

Regulation of RNA expression in nephron progenitor cells

by

Andrew Scott Clugston

B.S., Rochester Institute of Technology, 2008

M.S., Rochester Institute of Technology, 2010

Submitted to the Graduate Faculty of the
School of Medicine in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This thesis was presented

by

Andrew Scott Clugston

It was defended on

September 25, 2020

and approved by

Neil Hukriede, Professor, Integrative Systems Biology

Panagiotis Benos, Professor, Computational and Systems Biology

Michael Butterworth, Associate Professor, Department of Cell Biology

Maria Chikina, Assistant Professor, Computational and Systems Biology

Dissertation Directors:

Jacqueline Ho, M.D. MSc, Children's Hospital of Pittsburgh Department of Pediatrics

Dennis Kostka, Associate Professor, Department of Developmental Biology

Copyright © by Andrew Scott Clugston

2020

Regulation of RNA expression in nephron progenitor cells

Andrew Scott Clugston, PhD

University of Pittsburgh, 2020

To function, the mammalian kidney is dependent on its endowment of nephrons: the cellular structures within which filtration of waste as well as reabsorption of water, proteins, and important solutes take place. In mammals, these nephrons originate entirely during development and are never replaced, making an individual's nephron endowment at birth of critical importance for a long and healthy life. A low nephron endowment increases risk of hypertension, cardiovascular disease, and chronic kidney disease. Genetic abnormalities that affect the kidney are among the most common developmental defects observed, and hypoplastic kidneys (those with an abnormally low nephron endowment) are estimated to occur in 2% of all births. The causes of these conditions are varied as kidney development is a complex process that hinges on dynamic feedback mechanisms between over 20 cell types, but key among the cell types influencing nephron endowment are the nephron progenitor cells. Nephron progenitors are multipotent cells that differentiate into nephrons, but they also self-renew to maintain their population long enough to build a full complement of nephrons. They also propagate feedback mechanisms which shape the kidney, and they respond dynamically to the continually changing structure of the kidney by modulating their rates of self-renewal and differentiation to meet its needs based on signals from one another and from their surrounding cells. Over the course of nephrogenesis, nephron progenitors become increasingly likely to differentiate, which gradually decreases their rate of renewal until their population is ultimately depleted in a wave of differentiation, brought on in part by expression of miRNA in the *let-7* family. Understanding the genetics of nephron endowment

requires understanding the dynamics of the progenitor population, which we sought to do by identifying miRNA expressed in nephron progenitors, then measuring how their expression changes over time. In parallel, we also measured changes in chromatin accessibility that implicated regulatory features that may drive some of these changes. Finally, we measured heterogeneous gene expression of the progenitor population using single-cell RNA-seq, and observed parallels between the changes that define progenitors of different ages and progenitors of different predilections for self-renewal versus differentiation.

Table of Contents

1.0 Introduction.....	1
1.1 The mammalian kidney	1
1.2 Mammalian kidney development in brief	3
1.3 The nephron progenitor cell	5
1.4 Age-dependent changes in nephron progenitors	11
1.5 MicroRNA and the cessation of nephrogenesis	13
1.5.1 Canonical microRNA.....	13
1.5.2 MicroRNA in nephron progenitors	17
1.6 Enhancers and the cessation of nephrogenesis	18
1.6.1 Enhancer activity	18
1.6.2 Enhancer activity and chromatin accessibility.....	22
1.6.3 Assay for transposase-accessible chromatin sequencing (ATAC-seq).....	24
1.6.4 Enhancers in nephron progenitors.....	28
1.7 Enhancers of microRNA	30
2.0 Small non-coding RNA expression in mouse nephrogenic mesenchymal progenitors.....	33
2.1 Foreword	33
2.2 Summary	34
2.3 Introduction	35
2.4 Methods	37
2.4.1 Nephron progenitor isolation and total RNA preparation.....	37

2.4.2 Small RNA library preparation and sequencing	40
2.4.3 Small RNA-Seq data analysis and quality control.....	40
2.4.4 Data Records.....	41
2.4.5 miRNA Quantitative PCR and Locked Nucleic Acid in situ hybridization validation.....	41
2.4.6 Code Availability.....	42
2.5 Results.....	43
2.5.1 Cellular enrichment and total RNA quality control	43
2.5.2 Quality control of the cDNA libraries	46
2.5.3 Quality control of sRNA-Seq data	46
2.5.4 Locked Nucleic Acid section in situ hybridization and qPCR validation of the small RNA-Seq data.....	49
2.6 Discussion	53
3.0 Enhancers of microRNA in the cessation of nephrogenesis.....	56
3.1 Foreword	56
3.2 Summary	57
3.3 Introduction	58
3.4 Methods	61
3.4.1 Mouse strains.....	61
3.4.2 Nephron progenitor isolation.....	61
3.4.3 Chromatin accessibility library preparation	63
3.4.4 Chromatin accessibility sequencing	64
3.4.5 Identifying accessible chromatin regions	64

3.4.6	Changes in accessible chromatin regions.....	65
3.4.7	Transcription factor footprints.....	66
3.4.8	Nucleosomal configuration.....	67
3.4.9	Small RNA sequencing and analysis	67
3.4.10	Screening for regulatory elements affecting miRNA expression.....	68
3.5	Results.....	68
3.5.1	Nephron progenitor cell populations from E14.5 and P0 kidneys	68
3.5.2	Early and late nephron progenitors have distinct chromatin environments	69
3.5.3	Early and late nephron progenitors have distinct miRNA transcriptional profiles.....	74
3.5.4	Topologically associated domains help identify potential enhancer—miRNA relations.....	76
3.6	Discussion	81
4.0	Single cell RNA sequencing reveals differential cell cycle activity in key cell populations during nephrogenesis.....	84
4.1	Foreword	84
4.2	Summary	85
4.3	Introduction	86
4.4	Methods	90
4.4.1	Embryonic kidney collection and single-cell RNA sequencing.....	90
4.4.2	Data processing, quality control, and normalization.....	90
4.4.2.1	Alignment and read counting	90

4.4.2.2 Quality Control	91
4.4.2.3 Normalization.....	91
4.4.3 Identification of major structural components of the kidney	92
4.4.4 Nephron progenitor and descendant cell types	92
4.4.4.1 Selecting and characterizing NP lineage cells	92
4.4.4.2 Pseudotime Analysis of NP cells	93
4.4.4.3 Enrichment Analysis for NP cells.....	94
4.4.4.4 Immunohistochemical staining.....	94
4.4.4.5 In situ hybridization	95
4.4.5 Reproducibility and data availability.....	95
4.5 Results.....	95
4.5.1 Single cell gene expression identifies anatomical structures and cell lineages in the developing kidney	95
4.5.2 Stratification of cell-types in the nephron progenitor lineage	97
4.5.3 Transcriptional dynamics across nephron progenitor cell differentiation.....	100
4.5.3.1 Cell cycle activity distinguishes between two populations of nephron progenitor cells.....	100
4.5.3.2 <i>Birc5</i> expression in the tubular interconnection zone	103
4.5.3.3 Conserved features in mouse and human podocyte development...	104
4.5.3.4 Gene expression differences between proximal and distal tubular cells	106
4.5.4 Expression of known lineage-marker genes in unexpected cell types	107
4.6 Discussion	110

5.0 Discussion and future directions.....	115
5.1 miRNA expression contributes to age-dependent changes in nephron progenitors	115
5.2 Chromatin accessibility contributes to age-dependent changes in nephron progenitors	117
5.3 Gene expression reflects heterogeneity among nephron progenitor cells	119
5.4 Enhancers of miRNA	121
5.5 Nephron progenitor cells	123
6.0 Non-thesis work.....	124
7.0 Publications	125
Appendix.....	126
Supplemental figures.....	126
Bibliography	138

List of Tables

Table 1. Primer sequences used for quantitative PCR.....	42
Table 2. Primer sequences used for novel miRNA quantitative PCR.....	42
Table 3. Exiqon miRCURY LNA detection probes used for miRNA section in situ hybridisation.....	44
Table 4. Sample IDs and RIN score of RNA samples used for small RNA-Sequencing.	44
Table 5. Quantitative PCR profile of novel miRNA assayed	50
Table 6. RT-PCR primers.	62
Table 7. Co-expression of stromal marker genes and nephron progenitor marker genes in cortical/stromal cells.	108
Table 8. Ureteric bud / collecting duct lineage genes are expressed in distal tubular cells..	109

List of Figures

Figure 1. Major structures of the mammalian kidney and nephron.	2
Figure 2. Reciprocal induction of the ureteric bud and metanephric mesenchyme.	5
Figure 3. The nephron progenitor niche.	10
Figure 4. miRNA processing.	16
Figure 5. Enhancer-promoter interactions.	22
Figure 6. ATAC-seq measures chromatin accessibility.	28
Figure 7. Periods of nephrogenesis measured.	32
Figure 8. Schematic pipeline illustrating the workflow of nephron progenitor isolation and bioinformatics analysis of the small RNA-Seq dataset.	39
Figure 9. Quality control of the input samples demonstrated an enrichment of nephron progenitors in the cellular isolate with good quality RNA, and successful cDNA library synthesis.	45
Figure 10. Quality control of the small RNA-Seq dataset showed good sequencing quality reads and congruence of the biological samples.	48
Figure 11. Small RNA-Seq and miRDeep2 analysis revealed novel unannotated miRNAs expressed in the developing kidney.	51
Figure 12. Integrative Genome Viewer visualisation of Six2 binding upstream of expressed miRNAs in nephron progenitors.	53
Figure 13. Differences in chromatin accessibility characterize E14.5 vs. P0 nephron progenitor cells.	71
Figure 14. ATAC-seq profiles at known and putative regulatory loci.	73

Figure 15. Substantial changes in miRNA expression between E14.5 and P0 nephron progenitor cells.....	76
Figure 16. Finding candidate enhancer—target-miRNA pairs.....	79
Figure 17. Three potential enhancer—target-miRNA pairs.....	80
Figure 18. Developing embryonic day 14.5 mouse kidney cell types.....	89
Figure 19. Cell types of the nephron progenitor lineage.	99
Figure 20. Transcriptional signatures of self-renewing, primed and differentiating nephron progenitor cells.....	102
Figure 21. Transcriptional signatures of podocytes and tubular cells.....	105
Figure 22. Expression of lineage-marker genes in unexpected cell types.....	109
Supplemental figure 1. Nephron progenitor isolation.....	127
Supplemental figure 2. ATAC-seq quality control.....	128
Supplemental figure 3. Genomic Region Enrichment of Annotations Tool (GREAT) analysis results.....	129
Supplemental figure 4. DIANA miRPath results.....	130
Supplemental figure 5. Heatmap of marker genes that distinguish between kidney cell types.....	132
Supplemental figure 6. Differentiation lineages for nephron progenitor cells.....	133
Supplemental figure 7. Differentially expressed genes during nephron progenitor cell differentiation.....	135
Supplemental figure 8. Differentially expressed genes in the podocyte and tubular lineages.....	136

Supplemental figure 9. Consistent *Birc5* expression in distal tubular cells and a subpopulation of cells from the ureteric bud..... 137

1.0 Introduction

1.1 The mammalian kidney

The evolution of the kidney allowed multicellular organisms to balance their internal aqueous environment in the face of changing osmotic surroundings, filtering waste and balancing pH, electrolytes, and water content. In mammals, the multicellular structure in which these requisite functions take place is the nephron¹, and these may number more than 1 million in a healthy human kidney (Figure 1)². This enables the kidneys to filter on the order of 180 liters of fluid per day³, but it is also important for longevity as mammalian nephrons that are lost or damaged cannot be replaced after birth⁴. A reduced nephron endowment places an individual at risk of hypertension, cardiovascular disease, and chronic kidney disease⁵, and nephron loss increases strain on the remaining nephron population which risks further damage and loss, potentiating a downward spiral toward kidney failure with few available treatments besides dialysis and kidney transplant⁶. Congenital causes of a reduced nephron endowment are common⁷, and overall nephron count in humans can vary widely⁵, but the development of the mammalian kidney is a complex process requiring the coordinated differentiation and interaction of more than 20 distinct cell types across three dimensions and over time^{3,8}. Genetic malformations of the kidney are among the most common genetic diseases, occurring three to six times for every 1,000 live births and causing between 34 and 59% of pediatric chronic kidney disease cases⁹. While often difficult to differentiate from co-occurring disease types¹⁰, renal hypoplasia (small kidneys with low nephron number but otherwise normal morphology) are expected to occur in approximately 2% of births^{10,11}.

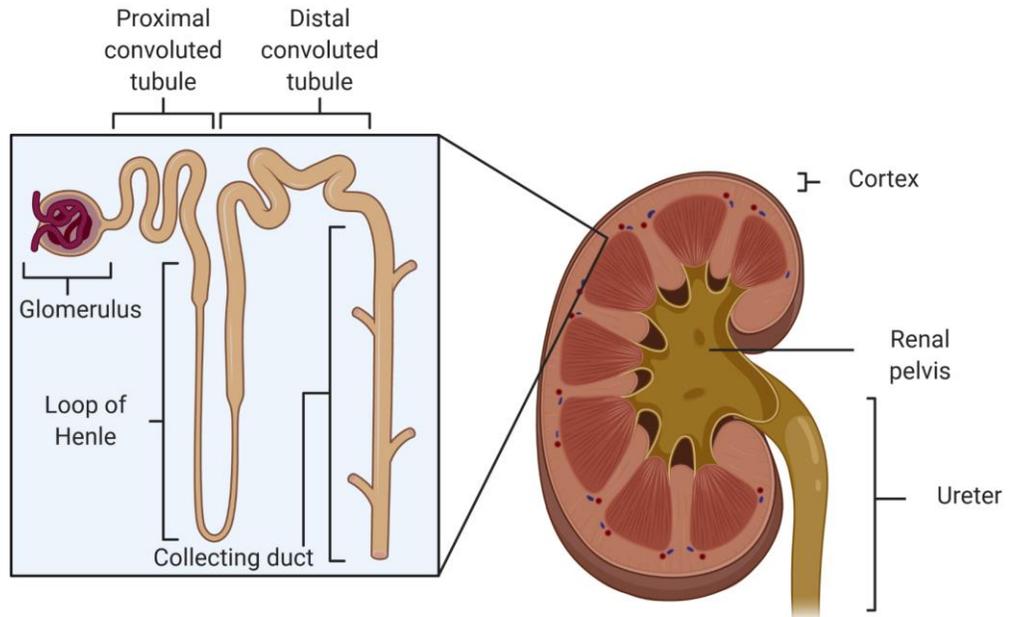


Figure 1. Major structures of the mammalian kidney and nephron.

The human kidney comprises an outer cortex containing hundreds of thousands to more than 1 million nephrons (inset) surrounding tree-like branches of collecting ducts extending outward from the renal pelvis. Filtration takes place in the glomerulus, and reabsorption of water, proteins, sugars, and other solutes takes place while filtrate travels through the proximal convoluted tubule, the loop of Henle, and distal convoluted tubule before finally flowing through the collecting ducts, into the renal pelvis, and out through the ureter. Image created with BioRender.com.

1.2 Mammalian kidney development in brief

The kidney arises from the intermediate mesoderm¹² around embryonic day 9.5 (E9.5) in mice⁸ (E35 in humans¹³), emerging as an epithelial growth called the ureteric bud from the mesonephric duct (Figure 2, left)¹⁴. Initiating the formation of the metanephric kidney, the ureteric bud extends into an adjacent layer of loose, undifferentiated intermediate mesoderm cells known as the metanephric mesenchyme¹⁵, attracted by their secretion of Glial cell line-derived neurotrophic factor (Gdnf; Figure 2, center) which interacts with RET receptor tyrosine kinases (Ret) on the bud's surface¹⁶. As this invagination takes place, the two tissues begin to exchange reciprocal developmental signals that drive the induction of new nephrons¹⁷. Ongoing secretion of Gdnf from the mesenchymal cells instructs the tips of the extending ureteric bud to branch (Figure 2, right)¹⁶, which ultimately shapes the tree-like network of collecting ducts that will end in the pelvis, and which will funnel urine down the ureters and toward the bladder (Figure 1)⁸. Simultaneously, cells at the tips of the ureteric buds secrete a mix of signals including Wnt family members 9b and 11 (Wnt9b, Wnt11) and Bone morphogenic protein 7 (Bmp7), which prevent cells in the metanephric mesenchyme from undergoing apoptosis¹⁸ and cause them to proliferate and condense around the ureteric tip, forming a transient structure known as the cap mesenchyme⁸. After each of the ureteric bud's many branching events, cells of the cap mesenchyme will be induced by increased concentrations of Wnt9b emanating from just beneath the newly formed branch to coalesce into pre-tubular aggregates. These aggregates are collections of cells that will polarize and differentiate into the renal vesicle, a structure marking the beginning of the construction of a new nephron¹⁹. The tips of the ureteric bud will branch approximately twelve times in the mouse kidney over the course of nephrogenesis²⁰, doubling the number of bud tips per kidney each time¹⁷. Each branching event requires the existing cap mesenchyme to not only commit cells for early nephrons, but also

to cover each new bud tip¹⁵. To compensate for this drastic increase in surface area to cover, the cells of the cap mesenchyme are capable of self-renewal and will roughly double the cap mesenchyme's volume for each branching event¹⁷, increasing its initial size by several hundred fold before nephrogenesis is complete²¹. Healthy kidney formation hinges on the proliferation and maintenance of this cap mesenchyme population, but also on its timely depletion: Wilms tumor, the most common form of kidney cancer in children, is a consequence of residual undifferentiated mesenchymal cells²². For this reason, nephrogenesis ceases when the cap mesenchyme is exhausted following a final wave of differentiation¹⁷, roughly at birth in humans⁴ and between postnatal day 2 (P2) and P4 in mice¹⁷. The cells that constitute the cap mesenchyme and will differentiate into most components of the mature nephron are nephron progenitor cells²³, and the behavior of the cap mesenchyme is the emergent product of their interactions with each other and with the surrounding ureteric bud and renal stromal cells of the developing kidney^{17,24}.

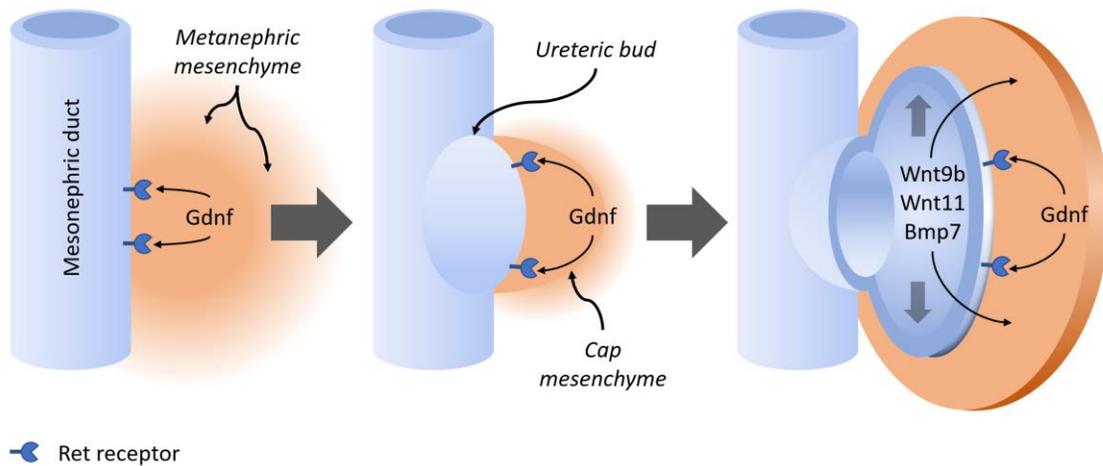


Figure 2. Reciprocal induction of the ureteric bud and metanephric mesenchyme.

Cells of the metanephric mesenchyme secrete Gdnf, which interacts with Ret proteins on the surface of mesonephric duct cells to induce ureteric bud outgrowth from the mesonephric duct. In response to signals from the ureteric bud, the metanephric mesenchyme condenses around the ureteric tip to form the cap mesenchyme. Ongoing expression of Gdnf in the cap mesenchyme induces branching to shape the ureteric tree, while the ureteric bud's expression of signaling molecules Wnt9b, Wnt11, and Bmp7 induces cap mesenchyme survival and growth.

1.3 The nephron progenitor cell

The cap mesenchyme is composed of nephron progenitor cells. The nephron progenitor cell is a multipotent mesenchymal cell committed to either “self-renew” by dividing into new nephron progenitors or to differentiate into the constituent cell types of the nephron²⁵. It remains motile and migrates about the cap mesenchyme, guided by an assortment of poorly characterized attractive and repulsive cues, occasionally adhering to the cells of the ureteric tip, occasionally breaking free to resume migration. This swarming behavior among the population of progenitors makes the cap mesenchyme a fluid structure that is continuously remodeled in reaction to the

constantly expanding structure of the ureteric tips²⁶. Within this structure, the nephron progenitor is dependent on signaling from the ureteric bud for cues to promote its survival, to the point that without them it will undergo apoptosis²⁷. The nephron progenitor also signals the ureteric bud, secreting Gdnf to promote the bud's induction and the branching of the ureteric tip²⁸. In response to this signal, the tip of the ureteric bud releases secreted growth factors that include Wnt11, Wnt9b, and Bmp7. A progenitor near the ureteric bud tip will proliferate in response to Wnt9b and Bmp7⁸, and will secrete additional Gdnf in response to Wnt11²⁹. This establishes a Wnt11/Gdnf feedback loop between nephron progenitors in the cap mesenchyme and the epithelial cells of the ureteric tip that will simultaneously drive the proliferation of nephron progenitors and the branching of the collecting duct network³⁰.

After each branching event of the ureteric bud, epithelial cells of the ureteric bud stalk below the branch "armpit" secrete a unique mixture of signals: Wnt9b is expressed in a greater concentration, and Wnt11 secretion is stopped entirely¹⁹. Combined with signals from the surrounding stromal cells³¹, a nephron progenitor entering this region is no longer induced to proliferate but instead to express Wnt family member 4 (*Wnt4*)¹⁹. *Wnt4* expression is a mark of nephron progenitor cells committing to a mesenchymal-to-epithelial transition³², which will transform them from a migrating mesenchymal type into an epithelial cell^{33,34}. This mechanism is regulated, however, because a differentiated nephron progenitor loses the ability to self-renew, and waves of differentiation among nephron progenitors would rapidly deplete the cap mesenchyme and preclude the formation of additional nephrons²⁴.

The different environments surrounding the ureteric bud and tip have different effects on the nephron progenitor cell, but the nephron progenitor cell's swarming behavior allows it to move back and forth among these regions multiple times²⁶ (Figure 3). Recent work by Lawlor et. al

(2019) showed that while some changes may accumulate based on the cell's migratory path, key consequential changes that result from the cell's wanderings are reversible: progenitor cells can be primed for differentiation by Bmp7, migrate to the base of the ureteric bud, and can still "escape" differentiation by migrating back into the cap mesenchyme. They also demonstrated that computer models of cap mesenchyme behavior which assume migrating progenitor cells return to a "base state" of gene expression are more in line with *in vivo* observations than models which assume cells experience cumulative sensitization. Moreover, these models indicate that the rate of cap mesenchyme depletion is a product of individual progenitor sensitivities as well as their stochastic migration patterns: a progenitor's sensitivity determines how long it needs to remain in the Wnt9b^{high} branch niche to be "captured," and its migration pattern determines the likelihood that this will happen³⁵.

Nephron progenitors of the cap mesenchyme are not a homogeneous population, nor are they homogeneously receptive to signals for differentiation²⁰. Wnt signaling through Wnt9b will eventually initiate the progenitor's differentiation program, but progenitors are at first predisposed to self-renew and are marked by their coexpression of Six Homeobox 2 (*Six2*) and Cpb/P300 Interacting Transactivator 1 (*Cited1*), denoted as the *Six2*⁺/*Cited1*⁺ population of progenitors³⁶. *Six2* is a marker of nephron progenitor multipotency³³, and kidneys formed without it are non-functional: nephron progenitors of the cap mesenchyme are depleted prematurely, and many of the pre-tubular aggregates that form do so ectopically and succumb to apoptosis³⁷. The role of *Cited1* is unclear and it is not essential for nephrogenesis³⁸, but nephron progenitor cells are receptive to differentiation signals from Wnt9b after transitioning to a "primed," *Cited1*-negative state (*Six2*⁺/*Cited1*⁻)³⁹. This transition is accomplished by Bmp7, which like Wnt9b is expressed by the ureteric bud and plays a dual role with nephron progenitors²¹. Among the secretions of the ureteric

bud tip both Bmp7 and Wnt9b not only promote proliferation but are required for it, with Bmp7 acting through the c-Jun N-terminal kinase (JNK) pathway⁴⁰ and Wnt9b via canonical Wnt signaling (promoting the accumulation of β -catenin)¹⁹. However, Bmp7 also acts through Smad signaling to prime progenitor cells for differentiation, transitioning them from the Wnt9b-resistant *Six2*⁺/*Cited1*⁺ state to the differentiation-primed *Six2*⁺/*Cited1*⁻ state³⁹. *Six2* expression is only down-regulated after nephron progenitors incorporate into a pre-tubular aggregate, at which point Notch signaling will begin to repress *Six2* expression via Notch2⁴¹.

Besides the ureteric bud, the renal stromal cells surrounding nephron progenitors also affect their receptiveness to signals for differentiation. The stromal cells that surround the cap mesenchyme have been shown to influence nephron progenitors through their expression of a variety of proteins, including *FAT Atypical Cadherin 4 (Fat4)*³¹ and *Forkhead Box D1 (Foxd1)*⁴². Fat4 is found on the surface of the stromal cell, and has the effect of upregulating gene targets of Wnt9b in nephron progenitors, thereby increasing Wnt9b's efficacy within the progenitor³¹. Foxd1 is a marker for the stromal cell lineage⁴³, and it impacts nephron progenitors by inhibiting Decorin (Dcn), a protein which obstructs the Smad pathway through which Bmp7 primes nephron progenitors for differentiation⁴². Both mechanisms represent a pro-differentiation signal originating from the *Foxd1*⁺ stromal cells⁴².

Nephron progenitor cells are also responsive to the growth factor Fibroblast Growth Factor 20 (Fgf20), which promotes their proliferation and survival. This signal is in fact an *autocrine* signal, both secreted *and* perceived by the nephron progenitor population. In mice Fgf20 is partially redundant with Fibroblast Growth Factor 9 (Fgf9) produced by the ureteric bud (normal development takes place as long as *at least* two functional copies of *Fgf9* remain), but human

kidneys developed with truncated transcripts for *Fgf20* fail to form entirely⁴⁴. These FGF signals are key components of cell media used to maintain multipotent nephron progenitors *in vitro*⁴⁵.

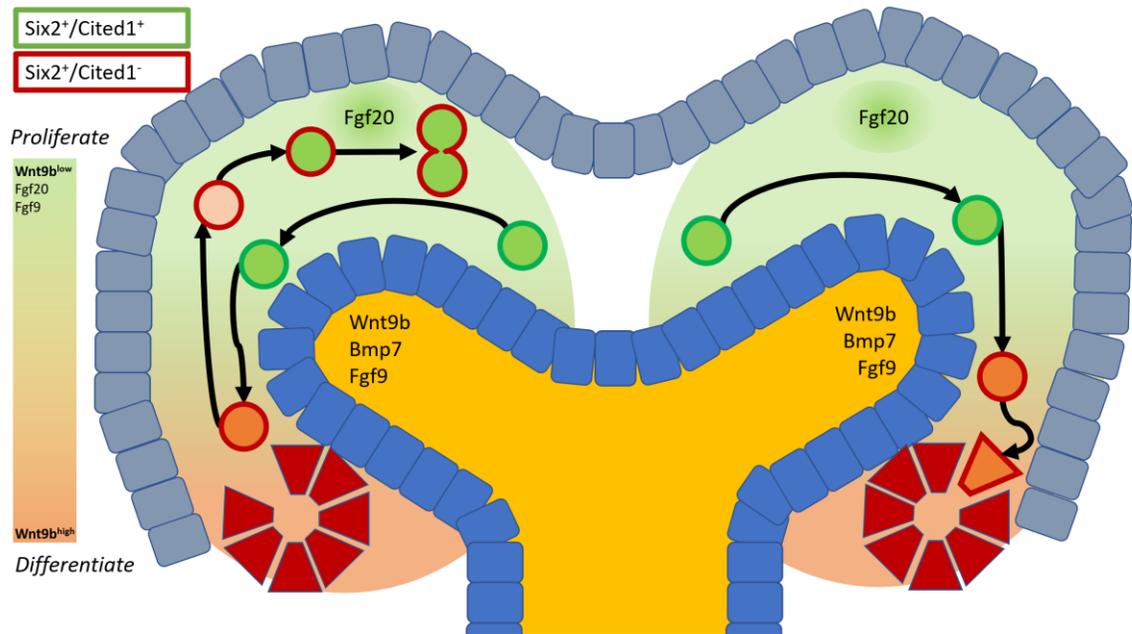


Figure 3. The nephron progenitor niche.

Nephron progenitors (circular cells) migrate freely about the cap mesenchyme, sandwiched between *Foxd1*⁺ stromal cells (gray) and epithelial cells of the ureteric bud (blue). Growth factor signals in the ureteric tip region instigate proliferation (green fill), while those beneath the tip branch promote differentiation (red fill) and transitioning from the Six2⁺/Cited1⁺, self-renewing niche (dark green outline) into the Six2⁺/Cited1⁻, pro-differentiation niche (red outline). Extended exposure of a Six2⁺/Cited1⁻ progenitor cell to high concentrations of Wnt9b (Wnt9b^{high}) leads to engraftment and differentiation (right side), but progenitor cells can also migrate back into the tip region's Wnt9b^{low} region and will revert to a baseline in gene expression (left side).

1.4 Age-dependent changes in nephron progenitors

The age at which cessation of nephrogenesis occurs directly affects nephron endowment, and the timing of this event is dependent on the replenishment of the nephron progenitor population⁴⁶. This is in turn influenced by the progenitor cell's exposure to signals for renewal and differentiation³⁵, but also the progenitor cell's age. Work by Chen et. al (2015) demonstrated that nephron progenitors isolated from mouse kidneys early (E12.5) and late (P0) in nephrogenesis display different capacities to engraft and to differentiate when transplanted into a separate E12.5 kidney, suggesting that there are intrinsic differences between “early” and “late” nephron progenitors. They showed that early progenitors engraft and renew more successfully, while late progenitors engraft less frequently and typically differentiate more quickly. Demonstrating the importance of the surrounding niche, however, they observed that some late nephron progenitors both engraft and remain in the undifferentiated state for several days, which requires that these cells have maintained multipotency several days longer than would have been expected in their original niche. They also demonstrated that older progenitors are more likely to proliferate if they have physical contact with young progenitors, which indicates that juxtacrine signaling between nephron progenitor cells is important in determining each progenitor's behavior. Finally, they showed that *Cited1*⁺ progenitors isolated at P0 are better able to maintain multipotency than *Cited1*⁻ progenitors isolated at E12.5 (though they are still less able to engraft after transplant), and “P0-like” progenitors can be isolated from cells collected at E14.5, but “E14.5-like” progenitors cannot be found at P0. While the surrounding cellular environment of nephron progenitors influences their propensity to differentiate or self-renew, an important portion of this balance is attributable to changes within the nephron progenitor cell over the course of nephrogenesis⁴⁷.

Gene expression profiles of progenitors isolated at multiple stages in development suggest that many transcriptional changes during this aging process are continuous, and a significant number of genes that exhibit increased expression with age are associated with ribosomal biogenesis, a key target of the mammalian target of rapamycin (mTOR) pathway⁴⁷. mTOR is a protein kinase in the Phosphoinositide-3-kinase (PI3K)-related kinase family, and within it mTOR modulates cell growth, survival, and proliferation of mammalian cells in response to environmental and metabolic conditions^{48,49}. In particular, mTOR complexed with Raptor (mTOR complex 1, or mTORC1) responds to growth factors and optimal growth conditions (minimal stress, ample energy, nutrient, and oxygen resources) by promoting cell growth-related functions (restricting protein catabolism, promoting nucleotide synthesis, glucose metabolism, and mitochondrial biogenesis, etc.). mTORC2 (mTOR complex 2, mTOR complexed with Rictor) responds to growth factors by transducing functions such as glucose metabolism, cell migration, cytoskeletal rearrangement, and resistance to apoptosis⁴⁹. Hemizygous deletion of *mTOR* in nephron progenitors severely limits nephron endowment, and homozygous deletion results in death immediately after birth⁵⁰. Further, hemizygous deletion of the mTORC1 repressor Hamartin (*Tsc1*) has the effect of increasing progenitor endowment by extending the duration of nephrogenesis. Interestingly, this activity appears to be independent of mTOR: it is instead dependent on its mTORC1 *complement*, Raptor. The mechanism at work appears to be complex, but it is thought that a reduction in *Tsc1* leads to an increase in free Raptor and an increase in mTORC2 activity relative to mTORC1⁵⁰. While an increase in mTORC2 activity is a counter-intuitive consequence of an increased abundance of Raptor, the authors hypothesize that free Raptor decreases phosphorylated Protein Kinase B (Akt)⁵¹ and increases glycolysis among nephron progenitor cells (a known trait in proliferating nephron progenitors^{52,53}), which

predominantly benefits the actions of mTORC2. Though the molecular mechanism is still uncertain, a hemizygous deletion of *Tsc1* results in a 25% increase in progenitor endowment in mice⁵⁰, and further implicates mTOR in aging- and development-related pathways modulating the nephron progenitor “aging” process.

1.5 MicroRNA and the cessation of nephrogenesis

1.5.1 Canonical microRNA

MicroRNA are pervasive and conserved throughout the animal kingdom⁵⁴: they are short (typically around 22 nucleotides long), non-coding strands of RNA that guide the RNA-induced silencing complex (RISC) to post-transcriptionally repress mRNA, either through cleavage of its 3' untranslated region (3' UTR) or by obstructing its translation⁵⁵. A typical miRNA is transcribed within a primary miRNA (pri-miRNA) transcript that may contain more than one miRNA sequence (Figure 4). The miRNA sequence within the pri-miRNA folds against a nearby complementary sequence in the transcript to form a “hairpin” structure, with the miRNA sequence occupying one of the two “arms” of the hairpin. The hairpins of the pri-miRNA transcript are released as “pre-miRNA” hairpins by the Drosha enzyme within a microprocessor complex, then exported from the nucleus into the cytoplasm by exportin 5. In the cytoplasm, the enzyme Dicer will cleave the hairpin’s “loop” segment to produce a miRNA duplex containing the complementary 3' and 5' strands of the miRNA hairpin, one or both of which may be a viable miRNA, typically identified with the “3p” or “5p” suffix to indicate the sequence direction relative the genome. This duplex will be loaded into RISC such that the “mature” miRNA strand is bound,

and the unbound “passenger” strand will be discarded. The bound miRNA fragment will then direct RISC to target mRNA containing miRNA response elements (MREs): sequences of RNA roughly complementary to an approximately 8bp long “seed sequence” within the RISC-bound miRNA. MREs that are an exact match to the seed sequence will be cleaved, and MREs with mismatches may still cause the mRNA’s translation to be obstructed by the bound RISC⁵⁵.

miRNA expression may not only serve to repress a gene’s production, but can also stabilize its transcript levels by repressing rapid increases in transcript⁵⁶. Schmiedel et. al (2015) reported that expression of miRNA to target a lowly expressed gene can have a stabilizing effect on that gene’s transcript number as sudden influxes of transcript are muted by proportional increases in miRNA binding. This effect is bolstered by the inclusion of more miRNAs and MREs. The opposite is true for genes which see high rates of transcription, as variations in targeting miRNA concentration may *increase* transcriptional noise⁵⁶. Interestingly, such an increase in noise may be intentional in certain systems, as when homogeneous, non-polarized cell populations need to differentiate into heterogeneous cell types. In a recent preprint describing mouse embryonic stem cells (ESCs), Chakraborty et. al (2019; preprint) report that miR-182 has a miniscule effect on the transcripts of several pluripotency factors, but its expression introduces transcriptional noise that results in the heterogeneous binding of these factors at their target enhancers. Without miR-182 these cells rarely transition out of the same ground state and remain relatively homogeneous. With miR-182, seemingly minute and inconsequential changes to these transcript populations can vault cells into different states that anticipate separate developmental trajectories⁵⁷.

As post-transcriptional repressors *in vivo*, the effects miRNA have are often modest, sometimes to a degree that can be difficult to discern with an assay or observed as a phenotype after a genetic knockout⁵⁶. Moreover, their short seed sequence length combined with the viability

of mismatched MREs ensures that miRNAs rarely if ever act in an environment with a single target mRNA, and target mRNAs rarely if ever find themselves targeted by a singular miRNA antagonist⁵⁸. This promiscuous nature also makes miRNA a uniquely tunable component within dynamic cellular mechanisms: natural examples have been identified that add or remove MREs by changing transcript isoforms⁵⁹, bias the preferred miRNA strand selection within RISC⁶⁰, incorporate multiple identical or mismatched MREs for an additive effect on a targeting miRNA's efficacy⁵⁶, and adjust miRNA efficacy stoichiometrically by introducing or subtracting "sponge" molecules with competitive MREs⁶¹. The overall repressive effects of a particular miRNA are often more apparent in the context of its co-expressed miRNA and the sum of their effects across all of their mRNA targets⁶².

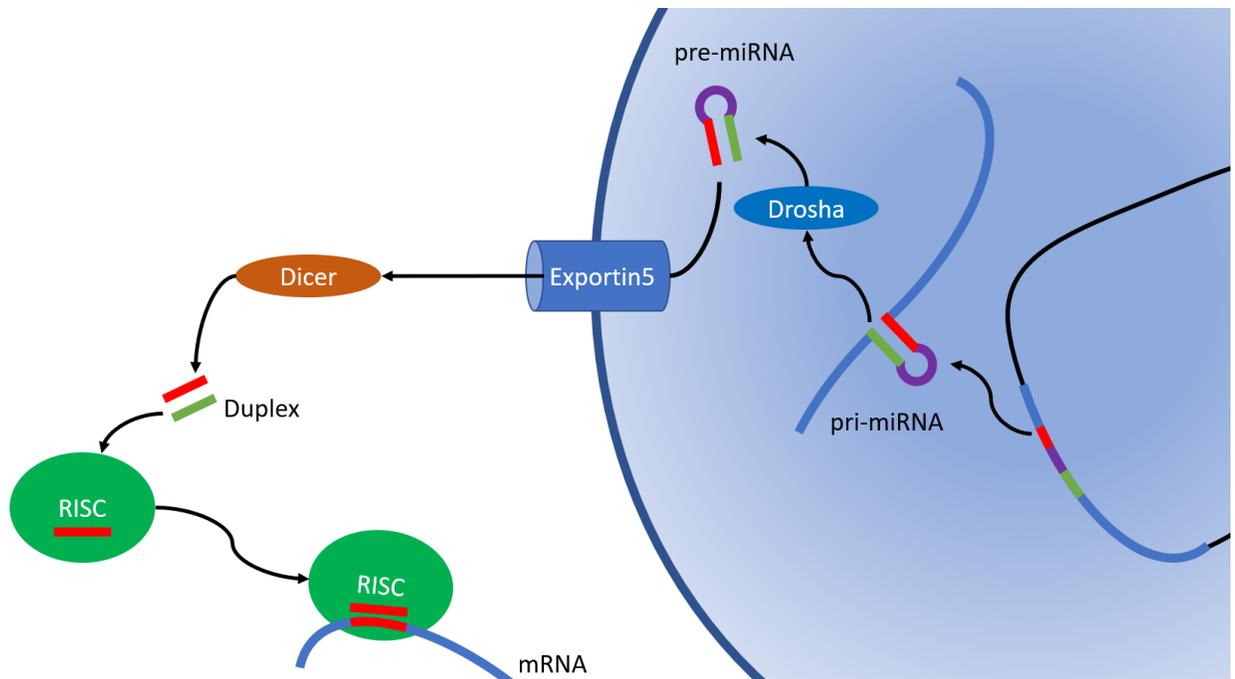


Figure 4. miRNA processing.

Processing of miRNA begins in the nucleus, where pri-miRNA containing one or more miRNA are transcribed and their miRNA fold into “hairpin” loops, with the miRNA (red) and a nearby complementary strand (green) making up the “arms” of the hairpin. A microprocessor complex containing Drosha frees the hairpin loop from the rest of the pri-miRNA to produce a pre-miRNA, which is then exported from the nucleus into the cytoplasm. Once there, Dicer cleaves the loop segment (purple) from the pre-miRNA to form the miRNA duplex. Either strand of this duplex may be a viable miRNA, and can be incorporated into the RNA induced silencing complex (RISC) while the opposite strand is degraded. Once incorporated into RISC, the complex will cleave or obstruct translation of mRNA containing a sequence in its 3’ untranslated region (UTR) that is approximately complementary to the incorporated miRNA strand (MRE).

1.5.2 MicroRNA in nephron progenitors

Expression of miRNA is essential in nephron progenitors for proper kidney development. Conditional loss of Dicer leaves a cell unable to process canonical miRNA, and in nephron progenitors this has the effect of depleting the cap mesenchyme prematurely due to progenitor apoptosis. This is caused by the pro-apoptotic protein Bcl-2L11, otherwise known as Bim, which is normally kept in check through miRNA expression (specifically by *miR-10a*, *miR-106b*, and *miR-17-5p*)⁶³. Conditional deletion of the hypoxia-responsive miRNA *miR-210* in nephron progenitors leads to a sex-specific 35% reduction in nephron endowment in male mice⁶⁴. Kidneys with the nephron progenitor-specific deletion of the *miR-17~92* cluster (which includes the miRNA *miR-17*, *miR-18a*, *miR-19a*, *miR-20a*, *miR-19b-1*, and *miR-92a-1*)⁶⁵ maintain their population of nephron progenitor cells and manage to avoid apoptosis, but see a reduction in nephron endowment and an increased incidence of chronic kidney disease⁶⁶. This appears to be partly a consequence of the loss of the *miR-17~92* member *miR-19b*, which affects nephron progenitor proliferation through its targeting of *Cystic Fibrosis Transmembrane Conductance Regulator (Cftr)*. Other genes affected by this *miR-17~92* deletion are associated with developmental pathways related to important processes that influence the nephron progenitor cell and its behavior within the cap mesenchyme, including cell motility, proliferation, and cell-cell adhesion⁶⁷.

The *let-7* family of miRNA is directly integrated into the nephron progenitor aging mechanism. The protein Lin-28 Homolog B (Lin28b) is a known repressor of the *let-7* family⁶⁸, and its expression in nephron progenitors decreases starting at E14.5 in mice (expression of its paralog, *Lin28a*, ceases in mouse nephron progenitors by E14.5)⁴⁶. Ectopic expression of *Lin28b* later in nephrogenesis blocks the terminal wave of differentiation that should deplete the cap

mesenchyme, causing a Wilms tumor-like condition that results from residual undifferentiated mesenchyme⁶⁹. This condition can be rescued through expression of Lin28b-resistant *let-7*⁶⁹, indicating that Lin28b's effects on nephron progenitors are predominantly the result of its repression of the *let-7* family⁷⁰. The *let-7* family is a documented repressor of cell proliferation in the animal kingdom^{70,71}, and expression of *Lin28b* early in nephrogenesis checks this behavior to allow the nephron progenitor population to repopulate the cap mesenchyme. As nephrons age, their expression of *Lin28b* is gradually reduced until sufficient *let-7* transcripts are generated that they can trigger an anti-*Lin28* feedback mechanism, wherein these same *let-7* miRNA target transcripts of *Lin28b* for destruction^{46,72}. This *let-7*/Lin28 mechanism is not without precedent in the animal kingdom: it is also employed to balance multipotency and differentiation in ESCs⁷³. Mutations that result in over-expression of *Lin28* are common in cancers as it presents a uniquely comprehensive way to down-regulate the *let-7* family of miRNAs⁷⁴, whose reduced expression is a prognostic indicator of proliferative cancers with poor outcomes⁷⁵.

1.6 Enhancers and the cessation of nephrogenesis

1.6.1 Enhancer activity

Orchestrating the lineage-defining transcriptional networks that govern development in metazoans requires the binding of key transcription factors at cis-regulatory enhancers⁷⁶. Enhancers are genomic sequences of DNA that facilitate transcription at gene promoters, and they respond to many of the same regulatory cues as gene promoters: different enhancers can be activated or inactivated by different combinations of molecular inputs, including the binding of

transcription factors and other ligands, or by the complexing of transcriptional machinery⁷⁷. To activate transcription of a gene, enhancers typically⁷⁸ incorporate into a complex with their target promoters. This is made possible by the precise looping patterns employed to compact the genome into the limited three-dimensional space of the nuclear envelope. Chromatin regions that appear distant in a linear view of the genome may in fact loop back into close contact, allowing a promoter and an enhancer to interact physically despite being separated by a megabase or more of sequence^{79–81} (Figure 5).

The mechanisms by which enhancers facilitate transcription are uncertain and likely numerous⁸², but a typical model holds that at the time of transcription a complex is formed between the enhancer, promoter, RNA Polymerase II (RNAPII) enzyme, and any required transcription factors or ligands. Once formed, transcription occurs by drawing the gene through this transcription complex^{83–85}. In this model an enhancer is “active” when it can facilitate the right confluence of these factors, thereby increasing the rate of transcription at its target promoters⁸⁵. Active enhancers can show preference for a specific promoter or type of promoter⁸⁶, but more often they are promiscuous and can interact with multiple promoters within reach⁸⁰. For instance, inverting a region of chromatin to swap enhancers for *Six2* (expressed in nephron progenitors) and *SIX Homeobox 3* (*Six3*; expressed in the lens of the developing eye) results in the reversal of their expression patterns, with *Six3* expressed in nephron progenitors and *Six2* expressed in the lens of the developing eye⁸⁷. Both the frequency with which an enhancer and promoter meet⁸⁸ and their specificity are partially mediated by the genome’s local topology—its physical arrangement in space⁸⁹. When chromatin is arranged into the nucleus, it is bundled into units termed “topologically associated domains” (TADs), often demarcated by genomic regions with a high density of binding sites for CCCTC-Binding Factor (CTCF) and cohesin⁹⁰, the proteins which work in tandem to

form most loop structures through loop extrusion⁹¹⁻⁹⁴. Sequencing methods based on chromatin conformation capture assays such as Hi-C have revealed TAD interactions in detail, and have shown that they are not only consistent between cell types but also conserved between organisms⁷⁹. Promoters and enhancers within the same TAD have a greater likelihood of interacting⁸⁸, and are insulated from features outside of that TAD⁸⁰. The aforementioned *Six2* and *Six3* genes, for instance, are located in different TADs separated by a CTCF-marked TAD boundary in order to prevent cross-talk: the genomic inversion used to swap these enhancers does so by swapping their TADs of residence⁸⁷. Disruptions to TAD boundaries can be severe: F-syndrome, for example, is a limb malformation syndrome in humans that can result from chromatin rearrangements that disrupt a TAD boundary that normally insulates the enhancer for Ephrin Type-A Receptor 4 (EPHA4) from the promoter for Wingless Family Member 6 (WNT6), allowing the two to interact ectopically⁹⁵. But current data suggest that for most genes even severe disruptions to TAD boundaries have little effect on their transcription, which indicates that other mechanisms for specificity and activity of enhancers are usually at play⁹⁶.

Other mechanisms influencing enhancer activity are likely comprised of smaller, sub-TAD scale architectural features of the chromatin environment including the formation, movement, and alteration of insulated loops of chromatin. Architectural changes such as these could allow for more minute and malleable changes to enhancer/promoter interactions, and are thought to be essential for fine-grained modulation of enhancer activity⁸⁵. The complex and unique dynamics that may govern an enhancer's activity are illustrated by the enhancer-dependent expression patterns of *β -globulin* in differentiating mouse erythrocytes: *β -globulin* expression increases as erythrocytes differentiate, and this increase occurs while the chromatin loop containing the *β -globulin* locus is physically relocated from the nuclear periphery to the nuclear interior, possibly

to compensate for the decreasing availability of transcriptional machinery in the nuclear periphery as cells mature. For this enhancer, activity requires not only joining the transcription complex but also following the gene's chromatin loop to its new compartment⁹⁷. Described in this level of detail, an enhancer/promoter mechanism can reveal much about how changes in the cellular environment result in changes in enhancer activity, but these details are not easily generalized across enhancers.

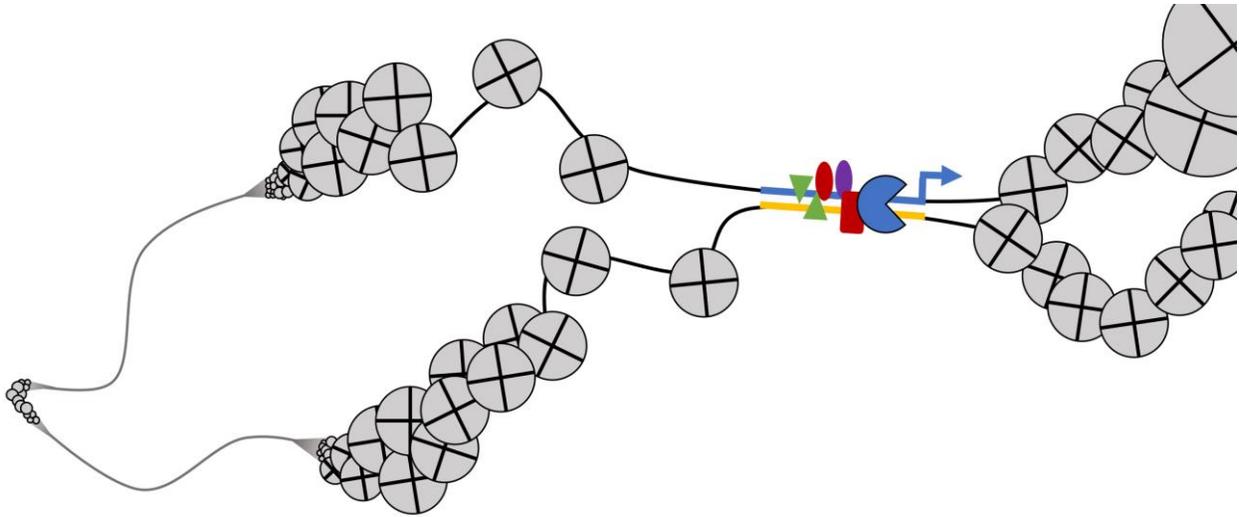


Figure 5. Enhancer-promoter interactions.

Enhancers promote gene transcription at promoters that are up to a megabase away, which is possible due to the formation of loops: enhancer sequences (yellow) can make physical contact with a target promoter (blue), facilitating the formation of the transcription complex including RNAPII and the requisite ligands and transcription factors.

1.6.2 Enhancer activity and chromatin accessibility

In addition to chromatin architecture, enhancer activity is also dependent on the chromatin itself. Chromatin fibers comprise DNA wrapped around nucleosomes: octameric, cylindrical protein structures approximately 5.5nm high and 11nm across, assembled from histones and wrapped in 146bp of DNA^{98,99}. When tightly wound and compacted into condensed chromatin, DNA is inaccessible to most binding proteins that may cause damage or induce errant activity. This prevents activity outside of specific genomic regions, and is a crucial means by which the cell implements and maintains cell lineage and identity¹⁰⁰. In the average fully differentiated human cell, between 97 and 98% of total chromatin is condensed in this fashion¹⁰¹. Permissive chromatin, by comparison, maintains its nucleosomal associations but is less compacted, exposing its

nucleosomal proteins to competition from DNA binding proteins like transcription factors, RNA Polymerase II (RNAPII), and chromatin modifiers such as Barrier to Autointegration Factor 1 (BAF)¹⁰². Condensed chromatin can be rendered permissible through the binding of pioneering factors, a unique class of transcription factor capable of binding to DNA within condensed chromatin and facilitating the subsequent binding of additional proteins and chromatin remodelers, acting as a “wedge” to reduce the activation energy required for subsequent DNA/protein interactions¹⁰³. Permissible chromatin becomes “accessible” when the combined efforts of incoming proteins successfully dislodge histones from the DNA molecule¹⁰⁴, and enhancer sequences frequently contain multiple transcription factor binding sites to facilitate this by increasing the number of competitive proteins^{105,106}. These accessible regions experience either the persistent binding of non-histone proteins or the rapid turnover between bound proteins and histones^{100,101,107}. In total, 5% of detected accessible chromatin regions in humans fall within 2.5kb of a transcription start site and are generally consistent among most cell types, while the remaining 95% are typically cell type-specific and fall in the intronic and intergenic regions characteristic of enhancers¹⁰¹.

The rates at which proteins are able to dislodge nucleosomes—and thus maintain chromatin accessibility—can be altered by changing the nucleosome’s affinity for the DNA strand. To modulate activity, the chromatin surrounding an enhancer is frequently altered to increase or decrease the activation energy required to dissociate the nucleosome, using histone modifications or variants that can have a variety of effects¹⁰⁸. In nucleosomes near promoters, enhancers, and gene bodies, for instance, histone 3 is frequently replaced with histone variant H3.3, and this is thought to impose more accessibility by increasing the rate at which the nucleosome is disassembled, leading to increased activity^{109–111}. Modifications to the N-terminal tails of histones can also affect

local activity levels⁸², and patterns of their usage have proven consistent enough that they are used to identify promoters and enhancers genome-wide using chromatin immunoprecipitation sequencing (ChIP-seq) to locate certain characteristic histone marks¹¹². Enhancer activity levels, for instance, can be approximated based on specific histone modifications: enhancers marked with H3K4me1 typically exhibit neutral or intermediate activity unless also marked with H3K27ac or H3K27me3, in which case they are likely active or poised (not condensed, but resistant to competitive binding), respectively⁸⁵. However, identifying enhancers and assessing their activity using ChIP-seq typically requires between 100,000 and 500,000 cells, making the technique difficult to apply in cases with limited available material¹¹³. Further, ChIP-seq for the active H3K27ac histone mark in humans and mice identifies active enhancers with 41% and 48% success rates, respectively, and only 12% of mouse regions marked “active” in this way demonstrate activity in a generalized enhancer assay⁸⁵. The necessity of these modifications is also uncertain⁸²: despite its relative ubiquity, the H3K27ac modification can be effectively blocked in mouse ESCs by mutating histone H3.3 lysine 27 with arginine, with little effect on gene transcription¹¹⁴. This further supports a model of enhancer activity as a product of the nuances of the local chromatin environment rather than the effect of a singular feature, mark, or complex⁸⁵. The cumulative effects of these chromatin features are reflected in chromatin accessibility, which can be measured using assay for transposase-accessible chromatin sequencing (ATAC-seq)¹⁰⁰.

1.6.3 Assay for transposase-accessible chromatin sequencing (ATAC-seq)

Activity levels of enhancers and promoters are a consequence of diverse chromatin characteristics ranging from TAD boundaries to architectural features and histone modifications, many of which are beyond the reach of a singular genome-wide assay. However their

consequences—rates of transcription factor binding, chromatin remodeling, nucleosome dissolution, etc.—are all reflected in changes to chromatin accessibility¹¹⁵. Assay for transposase-accessible chromatin sequencing (ATAC-seq) provides a means with which this accessibility can be measured genome-wide, and it can be used to approximate the activity levels of regulatory features such as enhancers or promoters^{102,115,116}. ATAC-seq uses a hyperactivated Tn5 transposase enzyme designed to digest DNA for next-generation sequencing libraries by cutting double-stranded DNA and inserting sequencing primers irrespective of sequence¹¹⁷, though with some sequence bias¹¹⁸. By subjecting the intact nuclear envelope to a concentration of Tn5 enzymes, regions of accessible and permissible chromatin are exposed to a higher incidence of transposition events than regions of condensed chromatin due to the protection the latter offers from enzymatic access. Accessible chromatin regions are therefore sequenced more frequently, and are thus identifiable as genomic locations with significantly increased read depth¹¹⁵ (as calculated by enrichment algorithms such as the Model-based Analysis of ChIP-seq; MACS2)¹¹⁹. Rather than requiring hundreds of thousands or millions of cells, ATAC-seq protocols have been performed with samples numbering in the thousands or even hundreds of cells, allowing for analysis of considerably smaller biological samples¹¹⁶.

Paired-end sequencing of ATAC-seq reads (sequencing both the beginning and end of the read to reveal its length) reveals a great deal of information about the local chromatin environment beyond accessibility (Figure 6). For instance, the lengths of these reads can be used to discern sequencing library quality as well as to describe the nucleosomal positioning. Nucleosomes provide protection from the Tn5 enzyme, but this protection is incomplete even when fully assembled: protection is greatest at the nucleosome's dyad position (the central position of the associated DNA strand)⁹⁹, and is reduced with distance from that point. Crucially, any DNA

fragment sequenced during ATAC-seq must result from two separate transposition events on the same DNA molecule in order for the fragment to be sequenced, referred to as “double cut” cleavage¹⁰⁰. Most cleavage events occur in regions of low protection, such as those directly between neighboring nucleosomes (the furthest point from either nucleosomal dyad), or in accessible regions where the nucleosome is frequently dissociated, such as a transcription start site. Consequently, the majority of ATAC-seq reads are less than 146bp in length (“sub-nucleosomal”), resulting from pairs of transposition events taking place on the same DNA strand and in the same accessible region. Longer reads are less frequent and typically result from insertion events that occur in adjacent accessible regions; given the abundance of nucleosomes in the genome the majority of reads spanning adjacent accessible regions originate from spans of nucleosomes. This pattern provides a useful quality control metric for ATAC-seq sequencing libraries as a successful ATAC-seq transposition reaction should yield read lengths that exhibit a characteristic profile: the majority of reads should be sub-nucleosomal in length, and an abundance of larger reads should be in multiples of 146bp, the distance between nucleosomes¹¹⁵. Further, read length distributions along the genome can be used to identify nucleosomal positions: few transposition events occur at or near the nucleosomal dyad, an abundance of short reads will occur on either side of the nucleosome, and longer reads should span the nucleosome rather than start or end on or near the dyad. The software package NucleoATAC leverages these patterns in ATAC-seq read alignment with the expected periodicity of nucleosomes positioned every 146bp to calculate nucleosomal occupancy, ultimately predicting nucleosome positions and identifying nucleosome-free regions (NFRs) that may result from frequent nucleosomal dissolution¹²⁰.

ATAC-seq can also reveal information about bound proteins. Non-nucleosomal bound proteins provide protection from the Tn5 enzyme during transposition, though considerably less

than nucleosomes given their shorter residence times and smaller genomic footprint^{100,121}. The protection they do afford, however, can result in a detectable transcription factor footprint: binding proteins with sufficiently long residence times can cause a broad peak in accessibility with widespread transposition events, save for a narrow region of protection at the binding site itself. The size of the transposase enzyme and its sequence biases can make identifying such footprints difficult using ATAC-seq data, but the recently published HINT-ATAC footprinting algorithm is able to correct for these biases and search for these footprints, then use motif matching methods to search databases of known transcription factor binding motifs to identify the transcription factors likely to have made these footprints¹¹⁸.

The ATAC-seq assay provides an appreciable amount of information about the chromatin environment, which in turn can describe in some detail enhancers and their accessibility genome-wide. These include both known and novel enhancers, as regions of activity need not be attributable to a pre-defined annotated feature to be detected. ATAC-seq measurements made from transgenic nephron progenitor cells suggest that extensive chromatin changes occur between early and late nephrogenesis, including changes in activity at or near key regulatory features¹²². By leveraging chromatin accessibility data and identifying chromatin changes between early and late nephrogenesis, it should be possible to locate some of the features either responding to or accumulating these differences, and their location with regard to the chromatin environment could hint at the promoters with which they may interact.

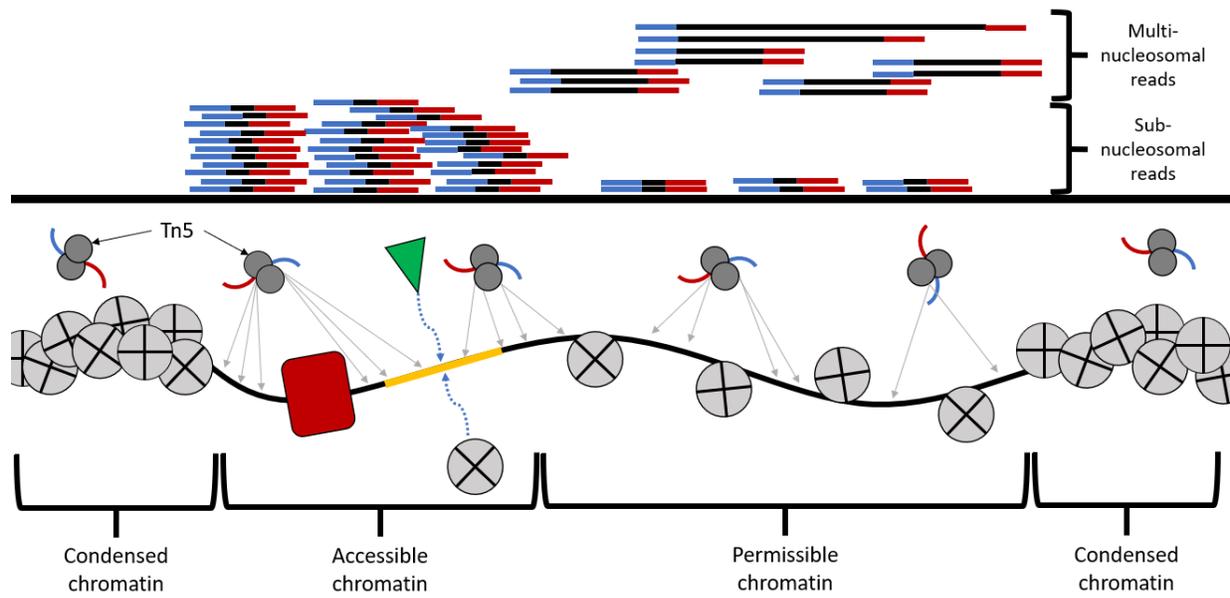


Figure 6. ATAC-seq measures chromatin accessibility.

DNA wrapped around nucleosomes (grey circles) is predominantly found in one of three primary states: condensed, permissible, and accessible. In condensed chromatin, the nucleosomes are packed into a dense structure that is largely inaccessible to DNA binding proteins, as well as Tn5 transposase. Permissible chromatin is chromatin which retains its association with nucleosomes, but is not densely packed, leaving DNA binding proteins free to compete with nucleosomes for binding. Tn5 is more likely to access regions DNA furthest from the center of an attached nucleosome. Accessible chromatin is chromatin with nucleosomes that are frequently dissociated, either due to the presence of DNA binding proteins with high binding affinity and long residence times (red rectangle), or due to DNA binding proteins with sufficient binding affinity to occasionally dislodge bound nucleosomes (green triangle). Tn5 transposase is most capable of inserting primers in these regions. Bound proteins do provide protection from Tn5, but their small footprints and low residence times typically allow for greater signal than is permitted by bound nucleosomes. An abundance of sub-nucleosomal (<146bp) read lengths in ATAC-seq data typically indicate open, accessible regions, while longer fragments usually result from Tn5 insertions made between adjacent nucleosomes.

1.6.4 Enhancers in nephron progenitors

The array of regulatory signals that induce nephron progenitor predilection to survive, proliferate, and differentiate do so largely through their influence on key enhancers in the genome,

helping to drive or repress important down-stream signals^{33,87}. Expression of the multipotency-preserving transcription factor *Six2*, for instance, is associated with multiple enhancer sequences¹²³, including one enhancer 60kb upstream that is bound by both *Six2* and β -catenin³³. *Six2* binds its own enhancer and promoter to form an autoregulatory loop¹²⁴, and β -catenin is the key transducer of canonical Wnt signaling: Wnt signals (Wnt9b in this case) stabilize β -catenin levels by preventing its degradation in the cytoplasm, allowing it to enter into regulatory complexes throughout the genome¹²⁵. *Six2* and β -catenin, in turn, can be found binding enhancers for *Bmp7*, *Wnt4*, *Fgf8*, and *Gdnf*, among others³³. O'Brien et. al (2018) performed chromatin immunoprecipitation sequencing (ChIP-seq) in nephron progenitors to show that many nephrogenesis-regulating growth factors bind to the genome in complexes which the authors describe as “regulatory hot spots,” suggesting that activity levels in these hot spots are responsive to several transcription factors. These multi-transcription factor complexes can be found bound to enhancers for several genes that are expressed in nephron progenitors and regulate their differentiation and renewal, including *Six2*, *Fgf9*, and *Wnt4*, as well as enhancers for genes that mark different cell lineages and are not expressed in nephron progenitors, including *Foxd1* and *Wnt11*. Thus these complexes appear to both activate and repress gene expression, and do so in response to several transcription factor inputs⁸⁷. As evident from *Six2*'s activation and repression in response to changing concentrations of Wnt9b/ β -catenin at the same enhancer, their responses can also be multifaceted and nuanced⁴¹. Complex regulatory networks in the nephron progenitor cell allow it to react to the multifactorial inputs of its surrounding niche⁸⁷, but there is also evidence that intrinsic and age-dependent changes in the nephron progenitor cell can be traced to changes in their regulatory features¹²². How these changes accumulate is uncertain, but measuring changes

in chromatin accessibility over the course of development may reveal known and novel regulatory enhancers that are influenced by the nephron progenitor's age¹⁰⁰.

1.7 Enhancers of microRNA

As components of the nephron progenitor's transcriptional network, enhancers are a means by which the nephron progenitor cell can dynamically respond to inputs from its surroundings to modulate differentiation and proliferation. Given the fundamental dependence of an enhancer's activity on its chromatin environment^{100,121}, and the changes observed in this chromatin environment in the nephron progenitor cell over the course of development¹²², many of the age-dependent changes that differentiate an "early" and "late" nephron progenitor likely coincide with differences in enhancer activity. Chromatin accessibility data from ATAC-seq allows for a genome-wide and unbiased approximation of chromatin accessibility that is agnostic of existing genomic annotations, making it possible to identify and quantify accessibility around novel features as well as known¹¹⁵.

Enhancers play important roles in the nephron progenitor's development changes over time. Expression of miRNA is an important component as well, resisting apoptosis by targeting *Bim* transcripts⁶³ and modulating cell cycle and proliferation through the *miR-17~92*⁶⁷ cluster. Crucially, age-dependent expression of the *let-7* miRNA family determines the timing of the final wave of nephron differentiation that depletes the nephron progenitor population and ceases nephrogenesis⁴⁶. Instances of enhancer-regulated miRNA have been noted in other systems, in particular within cell type-dependent activities of super enhancers. Super enhancers are clusters of highly active enhancers that reside in domains with multiple genes that influence cell identity, and

groups of these enhancers have been shown to cooperate in promoting miRNA expression through recruitment of Drosha (to promote pri-miRNA processing)¹²⁶. Having identified miRNA with age-dependent changes in expression in the nephron progenitor, identifying enhancers with similar age-dependent activity patterns in the same chromatin environment could tie these miRNA into the regulatory networks of the nephron progenitor. Doing so using ATAC-seq could, in turn, reveal important facets of the chromatin environment of the enhancer (and miRNA promoter, if known) including locations of transcription factor binding sites and instances of nucleosomal rearrangement or dissolution. Given the variety of roles that miRNA appear to fill in the developing nephron progenitor and the importance of enhancers when conferring behavioral changes responsive to the nephron progenitor age and niche, the existence of enhancer-driven miRNA expression seems likely, and could be screened from the appropriate data sets.

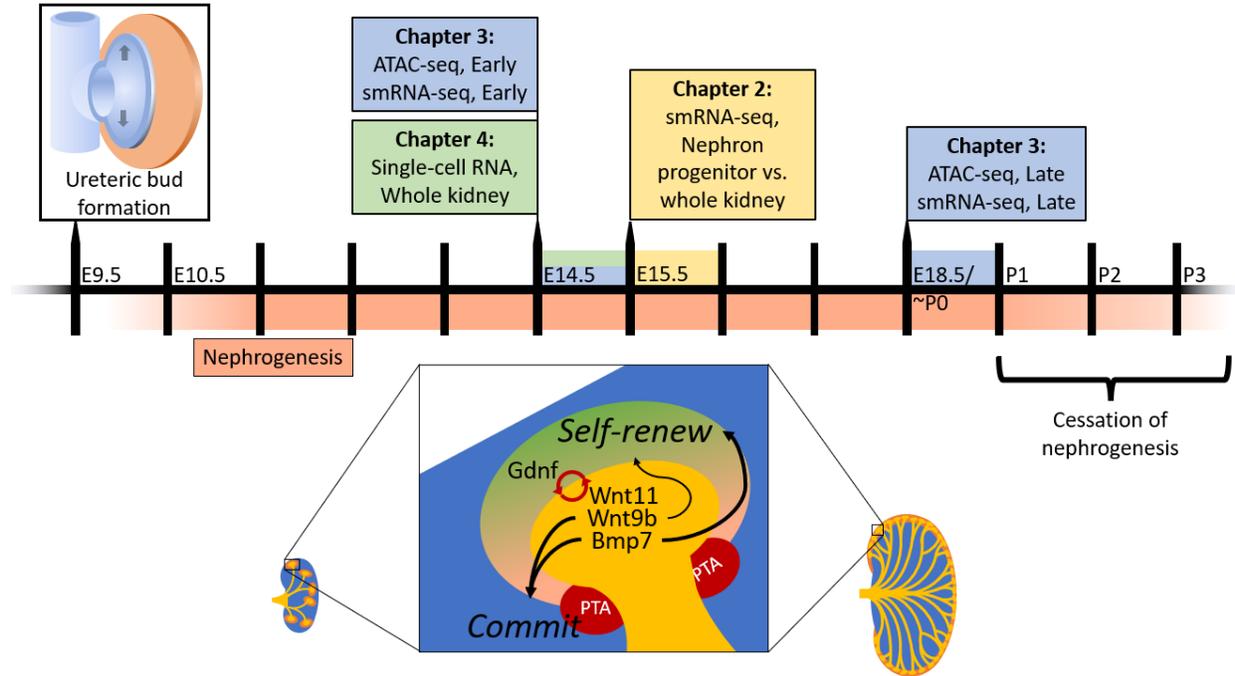


Figure 7. Periods of nephrogenesis measured.

Mouse nephrogenesis begins after E9.5 following the formation of the ureteric bud and proceeds through birth (P0) for up to three days (P3). The publications featured in this document sampled kidney development at different stages between E14.5 and P0. Chapter 2.0 details smRNA-seq data collected from nephron progenitors and whole-kidney samples at E15.5, chapter 3.0 describes smRNA-seq and ATAC-seq measurements taken from nephron progenitors collected at E14.5 and P0, and chapter 4.0 observes single cell sequencing of a full E14.5 kidney. The reciprocal interactions between the ureteric bud and nephron progenitor population which drive nephrogenesis are ongoing throughout each of these periods measured (lower panel).

2.0 Small non-coding RNA expression in mouse nephrogenic mesenchymal progenitors

2.1 Foreword

This chapter is adapted from a previously published work on which I was a co-first author, and it has been adapted and modified with permission from Scientific Data. For this project, I performed the bioinformatic analysis of the smRNA-seq data, while the majority of sample processing and in situ hybridizations were performed by Yu Leng Phua with aid from Kevin Hong Chen. Yu Leng Phua and I drafted the manuscript. Jacqueline Ho and Dennis Kostka contributed to the experimental design and editing the manuscript.

Citation: Yu Leng Phua*, Andrew Clugston*, Kevin Hong Chen, Dennis Kostka, & Jacqueline Ho (2018), Small non-coding RNA expression in mouse nephrogenic mesenchymal progenitors. Scientific Data, 10.1038/sdata.2018.218. * first co-authors

Both the maintenance and elimination of the nephron progenitor population is dependent on the expression of miRNA: loss of miRNA processing causes the premature depletion of the progenitor population due to apoptosis⁶³, elimination of the *miR-17~92* cluster leads to a reduction in nephron endowment due to cell cycle dysregulation⁶⁷, loss of *miR-210* leads to a sex-dependent reduction in nephron endowment⁶⁴, and obstruction of *let-7* activity leads to a failure to differentiate among mesenchymal cells reminiscent of Wilms tumor⁴⁶. To identify other miRNA of importance to the nephron progenitor cell's identity and lineage, we chose to compare expression of miRNA in nephron progenitors with that of the rest of the kidney. We used small

RNA sequencing (smRNA-seq) to measure expression of miRNA genome-wide in wild-type nephron progenitors, and used two methods to map and quantify miRNA expression in both progenitor and whole-kidney samples. Data processing was done with a custom smRNA-seq pipeline, and one method of quantification made use of the miRDeep2 software package to identify novel miRNA based on characteristic patterns of mature/passenger strand reads in the genome. Our analysis detected 162 differentially expressed miRNA and we identified 49 novel miRNA species, four of which we tested for *in vivo* expression experimentally.

2.2 Summary

MicroRNAs (miRNAs) are small non-coding RNAs that are essential for the regulation of gene expression and play critical roles in human health and disease. Here we present comprehensive miRNA profiling data for mouse nephrogenic mesenchymal progenitors, a population of cells enriched for nephron progenitors that give rise to most cell-types of the nephron, the functional unit of the kidney. We describe miRNA expression in nephrogenic mesenchymal progenitors, with 162 miRNAs differentially expressed in progenitors when compared to whole kidney. We also annotated 49 novel miRNAs in the developing kidney and experimentally validated 4 of them. Our data is available as a public resource, so that it can be integrated into future studies and analyzed in the context of other functional and epigenomic data in kidney development. Specifically, it will be useful in the effort to shed light on molecular mechanisms underlying processes essential for normal kidney development, like nephron progenitor specification, self-renewal, and differentiation.

2.3 Introduction

MicroRNAs (miRNAs) are small, endogenously synthesized non-coding RNAs (~22 nucleotides long) that are critical in a wide variety of biological processes, where they primarily act to fine-tune gene expression at the post-transcriptional and translational level¹²⁷. Biosynthesis of miRNAs begins in the nucleus, where a primary miRNA (pri-miRNA) transcript is first transcribed by RNA polymerase II^{127,128}. Subsequently, the endoribonuclease enzyme Drosha-Dgcr8 complex processes pri-miRNAs into individual hairpin-shaped stem loop precursor miRNAs (pre-miRNAs), which are then exported into the cytoplasm via Exportin5¹²⁹⁻¹³¹. In the final steps of miRNA biogenesis, Dicer processes pre-miRNA into mature miRNA, with the miRNA loaded into the RNA-induced silencing complex (RISC)¹³²⁻¹³⁴. mRNA 3'UTR target recognition by miRNA is primarily dependent on the miRNA seed sequence located within nucleotide position¹²⁸⁻¹³⁴ of the miRNA, and complementary Watson-Crick base pairing between the miRNA and mRNA allows the RISC complex to catalyze the process of mRNA degradation and/or translational inhibition^{135,136}.

Nephron progenitors are multipotent cells that undergo a mesenchymal to epithelial transition to subsequently differentiate into glomerular podocytes, proximal tubules, loops of Henle and distal tubule in the developing kidney⁸. They also self-renew throughout kidney development, and the number of nephron progenitors is one of the primary determinants of nephron endowment at birth¹³⁷. Since no new nephrons are formed postnatally in humans, low nephron number increases an individual's risk to develop chronic kidney disease in adulthood¹³⁸. While the role of many protein-coding genes in nephron progenitor self-renewal, maintenance and differentiation is well established, the precise role of small non-coding RNAs remains ill-defined. There are several lines of evidence that point to the importance of miRNAs in kidney development

and disease. Mutations in DROSHA and DICER1 have been identified and implicated in several human kidney disorders, including pediatric Wilms Tumor and multilocular cystic renal tumours¹³⁹. Conditional ablation of Dicer-dependent miRNAs in mouse nephron progenitor or ureteric bud lineages during renal development resulted in severe renal hypodysplasia (small kidneys) and collecting duct cyst formation, due to aberrant progenitor apoptosis and attenuated cilium length respectively^{63,140}. Also, conditional ablation of Dicer or Drosha in glomerular podocytes and renal stroma results in early postnatal death and a wide variety of renal anomalies including podocyte foot process effacement, marked proteinuria, and glomerular aneurysms^{141–145}.

However, more precise characterization of the role of small RNAs in nephron progenitors during kidney development has been hampered by the lack of comprehensive miRNA expression datasets in this context. Given the biomedical relevance of this system, we have performed high throughput small RNA Sequencing (sRNA-Seq) in three biological replicates of embryonic day 15.5 (E15.5; Figure 7) nephrogenic mesenchymal cells enriched for nephron progenitors and whole kidney (Figure 8). Using an adjusted p-value cutoff of 0.05, we identified a total of 162 miRNAs (5p and 3p strand inclusive) out of 792 detectable miRNAs to be differentially expressed in this population when compared to whole kidney. Among the top differentially expressed miRNAs are members of the epithelial-specific miR-200 family, consisting of miRs-200a, 200b, 200c, 141 and 429¹⁴⁶. Levels of miR-200 family miRNAs¹⁴⁶ were significantly lower in nephron progenitors, as might be predicted given that nephron progenitors undergo a mesenchymal to epithelial transition upon differentiation. Furthermore, we uncovered 49 novel miRNA species expressed in the developing kidney. Of these miRNAs, 4 were validated via quantitative real-time PCR (qPCR), with 3 via section in situ hybridization. In general, the data resource will be useful for researchers studying miRNA related biology in kidney/nephron development.

2.4 Methods

2.4.1 Nephron progenitor isolation and total RNA preparation

Nephrogenic mesenchymal cells enriched for nephron progenitors and whole kidney samples were isolated from 3 litters of E15.5 CD1 mouse embryos (Charles River Laboratories) in accordance to a published protocol using a negative selection approach⁴⁵. Briefly, intact embryonic kidneys were subjected to limited digestion, followed by incubation with a cocktail of monoclonal biotinylated antibodies (eBioscience; CD140a #13-1401-82, CD105 #13-1051-82, Epcam #13-5791-82 and Ter119 #13-5921-82), and magnetic activated cell sorted using Dynabeads MyOne Streptavidin C1 magnetic beads (Thermo; #65001) to deplete unwanted cell types. To minimize undesired gene expression changes, total RNA from nephron progenitors were immediately processed in QIAzol Lysis Reagent (Qiagen; #79306) and purified using a miRNeasy Micro Kit (Qiagen; #217084). For whole kidney samples, remnant kidneys left from the limited digestion step were homogenized in QIAzol Lysis Reagent, clarified by centrifugation at 13,000rpm for 5 minutes, and subsequently purified using a miRNeasy Mini Kit (Qiagen; #217004). Purified total RNA samples were stored at -80°C until further processing, and freeze thawing of samples was limited to no more than 2 cycles. Quantitative PCR (qPCR) was carried out using standard SYBR Green detection on a BioRad CFX96 Real Time PCR Instrument to determine enrichment of nephron progenitors from the isolation. Primers used include *Six2*, *Cited1* (nephron progenitors), *Pdgfr β* (renal stroma), *VE-Cad* (endothelial), *Calb*, *Epcam* (epithelial tubules) and *Synpo* (podocytes) (Table 1). Statistical analysis was performed using an unpaired Student's t-test, and genes with a p-value of <0.05 were considered statistically significant.

To account for biological heterogeneity, a total of 3 biological replicates were collected from 3 litters of wildtype CD1 embryos. All experimental procedures were performed in accordance with the University of Pittsburgh Institutional Animal Care and Use Committee guidelines, which adheres to the NIH Guide for the Care and Use of Laboratory Animals. sRNA-Seq experimental workflow and quality control standards were carried out in accordance with ENCODE guidelines (<https://www.encodeproject.org/rna-seq/small-rnas/>)

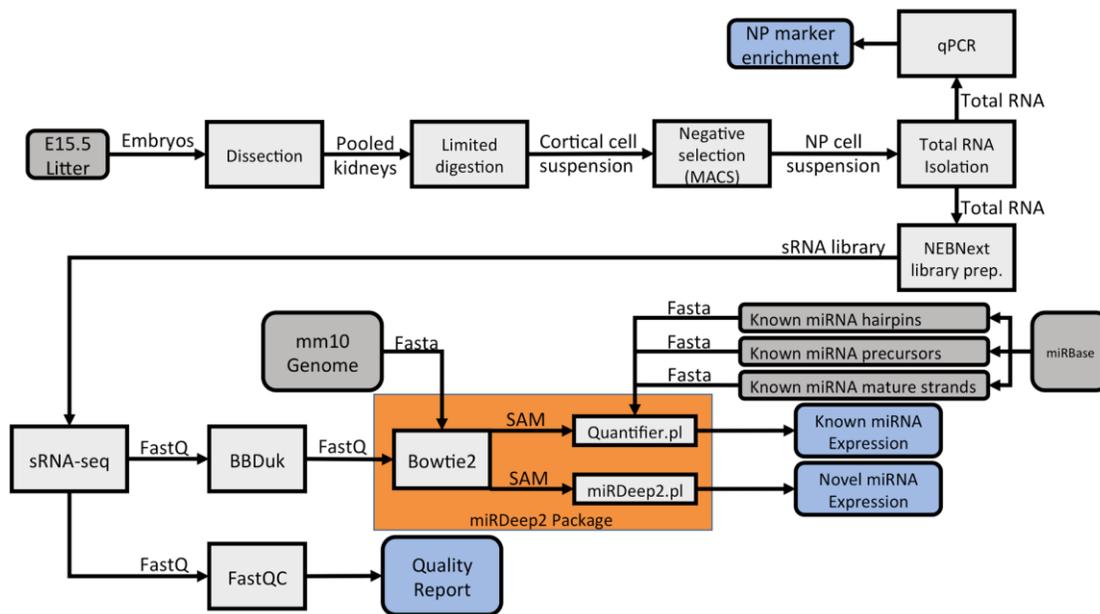


Figure 8. Schematic pipeline illustrating the workflow of nephron progenitor isolation and bioinformatics analysis of the small RNA-Seq dataset.

For nephron progenitor isolation, embryonic day 15.5 (E15.5) CD1 mouse kidneys were subjected to limited digestion, followed by negative cell selection through Magnetic Activated Cell Sorting (MACS). Total RNA from the isolate was extracted and subjected to qPCR analysis for enrichment of nephron progenitor markers. Following verification of nephron progenitor enrichment, total RNA was used as an input for the NEBNextMultiplex Small RNA Library Prep to generate libraries for the sRNA-Seq. The sRNA-Seq dataset was analysed in accordance with the pipeline, with the fastq files first analysed by FastQC to determine the quality of the sequencing reads, followed by adaptor removal using the BBduk package, and finally aligned, quantified and annotated to the mus musculus mm10 genome using the miRDeep2 package.

2.4.2 Small RNA library preparation and sequencing

Because miRNAs can be stably bound to mRNAs¹⁴⁷, total RNA instead of size-selected purified small RNA was used as the initial input for cDNA library synthesis. The NEBNext Multiplex Small RNA Library Prep Set for Illumina (NEB; #E7300S) was used to synthesize barcoded cDNA libraries from 100ng total RNA in accordance with the manufacturer's protocol and size selection performed using the Agencourt AMPure XP beads (Beckman Coulter; #A63881). The purified cDNA libraries were then pooled, normalized and multiplex sequenced at 1X50bp with the use of a NextSeq 500/550 Mid Output v2 Kit (150 cycles) (Illumina; FC-404-2001) on an Illumina NextSeq 550 System at the Rangos Research Center, yielding approximately 18 million reads per sample.

2.4.3 Small RNA-Seq data analysis and quality control

Adapter trimming was performed using the BBmap software (<https://sourceforge.net/projects/bbmap/>) and reads were aligned to the mouse genome (mm10) using Bowtie2¹⁴⁸. miRNA specific analysis was performed with miRDeep2 software¹⁴⁹. Expression of known miRNAs (obtained from miRBase v21¹⁵⁰) was quantified, and miRNAs with more than 10 counts across all sample conditions were subsequently analyzed using DESeq2¹⁵¹ to identify differential expression between nephron progenitor and whole kidney samples. miRNAs with an Benjamini-Hochberg adjusted p-value of 5% or less are reported. Novel miRNAs were annotated with the standard miRDeep2 parameters, and read alignment data was visualized with the Integrative Genomics Viewer software available from the Broad Institute¹⁵².

2.4.4 Data Records

The sRNA-Seq FastQ files were deposited at NCBI Sequence Read Archive (Data Citation 1). This accession contains all 6 FastQ files from the sRNA-Seq. The raw FastQ files were subsequently processed and deposited into NCBI Gene Expression Omnibus (Data Citation 2).

2.4.5 miRNA Quantitative PCR and Locked Nucleic Acid in situ hybridization validation

For the validation of differentially expressed miRNAs, cDNA synthesis was carried out with the use of TaqMan Advanced miRNA cDNA Synthesis Kit (Thermo; #A28007) in accordance with the manufacturer's protocol using equivalent amounts of total RNA. qPCR was carried out with the use of TaqMan Advanced miRNA Assay probes (Thermo; #A25576; 477952_miR, 477970_miR, 477885_miR, 478008_miR) and TaqMan Universal Master Mix II, no UNG (Thermo; #4440040) on a BioRad CFX96 Real Time PCR Instrument using FAM detection.

For the validation of novel miRNAs, cDNA synthesis from total RNA was performed with the use of NCode VILO miRNA cDNA Synthesis Kit (Thermo; #A11193050) in accordance to the manufacturer's protocol using equivalent amounts of total RNA, and qPCR was performed using standard SYBR Green detection on a BioRad CFX96 Real Time PCR Instrument. The universal reverse primer is supplied with the NCode VILO miRNA cDNA Synthesis Kit, and the forward primer sequence is listed in Table 2. Locked Nucleic Acid miRNA in situ hybridization on E15.5 embryonic kidney cryosections was performed as previously described with the use of custom-designed LNA detection probes¹⁴³ (Table 3) (Exiqon).

Table 1. Primer sequences used for quantitative PCR

Gene	Forward (5'-3')	Reverse (5'-3')
<i>Gapdh</i>	AGGTCGGTGTGAACGGATTTG	TGTAGACCATGTAGTTGAGGTCA
<i>Six2</i>	TCTGCTCGGTATCCTTTGGG	TTAAAAATCGGGGTGGTGGTG
<i>Cited1</i>	AACCTTGGAGTGAAGGATCGC	GTAGGAGAGCCTATTGGAGATGT
<i>Ve-Cad</i>	CACTGCTTTGGGAGCCTTC	GGGGCAGCGATTCATTTTTCT
<i>Pdgfrβ</i>	TTCCAGGAGTGATACCAGCTT	AGGGGGCGTGATGACTAGG
<i>Calb</i>	ATGATCAGGATGGCAACGGA	GTTCGGTACAGCTTCCCTCC
<i>Epcam</i>	GCGGCTCAGAGAGACTGTG	CCAAGCATTTAGACGCCAGTTT
<i>Synpo</i>	CCTGCCCGTAACTTCCGTG	GAGCGGCGGTAGGGAAAAG

Table 2. Primer sequences used for novel miRNA quantitative PCR

miRNA	Mature forward sequence (5'-3')	Star forward sequence (5'-3')
<i>U6</i>	GTGCTCGCTTCGGCAGC	NA
<i>chr1_100950463-100950527</i>	GCGGAAAGCTGAAACTTAGAGG	TGAGTCTCCAGCCCTACCCTAA
<i>chr8_66396466-66396533</i>	GCGGTTCTTAGTTGGTGGAGC	GCGCTGATTTGACTGAGAATGT
<i>chr11_108839740-108839780</i>	GCGTGAGAAGAACTTTGAAGAGC	GCGTCCTCAGAGTTAGTGCAGA
<i>chr13_23436971-23437058</i>	GTCTAGGGGTATGATTCTCGCAA	GCGTTTTTATCGTGTTTCCTGT

2.4.6 Code Availability

Software and settings used in the processing of this data are listed as follows:

- 1) BBDuk: from the BBDuk version 35.59 package, adapter references for Illumina indices and the Nextera smRNA library prep set, quality trimming right to left (qtrim=30), kmer trimming right to left (ktrim=r) then left to right (ktrim=l), minimum kmers of 11 (mink=11), hamming distance of 1 bp (hdist=1), minimum final read length of 18bp (minlen=18).
- 2) Bowtie 2: version 2.2.3, parameters of -threads 8, -q, -U, -S, --sensitive-local.

- 3) miRDeep2 version 2.0.0.8, using the `bwa_sam_converter.pl` script to convert Bowtie2's SAM output formats, and the `miRDeep2.pl` script using default parameters for the mouse species (-t Mouse) and miRbase v18 identifiers (-P). Reference data included a pre-built Bowtie2 index for the mm10 genome and miRNA mature and hairpin sequences available in Illumina's mm10 iGenomes package. A FASTA with miRNA "other" sequences were provided from miRbase version 21.
- 4) R: version 3.4.3, data analysis was performed using the R programming language, including packages `mirbase.db`, `DESeq2`, and `GenomicRanges`.

2.5 Results

2.5.1 Cellular enrichment and total RNA quality control

Our overall goal was to quantify small RNA expression in mouse nephron progenitors and whole kidney. We used a previously published protocol⁴⁵ to obtain an enriched fraction of nephron progenitor cells from whole embryonic kidney samples. Confirming that this approach was successful, qPCR analysis of the nephron progenitor-enriched isolate showed high enrichment of nephron progenitor transcripts (*Six2*, *Cited1*) in comparison to whole kidney samples, with evidence of minimal cellular contamination from other key renal cell lineages including the endothelium (*VE-Cad*), epithelial tubules (*Calb*, *Epcam*) and podocytes (*Synpo*) (Figure 9A). An inherent limitation of this protocol⁴⁵ is minor contamination from the renal stroma (*Pdgfr β*), which was indeed detected in our isolate. Although the *Six2*-TGC transgenic mouse³⁴ provides the opportunity for isolating a more pure fraction of nephron progenitor cells, these mice are known

to have renal hypoplasia³², suggesting that their nephron progenitors may not be entirely normal. Taken together, our small-RNA sequencing dataset represents a highly enriched fraction of wild-type nephron progenitors, with minor contamination from renal stroma.

Following cellular isolation and total RNA extraction, the quality of the total RNA was assessed on an Agilent 4200 TapeStation System (Agilent; #G2991AA) using a High Sensitivity RNA ScreenTape (Agilent; #5067-5579) to determine the RNA Integrity Number (RIN) score, which is generated based on the electrophoretic profile of 18S and 28S ribosomal RNA (rRNA)¹⁵³ (Figure 9B). All RNA samples used for the sRNA-Seq had a RIN score of above 9 (

Table 4) and were therefore considered to be good quality intact RNA with minimal degradation. Retention of small RNA in the total RNA isolation was also evident by a distinct broad peak to the right of the lower molecular marker (Figure 9B). To ensure equivalent amounts of total RNA input for the cDNA library construction, the Qubit RNA BR Assay Kit (Thermo; #Q10210) was used for RNA quantitation.

Table 3. Exiqon miRCURY LNA detection probes used for miRNA section in situ hybridisation.

LNA Probe	Product No.	Product sequence 5'-3'
Scramble-miR	699003-300	Scrambled sequence
miR-125b-5p	611756-300	CACAAGTTAGGGTCTCAGGGA
miR-615-3p	615721-300	AAGAGGGAGACCCAGGCTC
miR-10a-5p	YD00612528	ACAAATTCGGATCTACAGGGTA
chr1_100950463-100950527	Custom	CCTCTAAGTTTCAGCTTTC
chr9_113756831-113756880	Custom	GTCCTGTATTGTTATTTTT
chr11_108839740-108839780	Custom	GCTCTCAAAGTTCTTCTCA

Table 4. Sample IDs and RIN score of RNA samples used for small RNA-Sequencing.

Sample	RIN	Fastq file name
MNP1	10	MNP1.fastq
MNP2	10	MNP2.fastq
MNP3	10	MNP3.fastq
MWK1	9.9	MWK1.fastq
MWK2	9.8	MWK2.fastq
MWK3	10	MWK3.fastq

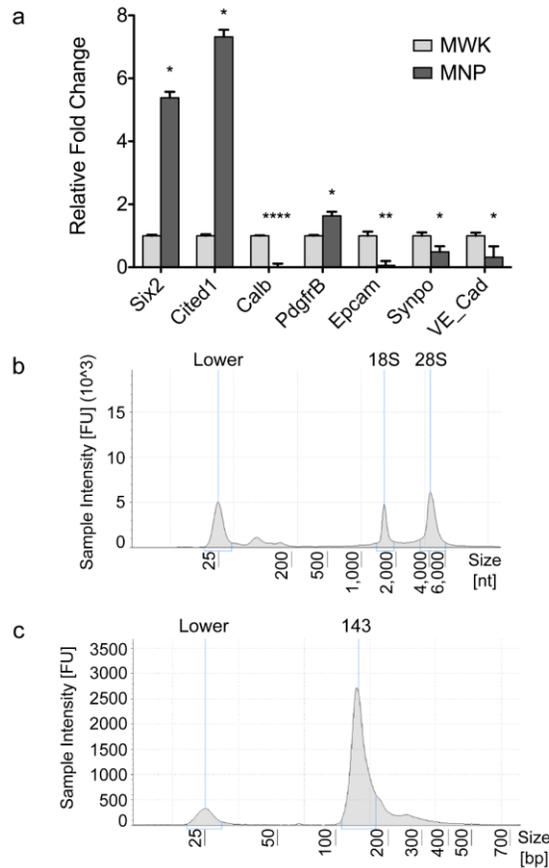


Figure 9. Quality control of the input samples demonstrated an enrichment of nephron progenitors in the cellular isolate with good quality RNA, and successful cDNA library synthesis.

(a) qPCR analysis showed that the cells in the isolate are enriched for nephron progenitors (*Six2*, *Cited1*) with minimal cellular contamination from ureteric bud (*Calb*), podocytes (*Sympo*), endothelial (*VE-Cad*) and epithelial cells (*Epcam*), but minor contamination from the renal stroma lineage (*Pdgfrβ*). (b) TapeStation analysis of the total RNA showed good quality RNA traces with distinct 18S and 28S rRNA peaks with retention of small RNAs as evident by the broad peak to the right of the lower molecular marker. (c) TapeStation analysis showed that the purified cDNA libraries exhibited a distinct peak at ~140-150 nucleotides (adaptor ligated small RNA products), indicating that both the cDNA library construction and cleanup are successful. N=3, * p-value <0.05, ** p-value <0.01, **** p-value <0.0001.

2.5.2 Quality control of the cDNA libraries

cDNA libraries were assayed on an Agilent 4200 TapeStation System (Agilent; #G2991AA) using a High Sensitivity D1000 ScreenTape (Agilent; #5067-5584) and were verified to exhibit a distinct peak at ~140-150 nucleotides that corresponds to adapter-ligated constructs derived from small RNAs (Figure 9C). Moreover, TapeStation analysis verified that the final cDNA libraries for sequencing were depleted of large molecular weight products that one would expect be indicative of larger RNA molecules like messenger or ribosomal RNA.

2.5.3 Quality control of sRNA-Seq data

Purified cDNA libraries for all samples were normalized and multiplex sequenced using a single flow cell on the Illumina NextSeq550 System to minimize technical variability from the sequencing. Following sequencing, reads for each sample was evaluated using the FastQC software (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and all sequenced libraries had an average quality score of above 30, indicating high quality base of the sequenced product (Figure 10A). Percentage of aligned reads was consistently above 70%, with 13 million reads mapped per sample on average. For miRDeep2 analysis, aligned reads shorter than 18bp were discarded in keeping with the software's requirements¹⁴⁹. Principal component analysis (PCA) revealed a clear distinction between and separation of the nephron progenitor (MNP) and whole kidney (MWK) samples (Figure 10B), as did hierarchical clustering of samples based on their miRNA abundances (Figure 10C). PCA and hierarchical clustering analysis both indicate that samples enriched for nephron progenitors are transcriptionally distinct from whole kidney samples.

We identified a total of 162 miRNAs (5p and 3p strand inclusive) out of 792 detectable miRNAs to be statistically differentially expressed in nephrogenic mesenchymal cells enriched for nephron progenitors relative to whole kidney (Figure 10D), and validated the differential profile of miR-210, miR-125b and miR-30c by qPCR (Figure 11A). Among the top differentially expressed miRNAs are members of the epithelial-specific miR-200 family, consisting of miRs-200a, 200b, 200c, 141 and 429¹⁴⁶. Considering that nephron progenitors are mesenchymal in nature and that they undergo a mesenchymal to epithelial transition upon differentiation³⁴, the under-expression of the miR-200 cluster in nephron progenitors not only serves as an excellent internal validation of the phenotypic characteristics of these cells, but also supports the overall integrity of the miRNA profiling dataset.

We explored the use of alternative bioinformatics approaches and obtained comparable results with choices such as Rsubread for read mapping to the mm10 genome (parameters: maximum allowed for 3 nucleotides mismatch and 1 insertion-deletion)¹⁵⁴, featureCounts which outputs number of reads assigned to genomic features¹⁵⁵ and limma-voom for differential expression statistical analysis¹⁵⁶.

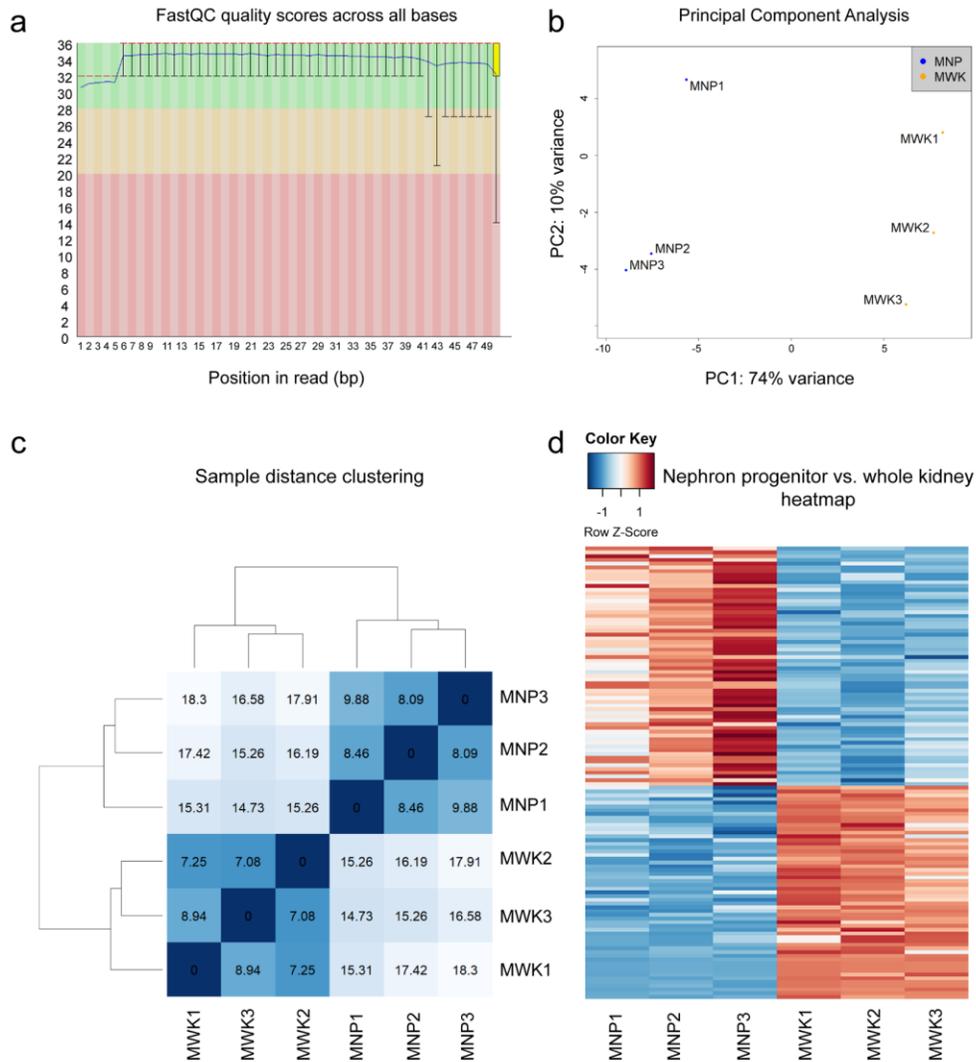


Figure 10. Quality control of the small RNA-Seq dataset showed good sequencing quality reads and congruence of the biological samples.

FastQC analysis of a sequenced library (MNP1) showing that each sequenced base had a mean value score of >30, indicating good quality sequencing. (b) Principal Component Analysis of the sRNA-Seq dataset showed clear separation of the nephron progenitor (MNP) and whole kidney (MWK) samples. (c) Hierarchical clustering correctly grouped samples based on their tissue of origin. (d) Heatmap representation of differentially expressed miRNAs in nephron progenitors vs. whole kidney samples with a false discovery rate of 0.05.

2.5.4 Locked Nucleic Acid section in situ hybridization and qPCR validation of the small RNA-Seq data

The additional novelty of this dataset stems from its inherent ability to identify novel unannotated miRNA transcripts without a priori sequence knowledge. A major challenge in novel miRNA identification is that presumed novel miRNA transcripts are in fact byproducts resulting from degraded mRNA transcripts, and hence not necessarily representative as bona fide novel miRNAs. To overcome this, the miRDeep2 package¹⁴⁹ was used because it includes a series of stringent core analysis modules including the RNAfold tool (<http://rna.tbi.univie.ac.at/>) and the miRDeep2 core algorithm. The RNAfold tool determines and predicts whether the novel miRNA sequence would exhibit a typical hairpin secondary structure, and the miRDeep2 core algorithm evaluates both the miRNA's secondary structure and the expected read mapping signature of its hairpin precursor. The core algorithm also checks that the sequencing reads which span the predicted hairpin structure contains specific recognition sites for subsequent processing by Dicer into mature miRNAs. Novel miRNAs were considered high confidence when both mature and star strands were detected in at least 2 independent samples, in conjunction with sequencing reads spanning across the putative chromosome coordinates independently verified on IGV Viewer. In this manner, we identified a total of 49 novel miRNAs, exemplified by the 20 nucleotides long novel miRNA (GGA AAG CTG AAA CTT AGA GG) depicted with a characteristic hairpin precursor secondary structure located within chr1: 100950463-100950527 (Figure 11B). These novel miRNAs will be submitted for names through miRBase in accordance with their criteria.

As a confirmation that the novel miRNAs were not an artefact from sRNA sequencing, qPCR was performed on 4 candidate novel miRNAs. The qPCR analysis depicted distinct detection of mature (low CT) over star strand (high CT) transcripts (Figure 11C and Table 5),

thereby validating the overall approach in novel miRNA discovery through sRNA-Seq and miRDeep2 identification. Using miRNA Locked Nucleic Acid in situ hybridization for spatial localization of miRNA expression sites, both annotated (Figure 11F-H) (miRs-10a, 125b, 615) and novel miRNAs (Figure 11I-K) (chr1_100950463- 100950527, chr9_113756831-113756880, chr11_108839740-108839780) were found to be expressed in the developing mouse kidney. miR-10a, chr9_113756831-113756880 and chr11_108839740-108839780 exhibited a distinct expression pattern within the nephrogenic zone, encompassing nephron progenitors, renal stroma, and the ureteric bud epithelium. Together, these assays validate both the sRNA-Seq dataset and the approach for novel miRNA discovery in the developing mouse kidney.

Table 5. Quantitative PCR profile of novel miRNA assayed

miRNA	Mature C_T	Star C_T
chr1_100950463-100950527	20.25	29.01
chr8_66396466-66396533	7.04	37.65
chr11_108839740-108839780	26.59	34.48
chr13_23436971-23437058	20.60	30.65

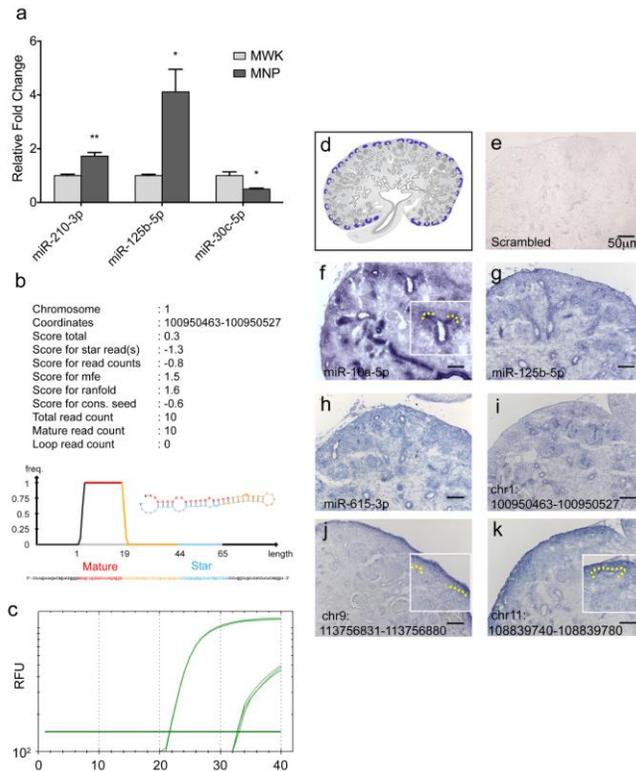


Figure 11. Small RNA-Seq and miRDeep2 analysis revealed novel unannotated miRNAs expressed in the developing kidney.

(a) qPCR analysis validated the differential expression profile of miRs-210, 125b and 30c in nephron progenitor when compared to whole kidney. N=3, * p-value <0.05, ** p-value <0.01. (b) miRDeep2 output for chr1_100950463-10095052, a novel miRNA discovered in our sRNA-Seq data; high frequency (freq) of reads mapped to the mature region of the predicted pre-miRNA structure. (c) qPCR validation of chr1_100950463-100950527 showed a clear distinction between and enrichment of mature (low CT) over star strand (High CT) transcripts. RFU: Relative Fluorescence Unit. (d) For spatial orientation purposes, schematic representative image of nephron progenitors is highlighted in blue. (e-k) Locked Nucleic Acid section in situ hybridisation was used to validate the spatial expression pattern of miRNAs in the developing kidney. miRs-125b, 615 and chr1_100950463-100950527 exhibited a ubiquitous expression pattern in the developing kidney. miR-10a, chr9_113756831-113756880 and chr11_108839740-108839780 appear spatially enriched in the nephrogenic zone, comprising nephron progenitors, renal stroma and the ureteric bud. Schematic image in (d) was designed by Kylie Georgas

(University of Queensland) and are publicly available for use from the GUDMAP database (<http://www.gudmap.org/Schematics>).

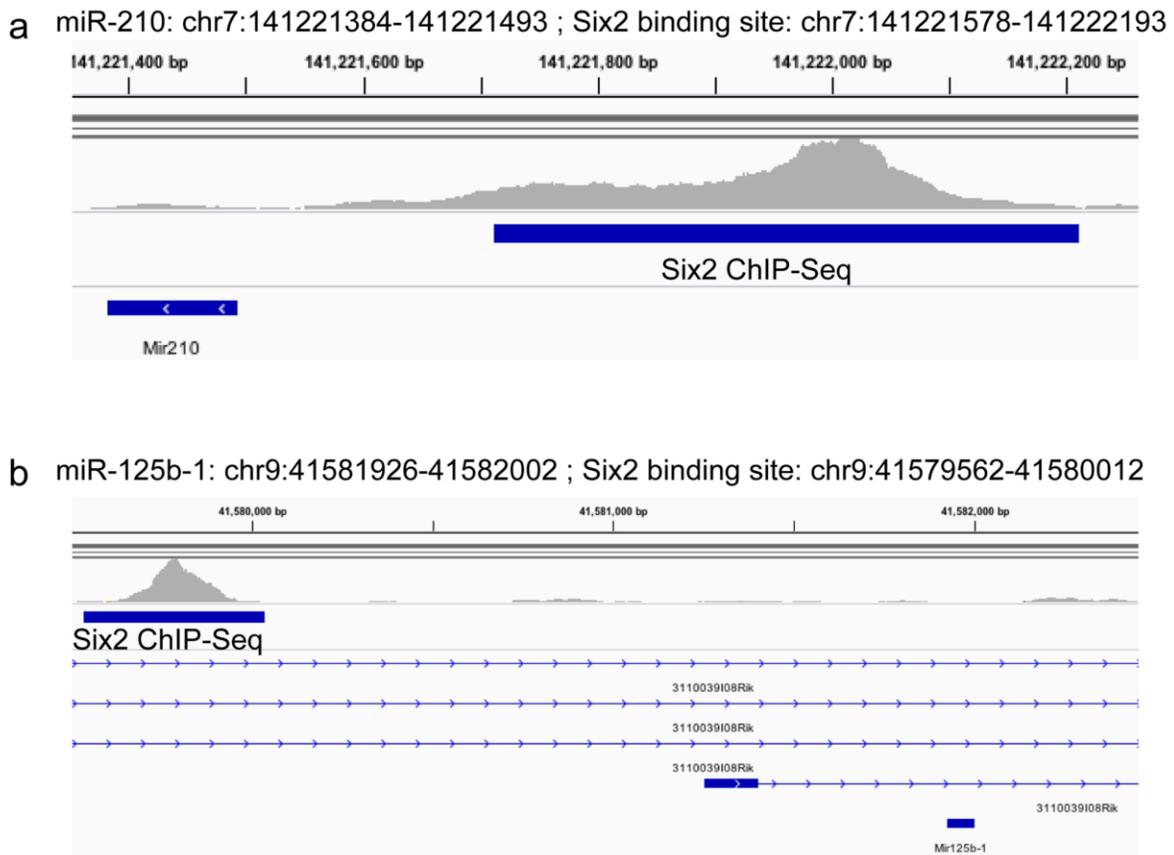


Figure 12. Integrative Genome Viewer visualisation of Six2 binding upstream of expressed miRNAs in nephron progenitors.

By combining a publicly available Six2 ChIP-Seq dataset (SRA Study: SRP064623; GEO: GSE73867) with the current sRNA-Seq results, evidence for Six2 binding was observed in the promoter region of miRs-210 and 125b, miRNAs that are enriched in nephron progenitors.

2.6 Discussion

The sRNA-Seq data presented here represents a comprehensive resource for miRNA expression in developing nephrogenic mesenchymal cells enriched for nephron progenitors and whole kidney. This resource can be utilized to predict regulatory networks downstream of these

expressed miRNAs. For example, this dataset can be integrated with other published RNA-Seq data containing transcriptomic information (mRNA and long non-coding RNA) from nephron progenitors and whole kidney for the analysis of miRNA-mRNA interactions using algorithms such as TargetScan¹⁵⁷ or DIANA-Tools¹⁵⁸. Such analysis would allow one to uncover potential candidate miRNAs and/or downstream target mRNAs for future studies. For example, our sRNA-Seq dataset supports the expression of the highly conserved let-7 family members (let-7a, 7b, 7c, 7d, 7e, 7f, 7g, 7h, 7i, 7j, 7k and miR-98) in the developing kidney with let-7c, 7d, 7e being differentially enriched in nephron progenitors compared to whole kidney during development. The pathogenesis of Wilms tumor has been linked to overexpression of Lin28, a negative regulator of let-7 microprocessing⁶⁹. Wilms tumor is the most common pediatric kidney cancer and arises from a failure of embryonic kidney tissues to terminally differentiate. Thus, de-repression of let-7 downstream targets including Hmga2^{159,160}, Ras¹⁶¹ and Myc¹⁶² oncogenes may be implicated in the malignant transformation of nephron progenitors in Wilms Tumour⁶⁹. This integrative analysis approach allows for future hypothesis testing studies to elucidate the biological function of key miRNAs in nephron progenitors.

Finally, the sRNA-Seq dataset can be complemented with published ChIP-Seq datasets to infer potential relationships between expressed miRNAs and cis-regulatory regions in nephron progenitors. The transcription factor Six2 is known to be critical in maintaining self-renewal and has been shown to regulate target genes that are indispensable for maintaining multipotency of these cells during nephrogenesis³⁷. By analyzing available Six2 ChIP-Seq data (SRA Study: SRP064623; GEO: GSE73867)¹⁶³ with our current sRNA-Seq dataset, it is now possible to curate candidate miRNAs that are potentially regulated by Six2. Two examples are Six2 binding within the promoter regions of miRs-210 and 125b, miRNAs that are enriched in nephron progenitors

based on this sRNA-Seq dataset (Figure 12). This is an example of how to use this resource to identify putative Six2-regulated miRNAs that may play important roles in nephron progenitor maintenance and differentiation during kidney development.

3.0 Enhancers of microRNA in the cessation of nephrogenesis

3.1 Foreword

This chapter is adapted from a manuscript where I am the first author. I performed processing steps from sample collection through to ATAC- and smRNA-sequencing (with the exception of smRNA-seq library preparation, which was performed by the Health Sciences Sequencing Core at Children's Hospital of Pittsburgh). I developed data processing workflows for sequencing quality control, data analysis and data visualization. Yu Leng Phua contributed to study design and experimental preparations. Andrew Bodnar performed assays to determine embryo sexes. Debora Malta Cerqueira contributed to study design, and performed experiments in preparation for enhancer activity assays. Kristy Boggs helped implement the ATAC-seq protocol in our laboratory, and Andreas Pfenning made available his laboratory resources and staff for preliminary sequencing of our ATAC-seq libraries for quality control. I drafted the manuscript. Jacqueline Ho and Dennis Kostka contributed to the experimental design and editing of the manuscript.

Citation: Andrew Clugston, Andrew Bodnar, Debora Malta Cerqueira, Yu Leng Phua, Andreas Pfenning, Kristy Boggs, Jacqueline Ho, Dennis Kostka. Changes in miRNA expression and chromatin accessibility in nephron progenitors (2020, unpublished).

With a comprehensive inventory of miRNA expression in nephron progenitors, we next sought to identify which of these miRNA are most likely associated with the nephron progenitor's

aging mechanisms. We performed smRNA-seq of nephron progenitors at two points of development (E14.5 and P0), which revealed 114 individual differentially expressed miRNA. We also sought to tie the expression changes of these miRNA to regulatory enhancers in the genome. We observed chromatin accessibility using ATAC-seq from the same samples of progenitor cells. Data processing for smRNA-seq samples was done with a custom analysis pipeline, and ATAC-seq data was processed using a custom NextFlow¹⁶⁴ pipeline to ensure data met or exceeded published ENCODE quality standards¹⁶⁵. Using TAD annotations from mouse ESCs¹⁶⁶, we then sought enhancers and miRNA which simultaneously change over the course of nephrogenesis and shared the same TAD to identify potential enhancer-dependent miRNA involved in the nephron progenitor aging process.

3.2 Summary

Mammalian nephrons develop from a multipotent and self-renewing population of nephron progenitors during kidney development. Nephron progenitors that exist early in nephrogenesis are known to be transcriptionally distinct from those in later stages of nephrogenesis. We hypothesize that changes in chromatin accessibility and microRNA expression contribute to these distinct transcriptomes. Nephron progenitor cells were isolated from mouse kidneys at embryonic day 14.5 (E14.5) and postnatal day 0 (P0; Figure 7). In parallel, these progenitors were assayed for transposase-accessible chromatin (ATAC-seq) and small RNA expression (smRNA-seq) to quantify miRNA expression. A total of 46,374 regions of accessible chromatin were identified based on the irreproducible discovery rate (FDR = 0.1), 2,103 of which underwent a statistically significant change in read depth between age groups (FDR = 0.1). 1,104 known miRNAs were

detected across all nephron progenitor samples, of which 114 underwent significant changes in expression between the measured time points ($p_{\text{adj.}} \leq 0.05$). Among significantly changing miRNA, most members of the let-7 family see a significant increase, which confirms previously reports that this family sees increased expression in progenitors which coincides with a decrease in expression of *Lin28b*. Changing regions of chromatin and also predicted miRNA gene targets were enriched in pathways affecting cell migration, extracellular matrix interactions, and key developmental signaling pathways such as Notch and TGF- β . Among regions of changing chromatin and miRNA with changing expression, we identified 33 possible miRNA-enhancer dependencies for further study. These data represent the first measurement of miRNA expression changes in nephron progenitors over time, and the first of such measured in parallel with chromatin accessibility.

3.3 Introduction

The functional unit of the kidney is the nephron, which serves to filter waste and maintain normal homeostasis of water, acid-base, and electrolytes in the body. Nephron number varies widely in humans (typically between 200 thousand and greater than one million nephrons²) and is established prior to birth (birth in humans⁴, approximately post-natal day 2-3 in mice¹⁷). Nephron number is determined in part by the population of nephron progenitor cells¹⁶⁷, and because nephrons cannot regenerate after birth decreased nephron endowment is associated with an increased risk of chronic kidney disease and hypertension⁵. During kidney development, one subset of nephron progenitor cells commits to differentiate into early developing nephrons (renal vesicle), and another subset continues to self-renew. In the latter stages of embryonic development,

nephron progenitors increase their propensity to differentiate, which gradually depletes their population and marks the eventual cessation of nephrogenesis^{17,50,168}.

Nephron progenitor differentiation is regulated by a series of coordinated events. Roughly, Bmp7-pSmad1/5/8 signaling induces the initial exit of self-renewing Cited1⁺/Six2⁺ nephron progenitors into a primed Cited1⁻/Six2⁺ state, followed by Wnt9b/ β -catenin induction of differentiation^{21,168}. This is influenced by migration of nephron progenitors relative to Wnt9b secretion from the ureteric bud, prolonged exposure to which results in the differentiation of progenitors into *Wnt4*-expressing renal vesicles^{35,169}. Transcriptional changes in nephron progenitors that are dependent on the stage of embryo development have been shown to be associated with genes and pathways associated with organism development and stem cell aging, specifically including *mTor* and its repressor *hamartin*^{47,50}. The general idea is that these changes may desensitize nephron progenitors to signals for self-renewal and in this way contribute to the cessation of nephrogenesis⁴⁷.

Recent studies have implicated non-coding RNA, such as micro-RNA (miRNA), in regulating nephrogenesis. miRNA are short, non-coding RNA molecules that target messenger-RNA transcripts (mRNA) for impeded translation or degradation, in both cases via the RNA-induced silencing complex (RISC)⁵⁴. Loss of miRNA processing in nephron progenitors results in premature depletion of the nephron progenitor population⁶³, global removal of the hypoxia-responsive *miR-210* causes a significant decrease in nephron endowment in male mice¹⁷⁰, and deletion of the *miR17hg* cluster (miR-17~92) in nephron progenitors impairs proliferation and reduces nephron endowment⁶⁶. Further, the protein Lin28b is a known repressor of the *let-7* family of miRNA⁶⁸, and its expression in nephron progenitors decreases as nephrogenesis proceeds. This reduction coincides with the gradual increase in expression of the *let-7* family, and interestingly

ectopic treatment with *Lin28b* is sufficient to prolong nephrogenesis⁴⁶. This suggests that miRNA in the *let-7* family may play an important role in the timing the cessation of nephrogenesis.

In addition to transcriptional changes, changes in chromatin accessibility have also been observed in cultured nephron progenitors during embryogenesis, suggesting a developmentally-timed opening or closing of gene regulatory sequences like promoters and enhancers¹²². Consistent with this idea, transcription factors that regulate nephron progenitors including *Six2*, *Hoxd11*, *Osr1*, and *Wt1* have been shown to bind enhancer sequences located in “regulatory hot-spots” in the genome⁸⁷, and the multipotent nephron progenitor marker *Six2* co-binds enhancers with β -catenin to drive expression of markers of self-renewal and differentiation³³. Expression of the transcripts for *Six2* and for the renal vesicle marker *Wnt4* are regulated by enhancers as well⁸⁷.

In this study, we sought to identify nephron progenitor enhancers and miRNA that contribute to the cessation of nephrogenesis by observing changes in miRNA expression and chromatin accessibility between embryonic day 14.5 (E14.5) and post-natal day zero (P0; Figure 7). Enhancers that regulate long non-coding RNA expression during nephrogenesis have been identified¹⁷¹; however no enhancers regulating miRNA expression are currently known in nephron progenitors. We therefore produced matched smRNA-seq and ATAC-seq libraries of nephron progenitor cells at the two different time points. In our analysis we identified 2,406 regions of changing chromatin accessibility and 114 miRNAs with changing expression in nephron progenitor cells. Ontological enrichment among changing regions of chromatin and among predicted miRNA gene targets suggest these changes may affect cell migration, extracellular matrix interactions, and key developmental signaling pathways such as Notch and TGF- β . Among regions of changing chromatin and miRNA with changing expression, we identified 37 possible miRNA-enhancer dependencies for further study.

3.4 Methods

3.4.1 Mouse strains

Wildtype CD-1 time-mated pregnant females were obtained from Charles River Laboratories, Inc. (Wilmington, MA, USA, RRID:MGI:5659424) and kidneys were collected from litters at embryonic day E14.5 or P0. All animals were housed in the vivarium at the Rangos Research Center at the UPMC Children's Hospital of Pittsburgh (Pittsburgh, PA, USA) and all animal experiments were carried out in accordance with the policies of the Institutional Animal Care and Use Committee at the University of Pittsburgh. The sex for each embryo or pup was identified by performing PCR on genomic DNA isolated from tail clippings using the following primers for the Y-chromosome gene *Sry*: SryF 5'-GATGATTTGAGTGGAAATGTGAGGTA-3' and SryR 5'-CTTATGTTTATAGGCATGCACCATGTA-3', as previously published¹⁷².

3.4.2 Nephron progenitor isolation

Kidneys were dissected from litters of 8-12 E14.5 embryos or P0 pups, and each litter was considered one sample. One kidney per sample was used for total RNA isolation to be used as a “whole kidney control” in subsequent analyses (see below). A total of three samples were collected at each time point. Cortical cells were then isolated from each sample using magnetic-activated cell sorting (MACS) as previously published¹⁷³. Briefly, kidneys underwent partial digestion with 2mM collagenase A (Roche 11088793001) and 3.5mM pancreatin (Sigma P1625) in phosphate-buffered saline (PBS) for 15 min at 37°C. The digestion reaction was halted using cold 100% fetal bovine serum (FBS), and the cell suspension was collected and resuspended in cold Dulbecco's

phosphate-buffered saline (DPBS) with 1mM phenylmethylsulfonyl fluoride (PMSF) protease inhibitor (Sigma-Aldrich 10837091001). The cells were washed and resuspended in ice-cold isolation buffer consisting of 2% FBS in PBS, then incubated with magnetic beads (Dynabeads FlowComp Flexi kit, ThermoFisher 11061D) biotinylated to antibodies for Integrin alpha 8 (Itga8, Supplemental figure 1A, R&D Systems AF4076) using the DSB-X Biotin Protein Labeling Kit (ThermoFisher D-20655). Nephron progenitors bound to these beads were isolated as previously published¹⁷⁴ and pooled across kidneys for each sample. An aliquot of 50,000 nephron progenitor cells was immediately processed through the chromatin accessibility library preparation protocol (see below), and total RNA was extracted from the remaining cell suspension using the Qiagen miRNeasy Mini Kit (Qiagen 217004).

To confirm enrichment of nephron progenitor cells relative to other cell types, total RNA was isolated from nephron progenitors as well as from whole kidneys (whole kidney control, see above), and quantitative reverse-transcription PCR (qRT-PCR) for the following markers was performed: nephron progenitor markers *Six2* and *Cited1*²², and for *Lhx1*, *Pdgfrβ*, *Pecam* and *Calb* that mark renal vesicle¹⁷⁵, renal stroma⁴⁷, endothelial³⁴, and ureteric bud¹³⁷ cells, respectively; also see Table 1. *Gapdh*¹⁷⁶ was used as housekeeping gene and relative quantification was calculated via the $2^{-\Delta\Delta Ct}$ method¹⁷⁷.

Table 6. RT-PCR primers.

Transcri	Cell type	Forward primer sequence	Reverse primer sequence
<i>Six2</i>	Nephron	GCAGGACTCCATACTCAA	GATACCGAGCAGACCATT
<i>Cited1</i>	Nephron	AACCTTGAGTGAAGGATC	GTAGGAGAGCCTATTGGAG
<i>Lhx1</i>	Renal	CTACATCATAGACGAGAAC	TCATTACTACCACCTTCCTT
<i>Pdgfrβ</i>	Renal	TTCCAGGAGTGATACCAGC	AGGGGGCGTGATGACTAGG
<i>Pecam</i>	Endothelial	ATTGCGGTGGTTGTCATTG	TAACCGTAATGGCTGTTGGC
<i>Calb</i>	Ureteric	ATGATCAGGATGGCAACG	GTTCGGTACAGCTTCCCTCC
<i>Gapdh</i>	Housekeepi	AGGTCGGTGTGAACGGATT	TGTAGACCATGTAGTTGAG

3.4.3 Chromatin accessibility library preparation

Approximately 50,000 nephron progenitor cells were used to generate each sequencing library for the assay for transposase-accessible chromatin (ATAC-seq) using the Nextera DNA Flex Library Prep Kit (Illumina FC-121-1030) with modifications according to a previously published method¹⁷⁸. In brief, cells were suspended in an ice-cold non-ionic lysis buffer consisting of 10mM Tris, 10mM NaCl, 3mM MgCl₂, 1mM PMSF protease inhibitor, and 0.05% Triton X-100 by volume for 10 minutes to lyse cell membranes while leaving nuclear membranes intact. Intact nephron progenitor nuclei were resuspended in transposition buffer containing the transposase enzyme and incubated at 37°C for 30 minutes. Free genomic DNA released by the transposition process was purified using the MinElute PCR Purification kit (Qiagen 28004) and then indexed using forward (i7) and reverse (i5) index primers from the Nextera Index Kit (Illumina FC-121-1011). Index ligation and fragment amplification were achieved using the method's PCR amplification thermal cycling program.

To determine the optimal number of amplification cycles required for the ATAC-seq library, a qPCR side reaction was performed using SsoAdvanced SYBR Supermix (BioRad 1725274) and a 96-well C100 Thermal Cycler (Bio-Rad) to calculate the normalized reporter value (R_n) for each cycle. The cycle number at which the reaction reached one-third of its maximum fluorescence was identified as the optimal number of amplification cycles remaining, and the remaining volume of the indexed ATAC-seq transposition reaction was subjected to that number of additional cycles in the final amplification program. Final libraries were purified using Ampure XP magnetic beads (Beckman Coulter A63881).

3.4.4 Chromatin accessibility sequencing

Paired-end sequencing of the ATAC-seq library was performed on an Illumina NextSeq500 by the Health Sciences Sequencing Core at UPMC Children’s Hospital of Pittsburgh, multiplexed with library concentrations expected to yield 90 million paired end reads per sample. Reads were quality trimmed using TrimGalore¹⁷⁹ (version 0.4.3) in “--paired” mode and with otherwise default settings. Reads were aligned to the mm10 genome¹⁸⁰ using Bowtie2¹⁸¹ (version 2.3.1) with the settings “--local -q -X 2000 --m”. Reads resulting from PCR duplicates of the same DNA fragment were marked using Picard Tools’¹⁸² MarkDuplicates function (version 2.10.9) with default settings. Reads that mapped ambiguously to more than one location in the genome were randomly assigned to one of these locations if there were fewer than four possibilities, and were otherwise eliminated¹⁸³ (using a Python script available at https://github.com/kundajelab/atac_dnase_pipelines,¹⁸⁴ with the settings “--paired-end -k”). Samtools¹⁸⁵ (version 1.3.1) was used filtering out duplicated, unmapped, orphaned, and mitochondrial reads, as well as the sorting and indexing of BAM files. To address differences in library sizes after sequencing and filtration steps, we performed down-sampling and used Samtools to randomly select reads from each library until each had the same number of reads (with the argument “-s *n*”, where *n* is the ratio of reads in the smallest sample library to the reads in the library being sampled).

3.4.5 Identifying accessible chromatin regions

BAM files containing filtered ATAC-seq reads were converted into paired-end BED format using Bedtools’¹⁸⁶ (version 2.26.0) ‘bamtoBED’ function with settings “-bedpe -mat1”.

Regions of accessible chromatin were identified using the Model-based Analysis of ChIP-seq tool (MACS2, version 2.1.1.20160309) broad peak calling algorithm¹¹⁹, with settings “--format BEDPE --g mm -p 0.01 --broad --shift 37 --extsize 75 --keep-dup all --nomodel”. To identify high-confidence accessible regions, pair-wise comparisons (of each combination of replicates from a given time point) were performed and the irreproducible discovery rate (IDR)¹⁸⁷ was calculated using a publicly available Python implementation (GitHub <https://github.com/nboley/idr>)¹⁸⁸ with settings “--input-file-type broadPeak --rank p.value --soft-idr-threshold 0.1 --output-file-type bed”. Specifically, for each time point a concatenated BED file of all accessible regions was submitted through the “--peak-list” argument for each comparison (for instance, when comparing two E14.5 samples, all E14.5 broadPeaks between all three replicates are used). Regions of accessibility that were consistent by IDR in at least two pair-wise replicate comparisons (with a minimum of 20% overlap) were combined into a unified set of high-confidence accessible regions (“IDR regions”).

3.4.6 Changes in accessible chromatin regions

The accessibility of each IDR region within a given sample was quantified by counting the number of transposition events that occur while normalizing for sequence GC-content across samples. This was achieved using a custom script (available at the accompanying software repository, see supplemental data). Changes in this accessibility were then determined by comparing these values across time points using the Limma-voom (version 3.42.2) software package¹⁸⁹. To account for embryo sex, the fraction of female embryos in each sample was included as a cofactor in linear models used for differential accessibility analysis, and it was removed as a batch effect in visualizations using the Limma package’s `removeBatchEffect`

function. Differentially accessible regions of chromatin (DARs) were deemed to be significantly opening (increased accessibility between E14.5 and P0) or closing (decreased accessibility between E14.5 and P0) controlling the false discovery rate at 10%. Functional enrichment analyses of opening and closing genomic regions was determined by submitting respective coordinates to the Genomic Regions Enrichment of Annotations Tool (GREAT, available at <http://great.stanford.edu/public/html/>)¹⁹⁰. Annotations of known and predicted enhancers were retrieved from the FANTOM5 repository¹⁹¹.

3.4.7 Transcription factor footprints

Transcription factor footprints were identified by pooling all BAM files from the same time point, then running the Regulatory Genomics Toolbox's Hmm-based Identification of TF Footprints (HINT) software (version 0.12.3), specifically the footprinting model for ATAC-seq data (available at www.regulatory-genomics.org)¹⁹². This software was run with settings "rgt-hint footprinting --organism=mm10 --atac-seq --paired-end". Footprints were then annotated with known mouse binding motifs from the HOCOMOCO 11 database¹⁹³ using HINT's motif matching function with the settings "rgt-motifanalysis matching --organism=mm10". Footprints with HINT scores below 10 were excluded. Finally, differential activity patterns among motif-matched footprints was measured using the RGT-HINT program's differential function, with settings "rgt-hint differential --organism=mm10 --bc --nc 12 --output-profiles". Significantly changing activity levels were identified based on a p-value cutoff of 0.05 (Friedman-Nemeny method).

3.4.8 Nucleosomal configuration

Aligned reads from the same time point were pooled into a single BAM file, and the NucleoATAC package (version 0.3.4) was used to identify patterns in read lengths suggestive of both bound nucleosomal dyads and of nucleosome-free regions (NFRs)¹²⁰. This software was executed using default settings. Nucleosomes were annotated as the 146bp region centered around each called nucleosome dyad.

3.4.9 Small RNA sequencing and analysis

Total RNA isolated from the nephron progenitor cells remaining after removing the ATAC-seq fraction was submitted to the Health Sciences Sequencing Core at UPMC Children's Hospital of Pittsburgh for library preparation using the QIAseq miRNA library preparation kit (Qiagen 331502). Single-end sequencing of smRNA-seq libraries was performed on an Illumina NextSeq500, and libraries were sequenced to a depth of approximately 50 million single-ended reads per library. Sequenced reads were then aligned to the mouse mm10 genome using the Rsubread package¹⁹⁴ (version 2.0.1), and annotated to known miRNA listed in miRBase (version 22)¹⁹⁵ with the Rsubread package's featureCounts function. Differential expression of known miRNA was measured using the DESeq2 R package (version 1.26.0)¹⁹⁶. The fraction of aligned reads annotated to known miRNA as well as the fraction of female embryos per sample were included as cofactors in the testing model submitted to DESeq2. Differentially expressed miRNA were then identified while controlling the false discovery rate at 5%, and their direction of change (increasing vs. decreasing) was determined based on their fold change value ("increasing" indicating higher expression at P0 compared with E14.5). Enrichment of regulatory pathways

among predicted gene targets of miRNA with significantly increasing and decreasing expression were calculated using DIANA tools miRPath version 3.0⁶².

3.4.10 Screening for regulatory elements affecting miRNA expression

To screen for potential enhancers of miRNA, we annotated miRNA and DARs with the topologically associated domains (TADs) in which they are located based on mm10 annotations from mouse embryonic stem cells, downloaded from the 3D Genome Browser (<http://promoter.bx.psu.edu/hi-c>)¹⁶⁶. Regions of chromatin accessibility and miRNA were considered candidate enhancer—miRNA pairs if they occupied the same TAD and showed consistent changes in accessibility and expression (increasing accessibility with increasing miRNA expression and vice versa). DARs were only considered as possible “regulatory features” if they did not overlap a known promoter or exon end. Potential pairs of miRNA and putative regulatory elements were prioritized if they 1) fell within the same TAD and 2) experienced significant changes between E14.5 and P0 (increasing miRNA expression and increasing IDR region accessibility, and vice versa).

3.5 Results

3.5.1 Nephron progenitor cell populations from E14.5 and P0 kidneys

Nephron progenitors were isolated by positive selection for *Itgα8* expression from litters of mice collected at either E14.5 or P0 (Supplemental figure 1A), and split into fractions for

ATAC-seq and smRNA-seq (Supplemental figure 1B). Significant enrichment for the nephron progenitor markers *Six2* and *Cited1*²², compared to that for markers of ureteric bud (*Calbindin*)¹³⁷, endothelial cells (*Pecam*)³⁴, renal stroma (*Pdgfrβ*)⁴⁷, and renal vesicle (*Lhx1*)¹⁷⁵, was confirmed in all nephron progenitor samples using qPCR (Supplemental figure 1C). Litters used for E14.5 samples ranged from 27-50% female, and litters for P0 samples ranged from 31-75% female; on average E14.5 samples were 38% female, and P0 samples were 49% female (Supplemental figure 1D). These matched nephron progenitor cell populations allowed us to analyze DNA accessibility and small RNA expression across two developmental time points (E14.5 and P0).

3.5.2 Early and late nephron progenitors have distinct chromatin environments

Quality control of ATAC-seq data followed the ENCODE project's ATAC-seq guidelines¹⁹⁷ (Supplemental figure 2A, B), and sample libraries were randomly down-sampled to each contain 31 million paired-end reads. Using the irreproducible discovery rate (IDR)¹⁹⁷ as a criterion for consistently detected “peaks” (see Methods), we identified 46,374 regions of accessible chromatin (with a false discovery rate of 0.1, see Supplemental figure 2C). Principal component analysis (normalized for sex differences, see Methods), showed that samples characterized by accessible regions group by developmental time point (Figure 13A). We next compared chromatin accessibility between E14.5 and P0 and found 2,406 differentially accessible regions (FDR = 0.1), with 1,323 (55%) showing increased accessibility at P0 compared with E14.5. Hierarchical clustering of differentially accessible regions of chromatin (DARs) revealed a clear distinction between chromatin regions that open and close between E14.5 and P0 (Figure 13B).

Enrichment analysis of opening and closing DARs using the Genomic Regions Enrichment of Annotations Tool (GREAT)¹⁹⁰ highlights GO biological processes closely tied to nephron

progenitor development. Opening DARs are enriched for “Positive regulation of Notch Signaling pathway” (GO:0045747, FDR = $2e^{-5}$) and “Negative regulation of ERK1 and ERK2 cascade” (GO:0070373, FDR = $6e^{-4}$, Figure 13C, Supplemental figure 3). Notch signaling promotes nephron progenitor differentiation⁴¹, and ERK1 and ERK2 signaling cascades frequently stimulate entry into the cell cycle¹⁹⁸. Interestingly, regulatory pathways associated with neural development (Forebrain development, Telencephalon development, etc.) are among the most significantly implicated in regions of opening chromatin. A potential reason could be the important roles of Wnt/ β -catenin signaling in differentiation of nephron progenitors, and cortical neural precursor cells¹⁹⁹. Regions with a decrease in accessibility are enriched for “Positive regulation of cell differentiation” (GO: GO:0045597, FDR = $8e^{-8}$), “Regulation of cellular component biogenesis” (GO:0044087, FDR = $8e^{-5}$), and “Stem cell differentiation” (GO:0048863, FDR = $6e^{-4}$) (Supplemental figure 3). These results show that E14.5 and P0 nephron progenitor cells are characterized by distinct chromatin environments, and that many of the differences between them implicate key developmental processes.

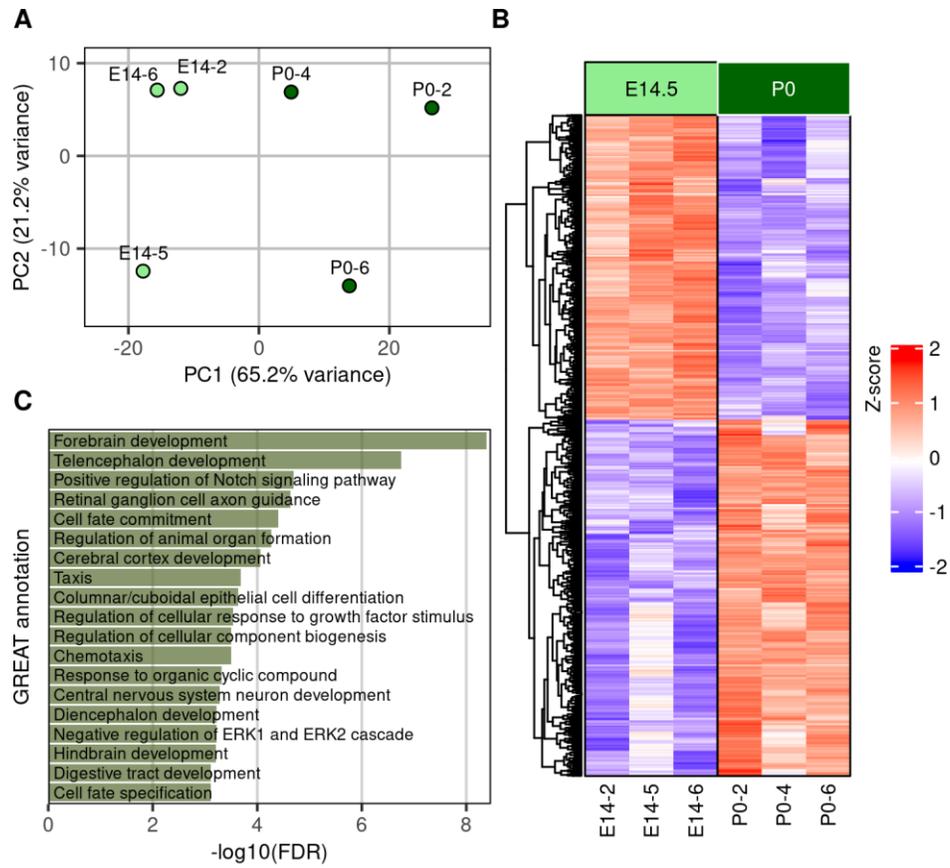


Figure 13. Differences in chromatin accessibility characterize E14.5 vs. P0 nephron progenitor cells.

A) Principal component analysis of ATAC-seq samples highlights developmental time points as a major source of variation. **B)** Heatmap showing significantly changing IDR regions among ATAC-seq samples. E14.5 samples/columns are annotated in light green, and P0 samples/columns are annotated with dark green. **C)** GO Biological Process terms enriched among genomic regions that open between E14.5 and P0 according to GREAT, $FDR \leq 0.001$.

Next, we observed the chromatin environment for select regions of the genome. The *Six2* transcription factor is critical for maintaining the multipotency of nephron progenitor cells³⁷. In all samples, we observed accessible chromatin surrounding the *Six2* promoter and a published *Six2* enhancer ~50kb upstream of the *Six2* transcript⁸⁷ (Figure 14A, B, blue-shaded areas). Next, *Gdnf* codes for a protein that is secreted by nephron progenitors to promote ureteric branching^{16,201}. *Gdnf* expression in nephron progenitors is known to increase as they differentiate, and *Gdnf* has an enhancer 113kb upstream that is responsive to Wnt signaling over this same period³³. We observe that both this enhancer and the promoter of *Gdnf* fall within regions of accessible chromatin (Figure 14D,E), and we note that while the *Gdnf* promoter remains consistently accessible, this enhancer significantly increases in accessibility from E14.5 to P0. Increasing accessibility at the *Gdnf* enhancer could be one means by which progenitors are primed for differentiation.

Our data reveal several novel DARs that may contain relevant gene-regulatory features. One specific region on chromosome 17 is conserved in mammals and is significantly more accessible at P0 compared with E14.5. It also appears to undergo nucleosomal configuration changes in the form of a nucleosome-free region (NFR) that appears adjacent to its conserved sequence (Figure 14C). A second DAR on chromosome 4 is highly conserved in mammals and undergoes a significant reduction in accessibility between E14.5 and P0 (Figure 14F). Whether these regions have a role in nephron progenitor differentiation, or nephrogenesis in general, is currently unknown. These examples show that our data confirms known regulatory elements with a role in affecting nephron progenitor differentiation, and that it highlights compelling genomic regions of as-of-yet unknown function.

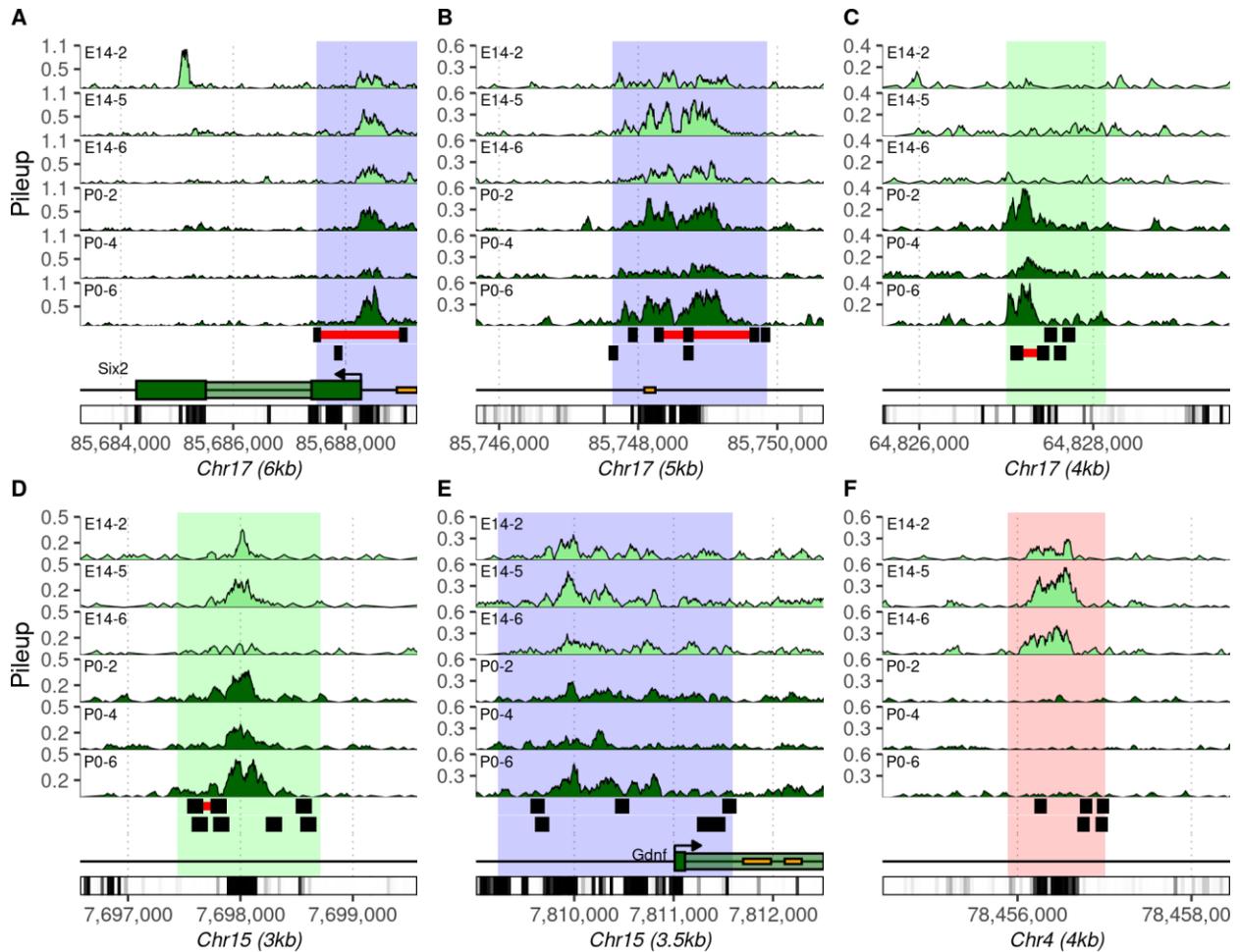


Figure 14. ATAC-seq profiles at known and putative regulatory loci.

A) Accessible chromatin at the promoter of the *Six2* transcription factor. **B)** An enhancer for *Six2* with enhancer location depicted in orange³³. **C)** An unannotated intergenic IDR with increased accessibility at P0, characterized by gain of a nucleosome free region and sequence conservation across mammals. **D)** Differentially accessible enhancer for *Gdnf* with increased accessibility at P0. **E)** Promoter of *Gdnf* is accessible, but it does not significantly change over time. **F)** Unannotated intergenic DAR that is highly conserved and shows significantly reduced accessibility between E14.5 and P0. IDR regions that are stable between time points are highlighted in blue, DARs which show increasing or decreasing accessibility with FDR controlled at 0.1 are highlighted in green and red, respectively. The top three panels in each plot show pileup in each E14.5 sample (light green), and the next three panels show pileup in each P0 sample (dark green). The remaining tracks from top to bottom indicate nucleosomal occupancy detected using NucleoATAC with nucleosomal regions annotated in black and nucleosome-depleted regions (NFRs) in red,

genomic annotations, and PhastCon60 conservation. Orange blocks in annotation tracks indicate mm10 coordinates for regulatory elements identified by the FANTOM5 consortium¹⁹¹.

3.5.3 Early and late nephron progenitors have distinct miRNA transcriptional profiles

Small-RNA sequencing detected 1,104 known miRNA transcripts, and 114 miRNA that show a significant change in expression between E14.5 and P0 samples (FDR = 0.05). Exactly half of these changing miRNA were increasing in expression between E14.5 and P0. Principal component analysis highlights greater homogeneity in miRNA expression among E14.5 samples compared with P0 (Figure 15A), and hierarchical clustering reveals a clear grouping of miRNA into those of increasing and decreasing expression between E14.5 and P0, respectively (Figure 15B). Members of the *let-7* family of miRNA are among the most highly expressed miRNA detected, and we note that all ten significantly changing members of this family exhibit a significant increase in expression (Figure 15C). This is in line with results showing that a reduction in *Lin28b* expression leads to a broad increase in *let-7* family expression over the course of nephrogenesis⁴⁶. Other miRNA significantly decreasing in expression include *miR-429-3p*, a member of the *miR-200* family known to affect podocyte differentiation²⁰², and both *miR-125a-5p* and *miR-125b-5p*, which have been found to repress *Lin28* transcripts in different cellular environments^{203,204}. Among miRNA with decreasing expression, the angiogenesis²⁰⁵ and proliferation-promoting²⁰⁶ *miR-126a* is among the those with the most substantially reduced expression. Gene transcripts that were the predicted targets of up-regulated miRNA according to TargetScan²⁰⁷ include *Sox11*, whose protein promotes expression of *Wnt4* to induce mesenchymal-to-epithelial transitions²⁰⁸, and *Bach2*, a proposed molecular link between the MAPK/AP1 and

Six2/ β -catenin pathways for self-renewal and differentiation in nephron progenitors, respectively¹²².

The DIANA miRPath tool enables KEGG pathway analysis of miRNAs based on their respective gene targets as predicted by microT-CDS⁶². We find that miRNA with the greatest changes in expression between E14.5 and P0 (up or down) are those that regulate pathways with key roles in nephrogenesis (Figure 15, Supplemental figure 4). For instance, Tgf- β has been shown to be secreted by stromal cells to instigate nephron progenitor differentiation²⁰⁹, and we note that several genes within the “TGF- β signaling pathway” (KEGG pathway mmu04350) are predicted targets of miRNA with reduced expression over time (FDR $1.45e^{-6}$). The “MAPK signaling pathway” (mmu04010) plays important roles in organizing and priming nephron progenitors for differentiation, and regulates their interactions with the extracellular matrix (ECM) via the genes proteins Pax2 and Itg α 8²¹⁰. We note a significant number of up- and down-regulated miRNA target this pathway (FDR of $3.5e^{-6}$ and $3.0e^{-4}$ for up and down lists, respectively). Interestingly, this analysis suggests that the ECM appears to be directly affected by changing miRNA expression: the KEGG pathway “Extracellular matrix (ECM) receptor interaction” (mmu04512) is the most significantly enriched pathway identified, and is more significantly targeted by miRNA which increase in expression rather than decrease (FDR $7.4e^{-28}$ versus $1.1e^{-2}$, respectively). ECM dynamics play a variety of roles in development²¹¹ and in determining cellular identity²¹², and are a crucial component in the regulation of nephrogenesis²¹³.

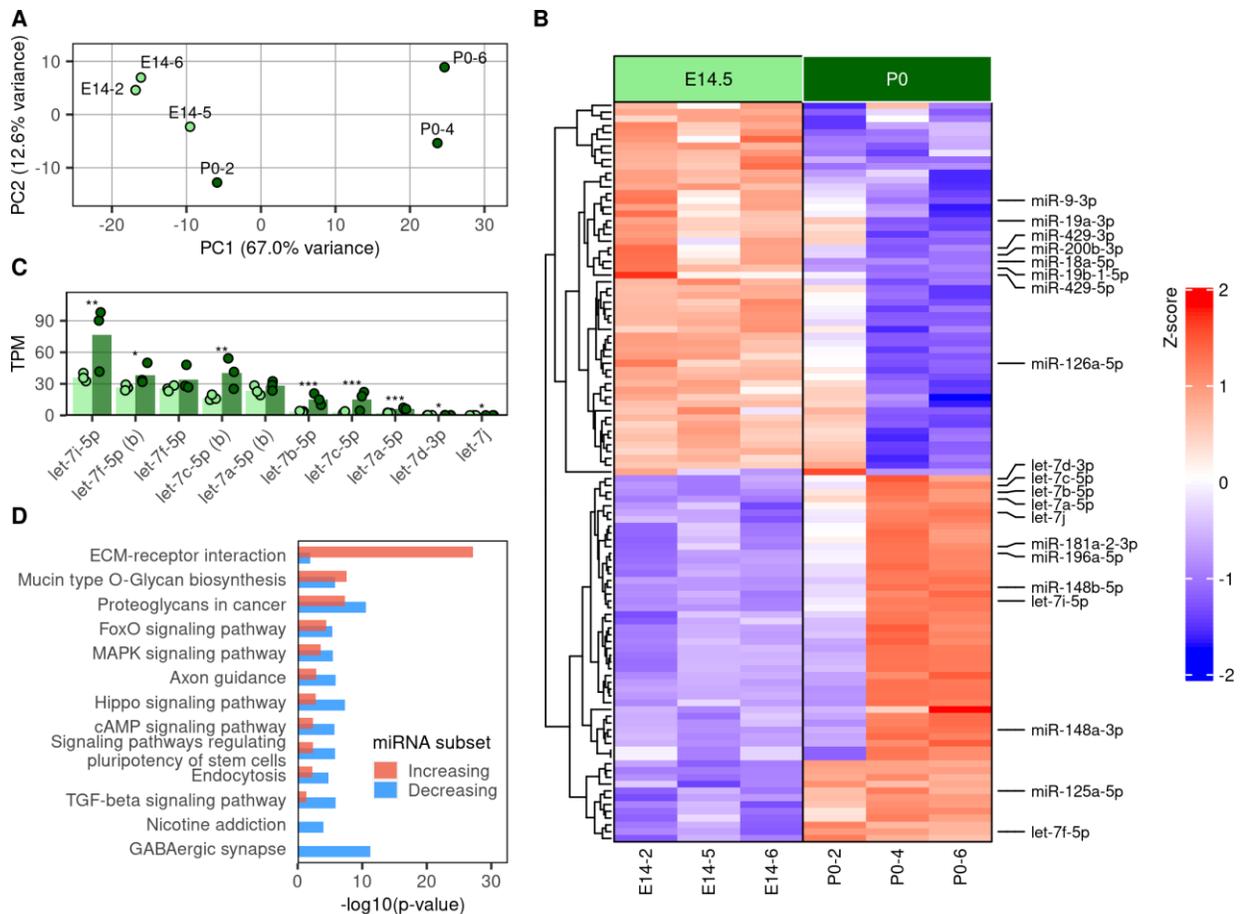


Figure 15. Substantial changes in miRNA expression between E14.5 and P0 nephron progenitor cells.

A) Principal component analysis shows that developmental time point is a major contributor to miRNA expression variation. **B)** Heatmap depicting relative expression of all significantly changing miRNA. Columns containing E14.5 samples on the left are annotated in light green, and columns with P0 samples on the right are annotated in dark green. **C)** Significantly changing members of the let-7 family of miRNA are all increasing between E14.5 and P0 timepoints. **D)** Enrichment of genes from specific KEGG pathways among predicted gene targets of up- and down-regulated miRNA (red and blue bars, respectively). Minimum p-value included is $1e^{-5}$.

3.5.4 Topologically associated domains help identify potential enhancer—miRNA relations

Topologically associated domains (TADs) are megabase-scaled regions of higher-order chromatin organization which result from the packaging of chromatin into the nucleus in three-

dimensional space⁷⁹. TADs are remarkably conserved between cell types, and enhancer-promoter interactions are typically restricted to elements within the same TAD⁸⁰. To screen for possible enhancers of miRNA, we anticipated that increasing/decreasing accessibility in a given DAR may be the result of increased/decreased activity in that location, and if this region contains a regulatory enhancer this change may correlate with increased/decreased expression of its targeted miRNA. In implementing this screen, 42,778 unique accessible regions are located inside annotated topologically associated domains, and among these 30,626 do not overlap a known promoter. Of these accessible regions 2,103 are DARs, with 1,180 and 923 opening and closing over time, respectively. Fifteen opening DARs share a TAD with one or more miRNA that increases expression, and nine closing DARs share a TAD with one or more miRNA that decreases expression. With this approach we identified matches among 24 DARs and 23 differentially expressed miRNA (Figure 16, A-C). The mean distances between screened enhancer/promoter pairs is 447 kb, and range from 27 kb to over 1Mbp

Amongst the miRNA-enhancer pairs annotated, we note three of particular interest. First, an opening DAR has been identified located 120 kb downstream of *let-7c-5p* and 73 kb downstream of *miR-125b-5p*. *Let-7c-5p* is one of the highest-expressed of the detected *let-7* miRNA repressed by *Lin28b*⁶⁸, and *miR-125b-5p* has been implicated in repression of *Lin28* transcripts in mouse cell types²⁰⁴. Within this putative enhancer, we detect DNA binding footprints for transcription factors known to play important roles in kidney development, including *Sox9*, a transcription factor known to mark highly proliferative “progenitor like” cells during kidney repair (Figure 17B)^{214,215}.

Second, *miR-9-3p* is the reverse complement to *miR-9-5p*, a miRNA which has recently been shown to protect from kidney fibrosis by targeting metabolic pathways²¹⁶. We detect a greater

than 10-fold change in *miR-9-3p* transcripts between E14.5 and P0 nephron progenitors, and have identified a potential enhancer approximately 250kb away (Figure 17A).

Third, we matched *miR-181a-2-3p* to a possible intergenic enhancer 131kb upstream on chromosome 2, within which we note a transcription factor footprint for the transcription factor Wt1, a key regulator of nephron progenitor survival²¹⁷.

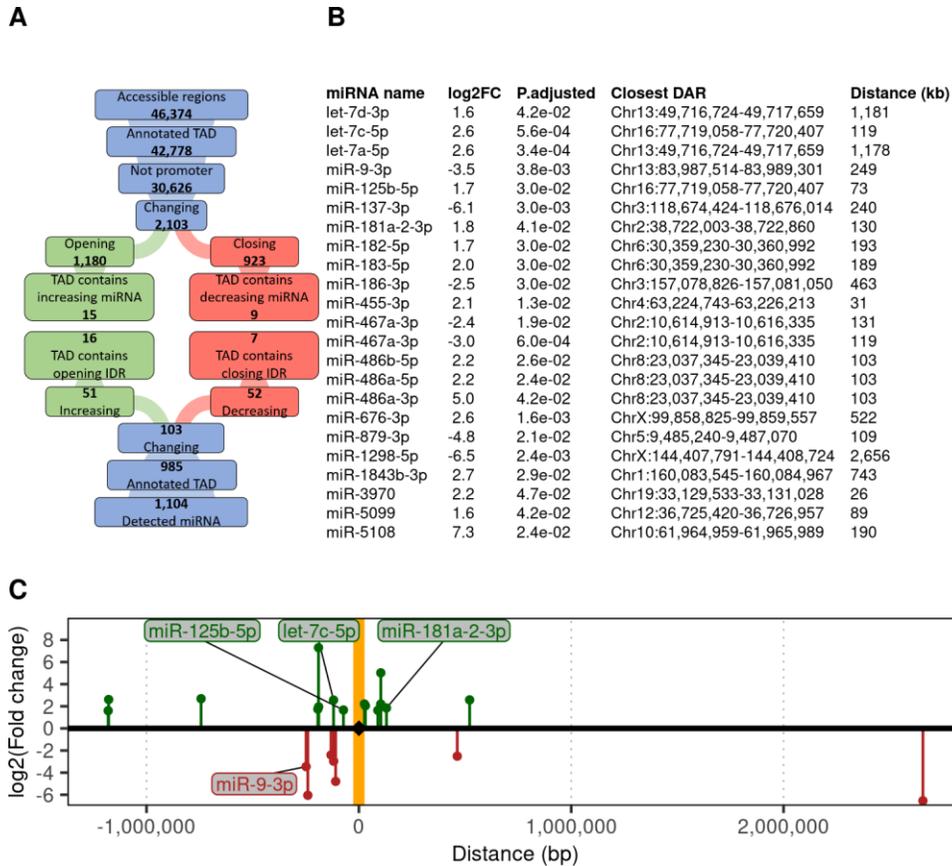


Figure 16. Finding candidate enhancer—target-miRNA pairs.

- A)** Flow chart denoting how regions of accessibility were prioritized for possible regulation of miRNA expression.
- B)** Table denoting miRNAs matched to one or more DAR in the same TAD, along with fold change of miRNA expression between E14.5 and P0 and adjusted p-values
- C)** Genomic locations of each screened miRNA relative to their candidate enhancer (orange vertical stripe). The scale of their log₂ fold change is shown on the y-axis, and miRNA with increasing or decreasing expression are highlighted in green and red, respectively.

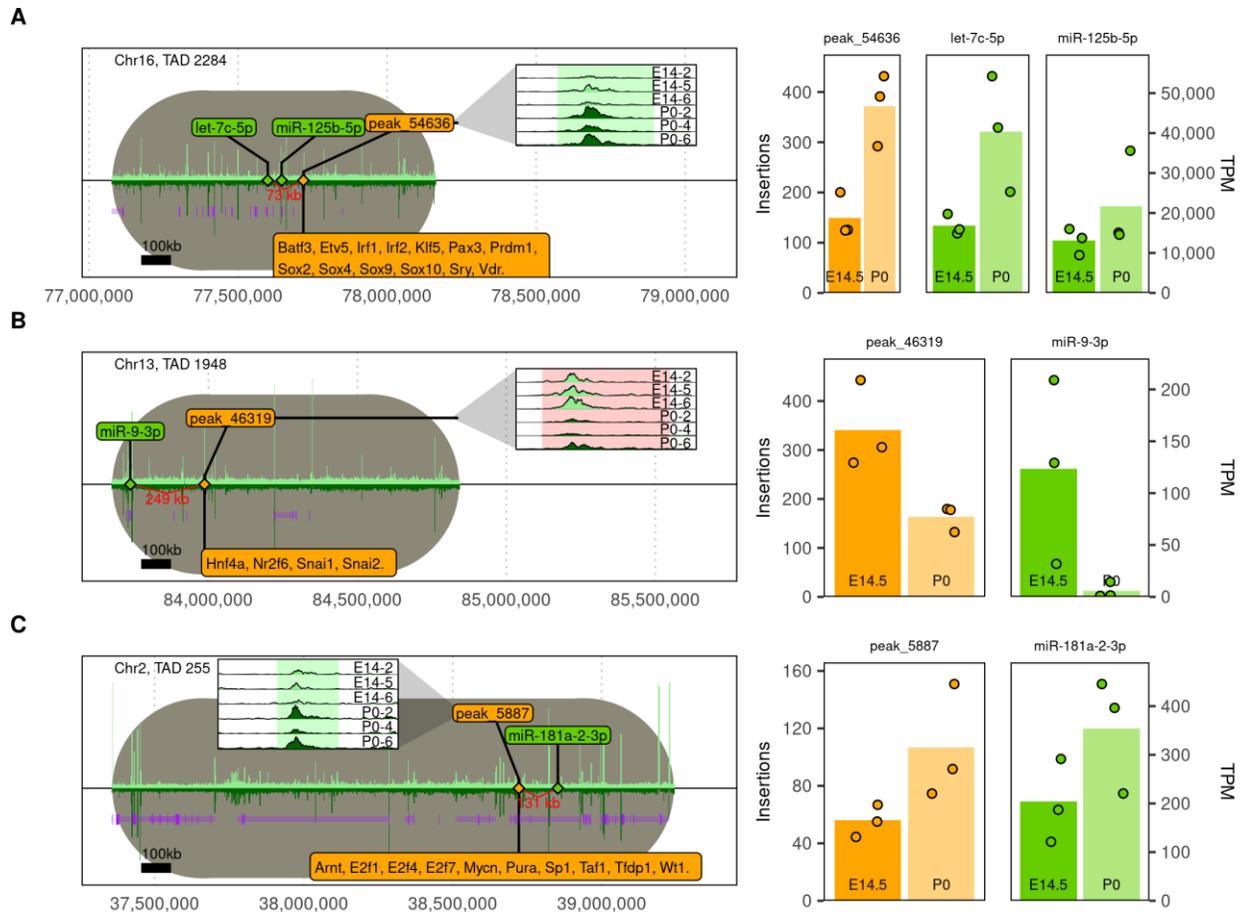


Figure 17. Three potential enhancer—target-miRNA pairs.

TADs are highlighted in gray, aligned by starting points and drawn in the same scale to illustrate relative size.

Known genes from Ensembl are shown in purple, and darker regions indicate exons. Relative accessibility at E14.5 and P0 are shown as bar plots above and below the horizontal black line, respectively, as well as in inset plots illustrating pileup in the specified DAR. Transcription factor footprints identified by HINT are listed in orange banners. Bar plots in separate panels on the right indicate miRNA expression (green) and normalized Tn5 insertion events in proposed enhancer regions (orange). **A)** *miR-9-3p* is matched to a DAR 249kb away. **B)** The let-7 family member *let-7c-5p* is very highly expressed in nephron progenitor cells and is matched to “peak_54636” 119kb away in an intergenic region. **C)** *miR-181a-2-3p* is located 131kb away from “peak_5887,” a possible intergenic enhancer.

3.6 Discussion

In this study, we sought to define miRNA expression changes and changes in chromatin accessibility in nephron progenitor cells during embryogenesis genome-wide, and to link miRNA expression changes with alterations in accessibility. We identified 114 miRNAs differentially expressed between E14.5 and P0, together with 2,406 differentially accessible regions (DARs). We next used this data to associate DARs (i.e., putative regulatory regions) with differentially expressed miRNA based on genomic distance and presence in the same topologically associated domain¹⁶⁶. This resulted in 37 DARs (possible enhancers) with one or more of 18 potential close-by miRNA targets, including several members of the *let-7* family. We have produced a unique dataset that observes the dynamics of chromatin accessibility changes and the miRNA transcriptome of wild-type progenitors in parallel as they approach cessation of nephrogenesis, and we have documented changes in each that suggest how both may guide progenitor renewal and differentiation.

Prior studies have also observed changes in chromatin over the course of nephrogenesis⁸⁷, including changes in chromatin markers as well as accessibility in nephron progenitor cells cultured from different time points^{122,171}. We identify several thousand unannotated accessible regions that are highly conserved in mammals and which may portend novel regulatory features. Among regions of chromatin which change between E14.5 and P0, GREAT analysis¹⁹⁰ revealed a significant enrichment for GO biological processes related to cell motility, cell migration, locomotion, and chemotaxis, in addition to epithelial cell differentiation and Notch signaling (the latter is known to repress *Six2* expression and promote nephron progenitor cell differentiation⁴¹). Changes affecting any of these processes may affect the rate of progenitor cell differentiation,

either directly or stochastically through changes to their migratory behavior in spatial relation to the Wnt9b inductive signal¹⁶⁸.

To our knowledge, this work represents the first description of miRNA expression in nephron progenitors over time. We observe broad increases in expression of the *let-7* family as progenitors age, consistent with previous findings⁴⁶. We observed a total of 114 miRNAs that exhibit a significant change in expression over this period of development, suggesting that the role of miRNAs in nephron progenitor aging is not limited to the *let-7* family. miRNAs with significantly decreased expression in P0 progenitors include two members of the *miR-200* family, *miR-200b-3p* and *miR-429-3p*. This suggests these miRNA could play a role in nephron progenitor cell differentiation, consistent with the knowledge that the *miR-200* family is known to be expressed in nephron progenitors²¹⁸, has been implicated in podocyte differentiation²⁰², and is well-described in regulating epithelial-to-mesenchymal transitions^{219,220}.

Pathway analyses of gene targets of miRNA that change expression between E14.5 and P0 implicate genes in a variety of known KEGG pathways, including MAPK and Tgf- β signaling, key regulators of nephron progenitor proliferation and differentiation, respectively^{122,210}. Potentially congruent with our ATAC-seq findings, up-regulated miRNA targets are significantly enriched for KEGG pathways that affect the extracellular matrix, actin cytoskeleton, and focal adhesion, all crucial facets of cellular migration, differentiation, branching, attachment, polarization, and proliferation²¹¹. ECM-associated genes are targeted by some of the most highly expressed miRNAs we detect, including *miR-196a-5p*, *miR-148a-3p*, *miR-125a-5p*, and *let-7f-5p*. *miR-196a-5p* in particular targets *Colla1*, *Colla2*, and *Col3a1*. These are all up-regulated in *Pax2*-deficient nephron progenitors, which itself results in transdifferentiation into the renal stromal lineage²²¹.

By combining ATAC-seq and smRNA-seq in matched samples we screened for possible enhancer-miRNA relationships and interactions. We measured a significant increase in expression of *miR-125a-5p*, a miRNA which is known to repress proliferation by targeting *Lin28* transcripts²⁰³. We then matched this miRNA to a possible enhancer with several transcription factor footprints ascribed to the *Sox* family. *Sox* family members are known regulators of cell fate in stem and progenitor cells,²²² and *Sox9* in particular marks highly proliferative multi-lineage progenitor cells in mouse kidneys^{214,215}. The same enhancer is also matched to *let-7c-5p*, another miRNA known to target *Lin28* transcripts⁴⁶. This feature may regulate expression of *miR-125a-5p* or *let-7c-5p* (or both) to repress *Lin28*, thereby repressing progenitor proliferation and promoting differentiation.

Our study has identified changing chromatin regions and miRNA expression that may control nephron progenitor differentiation and renewal. Ontological analyses for both the ATAC-seq and smRNA-seq data provide an enticing view of the mechanisms by which they may do this, and they reveal key transcripts or regulatory regions to interrogate in future.

4.0 Single cell RNA sequencing reveals differential cell cycle activity in key cell populations during nephrogenesis

4.1 Foreword

This chapter is adapted from a work in submission for publication in which I was co-first author. I performed the single-cell sample preparation and aided Abha Bais with the data analysis. Abha Bais contributed to study design along with Débora Malta Cerqueira, who performed histological assays and drafted the manuscript. Jacqueline Ho and Dennis Kostka contributed to the experimental design and editing the manuscript

Citation: Abha Bais*, Débora M. Cerqueira*, Andrew Clugston *, Jacqueline Ho and Dennis Kostka. Single cell RNA sequencing reveals differential cell cycle activity in key cell populations during nephrogenesis. (2020 bioRxiv; doi:10.1101/2020.09.16.300293) * first co-authors

Our work revealed that both miRNA expression and chromatin accessibility exhibit distinct changes in nephron progenitors that match with their stage of development at the time of isolation. Importantly, ontological analyses of predicted miRNA target genes (DIANA miRPath 3.0⁶², <http://snf-515788.vm.okeanos.grnet.gr/>) and enrichment analysis of genomic annotations near regions of changing chromatin accessibility (Genomic Region Enrichment of Annotations Tool; GREAT¹⁹⁰, <http://great.stanford.edu/public/html/>) implicate their respective changes in functions related to cell fate, cell movement, and extracellular matrix (ECM) interactions. We sought to identify gene expression differences that may reflect key developmental lineages, intending to

identify signatures in gene expression among nephron progenitors, the nephron progenitor niche, and other lineages in the developing kidney. We accomplished this using single-cell RNA sequencing of a wild-type E14.5 kidney (Figure 7), which enabled us to look at gene expression in all of the cell types of the E14.5 kidney simultaneously. Clustering analysis allowed us to annotate cells into key cell types of the kidney, including nephron progenitors, which were in turn able to be grouped by gene expression patterns into “primed” and “self-renewing” populations, which are known to parallel differences resulting from developmental stages⁴⁷.

4.2 Summary

The kidney is a complex organ composed of more than 30 terminally differentiated cell types that all are required to perform its numerous homeostatic functions. Defects in kidney development are a significant cause of chronic kidney disease in children, which can lead to kidney failure that can only be treated by transplant or dialysis. A better understanding of molecular mechanisms that drive kidney development is important for de-signing strategies to enhance renal repair and regeneration. In this study, we profiled gene expression in the developing mouse kidney at embryonic day 14.5 (Figure 7) at single cell resolution. Consistent with previous studies, clusters with distinct transcriptional signatures clearly identify major compartments and cell types of the developing kidney. Cell cycle activity distinguishes between the “primed” and “self-renewing” sub-populations of nephron progenitors, with increased expression of the cell cycle related genes *Birc5*, *Cdca3*, *Smc2* and *Smc4* in “primed” nephron progenitors. Augmented *Birc5* expression was also detected in immature distal tubules and a sub-set of ureteric bud cells, suggesting that *Birc5*

might be a novel key molecule required for early events of nephron patterning and tubular fusion between the distal nephron and the collecting duct epithelia.

4.3 Introduction

The mammalian kidney has evolved to provide critical adaptive regulatory mechanisms, such as the excretion of waste, and the maintenance of water, electrolyte, and acid-base homeostasis to the body. These functions require the coordinate development of specific cell types within a precise three-dimensional pattern. Defects in kidney development are amongst the most common malformations at birth. Congenital anomalies of the kidney and urinary tract (CAKUTs) represent more than 20 percent of birth defects over-all²²³, and they account for a large fraction of chronic kidney disease and renal failure in children⁹. For example, the number of nephrons formed at birth is thought to be an important determinant of renal function, because reduced nephron numbers are often observed in humans with primary hypertension and chronic kidney disease^{224,225}. An estimated 37 million people in the United States (~15% of the population) have chronic kidney disease (CKD)^{226,227} that can lead to kidney failure requiring transplant or dialysis. Development of strategies to enhance renal repair or regeneration are needed to reduce the morbidity and mortality associated with kidney disease, and they are dependent on a better understanding of the molecular genetic processes that govern kidney development.

Nephrons form the functional units of the kidney and are derived from a nephron progenitor (NP) cell population, also known as cap mesenchyme. These cells are capable of self-renewal, which is necessary to generate an appropriate number of nephrons during the course of embryogenesis and development. They are also multipotent, that is they have the ability to

differentiate into the multiple cell types of the mature nephron^{23,34}. More specifically, multipotent Cbp/P300-Interacting Transactivator 1 (*Cited1*)-positive/*Sine Oculis Homeobox Homolog 2* (*Six2*)-positive nephron progenitors give rise to multiple nephron segments, and are termed “self-renewing” nephron progenitors⁸. The transition of nephron progenitors into epithelialized structures is dictated by a series of tightly orchestrated signaling events. Of this, Bone morphogenetic protein 7 (*Bmp7*) induces the initial exit of *Cited1*⁺/*Six2*⁺ cells into a *Cited1*⁻/*Six2*⁺ state, which marks nephron progenitors “primed” for differentiation by ureteric bud-derived Wnt family member 9b (*Wnt9b*)/ β -catenin signaling. Conversely, remaining *Cited1*⁺/*Six2*⁺ nephron progenitors are kept in an undifferentiated and self-renewing state in response to Fibroblast growth factor 9 (*FGF9*), Wnt and BMP7 signals^{30,33,36,39,40,44,228–230}.

Upon *Wnt9b*/ β -catenin stimulation, nephron progenitors undergo a mesenchymal to epithelial transition to form pre-tubular aggregates, which then proceed to develop sequentially into polarized epithelial renal vesicles, comma-, and then S-shaped body structures. Cells in the proximal portion of the S-shaped body differentiate into podocytes (glomerular development), while its mid- and distal portions give rise to tubular segments of the nephron, which are subdivided into proximal tubules, loops of Henle and distal tubules²³¹, Figure 18A. During the S-shaped stage of glomerular development, developing podocytes secrete vascular endothelial growth factor A (*VEGF-A*), which attracts invading endothelial cells into the cleft of the S-shaped body. Platelet-derived growth factor- β (*PDGF* β) signal produced by endothelial cells mediates the recruitment of mesangial cells, which invade the developing glomerulus and attach to the forming blood vessels. By the end of maturation, the glomerulus consists of four specified cell types: the fenestrated endothelium, mesangial cells, podocytes, and parietal epithelial cells of the Bowman’s capsule^{232–236}.

Single cell RNA sequencing (scRNA-seq) technology offers the ability to comprehensively identify the transcriptional and (inferred) cellular composition of the developing kidney. Recent studies in developing mouse^{175,200,237–240} and human kidneys^{241–243} have contributed to our understanding of subpopulations of nephron progenitors and stromal cells, lineage fidelity, novel receptor-ligand pathways, and differences between mouse and human kidney development. scRNA-seq also has the potential to inform improvements in our ability to culture nephron progenitor cells^{45,244,245} and produce higher-fidelity human kidney organoids^{246,247}, and to develop novel strategies for enhancing renal repair and regeneration.

In this study, we used scRNA-seq to interrogate cell types and transcriptomes within 4,183 cells from one kidney pair of an E14.5 female mouse embryo. Clustering identified eleven clusters corresponding to the major components/cell-types of the developing kidney and revealed expression of known lineage markers in unexpected cell types (e.g. renal stromal markers in nephron progenitors). Pseudotime analysis was utilized to describe transcriptional dynamics as nephron progenitors differentiate. Notably, we find that cell cycle activity distinguishes between the “primed” and “self-renewing” sub-populations of nephron progenitors, with increased levels of the cell cycle related genes *Survivin (Birc5)*, *Cell division cycle-associated protein 3 (Cdca3)*, and *Structural maintenance of chromosomes proteins 2 (Smc2)* and *4 (Smc4)* in the “primed” sub-population. Moreover, increased *Birc5* expression was also observed in immature distal tubules and in a subset of ureteric bud cells, suggesting its involvement in fusion between the distal nephron and the collecting duct epithelia, perhaps by promoting cell survival in these cells.

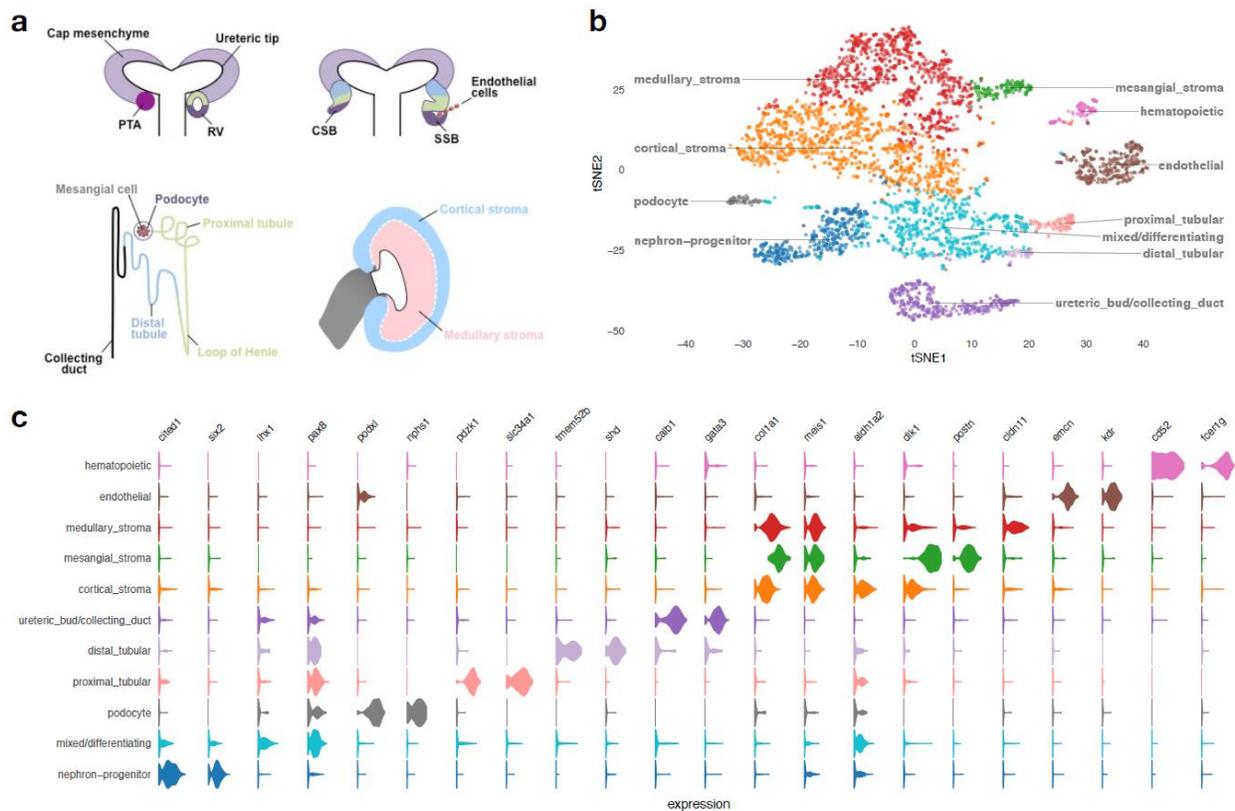


Figure 18. Developing embryonic day 14.5 mouse kidney cell types.

A) Schematic illustration of nephron induction and patterning. In response to signals from the ureteric bud, the metanephric mesenchyme condenses and forms a cap of nephron progenitors (cap mesenchyme) around the ureteric bud tips. Next, a sub-population of nephron progenitors undergoes a mesenchymal to epithelial transition to form pre-tubular aggregates (PTA), which develop sequentially into renal vesicles (RV), comma-shaped body (CSB) and S-shaped body (SSB). Endothelial cells are attracted into the cleft of the SSB. Color-coded map indicates the cell fate relationship of progenitor regions in SSB structure (upper right) and adult nephron structure (lower left). Schematic of a lateral view of the metanephric kidney depicting the cortical and medullary stroma (lower right). **B)** tSNE plot showing the eleven cell clusters in the embryonic mouse kidney, with cell clusters corresponding to major components indicated by color. **C)** Violin plots of gene expression for known lineage-associated genes (columns), stratified by cluster (rows). Our data clearly identifies cells from the major structural components of the developing kidney.

4.4 Methods

4.4.1 Embryonic kidney collection and single-cell RNA sequencing

Timed pregnant wild-type CD1 female mice used in this study were obtained from Charles River Laboratories (Wilmington, MA, USA). The date on which the plug was observed was considered embryonic day 0.5 (E0.5). All experimental procedures were performed in accordance with the University of Pittsburgh Institutional Animal Care and Use Committee guidelines (IACUC protocol #17091432), which adheres to the NIH Guide for the Care and Use of Laboratory Animals.

We harvested two kidneys at E14.5 (Figure 7) and generated a single cell suspension using 0.05% trypsin at 37°C for 10 minutes. Kidneys were mechanically dissociated with pipetting at 5 and 10 min. 3% fetal calf serum in PBS was added to halt the trypsin. The cell suspension was filtered using a 40 µm filter and pelleted. The cells were resuspended in 90% FCS in DMSO and frozen, prior to shipment to GENEWIZ Inc. Single-cell library preparation and sequencing was performed by GENEWIZ Inc. using the 10X Genomics Inc. Chromium 3' Single Cell v2 library preparation kit. Cells exhibited high viability after freezing and thawing (>90%).

4.4.2 Data processing, quality control, and normalization

4.4.2.1 Alignment and read counting

Sequencing data was processed using the CellRanger count pipeline of the Cell Ranger software (version 2.2.0) (www.10xgenomics.com) to perform alignments and yield bar-code and UMI counts, such that the cell detection algorithms are bypassed and counts for 10,000 cells are

returned (force-cells=10000 option). The mouse reference genome (GRCm38.p4) and transcript annotations from Ensembl (version 84) were used²⁴⁸.

4.4.2.2 Quality Control

The Bioconductor²⁴⁹ R package DropletUtils^{250,251} was used to detect and remove empty droplets with default parameters at an FDR of 0.01, yielding a total of 5,887 non-empty droplets. Multiple quality control (QC) metrics were calculated using the R package scater²⁵² and cells with at least 1,000 detected features, and percentage of mitochondrial counts less than three times the median absolute deviation (MAD) from the median value were considered, resulting in a total of 4,402 cells. We excluded putative doublets as the top 5% cells ranked by the hybrid score from the R package scds²⁵³, further filtering out 220 droplets. Finally, only genes with at least three or more counts in at least three samples were considered, yielding a digital gene expression matrix comprising 11,155 genes in 4,183 cells/droplets.

4.4.2.3 Normalization

We normalized the data using size factors calculated using the deconvolution method implemented in the computeSumFactors function in the R package scran²⁵⁴ after performing clustering using the quickcluster function on endogenous features with an average count ≥ 0.1 , (min.mean=0.1 option) yielding log-transformed normalized expression data. Feature selection and dimension reduction were performed using scran procedures. Briefly, we fit a mean-variance trend to the gene variances using the trendVar function and identified the biological component of the total variance with decomposeVar. All genes with an FDR < 0.01 and proportion of biological variance of at least 25% are considered as highly variable genes (HVG). Principal component

analysis (PCA) was then performed using denoisePCA and two-dimensional representation was then derived using runTSNE (see Figure 18).

4.4.3 Identification of major structural components of the kidney

Cells were grouped into clusters using the scran R package by building a shared k-nearest-neighbors graph using buildSNNGraph (with use.dimred=PCA and k=25 options), followed by clustering with the Walktrap community finding algorithm as implemented in the iGraph package (<https://igraph.org>), cutting the graph at 10 clusters. We used the expression of a curated list of marker genes for major components of the developing kidney (see Figure 18B) to assign cluster labels. Cluster-specific markers were derived using the findMarkers function. We note that at this resolution tubular distal cells were grouped in the mixed/differentiating group; specific analysis of nephron progenitor descendant cell types then revealed distinct groups of distal vs. proximal tubular cells (see below).

4.4.4 Nephron progenitor and descendant cell types

4.4.4.1 Selecting and characterizing NP lineage cells

Focusing on nephron progenitor and descendant cell types (termed “nephron-progenitor”, “mixed/differentiating” (at that point “tubular_dist” cells as well) , “podocytes and “tubular_prox” in Figure 18) and requiring expression of each gene with at least three counts in three cells yielded a gene expression matrix of 9,611 genes across 1,273 cells. Following the same procedure as before we derived a low-dimensional representation and identified six clusters of cells, corresponding to two types of nephron progenitor cells (“self-renew” and “primed”),

“mixed/differentiating” cells as well as distal tubular cells and proximal tubular cells. Cluster-specific marker genes were derived as before. Enrichment analysis for Gene Ontology terms enriched between “self-renew” and “primed” and between “primed” and “mixed/differentiating” (Figure 20B, C) were performed using the topGO function of the Limma Bioconductor package¹⁵⁶ with default parameters.

4.4.4.2 Pseudotime Analysis of NP cells

Pseudotime analysis was performed using slingshot²⁵⁵, using cluster labels and principal components derived as described above (via the clusterLabels and reducedDim options). This recovered three lineages (to podocytes, distal-, and proximal tubular cells), with cells in “self-renew”, “primed” and early “differentiating/mixed” being shared (see Supplemental figure 5).

Next, we fitted a multinomial log-linear model (using the nnet package²⁵⁶) relating pseudotime with the annotated clusters. For cells with more than one annotated lineage (in the “self-renew”, “primed” and early “differentiating/mixed” clusters) lineage-pseudotimes from slingshot were averaged. This enabled us to define NP-cells as cells with annotated pseudotime less than the (pseudo)timepoint between “primed” and “mixed/differentiating” where the probability of the “primed” cluster has declined to 50% (i.e., 50% probability for “mixed/differentiating”, see Supplemental figure 6). These cells contain all “self-renew” cells, 20 “differentiating” cells and all but 15 “primed” cells and were used in analyses for Figure 22.

Focusing on so-defined 450 bona fide NP-cells, we used SAVER²⁵⁷ to impute gene expression values and then applied generalized additive models, as implemented in the mgcv package²⁵⁸, to screen for pseudotime-associated genes (Figure 20H) by modeling gene expression as a smooth function of pseudotime. Focusing on high-quality pseudotime-associated genes (FWER<1%, Bonferroni correction) we further required high variability (as measured by median

absolute deviation) and substantial Spearman correlation (for each gene with pseudotime, $|\rho| > 0.3$). This yielded 175 genes with overall decreasing expression across pseudotime (down-regulated), and 395 with increasing expression (up-regulated).

4.4.4.3 Enrichment Analysis for NP cells

We used MsigDB (v7.0)²⁵⁹ and hypergeometric tests to screen for annotated gene sets enriched for up- or down-regulated genes, focusing on Gene Ontology and Hallmark gene sets. We then screened for regulatory modules in time-varying genes using SCENIC²⁶⁰, where we used default options including GENIE3²⁶¹ for network inference.

4.4.4.4 Immunohistochemical staining

Kidneys dissected from embryonic day 14.5 (E14.5) and postnatal day 0 (P0) mice were fixed overnight in 4% paraformaldehyde, embedded in paraffin, and sectioned at 4 μm . After deparaffinization, rehydration, and permeabilization in PBS-Tween (PBS-T), antigen retrieval was performed by boiling the slides in 10 mM sodium citrate pH 6.0 buffer for 30 min. Next, sections were blocked in 3% bovine serum albumin (BSA) and incubated overnight with antibodies recognizing Birc5 (#2808, Cell Signaling Technology, Danvers, MA, USA), Cyclin D1 (#2978, Cell Signaling) and Neural cell adhesion molecule (C9672, Sigma-Aldrich, St. Louis, MO, USA) at the dilutions recommended by the manufactures. On the next day, sections were washed with PBS-T, incubated with secondary antibodies at the dilution of 1:200, washed again with PBS-T, and mounted in Fluoro Gel with DABCO (Electron Microscopy Science, Hatfield, PA) before being visualized with a Leica DM2500 microscope and photographed with a Leica DFC 7000T camera using LAS X software (Leica, Buffalo Grove, IL, USA). Goat anti-rabbit 594 (#111-515-

047) and donkey anti-mouse 488 (#715-545-151) antibodies were purchased from Jackson ImmunoResearch Laboratories (West Grove, PA, USA).

4.4.4.5 In situ hybridization

Kidneys were harvested from P0 pups, fixed in 4% paraformaldehyde overnight, treated with 30% sucrose/PBS and embedded in Tissue-Tek Optimal Cutting Temperature Compound (OCT; Sakura, Torrance, CA, USA). In situ hybridization was conducted on 10 μ m cryosections as described²⁶². To generate sense and antisense probes, plasmids were linearized and transcribed as follows: pGEM®-T Easy-RSPO1-SacII/SP6 and pGEM®-T Easy-RSPO1-SaI/T7.

4.4.5 Reproducibility and data availability

Scripts used for data processing and data analysis are available on github: kostkalab/wksc. The data itself is available on GEO: (link to GSE).

4.5 Results

4.5.1 Single cell gene expression identifies anatomical structures and cell lineages in the developing kidney

New nephrons are induced in response to signals from the ureteric bud through-out nephrogenesis until approximately postnatal day 3 in mice¹⁷. We chose to perform scRNA-seq at embryonic day 14.5 (E14.5), a timepoint at which there is active nephron induction and varying

degrees of nephron maturation, to comprehensively interrogate single cell transcriptomes spanning different stages of differentiation during kidney development at mid-gestation. Using one kidney pair from an E14.5 female mouse embryo processed using the 10X Chromium platform and Illumina sequencing, our dataset consists of 4,183 high-quality kidney cells, with a median number of 2,789 genes detected per cell. Grouping cells into eleven clusters (see Methods) reveals major components/cell-types of the developing kidney (Figure 18A-C, Supplemental figure 5). Clusters and key markers are consistent with prior single cell analyses of the developing mouse kidney^{175,238,239,263}. We observe clear separation of cells of the hematopoietic (*Cd52*, *Fcer1g*), ureteric bud/collecting duct (*Calb1*, *Gata3*), and endothelial (*Emcn*, *Kdr*) lineages from other cells of the developing kidney (nephron progenitors, mixed/differentiating cells, podocytes, tubular cells and stromal cells). Stromal lineages are marked by expression of *Colla1* and *Meis1*, while cells derived from the nephron progenitor lineage express established marker genes associated with progressive stages of nephron differentiation. Thus, *Cited1* and *Six2* identify nephron progenitors, *Lhx1* and *Pax8* mark mixed/differentiating cells, *Fxyd2* and *Hnf4a* mark tubular cells, and podocytes are marked by *Podxl* and *Nphs1*.

Consistent with other reports, we identify at least three major stromal clusters, which we identify as medullary stroma (*Colla1*, *Meis1*, and *Cldn11*), cortical stroma (*Colla1*, *Meis1*, *Aldh1a2* and *Dlk1*) and mesangial stroma (*Colla1*, *Meis1*, *Dlk1*, and *Postn*)^{237,239,264}. Analyses of in situ hybridization data at E14.5 from other reports, as well as GUDMAP (The GenitoUrinary Development Molecular Anatomy Project) and Eurexpress public resources facilitated identification and assignment of these clusters^{238,265-267}.

Taken together, these results show that our scRNA-seq data successfully captured major cell types that are expected to be present in the developing kidney at E14.5, including progenitor cells and their derivatives as well as mature cell populations.

4.5.2 Stratification of cell-types in the nephron progenitor lineage

Next we focused on nephron progenitor cells and their descendant/derived cell types (mixed/differentiating, podocytes and tubular cells). Selecting those cell types yielded 1,727 cells for further analysis. We are able to clearly distinguish between proximal and distal tubular cells and podocytes, and pseudotime analysis allows us to assess the level of lineage commitment (Figure 19A). Nephron progenitor cells (marked by *Six2* and *Cited1* expression) clearly separate into two sub-groups, “self-renewing” and “primed” (Figure 19B), see below for more details. Mixed/differentiating cells express transcription factors like *Pax8* and *Lhx1*, which are associated with nephron development and encompass nephron progenitor cells differentiating into tubular cells and podocytes. We note heterogeneity in the mixed/differentiating cell cluster, which likely contains cells with different degrees of differentiation, like pre-tubular aggregate, renal vesicle, comma-, and S-shaped bodies. Pseudotime analysis on this subset of cells reconstructs three lineages: differentiation into podocytes, and into proximal and distal tubular cells (Supplemental figure 6). This enabled us to distinguish between mature and immature podocytes and proximal/distal tubular cells (Figure 19, see Methods). Overall, our data clearly shows the major differentiation trajectories of nephron progenitor cells. For these three lineages, podocytes are marked by increasing expression of *Podxl* and *Nphs1*²⁶⁸, proximal tubular cells by *Pdzk1* and *Slc34a1*²⁶⁹, and distal tubular cells by *Tmem52b* and *Shd*²³⁹ (Supplemental figure 7 shows a

heatmap of genes with pronounced expression differences during nephron progenitor cell differentiation).

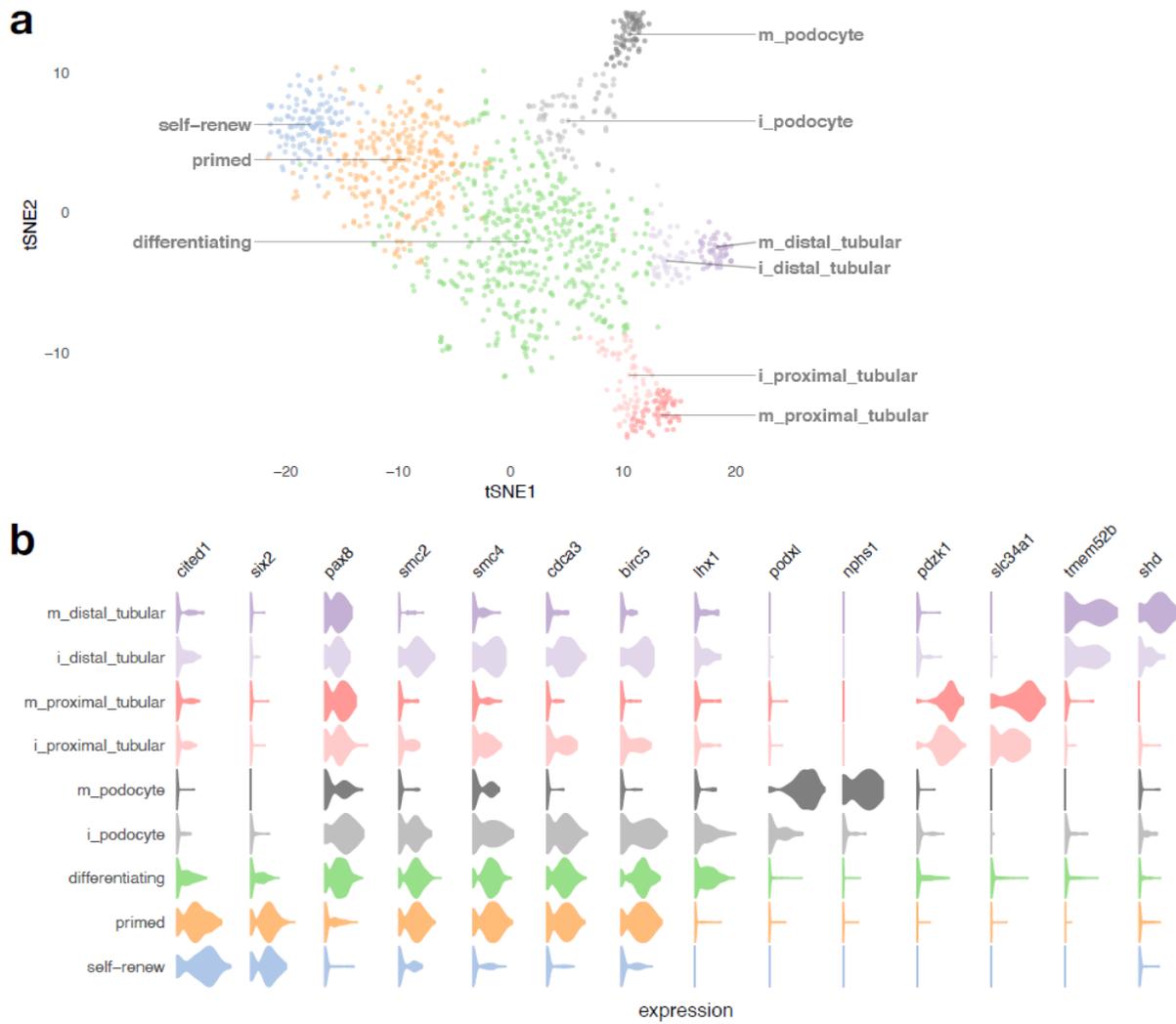


Figure 19. Cell types of the nephron progenitor lineage.

A) tSNE plot of NP-derived cells, with clusters corresponding to cell types annotated in colors. The prefix “i_” indicates immature cells, while “m_” indicates mature cells. **B)** Violin plots of gene expression for known lineage-associated genes (columns), stratified by cluster (rows). We observe two types of NP cells (“self-renewing” and “primed”) and clear separation of distal and proximal tubular cells and podocytes in our data.

4.5.3 Transcriptional dynamics across nephron progenitor cell differentiation

4.5.3.1 Cell cycle activity distinguishes between two populations of nephron progenitor cells

Comparing gene expression between “self-renewing” with “primed” NP cells yielded cell cycle as a main difference between the two types of nephron progenitor cells (Figure 20). We find that cell cycle-related genes like *Birc5*, *Cdca3*, *Smc2* and *Smc4* are up-regulated between “primed” and “self-renewing” nephron progenitor cells. Immunofluorescence analysis on kidney sections from E14.5 and P0 mice indeed corroborates these observations, showing increased expression of *Birc5* in “primed” nephron progenitors as well as pre-tubular aggregates/renal vesicles, but negligible or absent expression in the “self-renewing” nephron progenitor cells (Figure 20D sub-panels α , α' , β , β' and panel E). These results corroborate previous findings that demonstrated that the committed nephron progenitor cells are more proliferative (fast-cycling population) and more likely to differentiate than the slow-cycling, self-renewing NP population²⁰. Next, comparing primed nephron progenitor cells with mixed/differentiating cells, we observe up-regulation of transcription factors associated with differentiation (*Lhx1*, *Pax8*), and down-regulation of nephron progenitor-associated genes like *Cited1*, *Six2*, *Eya1*, *Crym*, *Meis2*, *Rspo1* and others (Figure 19b and 3a). In situ hybridization analysis confirms the expression of *Rspo1*²⁴³ primarily in nephron progenitors (Figure 20F, G).

To better understand these transcriptional changes occurring between self-renewing and primed nephron progenitor cells, we performed two types of gene set enrichment analyses. Analyzing up-regulated and down-regulated genes separately and performing enrichment analysis across Gene Ontology and Hallmark gene sets from MSigDB^{259,270} yielded “EPITHELIAL_MESENCHYMAL_TRANSITION” (FDR-adjusted p-value: 6.2E-5) as the most enriched hallmark gene set for the down-regulated genes, while gene sets enriched for up-regulated

genes included “E2F_TARGETS” as the most enriched term as well as a multitude of gene sets associated with cell cycle/replication. We then used SCENIC²⁶⁰ to gain some insight into gene regulation driving the transcriptional changes we observe across pseudotime between the two nephron progenitor cell types. Figure 20H depicts the activity of inferred regulatory modules across pseudotime for 20 recovered transcription factors and the number of their respective target genes in each module. We observe three regulatory modules of down-regulated genes, attributed to the transcription factors Hmgn3, Maf and Junb. These findings corroborate previous studies showing that these transcription factors are critical regulators of gene expression, controlling transition from a pluripotent to differentiated state in nephron pro-genitor and human embryonic stem cells^{122,271,272}. For up-regulated genes, we observe modules associated with cell cycle-related transcription factors like E2f8, Hcf1, Ezh2, Kdm2b and Mybl2, which have previously been implicated in specific aspects of cell cycle progression and cell fate decision in stem and progenitor cells^{273–278}.

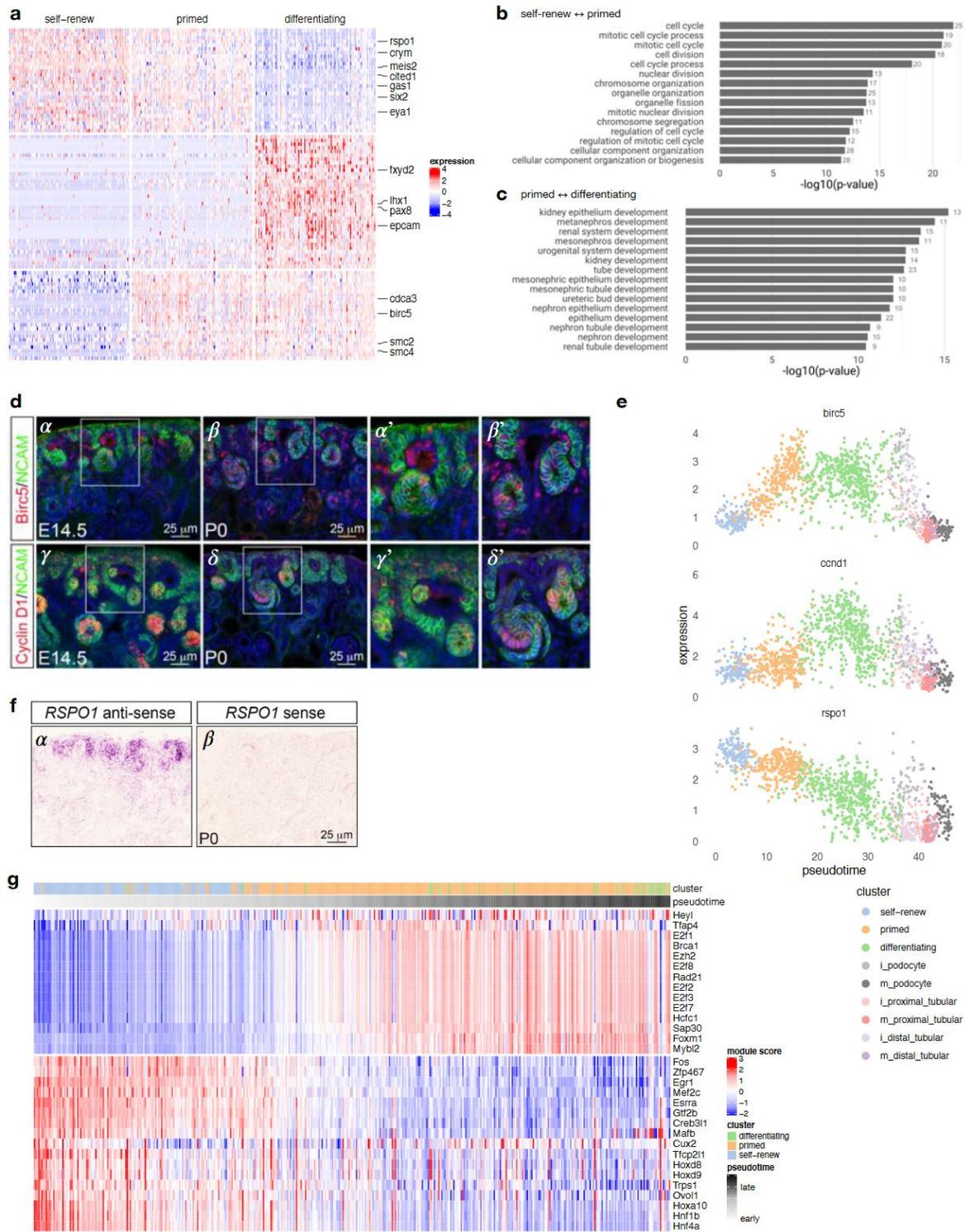


Figure 20. Transcriptional signatures of self-renewing, primed and differentiating nephron progenitor cells.

A) Differentially expressed genes on a heatmap of 100 random cells for each of the “self-renew”, “primed” and differentiating clusters, with key genes annotated on the right. **B,C)** Ten most-enriched Gene Ontology terms for

genes differentially expressed between self-renewing and primed NP cells, and between primed NP cells and differentiating cells, respectively. **D**) Immunofluorescence on kidney sections from embryonic day 14.5 (E14.5) and postnatal day 0 (P0) mice using anti-*Birc5* (α - β) and anti-Cyclin D1 (γ - δ) antibodies (red). Nephron progenitors and their early epithelial derivatives were detected using an antibody against anti-Neural cell adhesion molecule (NCAM; green). Nuclei were counterstained with DAPI (blue). Scale bar, 25 μ m. The sub-panels α , β , γ and δ are close-ups of the areas indicated by the white boxes. **E**) Expression of *Birc5* and *Ccnd1* across pseudotime; colors indicate cell clusters, see Figure 19 for the color key. **F**) In situ hybridization on cryosections of P0 kidneys confirms the expression of *Rspo1* in nephron progenitors and their early epithelial derivatives (α). No signal was detected with sense probe hybridization (β). Images are representative of three independent experiments. Scale bar 25 μ m. (g) Expression of *Rspo1* across pseudotime, similar as E. **H**) Inferred regulatory module activity based on SCENIC²⁶⁰ across pseudotime for self-renewing and primed nephron progenitor cells.

4.5.3.2 *Birc5* expression in the tubular interconnection zone

The cell cycle-related genes *Birc5*, *Ccnd1* and *Tubal1* were up-regulated in immature distal tubules (Figure 19B). Immunostaining analysis confirmed augmented expression of *Birc5* and *CyclinD1* in the distal renal vesicle domain (Figure 20D). Interestingly, increased *Birc5* was also observed in a subset of ureteric bud cells (Supplemental figure 9 and Figure 20D) located in the region of interconnection between the late renal vesicle and the adjacent ureteric bud tips. The fusion between the nephron and the collecting system is required for the formation of a functional renal network. Studies in mouse models have demonstrated that this process is driven by preferential cell division within the distal renal vesicle domain²⁷⁹. Therefore, *Birc5* may contribute to tubular interconnection by regulating proliferation in the late renal vesicle and cell survival in the adjacent ureteric tip cells.

4.5.3.3 Conserved features in mouse and human podocyte development

In the podocyte lineage, the genes most significantly defining the cluster are *Pax8*, *Podxl* and *Nphs1*. *Podxl* and *Nphs1* (in combination with *Synpo*, *Nphs2* and *VEGF-A*) are restricted to a subpopulation of mature podocytes^{200,241,280}, which is largely consistent with our observations (Figure 19B and Figure 21A). In a subpopulation of early podocytes, *WT1* and *Mafb* expression has been reported to overlap with the immature marker *Pax8*^{238,241}, and is expressed in parietal epithelial cells²⁸¹, also consistent with our findings (Figure 19B and Figure 21A). Similar to previous scRNA-seq analysis in human fetal kidney²⁴¹, we observe enrichment in the PDZ domain proteins *Magi2*, *Slc9a3r2* and *Pard3b* in mature podocytes (Figure 21A). We also observe podocyte-specific activity of *Cldn5* (while the claudins *Cldn6* and *Cldn7* are expressed in tubular lineages, Figure 21B). Further, the gene *Sparc* (a cystine-rich matrix-associated protein) and the Tissue-Type Plasminogen Activator *Plat* are expressed specifically in the podocyte lineage, as is *Robo2*, a gene known to be expressed and colocalized with nephrin on the basal surface of mouse podocytes²⁸², while the cell cycle regulator *Gas1* (Growth Arrest Specific 1) is expressed in undifferentiated cells and mature podocytes, but less so in mature tubular cells (Figure 21A). Together, these findings further define the gene expression profile of the podocyte lineage and suggest substantial conservation between mouse and human developing podocytes.

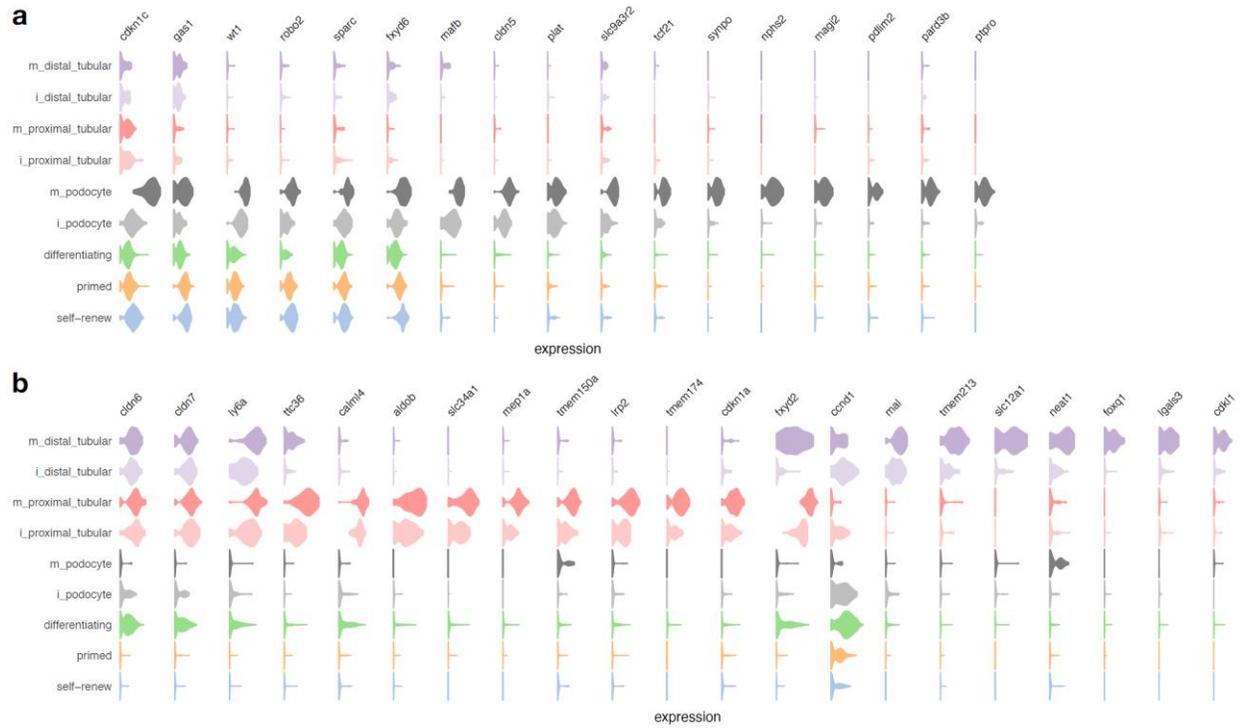


Figure 21. Transcriptional signatures of podocytes and tubular cells.

A) Violin plot of genes expressed in podocytes (rows are clusters and columns denote genes). **B)** Same as (A), but for proximal and distal tubular cells.

4.5.3.4 Gene expression differences between proximal and distal tubular cells

In addition to their respective marker genes *Pdzk1*, *Slc34a1*, *Tmem52b* and *Shd* (see above), we observe that tubular lineages express the claudins *Cldn6* and *Cldn7*, as well as the Lymphocyte Antigen 6 Complex Locus A (*Ly6a*, aka *Sca-1*), Figure 21B. *Ly6a* is a member of the murine L6 family and has been reported to mark cancer and tissue-resident stem cells in mice²⁸³; however there is no known direct human ortholog for *Ly6a*, and also the function of the LU domain, which characterizes *Ly6a*'s superfamily of proteins, is currently unknown in humans²⁸³.

We note that *Mep1a*, *Aldob* and *Tmem174* mark proximal tubular cells in our data (Figure 21B) and have been reported amongst the most down-regulated genes after p53 conditional deletion in nephron progenitor cells²⁸⁴. Of the other three reported top down-regulated genes two (*Pck1* and *Cyp2d26*) also show proximal tubular cell specific expression (data not shown), while *Reg8* expression was not detected in our data. This is in line with the observation of fewer proximal tubular cells in P0 mutant kidneys reported in Li et. al (2015)²⁸⁴. With respect to the cell cycle, we find that Cyclin-Dependent Kinase Inhibitor 1 A (*Cdkn1a*, aka P21) is active specifically in proximal tubular cells, while Cyclin-Dependent Kinase Inhibitor 1 C (*Cdkn1c*, aka P57) is primarily expressed in podocytes (Figure 21A,B). For distal tubular cells, we do not observe a selectively active kinase inhibitor but note that Cyclin-Dependent Kinase-like 1 (*Cdkl1*) is specifically expressed in this cell type. These findings pinpoint lineage-specific gene expression differences between the proximal vs. distal tubular lineages, and they point towards lineage-specific control of the cell cycle across nephron progenitor differentiation.

Recently published scRNA-seq papers have described differences in gene expression across a variety of proximal tubule transcripts and lncRNAs in different sexes in the adult mouse kidney^{285,286}. We observe that female-enriched markers, including *Gm4450*, *Lrp2*, *Sultd1*, *Aadat*,

Hao2 were highly expressed in our proximal tubular cluster, while most of male-enriched markers (*Slc22a12*, *Cndp2*, *Cesf1*, etc.) were absent or expressed at low levels. This data suggests that sexually dimorphic gene expression in proximal tubule may occur at or before E14.5.

4.5.4 Expression of known lineage-marker genes in unexpected cell types

Expression of known lineage-marker genes in unexpected cell types has been reported based on the analysis of scRNA-seq data, for example that stromal cells express *Gdnf*²³⁸. Consistent with this report, we found that nephron progenitor markers (*Six2*, *Cited1*, *Crym*) are expressed in cells in the stromal cluster, and that stromal markers are present in the nephron progenitor cluster (*Meis1*, *Foxd1*, *Crabp1*). We also confirm that *Gdnf* is expressed in the stromal cluster (in addition to nephron progenitor cells), and that *Aldh1a2* RNA is present in stromal and nephron progenitor clusters (Figure 22).

We note that nephron progenitor marker genes are not homogeneously expressed across different stromal cell types. For instance, *Cited1* is detected (five or more reads) in about 7% of cortical stromal cells, but in less than 1% of other stromal cell types. We find similar enrichment (expression in ~7% vs less than 1% of cells) for *Six2* and *Crym* in cortical stroma, whereas *Gdnf* is more modestly enriched in cortical stroma (expressed in ~3% vs less than 1% of cells, respectively). We next focused on cortical stroma and looked at co-expression of nephron progenitor and stromal marker genes in the same cells (binary expression “on” vs. “off”) and find significant positive association between the expression of stromal- and nephron-progenitor genes (Fisher exact test, Table 7). This analysis demonstrates wide-spread co-expression of nephron-progenitor and stromal markers in the same cortical/stromal cells, and on average we observe

higher odds ratios of association for *Coll1a1* expression with nephron progenitor lineage-markers, compared with *Meis1* (Table 7).

Further on, the cluster we identified as distal tubular cells contains cells with a distal-like expression profile, as characterized by the expression of *Tmem52b* and *Shd*²³⁹. However, despite the distinct lineage origins, cells from this cluster and from the *ureteric_bud/collecting_duct* cluster exhibit some transcriptional congruence^{239,287}. Specifically, *Calb1*, *Wdfc2* and *Mal*²⁰⁰, which are thought to mark the ureteric but lineage, and *Mecom*²⁸⁶, which is thought to mark distal tubular cells, are expressed in a significant fraction of cells in both these clusters, but absent in proximal tubular cells (Table 8).

Table 7. Co-expression of stromal marker genes and nephron progenitor marker genes in cortical/stromal cells.

The *cs_gene* and *np_gene* columns show the cortical/stromal and nephron-progenitor marker genes, respectively; the *#cells* column shows the number of cells expressing both genes in the stromal/cortical cluster (which contains 1,085 cells overall), while the *odds_ratio* and *p_value* columns contain odds ratio and *p_value* of a corresponding Fisher exact test.

cs_gene	np_gene	#cells	odds_ratio	p_value
<i>coll1a1</i>	<i>cited1</i>	237	8.32	8.29E-13
<i>coll1a1</i>	<i>crym</i>	260	5.87	1.27E-11
<i>coll1a1</i>	<i>gdnf</i>	407	1.73	1.87E-03
<i>coll1a1</i>	<i>six2</i>	237	4.71	3.00E-09
<i>meis1</i>	<i>cited1</i>	214	2.14	2.12E-04
<i>meis1</i>	<i>crym</i>	238	2.21	4.33E-05
<i>meis1</i>	<i>gdnf</i>	402	2.39	5.94E-08
<i>meis1</i>	<i>six2</i>	221	2.42	1.64E-05

Table 8. Ureteric bud / collecting duct lineage genes are expressed in distal tubular cells.

Shown is the percentage of cells expressing ureteric bud / collecting duct (ub/cd) lineage marker genes *Calb1*, *Mal*, *Mecom* and *Wfdc2* for cells from immature and mature distal and proximal tubular clusters, and also from the ub/cd cluster. We see that in the tubular distal lineage, in contrast to the proximal lineage, a significant fraction of cells express these ub/cd marker genes.

Cluster	<i>Calb1</i>	<i>Mal</i>	<i>Mecom</i>	<i>Wfdc2</i>
i_tubular_prox	3.0	3.0	1.5	48.5
m_tubular_prox	1.6	1.6	4.8	41.3
i_tubular_dist	27.0	77.8	88.9	98.4
m_tubular_dist	14.3	91.8	87.8	100.0
ub/collecting-duct	83.6	52.0	71.0	99.7

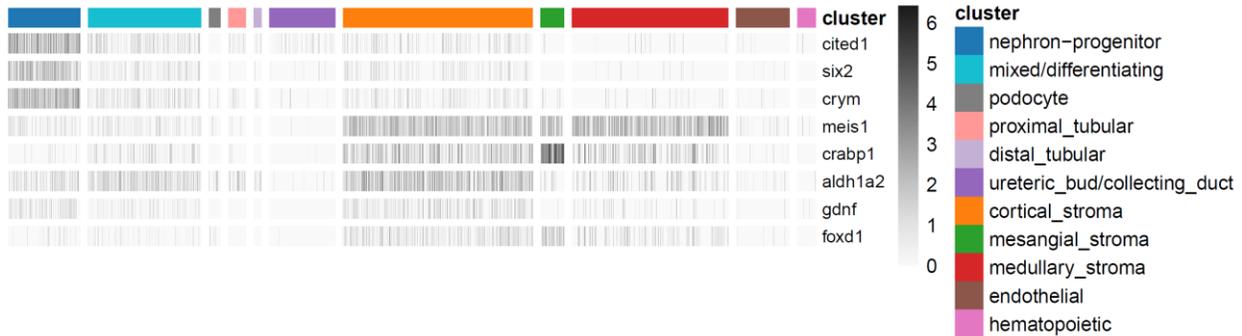


Figure 22. Expression of lineage-marker genes in unexpected cell types.

Heatmap of gene expression (gray scale) of known lineage-marker genes (rows) across cells (columns), ordered by cell clusters (color index). We observe the expression of cap mesenchyme markers (*Cited1*, *Six2*, *Crym*, *Gdnf*) in stromal cells and vice versa, consistent with previous reports²³⁸.

4.6 Discussion

Over 30 terminally differentiated nephron cell types are required for the function of the mammalian kidney. The advent of scRNA-seq technology has made it possible to explore the cellular heterogeneity of the kidney and precisely identify the transcriptional signatures that define each of its cell types. In this study, we have performed scRNA-seq analysis of the developing mouse kidney at E14.5, a time point in which there is active nephron induction and varying degrees of nephron maturation. Major transcriptional clusters—corresponding to nephron progenitors, mixed/differentiating cells, podocytes, differentiated tubules (proximal and distal), ureteric epithelium, stroma (medullary, mesangial, and cortical), hematopoietic and endothelial lineages—are identified within the whole kidney analysis, and are consistent with prior single cell analyses of the developing mouse kidney^{175,237–239}. We find that cell cycle activity distinguishes between “primed” and “self-renewing” sub-populations of nephron progenitors. Furthermore, augmented *Birc5* expression occurs in immature distal tubules and a subset of ureteric bud cells, suggesting that *Birc5* might be a novel key molecule required for early events of tubular fusion between the distal nephron and the collecting duct epithelia.

Birc5 (also known as Survivin) has been implicated in a number of kidney conditions, including autosomal-dominant polycystic kidney disease, acute kidney injury and renal cell carcinomas^{288–292}, however its role in context of normal kidney development is still unknown. In normal tissues, transcription of *Birc5* is tightly regulated in a cell cycle-dependent manner, reaching a peak in the G2/M phase^{293–295}, followed by a rapid decline at the G1 phase²⁹⁶. *Birc5* targets the chromosomal passenger complex (CPC) to the centromere, ultimately enabling proper chromosome segregation and cytokinesis^{297–303}. *Birc5* also plays a role as an inhibitor of programmed cell death. Although this mechanism is not completely understood, it seems to require

cooperation with other molecules (such as XIAP and HBXIP) and results in inhibition of caspase-9^{304–307}. Our data suggest that Birc5 might be a novel key molecule required for early events of nephron patterning and fusion, by regulating cell survival and/or proliferation in late renal vesicle and the adjacent ureteric tip.

All nephron segments derive from a multipotent self-renewing nephron progenitor population, which co-expresses the transcription factor *Six2* and the transcriptional activator *Cited1*. Previous studies have identified two subtypes of nephron progenitors, with *Cited1*⁺/*Six2*⁺ progenitors transitioning to a *Cited1*⁻/*Six2*⁺ primed state as nephrogenesis proceeds^{29,39,45}. Recent studies using time-lapse imaging and scRNA-seq analyses have indicated, however, that the nephron progenitor compartment is more heterogeneous than initially supposed^{26,35,175,238,239,308}. Moreover, differences in cell cycle length within progenitors appear to play a role in the sub-compartmentalization of the progenitor population²⁰. In agreement, our scRNA-seq analysis shows separation of nephron progenitor cells into a “self-renewing” and “primed” subpopulation, both co-expressing *Six2* and *Cited1*, but distinguished by higher cell cycle activity in the “primed” cells. Studies in mice have demonstrated that the committed nephron progenitors are more proliferative, exhibiting preferential exit from the cap mesenchyme compartment and differentiation into early nephrons²⁰. Intriguingly, in the human renal cap mesenchyme, the “self-renewing” nephron progenitors exhibit a greater proliferative activity, compared to the committed progenitor population²⁴². Although it is still unclear what drives these species-specific differences, this may be related to unique transcription factor expression in the human fetal kidney (such as continuous *Six1* expression in cap mesenchyme throughout nephrogenesis)²⁴².

We find that the transcriptional profile of “primed” nephron progenitors represents an intermediate/transitional state between self-renewing NP and mixed/differentiating cells (pre-

tubular aggregates/renal vesicles), with lower levels of *Cited1* and increased expression of early commitment markers like *Lhx1* and markers of renal epithelia like *Pax8*. These findings are consistent with previous scRNA-seq analyses of developing human and mouse kidneys^{200,243}, but are in contrast to other studies on nephron progenitor subpopulations where *Cited1* expression seems to be turned off prior to the activation of pre-tubular aggregate genes^{39,45,308}. Such discrepancies might be due to differences in the technical sensitivity of the methods applied in each study (scRNA-seq versus immunofluorescence or in situ hybridization). They also highlight the importance of further analysis to confirm whether these nephron progenitor subpopulations coincide with distinct spatial domains within the developing kidney.

Our approach successfully identified a number of cell types in the developing kidney. Consistent with previously reported expression patterns, we observe *Podxl*, *Synpo*, *Nphs1*, and *Nphs2* expression in mature podocytes, whereas *WT1* and *Mabf* are also expressed in a subpopulation of early podocytes^{241,280}. Indeed, we also observed several PDZ domain proteins (*Magi2*, *Slc9a3r2* and *Pard3b*)²⁴¹ expressed in human developing podocytes in our data, suggesting that podocyte identity is conserved in the mouse and human developing kidney. Cells in the proximal tubular cluster are characterized by the specific expression of known proximal tubule markers, such as *Pdzk1* and *Slc34a1*^{238,286}. The scaffold protein *Pdzk1* is essential for the proper localization of interacting proteins, such as the sodium-phosphate transporter NaPi-IIa (encoded by *Slc34a1*), in the brush border of the proximal tubular cells³⁰⁹⁻³¹¹. Interestingly, mutations in *Slc34a1* have been linked to nephrocalcinosis and Fanconi renal tubular syndrome^{312,313}. Further on, we observe that *Cldn5* marks the podocyte lineage, while *Cldn6* and *Cldn7* are expressed in mixed/differentiating cells and both tubular lineages, but absent in

podocytes. Genes specifically expressed in distal tubular cells in our data include Galectin 3 (*Lgals3*), *Slc12a1* and the long non-coding RNA *Neat1*.

The formation of a fully functional nephron entails fusion between the late renal vesicle and the adjacent ureteric tip. An elegant study using 3D modeling of nephrons and Six2-eGFP-Cre x R26R-lacZ mice demonstrated that this connecting segment of the nephron is derived from the cap mesenchyme (not the ureteric epithelium), and the process of fusion is likely driven by preferential cell division within the distal renal vesicle domain²⁷⁹. In line with this, our data identified augmented expression of cell cycle-related molecules, such as Cyclin D1 (*Ccnd1*), *Birc5* and *Tuba1a*, in immature distal tubules. Interestingly, high *Birc5* expression was also detected in ureteric bud cells located in the region where the ureteric tip connects with the distal portion of the renal vesicle (see Supplemental figure 9).

In line with other scRNA-seq studies, we identify three stromal clusters in our dataset: cortical, medullary and mesangial^{238,263}. The genes most significantly defining the mesangial stroma cluster are *Dlk1* and *Postn*^{239,263}. Cells in cortical stroma express high levels of *Aldh1a2* and *Dlk1*, while medullary stroma cluster contains cells with increased expression of *Cldn11*. The absence of an expression profile consistent with a loop of Henle signature in our scRNA-seq data is likely due to a low-abundance of these cell populations at E14.5³¹⁴. In addition, the lack of information on the cell diversity and identity within the loops of Henle continues to hinder the annotation of this segment²⁸⁶.

In summary, this study provides an in-depth transcriptional profile of the developing mouse kidney at mid-gestation. Major main transcriptional clusters are identified, and are consistent with prior single cell analyses of the developing mouse kidney^{175,238,239,263}. Notably, we find that cell cycle activity distinguishes between the “primed” and “self-renewing” sub-populations of nephron

progenitors, with increased levels of the cell cycle related genes *Birc5*, *Cdca3*, *Smc2* and *Smc4* in the “primed” sub-population. Finally, *BirC5* expression in immature distal tubules and ureteric bud cells may contribute to early events of tubular fusion between the distal nephron and the collecting duct epithelia.

5.0 Discussion and future directions

5.1 miRNA expression contributes to age-dependent changes in nephron progenitors

Our work in chapter 2.0 demonstrated that miRNA expression in nephron progenitors differs from miRNA expression in the surrounding kidney, suggesting this is an important facet of the nephron progenitor cell type's identity and transcriptome. In chapter 3.0 we showed that this transcriptional signature is not static over time but changes with progenitor age, with 114 different miRNA undergoing a significant change in expression between E14.5 and P0 (Figure 7). Key among these changes were not only the increasing expression levels of the *let-7* family members, but also the decreased expression of miRNA in the miR-200 family (*miR-429-3p* and *miR-200b-3p*), which is known to inhibit epithelial-to-mesenchymal transitions^{219,220}. In chapter 2.0 we also noted that miR-200 family members exhibit lower expression in nephron progenitors than they do in the rest of the kidney. Repressing the miR-200 family likely promotes the mesenchymal fate of progenitors. However, it remains unclear why a measurable decrease in expression of miR-200 family members occurs as the rate of progenitor differentiation increases with the age of the progenitor population. Pathway analysis of predicted miRNA gene targets⁶² implicated changing miRNA in modulation of pluripotency and cell differentiation: both increasing and decreasing miRNA are enriched for gene targets involved in signaling pathways FoxO (which is downstream of mTOR signaling to promote proliferation³¹⁵) and MAPK (integral for interpretation of juxtacrine signals for differentiation in nephron progenitors²¹⁰). Interestingly, genes associated with the “extracellular matrix interactions” KEGG pathway are far more likely to be targeted by an increasing miRNA than a decreasing one, suggesting that the changes to the miRNA

transcriptome have a distinctly repressive effect as progenitors age. The functions of the ECM are varied and include conveying juxtacrine signals, accumulating growth factors, and enabling cell engraftment and migration^{211,316}. In nephron progenitor cells, expression of ECM genes including *Flrt3*, *Fnl*, and *Thbs1* are diminished as progenitor cells age, and this is thought to contribute to the difficulty older progenitors exhibit when engrafting to other cells in the niche, as well as affect their arrangement among the nephron progenitor population as they approach differentiation⁴⁷. *Flrt3* in particular responds to FGF signaling³¹⁷ and allows for homotypic cell-cell adhesion³¹⁸, so its reduced expression could diminish the influence of *Fgf20* on the progenitor cell as well as mitigate its migration through the cap mesenchyme by reducing its adherence to the surrounding progenitor cells. Gradually increasing the expression of miRNA to target genes in the ECM could potentially affect progenitor differentiation rates by altering the progenitor cell's migration and sorting behavior, which plays a direct and important role in the population-wide rate of nephron progenitor commitment³⁵. It may be possible to test this idea by observing the migration and sorting behavior of young and old progenitors with and without exposure to a cocktail of miRNA predicted to influence this pathway⁶². The ECM receptor pathway-associated miRNA most expressed and most significantly increasing in our smRNA-seq data include *miR-196a-5p*, *miR-148a-3p*, *let-7f-5p*, and *miR-125a-5p*. Real-time cell tracking of nephron progenitor cells in the cap mesenchyme has been demonstrated³⁵, and so ectopic expression of one or more of these key miRNA species with a constitutively active promoter may mitigate some of the differences normally observed in progenitor migration or engraftment. Similarly, transplantation of these progenitors may reveal an improvement in the percentage of cells that successfully engraft.

5.2 Chromatin accessibility contributes to age-dependent changes in nephron progenitors

Along with the changes in nephron progenitors' miRNA expression, our work in chapter 3.0 showed that concurrent changes in chromatin accessibility can be observed as well. We developed a custom NextFlow data processing pipeline to allow parallelized alignment, quantification, and quality control of ATAC-seq data that complies with ENCODE standards for ATAC-seq libraries¹⁶⁵. Our work demonstrated that wild-type nephron progenitors isolated at E14.5 and P0 can be clearly distinguished by their ATAC-seq signals, and we noted 46,374 regions deemed accessible using the irreproducible discovery rate as a conservative standard¹⁸⁸. Among these, 2,103 regions exhibit a significant increase or decrease in accessibility as nephron progenitors age. Many of these changes can be tied to known enhancer elements, including the enhancer of the *Gdnf* gene. This enhancer has been shown to be upregulated by Wnt signaling³³, and our data suggests that this increase is reflected in an increase in accessibility of the chromatin around this enhancer (Figure 14D). The genome-wide nature of our ATAC-seq data set allows us to identify thousands of unannotated regions of accessibility, many of which are both conserved and exhibit signs of transcription factor binding or changes to the local nucleosomal configuration. Enrichment of genomic annotations using the Genomic Region Enrichment of Annotations Tool (GREAT)¹⁹⁰ suggests that chromatin regions that see either an increase or decrease in levels of accessibility and activity are particularly common among genes controlling stem cell differentiation and chemotaxis. Regions that see an increase in accessibility are very often near genes that affect cell fate commitment, chemotaxis, migration, and cell fate specification, as well as several ECM-associated components such as adherens and anchoring junctions. In particular, opening chromatin regions are enriched for Notch signaling as well, which could indicate that *Six2*'s gradual reduction in expression over the course of nephron differentiation^{41,319} is in part

attributable to enhancer changes that facilitate Notch signaling. Moreover, that this effect is seen between early and late progenitors could imply that reduced expression of *Six2* is not just a consequence of differentiation but also part of progenitor aging. However, this may also be a consequence of an increased fraction of committed progenitors isolated at later time points, and moreover an increase in accessibility associated with Notch signaling may in fact be an *effect* of increased Notch signaling rather than a cause.

Regions of chromatin that see any change—either increasing or decreasing accessibility—are most significantly associated with GO annotations associated with the cell cycle: GO processes including “Mitotic cell cycle”, “Cell cycle process”, and “DNA metabolic process” are among the most significantly enriched, as are GO cellular components “Centrosome” and “Microtubule organizing center.” Rates of cell cycling and the effective lengths of the cell cycle have been suggested to differ between self-renewing and differentiated nephron progenitors previously²⁰, and our findings could implicate changes in the chromatin landscape with modulation of these functions. Such functional claims need to be verified, however, as changes in accessibility over a putative regulatory feature are not necessarily causative and may in fact reflect changes to the wider chromatin environment. While the annotation of specific accessible regions as enhancers with specific gene targets will take time, our analysis suggests that they affect many of the key processes that drive nephron progenitor behavior within its niche. Future work should benefit from the availability of transcription factor footprinting data that our ATAC-seq analyses make available when deciding which regulatory features are worth investigating for enhancer activity.

5.3 Gene expression reflects heterogeneity among nephron progenitor cells

Our efforts in chapter 4.0 to perform single-cell sequencing of the E14.5 mouse kidney revealed a great deal about the heterogeneity of the cells in the kidney, in particular the nephron progenitor population. We were able to separate the progenitor population into “self-renewing” and “primed” populations based on their expression levels of *Six2* and *Cited1* and noted several developmental genes that are more highly expressed in differentiating progenitors than self-renewing. Crucially, we note that a considerable amount of heterogeneity exists in the differentiating progenitor population. This is to be expected as progenitors must differentiate into podocyte, proximal tubular, and distal tubular cell types, and differentiating cells are likely found in different stages of the differentiation process at the time of isolation. Gene expression in our single-cell data also reveals that early and late nephron progenitors are clearly able to be differentiated by cell cycle associated genes such as *Birc5*, *Cdca3*, and *Smc4*, which is in agreement with published findings that cell cycle changes are a major differentiator between self-renewing and differentiating nephron progenitors⁸⁷, as well as our findings from chapter 3.0 that implicate many of our measured chromatin changes with changes to progenitor proliferation and cell cycle.

Having measured miRNA expression changes with progenitor age in chapter 3.0, measuring their predicted effects through the targeting of lineage marking genes could reveal miRNA-induced transition preferences: whether older or younger progenitors have changing penchants for differentiation into one cell type or another). Expression of miRNA targeting highly expressed transcripts introduces transcriptional noise⁵⁶, which a current pre-print suggests is used to generate variability among homogenous stem cell populations⁵⁷. The data we have so far generated goes a long way toward testing whether any miRNA play such a role in nephron

progenitors: it identifies miRNA that are highly expressed in late nephron progenitors, genes that are highly expressed and / or exhibit high-variance in expression in differentiating progenitors, and chromatin accessibility data describing accessibility and transcription factor footprints for potential noise-inducing genes.

Given that canonical miRNA processing requires exporting of the miRNA molecule from the nucleus, the presence of miRNA within the nucleus of mammalian cells alludes to a host of non-canonical functions for these nuclear miRNA, mechanisms for which are hypothesized to include miRNA interactions with nuclear RNAs, promoters, and DNA³²⁰. A 2016 study (Xiao et. al) demonstrated that the miRNA *miR-24-1* is present in the mammalian nucleus as well as the cytoplasm, and indicated that nuclear *miR-24-1* exhibits transcription factor-like activity by activating enhancers and causing an increase in expression of target genes. Moreover, this activity increase is predicated on the integrity of the *miR-24-1* seed sequence as well as a target sequence within the enhancer³²¹. Having measured changing miRNA expression levels and chromatin accessibility in parallel, our data appears uniquely situated to describe similar non-canonical miRNA: seed sequences for miRNA that increase or decrease expression over time can be identified, and chromatin regions that increase or decrease in accessibility (to include not only distal regions but also gene promoters) can be searched for their target sequences using software developed for this purpose³²⁰. By cross-referencing identified promoters and enhancers with gene expression differences between renewing and differentiating progenitors, it may be possible to broadly describe non-canonical miRNA gene activation mechanisms. Our data do not discern nuclear versus cytoplasmic miRNA concentrations, however, which is an important caveat for a miRNA's ability to function as a transcription factor. As such, these potential mechanisms would

need to be cross-referenced with measurements of nuclear miRNA transcripts in nephron progenitors.

5.4 Enhancers of miRNA

Having confirmed the importance of miRNA expression in chapter two and identified parallel changes in miRNA expression and chromatin accessibility over time during nephrogenesis, we sought to identify enhancers that could modulate the expression of these miRNA with the aging of the progenitor cell. Using available TAD annotations¹⁶⁶ to identify changing accessible regions and miRNA that share the same domain, we then filtered to include only changing miRNA and enhancers which exhibit the same relative changes over time: enhancers with increasing accessibility that share a TAD with miRNA that increase expression, and vice versa. Our conservative screen filtered changing regions of chromatin accessibility down to 24 otherwise unannotated regions that match changes with a miRNA (Figure 16), and from among these we were able to identify promising candidates for *in vivo* assays based on the anticipated value of the miRNA of interest and also promising details of the chromatin environment of the enhancer visible in our ATAC-seq signal, such as the presence of transcription factor footprints with motifs attributed to nephrogenesis-linked transcription factor. Future efforts to confirm these miRNA-enhancer dependencies can be accomplished both *in vivo* and *in vitro*. Confirmation of miRNA expression changes should be possible using simple progenitor isolation methods followed by quantitative PCR, but confirming the enhancer activity of a chosen genomic region requires a reporter assay. One such assay has been published³²² that allows for rapid enhancer activity testing, and should be readily adaptable to testing for mouse enhancers: the

putative enhancer's DNA sequence is cloned into a plasmid along with a weak promoter driving a fluorescent gene reporter, and after transfection into an appropriate cell type true enhancers should drive expression of the fluorescent reporter. Nephron progenitor cells have proven capable of surviving in a multipotent state for extended periods of time given the right growth factor signals⁴⁵, and recent advances allow for them to be incorporated into three-dimensional beads *in vitro*^{246,323,324}, so it may be possible to isolate and passage wild-type nephron progenitor cells for use in this assay. Alternatively, *Xenopus laevis* embryos are both translucent and comparatively easy to transfect, and nephrogenesis in *Xenopus* is sufficiently similar to mammalian nephrogenesis that they serve as a common model system. Transfecting into *Xenopus* has the added benefit of allowing fluorescent reporters to be observed over time, perhaps demonstrating the effects of transcription factors in solution on time dependent enhancer activity.

Building a case for enhancer-miRNA dependencies genome-wide could benefit from additional information describing the chromatin environment. Hi-C data would presumably reveal a TAD layout to similar to that of the mouse ESCs publicly available¹⁶⁶, but it may also reveal sub-TAD architecture through chromatin interactions that place miRNA and enhancers loci together three-dimensional space, lending credence to potential interactions as well as revealing others that our screening method from chapter 4.0 may have proven too stringent to identify³²⁵. Confirming that our data's measured increases in accessibility correlate with H3K27ac histone marks for activity, as well, could also improve our map of the nephron progenitor's chromatin landscape³²⁶. However, a key advantage to ATAC-seq was the fairly minimal requirements for biological starting material, isolating sufficient progenitor cells to be adequately assayed through these methods could be a limitation.

5.5 Nephron progenitor cells

Our efforts throughout chapters 2.0, 3.0, and 4.0 have built upon our prior understanding of the nephron progenitor population by elaborating on the mechanisms that drive their behavior changes over time. We confirmed the importance of miRNA expression in the nephron progenitor's regulatory programs, and demonstrated more than one hundred individual miRNA whose expression appears to change along with the age of the nephron progenitor. At the same time, we observed chromatin changes in matched wild-type nephron progenitors which demonstrate that many of the changes that explain the differences in behavior between early and late progenitors may be the result of a changing regulatory landscape. Combined we see a picture of a cell which transforms over the course of nephrogenesis, responding to an array of inputs from its surroundings and accumulating changes as it does so, either because of or alongside these environmental cues. Such changes are visible in both the miRNA transcriptome and the chromatin environment, and their consequences are visible in the heterogeneity of the differentiating nephron progenitor population. By understanding this dynamic nature of the nephron progenitor cell, future work will be better equipped to translate between the dynamic systems that orchestrate the kidney's development and its fitness after birth, and perhaps even to intervene when these systems go awry with the delicate touch required for preventative treatment.

6.0 Non-thesis work

2018-NCBI Hackathon Pittsburgh 2018 -Team “Viruspy”

2018-Best poster presentation - Biomedical Graduate Students Association Symposium,
University of Pittsburgh

2019-2020-NIH NIDDK T32 Training Grant Recipient

2019-American Society of Nephrology (ASN) 2019 Kidney STARS recipient

2019-American Society of Nephrology (ASN) oral abstract presentation

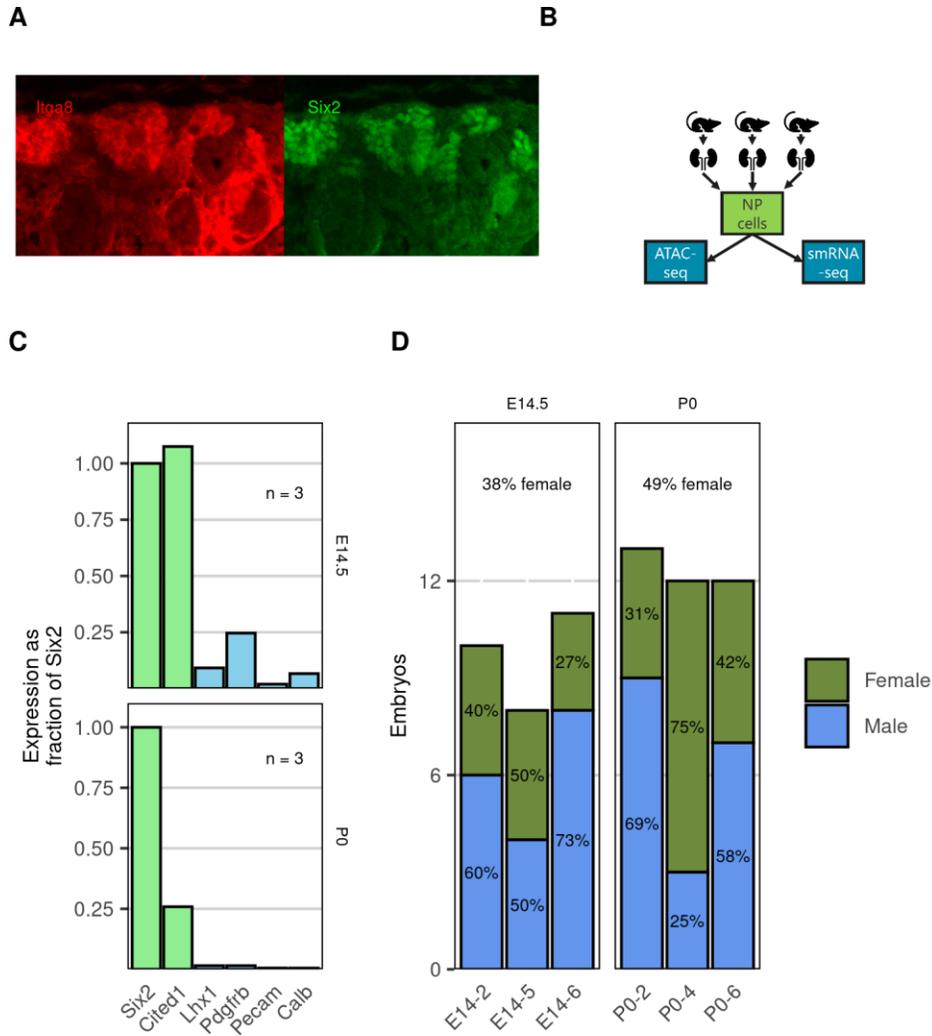
2020 - NCBI Hackathon Pittsburgh 2020-Team “OMOPomics”

7.0 Publications

- a. Cargill, K., Hemker, S. L., Clugston, A., Murali, A., Mukherjee, E., Liu, J., ... Sims-Lucas, S. (2019). **Von hippel-lindau acts as a metabolic switch controlling nephron progenitor differentiation.** *Journal of the American Society of Nephrology*, 30(7), 1192–1205. <https://doi.org/10.1681/ASN.2018111170>
- b. Hemker, S. L., Cerqueira, D. M., Bodnar, A. J., Cargill, K. R., Clugston, A., Anslow, M. J., ... Ho, J. (2020). **Deletion of hypoxia-responsive microRNA-210 results in a sex-specific decrease in nephron number.** *FASEB Journal*, 00(November 2019), 1–18. <https://doi.org/10.1096/fj.201902767R>
- c. Phua, Y. L., Chen, K. H., Hemker, S. L., Marrone, A. K., Bodnar, A. J., Liu, X., ... Ho, J. (2019). **Loss of miR-17~92 results in dysregulation of Cftr in nephron progenitors.** *American Journal of Physiology - Renal Physiology*, 316(5), F993–F1005. <https://doi.org/10.1152/ajprenal.00450.2018>
- d. Phua YL, Clugston A, Chen KH, Kostka D, Ho J. **Small non-coding RNA expression in developing mouse nephron progenitor cells.** *Sci Data*. 2018;5:180218. doi:<https://doi.org/10.1038/sdata.2018.218>
- e. Sathish Babu Vasamsetti, Emilie Coppin, Xinyi Zhang, Jonathan Florentin, Sasha Koul, Matthias Götberg, Andrew S. Clugston, Floyd Thoma, John Sembrat, Grant C. Bullock, Dennis Kostka, Claudette M. St. Croix, Ansuman Chattopadhyay, Mauricio Rojas, Suresh Mu, P. D. (2020). **Apoptosis of hematopoietic progenitor-derived adipose tissue resident macrophages causes insulin resistance after a remote organ injury.** *Science Translational Medicine*.

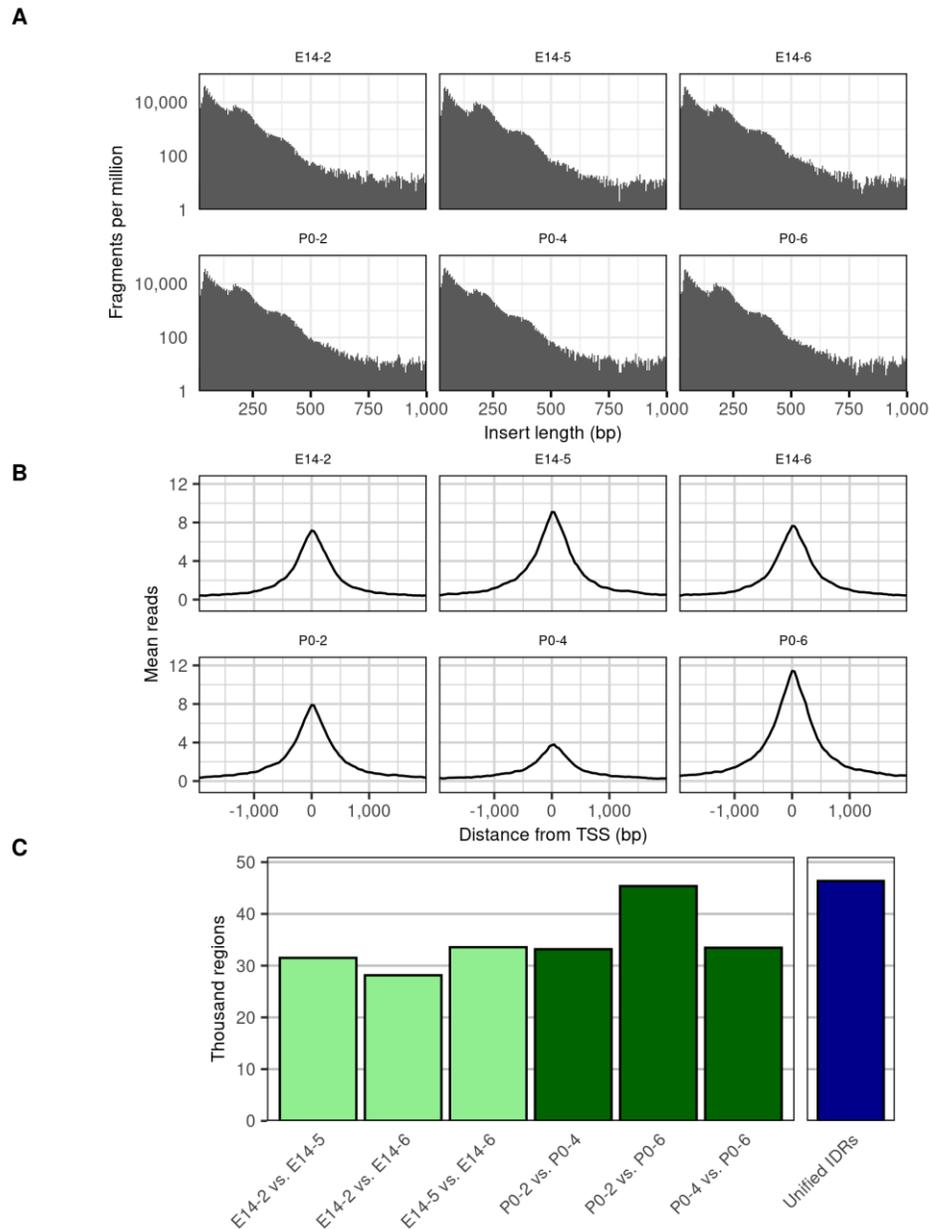
Appendix

Supplemental figures



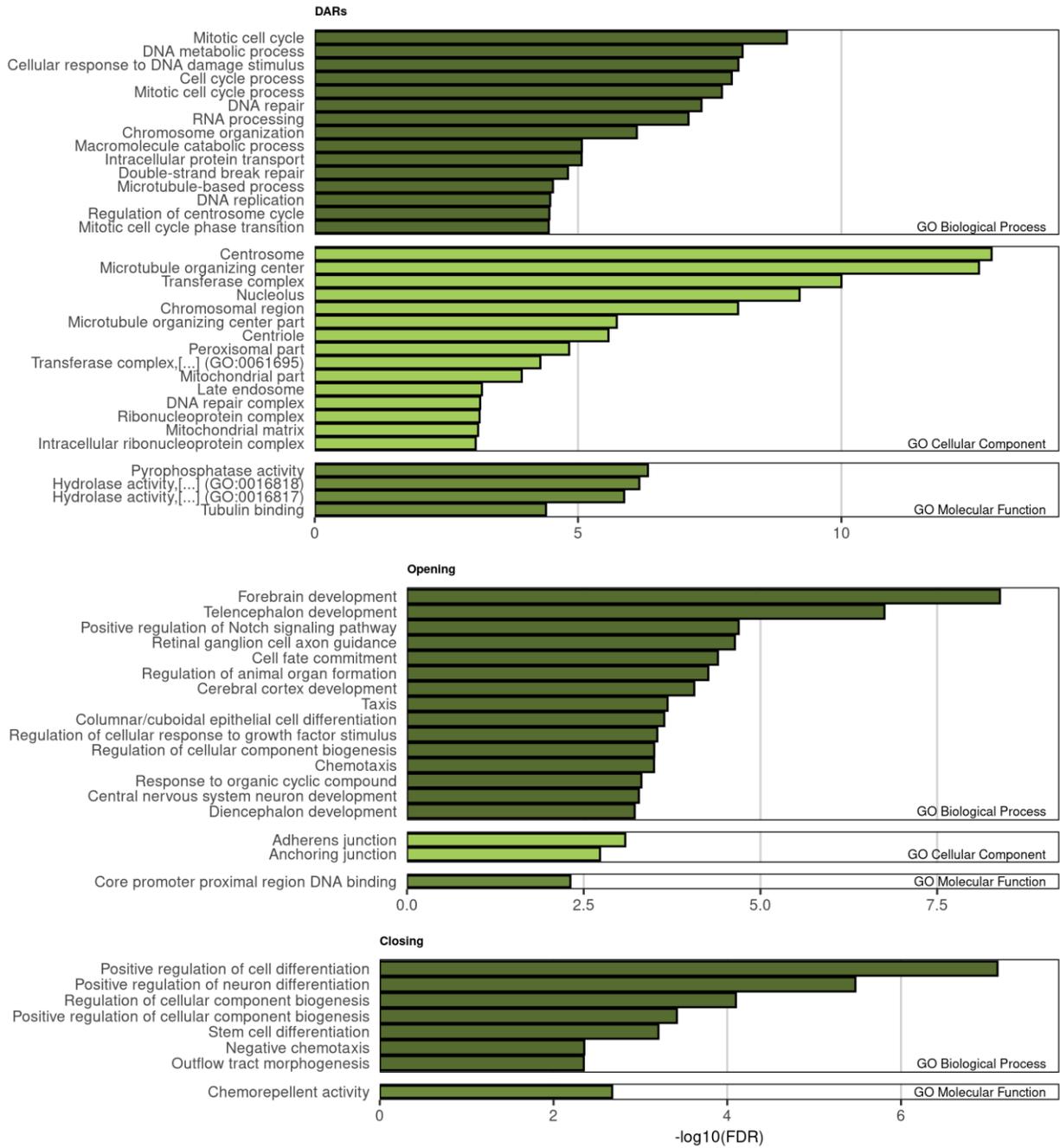
Supplemental figure 1. Nephron progenitor isolation.

A) Fluorescent in situ hybridization shows co-expression of Six2 and Itga8 surface markers in developing mouse NPs. **B)** Nephron progenitors were isolated from pooled kidneys collected from the same litter at E14.5 and P0. **C)** Quantitative PCR (qPCR) was used to confirm enrichment of NP-specific markers Six2 and Cited1 relative to markers for ureteric bud (Calb), renal vesicle (Lhx1) renal stroma (Pdgfrb), and endothelial (Pecam) cell types. **D)** The sex of each embryo was determined by PCR and tallied for each sample pool.



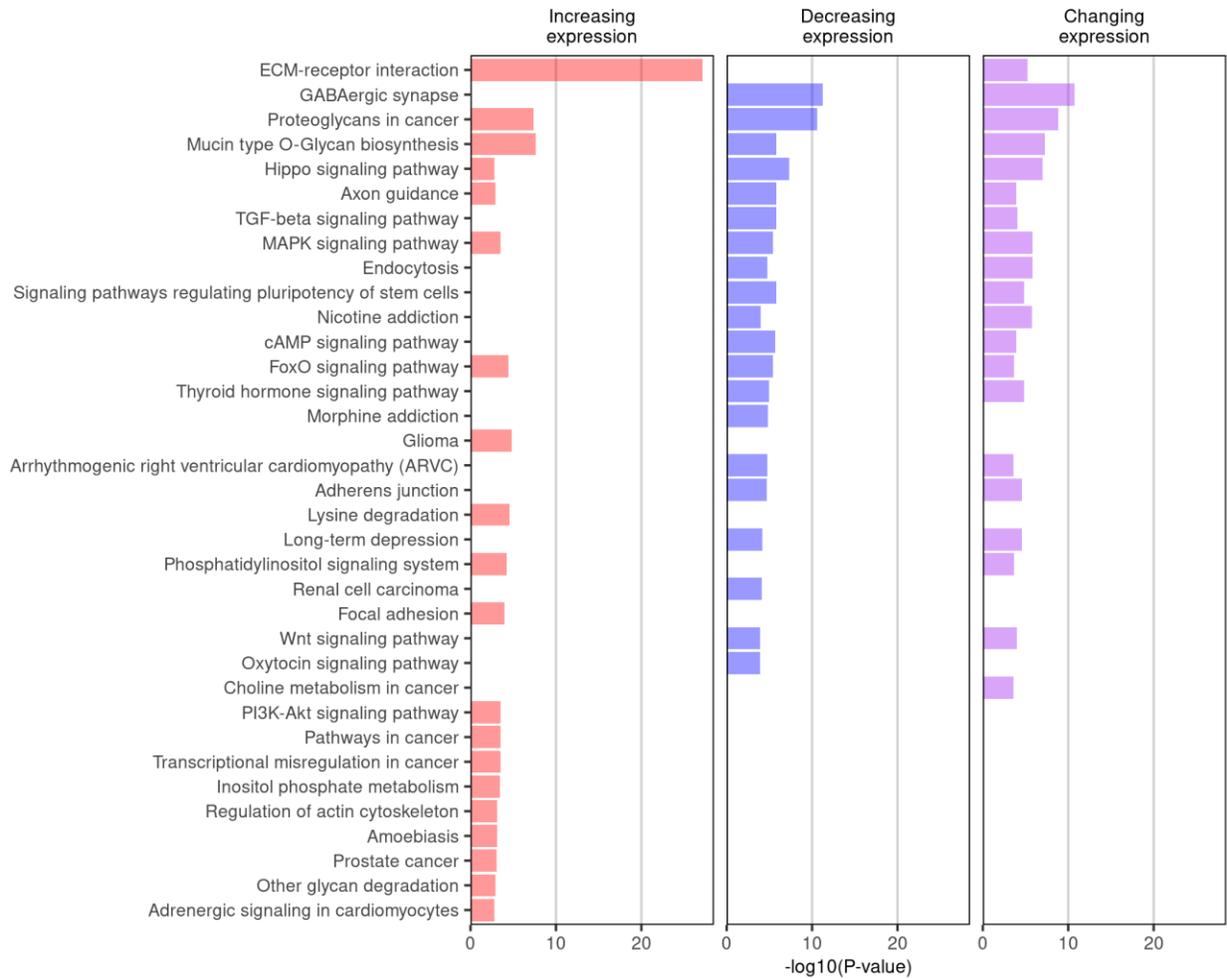
Supplemental figure 2. ATAC-seq quality control.

A) The frequency of read lengths resulting from ATAC-seq exhibit periodicity corresponding to sub-nucleosomal, mono-nucleosomal, and di-nucleosomal distances. **B)** Tn5 insertion events are enriched at transcription start sites (TSSs). **C)** Of all peaks in ATAC-seq signal identified by MACS2, only those which are consistent between at least one pair of replicates using the irreproducible discovery rate (IDR; FDR = 0.1) are considered. The total number of IDR peaks identified in pairwise comparisons of peak sets from each sample are shown in green, and the union of all such IDR peaks is shown in dark blue.



Supplemental figure 3. Genomic Region Enrichment of Annotations Tool (GREAT) analysis results.

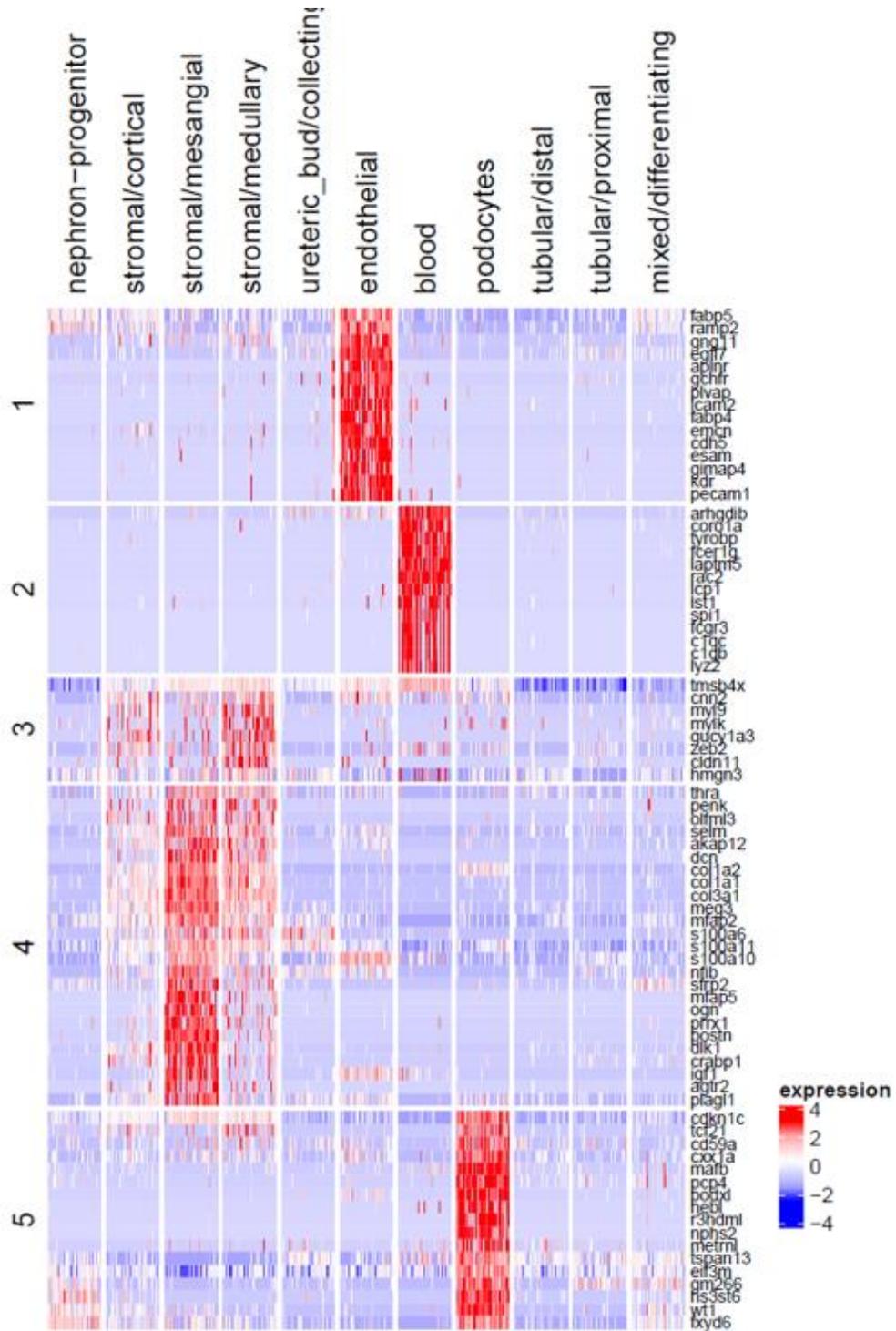
Sets of DARs were subjected to GREAT analysis to identify enriched GO terms. Sets submitted included all changing regions (DARs), all opening DARs (Opening), and all closing DARs (Closing).

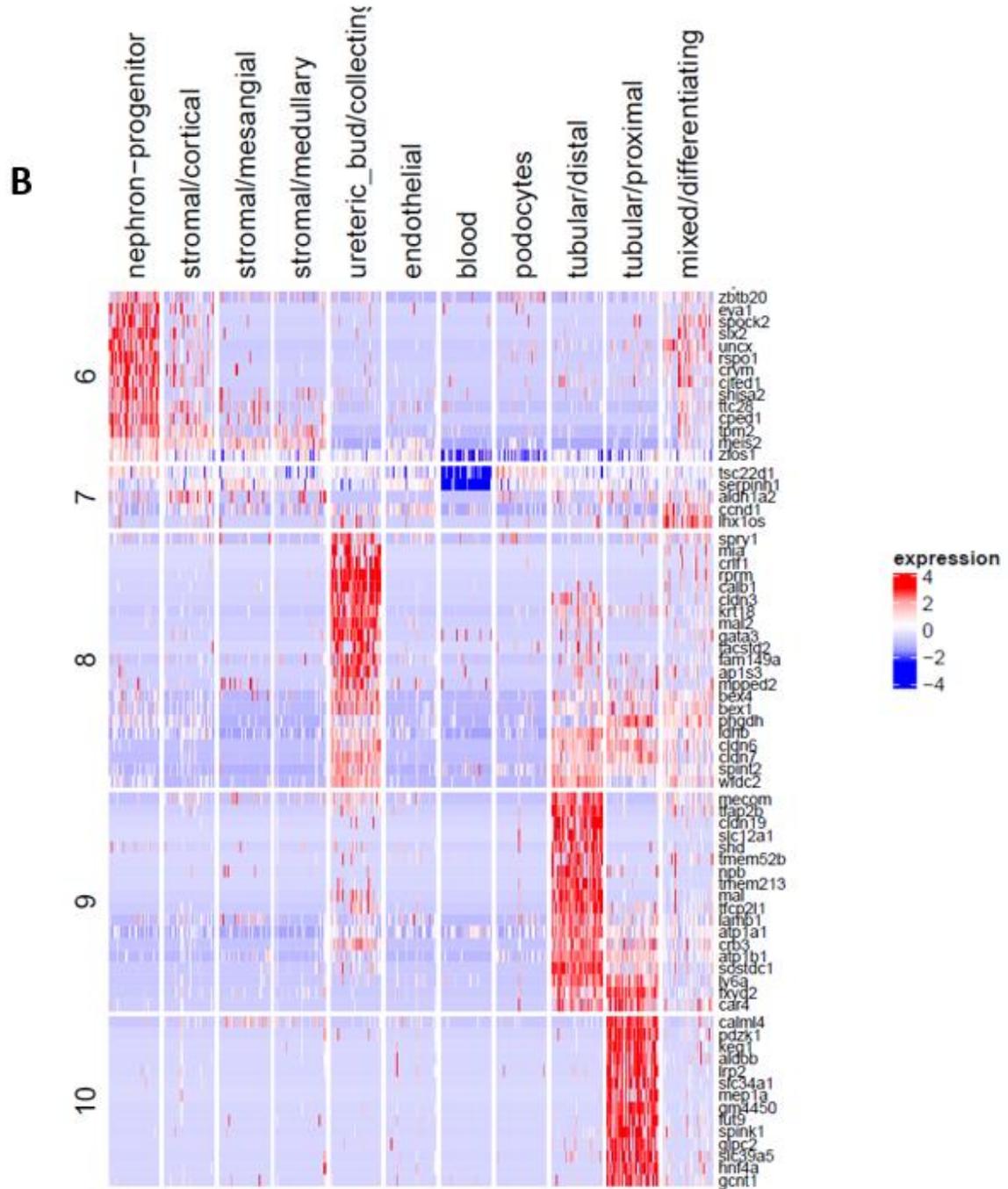


Supplemental figure 4. DIANA miRPath results.

DIANA miRPath results for sets of increasing, decreasing, and changing (increasing and decreasing) miRNA.

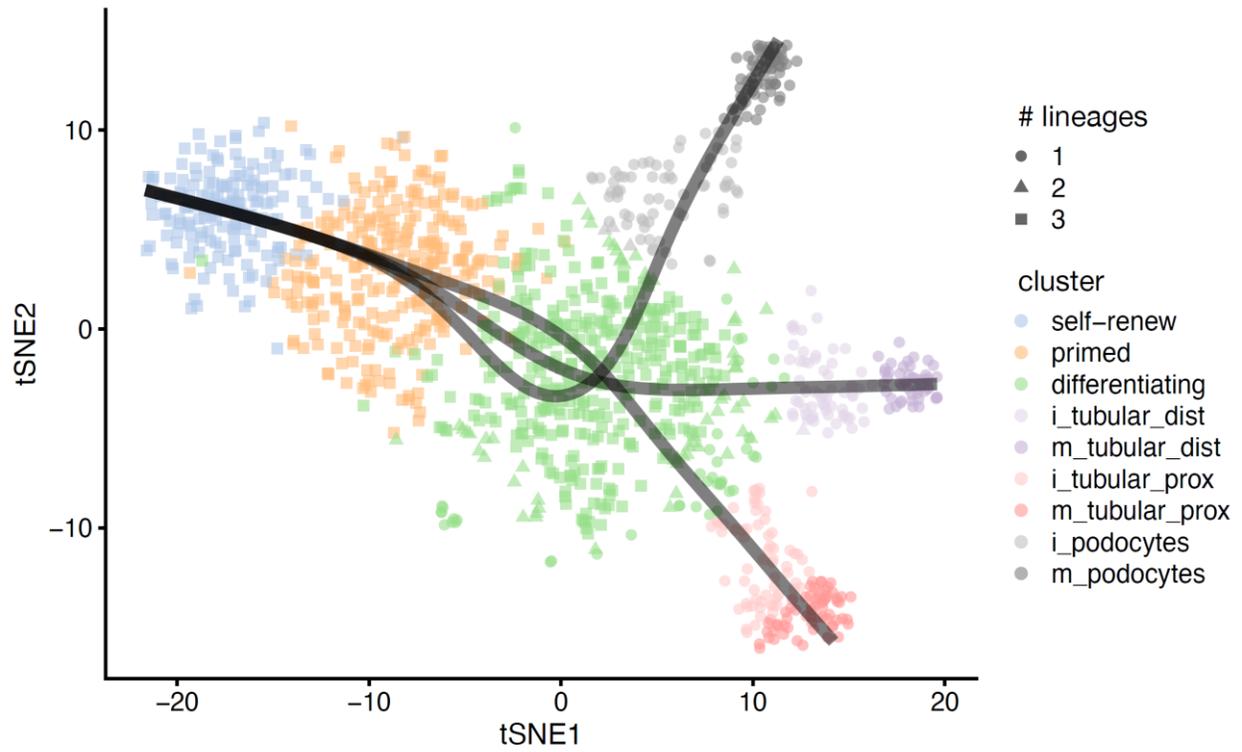
A





Supplemental figure 5. Heatmap of marker genes that distinguish between kidney cell types.

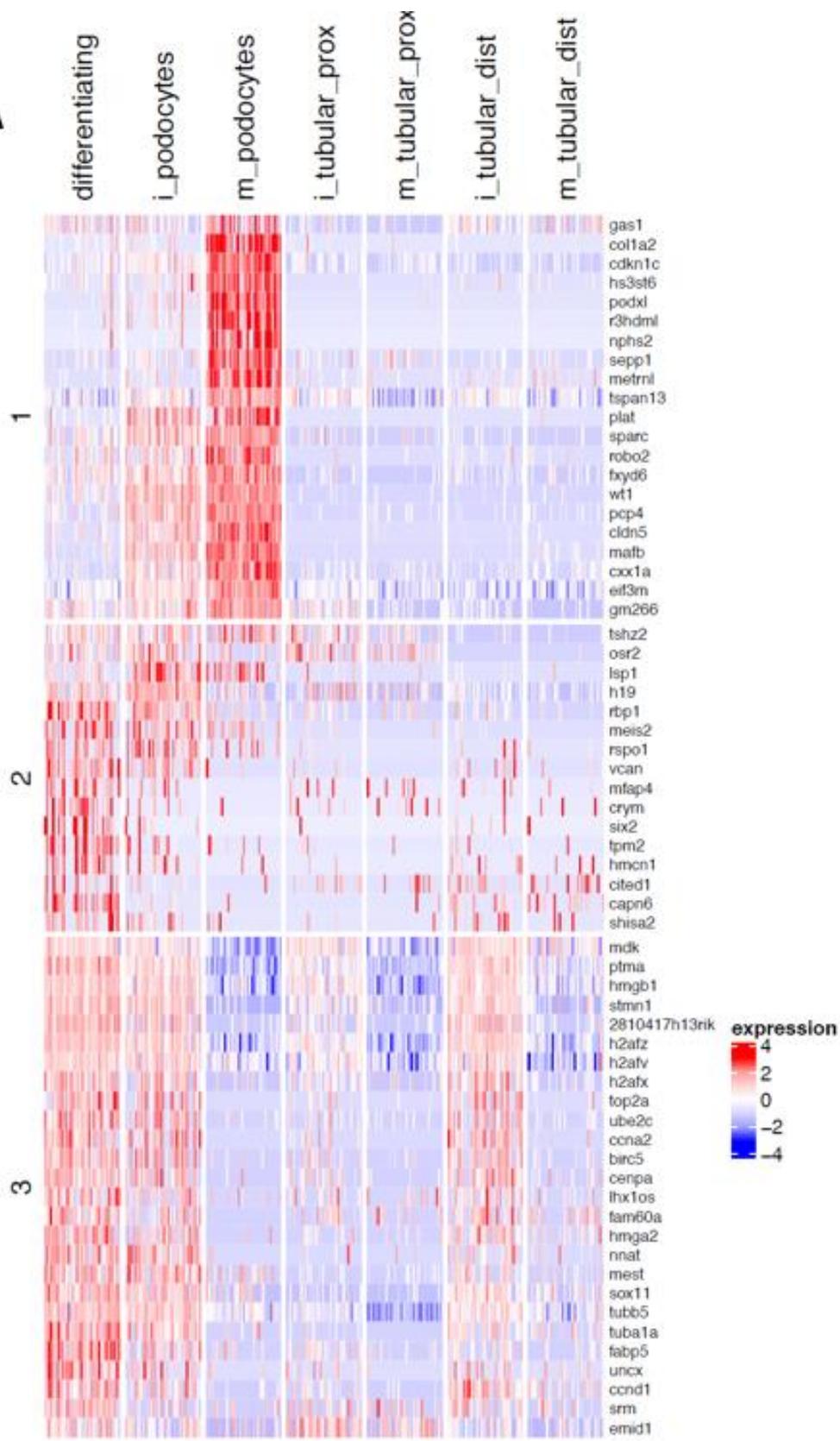
Rows are genes (fifteen top-most marker genes have been selected for each cluster), and columns are cells grouped by cell-types. **A)** Clusters 1-5. **B)** Clusters 7-10.

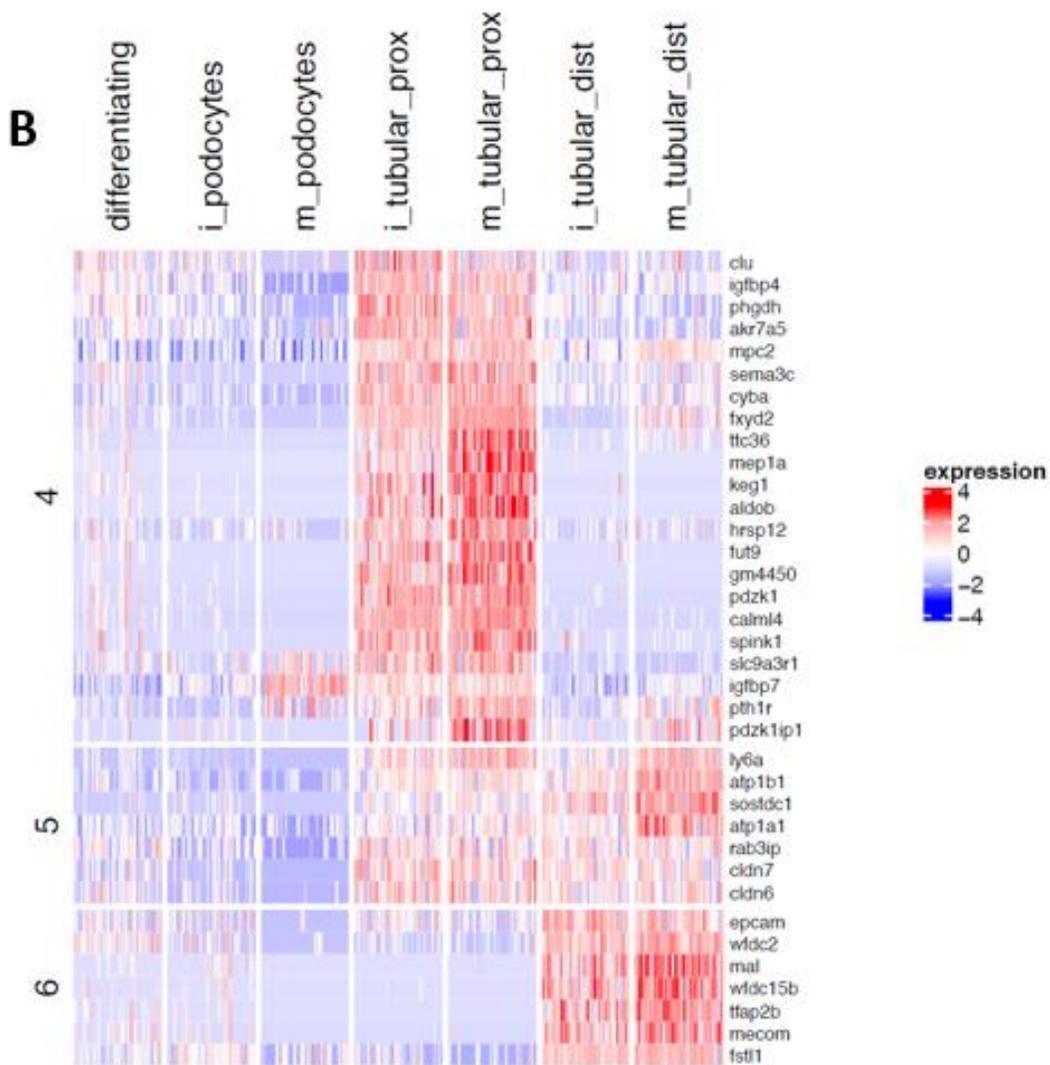


Supplemental figure 6. Differentiation lineages for nephron progenitor cells.

tSNE embedding for nephron progenitor cells is shown. Differentiation lineages inferred by the slingshot R package are displayed in gray.

A





Supplemental figure 7. Differentially expressed genes during nephron progenitor cell differentiation.

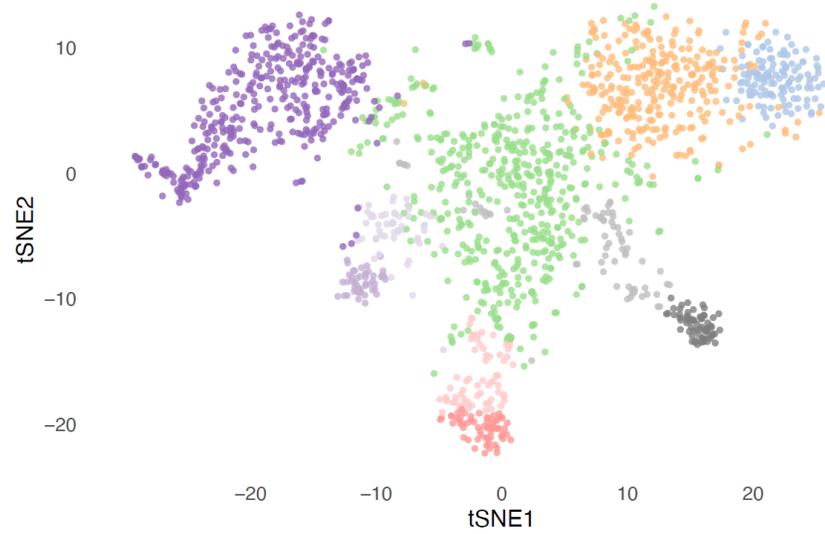
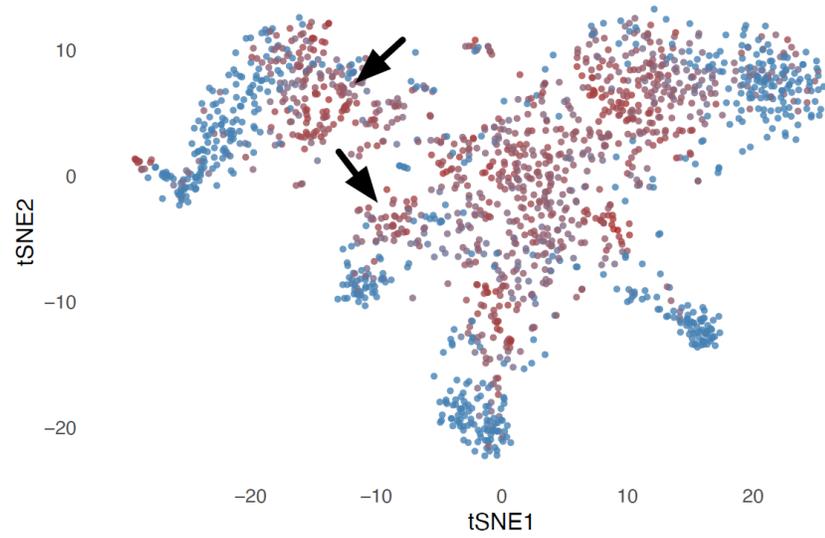
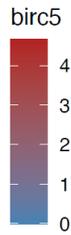
Heatmap of top differentially expressed genes between annotated clusters (column labels). **A)** Clusters 1-3. **B)**

Clusters 4-6.

a

cluster

- differentiating
- i_podocytes
- i_tubular_dist
- i_tubular_prox
- m_podocytes
- m_tubular_dist
- m_tubular_prox
- primed
- self-renew
- ureteric_bud/collecting_duct

**b**

Supplemental figure 9. Consistent *Birc5* expression in distal tubular cells and a subpopulation of cells from the ureteric bud.

Shown are tSNE plots of nephron progenitor derived cells and cells of the ureteric bud / collecting duct (UB/CD) cluster (see Figure 18). **A)** Cell-type annotations for depicted cells. **B)** *Birc5* expression. Arrows denote early distal tubular cells and UB/CD cells with similar *Birc5* expression.

Bibliography

1. Aboul Mahasen LM. Evolution of the Kidney. *Anat Physiol Biochem Int J*. 2016. doi:10.19080/apbij.2016.01.555554
2. Hughson M, Farris AB, Douglas-Denton R, Hoy WE, Bertram JF. Glomerular number and size in autopsy kidneys: The relationship to birth weight. *Kidney Int*. 2003. doi:10.1046/j.1523-1755.2003.00018.x
3. Quaggin SE. Kindling the Kidney. *N Engl J Med*. 2016. doi:10.1056/NEJMcibr1514351
4. Hinchliffe SA, Sargent PH, Howard C V., Chan YF, Van Velzen D. Human intrauterine renal growth expressed in absolute number of glomeruli assessed by the disector method and cavalieri principle. *Lab Investig*. 1991.
5. Bertram JF, Douglas-Denton RN, Diouf B, Hughson MD, Hoy WE. Human nephron number: Implications for health and disease. In: *Pediatric Nephrology*. Vol 26. ; 2011:1529-1533. doi:10.1007/s00467-011-1843-8
6. Kriz W, LeHir M. Pathways to nephron loss starting from glomerular diseases - Insights from animal models. *Kidney Int*. 2005. doi:10.1111/j.1523-1755.2005.67097.x
7. Luyckx V, Shukha K, Brenner BM. Low Nephron Number and Its Clinical Consequences. *Rambam Maimonides Med J*. 2011;2(4):e0061. doi:10.5041/rmmj.10061
8. Little MH, McMahon AP. Mammalian kidney development: Principles, progress, and projections. *Cold Spring Harb Perspect Biol*. 2012;4(5):3. doi:10.1101/cshperspect.a008300
9. Yosypiv I V. Congenital anomalies of the kidney and urinary tract: A genetic disorder? *Int J Nephrol*. 2012. doi:10.1155/2012/909083

10. Cain JE, Di Giovanni V, Smeeton J, Rosenblum ND. Genetics of renal hypoplasia: Insights into the mechanisms controlling nephron endowment. *Pediatr Res.* 2010. doi:10.1203/00006450-201011001-00175
11. Hiraoka M, Tsukahara H, Ohshima Y, Kasuga K, Ishihara Y, Mayumi M. Renal aplasia is the predominant cause of congenital solitary kidneys. *Kidney Int.* 2002. doi:10.1046/j.1523-1755.2002.00322.x
12. O'Brien LL, McMahon AP. Induction and patterning of the metanephric nephron. *Semin Cell Dev Biol.* 2014;36:31-38. doi:10.1016/j.semcdb.2014.08.014
13. Takasato M, Little MH. The origin of the mammalian kidney: Implications for recreating the kidney in vitro. *Dev.* 2015. doi:10.1242/dev.104802
14. Saxen L. *Organogenesis of the Kidney.* Helsinki: University of Helsinki; 1987. doi:https://doi.org/10.1017/CBO9780511565083
15. Craik C. Patterning a complex organ: branching morphogenesis and nephron segmentation in kidney development. *Dev Cell.* 2010;18(5):698-712. doi:10.1016/j.devcel.2010.04.008
16. Vega QC, Worby CA, Lechner MS, Dixon JE, Dressler GR. Glial cell line-derived neurotrophic factor activates the receptor tyrosine kinase RET and promotes kidney morphogenesis. *Proc Natl Acad Sci U S A.* 1996;93(20):10657-10661. doi:10.1073/pnas.93.20.10657
17. Hartman HA, Lai HL, Patterson LT. Cessation of renal morphogenesis in mice. *Dev Biol.* 2007;310(2):379-387. doi:10.1016/j.ydbio.2007.08.021
18. Bard JBL, Ross ASA. LIF, the ES-cell inhibition factor, reversibly blocks nephrogenesis in cultured mouse kidney rudiments. *Development.* 1991.
19. Carroll TJ, Park JS, Hayashi S, Majumdar A, McMahon AP. Wnt9b plays a central role in

- the regulation of mesenchymal to epithelial transitions underlying organogenesis of the mammalian urogenital system. *Dev Cell*. 2005. doi:10.1016/j.devcel.2005.05.016
20. Short KM, Combes AN, Lefevre J, et al. Global quantification of tissue dynamics in the developing mouse kidney. *Dev Cell*. 2014;29(2):188-202. doi:10.1016/j.devcel.2014.02.017
 21. McMahon AP. Development of the Mammalian Kidney. *Curr Top Dev Biol*. 2016;116(8):1477-1490. doi:10.1161/CIRCRESAHA.116.303790.The
 22. Huang L, Mokkaapati S, Hu Q, Ruteshouser EC, Hicks MJ, Huff V. Nephron Progenitor But Not Stromal Progenitor Cells Give Rise to Wilms Tumors in Mouse Models with β -Catenin Activation or Wt1 Ablation and Igf2 Upregulation. *Neoplasia*. 2016;18(2):71-81. doi:10.1016/j.neo.2015.12.001
 23. Boyle S, Misfeldt A, Chandler KJ, et al. Fate mapping using Cited1-CreERT2 mice demonstrates that the cap mesenchyme contains self-renewing progenitor cells and gives rise exclusively to nephronic epithelia. *Dev Biol*. 2008;313(1):234-245. doi:10.1016/j.ydbio.2007.10.014
 24. O'Brien LL. Nephron progenitor cell commitment: Striking the right balance. *Semin Cell Dev Biol*. July 2018. doi:10.1016/J.SEMCDB.2018.07.017
 25. Sariola H. Nephron induction revisited: From caps to condensates. *Curr Opin Nephrol Hypertens*. 2002. doi:10.1097/00041552-200201000-00003
 26. Combes AN, Lefevre JG, Wilson S, Hamilton NA, Little MH. Cap mesenchyme cell swarming during kidney development is influenced by attraction, repulsion, and adhesion to the ureteric tip. 2016. doi:10.1016/j.ydbio.2016.06.028
 27. Grobstein C. Inductive epithelio-mesenchymal interaction in cultured organ rudiments of

- the mouse. *Science* (80-). 1953. doi:10.1126/science.118.3054.52
28. Costantini F, Shakya R. GDNF/Ret signaling and the development of the kidney. *BioEssays*. 2006. doi:10.1002/bies.20357
 29. O'brien LL, Combes AN, Short KM, et al. Wnt11 directs nephron progenitor polarity and motile behavior ultimately determining nephron endowment. *Elife*. 2018. doi:10.7554/eLife.40392
 30. Majumdar A, Vainio S, Kispert A, McMahon J, McMahon AP. Wnt11 and Ret/Gdnf pathways cooperate in regulating ureteric branching during metanephric kidney development. *Development*. 2003. doi:10.1242/dev.00520
 31. Das A, Tanigawa S, Karner CM, et al. Stromal-epithelial crosstalk regulates kidney progenitor cell differentiation. *Nat Cell Biol*. 2013. doi:10.1038/ncb2828
 32. Park J-S, Valerius MT, McMahon AP. Wnt/beta-catenin signaling regulates nephron induction during mouse kidney development. *Development*. 2007;134:2533-2539. doi:10.1242/dev.006155
 33. Park JS, Ma W, O'Brien LL, et al. Six2 and Wnt Regulate Self-Renewal and Commitment of Nephron Progenitors through Shared Gene Regulatory Networks. *Dev Cell*. 2012;23(3):637-651. doi:10.1016/j.devcel.2012.07.008
 34. Kobayashi A, Valerius MT, Mugford JW, et al. Six2 defines and regulates a multipotent self-renewing nephron progenitor population throughout mammalian kidney development. *Cell Stem Cell*. 2008;3(2):169-181. doi:10.1016/j.stem.2008.05.020
 35. Lawlor KT, Zappia L, Lefevre J, et al. Nephron progenitor commitment is a stochastic process influenced by cell migration. *Elife*. 2019. doi:10.7554/eLife.41156
 36. Karner CM, Das A, Ma Z, et al. Canonical Wnt9b signaling balances progenitor cell

- expansion and differentiation during kidney development. *Development*. 2011.
doi:10.1242/dev.057646
37. Self M, Lagutin O V, Bowling B, et al. Six2 is required for suppression of nephrogenesis and progenitor renewal in the developing kidney. *EMBO J*. 2006;25(21):5214-5228.
doi:10.1038/sj.emboj.7601381
 38. Boyle S, Shioda T, Perantoni AO, De Caestecker M. Cited1 and Cited2 are differentially expressed in the developing kidney but are not required for nephrogenesis. *Dev Dyn*. 2007.
doi:10.1002/dvdy.21242
 39. Brown AC, Muthukrishnan SD, Guay JA, et al. Role for compartmentalization in nephron progenitor differentiation. *Proc Natl Acad Sci U S A*. 2013;110(12):4640-4645.
doi:10.1073/pnas.1213971110
 40. Blank U, Brown A, Adams DC, Karolak MJ, Oxburgh L. BMP7 promotes proliferation of nephron progenitor cells via a JNK-dependent mechanism. *Development*. 2009.
doi:10.1242/dev.036335
 41. Chung E, Deacon P, Marable S, Shin J, Park J-S. Notch signaling promotes nephrogenesis by downregulating Six2. *Development*. 2016;(September):dev.143503.
doi:10.1242/dev.143503
 42. Fetting JL, Guay JA, Karolak MJ, et al. FOXD1 promotes nephron progenitor differentiation by repressing decorin in the embryonic kidney. *Development*. 2014;141(1):17-27. doi:10.1242/dev.089078
 43. Kobayashi A, Mugford JW, Krautzberger AM, Naiman N, Liao J, McMahon AP. Identification of a multipotent self-renewing stromal progenitor population during mammalian kidney organogenesis. *Stem Cell Reports*. 2014.

doi:10.1016/j.stemcr.2014.08.008

44. Barak H, Huh S-H, Chen S, et al. FGF9 and FGF20 Maintain the Stemness of Nephron Progenitors in Mice and Man. *Dev Cell*. 2012;22(6):1191-1207. doi:10.1016/j.devcel.2012.04.018
45. Brown AC, Muthukrishnan SD, Oxburgh L. A Synthetic Niche for Nephron Progenitor Cells. *Dev Cell*. 2015. doi:10.1016/j.devcel.2015.06.021
46. Yermalovich A V, Osborne JK, Sousa P, et al. Lin28 and let-7 regulate the timing of cessation of murine nephrogenesis. *Nat Commun*. 2019;10(168). doi:10.1038/s41467-018-08127-4
47. Chen S, Brunskill EW, Potter SS, et al. Intrinsic Age-Dependent Changes and Cell-Cell Contacts Regulate Nephron Progenitor Lifespan. *Dev Cell*. 2015;35(1):49-62. doi:10.1016/j.devcel.2015.09.009
48. Laplante M, Sabatini DM. mTOR signaling at a glance. *J Cell Sci*. 2009. doi:10.1242/jcs.051011
49. Saxton RA, Sabatini DM. mTOR Signaling in Growth, Metabolism, and Disease. *Cell*. 2017;168(6):960-976. doi:10.1016/j.cell.2017.02.004
50. Volovelsky O, Nguyen T, Jarmas AE, et al. Hamartin regulates cessation of mouse nephrogenesis independently of Mtor. *Proc Natl Acad Sci U S A*. 2018. doi:10.1073/pnas.1712955115
51. Kim K, Qiang L, Hayden MS, Sparling DP, Purcell NH, Pajvani UB. MTORC1-independent Raptor prevents hepatic steatosis by stabilizing PHLPP2. *Nat Commun*. 2016. doi:10.1038/ncomms10255
52. Liu J, Edgington-Giordano F, Dugas C, et al. Regulation of nephron progenitor cell self-

- renewal by intermediary metabolism. *J Am Soc Nephrol.* 2017. doi:10.1681/ASN.2016111246
53. Cargill K, Hemker SL, Clugston A, et al. Von hippel-lindau acts as a metabolic switch controlling nephron progenitor differentiation. *J Am Soc Nephrol.* 2019;30(7):1192-1205. doi:10.1681/ASN.2018111170
54. Bartel DP. MicroRNA Target Recognition and Regulatory Functions. *Cell.* 2009;136(2):215-233. doi:10.1016/j.cell.2009.01.002.MicroRNA
55. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Front Endocrinol (Lausanne).* 2018. doi:10.3389/fendo.2018.00402
56. Schmiedel JM, Klemm SL, Zheng Y, et al. MicroRNA control of protein expression noise. *Science (80-).* 2015;348(6230):128-132. doi:10.1126/science.aaa1738
57. Chakraborty M, Hu S, Del Giudice M, et al. Networks of enhancers and microRNAs drive variation in cell states. *Networks Enhanc microRNAs drive Var cell states.* 2019;1:668145. doi:10.1101/668145
58. Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. The impact of microRNAs on protein output. *Nature.* 2008;455(7209):64-71. doi:10.1038/nature07242
59. Nam JW, Rissland OS, Koppstein D, et al. Global analyses of the effect of different cellular contexts on microRNA targeting. *Mol Cell.* 2014. doi:10.1016/j.molcel.2014.02.013
60. Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. *Cell.* 2003. doi:10.1016/S0092-8674(03)00759-1
61. Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: Competitive inhibitors of small RNAs in mammalian cells. *Nat Methods.* 2007. doi:10.1038/nmeth1079

62. Vlachos IS, Zagganas K, Paraskevopoulou MD, et al. DIANA-miRPath v3.0: Deciphering microRNA function with experimental support. *Nucleic Acids Res.* 2015. doi:10.1093/nar/gkv403
63. Ho J, Pandey P, Schatton T, et al. The pro-apoptotic protein Bim is a microRNA target in kidney progenitors. *J Am Soc Nephrol.* 2011;22(6):1053-1063. doi:10.1681/ASN.2010080841
64. Hemker SL. Hypoxia-induced microRNA-210 regulation of kidney development. 2020.
65. Mogilyansky E, Rigoutsos I. The miR-17/92 cluster: A comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death Differ.* 2013. doi:10.1038/cdd.2013.125
66. Marrone AK, Stolz DB, Bastacky SI, Kostka D, Bodnar AJ, Ho J. MicroRNA-17~92 Is Required for Nephrogenesis and Renal Function. *J Am Soc Nephrol.* 2014:1-13. doi:10.1681/ASN.2013040390
67. Leng Phua Y, Hong Chen K, Hemker SL, et al. Loss of miR-17~92 results in dysregulation of Cftr in nephron progenitors. *Am J Physiol Ren Physiol.* 2019;316:993-1005. doi:10.1152/ajprenal.00450.2018.-We
68. Heo I, Joo C, Cho J, Ha M, Han J, Kim VN. Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Mol Cell.* 2008. doi:10.1016/j.molcel.2008.09.014
69. Urbach A, Yermalovich A, Zhang J, et al. Lin28 sustains early renal progenitors and induces Wilms tumor. *Genes Dev.* 2014. doi:10.1101/gad.237149.113
70. Johnson CD, Esquela-Kerscher A, Stefani G, et al. The let-7 microRNA represses cell proliferation pathways in human cells. *Cancer Res.* 2007. doi:10.1158/0008-5472.CAN-07-1083

71. Abbott AL, Alvarez-Saavedra E, Miska EA, et al. The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in *Caenorhabditis elegans*. *Dev Cell*. 2005;9(3):403-414. doi:10.1016/j.devcel.2005.07.009
72. Zhu H, Ng SC, Segr A V., et al. The Lin28/let-7 axis regulates glucose metabolism. *Cell*. 2011. doi:10.1016/j.cell.2011.08.033
73. Viswanathan SR, Daley GQ. Lin28: A MicroRNA Regulator with a Macro Role. *Cell*. 2010. doi:10.1016/j.cell.2010.02.007
74. Viswanathan SR, Powers JT, Einhorn W, et al. Lin28 Enhances Tumorigenesis and is Associated With Advanced Human Malignancies. *Nat Genet*. 2009. doi:10.1038/ng.392.Lin28
75. Takamizawa J, Konishi H, Yanagisawa K, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res*. 2004. doi:10.1158/0008-5472.CAN-04-0637
76. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol*. 2015;16(3):144-154. doi:10.1038/nrm3949
77. Ackermann AM, Wang Z, Schug J, Najj A, Kaestner KH. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol Metab*. 2016;5(3):233-244. doi:10.1016/j.molmet.2016.01.002
78. Benabdallah N, Williamson I, Illingworth R, et al. PARP mediated chromatin unfolding is coupled to long-range enhancer activation. 2017:155325. doi:10.1101/155325
79. Dixon JR, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012;485(7398):376-380. doi:10.1038/nature11082

80. Symmons O, Uslu VV, Tsujimura T, et al. Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* 2014. doi:10.1101/gr.163519.113
81. Meuleman W, Peric-Hupkes D, Kind J, et al. Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res.* 2013;23(2):270-280. doi:10.1101/gr.141028.112
82. Calo E, Wysocka J. Modification of Enhancer Chromatin: What, How, and Why? *Mol Cell.* 2013;49(5):825-837. doi:10.1016/j.molcel.2013.01.038
83. Lee K, Hsiung CCS, Huang P, Raj A, Blobel GA. Dynamic enhancer–gene body contacts during transcription elongation. *Genes Dev.* 2015. doi:10.1101/gad.255265.114
84. Papantonis A, Larkin JD, Wada Y, et al. Active RNA polymerases: Mobile or immobile molecular machines? *PLoS Biol.* 2010. doi:10.1371/journal.pbio.1000419
85. Schoenfelder S, Fraser P. Long-range enhancer–promoter contacts in gene expression control. *Nat Rev Genet.* 2019. doi:10.1038/s41576-019-0128-0
86. Furlong EEM, Levine M. Developmental enhancers and chromosome topology. *Science (80-).* 2018. doi:10.1126/science.aau0320
87. O’Brien LL, Guo Q, Bahrami-Samani E, et al. *Transcriptional Regulatory Control of Mammalian Nephron Progenitors Revealed by Multi-Factor Cistromic Analysis and Genetic Studies.* Vol 14.; 2018. doi:10.1371/journal.pgen.1007181
88. Symmons O, Pan L, Remeseiro S, et al. The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. *Dev Cell.* 2016. doi:10.1016/j.devcel.2016.10.015
89. Kind J, van Steensel B. Genome-nuclear lamina interactions and gene regulation. *Curr Opin Cell Biol.* 2010. doi:10.1016/j.ceb.2010.04.002

90. Pugacheva EM, Kubo N, Loukinov D, et al. CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. *Proc Natl Acad Sci U S A*. 2020. doi:10.1073/pnas.1911708117
91. Parelho V, Hadjur S, Spivakov M, et al. Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. *Cell*. 2008. doi:10.1016/j.cell.2008.01.011
92. Wendt KS, Yoshida K, Itoh T, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*. 2008. doi:10.1038/nature06634
93. Davidson IF, Bauer B, Goetz D, Tang W, Wutz G, Peters JM. DNA loop extrusion by human cohesin. *Science (80-)*. 2019. doi:10.1126/science.aaz3418
94. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep*. 2016. doi:10.1016/j.celrep.2016.04.085
95. Lupiáñez DG, Kraft K, Heinrich V, et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*. 2015. doi:10.1016/j.cell.2015.04.004
96. Rao SSP, Huang SC, Glenn St Hilaire B, et al. Cohesin Loss Eliminates All Loop Domains. *Cell*. 2017. doi:10.1016/j.cell.2017.09.026
97. Ragoczy T, Bender MA, Telling A, Byron R, Groudine M. The locus control region is required for association of the murine β -globin locus with engaged transcription factories during erythroid maturation. *Genes Dev*. 2006;20(11):1447-1457. doi:10.1101/gad.1419506
98. Dixon JR, Gorkin DU, Ren B. Chromatin Domains: The Unit of Chromosome Organization. *Mol Cell*. 2016;62(5):668-680. doi:10.1016/j.molcel.2016.05.018

99. Cutter AR, Hayes JJ. A brief review of nucleosome structure. *FEBS Lett.* 2015. doi:10.1016/j.febslet.2015.05.016
100. Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet.* 2019. doi:10.1038/s41576-018-0089-8
101. Thurman RE, Rynes E, Humbert R, et al. The accessible chromatin landscape of the human genome. *Nature.* 2012;489(7414):75-82. doi:10.1038/nature11232
102. Bao X, Rubin AJ, Qu K, et al. A novel ATAC-seq approach reveals lineage-specific reinforcement of the open chromatin landscape via cooperation between BAF and p63. *Genome Biol.* 2015;16(284). doi:10.1186/s13059-015-0840-9
103. Zaret KS, Carroll JS. Pioneer transcription factors: Establishing competence for gene expression. *Genes Dev.* 2011;25(21):2227-2241. doi:10.1101/gad.176826.111
104. Spitz F, Furlong EEM. Transcription factors: From enhancer binding to developmental control. *Nat Rev Genet.* 2012. doi:10.1038/nrg3207
105. Burz DS, Rivera-Pomar R, Jäckle H, Hanes SD. Cooperative DNA-binding by Bicoid provides a mechanism for threshold-dependent gene activation in the *Drosophila* embryo. *EMBO J.* 1998. doi:10.1093/emboj/17.20.5998
106. Long HK, Prescott SL, Wysocka J. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell.* 2016;167(5). doi:10.1016/j.cell.2016.09.018
107. Felsenfeld G, Boyes J, Chung J, Clark D, Studitsky V. Chromatin structure and gene expression. *Proc Natl Acad Sci U S A.* 1996. doi:10.1073/pnas.93.18.9384
108. Schones DE, Cui K, Cuddapah S, et al. Dynamic Regulation of Nucleosome Positioning in the Human Genome. *Cell.* 2008;132(5):887-898. doi:10.1016/j.cell.2008.02.022
109. Venkatesh S, Workman JL. Histone exchange, chromatin structure and the regulation of

- transcription. *Nat Rev Mol Cell Biol.* 2015. doi:10.1038/nrm3941
110. Henikoff S. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet.* 2008. doi:10.1038/nrg2206
 111. Janicki SM, Tsukamoto T, Salghetti SE, et al. From silencing to gene expression: Real-time analysis in single cells. *Cell.* 2004. doi:10.1016/S0092-8674(04)00171-0
 112. Heintzman ND, Stuart RK, Hon G, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet.* 2007. doi:10.1038/ng1966
 113. University of Alabama. Epigenetics Protocols - Methods in Molecular Biology 2nd Ed. In: *Methods in Molecular Biology.* Vol 2. ; 2011:1-300. doi:10.1007/978-1-61779-316-5
 114. Zhang T, Zhang Z, Dong Q, Xiong J, Zhu B. Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biol.* 2020. doi:10.1186/s13059-020-01957-w
 115. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol.* 2015;2015(January):21.29.1-21.29.9. doi:10.1002/0471142727.mb2129s109
 116. Jason D. Buenrostro, Paul G. Giresi, Lisa C. Zaba, Howard Y. Chang, and Greenleaf WJ. Transposition of native chromatin for multimodal regulatory analysis and personal epigenomics. *Nat Methods.* 2012;100(2):130-134. doi:10.1016/j.pestbp.2011.02.012.Investigations
 117. Adey A, Morrison HG, Asan, et al. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* 2010. doi:10.1186/gb-2010-11-12-r119

118. Li Z, Schulz MH, Look T, Begemann M, Zenke M, Costa IG. Identification of transcription factor binding sites using ATAC-seq. *Genome Biol.* 2019;20(1):1-24. doi:10.1186/s13059-019-1642-2
119. Yong Zhang, Tao Liu, Clifford A Meyer, Jérôme Eeckhoute, David S Johnson, Bradley E Bernstein, Chad Nusbaum, Richard M Myers, Myles Brown, Wei Li XSL. Model-based Analysis of ChIP-Seq. *Genome Biol.* 2008;9(9):137. doi:10.1038/ng1901
120. Schep AN, Buenrostro JD, Denny SK, Schwartz K, Sherlock G, Greenleaf WJ. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Res.* 2015;25(11):1757-1770. doi:10.1101/gr.192294.115
121. Tsompana M, Buck MJ. Chromatin accessibility: a window into the genome. *Epigenetics Chromatin.* 2014;7(1). doi:10.1186/1756-8935-7-33
122. Hilliard S, Song R, Liu H, et al. Defining the dynamic chromatin landscape of mouse nephron progenitors. *Biol Open.* 2019;8(5). doi:10.1242/bio.042754
123. Suzuki N, Hirano K, Ogino H, Ochi H. Identification of distal enhancers for Six2 expression in pronephros. *Int J Dev Biol.* 2015. doi:10.1387/ijdb.140263ho
124. Brodbeck S, Besenbeck B, Englert C. The transcription factor Six2 activates expression of the Gdnf gene as well as its own promoter. *Mech Dev.* 2004;121(10):1211-1222. doi:10.1016/J.MOD.2004.05.019
125. Nelson WJ, Nusse R. Convergence of Wnt, β -Catenin, and Cadherin pathways. *Science (80-)*. 2004. doi:10.1126/science.1094291
126. Suzuki HI, Young RA, Sharp PA, et al. Super-Enhancer-Mediated RNA Processing Revealed by Integrative MicroRNA Network Analysis. *Cell.* 2017;168(6):1000-1014.e15. doi:10.1016/j.cell.2017.02.015

127. Ho J, Kreidberg JA. The long and short of Micrnas in the kidney. *J Am Soc Nephrol.* 2012. doi:10.1681/ASN.2011080797
128. Lee Y, Kim M, Han J, et al. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 2004. doi:10.1038/sj.emboj.7600385
129. Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* 2004. doi:10.1101/gad.1262504
130. Kim YK, Kim B, Kim VN. Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. *Proc Natl Acad Sci U S A.* 2016. doi:10.1073/pnas.1602532113
131. Lee Y, Ahn C, Han J, et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature.* 2003. doi:10.1038/nature01957
132. Bernstein E, Kim SY, Carmell MA, et al. Dicer is essential for mouse development. *Nat Genet.* 2003. doi:10.1038/ng1253
133. MacRae IJ, Ma E, Zhou M, Robinson C V., Doudna JA. In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A.* 2008. doi:10.1073/pnas.0710869105
134. Rivas F V., Tolia NH, Song JJ, et al. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol.* 2005. doi:10.1038/nsmb918
135. Grimson A, Farh KKH, Johnston WK, Garrett-Engle P, Lim LP, Bartel DP. MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Mol Cell.* 2007. doi:10.1016/j.molcel.2007.06.017
136. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell.* 2005. doi:10.1016/j.cell.2004.12.035
137. Cebrian C, Asai N, D'Agati V, Costantini F. The Number of Fetal Nephron Progenitor Cells

- Limits Ureteric Branching and Adult Nephron Endowment. *Cell Rep.* 2014;7(1):127-137.
doi:10.1016/j.celrep.2014.02.033
138. Brenner BM, Garcia DL, Anderson S. Glomeruli and blood pressure less of one, more the other? *Am J Hypertens.* 1988. doi:10.1093/ajh/1.4.335
139. Foulkes WD, Priest JR, Duchaine TF. DICER1: Mutations, microRNAs and mechanisms. *Nat Rev Cancer.* 2014. doi:10.1038/nrc3802
140. Nagalakshmi VK, Ren Q, Pugh MM, Valerius MT, McMahon AP, Yu J. Dicer regulates the development of nephrogenic and ureteric compartments in the mammalian kidney. *Kidney Int.* 2011;79(3):317-330. doi:10.1038/ki.2010.385
141. Harvey SJ, Jarad G, Cunningham J, et al. Podocyte-specific deletion of dicer alters cytoskeletal dynamics and causes glomerular disease. *J Am Soc Nephrol.* 2008. doi:10.1681/ASN.2008020233
142. Ho J, Ng KH, Rosen S, Dostal A, Gregory RI, Kreidberg J a. Podocyte-specific loss of functional microRNAs leads to rapid glomerular and tubular injury. *J Am Soc Nephrol.* 2008;19(11):2069-2075. doi:10.1681/ASN.2008020162
143. Phua YL, Chu JYS, Marrone AK, Bodnar AJ, Sims-Lucas S, Ho J. Renal stromal miRNAs are required for normal nephrogenesis and glomerular mesangial survival. *Physiol Rep.* 2015. doi:10.14814/phy2.12537
144. Shi S, Yu L, Chiu C, et al. Podocyte-selective deletion of dicer induces proteinuria and glomerulosclerosis. *J Am Soc Nephrol.* 2008. doi:10.1681/ASN.2008030312
145. Zhdanova O, Srivastava S, Di L, et al. The inducible deletion of Droscha and microRNAs in mature podocytes results in a collapsing glomerulopathy. *Kidney Int.* 2011. doi:10.1038/ki.2011.122

146. Park SM, Gaur AB, Lengyel E, Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev.* 2008. doi:10.1101/gad.1640608
147. Cloonan N, Wani S, Xu Q, et al. MicroRNAs and their isomiRs function cooperatively to target common biological pathways. *Genome Biol.* 2011. doi:10.1186/gb-2011-12-12-r126
148. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357-359. doi:10.1038/nmeth.1923
149. Friedländer MR, MacKowiak SD, Li N, Chen W, Rajewsky N. MiRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 2012;40(1):37-52. doi:10.1093/nar/gkr688
150. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 2006. doi:10.1093/nar/gkj112
151. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550. doi:10.1186/s13059-014-0550-8
152. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011. doi:10.1038/nbt.1754
153. Schroeder A, Mueller O, Stocker S, et al. The RIN: An RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol.* 2006. doi:10.1186/1471-2199-7-3
154. Liao Y, Smyth GK, Shi W. The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 2013. doi:10.1093/nar/gkt214
155. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for

- assigning sequence reads to genomic features. *Bioinformatics*. 2014.
doi:10.1093/bioinformatics/btt656
156. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
doi:10.1093/nar/gkv007
157. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of Mammalian MicroRNA Targets. *Cell*. 2003. doi:10.1016/S0092-8674(03)01018-3
158. Reczko M, Maragkakis M, Alexiou P, Grosse I, Hatzigeorgiou AG. Functional microRNA targets in protein coding sequences. *Bioinformatics*. 2012.
doi:10.1093/bioinformatics/bts043
159. Yong SL, Dutta A. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes Dev*. 2007. doi:10.1101/gad.1540407
160. Mayr C, Hemann MT, Bartel DP. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science (80-)*. 2007. doi:10.1126/science.1137999
161. Johnson SM, Grosshans H, Shingara J, et al. RAS is regulated by the let-7 microRNA family. *Cell*. 2005. doi:10.1016/j.cell.2005.01.014
162. Sampson VB, Rong NH, Han J, et al. MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. *Cancer Res*. 2007. doi:10.1158/0008-5472.CAN-07-2462
163. O'Brien LL, Guo Q, Lee Y, et al. Differential regulation of mouse and human nephron progenitors by the Six family of transcriptional regulators. *Development*. 2016;143(4):595-608. doi:10.1242/dev.127175
164. DI Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow

- enables reproducible computational workflows. *Nat Biotechnol.* 2017.
doi:10.1038/nbt.3820
165. Encode Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489(7414):57-74. doi:10.1038/nature11247
166. Wang Y, Song F, Zhang B, et al. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biol.* 2018;19(1):151. doi:10.1186/s13059-018-1519-9
167. Cebrian C, Asai N, D'Agati V, Costantini F. The number of fetal nephron progenitor cells limits ureteric branching and adult nephron endowment. *Cell Rep.* 2014;7(1):127-137. doi:10.1016/j.celrep.2014.02.033
168. Rumballe BA, Georgas KM, Combes AN, Ju AL, Gilbert T, Little MH. Nephron formation adopts a novel spatial topology at cessation of nephrogenesis. *Dev Biol.* 2011. doi:10.1016/j.ydbio.2011.09.011
169. Stark K, Vainio S, Vassileva G, McMahon AP. Epithelial transformation of metanephric mesenchyme in the developing kidney regulated by Wnt-4. *Nature.* 1994. doi:10.1038/372679a0
170. Kulshreshtha R, Ferracin M, Wojcik SE, et al. A MicroRNA Signature of Hypoxia. *Mol Cell Biol.* 2006;27(5):1859-1867. doi:10.1128/MCB.01395-06
171. Nishikawa M, Yuri S, Kimura H, et al. Comprehensive analysis of chromatin signature and transcriptome uncovers functional lncRNAs expressed in nephron progenitor cells. *Gene Regul Mech.* 2019;1862(1):58-70. doi:10.1016/j.bbagr.2018.09.006
172. McFarlane L, Truong V, Palmer JS, Wilhelm D. Novel PCR assay for determining the genetic sex of mice. *Sex Dev.* 2013;7(4):207-211. doi:10.1159/000348677

173. Brown AC, Blank U, Adams DC, et al. Isolation and culture of cells from the nephrogenic zone of the embryonic mouse kidney. *J Vis Exp*. 2011. doi:10.3791/2555
174. Hemker SL, Cerqueira DM, Bodnar AJ, et al. Deletion of hypoxia-responsive microRNA-210 results in a sex-specific decrease in nephron number. *FASEB J*. 2020;34(4):5782-5799. doi:10.1096/fj.201902767R
175. Brunskill EW, Park J-S, Chung E, Chen F, Magella B, Potter SS. Single cell dissection of early kidney development: multilineage priming. *Development*. 2014;141(15):3093-3101. doi:10.1242/dev.110601
176. Barber RD, Harmer DW, Coleman RA, Clark BJ. GAPDH as a housekeeping gene: Analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol Genomics*. 2005. doi:10.1152/physiolgenomics.00025.2005
177. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods*. 2001. doi:10.1006/meth.2001.1262
178. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013;10(12):1213-1218. doi:10.1038/nmeth.2688
179. BabrahamLab. Trim Galore. *Bioinformatics*. 2014. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore.
180. Church DM, Schneider VA, Graves T, et al. Modernizing reference genome assemblies. *PLoS Biol*. 2011. doi:10.1371/journal.pbio.1001091
181. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.

doi:10.1186/gb-2009-10-3-r25

182. Broad Institute. Picard tools. <https://broadinstitute.github.io/picard/>. 2016.
183. ENCODE. ATAC-seq Data Analysis Pipeline. :5-6. <https://www.encodeproject.org/atac-seq/>.
184. Lee J. Assign multimappers. 2017. https://github.com/kundajelab/atac_dnase_pipelines.
185. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
186. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842. doi:10.1093/bioinformatics/btq033
187. Li Q, Brown JB, Huang H, Bickel PJ. Measuring reproducibility of high-throughput experiments. *Ann Appl Stat*. 2011;5(3):1752-1779. doi:10.1214/11-AOAS466
188. Boleu N, Kundaje A, Bickel PJ, Li Q. Irreproducible Discovery Rate. 2015. <https://github.com/nboley/idr>.
189. Ritchie ME, Diyagama D, Neilson J, et al. Empirical array quality weights in the analysis of microarray data. *BMC Bioinformatics*. 2006;7(1):261. doi:10.1186/1471-2105-7-261
190. McLean CY, Bristor D, Hiller M, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010. doi:10.1038/nbt.1630
191. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014;507(7493):455-461. doi:10.1038/nature12787
192. Li Z, Schulz MH, Zenke M, et al. Identification of Transcription Factor Binding Sites using ATAC-seq. *bioRxiv*. July 2018:362863. doi:10.1101/362863
193. Kulakovskiy I V., Vorontsov IE, Yevshin IS, et al. HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-

- Seq analysis. *Nucleic Acids Res.* 2018. doi:10.1093/nar/gkx1106
194. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* 2019. doi:10.1093/nar/gkz114
195. Kozomara A, Griffiths-Jones S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 2014;42(D1). doi:10.1093/nar/gkt1181
196. Love MI, Anders S, Huber W. *Differential Analysis of Count Data - the DESeq2 Package.* Vol 15.; 2014. doi:110.1186/s13059-014-0550-8
197. ENCODE. ATAC-seq Data Standards – ENCODE. <https://www.encodeproject.org/data-standards/atac-seq/>. Accessed March 4, 2017.
198. Buscà R, Pouysségur J, Lenormand P. ERK1 and ERK2 map kinases: Specific roles or functional redundancy? *Front Cell Dev Biol.* 2016. doi:10.3389/fcell.2016.00053
199. Hirabayashi Y, Itoh Y, Tabata H, et al. The Wnt/ β -catenin pathway directs neuronal differentiation of cortical neural precursor cells. *Development.* 2004. doi:10.1242/dev.01165
200. Combes AN, Phipson B, Lawlor KT, et al. Single cell analysis of the developing mouse kidney provides deeper insight into marker gene expression and ligand-receptor crosstalk. *Development.* 2019. doi:10.1242/dev.178673
201. Sanchez MP, Silos-Santiago I, Frisen J, He B, Lira SA, Barbacid M. Renal agenesis and the absence of enteric neurons in mice lacking GDNF. *Nature.* 1996;382(6586):70-73. doi:10.1038/382070a0
202. Li Z, Yin H, Hao S, et al. MiR-200 family promotes podocyte differentiation through repression of RSAD2. *Sci Rep.* 2016. doi:10.1038/srep27105
203. Potenza N, Panella M, Castiello F, Mosca N, Amendola E, Russo A. Molecular mechanisms

- governing microRNA-125a expression in human hepatocellular carcinoma cells. *Sci Rep*. 2017. doi:10.1038/s41598-017-11418-3
204. Chaudhuri AA, So AYL, Mehta A, et al. Oncomir miR-125b regulates hematopoiesis by targeting the gene Lin28A. *Proc Natl Acad Sci U S A*. 2012. doi:10.1073/pnas.1200677109
205. Wang S, Aurora AB, Johnson BA, et al. The Endothelial-Specific MicroRNA miR-126 Governs Vascular Integrity and Angiogenesis. *Dev Cell*. 2008. doi:10.1016/j.devcel.2008.07.002
206. Schober A, Nazari-Jahantigh M, Wei Y, et al. MicroRNA-126-5p promotes endothelial proliferation and limits atherosclerosis by suppressing Dlk1. *Nat Med*. 2014. doi:10.1038/nm.3487
207. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *Elife*. 2015. doi:10.7554/eLife.05005
208. Murugan S, Shan J, Kühl SJ, et al. WT1 and Sox11 regulate synergistically the promoter of the Wnt4 gene that encodes a critical signal for nephrogenesis. *Exp Cell Res*. 2012. doi:10.1016/j.yexcr.2012.03.008
209. Rowan CJ, Li W, Martirosyan H, et al. Hedgehog-GLI signaling in Foxd1-positive stromal cells promotes murine nephrogenesis via TGF β signaling. *Development*. 2018. doi:10.1242/dev.159947
210. Ihermann-Hella A, Hirashima T, Kupari J, et al. Dynamic MAPK/ERK Activity Sustains Nephron Progenitors through Niche Regulation and Primes Precursors for Differentiation. *Stem Cell Reports*. 2018. doi:10.1016/j.stemcr.2018.08.012
211. Rozario T, Desimone DW. The Extracellular Matrix In Development and Morphogenesis: A Dynamic View. *Dev Biol*. 2010;341(1):126-140. doi:10.1016/j.ydbio.2009.10.026.The

212. Närvä E, Stubb A, Guzmán C, et al. A Strong Contractile Actin Fence and Large Adhesions Direct Human Pluripotent Colony Morphology and Adhesion. *Stem Cell Reports*. 2017. doi:10.1016/j.stemcr.2017.05.021
213. Linton JM, Martin GR, Reichardt LF. The ECM protein nephronectin promotes kidney development via integrin $\alpha 8\beta 1$ -mediated stimulation of Gdnf expression. *Development*. 2007. doi:10.1242/dev.005033
214. Kang HM, Huang S, Reidy K, Han SH, Chinga F, Susztak K. Sox9-Positive Progenitor Cells Play a Key Role in Renal Tubule Epithelial Regeneration in Mice. *Cell Rep*. 2016. doi:10.1016/j.celrep.2015.12.071
215. Kumar S, Liu J, Pang P, et al. Sox9 Activation Highlights a Cellular Pathway of Renal Repair in the Acutely Injured Mammalian Kidney. *Cell Rep*. 2015. doi:10.1016/j.celrep.2015.07.034
216. Fierro-Fernández M, Miguel V, Márquez-Expósito L, et al. MiR-9-5p protects from kidney fibrosis by metabolic reprogramming. *FASEB J*. 2020;34(1):410-431. doi:10.1096/fj.201901599RR
217. Kreidberg JA, Sariola H, Loring JM, et al. WT-1 is required for early kidney development. *Cell*. 1993;74(4):679-691. doi:10.1016/0092-8674(93)90515-R
218. Phua YL, Clugston A, Chen KH, Kostka D, Ho J. Small non-coding RNA expression in developing mouse nephron progenitor cells. *Sci Data*. 2018;5:180218. doi:https://doi.org/10.1038/sdata.2018.218
219. Gregory PA, Bert AG, Paterson EL, et al. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol*. 2008;10(5):593-601. doi:10.1038/ncb1722

220. Mongroo PS, Rustgi AK. The role of the miR-200 family in epithelial-mesenchymal transition. *Cancer Biol Ther.* 2010;10(3):219-222. doi:10.4161/cbt.10.3.12548
221. Naiman N, Fujioka K, Fujino M, et al. Repression of Interstitial Identity in Nephron Progenitor Cells by Pax2 Establishes the Nephron-Interstitium Boundary during Kidney Development. *Dev Cell.* 2017;41(4):349-365.e3. doi:10.1016/j.devcel.2017.04.022
222. Sarkar A, Hochedlinger K. The Sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. *Cell Stem Cell.* 2013. doi:10.1016/j.stem.2012.12.007
223. Capone VP, Morello W, Taroni F, Montini G. Genetics of congenital anomalies of the kidney and urinary tract: The current state of play. *Int J Mol Sci.* 2017. doi:10.3390/ijms18040796
224. Keller G, Zimmer G, Mall G, Ritz E, Amann K. Nephron Number in Patients with Primary Hypertension. *N Engl J Med.* 2003;348:101-108. doi:10.1056/NEJMoa1005372
225. Hoy WE, Hughson MD, Singh GR, Douglas-Denton R, Bertram JF. Reduced nephron number and glomerulomegaly in Australian Aborigines: A group at high risk for renal disease and hypertension. *Kidney Int.* 2006. doi:10.1038/sj.ki.5000397
226. Humphreys BD. Mechanisms of Renal Fibrosis. *Annu Rev Physiol.* 2018. doi:10.1146/annurev-physiol-022516-034227
227. Coresh J, Selvin E, Stevens LA, et al. Prevalence of chronic kidney disease in the United States. *J Am Med Assoc.* 2007. doi:10.1001/jama.298.17.2038
228. Brown AC, Adams D, de Caestecker M, Yang X, Friesel R, Oxburgh L. FGF/EGF signaling regulates the renewal of early nephron progenitors during embryonic development. *Development.* 2011;138(23). <http://dev.biologists.org/content/138/23/5099>. Accessed June 3, 2017.

229. Muthukrishnan SD, Yang X, Friesel R, Oxburgh L. Concurrent BMP7 and FGF9 signalling governs AP-1 function to promote self-renewal of nephron progenitor cells. *Nat Commun.* 2015. doi:10.1038/ncomms10027
230. Di Giovanni V, Walker KA, Bushnell D, et al. Fibroblast growth factor receptor-Frs2 α signaling is critical for nephron progenitors. *Dev Biol.* 2015. doi:10.1016/j.ydbio.2015.01.018
231. Dressler GR. *The Cellular Basis of Kidney Development.* Vol 22.; 2006:509-529. doi:10.1146/annurev.cellbio.22.010305.104340
232. Schell C, Wanner N, Huber TB. Glomerular development - Shaping the multi-cellular filtration unit. *Semin Cell Dev Biol.* 2014. doi:10.1016/j.semcd.2014.07.016
233. Lindahl P, Hellström M, Kalén M, et al. Paracrine PDGF-B/PDGF-R β signaling controls mesangial cell development in kidney glomeruli. *Development.* 1998.
234. Betsholtz C, Lindblom P, Bjarnegard M, Enge M, Gerhardt H, Lindahl P. Role of platelet-derived growth factor in mesangium development and vasculopathies: Lessons from platelet-derived growth factor and platelet-derived growth factor receptor mutations in mice. *Curr Opin Nephrol Hypertens.* 2004. doi:10.1097/00041552-200401000-00007
235. Robert B, Zhao X, Abrahamson DR. Coexpression of neuropilin-1, Flk1, and VEGF164 in developing and mature mouse kidney glomeruli. *Am J Physiol - Ren Physiol.* 2000. doi:10.1152/ajprenal.2000.279.2.f275
236. Abrahamson DR. Glomerulogenesis in the developing kidney. *Semin Nephrol.* 1991. doi:10.5555/uri:pii:027092959190060W
237. Adam M, Potter AS, Potter SS. Psychrophilic proteases dramatically reduce single cell RNA-seq artifacts: A molecular atlas of kidney development. *Development.*

- 2017;1:dev.151142. doi:10.1242/dev.151142
238. Magella B, Adam M, Potter AS, et al. Cross-platform single cell analysis of kidney development shows stromal cells express Gdnf. *Dev Biol.* 2018. doi:10.1016/j.ydbio.2017.11.006
239. Correction: Single cell analysis of the developing mouse kidney provides deeper insight into marker gene expression and ligand-receptor crosstalk (Development, (2019) 146, 12, 10.1242/dev.178673). *Dev.* 2019. doi:10.1242/dev.182162
240. England AR, Chaney CP, Das A, et al. Identification and characterization of cellular heterogeneity within the developing renal interstitium. *Development.* 2020. doi:10.1242/dev.190108
241. Menon R, Otto EA, Kokoruda A, et al. Single-cell analysis of progenitor cell dynamics and lineage specification in the human fetal kidney. *Dev.* 2018. doi:10.1242/dev.164038
242. Wang P, Chen Y, Yong J, et al. Dissecting the Global Dynamic Molecular Profiles of Human Fetal Kidney Development by Single-Cell RNA Sequencing. *Cell Rep.* 2018. doi:10.1016/j.celrep.2018.08.056
243. Lindström NO, Guo J, Kim AD, et al. Conserved and divergent features of mesenchymal progenitor cell types within the cortical nephrogenic niche of the human and mouse kidney. *J Am Soc Nephrol.* 2018;29(3):806-824. doi:10.1681/ASN.2017080890
244. Tanigawa S, Taguchi A, Sharma N, Perantoni AO, Nishinakamura R. Selective In Vitro Propagation of Nephron Progenitors Derived from Embryos and Pluripotent Stem Cells. *Cell Rep.* 2016. doi:10.1016/j.celrep.2016.03.076
245. Li Z, Araoka T, Wu J, et al. 3D Culture Supports Long-Term Expansion of Mouse and Human Nephrogenic Progenitors. *Cell Stem Cell.* 2016. doi:10.1016/j.stem.2016.07.016

246. Takasato M, Little MH. A strategy for generating kidney organoids: Recapitulating the development in human pluripotent stem cells. *Dev Biol.* 2016. doi:10.1016/j.ydbio.2016.08.024
247. Morizane R, Bonventre J V. Generation of nephron progenitor cells and kidney organoids from human pluripotent stem cells. *Nat Protoc.* 2017. doi:10.1038/nprot.2016.170
248. Zerbino DR, Achuthan P, Akanni W, et al. Ensembl 2018. *Nucleic Acids Res.* 2018. doi:10.1093/nar/gkx1098
249. Huber W, Carey VJ, Gentleman R, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods.* 2015. doi:10.1038/nmeth.3252
250. Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, Marioni JC. EmptyDrops: Distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* 2019. doi:10.1186/s13059-019-1662-y
251. Griffiths JA, Richard AC, Bach K, Lun ATL, Marioni JC. Detection and removal of barcode swapping in single-cell RNA-seq data. *Nat Commun.* 2018. doi:10.1038/s41467-018-05083-x
252. McCarthy DJ, Campbell KR, Lun ATL, Wills QF. Scater: Pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics.* 2017. doi:10.1093/bioinformatics/btw777
253. Bais AS, Kostka D. Scds: Computational annotation of doublets in single-cell RNA sequencing data. *Bioinformatics.* 2020. doi:10.1093/bioinformatics/btz698
254. Lun ATL, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research.* 2016. doi:10.12688/f1000research.9501.2

255. Street K, Risso D, Fletcher RB, et al. Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*. 2018. doi:10.1186/s12864-018-4772-0
256. Venables WN, Ripley BD. *Modern Applied Statistics with S Fourth Edition* By.; 2002. doi:10.2307/2685660
257. Huang M, Wang J, Torre E, et al. SAVER: Gene expression recovery for single-cell RNA sequencing. *Nat Methods*. 2018. doi:10.1038/s41592-018-0033-z
258. Wood SN, Pya N, Säfken B. Smoothing Parameter and Model Selection for General Smooth Models. *J Am Stat Assoc*. 2016. doi:10.1080/01621459.2016.1180986
259. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst*. 2015. doi:10.1016/j.cels.2015.12.004
260. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: Single-cell regulatory network inference and clustering. *Nat Methods*. 2017. doi:10.1038/nmeth.4463
261. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring regulatory networks from expression data using tree-based methods. *PLoS One*. 2010. doi:10.1371/journal.pone.0012776
262. In situ hybridization; a practical approach. *Gen Pharmacol Vasc Syst*. 1993. doi:10.1016/0306-3623(93)90403-k
263. Adam M, Potter AS, Potter SS. Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: A molecular atlas of kidney development. *Dev*. 2017. doi:10.1242/dev.151142
264. England AR, Chaney CP, Das A, et al. Identification and characterization of cellular heterogeneity within the developing renal interstitium. *Development*. 2020.

doi:10.1242/dev.190108

265. Marlier A, Gilbert T. Expression of retinoic acid-synthesizing and -metabolizing enzymes during nephrogenesis in the rat. *Gene Expr Patterns*. 2004;5(2):179-185. doi:10.1016/j.modgep.2004.08.005
266. Diez-Roux G, Banfi S, Sultan M, et al. A high-resolution anatomical atlas of the transcriptome in the mouse embryo. *PLoS Biol*. 2011. doi:10.1371/journal.pbio.1000582
267. Harding SD, Armit C, Armstrong J, et al. The GUDMAP database - an online resource for genitourinary research. *Development*. 2011. doi:10.1242/dev.063594
268. Takemoto M, He L, Norlin J, et al. Large-scale identification of genes implicated in kidney glomerulus development and function. *EMBO J*. 2006. doi:10.1038/sj.emboj.7601014
269. Lee JW, Chou CL, Knepper MA. Deep sequencing in microdissected renal tubules identifies nephron segment-specific transcriptomes. *J Am Soc Nephrol*. 2015. doi:10.1681/ASN.2014111067
270. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011. doi:10.1093/bioinformatics/btr260
271. Yoffe Y, David M, Kalaora R, et al. Cap-independent translation by DAP5 controls cell fate decisions in human embryonic stem cells. *Genes Dev*. 2016. doi:10.1101/gad.285239.116
272. Chen PM, Lin CH, Li NT, et al. C-Maf regulates pluripotency genes, proliferation/self-renewal, and lineage commitment in ROS-mediated senescence of human mesenchymal stem cells. *Oncotarget*. 2015. doi:10.18632/oncotarget.6178
273. Julian LM, Blais A. Transcriptional control of stem cell fate by E2Fs and pocket proteins. *Front Genet*. 2015. doi:10.3389/fgene.2015.00161

274. Hu T, Ghazaryan S, Sy C, Wiedmeyer C, Chang V, Wu L. Concomitant inactivation of Rb and E2f8 in hematopoietic stem cells synergizes to induce severe anemia. *Blood*. 2012. doi:10.1182/blood-2011-10-388231
275. Jolly LA, Nguyen LS, Domingo D, et al. HCFC1 loss-of-function mutations disrupt neuronal and neural progenitor cells of the developing brain. *Hum Mol Genet*. 2015. doi:10.1093/hmg/ddv083
276. Chen YH, Hung MC, Li LY. EZH2: A pivotal regulator in controlling cell differentiation. *Am J Transl Res*. 2012.
277. Farcas AM, Blackledge NP, Sudbery I, et al. KDM2B links the polycomb repressive complex 1 (PRC1) to recognition of CpG islands. *Elife*. 2012. doi:10.7554/eLife.00205
278. Papetti M, Augenlicht LH. MYBL2, a link between proliferation and differentiation in maturing colon epithelial cells. *J Cell Physiol*. 2011. doi:10.1002/jcp.22399
279. Georgas K, Rumballe B, Valerius MT, et al. Analysis of early nephron patterning reveals a role for distal RV proliferation in fusion to the ureteric tip via a cap mesenchyme-derived connecting segment. *Dev Biol*. 2009. doi:10.1016/j.ydbio.2009.05.578
280. Bariety J, Mandet C, Hill GS, Bruneval P. Parietal podocytes in normal human glomeruli. *J Am Soc Nephrol*. 2006. doi:10.1681/ASN.2006040325
281. Ohse T, Vaughan MR, Kopp JB, et al. De novo expression of podocyte proteins in parietal epithelial cells during experimental glomerular disease. *Am J Physiol - Ren Physiol*. 2010. doi:10.1152/ajprenal.00428.2009
282. Fan X, Li Q, Pisarek-Horowitz A, et al. Inhibitory effects of Robo2 on nephrin: A crosstalk between positive and negative signals regulating podocyte structure. *Cell Rep*. 2012. doi:10.1016/j.celrep.2012.06.002

283. Upadhyay G. Emerging Role of Lymphocyte Antigen-6 Family of Genes in Cancer and Immune Cells. *Front Immunol.* 2019. doi:10.3389/fimmu.2019.00819
284. Li Y, Liu J, Li W, et al. p53 enables metabolic fitness and self-renewal of nephron progenitor cells. *J Cell Sci.* 2015. doi:10.1242/jcs.171835
285. Wu H, Lai CF, Chang-Panesso M, Humphreys BD. Proximal tubule translational profiling during kidney fibrosis reveals proinflammatory and long noncoding RNA expression patterns with sexual dimorphism. *J Am Soc Nephrol.* 2020. doi:10.1681/ASN.2019040337
286. Ransick A, Lindström NO, Liu J, et al. Single-Cell Profiling Reveals Sex, Lineage, and Regional Diversity in the Mouse Kidney. *Dev Cell.* 2019. doi:10.1016/j.devcel.2019.10.005
287. Combes AN, Zappia L, Er PX, Oshlack A, Little MH. Single-cell analysis reveals congruence between kidney organoids and human fetal kidney. *Genome Med.* 2019. doi:10.1186/s13073-019-0615-0
288. AbouAlaiwi WA, Ratnam S, Booth RL, Shah J V., Nauli SM. Endothelial cells from humans and mice with polycystic kidney disease are characterized by polyploidy and chromosome segregation defects through survivin down-regulation. *Hum Mol Genet.* 2011. doi:10.1093/hmg/ddq470
289. Chen J, Chen JK, Conway EM, Harris RC. Survivin mediates renal proximal tubule recovery from AKI. *J Am Soc Nephrol.* 2013. doi:10.1681/ASN.2013010076
290. Zhou D, Li Y, Lin L, Zhou L, Igarashi P, Liu Y. Tubule-specific ablation of endogenous B-catenin aggravates acute kidney injury in mice. *Kidney Int.* 2012. doi:10.1038/ki.2012.173
291. Byun SS, Yeo WG, Lee SE, Lee E. Expression of Survivin in Renal Cell Carcinomas: Association with Pathologic Features and Clinical Outcome. *Urology.* 2007. doi:10.1016/j.urology.2006.09.024

292. Parker AS, Lohse CM, Leibovich BC, Cheville JC, Sheinin YM, Kwon ED. Comparison of digital image analysis versus visual assessment to assess survivin expression as an independent predictor of survival for patients with clear cell renal cell carcinoma. *Hum Pathol*. 2008. doi:10.1016/j.humpath.2007.12.009
293. Kobayashi K, Hatano M, Otaki M, Ogasawara T, Tokuhisa T. Expression of a murine homologue of the inhibitor of apoptosis protein is related to cell proliferation. *Proc Natl Acad Sci U S A*. 1999. doi:10.1073/pnas.96.4.1457
294. Li F, Altieri DC. Transcriptional analysis of human survivin gene expression. *Biochem J*. 1999. doi:10.1042/0264-6021:3440305
295. Li F, Ambrosini G, Chu EY, et al. Control of apoptosis and mitotic spindle checkpoint by survivin. *Nature*. 1998. doi:10.1038/25141
296. Zhao J, Tenev T, Martins LM, Downward J, Lemoine NR. The ubiquitin-proteasome pathway regulates survivin degradation in a cell cycle-dependent manner. *J Cell Sci*. 2000.
297. Uren AG, Wong L, Pakusch M, et al. Survivin and the inner centromere protein INCENP show similar cell-cycle localization and gene knockout phenotype. *Curr Biol*. 2000. doi:10.1016/S0960-9822(00)00769-7
298. Lens SMA, Wolthuis RMF, Klompmaker R, et al. Survivin is required for a sustained spindle checkpoint arrest in response to lack of tension. *EMBO J*. 2003. doi:10.1093/emboj/cdg307
299. Carvalho A, Carmena M, Sambade C, Earnshaw WC, Wheatley SP. Survivin is required for stable checkpoint activation in taxol-treated HeLa cells. *J Cell Sci*. 2003. doi:10.1242/jcs.00612
300. Rajagopalan S, Balasubramanian MK. *Schizosaccharomyces pombe* Bir1p, a nuclear

- protein that localizes to kinetochores and the spindle midzone, is essential for chromosome condensation and spindle elongation during mitosis. *Genetics*. 2002.
301. Yue Z, Carvalho A, Xu Z, et al. Deconstructing Survivin: Comprehensive genetic analysis of Survivin function by conditional knockout in a vertebrate cell line. *J Cell Biol*. 2008. doi:10.1083/jcb.200806118
 302. Speliotes EK, Uren A, Vaux D, Horvitz HR. The survivin-like *C. elegans* BIR-1 protein acts with the aurora-like kinase AIR-2 to affect chromosomes and the spindle midzone. *Mol Cell*. 2000. doi:10.1016/S1097-2765(00)00023-X
 303. Yang D, Welm A, Bishop JM. Cell division and cell survival in the absence of survivin. *Proc Natl Acad Sci U S A*. 2004. doi:10.1073/pnas.0406665101
 304. Marusawa H, Matsuzawa S ichi, Welsh K, et al. HBXIP functions as a cofactor of survivin in apoptosis suppression. *EMBO J*. 2003. doi:10.1093/emboj/cdg263
 305. Dohi T, Okada K, Xia F, et al. An IAP-IAP complex inhibits apoptosis. *J Biol Chem*. 2004. doi:10.1074/jbc.C400236200
 306. Mita AC, Mita MM, Nawrocki ST, Giles FJ. Survivin: Key regulator of mitosis and apoptosis and novel target for cancer therapeutics. *Clin Cancer Res*. 2008. doi:10.1158/1078-0432.CCR-08-0746
 307. Wheatley SP, Altieri DC. Survivin at a glance. *J Cell Sci*. 2019. doi:10.1242/jcs.223826
 308. Mugford JW, Yu J, Kobayashi A, McMahon AP. High-resolution gene expression analysis of the developing mouse kidney defines novel cellular compartments within the nephron progenitor population. *Dev Biol*. 2009. doi:10.1016/j.ydbio.2009.06.043
 309. Brent Thomson R, Wang T, Thomson BR, et al. Role of PDZK1 in membrane expression of renal brush border ion exchangers. *Proc Natl Acad Sci U S A*. 2005.

doi:10.1073/pnas.0506578102

310. Anzai N, Miyazaki H, Noshiro R, et al. The multivalent PDZ domain-containing protein PDZK1 regulates transport activity of renal urate-anion exchanger URAT1 via its C terminus. *J Biol Chem*. 2004. doi:10.1074/jbc.M406724200
311. Sayer JA. Progress in understanding the genetics of calcium-containing nephrolithiasis. *J Am Soc Nephrol*. 2017. doi:10.1681/ASN.2016050576
312. Fearn A, Allison B, Rice SJ, et al. Clinical, biochemical, and pathophysiological analysis of SLC34A1 mutations. *Physiol Rep*. 2018. doi:10.14814/phy2.13715
313. Chau H, El-Maadawy S, McKee MD, Tenenhouse HS. Renal calcification in mice homozygous for the disrupted type IIa Na/Pi cotransporter gene *Npt2*. *J Bone Miner Res*. 2003. doi:10.1359/jbmr.2003.18.4.644
314. Chang CH, Davies JA. In developing mouse kidneys, orientation of loop of Henle growth is adaptive and guided by long-range cues from medullary collecting ducts. *J Anat*. 2019. doi:10.1111/joa.13012
315. Mori S, Nada S, Kimura H, et al. The mTOR pathway controls cell proliferation by regulating the FoxO3a transcription factor via SGK1 kinase. *PLoS One*. 2014. doi:10.1371/journal.pone.0088891
316. Lu P, Weaver VM, Werb Z. The extracellular matrix: A dynamic niche in cancer progression. *J Cell Biol*. 2012. doi:10.1083/jcb.201102147
317. Böttcher RT, Pollet N, Delius H, Niehrs C. The transmembrane protein XFLRT3 forms a complex with FGF receptors and promotes FGF signalling. *Nat Cell Biol*. 2004. doi:10.1038/ncb1082
318. Karaulanov EE, Böttcher RT, Niehrs C. A role for fibronectin-leucine-rich transmembrane

- cell-surface proteins in homotypic cell adhesion. *EMBO Rep.* 2006. doi:10.1038/sj.embor.7400614
319. Boyle SC, Kim M, Valerius MT, McMahon AP, Kopan R. Notch pathway activation can replace the requirement for Wnt4 and wnt9b in mesenchymal-to-epithelial transition of nephron stem cells. *Development.* 2011. doi:10.1242/dev.070433
320. Liu H, Lei C, He Q, Pan Z, Xiao D, Tao Y. Nuclear functions of mammalian MicroRNAs in gene regulation, immunity and cancer. *Mol Cancer.* 2018. doi:10.1186/s12943-018-0765-5
321. Xiao M, Li J, Li W, et al. MicroRNAs Activate Gene Transcription Epigenetically as an Enhancer Trigger. *RNA Biol.* February 2016:00-00. doi:10.1080/15476286.2015.1112487
322. Uchikawa M, Ishida Y, Takemoto T, Kamachi Y, Kondoh H. Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev Cell.* 2003;4(4):509-519. doi:10.1016/S1534-5807(03)00088-1
323. Dapkunas A, Rantanen V, Gui Y, et al. Simple 3D culture of dissociated kidney mesenchyme mimics nephron progenitor niche and facilitates nephrogenesis Wnt-independently. *Sci Rep.* 2019;9(1):1-10. doi:10.1038/s41598-019-49526-x
324. Takasato M, Er PX, Chiu HS, Little MH. Generation of kidney organoids from human pluripotent stem cells. *Nat Protoc.* 2016;11(9):1681-1692. doi:10.1038/nprot.2016.098
325. Pal K, Forcato M, Ferrari F. Hi-C analysis: from data generation to integration. *Biophys Rev.* 2019. doi:10.1007/s12551-018-0489-1
326. Creighton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A.* 2010;107(50):21931-21936. doi:10.1073/pnas.1016071107