

Validation of reference genes in pansy for accurate transcript normalization during its flowering stages

Yuan Zeng, Jing Li, Sheng Gong, Ke Zhang, Haiyan Sun, Guopeng Zhu, Xiqiang Song and Jian Wang*

Key Laboratory of Protection and Development Utilization of Tropical Crop Germplasm Resources (Hainan University), Ministry of Education, Haikou 570228, China; College of Horticulture and Landscape Architecture, Hainan University, Haikou 570228, China

(Received: May 2015; Revised: April 2016; Accepted: April 2016)

Abstract

Quantitative real-time PCR (qRT-PCR) is a perfect method for rapid and accurate quantification of gene expression in different organs or at different development stages in plants. Suitable reference genes are important in this method. In order to obtain more accurate genes expression data in pansy (Viola × wittrockiana Gams.) flower, the expression stability of four housekeeping genes, β**-actin (ACT), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), tubulin (TUB) and 18S ribosomal proteins (18S) at seven different floral development stages in pansy were studied in this experiment. The results showed that TUB and 18S genes were the top two stable genes. Additionally, the expression pattern of VwbHLH1 gene was studied with the two most stable reference genesTUB and 18S, and the worst gene GAPDH respectively. The gene expression data were very different when various reference genes were applied. The present study proved that selection of reference genes was definitely important to obtain precise experimental data in qRT-PCR even for the same organ but at different development stages. This study would provide guidelines to obtain more accurate gene expression results for future molecular mechanism study in pansy.**

Key words: Quantitative real-time PCR, reference gene, pansy, normalization, gene expression

Introduction

Pansy (*Viola* \times *wittrockiana* Gams.) is popular for its colorful petals with large area of black, red or purple patches. However, there was no systemic study on the molecular mechanism of anthocyanin accumulation and blotches occurrence in pansy petal so far. Anthocyanin biosynthesis in plants has been known to be mainly controlled by two groups of genes, structural genes and transcriptional regulators. Previous studies indicated that blotches on petals in pansy were the results of differential expressions of structural genes required in anthocyanin biosynthesis in acyanic and cyanic flower areas (Li et al. 2014). However, the type and function of transcript factors probably associated to the anthocyanin accumulation and blotches occurrence in pansy were not clear. Some researchers have proved the transcriptional regulators families containing R2R3-MYB domains, bHLH (basic helix–loop–helix) domains and conserved WD40 repeats (Petroni and Tonelli 2011) could activate the expression of structural genes to regulate anthocyanidins accumulation. We have separated VwbHLH1 gene that probably regulated the anthocyanin accumulation from pansy petals and attempted to achieve the accurate expression data of this gene in the blotched and non-blotched area of petal to confirm the possible function for the blotch occurrence.

Quantitative real time polymerase chain reaction (qRT-PCR) is presently an excellent technique for determination of gene expression for its rapid, sensitive, and reliable result in biological samples. In order to reduce the random error results from the effect of RNA quality, primer design, PCR amplification efficiency and normalization, housekeeping genes involved in basic metabolism and maintenance of the cells are used as internal references. An ideal reference gene should exhibit constant expression level in all plant tissues, materialsor at different development stages andenvironmental changes (Suzuki et al. 2000).

*Corresponding author's e-mail: wjhnau@163.com

Published by the Indian Society of Genetics & Plant Breeding, F2, First Floor, NASC Complex, PB#11312, IARI, New Delhi 110 012 Online management by indianjournals.com

Researchers have reported that the relative transcript levels of some housekeeping genes varied significantly under different experimental conditions (Czechowski et al. 2005; Schmittgen et al. 2000; Suzuki et al. 2000). Gutierrez et al. (2008) found that genes commonly used as references might be expressed unstably during the different development stages of Arabidopsis. Paolacci et al. (2009) found that up to 100-fold variations occurred in the expression level of a target gene in wheat only for the expression variations of the reference genes, and the deviation could result in huge potential misinterpretation of the results. Therefore, it is necessary to select an appropriate inner reference gene for qRT-PCR to decrease the errors and avoid misinterpretations when measuring RNA expression level with this technique. However, systematic research on the reference genes in pansy is lacking presently. In the present study, we chose four traditionally housekeeping genes, β-actin(ACT), glyceraldehyde-3-phosphate dehydrogenase(GAPDH), tubulin(TUB), and 18S ribosomal proteins (18S) , which were widely used in qRT-PCR, to validate the candidate reference genes for more accurate qRT-PCR data in pansy. The expression pattern of VwbHLH1 gene was also studied by different reference genes to verify the validity of appropriate housekeeping gene and to discuss the possible effect of this gene for the flower blotch occurrence in pansy.

Materials and methods

Plant materials

Viola × wittrockiana Gams. 'Mengdie' (yellow petals with purple blotches, Fig. 1) plants were grown at the horticultural farm of Hainan University (Latitude: 20.03N, longitude: 110.33E), Haikou, Hainan Province, China. Petal samples were collected from pansy flowers at seven different developmental stages according to Li et al. (2014). The petal tissues from stage II to VII were divided into two parts i.e., cyanic blotches and acyanic areas for total RNA extraction, but the whole petals at stage I were sampled for the acyanic buds at this stage (Fig. 1). Thus, a total of 13 petal samples were collected and immediately frozen in liquid nitrogen and stored at 80° until use.

RNA extraction and cDNA synthesis

Total RNA of 13 samples was extracted using the modified Trizol method (Li et al. 2013). An aliquot of RNA was quantified by light absorbance at 260 nm and 280 nm, and then electrophoretically separated

Fig. 1. Pansy flowers at seven development stages (I-VII). A = Full bud, B = Morphological and anatomical structure of lower three cyanic blotched petals corresponding to A. c = cyanic blotch arear; a = acyanic non-blotched area. (Li et al. 2014)

on a 1.2% agarose gel to check integrity. First-strand cDNA was synthesized from 1.0 µg of total RNA using a RevertAid First Strand cDNA Synthesis Kit (Trans-Gene, China) according to the manufacturer's instructions. All cDNA samples were stored at -20° before being used as templates in qRT-PCR.

Primer design

The specific primers (Table 1) of 18S, ACT and TUB for qPCR were designed based on the cDNA sequences of homologous plant species, and the primer of GAPDH was quoted from Gao et al. (2012). The primer sequence of VwbHLH1 was designed based on the transcriptome dataof pansy petals. All of the primers were designed by Primer Premier 5.0 Software. The

Table 1. Primer sequences and length of product for the four candidate reference genes

primers were further confirmed with a dissociation curve analysis after amplification. Standard curves based on a 10-fold dilution series from pooled cDNA were prepared to calculate the gene-specific PCR efficiency and regression coefficient (R^2) for each gene.

Quantitative real-time PCR

qRT-PCR was carried out in a Mx3005P QPCR System (Agilent, Palo Alto, CA, USA) using the SYBR premix Ex TaqTM II (TAKARA, JAPAN). The reaction was performed in triplicate in a total volume of 20 µl containing 10 µl SYBR Green Master, 2 µl template cDNA, 0.8 μ each of the primers (10 μ M), 6.4 μ PCRgrade water. The PCR program was 95° for 30s, followed by 40 cycles at 95° for 30 s, 60 $^{\circ}$ for 34 s. Melting curve analysis of amplification products was performed at the end of each PCR to confirm that a single product was amplified and detected. Templates for standard curve were harvested and purified from PCR products by the primers of all tested genes. And the templates were diluted to 10 concentration gradients series by $5 \times$ variance for every gradient to establish the standard curve.

Reference gene validation and data analyses

The expression stability of the candidate genes was estimated by three statistical approaches: 1) threshold cycle (Ct) variation range and coefficient of variation; 2) geNorm v3.4 (medgen.ugent.be/genorm/) (Vandesompele et al. 2002), and 3) NormFinder v20 (Andersen et al. 2004). Ct values of all the programs raw were transformed to relative quantities according to the standard curve equation of each gene. geNorm defines the reference gene stability measure M as the average pairwise variation of a particular gene with all other reference genes. Genes with the lowest M value has the most stable expression. NormFinder assesses the expression stability value (SV) of a gene by evaluating its expression variation within tissues or treatments compared to variation among tissues/ treatments. The candidate gene with the lowest stability value is the most stable gene within the groups studied.

Gene expression of VwbHLH1

Partial cDNA sequence of VwbHLH1 was obtained from transcriptome sequence data of pansy petals from stage V (Fig. 1) and used as a target gene to demonstrate the usefulness of the validated candidate reference gene in RT-qPCR. Primer pair (forward: 5'- ATCACTTGTCCCGTTCGTG -3', reverse:5'- CCGCTTCCCTCAACTATCC-3') of VwbHLH1 was designed base on the cDNA sequence using the Primer Premier 5.0 Software, and the amplification length was 189bp. The 13 petal cDNA samples from stage I to stage VII (Fig. 1) were analyzed by qRT-PCR for the expression of VwbHLH1. The target gene was normalized by the top2 suitable and 1 worst reference genes based on geNorm pair wise variation and NormFinder analysis results for comparison purpose. The relative gene expression data was generated by triplicate experiments and subjected to variance analysis by using general linear model and the means compared using Duncan's multiple range test (Duncan) method at 0.05 level using SAS (version 9.0).

Results

Standard curve equation

The amplification specificity of each primer was confirmed by melting curve analysis.The amplification efficiency and correlation coefficients (R²) were calculated from the slopes of the standard curves generated by serial template dilution (Table 2).

Table 2. Amplification efficiency of qRT-PCR primers used in qRT-PCR analysis

Primer	Standard curve equation	Regre- ssion coeffi-	Amplifi- cation effici-
		cient (R^2)	ency (%)
ACT	$Y = -3.433 \times LOG(X) + 26.18$	0.983	95.6
TUB	$Y = -3.234*LOG(X) + 27.19$	1.000	103.8
18S	$Y = -3.334*LOG(X) + 25.98$	0.997	99.5
	$GAPDH$ Y = -3.610*LOG(X) + 25.26	0.989	89.2
	$VwblLH1 Y = -3.138*LOG(X) + 20.13$	0.995	108.3

Accordingly, the qPCR efficiencies were 95.6%, 103.8%, 99.5% and 89.2%, while the regression coefficients were 0.983, 1.000, 0.997 and 0.989 for ACT,TUB, 18S and GAPDH respectively. Among the four housekeeping genes, the GAPDH had the most offset of amplification efficiency, indicating that the amplification was the most inefficient with this gene as reference gene. The amplification efficiency of VwbHLH1 gene was 108.3, with a R^2 value of 0.995, which demonstrated the PCR system could be used in the calculation of relative expression quantity.

Expression stability of the selected candidate reference genes

The stability of gene expression was determined by quantifying the mRNA level with qRT-PCR (Ermei Chang et al. 2012). For each gene, the cycle threshold (Ct) value was calculated, which represented the cycle times at which a significant increase of the PCR product occurred. In general, this is marked by the middle of the exponential phase of amplification (Bustin 2002). The scatter plots exhibited the expression levels of candidate reference genes in the tested samples (n=13) (Fig. 2). Values are given as cycle threshold

Fig. 2. Expression levels of candidate reference genes in 13 samples of petals

numbers (Ct values) with a mean of three duplicate samples. The upper and lower points represented the maximum and minimum values of threshold numbers. The geNorm 18S, TUB and ACT exhibited relatively low Ct values that ranged from 17.42 to 22.2, from 28.14 to 28.78, and from 17.02 to 21.41, respectively. However, the Ct values of GAPDH ranged from 33.91 to 39.29, which were too high to be out of the appropriate scope.

Expression profiling of candidate reference genes

qRT-PCR experiment was completed with these four housekeeping genes, then the data were analyzed by geNorm and NormFinder (Fig. 3). The program geNorm computed M values to assess the expression stability of a gene based on the average pairwise variation between all studied genes, and the lowest M values are produced by genes with the most stable expression (Ermei Chang et al. 2012). The M values of 18S, TUB, ACT, and GAPDH were 0.175, 0.179, 0.211 and 0.423 respectively (Table 3), indicating that the 18S and TUB would be the most stable reference genes and the

Fig. 3. Determination of the optimal number of reference genes required for effective normalization by pairwise variation (V) using geNorm

Table3. Ranking of the 4 candidate reference genes according to their expression stability value given by geNorm and Normfinder

 $M =$ The expression stability value by the program geNorm; $SV =$ The stability value according to NormFinder program

GAPDH was the worst reference gene. The NormFinder program is a Visual Basic application tool for Microsoft Excel used to determine expression stabilities of reference genes with stability value (SV) (Xu et al. 2011). The SV of 18S, TUB, ACT, and GAPDH were 0.007, 0.007, 0.078 and 0.291 respectively in our experiments. The similar results were discovered that the 18S and TUB were the most stable reference genes with the lowest SV and the GAPDH was the worst reference gene with the highest SV.

The pair wise variation (V_n/V_{n+1}) between the normalization factors \textsf{NF}_{n} and $\textsf{NF}_{\mathsf{n+1}}$ was analyzed by geNorm program to determine the minimum number of reference genes required for RT-qPCR data normalization. According to the pairwise variation analysis from geNorm (Fig. 3), two reference genes were enough to apply for this qRT-PCR experiment, as V2/3 was 0.0325, which was under the cut-off level of 0.15.

Evaluation of selected reference genes based on the expression level of VwbHLH1

Relative expressions of VwbHLH1 were analyzed using the top 2 reference genes, 18S and TUB, and the worst gene GAPDH to confirm accurate normalization. The expression levels were calculated using the standard curve equation (Table 2). The significant differences calculated by SAS 9.0 are demonstrated in Fig. 4. According to the experiment with the 2 best reference genes, the expression level of VwbHLH1 peaked at stage II and reached its minimum value at stage VII (Fig. 4i). The expression levels in cyanic areas of stage IV and VI were higher than that in the acyanic areas, but opposite expression patterns were observed in stage III and V (Fig.4i). However, the variation tendency of VwbHLH1 was quite different when the worst gene GAPDH was used as reference gene. The relative expression level normalized by GAPDH was higher than 18S and TUB (Fig.4i and Fig.4ii), and the

Fig. 4. Relative expression quantity of VwbHLH1 of different development stages in pansy A and C, respectively represent the acyanic nonblotched and cyanitic blotch areas. Mean and SD deviation values of two reference genes are presented. Columns with different letters at each time point indicate significant differences at P < 0.05

expression quantity by GAPDH in the cyanic areas at stage II, IV and VI were significantly higher than that in the acyanic areas, which were very different with the results by 18S and TUB (Fig. 4i and Fig. 4ii). These data confirmed that the obvious errors would be produced when the wrong reference gene, such as GAPDH, was used in the real-time PCR experiment.

Discussion

It is currently accepted that the stability of potential reference genes must be systematically determined prior to their use (Guenin et al. 2009). A lot of experiments have proved that the stabilities of reference genes relied on the species of plant and the organs and even on the development stages. In the present study, GAPDH expressed most unsteadily, however, in eggplant, GAPDHand 18S had a good performance under most experimental conditions (Zhou et al. 2014), whereas TUB exhibited poor stability. On the contrary, GAPDH and 18S ranked bottom in stability value of strawberry (Galli et al. 2015). In other studies, GAPDH showed an excellent stability during carrot tap root secondary growth, yet TUB was unsteadily compared to other gene included in the study (Campos et al. 2015). Moreover, TUB also ranked in the lower positions of stability of many other plant materials, such as citrumelo, kumquat, orange, and several virus-infected monocots such as brachypodium, barley, sorghum (Hu et al. 2014; Paolacci et al. 2009; Wu et al. 2014; Zhang et al. 2013). ACT performed quite moderate in our study, however, it was the optimal reference gene in cherry, kumquat orchardgrass, etc (Hu et al. 2014; Huang et al. 2014; Ye et al. 2015). Consequently, validating a set of candidate reference genes for specific experimental system should be very necessary, because the expression level of those genes cannot be assumed to remain constant under all possible plants and development stages. The present study showed, the huge gene expression variations occurred when the worst stable reference gene GAPDH was used for the VwbHLH1 expression, indicating that the optimal reference gene was also necessary and important even to the same organs at the different development stages.

There was no report available on the validation of reference genes in pansy before, although we have used the ACT gene as the reference gene for the realtime PCR in the previous study (Li et al. 2014). In this study the ACT gene was proved to be a stable gene at different floral development stages in pansy (M

values = 0.211 and SV = 0.078), which was lower than the cut-off level of 0.5 and 0.15 (Table 3). Moreover, 18S and TUB genes exhibited better application potentiality for the quantitative gene expression study for their higher stability. These results provided guidelines to obtain more accurate qRT-PCR results for future molecular mechanism study in pansy.

VwbHLH1 cDNA sequence had an identity of 80% with FabHLH3 from strawberry (Fragariax ananassa), which could be able to regulate the anthocyanin biosynthesis in strawberry (data not shown). However, FabHLH3 did not directly impact on anthocyanin biosynthesis. It interacted with FaMYB transcript factors family, which can promote or inhibit the anthocyanin biosynthesis. In FaMYB family of strawberry, FaMYB1 was confirmed as a repressor in the regulation of the anthocyanin biosynthetic pathway (Aharoni et al. 2001); FaMYB5 was proved to possibly play an inhibitory role in flavonoid biosynthesis; FaMYB9 and FaMYB11 had a positive effect on proanthocyanidin biosynthesis; FabHLH3 displayed an interaction with the Arabidopsis and strawberry TTG1 proteins (Schaart et al. 2013). Reasonably, we speculated that VwbHLH1 could cooperate with other transcript factors, such as MYB or WD40 protein to promote or repress the expression of anthocyanin biosynthesis structural genes in pansy flower. However, in the present experiment, the expression levels of VwbHLH1 in cyanic part were not always significantly higher or lower than that in the acyanic part in petals of pansy, indicating that VwbHLH1 would be not directly impact on anthocyanin biosynthesis as the FabHLH3 from strawberry. The specific mechanism should be further investigated using more genetic and molecular approaches.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 31060265, 31260488) and was also supported in part by the Academic Discipline Construction Project Plan in the Central and Western Regions of Hainan University (ZXBJH-XK008) and the State Key Subject of Botany at Hainan University (071001).

References

Andersen C. L., Jensen J. L. and Ørntoft T. F. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res., **64**: 5245-5250.

- Aharoni A., Vos C. H. R. D., Martina Wein, et al. 2001. The strawberry FaMYB1, transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. Plant J., **28**: 319-332.
- Bustin S. A. 2002. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J. Mol. Endocrinol., **25**: 169-193.
- Campos M. D., Frederico A. M., Nothnagel T. et al. 2015. Selection of suitable reference genes for reverse transcription quantitative real-time PCR studies on different experimental systems from carrot (Daucuscarota L.). Sci. Hortic., **186**: 115-123.
- Czechowski T., Stitt M., Altmann T. et al. 2005. Genomewide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol., **139**: 5-17.
- Chang E., Shi S., Liu J. et al. 2012. Selection of reference genes for quantitative gene expression studies in Platycladusorientalis (Cupressaceae) using realtime PCR. Plos One, **7**: 65-65.
- Galli V., Borowski J. M., Perin E. C., et al. 2015. Validation of reference genes for accurate normalization of gene expression for real time-quantitative PCR in strawberry fruits using different cultivars and osmotic stresses. Gene, **554**: 205-214.
- Gao Z. H., Wei J. H., Yang Y. et al. 2012. Selection and validation of reference genes for studying stressrelated agarwood formation of Aquilaria sinensis. Plant Cell Rep., **31**: 1759-1768.
- Gutierrez L., Mauriat M., Guenin S. et al. 2008. The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcriptionpolymerase chain reaction (RT-PCR) analysis in plants. Plant Biotechnol. J., **6**: 609-61.
- Guenin S., Mauriat M., Pelloux J. et al. 2009. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. J. Exp. Bot., **60**: 487-493.
- Hu Y., Chen H., Luo C. et al. 2014. Selection of reference genes for real-time quantitative PCR studies of kumquat in various tissues and under abiotic stress. Sci. Hortic., **174**: 207-216.
- Huang L., Yan H., Jiang X. et al. 2014. Reference gene selection for quantitative real-time reversetranscriptase PCR in orchardgrass subjected to various abiotic stresses. Gene, **553**: 158-165.
- Li Q., Wang J., Sun H. Y. and Shang X. 2014. Flower color patterning in pansy (Viola x wittrockiana Gams.) is caused by the differential expression of three genes from the anthocyanin pathway in acyanic and cyanic flower areas. Plant Physiol. Biochem., **84**: 134-141.
- Li Q., Wang J., Zhao Z. X. and Shang X. 2013. Comparison analysis of three methods of RNA isolation methods

using pansy petals as materials. Jiangsu Agric. Sci., **41**: 3.

- Paolacci A. R., Tanzarella O. A., Porceddu E. and Ciaffi M. 2009. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. BMC Mol. Boil., **10**: 11.
- Petroni K. and Tonelli C. 2011. Recent advances on the regulation of anthocyanin synthesis in reproductive organs. Plant Sci., **181**: 219-229.
- Schaart J. G., Dubos C., Romero De La Fuente I. et al. 2013. Identification and characterization of MYBbHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (Fragaria xananassa) fruits. New Phytol., **197**: 454- 467.
- Schmittgen T. D. and Zakrajsek B. A. 2000. Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. J. Biochem. Biophys. Methods, **46**: 69-81.
- Suzuki T., Higgins P. J. and Crawford D. R. 2000. Control selection for RNA quantitation. Biotechniques, **29**: 332-7.
- Vandesompele J., Preter K.D., Pattyn F. et al. 2002. Accurate normalization of real-time quantitative RT-

PCR data by geometric averaging of multiple internal control genes. Genome Biol., **3**: research0034.1 research0034.11.

- Wu J., Su S., Fu L. et al. 2014. Selection of reliable reference genes for gene expression studies using quantitative real-time PCR in navel orange fruit development and pummelo floral organs. Sci. Hortic., **176**: 180-188.
- Xu M., Zhang B., Su X. et al. 2011. Reference gene selection for quantitative real-time polymerase chain reaction in Populus. Anal. Biochem., **408**: 337-339.
- Ye X., Zhang F., Tao Y. et al. 2015. Reference gene selection for quantitative real-time PCR normalization in different cherry genotypes, development stages and organs. Sci. Hortic., **181**: 182-188.
- Zhang K., Niu S., Di D. et al. 2013. Selection of reference genes for gene expression studies in virus-infected monocots using quantitative real-time PCR. J. Biotech., **168**: 7-14.
- Zhou X., Liu J. and Zhuang Y. 2014. Selection of appropriate reference genes in eggplant for quantitative gene expression studies under different experimental conditions. Sci. Hortic., **176**: 200-207.