the microenvironment through formation of aligned fibrillar structures (Edeling et al., 2016; Kim et al., 2022). Alterations in mechano-transduction further drive the profibrotic phenotype, promoting myofibroblast differentiation, EMT and Endo-MT as well as immune cell recruitment (Edeling et al., 2016; Henderson et al., 2020; Kim et al., 2022). Additionally, the tissue environment turns more acidic as there are increasing concentrations of Lactate dehydrogenase (LDH), due to tissue injury (freed up by dead cells), leading to an increase in lactic acid and, in turn, to activation of latent TGF-β

(Henderson et al., 2020; Kim et al., 2022). Therefore, LDH could be a possible early biomarker of fibrosis (Kim et al., 2022). Moreover, the cells undergo metabolic changes which are required for proliferation and collagen synthesis including enhanced aerobic glycolysis and altered fatty acid oxidation, in contrast to reduced oxidative phosphorylation (Henderson et al., 2020; R. Huang et al., 2023; Kim et al., 2022).

1.2.1. Collagens

Among core matrisome components including glycoproteins and proteoglycans, collagens are the most abundant proteins (Naba et al., 2016). In humans, there are 28 genes that encode collagens, that can be roughly divided into fibrillar (type I, II, III) and non-fibrillar types (An et al., 2016; Ricard-Blum, 2011). Collagens are synthesized as procollagens and later undergo extensive posttranslational modifications including hydroxylation, glycosylation and sulfation as well as cleavage my proteinases like A Disintegrin And Metalloproteinase with Thrombospondin motifs (ADAMTS), Bone Morphogenetic Protein-1 (BMP-1) and lysyl oxidases for maturation and cross-linking (Ricard-Blum, 2011). Especially in the context of tissue homeostasis, wound healing and fibrosis, the degradation of collagen via matrix metalloproteinases (MMPs) plays a central role (An et al., 2016; Naba et al., 2016; Rasmussen et al., 2019; Ricard-Blum, 2011).

Collagens and bioactive fragments generated by collagen cleavage interact with cell surface receptors and thereby participate in guiding processes including cell growth, differentiation, migration, development, angiogenesis and tissue repair (Ricard-Blum, 2011). In the context of CKD, a recent study further investigated collagen turnover profiles of CKD patients in stages 2 - 5 and a healthy reference group (Rasmussen et al., 2019). The study monitored collagen synthesis and degradation fragments found in the urine and serum and thereby demonstrated that subgroups of patients dependent on disease severity/stage can be identified using the proposed collagen turnover profiles. For instance, in patients at a moderate disease stage, markers of collagen degradation were higher than in patients with an advanced disease stage (Rasmussen et al., 2019). This provides evidence that the molecular mechanisms change over the course of fibrosis progression.

It has long been believed that the scarring is irreversible and permanent while recent papers have pointed to the fact that the scar tissue could regress under certain circumstances (Rockey et al., 2015). Effector cells would have to be eradicated and

the balance between synthesis and degradation of the matrix restored. Nevertheless, the aim of current treatments for CKD focuses on eliminating initial triggers (Rockey et al., 2015), like lowering the blood pressure or managing diabetes and blood glucose levels. The problem that I see in the case of CKD is, that, as the disease progresses, functional units of the kidney are lost and are replaced by fibrotic scar tissue. Given the complexity of the renal architecture and its diversity of highly specialized cells, additional measures are necessary to replace the lost endothelial and epithelial cells.

1.2.2. Accelerated ECM Deposition in Cellular Systems with Macro Molecular Crowding Agents

Deposition of ECM and collagen fibril formation are long processes. To study fibrosis in vitro, deposition of ECM can be accelerated using a so-called scar-in-a-jar model. This model uses macromolecular crowing (MC) agents that limit the available extracellular space and thereby support protein folding and complex stabilization with required enzymes. As a result, there is enhanced conversion of pro-collagen I to collagen I, improved assembly of collagen fibers, and stabilization of ECM components (Chen et al., 2009; Coentro et al., 2021; Hinz & Lagares, 2021; Rønnow et al., 2020). This set-up has proved particularly valuable, especially in the context of microscopy-based high-throughput screens to evaluate drug effects on the ECM, as demonstrated by Khan et al. who used a combination of Ficoll with different molecular weights (Khan et al., 2023), that I also applied in this thesis.

1.3. Omics Techniques

The human genome contains 20 000 to 25 000 genes, leading to a variety of 100 000 transcripts. Protein synthesis and post-translational modifications yield more than 1 million protein species (Virág et al., 2020). Omics is a term used to measure molecules in a high-throughput way. For instance, genomics methods like whole-exome sequencing (WES) and whole-genome sequencing (WGS) are used to provide a holistic view of the genome of a species. In the context of kidney fibrosis, genome sequencing is not widely used as only approximately 10% of end-stage renal disease in adults is associated to inherited kidney disease and hence 90% of cases cannot be explained by genetic variances (Saez-Rodriguez et al., 2019). Furthermore, epigenetic changes can be investigated using epigenomics which includes ChIP-seq (chromatin

immunoprecipitation sequencing) to identify TF binding and histone modifications, bisulfite sequencing to study DNA methylation, and ATAC-seq (Assay of Transposase Accessible Chromatin Sequencing) to assess the chromatin accessibility (Lindenmeyer et al., 2021).
Monitoring RNAs was becoming popular by the availability of microarrays which was soon replaced by RNA sequencing. Transcriptomics, with expanded (and theoretically whole) genome coverage, has provided new insights and discoveries. Currently, single cell RNA sequencing is used more and more often. This technique is especially useful to investigate the mRNA expression profiles in tissues with heterogenous cell populations, including the kidney (Kuppe et al., 2021). However, single cell RNA sequencing (seq) is still more expensive than transcriptomics and cannot provide full transcript coverage. Additionally, the spatial information is lost, which could be informative especially in the context of cell-cell interactions. This gap could be covered by combining single cell RNA seq with spatial transcriptomics, which is an approach that is currently gaining popularity. In general, information generated by transcriptomics and related techniques is limited due to the discrepancy between transcript number and protein abundance as well as activity (Lindenmeyer et al., 2021; Saez-Rodriguez et al., 2019).

Mass spectrometry (MS) techniques facilitated the high-throughput analysis of the cellular proteome, including protein abundance, post-translational modifications, and protein-protein interactions (Lindenmeyer et al., 2021). The proteome provides further information such as response to environmental factors and signaling mechanisms that are not always immediately translated into gene expression and mRNA abundance changes, especially in dynamic processes. Activity and post-translational modifications as well as binding to different molecules can only be addressed in a high-throughput manner using MS. However, currently available untargeted techniques are heavily biased towards the most abundant proteins and therefore do not cover the entire proteome. This could be overcome by applying targeted proteomics (Saez-Rodriguez et al., 2019). In the context of kidney diseases, proteomics applications have already facilitated identification of new biomarkers and signatures (e.g. CKD273 (Argilés et al., 2013; Cañadas-Garre et al., 2019; Rodríguez-Ortiz et al., 2018; Verbeke et al., 2021)) and characterization of fibrillar glomerulopathy (Saez-Rodriguez et al., 2019). One example includes the proposed predictive biomarkers for diabetic kidney disease

progression. For this, 17 proteins described as the “kidney risk inflammatory signature” (KRIS), were shown to be associated with renal failure development, including 6 members of the TNFR superfamily (Niewczas et al., 2019).
Moreover, by studying metabolomic changes, we are gaining further insights into disease mechanisms, providing new treatment options and early detection of CKD (Lindenmeyer et al., 2021).

Since omics measurements result in a huge amount of data, there are approaches to compress data including functional analysis to generate an overview of biological processes (like pathways) that are changing, driven by the measured factors (Saez-Rodriguez et al., 2019). Another approach involves network-based analysis, including protein-protein interaction (PPI) and gene regulatory networks (GRN). Those networks use the measured factors/biomolecules as nodes and depict interactions as edges. In the case of a PPI, the edges are based on physical interactions. As we know it from literature, there are factors like proteins that are involved in several biological processes and signaling pathways, such as p53. Such factors might represent nodes in a network with high connectivity (Lindenmeyer et al., 2021).

Even though omics methods, and especially the integration of several methods, could help gain further information on kidney disease, there are still many challenges remaining. Sample generation, data collection and storage are a few of the challenging aspects, limiting the potential to integrate generated data. Currently, there are efforts taken to overcome such burdens e.g. by generating standard operating procedures for sample and data collection and analysis as provided by the KPMP (Lindenmeyer et al., 2021).
Additionally, integration of omics data with other relevant data, including patient history, kidney function measurements and treatment regimens will be necessary to provide a holistic overview and link the molecular findings to relevant clinical phenotypes. In the NEPTUNE study cohort, for instance, not only omics of biopsy samples are acquired but also blood and urine samples, all supplemented with histopathological scoring and morphometry. NEPTUNE also provides longitudinal data due to comprehensive follow ups (Lindenmeyer et al., 2021). It is anticipated that such comprehensive integration of data will provide better patient stratification due to increased understanding of disease mechanisms. This, in turn, will result in more effective clinical study design,

targeting a subset of patients and improve diagnostics as well as personalized medicine (Henderson et al., 2020; Lindenmeyer et al., 2021).

2. Aims and Objectives

Despite progress in understanding CKD and even though several inhibitors of fibrosis, inflammation and oxidative stress related pathways have shown to be effective in animal models there is no clinically approved anti-fibrotic treatment for CKD yet (Prakash & Pinzani, 2017; Yamashita & Kramann, 2024). In absence of effective treatments, to compensate for reduced kidney function, patients have to undergo dialysis or kidney transplantation (Edeling et al., 2016).

Even if we were able to provide anti-fibrotic therapies that reverse fibrosis, it is still unlikely that the lost functional units of the kidney would regenerate (Yamashita & Kramann, 2024). On the other hand, it has been shown that early diagnosis could have a major impact in slowing down disease progression (Zürbig et al., 2019). Clinical symptoms only appear after irreversible damage of kidney function. Therefore, CKD is rarely diagnosed in early stages and ultimately patients will suffer from kidney failure (Mihai et al., 2018; Walraven & Hinz, 2018). Current diagnostic indicators of the molecular processes do not facilitate early detection, have limitations, and are not correlating with actual renal damage.

Based on these information, the following goals were set:

- One of the aims of this study was to establish a cellular fibrosis model and use an unbiased approach to study the change of molecular signatures and the effect on ECM deposition over time.
- Another goal was to identify potential targets that could provide new treatment options.
- Additionally, I aimed to propose potential biomarkers that could be taken further into clinical trials and might facilitate early diagnosis or correlate with disease progression.

3. Materials and Methods

3.1. Materials

3.1.1. Mammalian Cell Culture

The human kidney PDGFRβ^{+} cells was provided by the Kramann lab (Kuppe et al., 2021). Cells were cultured in DMEM 31885 (Gibco) supplemented with 5% (v/v) fetal bovine serum (Gibco) at 5% CO2 and 37°C.

3.1.2. Cell Culture Reagents

Table 2 Cell culture reagents. The supplier, identifier and working concentration are stated for the cell culture reagents used in this thesis.

Reagent	Supplier	Identifier	Working concentration
DMEM (Dulbecco's Modified Eagle Medium) 1 g/L D-glucose, L-Glutamine, 110 mg/L Sodium Pyruvate	Gibco	31885	
DMEM (Dulbecco's Modified Eagle Medium) 1 g/L D-glucose, no glutamine, no phenol red	Gibco	11880	
FBS (Fetal Bovine Serum)	Gibco	A5256701	5% (v/v)
L-Glutamine solution	Sigma Life science	G7513	
Ficoll® PM 770 kDa	Sigma-Aldrich	F2878	18.75 g / 500 ml
Ficoll® PM 400 kDa	Sigma-Aldrich	F4375	12.5 g / 500 ml
Recombinant Human TGF-beta 1 Protein	R&D systems	240-B-010	10 ng/ml
BSA (Bovine Serum Albumin)	Sigma-Aldrich	A2153	
L-Ascorbic Acid 2-phosphate (magnesium salt hydrate)	Cayman Chemical Company	16457	500 µM
Trypsin-EDTA 0.05%	Gibco	25300-054	
AmbionTM Nuclease-Free Water, DEPC-Treated	Life Technologies Corp.	AM9906	
DMSO (Dimethyl sulfoxide)	Sigma-Aldrich	D2438	
ScreenFect® Dilution Buffer	ScreenFect	S-2001	
ScreenFect® siRNA	ScreenFect	S-4001	

3.1.3. Drugs

Table 3 Drugs used for cell culture. For each drug compound tested (reagent), the supplier and identifier are shown.

Reagent	Supplier	Identifier
Dextromethorphan	Sigma Aldrich	D2531
Pirfenidone	Sigma Aldrich	P2116
Nintedanib	Sigma Aldrich	SML2848
LY2109761	MedChemExpress	HY-12075

3.1.4. Buffers and Solutions

Table 4 Buffers. Listed are Buffers and solutions that were used for immunofluorescence fixation, staining and washing as well as for western blotting and mass spectrometry.

Solution	Source	Composition
Phosphate Buffered Saline (PBS)	Media kitchen, EMBL	2.7 mM KCl 1.4 mM KH_2PO_4 4.8 mM Na_2HPO_4 137 mM NaCl in ddH2O pH 7.4, autoclaved
4% PFA fixation buffer	self-generated	16% PFA diluted in 10X PBS + ddH_2O
Triton X-100 permeabilization buffer	self-generated	1% Triton X-100 (v/v) in 10X PBS --> diluted 1:10 in ddH_20
Pierce™ RIPA Buffer	Thermo Scientific, REF#89900	with complete mini EDTA-free protease inhibitors (Roche)
Sample buffer 2X	self-generated	200 mM Tris-HCL 25% glycerol (v/v) 11.25% SDS (v/v) 325 mM DTT 0.0125% (w/v) bromphenol blue, pH 6.8
NuPAGE MOPS SDS running buffer (20X)	Thermofisher Scientific NP0001	Diluted to 1X buffer with ddH_2O
Western blot transfer buffer 10X	self-generated	2.9% Glycin (w/v) 5.8% Trizma Base (w/v) 0.5% BSA (w/v) in ddH_2O
Western blot transfer buffer 1X	self-generated	10X Transfer buffer diluted to 1X with ddH_2O 20% Methanol (v/v) 0.1% SDS (v/v)
1X TBS-T western blot wash buffer	self-generated	10X TBS (ThermoScientific, #J60764.K2) diluted to 1X with ddH_2O 0.2% Tween
Western blot blocking buffer	self-generated	1X TBS-T 5% BSA (w/v)
1% SDS buffer	self-generated	1% SDS 50 mM HEPES DEPC-treated water
100% TCA (Trichloroacetic acid)	self-generated	TCA and ddH_2O

continued on the next page

proteomics lysis buffer	generated by PCF	100 mM Tris-HCl pH 8.5 7 M Urea 1% Triton 5 mM Tris(2-carboxyethyl)phosphin-hydrochlorid 30 mM chloroacetamide 10 U/ml DNaseI (Sigma-Aldrich) 1 mM magnesium chloride 1 mM sodium orthovanadate phosphoSTOP phosphatase inhibitors (Sigma-Aldrich) complete mini EDTA-free protease inhibitors (Roche)
Protein digestion buffer	generated by PCF	100 mM Tris-HCl pH 8.5 1% sodium deoxycholate (Sigma-Aldrich) 5 mM Tris(2-carboxyethyl)phosphin-hydrochlorid 30 mM chloroacetamide
IMAC loading solvent	generated by PCF	70% acetonitrile 0.07% TFA

3.1.5. Oligonucleotides

3.1.5.1. Primers

Table 5 Primers. Primers used for RT-qPCR are shown in the table below with their corresponding sequences and order number.

Gene	Primer name	Sigma number	Primer sequence 5'→3'
GAPDH	GAPDH_F	8820603596-000010	CATGAGAAGTATGACAACAGCCT
	GAPDH_R	8820603596-000020	AGTCCTTCCACGATACCAAAGT
E2F1	E2F1_F1	8820713358-000310	ACGTGACGTGTCAGGACCT
	E2F1_R1	8820713358-000320	GATCGGGCCTTGTTTGCTCTT
SMAD1	SMAD1_F1	8820713358-000430	CTCATGTCATTTACTGCCGTGT
	SMAD1_R1	8820713358-000440	TATTCGCTGTGTCTTGGAACC
HNF4G	HNF4G_F2	8820713358-000370	TTGCAGGTTCAGTCGGCAAT
	HNF4G_R2	8820713358-000380	TTTCATTCCCGCTCTAAAACACT
FLI1	FLI_F1	8820713358-000390	CCAACGAGAGGAGAGTCATCG
	FLI_R1	8820713358-000400	TTCCGTGTTGTAGAGGGTGGT

continued on the next page

NR4A1	NR4A1_F1	8820755574-000010	GGACAACGCTTCATGCCAGCAT
	NR4A1_R1	8820755574-000020	CCTTGTTAGCCAGGCAGATGTAC
BHLHE40	BHLHE40_F1	8820713358-000190	GACGGGGAATAAAGCGGAGC
	BHLHE40_R1	8820713358-000200	CCGGTCACGTCTCTTTTTCTC
SERPINE1	SERPINE1_F1	8820713358-000070	AGTGGACTTTTCAGAGGTGGA
	SERPINE1_R1	8820713358-000080	GCCGTTGAAGTAGAGGGCATT
SPHK1	SPHK1_F2	8820713358-000130	GGCTGCTGTCACCCATGAA
	SPHK1_R2	8820713358-000140	TCACTCTCTAGGTCCACATCAG
ID3	ID3_F	8820713358-000150	GAGAGGCACTCAGCTTAGCC
	ID3_R	8820713358-000160	TCCTTTTGTCGTTGGAGATGAC
ID2	ID2_F	8820713358-000170	AGTCCCGTGAGGTCCGTTAG
	ID2_R	8820713358-000180	AGTCGTTCATGTTGTATAGCAGG
SMAD7	SMAD7_F2	8820713358-000250	GGACAGCTCAATTCGGACAAC
	SMAD7_R2	8820713358-000260	GTACACCCACACACCATCCAC
IGFBP3	IGFBP3_F1	8820713358-000270	AGAGCACAGATACCCAGAACT
	IGFBP3_R1	8820713358-000280	GGTGATTCAGTGTGTCTTCCATT
COL1A1	collagen I – FW	8812680987 0030	CAGCCGCTTCACCTACAGC
	collagen I_Rev	8812680987 0040	TTTTGTATTCAATCACTGTCTTG
COL1A2	COL1A2_F4	8820713358-000050	GAGCGGTAACAAGGGTGAGC
	COL1A2_R4	8820713358-000060	CTTCCCCATTAGGGCCTCTC
fibronectin	fibronectin – FW	8812680987 0050	GGAGAATTCAAGTGTGACCCTCA
	fibronectin – Rev	8812680987 0060	TGCCACTGTTCTCCTACGTGG
Col3	COL3fwd		GGAGCTGGCTACTTCTCGC
	COL3rev	8814570718-000140	GGGAACATCCTCCTTCAACAG
vimentin	vimentin fwd		CCAGGCAAAGCAGGAGTC
	vimentin rev		GGGTATCAACCAGAGGGAGT
aSMA/ ACTA2	α – SMA_FW	8812680987 0010	CCGACCGAATGCAGAAGGA
	α – SMA_Rev	8812680987 0020	ACAGAGTATTTGCGCTCCGAA

3.1.5.2. siRNAs

Table 6 siRNAs. This list contains information about siRNAs with their respective targets and sequences.

Target gene	ENSEMBL	Sense siRNA sequence 5'-3'	Antisense siRNA sequence 5'-3'	siRNA ID
SMAD1	ENSG00000170365	CCAAUAGCAGUUACCCAAAtt	agGGUUAUCGUCAAUGGGUUU	s8395
SMAD1	ENSG00000170365	GGAUAAAGUUCUUACUCAAtt	gaCCUAUUUCAAGAAUGAGUU	s8396
SMAD1	ENSG00000170365	CAGUCUAUGAGCUUACAAAtt	ctGUCAGAUACUCGAAUGUUU	s8394
E2F1	ENSG00000101412	AGAUCUCCCUUAAGAGCAAtt	agUCUAGAGGGAAUUCUCGUU	s223454
E2F1	ENSG00000101412	GGACCUUCGUAGCAUUGCAtt	gtCCUGGAAGCAUCGUAACGU	s4406
BHLHE40	ENSG00000134107	CGAACAUCUCAAACUUACAtt	ggGCUUGUAGAGUUUGAAUGU	s16282
FLI1	ENSG00000151702	CAAACGAUCAGUAAGAAUAtt	gtGUUUGCUAGUCAUUCUUAU	s5266
FLI1	ENSG00000151702	ACAGCUAUAUGGACGAGAAtt	gtUGUCGAUAUACCUGCUCUU	s5267
HNF4G	ENSG00000164749	GGGCUAAGCGAUCCAGUAAtt	ttCCCGAUUCGCUAGGUCAUU	s6701
NR4A1	ENSG00000123358	GCACCUUCAUGGACGGCUAtt	gtCGUGGAAGUACCUGCCGAU	s6678
COL1A1	ENSG00000108821	GCGAUGACGUGAUCUGUGAtt	caCGCUACUGCACUAGACACU	s3275
COL1A1	ENSG00000108821	CAAUCACCUGCGUACAGAAtt	tgGUUAGUGGACGCAUGUCUU	s3277
COL1A1	ENSG00000108821	AGGUUUCAGUGGUUUGGAUtt	tcUCCAAAGUCACCAAACCUA	s3276
Neg9		UACGACCGGUCUAUCGUAGtt	CUACGAUAGACCGGUCGUAtt	s444246
ID2	ENSG00000115738	ccCUGGAAGUCAACCUCGACU	ggGCUACUCGGACGAUAUGUU	s7108
ID3	ENSG00000117318	GGAAGGUGACUUUCUGUAAtt	gtCCUUCCACUGAAAGACAUU	s7112
GM130/GOLGA2	ENSG00000167110	GGUUCACAUUCAGACCAUAtt	gtCCAAGUGUAAGUCUGGUAU	s5942
SMAD2	ENSG00000175387	GGCUGUAAUCUGAAGAUCUtt	gtCCGACAUUAGACUUCUAGA	s8397
FN1	ENSG00000115414	GCCCGGUUGUUAUGACAAUtt	ttCGGGCCAACAAUACUGUUA	s5323

3.1.6. Antibodies and Fluorescent Dyes

3.1.6.1. Primary Antibodies and Fluorescent Dyes for Immunofluorescence and Western Blot

Table 7 Primary Antibodies. The table contains information about target, host, supplier, identifier, fixation used (for IF staining) and dilution of primary antibodies used in this thesis, either for immunofluorescence staining or western blot.

Target	Host	Supplier	Cat. No.	IF Fixation	IF Dilution	WB
Collagen Type I Antibody	rabbit polyclonal	Rockland	600-401-103-0.5	PFA	1:500	
Anti-Collagen, Type I antibody	mouse monoclonal	Sigma	C2456-.2ML	PFA	1:300	
Anti-Actin, α-Smooth Muscle antibody	mouse monoclonal	Sigma	A5228-100uL	PFA	1:100	
Vimentin Antibody	rabbit polyclonal	Proteintech	10366-1-AP	PFA	1:100	
Anti-Actin, α-Smooth Muscle antibody - FITC	mouse monoclonal	Sigma	F3777-.2ML	PFA	1:100	
anti-GFP	rabbit polyclonal	Origene	TP401	PFA	1:500	
phospho-SMAD2 (Ser465/Ser467) (E8F3R)	rabbit monoclonal	Cell Signaling	#18338			1:1000
Phospho-SMAD2 (Ser465/467)	rabbit polyclonal	Merck	AB3849-I			1:500
SMAD2	rabbit polyclonal	Proteintech	12570-1-AP			1:1000
Anti-alpha smooth muscle Actin antibody [1A4]	mouse monoclonal	Abcam	ab7817	MeOH	1:200 and 1:1000	
Tubulin-α AB-2	mouse monoclonal	Thermo Scientific	MS581			1:10000
Anti-Collagen I antibody	rabbit polyclonal	Abcam	AB34710			1:2000
Anti-HA tag	rabbit polyclonal to HA tag - ChIP Grade	Abcam	ab9110	PFA	1:500	
Anti-GM130	mouse	BD biosciences	610823	PFA	1:500	
Phalloidin Alexa Fluor 647		Invitrogen	A22287	PFA	1:200	

continued on the next page

CNA35-GFP		Protein expression core facility		PFA	1:250	
Hoechst H33342		Sigma	B2261	PFA/MeOH	1:1000	

3.1.6.2. Secondary Antibodies for Immunofluorescence and Western Blot

Table 8 Secondary Antibodies. Information about secondary antibodies, either for immunofluorescence staining or western blotting, is depicted in this list.

Target	Host	Supplier	Cat. No.	IF Fixation	IF Dilution	WB
Anti-mouse IgG secondary antibody, Oregon green 488	goat	Thermo Fisher Scientific	O-11033	PFA	1:400	
Anti-rabbit-IgG AlexaFluor 488	goat	Molecular Probes	A11008	PFA	1:400	
Anti-rabbit AlexaFluor 647	goat	Invitrogen	A21245	PFA	1:400	
Anti-mouse Peroxidase	rabbit	Sigma	A9044			1:8000
Anti-rabbit Peroxidase	goat	Sigma	A0545			1:16000

3.1.7. RNA Isolation and RT-qPCR Reagents

Table 9 RNA isolation and RT-qPCR reagents. Kits and reagents used for RNA isolation and RT-qPCR are listed in this table with their respective supplier and identifier.

Reagent	Supplier	Identifier
Total RNA Purification Kit	Norgen Biotek Corp.	17200
RNeasy® Mini Kit	Qiagen	74104
RNase-Free DNase Set	Qiagen	79256
DTT (DL-Dithiothreitol) solution	Sigma-Aldrich	3483-12-3
Ethanol	Merck	100983
Random Hexamer Primers	Invitrogen by Thermo Fisher Scientific	N8080127
RNaseOUT™ Recombinant Ribonuclease Inhibitor	Invitrogen by Thermo Fisher Scientific	10777019
dNTP mix, 10 mM each	Thermo Scientific	R0193
SuperScript™ IV Reverse Transcriptase	Invitrogen by Thermo Fisher Scientific	18090200
SuperScript™ III	Invitrogen by Thermo Fisher Scientific	18080093
SYBR™ Green PCR Master Mix	Applied Biosystems by Thermo Fisher Scientific	4309155
RNaseZap™ RNase Decontamination Solution	Thermo Fisher Scientific	AM9780

3.1.8. Other Reagents

Table 10 Reagents. Additional reagents that were not used for cell culture are summarized in this list.

Reagent	Source	Identifier
Agarose	Sigma-Aldrich	A-9539
Agarose Low-Gelling Temperature	Sigma-Aldrich	A9414
Ampicillin	Sigma-Aldrich	A-9393
Bovine Serum Albumin (BSA)	Sigma-Aldrich	A2153
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	D2438
DL-Dithiothreitol (DTT)	Sigma-Aldrich	D0632
EDTA-Free Protease Inhibitor Cocktail (PIC)	Roche	1.84E+09
Ethanol	Merck	1.00983
Ethylenediaminetetraacetic acid (Na_2EDTA)	Merck Millipore	324503
Glycerine	VWR	56-81-5
Glycine	Merck	1.04201
Hoechst 33342	Thermofisher Scientific	H21492
Kanamycin	Sigma-Aldrich	K-0254
Methanol	Merck	322415
Milk powder	Frema	-
Oligofectamine 2000 Transfection reagent	Thermofisher Scientific	12252011
Paraformaldehyde (PFA)	Thermofisher Scientific	50-980-491
Precision plus protein prestained standard marker	Bio-Rad	1610394
Sodium Chloride (NaCl)	Merck	1.06404
Sodium Dodecyl Sulfate (SDS) 20% Solution	Bio-Rad	1610418
Sodium Hydroxide (NaOH) Merck	Merck	1.06498
SYBR™ Gold Nucelic Acid Gel Stain	Thermofisher Scientific	S11494
SYBR™ Safe DNA Gel Stain	Thermofisher Scientific	S33102
Trichloroacetic acid (Tris-HCl)	Merck	1.0081
Triton X-100	Sigma-Aldrich	T8787
Trizma base	Sigma-Aldrich	T1503
Tween-20	Sigma-Aldrich	P7949

3.1.9. Laboratory Machines

Table 11 Laboratory Machines. Laboratory machines are listed in the table below.

Machine	Source
StepOne Real-Time PCR system	Thermo Fisher Scientific
QuantStudio 6 Flex	Applied Biosystems by Life Technologies
T100 Thermal Cycler	Bio-Rad Laboratories
Centrifuge 5417R	Eppendorf
Centrifuge 5408R	Eppendorf
Vortex-Genie 2	Scientific Industries
Mini-Centrifuge ROTILABO®	Carl Roth
Analytical balance TE124S-OCE	Sartorius
Tilt/roller mixer RS-TR05	Phoenix Instruments
ThermoMixer C	Eppendorf
Scale	VWR
Rotary shake	neoLab

Azure 280 Western blot imaging system	Azure Biosystems
ChemiDoc TM Touch Imaging System	Bio-Rad
Tissue culture incubator	Binder
Thermo Scientific Heraeus® BBD6220 incubator	Thermo Scientific
Water bath	GFLR
Electroporesis chamber	EMBL Heidelberg workshops
Magnetic stirring hotplate MR3001 K	Heidolph
Mini Trans-blot Cell Blotting system	Bio-Rad
Mini Trans-blot Cell system	Bio-Rad
Nanodrop 8000 Spectrophotometer	Thermo Scientific
Protein Gel Electrophoresis XCell SureLock System	Thermo Fisher Scientific
Infinite M1000 pro plate reader	Tecan

3.1.10. Equipment

Table 12 Equipment. Cell culture dishes and plates as well as other equipment are summarized in this table with the respective supplier and ifentifier.

Cell culture dishes / plates	Supplier	Identifier
NuclonTM Delta Surface 96-well plate	Thermo Scientific	167008
NuclonTM Delta Surface 24-well plate	Thermo Scientific	142475
NuclonTM Delta Surface 6-well plate	Thermo Scientific	140675
NuclonTM Delta Surface 10 cm dish	Thermo Scientific	150350
NuclonTM Delta Surface 15 cm dish	Thermo Scientific	168381
Glass bottom imaging plate 96-well	Zell-Kontakt	5241
Glass bottom imaging plate 24-well	Cellvis	P24-1.5H-N
Millipore® Stericup® Vacuum Filtration System	Millipore	S2GPU05RE
Cool Cell freezing container	Corning	CLS432004
Cryotubes	Thermo Scientific	363401
MicroAmp™ Fast Optical 96-Well Reaction Plate	Applied Biosystems by Life Technologies	4346906
MicroAmp™ Optical Adhesive Film	Thermo Fisher Scientific	4311971
Applied Biosystems™ MicroAmp™ Optical 384-Well Reaction Plate with Barcode	Applied Biosystems by Life Technologies	4343814
PCR strips of 8 tubes 0-2 ml	Ratiolab	8610040

3.1.11. Microscopes

Table 13 Microscopes. The widefield and confocal microscopes used in this study are stated with their corresponding sources.

Miroscope	Source
Automated widefield screening microscope Molecular Devices IXM	IXM
Confocal high-throughput Microscope LSM 900	Zeiss
Confocal microscope Nikon A1	Nikon

3.1.12. Software

Table 14 Software. The table contains software used and states the sources for each software.

Software	Source
Adobe Acrobat	Adobe Systems Incorporated, San Jose, USA
Adobe Illustrator	Adobe Systems Incorporated, San Jose, USA
APE A plasmid Editor	M. Wayne Davis, Jorgensen lab, Utah, USA
CellProfiler	Broad Institute, Cambridge, USA
Cytoscape	Institute of Systems Biology, Seattle, USA
Fiji ImageJ	Johannes Schindelin, Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany and others
Mendeley Desktop	Elsevier, Amsterdam, Netherlands
Microsoft Office	Microsoft Corporation, Redmond, USA
R 4.2.3 and 4.3.3	Comprehensive R Archive Network (CRAN)
StepOne Software v2.3	Thermo Fisher Scientific
QuantStudio™ Real-Time PCR Software v1.3	Thermo Fisher Scientific

3.2. Methods

3.2.1. Cell Biology

Cell Lines and Reagents

The human kidney PDGFRβ$^+$ mesenchymal cells were cultured in low glucose DMEM growth medium supplemented with 5% FBS. They were kept at 37°C and 5% CO_2 in a humidified incubator and regularly passaged when required, on average three times a week. The cells were regularly tested for mycoplasma and scored negative.

Cell Freezing

The initial vial received from the Kramann lab was passaged for 10 days before vials were frozen for further experiments. At 70% confluency, cells were washed with 0.5% Trypsin-EDTA solution (e.g. 1 ml per 10 cm^2 dish) and then incubated with fresh 0.5% Trypsin-EDTA solution at 37°C for 3-5 min until cells detached but before they started forming clumps. After neutralizing the enzyme activity with DMEM containing 5% FBS, cells were centrifuged at 300 g for 5 min. The cell pellets were subsequently resuspended in cell freezing medium (90% FBS with 10% DMSO) and aliquoted in 1 ml cryotubes. After 24 h in a freezing container at -80°C, the cells were then kept in liquid nitrogen tanks for long-term storage.

Plating Cells

The cells were thawed and cultured for 5-7 days before the start of an experiment. After trypsinization, the cells were centrifuged at 300 g for 5 min before plated for experiments. Initial experiments were performed to determine the cell number for plating to achieve the appropriate confluency for the intended plate format and culture time. To determine the cell number, the cells were counted using a hemocytometer and plated as indicated in the table below. The column “Culture time” describes the time from seeding of the cells until the end of the experiment, while the experimental time starts 24 h after cells have been seeded.

Table 15 Plate formats and corresponding cell numbers seeded. The cell number seeded per well is shown for the different plate formats used and culture times. Additionally, the culture volume per dish is shown.

Plate format	Culture time [hours]	Experimental time [hours]	Seeding cell number/well	Culture volume [ml]
10 cm dish	120	96	$0.490 * 10^6$	10
6-well plate	120	96	$0.060 * 10^6$	2
24-well plate	120	96	$0.015 * 10^6$	0.5
24-well plate	96	72	$0.020 * 10^6$	0.5
24-well plate	72	48	$0.025 * 10^6$	0.5
24-well plate	48	24	$0.050 * 10^6$	0.5
24-well plate	24	0-12	$0.075 * 10^6$	0.5
96-well plate	120	96	$0.0022 * 10^6$	0.2
96-well plate	96	72	$0.0027 * 10^6$	0.2
96-well plate	72	48	$0.0042 * 10^6$	0.2
96-well plate	48	0-24	$0.0047 * 10^6$	0.2

For the multi-omics and time point experiments, the medium of all wells was changed 24 hours after plating the cells. The day after cell seeding, the treatment of the longest time point was started (e.g. 96 h). In the 96 h condition, the medium of the control samples (96 h ctrl) was changed to low glucose DMEM, without phenol red and supplemented with 1% L-Glutamine, a mixture of Ficoll 70 and 400 as well as 500 µM ascorbic acid (ctrl condition containing molecular crowding). The same medium was used for the TGF-β treated samples, with the addition of 10 ng/ml TGF-β1. The medium of all the other conditions was changed to DMEM (phenol red-free) without FBS and supplemented with 1% L-Glutamine.

The following day, the medium was changed in the 96 and 72 h probes to DMEM+Ficoll+L-Glutamine+ascorbic acid (ctrl) and +TGF-β (TGF-β), while the medium of all other wells was changed to DMEM+L-Glutamine. In summary, the medium was changed every 24 h and the treatments were arranged so that all samples could be fixed/harvested for the intended downstream steps.

siRNA Transfections

The day before siRNA transfections, cells were plated as described in Table *15* and the previous section. On the day of the siRNA treatment, the cells were at 40-50% confluency.

For the transfection, the ScreenFect®siRNA protocol was used with the supplied reagents. For one well of a 24-well plate (1 reaction), 1 µl of ScreenFect®siRNA reagent was mixed with 39 µl of the recommended Dilution Buffer. In parallel, 0.17 µl

of siRNA stock solution (30 µM) were used, resulting in 5 pmol siRNA per well and mixed with 39 µl of Dilution buffer. Both reactions were mixed properly before combining the solutions with 10 rapid pipet strokes. After a 20 min incubation at room temperature, 420 µl of fresh DMEM (without FBS) was added. The cell medium was discarded before adding the total siRNA transfection suspension. For experiments in 6-well plates, the corresponding protocol proposed by the manufacturer was used. 5-6 hours after the transfection, the medium was changed to DMEM + 5% FBS and cells were cultured for the intended time.

For experiments in which knock-downs and TGF-β treatment (+ corresponding ctrls) were performed, the cells were cultured for 48 h in DMEM+5% FBS after siRNA transfection and before starting the TGF-β treatments. Afterwards, the medium was changed every 24 h to the TGF-β containing medium or the control (Ficoll). The cells were fixed/harvested at the corresponding time points (not all at once).

Nucleic Acid Handling

Primers delivered by Thermo Fisher arrived in a dried form and where therefore centrigued before resuspension in appropriate amounts of DEPC treated water to yield 100 µM stock solutions. Working solutions (10 µM) were generated by 1:10 dilutions and were used for RT-qPCR experiments.

For siRNA transfection, siRNAs were ordered also in dried form. After centrifugation, they were resuspended in appropriate amounts of nuclease free water to yield 30 µM stock solutions (e.g. 5 nM of siRNA + 166.66 µl of water). All reagents were stored at -20°C when not in use.

Drug Treatments

The day before drug treatments, the cells were plated as described previously (Table 15). On the day of the treatment, the cells had a confluency of about 50%. As a control, DMSO was added to the media containing the mixture of Ficoll 70 and 400 as well as 500 µM ascorbic acid. The same media (without DMSO), was supplemented with TGF-β and/ or Nintedanib (Nin), Pirfenidone (PFD), LY2109761 (LY), Dextromethorphan (Dex) (Table 3) and treatment was performed for 48 or 96 h before cells were fixed for imaging.

Immunofluorescence Assays

Immunofluorescence staining of structures and proteins of interest was performed with different antibodies and fluorescence dyes and different protocols. In general, cells were seeded on either 96-well plates with glass bottom (Zell-Kontakt) or 24-well plates with glass bottom and treated according to the experimental conditions described before. Before fixation, the cell medium was aspirated, cells were washed once with PBS and then fixed with 4% PFA fixation buffer containing 1:1000 Hoechst for 10-15 min at room temperature.The cells were consequently washed 3x with PBS and were directly used for or stored at 4°C until the start of the immunofluorescnec staining..

Extracellular Immunofluorescence Staining

For extracellular staining, which I mostly used to visualize the ECM, the cells were not permeabilized. Instead the cells were stained with the primary antibody anti-Col1 antibody (Rockland, 1:500 in PBS) for 1-1.5 h at room temperature. After 3x washes with PBS, the cells were further incubated with fluorescently labelled secondary antibodies in PBS for 30-45 min at room temperature. After the last 3 washes with PBS, the cells were kept in fresh PBS and either stored at 4°C or directly imaged.
Since the new batches of anti-Col1 antibody from Rockland did not work, the GFP-labeled dye CNA35 (produced in house, 1:250 in PBS) which directly binds to fibrillar collagen, was employed for the validation experiments (section 4.4, starting at page 115). For this, the cells were incubated after the fixation and washed at room temperature for 1-1.5 h. The cells were then washed 3x with PBS and directly used for imaging or stored at 4°C before imaging. Since the siRNA transfection resulted in a higher number of autofluorescent cells in the 488 channel, the cells were stained with an anti-GFP antibody followed by a secondary antibody conjugated to Alexa 647.

Intracellular Immunofluorescence Staining

For visualization of intracellular proteins (like GM130), the cells were permeabilized after the fixation for 15 min at room temperature with 0.1% Triton X-100. After this, cells were incubated with the primary antibody solution. Washing and secondary antibody staining steps were performed as described in the previous paragraph.

3.2.2. Microscopy

Wide-field Microscopy

Imaging of high-throughput experiments was performed with the Molecular Devices IXM microscope, an automated wide-field screening microscope. In general, 36 fields of view were acquired per well with a CFI P-Apo 20x Lambda/ 0.75 objective. All channels of interest were acquired per field of view, with the first channel (Hoechst / DAPI-5060C-NTE-ZERO Ex: 377/50-25 Em: 477/60-25) for the nuclear signal. Based on the nuclear signal, an off-set was set for the ECM signal either being in the GFP channel (GFP-3035D-NTE-ZERO Ex:472/30-25 Em: 520/35-25) for CNA35 staining or other channels if a secondary antibody labelled with another fluorophore was used (Cy3-4040C-NTE-ZERO Ex:531/40-25 Em: 593/40-25, TXRED-4040C-NTE-ZERO (mcherry) Ex:562/40-25 Em: 624/40-25, Cy5-4040C-NTE-ZERO 6 Ex:28/40-25 Em: 692/40-25).

Confocal Microscopy

The confocal microscopes Nikon A1 and LSM 900 (Zeiss) have been used for sample inspection and visualization. However, only images acquired by the LSM 900 were included in this thesis. Hence, I focused on providing information of the latter. Especially for visualization of changes of the cell cytoskeletons (actin and vimentin fibers), Z-stacks were acquired covering the entire cell thickness. For this, the parameters and number of Z-stacks was kept constant. Following filters and objectives were used in the experiments: Objective Plan-Apochromat 20x/0.8 M27 air (FWD=0.55mm), Lasers: 405 nm – 5 mW, 488 nm – 10 mW, 640 nm – 5 mW, confocal detectors: Gallium Arsenide Phosphid-PMT (GaAsP-PMT) for fluorescence and standard PMT as transmission detector, Filters: EX BP 385/30 for DAPI/Hoechst, EX BP 469/38 for FITC/Alexa Fluor 488, EX BP 631/33 for Cy5, QBS 405+493+575+653, EM QBP 425/30+514/30+592/25+709/100.

3.2.3. Image Analysis of Wide-Field Microscopy Images

Image inspection (i.e., background fluorescence, cellular autofluorescence, image quality assessment, etc.) was conducted using Image J (Schindelin et al., 2012). For analysis, including nuclei segmentation, counting of nuclei as well as quantification of

fluorescent signal of immunofluorescence staining, CellProfiler (Stirling et al., 2021) pipelines were developed. Subsequent downstream analysis was executed in R (R Core Team, 2021).

Nuclei Count and Immunofluorescence Intensity Measurements with CellProfiler

An example of the used CellProfiler pipeline and parameters for the analysis of the collagen wide-field images are shown in the appendix (**CellProfiler Pipeline**, page XXXV). The following is a brief description of the essential steps of the pipeline with the corresponding CellProfiler modeule names in italics. In a first step, the acquired images were loaded into CellProfiler (*Images*). Using a consitent file naming scheme, the image files were assigned to the different channels (*NamesAndTypes*), e.g. channel “C01.ome” as Nuclei or “C02.ome” as Collagen. Nuclei were seqmented using an intensity based thresholding , including an object size filter (*IdentifyPrimaryObjects*). In the following step, nuclei segmentation results were saved for quality assessment (*ConvertObjectsToImage* and *SaveImages*). For the Collagen/ECM staining (or other stainings like for GM130), the fluorescence was measured using the *MeasureImageIntensity* module. The resulting data were exported to a spreadsheet as CSV files (*ExportToSpreadsheet*). The code was executed on a virtual desktop computer at EMBL (jupyterhub.embl.de).

Further Analysis Steps Using R

All further analysis steps were performed in R, using the above CellProfiler measurements as input. The column names of the CellProfiler output table are given in italics. After import of concatenated CellProfiler CSV files, the well and position information of each image was extracted from *FileName_Collagen*. Based on this information, the treatments were assigned (time point, +/- TGF-β, siRNA treatment, etc.). The *Intensity_TotalIntensity_Collagen* values were multiplied by 65535 (in order to rescale to the original 16 bit value range) followed by cover slip background subtraction (value of cover slip background (MEAN_BG_65535*,* measured using Fiji ImageJ) times total area of image (*Intensity_TotalArea_Collagen*)). Based on the unstained control, the autofluorescence of the cells was calculated per cell and subtracted from all other images corresponding to the cell number. Additionally, subtraction of another value, “non-fibrillar Col I/ECM staining” was performed. In the R script this was called “unspecific antibody binding”. This step was included as the cover

slip background in the unstained (secondary antibody only) controls was usually lower than in the stained images. Hence, correction for any unspecific or non-fibrillar Col I/ECM binding of the primary antibody or fluorescent dye was conducted. The resulting intensity values per image were then normalized to the nuclei number of each image. Images with nuclei number lower than 20 nuclei (sometimes higher like 30 nuclei) were excluded for further analysis. The decision on taking a nuclei number cut off, was based on the fact that images with hardly any nuclei were often out of focus, or the fluorescent intesity was not correlating with actual cell number (e.g. when the intensity was influenced by dust particles in the area of interest). In the annex, I included figures for each experiment showing the normalized fluorescence per cell on the y-axis versus the nuclei number on the x-axis as a quality inspection and initial overview of the data. The data frames of the different replicates were subsequently merged and used to calculate ratios between the selected samples.

To show the difference between ECM deposition of one condition versus the other (e.g. TGF-β treated versus control), ratios were calculated (TGF-β divided by control) per respective replicate. More specifically, for the experiments including determination of cell numbers and comparison of experimental set-up A vs B, exactly the above-mentioned calculation was used, where, in the case of the cell number experiment, the Col I intensity per cell of the TGF-β stimulated sample was divided by the mean of the Col I per cell for the control sample, for each cell number and time point separately.

Following calculation of ratios was used for the TGF-β response curve and treatments with drug compounds as well as siRNA treated samples. Taking the example of siRNA treatments, I normalized the ECM staining calculated per cell for each sample treated with the siRNA to the ECM per cell of the siNeg9 (non-targeting siRNA) control by subtraction and then calculated the ratio by division, as follows:

$$\text{ratio} = \frac{\text{ECM/cell}_{(\text{siRNA})} - \text{ECM/cell}_{(\text{siNeg9})}}{\text{ECM/cell}_{(\text{siNeg9})}}$$

This was done for control and TGF-β treated samples, always normalized to the siNeg9 control (without TGF-β) sample, separately per plate/biological replicate. Additionally, I calculated the ratio between the TGF-β treated samples and therefore divided the

siRNA + TGF-β treated samples by the siNeg9 + TGF-β samples. For statistical analysis, t-test was used to indicate statistically significant differences.

For more detailed analysis of Col I deposition over time (Results section 4.2.2), I applied a linear mixed model, accounting for the factors treatment (TGF-β or control), time (time points 0, 12, 24, 48, 72 and 96 h) and their interaction as fixed effects, as well as random effects based on the applied experimental design, i.e., the factor plate (biological replicate).

$$\text{sqrt\_mean\_intensity} \sim \text{condition} * \text{time} + (1 \mid \text{plate})$$

First, I duplicated the values of the 0 h time point, as there is no 0 h TGF-β treatment. Shown by the graphical representation of the model matrix, generated with the ExploreModelMatrix R package (Soneson et al., 2020)(**Figure *3***), there is a fitted value for the control condition at 0 h (intercept that changes for each plate). Another parameter is conditionTGF (TGF-β treatment). At the 0 h time point the fitted value is the sum of intercept and the conditionTGF coefficient. Time is a further parameter which is determined for each time point (time12h, time24h, etc.). Moreover, there is an interaction term for time and condition (conditionTGF:time12h, conditionTGF:time24h, etc.). For instance, as a result of the additive model that I used, the treatment effect of e.g. TGF-β at 12 h is represented by the sum of the intercept, conditionTGF, time12h and the interaction of conditionTGF and time12h, represented as conditionTGF:time12h. The average intensity of Col I per cell for all images acquired under each condition was calculated, for each technical and biological replicate. Examining the residuals, I decided to transform the previously calculated averages of the Col I per cell values using square root transformation which resulted in the quality graphs below (**Figure *3***).

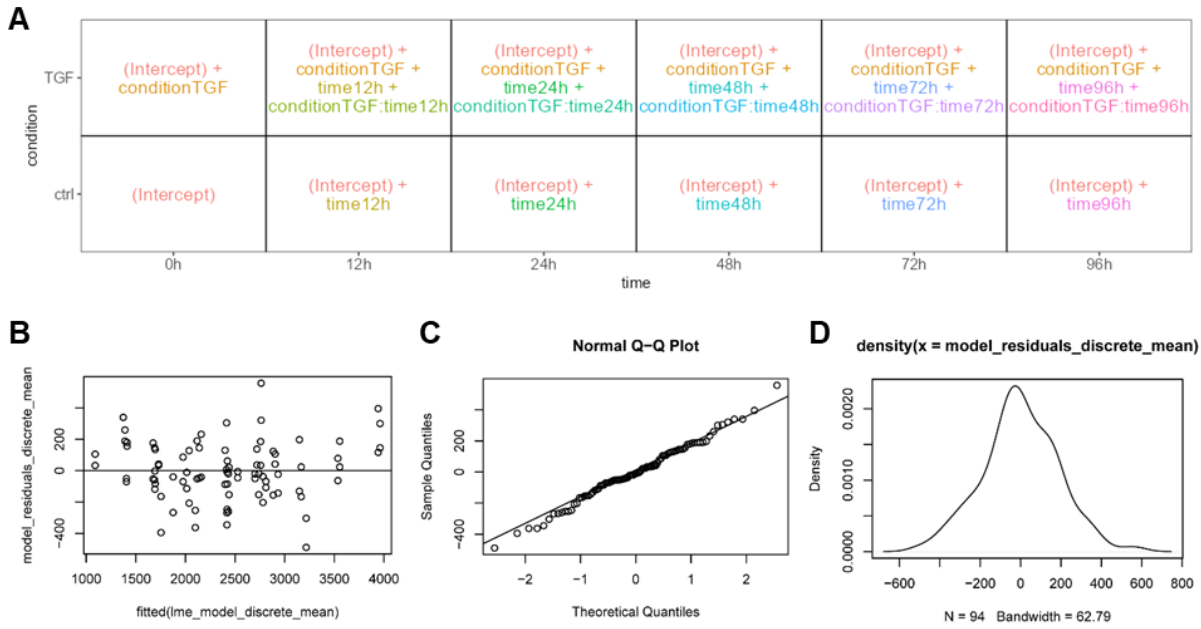

Figure 3 Graphical representation of model matrix and residuals. A) Model matrix representation. **B)** Residual plot showing relatively constant variance across fitted values. X-axis depicts fitted values vs model residuals on the y-axis. **C)** Normal Q-Q and **D)** density plots to inspect normal distribution of residuals.

Subsequently, analysis of variance was performed using type III sum of squares followed by a post hoc test (as stated in results 4.2.2) to show significant differences between factor levels. For visualization, the sqrt_mean_int data were normalized per plate by subtraction of the values from the 0 h time point. The resulting data points were plotted and the predicted values of the model were added via a line (Figure 25).

3.2.4. Biochemistry

Cell Lysis and Sample Preparation

For cell lysis, the cells were washed two times with ice-cold PBS and then lysed in Pierce™ RIPA buffer (Thermo Scientific) supplemented with a proteinase inhibitor cocktail for 5 min on ice. The lysates were subsequently centrifuged at 14000 g for 15 min at 4ºC. The supernatants were taken and used for further analysis or frozen at -80ºC until use. Protein samples were mixed with 2x sample buffer and incubated at 98ºC for 5 min.

SDS-PAGE and Western Blot

The protein samples were loaded on pre-cast NuPAGE™ 4-12% Bis-Tris Gels (Invitrogen Thermofisher Scientific). As markers either PageRule ™ Plus Prestained Protein Ladder (Thermo Scientific, Lot# 01114357, Product#26619) and/or Prestained Protein Marker MW 10-180 kDa (Proteintech, Cat No PL00001, Lot 39000001) were loaded to determine the molecular weight of the target proteins. The gels were then run using NuPAGE® MOPS running buffer (Life Technologies, LOT 1904556, REF NP001) at 90-100 V for 130 min.

The transfer of proteins to PVDF membranes (pore size 0.45 µM, Merck, Millipore), was performed after the membranes were activated in methanol. For the transfer, the sandwiches (including sponges, Whatman papers, PVDF membrane and SDS-PAGE gels) were transferred into a Mini-Protean II cell gel system (Bio-Rad) filled with Transfer buffer and kept in the cold room for the time of the transfer (1.5 h at 100 V). Subsequently, the membranes were incubated in blocking buffer (TBS-T with BSA and milk powder) for 1 h at room temperature. Between each of the following steps, the membranes were washed 3x with TBS-T buffer, 5-10 min each. The membranes were either incubated at room temperature for 1 h or at 4ºC over night with the primary antibodies (Table 7). The secondary antibodies (Table 8) were HRP-coupled and used for incubation of membranes for 45 min at room temperature. For visualization of the target proteins, the membranes were treated with chemiluminescence substrate (Pierce® ECL Plus Western Blotting Substrate, Thermofisher Scientific) according to the manufacturer's instructions. The Azure 280 (Biozym) or Biorad imager were used for image acquisition. Protein bands were quantified with Fiji ImageJ. All bands were normalized to the corresponding loading control. Additionally, the resulting values were taken to calculate the ratio of phosphorylated SMAD2 vs SMAD2 protein abundance.

RNA Isolation, Quantification and Purity Measurement

Except for the multi-omics experiment, the RNeasy kit (Qiagen) was used to extract RNA from the cultured cells. Therefore, the protocol provided by the manufacturer was followed, except for the DNase treatment step, where half of the recommended concentration was used (as initial experiments have shown that this resulted in sufficient digestion of DNA with the used amount of input material). RNA concentration and purity were measured with the Nanodrop 8000 Spectromphotometer using the absorbance measurements at 260 nm and ratio between 260 and 280 nm. The read-

outs were then used for adjusting RNA concentrations for the following step. In between steps, samples were stored at -20°C for short term storage or -80°C for long time storage.

Reverse Transcription

Total RNA (300 – 500 ng) was used for reverse transcription. In initial experiments (Results section 0 and 0), reverse transcription was performed using SuperScript III, while SuperScript IV was used for the validation experiments (Results section 4.4) as stated in the manufacturer's protocol. cDNA was generally diluted 1:10 after reverse transcription and used for RT-qPCR.

RT-qPCR

For the RT-qPCR experiments, there were differences between initial and validation experiments. For the initial experiments, a master mix was prepared with a total volume of 15 µl per reaction, containing 0.5 pmol/µl of each of the primers (forward and reverse), 10 µl of the Sybr Green mix and the rest water (3 µl). Based on optimization experiments for quantification of fibronectin mRNA levels, only one fourth of the fibronectin primer concentrations were used. Therefore, more water was added to the reaction to yield the same volume as other reactions. Primers are shown in Table 5. Then, 5 µl of the 1:10 diluted cDNA was added. Three technical replicates were prepared in a 96-well plate that was centrifuged before samples were measured in the StepOne Real-Time PCR machine. This was done using the manufacturer's instructions with following selections: experimental type 'Standard curve', reagent type 'SYBR® Green reagents', including melt curve, 'Standard 2 h', and 'cDNA'. The program ran roughly 2 hours, including a 10 min hot start phase at 95°C, 40 cycles consisting of i) 15 sec denaturation at 95°C, ii) 1 min for primer annealing and template extension at 60°C, followed by iii) the melt curve generation. Initial analysis was performed with the StepOne Software v2.3.

For later experiments (Results section 4.4), the reactions were prepared in 384 well plates. Therefore, two master mixes were prepared. One master mix contained the primers (1 µl each) with 3 µl Sybr Green mix per reaction. The second master mix contained 2 µl of Sybr Green mix, 2 µl of water and 1 µl of 1:10 diluted cDNA per reaction. For each condition, I prepared 3 to 4 technical replicates. Samples were run

at the QuantStudio 6 Flex according to manufacturer's instructions and analysed using QuantStudio Real-Time PCR Software v1.3.

RT-qPCR Analysis

Technical replicates were removed if they were outliers (only if standard deviation was > 0.5 and if it was obvious that the sample is an outlier). The $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen, 2001) was applied for analysis. In short, the average CT value of technical replicates was then normalized (by subtraction) to the average CT value of the housekeeping gene (GAPDH), this was done for each sample respectively (for instance: ΔCT_1 = CT control target – CT control GAPDH; ΔCT_2 = CT TGF-β target – CT TGF-β GAPDH). For calculating the difference in expression of mRNA abundance between the e.g. TGF-β treated sample and the control sample, the GAPDH-normalized CT value of the control was subtracted from the GAPDH-normalized CT value of the TGF-β treated sample resulting in the ΔΔCT value (such as $\Delta\Delta CT = \Delta CT_2 - \Delta CT_1$). mRNA expression was generally stated as percentages, so that the control was always set to 100% as calculated by $2^{-\Delta\Delta CT}$ times 100 (e.g. percentage of fold-change = $2^{-\Delta\Delta CT} * 100$).

To determine the statistical significance of the difference of mRNA expression of target genes induced by TGF-β treatment, statistical analysis of initial experiments was performed by using t-test against μ = 100 (control condition) for each gene and each time point separately.

For later experiments (Results section 4.4), the initial steps were the same. However, the sample to which each other ΔCT value was compared to was the siNeg9 control sample at 48 h post siRNA transfection. This was done to investigate the change of mRNA expression over time across the different conditions.

For the statistical analysis, a t-test was performed initially. This was done to compare samples at the last time point (96 h post siRNA transfection) plotted in the heatmap (**Figure *53***). Therefore, a t-test was conducted between the 96 h siRNA control and the 96 h siNeg9 control, as well as for the 96 h siRNA TGF-β and the 96 h siNeg9 TGF-β samples, respectively, to compare the untreated versus the TGF-β treated samples with each other and check the relevance of the knock-down of the siRNA targets in the presence and absence of TGF-β.

This was followed up by anova and post hoc tests for selected comparisons. For instance, one of the questions was to see whether the expression of COL1A1 mRNA

changes upon knock-down of a target gene in the presence and absence of TGF-β treatment compared to the control knock-down with siNeg9. Therefore, an anova was performed with testing the significance of the percentage of the fold-change (percentage) based on siRNA-based knock-down (siRNA) and ± TGF-β treatment (treatment).

$$\text{percentage} \sim \text{siRNA} * \text{treatment}$$

Especially when the interaction term of siRNA and treatment were significant, I followed up with a Tukey post hoc test of the model presented above to show the significance of the different interactions. To note, there seem to be more siNeg9 samples. However, for each biological replicate, controls treated with siNeg9 were generated, that were then loaded onto each plate. Since the sample samples were loaded as references samples on different plates used for RT-qPCR, there are more technical replicates of the siNeg9 samples, which explains the higher number of data points in the graphs in the annex of this thesis.

3.2.5. Multi-omics Experiment

Experimental Set-up

For the multi-omics experiment, including the imaging of Col I deposition over time, first, one vial of cells per day was thawed at four consecutive days, for the four biological replicates. Cells were cultured 6 days in DMEM + 5% FBS before they were plated at a 30-40% confluencey for the experiment.

For imaging, 0.015 mio cells were seeded per well of 24 well plates. Four biological replicates (plate A-D) were generated, by setting up the experiments at 4 consecutive days). Culture volume per plate was 0.5 ml. For the transcriptomics samples, 0.06 mio cells were seeded per well of a 6 well plate, again on three consecutive days and kept in 2 ml medium. Protoemics and phosphoproteomics samples were generated from 0.49 mio cells seeded in 10 cm dishes (8 ml culture volume), two dishes per condition. For this, four biological replicates were generated on four consecutive days. The supernatants of these plates were used for secretomics in the end (more detail in the following paragraphs).

In general, the day after seeding, the media of all samples of one replicate was changed to DMEM without FBS. Except for the sample with the longest treatment time point (96 h), the media was changed to phenol free DMEM without FBS but with Ficoll (mixture of Ficoll 70 and 400) and 500 µM ascorbic acid for the control condition. The later media containing Ficoll and ascorbic acid was further supplemented with 10 ng/ml TGF-β for the TGF-β treated samples. On the second day after seeding, the media of all samples was changed again. The still untreated samples were changed to DMEM without FBS, while media containing Ficoll and ascorbic acid was added to the 96 and 72 h samples and media containing Ficoll, ascorbic acid and TGF-β was added to TGF-β treated samples (72 and 96 h) after removing the old medium. So the media was changed every 24 h, with every day staring another treatment so that at the last day, all samples could be either fixed for imaging or collected for the corresponding omics measurements.
For imaging, the samples were fixed for 15 min with 4% PFA at room temperature, before they were washed and stained as explained in 3.2.1 decribing immunofluorescence staining of extracellular Col I. The other samples were processed as described in following sections.

RNA Sequencing

RNA Isolation

The cell culture supernatants were removed and the cells were washed once with PBS. For cell lysis (Cell Lysate Preparation from Cells Growing in a Monolayer) and RNA isolation, the Total RNA Purification Kit (Norgen Biotek Corporation, Product #17200) was used and procedures were carried out as described in the manufacturer's protocol. Additionally, the On-Column DNA Removal Protocol was performed with the Norgen's RNase-Free DNase I Kit (Product # 25710).

RNA Quality Control

The quality of the RNA samples was assessed with the Agilent Bioanalyzer with the RNA 6000 Nano Assay kit and following the manufacturer's instructions. This step was performed with the help of Ferris Jung from the Genomics Core Facility, EMBL Heidelberg. For each sample, 300 ng of RNA were used and the volume was adjusted to 50 µl.

Library Preparation

Ferris Jung prepared the libraries on a Beckman Coulter Automated Workstation Biomek i7 Hybrid (MC + Span-8). Therefore, an automated version of the NEBNext® Ultra™ II Directional RNA Library Prep Kit was used, following section 1 - Protocol for use with NEBNext Poly(A) mRNA Magnetic Isolation Module. An adaptor dilution of 1 to 20 was used, the samples were individual barcoded using unique dual indices during the PCR using 13 PCR cycles as per the manufacturer's protocol. The individual libraries were quantified using the Qubit HS DNA assay as per the manufacturer's protocol. Therefore, 1 µl of sample in 199 µl of Qubit working solution were used. The DNA HS Assay kit together with the Agilent Bioanalyzer was used to assess the quality and molarity of libraries. The latter was necessary to equimolarly combine the individual libraries into one pool for sequencing.

Sequencing

The pooled samples were loaded and sequenced by Ferris Jung on an Illumina NextSeq 2000 platform (Illumina, San Diego, CA, USA) using a P3 50 cycle kit, a read-length of 72 bp single-end reads and 650 pM final loading concentration.

Analysis

Sequencing reads were aligned by Jonathan Landry, Genomics Core Facility, EMBL Heidelberg, using STAR (version 2.7.9a) with default parameters on GRCh38. The gene count tables were produced during the alignment (--quantMode GeneCounts) using the annotation GRCh38.93. Following steps were carried out by me with the support of Mira Burtscher, with initial help of Martin Garrido-Rodriguez. ENSEMBL IDs were translated into gene symbols using the org.HS.eg.db R package. Data were filtered to exclude low expressed genes and normalised using TMM (Trimmed mean of M-values) + CPM (counts per million), by using the function calcNormFacots (calculating normalisation factors, as certain genes have higher read counts that need to be accounted for) and cpm of the edgeR R package (Smyth et al., 2018).
After principal component analysis and further investigations, the replicate A of the 24 h control sample was excluded due to an experimental error. Differential expression analysis was performed using edgeR, including multiple testing correction (BH) resulting in adjusted p-values. As differentially expressed genes, I considered genes

with an adjusted p-value < 0.05 and absolute log2 fold-change of > log(2) as significantly altered in expression. For visualisation of data, I used the statistical language R.

Mass Spectrometry

Sample Lysis for Whole Proteome and Phosphoproteome

On-plate cell lysis of samples cultured in 10 cm dishes was performed with 500 µl proteomics lysis buffer. The lysates were stored at -80°C before further processing by Jennifer Schwarz from the Proteomics Core Facility who performed further sample preparation, data acquisition and MS database search.

Sample Preparation and Labeling for Whole Proteome and Phosphoproteome

The samples were processed in the Proteomics Core Facility of EMBL Heidelberg by Jennifer Schwarz following the protocols described in Potel et al. (2018) for preparation and phosphopeptide enrichment (Potel et al., 2018). The frozen samples were thawed and sonicated with an unltrasonic sonifier (Branson) three times for 10 sec with intervals of cooling the samples on ice for 60 sec (continuous pulse, 50% duty cycle) until the viscosity was reduced. By centrifuging the samples at 17000 g at 8°C for 10 min, any residual cell debris was removed. The supernatants were harvested and supplemented by 1% benzonase (Merck Millipore) followed by incubation at RT for 1 h. Methanol/chloroform precipitation was performed. Therefore, 1 volume of sample was supplemented with 4 volumes of methanol, 1 volume of chloroform and 3 volumes of ultrapure water. After centrifugation at 4000 g for 15 min, the upper layer was removed. Then, 3 volumes of methanol were added to the rest of the samples and the samples were centrifuged again. After removal of the liquid phase, the protein precipitate was left to dry before the proteins were resuspended in digestion buffer. Subsequently, trypsin was added in a 1:50 ratio (w/w) and was incubated over night at room temperature for digestion of proteins. The digestion was stopped by addition of 1% TFA (final concentration). Precipitation of the sodium deoxycholate was performed at room temperature for 15 min after which the samples were centrifuged for 10 min at 17000 g at room temperature. Desalting of the supernatant was performed using the Oasis HLB 96-well plates 30 µM (Waters). Desalting buffer A contained MS-grade water (Chemsolute) with 0.1% formic acid and buffer B consited of 80% acetonitrile

(Chemsolute) in MS-grade water with 0.1% formic acid. Using a vacuum centrifuge, the eluted peptides were dried.

Posphopetides were enriched by first taking up the peptides in IMAC loading solvent. Before further proceding with phosphopeptide enrichment, a small aliquot of samples was used for full proteome analysis. The rest of the samples were enriched using an UltiMate 3000 RSLC LC system (Dionex) and ProPac IMAC-10 column 4 x 50 mm, P/N 063276 (Thermo Fisher Scientific). The peptides were then eluted with 0.4% dimethylamine (Sigma-Aldrich) for labelling with TMT16plex (Thompson et al., 2019) Isobaric Label Reagent (Thermo Fisher) in accordance with the manufacturer's protocols. Therefore, 0.8 mg reagent was dissolved in 42 µl of 100% acetonitrile and 4 µl of stock was added and incubated for 1 h at room temperature. Subsequently, the reaction was quenched with 5% hydroxylamine for 15 min at room temperature. Sampes wre then combined for further sample clean up using an OASIS® HLB µElution Plate (Waters). Using high-pH reversed-phase fractionation on an Agilent 1200 Infinity high-performance liquid chromatography system, including a Gemini C18 column (3 µm, 110 Å, 100 x 1.0 mm, Phenomenex), the labelled phosphoproteome and full proteome were separated into 48 fractions along with the LC separation that were then pooled into 12 fractions. The pooled fractions were dried using vacuum centrifugation and were subsequently reconstituted in 10 µl of 1% fromic acid with 4% acetonitrile. The samples were stored at -80°C before the LC-MS/MS analysis.

Whole Proteome and Phosphoproteome Data Acquisition

The UltiMate 3000 RSLC nano LC system (Dionex) fitted with a trapping cartridge (µ-Precolumn C18 PepMap 100, 5 µm, 300 µm i.d. x 5 mm, 100 Å) and an analytical column (nanoEase™ M/Z HSS T3 column 75 µm x 250 mm C18, 1.8 µm, 100 Å, Waters) was coupled to a Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo). With a constant flow of 0.05% trifluoroacetic acid at 30 µl/min, peptides were concentrated on the trapping column for 6 minutes. Then, peptides were eluted via the analytical column using a binary solution system at a constant flow rate of 0.3 µl/min. Solvent A consisted of 0.1% formic acid in water with 3% DMSO and solvent B contained of 0.1% formic acid in acetonitrile with 3% DMSO. The percentage of solvent B was increased as follows: from 2% to 4% in 4 min, to 8% in 2 min, to 25% in 64 min, to 40% in 12 min, to 80% in 4 min, followed by re-equilibration back to 2% B in 4 min (for the phosphoproteome). For the full proteome analysis, the steps were as follows:

from 2% to 8% in 4 min, to 28% in 104 min, to 40% in 4 min, to 80% in 4 min, followed by re-equilibration back to 2% B in 4 min.
The peptides were injected into the Fusion Lumos applying a Pico-Tip Emitter 360 µm OD x 20 µm ID; 10 µm tip (New Objective) and a defined spray voltage of 2.4 kV. The capillary temperature maintained at 275°C. Full mass scan was acquired with mass range 375 to1400 m/z for the phosphoproteome (375 to 1500 m/z for the full proteome), in profile mode in the orbitrap with resolution of 120000. The filling time was set at maximum of 50 ms for the full proteome with a limitation of 4 x 105 ions. Data dependent acquisition (DDA) was done with the resolution of the Orbitrap set to 30000, with a fill time of 110 ms for the phosphoproteome (94 ms for the full proteome) and a limitation of 1 x 105 ions. For this, normalized collision energy of 34 was applied. MS2 data was acquired in profile mode. Fixed first was mass was set 110 m/z.

Whole Proteome and Phosphoproteome MS Database Search

The acquired data were processed using MSFragger v3.8 (Kong et al., 2017), searching against the homo sapiens Uniprot proteome database (UP000005640, ID9606, 20594 entries, release October 2022) including common contaminants and reversed sequences. Fixed modifications that were included are Carbamidomethyl (C) and TMT16 (K). Variable modifcations were set as Acetyl (protein N-term), Oxidation (M) and TMT16 (N-term), and for the phosphoproteome phosphorylation of STY sites was considered. For the MS1 and MS2 scans a mass error tolerance of 20 ppm was set. Additional parameters included the use fo trypsin as protease, allowing for a maximum of two missed cleavages; a minimum peptide length of seven amino acids; at least two unique peptides were required for a protein identification. The false discovery rate on peptide and protein level was set to 0.01.

Sample Collection and TCA Precipitation for Secretome

The supernatants of the samples were collected and spun down at 300 g for 5 min at 4°C to remove any cells. Afterwards, the supernatants were transferred to new falcon tubes and snap frozen. Until further processing, the samples were stored at -80°C.
To enrich the proteins for further down stream analysis, TCA precipitation was used. Therefore, all the samples were transferred to 1.5 ml Eppendorf tubes. One part of ice-cold TCA (250 µl) was added to four parts (1 ml) of protein sample. The samples were vortexed and incubated on ice for 20-30 min. Subsequently, the samples were

centrifuged at 10000 g for 20 min at 4°C. Taking care not to disrupt the pellet, the supernatant was aspirated and the pellets were washed with 500 µl ice-cold 10% TCA. After vortexing and centrifugation at top speed at 4°C for 20 min, the pellets were further washed with 1 ml ice-cold acetone (stored at -20°C) to residual TCA. Then, the samples were vortexed and centrifuged at top speed at 4°C for 30 min. Afterwards, the supernatant was carefully removed and the samples were pooled for the last wash with acetone. Before dissolving the samples in 25 µl 1% SDS buffer, the pellets were left on ice to dry.

Protein Quantification and Normalization

To determine the concentration of the protein samples, the Pierce BCA Protein assay kit was used. Initially, the working reagent was prepared with 50 parts of BCA reagent A and 1 part of BCA reagent B. The diluted BSA standards were prepared as described by the manufacturer protocol (0, 25, 125, 250, 500, 750, 1000, 1500, 2000 µg/ml BSA concentration). Using the microplate procedure (plate), 25 µl per well of each prepared BSA standard solution was transferred in triplicates. Furthermore, 2 ml of the 1% SDS buffer was added per well of the BSA standards. Then, 2 µl of TCA precipitated samples were added per well in triplicates. 200 µl of the working reagent were then added to each well. The plate was covered and shaken for 1 min. Subsequently, the samples were incubated at 37°C for 30 min. After cooling the samples down to room temperature, the samples were measured using TECAN (562 nm absorbance). The average absorbance of the blank standard measurements was subtracted from all other values and a standard curve was established to determine the concentration of the samples of interest. Dilutions were made of the samples to yield an equal amount of protein across the samples.

Sample Preparation of Secretome Samples

The samples were further processed by Mandy Rettel, Proteomics Core Facility, EMBL Heidelberg, who also carried out the data acquisition and MS Database search. For the sample preparation including SP3 and TMT labeling, the reduction of disulphite bridges in cysteine containing proteins was first performed using 10 mM dithiothreitol in 50mM HEPES (pH 8.5) at 56°C for 30 min. The reduced cysteines were then alkylated with 20 mM 2-chloracetamide in 50 mM HEPES (pH 8.5) for 30 min at room temperature, in the dark. Sample preparation was conducted using the SP3 protocol

(Hughes et al., 2014, 2019) and addition of trypsin (sequencing grade, Promega) in an enzyme to protein ratio of 1:50. The digestion was incubated over night at 37°C. On the subsequent day, the peptide recovery was perforemd in a two-step collection. First, the supernatant of the overnight digestion was transferred on magnet to a new tube (first elution). Followed by a wash of beads in 50 mM HEPES buffer, the supernatant (second elution) is combined on magnet with the first elution in the new tube.
Peptide labelling was done using TMT11plex (Werner et al., 2014) Isobaric Label Reagent (ThermoFisher) following the instructions of the manufacturer. For this, samples were combined and cleaned using an OASIS® HLB µElution Plate (Walters). Offline high pH reverse phase fractionation was performed out using an Agilent 1200 Infinity high-performance liquid chromatography system, equipped with a Gemini C18 column (3 µm, 110 Å, 100 x 1.0 mm, Phenomenex) (Reichel et al., 2016).

Secretome Data Acquisition

An UltiMate 3000 RSLC nano LC system (Dionex) fitted with a trapping cartridge (µ-Precolumn C18 PepMap 100, 5 µm, 300 µm i.d. x 5 mm, 100 Å) and an analytical column (nanoEase™ M/Z HSS T3 column 75 µm x 250 mm C18, 1.8 µm, 100 Å, Waters). Trapping was carried out with a constant flow of 0.05% trifluoroacetic acid at 30 µl/min onto the trapping column for 6 min. Afterwards, peptides were eluted via the analytical column with a consistent flow of solvent A (0.1% formic acid, 3% DMSO in water) at 0.3 µl/min and increasing percentage of solvent B (0.1% formic acid, 3% DMSO in acetonitrile). The outlet of the analytical column was directly coupled to a Fusion Lumos (Thermo) mass spectrometer using the Nanospray Flex™ ion source in positive ion mode.
The peptides were introduced into the Fusion Lumos via a Pico-Tip Emitter 360 µm OD x 20 µm ID; 10 µm tip (CoAnn Technologies). For this, an applied spray voltage of 2.4 kV was used. Theerfore, capillary temperature was set at 275°C. Full mass scan was acquired with mass range 375-1500 m/z in profile mode in the orbitrap with resolution of 120000. The filling time was set at maximum of 50 ms with a limitation of 4 x 105 ions. Data dependent acquisition (DDA) involved quadrupole isolation at 0.7 m/z, with the resolution of the Orbitrap set to 30000, a fill time of 94 ms and a limitation of 1 x 105 ions. A normalized collision energy of 36 was applied and MS2 data was acquired in profile mode.

Secretome MS Database Search

First, raw files were converted to mzML format using MSConvert from Proteowizard (version 3.0.22129), and peak picking from the vendor algorithm. Afterweards, files were searched with MSFragger v3.7 (Kong et al., 2017) in Fragpipe v19.1 against the Swissprot Homo sapiens database (20,594 entries), which also contained common contaminants and reversed sequences. The standard settings of the Fragpipe TMT11 workflow were used and modifications such as Carbamidomethyl (C) and TMT11 (K) (fixed modification), Acetyl (Protein N-term), Oxidation (M), and TMT11 (N-term) (variable modifications) were included into the search parameters. For the full scan (MS1) a mass error tolerance of 10 ppm and for MS/MS (MS2) spectra of 0.02 Da was set. Furthermore, just like mentioned above, parameters including Trypsin (as protease with max. two missed cleavages) and a min. peptide length of 7 amino acids were set. The false discovery rate on peptide and protein level was set to 0.01.

Analysis of Whole Proteome and Phosphoproteome

The initial analysis of MS data was performed by Frank Stein from the Proteomics Core Facility, EMBL Heidelberg. Further analysis was supported by Mira Burtscher and Martin Garrido-Rodriguez. Initially, raw data files of FragPipe (Kong et al., 2017) were processed using R. Peptide spectral matches (PSMs) with a phosphorylation probabilty > 0.75 were taken for further analysis. For the whole proteome, only proteins with 2 unique peptides were taken for further analysis. Naming of the phosphopeptides includes * behind phosphorylated amino acids and the numbers 1-3 indicating the number of phosphorylation sites. The names are precided with gene symbols (translated from uniprot identifiers to make them comparable to transcriptomics data). Gene symbols, peptide sequence (with *) and number of phosphorylation sites are concatenated to generate a unique phosphopeptide ID. For all PSMs sharing the same phosphopeptide ID all raw TMT reporter ion intensities were summed. Phosphorylation intensities were further normalized to the input data (whole proteome), per replicate and condition. First, ‘vsn’ normalisation and imputation of missing values using the ‘knn’ method were performed followed by batch effect removal (between biological replicates/different MS runs) using ‘removeBatchEffects’ from the limma R package (Ritchie et al., 2015). Differential expression analysis was conducted using ‘lmFit’ of the limma R package. This was done separately for the input (whole proteome), phospho- and the normalized phosphoproteomics (phosphoproteomics normalize to

input) data. Hits were described as differentially expressed genes with an absolute log2 fold-change > log(1.5) and an adjusted p-value < 0.05.

Analysis of Secretome

Initial analysis of MS data was performed by Frank Stein from the proteomics core facility, EMBL Heidelberg. For the secretomics data, Frank Stein processed the raw output files of FragPipe (Kong et al., 2017) using the R programming language (ISBN 3-900051-07-0). Contaminants were filtered out (except for albumin, which was later removed by me). Translation of uniprot identifiers to gene symbols was performed to keep the naming consistent with the transcriptomics data. Only proteins with at least 2 unique peptides and that showed up in at least two out of three biological replicates were considered for further analysis. This resulted in 2188 proteins quantified. Further analysis was conducted by me with support of Mira Burtscher and initial help from Martin Garrido-Rodriguez. The log2 transformed raw TMT reporter ion intensities were 'vsn' normalized using the vsn R package (Huber et al., 2002) and missing values were imputed using 'knn'. This was further followed up by batch effect correction using the limma package function 'removeBatchEffects' (Ritchie et al., 2015). Differential expression analysis was performed using the limma package, and differentially expressed genes, were considered as hits if they had an adjusted p-value < 0.05 and absolute log2 fold-change of > log(1.5). Furthermore, the data were filtered for genes that are secreted or part of the extracellular matix. Hence, the data were filtered based on the MSigDB gene sets NABA_MATRISOME (M5889) and NABA_MATRISOME_ASSCIATED (M5885) (not performed for the principal component analysis). Additionally, I included following proteins that were neither part of the Naba matrisome nor the associated factors but were annotated as secreted or secreted via extracellular vesicles (Information provided by Uniprot): MARCKS, CDH11, NRP2, SDC4, TFPI, CPA4, and CACHD1. For visualisation of data, I used the statistical language R.

3.2.6. Function Analysis of Omics Data

Using MSigDB and decoupleR (Badia-i-Mompel et al., 2022), pathway enrichment analysis was performed, supported by Mira Burtscher. First, the species was specified as "Homo Sapiens" and the MSigDB "HALLMARK" gene set was chosen. All

differentially expressed genes, not just hits, were used as input. Pathway activity was inferred using the ‘wmean’ function of the decoupleR package.

3.2.7. Heatmaps of Omics Hits

Heatmaps of hits, defined as adjusted p-value < 0.05 and absolute log2 fold-change > log(1.5) for proteomics and secretomics while log(2) was used as a threshold for the transcriptomics data, were generated using the ‘Heatmap’ function of the ComplexHeatmap R package. More specifically, I selected hits that were shared in at least two of the three omics modalities. Since phosphoproteomics data show phosphorylation and not necessarily abundance of proteins, I generated a separate heatmap of the top 50 deregulated phosphopeptides.

3.2.8. Comparison of Omics Data to Single-Cell RNA Seq Data

Supplementary information of Kuppe et al. (2021) was used to link single-cell RNA sequencing data of CKD and healthy individuals to the acquired data set of this thesis. Thererfore, a supplementary table was used which included the specificity scores of analysed genes calculated by Kuppe et al., for each cell type (41586_2020_2941_MOESM4_ESM.xlsx, Human_PDGFBplus_Level_2_Specifi sheet). The top 7% of genes specific for myofibroblasts were filtered based on a specificity score threshold that we defined as > 0.2. This resulted in 1650 genes, that were then compared to the dataset of this thesis (Kuppe et al., 2021).

3.2.9. Footprint Analysis for Inference of TF and Kinase/Phosphatase Activity

For inference of TF and kinase/phosphatase activities, the decoupleR package (Badia-i-Mompel et al., 2022) was used in combination with the DOROTHEA database (Garcia-Alonso et al., 2019) to obtain TF-target interactions (with the confidence levels of A, B or C). The phosphosite-enzyme interactions were obtained from the OminpathR package (Türei et al., 2016). All differentially expressed transcripts and phosphosites were used as input for the respective activity inference. For this, the ‘wmean’ fucntion was used (calculating the mean of log2 fold-changes of i.e. targets of a TF and adding weight based on whether the interaction is activating or inhibiting).

3.2.10. Clustering and Correlation of Omics Data

Clustering

Clustering of omics modalities was performed on the hits (as defined above based, on the log2 fold and adjusted p-value thresholds) the ‘cclust’ function of the cclust R package. For the clustering the ‘euclidean’ distance was chosen, where cluster means are the centroids. The algorithm was defined as ‘neuralgas’.

Correlation

For correlation of omics modalities (excluding phosphoproteomics), the top 10% of differentially expressed features per omics modality were selected. Using the Hmisc R package’s function ‘rcorr’, a correlation matrix with Pearson correlation coefficients was generated and plotted using the ‘corrplot’ function from the corrplot R package.

3.2.11. Network Modeling

Following steps were performed by Mira Burtscher with input from me. For inference of signaling processes, the CARNIVAL algorithm combined upstream perturbation information (e.g. TGF-β stimulation) with measurements of downstream factors (e.g. TF activities) and a causal prior knowledge network (PKN) (A. Liu et al., 2019). For kinase/phosphatase and TF activity estimation, they are calculated using a footprinting method (Dugourd & Saez-Rodriguez, 2019). COSMOS extends CARNIVAL to multimodal datasets (Türei et al., 2016). All the network analyses were performed using the CARNIVAL method (ref to Liu et al.) as implemented in CORNETO (https://github.com/saezlab/corneto).

Prior Knowledge Parsing

The PKN contains the signed and directed protein-protein interactions (PPI) (node A leads to in-/activation of node B) that are then used to model signaling interactions between measurements and e.g. TGF-β input by minimizing an objective function in an integer linear programming (ILP) optimization.
In this study, we used a prior knowledge network which was provided as part of the OmnipathR package via the ‘import_all_interactions’ function. This initial PKN was filtered for trusted resources and interactions with a valid consensus signal.

Additionally, the core of the TGF-β signaling module was fixed to guide the optimisation (enforce TGF-β to SMAD1-5, support TGFβ to MAPK1, MAPK14, AKT1 and PI3K) following Park and Yoo (2022) (C. H. Park & Yoo, 2022). The final PKN (42 k interactions) was obtained by filtering for nodes which were measured in one of the omics modalities obtained as part of this study.

Input Node Filtering

As input nodes for the optimisation, TGF-β was used as upstream perturbation as well as the top transcription factors (maximum absolute enrichment score > 3, p-value < 0.03) and kinases/phosphatases (maximum absolute enrichment score > 3, p-value < 0.03) from the activity estimation analysis and the hits of the secretomics dataset.
Secreted proteins were filtered for 'M5889' and 'M5885' MSIGDB terms to ensure that they are indeed related to secretion processes. This list was manually extended with COL1A1, COL5A1, SERPINE2, SPARC, ITGB1, VIM, JUP, ACTA1, HSPG2, LOXL2, TNC, TGFBI, IGFBP3, IGFBP7, LTBP2, TAGLN, CCN2, LRATD2, MRC2, FN1, FBN1, BGN, ADGRG1, MMP2, ITGA11, AMIGO2, ADAM12, CPA4, DCDC2, PLEKHG4 and ITGB1BP1 which were significantly deregulated in the secretomics experiment and reported to be secreted before. The number of input nodes has been chosen to find an acceptable compromise of computational cost and information content for the network modeling step.

Optimization

In total, we computed four subnetworks inferred based on different timepoints and input modalities as shown in Figure *41* to account for the time-resolved nature of the obtained data.
To start the optimisation, TGF-β was taken as upstream input node of early affected enzymes (36 TFs, 22 kinases and phosphatases). This network was combined with a second network connecting the set of early affected enzymes (36 TFs, 22 kinases and phosphatases) and early secretomics hits (54 proteins). This combination represents early signaling events stimulated by TGF-β treatment which then lead to altered protein secretion up to 24 hours (early network). This ealry network contained 127 nodes and 156 edges.
To model assumed autostimulatory function of secreted factors, we modelled later signaling processes by inferring a first network between secreted proteins included in

the early network (43 proteins) and enzymes affected at a later timepoint (between 24 h and 96 h, 59 proteins). As last step, this network was combined with a fourth subnetwork connecting later affected enzymes with late secretomics hits (72 proteins), resulting in 160 nodes and 202 edges.
All in all, these combined networks result in one early and one late network describing the response of human derived resident PDGFRβ^{+} mesenchymal cells to TGF-β in a multimodal manner.

Clustering, Cluster Comparison and Cluster Enrichment

To obtain a more detailed understanding of biological processes at the network level, a clustering analysis was performed. All edges identified for the early and late network, respectively, were extracted and a fast and greedy clustering algorithm on the resulting graph was applied. Thereafter, a cluster-specific enrichment analysis, as described before, was carried out and all nodes related to the identified processes were extracted to gain mechanistic insights.
We performed a cluster-specific enrichment analysis using ReactomePA package (Gillespie et al., 2022). All nodes per cluster of the result network were used as input. As background, all proteins were included after differential expression analysis for the network-level enrichments. In the ‘enrichPathway’ function, the organism was specified as "Human", minGSSize was set to five and maxGSSize was determined as 500. For identification of significant pathways, the result for pathways were enriched with at least two detected genes and with a significance cutoff of 0.01 after multiple testing correction.

4. Results

4.1. Rationale and Characterization of Cellular Fibrosis Model

4.1.1. Selection of Cellular Model

As myofibroblasts are one of the major drivers of kidney fibrosis, I decided to dissect how mesenchymal cells, like pericytes and fibroblasts, differentiate into myofibroblasts. Therefore, I obtained human kidney-derived PDGFRβ^{+} cells, comprising different mesenchymal cells, from the lab of Prof. Dr. med. Rafael Kramann (Kuppe et al., 2021). To study a profibrotic response, I decided to stimulate the cells with TGF-β, a master regulator of fibrosis (Edeling et al., 2016; Meng et al., 2016; Sziksz et al., 2015; Van Linthout et al., 2014; Zhou et al., 2020).
Furthermore, a simple cellular model has the advantage, compared to patient-derived material, that it is not limited to the availability of human tissue, which is often a limiting factor. Additionally, a cellular model can be easily perturbed and therefore different compounds could be tested.

4.1.2. Model Characteristics

In the following section, I outlined initial experiments in which I characterized the cell lines response to TGF-β stimulation. Key objectives included observing the alterations in cell morphology, evaluating changes in expression of fibrotic genes at the mRNA level and assessing the deposition of extracellular matrix over time.

Extracellular Matrix Dynamics

Assuming that activated myofibroblasts are the main source of ECM, I examined whether the ECM deposition of the used cell line would be altered upon stimulation with TGF-β. In our lab, we are accelerating the ECM deposition by adding macromolecular crowding to the culture media, a mixture of neutral 70 and 400 kDa Ficoll (Chen et al., 2009; Coentro et al., 2021; Khan et al., 2023). Initially, I assessed the changes in ECM deposition using antibody staining of extracellular Collagen 1 (Col I) (*Figure 4*).

Based on the findings of previous experiments conducted in the lab with normal human lung fibroblasts (Khan et al., 2023), we expected a significant difference in ECM deposition of TGF-β treated cells compared to control cells after 48 h. My results indeed showed a significant difference between TGF-β treated and control cells on extracellular Col I synthesis (*Figure 4d*), affirming the suitability of the used cellular system to study fibrosis.

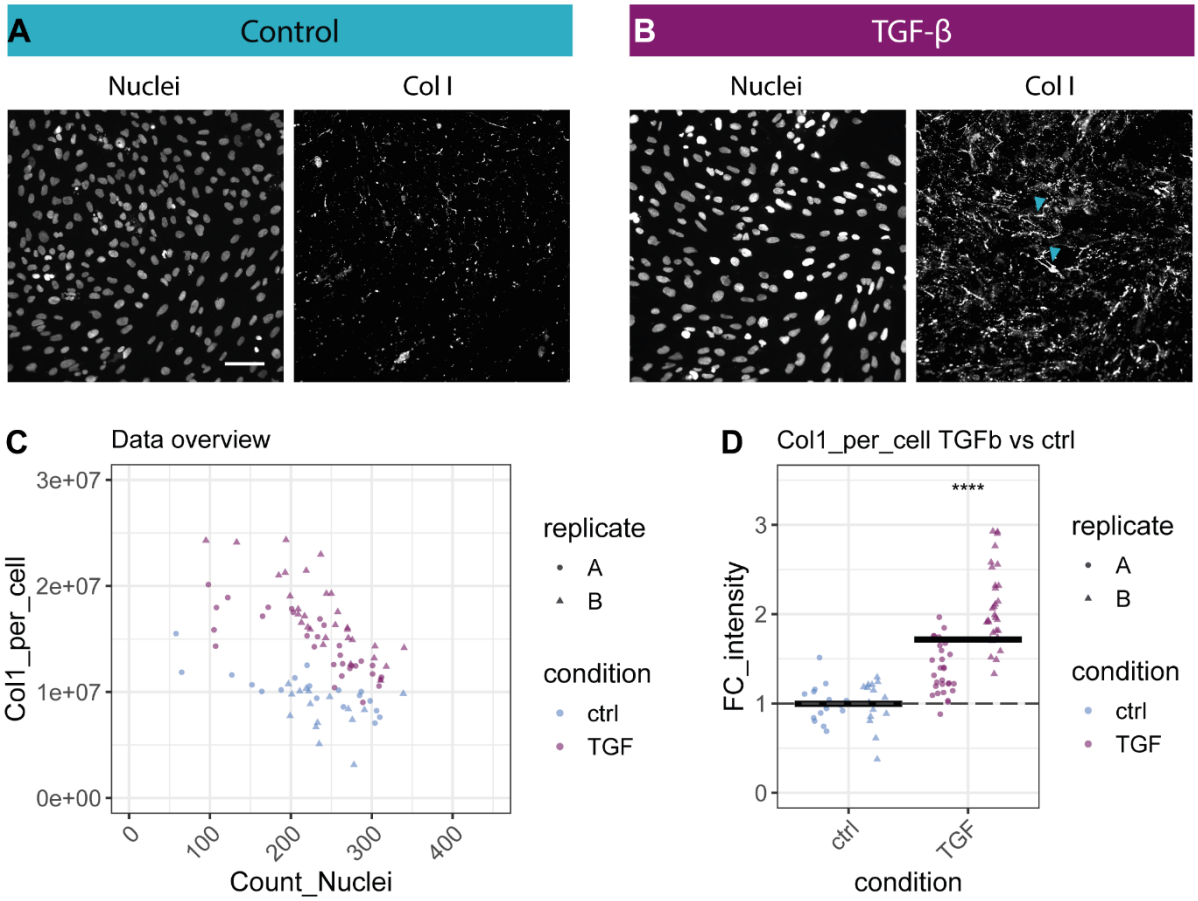

Figure 4 Immunofluorescence images of extracellular Col I at 48 h post TGF-β treatment and quantification. The day after cell seeding, media was changed to either DMEM + Ficoll + ascorbic acid + 0.5 FBS (control) or + TGF-β for the treated condition. **A, B)** Widefield microscopy images were acquired with the Molecular Devices IXM. Immunofluorescence staining of extracellular Col I and nuclei (Hoechst 33342) are shown 48 h after the medium change for the **A)** control and **B)** TGF-β treated condition. Blue arrows highlight fibrillar collagen. Scale bar = 100 µm. **C)** Data overview of immunofluorescence signal (Col I staining) after background subtraction and normalization to the cell number (Count_Nuclei) resulting in the values Col1_per_cell shown on the y-axis versus the Count_Nuclei on the x-axis. Each dot stands for individual images of the biological replicates (plates), blue dots depict control samples and purple dots TGF-β treated samples. Here, images with less than 50 nuclei were excluded from the analysis. **D)** The y-axis shows the FC_intensity values, calculated as fluorescence intensity (Col1_per_cell) of the TGF-β treated samples divided by the mean of the control treated samples. The median values for each sample are shown by the bars, dashed lines stand for the control sample used as reference in the t-test. TGF-β treatment resulted in increased Col I deposition.

Morphological Changes

The transition of mesenchymal cells to activated myofibroblasts entails morphological alterations. Using microscopy read-outs, I therefore tested whether the cells showed any morphological changes and expression of typical markers like alpha-smooth muscle actin (α-SMA, ACTA2), vimentin or an increased production of ECM (Hinz & Lagares, 2021; Walker et al., 2019).

Initially, I tested different fixation methods (PFA, methanol) and staining techniques. Even without TGF-β treatment, some of the PDGFRβ$^+$ cells, consisting of several mesenchymal cell types, showed a more intense staining for α-SMA. Nevertheless, the overall expression of α-SMA remained low, with no discernible increase upon TGF-β stimulation (data not shown). Additionally, while the cells were vimentin positive, I did not see an enhancement in fluorescent signal upon TGF-β treatment. Nevertheless, the vimentin fibers were aligned and more pronounced after 48 h of TGF-β stimulation in contrast to the control condition with shorter fibers (Figure *5*).

Moreover, visualization of the actin cytoskeleton, using phalloidin, revealed an increase in actin stress fibers and notable changes in cell morphology post TGF-β treatment. Upon TGF-β treatment for 48 h, the cells showed more pronounced fibers that were locally and globally organized when compared to the control where the fibers were more unorganized on the multicell scale (Figure *5*).

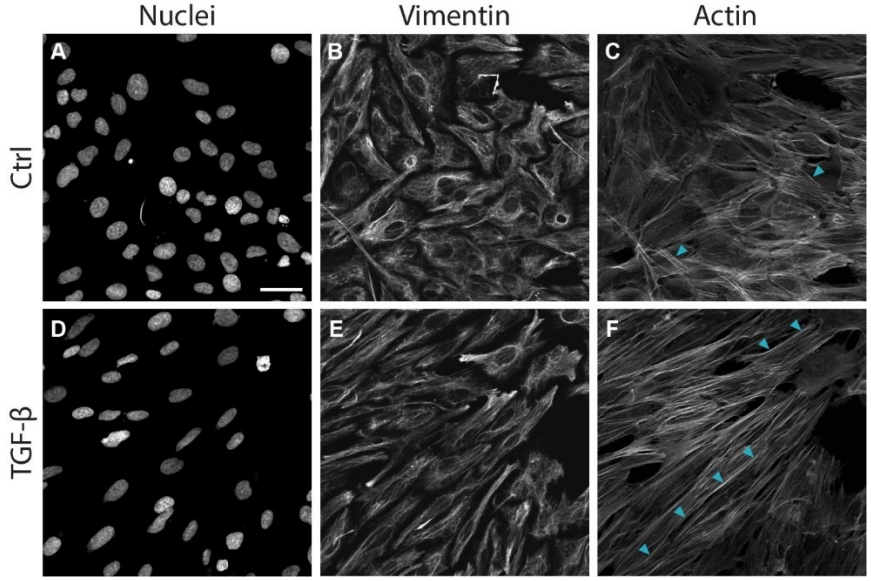

Figure 5 Morphological changes following 48 h of TGF-β treatment. Sum Z-projections of confocal images acquired at the Zeiss LSM 900 depict morphological alternations. Cells were seeded on day 0, media changed 24 h after to DMEM + 0.5% FBS + Ficoll for the control **(A-C)**, and the same media supplemented with 10 ng/ml TGF-β for the treatment condition **(D-F)**. Shown are images taken of the **A, D**) nuclei stained with Hoechst 33342, **B, E)** vimentin staining and **C, F)** phalloidin staining for actin filaments (highlighted by blue arrows). Scale bar: 50 µm. TGF-β treatment showed no effect on vimentin expression or structure, but led to more stress fibers.

Fibrotic Marker Gene Expression

To elucidate the transcriptional regulation underlying TGF-β-induced fibrosis, I used RT-qPCR to assess marker genes like COL1A1, ACTA2, VIM, FN1 (F. C. Chang et al., 2012; Edeling et al., 2016; Kendall & Feghali-bostwick, 2014; Kramann & Humphreys, 2014; Lotz-Jenne et al., 2016; Sziksz et al., 2015; Zeisberg & Kalluri, 2015).

Examination of different time points was conducted to determine the optimal time frame for analysis. Figure 6 illustrates the mRNA expression changes of TGF-β treated samples compared to control at each time point for individual genes. At 8 h, there was no significant change of any of the tested genes. The mRNA expression of VIM did not change over time while there was a significant increase of COL1A1 mRNA expression upon stimulation with TGF-β after 24 h. Moreover, ACTA2 mRNA expression levels increased, however, not significantly. These data collectively suggested a profibrotic response of the cells to TGF-β.

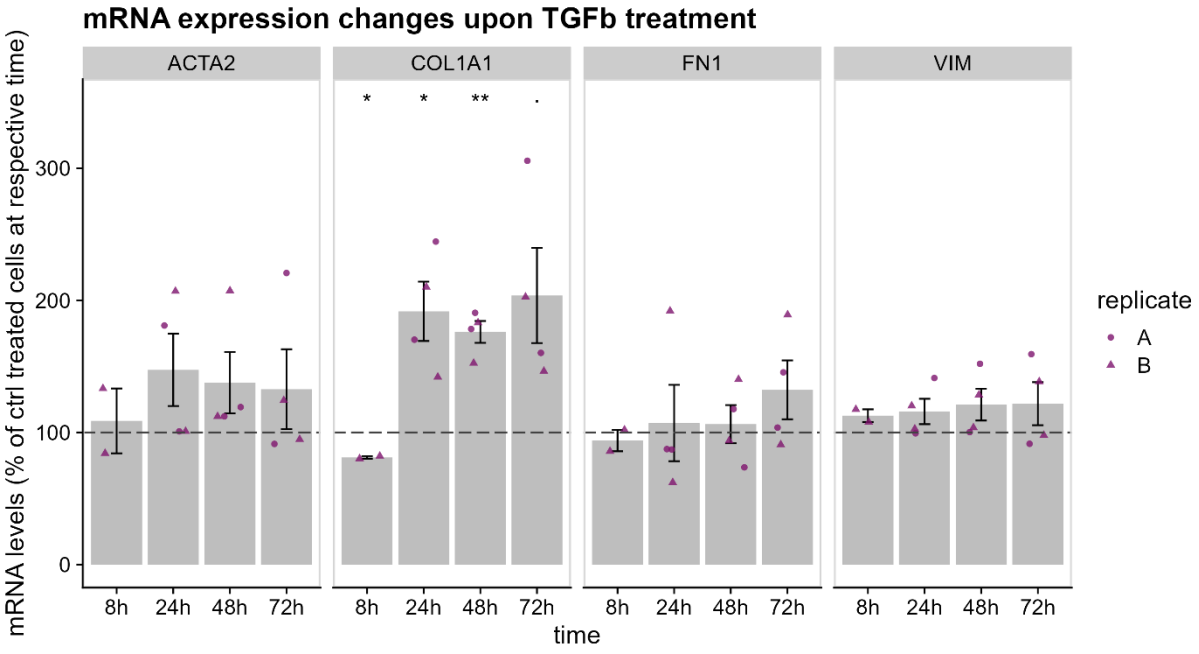

Figure 6 mRNA expression changes over time following TGF-β treatment. RT-qPCR quantification of fibrotic marker genes (ACTA2, COL1A1, FN1 and VIM) at various time points (8, 24, 48, 72 h) post TGF-β stimulation. Data were acquired and normalized as described in the methods section 3.2.4. In short, the CT values were normalized to the respective GAPDH CT values. Treated samples were then further normalize to the non-treated samples for each time point respectively. The percentage was calculated using the $2^{-\Delta\Delta CT}$ times 100 to obtain the percentages. The time post TGF-β treatment is shown on the x-axis, the mRNA expression levels as described above, on the y-axis. Bars denote the average values of two independent biological experiments with two technical replicates, while individual values are depicted as dots. Control treated samples, were normalized to 100% indicated by the dashed line, to which a t-test was performed.

4.1.3. Establishing Experimental Design and Time Points

Upon recognizing the variability in initial experiments regarding ECM deposition, I directed my efforts towards optimizing experimental parameters for consistency and reproducibility. Moreover, determining the most relevant time points for investigating the underlying mechanisms of the cellular system to a profibrotic stimulus became more prominent.

Furthermore, leveraging the cells for multi-omics analysis was deemed valuable to systematically explore mRNA and protein-level changes, as well as alterations in secreted factors, as crucial extracellular signaling molecules. However, to use secretomics as a read-out, I had to eliminate the inference of proteins from the fetal bovine serum (FBS) in the media. FBS, containing various growth factors, could potentially influence signaling, thereby elevating background signals. Consequently, an adjusted protocol without FBS during the TGF-β treatment period was tested. In the following section, I will describe the tested variables like cell number, media change after 24 h and 48 h and different time points, in more detail.

TGF-β Response Curve

Additionally, I assessing whether a different concentration of TGF-β used would be more suitable for the experimental set-up. Therefore, I treated cells with different concentrations of TGF-β for 48 h and compared the resulting ECM deposition, through immunofluorescence staining, to the control. All of the tested concentrations yielded a significantly higher expression of extracellular Col I (Figure 7). After consultation with our collaborators, that use the cells as well, we decided to stick with 10 ng/ml TGF-β.

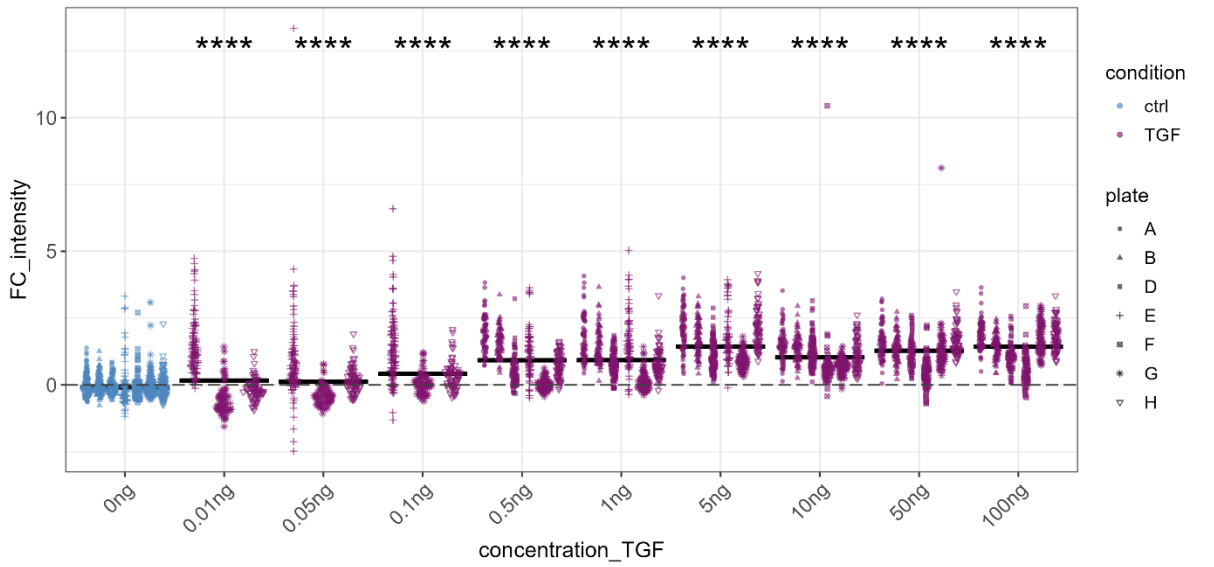

Figure 7 Quantification of immunofluorescence staining of the extracellular Col I for different concentrations (ng/ml) of TGF-β. Quantification of fluorescence intensity (FC_intensity), representing the immunofluorescence staining of the extracellular Col I, expressed by cells 48 h at the control condition (blue) or after TGF-β treatment (purple). Each dot stands for individual images of the biological replicates (plates). The x-axis shows the different concentrations of TGF-β tested. Images with less than 25 nuclei were excluded for the analysis. The y-axis shows the FC_intensity values, calculated as fluorescence intensity (after background correction and normalization to nuclei number) of the ctrl (0 ng) subtracted from all images. The result was further divided by the value of the ctrl (0 ng) condition. The median values for each sample are shown by the bars, dashed lines stand for the control sample used as reference in the t-test. All concentrations of TGF-β used yield a significant increase of extracellular deposited Col I.

Cell Number Determination

To establish optimal cell densities for subsequent experiments, different cell numbers were seeded per well of a 24-well plate. Subsequently, cells were treated at various time points, and fluorescence signal, indicative of extracellular Collagen 1 build-up, was analyzed following TGF-β treatment compared to control (Figure *8*).

In Figure *8*, it is shown that the Col I intensity generally increases with increasing cell number seeded, this is true for the control and the TGF-β treated conditions. However, when plotting the Col I per cell (normalized to the nuclei number), it became apparent that more cells did not produce more Collagen over all. Since cells were overgrowing in more densely seeded wells, there were also more autofluorescent cells. Based on these results and considering factors such as overgrowth and autofluorescence, the following cell numbers were fixed for each time point: 15k for 96 h, 20k for 72 h, 25k for 48 h, 50k for 24 h, and 75k for less than 24 h.

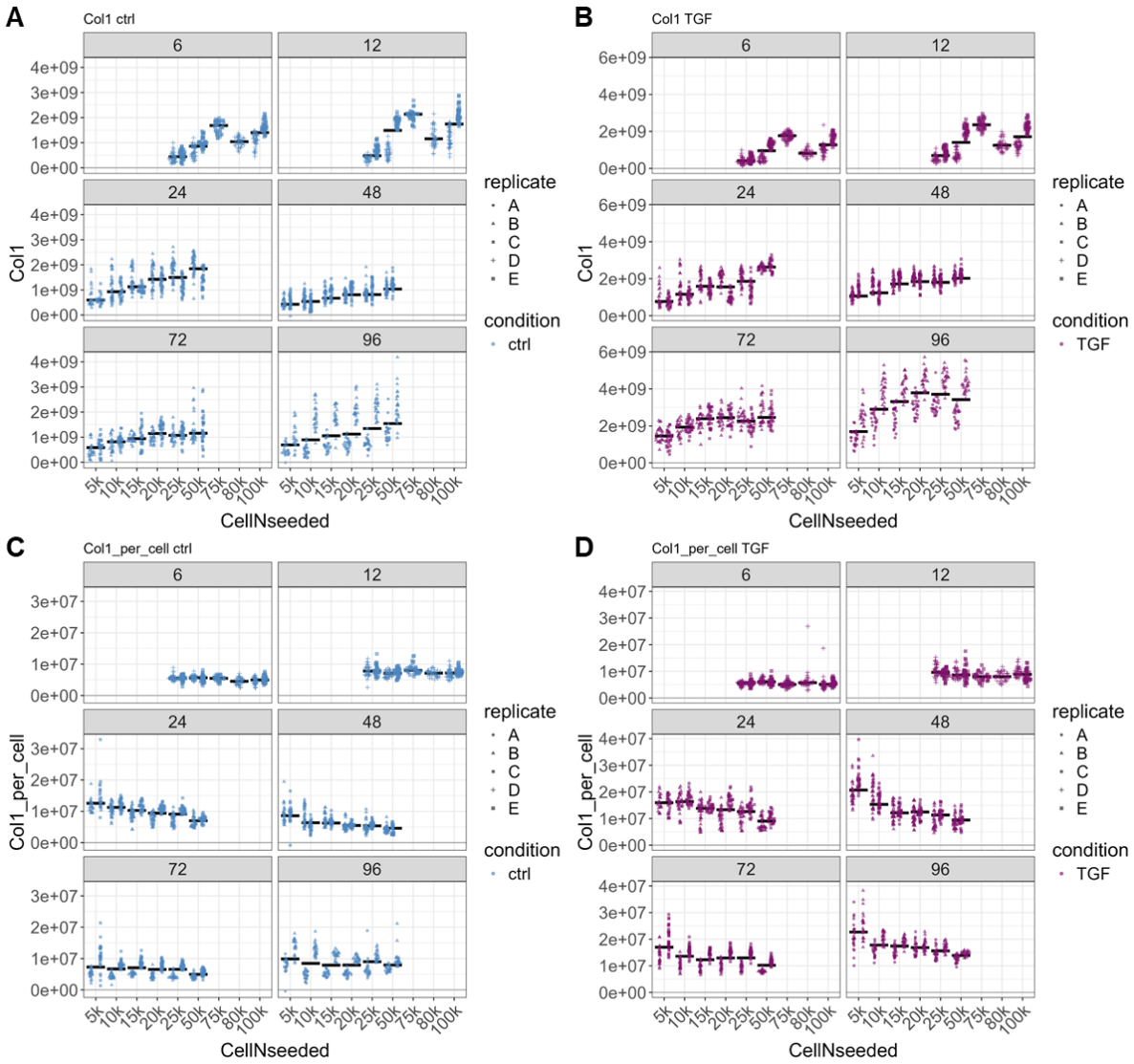

Figure 8 Change of immunofluorescence-stained extracellular Col I and Col I per cell across various nuclei numbers seeded and time points. Different cell numbers, depicted on the x-axis, were seeded for ± TGF-β treatment for 6, 12, 24, 48, 72 and 96 h. **A, B)** The y-axis shows the fluorescent signal after background correction for the extracellular Col I per image, for **A)** control (blue) and **B)** TGF-β treated (purple) samples at the different time points tested. **C, D)** On the y-axis, Col I signal normalized to the nuclei number is depicted for **C)** control (blue) and **D)** TGF-β treated (purple) samples at 6 – 96 h post treatment. While the fluorescence increases with more cells, this increase is not linear and results in less Col I produced per cell in wells with a denser cell population.

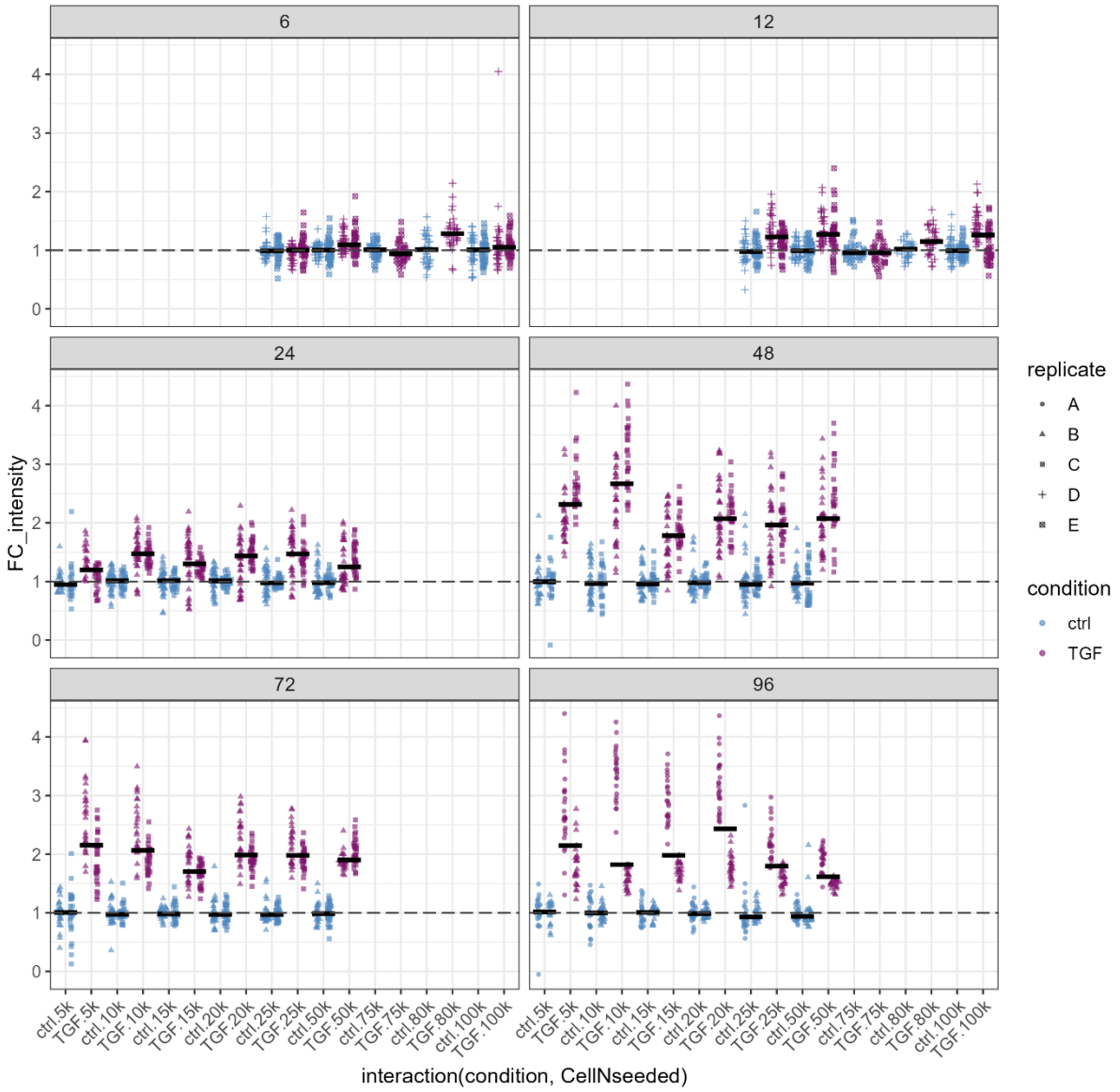

Figure 9 Analysis of extracellular Col I immunofluorescence-staining following TGF-β treatment at various cell numbers. The fluorescent intensity corresponding to the extracellular Col I per cell in TGF-β treated samples is normalized to the respective control conditions for each time point (6, 12, 24, 48, 72, and 96 h) and cell number. This means that after background correction, and exclusion of images with less than 30 nuclei, the average fluorescence intensity per cell of the control samples was subtracted from the TGF-β treated samples. The results were further divided by the average values of the control samples leading to the FC_intensity, shown on the y-axis. Cell number and treatment conditions (ctrl or TGF-β treated) are shown on the x-axis for each time point. The different shapes represent distinct biological replicates and the individual points depict the individual images taken for each condition (ctrl = blue, TGF-β = purple). Points from ctrl and TGF-β treated samples are shown in different shapes.

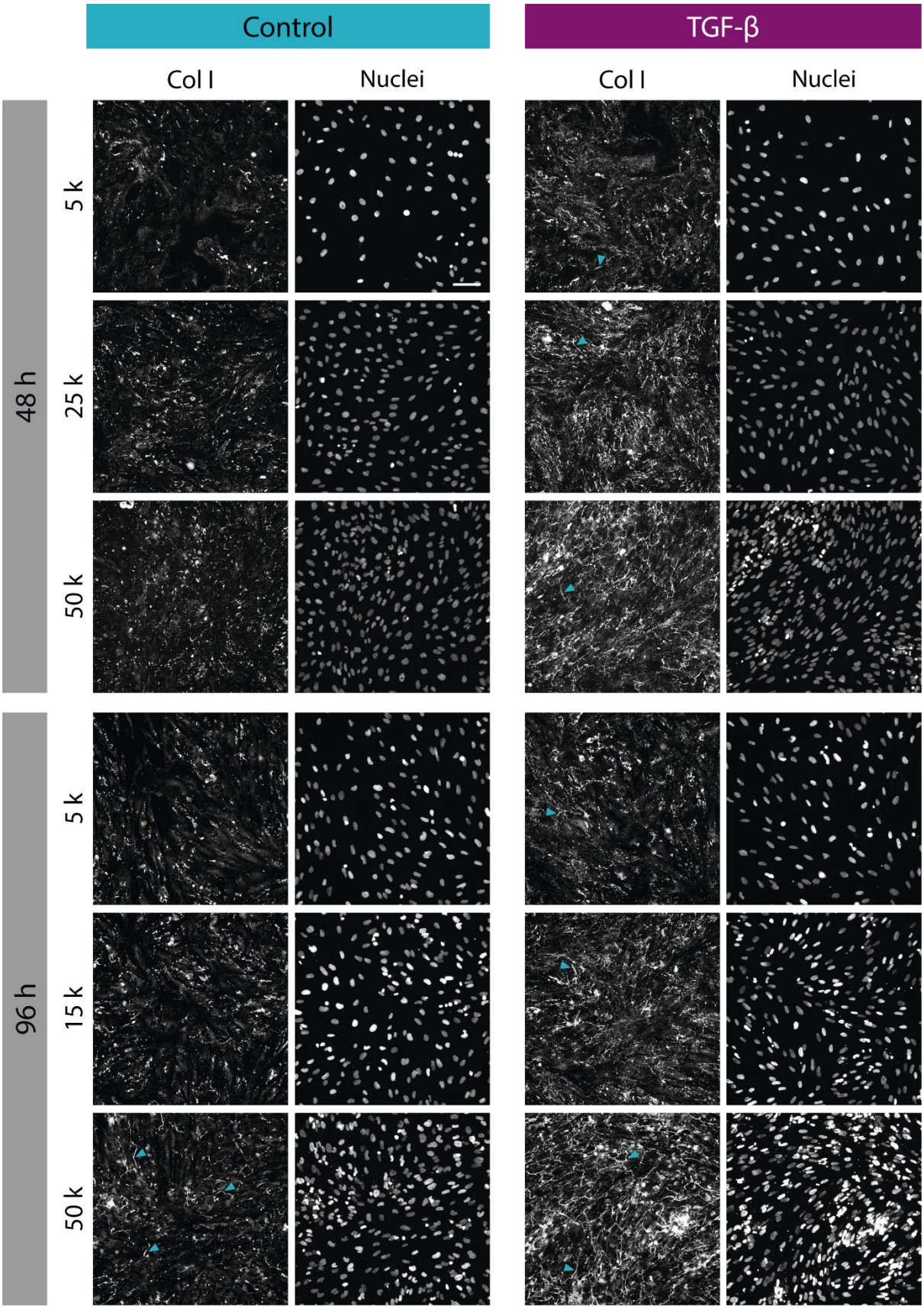

Figure 10 Immunofluorescence images of extracellular Col I at 48 and 98 h and different cell numbers seeded. Different cell numbers were seeded in 24 well plates. The following day, medium was changed either to DMEM + Ficoll + ascorbic acid (control) or + TGF-β for the treatment conditions. Staining of extracellular Col I and nuclei (Hoechst 33342) are shown for 48 h (5, 25, 50 k cells seeded) and 96 h (5, 15, 50 k cells seeded). Blue arrows highlight fibrillar collagen. Treatment with TGF-β

induced more Col I deposition with more Col I in conditions with higher cell numbers. Scale bar = 100 µm.

Fibrotic Response Over Time in Different Experimental Set-ups

Furthermore, alternative experimental set-ups were explored (**Figure *11***). Initially, I seeded different cell densities for the various time points and started treating all samples at the same time. Then I would harvest the samples accordingly. However, in the experimental set-up B, I seeded the same cell number for all of the samples at the same time. The day after seeding, I changed the medium of the samples with the longest treatment to medium containing molecular crowding agents and ascorbic acid for the control and the same media + TGF-β for the treatment. The media of all other conditions was changed to DMEM without FBS.

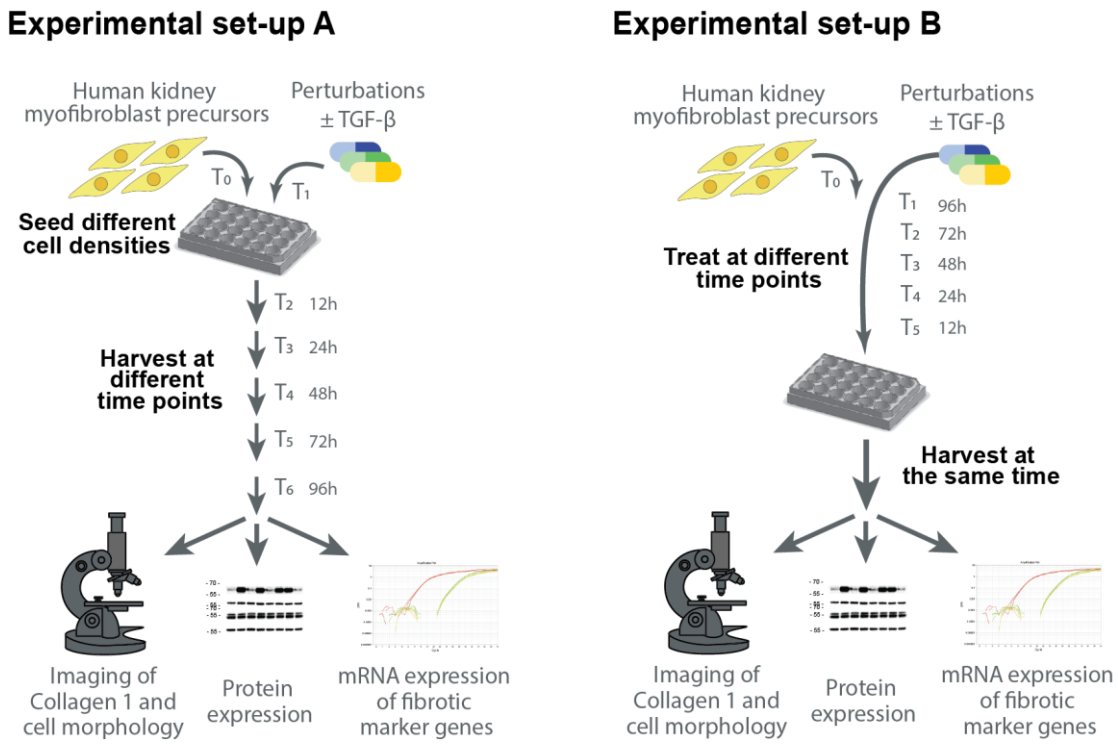

Figure 11 Illustration of experimental set-up A and B. In set-up A, the cells were seeded at the same time, however, at different cell densities according to the time of treatment. At the next day, the treatment ± TGF-β was started, and samples were harvested after 12, 24, 48, 72 and 96 h. In contrast to this, in the experimental set-up B, cells were seeded at the same density. On the next day, the treatment of the longest time point (96 h) was started, followed by treatment of the next time point (72 h) at the consecutive day, and so on. Samples for the different treatment conditions were then harvested all at the same day. Down-stream read-outs involved image acquisition and analysis of deposited Col I and morphological changes. In addition, the mRNA expression of fibrotic marker genes and relevant signaling proteins was assessed using RT-qPCR and western blotting.

Changes in Extracellular Collagen I

Both experimental set-ups revealed significant differences in extracellular Col I deposition as early as 12 hours post-TGF-β treatment, with deposition increasing over time, particularly evident in set-up B (t-test) (Figure *12*), where all samples underwent the same background correction. The advantage of set-up B was that the cells were seeded all at the same time and on the same plate which resulted in the same background subtraction of all samples and direct comparison between the time points.

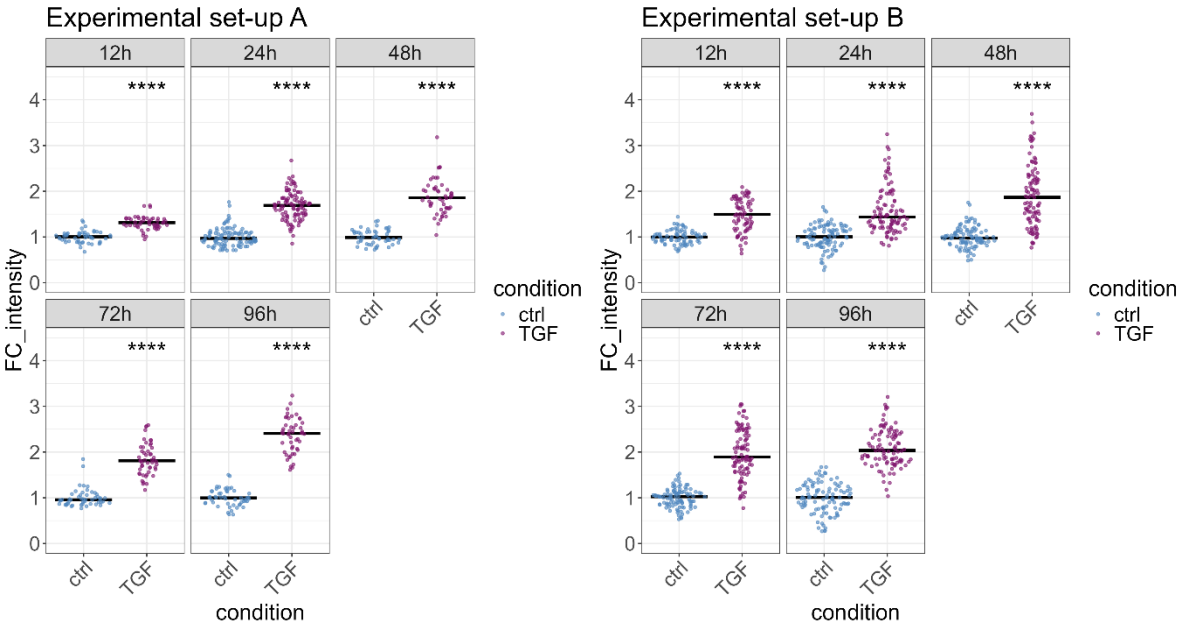

Figure 12 Analysis of extracellular Col I post-TGF-β treatment at various time points and different experimental set-ups. Immunofluorescence staining of extracellular Col I following treatment with TGF-β is compared and normalized to the respective control condition at designated time points (12, 24, 48, 72 and 96 h). Condition is indicated on the x-axis while the y-axis shows the FC_intensity (calculated as in the previous graph). Dots are color coded based on treatment (ctrl as blue, and TGF-β as purple). The bars indicate the median and individual data points denote the results from individual images taken. For set-up A, one biological replicate was tested, while set-up B involved two independent biological replicates. T-test was used to determine statistically significant differences between TGF-β and ctrl treated samples per time point.

Changes in mRNA Expression of Fibrotic Marker Genes

Furthermore, I analyzed the changes of fibrotic marker genes which varied over time in the two experimental set-ups. In Figure *13*, one can see that the expression of vimentin (VIM) did not change much in the TGF-β treated conditions compared to the controls at each time point or over time and there was no difference between the experimental set-up A and B. On the other hand, the mRNA expression of the other genes tested increased with TGF-β treatment. It was also shown that ACTA2 mRNA expression varied across the tested time points and in the two experimental set-ups.

Additionally, I want to note that the expression of COL5 increased in the experimental set-up A while it stayed constant over time in the set-up B.

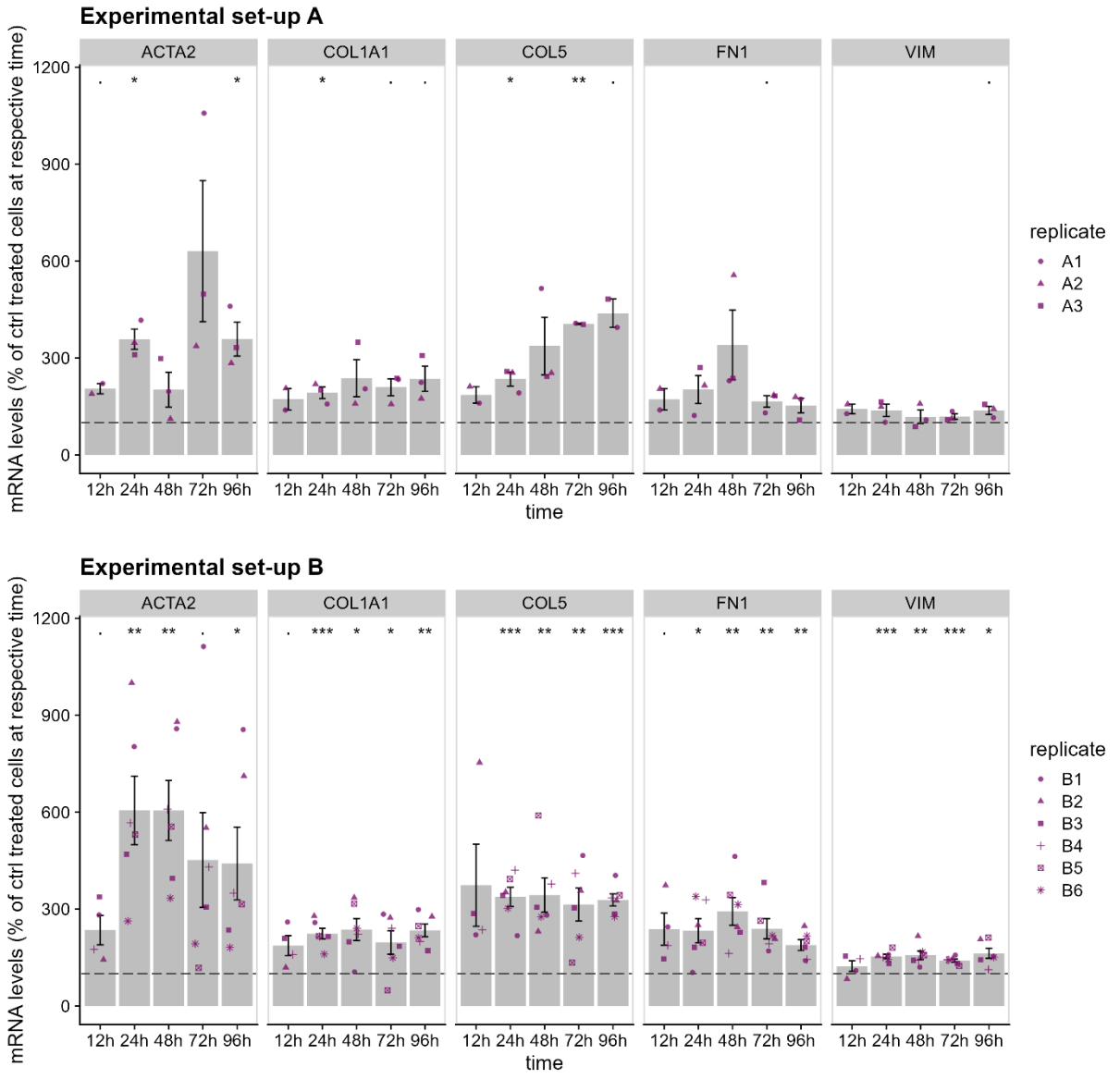

Figure 13 Comparison of fibrotic marker gene mRNA expression changes in different experimental set-ups. RT-qPCR quantification of fibrotic marker genes (ACTA2, COL1A1, COL5, FN1 and VIM) at various time points (12, 24, 48, 72 and 96 h) post TGF-β stimulation. Data were acquired and normalized as described in the methods section 3.2.4 **Figure 6**. The x-axis shows the time post TGF-β treatment, the y-axis the mRNA expression. Bars denote the average values of independent biological replicates with two technical replicates, while individual values are depicted as dots. Control treated samples, were normalized to 100 % indicated by the dashed line, on which a t-test was performed. The upper part of the figure shows the data of three biological replicates (A1-A3) for the experimental set-up A. The lower part of this figure summarizes the result of experiments with the set-up B, including six independent biological replicates (B1-B6).

Morphological Differences

I was intrigued to determine the onset of morphological changes induced by the TGF-β stimulation and wanted to know whether the effect is sustained over time. Therefore, I investigated these alterations using phalloidin staining to visualize the cells

cytoskeleton. In the figure below (Figure *14*), images taken from the experimental set-up B are shown as there all the samples were on the same plate and could be acquired with the same conditions. While the cytoskeleton appeared relatively relaxed in the control condition, cells treated with TGF-β showed an increase of stress fibers over time, this was definitely prominent after 48 h, whereas at earlier time points the inspection of images showed a heterogenous distribution of cells with stress fibers.

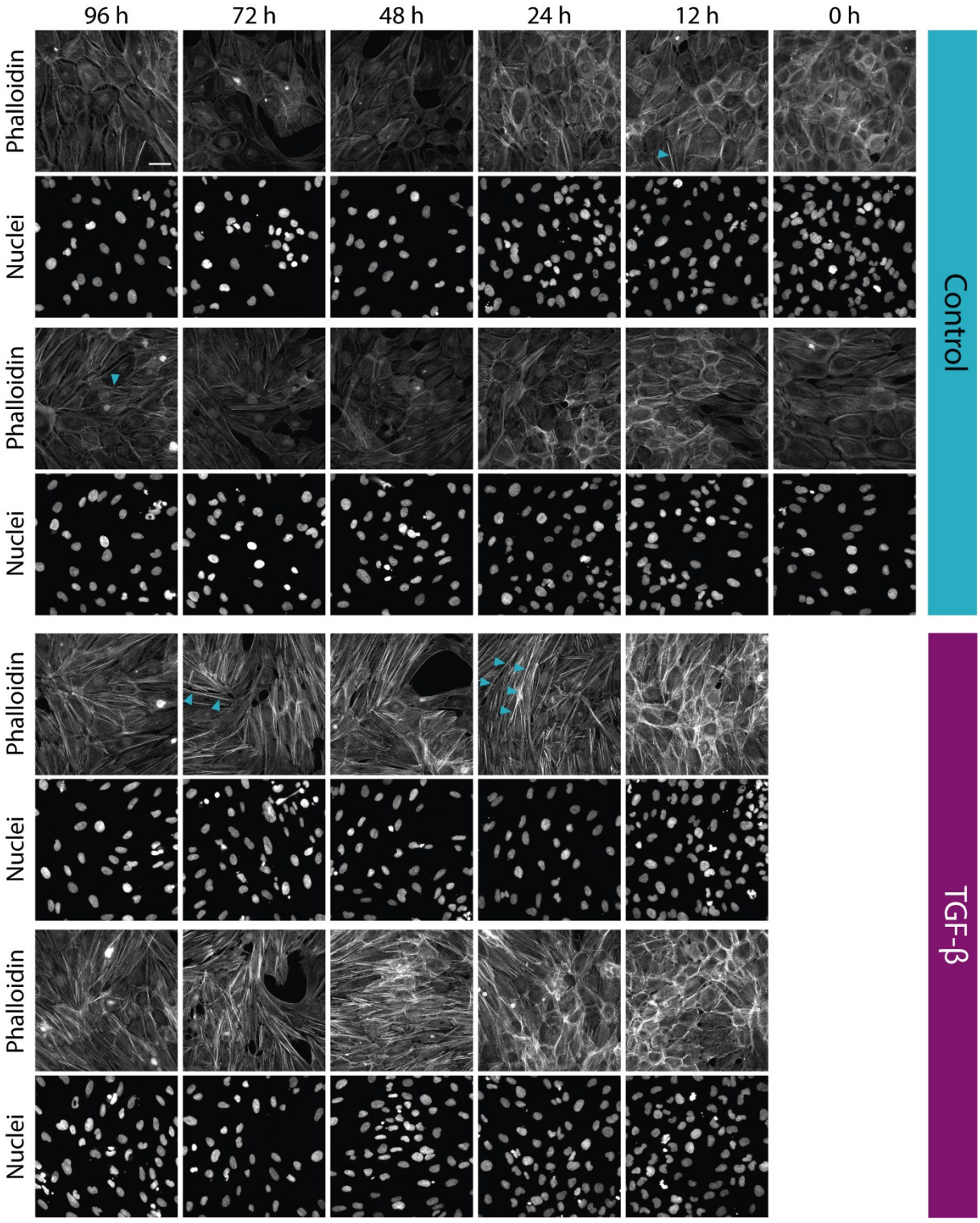

Figure 14 Morphological changes over time and with TGF-β treatment. Sum Z-projections of confocal images acquired at the Zeiss LSM 900 show immunofluorescence-stained cells at 0, 12, 24, 48, 72 and 96 h after ± TGF-β treatment with the experimental set-up B. Cells were seeded on day 0, and media of the longest treatment time point was changed to DMEM + Ficoll for the control condition, the same media + TGF-β for the treatment condition. Additionally, media of all other wells was changed to blank DMEM. The day after, the media was changed for all conditions as the day before, with the exception of the 72 h time point for which the treatment was started at that day. This continued until the cells were fixed at the last day and stained with Hoechst 33342 for nuclear staining and phalloidin for the staining of the actin cytoskeleton. Scale bar: 50 µm. Blue arrows highlighting examples of actin stress fibers.

Timely Dynamics of SMAD2 Phosphorylation

In an effort to capture early signaling responses, and since phosphorylation of proteins is an important factor in signal transduction, I wanted to add phosphoproteomics to the study. Initially, I examined the canonical downstream signal of TGF-β, SMAD2 and its phosphorylation (Friedman et al., 2013; Meng et al., 2016; Sziksz et al., 2015; Tie et al., 2022). Antibody validation for western blotting was performed, assessing SMAD2 phosphorylation at time points 24, 48, and 72 hours (**Figure *15***). The results demonstrated that TGF-β induced phosphorylation of SMAD2 across all tested time points. In the bar plot, the signals for phospho-SMAD2 (p-SMAD2) and SMAD2 were normalized to the loading control α-tubulin and normalized p-SMAD2 was further correlated to the SMAD2 signal.

The phenol free condition in the blot below refered to the DMEM used, which was without phenol red. This sample was just included because I wanted to see whether there was an effect on the signaling, as I wanted to use phenol free medium for imaging and the omics experiments.

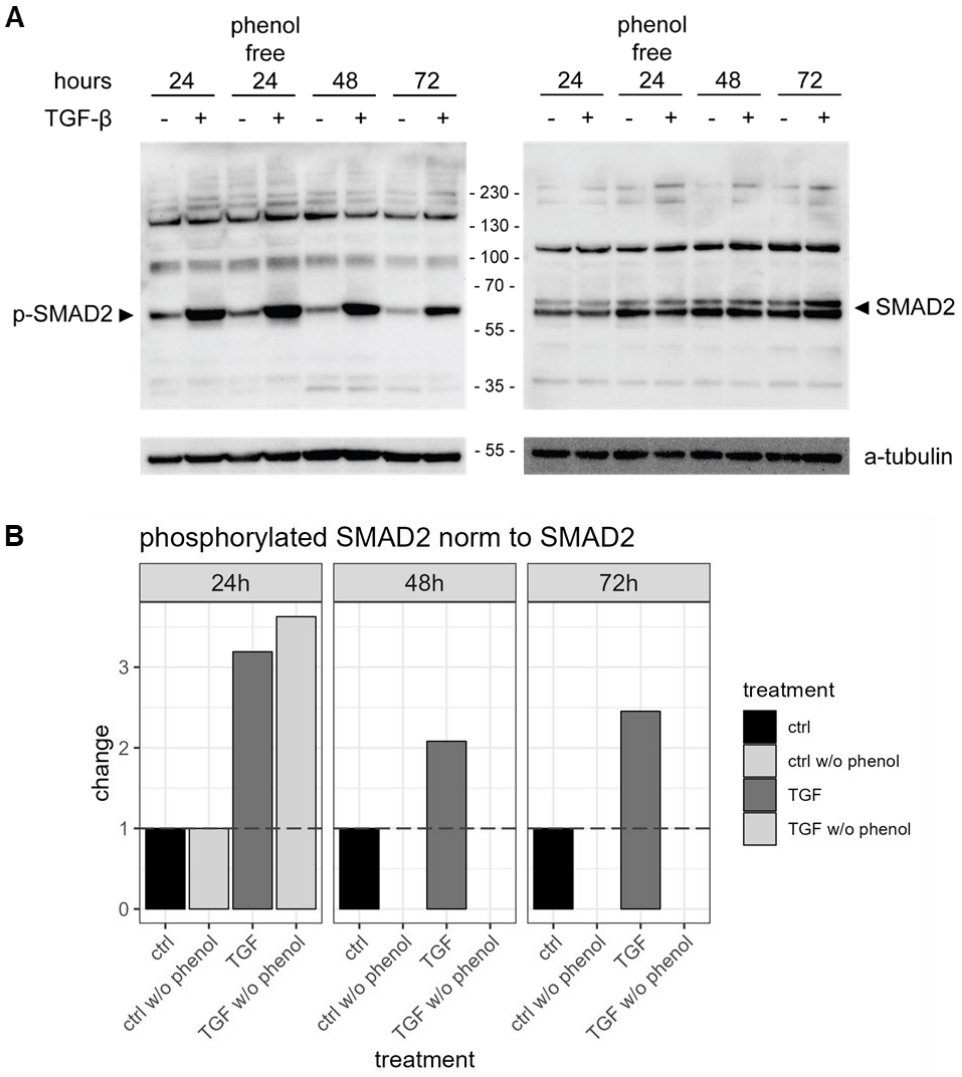

Figure 15 Analysis of SMAD2 phosphorylation and expression changes upon treatment with TGF-β for 24, 48 and 72 h. A) Western blots showing chemiluminescent signal of plots stained with phospho-SMAD2 (p-SMAD2) and SMAD2 antibodies and α-tubulin as loading control. Cells were treated ± TGF-β for the respective time points and loaded onto two gels in parallel. One membrane was used for staining with p-SMAD2 antibody and α-tubulin (after sodium azide treatment), while the second membrane was used to assess the change of SMAD2 protein expression through staining with an antibody against SMAD2. Following sodium azide treatment to deactivate the HRP activity of previous SMAD2 staining, the membrane was then stained for α-tubulin. **B)** Bar plots represent the change of chemiluminescence signal, corresponding to the respective staining, per time point. First, values for p-SMAD2 and SMAD2 were normalized to the loading controls. The resulting values for p-SMAD2 staining were then normalized to the respective signal for SMAD2. These results are depicted on the y-axis. The treatments are shown on the x-axis (ctrl in black, TGF-β in grey). In summary, stimulation with TGF-β increased the phosphorylation of SMAD2.

I further assessed the timeframe for SMAD2 phosphorylation following TGF-β treatment. The analysis was the same as above and showed that already 2 min after TGF-β treatment, phosphorylation of SMAD2 increased when compared to the control at that time (Figure *16*).

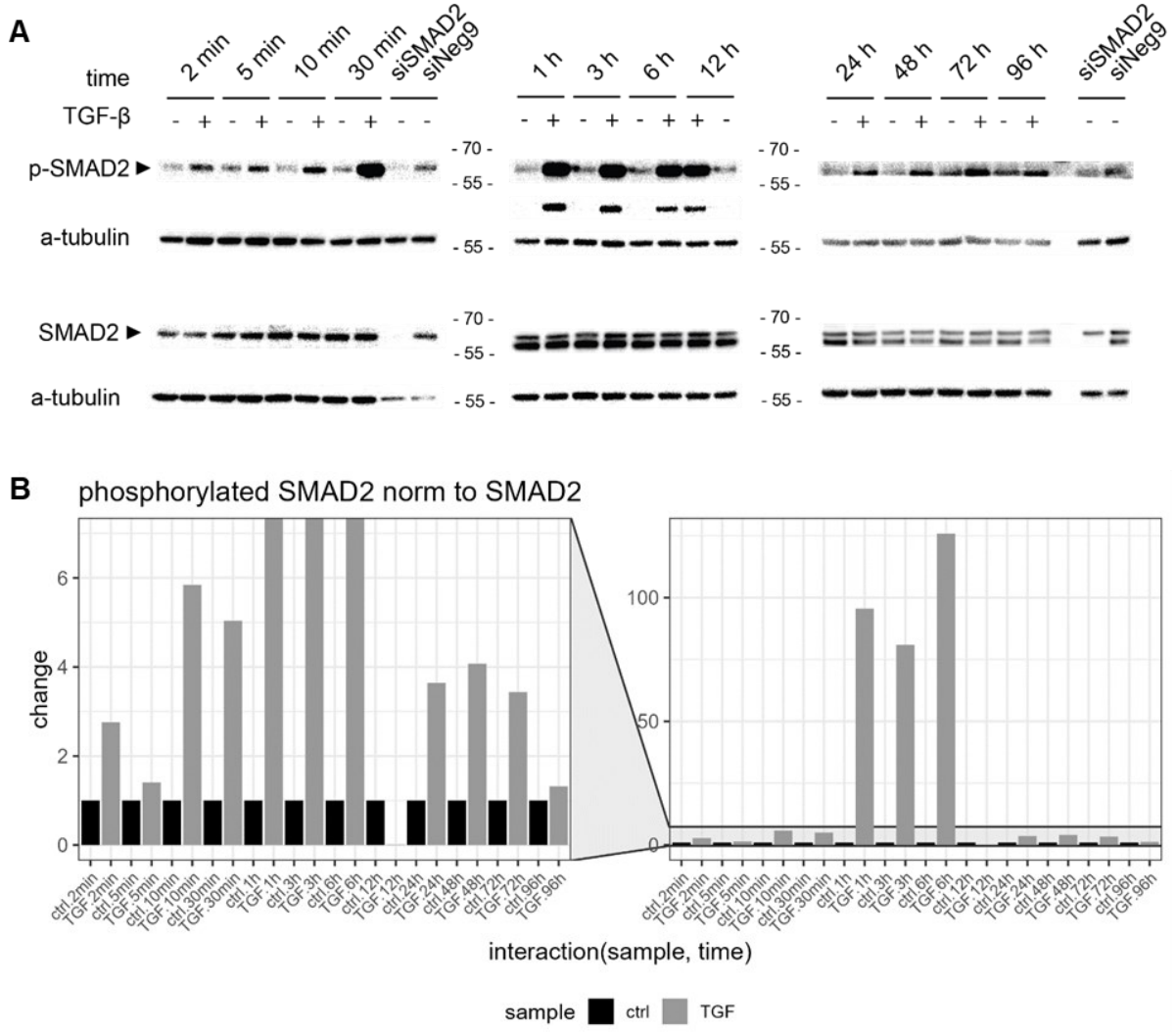

Figure 16 Analysis of SMAD2 phosphorylation and expression changes upon treatment with TGF-β after 2 min to 96 h. A) Western blots showing chemiluminescent signal of plots stained with phospho-SMAD2 (p-SMAD2) and SMAD2 antibodies and α-tubulin as loading control. Cells were treated ± TGF-β for the respective time points and loaded onto two gels in parallel. One membrane was used for staining with p-SMAD2 antibody and α-tubulin (after sodium azide treatment), while the second membrane was used to assess the change of SMAD2 protein expression through staining with an antibody against SMAD2. Following sodium azide treatment to deactivate the HRP activity of previous SMAD2 staining, the membrane was then stained for α-tubulin. Notably, since the signal for TGF-β treated samples at 1, 3, 6 and 12 h was overexposed in the first plot, right underneath the bands with a lower exposure are shown. **B)** Bar plots represent the change of chemiluminescence signal, corresponding to the respective staining, per time point. First, values for p-SMAD2 and SMAD2 were normalized to the loading controls. The resulting values for p-SMAD2 staining were then normalized to the respective signal for SMAD2. These results are depicted on the y-axis. The treatments are shown on the x-axis (ctrl in black, TGF-β in grey). Notably, the stimulation with TGF-β led to an increased phosphorylation of SMAD2 at all time points.

Replicating the previous experiments, one can see that already at the earliest time points tested, the phosphorylation of SMAD2 was increased, sustained over time and the highest at 30 min (**Figure *17***).

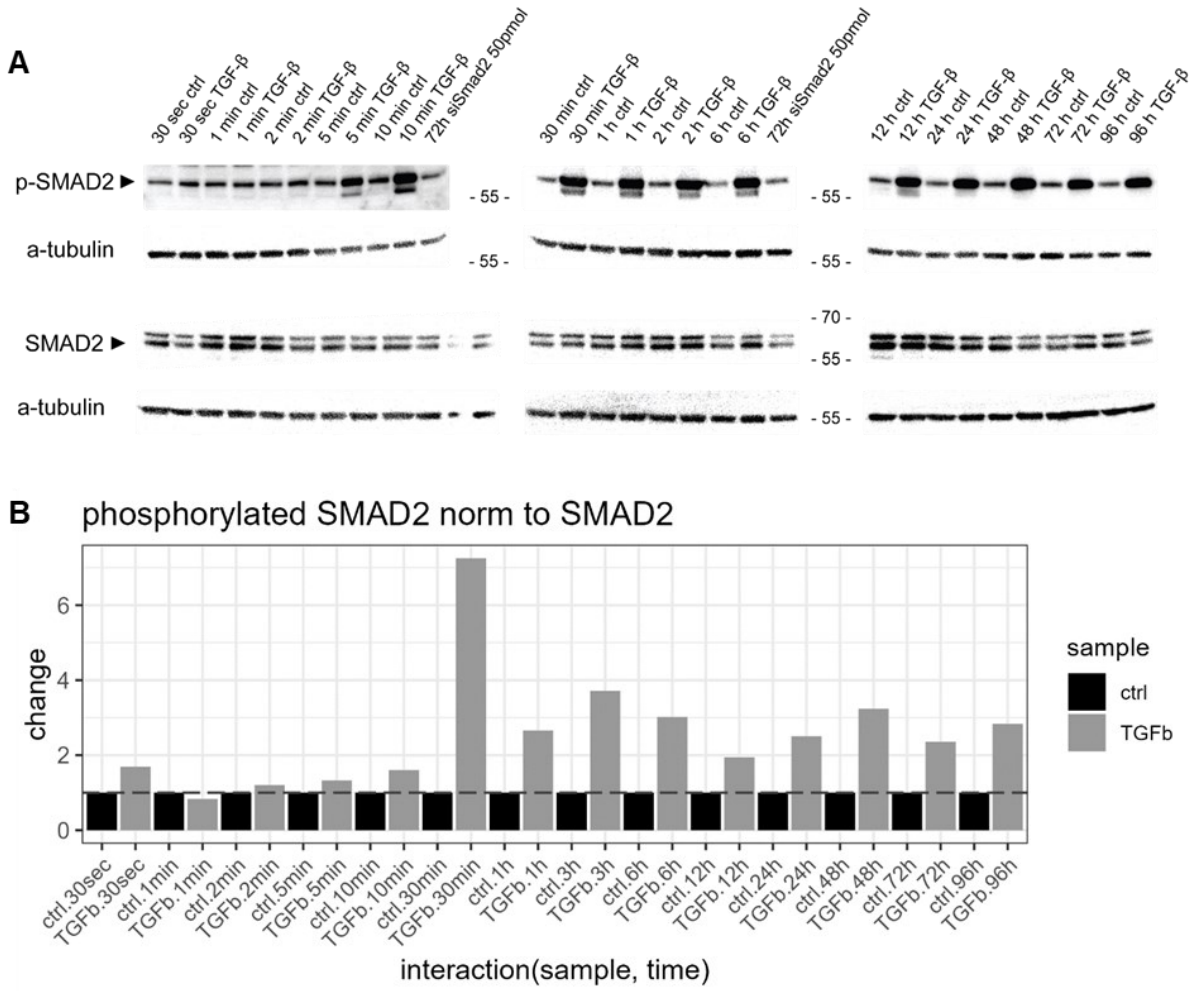

Figure 17 Analysis of SMAD2 phosphorylation and expression changes upon treatment with TGF-β after 30 sec to 96 h. A) Western blots showing chemiluminescent signal of plots stained with phospho-SMAD2 (p-SMAD2) and SMAD2 antibodies and α-tubulin as loading control. Cells were treated ± TGF-β for the respective time points and loaded onto two gels in parallel. One membrane was used for staining with p-SMAD2 antibody and α-tubulin (after sodium azide treatment), while the second membrane was used to assess the change of SMAD2 protein expression through staining with an antibody against SMAD2. Following sodium azide treatment to deactivate the HRP activity of previous SMAD2 staining, the membrane was then stained for α-tubulin. **B)** Bar plots represent the change of chemiluminescence signal, corresponding to the respective staining, per time point. First, values for p-SMAD2 and SMAD2 were normalized to the loading controls. The resulting values for p-SMAD2 staining were then normalized to the respective signal for SMAD2. These results are depicted on the y-axis. The treatments are shown on the x-axis (ctrl in black, TGF-β in grey). Increased phosphorylation of SMAD2 resulted from stimulation with TGF-β at almost all time points (except after 1 min).

4.1.4. Investigating Drug Effects on Extracellular Matrix Accumulation

To show the potential of the used cellular system for drug discovery purposes, I tested a few compounds that are used for the treatment of other fibrotic diseases. Two FDA approved drugs for the treatment of pulmonary fibrosis are Pirfenidone (PFD) and Nintedanib (Nin) (Bigaeva, Cavanzo, et al., 2020). The function of PFD is not entirely understood but is associated with inhibition of production as well as activity of TGF-β (Cho & Kopp, 2010).

Affected pathways by Nin are VEGF, FGF, PDGF as well as lymphocyte-specific protein tyrosine kinase and Src nonreceptor kinase signaling. The inhibited pathways are of therapeutic interest as all of them are involved in fibrosis (Bigaeva, Stribos, et al., 2020).

However, in the used cellular model, neither PFD nor Nin showed a decrease of extracellular Col I deposition after 48 h in the presence or absence of TGF-β (Figure *18*, Figure *19*). Only treatment with 0.05 µM and 0.5 µM PFD combined with TGF-β significantly decreased extracellular Col I deposition after 96 h when compared to the DMSO + TGF-β condition (Figure *18*).

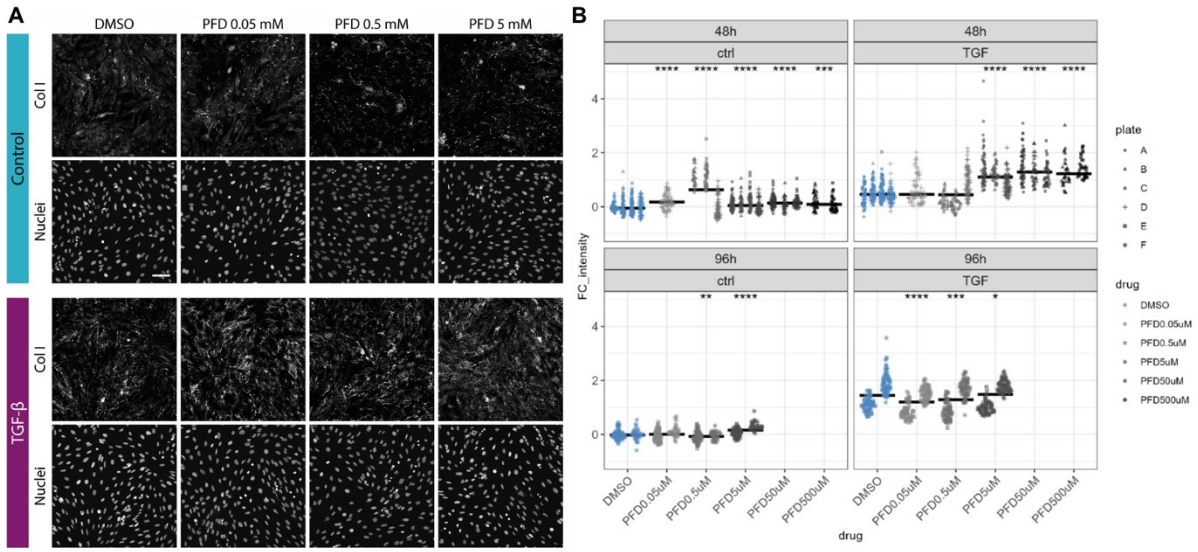

Figure 18 Assessing effects of Pirfenidone (PFD) on the accumulation of extracellular Col I over time. The day after cell seeding, the cells were treated with different concentrations of PFD (0.05, 0.5, 5, 50, 500 µM) either with or without TGF-β. Immunofluorescence staining after 48 (**A**) and 96 h of treatment was done for extracellular Col I. **B)** After background correction, the immunofluorescence signal of the DMSO control was first subtracted from all images. Subsequently, the values were divided by the signal of the DMSO control to result in the FC_intensity, plotted on the y-axis. This was done for each time point separately. The treatment and concentration of PFD used are indicated on the x-axis. Dots are color coded based on treatment (ctrl as blue, and PDF treated in grey). The bars indicate the median and individual data points denote the results from individual images taken. One to three biological replicates (plates) were tested for each condition. T-test was used to determine statistically significant differences between the PFD treated and DMSO samples per time point. This was also done separately for all TGF-β treated samples.

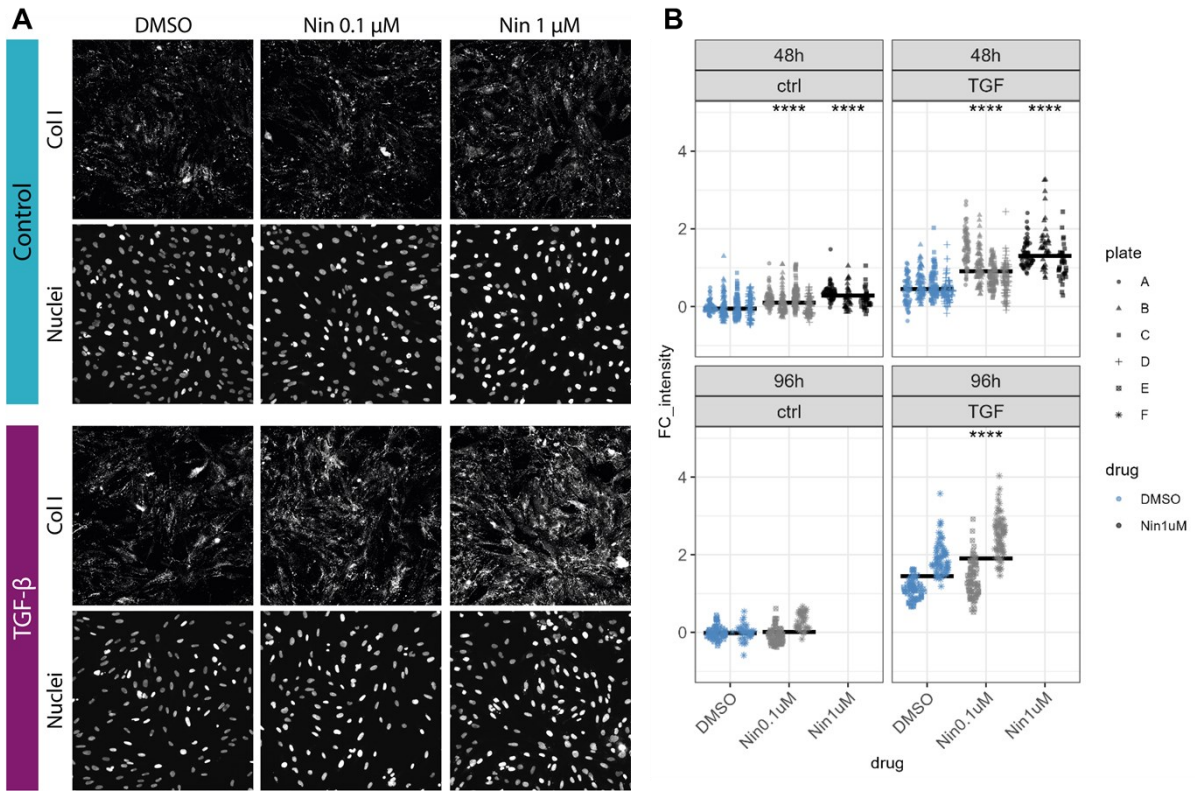

Figure 19 Assessing effects of Nintedanib (Nin) on the accumulation of extracellular Col I over time. The day after cell seeding, the cells were treated with different concentrations of Nin (0.1 and 1 µM) either with or without TGF-β. Immunofluorescence staining after 48 (**A**) and 96 h of treatment was done for extracellular Col I. **B)** After background correction, the immunofluorescence signal of the DMSO control was first subtracted from all images. Subsequently, the values were divided by the signal of the DMSO control to result in the FC_intensity, plotted on the y-axis. This was done for each time point separately. The treatment and concentration of Nin used are indicated on the x-axis. Dots are color coded based on treatment (ctrl as blue, and PDF treated in grey). The bars indicate the median and individual data points denote the results from individual images taken. One to three biological replicates (plates) were tested for each condition. T-test was used to determine statistically significant differences between the Nin treated and DMSO samples per time point. This was also done separately for all TGF-β treated samples.

Alternatively, I tested LY2109761 (LY), a TGF-β receptor type I/II (TβRI/II) dual inhibitor. Only when using 20 µM, the drug led to a reduction of extracellular Col I in the samples without TGF-β when compared to the DMSO treated control (Figure *20*). Nevertheless, in the TGF-β treated conditions, co-treatment with LY resulted in significant decrease of deposited Col I when compared to the TGF-β treated DMSO condition.

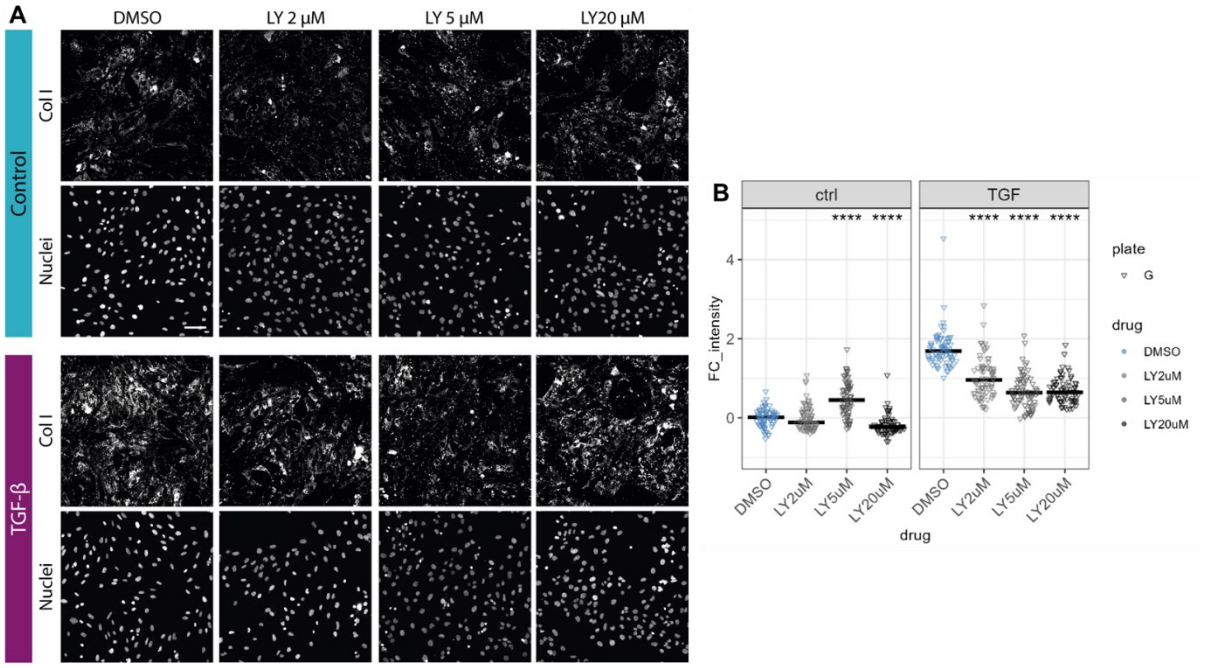

Figure 20 Quantification of effects by different concentrations of LY210976 (LY) on the accumulation of extracellular Col I after 48 h. The day after cell seeding, the cells were treated with different concentrations of LY (2, 5, 20 µM) either with or without TGF-β. **A)** Immunofluoresence staining 48 h post treatment was done for extracellular Col I. **B)** After background correction, the average immunofluorescence signal of the DMSO control was first subtracted from all images. Subsequently, the values were divided by the signal of the DMSO control to result in the FC_intensity, plotted on the y-axis. The treatment and concentration of LY used are indicated on the x-axis. Dots are color coded based on treatment (ctrl as blue, and LY treated in grey). The bars indicate the median and individual data points denote the results from individual images taken. The experiment included one biological replicate (plate). T-test was used to determine statistically significant differences between the LY treated and DMSO samples. This was also done separately for all TGF-β treated samples.

Additionally, I treated the cells with Dextromethorphan (Dex), a drug that showed up as a hit in my colleague's screen (Khan et al., 2023), leading to reduced Col I secretion in normal human lung fibroblasts (NHLF). Treatment with Dex showed the same potential in reduction of extracellular Col I deposition in our cellular model (Figure *21*). Dex treatment resulted in both conditions, with and without TGF-β, in a significant reduction of deposited Col I after 48 and 96 h compared to the respective controls.

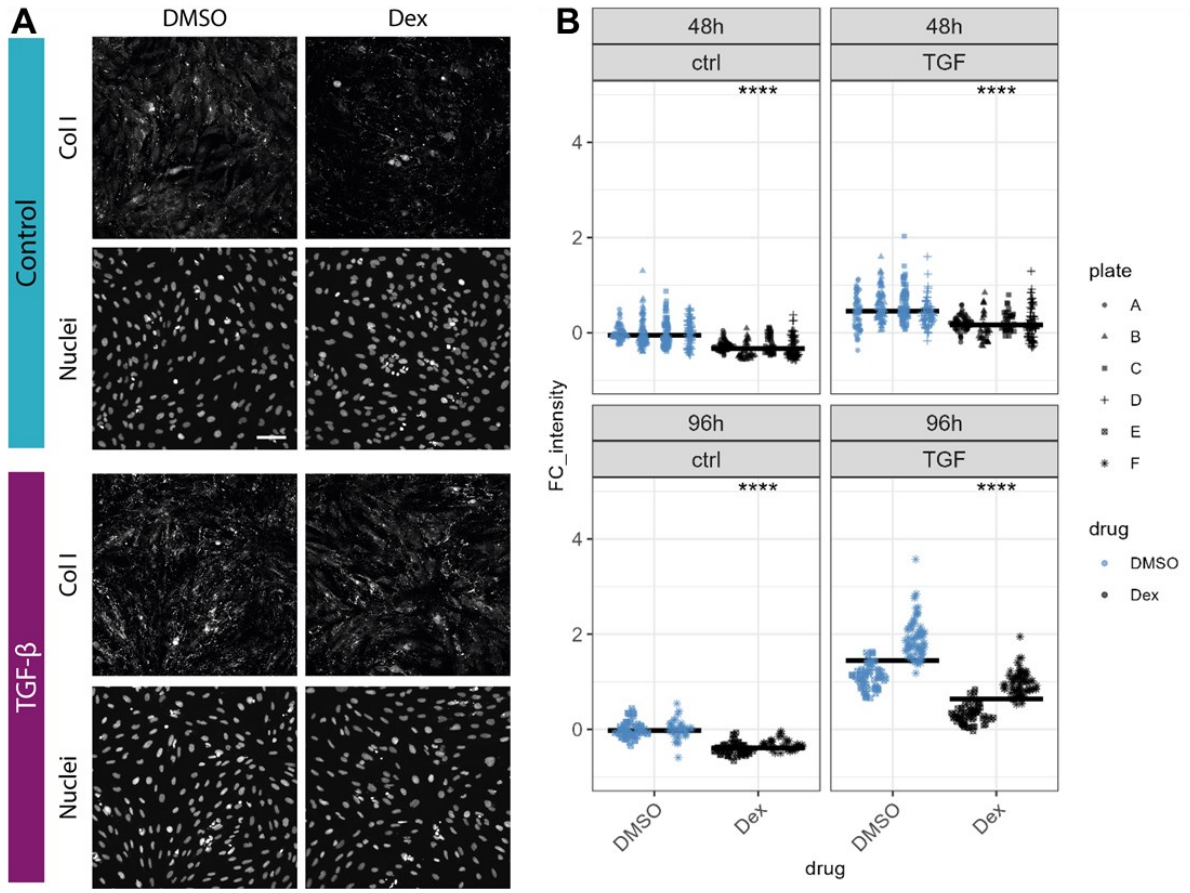

Figure 21 Quantification of extracellular Col I accumulation over time post Dextromethorphan (Dex). The day after cell seeding, the cells were treated with 10 µM of Dex and additionally either with or without TGF-β. Immunofluorescence staining 48 (**A**) and 96 h post treatment was done for extracellular Col I. **B)** After background correction, the average immunofluorescence signal of the DMSO control was first subtracted from all images. Subsequently, the values were divided by the signal of the DMSO control to result in the FC_intensity, plotted on the y-axis while the treatment is shown on the x-axis. Dots are color coded based on treatment (ctrl as blue, and Dex treated in black). The bars indicate the median and individual data points denote the results from individual images taken. The experiment included two to four biological replicates (plates). T-test was used to determine statistically significant differences between the Dex treated and DMSO samples. This was also done separately for all TGF-β treated samples.

4.2. Multi-omics Insights into Cellular Responses and Pathways

4.2.1. Motivation for and Experimental Set-up of Multi-omics Experiments

CKD is a disease that is driven by genetic, molecular and environmental factors. To further dissect the underlying mechanisms of a complex disease like CKD/kidney fibrosis and especially the contribution of the used cellular model to the pathology, I decided to provide a comprehensive view of the molecular mechanism. Aiming to use an unbiased approach and to integrate multiple read-outs, I chose to generate a multi-omics dataset. The read-outs involved whole RNA sequencing, proteomics including full proteomics (of the cell lysates), phosphoproteomics and secretomics (proteomics of the cell culture supernatants). Moreover, as established in previous sections, changes in the ECM were assessed using immunofluorescence staining of extracellular Col I.

Additionally, to understand the timely dynamics of the profibrotic signaling, cells were treated with TGF-β at the following time points 5 min, 1, 12, 24, 48, 72, and 96 h.

All data presented in the following section were generated from three to four biological replicates with the following set-up (set-up B in previous sections): Cells were plated at 30-40% confluency, which would lead to a confluency of roughly 90% at the collection point (after 5 days). The cells were grown for 18 h in the plates before changing the medium to 0% FBS-containing medium on day 1. Treatment for the longest duration (96 h), also began on day 1. On day 2, I started with the treatment for 72 h, and so forth. In general, the medium was changed every 24 h.

As controls, I collected control samples at each time point as well as an untreated 0 h control that was cultured in medium without molecular crowding and ascorbic acid (unlike treated and other control samples). At the end of the experiment, samples were either fixed (for imaging) or harvested for the omics experiments (**Figure 22**).

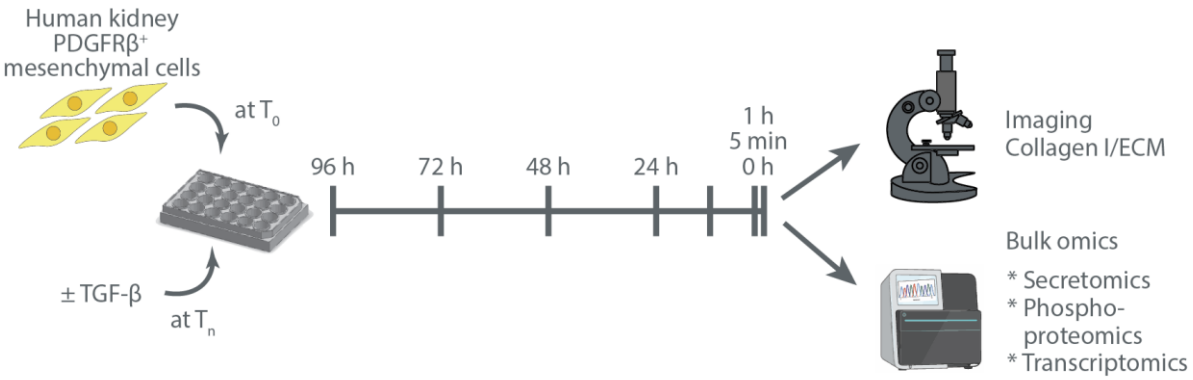

Figure 22 Set-up for multi-omics experiment. Cells were plated on day 0 at 30-40% confluency. At day 1, the medium of all conditions was changed to DMEM without FBS and Ficoll, except for the 96 h time point, where the medium was changed to DMEM containing Ficoll and ascorbic acid for the control. The latter medium was also used for the treatment condition and contained 10 ng/ml TGF-β. The media was generally changed every 24 h. On day 2, the 72 h treatment started. All samples of one biological replicate were harvested at the last day for the omics read-outs or fixed for the imaging of extracellular Col I. In general, three biological replicates were generated for transcriptomics and secretomics while four replicates were generated for proteomics, phosphor-proteomics and the imaging read-out.

4.2.2. Extracellular Matrix Dynamics

In addition to the multi-omics experiment, I investigated the effects of TGF-β treatment on extracellular matrix deposition through immunofluorescence staining of Col I. Therefore, I set up four biological replicates (plates) with two technical replicates each for multiple time points treated with and without TGF-β. Figure *23* shows representative images of the increase of extracellular Col I at the time points tested.

After image analysis with CellProfiler and rigorous background correction, including non-fibrillar Col I/ECM staining and autofluorescence subtraction (see Annex Figure 64), first, I calculated the fold-change between the samples stimulated with TGF-β and the respective controls. While the change of the control was set to 1, the average fold-change of the treated samples rangeed from 1.33 at 12 h, 1.57 at 24 h, 1.91 at 48 h, 1.92 at 72 h, to 2.47 at 96 h (Figure 24). Using a t-test, the increase of Col I post TGF-β treatment was shown to be significant at all time points compared to the respective control.

To further examine the increase of ECM over time, I normalized all samples to the respective 0 h control, per plate. In this case, the value of the 0 h control was subtracted from all images, and later divided by the 0 h control so that the 0 h ctrl is set to 0. In this graph (Figure 24 B), the increase of ECM in the controls is also visible at the different time points.

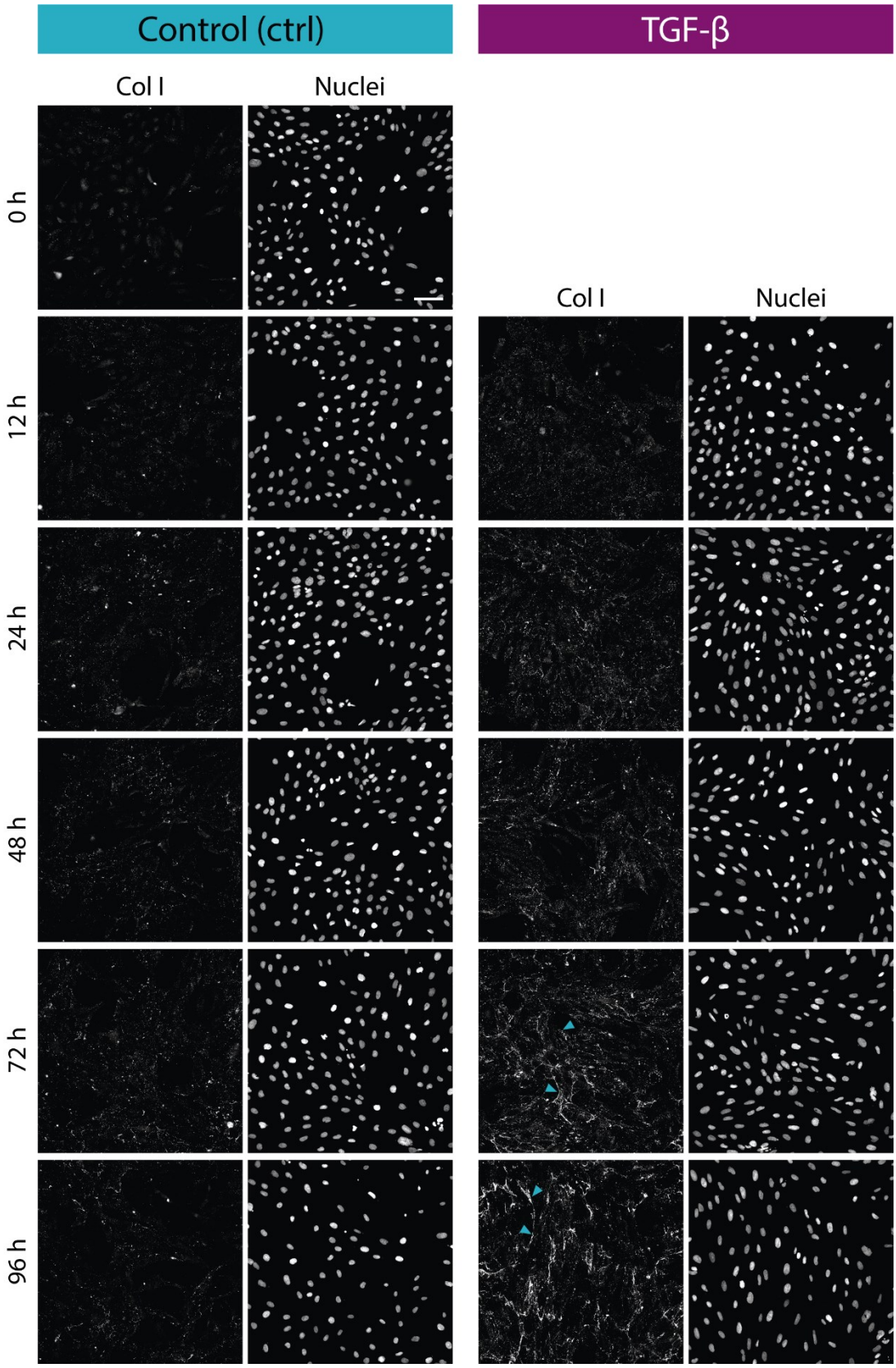

Figure 23 Immunofluorescence staining of deposited Col I over time. Widefield microscopy images were acquired with the Molecular Devices IXM. Staining of the extracellular Col I and nuclei stained with Hoechst 33342 are shown for the control and TGF-β treated samples over time (0 – 96 h). Scale bar = 100 µm.

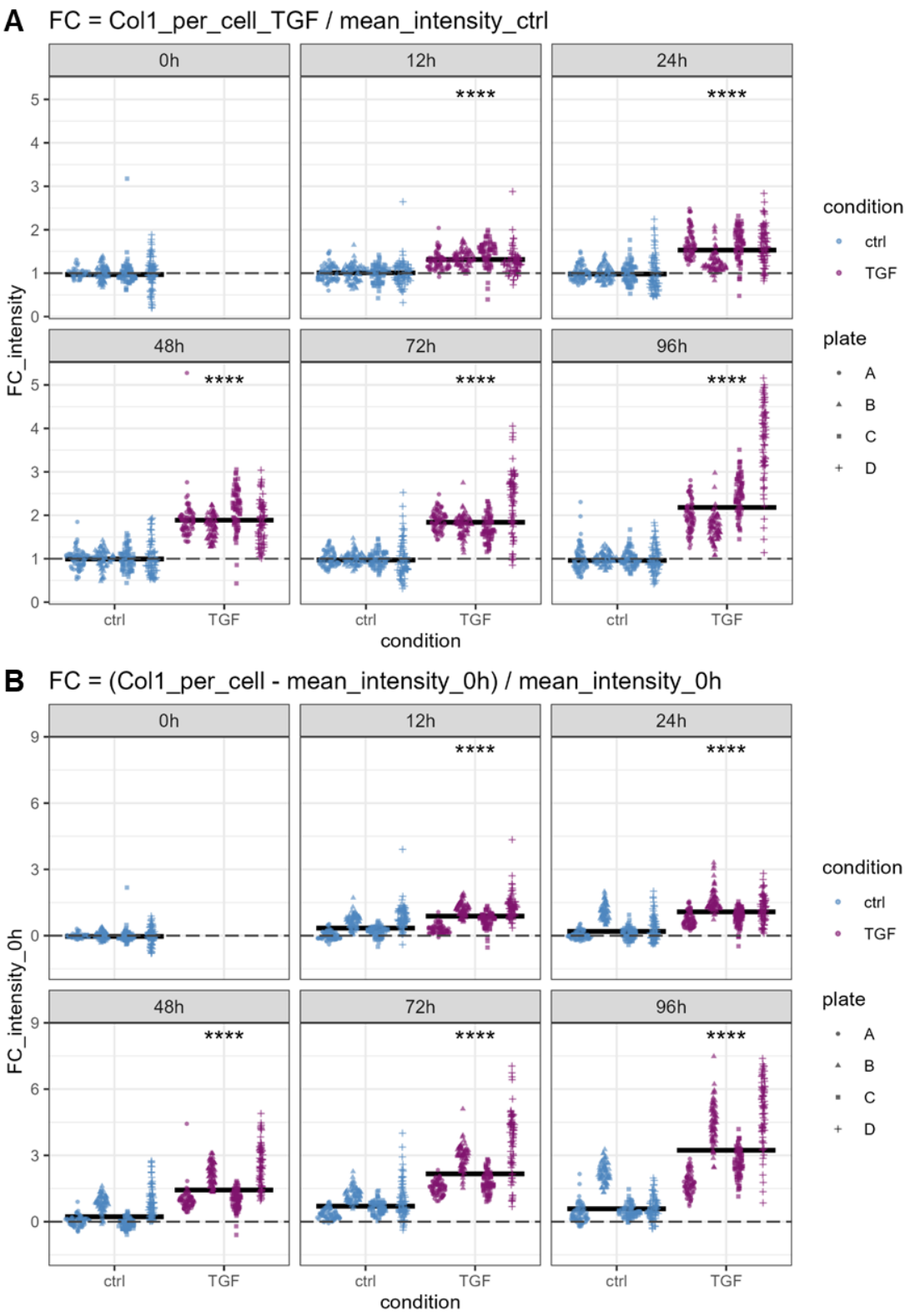

Figure 24 Temporal dynamics of extracellular Col I depicting fold-changes between A) TGF-β and control, B) alongside comparison to the baseline (0 h). Each of the data points represents the fold-change calculated for **A)** TGF-β treated sample compared to the corresponding controls at each of the time points, and **B)** each data point against the 0 h control. The data are presented for individual time points and biological replicates (plates A-D, denoted by distinct shapes). Control sample values are denoted in blue, while TGF-β treated samples are represented in purple. Statistical analysis was conducted utilizing the t-test. Notably, the fold-change between TGF-β stimulated samples and controls increased with time, while there is also an increase in deposited Col I in the control samples as time progresses.

To further analyze the changes of ECM over time and assess the differences between the samples statistically, I used a mixed linear model. While accounting for fixed effects like treatment (condition), time and their interaction, I also included the between-plate variability as random effects.

In general, the extracellular Col I deposition changed significantly over time and with TGF-β treatment (Table 16). More specifically, I asked myself at which time points there was a significant difference between TGF-β treated samples compared to the respective controls. The difference of deposited Col I was not significant between the 12 h control and TGF-β treated samples. Nevertheless, from the 24 h time point onwards, the post hoc test of the model indicated a significant increase of Col I deposition in the TGF-β stimulated samples when compared to the respective controls at each time point (Table 17). This effect was consistent across biological replicates, indicating robustness and reproducibility.

Based on previous experiments, I hypothesized that the accumulation of Col I would increase over time. While the early time points 12, 24, and 48 h were not significantly different from the 0 h time point, the later time points 72 and 96 h were (Figure 25, Table 17). However, the control samples largely grouped together, including the 12 h TGF-β time point. Only the 48 h, 72h and 96 h time point were significantly different from each other. Additionally, treatment with TGF-β for 12 and 24 h did not lead to a significant increase of deposited Col I when compared to the last control time points, 72 and 96 h. While the TGF-β treated samples, 12 and 24 h, as well as 24 h and 48 h grouped together, the 72 and 96 h samples post TGF-β treatment are significantly different from the other time points. Even though the biological replicates B and C show consistently higher per cell intensity (as shown by the intercept, Figure *65*) than plates A and D, the increase due to TGF-β treatment appeared constant for all biological replicates within one time point (Figure 25).

Overall, these results showed the suitability of the used cellular model to study one of the most widely used phenotypic read-outs for fibrosis and the importance that TGF-β plays in the profibrotic response of the cells.

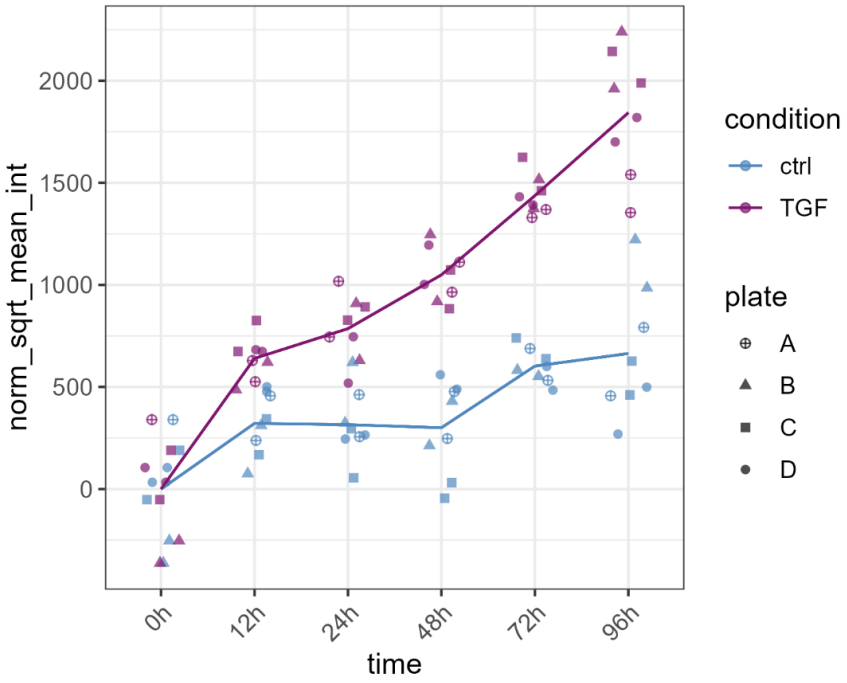

Figure 25 Temporal dynamics of TGF-β induced Collagen I deposition in cultured cells. Cells were cultured in glass-bottom 24 well-plates with four biological replicates (plate A-D, indicated by different shapes) with two technical replicates each. Cells were treated ± TGF-β for durations ranging from 0 to 96 h. Using CellProfiler 4.2.6, the intensity values for the Col I staining and nuclei count were obtained from the imaging data. After background correction (including subtraction of non-fibrillar Col I/ECM staining and cell autofluorescence correction), the Col I intensity values of each image were normalized to the corresponding cell number to result in the Col I intensity per cell. For statistical analysis, the mean was calculated for each sample, the data were square root transformed and the resulting sqrt_mean_int values for the 0 h time points were duplicated for the TGF-β condition. A mixed linear model (formula = sqrt_mean_int ~ condition*time + (1 | plate)) was applied for statistical analysis of the increase of deposited Col I in both control as well as TGF-β treated conditions over time. Shown are square root transformed mean Col I per cell (sqrt_mean_int) on x-axis, with the y-axis displaying the time. Shape of individual data points depict the two technical replicates per biological replicate (plate), and colors stand for the ctrl (blue) and treatment with TGF-β (purple). More specifically, for visualization, the data were normalized to the predicted value at the 0 h time point, done separately per plate. The predicted values (also normalized to the values at 0 h), are indicated by the line, separately for the control (blue) and TGF-β samples (purple).

Table 16 ANOVA analysis of Deviance Table (Type III Wald chisquare tests).

Response: sqrt_mean_int			
	Chisq	Df	Pr(>Chisq)
(Intercept)	37.814	1	7.783e-10 ***
condition	0.000	1	1
time	52.276	5	4.734e-10 ***
condition:time	79.919	5	8.727e-16 ***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1			

Table 17 Post hoc test of linear mixed model for pairwise comparisons between conditions and time points. For each condition and time, the estimated marginal means (emmeans, average response variable value across all levels of factors) together with the confidence intervals (lower confidence level and upper confidence level (CL)) are shown. The group column provides grouping labels based on significant differences between means.

condition	time	emmean	lower.CL	upper.CL	.group
TGF	0h	1670	-128	3469	a
ctrl	0h	1670	-128	3469	a
ctrl	48h	1970	143	3798	ab
ctrl	24h	1985	158	3813	abc
ctrl	12h	1992	164	3819	abc
ctrl	72h	2272	445	4100	bcd
TGF	12h	2310	482	4138	bcd
ctrl	96h	2334	506	4162	cd
TGF	24h	2456	628	4284	de
TGF	48h	2719	892	4547	e
TGF	72h	3107	1280	4935	f
TGF	96h	3513	1686	5341	g

Degrees-of-freedom method: kenward-roger
Confidence level used: 0.95
Conf-level adjustment: sidak method for 12 estimates
P value adjustment: sidak method for 66 tests
significance level used: alpha = 0.05

4.2.3. Multi-omics Read-outs

Transcriptomics

In transcriptomics, we identified a total of 14939 genes, with roughly 2880 differentially expressed genes (absolute log2 fold-change > log(1.5), adjusted p-value < 0.05, over at least one time point) post TGF-β treatment. Opting for a more stringed log2 fold-change cut-off of 2, resulted in 1442 differentially expressed genes (roughly 10% of all identified genes).

The PCA plot (**Figure *26***) showed a clear separation between treated and control samples and a time dependent effect. Early time points, 5 min (0.08 h) and 1 h cluster together with the 0 h point. Notably, there was an outlier in the transcriptomics data (24 h ctrl, rep A) which I excluded for analysis as this was caused by an error while performing the experiment.

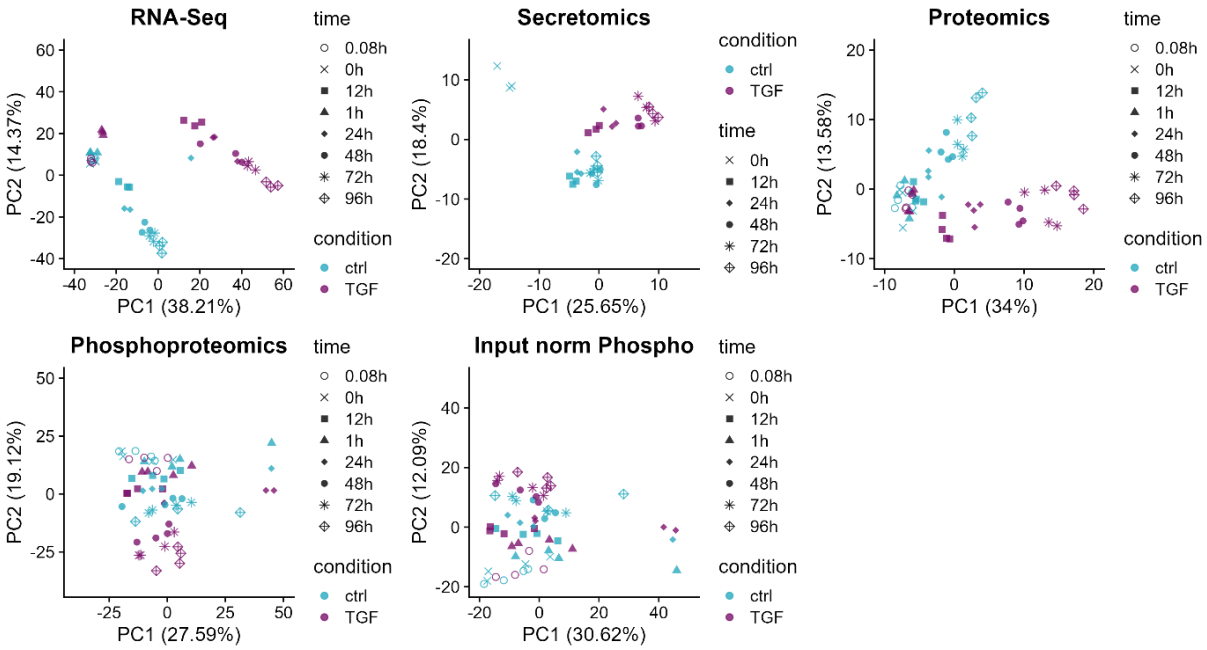

Figure 26 Principal component analysis (PCA) plots per omics modality. PCA scatter plots of dimensions PC1 and PC2. Samples are color coded by treatment, blue for control samples and purple for TGF-β treated samples. The various shapes indicate the treatment time. Included were 3 to 4 biological replicates.

Proteomics

I hypothesized that fibrosis, and the aberrant accumulation of ECM and secretion of signaling molecules, are influenced by mechanisms that are not only dependent on transcriptional regulation but can be driven by factors like protein abundance changes and alterations in signaling pathways, e.g., induced by differences in post-translational modifications (Mullenbrock et al., 2018). Therefore, I had a further look into the proteomics and phosphoproteomics of cell lysates and cell culture supernatants (thereafter referred as secretomcis).

Proteomics and Phosphoproteomics of Cell Lysates

To link the findings further to the signaling mechanisms, we performed quantitative MS/MS analysis of protein lysates of cells stimulated with and without TGF-β for the beforementioned time points. 5% of the cell lysates were taken as a full proteome input, and 95% were used for the IMAC phosphopeptide enrichment, followed by high pH reverse phase fractionation and quantitative MS/MS.

For the proteomics data, 7024 proteins were identified with 196 proteins being differentially regulated (absolute log2 fold-change > log(1.5), adjusted p-value < 0.05) when comparing TGF-β versus the control treatment at each time point separately.

In phosphoproteomics, 17255 phosphosites were identified resulting in 15186 phosphosites when mapped to the proteomics data of the whole cell lysates. This represents 794 differentially expressed phosphosites in the phosphoproteomics and 231 when normalized to the proteins measured in the full proteome runs.

In general, the PCA plots (Figure *26*) indicated a separation between treated and control samples over time. Especially in the PCA plot for the proteomics data one can see that the early time points like 5 min (0.08 h) and 1 h cluster together with the 0 h time point, as already seen for the transcriptomics results. Whereas most of the variance of the phosphoproteomics data and Input norm Phospho (phosphoproteomics normalized to the full proteome) is explained by the biological replicates (PC1).

Secretomics - Proteomics of Cell Culture Supernatants

On the one hand, I wanted to look at secretomics to identify potential biomarkers that could later on be assessed in the blood or urine of patients. On the other hand, it was evident that mesenchymal cells would secrete factors that further drive the profibrotic environment and thereby perform autocrine signaling as well as stimulating other cells that additionally promote the profibrotic environment. Therefore, to establish a coherent dataset, I collected cell culture supernatants of the plates used for proteomics.

Combined analysis of the secretomics spectra resulted in the identification of 2187 proteins, 219 of which were significantly differently expressed (absolute log2 fold-change > log(1.5), adjusted p-value < 0.05) after TGF-β treatment. We further filtered the data based on entries of the MSigDB gene sets NABA_MATRISOME and NABA_MATRISOME_ASSOCIATED and ended up with roughly 80 differentially expressed proteins in at least one time point after manual annotation (we added MARCKS, CDH11, NRP2, SDC4, TFPI, CPA4, CACHD1).
Unlike for the other omics, we did not measure the secreted proteins after 0.08 and 1 h post TGF-β treatment. In the PCA plot (Figure *26*), the 0 h time point was clearly separated from the other clusters. As already described for the other omics modalities, the control samples were separated from the TGF-β treated samples. Additionally, the time dependent effect was visible.

Summary and Overview of Multi-Omics Data

One approach to link the findings obtained from the multi-omics experiment was to look for similar patterns and similarly regulated pathways/factors that were shared across the omcis modalities. Therefore, I will summarize them in this section.

Increase of Significantly Deregulated Factors Over Time

The analysis showed that the number of significantly affected proteins or genes (absolute log2 fold-change > log(1.5), adjusted p-value < 0.05) increased over time in all the data modalities (Figure 27), indicating a time dependent effect of the TGF-β treatment.

More specifically, at the 5 min (0.08 h) time point of the transcriptomics data (rna), no genes were significantly deregulated whereas at 1 h post TGF-β stimulation, 98 genes were significantly deregulated. The number of differentially expressed transcripts increased over time with 470 at 12 h, 676 at 24 h, 624 at 48 h, 767 at 72 h, and 950 at 96 h.

While there were no differentially regulated proteins at 5 min and 1 h post TGF-β treatment, 11 proteins were up-/down-regulated at the 12 h time point, 35 at 24 h, 77 at 48 h, 123 at 72 h, and 181 proteins after 96 h. Interestingly, at 5 min 2 phosphosites were deregulated upon treatment with TGF-β followed by 29 phosphosites at 1 h. At the 12 h time point, with 2 phosphosites there were less phosphosites significantly deregulated than at 1 h. This number, however, increased again over time with 13 sites at 24 h, 26 at 48 h, 54 at 72 h, and 176 at 96 h.

Looking at the secretomics data, one can see that there were 23 differently expressed proteins after 12 h of stimulation with TGF-β, increasing to 22 at 24 h, 42 at 48 h, 51 at 72 h, and 58 proteins at 96 h post TGF-β treatment.

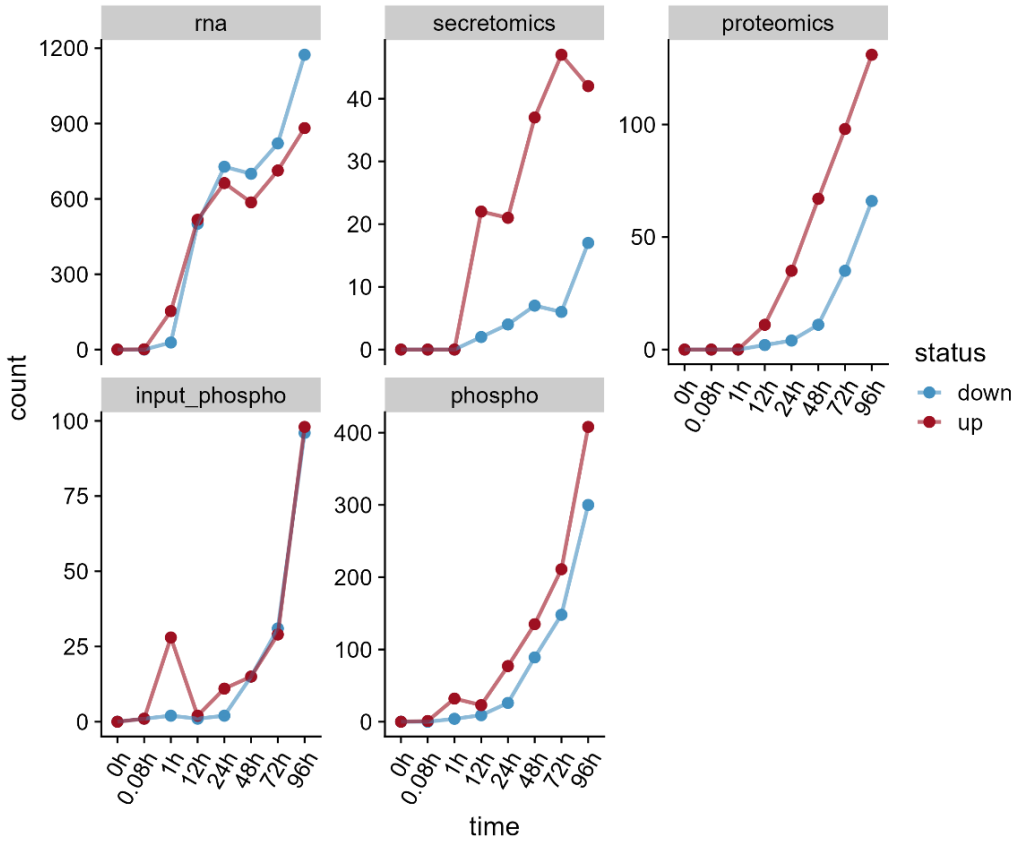

Figure 27 Summary of number of hits over time per omics modality. The number of hits (absolute log2 fold-change > log(1.5), adjusted p-value < 0.05), up (red) and down (blue) regulated genes is shown on the y-axis per time, depicted on the x-axis. This was done per omics modality (transcriptomics = rna, secretomics, proteomics, input_phospho = phosphoproteomics normalized to proteomics of the whole cell lysates, and phospho = phosphoproteomic).

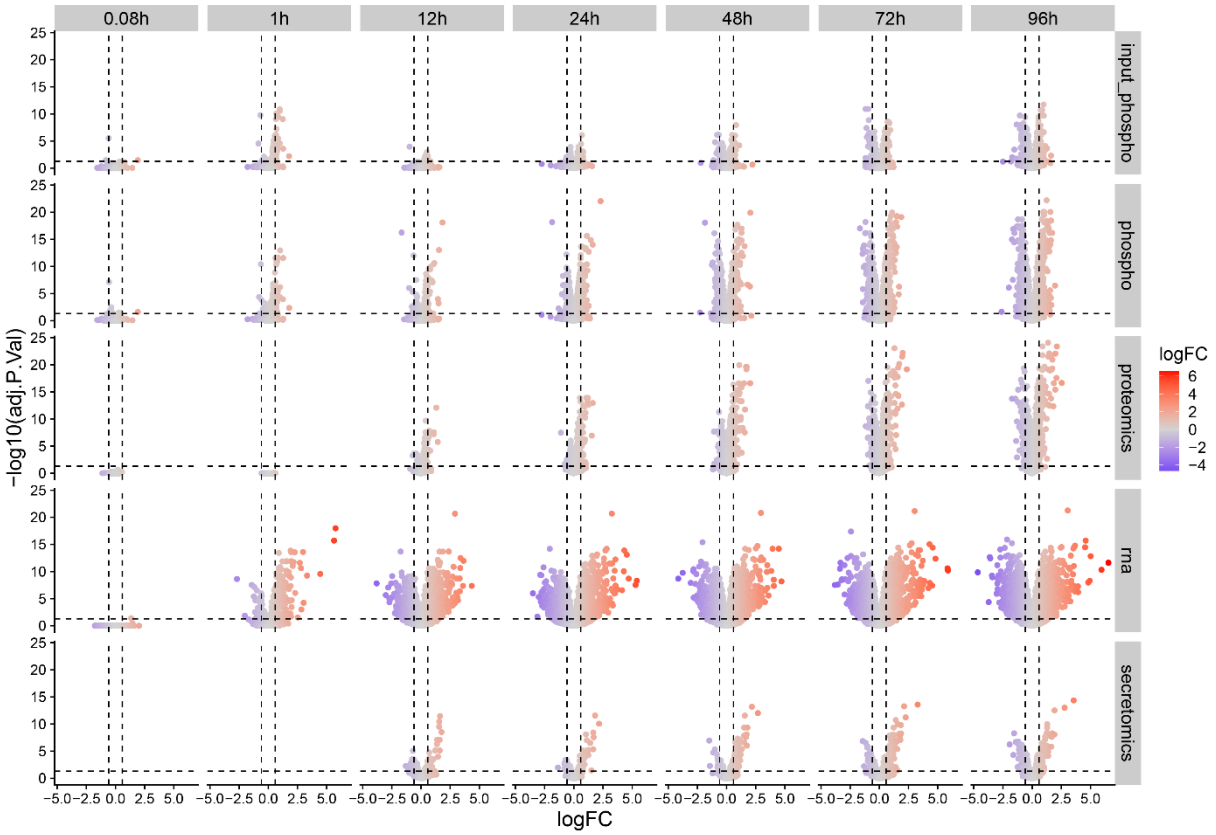

Figure 28 Volcano plot. Volcano plot with the log2 fold-change on the x-axis and adjusted p-value (-log10(adj.P.Val)) on the y-axis. In red, significantly upregulated factors with an absolute log2 fold-change > log(1.5) are shown. Blue dots represent significantly downregulated. This is displayed for each omics modality and per time point measured. In general, there is an increase of significantly deregulated factors among all modalities and with time.

Functional Analysis Indicates Pro-Fibrotic Response

In Figure *29*, using MSigDB and decoupleR (Badia-i-Mompel et al., 2022), it can be seen that treatment of TGF-β induced profibrotic pathways like **TGF-β signaling, epithelial to mesenchymal transition** (EMT), and **hypoxia**, starting from 12 h onwards in three omics modalities including transcriptomics, proteomics and secretomics. In contrast, anti-inflammatory pathways like **Interferon α and γ** response, were initially upregulated at 5 min in the transcriptomics data but were then continuously downregulated. From 72 h post TGF-β treatment, the mentioned interferon pathways were also downregulated in proteomics and secretomics. Another downregulated pathway that was shared among omics modalities was the **fatty acid metabolism** starting from 12 h onwards on proteomics and additionally transcriptomics at some time points.

Induced by TGF-β treatment, **TNFα signaling via NF-κB** was upregulated at 1 h post TGF-β treatment in proteomics and transcriptomics, whereas it stayed upregulated until 72 h in proteomics. Other pathways that were upregulated upon TGF-β stimulation

at the tested time points and across the omics modalities were **Myogenesis** (in proteomics, secretomics and RNA-seq at 12 to 72 h, and in proteomics alone at 96 h), **Coagulation** (in proteomics, secretomics and RNA-seq at 24 h, in proteomics and secretomics at 48 and 72 h, in proteomics alone at 96 h), **Angiogenesis** (in proteomics, secretomics and RNA-seq at 24 and 72 h, in proteomics only at 12, 48 and 96h), and **Apical junction** (in proteomics, secretomics and RNA-seq at 48 to 96 h, in RNA-seq and proteomics at 24h, at proteomics alone at 12 h).

Interestingly, **cholesterol homeostasis** was only downregulated in the transcriptomics data at 24 and 72 h and in transcriptomics and proteomics at 96 h. Furthermore, **Oxidative phosphorylation** was downregulated in proteomics and at some time points secretomics, starting 24 h after TGF-β treatment.

Additionally, the **E2F targets** pathway was downregulated in RNA-seq and proteomics after 72 h, and is downregulated in secretomics at 96 h post TGF stimulation. At 96 h, the **G2M checkpoin**t pathway was deregulated in transcriptomics. In line with this was the increase of **p53 pathway** that was upregulated starting at 48 until 96 h after TGF-β treatment in transcriptomics.

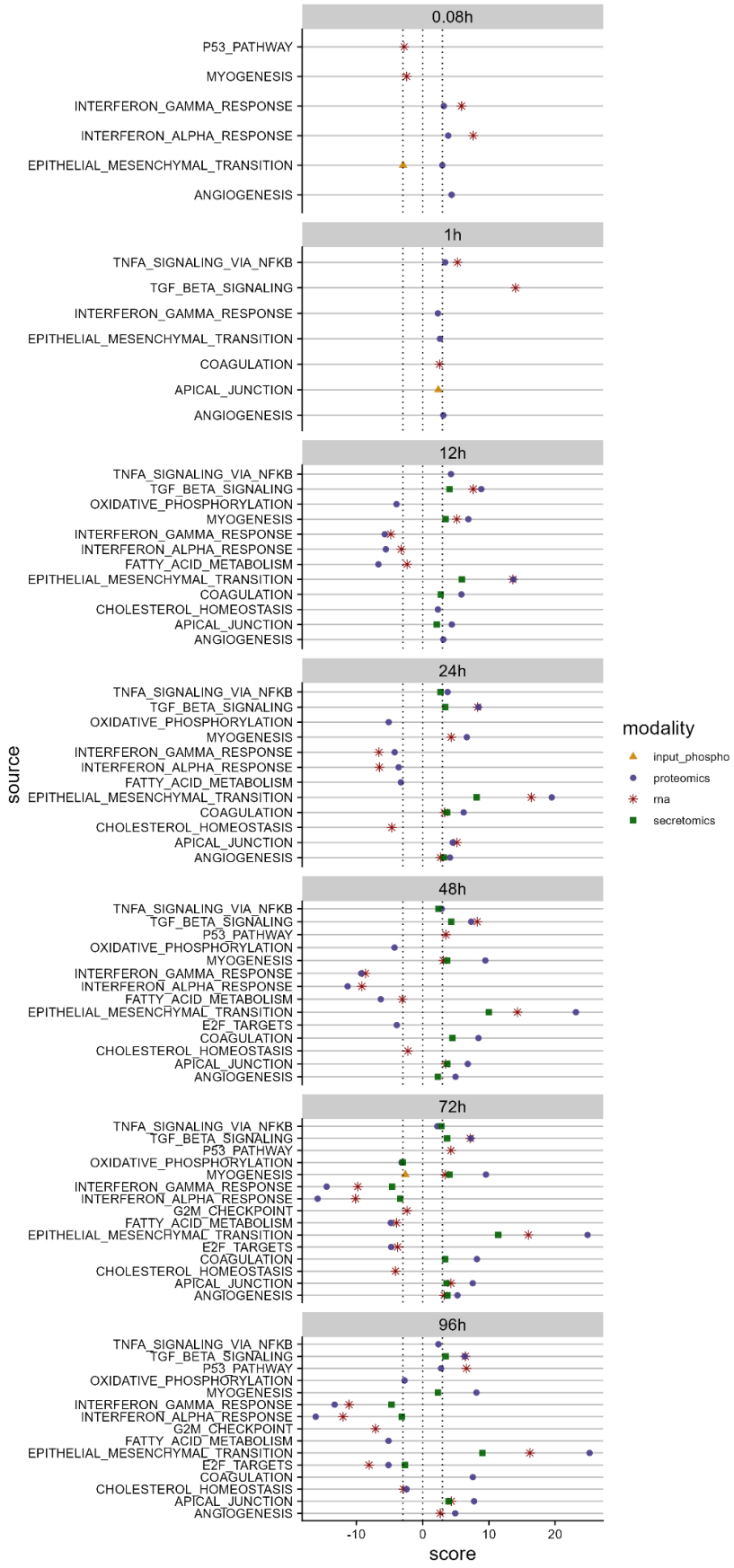

Figure 29 Functional analysis over time and for all omics modalities. A selection of pathways that are shared between the omics modalities is shown for the different time points. The color and shape of each dot depicts the modality: input_phospho = yellow triangle, protomics = blue dot, RNA (transcriptomics) = red star, secretomics = green square.

Multi-Omics Hits

In my thesis, I aimed to highlight secreted proteins that also were significantly regulated in other omics modalities. As already mentioned, I focused on secreted proteins because they have a great potential to serve as biomarkers as those could be potentially measured in the blood or urine of patients.

The heatmap below summarizes overlapping hits that showed up in at least two out of the three omics modalities, at one or more time points, as significantly deregulated (absolute log2 fold-change > log(1.5) (log(2) for transcriptomics), adjusted p-value < 0.05). White cells in the heatmap indicate that the corresponding genes/proteins were not significantly deregulated at the respective time point. I did not include the phosphoproteomics in there as the number and overlap of hits was low. For readability of the heatmap, I divided it into two figures (**Figure *30***, **Figure *31***). Since the hits were roughly divided into three cluster, I generated one heatmap with cluster two (**Figure *30***) and another heatmap with cluster one and three (**Figure *31***). Cluster one contained factors which expression was mainly downregulated upon treatment with TGF-β. Factors which expression was upregulated over time are summarized in cluster two. In cluster three, the factors were strongly upregulated over time.

Factors, which expression was upregulated in transcriptomics, proteomics and secretomics included TGFBI, SERPINE, IGFBP3, LTBP2, TNC, AMIGO2, LMCD1, ADAM12, COL4A1, BGN, MMP2, HTRA1, FBN1, COL5A1, CDH11 in cluster three, and TGFB1, MMP14, PLXDC2, CACHD1, THBS1, HMCN1, NRP2, SPARC, SERPINE2, COL7A1, MRC2, IGFBP7, SEMA3C, PTK7, MARCKS, CSPG4, COL4A2, CCN2, LTBP3, CTHRC1, CRIP2, GPC4, PXDN, and COL6A2 in cluster two. TFPI, MX1 as well as CPA4 were downregulated (cluster one). In the proteomics and transcriptomics data, factors such as PADI2, ITGA11, CA3, SLC29A1, COL1A1, FN1, ITGAV, TNS1, JUP**,** CMPK2 showed up as upregulated hits while factors like OAS3, OAS1, OAS2, MX2, STARD5 and DCDC2 were downregulated.

SRFP2, POSTN, GDF6, IL11, GREM1, INHBA, ADAM19, MMP11 and ADAMTS7 were significantly upregulated in the secretomics as well as transcriptomics data but showed no significant change in proteomics. Notably, COL1A2 was significantly downregulated in the last time point of the secretomics data but upregulated over time in the transcriptomics data. Similarly, expression of COL18A1 was decreased at the last time point in the secretomics data but upregulated in the proteomics and transcriptomics data.

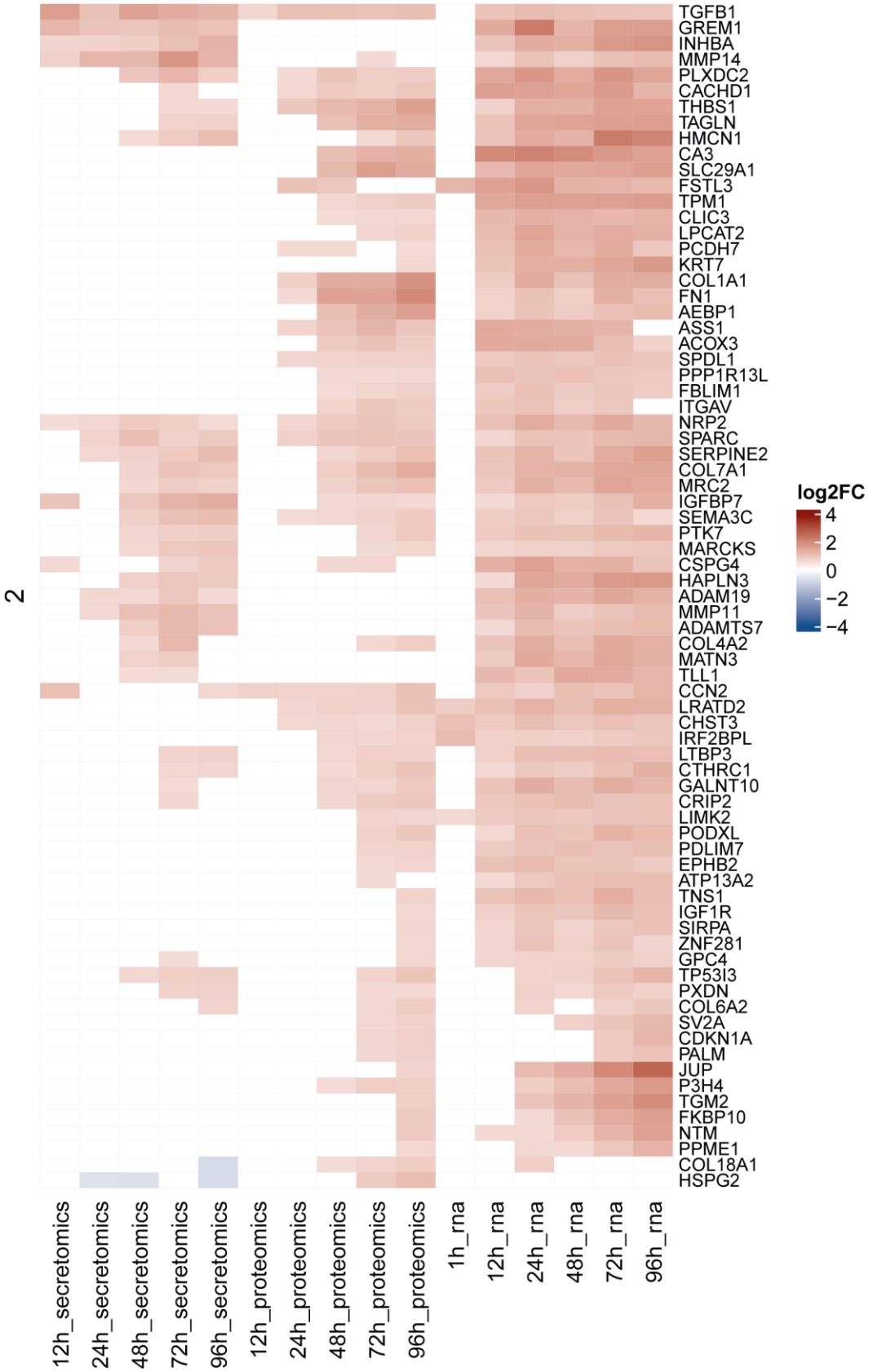

Figure 30 Heatmap of overlapping and upregulated factors defined as hits in at least two modalities across the time points tested. Heatmap summarises factors that are significantly upregulated in at least two out of the three omics modalities (secretomics, proteomics, transcriptomics = rna). Significant upregulation is defined as log2 fold-change > log(1.5 (or 2 for transcriptomics)) and adjusted p-value < 0.05, and is illustrated by red for upregulated factors and blue for downregulated factors. Non-significantly deregulated values are shown in white.

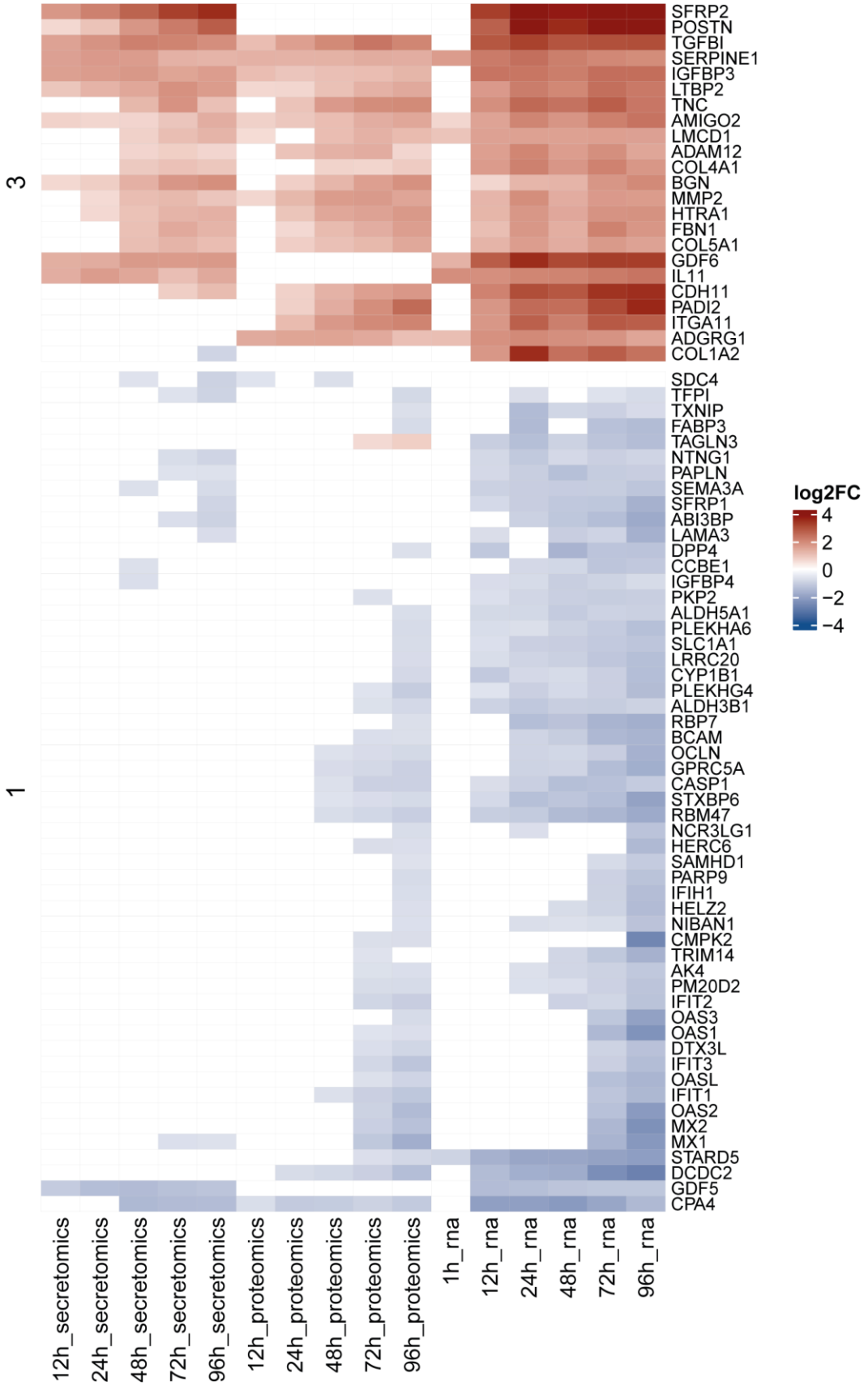

Figure 31 Heatmap of overlapping factors defined as hits in at least two modalities across the time points tested. Heatmap summarises factors that are significantly deregulated in at least two out of the three omics modalities (secretomics, proteomics, transcriptomics = rna). Significant deregulation is defined as absolute log2 fold-change > log(1.5 (or 2 for transcriptomics)) and adjusted p-value < 0.05,

and is illustrated by red for upregulated factors and blue for downregulated factors. Non-significantly deregulated values are shown in white.

Similar to the graph above, the heatmap below (**Figure *32***), summarizes significantly deregulated phosphosites, which were divided into four clusters. Cluster one includes factors that were downregulated. In comparison to this, cluster two contains hits that were only significantly downregulated after 96 h of TGF-β treatment. In cluster three, factors that were upregulated over time or at the last time point are shown, while cluster four highlights hits that were upregulated at earlier time points.

When comparing these hits to the other data modalities, some hits were shared, for instance CDK18, PTK2B, PARD3B, PPL and CRB2 were also significantly downregulated in the transcriptomics data, however, mostly from earlier time points on. Of note, EPHB2, which phosphorylation was decreased upon treatment with TGF-β, showed a significantly higher mRNA expression. **TNS1**, a protein that is essential in myofibroblast differentiation, showed an upregulated mRNA expression and phosphorylation (Bernau et al., 2017). The same was true for **ADAM12** (Nakamura et al., 2020; Sobecki et al., 2022), a disintegrin metalloprotease, which was further upregulated in the proteomics data.

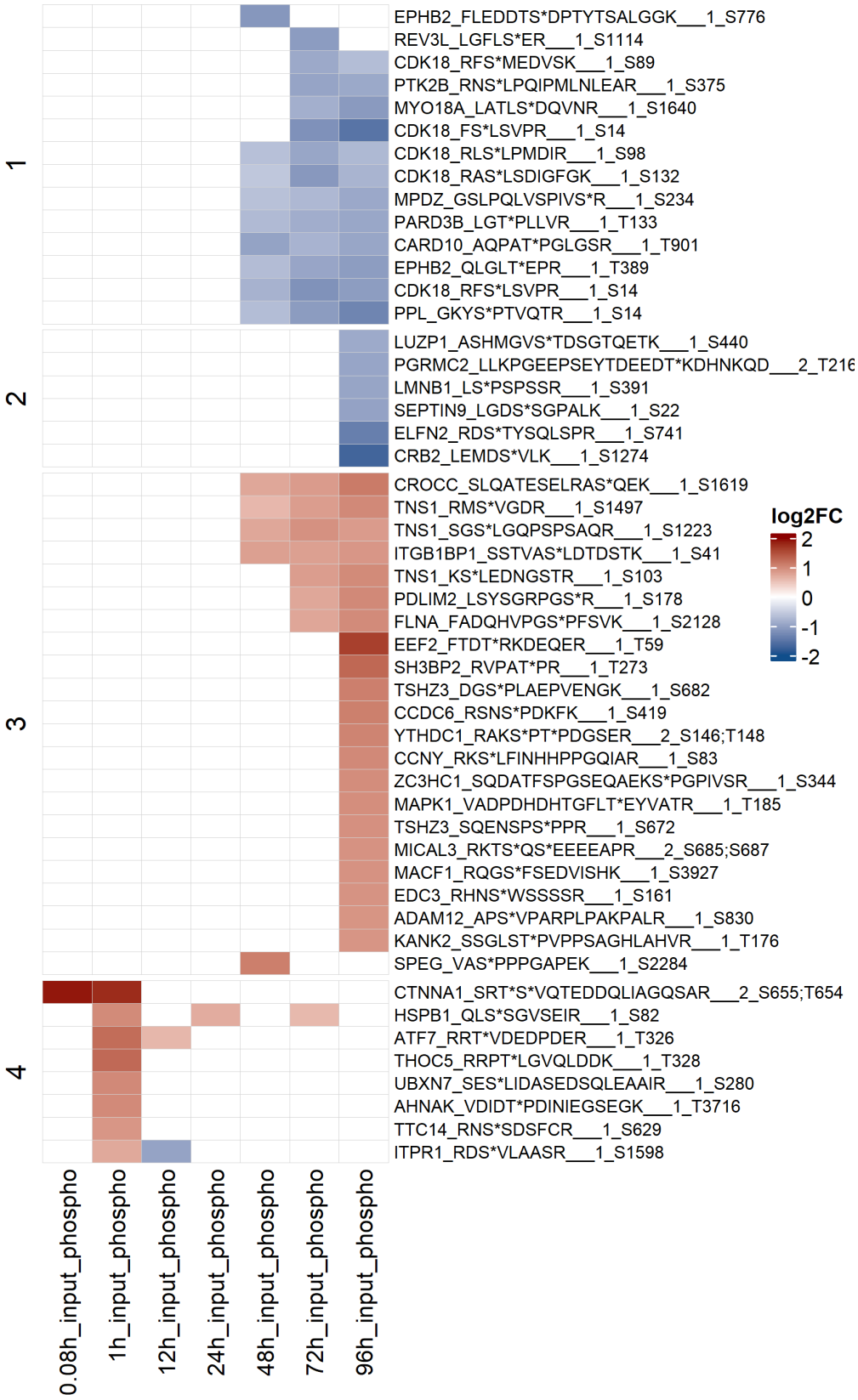

Figure 32 Heatmap of phosphoproteomic hits over time. This heatmap summarizes the top 50 phosphosites that are significantly deregulated in the phosphoproteomics data (normalized to total proteomics). Significant deregulation is defined as absolute log2 fold-change > log(1.5)adjusted p-value < 0.05, and is illustrated by red for upregulated factors and blue for downregulated factors. Non-significantly deregulated values are shown in white. Furthermore, the factors were divided into four clusters.

Having a closer look into which secretomics hits were part of which pathway, I am highlighting a few hits (absolute log2 fold-change > log(1.5), adjusted p-value < 0.05) here. One of the pathways that was upregulated already at early time points after TGF-β treatment, was **epithelial to mesenchymal transition**. Many of the factors involved in this pathway were dysregulated in the secretomics data including BGN, CCN2/CTGF, CDH11, COL4A2, COL4A1, COL5A1, FBN1, GREM1, HTRA1, IGFBP3, MMP14, POSTN, SERPINE1/PAI1, SERPINE2, SPARC, TGFBI, and TNC. Furthermore, several relevant metalloproteases, due to their ECM regulatory function, showed a change in protein abundance post-TGF-β treatment, e.g., ADAM12, HTRA1, MMP2, and MMP14. Additionally, core matrisome glycoproteins such as FBN1, HMCN1, IGFBP3, IGFBP7, POSTN, SPARC, THBS1, and TNC, as well as proteoglycans like BGN and HSPG2 and collagens including COL4A1, COL4A2, and COL5A1 were increased upon stimulation with TGF-β.

Having a closer look into the expression of COL1A1, it was downregulated in the secretomics data while it was upregulated in the proteomics and transcriptomics, with increasing log2 fold-change over time (Figure *30*), COL1A1 was not shown in the secretomics samples in the heatmap above since it did not meet the filtering criteria and was only significantly deregulated in proteomics from 24 h onwards. Therefore, the expression changes are summarized in the heatmap below (Figure *33*).

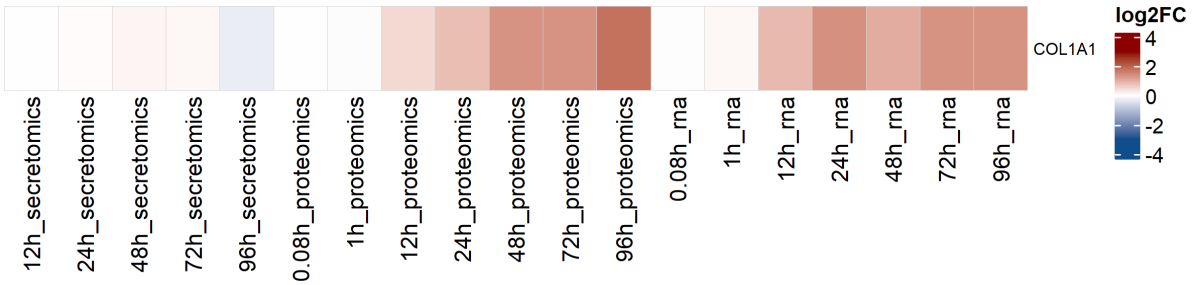

Figure 33 Log2 fold-change of COL1A1 in the different omics modalities over time. Over time, the log2 fold-change calculated between TGF-β treatment and control at each time point is shown for each data modality. While expression of COL1A1 is increased in the transcriptomics data, significantly increased in the proteomics data, it is downregulated in the secretomics data.

4.3. Activity Inference and Mechanistic Modeling

In my thesis, we applied COSMOS (Dugourd et al., 2021), a tool developed in the Saez group to extract mechanistic insights from complex datasets like multi-omics data. By using prior knowledge and the omics data generated in this study, it added another layer to connect and systematically interpret the obtained data by inferring causal links between the data modalities.

4.3.1. Activity Inference of Transcription Factors and Kinases/Phosphatases

Arguing that the changes of mRNA abundance and phosphorylation can be explained by modified activities of TFs and kinases/phosphatases, I initially inferred the activity of transcription factors based on known target gene expression changes. This can be compared to pathway inference, where differential expression of factors that are involved in a pathway are used to determine the activity of that pathway. Similarly, changes in phosphorylation of peptides were as a proxy for kinase and phosphatase activities (Dugourd et al., 2021; Dugourd & Saez-Rodriguez, 2019).

Transcription Factor Activity Inference

The following heatmap (Figure *34*) depicts the most active and inactive TFs over time, based on the changes of target transcripts in TGF-β treated samples in relation to their respective controls. These include TFs, whose function was well characterized in fibrosis e.g. SMAD4, SRF, TEAD4 and ETS1, as well as TFs, whose activity increase over time as summarized in cluster five.
In cluster four, TFs exhibited peak activity at 1 h post TGF-β treatment followed by a decline, examples include KMT2A, NFKB1, and SMAD1. TFs, whose activity was lowest at 1 h but then increased over time are shown in cluster three, for instance HIF1A, NR2F2, and STAT4. Cluster two comprises TFs, whose activity was reduced over time by TGF-β stimulation. The initial peak in activity of TFs in cluster one, should be taken with care as there were no significant differentially expressed transcripts (when setting log2 fold-change and adjusted p-value cut-offs) at that time (5 min). TFs in this cluster showed a reduced activity over time, similar to cluster two, e.g. GLI1, E2F4, and STAT2.

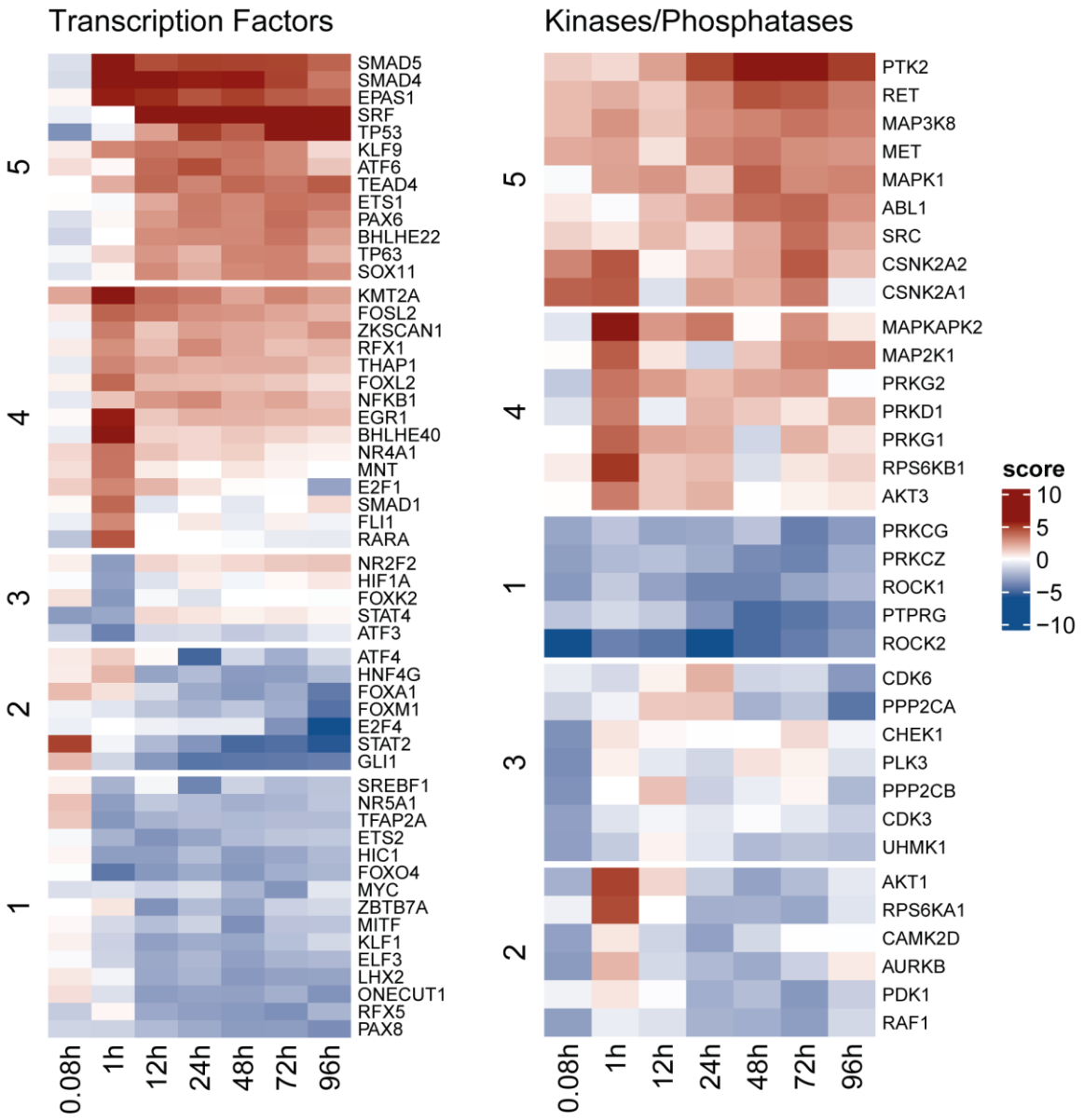

Figure 34 Transcription factor and kinase/phosphatase activity inference. Left: The activity of transcription factors (TFs) was inferred based on the mRNA expression changes of down-stream targets. **Right:** The deregulation of phosphosites was taken as a proxy for kinase and phosphatase activity. More active TFs and kinases/phosphatases are shown in red while more inactive ones are highlighted in blue.

Based on the activity profile, I chose six transcription factors for validation, as I saw their potential as modulators of fibrosis. Assuming that an inhibition of the initial activity of the selected TFs would lead to a reduction in ECM and a change in target genes. The following figure summarizes the target genes two of the selected transcription factors (Figure *35*). The depicted target genes were differentially expressed genes taken from the transcriptomics data. In red, hits that were upregulated are shown while blue indicates a decrease of mRNA expression. To bring an example, target genes of SMAD1 included ID2 and ID3, which were initially upregulated. Based on these findings, the activity of SMAD1 at 1 h post-TGF-β treatment was upregulated while its

activity decreased over time. Validation of selected TFs is shown in Results section 4.4.

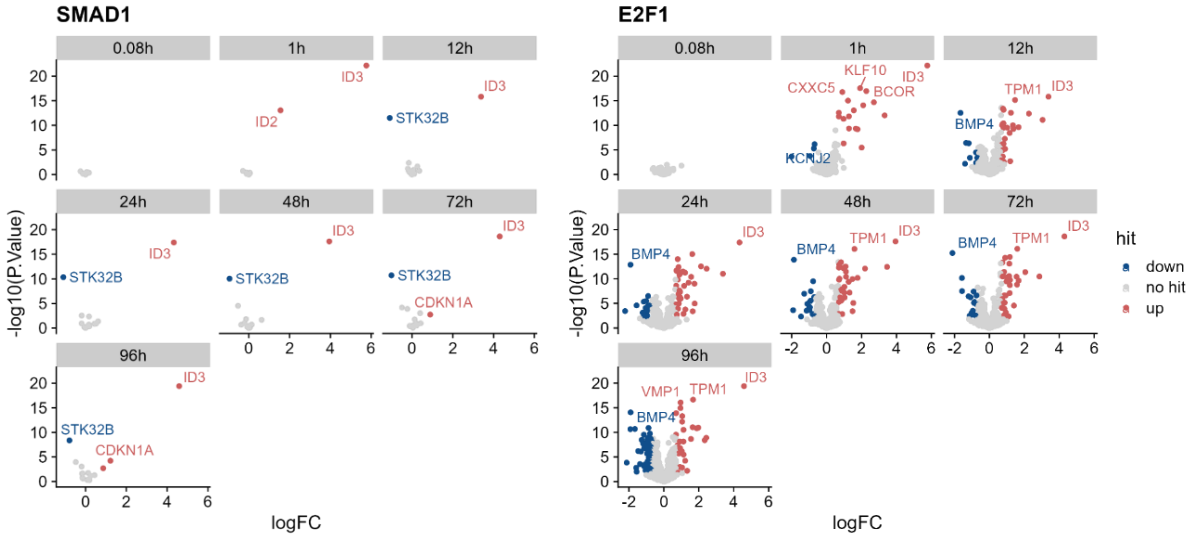

Figure 35 Differential target genes of SMAD1 and E2F1. Predicted downstream targets of SMAD1 (left) and E2F1 (right) are shown across the time points. The log2 fold-change of the genes is plotted on the x-axis against the -log10(P.Value) on the y-axis. Grey dots depict genes that are not significantly deregulated, blue dots indicate that these genes are significantly 'deactivated' by E2F1, while red highlights 'activated genes'.

Kinase/Phosphatase Activity Inference

Similar as for the activity inference of TFs, I used the phosphorylation of peptides as proxy for kinase and phosphatase activity. The activity of following kinases and phosphatases is illustrated in the heatmap above (**Figure *34***).

In cluster one, kinases which activity was decreased over time are shown. Cluster two and three contain kinases which activity was generally decreased but peaked at one or the other time point. Cluster five includes kinases that showed an increased activity over time while cluster four shows kinases which activity peaked at 1 h post-TGF treatment.

4.3.2. Selection of Time Points for Early and Late Network

To represent the changes over time, we decided to generate two networks, one that summarized the signaling network of the early response to TGF-β and a second network for the late responses that included the output of the early network as input to understand possible feedback mechanisms. However, to decide which time points should be included in either network, I performed clustering of hits and correlation of the data modalities, described in more detail in this section.

Clustering of Significantly Deregulated Features (Hits) Per Omics Modality

Using Neural Gas clustering, I examined the distribution of significantly deregulated features (hits, absolute log2 fold-change > log(1.5), adjusted p-value < 0.05) per omics modality. Even though two clusters were suggested for all modalities, as evaluated based on the elbow method, I divided the data into four clusters, as I saw that this represented the data better.

I divided the transcriptomics data into four clusters (Figure *36*), with the first and third cluster containing genes that showed an increased abundance over time, whereas the increase of factors in cluster one was less steep. A similar pattern can be seen for cluster two and four, in which there was a decrease of mRNA abundance of factors contained in the clusters over time. The number of factors in each cluster varied, with 1053 factors in cluster one, 436 in cluster two, 179 in cluster three and cluster four containing 1285 differentially regulated transcripts.

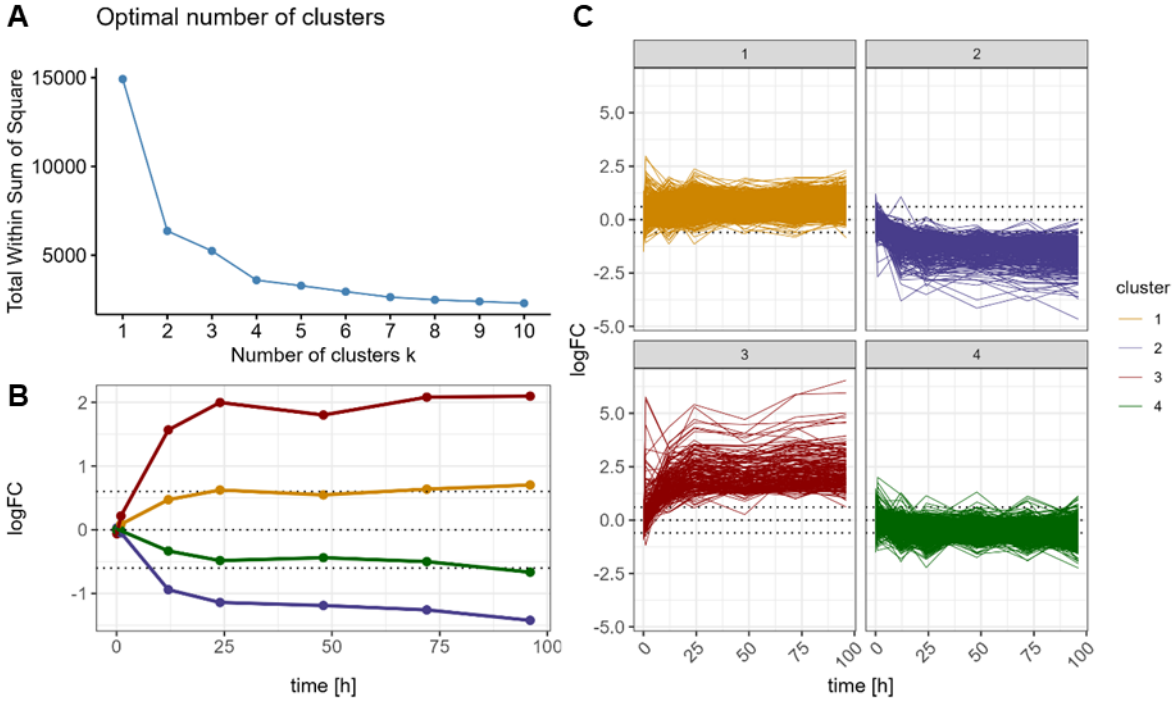

Figure 36 Clustering of significantly deregulated transcripts. A) Elbow method for graphical representation of optimal number of clusters. **B)** The median log2 fold-change for each cluster and time point are shown with lines representing the trend of log2 fold-change (y-axis) over time (x-axis). **C)** Each subplot represents a single cluster and lines depict the log2 fold-change per factor included in the cluster. Points and lines are color coded for each cluster.

Looking at the clusters for the proteomics data (Figure *37*), cluster one contained proteins with decreased expression over time while expression of factors in cluster two was increased over time. The average decrease in abundance changes of proteins in cluster three was visible at earlier time points than in cluster one. Similar to cluster two,

the factors in cluster four showed an increased expression over time which was however observed earlier on and generally resulted in a higher log2 fold-change than of factors in cluster two. Since there were less hits in the proteomics data, the clusters generally contained less features. Therefore, 62 proteins were contained in cluster one, and 111 in cluster two. Cluster three included 12 factors, while 28 factors were taken into account for cluster four.

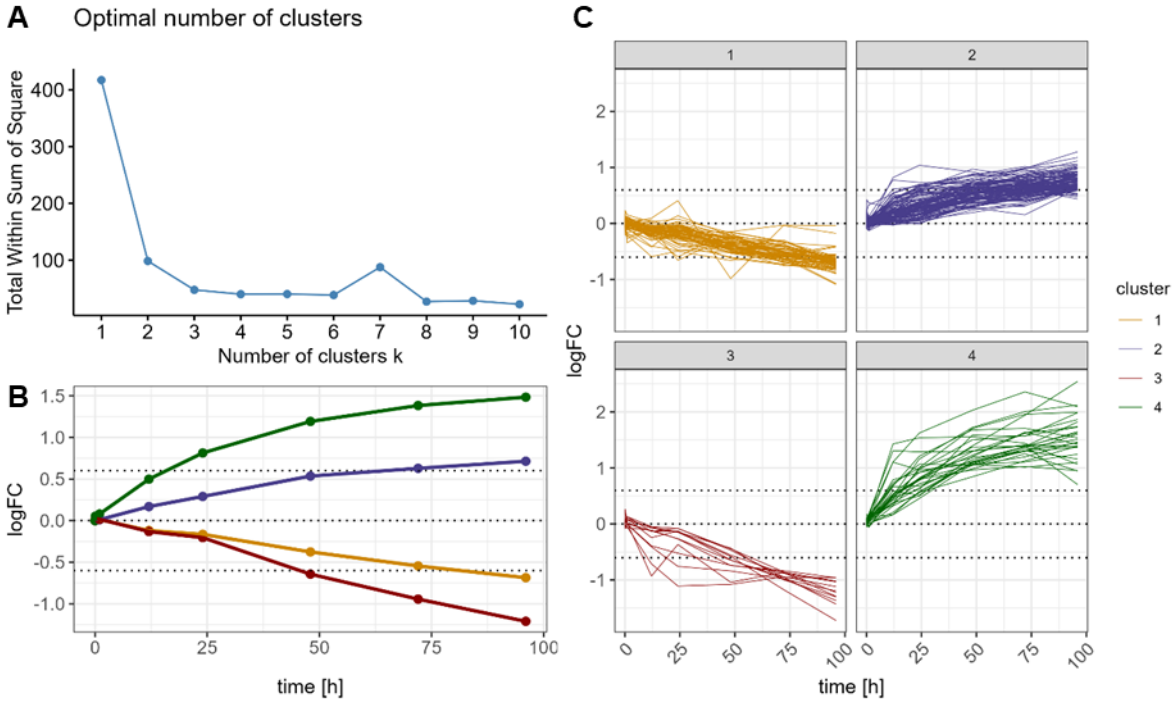

Figure 37 Clustering of significantly deregulated proteins in the proteomics dataset. A) Elbow method for graphical representation of optimal number of clusters. **B)** The median log2 fold-change for each cluster and time point are shown with lines representing the trend of log2 fold-change (y-axis) over time (x-axis). **C)** Each subplot represents a single cluster and lines depict the log2 fold-change per factor included in the cluster. Points and lines are color coded for each cluster.

For secretomics (Figure *38*), there were also four clusters, two of which showed an increased expression of proteins, namely cluster one and two. Whereas there was a steep increase of expression of factors contained in cluster two that was visible already 12 h post TGF-β treatment and afterwards rather stayed constant over time, there was a slow increase of expression of proteins in cluster one. Proteins in cluster three showed generally a decrease in expression over time. Cluster four contained proteins which expression was first increasing and then decreasing after 24 h of treatment, however the behavior of individual factors may differ from the average change. The number of features per cluster was as follows: 84 features in cluster one, 14 in cluster two, 95 for cluster three and in cluster four there were 42 proteins included.

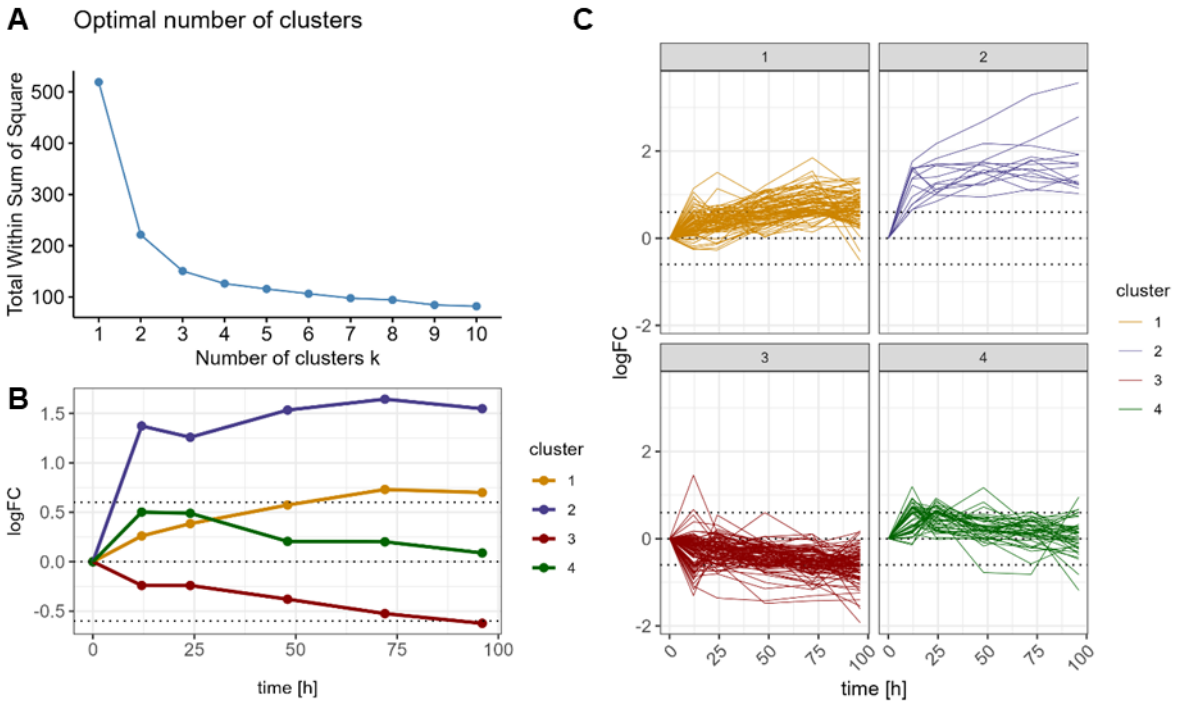

Figure 38 Clustering of significantly deregulated proteins in the secretomics dataset. A) Elbow method for graphical representation of optimal number of clusters. **B)** The median log2 fold-change for each cluster and time point are shown with lines representing the trend of log2 fold-change (y-axis) over time (x-axis). **C)** Each subplot represents a single cluster and lines depict the log2 fold-change per factor included in the cluster. Points and lines are color coded for each cluster.

For clustering of the phosphoproteomics data, I used the data that was abundance corrected using the full proteome data (Figure 39). The 75 factors of cluster one showed a general decrease in expression over time. Cluster two contained 42 factors that generally showed a decreased expression after 1 h of TGF-β treatment, and increased expression at the 24 h time point and then a decline in abundance over the later time points. While an increased log2 fold-change was observed for factors in cluster three 1 h post TGF-β treatment, the abundance of factors decreased until 24 h after TGF-β treatment and stayed at a constant log2 fold-change over time. Cluster three contained 31 factors whereas 106 factors were contained in cluster four. The latter showed that factors in this cluster generally show an increased phosphorylation over time.

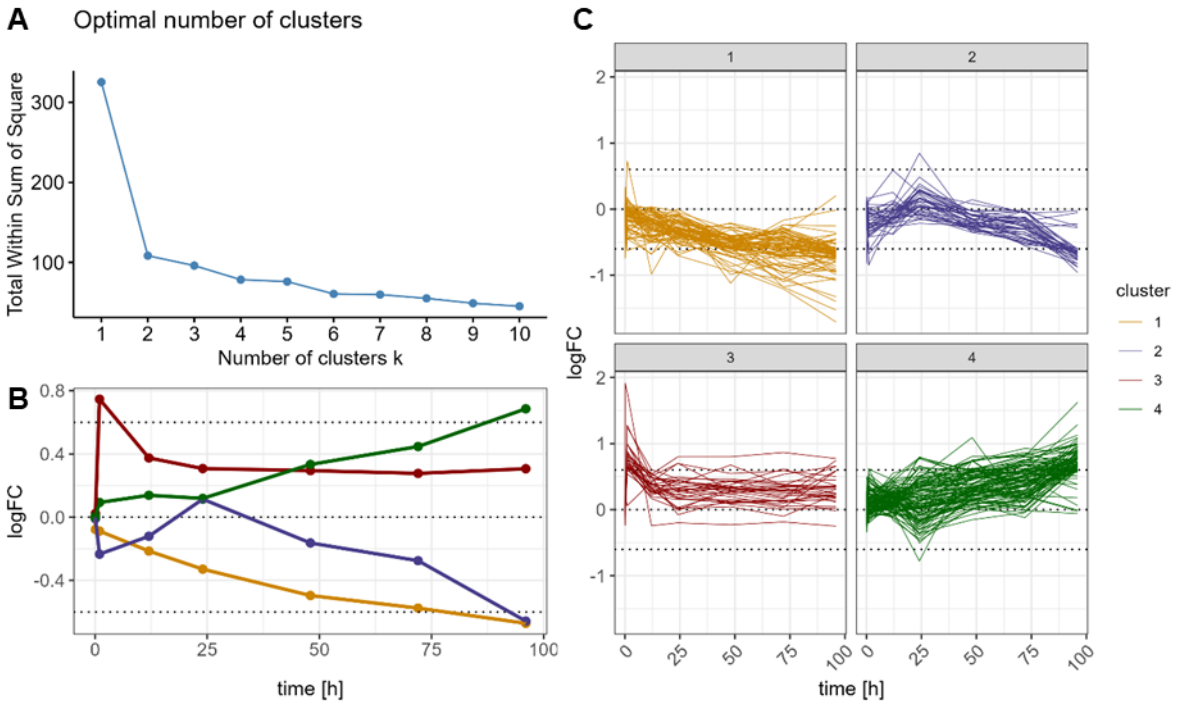

Figure 39 Clustering of significantly deregulated factors in the phosphoproteomics dataset. A) Elbow method for graphical representation of optimal number of clusters. **B)** The median log2 fold-change for each cluster and time point are shown with lines representing the trend of log2 fold-change (y-axis) over time (x-axis). **C)** Each subplot represents a single cluster and lines depict the log2 fold-change per factor included in the cluster. Points and lines are color coded for each cluster.

Correlation of Omics Modalities

Furthermore, the correlation of abundance changes across different data modalities was examined, excluding the phosphoproteomics as data mapping sites to genes blurs the signal and phosphorylation changes were not correlated with abundance changes. Therefore, I generated the graphical representation of the correlation matrix below (Figure *40*) with the top 10% variable features per omics modality. The 0.08 h transcriptomics (rna) data generally exhibited anti-correlation with other time points and modalities. Conversely, at the 1 h transcriptomics time point, the most significant correlation was observed with the 12 h proteomics and secretomics data. Subsequent transcriptomics data (from 12 to 96 h) demonstrated strong correlation with proteomics and secretomics data, particularly at time points ranging from 12 to 96 h.

The 0.08 h and 1 h proteomics data displayed strong correlation with each other but exhibited weaker correlation with other time points and data modalities. Conversely, longer time points exhibited robust correlation with each other and with other data modalities. With the exception of the 12 h and 96 h secretomics data, which displayed weaker correlation, the remaining data demonstrated consistent correlation across time points and modalities.

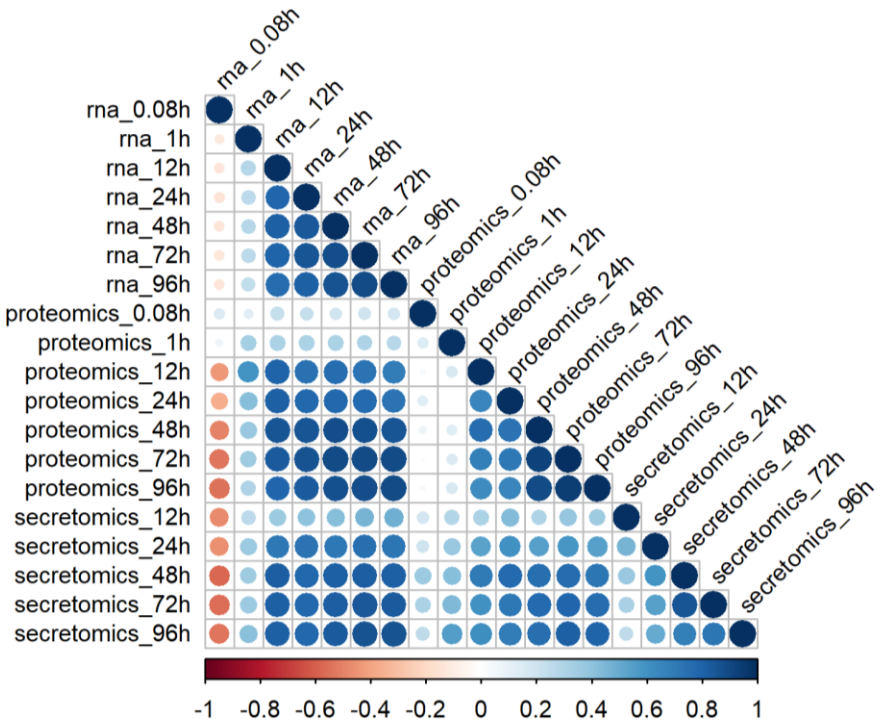

Figure 40 Correlogram representing Pearson correlation coefficients computed between time points and omics modalities. Correlation coefficients are colored according to the value, ordered based on omics modality and time of treatment (with TGF-β).

4.3.3. Network Set-up and Comparisons

In the previous section of this thesis, the changes of mRNA transcripts and phosphosites were taken as a proxy for activity of TFs and kinases/phosphatases, respectively. In this section, the top 30% of deregulated TFs and kinases/phosphatases were used as inputs for generating the network, offering a good balance of model size and complexity. A network modeling analysis allows for the generation of an integrated model of signaling upon TGF-β treatment spanning all obtained modalities. To account for time-dependent effects, I chose to separate the data into an early and a late response.

Based on the observations gained from clustering of the data and correlation matrix, I defined the time points for the early network to be as follows: 0.08, 1 and 12 h for the TFs and kinases/phosphatases as well as 12 and 24 h for the secretomics hits. The late network comprised 24 to 96 h of TFs and kinases/phosphatases and 48 to 96 h of the secretomics hits.

Since I initiated the fibrotic response with TGF-β, we took it as upstream starting node to model the early network. I hypothesized that stimulation with TGF-β leads to changes in activity of TFs and kinases/phosphatases, as intermediate mediators, that

further drive expression changes of secreted proteins. For the late network, TGF-β and the early secretomics hits were used as input, being linked to changes of TFs and kinases that further drive the expression of late secretomics hits (Figure *41*).

COL1A1 was not included in the network initially, as it was not a hit in the secretomics. However, based on the imaging data, that showed a clear increase of extracellular Col I over time, we included it as one of the hits in the secretomics data that was used for the network. The following data were the result of an optimized mechanistic network that Mira Burtscher generated.

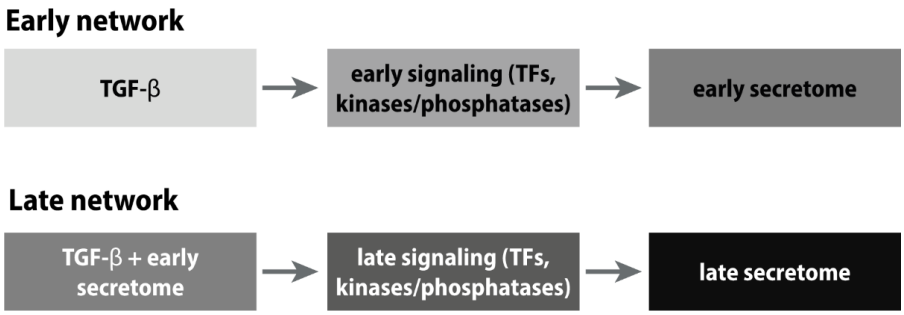

Figure 41 Graphical representation of network set-up. For the early network, TGF-β stimulation serves as the primary input to the network, triggering signaling cascades involving transcription factors (TFs), kinases and phosphatases, which induce changes in the early secretome. Subsequently, for the late network, accounting for autocrine signaling, the early secreted factors, alongside TGF-β stimulation were included. Induced by the upstream signals, activities of TFs, kinases and phosphatases resulting in alterations of secreted factors. Thereby, the networks describe the interplay between TFs, kinases/phosphatases and secreted factors, initiated by TGF-β stimulation.

In general, there were a similar number of edges and nodes in the networks. The network for the early signaling contained 127 nodes with 156 edges, whereas the late network had 160 nodes and 202 edges.

Pathway enrichment analysis of both networks, using the Reactome database, resulted in the figure below (Figure *42*). While most of the pathways were more significantly enriched in the late network, the following pathways were more significant in the early network including signaling by interleukins, diseases of signal transduction by growth factor receptors and second messengers, PIP3 activates AKT signaling, Intracellular signaling by second messengers, and signaling by GPCRs. Among the most important pathways, as shown by the enrichment analysis, were Signaling by receptor tyrosine kinases, signaling by interleukins, extracellular matrix organization, signaling by TGFB family members, signaling by NTRKs and ECM proteoglycans.

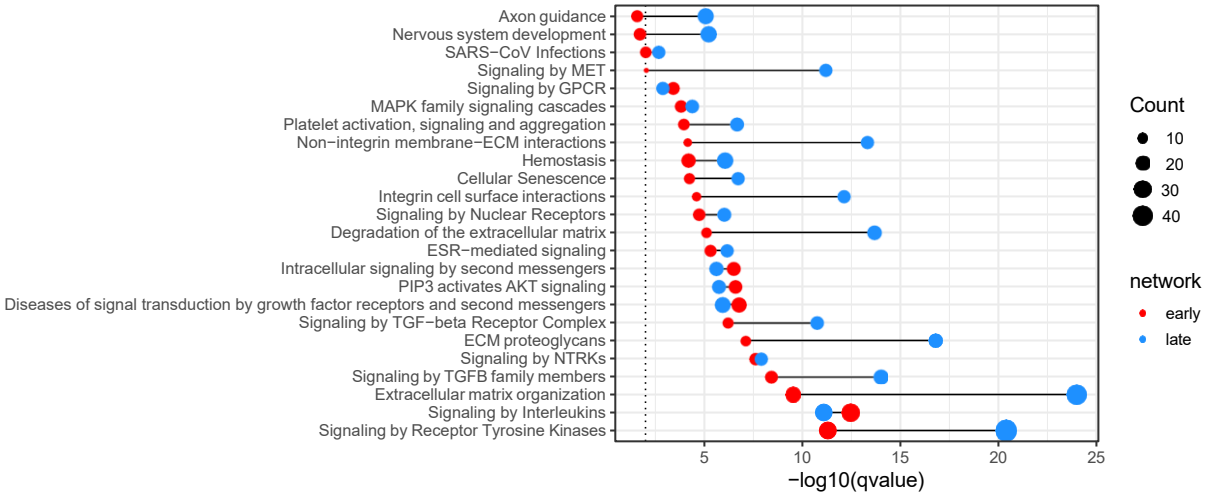

Figure 42 Pathway enrichment analysis comparing the early vs late network. Pathways are depicted on the y-axis, while the -log10(qvalue) is represented on the x-axis. The qvalue reflects the adjusted p-value (FDR) for multiple hypothesis testing. The intercept at the x-axis marks the threshold set at -log10(0.01), corresponding which is 2 in the graph. Size of the dots indicates the number of nodes found in each pathway, while the color distinguishes whether the data belongs to the early or the late network.

Early Network

To further dissect the information contained in these mechanistic networks, I employed fast & greedy clustering of the networks. For the early network, we defined six clusters (Figure *43*). Pathway enrichment was then performed for each cluster and the most significant pathways are shown in Figure *43*.

The most significant pathways of the first cluster included Extra-nuclear estrogen signaling, Interleukin-4 and Interleukin-13 signaling as well as Signaling by receptor tyrosine kinases. Cluster two contained factors that were enriched for the pathways Degradation of the extracellular matrix, Extracellular matrix organization, Interleukin-4 and Interleukin-13 signaling, Regulation of CDH11 Expression and Function, Signaling by receptor tyrosine kinases and Transcriptional regulation by RUNX3. Macroautophagy, Selective autophagy, and KEAP1-NFE2L2 pathway were enriched in cluster three. In cluster four, the pathways Signaling by NTRKs, Signaling by NTRK1 (TRKA), Nuclear Events (kinase and transcription factor activation) as well as Signaling by Receptor Tyrosine Kinases were enriched. Cluster five contained factors that were included in Degradation of the extracellular matrix, Extracellular matrix organization, CTLA4 inhibitory signaling and RAF activation. Whereas the pathways Pre-NOTCH Expression and Processing, Pre-NOTCH Transcription and Translation, and Transcriptional regulation by RUNX3 were significantly enriched in cluster six.

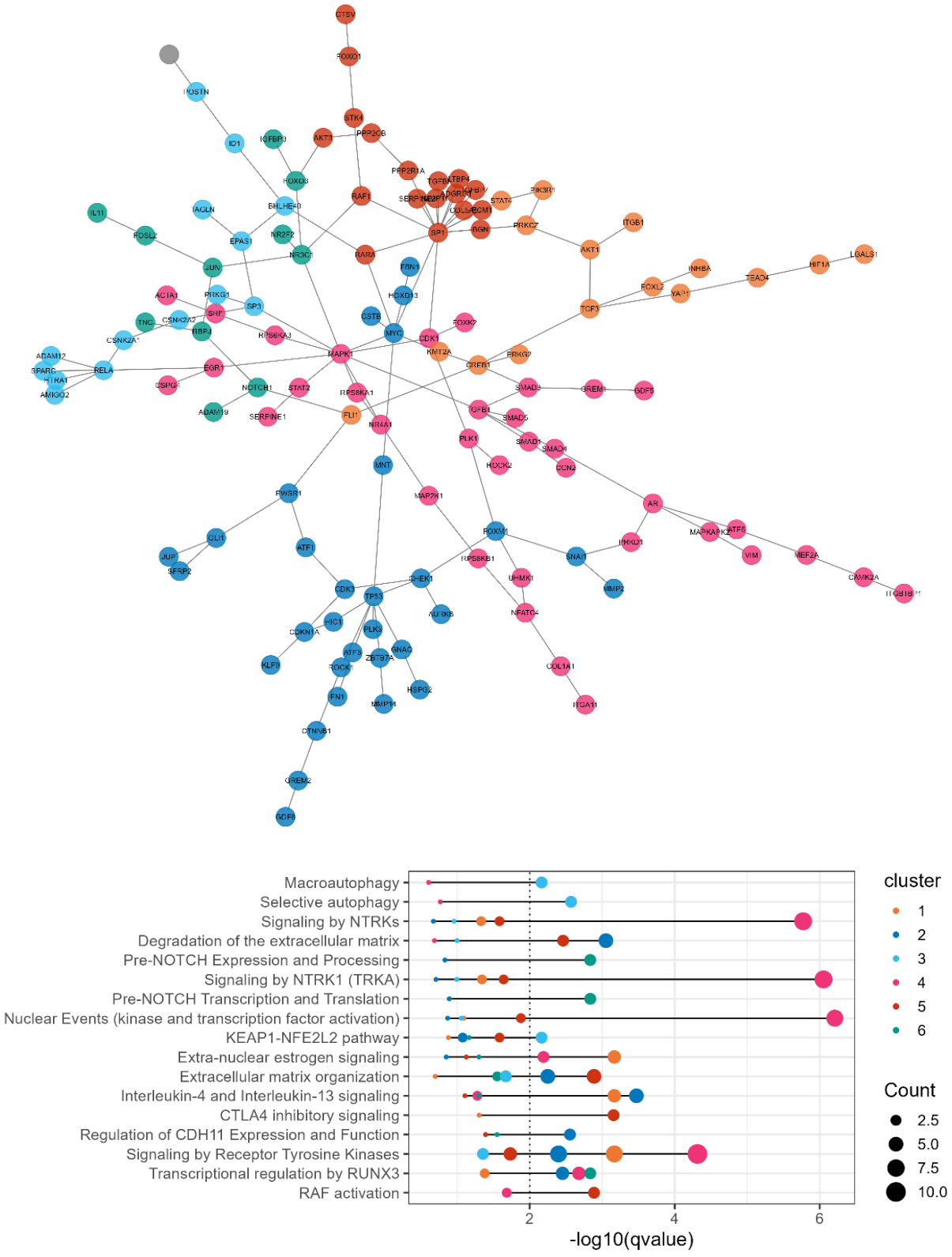

Figure 43 Early network with clustering of nodes and pathway enrichment analysis performed per cluster. Pathways are depicted on the y-axis, while the -log10(qvalue) is represented on the x-axis. The qvalue reflects the adjusted p-value (FDR) for multiple hypothesis testing. The intercept at the x-axis marks the threshold set at -log10(0.01), corresponding which is 2 in the graph. Size of the dots indicates the number of nodes found in each pathway, while the color distinguishes whether the data belongs to the early or the late network.

Late Network

Since the late network contained more nodes and edges, we divided the data into 8 clusters (Figure *44*). Cluster one exhibited enrichment in pathways such as Hemostasis, Cellular Senescence, Platelet activation, signaling and aggregation, Extracellular matrix organisation, Degradation of the extracellular matrix, Signaling by TGFB family members, Signal transduction by L1 along with Signaling by Receptor Tyrosine Kinases. Cluster two comprised factors enriched in pathways related to Platelet activation, signaling and aggregation, Extracellular matrix organization, signaling by BMP, Signaling by TGFB family members along others. In cluster three, enrichment was observed in pathways associated with Post-translational protein phosphorylation, Signaling by Interleukins, Signal transduction by L1, Signaling by Receptor Tyrosine Kinases, and Interleukin-6 family signaling. Cluster four was characterized by enrichment in pathways such as RHO GTPase Effectors, Hemostasis, Platelet activation, signaling and aggregation and Signaling by Receptor Tyrosine Kinases. Cluster five demonstrated involvement in pathways including Extracellular matrix organization, Degradation of the extracellular matrix, Regulation of beta-cell development, Formation of axial mesoderm, RUNX3 regulates WNT signaling. Pathways notably enriched in cluster 6 included Signaling by Interleukins, Zygotic genome activation (ZGA), Signaling by Receptor Tyrosine Kinases, Regulation of beta-cell development Formation of axial mesoderm, as well as Regulation of gene expression by Hypoxia-inducible Factor. In cluster seven, there was no significant enrichment observed in any pathway, as set by our threshold for the qvalue of -log10(0.01). Finally, cluster eight was notably enriched in pathways such as Extracellular matrix organization, Degradation of the extracellular matrix, Signaling by TGFB family members, and Collagen degradation.

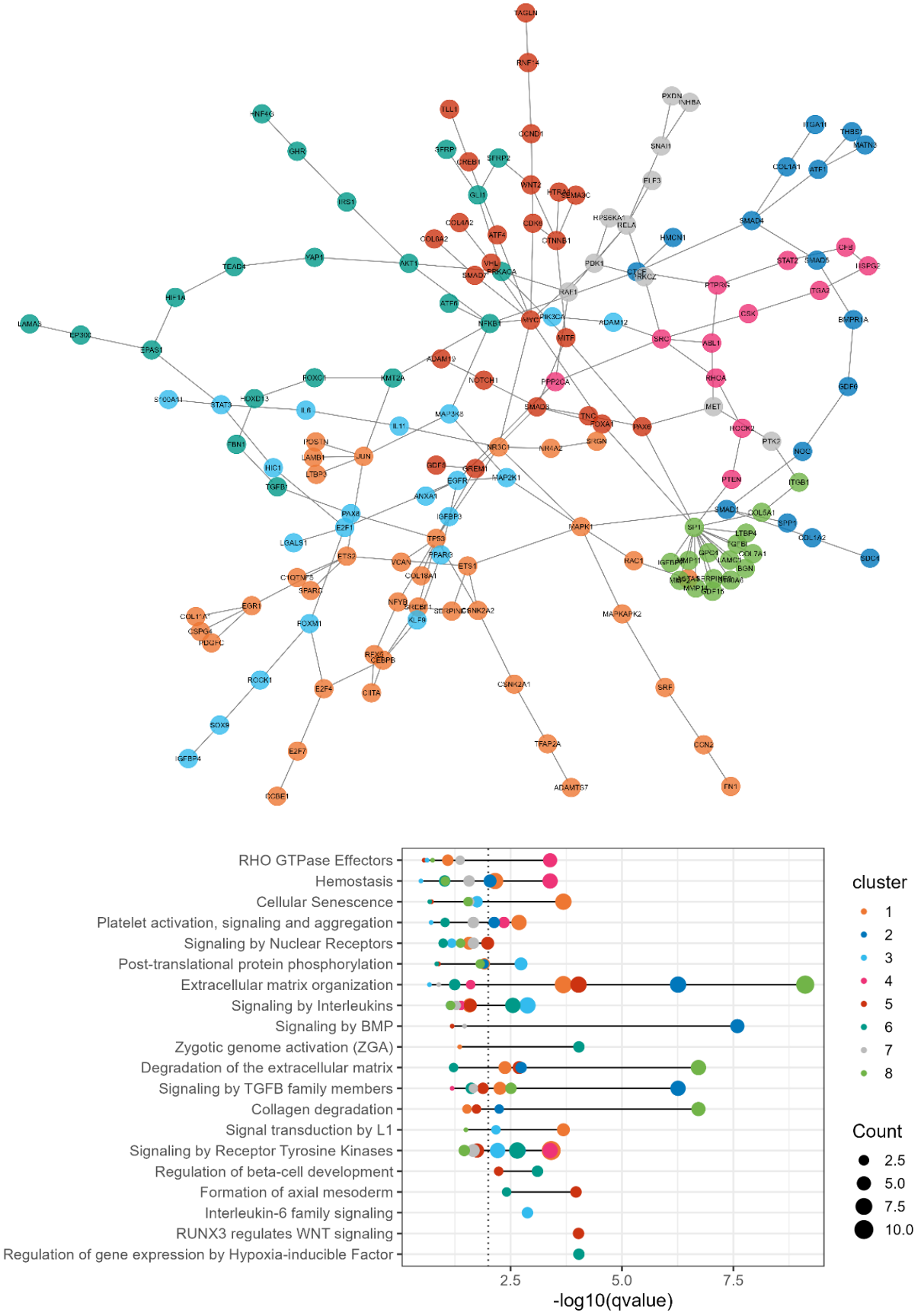

Figure 44 Late network with clustering of nodes and pathway enrichment analysis performed per cluster. Pathways are depicted on the y-axis, while the -log10(qvalue) is represented on the x-axis. The qvalue reflects the adjusted p-value (FDR) for multiple hypothesis testing. The intercept at the x-axis marks the threshold set at -log10(0.01), corresponding which is 2 in the graph. Size of the dots indicates

the number of nodes found in each pathway, while the color distinguishes whether the data belongs to the early or the late network.

The pathways Degradation of the extracellular matrix, Extracellular matrix organization and Signaling by Receptor Tyrosine Kinase were significantly enriched in clusters of the early as well as the late networks. In both, the early and late network, pathway activity enrichment for ECM organisation and degradation included SP1 and predicted downstream targets (Figure *45*). Utilizing prior knowledge, SP1 targets genes, as identified by the network, were found to encompass NABA matrisome core and associated factors, thereby demonstrating enrichment for the aforementioned pathways.

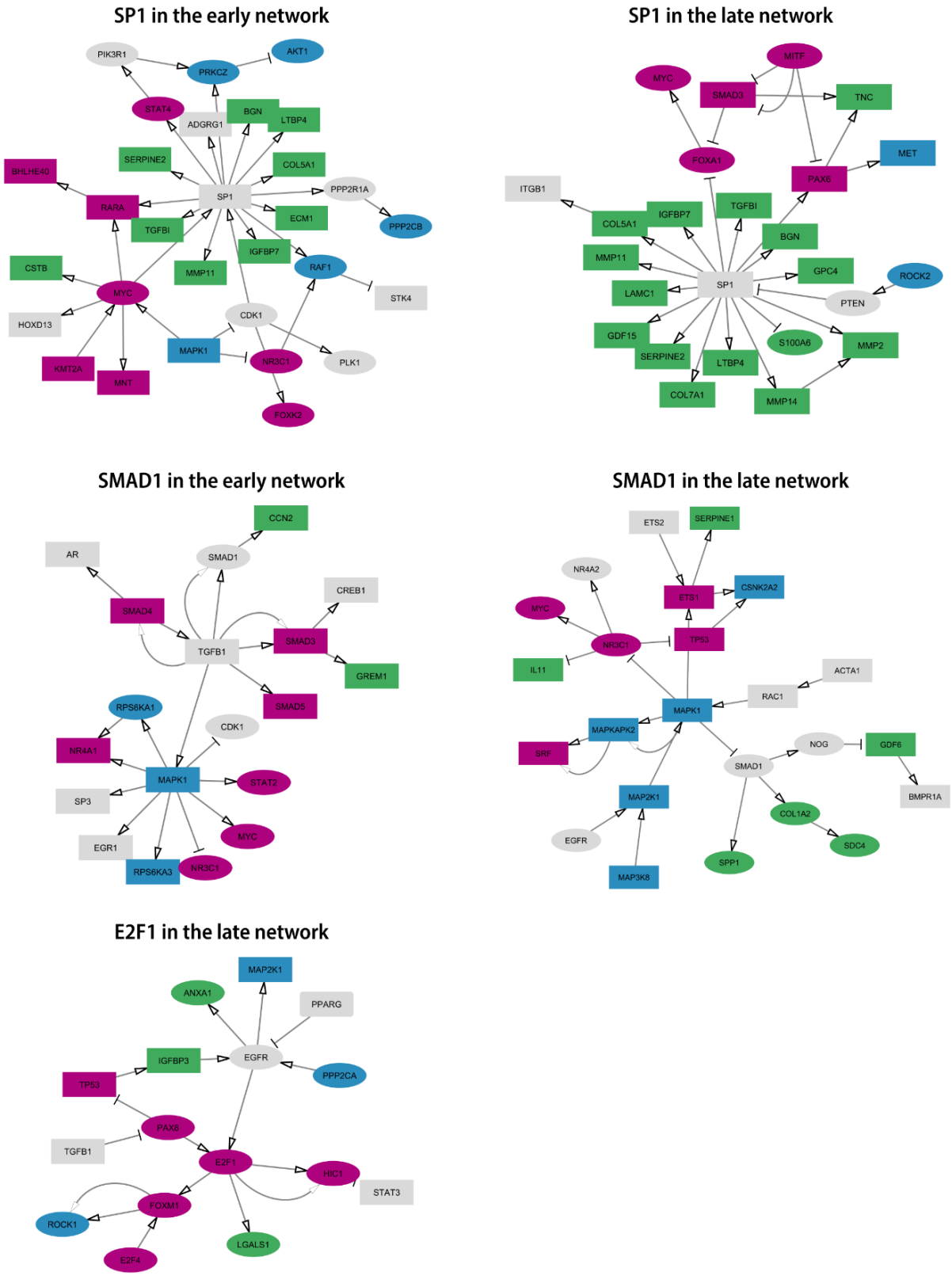

Figure 45 Examples from Networks. Shown are examples for SP1, SMAD1 in the early and late networks and E2F1 for the late network. The colors are representing the different modalities, green for secreted factors, blue for kinases/phosphatases, purple for TFs and grey as factors imputed by prior knowledge. The ellipse shape are factors which expression/activity is downregulated while squares indicate factors that are upregulated.

4.4. Validation of Multi-omics Hits

4.4.1. Setting up the siRNA Transfection Protocol

Initially, I tested a siRNA transfection protocol used in our lab to assess its suitability for the cells used and establish the concentration of siRNA needed for the knock-down. Therefore, I chose to knock-down GM130 (GOLGA2), a cis-Golgi matrix protein. Targeting this protein, gave the advantage that I knew that the siRNAs should work, as they have been widely used in our lab. Additionally, I was able to assess the knock-down at the protein level using an image-based read-out. As SMAD2 is integral of the canonical signaling pathway induced by TGF-β, I further chose to treat cells with siSMAD2 in order to assess the knock-down using western blot analysis.

Knock-down of GM130

I assessed the knock-down of GM130 48 h and 72 h post siRNA transfection. The plates for each time point were imaged separately. In Figure *46*, it is shown that the knock-down was already visible after 48 h and was still present at 72 h post siRNA treatment. The experiment also showed that the lowest amount of siRNA tested yielded a knock-down that led to a significant reduction in protein when compared to the corresponding siNeg9 control.

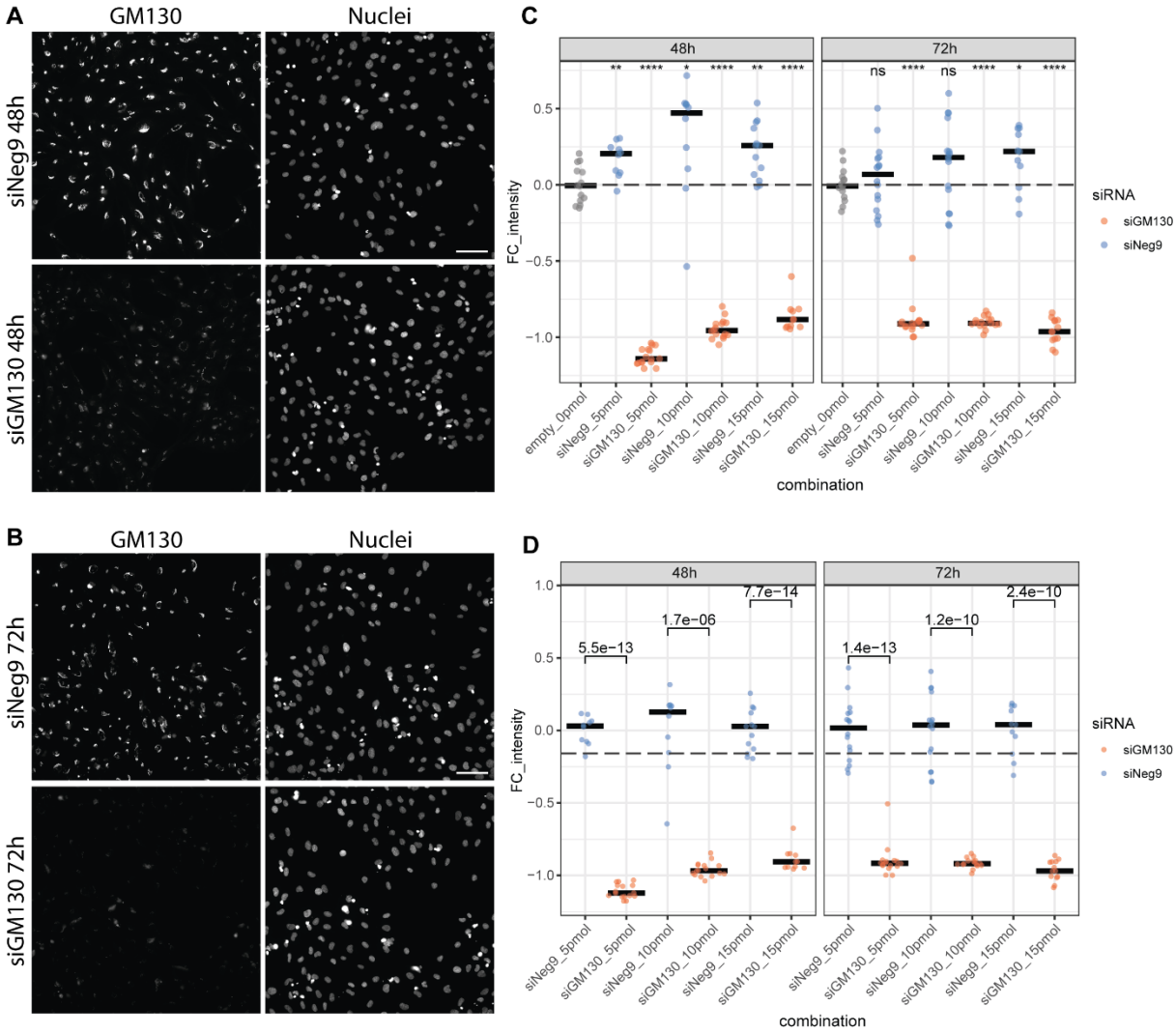

Figure 46 GM130 protein expression changes upon siGM130 transfection. A, B) Widefield images of nuclear and GM130 immunofluorescence staining of cells treated with siNeg9 (control siRNA) or siGM130 for **A)** 48 h and **B)** 72 h. Scale bar = 100 µm. **C)** Change of protein expression (FC_intensity) calculated per image as follows: Intensity of images of cells treated with siGM130 minus intensity of empty vector treated samples divided by empty vector treated samples at 48 h and 72 h post siRNA transfection. **D)** Changes of protein expression (FC_intensity) calculated per image as follows: Intensity of images of cells treated with siGM130 minus intensity of siNeg9 treated samples divided by siNeg9 treated samples. Colors represent siRNA treatment (grey = empty vector, blue = siNeg9, orange = siGM130), dots depict individual images. The x-axis shows combination of siRNA and concentration (pmol, per well of a 24 well plate) of siRNA used. The dashed line shows the intercept for the control group, the bars stand for median intensity change of each sample. n = 1. For statistical analysis, t-test was used. In summary, all concentrations of siGM130 tested resulted in a significant decrease of GM130 protein expression (shown via immunofluorescence staining) at 48 and 72 h post-siRNA treatment.

Knock-down of SMAD2

Furthermore, the siSMAD2 treatment led to a reduction in SMAD2 abundance and a reduced signal for the phospho-SMAD2, after 48 h as well as 72 h of siRNA treatment (*Figure 47*). Importantly, this experiment was conducted using 6 well plates, not like in the knock-down experiment for GM130, where a 24 well plate was used. Hence the concentration of siRNA used was adjusted.

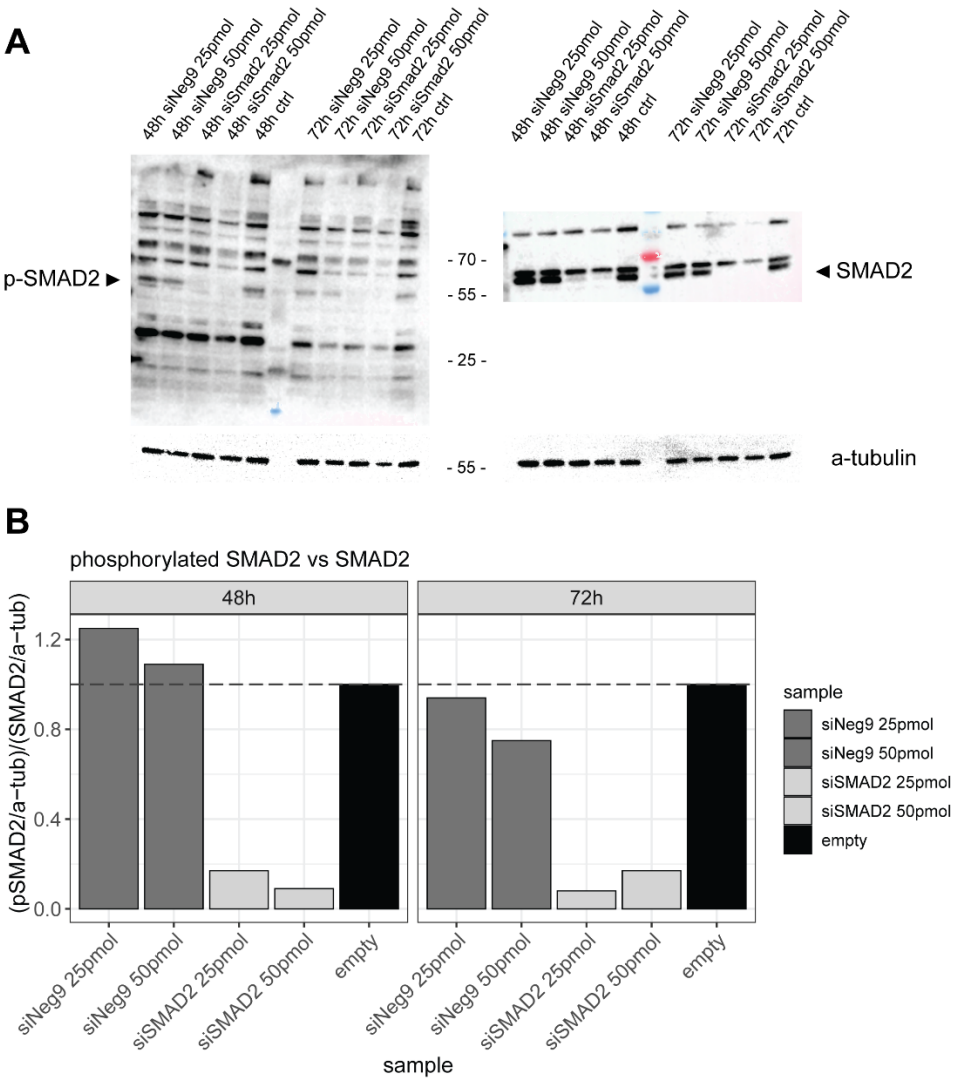

Figure 47 Western blot analysis of SMAD2 and phospho-SMAD2 protein expression after siSMAD2 knock-down. A) Western blot showing protein expression by staining of phospho-SMAD2, SMAD2 and the respective loading controls (alpha-tubulin) for each sample. Cells were treated, in a 6 well plate, with different concentrations of siNeg9 (control) and siSMAD2 for 48 to 72 h. **B)** Quantification of phosphor-SMAD 2 signal normalized to SMAD2 abundance (each normalized to the respective loading control), per time and sample. n = 1. Knock-down of SMAD2 lead to the expected downregulation of SMAD2 and phospho-SMAD2 protein expression after 48 and 72 h of siSMAD2 treatment.

Together, this data showed that the used protocol for siRNA treatment resulted in significant decrease of target genes that was maintained for the time points of interest.

4.4.2. Setting up the Experimental Protocol to Test Target Gene Involvement in ECM Deposition

My strategy was to knock-down genes of interest, aiming to test their effect on downstream targets. The hypothesis was that targeting hits identified in our multi-omics experiment could reveal modulators of ECM regulation, particularly relevant in the field of fibrosis. Furthermore, I aimed to assess the effect of the knock-down on ECM

deposition followed by subsequent TGF-β treatment. Therefore, the experimental protocol included a media change 48 h post-siRNA treatment, to control medium containing macromolecular crowding and ascorbic acid and with additional TGF-β for the treatment condition. After another 48 h of TGF-β treatment (in total after 96 h post-siRNA treatment), the cells were fixed and stained with CNA35, a fluorescent dye binding to fibrillar collagen (Figure *48*). An imaging based-read-out was used to assess the effect on the ECM deposition.

Moreover, I isolated mRNA at 48, 72 and 96 h post siRNA treatment and checked whether the mRNA abundance decreased upon the knock-down. In addition to that, I assessed the mRNA abundance changes of proposed down steam targets. The aim of this was to gain further information of how the knock-down influenced the ECM.

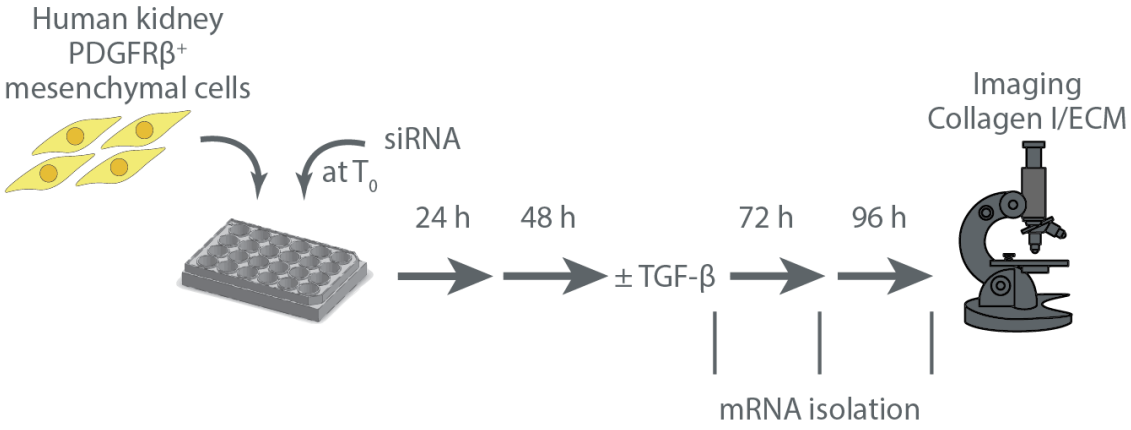

Figure 48 Experimental set-up for validation experiments using siRNA treatment. The cells were seeded at day 0. The day after, cells were transfected with the control siNeg9 and siRNAs against genes of interest. 6 h post transfection, the media was changed to new DMEM + 5 % FBS. After another 48 h, the first samples were taken for mRNA isolation while the treatment ± TGF-β

Knock-down of COL1A1

To start, I used siCOL1A1 treatment. To investigate the knock-down efficiency, I isolated mRNA at 48, 72 and 96 h post siRNA treatment and tested whether the mRNA abundance decreased upon the knock-down. In Figure *49*, the effect of the siCOL1A1 treatment is shown, leading to a significant decrease of COL1A1 mRNA over time with all the siRNAs used (siCOL1A1_75, siCOL1A1_76, siCOL1A1_77). Additionally, the COL1A1 abundance increased over time in the siNeg9 treated cells and was significantly increased by TGF-β stimulation (Figure *50*). Moreover, the expression of other genes, including E2F1, ID2, ID3 and SMAD1 were assessed to see whether there might be a feedback loop involved.

The imaging data, at the last time point (96 h post siRNA treatment and after 48 h of TGF stimulation), generally showed that ECM deposition was decreased in the siCOL1A1 conditions (Figure *50*). The reduction of ECM was significant for the control samples, except for siCOL1A1_77, as well as for the TGF-β treated samples. In Figure *50*A, the fold-change of fluorescence intensity per cell, correlating with the staining of the ECM signal normalized to the cell number, was calculated compared to the siNeg9 ctrl samples. In Figure *50*B, the fold-change of siCOL1A1 + TGF-β treated samples was calculated in comparison to the siNeg9 + TGF-β stimulated samples. Both graphs indicate that, while ECM deposition was increased upon stimulation with TGF-β, the deposition was significantly decreased upon siRNA knock-down, with all three siRNAs tested for COL1A1.

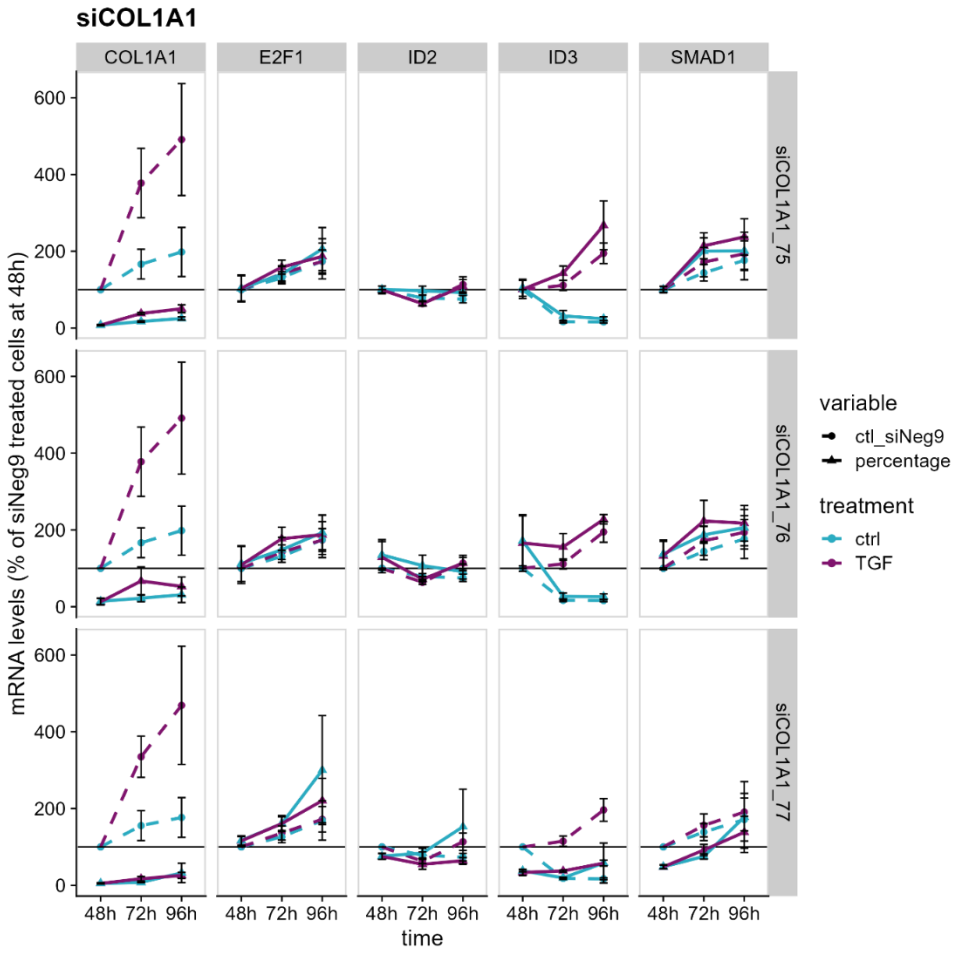

Figure 49 Change of mRNA expression levels at 48, 72 and 96 h post siCOL1A1 treatment. mRNA was isolated from cells transfected with three different siRNAs against COL1A1 at 48, 72, 96 h post siRNA transfection, correlating to 0, 24, 48 h treatment ± TGF-β. Expression levels of genes of interest are shown on the y-axis, that were calculated using the $2^{-\Delta\Delta CT}$ method. In short, CT values are first normalized to GADPH, then normalized to the siNeg9 ctrl at 48 h. Then the $2^{-\Delta\Delta CT}$ value is multiplied by 100 to obtain the resulting mRNA levels. Expression of COL1A1 (and other genes of interest) are shown over time, upon siNeg9 treatment (dashed lines), treatment with siCOL1A1 (continuous lines), and at control (blue) and TGF-β (violet) conditions. While TGF-β treatment leads to increase in COL1A1 mRNA expression levels, the abundance of transcripts is significantly decrease upon the knock-downs with siCOL1A1. n = 4. Code provided by Mira Burtscher was the starting point for this plot.

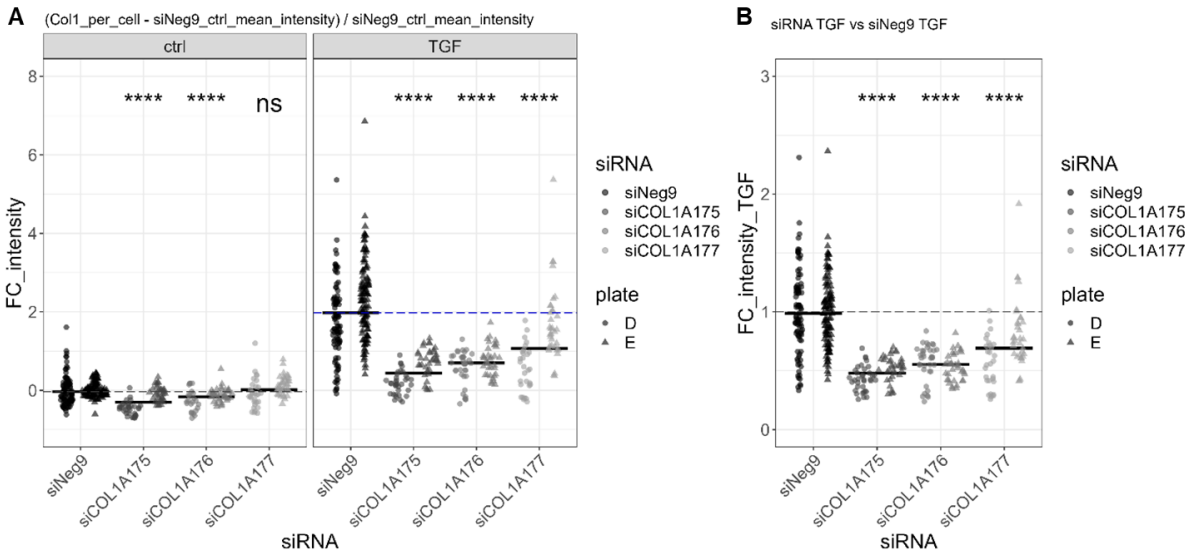

Figure 50 Quantification of immunofluorescent staining of the ECM upon siCOL1A1 knock-down. Quantification of fluorescence intensity (FC_intensity), correlating to ECM/CNA35 staining, expressed by cells 96 h post siRNA transfection ± 48 h TGF-β stimulation. Three siRNAs against COL1A1 were tested (grey) and compared to the siNeg9 controls (black). Each dot represents individual images of two biological replicates. The siRNAs are shown on the x-axis. The y-axis shows the FC_intensity, calculated as follows: **A)** Per-cell fluorescence intensity (after background correction and autofluorescence subtraction) of the siNeg9 ctrl was subtracted from all images. The result was further divided by the value of the siNeg9 ctrl condition. **B)** This time, the values of TGF-β treated samples were compared, siCOL1A1 treated samples vs siNeg9 resulting in the illustrated FC_intensity_TGF. The median values for each sample are shown by the bars, dashed lines stand for the control sample used as reference in the t-test. Summarizing, the siCOL1A1 treatment resulted in a decrease of deposited ECM.

Knock-down of FN1

Another siRNA that I tested was targeting FN1. Using qPCR, I showed that the knock-down was efficient and maintained over the tested time points. Surprisingly, while the effect of the knock-down on COL1A1 mRNA expression was not significant (**Figure *51***, two-way Anova **Table *24***), the ECM deposition was significantly increased after siFN1 treatment (**Figure *52***).

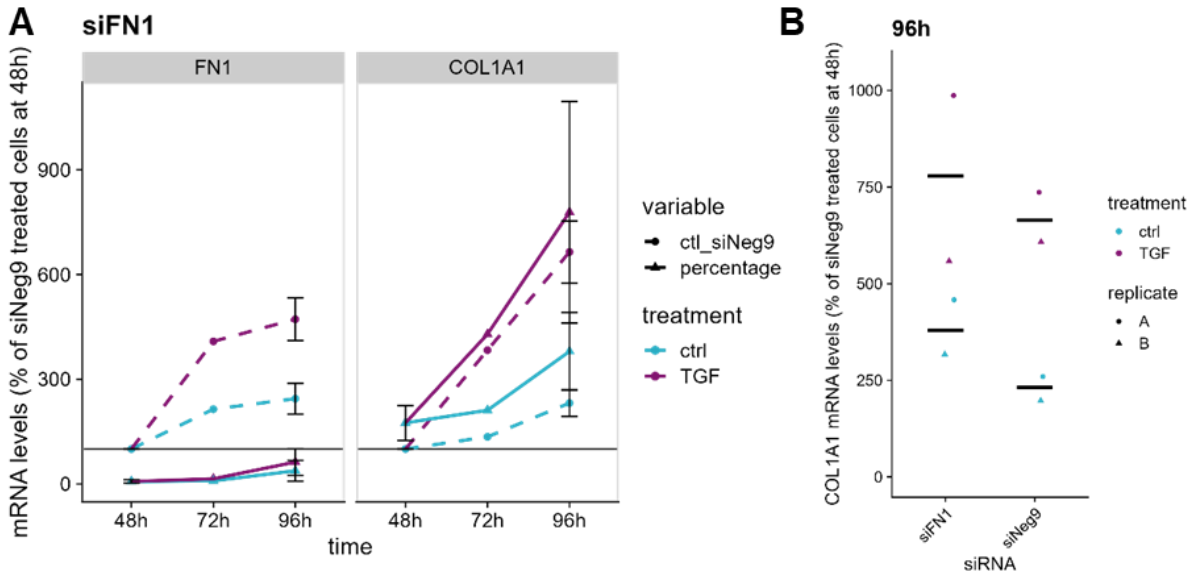

Figure 51 Change of mRNA expression levels at 48, 72 and 96 h post siFN1 treatment. mRNA was isolated from cells at 48, 72, 96 h post siRNA transfection with siNeg9 and siFN1, correlating to 0, 24, 48 h treatment ± TGF-β. **A)** Expression levels of genes of interest are shown on the y-axis, that were calculated using the $2^{-\Delta\Delta CT}$ method. In short, CT values are first normalized to GADPH, then normalized to the siNeg9 ctrl at 48 h. Then the $2^{-\Delta\Delta CT}$ value is multiplied by 100 to obtain the resulting mRNA levels. Expression of FN1 and COL1A1 are shown over time, upon siNeg9 treatment (dashed lines), treatment with siFN1 (continuous lines), and at control (blue) and TGF-β (violet) conditions. While TGF-β treatment leads to increased FN1 and COL1A1 mRNA expression levels, the abundance of COL1A1 transcripts is not significantly changed upon knock-down with siFN1 **(B)**. n = 2 for 48 and 96 h, n = 1 for 72 h. Code provided by Mira Burtscher was the starting point for this plot.

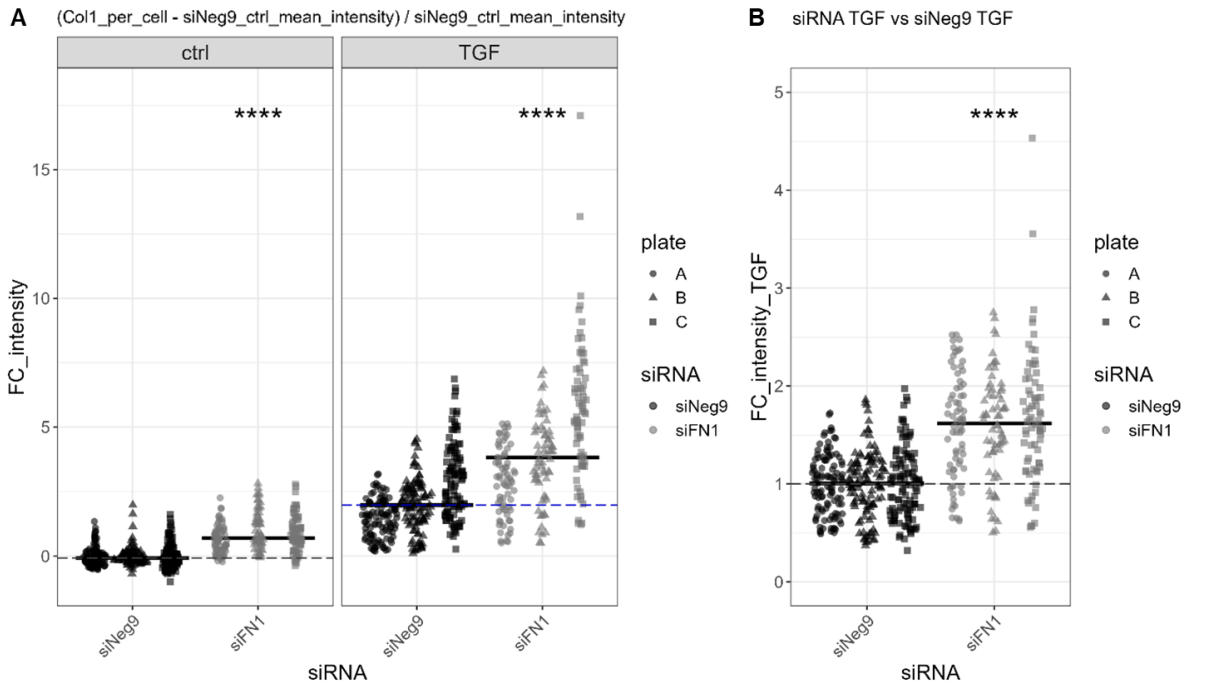

Figure 52 Quantification of immunofluorescence staining of the ECM upon siFN1 knock-down. Quantification of fluorescence intensity (FC_intensity), representing the CNA35 staining of the ECM, expressed by cells 96 h post siRNA transfection ± 48 h TGF-β treatment. Each dot stands for individual images of three biological replicates, siNeg9 treated samples are shown in black, grey dots depict samples treated with siFN1. Samples treated with the different siRNAs are indicated on the x-axis. The y-axis shows the FC_intensity, calculated as follows: **A)** Per-cell fluorescence intensity (after background correction and autofluorescence subtraction) of the siNeg9 ctrl was subtracted from all

images. The result was further divided by the value of the siNeg9 ctrl condition. **B)** This time, the values of TGF-β treated samples were compared, siFN1 treated samples vs siNeg9, resulting in the illustrated FC_intensity_TGF. The median values for each sample are shown by the bars, dashed lines stand for the control sample used as reference in the t-test. Knock-down of FN1 resulted in significantly more ECM deposited.

4.4.3. Knock-down of Transcription Factors for Evaluating Their Involvement in Fibrosis

For validation, I focused on transcription factors (TFs) that showed an increased activity 1 h post TGF-β treatment, inferred based on their target gene expression changes, as mentioned in section 4.3.1.

Gene expression analysis after TF Knock-down

Again, I isolated mRNA at the time points 48, 72 and 96 h post siRNA treatment, correlating to 0, 24 and 48 h after treatment ± TGF-β and monitored how the mRNA expression of the genes of interest and other target genes changed. Additionally, I tested a second set of siRNAs for E2F1 and FLI1, and a third siRNA for SMAD1. I am summarizing the changes of the last time point in the heatmap below (Figure *53*). For more information (across all time points tested), please have a closer look into the Annex (Figure *77* - Figure *81*).

Looking at the siSMAD1 treatment (96 h after transfection), a significant decrease of SMAD1 mRNA was visible for all the siRNAs tested, showing that the knock-down worked. This was true for the TGF-β treated and the control samples compared to the respective siNeg9 treatments. Another question that I wanted to answer, was whether knock-down of SMAD1 would affect COL1A1 expression. For the first siSMAD1 tested, COL1A1 expression was significantly decreased in the siSMAD1 + TGF-β treated sample when compared to the respective siNeg9 + TGF-β control at the same time point. The effect on COL1A1 abundance changes upon treatment with the second siSMAD1 were not significant while there was a significant increase of COL1A1 mRNA for the third siRNA tested (siSMAD1_C), in the control as well as the TGF-β treated sample. All three siRNAs against SMAD1 induced a significant decrease in ID2 and ID3 levels, proposed downstream targets (Figure *35*), upon co-stimulation with TGF-β, whereas there was rather an increase of both factors in the control samples (-TGF-β), which ws significant for siSMAD1_C. ID3 mRNA expression was significantly downregulated using siSMAD1_B, without TGF-β.

Knock-down of E2F1 was successfully confirmed by both siRNAs used (Figure *53*). The first siRNA for E2F1 lead to a significant increase of COL1A1 mRNA in the non-TGF-β condition. However, no significant increase of COL1A1 expression was measured in the other conditions. A proposed downstream target of E2F1, SERPINE1/PAI1, was significantly upregulated upon siE2F1 treatment with all of the tested siRNAs and at all conditions. The decreased mRNA expression of ID2 and ID3 was significant for both siRNAs tested in the conditions with TGF-β (Figure *53*). Interestingly, E2F1 knock-down lead to reduced SMAD1 mRNA with the first siE2F1 but did not result in significant deregulation with the second siE2F1.

Another TF, which effects I wanted to test on downstream targets is FLI1. FLI1 knock-down led to a significant decrease in FLI mRNA expression with both siRNAs targeting FLI1. For the first siFLI1 tested, there was a significant increase of COL1A1 expression observed in the non-TGF-β treated sample. However, this effect was not replicated using the second siRNA against FLI1. The only significant effect of the second siRNA tested was on the mRNA expression of ID3, which was increased upon KD of FLI1, in the control and TGF-β treated samples (not-tested for the first siFLI1).

Except for the significant increase of SERPINE1 (non-TGF-β treatment) and SMAD7 (with TGF-β treatment), siNR4A1 treatment did not result in any significant effect on other genes tested (Figure *53*).

Treatment with siHNF4G led to a significant decrease in HNF4G mRNA levels, further resulting in increased mRNA expression of BHLHE40 and decreased E2F1 mRNA abundance. Additionally, knock-down of HNF4G resulted in significantly decreased COL1A1 expression in TGF-β treated samples as well decreased SPHK1 mRNA levels. Significant increase of FLI1 mRNA expression of non-TGF-β stimulated cells was a further result of HNF4G knock-down.

The knock-down of siBHLHE40, another TF tested, resulted in a significant decrease of BHLHE40 mRNA levels. Treatment with siBHLHE40 and TGF-β led to a significant increase of COL1A1 mRNA levels and a significant decrease of SMAD1 mRNA levels (Figure *53*). However, increased mRNAs abundances of ID2 and SMAD7 resulted from knock-down of BHLHE40 (without TGF-β).

Notably, the expression of SMAD1, E2F1, FLI1, HNF4G and NR4A1 all increased over time and were unchanged upon treatment with TGF-β. Whereas the mRNA abundance of BHLHE40 increased over time but was further raised by TGF-β stimulation (Figure *77* - Figure *81*).

Figure 53 Heatmap summarizing log fold-changes of mRNA expression levels at 96 h post siRNA transfection compared to the siNeg9 at 96 h. The upper heatmap depicts the data obtained for the first set of siRNAs tested while the lower heatmaps represents data obtained from validations using an additional set of siRNAs for selected target genes. The mRNA expression levels ($2^{-\Delta\Delta CT}$ value after normalisation to corresponding CT values for GAPDH and the siNeg9 at 48 h post siRNA transfection) were used to calculate the average log fold-changes (logFC) between the combination of siRNAs of interest compared to the siNeg9 at 96 h. For each combination of siRNA and gene, there are two cells. The lower cell stands for the logFC of the siRNA treated control sample compared to the siNeg9 treated control. The upper cell shows the fold-change of the TGF-β treated samples, siRNA vs siNeg9. Violet indicates an increase in mRNA expression, while blue stands for decreased values. The x-axis shows the genes tested, the y-axis the siRNAs used. A t-test was performed to compare the percentages between siRNAs of interest and siNeg9. Code for this graph was initially provided by Sarah Kaspar and then adjusted by me.

Effect of TF Knock-down on ECM Deposition

Besides testing the mRNA expression of the selected TFs, Col I and possible target genes, I investigated the effect of the KDs on the ECM deposition, by staining fibrillar collagen with CNA35 and imaging at the last time point (96 h after siRNA transfection ± 48 h TGF-β treatment) (**Figure *54***, **Figure 55**).

The first siRNA that I tested for SMAD1 (siSMAD1a), had a significant impact on the ECM deposition, leading to less ECM in the control as well as the TGF-β treated samples. To confirm the effect, I tested a second siRNA (siSMAD1b) which led to an increase in the deposited ECM. A third siRNA (siSMAD1c) showed no effect on the ECM deposition when compared to the siNeg9 ctrl but led to a significant increase of ECM relative to siNeg9 TGF-β treated samples (**Figure *54*A**, **Figure *55*A**).

Both siRNAs used for the knock-down of E2F1 resulted in a significant increase of ECM in the control as well as TGF-β treated conditions (**Figure *54*B**, **Figure *55*B**). While siFLI1 treatment had no significant effect on the ECM deposition in the ctrl samples, it resulted in a significantly higher signal in the samples with TGF-β stimulation for both siRNAs tested (**Figure *54*D**, **Figure *55*D**).

Other TFs, for which I tested only one siRNA are NR4A1, HNF4G and BHLHE40 (**Figure *54*C**, **Figure *55*C**). The knock-down of NR4A1, for example, showed a significant increase of ECM in the ctrl treated samples compared to the siNeg9 ctrl. In **Figure *54*C**, however, when the TGF-β and siNR4A1 treated samples were normalized to the siNeg9 ctrl sample and then compared to the siNeg9 TGF-β treated sample, the difference was not significant. The TGF-β and siNeg9 stimulated samples, however, compared to the siNeg9 TGF-β treated sample, led to an increase in ECM deposition that was significant. Knock-down of HNF4G on the other hand, did not yield a significant change in ECM deposition in ctrl samples, but led to a significant decrease of ECM deposition under TGF-β conditions. A significant increase of ECM was observed in all conditions treated with siBHLHE40.

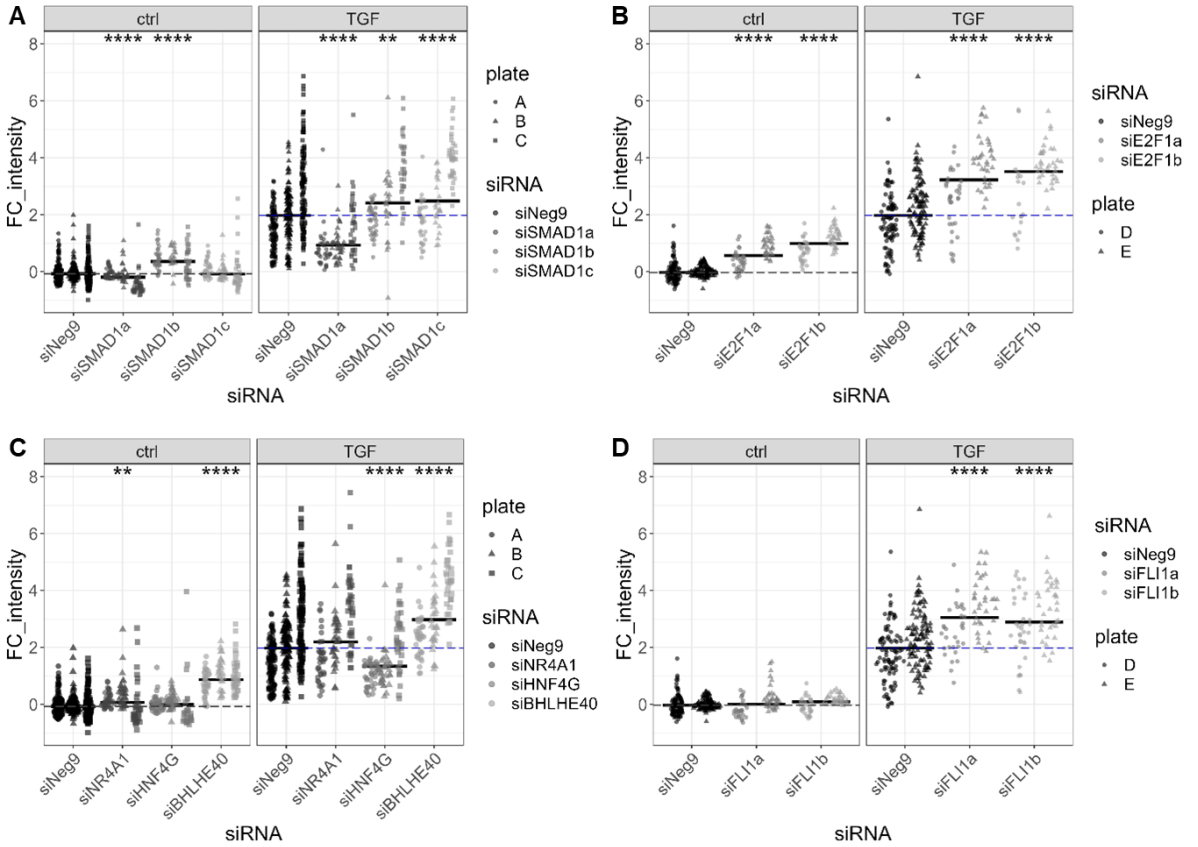

Figure 54 Quantification of immunofluorescence staining of the ECM post siRNA treatment. Quantification of fluorescence intensity (FC_intensity), representing the CNA35 staining of the ECM, expressed by cells 96 h post siRNA transfection ± 48 h TGF-β treatment. Each dot stands for individual images of two to three biological replicates, siNeg9 treated samples are shown in black, grey dots depict samples treated with the siRNAs of interest. The x-axis summarizes the samples. The y-axis shows the FC_intensity values, calculated as fluorescence intensity (after background correction, autofluorescence subtraction and normalisation to nuclei number) of the siNeg9 ctrl subtracted from all images. The result was further divided by the value of the siNeg9 ctrl condition. The median values for each sample are shown by the bars, dashed lines stand for the control sample used as reference in the t-test. ECM deposition changes after knock-down of **A)** SMAD1, **B)** E2F1, **C)** NR4A1, HNF4G and BHLHE40, and **D)** FLI1. Statistical analysis was performed using t-test.

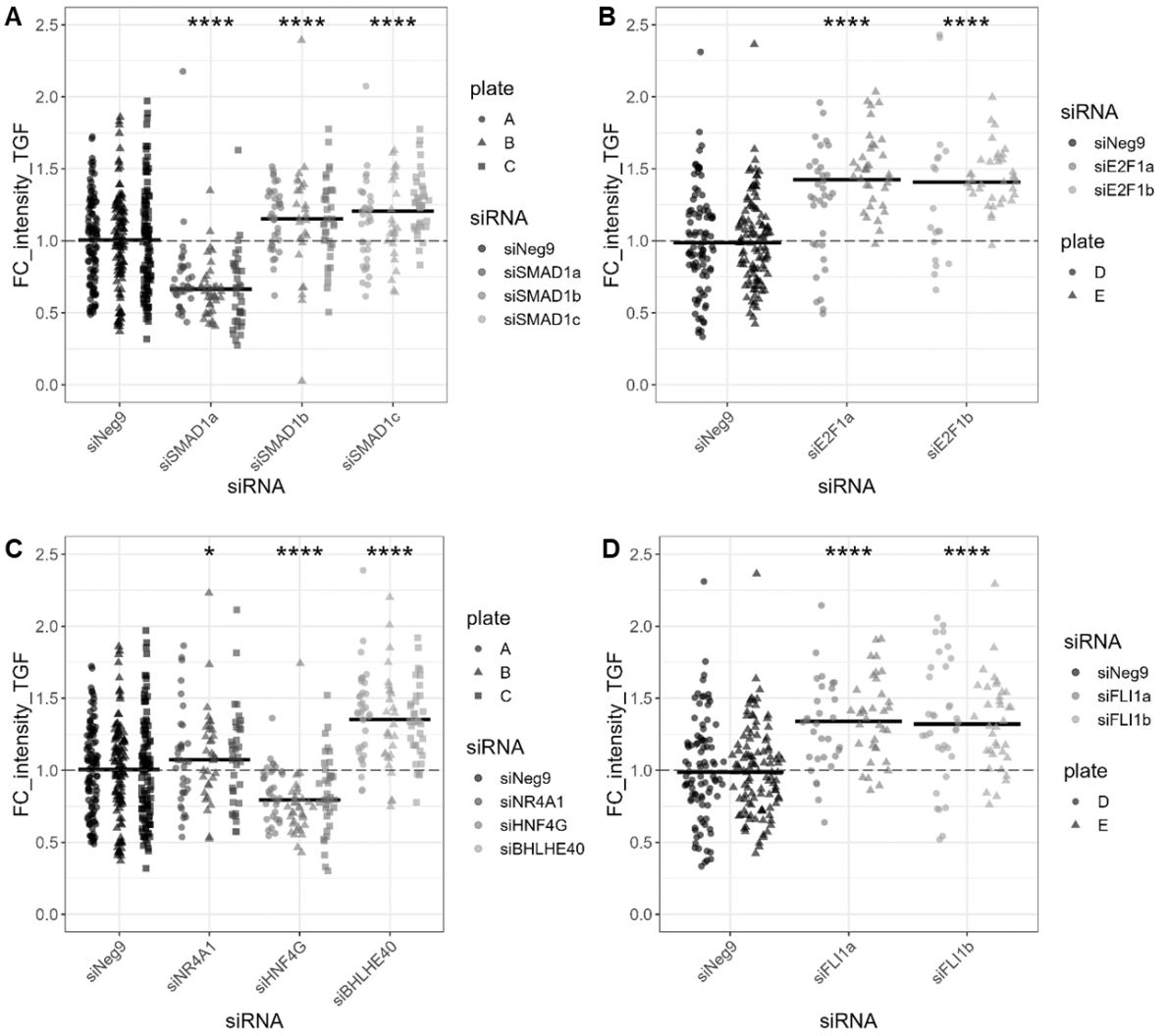

Figure 55 Quantification of immunofluorescence staining of the ECM post siRNA treatment in the TGF-β treated conditions. Quantification of fluorescence intensity (FC_intensity), representing the CNA35 staining of the ECM, expressed by cells 96 h post siRNA transfection ± 48 h TGF-β treatment. Each dot stands for individual images of two to three biological replicates, siNeg9 treated samples are shown in black, grey dots depict samples treated with the siRNAs of interest. The x-axis summarizes the samples. The y-axis shows the FC_intensity values, calculated as fluorescence intensity (after background correction, autofluorescence subtraction and normalisation to nuclei number) of the samples treated with the siRNA of interest and TGF-β, were normalized to the samples treated with siNeg9 and TGF-β. The median values for each sample are depicted by the bars, dashed lines stand for the control sample used as reference in the t-test. ECM deposition changes after knock-down of **A)** SMAD1, **B)** E2F1, **C)** NR4A1, HNF4G and BHLHE40, and **D)** FLI1. T-test was performed for statistical analysis.

5. Discussion

5.1. Initial Experiments Highlight the Fibrotic Response of the Cellular Model to TGF-β

I was aiming to set up a cellular model to study the mechanisms of fibrosis. First, I planned to investigate whether the cellular model shows a profibrotic response when stimulated with TGF-β, a major driver of fibrosis. ECM deposition, by activated myofibroblasts, and replacement of functional tissue is one of the major read-outs of kidney fibrosis. Therefore, I first examined the ECM deposition post TGF-β stimulation in the presence of macromolecular crowding, as already successfully applied in the lab (Khan et al., 2023). Following a 48 h exposure to TGF-β, ECM deposition was notably increased (Figure *4*), measured by immunofluorescence staining of extracellular Col I, a major component of the ECM (Karsdal et al., 2017).

Previous studies have indicated that differentiation into myofibroblasts involves morphological changes, increased expression of α-SMA and stress fiber formation (F. C. Chang et al., 2012; Edeling et al., 2016; Hinz & Lagares, 2021; Kramann et al., 2015; Sziksz et al., 2015; Zeisberg & Kalluri, 2015). The α-SMA antibodies tested in this project, showed a generally low expression of α-SMA. Only a few cells expressed a higher level of α-SMA. This could be explained by the heterogeneity of the used cellular system, comprising PDGFRβ^{+} cells, including tissue resident pericytes and fibroblasts. However, the treatment with TGF-β did not increase the abundance of α-SMA in the cells or the population of cells expressing α-SMA. Moreover, the expression of α-SMA alone is not a reliable marker to identify myofibroblasts but rather the colocalization of α-SMA and F-actin containing stress fibers. Therefore, I tested staining of stress fibers using phalloidin. The data showed F-actin staining and stress fibers in the control conditions which could be explained by the fact that fibroblastic cells express F-actin stress fibers in culture, due to the stiff surfaces they are cultured on (Hinz & Lagares, 2021). Nevertheless, the stress fibers expressed by TGF-β stimulated cells were more pronounced (Figure *5*), and were further locally and globally organized. In contrast, while the cells were vimentin positive, TGF-β stimulation did not lead to an increase of vimentin nor to any significant morphological changes. Initially,

I chose vimentin as it is often used in literature as a marker for activated myofibroblasts (Edeling et al., 2016; Lotz-Jenne et al., 2016; Zeisberg & Kalluri, 2015). However, it is also discussed that vimentin usually fails to discriminate myofibroblasts from endothelial and smooth muscle cells and might not be a suitable marker to use, at least not on its own (Hinz & Lagares, 2021).

To further test the fibrotic response to TGF-β, I assessed the mRNA expression of fibrotic marker genes. In initial experiments, there was a trend towards increased expression of ACTA2, FN1, and VIM which was, however, not significantly higher in the TGF-β treated samples compared to the respective control samples (**Figure *6***). Nevertheless, a significantly higher COL1A1 mRNA expression was induced by TGF-β at 24, 48 and 72 h post treatment.
In summary, the data highlight the fibrotic response of the used cellular model to TGF-β treatment, and show that the cells adjusted their morphology as well as increased expression of COL1A1 mRNA and protein.

5.2. Optimized Experimental Conditions

With regards to using multi-omics read-outs, I tested different experimental set-ups and optimized the protocol so that it can be used for the intended purpose. This meant that I had to get rid of the FBS used in the medium as it would lead to background signal in the MS runs for secretomics. Furthermore, the growth factors in the FBS could influence the signaling mechanisms, that I aimed to study in my project.
First, the TGF-β concentration was revised (**Figure 7**). While all concentrations used resulted in a significant increase of extracellular Col I, when compared to the untreated condition, I decided to stick with the 10 ng/ml as previously employed. Despite the common usage of 5 ng/ml or even 1 ng/ml TGF-β in several studies (Hinz & Lagares, 2021), 10 ng/ml are also used by the Kramann lab, who utilize this cellular system as well. Hence, the decision was taken to keep obtained data comparable.

In a study led by Zi et al. (2011), they showed that HaCaT cells (keratinocytes) stimulated with a TGF-β dose, that was less than a certain threshold, had almost no signaling response. Furthermore, they suggest that the concentration of molecules of TGF-β per cell determines whether the cells show a transient or a sustained long-term signaling response. They highlight that this could influence cell fate decisions like EMT,

migration or cell growth arrest (Zi et al., 2011). While the response to TGF-β can vary across cell types, this study highlights the effects of TGF-β depending on the dose used, meaning how many molecules of TGF-β are available per cell and whether the stimulation is persistent.

It was not within the scope of my project to determine the TGF-β molecules per cell. However, it raised the question whether I could further improve the cellular assay and the cells' response to TGF-β by optimizing the cell number. In general, the data showed that the fluorescence intensity increased with more cells, or more specifically, with more cells seeded (**Figure 8**Figure 7). This was true for the control and the treated samples, whereas the fluorescence intensity was higher in the TGF-β treated samples (**Figure 8**, **Figure 9**). Of note, especially for the 6 and 12 h time points, there seemed to be a drop of fluorescence intensity with a higher cell number seeded (**Figure 8**). This was due to cells overgrowing. When normalizing the Col I intensity values to the cell number per image, in some cases, it seemed that the resulting Col I intensity values decreased with increasing number of cells seeded (**Figure 8 c, d**).

To compare the effect of TGF-β treatment at each time and for each cell number seeded, I showed the FC_intensity, where the Col I per cell values were normalized to the corresponding control (for each cell number seeded separately, **Figure 9**). Generally, after 24 h of treatment with TGF-β, the difference between ctrl and treatment was increasing.

The decision to select the cell numbers, as described in results section 4.1.3, was based on the aspects above. On the one hand, I selected conditions where the difference of deposited Col I between control and TGF-β treated samples was significant and as high as possible. So, if the cells were seeded too sparsely, less fibrillar collagen was deposited. On the other hand, there were limitations for the imaging read-out. If too many cells were seeded, cells became too dense and either detached (often seen in the control condition) or overgrew (**Figure *10***). Moreover, more cells became autofluorescent, or by overgrowing were in a different focal plane, which led to false positive signals. The autofluorescence could have been due to apoptosis (Levitt et al., 2006).

In parallel, I planned the multi-omics experiment and was becoming aware that, to reduce the batch effect, it would be best to harvest all time points and samples of one biological replicate at the same time. To explore the potential impact of different

experimental set-ups on the cells' response, I compared the previously used set-up (A) and an alternative set-up (B), where all samples were harvested simultaneously (**Figure *11***).

Set-up B provided the advantage of comparing all samples as they have been cultured in one plate (**Figure *12***), which was especially advantageous for the imaging data (supplementary ***Figure 62***). In both set-ups, there was a significant increase of extracellular Col I in the TGF-β treated conditions when compared to the control at the corresponding time point.

Moreover, RT-qPCR quantification of fibrotic marker genes, showed that in both set-ups the expression of the tested genes increased post TGF-β stimulation (**Figure *13***). Notably, the number of biological replicates was higher for the set-up B which makes the changes significant, while the t-test did not show significance for most of the genes in set-up A.

Additionally, the morphological changes of the F-actin stress fibers, using phalloidin staining, have been shown for set-up B (**Figure *14***). As previously discussed, the cells in the control conditions also expressed actin stress fibers. However, after stimulation with TGF-β, the fibers became more intense, longer and globally organized. While this effect might not have been so prominent 12 h post TGF-β treatment, it was visible after 24 h.

Since the morphological modifications, alterations in mRNA expression and Col I deposition take time, I sought to explore more immediate responses to changes of the signaling pathways. Moreover, in order to incorporate relevant time points for phosphoproteomics, as a vital component of the signaling cascade and cellular response to TGF-β, I investigated the phosphorylation of a key downstream target, SMAD2 (Friedman et al., 2013; Meng et al., 2016; Sziksz et al., 2015; Tie et al., 2022). Even short stimulation with TGF-β, for instance 30 sec or 2 min, increased phosphorylation of SMAD2 (**Figure *15***, **Figure *16***, **Figure *17***). This effect persisted until the last time point (96 h) (**Figure *16***, **Figure *17***).

Based on these findings, I chose to include multiple time points for the multi-omics experiments. This decision was driven by several objectives. To start with, I aimed to resolve the early mechanisms underlying fibrosis, as gaining insights into early stages could potentially facilitate identification of therapeutic targets or inform strategies for

slowing down disease progression. Conversely, fibrosis is a long process, requiring investigation beyond the immediate response. As a consequence, I hypothesized that studying the time resolved response of the used cells would lead to valuable insights into the dynamic mechanisms underlying fibrosis. Combined with the multi-omics read-outs, this would result in a deeper understanding of fibrosis mechanism and progression.

This idea was supported by other studies that applied a time-resolved read-out and provided comprehensive insights by doing so. D'Souza et al. (2014), for instance, propose that a temporal regulation of proteins may be responsible for the context-dependent effects of TGF-β in their tested cellular system, keratinocytes (D'Souza et al., 2014).

5.3. Testing of Anti-fibrotic Compounds Show Cell Specific Response

To further highlight the relevance of the cellular system used in this study, I examined the effect of anti-fibrotic compounds including FDA approved drugs that are used in treatment of other fibrotic diseases. I chose Pirfenidone (PFD) and Nintedanib (Nin), as they are used as standard of care of IPF worldwide (Collins & Raghu, 2019; Khan et al., 2023; Walraven & Hinz, 2018). Even though treatment with PFD and Nin does not cure fibrosis, these compounds have shown to slow the rate of lung function decline. Moreover, their potential to treat other fibrotic diseases was further tested in smaller clinical trials also for kidney fibrosis (Cho & Kopp, 2010; Klinkhammer et al., 2017).

Proposed mechanisms of action of PFD included inhibition of the TNF-α translation and GLI transcription factor activity (Didiasova et al., 2015; Khan et al., 2023). Pathways affected by Nin are VEGF, FGF, PDGF as well as lymphocyte-specific protein tyrosine kinase and Src nonreceptor kinase signaling. The inhibited pathways are of therapeutic interest as all of them are involved in fibrosis (Bigaeva, Stribos, et al., 2020).

The concentrations used in my experiments were based on previous publications (Bon et al., 2019; Cannito et al., 2017; Cho & Kopp, 2010; Cui et al., 2020; Khan et al., 2023; Lehtonen et al., 2016; Nakayama et al., 2008; Sheu et al., 2019; Stahnke et al., 2017). However, in the used cellular model, neither PFD nor Nin showed a decrease of

extracellular Col I deposition after 48 h in the presence or absence of TGF-β (Figure *18*, Figure *19*). Only treatment with 0.05 and 0.5 µM PFD (the lowest concentrations tested) significantly decreased extracellular Col I deposition after 96 h (Figure *18*). As previously mentioned, phenotypes and signaling pathways of myofibroblasts might differ depending on the cell type they originate from. As an example, there are several studies that hint towards an anti-fibrotic potential in lung fibroblasts or animal models of pulmonary fibrosis. Nevertheless, the promising anti-fibrotic effects of PFD observed did not reflect the performance in human renal fibroblasts or an ex-vivo model of human PCKS, as shown by Bigaeva et al. (2020). Moreover, in clinical trials of patients with diabetic kidney disease and focal segmental glomerulosclerosis, use of PFD did not show a prevention of proteinuria or GFR (Bigaeva, Cavanzo, et al., 2020). Adversely, there is even a case report showing that PFD could lead to kidney injury (Kanbay et al., 2019).

It was also shown by Bigaeva et al. (2020), in human PCKSs, that Nin has an effect in preventing fibrosis onset rather than reversing fibrosis (Bigaeva, Stribos, et al., 2020). Other sources have applied pre-incubation with anti-fibrotic compounds prior TGF-β treatment (Hinz & Lagares, 2021). Even though it might be interesting to repeat the experiments with stimulation of anti-fibrotic compounds prior to incubation with TGF-β, I did not do this as my goal was rather focused to understand how to inhibit the already existing fibrotic effects of TGF-β stimulation.

Since I wanted to see how the used cellular model would respond to other anti-fibrotic stimuli, I further tested LY2109761 (LY), a dual inhibitor of TGF-β receptor type I/II, as it had been shown to decrease the fibrotic response in murine precision-cut kidney slices (Luangmonkong et al., 2018; Melisi et al., 2008; Stribos et al., 2017). In an experiment, I showed that treatment with the highest concentration tested (20 µM) led to a significant reduction of ECM deposition in the control samples compared to the untreated control (Figure *20*). In comparison, 48 h treatment with LY in the presence of TGF-β resulted in a significantly lower ECM deposition for all concentrations tested.

Finally, I tested a compound that showed promising results in decreasing extracellular Col I deposition in normal human lung fibroblasts, illustrated by my colleague (Khan et al., 2023). Dextromethorphan (Dex), led to a significant reduction of ECM also in the cellular system applied in my project, in the presence of TGF-β as well as its absence (Figure *21*). This effect was visible already after 48 h and was even more prominent after 96 h, further supporting its anti-fibrotic effects. The anti-fibrotic effects of Dex are

not fully understood yet, but Khan et al. (2023) propose that Dex leads to accumulation of Col I in the ER and an inhibited transport to the extracellular space.
In summary, I showed the relevance of the used cellular model to study fibrotic mechanisms induced by TGF-β stimulation and its application to screen for anti-fibrotic compounds.

5.4. Dynamic Fibrotic Response of Cellular Model Represented at Multiple Scales

5.4.1. Increased Deposition of Extracellular Col I Over Time and upon Stimulation with TGF-β

The immunofluorescence staining of Col I showed increased accumulation of ECM over time upon incubation with molecular crowding and ascorbic acid (**Figure *23***, **Figure 24**). After initial inspection of the data and statistical analysis with t-test (**Figure 24**), I opted for using a linear mixed model (as described in the methods section 3.2.3). This approach provided the advantage to investigate the effect of time, treatment and their interaction on the deposited Col I per cell while accounting for biological replicates as random effects (Piepho et al., 2003, 2004).
For the model, I decided to use the average per cell intensity of the Col I staining calculated for each sample. The reasoning behind this was, that even though the cells showed different densities across the images acquired for each sample, the media they are cultured in was the same. Therefore, the signaling factors secreted by the cells leading to autocrine signaling should have been distributed in the media. Additionally, my set-up averaged the response of the cells across the population. Since the cellular model contained different subpopulations of PDGFRβ$^+$ cells, I opted to study the average response of the population to the treatment.
Other models, for which I used all images or even took the cell number into consideration, led to similar results. Consequently, I decided to run the model with the (square root transformed) mean values for each sample of the Col I per cell value. Additionally, after investigation of model residuals (**Figure *3***), I transformed the Col I per cell values using square root transformation to stabilize the variance and achieve normal distribution of model residuals, which is a prerequisite for anova (Piepho, 2009). It is important to note that there was no 0 h TGF-β treatment and therefore I duplicated the values of the 0 h ctrl for the statistical analysis of the model. This adjustment was

necessary to optimize the model matrix (Figure 3). As a result, the Anova showed no significant main factor effect for the TGF-β condition at the 0 h time point (Table 16). Notably, the influence of time is significant as well as the interaction between time and condition (treatment).
While deposited Col I was significantly different between the untreated control at 0 h and the later control time points (72 and 96 h), there was no significant increase of ECM when comparing the different time points of the conditions incubated with molecular crowding and ascorbic acid (control conditions) (Table 17). This suggested a time-dependent increase of extracellular Col I, independent of external TGF-β stimulation but influenced by treatment with molecular crowding and ascorbic acid.
On the contrary, 24 h of TGF-β treatment resulted in a significant increase of extracellular Col I when compared to the respective control sample at 24 h. Additionally, over time the difference of deposited Col I was increasing and is significant between the tested time points treated with TGF-β (Figure 25, Table 17). Therefore, continuous TGF-β might exert a persistent influence on the ECM accumulation.

Additionally, when looking at the cell number, one could see that the cell number decreased with the later time points (Figure 64). Especially for the longer time points, the cell number was smaller in the TGF-β treated samples when compared to the respective controls. On the one hand, this can be explained by the fact that the cells were under starvation, since I did not add FBS over the treatment period. However, after the day of seeding, all samples have been kept under starvation. Therefore, one could conclude that the ascorbic acid and macromolecular crowding led to a switch in the cell cycle, or a fibrotic response that involved differentiation and halted proliferation. This effect could have been induced, on the one hand, by the macromolecular crowding agents, which boost the deposition of ECM. The changes of the matrix stiffness, variations in composition of matrix proteins and interaction of those with cell surface receptors regulate many cellular events, including proliferation and cell differentiation (An et al., 2016; Henderson et al., 2020). On the other hand, ascorbic acid is added to the medium as it is a cofactor required for various collagen-modifying enzymes, for instance prolyl and lysyl hydroxylases, thereby facilitating Collagen and ECM maturation (Hinz & Lagares, 2021).

In the TGF-β conditions, this effect was supplemented by the profibrotic effects that TGF-β exerts. As indicated in Figure *29*, pathways like E2F targets and G2M checkpoint were downregulated with p53 pathway being upregulated post TGF-β treatment at the later time points and across omics modalities. This could be indicative of induction of cellular senescence.
In summary, the data represent a time dependent effect of TGF-β and reflect the profibrotic response of the cells.

5.4.2. Dynamic Fibrotic Response of Cellular Model Represented at the Multi-omics Scale

Consistent with the imaging data and previous experiments, the analysis of the multi-omics data revealed a clear separation between control and TGF-β treated cells shown in the principal component analysis (PCA) and across the tested omics modalities (Figure *26*). However, most of the variance of the phosphoproteomics data could be explained by the batch effect. This suggested a profound effect of TGF-β on the changes at the molecular profile of the cells which was also supported by the increase of differentially expressed factors at all omics modalities over time (Figure *27*, Figure *28*). Notably, a higher proportion of significantly differentially expressed transcripts (10% of mRNAs measured), as set by the log2 fold-change and adjusted p-value cut-offs, compared to proteomics (1.5 - 5%) was observed. This indicated that transcriptional regulation played a major role in the response of TGF-β treatment. However, changes on the transcriptional level are not fully transferable to alterations in protein abundance. This discrepancy was also evident in the volcano plot (Figure *28*), where numerous mRNAs were annotated as hits with often higher log2 fold-changes than in the proteomics data. The obvious difference of the abundance changes across the omics modalities could be attributed to feedback loops or buffering effects (Kusnadi et al., 2022; Y. Liu et al., 2016; Timmers & Tora, 2018). What is meant with this is that protein expression levels are dependent on several factors that are far beyond changes on transcript levels. The most obvious explanation might be that there is a temporal delay on protein synthesis which can be influenced by factors like the mRNA sequence (e.g. influenced by codon composition or internal ribosome entry sites), or translation modulators (e.g. binding of proteins/micro-RNAs to regulatory elements). When looking at the data, one can see a delayed response, i.e. TGF-β signaling pathway that was upregulated 1 h post TGF-β treatment in the transcriptomics data but only became

more active in the proteomics and secretomics data after 12 h (Figure 29). However, the delay in protein synthesis cannot be the full explanation. By regulating the protein's half-life and abundance for example by degradation or via transport/export of proteins this effect could be further explained (Y. Liu et al., 2016). Nevertheless, further investigations are needed to confirm this hypothesis.

Another layer that is not necessarily dependent on the protein abundance is the activity of proteins and signaling molecules which can be regulated via post translational modifications (PTMs). This was one of the main reasons why I included the phosphoproteomics, as I argued that understanding this critical aspect of signaling would facilitate a better understanding of disease mechanisms.

Under steady state conditions, it is assumed that the relative protein and/or mRNA levels remain relatively stable over time (Lassé et al., 2023; Y. Liu et al., 2016). The data showed that the cells stimulated with TGF-β did not reach a steady state yet, which was indicated by the increasing hit number over time and did not reach a plateau 96 h post TGF-β treatment (Figure 27).

I hypothesized that enrichment of pathways that are observable at multiple omics layers would provide a comprehensive overview of the molecular mechanisms that drive the profibrotic response of the cells when stimulated with TGF-β. Therefore, I initially examined pathways that are differentially regulated across the measured modalities (Figure 29). Using MSigDB and decoupleR (Badia-i-Mompel et al., 2022) I showed that treatment of TGF-β induced profibrotic pathways such as EMT (Kendall & Feghali-bostwick, 2014; Walker et al., 2019), TGF-β (Yan et al., 2021), Hypoxia (Eddy, 2014; Gewin et al., 2017; Sziksz et al., 2015), and TNF-α (Dom et al., 2021; Granata et al., 2016; Mariani et al., 2018, 2023; Rinschen & Saez-Rodriguez, 2021), which were expected and previously linked to fibrosis. On the other side, pathways that exert an anti-fibrotic effect like Interferon α and γ were upregulated at 0.08 h post TGF-β in transcriptomics and secretomics, followed by downregulation at later time points (Inagaki et al., 2003; Oldroyd et al., 1999; Vu et al., 2019). The initial upregulation of those pathways can be explained by the fact that all differentially regulated factors (not just the ones that were defined as hits) were included in the pathway enrichment analysis. Notably, at the earliest time point, there were no hits (according to our log2 fold-change and p-value cut-off) detectable (Figure 27, Figure 28).

Other pathways that have been previously associated with fibrosis include Angiogenesis (Hinz & Lagares, 2020; Rasmussen et al., 2019), and Coagulation (Van Linthout et al., 2014). As already discussed, pathways that are involved in cell cycle regulation, like E2F targets, G2/M checkpoint (Gewin et al., 2017; Li et al., 2016), and p53 pathway (C. E. Higgins et al., 2019; S. P. Higgins et al., 2018; Samarakoon et al., 2013), may indicate the transition of the cells towards a senescent phenotype. Moreover, recent studies have shown the implication of p53 and the interaction with SMAD, leading to a TGF-β induced PAI-1/SERPINE1 transcription involving ROS generation (S. P. Higgins et al., 2018; Samarakoon et al., 2013). Additionally, Myogenesis, a pathway involved in fibroblast differentiation showed increased activity over time (Meran & Steadman, 2011).
Furthermore, the data showed the deregulation of metabolic pathways, previously linked to fibrosis. Fatty acid metabolism and oxidative phosphorylation exhibited similar trends as both pathways were downregulated in the proteomics data and additionally in secretomics or transcriptomics. In contrast, glycolysis was upregulated at 48 h and 72 h in proteomics and/or secretomics, however, this hardly passed the cut-off that I set for the enrichment score (data not shown). Previous studies described the effect of dysregulated glycolysis in experimental models of fibrosis, including kidney fibrosis, and that by inhibiting glycolysis, reduction of ECM accumulation could be yielded (Henderson et al., 2020; Lassé et al., 2023; Zhou et al., 2020). Moreover, it was proposed that upregulated glycolysis could compensate for reduced fatty acid oxidation during kidney injury, thereby enhancing progression of fibrosis (Henderson et al., 2020).

Aiming to identify biomarkers and possible drug targets, I had a further look into factors that changed significantly across the omics modalities (**Figure *30***, **Figure *31***, **Figure *32***). Validating my approach, some of the hits have already been proposed as biomarkers for fibrotic diseases. Of particular interest for me were hits measured in the secretomics dataset as those have the potential to serve as biomarkers detectable in blood or urine and thereby providing a valuable diagnostic and prognostic tool for kidney fibrosis. One of the hits is Tenascin-C (TNC), an extracellular matrix glycoprotein, which is produced by fibroblasts and, in an autocrine fashion, promotes the activation and proliferation (R. Huang et al., 2023; Moita et al., 2022). Cadherin-11 (CDH11), another hit in my study, was previously proposed as a non-invasive biomarker for kidney fibrosis (R.

Huang et al., 2023; Schmidt et al., 2021). Additionally, CDH11 might serve as a therapeutic target for lung fibrosis (Schneider et al., 2012). Both factors exhibited upregulation over time in the data set in all omics layers, except phosphoproteomics. Furthermore, biglycan (BGN), involved in collagen cross-linking showed increased expression during progressive renal scarring (Coleman et al., 2019). More precisely, the soluble biglycan was proposed as a biomarker for inflammatory renal diseases (Hsieha et al., 2014) and has been shown to be useful to monitor the function in liver cirrhosis patients (Genovese et al., 2014). The potential of latent transforming growth factor-β (TGF-β) binding protein-2 (LTBP2), another factor that was increased in the data, as prognostic blood biomarker has been shown recently for IPF as it might give insights into the differentiation of lung fibroblasts to myofibroblasts (Arif et al., 2023; Enomoto et al., 2018; Zou et al., 2021).

Besides biomarkers, the expression of numerous extracellular matrix core and associated proteins was changed, accompanied by proteases that are known to regulate ECM deposition and degradation, as mentioned in the results section. One such protease, carboxypeptidase A4 (CPA4), was found to be downregulated upon TGF-β stimulation in the dataset. Although its function in fibrosis remains poorly understood, its association to different cancer types has been demonstrated, where overexpressed CPA resulted in enhanced EMT and migration (Fu et al., 2019; Gao et al., 2020; Shao et al., 2020; F. Zhang et al., 2019). I hypothesized, that CPA4 could mediate an anti-fibrotic effect, either directly or through regulation of intermediate factors participate in remodeling of the ECM. My attempts to overexpress it in the cellular system have not been successful yet but it would be interesting to assess whether overexpression of CPA4 would result in a reduced ECM deposition.

Surprisingly, collagens were not upon the most regulated proteins in our secretomics data. This could be attributed to post-translational modifications such as hydroxylation, glycosylation or extensive crosslinking, which may render them undetectable or exclude them form analysis (Zhou et al., 2020). The more post-translational modifications are included in the analysis of Mass spectrometry data, the more likely it is to wrongly assign peptides. Additionally, I assumed that due to the use of macromolecular crowding in the media, most of the secreted collagens were quickly deposited. This assumption was supported by my imaging data that clearly showed an

increase of deposited Col I over time (Figure 23, Figure 24, Figure 25). Additionally, the proteomics data, including cell lysates and the ECM, showed a significant increase of COL1A1 and other Collagens, similar to the transcriptomics data (Figure 33).

While focusing on factors that were shared across omics modalities, I may have overlooked factors that could serve as relevant drug targets, captured only in one of the datasets. Additionally, the log2 fold-change thresholds used may have led to neglecting factors that are essential in fibrosis progression but whose effect was not necessarily driven by changes in abundance. Factors such as COL1A1 and BMP1, though deregulated in the dataset, did not meet the log2 fold-change threshold set, explaining their absence from the heatmap (Figure 30, Figure 31).

Notably, I was missing out on other factors that might have been secreted but were not measured via proteomics techniques such as secreted micro or long non coding RNAs, which however, could have been measured in the transcriptomics data. For example, when looking at the top 20 loading vectors for the first PC of the principal component analysis for the transcriptomics data (supplementary Figure 66), there was a long intergenic non-protein coding RNA (LINC01711) that contributed to the separation between control and TGF-β treated samples. LINC01711 has just recently been shown to promote TGF-β induced invasion in one of the central nervous system's most aggressive and malignant tumors (GBM) (Shree et al., 2023). Therefore, it might be interesting for future studies to have a look at the data that I generated. To support the transcriptomics data and identify non-protein biomarkers for the blood or urine, it would have been interesting to enrich RNAs of the cell culture supernatants and send them for sequencing.

Nevertheless, the data highlighted proteins that have been proposed as drug targets for fibrotic diseases, including the recent focus on a disintegrin and metalloprotease 12 (ADAM12) which becomes upregulated in fibrogenic cells (Arif et al., 2023; Sobecki et al., 2022). Therefore, it was suggested as a target in a vaccination-based immunotherapy approach, where cytotoxic CD8+ T cells were directed to attack collagen producing cells expressing ADAM12. Research indicates that ADAM12 vaccination reduced collagen staining in livers of mice and even prevented fibrosis in the lungs of mice that received the vaccine as a prophylactic (Sobecki et al., 2022). Conversely, significantly more cardiac fibrosis was observed in ADAM12 knock-out mice possibly due to enhanced focal adhesion and TGF-β signaling (Nakamura et al.,

2020). Of note, the role of ADAM12 in the kidney has not been extensively studies yet (Wozniak et al., 2021). In line with literature, my data showed that expression of ADAM12 was upregulated in transcriptomics, proteomics and secretomics (**Figure *31***) and its phosphorylation was enhanced (**Figure *32***). Therefore, it would be interesting to further investigate how ADAM12 is involved in kidney fibrosis.
SERPINE1, also called Plasminogen Activator Inhibitor-1 (PAI-1), is involved in ECM remodeling and is associated with increased fibrosis, also in the kidney (Ghosh & Vaughan, 2012; Moita et al., 2022; Samarakoon et al., 2013; Sciences et al., 2017; Shindo et al., 2018; Vaughan et al., 2017; Zhou et al., 2020). Studies have also shown that SERPINE1 is a marker for cellular senescence, which plays a significant role in driving fibrosis. Targeting SERPINE1 provides a promising strategy for therapeutics to reduce ECM deposition and slow down kidney fibrosis progression (Walraven & Hinz, 2018; Wang et al., 2021). Other hits that showed up in this study and have been previously proposed as drug targets include FN1, MMP2, ITGB5, ITGAV, TNC and CCN2 (**Figure *30***, **Figure *31***) (Moita et al., 2022; Samarakoon et al., 2013).

As mentioned in the results section, there were only a few hits that were shared between the phosphoproteomics data and other data modalities. Besides the shared hits, I will now highlight a few hits of interest from the phosphoproteomics data (**Figure *32***). Some factors that were upregulated at early time points were heat shock protein beta-1 (HSPB1), also known as HSP27, and inositol 1,4,5-trisphosphate receptor, type 1 (ITPR1) (**Figure *32***). HSPB1 was already linked to cell motility (Hoffman et al., 2022), EMT, chronic tubulointerstitial fibrosis (Vidyasagar et al., 2008), and hyperphosphorylation which induces actin stress fiber formation (Choi et al., 2016). Regarding kidney fibrosis, overexpression of HSPB1 has been associated with decreased expression of fibrotic marker genes, pro-inflammatory signals and oxidative stress, when overexpressed, thereby excreting an anti-fibrotic effect (X. Zhang et al. 2021). Valles et al. (2003) investigated the expression of HSPB1 in 22 patients with uretero-pelvic junction (UPJ) obstruction (that are occlusions in the urinary tract, at the base of the kidney) (Vallés et al., 2003). The authors found that UPJ obstructions showed tubulointerstitial fibrosis and that the duration of obstruction correlated with HSPB1 expression. In subsequent studies, they demonstrated the protective role of HSPB1 in reducing fibrosis (Vidyasagar et al., 2012). Although, ITPR1 can regulate

senescence in the liver, and thereby may regulate glucose homeostasis (Ziegler et al., 2021), its implications in fibrosis remain poorly understood.
Another protein that showed significant increase of phosphorylation at different sites, however at later time points, was Tensin-1 (TNS1). In human autosomal dominant polycystic kidney disease (ADPKD) tissue, the protein expression of TNS1 was significantly decreased compared to healthy control samples (C. Huang & Lo, 2023). In contrast, another study described the increased expression of TNS1 upon TGF-β stimulation and its essentiality in TGF-β induced myofibroblast differentiation and ECM deposition (Bernau et al., 2017).

In summary, the profibrotic response of the cells was represented at all omics layers, by an increase of profibrotic pathways and a decrease of anti-inflammatory pathways. Moreover, my approach identified previously proposed biomarkers and drug targets, including factors that have not yet been studied in the context of kidney fibrosis. Additionally, the phosphoproteomics data might provide valuable insights into kidney fibrosis progression. However, a deeper and more targeted approach is needed to extract further information.

5.4.3. Multi-omics Results Linked to Single-Cell Data of CKD

Two of the major questions in the field of chronic kidney disease are: i) What is the origin of myofibroblasts that are producing the excess of ECM, and by understanding this, ii) could we find drug targets to provide new treatment options for patients? Single cell sequencing methods have provided an instructive tool to investigate this further and showed that the majority of ECM producing cells in the kidney stem from resident PDGFRβ$^+$ pericytes and fibroblasts (Kuppe et al., 2021).
In this project, I used a cell culture model based on human kidney derived PDGFRβ$^+$ cells, and induced a fibrotic response using TGF-β. I argue that the dynamic data that I generated, provide a thorough characterization of this model, and that the model is a valuable tool to further investigate the molecular response of myofibroblast precursor cells to fibrotic stimuli, which could also involve mechanistic or oxidative stress and stimualtion with other growth factors. An additional application of the used cellular system could be to test perturbations, and a combination thereof, to further investigate important factors in the context of fibrosis but also test new and approved drug compounds.

In order to further demonstrate the validity of my ideas, together with Mira Burtscher, we compared the obtained data to single cell RNA sequencing data of Kuppe et al. (2021). Therfore, we used the supplementary table including the specificity scores of analysed genes for each cell type, calculated by Kuppe et al. (2021) (41586_2020_2941_MOESM4_ESM.xlsx, Human_PDGFBplus_Level_2_Specifi sheet). The top 7% of genes specific for myofibroblasts were filtered based on a specificty score cut-off that we defined as > 0.2 (Figure 56B). This resulted in 1650 genes, that were then compared to my dataset. In the volcano and bar plot (Figure 56C,D), one can see that there was an increase of deregulated genes over time in all multi-omics data that were also expressed in myofibroblasts of CKD patients.

This demonstrates that the treatment with TGF-β initiated a cellular response, leading to differentiation towards myofibroblasts over time, which is biologically relevant in the context of kidney disease.

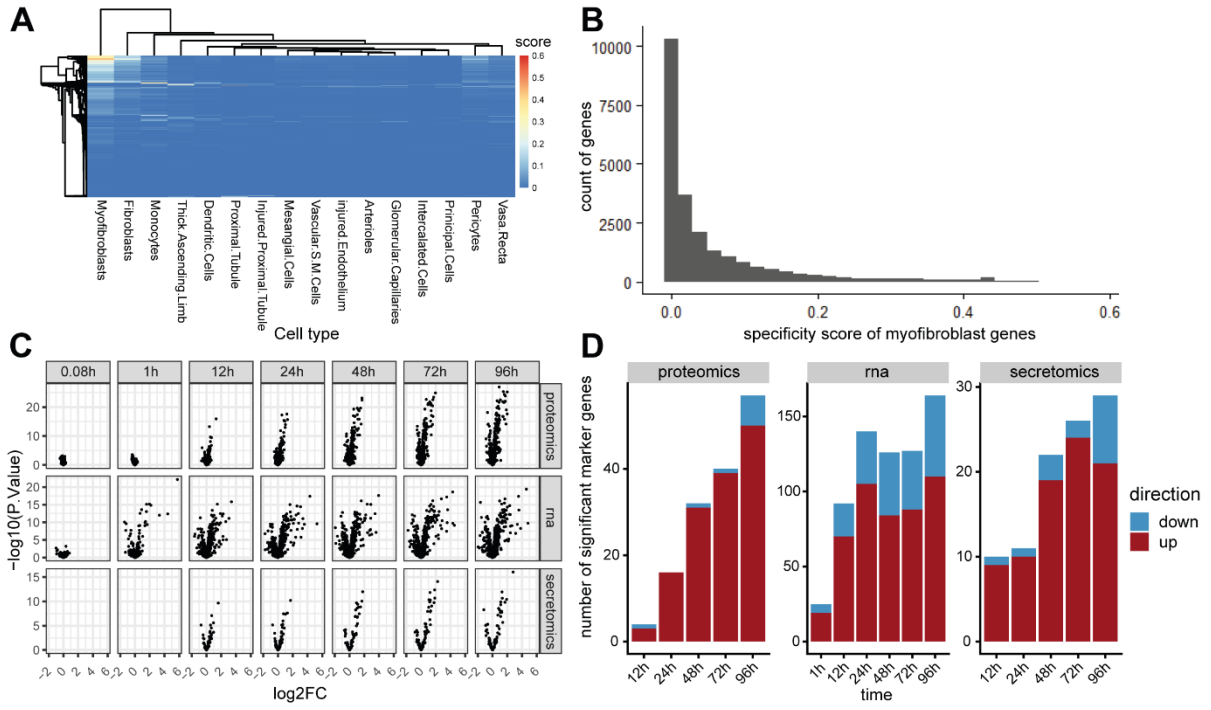

Figure 56 Comparison to myofibroblast specific genes. A) Heatmap illustrating the specificity of genes expressed in different cell types. **B)** Specificity score (x-axis) versus number of genes (y-axis) plotted for genes expressed in myofibroblasts. Cut-off taken at 0.2 **C)** Volcano plot showing filtered myofibroblast specific genes expressed in the dataset generated in this study **D)** Boxplot summarizing deregulated genes (absolute log2 fold-change > log(1.5) and adjusted p-value < 0.05, upregulated genes in red, downregulated genes in blue) among the myofibroblasts specific genes. Code for analysis and figure generated by Mira Burtscher and modified by me.

5.5. Causal Inference of Omics Data Using Mechanistic Modeling of fibrotic response

With the previous discussion sections I wanted to highlight that the fibrotic response to TGF-β involves not only the activation of a singular pathway, but rather the dysregulation of a bigger signal transduction network, leading to altered cellular behavior thereby directing disease context. How the cells sense, integrate and transmit information is especially important in a complex disease like kidney fibrosis. With the aim to further understand the molecular mechanisms of fibrosis, I included footprint-based activity estimation of TFs and kinases/phosphatases based on changes in the transcriptomics and phosphoproteomics data, and used those as signaling nodes in a mechanistic network generated using COSMOS (Causal Oriented Search of Multi-Omics Space) to connect the estimated activities and abundance changes of secreted factors (Dugourd et al., 2021; Dugourd & Saez-Rodriguez, 2019). I argued that, by using these methods, I would be able to link the datasets and provide functional insights by identifying causal pathways through integration of prior knowledge to connect the observed changes (Szalai & Saez-Rodriguez, 2020).

5.5.1. Activity Inference of TFs and Kinases/Phosphatases Shows Fibrotic Mechanisms and New Players

To prepare for the causal network analysis to identify mediators of fibrosis, I used the decoupleR R package to infer the activity of TFs and kinases/phosphatases (Badia-i-Mompel et al., 2022). These signaling molecules have the potential to modulate the expression of proteins and thereby contribute to fibrotic phenotypes. Here, I will address not only the changes in activity of TFs and kinases/phosphatases (**Figure *34***) but also the alterations in abundance (**Figure *71***) and explain the reasons behind their omission as significant findings (hits) in the previous sections.

Among the top regulated TFs was SMAD4 which is known to be implicated in the canonical TGF-β signaling pathway (Hinz & Lagares, 2020; Tie et al., 2022). But SMAD4 did not show up as a hit, neither in the transcriptomics nor proteomics data as its abundance did not change significantly (**Figure *71*Figure *71***). Therefore, I hypothesized that the activity changes of SMAD4 may not be driven by changes in abundance but rather by e.g. changes of posttranslational modifications, interaction

with other proteins or localization (e.g. from cytoplasm to nucleus). SMAD4 was not measured in the phosphoproteomics, hence I cannot make a definite statement about phosphorylation changes.

Interestingly, the activity of a mixed lineage leukemia protein 1 (MLL1), encoded by the gene KMT2A, increased from 5 min onwards. It has previously been shown that dysregulation of the menin/MLL1 pathway contributes to renal fibrinogenesis (Jin et al., 2022). KMT2A is another factor that was deregulated at the transcriptome, proteome and phosphoproteome level but did not pass the significance cut-offs (**Figure *71***).

Furthermore, one can see an upregulation of transcriptional enhancer factor TEF-3 (TEAD4) and serum response factor (SRF) activity, which are implicated in the signaling with the transcriptional co-activators yes-associated protein (YAP)/transcriptional co-activator with PDZ-binding motif (TAZ). The latter are mechanosensitive co-activators that regulate cell-ECM mechano-transduction, meaning that increased ECM stiffness promotes the nuclear transport of YAP/TAZ which bind to TEAD4 and SRF and thereby contribute to production of ECM proteins and profibrotic signaling pathways such as TGF-β, TNS1 and Wnt (Bernau et al., 2017; Hinz & Lagares, 2020; Kim et al., 2022). Looking at the data, TEAD4 mRNA abundance was not significantly deregulated, considering the absolute log2 fold-change > log(1.5) and adjusted p-value < 0.05. Furthermore, it was not detected in the proteomics data, which could be attributed to its low expression levels. While expression of SRF was only significantly upregulated (considering the absolute log2 fold-change cut-off < log(2) and adjusted p-value < 0.05) at 12 h post TGF-β treatment in the transcriptomics data set, its abundance did not change significantly in the proteomics and phosphoproteomics data.

Given the insights from previous studies, which underscore the importance of $Gli1^+$ cells in injury-induced organ fibrosis (Black et al., 2019; C. E. Higgins et al., 2021; Kramann, 2016; Kramann et al., 2013, 2015), it was unexpected to observe downregulated activity of GLI1 over time. Despite this, it is important to mention that GLI1 abundance on mRNA level was significantly increased between 12 – 48 h post TGF-β treatment, aligning with existing literature.

The abundance changes of factors like FOXA1, E2F1 and HIC1 were only measured in transcriptomics and phosphoproteomics but the proteins were not measured in proteomics. Due to the enrichment of phosphopeptides for the phosphoproteomics, also peptides of less abundant proteins could be measured. Nevertheless, factors that

only showed up in phosphoproteomics data but not in the proteomics data of the cell lysates were excluded from downstream analysis as we could not normalize them to changes in protein abundance. Notably, the estimated alteration of activity of FOXA1, E2F1 and HIC1 showed a similar pattern when compared to the mRNA expression changes (**Figure *34***, **Figure *71***).

According to the data, PTK2, the focal adhesion kinase 1, was one of the kinases which activity was increased at all time points measured (**Figure *34***). This kinase has been shown to be implicated in fibrinogenesis by participating e.g. in integrin signal transduction and MAPK signaling (Lagares & Kapoor, 2013). Looking at the mRNA and protein expression levels, the abundance did not change significantly (**Figure *71***). Casein Kinase 2 Alpha 1 and 2 (CSNK2A1, CSNK2A2), genes which encode catalytic subunits of the tetrameric casein kinase II (CK2), were shown to be upregulated in diabetic mouse models. Furthermore, inhibition of the kinases demonstrated amelioration of renal fibrosis in mice (Borgo et al., 2021). In my dataset, the kinase subunits demonstrated increased activity early on. While CSNK2A1 activity was downregulated at 12 h, CSNK2A2 maintained its increased activity. Neither of both kinase subunits showed a significant change of mRNA or protein abundance over time. Interestingly, a recent study demonstrated the potential of a small molecule inhibitor of CK2, CX-4945, in reducing liver fibrosis by inhibiting the activation of HSCs (hepatic stellate cells, which, together with activated portal fibroblasts, constitute more than 90% of the cells in experimental models of liver fibrosis that contribute to collagen production) (Fan et al., 2023). In light of these findings, combined with the fact that these catalytic subunits exhibited early activity, it was valid to assume that CK2 might have an important role in activation of the cells that I used. Therefore, it would be interesting to test the effect of CK2 inhibition in the context of kidney fibrosis in the human-derived PDGFRβ^{+} cells with the protocol that I applied earlier in my project (results section 4.1.4).

To highlight one of the SMAD independent pathways induced by TGF-β signaling, I am focusing on the activation of Rho GTPases and their effector kinases Rho coiled-coil kinases (ROCK) 1 and 2, players that are further involved in mechano-transduction (Walraven & Hinz, 2018). Linking it to the phosphoproteomics hits and activated TFs from before, studies already demonstrated that ROCK and actin-regulated MKL1/SRF

activation result in the TNS1 expression that might be involved in driving differentiation of myofibroblasts and ECM deposition (Bernau et al., 2017). These findings are supported by another study that showed the prevention of myofibroblast differentiation and ECM deposition in lung fibroblasts and precision cut lung slices resulting from the inhibition of ROCK1 and 2 (Wu et al., 2021). In this study, ROCK1 and ROCK2 both displayed reduced activities, unexpectedly (Figure *34*). While ROCK1 mRNA and protein abundance increased gradually over time following TGF-β stimulation, it did not meet the log2 fold-change and adjusted p-value thresholds mentioned before (Figure *71*). In contrast to this, ROCK2 expression decreased, with significance observed only at the 96 h mark in the transcriptomics data. It would be interesting to further investigate the discrepancies of my data with literature, to understand whether there are cell-type dependent effects and to further understand the role of ROCK1 and ROCK2 in the context of the used cellular system of this study.

In summary, the footprint-based methods estimated altered activity of TFs and kinases/phosphatases that were mostly not detected as hits in the multi-omics approach since those factors did not meet the selection criteria such as absolute log2 fold-change > log(1.5), an adjusted p-value < 0.05, and occurrence as hits across different omics modalities. Therefore, my observations underscore the estimated potential of footprint methods in enhancing our understanding by providing activity inferences of essential signaling molecules, such as TFs and kinases/phosphatases, based on changes of downstream targets. However, it is crucial to acknowledge that these tools are inherently biased by the prior knowledge used, since well-studied pathways and factors are favored, thereby overlooking less-well investigated factors (Garrido-Rodriguez et al., 2022; Hishikawa et al., 2018). Therefore, I decided to validate a set of TFs to confirm the results obtained using footprint analysis and find possible modulators of kidney fibrosis. BHLHE40, NR4A1, E2F1, SMAD1, FLI1 and HNF4G all showed an increased activity at 1 h post TGF-β treatment (Figure *34*). Based on this observation and initial screening of literature, I argued that inhibiting these TFs could influence ECM deposition, given their potential fibrogenic role. The data of the validations of section 4.4.3 will be discussed in 5.6.2.

5.5.2. Multi-omics Network to Study Mechanisms of Fibrosis

For further integration of the data, we generated two networks, one for the early mechanisms (**Figure *72***) and one for the late (**Figure *73***) to consider the temporal and multimodal information contained in the dataset. The separation of time points into early and late network was based on observations of the data, clustering and correlation of datasets for the different time points.

The rationale behind clustering hits rather than the most variable features stemmed from the fact that in the phosphoproteomics data, the lines depicting average log2 fold-change per cluster did not meet the threshold of log(1.5), and most features were not significantly deregulated (data not shown). For clustering of the hits, only two clusters were suggested (up and downregulated features) (**Figure *36***, **Figure *37***, **Figure *38***, **Figure *39***). However, I wanted to assess the changes of genes and proteins over time to decide which time points to include in either of the two networks (early vs late) and therefore opted for four clusters to capture temporal changes better. This supported the selection of time points for network inclusion (early vs late).
In addition to clustering, correlation analysis of datasets was performed (**Figure *40***). Correlation of the top 10% variable features per omics and hits yielded very similar results. Considering both approaches and the obtained results, I selected 0.08, 1, and 12 h for TFs and kinases/phosphatases, as well as 12 and 24 h for secretomics data for inclusion in the early network. The rationale behind this decision was that signaling and protein expression might delay the secretion, hence setting the cut-off at 24 h for the secretomics hits. For the late network, time points 24 to 96 h for TFs and kinases/phosphatases, and 48 to 96 hours for secretomics were included.

For the network, we used all transcripts measured to filter the prior knowledge for genes that are expressed in the used cellular system to generate a context specific network. The final networks actually consist of two networks each. In round one, to generate the first subnetwork, the upstream inputs (TGF-β, TGF-β & early secreted factors) were linked to downstream measurements (TFs, kinases/phosphatases). In round two, the TFs and kinases/phosphatases were used as upstream input and secreted factors as downstream measurements. In the post processing, a final network was generated based on both rounds.

These steps resulted in an early and a late network (Figure *41*), where the nodes constitute TFs, kinases/phosphatases and secreted factors and the edges represent the relationship between these factors, in our case labelled as “activators”/“inhibitors”. Additionally, we added COL1A1 as an enriched secreted factor. Even though COL1A1, based on the log2 fold-change and adjusted p-value thresholds, was not a hit in the secretomics, we included it as the imaging data clearly showed an increase of deposited Col I. In general, there are more nodes and edges in the late network than in the early network. This is in line with my expectations and previous sections of this thesis, showing that the number of hits per omics increase over time (Figure *27*). Directly downstream of TGF-β in the early network are SMAD1, 3, 4, 5 and MAPK1, as we constrained the PKN to enforce SMAD3 and 4 as well as MAPK1 downstream of TGF-β. In the late network, downstream of TGF-β is paired box protein PAX8.
Initially, I was surprised by the results as I expected to see pathways as we are used to see in text books such as SMAD2, 3, 4 complex formations, translocation into the nucleus and initiation of transcription. But in the network generated in this study the shortest connection was displayed. Nevertheless, I did not expect SMAD1 and 5 as they are associated with BMP signaling which is antagonistic to TGF-β signaling (Massagué, 2012; Matsubara et al., 2015; Pannu et al., 2007; Wei et al., 2020) (more to this in a later section).

The nodes of each network were used to perform pathway enrichment analysis using the Reactome Pathway Database. Besides pathways that were expected, including signaling by receptor tyrosine kinases and TGF-β family members, or extracellular matrix organisation and degradation, I was surprised by the appearance of the pathways axon guidance and nervous system development (Figure *42*). However, by looking into the nodes that were contained in the later pathways, there were several collagens (COL5A1, COL6A2, COL4A2), MMP2, integrins (ITGB1, ITGA2), and factors involved in MAPK singaling as well as signaling by receptor tyrosine kinases (MAP2K1, MAPK1, PTK2, MET, EGFR, AKT1, PIK3CA, PRKACA, RAC1, RHOA, ROCK1, ROCK2, YAP1, RPS6KA1, SRC, CREB1, LAMB1, LAMC1). In general, most of the pathways have a stronger significance in the late network as there are generally more factors that are taken into account due to the higher number of nodes. Nevertheless, the pathways with a stronger significance in the early network comprise signaling by interleukins, AKT signaling activated by PIP3, intracellular signaling by second

messengers, diseases of signal transduction by growth factor receptors and second messengers as well as signaling by GPCR.
A possible explanation might be that the pathways highlighted in the early network could represent the immediate response to TGF-β stimulation whereas the late network might capture better the mechanisms associated with fibrosis. In the context of fibrosis, pathways like ECM organization and degradation might become more prominent.

To further analyze and interpret each of the networks, we preformed clustering of the networks, and subsequent pathway enrichment analysis (**Figure *43***, **Figure *44***). For instance, the more prominent interleukin signaling in the early network could be explained by factors involved in Interleukin-4 and -13 signaling. Additionally, pathways that are important in development but become reactivated in renal injury involve Notch, Wnt and hedgehog (Edeling et al., 2016). Linking to this, pre-NOTCH expression and processing/transcription and translation showed up as enriched pathways in the early network. Interestingly, regulation of CDH11 expression and functions was shown to be enriched, linking to CDH11 which was a hit in all data modalities (**Figure *31***).
Supporting my hypothesis that the cells become senescent with time, the clustering of the late network showed enrichment of factors involved in cellular senescence. Another factor that is relevant in fibrosis is oxidative stress (Gewin et al., 2017; Rockey et al., 2015; Sziksz et al., 2015), involving hypoxia-inducible factor 1-alpha (HIF1A, with upregulated activity at 96 h in **Figure *34***) and shown by the enrichment of the pathway regulation of gene expression by hypoxia-inducible factor. Interestingly, there was signaling by BMP enriched in one cluster, which was mainly driven by BMPR1A, SMAD1, SMAD4 and SMAD5.
In general, clustering methods are not very robust and sometimes predicted pathways change when clustering methods change. Nevertheless, pathways such as degradation of the extracellular matrix, extracellular matrix organization and signaling by receptor tyrosine kinase were robustly detected as enriched in both, the early and late networks, even if the clustering methods changed. One of the reasons for the enrichment of ECM regulating pathways could be e.g. SP1 which was a node with many edges especially to ECM regulating enzymes and matrisome proteins (**Figure *45***).

Even though network generation tools offer valuable options to generate testable hypotheses, thereby advancing our understanding of complex diseases such as CKD, the results should be interpreted cautiously. From my point of view, a significant limitation of these methods lies in the inherent bias coming from prior knowledge. This research bias stems from certain well-studied biological contexts, such as cancer, and a small subset of extensively investigated proteins (e.g. p53). Therefore, it is easier to link the functions of these factors than of understudied proteins and pathways. At the network level, this means that a minority of extensively studied proteins drive the majority of interactions and often exhibit multifunctionality, potentially biasing interpretations (Garrido-Rodriguez et al., 2022). On the contrary, this could also highlight their intricate importance in regulating different pathways. Another example is that the functions of only 5% of known phosphorylations in human cells are understood, even though the contributions of phosphorylation events are essential in cellular signaling. Moreover, over 90% of these known phosphorylation events involve a selected group of well-examined kinases (Needham et al., 2019). In summary, these approaches are limited to our current knowledgebase and only provide an approximation of the reality.

5.6. Observations form Validation Experiments

5.6.1. Optimized siRNA Transfection Protocol Resulted in Efficient Knock-down of Target Genes

After confirming that the siRNA transfection protocol and reagents worked in the cellular model, by knock-down of GM130 and SMAD2 (**Figure *46***, **Figure *47***), I aimed to investigate the impact of target gene knock-downs on ECM deposition. For this RT-qPCR read-outs and immunofluorescence staining were used. Since new batches of anti-Col I antibody and other antibodies that I tried did not work, I utilized CNA35, a fluorescently labelled dye that binds fibrillar collagen (Baues et al., 2020; Klinkhammer & Boor, 2023; Krahn et al., 2006). However, due to increased autofluorescence, most likely coming from the stress induced by the siRNA transfection, an additional anti-GFP antibody (secondary 647) staining was employed to mitigate the signal stemming from autofluorescent (apoptotic) cells. In parallel, the mRNA expression levels of siRNA target genes and their downstream targets were monitored.

Initially, knock-down of COL1A1 was performed to optimize the protocol for further validation experiments. Assuming that the protein levels of target genes were efficiently downregulated at 48 h, I started the treatment with or without TGF-β and kept the cells under treatment for further 48 h. This set-up was chosen as changes in ECM deposition could be assessed after 48 h of TGF-β treatment. As expected, treatment with siCOL1A1, using three different siRNAs, resulted in significant decrease of COL1A1 mRNA levels and decreased ECM deposition shown by immunofluorescence staining (**Figure *50***).

Given the integral part of FN1 for collagen assembly into the ECM (Kendall & Feghali-bostwick, 2014; Moriya et al., 2011), I did not expect to see an increased COL1A1 mRNA expression and an increase of deposited ECM upon knock-down of FN1, which was further boosted by TGF-β treatment (**Figure *51***, **Figure *52***). However, it has also been previously shown that knock-out in liver of mice resulted in an increased collagen production, TGF-β activity and liver fibrosis (Kendall & Feghali-bostwick, 2014; Moriya et al., 2011).

To conclude, knock-down and subsequent TGF-β stimulation effectively maintained reduced mRNA and protein expression, while TGF-β stimulation induced increased COL1A1 mRNA and protein expression.

5.6.2. Observations from TF Validations

In order to identify factors potentially modulating ECM deposition, I selected a set of TFs estimated to display an upregulated activity 1 h post TGF-β treatment (**Figure *34***). I argued that these factors, besides established factors like SMAD4 and SRF, might influence the cells response to fibrotic stimuli such as TGF-β. By manipulating these factors, I could study their potential impact on ECM position. Based on these findings and initial literature research, I shortlisted the following TFs: Mothers against decapentaplegic homolog 1 (SMAD1), Transcription Factor E2F1, Friend leukaemia integration 1 transcription factor (FLI1), Nuclear Receptor Subfamily 4 Group A Member 1 (NR4A1), Hepatocyte nuclear factor 4-gamma (HNF4G) and Class E basic helix-loop-helix protein 40 (BHLHE40) for validation.

I confirmed the knock-down with RT-qPCR to monitor whether the mRNA expression of the selected TFs decreased upon siRNA treatment (**Figure *53***). This was the case for all of them. In general, the mRNA expression of all tested TFs increased over time

(siNeg9 treated samples), however, only BHLHE40 significantly increased with TGF-β treatment (Figure *81*), being in line with previous findings, while knock-down led to a significant reduction in expression levels over time (Figure *53*). To assess the potential fibrogenic effect of the selected TFs, I monitored the mRNA expression levels of COL1A1. The upregulation of COL1A1 mRNA upon TGF-β treatment was visible in all experiments, as anticipated by initial RT-qPCR and imaging-based experiments (Figure *53*, Figure *81*, Figure *54*, Figure *55*).

While I summarized the effects of the knock-down at the last time point on the different genes tested, in the heatmap (Figure *53*), I took a closer look into the data to assess the effect of interactions between siRNA and TGF-β treatment on the expression of COL1A1 mRNA levels and therefore used anova (Figure *77* - Figure *81* , Table *18* - Table *23*). This was performed for the last time point (96 h of siRNA ± 48 h TGF-β treatment).

No Effect of HNF4G and BHLHE40 on COL1A1 mRNA Expression Levels but Opposing Effects on ECM Accumulation

In contrast to the t-test results (Figure *53*), siHNF4G and siBHLHE40 treatment exerted no significant effect on COL1A1 mRNA expression, neither in the control nor TGF-β treated conditions, using the anova (compared to the corresponding siNeg9 samples) (Table *21*, Table *22*).

Comparing these data to the immunofluorescence staining of ECM, siBHLHE40 treatment induced a significantly higher deposition of ECM, which was further boosted by stimulation with TGF-β (Figure *54*, Figure *55*). These data indicate, that BHLHE40 might not regulate COL1A1/ECM deposition on the transcriptional level but might influence factors that control it. On the contrary, treatment with siHNF4G only led to a significant decrease of deposited ECM in the TGF-β treated condition when compared to the siNeg9 + TGF-β condition.

To date there is not much known about the role of HNF4G in the context of kidney fibrosis. A study on the PDGFR-β activation on renal mesenchymal cells mentioned HNF4G as a TF which expression is downregulated in a CKD patient's cohort (Buhl et al., 2020). Even though it would be interesting to further investigate the effect of HNF4G, especially since siRNA treatment led to a reduced ECM deposition in my assay, I did not follow up on this TF due to time limitations.

My hypothesis for BHLHE40, which is involved in cell cycle control (Kiss et al., 2020), stemmed from a paper highlighting the impact of circadian control on collagen homeostasis (J. Chang et al., 2020). Additionally, BHLHE40 was found in my early network (Figure *57*), predicted to inhibit ID1 which would inhibit periostin (POSTN), a secreted ECM protein and marker for myofibroblasts (Kuppe et al., 2021). I postulated, that by BHLHE40 knock-down, more ID1 would be present leading to less POSTN and hence less ECM deposition, since it could be affected by loss of POSTN that is involved in cross-linking of the ECM. Additionally, the protein was shown to be involved in cell adhesion, migration, differentiation and renal fibrosis (François & Chatziantoniou, 2018). Since the opposite was observed in initial experiments, and due to time constrains, I did not further validate BHLHE40. Nevertheless, it would be interesting to follow up on initial experiments, especially since there could be an effect when extending the time of the experiment and assessing the changes of ECM at a later time point.

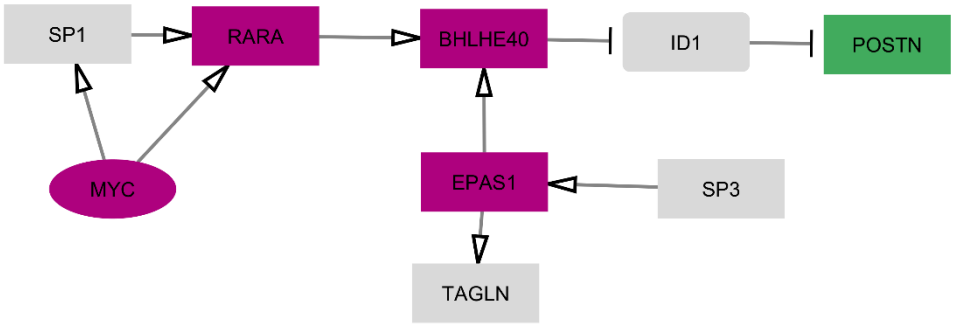

Figure 57 Zoom in on BHLHE40 with up- and downstream nodes in the early mechanistic network. The colors are representing the different modalities of the nodes, green for secreted factors, purple for TFs and grey as factors imputed by prior knowledge. The ellipse shape are factors which expression/activity is downregulated while squares indicate factors that are upregulated.

Knock-down Inhibition of FLI1 Resulting in Increased COL1A1 mRNA and ECM

The knock-down of FLI1 with the first siRNA (siFLI1a) resulted in a significant increase of COL1A1 mRNA expression in the TGF-β treated samples compared to the siNeg9 + TGF-β but not with the second siRNA tested (Figure *53*, Figure *80*, Table *20*). For the first siRNA, however, the distribution of data points for the FLI1 knock-down was high and the uneven distribution of sample numbers (between siFLI1 and siNeg9 treated samples) suggests the need to increase the number of replicates to make a definite statement (Figure *80*). The same is true for the second siRNA tested. Both siRNAs tested for FLI1 led to a significant increase of deposited ECM in the TGF-β treated samples (Figure *54*, Figure *55*). These data suggest that FLI1 might exert an antifibrotic

effect that gets inhibited upon knock-down, leading to more COL1A1 mRNA and an increased ECM deposition.

Having a closer look at the activity estimation of FLI1, it exhibited a bimodal behavior, being decreased at 0.08, 48 and 96 h, while increased at 1, 24 and 72 h (Figure *34*). Similarly, its abundance changed with time of TGF-β treatment (Figure *71*). On proteomics level, FLI1 abundance decreased at 0.05, and 12 h and increased at 1, 48, 72 and 96 h, however, with no log2 fold-change bigger than log(1.5). Despite none of the log2 fold-change values passed the cut-off in the transcriptomics data, the mRNA expression of FLI1 also varied across time with an increased expression at 24 and 72 h and a decreased expression at 1 h.

The involvement of FLI1 has been shown for a number of diseases, including cancer and fibrosis. Deficiency of FLI1 is implicated in activation of fibrotic pathways as FLI1 is a natural inhibitor of collagen genes (Kendall & Feghali-bostwick, 2014; Mikhailova et al., 2023). Therefore, I expected an increase of COL1A1 mRNA expression and ECM upon inhibition of FLI1 via siRNA treatment, which was confirmed in my experiments (Figure *53*), thereby further validating my model. In the early network, FLI1 proposetly regulated NOTCH1 and EWSR1 (Figure *58*). This hypothesis, however, needs to be validated experimentally.

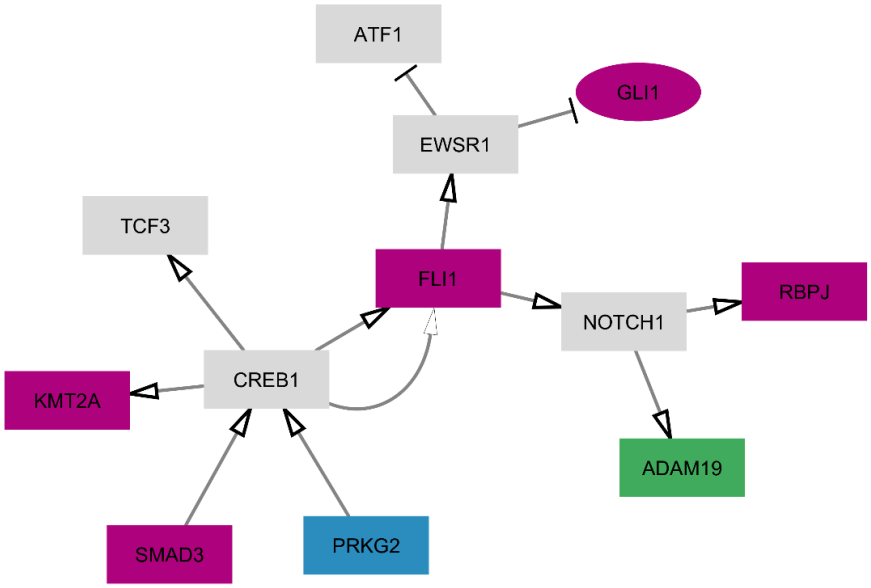

Figure 58 Zoom in on FLI1 with up- and downstream nodes in the early mechanistic network. The colors are representing the different modalities of the nodes, green for secreted factors, blue for kinases/phosphatases, purple for TFs and grey as factors imputed by prior knowledge. The ellipse shape are factors which expression/activity is downregulated while squares indicate factors that are upregulated.

Validation of SMAD1 leading to unexpected fibrotic response

Interestingly, knock-down of SMAD1 with the first siRNA significantly reduced the COL1A1 mRNA expression, in the TGF-β treated samples compared to the corresponding siNeg9 controls (Figure *53*, Figure *78*, Table *18*). The second siRNA targeting SMAD1, did not result in a significant change of COL1A1 mRNA expression, which can be attributed to a too low number of samples tested. I did not follow up further on this siRNA as I realized that it has an off-target effect on SMAD5 (Figure *76*).

This was also the reason why I tested a third siRNA, which, to my surprise, led to a significant increase of COL1A1 mRNA levels in the samples stimulated with TGF-β compared to the corresponding siNeg9 samples.
Consistent with the RT-qPCR data, the ECM significantly decreased upon treatment with siSMAD1a, especially when comparing the TGF-β treated condition to the siNeg9 TGF-β condition. On the contrary, siSMAD1_c treatment led to a significant increase of deposited ECM in the TGF-β treated condition (Figure *54*, Figure *55*).
To confirm either effect, an additional siRNA needs to be tested. Nevertheless, the data showed the importance of SMAD1 in the fibrotic response of the used cellular model. To further understand how COL1A1 expression and ECM deposition is regulated by the tested transcription factors, I monitored how other downstream targets and interactors of SMAD1 were influenced and if knock-down of possible interactors might have shown a similar effect. As ID2 and ID3 are proposed downstream targets of SMAD1, the mRNA expression of these factors was also assessed in this study. ID2 abundance was significantly decreased upon siSMAD1 + TGF-β (first and third siRNA tested) treatment compared to the siNeg9 + TGF-β samples while there was no significant effect on ID3 expression when using the first siSMAD1 but a significant decrease of ID3 upon siSMAD1_C + TGF-β treatment (Table *18*). To investigate an anticipated additive effect of inhibiting SMAD1, ID2 and ID3, I conducted double and triple knock-downs with siRNAs targeting ID2, ID3 and the initial siRNA against SMAD1. Preliminary data hinted towards an effect of ID2 and ID3 knock-down leading to decreased COL1A1 mRNA expression levels. However, double and triple knock-downs did not result in a further decrease of COL1A1 mRNA abundance (n = 1, data not shown).

Besides the estimated increased activity of SMAD1 at 1 h, it also showed an increased activity at 96 h post TGF-β treatment and a decreased activity at 12 and 48 h (Figure *34*). On mRNA level (Figure *71*), SMAD1 abundance decreased with increasing time of TGF-β treatment. However, this was only significant at the 96 h time point and did not pass the log2 fold-change cut-off at any time. In the proteomics data, its abundance increased but was not significant at any time point nor passed the log2 fold-change threshold.

In the literature, activation of SMAD1/SMAD5, e.g. by BMP-7, the 'inhibitor of TGF-β', is known to inhibit the SMAD3 and ERK dependent gene transcription, thereby exerting an anti-fibrotic effect (C. H. Park & Yoo, 2022). However, while canonical TGF-β signaling via SMAD2/SMAD3 does not involve SMAD1, TGF-β can also activate other pathways that lead to a fibrogenic response. It has been previously shown, for instance, that there is a CCN2-dependent pathway (CCN2 is one of our top upregulated hits in 3 out of 4 omics modalities used in this study) that induced phosphorylation of SMAD1 and ERK1/2 and requires the CCN2/β3 integrin protein complex to do so. The mentioned study demonstrated that blockage of CCN2 suspended collagen mRNA and protein levels induced by TGF-β, thereby suggesting the contribution of a CCN2/αvβ3 integrin/Src/Smad1 axis to the pro-fibrotic TGF-β signaling (Nakerakanti et al., 2011; Pannu et al., 2007). It would be interesting to further examine the importance of this pathway in the used cellular system. As described in the early network generated in this study, proposing TGF-β induced SMAD1 regulation leading to changes of CCN2 (CTGF), would be in line with parts of the feedback loops described in the study of Nakerakanti et al. (2011) (Nakerakanti et al., 2011) (Figure *59*Figure *45*). Alternatively, the effect of SMAD1 on ECM deposition could be controlled via other means, or could be a time-dependent effect e.g. as demonstrated in the late network (Figure *45*). Further experiments and a deeper literature research are needed to come up with a relevant hypothesis and confirm the effect of siSMAD1 knock-down on ECM deposition.

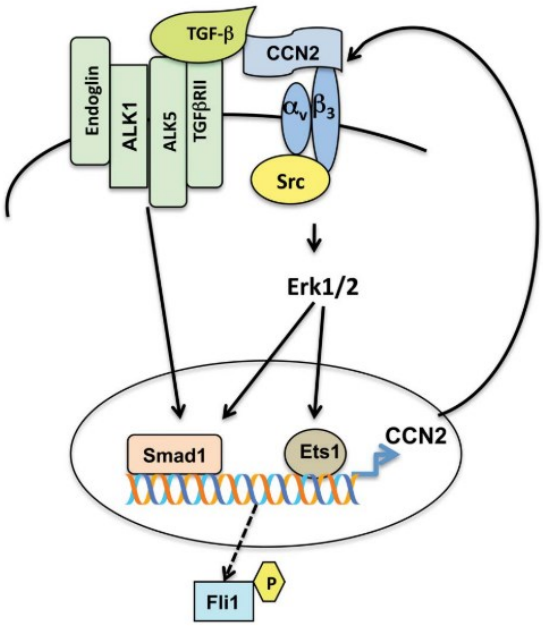

Figure 59 TGF-β/SMAD1/CCN2 positive feedback loop. Model proposed by Nakerakanti et al. (2011) for the positive loop between TGF-b/Smad1 and CCN2 (taken from Nakerakanti et al., 2011).

Inhibition of NR4A1 and E2F1 Leading to Increased COL1A1 mRNA and ECM via Similar Means?

For the NR4A1 knock-down, the post hoc table showed that the difference between siNR4A1 + TGF-β treated samples and siNeg9 + TGF-β treated samples was significant (Table *23*). However, also there it would be best to increase the number of replicates as the interaction term is not significant in the anova table. A similar effect was seen in the E2F1 knock-down samples, where the interaction term suggested a significant increase of COL1A1 mRNA expression in the control samples compared to the siNeg9 controls (Table *19*). For the second siRNA targeting E2F1 used, the samples showed a high distribution as well, leading to no significant differences when compared to the siNeg9 samples, neither in the untreated nor the TGF-β treated conditions (Figure *77*). While knock-down of E2F1 resulted in an enhanced ECM deposition, with both siRNAs and with as well as without TGF-β treatment, the siNR4A1 treatment only slightly increased the deposition of ECM (Figure *54*, Figure *55*).

E2F1 activity was estimated to be increased at 0.08 to 24 h post TGF-β treatment but is decreased at 96 h (Figure *34*). The transcriptomics data showed a non-significant decrease of E2F1 abundance over time. In comparison to this, the activity if NR4A1 was inferred as constantly active with a peak at 1 h after TGF-β stimulation. Except for the increased mRNA abundance of NR4A1 in the transcriptomics data at 1 h, which

was not significant, it showed a decreased abundance over time with no change in the proteomics data.

NR4A1 involvement in fibrosis was previously demonstrated by regulating TGF-β signaling through AKT- and HDAC-dependent mechanisms and in the context of myofibroblast activation in mice (Devos et al., 2023; Kuppe et al., 2021; Meng et al., 2016). However, due to the minor effect of NR4A1 knock-down on ECM deposition and COL1A1 mRNA expression, I did not follow up on this TF.

Unexpectedly, E2F1 knock-down with both siRNAs led to a significantly increased mRNA expression of SERPINE1 (PAI1), an inhibitor of ECM degradation (Bergheim et al., 2006) (**Figure *53***, **Table *19***). E2F1 has been previously shown in the context of fibrosis in other diseases like liver fibrosis. In Denechaud et al. (2017), it has been shown that E2F1-/- mice developed liver fibrosis (Denechaud et al., 2017). In contrary to this, Liao et al. (2021), demonstrated that TGF-β induced human cardiac fibroblasts showed upregulated E2F1 expression. Moreover, they demonstrated that knock-down of E2F1 led to decreased cell differentiation and suppressed COL1A1 expression (Liao, Xie, et al., 2021). Additional experiments are needed to evaluate the importance of E2F1 on abundance changes of COL1A1 to further test whether E2F1 affects the ECM deposition via direct transcriptional regulation or via modulating the activity of downstream targets that are relevant for the increased ECM. One possible way of regulating the turnover/degradation of the deposited ECM could be by controlling the expression of factors like SERPINE1, or, as suggested by the network, by controlling HIC1 and FOXM1/ROCK1 (Liao, Qi, et al., 2021).

6. Conclusion and Outlook

In this study, a comprehensive investigation of the dynamic, pro-fibrotic molecular response of PDFGRβ^{+} cells, the main precursors of myofibroblasts, to TGF-β was provided. After initial experiments that validated the suitability of the cellular system for this purpose, the cells were used for multi-omics analysis, including transcriptomics, proteomics, and phosphoproteomics. In all omics modalities, pro-fibrotic pathways and factors were shown to be enriched, while anti-inflammatory pathways were downregulated. Moreover, already known biomarkers for CKD and other fibrotic diseases were identified, further validating the approach. Additionally, the changes on deposited Col I and ECM were investigated over the course of the multi-omics experiment, showing a significant increase upon stimulation with TGF-β and accumulating over time.

Furthermore, advanced computational tools like footprint analysis were applied to gain further information on deregulated TFs and kinases/phosphatases, which inferred activities changed despite the fact that many factors did not show alterations in mRNA or protein expression levels post TGF-β treatment. While integration of the different omics modalities into an early and late mechanistic network supported the validity of earlier analysis, it also provided testable hypothesis to advance the understanding of kidney fibrosis.

Validation experiments focused on TFs, whose activity was estimated to be increased 1 h post TGF-β treatment, and therefore were identified as potential regulators of the fibrotic response. These TFs included BHLHE40, HNF4G, FLI1, SMAD1, NR4A1, and E2F1. Besides assessment of TF and downstream target gene expression, the effect of the siRNA treatment on deposited ECM, as a central part of fibrosis, was evaluated. Reduced FLI1 expression led to increased COL1A1 mRNA expression and ECM deposition confirming the previously anticipated anti-fibrotic role of FLI1. BHLHE40 and HNF4G, based on the data, might not control the mRNA expression of COL1A1 but rather the expression of factors that modulate ECM deposition via other means as knock-down of BHLHE40 resulted in increased ECM deposition while siHNF4G treatment led to a reduction showing their importance in ECM homeostasis. While there was a significant increase of COL1A1 mRNA expression via siRNA treatment targeting

NR4A1 or E2F1, the effect on ECM deposition was minor in the case of NR4A1 inhibition but led to a stronger increase of upon knock-down of E2F1. For SMAD1 knock-down, I observed both, an increase and decrease of COL1A1 mRNA expression and ECM deposition, depending in the siRNAs used.
Therefore, further experimental validations are needed to confirm the effects and dissect the mechanisms of ECM deposition in this cellular model. Understanding this could provide a deeper insight and facilitate identification of potential therapeutic targets and treatment options for preventing or halting the progression of fibrosis.

To further highlight the value of this cellular model, like easy access (compared to human tissue samples) and possibility to perturb the system, I conducted treatment with known anti-fibrotic compounds. However, not all of them showed the expected decrease in ECM which could be attributed to the experimental set-up or further highlighting the importance to understand the fibrotic mechanism of myofibroblasts arising from different tissues and possibly even different origins. Therefore, the cellular model offers a valuable tool to further investigate the effect of different compounds on cellular signaling and ECM deposition and could also be used to test combinatorial therapies. Additionally, it would be interesting to test other factors that induce a fibrotic response like induction of hypoxia, mechanical stress or stimulation with other growth factors such as PDGFRβ.

Additionally, single cell RNA expression profile of myofibroblasts derived from CKD patients was compared to the data that I generated. Thereby, an increase in deregulated genes over time in multi-omics data was shown, which were also expressed in myofibroblasts of CKD patients, further demonstrating the relevance of the cellular response to TGF-β treatment in driving differentiation towards myofibroblasts, which is crucial in the context of fibrosis. This could be taken further by mapping my data to different stages of fibroblast/myofibroblasts and inferring cell-cell crosstalk with other cell types in the kidney.

To conclude, the data provided valuable insights into the molecular mechanisms of kidney fibrosis and ECM deposition, proposing potential biomarkers as well as therapeutic targets and advancing our understanding of disease pathogenesis.

Nevertheless, further research is needed to validate identified targets and explore novel therapeutic interventions.

To add further weight to the findings of this project, it would be interesting to investigate the effect of proposed modulators of ECM in mouse models or human precision-cut kidney slices. This would add valuable insight on the effect on ECM deposition, cellular crosstalk and fibrosis progression in a more physiological context. Clinical relevance of proposed drug targets and biomarkers could be further assessed by integration of data obtained from longitudinal studies that track disease progression over time. In the case of CKD, an example of such a study is the Nephrotic Syndrome Study Network (NEPTUNE) (Lake et al., 2021; Mariani et al., 2023). For this multi-center prospective study, data on e.g. medical history, medication, eGFR and RNA sequencing are collected over time from patients suffering from proteinuric kidney diseases (Lake et al., 2021; Mariani et al., 2023). Overall, development of effective therapies will require integrated efforts of scientists, clinicians and industry partners to address the complex challenges of fibrotic diseases. The results of the current PhD thesis provide one piece of the basis for such efforts.

Bibliography

Adler, M., Mayo, A., Zhou, X., Franklin, R., Meizlish, M., Medzhitov, R., Kallenberger, S., & Alon, U. (2019). Principles of Cell Circuits for Tissue Repair and Fibrosis. *BioRxiv*, 1–22.

Alpern, R. J., Caplan, M. J., & Moe, O. W. (Eds.). (2013). *The Kidney Physiology and Pathology* (5th ed.).

An, B., Lin, Y. S., & Brodsky, B. (2016). Collagen interactions: Drug design and delivery. *Advanced Drug Delivery Reviews*, *97*, 69–84. https://doi.org/10.1016/j.addr.2015.11.013

Argilés, À., Siwy, J., Duranton, F., Gayrard, N., Dakna, M., Lundin, U., Osaba, L., Delles, C., Mourad, G., Weinberger, K. M., & Mischak, H. (2013). CKD273, a New Proteomics Classifier Assessing CKD and Its Prognosis. *PLoS ONE*, *8*(5). https://doi.org/10.1371/journal.pone.0062837

Arif, M., Basu, A., Wolf, K. M., Park, J. K., Pommerolle, L., Behee, M., Gochuico, B. R., & Cinar, R. (2023). *An Integrative Multiomics Framework for Identification of Therapeutic Targets in Pulmonary Fibrosis*. *2207454*, 1–20. https://doi.org/10.1002/advs.202207454

Badia-i-Mompel, P., Vélez Santiago, J., Braunger, J., Geiss, C., Dimitrov, D., Müller-Dott, S., Taus, P., Dugourd, A., Holland, C. H., Ramirez Flores, R. O., & Saez-Rodriguez, J. (2022). decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinformatics Advances*, *2*(1), 1–3. https://doi.org/10.1093/bioadv/vbac016

Baues, M., Klinkhammer, B. M., Ehling, J., Gremse, F., van Zandvoort, M. A. M. J., Reutelingsperger, C. P. M., Daniel, C., Amann, K., Bábíčková, J., Kiessling, F., Floege, J., Lammers, T., & Boor, P. (2020). A collagen-binding protein enables molecular imaging of kidney fibrosis in vivo. *Kidney International*, *97*(3), 609–614. https://doi.org/10.1016/j.kint.2019.08.029

Bernau, K., Torr, E. E., Evans, M. D., Aoki, J. K., Ngam, C. R., & Sandbo, N. (2017). Tensin 1 is essential for myofibroblast differentiation and extracellular matrix formation. *American Journal of Respiratory Cell and Molecular Biology*, *56*(4), 465–476. https://doi.org/10.1165/rcmb.2016-0104OC

Bigaeva, E., Cavanzo, N. P., Stribos, E. G. D., de Jong, A. J., Biel, C., Mutsaers, H. A. M., Jensen, M. S., Nørregaard, R., Leliveld, A. M., de Jong, I. J., Hillebrands, J. L., van Goor, H., Boersema, M., Bank, R. A., & Olinga, P. (2020). Predictive value of precision-cut kidney slices as an ex vivo screening platform for therapeutics in human renal fibrosis. In *Pharmaceutics* (Vol. 12, Issue 5). https://doi.org/10.3390/pharmaceutics12050459

Bigaeva, E., Stribos, E. G. D., Mutsaers, H. A. M., Piersma, B., Leliveld, A. M., de Jong, I. J., Bank, R. A., Seelen, M. A., van Goor, H., Wollin, L., Olinga, P., & Boersema, M. (2020). Inhibition of tyrosine kinase receptor signaling attenuates fibrogenesis in an ex vivo model of human renal fibrosis. *American Journal of Physiology - Renal Physiology*, *318*(1), F117–F134. https://doi.org/10.1152/ajprenal.00108.2019

Black, L. M., Lever, J. M., & Agarwal, A. (2019). Renal Inflammation and Fibrosis: A Double-edged Sword. *Journal of Histochemistry and Cytochemistry*, *67*(9), 663–681. https://doi.org/10.1369/0022155419852932

Bon, H., Hales, P., Lumb, S., Holdsworth, G., Johnson, T., Qureshi, O., & Twomey, B. M. (2019). Spontaneous Extracellular Matrix Accumulation in a Human in vitro Model of Renal Fibrosis Is Mediated by αv Integrins. *Nephron*, *142*(4), 329–350. https://doi.org/10.1159/000499506

Borgo, C., D'Amore, C., Sarno, S., Salvi, M., & Ruzzene, M. (2021). Protein kinase CK2: a potential therapeutic target for diverse human diseases. *Signal Transduction and Targeted Therapy*, *6*(1), 1–20. https://doi.org/10.1038/s41392-021-00567-7

Buhl, E. M., Djudjaj, S., Klinkhammer, B. M., Ermert, K., Puelles, V. G., Lindenmeyer, M. T., Cohen, C. D., He, C., Borkham-Kamphorst, E., Weiskirchen, R., Denecke, B., Trairatphisan, P., Saez-Rodriguez, J., Huber, T. B., Olson, L. E., Floege, J., & Boor, P. (2020). Dysregulated mesenchymal PDGFR-β drives kidney fibrosis. *EMBO Molecular Medicine*, *12*(3), 1–20. https://doi.org/10.15252/emmm.201911021

Cañadas-Garre, M., Anderson, K., McGoldrick, J., Maxwell, A. P., & McKnight, A. J. (2018). Genomic approaches in the search for molecular biomarkers in chronic kidney disease. In *Journal of Translational Medicine* (Vol. 16, Issue 1). BioMed Central. https://doi.org/10.1186/s12967-018-1664-7

Cañadas-Garre, M., Anderson, K., McGoldrick, J., Maxwell, A. P., & McKnight, A. J. (2019). Proteomic and metabolomic approaches in the search for biomarkers in chronic kidney disease. *Journal of Proteomics*, *193*(July), 93–122. https://doi.org/10.1016/j.jprot.2018.09.020

Cannito, S., Novo, E., & Parola, M. (2017). *Therapeutic pro-fibrogenic signaling pathways in fibroblasts*. *121*, 57–84. https://doi.org/10.1016/j.addr.2017.05.017

Chana, R. S., Martin, J., Rahman, E. U., & Wheeler, D. C. (2003). Monocyte adhesion to mesangial matrix modulates cytokine and metalloproteinase production. *Kidney International*, *63*, 889–898. https://doi.org/10.1046/j.1523-1755.2003.00828.x

Chang, F. C., Chou, Y. H., Chen, Y. T., & Lin, S. L. (2012). Novel insights into pericyte-myofibroblast transition and therapeutic targets in renal fibrosis. *Journal of the Formosan Medical Association*, *111*(11), 589–598. https://doi.org/10.1016/j.jfma.2012.09.008

Chang, J., Garva, R., Pickard, A., Yeung, C. Y. C., Mallikarjun, V., Swift, J., Holmes, D. F., Calverley, B., Lu, Y., Adamson, A., Raymond-Hayling, H., Jensen, O., Shearer, T., Meng, Q. J., & Kadler, K. E. (2020). Circadian control of the secretory pathway maintains collagen homeostasis. *Nature Cell Biology*, *22*(1), 74–86. https://doi.org/10.1038/s41556-019-0441-z

Chen, C. Z. C., Peng, Y. X., Wang, Z. B., Fish, P. V., Kaar, J. L., Koepsel, R. R., Russell, A. J., Lareu, R. R., & Raghunath, M. (2009). The Scar-in-a-Jar: Studying potential antifibrotic compounds from the epigenetic to extracellular level in a single well. *British Journal of Pharmacology*, *158*(5), 1196–1209. https://doi.org/10.1111/j.1476-5381.2009.00387.x

Cho, M. E., & Kopp, J. B. (2010). Pirfenidone: An anti-fibrotic therapy for progressive kidney disease. *Expert Opinion on Investigational Drugs*, *19*(2), 275–283. https://doi.org/10.1517/13543780903501539

Choi, S., Nam, J., Kim, B., Jang, J., Jin, Y., Lee, H., Park, S., Ji, Y. H., Cho, J., & Lee, Y. (2016). HSPB1 Inhibits the Endothelial-to-Mesenchymal Transition to Suppress Pulmonary Fibrosis and Lung Tumorigenesis. *Cancer Research*, *1*(28), 1019–1030. https://doi.org/10.1158/0008-5472.CAN-15-0952

Coentro, J. Q., May, U., Prince, S., Zwaagstra, J., Ritvos, O., Järvinen, T. A. H., & Zeugolis, D. I. (2021). *Adapting the Scar-in-a-Jar to Skin Fibrosis and Screening Traditional and Contemporary Anti-Fibrotic Therapies*. *9*(October), 1–18. https://doi.org/10.3389/fbioe.2021.756399

Coleman, W. B., Tsongalis, G. J., Willis, M. S., Yates, C. C., & Schisler, J. C. (2019). *Fibrosis in Disease*. https://www.springer.com/series/8176

Collins, B. F., & Raghu, G. (2019). *Antifibrotic therapy for fibrotic lung disease beyond idiopathic pulmonary fibrosis*. *153*. https://doi.org/10.1183/16000617.0022-2019

Cui, Y., Zhang, M., Leng, C., Blokzijl, T., Jansen, B. H., Dijkstra, G., & Faber, K. N. (2020). Pirfenidone Inhibits Cell Proliferation and Collagen I Production of Primary Human Intestinal Fibroblasts. *Cells*, *9*(3), 775.

D'Souza, R. C. J., Knittle, A. M., Nagaraj, N., Van Dinther, M., Choudhary, C., Ten Dijke, P., Mann, M., & Sharma, K. (2014). Time-resolved dissection of early phosphoproteome and ensuing proteome changes in response to TGF-β. *Science Signaling*, *7*(335), 1–17. https://doi.org/10.1126/scisignal.2004856

de Boer, I. H., Alpers, C. E., Azeloglu, E. U., Balis, U. G. J., Barasch, J. M., Barisoni, L., Blank, K. N., Bomback, A. S., Brown, K., Dagher, P. C., Dighe, A. L., Eadon, M. T., El-Achkar, T. M., Gaut, J. P., Hacohen, N., He, Y., Hodgin, J. B., Jain, S., Kellum, J. A., … Alfano, S. (2021). Rationale and design of the Kidney Precision Medicine Project. *Kidney International*, *99*(3), 498–510. https://doi.org/10.1016/j.kint.2020.08.039

Denechaud, P. D., Fajas, L., & Giralt, A. (2017). E2F1, a novel regulator of metabolism. *Frontiers in Endocrinology*, *8*(NOV), 1–8. https://doi.org/10.3389/fendo.2017.00311

Devarajan, P. (2010). The Use of Targeted Biomarkers for Chronic Kidney Disease. *Advances in Chronic Kidney Disease*, *23*(1), 1–7. https://doi.org/10.1161/CIRCULATIONAHA.110.956839

Devos, H., Zoidakis, J., Roubelakis, M. G., Latosinska, A., & Vlahou, A. (2023). Reviewing the Regulators of COL1A1. *International Journal of Molecular Sciences*, *24*(12), 10004. https://doi.org/10.3390/ijms241210004

Didiasova, M., Singh, R., Wilhelm, J., Kwapiszewska, G., & Wujak, L. (2015). *Pirfenidone exerts antifibrotic effects through inhibition of GLI transcription factors*. *8*, 1916–1928. https://doi.org/10.1096/fj.201600892RR

Dom, Z. I. M., Satake, E., Skupien, J., Krolewski, B., O'Neil, K., Willency, J. A., Dillon, S. T., Wilson, J. M., Kobayashi, H., Ihara, K., Libermann, T. A., Pragnell, M., Duffin, K. L., & Krolewski, A. S. (2021). Circulating proteins protect against renal

decline and progression to end-stage renal disease in patients with diabetes. *Science Translational Medicine*, *13*(600), 1–13. https://doi.org/10.1126/scitranslmed.abd2699

Dubin, R. F., & Rhee, E. P. (2020). Proteomics and metabolomics in kidney disease, including insights into etiology, treatment, and prevention. *Clinical Journal of the American Society of Nephrology*, *15*(3), 404–411. https://doi.org/10.2215/CJN.07420619

Dugourd, A., Kuppe, C., Sciacovelli, M., Gjerga, E., Gabor, A., Emdal, K. B., Vieira, V., Bekker-Jensen, D. B., Kranz, J., Bindels, E. M. J., Costa, A. S. H., Sousa, A., Beltrao, P., Rocha, M., Olsen, J. V., Frezza, C., Kramann, R., & Saez-Rodriguez, J. (2021). Causal integration of multi-omics data with prior knowledge to generate mechanistic hypotheses. *Molecular Systems Biology*, *17*(1), 1–17. https://doi.org/10.15252/msb.20209730

Dugourd, A., & Saez-Rodriguez, J. (2019). Footprint-based functional analysis of multiomic data. *Current Opinion in Systems Biology*, *15*, 82–90. https://doi.org/10.1016/j.coisb.2019.04.002

Eaton, D. C. (2012). Frontiers in renal and epithelial physiology - grand challenges. *Frontiers in Physiology*, *3 JAN*(January), 2011–2013. https://doi.org/10.3389/fphys.2012.00002

Eddy, A. A. (2014). Overview of the cellular and molecular basis of kidney fibrosis. *Kidney International Supplements*, *4*(1), 2–8. https://doi.org/10.1038/kisup.2014.2

Edeling, M., Ragi, G., Huang, S., Pavenstädt, H., & Susztak, K. (2016). Developmental signalling pathways in renal fibrosis: the roles of Notch, Wnt and Hedgehog. *Nature Reviews Nephrology*, *118*(24), 6072–6078. https://doi.org/10.1002/cncr.27633.Percutaneous

El-Achkar, T. M., Eadon, M. T., Kretzler, M., Himmelfarb, J., Lake, B., Zhang, K., Lecker, S., Morales, A., Bogen, S., Amodu, A. A., Beck, L., Henderson, J., Ilori, T., Maikhor, S., Onul, I., Schmidt, I., Verma, A., Waikar, S., Yadati, P., … Wilson, F. P. (2024). Precision Medicine in Nephrology: An Integrative Framework of Multidimensional Data in the Kidney Precision Medicine Project. *American Journal of Kidney Diseases*, *83*(3), 402–410. https://doi.org/10.1053/j.ajkd.2023.08.015

Enomoto, Y., Matsushima, S., Shibata, K., Aoshima, Y., Yagi, H., Meguro, S., Kawasaki, H., Kosugi, I., Fujisawa, T., Enomoto, N., Inui, N., Nakamura, Y., Suda, T., & Iwashita, T. (2018). LTBP2 is secreted from lung myofibroblasts and is a potential biomarker for idiopathic pulmonary fibrosis. *Clinical Science*, *132*(14), 1565–1580. https://doi.org/10.1042/CS20180435

Fan, J., Tong, G., Chen, X., Li, S., Yu, Y., Zhu, S., Zhu, K., Hu, Z., Dong, Y., Chen, R., Zhu, J., Gong, W., Hu, Z., Zhou, B., Chen, Y., Jin, L., & Cong, W. (2023). CK2 blockade alleviates liver fibrosis by suppressing activation of hepatic stellate cells via the Hedgehog pathway. *British Journal of Pharmacology*, *180*(1), 44–61. https://doi.org/10.1111/bph.15945

François, H., & Chatziantoniou, C. (2018). Renal fibrosis: Recent translational aspects. In *Matrix Biology* (Vols. 68–69, pp. 318–332). Elsevier B.V.

https://doi.org/10.1016/j.matbio.2017.12.013

Friedman, S. L., Sheppard, D., Duffield, J. S., & Violette, S. (2013). Therapy for fibrotic diseases: Nearing the starting line. *Science Translational Medicine*, *5*(167). https://doi.org/10.1126/scitranslmed.3004700

Fu, Y., Su, L., Cai, M., Yao, B., Xiao, S., He, Q., Xu, L., Yang, L., Zhao, C., Wan, T., Shao, L., Wang, L., & Huang, X. (2019). Downregulation of CPA4 inhibits non small–cell lung cancer growth by suppressing the AKT/c-MYC pathway. *Molecular Carcinogenesis*, *58*(11), 2026–2039. https://doi.org/10.1002/mc.23095

Gao, W., Guo, N., Zhao, S., Chen, Z., Zhang, W., Yan, F., Liao, H., & Chi, K. (2020). Carboxypeptidase A4 promotes cardiomyocyte hypertrophy through activating PI3K-AKT-mTOR signaling. *Bioscience Reports*, *40*(5), 1–10. https://doi.org/10.1042/BSR20200669

Garcia-Alonso, L., Holland, C. H., Ibrahim, M. M., Turei, D., & Saez-Rodriguez, J. (2019). Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Research*, *29*(8), 1363–1375. https://doi.org/10.1101/gr.240663.118

Garrido-Rodriguez, M., Zirngibl, K., Ivanova, O., Lobentanzer, S., & Saez-Rodriguez, J. (2022). Integrating knowledge and omics to decipher mechanisms via large-scale models of signaling networks. *Molecular Systems Biology*, *18*(7), e11036. https://doi.org/10.15252/msb.202211036

Genovese, F., Manresa, A. A., Leeming, D. J., Karsdal, M. A., & Boor, P. (2014). The extracellular matrix in the kidney: A source of novel non-invasive biomarkers of kidney fibrosis? *Fibrogenesis and Tissue Repair*, *7*(1), 1–14. https://doi.org/10.1186/1755-1536-7-4

Gewin, L., Zent, R., & Pozzi, A. (2017). Progression of chronic kidney disease: too much cellular talk causes damage. *Kidney International*, *91*(3), 552–560. https://doi.org/10.1016/j.kint.2016.08.025

Ghosh, A. K., & Vaughan, D. E. (2012). PAI-1 in tissue fibrosis. *Journal of Cellular Physiology*, *227*(2), 493–507. https://doi.org/10.1002/jcp.22783

Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., Griss, J., Sevilla, C., Matthews, L., Gong, C., Deng, C., Varusai, T., Ragueneau, E., Haider, Y., May, B., Shamovsky, V., Weiser, J., Brunson, T., Sanati, N., … D'Eustachio, P. (2022). The Reactome Pathway Knowledgebase 2022. *Nucleic Acids Research*, *50*(D1), D672–D678. https://doi.org/10.1093/nar/gkad1025

Granata, S., Dalla Gassa, A., Bellin, G., Lupo, A., & Zaza, G. (2016). Transcriptomics: A step behind the comprehension of the polygenic influence on oxidative stress, immune deregulation, and mitochondrial dysfunction in chronic kidney disease. *BioMed Research International*, *2016*. https://doi.org/10.1155/2016/9290857

Henderson, N. C., Rieder, F., & Wynn, T. A. (2020). Fibrosis: from mechanisms to medicines. *Nature*, *587*(7835), 555–566. https://doi.org/10.1038/s41586-020-2938-9

Higgins, C. E., Tang, J., Higgins, S. P., Gifford, C. C., Mian, B. M., Jones, D. M., Zhang, W., Costello, A., Conti, D. J., Samarakoon, R., & Higgins, P. J. (2021).

The Genomic Response to TGF-β1 Dictates Failed Repair and Progression of Fibrotic Disease in the Obstructed Kidney. *Frontiers in Cell and Developmental Biology*, *9*(July), 1–24. https://doi.org/10.3389/fcell.2021.678524

Higgins, C. E., Tang, J., Mian, B. M., Higgins, S. P., Gifford, C. C., Conti, D. J., Meldrum, K. K., Samarakoon, R., & Higgins, P. J. (2019). TGF-β1–p53 cooperativity regulates a profibrotic genomic program in the kidney: molecular mechanisms and clinical implications. *FASEB Journal*, *33*(10), 10596–10606. https://doi.org/10.1096/fj.201900943R

Higgins, S. P., Tang, Y., Higgins, C. E., Mian, B., Zhang, W., Czekay, R. P., Samarakoon, R., Conti, D. J., & Higgins, P. J. (2018). TGF-β1/p53 signaling in renal fibrogenesis. *Cellular Signalling*, *43*(October 2017), 1–10. https://doi.org/10.1016/j.cellsig.2017.11.005

Himmelfarb, J., & Ikizler, T. A. (2019). *Chronic Kidney Disease, Dialysis, and Transplantation*. https://doi.org/10.1016/B978-0-323-52978-5.18001-2

Hinz, B., & Lagares, D. (2020). Evasion of apoptosis by myofibroblasts: a hallmark of fibrotic diseases. *Nature Reviews Rheumatology*, *16*(1), 11–31. https://doi.org/10.1038/s41584-019-0324-5

Hinz, B., & Lagares, D. (2021). *Myofibroblasts*. https://doi.org/10.1007/978-1-0716-1382-5

Hishikawa, A., Hayashi, K., & Itoh, H. (2018). Transcription factors as therapeutic targets in chronic kidney disease. *Molecules*, *23*(5). https://doi.org/10.3390/molecules23051123

Hoffman, L. M., Jensen, C. C., & Beckerle, M. C. (2022). Phosphorylation of the small heat shock protein HspB1 regulates cytoskeletal recruitment and cell motility. *Molecular Biology of the Cell*, 1–14. https://doi.org/10.1091/mbc.E22-02-0057

Hsieha, L. T.-H., Nastasea, M.-V., Zeng-Brouwersa, J., V. Iozzob, R., & Schaefer, L. (2014). Soluble biglycan as a biomarker of inflammatory renal diseases Louise. *International Journal of Biochemistry and Cell Biology*, *23*(1), 1–7. https://doi.org/10.1016/j.biocel.2014.07.020

Huang, C., & Lo, S. H. (2023). Tensins in Kidney Function and Diseases. *Life*, *13*(1244), 1–13.

Huang, R., Fu, P., & Ma, L. (2023). Kidney fibrosis: from mechanisms to therapeutic medicines. *Signal Transduction and Targeted Therapy*, *8*(129). https://doi.org/10.1038/s41392-023-01379-7

Huber, W., von Heydebreck, A., Sültmann, H., Poustka, A., & Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, *18*(suppl_1), S96–S104. https://doi.org/10.1093/bioinformatics/18.suppl_1.S96

Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M., & Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead technology. *Molecular Systems Biology*, *10*(10), 1–14. https://doi.org/10.15252/msb.20145625

Hughes, C. S., Moggridge, S., Müller, T., Sorensen, P. H., Morin, G. B., & Krijgsveld, J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics

experiments. *Nature Protocols*, *14*(1), 68–85. https://doi.org/10.1038/s41596-018-0082-x

Inagaki, Y., Nemoto, T., Kushida, Z. M., Sheng, Y., Higa, K., Shirasaki, F., Kota, S., Fujii, M., Yamauchi, H., Nakao, A., Crombrugghe, B. De, Watanabe, T., & Okazaki, I. (2003). Interferon Alfa Down-regulates Collagen Gene Transcription and Suppresses Experimental Hepatic Fibrosis in Mice. *Hepatology*, *38*(4), 890–899.

Jin, B., Zhu, J., Zhou, Y., Liang, L., Yang, Y., Xu, L., Zhang, T., Li, P., Pan, T., Guo, B., Chen, T., & Li, H. (2022). Loss of MEN1 leads to renal fibrosis and decreases HGF-Adamts5 pathway activity via an epigenetic mechanism. *Clinical and Translational Medicine*, *12*(8). https://doi.org/10.1002/ctm2.982

Kanbay, A., Ecder Alişir, S., Yakar, H. I., Yildiz Sahin, B., & Ceylan, E. (2019). Acute Renal Failure Due To Pyrfenidone: a Case Report. *Chest*, *155*(4), 5A. https://doi.org/10.1016/j.chest.2019.02.025

Karsdal, M. A., Nielsen, S. H., Leeming, D. J., Langholm, L. L., Nielsen, M. J., Manon-Jensen, T., Siebuhr, A., Gudmann, N. S., Rønnow, S., Sand, J. M., Daniels, S. J., Mortensen, J. H., & Schuppan, D. (2017). The good and the bad collagens of fibrosis – Their role in signaling and organ function. *Advanced Drug Delivery Reviews*, *121*, 43–56. https://doi.org/10.1016/j.addr.2017.07.014

Ke, B., Fan, C., Yang, L., & Fang, X. (2017). *Matrix Metalloproteinases-7 and Kidney Fibrosis*. *8*(February), 1–7. https://doi.org/10.3389/fphys.2017.00021

Kendall, R. T., & Feghali-bostwick, C. A. (2014). Fibroblasts in fibrosis: novel roles and mediators. *Frontiers in Pharmacology*, *5*(May), 1–13. https://doi.org/10.3389/fphar.2014.00123

Khan, M. M., Zukowska, J., Galea, G., Tuechler, N., Halavatyi, A., Tischer, C., Haberkant, P., Stein, F., Jung, F., Landry, J., Khan, A. M., Oorschot, V., Becher, I., Neumann, B., Muley, T., Winter, H., Duerr, J., Mall, M. A., Savitski, M., & Pepperkok, R. (2023). Dextromethorphan inhibits collagen transport in the endoplasmic reticulum eliciting an anti-fibrotic response in vitro models of pulmonary fibrosis. *BioRxiv*.

Kim, K. P., Williams, C. E., & Lemmon, C. A. (2022). Cell-Matrix Interactions in Renal Fibrosis. *Kidney and Dialysis*, *2*(4), 607–624. https://doi.org/10.3390/kidneydial2040055

Kiss, Z., Mudryj, M., & Ghosh, P. M. (2020). Non-circadian aspects of BHLHE40 cellular function in cancer. *Genes and Cancer*, *11*(1–2), 1–18. https://doi.org/10.18632/genesandcancer.201

Klinkhammer, B. M., & Boor, P. (2023). Kidney fibrosis: Emerging diagnostic and therapeutic strategies. In *Molecular Aspects of Medicine* (Vol. 93). Elsevier Ltd. https://doi.org/10.1016/j.mam.2023.101206

Klinkhammer, B. M., Goldschmeding, R., Floege, J., & Boor, P. (2017). Treatment of Renal Fibrosis—Turning Challenges into Opportunities. *Advances in Chronic Kidney Disease*, *24*(2), 117–129. https://doi.org/10.1053/j.ackd.2016.11.002

Kobayashi, H., Looker, H. C., Satake, E., Saulnier, P. J., Md Dom, Z. I., O'Neil, K., Ihara, K., Krolewski, B., Galecki, A. T., Niewczas, M. A., Wilson, J. M., Doria, A.,

Duffin, K. L., Nelson, R. G., & Krolewski, A. S. (2022). Results of untargeted analysis using the SOMAscan proteomics platform indicates novel associations of circulating proteins with risk of progression to kidney failure in diabetes. *Kidney International*, *102*(2), 370–381. https://doi.org/10.1016/j.kint.2022.04.022

Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., & Nesvizhskii, A. I. (2017). MSFragger: Ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nature Methods*, *14*(5), 513–520. https://doi.org/10.1038/nmeth.4256

Krahn, K. N., Bouten, C. V. C., Van Tuijl, S., Van Zandvoort, M. A. M. J., & Merkx, M. (2006). Fluorescently labeled collagen binding proteins allow specific visualization of collagen in tissues and live cell culture. *Analytical Biochemistry*, *350*(2), 177–185. https://doi.org/10.1016/j.ab.2006.01.013

Kramann, R. (2016). Hedgehog Gli signalling in kidney fibrosis. *Nephrology Dialysis Transplantation*, *31*(12), 1989–1995. https://doi.org/10.1093/ndt/gfw102

Kramann, R., Dirocco, D. P., & Humphreys, B. D. (2013). Understanding the origin, activation and regulation of matrix-producing myofibroblasts for treatment of fibrotic disease. *Journal of Pathology*, *231*(3), 273–289. https://doi.org/10.1002/path.4253

Kramann, R., & Humphreys, B. D. (2014). Kidney Pericytes: Roles in Regeneration and Fibrosis. *Seminars in Nephrology*, *34*(4), 374–383. https://doi.org/10.1038/jid.2014.371

Kramann, R., Machado, F., Wu, H., Kusaba, T., Hoeft, K., Schneider, R. K., & Humphreys, B. D. (2018). Parabiosis and single-cell RNA sequencing reveal a limited contribution of monocytes to myofibroblasts in kidney fibrosis. *JCI Insight*, *3*(9). https://doi.org/10.1172/jci.insight.99561

Kramann, R., Schneider, R. K., Dirocco, D. P., Machado, F., Fleig, S., Bondzie, P. A., Henderson, J. M., Ebert, B. L., & Humphreys, B. D. (2015). Perivascular Gli1+ progenitors are key contributors to injury- induced organ fibrosis. *Cell Stem Cell*, *16*(1), 51–66. https://doi.org/10.1016/j.stem.2014.11.004.Perivascular

Krolewski, A., Skupien, J., Rossing, P., & Warram, J. (2017). Fast renal decline to ESRD: an unrecognized feature of nephropathy in diabetes. *Kidney International*, *91*(6), 29–39. https://doi.org/10.1016/j.kint.2016.10.046.Fast

Kuppe, C., Ibrahim, M. M., Kranz, J., Zhang, X., Smith, J. R., Dobie, R., Wilson-kanamori, J. R., Halder, M., & Xu, Y. (2021). Decoding myofibroblast origins in human kidney fibrosis. *Nature*, *589*, 281–286. https://doi.org/10.1038/s41586-020-2941-1

Kusnadi, E. P., Timpone, C., Topisirovic, I., Larsson, O., & Furic, L. (2022). Regulation of gene expression via translational buffering. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1869*(1), 119140. https://doi.org/10.1016/j.bbamcr.2021.119140

Lagares, D., & Kapoor, M. (2013). Targeting focal adhesion kinase in fibrotic diseases. *BioDrugs*, *27*(1), 15–23. https://doi.org/10.1007/s40259-012-0003-4

Lake, B. B., Menon, R., Winfree, S., Hu, Q., Ferreira, R. M., Kalhor, K., Barwinska, D., Otto, E. A., Ferkowicz, M., Diep, D., Plongthongkum, N., Knoten, A., Urata,

S., Naik, A. S., Eddy, S., Zhang, B., Wu, Y., Salamon, D., Williams, J. C., … consortium, for the K. (2021). An atlas of healthy and injured cell states and niches in the human kidney. *BioRxiv*, *619*(July), 2021.07.28.454201. https://doi.org/10.1038/s41586-023-05769-3

Lassé, M., El Saghir, J., Berthier, C. C., Eddy, S., Fischer, M., Laufer, S. D., Kylies, D., Hutzfeldt, A., Bonin, L. L., Dumoulin, B., Menon, R., Vega-Warner, V., Eichinger, F., Alakwaa, F., Fermin, D., Billing, A. M., Minakawa, A., McCown, P. J., Rose, M. P., … Rinschen, M. M. (2023). An integrated organoid omics map extends modeling potential of kidney disease. *Nature Communications*, *14*(1), 4903. https://doi.org/10.1038/s41467-023-39740-7

Lehtonen, S. T., Veijola, A., Karvonen, H., Lappi-blanco, E., & Sormunen, R. (2016). Pirfenidone and nintedanib modulate properties of fibroblasts and myofibroblasts in idiopathic pulmonary fibrosis. *Respiratory Research*, 1–12. https://doi.org/10.1186/s12931-016-0328-5

Levitt, J. M., Baldwin, A., Papadakis, A., Puri, S., Xylas, J., Münger, K., & Georgakoudi, I. (2006). Intrinsic fluorescence and redox changes associated with apoptosis of primary human epithelial cells. *Journal of Biomedical Optics*, *11*(6), 064012. https://doi.org/10.1117/1.2401149

Li, H., Peng, X., Wang, Y., Cao, S., Xiong, L., Fan, J., Wang, Y., & Zhuang, S. (2016). Atg5-mediated autophagy deficiency in proximal tubules promotes cell cycle G2/M arrest and renal fibrosis. *Autophagy*, *12*(9), 1472–1486. https://doi.org/10.1080/15548627.2016.1190071

Liao, R., Qi, Z., Tang, R., Wang, R., & Wang, Y. (2021). Methyl Ferulic Acid Attenuates Human Cardiac Fibroblasts Differentiation and Myocardial Fibrosis by Suppressing pRB-E2F1/CCNE2 and RhoA/ROCK2 Pathway. *Frontiers in Pharmacology*, *12*(August), 1–12. https://doi.org/10.3389/fphar.2021.714390

Liao, R., Xie, B., Cui, J., Qi, Z., Xue, S., & Wang, Y. (2021). E2F transcription factor 1 (E2F1) promotes the transforming growth factor TGF-β1 induced human cardiac fibroblasts differentiation through promoting the transcription of CCNE2 gene. *Bioengineered*, *12*(1), 6869–6877. https://doi.org/10.1080/21655979.2021.1972194

Lindenmeyer, M. T., Alakwaa, F., Rose, M., & Kretzler, M. (2021). Perspectives in systems nephrology. *Cell and Tissue Research*, *385*(2), 475–488. https://doi.org/10.1007/s00441-021-03470-3

Liu, A., Trairatphisan, P., Gjerga, E., Didangelos, A., Barratt, J., & Saez-Rodriguez, J. (2019). From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL. *NPJ Systems Biology and Applications*, *5*(1), 1–10. https://doi.org/10.1038/s41540-019-0118-z

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*, *165*(3), 535–550. https://doi.org/10.1016/j.cell.2016.03.014

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Lotz-Jenne, C., Lüthi, U., Ackerknecht, S., Lehembre, F., Fink, T., Stritt, M., Wirth,

M., Pavan, S., Bill, R., Regenass, U., Christofori, G., & Meyer-Schaller, N. (2016). A high-content EMT screen identifies multiple receptor tyrosine kinase inhibitors with activity on TGFβ receptor. *Oncotarget*, *7*(18), 25983–26002. https://doi.org/10.18632/oncotarget.8418

Lousa, I., Reis, F., Beirão, I., Alves, R., Belo, L., & Santos-Silva, A. (2021). New potential biomarkers for chronic kidney disease management — A review of the literature. *International Journal of Molecular Sciences*, *22*(1), 1–37. https://doi.org/10.3390/ijms22010043

Luangmonkong, T., Suriguga, S., Adhyatmika, A., Adlia, A., Oosterhuis, D., Suthisisang, C., de Jong, K. P., Mutsaers, H. A. M., & Olinga, P. (2018). In vitro and ex vivo anti-fibrotic effects of LY2109761, a small molecule inhibitor against TGF-β. *Toxicology and Applied Pharmacology*, *355*, 127–137. https://doi.org/10.1016/j.taap.2018.07.001

Mariani, L. H., Eddy, S., AlAkwaa, F. M., McCown, P. J., Harder, J. L., Nair, V., Eichinger, F., Martini, S., Ademola, A. D., Boima, V., Reich, H. N., El Saghir, J., Godfrey, B., Ju, W., Tanner, E. C., Vega-Warner, V., Wys, N. L., Adler, S. G., Appel, G. B., … Kretzler, M. (2023). Precision nephrology identified tumor necrosis factor activation variability in minimal change disease and focal segmental glomerulosclerosis. *Kidney International*, *103*(3), 565–579. https://doi.org/10.1016/j.kint.2022.10.023

Mariani, L. H., Martini, S., Barisoni, L., Canetta, P. A., Troost, J. P., Hodgin, J. B., Palmer, M., Rosenberg, A. Z., Lemley, K. V., Chien, H. P., Zee, J., Smith, A., Appel, G. B., Trachtman, H., Hewitt, S. M., Kretzler, M., & Bagnasco, S. M. (2018). Interstitial fibrosis scored on whole-slide digital imaging of kidney biopsies is a predictor of outcome in proteinuric glomerulopathies. *Nephrology Dialysis Transplantation*, *33*(2), 310–318. https://doi.org/10.1093/ndt/gfw443

Massagué, J. (2012). TGFβ signalling in context. *Nature Reviews Molecular Cell Biology*, *13*(10), 616–630. https://doi.org/10.1038/nrm3434

Matsubara, T., Araki, M., Abe, H., Ueda, O., Jishage, K. I., Mima, A., Goto, C., Tominaga, T., Kinosaki, M., Kishi, S., Nagai, K., Iehara, N., Fukushima, N., Kita, T., Arai, H., & Doi, T. (2015). Bone morphogenetic protein 4 and Smad1 mediate extracellular matrix production in the development of Diabetic nephropathy. *Diabetes*, *64*(8), 2978–2990. https://doi.org/10.2337/db14-0893

Melisi, D., Ishiyama, S., Sclabas, G. M., Fleming, J. B., Xia, Q., Tortora, G., Abbruzzese, J. L., & Chiao, P. J. (2008). LY2109761, a novel transforming growth factor β receptor type I and type II dual inhibitor, as a therapeutic approach to suppressing pancreatic cancer metastasis. *Molecular Cancer Therapeutics*, *7*(4), 829–840. https://doi.org/10.1158/1535-7163.MCT-07-0337

Meng, X. M., Nikolic-Paterson, D. J., & Lan, H. Y. (2016). TGF-β: The master regulator of fibrosis. *Nature Reviews Nephrology*, *12*(6), 325–338. https://doi.org/10.1038/nrneph.2016.48

Meran, S., & Steadman, R. (2011). Fibroblasts and myofibroblasts in renal fibrosis. *International Journal of Experimental Pathology*, *92*(3), 158–167. https://doi.org/10.1111/j.1365-2613.2011.00764.x

Mihai, S., Codrici, E., Popescu, I. D., Enciu, A. M., Albulescu, L., Necula, L. G.,

Mambet, C., Anton, G., & Tanase, C. (2018). Inflammation-related mechanisms in chronic kidney disease prediction, progression, and outcome. *Journal of Immunology Research*, *2018*. https://doi.org/10.1155/2018/2180373

Mikhailova, E. V., Romanova, I. V., Bagrov, A. Y., & Agalakova, N. I. (2023). Fli1 and Tissue Fibrosis in Various Diseases. *International Journal of Molecular Sciences*, *24*(3), 1–20. https://doi.org/10.3390/ijms24031881

Moita, M. R., Silva, M. M., Diniz, C., Serra, M., Hoet, R. M., Barbas, A., & Simão, D. (2022). Transcriptome and proteome profiling of activated cardiac fibroblasts supports target prioritization in cardiac fibrosis. *Frontiers in Cardiovascular Medicine*, *9*(December), 1–17. https://doi.org/10.3389/fcvm.2022.1015473

Moriya, K., Bae, E., Honda, K., Sakai, K., Sakaguchi, T., Tsujimoto, I., Kamisoyama, H., Keene, D. R., Sasaki, T., & Sakai, T. (2011). A fibronectin-independent mechanism of collagen fibrillogenesis in adult liver remodeling. *Gastroenterology*, *140*(5), 1653–1663. https://doi.org/10.1053/j.gastro.2011.02.005.A

Mullenbrock, S., Liu, F., Szak, S., Hronowski, X., Gao, B., Juhasz, P., Sun, C., Liu, M., McLaughlin, H., Xiao, Q., Feghali-Bostwick, C., & Zheng, T. S. (2018). Systems analysis of transcriptomic and proteomic profiles identifies novel regulation of fibrotic programs by miRNAs in pulmonary fibrosis fibroblasts. *Genes*, *9*(12). https://doi.org/10.3390/genes9120588

Mullins, L. J., Conway, B. R., Menzies, R. I., Denby, L., & Mullins, J. J. (2016). Renal disease pathophysiology and treatment: Contributions from the rat. *DMM Disease Models and Mechanisms*, *9*(12), 1419–1433. https://doi.org/10.1242/dmm.027276

Naba, A., Clauser, K. R., Ding, H., Whittaker, C. A., Carr, S. A., & Hynes, R. O. (2016). The extracellular matrix: Tools and insights for the “omics” era. *Matrix Biology*, *49*, 10–24. https://doi.org/10.1016/j.matbio.2015.06.003

Naba, A., Clauser, K. R., & Hynes, R. O. (2015). *Enrichment of Extracellular Matrix Proteins from Tissues and Digestion into Peptides for Mass Spectrometry Analysis*. *July*, 1–9. https://doi.org/10.3791/53057

Nakamura, Y., Kita, S., Tanaka, Y., Fukuda, S., Obata, Y., Okita, T., Kawachi, Y., Tsugawa-Shimizu, Y., Fujishima, Y., Nishizawa, H., Miyagawa, S., Sawa, Y., Sehara-Fujisawa, A., Maeda, N., & Shimomura, I. (2020). A disintegrin and metalloproteinase 12 prevents heart failure by regulating cardiac hypertrophy and fibrosis. *American Journal of Physiology. Heart and Circulatory Physiology*, *318*(2), H238–H251. https://doi.org/10.1152/ajpheart.00496.2019

Nakayama, S., Mukae, H., Sakamoto, N., & Kakugawa, T. (2008). Pirfenidone inhibits the expression of HSP47 in TGF-β1-stimulated human lung fibroblasts. *Life Sci*, *82*, 210–217. https://doi.org/10.1016/j.lfs.2007.11.003

Nakerakanti, S. S., Bujor, A. M., & Trojanowska, M. (2011). CCN2 is required for the TGF-β induced activation of Smad1 - Erk1/2 signaling network. *PLoS ONE*, *6*(7). https://doi.org/10.1371/journal.pone.0021911

Needham, E. J., Parker, B. L., Burykin, T., James, D. E., & Humphrey, S. J. (2019). Illuminating the dark phosphoproteome. *Science Signaling*, *12*(565), 1–18. https://doi.org/10.1126/scisignal.aau8645

Niewczas, M. A., Pavkov, M. E., Skupien, J., Smiles, A., Zaipul, I., Wilson, J. M., Park, J., Nair, V., Schlafly, A., Satake, E., Simeone, C. A., Shah, H., Qiu, C., Looker, H. C., Fiorina, P., Ware, C. F., Sun, J. K., Doria, A., Kretzler, M., … Andrzej, S. (2019). A Signature of Circulating Inflammatory Proteins and Development of End Stage Renal Disease in Diabetes. *Nature Medicine*, *25*(5), 805–813. https://doi.org/10.1038/s41591-019-0415-5.A

Nitta, K. (2015). *Chronic Kidney Disease - Recent Advances in Clinical and Basic Research*.

Nolte, M. A., Hoen, E. N. M. N.-, & Margadant, C. (2020). Integrins Control Vesicular Trafficking ; New Tricks for Old Dogs. *Trends in Biochemical Sciences*, 1–14. https://doi.org/10.1016/j.tibs.2020.09.001

Oldroyd, S. D., Thomas, G. L., Gabbiani, G., & Meguid El Nahas, A. (1999). Interferon-γ inhibits experimental renal fibrosis. *Kidney International*, *56*(6), 2116–2127. https://doi.org/10.1046/j.1523-1755.1999.00775.x

Pandey, A. K., & Loscalzo, J. (2023). Network medicine: an approach to complex kidney disease phenotypes. *Nature Reviews Nephrology*, *19*(July), 463–475. https://doi.org/10.1038/s41581-023-00705-0

Pannu, J., Nakerakanti, S., Smith, E., Ten Dijke, P., & Trojanowska, M. (2007). Transforming growth factor-β receptor type I-dependent fibrogenic gene program is mediated via activation of Smad1 and ERK1/2 pathways. *Journal of Biological Chemistry*, *282*(14), 10405–10413. https://doi.org/10.1074/jbc.M611742200

Park, C. H., & Yoo, T. H. (2022). TGF-β Inhibitors for Therapeutic Management of Kidney Fibrosis. *Pharmaceuticals*, *15*(12). https://doi.org/10.3390/ph15121485

Park, J., Shrestha, R., Qiu, C., Kondo, A., & Huang, S. (2018). *Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease*. *360*(6390), 758–763. https://doi.org/10.1126/science.aar2131.Single-cell

Piepho, H. P. (2009). Data transformation in statistical analysis of field trials with changing treatment variance. *Agronomy Journal*, *101*(4), 865–869. https://doi.org/10.2134/agronj2008.0226x

Piepho, H. P., Büchse, A., & Emrich, K. (2003). A Hitchhiker's Guide to Mixed Models for Randomized Experiments. *Journal of Agronomy and Crop Science*, *189*(5), 310–322. https://doi.org/10.1046/j.1439-037X.2003.00049.x

Piepho, H. P., Büchse, A., & Richter, C. (2004). A mixed modelling approach for randomized experiments with repeated measures. *Journal of Agronomy and Crop Science*, *190*(4), 230–247. https://doi.org/10.1111/j.1439-037X.2004.00097.x

Piersma, B., Bank, R. A., & Boersema, M. (2015). Signaling in fibrosis: TGF-β, WNT, and YAP/TAZ converge. *Frontiers in Medicine*, *2*(SEP), 1–14. https://doi.org/10.3389/fmed.2015.00059

Potel, C. M., Lin, M. H., Heck, A. J. R., & Lemeer, S. (2018). Defeating major contaminants in Fe3- immobilized metal ion affinity chromatography (IMAC) phosphopeptide enrichment. *Molecular and Cellular Proteomics*, *17*(5), 1028–1034. https://doi.org/10.1074/mcp.TIR117.000518

Prakash, J., & Pinzani, M. (2017). Fibroblasts and extracellular matrix: Targeting and therapeutic tools in fibrosis and cancer. *Advanced Drug Delivery Reviews*, *121*, 1–2. https://doi.org/10.1016/j.addr.2017.11.008

Rasmussen, D. G. K., Boesby, L., Nielsen, S. H., Tepel, M., Birot, S., Karsdal, M. A., Kamper, A. L., & Genovese, F. (2019). Collagen turnover profiles in chronic kidney disease. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-51905-3

Reichel, M., Liao, Y., Rettel, M., Ragan, C., Evers, M., Alleaume, A. M., Horos, R., Hentze, M. W., Preiss, T., & Millar, A. A. (2016). In planta determination of the mRNA-binding proteome of arabidopsis etiolated seedlings. *Plant Cell*, *28*(10), 2435–2452. https://doi.org/10.1105/tpc.16.00562

Reznichenko, A., Nair, V., Eddy, S., Tomilo, M., Slidel, T., Ju, W., Conway, P., Badal, S. S., Wesley, J., Liles, J. T., Moosmang, S., Julie, M., Quinn, C. M., Bitzer, M., Karihaloo, A., Breyer, M. D., & Kevin, L. (2021). Molecular Stratification of Chronic Kidney Disease. *MedRxiv*, 734–763. https://www.medrxiv.org/content/10.1101/2021.09.09.21263234v1%0Ahttps://www.medrxiv.org/content/10.1101/2021.09.09.21263234v1.abstract

Ricard-Blum, S. (2011). The Collagen Family. *Cold Spring Harbour Perspectives in Biology*, *3*, 1–20.

Rinschen, M. M., & Saez-Rodriguez, J. (2021). The tissue proteome in the multi-omic landscape of kidney disease. *Nature Reviews Nephrology*. https://doi.org/10.1038/s41581-020-00348-5

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Rockey, D. C., Bell, P. D., & Hill, J. A. (2015). Fibrosis-a common pathway to organ injury and failure. *New England Journal of Medicine*, *372*(12), 1138–1149. https://doi.org/10.1056/NEJMra1300575

Rodríguez-Ortiz, M. E., Pontillo, C., Rodríguez, M., Zürbig, P., Mischak, H., & Ortiz, A. (2018). Novel Urinary Biomarkers For Improved Prediction Of Progressive Egfr Loss In Early Chronic Kidney Disease Stages And In High Risk Individuals Without Chronic Kidney Disease. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-34386-8

Rønnow, S. R., Dabbagh, R. Q., Genovese, F., Nanthakumar, C. B., Barrett, V. J., Good, R. B., Brockbank, S., Cruwys, S., Jessen, H., Sorensen, G. L., Karsdal, M. A., Leeming, D. J., & Sand, J. M. B. (2020). Prolonged Scar-in-a-Jar: An in vitro screening tool for anti-fibrotic therapies using biomarkers of extracellular matrix synthesis. *Respiratory Research*, *21*(1), 1–14. https://doi.org/10.1186/s12931-020-01369-1

Saez-Rodriguez, J., Rinschen, M. M., Floege, J., & 1RWTHKramann, R. (2019). *Big science and big data in nephrology*. *November 2018*, 1326–1337. https://doi.org/10.1016/j.kint.2018.11.048

Samarakoon, R., Overstreet, J. M., & Higgins, P. J. (2013). TGF-β signaling in tissue fibrosis: Redox controls, target genes and therapeutic opportunities. *Cellular*

Signalling, *25*(1), 264–268. https://doi.org/10.1016/j.cellsig.2012.10.003

Satake, E., Saulnier, P. J., Kobayashi, H., Gupta, M. K., Looker, H. C., Wilson, J. M., Md Dom, Z. I., Ihara, K., O'Neil, K., Krolewski, B., Pipino, C., Pavkov, M. E., Nair, V., Bitzer, M., Niewczas, M. A., Kretzler, M., Mauer, M., Doria, A., Najafian, B., … Krolewski, A. S. (2021). Comprehensive search for novel circulating mirnas and axon guidance pathway proteins associated with risk of eskd in diabetes. *Journal of the American Society of Nephrology*, *32*(9), 2331–2351. https://doi.org/10.1681/ASN.2021010105

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Schmidt, I. M., Colona, M. R., Kestenbaum, B. R., Alexopoulos, L. G., Palsson, R., Srivastava, A., Liu, J., Stillman, I. E., Rennke, H. G., Vaidya, V. S., Wu, H., Humphreys, B. D., & Waikar, S. (2021). Cadherin-11, Sparc-related modular calcium binding protein-2, and Pigment epithelium-derived factor are promising non-invasive biomarkers of kidney fibrosis. *Kidney International*, *100*(3), 672–683. https://doi.org/10.1016/j.kint.2021.04.037.Cadherin-11

Schneider, D. J., Wu, M., Le, T. T., Cho, S.-H., Brenner, M. B., Blackburn, M. R., & Agarwal, S. K. (2012). CDH11 contributes to pulmonary fibrosis: potential role in TGF-b production and epithelial to mesenchymal transition. *The FASEB Journal*.

Schnieder, J., Mamazhakypov, A., Birnhuber, A., Wilhelm, J., Kwapiszewska, G., Ruppert, C., Markart, P., Wujak, L., Rubio, K., Barreto, G., Schaefer, L., & Wygrecka, M. (2016). Loss of LRP1 promotes acquisition of contractile-myofibroblast phenotype and release of active TGF-β1 from ECM stores. *Matrix Biology*, *August*. https://doi.org/10.1016/j.matbio.2019.12.001

Sciences, B., Douro, A., Real, V., & Technologies, D. (2017). *Pathophysiological Mechanisms of Renal Fibrosis : A Review of Animal Models and Therapeutic Strategies*. *22*, 1–22. https://doi.org/10.21873/invivo.11019

Shao, Q., Zhang, Z., Cao, R., Zang, H., Pei, W., & Sun, T. (2020). Cpa4 promotes emt in pancreatic cancer via stimulating pi3k-akt-mtor signaling. *OncoTargets and Therapy*, *13*, 8567–8580. https://doi.org/10.2147/OTT.S257057

Sheu, C. C., Chang, W. A., Tsai, M. J., Liao, S. H., Chong, I. W., & Kuo, P. L. (2019). Gene expression changes associated with nintedanib treatment in idiopathic pulmonary fibrosis fibroblasts: A next-generation sequencing and bioinformatics study. *Journal of Clinical Medicine*, *8*(3). https://doi.org/10.3390/jcm8030308

Shindo, T., Doi, S., Nakashima, A., Sasaki, K., Arihiro, K., & Masaki, T. (2018). TGF-β1 promotes expression of fibrosis-related genes through the induction of histone variant H3.3 and histone chaperone HIRA. *Scientific Reports*, *8*(1), 1–13. https://doi.org/10.1038/s41598-018-32518-8

Shree, B., Sengar, S., Tripathi, S., & Sharma, V. (2023). LINC01711 promotes transforming growth factor-beta (TGF-β) induced invasion in glioblastoma multiforme (GBM) by acting as a competing endogenous RNA for miR-34a and promoting ZEB1 expression. *Neuroscience Letters*, *792*(October 2022), 136937.

https://doi.org/10.1016/j.neulet.2022.136937

Smyth, G. K., Ritchie, M. E., Law, C. W., Alhamdoosh, M., Su, S., Dong, X., & Tian, L. (2018). RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research*, *5*(1), 1–29. https://doi.org/10.12688/f1000research.9005.3

Sobecki, M., Chen, J., Krzywinska, E., Nagarajan, S., Fan, Z., Nelius, E., Monné Rodriguez, J. M., Seehusen, F., Hussein, A., Moschini, G., Hajam, E. Y., Kiran, R., Gotthardt, D., Debbache, J., Badoual, C., Sato, T., Isagawa, T., Takeda, N., Tanchot, C., … Stockmann, C. (2022). Vaccination-based immunotherapy to target profibrotic cells in liver and lung. *Cell Stem Cell*, *29*(10), 1459-1474.e9. https://doi.org/10.1016/j.stem.2022.08.012

Soneson, C., Marini, F., Geier, F., Love, M. I., & Stadler, M. B. (2020). ExploreModelMatrix : Interactive exploration for improved understanding of design matrices and linear models in R [version 2 ; peer review : 3 approved]. *F1000Research*, *9*(512). https://doi.org/10.12688/f1000research.24187.2

Stahnke, T., Kowtharapu, B. S., Stachs, O., Schmitz, K., Wurm, J., Wree, A., Guthoff, R. F., & Hovakimyan, M. (2017). Suppression of TGF-β pathway by pirfenidone decreases extracellular matrix deposition in ocular fibroblasts in vitro. *PLoS ONE*, 1–20. https://doi.org/10.1371/journal.pone.0172592

Stirling, D. R., Swain-Bowden, M. J., Lucas, A. M., Carpenter, A. E., Cimini, B. A., & Goodman, A. (2021). CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics*, *22*(1), 1–11. https://doi.org/10.1186/s12859-021-04344-9

Stribos, E. G. D., Seelen, M. A., Goor, H. Van, Olinga, P., & Mutsaers, H. A. M. (2017). *Murine Precision-Cut Kidney Slices as an ex vivo Model to Evaluate the Role of Transforming Growth Factor- β 1 Signaling in the Onset of Renal Fibrosis*. *8*(December), 1–9. https://doi.org/10.3389/fphys.2017.01026

Szalai, B., & Saez-Rodriguez, J. (2020). Why do pathway methods work better than they should? *FEBS Letters*, *594*(24), 4189–4200. https://doi.org/10.1002/1873-3468.14011

Sziksz, E., Pap, D., Lippai, R., Béres, N. J., Fekete, A., Szabó, A. J., & Vannay, Á. (2015). *Fibrosis Related Inflammatory Mediators : Role of the IL-10 Cytokine Family*. *2015*.

R Core Team (2021). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*. https://www.r-project.org/

Thompson, A., Wölmer, N., Koncarevic, S., Selzer, S., Böhm, G., Legner, H., Schmid, P., Kienle, S., Penning, P., Höhle, C., Berfelde, A., Martinez-Pinna, R., Farztdinov, V., Jung, S., Kuhn, K., & Pike, I. (2019). TMTpro: Design, Synthesis, and Initial Evaluation of a Proline-Based Isobaric 16-Plex Tandem Mass Tag Reagent Set. *Analytical Chemistry*, *91*(24), 15941–15950. https://doi.org/10.1021/acs.analchem.9b04474

Tie, Y., Tang, F., Peng, D., Zhang, Y., & Shi, H. (2022). TGF-beta signal transduction: biology, function and therapy for diseases. *Molecular Biomedicine*, *3*(1). https://doi.org/10.1186/s43556-022-00109-9

Timmers, H. T. M., & Tora, L. (2018). Transcript Buffering: A Balancing Act between mRNA Synthesis and mRNA Degradation. *Molecular Cell*, *72*(1), 10–17.

https://doi.org/10.1016/j.molcel.2018.08.023

Türei, D., Korcsmáros, A., & Saez-Rodriguez, J. (2016). OmniPath : guidelines and gateway for literature-curated signaling pathway resources. *Nature Publishing Group*, *13*(12), 966–967. https://doi.org/10.1038/nmeth.4077

Vallés, P., Jorro, F., Carrizo, L., Manucha, W., Oliva, J., Cuello-Carrión, F. D., & Ciocca, D. R. (2003). Heat shock proteins HSP27 and HSP70 in unilateral obstructed kidneys. *Pediatric Nephrology*, *18*, 527–535. https://doi.org/10.1007/s00467-003-1096-2

Van Linthout, S., Miteva, K., & Tschöpe, C. (2014). Crosstalk between fibroblasts and inflammatory cells. *Cardiovascular Research*, *102*(2), 258–269. https://doi.org/10.1093/cvr/cvu062

Vaughan, D. E., Rai, R., Khan, S. S., Eren, M., & Ghosh, A. K. (2017). PAI-1 is a Marker and a Mediator of Senescence. *Physiology & Behavior*, *176*(1), 139–148. https://doi.org/10.1161/ATVBAHA.117.309451.PAI-1

Verbeke, F., Siwy, J., Van Biesen, W., Mischak, H., Pletinck, A., Schepers, E., Neirynck, N., Magalhães, P., Pejchinovski, M., Pontillo, C., Lichtinghagen, R., Brand, K., Vlahou, A., De Bacquer, D., & Glorieux, G. (2021). The urinary proteomics classifier chronic kidney disease 273 predicts cardiovascular outcome in patients with chronic kidney disease. *Nephrology, Dialysis, Transplantation : Official Publication of the European Dialysis and Transplant Association - European Renal Association*, *36*(5), 811–818. https://doi.org/10.1093/ndt/gfz242

Vidyasagar, A., Reese, S., Acun, Z., Hullett, D., & Djamali, A. (2008). HSP27 is involved in the pathogenesis of kidney tubulointerstitial fibrosis. *American Journal of Renal Physiology*, *53713*, 707–716. https://doi.org/10.1152/ajprenal.90240.2008.

Vidyasagar, A., Wilson, N. A., & Djamali, A. (2012). Heat shock protein 27 (HSP27): biomarker of disease and therapeutic target. *Fibrogenesis and Tissue Repair*, *5*(7), 1–7.

Virág, D., Dalmadi-Kiss, B., Vékey, K., Drahos, L., Klebovich, I., Antal, I., & Ludányi, K. (2020). Current Trends in the Analysis of Post-translational Modifications. *Chromatographia*, *83*(1), 1–10. https://doi.org/10.1007/s10337-019-03796-9

Vu, T. N., Chen, X., Foda, H. D., Smaldone, G. C., & Hasaneen, N. A. (2019). Interferon-γ enhances the antifibrotic effects of pirfenidone by attenuating IPF lung fibroblast activation and differentiation. *Respiratory Research*, *20*(1), 1–14. https://doi.org/10.1186/s12931-019-1171-2

Walker, E. J., Heydet, D., Veldre, T., & Ghildyal, R. (2019). Transcriptomic changes during TGF-β-mediated differentiation of airway fibroblasts to myofibroblasts. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-56955-1

Walraven, M., & Hinz, B. (2018). Therapeutic approaches to control tissue repair and fibrosis: Extracellular matrix as a game changer. *Matrix Biology*, *71–72*(2017), 205–224. https://doi.org/10.1016/j.matbio.2018.02.020

Wang, W. juan, Chen, X. mei, & Cai, G. yan. (2021). Cellular senescence and the senescence-associated secretory phenotype: Potential therapeutic targets for

renal fibrosis. *Experimental Gerontology*, *151*(May), 111403. https://doi.org/10.1016/j.exger.2021.111403

Wei, Q., Holle, A., Li, J., Posa, F., Biagioni, F., Croci, O., Benk, A. S., Young, J., Noureddine, F., Deng, J., Zhang, M., Inman, G. J., Spatz, J. P., Campaner, S., & Cavalcanti-Adam, E. A. (2020). BMP-2 Signaling and Mechanotransduction Synergize to Drive Osteogenic Differentiation via YAP/TAZ. *Advanced Science*, *7*(15), 1–15. https://doi.org/10.1002/advs.201902931

Werner, T., Sweetman, G., Savitski, M. F., Mathieson, T., Bantscheff, M., & Savitski, M. M. (2014). Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Analytical Chemistry*, *86*(7), 3594–3601. https://doi.org/10.1021/ac500140s

Wozniak, J., Floege, J., Ostendorf, T., & Ludwig, A. (2021). Key metalloproteinase-mediated pathways in the kidney. *Nature Reviews Nephrology*, *17*. https://doi.org/10.1038/s41581-021-00415-5

Wu, X., Verschut, V., Woest, M. E., Ng-Blichfeldt, J. P., Matias, A., Villetti, G., Accetta, A., Facchinetti, F., Gosens, R., & Kistemaker, L. E. M. (2021). Rho-Kinase 1/2 Inhibition Prevents Transforming Growth Factor-β-Induced Effects on Pulmonary Remodeling and Repair. *Frontiers in Pharmacology*, *11*(January), 1–12. https://doi.org/10.3389/fphar.2020.609509

Xu, J., Zhou, L., & Liu, Y. (2020). Cellular Senescence in Kidney Fibrosis: Pathologic Significance and Therapeutic Strategies. *Frontiers in Pharmacology*, *11*(December), 1–17. https://doi.org/10.3389/fphar.2020.601325

Yamashita, N., & Kramann, R. (2024). Mechanisms of kidney fibrosis and routes towards therapy. In *Trends in Endocrinology and Metabolism* (Vol. 35, Issue 1, pp. 31–48). Elsevier Inc. https://doi.org/10.1016/j.tem.2023.09.001

Yan, Z., Wang, G., Shi, X., & Chen, L. (2021). Advances in the Progression and Prognosis Biomarkers of Chronic Kidney Disease. *Frontiers in Pharmacology*, *12*(December), 1–20. https://doi.org/10.3389/fphar.2021.785375

Zeisberg, M., & Kalluri, R. (2015). Physiology of the renal interstitium. *Clinical Journal of the American Society of Nephrology*, *10*(10), 1831–1840. https://doi.org/10.2215/CJN.00640114

Zhang, F., Zhang, Y., Sun, L. xin, Chen, M., Ran, Y. liang, & Sun, L. chao. (2019). Carboxypeptidase A4 promotes migration and invasion of lung cancer cells, and is closely associated with lymph node metastasis. *Precision Radiation Oncology*, *3*(2), 44–51. https://doi.org/10.1002/pro6.1068

Zhang, X., Zhang, X., Huang, W., & Ge, X. (2021). The role of heat shock proteins in the regulation of fibrotic diseases. *Biomedicine & Pharmacotherapy*, *135*, 111067. https://doi.org/10.1016/j.biopha.2020.111067

Zhao, X., Kwan, J. Y. Y., Yip, K., Liu, P. P., & Liu, F. (2020). Targeting metabolic dysregulation for fibrosis therapy. *Nature Reviews Drug Discovery*, *19*(January). https://doi.org/10.1038/s41573-019-0040-5

Zhou, S., Yin, X., Mayr, M., Noor, M., Hylands, P. J., & Xu, Q. (2020). Proteomic landscape of TGF-β1-induced fibrogenesis in renal fibroblasts. *Scientific Reports*, *10*(1), 1–17. https://doi.org/10.1038/s41598-020-75989-4

Zi, Z., Feng, Z., Chapnick, D. A., Dahl, M., Deng, D., Klipp, E., Moustakas, A., & Liu,

X. (2011). Quantitative analysis of transient and sustained transforming growth factor-β signaling dynamics. *Molecular Systems Biology*, *7*. https://doi.org/10.1038/msb.2011.22

Ziegler, D. V, Vindrieux, D., Goehrig, D., Jaber, S., Collin, G., Griveau, A., Wiel, C., Bendridi, N., Djebali, S., Farfariello, V., Prevarskaya, N., Payen, L., Marvel, J., Aubert, S., Flaman, J., Rieusset, J., Martin, N., & Bernard, D. (2021). Calcium channel ITPR2 and mitochondria–ER contacts promote cellular senescence and aging. *Nature Communications*, *12*(720), 1–12. https://doi.org/10.1038/s41467-021-20993-z

Zou, M., Zou, J., Hu, X., Zheng, W., Zhang, M., & Cheng, Z. (2021). Latent Transforming Growth Factor-β Binding Protein-2 Regulates Lung Fibroblast-to-Myofibroblast Differentiation in Pulmonary Fibrosis via NF-κB Signaling. *Frontiers in Pharmacology*, *12*(December), 1–14. https://doi.org/10.3389/fphar.2021.788714

Zürbig, P., Siwy, J., & Mischak, H. (2019). Emerging urine-based proteomic biomarkers as valuable tools in the management of chronic kidney disease. *Expert Review of Molecular Diagnostics*, *19*(10), 853–856. https://doi.org/10.1080/14737159.2019.1657406

Figure and Table References

Figure 72 Author's chart
Figure 73 Author's chart
Figure 74 Author's chart
Figure 75 Author's chart
Figure 76 Author's chart, screenshots from bluegecko.embl.de
Figure 77 Author's chart
Figure 78 Author's chart
Figure 79 Author's chart
Figure 80 Author's chart
Figure 81 Author's chart
Table 1 Adjusted after Cañadas-Garre et al., 2018, 2019
Table 2 Author's table
Table 3 Author's table
Table 4 Author's table
Table 5 Author's table
Table 6 Author's table
Table 7 Author's table
Table 8 Author's table
Table 9 Author's table
Table 10 Author's table
Table 11 Author's table
Table 12 Author's table
Table 13 Author's table
Table 14 Author's table
Table 15 Author's table
Table 16 Author's table
Table 17 Author's table
Table 18 Author's table supported by Sarah Kaspar
Table 19 Author's table supported by Sarah Kaspar
Table 20 Author's table supported by Sarah Kaspar
Table 21 Author's table supported by Sarah Kaspar
Table 22 Author's table supported by Sarah Kaspar
Table 23 Author's table supported by Sarah Kaspar
Table 24 Author's table supported by Sarah Kaspar

Annex

Annex for Chapter 4.1 Rationale and Characterization of the cellular model

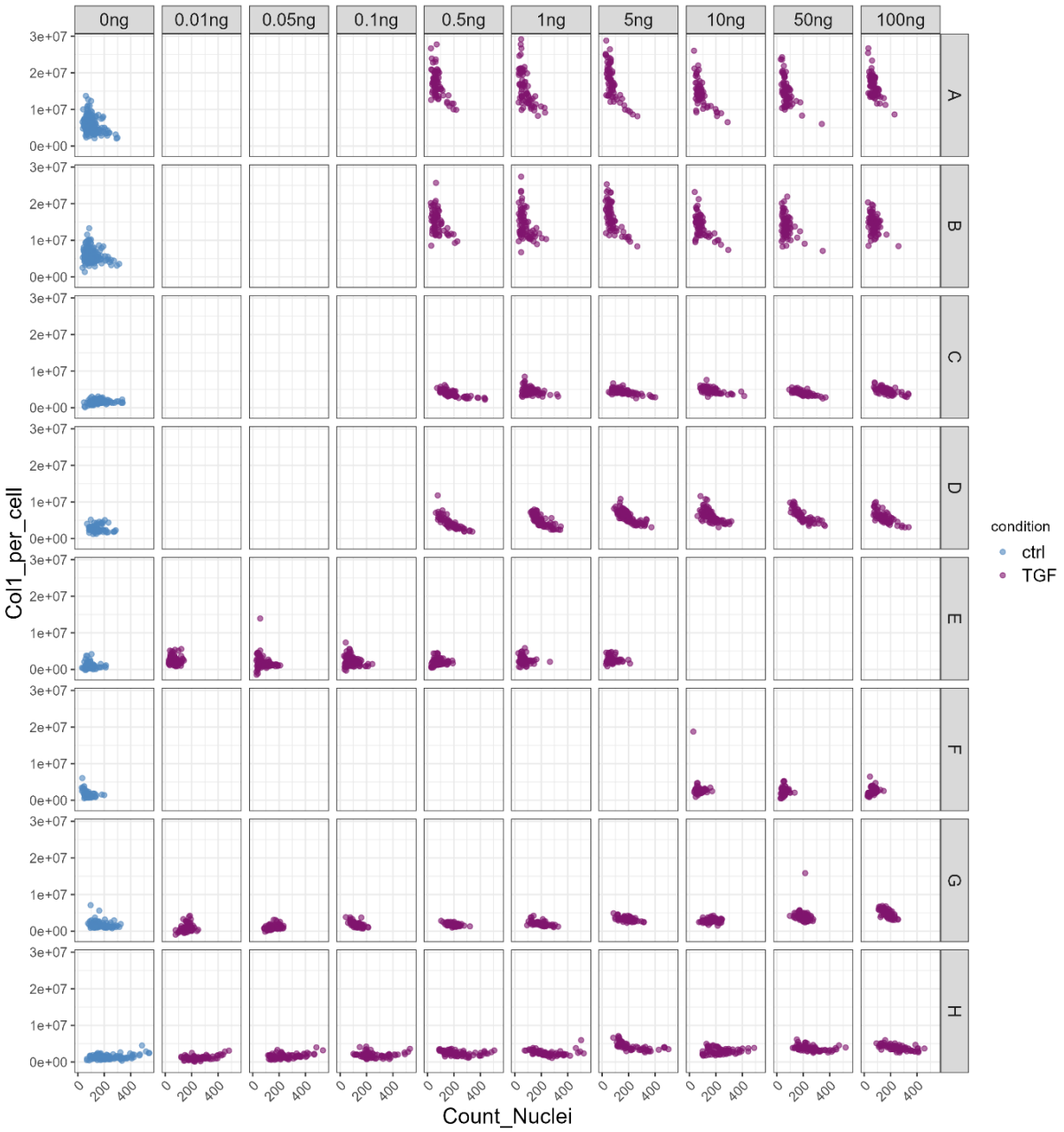

Figure 60 Col I per cell vs nuclei number per TGF-β concentration used and biological replicate. Data points are representative of each image acquired, colors depict the condition (blue = control (ctrl), purple = TGF-β treated). On the x-axis the nuclei number (Count_Nuclei) per image is depicted, including images with > 25 nuclei. The y-axis represents fluorescence intensity (Col I staining) normalized to the nuclei number. Data are shown after background subtraction, for each concentration of TGF-β (0, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 50, 100 ng/ml). A – H indicate the different biological replicates.

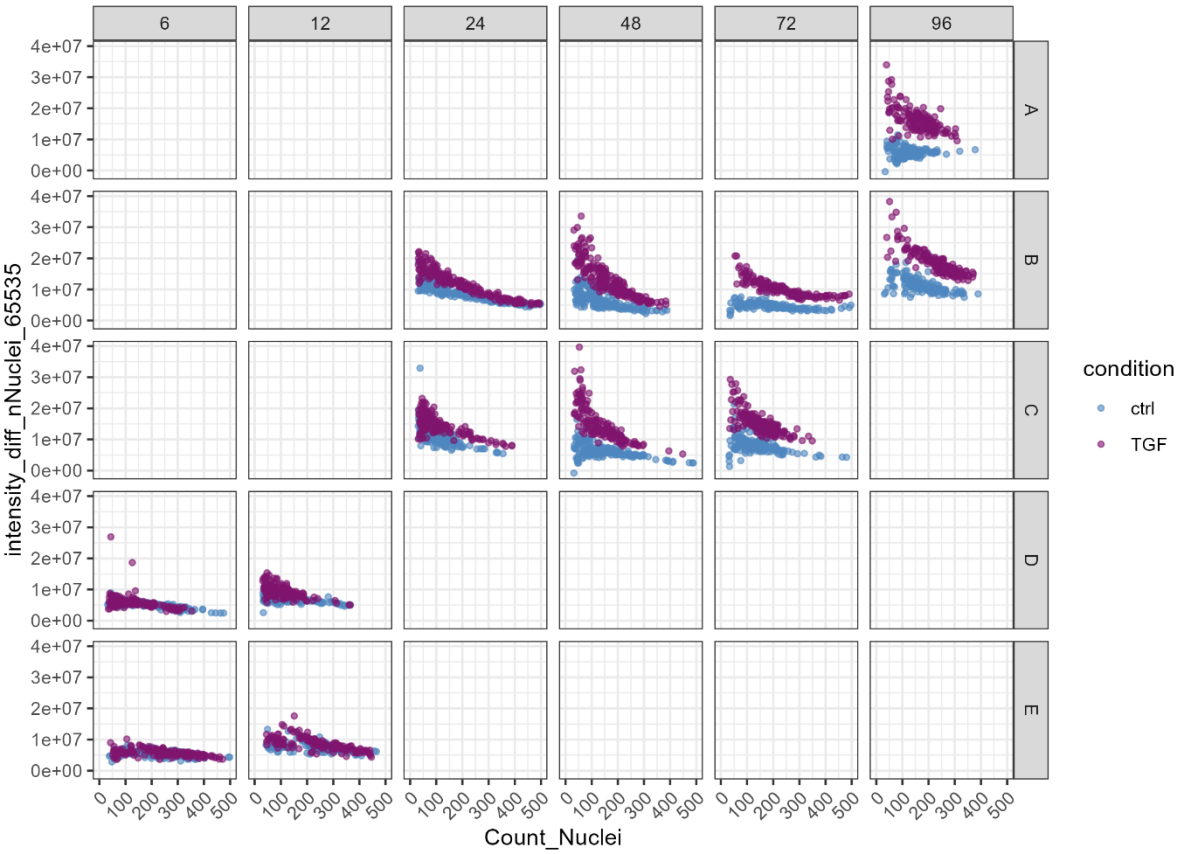

Figure 61 Col I per cell vs nuclei number per time point and biological replicate. Data points are representative of each image acquired, colors depict the condition (blue = control (ctrl), purple = TGF-β treated). On the x-axis the nuclei number (Count_Nuclei) per image is depicted, including images with > 30 nuclei. The y-axis represents fluorescence intensity (Col I staining) normalized to the nuclei number. Data are shown after background subtraction, per time point (6, 12, 24, 48, 72, 96 h). A – E indicate the different biological replicates/plates.

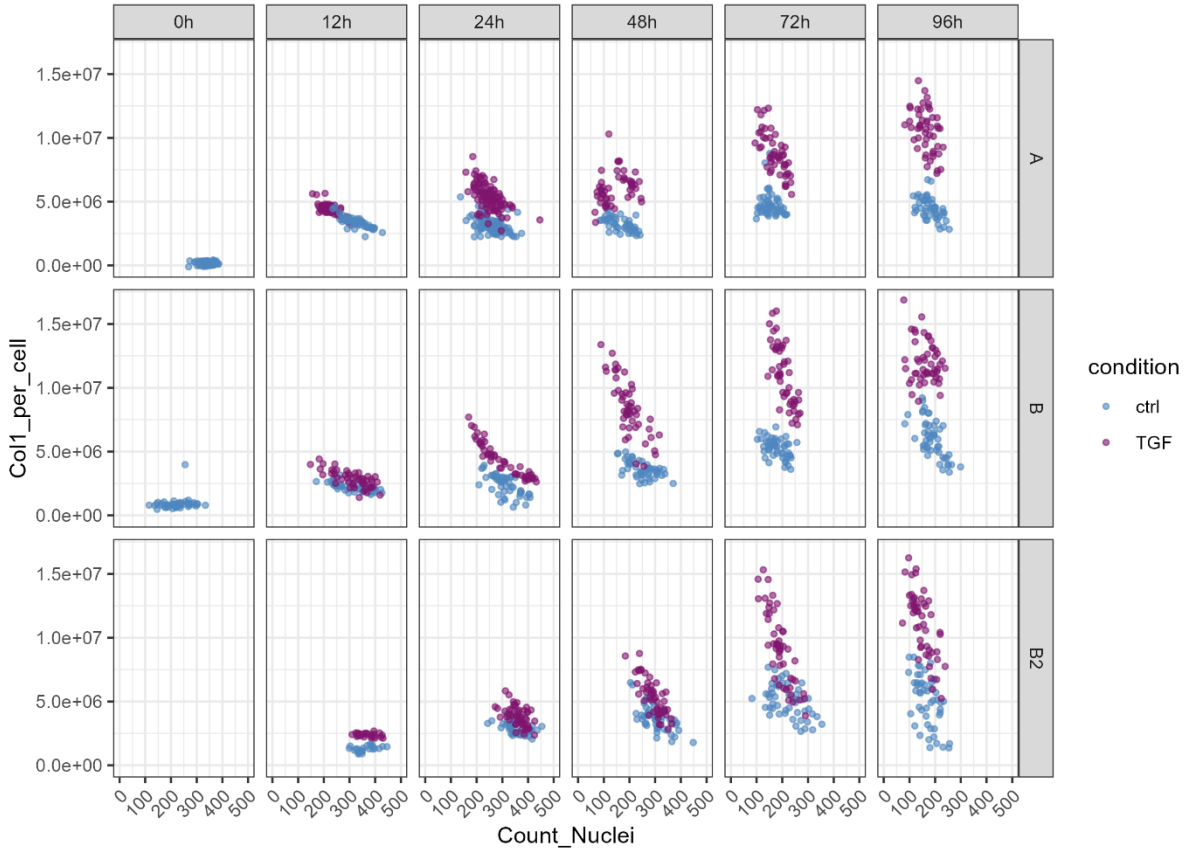

Figure 62 Col I per cell vs nuclei number per time point and biological replicate comparing experimental set-up A vs B. Data points are representative of each image acquired, colors depict the condition (blue = control (ctrl), purple = TGF-β treated). On the x-axis the nuclei number (Count_Nuclei) per image is depicted, including images with > 30 nuclei. The y-axis represents fluorescence intensity (Col I staining) normalized to the nuclei number. Data are shown after background subtraction, per time point (0, 12, 24, 48, 72, 96 h). For set-up B, two plates were imaged (B, B2). For set-up A, a separate plate per time point was imaged.

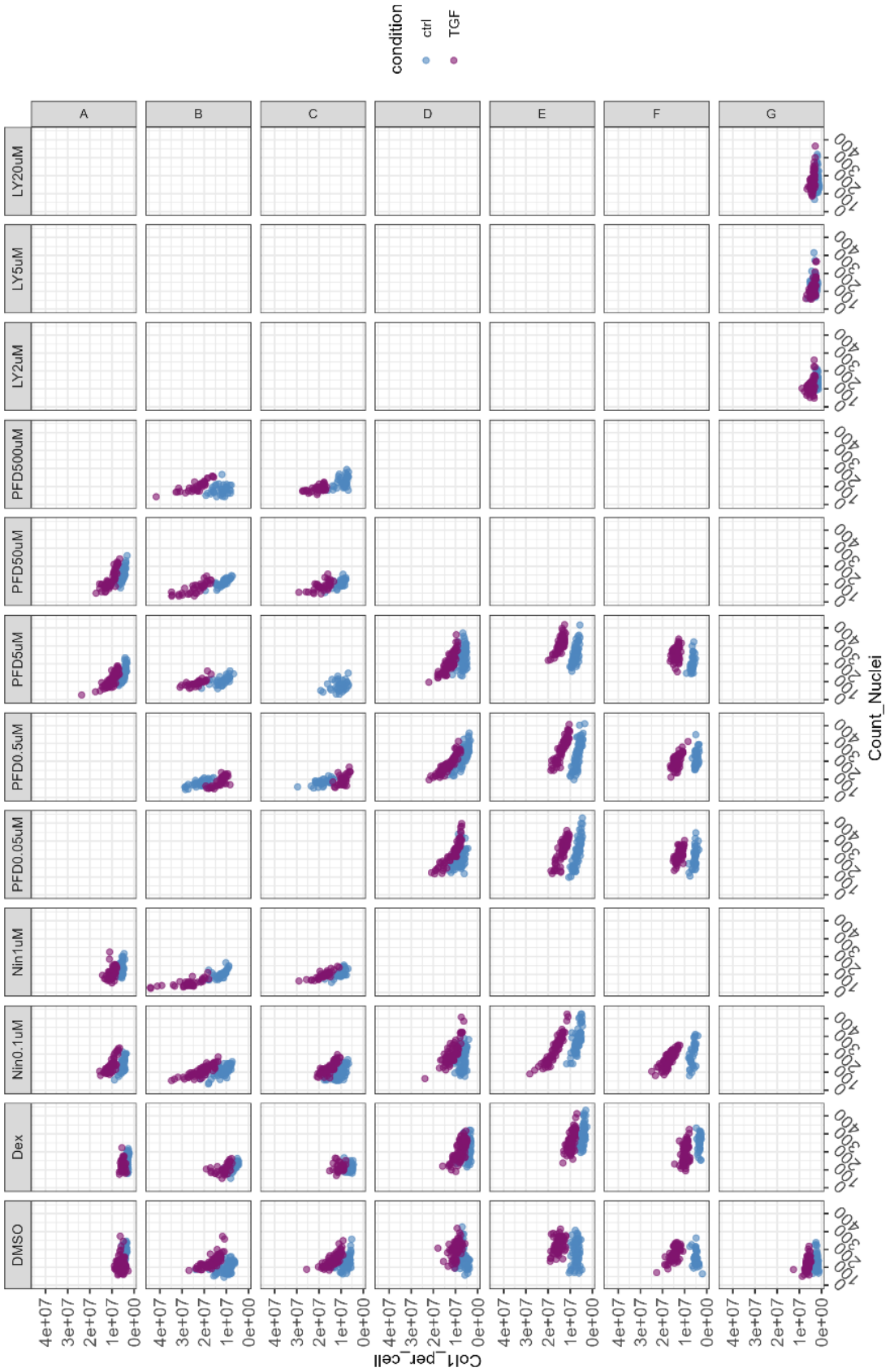

Figure 63 Col I per cell vs nuclei number per drug treatment and biological replicate. Data points are representative of each image acquired, colors depict the condition (blue = control (ctrl), purple = TGF-β treated). On the x-axis the nuclei number (Count_Nuclei) per image is depicted, including images with > 20 nuclei. The y-axis represents fluorescence intensity (Col I staining) normalized to the nuclei number. Data are shown after background subtraction, for each condition. A – G indicate the different biological replicates/plates.

Annex for Chapter 4.2 Multi-omics Insights into Cellular Responses and Pathways

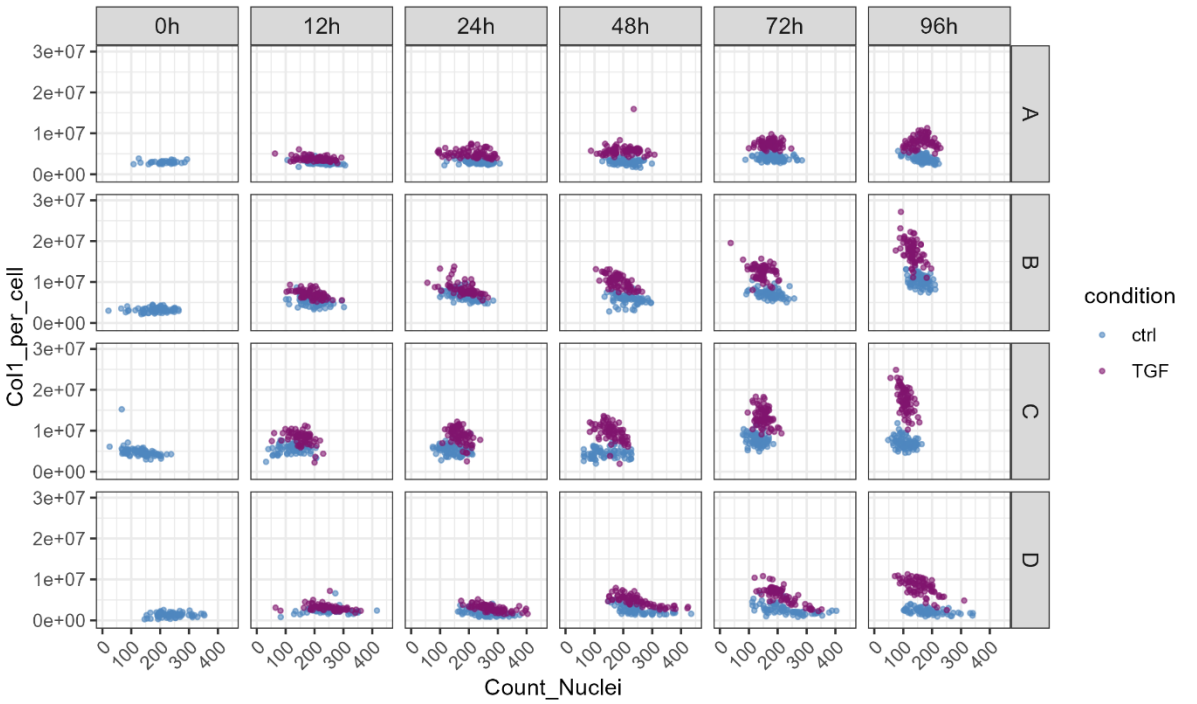

Figure 64 Extracellular Col I per cell versus cell number over time and ± TGF-β stimulation. Cells were cultured in glass-bottom 24 well-plates with four biological replicates (plate A-D) with two technical replicates each. Cells were treated ± TGF-β for durations ranging from 0 to 96 h, and subsequently fixed with PFA. Nuclear visualization was achieved with Hoechst staining, while the extracellular Col I was visualized using an anti-Col I antibody. Images of 36 positions per well were acquired using the wide field microscope Molecular Devices IXM. Initial assessment of nuclei number and intensity measurement, correlating with the Col I staining, was performed using CellProfiler 4.2.6. Background correction and subtraction of the cell autofluorescence were followed by normalizing the resulting intensity values to the cell number to obtain the Col I signal per cell (Col1_per_cell shown on the y-axis). Images with less than 20 nuclei were excluded from the analysis. The resulting data are plotted per biological replicate per time point against the cell number. Data points depict individual images, color coded for ctrl (blue) and TGF-β treated (purple). Over time, the Col I per cell intensity is increasing, and further influenced by TGF-β stimulation.

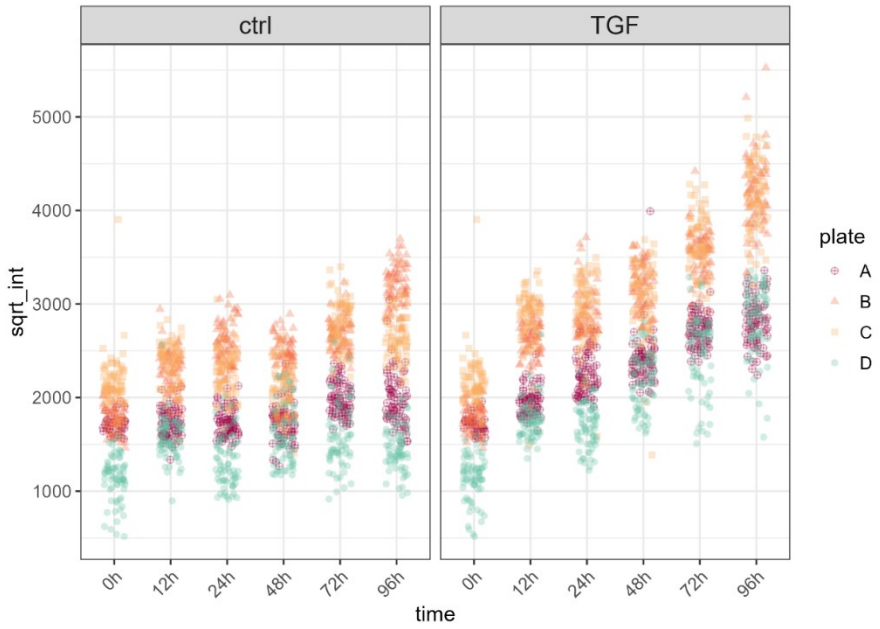

Figure 65 Col I per cell intensity values of each image transformed using square root transformation, per time and condition. Data points depict individual images, color- and shape-coded for plates (biological replicates). Over time, the Col I per cell intensity is increasing, and further influenced by TGF-β stimulation. However, there is a variance based on biological replicates.

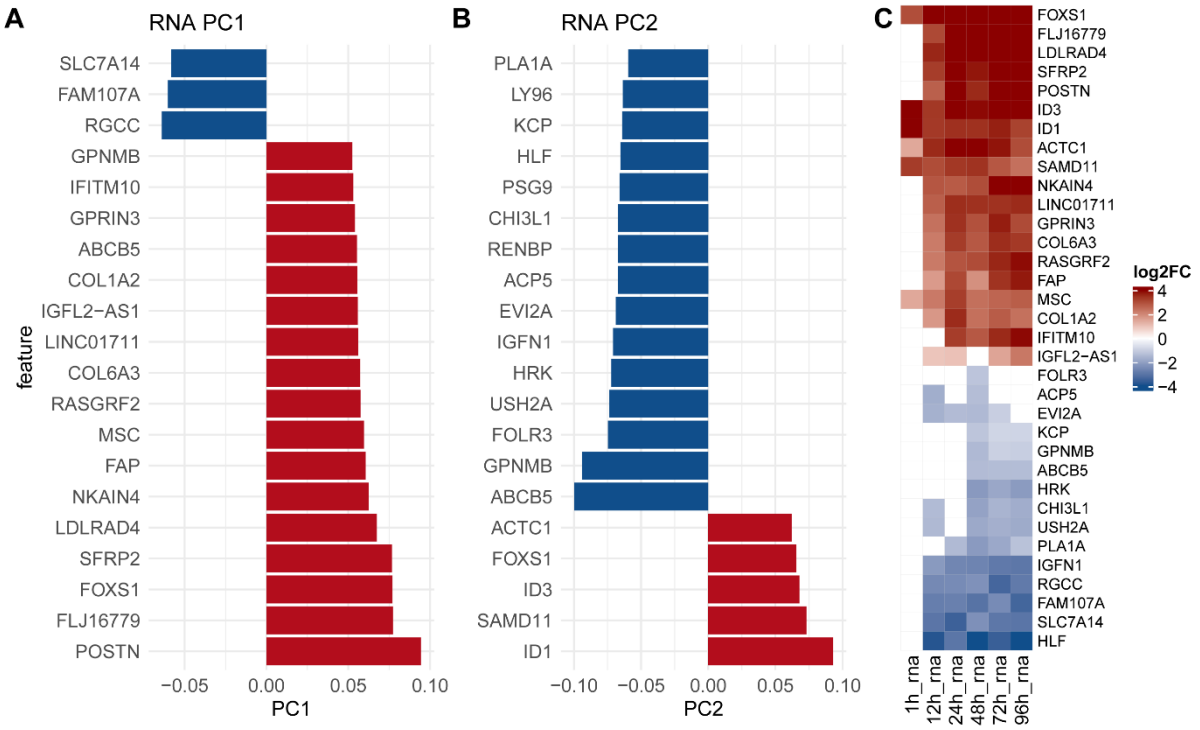

Figure 66 Loading values for principal components and abundance changes over time in transcriptomics data. The top 20 loading values for the first two principal components **A)** PC1 and **B)** PC2 from the principal component analysis are illustrated, alongside **C)** the corresponding abundance changes (log2FC) over time in the transcriptomics data.

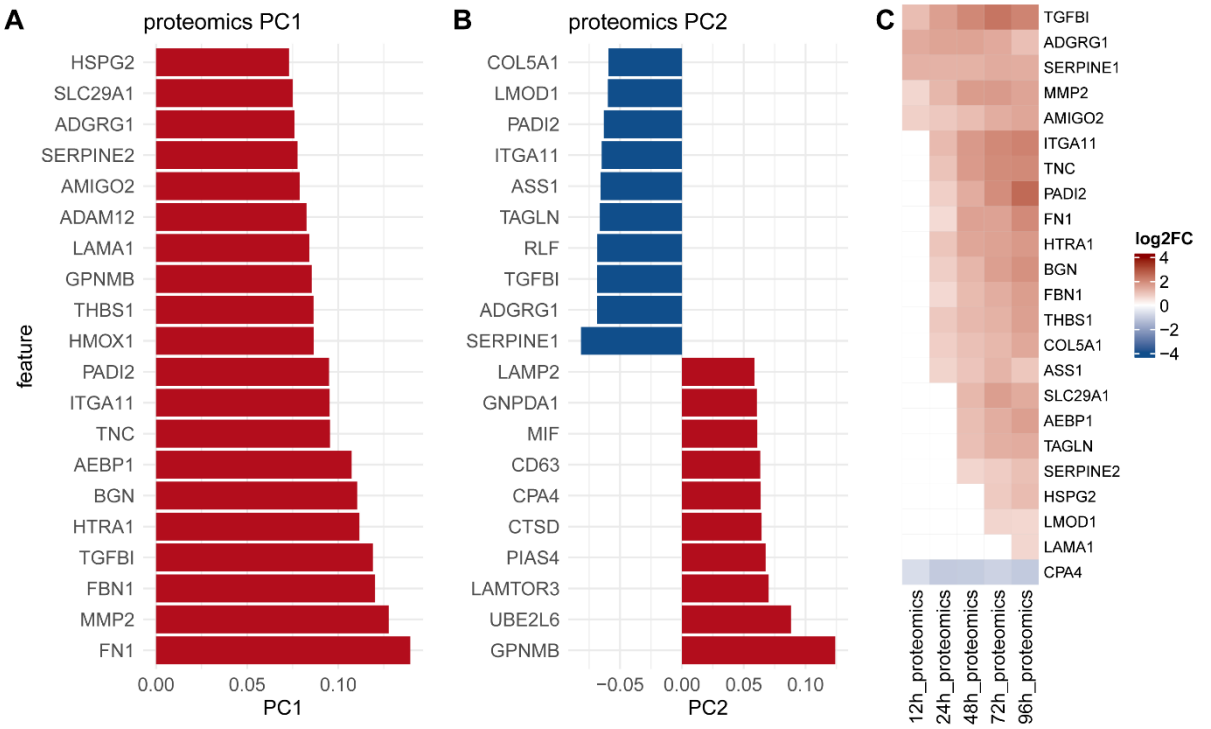

Figure 67 Loading values for the principal components and abundance changes over time in proteomics data. The top 20 loading values for the first two principal components **A)** PC1 and **B)** PC2 from the principal component analysis are illustrated, alongside **C)** the corresponding abundance changes (log2FC) over time in the proteomics data.

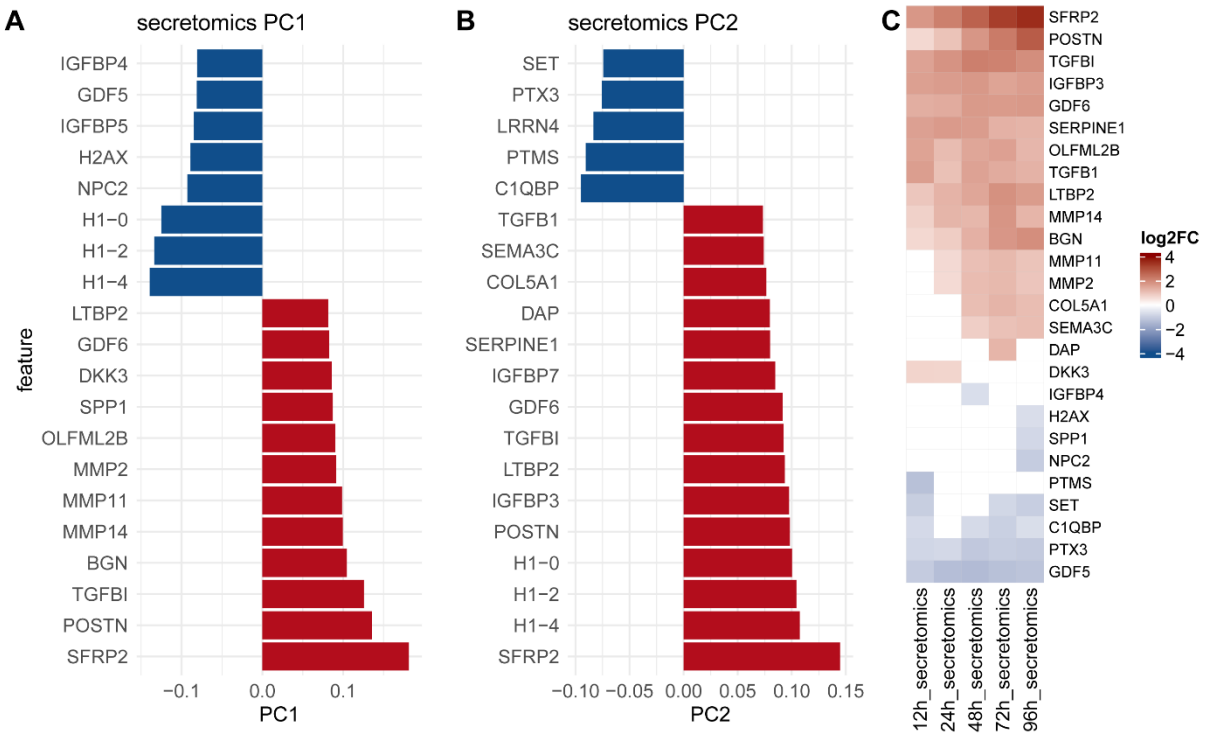

Figure 68 Loading values for the principal components and abundance changes over time in secretomics data. The top 20 loading values for the first two principal components **A)** PC1 and **B)** PC2 from the principal component analysis are illustrated, alongside **C)** the corresponding abundance changes (log2FC) over time in the secretomics data.

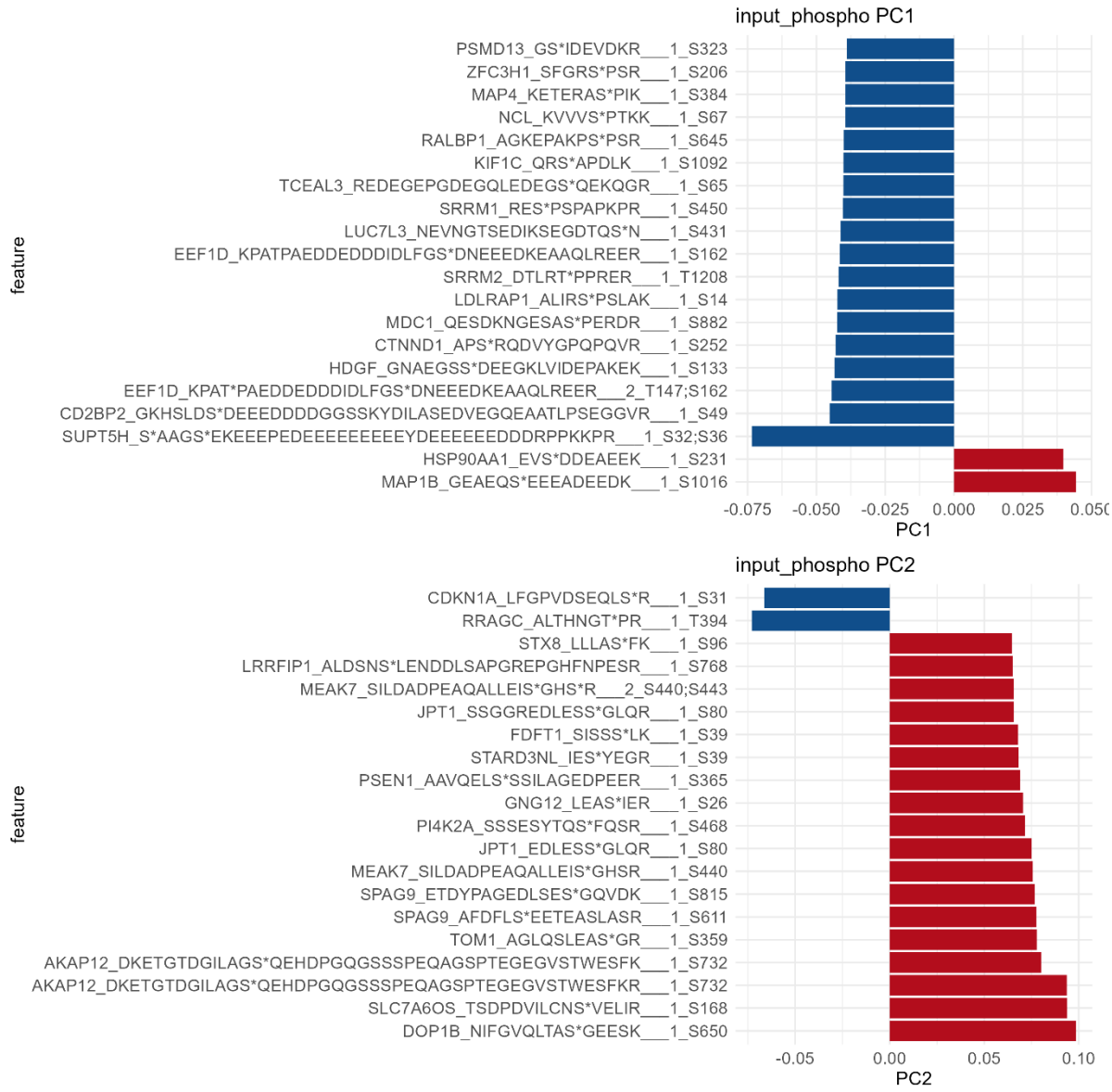

Figure 69 Loading values for the principal components and in phosphoproteomics data. The top 20 loading values for the first two principal components **upper)** PC1 and **lower)** PC2 from the principal component analysis are illustrated for the phosphoproteomics data that were normalized to the proteomics of the cell lysates.

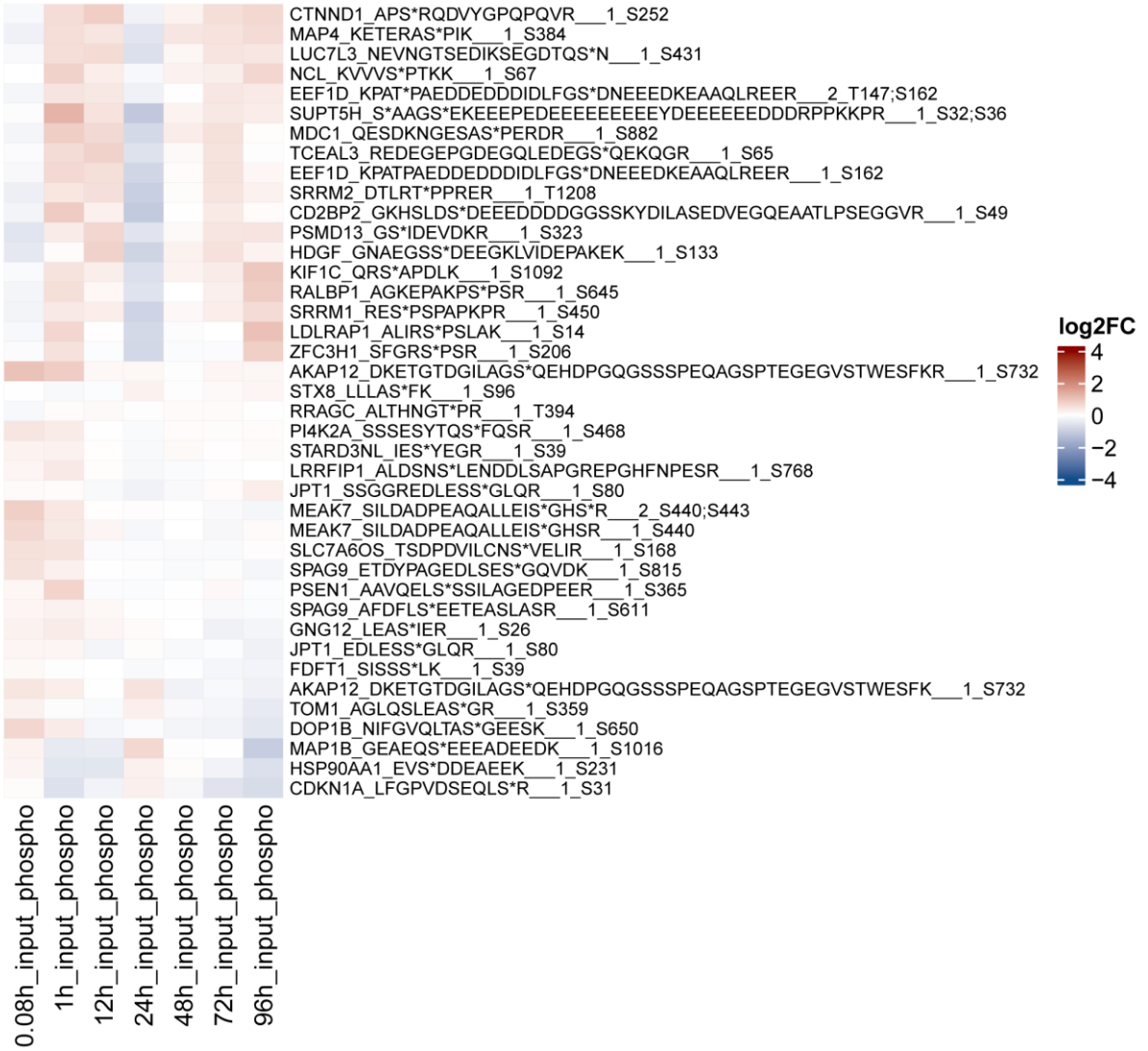

Figure 70 Abundance changes over time of loading values for PCA components in phosphoproteomics data. The abundance changes (log2FC) over time of top 20 loading values for first two principal components are illustrated for the phosphoproteomics data that were normalized to the proteomics of the cell lysates. Considering that none of the changes were either significant nor met the log2 fold-change thresholds, all of the cells are color coded based on the up or downregulation of the phosphor-site, unlike in the previous graphs, where white cells were representative for values that did not meet the log2 fold-change and/or adjusted p-value cut-offs.

Annex for Chapter 4.3 Multi-omics Insights into Cellular Responses and Pathways

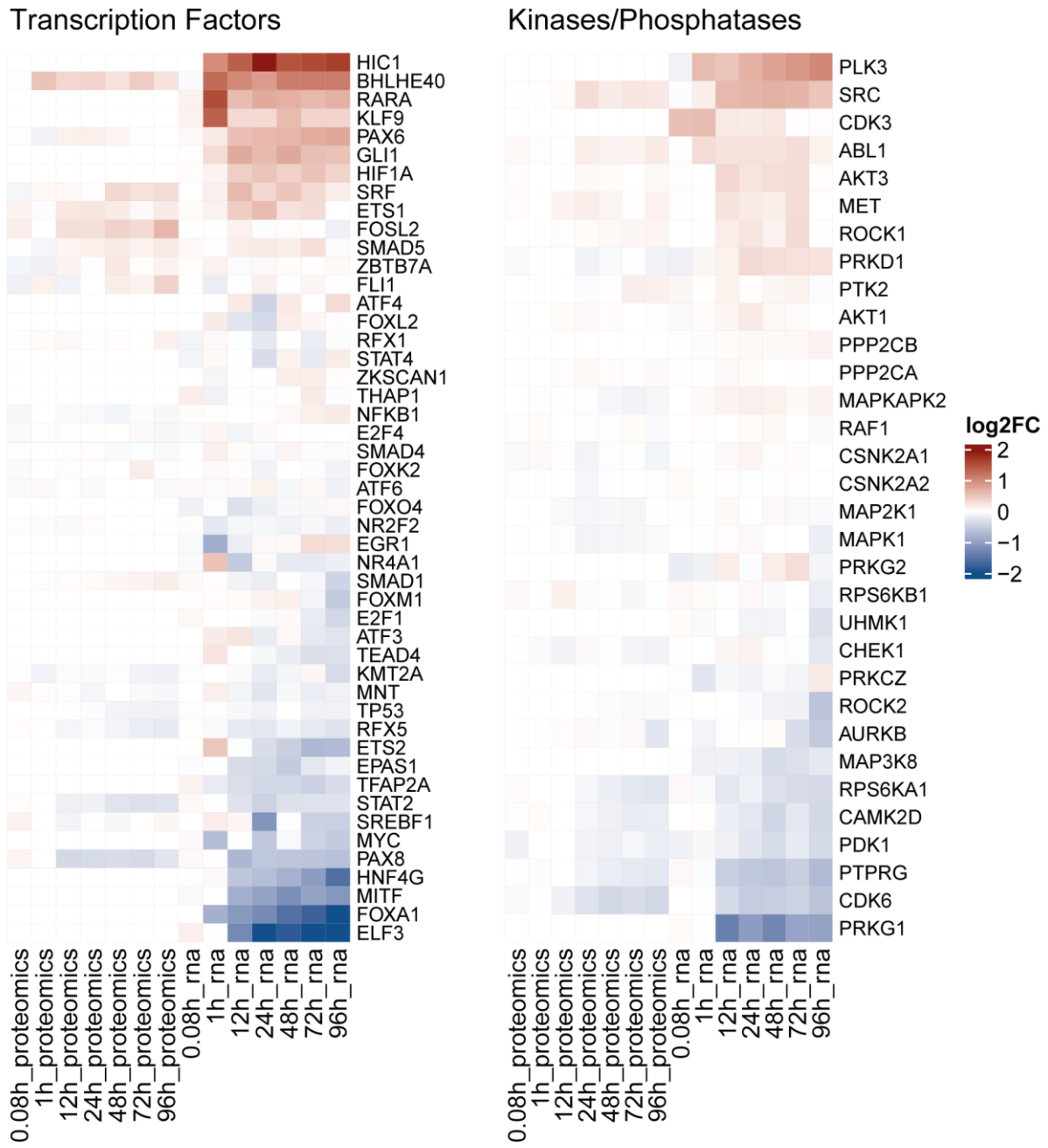

Figure 71 Heatmap of abundance changes of TFs and kinases/phosphatases shown in results section 4.3.1. Abundance changes (log2 fold-changes) of TFs (left) and kinases/phosphatases (right) are shown over time on mRNA and protein level.

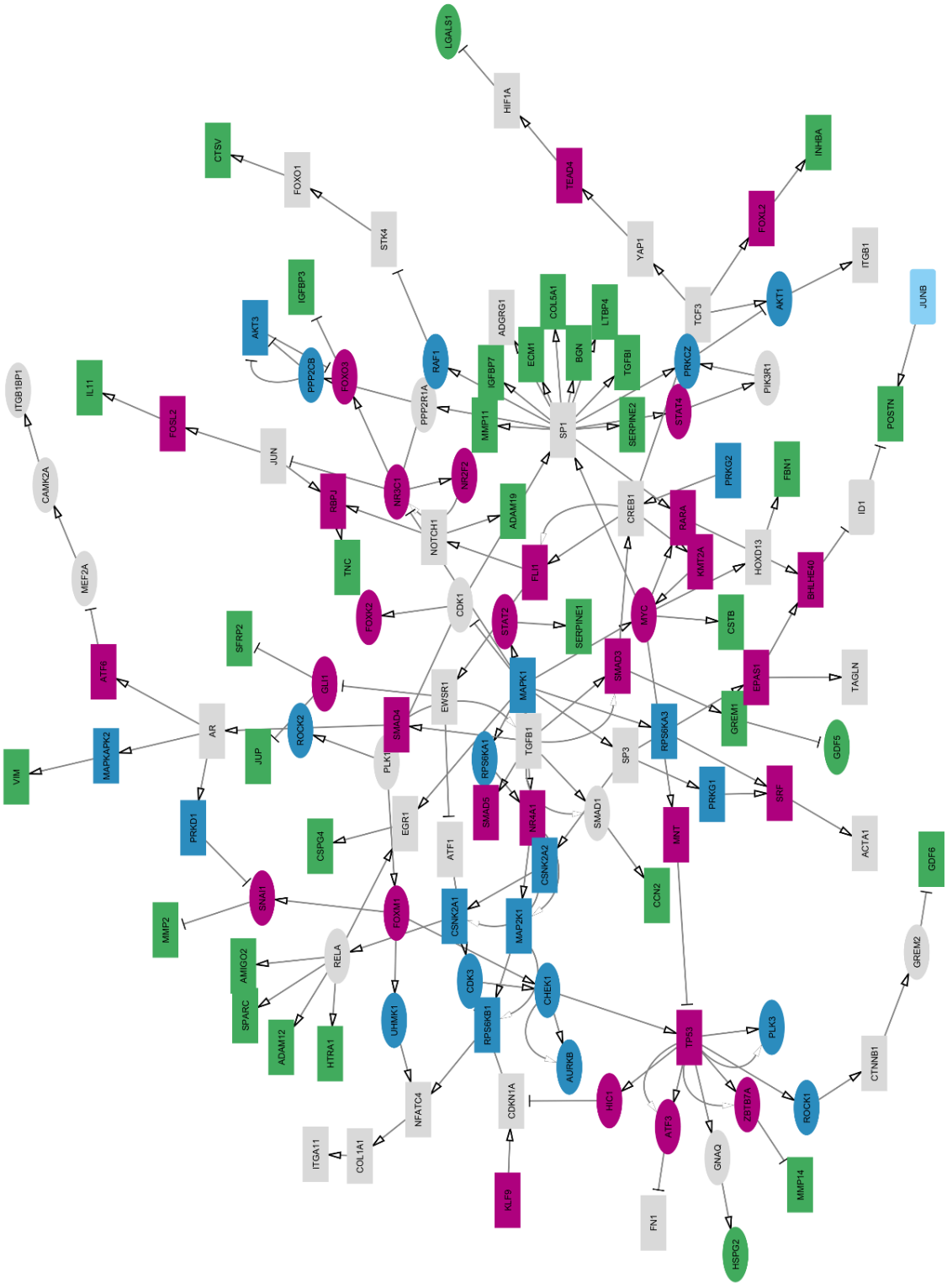

Figure 72 Early network. Colors are representing the different modalities, green for secreted factors, blue for kinases/phosphatases, purple for TFs, grey as factors imputed by prior knowledge. The ellipse shape are factors which expression/activity is downregulated while squares indicate factors that are upregulated.

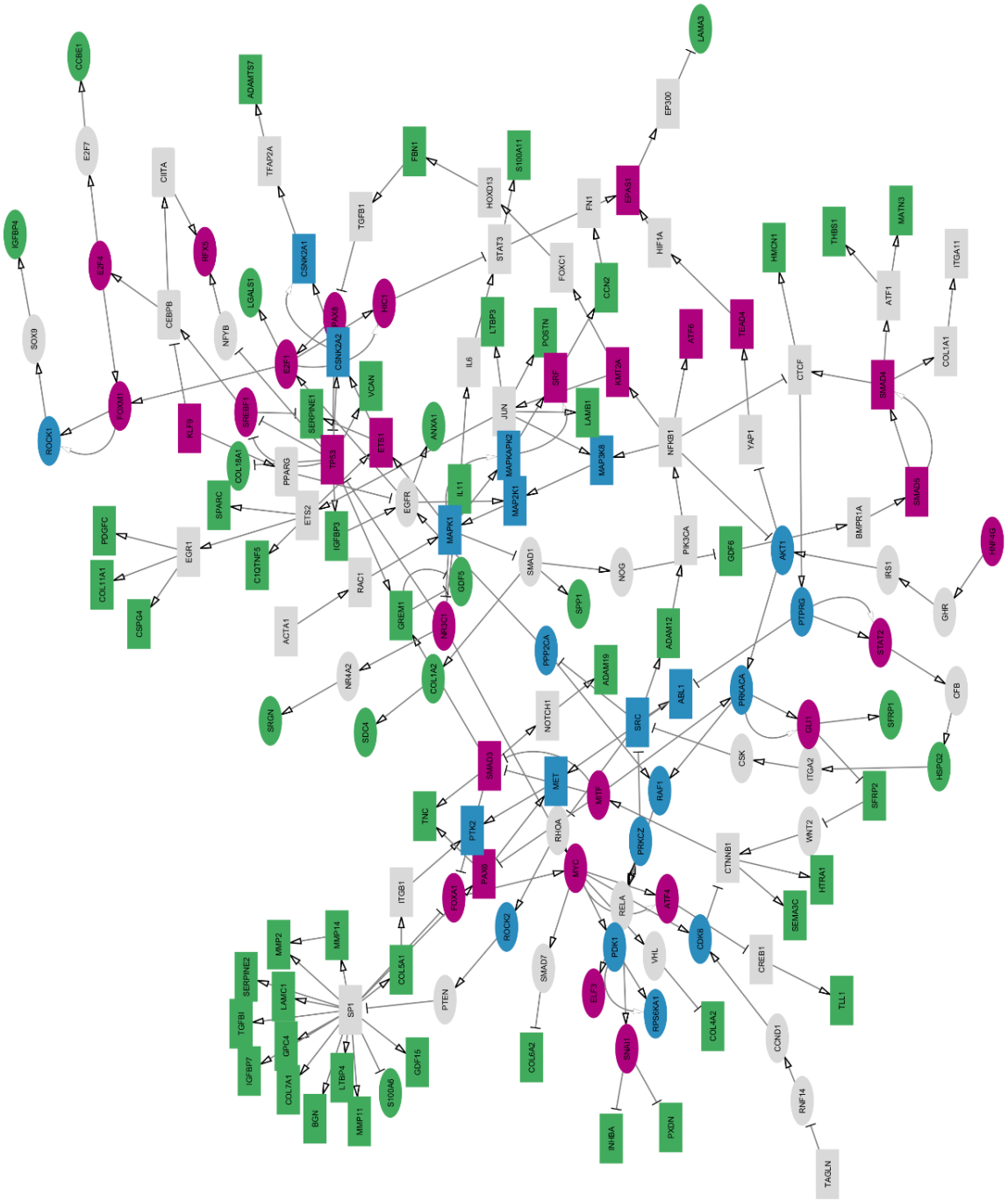

Figure 73 Late network. Colors are representing the different modalities, green for secreted factors, blue for kinases/phosphatases, purple for TFs, grey as factors imputed by prior knowledge. The ellipse shape are factors which expression/activity is downregulated while squares indicate factors that are upregulated.

Annex for Chapter 4.4. Validation of Multi-omics Hits

Quality Assessment Figures of Immunofluorescence Stainings

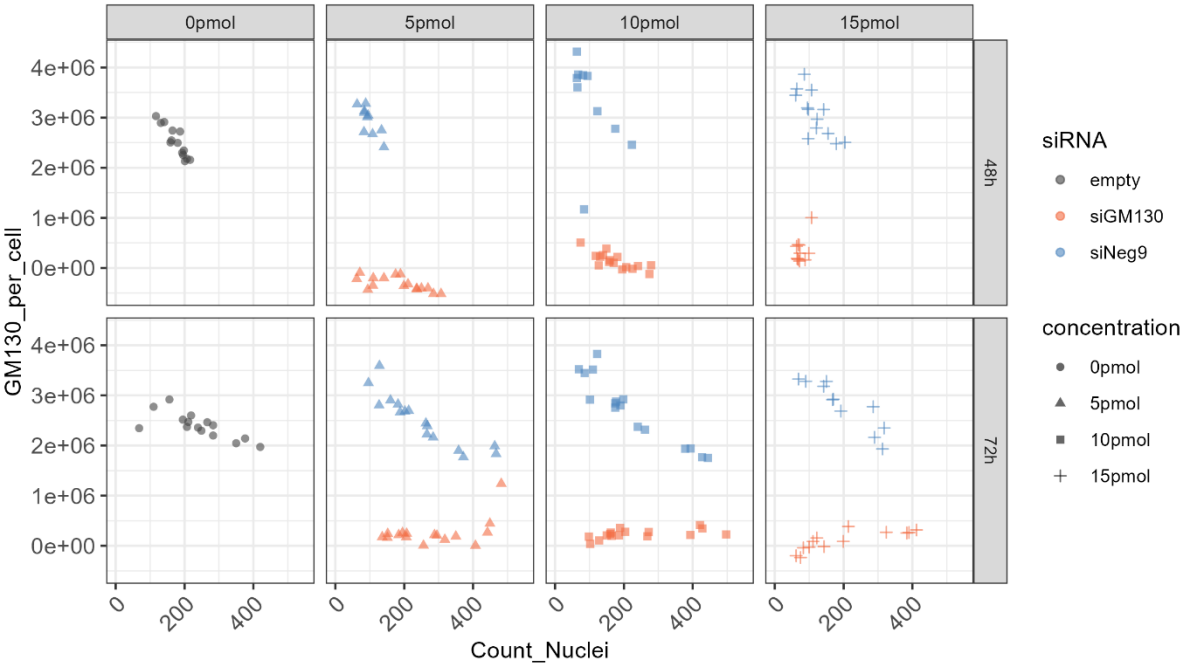

Figure 74 Fluorescence intensity vs nuclei number per time and siRNA concentration. Data points are representative of each image acquired, colors depict the siRNA used (grey = empty vector, blue = siNeg9, orange = siGM130). On the x-axis the nuclei number (Count_Nuclei) per image is depicted, with a cut-off at > 60 nuclei. The y-axis represents fluorescence intensity (GM130 staining) normalized to the nuclei number. Data are shown after background subtraction, for each concentration of siRNA and the time post siRNA transfection. n = 1.

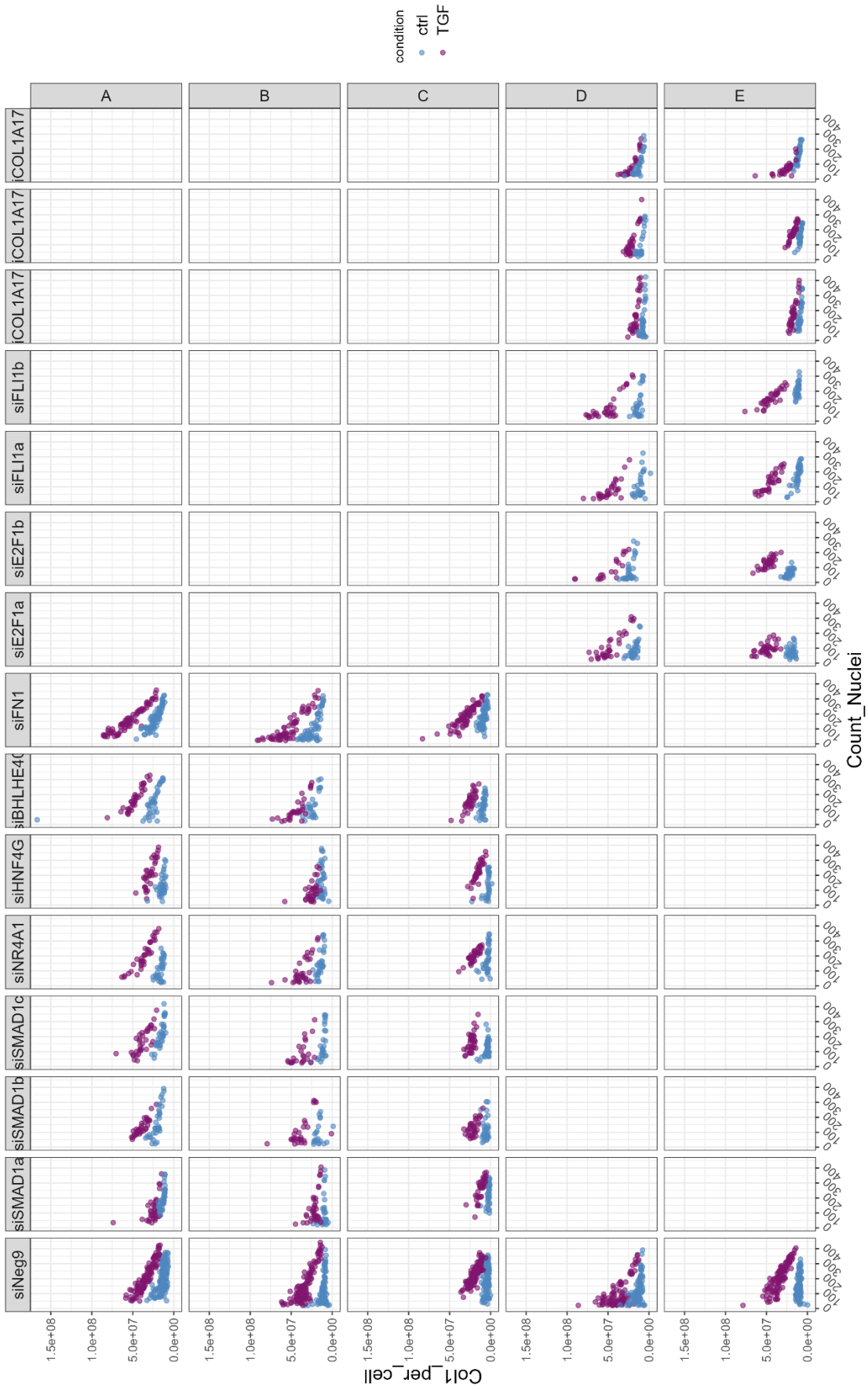

Figure 75 Fluorescence intensity vs nuclei number per siRNA and plate at 96 h. Data points are representative of each image acquired, colors depict the treatment (blue = control, purple = TGF-β treated). On the x-axis the nuclei number (Count_Nuclei) per image is depicted, with a cut-off at > 20

nuclei. The y-axis represents fluorescence intensity (Col I staining) normalized to the nuclei number. Data are shown after background subtraction, for each siRNA used, 96 h post siRNA transfection, for each biological replicate separately. n = 2-3.

Ambion database on Bluegecko

Ensembl 89_38

siRNA: GGAUAAAGUUCUUACUCAAtt
gaCCUAUUUCAAGAAUGAGUU

Supplier: **Ambion Silencer Select Pre-designed siRNA, ID# s8396**

Target gene	Mismatch	transcripts hit/total transcripts	Genome context
SMAD1	0	3/9	Ensembl or UCSC
SMAD5	2	0/8	Ensembl or UCSC

Target transcript	Alignment
SMAD1, transcript: ENST00000515385, chr 4	DSR00114447 21 CTGGATAAAGTTCTTACTCAA 1 ||||||||||||||||||| 1881 CTGGATAAAGTTCTTACTCAA 1901 ENST00000515385
SMAD1, transcript: ENST00000394092, chr 4	DSR00114447 21 CTGGATAAAGTTCTTACTCAA 1 ||||||||||||||||||| 1579 CTGGATAAAGTTCTTACTCAA 1599 ENST00000394092
SMAD1, transcript: ENST00000302085, chr 4	DSR00114447 21 CTGGATAAAGTTCTTACTCAA 1 ||||||||||||||||||| 1762 CTGGATAAAGTTCTTACTCAA 1782 ENST00000302085
SMAD5, transcript: ENST00000545620, chr 5	DSR00114447 21 CTGGATAAAGTTCTTACTCA 2 ||||||||||| |||||||| 1624 CTGGATAAAGTCCTTACTCA 1643 ENST00000545620
SMAD5, transcript: ENST00000509297, chr 5	DSR00114447 21 CTGGATAAAGTTCTTACTCA 2 ||||||||||| |||||||| 2629 CTGGATAAAGTCCTTACTCA 2648 ENST00000509297
SMAD5, transcript: ENST00000545279, chr 5	DSR00114447 21 CTGGATAAAGTTCTTACTCA 2 ||||||||||| |||||||| 1701 CTGGATAAAGTCCTTACTCA 1720 ENST00000545279

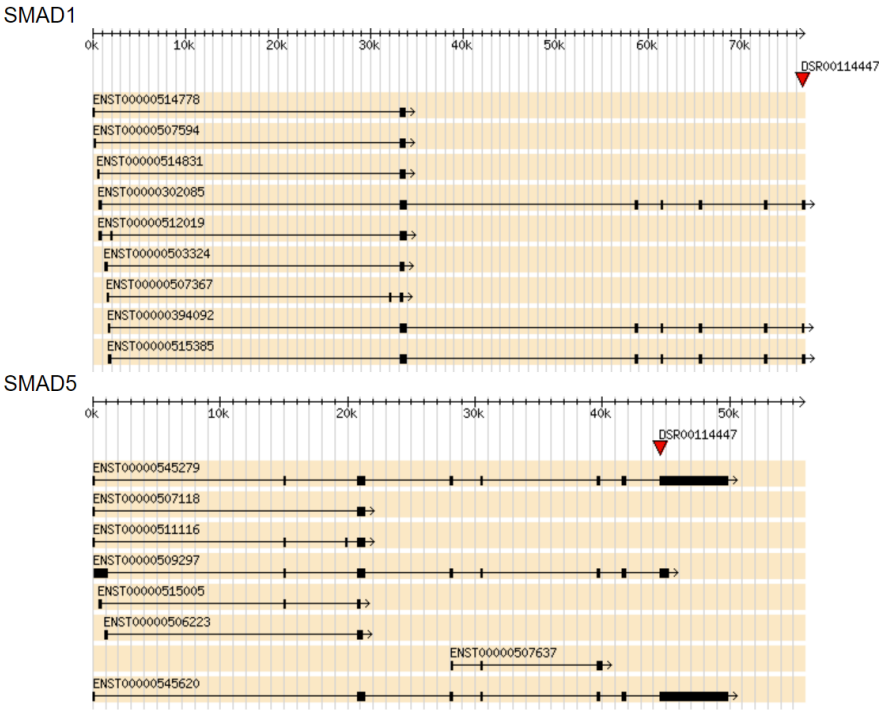

Figure 76 Off-target of second siRNA targeting SMAD1. The second siRNA used to target SMAD1 had an off-target effect on SMAD5. (source bluegecko.embl.de)

Figures for RT-qPCR Data Over Time

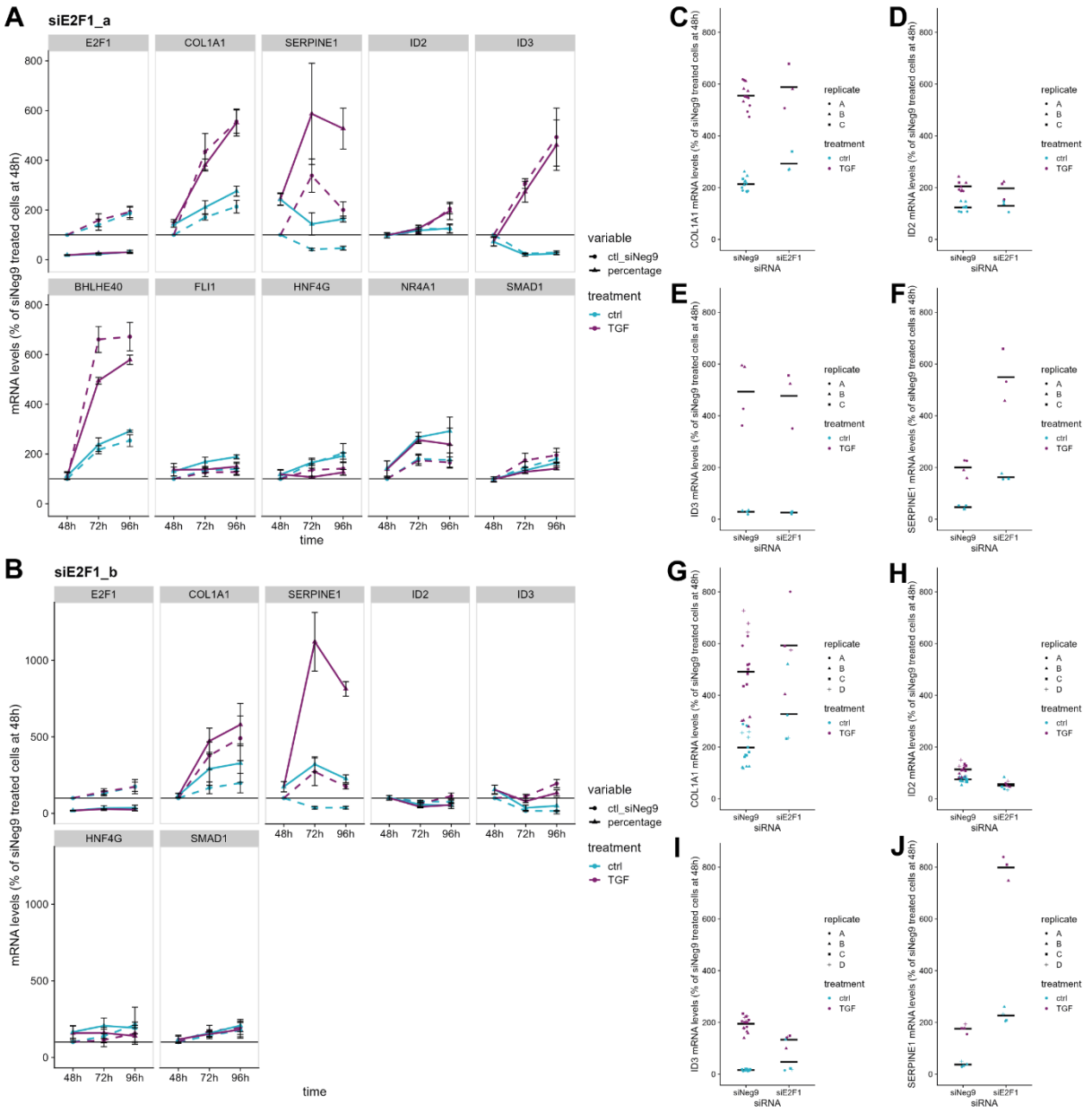

Figure 77 Change of mRNA expression levels at 48, 72 and 96 h post siE2F1 treatment. mRNA was isolated from cells transfected with two different siRNAs against E2F1 (a, b) at 48, 72, 96 h post siRNA transfection, correlating to 0, 24, 48 h treatment ± TGF-β. Expression levels of genes of interest are shown on the y-axis, that were calculated using the $2^{-\Delta\Delta CT}$ method. In short, CT values are first normalized to GADPH, then normalized to the siNeg9 ctrl at 48 h. Then the $2^{-\Delta\Delta CT}$ value is multiplied by 100 to obtain the resulting mRNA levels. Expression of genes of interest are shown over time, upon siNeg9 treatment (dashed lines), treatment with siE2F1 (continuous lines), and at control (blue) and TGF-β (violet) conditions. Overview of all read-outs is provided for **A)** the first siRNA (siE1F1_a) and the **B)** second siRNA (siE2F1_b) tested against E2F1. **C-F)** A closer look at the mRNA expression levels at the last time point (96 h post siRNA transfection) are shown for **C)** COL1A1, **D)** ID2, **E)** ID3 and **F)** SERPNE1 upon treatment with siE2F1_a. For treatment with siE2F1_b, the corresponding expression levels are highlighted in **G-J, G)** COL1A1**, H)** ID2**, I)** ID3**, J)** SERPINE1.

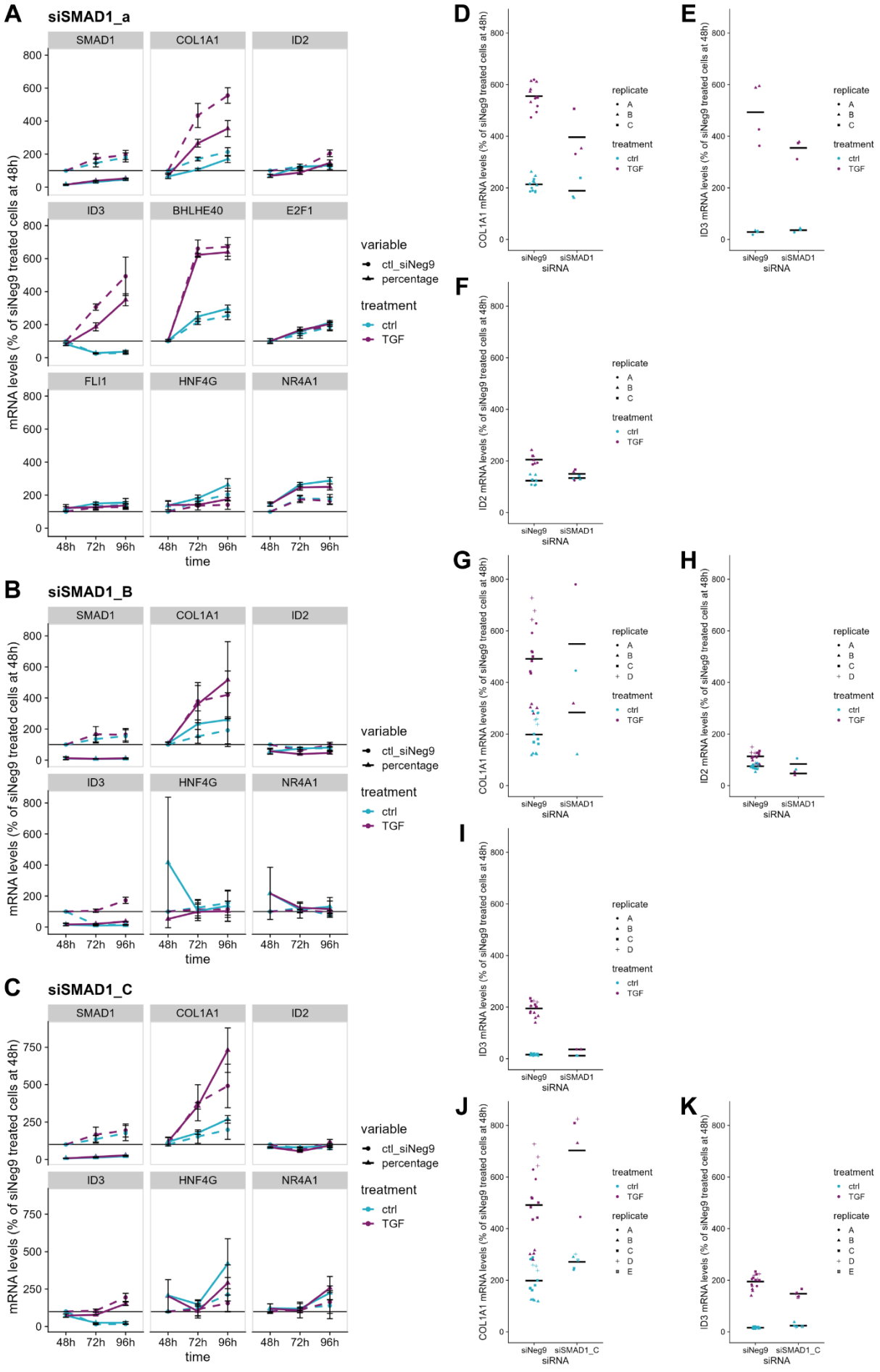
A
siSMAD1_a
SMAD1
COL1A1
ID2
ID3
BHLHE40
E2F1
FLI1
HNF4G
NR4A1
mRNA levels (% of siNeg9 treated cells at 48h)
48h 72h 96h
time
variable
ctl_siNeg9
percentage
treatment
ctrl
TGF
B
siSMAD1_B
C
siSMAD1_C
D
COL1A1 mRNA levels (% of siNeg9 treated cells at 48h)
siNeg9 siSMAD1
siRNA
replicate
A
B
C
E
ID3 mRNA levels (% of siNeg9 treated cells at 48h)
F
ID2 mRNA levels (% of siNeg9 treated cells at 48h)
G
H
I
J
siNeg9 siSMAD1_C
K

Figure 78 Change of mRNA expression levels at 48, 72 and 96 h post siSMAD1 treatment. mRNA was isolated from cells transfected with three different siRNAs against SMAD1 (a, B, C) at 48, 72, 96 h post siRNA transfection, correlating to 0, 24, 48 h treatment ± TGF-β. Expression levels of genes of interest are shown on the y-axis, that were calculated as in the previous graph. Expression of genes of interest are shown over time, upon siNeg9 treatment (dashed lines), treatment with siSMAD1 (continuous lines), and at control (blue) and TGF-β (violet) conditions. Overview of all read-outs is provided for **A)** the first siRNA (siSMAD1_a), the **B)** second siRNA (siSMAD1_B), and the **C)** third siRNA (siSMAD1_C) tested against SMAD1. **D-F)** A closer look at the mRNA expression levels at the last time point (96 h post siRNA transfection) are shown for **D)** COL1A1, **E)** ID3, and **F)** ID2 upon treatment with siSMAD1_a. For treatment with siSMAD1_B, the corresponding expression levels are highlighted in **G)** for COL1A1, **H)** for ID2, and **I)** for ID3. The expression levels of **J)** COL1A1 and **K)** ID3 are shown for the treatment with siSMAD1_C.

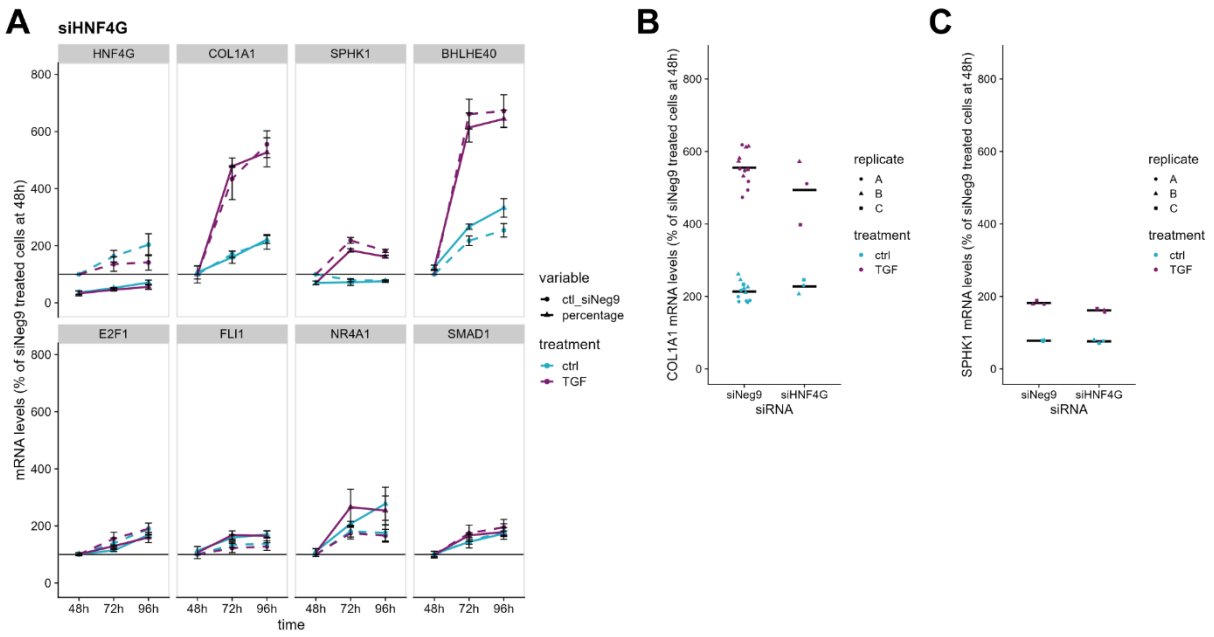

Figure 79 Change of mRNA expression levels at 48, 72 and 96 h post siHNF4G treatment. mRNA was isolated, from cells transfected with siRNAs, at 48, 72, 96 h post transfection, correlating to 0, 24, 48 h treatment ± TGF-β. Expression levels of genes of interest are shown on the y-axis (calculated as previously described) for siNeg9 treatment (dashed lines), treatment with siHNF4G (continuous lines), and at control (blue) and TGF-β (violet) conditions. **A)** Overview of all read-outs is provided. A closer look into target gene expression 96 h post siHNF4G treatment is shown for **B)** COL1A1 and **C)** SPHK1.

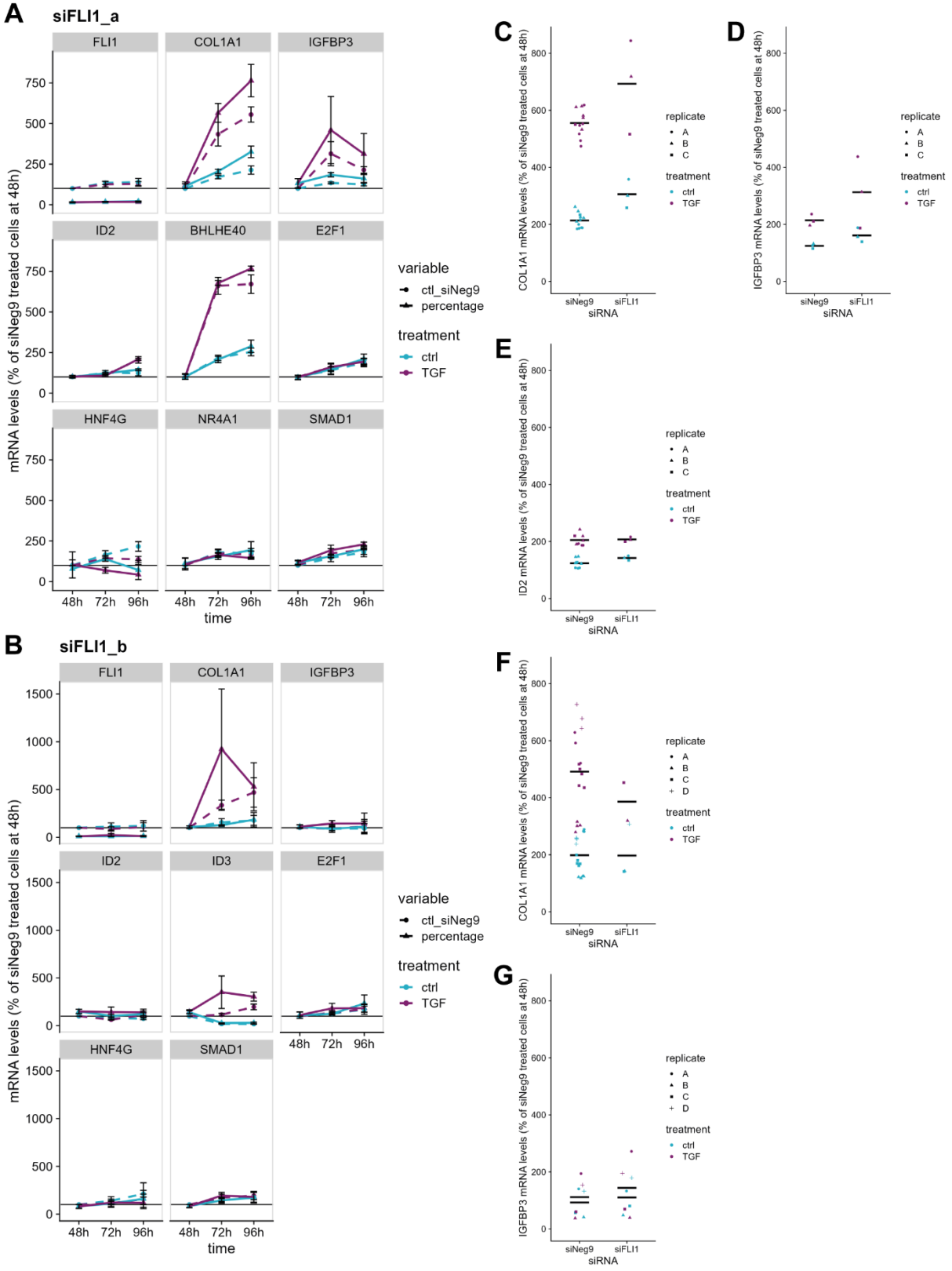

Figure 80 Change of mRNA expression levels at 48, 72 and 96 h post siFLI1 treatment. mRNA was isolated from cells transfected with two different siRNAs against FLI1 (a, b) at 48, 72, 96 h post siRNA transfection, correlating to 0, 24, 48 h treatment ± TGF-β. Expression levels of genes of interest are shown on the y-axis (calculated as previously described) for siNeg9 treatment (dashed lines), treatment with siFLI1 (continuous lines), and at control (blue) and TGF-β (violet) conditions. Overview of all read-outs is provided for **A)** the first siRNA (siFLI1_a) and the **B)** second siRNA (siFLI1_b) tested against FLI1. **C-E)** A closer look at the mRNA expression levels at the last time point (96 h post siRNA transfection) are shown for **C)** COL1A1, **D)** IGFBP3, and **E)** ID2 upon treatment with siFLI1_a. For

treatment with siFLI1_b, the corresponding expression levels are highlighted **F)** for COL1A1 and **G)** for IGFBP3.

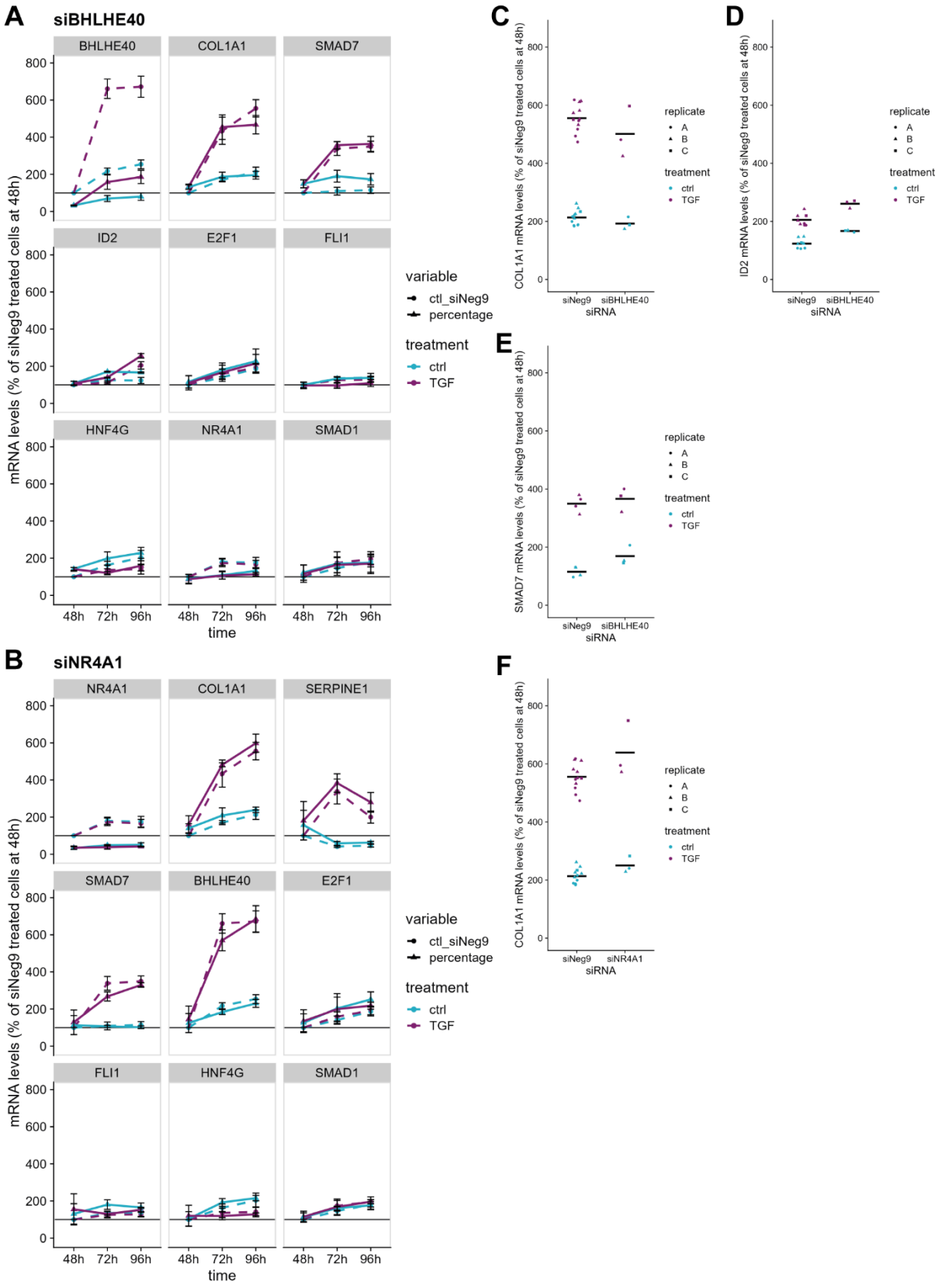

Figure 81 Change of mRNA expression levels at 48, 72 and 96 h post siBHLHE40 or siNR4A1 treatment. mRNA was isolated, from cells transfected with siRNAs, at 48, 72, 96 h post transfection, correlating to 0, 24, 48 h treatment ± TGF-β. Expression levels of genes of interest are shown on the y-

axis (calculated as previously described) for siNeg9 treatment (dashed lines), treatment with target siRNA (continuous lines), and at control (blue) and TGF-β (violet) conditions. Overview of all read-outs is provided for the treatment with **A)** siBHLHE40 and **B)** siNR4A1. Target gene expression 96 h post siBHLHE40 treatment is shown for **C)** COL1A1, **D)** ID2 and **E)** SMAD7. **F)** For siNR4A1, the expression changes of COL1A1 96 h after siRNA transfection are shown.

Anova and Post Hoc Tables for RT-qPCR Data

Table 18 Anova and post hoc results of RT-qPCR data obtained from siSMAD1 treated samples.

Anova siSMAD1a | gene COL1A1 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	40476.49	40476.49	19.74902	0.000146
treatment	1	745031.6	745031.6	363.5108	8.36E-17
siRNA:treatment	1	21530.1	21530.1	10.50482	0.003253
Residuals	26	53288.15	2049.544		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siSMAD1-siNeg9	-91.8292	-134.304	-49.3544	0.000146
TGF-ctrl	315.1786	281.1987	349.1585	3.01E-14
siSMAD1:ctrl-siNeg9:ctrl	-24.8558	-105.024	55.31192	0.829801
siNeg9:TGF-siNeg9:ctrl	341.968	291.2654	392.6705	3.11E-14
siSMAD1:TGF-siNeg9:ctrl	183.1653	102.9976	263.3331	7.08E-06
siNeg9:TGF-siSMAD1:ctrl	366.8237	286.656	446.9915	9.11E-12
siSMAD1:TGF-siSMAD1:ctrl	208.0211	106.6161	309.4262	3.66E-05
siSMAD1:TGF-siNeg9:TGF	-158.803	-238.97	-78.6349	6.04E-05

Anova siSMAD1a | gene ID3 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	2209.021	2209.021	6.836708	0.017548
treatment	1	22368.33	22368.33	69.22783	1.39E-07
siRNA:treatment	1	4664.238	4664.238	14.43537	0.001313
Residuals	18	5816.012	323.1118		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siSMAD1-siNeg9	-22.4996	-40.5781	-4.42116	0.017548
TGF-ctrl	63.77279	47.66986	79.87572	1.39E-07
siSMAD1:ctrl-siNeg9:ctrl	10.19421	-24.1999	44.5883	0.835885
siNeg9:TGF-siNeg9:ctrl	81.6058	56.20408	107.0075	2.17E-07
siSMAD1:TGF-siNeg9:ctrl	26.41231	-7.98178	60.8064	0.169376
siNeg9:TGF-siSMAD1:ctrl	71.41159	37.0175	105.8057	8.04E-05
siSMAD1:TGF-siSMAD1:ctrl	16.2181	-25.2627	57.69894	0.691104
siSMAD1:TGF-siNeg9:TGF	-55.1935	-89.5876	-20.7994	0.001335

Anova siSMAD1a | gene ID2 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	14731.95	14731.95	3.350955	0.097087
treatment	1	565130	565130	128.5455	4.97E-07

siRNA:treatment	1	18180.86	18180.86	4.135451	0.069388
Residuals	10	43963.43	4396.343		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siSMAD1-siNeg9	-65.5501	-145.337	14.23675	0.097087
TGF-ctrl	401.828	322.8594	480.7965	4.97E-07
siSMAD1:ctrl-siNeg9:ctrl	7.269873	-147.66	162.1994	0.998871
siNeg9:TGF-siNeg9:ctrl	464.2451	320.8083	607.6819	8.70E-06
siSMAD1:TGF-siNeg9:ctrl	325.875	170.9455	480.8045	0.000363
siNeg9:TGF-siSMAD1:ctrl	456.9752	302.0457	611.9047	2.01E-05
siSMAD1:TGF-siSMAD1:ctrl	318.6051	152.9785	484.2317	0.00074
siSMAD1:TGF-siNeg9:TGF	-138.37	-293.3	16.55939	0.083822

Anova siSMAD1b | gene COL1A1 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	18213.38	18213.38	1.064127	0.310521
treatment	1	713919.4	713919.4	41.71113	3.90E-07
siRNA:treatment	1	658.7527	658.7527	0.038488	0.845789
Residuals	30	513474	17115.8		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siSMAD1-siNeg9	71.83633	-70.3838	214.0565	0.310521
TGF-ctrl	289.8112	198.1674	381.4549	3.90E-07
siSMAD1:ctrl-siNeg9:ctrl	85.4982	-182.289	353.285	0.821121
siNeg9:TGF-siNeg9:ctrl	293.0257	163.1301	422.9214	5.54E-06
siSMAD1:TGF-siNeg9:ctrl	351.2002	83.41343	618.987	0.00644
siNeg9:TGF-siSMAD1:ctrl	207.5275	-60.2592	475.3143	0.173917
siSMAD1:TGF-siSMAD1:ctrl	265.702	-90.0319	621.4359	0.199439
siSMAD1:TGF-siNeg9:TGF	58.17446	-209.612	325.9612	0.934112

Anova siSMAD1b | gene ID2 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	2940.642	2940.642	11.18201	0.002228
treatment	1	7391.357	7391.357	28.10619	9.94E-06
siRNA:treatment	1	4990.959	4990.959	18.97849	0.000142
Residuals	30	7889.392	262.9797		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siSMAD1-siNeg9	-28.8649	-46.4937	-11.2361	0.002228
TGF-ctrl	29.4885	18.12884	40.84816	9.94E-06
siSMAD1:ctrl-siNeg9:ctrl	8.739697	-24.4537	41.93309	0.889983
siNeg9:TGF-siNeg9:ctrl	38.33664	22.23547	54.4378	2.16E-06
siSMAD1:TGF-siNeg9:ctrl	-28.1328	-61.3262	5.060548	0.119581
siNeg9:TGF-siSMAD1:ctrl	29.59694	-3.59645	62.79033	0.09396
siSMAD1:TGF-siSMAD1:ctrl	-36.8725	-80.9674	7.222301	0.127022
siSMAD1:TGF-siNeg9:TGF	-66.4695	-99.6629	-33.2761	3.78E-05

Anova siSMAD1b | gene ID3 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	23353.08	23353.08	67.65291	3.51E-09
treatment	1	219149.4	219149.4	634.8668	9.66E-22
siRNA:treatment	1	21029.11	21029.11	60.92045	1.04E-08
Residuals	30	10355.69	345.1896		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siSMAD1-siNeg9	-81.3431	-101.54	-61.1459	3.51E-09
TGF-ctrl	160.5686	147.554	173.5833	4.61E-14
siSMAD1:ctrl-siNeg9:ctrl	-4.15346	-42.1829	33.87593	0.990703
siNeg9:TGF-siNeg9:ctrl	178.7309	160.2839	197.1779	4.61E-14
siSMAD1:TGF-siNeg9:ctrl	20.19809	-17.8313	58.22748	0.482754
siNeg9:TGF-siSMAD1:ctrl	182.8844	144.855	220.9138	4.29E-13
siSMAD1:TGF-siSMAD1:ctrl	24.35155	-26.1675	74.87064	0.563516
siSMAD1:TGF-siNeg9:TGF	-158.533	-196.562	-120.503	1.36E-11

Anova siSMAD1c | gene COL1A1 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	196765.8	196765.8	14.69945	0.000488
treatment	1	1139339	1139339	85.1147	5.11E-11
siRNA:treatment	1	59448.25	59448.25	4.441101	0.042114
Residuals	36	481893.4	13385.93		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siSMAD1_C-siNeg9	161.9736	76.29315	247.654	0.000488
TGF-ctrl	337.541	263.3396	411.7424	5.09E-11
siSMAD1_C:ctrl-siNeg9:ctrl	72.94306	-87.9664	233.8525	0.617913
siNeg9:TGF-siNeg9:ctrl	293.0257	179.2456	406.8059	2.34E-07
siSMAD1_C:TGF-siNeg9:ctrl	544.0298	383.1204	704.9393	4.24E-10
siNeg9:TGF-siSMAD1_C:ctrl	220.0827	59.1732	380.9921	0.004003
siSMAD1_C:TGF-siSMAD1_C:ctrl	471.0868	274.0137	668.1598	1.06E-06
siSMAD1_C:TGF-siNeg9:TGF	251.0041	90.09465	411.9136	0.000922

Anova siSMAD1c | gene ID3 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	2887.847	2887.847	9.22323	0.004426
treatment	1	271945.4	271945.4	868.5415	8.45E-27
siRNA:treatment	1	5732.47	5732.47	18.30841	0.000133
Residuals	36	11271.81	313.1058		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siSMAD1_C-siNeg9	-19.6226	-32.7266	-6.51862	0.004426
TGF-ctrl	164.9077	153.5593	176.256	0
siSMAD1_C:ctrl-siNeg9:ctrl	8.023913	-16.5856	32.63342	0.816111

siNeg9:TGF-siNeg9:ctrl	178.7309	161.3294	196.1325	0
siSMAD1_C:TGF-siNeg9:ctrl	131.4618	106.8523	156.0713	0
siNeg9:TGF-siSMAD1_C:ctrl	170.707	146.0975	195.3165	0
siSMAD1_C:TGF-siSMAD1_C:ctrl	123.4379	93.29753	153.5783	2.29E-12
siSMAD1_C:TGF-siNeg9:TGF	-47.2691	-71.8786	-22.6596	5.07E-05

Table 19 Anova and post hoc results of RT-qPCR data obtained from siE2F1 treated samples.

Anova siE2F1a | gene COL1A1 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	15271.16	15271.16	8.028083	0.008785
treatment	1	830068.2	830068.2	436.3685	8.95E-18
siRNA:treatment	1	2588.172	2588.172	1.360608	0.25402
Residuals	26	49457.67	1902.218		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siE2F1-siNeg9	56.40472	15.48495	97.3245	0.008785
TGF-ctrl	332.6797	299.9438	365.4155	3.00E-14
siE2F1:ctrl-siNeg9:ctrl	79.62546	2.392799	156.8581	0.041514
siNeg9:TGF-siNeg9:ctrl	341.968	293.1217	390.8142	3.04E-14
siE2F1:TGF-siNeg9:ctrl	375.1519	297.9193	452.3846	2.38E-12
siNeg9:TGF-siE2F1:ctrl	262.3425	185.1098	339.5752	5.28E-09
siE2F1:TGF-siE2F1:ctrl	295.5265	197.834	393.2189	5.14E-08
siE2F1:TGF-siNeg9:TGF	33.18399	-44.0487	110.4166	0.645285

Anova siE2F1a | gene ID2 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	1.876218	1.876218	0.00396	0.950517
treatment	1	33234.56	33234.56	70.1468	1.27E-07
siRNA:treatment	1	219.8132	219.8132	0.463951	0.504453
Residuals	18	8528.143	473.7857		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siE2F1-siNeg9	-0.65572	-22.5473	21.23584	0.950517
TGF-ctrl	77.73446	58.23514	97.23379	1.27E-07
siE2F1:ctrl-siNeg9:ctrl	6.441736	-35.2067	48.09015	0.971239
siNeg9:TGF-siNeg9:ctrl	81.6058	50.84641	112.3652	3.39E-06
siE2F1:TGF-siNeg9:ctrl	73.85263	32.20422	115.501	0.000481
siNeg9:TGF-siE2F1:ctrl	75.16407	33.51565	116.8125	0.000398
siE2F1:TGF-siE2F1:ctrl	67.41089	17.18102	117.6408	0.006641
siE2F1:TGF-siNeg9:TGF	-7.75317	-49.4016	33.89524	0.95164

Anova siE2F1a | gene ID3 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	319.4707	319.4707	0.048944	0.829364
treatment	1	736461.9	736461.9	112.8278	9.12E-07
siRNA:treatment	1	142.8153	142.8153	0.02188	0.885348
Residuals	10	65273.09	6527.309		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siE2F1-siNeg9	-9.65292	-106.872	87.56643	0.829364
TGF-ctrl	458.7131	362.4909	554.9353	9.11E-07
siE2F1:ctrl-siNeg9:ctrl	-3.1989	-191.979	185.5809	0.999946
siNeg9:TGF-siNeg9:ctrl	464.2451	289.469	639.0212	5.07E-05
siE2F1:TGF-siNeg9:ctrl	448.1382	259.3584	636.9179	0.000133
siNeg9:TGF-siE2F1:ctrl	467.444	278.6642	656.2238	9.31E-05
siE2F1:TGF-siE2F1:ctrl	451.3371	249.523	653.1511	0.00022
siE2F1:TGF-siNeg9:TGF	-16.107	-204.887	172.6728	0.993369

Anova siE2F1a | gene SERPINE1 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	185069	185069	76.01301	5.50E-06
treatment	1	225101.4	225101.4	92.45545	2.27E-06
siRNA:treatment	1	46511.79	46511.79	19.10369	0.001397
Residuals	10	24347.01	2434.701		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siE2F1-siNeg9	232.3326	172.957	291.7083	5.50E-06
TGF-ctrl	253.6034	194.8367	312.3701	2.27E-06
siE2F1:ctrl-siNeg9:ctrl	115.8597	0.564529	231.1549	0.048823
siNeg9:TGF-siNeg9:ctrl	153.7695	47.02688	260.5121	0.006053
siE2F1:TGF-siNeg9:ctrl	502.575	387.2798	617.8702	5.41E-07
siNeg9:TGF-siE2F1:ctrl	37.90977	-77.3854	153.205	0.749707
siE2F1:TGF-siE2F1:ctrl	386.7153	263.4596	509.971	1.15E-05
siE2F1:TGF-siNeg9:TGF	348.8055	233.5103	464.1007	1.60E-05

Anova siE2F1b | gene COL1A1 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	84138.95	84138.95	5.858727	0.020991
treatment	1	782741.4	782741.4	54.50351	1.47E-08
siRNA:treatment	1	1275.014	1275.014	0.088781	0.767546
Residuals	34	488284.3	14361.3		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siE2F1-siNeg9	115.421	18.5131	212.3288	0.020991
TGF-ctrl	287.0433	208.028	366.0585	1.47E-08
siE2F1:ctrl-siNeg9:ctrl	129.6293	-52.5046	311.7633	0.237974
siNeg9:TGF-siNeg9:ctrl	293.0257	174.8416	411.2098	6.37E-07
siE2F1:TGF-siNeg9:ctrl	394.2383	212.1044	576.3723	7.90E-06
siNeg9:TGF-siE2F1:ctrl	163.3964	-18.7376	345.5304	0.091758
siE2F1:TGF-siE2F1:ctrl	264.609	35.74646	493.4716	0.018248
siE2F1:TGF-siNeg9:TGF	101.2126	-80.9213	283.3466	0.448142

Anova siE2F1b | gene ID2 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	10411.09	10411.09	40.73559	2.76E-07
treatment	1	9444.917	9444.917	36.95522	6.80E-07
siRNA:treatment	1	1650.048	1650.048	6.456159	0.01579
Residuals	34	8689.629	255.5773		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siE2F1-siNeg9	-40.6008	-53.5285	-27.673	2.76E-07
TGF-ctrl	31.53097	20.99012	42.07181	6.80E-07
siE2F1:ctrl-siNeg9:ctrl	-24.4373	-48.7345	-0.14017	0.048257
siNeg9:TGF-siNeg9:ctrl	38.33664	22.57057	54.1027	9.32E-07
siE2F1:TGF-siNeg9:ctrl	-18.4276	-42.7247	5.869536	0.190771
siNeg9:TGF-siE2F1:ctrl	62.77395	38.4768	87.07109	2.80E-07
siE2F1:TGF-siE2F1:ctrl	6.009705	-24.5212	36.54056	0.95077
siE2F1:TGF-siNeg9:TGF	-56.7642	-81.0614	-32.4671	1.99E-06

Anova siE2F1b | gene ID3 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	1432.704	1432.704	2.218128	0.145617
treatment	1	240837.8	240837.8	372.8676	6.69E-20
siRNA:treatment	1	13560.08	13560.08	20.99385	5.94E-05
Residuals	34	21960.84	645.9071		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siE2F1-siNeg9	-15.0614	-35.613	5.490303	0.145617
TGF-ctrl	159.2211	142.464	175.9782	0
siE2F1:ctrl-siNeg9:ctrl	31.27449	-7.35145	69.90043	0.147411
siNeg9:TGF-siNeg9:ctrl	178.7309	153.6671	203.7947	0
siE2F1:TGF-siNeg9:ctrl	117.3337	78.70774	155.9596	8.44E-09
siNeg9:TGF-siE2F1:ctrl	147.4564	108.8305	186.0824	3.15E-11
siE2F1:TGF-siE2F1:ctrl	86.05918	37.52331	134.5951	0.000182
siE2F1:TGF-siNeg9:TGF	-61.3972	-100.023	-22.7713	0.000767

Anova siE2F1b | gene SERPINE1 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	682433.4	682433.4	853.0213	1.62E-12
treatment	1	524904.1	524904.1	656.1144	7.63E-12
siRNA:treatment	1	200359.4	200359.4	250.4432	2.10E-09
Residuals	12	9600.23	800.0192		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siE2F1-siNeg9	413.0476	382.2342	443.8611	9.37E-13
TGF-ctrl	362.2513	331.4379	393.0648	4.72E-12
siE2F1:ctrl-siNeg9:ctrl	189.24	129.8613	248.6188	3.40E-06
siNeg9:TGF-siNeg9:ctrl	138.4437	79.06499	197.8225	8.20E-05
siE2F1:TGF-siNeg9:ctrl	775.299	715.9202	834.6777	5.54E-13

siNeg9:TGF-siE2F1:ctrl	-50.7963	-110.175	8.582432	0.103441
siE2F1:TGF-siE2F1:ctrl	586.0589	526.6802	645.4377	4.14E-12
siE2F1:TGF-siNeg9:TGF	636.8552	577.4765	696.234	2.15E-12

Table 20 Anova and post hoc results of RT-qPCR data obtained from siFLI1 treated samples.

Anova siFLI1a | gene COL1A1 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	63182.47	63182.47	17.99572	0.000248
treatment	1	924277.2	924277.2	263.2539	4.05E-15
siRNA:treatment	1	2475.171	2475.171	0.704982	0.408769
Residuals	26	91285.28	3510.972		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siFLI1-siNeg9	114.7302	59.13761	170.3228	0.000248
TGF-ctrl	351.0512	306.5772	395.5253	3.43E-14
siFLI1:ctrl-siNeg9:ctrl	92.02202	-12.9043	196.9484	0.100874
siNeg9:TGF-siNeg9:ctrl	341.968	275.6067	408.3292	6.30E-13
siFLI1:TGF-siNeg9:ctrl	479.4063	374.48	584.3326	9.42E-12
siNeg9:TGF-siFLI1:ctrl	249.9459	145.0196	354.8723	3.61E-06
siFLI1:TGF-siFLI1:ctrl	387.3843	254.6618	520.1068	1.01E-07
siFLI1:TGF-siNeg9:TGF	137.4383	32.512	242.3647	0.006873

Anova siFLI1a | gene ID2 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	467.6677	467.6677	1.681409	0.211115
treatment	1	32703.29	32703.29	117.5784	2.53E-09
siRNA:treatment	1	296.3599	296.3599	1.065505	0.315634
Residuals	18	5006.527	278.1404		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siFLI1-siNeg9	10.35248	-6.42079	27.12575	0.211115
TGF-ctrl	77.11065	62.17031	92.051	2.53E-09
siFLI1:ctrl-siNeg9:ctrl	18.59359	-13.3173	50.50452	0.379023
siNeg9:TGF-siNeg9:ctrl	81.6058	58.03802	105.1736	7.03E-08
siFLI1:TGF-siNeg9:ctrl	83.71718	51.80625	115.6281	3.95E-06
siNeg9:TGF-siFLI1:ctrl	63.01222	31.10128	94.92315	0.000145
siFLI1:TGF-siFLI1:ctrl	65.12359	26.63756	103.6096	0.000785
siFLI1:TGF-siNeg9:TGF	2.111376	-29.7996	34.02231	0.997592

Anova siFLI1a | gene IGFBP3 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	13722.84	13722.84	3.27228	0.108065
treatment	1	43733.44	43733.44	10.42846	0.012069
siRNA:treatment	1	2885.706	2885.706	0.688111	0.430862
Residuals	8	33549.3	4193.662		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siFLI1-siNeg9	67.63342	-18.5842	153.8511	0.108065
TGF-ctrl	120.7386	34.52097	206.9563	0.012069
siFLI1:ctrl-siNeg9:ctrl	36.61887	-132.706	205.9435	0.897123
siNeg9:TGF-siNeg9:ctrl	89.72407	-79.6006	259.0487	0.38444
siFLI1:TGF-siNeg9:ctrl	188.372	19.0474	357.6967	0.030233
siNeg9:TGF-siFLI1:ctrl	53.1052	-116.219	222.4298	0.751523
siFLI1:TGF-siFLI1:ctrl	151.7532	-17.5715	321.0778	0.079922
siFLI1:TGF-siNeg9:TGF	98.64796	-70.6767	267.9726	0.313215

Anova siFLI1b | gene COL1A1 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	7260.288	7260.288	0.399652	0.53176
treatment	1	843253	843253	46.41795	1.05E-07
siRNA:treatment	1	7687.466	7687.466	0.423166	0.520005
Residuals	32	581328.9	18166.53		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siFLI1-siNeg9	38.10587	-84.6742	160.8859	0.53176
TGF-ctrl	306.096	214.5812	397.6109	1.05E-07
siFLI1:ctrl-siNeg9:ctrl	-1.10501	-232.063	229.853	0.999999
siNeg9:TGF-siNeg9:ctrl	293.0257	159.6821	426.3694	7.17E-06
siFLI1:TGF-siNeg9:ctrl	370.3425	139.3845	601.3005	0.000727
siNeg9:TGF-siFLI1:ctrl	294.1307	63.17277	525.0887	0.008232
siFLI1:TGF-siFLI1:ctrl	371.4475	73.28203	669.6129	0.009992
siFLI1:TGF-siNeg9:TGF	77.31675	-153.641	308.2747	0.801223

Anova siFLI1b | gene IGFBP3 | 96 h

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	2485.909	2485.909	0.425354	0.526573
treatment	1	2762.86	2762.86	0.472742	0.504801
siRNA:treatment	1	224.5693	224.5693	0.038425	0.847872
Residuals	12	70131.92	5844.327		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siFLI1-siNeg9	24.92945	-58.3537	108.2126	0.526573
TGF-ctrl	26.28146	-57.0017	109.5646	0.504801
siFLI1:ctrl-siNeg9:ctrl	17.43663	-143.053	177.9266	0.987814
siNeg9:TGF-siNeg9:ctrl	18.78864	-141.701	179.2786	0.98487
siFLI1:TGF-siNeg9:ctrl	51.21091	-109.279	211.7009	0.78061
siNeg9:TGF-siFLI1:ctrl	1.352014	-159.138	161.842	0.999994
siFLI1:TGF-siFLI1:ctrl	33.77428	-126.716	194.2643	0.922124
siFLI1:TGF-siNeg9:TGF	32.42226	-128.068	192.9123	0.930173

Table 21 Anova and post hoc results of RT-qPCR data obtained from siBHLHE40 treated samples.

Anova siBHLHE40 | gene COL1A1 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	6845.935	6845.935	3.730343	0.064402
treatment	1	843121.3	843121.3	459.416	4.75E-18
siRNA:treatment	1	1339.775	1339.775	0.730042	0.400674
Residuals	26	47715.26	1835.202		
Post hoc tukey multiple comparisons of means, 95% family-wise confidence level					
	diff	lwr	upr	p adj	
siBHLHE40-siNeg9	-37.7655	-77.9581	2.426957	0.064402	
TGF-ctrl	335.2852	303.1312	367.4392	3.00E-14	
siBHLHE40:ctrl-siNeg9:ctrl	-21.0587	-96.9187	54.80133	0.870888	
siNeg9:TGF-siNeg9:ctrl	341.968	293.9899	389.946	3.02E-14	
siBHLHE40:TGF-siNeg9:ctrl	287.4955	211.6355	363.3555	5.50E-10	
siNeg9:TGF-siBHLHE40:ctrl	363.0266	287.1666	438.8866	3.33E-12	
siBHLHE40:TGF-siBHLHE40:ctrl	308.5542	212.5981	404.5103	1.58E-08	
siBHLHE40:TGF-siNeg9:TGF	-54.4724	-130.332	21.38757	0.225057	

Anova siBHLHE40 | gene ID2 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	10638.19	10638.19	37.07685	9.40E-06
treatment	1	39736.06	39736.06	138.4905	6.90E-10
siRNA:treatment	1	168.8152	168.8152	0.588365	0.452996
Residuals	18	5164.607	286.9226		
Post hoc tukey multiple comparisons of means, 95% family-wise confidence level					
	diff	lwr	upr	p adj	
siBHLHE40-siNeg9	49.37528	32.33926	66.41129	9.40E-06	
TGF-ctrl	84.99846	69.82408	100.1728	6.90E-10	
siBHLHE40:ctrl-siNeg9:ctrl	43.15541	10.7446	75.56621	0.007081	
siNeg9:TGF-siNeg9:ctrl	81.6058	57.66883	105.5428	8.90E-08	
siBHLHE40:TGF-siNeg9:ctrl	137.201	104.7901	169.6118	3.01E-09	
siNeg9:TGF-siBHLHE40:ctrl	38.4504	6.039589	70.8612	0.016968	
siBHLHE40:TGF-siBHLHE40:ctrl	94.04555	54.95664	133.1345	1.26E-05	
siBHLHE40:TGF-siNeg9:TGF	55.59515	23.18435	88.00596	0.000682	

Anova siBHLHE40 | gene SMAD7 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	4235.129	4235.129	4.715241	0.055034
treatment	1	167079.4	167079.4	186.0202	8.69E-08
siRNA:treatment	1	1171.53	1171.53	1.304339	0.280031
Residuals	10	8981.788	898.1788		
Post hoc tukey multiple comparisons of means, 95% family-wise confidence level					
	diff	lwr	upr	p adj	
siBHLHE40-siNeg9	35.14607	-0.91736	71.2095	0.055034	

TGF-ctrl	218.4879	182.7944	254.1814	8.73E-08
siBHLHE40:ctrl-siNeg9:ctrl	53.6311	-16.3966	123.6588	0.152464
siNeg9:TGF-siNeg9:ctrl	234.3322	169.4992	299.1653	3.15E-06
siBHLHE40:TGF-siNeg9:ctrl	250.9933	180.9656	321.0209	3.40E-06
siNeg9:TGF-siBHLHE40:ctrl	180.7011	110.6734	250.7288	6.53E-05
siBHLHE40:TGF-siBHLHE40:ctrl	197.3622	122.4994	272.2249	5.42E-05
siBHLHE40:TGF-siNeg9:TGF	16.66104	-53.3666	86.68872	0.883755

Table 22 Anova and post hoc results of RT-qPCR data obtained from siHNF4G treated samples.

Anova siHNF4G | gene COL1A1 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	2714.393	2714.393	1.474852	0.235497
treatment	1	800905.1	800905.1	435.1678	9.26E-18
siRNA:treatment	1	6917.182	6917.182	3.758417	0.06347
Residuals	26	47851.73	1840.451		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siHNF4G-siNeg9	-23.7802	-64.0302	16.46972	0.235497
TGF-ctrl	326.7833	294.5834	358.9833	3.00E-14
siHNF4G:ctrl-siNeg9:ctrl	14.18134	-61.7871	90.14974	0.955462
siNeg9:TGF-siNeg9:ctrl	341.968	293.9213	390.0146	3.02E-14
siHNF4G:TGF-siNeg9:ctrl	280.2262	204.2578	356.1946	9.70E-10
siNeg9:TGF-siHNF4G:ctrl	327.7866	251.8182	403.755	3.35E-11
siHNF4G:TGF-siHNF4G:ctrl	266.0448	169.9516	362.1381	2.68E-07
siHNF4G:TGF-siNeg9:TGF	-61.7418	-137.71	14.22663	0.141766

Anova siHNF4G | gene SPHK1 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	366.6757	366.6757	18.04778	0.002806
treatment	1	26984.94	26984.94	1328.199	3.52E-10
siRNA:treatment	1	264.0235	264.0235	12.99524	0.006933
Residuals	8	162.5356	20.31694		

Post hoc tukey multiple comparisons of means, 95% family-wise confidence level

	diff	lwr	upr	p adj
siHNF4G-siNeg9	-11.0556	-17.0566	-5.05448	0.002806
TGF-ctrl	94.84187	88.84081	100.8429	6.73E-10
siHNF4G:ctrl-siNeg9:ctrl	-1.6743	-13.4599	10.11132	0.966774
siNeg9:TGF-siNeg9:ctrl	104.2231	92.4375	116.0087	1.44E-08
siHNF4G:TGF-siNeg9:ctrl	83.78632	72.0007	95.57195	5.12E-08
siNeg9:TGF-siHNF4G:ctrl	105.8974	94.1118	117.6831	1.33E-08
siHNF4G:TGF-siHNF4G:ctrl	85.46063	73.675	97.24625	4.40E-08
siHNF4G:TGF-siNeg9:TGF	-20.4368	-32.2224	-8.65118	0.002402

Table 23 Anova and post hoc results of RT-qPCR data obtained from siNR4A1 treated samples.

Anova siNR4A1 | gene COL1A1 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	17451.86	17451.86	8.795745	0.006395
treatment	1	925157.9	925157.9	466.2801	3.95E-18
siRNA:treatment	1	2567.137	2567.137	1.293838	0.265716
Residuals	26	51587.25	1984.125		
Post hoc tukey multiple comparisons of means, 95% family-wise confidence level					
	diff	lwr	upr	p adj	
siNR4A1-siNeg9	60.29762	18.50616	102.0891	0.006395	
TGF-ctrl	351.2184	317.7853	384.6516	3.00E-14	
siNR4A1:ctrl-siNeg9:ctrl	37.17144	-41.7065	116.0493	0.575543	
siNeg9:TGF-siNeg9:ctrl	341.968	292.0812	391.8547	3.06E-14	
siNR4A1:TGF-siNeg9:ctrl	425.3918	346.5139	504.2697	2.38E-13	
siNeg9:TGF-siNR4A1:ctrl	304.7965	225.9186	383.6744	3.64E-10	
siNR4A1:TGF-siNR4A1:ctrl	388.2203	288.4468	487.9938	3.14E-10	
siNR4A1:TGF-siNeg9:TGF	83.4238	4.545907	162.3017	0.035283	

Table 24 Anova and post hoc results of RT-qPCR data obtained from siFN1 treated samples (n=2).

Anova siFN1 | gene COL1A1 | 96 h					
	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
siRNA	1	34331	34331	1.119	0.3497
treatment	1	346017	346017	11.281	0.0283
siRNA:treatment	1	577	577	0.019	0.8976
Residuals	4	122687	30672		

CellProfiler Pipeline

CellProfiler Pipeline: http://www.cellprofiler.org
Version:5
DateRevision:426
GitHash:
ModuleCount:11
HasImagePlaneDetails:False

Images:[module_num:1|svn_version:'Unknown'|variable_revision_number:2|show_window:False|notes:['To begin creating your project, use the Images module to compile a list of files and/or folders that you want to analyze. You can also specify a set of rules to include only the desired files in your selected folders.']|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
    :
    Filter images?:Custom
    Select the rule criteria:and (extension does istif)

Metadata:[module_num:2|svn_version:'Unknown'|variable_revision_number:6|show_window:False|notes:['The Metadata module optionally allows you to extract information describing your images (i.e, metadata) which will be stored along with your measurements. This information can be contained in the file name and/or location, or in an external file.']|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
    Extract metadata?:Yes
    Metadata data type:Text
    Metadata types:{}
    Extraction method count:3
    Metadata extraction method:Extract from file/folder names
    Metadata source:File name
    Regular expression to extract from file name:--W(?P<WellNum>.*)--P(?P<PosNum>.*)--.*--.*--.*
    Regular expression to extract from folder name:(?P<Date>[0-9]{4}_[0-9]{2}_[0-9]{2})$
    Extract metadata from:All images
    Select the filtering criteria:and (file does contain "")
    Metadata file location:Elsewhere...|
    Match file and image metadata:[]
    Use case insensitive matching?:No
    Metadata file name:None
    Does cached metadata exist?:No
    Metadata extraction method:Extract from file/folder names
    Metadata source:File name
    Regular expression to extract from file name:^(?P<RootImageName>.*)--T.*$
    Regular expression to extract from folder name:(?P<Date>[0-9]{4}_[0-9]{2}_[0-9]{2})$
    Extract metadata from:All images
    Select the filtering criteria:and (file does contain "")
    Metadata file location:Elsewhere...|
    Match file and image metadata:[]
    Use case insensitive matching?:No

Metadata file name:None
Does cached metadata exist?:No
Metadata extraction method:Extract from file/folder names
Metadata source:Folder name
Regular expression to extract from file name:^(?P<Plate>.*)_(?P<Well>[A-P][0-9]{2})_s(?P<Site>[0-9])_w(?P<ChannelNumber>[0-9])
Regular expression to extract from folder name:(?P<RootFolder>.*)[\\/](?P<Year>.*)[\\/](?P<Experiment>.*)[\\/](?P<Date>.*)[\\/](?P<Well>.*)[\\/](?P<Position>.*)$
Extract metadata from:All images
Select the filtering criteria:and (file does contain "")
Metadata file location:Elsewhere...|
Match file and image metadata:[]
Use case insensitive matching?:No
Metadata file name:None
Does cached metadata exist?:No

NamesAndTypes:[module_num:3|svn_version:'Unknown'|variable_revision_number:8|show_window:False|notes:['The NamesAndTypes module allows you to assign a meaningful name to each image by which other modules will refer to it.']|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
Assign a name to:Images matching rules
Select the image type:Grayscale image
Name to assign these images:DNA
Match metadata:[]
Image set matching method:Order
Set intensity range from:Image metadata
Assignments count:2
Single images count:0
Maximum intensity:255.0
Process as 3D?:No
Relative pixel spacing in X:1.0
Relative pixel spacing in Y:1.0
Relative pixel spacing in Z:1.0
Select the rule criteria:and (file does contain "C01.ome.tif")
Name to assign these images:Nuclei
Name to assign these objects:Cell
Select the image type:Grayscale image
Set intensity range from:Image metadata
Maximum intensity:255.0
Select the rule criteria:and (file does contain "C02.ome.tif")
Name to assign these images:Collagen
Name to assign these objects:Nucleus
Select the image type:Grayscale image
Set intensity range from:Image metadata
Maximum intensity:255.0

Groups:[module_num:4|svn_version:'Unknown'|variable_revision_number:2|show_window:False|notes:['The Groups module optionally allows you to split your list of images into image subsets (groups) which will be processed independently of each

other. Examples of groupings include screening batches, microtiter plates, time-lapse movies, etc.']|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
    Do you want to group your images?:No
    grouping metadata count:1
    Metadata category:None

IdentifyPrimaryObjects:[module_num:5|svn_version:'Unknown'|variable_revision_number:15|show_window:False|notes:[]|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
    Select the input image:Nuclei
    Name the primary objects to be identified:Nuclei
    Typical diameter of objects, in pixel units (Min,Max):30,200
    Discard objects outside the diameter range?:Yes
    Discard objects touching the border of the image?:Yes
    Method to distinguish clumped objects:Shape
    Method to draw dividing lines between clumped objects:Intensity
    Size of smoothing filter:0
    Suppress local maxima that are closer than this minimum allowed distance:10
    Speed up by using lower-resolution image to find local maxima?:No
    Fill holes in identified objects?:After both thresholding and declumping
    Automatically calculate size of smoothing filter for declumping?:No
    Automatically calculate minimum allowed distance between local maxima?:No
    Handling of objects if excessive number of objects identified:Continue
    Maximum number of objects:500
    Use advanced settings?:Yes
    Threshold setting version:12
    Threshold strategy:Global
    Thresholding method:Minimum Cross-Entropy
    Threshold smoothing scale:1.0
    Threshold correction factor:1.0
    Lower and upper bounds on threshold:0.0,1.0
    Manual threshold:0.0
    Select the measurement to threshold with:None
    Two-class or three-class thresholding?:Two classes
    Log transform before thresholding?:No
    Assign pixels in the middle intensity class to the foreground or the background?:Foreground
    Size of adaptive window:50
    Lower outlier fraction:0.05
    Upper outlier fraction:0.05
    Averaging method:Mean
    Variance method:Standard deviation
    # of deviations:2.0
    Thresholding method:Minimum Cross-Entropy

ConvertObjectsToImage:[module_num:6|svn_version:'Unknown'|variable_revision_number:1|show_window:False|notes:[]|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:True]
    Select the input objects:Nuclei
    Name the output image:NucleiObjectImage
    Select the color format:Color

Select the colormap:Default

SaveImages:[module_num:7|svn_version:'Unknown'|variable_revision_number:16|show_window:False|notes:[]|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
Select the type of image to save:Image
Select the image to save:NucleiObjectImage
Select method for constructing file names:From image filename
Select image name for file prefix:Nuclei
Enter single file name:OrigBlue
Number of digits:4
Append a suffix to the image file name?:No
Text to append to the image name:
Saved file format:jpeg
Output file location:Same folder as image|
Image bit depth:8-bit integer
Overwrite existing files without warning?:Yes
When to save:Every cycle
Record the file and path information to the saved image?:No
Create subfolders in the output folder?:No
Base image folder:Elsewhere...|
How to save the series:T (Time)
Save with lossless compression?:Yes

MeasureImageQuality:[module_num:8|svn_version:'Unknown'|variable_revision_number:6|show_window:False|notes:[]|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
Calculate metrics for which images?:Select...
Image count:1
Scale count:1
Threshold count:1
Select the images to measure:Collagen, Nuclei
Include the image rescaling value?:No
Calculate blur metrics?:Yes
Spatial scale for blur measurements:20
Calculate saturation metrics?:No
Calculate intensity metrics?:No
Calculate thresholds?:No
Use all thresholding methods?:No
Select a thresholding method:Otsu
Typical fraction of the image covered by objects:0.1
Two-class or three-class thresholding?:Two classes
Minimize the weighted variance or the entropy?:Weighted variance
Assign pixels in the middle intensity class to the foreground or the background?:Foreground

MeasureObjectSizeShape:[module_num:9|svn_version:'Unknown'|variable_revision_number:3|show_window:False|notes:[]|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
Select object sets to measure:Nuclei
Calculate the Zernike features?:No

Calculate the advanced features?:No

MeasureImageIntensity:[module_num:10|svn_version:'Unknown'|variable_revision_number:4|show_window:False|notes:[]|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:False]
Select images to measure:Collagen
Measure the intensity only from areas enclosed by objects?:No
Select input object sets:
Calculate custom percentiles:No
Specify percentiles to measure:10,90

ExportToSpreadsheet:[module_num:11|svn_version:'Unknown'|variable_revision_number:13|show_window:False|notes:[]|batch_state:array([], dtype=uint8)|enabled:True|wants_pause:True]
Select the column delimiter:Comma (",")
Add image metadata columns to your object data file?:Yes
Add image file and folder names to your object data file?:No
Select the measurements to export:Yes
Calculate the per-image mean values for object measurements?:No
Calculate the per-image median values for object measurements?:No
Calculate the per-image standard deviation values for object measurements?:No
Output file location:Elsewhere...|/g/pepperkok/tuechler/Experiments_2022/CP_outputs/NT22_002
Create a GenePattern GCT file?:No
Select source of sample row name:Metadata
Select the image to use as the identifier:None
Select the metadata to use as the identifier:None
Export all measurement types?:No
Press button to select measurements:Image|ImageQuality_PowerLogLogSlope_Collagen,Image|ImageQuality_PowerLogLogSlope_Nuclei,Image|Intensity_TotalIntensity_Collagen,Image|Intensity_TotalArea_Collagen,Image|PathName_Collagen,Image|PathName_Nuclei,Image|Metadata_Well,Image|Metadata_Year,Image|Metadata_FileLocation,Image|Metadata_PosNum,Image|Metadata_Date,Image|Metadata_Experiment,Image|FileName_Collagen,Image|FileName_Nuclei,Image|Count_Nuclei,Nuclei|Number_Object_Number
Representation of Nan/Inf:NaN
Add a prefix to file names?:No
Filename prefix:MyExpt_
Overwrite existing files without warning?:Yes
Data to export:Image
Combine these object measurements with those of the previous object?:No
File name:\g<RootImageName>_image.csv
Use the object name for the file name?:No

List of Tables