Validation of the Reference Genes for the Gene Expression Studies in Different Cell Lines of Pig

Chun-Di Xie,1 Bingyuan Wang,1 Zhao-Ji Shen,1,2 Wen-Ye Yao,1 Hong Ao,1 Bugao Li,3 Yangli Pei,2 and Rong Zhou1

1State Key Laboratory of Animal Nutrition; Key Laboratory of Animal Genetics Breeding and Reproduction, Ministry of Agriculture and Rural Affairs, Institute of Animal Science, Chinese Academy of Agricultural Sciences, 100193 Beijing, China
2Guangdong Provincial Key Laboratory of Animal Molecular Design and Precise Breeding, College of Life Science and Engineering, Foshan University, 528231 Foshan, China
3College of Animal Science, Shanxi Agricultural University, 030801 Shanxi, China

Correspondence should be addressed to Yangli Pei; peiyangli@163.com and Rong Zhou; zhourong03@caas.cn

Received 13 April 2021; Accepted 20 July 2021; Published 19 August 2021

Academic Editor: Pengjun Shi

Copyright © 2021 Chun-Di Xie et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Reverse transcription quantitative real-time polymerase chain reaction is one of the important methods to investigate gene expression in cells and tissues. However, if the data cannot be normalized with appropriate reference genes, the results may be unreliable. In this study, we detected the expression of 15 reference genes in three pig cell lines. The results showed that SDHA and ALDOA were the most stable reference genes in 3D4/21 cells. TOP2B, TBP, and PPIA were the most stable reference genes in PK-15 cells. SDHA and ALDOA were the most stable reference genes in IPEC-J2 cells. In addition, each cell line only needs to use two reference genes to standardize the expression of target genes. Taken together, this study provides a reference for different pig cell lines to select reference genes and also provides a theoretical basis for the use of these cell lines in related functional researches.

1. Introduction

Reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) is frequently used to detect gene expression in cells and tissues due to its high sensitivity, specificity, and accuracy [1–4]. However, the results of RT-qPCR can be affected by some factors, including varying quality and quantity of RNA, different sample amounts, enzymatic efficiency in reverse transcription steps, and PCR amplification efficiency [5, 6]. Therefore, it is necessary to select a proper reference gene as an internal control to correct and normalize the expression of target gene. Excellent reference genes need to be expressed stably under all conditions. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), β-actin (ACTB), and 18S ribosomal RNA (18S) are considered to be stable expression in different conditions and tissues and are widely used as reference genes [7]. However, it has been proved that the expression of these genes was not as stable as initially thought. Many studies showed that these genes had a great different expression level in different experimental conditions [8–15]. Therefore, it is necessary to select suitable reference genes according to different tissues, cells, and experiments.

Pig cell lines are often used as cell models to explore the mechanism of gene function and immune disease. Porcine alveolar macrophages (3D4/21 cells), isolated from the lung of Landrace pig and immortalized with SV40 large T antigen transformed with pSV3neo, could secrete cytokines and were used to study the mechanism of Streptococcus suis [16] and swine fever virus [17]. Porcine kidney 15 cell line (PK-15) is a clone of PK-1a cell, which is often used to study circovirus [18]. And intestinal porcine epithelial cells (IPEC-J2) are most commonly used to study porcine epidemic diarrhea-related diseases [19]. In these researches, a large number of gene expression verification are usually involved, and the inaccurate results of RT-qPCR will lead to the failure...
of the experiment. Therefore, it is of great importance to select an appropriate reference gene for these cell lines. At present, most of the studies on reference genes are focused on tissues, but few on cell lines, especially on these three types of pig cell lines.

The aim of this study was the selection of suitable reference genes for expression studies in pig cell lines using quantitative RT-qPCR. In this study, the expression level of 15 candidate genes in porcine 3D4/21, PK-15, and IPEC-J2 cell lines was detected. The expression stability of genes was analyzed and evaluated by geNorm [6], NormFinder [20], and BestKeeper [21]. The findings screened out the most stable reference genes for these three porcine cell lines and provided the reference for carrying out relevant experiments to select reference genes.

2. Materials and Methods

2.1. Cell Culture. PK15, IPEC-J2, and 3D4/21 cells were seeded at the concentration of $1 \times 10^5$ cells per well in 12 well plates containing 1 mL of DMEM (Gibco, 12800082) or RPMI-1640 (Gibco, 42401042), respectively, and supplemented with 10% fetal bovine serum (ExCell Bio, FND500). Cells were cultured in a highly humidified atmosphere of 95% air and 5% CO$_2$ at 37°C.

2.2. Isolation of RNA and Real-Time Quantitative PCR (RT-qPCR). Cell were homogenized in 500 μL of RNA isolater (Vazyme, R401-01) according to the manufacturer’s instructions, purified by DNaseI, and quantified by spectrophotometry. Each cell had three biological repetitions. The cDNA for qPCR analysis was synthesized using the HiScript III 1st Strand cDNA Synthesis Kit (+gDNA wiper) (Vazyme, R312-01). Prior to qPCR amplification, cDNA was diluted to 300 ng/μL. The reaction mixture for the qPCR step was ChamQ Universal SYBR qPCR Master Mix (Vazyme, Q711-02). The 15 μL RT-qPCR reaction mixture encompassed 1 μL of cDNA template, 7.2 μL of 2× SYBR premix Ex Taq, 0.3 μL of 50× ROX Reference Dye II, 0.3 μL of each forward and reverse primer, and 5.9 μL of double-distilled water. Each sample was performed in triplicates. And the RT-qPCR conditions were as follows: 95°C for 5 min, followed by 40 cycles of 95°C for 5 s and 60°C for 34 s. RT-qPCR was performed using TB Green Premix Ex Taq in a QuantStudio 3 (Thermo Fisher Scientific, QuantStudio 3).

| Gene       | Gene name                                                                 | Primer sequence (5′-3′) | Amplicon size (bp) | Reference |
|------------|---------------------------------------------------------------------------|-------------------------|--------------------|-----------|
| B2M        | Beta-2-microglobulin                                                       | F: TTCACACCCTGTCAGTAG   | 166                | [23]      |
|            |                                                                           | R: CCAGATACATAGTCCTCTAG |                    |           |
| TBP        | TATA box binding protein                                                  | F: GATGGACGTCTGTTAGAGG  | 124                | [23]      |
|            |                                                                           | R: AGCAGCACATTCAGAGCA   |                    |           |
| YWHAZ      | Tyrosine 3-monoxygenase/tryptophan 5-monoxygenase activation protein, zeta polypeptide | F: ATGCAAACCAACTTCATCT  | 178                | [24]      |
|            |                                                                           | R: GCATTATTAGCGTGCTGTCTT|                    |           |
| PPIA       | Peptidylprolyl isomerase A (cyclophilin A)                               | F: CACAAAGGTTCGCCAGTTTTT| 171                | [25]      |
|            |                                                                           | R: TGGTCACATCGCAACATTGGA|                    |           |
| RPL4       | Ribosomal protein L4                                                      | F: TGCCGATCGAGATGGAAA   | 178                | NCBI      |
|            |                                                                           | R: CAGAACAGCCTCTTTGATG  |                    |           |
| SDHA       | Succinate dehydrogenase complex, subunit A                               | F: CTACAGGGGCAAGTCTGAGA | 141                | [22]      |
|            |                                                                           | R: AAGACAAGGATGCCAGAG   |                    |           |
| HPRT1      | Hypoxanthine phosphoribosyl transferase 1                                 | F: CGAGGAATTTGGGAAAAAT  | 181                | [23]      |
|            |                                                                           | R: CTATTTCTGCTGATGTTGATG|                    |           |
| TOP2B      | Topoisomerase II beta                                                     | F: AAAGGGGAGAGTCAATGAT  | 115                | [1]       |
|            |                                                                           | R: ACATCTCTGCTTTCAGGC   |                    |           |
| ALDOA      | Aldolase A, fructose-bisphosphate                                         | F: GAACAAAGCGGGAGCAAC   | 142                | [26]      |
|            |                                                                           | R: ATGATGCGGGAGGAGGA    |                    |           |
| HSPCB (HSP90AB1) | Heat shock 90 kDa protein 1, beta                                   | F: GGCAGAAGCAGAGAACG    | 131                | [26]      |
|            |                                                                           | R: CAGACTGGAGGTATGGTAG  |                    |           |
| PGK1       | Phosphoglycerate kinase 1                                                | F: AGATAAAGCAAACAGAGG   | 126                | [27]      |
|            |                                                                           | R: TGGCAGGATAGGGATACC  |                    |           |
| POLR2G     | Polymerase (RNA) II (DNA directed) polypeptide G                          | F: CTCAAAGTCAAAAGGTCCAG | 181                | [1]       |
|            |                                                                           | R: GCCAAACAACTTCAGGCG   |                    |           |
| ACTB       | Beta actin                                                                | F: GGAACCTTGCAGAGGAGATT| 138                | NCBI      |
|            |                                                                           | R: AGGAAGGGCGCTGGAAGAG  |                    |           |
| GAPDH      | Glyceraldehyde-3-phosphate dehydrogenase                                 | F: TCGGAGTGAAAGGATTTGCC| 189                | NCBI      |
|            |                                                                           | R: TGACAAGCTTCGAGCTTCCCCCC|                                   |           |
| RPS18      | Ribosomal protein S18                                                    | F: GTACACGTGAAACCCCATTT| 151 bp              | NCBI      |
|            |                                                                           | R: CCATCAATCGTAGATGCG   |                    |           |
In this program, reference genes with high coefficient of variation (CV) and standard deviation (SD) value are more stable [21]. The correlation coefficient (R) can be used to correlate the CT value of gene expression. The program can obtain genes. However, the program can only select the most suitable gene expression, and the criterion is same with geNorm value. The gene-stability measure in geNorm for control gene determine the optimal number according to the standardized factor by increasing a new reference gene and software can also calculate the paired variations of the genes. On the contrary, the genes with low M value are stably expressed and would not be proper reference genes. On the contrary, the genes with low M values are stably expressed and would be suitable reference genes. The software can also calculate the paired variations (V) of the standardized factor by increasing a new reference gene and determine the optimal number according to the VN/VN + 1 value. The gene-stability measure in geNorm for control gene is the arithmetic mean of all pairwise variations [6].

NormFinder can calculate the stable value of reference gene expression, and the criterion is same with geNorm [20]. However, the program can only select the most suitable genes.

The BestKeeper (version 1) program can directly calculate the CT value of gene expression. The program can obtain the correlation coefficient (R), standard deviation (SD), and coefficient of variation (CV) of pairing between each gene. In this program, reference genes with high R value and low CV and SD value are more stable [21].

3. Results

3.1. The Expression of Candidate Genes in Three Cell Lines Detected by RT-qPCR. The expression of 15 candidate genes was detected by RT-qPCR. The CT value was used to detect the stability of gene expression (Figure 1). The results showed that the average CT value of RPS18 was the lowest (12.33), while the average CT value of TATA box binding protein (TBP) was the highest (22.65). These candidate genes are highly expressed in all cell lines, but their expression level varies greatly. Therefore, it is necessary to evaluate the stability of gene expression and determine the appropriate number of internal reference genes for accurate gene expression profile analysis in different cells.

3.2. Expression Stability of Candidate Genes Analyzed by geNorm. The M values of 15 candidate genes were calculated by geNorm (v3.5) program. The results are shown in Figure 2. The order of gene stability in all cell lines is (from the most stable to the least stable): Ribosomal protein L4(RPL4)/TBP, Heat shock 90kDa protein 1, beta (HSPCB), Topoisomerase II beta(TOP2B), GAPDH, ACTB, Hypoxantine phosphoribosyl transferase 1 (HPRT1), Phosphoglycerate kinase 1 (PGK1), Aldolase A, fructose-bisphosphatase (ALDOA), Polymerase (RNA) II (DNA directed) polypeptide G (POLR2G), Succinate dehydrogenase complex subunit A (SDHA), Tyrosine 3-monoxygenase/tryptophan 5-monoxygenase activation protein, zeta polypeptide (YWHAZ), beta-2-microglobulin (B2M), Peptidylprolyl isomerase A (cyclophilin A) (PPIA), and RPS18 (Figure 2(a)). In PK-15 cell line, SDHA and B2M were the most stable genes, and RPS18 was the least stable gene (Figure 2(b)). B2M and TBP with low M values were identified as the two most stable genes in 3D4/21 cell line (Figure 2(c)). B2M and TOP2B which had low M values were identified as the two most stable genes in IPEC-J2 cell line (Figure 2(d)). Specially, RPS18 was the least stable in any cell types.

One reference gene is usually not enough for gene expression analysis in all cell types. Therefore, geNorm is used to analyze the optimal number of reference genes. GeNorm computationally introduces a new reference gene. Then, the paired variation value of the factor was standardized. Thereafter, the optimal number of reference genes was determined by the value of VN/VN + 1 is less than 0.15, N is the most suitable reference gene number. The results showed that VN/VN + 1 was less than 0.15 in all three cells, so the two reference genes were the best combination for gene expression analysis of the three cells (Figure 3(a)). The optimal number of reference genes for PK-15, 3D4/21, and IPEC-J2 cells is shown in Figures 3(b)–3(d), and the number is two.

3.3. Analysis of Gene Stability Value by NormFinder Program. The NormFinder program calculated the stability of gene expression (stability value (SV)), which was used to rank the genes. According to NormFinder analyses (Table 2), we found that POLR2G was the most stable in 3D4/21 cells. PPIA and TBP were the most stable in PK-15 cells. POLR2G and RPL4 were the most stable in IPEC-J2 cells. These results were not completely consistent with the results analyzed by geNorm. Maybe the algorithms of the two software are different. Although the two algorithms are different, RPS18 was the least stable in both software.
Figure 2: Continued.
3.4. Analysis of Gene Stability Value by BestKeeper Program.

The BestKeeper program is an Excel-based software tool. On the basis of the correlation coefficients ($R$), CV, and SD values, the optimal reference gene was determined (Table 3). The optimal reference gene was selected by high $R$ value ($\geq 0.900$) and low CV and SD value. This program can calculate $R$ values for up to 10 genes. Therefore, we selected 10 reference genes for analysis according to the results of geNorm and NormFinder. In all samples with higher $R$ values ($\geq 0.900$), the SDHA and ALDOA genes had the lowest CV and SD values in 3D4/21 cells, the TOP2B gene had the lowest CV and SD values in PK-15 cells, and the PGK1 gene had the lowest CV and SD values in IPEC-J2 cells. These results indicate that the most stable reference genes in 3D4/21, PK-15, and IPEC-J2 are SDHA/ALDOA, TOP2B, and PGK1, respectively.

4. Discussion

RT-qPCR is one of the most commonly used methods in molecular biology research for gene expression. Appropriate reference genes are very important for the reliability and repeatability of gene expression results. However, the idealized reference gene does not exist. Therefore, it is necessary to select suitable reference genes under different experimental conditions. At present, the selection of reference genes has been studied in different species, experimental conditions, cells, and tissues. However, no systematic analysis of suitable endogenous control genes exists for different types of pig cell lines.

At present, there are only some studies in human cells. RefFinder was used to select the optimal reference genes of human reticulocyte [28]; MPP1 and GAPDH were predicted...
Figure 3: Continued.
as the best reference genes of reticulocyte through comprehensive sequencing. The reference genes of nine hepatocellular carcinoma (HCC) cell lines were systematically evaluated, revealing that TFG and SFRS4 were the most reliable reference genes [29].

As an increasing study exploring mRNA expression in pig cells has been published, there has been a greater interest in evaluating the commonly used, widely expressed housekeeping genes for comparisons between different types of pig cells. The expression stability of four genes (EEF1A1, GAPDH, HPRT1, and TOP2B) in pig tissues was evaluated by geNorm [30], and the results showed that EEF1A1 and TOP2B were the most stable genes in kidney tissues, while HPRT1 and TOP2B were the most stable genes in lung tissues. However, in our study, the result showed that B2M and SDHA were the most stable reference genes in PK-15 cells, which isolated from kidney tissues. TBP and B2M were the most stable reference genes in 3D4/21 cells, which is different from the results in corresponding tissues, indicating that the reference genes suitable for tissues may not be the right housekeeping genes in cells. Furthermore, even though SDHA and ALDOA in 3D4/21 cells, TOP2B, TBP, and PPIA in PK-15 cells, and SDHA and ALDOA in IPEC-J2 cells are recommended as reference gene in our study, the PK-15 cells and IPEC-J2 cells, which are both epithelial cells, owned dissimilar recommended reference genes.

In addition, the ranking of reference genes was different according to the geNorm NormFinder and BestKeeper analyses. B2M/TBP, SDHA/B2M, and TOP2B/B2M were the most stable reference genes in 3D4/21, PK-15, and IPEC-J2, by
geNorm, respectively. NormFinder identified that the expression of POLR2G in 3D4/21, PPIA/TBP in PK-15, and POLR2G/RPL4 in IPEC-J2 was highly stable. BestKeeper evaluates the expression variation for each single reference gene and showed that the most stable reference genes were SDHA/ALDOA in 3D4/21, TOP2B in PK-15, and PGK1 in IPEC-J2. The published work conducted on reference genes in 7 different porcine tissues has analyzed the stability of 15 candidate genes (ACTB, GAPDH, HPAR1, ALDOA, B2M, HSPCB, PPIA, PGK1, POLR2G, RPL4, TBP, RPS18, SDHA, TOP2B) by NormFinder, geNorm, and BestKeeper, and the results were also inconsistent, which may lead by the software algorithm deviation [1]. As we all know, instead of SDHA, ALDOA, TOP2B, TBP, and PPIA, ACTB, GAPDH, and RPS18 were the most popular reference genes in the mRNA expression-related study; it may due to the lack of systematic analysis of suitable endogenous control genes that exists under different conditions.

5. Conclusion

This investigation found evidence that there can be variation in the expression of commonly used reference genes in different type of pig cells. In general, we recommend taking SDHA and ALDOA in 3D4/21 cells, TOP2B, TBP, and PPIA in PK-15 cells, and SDHA and ALDOA in IPEC-J2 cells as reference genes to normalize the mRNA expression level.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

Chun-Di Xie and Bingyuan Wang contributed equally to this work.

References

[1] S. J. Park, S. G. Kwon, J. H. Hwang, D. H. Park, T. W. Kim, and C. W. Kim, "Selection of appropriate reference genes for RT-qPCR analysis in Berkshire, Duroc, Landrace, and Yorkshire pigs," *Gene*, vol. 558, no. 1, pp. 152–158, 2015.

[2] K. Mishima, Y. Kato, M. K. Kaneko, R. Nishikawa, T. Hirose, and M. Matsutani, "Increased expression of podoplanin in malignant astrocytic tumors as a novel molecular marker of malignant progression," *Acta Neuropathologica*, vol. 111, no. 5, pp. 483–488, 2006.

[3] C. Blázquez, M. Salazar, A. Carracedo et al., "Cannabinoids inhibit glioma cell invasion by down-regulating matrix metalloproteinase-2 expression," *Cancer Research*, vol. 68, no. 6, pp. 1945–1952, 2008.

[4] Y. Ruano, M. Mollejo, F. I. Camacho et al., "Identification of survival-related genes of the phosphatidylinositol 3'-kinase signaling pathway in glioblastoma multiforme," *Cancer*, vol. 112, no. 7, pp. 1575–1584, 2008.

[5] S. J. Park, J. W. Huh, Y. H. Kim et al., "Selection of internal reference genes for normalization of quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis in the canine brain and other organs," *Molecular Biotechnology*, vol. 54, no. 1, pp. 47–57, 2013.

[6] J. Vandesompele, K. de Preter, F. Pattyn et al., "Accurate normalization of real-time quantitative RT-PCR data by
geometric averaging of multiple internal control genes,” *Genome Biology*, vol. 3, no. 7, 2002.

[7] Z. Tang, Y. Li, P. Wan et al., “LongSAGE analysis of skeletal muscle at three prenatal stages in Tongcheng and Landrace pigs,” *Genome Biology*, vol. 8, no. 6, p. R115, 2007.

[8] L. Beekman, T. Tohver, R. Dardari, and R. Légueillète, “Evaluation of suitable reference genes for gene expression studies in bronchoalveolar lavage cells from horses with inflammatory airway disease,” *BMC Molecular Biology*, vol. 12, no. 1, p. 5, 2011.

[9] M. Jain, A. Nijhawan, A. K. Tyagi, and J. P. Khurana, “Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations,” *BMC Genomics*, vol. 9, no. 1, p. 7, 2008.

[10] X. Zhou, J. Liu, and Y. Zhuang, “Selection of reference genes for normalization of real-time PCR data in minipig heart failure model and evaluation of TNF-α mRNA expression,” *Journal of Biotechnology*, vol. 153, no. 3-4, pp. 92–99, 2011.

[11] Y. Wang, Y. Zhao, J. Li et al., “Evaluation of housekeeping genes for normalizing real-time quantitative PCR assays in pig skeletal muscle at multiple developmental stages,” *Gene*, vol. 565, pp. 235–241, 2015.

[12] C. P. P. Tricarico, P. Pinzani, S. Bianchi et al., “Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies,” *Analytical Biochemistry*, vol. 390, no. 2, pp. 293–300, 2002.

[13] J. B. de Kok, R. W. Roelofs, B. A. Giesendorf et al., “Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes,” *Laboratory Investigation*, vol. 85, no. 1, pp. 154–159, 2005.

[14] S. Selvey, E. W. Thompson, K. Matthaei, R. A. Lea, M. G. Irving, and L. R. Griffiths, “β-Actin—an unsuitable internal control for RT-PCR,” *Molecular & Cellular Probes*, vol. 15, no. 5, pp. 307–311, 2002.

[15] E. M. Glare, M. Divjak, M. J. Bailey, and E. H. Walters, “β-Actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalising mRNA levels,” *Thorax*, vol. 57, no. 9, pp. 765–770, 2002.

[16] A. de Greeff, L. Benga, P. J. Wichgers Schreur, P. Valentin-Weigand, J. M. J. Rebel, and H. E. Smith, “Involvement of NF-kB and MAP-kinases in the transcriptional regulation of alveolar macrophages by *Streptococcus suis*,” *Veterinary Microbiology*, vol. 141, no. 1-2, pp. 59–67, 2010.

[17] X. Gong, X. Li, X. You et al., “AIIF1 was identified as an up-regulated gene contributing to CSFV Shimen infection in porcine alveolar macrophage 3D4/21 cells,” *PeerJ*, vol. 8, article e8543, 2020.

[18] J. Zhang, X. Lei, D. Wang et al., “Inhibition of Abi or Src tyrosine kinase decreased porcine circovirus type 2 production in PK15 cells,” *Research in Veterinary Science*, vol. 124, pp. 1–9, 2019.

[19] X. Shen, L. Yin, X. Pan, R. Zhao, and D. Zhang, “Porcine epidemic diarrhea virus infection blocks cell cycle and induces apoptosis in pig intestinal epithelial cells,” *Microbial Pathogenesis*, vol. 147, article 104378, 2020.

[20] C. L. Andersen, J. L. Jensen, and T. F. Ørntoft, “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 Research*, vol. 64, no. 15, pp. 5245–5250, 2004.