

IDENTIFICATION AND VALIDATION OF NEW REFERENCE GENES FOR NORMALIZATION OF GENE EXPRESSION IN FLOWER AND BERRY DEVELOPMENTAL STAGES OF INTERSPECIFIC GRAPE HYBRID *V. VINIFERA* (L.) × *V. LABRUSCA* (L.)

Siddhi CHAVAN^{1,2}, Anil KHATAL³, Satish PHALAKE^{1,2}, Sujata TETALI^{1,2}, Ravindra PATIL^{1,2*} 

¹Genetics and Plant Breeding Group, Agharkar Research Institute, G. G. Agarkar Road, Pune 411004, India

²Savitribai Phule Pune University, Pune, 411007, India

³Department of Biotechnology, Ahmednagar College, Ahmednagar, 414001, India

Received: February 2024; Accepted: April 2024

ABSTRACT

In quantitative reverse transcription polymerase chain reaction (qRT-PCR), normalizing target gene expression using a reference gene is an indispensable step to control the variability of RNA extraction yield, RNA integrity, reverse transcription efficiency, and PCR amplification. In the present study, we identified candidate reference genes with stable expression during grapes' flowering and berry development stages. Ten genes, including *ACT*, *CYP5*, *RLI*, *TUB*, *UBC*, *UBC17*, *UBC60*, *UFD1*, *VAG*, and *ZNF* with relatively stable expression, were selected based on RNAseq data generated earlier in grape hybrid 'ARI 516'. The expression of these candidate genes was tested at different stages of flowering and grape berry development. Five different algorithms such as RefFinder, geNorm, NormFinder, BestKeeper, and the comparative ΔC_q method were used to test the expression stability of candidate genes. A comprehensive ranking obtained by RefFinder showed that *UBC17*, *RLI*, and *ZNF* were the most stable reference genes during flower and berry development stages. *UBC17*, *RLI*, and *ZNF* were calibrators to normalize the expression of *VvAGLII* as a target gene to validate the worthiness of identified reference genes. The result demonstrated that newly identified reference genes could be successfully used to normalize the expression of the target gene accurately. These reference genes will provide more choices for selecting appropriate reference genes to normalize gene expression in grapes.

Key words: algorithms, expression stability, relative expression, quantitative RT-PCR, *UBC17*, *RLI*, *ZNF*

INTRODUCTION

Grapevine (*Vitis vinifera* L.) is considered one of the major fruit crops in the world based on the area under cultivation and its economic value. Grape is a globally cultivated fruit with 80.1 million tons of produce harvested from 7.3 million hectares of area in 2022 (<https://wineindustryadvisor.com/2023/06/06/oiv-annual-grape-production-assessment-2022>). The grape berries are eaten fresh, processed to make wine, vinegar, juice, or dried to produce raisins. The grape berry is an independent biochemical factory. Berry development involves a cascade of biochemical and physiological processes determining the harvest's taste, aroma, and other sensory characteristics (Kuhn et al. 2014). Beyond

the primary metabolites such as sugars, organic acids, anthocyanins, and amino acids, the berry can synthesize precursors of aroma and flavor that define a particular wine. After the fruit has set, the berries undergo a double sigmoidal growth pattern, divided into three distinct stages. The first stage involves rapid growth and accumulation of organic acids and phenolic compounds. The veraison stage is characterized by slower growth and accumulation of sugars and pigments. In the ripening stage, the berries reach their mature size and color, the sugar concentration increases, organic acid production decreases, and volatile aroma precursors are synthesized that contribute flavor and aroma (Fasoli et al. 2012; Sweetman et al. 2012; González-Agüero et al. 2013; Kuhn et al. 2014). The study of genes

*Corresponding author

e-mail: rmpatil@aripune.org

involved in these biochemical and physiological processes has gained importance in grape improvement programs due to the economic value of grape berries. Therefore, recent studies have reported examining the expression of these genes at various key stages of berry development (Deluc et al. 2007; Zenoni et al. 2010; Sweetman et al. 2012; Muñoz-Espinoza et al. 2016; Massonnet et al. 2017; Ma et al. 2020). Similarly, flowering or anthesis in grapes is one of the critical stages determining the quantity and quality of the harvest in grape cultivation. Therefore, these stages have been extensively studied by researchers to understand the molecular mechanism underlying parthenocarpy, ovule development, stenospermocarpy, and development of berries (Zenoni et al. 2010; Mejía et al. 2011; Sweetman et al. 2012; González-Agüero et al. 2013; Kuhn et al. 2014; Nwafor et al. 2014; Alva et al. 2015; Wang et al. 2015, 2016; Royo et al. 2016).

The deployment of high-throughput techniques such as RNAseq and microarray to measure gene expression profiles has become popular in the research community. However, the occurrence of artifacts is possible during an array or RNA sequencing experiments. Therefore, the gene expression results obtained from these techniques have to be validated using independent laboratory approaches such as quantitative reverse transcription polymerase chain reaction (qRT-PCR), northern blot, *in situ* hybridization, and ribonuclease protection assay (Chuaqui et al. 2002). Among these, qRT-PCR is widely preferred to detect a change in gene expression and validate the RNAseq data due to its accuracy, sensitivity, and rapid assay time. In a qRT-PCR experiment, the normalization of expression of a target gene by using a reference gene is an indispensable step that controls variation in RNA extraction yield, RNA integrity, reverse transcription efficiency, and PCR amplification (Bustin et al. 2009). Normalization of expression using a stably expressed reference gene allows comparison of RNA concentrations in different samples. Various statistical algorithms, including BestKeeper, geNorm, NormFinder, ΔC_q , and RefFinder, have been developed to assess the expression stability of reference genes. The quantification cycle (C_q) value indicates the position of the amplification curve concerning the cycle axis and is inversely correlated with the quantity of the initial RNA template. The ΔC_q method calculates the relative expression levels between a candidate reference gene and other candidate reference genes within each sample. It compares

the standard deviation (SD) of ΔC_q to estimate the most stable reference genes (Silver et al. 2006). Consistent ΔC_q values between different RNA samples mean that gene expression remains stable in these samples. Therefore, it can be concluded that both genes are stably expressed among those samples. Evaluating the most stable genes in the geNorm algorithm is based on sequentially eliminating genes with the highest M -value (gene-stability measure). It determines the optimal number of reference genes by calculating pairwise variation between them and the other reference genes (Vandesompele et al. 2002). NormFinder ranks candidate reference genes by calculating their stability value (SV) among samples in the given groups and selecting the ones with a lower stability value, considered to have higher expression stability (Andersen et al. 2004). The SV adds intra- and intergroup variation in expression and, therefore, represents a practical measure that determines the precise amount of systematic error that will be introduced when using the investigated gene. In the BestKeeper algorithm, the stability of the candidate genes is calculated based on the SD of their C_q values, plus the coefficient of variance (CV), Pearson correlation coefficient (r), and the probability p -value, which are also important parameters for stability analysis (Pfaffl et al. 2004). These four algorithms were integrated into the online program RefFinder, which provides a comprehensive ranking of candidate reference genes based on the geometric mean (GM) values of the results obtained from the four algorithms described above (Xie et al. 2012).

Reference genes have been identified in grapes under abiotic stress conditions (Coito et al. 2012; Borges et al. 2014; Luo et al. 2018) as well as under biotic stresses such as downy mildew (Selim et al. 2012; Monteiro et al. 2013). Reference genes were also identified using heterologous species such as *Arabidopsis* (Tashiro et al. 2016). Genes such as β -actin, tubulin, elongation factor-1 α , polyubiquitin, and glyceraldehyde-3-phosphate dehydrogenase are commonly used as reference genes for normalization of qRT-PCR data in grapes (Zenoni et al. 2010; Sweetman et al. 2012; González-Agüero et al. 2013; Borges et al. 2014; Upadhyay et al. 2015). These housekeeping genes are assumed to have constant expression levels among control and treated samples regardless of experimental conditions, developmental stages, tissues, organs, or stress treatments. However, it is observed that the expression of housekeeping genes may vary across genotypes, tissues, growth stages, physiological conditions, and also in response to the

application of phytohormones (González-Agüero et al. 2013; Monteiro et al. 2013; Borges et al. 2014; Upadhyay et al. 2015; Katayama-Ikegami et al. 2016). Although cultivated grapes consist of various *Vitis* species and their interspecific hybrids, it is essential to note that it cannot be assumed that the reference genes will perform equally well in different species (Gutierrez et al. 2008; Graeber et al. 2011). Together, these findings suggest that an experimentally suitable reference gene may not be helpful in normalizing gene expression in a different set of experiments with variation in species, interspecific hybrids, growth stage, or physiological parameters. In the present study, a panel of 10 potential reference genes was selected from the RNAseq data generated for critical flowering and berry development stages of 'ARI 516'. These genes were tested for expression stability at different growth stages using five algorithms to identify the most stable combination of reference genes. It would help improve the accuracy of gene expression analysis for flowering and berry development stages in grapes.

MATERIALS AND METHODS

Plant material and sample collection

Eight-year-old 'ARI 516' grapevines were selected for the study. The grapevines were planted at the experimental farm of Agharkar Research Institute (ARI), Pune (18°31' N, 73°55' E) at a spacing of 2.4 m between rows and 1.5 m between vines. 'ARI 516' is a high-yielding, disease-resistant, multipurpose cultivar with purple-black berries and a unique musky flavor developed at ARI (Tetali et al. 2020). It is an interspecific hybrid developed from a cross of 'Beauty seedless' (*V. vinifera* L.) × 'Catawba' (*V. labrusca* L.). Tissue samples were collected at different development stages such as E-L 12 (inflorescence visible), E-L 15 (flower development stage), E-L 33 (hard and green berries), E-L 36 (berries with intermediate sugar content), and E-L 37 (not quite ripe berries) (Coombe 1995). E-L 12 and E-L 15 are critical inflorescence development stages during which reproductive success and harvest potential of grapevine are determined. On the other hand, E-L 33, E-L 36, and E-L 37 are the critical berry developmental stages during which the sugar concentration in berry increases, organic acid production decreases, and volatile aroma precursors are synthesized, contributing to the quality of the grape harvest. Samples were immediately frozen

in liquid nitrogen to avoid RNA degradation. Three independent samples were collected as biological replicates for each developmental stage.

RNA extraction

Total RNA extraction was performed for each sample from developing flowers/berries in triplicate (technical replicates) using the Spectrum Plant Total RNA kit (Sigma-Aldrich) following the manufacturer's protocol. RNA purity and concentration were measured at 260/280 nm using a spectrophotometer Nanodrop 8000, and for RNA integrity, the samples were checked on 1.2% denaturing formaldehyde gel and visualized using UV transilluminator. First-strand cDNA was synthesized from 1 µg of DNase I treated RNA using PrimeScript 1st strand synthesis Kit (TAKARA). After the first strand synthesis, all cDNA samples were diluted 5-fold with RNase-free water and used as templates in qPCR analysis.

Selection of candidate reference genes

We used earlier generated transcriptome data of flower and berry development stages of grape 'ARI 516' (unpublished data) to select the potential reference genes for this study. Ten genes showing relatively stable expression in transcriptome analysis (BioProject PRJNA1085542 at NCBI database) were selected, including *actin* (*ACT*), *peptidyl-prolyl cis-trans isomerase* (*CYP5*), *RNase L inhibitor protein 1/ABC transporter E1* (*RLI*), *tubulin* (*TUB*), *ubiquitin-conjugating enzyme* (*UBC*), *ubiquitin-conjugating enzyme E2-17* (*UBC17*), *ubiquitin-60S ribosomal protein* (*UBC60*), *ubiquitin fusion degradation protein-1* (*UFD1*), *vacuolar ATPase subunit G* (*VAG*), and *zinc finger-(C2H2 type) family protein* (*ZNF*). Expression stability was also studied using the Grape eFP Browser, the Grapevine Expression Atlas, which reveals a profound transcriptome shift that puts the entire plant into a maturation program (Fasoli et al. 2012).

Primer design and amplification efficiency analysis for qRT-PCR

Gene-specific primers were designed using Primer3Web (version 4.1.0) with primer-pick criteria – T_m values ranging from 55°C to 61°C, GC percentage of 45–55%, and product length of 100–180 bp. To detect the specificity of each primer, PCR with 10 µl of reaction volume was conducted using 2 µl of 5-fold diluted cDNA, 1× Taq Buffer (Kapa Biosystems), 10U Taq polymerase (KAPA BIOSYSTEM), 0.5 mM MgCl₂, 2.5 nM primers each and 2.5 mM dNTP using Eppendorf Master Cycler PCR System. The PCR thermal cycling protocol included

initial denaturation for 5 minutes at 95°C, followed by 35 cycles of 30 seconds at 95°C, 30 seconds at 55–60°C and 30 seconds at 72°C, and a final extension for 5 minutes at 72°C. Amplification products were assessed using a 2.5% metaphor agarose gel and visualized using a UV transilluminator. The amplification efficiency of each primer pair was determined from the slope of the standard curve. Standard curves were generated by performing qRT-PCR on each gene using a series of 10-fold diluted cDNA samples in duplicates (10–1000-fold dilutions). The PCR amplification efficiency of each primer pair was calculated using the formula $E (\%) = (10^{-1/\text{slope}} - 1) \times 100\%$.

Quantitative reverse transcription PCR

Reactions were carried out with KAPA SYBR FAST qPCR Master Mix (2×) Universal on PCR Max Eco 48 Real-Time PCR System. The PCR conditions were: 95°C for 10 minutes as an initial step, followed by 45 cycles of 95°C for 30 seconds, 52–64°C for 30 seconds and 72°C for 30 seconds. Finally, a post-PCR melting curve analysis at 95°C for 15 seconds, 55°C for 15 seconds, and 95°C for 15 seconds was performed to verify the specificity of amplification to rule out the primer-dimers, nonspecific product formation, and confirmation of a single peak. Each sample was examined in three technical replicates and analyzed using the EcoStudy Software Version 5.2.12, Cole-Parmer.

Analysis of the expression stability of reference genes

RefFinder (Xie et al. 2012), geNorm (Vandesompele et al. 2002), NormFinder (Andersen et al. 2004), BestKeeper (Pfaffl et al. 2004), and the comparative ΔC_q method (Silver et al. 2006) were used to assess the stability of reference gene expression. The mean C_q value of each sample and primer was used as input to estimate gene expression stability.

Validation of selected reference candidate genes

In *Vitis vinifera*, the *VvAGL11* gene, which encodes Agamous-like MADS-box protein AGL11, is well documented for its expression in developing berries as it plays a significant and direct role in seed development (Mejía et al. 2011). Therefore, to confirm the suitability of the identified reference genes, they were used as a calibrator to measure the relative expression of the *VvAGL11* gene at different stages of berry development. Relative expression of the *VvAGL11* gene was calculated using the $2^{-\Delta\Delta C_q}$ method. As detailed above, the qRT-PCR analysis was conducted in the three biological replicates.

RESULTS

Quantification cycle (C_q) values of candidate reference genes

Primers targeting ten genes, including *ACT*, *CYP5*, *RLI*, *TUB*, *UBC*, *UBC17*, *UBC60*, *UFD1*, *VAG* and *ZNF* were used for qRT-PCR analysis. All the primer pairs' amplification efficiency values ranged from 93.2% to 99.09%, with a correlation coefficient (R^2) of 0.99 and higher (Table 1). The expression levels of 10 selected potential reference genes were measured by the C_q value. The expression stability of ten selected candidate reference genes was assessed based on C_q values at different developmental stages of 'ARI 516' flowers and berries. Other genes showing C_q values above the cutoff ($C_q > 29$ or between 29 and 34) were excluded from further stability analysis because genes showing higher C_q values cannot be considered as reference genes. The expression profiles of all genes at five different developmental stages are shown in Figure 1. Based on C_q values, *ACT* showed high transcript abundance. In contrast, *RLI* and *UFD* showed the lowest transcript abundance. Among the candidate reference genes tested, *UBC17*, *ZNF*, *UBC60*, and *TUB* showed intermediate transcript abundance. Among the different developmental stages analyzed, ten candidate reference genes showed a C_q range of 20–26 cycles (Fig. 1). *ACT* showed average C_q values below 21 cycles, and *UBC60*, *CYP5*, and *UBC* were below 23. *ZNF* and *VAG* showed C_q below 24 cycles. *TUB*, *UBC17*, *UFD1*, and *RLI* were the least expressed genes with a C_q value between 24 and 26 cycles.

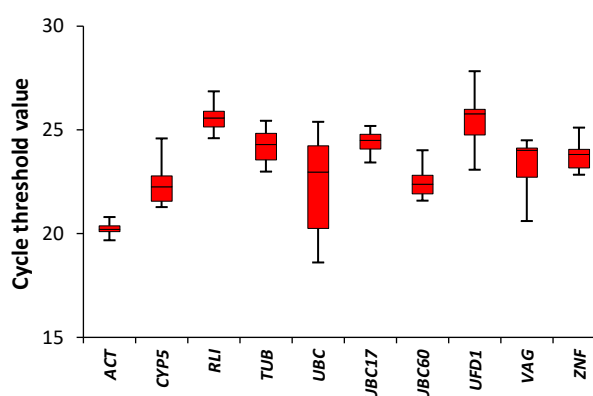


Figure 1. Distribution of C_q values for ten candidate reference genes observed in grape tissue samples of flowering and berry development stages

Red color – boxplot for C_q values; Dotted line in each box – median C_q values; The top and the bottom lines of the box – the first and third quartiles, respectively; The top and bottom whiskers – 5th and 95th percentiles (minimum and maximum range of C_q values, respectively)

Analysis of expression stability of candidate reference genes at various stages of development

Four different statistical algorithms (geNorm, NormFinder, BestKeeper, and ΔC_q method) were applied to evaluate the expression stability of ten reference candidate genes in different developmental stages (E-L 12, E-L 15, E-L 33, E-L 36, and E-L 37). The final ranking was based on the geometric mean of the four algorithms, the smaller the geometric mean, the greater the stability of the reference gene expression (Rodrigues et al. 2014). The geNorm algorithm using mean pairwise variation showed that all ten candidate genes were suitable as reference genes ($MV < 1.5$), with *ZNF*, *UBC17* ($MV = 0.37$), and *UBC60* ($MV = 0.42$) being the most stable genes, whereas *UFD1* and *UBC* were the least stable genes. All genes showed an *M* score below the cutoff value of 1.5. The most stable reference gene was *UBC17*, consistent with NormFinder, BestKeeper, and ΔC_q . The results of the BestKeeper analysis showed that *ACT*, *RLI*, *ZNF*, *UBC17*, *UBC60*, *TUB*, and *CYP5* were the most stable genes, all with $SD < 1$ (Table 2) and thus qualifying as reference genes. In addition, according to NormFinder analysis, *RLI* and *UBC17* were the most stable genes based on their SV value of 0.483 and 0.489, whereas *CYP5*, *UFD1*, and *UBC* were the most unstable genes ($SV > 1$). According to NormFinder analysis, the lower the SV, the higher the stability, and the higher the SV, the lower the stability of that gene (Table 2). We also used the ΔC_q algorithm to rank the expression stability of the ten

candidate reference genes. *RLI* was the most stable gene based on its SD value, similar to NormFinder results. According to ΔC_q analysis, the second most stable gene followed by *RLI* was *UBC17*. Finally, mean C_q values were analyzed using the online RefFinder tool and a comprehensive ranking was obtained based on the rankings obtained for each gene using the previous four algorithms. *UBC17*, *RLI*, and *ZNF* proved to be the most stable genes, all with $GM < 3$ (Table 2). The geometric mean ranking for all ten candidate reference genes from the most to the least stable was *UBC17*, *RLI*, *ZNF*, *ACT*, *UBC60*, *TUB*, *CYP5*, *VAG*, *UFD1*, and *UBC* (Table 2).

Validation of selected candidate reference genes

In *Vitis vinifera*, the VIT_18s0041g01880 gene encoding Agamous-like MADS-box protein *AGL11* plays a significant and direct role in seed development. Mutation in this gene leads to seedlessness in fleshy fruits (Mejía et al. 2011). To validate the suitability of *UBC17*, *ZNF*, and *RLI* as reference genes, they were used as calibrators to normalize the expression of *VvAGL11*. As illustrated in Figure 2, the expression levels of *VvAGL11* were higher in E-L 33 and E-L 36 stages. In contrast, the expression level was the lowest or the gene was downregulated at stage E-L 37. The expression levels of *VvAGL11* normalized with new reference genes by qRT-PCR analysis agreed with the earlier report (Mejía et al. 2011) and our earlier transcriptome data (SRA: SUB14289652 at NCBI). The result demonstrated that the identified genes could be successfully used to accurately normalize the expression of the target gene.

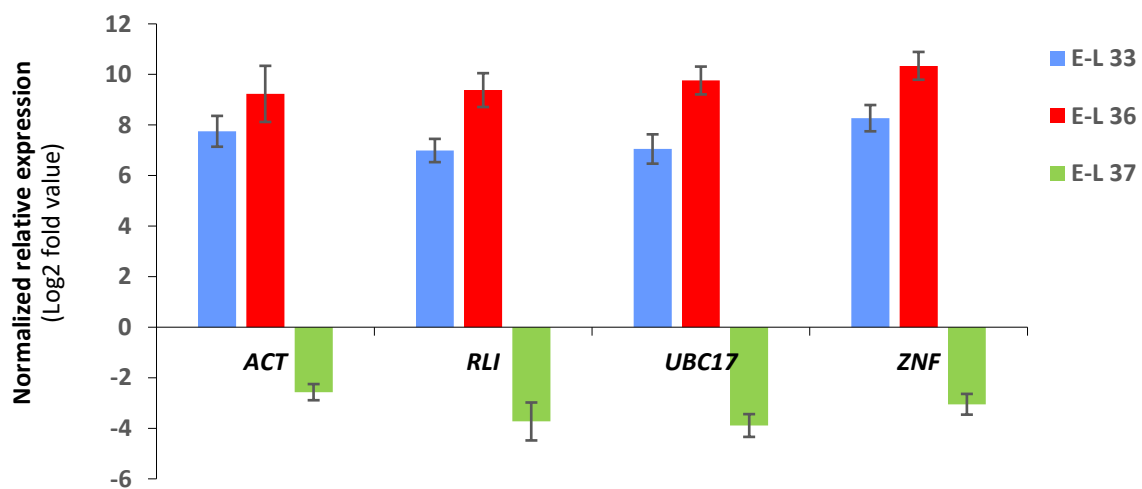


Figure 2. Validation of the *UBC17*, *RLI*, and *ZNF* reference genes

The newly identified reference genes were used to normalize the expression of the *VvAGL11* target gene at three different stages of berry formation (E-L 33, E-L 36, and E-L 37). The expression values of *VvAGL11* at the growth stages E-L 33, E-L 36, and E-L 37 were normalized against the expression values at growth stage E-L 15. Error bars – standard deviation of the log₂ fold expression of *VvAGL11*

Table 1. Details of candidate reference genes, their primer sequences and amplification characteristics for qRT-PCR

Gene symbol	Gene ID	Gene name	Primer Sequence (5' to 3')	Amplicon length (bp)	T (°C)	E (%)	R ²	Reference
<i>ACT</i>	VIT_13s0019g01240	<i>actin</i>	F: GACTACCTACAACTCCATCAT R: TCATTCTGTGAGCAATACCA	113	58	99.09	0.996	Borges et al. (2014)
<i>CYP5</i>	VIT_03s0038g01930	<i>peptidyl-prolyl cis-trans isomerase</i>	F: ACAAACGGATCGAGTTCCT R: TCCTGATCTTCACGGTCTT	161	60	93.2	0.997	this study
<i>RLI</i>	VIT_02s0087g00770	<i>RNase L inhibitor protein 1/ABC transporter EI</i>	F: GGTTTATCTCCACGCAAG R: GATTATCCAGTCAACAATGA	133	61	95.83	0.999	this study
<i>TUB</i>	VIT_06s0004g00480	<i>tubulin</i>	F: TGATAACTGCACTGGCTTGC R: TGGGGAGAGGATAGATAGTGA	150	61	98.64	0.998	this study
<i>UBC</i>	VIT_08s0040g00040	<i>ubiquitin-conjugating enzyme</i>	F: CATAAGGGCTATCAGGAGGAC R: TGGCGGTCGGAGTTAGG	174	61	95.51	0.999	Borges et al. (2014)
<i>UBC17</i>	VIT_19s0015g01190	<i>ubiquitin-conjugating enzyme E2-17</i>	F: GTTCCACTGGCAAGCAACTA R: GTGGCTTGAATGGGTAATCC	104	60	98.19	0.99	this study
<i>UBC60</i>	VIT_16s0098g01190	<i>ubiquitin-conjugating enzyme E2-60</i>	F: CATCTGGTCTGAGGCTTCG R: GTGGCCACACTCTCTTTTCC	150	61	98.82	0.999	this study
<i>UFD1</i>	VIT_03s0038g00750	<i>ubiquitin fusion degradation protein-1</i>	F: AGTTCAAACCCCTCTCTGTGA R: CTTCCCTGAGACTGACGAG	173	62	95.97	0.998	this study
<i>VAG</i>	VIT_08s0040g01910	<i>vacuolar ATPase subunit G</i>	F: TTGCTGTGTCTCTTGTTC R: TCAATGCTGCCAGAAGTG	161	59	96.94	0.997	Borges et al. (2014)
<i>ZNF</i>	VIT_04s0023g00440	<i>zinc finger (C2H2 type) family protein</i>	F: AAAGCGAAACCAAGAAAGCA R: GGATTTCATCGATTGAGG	184	61	95.18	0.994	this study
<i>PvAGL11</i>	VIT_18s0041g01880*	<i>Agamous-like MADS-box protein AGL11</i>	F: GCAGAGTTGCCCTCATCGT R: AAGCCAAAGGAATCACCATT	232	58	-	-	Mejia et al. (2011)

* used to validate selected candidate reference genes

T (°C) – annealing temperature; E (%) – amplification efficiency; R² – regression coefficient

Table 2. Expression stability ranking of candidate reference genes at the stages of flowering and grape berry development

Gene	RefFinder (comprehensive ranking)		geNorm		NormFinder		BestKeeper		ΔC_q	
	rank	GM	rank	MV	rank	SV	rank	SD	rank	SD
<i>UBC17</i>	1	1.68	1	0.375	2	0.489	2	0.437	2	1.04
<i>RLI</i>	2	1.86	3	0.468	1	0.483	3	0.54	1	1.03
<i>ZNF</i>	3	2.78	1	0.375	5	0.735	4	0.542	3	1.11
<i>ACT</i>	4	3.22	5	0.63	3	0.601	1	0.294	6	1.19
<i>UBC60</i>	5	4.61	2	0.422	6	0.776	5	0.624	5	1.15
<i>TUB</i>	6	4.68	4	0.525	4	0.65	6	0.683	4	1.13
<i>CYP5</i>	7	7.24	6	0.71	8	1.195	7	0.936	7	1.41
<i>VAG</i>	8	7.74	7	0.911	7	0.993	8	1.054	8	1.47
<i>UFD1</i>	9	9	8	1.092	9	1.329	9	1.095	9	1.67
<i>UBC</i>	10	10	9	1.365	10	2.372	10	2.029	10	2.46

The ranks were determined using the following parameters: GM – geometric mean, SV – stability value, MV – *M*-value, and SD – standard deviation

DISCUSSION

The qRT-PCR has been extensively used to quantify the expression of genes as well as to validate the data obtained in RNAseq- and microarray-based transcriptome studies. According to MIQE (minimum information for publication of quantitative real-time PCR experiments) guidelines, a reliable qRT-PCR quantification assay needs suitable reference genes as internal controls (Zhou et al. 2018). Normalization with multiple reference genes has become a gold standard in qPCR expression analysis and a requisite according to MIQE guidelines. Earlier known reference genes such as *GAPDH*, *EF1-alpha*, and *ubiquitin* (involved in fundamental cellular processes) or those encoding *actin* and *tubulin* (involved in maintaining cellular structure) were often used in qRT-PCR experiments. However, some reports have shown that the thoughtless use of these genes may be inappropriate for normalization in qRT-PCR experiments due the variability of their expression under different experimental conditions and at different developmental stages (González-Agüero et al. 2013; Monteiro et al. 2013; Borges et al. 2014; Upadhyay et al. 2015; Katayama-Ikegami et al. 2016). Some of these reference genes also showed species- or genotype-dependent variation in their gene expression (Graeber et al. 2011; Gutierrez et al.

2008; Monteiro et al. 2013). Therefore, it is important to check the stability of reference gene expression in species or interspecific hybrids, especially in grapes whose commercial cultivation includes several cultivars belonging to different species and interspecific hybrids. Most of the reference genes in grapes have been reported and validated in the genotypes of *Vitis vinifera*. In the present study, reference genes were validated on an interspecific hybrid developed from *V. vinifera* × *V. labrusca*, which implies that the results obtained have a broader application in grapes.

Ten potential reference genes were selected from an RNAseq database of ‘ARI 516’ based on their expression pattern in flowering and berry development stages. These genes include *ACT*, *RLI*, *ZNF*, *UBC17*, *UBC60*, *TUB*, *CYP5*, *VAG*, *UFD1*, and *UBC*. The expression stability of these genes was analyzed using geNorm, NormFinder, BestKeeper, ΔC_q , and RefFinder packages. These packages produced stability ranking with few differences due to differences in algorithms. The comprehensive ranking by RefFinder showed that the most stable reference genes were *UBC17*, *RLI*, and *ZNF*, followed by *ACT*, *UBC60*, and *TUB* with intermediate stability, whereas *UBC*, *UFD1*, *VAG*, and *CYP5* were the least stable candidates.

The BestKeeper algorithm calculates Pearson correlation coefficient (r) and standard deviation (SD), with lower values considered more stable and values that surpass the cutoff value ($SD = 1$) deemed to be unstable. The results of the BestKeeper analysis showed that *ACT*, *RLI*, *ZNF*, *UBC17*, *UBC60*, *TUB*, and *CYP5* were the most stable genes, all with $SD < 1$ and thus qualifying as reference genes, whereas *VAG*, *UFDI*, and *UBC* were not suitable as a reference gene ($SD > 1$) (Rodrigues et al. 2014; Zhou et al. 2018). In the present study, *UBC17* was the most stable gene expressed by RefFinder and geNorm. *UBC17* was recently identified as the most stable reference gene in *Toona ciliata* leaves under the treatment of methyl jasmonate (Song et al. 2020), and also in *Stipa grandis* under cold stress, treatment of jasmonic acid and salicylic acid (Wan et al. 2017). *RLI* was the second most stable gene according to comprehensive ranking and the top most stable according to NormFinder and Delta Cq method. It is a member of the large family of ATP-binding cassette (ABC) proteins that function as membrane transporters. To our knowledge, *RLI* has been reported for the first time as a stable reference gene in plants. Our results suggest that *UBC17*, *RLI*, and *ZNF* were more stable reference genes than the most commonly used reference gene *ACT*. According to BestKeeper algorithm, *ACT* ranked first among all the reference genes; however, the comprehensive ranking of *ACT* was 4th among the reference genes tested. Therefore, caution should be exercised when using *ACT* as a reference gene to normalize gene expression in grape flower development and berry formation. These results agree with previous studies that recommended *ACT* as an unreliable gene for normalizing gene expression (Monteiro et al. 2013; Upadhyay et al. 2015). *TUB* was not very stable in previously reported studies (Borges et al. 2014; Rodrigues et al. 2014; Upadhyay et al. 2015), and the same was observed in our study.

CONCLUSION

Using five different algorithms, we identified *UBC17*, *ZNF*, and *RLI* as novel reference genes with stable expression in flower and berry development in interspecific hybrids of cultivated grapes. Our results also demonstrated the reliability of *UBC17*, *ZNF*,

and *RLI* as reference genes to normalize the transcript of *VvAGLII*, which further validates the suitability of newly identified reference genes for normalization of gene expression in flower and berry development in grapes. The reference genes identified in the present study and previous studies will provide more choices for selecting appropriate genes to normalize gene expression in grapes.

Acknowledgments

The research was performed with financial support from the New Delhi Science and Engineering Research Board under research grant CRG/2018/001239.

Author contributions

The authors confirm their contribution to the paper as follows: study conception and design – R. Patil and S. Tetali; data collection and analysis – S. Chavan and A. Khatal; interpretation of results – R. Patil, S. Chavan, and A. Khatal, plant tissue collection S. Chavan, S. Tetali, and S. Phalake, draft manuscript preparation – S. Chavan, A. Khatal, and R. Patil. All authors reviewed the results and approved the final version of the manuscript.

Conflict of interest

The authors declare no conflict of interest, financial or other aspects.

REFERENCES

- Alva O., Roa-Roco R.N., Pérez-Díaz R., Yáñez M., Tapia J., Moreno Y., et al. 2015. Pollen morphology and boron concentration in floral tissues as factors triggering natural and GA-induced parthenocarpic fruit development in grapevine. PLoS ONE 10(10): e0139503; 18 p. DOI: 10.1371/journal.pone.0139503.
- Andersen C.L., Jensen J.L., 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 Research 64(15): 5245–5250. DOI: 10.1158/0008-5472.can-04-0496.
- Borges A.F., Fonseca C., Ferreira R.B., Lourenço A.M., Monteiro S. 2014. Reference gene validation for quantitative RT-PCR during biotic and abiotic stresses in *Vitis vinifera*. PLoS ONE 9(10): e111399; 9 p. DOI: 10.1371/journal.pone.0111399.

- Bustin S.A., Benes V., Garson J.A., Hellemans J., Huggett J., Kubista M. et al. 2009. The MIQE guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry* 55(4): 611–622. DOI: 10.1373/clinchem.2008.112797.
- Chuaqui R.F., Bonner R.F., Best C.J.M., Gillespie J.W., Flaig M.J., Hewitt S.M. et al. 2002. Post-analysis follow-up and validation of microarray experiments. *Nature Genetics* 32(Supplement 4): 509–514. DOI: 10.1038/ng1034.
- Coito J.L., Rocheta M., Carvalho L., Amâncio S. 2012. Microarray-based uncovering reference genes for quantitative real time PCR in grapevine under abiotic stress. *BMC Research Notes* 5: 220; 12 p. DOI: 10.1186/1756-0500-5-220.
- Coombe B.G. 1995. Adoption of a system for identifying grapevine growth stages. *Australian Journal of Grape and Wine Research* 1(2): 104–110. DOI: 10.1111/j.1755-0238.1995.tb00086.x.
- Deluc L.G., Grimplet J., Wheatley M.D., Tillett R.L., Quilici D.R., Osborne C. et al. 2007. Transcriptomic and metabolite analyses of Cabernet Sauvignon grape berry development. *BMC Genomics* 8: 429; 42 p. DOI: 10.1186/1471-2164-8-429.
- Fasoli M., Dal Santo S., Zenoni S., Tornielli G.B., Farina L., Zamboni A. et al. 2012. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. *Plant Cell* 24(9): 3489–3505. DOI: 10.1105/tpc.112.100230.
- González-Agüero M., García-Rojas M., Di Genova A., Correa J., Maass A., Orellana A., Hinrichsen P. 2013. Identification of two putative reference genes from grapevine suitable for gene expression analysis in berry and related tissues derived from RNA-Seq data. *BMC Genomics* 14: 878; 12 p. DOI: 10.1186/1471-2164-14-878.
- Graeber K., Linkies A., Wood A.T.A., Leubner-Metzger G. 2011. A guideline to family-wide comparative state-of-the-art quantitative RT-PCR analysis exemplified with a Brassicaceae cross-species seed germination case study. *Plant Cell* 23(6): 2045–2063, DOI: 10.1105/tpc.111.084103.
- Gutierrez L., Mauriat M., Guénin S., Pelloux J., Lefebvre J.-F., Louvet R. et al. 2008. The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnology Journal* 6(6): 609–618. DOI: 10.1111/j.1467-7652.2008.00346.x.
- Katayama-Ikegami A., Katayama T., Takai M., Sakamoto T. 2016. Reference gene validation for gene expression studies using quantitative RT-PCR during berry development of ‘Aki Queen’ grapes. *Vitis* 55(4): 157–160. DOI: 10.5073/vitis.2016.55.157-160.
- Kuhn N., Guan L., Dai Z.W., Wu B.-H., Lauvergeat V., Gomès E. et al. 2014. Berry ripening: recently heard through the grapevine. *Journal of Experimental Botany* 65(16): 4543–4559. DOI: 10.1093/jxb/ert395.
- Luo M., Gao Z., Li H., Li Q., Zhang C., Xu W. et al. 2018. Selection of reference genes for miRNA qRT-PCR under abiotic stress in grapevine. *Scientific Reports* 8: 4444; 11 p. DOI: 10.1038/s41598-018-22743-6.
- Ma L., Sun L., Guo Y., Lin H., Liu Z., Li K., Guo X. 2020. Transcriptome analysis of table grapes (*Vitis vinifera* L.) identified a gene network module associated with berry firmness. *PLoS ONE* 15(8): e0237526; 15 p. DOI: 10.1371/journal.pone.0237526.
- Massonnet M., Fasoli M., Tornielli G.B., Altieri M., Sandri M., Zuccolotto P. et al. 2017. Ripening transcriptomic program in red and white grapevine varieties correlates with berry skin anthocyanin accumulation. *Plant Physiology* 174(4): 2376–2396. DOI: 10.1104/pp.17.00311.
- Mejía N., Soto B., Guerrero M., Casanueva X., Houel C., de los Ángeles Miccono M. et al. 2011. Molecular, genetic and transcriptional evidence for a role of *VvAGL1* in stenospermocarpic seedlessness in grapevine. *BMC Plant Biology* 11: 57; 18 p. DOI: 10.1186/1471-2229-11-57.
- Monteiro F., Sebastiana M., Pais M.S., Figueiredo A. 2013. Reference gene selection and validation for the early responses to downy mildew infection in susceptible and resistant *Vitis vinifera* cultivars. *PLoS ONE* 8(9): e72998; 10 p. DOI: 10.1371/journal.pone.0072998.
- Muñoz-Espinoza C., Di Genova A., Correa J., Silva R., Maass A., González-Agüero M. et al. 2016. Transcriptome profiling of grapevine seedless segregants during berry development reveals candidate genes associated with berry weight. *BMC Plant Biology* 16: 104; 17 p. DOI: 10.1186/s12870-016-0789-1.
- Nwafor C.C., Gribaudo I., Schneider A., Wehrens R., Grando M.S., Costantini L. 2014. Transcriptome analysis during berry development provides insights into co-regulated and altered gene expression between a seeded wine grape variety and its seedless somatic variant. *BMC Genomics* 15: 1030; 22 p. DOI: 10.1186/1471-2164-15-1030.

- Pfaffl M.W., Tichopad A., Prgomet C., Neuvians T.P. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pairwise correlations. *Biotechnology Letters* 26: 509–515. DOI: 10.1023/b:bile.0000019559.84305.47.
- Rodrigues T.B., Khajuria C., Wang H., Matz N., Cunha Cardoso D., Valicente F.H. et al. 2014. Validation of reference housekeeping genes for gene expression studies in Western corn rootworm (*Diabrotica virgifera virgifera*). *PLoS ONE* 9(10): e109825; 8 p. DOI: 10.1371/journal.pone.0109825.
- Royo C., Carbonell-Bejerano P., Torres-Pérez R., Nebish A., Martínez Ó., Rey M. et al. 2016. Developmental, transcriptome, and genetic alterations associated with parthenocarpy in the grapevine seedless somatic variant Corinto blanco. *Journal of Experimental Botany* 67(1): 259–273. DOI: 10.1093/jxb/erv452.
- Selim M., Legay S., Berkelmann-Löhnertz B., Langen G., Kogel K.-H., Evers D. 2012. Identification of suitable reference genes for real-time RT-PCR normalization in the grapevine-downy mildew pathosystem. *Plant Cell Reports* 31(1): 205–216. DOI: 10.1007/s00299-011-1156-1.
- Silver N., Best S., Jiang J., Thein S.L. 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Molecular Biology* 7: 33; 9 p. DOI: 10.1186/1471-2199-7-33.
- Song H., Mao W., Duan Z., Que Qi., Zhou W., Chen X., Li P. 2020. Selection and validation of reference genes for measuring gene expression in *Toona ciliata* under different experimental conditions by quantitative real-time PCR analysis. *BMC Plant Biology* 20: 450; 14 p. DOI: 10.1186/s12870-020-02670-3.
- Sweetman C., Wong D.C.J., Ford C.M., Drew D.P. 2012. Transcriptome analysis at four developmental stages of grape berry (*Vitis vinifera* cv. Shiraz) provides insights into regulated and coordinated gene expression. *BMC Genomics* 13: 691; 25 p. DOI: 10.1186/1471-2164-13-691.
- Tashiro R.M., Philips J.G., Winefield C.S. 2016. Identification of suitable grapevine reference genes for qRT-PCR derived from heterologous species. *Molecular Genetics and Genomics* 291(1): 483–492. DOI: 10.1007/s00438-015-1081-z.
- Tetali S., Karkamkar S.P., Phalake S.V. 2020. Mutation breeding for inducing seedlessness in grape variety ARI 516. *International Journal of Minor Fruits, Medicinal and Aromatic Plants* 6(2): 67–71.
- Upadhyay A., Jogaiah S., Maske S.R., Kadoo N.Y., Gupta V.S. 2015. Expression of stable reference genes and *SPINDLY* gene in response to gibberellic acid application at different stages of grapevine development. *Biologia Plantarum* 59(3): 436–444. DOI: 10.1007/s10535-015-0521-2.
- Vandesompele J., De Preter K., Pattyn F., Poppe B., Van Roy N., De Paepe A., Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3(7): research0034.1; 12 p. DOI: 10.1186/gb-2002-3-7-research0034.
- Wan Q., Chen S., Shan Z., Yang Z., Chen L., Zhang C. et al. 2017. Stability evaluation of reference genes for gene expression analysis by RT-qPCR in soybean under different conditions. *PLoS ONE* 12(12): e0189405; 14 p. DOI: 10.1371/journal.pone.0189405.
- Wang L., Yin X., Cheng C., Wang H., Guo R., Xu X. et al. 2015. Evolutionary and expression analysis of a MADS-box gene superfamily involved in ovule development of seeded and seedless grapevines. *Molecular Genetics and Genomics* 290(3): 825–846. DOI: 10.1007/s00438-014-0961-y.
- Wang L., Hu X., Jiao C., Li Z., Fei Z., Yan X. et al. 2016. Transcriptome analyses of seed development in grape hybrids reveals a possible mechanism influencing seed size. *BMC Genomics* 17: 898. DOI: 10.1186/s12864-016-3193-1.
- Xie F., Xiao P., Chen D., Xu L., Zhang B. 2012. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Molecular Biology* 80(1): 75–84. DOI: 10.1007/s11103-012-9885-2.
- Zenoni S., Ferrarini A., Giacomelli E., Xumerle L., Fasoli M., Malerba G. et al. 2010. Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. *Plant Physiology* 152(4): 1787–1795. DOI: 10.1104/pp.109.149716.
- Zhou L., Chen F., Ye J., Pan H. 2018. Selection of reliable reference genes for RT-qPCR analysis of *Bursaphelenchus mucronatus* gene expression from different habitats and developmental stages. *Frontiers in Genetics* 9: 1–11. DOI: 10.3389/fgene.2018.00269.