Selection of optimized candidate reference genes for qRT-PCR normalization in rice (Oryza sativa L.) during Magnaporthe oryzae infection and drought

R. Bevitori¹, M.B. Oliveira², M.F. Grossi-de-Sá³, A.C. Lanna¹, R.D. da Silveira² and S. Petrofeza²

¹Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, Brasil
²Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, GO, Brasil
³Laboratório de Interação Molecular Planta-Praga, Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, Brasil

Corresponding author: S. Petrofeza
E-mail: petrofez@uol.com.br

Received November 11, 2013
Accepted June 17, 2014
Published November 27, 2014
DOI http://dx.doi.org/10.4238/2014.November.27.7

ABSTRACT. Drought and rice blast disease caused by Magnaporthe oryzae are two of the most serious threats to global rice production. To explore the mechanisms underlying gene expression induced in rice by stresses, studies involving transcriptome analyses have been conducted over the past few years. Thus, it is crucial to have a reliable set of reference genes to normalize the expression levels of rice genes affected by different stresses. To identify potential reference genes for studies of the differential expression of target genes in rice under M. oryzae infection and drought conditions, the present study evaluated five housekeeping genes for the normalization of gene expression. The stability of the expression of these genes was assessed using the analytical software packages geNorm and NormFinder. For all samples analyzed, the stability rank was UBQ5.
> GAPDH > eIF-4α > β-TUB > 18S rRNA. The data showed that the UBQ5, GAPDH, and eIF-4α genes are appropriate, high-performing reference genes and will be highly useful in future expression studies of fungal infections and drought in rice.

**Key words:** Housekeeping genes; Gene expression; Stable expression; Real-time quantitative reverse transcription polymerase chain reaction; Normalization

**INTRODUCTION**

Rice blast disease, which is caused by the ascomycetous hemibiotrophic fungus *Magnaporthe oryzae*, is a major constraint on rice production worldwide, causing yield losses of up to 100% that are influenced by cultivar susceptibility, environmental conditions, and disease management systems. Due to the unrestricted public availability of genome sequences, genetic stocks, and improved research tools, this pathosystem has become the premier model for understanding the molecular basis of plant-fungal interactions (Wang et al., 2010). Although significant efforts are underway to develop rice varieties with durable resistance to blast, and more than 40 major genes and quantitative trait loci for blast resistance have been identified (Ballini et al., 2008), the molecular basis of the defense response to rice blast remains poorly characterized.

Rice is also subject to drought stress, which can negatively affect plant growth and productivity, and drought has become an increasingly severe problem worldwide (Passioura, 2007). Abiotic stress is, therefore, a major constraint on rice production in water-limited environments (Bernier et al., 2008). Due to tremendous efforts over the past decade, rice drought tolerance has been scrutinized using the quantitative trait loci (QTL) mapping approach (Fu et al., 2007), and the involvement of many genes in the plant responses to drought have been identified. Understanding the molecular basis of drought stress tolerance will contribute to the breeding of elite rice cultivars with reduced water requirements.

In plants, as in other organisms, environmental conditions induce the expression of a variety of genes, and the products of those genes are thought to promote stress tolerance, regulate gene expression, and activate various pathways (Shinozaki et al., 2003). Candidate gene identification through rice functional genomics has great potential for the development of more durable and resistant/tolerant cultivars (Khush and Jena, 2009). Studying alterations in the expression of candidate genes is important for understanding the mechanisms involved in plant responses to rice blast disease and drought.

Analysis of gene expression requires sensitive, precise, and reproducible quantification of specific messenger RNAs (mRNAs). Among the methods commonly used to determine levels of gene expression, real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) is the method of choice, and due to its high sensitivity, specificity, and ability to provide reproducible quantification of nucleic acids, this technique has been considered a useful and available method for evaluating changes in gene expression (Bustin et al., 2005, 2009; Artico et al., 2010). However, the accuracy of qRT-PCR is influenced by a number of variables, including variability in the RNA samples, the amount of RNA, and the RNA extraction protocol. For this reason, it is essential to normalize qRT-PCR data using suitable internal reference genes so that target-gene expression analyses are accurate and reliable (Andersen et al., 2004; Czechowski et al., 2005). Thus, the particular reference gene acts as a control to which com-
Selection of reference genes in rice

Comparisons are made for the validation of changes in the expression of a target gene.

Ideally, the reference gene should be expressed at constant levels in all tissues and at all developmental stages, regardless of the experimental conditions or treatments (Vandesompele et al., 2002; Radonic et al., 2004). The choice of reference gene(s) for normalization is critical to ensure the reliability and validity of qRT-PCR assays.

Reference genes that are related to basic and structural cell processes [e.g., albumin, β-tubulin (TUB), β-actin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ribosomal subunits, ubiquitin, and cyclophilin], called housekeeping genes, have been used as normalizers for quantitative assays (Tricarico et al., 2002). However, many studies reported variable expression levels of these types of genes under different conditions and in different tissues (Livak and Schmittgen, 2001; Czechowski et al., 2005).

Various studies of different plant species have defined reference genes for different organs (Nicot et al., 2005; Jain et al., 2006; Gutierrez et al., 2008; Hu et al., 2009; Qi et al., 2010; Manoli et al., 2012) or different environmental conditions (Nicot et al., 2005; Reid et al., 2006; Barsalobres-Cavallari et al., 2009; Hu et al., 2009). However, few studies regarding the validation of reference genes in rice have been reported (Kim et al., 2003; Jain et al., 2006; Narsai et al., 2010; Wang et al., 2010). Most reference genes used for validating gene expression in rice are genes used in other plants or the orthologs of defined reference genes, such as those of Arabidopsis, including 18S ribosomal RNA (18S rRNA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ubiquitin 5 (UBQ5), and elongation factor 4 alpha (EF-1α) (Kim et al., 2003; Jain et al., 2006). Recently, novel reference genes for transcript normalization in rice have been identified by a genome-wide analysis of whole-genome GeneChip data, such as endo-1,4-beta glucanase, the AP-2 complex subunit, and SKP1-like protein 1A (Narsai et al., 2010).

The current study assessed the effects of compatible and incompatible M. oryzae isolates, and drought conditions on the performance of five candidate reference genes commonly used in expression studies via qRT-PCR analyses of rice leaf samples collected over the course of infection. For this purpose, genes reported to have stable expression in plants grown under various environmental conditions (Jain et al., 2006) such as 18S rRNA, UBQ5, GAPDH, β-tubulin (β-TUB), and eukaryotic initiation factor 4 alpha (eIF-4α) were used. The data showed that the UBQ5, GAPDH, and eIF-4α genes are high-performing reference genes under both conditions. This study will benefit future studies on gene expression in rice under fungal infection and drought stress conditions.

MATERIAL AND METHODS

Candidate reference gene selection and primer design

The reference genes used in this study were obtained through bibliographic reviews of studies involving rice or related plants under biotic or abiotic stresses and subsequent in silico analysis using the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) of the National Center for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi). For example, the BLASTN tool was used to obtain the expressed sequence tag (EST) homologs in rice of a previously selected EST. Subsequently, we searched the NCBI non-redundant protein sequence database using the BLASTX tool to confirm the function of the EST. The name of the gene and gene product, accession number, primer sequences, and expected size of the amplicon are provided in Table 1.
The Primer3 web tool (http://frodo.wi.mit.edu/primer3/) was used to design the primers, with the following parameters: (a) product size range of 100-200 bp, (b) primer size of 18-22 bp, and (3) primer temperature of 57°-65°C. Amplification of single products of the expected size was verified by electrophoresis on 2% agarose gel and sequencing.

**Plant material and *M. oryzae* infection**

Seeds of the cultivar Metica were grown in plastic trays (15 x 30 x 10 cm) containing 3 kg field soil fertilized with macro- and micronutrients, and watered as necessary. Py-435 (avirulent, incompatible reaction) and Py-1059 (virulent, compatible reaction) fungi were used to induce resistance and susceptibility, respectively. Three trays were planted for each fungal isolate. For inoculation of the avirulent and virulent isolates of *M. oryzae*, 22-day-old plants were sprayed with an aqueous spore suspension at a concentration of 3 x 105 conidia/mL until run-off using an atomizer connected to an air compressor, as described by Prabhu et al. (1992). The inoculated plants were incubated at 25°C in a humidified plastic chamber in the dark to induce and maintain dew formation on the plants for periods of 4 and 24 h. The same procedure was performed for the control plants. Next, fungus-treated and mock-treated leaves were collected. The penultimate leaf of each plant from all of the replicates for each treatment was cut because it is the first leaf to exhibit disease symptoms; each of these leaves was flash-frozen in liquid nitrogen and stored at -80°C until use for RNA extraction.

**Plant material and drought treatment**

For the abiotic stress treatment, the Brazilian cultivars BRS Douradão and BRS Primavera, which exhibit different levels of drought tolerance (Lorençoni et al., 2010; Guimarães et al., 2011), were planted in pots containing native soil from the Cerrado, and fertilizer was applied to adequately nourish the rice plants. The experiment was designed according to Lanna et al. (2013) and consisted of a randomized block of split plots with six replicates for each cultivar. The plants were watered until they reached vegetative stage V3 (the beginning of tillering, as classified by Counce et al., 2000). Irrigation was discontinued when the plants reached stage V3 (11 and 12 days after planting for the BRS Primavera and BRS Douradão cultivars, respectively) and was restarted when the plants reached vegetative growth stage V6 (29 and 31 days after planting for the BRS Primavera and BRS Douradão cultivars, respectively). To establish the amount of water in the pot for each water treatment, the field capacity of the soil was determined and used to maintain two water treatments as follows: 1) one set of plants was watered normally (control), and 2) a second set of plants was subjected to 70% water restriction (or provided with soil that was 30% water). The leaves from the plants exposed to each water treatment were collected at stage V6, and RNA extraction was performed as described below.

**Total RNA extraction and cDNA synthesis**

Total RNA was extracted from the leaves using an RNeasy Kit (Qiagen, Valencia, CA, USA). The RNA was treated with RNase-free DNase (Qiagen) according to manufacturer instructions, and the RNA quality was determined using a 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). The first-strand cDNAs were synthesized using 1 µg
DNase-treated total RNA, specific primers, and qRT-PCR SuperMix Kit components (Invitrogen, Grand Island, NY, USA) according to manufacturer instructions.

**Quantitative cDNA amplification by qRT-PCR**

The quantitative amplification reactions for the candidate reference genes were performed using a StepOnePlus Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) with 10-μL one-step reaction mixtures containing 0.4 μM each primer, 6 μL SYBR Green PCR Master Mix (2X), and 0.2 μL cDNA (1:10 dilution). The reaction conditions were as follows: 10 min at 42°C and, after an initial denaturation at 95°C for 10 min, 40 cycles at 95°C for 15 s and 60°C for 1 min. To evaluate the specificity of the PCR product, each data point on the melting curves was analyzed. Three technical replicates were performed. The cycle threshold (Ct) value was defined as the cycle in which there was a significant increase in the amount of the PCR product. The relative quantities of each transcript were determined by interpolation using standard curves. Both the Ct values and relative quantities of each sample were acquired using the StepOnePlus Real-Time PCR System.

The real-time PCR efficiency ($E$) was determined for each gene and each stress condition using the slope of a linear regression model (Pfaffl, 2001). Thus, the cDNA samples were bulked for use as PCR templates. $E$ was calculated using the formula $E = 10^{\text{slope}}$ (Radonic et al., 2004), and the correlation coefficient ($R^2$) was calculated for each gene. For each gene, the PCR efficiency was determined using the associated Ct, which was obtained by serially diluting the bulk cDNA until the detection level reached a particular threshold.

**Determination of the stability of the expression of the reference genes**

To select a suitable reference gene, the stability of each candidate reference gene mRNA expression under the two experimental conditions was statistically analyzed using two programs: geNorm (Version 3.5) (Vandesompele et al., 2002) and NormFinder (Andersen et al., 2004).

**RESULTS AND DISCUSSION**

**Expression profiling of candidate reference genes**

cDNA synthesized from total RNA, which was isolated from the rice leaves under biotic and abiotic stresses, was used to compare the expression levels of five different reference genes. Amplification of the specific transcript was confirmed by the appearance of a single peak in the melting curve analysis following completion of the amplification reaction. Amplification curves for each gene were generated, and all of the samples for the biotic or abiotic stress conditions were grouped.

The stability of gene expression was determined by quantifying the associated mRNA levels using qRT-PCR, and the Ct value was calculated. Transcripts of GAPDH, β-TUB, 18S rRNA, elf-4α, and UBQ5 were detected in all of the tested samples. 18S rRNA was the most abundantly transcribed, with Ct values ranging from 18.9 to 22.7. GAPDH, elf-4α, and β-TUB showed the highest Ct values, and UBQ5 yielded moderate Ct values (Figure 1).
The PCR efficiencies for the reference genes tested in this study ranged from 90 to 99%, and the linear correlation values of the five constructed standard curves were all >0.998 (Table 1).

Expression stabilities of the candidate reference genes

Herein, the expression stabilities of five candidate reference genes in rice samples were evaluated and analyzed using the geNorm program. This software calculates a gene stability measurement (M) value based on the average pairwise variation among all of the genes tested. The gene with the lowest M value has the most stable expression, and the gene with the highest M value has the least stable expression. An average expression stability value (M) of 1.5 indicates that a gene’s expression is relatively stable (Vandesompele et al., 2002).

The geNorm analysis showed that the most stably expressed gene throughout the abiotic (Figure 2A) and biotic treatments was the UBQ5 gene (Figure 2B). GAPDH and eIF-4α also showed high expression stabilities in samples under biotic and abiotic conditions, respectively. 18S rRNA was the least stably expressed in all of the groups. All of the reference genes were suitable for expression analysis under these experimental conditions because these genes all reached M values lower than the recommended cutoff value of 1.5. Therefore, each of these genes may be used as a single reference gene for multiple treatments.

Furthermore, the same data sets from the geNorm analyses were analyzed using the NormFinder program (Andersen et al., 2004) to determine the expression stabilities of the candidate reference genes. As for the geNorm method, the gene with the lowest M value has the most stable expression, and the gene with the highest M value has the least stable expression. NormFinder showed that eIF-4α and UBQ5 were the most suitable reference genes (Figure 3). When the results of the geNorm and NormFinder analyses were compared, minimal differ-
<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>GenBank accession No.</th>
<th>Primer sequences (5'→3')</th>
<th>Amplicon length (bp)</th>
<th>Biotic stress¹</th>
<th>Abiotic stress¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>18S rRNA</td>
<td>18S ribosomal RNA</td>
<td>AK059783</td>
<td>CTACGTCCCTGGCCCTTGTACA ACACCTACCGGACCATCTCA</td>
<td>65</td>
<td>92.3</td>
<td>0.0228</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AGGCCAGCAGCTTACATTTTC</td>
<td></td>
<td>0.9987</td>
<td>0.0136</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate</td>
<td>AK064960</td>
<td>CGAAGGCGCTGGAAGATCTGAGT</td>
<td>79</td>
<td>90.66</td>
<td>0.0287</td>
</tr>
<tr>
<td></td>
<td>dehydrogenase</td>
<td></td>
<td>CGAAGGCGCTGGAAGATCTGAGT</td>
<td></td>
<td>0.9999</td>
<td>0.0127</td>
</tr>
<tr>
<td>eIF-4α</td>
<td>Elongation factor 4 alpha</td>
<td>AK073620</td>
<td>TGTGCTGATGGAAGCTGAGT</td>
<td>76</td>
<td>90.78</td>
<td>0.0176</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GGAAGGCGCTGGAAGATCTGAGT</td>
<td></td>
<td>0.9994</td>
<td>0.0167</td>
</tr>
<tr>
<td>TUB</td>
<td>Beta-tubulin</td>
<td>AK072502</td>
<td>GCTGACCACACCTAGCTTTGG</td>
<td>82</td>
<td>99.25</td>
<td>0.0534</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AGGGAGCTTTGACCTTATGAGT</td>
<td></td>
<td>0.9996</td>
<td>0.0354</td>
</tr>
<tr>
<td>UBQ5</td>
<td>Ubiquitin 5</td>
<td>AK061988</td>
<td>ACCACTTCTGACCGCCACTACT</td>
<td>76</td>
<td>94.35</td>
<td>0.0789</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AGCCTAAGCCTGCTGCTGT</td>
<td></td>
<td>0.9996</td>
<td>0.0579</td>
</tr>
</tbody>
</table>

¹Gene name based on similarity to *Arabidopsis* proteins. ²Average amplicon efficiency for all of the samples under biotic stress. ³Average amplicon efficiency for all of the samples under abiotic stress. *Standard deviation. **Correlation coefficient of the slope of the SD.
ences in the reference gene rankings were observed, but the UBQ5, GAPDH, and elF-4a genes were always among the three most stable genes chosen by both algorithms, indicating that the results were consistent. The most consistent result of the consensus rankings was that under all of the conditions, 18S rRNA ranked at or near the bottom.

Figure 2. Average expression stability values (M) of the five reference genes analyzed using the geNorm program. The greater the M value, the lower the stability. A. Biotic; B. abiotic stresses.

The gene 18S rRNA has been widely used as a reference gene for gene expression analyses for a number of species (Teste et al., 2009; de Almeida et al., 2010). However, herein, the data demonstrated that 18S displayed the lowest expression stability, with the lowest average gene ranking across all of the samples analyzed. Furthermore, the gene Ct value was lower than the value of the other candidates genes tested, implying that the expression level of this gene was several orders of magnitude higher than the expression of the other genes. Therefore, the results of this study suggest that 18S should not be used as a reference in qRT-PCR studies

Figure 3. Gene expression stabilities of the five candidate genes obtained using the NormFinder program.

The gene 18S rRNA has been widely used as a reference gene for gene expression analyses for a number of species (Teste et al., 2009; de Almeida et al., 2010). However, herein, the data demonstrated that 18S displayed the lowest expression stability, with the lowest average gene ranking across all of the samples analyzed. Furthermore, the gene Ct value was lower than the value of the other candidates genes tested, implying that the expression level of this gene was several orders of magnitude higher than the expression of the other genes. Therefore, the results of this study suggest that 18S should not be used as a reference in qRT-PCR studies.
of rice under biotic or abiotic stress conditions. The strongest argument against the use of total rRNA for normalization is the high content of 18S or 28S rRNA molecules in comparison to the molecules of the target mRNA transcripts (Schmittgen and Zakrajsek, 2000; Kim et al., 2003).

**Determination of the optimal number of reference genes**

Although most studies use only a single gene as an internal control for normalization, the use of two or more reference genes for normalization may improve the reliability of the results (Tricarico et al., 2002; Bustin et al., 2005). In the present study, the optimal number of internal control genes for normalization was determined using the geNorm program, which also calculates the pairwise variation, $V_n/V_{n+1}$, to determine the ideal number of reference genes required for accurate normalization (Vandesompele et al., 2002). The $V_n$ values are calculated by the stepwise inclusion of additional reference genes until the $(n + 1)$ gene makes no significant contribution to the newly calculated normalization factor. In this study, the paired-variable $V_2/3$, $V_3/4$, and $V_4/5$ values obtained for both the biotic and abiotic treatments had $V$ values below the cutoff of 0.15, as suggested by Vandesompele et al. (2002) (Figure 4A, B). The results suggest that $UBQ5$ and $eIF-4α$ plus the third most stable gene, $GAPDH$, are appropriate candidates to produce reliable normalization of the target gene expressions levels during biotic and abiotic stresses in rice.

![Figure 4. Determination of the pairwise variation (V) for the optimal number of control genes. The cutoff value for significance was 0.15.](image-url)
Similar results have been observed by Kim et al. (2003). These authors reported that a group of genes, including 18S and 28S rRNA, UBC, UBQ5, and EF-1α, showed the most stable expression in rice seedlings grown under a variety of environmental conditions involving stress and hormone treatments. Furthermore, due to the genes highly stable expression levels, the use of UBQ5 or eIF-4α as an internal control may be sufficient for gene expression studies in rice, although the use of both genes is recommended for more reliable results (Kim et al., 2003).

The present study provides a set of reference genes that can be used to evaluate the expression of target genes in rice plants cultivated under biotic and abiotic stresses.

ACKNOWLEDGMENTS

Research supported by the Brazilian Government (CNPq, CAPES, and Embrapa). The funding agency had no role in the study’s design, data collection and analysis, the decision to publish, or preparation of the manuscript. We would also like to acknowledge Dr. Cláudio Brondani for permission to collect the leaves of his water deficit experiments.

REFERENCES

Selection of reference genes in rice


