



# Abstract

It is estimated that around 10% of the global population is affected by chronic kidney disease (CKD), placing a significant burden on healthcare systems worldwide (Francis et al. 2024). Among the causes of CKD, 90% are attributed to glomerular diseases, in which podocytes are damaged and lost (Wiggins 2007). Podocytes are terminally differentiated visceral epithelial cells, which play a crucial role in establishing the selective permeability in the glomerulus. Recent large-scale transcriptomic approaches in mice and humans have demonstrated that alterations in the glomerular transcriptional program are a pivotal feature of numerous diseases affecting podocytes. In this thesis, I investigate the transcriptional regulation of healthy and damaged podocytes through multiple approaches. First, I analyzed novel bulk RNA sequencing data from glomeruli to characterize transcriptional changes resulting from perturbation of the *Wt1* transcription factor, a critical regulator of podocyte biology. I then leveraged novel ChIP-seq data to explore how podocyte damage rewires the transcriptional network, focusing on interactions regulated by *Wt1* and co-regulated by *Tead1*. Second, using single-nucleus RNA sequencing (snRNA-seq), I distinguished podocyte-specific transcriptional changes from shifts in cellular composition induced by podocyte injury. Third, I developed a universal metric of podocyte health, termed the podocyte damage score (PDS), using transcriptomic data from published sources. Applying the PDS to single-cell RNA sequencing datasets from various podocyte damage models allowed us to identify both universal and model-specific features of the transcriptional response to injury. Additionally, by integrating these findings with a podocyte transcriptional regulatory network (TRN) constructed from podocyte-specific ATAC-seq data and transcription factor motifs, I characterized the transcriptional regulators involved in podocyte transcriptome rewiring under damage conditions. In conclusion, this research advances our understanding of gene regulation in healthy and damaged podocytes and establishes methodologies for studying cell-specific mechanisms of disease at the single-cell level. It is my hope that these findings will contribute to the development of new therapeutic strategies for CKD.

# Acknowledgements

A doctoral thesis in the life sciences might have a single author, but almost always has a great many direct and indirect contributors. Mine is no exception, and so I have a number of people to thank. First, my sincere thanks to Andreas Beyer, who first welcomed me into his research group in 2015 as a master student and then in 2017 as a PhD candidate. He is an enthusiastic scientist and an understanding person, who was consistently supportive along the way. I am also thankful to Dr. Martin Kann, a co-supervisor of the project, who taught me a lot about kidney biology and the clinician's approach to science. My unwavering gratitude goes to Mahdieh Rahmatollahi and He Chen who did all the hard work in the wet lab to generate the data for this project. Without their work and their feedback on my analysis this research would not be possible.

My sincere thanks to Antonis, Carolina, Corina, Denis, Erik, Fabian, Jan, Jonathan, Luise, Mathiew, Paula, Robert, Ronja and all other group members for their help, support and fun during countless coffee breaks. Thanks also to Achim, his group and the rest of the bioinformatics community of CECAD and MPI for their patient help and advice. Thanks to all my friends from Belarus, all who I met in Cologne, Will, Davi and all the other people I have failed to mention here whose assistance, friendship and support I have relied on throughout this lengthy process. My deepest and sincerest thanks go to Nastassia and other past romantic partners without whom the adventure before 2020 would have been a great deal more difficult and a great deal less fun.

Thanks and support to all Belarusian people who were brave to stand for Humanity and against violence in 2020: we search so we shall find. I also extend my solidarity to my friends from Ukrainians and all other people who have been displaced, injured, lost their nearest or continue to suffer the brutal consequences of war.

My parents and family, of course, have been a constant source of love and support. Mom and Dad, thank you for everything. My love and thanks also to my siblings, cousins, nephews and nieces and to the rest of my big family, for their support and belief in me and my way. Finally, I want to express gratitude to my wife and my daughter, who have been the ultimate sources of love, growth and inspiration throughout the second half of this journey. Thank you for everything.

Творачы зруйнуем.<sup>1</sup>

---

<sup>1</sup> Essay "Adviečnym šliacham" (Eternal way). Ihnat Abdziralovič. Viľnia, 1921

# Contents

<b>Abstract</b>	<b>1</b>
<b>Acknowledgements</b>	<b>2</b>
<b>Contents</b>	<b>3</b>
<b>List of Abbreviations</b>	<b>6</b>
<b>Chapter 1. General Introduction</b>	<b>7</b>
1.1 Overview	8
1.2 Podocytes in Health and Disease	9
1.2.1 Podocytes: Highly Specialized Cells of the Kidney Glomerulus	9
1.2.2 Podocyte Damage as a Major Cause of Kidney Disease	10
1.2.3 Mechanisms in Podocyte Disease: The Role of the Podocyte Nucleus	11
1.3 Transcriptional Regulatory Networks	12
1.3.1 Types of Gene Regulatory Networks	12
1.3.2 Building and Interpreting Transcriptional Regulatory Networks	13
1.3.3 Changes in Transcriptional Regulatory Networks in Disease	13
1.3.4 The Role of Transcriptional Regulatory Networks in Therapeutics	14
1.4 Transcriptional Regulation in Podocytes	14
1.4.1 The Podocyte Transcription Factor Repertoire	14
1.4.2 WT1 as a Core Regulator of Podocyte Transcriptional Programs	15
1.4.3 Analysis of Transcriptional Networks in Glomerular Disease	16
1.5 Single-Cell RNA Sequencing: A Tool for Studying Cellular Damage	17
1.5.1 Challenges in Understanding Chronic Kidney Disease	17
1.5.2 Advancements Through Single-Cell RNA Sequencing	18
1.5.3 Limitations of Current Analytical Approaches	19
1.5.4 Quantifying Cellular Damage Using Single-Cell Data	19
1.6 Goals of the thesis	21
<b>Chapter 2. Materials and Methods</b>	<b>22</b>
2.1 Biological Materials.	23
2.2 Biochemistry and Molecular Biology Methods	24
2.2.1 Albumin and creatinine (ACR) determination with ELISA	24
2.2.2 Isolation of the glomeruli from mouse kidney	24
2.2.3 Bulk RNA-seq	24
2.2.4 Wt1 chromatin immunoprecipitation	25
2.2.5 Tead1 chromatin immunoprecipitation	25
2.2.6 Single nuclei RNA-seq	25

2.2.7 ATAC-seq	26
2.2.8 Library preparation and sequencing	26
2.3 Computational Methods	27
2.3.1 Analysis of Sequencing Data	27
2.3.1.1 Chip-seq	27
2.3.1.2 ATAC-seq	28
2.3.1.3 Bulk RNAseq	28
2.3.1.4 Single nuclei RNAseq	29
2.3.2 Differential Expression and Cell-type Abundance Analysis	29
2.3.3 Gene Set Functional Annotation and Visualization	30
2.3.4 Podocyte Damage Score Methods	30
2.3.4.1 Algorithm for generating single-cell damage signature	30
2.3.4.2 Gene set activity analysis in sc and sn RNAseq data	31
2.3.4.3 Randomisation test	32
2.3.4.4 Unsupervised damage signatures	32
2.3.4.5 Protein expression analysis	32
2.3.4.6 Spatial transcriptomics analysis	32
2.3.4.7 Analysis of mouse kidney circadian transcriptome	33
2.3.4.8 Analysis of circadian regulation in snRNA-seq	33
<b>Chapter 3. A Wt1 Transcriptional Regulatory Network in Podocytes</b>	<b>34</b>
3.1 Introduction	35
3.2 Transcriptomic Analysis of Wt1ko/wt Model of Genetic FSGS	39
3.2.1 Progressive FSGS is Characterized by Cell Junction Disassembly and Increased Immune Activation	40
3.2.2 Wt1 and Nphs2 Models of Genetic FSGS Share Transcriptional Changes	44
3.3 Chip-seq Analysis of Transcriptional Rewiring in Damaged Podocytes	49
3.3.1 WT1 and TEAD1 Chip-seq Analysis Reveal Coordinated and Independent Gene Regulation in Healthy Podocytes	50
3.3.2 WT1 Knock-Down Differentially Impacts WT1 and TEAD1 Binding	54
3.3.3 WT1 and TEAD1 have Common and Distinct Functions in Regulating Podocyte Structure and Survival in Podocyte Injury	58
3.3.4 Integrative Analysis of Chip-seq and RNA-seq: Altered Binding of WT1 and TEAD1 in Podocytes Affects Glomerular Transcriptome	61
3.4 Discussion	65
<b>Chapter 4. Single-cell Analysis of Podocyte Transcriptional Response to Damage</b>	<b>68</b>
4.1 Introduction	69
4.2 Single Nuclei RNA-seq Allows in-depth Characterization of Transcriptional Response to Podocyte Damage	73

4.2.1 Cell-type Composition of Glomeruli is Changed by FSGS Progression.	74
4.2.2 Transcriptional Changes in Podocytes of Wt1 and Nphs2 FSGS models.	76
4.2.3 Limitations of Glomerular Transcriptomics for Studying Podocyte Damage	80
4.2.4 Discussion	85
4.3 Single Cell Damage Score in Podocytes	88
4.3.1 Deriving Universal Podocyte Damage Signature	88
4.3.2 Testing and Validating the Podocyte Damage Score	89
4.3.3 Characterizing Cellular Changes in Disease Progression with PDS	95
4.3.4 Discussion	103
<b>General Discussion</b>	<b>107</b>
Integrative Analysis of Podocyte Transcriptional Networks	107
Single-Cell Resolution of Podocyte Injury	107
Development of a Universal Podocyte Damage Score	108
Implications for Podocyte Biology and FSGS Pathogenesis	109
Limitations of the Study	109
Future Research Directions	110
Conclusion	110
<b>References</b>	<b>111</b>
<b>Appendix A: Supplementary Figures</b>	<b>122</b>
<b>Appendix B: Supplementary Tables</b>	<b>138</b>

# List of Abbreviations

CKD - chronic kidney disease

DEG - differentially expressed genes

DE differential expression

DBG - differentially bound genes

DS - damage signature

FSGS - focal segmental glomerulosclerosis

GEO - Gene Expression Omnibus

GRN - gene regulatory network

KPMP - kidney precision medicine project

PA - pathway activity

PCA - principal component analysis

PC - principal component

PDS - podocyte damage score

sc - single cell

sn - single nucleus

RNA - ribonucleic acid

RNA-seq - RNA sequencing

TF - transcription factor

TRN - transcriptional regulatory network

UMAP - Uniform Manifold Approximation and Projection for Dimension Reduction

# **Chapter 1. General Introduction**

## 1.1 Overview

Chronic kidney disease (CKD) is a major global health concern, affecting approximately 10% of the world's population and placing a significant burden on healthcare systems worldwide. The prevalence of CKD has increased by 33% between 1990 and 2017, and it is projected to continue rising due to factors such as aging populations and increasing rates of diabetes and hypertension (Kovesdy 2022; Wong et al. 2024; GBD Chronic Kidney Disease Collaboration 2020).

Among the causes of CKD, approximately 90% are attributed to glomerular diseases, where damage and loss of podocytes play a critical role (Reiser and Sever 2013). Podocytes are terminally differentiated visceral epithelial cells that are essential for maintaining the selective permeability of the glomerular filtration barrier in the kidney. They possess a unique architecture with interdigitating foot processes connected by slit diaphragms, which prevent the leakage of proteins into the urine (Pavenstädt, Kriz, and Kretzler 2003).

Damage to podocytes disrupts this filtration barrier, leading to proteinuria—a hallmark of many kidney diseases—and can ultimately result in focal segmental glomerulosclerosis (FSGS), a severe form of glomerular scarring. FSGS is characterized by the segmental scarring of some glomeruli and is a frequent pathological event in CKD. The progression of FSGS often leads to end-stage renal disease, necessitating dialysis or kidney transplantation (Fogo 2015; Reidy and Kaskel 2007).

Recent large-scale transcriptomic approaches in mice and humans have demonstrated that alterations in the glomerular transcriptional program are pivotal features of numerous diseases affecting podocytes. Changes in gene expression within podocytes can influence key cellular processes such as cytoskeletal organization, cell adhesion, and survival, all of which are essential for maintaining podocyte structure and function (Fu et al. 2021; Ettou et al. 2020a; Lake et al. 2023; Clair et al. 2024).

Understanding the mechanisms underlying podocyte injury and the progression of glomerular diseases is essential for developing effective therapies. One crucial aspect is the regulation of gene expression by transcription factors (TFs), which control the transcriptional programs that maintain podocyte identity and function. Dysregulation of TFs can lead to aberrant expression of target genes, contributing to podocyte injury and disease progression (Susan E. Quaggin 2002; Rascole et al. 2007).

For instance, the Wilms' tumor 1 (WT1) transcription factor is a key regulator of podocyte gene expression and has been implicated in both the development and maintenance of podocytes. Mutations in WT1 can lead to a range of kidney diseases, including FSGS (Kann et al. 2015; Lipska et al. 2014). Similarly, TEA domain transcription factor 1 (TEAD1) has been identified as a crucial regulator of podocyte biology, influencing cellular pathways involved in podocyte function and disease (Burt et al. 2022; Chen et al. 2024).

To elucidate the complex gene regulatory mechanisms in podocytes, transcriptional regulatory networks (TRNs) provide a powerful framework. TRNs represent the interactions

between TFs and their target genes, allowing for the visualization and analysis of gene regulation at a systems level. By modeling these networks, researchers can identify key regulatory nodes and pathways that may be altered in disease states.

Since the advent of high-throughput measurement technologies in the late 1990s, reconstructing the structure of TRNs has been a central computational problem in systems biology (Lee et al. 2002; Tan, Tegner, and Ravasi 2008). Significant progress has been made over the past two decades, with many computational tools now available for network reconstruction. However, the complexity of biological systems and the limitations of current technologies mean that fully accurate TRN models are still challenging to obtain (Unger Avila et al. 2024).

In this thesis, we aim to update a TRN model for podocytes so that it can capture the key aspects of gene regulation during disease progression at the single-cell level. By integrating data from various sources, including high-throughput sequencing and chromatin accessibility assays, we can construct a comprehensive network that reflects the dynamic regulatory landscape of podocytes under healthy and diseased conditions.

## 1.2 Podocytes in Health and Disease

### 1.2.1 Podocytes: Highly Specialized Cells of the Kidney Glomerulus

Podocytes are terminally differentiated visceral epithelial cells that play a crucial role in establishing the selective permeability of the glomerular filtration barrier in the kidney. Structurally, podocytes consist of a central cell body and numerous cytoplasmic extensions known as foot processes (FPs). These foot processes interdigitate with those of adjacent podocytes, forming specialized cell-cell junctions called slit diaphragms (SDs). The slit diaphragm is a critical component of the filtration barrier, functioning as a size-selective and charge-selective filter that prevents the passage of proteins into the urine (Pavenstädt, Kriz, and Kretzler 2003; K. Tryggvason and Wartiovaara 2001).

The cytoskeleton of podocyte foot processes is composed of microtubules, intermediate filaments, and a dynamic network of actin filaments. The actin cytoskeleton, in particular, is essential for maintaining the complex architecture of foot processes and for the contractile functions of podocytes (Drenckhahn and Franke 1988; Ichimura, Kurihara, and Sakai 2003). The integrity of the slit diaphragm and foot processes is supported by various structural and signaling proteins, including nephrin, podocin, and CD2-associated protein (Schwarz et al. 2001; Huber and Benzing 2005).

Together with the glomerular basement membrane (GBM) and the fenestrated endothelium, podocytes form the tri-layered glomerular filtration barrier. The GBM is a specialized

extracellular matrix that provides structural support and contributes to the selective filtration properties of the barrier through its negatively charged components, such as heparan sulfate proteoglycans (Haraldsson, Nyström, and Deen 2008; Karl Tryggvason, Patrakka, and Wartiovaara 2006).

Podocytes express specific markers that reflect their specialized functions. Key podocyte markers include nephrin (NPHS1), podocin (NPHS2), synaptopodin (SYNPO), and Wilms' tumor 1 (WT1). These proteins are integral to the structure and signaling mechanisms of podocytes and are often used to identify podocytes in research and diagnostic settings (Dong et al. 2015; Schmid et al. 2003).

### 1.2.2 Podocyte Damage as a Major Cause of Kidney Disease

Podocyte injury is a central event in the pathogenesis of many glomerular diseases, leading to proteinuria and progressive kidney dysfunction. Damage to podocytes results in the effacement (flattening and broadening) of foot processes and disruption of the slit diaphragm, compromising the integrity of the glomerular filtration barrier. This allows proteins, such as albumin, to leak into the urine—a hallmark of glomerular disease.

Persistent podocyte injury can lead to detachment of podocytes from the GBM and subsequent loss of these cells, as podocytes are terminally differentiated and have limited capacity for regeneration. The loss of podocytes triggers a cascade of events, including the formation of sclerotic lesions and excessive deposition of extracellular matrix components, culminating in focal segmental glomerulosclerosis (FSGS). FSGS is characterized by segmental scarring of some glomeruli and is a common pathological finding in chronic kidney disease (CKD). During the progression of FSGS, the remaining healthy glomeruli undergo hyperfiltration to compensate for the loss of function, which increases shear stress and can exacerbate podocyte damage, creating a vicious cycle. If left untreated, FSGS can lead to end-stage renal disease, necessitating dialysis or kidney transplantation (Kriz, Gretz, and Lemley 1998; Mundel and Shankland 2002)(Kerjaschki 2001; Pavenstädt, Kriz, and Kretzler 2003; Mundel and Reiser 2010; Rosenberg and Kopp 2017).

Early detection of podocyte injury is crucial for preventing disease progression. Microalbuminuria, a condition characterized by the excretion of small amounts of albumin in the urine, is an early indicator of podocyte damage and CKD. Recent studies have linked microalbuminuria with aging and have demonstrated that interventions targeting the renin-angiotensin system, such as angiotensin-converting enzyme inhibitors or angiotensin receptor blockers, can reduce mortality risk and biological age in patients with CKD (Fong et al. 2024).

Current therapeutic strategies for FSGS include the use of corticosteroids and immunosuppressive agents like calcineurin inhibitors (CNIs). While these treatments can induce remission of proteinuria in some patients, they are not curative and may have significant side

effects (Campbell and Tumlin 2018). The narrow window for effective intervention underscores the need for novel therapies that target the underlying mechanisms of podocyte injury (Malaga-Diequez et al. 2015; Trachtman 2020).

Genetic factors play a significant role in the susceptibility to podocyte injury and the development of FSGS. Mutations in genes encoding key podocyte proteins, such as NPHS1 (nephrin), NPHS2 (podocin), ACTN4 (alpha-actinin-4), and TRPC6 (transient receptor potential cation channel 6), have been identified in hereditary forms of FSGS (Kaplan et al. 2000; Winn et al. 2005; Akchurin and Reidy 2015; Hall et al. 2015). These genetic insights highlight the importance of podocyte-specific pathways in maintaining glomerular function.

### 1.2.3 Mechanisms in Podocyte Disease: The Role of the Podocyte Nucleus

While extensive research has focused on the structural components of podocytes, emerging evidence indicates that nuclear processes play a pivotal role in podocyte function and disease. Transcriptional regulation within the podocyte nucleus orchestrates the expression of genes critical for maintaining podocyte identity, cytoskeletal integrity, and response to injury (Lake et al. 2023; Yang et al. 2021).

Mutations in genes encoding nuclear proteins and transcription factors have been implicated in podocyte diseases. Notably, mutations in WT1 (Wilms' tumor 1), a zinc-finger transcription factor highly expressed in podocytes, cause a spectrum of glomerular diseases, including Denys-Drash syndrome, Frasier syndrome, and isolated FSGS. WT1 regulates the expression of genes involved in podocyte differentiation and function, and its dysregulation leads to podocyte dysfunction and disease progression (Guo et al. 2002; Morrison et al. 2008).

Other transcription factors, such as LMX1B (LIM homeobox transcription factor 1 beta), MAFB, FOXC2, and TCF21, have been identified as essential regulators of podocyte development and maintenance (Rohr et al. 2002; Sadl et al. 2002; White et al. 2010; Maezawa et al. 2014). Mutations in these factors can result in podocyte abnormalities and glomerular disease. For example, mutations in LMX1B cause Nail-Patella syndrome, which includes nephropathy due to podocyte defects (Sato et al. 2005).

Moreover, epigenetic regulators and chromatin remodelers, such as SMARCA1, have been associated with podocyte diseases. Mutations in SMARCA1 lead to Schimke immuno-osseous dysplasia, characterized by immunodeficiency and nephrotic syndrome due to podocyte dysfunction (Lefevre et al. 2010; Liu et al. 2020).

Canonical signaling pathways activated upon podocyte damage, such as Notch, Wnt, and TGF- $\beta$  pathways, converge in the nucleus to modulate gene expression. Aberrant activation of these pathways contributes to podocyte injury and sclerosis. For instance, activation of Notch signaling in podocytes induces dedifferentiation and apoptosis, promoting glomerulosclerosis (H. Kato and Susztak 2012; D. Wang et al. 2011).

As mentioned previously, large-scale transcriptomic studies have revealed that changes in glomerular transcriptional programs are key features of CKD and podocyte disease processes. These studies highlight the dynamic nature of gene expression changes in podocytes during disease progression and underscore the importance of nuclear mechanisms in podocyte pathology. Understanding the transcriptional regulatory networks in podocytes is crucial for identifying novel therapeutic targets. Investigating the roles of transcription factors and chromatin modifiers in podocyte health and disease can provide insights into the molecular mechanisms underlying podocyte injury and suggest strategies for intervention.

## 1.3 Transcriptional Regulatory Networks

Gene expression in cells is a complex and tightly regulated process. This regulation is orchestrated by intricate systems known as gene regulatory networks (GRNs), which ensure that genes are expressed at the right time, location, and levels to maintain proper cellular function. The control of gene expression is fundamental to all biological processes, from development to cellular homeostasis, and its dysregulation is implicated in many diseases, including cancer, neurodegeneration, and chronic degenerative diseases (Macneil and Walhout 2011; Lee et al. 2002).

A transcriptional regulatory network (TRN) is a subset of GRNs that focuses on the role of transcription factors (TFs) in regulating gene expression. Transcription factors are proteins that bind to specific DNA sequences, thereby controlling the transcription of genetic information from DNA to mRNA. TRNs map out the relationships between TFs and their target genes, representing a directed, causal network. They can reveal how a single transcription factor might influence multiple downstream targets and how different TFs might work in combination to regulate complex gene expression patterns (Walhout 2006; Busser, Bulyk, and Michelson 2008; Qian et al. 2003).

### 1.3.1 Types of Gene Regulatory Networks

Gene regulatory networks are categorized based on the nature of the interactions they describe. Three common types of networks are:

- **Signal Transduction Networks:** These networks represent interactions between biomolecules that mediate cellular signaling pathways. A classic example is kinase networks, where proteins interact through phosphorylation to transmit signals. These signals often lead to changes in the activity of transcription factors, thereby influencing gene expression (Pawson 1995).

- **Transcription Factor Networks:** These networks focus on the direct regulation of genes by TFs. They map which genes are under the control of specific TFs, which can either activate or repress gene expression. These networks are critical in understanding cellular differentiation and response to stimuli (Qian et al. 2003; Bonneau et al. 2007).
- **Gene Coexpression Networks:** These are statistical networks based on gene expression data. They model relationships between genes that are coexpressed across different conditions or tissues. Unlike transcription factor networks, coexpression networks do not necessarily imply a direct regulatory interaction but suggest that the genes are part of a common functional module (Stuart et al. 2003).

### 1.3.2 Building and Interpreting Transcriptional Regulatory Networks

The construction of TRNs is typically data-driven, relying on a variety of high-throughput technologies. Advances in genomics, transcriptomics, and epigenomics have allowed researchers to map transcription factor binding sites and correlate these sites with gene expression data (Villar et al. 2015). Key data sources for constructing TRNs include:

- **ChIP-seq (Chromatin Immunoprecipitation Sequencing):** This method allows for the identification of transcription factor binding sites across the genome (Johnson et al. 2007).
- **ATAC-seq (Assay for Transposase-Accessible Chromatin Sequencing):** Used to assess chromatin accessibility, indicating potential regulatory regions (Pique-Regi et al. 2011).
- **RNA-seq:** This is used to measure gene expression levels and can help identify genes that are regulated by specific TFs based on changes in expression under different conditions.

The integration of these diverse data types is critical for building accurate and context-specific TRNs. Computational algorithms such as Inferrelator or SCENIC (Single-cell regulatory network inference and clustering) can infer regulatory interactions at the single-cell level, allowing for the analysis of gene regulation in different cell types and disease states (Aibar et al. 2017; Jackson et al. 2020). These approaches help researchers move beyond simply identifying which genes are expressed in a cell to understanding how and why these genes are being expressed.

### 1.3.3 Changes in Transcriptional Regulatory Networks in Disease

In many diseases, particularly chronic and degenerative conditions, TRNs undergo significant alterations. These changes can be triggered by mutations, environmental stress, or aging, which disrupt the normal regulation of gene expression (Lee and Young 2013). For example, in cancer, mutations in transcription factors or in the regulatory regions of target genes can lead to the misregulation of critical cellular processes, such as cell cycle control or apoptosis, contributing to tumorigenesis (Bradner, Hnisz, and Young 2017; Brach, Kauer, and Herrmann 1996).

In neurodegenerative diseases, such as Alzheimer's or Parkinson's, changes in TRNs can result in the loss of neuronal function and cell death (Pearl et al. 2019). Similarly, in chronic kidney disease, alterations in TRNs may disrupt the normal function of podocytes, leading to proteinuria and glomerulosclerosis (Kann et al. 2015). Identifying these changes in TRNs not only helps in understanding disease mechanisms but also offers potential therapeutic targets. By targeting specific TFs or their downstream pathways, it may be possible to restore normal gene expression patterns and slow or reverse disease progression (Hayashi et al. 2014).

### 1.3.4 The Role of Transcriptional Regulatory Networks in Therapeutics

Targeting transcriptional regulatory networks offers a promising avenue for developing novel therapies. Since TFs often regulate multiple genes, modulating a single TF can have broad effects on cellular function (Fisher and Kelly 2011). Small molecules that inhibit or activate specific TFs are being explored as potential treatments for various diseases. In cancer, drugs that target transcription factors such as MYC or p53 are already under investigation (“MYC on the Path to Cancer” 2012; Hollstein et al. 1991). Additionally, epigenetic therapies, which modify the accessibility of DNA to transcription factors, are showing promise in the treatment of certain cancers and genetic diseases (“Cancer Epigenetics: From Mechanism to Therapy” 2012).

In conclusion, TRNs provide a powerful framework for understanding the regulation of gene expression in both health and disease. Advances in high-throughput technologies and computational methods are allowing for the construction of increasingly detailed and accurate TRNs. By studying changes in these networks in disease states, researchers can identify new therapeutic targets and develop more effective treatments for a wide range of conditions (Konda et al. 2023).

## 1.4 Transcriptional Regulation in Podocytes

### 1.4.1 The Podocyte Transcription Factor Repertoire

The functionality and identity of podocytes are maintained by a complex network of transcription factors (TFs) that regulate gene expression essential for their development, differentiation, and maintenance. Transcription factors often function through cooperation with other proteins, forming intricate networks that allow for combinatorial regulation of gene expression (Susan E. Quaggin 2002). This cooperative interaction enhances the regulatory capacity of TFs, enabling a higher number of unique expression states without necessitating an increase in the number of distinct TFs (M. Kato et al. 2004). Additionally, such networks provide robustness to the system, allowing certain regulatory functions to persist even if one TF becomes non-functional (Macneil and Walhout 2011).

The significance of individual transcription factors in podocyte specification and differentiation has been established over the past few decades. Pod-1/Tcf21 was one of the first TFs identified in glomerular epithelial cells during kidney development (S. E. Quaggin, Vanden Heuvel, and Igarashi 1998). Subsequently, the roles of Wilms' tumor 1 (WT1), LIM homeobox transcription factor 1 beta (LMX1B), and v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (MAFB) were elucidated, highlighting their importance in both developing and mature podocytes (Susan E. Quaggin 2002; Miner et al. 2002; Moriguchi et al. 2006). These TFs, along with members of the FOXO family, have been confirmed as crucial regulators of podocyte gene expression (Yang et al. 2021).

Advancements in genomic technologies have allowed for the reconstruction of transcriptional regulatory networks in podocytes. In *Xenopus*, WT1 was identified as a master regulator within such a network (White et al. 2010). Further studies expanded on the role of WT1, revealing a comprehensive gene regulatory network in podocytes with WT1 at its core (Kann et al. 2015). Analysis of WT1 chromatin immunoprecipitation sequencing (ChIP-seq) data in mice suggested that podocytes possess an intricate network of transcription factors, including WT1, TEA domain transcription factor 1 (TEAD1), MAFB, FOX proteins, TCF21, and LMX1B (Kann et al. 2015; Rahmatollahi 2020). These transcription factors cooperatively bind to podocyte-specific enhancers to drive gene expression essential for podocyte function. However, the specific podocyte pathways co-regulated by these master TFs remain to be fully elucidated.

### 1.4.2 WT1 as a Core Regulator of Podocyte Transcriptional Programs

WT1 is an indispensable transcription factor for podocyte development and function. It is a DNA-binding zinc-finger protein belonging to the early growth response (EGR) family. WT1 expression is upregulated early during kidney development, marking the emergence of nephron

epithelia and podocyte precursors, suggesting its role as a pioneer transcription factor (Dressler 1995; Rasclé et al. 2007). Pioneer TFs initiate tissue-specific transcriptional programs by binding to regulatory regions of target genes and activating their transcription, often leading to the activation of additional TFs that stabilize the gene expression patterns (Zaret 2020).

In mature podocytes, WT1 expression is maintained at high levels throughout life. Decreased WT1 levels, due to mutations or disease states, result in failed maintenance of podocyte identity and function (Guo et al. 2002 (Morrison et al. 2008)). Mutations in WT1 lead to various renal diseases, ranging from severe congenital disorders such as Wilms' tumor (nephroblastoma), Denys-Drash syndrome, and Frasier syndrome, to less severe but persistent conditions like minimal change disease and focal segmental glomerulosclerosis (FSGS) (Lipska et al. 2014).

WT1 regulates gene expression by binding to promoters and enhancers of target genes, functioning as either a transcriptional activator or repressor depending on the recruitment of co-activators like CBP/p300 or co-repressors such as BASP1 in a chromatin-specific context (Green et al. 2009). Prior to comprehensive studies, only a few target genes of WT1 in podocytes had been identified (Guo et al. 2002).

Our group characterized the WT1-dependent transcriptional network in podocytes by performing ChIP-seq in wild-type mouse kidneys and mRNA sequencing (RNA-seq) of fluorescence-activated cell sorting (FACS)-sorted mouse podocytes *in vivo* (Kann et al. 2015; Rahmatollahi 2020). We identified approximately 14,500 reproducible WT1 binding sites in the podocyte genome, encompassing all previously known WT1 target genes in podocytes.

We classified WT1 target genes into two categories based on their binding patterns:

- **Class 1:** Genes bound exclusively at their transcription start site (TSS).
- **Class 2:** Genes bound at their TSS and additional sites within intronic or intergenic regions.

Both classes exhibited increased podocyte mRNA expression levels compared to unbound genes, with class 2 genes displaying higher expression levels. Functional annotation revealed that class 1 genes predominantly consisted of housekeeping genes, whereas class 2 genes represented a specialized podocyte gene toolkit essential for podocyte-specific functions.

Analysis of transcription factor binding motifs within non-TSS WT1 peaks (putative enhancer sites) uncovered a transcriptional network involving WT1 and other TFs with established roles in podocytes, including FOX proteins, MAFB, TCF21, and LMX1B. This supports the concept of WT1 functioning as a pioneer TF orchestrating a network crucial for podocyte identity and function.

Notably, the most significantly enriched TF binding motif within these enhancer regions was that of TEAD1, a downstream effector of the Hippo signaling pathway. The Hippo pathway is well-known for its role in regulating organ growth and cell proliferation. TEAD1's involvement suggests that Hippo signaling may play a significant role in podocyte biology, particularly in regulating cell-matrix interactions vital for podocyte structure and function (Bonse et al. 2018; Chen et al. 2024).

### 1.4.3 Analysis of Transcriptional Networks in Glomerular Disease

Building upon our understanding of the WT1-dependent transcriptional network in healthy podocytes, we sought to investigate how this network is altered in podocyte disease. We utilized a heterozygous WT1 knockout mouse model ( $WT1^{+/-}$ ), which serves as a model for genetic FSGS. In humans, haploinsufficiency of WT1 leads to reduced WT1 expression levels and is associated with FSGS and other glomerular diseases (Guo et al. 2002).

By backcrossing  $WT1^{+/-}$  mice into an FVB background, we generated a consistent phenotype characterized by proteinuria, foot process effacement, and histological features of FSGS beginning at four weeks of age (Kann et al. 2015). We performed WT1 ChIP-seq analyses on  $WT1^{+/-}$  mice and wild-type controls at eight weeks of age, an early stage of disease progression.

Our results demonstrated that a 40% reduction in WT1 expression in podocytes led to a significant loss of WT1 occupancy at approximately one-third of WT1 binding sites in  $WT1^{+/-}$  mice compared to controls (Kann et al. 2015). This loss of binding was independent of the initial strength of the binding sites in wild-type mice, suggesting a regulated process rather than random loss at weaker sites.

Gene ontology analysis of the lost WT1 binding sites revealed significant enrichment of genes involved in transforming growth factor-beta (TGF- $\beta$ ) and Notch signaling pathways (Kann et al. 2015). These pathways are known contributors to podocyte injury and glomerulosclerosis. For instance, WT1 was no longer bound near the promoters of *Deltex-2* and *Deltex-4*, positive regulators of Notch signaling. The upregulation of these genes upon WT1 loss suggests that WT1 may normally repress components of the Notch pathway to maintain podocyte integrity (Kann et al. 2015).

These findings provide evidence that alterations in WT1 binding contribute to the activation of pathogenic signaling pathways in podocyte disease. They also highlight the dynamic response of nuclear TFs to podocyte injury, emphasizing the importance of transcriptional regulation in disease progression. While ChIP-seq experiments have been instrumental in uncovering transcriptional networks, they are resource-intensive and technically challenging, especially when profiling multiple TFs in specific cell types like podocytes.

## 1.5 Single-Cell RNA Sequencing: A Tool for Studying Cellular Damage

### 1.5.1 Challenges in Understanding Chronic Kidney Disease

Despite extensive research over several decades, our understanding of the molecular and cellular processes driving CKD—particularly in relation to personalized diagnostics and treatment—remains limited (Delrue and Speeckaert 2024). Several factors contribute to this slow progress.

Firstly, the gradual progression of CKD in non-hereditary cases makes it challenging to study in humans. The disease often develops over many years or even decades, making it difficult to capture early-stage molecular events and understand the initiation of the disease process. Animal models, while invaluable for studying disease mechanisms, frequently represent rare monogenic conditions and may not adequately reflect the sporadic, slow-progressing nature of CKD in humans (Eddy et al. 2012; “Using Genetic and Species Diversity to Tackle Kidney Disease” 2020). As a result, findings from these models may have limited applicability to the broader patient population.

Secondly, CKD involves complex molecular and cellular processes, including poorly understood interactions between different cells and cell subtypes within the kidney (Kuppe et al. 2019). The kidney's intricate architecture and the dynamic interplay between podocytes, endothelial cells, mesangial cells, and immune cells add layers of complexity to disease mechanisms (Kirita et al. 2020). This complexity poses significant challenges in identifying specific pathways and targets for therapeutic intervention.

Common to all degenerative diseases, including CKD, is the accumulation of damage within and around cells, leading to a progressive decline in cellular functions accompanied by chronic inflammatory signals (López-Otín et al. 2013). Over time, this damage results in the loss of cellular function, cell death, cellular senescence, or neoplastic transformation (Campisi and d’Adda di Fagagna 2007). In CKD, the progressive decline of kidney filtration capacity can be traced to damage and loss of podocytes, the specialized epithelial cells that form the glomerular filtration barrier. For instance, Lake et al. (2023), Menon et al. (2020), Mariani et al. (2023), McNulty et al. (2022), and Hodgin et al. (2022) have all highlighted the central role of podocyte damage in CKD progression (Lake et al. 2023; Menon et al. 2020; Mariani et al. 2023; McNulty et al. 2022; Hodgin et al. 2022). These studies underscore the importance of understanding podocyte biology and the molecular mechanisms underlying their injury.

### 1.5.2 Advancements Through Single-Cell RNA Sequencing

Single-cell RNA sequencing (scRNA-seq) and single-nucleus RNA sequencing (snRNA-seq), collectively referred to as scRNA-seq, have substantially advanced our understanding of degenerative diseases, including CKD (Tang et al. 2009; Habib et al. 2016). These technologies enable the profiling of gene expression at the individual cell level, allowing

for the characterization of different cell types and subpopulations within diseased tissues (Potter 2018).

A key feature of scRNA-seq technologies is their ability to identify genes that are differentially expressed in specific cell types, aiding in the identification of disease-associated genes (Villani et al. 2017). For example, Young et al. (2018) utilized scRNA-seq to uncover cell-type-specific gene expression changes in kidney disease, providing insights into disease mechanisms. Additionally, scRNA-seq data can be employed to reconstruct transcriptional regulatory networks using tools like SCENIC (Single-Cell Regulatory Network Inference and Clustering) (Aibar et al. 2017). The SCENIC algorithm enables the identification of regulatory events and key transcription factors driving cellular phenotypes by integrating gene expression data with known transcription factor motifs (Aibar et al. 2017).

Furthermore, scRNA-seq data have been integrated with genome-wide association study (GWAS) findings to identify subpopulations of cells that potentially contribute to disease phenotypes (Wang et al. 2022). For instance, Wang et al. (2022) correlated scRNA-seq data with GWAS results to uncover disease-associated cell types in CKD. This integrative approach enhances our understanding of the genetic basis of CKD and aids in identifying potential therapeutic targets.

More recently, analytical approaches developed for scRNA-seq data have been adapted for spatial transcriptomics, enabling the investigation of molecular mechanisms of degenerative diseases within the tissue context (Ståhl et al. 2016). Spatial transcriptomics combines gene expression profiling with spatial information, allowing researchers to study gene expression patterns while preserving the spatial relationships between cells (Ståhl et al. 2016). This advancement provides a more comprehensive understanding of cellular interactions and the microenvironment's role in disease progression.

### 1.5.3 Limitations of Current Analytical Approaches

Despite the significant progress made possible by scRNA-seq and spatial transcriptomics, existing analytical approaches have limitations, particularly in their ability to map the gradual transition of cellular states as they progress from healthy to damaged conditions contributing to disease phenotypes (Trapnell 2015). Traditional analyses often involve comparing averaged gene expression levels between populations of cells from diseased and healthy tissues. While useful, this approach overlooks subtle variations in cellular states within individual samples and fails to capture the heterogeneity of cell populations (Kiselev, Andrews, and Hemberg 2019).

Over time, virtually all organs develop subpopulations of damaged cells that may serve as early indicators of emerging diseases (Schumacher et al. 2021). However, current tools lack the sensitivity to identify such rare cells in scRNA-seq data from clinically healthy subjects. This limitation hinders early detection and intervention, which are crucial for preventing disease progression.

Moreover, many analytical methods focus on identifying differences between predefined groups (e.g., healthy vs. diseased) rather than capturing continuous transitions between cellular states. This binary approach does not adequately reflect the dynamic nature of disease processes, where cells gradually accumulate damage and undergo phenotypic changes (Saelens et al. 2019).

#### 1.5.4 Quantifying Cellular Damage Using Single-Cell Data

To address these challenges, we propose quantifying the damage of individual disease-causing cells using generic, cross-model molecular markers. By developing a "damage score" based on the expression of specific marker genes consistently associated with cellular injury, we can sort cells according to their degree of damage. This approach facilitates the identification of molecular processes associated with progressive cellular injury and allows for the analysis of continuous transitions between cellular states.

Utilizing these generic damage markers enables the study of molecular mechanisms in disease-causing cells across multiple animal models and even under unperturbed control conditions. This cross-model applicability enhances the generalizability of findings and supports the identification of common pathways involved in cellular damage.

Our approach also enables the sequential ordering of events, distinguishing early versus late-stage cellular responses during pathogenesis. By aligning cells along a damage continuum, we can identify genes and pathways that are activated or repressed at different stages of disease progression. This temporal resolution is essential for understanding the dynamics of disease mechanisms and for identifying potential therapeutic targets that may be effective at specific stages.

For example, in CKD, applying this damage scoring method to podocytes in scRNA-seq data allows for the detection of early molecular changes preceding overt cellular injury. Identifying such early changes is critical for developing interventions aimed at preserving podocyte function and preventing progression to irreversible kidney damage.

Our proposed method addresses the limitations of current analytical approaches by focusing on the gradual accumulation of cellular damage and the associated molecular changes at the single-cell level. It provides a powerful tool for dissecting the complex cellular heterogeneity underlying degenerative diseases and has the potential to advance personalized diagnostics and therapeutics.

## 1.6 Goals of the thesis

The overarching aim of this thesis is to deepen our understanding of the transcriptional regulatory networks (TRNs) governing podocyte function and their dysregulation during focal segmental glomerulosclerosis (FSGS). The work is structured around two major goals: (1) reconstructing the *Wt1*-governed transcriptional network in podocytes and (2) leveraging podocyte heterogeneity to uncover universal mechanisms of transcriptional dysregulation and damage across different FSGS models. The research integrates bulk RNA sequencing (RNA-seq), single-nucleus RNA sequencing (snRNA-seq), chromatin immunoprecipitation sequencing (ChIP-seq), and assay for transposase-accessible chromatin sequencing (ATAC-seq) to build a comprehensive map of transcriptional changes in healthy and diseased podocytes.

### **Goal 1: Reconstruct the *Wt1*-governed transcriptional regulatory network in podocytes**

- **Task 1.1: Characterize transcriptional changes in glomeruli during genetic FSGS**
  - Perform bulk RNA-seq analysis on glomeruli from *Wt1ko/wt* and *Nphs2mut* mice.
  - Assess gene expression changes at both early and late stages of FSGS to understand the progression of the disease.
- **Task 1.2: Investigate the roles of *Wt1* and *Tead1* in rewiring the podocyte transcriptional network during damage**
  - Conduct ChIP-seq experiments for *Wt1* and *Tead1* in glomeruli from healthy and *Wt1ko/wt* mice at an early stage of FSGS (8 weeks old).
  - Integrate ChIP-seq and RNA-seq data from the *Wt1ko/wt* experiments to analyze the transcriptional regulatory network rewiring in podocytes under damage.

### **Goal 2: Leverage podocyte heterogeneity to study drivers of transcriptional rewiring**

- **Task 2.1: Characterize transcriptional changes in individual podocytes during genetic FSGS**
  - Perform single-nucleus RNA-seq on glomeruli from *Wt1ko/wt* and *Nphs2mut* mice to investigate podocyte-specific transcriptional changes.
- **Task 2.2: Develop a framework to assess podocyte damage in a model-agnostic manner**
  - Derive a universal transcriptomic signature of podocyte damage using public datasets.
  - Test and validate this signature to assess podocyte damage across various FSGS models.
- **Task 2.3: Identify universal mechanisms of transcriptional regulation in FSGS**
  - Align podocytes from different models by their Podocyte Damage Score (PDS) to identify common features of disease progression.
  - Reconstruct the podocyte transcriptional regulatory network with ATACseq and predict transcriptional regulators involved in the common podocyte damage response.

# **Chapter 2. Materials and Methods**

## 2.1 Biological Materials.

Wt1 and Nphs2 murine models of genetic focal segmental glomerulosclerosis (FSGS) were used to generate original data for the research. In the Wt1 model podocyte damage and FSGS is caused by deletion of exon 1 in one of the Wt1 alleles. This hetero-zygous deletion of Wt1 does not compromise the kidney development as evaluated by the morphological as well as functional analysis of the mice at the postnatal day 6 (Menke et al. 2003). In the Nphs2 model disease is caused by Nphs2<sup>R231Q/A286V</sup> compound heterozygous mutation in the podocin gene (Butt et al. 2020). Mice with this mutation develop progressive albuminuria starting at 2 to 4 weeks of age with kidney failure at around 20 weeks of age.

Mice were housed and maintained in CECAD in vivo Research Facility. All mouse experiments were performed with approval from The Animal Care Committee of the University of Cologne and LANUV NRW (Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, State Agency for Nature, Environment and Consumer Protection North Rhine-Westphalia) for this study.

All wet lab experimental work and sample preparations for bulk RNA-seq and CHip-seq, including microscopy, staining assays, measures of kidney function etc., was performed in the Nephrolab (University of Cologne, Germany) by Mahdiah Rahmatollahi, details on mouse lines and the experimental protocols can be found in her doctoral thesis (Rahmatollahi 2020). All wet lab experimental work and sample preparations for bulk ATAC-seq and single nucleus RNA-seq was performed in the Nephrolab by He Chen; details on mouse lines and experimental protocols can be found in her doctoral thesis (He Chen. 2024. Transcriptional Regulation and Epigenetic Landscape of the Glomerulus. Doctoral thesis, University of Cologne, unpublished). Figure 1 below presents an overview of experimental data used in the study.

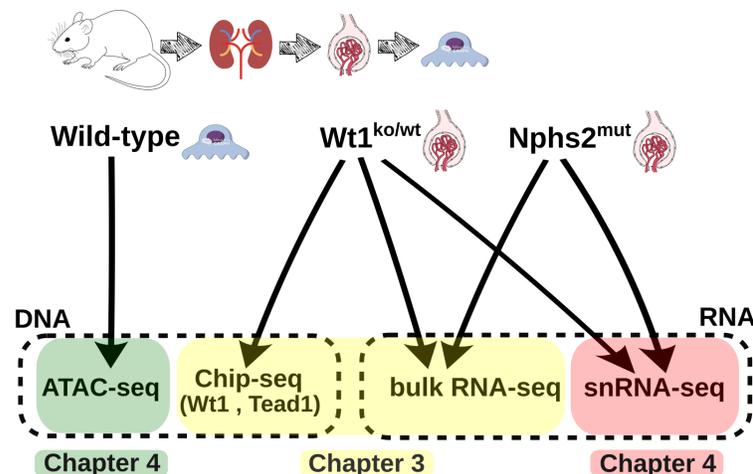


Figure 1: Overview of experimental data used in the study.

## 2.2 Biochemistry and Molecular Biology Methods

### 2.2.1 Albumin and creatinine (ACR) determination with ELISA

Urines from mice were first analyzed by coomassie urinary analysis to approximate the dilution for albumin ELISA (usually ranging between 1:100 to 1:100,000). All steps were carried out at room temperature. First, wells were coated for 1 h with 100  $\mu$ l of anti-mouse albumin coating antibody at 1:10,000 in coating buffer (0.05 m carbonate-bicarbonate, pH 9.6), then washed 5 times with 200  $\mu$ l of wash solution (50 mm Tris, 210.14 m NaCl, 0.05% Tween-20, pH 8.0) before incubating for 30 min with 200  $\mu$ l of blocking solution (50 mm Tris, 0.14M NaCl, 1% BSA, pH 8.0). The wells were washed again 5 times with 200  $\mu$ l of wash solution before incubation for 1 h with 100  $\mu$ l of standards or samples diluted in diluent buffer (50 mm Tris, 0.14 m NaCl, 0.05% w/v Tween-20, 1% w/v BSA, pH 8.0). Next, the wells were washed 5 times with 200  $\mu$ l of wash solution before incubation with 100  $\mu$ l of HRP detection antibody diluted 1:25,000 in the diluent buffer for 1 h. Finally, the wells were washed 5 times with 200  $\mu$ l of wash solution and developed with 100  $\mu$ l of substrate solution (100  $\mu$ g/ml TMB, 48 mm sodium acetate, 0.01% v/v hydrogen peroxide, pH 5.2) for 15 min in the dark. The reaction was stopped by adding 100  $\mu$ l of stop solution (0.18 m sulfuric acid) and the absorbance was measured at 450 nm. For the creatinine urinary colorimetric ELISA, urines were diluted 1:20 in water and assayed according to the manufacturer's protocol. All samples were measured in triplicates.

### 2.2.2 Isolation of the glomeruli from mouse kidney

Glomeruli were isolated from the mouse kidneys as described before (Boerries et al. 2013). In brief, kidneys were dissected together with the abdominal aorta after cervical dislocation of the mouse and each kidney was perfused with 1-2 ml of magnetic beads solution consisting of Magnetic Dynabeads in 1x Hank's Buffered Salt Solution (HBSS). Renal capsules were removed, and kidneys were minced in 1-mm 3 pieces using a scalpel. Digestion was done in 3 ml of the digestion solution in 37 °C for 15 minutes on a plate shaker. To facilitate digestion, kidney pieces were triturated using a cut pipette filter tip. Digested kidneys were meshed twice through a 100- $\mu$ m cell strainer and centrifuged. Glomeruli were collected after re-suspending the pellet using DynaMag Magnet.

### 2.2.3 Bulk RNA-seq

Isolated glomeruli were homogenized in 700  $\mu$ l TRI reagent and 140  $\mu$ l chloroform was added to the homogenate. After centrifugation, the aqueous phase was collected, and total RNA was extracted using the miRNeasy RNA extraction kit. RNA quality was assessed on RNA screen tape and all the samples showing RNA integrity number (RIN) > 8 were submitted for sequencing.

#### 2.2.4 Wt1 chromatin immunoprecipitation

Wt1 chromatin immunoprecipitation (ChIP) was performed by Dr. Maximillian Lenz in Nephrolab, University of Cologne. Mouse kidneys from 4-week old mice were processed by dissecting the cortex from medulla and mincing. Crosslinking was performed using 1% formaldehyde in PBS followed by quenching with 125 mM glycine. Further tissue disruption was carried out using a rotor-stator homogenizer in sterile 0.3 mM NaCl RIPA supplemented 42 with protease inhibitor. The pellet was sonicated in the sonication buffer on a probe-tip sonicator to achieve an average chromatin size ranging between 200-600 bp. Sonication efficiency was checked on a 2% agarose gel for each ChIP round. After sonication, sample was filled up to 1 ml with 0.3 RIPA, the Protease inhibitor was refreshed, and IP was carried out overnight by adding 3 µg of Wt1 C19 antibody or IgG isotype control. Samples were incubated with Protein G Dynabeads the next day for 2-3 hours at 4°C. After washing and elution, de-cross linking was done in SDS elution buffer at 65°C overnight. DNA extraction was completed using phenol/chloroform/isoamylalcohol 25:24:1. Following centrifugation, the aqueous phase containing the purified DNA is transferred to clean tubes and DNA is recovered by alcohol precipitation.

#### 2.2.5 Tead1 chromatin immunoprecipitation

Mouse kidneys from 8 to 10-week old mice were used for isolation of glomeruli. Crosslinking was performed using 1% formaldehyde in PBS followed by quenching with 125 mM glycine. Sonication was completed in the ice-cold fresh nuclear lysis buffer (NLB) supplemented with protease inhibitors on a Covaris M220 Focused-ultrasonicator for 5 minutes. A total of 8 mice were used as input per immunoprecipitation round and the sonication efficiency was checked on a 2% agarose gel per round. Chromatin buffer was added to samples in a 2:3 ratio, protease inhibitor was refreshed, and IP was carried out at 4°C overnight by adding 6 µg of Tead1 antibody or IgG isotype control priorly bound to Protein G Dynabeads. After washing and elution, de-cross linking was done in SDS elution buffer at 65°C overnight. DNA extraction was completed using phenol/chloroform/isoamylalcohol 25:24:1. Following centrifugation, the aqueous phase containing the purified DNA is transferred to clean tubes and DNA is recovered by alcohol precipitation.

#### 2.2.6 Single nuclei RNA-seq

Glomeruli were isolated from mice kidneys and subjected to isolation of nuclei using Sigma EZ lysis buffer supplemented with EDTA-free protease inhibitors and 0.1% murine RNase inhibitor. Isolated nuclei were washed twice in lysis buffer and counted with a hemocytometer. After centrifugation at 1000g for 10 min, nuclei were resuspended in PBS supplemented with 2%

BSA to reach an approximate concentration of 1000 nuclei/  $\mu\text{l}$ . Prior to immediate processing, the samples were further filtered through a 10  $\mu\text{m}$  cell strainer.

### 2.2.7 ATAC-seq

ATAC-seq was performed as described in (Buenrostro et al. 2013). Briefly, 50,000 GFP+ or tdTomato+ nuclei were sorted by FACS into lysis buffer and centrifuged at 1000 g for 10 min at 4 °C. The pellet was resuspended in 50  $\mu\text{l}$  of transposition reaction mix (25  $\mu\text{l}$  of transposition buffer, 2.5  $\mu\text{l}$  of TDE1, and 22.5  $\mu\text{l}$  of nuclease-free water). The reaction was carried out at 37 °C for 30 min. Immediately following transposition, DNA was purified using Zymo DNA clean concentrator-5 per the manufacturer's instructions. The transposed DNA was eluted in 10  $\mu\text{l}$  of elution buffer.

### 2.2.8 Library preparation and sequencing

Library preparation and sequencing of ATAC-seq, Chip-seq, bulk and single-nuclei RNA-seq experiments were performed by the Cologne Centre of Genomics (CCG). 12 cycles of PCR amplification were used for ATAC-seq samples. Cycle numbers of PCR amplification used for Chip-seq samples were determined by qPCR, single-end libraries were constructed according to the standardized protocols and sequencing was performed using Illumina HiSeq sequencer. For bulk RNA-seq experiments, Ribo-minus libraries were constructed according to the standardized protocols and paired-end sequencing was done on an Illumina HiSeq sequencer. For snRNA-seq, libraries and sequencings were prepared per the manufacturer's instructions supported by 10X Genomics.

## 2.3 Computational Methods

Data analysis and visualization was done in R (“Website,” n.d.-a), unless specified otherwise.

### 2.3.1 Analysis of Sequencing Data

#### 2.3.1.1 Chip-seq

**Mapping and peak calling.** Quality of raw sequencing data was assessed with FASTQC and then reads were mapped to the mouse genome mm10 using BWA mem (Li and Durbin 2009) with default parameters, no trimming is done before mapping. Genomic regions from the ENCODE blacklist were excluded from the analysis (Li and Durbin 2009; Amemiya, Kundaje, and Boyle 2019). Chip-seq sample quality was assessed with cross-correlation plots: samples with relative strand cross-correlation coefficient (RelCC) around or larger than 1 were considered. Peak calling was done by MACS2 using low-confidence parameters (p-value <0.01), consistency of peaks was assessed with IDR analysis (T. Bailey et al. 2013). A consensus peakset was calculated with MSPC software (Jalili et al. 2015): a peak has to be called in minimum of 2 replicates and the combined q-value has to be smaller than 0.05 to get selected into the consensus peakset.

**Differential binding** analysis and plotting was done with the DiffBind R package v3.10.1 using GLM implemented in edgeR method, see `ChIPseq_DiffBind.r` script for the details.

**Identification and annotation of target genes.** Targets of WT1 and TEAD1 wild-type peaks were inferred using TFtargetCaller v0.7 R package with ClosestGene method, using TSS positions of all mouse genes from *Ensembl* database v.79. Sets of peaks differentially bound in  $Wt1^{ko/wt}$  glomeruli were annotated, due to their limited size, with `annotatePeakInBatch()` function from ChIPpeakAnno v3.24.2 R package using the following parameters: `multiple = T`, `output="both"`, `maxgap=2000`, `AnnotationData= EnsDb.Mmusculus.v79`. With this approach each peak is assigned a) the nearest, by the distance to TSS, feature and b) any features that overlap the peak (maximum gap allowed is 2000 bp) that are not the nearest features.

Peaks were classified as cis- and trans-regulatory regions if the log10 distance to the nearest TSS was smaller or bigger, respectively, than threshold value of 3.5. The threshold was chosen based on the empirical distribution of distances between all peaks and TSS of their annotated target genes.

**Motif enrichment.** Peaks were limited to +/-100 bp around the peak center for the motif enrichment. De-novo motif analysis was performed with HOMER. Known-motif enrichment was tested with `runAme(method = "fisher")` from `memes` R package v.0.99.11.9000. `Ame` was run in discriminative mode, to discover motifs enriched relative to shuffled input, or a set of provided control sequences. 110 PWM motifs of TFs expressed in podocytes were tested.

### 2.3.1.2 ATAC-seq

**Mapping** was done using BWA mem with default parameters. Data quality was assessed by calculating the distance between the R1 and R2 read pairs, so called insert sizes. Peaks were called using *Genrich* tool (“GitHub - jsh58/Genrich: Detecting Sites of Genomic Enrichment,” n.d.). Processed ATACseq data and known transcription factor binding motif was used to reconstruct the podocyte transcriptional regulatory network.

#### **Generation of podocyte TRN**

There are more than thousand mouse genes with annotated transcription factor activity (Sun, Wang, and Sun 2017) but not all of them are expressed in podocytes and/or detected in sc/snRNAseq data. To address this issue we selected TFs with detection rate above 5% in at least one of the analyzed sc/snRNAseq datasets, which resulted in 110 Transcription factors that were used to construct the podocyte transcription regulatory network (TRN). TRN is a directed graph representing interactions between transcription factors and their target genes. The TRN was constructed using the podocyte ATACseq data and known transcription factor binding motifs in following 3 steps:

**Footprinting** is performed on the ATACseq peaks with HINT (Yan et al. 2020) to identify and locate TF binding events more precisely.

**Motif scanning** is done with the Fimo tool from MEME-suite on ATACseq footprints, scanning for sequences that are matched by TF position weight matrices (PWM) (T. L. Bailey 1994).

**Associating potential binding events with genes** is the last necessary step in generating TRN, which was done with ClosestGene approach implemented in TFtargetCaller R package (Sikora-Wohlfeld et al. 2013), Target genes were called for each tested TF and genes q-value <0.3 were considered putative targets of the respective TF.

### 2.3.1.3 Bulk RNAseq

**Mapping.** STAR (Dobin et al. 2013) was used for mapping reads. Alignment of RNAseq data to the reference genome showed that between 30-40% of reads were mapped to the intronic regions and 25-30% to the coding exonic regions. According to literature, the changes in the intronic read counts are not merely due to the technical artifacts. In fact, they directly reflect changes in transcriptional activity (Gaidatzis et al. 2015). Hence, both exonic and intronic reads were utilized in bulk and single cell/nuclei transcriptome analyses to increase the statistical power of detecting transcriptional changes.

**Removal of unwanted variation.** For the  $Wt1^{ko/wt}$  dataset removal of unwanted variation using replicate samples (RUVs) was done using the RUVseq R package (Risso et al. 2014). RUVs estimate the factors of unwanted variation using an empirically defined set of invariable genes, i.e. control genes, and design matrix: it basically removes variation that is not associated

with conditions of interest. Differential expression analysis was performed on the results of RUVs, when applied.

#### 2.3.1.4 Single nuclei RNAseq

**Mapping.** STARsolo v2.7.3 (Dobin et al. 2013) was used to map reads to the mouse genome and then gene abundances were quantified using both exonic and intronic reads and pre-mRNA full gene models. Sequence files from 2 sequencing lanes were pulled in one sample.

**Removal of ambient RNA.** Ambient RNA may be an issue in snRNA-seq data analysis, masking rare cell-types and affecting the downstream analysis (Caglayan, Liu, and Konopka 2022). Ambient RNA was removed from the raw read count (UMI) matrices using R package decontX (Dobin et al. 2013; “decontX,” n.d.).

**General processing of sc/snRNAseq count tables.** Count matrices corrected for ambient RNA were loaded and analyzed in Seurat v4.0.1 R package (Hao et al. 2021). Cells with less than 200 non-zero features were filtered out. Cells with more than 1% of mitochondrial RNA were also filtered out from snRNA-seq samples. Clustering and dimensionality reduction was performed on the filtered data.

**Doublet removal.** Cluster-based artificial doublets were identified in pre-processed and clustered Seurat objects using scDblFinder (“scDblFinder,” n.d.) from the same-named R package (?). Cells annotated as “singlets” by the algorithm were removed from the downstream analysis and remaining cells were re-clustered. The resulting object was used for visualization, cell-type annotation and differential expression (DE) analysis, if needed.

**Cell-type annotation** was performed on cluster level, manually, by using kidney cell-type marker sets published by Humphreys lab (“scDblFinder,” n.d.; Kirita et al. 2020).

### 2.3.2 Differential Expression and Cell-type Abundance Analysis

Differential expression/abundance analysis of original data was done using DESeq( test="LRT" ) function from R package DESeq2 v.1.42.0 (Love, Huber, and Anders 2014). Differential cell-type abundance analysis was performed on a cell-type proportion matrix, calculated after annotating clusters with cell-types.

Effects of genotype in a specific age, general effect of genotype and effect of disease progression (age:genotype) were tested. Effects of batch and sex were controlled for, when respective annotations available. Analysis of the disease progression focuses on changes between control and mutant samples that depend on age of mice. This, combined with the analysis of genotype effects that do not depend on time, allows to separate, to some extent, primary and secondary effects of the mutation.

### 2.3.3 Gene Set Functional Annotation and Visualization

Gene sets were annotated with *Gene Ontology* (GO) terms and a collection of pathways, which includes Reactome (Croft et al. 2011), KEGG (Kanehisa and Goto 2000) pathways and 50 Hallmark mouse gene sets (MSigDB) (Liberzon et al. 2015). Significance of the overlap was tested with a one-sided Fisher test.

GO annotation was in most cases, unless specified otherwise, performed using a custom R function `sf.clusterGoByGeneset()` provided in `justGO_ROBERT.r` script, developed by Robert Sehlke. The function addresses redundancy of significantly enriched GO terms by clustering terms based on their similarity, cutting the tree at a chosen height and then selecting the most representative term from each cluster (see the code for details).

Gene set enrichment analysis (GSEA) of the ranked lists was performed with `fgseaMultilevel()` function of `fgsea v1.17.1` R package (“Fgsea,” n.d.) using aforementioned collection of pathways.

#### Visualization

Results of the enrichment were routinely visualized using barplots, where X-axis shows  $-\log_{10}$  p-value of the enrichment test and Y axis shows term labels. Labels of 50 Hallmark mouse gene sets are marked with prefix H50 on all pathway histograms. Terms are hierarchically clustered by similarity of genes annotated under each term. Row annotation bar on the right side of GO histograms always color-code terms by GO category: BP - biological process, MF - molecular function, CC - cellular component.

In addition, we used a so-called **2D GO plot** to highlight differences between results of 2 differential expression tests. To produce a plot for each of the 2 tests being contrasted, we calculated the average log fold change of genes annotated under each GO term significantly ( $p < 0.05$ ) enriched in at least one of the 2 DE tests. We then used a scatter plot to compare average expression changes of GO terms in one test with average expression changes of GO terms in another DE test against each other. The relationship between test-specific expression changes of GO terms was modeled using a simple linear model. 2D GO plot highlights functional differences between the tests by identifying GO terms that deviate significantly from the expected relationship. Specifically, GO terms located outside the 95% prediction interval of the regression line exhibit substantially greater average expression changes in either one or the other DE test.

### 2.3.4 Podocyte Damage Score Methods

#### 2.3.4.1 Algorithm for generating single-cell damage signature

To generate a universal transcriptome signature of podocyte damage we collected over 50 published transcriptomic datasets that represent different models of podocyte damage and

include microarray, bulk and single-cell RNAseq datasets. After initial analysis and quality checks we selected and analyzed 37 datasets, described in [Supplementary Table 4](#). Each selected dataset was subject to differential expression analysis, always comparing disease versus control samples. Microarray data was processed and DE tested using the oligo (“Oligo,” n.d.) package. Non-normalized bulk RNAseq count tables were subject to DE analysis with DESeq2. Single cell RNAseq UMI count tables were loaded and processed in Seurat: podocytes were extracted based on cell-type markers, non-parametric Wilcoxon rank sum test was used for robust DE testing of the damage effect in podocytes.

Results of DE analysis in individual studies were combined to generate a universal gene signature of podocyte damage. P-values for differential expression may not be comparable between studies, due to variation in sample numbers and/or variable quality of the data. We therefore decided to base the selection process on ranks of p-values rather than defining a common p-value threshold across studies. The damage signature was generated in following steps:

1. Genes differentially expressed in at least 75% of the studies ( $N > 27$ ) are selected.
2. Genes are further filtered by the consistency of expression change (disease versus healthy): the direction of change should be same in at least 75% of the studies
3. Selected genes are ranked in each study by p-value of differential expression.
4. Ranks are aggregated by calculating the average rank of each gene across all studies.
5. Genes are sorted by the average rank, and the top  $N$  ( $N = 42$  for the final score) are selected as a universal damage signature.

The damage signature is used to calculate the **podocyte damage score** (PDS) for each individual cell with AUCell R package (“AUCell,” n.d.). The AUCell score is calculated separately for up and down-regulated genes and then the latter is subtracted from the former to get the final podocyte damage score.

#### 2.3.4.2 Gene set activity analysis in sc and sn RNAseq data

AUCell is also used to calculate activity of any gene signature, e.g. pathway genes or cell-type markers, in individual cells of single cell and single-nuclei RNA-seq data.

To select pathways that change along the podocyte damage axis we calculate Spearman correlation of pathway activities and PDS ( $q$ -value  $< 0.05$ ). For visualization, cells are ordered by PDS and the pathway activity (PA) signal is smoothed using a moving average window of width=500 and a step=250 cells. The results are visualized using a heatmap, called pathway fingerprint, where X-axis show PDS, Y-axis show pathways and smoothed PA is encoded by color of cells. Alternatively, results are visualized for individual features, which allows to compare side by side different damage models. Reactome and KEGG pathway databases were used for the analysis.

#### 2.3.4.3 Randomisation test

To assess the sensitivity of our approach to the inclusion of low-quality or irrelevant studies, we implemented a randomization procedure. In this process, a certain percentage of studies had their p-values and log fold changes (LFCs) shuffled, effectively simulating noise in the differential expression results. We then generated a signature from this partially randomized data and used it to calculate a damage score in a test dataset.

Since the choice of which studies to randomize could influence the damage score, we performed 50 rounds of subsampling for each level of randomness. In each round, different studies were selected for randomization. After each randomization round, we calculated the average disease score across both control and experimental samples or cells, resulting in 50 data points for both control and experimental groups at each level of randomness.

This approach allowed us to evaluate the robustness of the damage score and understand how randomizing portions of the data affected its reliability. The multiple rounds of randomization provided a comprehensive view of the variability in score outcomes, ensuring that the results were not overly dependent on any specific subset of studies.

#### 2.3.4.4 Unsupervised damage signatures

As an alternative to the supervised approach of aggregating DE results, the gene signature can be generated in an unsupervised fashion, by trajectory analysis of single cell data, where trajectory reflects progression of the underlying biological process - extent of a cell damage. To this end we extracted podocytes from 8 scRNAseq damage models (see Supplementary Table 1), integrated the data and inferred the trajectory by fitting a single principal curve with slingshot R package (“Slingshot,” n.d.) a cluster of control cells was treated as a cluster of origin, cells of each damage model were treated as a cluster. Next we test for significant differences in gene expression with respect to pseudotime. The “pseudotime” is a number describing the relative position of a cell along the trajectory, i.e. cells with larger values are considered to be more damaged than cells with smaller values.

#### 2.3.4.5 Protein expression analysis

PXD016238 and PXD018326 proteomics datasets were used to analyze expression of podocyte damage markers on the protein level. Both datasets were generated in the Nephrolab. The LFQ data was processed with the R package DEP (Arne Smits [cre, aut], Wolfgang Huber [aut] 2017). Only proteins expressed in more than 50% of replicates of at least one condition were analyzed.

#### 2.3.4.6 Spatial transcriptomics analysis

Both sequence- (10X visium) and image-based (Slide-seqV2) transcriptomic datasets were used to validate the podocyte damage score in murine and human samples. Spatial coordinate and count tables were loaded and analyzed with Seurat.

Slide-seqV2 spatial transcriptomics mouse dataset (Marshall et al. 2022) was retrieved from GEO database entry GSE190094. Only BTBR and control samples of sufficient quality were used for correlation analysis of PDS and glomerular morphological features. Podocyte KNN filtration, glomerular identification and size estimation was performed using the publication code.

10X visium spatial transcriptomics data for sample S-1905-017750 (patient 29-10282) was retrieved from the public database of kidney precision medicine project. The results here are in whole or part based upon data generated by the Kidney Precision Medicine Project (“Website,” n.d.-b).

#### 2.3.4.7 Analysis of mouse kidney circadian transcriptome

To identify kidney specific circadian genes we used bulk kidney RNAseq data from young and old mice that was previously produced and analyzed by the Nephrolab. Cycling genes were identified from the count matrix with `meta2d(cycMethod = "JTK")` function from `MetaCycle v1.2.0` R package (“GitHub - gangwug/MetaCycle: An Integrated R Package to Evaluate Periodicity in Large Scale Data,” n.d.). Overall 317 genes with BH adjusted p-value <0.01 in either young or old mice were called circadian. Estimation of acrophase and acrophase uncertainty for circadian genes was done by fitting a cosinor model with `cglm(Y ~ amp_acro(time_col=Time, period = 24 ))` function of `GLMMcosinor 0.2.0.9` R package (Hercz 2013; “GitHub - ropensci/GLMMcosinor: An R Package for Flexible Cosinor Modelling Using the `glimmTMB` Framework,” n.d.), where Y is rlog-normalised expression values of a gene.

#### 2.3.4.8 Analysis of circadian regulation in snRNA-seq

**Circadian rhythm disruption (CRD)** in individual podocytes was calculated using `cal_CRDscore()` from the `CRDscore` R package (He et al. 2022). A set of cycling genes identified from the kidney circadian bulk RNAseq was used to calculate CRD.

**Estimation of circadian time in individual podocytes** was done using a `Tempo` algorithm, implemented as a python package (Auerbach, FitzGerald, and Li 2022). Estimates of gene acrophases, required by the algorithm, were taken from the analysis of bulk circadian RNAseq data from the mouse kidney dataset. `Clock` or `Arntl` were used as reference genes.

# **Chapter 3. A Wt1 Transcriptional Regulatory Network in Podocytes**

## 3.1 Introduction

### **WT1 is a core regulator of the podocyte transcriptional programs**

The critical role of specific transcription factors in podocyte specification and differentiation has been well established over the past few decades. Research in the early 2000s identified transcription factors such as TCF21, WT1, LMX1B, MAFB, and members of the FOX family as crucial regulators of podocyte gene expression (S. E. Quaggin et al. 1999; Palmer et al. 2001; Rohr et al. 2002; Moriguchi et al. 2006). Advancements in our understanding of individual factors and methodological innovations led to the reconstruction of the core transcription factor regulatory network in podocytes of *Xenopus* by the end of the 2000s, where WT1 emerged as a master regulator .

Further expanding on the role of WT1, Kann et al. (2015) (Kann et al. 2015) revealed a comprehensive Wt1 gene regulatory network in podocytes by using chromatin immunoprecipitation followed by sequencing (ChIP-seq). Chip-seq provides direct evidence of transcription factor (TF) binding to DNA and is a powerful tool for reconstructing transcriptional regulatory networks by identifying genes directly targeted by TFs of interest (Johnson et al. 2007). Kann et al. (2015) also demonstrated that Tead1, a downstream effector in the hippo signaling pathway and an important regulator of the podocyte transcriptome, is a target of WT1 binding. Overall, the results suggested that WT1 together with several pioneer TFs initiate a tissue-specific TF network by binding to further TF genes and subsequently activating their transcription.

Building on insights into the WT1-dependent transcriptional regulatory network (TRN) in healthy podocytes *in vivo*, we aimed to modulate this system and investigate its response in the context of focal segmental glomerulosclerosis (FSGS), a podocyte disease. To this end, we investigated rewiring of glomerular and podocyte gene regulation in heterozygous Wt1 knock-out murine model of genetic FSGS, further denoted as Wt1<sup>ko/wt</sup>, using both Chip-seq and RNA-seq experiments. In addition, we profiled the glomerular transcriptome of Nphs2 mutant murine model of genetic FSGS, further denoted as Nphs2<sup>mut</sup>, to delineate common and model-specific features of transcriptome remodeling under FSGS. Figure 3.1 gives an overview of the experimental data generated for this chapter. All wet lab experimental work for this chapter was performed by Mahdiah Rahmatollahi.

Chapter 3 is split into several parts, according to a specific research question and an experiment performed to address it. Below is a short summary of each part.

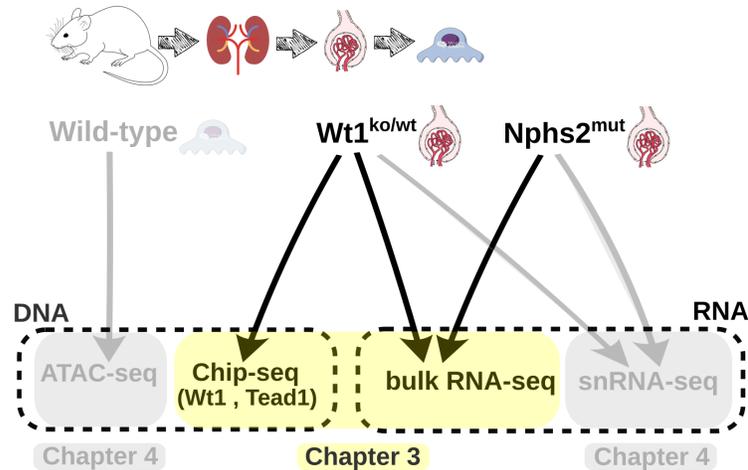


Figure 3.1: Overview of experimental data used in chapter 3<sup>2</sup>.

### Transcriptomic analysis of $Wt1^{ko/wt}$ model of genetic FSGS

To characterize transcriptional changes in glomeruli caused by heterozygous deletion of  $Wt1$ , we performed a bulk RNA-seq experiment on glomeruli samples from wild-type and  $Wt1^{ko/wt}$  mice. This published (Johnson et al. 2007; Menke et al. 2003) murine model of genetic FSGS has a normal embryonic development and shows consistent phenotype of proteinuria, foot process effacement and histologic features of FSGS starting at age 4 weeks, with advanced proteinuria observed at age 12 weeks. We therefore attempted to differentiate early and late effects of  $Wt1$  perturbation on transcriptome by generating and analyzing control and mutant samples from 4 (early FSGS) and 12 (late FSGS) week old mice.

Additionally, we sought to distinguish primary effects of  $Wt1$  perturbation from secondary effects due to disease progression by comparing  $Wt1^{ko/wt}$  with another published model of genetic FSGS in which disease is caused by  $Nphs2$  R231Q/A286V mutation in the podocin gene (Butt et al. 2020). Murine kidneys with this mutation, further referred as  $Nphs2^{mut}$ , undergo normal embryonic development, start showing FSGS features between 2 and 4 weeks of age and suffer from advanced glomerulo-sclerosis by the age of 8 weeks, which makes  $Nphs2^{mut}$  mutation more aggressive than  $Wt1^{ko/wt}$ .

Transcriptomic analysis of the disease stages showed progressive alteration of glomerular transcriptome with changes mostly consistent between 2 stages and the magnitude of changes being higher at the later disease stage. Comparison of  $Wt1^{ko/wt}$  and  $Nphs2^{mut}$  models of genetic FSGS showed that the overall effect of  $Nphs2^{mut}$  mutation on glomerular transcriptome is much stronger, compared to  $Wt1^{ko/wt}$ , but it affects mostly the same biological processes. Overall, results of differential expression analysis shed some light on transcriptional rewiring in sick

<sup>2</sup> All samples were taken from murine glomeruli.

podocytes but also showed intrinsic limitations of bulk glomerular RNAseq transcriptomics in studying podocyte gene regulatory network.

### **Chip-seq analysis of transcriptional rewiring in damaged podocytes**

To overcome the limitations of bulk transcriptomic analysis and gain insights into binding dynamics of podocyte transcription factors (TFs), we employed chromatin immunoprecipitation followed by sequencing (ChIP-seq) (Johnson et al. 2007; P. J. Park 2009). In this method, proteins are cross-linked to DNA within the cell, the chromatin is fragmented, and an antibody specific to the TF of interest is used to immunoprecipitate the TF-DNA complexes. The bound DNA is then sequenced and mapped back to the genome, providing a comprehensive profile of TF binding sites. Chip-seq is a powerful tool for reconstructing transcriptional regulatory networks by identifying genes directly targeted by TFs of interest.

Previous ChIP-seq analyses of WT1 binding in healthy podocytes established WT1 as a master regulator of the podocyte transcriptional regulatory network (TRN) (Guo et al. 2002; Kann et al. 2015). These studies also showed that TEAD1, a downstream effector in the Hippo signaling pathway and another key regulator of the podocyte transcriptome, is a target of WT1 binding. Building on this knowledge, we investigated transcriptional rewiring in the heterozygous *Wt1* knock-out mouse model of genetic FSGS by conducting WT1 and TEAD1 ChIP-seq experiments in both wild-type and *Wt1*<sup>ko/wt</sup> glomeruli of 8 week old mice, at which age kidneys of mutant animals show signs of the early stage of FSGS. Our results demonstrated that nuclear TFs in podocytes respond dynamically to injury, with changes in binding patterns reflecting functional adaptations or maladaptations. Importantly, our study confirms that ChIP-seq analysis of disease-affected podocytes is both feasible and informative, providing valuable insights into the molecular mechanisms driving disease progression.

We also predicted target genes of differential binding, which allowed us to elaborate the link between rewiring of *Wt1*-controlled TRN and changes in podocyte biology associated with FSGS. By integrating the functional analyses of Chip-seq from both wild-type and FSGS glomeruli, we gained a better understanding of how transcriptional regulation by WT1 and TEAD1 maintains podocyte function under normal conditions and how alterations in their binding contribute to disease pathology.

In the final part of this chapter, we aimed to further differentiate the primary effects of *Wt1* knockdown from secondary effects due to disease progression by integrating our ChIP-seq data with RNA-seq data. We correlated changes in gene expression with changes in WT1 and TEAD1 binding in *Wt1*<sup>ko/wt</sup> mice. Additionally, we compared the functional annotations of differentially expressed genes (DEGs) and differentially bound genes (DBGs).

Our integrative analysis revealed that targets of differential binding are more enriched in podocyte-specific biological processes, whereas the glomerular transcriptome may reflect signals from various cell types within the glomerulus. Importantly, the integration of RNA-seq and ChIP-seq data demonstrated that altered binding of WT1 and TEAD1 in the podocytes of

Wt1<sup>ko/wt</sup> mice directly affects the glomerular transcriptome by modulating key biological processes, ultimately leading to podocyte injury.

These findings highlight the critical roles of WT1 and TEAD1 in regulating gene expression related to podocyte structure and function. Moreover, our study underscores the value of integrating different data modalities to gain a comprehensive understanding of the molecular mechanisms underlying podocyte damage and FSGS progression.

## 3.2 Transcriptomic Analysis of Wt1<sup>ko/wt</sup> Model of Genetic FSGS

To study effects of Wt1<sup>ko/wt</sup> mutation and ensued progression of focal segmental glomerulosclerosis on glomerular transcriptome, we performed bulk RNA sequencing of isolated mouse glomeruli from wild-type and Wt1<sup>ko/wt</sup> mice, sampled at 4 weeks (early, non-sclerotic FSGS) and 12 weeks (late, sclerotic FSGS), in 3 biological replicates per group. Experimental design is visually summarized in Figure 3.2.1.A, complete information for bulk RNA-seq samples is provided in [Supplementary Table 2](#). Principal component analysis (PCA) of RNA-seq data from Wt1 heterozygous knock-outs and control mice at day 4 and day 12 age analysis showed that PC1 represents age differences and accounts for most variation in the data, whereas the effect of Wt1<sup>ko/wt</sup> mutation is best represented by PC3. PC2 could not be attributed to any experimental factor and, since the batch information was not available, was likely representing a batch effect. Upon removal of unwanted variation (see Material and methods 2.3.2) and performing PCA on the corrected data, the 2nd PC separates samples by the genotype (Figure 3.2.1 panel B). DE analysis of the denoised counts showed an improvement in distribution of p-values and increase in statistical power, yielding a comparable number of genes differentially expressed at 4 and 12 weeks ([Suppl.Fig.3.1 B](#)).

Differential expression analysis revealed that the effect of Wt1<sup>ko/wt</sup> mutation on the transcriptome is relatively minor compared to the effect of age: 738 and 7725 genes are significantly differentially expressed (FDR adjusted p-value <0.05) while testing the respective effects. Comparison of mutant and control samples separately in 2 disease stages found 623 and 751 genes differentially expressed in 4 and 12 week old mice, respectively, of which 332 genes are common for both comparisons. Notably, the mutation effect on the glomerular transcriptome gets slightly bigger at the later stage of FSGS (12 week), as reflected by a slightly wider distribution of LFC values on Figure 3.2.1 D. Overall, the genotype effect is relatively mild and consistent across both FSGS stages, as indicated by high correlation between vectors of gene expression log<sub>2</sub> fold changes (LFC), Pearson correlation coefficient = 0.75. These results are in line with Albumin/Creatinine measures of kidney function in mice of matching age and genotype that show a strong increase in Alb/Cre between 1 and 4 week-old and then no change between 4 and 12 week old Wt1<sup>ko/wt</sup> mice (see [Suppl.Fig.3.1 A](#)).

To analyze the effect of disease progression, while accounting for the effect of normal glomerular maturation and aging, we tested an interaction between age and genotype. Interaction analysis revealed just 9 genes that are differentially affected by Wt1 knock-down depending on age ([Suppl.Fig.3.1 B](#)), among which *Lamc2* and *P4ha3* are much stronger upregulated in Wt1<sup>ko/wt</sup> mice at 4 weeks, compared to 12 weeks. These few genes are involved in *collagen fibril organization* and *extracellular matrix organization* GO terms.

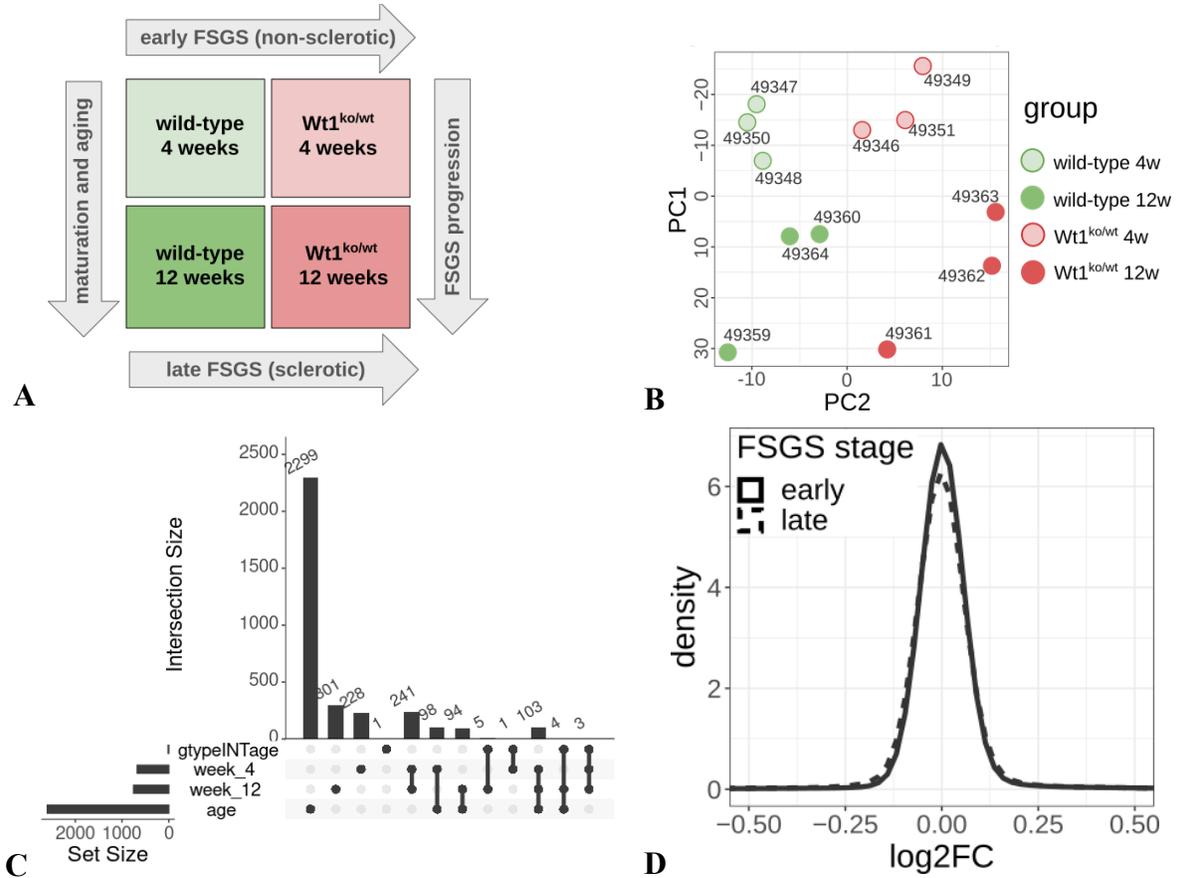


Figure 3.2.1: Analysis of bulk RNAseq data from glomeruli of Wt1 ko/wt mice.

(A) Scheme of the experimental design, arrows show pairwise comparisons between the groups. (B) Principal component analysis of the counts after RUV correction. (C) Comparison of sets of differentially expressed genes (in rows), where columns show number of unique or shared genes. (D) distribution of log<sub>2</sub> fold changes of gene expression in the analysis of early and late FGS.

### 3.2.1 Progressive FGS is Characterized by Cell Junction Disassembly and Increased Immune Activation

To characterize functional consequences of transcriptional changes in glomeruli of Wt1<sup>ko/wt</sup> mice, we annotated respective sets of differentially expressed genes with gene ontologies (GO) and molecular functions (Material and Methods 2.3.3), including KEGG, Reactome and focal adhesion<sup>3</sup> pathways. Results confirmed a high degree of functional similarity of genes perturbed at 4 and 12 weeks of FGS (Fig.3.2.2 A). The strongest down-regulated function is *focal adhesion in podocytes*, as seen in (Fig.3.2.2 B), which is especially strongly affected in late FGS and reflects podocyte detachment. Most other significantly enriched functions are

<sup>3</sup> A set of genes encoding podocyte-enriched focal adhesion complex is taken from (Schell et al. 2017) (Figure 1e)

upregulated at both stages of FSGS. Strongly enriched and upregulated are genes related to immune function, like *cytokine activity* and *humoral immune response*. Traditionally, immune system involvement in FSGS is associated with the adaptive immune system and T-cells. However, podocytes express Toll-like receptors which recognize pathogens, certain exo- and endogenous signals and are an integral part of the cellular innate immune defense system. It has been proposed that the podocyte's role in innate immunity may predispose it to injury, depending on chronicity (Issa et al. 2024; Burke et al. 2023).

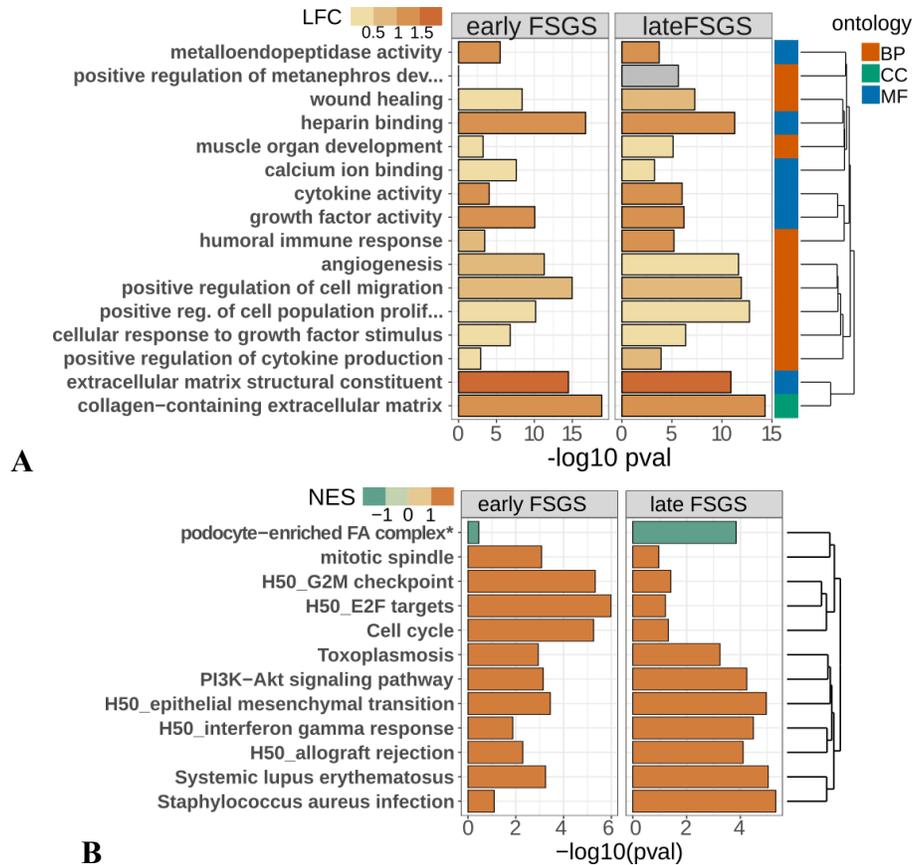


Figure 3.2.2: Histogram of GO (A) and pathway (B) annotations of genes differentially expressed in early (4 weeks) and late (12 weeks) stages of FSGS, compared to the wild-type. X-axis shows strength of the enrichment  $-\log_{10}(p\text{-value})$ , color of the bars shows average expression  $\log_2\text{FC}$  (panel A) or normalized enrichment score of GSEA (panel B). Union of top 10 most significantly enriched categories in each comparison was used for each figure. Terms are clustered by similarity of genes annotated under each term. Row annotation bar on the right side of GO histogram (panel B) color-code terms by GO category: BP - biological process. MF - molecular function, CC - cellular component.

Another group of upregulated processes, including *metalloendopeptidase activity*, contains terms related to the integrity of the extracellular matrix. Alteration in matrix metalloproteinase (MMP) activity was previously described in a number of renal diseases (Tveita, Rekvig, and Zykova 2008), it is also an important predictor of early diabetic nephropathy (Ghazala and

Taghreed U. Mohammd, n.d.). An increase in MMP activity can be triggered by inflammation signals and, in turn, continuous matrix degradation propagates inflammation signals, leading to chronic inflammation. Among other functions affected by *Wt1* knock-down are *calcium ion binding* and *heparin binding*. Tight  $\text{Ca}^{2+}$  regulation is necessary for maintaining glomerular integrity and filtration barrier (Vassiliadis et al. 2011) while specific endothelial heparin-binding EGF-like factor is a critical mediator in the initiation and progression of chronic kidney diseases caused by transactivation of EGF receptor (EGFR) by angiotensin II (Zeng et al. 2016). At the same time, studies have shown that the formation of stress fibers, which is necessary for normal podocyte function, is directly dependent on binding of syndecan-4 through its HS chains to the heparin-binding domain of fibronectin (Saoncella et al. 1999).

Upon general description of functions affected by the mutation we wanted to focus on differences between 2 stages of FSGS. Although only 9 individual genes were found to be significantly affected by the mutation in a disease stage specific manner, there might be a cumulative change on the level of functions and processes, detectable with the same dataset. To test this hypothesis, we performed what we called "2D GO analysis" (Material and Methods 2.3.3). First we calculated the average log fold change (LFC) of genes annotated under each Gene Ontology (GO) term, effectively summarizing the collective expression changes of genes within each biological category. We then compared these average expression changes of GO terms between early-stage and late-stage FSGS. By plotting the average LFCs of GO terms from the early stage against those from the late stage, we modeled the relationship between stage-specific expression changes using a simple linear regression. The figure 3.2.3 below illustrates results of this analysis: labeled terms have much greater average expression change in early or in late FSGS, correspondingly, and characterize FSGS progression in terms of immunity, developmental processes, extracellular matrix and other categories. Close inspection of 2D GO plot shows that *Vascular endothelial growth factor receptor binding* and *laminin complex* are stronger upregulated at the early FSGS. Several immune-related terms, like *complement activation*, and *cell junction disassembly* are stronger upregulated in late FSGS. *Fatty acid derivative biosynthetic processes* are stronger downregulated in late FSGS. Some developmental terms, such as *cell proliferation* or *regulation of metanephros development* are stronger upregulated at 4 weeks, while *mesenchymal cell development* is more downregulated at 12 weeks. Below we attempt to interpret observed functional changes in the context of FSGS progression.

VEGF is crucial for maintaining glomerular endothelial cell function and the integrity of the glomerular filtration barrier. Laminin complex is a member of GBM, its interaction with dystroglycan controls the shape of podocyte foot processes (Kojima and Kerjaschki 2002). Increased VEGF receptor binding and laminin complex expression in glomeruli of 4 week old mutants may represent attempts to compensate for *Wt1* haploinsufficiency and maintain glomerular integrity in early genetic FSGS. On the other hand, increase in the activity of cell junction disassembly and innate immunity genes in glomeruli at 12 weeks (late FSGS) likely contributes to podocyte detachment later in the disease progression. Downregulation of *fatty acid*

*derivative biosynthetic processes* at 12 weeks may also reflect progressive glomerular sclerosis, as certain fatty acid derivatives regulate the expression and activity of matrix metalloproteinases (MMPs), affecting integrity of the ECM and cell adhesion. Upregulation of cell-proliferation and levels of mesenchymal markers was previously associated with podocyte damage and was suggested to be mediated by Tead1 transcription factor (Xie et al. 2019). Interestingly, the magnitude of the effect is stronger earlier in the disease in the Wt1 FSGS model.

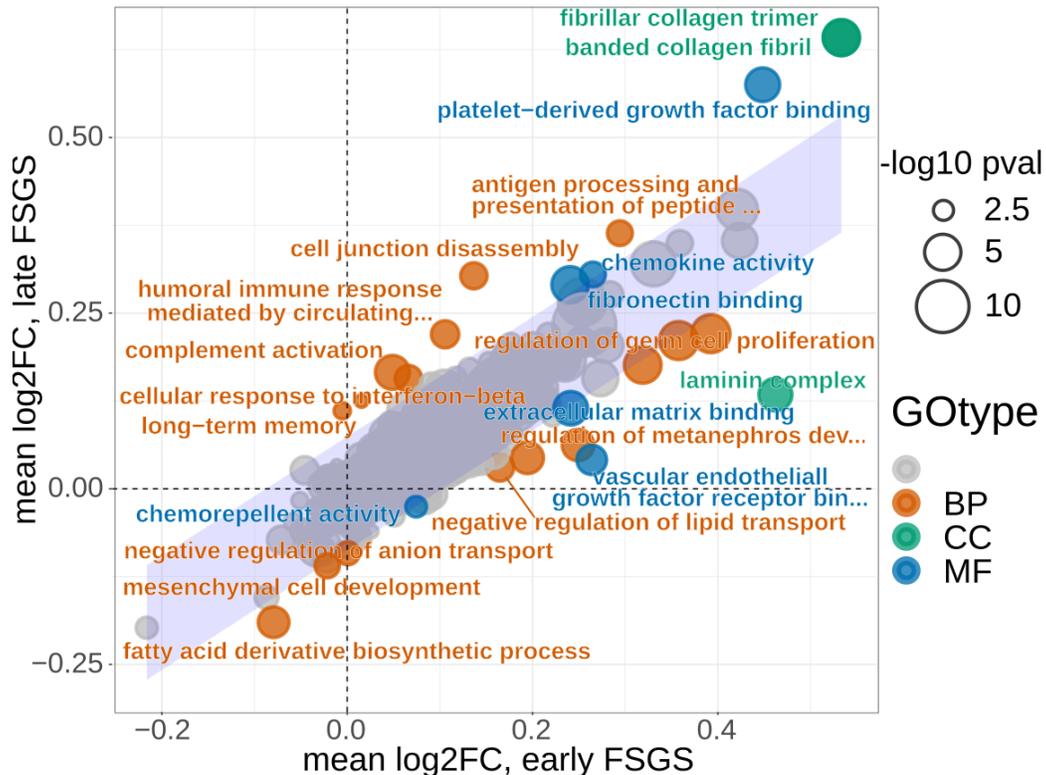


Figure 3.2.3: 2D GO plot compares 2 stages of FSGS, combining results of DE and functional annotation. Each circle represents a GO term, its color and size show GO category and  $-\log_{10}$  of the enrichment significance (max), respectively. X and Y axes show average LFC of genes in the GO term at early (4w) and late (12w) FSGS, respectively. The blue ribbon depicts the 99% prediction interval (PI) of the regression line: terms located outside the PI have greater than expected expression change of the contributing genes either in early or in late FSGS.

In summary, the GO terms exclusively enriched in early-stage FSGS likely reflect processes linked to Wt1 haploinsufficiency, while those enriched in late-stage FSGS are indicative of the increasing inflammation and sclerosis associated with advanced podocyte damage. GO terms enriched at both early and late stages may represent processes broadly involved in podocyte injury. As expected, the majority of the annotated GO terms align with key functions previously established as critical for podocyte biology.

### 3.2.2 Wt1 and Nphs2 Models of Genetic FSGS Share Transcriptional Changes

The Nphs2 R231Q/A286V mutation in the podocin gene induces podocyte damage and serves as a genetic model for FSGS. Mice with this mutation develop progressive albuminuria starting at 2 to 4 weeks of age with kidney failure at around 20 weeks of age. To investigate consequences of the mutation and ensued disease progression on transcriptional level, we generated and analyzed bulk RNA-seq data from glomeruli of healthy and Nphs2<sup>mut</sup> mice, sampled at 4 weeks (early, non-sclerotic FSGS) and 8 weeks (late, sclerotic FSGS) of age. 3 and 5 biological replicates per genotype were generated from 4 and 8 week old mice, correspondingly. Complete information for bulk RNA-seq samples is provided in [Supplementary Table 2](#). PCA analysis of rlog-normalized RNA-seq counts showed that samples are well separated by age and condition with 2 first PCs and that 8 weekers are substantially different from both control and 4 week old experimental samples ([Suppl.Fig3.2 A](#)).

Differential expression analysis revealed 6704 genes that are significantly differentially expressed (FDR adjusted p-value <0.05) between all control and Nphs2<sup>mut</sup> samples. Comparison of mutant and control samples separately in 2 disease stages found 4252 and 6954 genes differentially expressed in 4 and 8 week old mice, respectively, of which 2106 genes are common for both stages ([Suppl.Fig3.2 B](#)). The high number of DE genes reflect a strong effect of Nphs2 mutation on glomerular transcriptome, while higher number of DE genes at the later stage of FSGS reflects the effect of disease progression.

Mice with the Nphs2 mutation develop severe phenotype faster compared to Wt1<sup>ko/wt</sup> mice, so that at the age of 4 weeks they are already very proteinuric. The faster disease progression in Nphs2<sup>mut</sup>, compared to Wt1<sup>ko/wt</sup> mice, is clearly reflected on the transcriptome level by a greater number of DEGs and a greater magnitude of changes (broader distributions of LFC) in the respective tests ([Fig.3.2.4 A](#)).

Despite a different degree of the transcriptome remodeling, caused by each mutation, patterns of gene expression changes (log<sub>2</sub>FC vectors) significantly correlate between 2 models ([Fig.3.2.4 B](#)). Interestingly, there is convergence of the models at the later disease stage. log<sub>2</sub> fold changes of gene expression in Nphs2<sup>mut</sup> at 8 weeks (late FSGS) correlate stronger with LFCs from corresponding analysis of Wt1<sup>ko/wt</sup> experiment, than with LFCs in Nphs2 at 4 weeks.

Analysis of interaction between genotype and age showed that 3332 genes are associated with the disease progression, i.e. the effect of Nphs2 mutation on their expression significantly (q-value < 0.05) depends on the age of mice. This is different from Wt1<sup>ko/wt</sup> model where the effect of disease progression on the glomerular transcriptome is relatively mild. Interestingly, decrease of Wt1 expression at the age of 4 weeks (early FSGS) does not reach the significance level (FDR-adjusted p-value < 0.05) in either of 2 models ([Fig.3.2.4 C](#)). Similarly, Nphs2 gene expression is significantly decreased in Nphs2<sup>mut</sup> mice only at the later and not at the early stage of FSGS where it even increases ([Fig.3.2.4 C](#)).

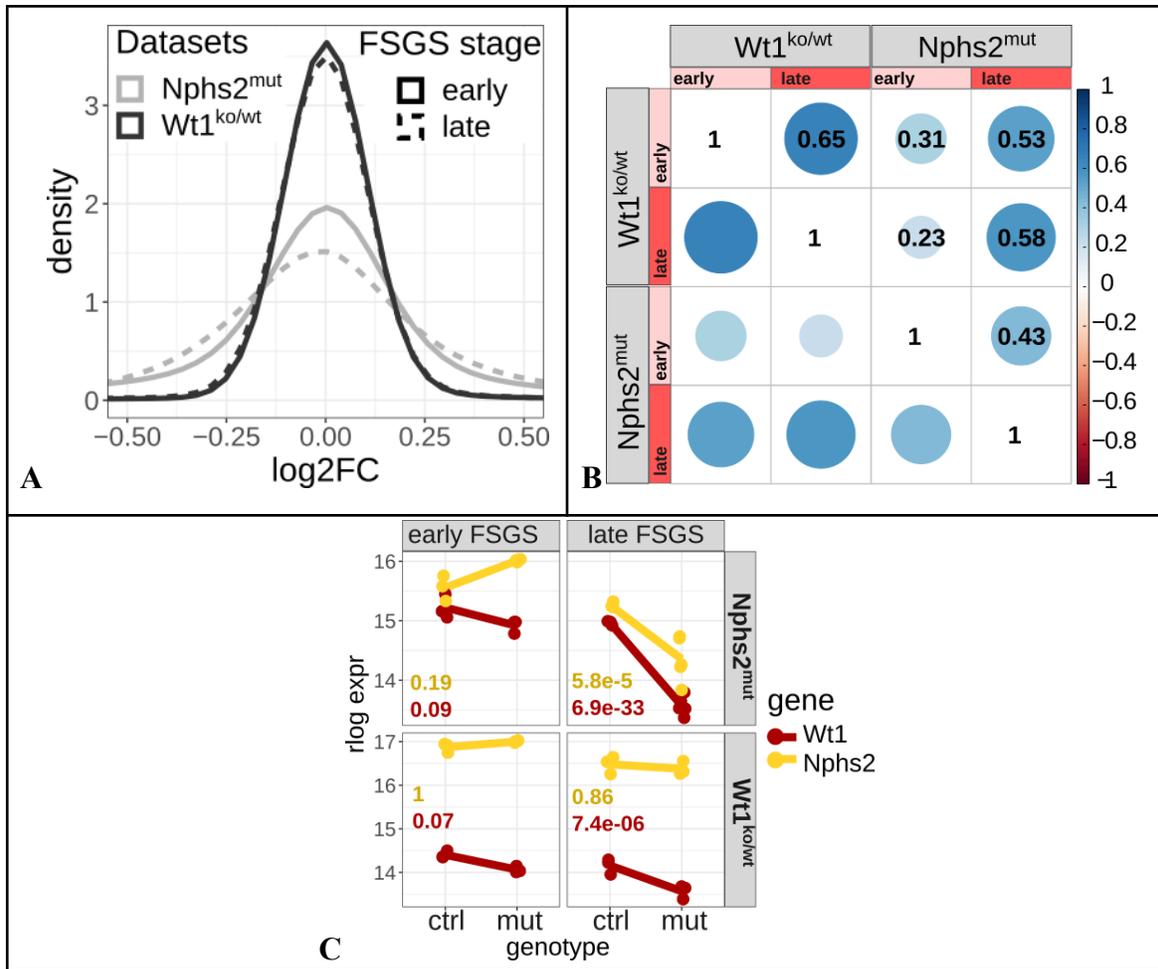


Figure 3.2.4: Comparison of differential expression in Wt1<sup>ko/wt</sup> and Nphs2<sup>mut</sup> bulkRNA-seq datasets. (A) Distributions of log<sub>2</sub> fold changes in the 4 comparisons. (B) Correlation heatmap of differential expression results, calculated using spearman correlation of shrunk log<sub>2</sub>FC, only the top 30% of variable genes was used. Early and late FSGS is denoted by light and dark red annotation bars, correspondingly. (C) Expression of Wt1 and Nphs2 genes, denoted by red and yellow, in glomerular samples of 2 models (panel row) at 2 disease stages (panel columns). Adjusted p-values show results of DEseq2 test between control and mutant samples.

To compare functional consequences Wt1 and Nphs2 mutations, we annotated genes differentially expressed between mutant and control mice, at 4 and at 12 weeks, with biological pathways and GO terms. Results show that many biological processes related to extracellular matrix, collagen biology and immunity are affected in both disease stages of Wt1 and Nphs2 models (Fig.3.2.5 A,B). There are common trends of increased innate and adaptive immune responses and a striking downregulation of *podocyte-enriched focal adhesion* genes in late FSGS, in both models. These trends may reflect a positive feedback loop among the glomerulus, platelets, and immune cells that may contribute to persistent glomerular damage. At the same time, there are pathways and ontologies that differentiate 2 FSGS models. For instance,

cell-cycle related gene sets are stronger upregulated at early FSGS or at late FSGS in *Wt1*<sup>ko/wt</sup> and *Nphs2*<sup>mut</sup>, correspondingly (Fig.3.2.5 B). There is also a difference in ranking of significant terms: innate and adaptive immune responses are much stronger in *Nphs2*<sup>mut</sup> glomeruli, which reflects their infiltration by immune cells already in 4 week old mice, while *collagen-containing extracellular matrix* has higher rank and upregulation in *Wt1*<sup>ko/wt</sup> glomeruli at both disease stages (Fig.3.2.5 A).

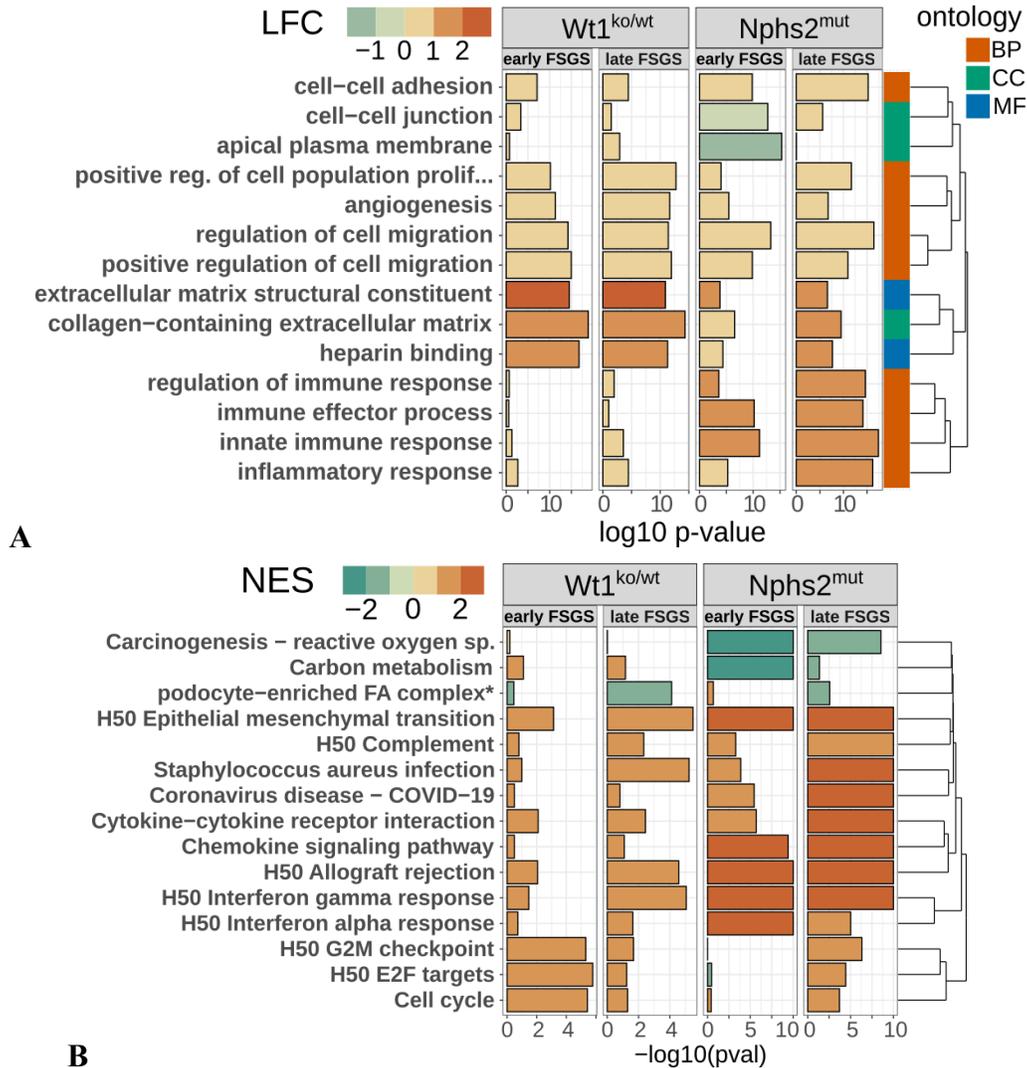


Figure 3.2.5: Functional annotation of differentially expressed genes.

Histograms of GO (A) and biological pathway (B) annotations of genes differentially expressed in early (4 weeks) and late (4 weeks) stages of FSGS in *Wt1*<sup>ko/wt</sup> and *Nphs2*<sup>mut</sup> models of FSGS. X-axis shows -log<sub>10</sub> p-value of the enrichment test, color of the bars shows average expression log<sub>2</sub> fold change (panel A) or normalised enrichment score of GSEA (panel B). Union of top 10 most significantly enriched categories in each comparison was used for each figure.

To highlight differences in biological effects of *Nphs2* and *Wt1* perturbations on glomerular transcriptome we performed 2D GO analysis (see the previous sub-chapter) and visualized the

result on Figure 3.2.6. GO terms that best differentiate effects of genotypes on glomerular transcriptome can be split in 3 groups: 1) processes related to innate immunity and 2) mitochondria-related processes that are stronger up- and down-regulated in  $Nphs2^{mut}$ , correspondingly; 3) functions related to morphogenesis and glomerular development that are upregulated in  $Wt1^{ko/wt}$ .

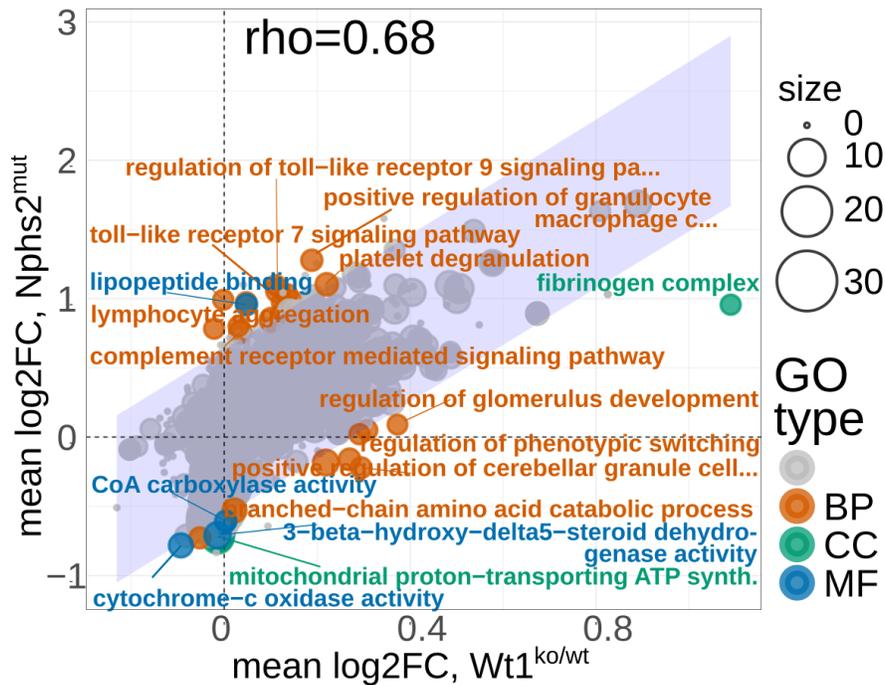


Figure 3.2.6: 2D GO plot shows functional comparison of transcriptomic changes in glomeruli of  $Wt1^{ko/wt}$  and  $Nphs2^{mut}$  mice. X and Y axis show mean gene expression log2FC of GO terms, where log2FC are taken from the analysis of the genotype effect of  $Wt1^{ko/wt}$  and  $Nphs2^{mut}$  mutations, respectively. Each circle represents one GO term, color and size of the circle show its GO category and significance of the Fisher test, correspondingly, GO terms significant ( $pval < 0.05$ ) in early and/or late FSGS are used. The blue ribbon depicts the 99% prediction interval (PI) of the regression line: terms located outside the PI have greater than expected expression change of the contributing genes in a respective model.

Early activation of innate immunity in  $Nphs2^{mut}$  is expected and caused by the mutation in the podocin gene that disrupts the slit diaphragm, leading to exposure of antigens. Downregulated mitochondrial processes are more difficult to interpret, it may reflect impaired energy metabolism contributing to podocyte injury and apoptosis. Upregulation of developmental and morphogenesis pathways in  $Wt1^{ko/wt}$  data emphasizes transcriptional dysregulation caused by  $Wt1$  haploinsufficiency and may represent an attempt by podocytes to compensate for  $WT1$  deficiency but may also be an artifact of non-podocyte signal in bulk glomerular RNAseq. To sum up, the functional annotation of differentially expressed genes in the two genetic models of FSGS reveals lots of commonalities and some distinctions in the molecular pathways and functions affected. Alterations in ECM and collagen biology in both

models indicate shared mechanisms of GBM remodeling and scarring. Increased immune response and podocyte detachment suggest that inflammation and loss of adhesion contribute to disease progression in both models. *Nphs2*<sup>mut</sup> is characterized by strong innate immune activation and mitochondrial dysfunction, which contributes to inflammatory damage and energy deficits. *Wt1*<sup>ko/wt</sup> exhibits upregulation of developmental and morphogenesis pathways, reflecting transcriptional dysregulation due to WT1 haploinsufficiency.

### Analysis of transcription factor expression

Finally, to investigate the broader spectrum of possible regulators of expression changes we scrutinized expression of individual transcription factors (TF) in our bulk RNAseq samples. Among 110 TFs that were previously found expressed in podocytes, *Wt1* and *Zfp423* are significantly downregulated in both models at both stages (Fig.3.2.7), supporting the crucial role of *Wt1* in the podocyte gene regulatory network. Interestingly, transcription factor *Zfp423* is differentially expressed in all 4 comparisons, this TF has been previously shown to control proliferation and differentiation of neural precursors (Alcaraz et al. 2006) but not much information is available regarding its role in glomerular damage response. At later FSGS stages in both models *E2f1*, *Lmx1b*, *Tcf21* and *Foxd1* are downregulated while *Maff* and *Egr1* are upregulated. *Egr1* is a TF that shows consistent stage dependent behavior in both models: downregulated in early and upregulated in late FSGS

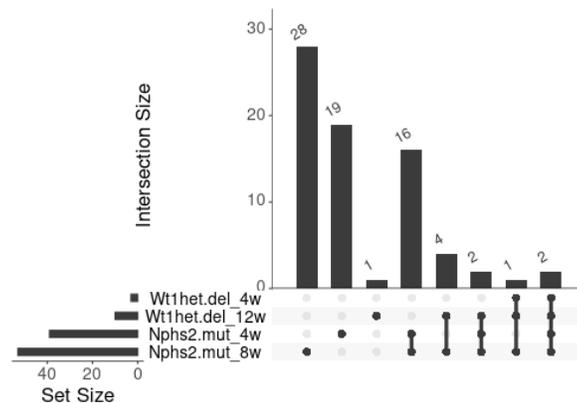


Figure 3.2.7: Overlap of transcription factors expressed in wild-type podocytes and significantly differentially expressed (q-value <0.1) in 2 disease stages of 2 FSGS models.

Overall results of differential expression analysis of bulk glomerular RNA-seq in 2 models of genetic FSGS shed some light on transcriptional rewiring in damaged podocytes, highlighting similarity of transcriptional responses but also presenting features specific to each individual mutation. Despite the progress, the mixture of signals from various cell types in bulk glomerular samples makes it difficult to disentangle transcriptional changes in podocytes from transcriptional changes happening in other glomerular cells, from changes in cell-type composition. This issue is addressed in the following chapters by performing Chip-seq (Chapter 3.3) and single-nuclei RNAseq (Chapter 4.2) experiments.

### 3.3 Chip-seq Analysis of Transcriptional Rewiring in Damaged Podocytes

Differential expression analysis of  $Wt1^{ko/wt}$  bulk RNA-seq glomerular data has several limitations when used to characterize remodeling of  $Wt1$  transcriptional regulatory network (TRN) in damaged podocytes, as it is likely confounded by (1) changes in cell-type composition in the damaged glomeruli and (2) a mixture of signals from primary and secondary, down-stream effects, mediated via other transcription factors or coming from other cell-types. To address these limitations and characterize rewriting of podocyte TRN, we performed Chip-seq analysis of WT1 itself and TEAD1 in wild-type and  $Wt1^{ko/wt}$  glomeruli.

The heterozygous  $Wt1$  knockout mouse model aka  $Wt1^{ko/wt}$ , a murine model of genetic FSGS, has a consistent phenotype of proteinuria, foot process effacement and histologic features of FSGS starting at the age of 4 weeks. We conducted WT1 and TEAD1 ChIP-seq experiments on glomeruli from  $Wt1^{ko/wt}$  and wild-type mice at the early disease stage, at the age of 8 weeks. 3 replicates per each condition per protein and corresponding mock IPs were sequenced. Peaks in individual replicates were called with MACS2 and subject to cross-correlation analysis. Results of clustering of individual peak-sets show grouping of biological replicates by the genotype in both TF experiments with the exception of replicates #2 from WT1 experiment (Figure 3.3.1 A). Results of cross-correlation and clustering analysis prompted us to exclude replicates #2 from the downstream analysis due to the low quality. On the other note, in both experiments samples from  $Wt1^{ko/wt}$  mice have higher biological variability than the wild-type samples.

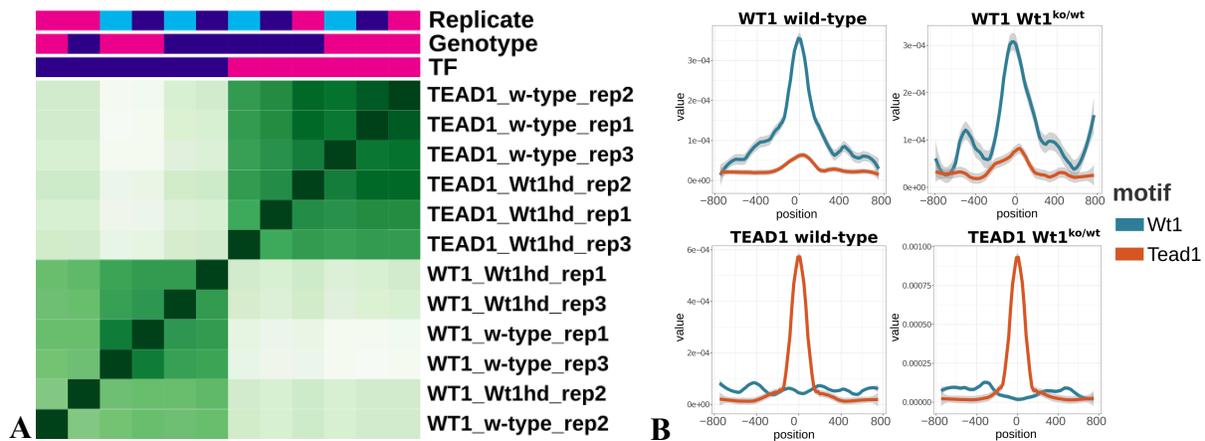


Figure 3.3.1: Quality of WT1 and TEAD1 Chip-seq experiments.

(A) Hierarchical clustering of biological samples from WT1 and TEAD1 experiments. Spearman correlation of Chip signals is used as the distance measure and reflected by the heatmap color, signal is normalized by input control. (B) Density of  $Wt1$  and  $Tead1$  primary motifs around centers of WT1 and TEAD1 peaks called by MACS2 in wild-type and  $Wt1^{ko/wt}$  samples.

Consensus peak-sets were derived with MSPC software and were used for motif analysis and peak annotation. 16402 and 8278 consensus WT1 peaks were called in wild-type and  $Wt1^{ko/wt}$  background; 31382 and 9206 consensus TEAD1 peaks were called in wild-type and  $Wt1^{ko/wt}$  background, correspondingly. Despite a smaller number of consensus peaks for  $Wt1^{ko/wt}$  samples, due to higher variability, de-novo motif analysis of wild-type and  $Wt1^{ko/wt}$  consensus peaks shows similarity of their motif composition in WT1 and TEAD1 experiments (Suppl.Fig.3.4). Moreover, the most enriched de-novo motif in each consensus peakset have the highest similarity with the known primary motifs of the corresponding IP-ed protein. Densities of primary  $Wt1$  and  $Tead1$  motifs around centers of WT1 and TEAD1 peaks from wild-type and  $Wt1^{ko/wt}$  samples are shown on Figure 3.3.1 B. Consensus Chip-seq peak sets are provided in the github folder (“Website,” n.d.-c). Overall, initial analysis of the peak-sets showed that WT1 and TEAD1  $Wt1^{ko/wt}$  experiments worked as intended and the data is suitable for differential binding analysis.

### 3.3.1 WT1 and TEAD1 Chip-seq Analysis Reveal Coordinated and Independent Gene Regulation in Healthy Podocytes

Understanding the transcriptional regulatory network governed by  $Wt1$  in the context of healthy podocytes is essential for appreciating how its disruptions might lead to disease. Upon identification of peaks and before differential binding analysis of FSGS samples, we performed comparative analysis of WT1 and TEAD1 peaks identified in wild-type glomeruli, which we further refer to as WT1wt and TEAD1wt peaks.

To get a visual impression of the TFs binding behavior we manually screened several individual podocyte genes in a genome browser, visualizing input-normalized Chip-seq signals from the two transcription factors. An example of WT1 and TEAD1 signals over the podocin gene in healthy glomeruli is shown on Figure 3.3.2 A and is representative of several other podocyte genes. As evident from the example, both TFs often exhibit similar binding profiles, with overlapping peaks of similar intensities and closely located summits. Checked globally, we found a significant overlap between WT1wt and TEAD1wt peak sets resulting in 10640 shared peaks (Fig.3.3.2 B), which suggests co-binding of these TFs.

Next, we have separately analyzed sites co-bound by two TFs and uniquely bound by each of the TFs. Genome region annotation shows that co-bound regions are more enriched in promoters/exons/5-prime-UTR (different annotation levels), compared to solo WT1 and TEAD1 binding sites that are mostly positioned in intronic and intergenic space, i.e. putative enhancers (Fig.3.3.2 C,D). Analysis of Chip signal with the aid of metagene plots shows that co-bound regions have strong and perfectly centered IP signals from both proteins (Fig.3.3.2 D, 2nd row), providing additional evidence of WT1 and TEAD1 co-binding. Interestingly, regions uniquely bound by WT1 or TEAD1 have, beside a strong signal from the corresponding IP-ed protein, a very weak signal from the other TF protein (Fig.3.3.2 D, 3rd row). The nature of this weak signal is interesting, especially in the case of non-specific TEAD1 signals in regions bound only by

WT1, as these peaks are not enriched in the TEAD1 motif (Fig.3.3.6 A). We hypothesize that if not directly co-binding, one TF can affect or be affected by the chromatin state, allowing a low-level binding of another TF at the same position: a molecule of a primary TF can be connected, possibly via cofactor, to a molecule of a secondary TF, which physically binds at a region distant from the site of the primary TF binding. In this scenario, a residual binding of the second TF at a given location is a “shadow” of a primary binding event at another location. It may as well be an artifact of the experiment or the analysis pipeline.

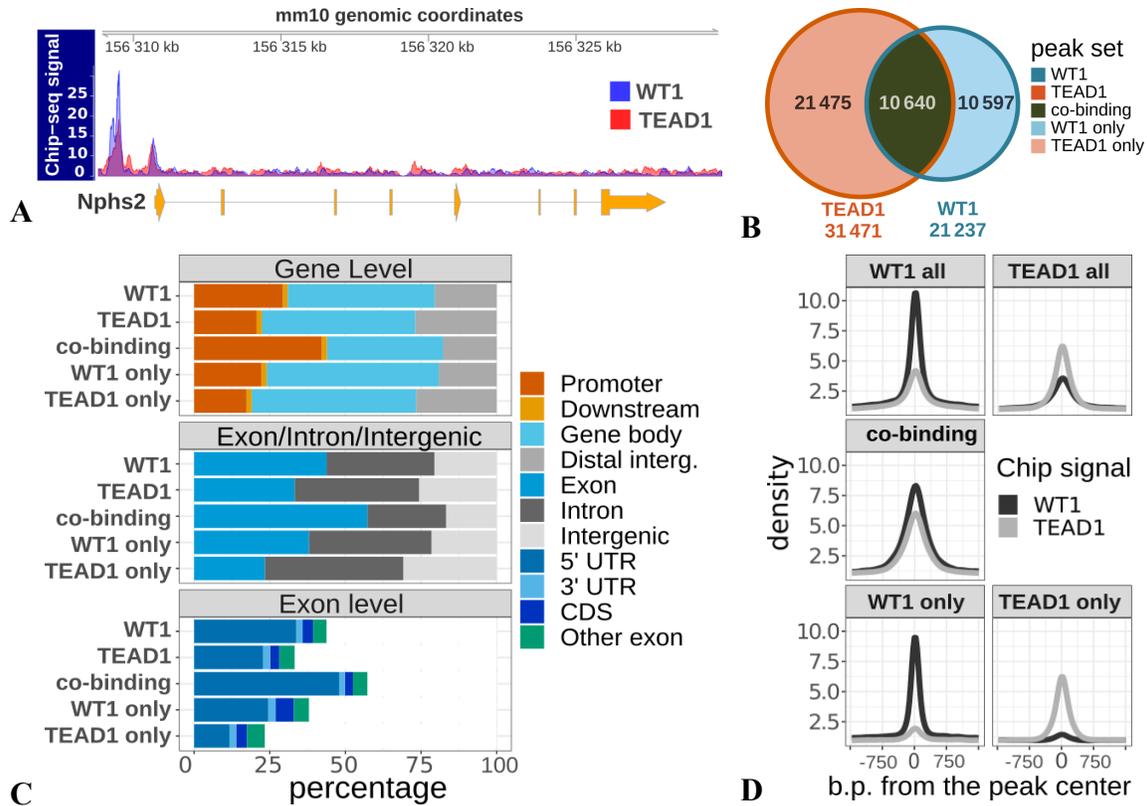


Figure 3.3.2: Integrative analysis of Wt1 and Tead1 Chip-seq datasets.

(A) Input-normalized Chip signal of WT1 and TEAD1 binding around the podocin gene shows clear overlap of signals from both TFs at Nphs2 promoter. (B) Overlap of Tead and Wt1 peaks in the wild-type glomeruli. (C) Genome region annotation of wild-type and differentially bound peaks. (D) Chip signal in sets and subsets of Wt1 and Tead1 Chip-seq peaks.

### Motif analysis of differentially bound regions

Figure 3.3.3 depicts results of known-motif enrichment analysis of regions co-bound and uniquely bound by WT1 and TEAD1. Results show that all 3 sets of regions have similar motif compositions and are enriched in Wt1 and other homeobox TFs, several members of Fox, Irf and Ets families, Rela and other primary motifs. At the same time, regions uniquely bound by each TF also exhibit unique characteristics. Regions bound only by TEAD1 lack enrichment in Lmx1b and a number of Hox motifs and are more enriched in Jun/Fos motifs, compared to

regions co-bound and bound only by WT1. Regions bound only by WT1 and co-bound by both TFs are characterized by lack of enrichment in Tead1 and Nfia motifs, correspondingly.

Strong association of TEAD1 with Jun/Fos motifs implies its role in enabling podocytes to respond to mechanical stress and environmental changes since Jun/Fos proteins form the AP-1 transcription factor complex, regulating genes involved in response to stress and extracellular stimuli. Lack of Lmx1b and Hox Motifs in regions bound only by TEAD1 indicates that TEAD1 regulates genes independently of the developmental pathways governed by Lmx1b and Hox genes. Hence TEAD1 may focus on dynamic cellular processes rather than maintaining podocyte identity. Lack of Tead1 motif in regions bound only by WT1 suggests that WT1 can independently regulate genes without TEAD1's involvement, focusing on the core aspects of podocyte structure and function. Overall, the absence of certain motifs in regions bound by one TF suggests that WT1 and TEAD1 can independently regulate distinct sets of genes, allowing for specialized control over different cellular processes. Regions co-bound by both TFs may represent a critical set of genes regulated collaboratively, potentially essential for fundamental podocyte functions. This specialization within the transcriptional network enables podocytes to maintain their unique structural and functional properties while adapting to physiological demands.

Additional angle of the analysis was provided by classifying peaks as cis- or trans-regulatory elements, based on distance to TSS, and testing motif enrichment separately in these subsets. Results of such analysis are shown on heatmap in Figure 3.3.3 B, where pink and green colors show motif enrichment in regions bound by the TF and classified as cis or trans-regulatory elements, respectively. The heatmap presents a similar pattern of cis/trans motif enrichment in WT1 and TEAD1 peaks, except for the striking difference in Wt1 and Klf6/15 TF motifs that are relatively enriched in trans- and cis-regulatory regions in WT1wt and TEAD1wt regions, correspondingly. This observation is supporting the role of WT1 as a cofactor necessary for expression of the core TEAD1 target gene.

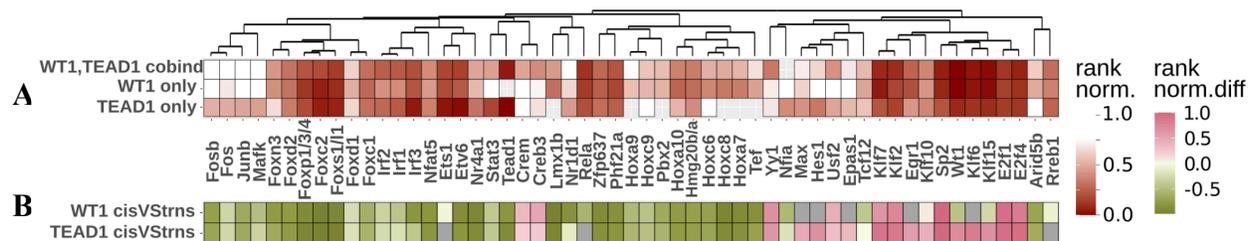


Figure 3.3.3: Known motif enrichment analysis of genomic regions co- and uniquely bound by TEAD1 and WT1 in wild-type glomeruli, (A) tested against the shuffled background, (B) tested in cis- and trans-regulatory elements in a discriminative mode. Color on the first heatmap shows the normalized rank of enrichment and color on the second heatmap shows the difference between enrichments, i.e. pink and green colors show relative enrichment in cis and trans-regulatory peaks, respectively.

In summary, CHIP-seq analysis revealed a complex interaction between WT1 and TEAD1 in healthy podocytes and provided necessary context for analysis of TRN rewiring under podocyte damage, which is described in the next sub-chapter.

### 3.3.2 WT1 Knock-Down Differentially Impacts WT1 and TEAD1 Binding

Having analyzed WT1-dependent transcriptional network in healthy podocytes in the previous chapter, here we investigated rewiring of the network in  $Wt1^{ko/wt}$  mouse model of genetic FSGS. To this end we performed differential binding analysis of WT1 and TEAD1 transcription factors in  $Wt1^{ko/wt}$  mice. PCA analysis of biological samples (Fig. 3.3.4 A,B) confirmed genotype as the main grouping variable, both in WT1 and TEAD1 experiments, so we proceeded with differential binding analysis using DiffBind R package (“Website,” n.d.-c; Rory Stark<rory.stark@cruk.cam.ac.uk>, Gord Brown<gdbzork@gmail.com> 2017) and identified 2106 WT1 and 655 TEAD1 peaks that are significantly ( $q$ -value < 0.05) differentially bound in  $Wt1^{ko/wt}$  compared to wild-type mice. WT1 and TEAD1 differentially bound regions are further referred to as WT1db and TEAD1db, correspondingly. Sets of Differentially bound peaks are provided in the supplementary github folder.

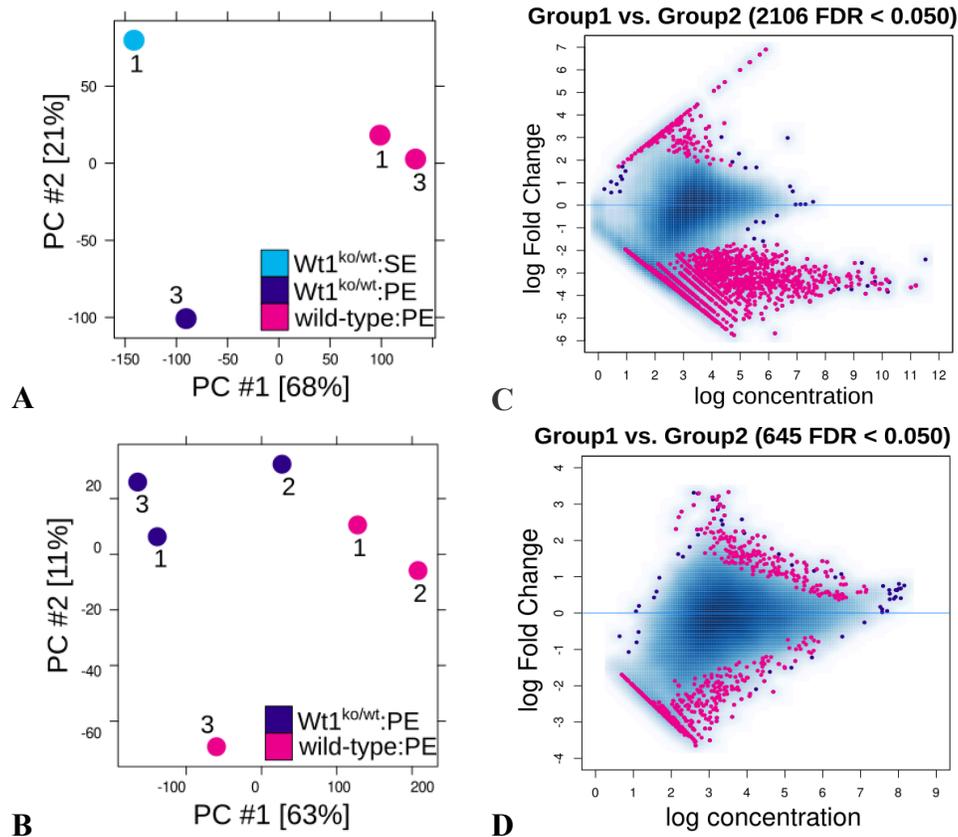


Figure 3.3.4: Effect of  $Wt1^{ko/wt}$  on WT1 and TEAD1 binding.

PCA plots show grouping of wild-type and  $Wt1^{ko/wt}$  samples in WT1 (A) and TEAD1 (B) experiments, plots are made using RPKM of ChIP divided by RPKM of input, dots are labeled by replicates and coloured by genotype. MA plots of changes in normalized ChIP signal show that a number of WT1 (C) and TEAD1 (D) binding sites undergo significant change upon  $Wt1$  knock-down. Sites identified as significantly differentially bound (FDR-adjusted  $p$ -value < 0.05) shown in magenta.

Analysis of MA plots, showing log<sub>2</sub> fold change of reads (Y-axis) relative to the total number of reads (X-axis) at the given peak, suggests that the overall effect of Wt1 knock-down on WT1 binding is stronger than on TEAD1 binding, as can be seen on Figure 3.3.4 C,D. There is also a difference in the ratio of up and down-bound regions: the absolute majority of WT1db peaks (1931 of 2106) lose Chip signal while the bias towards down-binding is weaker for TEAD1db where 428 and 227 peaks lose and gain binding, respectively. In addition, the average Chip signal in peaks that lose WT1 binding is higher than the average signal in all WT1 peaks (Fig.3.3.4 C), whereas Chip signal in TEAD1 peaks that lose binding is lower compared to all TEAD1 peaks (Fig.3.3.4.D).

Distribution of distances from peaks to the nearest TSS (Figure 3.3.5, panel A,B) implies that most changes in binding of both TFs occur in trans-regulatory regions. Indeed, genome region annotation of differentially bound peaks reveals depletion in promoter and enrichment in intronic/intergenic regions ([Supplementary Figure 3.4](#)), compared to the wild-type. This bias is relatively stronger in WT1db compared to TEAD1db: only 6% and 16% of differentially bound peaks are annotated as promoter regions, compared to 32% and 23% of wild-type binding in WT1 and TEAD1 experiments, respectively. Separate analysis of subsets of differentially bound peaks that gain (up-bound) and lose (down-bound) Chip signal show differences in their genomic composition, especially strong in WT1 experiment: up-bound peaks are more similar to wild-type binding while down-bound peaks are more enriched in trans-regulatory regions (Figure 3.3.5 C,D).

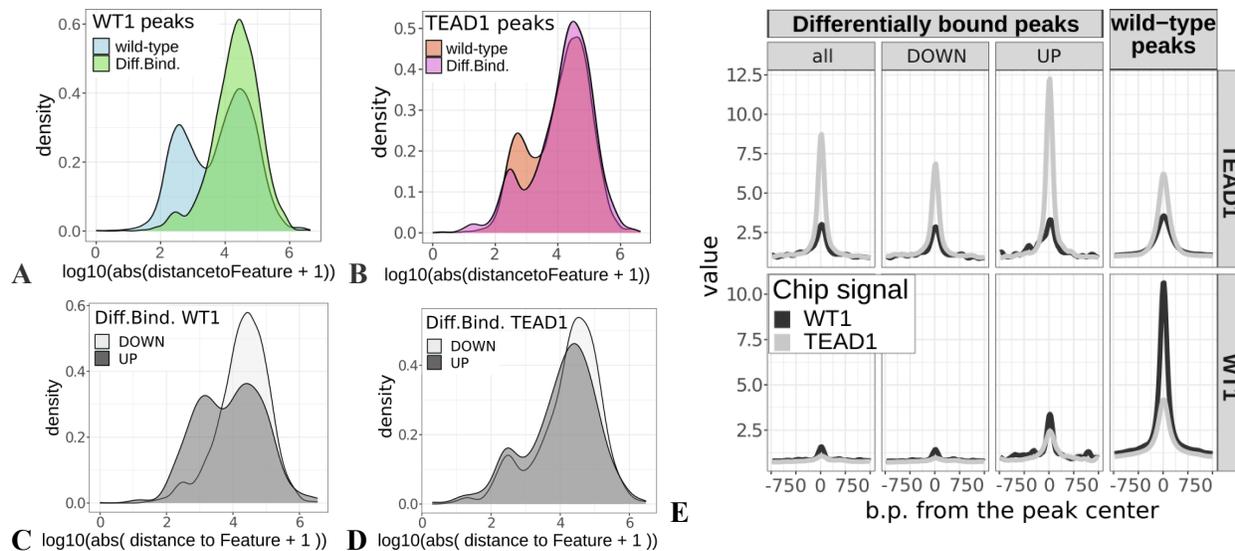


Figure 3.3.5: Annotation of WT1 and TEAD1 differential binding in wild-type vs Wt1<sup>ko/wt</sup> glomeruli. Distribution of distances between the center of each peak and the nearest annotated TSS for all differentially bound peaks (A,B) and for Up- Down-bound subsets (C,D) of TEADdb and WT1db, respectively. (E) Normalized WT1 and TEAD1 Chip signals in wild-type and differentially bound peaks.

The analysis of wild-type WT1 and TEAD1 ChIP-seq signals in regions differentially bound by these TFs upon Wt1 knock-down (Figure 3.3.5 E) reveals several findings. First, it is evident that signals from both TFs are present in WT1db and in TEAD1db regions, with higher signal from the corresponding IP-ed protein. Second, both signals are markedly diminished in WT1db, compared to the average wild-type levels, while in TEAD1db both signals are approximately on the wild-type levels. Moreover, an additional layer of complexity is observed when comparing peaks down- and up-bound in Wt1<sup>ko/wt</sup> glomeruli: down-bound regions exhibit relatively weaker ChIP signals compared to up-bound regions, with a much stronger difference in the case of WT1.

### **Motif analysis of differentially bound regions**

Further characterisation of podocyte TRN rewiring under damage involved motif analysis of differentially bound regions. Figure 3.3.6 A shows density of primary Wt1 and Tead1 motifs in regions differentially bound by WT1 and TEAD1 in glomeruli of Wt1<sup>ko/wt</sup> mice. Results show that TEAD1db regions are enriched in both primary motifs, while WT1db regions are enriched only in Wt1 motif. In both sets of db regions the peak of Wt1 motif enrichment is shifted ~100b.p. from the center. To find other transcription factors that may be involved in rewiring of podocyte TRN upon damage we performed motif enrichment analysis of differentially bound regions using known motifs of TFs expressed in podocytes. Results show that motif composition of genomic regions differentially bound by WT1 upon Wt1 knock-down resembles motif composition of WT1 wild-type peaks but has notable differences. Both sets of genomic regions are enriched in zinc finger protein motifs like Wt1, Sp2, Klf, Zfp637 etc (Fig.3.3.6 B) and both have unique motifs, e.g. Nr6a1 motif is enriched only in WT1db and Tead1 - only in WT1wt.

To investigate these differences further we tested motif enrichment of differentially bound regions in a discriminative mode, using a subset of WT1wt peaks selected to have the same peak-to-TSS distance distribution as WT1db peakset. This was done to control for genomic element composition differences between WT1db and all WT1wt peaks. Results on Figure 3.3.6 C reveal that most motifs are differentially enriched between the peaksets: Wt1, Klf and similar motifs are depleted (blue bars) while Lmx1b, Zfp637, Fox and Hox motifs are enriched (orange bars) in WT1db, compared to the matched WT1wt peaks. Motif enrichment of TEAD1db peaks is quite similar to TEAD1wt peaks: both are enriched in Tead1, Klf, Etv6, Fox, Wt1 motifs (Fig.3.3.6 B). Motif enrichment of TEAD1db in a discriminative mode against matched TEAD1wt peak-set (see explained for WT1 above) shows almost no differences in enrichment between wild-type and differential bound peaks, except for a relative enrichment in the primary Tead1 motif.

Comparative enrichment of up- and down-bound peaks against each other reveal minor differences between these subsets both in WT1db and in TEAD1db peaks. Up-bound peaks are relatively enriched in the corresponding primary motifs (Fig.3.3.6 C,D yellow bars) and down-bound - in motifs of other podocyte-expressed TFs, like Lmx1b, Zfp637, Foxc2 etc (Fig.3.3.6 C,D green bars).

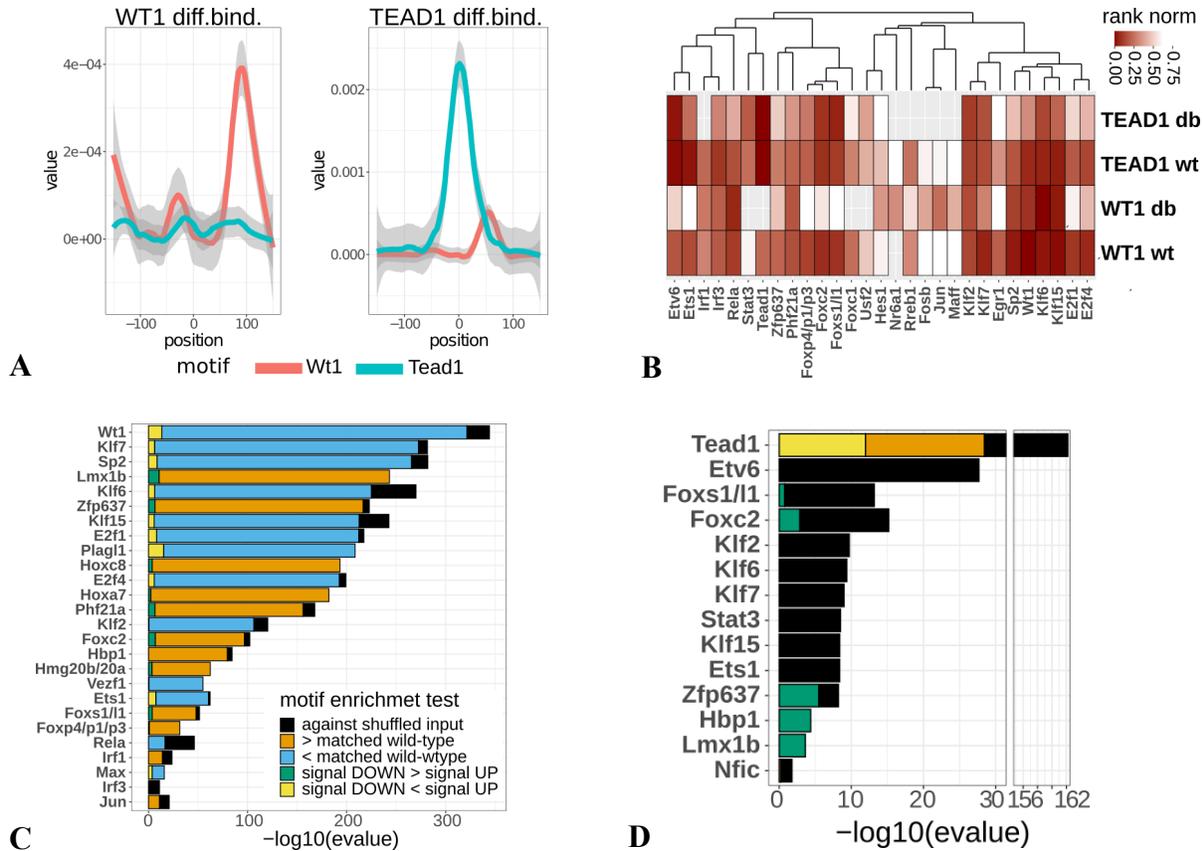


Figure 3.3.6: Motif analysis of differentially bound peaks.

(A) Density of selected primary motifs around centers of differentially bound regions. (B) Heatmap of motifs most enriched in regions bound by WT1 and TEAD1 in wild-type and differentially bound in  $Wt1^{ko/wt}$  glomeruli. The enrichment is tested against the shuffled input, color shows the normalized rank of enrichment, motifs are clustered by the PWM similarity. Union of 20 motifs per peak-set is used for the figure. Barplots show relative motif enrichment of WT1 (C) and TEAD1 (D) differentially bound regions. Colors of bars represent various tests performed: black color shows motif enrichment of differentially bound regions compared to shuffled input, orange and blue denote motif enrichment or depletion of differentially bound relative to wild-type bound regions, green and yellow show motif enrichment of up- and down-bound regions relative to each other. X-axis shows  $\log_{10}$  transformed e-value of enrichment (p-value divided by the number of motifs tested), rows show TF motifs tested, union of top 10 enriched motifs per test is shown.

To sum up, both TEAD1 and, especially, WT1 binding are affected by  $Wt1$  knock-down but in different ways. Given that majority of WT1db regions are 1) annotated as trans-regulatory elements (Fig.3.3.5 A,B,E), 2) lose WT1 signal (Fig.3.3.5 F), 3) depleted in  $Wt1$  while enriched in other podocyte specific TF motifs (Fig.3.3.6 C), compared to WT1wt, we can assume that decrease in WT1 concentration preferentially affects WT1 target genes that are regulated via podocyte specific enhancers collaboratively bound by WT1 and other TFs, such as *Lmx1b*, *Zfp637* and *Foxc2*. This would imply that gene regulatory circuits under control of these TFs are directly affected by the  $Wt1$  knock-down. *Lmx1b* and *Foxc2* cooperatively regulate podocin

expression through a shared enhancer motif (<https://doi.org/10.1681/ASN.2012080823>). Overall, these TFs can regulate actin-related genes to maintain cytoskeletal dynamics.

In contrast, there is no global loss of TEAD1 binding upon Wt1 knock-down, TEAD1db is relatively enriched in the Tead1 primary motif and not much different in motif composition, compared to TEAD1wt. This suggests that podocyte-specific enhancer regions under control of TEAD1 may play a compensatory role in transcriptome remodeling upon the induced podocyte damage. This hypothesis is further explored in the next subchapter by analyzing the functional impact of differential binding.

### 3.3.3 WT1 and TEAD1 have Common and Distinct Functions in Regulating Podocyte Structure and Survival in Podocyte Injury

To build on results of motif analysis and further speculate which biological functions are controlled by WT1 and TEAD1 in healthy podocytes and what aspects of podocyte biology may be directly affected by the binding changes of TFs in Wt1<sup>ko/wt</sup> model of genetic FSGS, we first inferred target genes of wild-type and differential binding of both TFs (Materials and Methods 2.3.1.1) and then performed functional annotation of the target gene sets with GO and collection of pathways, using the whole genome as background.

Results of the analysis show that genes predicted as targets of co-binding and individual binding of WT1 and TEAD1 in wild-type glomeruli mostly share functional annotations (Fig.3.3.7, first three columns), including a group of GO terms related to *actin cytoskeleton organization and remodeling*, *regulation of cell adhesion and cell-junction organization*, *morphogenesis* and *response to growth factor stimuli* and *glomerulus development*. When focusing on differences, co-binding targets are characterized by higher enrichment in genes involved in *focal adhesion* - a core podocyte function, targets of solo WT1 binding are more enriched in genes related to *Rho GTPase binding* and *glomerulus development*, targets of solo TEAD1 binding are more enriched in genes involved in *blood vessel morphogenesis* and in *regulation of cell adhesion* genes. Rho GTPases are pivotal regulators of the actin cytoskeleton and influence cell shape, motility, and differentiation. WT1 may independently regulate genes that control cytoskeletal dynamics and podocyte morphogenesis, underscoring its crucial role in the development and maintenance of podocyte architecture. TEAD1 may specifically regulate genes that facilitate interactions between podocytes and endothelial cells or contribute to the vascularization processes within the glomerulus. Additionally, by controlling genes related to cell adhesion, TEAD1 might influence podocyte adherence and communication with neighboring cells.

Functions enriched in predicted targets of differential binding of WT1 and TEAD1 in Wt1<sup>ko/wt</sup> are more heterogeneous (Fig.3.3.7, columns 4-7): some functions are shared with targets of wild-type binding but there are also unique enrichments. One of the most enriched and uniquely significant terms for TEAD1db is *negative regulation of anoikis*. Anoikis is a cell death

mechanism which occurs when the integrin-mediated interaction between a cell and the extracellular matrix (ECM) is disrupted, causing cells to no longer be anchored to the ECM, and apoptosis to be triggered (Frisch and Francis 1994). *Focal adhesion* is enriched in targets of wild-type binding and TEAD1db but not WT1db.

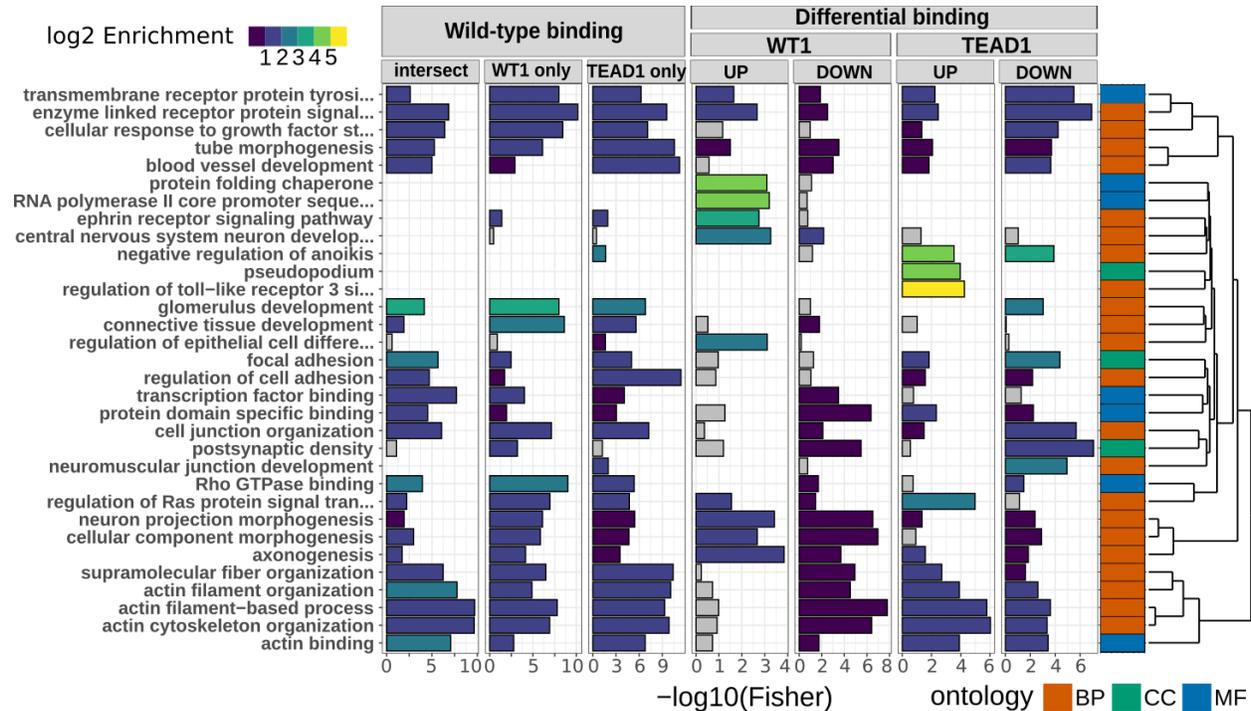


Figure 3.3.7: Functional annotation of predicted target genes of wild-type and  $Wt1^{ko/wt}$  differential binding of WT1 and TEAD1. Y-axis show GO terms, X-axis shows  $-\log_{10}$  p-value of the enrichment, bar color shows  $\log_2$  enrichment score. Union of top 10 most significantly enriched categories in each test was used for the figure.

Increasing complexity of the picture, we found differences in functional annotation of genes affected by up or down-binding of each TF. *Postsynaptic density*, *cell-junction organization* are more enriched in genes down-bound, *regulation of Ras protein transduction* are more enriched in genes up-bound by both TFs. *Regulation of epithelial cell differentiation*, *ephrin receptor signaling pathway*, *protein folding chaperone* are enriched only in targets up-bound by WT1. *Pseudopodium* and *regulation of toll-like receptor 3 signaling pathways* are enriched only in targets up-bound by TEAD1. Complex also seems to be regulation of genes related to *actin cytoskeleton organization*: they are relatively enriched in targets down-bound by WT1 and up-bound by TEAD1 in glomeruli of  $Wt1^{ko/wt}$  mice. The increased binding of TEAD1 to genes involved in focal adhesion, and actin cytoskeleton organization suggests that TEAD1 may partially compensate for the loss of WT1 function to maintain podocyte survival and attachment. Up-binding of genes involved in survival signaling pathways (e.g., Ras signaling, regulation of

anoikis, regulation of TLR) reflect cellular response to stress and indicate attempts to regulate apoptosis. On the other hand, enhanced regulation of TLR3 signaling by TEAD1 may lead to increased inflammatory responses within the glomerulus, potentially exacerbating podocyte injury.

To summarize, co-binding and shared functional annotations of TEAD1 and WT1 target genes highlight their collaborative roles in essential podocyte functions, such as maintaining the cytoskeleton, cell adhesion, and responding to growth factors. This collaboration ensures the preservation of podocyte structure and the integrity of the glomerular filtration barrier. The distinct enrichment patterns in their individually bound targets suggest that WT1 and TEAD1 also have specialized roles. WT1 appears to have a dominant role in regulating genes associated with cytoskeletal dynamics and glomerular development, which are crucial for podocyte differentiation and morphological maintenance. TEAD1 seems to focus on genes involved in vascular interactions and the modulation of cell adhesion, potentially affecting how podocytes interact with the glomerular capillary network and adapt to physiological changes. The differential binding patterns of WT1 and TEAD1 in the *Wt1* model of genetic FSGS highlight a complex interplay between these transcription factors in the context of podocyte injury. WT1 Haploinsufficiency leads to decreased regulation of genes essential for podocyte structure and function which results in impaired cytoskeletal organization, cell adhesion, and differentiation processes. TEAD1 response shows increased binding to genes involved in survival, adhesion, and cytoskeletal maintenance. It may represent an adaptive response to preserve podocyte integrity in the face of WT1 deficiency.

To conclude, change in WT1 concentration causes rewiring of the podocyte TRN by alteration in binding of WT1 and TEAD1 in podocyte-specific enhancer regions, which pushes the podocyte transcriptome away from the wild-type state. Distinct sequence and genomic characteristics of WT1 and TEAD1 differentially bound regions, including the difference between up and down-bound regions, indicate that these regions mediate specific effects of *Wt1* knock-down on the gene expression and hence cellular phenotypes. These specific regulatory functions are realized through binding changes at intergenic and intronic regions of putative podocyte-specific enhancers. By integrating the functional analyses of Chip-seq from both wild-type and FSGS glomeruli, we gained a better understanding of how transcriptional regulation by WT1 and TEAD1 maintains podocyte function under normal conditions and how alterations in their binding contribute to disease pathology.

### 3.3.4 Integrative Analysis of Chip-seq and RNA-seq: Altered Binding of WT1 and TEAD1 in Podocytes Affects Glomerular Transcriptome

To get an integrated view of GRN rewiring upon genetically induced FSGS we combined transcriptomic and Chip-seq evidence in multiple ways presented below. First, we tested overlaps between sets of differentially expressed and TF target genes. We found that genes differentially expressed in  $Wt1^{ko/wt}$  and  $Nphs2^{mut}$  models of FSGS significantly overlap with genes bound by WT1 and TEAD1 in wild type glomeruli and with genes differentially bound by both TFs under  $Wt1$  haploinsufficiency. There is a couple of specific observations: 1) overlap between TF target genes and DEGs is greater at the later stage of the disease, in both FSGS models; 2) genes differentially bound by both TFs have more significant overlap with  $Nphs2^{mut}$  DEG compared to  $Wt1^{ko/wt}$  DEG.

Next, we analyzed expression of TFs and their targets, identified in Chip-seq experiments. We observed that  $Wt1$  and  $Tead1$  glomerular mRNA levels respond differently to the podocyte damage (Fig.3.4.1 A-B, upper row):  $Wt1$  mRNA level is decreased at both stages of FSGS in both models, while  $Tead1$  mRNA level is increased at both FSGS stages of  $Wt1^{ko/wt}$  and increased at the later FSGS stage in  $Nphs2^{mut}$ . We also checked how expression of TF mRNA correlates with average expression of predicted TF targets. Average expression of all WT1 target genes follow expression of  $Wt1$  gene in both genetic models of FSGS (Fig.3.4.1 A, upper and middle row), while all TEAD1 targets don't follow expression of TF's mRNA (Fig.3.4.1 B, upper and middle row). Correlation between expression of the WT1 coding gene and all WT1 targets implies that  $Wt1$  is indeed the master regulator of the identified WT1 target genes. Lack of correlation between expression of TEAD1 coding gene and TEAD1 target genes, implies that either identification of target genes has failed or, more likely, that these genes are co-regulated by other TFs. Expression of genes predicted to be coregulated by the 2 TF correlates with expression of  $Wt1$  but weaker than expression of genes bound only by WT1 ([Suppl.Fig.3.5](#)). This makes sense as  $Tead1$  is also involved in regulation of these genes and, since its mRNA lvl is changing in the direction opposite to  $Wt1$  mRNA, it would compensate, to some extent, for changes in expression of co-bound genes caused by the decrease in WT1 level.

Interestingly, the behavior of targets of differential binding is reversed: WT1db targets change (Fig.3.4.1 A, bottom row) in the direction opposite to changes in  $Wt1$  mRNA level, while TEAD1db targets change (Fig.3.4.1 B, bottom row) in the same direction as  $Tead1$  mRNA level in both FSGS models. To interpret behavior of differential binding targets we should consider changes in TF mRNA levels together with changes in IP signal at the differentially bound sites (Fig.). Given that expression of  $Wt1$  mRNA decreases and most WT1db regions lose WT1 IP signal in podocytes, the increased expression of WT1db targets in glomeruli of both FSGS models may have following explanations: 1) WT1 acts as a repressor at WT1db regions, so removal of  $Wt1$  activates WT1db targets; 2) another TF binds WT1db regions and activates expression of the targets; 3) podocyte signature in bulk glomerular RNA-seq is heavily confounded so the observed increase in WT1db targets expression comes from another cell-type.

Situation is different for TEAD1: expression of *Tead1* mRNA increases, a significant amount of TEAD1db regions gain IP signal and targets of these regions exhibit increased expression in glomeruli of FSGS mice (Fig.3.4.1.B, bottom row). This implies that *Tead1* is recruited to a number of TEAD1db sites and plays an active role in rewiring of GRN under podocyte damage.

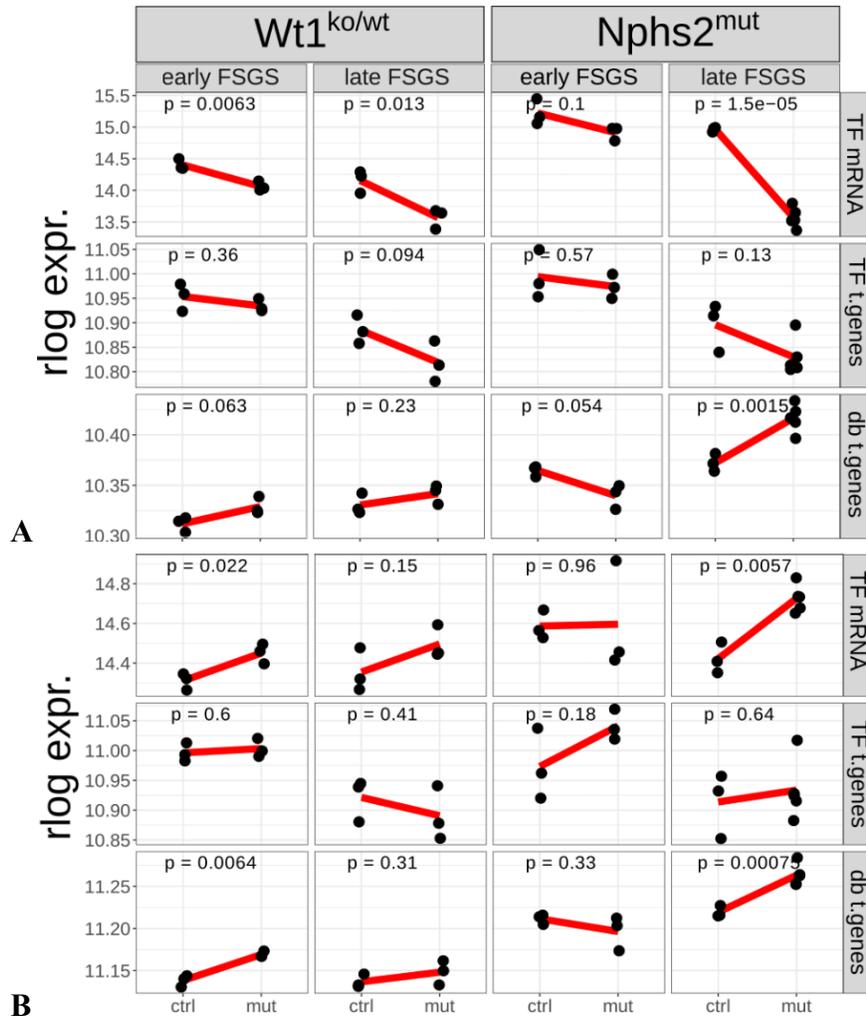


Figure 3.4.1: Correspondence between bulk RNA-seq and Chip-seq analysis of genetic FSGS models. (A) Overlap between genes differentially expressed in bulk RNAseq (columns) and target genes of TFs inferred from Chip-seq (rows). (B) Expression of *Wt1* and its target genes in glomerular bulk RNA-seq. Y-axis shows rlog-normalised expression of transcription factor mRNA, mean expression of their target genes (TF t.genes) or mean expression of targets of differential binding (db t.genes), as was defined based on Chip-seq. X-axis shows the genotype of samples. Complete results see in [Supplementary Figure 3.5](#).

One putative TF-targets-function link that we found in the data was correlation between *Tead1* mRNA, TEAD1 differentially bound genes (Figure 3.4.1 B) and cell-cycle genes (Figure 3.2.5 B). There is upregulation of all three components in glomerular transcriptome of both

FSGS models, but at different disease stages: early in *Wt1*<sup>ko/wt</sup> and late in *Nphs2*<sup>mut</sup>. It has been previously described that injury may cause podocytes to re-enter the cell cycle, which was at least partly mediated by YAP/TEAD1 signaling. One of the consequences of re-entry into the cell cycle was over-expression of mesenchymal markers in podocytes (Xie et al. 2019). Our results firmly support this observation, albeit leaving a puzzle of different timing of cell cycle upregulation in different genetic models.

Finally, to systematically understand functional relation between FSGS-induced changes in glomerular transcriptome and changes in *Wt1*/*Tead1* podocyte TRN, we compared functional annotations of genes differentially expressed in FSGS glomeruli with functional annotations of genes predicted as differentially bound by *WT1*/*TEAD1* in podocytes of FSGS glomeruli (Figure 3.4.2).

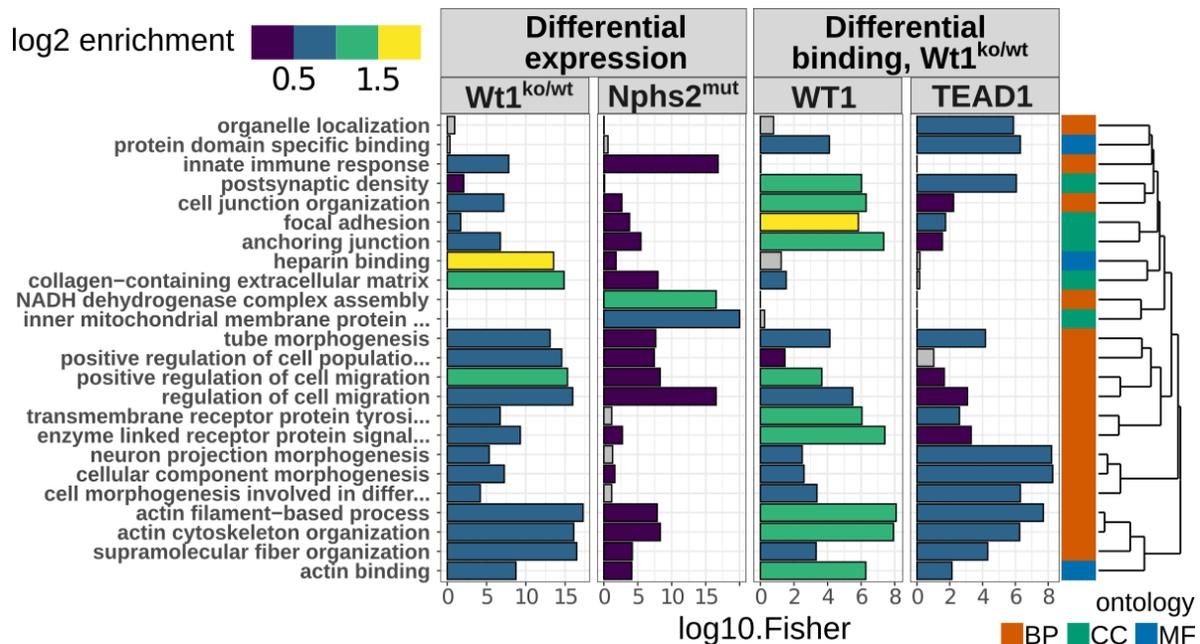


Figure 3.4.2: Integration of RNA-seq and ChIP-seq functional annotations highlights key GO terms contributing to FSGS. X-axis shows  $-\log_{10}$  p-value of the enrichment test, color of the bars shows  $\log_2$  enrichment score. Union of top 10 GO terms most significantly enriched in each test was used.

Results presented on the Figure can be summarized in following points: 1) many podocyte-related GO annotations are significant and highly ranked in all 4 tests, including processes, like *actin cytoskeleton organization and regulation of cell migration*. 2) Certain podocyte-specific processes, such as *focal adhesion* and *post-synaptic density* (protein complex that regulates cellular adhesion and controls receptor clustering), show much higher rankings in the differentially bound compared to the differentially expressed gene sets. 3) There are several functions, like *innate immune response*, *heparin binding* and *mitochondrial function*, that are significantly enriched only in genes differentially expressed and not in differentially bound genes. The fact that sets of DEG and DBG have overlapping but non-identical functions can be

due to (1) a mixture of primary and secondary effects of podocyte damage on transcription and/or (2) due to a mixture of signals from various cell-types in glomerular transcriptome. This confirms the earlier observation (Chapter 3.2) about limitations of bulk glomerular transcriptomics in capturing podocyte-specific transcriptional effects of FSGS .

To sum up, altered binding of WT1 and TEAD1 in  $Wt1^{ko/wt}$  model of FSGS directly affects glomerular transcriptome, modulating key biological processes, leading to podocyte injury, detachment, and eventual glomerulosclerosis. Furthermore, integrating bulk RNA-seq with ChIP-seq improved our understanding of podocyte-specific effects of  $Wt1$  mutation on transcription. The results further highlight the critical roles of WT1 and TEAD1 in regulating gene expression related to podocyte structure and function upon the podocyte damage.

## 3.4 Discussion

Our study provides an in-depth investigation into the WT1 transcriptional regulatory network (TRN) in podocytes, especially under pathological conditions like focal segmental glomerulosclerosis (FSGS). WT1 is a central transcription factor that orchestrates podocyte-specific gene expression, playing a vital role in the development, differentiation, and maintenance of podocyte structure and function. The loss or dysfunction of WT1 leads to podocyte injury, glomerular damage, and the progression of kidney diseases like FSGS.

### **Key findings and their implications**

**WT1 as a Master Regulator of Podocyte Function:** WT1 emerged as a core regulator of podocyte differentiation and maintenance. In healthy podocytes, WT1 controls essential genes related to the slit diaphragm, focal adhesion, and extracellular matrix (ECM). These components are crucial for maintaining the glomerular filtration barrier and ensuring proper kidney function. WT1 haploinsufficiency in the *Wt1ko/wt* FSGS model leads to disrupted transcriptional control over these podocyte-enriched proteins, resulting in impaired cytoskeletal organization, cell adhesion and podocyte differentiation. This disruption causes podocyte detachment and foot process effacement, culminating in proteinuria and glomerulosclerosis.

**Dynamic Response of WT1 and TEAD1 to Podocyte Damage:** Our ChIP-seq analyses revealed that both WT1 and TEAD1, another important transcription factor, show significant changes in DNA binding during podocyte injury. WT1 binding is globally diminished in the *Wt1ko/wt* FSGS model, leading to the loss of regulation over essential podocyte genes. Interestingly, TEAD1, which is part of the Hippo signaling pathway, demonstrated increased binding at regions regulating cell adhesion, cytoskeletal maintenance, and survival genes. To hypothesize, TEAD1 may play a compensatory role, attempting to maintain podocyte integrity in the face of WT1 dysfunction. Alternatively, binding of TEAD1 to the genomic sites vacated by *Wt1* could be a non-adaptive, side-effect of losing *Wt1*.

**Progressive Alterations in ECM and Immune Response:** In both the *Wt1ko/wt* and *Nphs2mut* FSGS models, we observed significant transcriptional changes in genes related to the extracellular matrix (ECM) and collagen biology. This may reflect an attempt to repair the glomerular basement membrane (GBM) following injury, or maladaptive remodeling of GBM or combination of both. As FSGS progresses, the innate immune response is also upregulated, with increased complement activation and immune cell infiltration in damaged glomeruli. These immune-mediated pathways exacerbate podocyte damage, promoting fibrosis and glomerulosclerosis.

**Limitations of Bulk RNA-Seq in Capturing Podocyte-Specific Changes:** Through RNA-seq analyses, we identified global transcriptional changes in FSGS glomeruli. However, these findings underscored the limitation of bulk glomerular transcriptomics, where the signal reflects contributions from various glomerular cell types. Despite the insights gained into

podocyte-specific injury, bulk RNA-seq data obscures more subtle changes in the podocyte-specific transcriptome. Integrating ChIP-seq data with RNA-seq partially addressed this issue by showing how binding changes in WT1 and TEAD1 directly affect the glomerular transcriptome, particularly genes involved in podocyte biology.

### **Findings in the context in FSGS etiology and treatment**

**Podocyte-Centric Pathogenesis of FSGS:** The findings from the *Wt1ko/wt* model highlight the central role of transcriptional dysregulation in podocyte biology as a driving force in FSGS. WT1 deficiency impairs the expression of key podocyte genes, triggering a cascade of cellular dysfunction that leads to progressive glomerular injury. Additionally, the compensatory response of TEAD1 suggests that podocytes attempt to adapt to WT1 loss, although these efforts are ultimately insufficient to prevent disease progression. TEAD1 may, therefore, represent a novel target for therapeutic interventions aimed at reinforcing podocyte resilience under stress.

**Potential Therapeutic Approaches for WT1-Related FSGS:** Targeting the WT1 transcriptional network offers a promising approach for managing FSGS. Gene therapies or small molecules that restore WT1 function or enhance its downstream effects could help maintain podocyte integrity. The intersection of ECM remodeling and immune activation suggests that therapeutic strategies must address both structural and inflammatory aspects of the disease. Perhaps anti-inflammatory drugs that modulate immune responses could help alleviate ongoing glomerular damage and delay the progression to chronic kidney disease.

### **Limitations of the Research**

**Model-Specific Insights:** Although the *Wt1ko/wt* and *Nphs2mut* models provide valuable insights into the molecular mechanisms of FSGS, they represent genetic models of the disease, which may not capture the full spectrum of sporadic or secondary FSGS in humans. The findings from these models should be validated in additional genetic and environmental contexts to better understand the heterogeneity of FSGS.

**Incomplete Understanding of Compensatory Mechanisms:** While TEAD1 showed increased binding at critical gene loci in WT1-deficient podocytes, the role of TEAD1 remains incompletely understood. Further studies are needed to clarify whether TEAD1 activation can genuinely prevent disease progression or merely delays podocyte injury. The role of other transcription factors and signaling pathways in this compensatory network should also be explored.

**Bulk RNA-Seq Limitations:** The use of bulk RNA-seq in glomeruli limits our ability to fully dissect podocyte-specific transcriptional changes from those occurring in other glomerular cells. While integrating RNA-seq with ChIP-seq data improved our understanding of podocyte-specific effects, single-cell transcriptomics would offer a more refined view of transcriptional changes in individual cell types within the glomerulus.

## **Conclusion**

Our research provides insights into the WT1 transcriptional network in podocytes and how its dysregulation contributes to the pathogenesis of FSGS. WT1 is essential for maintaining podocyte differentiation and function, and its loss leads to widespread transcriptional and structural defects in podocytes, culminating in glomerular sclerosis. The compensatory role of TEAD1, ECM remodeling, and immune activation further highlight the complex network of processes that drive FSGS progression. These findings suggest avenues for novel therapeutic strategies targeting WT1 function, TEAD1-mediated compensatory mechanisms, and the ECM-immune feedback loop to protect podocytes and slow the progression of FSGS. However, the limitations of genetic models and bulk transcriptomic approaches underscore the need for further research to refine our understanding of FSGS etiology.

# **Chapter 4. Single-cell Analysis of Podocyte Transcriptional Response to Damage**

## 4.1 Introduction

In the preceding chapter, we employed bulk RNA sequencing (RNA-seq) to analyze the glomerular transcriptomes of genetically engineered mouse models of focal segmental glomerulosclerosis (FSGS): the *Wt1* heterozygous knockout ( $Wt1^{ko/wt}$ ) and the *Nphs2* mutant ( $Nphs2^{mut}$ ) mice. While bulk RNA-seq provided valuable insights into global gene expression changes associated with podocyte injury and disease progression, it lacked the resolution to discern cell-type-specific transcriptional dynamics due to the heterogeneity of glomerular cell populations. Understanding the molecular mechanisms underlying podocyte dysfunction necessitates a focused examination of podocyte-specific transcriptomic alterations. Therefore, in this chapter, we transition to single-nucleus RNA sequencing (snRNA-seq) to dissect the transcriptional reprogramming occurring specifically within podocytes during FSGS progression. Specifically, we performed snRNA-seq analysis of previously described *Nphs2* mutant and *Wt1* heterozygous knockout mouse models of genetic FSGS (Figure 4.1, red). All experimental wet lab work necessary to generate data for this chapter was performed by He Chen.

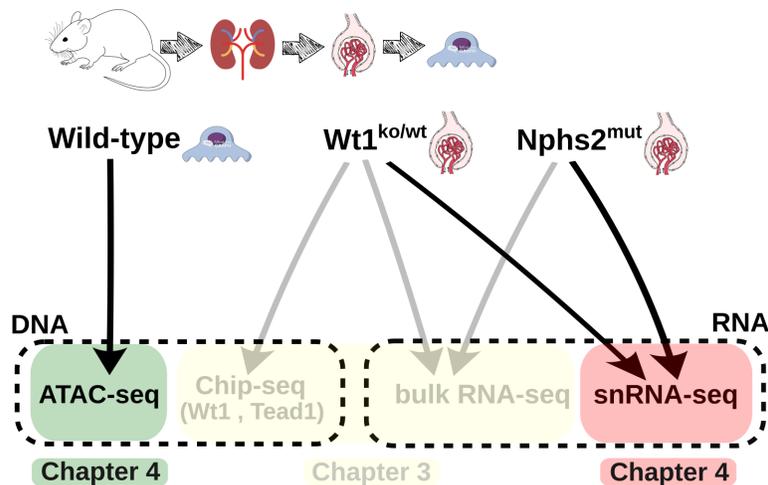


Figure 4.1: Experimental data used in chapter 4. Samples were taken from glomeruli of control,  $Wt1^{ko/wt}$  and  $Nphs2^{mut}$  experimental animals.

Single-cell transcriptomics is a rapidly evolving technology that allows for the quantification of gene expression at the individual cell level, capturing cellular heterogeneity and revealing rare cell populations. This technology has been instrumental in renal research, enabling the characterization of complex cell types within the kidney and their responses to injury. Specifically, single cell RNA-seq has already provided insights into the transcriptional alterations of podocytes and their interactions with other glomerular cells in FSGS (J. Park et al. 2018). By comparing healthy and diseased states at single-cell resolution, researchers could identify specific pathways and regulatory networks implicated in podocyte dysfunction. Furthermore, single-cell transcriptomics can aid in the development of personalized medicine approaches by

identifying population and patient-specific molecular signatures. However, existing analytical approaches for scRNA-seq and spatial transcriptomics data have limitations, particularly in mapping the gradual transition of cellular states from healthy to diseased phenotypes. Analyses that rely on averaged gene expression levels between populations of cells from diseased and healthy tissues may overlook subtle variations within individual samples. Over time, virtually all organs develop subpopulations of damaged cells that may serve as early indicators of emerging diseases. Yet, few tools are currently available to identify such rare cells in scRNA-seq data from clinically healthy subjects.

In the second part of this chapter we address these limitations and, capitalizing on publically available data, we derived a universal transcriptomic signature of podocyte damage that can be applied to (spatial)single-cell/nuclei RNA-seq to quantify damage at the level of individual podocytes. We then applied this podocyte damage score (PDS) to our snRNA-seq data and public scRNA-seq datasets to delineate both model-specific and universal features of transcriptional dysregulation in damaged podocytes. Furthermore, using bulk ATAC-seq data from FACS-sorted podocytes, we constructed a podocyte transcriptional regulatory network (Figure 4.1, green).. This network was instrumental in interpreting the results of the PDS analysis and in identifying novel transcription factors involved in the rewiring of the podocyte transcriptome during FSGS progression. Below is a short summary of each of the 2 parts of Chapter 4.

### **snRNA-seq allows in-depth characterization of transcriptional regulation in podocytes.**

Single-nuclei (sn) RNAseq is a variation of single-cell (sc) RNA-seq method which allows to sequence RNA molecules contained in nuclei (Grindberg et al. 2013). Studies comparing sc and sn RNA-seq report that the two technologies, although profile different RNA fractions, correlate well with each other and detect sufficient genes for adequate representation of cell populations. However, snRNA-seq reduces cellular and stress response biases and provides enhanced detection of nuclear transcripts, compared to scRNA-seq, making it a better suited technology to study transcriptional regulation(Wu et al. 2019; Bakken et al. 2018). These considerations motivated our choice in favor of single-nuclei sequencing.

We conducted snRNA-seq on isolated glomeruli from wild-type, *Wt1ko/wt*, and *Nphs2mut* mice to achieve a high-resolution, cell-type-specific transcriptome analysis. A critical observation was the alteration in cell-type composition associated with FSGS progression. Specifically, there was a significant reduction in the proportion of podocytes in both mutant models compared to wild-type controls, which closely paralleled the decline in kidney function. In particular, podocyte loss was more rapid and plateaued earlier in *Nphs2* mutant mice, reaching a critical threshold by week 8, whereas the *Wt1* knockout mice demonstrated a slower, more gradual podocyte depletion, stabilizing by week 12. This loss of podocytes was strongly associated with the development of proteinuria, highlighting the direct relationship between podocyte integrity and glomerular dysfunction.

At the transcriptional level, both models displayed a substantial number of differentially expressed genes (DEGs) involved in key biological processes such as cytoskeletal organization,

cell adhesion, and stress response pathways. Strikingly, despite their distinct genetic mutations, the transcriptional profiles of podocytes converged at later stages of disease progression in both models. This convergence underscores the potential for diverse genetic insults to elicit common maladaptive or adaptive responses in podocytes as the disease advances. These findings not only elucidate the molecular mechanisms driving podocyte injury but also identify transcriptional changes that may serve as therapeutic targets to preserve podocyte function in glomerular diseases.

Finally, by integrating our snRNA-seq data with the bulk RNA-seq results, described in Chapter 3, we assessed the limitations of bulk transcriptomics in capturing cell-type-specific changes. Principal component analysis (PCA) demonstrated that sequencing technology and sample type (bulk vs. single-nucleus; glomerular vs. podocyte) were major sources of variation, with cell-type-specific transcriptomic changes potentially masked in bulk analyses due to the heterogeneous cell populations. This underscores the importance of single-cell approaches in accurately delineating the molecular events within specific cell types, such as podocytes, during disease progression.

### **Single-cell damage score in podocytes**

One of the important results of bulk RNAseq analysis that we confirmed in single-nuclei RNAseq was a high degree of similarity between Wt1 and Nphs2 models of FSGS on the transcription regulation level. Advanced transcriptome analyses, particularly single-cell RNA sequencing, have revealed that injured podocytes, regardless of the underlying cause, activate common genetic programs leading to alterations in cell structure, function, and survival. For instance, Kuppe et al. (2021) identified a shared profibrotic transcriptional signature in podocytes from patients with various kidney diseases, suggesting a universal maladaptive response mechanism (Kuppe et al. 2021). These transcriptional changes often involve the upregulation of stress response genes, cytoskeletal remodeling proteins, and extracellular matrix components, contributing to podocyte dysfunction and loss (J. Park et al. 2018).

With this in mind, we hypothesized that different causes of FSGS converge on common mechanisms of transcriptional regulation. We propose that a compact transcriptome signature can be used to calculate damage in individual podocytes, regardless of the underlying disease etiology. We also assumed that this gene signature can be derived in a supervised manner, by meta-analysis of transcriptomic data of various models of kidney damage. To test our hypothesis, we collected studies of various podocyte damage models, performed their differential expression analysis and combined results from multiple models using ranks into a single gene signature. This universal transcriptomic signature of FSGS was then applied to single cell RNA-seq data to calculate damage score of individual podocytes, so called podocyte damage score. Rounds of validation showed that this score is applicable for a wide range of podocyte damage models and that it correlates with measures of podocyte morphology, glomerular and kidney health in mice and in Human data..

## Characterizing Cellular Changes in Disease Progression with PDS

To demonstrate the potential of PDS in characterizing cellular changes in disease progression, we applied PDS to eight diverse models of podocyte damage, each profiled with single cell RNA sequencing. Among the models are *Nphs2* and *Wt1* datasets analyzed in the previous chapter; conditional podocyte *Pdss2* knockout snRNA-seq data generated by the Nephrolab; and four models of podocyte damage from a published study (GEO accession GSE146912) (Chung et al. 2020). The dataset includes models of diabetic nephropathy (*Btbr*), conditional podocyte *Cd2ap* knockout (*Cd2ap*), the Adriamycin toxicity model (*Doxo*), and nephrotoxic serum nephritis (*Nephritis*), each representing distinct mechanisms of podocyte injury ([Supplementary Table 1](#)). Here is a short summary of the models: **Wt1ko/wt** model results in podocyte injury, proteinuria, and glomerulosclerosis, mimicking human syndromes like Denys-Drash and Frasier, which lead to nephrotic syndrome and kidney failure. Mutations in *Nphs2* disrupt the slit diaphragm, causing podocyte injury, proteinuria, and rapid progression to FSGS, similar to autosomal recessive FSGS in humans. Loss of *Pdss2* in podocytes causes mitochondrial dysfunction, leading to FSGS-like phenotypes with progressive podocyte injury and glomerular damage. The *Btbr* mouse develops insulin resistance and hyperglycemia, leading to progressive kidney damage and glomerulosclerosis, similar to human diabetic nephropathy. *Cd2ap* knockout disrupts the slit diaphragm, causing podocyte detachment, proteinuria, and severe glomerulosclerosis resembling FSGS. **Adriamycin Toxicity Model** induces podocyte damage through Adriamycin toxicity, leading to nephrotic syndrome, podocyte loss, and glomerulosclerosis. In **Nephritis model the damage is** Induced by nephrotoxic serum, this model causes immune-mediated podocyte injury, proteinuria, and glomerular damage, mimicking immune-related glomerulonephritis.

By sorting cells according to their extent of damage, we can more effectively identify molecular processes involved in progressive cellular injury. Employing these generalized damage markers enables the study of disease-related molecular mechanisms in affected cells across various animal models, as well as in unperturbed control conditions. Results of our analysis presented in this chapter shows that this approach allows for the sequential mapping of molecular events, facilitating the distinction between early and late-stage cellular responses during pathogenesis, which is crucial for the rational development of targeted therapies.

## 4.2 Single Nuclei RNA-seq Allows in-depth Characterization of Transcriptional Response to Podocyte Damage

To study effects of genetically-caused podocyte damage and ensued progression of focal segmental glomerulosclerosis on podocyte transcriptome, we performed single-nuclei RNA sequencing of isolated mouse glomeruli from wild-type,  $Wt1^{ko/wt}$  mice and  $Nphs2^{mut}$  mice.  $Wt1^{ko/wt}$  snRNA-seq experiment includes a total of 10 samples from 4, 12-13 and 25 week old mice.  $Nphs2^{het.mut.}$  snRNA-seq experiment includes a total of 14 samples from 4, 6, 8 and 12 week old mice. Complete information for single-nuclei RNA-seq samples is provided in [Supplementary table 3](#).

After mapping reads to mouse genome and quantification of gene expression, we removed ambient RNA from raw counts (Material and Methods 2.3.1.4), the corrected count matrices were used to produce QC plots ([Supplementary Figure 4.1](#)). QC plots confirmed good quality of individual samples: low levels of mitochondrial and ribosomal RNAs and an average of 1800 features being expressed per cell. After filtering out a small number of low quality cells and doublets (Material and Methods 2.3.1.4) we ended up with a total of 72133 and 96789 cells in  $Wt1^{ko/wt}$  and  $Nphs2^{mut}$  datasets, respectively. The processed data underwent a dimensionality reduction, the results of which are visualized in Figure 4.2.1. These UMAP plots present a two-dimensional view of the  $Wt1^{ko/wt}$  (panels A, C) and  $Nphs2^{mut}$  (panels C, D) snRNA-seq datasets, wherein each dot represents an individual cell. Panels A and B depict cells labeled by the combination of age and genotype (group-level labels). These panels show that, in both datasets, variation between batches does not dominate cell clustering and that data integration is not necessary.

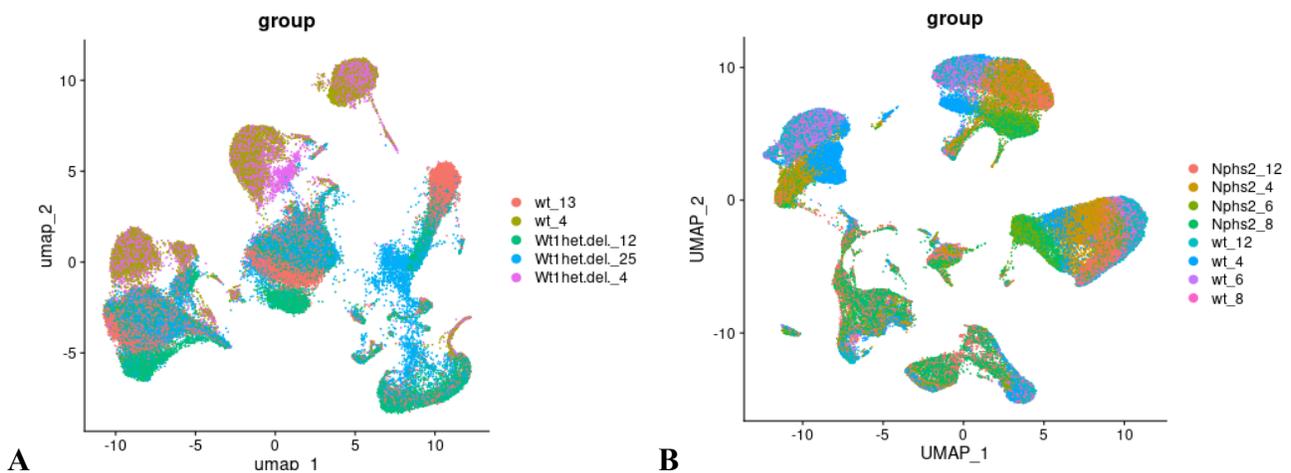


Figure 4.2.1: UMAPs of the pre-processed and filtered glomeruli snRNA-seq data from  $Wt1^{ko/wt}$  (A) and  $Nphs2^{mut}$  (B) experiments. Cells are coloured by combination of genotype and age information.

#### 4.2.1 Cell-type Composition of Glomeruli is Changed by FSGS Progression.

Next step of the analysis was annotation of cell clusters with known glomerular cell-types. The annotation was performed manually, by visual inspection of expression of published kidney cell-type markers (Kirita et al. 2020). Results of cell-type annotation (Fig.4.2.2 A,B) show a similar cell-type composition of both datasets with three biggest clusters of cells representing endothelial, mesangial and podocyte cells of glomeruli. Another identified glomerular cell-type is parietal epithelial cells (PEC). Non-glomerular kidney cells are presented by cells of juxtaglomerular apparatus (JGA) and by a substantial population of cells from collecting ducts: proximal tubule (PT), thick ascending limb (TAL), Intercalated (IC) and principal (PC) cells. Notably, there is also a cluster of immune cells and a small cluster of proliferating cells in both datasets. We summarized cell-type annotations by calculating the proportion of each cell type in individual samples in [Supplementary Figure 4.2](#). Results show that three most abundant glomerular cell-types comprise up to 92% all cells in glomerular snRNA-seq samples. Another observation is a systematic difference in detection rate of specific cell-types between *Wt1* and *Nphs2* experiments. For instance, the average podocyte fraction in sequenced controls of *Wt1<sup>ko/wt</sup>* and *Nphs2<sup>mut</sup>* experiments is 0.39 and 0.2, correspondingly, - almost 2 fold difference. Given the same genotype, similar sequencing depth and quality of samples in both experiments, this bias likely represents a technical batch effect or a strain difference. We should also keep in mind systematic differences in cell-type detection rates between scRNA-seq and snRNA-seq (Denisenko et al. 2020). These technical and methodological biases make it difficult to give a biological interpretation of cell-type composition for specific samples. However, we expect that the change in cell-type composition between samples, caused by genotype or disease progression, should not be affected by a technical factor to a great extent.

To describe changes in cellular composition caused by genetic FSGS we performed differential cell-type abundance analysis (Material and Methods 2.3.2). The analysis failed to identify the significant effect of disease progression on detection rates of glomerular cell-types, likely due to insufficient statistical power. However, results show an insignificant in *Wt1<sup>ko/wt</sup>* and a highly significant in *Nphs2<sup>mut</sup>* samples decrease in podocyte detection rate, compared to wild-type samples (Fig.4.2.2 C). There is also a significant increase in detection rates of immune and proliferating cells in *Nphs2<sup>mut</sup>*, compared to wild-type samples (Fig.4.2.2 C), which is likely related to the highly inflammatory properties of the podocin mutation.

We further investigated changes in podocyte detection rate over FSGS progression by visualizing podocyte fraction in glomerular snRNA-seq samples. Figure 4.2.2 D (left side) shows virtually no change of the podocyte fraction in *Wt1<sup>ko/wt</sup>* and 2 fold decrease in *Nphs2<sup>mut</sup>* mice at the age of 4 weeks (early FSGS). Both models exhibit a further decrease in podocyte fraction later in the disease, though the loss is much faster in *Nphs2<sup>mut</sup>* mice, plateauing at the 8th week of the disease, at which point podocytes are only 2% of all sequenced cells. The loss is much slower in *Wt1<sup>wt/ko</sup>* model, where the podocyte fraction decreases 2 times by 12th week and then stays almost unchanged till 25th week of mutants age.

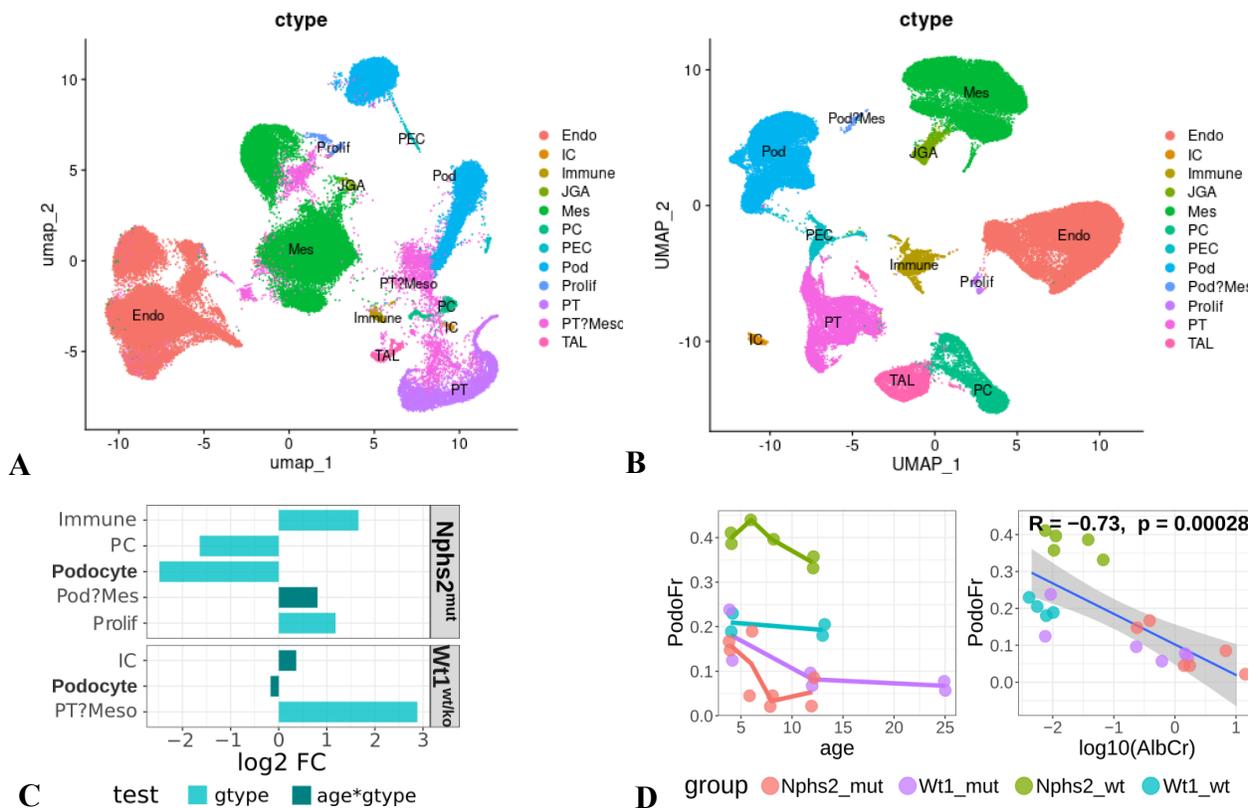


Figure 4.2.2: Cell-type analysis of glomeruli snRNA-seq samples.

UMAPs of *Wt1<sup>ko/wt</sup>* (**A**) and *Nphs2<sup>mut</sup>* (**B**) experiments where cells are coloured by kidney cell-types. (**C**) Results of differential cell-type abundance analysis show log<sub>2</sub> fold change (X-axis) in a cell-type (Y-axis) abundance caused by genotype or disease progression (age:gtype), only significant results (q-value<0.1) are shown. (**D**) Proportion of podocyte cells (Y-axis) in snRNA-seq samples depending on age or proteinuria (X-axes). Blue line depicts results of linear regression, spearman correlation was calculated. Cell-type labels assigned per cluster: Endo - endothelial, IC - intercalated, Immune, JGA - juxtaglomerular apparatus, Mes - mesangium, PC - principal, PEC - parietal epithelial, Pod - podocyte, Prolif - proliferating, PT - proximal tubule, TAL - thick ascending limb cells.

Lastly, to test how much of FSGS phenotype can be explained by the observed total podocyte fraction we have correlated it with Albumin/Creatinine measure of kidney function (Figure 4.2.2 D, right side), available for some of the sequenced mice (see last column in [Supplementary table 3](#), last column). The observed spearman correlation of -0.73 is highly significant, suggesting that a total fraction of podocyte cells in a glomerular snRNA-seq sample may serve as a reliable proxy of podocyte damage in FSGS study.

Overall, results of cell-type analysis show changes in cell-type composition of glomerular snRNA-seq samples in mice with genetic FSGS. The observed decrease in podocyte fraction correlates with kidney function and likely reflects transformation and loss of damaged podocytes.

## 4.2.2 Transcriptional Changes in Podocytes of Wt1 and Nphs2 FSGS models.

Podocytes identified in both analyzed datasets were subject to differential expression (DE) analysis to test effects of genotype and disease progression, while controlling for effects of batch and sex, when applicable. DE analysis revealed 636 and 2139 genes that are significantly (FDR adjusted  $p$ -value  $<0.05$ ) affected by genotype in podocytes of Wt1 and Nphs2 snRNA-seq experiments, correspondingly. The higher number of genes whose expression is affected in podocytes by Nphs2 mutation, is in line with the functional phenotype.

DE analysis of FSGS stages was done by testing genotype effects in samples from control and mutant mice of the same age, ages matching analysis of bulk glomerular transcriptome (Chapter 3.2). There are 469 (4 weeks) and 1377 (12 weeks) differentially expressed genes (DEG) in Wt1ko/wt podocytes; 912 (4 weeks) and 1589 (12 weeks) DEG in Nphs2mut podocytes. Higher number of DE genes at the later stage of FSGS reflects the effect of disease progression in both models.

Analysis of the effect size of the mutation on the podocyte transcriptome shows that it is indeed smaller at the earlier stage of FSGS, in both models, but the difference is bigger in Wt1ko/wt (Figure 4.2.3 B). It seems that at the age of 4 weeks the podocyte transcriptome is significantly transformed in Nphs2<sup>mut</sup> and only starts changing in Wt1<sup>ko/wt</sup> mice. However, at the later disease stage the degree of transcriptome remodeling in both models becomes comparable (Figure 4.2.3 A, dashed lines). This is confirmed by correlation analysis of gene expression changes ( $\log_2FC$ ) in podocytes, caused by mutation. As can be seen on Figure 4.2.3 B, Wt1 and Nphs2 mutation effects significantly correlate at different disease stages, showing convergence of the effects at the later disease stage (Fig.4.2.3 B). Incidentally, this replicates results of bulk glomerular transcriptome analysis.

Analysis of interaction between genotype and age found 260 and 426 genes that are significantly associated with the disease progression in podocytes of Wt1 and Nphs2 experiments. The results are in line with the analysis of podocyte fraction and functional Alb/Cre measures of kidney function (see Figure 4.2.2 D), confirming faster progression of disease in Nphs2 mutants on transcriptome level. Analysis of the disease progression focuses on changes between control and mutant samples that specifically depend on age of mice. This, combined with the analysis of general genotype effect, which does not depend on time, allows to separate, to some extent, primary and secondary effects of the mutation.

Analysis of Wt1 and Nphs2 mRNA levels in podocytes shows that Wt1 expression is strongly decreased in older mice in both mutants, compared to controls, while it is less decreased (Nphs2mut) or even increased (Wtko/wt) in 4 week old mice (Fig.4.2.3 C). Interestingly, Nphs2 gene expression is not significantly decreased in mutants, compared to controls, at any stage of FSGS. Moreover, Nphs2 expression is strongly up-regulated in Nphs2mut in 4 week old mice. These results are at odds with bulk glomerular transcriptome analysis, where we observed much stronger differences between controls and mutants (Fig.3.2.4 C), or even different direction of change of the knocked-out genes. This discrepancy is likely caused by the decrease in the

podocyte fraction (Fig.4.2.2 D) in bulk glomerular samples from sick mice, which would confound expression changes in podocyte cells.

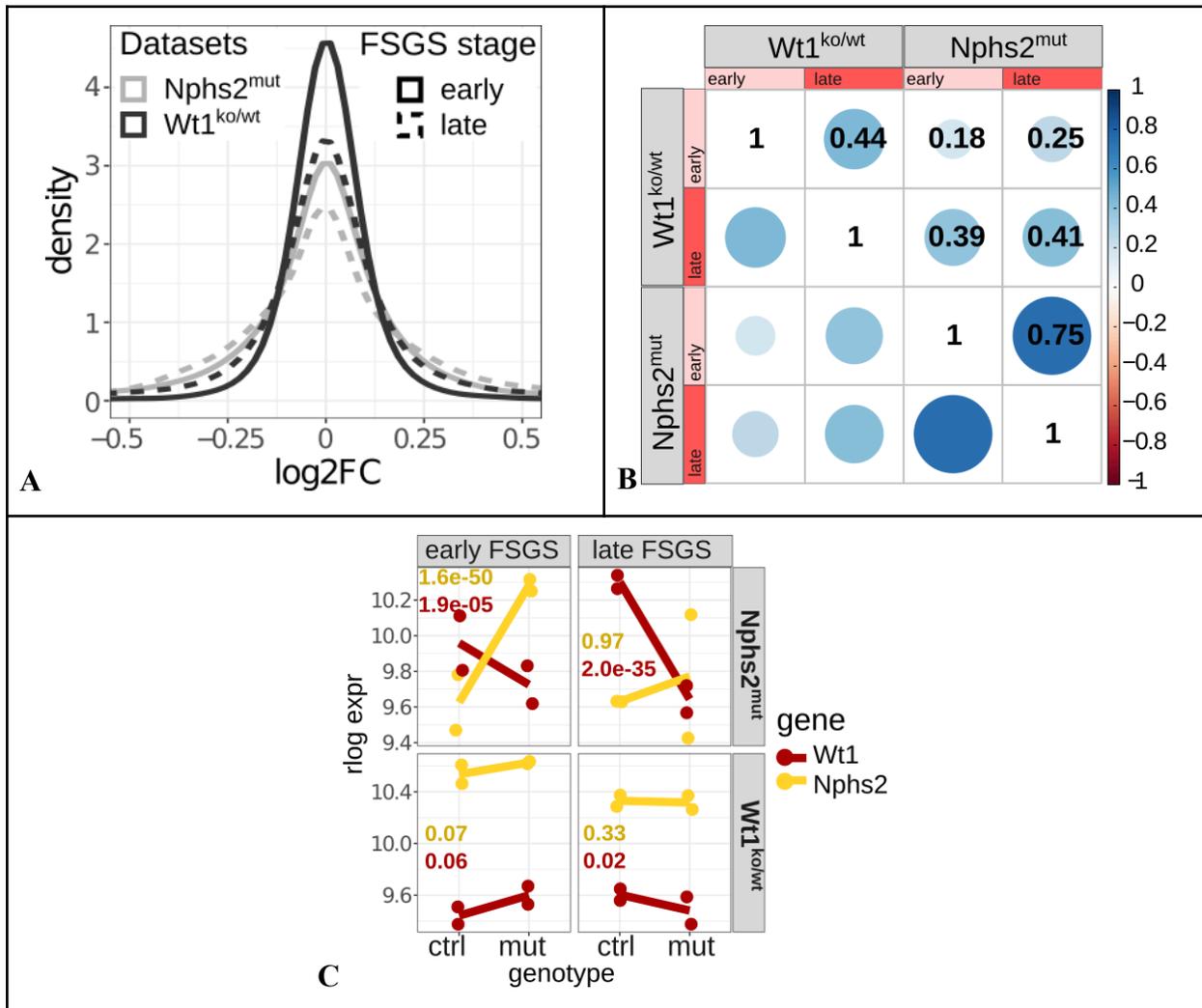


Figure 4.2.3: Results of differential expression analysis of snRNA-seq podocyte samples. (A) Distribution of log<sub>2</sub> fold changes of gene expression. LFCs are calculated between control and mutant podocyte samples in Wt1 and Nphs2 models of FSGS from 4 (early FSGS) and 8/12 (late FSGS) week old mice. Model is denoted by the line color, the stage of FSGS is denoted by the line type. (B) Correlation heatmap of differential expression results, calculated using spearman correlation of shrunk log<sub>2</sub>FC, only the top 30% of variable genes was used. Early and late FSGS is denoted by light and dark red, correspondingly. (C) Expression of Wt1 and Nphs2 genes, denoted by red and yellow, in podocytes, aggregated per sample. Panel rows and columns separate samples by model and disease stage. Adjusted p-values show results of DEseq2 comparison of control and mutant samples.

The observed up-regulation of knocked-out Wt1 and Nphs2 in the podocyte transcriptome earlier in disease progression can be explained in the paradigm of transcriptional network. My explanation invokes the notion of an adaptive response that goes rogue in a maladaptive

environment. Let's assume that deficiency of functional *Wt1* or *Nphs2*, key podocyte genes, elicits a damage response program that normally, in healthy podocytes, increases production of functional *Nphs2* and *Wt1*. Attempts to execute this transcriptional program in a disease context, when restoration of functional *Nphs2* or *Wt1* levels is not attainable, results in disbalance of expressed proteins, which results in further dysregulation of the podocyte TRN, thus feeding the vicious cycle of disease. This explanation has other published experimental evidences that support it (Ettou et al. 2020b). Since decrease in *Wt1* expression is quite a universal feature of podocyte damage of any etiology, an increase in expression of functional *Wt1*, or its targets, in the damaged cells may help to restore TRN state and, arguably, have a protective effect against podocyte damage.

### Functional annotation of DE genes

To characterize effects of *Wt1* and *Nphs2* mutations on podocyte cellular functions, we annotated respective sets of differentially expressed genes with biological pathways and GO terms. Results show that general mutation and disease progression effects on the podocyte transcriptome share multiple annotations ([Supplementary Figure 4.3](#)) but also have many functions specific to each genotype and disease course (Fig.4.2.4). In addition, the average expression of multiple common GO terms and biological pathways may change in different directions ([Supplementary Figure 4.3](#), Fig.4.2.4). Below we attempt to untangle these nuances.

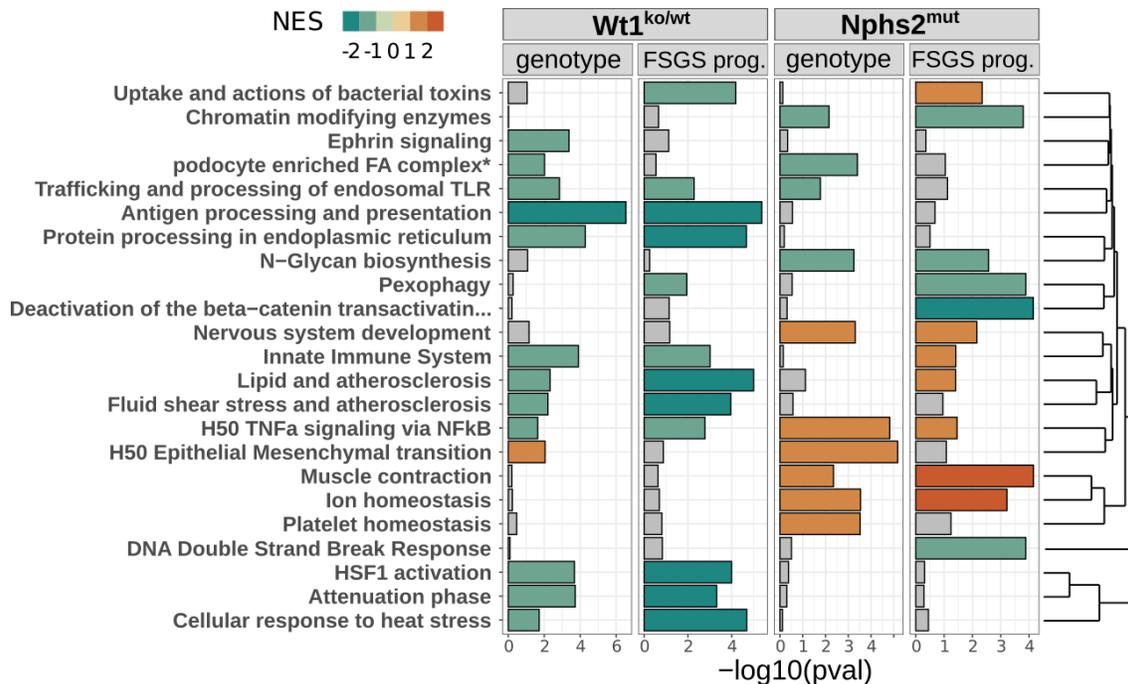


Figure 4.2.4: Functional annotations of genes associated with general effect of genotype and disease progression (FSGS prog) in podocytes of *Nphs2*<sup>mut</sup> and *Wt1*<sup>ko/wt</sup> snRNA-seq experiments. Barplot shows results of pathway GSEA. Color shows normalized enrichment score. X-axis on both plots show  $-\log_{10}$  p-value of fisher test for enrichment.

A closer look at GO term enrichment ([Supplementary Figure 4.3](#)) reveals a significant overlap in functions related to cell junctions, actin cytoskeleton, cell morphogenesis, regulation of cell migration, and angiogenesis between both models. These biological processes are key in maintaining podocyte structure and function, as they are heavily involved in the regulation of cell shape, motility, and interactions with the glomerular basement membrane—all critical functions for podocyte health and the integrity of the glomerular filtration barrier. Interestingly, genes involved in post-synaptic membrane organization are significantly enriched and upregulated specifically in the Wt1 knockout model, suggesting a model-specific effect on synaptic-like structures involved in podocyte cell-cell interactions. Given that many of these genes are implicated in tight-junction organization, this finding is particularly important for understanding how Wt1 loss affects podocyte integrity and contributes to disease progression. 2D GO analysis in [Supplementary Figure 4.4](#) provides an additional view into differences between Wt1 and Nphs2 mutation effects on the podocyte transcriptome.

The results of Gene Set Enrichment Analysis (GSEA) (Fig. 4.2.4) provide further insights into the functional consequences of these genetic mutations on podocyte biology. The focal adhesion complex and trafficking and processing of endosomal TLR are consistently downregulated, while epithelial-mesenchymal transition (EMT) is stably upregulated in both models. This suggests a common pathophysiological response of podocytes to both mutations, where impairment of adhesion complexes and enhanced EMT are indicative of podocyte detachment and phenotypic shift, which are known to contribute to podocyte depletion and glomerular dysfunction.

We also observed notable divergence in pathways between the two models, indicating differential transcriptional responses to podocyte damage. Specifically, TNF-alpha signaling via NFkB and innate immune response pathways showed opposing patterns of regulation: both were downregulated in the Wt1 model but upregulated in the Nphs2 model. This suggests that although both mutations contribute to FSGS progression, they do so via distinct inflammatory responses, potentially driven by differences in how each mutation affects podocyte resilience and signaling cascades.

Further, antigen processing and presentation, protein processing in the endoplasmic reticulum (ER), fluid shear stress, cellular response to heat stress, and HSF1 activation were all downregulated in the Wt1 mutation model, reflecting a general impairment in stress response mechanisms and ER function. These processes are essential for proper protein folding and cellular stress management, indicating that the absence of Wt1 disrupts cellular homeostasis and contributes to podocyte vulnerability. In contrast, the Nphs2 mutation specifically led to the downregulation of N-glycan biosynthesis, which is a necessary step for producing functional podocin and other tight junction proteins. The disruption of N-glycosylation likely reflects a substrate limitation caused by the mutation, contributing to compromised podocyte junctions (Reily et al. 2019).

Interestingly, several functions, such as muscle contraction, platelet homeostasis, and nervous system development, were continuously upregulated in the Nphs2 model but not in the

Wt1 model. This suggests that Nphs2 mutations may induce broader systemic changes in cellular signaling that are absent in the Wt1 context. These upregulated pathways could represent compensatory mechanisms or maladaptive changes contributing to the progression of FSGS in Nphs2 mutants. Additionally, chromatin modifying enzymes, which regulate gene expression epigenetically, were consistently downregulated in Nphs2 mutants, potentially reflecting altered chromatin architecture and transcriptional regulation as part of the podocyte response to the loss of functional podocin.

In summary, our analysis highlights both convergent and divergent biological effects of Wt1 and Nphs2 mutations on podocyte cellular function, particularly regarding signaling pathways, immune response, and cellular stress adaptation. Shared disruptions in adhesion, EMT, and cytoskeleton dynamics underscore common pathways leading to podocyte dysfunction. However, the contrasting regulation of inflammatory pathways and stress responses between the two models reveals genotype-specific effects, providing insights into how different podocyte injuries can differentially drive disease progression.

#### 4.2.3 Limitations of Glomerular Transcriptomics for Studying Podocyte Damage

##### **PCA analysis of bulk and single-nuclei RNA-seq samples**

To get a holistic view of differences between glomerular and podocyte samples, we generated pseudo-bulk glomerular and podocyte RNA-seq from snRNA-seq and compared them to bulk glomerular RNA-seq. We used PCA as the main tool for comparison and analysis of variation between the samples. PCA plot on Figure 4.2.5 A shows that the sequencing technology, as expected, is a largest source of variation, clearly separating bulk and pseudo-bulk samples along PC1. More interestingly, PC1 also separates podocyte and glomerular pseudo-bulk samples, showing that the sample type explains a great deal of variation in the data. This distinction suggests that despite the similarity in sequencing technology used for the pseudo-bulk samples, the underlying biology of podocytes and the glomerular compartment drives a significant portion of the observed variation.

To decide how many principal components to consider, we performed an analysis of variation (Camargo 2022) for the first 20 PCs, which showed that the first seven contain non-random levels of variation ([Suppl.Fig.4.5 A](#)). We correlated loadings of the first 7 PCs with sample-level experimental variables to understand how much each variable contributes to the variation in the data. Results on Figure 4.2.5 B shows that the first principal component, which accounts for 51.2% of variation in the data, represents a combination of a sample type (glomerular VS podocyte) and a sequencing type (bulk vs single-nuclei). The second PC, which accounts for 14.7% of variation, mostly represents the effect of a sample type. The third PC mostly explains experimental (Wt or Nphs2) effects, the 4th - age of mice, PC5 to PC7 mostly represent the overall genotype effect (wild-type or mutant).

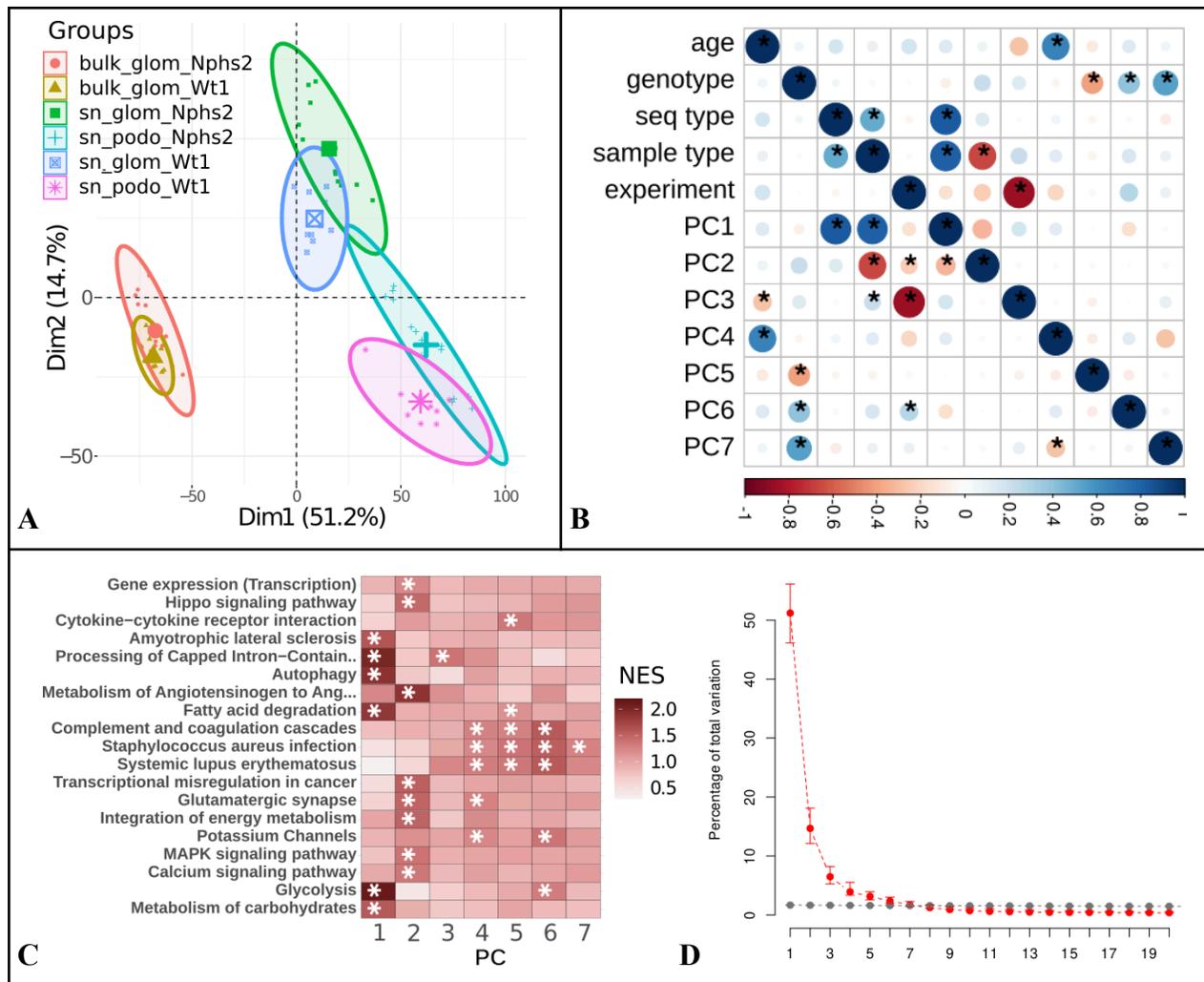


Figure 4.2.5: Comparison of bulk and single-nuclei glomerular RNAseq datasets.

(A) PCA analysis of bulk glomerular, pseudo-bulk glomerular and pseudo-bulk podocyte samples. Pseudo-bulk was generated from snRNA-seq. (B) Spearman correlation heatmap of sample level covariates and PCs. “Genotype” variable has 2 classes: control or mutant, “experiment” variable also has 2 classes: Nphs2 or Wt1. (C) Functional annotation of the 7 first loadings with KEGG and Reactome pathways using gene set enrichment test. Significant results ( $q$ -value $<0.05$ ) are marked with an asterisk. Color shows normalized enrichment score (NES). (D) Bootstrapped and randomized percentage of variation for the first 20 PC's of PCA analysis of bulk and pseudo-bulk single-nuclei RNA-seq. Whiskers show 95% CI.

In order to understand which aspects of biology are reflected in the analyzed PCs and, by extension, in the experimental variables, we performed pathway annotation of PC loadings (Suppl.Fig.4.5 B). The results show that genes contributing to the first PC (sequencing and sample type effects) are enriched in metabolic functions, like *glycolysis*, in *fatty acid metabolism* and *autophagy*. This suggests that differences in cellular metabolic activities are not only a consequence of sequencing artifacts but also may reflect distinct metabolic states of podocytes

versus glomerular cells. The second PC (sample type effect) is represented by *metabolism of Angiotensinogen to Angiotensin, Hippo, MAPK, Calcium signaling pathways* and genes related to *synapses*. The enrichment of these pathways in PC2 suggests that podocytes and other glomerular cells have distinct roles in maintaining kidney function, particularly with regard to signaling processes that regulate cell communication, stress responses, and homeostasis. The presence of synapse-related genes may also indicate the importance of specialized cell-cell contacts and signaling mechanisms that are unique to podocytes, which contribute to their ability to adapt and respond to changes in the glomerular environment. 3rd PC (batch and model effects) is represented only by one term *Processing of Capped Intron-Containing Pre-mRNA*. PC 4 (age effect) and PCs 5-7 (genotype/damage effect) are enriched in *immune function* genes. Additionally, PC 4 is enriched in *synapse* and *potassium channel* genes; PC 5 is enriched in *Cytokine-cytokine receptor interaction*; PC6, which also represents a model effect, is enriched in *glycolysis* and *potassium channels*. PC6 is particularly interesting, as it correlates both with the experimental condition (Nphs2 or Wt1) and the genotype (control or mutant). This dual correlation suggests that PC6 may represent the differential effects of Wt1 and Nphs2 mutations on the transcriptomes of glomerular and podocyte samples. The enrichment observed in glycolysis and potassium channel pathways for PC6 highlights potential shared and distinct metabolic and ion channel-related changes that are characteristic of these specific genetic mutations.

In conclusion, our PCA analysis highlights the complex interplay of sequencing technology, biological sample type, experimental batch, and genotype in shaping the transcriptional landscape of glomerular and podocyte samples. By integrating pathway analysis, we were able to link specific PCs to meaningful biological processes. Results show that differences in underlying biology of podocytes and the glomerular compartment drives a significant portion (PC1 and PC2) of the observed variation in analyzed samples, significantly exceeding effects of mutation and highlighting the intrinsic biological differences between these two cell populations.

### **Comparison of bulk and snRNA-seq differential expression**

In this analysis, we compared the effects of Wt1 and Nphs2 mutations on glomerular and podocyte transcriptomes using differential expression analysis across bulk glomerular, single-nuclei glomerular, and podocyte RNA-seq samples. Our goal was to determine whether these mutations induce unique or overlapping transcriptional changes in podocytes and the broader glomerular compartment.

Comparing the results of differential expression (DE) analysis between mutant and control cells (Suppl. Fig. 4.5 B) showed that the number of differentially expressed genes (DEGs) due to the Nphs2 mutation in podocytes and all glomerular cells was comparable. In contrast, the Wt1 mutation resulted in fewer DEGs in all glomerular cells compared to podocytes, where the number of DEGs was similar to that observed in the Nphs2 experiment. This suggests that Wt1 mutation effects are more pronounced in podocytes than in the broader glomerular population, whereas Nphs2 effects are consistently evident across both podocytes and glomerular cells.

An analysis of effect sizes, presented in Figure 4.2.6 A, further supports these findings by comparing log fold changes (LFC) in gene expression. The results indicated that the *Nphs2* mutation had a greater impact on both the glomerular and podocyte transcriptomes compared to the *Wt1* mutation. Moreover, changes in the podocyte transcriptome were more pronounced than changes in the glomerular transcriptome in the *Wt1* experiment, while the opposite pattern was observed for the *Nphs2* mutation, where glomerular changes surpassed those observed in podocytes. These observations suggest that the effects of the *Wt1* mutation are primarily driven by changes in podocytes, while the *Nphs2* mutation affects multiple cell populations within the glomerulus.

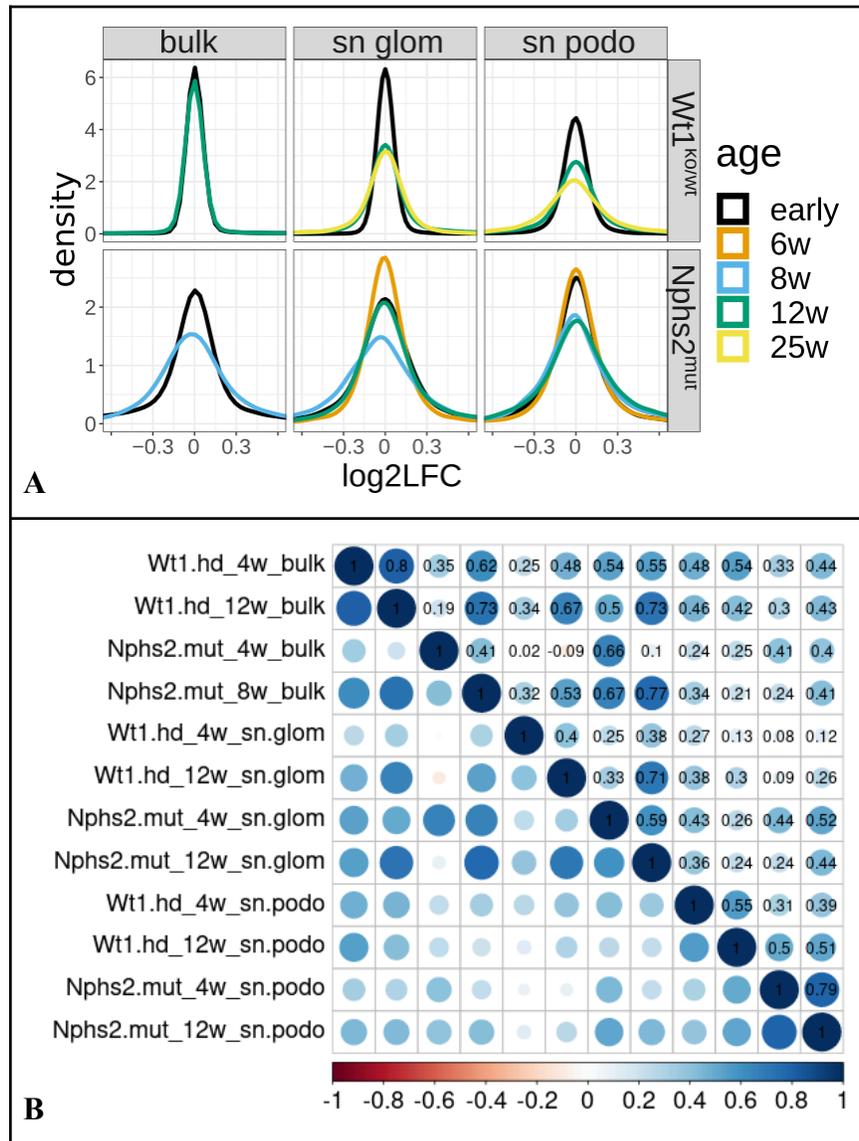


Figure 4.2.6: Comparison of bulk and single-nuclei glomerular RNA-seq datasets. (A) Distributions of LFC values show effect size of genotypes (panel rows) in bulk glomerular, snRNA-seq glomerular and podocyte samples (panel columns), as estimated by differential expression analysis. Line colors show ages of mice. (B) Heatmap of Spearman correlation between LFCs from DE

analysis of control and mutant cells(samples) in mice of specific age. 1249 genes within the upper variance quartile were used.

Taken together, the evidence suggests that differential expression observed in the bulk glomerular transcriptome in response to the *Wt1* mutation is largely driven by podocytes, where the mutation occurs. This interpretation is supported by the higher number of DEGs and larger effect sizes seen specifically in podocytes compared to all glomerular cells. Conversely, the changes in the glomerular transcriptome due to the *Nphs2* mutation appear to reflect a more complex combination of signals from different cell populations, indicating that multiple cell types within the glomerulus contribute to the observed transcriptomic changes.

To further explore these relationships, we performed a correlation analysis of DE results between glomerular and podocyte transcriptomes using log<sub>2</sub> fold change vectors (Fig 4.2.6 B). The results showed that the *Wt1* mutation effects in the glomerular transcriptome correlated more strongly with those in the podocyte transcriptome, which is consistent with the notion that podocytes are the primary source of transcriptomic changes under *Wt1* mutation. In contrast, the correlation between glomerular and podocyte transcriptomes for the *Nphs2* mutation was weaker, suggesting that the transcriptomic response in glomerular cells involves contributions from multiple cell types, resulting in a more heterogeneous signal.

To conclude, this analysis suggests that bulk glomerular transcriptome may be a reasonable proxy of the podocyte transcriptome in FSGS studies but only when the primary podocyte damage is relatively mild and the expected rate of disease progression is low. However, for studies of specific aspects of podocyte TRN we strongly recommend using FACS sorted podocytes or single-cell technology. The latter also allows analysis of cell-type interactions and analysis of podocyte damage trajectories.

#### 4.2.4 Discussion

In this chapter we employed single-nucleus RNA sequencing (snRNA-seq) to dissect the transcriptional landscape of podocytes in two genetic mouse models of focal segmental glomerulosclerosis (FSGS): the heterozygous *Wt1* knockout (*Wt1*<sup>ko/wt</sup>) and the *Nphs2* mutant (*Nphs2*<sup>mut</sup>) mice. By analyzing isolated glomeruli from these models at various stages of disease progression, we aimed to understand how genetically induced podocyte damage influences cellular composition and gene expression within the kidney, thereby contributing to the pathogenesis of FSGS.

##### **Findings and their implications for Podocyte Biology and Kidney Disease**

Cell-type annotation revealed a consistent cellular composition across both models, with the major glomerular cell types—podocytes, endothelial cells, and mesangial cells—forming the largest clusters. Notably, we observed a significant decrease in the proportion of podocytes in the *Nphs2*<sup>mut</sup> mice compared to wild-type controls, even at early disease stages. This reduction was less pronounced in the *Wt1*<sup>ko/wt</sup> mice and became more evident at later stages. The decline in podocyte numbers correlated strongly with measures of kidney function, such as the albumin-to-creatinine ratio, underscoring the critical role of podocyte integrity in maintaining glomerular filtration. These findings provide valuable insights into the cellular dynamics underlying FSGS progression. The differential rates of podocyte loss between the two models reflect the severity and kinetics of disease manifestation associated with each genetic mutation. In *Nphs2*<sup>mut</sup> mice, the rapid decline in podocyte numbers suggests a more aggressive disease course, likely due to the critical role of podocin in maintaining slit diaphragm integrity. The *Wt1*<sup>ko/wt</sup> mice exhibited a slower progression, consistent with the role of WT1 as a transcription factor essential for podocyte differentiation and maintenance.

The differential expression (DE) analysis of podocytes further highlighted the distinct and shared molecular pathways affected by each mutation. In *Nphs2*<sup>mut</sup> podocytes, a higher number of genes were differentially expressed compared to *Wt1*<sup>ko/wt</sup> podocytes, aligning with the observed phenotypic severity. Interestingly, already at the age of 4 weeks *Nphs2*<sup>mut</sup> podocytes showed significant transcriptional changes, whereas *Wt1*<sup>ko/wt</sup> podocytes exhibited fewer alterations, suggesting that the *Nphs2* mutation induces earlier and more profound transcriptomic disruptions.

Functional annotation of DE genes revealed common pathways implicated in cytoskeletal organization, cell adhesion, and stress responses—processes vital for podocyte function and survival. Both models demonstrated upregulation of epithelial-mesenchymal transition (EMT) pathways and downregulation of focal adhesion components, indicating a shared maladaptive response contributing to podocyte detachment and loss. However, there were notable differences; for instance, inflammatory pathways such as TNF- $\alpha$  signaling via NF- $\kappa$ B were upregulated in *Nphs2*<sup>mut</sup> podocytes but downregulated in *Wt1*<sup>ko/wt</sup> podocytes, suggesting mutation-specific inflammatory responses.

The expression patterns of the mutated genes themselves provided additional insights. The upregulation of *Wt1* and *Nphs2* transcripts in early disease stages may reflect a compensatory mechanism attempting to restore functional protein levels. However, the lack of functional protein due to the mutations could lead to further dysregulation of the podocyte transcriptional regulatory network (TRN), exacerbating the disease process. This notion aligns with the concept of a maladaptive feedback loop, where initial compensatory responses become detrimental in the context of persistent genetic defects.

### **Limitations of Bulk Transcriptomics and Advantages of Single-Cell Approaches**

Our comparative analysis between bulk glomerular RNA-seq and snRNA-seq data underscored the limitations of bulk transcriptomics in resolving cell-type-specific changes. Principal component analysis revealed that sample type (glomerular or podocyte) and sequencing technology were major sources of variation, with the inherent heterogeneity of glomerular samples potentially masking podocyte-specific transcriptional alterations. The bulk RNA-seq data showed discrepancies in the expression of *Wt1* and *Nphs2* genes, compared to the snRNA-seq data, likely due to changes in cell-type composition and the proportion of podocytes in diseased glomeruli.

The snRNA-seq approach allowed us to isolate and analyze podocyte-specific transcriptomes, providing a clearer picture of the molecular events occurring within this critical cell type during FSGS progression. By controlling for cell-type composition, we could more accurately attribute observed transcriptional changes to podocyte-specific responses rather than to shifts in the overall cellular makeup of the glomerulus.

### **Broader Context in Kidney Disease Research**

Our study contributes to the broader understanding of podocyte biology in the context of glomerular diseases. Podocyte loss is a hallmark of many forms of CKD, and understanding the mechanisms driving this loss is essential for developing effective therapies. The identification of shared pathways between different genetic models suggests the existence of common therapeutic targets that could be exploited across various forms of FSGS. For instance, interventions aimed at stabilizing the cytoskeleton or enhancing cell adhesion might mitigate podocyte detachment and loss.

Moreover, the mutation-specific differences observed highlight the importance of personalized approaches in treating kidney diseases. Understanding the distinct molecular responses elicited by different genetic mutations can inform the development of targeted therapies that address the specific pathways disrupted in each case.

### **Limitations of the Study**

Despite the valuable insights gained, our study has several limitations. Technical biases inherent to snRNA-seq, such as differences in cell-type capture efficiency and transcriptional dropout, may affect the interpretation of cell-type proportions and gene expression levels. The observed discrepancies in podocyte fractions between the two models' control samples suggest

potential batch effects or strain differences that could influence results. While we attempted to account for these factors, they may still impact the generalizability of our findings.

The sample sizes, particularly for certain age groups and genotypes, may have limited the statistical power of our analyses. The inability to detect significant changes in some cell-type abundances could be due to insufficient numbers of replicates or cells captured. Future studies with larger cohorts and more uniform sampling across disease stages would strengthen the conclusions drawn.

Additionally, while our functional annotations provide hypotheses about the pathways involved in podocyte damage, further experimental validation is necessary to confirm the roles of specific genes and pathways. *In vitro* studies or *in vivo* models with targeted interventions could elucidate the causal relationships between gene expression changes and podocyte dysfunction.

### **Future Directions**

Building on our findings, future research could explore the therapeutic potential of modulating the shared pathways identified, such as EMT and focal adhesion signaling. Investigating the upstream regulators of these pathways may reveal novel drug targets. Moreover, longitudinal studies tracking individual podocytes over time could provide insights into the temporal dynamics of gene expression changes during disease progression.

Integrating additional omics data, such as proteomics or epigenomics, could offer a more comprehensive understanding of the molecular alterations in podocytes. Employing spatial transcriptomics would also enable the examination of podocyte interactions with neighboring cells within the glomerular architecture, shedding light on the microenvironmental factors contributing to disease.

### **Conclusion**

This study underscores the value of single-cell transcriptomic approaches in unraveling the complex molecular underpinnings of podocyte injury in FSGS. By dissecting the cell-type-specific responses to genetic mutations, we have highlighted both common and distinct mechanisms driving disease progression. These insights improve our understanding of podocyte biology and may, potentially, assist development of targeted therapies aimed at preserving podocyte function and preventing CKD progression.

### 4.3 Single Cell Damage Score in Podocytes

Based on results of our analysis, described in previous chapters, and published research on transcriptional dysregulation in damaged podocytes, we hypothesized that different causes of podocyte damage converge on common mechanisms of transcriptional regulation. Cellular damage can be defined in various ways. In this study, we adopt an operational definition tailored to our specific objectives. We define cellular damage as a pathological cellular state that is not specific to any one disease-causing mechanism but rather represents a common endpoint resulting from a broad spectrum of disease entities and mechanisms. Such cellular damage states are well established in histopathology. Here, we develop an approach for the molecular characterization of these states.

#### 4.3.1 Deriving Universal Podocyte Damage Signature

We propose that a unified transcriptome signature can be used to calculate damage in individual podocytes, regardless of the underlying disease etiology. We also assumed that this gene signature can be derived in a supervised manner, by meta-analysis of transcriptomic data of various models of kidney damage (Figure 4.3.1, development). To test our concept, we developed a Podocyte Damage Score (PDS)—a single-cell damage metric for podocytes—to track cellular damage during the progression of focal segmental glomerulosclerosis (FSGS). The key idea behind the PDS is to identify genes whose expression is associated with increasing damage across a wide range of kidney damage models. By incorporating diverse disease models in the gene selection process, we ensure that the marker genes are not merely indicative of one specific intervention but reflect damage to the target cell type in general. Furthermore, our approach guarantees that the selected marker genes are expressed within podocytes themselves, thereby excluding genes specific to tissue-resident immune cells that might only indicate inflammation rather than direct podocyte damage.

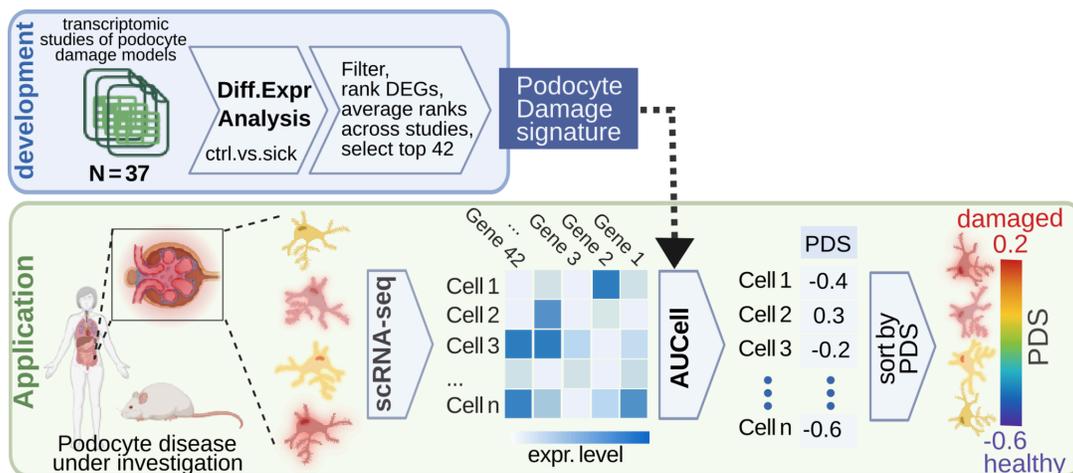


Figure 4.3.1: Damage score generation and application schemes.

To generate the damage score we collected and analyzed gene expression data from 37 studies covering multiple FSGS disease models, limiting our selection to the glomerular and podocyte specific data, when possible (see [Suppl. Table 4](#)). We performed differential expression analysis in each of these models, always comparing the diseased/experimental condition versus the respective control. Using results of DE analysis we first ranked genes by their p-values for differential expression in each of these models and then averaged ranks across all disease models to obtain one global gene ranking for podocytes. Next, we applied two additional filters: first, we removed genes that were not expressed in the target cell type, i.e. podocytes. Second, we only kept genes with a consistent fold change (same sign) in at least 75% of all studies. Finally, we chose the top  $n$  genes as podocyte damage marker genes. In order to quantify the damage of an individual cell (or bulk sample) we tested the ranking of the marker genes in a given cell using AUCCell (Aibar et al. 2017). Scores for genes with positive and negative log-fold changes were combined considering the direction of effects so that experimental samples/cells from damaged kidneys are expected to have a higher podocyte damage score than controls (Figure 4.3.1, application).

#### 4.3.2 Testing and Validating the Podocyte Damage Score

Figure 4.3.2 shows an example of the score application to single-nuclei RNA-seq data, *Wt1* model of genetic FSGS, and illustrates the ability of the score to show gradient of the damage, consistent with the genotype labels. Next, we ensured robustness of the score by extensive testing and validation.

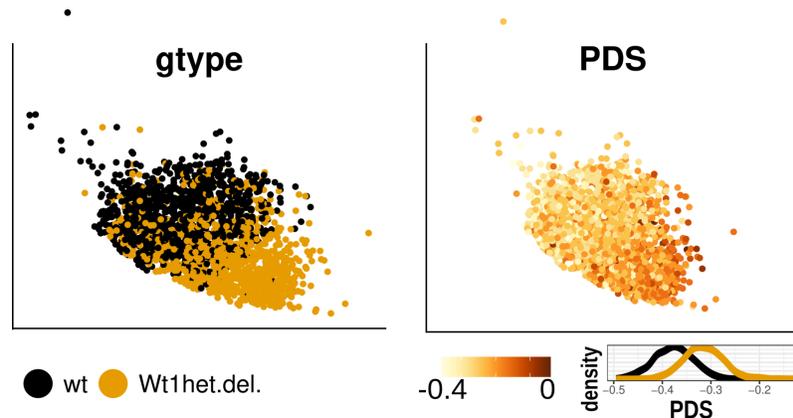


Figure 4.3.2: UMAPs of podocytes from *Wt1*<sup>ko/wt</sup> snRNAseq data show application of podocyte damage score (PDS) to single cell RNA-seq. Black and orange colors on the left panel denote control and experimental conditions, correspondingly. On the right panel cells are coloured by PDS.

#### Robustness tests

To establish the robustness of our approach we first confirmed that varying the number of top genes used for the damage scores had little influence on the outcome. Varying the top  $n$  genes

between 10 and 100 genes affected the damage scores only minimally. Actually, using as little as 5 genes allows to differentiate populations of healthy and sick podocytes but results are less stable, with high variance and many outliers within the groups. On the other hand, using more than 100 genes actually decreases correlation between PDS and proteinuria as we shown in the results of functional validation. To conclude, gene set size to use with AUCCell for calculating PDS can be picked in the range between 10 and 100 genes

Likewise, randomizing (Material and methods 2.3.4.2) parts of the input data had little influence on the ability of the damage scores to distinguish healthy from diseased samples (Figure 4.3.3 A). Even when randomizing (by shuffling the p-values and LFC of the differentially expressed genes per disease model) up to 70% of all input studies, the resulting damage scores were still able to distinguish experimental from control samples in disease models that were not used for developing the damage scores. Thus, the damage scores are remarkably robust to changes of the number of marker genes and variation in the training data - the choice of the top  $n$  genes used for computing the damage scores is to an extent arbitrary. We decided to continue all subsequent work with the top 42 marker genes, listed in the [Supplementary table 5](#).

### Model Cross-validation

A central aim of our study was to develop scores that measure cellular damage across disease conditions instead of reflecting molecular adaptations in specific disease models. We tested this notion using a ‘leave one model out’ cross validation: first we removed all samples of one disease model from the training data, then performed the procedure described above to determine damage marker genes on the remaining samples and, finally, tested the resulting damage scores on the left out disease model samples. In this way we cross-validated 4 models of podocyte damage that had 5 or more datasets available, Namely, diabetic, slit diaphragm damage, transcription factor perturbation and toxic damage models. In each of the cross-validation rounds the damage score was able to separate experimental samples from control samples in the unseen disease model samples (Figure 4.3.3 B).

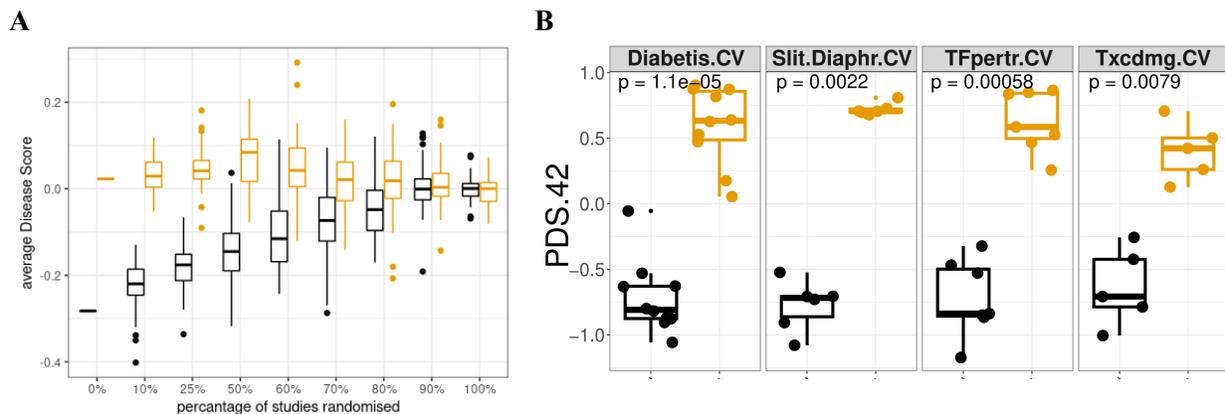


Figure 4.3.3: Robustness tests of the podocyte damage score.

(A) Results of the damage signature randomisation test, X-axis shows levels of randomisation - percent of

the datasets randomized. Boxplots show results of 50 rounds of randomisation at the respective randomness level. scRNAseq nephritis data from GSE146912 was used as a test dataset. In both panel A and B the Y-axis shows PDS; black and orange colors denote control and experimental conditions, respectively. **(B)** Model cross validation of PDS, each dot represents mean score across all samples/cells of a study, scores are scaled and centered per study, p-values of Wilcoxon rank test are shown..

In addition we cross-validated transcriptome quantification technologies, to make sure the signature is not specific to any transcriptomic platform: [Supplementary Figure 4.7](#) shows that transcriptome quantification platform has a very little effect. Altogether, these results demonstrated that the damage scores are remarkably robust against diverse variations in the data and that they capture generic molecular changes operating across disease models.

### **Supervised VS unsupervised damage signatures**

As an alternative to the supervised approach of aggregating DE results, the gene signature can be generated in an unsupervised fashion, by trajectory analysis of single cell data (Material and methods 2.3.4.3). Damage score calculated with the pseudo-time, unsupervised damage signature performs comparably to the original damage score, generated in a supervised fashion ([Supplemental Figure 4.8 A](#)). Moreover, the pseudotime damage signature is similar to the supervised podocyte damage signature, showing convergence of the approaches ([Supplemental Figure 4.8 B](#)). We speculate that unsupervised approaches may become more powerful, as more single cell/nuclei studies of podocyte damage become available. For now, the unsupervised approach for the signature generation has drawbacks: the pseudo-time signature is affected by the data integration process, and it poses difficulties for combining bulk and single cell studies.

### **Gene expression of the podocyte damage markers**

The majority of marker genes (39 from 42) were found to be downregulated in response to damage, according to the differential expression analysis conducted on public datasets to generate the damage signature ([Supplementary Table 5](#)). We also characterized expression levels of the 42 damage markers in our bulk and single-nuclei RNA-seq profiles of *Wt1* and *Nphs2* models of genetic FSGS ([Supplementary Figure 4.6](#)). Among the damage signature genes, only *Ankrd1*, *F2r*, *Nexn* were consistently upregulated in public datasets, and only Nexilin (*Nexn*) was consistently upregulated in our data. The set of downregulated genes includes known regulators of podocyte biology, such as *Wt1* and *Mafb*, and other essential podocyte genes that show high expression levels in podocytes at the transcript and protein levels. Notably, the marker set is enriched in highly cell type-specific genes. These results suggest that cellular damage in podocytes proceeds in conjunction with cell identity loss.

### **Protein expression of the podocyte damage markers**

To explore behavior of podocyte damage markers on the protein level we analyzed (Material and Methods 2.3.4.4) two proteomics datasets: a study of *Nphs2*<sup>mut</sup> model of podocyte damage and a protein atlas of glomerular cells (Butt et al. 2020) (Hatje et al. 2021)). Results of the

analysis are visualized with expression plots, where we show detected proteins of the damage markers, whose transcripts are downregulated in damaged podocytes according to the damage signature ([Supplementary Table 5](#)). Figure 4.3.4 shows expression of the damage markers on protein level in the Podocin mutant mice. Not surprisingly, upon damage many damage markers change on protein level in the same direction as on the transcript level. For example, *Magi2* and *Ddn* and other damage markers are consistently down-regulated on the protein level upon podocyte damage.

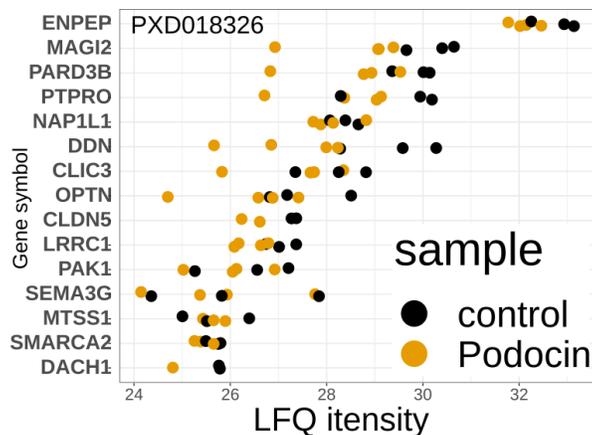


Figure 4.3.4: Expression (X-axis) of podocyte damage score markers (Y-axis) on the protein level. Data from PXD018326.

[Supplemental Figure 4.9](#) shows protein expression levels of PDS markers across 3 main cell-types of healthy glomeruli ((Hatje et al. 2021). Notably most of the damage markers have higher expression levels in the podocyte cell, compared to mesangial and endothelial.

### Morphological validation

To validate the association of the PDS with histologically detectable damage we used two morphological features: first, the slit diaphragm length is a proxy for damage to the primary filtration barrier, with shorter length indicating greater damage. Second, podocyte damage leads to a swelling of glomeruli, thus larger glomeruli being indicative of greater damage.

We used STED imaging to assess and localize expression of *Thsd7a*, a member of the PDS signature and a slit-diaphragm protein marker, which confirmed the association of this PDS gene with podocyte damage ([Supplementary figure 4.10](#)). Second, we used spatial transcriptomic data from mouse kidney to calculate and then test a link between glomerular area and glomerular PDS. We observed a significant correlation between the two measures on the individual sample level (Figure 4.3.5) and then confirmed a highly significant correlation between PDS and the area-adjusted glomerular count both in control and damaged glomeruli ([Supplementary Figure 4.11](#)) by analyzing multiple samples from healthy and damaged kidneys (Materials and Methods 2.3.4.5).

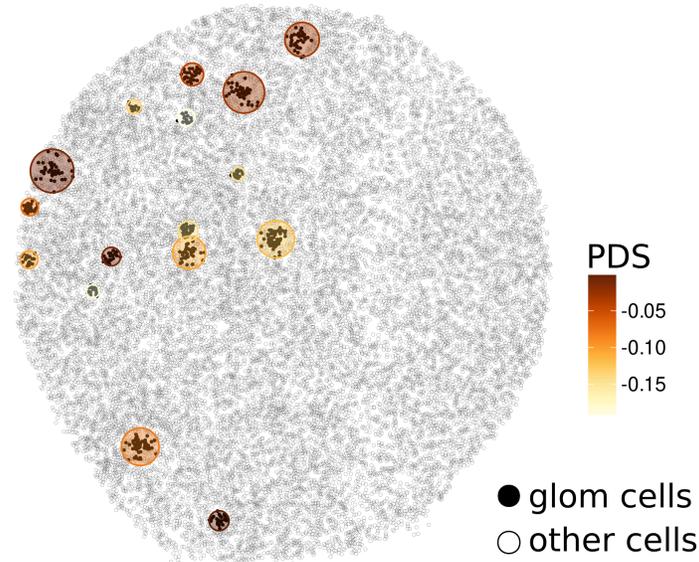


Figure 4.3.5: Spatial validation of PDS using murine kidney Slide-seqV2 spatial transcriptomics data. Spatial plot shows a section of the diabetic mouse kidney (GSM5713367 GEO sample (PMID: 35372810)), where glomeruli are colored by the estimated PDS. Correlation analysis of PDS and glomerular morphological metrics see in [Supplementary Figure 4.11](#).

### Functional validation

To test if the PDS also reflects functional decline of podocytes, we correlated the PDS with albuminuria measurements in mice. The Albumin/Creatinine ratio in Mg/Mg units (ACR) quantifies excess proteins in the urine and is an effective and established measure of kidney function and podocyte damage in particular (Hoefield et al. 2011). We analyzed 11 mouse studies with combined bulk tissue transcriptomic and Albumin/Creatinine measurements, and found a significant positive correlation (Spearman=0.65) between PDS and proteinuria ([Supplementary Figure 4.12 A](#)). The correlation was even stronger (Spearman=0.8) for our single nuclei data, where we aggregated PDS across cells of each sample and then correlated these pseudo-bulk PDS with proteinuria (Figure 4.3.6 A).

### Validation in Human data

Next, we tested if the PDS that was developed on mouse data could also be applied to human data. To transfer the damage signature to humans we restricted the 42 marker genes from mouse to 40 one-to-one orthologous genes in humans. We used publicly available sc/snRNA-seq data and anonymised patient data from the Kidney Precision Medicine Project (KPMP) to calculate and correlate PDS with clinical endpoints. This analysis resulted in a significant correlation between the PDS and Albuminuria measurements (ACR) in a diverse (by sex, age, race) human population (Figure 4.3.6 B). In addition, we showed that it is possible to calculate PDS in human urine scRNAseq data (NEPTUNE study, GEO accession GSE176465) (Latt et al. 2022) and that it correlates with a measure of kidney function called Urinary Protein Excretion Estimation

(UPCR) ([Supplementary Figure 4.12 C](#)). Our finding is in line with the recent publication showing that monitoring urinary podocyte markers can help determine the level of current disease activity and response to therapy in progressive glomerular diseases (Fukuda et al. 2024). To conclude, PDS developed on mouse data discriminates between control and experimental samples and reflects kidney function in the human data.

### Morphological validation in Human data

Lastly, we examined the relationship between the podocyte damage signature (PDS) and glomerular morphology using human data. We downloaded and analyzed 10X Visium spatial transcriptomics data from a human kidney exhibiting signs of FSGS (Materials and Methods 2.3.4.5). Glomeruli were annotated manually, and the PDS was calculated for all glomerular regions. We then overlaid the calculated PDS on the corresponding histological image, revealing a correlation between PDS and glomerular morphology at the level of the sequenced glomerular spots (Figure 4.3.6 C). A meta-analysis of all glomerular spots confirmed a highly significant correlation between PDS and the extent of glomerular damage (Supplementary Figure 4.12 D).

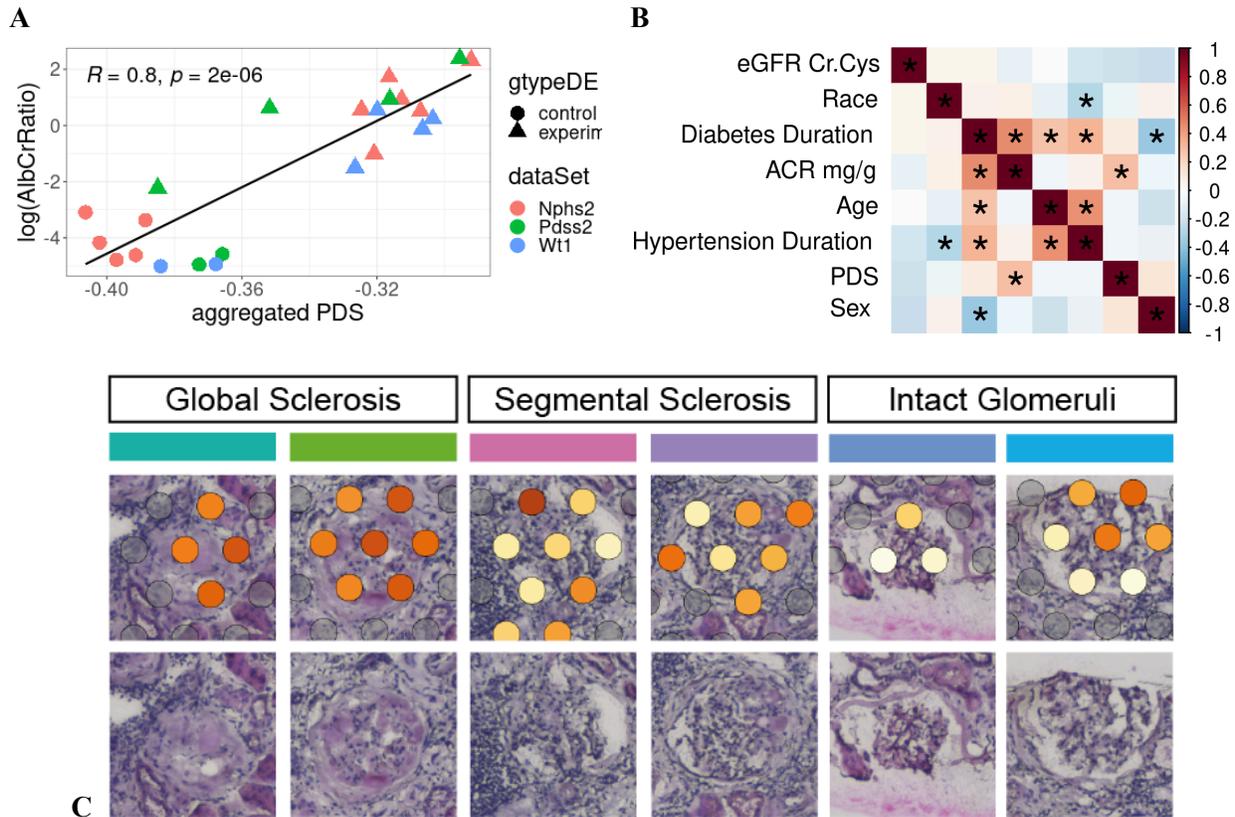


Figure 4.3.6: Functional and morphological validation of PDS in mouse and human data. (A) Association of albuminuria (Y-axis) with the PDS (X-axis) for three different disease models (Nphs2, Pdss2, Wt1). Cells of each sample were aggregated ('pseudo-bulk'), dot color and shape represent the study and experimental status of the sample, correspondingly. (B) Spearman correlation between PDS and

clinical traits for AKI and CKD patients, asterisks denote correlations with p-value < 0.05. In 54 samples with available albuminuria measurements and sc (N=17) and/or snRNA-seq (N=37) data were used.

(C) Histological image with overlaid results of PDS analysis shows damage gradient in individual human glomeruli, annotated manually. 10x visium spatial transcriptomics of a kidney sample from a human with FSGS, patient 29-10282, KPMP. See the full image in [Supplementary Figure 4.13](#).

The established correlations between the podocyte damage score and both glomerular function and morphology demonstrate that the PDS developed on murine data can be effectively applied to human datasets. Building on this, we next characterized the cellular changes that occur in various mouse models of FSGS, using the PDS as a proxy for podocyte damage.

### 4.3.3 Characterizing Cellular Changes in Disease Progression with PDS

Podocyte damage score offers an unprecedented ability to compare transcriptomic changes in podocytes across various datasets and models of podocyte injury. To harness this potential, we used PDS as a damage metric to compare eight diverse models of podocyte damage, each profiled with single cell RNA sequencing. The list includes the *Nphs2* and *Wtl* datasets analyzed in the previous chapter; conditional podocyte *Pdss2* knockout snRNA-seq data generated by the Nephrolab; and four models of podocyte damage from a published study (GEO accession GSE146912) (Chung et al. 2020). The dataset includes models of diabetic nephropathy (*Btbr*), conditional podocyte *Cd2ap* knockout (*Cd2ap*), the Adriamycin toxicity model (*Doxo*), and nephrotoxic serum nephritis (*Nephritis*) (see description of models in [Supplementary Table 1](#)). Each model represents a distinct mechanism of podocyte injury, ranging from genetic mutations to toxic or immune-mediated damage, offering valuable insights into different pathways leading to glomerular disease.

#### **Changes in activity of pathways under podocyte damage**

We processed and analyzed *Pdss2* and public datasets using Seurat, extracted podocytes from all control and experimental samples and calculated podocyte damage score for all podocyte cells. By calculating PDS and sorting cells according to their extent of damage, we were aiming to compare damage-related changes in various pathways in an unbiased way, within and between models. Towards this goal we calculated pathway activities (PA) using AUCell (Methods 2.3.4.2) for all KEGG and Reactome pathways in 8 datasets of ([Supplementary Table 1](#)). We then sorted all cells in each study along the PDS axis (PDS is increasing from left to right) and colored cells by changes in activity of a specific pathway. Next we combine results for several pathways in so-called “pathway fingerprint” heatmaps, to get a comprehensive view of pathway changes in a particular model of podocyte damage. Extended pathway fingerprints for all 8 analyzed podocyte damage models can be viewed in [Supplementary Figure 4.14](#).

Figure 4.3.7 A shows pathway fingerprints of Nphs2, Wt1 and Pdss2 models of podocyte damage and provide a rich source of information, allowing us to discern common and unique patterns of pathway changes across the models. Multiple pathways have the same direction of change with increasing podocyte damage across 3 models: *Ephrin signaling*, *Hippo-signaling*, *MAPK*, *Cell-adhesion* are down- while *Oxidative phosphorylation*, *Translation elongation* are up-regulated in 3 datasets. However, several pathways show model specific behavior. For instance, “Regulation of actin cytoskeleton” (KEGG) is upregulated in Nphs2 but downregulated in Pdss2 and Wt1 models upon the damage. Interestingly, two similar pathways may behave concordant in one model but discordant in another. An example would be behavior of *Thermogenesis* and *Oxidative phosphorylation* KEGG pathways: both are increasing in Pdss2 model but have different directions of change in Nphs2 and Wt1 models.

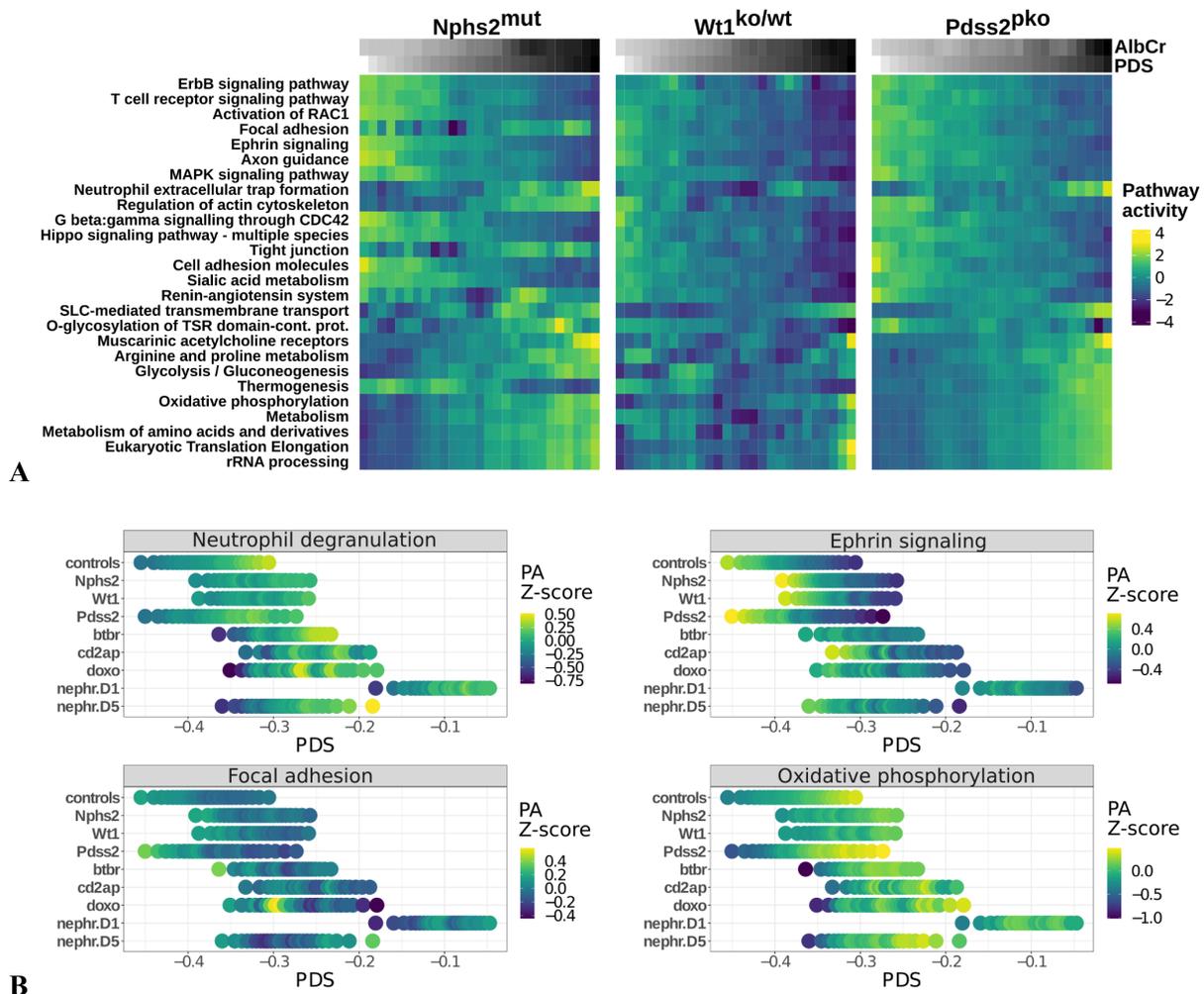


Figure 4.3.7: Analysis of changes in pathway activities caused by podocyte damage. (A) "pathway fingerprint" heatmaps of Nphs2<sup>mut</sup>, Wt1<sup>ko/wt</sup> and Pdss2<sup>pko</sup> KFO models of FSGS. Rows are pathways, columns are PDS bins and the heatmap color shows z-score of pathway activity. A selection of

KEGG and Reactome pathways, each significantly ( $q\text{-value} < 0.1$ ) correlated with PDS in at least one of the 3 studies, is used. Rows are clustered by pathway similarity. Columns of the heatmap are ordered from low (left) to high (right) PDS. Column annotation shows albuminuria (AlbCr) and PDS, averaged across cells of each bin. Heatmaps for all studied podocyte damage models are shown in [Supplementary Figure 4.14](#). **(B)** Comparison of pathway activity changes in individual pathways shows response to the damage ( x-axis ) in various FSGS models ( y-axis ). The heatmap color shows the z-score of pathway activity. Only experimental cells are shown for FSGS models.

When focusing on one pathway or gene at a time, PDS allows to align podocytes with very different degree and etiology of damage and to compare side by side multiple different damage models. Figure 4.3.7 B illustrates this concept by comparing activities (dot color) of individual pathways in podocytes aligned by PDS (X-axis) across 8 models (Y-axis). Results reinforce our earlier findings by showing that *Neutrophil degranulation* and *oxidative phosphorylation* are upregulated while *Ephrin signaling* is downregulated across all 8 models, therefore supporting the hypothesis of convergent podocyte transcriptome remodeling under damage. Strangely, Focal adhesion did not exhibit consistent behavior across models, despite its genes being strongly implicated in the response to podocyte damage. This seemingly contradictory result may be attributed to the nature of pathway-level analysis, which aggregates signals from multiple genes into a single value per cell. While this approach offers an overview, it may obscure finer details, particularly when gene-specific responses vary within the same pathway. This realization prompted us to investigate the behavior of individual genes within the Focal adhesion pathway, aiming to uncover more granular insights that could explain the observed inconsistencies.

To investigate gene-level changes in biological processes induced by podocyte damage, we first calculated the Spearman correlation between PDS and the expression levels of individual genes, separately for control and experimental cells. In the next step, we mapped these correlations onto the nodes of pathway diagrams, as illustrated for the Focal Adhesion network (Figure 4.3.8). This representation allows us to compare FSGS models at the gene level, while keeping the pathway context. Analysis of *Focal adhesion* pathway shows that some genes behave consistently across all models either being up- (Vim, Actn1) or down-regulated (Itgb5, Itga3, Dpp4, Mme, Epb41I5, Parva, Arhgap24) as PDS increases. However, other genes, like Itga2 or Cd151, show model specific behavior.

### Bars of barplots:

1. mean of controls
2. Nphs2 mut.
3. Wt1 het.del.
4. Pdss2 mut.
5. Btbr mut.
6. Cd2ap mut.
7. Doxorubicin
8. Nephritis day5

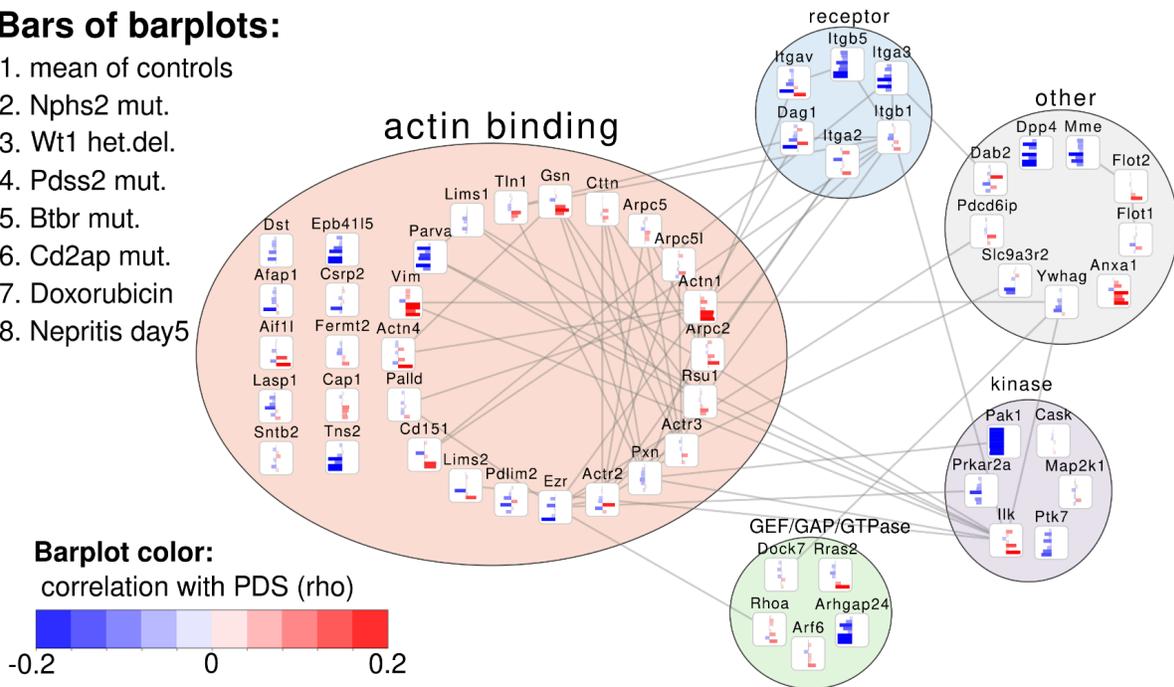


Figure 4.3.8: Integration of Focal adhesion PPI network from C.Shell et.al with PDS analysis shows damage-related gene level change in pathway context. Bars inside each node show spearman correlation (rho) between PDS and expression level of the gene. Red and blue colors depict positive and negative correlations, correspondingly. Saturation of the color and length of the bar reflect the strength of the correlation. Each bar represents, from top to bottom: an average over controls, Wt1het.del, Nphs2.mut, Pdss2, Lmx1b, doxorubicin, Cd2ap, Btbr, Nephritis\_day1 and day5 models of FSGS. Examples of integrating PDS gene-level analysis with Slit Diaphragm PPI network and KEGG Focal adhesion diagram are shown in [Supplementary Figure 4.15 A](#) and [B](#), respectively.

In summary, our analysis of pathway changes using podocyte damage score (PDS) across eight different models of podocyte injury allowed for a comprehensive examination of pathway-level and gene-specific responses. Highlighting convergent and divergent responses to podocyte damage on both levels of analysis, providing valuable insights for understanding the mechanisms underlying FSGS and potential therapeutic targets.

### Gene-correlation analysis of podocyte damage using PDS

The number of genes that significantly ( $q\text{-value} < 0.1$ ) correlate with PDS differs greatly between studies and between samples within each study. Generally, the number of genes that correlate with PDS (Npds.cor) depends on a specific dataset. Trivially, the number of significant PDS correlates scales with the sample size of a dataset, as statistical power increases. Further, we calculated Npds.cor in individual biological samples and correlated with other sample-level estimates. The analysis showed that Npds.cor correlates with expression variance but only between control samples, not between damaged samples ([Supplementary Figure 4.16](#)). The other

variable that predicts the *Npds.cor* is a damage effect size. For example, both the damage and *Npds.cor* in *Wt1ko/wt* samples are much lower than in Nephrotoxic Nephritis samples. From this analysis we may once again conclude that the level of transcriptome perturbation upon podocyte damage, measured by PDS, directly depends on severity of the damage phenotype.

To further characterize mechanisms of divergent transcriptional response to podocyte damage we selected genes that significantly ( $q\text{-value} < 0.1$ ) correlate with PDS across the majority of all analyzed studies. Based on these results we selected 157 genes that correlate with PDS in 7 and more tests, where correlations were tested separately in control and experimental podocytes of each study. To further characterize this extended damage signature we clustered the genes based on their correlation in control and experimental cells and visualized the result as a heatmap (Figure 4.3.9 A). Hierarchical clustering produced 4 major clusters: a middle-size cluster containing the majority of PDS markers, a small and a big cluster with mixed correlations in control and experimental cells and one small cluster with genes correlating only in experimental cells. Note that this comparison would not be possible with conventional differential expression analysis, as no differential expression can be revealed for cells of one sample.

Next we characterize 4 clusters of the extended damage signature by annotating them with Gene Ontologies (Figure 4.3.9 B). Overall, enrichments were often shared between clusters. Since the “damage signature” cluster consists mostly of the 42 universal damage signature genes, it is expectedly enriched in core podocyte functions and implicated in podocyte damage response. “Mixed big” cluster is enriched in cell junction assembly, regulation of actin cytoskeleton, regulation of cell shape and motility, glycosylation. Genes from the “Mixed small” cluster correlate more with PDS in healthy podocytes and are slightly enriched in *Actin depolymerisation*, *interleukin 1* etc. Finally, “experiment only” cluster contains genes that correlate with PDS only in podocytes from damaged samples, it is enriched in *cell-cell adhesion*, *regulation of inflammation*, *leukocyte regulation* etc. By identifying and clustering genes that significantly correlate with PDS across multiple studies, we uncovered distinct gene clusters enriched in specific biological functions, offering deeper insights into the mechanisms underlying podocyte damage and FSGS progression.

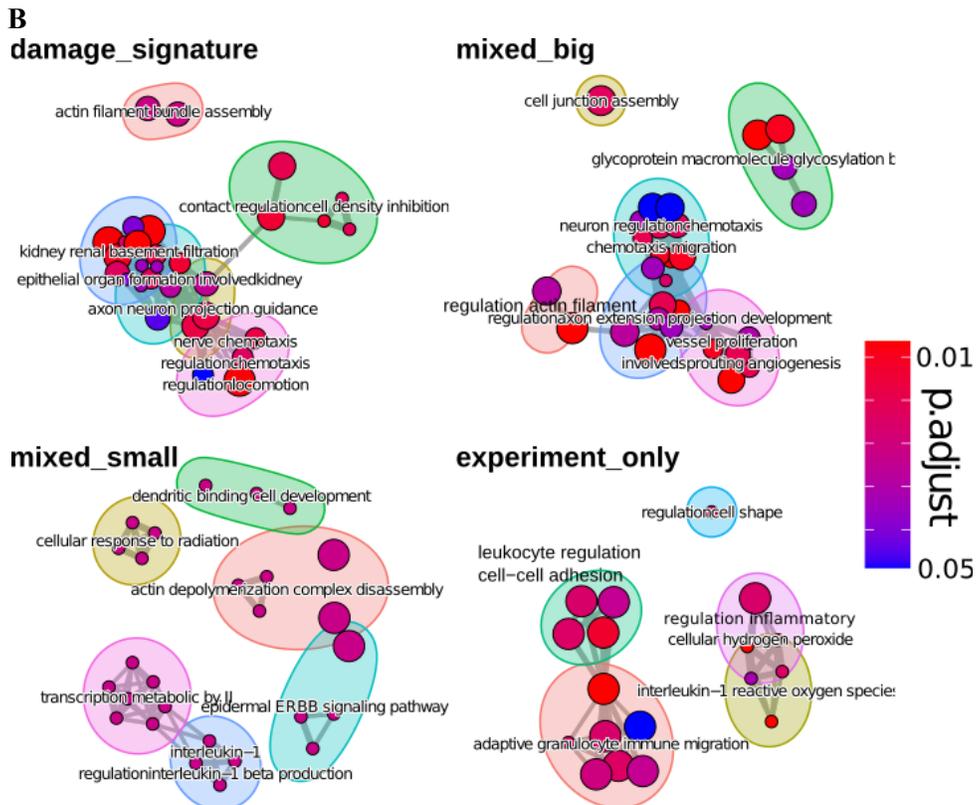
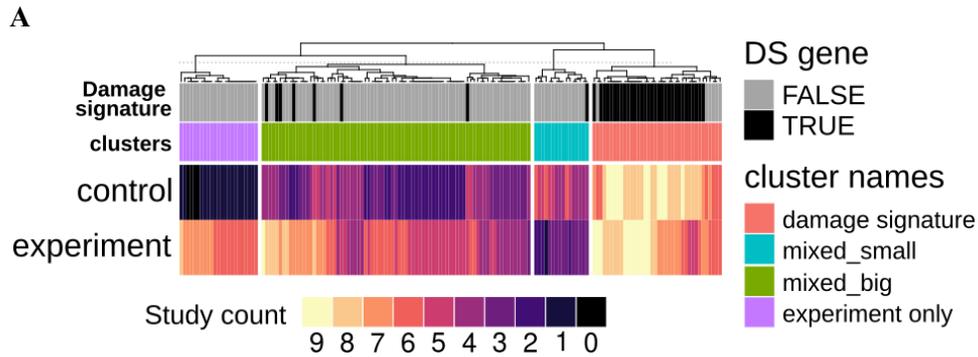


Figure 4.3.9: PDS reveals mechanisms of FSGS

(A) The count heatmap shows genes (in columns) that correlate with PDS in more than 7 conditions colored by the number of studies in which they correlate, results split into control and experimental samples (in rows). (B) GO functional annotation of clusters of genes that highly correlate with PDS. Annotation was performed with ClusterProfiler v4.0 and results summarized in a network diagram by enrichplot R packages (Yu 2018; Yu et al. 2012).

## Transcription factors involved in rewiring of podocyte transcriptome under damage

Finally, we aimed to identify the transcription factors (TFs) driving the transcriptional changes in key podocyte pathways and gene sets correlated with the podocyte damage score (PDS). To gain an unbiased perspective on gene regulation in podocytes, we first reconstructed a podocyte transcriptional regulatory network (TRN) using ATAC-seq data and transcription factor motifs (see Material and Methods 2.3.1.2). This network was then used to identify regulators of transcriptional changes associated with disease progression.

To streamline the analysis, we clustered highly similar TF motifs (Supplementary Figure 4.17 A), which reduced the number of motifs used to interpret PDS-related results. We combined the predicted target genes for all motifs within each cluster, forming a TF regulon, and named the 14 resulting regulons after a representative transcription factor for each cluster.

Next, to predict potential regulators of the gene sets of interest, we tested the significance of their overlap with the target gene sets of these TFs. The results, shown in Figure 4.3.10, indicate that TFs from the *Wt1*, *VeZF1*, and *Irf1* motif clusters are predicted to bind regulatory elements of most genes involved in key podocyte functions as well as genes correlated with PDS. Notably, it is primarily the gene sets correlated with PDS that show statistically significant enrichment in TF targets. Furthermore, these genes are regulated by multiple TFs, both cell-type-specific and general, with the strongest enrichment observed in the **Fos** TF cluster.

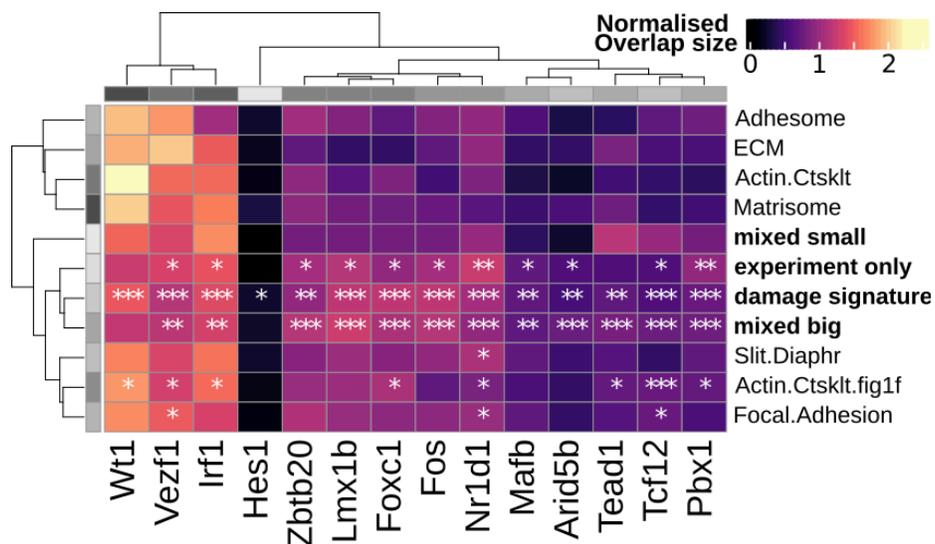


Figure 4.3.10: Results of the podocyte TRN (left) and motif enrichment (right) analysis show putative TF regulators of key podocyte functions. Heatmap color intensity shows the fraction of genes in genesets (rows) regulated by TF clusters (columns). Significance of the overlap was calculated with Fisher exact test, FDR adjusted, and depicted on the heatmap with asterisks: \*\*\*  $q < 0.001$ , \*\*  $q < 0.01$ , \*  $q < 0.1$ . Genes expressed in sc/sn RNA-seq podocyte data were used as the background for the test. Row and column annotation bars reflect size of gene sets, darker color means bigger.

Based on Figure 4.3.10, we also identified a cluster of motifs representing a family of nuclear receptor TFs related to circadian regulation as potential regulators of several key podocyte functions involved in FSGS progression. Circadian regulation has recently been linked to podocyte damage, though this connection is still emerging (L. Wang et al. 2024). Given its potential importance, we decided to investigate this link further.

We performed analysis of circadian (dis)-regulation in FSGS. Using 2 independent methods (Materials and Methods 2.3.4.8) we detected the increase in circadian dysregulation (Figure 4.3.11 A) and an increase in the error of estimated circadian time (Figure 4.3.11 B) with increase in PDS. Our results suggest that circadian dysregulation may be one of the common mechanism involved in rewiring transcription regulatory network under podocyte damage

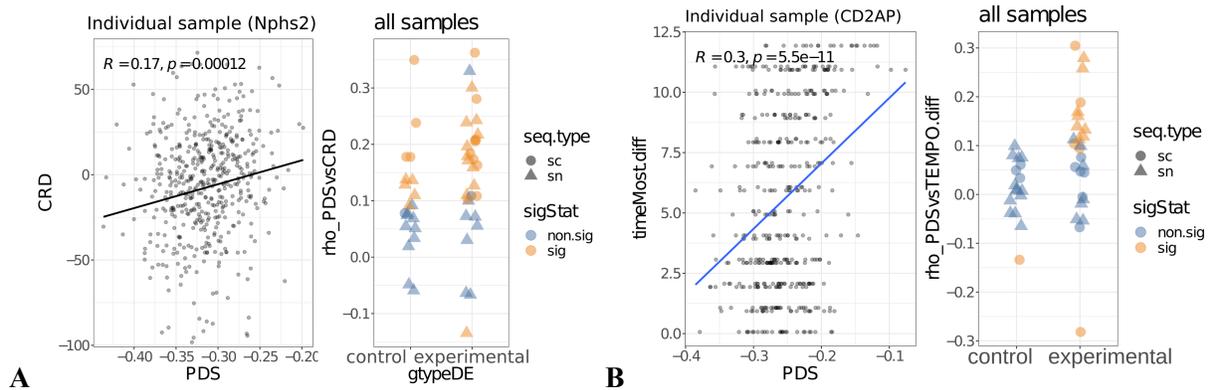


Figure 4.3.11: Association between podocyte damage and circadian regulation

(A) Scatterplot shows significant spearman correlation between circadian rhythm disruption (CRD) and PDS in a sample of Nphs2 mut. snRNAseq study; the dotplot to the right shows correlation coefficients over all individual samples. CRD was calculated using a CRD score algorithm (He et al. 2022).

(B) Scatterplot shows significant spearman correlation between the deviation from predicted circadian time and PDS in an individual sample; the dotplot to the right shows correlation coefficients in all analysed samples. Circadian time was calculated using a Tempo algorithm (Auerbach, FitzGerald, and Li 2022), Clock was used as a reference gene. Results with Arntl as the reference are in [Supplementary Figure 4.17 C](#).

Our exploration of circadian dysregulation as a novel mechanism involved in FSGS progression offers new avenues for understanding how transcriptional reprogramming under podocyte damage may be influenced by disruptions in circadian rhythms. These findings contribute valuable knowledge to the understanding of podocyte damage and suggest potential therapeutic targets for treating glomerular diseases like FSGS.

#### 4.3.4 Discussion

In this chapter, we developed and validated a universal Podocyte Damage Score (PDS) to quantify podocyte injury at the single-cell level across diverse models of podocyte damage. By integrating data from multiple datasets and employing both supervised and unsupervised approaches, we demonstrated that the PDS is a robust and generalizable metric that reflects the severity of podocyte damage irrespective of the underlying disease etiology.

##### **Summary of Findings**

We hypothesized that different causes of podocyte damage converge on common mechanisms of transcriptional regulation, leading to a shared pathological cellular state. To test this, we derived a universal podocyte damage signature by performing a meta-analysis of transcriptomic data from 37 studies encompassing various focal segmental glomerulosclerosis (FSGS) models. This resulted in the identification of 42 marker genes consistently associated with podocyte damage.

Using these markers, we developed the PDS and validated its robustness through extensive testing. The score remained effective even when varying the number of marker genes or randomizing portions of the input data. Cross-validation across different disease models confirmed that the PDS could distinguish between healthy and damaged podocytes in unseen datasets. Additionally, we showed that the PDS is not platform-specific and performs consistently across different transcriptomic technologies.

Functional validation demonstrated a strong correlation between the PDS and albuminuria levels, a key clinical indicator of kidney function. Morphological validation using spatial transcriptomics and proteomics data further supported the association between high PDS values and structural damage in podocytes. Importantly, we extended the applicability of the PDS to human data, confirming its relevance across species.

By applying the PDS to eight diverse models of podocyte damage profiled by single-cell RNA sequencing, we compared damage-related changes in various pathways and genes. We observed both convergent and divergent responses across models. Pathways like oxidative phosphorylation and translation elongation were consistently upregulated with increasing podocyte damage, suggesting common adaptive or maladaptive responses. Conversely, pathways such as focal adhesion exhibited model-specific behaviors, highlighting the complexity of podocyte responses to different injury mechanisms.

Gene-level analysis revealed clusters of genes that correlated with PDS across multiple studies. These clusters were enriched in specific biological functions, such as core podocyte activities, cell junction assembly, actin cytoskeleton regulation, and immune responses. This provided deeper insights into the molecular mechanisms underlying podocyte damage and FSGS progression.

Finally, we reconstructed a podocyte transcriptional regulatory network (TRN) using ATAC-seq data and identified key transcription factors (TFs) potentially driving the

transcriptional changes associated with podocyte damage. Notably, we found that TFs related to circadian regulation may play a significant role. Subsequent analysis revealed increased circadian dysregulation correlating with higher PDS values, suggesting that disruptions in circadian rhythms may contribute to the transcriptional reprogramming of damaged podocytes.

### **Broader Context and Implications**

Our findings have significant implications for understanding podocyte biology and the pathogenesis of glomerular diseases like FSGS. The development of the PDS provides a tool for quantifying podocyte damage at the single-cell level, enabling comparisons across different disease models and species. This contributes to a more unified understanding of the molecular responses to podocyte injury, which is crucial given the heterogeneity of causes leading to podocyte damage.

The consistent upregulation of pathways like oxidative phosphorylation across multiple models suggests a universal metabolic response to podocyte injury. This could reflect increased energy demands or compensatory mechanisms in damaged cells. Targeting such convergent pathways may offer therapeutic opportunities applicable to a range of glomerular diseases.

Conversely, the model-specific behaviors observed in certain pathways, such as the divergent regulation of the actin cytoskeleton and focal adhesion, underscore the importance of context in disease mechanisms. These differences highlight the need for personalized approaches when developing therapeutic interventions, as treatments effective for one type of podocyte injury may not be suitable for another.

The identification of circadian dysregulation as a potential mechanism in podocyte damage introduces a novel aspect to glomerular disease research. Circadian rhythms regulate numerous physiological processes, and their disruption has been linked to various pathologies. Our findings suggest that circadian genes may be integral to maintaining podocyte homeostasis, and their dysregulation could exacerbate or even initiate damage. This opens new avenues for research and potential therapeutic strategies targeting circadian pathways.

### **Limitations**

Because we construct the damage signature by selecting genes that consistently change across models and techniques, the damage score may be principally limited in the abilities (1) to characterize model-specific changes in the early or late stages of FSGS and (2) to detect model specific changes that drive expression in the direction opposite to the damage-score axis. We also expect that with a sufficient amount of high quality datasets, one can optimize the damage signature using unsupervised ML techniques. However, an unsupervised approach necessitates integration and normalization of the datasets, which introduces another source of bias. There are few more general limitations of our study that we should point:

**Transcriptomic Focus:** The PDS is based on gene expression data and may not capture post-transcriptional modifications, protein activity, or cellular functions that are critical in

podocyte biology. Gene expression changes do not always correlate directly with protein levels or functional outcomes.

**Marker Gene Set:** The 42-marker gene set, while effective, may not encompass all relevant genes involved in podocyte damage. As new data emerge, the damage signature may need to be updated to remain comprehensive and accurate.

**Computational Predictions:** The identification of TFs and the role of circadian dysregulation are based on computational analyses and correlations. Experimental validation is necessary to confirm these associations and understand the underlying mechanisms.

**Sample Size and Diversity:** Some analyses, particularly those involving human data, were limited by sample size. Larger and more diverse datasets would strengthen the conclusions and enhance the generalizability of the findings.

## **Future Directions**

**Experimental Validation:** Functional studies are needed to validate the roles of identified TFs, especially those related to circadian regulation, in podocyte damage. This could involve manipulating these TFs in podocyte cultures or animal models and assessing the effects on gene expression and cellular function.

**Proteomic and Epigenetic Integration:** Incorporating proteomic and epigenetic data could provide a more comprehensive understanding of podocyte damage mechanisms. This multi-omics approach would help link gene expression changes to protein function and regulatory modifications.

**Therapeutic Exploration:** The consistent pathways identified across models, such as oxidative phosphorylation, present potential targets for therapeutic intervention. Investigating compounds or treatments that modulate these pathways could lead to new strategies for preventing or mitigating podocyte damage.

**Circadian Rhythm Research:** Further exploration of circadian dysregulation in podocytes is warranted. Understanding how circadian genes influence podocyte function and how their disruption contributes to disease could reveal novel treatment approaches, possibly involving chronotherapy or circadian rhythm modulators.

**Clinical Application of PDS:** Applying the PDS to larger, more diverse human cohorts could enhance its utility as a diagnostic or prognostic tool. Correlating PDS with clinical outcomes may aid in patient stratification and personalized medicine approaches in nephrology.

**Expansion to Other Cell Types:** While our focus was on podocytes, the methodology is currently being adapted to study damage in other parenchymal cell types. Developing cell-type-specific damage scores could improve our understanding of complex organ pathologies.

## **Conclusion**

In conclusion, we successfully developed a universal Podocyte Damage Score that quantifies podocyte injury at the single-cell level across various disease models and species. Our comprehensive analysis revealed both convergent and divergent molecular responses to podocyte

damage, providing insights into the underlying mechanisms of glomerular diseases like FSGS. The identification of circadian dysregulation as a potential contributor to podocyte injury introduces a novel aspect to kidney disease research, highlighting new potential therapeutic targets. While limitations exist, our study advances the understanding of podocyte biology and offers a foundation for future research aimed at developing effective treatments for glomerular diseases.

# General Discussion

The podocyte is a specialized epithelial cell integral to the kidney's glomerular filtration barrier. Damage to podocytes is a critical event in the pathogenesis of various glomerular diseases, notably focal segmental glomerulosclerosis (FSGS). This thesis presents a comprehensive investigation into the molecular mechanisms underlying podocyte injury, focusing on transcriptional regulatory networks (TRNs) and employing both bulk and single-cell genomic approaches.

## Integrative Analysis of Podocyte Transcriptional Networks

Our studies leveraged bulk glomerular RNA sequencing (RNA-seq) and chromatin immunoprecipitation sequencing (ChIP-seq) to explore the transcriptional landscape of podocytes under pathological conditions. The findings solidified the role of **Wilms' tumor 1 (WT1)** as a master regulator of podocyte function. WT1 is crucial for podocyte differentiation and maintenance, regulating genes associated with the slit diaphragm, cytoskeletal organization, and extracellular matrix (ECM) components essential for glomerular filtration barrier integrity.

In the **Wt1** FSGS mouse model, WT1 haploinsufficiency led to a global reduction in WT1 DNA binding activity, resulting in impaired transcriptional control over key podocyte-specific genes. This disruption manifested as cytoskeletal disorganization, loss of cell adhesion, and dedifferentiation of podocytes, ultimately leading to foot process effacement and proteinuria.

Interestingly, ChIP-seq analyses revealed that **TEA domain transcription factor 1 (TEAD1)** exhibited increased DNA binding in response to WT1 deficiency. TEAD1, a component of the Hippo signaling pathway, may play a compensatory role by regulating genes involved in cell adhesion and survival. This suggests that podocytes might activate alternative transcriptional programs to mitigate the effects of WT1 loss, although the efficacy of such compensation requires further investigation.

However, the bulk RNA-seq approach presented limitations. While it provided insights into global transcriptional changes within the glomerulus, it lacked the resolution to distinguish cell-type-specific alterations, particularly subtle changes within podocytes. This limitation underscored the need for single-cell approaches to dissect the complex cellular heterogeneity of the kidney.

## Single-Cell Resolution of Podocyte Injury

To address this, we employed **single-nucleus RNA sequencing (snRNA-seq)** to analyze podocyte-specific transcriptional changes in *Wt1* and *Nphs2* mouse models of genetic FSGS. The *Nphs2* gene encodes podocin, a critical component of the slit diaphragm, and mutations in *Nphs2* are associated with steroid-resistant nephrotic syndrome in humans (Boute et al. 2000).

Our snRNA-seq analyses revealed a significant decrease in podocyte numbers in Nphs2<sup>mut</sup> mice compared to wild-type controls, correlating with elevated albuminuria levels. This decline was evident even at early disease stages, highlighting the aggressive nature of the disease caused by Nphs2 mutations. In contrast, the Wt1<sup>ko/wt</sup> mice showed a less pronounced reduction in podocyte numbers, with changes becoming more apparent at later stages.

Differential expression analyses indicated that Nphs2<sup>mut</sup> podocytes exhibited more extensive transcriptional alterations than Wt1<sup>ko/wt</sup> podocytes, aligning with the observed phenotypic severity. Both models demonstrated dysregulation in pathways critical for podocyte structure and function, such as cytoskeletal organization, cell adhesion, and stress responses. Notably, **epithelial-mesenchymal transition (EMT)** pathways were upregulated, suggesting a maladaptive response contributing to podocyte detachment and loss.

However, there were distinct differences between the models. Inflammatory pathways, such as **TNF-alpha signaling via NF-κB**, were upregulated in Nphs2<sup>mut</sup> podocytes but downregulated in Wt1<sup>ko/wt</sup> podocytes. This indicates mutation-specific inflammatory responses that may influence disease progression and severity.

The snRNA-seq approach provided a higher resolution of the transcriptional changes occurring in podocytes, overcoming the limitations of bulk RNA-seq. By isolating nuclei, we could focus on podocyte-specific gene expression without the confounding effects of other glomerular cell types.

## Development of a Universal Podocyte Damage Score

Recognizing the need for a quantitative measure of podocyte injury, we developed a **Podocyte Damage Score (PDS)**. By performing a meta-analysis of transcriptomic data from 37 studies encompassing various FSGS models, we identified a set of 42 marker genes consistently associated with podocyte damage. The PDS proved to be a robust and generalizable metric, effectively distinguishing between healthy and damaged podocytes across multiple datasets, disease models, and species.

Functional validation demonstrated a strong correlation between the PDS and clinical indicators of kidney function, such as albuminuria levels. Morphological validation using spatial transcriptomics and proteomics data further supported the association between high PDS values and structural damage in podocytes.

Applying the PDS to diverse models of podocyte damage profiled by single-cell RNA sequencing allowed us to compare damage-related changes in various pathways and genes. We observed both convergent and divergent responses across models:

- **Convergent Responses:** Pathways like **oxidative phosphorylation** and **translation elongation** were consistently upregulated with increasing podocyte damage. This suggests a universal metabolic response to injury, possibly reflecting increased energy demands or stress adaptation mechanisms.

- **Divergent Responses:** Pathways such as **focal adhesion** and **actin cytoskeleton regulation** exhibited model-specific behaviors. These differences highlight the complexity of podocyte responses to different injury mechanisms and the importance of context in disease progression.

Furthermore, we reconstructed a podocyte transcriptional regulatory network using ATAC-seq data and identified key transcription factors potentially driving the observed transcriptional changes. Notably, transcription factors related to **circadian regulation** emerged as significant, suggesting that circadian dysregulation may contribute to the transcriptional reprogramming of damaged podocytes.

## Implications for Podocyte Biology and FSGS Pathogenesis

Our findings have significant implications for understanding podocyte biology and the pathogenesis of glomerular diseases:

**Podocyte-Centric Pathogenesis:** The central role of podocyte injury in FSGS underscores the importance of maintaining podocyte health to prevent disease progression. The distinct transcriptional responses observed in different genetic models emphasize the need for personalized therapeutic approaches.

**Therapeutic Targets:** The consistent upregulation of metabolic pathways across models suggests potential targets for therapeutic intervention. Modulating them may help mitigate podocyte damage. Additionally, targeting TF compensatory mechanisms, such as TEAD1 activity, increasing levels of Wt1 or its targets, could reinforce podocyte resilience.

**Circadian Dysregulation:** The association between circadian rhythm disruptions and podocyte damage introduces a novel aspect to kidney disease research. Circadian genes regulate various physiological processes, and their dysregulation may exacerbate podocyte injury. Targeting circadian pathways could offer new therapeutic avenues.

**Limitations of Bulk Transcriptomics:** Our work highlights the limitations of bulk RNA-seq in capturing cell-type-specific changes in heterogeneous tissues like the kidney. Single-cell approaches are essential for dissecting the complex cellular interactions and identifying precise molecular targets.

## Limitations of the Study

While our research provides valuable insights, several limitations should be acknowledged:

**Transcriptomic Focus:** Our analyses are based on gene expression data and may not capture post-transcriptional modifications or protein-level changes. Integrating proteomic data could provide a more comprehensive understanding.

**Marker Gene Set:** The 42-marker gene set used for the PDS may not include all relevant genes associated with podocyte damage. As new data emerge, the marker set may need refinement.

**Technical Biases:** Single-cell sequencing technologies have inherent biases, such as varying capture efficiencies and transcriptional dropout, which could affect cell-type proportion estimations and gene expression measurements.

**Computational Predictions:** The identification of transcription factors and pathways is based on computational analyses. Experimental validation is essential to confirm their roles in podocyte injury.

## Future Research Directions

**Experimental Validation:** Functional studies to validate the roles of identified transcription factors, particularly those related to circadian regulation, in podocyte injury.

**Multi-Omics Integration:** Incorporating proteomic, epigenetic, and metabolomic data to provide a holistic view of podocyte biology and disease mechanisms.

**Therapeutic Exploration:** Investigating compounds or interventions that modulate convergent pathways identified, such as oxidative phosphorylation, to develop effective treatments.

**Studying very early stages of the damage:** Tracking gene expression changes over very early stages of disease development would facilitate understanding the temporal dynamics of podocyte injury and identify early biomarkers.

**Clinical Application of PDS:** Applying the PDS in clinical settings, especially to urinary samples, may aid in patient stratification, prognosis, and personalized medicine approaches in nephrology.

## Conclusion

In conclusion, this thesis advances the understanding of podocyte injury mechanisms in FSGS by integrating bulk and single-cell transcriptomic analyses. The characterisation of WT1 master regulator and the development of a universal Podocyte Damage Score provide valuable tools for studying podocyte biology. Our findings highlight both common and distinct molecular responses to podocyte damage, emphasizing the complexity of glomerular diseases and the need for tailored therapeutic strategies.

The discovery of circadian dysregulation's potential role in podocyte injury opens new avenues for research and underscores the intricate interplay between cellular processes and disease. Addressing the limitations of current approaches and expanding upon these findings will be crucial for developing effective therapies to preserve podocyte function and prevent chronic kidney disease progression.

# References

- Aibar, Sara, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, et al. 2017. “SCENIC: Single-Cell Regulatory Network Inference and Clustering.” *Nature Methods* 14 (11): 1083–86.
- Akchurin, Oleh, and Kimberly J. Reidy. 2015. “Genetic Causes of Proteinuria and Nephrotic Syndrome: Impact on Podocyte Pathobiology.” *Pediatric Nephrology* 30 (2): 221–33.
- Alcaraz, Wendy A., David A. Gold, Eric Raponi, Peter M. Gent, Dorothy Concepcion, and Bruce A. Hamilton. 2006. “Zfp423 Controls Proliferation and Differentiation of Neural Precursors in Cerebellar Vermis Formation.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (51): 19424–29.
- Amemiya, Haley M., Anshul Kundaje, and Alan P. Boyle. 2019. “The ENCODE Blacklist: Identification of Problematic Regions of the Genome.” *Scientific Reports* 9 (1): 9354.
- Arne Smits [cre, aut], Wolfgang Huber [aut]. 2017. *DEP*. Bioconductor. <https://doi.org/10.18129/B9.BIOC.DEP>.
- “AUCell.” n.d. Bioconductor. Accessed October 7, 2024. <http://bioconductor.org/packages/AUCell/>.
- Auerbach, Benjamin J., Garret A. FitzGerald, and Mingyao Li. 2022. “Tempo: An Unsupervised Bayesian Algorithm for Circadian Phase Inference in Single-Cell Transcriptomics.” *Nature Communications* 13 (1): 6580.
- Bailey, Timothy, Pawel Krajewski, Istvan Ladunga, Celine Lefebvre, Qunhua Li, Tao Liu, Pedro Madrigal, Cenny Taslim, and Jie Zhang. 2013. “Practical Guidelines for the Comprehensive Analysis of ChIP-Seq Data.” *PLoS Computational Biology* 9 (11): e1003326.
- Bailey, Timothy L. 1994. *Fitting a Mixture Model by Expectation Maximization to Discover Motifs in Bipolymers*.
- Bakken, Trygve E., Rebecca D. Hodge, Jeremy A. Miller, Zizhen Yao, Thuc Nghi Nguyen, Brian Aevermann, Eliza Barkan, et al. 2018. “Single-Nucleus and Single-Cell Transcriptomes Compared in Matched Cortical Cell Types.” *PloS One* 13 (12): e0209648.
- Boerries, Melanie, Florian Grahammer, Sven Eiselein, Moritz Buck, Charlotte Meyer, Markus Goedel, Wibke Bechtel, et al. 2013. “Molecular Fingerprinting of the Podocyte Reveals Novel Gene and Protein Regulatory Networks.” *Kidney International* 83 (6): 1052–64.
- Bonneau, Richard, Marc T. Facciotti, David J. Reiss, Amy K. Schmid, Min Pan, Amardeep Kaur, Vesteinn Thorsson, et al. 2007. “A Predictive Model for Transcriptional Control of Physiology in a Free Living Cell.” *Cell* 131 (7): 1354–65.
- Bonse, Jakob, Dirk Oliver Wennmann, Joachim Kremerskothen, Thomas Weide, Ulf Michgehl, Hermann Pavenstädt, and Beate Vollenbröker. 2018. “Nuclear YAP Localization as a Key Regulator of Podocyte Function.” *Cell Death & Disease* 9 (9): 850.
- Boute, N., O. Gribouval, S. Roselli, F. Benessy, H. Lee, A. Fuchshuber, K. Dahan, M. C. Gubler, P. Niaudet, and C. Antignac. 2000. “NPHS2, Encoding the Glomerular Protein Podocin, Is Mutated in Autosomal Recessive Steroid-Resistant Nephrotic Syndrome.” *Nature Genetics* 24 (4): 349–54.
- Brach, M. A., M. Kauer, and F. Herrmann. 1996. “Contribution of Transcription Factors to

- Oncogenesis.” *Cytokines and Molecular Therapy* 2 (2): 81–87.
- Bradner, James E., Denes Hnisz, and Richard A. Young. 2017. “Transcriptional Addiction in Cancer.” *Cell* 168 (4): 629–43.
- Buenrostro, Jason D., Paul G. Giresi, Lisa C. Zaba, Howard Y. Chang, and William J. Greenleaf. 2013. “Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position.” *Nature Methods* 10 (12): 1213–18.
- Burke, George W., 3rd, Alla Mitrofanova, Antonio Fontanella, Gaetano Ciancio, David Roth, Phil Ruiz, Carolyn Abitbol, Jayanthi Chandar, Sandra Merscher, and Alessia Fornoni. 2023. “The Podocyte: Glomerular Sentinel at the Crossroads of Innate and Adaptive Immunity.” *Frontiers in Immunology* 14 (July):1201619.
- Burt, Morgan A., Titilola D. Kalejaiye, Rohan Bhattacharya, Nikolaos Dimitrakakis, and Samira Musah. 2022. “Adriamycin-Induced Podocyte Injury Disrupts the YAP-TEAD1 Axis and Downregulates Cyr61 and CTGF Expression.” *ACS Chemical Biology* 17 (12): 3341–51.
- Busser, Brian W., Martha L. Bulyk, and Alan M. Michelson. 2008. “Toward a Systems-Level Understanding of Developmental Regulatory Networks.” *Current Opinion in Genetics & Development* 18 (6): 521–29.
- Butt, Linus, David Unnersjö-Jess, Martin Höhne, Aurelie Edwards, Julia Binz-Lotter, Dervla Reilly, Robert Hahnfeldt, et al. 2020. “A Molecular Mechanism Explaining Albuminuria in Kidney Disease.” *Nature Metabolism* 2 (5): 461–74.
- Caglayan, Emre, Yuxiang Liu, and Genevieve Konopka. 2022. “Ambient RNA Analysis Reveals Misinterpreted and Masked Cell Types in Brain Single-Nuclei Datasets.” *bioRxiv*. <https://doi.org/10.1101/2022.03.09.483658>.
- Camargo, Arley. 2022. “PCAtest: Testing the Statistical Significance of Principal Component Analysis in R.” *PeerJ* 10 (February):e12967.
- Campbell, Kirk N., and James A. Tumlin. 2018. “Protecting Podocytes: A Key Target for Therapy of Focal Segmental Glomerulosclerosis.” *American Journal of Nephrology* 47 Suppl 1 (Suppl. 1): 14–29.
- Campisi, Judith, and Fabrizio d’Adda di Fagagna. 2007. “Cellular Senescence: When Bad Things Happen to Good Cells.” *Nature Reviews. Molecular Cell Biology* 8 (9): 729–40.
- “Cancer Epigenetics: From Mechanism to Therapy.” 2012. *Cell* 150 (1): 12–27.
- Chen, Jianchun, Xiaoyong Wang, Qian He, Hai-Chun Yang, Agnes B. Fogo, and Raymond C. Harris. 2024. “Inhibition of Transcriptional Coactivator YAP Impairs the Expression and Function of Transcription Factor WT1 in Diabetic Podocyte Injury.” *Kidney International* 105 (6): 1200–1211.
- Chung, Jun-Jae, Leonard Goldstein, Ying-Jiun J. Chen, Jiyeon Lee, Joshua D. Webster, Merone Roose-Girma, Sharad C. Paudyal, Zora Modrusan, Anwasha Dey, and Andrey S. Shaw. 2020. “Single-Cell Transcriptome Profiling of the Kidney Glomerulus Identifies Key Cell Types and Reactions to Injury.” *Journal of the American Society of Nephrology: JASN* 31 (10): 2341–54.
- Clair, Jeremy, Hasmik Soloyan, Paolo Cravedi, Andrea Angeletti, Fadi Salem, Laith Al-Rabadi, Roger E. De Filippo, et al. 2024. “The Spatially Resolved Transcriptome Signatures of Glomeruli in Chronic Kidney Disease.” *JCI Insight* 9 (6). <https://doi.org/10.1172/jci.insight.165515>.
- Croft, David, Gavin O’Kelly, Guanming Wu, Robin Haw, Marc Gillespie, Lisa Matthews, Michael Caudy, et al. 2011. “Reactome: A Database of Reactions, Pathways and Biological

- Processes.” *Nucleic Acids Research* 39 (Database issue): D691–97.
- “decontX.” n.d. Bioconductor. Accessed October 7, 2024. <http://bioconductor.org/packages/decontX/>.
- Delrue, Charlotte, and Marijn M. Speeckaert. 2024. “Decoding Kidney Pathophysiology: Omics-Driven Approaches in Precision Medicine.” *Journal of Personalized Medicine* 14 (12): 1157.
- Denisenko, Elena, Belinda B. Guo, Matthew Jones, Rui Hou, Leanne de Kock, Timo Lassmann, Daniel Poppe, et al. 2020. “Systematic Assessment of Tissue Dissociation and Storage Biases in Single-Cell and Single-Nucleus RNA-Seq Workflows.” *Genome Biology* 21 (1): 130.
- Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. “STAR: Ultrafast Universal RNA-Seq Aligner.” *Bioinformatics* 29 (1): 15–21.
- Dong, Lihua, Stefan Pietsch, Zenglai Tan, Birgit Perner, Ralph Sierig, Dagmar Kruspe, Marco Groth, et al. 2015. “Integration of Cistromic and Transcriptomic Analyses Identifies Nphs2, Mafb, and Magi2 as Wilms’ Tumor 1 Target Genes in Podocyte Differentiation and Maintenance.” *Journal of the American Society of Nephrology: JASN* 26 (9): 2118–28.
- Drenckhahn, D., and R. P. Franke. 1988. “Ultrastructural Organization of Contractile and Cytoskeletal Proteins in Glomerular Podocytes of Chicken, Rat, and Man.” *Laboratory Investigation; a Journal of Technical Methods and Pathology* 59 (5): 673–82.
- Dressler, G. R. 1995. “Transcription Factors in Renal Development: The WT1 and Pax-2 Story.” *Seminars in Nephrology* 15 (4): 263–71.
- Eddy, Allison A., Jesús M. López-Guisa, Daryl M. Okamura, and Ikuyo Yamaguchi. 2012. “Investigating Mechanisms of Chronic Kidney Disease in Mouse Models.” *Pediatric Nephrology (Berlin, Germany)* 27 (8): 1233–47.
- Ettou, Sandrine, Youngsook L. Jung, Tomoya Miyoshi, Dhawal Jain, Ken Hiratsuka, Valerie Schumacher, Mary E. Taglienti, Ryuji Morizane, Peter J. Park, and Jordan A. Kreidberg. 2020a. “Epigenetic Transcriptional Reprogramming by WT1 Mediates a Repair Response during Podocyte Injury.” *bioRxiv*. bioRxiv. <https://doi.org/10.1101/2020.02.18.954347>.
- . 2020b. “Epigenetic Transcriptional Reprogramming by WT1 Mediates a Repair Response during Podocyte Injury.” *Science Advances* 6 (30): eabb5460.
- “Fgsea.” n.d. Bioconductor. Accessed October 7, 2024. <http://bioconductor.org/packages/fgsea/>.
- Fogo, Agnes B. 2015. “Causes and Pathogenesis of Focal Segmental Glomerulosclerosis.” *Nature Reviews. Nephrology* 11 (2): 76–87.
- Fong, Sheng, Kamil Pabis, Djakim Latumalea, Nomuundari Dugersuren, Maximilian Unfried, Nicholas Tolwinski, Brian Kennedy, and Jan Gruber. 2024. “Principal Component-Based Clinical Aging Clocks Identify Signatures of Healthy Aging and Targets for Clinical Intervention.” *Nature Aging* 4 (8): 1137–52.
- Francis, Anna, Meera N. Harhay, Albert C. M. Ong, Sri Lekha Tummalapalli, Alberto Ortiz, Agnes B. Fogo, Danilo Fliser, et al. 2024. “Chronic Kidney Disease and the Global Public Health Agenda: An International Consensus.” *Nature Reviews. Nephrology*, April. <https://doi.org/10.1038/s41581-024-00820-6>.
- Frisch, S. M., and H. Francis. 1994. “Disruption of Epithelial Cell-Matrix Interactions Induces Apoptosis.” *The Journal of Cell Biology* 124 (4): 619–26.
- Fu, Jia, Zhengzi Yi, Minchao Cai, Weijie Yuan, Weijia Zhang, Kyung Lee, and John Cijiang He. 2021. “Global Transcriptomic Changes in Glomerular Endothelial Cells in Mice with

- Podocyte Depletion and Glomerulosclerosis.” *Cell Death & Disease* 12 (7): 687.
- Fukuda, Akihiro, Yuji Sato, Hirotaka Shibata, Shouichi Fujimoto, and Roger C. Wiggins. 2024. “Urinary Podocyte Markers of Disease Activity, Therapeutic Efficacy, and Long-Term Outcomes in Acute and Chronic Kidney Diseases.” *Clinical and Experimental Nephrology* 28 (6): 496–504.
- Gaidatzis, Dimos, Lukas Burger, Maria Florescu, and Michael B. Stadler. 2015. “Analysis of Intronic and Exonic Reads in RNA-Seq Data Characterizes Transcriptional and Post-Transcriptional Regulation.” *Nature Biotechnology* 33 (7): 722–29.
- GBD Chronic Kidney Disease Collaboration. 2020. “Global, Regional, and National Burden of Chronic Kidney Disease, 1990-2017: A Systematic Analysis for the Global Burden of Disease Study 2017.” *The Lancet* 395 (10225): 709–33.
- Ghazala, Ahmed M., and \*. Taghreed U. Mohammb. n.d. “The Role of Metalloendopeptidase (MEP) as a Vital Predictor of Early Diabetic Nephropathy and Its Relationship to Some Other Biochemical Variables.” Accessed August 2, 2022. [http://www.echemcom.com/article\\_139840\\_af2d1783871608dcc97e77703292a08d.pdf](http://www.echemcom.com/article_139840_af2d1783871608dcc97e77703292a08d.pdf).
- “GitHub - gangwug/MetaCycle: An Integrated R Package to Evaluate Periodicity in Large Scale Data.” n.d. GitHub. Accessed October 7, 2024. <https://github.com/gangwug/MetaCycle>.
- “GitHub - jsh58/Genrich: Detecting Sites of Genomic Enrichment.” n.d. GitHub. Accessed October 7, 2024. <https://github.com/jsh58/Genrich>.
- “GitHub - ropensci/GLMMcosinor: An R Package for Flexible Cosinor Modelling Using the glmmTMB Framework.” n.d. GitHub. Accessed October 7, 2024. <https://github.com/ropensci/GLMMcosinor>.
- Green, Laura M., Kate J. Wagner, Hayley A. Campbell, Kelly Addison, and Stefan G. E. Roberts. 2009. “Dynamic Interaction between WT1 and BASP1 in Transcriptional Regulation during Differentiation.” *Nucleic Acids Research* 37 (2): 431–40.
- Grindberg, Rashed V., Joyclyn L. Yee-Greenbaum, Michael J. McConnell, Mark Novotny, Andy L. O’Shaughnessy, Georgina M. Lambert, Marcos J. Araúzo-Bravo, et al. 2013. “RNA-Sequencing from Single Nuclei.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (49): 19802–7.
- Guo, Jian-Kan, Aswin L. Menke, Marie-Claire Gubler, Alan R. Clarke, David Harrison, Annette Hammes, Nicholas D. Hastie, and Andreas Schedl. 2002. “WT1 Is a Key Regulator of Podocyte Function: Reduced Expression Levels Cause Crescentic Glomerulonephritis and Mesangial Sclerosis.” *Human Molecular Genetics* 11 (6): 651–59.
- Hall, Gentzon, Rasheed A. Gbadegesin, Peter Lavin, Guanghong Wu, Yangfan Liu, Edwin C. Oh, Liming Wang, et al. 2015. “A Novel Missense Mutation of Wilms’ Tumor 1 Causes Autosomal Dominant FSGS.” *Journal of the American Society of Nephrology: JASN* 26 (4): 831–43.
- Hao, Yuhan, Stephanie Hao, Erica Andersen-Nissen, William M. Mauck 3rd, Shiwei Zheng, Andrew Butler, Maddie J. Lee, et al. 2021. “Integrated Analysis of Multimodal Single-Cell Data.” *Cell* 184 (13): 3573–87.e29.
- Haraldsson, Börje, Jenny Nyström, and William M. Deen. 2008. “Properties of the Glomerular Barrier and Mechanisms of Proteinuria.” *Physiological Reviews* 88 (2): 451–87.
- Hatje, Favian A., Uta Wedekind, Wiebke Sachs, Desiree Loreth, Julia Reichelt, Fatih Demir, Christopher Kosub, et al. 2021. “Tripartite Separation of Glomerular Cell Types and Proteomes from Reporter-Free Mice.” *Journal of the American Society of Nephrology: JASN* 32 (9): 2175–93.

- Hayashi, Kaori, Hiroyuki Sasamura, Mari Nakamura, Tatsuhiko Azegami, Hideyo Oguchi, Yusuke Sakamaki, and Hiroshi Itoh. 2014. "KLF4-Dependent Epigenetic Remodeling Modulates Podocyte Phenotypes and Attenuates Proteinuria." *The Journal of Clinical Investigation* 124 (6): 2523–37.
- He, Lei, Yixian Fan, Yue Zhang, Tongtao Tu, Quan Zhang, Fahu Yuan, and Chao Cheng. 2022. "Single-Cell Transcriptomic Analysis Reveals Circadian Rhythm Disruption Associated with Poor Prognosis and Drug-Resistance in Lung Adenocarcinoma." *Journal of Pineal Research* 73 (1): e12803.
- Hercz, Daniel. 2013. *Flexible Modeling with Generalized Additive Models and Generalized Linear Mixed Models: Comprehensive Simulation and Case Studies*.
- Hoefield, Richard A., Philip A. Kalra, Patricia G. Baker, Ines Sousa, Peter J. Diggle, Martin J. Gibson, Donal J. O'Donoghue, Rachel J. Middleton, and John P. New. 2011. "The Use of eGFR and ACR to Predict Decline in Renal Function in People with Diabetes." *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association* 26 (3): 887–92.
- Hollstein, Monica, David Sidransky, Bert Vogelstein, and Curtis C. Harris. 1991. "p53 Mutations in Human Cancers." *Science*, July. <https://doi.org/10.1126/science.1905840>.
- Huber, Tobias B., and Thomas Benzing. 2005. "The Slit Diaphragm: A Signaling Platform to Regulate Podocyte Function." *Current Opinion in Nephrology and Hypertension* 14 (3): 211–16.
- Ichimura, Koichiro, Hidetake Kurihara, and Tatsuo Sakai. 2003. "Actin Filament Organization of Foot Processes in Rat Podocytes." *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society*, December. <https://doi.org/10.1177/002215540305101203>.
- Issa, Wadih, Rachel Njeim, Arianna Carrasco, George W. Burke, and Alla Mitrofanova. 2024. "Role of the Innate Immune Response in Glomerular Disease Pathogenesis: Focus on Podocytes." *Cells* 13 (13). <https://doi.org/10.3390/cells13131157>.
- Jackson, Christopher A., Dayanne M. Castro, Giuseppe-Antonio Saldi, Richard Bonneau, and David Gresham. 2020. "Gene Regulatory Network Reconstruction Using Single-Cell RNA Sequencing of Barcoded Genotypes in Diverse Environments." *eLife* 9 (January). <https://doi.org/10.7554/eLife.51254>.
- Jalili, Vahid, Matteo Matteucci, Marco Masseroli, and Marco J. Morelli. 2015. "Using Combined Evidence from Replicates to Evaluate ChIP-Seq Peaks." *Bioinformatics* 31 (17): 2761–69.
- Johnson, David S., Ali Mortazavi, Richard M. Myers, and Barbara Wold. 2007. "Genome-Wide Mapping of in Vivo Protein-DNA Interactions." *Science* 316 (5830): 1497–1502.
- Kanehisa, M., and S. Goto. 2000. "KEGG: Kyoto Encyclopedia of Genes and Genomes." *Nucleic Acids Research* 28 (1): 27–30.
- Kann, Martin, Sandrine Ettou, Youngsook L. Jung, Maximilian O. Lenz, Mary E. Taglienti, Peter J. Park, Bernhard Schermer, Thomas Benzing, and Jordan A. Kreidberg. 2015. "Genome-Wide Analysis of Wilms' Tumor 1-Controlled Gene Expression in Podocytes Reveals Key Regulatory Mechanisms." *Journal of the American Society of Nephrology: JASN* 26 (9): 2097–2104.
- Kaplan, J. M., S. H. Kim, K. N. North, H. Rennke, L. A. Correia, H. Q. Tong, B. J. Mathis, et al. 2000. "Mutations in ACTN4, Encoding Alpha-Actinin-4, Cause Familial Focal Segmental Glomerulosclerosis." *Nature Genetics* 24 (3): 251–56.
- Kato, Hideki, and Katalin Susztak. 2012. "Repair Problems in Podocytes: Wnt, Notch, and

- Glomerulosclerosis.” *Seminars in Nephrology* 32 (4): 350–56.
- Kato, Mamoru, Naoya Hata, Nilanjana Banerjee, Bruce Futcher, and Michael Q. Zhang. 2004. “Identifying Combinatorial Regulation of Transcription Factors and Binding Motifs.” *Genome Biology* 5 (8): 1–13.
- Kerjaschki, D. 2001. “Caught Flat-Footed: Podocyte Damage and the Molecular Bases of Focal Glomerulosclerosis.” *The Journal of Clinical Investigation* 108 (11): 1583–87.
- Kirita, Yuhei, Haojia Wu, Kohei Uchimura, Parker C. Wilson, and Benjamin D. Humphreys. 2020. “Cell Profiling of Mouse Acute Kidney Injury Reveals Conserved Cellular Responses to Injury.” *Proceedings of the National Academy of Sciences of the United States of America* 117 (27): 15874–83.
- Kojima, Kenichiro, and Donscho Kerjaschki. 2002. “Is Podocyte Shape Controlled by the Dystroglycan Complex?” *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association* 17 Suppl 9:23–24.
- Konda, Prathyusha, Simon Garinet, Eliezer M. Van Allen, and Srinivas R. Viswanathan. 2023. “Genome-Guided Discovery of Cancer Therapeutic Targets.” *Cell Reports* 42 (8): 112978.
- Kovesdy, Csaba P. 2022. “Epidemiology of Chronic Kidney Disease: An Update 2022.” *Kidney International. Supplement* 12 (1): 7–11.
- Kriz, W., N. Gretz, and K. V. Lemley. 1998. “Progression of Glomerular Diseases: Is the Podocyte the Culprit?” *Kidney International* 54 (3): 687–97.
- Kuppe, Christoph, Mahmoud M. Ibrahim, Jennifer Kranz, Xiaoting Zhang, Susanne Ziegler, Javier Perales-Patón, Jitske Jansen, et al. 2021. “Decoding Myofibroblast Origins in Human Kidney Fibrosis.” *Nature* 589 (7841): 281–86.
- Kuppe, Christoph, Katja Leuchtle, Anton Wagner, Nazanin Kabgani, Turgay Saritas, Victor G. Puelles, Bart Smeets, et al. 2019. “Novel Parietal Epithelial Cell Subpopulations Contribute to Focal Segmental Glomerulosclerosis and Glomerular Tip Lesions.” *Kidney International* 96 (1): 80–93.
- Lake, Blue B., Rajasree Menon, Seth Winfree, Qiwen Hu, Ricardo Melo Ferreira, Kian Kalhor, Daria Barwinska, et al. 2023. “An Atlas of Healthy and Injured Cell States and Niches in the Human Kidney.” *Nature* 619 (7970): 585–94.
- Latt, Khun Zaw, Jurgen Heymann, Joseph H. Jessee, Avi Z. Rosenberg, Celine C. Berthier, Arnon Arazi, Sean Eddy, et al. 2022. “Urine Single-Cell RNA Sequencing in Focal Segmental Glomerulosclerosis Reveals Inflammatory Signatures.” *Kidney International Reports* 7 (2): 289–304.
- Lee, Tong Ihn, Nicola J. Rinaldi, François Robert, Duncan T. Odom, Ziv Bar-Joseph, Georg K. Gerber, Nancy M. Hannett, et al. 2002. “Transcriptional Regulatory Networks in *Saccharomyces Cerevisiae*.” *Science* 298 (5594): 799–804.
- Lee, Tong Ihn, and Richard A. Young. 2013. “Transcriptional Regulation and Its Misregulation in Disease.” *Cell* 152 (6): 1237–51.
- Lefevre, Gaelle M., Sanjeevkumar R. Patel, Doyeob Kim, Lino Tessarollo, and Gregory R. Dressler. 2010. “Altering a Histone H3K4 Methylation Pathway in Glomerular Podocytes Promotes a Chronic Disease Phenotype.” *PLoS Genetics* 6 (10): e1001142.
- Liberzon, Arthur, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P. Mesirov, and Pablo Tamayo. 2015. “The Molecular Signatures Database (MSigDB) Hallmark Gene Set Collection.” *Cell Systems* 1 (6): 417–25.
- Li, Heng, and Richard Durbin. 2009. “Fast and Accurate Short Read Alignment with

- Burrows-Wheeler Transform.” *Bioinformatics* 25 (14): 1754–60.
- Lipska, Beata S., Bruno Ranchin, Paraskevas Iatropoulos, Jutta Gellermann, Anette Melk, Fatih Ozaltin, Gianluca Caridi, et al. 2014. “Genotype-Phenotype Associations in WT1 Glomerulopathy.” *Kidney International* 85 (5): 1169–78.
- Liu, Min, Zhe Qiao, Yang Zhang, Ping Zhan, and Fan Yi. 2020. “Histone Deacetylases Take Center Stage on Regulation of Podocyte Function.” *Kidney Diseases (Basel, Switzerland)* 6 (4): 236–46.
- López-Otín, Carlos, Maria A. Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. 2013. “The Hallmarks of Aging.” *Cell* 153 (6): 1194–1217.
- Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550.
- Macneil, Lesley T., and Albertha J. M. Walhout. 2011. “Gene Regulatory Networks and the Role of Robustness and Stochasticity in the Control of Gene Expression.” *Genome Research* 21 (5): 645–57.
- Maezawa, Yoshiro, Tuncer Onay, Rizaldy P. Scott, Lindsay S. Keir, Henrik Dimke, Chengjin Li, Vera Eremina, et al. 2014. “Loss of the Podocyte-Expressed Transcription Factor Tcf21/Pod1 Results in Podocyte Differentiation Defects and FSGS.” *Journal of the American Society of Nephrology: JASN* 25 (11): 2459–70.
- Malaga-Dieguez, Laura, Diana Bouhassira, Debbie Gipson, and Howard Trachtman. 2015. “Novel Therapies for FSGS: Preclinical and Clinical Studies.” *Advances in Chronic Kidney Disease* 22 (2): e1–6.
- Marshall, Jamie L., Teia Noel, Qingbo S. Wang, Haiqi Chen, Evan Murray, Ayshwarya Subramanian, Katherine A. Vernon, et al. 2022. “High-Resolution Slide-seqV2 Spatial Transcriptomics Enables Discovery of Disease-Specific Cell Neighborhoods and Pathways.” *iScience* 25 (4): 104097.
- Menke, Aswin L., Annemieke IJpenberg, Stewart Fleming, Allyson Ross, Claire N. Medine, Charles E. Patek, Lee Spraggon, Jeremy Hughes, Alan R. Clarke, and Nicholas D. Hastie. 2003. “The wt1-Heterozygous Mouse; a Model to Study the Development of Glomerular Sclerosis.” *The Journal of Pathology* 200 (5): 667–74.
- Miner, Jeffrey H., Roy Morello, Kaya L. Andrews, Cong Li, Corinne Antignac, Andrey S. Shaw, and Brendan Lee. 2002. “Transcriptional Induction of Slit Diaphragm Genes by Lmx1b Is Required in Podocyte Differentiation.” *The Journal of Clinical Investigation* 109 (8): 1065–72.
- Moriguchi, Takashi, Michito Hamada, Naoki Morito, Tsumoru Terunuma, Kazuteru Hasegawa, Chuan Zhang, Tomomasa Yokomizo, et al. 2006. “MafB Is Essential for Renal Development and F4/80 Expression in Macrophages.” *Molecular and Cellular Biology* 26 (15): 5715–27.
- Morrison, Avril A., Rebecca L. Viney, Moin A. Saleem, and Michael R. Lodomery. 2008. “New Insights into the Function of the Wilms Tumor Suppressor Gene WT1 in Podocytes.” *American Journal of Physiology. Renal Physiology* 295 (1): F12–17.
- Mundel, Peter, and Jochen Reiser. 2010. “Proteinuria: An Enzymatic Disease of the Podocyte?” *Kidney International* 77 (7): 571–80.
- Mundel, Peter, and Stuart J. Shankland. 2002. “Podocyte Biology and Response to Injury.” *Journal of the American Society of Nephrology: JASN* 13 (12): 3005–15.
- “MYC on the Path to Cancer.” 2012. *Cell* 149 (1): 22–35.
- “Oligo.” n.d. Bioconductor. Accessed October 7, 2024. <http://bioconductor.org/packages/oligo/>.

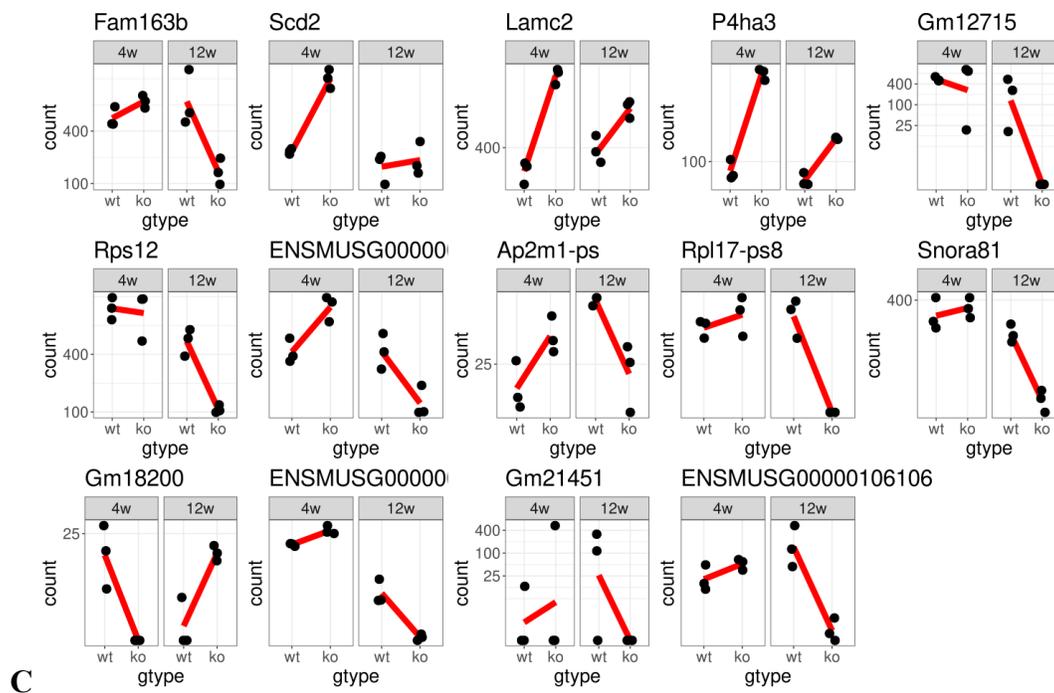
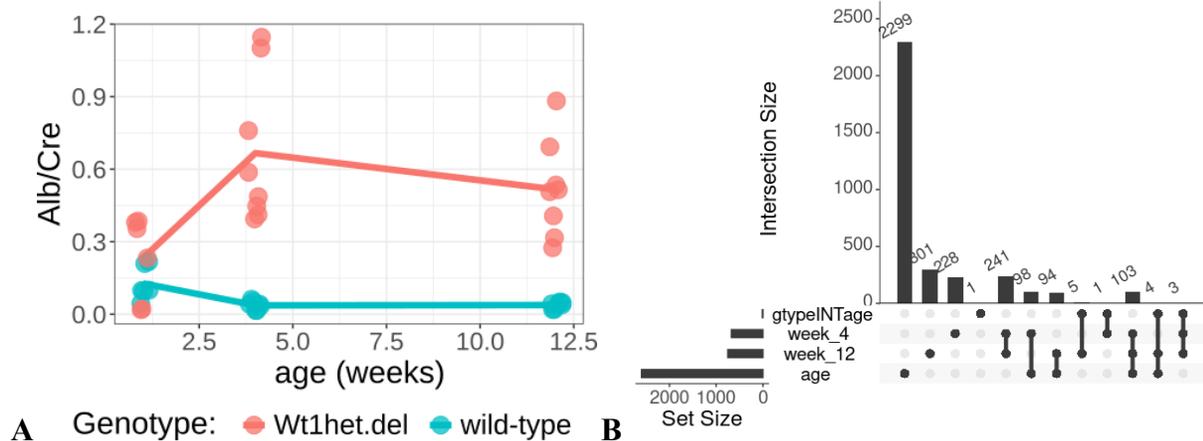
- Palmer, R. E., A. Kotsianti, B. Cadman, T. Boyd, W. Gerald, and D. A. Haber. 2001. "WT1 Regulates the Expression of the Major Glomerular Podocyte Membrane Protein Podocalyxin." *Current Biology: CB* 11 (22): 1805–9.
- Park, Jihwan, Rojesh Shrestha, Chengxiang Qiu, Ayano Kondo, Shizheng Huang, Max Werth, Mingyao Li, Jonathan Barasch, and Katalin Suszták. 2018. "Single-Cell Transcriptomics of the Mouse Kidney Reveals Potential Cellular Targets of Kidney Disease." *Science* 360 (6390): 758–63.
- Park, Peter J. 2009. "ChIP-Seq: Advantages and Challenges of a Maturing Technology." *Nature Reviews. Genetics* 10 (10): 669–80.
- Pavenstädt, Hermann, Wilhelm Kriz, and Matthias Kretzler. 2003. "Cell Biology of the Glomerular Podocyte." *Physiological Reviews* 83 (1): 253–307.
- Pawson, T. 1995. "Protein Modules and Signalling Networks." *Nature* 373 (6515): 573–80.
- Pearl, Jocelynn R., Carlo Colantuoni, Dani E. Bergey, Cory C. Funk, Paul Shannon, Bijoya Basu, Alex M. Casella, et al. 2019. "Genome-Scale Transcriptional Regulatory Network Models of Psychiatric and Neurodegenerative Disorders." *Cell Systems* 8 (2): 122–35.e7.
- Pique-Regi, Roger, Jacob F. Degner, Athma A. Pai, Daniel J. Gaffney, Yoav Gilad, and Jonathan K. Pritchard. 2011. "Accurate Inference of Transcription Factor Binding from DNA Sequence and Chromatin Accessibility Data." *Genome Research* 21 (3): 447–55.
- Qian, Jiang, Jimmy Lin, Nicholas M. Luscombe, Haiyuan Yu, and Mark Gerstein. 2003. "Prediction of Regulatory Networks: Genome-Wide Identification of Transcription Factor Targets from Gene Expression Data." *Bioinformatics (Oxford, England)* 19 (15): 1917–26.
- Quaggin, S. E., L. Schwartz, S. Cui, P. Igarashi, J. Deimling, M. Post, and J. Rossant. 1999. "The Basic-Helix-Loop-Helix Protein pod1 Is Critically Important for Kidney and Lung Organogenesis." *Development* 126 (24): 5771–83.
- Quaggin, S. E., G. B. Vanden Heuvel, and P. Igarashi. 1998. "Pod-1, a Mesoderm-Specific Basic-Helix-Loop-Helix Protein Expressed in Mesenchymal and Glomerular Epithelial Cells in the Developing Kidney." *Mechanisms of Development* 71 (1-2): 37–48.
- Quaggin, Susan E. 2002. "Transcriptional Regulation of Podocyte Specification and Differentiation." *Microscopy Research and Technique* 57 (4): 208–11.
- Rahmatollahi, Mahdieh. 2020. *Transcriptional Regulation in Podocytes*.
- Rasclé, Anne, Hani Suleiman, Tanja Neumann, and Ralph Witzgall. 2007. "Role of Transcription Factors in Podocytes." *Nephron. Experimental Nephrology* 106 (2): e60–66.
- Reidy, Kimberly, and Frederick J. Kaskel. 2007. "Pathophysiology of Focal Segmental Glomerulosclerosis." *Pediatric Nephrology* 22 (3): 350–54.
- Reily, Colin, Tyler J. Stewart, Matthew B. Renfrow, and Jan Novak. 2019. "Glycosylation in Health and Disease." *Nature Reviews. Nephrology* 15 (6): 346–66.
- Reiser, Jochen, and Sanja Sever. 2013. "Podocyte Biology and Pathogenesis of Kidney Disease." *Annual Review of Medicine* 64:357–66.
- Risso, Davide, John Ngai, Terence P. Speed, and Sandrine Dudoit. 2014. "Normalization of RNA-Seq Data Using Factor Analysis of Control Genes or Samples." *Nature Biotechnology* 32 (9): 896–902.
- Rohr, Claudia, Jürgen Prestel, Laurence Heidet, Hiltraud Hosser, Wilhelm Kriz, Randy L. Johnson, Corinne Antignac, and Ralph Witzgall. 2002. "The LIM-Homeodomain Transcription Factor Lmx1b Plays a Crucial Role in Podocytes." *The Journal of Clinical Investigation* 109 (8): 1073–82.
- Rory Stark<rori.stark@cruk.cam.ac.uk>, Gord Brown<gdbzork@gmail.com>. 2017.

- DiffBind*. Bioconductor. <https://doi.org/10.18129/B9.BIOC.DIFFBIND>.
- Rosenberg, Avi Z., and Jeffrey B. Kopp. 2017. "Focal Segmental Glomerulosclerosis." *Clinical Journal of the American Society of Nephrology: CJASN* 12 (3): 502–17.
- Sadl, Virginia, Fuzi Jin, Joanna Yu, Shiyong Cui, Douglas Holmyard, Susan Quaggin, Greg Barsh, and Sabine Cordes. 2002. "The Mouse Kreisler (Krm11/MafB) Segmentation Gene Is Required for Differentiation of Glomerular Visceral Epithelial Cells." *Developmental Biology* 249 (1): 16–29.
- Saoncella, S., F. Echtermeyer, F. Denhez, J. K. Nowlen, D. F. Mosher, S. D. Robinson, R. O. Hynes, and P. F. Goetinck. 1999. "Syndecan-4 Signals Cooperatively with Integrins in a Rho-Dependent Manner in the Assembly of Focal Adhesions and Actin Stress Fibers." *Proceedings of the National Academy of Sciences of the United States of America* 96 (6): 2805–10.
- Sato, Utako, Sachiko Kitanaka, Takashi Sekine, Shori Takahashi, Akira Ashida, and Takashi Igarashi. 2005. "Functional Characterization of LMX1B Mutations Associated with Nail-Patella Syndrome." *Pediatric Research* 57 (6): 783–88.
- "scDbfFinder." n.d. Bioconductor. Accessed October 7, 2024. <http://bioconductor.org/packages/scDbfFinder/>.
- Schell, Christoph, Manuel Rogg, Martina Suhm, Martin Helmstädter, Dominik Sellung, Mako Yasuda-Yamahara, Oliver Kretz, et al. 2017. "The FERM Protein EPB41L5 Regulates Actomyosin Contractility and Focal Adhesion Formation to Maintain the Kidney Filtration Barrier." *Proceedings of the National Academy of Sciences of the United States of America* 114 (23): E4621–30.
- Schmid, Holger, Anna Henger, Clemens D. Cohen, Karin Frach, Hermann-Josef Gröne, Detlef Schlöndorff, and Matthias Kretzler. 2003. "Gene Expression Profiles of Podocyte-Associated Molecules as Diagnostic Markers in Acquired Proteinuric Diseases." *Journal of the American Society of Nephrology: JASN* 14 (11): 2958–66.
- Schwarz, K., M. Simons, J. Reiser, M. A. Saleem, C. Faul, W. Kriz, A. S. Shaw, L. B. Holzman, and P. Mundel. 2001. "Podocin, a Raft-Associated Component of the Glomerular Slit Diaphragm, Interacts with CD2AP and Nephrin." *The Journal of Clinical Investigation* 108 (11): 1621–29.
- Sikora-Wohlfeld, Weronika, Marit Ackermann, Eleni G. Christodoulou, Kalaimathy Singaravelu, and Andreas Beyer. 2013. "Assessing Computational Methods for Transcription Factor Target Gene Identification Based on ChIP-Seq Data." *PLoS Computational Biology* 9 (11): e1003342.
- "Slingshot." n.d. Bioconductor. Accessed October 7, 2024. <http://bioconductor.org/packages/slingshot/>.
- Stuart, Joshua M., Eran Segal, Daphne Koller, and Stuart K. Kim. 2003. "A Gene-Coexpression Network for Global Discovery of Conserved Genetic Modules." *Science (New York, N.Y.)* 302 (5643): 249–55.
- Sun, Kun, Huating Wang, and Hao Sun. 2017. "mTFkb: A Knowledgebase for Fundamental Annotation of Mouse Transcription Factors." *Scientific Reports* 7 (1): 3022.
- Tan, Kai, Jesper Tegner, and Timothy Ravasi. 2008. "Integrated Approaches to Uncovering Transcription Regulatory Networks in Mammalian Cells." *Genomics* 91 (3): 219–31.
- Trachtman, Howard. 2020. "Emerging Drugs for Treatment of Focal Segmental Glomerulosclerosis." *Expert Opinion on Emerging Drugs* 25 (3): 367–75.
- Tryggvason, Karl, Jaakko Patrakka, and Jorma Wartiovaara. 2006. "Hereditary Proteinuria

- Syndromes and Mechanisms of Proteinuria.” *The New England Journal of Medicine* 354 (13): 1387–1401.
- Tryggvason, K., and J. Wartiovaara. 2001. “Molecular Basis of Glomerular Permeability.” *Current Opinion in Nephrology and Hypertension* 10 (4): 543–49.
- Tveita, Anders, Ole Petter Rekvig, and Svetlana N. Zykova. 2008. “Glomerular Matrix Metalloproteinases and Their Regulators in the Pathogenesis of Lupus Nephritis.” *Arthritis Research & Therapy* 10 (6): 1–8.
- Unger Avila, Paula, Tsimafei Padvitski, Ana Carolina Leote, He Chen, Julio Saez-Rodriguez, Martin Kann, and Andreas Beyer. 2024. “Gene Regulatory Networks in Disease and Ageing.” *Nature Reviews. Nephrology* 20 (9): 616–33.
- “Using Genetic and Species Diversity to Tackle Kidney Disease.” 2020. *Trends in Genetics* 36 (7): 499–509.
- Vassiliadis, John, Christina Bracken, Douglas Matthews, Stephen O’Brien, Susan Schiavi, and Stefan Wawersik. 2011. “Calcium Mediates Glomerular Filtration through Calcineurin and mTORC2/Akt Signaling.” *Journal of the American Society of Nephrology: JASN* 22 (8): 1453–61.
- Walhout, Albertha J. M. 2006. “Unraveling Transcription Regulatory Networks by Protein-DNA and Protein-Protein Interaction Mapping.” *Genome Research* 16 (12): 1445–54.
- Wang, Dan, Chunsun Dai, Yingjian Li, and Youhua Liu. 2011. “Canonical Wnt/ $\beta$ -Catenin Signaling Mediates Transforming Growth Factor- $\beta$ 1-Driven Podocyte Injury and Proteinuria.” *Kidney International* 80 (11): 1159–69.
- Wang, Lulu, Han Tian, Haiyan Wang, Xiaoming Mao, Jing Luo, Qingyun He, Ping Wen, et al. 2024. “Disrupting Circadian Control of Autophagy Induces Podocyte Injury and Proteinuria.” *Kidney International* 105 (5): 1020–34.
- “Website.” n.d.-a. R Core Team (2022) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org>.
- . n.d.-b. <https://www.kpmp.org>.
- . n.d.-c. <https://github.com/nevelsk90/thesis.2024>.
- White, Jeffrey T., Bo Zhang, Débora M. Cerqueira, Uyen Tran, and Oliver Wessely. 2010. “Notch Signaling, *wt1* and *foxc2* Are Key Regulators of the Podocyte Gene Regulatory Network in *Xenopus*.” *Development* 137 (11): 1863–73.
- Wiggins, R. C. 2007. “The Spectrum of Podocytopathies: A Unifying View of Glomerular Diseases.” *Kidney International* 71 (12): 1205–14.
- Winn, Michelle P., Peter J. Conlon, Kelvin L. Lynn, Merry Kay Farrington, Tony Creazzo, April F. Hawkins, Nikki Daskalakis, et al. 2005. “A Mutation in the TRPC6 Cation Channel Causes Familial Focal Segmental Glomerulosclerosis.” *Science* 308 (5729): 1801–4.
- Wong, Germaine, Amélie Bernier-Jean, Brad Rovin, Pierre Ronco, and Editors of *Kidney International*. 2024. “Time for Action: Recognizing Chronic Kidney Disease as a Major Noncommunicable Disease Driver of Premature Mortality.” *Kidney International* 105 (6): 1144–46.
- Wu, Haojia, Yuhei Kirita, Erinn L. Donnelly, and Benjamin D. Humphreys. 2019. “Advantages of Single-Nucleus over Single-Cell RNA Sequencing of Adult Kidney: Rare Cell Types and Novel Cell States Revealed in Fibrosis.” *Journal of the American Society of Nephrology: JASN* 30 (1): 23–32.
- Xie, Kewei, Chenqi Xu, Minfang Zhang, Minzhou Wang, Lulin Min, Cheng Qian, Qin Wang, et al. 2019. “Yes-Associated Protein Regulates Podocyte Cell Cycle Re-Entry and

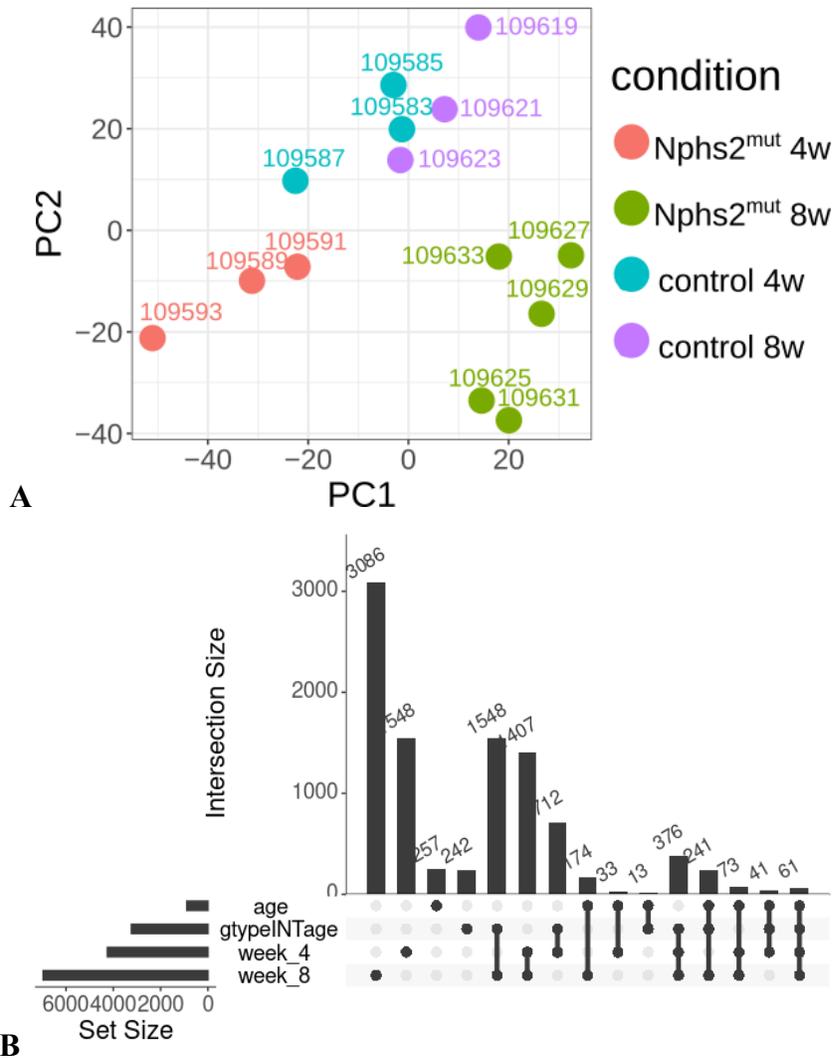
- Dedifferentiation in Adriamycin-Induced Nephropathy.” *Cell Death & Disease* 10 (12): 915.
- Yan, Feng, David R. Powell, David J. Curtis, and Nicholas C. Wong. 2020. “From Reads to Insight: A Hitchhiker’s Guide to ATAC-Seq Data Analysis.” *Genome Biology* 21 (1): 22.
- Yang, Jingping, Difei Zhang, Masaru Motojima, Tsutomu Kume, Qing Hou, Yu Pan, Aiping Duan, et al. 2021. “Super-Enhancer-Associated Transcription Factors Maintain Transcriptional Regulation in Mature Podocytes.” *Journal of the American Society of Nephrology: JASN* 32 (6): 1323–37.
- Yu, Guangchuang. 2018. *Enrichplot*. Bioconductor.  
<https://doi.org/10.18129/B9.BIOC.ENRICHPLOT>.
- Yu, Guangchuang, Li-Gen Wang, Yanyan Han, and Qing-Yu He. 2012. “clusterProfiler: An R Package for Comparing Biological Themes among Gene Clusters.” *Omics: A Journal of Integrative Biology* 16 (5): 284–87.
- Zaret, Kenneth S. 2020. “Pioneer Transcription Factors Initiating Gene Network Changes.” *Annual Review of Genetics* 54 (November):367–85.
- Zeng, Fenghua, Lance A. Kloefer, Charlene Finney, André Diedrich, and Raymond C. Harris. 2016. “Specific Endothelial Heparin-Binding EGF-like Growth Factor Deletion Ameliorates Renal Injury Induced by Chronic Angiotensin II Infusion.” *American Journal of Physiology. Renal Physiology* 311 (4): F695–707.

# Appendix A: Supplementary Figures

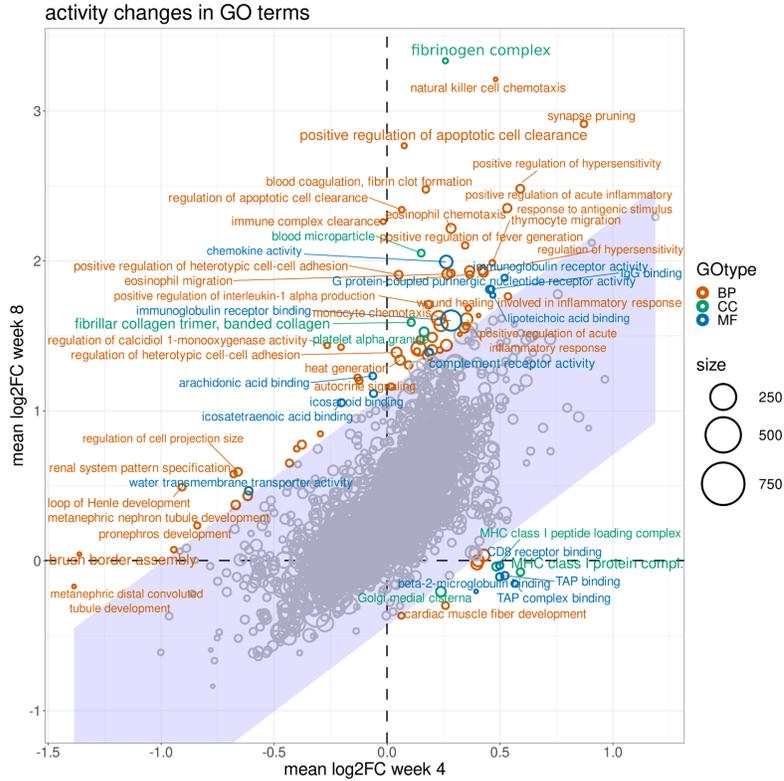


Supplementary Figure 3.1: Effect of FSGS progression in Wt1 ko/wt mice.

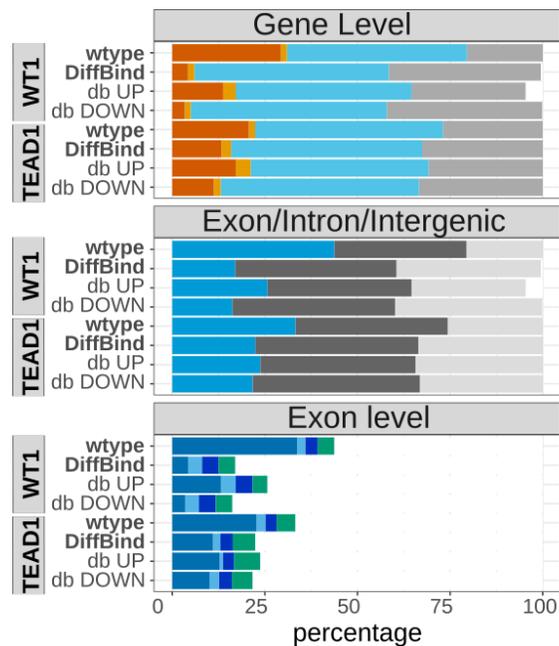
(A) Albumin/Creatinine ratio measures kidney function in healthy (turquoise) and Wt1ko/wt (red) mice at the age of 1, 4 (early FSGS) and 12 (late FSGS) weeks. Each dot represents a biological sample, samples were not used for sequencing. (B) Comparison of sets of differentially expressed genes (in rows), where columns show number of unique or shared genes. (C) Expression of genes that have significantly different Wt1<sup>ko/wt</sup> effect at 4 and 12 weeks, compared to controls. Counts are sample-wise normalized by the library size.



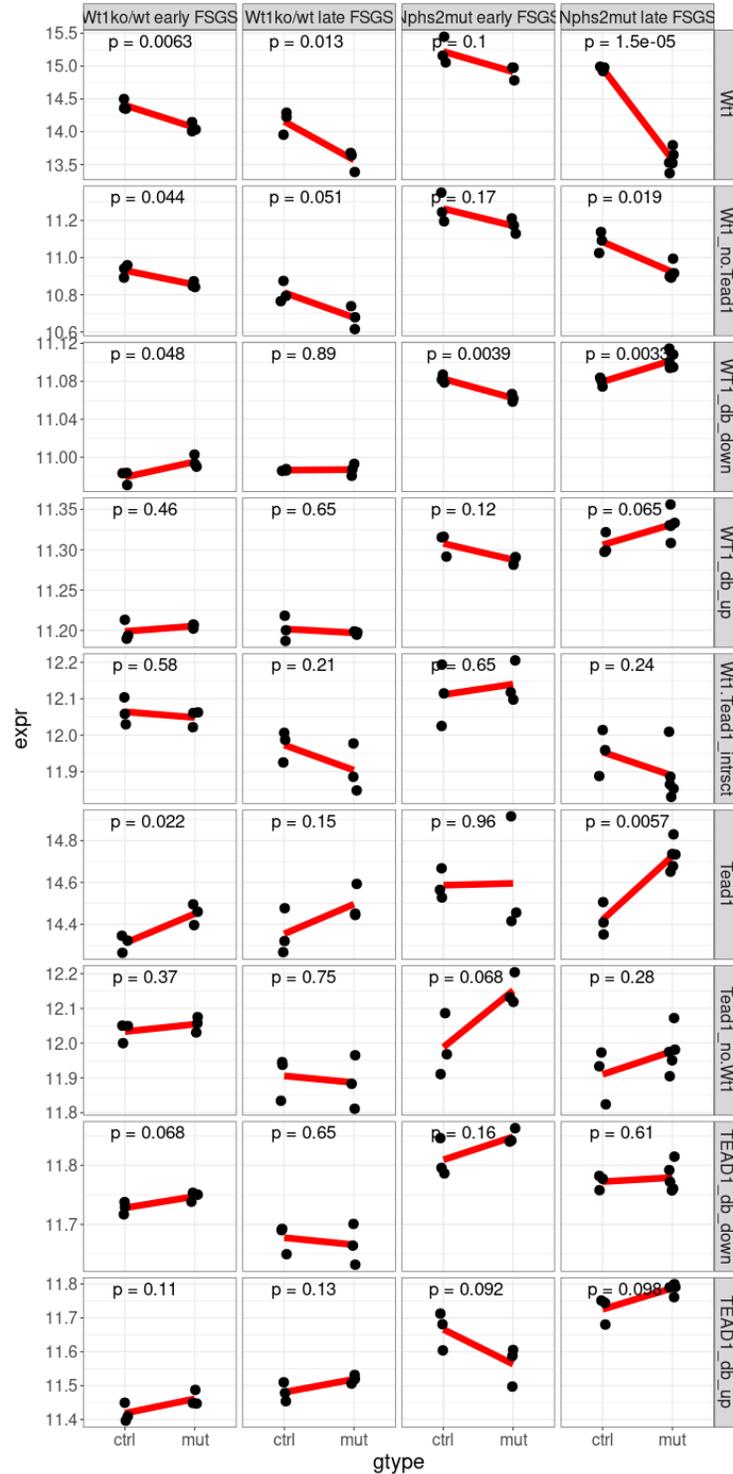
Supplementary Figure 3.2.: Analysis of bulk RNAseq data from glomeruli of Nphs2 mutant mice. (A) PCA plot of Nphs2<sup>mut</sup> bulk RNAseq samples, counts were log transformed, top 30% of variable genes were used. (B) Overlap of genes differentially expressed (FDR adjusted p-value < 0.05) in 4 comparisons: effect of the genotype (Nphs2mut) at 4 weeks, effect of the genotype at 8 weeks, effect of the age, and interaction between the age and the genotype.



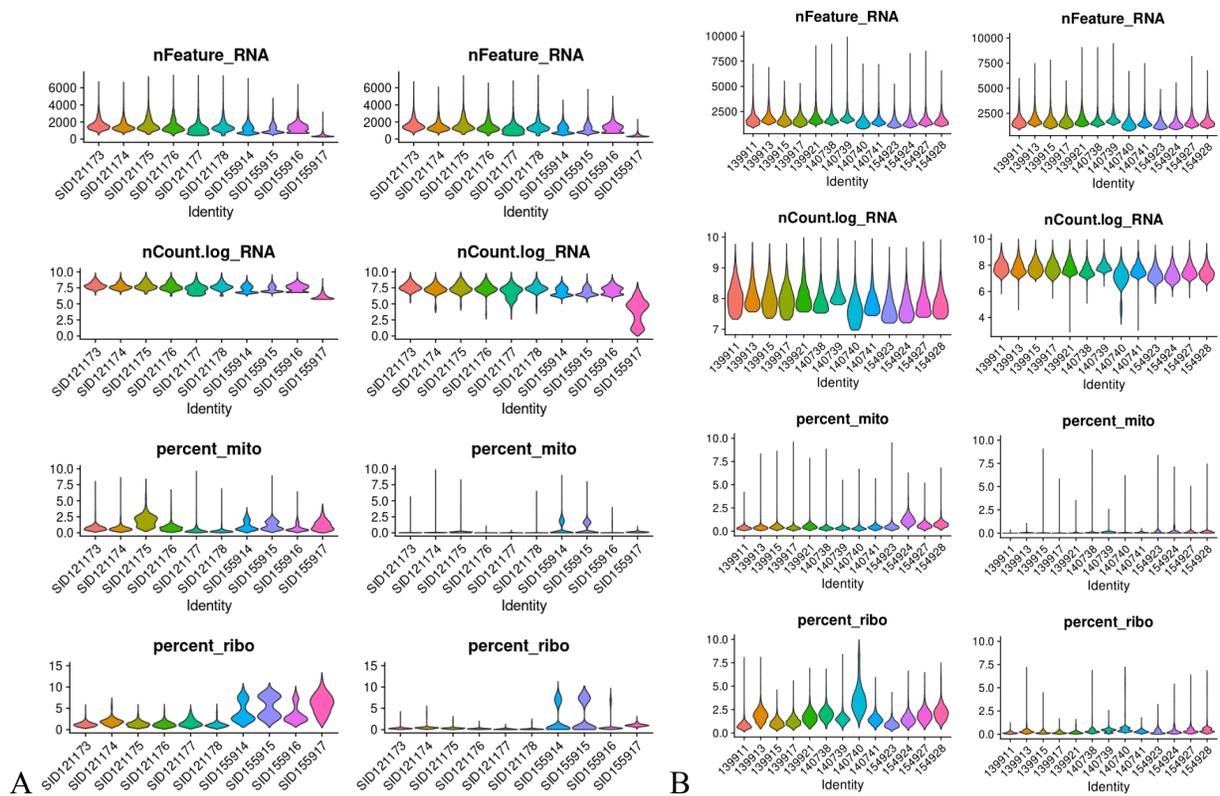
Supplementary Figure 3.3: 2D GO plot shows functional comparison of early and late FSGS in Nphs2<sup>mut</sup> model. Size of a circle reflects the number of genes in the term, color of a circle and a label reflects type of ontology. GO annotation was performed using the Roberts function to address the redundancy issue.



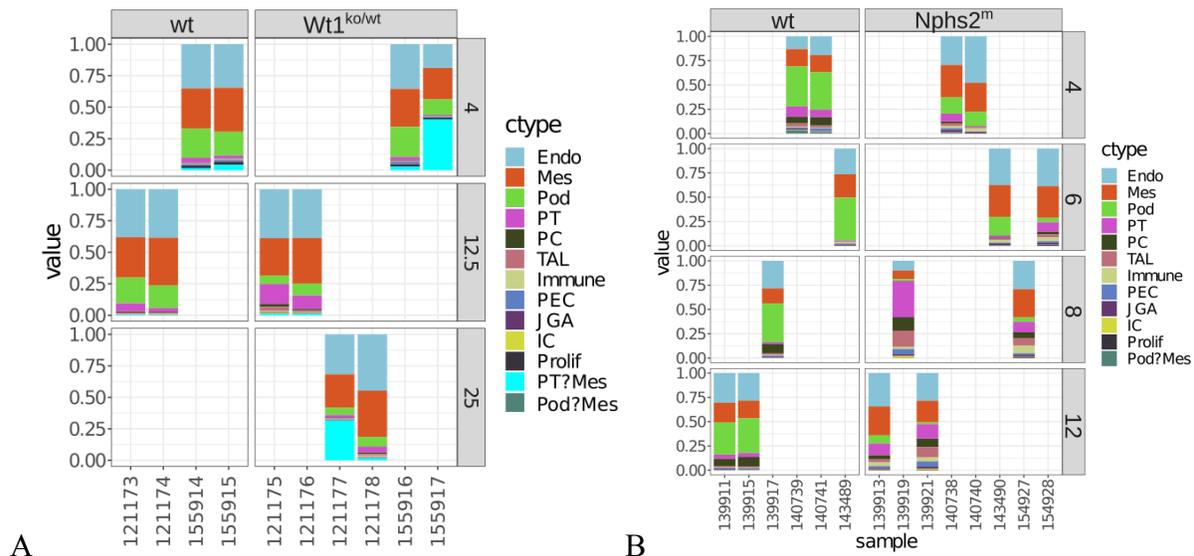
Supplementary Figure 3.4: Genome region annotation of wild-type and differentially bound peaks.



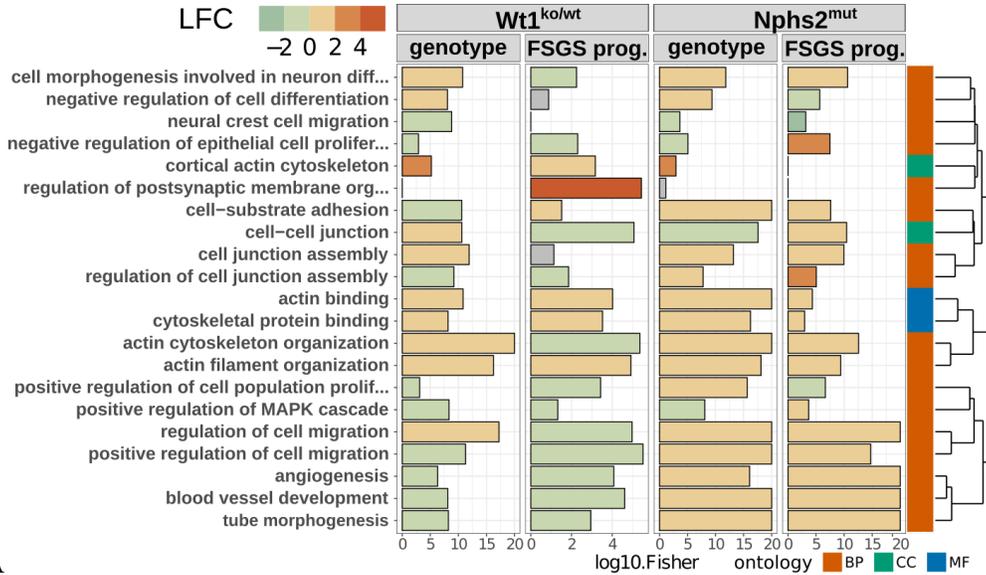
Supplementary Figure 3.5: Expression of WT1 and TEAD1 transcription factors and their target genes in glomerular bulk RNA-seq. Y-axis shows rlog-normalised mRNA expression levels of transcription factors (TF mRNA) or mean expression of unique and common TF target genes, mean expression of targets of up and down-bound regions (db t.genes up, db t.genes down), as was defined based on Chip-seq. X-axis shows sample genotype.



Supplementary Figure 4.1: Various QC plots of count matrices of (A)  $Wt1^{ko/wt}$  and (B)  $Nphs2^{mut}$  snRNA-seq samples (rows). Left and right columns show results before and after ambient RNA removal.

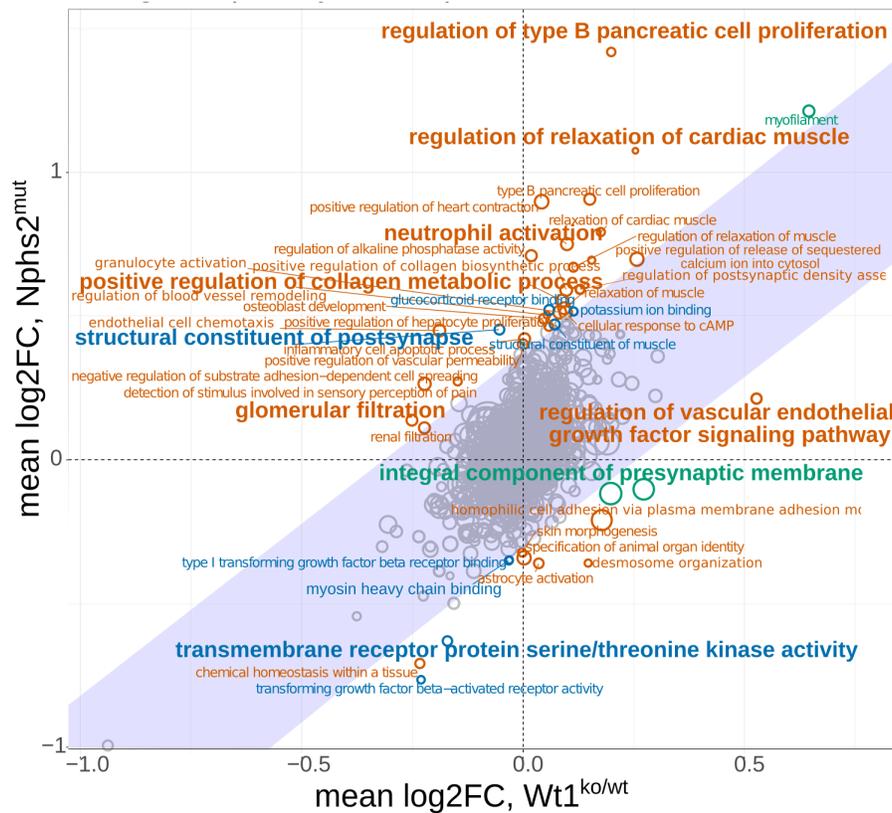


Supplementary Figure 4.2: Cell-type composition of individual samples from  $Wt1^{ko/wt}$  (C) and  $Nphs2^{mut}$  (B) datasets, correspondingly. Color encodes cell-types, X and Y axes show sample IDs and cell-type proportion. Columns and rows of panels separate genotype and age of samples, respectively.

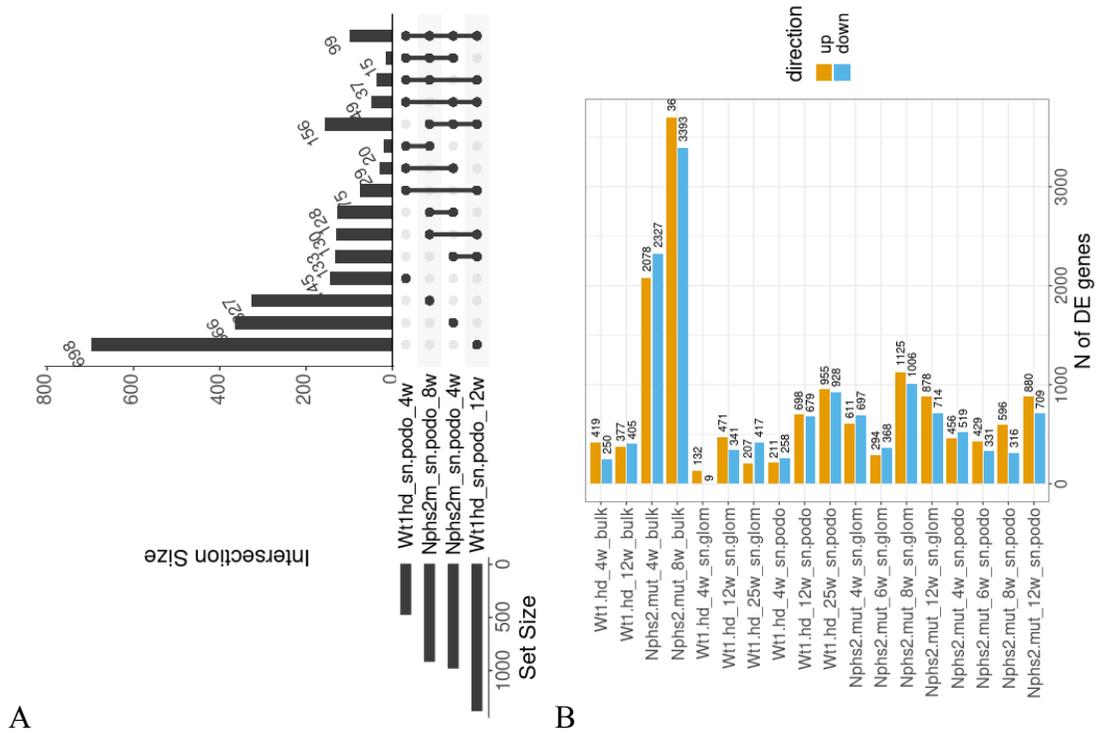


A

Supplementary Figure 4.3: GO annotation of genes associated with general effect of genotype and disease progression (interaction of time and genotype) effect in Nphs2<sup>mut</sup> and Wt1<sup>ko/wt</sup> experiments. Color shows mean LFC of all genes of the term.

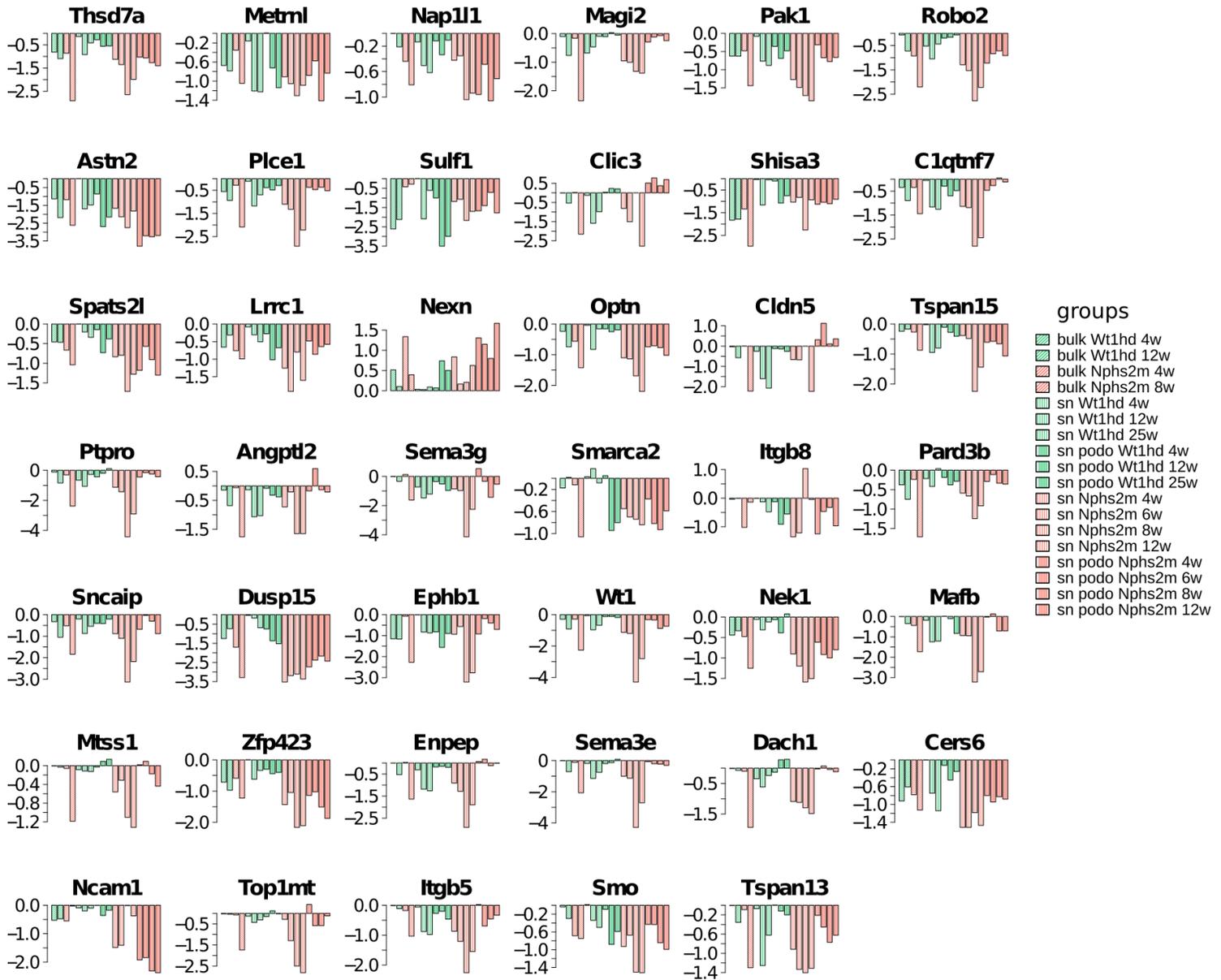


Supplementary Figure 4.4: 2D GO plots of genotype effects in podocytes of Wt1<sup>ko/wt</sup> (X-axis) and Nphs2<sup>mut</sup> (Y-axis). Change of activity in GO term is calculated as the mean LFC of all genes included in the term.

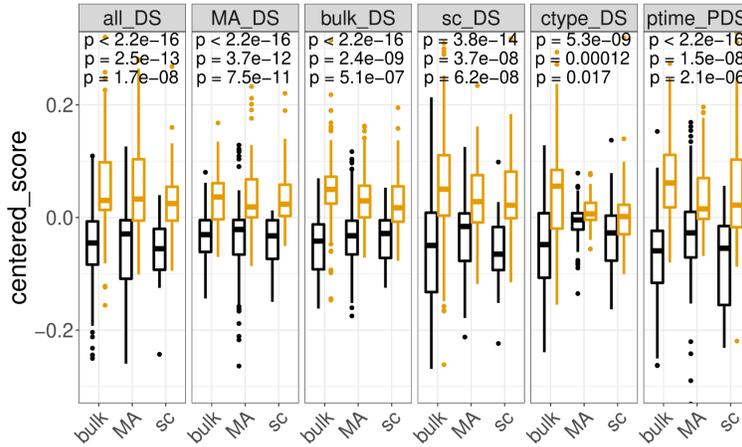


Supplementary Figure 4.5: Differential expression analysis of bulk and snRNA-seq data. of Wt1 and Nphs2mut models of genetic FSGS.

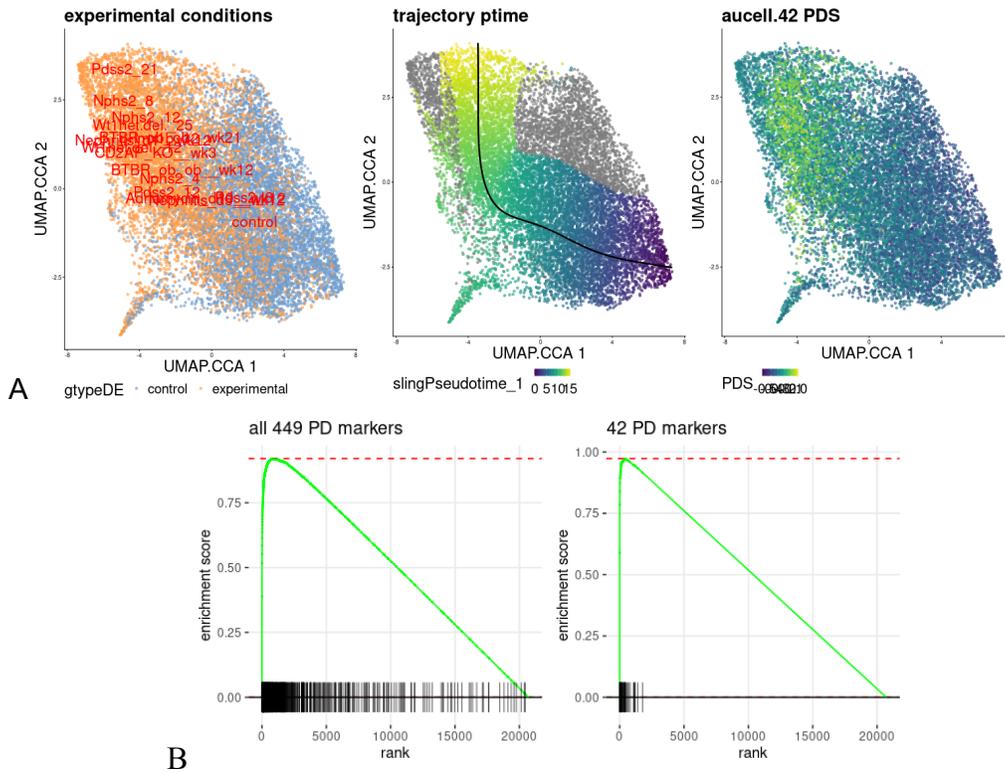
(A) Overlap of genes differentially expressed (FDR adjusted p-value < 0.05) between mutant and control podocytes in snRNA-seq samples from 4 (early FSGS) and 8/12 (late FSGS) week old mice, from Wt1 and Nphs2 models of FSGS. (B) Numbers of significantly (q-value < 0.05) up and down-regulated genes across age specific test of genotype effect in bulk and sn RNA-seq samples.



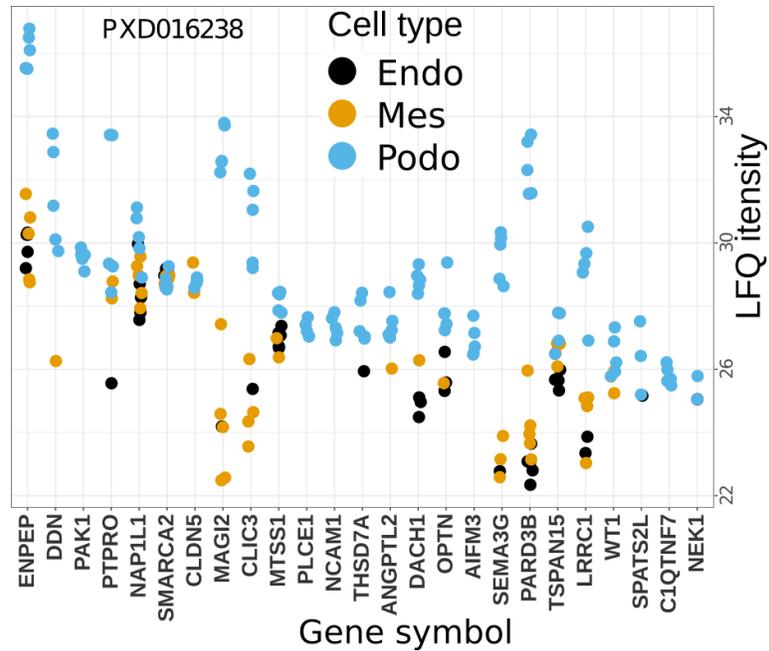
Supplementary Figure 4.6: DE expression between control and experimental samples of 42 podocyte damage markers in bulk and snRNAseq datasets from Wt1ko/wt (green) and Nphs2mut (red) datasets. Y-axis shows log fold-change of genes, obtained with lfcShrink() from results of DEseq2 analysis, x-axis shows groups of samples.



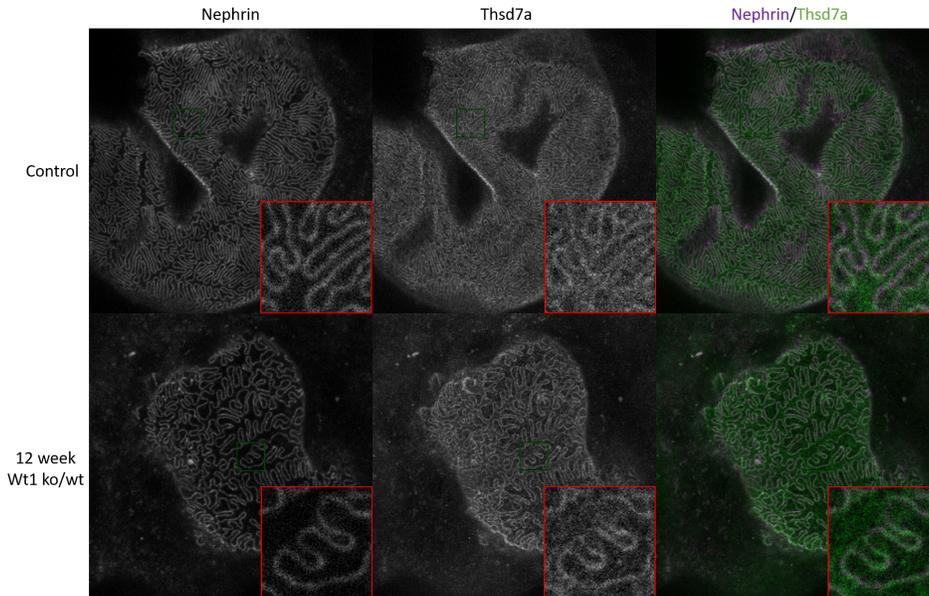
Supplemental Figure 4.7: Damage score robustness tests: cross-validation of various transcriptome abundance measurement platforms.



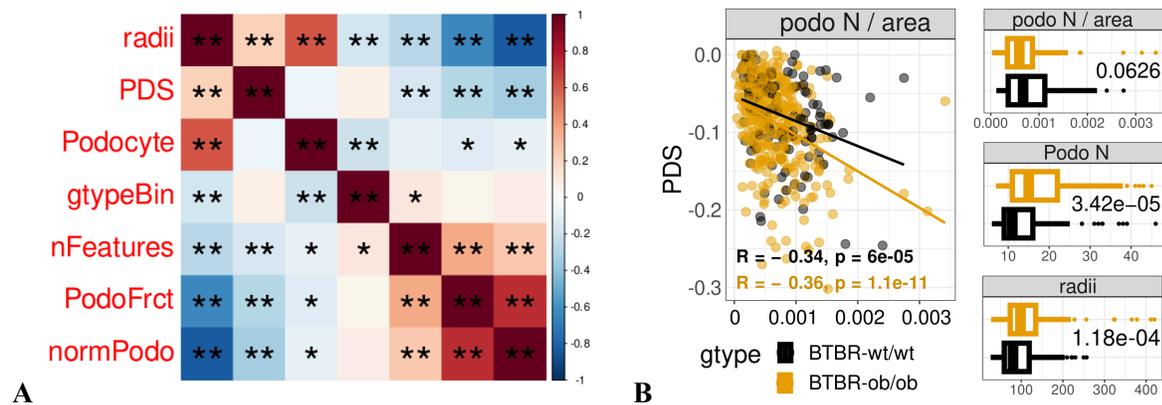
Supplemental Figure 4.8: Comparison of PDS with an unsupervised sling pseudo-time trajectory. (A) UMAPs showing application of PDS and pseudotime to scRNA-seq datasets. (B) GSEA, testing enrichment genes differentially expressed along the pseudotime in podocyte damage markers.



Supplemental Figure 4.9: Expression (Y-axis) of podocyte damage markers (X-axis) on protein level in 3 types of mouse glomerular cells (color-coded). Non-scaled protein expression values (LFQ) plotted for each sample (dot). Data taken from PXD016238 (Hatje et al. 2021).

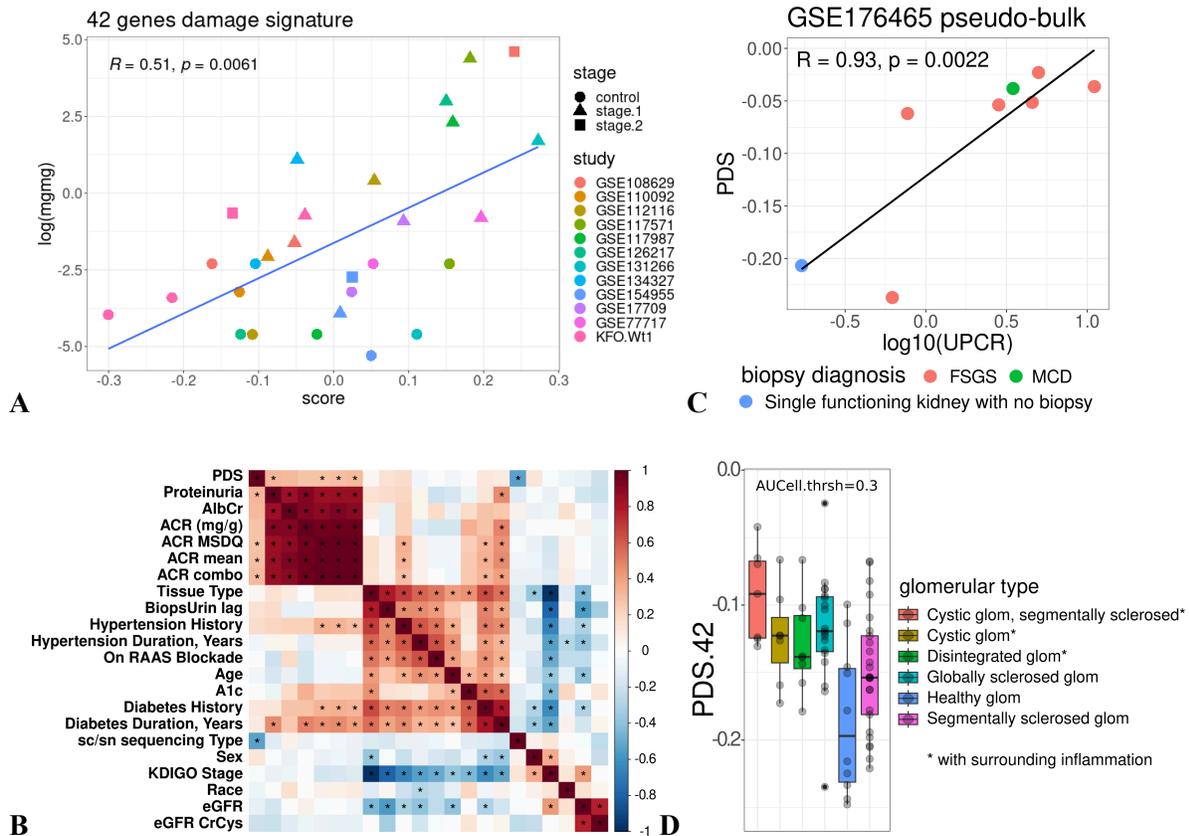


Supplementary Figure 4.10: STED images of nephrin and Thsd7a co-localization at the podocyte slit diaphragm in control and 12-week Wt1 het.del animals. The localization of Thsd7a is similar to that of nephrin.



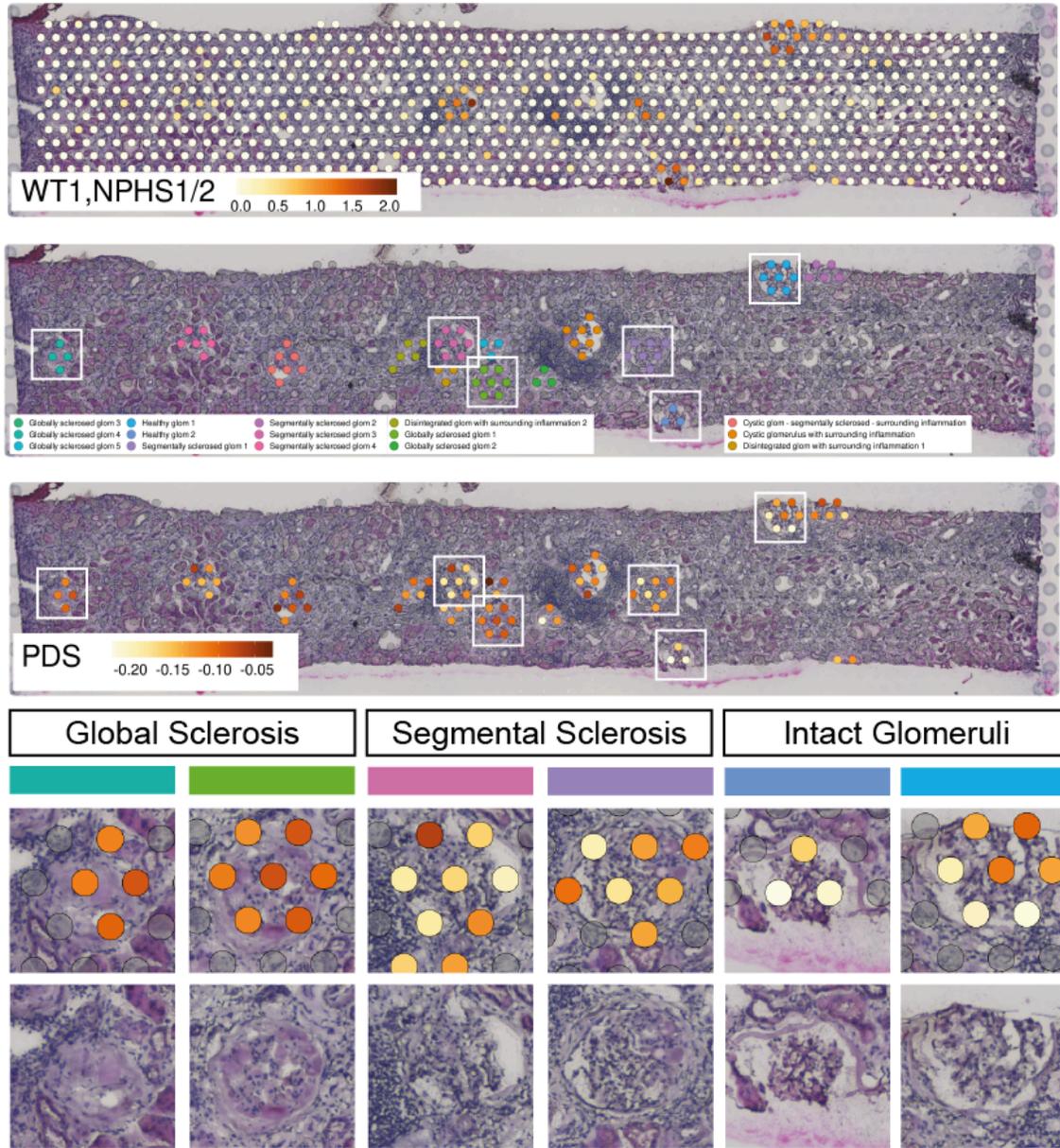
Supplementary Figure 4.11: Spatial validation of PDS.

(A) Correlation heatmap of PDS and glomeruli metrics estimated from Slide-seqV2 spatial transcriptomics mouse dataset GSE190094 (PMID: 35372810). Podocyte stands for podocyte spot count, nFeatures - average number of features detected in podocyte spots, PodoFrct - fraction of podocyte spots to other glomerular cell-type spots, normPodo - podocyte count adjusted by the glom area. (B) Scatterplot shows correlation between the PDS and glomerular morphology in control (black) and diabetic (orange) mice, based on the same dataset as in the previous panel.

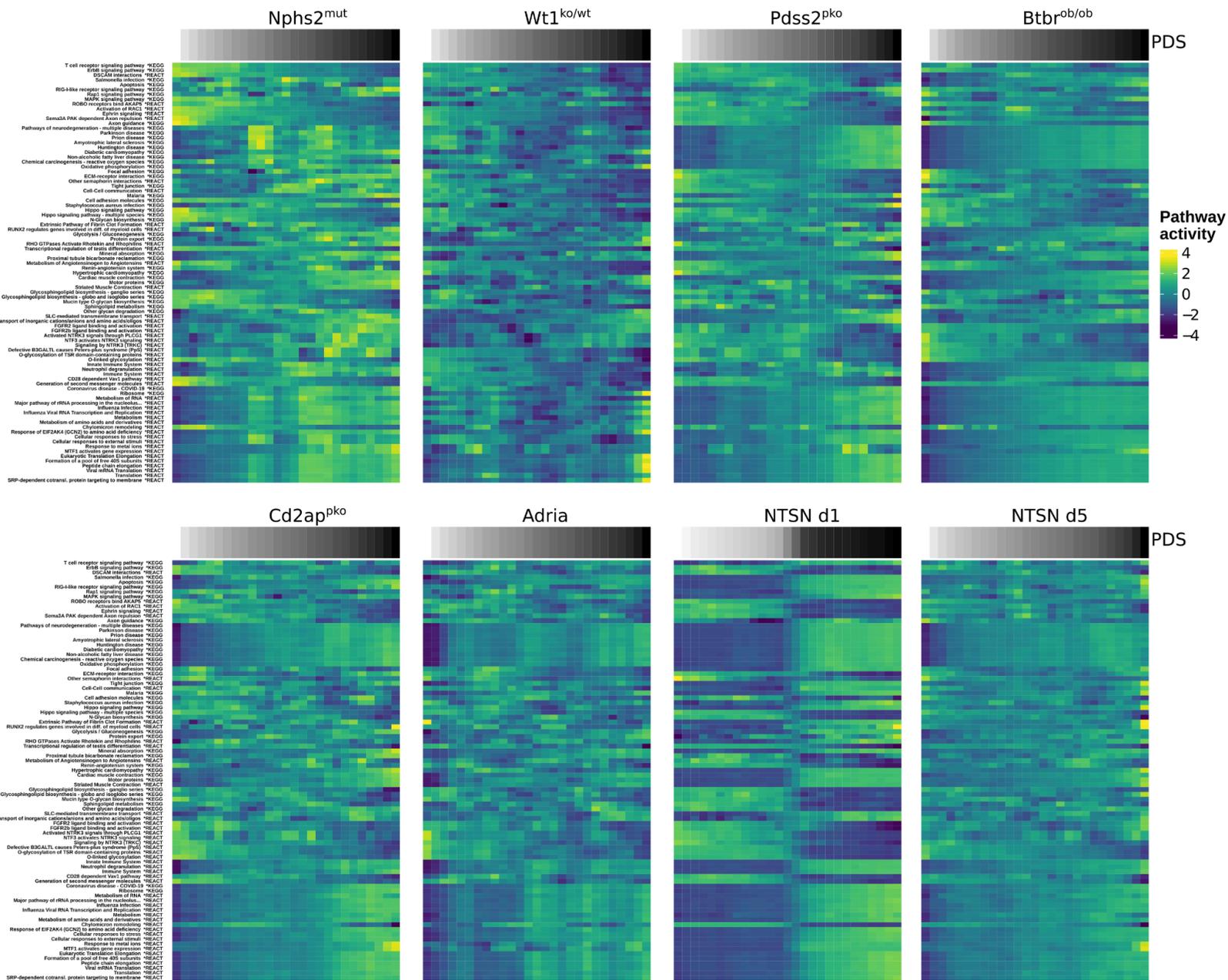


Supplemental Figure 4.12: Functional validation of PDS in mouse and human data.

(A) Correlation of proteinuria and PDS in bulk microarray and bulkRNAseq data. (B) Heatmap of spearman correlation between PDS and clinical traits from patients with acute kidney injury (AKI) and CKD. Anonymised patient data, sc and sn RNA-seq data used in the analysis are taken from KPMP. Stars denote cells with p-value of spearman correlation  $< 0.05$ . (C) Relation between PDS and a clinical trait called Urinary Protein Excretion Estimation (UPCR) in patients from GSE176465. (D) Boxplot shows PDS in different types of glomeruli, based on KPMP spatial transcriptomics dataset, patient ID 29-10282.

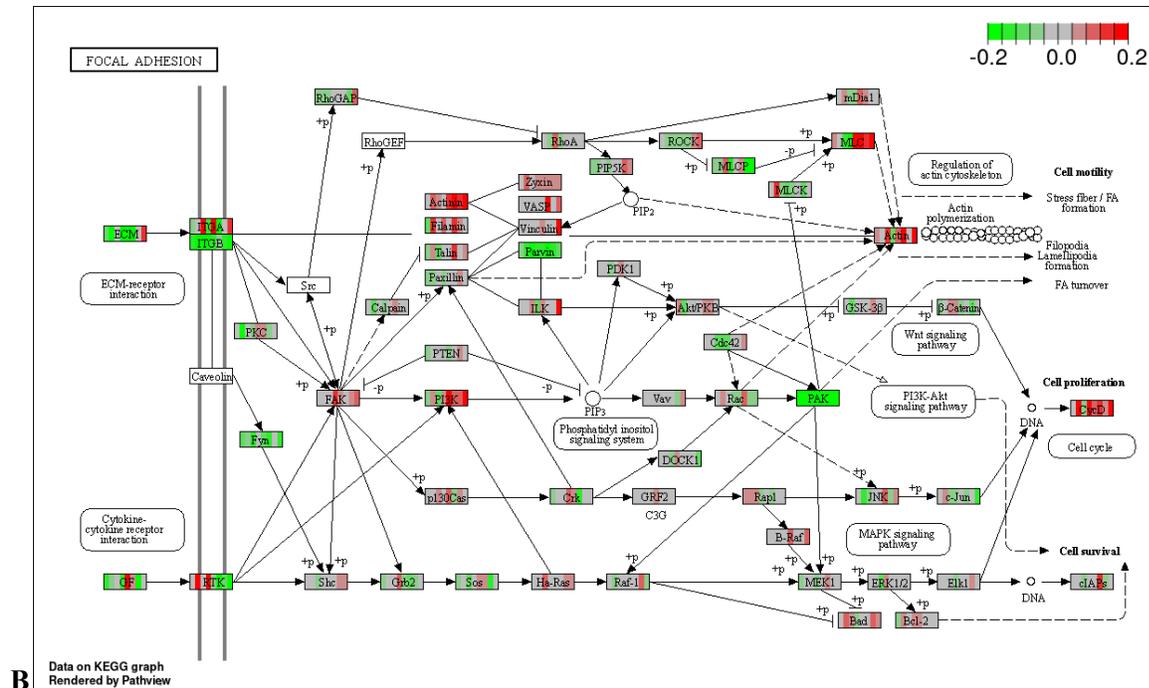
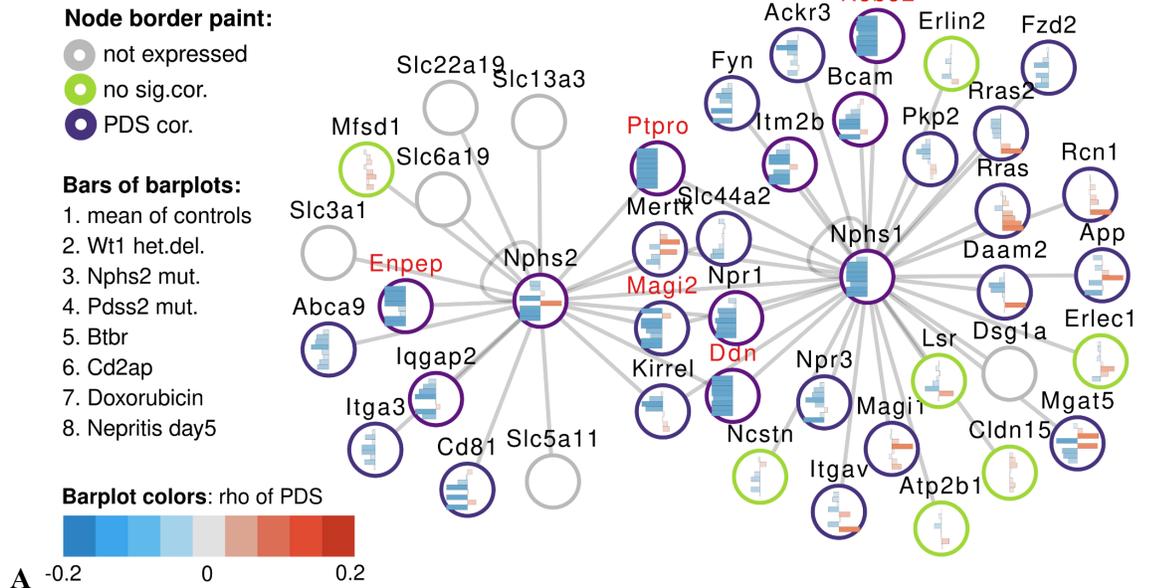


Supplementary Figure 4.13: Morphological validation of PDS in Human data. PDS gradient in Individual human glomeruli of a human kidney sample. KPMP spatial transcriptomics dataset, patient 29-10282.

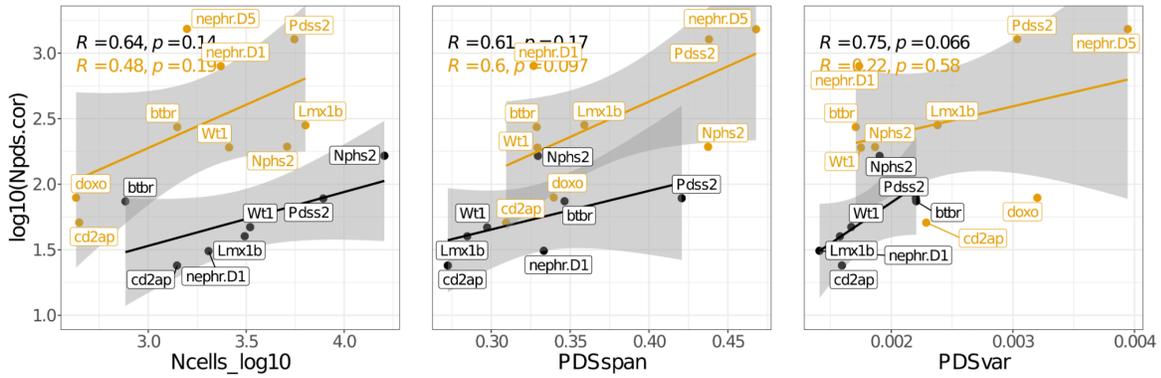


Supplemental Figure 4.14: Pathway fingerprints of the 8 models of FSGS. Rows are pathways, columns are PDS bins and the heatmap color shows normalized pathway activity score. Pathway activities are calculated with AUCell, cells are ordered by PDS and moving average window is used to smooth the signal. Columns of the heatmap are bins of the averaged pathway activity scores, ordered from low (left) to high (right) PDS. Union of 5 most correlated with PDS pathways (KEGG and Reactome) across all studies is used. Column annotation shows smoothed PDS.

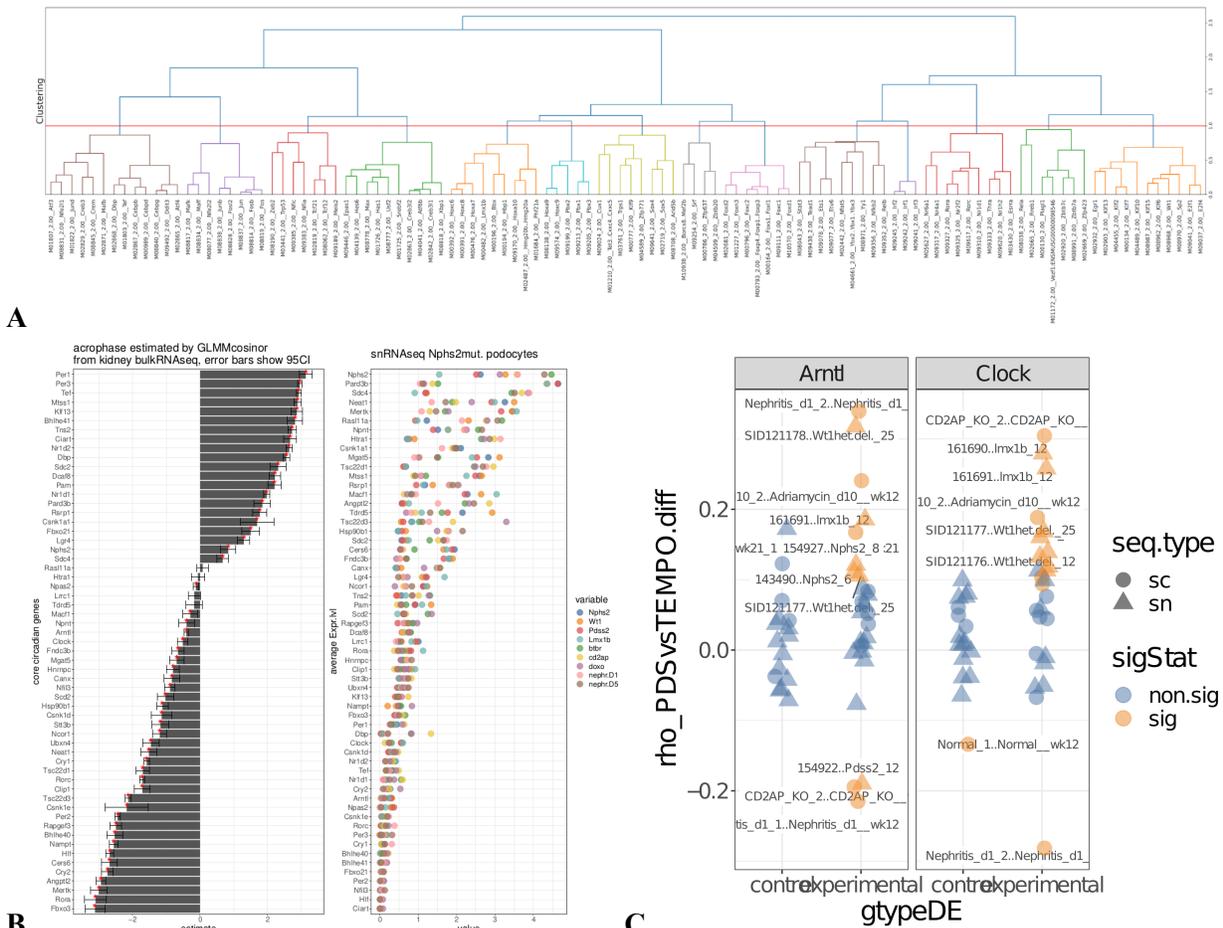
## A slit-diaphragm-associated protein network, PMID: 36307401



Supplementary Figure 4.15: Damage-associated gene level changes in pathway context  
**(A)** PDS mapped on Slit Diaphragm PPI network: bars inside nodes reflect correlation of the gene expression with the score across FSGS models. **(B)** KEGG Focal adhesion pathway map. Vertical bars inside nodes reflect correlation of the gene expression with the score across FSGS models.



Supplementary Figure 4.16: Scatter plots show relation between the number of genes that significantly (FDR.adj.  $p < 0.1$ ) correlate with PDS and various sample metrics.



Supplementary Figure 4.17: Circadian gene analysis in podocyte damage context. (A) Hierarchical Clustering of PWM motifs of 110 transcription factors expressed in podocytes. Red line shows the height at which the tree was cut. (B) Acrophase estimate and expression of highly circadian genes in sc and snRNAseq datasets. (C) Correlations between the variation in circadian time and PDS. Circadian time was calculated using a Tempo algorithm with Arntl or Clock as a reference gene.

# Appendix B: Supplementary Tables

Model colloquial	Explanation	Abbreviation	Full description for methods	Datasets	Description of the model
Nphs2	Nphs2 <sup>R231Q/A286V</sup> mutant	Nphs2 <sup>mut</sup>	Nphs2 <sup>R231Q/A286V</sup> compound heterozygous mutant mice (PMID: 32694662)	BulkRNA-seq, snRNA-seq	The <i>Wt1</i> gene encodes the Wilms' tumor 1 protein, a transcription factor critical for podocyte development and function. Heterozygous knockout of <i>Wt1</i> in podocytes leads to podocyte injury, resulting in proteinuria and progressive glomerulosclerosis. This model mimics human conditions like Denys-Drash and Frasier syndromes, which are associated with <i>WT1</i> mutations and lead to nephrotic syndrome and kidney failure.
Wt1	Wt1 heterozygous	Wt1 <sup>ko/wt</sup>	Wt1 heterozygous whole-body knockout (PMID: 12898605)	BulkRNA-seq, Chip-seq, snRNA-seq	<i>Nphs2</i> encodes podocin, a protein essential for the structure and function of the slit diaphragm in podocytes. Mutations in <i>Nphs2</i> disrupt podocin function, leading to structural defects in the filtration barrier, which causes podocyte injury, proteinuria, and rapid progression to FSGS-like glomerulosclerosis. This model mimics autosomal recessive forms of FSGS seen in humans with <i>NPHS2</i> mutations
Pdss2	Conditional podocyte Pdss2 knock-out	Pdss2 <sup>pko</sup>	Nphs2 <sup>Cre</sup> x Pdss2 <sup>fl/fl</sup> conditional podocyte knockout mice	snRNA-seq (KFO)	This model involves the conditional knockout of the <i>Pdss2</i> gene, which is crucial for coenzyme Q biosynthesis. Loss of <i>Pdss2</i> function in podocytes leads to mitochondrial dysfunction, resulting in podocyte injury, proteinuria, and progressive glomerular damage, ultimately leading to FSGS-like phenotypes.
Btbr	Diabetic nephropathy	Btbr <sup>ob/ob</sup>	BTBR.Cg- <i>Lep<sup>ob</sup>/Wi scJ</i> , Data taken from Chung et al. (PMID: 32651223)	scRNA-seq (public)	The <i>Btbr</i> (black and tan brachyury) mouse model is used to study type 2 diabetic nephropathy. These mice develop insulin resistance, hyperglycemia, and progressive kidney damage, including glomerulosclerosis and podocyte loss, mimicking the chronic kidney disease observed in human diabetic nephropathy.
Cd2ap	Conditional podocyte Cd2ap knock-out	Cd2ap <sup>pko</sup>	Nphs2 <sup>Cre</sup> x Cd2ap <sup>fl/fl</sup> , data taken from Chung et al. (PMID: 32651223)	scRNA-seq (public)	The <i>Cd2ap</i> gene is essential for maintaining the integrity of the slit diaphragm in podocytes. Conditional knockout of <i>Cd2ap</i> leads to podocyte detachment and loss, resulting in proteinuria and severe glomerulosclerosis, closely resembling FSGS phenotypes.
Doxo	Adriamycin toxicity model	Adria	Data taken from Chung et al. (PMID: 32651223)	scRNA-seq (public)	The <i>Doxo</i> model, induced by the administration of Adriamycin (doxorubicin), causes podocyte damage due to direct cytotoxic effects. It is used to model nephrotic syndrome and leads to podocyte loss, proteinuria, and glomerulosclerosis.
Nephritis	Nephrotoxic serum nephritis	NTSN	Data taken from Chung et al. (PMID: 32651223)	scRNA-seq (public)	This model is induced by administering nephrotoxic serum to mice, which triggers an immune response that causes inflammation and podocyte injury. It mimics immune-mediated glomerulonephritis, characterized by proteinuria, podocyte loss, and glomerular damage.

Supplementary table 1. Podocyte damage models used in the podocyte damage score study.

Eartag	CCG Sample ID	Age weeks	Genotype	Alb/Cre measures
AKK192	49346	4	mut	
AKK193	49347	4	wt	0,03271186
AKK194	49348	4	wt	
AKK195	49349	4	mut	0,486439025
AKK196	49350	4	wt	
AKK197	49351	4	mut	
AKK179	49359	12	wt	0,019362681
AKK178	49360	12	wt	0,019057877
AKK180	49361	12	mut	0,506674524
AKK181	49362	12	mut	0,514788289
AKK183	49363	12	mut	0,534743659
AKK185	49364	12	wt	
BDP_1	109583	4	wt	
BDP_3	109585	4	wt	
BDP_242	109587	4	wt	
BDP_255	109589	4	mut	
BDP_256	109591	4	mut	
BDP_266	109593	4	mut	
BDP_226	109619	8	wt	
am-913	109621	8	wt	
am-916	109623	8	wt	
BDP_207	109625	8	mut	
BDP_210	109627	8	mut	
BDP_227	109629	8	mut	
BDP_229	109631	8	mut	
BDP_231	109633	8	mut	

Supplementary table 2. Glomeruli samples subjected to bulk RNA sequencing.

CCG Sample ID	Sample Name	Genotype	line	Strain	run	Age in weeks	Gender	Alb/Cr Ratio
143483	CDD_177	CoQ2	Coq2.S96N/Podocin2A-iCre2A-mT	C57BL6/Nrj	1	6	f	
143484	CDD_190	CoQ2	Coq2.S96N/Podocin2A-iCre2A-mT	C57BL6/Nrj	1	6	m	
143485	CDC_46	Pdss2	Pdss2.V117M/Podocin2A-iCre2A-mT	C57BL6/Nrj	1	6	m	0.108
143486	CDC_47	Pdss2	Pdss2.V117M/Podocin2A-iCre2A-mT	C57BL6/Nrj	1	6	f	0.108
143487	CDD_188	wt	Coq2.S96N/Podocin2A-iCre2A-mT	C57BL6/Nrj	1	6	f	
143488	CDD_192	wt	Coq2.S96N/Podocin2A-iCre2A-mT	C57BL6/Nrj	1	6	m	
143489	BDP_399	wt	Podocin.R231Q/Podocin.A286V	C57BL6/Nrj	1	6	m	
143490	BDP_401	Nphs2	Podocin.R231Q/Podocin.A286V	C57BL6/Nrj	1	6	f	
146985	BWV_175	Pdss2	Pdss2.V117M	C57BL6/Nrj	2	21	f	10.944
146986	BWV_170	Pdss2	Pdss2.V117M	C57BL6/Nrj	2	21	m	10.944
146987	BWV_180	wt	Pdss2.V117M	C57BL6/Nrj	2	21	m	
146988	BWV_187	wt	Pdss2.V117M	C57BL6/Nrj	2	21	f	
154921	BWV_292	Pdss2	Pdss2.V117M	C57BL6/Nrj	3	12	m	1.853
154922	BWV_294	Pdss2	Pdss2.V117M	C57BL6/Nrj	3	12	f	2.561
154923	Am_7855	wt	breeder	C57BL6/Nrj	3	14	m	0.010
154924	Am_7856	wt	breeder	C57BL6/Nrj	3	14	f	0.007
154925	Ac_7788	wt	breeder	CD1	3	16	f	0.014
154926	Ac_7789	wt	breeder	CD1	3	16	f	0.008
154927	BDP_483	Nphs2	Podocin.R231Q/Podocin.A286V	C57BL6/Nrj	3	8	m	1.752
154928	BDP_495	Nphs2	Podocin.R231Q/Podocin.A286V	C57BL6/Nrj	3	6	m	1.667
140739	BDP398	wt	Podocin wt/wt	C57BL6/Nrj	4	4	f	0.0084
140738	BDP397	Nphs2	Podocin R231Q/A286V	C57BL6/Nrj	4	4	f	0.368
140740	BDP396	Nphs2	Podocin R231Q/A286V	C57BL6/Nrj	4	4	m	2.542
140741	BDP392	wt	Podocin wt/wt	C57BL6/Nrj	4	4	m	0.034
139919	BDP352	Nphs2	Podocin R231Q/A286V	C57BL6/Nrj	4	8	f	
139917	BDP351	wt	Podocin wt/wt	C57BL6/Nrj	4	8	f	0.001
139921	BDP342	Nphs2	Podocin R231Q/A286V	C57BL6/Nrj	4	12	f	10.280
139913	BDP341	Nphs2	Podocin R231Q/A286V	C57BL6/Nrj	4	12	m	5.584
139915	BDP340	wt	Podocin wt/wt	C57BL6/Nrj	4	12	m	0.016
139911	BDP339	wt	Podocin R231Q/wt	C57BL6/Nrj	4	12	m	0.045
121173	AKK_834	wt	Wt1_KO	FVB/N	5	13	m	0.007
121174	AKK_835	wt	Wt1_KO	FVB/N	5	13	m	0.007
121175	AKK_843	Wt1het.del.	Wt1_KO	FVB/N	5	12	m	1.270
121176	AKK_844	Wt1het.del.	Wt1_KO	FVB/N	5	12	m	0.223
121177	AKK_827	Wt1het.del.	Wt1_KO	FVB/N	5	25	m	0.888
121178	AKK_830	Wt1het.del.	Wt1_KO	FVB/N	5	25	m	1.715
155914	AKK957	wt	Wt1_KO	FVB/N	6	4	m	0.005
155915	AKK962	wt	Wt1_KO	FVB/N	6	4	m	0.011
155916	AKK963	Wt1het.del.	Wt1_KO	FVB/N	6	4	m	0.011
155917	AKK964	Wt1het.del.	Wt1_KO	FVB/N	6	4	m	0.007

Supplementary table 3. Glomeruli samples subjected to single nuclei RNA sequencing.

GSE ID	Seq. type	Model	Model group	sample type
GSE138774	bulk	Arhgef7 knockout	actin cytoskeleton	glomeruli
GSE136138	bulk	aged mice	ageing	podocytes
GSE43061	MA	Ercc1 progeria	ageing / DNA damage	glomeruli
GSE134327	bulk	streptozotocin eNOS-/- - samples GSM3942312-3 control, GSM3942314-6 streptozotocin, GSM3942320-1 control	diabetes	glomeruli
GSE123853	bulk	db/db datasets GSM3514331-6 control, GSM3514343-8 disease	diabetes	glomeruli
GSE77717	bulk	db/db	diabetes	glomeruli
GSE79291	bulk	streptozotocin	diabetes	glomeruli, podocytes
GSE106841	MA	ob/ob (leptin), 4,8,12,16 weeks !	diabetes	glomeruli
GSE112116	MA	streptozotocin	diabetes	glomeruli
GDS3992	MA	OVE26	diabetes	glomeruli
GSE131266	MA	ob/ob mice - only datasets GSM3768235-40	diabetes	glomeruli
GSE36209	MA	OVE26	diabetes	podocytes
GSE168676	bulk	streptozotocin + podocyte-specific Dach1 knockout	diabetes (STZ), transcription factor (Dach1), combined damage ( TF + Diabetes)	glomeruli
GSE104624	MA	CD151 knockout	focal adhesion	glomeruli
GSE181690	bulk	Parva KO	focal adhesion	glomeruli
GSE126217	bulk	Cosmc KO	glycan metabolism	glomeruli
GSE117987	bulk	HIV adriamycin	HIV-mediated damage + toxic damage	glomeruli
GSE56236	MA	anti-GBM nephritis / nephrotoxic serum	immunologic damage	glomeruli
GSE119049	bulk	light chain deposition disease (_GLO_FRA datasets: disease, _GLO_WT dataset control, _GLO_DH datasets control)	light chain deposition	glomeruli
GSE164273	sn	Pdss KO	metabolism, oxidative phosphorylation	whole kidney
KFO	bulk	Nphs2-R231Q/A286V	slit diaphragm	glomeruli
KFO	sc	Nphs2-R231Q/A286V	slit diaphragm	
GSE123179	bulk	Cd2apFyn	slit diaphragm	glomeruli
GSE110092	bulk	shroom3 knockdown	slit diaphragm	glomeruli
GSE63272	MA	Cd2ap	Slit diaphragm	podocytes
GSE154955	bulk	adriamycin	toxic damage	podocytes
KFO	bulk	adriamycin	toxic damage	glomeruli
GSE108629	MA	NEP25	toxic damage	podocytes

KFO	sc	Wt1 het.del	Transcription factor	podocytes
GSE96044	bulk	MafB KO	Transcription factor	glomeruli
GSE174102	sc	CTCF KO	Transcription factor	all glom cells
KFO	bulk	Wt1 het.del	Transcription factor	glomeruli
GSE18358	MA	Dendrash Wt1 mut	Transcription factor	glomeruli
GSE117571	MA	Foxc1/2	Transcription factor	glomeruli, primiry podocytes
GSE17709	MA	PTIP KO	Transcription factor	glomeruli
GSE146912	sc	4 models: nephrotoxic serum nephritis, diabetes, doxorubicin toxicity, CD2AP deficiency	various groups depending on the dataset.	all glom cells

Supplementary table 4. Transcriptomic datasets used to generate universal podocyte damage signature.

gene_symbol	mean_rank	direction_foldchange
Thsd7a	88.6847826086957	-1
Aifm3	90.0652173913043	-1
Metrn1	105.282608695652	-1
Nap111	108.195652173913	-1
Magi2	108.95652173913	-1
F2r	109.521739130435	1
Pak1	109.673913043478	-1
Robo2	113.065217391304	-1
Gja3	114.739130434783	-1
Astn2	116.684782608696	-1
Plce1	116.739130434783	-1
Sulf1	116.913043478261	-1
Clic3	119.173913043478	-1
Shisa3	119.902173913043	-1
Ankrd1	120.29347826087	1
C1qtnf7	120.804347826087	-1
Spats21	121.391304347826	-1
Lrrc1	122.260869565217	-1
Nexn	123.010869565217	1
Optn	123.217391304348	-1
Cldn5	123.282608695652	-1
Tspan15	124.217391304348	-1
Ptpro	125.804347826087	-1
Angptl2	126.804347826087	-1
Sema3g	128.608695652174	-1
Smarca2	128.826086956522	-1
Fgfbp1	129.163043478261	-1
Itgb8	129.45652173913	-1
Pard3b	130.304347826087	-1
Snaip	130.369565217391	-1
Dusp15	132.413043478261	-1
Ephb1	132.434782608696	-1
Wt1	132.478260869565	-1
Nek1	134.869565217391	-1
Ddn	136.282608695652	-1
Mafb	136.554347826087	-1
Mtss1	136.95652173913	-1
Zfp423	137.423913043478	-1
Enpep	138.271739130435	-1
Sema3e	138.434782608696	-1
Dach1	139.45652173913	-1

Supplementary table 5. First 42 markers of the podocyte damage signature, used for calculation of podocyte damage score.