HYPOXANTHINE GUANINE PHOSPHORIBOSYL TRANSFERASE IS THE MOST STABLE REFERENCE GENE FOR GENE EXPRESSION ANALYSIS BY QUANTITATIVE PCR IN PERIPHERAL BLOOD MONONUCLEAR CELLS FROM WOMEN WITH THE POLYCYSTIC OVARY SYNDROME

HIPOKSANTIN GUANIN FOSFORIBOZIL TRANSFERAZA JE NAJSTABILNIJI REFERENTNI GEN ZA ANALIZU EKSPRESIJE GENA KVANTITATIVNIM PCR-OM U PERIFERNIM MONONUKLEARNIM ĆELIJAMA ŽENA SA SINDROMOM POLICISTIČNIH JAJNIKA

Danijela Vojnović Milutinović1, Djuro Macut2, Ivana Božić Antić2, Jelica Bjekić Macut3, Marina Nikolić1, Gordana Matić1, Jelena Nestorov1

1Department of Biochemistry, Institute for Biological Research »Siniša Stanković«, University of Belgrade, Belgrade, Serbia
2Clinic for Endocrinology, Diabetes and Metabolic Diseases, Clinical Center of Serbia and Faculty of Medicine, University of Belgrade, Belgrade, Serbia
3CHC Bežanijska kosa, Belgrade, Serbia

Summary

Background: The polycystic ovary syndrome (PCOS) is a frequent endocrine disorder that affects women of reproductive age. As the syndrome is strongly associated with obesity, it is of interest to examine the gene expression differences that accompany its development and the associated metabolic disturbances. Real-time RT PCR is a standard method for studying changes in gene expression. However, to obtain accurate and reliable results, validation of reference genes is obligatory. The aim of this study was to identify a suitable reference for the normalization of gene expression in peripheral blood mononuclear cells (PBMCs) from obese and normal-weight women with PCOS.

Methods: The expression stability of four potential reference genes: hypoxanthine guanine phosphoribosyl transferase 1 (HPRT), β-actin (BA), β2-microglobulin (B2M) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), was assessed in PBMCs from healthy women, and from normal-weight and obese women with PCOS. The variabilit-
ity in the expression of potential reference genes was ana-
lized by the TaqMan real-time RT PCR method, using GeNorm and NormFinder software packages.

**Results:** Direct comparison of cycle threshold (Ct) values showed inter-individual variations for all validated genes, the Ct values of HPRT being less variable than those of BA, GAPDH and B2M. Both software packages pointed to HPRT as the most steadily expressed gene in the PBMCs of women with PCOS and healthy controls.

**Conclusions:** Cross-validation of the expression stability of four potential reference genes identified HPRT as the most stable reference, suitable for further investigations of gene expression in PBMCs from women with PCOS.

**Keywords:** HPRT, PCOS, peripheral blood mononuclear cells, real-time RT PCR, reference gene

---

**Introduction**

Polycystic ovary syndrome (PCOS) is a frequent endocrinopathy that affects 6–8% of women of reproductive age (1). The major determinants of PCOS are hyperandrogenism, chronic anovulation and polycystic ovaries (2). Usually, PCOS is associated with metabolic disturbances such as abdominal obesity, hyperinsulinemia and dyslipidemia (3). Approximately 40–60% of women with PCOS are overweight or obese, and obesity has profound effects on both the pathophysiology and the clinical manifestation of this disorder (4–6). The mechanisms underlying the PCOS remain largely unclear. Therefore, it is of interest to examine the differences in gene expression that may contribute to the different pathophysiological aspects of PCOS and the associated metabolic disturbances including obesity.

Real-time RT PCR is a convenient method for studying changes in gene expression, often used in biomedical research (7, 8). However, its accuracy and reliability critically depend on the selection of a normalization strategy (9, 10). Normalizing target gene expression to a stably expressed reference gene is the most commonly used normalization approach. The use of an inappropriate reference gene can lead to under- or over-estimation of the target gene expression level and to misinterpretation of the results, that might have significant implications (10–12). Therefore, to obtain an accurate result of the comparison of mRNA levels between different samples, tissues, or pathophysiological conditions, it is crucial to choose a suitable reference gene.

Due to the complex nature of PCOS, it is of great importance to clearly define experimental groups and to include validation of reference genes in the process of evaluating target gene expression. In this study, we examined four candidate reference genes: \( \beta_2 \)-microglobulin (B2M), \( \beta \)-actin (BA), hypoxanthine guanine phosphoribosyl transferase1 (HPRT) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), in order to identify a suitable reference for the normalization of gene expression in peripheral blood mononuclear cells (PBMCs) from obese and normal-weight women with PCOS.

**Materials and Methods**

**Subjects**

This study was performed on three groups of subjects: (a) normal-weight women with PCOS (BMI < 25 kg/m²), (b) obese women with PCOS (BMI ≥ 30 kg/m²), and (c) normal-weight controls. Eighteen women participated in the study (6 per group).

The PCOS diagnosis was established according to the Rotterdam/ESHRE consensus criteria (13) including 2 of the following 3 criteria: 1) oligomenorrhea or anovulation; 2) clinical and/or biochemical signs of hyperandrogenism, and 3) polycystic ovaries. Oligomenorrhea was defined as bleeding intervals between 35 days and 6 months and amenorrhea was defined as a bleeding interval greater than 6 months (14). Level of serum progesterone of less than 10 nmol/L was used for defining anovulation. In women with regular menstrual cycles, anovulation was established when two consecutive serum progesterone levels in the luteal phase were less than 10 nmol/L. Modified Ferriman-Gallwey (F-G) score higher than 6 was used to diagnose hirsutism (15). Based on a previous report by Macut et al. (16), hyperandrogenemia was defined by the level of serum total testosterone higher than 2 nmol/L. Ovarian morphology was examined by transvaginal ultrasonography, and presence of increased ovarian volume and/or at least twelve follicular cysts measuring 2–9 mm were used for confirming their polycystic structure (17). All the subjects with regular menstrual cycles were evaluated during the early follicular phase. All relevant drugs (oral contraceptives, glucocorticoids, antiandrogens, ovulation
inducing agents, antidiabetic and antiobesity drugs, or other hormonal therapy) were excluded or prohibited for at least 3 months preceding the study. Excluding criteria in this study were: hyperglycemia, thyroid disease, pregnancy, and prolactin-, cortisol- or androgen-secreting tumors. Study participants were consecutively recruited during a one-year period from the Outpatient Endocrinology Clinic, where they were referred for evaluation of menstrual cycle disturbances, hirsutism, acne or fertility problems.

All the subjects gave written informed consent for participating in this study and the Ethical Committee of the Faculty of Medicine, University of Belgrade, approved the protocol.

**Biochemical and hormonal measurements**

After overnight fasting, blood samples were drawn at 08:00 h and stored at –20 °C until use. Serum total testosterone (nmol/L) and sex hormone-binding globulin (SHBG, nmol/L) concentrations were determined by radioimmunoassay using TESTO-CT2 and SHBG-RIACT kits (CIS bio international, Gif-Sur-Yvette Cedex, France). The intra- and inter-assay coefficients of variation were: 4.5 and 5.1% for testosterone, and 3.9 and 4.7% for SHBG. Standard formula \[
\frac{\text{testosterone/SHBG}}{100}
\] was used for calculation of the free androgen index (FAI).

**RNA isolation and cDNA synthesis**

PBMCs were prepared from fresh blood by centrifugation through Ficoll-Paque PLUS (Amersham, UK). Total mRNAs were isolated from PBMCs by an RNeasy mini kit (Qiagen). Concentration and purity were tested spectrophotometrically (OD 260/280 higher than 1.8 was considered acceptable). The integrity of RNA was assessed by agarose gel electrophoresis. After adding the RNase inhibitor (Applied Biosystems), samples were frozen at –80 °C until use. cDNA was synthesized from 2 μg of RNA after removal of DNA by DNase I treatment (Fermentas). Reverse transcription reaction was performed using a High Capacity cDNA Reverse Transcription kit (Applied Biosystems). cDNA was stored at –80 °C until use.

**Real-time PCR**

Real-time PCR was performed using the ABI Prism 7000 Sequence Detection System (Applied Biosystems). The reaction mixture contained 1 × TaqMan Universal Master Mix with AmpErase UNG, 1 × Assay Mix (Applied Biosystems) and cDNA template (20 ng of RNA converted to cDNA). Primers and probes for the glucocorticoid receptor (GR, Hs00230813_m1), BA (Hs99999903_m1), B2M (Hs99999907_m1), GAPDH (Hs99999905_m1) and HPRT (Hs99999909_m1) were obtained from Applied Biosystems.

The cycle conditions were: 95 °C for 10 min and subsequently 40 cycles at 95 °C for 15 s and 60 °C for 90 s. All the reactions were run in triplicates. For the evaluation of PCR efficiencies of all the examined genes and GR as a target gene, serial dilutions of one randomly chosen cDNA were made and amplified, and standard curves were constructed.

**Data analysis**

Data are presented as mean values ± SD. Between group comparisons of anthropometrical and basal biochemical parameters were made by non-parametric Kruskal-Wallis H test and subsequently by Mann-Whitney U test (SPSS v13.0, SPSS INC., Chicago, IL, USA). A value of P<0.05 was considered statistically significant.

**Results**

All the anthropometrical and hormonal parameters are presented in Table I. In comparison to nor-

<table>
<thead>
<tr>
<th>Variable</th>
<th>Controls</th>
<th>Normal-weight PCOS</th>
<th>Obese PCOS</th>
<th>(p_a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>29.17±4.16</td>
<td>23.50±3.93</td>
<td>28.67±5.05</td>
<td>0.21</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.67±2.88b</td>
<td>22.97±2.47b</td>
<td>32.82±2.80</td>
<td>0.003</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>80.00±8.74b</td>
<td>80.83±13.53c</td>
<td>102.17±7.55</td>
<td>0.012</td>
</tr>
<tr>
<td>Waist-to-hip ratio</td>
<td>0.81±0.05</td>
<td>0.83±0.04</td>
<td>0.88±0.07</td>
<td>0.13</td>
</tr>
<tr>
<td>Testosterone (nmol/L)</td>
<td>1.58±0.59</td>
<td>3.18±0.52d</td>
<td>2.55±0.78e</td>
<td>0.01</td>
</tr>
<tr>
<td>SHBG (nmol/L)</td>
<td>56.50±13.86</td>
<td>32.84±9.89e</td>
<td>19.02±9.41d</td>
<td>0.005</td>
</tr>
<tr>
<td>FAI</td>
<td>3.06±1.57</td>
<td>10.01±1.91d</td>
<td>17.50±6.76d</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Values are means ± S.D. BMI, body mass index; SHBG, sex hormone-binding globulin; FAI, free androgen index. aNonparametric Kruskal-Wallis H test followed by Mann-Whitney U test if \(P<0.05\); b\(P<0.01\) vs obese PCOS; c\(P<0.05\) vs obese PCOS; d\(P<0.01\) vs Controls; e\(P<0.05\) vs Controls.
mal-weight PCOS women and healthy controls, the obese women with PCOS had a significantly higher BMI and waist circumference. As expected, total testosterone and FAI were significantly higher, while SHBG was significantly lower in both PCOS groups than in controls.

In order to evaluate the expression stability of selected potential reference genes in PBMCs from obese and normal-weight women with PCOS, first we calculated the PCR amplification efficiencies. For the construction of a standard curve, serial dilutions of one randomly chosen cDNA sample were amplified. The slope of a standard curve (semi-log regression line of Ct vs log of cDNA input) is used for estimating the amplification efficiency of the PCR reaction (Figure 1). Slope of –3.32 corresponds to 100% efficiency. Amplification efficiencies (E) of the candidate reference genes and GR, as a target gene, were calculated using the formula: $E = (10^{-1/slope} - 1) \times 100$, and are presented in Table II. Amplification efficiency values ranged from 91% to 104% (slope –3.23 to –3.55), and $R^2$ values of the standard curves were over 0.993.

For relative quantification by the ΔΔCt method, the amplification efficiencies of the reference and target gene should be approximately equal. The easiest way to check whether two genes have similar amplification efficiencies is to examine how the Ct value ($Ct_{target\ gene} - Ct_{reference\ gene}$) varies with sample dilution. The two genes have equal efficiencies if the slope of semi-log plot (Ct vs log of cDNA input) is less than 0.1. In this study, all the examined reference genes performed well, as compared to the GR as a target gene (data not shown).

Expression stability of potential reference genes was analyzed by the direct comparison of Ct values and by using GeNorm (18) and NormFinder (19) software. First, the individual Ct values of the examined reference genes were compared. Expression profiles of B2M, BA, GAPDH and HPRT are presented in Figure 2. Raw data displayed a wide expression range, with Ct values between 21 and 31 (Figure 2). B2M and BA were the most abundant genes, followed by GAPDH, while the least abundant gene in the PBMCs of all the examined subjects was HPRT. Each gene showed marked inter-individual variations, with stan-

Figure 1 Amplification efficiencies of candidate reference genes (B2M, BA, GAPDH and HPRT) and GR as a target gene.

Table II Gene symbol, molecular function and amplification efficiency of GR and four candidate reference genes.

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Gene Name</th>
<th>Function</th>
<th>Amplification efficiency E (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA</td>
<td>β-Actin</td>
<td>Cytoskeletal structural protein</td>
<td>96</td>
</tr>
<tr>
<td>B2M</td>
<td>β2-Microglobulin</td>
<td>β-Chain of major histocompatibility complex class I molecule</td>
<td>91</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3 phosphate dehydrogenase</td>
<td>Glycolytic enzyme</td>
<td>104</td>
</tr>
<tr>
<td>HPRT</td>
<td>Hypoxanthine guanine phosphoribosyl transferase 1</td>
<td>Glycosyl transferase</td>
<td>97</td>
</tr>
<tr>
<td>GR</td>
<td>Glucocorticoid receptor</td>
<td>Transcription factor</td>
<td>101</td>
</tr>
</tbody>
</table>
standard deviations ranging from 0.9 for HPRT and B2M to 1.2 for GAPDH (Figure 3).

The expression stability of the selected reference genes was further assessed using the GeNorm and the NormFinder software (Table III). GeNorm calculates the stability value M using a pair-wise comparison approach, thus grading the potential reference genes according to the similarity of expression profiles across the sample set (18). Low M value corresponds to high expression stability. All the investigated genes exhibited high expression stability with M-values less than 0.44, which is below the algorithm defined cut-off value of 1.5 (Table III) (18). After a step-wise exclusion of the least stable genes (highest M-values), M-values were recalculated. GeNorm identified the combination of HPRT

---

**Figure 2** Direct comparison of Ct values. Ct values of each gene and each individual subject are presented. C – controls, P – normal-weight women with PCOS, OP – obese women with PCOS.

**Figure 3** The Ct value ranges of potential reference genes. Box represents second and third quartiles, while the median is shown as a central line, and whiskers extend to smallest and largest non-outlier values.

**Table III** Expression stability values of BA, B2M, GAPDH and HPRT calculated by GeNorm and NormFinder software.

<table>
<thead>
<tr>
<th>Gene</th>
<th>GeNorm Stability value (M)</th>
<th>NormFinder Stability value (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA</td>
<td>0.368</td>
<td>0.024</td>
</tr>
<tr>
<td>B2M</td>
<td>0.373</td>
<td>0.038</td>
</tr>
<tr>
<td>GAPDH</td>
<td>0.435</td>
<td>0.031</td>
</tr>
<tr>
<td>HPRT</td>
<td>0.338</td>
<td>0.023</td>
</tr>
</tbody>
</table>

**Figure 4** GeNorm output: gene expression stability of BA, GAPDH, B2M and HPRT. Reference genes are ranked in order of their expression stability and presented at x axis. Stability values (M) are presented at y axis.
and B2M as the most stable one, with the combined M-value equal to 0.26 (Figure 4).

NormFinder uses a model-based approach for identifying the most stable genes based on the least inter- and intra-group expression variations (19). In contrast to GeNorm, NormFinder pointed to the HPRT and BA combination as the most stable one in PBMCs of women with PCOS and healthy controls, with its stability value of 0.022. As for the individual genes, HPRT was identified as the most stable gene, followed by BA, while B2M was identified as the least stable one (Table III).

Both software packages pointed to HPRT as the most stable gene in PBMCs of women with PCOS and healthy controls, while showing differences in the ranking order of all the other potential references, as well as in the most stable combinations (Table III).

**Discussion**

By using the TaqMan real-time RT PCR method, we determined variability in the expression of four potential reference genes: BA, B2M, GAPDH and HPRT, in order to identify the most stable reference for further investigations of gene expression in PBMCs from normal-weight and obese women with PCOS. We analyzed amplification efficiencies and gene expression stabilities of the potential reference genes. Our study shows that HPRT can be used as a reference gene for the normalization of gene expression in PBMCs from normal-weight and obese women with PCOS.

Real-time PCR is the standard method for studying gene expression, but the need for accurate data normalization is vital. Thus, the selection of an appropriate endogenous control represents an important step in the analysis of gene expression (10, 12). Currently, for the determination of target gene expression in different pathophysiological states, most of the researchers use relative quantification methods, either by using the standard curve method or the $2^{-\Delta\Delta C_T}$ method. Both of these methods require a properly selected reference gene for the normalization of target gene expression.

Validation of reference genes represents a time-consuming and expensive procedure. Yet, it has been well documented that using non-validated reference genes may result in incorrect data and misinterpretation of the results (10, 12). Therefore, before performing RT PCR experiments, it is important to carefully examine all the available data on the suitability of various reference genes that were previously used in similar experimental conditions, in order to select potential candidates that are going to be validated for the specific experimental design. The potential candidate genes should be selected so as to avoid possible co-regulations, and the best way is to use genes belonging to different functional classes. The gene(s) which is (are) going to be used as a reference in a particular experimental design should be expressed at a constant level between the samples through all experimental conditions or disease stages (20). In this study, we investigated four potential reference genes that belong to different functional classes. Namely, HPRT is an enzyme involved in nucleotide metabolism, GAPDH is an essential glycolytic enzyme, B2M is the small subunit of a major histocompatibility complex class I molecule and BA is a ubiquitous cytoskeleton protein.

Increasing concern regarding normalization using traditional reference genes has led to the development of several statistical models and software packages for the analysis of candidate gene stability, in order to help identify the best reference genes (18, 19, 21). GeNorm is the most frequently used software for the analysis of candidate gene stability. This algorithm operates on the assumption that the ratio of two ideal reference genes should be constant regardless of the experimental conditions. GeNorm uses a pair-wise comparison approach and ranks genes according to the similarity of their expression profiles, rather than minimal variation (18). On the other hand, NormFinder uses a model-based approach for the identification of the most stable genes based on the least inter- and intra-group expression variations (19).

In this study, we analyzed the expression stability of four commonly used reference genes (HPRT, BA, GAPDH and B2M) and showed that HPRT can be used as a suitable reference for further studies of gene expression in PBMCs of women with PCOS, and, in addition, for studies focusing on the connection of PCOS and obesity. All the examined reference genes in this study had amplification efficiencies close to 100%, and all performed well as compared to the GR, taken as an example target gene. The direct comparison of Ct values of BA, B2M, GAPDH and HPRT showed marked inter-individual variations for all four validated genes, whereas the Ct values of HPRT were among those less variable. The expression stability was further assessed by GeNorm and NormFinder and both softwares identified HPRT as the most stable gene, while the ranking order of other examined genes differed. Namely, GeNorm identified GAPDH, while NormFinder pointed to B2M as the least stable gene. Moreover, the best combination calculated by GeNorm was HPRT and B2M, while HPRT and BA appeared to be the best combination according to NormFinder. The observed discrepancy in the ranking of gene stabilities between the two algorithms is not unusual and may be explained by the different methodologies used to calculate the gene stability.

Although the GeNorm and NormFinder algorithms significantly simplified the time-consuming
process of data analysis during the selection of suitable reference genes, these approaches are still underutilized. Instead, in most of the previous studies, traditional housekeeping genes were used as references, without any validation. To date, only a few research groups used validated reference genes for gene expression studies in various tissues from women with PCOS, while most of the authors used GAPDH and 18S RNA without prior validation. Sadek et al. (22) validated nine reference genes and suggested using tyrosine-3 monooxygenase/tryptophan-5 monooxygenase activation protein, zeta polypeptide (YWHAZ), cytochrome c1 (CYC1) and BA as suitable references for gene expression studies on endometrial tissue from women with PCOS. HPRT and B2M were not among the nine validated endogenous controls. GAPDH was shown as unsuitable, while BA was one of the most stable genes (22). Lindholm et al. (23) investigated the gene expression pattern of inflammation markers in abdominal superficial subcutaneous tissue from overweight and lean patients with PCOS and overweight controls. The authors used the NormFinder algorithm to examine the expression stability of low density lipoprotein receptor-related protein 10 (LRP10), peptidylprolyl isomerase A (PPIA) and acidic ribosomal phosphoprotein P0 (RPLP0) and identified RPLP0 as the best reference. Up to now, there have been no reports on reference gene stability in the PBMCs of patients with PCOS. Our study shows that HPRT can be used for the normalization of gene expression in PBMCs from normal-weight and obese women with PCOS.

**Conclusion**

Cross-validation of the expression stability of four potential reference genes by the direct comparison of Ct values and by using two software packages, GeNorm and NormFinder, identified HPRT as the most stable reference gene for gene expression studies in PBMCs from normal-weight and obese women with the polycystic ovary syndrome.

**Acknowledgements.** This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, Grant III41009.

**Conflict of interest statement**

The authors stated that they have no conflicts of interest regarding the publication of this article.

**References**


Received: November 4, 2013
Accepted: December 28, 2013