Published in final edited form as:

J Med Primatol. 2018 June ; 47(3): 157–171. doi:10.1111/jmp.12340.

The nonhuman primate kidney transcriptome in fetal development

Kimberly D Spradling-Reeves¹, Jeremy P Glenn², Kenneth J Lange³, Natalia Kuhn²,
Jacqueline J Coalson⁴, Mark J Nijland⁵, Cun Li⁶, Peter W Nathanielsz⁶, and Laura A Cox^{1,7,*}

¹Department of Internal Medicine, Wake Forest University Health Sciences, Winston-Salem, NC, USA

²Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX, USA

³Pharmaceuticals and Bioengineering Department, Southwest Research Institute, San Antonio, TX

⁴Department of Pathology, University of Texas Health Science Center, San Antonio, TX, USA

⁵Department of Obstetrics and Gynecology, University of Texas Health Science Center, San Antonio, TX, USA

⁶Department of Animal Science, University of Wyoming, Laramie, WY, USA

⁷Southwest National Primate Research Center, Texas Biomedical Research Institute, San Antonio, TX, USA

Abstract

Background—Little is known about the repertoire of nonhuman primate kidney genes expressed throughout development. The present work establishes an understanding of the primate renal transcriptome during fetal development in the context of renal maturation.

Methods—The baboon kidney transcriptome was characterized at 60 days gestation (DG), 90DG, 125DG, 140DG, 160DG and adulthood (6–12 years) using gene arrays and validated by QRT-PCR. Pathway and cluster analyses were used to characterize gene expression in the context of biological pathways.

Results—Pathway analysis indicated activation of pathways not previously reported as relevant to kidney development. Cluster analysis also revealed gene splice variants with discordant expression profiles during development.

Conclusions—This study provides the first detailed genetic analysis of the developing primate kidney, and our findings of discordant expression of gene splice variants suggest that gene arrays likely provide a simplified view and demonstrate the need to study the fetal renal proteome.

Keywords

baboon; gene expression; array; ontogeny; pathway analysis; cluster analysis

*Corresponding Author: Laura A. Cox, PhD, Department of Internal Medicine, Wake Forest University Health Sciences, Winston-Salem, NC 27157, Phone: 336-713-7506, laurcox@wakehealth.edu.

Introduction

The kidney plays a vital role in blood pressure and body fluid homeostasis during fetal and postnatal life. (1,2) In spite of the importance of fetal kidney development and growth, little is known about the repertoire of genes expressed during metanephric development in the nonhuman primate kidney. Understanding the composition and activity of the renal transcriptome at different stages of fetal development will provide a framework of normal expression profiles that facilitates the identification of alterations in development trajectory that ultimately may influence postnatal kidney function.

Current information on the genetics of kidney development primarily derives from studies in rodents. A comprehensive study of five fetal time points during mouse kidney development was performed and shows genes and pathways activated at each of these time points. (3) More recently, the Genitourinary Developmental Molecular Anatomy Project (4) has greatly enhanced understanding of the complexities in regional expression profiles in the developing mouse kidney using several techniques, including gene array analysis. (5) Based on morphological studies, there are similarities between rodent and baboon development of the early kidney. Kidney development begins in the mouse embryo at approximately embryonic day (E) 8.5–E9.5. The ureteric bud begins to form at E10.5–E11.25, which corresponds to approximately 45 days gestation (DG) in the baboon kidney. (6) The definitive metanephric kidney develops between E13.5–E15 in the mouse embryo. Similarly, the metanephric kidney is present by the end of the first trimester in humans (7) and rhesus macaques. (8) The lack of a detailed description of baboon kidney development from conception to adulthood at multiple time points has limited comparisons that can be made between baboons and rodents regarding developmental stages. (6)

Although the timeline of early rodent renal development appears similar to primates, later stages clearly differ with rodent kidney development continuing after birth, while simultaneously producing urine, and primates possessing much more mature kidneys at birth (see (7) for multispecies comparison). Batchelder et al (8) examined targeted gene expression of nine biomarkers of nephrogenesis using QRT-PCR in samples pooled from fetal rhesus monkeys during the first, second and third trimesters, at term, and in postnatal juveniles at 3–6 months of age and adults. In addition to temporal expression, these studies also described regional and cell type expression of the biomarkers of interest. Li et al. (9) used gene arrays to study expression in human fetal kidneys; however, they used pooled RNA samples from 13- to 24-week-old human fetal kidneys and did not provide a timeline of expression of the genome through fetal development.

Our goal was to expand available data on primate kidney development, by characterizing the genetic framework of fetal baboon kidney development, using whole genome expression profiling at multiple developmental time points. Multiple RNA samples were assessed at each time point from the end of the first third (60DG), the mid way point (90DG), the beginning (125DG) and near the end (160DG) of the last third of gestation (G), and in adults. Because of the extensive physiological and genetic similarities between baboons and humans, the genes and genetic pathways characterized at different stages of fetal baboon

kidney development can provide insights into the importance of particular genes and/or gene families in normal human kidney development.

Materials and Methods

Animal Care and Tissue Collection

All procedures were approved by the Institutional Animal Care and Use Committee at Texas Biomedical Research Institute and conducted in Association for Assessment and Accreditation of Laboratory Animal Care-approved facilities. Baboon (*Papio hamadryas*) fetuses were obtained by performing Cesarean sections on pregnant females at 60DG (0.3G, n=1 male, n=2 undetermined sex), 90DG (0.5G, n=3 males, n=3 females), 125DG (0.7G, n=3 males, n=3 females), 140DG (0.8G, n=3 males, n=3 females), and 160DG (0.9G, n=3 males, n=3 females) under isoflurane anesthesia (2%, 2 l/min). Techniques used and post-operative maintenance of the dams have been previously described in detail. (10) In addition, samples were collected opportunistically at necropsy from animals euthanized for health reasons unrelated to kidney function at postnatal day 2 (PN2; n=3 males, n=3 females) and at 6–12 years old (n=3 males, n=2 females). Kidney weights and body weights were collected at necropsy from the animals used in this study as well as other studies at the following time points: 90DG (n=5 males, n=5 females), 125DG (n=15 males, n=13 females), 140DG (n=11 males, n=17 females), 160DG (n=5 males, n=12 females), 175DG (n=15 males, n=8 females) and PN2 (n=12 males, n=14 females). Kidneys were collected and cut in half longitudinally. Aliquots were immediately frozen in liquid nitrogen and stored at –80°C until used for RNA extractions.

Gene Array Analysis

RNA was isolated from renal tissues using TRIzol Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions and as described in Cox et al. (11) For kidneys collected at 60d, the entire kidney was used; for those collected at later time points, a kidney wedge was used that included cortex and medulla. Total RNA samples for 60DG (n=1 male, n=1 undetermined sex), 90DG (n=3 males), 125DG (n=3 males), 160DG (n=3 males) and adult (6–12 years of age; n=3 males) were sent to Genome Explorations, Inc. (Memphis, TN) for RNA quality check, cRNA synthesis, and determination of gene expression profiles by interrogation of the Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA). Details of this protocol are published in Cox et al. (11) Gene expression was detected using GCOS software (Affymetrix; GEO accession number GSE65162). Array data were all-median normalized and log₂-transformed using GeneSifter software (GeneSifter.Net, VizX Labs, Seattle, WA), and statistical analyses of array data were performed by ANOVA assuming unequal variance (after box plot inspection). Cluster analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and gene ontology (GO) category analysis were then performed using GeneSifter software. Previous experiments using the Affymetrix Human Genome U133 Plus 2.0 Array show quality signal for 16,186 probes when using baboon cRNA; therefore, we used an *in silico* “baboon array” with KEGG pathways (www.genome.jp/kegg/) (12) and GO categories (<http://www.geneontology.org/>) (13) based on the maximum of 16,186 genes. (11)

QRT-PCR Quantification of Target Gene Abundance

Total RNA from 60d (0.3G, n=3, 1 male, 2 undetermined), 90d (0.5G, n=3 males, n=3 females), 125d (0.7G, n=2 males, n=3 females), 140d (0.8G, n=3 males, n=3 females) and 160d (0.9G, n=3 males, n=3 females) male and female renal tissues was used to measure mRNA levels of target genes with the Assay-on-Demand system (Applied Biosystems (ABI), San Francisco, CA) according to the manufacturer's instructions. Due to limited amounts of RNA from 60DG samples, different kidneys were used for QRT-PCR as used for gene arrays for this time point. Although no baboon-specific primers and probes were available, we have successfully used the Assay-on-Demand system for more than 80 different baboon genes (e.g., (11,14)). Gene-specific primers for *CCNB1*, *MAPK1*, *TP53*, *SMAD4*, *EIF4e*, and *LRP5* were provided by the manufacturer (CCNB1, Hs00259126_m1; MAPK1, Hs01052196_m1; TP53, Hs01034253_m1; SMAD4, Hs00232068_m1; EIF4e, Hs00913390_m1; and LRP5, Hs00182031_m1). 18S rRNA (Hs99999901_s1) and *MRPL48* (Hs00740658_m1) were quantified as endogenous controls, and all samples were assayed in triplicate.

The comparative threshold cycle (Ct) method (described in User Bulletin 2 for ABI PRISM® 7700 Sequence Detection Systems) was employed for relative quantification of gene expression.

Statistical Analyses

Kidney and Body Weights

Weights were analyzed in females and males among time points by ANOVA, adjusting p-values for multiple testing (GraphPad Prism 7.0a Software).

Gene Expression and Pathway Enrichment

Array data were all-median normalized and log₂-transformed (GeneSifter software). Statistical analyses of array data were performed by ANOVA (GeneSifter software). Differentially expressed genes were overlaid onto Ontological pathways (<http://www.geneontology.org/>) (Ashburner et al. 2000) and KEGG pathways (<http://www.genome.jp/kegg/>) (Kanehisa et al. 2004) for gene set enrichment analysis (GeneSifter software). Pathways were considered significantly altered if the z-score for that pathway was less than -2 or greater than 2. z-Scores were calculated in GeneSifter using the following formula:

$$z - \text{score} = (r - n(R/N)) / \sqrt{((n(R/N)(1 - (R/N))(1 - ((n - 1)/(N - 1))))}$$

where R = total number of genes meeting selection criteria, N = total number of genes measured, r = number of genes meeting selection criteria with the specified GO term, and n = total number of genes measured with the specific GO term (Doniger et al. 2003).

Cluster analyses were performed using the Partitioning Around Medoids (PAM) method (15) to reveal the different patterns of gene expression within the dataset. For example, which genes show increased abundance through development, which show decreased abundance

through development, and which increase at specific time points and then decrease at later time points. These results are suggestive of genes that are important at specific time points or throughout development. Silhouette scores of clusters are used to assess quality of clusters generated by the clustering algorithm. The number of clusters that best match a dataset is based on the greatest silhouette scores for the dataset. Overlaying genes within each cluster (i.e., genes with similar expression patterns through development) reveals whether pathways of genes are enriched for this expression profile.

Results

Morphometric Measurements of Kidneys

Kidney and total body weights for fetal baboons from 90DG to PN2 are shown in Figure 1. Fetal body weights show a steady increase from 90DG to 175DG for both males and females, but the profile of renal growth differed between the two sexes. Female kidney weights increased from 90DG to 140DG, plateaued from 140DG to 160DG and again increased from 160DG to 175DG; whereas, male kidney weights increased from 90DG to 160DG and plateaued from 160DG to PN2. Pairwise analysis showed significant differences in body weight between females and males at 125DG, 160DG and PN2. There were also significant differences in body weight between females at each consecutive time point and in males at each consecutive time point, except 175DG and PN2. Pairwise analysis showed marginal differences in kidney weight between females and males at 125DG and significant differences between males and females at 160DG. Comparison of kidney weights at consecutive time points showed significant differences in both sexes from 90DG to 140DG. There were also significant differences in kidney weights from 140DG to 160DG in males and from 160DG to 175DG in females.

Whole Genome Expression Profiling in Male Kidneys

Whole genome expression profiling was performed on kidneys from 60DG baboons of undetermined sex and male baboon kidneys from 60DG, 90DG, 125DG, 160DG and adults. There were 11,331 genes that passed quality filter and were denoted as “expressed” in at least one time point studied. Statistical analysis of the dataset by ANOVA showed 4,698 differentially expressed genes across the developmental time course (Table S1). The top 20 differentially expressed genes are shown in Table 1. Because some genes on the array were represented by more than one probe in order to detect expression for annotated splice variants, some genes appeared more than once on the list of expressed genes.

Pathway Analysis of Differentially Expressed Genes Across Development

Overlaying the differentially expressed genes onto KEGG pathways to provide biological context showed 31 significant pathways including adherens junction, cell cycle and MAPK signaling pathways (Table 2). GO analysis of differentially expressed genes identified 409 biological processes (Table S2) containing significant numbers of differentially expressed genes.

Pathway Analysis of Non-Differentially Expressed Genes

Of the 11,331 genes expressed during kidney development, 6,633 do not show changes in expression from 60DG to adult (Table S3). Overlaying these genes onto KEGG pathways showed 31 significant pathways (Table 3). Eight of these pathways were also significant for differentially expressed genes: adherens junction, complement and coagulation cascades, cytokine-cytokine receptor interaction, gap junction, MAPK signaling, neuroactive ligand-receptor interaction, prostate cancer and ubiquitin mediated proteolysis. In addition, abundance of 37 genes in the VEGF signaling pathway and 63 genes in the Wnt signaling pathway did not change through development.

Branching Morphogenesis Genes

Genes known to play a role in kidney branching morphogenesis include frizzled homolog 1 (*FZD1*), frizzled homolog 7 (*FZD7*), glypican 3 (*GPC3*), homeobox A11 (*HOXA11*), integrin alpha 6 (*ITGA6*), matrix metalloproteinase 2 (*MMP2*), secreted frizzled-related protein 1 (*SFRP1*), TIMP metalloproteinase inhibitor 1 (*TIMP1*), Wilms tumor 1 (*WT1*) and wingless-type MMTV integration site family member 5A (*WNT5A*). We observed three expression profiles for these genes: 1) Gene expression decreased from 60DG to adult (Table S4); 2) gene expression increased from 60DG to 160DG and then decreased from 160DG to adult (Table S4); and 3) gene expression increased from 60DG to 90DG and then decreased from 90DG to adult (Table S4). Genes that increased in expression from 60DG to 90DG and decreased from 90DG to adult included: *FZD1*, *FZD7*, *GPC3*, *HOXA11*, *MMP2*, *TIMP1* and *WNT5A*. *WT1* showed marginal decrease in expression (p-value = 0.07). *FZD1*, *FZD7* and *WT1* were each represented by two probes on the array detecting different splice variants of the gene. Interestingly, although both splice variants of *WT1* decreased from 160DG to adult, one splice variant showed a much greater decrease (1.0 to 0.3 versus 1.9 to -1.2). *SFRP1* was represented by three different probes on the array, with all three variants showing an increase from 90DG to 160DG and a decrease from 160DG to adult. *ITGA6* did not show expression changes from 60DG to adult (Table S4).

mTOR Signaling Genes

Fifty-one gene variants in the mTOR signaling pathway were expressed during kidney development. Twenty-three of these variants were differentially expressed during development; 10 showed marginal differential expression ($0.05 < p < 0.1$), and 18 did not show changes in expression. Among the differentially expressed genes, we observed two general patterns of expression: those genes that decreased in expression from 60DG to adult and those that increased from 60DG to adult. The majority of genes in the former category showed an incremental decrease at each time point measured; whereas, the gene expression patterns in the latter category were far more variable. For example, one variant of phosphoinositide-3-kinase catalytic beta polypeptide (*PIK3CB*) showed an oscillating expression pattern through development while phosphoinositide-3-kinase catalytic delta polypeptide (*PIK3CD*) increased from 60DG to 160DG and then increased dramatically from 160DG to adult. And similar to *WT1*, genes with multiple probes on the array that detect different splice variants show different expression profiles. For example, two *EIF4E*

splice variants were not differentially expressed during kidney development, but one *EIF4E* splice variant was differentially expressed with an oscillating expression pattern (Table S4).

Cell Cycle Genes

We detected expression of 150 genes related to cell cycle regulation; 77 of these genes were differentially expressed through development (Table S4). The differentially expressed genes fit into six general expression profiles: 1) Genes that increased expression from 60DG to adult. This group included one variant of SMAD family member 4 (*SMAD4*). 2) Gene expression increased from 60DG to 90DG and decreased from 90DG to adult (n=19). This group included tumor protein 53 (*TP53*), cyclin-dependent kinase 4 (*CDK4*) and cell division cycle 16 (*CDC16*). 3) Gene expression decreased from 60DG to 90DG, increased from 90DG to 125DG and decreased from 125DG to adult (n=24). This group included cyclin-dependent kinase 6 (*CDK6*), two splice variants of SMAD family member 2 (*SMAD2*) and one variant of cell division cycle 2 G1 to S and G2 to M (*CDC2*). 4) Gene expression decreased from 60DG to 125DG, increased from 125DG to 160DG and decreased from 160DG to adult (n=4). This group included cyclin E1 (*CCNE1*), E2F transcription factor 5 (*E2F5*) and cyclin D2 (*CCND2*). 5) Gene expression decreased from 60DG to 160DG and increased from 160DG to adult (n=6). This group included MAD2 mitotic arrest deficient-like 1 (*MAD2L1*) and cyclin B2 (*CCNB2*). 6) Gene expression decreased at each time point from 60DG to adult (n=19). This group included Cyclin B1 (*CCNB1*), one variant of cyclin A2 (*CCNA2*) and 2 variants of *CDC2* (Table S4).

Renin-Angiotensin System Genes

We detected expression in 9 of the 17 genes in the renin angiotensin system; 5 of these genes were differentially expressed. Among the differentially expressed genes were angiotensin II receptor type 1 (*AGTR1*) and angiotensin II receptor type 2 (*AGTR2*), both of which also had splice variants that were not differentially expressed. In addition, angiotensin I converting enzyme 2 (*ACE2*) was represented by two different splice variants, both of which were differentially expressed. *AGTR2* showed increased expression from 60DG to 90DG and decreased from 90DG to adult. One variant of *ACE2* decreased from 60DG to 90DG and the other decreased slightly from 160DG to adult. Membrane metallo-endopeptidase (*MME*) decreased slightly from 160DG to adult, Alanyl aminopeptidase (*ANPEP*) decreased from 90DG to 125DG and one variant of *AGTR1* decreased from 60DG to 90DG (Table S4).

Apoptosis Genes

KEGG pathway analysis of apoptosis signaling showed 83 genes expressed in the baboon kidney during development; 25 of these genes were differentially expressed including caspase 7 (*CASP7*), caspase 9 (*CASP9*) and *TP53*. Five of these genes have two splice variants that were differentially expressed and two of these genes have a splice variant that was not differentially expressed. There were three general expression patterns: 1) Gene expression increased from 60DG to adult and included genes such as tumor necrosis factor superfamily, member 10 (*TNFSF10*) and apoptosis-inducing factor mitochondrion-associated 1 (*AIFM1*) genes (n=9); 2) gene expression decreased from 60DG to 160DG and

increased from 160DG to adult (n=3); and 3) gene expression decreased from 60DG to adult and included two splice variants of *TP53* and one variant of *CASP7* (n=18) (Table S4).

Sodium Ion Transport Genes

GO analysis of the dataset showed 24 of 61 sodium ion transport gene variants that were differentially expressed. Four of these genes have two splice variants that were differentially expressed. Twenty genes show increased expression from 60DG to adult, including 12 solute carrier genes and one sodium transporter; four genes show decreased expression from 60DG to adult, including three solute carrier genes (Table S4).

Cluster Analysis of Differentially Expressed Genes

Cluster analysis by time point showed the adult kidney samples least similar to the other time points and the 125DG and 160DG time points most similar (Figure 2). Cluster analysis of differentially expressed genes indicated that two clusters best defined the dataset with a mean silhouette width of 0.408. Figure 3 shows a graphical representation of each cluster, and Figure 4 shows the heat map of gene expression profiles. Cluster 1 shows an increase in gene expression from 60DG to adult, and cluster 2 shows a decrease in gene expression from 60DG to adult.

Pathway Analysis of Cluster 1 & Cluster 2 Genes

KEGG pathway analysis of cluster 1 genes (Table S5) showed 56 pathways with significant z-scores and included cell cycle, MAPK signaling, p53 signaling and oxidative phosphorylation (Table 4). Analysis of cluster 2 genes (Table S6) showed 48 pathways and included adipocytokine signaling, ribosome assembly and toll-like receptor signaling (Table 5). In addition, there were 14 pathways common to both clusters including: citrate cycle, fatty acid metabolism, notch signaling, peroxisome, spliceosome and Wnt signaling. GO analysis of cluster 1 genes included 865 categories (Table S7), and cluster 2 genes included 827 categories (Table S8).

Gene Expression Unique to a Developmental Time Point

We evaluated gene expression at each time point and identified genes with expression unique to a specific time point. We found 73 genes expressed at 60DG that were not expressed at any other time point analyzed by gene array. Included in this list were *CCNA2*, H3 histone 3B (*H3F3B*) and solute carrier family 39 (zinc transporter) 9 (*SLC39A9*). We found 13 genes unique to 90DG, including apolipoprotein A-1 (*APOA1*) and TP53 binding protein 2 (*TP53BP2*). Twenty genes were unique to 125DG, including cell division cycle 42 (*CDC42*) and janus kinase 2 (*JAK2*). Seventeen genes were unique to 160DG, including insulin-like growth factor 1 (*IGF1*), prostaglandin E receptor 3 (*PTGER3*) and thrombomodulin (*THBD*). And 77 genes were unique to adult, including three cytochrome P450 genes (*CYP2C9*, *CYP2D6*, *CYP7B1*), estrogen receptor 1 (*ESR1*), carboxylesterase 2 (*CES2*), sodium channel voltage-gated type I beta (*SCN1B*) and eight solute carrier genes (*SLC10A2*, *SLC13A3*, *SLC13A3*, *SLC17A3*, *SLC22A6*, *SLC22A6*, *SLC22A18*, *SLCO4C1*) (Table S9).

Gene Expression Profile Validation

We validated expression profiles for genes central to top-ranking pathways, including *MAPK1*, *TP53*, *CCNG1*, *SMAD4*, *EIF4E* and *LRP5* using QRT-PCR (Figure 5). In addition to the time points included in the gene array analyses, kidneys from 140DG, 175DG and PN2 were included. Also included were RNA samples from kidneys of female baboons at each time point, with the exception of 60DG. QRT-PCR results confirm the array results and provide additional profile information with the addition of three time points and samples from females. For example, *MAPK1*, *SMAD4*, *LRP5* and *TP53* expression was significantly higher in males than females (p-values = 0.05) at 140DG, and *EIF4E* was significantly higher in males than females at 160DG.

Discordant Expression of Gene Splice variants

There were 1,873 of 7,182 annotated genes detected on the array represented by two or more probes that detected more than one splice variant. Of the genes with multiple splice variants detected on the array, 808 genes had at least one splice variant that was not differentially expressed and at least one splice variant that was differentially expressed; 108 genes had at least one splice variant that was differentially expressed in cluster 1 and at least one that was differentially expressed in cluster 2 (Table S10); and 34 genes included splice variants in clusters 1 and 2 and the non-differentially expressed group (Table 6). Genes in the latter category included translation initiation factors, ribosomal subunits, ubiquitin and predicted genes. An example of discordant gene expression for different splice variants is shown for the gene encoding the signal transduction protein *SMAD4*, which plays a central role in cell cycle regulation and Wnt signaling (Figure 6).

Overlaying cluster 1, cluster 2 and non-differentially expressed genes onto KEGG pathways revealed multiple gene splice variants and gene family members in the signaling cascade with discordant expression profiles. For example, the Wnt signaling pathway included 25 genes that were down-regulated during development, two genes that were up-regulated during development and 18 genes that showed no expression changes. Three genes, including *SMAD4*, had one splice variant that was up-regulated during development and one variant down-regulated during development. In addition, the frizzled family of genes, which is annotated as a single entity in the Wnt signaling pathway, shows some family members up-regulated and some down-regulated during development (Figure 7).

Discussion

We detected expression of 11,331 transcripts, including 7,182 annotated genes, during baboon kidney development. Of these transcripts, 4,698 (41%) were differentially expressed during development from 60DG to adult. Pathway analyses of gene expression data showed significant z-scores for pathways known to be relevant to kidney development such as cell cycle, citrate cycle, DNA replication, gap junction, MAPK signaling, ribosome, spliceosome, ubiquitin-mediated proteolysis, Wnt signaling and toll-like receptor signaling pathways. Also included were pathways with less obvious roles in kidney development such as antigen processing and presentation, cytokine-cytokine receptor interaction, neuroactive ligand-receptor interaction and NOD-like receptor signaling pathways. Identification of

genes expressed at only one time point revealed a small number of genes unique to the time points studied.

Interestingly, 6,633 (58.5%) expressed genes did not show changes in expression from 60DG to adult. The process of fetal kidney development is very dynamic with dramatic structural and cellular changes from the end of the first trimester (60dGA) until term. In spite of these differences, there are pathways enriched with genes whose expression is maintained throughout this developmental trajectory, suggesting that these pathways play essential roles in kidney function. Some pathways in this category are probably expected, such as adherens junction. However, other pathways are less likely predicted to include genes that do not change during fetal development such as MAPK signaling, insulin signaling and WNT signaling, steroid hormone biosynthesis, ubiquitin mediated proteolysis, VEGF signaling and Wnt signaling pathways. Overlaying non-differential gene expression profiles onto KEGG pathways showed pathways in common with differentially expressed genes and including MAPK signaling pathway, and ubiquitin mediated proteolysis pathways, suggesting still unknown roles of genes in these pathways during fetal kidney development.

We identified genes unique to each time point and found the greatest number at 60DG and adult. The relevance of some of these genes to kidney development (e.g., 60DG: exosome component 10 (*EXOSC10*) and fizzy/cell division cycle 20 related 1 (*FZR1*)) is not currently known. Higher resolution analysis using methods such as RNA-Seq may reveal splice variants unique to specific fetal time points.

Cluster analysis of differentially expressed genes revealed two clusters that best fit the data with one cluster containing genes with expression increasing from 60DG to adult and one cluster containing genes decreasing from 60DG to adult. KEGG pathway analysis of cluster 1 genes showed pathways common to the whole dataset of differentially expressed genes but also revealed additional pathways not seen in the complete dataset of differentially expressed genes (e.g., oxidative phosphorylation and Wnt signaling pathways). Interestingly, pathways and genes overlapped between clusters 1 and 2. That is, some pathways were statistically significant in both clusters. In many cases, one splice variant for a gene was found in cluster 1 with a different splice variant of that gene in cluster 2.

We found 1,873 genes with expression of two or more splice variants. Many of the genes that have been previously described as playing roles in kidney development showed discordant expression between splice variants. *WT1* was represented by two probes on the array, allowing detection of different splice variants. Although both splice variants of *WT1* decreased from 160DG to adult, one splice variant showed a much greater decrease than the other (1.0 to 0.3 versus 1.9 to -1.2). These results suggest that different splice variants of *WT1* may play different roles in kidney development or one may be more important than the other. Additional genes with discordant expression between splice variants included *EIF4E* and *SMAD4*. Two splice variants of the translational regulatory factor *EIF4E* were not differentially expressed while one splice variant showed decreasing expression during kidney development. The transcription regulatory factor *SMAD4* has been shown to be expressed in mice at E12 until after birth in a decreasing expression pattern (16). In our

study, we found two splice variants of *SMAD4* - one variant increased and one decreased during development. Our results suggest that different splice variants of some genes may play different roles or may function at different times during kidney development.

The phenomenon of discordant expression among splice variants for a gene is more complex when evaluated in the context of signaling pathways. For example, the Wnt signaling pathway, which is central to kidney development (Figure 7), included 25 genes that were down-regulated during development, two genes that were up-regulated during development and 18 genes that showed no expression changes. Three genes, including *SMAD4*, had one up-regulated splice variant during development and one down-regulated variant during development. In addition, the frizzled family of genes, which is annotated as a single entity in the Wnt signaling pathway, showed some family members up-regulated and some down-regulated during development. Our results showing discordant expression profiles for genes and gene splice variants within a signaling pathway suggest that understanding the molecular mechanisms directing kidney development and responding to environmental factors during kidney development may depend not only on gene expression but specific splice variants expressed at specific time points during development. If this is the case, then our study using gene arrays, which included major splice variants of some but not all genes and did not include minor splice variants, may not detect “minor” variants that play important roles during fetal renal development. Consequently, our study provides a preliminary glimpse of the complexity of gene expression during normal primate kidney development.

We selected genes in enriched pathways based on known activities of the gene products with highest priority assigned to genes with roles in multiple signaling pathways and specifically with roles in pathways known to be central to fetal kidney development, including MAPK1, TP53, SMAD4, EIF4E, and LRP5. MAPK1, which plays roles in 72 pathways, including adherens junction, apoptosis, chemokine signaling, circadian entrainment, dorso-ventral axis formation, GnRH signaling, insulin signaling, mTOR signaling, and VEGF signaling, encodes a member of the MAP kinase family (17). These kinases regulate multiple biochemical signals and are involved in a wide variety of cellular processes such as proliferation, differentiation, transcription regulation and development. Specific to renal development, MAPK1 regulates branching morphogenesis (18). Two alternatively spliced transcript variants detected in this study encode the same protein, but have different untranslated regions, suggesting different regulation of the two variants (19) (GenBank accession: DQ399292 and BC099905). TP53 plays roles in 42 pathways and encodes a tumor suppressor protein that regulates cycle arrest, apoptosis, senescence, DNA repair, and metabolism (17). TP53 has been shown to regulate nephrogenesis and contribute to nephron number through interactions with Pax2 (20). In addition, CCNG1, which influences multiple pathways through regulation of cell cycle signaling (17), has growth inhibitory activity mechanistically linked to TP53 signaling (21). SMAD4 is a signal transduction protein that plays roles in 16 pathways including adherens junction, cell cycle, TGF-beta signaling, and Wnt signaling (17). SMAD4 is known to impact cell fate during kidney development (16). EIF4E is a rate-limiting protein for initiation of mRNA translation, which specifically plays roles in seven signaling pathways (e.g., HIF-1 signaling and insulin signaling (17) and influences all signaling pathways regulated by protein abundance (22). In addition, EIF4E is

a rate limiting step in mTOR signaling, which plays a central role in energy sensing in the developing fetal kidney (23,24). LRP5 acts as a co-receptor with Frizzled protein family members for transducing signals by Wnt proteins, influencing mTOR signaling and Wnt signaling pathway (16), which are both known to be central to kidney development (25).

Inclusion of additional time points allowed us to better define the trajectory of gene expression for these key genes. *MAPK1* is involved in initiation and regulation of meiosis, mitosis, and postmitotic functions in differentiated cells by phosphorylating numerous transcription factors. We saw an interesting expression profile for *MAPK1* with high expression at 60DG that decreased at 90DG and then increased until 140DG–160DG with another decrease in adult. *TP53* is known to induce growth arrest or apoptosis, depending on the physiological status of the cell, and is a component of MAPK signaling, cell cycle regulation, p53 signaling, apoptosis signaling and Wnt signaling pathways. We observed increased expression from 60DG to 90DG and decreased expression from 90DG to adult. *CCNB1*, which is essential for control of the cell cycle at the G2/M transition, is a component of cell cycle regulation and p53 signaling pathways. *CCNB1* was most abundant at 60DG with subsequent decreased expression to the adult time point. *SMAD4* is a component of cell cycle regulation, Wnt signaling, TGF-beta signaling and adherens junction signaling pathways. We validated expression of the *SMAD4* splice variant that showed highest expression at 60DG with decreased expression until 160DG.

EIF4E is a rate limiting eukaryotic translation initiation factor and is a component of mTOR signaling and insulin signaling pathways. We found that *EIF4E* differed at 60DG between the gene arrays and QRT-PCR. This difference in expression profiles may be due to the location of the gene array probe versus the QRT-PCR primers. The gene array probe and PCR primers both anneal to major splice variants of *EIF4E*. However, they do not anneal to the same minor alleles. It is possible that currently annotated “minor” alleles are major alleles early in primate kidney development. Inclusion of additional time points for the QRT-PCR analysis shows that *EIF4E* expression does not peak until 175DG with a gradual decrease at PN2 to the adult rather than the precipitous change seen with gene arrays from 160DG to adult. *LRP5* binds and internalizes ligands in the process of receptor-mediated endocytosis and is a component of Wnt signaling pathway. Similar to *EIF4E*, we found some differences between the gene array expression profile and the QRT-PCR expression profile, most notably at 60DG and adult time points. And similar to *EIF4E*, the gene array probe and QRT-PCR primers anneal to the same major alleles but different minor alleles of the gene. Results for these two genes reiterate that this transcriptome analysis is a first step towards identifying the complexity of gene splice variants that are expressed during renal development.

Evaluation of kidney weights between females and males suggested different developmental trajectories; therefore, we included RNA samples from female kidneys for our targeted analyses of gene expression profiles. Our morphometric data showed a lag in body weight and kidney weights by females compared with males at 160DG with catch up by females to males at 175DG. *CCNB1* expression profiles were very similar between females and males showing decreased expression from 140DG to 160DG. However, *TP53*, *SMAD4*, *EIF4E* and *LRP5*, which all play roles related to cell proliferation, showed markedly different

expression profiles in females than males. *TP53* expression peaked at 125DG in females; whereas it did not show a marked peak expression from 90DG to 140DG followed by decreased expression in males. *SMAD4* also peaked at 125DG in females as opposed to the constant expression from 90DG to 125DG in males. In addition, *SMAD4* showed a much greater decrease from 125DG to 140DG in females than males. *EIF4E* expression also peaked at 125DG in females as opposed to 175DG in males. Similarly, *LRP5* expression peaked at 125DG–140DG in females, and peaked at 140DG in males with a more precipitous decrease from 140DG to 175DG in males than females. Taken together, kidney weights and targeted analysis of female kidney gene expression indicate different timing of molecular mechanisms that regulate kidney development. These results also suggest different windows of vulnerability between females and males. Future studies should be undertaken to analyze the female primate kidney transcriptome in detail.

Acknowledgments

Funding: The authors acknowledge grant support from NIH P01 HD021350 and facilities support from NIH Research Facilities Improvement Program Grants C06 RR013556 and C06 RR017515.

The authors acknowledge technical support by Ms. Ashley Melchoir.

References

1. Ingelfinger JR, Woods LL. Perinatal programming, renal development, and adult renal function. *American Journal of Hypertension*. 2002 Feb; 15(2 Pt 2):46S–9S. [PubMed: 11866229]
2. Rasch R, Skriver E, Woods LL. The role of the RAS in programming of adult hypertension. *Acta Physiologica Scandinavica*. 2004 Aug; 181(4):537–42. [PubMed: 15283768]
3. Schwab K, Patterson LT, Aronow BJ, Luckas R, Liang H-C, Potter SS. A catalogue of gene expression in the developing kidney. *Kidney Int*. 2003 Nov; 64(5):1588–604. [PubMed: 14531791]
4. McMahon AP, Aronow BJ, Davidson DR, Davies JA, Gaido KW, Grimmond S, et al. GUDMAP: The genitourinary developmental molecular anatomy project. *J Am Soc Nephrol*. 2008 Apr; 19(4):667–71. [PubMed: 18287559]
5. Potter SS, Brunskill EW, Patterson LT. Microdissection of the gene expression codes driving nephrogenesis. *Organogenesis*. 2010; 6(4):263–9. [PubMed: 21220959]
6. Gubhaju L, Black MJ. The baboon as a good model for studies of human kidney development. *Pediatr Res*. 2005 Sep; 58(3):505–9. [PubMed: 16148064]
7. Matsell DG, Tarantal AF. Experimental models of fetal obstructive nephropathy. *Pediatr Nephrol*. 2002 Jul; 17(7):470–6. [PubMed: 12172756]
8. Batchelder CA, Lee CCI, Martinez ML, Tarantal AF. Ontogeny of the kidney and renal developmental markers in the rhesus monkey (*macaca mulatta*). *Anat Rec (Hoboken)*. 2010 Nov; 293(11):1971–83. [PubMed: 20818613]
9. Li W, Kessler P, Williams BRG. Transcript profiling of wilms tumors reveals connections to kidney morphogenesis and expression patterns associated with anaplasia. *Oncogene*. 2005 Jan 13; 24(3):457–68. [PubMed: 15531917]
10. Schlambritz-Loutsevitch NE, Howell K, Rice K, Glover EJ, Nevill CH, Jenkins SL, et al. Development of a system for individual feeding of baboons maintained in an outdoor group social environment. *J Med Primatol*. 2004 Jun; 33(3):117–26. [PubMed: 15102068]
11. Cox LA, Nijland MJ, Gilbert JS, Schlambritz-Loutsevitch NE, Hubbard GB, McDonald TJ, et al. Effect of 30 per cent maternal nutrient restriction from 0.16 to 0.5 gestation on fetal baboon kidney gene expression. *J Physiol*. 2006 Apr 1; 572(Pt 1):67–85. [PubMed: 16513668]
12. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res*. 2004 Jan 1; 32(Database issue):D277–80. [PubMed: 14681412]

13. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: Tool for the unification of biology. The gene ontology consortium. *Nat Genet.* 2000 May; 25(1):25–9. [PubMed: 10802651]
14. Karere GM, Glenn JP, Birnbaum S, Rainwater DL, Mahaney MC, VanderBerg JL, Cox LA. Identification of candidate genes encoding an LDL-C QTL in baboons. *J Lipid Res.* 2013 Apr 17;54:1776–85. [PubMed: 23596326]
15. Dudoit S, Fridlyand J. A prediction-based resampling method for estimating the number of clusters in a dataset. *Genome Biol.* 2002 Jun 25;3(7):RESEARCH0036. [PubMed: 12184810]
16. Vrljicak P, Myburgh D, Ryan AK, van Rooijen MA, Mummery CL, Gupta IR. Smad expression during kidney development. *Am J Physiol Renal Physiol.* 2004 Apr; 286(4):F625–33. [PubMed: 14656760]
17. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 2017 Jan 4; 45(D1):D353–61. [PubMed: 27899662]
18. Fisher CE, Michael L, Barnett MW, Davies JA. Erk MAP kinase regulates branching morphogenesis in the developing mouse kidney. *Development.* 2001 Nov; 128(21):4329–38. [PubMed: 11684667]
19. Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, et al. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci U S A.* 2002 Dec 24; 99(26):16899–903. [PubMed: 12477932]
20. Saifudeen Z, Liu J, Dipp S, Yao X, Li Y, McLaughlin N, et al. A p53-pax2 pathway in kidney development: Implications for nephrogenesis. *PLoS One.* 2012; 7(9):e44869. [PubMed: 22984579]
21. Zhao L, Samuels T, Winckler S, Korgaonkar C, Tompkins V, Horne MC, Quelle DE. Cyclin G1 has growth inhibitory activity linked to the arf-mdm2-p53 and pRB tumor suppressor pathways. *Mol Cancer Res.* 2003 Jan; 1(3):195–206. [PubMed: 12556559]
22. Duncan R, Milburn SC, Hershey JW. Regulated phosphorylation and low abundance of heLa cell initiation factor eIF-4F suggest a role in translational control. Heat shock effects on eIF-4F. *J Biol Chem.* 1987 Jan 5; 262(1):380–8. [PubMed: 3793730]
23. Nijland MJ, Schlambitz-Loutsevitch NE, Hubbard GB, Nathanielsz PW, Cox LA. Non-human primate fetal kidney transcriptome analysis indicates mammalian target of rapamycin (mTOR) is a central nutrient-responsive pathway. *J Physiol.* 2007 Mar 15; 579(Pt 3):643–56. [PubMed: 17185341]
24. González A, Hall MN. Nutrient sensing and TOR signaling in yeast and mammals. *EMBO J.* 2017 Feb 15; 36(4):397–408. [PubMed: 28096180]
25. Carroll TJ, McMahon AP. Secreted molecules in metanephric induction. *J Am Soc Nephrol.* 2000 Nov; 11(Suppl 16):S116–9. [PubMed: 11065341]
26. Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, et al. The UCSC genome browser database: 2015 update. *Nucleic Acids Res.* 2015 Jan; 43(Database issue):D670–81. [PubMed: 25428374]

Figure 1. Kidney and body weight measures during baboon development

A) body weights and B) kidney weights for females and males at 90DG, 125DG, 140DG, 160DG, 175DG and PN2 are shown in gm. The * denotes $p < 0.05$, ** denotes $p < 0.01$ and *** denotes $p < 0.001$.

Figure 2. Cluster analysis of gene expression data

Similarities are denoted by the lengths of the horizontal lines in arbitrary units. For 60d, n=3, 1 male, 2 undetermined; for 90d – Adult, n=3 males per group.

Figure 3. Cluster analysis of gene expression for differentially expressed genes

The lines represent the median gene expression value for all genes in that cluster. The y-axis denotes the relative gene expression value (log₂-transformed), and the x-axis shows the time points included in the gene array expression analysis. A) Cluster 1 consists of 1,760 genes and has a mean silhouette width of 0.383. B) Cluster 2 consists of 2,938 genes and has a mean silhouette width of 0.432. For 60d, n=3, 1 male, 2 undetermined; for 90d – Adult, n=3 males per group.

Figure 4. Heat map of cluster analysis

Each column represents one time point, and each row represents expression of a specific gene. Genes are clustered by expression profiles for all time points. Clusters are indicated to the left of the heat map. For 60d, n=3, 1 male, 2 undetermined; for 90d – Adult, n=3 males per group.

Figure 5. Validation of gene array expression profiles by QRT-PCR

Gene expression profiles for *MAPK1*, *TP53*, *CCNB1*, *SMAD4*, *EIF4E* and *LRP5* are shown. The left graph shows the array data, and the right graph shows the QRT-PCR data. For gene array graphs, relative expression of log₂-transformed values are shown. For QRT-PCR graphs, RQ values are shown on the y-axis. The time points are shown on the x-axis. Females are shown with white bars and males with hashed bars. The * denotes $p < 0.05$ for comparison with 90d for females and \$ denotes $p < 0.05$ and Ψ denotes $p = 0.06$ for comparison with 90d for males. For 60d, $n = 3$, 1 male, 2 undetermined; 90d, 140d, 160d, 175d, PN and Adult, $n = 3$ females and $n = 3$ males per group; and 125d $n = 2$ females and $n = 3$ males per group.

Figure 6. Gene array expression of two *SMAD4* splice variants
A) The *SMAD4* gene alignment showing the splice variants of the gene detected by each specific Affymetrix probe are shown (UCSC Genome Browser, (26)). B) Expression profiles for the two *SMAD4* splice variants detected on the gene array. For 60d, n=3, 1 male, 2 undetermined; for 90d – Adult, n=3 males per group.

Figure 7. WNT signaling pathway annotated by clusters

Genes that are differentially expressed and included in cluster 1, which are up-regulated from 60DG to adult, are denoted by red boxes; genes that are differentially expressed and included in cluster 2, which are down-regulated from 60DG to adult, are denoted by green boxes; genes that are differentially expressed and gene splice variants or gene family members are found in cluster 1 and cluster 2 are denoted by blue boxes; genes that are not differentially expressed are denoted by gray boxes with black lines; and genes that were not detected on the array are denoted by black boxes with white font. This pathway was modified from the original KEGG pathway by adding color annotation. For 60d, n=3, 1 male, 2 undetermined; for 90d – Adult, n=3 males per group.

Table 1

Top 20 differentially expressed kidney genes from 60DG to adult.

Gene Name	Gene ID	p-value	60DG	90DG	125DG	160DG	Adult
casein kinase 2, alpha 1 polypeptide	CSNK2A1	0.00E+00	1.851±0.01	2.373±0.04	1.567±0.18	1.273±0.05	0.245±0
Gamma-glutamyltransferase light chain 1	GGTLC1	0.00E+00	0.135±0.01	0.218±0.07	0.304±0.05	0.214±0.13	2.013±0.02
Poly(A) binding protein interacting protein 1	PAIP1	0.00E+00	2.708±0.01	2.547±0.01	2.355±0.08	2.07±0.07	0.959±0.01
Pentraxin-related gene, rapidly induced by IL-1 beta	PTX3	1.00E-06	-0.652±0	-1.006±0.6	-1.666±0.01	-1.991±0.3	-1.485±0.5
SH3 domain binding glutamic acid-rich protein like 3	SH3BGR1.3	2.00E-06	1.558±0.04	1.77±0.27	1.043±0.3	0.771±0.21	-2.073±0.04
F-box and leucine-rich repeat protein 7	FBXL7	3.00E-06	0.808±0	1.155±0.09	0.691±0.09	0.396±0.11	0.031±0.01
HRAS-like suppressor 2	HRASLS2	3.00E-06	-3.08±0	-1.26±0.93	-1.405±0.46	-0.983±0.31	1.167±0.06
Splicing factor proline/glutamine-rich	SFPQ	4.00E-06	3.751±0.02	3.002±0.24	2.76±0.15	2.964±0.19	1.847±0.03
Acyl-CoA synthetase long-chain family member 1	ACSL1	5.00E-06	-0.21±0.04	-0.02±0.2	0.498±0.15	1.336±0.19	3.893±0.07
Splicing factor, arginine/serine-rich 3	SFRS3	5.00E-06	4.086±0	3.653±0.09	3.482±0.15	3.303±0.14	2.869±0.02
SRY (sex determining region Y)-box 4	SOX4	5.00E-06	1.747±0.03	1.981±0.27	1.878±0.37	1.303±0.35	-4.615±0.1
TRM5 tRNA methyltransferase 5 homolog	TRMT5	5.00E-06	4.984±0.15	4.906±0.02	4.585±0.03	4.244±0.05	3.115±0.01
Cyclin-dependent kinase 2 associated protein 1	CDK2AP1	6.00E-06	2.956±0.02	2.403±0.16	2.763±0.14	2.487±0.08	1.172±0.03
Forkhead box D1	FOXDI	6.00E-06	2.22±0.09	1.397±0.11	1.471±0.12	0.853±0.23	-3.504±0.07
Steroid sulfatase (microsomal), isozyme S	STS	6.00E-06	-0.347±0.01	-0.831±0.15	-1.08±0.01	-0.63±0.42	-0.349±0.28
Elastin	ELN	7.00E-06	1.058±0.05	3.126±0.17	2.7±0.11	1.998±0.34	-1.952±0.06
Gamma-glutamyltransferase 1	GGT1	7.00E-06	-0.086±0.02	0.497±0.18	0.49±0.03	0.696±0.28	2.678±0.05
Microsomal glutathione S-transferase 2	MGST2	7.00E-06	0.951±0.02	1.465±0.15	1.829±0.04	1.635±0.19	2.289±0.02
Hemoglobin, gamma A	HGB1	8.00E-06	5.505±0.12	5.866±0.17	4.737±0.44	3.246±0.17	-1.555±0.1
Thymine-DNA glycosylase	TDG	8.00E-06	1.721±0.02	1.005±0.08	0.943±0.19	0.414±0.02	-0.067±0.35

Table 2

KEGG pathways of differentially expressed genes during kidney development.

KEGG Pathway	Genes	Gene Set	z-score
Acute myeloid leukemia	26	50	2.1
Adherens junction	33	64	2.31
Antigen processing and presentation	34	57	3.44
Arginine and proline metabolism	23	44	2.01
Cell cycle	57	111	3.01
Citrate cycle (TCA cycle)	18	29	2.72
Complement and coagulation cascades	14	59	-2.23
Cytokine-cytokine receptor interaction	38	164	-3.93
DNA replication	20	31	3.09
Endometrial cancer	26	46	2.65
Gap junction	42	79	2.87
Glioma	28	55	2.04
Glycosphingolipid biosynthesis - ganglio series	1	13	-2.24
MAPK signaling pathway	61	201	-2.22
Neuroactive ligand-receptor interaction	29	165	-5.46
NOD-like receptor signaling pathway	10	45	-2.16
Non-small cell lung cancer	26	50	2.1
Oocyte meiosis	43	87	2.28
Primary immunodeficiency	2	19	-2.45
Progesterone-mediated oocyte maturation	35	69	2.25
Propanoate metabolism	15	26	2.11
Prostate cancer	40	79	2.4
Protein export	11	18	2.05
Proximal tubule bicarbonate reclamation	14	20	2.99
Ribosome	63	81	7.53
Spliceosome	58	104	3.86
Toll-like receptor signaling pathway	21	85	-2.5
Tryptophan metabolism	18	28	2.91
Ubiquitin mediated proteolysis	53	107	2.56
Valine, leucine and isoleucine degradation	22	39	2.42
Vibrio cholerae infection	27	47	2.81

Table 3

KEGG pathways of genes not differentially expressed during kidney development.

KEGG Pathways	Genes	Gene Set	z-score
Adherens junction	38	64	2.03
Alzheimer's disease	78	130	3.07
Autoimmune thyroid disease	12	39	-2.02
Complement and coagulation cascades	17	59	-2.8
Cytokine-cytokine receptor interaction	54	164	-3.65
Drug metabolism - cytochrome P450	16	49	-2
Endocytosis	101	158	4.41
Gap junction	28	79	-2.05
Glycine, serine and threonine metabolism	15	22	2.01
Hematopoietic cell lineage	14	51	-2.79
Insulin signaling pathway	63	107	2.54
MAPK signaling pathway	109	201	2.17
Metabolism of xenobiotics by cytochrome P450	12	46	-2.84
N-Glycan biosynthesis	23	36	2.06
Neuroactive ligand-receptor interaction	57	165	-3.24
Neurotrophin signaling pathway	62	102	2.87
Other glycan degradation	9	11	2.33
Parkinson's disease	56	94	2.51
Prostate cancer	46	79	2.05
Proteasome	25	36	2.73
Pyrimidine metabolism	39	64	2.28
Renal cell carcinoma	39	63	2.42
Retinol metabolism	9	34	-2.39
Small cell lung cancer	44	73	2.33
Steroid hormone biosynthesis	7	35	-3.2
Systemic lupus erythematosus	28	80	-2.14
Thyroid cancer	18	27	2.07
Tyrosine metabolism	8	30	-2.22
Ubiquitin mediated proteolysis	61	107	2.14
VEGF signaling pathway	37	55	3.06
Wnt signaling pathway	63	108	2.43

Table 4

KEGG pathways of differentially expressed genes in cluster 1.

KEGG Pathway	Genes	Gene Set	z-score
Alanine, aspartate and glutamate metabolism	9	24	2.72
Aldosterone-regulated sodium reabsorption	14	33	3.96
Allograft rejection	10	29	2.56
Antigen processing and presentation	24	57	5.15
Arginine and proline metabolism	19	44	4.71
Ascorbate and aldarate metabolism	5	12	2.31
Autoimmune thyroid disease	15	39	3.64
Axon guidance	6	97	-2.83
beta-Alanine metabolism	9	20	3.38
Biotin metabolism	2	2	3.15
Butanoate metabolism	13	29	4.06
Cell cycle	7	111	-3
Citrate cycle (TCA cycle)	17	29	6.05
Collecting duct acid secretion	8	21	2.62
D-Arginine and D-ornithine metabolism	1	1	2.23
D-Glutamine and D-glutamate metabolism	3	4	3.12
Drug metabolism - cytochrome P450	17	49	3.38
Drug metabolism - other enzymes	10	30	2.43
Fatty acid metabolism	19	38	5.51
Folate biosynthesis	4	9	2.22
Glutathione metabolism	11	38	2.02
Glycerolipid metabolism	11	31	2.8
Glycolysis / Gluconeogenesis	17	48	3.48
Glyoxylate and dicarboxylate metabolism	6	10	3.66
Graft-versus-host disease	10	28	2.69
Histidine metabolism	8	22	2.46
Limonene and pinene degradation	5	8	3.46
Lysosome	27	95	3.07
MAPK signaling pathway	23	201	-2.09
Metabolic pathways	216	806	8.65
Metabolism of xenobiotics by cytochrome P450	13	46	2.1
Natural killer cell mediated cytotoxicity	21	80	2.29
Nitrogen metabolism	7	17	2.7
Notch signaling pathway	1	35	-2.22
Oocyte meiosis	7	87	-2.21
Oxidative phosphorylation	28	99	3.1
p53 signaling pathway	3	53	-2.18
Parkinson's disease	23	94	2.02
Pathways in cancer	32	262	-2.06

KEGG Pathway	Genes	Gene Set	z-score
Pentose and glucuronate interconversions	6	13	2.84
Peroxisome	27	64	5.49
Phenylalanine metabolism	6	13	2.84
PPAR signaling pathway	19	54	3.64
Primary bile acid biosynthesis	6	11	3.36
Propanoate metabolism	15	26	5.6
Proximal tubule bicarbonate reclamation	13	20	5.78
Pyruvate metabolism	14	32	4.1
Retinol metabolism	15	34	4.28
Shigellosis	3	53	-2.18
Spliceosome	7	104	-2.78
Steroid hormone biosynthesis	12	35	2.78
Tryptophan metabolism	16	28	5.73
Tyrosine metabolism	11	30	2.92
Ubiquitin mediated proteolysis	10	107	-2.09
Valine, leucine and isoleucine degradation	22	39	6.66
Wnt signaling pathway	6	108	-3.17

Table 5

KEGG pathways of differentially expressed genes in cluster 2.

KEGG Pathway	Genes	Gene Set	z-score
Adherens junction	29	64	4.46
Adipocytokine signaling pathway	5	57	-2.47
Arachidonic acid metabolism	1	36	-2.83
Bacterial invasion of epithelial cells	19	53	2.39
Butanoate metabolism	1	29	-2.45
Cell cycle	51	111	6.08
Chronic myeloid leukemia	21	62	2.21
Citrate cycle (TCA cycle)	1	29	-2.45
Complement and coagulation cascades	2	59	-3.52
Cytokine-cytokine receptor interaction	15	164	-4.14
DNA replication	13	31	2.64
Dorso-ventral axis formation	8	18	2.26
Drug metabolism - cytochrome P450	4	49	-2.39
ECM-receptor interaction	24	71	2.35
Endometrial cancer	18	46	2.76
Ether lipid metabolism	1	25	-2.21
Fatty acid metabolism	2	38	-2.54
Focal adhesion	52	172	2.56
Gap junction	32	79	3.93
Glioma	20	55	2.52
Hematopoietic cell lineage	5	51	-2.16
Leishmaniasis	7	60	-2
Melanogenesis	24	67	2.68
Metabolic pathways	119	806	-5.84
Neuroactive ligand-receptor interaction	9	165	-5.32
Notch signaling pathway	13	35	2.12
Oocyte meiosis	37	87	4.59
Pathogenic Escherichia coli infection	24	52	4.16
Pathways in cancer	79	262	3.17
Peroxisome	3	64	-3.42
Porphyryn and chlorophyll metabolism	1	27	-2.33
Progesterone-mediated oocyte maturation	25	69	2.81
Propanoate metabolism	1	26	-2.27
Prostate cancer	30	79	3.38
Protein export	8	18	2.26
Regulation of actin cytoskeleton	48	162	2.29
Ribosome	59	81	11.05
RIG-I-like receptor signaling pathway	5	49	-2.05
Spliceosome	53	104	7.13

KEGG Pathway	Genes	Gene Set	z-score
Steroid hormone biosynthesis	2	35	-2.37
Thyroid cancer	11	27	2.31
Tight junction	33	104	2.34
Toll-like receptor signaling pathway	9	85	-2.63
Type I diabetes mellitus	1	35	-2.78
Tyrosine metabolism	1	30	-2.51
Ubiquitin mediated proteolysis	45	107	4.98
Valine, leucine and isoleucine degradation	1	39	-2.98
Wnt signaling pathway	42	108	4.21

Table 6

Genes with splice variants in clusters 1 and 2 and non differentially expressed gene categories.

Gene Title
Activin A receptor, type IB
Adaptor-related protein complex 2, alpha 2 subunit
Alpha thalassemia/mental retardation syndrome X-linked (RAD54 homolog, <i>S. cerevisiae</i>)
BTG family, member 3
Calmodulin 1 (phosphorylase kinase, delta)
Catenin (cadherin-associated protein), alpha 1, 102kDa
Family with sequence similarity 134, member A
Guanine nucleotide binding protein (G protein), alpha 11 (Gq class)
Heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa)
Nucleophosmin (nucleolar phosphoprotein B23, numatrin)
Phosphatidylinositol 4-kinase, catalytic, beta
Ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)
Amyloid beta (A4) precursor-like protein 2
BAT2 domain containing 1
Calmodulin regulated spectrin-associated protein 1-like 1
Dihydrofolate reductase
Eukaryotic translation initiation factor 1
Eukaryotic translation initiation factor 5B
G-rich RNA sequence binding factor 1
Glutaminase
Insulin-like growth factor binding protein 5
KIAA0494
KIAA1033
Neurofibromin 1
Neuronal PAS domain protein 2
PPPDE peptidase domain containing 1
Plakophilin 4
Prolyl endopeptidase-like
RAB GTPase activating protein 1
RAD17 homolog (<i>S. pombe</i>)
Ribosomal protein L11
Ribosomal protein L5
Small nuclear ribonucleoprotein polypeptide N
V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog