

OPEN

# Identification of housekeeping genes for microRNA expression analysis in kidney tissues of *Pkd1* deficient mouse models

J. J. Muñoz<sup>1,4</sup>, A. C. Anauate<sup>1,4</sup>, A. G. Amaral<sup>2</sup>, F. M. Ferreira<sup>3</sup>, R. Meca<sup>1</sup>, M. S. Ormanji<sup>1</sup>, M. A. Boim<sup>1</sup>, L. F. Onuchic<sup>2</sup> & I. P. Heilberg<sup>1\*</sup>

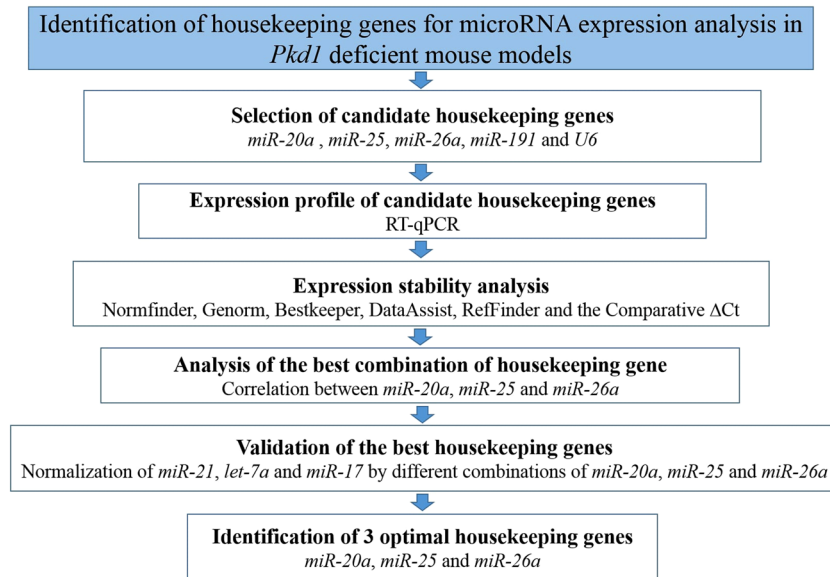
Polycystic kidney disease is a complex clinical entity which comprises a group of genetic diseases that leads to renal cyst development. We evaluated the most suitable housekeeping genes for microRNA expression by RT-qPCR analyses of kidney tissues in *Pkd1*-deficient mouse models from a panel of five candidate genes (*miR-20a*, *miR-25*, *miR-26a*, *miR-191* and *U6*) and 3 target genes (*miR-17*, *miR-21* and *let-7a*) using samples from kidneys of cystic mice (*Pkd1<sup>flox/flox</sup>:Nestin<sup>cre</sup>*, CY), non-cystic controls (*Pkd1<sup>flox/flox</sup>*, NC), *Pkd1*-haploinsufficient (*Pkd1<sup>+/-</sup>*, HT), wild-type controls (*Pkd1<sup>+/+</sup>*, WT), severely cystic mice (*Pkd1<sup>VIV</sup>*, SC), wild-type controls (CO). The stability of the candidate genes was investigated using NormFinder, GeNorm, BestKeeper, DataAssist, and RefFinder software packages and the comparative  $\Delta$ Ct method. The analyses identified *miR-26a* as the most stable housekeeping gene for all kidney samples, *miR-20a* for CY and NC, *miR-20a* and *miR-26a* for HT and WT, and *miR-25* and *miR-26a* for SC and CO. Expression of *miR-21* was upregulated in SC compared to CO and trends of *miR-21* upregulation and *let-7a* downregulation in CY and HT compared to its control kidneys, when normalized by different combinations of *miR-20a*, *miR-25* and *miR-26a*. Our findings established *miR-20a*, *miR-25*, and *miR-26a* as the best housekeeping genes for miRNA expression analyses by RT-qPCR in kidney tissues of *Pkd1*-deficient mouse models.

Autosomal dominant polycystic kidney disease (ADPKD) is the most frequent monogenic renal disorder, characterized by bilateral development of multiple cysts formation and gradual enlargement of the kidneys<sup>1</sup>. This disease is the fourth leading cause of kidney failure and affects about 1/400 to 1/1000 people worldwide<sup>1</sup>. The genetic and non-genetic factors determines the varying rates at which ADPKD progresses to end-stage renal disease<sup>2</sup>. ADPKD is caused by mutations in *PKD1* or *PKD2*, genes that encode polycystin-1 (PC1) and polycystin-2 (PC2), respectively. *PKD1* mutations are responsible for ~78% of the affected families and are associated with higher disease severity compared to mutations in *PKD2*<sup>3</sup>. PC1 regulates several aspects of tissue morphogenesis and function, which when disturbed promote abnormal cell proliferation, cyst development and transepithelial fluid secretion<sup>2,4-7</sup>.

Several studies have shown that microRNAs (miRNAs) are involved in the pathogenesis of ADPKD<sup>8-10</sup>. Importantly, miRNAs have been proposed as an acquired factor for the disease establishment and responsible for reducing or increasing cyst formation in ADPKD by regulating the expression of multiple target genes associated with cytogenesis<sup>11-19</sup>. A common set of miRNAs is aberrantly expressed in several murine models of polycystic kidney disease (PKD). Three miRNAs have their expression particularly altered in this setting; *let-7a* is downregulated while *miR-17* and *miR-21* are upregulated in kidney cysts and have been associated with PKD progression in mouse models<sup>20-24</sup>.

The heterozygous mouse (HT) is a model heterozygous for a null mutation in *Pkd1*, therefore represented by the genotype *Pkd1<sup>+/-</sup>*. This model reproduces the systemic *Pkd1*-haploinsufficient background found in ADPKD1 patients but virtually with no renal cysts by 15 weeks (wk) of life. Wild-type littermates (*Pkd1<sup>+/+</sup>*; WT)

<sup>1</sup>Nephrology Division, Department of Medicine, Universidade Federal de São Paulo, São Paulo, Brazil. <sup>2</sup>Divisions of Molecular Medicine and Nephrology, University of São Paulo School of Medicine, São Paulo, Brazil. <sup>3</sup>Laboratory of Immunology, Heart Institute, University of São Paulo School of Medicine, São Paulo, Brazil. <sup>4</sup>These authors contributed equally: J. J. Muñoz and A. C. Anauate \*email: ita.heilberg@gmail.com



**Figure 1.** Workflow diagram illustrating strategy for identification of housekeeping normalizer miRNAs for RT-qPCR. *Pkd1*, polycystic kidney disease 1 gene; miR, microRNA. For the sake of comparison, we included housekeeping genes selected from the literature, i.e., *miR-20a*, *miR-25*, *miR-26a*, *miR-191* and *U6*.

were used as their controls<sup>25,26</sup>. Homozygous mice for a *Pkd1*-floxed allele, in turn, with a mosaic pattern of gene inactivation driven by a Nestin-Cre transgene with consequent excision of exons 2–4 (*Pkd1*<sup>flox/flox</sup>; *Nestin*<sup>cre</sup>; CY), develop a renal cystic phenotype<sup>25,27,28</sup>. The *Pkd1*<sup>flox/flox</sup>; *Nestin*<sup>cre</sup> mice present cystic kidneys, reproducing the ADPKD phenotype, however do not have a systemic *Pkd1*-haploinsufficiency background. Noncystic animals (*Pkd1*<sup>flox/flox</sup>; NC) were their controls<sup>25,27–29</sup>. The third model comprised a homozygous mouse for a *Pkd1*-knockin hypomorphic allele that prevents PC1 cleavage at the GPS site (*Pkd1*<sup>V/V</sup>; SC). This mouse develops early-onset, post-natal, massive renal cystic disease, characterized by distal nephron involvement, uremia and early mortality<sup>30</sup>. Age-matched wild-type littermates (*Pkd1*<sup>+/+</sup>; WT) were used as their controls.

All three mouse models were maintained on the C57BL/6 strain background, an important requirement for the performance of our study. It must be noted that the WT and CO controls harbor the same genotype (*Pkd1*<sup>+/+</sup>) but are analyzed at different ages. NC controls, in turn, homozygous for the *Pkd1*<sup>flox</sup> allele but without the *Nestin*<sup>cre</sup> transgene, do not develop renal cysts<sup>25,30</sup>.

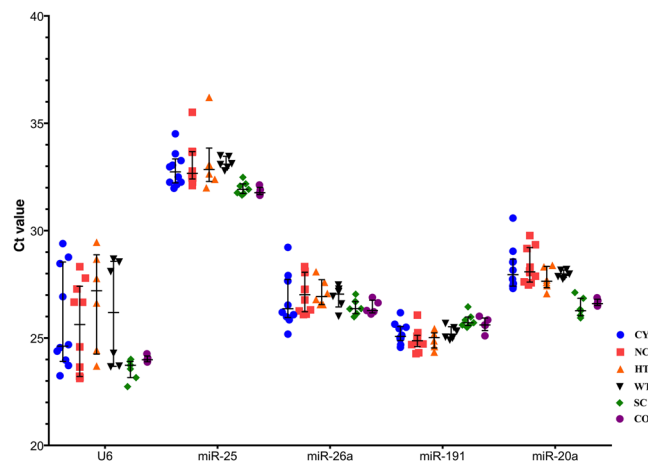
The regulatory function of miRNAs in ADPKD supports their use as potential prognostic and predictive biomarkers as well as therapeutic targets. Their identification, however, requires a reliable and quantitative assessment of miRNA expression. Different methodologies can be used to evaluate miRNA expression, but reverse transcription-quantitative polymerase chain reaction (RT-qPCR) remains the gold-standard for specific detection of selected sets of miRNAs<sup>31,32</sup>. A crucial step to ensure accurate and suitable quantification of PCR data is the normalization of expression levels. This normalization aims to differentiate true biological variations, explain the investigated phenotype and identify non-specific experimentally-induced alterations<sup>33</sup>. In fact, factors such as sample collection and preservation, amount of raw material, enzyme efficiency and RNA integrity can artefactually influence expression levels. Currently, the accepted method of miRNA gene expression normalization is the use of internal reference genes (housekeeping genes)<sup>33</sup>. Ideal housekeeping genes should show no (or minimal) expression variation in the sample under investigation in response to experimental treatment/disease condition.

Despite the growing number of studies investigating miRNA expression in animal models and human polycystic kidney disease<sup>20–24,34</sup>, to the best of our knowledge there is no current consensus on housekeeping genes to be employed in PKD/ADPKD kidney tissue samples. This limitation may restrict study comparisons and most importantly, lead to ambiguous data interpretation and misleading biological conclusions. Several miRNAs currently used as housekeeping genes in other diseases settings display unstable expression profiles, a reality that makes them not suitable for RT-qPCR data normalization<sup>35,36</sup>. *U6* expression profiles, therefore, require appropriate assessment in the PKD scenario, since it has been used as a housekeeping gene in research involving PKD samples<sup>37–39</sup>.

The present study aimed to quantitatively evaluate the performance of four candidate miRNA housekeeping genes (*miR-20a*, *miR-25*, *miR-26a* and *miR-191*) and the *U6* small nuclear RNA and identify the most suitable one/sets for miRNA expression normalization by RT-qPCR in kidney tissue of different *Pkd1*-deficient mouse models.

## Results

**Candidate housekeeping miRNAs: quantitative expression profiles.** Here we applied a stepwise strategy for the identification of the optimal miRNAs as housekeeping genes for miRNA expression by RT-qPCR analyses of kidney tissues in *Pkd1*-deficient mouse models. The scheme workflow is shown in Fig. 1.



**Figure 2.** The Ct values of five candidate housekeeping genes. A lower threshold value (Ct) indicates a higher gene expression level. The median values are expressed as horizontal lines, and the error bars represent interquartile range. The Ct values of *miR-25* were the highest, indicating the lowest expression levels.

The tissue samples were classified into 7 distinct groups: (1) CY, including the cystic kidney samples,  $n = 10$ ; (2) NC, non-cystic kidney samples,  $n = 10$ ; (3) HT, *Pkd1*-haploinsufficient kidney samples,  $n = 6$ ; (4) WT, wild-type kidney samples,  $n = 6$ ; (5) SC, severely cystic kidney samples,  $n = 7$ ; (6) CO, early-life, wild-type kidney samples,  $n = 5$ ; and (7) All, including all kidney samples,  $n = 44$ .

Before testing the expression stability of the five candidate housekeeping genes, all cDNA samples were normalized at the RNA level. After RT-qPCR cycling, the median Ct values of triplicate reactions were acquired, representing raw expression data. Expression of the *miR-20a*, *miR-25*, *miR-26a*, *miR-191* and *U6* candidate housekeeping genes is shown in Fig. 2.

The median Ct values for *miR-20a* [27.66 (1.26)], *miR-26a* [26.66 (1.07)] and *miR-191* [25.17 (0.70)] lied between 25 and 30, evincing moderate expression. Ct reached its highest value for *miR-25* [32.50 (1.02)], lying above 30, which makes it the lowest expressed gene among the five candidates. Ct values for *U6*, on the other hand, was the lowest one [24.28 (3.95)], placing it as the highest expressed gene (Fig. 2).

Here, we included RT-qPCR analyses of kidney tissues from three mouse models with distinct patterns of *Pkd1* deficiency and correlated them with their respective control tissues. The CY model presented some smaller variations in Ct values compared to its control NC (Fig. 2). *miR-26a* varied by more than 0.5 cycles in CY group but did not exceed 1.0 and Ct varied less than that for the other genes, except for *U6*. The HT model showed higher Ct variability compared to its control WT (Fig. 2). *U6* Ct values varied by more than 1.0 cycle between HT and WT. Expression differences were far lower between SC and its control (CO kidneys); only the *miR-20a* difference exceeded 0.3 cycle.

**Expression stability analysis of the housekeeping miRNAs.** The stability values of the candidate housekeeping genes were obtained applying six software packages (Supplementary Table 2). The top-ranked genes (associated with the smallest stability value) are the most stably expressed ones in the set of kidney tissues samples from the three *Pkd1*-deficient mouse models. Apart from *U6* in the HT group, all candidate housekeeping miRNAs presented M values below 1.5, the GeNorm set threshold, findings that are consistent with stability<sup>40</sup>. The BestKeeper software points out inconsistency when SD is higher than 1.0. The SD value was higher than 1.0 only for *U6* expression in the All, CY, NC, HT and WT sample groups (Supplementary Table 2).

Based on the different utilized algorithms and a visual inspection of all ranks generated by these analyses, *miR-26a* seems to be the best housekeeping gene for the All, HT and CO groups; *miR-20a* for the CY, NC and WT groups; and *miR-25* for the SC group (Table 1).

NormFinder recommends a SD value lower than 0.5 for genes to be considered relatively stable. Only *miR-20a*, *miR-25* and *miR-26a* had an SD value below 0.5, while *miR-191* and *U6* showed it above 0.5 in the All, CY and NC groups. However, *miR-25* presented SD higher than 0.5 in HT and WT and *miR-191*-related SD lower than 0.5 in SC and CO. While these data represent only a selection of possible tissue pairs (*Pkd1* deficiency kidney tissues versus controls), they illustrate that optimal housekeeping genes can significantly vary between kidneys of mouse models with distinct patterns of *Pkd1* deficiency (Tables 1, 2 and Supplementary Table 2).

GeNorm indicated that *miR-20a*, *miR-25*, *miR-26a*, *miR-191* and *U6* presented M values below 1.5 between the tissue groups, except for HT tissues that showed high variability ( $M > 1.5$ ) for *U6* (Supplementary Table 2).

The evaluation of relative expression stability by the BestKeeper software defines the genes that display SD higher than 1.0 as unstable. Again, *miR-20a*, *miR-25* and *miR-26a* presented SD lower than 1.0 and CV below 3.0 in the All, CY, NC, HT and WT groups (Supplementary Table 2). However, *U6* showed SD below 1.0 only for the SC and CO groups, thus it is not a suitable housekeeping gene for the All, CY, NC, HT and WT sample groups (Supplementary Table 2). Based on these results, *miR-20a*, *miR-25* and *miR-26a* were ranked as the most stable candidate housekeeping genes, whereas *miR-191* and *U6* were deemed least stable.

Groups	Best housekeeping genes by software						Best housekeeping gene by visual inspection
	NormFinder	GeNorm	RefFinder	$\Delta$ Ct method	Bestkeeper	DataAssist	
All	<i>miR-26a</i>	<i>miR-26a</i>	<i>miR-26a</i>	<i>miR-26a</i>	<i>miR-191</i>	<i>miR-26a</i>	<i>miR-26a</i>
CY	<i>miR-26a</i>	<i>miR-25</i>	<i>miR-20a</i>	<i>miR-20a</i>	<i>miR-191</i>	<i>miR-20a</i>	<i>miR-20a</i>
NC	<i>miR-26a</i>	<i>miR-26a</i>	<i>miR-20a</i>	<i>miR-20a</i>	<i>miR-191</i>	<i>miR-20a</i>	<i>miR-20a</i>
HT	<i>miR-26a</i>	<i>miR-26a</i>	<i>miR-20a</i>	<i>miR-20a</i>	<i>miR-191</i>	<i>miR-26a</i>	<i>miR-26a</i>
WT	<i>miR-20a</i>	<i>miR-25</i>	<i>miR-25</i>	<i>miR-26a</i>	<i>miR-20a</i>	<i>miR-20a</i>	<i>miR-20a</i>
SC	<i>miR-25</i>	<i>miR-25</i>	<i>miR-191</i>	<i>miR-191</i>	<i>miR-25</i>	<i>miR-25</i>	<i>miR-25</i>
CO	<i>miR-20a</i>	<i>miR-26a</i>	<i>miR-26a</i>	<i>miR-26a</i>	<i>miR-25</i>	<i>miR-20a</i>	<i>miR-26a</i>

**Table 1.** Best housekeeping miRNA for each group of samples yielded by software analyses. All, all samples; CY, cystic; NC, non-cystic; HT, haploinsufficient; WT, wild-type; SC, severely cystic phenotype; CO, severely cystic phenotype controls.

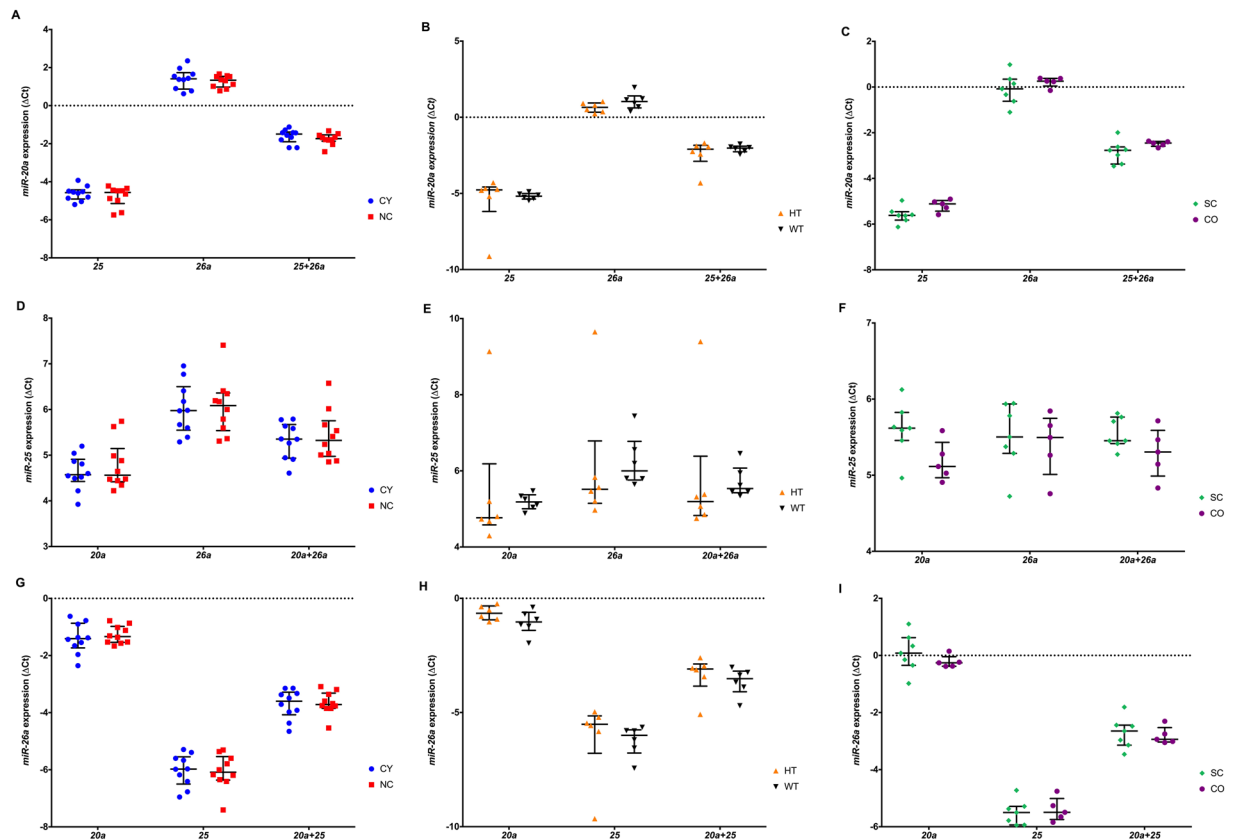
Groups	Best pair of housekeeping genes by software						Best pair of housekeeping gene by visual inspection	Best trio of housekeeping genes by visual inspection
	NormFinder	GeNorm	RefFinder	$\Delta$ Ct method	Bestkeeper	DataAssist		
All	<i>miR-25 + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-26a + miR-191</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-25 + miR-26a</i> or <i>miR-20a + miR-26a + miR-191</i>
CY	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-25</i>	<i>miR-20a + miR-25</i>	<i>miR-20a + miR-25</i>	<i>miR-25 + miR-191</i>	<i>miR-20a + miR-25</i>	<i>miR-20a + miR-25</i>	<i>miR-20a + miR-25 + miR-26a</i> or <i>miR-20a + miR-26a + miR-191</i>
NC	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-191</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a + miR-191</i>
HT	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-191</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a + miR-191</i>
WT	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-25</i>	<i>miR-25 + miR-26a</i>	<i>miR-25 + miR-26a</i>	<i>miR-20a + miR-25</i>	<i>miR-20a + miR-25</i>	<i>miR-20a + miR-25</i>	<i>miR-20a + miR-25 + miR-26a</i>
SC	<i>miR-25 + miR-191</i>	<i>miR-25 + miR-191</i>	<i>miR-25 + miR-191</i>	<i>miR-25 + miR-191</i>	<i>miR-25 + miR-191</i>	<i>miR-25 + miR-191</i>	<i>miR-25 + miR-191</i>	—
CO	<i>miR-20a + U6</i>	<i>miR-26a + miR-191</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>U6 + miR-25</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-26a</i>	<i>miR-20a + miR-25 + miR-26a</i> or <i>miR-20a + miR-26a + miR-191</i> or <i>miR-20a + miR-26a + U6</i>

**Table 2.** Best combination of housekeeping miRNAs for each group of samples yielded by software analyses. All, all samples; CY, cystic; NC, non-cystic; HT, haploinsufficient; WT, wild-type; SC, severely cystic phenotype; CO, severely cystic phenotype controls.

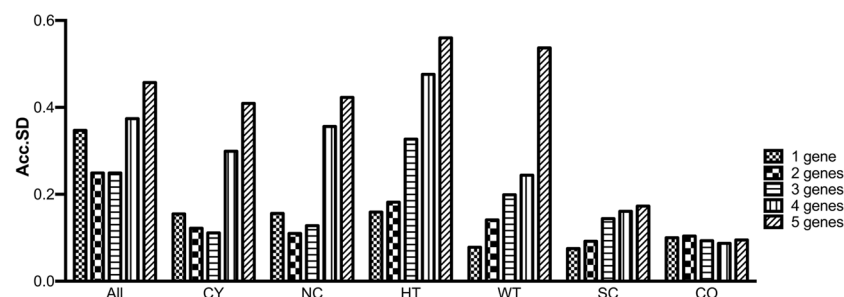
**Analysis of the best combination of housekeeping miRNAs.** The GeNorm software package recommends at least two genes as combination of housekeeping genes. Table 2 shows the best combination of housekeeping miRNAs for each model/control group pair, based on the different software packages and a visual inspection of all ranks generated by such analyses. The comparisons of *Pkd1*-deficiency kidney tissues versus their respective controls identified *miR-20a* for the CY and NC group pair, *miR-20a* and *miR-26a* for HT and WT, and *miR-25* and *miR-26a* for SC and CO as the most stable housekeeping genes. The comparison including all groups, in turn, revealed the *miR-20a + miR-26a* pair as the most stable housekeeping gene selection (Table 2).

To evaluate the effects of the best candidate housekeeping genes determined by the different algorithms, the expression levels of the top three candidate miRNAs (*miR-20a*, *miR-25* and *miR-26a*) were normalized by each other (Fig. 3). *miR-20a*, *miR-25* and *miR-26a* did not differ between the groups ranking for the best ones when normalized by each other (Fig. 3): All comparisons showed no statistically differential expression, but when *miR-20* and *miR-26* were used as housekeeping genes we observed a more cohesive distribution and equivalent expression in the CY vs NC, HT vs WT and SC vs CO group pairs. These results indicate that *miR-20* and *miR-26a* are the most suitable genes to be used together as housekeeping genes in the assessed experimental scenario.

**Determination of the suitable number of housekeeping genes.** After ranking the candidate housekeeping genes according to their stability, a second approach was applied to determine the optimal number of reference genes to be used in each dataset. This analysis was performed using the GenEx software package. The optimal number of reference genes was calculated using the Acc.SD for the five candidate housekeeping genes (Fig. 4). Estimating the use of the ideal number of genes for normalization, one gene seems to be the best number in HT (*miR-26a*), WT (*miR-20a*), SC (*miR-25*), and CO (*miR-26a*) groups. The analysis showed that two (*miR-20a + miR-26a*) is the optimal number of references to be considered for normalization of miRNAs gene expression when using all the samples and the NC group. Three (*miR-20a + miR-25 + miR-26a*), in turn, was the optimal number of housekeeping genes to be applied to the CY group. We did not observe difference among the numbers of housekeeping genes to be used for the CO group.

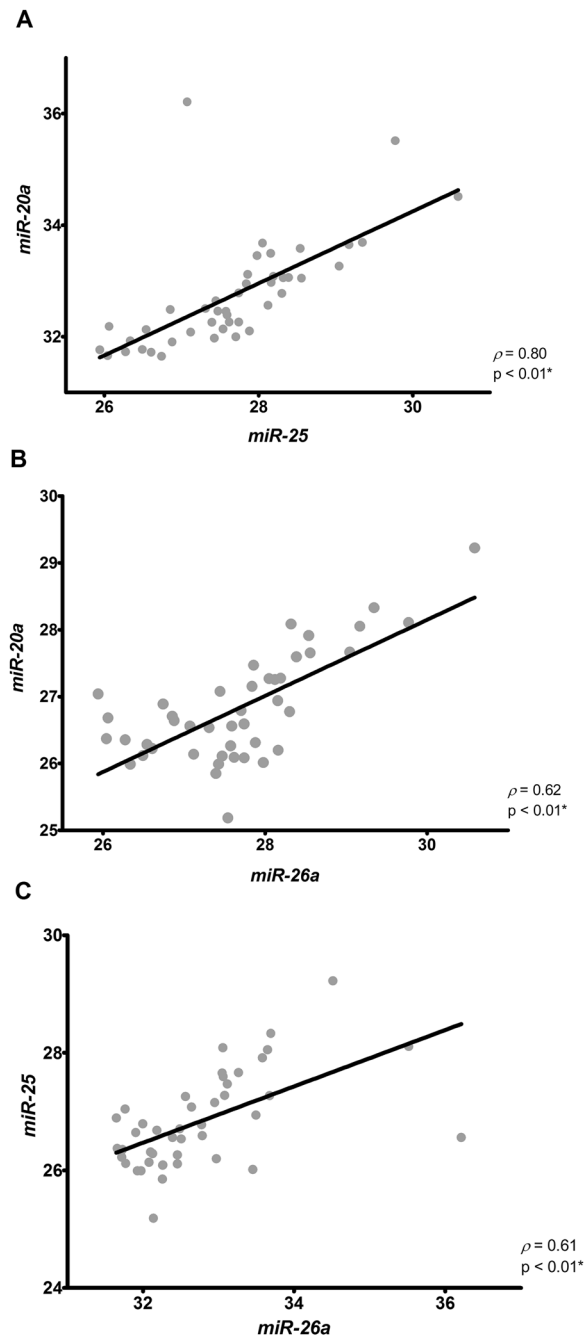


**Figure 3.** The  $\Delta C_t$  values of *miR-20a* (A–C), *miR-25* (D–F) and *miR-26a* (G–I) candidate housekeeping genes normalized by different combinations of each other. A lower threshold value ( $C_t$ ) indicates a higher gene expression. The median values are expressed as horizontal lines, and the error bars represent interquartile range. CY, cystic; NC, non-cystic; HT, haploinsufficient; WT, wild-type; SC, severely cystic phenotype; CO, severely cystic phenotype controls. 20a, target expression normalized by *miR-20a*; 25, target expression normalized by *miR-25*; 26a, target expression normalized by *miR-26a*; 25 + 26a, target expression normalized by *miR-25* + *miR-26a*; 20a + 25, target expression normalized by *miR-20a* + *miR-25*; 20a + 26a, target expression normalized by *miR-20a* + *miR-26a*. \* $p < 0.008$  by Mann-Whitney test, followed by Bonferroni correction.



**Figure 4.** Optimal number of reference genes according to GenEx analyses. Accumulated standard deviation (Acc.SD) for the five candidate housekeeping genes in all groups of samples allows to estimate the ideal number of genes for normalization. Acc.SD was calculated with the GenEx software package. Lower values of Acc.SD indicate the optimal number of reference genes. CY, cystic; NC, non-cystic; HT, *Pkd1*-haploinsufficient; WT, wild-type; SC, severely cystic phenotype; CO, severely cystic phenotype controls.

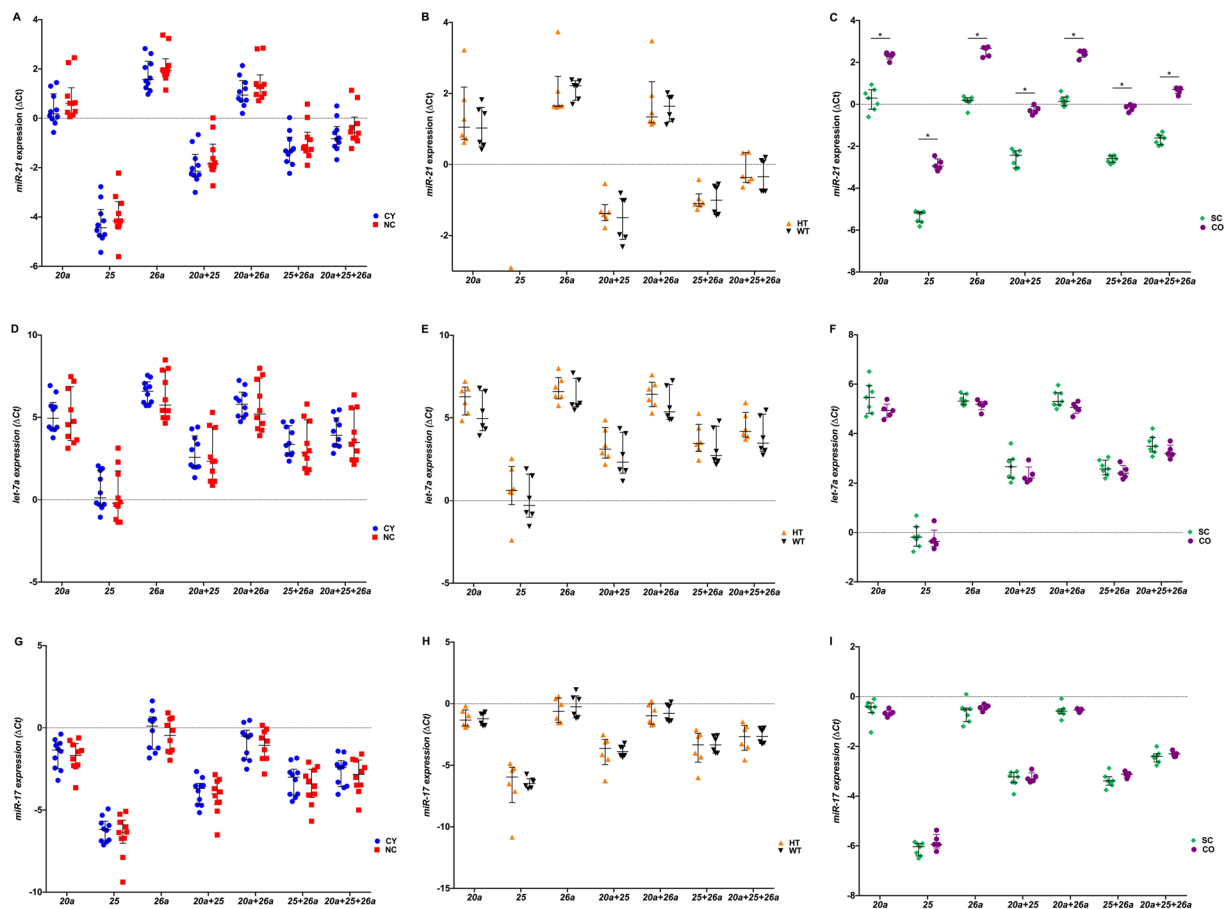
**Correlation between the top three candidate housekeeping miRNAs expression.** Correlation analyses were performed using the miRNAs expression data from all evaluated kidney samples. The expression level of the three best candidate housekeeping genes demonstrated a strong correlation between *miR-20a* and *miR-25* ( $\rho = 0.80$ ,  $p < 0.001$ , Fig. 5). Additionally, a moderate correlation was observed between *miR-20a* and *miR-26a* ( $\rho = 0.62$ ,  $p < 0.001$ , Fig. 5), as well as between *miR-25* and *miR-26a* ( $\rho = 0.61$ ,  $p < 0.001$ , Fig. 5). These results suggest that, in addition to *miR-25* and *miR-26a* having shown a moderate correlation, they are still correlated in all the samples herein evaluated and can be used together as suitable housekeeping genes. Our results



**Figure 5.** Correlation matrix between the expression of *miR-20a* and *miR-25* (A), *miR-20a* and *miR-26a* (B), and between *miR-25* and *miR-26a* (C) candidate housekeeping genes.  $\rho$ : Spearman's rank correlation coefficient. \* $p < 0.05$ .

showed that CT data dispersion of *miR-20a* and *miR-26a* increases substantially after normalization by *U6*. These results, were represented in Supplementary Fig. 3.

**Validation of the best candidate housekeeping genes for normalizing *miR-17*, *miR-21* and *let-7a* target genes.** In order to validate the three best candidate housekeeping genes stability, the relative expression of *miR-17*, *miR-21* and *let-7a* target genes was assessed using different combinations of *miR-20a*, *miR-25* and *miR-26a* (Fig. 6). The expression levels of *miR-21* were consistent with upregulation in SC compared to its control CO when normalized by different combinations of the best candidate housekeeping genes (Fig. 6C), trends of upregulation was observed in the CY and HT groups when compared with their respective controls (Fig. 6A,B). Trends of *let-7a* downregulation in CY, HT and SC compared to its control kidneys, was observed when their expression was normalized by different combinations of *miR-20a*, *miR-25* and *miR-26a* (Fig. 6D,F).



**Figure 6.** The  $\Delta C_t$  values of *miR-21* (A–C), *let-7a* (D–F) and *miR-17* (G–I) target genes normalized by different combinations of the three best candidate housekeeping genes (*miR-20a*, *miR-25* and *miR-26a*). A lower threshold value ( $C_t$ ) indicates a higher gene expression. The median values are represented as horizontal lines, and the error bars represent interquartile range. CY, cystic; NC, non-cystic; HT, haploinsufficient; WT, wild-type; SC, severely cystic phenotype; CO, severely cystic phenotype controls. *20a*, target expression normalized by *miR-20a*; *25*, target expression normalized by *miR-25*; *26a*, target expression normalized by *miR-26a*; *25 + 26a*, target expression normalized by *miR-25 + miR-26a*; *20a + 25*, target expression normalized by *miR-20a + miR-25*; *20a + 26a*, target expression normalized by *miR-20a + miR-26a*, target expression normalized by *miR-20a + miR-25 + miR-26a*. \* $p < 0.008$  by Mann-Whitney test, followed by Bonferroni correction.

Moreover, *miR-17* did not present expression statistically different in the CY vs NC, HT vs WT and SC vs CO groups (Fig. 6G,I).

When *U6* was used as housekeeping gene, the target genes did not differ between CY vs NC and HT vs WT, except for *miR-21* that showed downregulation in CY vs CO, HT vs CO and SC vs CO (Supplementary Fig. 2). These results also suggest that the SC group and its respective control CO have an expression profile slightly different from the other studied ADPKD-related mouse models. Taken together, the use of *miR-25* and *miR-26a* showed to be the most suitable pair of housekeeping genes among all considered sample groups.

To evaluate whether the severity of renal cystic phenotype was a significant modifier of expression of potential housekeeping genes, we investigated potential correlations between the expression levels of *miR-20a*, *miR-25* and *miR-26a* and kidney weight/body weight (KW/BW) in the CY group. The expression levels of *miR-20a*, *miR-25* and *miR-26a* were not correlated with the KW/BW ( $\rho = -0.52$ ,  $p < 0.18$ , Supplementary Fig. 3A;  $\rho = -0.24$ ,  $p < 0.57$ , Supplementary Fig. 3B; and  $\rho = -0.14$ ,  $p = 0.74$ , Supplementary Fig. 3C, respectively). The lack of correlation observed between the expression of the evaluated candidate housekeeping genes and kidney weight suggests that the cystic burden does not significantly influence the expression level of the analyzed housekeeping genes. These results provide support to use these miRNAs as controls in studies involving animals with different severities of renal cystic phenotype.

## Discussion

In recent years, miRNAs have emerged as key players in tightly-controlled biological processes such as proliferation<sup>37,41,42</sup>, apoptosis<sup>43–46</sup> and metabolism<sup>47</sup>. MiRNAs are known to be deregulated in numerous kidney diseases including ADPKD<sup>41–43</sup> and there is evidence to suggest that miRNA expression profiles may be more accurate in classifying kidney disease progression than mRNA expression profiles<sup>11</sup>. The relative miRNA expression

quantification is usually compared with a stably expressed housekeeping miRNA from the same sample<sup>33,48</sup>. The high sensitivity of RT-qPCR demands appropriate normalization to correct for non-biological variation and the use of housekeeping genes remains the most commonly used method<sup>49</sup>. However, there are no universally accepted housekeeping transcripts for miRNA data normalization by RT-qPCR analyses. In this scenario, the present study describes the first assessment of candidate housekeeping genes for the normalization of miRNA expression by RT-qPCR in kidney tissue of mouse models with distinct patterns of *Pkd1* deficiency. This is a significant problem to be solved, since different animal models orthologous to ADPKD have been and will keep being used to address distinct questions related to ADPKD pathogenesis and interventional studies.

Four candidate miRNA housekeeping genes (*miR-20a*, *miR-25*, *miR-26a* and *miR-191*) and the *U6* small nuclear RNA were selected from previous studies<sup>33,37–39,50–52</sup>. We used six algorithms (NormFinder, GeNorm, BestKeeper, DataAssist, the comparative  $\Delta\text{Ct}$  method and RefFinder) to identify the suitable housekeeping genes for relative quantitation of miRNA in fresh-frozen kidney tissues. Each algorithm ranked the best candidate reference genes. The software packages recommended *miR-20a*, *miR-25* and *miR-26a* as the most stable housekeeping genes among the tissue groups.

The expression of small nuclear (snRNAs) and nucleolar (snoRNAs) may exhibit tissue-specific and developmental regulation<sup>53</sup>, emphasizing the need for validation of commercially-available control assays. Although *U6* is the most commonly used gene to normalize miRNA RT-qPCR data<sup>37–39</sup>, we showed that it was the least stable candidate housekeeping gene in the evaluated samples. This result was also observed in other tissues and diseases<sup>35,36,54,55</sup>.

Since *miR-20a*, *miR-25* and *miR-26a* were the most suitable candidate genes, they were selected for normalization of *miR-17*, *miR-21* and *let-7a* target genes. The differences in *miR-21* expression detected between the tissue groups markedly varied depending on which single housekeeping gene was used for normalization. These results corroborate previous studies that associated *miR-21* with increase in cystogenesis and kidney size<sup>20,38</sup>. On the other hand, Gee *et al.* observed that *miR-21* presented variable expression in breast cancer and head and neck squamous cell carcinoma samples when normalized by *U6*, *U44*, *U48* and *U43*<sup>54</sup>. In renal cell carcinoma, *miR-28*, *miR-103* and *miR-106a* were proven to be more stably expressed than *U6*<sup>56</sup>. Therefore, these findings indicate that the use of unvalidated housekeeping genes may lead to inaccurate and unreliable results, and the use of snRNAs for normalization of expression of miRNAs might introduce bias in the associations between miRNA and the pathology or outcome<sup>50,54</sup>.

The current observed differences in *miR-17* expression detected between the tissue groups were at variance with many studies<sup>21,24,37,47</sup>, depending on which single housekeeping gene and animal models with *Pkd1* deficiency had been used for normalization. These findings draw attention to the potential effects of the housekeeping gene choice on the outcome of a study and demonstrates the need for validation of candidate housekeeping genes to generate reliable expression data. Even though our analysis suggested that the SC group and its respective control CO have an expression profile slightly different from the other *Pkd1*-deficient models, the use of *miR-25* and *miR-26a* showed to be the most suitable pair of housekeeping genes among all.

Finally, we assessed the correlation between the best candidate housekeeping genes and KW/BW ratio in the CY group of samples. The detected absence of correlation between the expression levels of the best ones and the KW/BW ratio in the CY group suggests that the level of cystic involvement does not lead to significant changes in housekeeping gene expression. Therefore, *miR-20a*, *miR-25* and *miR-26a* can be used as housekeeping genes in groups that include different stages of cystic burden in murine models orthologous to ADPKD.

One essential observation is the careful selection and validation of appropriate reference genes as a basis for normalizing the variability between samples in the corresponding study designs to each animal model used in the literature for experimental ADPKD research, amongst them *Ksp/cre* mice, *Pkhd1/cre* mice, *Kif3a*<sup>flx/flx</sup> mice, *Pkd2*<sup>flx/flx</sup> mice, *Pkhd1*<sup>-/-</sup> mice, *Hnf-1 $\beta$* <sup>flx/flx</sup> mice, *miR-17~92*<sup>flx/flx</sup>.

Based on current findings, we propose an appropriate selection of the best housekeeping genes for each comparison involving one or more mouse models, following the approach adopted in the present study. This procedure consisted in evaluating reference genes in three specific mouse models, namely *Pkd1*<sup>flx/flx</sup>; *Nestin*<sup>cre</sup>(CY) *Pkd1*<sup>+/-</sup> (HT) and *Pkd1*<sup>V/V</sup> (SC). The best housekeeping gene selection for each model/control group pair should be used when comparing a specific model with its respective control, while in comparisons including more than one model the best choice should be based on the whole analysis. Such specific approach enables a more proper and reliable normalization for future studies of dysregulated miRNAs within the ADPKD progression cascade. This task obviously implies reliable comparisons between expression data in different stages of the disease, including mild to severe PKD.

One limitation of the present study relies on the selection of only 5 candidate housekeeping genes for validation based on previous studies, but we recognize that more and other genes could also be suitable for accurate miRNA expression normalization by RT-qPCR in ADPKD in kidney tissue samples of orthologous mouse models and that the analysis of candidate housekeeping genes should be further tested using other models as well.

Previous studies suggested that, unlike mRNAs, the miRNA fraction present in FFPE tissues is relatively unaffected by the fixation process and that miRNAs extracted from these tissues may be accurately profiled using RT-qPCR<sup>57–59</sup>. In this context, the housekeeping genes identified in this study may also prove useful for miRNA RT-qPCR analysis of FFPE kidney tissues. Our findings will allow further analysis in kidney tissue of *Pkd1*-deficient mouse models gene expression to elucidate the role of different regulatory miRNAs in different scenarios of or related to ADPKD.

Normalizing to a suitable housekeeping gene, therefore, can eliminate differences due to sampling and quality of RNA and identify real changes in miRNA expression levels. In the present study, we analyzed a series of candidates and identified the most suitable housekeeping genes to be used for miRNA expression normalization in RT-qPCR studies in kidney tissues of mouse models orthologous to ADPKD and their respective controls.



The housekeeping gene spectrum and data generated by our work should therefore be employed in miRNA-related studies involving *Pkd1*-deficient mouse models. Among the genes currently used in this field, appropriate combinations of the *miR-20a*, *miR-25*, and *miR-26a* housekeeping genes offer increased accuracy and resolution in the quantitation of gene expression data, favoring the detection of smaller changes in miRNA expression than otherwise possible. The proper selection of the best housekeeping genes to be used in each of these scenarios should follow the guidelines specific to each comparison. It must be noted that other *Pkd1*-deficient models distinct from the ones analyzed in the current study may display slight expression differences of reference genes. Even in these cases, however, our guidelines would still be the best available housekeeping controls for such studies.

## Methods

**ADPKD mouse models.** We used three mouse models with distinct *Pkd1*-deficiency profiles in the current study, all generated and maintained on the C57BL/6 background. Only kidneys from male animals were analyzed in order to avoid potential gender-related experimental variability. Two models were evaluated at 10–12 weeks of age, including a renal cystic mouse (*Pkd1*<sup>flx/flx</sup>;*Nestin*<sup>cre</sup>, CY, n = 10) and its corresponding non-cystic control (*Pkd1*<sup>flx/flx</sup>, NC, n = 10), and a *Pkd1*-haploinsufficient mouse (*Pkd1*<sup>+/-</sup>, HT, n = 6) and its respective wild-type control (*Pkd1*<sup>+/+</sup>, WT, n = 6). The third model was assessed at 15 days of life due to its severely renal cystic phenotype (*Pkd1*<sup>VV</sup>, SC, n = 7) along with its wild-type control (CO, n = 5). These animals were genotyped using specific PCRs. The CY mouse is homozygous for a *Pkd1*-floxed allele (*Pkd1*<sup>flx</sup>) and displays a mosaic pattern of gene inactivation, induced by a Nestin-Cre transgene through excision of exons 2–4 (*Pkd1*<sup>flx/flx</sup>;*Nestin*<sup>cre</sup>)<sup>25,27–29</sup>. The HT model is heterozygous for a *Pkd1* null allele, characterized by early transcriptional interruption<sup>25,26</sup>, and develops no renal cysts by 12 weeks of age. The SC model is homozygous for the *Pkd1* knockin T3041V allele, which prevents the autoproteolytic cleavage of PC1 at the GPS site<sup>25,30</sup>. *Pkd1*<sup>VV</sup> animals have no gross phenotype by post-natal day (P) 6 but develop rapid and progressive distal nephron dilatation thereafter. This severe renal phenotype, that eventually leads to renal failure, is apparently responsible for the early mortality that occurs between the 2nd and 6th week<sup>30</sup>. Animals were fed *ad libitum* and housed at constant ambient temperature in a 12-hour light cycle.

Animal procedures were approved by the Internal Biosafety Commission of Genetically Modified Organisms of the University of São Paulo School of Medicine and by the Universidade Federal de São Paulo (UNIFESP) Research Ethics Committees.

SC (*Pkd1*<sup>VV</sup>) and its wild-type control (CO) animals were euthanized by adopting cervical dislocation and the other animal groups were euthanized with intraperitoneal thiopental (0.4 mg/g of body weight); their kidneys were appropriately harvested for RT-qPCR analyses. All experiments were conducted in accordance with international standards of animal care and experimentation. Both kidneys were collected and stored at –80 °C for further use.

**Housekeeping genes.** We selected these genes based on the observation that they were among the miRNAs with the lowest variances in a previous study of ours (TLDA Taqman Array, unpublished data) and that miRNA expression studies showed minor evidence of differential expression<sup>10,20,37–39</sup>. These genes not only behaved stably but also presented expression stability in transcriptomic analysis. Of note, *U6* and *miRNA191* presented stable expression, also documented in a previous report<sup>33</sup>. All these five miRNAs are constitutively expressed in kidney tissue of mouse models orthologous to ADPKD, have independent cellular functions, and are assumed not to be co-regulated.

**RNA extraction.** Renal tissue lysis was performed using zirconia beads (Interphase, USA) and the Precellys (BioAmerica, USA) homogenizer. TRIzol (Life Technologies, USA) was employed for total RNA extraction according to the manufacturer's protocol. The RNA quantity and purity were determined using the NanoVue spectrophotometer (GE Healthcare Life Sciences, USA) and analyzed with the Agilent 2100 Bioanalyzer 6000 Nanochip (Agilent Technologies Inc., Waldbronn, BW, Germany). Total RNA was stored at –80 °C until further use.

**cdNA preparation and RT-qPCR.** Complementary DNA (cdNA) was synthesized from 1 µg of total RNA with an oligonucleotide pool for each evaluated gene (Supplementary Table 1), using the TaqMan<sup>®</sup> microRNA reverse transcription kit (Life Technologies) according to the manufacturer's instructions. Detection of the expression range of the evaluated genes was achieved by using TaqMan<sup>®</sup> assays and the QuantStudio<sup>®</sup> 7 Flex real-time PCR system (Applied Biosystems, USA). The used primers are shown in Supplementary Table 1. *miR-17*, *miR-21* and *let-7a* were employed as target genes<sup>20–24</sup>. All reactions were run in triplicate. The expression of the candidate housekeeping genes is represented by the original cycle threshold (Ct) value.

**Analysis of housekeeping gene expression stability.** Cts of RT-qPCR were manually settled as 0.02 while the mean Ct values of the three technical replicates were imported to six algorithms (Fig. 1): NormFinder (version 0.953; <https://moma.dk/normfinder-software>)<sup>60</sup>, GeNorm (<https://genorm.cmgg.be/>)<sup>40</sup>, BestKeeper (version 1.0; <https://www.gene-quantification.de/bestkeeper.html>)<sup>61</sup>, DataAssist<sup>™</sup> (version 3.01; <https://www.thermofisher.com/br/en/home/technical-resources/software-downloads/dataassist-software.html>), the comparative  $\Delta$ Ct method<sup>62</sup> and RefFinder (<https://www.heartcure.com.au/reffinder/>)<sup>63</sup>, following the authors' recommendations. These software packages were used to determine the relative expression stability of the candidate housekeeping genes and to generate a ranking for the best ones.

The optimal number of reference genes was selected using the GenEx software package. GenEx is a software for the processing and analysis of RT-qPCR data and provides methods to select and validate housekeeping genes, classify samples, group genes and monitor time-dependent processes. GenEx calculates Accumulated standard deviation (Acc.SD); lower Acc.SD values, in turn, indicate the optimal number of reference genes (<https://www.biomcc.com/genex-software.html>).

**Statistical analysis.** The Shapiro-Wilk test showed that the Ct values of the candidate housekeeping genes were not normally distributed. Therefore, results are shown in median and interquartile range (IQ). Initially, the Kruskal-Wallis test was used to evaluate possible expression differences of *miR-17*, *miR-20a*, *miR-21*, *miR-25*, *miR-26a*, *miR-191*, *U6*, and *let-7a* among the sample groups. When the Kruskal-Wallis test showed significance ( $p < 0,05$ ), the Mann-Whitney test, followed by Bonferroni correction, was used for all comparisons between two groups. Lastly, the Spearman correlation test was used to verify potential correlations among the expression levels of *miR-20a*, *miR-25* and *miR-26a*, and between the expression levels of these miRNAs in all kidney samples evaluated herein and kidney weight. A value between 0.30–0.50 was determined as a weak correlation, 0.50–0.70 as moderate, 0.70–0.90 as strong, and 0.90–1.00 as a very strong correlation<sup>64</sup>.

## Data availability

All data including supporting datasets are made available as main figures or Supplementary Information Files.

Received: 15 August 2019; Accepted: 21 December 2019;

Published online: 14 January 2020

## References

1. Grantham, J. J., Mulamalla, S. & Swenson-Fields, K. I. Why kidneys fail in autosomal dominant polycystic kidney disease. *Nat. reviews. Nephrology* **7**, 556–566, <https://doi.org/10.1038/nrneph.2011.109> (2011).
2. Grantham, J. J. & Torres, V. E. The importance of total kidney volume in evaluating progression of polycystic kidney disease. *Nat. reviews. Nephrology* **12**, 667–677, <https://doi.org/10.1038/nrneph.2016.135> (2016).
3. Corne-Le Gall, E., Torres, V. E. & Harris, P. C. Genetic Complexity of Autosomal Dominant Polycystic Kidney and Liver Diseases. *J. Am. Soc. Nephrol.* **29**, 13–23, <https://doi.org/10.1681/ASN.2017050483> (2018).
4. Jiang, S. T. *et al.* Defining a link with autosomal-dominant polycystic kidney disease in mice with congenitally low expression of Pkd1. *Am. J. Pathol.* **168**, 205–220, <https://doi.org/10.2353/ajpath.2006.050342> (2006).
5. Lantinga-van Leeuwen, I. S. *et al.* Lowering of Pkd1 expression is sufficient to cause polycystic kidney disease. *Hum. Mol. Genet.* **13**, 3069–3077, <https://doi.org/10.1093/hmg/ddh336> (2004).
6. Rossetti, S. *et al.* Incompletely penetrant PKD1 alleles suggest a role for gene dosage in cyst initiation in polycystic kidney disease. *Kidney Int.* **75**, 848–855, <https://doi.org/10.1038/ki.2008.686> (2009).
7. Paterson, A. D. *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.* **16**, 755–762, <https://doi.org/10.1681/asn.2004090758> (2005).
8. Tsai, Y. C. *et al.* Safe Nanocomposite-Mediated Efficient Delivery of MicroRNA Plasmids for Autosomal Dominant Polycystic Kidney Disease (ADPKD) Therapy. *Adv. Healthc. Mater.* **8**, e1801358, <https://doi.org/10.1002/adhm.201801358> (2019).
9. Kocyyigit, I. *et al.* Serum micro-rna profiles in patients with autosomal dominant polycystic kidney disease according to hypertension and renal function. *BMC Nephrol.* **18**, 179, <https://doi.org/10.1186/s12882-017-0600-z> (2017).
10. Kim, D. Y. *et al.* Impact of miR-192 and miR-194 on cyst enlargement through EMT in autosomal dominant polycystic kidney disease. *FASEB J.* **33**, 2870–2884, <https://doi.org/10.1096/fj.201800563RR> (2019).
11. Wang, E. *et al.* Progressive renal distortion by multiple cysts in transgenic mice expressing artificial microRNAs against Pkd1. *J. Pathol.* **222**, 238–248, <https://doi.org/10.1002/path.2765> (2010).
12. Dweep, H., Sticht, C., Kharkar, A., Pandey, P. & Gretz, N. Parallel analysis of mRNA and microRNA microarray profiles to explore functional regulatory patterns in polycystic kidney disease: using PKD/Mhm rat model. *PLoS One* **8**, e53780, <https://doi.org/10.1371/journal.pone.0053780> (2013).
13. Pandey, P. *et al.* Microarray-based approach identifies microRNAs and their target functional patterns in polycystic kidney disease. *BMC Genomics* **9**, 624, <https://doi.org/10.1186/1471-2164-9-624> (2008).
14. Pandey, P., Qin, S., Ho, J., Zhou, J. & Kreidberg, J. A. Systems biology approach to identify transcriptome reprogramming and candidate microRNA targets during the progression of polycystic kidney disease. *BMC Syst. Biol.* **5**, 56, <https://doi.org/10.1186/1752-0509-5-56> (2011).
15. Park, E. Y., Woo, Y. M. & Park, J. H. Polycystic kidney disease and therapeutic approaches. *BMB Rep.* **44**, 359–368, <https://doi.org/10.5483/BMBRep.2011.44.6.359> (2011).
16. Tan, Y. C., Blumenfeld, J. & Rennert, H. Autosomal dominant polycystic kidney disease: genetics, mutations and microRNAs. *Biochim. Biophys. Acta* **1812**, 1202–1212, <https://doi.org/10.1016/j.bbdis.2011.03.002> (2011).
17. Chatterjee, S., Verma, S. P. & Pandey, P. Profiling conserved biological pathways in Autosomal Dominant Polycystic Kidney Disorder (ADPKD) to elucidate key transcriptomic alterations regulating cystogenesis: A cross-species meta-analysis approach. *Gene* **627**, 434–450, <https://doi.org/10.1016/j.gene.2017.06.059> (2017).
18. Chu, A. S. & Friedman, J. R. A role for microRNA in cystic liver and kidney diseases. *J. Clin. Invest.* **118**, 3585–3587, <https://doi.org/10.1172/jci36870> (2008).
19. Sun, H. *et al.* MicroRNA-17 post-transcriptionally regulates polycystic kidney disease-2 gene and promotes cell proliferation. *Mol. Biol. Rep.* **37**, 2951–2958, <https://doi.org/10.1007/s11033-009-9861-3> (2010).
20. Lakhia, R. *et al.* MicroRNA-21 Aggravates Cyst Growth in a Model of Polycystic Kidney Disease. *J. Am. Soc. Nephrol.* **27**, 2319–2330, <https://doi.org/10.1681/asn.2015060634> (2016).
21. Yheskel, M. & Patel, V. Therapeutic microRNAs in polycystic kidney disease. *Curr. Opin. Nephrol. Hypertens.* **26**, 282–289, <https://doi.org/10.1097/mnh.0000000000000333> (2017).
22. Kato, M., Arce, L. & Natarajan, R. MicroRNAs and their role in progressive kidney diseases. *Clin. J. Am. Soc. Nephrol.* **4**, 1255–1266, <https://doi.org/10.2215/cjn.00520109> (2009).
23. Ben-Dov, I. Z. *et al.* Urine microRNA as potential biomarkers of autosomal dominant polycystic kidney disease progression: description of miRNA profiles at baseline. *PLoS One* **9**, e86856, <https://doi.org/10.1371/journal.pone.0086856> (2014).
24. Patel, V. *et al.* miR-17–92 miRNA cluster promotes kidney cyst growth in polycystic kidney disease. *Proc. Natl Acad. Sci. USA* **110**, 10765–10770, <https://doi.org/10.1073/pnas.1301693110> (2013).
25. Balbo, B. E. *et al.* Cardiac dysfunction in Pkd1-deficient mice with phenotype rescue by galectin-3 knockout. *Kidney Int.* **90**, 580–597, <https://doi.org/10.1016/j.kint.2016.04.028> (2016).
26. Bastos, A. P. *et al.* Pkd1 haploinsufficiency increases renal damage and induces microcyst formation following ischemia/reperfusion. *J. Am. Soc. Nephrol.* **20**, 2389–2402, <https://doi.org/10.1681/asn.2008040435> (2009).
27. Fonseca, J. M. *et al.* Renal cyst growth is the main determinant for hypertension and concentrating deficit in Pkd1-deficient mice. *Kidney Int.* **85**, 1137–1150, <https://doi.org/10.1038/ki.2013.501> (2014).
28. Piontek, K. B. *et al.* A functional floxed allele of Pkd1 that can be conditionally inactivated *in vivo*. *J. Am. Soc. Nephrol.* **15**, 3035–3043, <https://doi.org/10.1097/01.asn.0000144204.01352.86> (2004).

29. Shillingford, J. M., Piontek, K. B., Germino, G. G. & Weimbs, T. Rapamycin ameliorates PKD resulting from conditional inactivation of Pkd1. *J. Am. Soc. Nephrol.* **21**, 489–497, <https://doi.org/10.1681/asn.2009040421> (2010).
30. Yu, S. *et al.* Essential role of cleavage of Polycystin-1 at G protein-coupled receptor proteolytic site for kidney tubular structure. *Proc. Natl Acad. Sci. USA* **104**, 18688–18693, <https://doi.org/10.1073/pnas.0708217104> (2007).
31. Yan, J., Zhang, N., Qi, C., Liu, X. & Shangguan, D. One-step real time RT-PCR for detection of microRNAs. *Talanta* **110**, 190–195, <https://doi.org/10.1016/j.talanta.2013.02.028> (2013).
32. Mestdagh, P. *et al.* Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. *Nat. Methods* **11**, 809–815, <https://doi.org/10.1038/nmeth.3014> (2014).
33. Schwarzenbach, H., da Silva, A. M., Calin, G. & Pantel, K. Data Normalization Strategies for MicroRNA Quantification. *Clin. Chem.* **61**, 1333–1342, <https://doi.org/10.1373/clinchem.2015.239459> (2015).
34. Li, J. Y., Yong, T. Y., Michael, M. Z. & Gleadle, J. M. Review: The role of microRNAs in kidney disease. *Nephrology (Carlton)* **15**, 599–608, <https://doi.org/10.1111/j.1440-1797.2010.01363.x> (2010).
35. Aboudehen, K. *et al.* Long noncoding RNA. *J. Biol. Chem.* **293**, 9388–9398, <https://doi.org/10.1074/jbc.RA118.001723> (2018).
36. Morata-Tarifa, C. *et al.* Validation of suitable normalizers for miR expression patterns analysis covering tumour heterogeneity. *Sci. Rep.* **7**, 39782, <https://doi.org/10.1038/srep39782> (2017).
37. Shin, Y., Kim, D. Y., Ko, J. Y., Woo, Y. M. & Park, J. H. Regulation of KLF12 by microRNA-20b and microRNA-106a in cystogenesis. *FASEB J.* **32**, 3574–3582, <https://doi.org/10.1096/fj.201700923R> (2018).
38. Woo, Y. M. *et al.* Profiling of miRNAs and target genes related to cystogenesis in ADPKD mouse models. *Sci. Rep.* **7**, 14151, <https://doi.org/10.1038/s41598-017-14083-8> (2017).
39. Patel, V. *et al.* MicroRNAs regulate renal tubule maturation through modulation of Pkd1. *J. Am. Soc. Nephrol.* **23**, 1941–1948, <https://doi.org/10.1681/asn.2012030321> (2012).
40. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, RESEARCH0034, <https://doi.org/10.1186/gb-2002-3-7-research0034> (2002).
41. Yheskel, M., Lakhia, R., Cobo-Stark, P., Flaten, A. & Patel, V. Anti-microRNA screen uncovers miR-17 family within miR-17~92 cluster as the primary driver of kidney cyst growth. *Sci. Rep.* **9**, 1920, <https://doi.org/10.1038/s41598-019-38566-y> (2019).
42. O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839–843, <https://doi.org/10.1038/nature03677> (2005).
43. de Stephanis, L. *et al.* MicroRNA501-5p induces p53 proteasome degradation through the activation of the mTOR/MDM2 pathway in ADPKD cells. *J. Cell Physiol.* **233**, 6911–6924, <https://doi.org/10.1002/jcp.26473> (2018).
44. Duplomb, L. *et al.* A constitutive BCL2 down-regulation aggravates the phenotype of PKD1-mutant-induced polycystic kidney disease. *Hum. Mol. Genet.* **26**, 4680–4688, <https://doi.org/10.1093/hmg/ddx349> (2017).
45. Chan, J. A., Krichevsky, A. M. & Kosik, K. S. MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res.* **65**, 6029–6033, <https://doi.org/10.1158/0008-5472.can-05-0137> (2005).
46. Cimmino, A. *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl Acad. Sci. USA* **102**, 13944–13949, <https://doi.org/10.1073/pnas.0506654102> (2005).
47. Hajarnis, S. *et al.* microRNA-17 family promotes polycystic kidney disease progression through modulation of mitochondrial metabolism. *Nat. Commun.* **8**, 14395, <https://doi.org/10.1038/ncomms14395> (2017).
48. Meyer, S. U., Pfaffl, M. W. & Ulbrich, S. E. Normalization strategies for microRNA profiling experiments: a 'normal' way to a hidden layer of complexity? *Biotechnol. Lett.* **32**, 1777–1788, <https://doi.org/10.1007/s10529-010-0380-z> (2010).
49. Bustin, S. A. *et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **55**, 611–622, <https://doi.org/10.1373/clinchem.2008.112797> (2009).
50. Peltier, H. J. & Latham, G. J. Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* **14**, 844–852, <https://doi.org/10.1261/rna.939908> (2008).
51. Shen, Y. *et al.* Identification of miR-23a as a novel microRNA normalizer for relative quantification in human uterine cervical tissues. *Exp. & Mol. Med.* **43**, 358–366, <https://doi.org/10.3858/emmm.2011.43.6.039> (2011).
52. Li, Y. *et al.* Identification of endogenous controls for analyzing serum exosomal miRNA in patients with hepatitis B or hepatocellular carcinoma. *Dis. markers* **2015**, 893594, <https://doi.org/10.1155/2015/893594> (2015).
53. Tu, C. *et al.* Evaluating the potential of housekeeping genes, rRNAs, snRNAs, microRNAs and circRNAs as reference genes for the estimation of PMI. *Forensic Sci. Med. Pathol.* **14**, 194–201, <https://doi.org/10.1007/s12024-018-9973-y> (2018).
54. Gee, H. E. *et al.* The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. *Br. J. Cancer* **104**, 1168–1177, <https://doi.org/10.1038/sj.bjc.6606076> (2011).
55. Wang, J. *et al.* Altered serum microRNAs as novel diagnostic biomarkers for atypical coronary artery disease. *PLoS One* **9**, e107012, <https://doi.org/10.1371/journal.pone.0107012> (2014).
56. Wotschovsky, Z. *et al.* Reference genes for the relative quantification of microRNAs in renal cell carcinomas and their metastases. *Anal. Biochem.* **417**, 233–241, <https://doi.org/10.1016/j.ab.2011.06.009> (2011).
57. de Biase, D. *et al.* miRNAs expression analysis in paired fresh/frozen and dissected formalin fixed and paraffin embedded glioblastoma using real-time pCR. *PLoS One* **7**, e35596, <https://doi.org/10.1371/journal.pone.0035596> (2012).
58. Klopffleisch, R., Weiss, A. T. & Gruber, A. D. Excavation of a buried treasure—DNA, mRNA, miRNA and protein analysis in formalin fixed, paraffin embedded tissues. *Histol. Histopathol.* **26**, 797–810, <https://doi.org/10.14670/hh-26.797> (2011).
59. Hajarnis, S. *et al.* Suppression of microRNA Activity in Kidney Collecting Ducts Induces Partial Loss of Epithelial Phenotype and Renal Fibrosis. *J. Am. Soc. Nephrol.* **29**, 518–531, <https://doi.org/10.1681/asn.2017030334> (2018).
60. Andersen, C. L., Jensen, J. L. & Ørntoft, T. F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* **64**, 5245–5250, <https://doi.org/10.1158/0008-5472.can-04-0496> (2004).
61. Pfaffl, M. W., Tichopad, A., Prgomet, C. & Neuvians, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**, 509–515 (2004).
62. Silver, N., Best, S., Jiang, J. & Thein, S. L. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* **7**, 33, <https://doi.org/10.1186/1471-2199-7-33> (2006).
63. Xie, F., Xiao, P., Chen, D., Xu, L. & Zhang, B. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant. Mol. Biol.* <https://doi.org/10.1007/s11103-012-9885-2> (2012).
64. Mukaka, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* **24**, 69–71 (2012).

## Acknowledgements

This study received funding from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) 309045/2018-5 (I.P.H.); and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) 2018/09135-0 (J.J.M.), 2011/21593-4 (R.M.) 2015/23345-9 (M.A.B. and I.P.H.), 2015/17152-3 (L.F.O.).

### Author contributions

J.J.M., A.C.A. and I.P.H. designed the study, conceptualization, methodology, validation, formal analysis and investigation. A.G.A., F.M.F. and R.G.M. were responsible for sample collection. J.J.M., A.C.A. and M.S.O. performed the experiments. J.J.M., A.C.A., L.F.O., I.P.H. and A.G.A. analyzed the data. J.J.M., A.C.A., L.F.O. and I.P.H. wrote the first draft of the manuscript. M.A.B. provided study material. All authors gave intellectual input and reviewed the last version of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-019-57112-4>.

**Correspondence** and requests for materials should be addressed to I.P.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020