
Article:
Magayr, T.A., Song, X., Streets, A.J. orcid.org/0000-0002-4328-044X et al. (10 more authors) (2020) Global microRNA profiling in human urinary exosomes reveals novel disease biomarkers and cellular pathways for Autosomal Dominant Polycystic Kidney Disease. Kidney International. ISSN 0085-2538

https://doi.org/10.1016/j.kint.2020.02.008

Global microRNA profiling in human urinary exosomes reveals novel disease biomarkers and cellular pathways for Autosomal Dominant Polycystic Kidney Disease

Tajdida A Magayr[1#], Xuewen Song[2#], Andrew J Streets[1#], Laura Vergoz[1#], Lijun Chang[1], Manoj K Valluru[1], Hsiu L Yap[3], Morgane Lannoy[1], Amirreza Haghighi[2], Roslyn J Simms[1], Frederick WK Tam[3], York Pei[2*], Albert CM Ong[1*]

[1]Kidney Genetics Group, Academic Nephrology Unit, University of Sheffield Medical School, Sheffield, UK

[2]Division of Nephrology, University Health Network and University of Toronto, Toronto, Ontario, Canada

[3]Centre for Inflammatory Disease, Department of Immunology and Inflammation, Imperial College London, Hammersmith Hospital, Du Cane Road, London, UK

# equal contribution

*Address correspondence to:

Albert CM Ong
Academic Unit of Nephrology
Department of Infection, Immunity and Cardiovascular Disease
University of Sheffield Medical School
Beech Hill Road, Sheffield S10 2RX, UK
Tel: +44 114 215 9542
Fax: +44 114 271 3892
E-mail: a.ong@sheffield.ac.uk

York Pei
Division of Nephrology
University Health Network
8N838, 585 University Avenue
Toronto, Ontario, Canada M5G 2N2
Tel: 416-340-4257
Fax: 416-340-4999
E-mail: York.Pei@uhn.ca

Running title: Global microRNA profiling in ADPKD urine exosomes

ABSTRACT

MicroRNAs (miRNAs) play an important role in regulating gene expression in health and disease but their role in modifying disease expression in ADPKD remains uncertain. In this study, we profiled human urinary exosome miRNA by global small RNA-sequencing in an initial discovery cohort of ADPKD patients with early (eGFR >60ml/min/1.73m^2, n=7) and late (eGFR <60ml/min/1.73m^2, n=9) disease and compared their differential expression with age and sex matched healthy controls (n=6). Two candidate miRNA families identified (miR-192/miR-194-2 and miR-30) were selected for confirmatory testing in a validation cohort (n=60) by quantitative PCR. We confirmed that miR-192-5p, miR-194-5p, miR-30a-5p, miR-30d-5p and miR-30e-5p were significantly downregulated in human urine exosomes, in murine *Pkd1* cystic kidneys and in human *PKD1* cystic kidney tissue. All 5 miRNAs showed significant correlations with baseline eGFR and ultrasound-determined mean kidney length (MKL) and improved the diagnostic performance (AUC) of MKL for the rate of disease progression. Finally, inverse correlations of these two miRNA families with increased expression in their predicted target genes in human *PKD1* cystic tissue identified dysregulated pathways and transcriptional networks including novel interactions between miR-194-5p and two potentially relevant candidate genes, *PIK3R1* and *ANO1*. Our results identify a subset of urinary exosomal miRNAs that could serve as novel biomarkers of disease progression and suggest new therapeutic targets in ADPKD.

Keywords: ADPKD, microRNA, urine exosomes, PKD1

INTRODUCTION

Autosomal-dominant polycystic kidney disease (ADPKD) is one of the most common human monogenic diseases with an estimated prevalence of 1:1000 to 1:2500. It is characterized by the progressive development and enlargement of focal kidney cysts typically resulting in end-stage renal disease (ESRD) by the fifth decade [1]. Mutations of *PKD1* and *PKD2* account for ~85% and ~15% of ADPKD, respectively and together, they account for ~10% of prevalent patients with ESRD [1]. Polycystin-1 (PC1) and polycystin-2 (PC2), the proteins encoded by *PKD1* and *PKD2*, function as a receptor-ion channel complex that maintains normal tubular structure and function through multiple signaling pathways [2]. The renal phenotype of ADPKD is however highly variable and this variability is only partly accounted for by locus and allelic factors. These differences are particularly exemplified by striking intra-familial variability in the age of onset of ESRD observed in some families [3 4]. The modifying role of environmental factors acting through somatic mutations, epigenetic changes and non-coding regulatory RNAs are only just being explored.

MicroRNAs (miRNAs) are ~22nt non-coding regulatory RNAs that function as inhibitors of post-transcriptional gene expression in plants and animals [5]. Several lines of evidence have revealed the role of miRNAs in experimental models of PKD[6-16]. First, differentially expressed miRNAs have been identified in animal models[6-9]. Second, the miR-17~92 miRNA cluster is upregulated in the kidney of several models and overexpression of miR-17~92 miRNA cluster leads to cyst formation in mice[12]. Specifically, miR-17 represses *PKD1*, *PKD2* and *PPARA*[10-12, 16], whereas miR-17~92 inhibits *PKHD1* through the

transcriptional factor *HNF1B* [12]. Third, several other miRNAs have been implicated in the pathogenesis of PKD, including: miR-15a [7], miR-21 [13] and miR-200 miRNA family members [15]. However, there have been few consistent findings between different models and importantly, a relative lack of confirmatory human studies [17].

In this study, we have taken a systems biology approach to discover dysregulated miRNAs in human ADPKD. Starting with human urinary exosomes, we first identified differentially expressed miRNA between ADPKD patients and healthy controls by global miRNA-sequencing. Members of two kidney-enriched miRNA families (miR-192/miR-194-2 and miR-30) were then validated in a larger patient cohort and in *Pkd1* mouse kidneys of different ages. We then confirmed these findings in an independent global miRNA microarray analysis performed on normal and *PKD1* human cystic tissue. By integrating these miRNAs with previously identified dysregulated genes/gene pathways in ADPKD[18], we found a significant inverse correlation between expression of these down-regulated miRNAs and their predicted target genes including previously unreported interactions between miR-194-5p and two candidate genes, *PIK3R1* and *ANO1*. The use of global miRNA profiling has allowed us to identify a subset of urinary exosomal miRNAs as potential novel biomarkers of disease progression and indicate new therapeutic targets in ADPKD.

RESULTS

Differential expression of a subset of miRNAs in ADPKD urine exosomes by global miRNA sequencing

Urine exosome microRNA were extracted from a discovery cohort (n=22) whose baseline characteristics are shown in **Table 1;** they included healthy controls with normal kidney function (n=6), and ADPKD patients with early (eGFR >60ml/min/1.73m^2, n=7) and late disease (eGFR <60ml/min/1.73m^2, n=9). Although the mean age of the controls was lower than both patient groups, these differences were not statistically significant. There were expected differences in eGFR and serum uric acid between the groups. Mean kidney length and historical annualised eGFR decline (ml/min/yr over 5 years) were also significantly different between patients with early and late disease. Otherwise, all three groups were well matched.

Principal Component Analysis (PCA) of the miRNA sequencing data demonstrated clustering of individual patient groups with the clearest separation between controls and ADPKD patients with late disease; patients with early disease overlapped with both controls and those with late disease (**Figure 1A**). Heat maps further displayed separation of the three groups with the lowest expression of multiple individual miRNAs evident in the ADPKD group with late disease (**Figure 1B**).

A total of 23 miRNAs were identified by differential expression analysis to be significantly different between the three groups (**Figure 1C**). Of interest, all were significantly downregulated in ADPKD patients with late disease compared to the healthy controls. An intermediate pattern was observed in patients with early disease with a marked decrease in certain miRNA whereas others were less altered compared to controls (**Figure 1C**). In particular, expression of seven kidney-enriched miRNAs, ie the miR-194-2 cluster (miR-

192-5p, miR-194-2-5p), miR-194-1-5p and four members of the miR-30 family (miR-30a-5p, miR-30b-5p, miR-30d-5p, miR-30e-5p) were decreased (**Figure 1C**). Expression of all 7 miRNAs was lower in ADPKD patients compared to controls, with a significant change in patients with late disease. In patients with early disease, only miR194-1-5p and miR194-2-5p were significantly lower than controls (**Figure 1D**). Although these paralogs result in an identical mature miR-194 transcript, they are located on separate gene loci ie the miR-192-miR-194-2 cluster (11q13.1) and the miR-215-miR194-1 cluster (1q41.1) [19]. Interestingly, miR-192, miR-194 and miR-215 have been reported to be part of a human kidney-specific miRNA cluster [19] and miR-192 and miR-194 are also highly enriched in normal rat kidney with increased cortical expression relative to medulla [20].

Although less highly expressed, members of the miR-30 family display a restricted expression pattern in the amphibian pronephric kidney and mouse metanephric kidney where they play an essential role in promoting *Xenopus* terminal nephron differentiation (in particular miR-30a-5p but also miR-30d-5p, miR-30e-5p) [21] and in the normal maintenance of mouse podocytes [22]. Since the most abundant and tissue-specific miRNAs in the cell are likely to mediate target mRNA suppression [23], we chose to study five candidate miRNAs representing the miR-194-2 cluster and the miR-30 family.

Validation of differentially expressed kidney enriched exosomal miRNAs by qPCR

To validate these findings, patients were recruited from a tertiary specialist PKD clinic following-up over 500 patients [24]. A validation cohort of 40 patients (n=20 with eGFR>60, n=20 with eGFR<60) was selected based on matching for age, gender, baseline eGFR and

availability of urine samples. The sample size of 40 patients was calculated based on a historical eGFR slope < or > 3 ml/min/yr (over 5 years), and provided 80% power to detect an odds ratio of 2.5 and type 1 error (alpha level) of 0.05 using a z test. The baseline clinical characteristics of this cohort, which includes 20 healthy controls, are shown in **Table 2**. Overall, all three groups were well matched with no difference in baseline eGFR between controls and ADPKD patients with early disease. As expected, patients with late disease also had significantly bigger kidneys as evidenced by increased mean kidney length (MKL), more rapid historical eGFR decline and increased proteinuria (**Table 2**). Total Kidney Volume (TKV) or height-adjusted TKV (ht-TKV) measured from renal MRI was only available in 50% of patients in this cohort performed for clinical or other indications. MKL was significantly correlated with both TKV and ht-TKV (**Figure S1A**). In this subgroup, there was an enrichment of patients with Mayo Imaging Class 1C-E in patients with late disease (7/11) compared to those with early disease (4/9) (**Table 2**) [25]. There was however no difference in the frequency of *PKD1* mutations between early and late groups (**Table 2**).

Unlike the discovery cohort, serum uric acid was not different between the control and patient groups. There was a higher frequency of hypertension and use of angiotensin receptor blockers (ARB) or Angiotensin Converting Enzyme inhibitor (ACEi) in the group with late disease. There were no differences in relevant co-morbidities between early and late patient groups: type 2 diabetes mellitus (2 early v 1 late) or gout (3 early v 2 late).

miRNA expression for all five miRNAs was validated by qPCR in total urine exosome miRNA using Taqman-specific primers: primers for miR-194-5p detected transcripts from both miR-194-1 and miR-194-2 paralogues. As shown in **Figure 2A**, expression of all five miRNAs was significantly reduced in ADPKD patients with late disease compared to healthy controls. In patients with early disease, only miR-192-5p showed a significant decrease when compared to healthy controls (fold change 0.54) (**Figure 2A**). To investigate whether these changes were specific to ADPKD or reflected changes common to Chronic Kidney Disease (CKD), we analysed urine exosome samples from a separate CKD patient cohort (n=24) with Type 2 diabetes mellitus (T2DM) alongside a group of healthy controls (n=10). Although the mean age of this cohort was higher than the ADPKD patients with late disease in the validation cohort (63.3 ± 10.7 v 54.0 ± 9.8 yrs), their eGFR was comparable (37.4 ± 13.4 v 34.8 ± 12.6 ml/min/1.73m^2) (**Table 2**). In the CKD group, changes for these five miRNAs showed a different pattern to ADPKD but did not reach statistical significance (**Figure 2B**). However, there was a trend towards decreased miR-30a, miR-30d and miR-30e expression with an opposite trend to increased miR-192 and miR-194 expression compared to healthy controls (eGFR>90 ml/min/1.73m^2).

Validation of differentially expressed kidney enriched miRNAs in a *Pkd1* mouse model

To validate that our findings in urine exosomes were an accurate reflection of the cystic kidney, we isolated miRNAs from *Pkd1*^{nl,nl} mouse kidneys at different ages (2, 4 and 10 weeks) after birth, representing early, middle and late disease [26]. Differential expression of

all five miRNAs was then measured by qPCR in mutant kidneys and their normal littermates at the three time points.

As in human ADPKD urine exosomes, expression of all five miRNAs was significantly decreased in *Pkd1* mouse kidneys compared to wild-type littermates (**Figure 2C**). However, the temporal expression pattern for each miRNA was not identical. Changes in the miR-30 miRNA were observed in early disease (2 weeks) but were no longer significant in late disease (10 weeks). In contrast, expression of miR-192-5p and miR-194-5p were significantly decreased at all time points measured.

Profiling of differentially expressed kidney enriched miRNAs in human *PKD1* cystic and control kidney tissue

Next, we interrogated an independent miRNA microarray study that profiled the expression of 1,145 miRNAs in 16 renal cysts of different sizes and 6 minimally cystic tissue (MCT) derived from 4 polycystic kidneys or 4 normal renal cortical samples. By using the top 200 or 500 most variable miRNAs across all samples in an unsupervised hierarchical cluster analysis[27, 28] a matrix of Pearson correlation coefficients that indicate the degree of overall similarity between any two samples was derived (**Figure S2A**). We found that the MCT and normal renal cortical samples showed the most homogeneous pattern of global miRNA expression, while cyst samples exhibited greater heterogeneity, with 11 of 16 cyst samples clustered as a single group and others resembling the MCT group; the cyst clustering profile was independent of cyst size and patients. We also noted all technical replicates

clustered together, supporting the robustness and reproducibility of the miRNA gene expression profiles.

Differentially expressed miRNAs in *PKD1* renal cysts compared with MCT and normal renal cortical samples

To identify the miRNAs that best discriminated the cysts from MCT and normal renal cortical samples, we used SAM multiclass analysis[29]. At FDR ≤ 0.1% to control for multiple testing, 158 miRNAs with medium to high expression levels were identified as significant (**Table S1**) and were then subjected to supervised hierarchical clustering. We found all cyst samples clustered as a single group, while the MCT and normal renal cortical samples once again clustered as a second group (**Figure S2B**).

The examination of 158-miRNA gene clusters revealed 4 distinct miRNA expression patterns (**Figure 3A**): two major sets whose expressions were down or up-regulated solely in cysts vs. MCT and normal kidney samples, including 81 down and 53 up-regulated miRNAs, and two minor sets whose expressions were down or up-regulated in both cysts and MCT vs. normal kidney samples, including 18 down and 6 up-regulated miRNAs. It had been previously demonstrated that miRNAs in genomic clusters can be co-expressed [30]. Among the 158 dysregulated miRNAs, 76 were located in 23 genomic clusters with at least 2 miRNAs within a proximal 50kb distance on a chromosome **(Table S2).** In each of 19 genomic clusters, all differentially expressed miRNAs displayed concordance in their expression pattern suggesting polycistronic transcriptional regulation and co-expression of these miRNAs.

miR-192 and miR-194 have been shown to be more highly enriched in the renal cortex than the medulla [20]. In renal cysts, these miRNAs were among the miRNAs with the highest rank and fold changes based on SAM analysis (**Table S1**). According to the expression level of miRNAs in our 4 normal kidney tissue samples, the average signal intensity for miR-192-5p, miR-194-5p and miR-215-5p was 28000, 34000 and 4800, respectively, suggesting the mature forms of miR-192 and miR-194 are mainly from the mir-194-2-192 cluster in normal kidney. We confirmed these findings by qPCR in patient samples representing minimally cystic tissue (MCT), small cysts (SC <1ml), medium cysts (MC 10-30ml) and large cysts (LC >50ml). As shown in **Figure 3C**, significant decreases in all 5 miRNA were found even in MCT (non-significant reduction for miR-30a-5p) compared to normal tissue indicating that these changes were detectable in human kidney tissue at the earliest stages of disease.

miRNA expression correlates with markers of disease and can predict progression

We decided to use MKL was a surrogate marker of kidney volume instead of MRI-TKV in this cohort as TKV measurements were only available in 50% of the validation cohort (**Table 2**). MKL has also been shown to be as reliable as htTKV as a prognostic marker for GFR decline in the CRISP cohort [31]. In ROC analysis, we found that MKL was highly predictive of baseline eGFR (AUC 0.905) although less so in discriminating between slow and rapid progressors (AUC 0.634, see below) (**Figure S1B**).

All five microRNAs showed significant negative correlations with baseline eGFR and positive correlations with MKL indicating that microRNA expression decreased with disease progression (**Figure S3**). Linear regression analysis was performed to identify significant correlations between individual microRNAs and other clinical variables (**Table S3**). In univariate analysis, miR-30a-5p, miR-30e-5p and miR-192-5p each showed significant correlations with baseline eGFR, eGFR slope and MKL. Additionally, miR-30e-5p and miR-192-5p were significantly correlated with PCR whereas miR-30d-5p was significantly correlated with serum uric acid. Correlations between miR-194-5p with baseline eGFR (p=0.055) and MKL (p=0.069) just failed to reach significance (**Table S3**). In multiple regression analysis, baseline eGFR remained significantly associated with the miR-30-family while MKL remained significantly associated with miR-192-5p. MiR-30d-5p was also significantly associated with serum uric acid levels (**Table 3**).

The ability of urine exosomal microRNAs to discriminate between ADPKD patients with slowly versus rapidly progressive disease (annual eGFR slope < or > 3 ml/min/year over 5 years) in this cohort was next calculated using ROC curve analysis [32] (**Figure S4**). MiR-30e-5p showed the highest AUC (0.826) whereas MKL had the lowest ROC value (0.634). In this cohort, a combination of all five microRNAs (AUC 0.889) proved superior to MKL but the combination of both miRNA and MKL further improved the ability to discriminate patients with a slowly progressive disease from those with rapidly progressive disease (AUC 0.914) (**Table 4**).

Identification of predicted target genes of individual exosomal miRNAs

Next, we sought to identify the predicted target genes of miR-30 and miR192/miR194 families using the miRWalk database since these could potentially identify new therapeutic targets or disease biomarkers. A total of 11 significantly enriched pathways were identified using the DAVID pathway program (**Table S4**). Of interest, the top 4 pathways common to both miR families were related to Wnt signaling, MAPK signaling, ErbB signaling and Calcium signaling.

In silico predictions of novel miR-mRNA target genes

Our previous global mRNA profiling on cysts of different size and minimally cystic tissue from *PKD1* human polycystic kidneys identified 6764 (up: 4515; down: 2249) differentially regulated genes and 212 (128 up- and 84 down-regulated) pathways from a total of 20488 unique genes with gene symbols in renal cysts respectively[18]. To test which miRNA target genes were inversely correlated to the expression of the respective miRNAs in these samples, we compared the predicted miRNA target genes identified by TargetScan 5.1 [33] to the differentially regulated mRNAs [18] and obtained the number of predicted mRNAs that showed increased expression in *PKD1* renal cysts. At the $p < 4.67xE-04$ level using a two-tailed Fisher's Exact Test with the Bonferroni Method for multiple testing correction, 56 down-regulated miRNAs from 42 families had a significant number of target mRNAs that were up-regulated **(Figure 3B, Table S5).** Using GSEA [34], we tested 6319 target genes predicted for the 42 down-regulated miRNA families and observed enrichment for 50 gene pathways, 47 of which overlapped with the previously identified 128 up-regulated pathways in renal cysts [18] including mitogenic signaling pathways via RTKs, cell cycle, Wnt, Notch, TGFβ and EMT (**Table S6**). Specific miRNA-target genes for the miR-

192/miR-194-2 cluster and members of the mir-30 family (miR-30a-5p, miR-30b-5p, miR-30c-5p, miR-30d-5p and miR-30e-5p) were then identified and integrated with the differentially up-regulated genes from our previous study (**Figure 4**).

miR-192 and miR-194 regulate proliferation in ADPKD cyst assays.

Both miR-192 and miR-194 showed significant decreases at all time points in mouse ADPKD tissue and the strongest down-regulation in human ADPKD tissue. We therefore tested the effect of miR-192 and 194 mimics in a 3D matrigel cyst assay. Human ADPKD cells (OX161 c1) were transfected with miR mimics or miR negative control mimics and cultured in matrigel for 12 days to form cysts. Expression of both miR-mimics led to a significant decrease in cyst area compared to scrambled miR controls (**Figure 5A**). This was associated with a significant reduction in proliferation (Ki67 positive cysts) (**Figure 5B**) but not in apoptosis (cleaved-caspase 3 positive cysts) (**Figure 5C**).

Identification of a miR-194-5p/*PIK3R1* functional interaction

One of the genes predicted to be regulated by both miR-30 and miR-194 was *PIK3R1*. *PIK3R1* encodes p85α, the major regulatory subunit of phosphoinositide 3-kinase (PI3K) which associates with the catalytic subunit encoded by *PIK3CA*. PI3K is a master regulator of cellular function, being activated by multiple inputs to in turn regulate proliferation, apoptosis, vesicular transport and polarity amongst many reported functions[35]. We found that the expression of *PIK3R1* was highly upregulated by several fold (4-6x) in human and mouse cystic cells or kidneys compared to controls (**Figure 6A**). This was confirmed by immunoblotting for PI3K-p85α in human cystic cells (**Figure 6B, S5**). Of interest,

expression of *PIK3CA* and its protein (p110α) were similarly increased in human cystic cells although with a lower fold change (**Figure 6B, C**). Examination of the 3'UTR of *PIK3R1* revealed a conserved seed sequence for miR-194-5p, a finding not previously described in the DIANA-TarBase v8 database. Using a luciferase reporter assay, we demonstrate that binding of a miR-194-5p mimic to this sequence significantly reduced luciferase activity, confirming that *PIK3R1* is a functional mRNA target of miR-194-5p (**Figure 6D**). Mutation of the seed sequence eliminated the interaction with miR-194-5p demonstrating the specificity of this interaction (**Figure 6E**).

In a previous paper, we reported that human *PKD1* cystic cells showed a hyperproliferative response to IGF-1 compared to non-cystic controls and that this was mediated by both PI3K and ERK signaling pathways[36]. In the microarray of human ADPKD cysts reported in this paper, the IGF-1 signaling pathway was also highly upregulated ranking 9th overall (**Table S6**). To investigate the functional effects of *PIK3R1* in a relevant system, we tested the effect of IGF-1 on cyst formation in OX161c1, a human cystic line which expresses high levels of *PIK3R1* and PI3K-p85α (**Figure 6A, B**). IGF-1 stimulated cyst growth in these cells, a response that was completely inhibited by either knockdown of *PIK3R1* (**Figure S5**) or the chemical PI3K inhibitor LY29004 (**Figure S6**). These findings confirm that the upregulation of *PIK3R1* is functionally relevant to cyst expansion in ADPKD.

Identification of a miR-194-5p/*ANO1* functional interaction

To further validate the usefulness of this approach to identify new therapeutic targets, we investigated the possible interaction between mir-194-5p and the calcium-activated

chloride channel, anoctamin-1 (ANO1), a prediction also not described in the DIANA-TarBase v8 database. ANO1 had been identified as a potential contributor to epithelial chloride (and hence fluid) secretion in *in vitro* cyst models although its functional role *in vivo* has not yet been proven[37]. In a recent microarray study, we found that *ANO1* expression was highly up-regulated in human *PKD1* cystic epithelial cells[38]; increased expression of *ANO1* was also observed in the *PKD1* renal cyst microarray (with 3.7-fold vs MCT, data not shown). *In silico* analysis predicted that *ANO1* mRNA was a predicted target for miR-194-5p although this had not been experimentally validated.

We first confirmed that *ANO1* expression was differentially increased in cellular and mouse models of ADPKD. *Ano1* expression was increased by 3.7 and 2.5-fold respectively in 4 week and 10 week *Pkd1* mice compared to their wild-type littermates (n=6 per group) (**Figure 7A**). In mouse embryonic kidney cells (MEK) derived from E15.5 collecting ducts of a different *Pkd1* mutant[39], a 2.9 fold increase in *Ano1* was also observed in mutant compared to wild-type cells (n=3, data not shown). We next profiled *ANO1* expression in four human *PKD1* cystic lines (OX161C1, OX938, SKI001, SKI002) and compared them to 5 human control kidney lines (UCL93, RFH94, CL5, CL8, CL11) [36]. Overall, the mean fold changes in *PKD1* cystic cells were significantly different from control cells although there was a greater variation in *ANO1* expression between the four lines compared to *PIK3R1* (**Figure 7B**).

The 3'UTR of *ANO1* mRNA contains two predicted binding seed sequences for miR-194-5p. By deletion analysis, we showed that the first of these (993-999bp) bound to miR-194-

5p in luciferase reporter assays, while the second (1258-1264bp) was neutral (**Figure 7C**). Mutation of the seed sequence eliminated the interaction with miR-194-5p demonstrating the specificity of this interaction (**Figure 7C**). Blocking ANO1 expression (by siRNA) or action (with the specific antagonist CaCCinh-A01) inhibited the growth of human *PKD1* cystic cells (SKI001) in 3D cyst assays (**Figures 7D, S5**). Our results extend previous *in vitro* data using dog (MDCK) and embryonic mouse kidneys to human disease cells[37]. Importantly, we identify and confirm for the first time a hitherto unknown miR-194-5p/*ANO1* interaction. Interestingly, *ANO1* is located close to the miR-192-miR194-2 cluster on 11q.13. Overall, these findings support a translational strategy to target ANO1 in ADPKD.

DISCUSSION

In this study, we have taken an unbiased approach to identify differentially expressed miRNA from human ADPKD urine exosomes using global miRNA sequencing. Unlike experimental models, we were able to correlate our findings with commonly used clinical markers of disease such as baseline eGFR, eGFR slope and MKL as a relevant surrogate for total kidney volume (TKV) to test their potential as disease biomarkers.

Surprisingly, we only identified 23 exosomal miRNAs that were significantly altered in ADPKD compared to healthy controls in a discovery cohort. Nevertheless, the identification of several members of the kidney-enriched miR-30 family and the miR-192/miR-194 gene cluster led us to validate these miRNAs in a larger validation cohort. The changes determined in urine exosome miRNA were confirmed in kidney miRNA

isolated from an established mouse *Pkd1* model showing that they were reflective of intrinsic changes in the diseased kidney. We also confirmed these changes in a unique sample set derived from human ADPKD cysts including tissue from minimally cystic tissue (MCT) and cysts of different sizes. These findings revealed that for these five differentially expressed miRNAs, a decrease was already detectable in MCT indicating these are likely to be primary or early changes in the disease process. Furthermore, by selecting potential mRNA targets recognized by these miRNA from previously identified upregulated genes in the same tissues[18], we were able to identify several enriched pathways and potential target genes. Finally, as an example of potential identification of novel therapeutic targets, we provide the first evidence of a functional interaction between miR-194-5p and two genes, *PIK3R1* and *ANO1*.

During the writing of this manuscript, another group reported that miR-192-5p and miR-194-5p were significantly downregulated in human end-stage ADPKD kidney tissue by differential miR microarray [40]. Of interest, these changes were negatively correlated to hypermethylation of the common promoter for miR-192/194-2 in ADPKD compared to non-ADPKD kidneys. In this study, a total of 65 miRs were differentially expressed by at least 1.5-fold. Of relevance to our results, miR-30b-5p and miR-30d-5p were also found to be downregulated although interestingly, without promoter hypermethylation. We conclude that apart from promoter hypermethylation, other factors may result in downregulation of mature miRNA expression in ADPKD. These could include the deregulation of transcriptional control (eg p53, HNF, TAZ/YAP) or secondary changes in the miRNA biogenesis pathway (eg Dicer)[41-43]. Of interest, several key transcription factors implicated in other ADPKD models (p53, HNF1A, HNF1B, HNF4A) have also been

shown to regulate the miR-192/miR-194-2 cluster (Figure 4) [44, 45]. Otherwise there was limited overlap with the miRs that we identified in our study. Of interest, we also detected differential expression of two miR reported to be more enriched in renal medulla (miR-27b, miR-99a) [20] although these were not detected in this study. These differences could relate to the stage of disease studied (end-stage kidneys) or be secondary to changes in renal function.

To investigate whether these miRNAs were specifically altered in ADPKD or common to other forms of CKD, we analysed an independent cohort comprised of patients with CKD secondary to T2DM. Our results indicate that in contrast to ADPKD, miR-192-5p and miR-194-5p showed a trend of increased expression relative to controls. This suggests that the changes we have detected are likely to be specific to ADPKD for these two miRNA. Of interest, miR-192-5p expression was increased in the renal cortex and glomeruli of streptozotocin-induced diabetic mice where it was functionally linked to diabetic glomerular hypertrophy, mesangial expansion and proteinuria [46]. Similar increases in miR-192-5p have been reported in studies of renal biopsy tissue from patients with IgA nephropathy and hypertensive nephrosclerosis [47, 48] probably relevant to glomerular-specific changes in miR-192-5p expression (see below).

For miR-30a-5p, miR30b-5p and miR-30e-5p, we noted a similar non-significant trend to decreased expression compared to controls. This could indicate that changes for the miR30 family may be common to both ADPKD and T2DM although further confirmatory studies will be required. Nonetheless, they could point to clinically and biologically important pathways common to CKD regulated by miR-30 members. A

previous study in microdissected nephron segments obtained from male adult rat kidneys showed miR-192-5p and miR-194-5p were highly enriched in proximal tubule (PT) segments compared to the medullary thick ascending limb (mTAL) while miR-30a-5p and miR-30e-5p showed the reverse; all four miRNA were enriched in tubular segments compared to glomeruli [49]. Our results suggest the likelihood of exosomes arising from multiple cysts of different segmental origin or non-cystic tubules (and glomeruli) that might be secondarily involved in the disease process (eg due to compression, ischaemia or inflammation) in a non-cell autonomous manner. The use of urine exosomes could therefore provide a global picture of the many processes determining disease progression *in vivo*, both cystic and non-cystic, in ADPKD kidneys [50]. A previous study linked miR-192-5p in regulating the expression and function of the β1 subunit of Na^+/K^+-ATPase in rat PT [49] and WNK1 expression in mouse distal tubule (DT) segments [51]. Given their importance in regulating tubular sodium reabsorption, the reduction in miR-192-5p observed may also be relevant to the pathophysiology of hypertension observed early in ADPKD [52].

One striking difference between our results and those reported from rodent models was the lack of common differentially expressed miRNAs. We did not identify upregulated miRNA such as those reported in murine *Pkd1* models (miR17-92 cluster) or miR-21[12, 13]. Possible explanations are differences in species, pathogenic mutations, disease pathogenesis, disease stage, tissue type or methodology used. For instance, a study in a mouse model of PKD (*Kif3a-KO*) found up-regulation of miR-17~92 in 21d-old and 28d-old (cyst growth) but not in 10d-old and 14d-old mice (cyst initiation)[12]. The discrepancy in the expression

of miR-17~92 cluster between our results and this study may due to differences in species (human vs mouse), mutation types (*PKD1 vs Kif3A*) as well as disease stages (late vs middle).

In summary, we have identified a novel subset of urinary exosomal miRNAs that could serve as potential biomarkers of disease progression and suggest new therapeutic targets in ADPKD. These results are clearly relevant to human disease but were also confirmed in a mouse model of ADPKD. By taking a global approach using urinary exosomes and human cystic tissue, we have also identified several other miRNA families for future study. Integrating these findings within a wider gene regulatory network (eg derived from the microarray dataset of upregulated genes) will be the next step since typically, multiple miRNAs may regulate one mRNA, and one miRNA may target several hundred different genes, thus reflecting cooperative translational control and target multiplicity [53]. The changes we have described make an important contribution to understanding disease pathogenesis, defining new disease biomarkers and identifying new drug targets to slow disease progression in ADPKD.

MATERALS and METHODS

Study Participants and Ethical Approval

Ethical approval for this study was obtained from the National Research Ethics Service Committee Yorkshire and The Humber Bradford (REC12/YH/0297) and from the REB of the Toronto General Hospital (02-0036).

Isolation of urinary extracellular vesicles and extraction of miRNAs

Urinary extracellular vesicles (UEVs) or exosomes were isolated by differential centrifugation as previously reported[38]. In brief, 10 ml of urine were initially centrifuged at 17,000 g for 15 min. The supernatant was retained and the pellet was resuspended in 200 µl isolation buffer (250 mM sucrose, 10 mM triethanolamine pH 7.6, 50 µl of DTT) to reduce the highly abundant Tamm-Horsfall protein. Following incubation at room temperature for 5 min, the sample was briefly vortexed and centrifuged again at 17,000 g for 15 min. This supernatant was then combined with the initial supernatant and centrifuged once more at 170,000 g for 150 min. Urinary exosomal associated microRNAs were isolated from the exosome pellet using a miRCURY™ RNA cell and plant isolation kit (Exiqon) according to the manufacturers protocol. Due to the discontinuation of this kit, we used the miRNeasy Micro kit (Qiagen) for the later study on patients with CKD.

Differential expression of miRNAs

The raw reads obtained were processed using bcl2fastq conversion software (Illumina). The data was de-multiplexed according to their specific barcode codes into FASTQ format. Before proceeding to identify differentially expressing microRNAs, an overview of the

data quality was performed. The quality of generated data was checked using the default settings of Strand NGS. The quality check of FASTQ data is based on a PHRED quality score [54-56] and all the sequencing reads from this study had a quality score > 30, denoting that the base call accuracy of overall reads is 99.9%. The size distributions of sequencing reads were evaluated for all samples showing reads were compatible with the size range for miRNA (19-24nt). A total of 74.22% of the mean sequencing reads were mapped to the human genome (hg39). To detect and profile miRNAs in our small RNA high-throughput sequencing data we used Strand NGS software by creating a small RNA analysis experiment. Default parameters were used for alignment quality controls and for further data analysis steps. The expression level of the raw counts was normalised between samples using the default DeSeq normalisation algorithm. Normalised counts were then log transformed to yield normalised signal values used for all the statistical tests and plots. False discovery rate (FDR) cutoff value for differentially expressed miRNAs between control and treated samples was set at FDR = 0.05 and the expression threshold was adjusted as $p<0.05$. Data has been deposited under ArrayExpress accession number E-MTAB-8109.

ROC analysis

A ROC curve calculation relies on the fraction of true positives (study sensitivity) and false positives (1 - sensitivity). The area under the curve (AUC), confidence interval (CI), p values of all calculated ROC curves and the ROC curves for each individual microRNA

are displayed together with cut-off values of the differentially expressed microRNAs together with their corresponding sensitivity and specificity (**Table 4, Figure S3**).

Statistical analysis

Data are presented as means values ± SE. Student's t-test or ANOVA was used for statistical analysis with a $P < 0.05$ indicating statistical significance. The degree of significance as determined by GraphPad Prism software is denoted as follows: *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$.

See the Online Supplement for description of other methods used in this study.

ACKNOWLEDGEMENTS

TAM was supported by a PhD scholarship from the Libyan Government. We thank Paul Heath and Shuang Feng for assistance, Dorien Peters for access to *Pkd1* mouse tissues and all the patients who participated in this study. This work was supported by grants from the European Union (EU-FP7/2007-2013, grant agreement no. 317246, TranCYST) and the Sheffield Kidney Research Foundation to ACMO, the Canadian Institutes of Health Research (MOP 77806) to YP. FWKT is supported by the Ken and Mary Minton Chair of Renal Medicine. Part of this work was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre, Imperial College Healthcare NHS Trust and Imperial College London.

CONFLICT OF INTEREST STATEMENT. FWKT has received research project grants from AstraZeneca Limited, Baxter Biosciences, Boehringer Ingelheim, MedImmune and Rigel Pharmaceuticals, and has consultancy agreements with Rigel Pharmaceuticals, Novartis and Baxter Biosciences.

FIGURE LEGENDS

Figure 1. miRNAseq identifies differential expression of miRNAs from ADPKD urinary exosomes

(A) Principal component analysis of the miRNAseq data from healthy control, early (GFR>60) and late (GFR<60) ADPKD patients generated using ClustVis software. (B) Heat map showing expression of differentially expressed miRNAs in individual healthy control, early (GFR>60) and late (GFR<60) ADPKD patients generated using ClustVis software. (C) Expression of differentially expressed miRNAs in individual healthy control, early (GFR>60) and late (GFR<60) ADPKD patients. (D) Expression of individual kidney enriched miRNAs by DeSeq normalised signal values between healthy control, early (GFR>60) and late (GFR<60) ADPKD patients.

Figure 2 Quantitative PCR (qPCR) validation of differentially expressed kidney enriched miRNAs in human urine exosomes and *Pkd1* mouse kidney

(A) qPCR showed significantly lower expression of all 5miRNAs in late ADPKD (GFR<60) compared to healthy controls. (B) qPCR showed no significant difference in all 5 miRNA expression between healthy controls and a non ADPKD CKD cohort. (C) Differential expression of miR-30a-5p, miR-30d-5p, miR-30e-5p, miR-192-5p and miR-194-5p in *Pkd1* mice *($Pkd1^{nl,nl}$)* at 2, 4 and 10 week old. Values are expressed as fold change over $Pkd1^{wt,wt}$ at the same age. Values are expressed as means ±SEM, *n*=5-6 animals per group. **P*<0.05; ***P*<0.01; ****P*<0.001; *****P*<0.0001 compared to $Pkd1^{wt,wt}$ at the same age. Statistical significance was determined using an unpaired t-test.

Figure 3 Quantitative PCR (qPCR) validation of differentially expressed kidney enriched miRNAs in human tissue samples

(A) At FDR of 0.1%, SAM identified 158 miRNAs with differential expression among the 3 groups: cysts, MCT and normal renal cortical samples. The 158-miRNA gene clustering depicts 4 groups of miRNAs. Abbreviations: SC (small cysts); MC (medium cysts); LC (large cysts); MCT (minimally cystic tissue); KIDNEY (normal renal cortical tissue), and re (technical replicate). (B) Pathways/genes predicted to be affected by 42 down-regulated miRNA families in PKD1 renal cysts (NOM p $\leqslant$ 0.05 and FDR $\leqslant$ 0.25 by GSEA). (C) Differential expression of miR-30a-5p, miR-30d-5p, miR-30e-5p, miR-192-5p and miR-194-5p in human tissue samples was determined by Taqman qPCR. Values are expressed as fold change over normal renal cortex tissue. Statistical significance was determined using an unpaired t-test. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001

Figure 4 Schematic summary of the aberrantly activation of pathways (Growth factors/RTKs, Notch, Wnt/β-catenin, and TGFβ signaling pathways) in PKD1 renal cysts predicted to be affected by the selected down-regulation of miRNAs (miR-192/194 and miR-30 family members, green letters in colored boxes).

All experimentally supported miRNA-target gene pairs were obtained from DIANA-TarBase v8 (http://www.microrna.gr/tarbase). Genes whose expression was up or down regulated are also shown in red or green letters with fold changes respectively (cyst vs. minimal cystic tissue; the number in bracket denotes fold-change). Genes that were not differentially expressed are shown in black. * denotes miRNA negatively regulates a given

target gene expression at the translational level. The mechanisms known to contribute to the downregulation of miR-192/194 are summarized in rectangle box.

Figure 5 Inhibition of cyst growth in 3D matrigel cultures by miR-192 and 194 mimics. (A) Representative images of day 12 cysts following transfection with 50nM miR-192/194 or negative control mimics. Mean cyst area was significantly reduced following transfection with miR mimics. (B) The number of cysts stained with the proliferation marker Ki67 was significantly reduced in cyst transfected with miR mimics indicating a reduction in the rate of proliferation. (C) No significant difference in the number of cleaved caspase-3 stained cysts was observed between miR mimics and negative controls.

Figure 6 The miR-194-5p target gene *PIK3R1* is increased in ADPKD models and mediated IGF-1-mediated cyst growth

(A) Significant over-expression of *PIK3R1* or *Pik3r1* in ADPKD models detected by qPCR. Left panel: normal (WT, N=2) v *PKD1* cystic human cells (N=4) - fold-change of 3.9 (n=3); Middle panel: *Pkd1wt,wt* (N=10) v *Pkd1nl,nl* (N=10) mice at 4 and 10 weeks - fold change of 3.3 and 2.26 respectively (n=3); Right panel: MEK wild-type (+/+) v MEK mutant (-/-) cells - fold change of 5.8 (n=3). (B) Over-expression of p85-α and p110-α in human *PKD1* cystic cells (n=4) compared to non-cystic cells (n=5) by Western blotting with a p85α/actin ratio value of 5.7-fold compared to WT and a p110α/actin ratio of 2.4 fold. (C) Increased expression of *PIK3CA* by qPCR in four human *PKD1* cystic cell lines compared to two normal cell lines (n=3, 2.3-fold increase). (D) A predicted miR-194-5p site was identified in *PIK3R1* between bp 2094-2100 in the 3'UTR (seed sequences shown in red) and cloned

into the pmirGLO luciferase reporter plasmid. Significant suppression of luciferase activity was observed in cells transfected with *PIK3R1* 3'UTR and a miR-194-5p mimic. The non-targeting miRNA mimic used as a negative control had no effect on luciferase activity. (E). Mutation of two bp in the miR-194-5p seed sequence in the 3'UTR of *PIK3R1* eliminated the suppression of luciferase activity demonstrating a specific interaction between miR-194-5p and the 3'UTR of *PIK3R1*. Mann Whitney test and SEM, **** $p < 0.0001$. Unpaired t-test and SEM, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Figure 7 The miR-194-5p target gene *ANO1* is overexpressed in ADPKD models and its inhibition reduces cyst growth

(A) *Ano1* expression by qPCR was significantly increased in 4 week (3.7-fold) and 10 week (2.5-fold) *Pkd1* mice ($Pkd1^{nl/nl}$) compared to control littermates (n= 6 per group). *p<0.05; **p<0.01; (B) Expression of *ANO1* was significantly increased in human *PKD1* cystic cell lines (n=4) compared to control non-cystic lines (n=5) with wide variation shown by fold changes of 1,059 (OX161c1), 20 (SKI001) and 200 (OX938). *p<0.05; (C) Two predicted miR-194-5p sites were identified in *ANO1* by TargetScan between bp 993-999 and 1258-1264 in the 3'UTR of *ANO1* (seed sequences shown in red) and cloned into the pmirGLO luciferase reporter plasmid. Significant suppression of luciferase activity was only seen in cells transfected with *ANO1* 3'UTR between bp 993-999 and a miR-194-5p mimic. The non-targeting miRNA mimic used as a negative control had no effect on luciferase activity. Mutation of two bp in the miR-194-5p seed sequence in the 3'UTR of *ANO1* eliminated the suppression of luciferase activity demonstrating a specific interaction between miR-194-5p and the 3'UTR of *ANO1* between bp 993-999. (D) 3D matrigel cyst assays carried

out in the *PKD1* line, SKI-001, in the presence of forskolin (10μM). Cells were treated with the ANO1 selective antagonist CaCCinh-A01 (10 μM, A01), DMSO (control), a control scrambled siRNA (50 nM, Neg) or anti-ANO1 siRNA (50 nM, siRNA). Cyst areas were measured after 7, 12 and 19 days (n=80-179 cysts per time-point). ****/$$$$ Mann-Whitney test and SEM. $p < 0.0001$. Representative images of cysts incubated with DMSO (control), CaCCinh-A01 (A01), negative control siRNA (Neg) or anti-ANO1 siRNA (siRNA) at 7, 12 and 19 d respectively. Scale bar: 25 μM

Table 1 Basic demographics and clinical features of the discovery cohort

Characteristic	Healthy controls (Mean ± SEM)	ADPKD patients with eGFR > 60 (Mean ± SEM)	ADPKD patients with eGFR < 60 (Mean ± SEM)	P value
Number	6	7	9	-
Gender (F:M)	3:3	4:3	4:5	0.881
Age (years)	46.00 ± 11.64	54.43 ± 10.98	53.67 ± 9.179	0.2987
BMI (kg/m^2)	28.22 ± 3.513	27.00 ± 2.708	25.33 ± 2.646	0.1872
eGFR (ml/min/1.73 m^2)	96.95 ± 24.55	99.71 ± 34.14	34.33 ± 17.56	**< 0.0001**
MKL (cm)	nd	13.28 ± 2.126	18.42 ± 2.041	**0.0004**
eGFR slope (ml/min/1.73 m^2)	nd	-0.3333 ± 3.470	-4.829 ± 2.105	**0.0151**
Calcium (mmol/l)	2.348 ± 0.04119	2.333 ± 0.05376	2.283 ± 0.07297	0.1106
Phosphate (mmol/l)	1.083 ± 0.1938	1.127 ± 0.1262	1.237 ± 0.1982	0.2437
Uric acid (µmol/l)	375.0 ± 90.40	259.0 ± 29.69	414.0 ± 133.4	**0.0186**
Cholesterol (mmol/l)	5.567± 0.9688	4.686 ± 0.8821	4.533 ± 0.9836	0.1259
PCR (mg/mmol)	nd	11.00 ± 1.732	24.88 ± 18.4	0.2399

nd = not determined

MKL = mean kidney length (ultrasound)

PCR = Protein Creatinine Ratio

Table 2 Basic demographics and clinical features of the validation cohort

Characteristic	Healthy controls (Mean ± SEM)	ADPKD patients with eGFR > 60 (Mean ± SEM)	ADPKD patients with eGFR < 60 (Mean ± SEM)	P value
Number	20	20	20	-
Gender (M:F)	10:10	10:10	12:8	0.765
Age (years)	46.15 ± 12.62	47.80 ± 11.79	54.00 ± 9.830	0.0829
BMI (kg/m^2)	25.94 ± 3.475	30.12 ± 4.554	26.56 ± 3.974	**0.0064**
eGFR (ml/min per 1.73 m^2)	92.16 ± 23.25	95.45 ± 24.47	34.80 ± 12.64	**< 0.0001**
MKL (cm)	nd	13.82 ± 2.243	18.61 ± 3.440	**< 0.0001**
eGFR slope (ml/min/year)	nd	-0.2429 ± 3.283	-3.688 ± 2.085	**0.0017**
Calcium (mmol/l)	2.320 ± 0.07152	2.333 ± 0.06789	2.291 ± 0.08811	0.2232
Phosphate (mmol/l)	1.097 ± 0.1547	1.097 ± 0.1547	1.157 ± 0.1945	0.4468
Uric acid (µmol/l)	311.0 ± 74.15	327.3 ± 90.32	390.5 ± 136.5	0.0559
Cholesterol (mmol/l)	5.330 ± 0.9836	5.211 ± 1.149	4.675 ± 0.8735	0.1001
PCR (mg/mmol)	nd	11.08 ± 6.487	25.47 ± 21.82	**0.0361**
PKD1/PKD2/NMD	nd	13/3/4	14/2/4	
Mayo htTKV Class (1AB/1CDE/2)	nd	6/4/1 (n=11)	2/7/0 (n=9)	
Hypertension (Y/N)	0/20	15/5	19/1	
ARB/ACEi (Y/N)	0/20	10/10	15/5	

nd = not determined

PCR = Protein Creatinine Ratio

MKL = mean kidney length (ultrasound)

NMD = no mutation detected

HtTKV = height adjusted Total Kidney Volume (Magnetic Resonance Imaging)

ARB = Angiotensin Receptor Blocker, ACEi = Angiotensin Converting Enzyme Inhibitor

Table 3 Multivariate analysis of selected differentially expressed microRNAs
Multivariate regression model adjusted to the variables which have significant correlations with the corresponding differentially expressed microRNAs.

Variables	miR-30a-5p Standardized β (P value)	miR-30d-5p Standardized β (P value)	miR-30e-5p Standardized β (P value)	miR-192-5p Standardized β (P value)	miR-194-5p Standardized β (P value)
eGFR (ml/min/1.73m^2)	**-0.561 (0.01)**	**-0.317 (0.036)**	**-0.386 (0.04)**	-0.290 (0.12)	
eGFR slope (ml/min/year)	-0.215 (0.16)		-0.214 (0.16)	-0.236 (0.12)	
MKL (cm)	0.075 (0.655)		0.134 (0.442)	**0.533 (0.000)**	
Uric acid (µmol/l)		**0.394 (0.010)**			
PTH (mmol/l)	-0.035 (0.842)				
PCR (mg/mmol)					

PTH = Parathyroid hormone
PCR = Protein Creatinine Ratio

Table 4 ROC analysis for the differentially expressed microRNAs for dichotomised eGFR slopes. A perfect diagnostic test has AUC of 1 while a value of 0.5 indicates weak prediction ability. 95% CI indicates 95 confidence limits of the true AUC of the population. The cut-off value for each differentially expressed microRNAs, beside mean kidney length, were calculated based on the value that provided maximum sensitivity and specificity for ADPKD progression. Receiver Operating Curve (ROC), Area Under the Curve (AUC), CI; Confidence Intervals (CI), Mean Kidney Length (MKL).

Variables	AUC	95% CI	P value	Cut-off value (delta CT)	Sensitivity	Specificity
miR-30a-5p	0.770	0.620-0.920	**0.004**	-2.49	77.3%	70.6%
miR-30d-5p	0.765	0.620-0.911	**0.004**	0.088	72.7%	64.7%
miR-30e-5p	0.826	0.695-0.956	**0.000**	-1.89	72.7%	82.4%
miR-192-5p	0.785	0.644-0.926	**0.002**	1.03	81.8%	70.6%
miR-194-5p	0.677	0.510-0.844	0.057	1.42	72.7%	41.2%
MKL	0.634	0.453-0.815	0.157	14.55 cm	72.7%	41.2%
Combination of microRNAs	0.889	0.790-0.988	**0.000**	/	/	/
Combination of microRNAs and MKL	0.914	0.820-1	**0.000**	/	/	/

SUPPLEMENTARY MATERIAL

Supplementary Methods

Supplementary Figures and Tables

Supplementary Figure S1: (A) Linear correlations between MKL and TKV or ht-TKV in validation cohort where both measurements were available (n=19). (B) ROC analysis for MKL in predicting baseline eGFR (AUC 0.905) or discriminating eGFR slope (AUC 0.634).

Supplementary Figure S2: miRNA expression profiles of normal kidneys, MCTs and renal cysts based on the standardized signal intensities of (A) the top 500 most variable miRNAs and (B) 158 dysregulated miRNAs, derive a matrix of Pearson correlation coefficients that indicate the degree of overall similarity between any two samples.

Supplementary Figure S3: miRNA expression correlates with eGFR and mean kidney length.

Supplementary Figure S4: ROC analysis of differentially expressed individual microRNAs to predict rapid disease progression in ADPKD.

Supplementary Figure S5: (A) Knockdown of *PIK3R1* or *ANO1* in the human cystic lines, OX161c1 or SKI-001 respectively, measured by qPCR; (B) Uncropped blots of PI3K-p85α and PI3K-p110α following IGF-1 stimulation (see Figure 5B)

Supplementary Figure S6: Knockdown of *PIK3R1* or chemical inhibition of PI3K inhibits IGF-1 stimulated cyst growth.

Supplementary Table 1: Differentially expressed miRNAs in human PKD1 renal cysts identified by SAM (FDR ≤ 0.1%)

Supplementary Table 2: Dysregulated miRNA clusters within a proximal 50kb distance on a chromosome in PKD1 renal cysts.

Supplementary Table 3: Linear regression analysis for selected clinical and biochemical variables identifying independent associations with differentially expressed microRNAs.

Supplementary Table 4: Pathway analysis using DAVID based on changes in the five differentially expressed miRNA.

Supplementary Table 5: Selected down-regulated miRNAs in human PKD1 renal cysts identified by SAM (FDR <0.1%), whose expression were also inversely correlated with the expression of their target genes ($p < 4.67E-04$).

Supplementary Table 6: Pathways predicted to be affected by 42 downregulated miRNA families as identified by GSEA at NOM $p \leq 0.05$ and FDR ≤ 0.25.

Supplementary material is linked to the online version of the paper at www.kidney-international.org.

REFERENCES

1. Ong AC, Devuyst O, Knebelmann B*, et al.* Autosomal dominant polycystic kidney disease: the changing face of clinical management. *Lancet* 2015; **385:** 1993-2002.

2. Ong AC, Harris PC. A polycystin-centric view of cyst formation and disease: the polycystins revisited. *Kidney Int* 2015; **88:** 699-710.

3. Paterson AD, Magistroni R, He N*, et al.* Progressive loss of renal function is an age-dependent heritable trait in type 1 autosomal dominant polycystic kidney disease. *J Am Soc Nephrol* 2005; **16:** 755-762.

4. Pei Y. Nature and nurture on phenotypic variability of autosomal dominant polycystic kidney disease. *Kidney Int* 2005; **67:** 1630-1631.

5. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell* 2009; **136:** 215-233.

6. Pandey P, Brors B, Srivastava PK*, et al.* Microarray-based approach identifies microRNAs and their target functional patterns in polycystic kidney disease. *BMC Genomics* 2008; **9:** 624.

7. Lee SO, Masyuk T, Splinter P*, et al.* MicroRNA15a modulates expression of the cell-cycle regulator Cdc25A and affects hepatic cystogenesis in a rat model of polycystic kidney disease. *J Clin Invest* 2008; **118:** 3714-3724.

8. Pandey P, Qin S, Ho J*, et al.* Systems biology approach to identify transcriptome reprogramming and candidate microRNA targets during the progression of polycystic kidney disease. *BMC Syst Biol* 2011; **5:** 56.

9. Dweep H, Sticht C, Kharkar A*, et al.* Parallel analysis of mRNA and microRNA microarray profiles to explore functional regulatory patterns in polycystic kidney disease: using PKD/Mhm rat model. *PloS one* 2013; **8:** e53780.

10. Tran U, Zakin L, Schweickert A*, et al.* The RNA-binding protein bicaudal C regulates polycystin 2 in the kidney by antagonizing miR-17 activity. *Development* 2010; **137:** 1107-1116.

11. Sun H, Li QW, Lv XY*, et al.* MicroRNA-17 post-transcriptionally regulates polycystic kidney disease-2 gene and promotes cell proliferation. *Mol Biol Rep* 2010; **37:** 2951-2958.

12. Patel V, Williams D, Hajarnis S*, et al.* miR-17~92 miRNA cluster promotes kidney cyst growth in polycystic kidney disease. *Proc Natl Acad Sci U S A* 2013; **110:** 10765-10770.

13. Lakhia R, Hajarnis S, Williams D*, et al.* MicroRNA-21 Aggravates Cyst Growth in a Model of Polycystic Kidney Disease. *J Am Soc Nephrol* 2016; **27:** 2319-2330.

14. Ben-Dov IZ, Whalen VM, Goilav B*, et al.* Cell and Microvesicle Urine microRNA Deep Sequencing Profiles from Healthy Individuals: Observations with Potential Impact on Biomarker Studies. *PLoS One* 2016; **11:** e0147249.

15. Patel V, Hajarnis S, Williams D*, et al.* MicroRNAs regulate renal tubule maturation through modulation of Pkd1. *Journal of the American Society of Nephrology : JASN* 2012; **23:** 1941-1948.

16. Hajarnis S, Lakhia R, Yheskel M*, et al.* microRNA-17 family promotes polycystic kidney disease progression through modulation of mitochondrial metabolism. *Nat Commun* 2017; **8:** 14395.

17. Kim DY, Woo YM, Lee S*, et al.* Impact of miR-192 and miR-194 on cyst enlargement through EMT in autosomal dominant polycystic kidney disease. *FASEB J* 2019; **33:** 2870-2884.

18. Song X, Di Giovanni V, He N*, et al.* Systems biology of autosomal dominant polycystic kidney disease (ADPKD): computational identification of gene expression pathways and integrated regulatory networks. *Hum Mol Genet* 2009; **18:** 2328-2343.

19. Sun Y, Koo S, White N*, et al.* Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs. *Nucleic Acids Res* 2004; **32:** e188.

20. Tian Z, Greene AS, Pietrusz JL*, et al.* MicroRNA-target pairs in the rat kidney identified by microRNA microarray, proteomic, and bioinformatic analysis. *Genome Res* 2008; **18:** 404-411.

21. Agrawal R, Tran U, Wessely O. The miR-30 miRNA family regulates Xenopus pronephros development and targets the transcription factor Xlim1/Lhx1. *Development* 2009; **136:** 3927-3936.

22. Wu J, Zheng C, Fan Y*, et al.* Downregulation of microRNA-30 facilitates podocyte injury and is prevented by glucocorticoids. *J Am Soc Nephrol* 2014; **25:** 92-104.

23. Mullokandov G, Baccarini A, Ruzo A*, et al.* High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries. *Nat Methods* 2012; **9:** 840-846.

24. Thong KM, Ong AC. The natural history of autosomal dominant polycystic kidney disease: 30-year experience from a single centre. *QJM : monthly journal of the Association of Physicians* 2013; **106:** 639-646.

25. Irazabal MV, Rangel LJ, Bergstralh EJ*, et al.* Imaging classification of autosomal dominant polycystic kidney disease: a simple model for selecting patients for clinical trials. *J Am Soc Nephrol* 2015; **26:** 160-172.

26. Happe H, van der Wal AM, Salvatori DC*, et al.* Cyst expansion and regression in a mouse model of polycystic kidney disease. *Kidney international* 2013; **83:** 1099-1108.

27. Eisen MB, Spellman PT, Brown PO*, et al.* Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998; **95:** 14863-14868.

28. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 2001; **98:** 31-36.

29. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001; **98:** 5116-5121.

30. Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* 2005; **11:** 241-247.

31. Bhutani H, Smith V, Rahbari-Oskoui F*, et al.* A comparison of ultrasound and magnetic resonance imaging shows that kidney length predicts chronic kidney disease in autosomal dominant polycystic kidney disease. *Kidney Int* 2015.

32. Rosansky SJ, Glassock RJ. Is a decline in estimated GFR an appropriate surrogate end point for renoprotection drug trials? *Kidney Int* 2014; **85:** 723-727.

33. Friedman RC, Farh KK, Burge CB*, et al.* Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009; **19:** 92-105.

34. Subramanian A, Tamayo P, Mootha VK*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; **102:** 15545-15550.

35. Fruman DA, Chiu H, Hopkins BD*, et al.* The PI3K Pathway in Human Disease. *Cell* 2017; **170:** 605-635.

36. Parker E, Newby LJ, Sharpe CC*, et al.* Hyperproliferation of PKD1 cystic cells is induced by insulin-like growth factor-1 activation of the Ras/Raf signalling system. *Kidney Int* 2007; **72:** 157-165.

37. Buchholz B, Faria D, Schley G*, et al.* Anoctamin 1 induces calcium-activated chloride secretion and proliferation of renal cyst-forming epithelial cells. *Kidney Int* 2014; **85:** 1058-1067.

38. Streets AJ, Magayr TA, Huang L*, et al.* Parallel microarray profiling identifies ErbB4 as a determinant of cyst growth in ADPKD and a prognostic biomarker for disease progression. *Am J Physiol Renal Physiol* 2017; **312:** F577-F588.

39. Nauli SM, Alenghat FJ, Luo Y*, et al.* Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nat Genet* 2003; **33:** 129-137.

40. Kim DY, Woo YM, Lee S*, et al.* Impact of miR-192 and miR-194 on cyst enlargement through EMT in autosomal dominant polycystic kidney disease. *FASEB J* 2018**:** fj201800563RR.

41. Pichiorri F, Suh SS, Rocci A*, et al.* Downregulation of p53-inducible microRNAs 192, 194, and 215 impairs the p53/MDM2 autoregulatory loop in multiple myeloma development. *Cancer Cell* 2010; **18:** 367-381.

42. Hino K, Tsuchiya K, Fukao T*, et al.* Inducible expression of microRNA-194 is regulated by HNF-1alpha during intestinal epithelial cell differentiation. *RNA* 2008; **14:** 1433-1442.

43. Chaulk SG, Lattanzi VJ, Hiemer SE*, et al.* The Hippo pathway effectors TAZ/YAP regulate dicer expression and microRNA biogenesis through Let-7. *J Biol Chem* 2014; **289:** 1886-1891.

44. Krutzfeldt J, Rosch N, Hausser J*, et al.* MicroRNA-194 is a target of transcription factor 1 (Tcf1, HNF1alpha) in adult liver and controls expression of frizzled-6. *Hepatology* 2012; **55:** 98-107.

45. Morimoto A, Kannari M, Tsuchida Y*, et al.* An HNF4alpha-microRNA-194/192 signaling axis maintains hepatic cell function. *J Biol Chem* 2017; **292:** 10574-10585.

46. Putta S, Lanting L, Sun G*, et al.* Inhibiting microRNA-192 ameliorates renal fibrosis in diabetic nephropathy. *J Am Soc Nephrol* 2012; **23:** 458-469.

47. Wang G, Kwan BC, Lai FM*, et al.* Intrarenal expression of microRNAs in patients with IgA nephropathy. *Lab Invest* 2010; **90:** 98-103.

48. Wang G, Kwan BC, Lai FM*, et al.* Intrarenal expression of miRNAs in patients with hypertensive nephrosclerosis. *Am J Hypertens* 2010; **23:** 78-84.

49. Mladinov D, Liu Y, Mattson DL*, et al.* MicroRNAs contribute to the maintenance of cell-type-specific physiological characteristics: miR-192 targets Na+/K+-ATPase beta1. *Nucleic Acids Res* 2013; **41:** 1273-1283.

50. Chang MY, A CMO. Targeting new cellular disease pathways in autosomal dominant polycystic kidney disease. *Nephrol Dial Transplant* 2018; **33:** 1310-1316.

51. Elvira-Matelot E, Zhou XO, Farman N*, et al.* Regulation of WNK1 expression by miR-192 and aldosterone. *J Am Soc Nephrol* 2010; **21:** 1724-1731.

52. Baker MA, Wang F, Liu Y*, et al.* MiR-192-5p in the Kidney Protects Against the Development of Hypertension. *Hypertension* 2019; **73:** 399-406.

53. Enright AJ, John B, Gaul U*, et al.* MicroRNA targets in Drosophila. *Genome Biol* 2003; **5:** R1.

54. Ewing B, Hillier L, Wendl MC*, et al.* Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 1998; **8:** 175-185.

55. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 1998; **8:** 186-194.

56. Cock PJ, Fields CJ, Goto N*, et al.* The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res* 2010; **38:** 1767-1771.

Figure 1

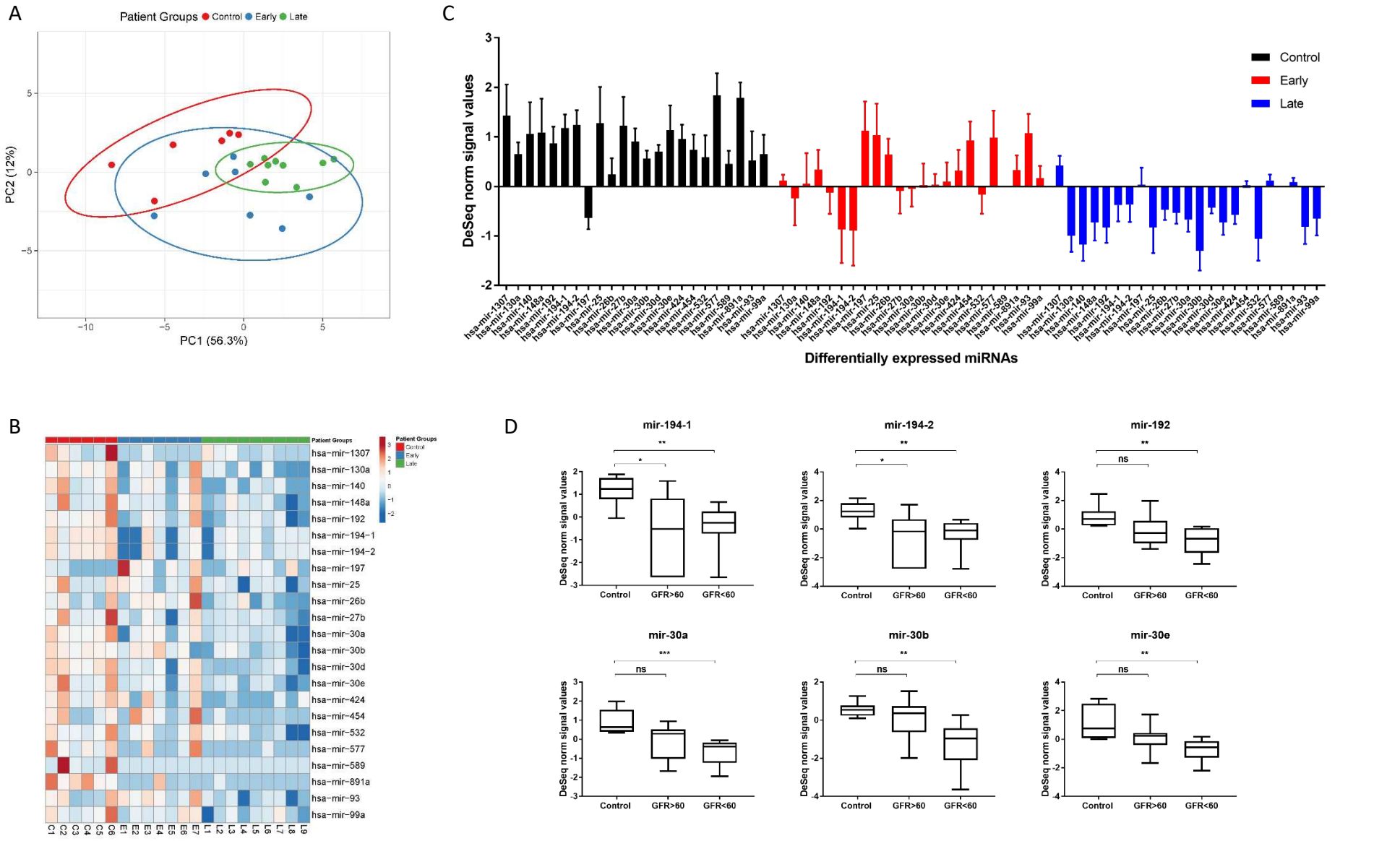

A
Patient Groups Control Early Late
PC2 (12%)
PC1 (56.3%)
C
DeSeq norm signal values
Control
Early
Late
Differentially expressed miRNAs
B
Patient Groups
hsa-mir-1307
hsa-mir-130a
hsa-mir-140
hsa-mir-148a
hsa-mir-192
hsa-mir-194-1
hsa-mir-194-2
hsa-mir-197
hsa-mir-25
hsa-mir-26b
hsa-mir-27b
hsa-mir-30a
hsa-mir-30b
hsa-mir-30d
hsa-mir-30e
hsa-mir-424
hsa-mir-454
hsa-mir-532
hsa-mir-577
hsa-mir-589
hsa-mir-891a
hsa-mir-93
hsa-mir-99a
C1 C2 C3 C4 C5 C6 E1 E2 E3 E4 E5 E6 E7 L1 L2 L3 L4 L5 L6 L7 L8 L9
D
mir-194-1
mir-194-2
mir-192
mir-30a
mir-30b
mir-30e
DeSeq norm signal values
Control
GFR>60
GFR<60
ns

Figure 2

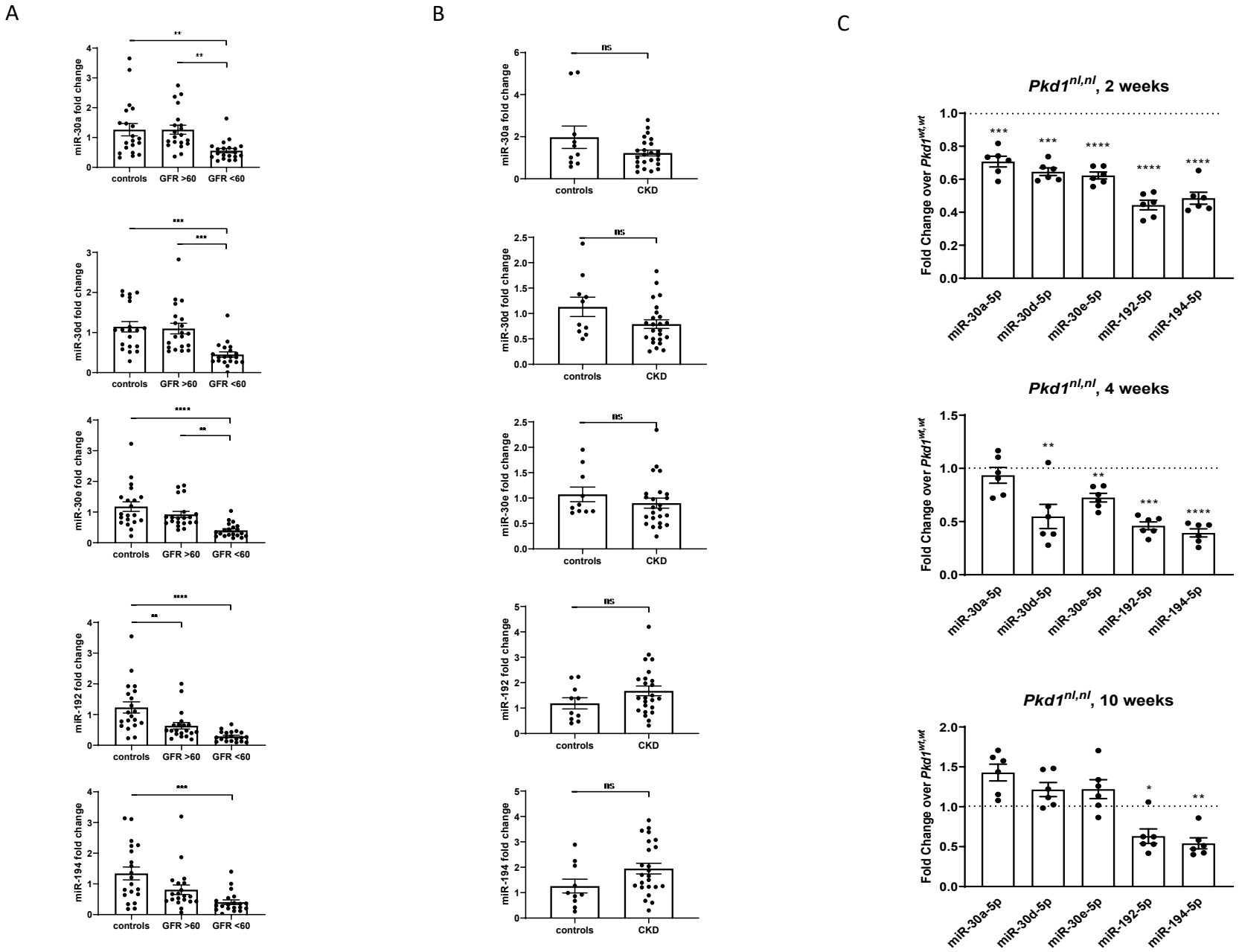
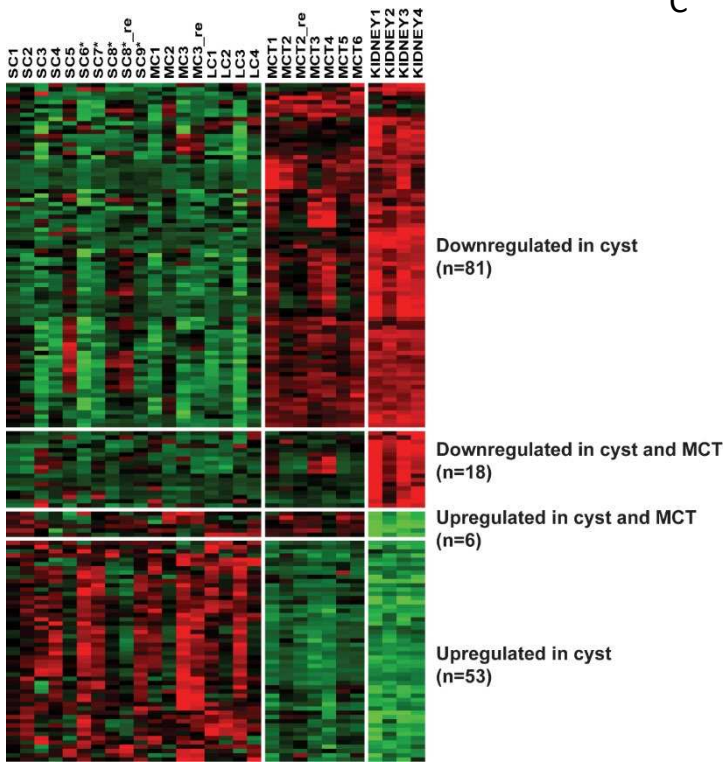
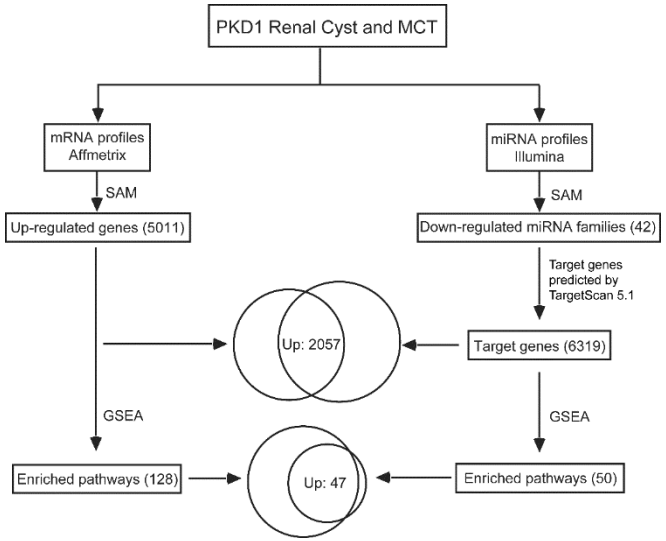

A
B
C
miR-30a fold change
miR-30d fold change
miR-30e fold change
miR-192 fold change
miR-194 fold change
controls
GFR >60
GFR <60
CKD
ns
Pkd1nl,nl, 2 weeks
Pkd1nl,nl, 4 weeks
Pkd1nl,nl, 10 weeks
Fold Change over Pkd1wt,wt
miR-30a-5p
miR-30d-5p
miR-30e-5p
miR-192-5p
miR-194-5p

Figure 3

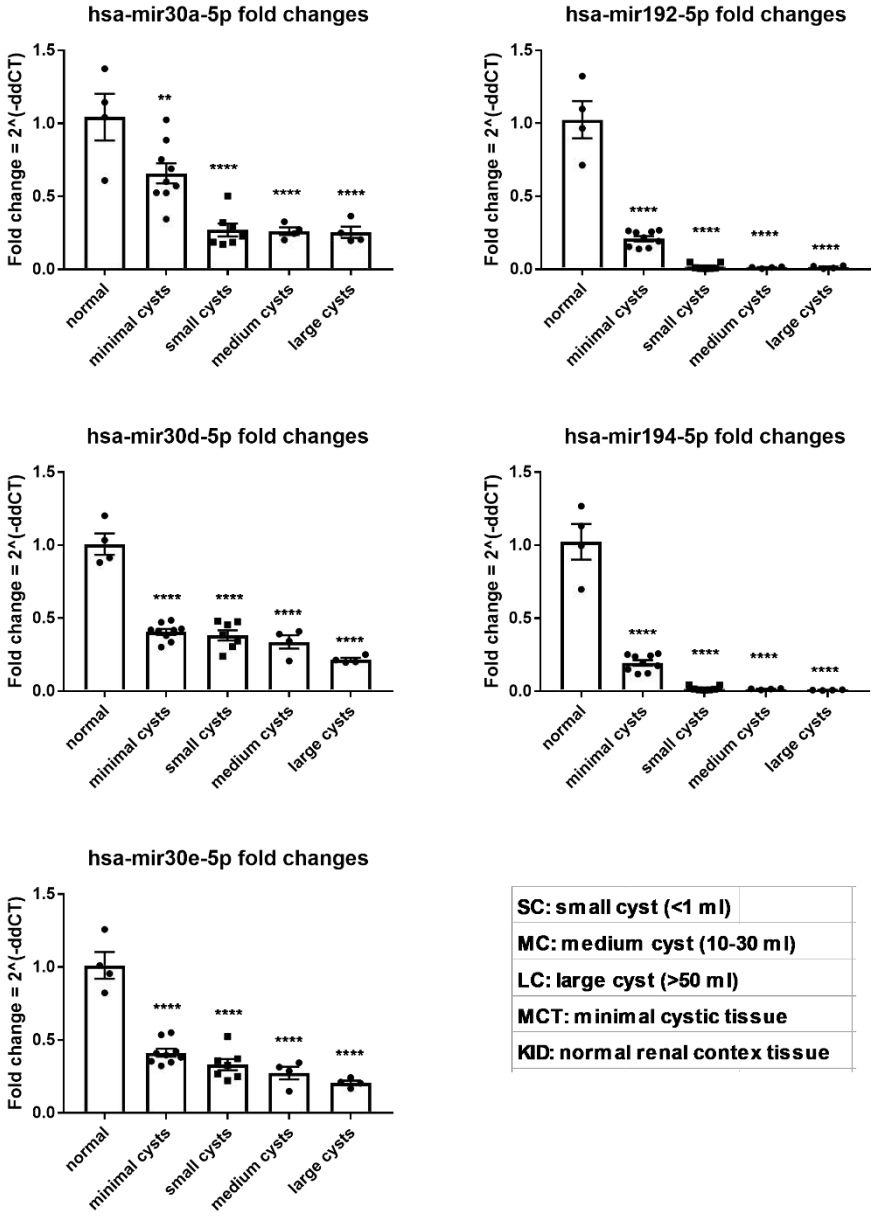
A



B



C



SC: small cyst (<1 ml)
MC: medium cyst (10-30 ml)
LC: large cyst (>50 ml)
MCT: minimal cystic tissue
KID: normal renal contex tissue

Figure 4

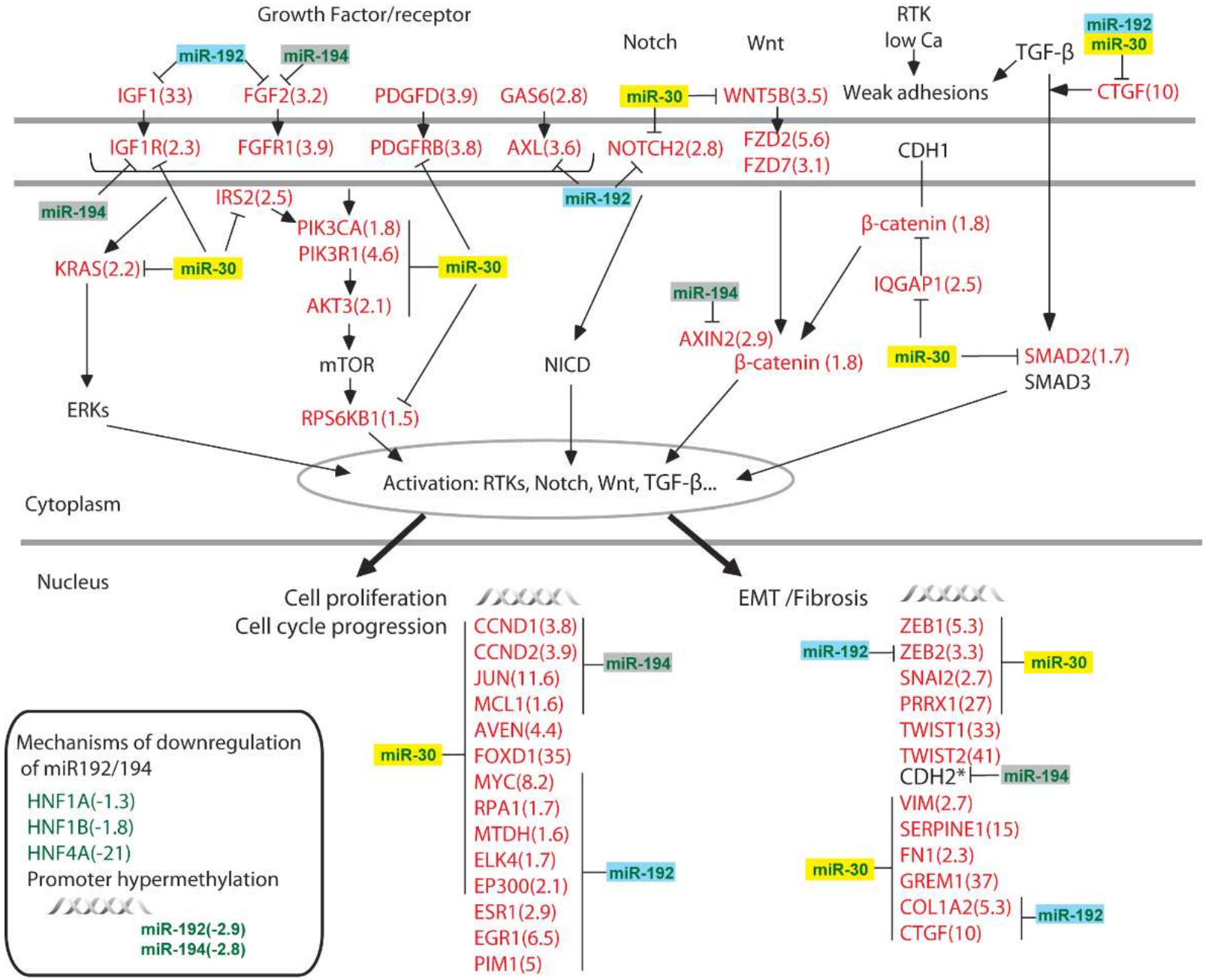
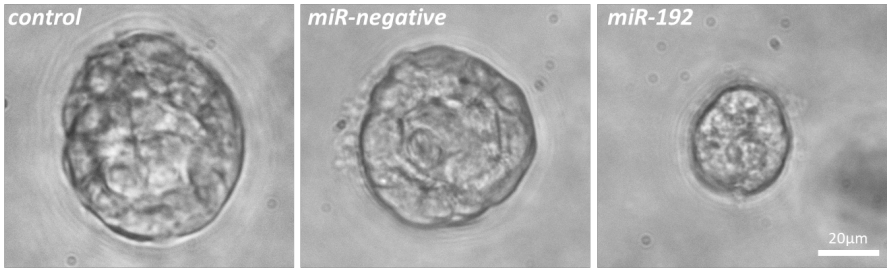
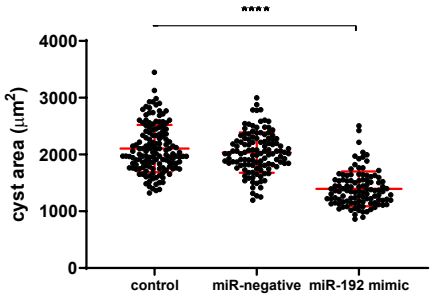
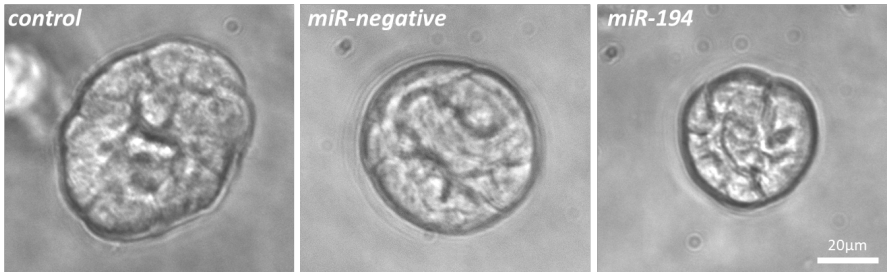

Growth Factor/receptor
Notch
Wnt
RTK
low Ca
TGF-β
miR-192
miR-30
miR-194
IGF1(33)
FGF2(3.2)
PDGFD(3.9)
GAS6(2.8)
WNT5B(3.5)
Weak adhesions
CTGF(10)
IGF1R(2.3)
FGFR1(3.9)
PDGFRB(3.8)
AXL(3.6)
NOTCH2(2.8)
FZD2(5.6)
FZD7(3.1)
CDH1
IRS2(2.5)
PIK3CA(1.8)
PIK3R1(4.6)
β-catenin (1.8)
KRAS(2.2)
AKT3(2.1)
IQGAP1(2.5)
AXIN2(2.9)
mTOR
NICD
β-catenin (1.8)
SMAD2(1.7)
SMAD3
ERKs
RPS6KB1(1.5)
Activation: RTKs, Notch, Wnt, TGF-β...
Cytoplasm
Nucleus
Cell proliferation
Cell cycle progression
CCND1(3.8)
CCND2(3.9)
JUN(11.6)
MCL1(1.6)
AVEN(4.4)
FOXD1(35)
MYC(8.2)
RPA1(1.7)
MTDH(1.6)
ELK4(1.7)
EP300(2.1)
ESR1(2.9)
EGR1(6.5)
PIM1(5)
EMT /Fibrosis
ZEB1(5.3)
ZEB2(3.3)
SNAI2(2.7)
PRRX1(27)
TWIST1(33)
TWIST2(41)
CDH2*
VIM(2.7)
SERPINE1(15)
FN1(2.3)
GREM1(37)
COL1A2(5.3)
CTGF(10)
Mechanisms of downregulation of miR192/194
HNF1A(-1.3)
HNF1B(-1.8)
HNF4A(-21)
Promoter hypermethylation
miR-192(-2.9)
miR-194(-2.8)

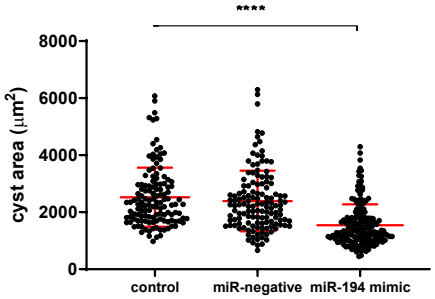
Figure 5

A

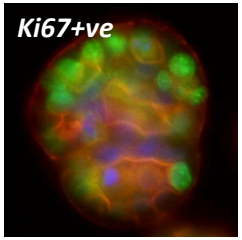

control
miR-negative
miR-192
20µm

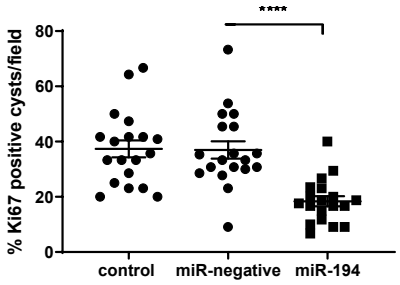

****
cyst area (µm²)
4000
3000
2000
1000
0
control
miR-negative
miR-192 mimic

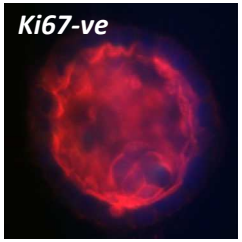

control
miR-negative
miR-194
20µm

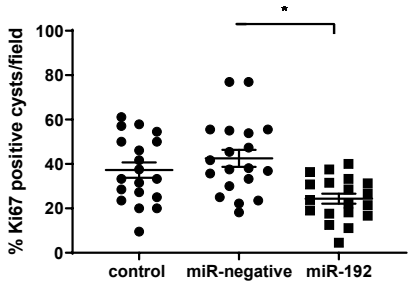

****
cyst area (µm²)
8000
6000
4000
2000
0
control
miR-negative
miR-194 mimic

B

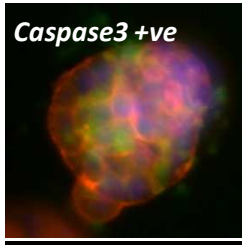

Ki67+ve

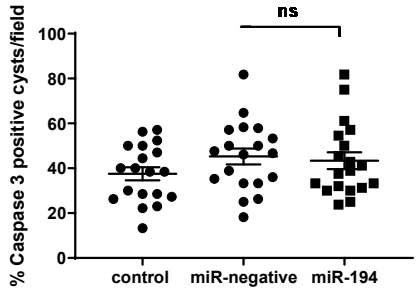

****
% Ki67 positive cysts/field
80
60
40
20
0
control
miR-negative
miR-194

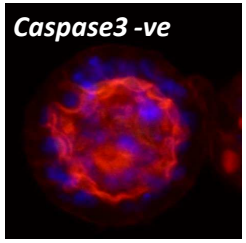

Ki67-ve


*
% Ki67 positive cysts/field
100
80
60
40
20
0
control
miR-negative
miR-192

C


Caspase3 +ve


ns
% Caspase 3 positive cysts/field
100
80
60
40
20
0
control
miR-negative
miR-194


Caspase3 -ve

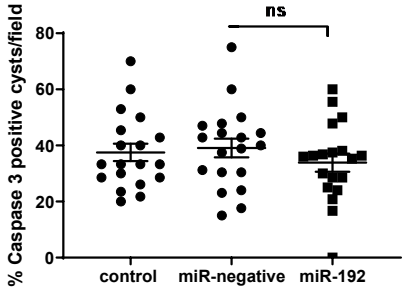

ns
% Caspase 3 positive cysts/field
80
60
40
20
0
control
miR-negative
miR-192

Figure 6

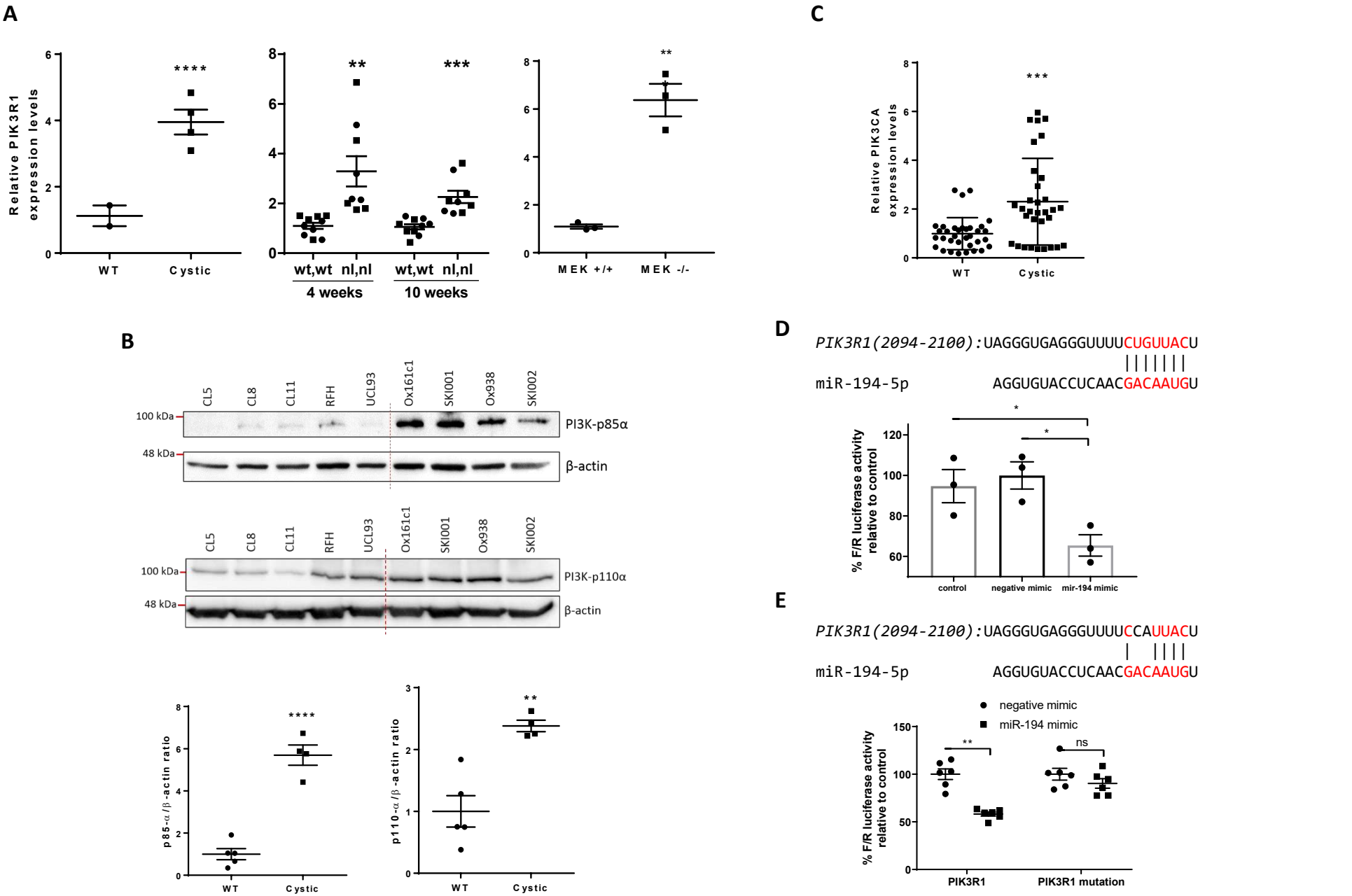
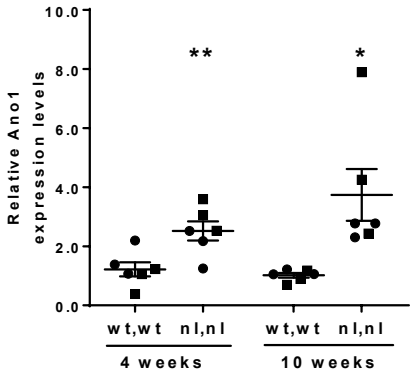
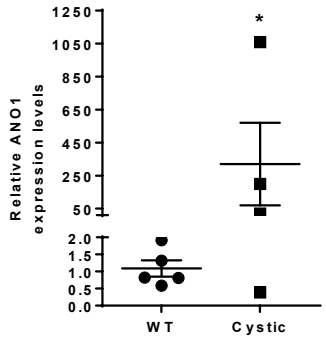

A
Relative PIK3R1 expression levels
****
WT
Cystic
**
***
wt,wt nl,nl
4 weeks
wt,wt nl,nl
10 weeks
**
MEK +/+
MEK -/-
B
CL5 CL8 CL11 RFH UCL93 Ox161c1 SKI001 Ox938 SKI002
100 kDa
PI3K-p85α
48 kDa
β-actin
CL5 CL8 CL11 RFH UCL93 Ox161c1 SKI001 Ox938 SKI002
100 kDa
PI3K-p110α
48 kDa
β-actin
p85-α/β-actin ratio
****
WT
Cystic
p110-α/β-actin ratio
**
WT
Cystic
C
Relative PIK3CA expression levels
***
WT
Cystic
D
PIK3R1(2094-2100):UAGGGUGAGGGUUUUCUGUUACU
miR-194-5p AGGUGUACCUCAACGACAAUGU
% F/R luciferase activity relative to control
*
*
control
negative mimic
mir-194 mimic
E
PIK3R1(2094-2100):UAGGGUGAGGGUUUUCCAUUACU
miR-194-5p AGGUGUACCUCAACGACAAUGU
negative mimic
miR-194 mimic
% F/R luciferase activity relative to control
**
ns
PIK3R1
PIK3R1 mutation

Figure 7

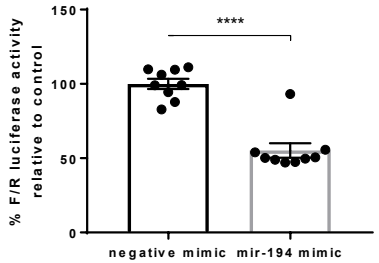
A



B

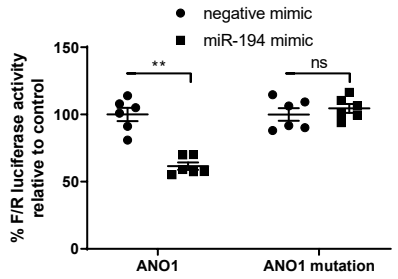


C



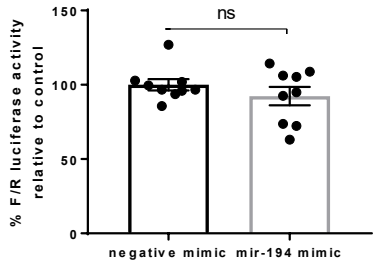
ANO1(993-999): AGGAAUUUAAAAUACUGUUACU

miR-194-5p AGGUGUACCUCAACGACAAUGU



ANO1(993-999): AGGAAUUUAAAAUACCAUUACU

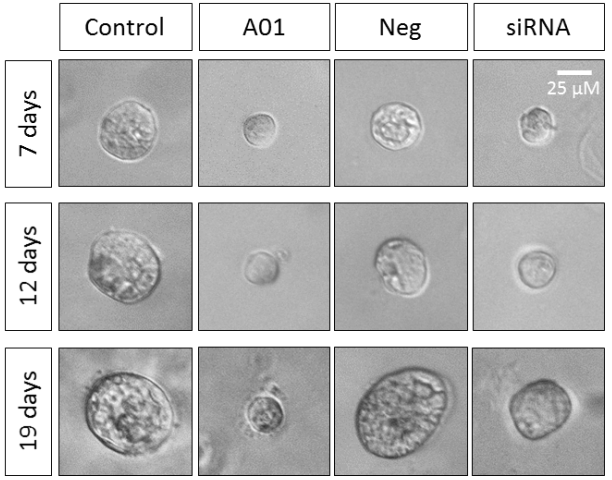
miR-194-5p AGGUGUACCUCAACGACAAUGU

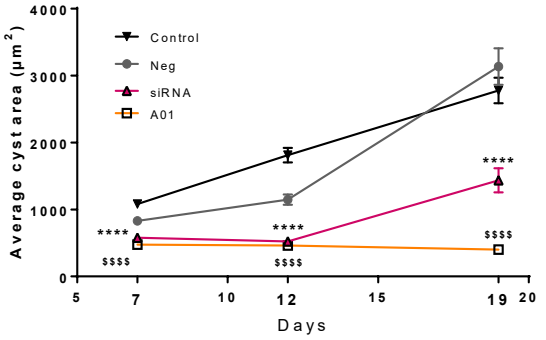


ANO1(1258-1264): AUCUCAUUGAUGUGAUGUUACAU

miR-194-5p AGGUGUACCUCAACGACAAUGU

D





Supplementary information (Magayr et al, KI)

Supplementary methods

Study Participants and Ethical Approval

Ethical approval for this study was obtained from the National Research Ethics Service Committee Yorkshire and The Humber Bradford (REC12/YH/0297) and from the REB of the Toronto General Hospital (02-0036). Ethical approval for the urine exosome samples from the CKD patient cohort with Type 2 diabetes was obtained from the National Research Ethics Service Committees – West London & GTAC (REC 04/Q0406/25), and London Bloomsbury (REC 17/LO.2083). All participants gave their signed informed consent at the time of recruitment or donation of their nephrectomized kidneys for the human microarray study. ADPKD patients were recruited through the Sheffield Kidney Institute and healthy normal volunteers were recruited from laboratory staff. Following collection, urine samples were centrifuged at 1,000 g for 10 min at 4°C. Cell pellets were discarded and cell free supernatants stored at −80°C until further analysis.

Library preparation and small RNA sequencing (miRNA-Seq)

The Illumina® TruSeq Small RNA Sample Preparation Protocol was used to generate the small RNA libraries. This protocol works well for studying miRNA because it takes advantage of the structure of most small RNA molecules by ligating specific adapters to the 5`-phosphate and 3`-hydroxyl group, which are molecular signatures of their biogenesis pathway. The 3`and 5`adapters were ligated to the RNA and used as templates for reverse transcription to cDNA. To amplify the obtained cDNA, ligated samples underwent 13 cycles of PCR. For the size selection of amplified cDNA libraries, PCR products were then

run on a 6% TBE gel with a custom ladder. The small RNA of approximately 140–160 bp in size was excised from the gel and incubated overnight. The incubated gel was eluted using a spin column. The obtained sequence libraries were then validated using an Agilent Bioanalyzer. Sequencing analysis was carried out using an Illumina HiSeq 2000 sequencing platform with 50 nt single reads (Illumina, USA).

Renal Cyst and Control Tissues

Renal cysts of different sizes were obtained from 4 polycystic kidneys removed for medical reasons. The surgical technique was described in detail in our previous article[S1]. Small cysts (SC) were defined as less than 1 mL, SC* between 1-5 mL, medium cysts (MC) between 10-25 mL, and large cysts (LC) greater than 50 mL. Minimally cystic tissue (MCT), which might have contained a few microscopic cysts from the renal cortex, was obtained as PKD control tissue from the same kidneys. Additionally, non-cancerous renal cortical tissue from four nephrectomized kidneys with isolated renal cell carcinoma was used as normal control tissue. For the microarray study, 16 cysts (SC, n=5; SC*, n=5 (including 1 duplicate); MC, n=4 (including 1 duplicate), and LC, n=4), 7 MCT (including 1 duplicate) and 4 normal renal cortical tissue samples were used. All the study patients were shown to have *PKD1* by DNA linkage or documentation of a pathogenic mutation identified through DNA sequencing by Athena Diagnostics™. Informed consent was obtained from all patients and the Institutional Review Board of the hospital where the nephrectomy was performed approved the research protocol used for this study.

Microarray Analysis of Cyst and Control Tissues

200 ng of total RNA were labelled and hybridized onto Illumina Human miRNA BeadChip V2 (12-array BEADChip format) according to the manufacturer's protocol at the Microarray and Gene Expression Facility of the Centre for Applied Genomics in the Hospital for Sick Children, Toronto. This array comprises 1145 assays, detecting more than 97% of the miRNAs described in the Sanger miRBase v12 (http://www.mirbase.org/) [S2]. The miRNA names used in this manuscript were based on miRBase v12; the revised miRNA names based on miRBase v22.1 have been provided in the GEO submission (GSE133530). We repeated some samples in duplicate for a total of 36 expression profiles. We also arranged each of these replicated in separate arrays to reduce potential batch effect. Before data processing, we checked that all hybridization passed internal QC metrics (including array hybridization controls, negative controls, polyadenylation controls, oligo annealing controls, mismatch controls, extension controls, and contamination detection controls) and kept the 29 arrays met the Illumina acceptance criteria.

QPCR validation of differentially expressed miRNA

miRNA expression in isolated exosomes purified from human urine samples and human tissue samples was determined by Taqman qPCR according to the manufacturer's protocol (Life Technologies). Following RNA extraction, cDNA was synthesized from ADPKD, normal cells, tissue samples and exosomes using specific miRNA RT primers. QPCR was carried out on an ABI7900 qPCR machine. NormFinder was used to identify suitable universally expressed microRNAs in or RNAseq experiment which assesses the stability of differentially expressed microRNA based on their expression variability and stability

values. Based on this, miR-191-5p was the most universally expressed microRNAs among the samples with a stability value of 94.714. Normalisation of expression was therefore carried out relative to urinary expression of miR-191-5p (an endogenous control) and cel-miR-39-3p (an exogenous spike in control).

Analysis of miRNA expression in mouse kidney

To purify total RNA, kidneys were mechanically homogenized in 1 ml of Qiazol® Lysis Reagent (Qiagen) using pre-filled Triple-pure zirconium beads (Benchmark Scientific) in a benchtop microtube homogenizer (Benchmark Scientific) at 4 °C. Purification of total RNA were processed using the miRNeasy kit (Qiagen) according to the manufacturer's instructions. Residual DNA fragments were digested after elution of total RNA performing a DNase treatment (Invitrogen). Total RNA was quantified on a NanoDrop™ 2000/2000c Spectrophotometer (Thermo Fisher Scientific) and 250 ng RNA were retrotranscribed using TaqMan™ MicroRNA Reverse Transcription Kit (Applied Biosystem) and amplified using TaqMan qPCR master mix No Amp UNG (Applied Biosystem) and the following specific primers to each miRNA: miR-30a-5P (000417), miR-30d-5p (000420, miR-30e-5p (002223), miR-192-5p (000491) and miR-194-5p (000493). Results were expressed as fold change over $Pkd1^{wt,wt}$ at the same age following the $2^{-\Delta\Delta C_T}$ Livak Method and using U6 snRNA (001973) as a reference gene.

Western blotting

Total cell lysates were prepared and processed for Western blotting as described [S3]. Cells were solubilized in detergent lysis buffer (50 mM Tris, 0.14 M NaCl, 1% Triton X-100, and 0.5% NP40) supplemented with Complete protease inhibitors and PhosStop phosphatase inhibitors (Roche Diagnostics, Mannheim, Germany). ECL detection and quantification using a Biorad Chemidoc XRS+ system running Image Lab automated image capture and analysis software. All quantification was carried out on non-saturated bands as determined by the software from 3 independent experiments.

Matrigel 3D cyst assays

3D matrigel cyst assays were performed as previously described [S4]. In brief, OX161/C1 cells ($1x10^5$/well) were mixed with 70 µl Matrigel (Becton Dickinson, UK), plated into 96 well plates in triplicate and incubated for 30 min at 37°C to facilitate gel formation. Media was replaced every 2 d. The average cyst area was calculated by measuring cyst areas in individual wells on day 12. At least 130 cysts were measured in triplicate wells at each time-point.

Immunofluorescence staining

Immunofluorescence staining of cysts was performed as previously described [S4] . Antibodies to Ki67 and cleaved-caspase 3 (Cell Signalling, USA) were used at 1/100 dilution. Alexa Fluor 488 or 594 secondary antibodies (Invitrogen, USA) were used at 1/1000 dilution. Slides were viewed using an Imaging Systems inverted IX71 microscope (Olympus, Tokyo, Japan) configured for multifluorescence image capture. Images were

acquired using SimplePCI imaging software (Compix, Hamamatsu, PA). At least 300 cysts from 19 random fields were quantified and expressed as a percentage positive cysts per field.

Luciferase reporter assays

Two predicted miR-194-5p sites in *ANO1* were identified by TargetScan between bp 993-999 and 1258-1264 in the 3'UTR of *ANO1*. A predicted miR-194-5p site in *PIK3R1* was identified by TargetScan between bp 2094-2100 (seed sequences shown in red). The *ANO1* 3'UTR containing each sequence and the *PIK3R1* 3'UTR was cloned into a pmirGLO luciferase reporter vector and transfected into HEK cells with either scrambled control miRNA or miR-194-5p mimics (50nM). Changes in luciferase activity were measured as previously described [S4]. Site directed mutagenesis of seed sequences was carried out as described previously [S4] .

***In silico* prediction of miRNA-target pairs**

The miRNA-target pairs for seven mature miRNAs (i.e. miR-192-5p, miR-194-5p, miR-30a-5p, miR-30b-5p, miR-30c-5p, miR-30d-5p and miR-30e-5p) were obtained using an online database – DIANA-TarBase v8 (http://www.microrna.gr/tarbase), which constitutes an extensive repository, integrating ~670,000 unique experimentally supported miRNA–gene interactions from low-/high-throughput experiments [S5].

Microarray Data Analysis

Data Preprocessing and Normalization

Scanned raw data images were processed with BeadStudio 3.4 (Illumina), a software package that performs visualization and normalization. We used the "Quantile" normalization method, as previously suggested for Illumina miRNA microarray experiments. The normalized intensities and detection p-values were exported and further adjusted for batch effect with the R package ComBat, using the non-parametric empirical Bayes method [S6]; ComBat successfully eliminated batch effects as evidenced by hierarchical clustering and improved Pearson correlation coefficient for the technical replicates. After normalization, the samples with replicates were averaged prior to subsequent analysis. As some miRNAs were either not expressed or non-detectable, we filtered out probes with median detection p value > 0.0001. This procedure reduced the number of miRNA probes from 1146 to 829 with detectable levels. Additionally, previous work suggests that only the most abundant microRNAs in a cell mediate target suppression [S7], we gave higher priority to higher expressed miRNAs. According to the average expression level of miRNAs in any of 3 renal groups, we assigned each miRNA into highly expressed (H): signal intensity between 10000 to 40000 (156 out of 829), moderately expressed (M): between 1000 to 10000 (328 out of 829), and low expressed (L): <1000 (345 out of 829). Microarray data is MIAME compliant and available in Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) (ID: GSE133530).

Hierarchical Cluster Analysis (HCA)

The detail procedure for HCA has been described in our previous article [S1]. In brief, we performed HCA based on miRNA expression from the top 200, or 500 most variable miRNAs (with the largest CV) across all samples to define the similarity of miRNA expression patterns between renal cysts of varying size, MCT and normal renal cortical tissue. HCA was performed to assess the relatedness of each case to every other case using dCHIP [S8]. The correlation coefficient matrix containing all pairwise correlation coefficients was created by dChip for heat-map visualization. The range of observed correlation coefficients was -0.8 to 0.8. Positive correlation is indicated by red and negative correlation by green. The degree of correlation is represented by color intensity.

Statistical Analysis of Individual Differentially Expressed miRNAs

Significance Analysis of Microarrays (SAM) was used to provide conventional measures of statistical significance for the individual miRNAs that are differentially expressed between classes [S9]. The SAM software works as a Microsoft Excel add-in and is available via http://www-stat.stanford.edu/~tibs/SAM/. We used a false discovery rate (FDR) $\leq$ 0.1% for these individual miRNA comparisons. To reduce the noise, we only focused on miRNAs with medium to high average expression.

Bioinformatic Analysis

In silico sequence analysis predict the presence of hundreds of thousands of miRNA-target pairs, only a small number of which have been experimentally validated, which usually

have the evidence of reciprocal expression between the miRNAs and the targets. We used TargetScan 5.1 (http://www.targetscan.org/) [S10] for prediction of potential miRNA target genes, as it was shown to be superior to other target predictors [S11]. By TargetScan, miRNAs that belong to the same family (a set of miRNAs sharing the same seed sequence, nt 2-8) were grouped together and then each family was then further assigned into 3 TargetScan families: broadly conserved, conserved or poorly conserved family, according to the evolutionary conservation among species. For each predicted target gene of highly conserved miRNA families, TargetScan calculated an Aggregate P_{CT} (the probability of conserved targeting) score, which reflects the probability of preferential conservation and has the advantage of identifying targeting interactions that are not only more likely to be effective, but also those that are more likely to be consequential for the animal [S10].

We defined a genomic cluster of miRNAs as a maximal segment such that every two consecutive miRNAs were separated by <50 kb. Genomic positions of miRNAs were taken from miRBase [S2].

Correlation analysis of the miRNA and mRNA expression profiles was performed by using the lists of differentially expressed miRNAs and mRNAs. The predicted mRNA targets of each of the dysregulated miRNAs were obtained using TargetScan 5.1. For a novel miRNA, only the predicted genes with conserved sites on the 3' UTR were considered as targets. A two-tailed Fisher's Exact Test was performed to determine whether the number of target mRNAs that changed expression was greater than would be expected by chance. Fisher's Exact Test computed 107 unadjusted p-values from the dysregulated miRNAs that had predicted targets; all p-values were adjusted based on the Bonferroni method

accounting for multiple testing. P values less than 4.67E-04 (0.05 divided by 107) were considered to be significant.

Pathway analysis based on the predicted target genes from all 42 down-regulated miRNA families, which had significantly inverse correlation with the expression of their target genes, was conducted using gene set enrichment analysis (GSEA) (http://www.broad.mit.edu/gsea/) [S12]. For purpose of comparison, we tested the same set of pathways as we used in our previous study [S1] and defined overrepresented pathways by a NOM p-value ≤ 0.05 with a FDR ≤ 0.25.

To test individual miRNA from 42 downregulated miRNA families possibly targeting multiple enriched pathways or multiple components of a pathway, we used a web-based tool – GENECODIS 2.0, which uses the hypergeometric distribution to determine whether individual pathway or combinations of pathways are significantly associated to a list of genes under study with respect to a reference list (http://genecodis.genyo.es/) [S13]. The *p* values computed for each pathway were adjusted using the FDR method of Hochberg and Benjamini for multiple hypothesis testing, and corrected *p* values < 0.05 ($-\log_2$(p-value) >4.32) were considered significant.

The experimentally validated miRNA-target pairs were obtained using the online DIANA-TarBase v8 database [S5].

Supplementary figures

Figure S1 Linear correlations between MKL and TKV or ht-TKV in validation cohort (A) where both measurements were available (n=19). (B) ROC analysis for MKL in predicting baseline eGFR (AUC 0.905) or discriminating eGFR slope (AUC 0.634).

Figure S2 miRNA expression profiles of normal kidneys, MCTs and renal cysts based on the standardized signal intensities of (A) the top 500 most variable miRNAs and (B) 158 dysregulated miRNAs, derive a matrix of Pearson correlation coefficients that indicate the degree of overall similarity between any two samples. Positive correlation is indicated by red and negative correlation by green, with deeper colors indicating higher positive (red) or negative (green) correlations. Abbreviations: SC (small cysts); MC (medium cysts); LC (large cysts); MCT (minimally cystic tissue); KIDNEY (normal renal cortical tissue), and re (technical replicates).

Figure S3 miRNA expression correlates with eGFR and mean kidney length. Significant correlations as determined by Spearman correlation coefficient calculations were observed between miRNA expression (dCT values) and (A) mean kidney length and (B) eGFR.

Figure S4 ROC analysis of differentially expressed individual microRNAs to predict rapid disease progression in ADPKD.

ROC curves for the differentially expressed individual microRNAs to discriminate between patients with slow or rapid disease progression (annualized eGFR of < or > 3ml/min/year over 5 years). A perfect diagnostic test has AUC of 1 while a value of 0.5 indicates weak prediction ability.

Figure S5 Knockdown of *PIK3R1* or *ANO1* confirmed by QPCR

(A) Knockdown of *PIK3R1* or *ANO1* in the human cystic lines, OX161c1 or SKI-001 respectively, measured by QPCR; (B) Uncropped blots of PI3K-p85a and PI3K-p110a following IGF-1 stimulation (see Figure 5B).

Figure S6 Knockdown of *PIK3R1* or chemical inhibition of PI3K inhibits IGF-1 stimulated cyst growth

(A) OX161c1 cells were transfected with a negative control siRNA or *PIK3R1* siRNA before being plated in 3D Matrigel and cultured in the presence or absence of IGF-1 (200ng/ml) for 7 days. Knockdown of *PIK3R1* abolished the IGF-1 response in transfected cells. Representative pictures of cysts for each condition indicate changes in cyst area after transfection. (B) OX161c1 cells grown in 3D Matrigel were treated with IGF-1 (200 ng/mL for 7 days). Control cells showed an increase in cyst area compared to untreated control cells (158% at 7 days). The PI3K inhibitor LY294002 at 20 μM or 50 μM completely inhibited the response to IGF-1 treatment with significantly reduced cyst area (56% at day 7) compared to the untreated control condition. Scale bar: 20 μm. Unpaired t-test and SEM, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Supplementary Table 1. Differentially expressed miRNAs in human PKD1 renal cysts identified by SAM (FDR ≤ 0.1%)

	Dysregulated miRNAs in cysts (n=158)	Rank by SAM score	Fold changes cyst/MCT	Fold changes cyst/kid	Expression levels #	miRNA families (n=105)	Seed + m8 sequence	No. of predicted targets##	No. of predicted targets on arrays	No. of predicted targets inversely altered in cysts	Inverse correlation p value^
Class 1↑ n=19	hsa-let-7b↑	13	1.2	1.4	H	let-7/98	GAGGUAG	819	784	97↓	2.0E-01
	hsa-let-7c↑	25	1.6	1.9	H						
	hsa-let-7d↑	28	1.3	1.5	H						
	hsa-let-7e↑	81	1.2	1.3	H						
	hsa-let-7f↑	84	1.0	1.1	H						
	hsa-let-7i↑	12	1.1	1.5	H						
	hsa-miR-23a↑	38	1.2	1.5	H	miR-23ab	UCACAUU	838	801	108↓	2.4E-02
	hsa-miR-34a↑	85	1.1	1.4	H	miR-34a/34b-5p/34c/34c-5p/449/449abc/699	GGCAGUG	469	454	58↓	2.2E-01
	hsa-miR-100↑	59	1.2	1.4	H	miR-99ab/100	ACCCGUA	40	39	6↓	4.4E-01
	hsa-miR-125a-5p↑	93	1.1	1.1	H	miR-125/351	CCCUGAG	604	573	83↓	8.1E-03
	hsa-miR-125b↑	3	1.0	1.3	H						
	hsa-miR-132↑	81	1.2	1.8	H	miR-132/212	AACAGUC	283	268	33↓	4.9E-01
	hsa-miR-146a↑	68	0.9	1.3	H	miR-146	GAGAACU	130	123	13↓	1.0E+00
	hsa-miR-152↑	16	1.4	1.5	H	miR-148/152	CAGUGCA	536	507	60↓	5.2E-01
	hsa-miR-155↑	97	1.1	2.2	H	miR-155	UAAUGCU	281	266	31↓	6.9E-01
	hsa-miR-196a	33	1.4	1.5	H	miR-196ab	AGGUAGU	211	199	27↓	2.5E-01
	hsa-miR-214↑	27	1.2	1.7	H	miR-214/761	CAGCAGG	407	389	54↓	7.1E-02

	hsa-miR-221↑	21	1.3	1.4	H	miR-221/222	GCUACAU	307	293	34↓	7.1E-01
	hsa-miR-222↑	104	1.3	1.6	H						
Class 1↓ n=42	hsa-miR-9↓	17	-4.2	-4.4	M	miR-9	CUUUGGU	936	895	316↑	**1.6E-20**
	hsa-miR-17↓	77	-1.1	-1.3	H	miR-17-5p/20/93.mr/106/519.d	AAAGUGC	990	935	337↑	**1.5E-23**
	hsa-miR-20a↓	98	-1.2	-1.2	H						
	hsa-miR-20b↓	58	-1.7	-2.5	H						
	hsa-miR-106a↓	91	-1.1	-1.3	H						
	hsa-miR-106b↓	73	-1.4	-1.6	H						
	hsa-miR-18a↓	44	-2.1	-2.7	M	miR-18ab	AAGGUGC	189	178	70↑	**2.3E-07**
	hsa-miR-18b↓	22	-2.0	-5.1	M						
	hsa-miR-19a↓	66	-1.9	-4.9	M	miR-19	GUGCAAA	938	879	312↑	**1.1E-20**
	hsa-miR-19b↓	104	-1.2	-2.4	H						
	hsa-miR-30a↓	57	-1.4	-1.4	H	miR-30a/30a-5p/30b/30b-5p/30cde/384-5p	GUAAACA	1080	1042	382↑	**2.7E-28**
	hsa-miR-30b↓	36	-1.3	-1.4	H						
	hsa-miR-30c↓	42	-1.2	-1.2	H						
	hsa-miR-30d↓	88	-1.1	-1.2	H						
	hsa-miR-30e↓	26	-1.4	-1.9	H						
	hsa-miR-101↓	76	-1.3	-1.8	H	miR-101	ACAGUAC	607	578	199↑	**4.2E-12**
	hsa-miR-107↓	23	-1.5	-4.2	H	miR-103/107	GCAGCAU	471	448	147↑	**4.5E-19**
	hsa-miR-124↓	15	-1.8	-13.4	M	miR-124/506	AAGGCAC	1299	1253	436↑	**1.2E-26**
	hsa-miR-506↓	72	-4.7	-4.4	M						
	hsa-miR-135a↓	19	-1.9	-2.8	H	miR-135	AUGGCUU	510	490	184↑	**3.2E-15**
	hsa-miR-135b↓	95	-2.6	-1.4	H						

	hsa-miR-138↓	10	-2.1	-5.5	M	miR-138	GCUGGUG	388	367	134↑	**2.0E-10**
	hsa-miR-140-5p↓	96	-1.5	-1.7	H	miR-140/140-5p/876-3p	AGUGGUU	251	242	93↑	**7.3E-09**
	hsa-miR-141↓	46	-2.4	-3.5	H	miR-141/200a	AACACUG	531	504	198↑	**8.8E-19**
	hsa-miR-200a↓	92	-1.8	-2.0	H						
	hsa-miR-142-3p↓	52	-2.2	-1.8	H	miR-142-3p	GUAGUGU	250	242	92↑	**1.9E-08**
	hsa-miR-144↓	91	-0.7	-6.6	M	miR-144	ACAGUAU	647	619	225↑	**2.0E-16**
	hsa-miR-187↓	2	-2.6	-8.3	M	miR-187	CGUGUCU	4	2	1↑	3.9E-01
	hsa-miR-190↓	5	-2.5	-7.2	M	miR-190	GAUAUGU	117	112	40↑	8.5E-04
	hsa-miR-192↓	4	-2.9	-4.7	H	miR-192/215	UGACCUA	89	82	27↑	2.3E-02
	hsa-miR-215↓	8	-3.8	-9.1	M						
	hsa-miR-194↓	6	-2.8	-3.6	H	miR-194	GUAACAG	258	248	89↑	**6.1E-07**
	hsa-miR-204↓	35	-1.9	-2.8	H	miR-204/211	UCCCUUU	461	435	148↑	**7.6E-09**
	hsa-miR-206↓	99	-1.4	-2.4	M	miR-1/206	GGAAUGU	584	557	197↑	**3.4E-13**
	hsa-miR-216a↓	72	-0.7	-2.3	M	miR-216/216a	AAUCUCA	161	153	37↑	4.9E-01
	hsa-miR-217↓	55	-0.9	-5.6	M	miR-217	ACUGCAU	237	224	74↑	**1.3E-04**
	hsa-miR-223↓	56	-1.5	-1.0	H	miR-223	GUCAGUU	202	189	70↑	**2.4E-06**
	hsa-miR-301a↓	36	-2.4	-4.5	H	miR-130/301	AGUGCAA	724	684	229↑	**2.6E-12**
	hsa-miR-338-3p↓	23	-2.1	-3.3	H	miR-338/338-3p	CCAGCAU	198	191	69↑	**9.2E-06**
	hsa-miR-455-5p↓	37	-2.1	-3.4	M	miR-455/455-5p	AUGUGCC	135	129	48↑	**1.0E-04**
	hsa-miR-503↓	87	-1.4	-2.4	M	miR-503	AGCAGCG	279	267	112↑	**2.8E-13**
	hsa-miR-551b↓	34	-1.9	-6.4	M	miR-551ab	CGACCCA	5	4	2↑	2.1E-01
Class 1	hsa-miR-15a↓	102	-1.2	-1.2	H		AGCAGCA	968	924	325↑	**8.8E-21**

other n=6	hsa-miR-424↓	21	-1.3	-1.5	H	miR-15/16/195/424/497					
	hsa-miR-15b↑	94	1.2	1.2	H			968	924	117↓	9.5E-02
	hsa-miR-195↑	71	1.0	1.3	H						
	hsa-miR-25↑	64	1.3	1.4	H	miR-25/32/92/92ab/363/367	AUUGCAC	692	661	66↓	4.5E-01
	hsa-miR-363↓	89	-1.5	-1.9	H			692	661	246↑	**2.4E-19**
Class 2↑ n=8	hsa-miR-28-5p↑	106	1.1	1.2	H	miR-28/28-5p/708	AGGAGCU	111	100	17↓	7.5E-02
	hsa-miR-134↑	70	1.6	3.6	M	miR-134	GUGACUG	96	92	7↓	4.0E-01
	hsa-miR-197↑	90	1.2	1.6	H	miR-197	UCACCAC	140	134	18↓	3.3E-01
	hsa-miR-224↑	82	2.5	9.9	M	miR-224	AAGUCAC	236	221	18↓	1.9E-01
	hsa-miR-320abcd↑	50	1.4	1.7	H	miR-320/320abcd	AAAGCUG	539	510	52↓	6.2E-01
	hsa-miR-342-3p↑	69	1.0	1.5	H	miR-342/342-3p	CUCACAC	169	156	21↓	3.0E-01
	hsa-miR-361-5p↑	47	1.1	1.2	H	miR-361/361-5p	UAUCAGA	138	130	19↓	2.0E-01
	hsa-miR-379↑	79	1.5	2.9	M	miR-379	GGUAGAC	62	58	5↓	8.3E-01
Class 2↓ n=7	hsa-miR-340↓	62	-2.1	-1.7	M	miR-340/340-5p	UAUAAAG	963	930	329↑	**2.3E-21**
	hsa-miR-362-3p↓	43	-2.1	-2.4	H	miR-329/362-3p	ACACACC	221	209	84↑	**3.6E-09**
	hsa-miR-378↓	40	-1.1	-2.0	H	miR-378/422a	CUGGACU	106	100	36↑	1.5E-03
	hsa-miR-450a↓	57	-1.8	-2.1	H	miR-450a/450a-5p	UUUGCGA	3	3	0↑	1.0E+00
	hsa-miR-542-3p↓	14	-2.6	-3.0	H	miR-542/542-3p	GUGACAG	156	153	46↑	1.9E-02
	hsa-miR-708↓	24	-1.6	-1.9	M	miR-28/28-5p/708	AGGAGCU	111	100	27↑	2.3E-01
	hsa-miR-874↓	63	-1.4	-2.2	M	miR-874	UGCCCUG	152	145	34↑	6.9E-01
Class 3↑ n=19	hsa-miR-28-3p↑	105	1.1	1.3	H	miR-28/28-3p	ACUAGAU	77	72	8↓	1.0E+00
	hsa-miR-151-3p↑	39	1.2	1.4	H	miR-151-3p	UAGACUG	76	71	9↓	5.7E-01
	hsa-miR-151-5p↑	65	1.1	1.2	H	miR-151	CGAGGAG	6	6	2↓	1.3E-01

	hsa-miR-193a-5p↑	37	1.7	2.9	H	miR-193a/193a-5p	GGGUCUU	90	82	11↓	4.8E-01
	hsa-miR-199a-3p↑ hsa-miR-199b-3p↑	18	1.1	1.3	H	miR-199/199-3p	CAGUAGU	294	284	39↓	1.5E-01
	hsa-miR-342-5p↑	84	1.3	2.9	M	miR-342-5p	GGGGUGC	63	60	5↓	6.8E-01
	hsa-miR-432↑	83	2.8	5.6	M	miR-432	CUUGGAG	133	126	19↓	1.5E-01
	hsa-miR-452↑	74	1.0	4.0	M	miR-452.h	ACUGUUU	205	199	19↓	5.7E-01
	hsa-miR-548d-5p↑	100	2.0	1.8	M	miR-548a-5p/548b-5p/548c-5p/548d-5p/548hij/559	AAAGUAA	443	431	42↓	4.4E-01
	hsa-miR-574-3p↑	54	1.5	1.7	H	miR-574/574-3p	ACGCUCA	6	6	0↓	1.0E+00
	hsa-miR-574-5p↑	84	3.2	5.1	M	miR-574-5p	GAGUGUG	89	87	12↓	3.9E-01
	hsa-miR-576-5p↑	80	1.9	2.6	M	miR-576-5p	UUCUAAU	392	364	55↓	1.4E-02
	hsa-miR-625↑	101	1.4	2.3	M	miR-625	GGGGGAA	70	67	7↓	1.0E+00
	hsa-miR-628-3p↑	53	1.8	1.8	M	miR-628-3p	CUAGUAA	71	67	7↓	1.0E+00
	hsa-miR-628-5p↑	99	2.5	2.4	M	miR-628-5p	UGCUGAC	157	152	18↓	7.0E-01
	hsa-miR-664↑	98	1.1	1.3	H	miR-664.hr	AUUCAUU	406	386	55↓	4.8E-02
	hsa-miR-923↑	91	4.9	5.5	H	miR-923	UCAGCGG	3	3	0↓	1.0E+00
	hsa-miR-1246↑	103	1.5	2.2	H	miR-1246	AUGGAUU	114	106	7↓	2.1E-01
	hsa-miR-1299↑	104	2.4	1.8	M	miR-1299	UCUGGAA	389	372	40↓	1.0E+00
Class 3↓ n=29	hsa-miR-129-3p↓	29	-4.8	-2.3	M	miR-129-3p	AGCCCUU	286	274	94↑	**3.1E-06**
	hsa-miR-142-5p↓	70	-3.4	-4.0	M	miR-142-5p	AUAAAGU	537	520	191↑	**1.3E-14**
	hsa-miR-188-5p↓	9	-1.8	-3.8	M	miR-188/188-5p	AUCCCUU	152	141	42↑	3.2E-02
	hsa-miR-337-5p↓	80	-2.3	-2.8	M	miR-337-5p.h	AACGGCU	13	13	3↑	1.0E+00
	hsa-miR-339-3p↓	48	-1.7	-2.3	M	miR-339-3p	GAGCGCC	11	11	5↑	7.3E-02

hsa-miR-362-5p↓	28	-2.1	-3.4	M	miR-362-5p	AUCCUUG	93	89	31↑	6.6E-03
hsa-miR-450b-5p↓	47	-2.0	-2.1	M	miR-450b/450b-5p	UUUGCAA	510	484	195↑	**8.2E-20**
hsa-miR-500↓	32	-1.8	-2.1	H	miR-500.h	AAUCCUU	179	167	65↑	**7.5E-07**
hsa-miR-501-3p↓	73	-1.5	-1.7	M	miR-500/501-3p/502/502-3p	AUGCACC	151	145	57↑	**2.6E-06**
hsa-miR-501-5p↓	74	-1.9	-2.2	M	miR-501/501-5p	AUCCUUU	150	145	56↑	**6.9E-06**
hsa-miR-507↓	41	-1.2	-3.0	M	miR-507/557	UUUGCAC	343	323	127↑	**2.2E-12**
hsa-miR-508-3p↓	43	-5.5	-5.2	M	miR-508-3p	GAUUGUA	218	210	75↑	**5.2E-06**
hsa-miR-509-3-5p↓	43	-12.4	-12.7	M	miR-509-5p/509-3-5p	ACUGCAG	265	247	88↑	**9.1E-07**
hsa-miR-509-3p↓	31	-9.1	-7.2	M	miR-509-3p.h	GAUUGGU	87	82	24↑	1.4E-01
hsa-miR-513a-3p↓	60	-4.9	-4.3	M	miR-513a-3p	AAAUUUC	468	453	155↑	**1.9E-09**
hsa-miR-514↓	7	-13.0	-14.6	H	miR-514	UUGACAC	190	178	61↑	**1.7E-04**
hsa-miR-532-5p↓	89	-1.3	-1.4	H	miR-532/532-5p	AUGCCUU	130	125	42↑	3.2E-03
hsa-miR-542-5p↓	41	-1.6	-3.5	M	miR-542-5p.h	CGGGGAU	3	3	1↑	5.3E-01
hsa-miR-582-5p↓	94	-1.8	-2.4	M	miR-582-5p.h	UACAGUU	484	463	178↑	**7.4E-16**
hsa-miR-618↓	51	-1.3	-2.3	M	miR-618	AACUCUA	123	117	48↑	**4.9E-06**
hsa-miR-660↓	107	-1.9	-1.9	M	miR-660	ACCCAUU	89	85	23↑	2.9E-01
hsa-miR-720↓	20	-1.3	-1.6	H	miR-720.h	CUCGCUG	5	5	0↑	1.0E+00
hsa-miR-876-3p↓	54	-1.7	-4.7	M	miR-876-3p	GGUGGUU	93	89	31↑	6.6E-03
hsa-miR-888↓	36	-1.1	-8.1	M	miR-888	ACUCAAA	235	221	80↑	**1.7E-06**
hsa-miR-891a↓	80	-0.7	-4.5	M	miR-891a	GCAACGA	4	4	1↑	1.0E+00
hsa-miR-1201↓	106	-1.1	-1.2	H	miR-1201	GCCUGAU	92	88	38↑	**1.3E-05**
hsa-miR-1274b↓	91	-1.2	-1.3	H	miR-1274b	CCCUGUU	128	120	36↑	4.6E-02

	hsa-miR-1307↓	61	-1.9	-2.0	M	miR-1307	CUCGGCG	3	3	2↑	1.2E-01
	hsa-miR-1280↓	83	-1.2	-1.7	H	miR-1840	CACGUGA	50	47	11↑	8.6E-01
Others ↑ (n=10) ↓ (n=18)	hsa-miR-9*↓	88	-2.1	-2.1	M			N/A			
	hsa-miR-16-1*↓	46	-2.2	-2.5	M			N/A			
	hsa-miR-20a*↓	63	-1.7	-2.4	M			N/A			
	hsa-miR-24-1*↓	38	-2.3	-4.0	M			N/A			
	hsa-miR-29a*↓	30	-2.0	-2.7	H			N/A			
	hsa-miR-29c*↓	95	-1.1	-1.4	H			N/A			
	hsa-miR-30c-2*↓	100	-1.3	-1.6	M			N/A			
	hsa-miR-30d*↓	11	-1.7	-3.8	M			N/A			
	hsa-miR-129*↓	15	-4.5	-4.2	M			N/A			
	hsa-miR-132*↓	92	-1.9	-1.4	M			N/A			
	hsa-miR-154*↓	86	-2.2	-2.4	M			N/A			
	hsa-miR-181a*↓	98	-1.5	-1.5	M			N/A			
	hsa-miR-192*↓	1	-9.7	-32.2	M			N/A			
	hsa-miR-194*↓	45	-1.2	-2.0	M			N/A			
	hsa-miR-378*↓	71	-1.0	-2.5	M			N/A			
	hsa-miR-424*↓	91	-1.5	-2.1	M			N/A			
	HS_100↓	83	-1.1	-1.7	H			N/A			
	HS_97↓	93	-1.1	-2.2	M			N/A			
	HS_108.1↑	79	2.4	3.4	M			N/A			
	HS_204.1↑	49	1.2	1.2	H			N/A			
	HS_244↑	98	1.4	1.5	M			N/A			

	hsa-let-7d*↑	60	1.5	1.7	H			N/A			
	hsa-miR-10a*↑	67	1.2	1.4	M			N/A			
	hsa-miR-199a*:9.1↑	29	1.2	1.4	H			N/A			
	hsa-miR-23b*↑	96	2.1	2.3	M			N/A			
	hsa-miR-493*↑	75	1.6	2.8	M			N/A			
	hsa-miR-664*↑	78	2.4	3.1	M			N/A			
	solexa-9029-92↑	102	1.2	1.2	M			N/A			

TargetScan classifies miRNAs into families based on identical seed region

Class 1: broadly conserved miRNA family across most vertebrates, usually to zebrafish

Class 2: conserved miRNA family across most mammals, but usually not beyond placental mammals

Class 3: poorly conserved miRNA family

Kidney enriched miRNAs are in bold and italic

According to the average expression level of miRNAs in any of 3 renal groups, dysregulated miRNAs were either highly expressed (H): signal intensity between 10000 – 40000, or moderately expressed (M): signal intensity between 1000 – 10000

##For a novel miRNA, only the conserved predicted targets were considered as targets

^A two-tailed Fisher's Exact Test was used to determine if the number of predicted target genes that actually changed expression in *PKD1* renal cysts was greater than would be expected by chance. All p-values were adjusted based on the Bonferroni method accounting for multiple testing. $p < 4.67E\text{-}04\ (0.05/107)$ were considered as significance

Dysregulated miRNA/mRNA: Up (↑) and Down (↓) in PKD1 renal cysts vs. MCT and normal kidney samples

Supplementary Table 2. Dysregulated miRNA clusters within a proximal 50kb distance on a chromosome in PKD1 renal cysts

miRNA Clusters# (n=23)	**Chromosome**	**Dysregulated miRNAs##**							
mir-30e (2)	1p34.2	miR-30c↓	miR-30e↓						
mir-215-194-1 (2)	1q42.1	miR-194↓	miR-215↓						
mir-216a (3)	2p16.1	miR-216a↓	miR-217↓						
mir-30a (2)	6q13	miR-30a↓	miR-30c↓	miR-30c-2*↓					
mir-30b (2)	8q24.22	miR-30b↓	miR-30d↓	miR-30d*↓					
mir-194-2-192 (2)	11q13.1	miR-192↓	miR-192*↓	miR-194↓	miR-194*↓				
mir-15a (2)	13q14.2	miR-15a↓	miR-16-1*↓						
mir-17-92 (6)	13q31.3	miR-17↓	miR-18a↓	miR-19a↓	miR-20a↓	miR-20a*↓	miR-19b↓		
mir-188 (7)	Xp11.23	miR-532-5p↓	miR-188-5p↓	miR-660↓	miR-362-3p↓	miR-362-5p↓	miR-500↓	miR-501-3p↓	miR-501-5p↓
mir-424 (6)	Xq26.3	miR-424↓	miR-424*↓	miR-503↓	miR-542-3p↓	miR-542-5p↓	miR-450a↓	miR-450b-5p↓	
mir-106a-363 (6)	Xq26.2	miR-106a↓	miR-18b↓	miR-20b↓	miR-19b↓	miR-363↓			
mir-888 (6)	Xq27.3	miR-888↓	miR-891a↓						
mir-506 (14)	Xq27.31-Xq28	miR-514↓	miR-509-3-5p↓	miR-509-3p↓	miR-508-3p↓	miR-507↓	miR-506↓	miR-513a-3p↓	
mir-199a-1 (3)	1q24.3	miR-214↑	miR-199a-3p↑						
let-7d (3)	9q22.32	let-7d↑	let-7d*↑	let-7f↑					

mir-100 (3)	11q24.1	miR-100↑	miR-125b↑						
mir-99b (3)	19q13.41	let-7e↑	miR-125a-5p↑						
mir-221 (2)	Xp11.3	miR-221↑	miR-222↑						
mir-224 (2)	Xq28	miR-224↑	miR-452↑						
mir-106b-25 (3)	7q22.1	miR-106b↓	miR-25↑						
mir-23b (3)	9q22.32	miR-24-1*↓	miR-23b*↑						
mir-127 (10)	14q32.2	miR-337-5p↓	miR-493*↑	miR-432↑					
mir-134 (41)	14q32.31	miR-154*↓	miR-134↑	miR-379↑					

The numbers in brackets are the total numbers of miRNA genes within a proximal 50kb distance on a chromosome

Dysregulated miRNA: Up (↑) and Down (↓) in PKD1 renal cysts

Supplementary Table 3 Linear regression analysis for selected clinical and biochemical variables identifying independent associations with differentially expressed microRNAs. Standardized β (standardized coefficients beta) denotes that an alteration of 1SD in the clinical parameters (independent variable) will lead to a one measurement change in the corresponding differentially expressed microRNAs (dependent factor).

Variables	MiR-30a-5p Standardized β (P value)	MiR-30d-5p Standardized β (P value)	MiR-30e-5p Standardized β (P value)	MiR-192-5p Standardized β (P value)	MiR-194-5p Standardized β (P value)
Age (year)	0.128 (0.338)	0.133 (0.4)	0.133 (0.318)	0.091 (0.498)	0.164 (0.218)
BMI (kg/m^2)	-0.124 (0.37)	-0.133 (0.418)	-0.065 (0.63)	0.061 (0.907)	0.031 (0.924)
eGFR ($ml/min/1.73m^2$)	-0.577 (**0.00**)	-0.444 (**0.002**)	-0.550 (**0.00**)	-0.495 (**0.00**)	-0.288 (0.055)
eGFR slope (ml/min/year)	-0.415 (**0.01**)	-0.275 (0.085)	-0.399 (**0.01**)	-0.369 (**0.01**)	-0.234 (0.146)
MKL (cm)	0.432 (**0.01**)	0.232 (0.154)	0.436 (**0.006**)	0.533 (**0.000**)	0.294 (0.069)
Cholesterol (mmol/l)	-0.186 (0.162)	-0.190 (0.15)	-0.222 (0.09)	-0.157 (0.24)	-0.066 (0.621)
Uric acid (µmol/l)	0.103 (0.464)	0.466 (**0.000**)	0.137 (0.328)	0.027 (0.846)	0.123 (0.328)
PTH (mmol/l)	0.362 (**0.036**)	0.120 (0.499)	0.301 (0.084)	0.322 (0.064)	-0.032 (0.859)
PCR (mg/mmol)	0.358 (0.057)	0.289 (0.128)	0.372 (**0.047**)	0.458 (**0.013**)	0.192 (0.317)

Supplementary Table 4 Pathway analysis using DAVID based on changes in the five differentially expressed miRNA

Significantly enriched pathways	miR-30a P value (fold enrichment)	miR-30d P value (fold enrichment)	miR-30e P value (fold enrichment)	miR-192 P value (fold enrichment)	miR-194 P value (fold enrichment)
Wnt signaling	0.02 (1.5)	0.01 (1.5)	0.02(1.5)	ns	4.45E-04 (2)
MAPK signaling	0.02 (1.3)	0.03 (1.3)	0.01 (1.4)	0.04 (1.5)	ns
ErbB signaling	0.02 (1.7)	0.01 (1.8)	0.007 (1.8)	ns	ns
Calcium signaling	0.04 (1.4)	0.04 (1.4)	0.03 (1.4)	ns	ns
Apoptosis	0.003(1.98)	0.006 (1.9)	0.01 (1.7)	0.01 (2.3)	ns
p53 signaling	ns	ns	0.03 (1.7)	ns	ns
Cell cycle	ns	ns	ns	0.03 (1.9)	ns
Endocytosis	ns	ns	ns	ns	0.002 (1.7)
Focal adhesion	ns	ns	ns	ns	0.01 (1.5)
JAK-STAT signaling	ns	ns	ns	ns	0.02 (1.6)
mTOR signaling	ns	ns	ns	ns	0.01 (2.3)

Supplementary Table 5. Selected down-regulated miRNAs in human PKD1 renal cysts identified by SAM (FDR <0.1%), whose expression were also inversely correlated with the expression of their target genes (p < 4.67E-04)

Down-regulated miRNAs (n=56)		Seed + m8 (42 families)	Ranking by SAM score	Fold changes cyst/MCT	Fold changes cyst/kid	Expression levels #	Inverse correlation p value^
Class 1 (n=39)	hsa-miR-9	CUUUGGU	17	-4.2	-4.4	M	1.6E-20
	hsa-miR-15a	AGCAGCA	102	-1.2	-1.2	H	8.8E-21
	hsa-miR-424		21	-1.3	-1.5	H	
	hsa-miR-17	AAAGUGC	77	-1.1	-1.3	H	1.5E-23
	hsa-miR-20a		98	-1.2	-1.2	H	
	hsa-miR-20b		58	-1.7	-2.5	H	
	hsa-miR-106a		91	-1.1	-1.3	H	
	hsa-miR-106b		73	-1.4	-1.6	H	
	hsa-miR-18a	AAGGUGC	44	-2.1	-2.7	M	2.3E-07
	hsa-miR-18b		22	-2.0	-5.1	M	
	hsa-miR-19a	GUGCAAA	66	-1.9	-4.9	M	1.1E-20
	hsa-miR-19b		104	-1.2	-2.4	H	
	hsa-miR-30a	GUAAACA	57	-1.4	-1.4	H	2.7E-28
	hsa-miR-30b		36	-1.3	-1.4	H	
	hsa-miR-30c		42	-1.2	-1.2	H	
	hsa-miR-30d		88	-1.1	-1.2	H	
	hsa-miR-30e		26	-1.4	-1.9	H	
	hsa-miR-101	ACAGUAC	76	-1.3	-1.8	H	4.2E-12
	hsa-miR-107	GCAGCAU	23	-1.5	-4.2	H	4.5E-19
	hsa-miR-124	AAGGCAC	15	-1.8	-13.4	M	1.2E-26
	hsa-miR-506		72	-4.7	-4.4	M	
	hsa-miR-135a	AUGGCUU	19	-1.9	-2.8	H	3.2E-15
	hsa-miR-135b		95	-2.6	-1.4	H	
	hsa-miR-138	GCUGGUG	10	-2.1	-5.5	M	2.0E-10
	hsa-miR-140-5p	AGUGGUU	96	-1.5	-1.7	H	7.3E-09
	hsa-miR-141	AACACUG	46	-2.4	-3.5	H	8.8E-19
	hsa-miR-200a		92	-1.8	-2.0	H	
	hsa-miR-142-3p	GUAGUGU	52	-2.2	-1.8	H	1.9E-08
	hsa-miR-144	ACAGUAU	91	0.7	-6.6	M	2.0E-16

	hsa-miR-194	GUAACAG	6	-2.8	-3.6	H	6.1E-07
	hsa-miR-204	UCCCUUU	35	-1.9	-2.8	H	7.6E-09
	hsa-miR-206	GGAAUGU	99	-1.4	-2.4	M	3.4E-13
	hsa-miR-217	ACUGCAU	55	0.9	-5.6	M	1.3E-04
	hsa-miR-223	GUCAGUU	56	-1.5	1.0	H	2.4E-06
	hsa-miR-301a	AGUGCAA	36	-2.4	-4.5	H	2.6E-12
	hsa-miR-338-3p	CCAGCAU	23	-2.1	-3.3	H	9.2E-06
	hsa-miR-363	AUUGCAC	89	-1.5	-1.9	H	2.4E-19
	hsa-miR-455-5p	AUGUGCC	37	-2.1	-3.4	M	1.0E-04
	hsa-miR-503	AGCAGCG	87	-1.4	-2.4	M	2.8E-13
Class 2 (n=2)	hsa-miR-340	UAUAAAG	62	-2.1	-1.7	M	2.3E-21
	hsa-miR-362-3p	ACACACC	43	-2.1	-2.4	H	3.6E-09
Class 3 (n=15)	hsa-miR-129-3p	AGCCCUU	29	-4.8	-2.3	M	3.1E-06
	hsa-miR-142-5p	AUAAAGU	70	-3.4	-4.0	M	1.3E-14
	hsa-miR-450b-5p	UUUGCAA	47	-2.0	-2.1	M	8.2E-20
	hsa-miR-500	AAUCCUU	32	-1.8	-2.1	H	7.5E-07
	hsa-miR-501-3p	AUGCACC	73	-1.5	-1.7	M	2.6E-06
	hsa-miR-501-5p	AUCCUUU	74	-1.9	-2.2	M	6.9E-06
	hsa-miR-507	UUUGCAC	41	-1.2	-3.0	M	2.2E-12
	hsa-miR-508-3p	GAUUGUA	43	-5.5	-5.2	M	5.2E-06
	hsa-miR-509-3-5p	ACUGCAG	43	-12.4	-12.7	M	9.1E-07
	hsa-miR-513a-3p	AAAUUUC	60	-4.9	-4.3	M	1.9E-09
	hsa-miR-514	UUGACAC	7	-13.0	-14.6	H	1.7E-04
	hsa-miR-582-5p	UACAGUU	94	-1.8	-2.4	M	7.4E-16
	hsa-miR-618	AACUCUA	51	-1.3	-2.3	M	4.9E-06
	hsa-miR-888	ACUCAAA	36	-1.1	-8.1	M	1.7E-06
	hsa-miR-1201	GCCUGAU	106	-1.1	-1.2	H	1.3E-05

TargetScan classifies miRNAs into families based on identical seed region; for a novel miRNA, only the conserved predicted target genes were considered as targets.

Class 1: broadly conserved miRNA family across most vertebrates, usually to zebrafish

Class 2: conserved miRNA family across most mammals, but usually not beyond placental mammals

Class 3: poorly conserved miRNA family

highly expressed (H): signal intensity 10000 – 40000, moderately expressed (M): 1000 – 10000.

^ A Fisher's Exact Test was used to determine if the number of predicted target genes that actually changed expression in PKD1 renal cysts was greater than would be expected by chance. All p-values were adjusted based on the Bonferroni method accounting for multiple testing; $p < 4.67E-04$ (0.05/107) were considered as significance.

Supplementary Table 6. Pathways predicted to be affected by 42 downregulated miRNA families as identified by GSEA at NOM p ≤ 0.05 and FDR ≤ 0.25

Up-regulated (50)	**NES**	**NOM p-val**	**FDR q-val**	**Ranking by NES**
Proliferation (1)				
SMITH_HTERT_UP*	1.65	0.011	0.182	16
Mitogenic signaling pathways via RTKs (7)				
IGF1_NIH3T3_UP*	1.76	0.005	0.113	9
MAPKPATHWAY*	1.58	0.011	0.217	27
HDACPATHWAY*	1.57	0.023	0.208	28
ST_JNK_MAPK_PATHWAY*	1.53	0.024	0.220	37
SIG_INSULIN_RECEPTOR_PATHWAY_IN_CARDIAC_MYOCYTES*	1.47	0.037	0.221	48
FCER1PATHWAY*	1.49	0.050	0.243	43
NGFPATHWAY*	1.54	0.042	0.241	31
Cell cycle (6)				
IGLESIAS_E2FMINUS_UP*	1.82	<0.001	0.095	5
LEE_E2F1_UP*	1.70	0.007	0.129	13
RACCYCDPATHWAY*	1.68	0.012	0.148	15
HSA04110_CELL_CYCLE*	1.58	0.017	0.220	25
VERNELL_PRB_CLSTR1*	1.55	0.022	0.237	30
G1PATHWAY	1.53	0.046	0.224	34
Wnt pathway (3)				
WNT DIRECT TARGETS*	1.69	0.014	0.141	14
ST_WNT_BETA_CATENIN_PATHWAY*	1.61	0.024	0.212	20
WNT_SIGNALING*	1.47	0.038	0.224	47
Notch pathway (2)				
NGUYEN_KERATO_DN*	1.47	0.036	0.225	46
HSA04330_NOTCH_SIGNALING_PATHWAY*	1.58	0.030	0.227	24
BMP/TGFβ/Activin pathway (5)				
TGFBETA_ALL_UP*	1.78	0.004	0.113	7
TGFBETA_EARLY_UP*	1.96	<0.001	0.038	2
HSA04350_TGF_BETA_SIGNALING_PATHWAY*	1.72	<0.001	0.122	12
ALKPATHWAY*	1.56	0.042	0.217	29
TGFBETA_C1_UP*	1.52	0.043	0.219	39
Epithelial-mesenchymal transition (2)				
EMT_UP*	1.98	<0.001	0.041	1
JECHLINGER_EMT_UP*	1.93	<0.001	0.040	3
Hypoxia pathway (3)				

HYPOXIA_REVIEW*	1.63	0.009	0.205	17
HIFPATHWAY*	1.48	0.047	0.223	45
HYPOXIA_FIBRO_UP*	1.61	0.023	0.202	21
Aging (1)				
AGEING_KIDNEY_SPECIFIC_UP*	1.75	<0.001	0.117	10
Ca^{2+} mediated signaling pathway (3)				
NDKDYNAMINPATHWAY*	1.50	0.046	0.232	41
CALCINEURINPATHWAY*	1.73	0.005	0.122	11
NFATPATHWAY*	1.42	0.042	0.243	50
Genomic integrity (1)				
PARP_KO_UP*	1.60	0.015	0.205	23
Cytokine mediated JAK-STAT, MEK/ERK cascades (4)				
IL2RBPATHWAY*	1.63	0.017	0.202	18
TPOPATHWAY*	1.61	0.013	0.216	19
GHPATHWAY*	1.49	0.042	0.245	42
IL2PATHWAY	1.52	0.045	0.222	38
MYC regulated genes (1)				
LEE_MYC_UP*	1.53	0.041	0.215	36
Immune and inflammatory response (2)				
PAR1PATHWAY*	1.54	0.042	0.235	32
NTHIPATHWAY*	1.49	0.047	0.236	44
p53 pathway (1)				
P53GENES_ALL*	1.77	0.004	0.112	8
Oxidative damage (1)				
ARENRF2PATHWAY*	1.86	0.003	0.068	4
Differentiation (1)				
ETSPATHWAY*	1.54	0.032	0.231	33
Nuclear receptor signaling (4)				
GCRPATHWAY*	1.58	0.018	0.212	26
CARM_ERPATHWAY*	1.60	0.020	0.204	22
BREAST_CANCER_ESTROGEN_SIGNALING*	1.44	0.038	0.229	49
RARRXRPATHWAY*	1.51	0.048	0.232	40
Wilms' tumor signature (1)				
LI_FETAL_VS_WT_KIDNEY_DN*	1.81	<0.001	0.092	6
Cell communication (1)				
HSA01430_CELL_COMMUNICATION	1.53	0.025	0.220	35

* Overlapping pathways (n=47) between using predicted miRNA target genes (6319) and all genes.

Supplementary References

S1. Song X, Di Giovanni V, He N*, et al.* Systems biology of autosomal dominant polycystic kidney disease (ADPKD): computational identification of gene expression pathways and integrated regulatory networks. *Hum Mol Genet* 2009; **18:** 2328-2343.

S2. Griffiths-Jones S, Bateman A, Marshall M*, et al.* Rfam: an RNA family database. *Nucleic Acids Res* 2003; **31:** 439-441.

S3. Newby LJ, Streets AJ, Zhao Y*, et al.* Identification, characterization, and localization of a novel kidney polycystin-1-polycystin-2 complex. *J Biol Chem* 2002; **277:** 20763-20773.

S4. Streets AJ, Magayr TA, Huang L*, et al.* Parallel microarray profiling identifies ErbB4 as a determinant of cyst growth in ADPKD and a prognostic biomarker for disease progression. *Am J Physiol Renal Physiol* 2017; **312:** F577-F588.

S5. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S*, et al.* DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res* 2018; **46:** D239-D245.

S6. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; **8:** 118-127.

S7. Mullokandov G, Baccarini A, Ruzo A*, et al.* High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries. *Nat Methods* 2012; **9:** 840-846.

S8. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 2001; **98:** 31-36.

S9. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001; **98:** 5116-5121.

S10. Friedman RC, Farh KK, Burge CB*, et al.* Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009; **19:** 92-105.

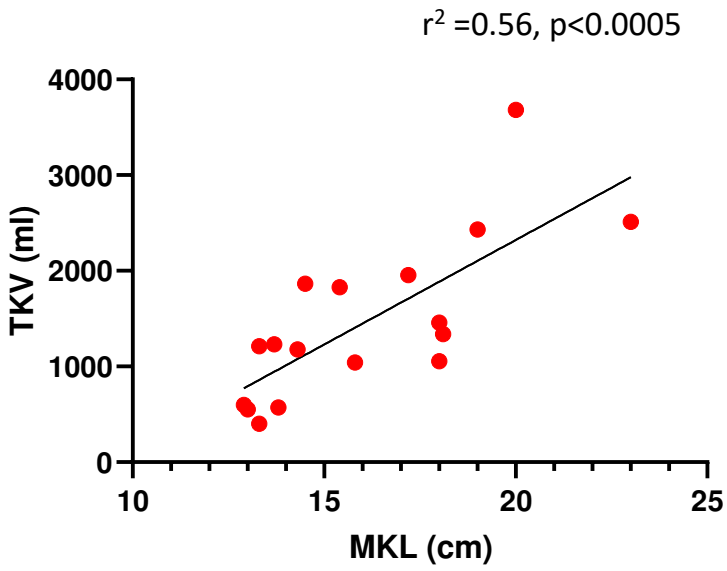
S11. Baek D, Villen J, Shin C*, et al.* The impact of microRNAs on protein output. *Nature* 2008; **455:** 64-71.

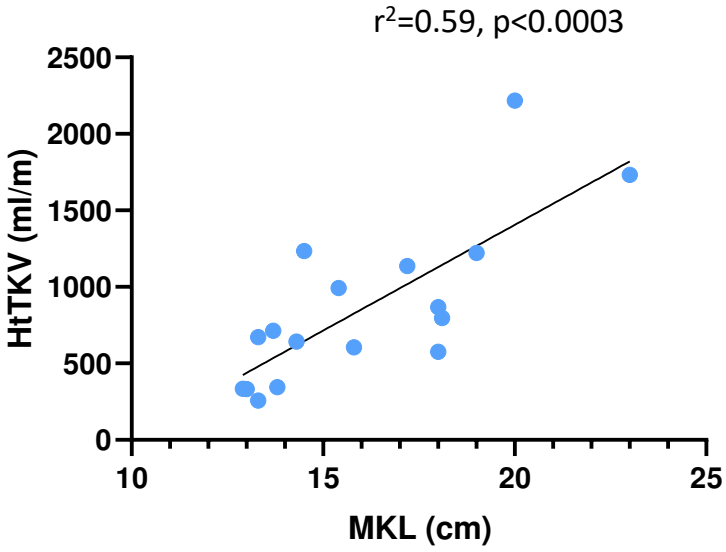
S12. Subramanian A, Tamayo P, Mootha VK*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; **102:** 15545-15550.

S13. Nogales-Cadenas R, Carmona-Saez P, Vazquez M*, et al.* GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. *Nucleic Acids Res* 2009; **37:** W317-322.

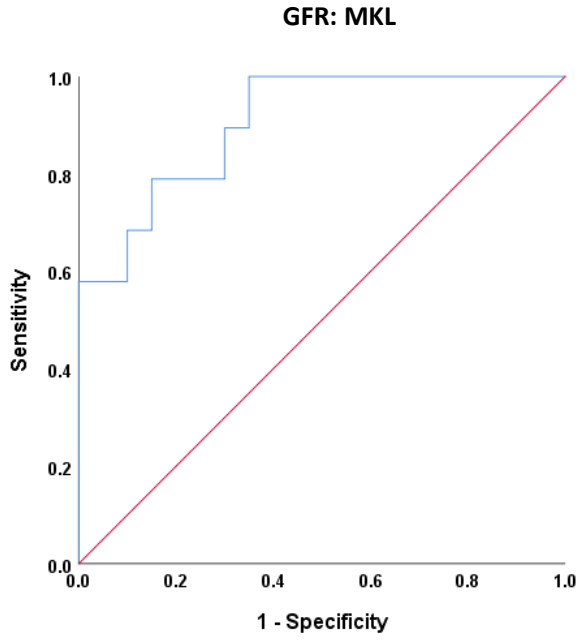
Figure S1

A


r² =0.56, p<0.0005
TKV (ml)
4000
3000
2000
1000
0
10
15
20
25
MKL (cm)


r²=0.59, p<0.0003
HtTKV (ml/m)
2500
2000
1500
1000
500
0
10
15
20
25
MKL (cm)

B


GFR: MKL
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1 - Specificity

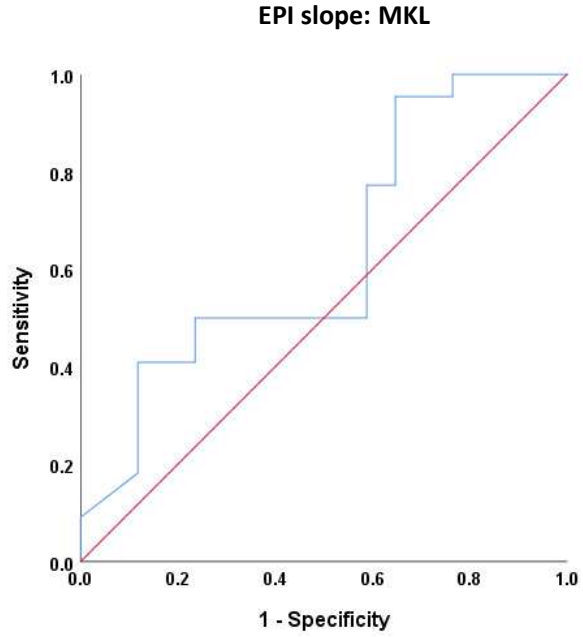

EPI slope: MKL
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1 - Specificity

Figure S2

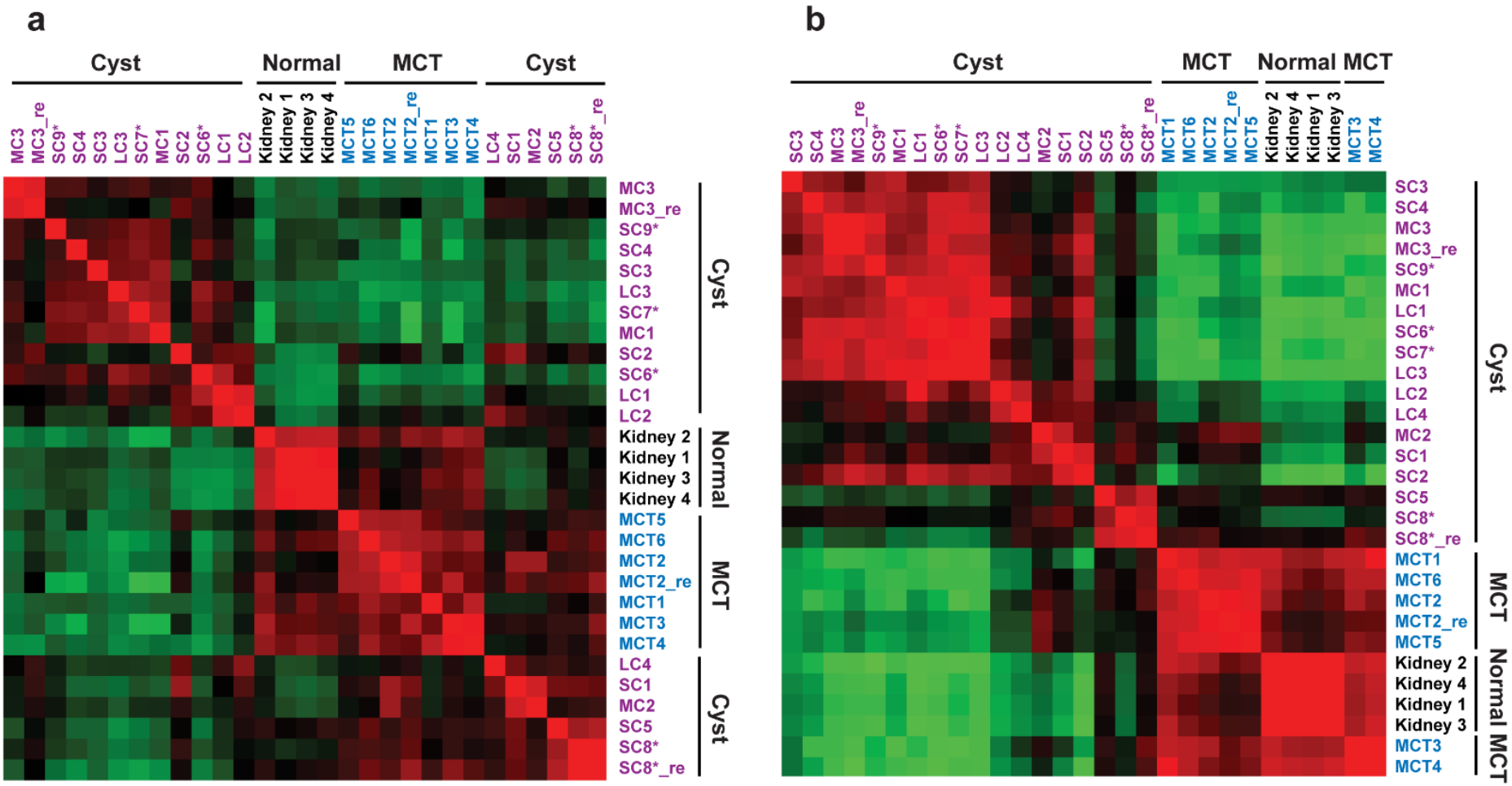

a
Cyst
Normal
MCT
Cyst
MC3
MC3_re
SC9*
SC4
SC3
LC3
SC7*
MC1
SC2
SC6*
LC1
LC2
Kidney 2
Kidney 1
Kidney 3
Kidney 4
MCT5
MCT6
MCT2
MCT2_re
MCT1
MCT3
MCT4
LC4
SC1
MC2
SC5
SC8*
SC8*_re
Cyst
Normal
MCT
Cyst
b
Cyst
MCT
Normal
MCT
SC3
SC4
MC3
MC3_re
SC9*
MC1
LC1
SC6*
SC7*
LC3
LC2
LC4
MC2
SC1
SC2
SC5
SC8*
SC8*_re
MCT1
MCT6
MCT2
MCT2_re
MCT5
Kidney 2
Kidney 4
Kidney 1
Kidney 3
MCT3
MCT4
Cyst
MCT
Normal
MCT

Figure S3

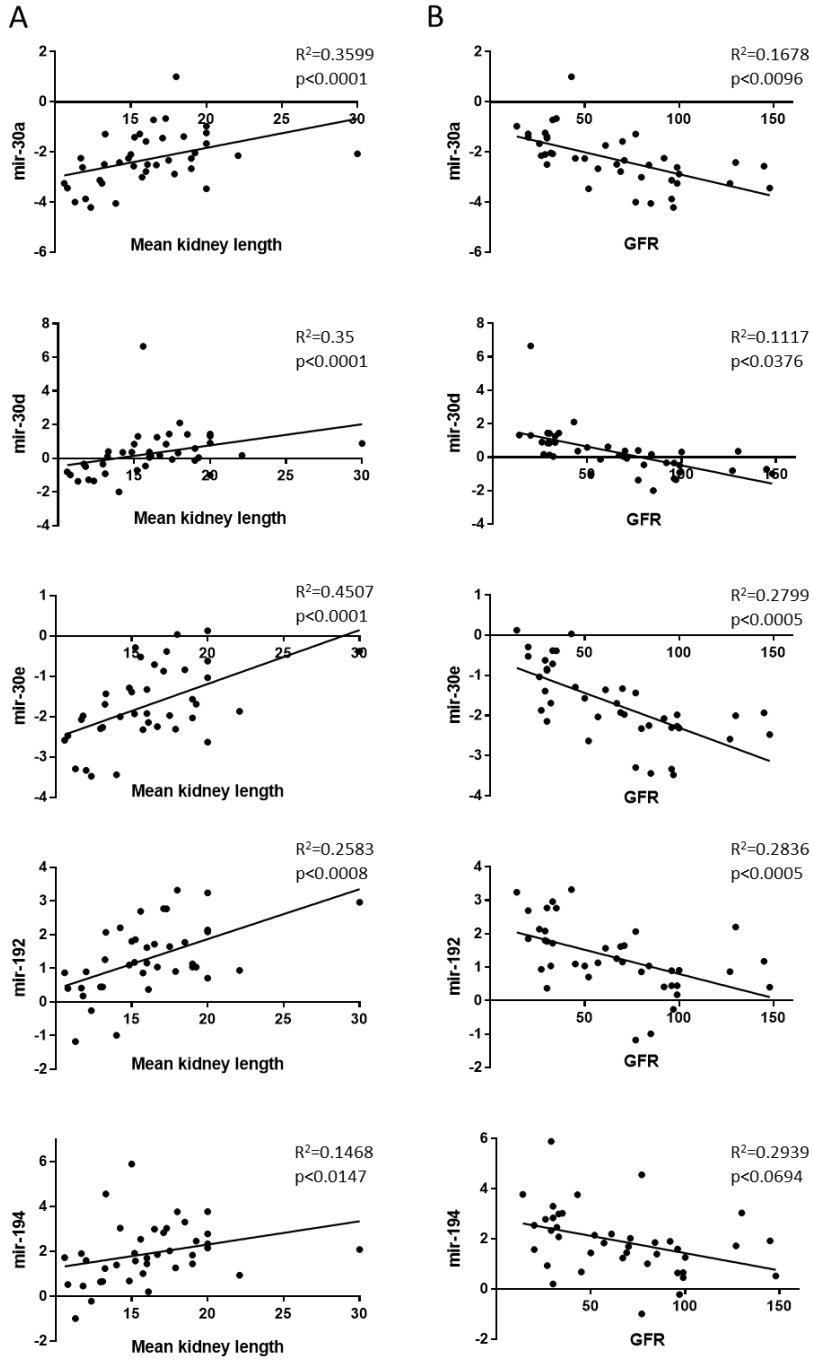
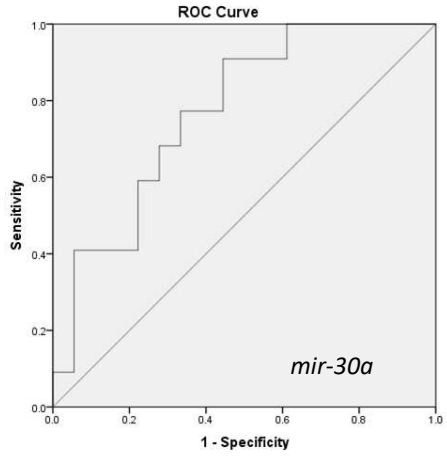
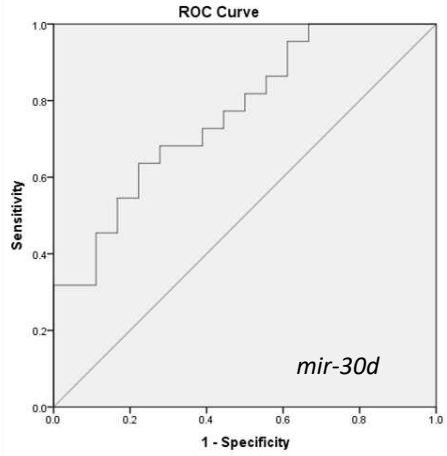
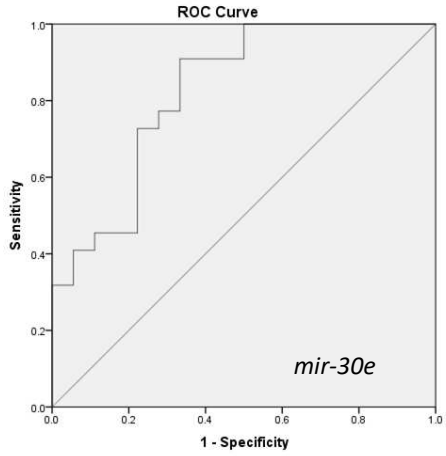
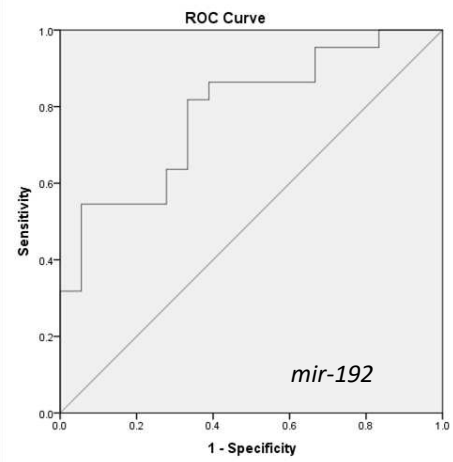

A
B
mir-30a
R²=0.3599
p<0.0001
Mean kidney length
mir-30a
R²=0.1678
p<0.0096
GFR
mir-30d
R²=0.35
p<0.0001
Mean kidney length
mir-30d
R²=0.1117
p<0.0376
GFR
mir-30e
R²=0.4507
p<0.0001
Mean kidney length
mir-30e
R²=0.2799
p<0.0005
GFR
mir-192
R²=0.2583
p<0.0008
Mean kidney length
mir-192
R²=0.2836
p<0.0005
GFR
mir-194
R²=0.1468
p<0.0147
Mean kidney length
mir-194
R²=0.2939
p<0.0694
GFR

Figure S4


ROC Curve
Sensitivity
1 - Specificity
mir-30a


ROC Curve
Sensitivity
1 - Specificity
mir-30d


ROC Curve
Sensitivity
1 - Specificity
mir-30e


ROC Curve
Sensitivity
1 - Specificity
mir-192

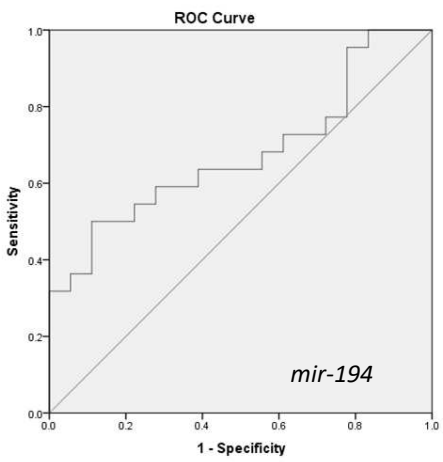

ROC Curve
Sensitivity
1 - Specificity
mir-194

Figure S5

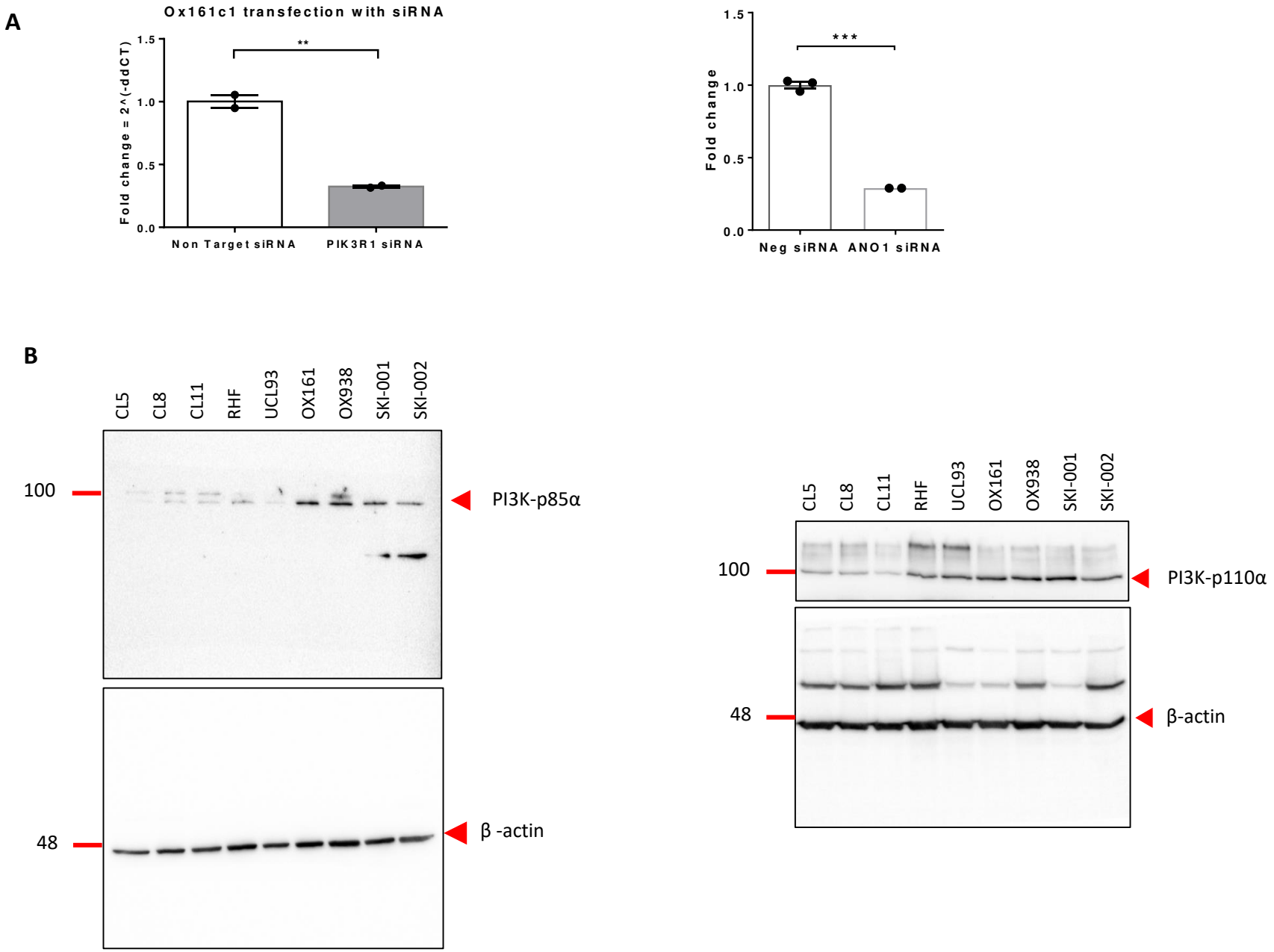
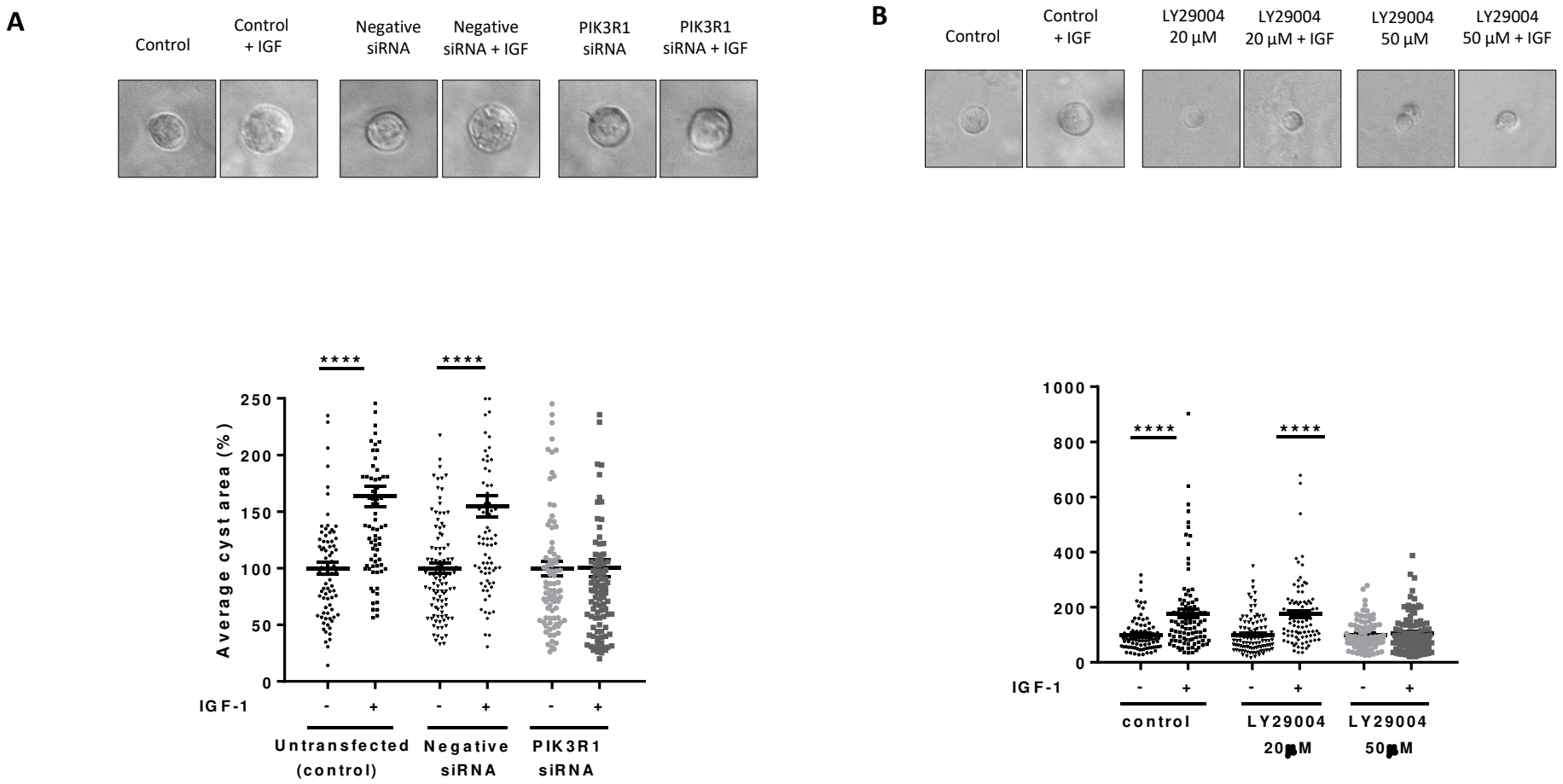

A
Ox161c1 transfection with siRNA
Fold change = 2^(-ddCT)
1.5
1.0
0.5
0.0
**
Non Target siRNA
PIK3R1 siRNA
Fold change
1.5
1.0
0.5
0.0
***
Neg siRNA
ANO1 siRNA
B
CL5
CL8
CL11
RHF
UCL93
OX161
OX938
SKI-001
SKI-002
100
PI3K-p85α
48
β -actin
CL5
CL8
CL11
RHF
UCL93
OX161
OX938
SKI-001
SKI-002
100
PI3K-p110α
48
β-actin

Figure S6


A
Control
Control + IGF
Negative siRNA
Negative siRNA + IGF
PIK3R1 siRNA
PIK3R1 siRNA + IGF
B
Control
Control + IGF
LY29004 20 µM
LY29004 20 µM + IGF
LY29004 50 µM
LY29004 50 µM + IGF
****
****
Average cyst area (%)
250
200
150
100
50
0
IGF-1
- + - + - +
Untransfected (control)
Negative siRNA
PIK3R1 siRNA
1000
800
600
400
200
0
control
LY29004 20µM
LY29004 50µM