scientific reports

OPEN

Determination of reference genes as a quantitative standard for gene expression analysis in mouse mesangial cells stimulated with TGF-β

Bruno Aristides dos Santos Bronel, Ana Carolina Anauate, Edgar Maquigussa, Mirian Aparecida Boim & Antônio da Silva Novaes✉

Reverse transcription-quantitative polymerase chain reaction (RT-PCR) is the gold standard technique for gene expression analysis, but the choice of quantitative reference genes (housekeeping genes, HKG) remains challenging. Identify the best HKG is essential for estimating the expression level of target genes. Therefore, the aim of this study was to determine the best HKG for an in vitro model with mouse mesangial cells (MMCs) stimulated with 5 ng/mL of TGF-β. Five candidates HKG were selected: *Actb*, *Hprt*, *Gapdh*, *18S* and *Ppia*. After quantitative expression, the best combination of these genes was analyzed in silico using six software programs. To validate the results, the best genes were used to normalize the expression levels of *fibronectin*, *vimentin* and *α-SMA*. In silico analysis revealed that *Ppia*, *Gapdh* and *18S* were the most stable genes between the groups. GenEX software and Spearman's correlation determined *Ppia* and *Gapdh* as the best HKG pair, and validation of the HKG by normalizing *fibronectin*, *vimentin* and *α-SMA* were consistent with results from the literature. Our results established the combination of *Ppia* and *Gapdh* as the best HKG pair for gene expression analysis by RT-PCR in this in vitro model using MMCs treated with TGF-β.

Transforming growth factor-β (TGF-β) is a major renal profibrogenic cytokine and plays a critical role in mesangial dysfunction in many pathophysiological conditions characterized by excessive accumulation of extracellular matrix (ECM) proteins, mesangial cell (MCs) hypertrophy, and proliferation[1,2]. The interaction of TGF-β and its receptors forms a heterodimeric complex, which is translocated into the nucleus and regulates transcription of target genes, such as *fibronectin*, *vimentin* and *α-smooth muscle actin* (*α-SMA*)[1,3–5].

The reverse transcription-quantitative polymerase chain reaction (RT-PCR) is the gold standard method to identify changes in mRNA expression levels[6–8]. Considering the many steps of RT-PCR, and that several factors can influence expression levels, the normalization of target genes is crucial for accurate gene expression quantification. Currently, the most accepted method of target gene expression normalization by RT-PCR technique is through quantification of very stable endogenous housekeeping genes (HKG)[9,10].

HKG or reference genes, are genes required for maintenance of basal cellular functions[11,12] and the ideal reference genes are expected to be expressed in all cell types and should show minimal variation in the expression, regardless of cell cycle state, developmental stage, external stimuli and physiological condition[10,12,13]. Examples of the most used reference genes include *Actin beta* and *Gapdh*[14]. Although the use of HKG is the most accurate method for normalizing mRNA expression levels, it is well known that the expression levels of even the most stable HKG can change depending on cell type and experimental conditions and design[8,15].

Despite the growing number of studies investigating the reference genes for renal disease models[10,16,17], to our knowledge there are no detailed reports selecting the most stable gene recommended for the frequently used in vitro model of TGF-β-induced fibrogenesis in MCs. This limitation may lead to non-reproducible data. Therefore, the present study aimed to evaluate the performance of five frequently used reference genes (*Actb, Hprt, Gapdh, 18S* and *Ppia)* in renal models[18–29] and to identify the most stable ones and the optimal number of genes for normalization the expressions of target genes by RT-PCR in MCs treated with TGF-β model.

Renal Division, Department of Medicine, Universidade Federal de São Paulo, Rua Pedro de Toledo, 781, São Paulo, SP 04039-032, Brazil. ✉email: antonio.novaes52@gmail.com
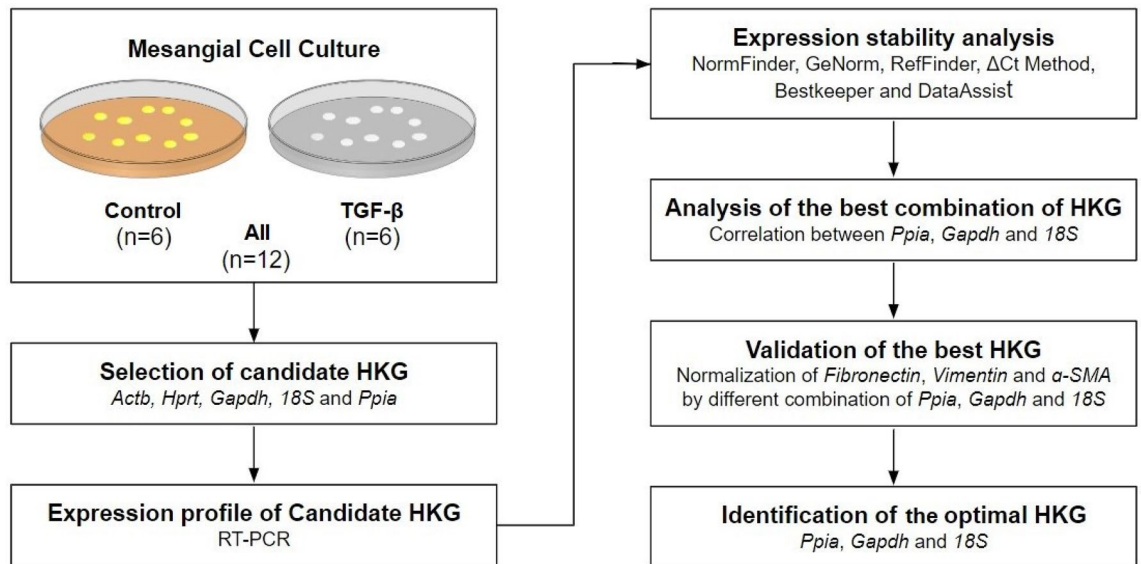
Figure 1. Workflow diagram illustrating the strategy for identification of housekeeping normalizer for RT-PCR. *Actb, Hprt, Gapdh, 18S* and *Ppia* housekeeping genes (HKG) selected from the literature for comparison.

Results

Candidate housekeeping genes.   First, we followed a stepwise strategy to identify the best HKG expression by RT-PCR analysis. The workflow diagram is shown in Fig. 1. The samples were classified into three groups: (1) control cells (n = 6); (2) cells treated with TGF-β (n = 6); and (3) All, which includes all cells (n = 12). The Ct values of the five candidate HKG ranged between 23.511–9.387 and are expressed as [median (interquartile range)]. Ct values are inversely proportional to gene expression, and the mean highest Ct value among the candidate genes was achieved by *Hprt* [20.876 (2.05)], indicating the least expressed gene. In contrast, the lowest Ct value was obtained for *18S* [10.232 (0.50)], indicating the highest expression among the candidate genes. *Gapdh* [17.948 (2.41)], followed by *Actb* [15.986 (2.44)] and *Ppia* [15.514 (1.54)], showed moderate expression. The median Ct values of triplicate reactions according to each gene is shown in Fig. 2A. Furthermore, no statistically significant differences were found between the control group and the TGF-β-treated group. Thus, these data suggest that the endogenous genes selected showed good stability and that their expression did not change when treated with TGF-β.

Stability analysis of housekeeping genes.   Stability analysis of the five candidate HKG were determined using six software packages (Supplementary Table 1). Genes with the smallest stability value (SV) had the most stable expression. Following NormFinder criteria (SD < 0.5), only *Ppia* and *Gapdh* had a SD value below 0.5 in the All and control groups, while in the TGF-β-treated group, only *Ppia* respects this criterion (Fig. 2B and Supplementary Table 1). The GeNorm software defines an M value < 1.5; therefore, the genes with lower variability were: *Hprt*, *18S* and *Ppia* for All samples; *Gapdh*, *Ppia*, *Hprt* and *18S* for the control group; and *18S, Hprt* and *Ppia* for the TGF-β-treated group (Fig. 2C and Supplementary Table 1). According to Bestkeeper criteria, when considering all three experimental groups, only *18S* fits the parameters, although the CV exceeds 3.0 in the TGF-β group (Fig. 2D and Supplementary Table 1).

In the RefFinder and ΔCt method analysis, the most stable gene in the All group was *Ppia*, while in the control group it was *Gapdh*. However, in the TGF-β group, the most stable genes were *18S* and *Ppia* according to Ref-Finder and ΔCt method analysis, respectively (Fig. 2E–F and Supplementary Table 1). The evaluation of the most stable gene on DataAssist software identified that *Ppia* as the best HKG in the All, control, and TGF-β-treated groups (Fig. 2G and Supplementary Table 1). In all software analysis, the less stable gene was *Actb* (Fig. 2B–G and Supplementary Table 1). Based on software analysis and a qualitative inspection of all ranks generated, the best HKG for the All group was *Ppia*, whereas it was *Gapdh* for the control group and *Ppia* or *18S* for the TGF-β-treated group (Fig. 2B–G and Supplementary Table 1).

Analysis of the best combination of housekeeping genes.   The GeNorm software recommends at least two genes for gene expression normalization, and the best combination of HKG for each group/software package is shown in Table 1. In the All group, the best pair of HKG is *Ppia + Gapdh* or *Ppia + 18S*; in the control group is *Gapdh + Ppia* and in the TGF-β group is *Ppia + 18S* (Table 1). To analyze the effects of the best candidate HKG, the expression levels of the top three (*Ppia*, *Gapdh*, and *18S*) were normalized by each other (Fig. 3). All comparisons showed no statistically differential expression, meaning that *Ppia*, *Gapdh*, and *18S* did not differ between the groups when normalized by each other (Fig. 3).
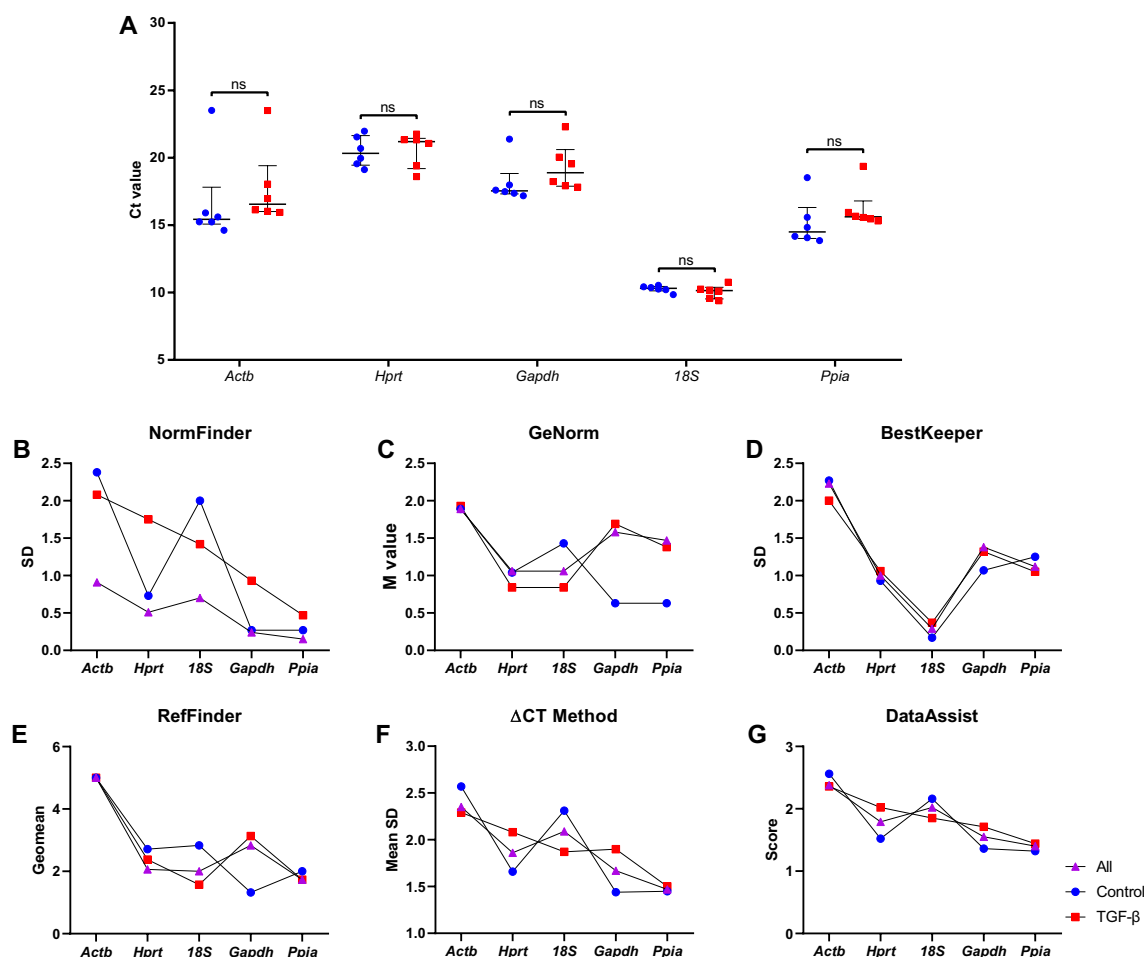
Figure 2. Ct values of five candidate housekeeping genes (**A**) and stability analysis of housekeeping genes by six different software (**B–G**). A lower cycle threshold (Ct) value indicates a higher gene expression (**A**). The median values are expressed as horizontal lines, and the error bars represent the interquartile range. The Ct values of *18S* were the lowest, indicating the highest expression levels. Best housekeeping gene for each group of samples yielded by software analysis (**B–G**). All, all samples. ns, non-significant.

Groups	NormFinder	GeNorm	RefFinder	ΔCt method	Bestkeeper	DataAssist	Best HKG Pair
All	*Ppia/Gapdh*	*Hprt/18S*	*Ppia/18S*	*Ppia/Gapdh*	*18S/Hprt*	*Ppia/Gapdh*	*Ppia/Gapdh* or *Ppia/18S*
Control	*Gapdh/Ppia*	*Gapdh/Ppia*	*Gapdh/Ppia*	*Gapdh/Ppia*	*18S/Hprt*	*Ppia/Gapdh*	*Gapdh/Ppia*
TGF-β	*Ppia/Gapdh*	*18S/Hprt*	*18S/Ppia*	*Ppia/18S*	*18S/Hprt*	*Ppia/Gapdh*	*Ppia/18S*

Table 1. The best combination of housekeeping genes for each group of samples yielded by software analysis. All, all samples. HKG, housekeeping genes.

Determination of the suitable number of housekeeping genes. After rating the candidate HKG by their stability values, the optimal number of candidate genes to be used in each dataset must be established. The Acc.SD results showed that one gene (*Ppia* or *18S*) is the optimal number of HKG for normalization of gene expression in TGF-β-treated samples (Fig. 4). Two genes are required for normalization in the All group (*Ppia + Gapdh* or *Ppia + 18S*) and in the control group (*Gapdh + Ppia*) (Fig. 4).

Correlation between the top three candidate housekeeping gene expressions. Correlation analysis were performed using the gene expression data from all samples. The expression levels of the three best candidate HKG showed a strong correlation between *Ppia* and *Gapdh* ($\rho = 0.804$, p = 0.002) (Fig. 5). Also, no statistically significant correlation was found between *Ppia* and *18S* ($\rho = -0.392$, p = 0.208) and *Gapdh* and *18S* ($\rho = -0.580$, p = 0.052) (Fig. 5). These results suggest that besides *Ppia* and *Gapdh* showed a strong correlation, they are correlated in all the samples and can be used together as suitable HKG.
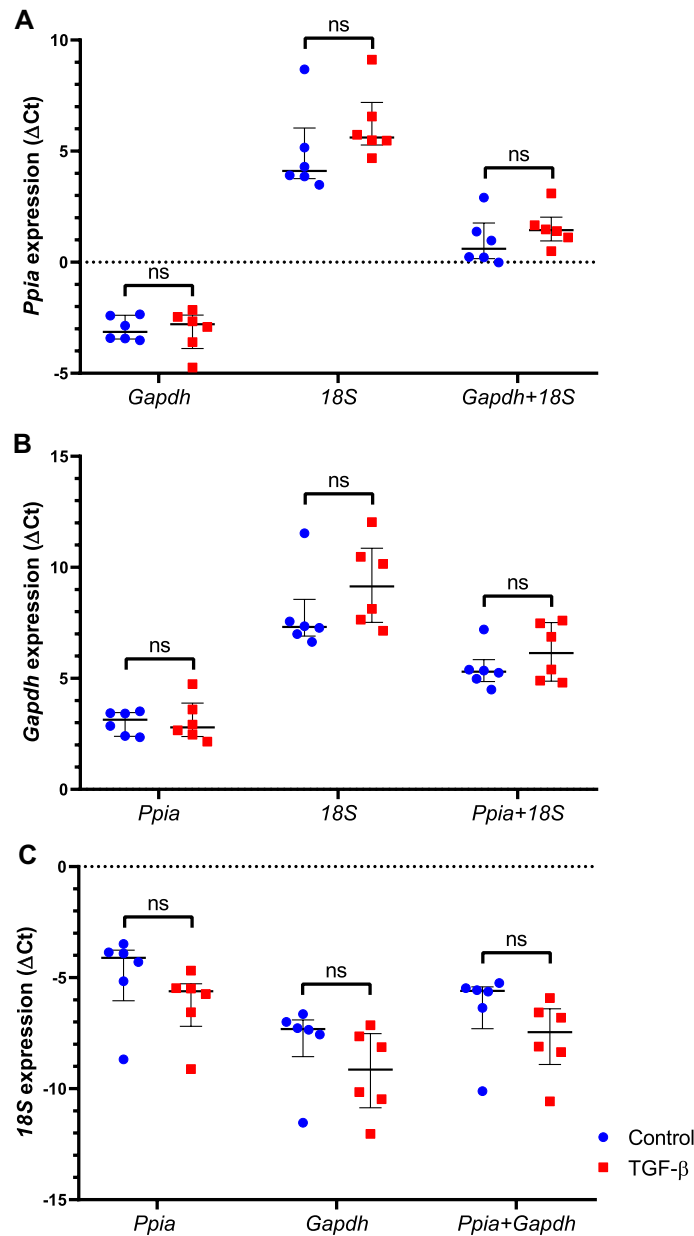
Figure 3. The ΔCt values of *Ppia* (**A**), *Gapdh* (**B**) and *18S* (**C**) candidate housekeeping genes were normalized by combinations of each other. A negative ΔCt value indicates that the target gene is more abundant than the HKG. The median values are expressed as horizontal lines, and the error bars represent the interquartile range. *Ppia*, target expression normalized by *Gapdh*, *18S* and *Gapdh* + *18S*; *Gapdh*, target expression normalized by *Ppia*, *18S* and *Ppia* + *18S*; *18S*, target expression normalized by *Ppia*, *Gapdh* and *Ppia* + *Gapdh*. ns, non-significant.

Validation of the best candidate housekeeping genes for normalizing target genes of *fibronec-tin, vimentin,* and *α-SMA*. To validate the stability of the top three candidate HKG, the relative expression of *fibronectin*, *vimentin,* and *α-SMA* target genes was normalized using different combinations of *Ppia*, *Gapdh* and *18S* (Fig. 6). The expression levels of *fibronectin*, *vimentin,* and *α-SMA* target genes were consistent with upregulation in the TGF-β group relative to controls (Fig. 6). The normalization of target genes by the two less stable genes (*Actb* and *Hprt*) was also evaluated (Supplementary Fig. 1). The use of these HKG, whether alone or in combination, was not able to demonstrate the statistically significant difference that was expected between controls and TGF-β-treated samples of *fibronectin* and *α-SMA* (Supplementary Fig. 1).
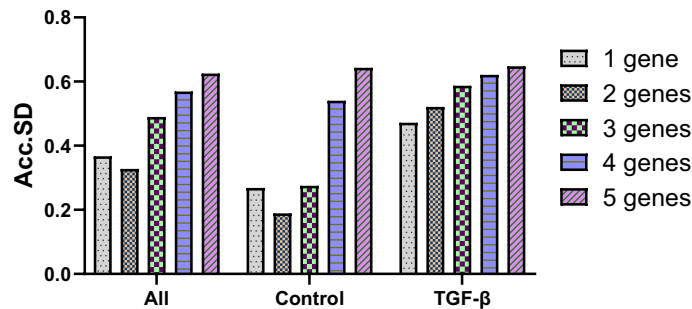
Figure 4. Optimal number of housekeeping genes according to GenEx software analysis. Accumulated standard deviation (Acc.SD) for the five candidate reference genes in all groups to estimate the ideal number of genes for normalization. Lower values of Acc.SD indicate the optimal number of reference genes. All, all samples.

Discussion

The broadly accepted method used to normalize gene expression through RT-PCR technology involves the expression of endogenous HKG. However, the utility of HKG must be validated for specific experimental conditions, since the expression of these endogenous genes can vary depending on experimental conditions[8–10,15]. In vitro systems, including cultured MCs, constitute an useful model to study many pathophysiological states affecting the glomeruli, such as glomerulosclerosis[1]. Therefore, we aimed to determine the most stable reference genes for mRNA quantification in studies performed in vitro, mimicking the in vivo glomerular fibrosis using MMCs treated with TGF-β[10,13,15,30].

Since each algorithm ranked the best candidate HKG, the software packages recommended *Ppia*, *Gapdh* and *18S* as the most stable reference genes between the groups. Peptidylprolyl isomerase A (*Ppia*), a highly abundant protein in the cytoplasm, takes part in various intracellular functions, including a homeostatic role in protein folding and trafficking, intracellular signaling, transcription, inflammation, apoptosis, and regulation of other proteins[31–33]. Glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) catalyzes the sixth reaction of anaerobic glycolysis, which produces ATP and pyruvate. Other than metabolic functions, this enzyme has been implicated in non-metabolic processes, such as apoptosis induction, DNA repair, cellular proliferation, and transcriptional activation[34–36]. Small subunit 18S ribosomal RNA (*18S*) is the smallest component of eukaryotic cytoplasmic ribosomes and is used as one of the molecular markers[37,38].

The other two genes considered in this study (*Hprt* and *Actb*) are also commonly used as reference genes. Hypoxanthine phosphoribosyltransferase (*Hprt*) is responsible for purine metabolism, and deficiency of this gene dysregulates cell cycle-controlling functions and cell proliferation mechanisms[39,40]. Actin beta (*Actb*), which is highly abundant in eukaryotic cells, is essential for a variety of cellular functions and is involved in maintaining the cell's structure, integrity, and motility[41]. Although extensively used as reference genes[14], *Hprt* and *Actb* ranked as the least stable in this study; however, further studies are needed to better delineate the interactions of these genes with TGF-β.

Since *Ppia*, *Gapdh*, and *18S* were the most suitable candidate reference genes, we normalized them by each other, resulting in no statistically significant differences between groups, which suggests that these genes are good choices for our experimental conditions. After determining the candidate HKG by their stability values, we established the optimal number of reference genes using GenEx software. According to calculated Acc.SD, the optimal number of HKG in this model is the combination of two genes. When used together, *Ppia* and *Gapdh* showed a strong correlation, indicating that all samples were correlated and validating the best pair of HKG.

It is well demonstrated that TGF-β stimulates production of *fibronectin*, *vimentin*, and *α-SMA* in cultured MCs[2,42]; thus, the best HKG combinations herein determined were used to normalize these target genes. Several studies have reported that *Ppia*[14,17,43–48], *Gapdh*[44,49,50], and *18S*[45,51–53] are suitable reference genes and could be used as normalizers of target genes in different models. In the present study, the top three candidate reference genes, whether used alone or in combination, showed the expected increase in the expression of the target genes in the TGF-β-treated group. In contrast, the less stable HKG, employed alone or in combination, did not yield these expected differences, indicating that the in silico analysis selected the better, more stable HKG for this in vitro fibrosis model; they also revealed that an inadequate choice of the endogenous standard HKG could influence the results.

It is important to mention that other non-tested genes can also be used for normalization of the expression of target genes, and additional studies are needed to identify additional candidate genes. Furthermore, this study is specific to MMCs stimulated with TGF-β; thus, the conclusions drawn from our study are not transferable to other models that employ MMCs.

Validating gene expression stability of reference genes is crucial for reliable normalization of RT-PCR data. The work herein presented will serve as a reference for future studies using MMCs stimulated with TGF-β and allow a higher reliability and reproducibility in the identification of gene expression alterations.

Among the reference genes tested in this study, the combination of *Ppia* and *Gapdh* was the best HKG pair and should, therefore, be used as HKG in gene expression analysis in TGF-β-treated MMCs models.
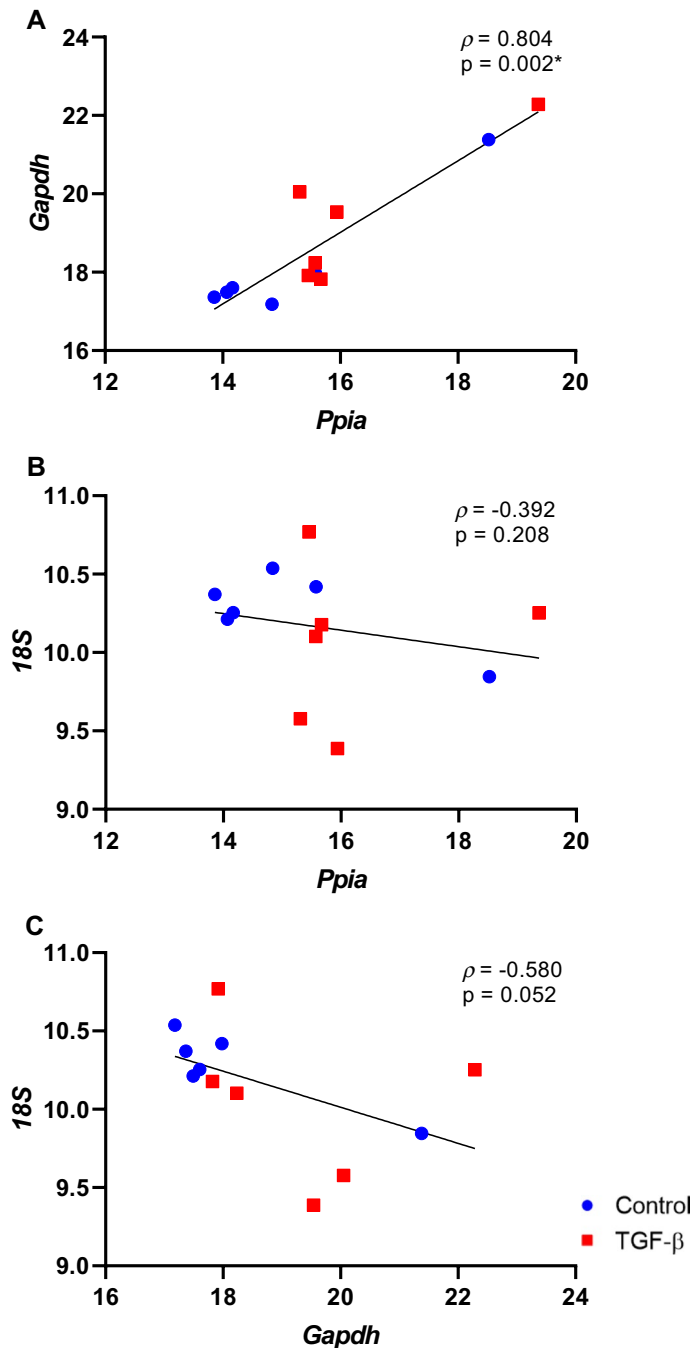
Figure 5. Correlation matrix between the expression of *Ppia* and *Gapdh* (**A**), *Ppia* and *18S* (**B**) and *Gapdh* and *18S* (**C**) candidate housekeeping genes. ρ: Spearman's rank correlation coefficient. *$p < 0.05$.

Methods

Mesangial cell culture.   Mouse mesangial cells (MMCs) were purchased from ATCC (CRL 1927), and the recombinant mouse transforming growth factor-beta (TGF-β) was obtained from R&D Systems (USA). MMCs were grown at 37ºC in plastic flasks in Dulbecco's Modified Eagle's medium/Ham's F12 Medium (DMEM/F12; 3:1 mixture; Invitrogen Corporation, Gaithersburg, MD, USA) containing 10% fetal bovine serum (FBS), penicillin (50 U/mL), and 2.6 g HEPES. The culture flasks were maintained in a 95% air atmosphere and 5% CO_2 humidified environment. At confluence, cells were exposed to DMEM/F12 medium containing no FBS for 24 h according to the following experimental groups: control, cultured in DMEM/F12 medium and TGF-β group, cultured in DMEM/F12 containing 5 ng/mL of recombinant TGF-β. After 24 h of incubation, cells were rinsed twice with PBS, and 1 mL of the commercial kit (TRIzol, Gibco BRL, Rockland, MD, USA) was added to isolate total RNA and evaluate the mRNA expression.
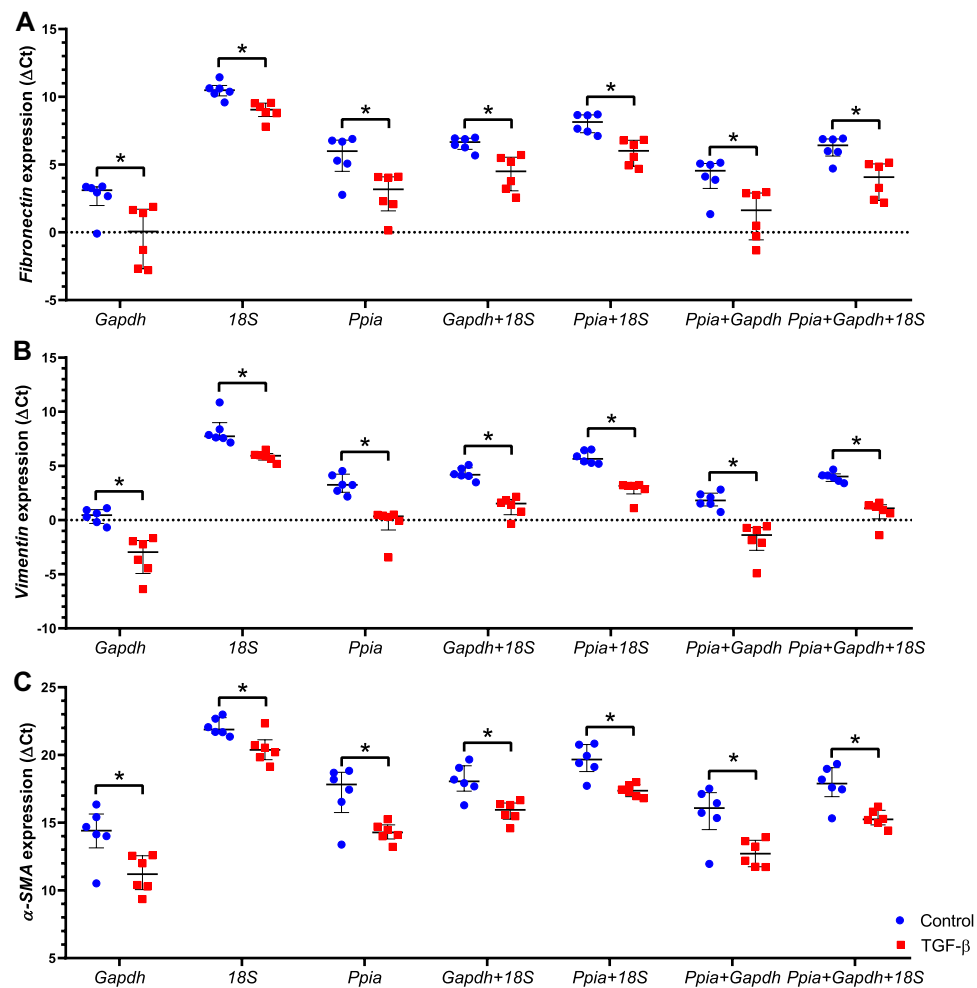
Figure 6. The ΔCt values of *fibronectin* (**A**), *vimentin* (**B**), and *α-SMA* (**C**) target genes normalized by different combinations of the three best candidate housekeeping genes (*Ppia*, *Gapdh,* and *18S*). A negative ΔCt value indicates that the target gene is more abundant than the HKG. The median values are expressed as horizontal lines, and the error bars represent the interquartile range. $p < 0.05$ by Mann–Whitney *U* test: * vs control group.

Housekeeping genes.   The selection of the candidate HKG was based on previous studies that used RT-PCR for gene expressions in kidney diseases models. Then, five genes were selected (*Actb, Hprt, Gapdh, 18S* and *Ppia)* being constitutively expressed in kidney cells with independent cellular functions.

mRNA expression by RT-PCR.   The mRNA expression levels were estimated by quantitative RT-PCR. The total RNA was purified from MMCs using TRIzol kit according to the manufacturer's instructions. The RNA quantity and purity were determined using the NanoVue spectrophotometer (GE Healthcare Life Sciences, USA). A mass of 2 µg of total RNA was treated with DNase (Promega, Madison, WI, USA) to prevent genomic DNA contamination and DNase inactivation was performed according to manufacturer's instructions. The RNA pellet was resuspended in RNase-free water and reverse transcribed into cDNA using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). RT-PCR amplification was performed in triplicate using SYBR Green (Applied Biosystems) in the QuantStudio (TM) 7 Flex System (Applied Biosystems), with specific primers for each gene as follows (sense and antisense, respectively): *Fibronectin* (5'' acactaacgtaaattgcccca 3' and 5' gctaacatcactggggtgtggat 3'), *Vimentin* (5' aggtggatcagctcaccaatgaca 3' and 5' tcaaggtcaagacgtgccagagaa 3'), *α-SMA* (5' tattgtgctggactctggagatgg 3' and 5' agtagtcacgaaggaatagccacg 3'), *Actb* (5' cctcatgccaacacagtgc 3' and 5' acatctgctggaaggtggac 3'), *Hprt* (5' ctcatggactgattatggacaggac 3' and 5' gcaggtcagcaaagaacttatagcc 3'), *Gapdh* (5' ggtggtctcctctgactttaaca 3' and 5' accaggaaatgagccttgacaaag 3'), *18S* (5' gactgtctcgccggtgtc 3' and 5' ggagagc-cggaacgtcga 3') and *Ppia* (5' caggtccatctacggagaga 3' and 5' catccagccattcagtcttg 3'). The relative gene expression was calculated using the PCR conditions under which the amplification curve was logarithmic.

Analysis of housekeeping gene expression stability.   To define the best housekeeping gene and the best combination, we evaluate the cycle thresholds (Ct) value of RT-PCR in five different software applications: DataAssist (version 3.1; https://www.thermofsher.com/br/en/home/technical-resources/sofware-downloads/

dataassist-sofware.html), Bestkeeper (version 1.0; https://www.gene-quantifcation.de/bestkeeper.html), Ref-Finder and the comparative ΔCt method (https://www.heartcure.com.au/refnder/), GeNorm (https://genorm.cmgg.be/), and NormFinder (version 0.953; https://moma.dk/normfnder-sofware), following the authors' recommendations. These software packages determine the relative expression stability of the candidate HKG and generate a rank of the best genes[16]. NormFinder is a freely available tool and recommends a standard deviation less than 0.5 (SD < 0.5). GeNorm software calculates the gene stability measure (M value) and recommends that this value falls below 1.5 (M < 1.5). BestKeeper evaluates the SD and coefficient of variation (CV) of the samples, and a SD of less than 1.0 (SD < 1.0) and a CV of less than 3.0 (CV < 3.0) are required. DataAssist shows the Ct values of the candidate genes for all samples and organizes them by score. The comparative ΔCt method was used to calculate the mean SD of the samples. RefFinder software includes all the above software and calculates the geometric mean (Geomean). The optimal number of HKG was evaluated using the GenEx software package, which calculates the accumulated standard deviation (Acc.SD) of sample groups and estimates the ideal number of genes for normalization. The following groups were evaluated: control, TGF-β-treated and All (control + TGF-β samples).

Statistical analysis. The test of normality (Shapiro–Wilk test) showed that Ct values of HKG were not normally distributed; hence, the median values are expressed as horizontal lines, and the error bars represent interquartile range. All groups were analyzed using the Mann–Whitney U test and the Spearman's correlation. Values between 0.30–0.50 were considered as a weak correlation, 0.50–0.70 as moderate, 0.70–0.90 as strong and 0.90–1.00 as very strong correlation.

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

References
1. Nogueira, A., Pires, M. J. & Oliveira, P. A. Pathophysiological mechanisms of renal fibrosis: A review of animal models and therapeutic strategies. *In Vivo* **31**, 1–22. https://doi.org/10.21873/invivo.11019 (2017).
2. Simonson, M. S. Phenotypic transitions and fibrosis in diabetic nephropathy. *Kidney Int.* **71**, 846–854. https://doi.org/10.1038/sj.ki.5002180 (2007).
3. Leask, A. & Abraham, D. J. TGF-beta signaling and the fibrotic response. *FASEB J.* **18**, 816–827. https://doi.org/10.1096/fj.03-1273rev (2004).
4. Padgett, R. W. & Reiss, M. TGFbeta superfamily signaling: Notes from the desert. *Development* **134**, 3565–3569. https://doi.org/10.1242/dev.005926 (2007).
5. Rahimi, R. A. & Leof, E. B. TGF-beta signaling: A tale of two responses. *J. Cell Biochem.* **102**, 593–608. https://doi.org/10.1002/jcb.21501 (2007).
6. Derveaux, S., Vandesompele, J. & Hellemans, J. How to do successful gene expression analysis using real-time PCR. *Methods* **50**, 227–230. https://doi.org/10.1016/j.ymeth.2009.11.001 (2010).
7. Ho-Pun-Cheung, A. *et al.* Reverse transcription-quantitative polymerase chain reaction: Description of a RIN-based algorithm for accurate data normalization. *BMC Mol. Biol.* **10**, 31. https://doi.org/10.1186/1471-2199-10-31 (2009).
8. 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).
9. 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).
10. Muñoz, J. J. *et al.* Identification of housekeeping genes for microRNA expression analysis in kidney tissues of Pkd1 deficient mouse models. *Sci. Rep.* **10**, 231. https://doi.org/10.1038/s41598-019-57112-4 (2020).
11. Caracausi, M. *et al.* Systematic identification of human housekeeping genes possibly useful as references in gene expression studies. *Mol. Med. Rep.* **16**, 2397–2410. https://doi.org/10.3892/mmr.2017.6944 (2017).
12. Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends Genet.* **29**, 569–574. https://doi.org/10.1016/j.tig.2013.05.010 (2013).
13. Wang, Z., Lyu, Z., Pan, L., Zeng, G. & Randhawa, P. Defining housekeeping genes suitable for RNA-seq analysis of the human allograft kidney biopsy tissue. *BMC Med. Genomics* **12**, 86. https://doi.org/10.1186/s12920-019-0538-z (2019).
14. Jung, M. *et al.* In search of suitable reference genes for gene expression studies of human renal cell carcinoma by real-time PCR. *BMC Mol Biol* **8**, 47. https://doi.org/10.1186/1471-2199-8-47 (2007).
15. Huggett, J., Dheda, K., Bustin, S. & Zumla, A. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* **6**, 279–284. https://doi.org/10.1038/sj.gene.6364190 (2005).
16. Hosni, N. D., Anauate, A. C. & Boim, M. A. Reference genes for mesangial cell and podocyte qPCR gene expression studies under high-glucose and renin-angiotensin-system blocker conditions. *PLoS ONE* **16**, e0246227. https://doi.org/10.1371/journal.pone.0246227 (2021).
17. Muñoz, J. J. *et al.* Ppia is the most stable housekeeping gene for qRT-PCR normalization in kidneys of three Pkd1-deficient mouse models. *Sci. Rep.* **11**, 19798. https://doi.org/10.1038/s41598-021-99366-x (2021).
18. Guan, Q., Nguan, C. Y. & Du, C. Expression of transforming growth factor-beta1 limits renal ischemia-reperfusion injury. *Transplantation* **89**, 1320–1327. https://doi.org/10.1097/TP.0b013e3181d8e9dc (2010).
19. Morrissey, J. *et al.* Transforming growth factor-beta induces renal epithelial jagged-1 expression in fibrotic disease. *J. Am. Soc. Nephrol.* **13**, 1499–1508. https://doi.org/10.1097/01.asn.0000017905.77985.4a (2002).
20. Wu, X. *et al.* Exosomes from high glucose-treated glomerular endothelial cells trigger the epithelial-mesenchymal transition and dysfunction of podocytes. *Sci. Rep.* **7**, 9371. https://doi.org/10.1038/s41598-017-09907-6 (2017).
21. Cappelli, C. *et al.* The TGF-β profibrotic cascade targets ecto-5'-nucleotidase gene in proximal tubule epithelial cells and is a traceable marker of progressive diabetic kidney disease. *Biochim. Biophys. Acta Mol. Basis Dis.* **1866**, 165796. https://doi.org/10.1016/j.bbadis.2020.165796 (2020).
22. Castro, N. E., Kato, M., Park, J. T. & Natarajan, R. Transforming growth factor β1 (TGF-β1) enhances expression of profibrotic genes through a novel signaling cascade and microRNAs in renal mesangial cells. *J. Biol. Chem.* **289**, 29001–29013. https://doi.org/10.1074/jbc.M114.600783 (2014).

23. Takakura, K., Tahara, A., Sanagi, M., Itoh, H. & Tomura, Y. Antifibrotic effects of pirfenidone in rat proximal tubular epithelial cells. *Ren. Fail.* **34**, 1309–1316. https://doi.org/10.3109/0886022X.2012.718955 (2012).

24. Deshpande, S. *et al.* Reduced autophagy by a microRNA-mediated signaling cascade in diabetes-induced renal glomerular hypertrophy. *Sci. Rep.* **8**, 6954. https://doi.org/10.1038/s41598-018-25295-x (2018).

25. Huang, F. *et al.* High glucose and TGF-β1 reduce expression of endoplasmic reticulum-resident selenoprotein S and selenoprotein N in human mesangial cells. *Ren. Fail.* **41**, 762–769. https://doi.org/10.1080/0886022X.2019.1641413 (2019).

26. Jia, Y. *et al.* Dysregulation of histone H3 lysine 27 trimethylation in transforming growth factor-β1-induced gene expression in mesangial cells and diabetic kidney. *J. Biol. Chem.* **294**, 12695–12707. https://doi.org/10.1074/jbc.RA119.007575 (2019).

27. Masola, V. *et al.* In vitro effects of interleukin (IL)-1 beta inhibition on the epithelial-to-mesenchymal transition (EMT) of renal tubular and hepatic stellate cells. *J. Transl. Med.* **17**, 12. https://doi.org/10.1186/s12967-019-1770-1 (2019).

28. Ma, J. *et al.* Up-regulation of microRNA-93 inhibits TGF-β1-induced EMT and renal fibrogenesis by down-regulation of Orai1. *J. Pharmacol. Sci.* **136**, 218–227. https://doi.org/10.1016/j.jphs.2017.12.010 (2018).

29. Suzuki, Y. *et al.* Transforming growth factor-β induces vascular endothelial growth factor-C expression leading to lymphangiogenesis in rat unilateral ureteral obstruction. *Kidney Int.* **81**, 865–879. https://doi.org/10.1038/ki.2011.464 (2012).

30. Biederman, J., Yee, J. & Cortes, P. Validation of internal control genes for gene expression analysis in diabetic glomerulosclerosis. *Kidney Int.* **66**, 2308–2314. https://doi.org/10.1111/j.1523-1755.2004.66016.x (2004).

31. Leong, K. G., Ozols, E., Kanellis, J., Nikolic-Paterson, D. J. & Ma, F. Y. Cyclophilin A promotes inflammation in acute kidney injury but not in renal fibrosis. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms21103667 (2020).

32. Nigro, P., Pompilio, G. & Capogrossi, M. C. Cyclophilin A: A key player for human disease. *Cell Death Dis.* **4**, e888. https://doi.org/10.1038/cddis.2013.410 (2013).

33. Ramachandran, S. *et al.* Plasma level of cyclophilin A is increased in patients with type 2 diabetes mellitus and suggests presence of vascular disease. *Cardiovasc. Diabetol.* **13**, 38. https://doi.org/10.1186/1475-2840-13-38 (2014).

34. Nicholls, C., Li, H. & Liu, J. P. GAPDH: A common enzyme with uncommon functions. *Clin. Exp. Pharmacol. Physiol.* **39**, 674–679. https://doi.org/10.1111/j.1440-1681.2011.05599.x (2012).

35. Seidler, N. W. GAPDH and intermediary metabolism. *Adv. Exp. Med. Biol.* **985**, 37–59. https://doi.org/10.1007/978-94-007-4716-6_2 (2013).

36. Tarze, A. *et al.* GAPDH, a novel regulator of the pro-apoptotic mitochondrial membrane permeabilization. *Oncogene* **26**, 2606–2620. https://doi.org/10.1038/sj.onc.1210074 (2007).

37. Meyer, A., Todt, C., Mikkelsen, N. T. & Lieb, B. Fast evolving 18S rRNA sequences from Solenogastres (Mollusca) resist standard PCR amplification and give new insights into mollusk substitution rate heterogeneity. *BMC Evol. Biol.* **10**, 70. https://doi.org/10.1186/1471-2148-10-70 (2010).

38. Moore, P. B. & Steitz, T. A. The involvement of RNA in ribosome function. *Nature* **418**, 229–235. https://doi.org/10.1038/418229a (2002).

39. Kang, T. H., Park, Y., Bader, J. S. & Friedmann, T. The housekeeping gene hypoxanthine guanine phosphoribosyltransferase (HPRT) regulates multiple developmental and metabolic pathways of murine embryonic stem cell neuronal differentiation. *PLoS ONE* **8**, e74967. https://doi.org/10.1371/journal.pone.0074967 (2013).

40. Townsend, M. H., Robison, R. A. & O'Neill, K. L. A review of HPRT and its emerging role in cancer. *Med. Oncol.* **35**, 89. https://doi.org/10.1007/s12032-018-1144-1 (2018).

41. Bugyi, B. & Kellermayer, M. The discovery of actin: "to see what everyone else has seen, and to think what nobody has thought". *J. Muscle Res. Cell Motil.* **41**, 3–9. https://doi.org/10.1007/s10974-019-09515-z (2020).

42. Schnaper, H. W., Hayashida, T., Hubchak, S. C. & Poncelet, A. C. TGF-beta signal transduction and mesangial cell fibrogenesis. *Am. J. Physiol. Renal. Physiol.* **284**, F243-252. https://doi.org/10.1152/ajprenal.00300.2002 (2003).

43. Bjerregaard, H., Pedersen, S., Kristensen, S. R. & Marcussen, N. Reference genes for gene expression analysis by real-time reverse transcription polymerase chain reaction of renal cell carcinoma. *Diagn. Mol. Pathol.* **20**, 212–217. https://doi.org/10.1097/PDM.0b013e318212e0a9 (2011).

44. Cui, X., Zhou, J., Qiu, J., Johnson, M. R. & Mrug, M. Validation of endogenous internal real-time PCR controls in renal tissues. *Am. J. Nephrol.* **30**, 413–417. https://doi.org/10.1159/000235993 (2009).

45. Schmid, H. *et al.* Validation of endogenous controls for gene expression analysis in microdissected human renal biopsies. *Kidney Int.* **64**, 356–360. https://doi.org/10.1046/j.1523-1755.2003.00074.x (2003).

46. Ma, Y., Dai, H., Kong, X. & Wang, L. Impact of thawing on reference gene expression stability in renal cell carcinoma samples. *Diagn. Mol. Pathol.* **21**, 157–163. https://doi.org/10.1097/PDM.0b013e31824d3435 (2012).

47. Ma, Y. *et al.* Renal tissue thawed for 30 minutes is still suitable for gene expression analysis. *PLoS ONE* **9**, e93175. https://doi.org/10.1371/journal.pone.0093175 (2014).

48. Zinzow-Kramer, W. M., Horton, B. M. & Maney, D. L. Evaluation of reference genes for quantitative real-time PCR in the brain, pituitary, and gonads of songbirds. *Horm. Behav.* **66**, 267–275. https://doi.org/10.1016/j.yhbeh.2014.04.011 (2014).

49. Barber, R. D., Harmer, D. W., Coleman, R. A. & Clark, B. J. GAPDH as a housekeeping gene: Analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol. Genomics* **21**, 389–395. https://doi.org/10.1152/physiolgenomics.00025.2005 (2005).

50. Gholami, K., Loh, S. Y., Salleh, N., Lam, S. K. & Hoe, S. Z. Selection of suitable endogenous reference genes for qPCR in kidney and hypothalamus of rats under testosterone influence. *PLoS ONE* **12**, e0176368. https://doi.org/10.1371/journal.pone.0176368 (2017).

51. Bas, A., Forsberg, G., Hammarström, S. & Hammarström, M. L. Utility of the housekeeping genes 18S rRNA, beta-actin and glyceraldehyde-3-phosphate-dehydrogenase for normalization in real-time quantitative reverse transcriptase-polymerase chain reaction analysis of gene expression in human T lymphocytes. *Scand. J. Immunol.* **59**, 566–573. https://doi.org/10.1111/j.0300-9475.2004.01440.x (2004).

52. Banda, M., Bommineni, A., Thomas, R. A., Luckinbill, L. S. & Tucker, J. D. Evaluation and validation of housekeeping genes in response to ionizing radiation and chemical exposure for normalizing RNA expression in real-time PCR. *Mutat. Res.* **649**, 126–134. https://doi.org/10.1016/j.mrgentox.2007.08.005 (2008).

53. Granfar, R. M., Day, C. J., Kim, M. S. & Morrison, N. A. Optimised real-time quantitative PCR assays for RANKL regulated genes. *Mol. Cell Probes* **19**, 119–126. https://doi.org/10.1016/j.mcp.2004.10.003 (2005).

Acknowledgements

Author contributions

B.B., A.C.A. and A.S.N. designed the study. B.B., E.M. and A.S.N. performed the experiments. B.B. and A.C.A. analyzed the data. B.B., M.A.B. and A.S.N. wrote the first draft of the manuscript. M.A.B .and A.S.N. provided

financial support, study materials, and the final approval of the manuscript. All authors reviewed the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-19548-z.

Correspondence and requests for materials should be addressed to A.S.N.

Reprints and permissions information is available at www.nature.com/reprints.

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