Identification of suitable reference genes by quantitative real-time PCR for gene expression normalization in sunflower

Ana Claudia Ochogavía[1*], María Angelina Novello[1], Liliana Amelia Picardi[1,2], Graciela María Nestares[1]

[1]IICAR, UNR, CONICET (Instituto de Investigaciones en Ciencias Agrarias de Rosario, Universidad Nacional de Rosario, Consejo Nacional de Investigaciones Científicas y Técnicas), Zavalla, Santa Fe, Argentina
[2]CIUNR (Consejo de Investigaciones de la Universidad Nacional de Rosario), Argentina

*Correspondence: anaochogavia@conicet.gov.ar

Abstract

Quantitative real-time PCR (qPCR) is currently the most accurate method for detecting differential gene expression, but depends greatly on normalization with stably expressed housekeeping genes. Transcriptomics analyses and experimental validation in different plant species have shown that the reliability of these endogenous controls can be influenced by, growth conditions and organs/tissues examined. Thus, reliable validation of reference genes is required to ensure proper normalization. This paper presents a systematic comparison of ten potential reference genes in sunflower: five commonly used genes (Actin, Elongation Factor1, Plastid-encode RNA polymerase, Tubulin, and Ubiquitin, as ACT, EF1, PEP, TUB, and UBQ respectively), as well as five new candidates (Translation initiation factor, MicroRNA precursors 171 and 156, Ask-interacting protein, and Protein of unknown function, as ETIF5, MIR171, MIR156, SKIP, and UNK2 respectively). Reference gene expression stability was examined by qPCR across 20 biological samples, representing different tissues at various developmental stages. Expression of all 10 genes was variable to some extent, but that of ACT, UNK2, and EF1 was overall the most stable. A combination of ETIF5/UNK2/EF1 would be appropriate to use as a reference panel for normalizing gene expression data among vegetative tissues, whereas the combination of ACT/MIR156/UNK2 is most suitable for reproductive tissues. Reference genes selected in this study were further validated by examining relative expression of *ahas1*, one of three acetohydroxyacid synthase genes of sunflower. Our identification and validation of suitable normalizer genes will be of use to ensure accurate results in future transcriptomics studies in this crop.

Key words: *Helianthus annuus* L.; qPCR; Reference genes panel; Stably expressed genes; UNK2.
Abbreviations: ACT_Actin; AHAS_Acetohydroxyacid synthase; CV_coefficient of variation; EF1_elongation factor 1; ETIF5_Eukaryotic translation initiation Factor 5; M_average expression stability value; MIR/miRNA_microRNA; NF_normalization factor; PEP_plastid-encoded RNA polymerase; qRT–PCR_Real-time quantitative polymerase chain reaction; SKIP_SKIP1/Ask-Interacting Protein; Sv_Standard variation value; TUB_Tubulin subunit β; UBQ_Ubiquitin; V_pairwise variation value.

Introduction

Gene expression studies require sensitive, precise, and reproducible measurements of transcript levels. These studies are affected by a number of experimental conditions, including variations in sample quality, RNA stability, as well as differences in cDNA synthesis and PCR efficiency. Implementation of transcriptomics techniques, such as quantitative real-time PCR (qPCR), requires amplification of a reference gene to evaluate experimental variability (Bustin 2002). Given that inappropriate normalization can radically change the interpretation of gene expression patterns and thus the identification of gene function, a suitable reference gene needs to be expressed at a constant level in different tissues and under different experimental conditions (Andersen et al., 2004).

In plant species, the number of validated reference genes is still limited. Most genes used as a reference in qPCR assays are involved in cytoskeleton structure, such as Actin and Tubulin, or in principal cellular processes, like Elongation Factor 1 or Ubiquitin (Bomal et al., 2008; Bracha-Drori et al., 2008; Cuevas et al., 2008; Chen et al., 2011; Wan et al., 2010). Many of these genes were adopted from traditional

techniques, such as Northern blot, and are assumed to be invariable because of their constitutive cellular function (Jain et al., 2006) but it has been shown that some of them may change its expression in different tissues and/or experimental conditions (Kim et al., 2003; Nicot et al., 2005; Wan et al., 2010). Nonetheless these genes are being widely used but performing a preliminary stability study under particular experimental conditions (Fernandez et al., 2011; Hewezi et al., 2006; Ochogavia et al., 2014).

Numerous candidate reference gene studies have been achieved in different species of plants, including potato (Nicot et al., 2005), tomato (Exposito-Rodriguez et al., 2008), soybean (Hu et al., 2009), peanut (Chi et al., 2012), and sesame (Wei et al., 2013), among others. Most of these studies compare traditionally used reference genes among plant tissues, developmental stages, genotypes, and/or experimental conditions. No universally accurate normalizing genes were discovered, thus high-throughput studies were aimed to find new reference genes. Novel stably expressed candidates included transcription factors, protein-coding sequences of unknown function, and small non-coding

RNAs, specifically precursors of microRNA (miRNAs). Several of these nontraditional genes were reported to be the best ranked as normalizing genes in different tissues and experimental conditions.

Sunflower (*Helianthus annuus* L.) is the second most important oilseed crop in the world. It is also regarded as an important source of biodiesel but weed competition causes substantial yield loss in this crop. Group B herbicides are specific inhibitors of the acetohydroxyacid synthase (AHAS) and control a wide spectrum of grasses and broadleaf weeds at low application rates. In sunflower, this enzyme is coded by three paralog genes (*ahas1*, *ahas2* and *ahas3*) (Kolkman et al., 2004), and point mutations on one of them (*ahas1*) confers herbicide resistance (Sala et al., 2012). We recently reported tissue-specific regulation of *ahas* genes in a susceptible line of sunflower (Ochogavia et al., 2014). Further expression studies of *ahas* genes in cultivars carrying resistant alleles could be useful for the characterization of mechanisms regulating *ahas* gene expression. However, only a limited number of traditionally used reference genes have been applied as normalizers in this species (Breccia et al., 2013; Fernandez et al., 2011; Ochogavia et al., 2014) and no previous comparison on nontraditional reference genes were reported in this species. The aim of this work was to evaluate the stability of ten genes: five traditionally used and five new putative genes in different sunflower tissues and developmental stages. Identification and consistent validation of new stable reference genes in a wide range of tissues and developmental stages will be helpful for future transcriptomics studies in sunflower.

Results

Identification and characterization of sunflower candidate reference genes

A group of traditional and novel protein-coding reference genes, as well as miRNAs and unknown expressed sequences were interrogated to select the most reliable genes for use as normalizers in ten different sunflower tissues and developmental stages. Five protein-coding, traditional reference genes (ACT, EF1, PEP, TUB, and UBQ) that have been reported to be good potential candidates in previously published studies (Czechowski et al., 2005; Fernandez et al., 2011; Jain et al., 2006) were selected as candidates. Two protein-coding novel reference genes (elongation initiation factor gamma subunit 5 ETIF5 and SKIP1/Ask-Interacting Protein- SKIP) that have been reported as the most stable in recently published articles (Borowski et al., 2014; Hu et al., 2009) were also included in the study. Sunflower protein-coding candidate reference genes were extracted from expressed sequences of *Helianthus annuus* L. in the GenBank database. Homology of these seven protein-coding genes in sunflower and *Arabidopsis* was all high with E values of\1e-20. An unknown expressed coding sequence was also studied (UKN2). This sequence has been identified as highly stable by microarray analysis in soybean (Libault et al., 2008) and its stability was then confirmed in different tissues, developmental stages, genotypes, and environmental conditions (Hu et al., 2009). The amino acid sequence of this unknown protein was compared with the ESTs translated sunflower database to identify sequences with high sequence identity. Homology of UKN2 aminoacid sequence in sunflower was very high with E values of\1e-20. Moreover,

two miRNA stem loop sequences that have been recently reported as highly stable in related species (Borowski et al., 2014) were evaluated. *Helianthus annuus* miRNA precursor sequences were identified from the MIR156 and MIR170/171 gene family sequences as *ha-MIR156a* and *ha-MIR170/171a* on the miRBASE. All primer sequences and relevant information regarding the genes are presented in Table 1.

Gene expression profiles

The expression stabilities of these potential sunflower normalizer genes were assessed by qPCR on two biological replicates of ten tissue samples obtained from different developmental stages. Samples were also categorized into vegetative and reproductive tissues. Vegetative tissues correspond to: 8 and 15 days post-germination leaves (L8 and L15 respectively); 8 and 15 days post-germination roots (R8 and R15 respectively). Reproductive tissues correspond to 9 and 25 post-pollination embryos (E9 and E25 respectively) and tissues from dissected immature flowers, i.e. style, stigma and ovary tissue (F-E1f) and pistils (F-E1m), and tissues from dissected mature flowers, i.e. style, stigma and ovary tissue (F-E3f) and pistils (F-E3m).

The specificity of amplified sequences was verified by the presence of a single peak in qPCR melting curves and a single band with the expected size in an electrophoresis assay (Fig. 1).

### Expression stability analysis of candidate reference genes

Expression levels of selected genes were compared and results were presented as quantification cycle values (Cq). Fig. 2 represents gene variability among samples. The mean Cq values ranged from 21.9 to 31.3, with most lying between 24 and 29. With regard to gene expression variation, MIR156 showed the highest value (37.1 cycles) while MIR171 the lowest (16 cycles) when all tissues were considered. No amplification was detected in the absence of templates. Expression variation was slightly lower when only vegetative or reproductive tissues were considered. In vegetative tissue, MIR156 showed the highest value (37.1 cycles), whereas the lowest value was observed for PEP (18.6 cycles). For reproductive tissue, UNK2 showed the highest value (37 cycles), whereas the lowest value was observed for MIR171 (16 cycles).

A preliminary analysis was performed to calculate Cq value variability through the conventional coefficient of variance (CV) statistical test. Tubulin gene, TUB, showed the least variation (CV of 8.8%), while microRNA precursor 171, MIR171, was the most variable among all samples (19.2%). However, plastid-encoded RNA polymerase, PEP, showed the least variation (CV of 4.1%) in leaves and roots and TUB had the lowest variability (CV of 9.3%) in embryos and flowers.

The average expression stability M value was calculated for all candidate genes with the geNorm program (Vandesompele et al., 2002). We independently analyzed ten candidate gene expression levels in i) all tissues, ii) vegetative tissues, and iii) reproductive tissues. When all samples were analyzed together, the geNorm M value was below 1.5 for elongation factor 1α (EF1), eukaryotic translation initiation factor 5 (ETIF5), and actin (ACT), which classifies these genes as the

Table 1. Gene ID and primer sequences used in gene expression analysis.

| Gene	Accession number	Primer sequence (5´-3´; Forward/Reverse)	Fragment size	Tm °C (F/R)
ACT	DY915068	CAGGCCGTGCTTTCCCTCTA/GGTCACGACCAGCGAGATCA	145 pb	64.8/64.8
EF1	DY907212	TGCCCAAGAAGTTGCTGGTG/ACGTGCCCAGGTGAGTCGAT	189 pb	64.6/65.3
ETIF5	DY929865	GGCCGATGCAGGTGCTTCCA/AGCATGACCGTGCTTGCCAGT	136 pb	65.6/65.3
MIR156	MI0022228	AACGAGATGTTGGTTCGGTTCAA/GCACGTGATATTGGCACGGCT	91 pb	60.9/63.1
MIR171	EE650377	ACGAGATGTTGGTTCGGTTCAA/CACGTGATATTGGCACGGCTCA	93 pb	60.9/63.1
PEP	DY912398	ACCGGGCCAATGATGAGGTGT/AGGAGCAAGCTGGCCCAACA	126 pb	64.2/64.5
SKIP	GE522343	ACGCTCGAACCTGAACACCGC/AAGAACATGCCCGCCGCCTC	135 pb	65.7/65.8
TUB	DY921120	TCTGCCACCATGTCGGGAGTT/GTAACGCCCGTGTCGTGGGTC	252 pb	64.1/65.3
UBQ	CX947085	CCTCTTCTTCGCGCCACCGC/CACCGGACCAACAGCGTCCC	131 pb	66.4/65.8
UKN2	GE512112	AGGAGGTAGCCGTCGTCCAGC/AGCCTTGGAGTTCAATTGGGCCG	142 pb	65.9/65.7
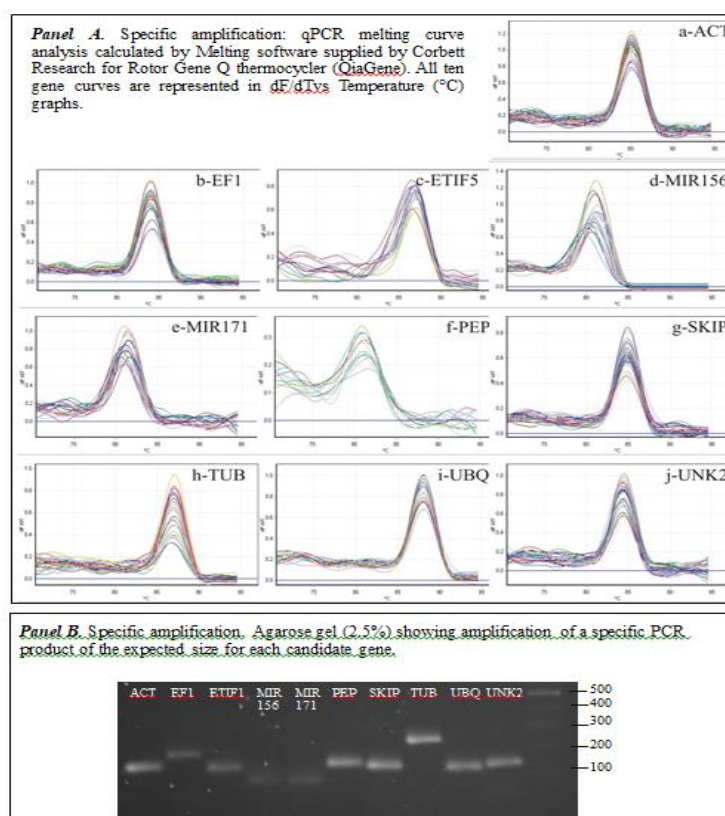
Fig 1. *Panel A.* Specific amplification: qPCR melting curve analysis calculated by Melting software supplied by Corbett Research for Rotor Gene Q thermocycler (QiaGene). All ten gene curves are represented in dF/dTvs Temperature (°C) graphs. Curves are named as a-ACT; b-EF1; c-ETIF5; d-MIR156; e-MIR171; f-PEP; g-SKIP; h-TUB; i-UBQ; j-UNK2. *Panel B.* Specific amplicon size: agarose gel (2.5 %) showing amplification of a specific PCR product of the expected size for the ten reference candidates. Line 1:MIR156; Line 2: MIR171; Line 3:TUB; Line 4: UNK2; Line 5: UBQ; Line 6: ACT; Line 7: PEP; Line 8: SKIP, Line 9: ETIF5; Line 10: EF1; Line 11: 100bp molecular marker.

Table 2. Gene scoring according to reference genes selection methods: NormFinder (NF), geNorm (GN) and BestKeeper (BK)

| Ranking	All Tissues			Vegetative Tissue			Reproductive Tissue		
	NF (Sv)	GN (M)	BK (R^2)	NF (Sv)	GN (M)	BK (R^2)	NF (Sv)	GN (M)	BK (R^2)
1	UNK2 (0.04)	EF1 (1.26*)	ACT (0.93)	SKIP (0.04)	ETIF5 (1.13*)	ETIF5 (0.94)	UNK2 (0.05)	EF1 (1.25*)	MIR156 (0.96)
2	SKIP (0.06)	ETIF5 (1.32*)	UNK2 (0.91)	UNK2 (0.04)	EF1 (1.25*)	UBQ (0.89)	MIR156 (0.05)	UBQ (1.29*)	UNK2 (0.94)
3	ACT (0.06)	ACT (1.40*)	EF1 (0.91)	ETIF5 (0.05)	UNK2 (1.33*)	UNK2 (0.87)	ACT (0.07)	ACT (1.29*)	ACT (0.95)
4	EF1 (0.07)	UBQ (1.62)	SKIP (0.89)	PEP (0.05)	UBQ (1.50)	EF1 (0.87)	SKIP (0.07)	MIR156 (1.76)	EF1 (0.93)
5	TUB (0.09)	UNK2 (1.82)	UBQ (0.86)	ACT (0.05)	ACT (1.7)	ACT (0.86)	EF1 (0.07)	SKIP (2.02)	SKIP (0.90)
6	PEP (0.10)	SKIP (2.19)	MIR171 (0.78)	EF1 (0.06)	TUB (1.84)	SKIP (0.74)	PEP (0.08)	UNK2 (2.11)	UBQ (0.87)
7	MIR156 (0.11)	TUB (2.26)	MIR156 (0.77)	TUB (0.07)	PEP (1.98)	TUB (0.49)	TUB (0.10)	MIR171 (2.25)	MIR171 (0.87)
8	UBQ (0.11)	PEP (2.41)	ETIF5 (0.71)	UBQ (0.09)	SKIP (2.12)	PEP (0.25)	UBQ (0.12)	PEP (2.41)	PEP (0.84)
9	ETIF5 (0.12)	MIR171 (2.69)	PEP (0.67)	MIR156 (0.09)	MIR156 (2.35)	MIR171 (0.22)	MIR171 (0.12)	TUB (2.67)	TUB (0.77)
10	MIR171 (0.12)	MIR156 (2.83)	TUB (0.68)	MIR171 (0.10)	MIR171 (2.48)	MIR156 (-0.18)	ETIF5 (0.15)	ETIF5 (2.93)	ETIF5 (0.69)

* indicates geNorm suitable reference genes (M values below 1.5)
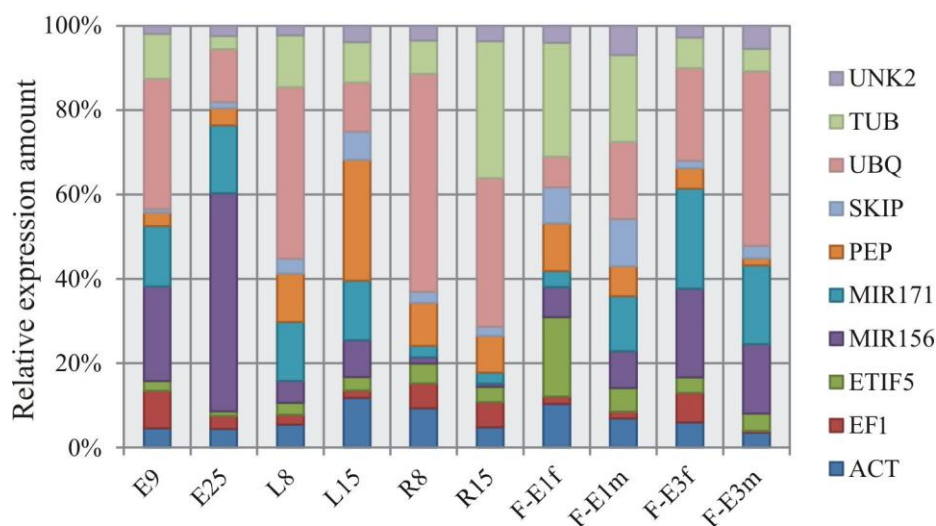
Fig. 2. Relative distribution of transcript levels from candidate genes among samples. Two biological and three technical replicates of each sample are represented. Transcript quantities are represented as percentages of the aggregated 10-transcript pool for each sample. Initials indicate tissue type: embryos (E), leaves (L), roots (R) and flowers (F). Numbers indicate developmental stages, 9 and 25 are days post-pollination, 8 and 15 are days post-germination. Flower tissues: ovary tissue of immature flowers (F-E1f), pistils of immature flowers (F-E1m), ovary tissue of mature flowers (F-E3f) and ovaries from mature flowers (F-E3m). Table 1 provides full gene names.

Table 3. Optimal sunflower reference genes in three sample groups by three methods.

Tissue groups	Optimal reference genes			
	NormFinder	geNorm	BestKeeper	Common genes
All tissues	UNK2	EF1	ACT	ACT
	SKIP	ETIF5	UNK2	EF1
	ACT	ACT	EF1	UNK2
Vegetative Tissues	SKIP	ETIF5	ETF5	ETIF5
	UNK2	EF1	UBQ	UNK2
	ETIF5	UNK2	UNK2	SKIP/EF1/UBQ
Reproductive Tissues	UNK2	EF1	MIR156	ACT
	MIR156	UBQ	UNK2	MIR156
	ACT	ACT	ACT	UNK2
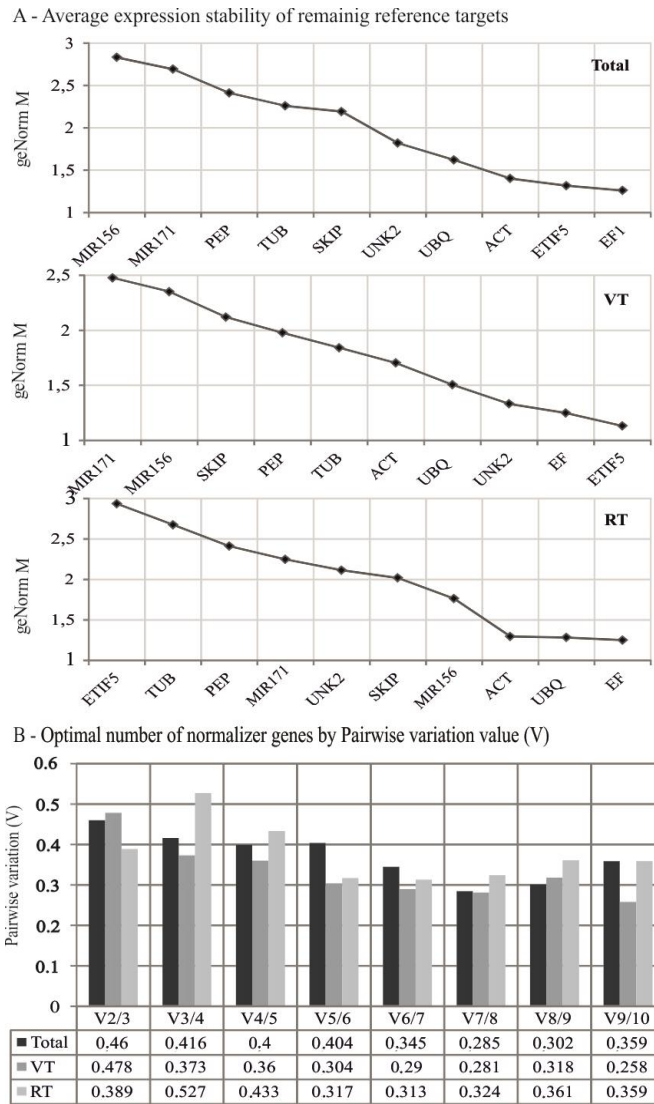
Fig 3. Results of analysis of the ten candidate reference genes by *geNorm* software. *Panel A:* Average expression stability value (M) from five tissues of sunflower and two developmental stages in each one. All tissues were analyzed together (Total); vegetative tissue (VT) and reproductive tissue (RT) sub-groups are also independently represented. A lower average expression stability M value indicates more stable expression. *Panel B:* Pairwise variation ($Vn/Vn + 1$) was analyzed to determine the optimal number of reference genes required for qPCR data normalization in a sample pool. Total refers to all tissues analyzed together; VT and RT are the sub-group of vegetative tissue and reproductive tissue respectively.
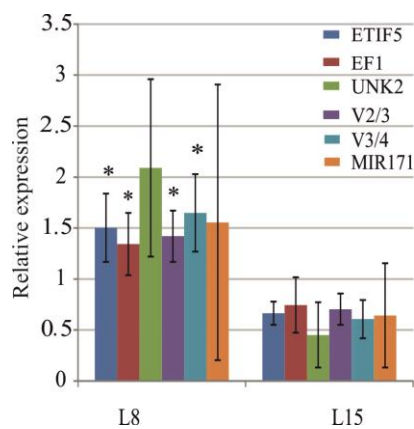
Fig. 4. Expression levels of *ahas1* in leaves of 8 days (L8) and 15 days after germination (L15). Relative expression was calculated with respect to the mean of each developmental stage. Reference genes used for normalize were ETIF5, EF1, UNK2, ETIF/UNK2, ETIF/UNK2/EF1 (the most stable reference genes of vegetative tissues), and MIR171 (the most unstable reference gene in these tissues). Asterisk (*) indicates significant differences ($p \leq 0.05$) between the two leaf developmental stages.

most stably expressed ones (Fig. 3-A). However, when vegetative and reproductive tissues were evaluated independently, different genes were identified as the best candidates (see Table 2). We further evaluate the optimal number of genes required for accurate normalization, pairwise variations ($Vn/Vn + 1$) were calculated with geNorm function. Accordingly to the trend observed in Fig. 3-B, the best three geNorm candidate genes should be required for the vegetative tissue sub-group normalization and the best fives in the reproductive tissue pool. However, none of the calculated V values were below the suggested cut-off number 0.15. Notwithstanding the threshold of V value has been proposed as a guidance of the minimal number of reference genes it has also been recommended not be consider as a strict cut-off number (Vandesompele et al., 2002). Furthermore, it was also proposed that normalization with the best three genes should be more accurate than the use of a unique reference gene. The sample pool was also analyzed using the BestKeeper and NormFinder software. Accordingly to the first application, the three best ranked genes were ACT, EF1, and the unknown protein 2 (UKN2) when all sunflower tissues were analyzed (Table 2). For NormFinder, best stability value was obtained for UNK2 when all the samples were analyzed as a group (Table 2). Overall, the three analysis strategies allowed the identification of different best normalizer genes in each tissue group. Both novel and traditional candidate genes were almost equally represented in the top half of rankings and some genes were specifically selected by the three algorithms. Table 3 shows the consensus of the three best genes in different groups of tissues analyzed by three statistical strategies. When all tissues were analyzed together, ACT was included in the three best ranks of the three statistical strategies; and UNK2 and EF1 were also included in the top three rankings, but only in two of the statistical strategies. Moreover, tissue-specific groups are comprised of different candidate genes. In vegetative tissues, the three analyses included ETIF5 and UKN2 in the top three best gene candidates, but the other third best gene could be SKIP, EF1, or UBQ. Likewise, in reproductive tissues the three rankings included ACT in the third position and two of them selected UNK2 and MIR156 as best-ranked genes.

Validation of selected genes

Selected genes were validated comparing the expression profile of *ahas1* to test if normalization by different reference genes modified qPCR expression levels. When the *ahas1* expression was normalized using the three best candidates for vegetative tissue: ETIF5, UNK2, and EF1 individually, the differences of transcript abundance gradually decreased as the standard error bars increased (Fig. 4). It is worth noting that when either two or three of the most stable genes were employed together for normalization, similar expression patterns were obtained. The expression profile achieved from normalization using the worst ranked genes differed from the other normalization strategies, and normalization combining more genes resulted in improved accuracy.

We present a systematic comparison and validation of expression stability from commonly used reference genes (ACT, EF1, PEP, TUB, and UBQ), and new candidates (ETIF5, MIR171, MIR156, SKIP, and UNK2). Expression levels of ACT, UNK2, and EF1 were overall the most stable in all plant tissues. A combination of ETIF5/UNK2/EF1 and ACT/MIR156/UNK2 would be appropriate to use as a normalizer in vegetative and reproductive tissues respectively.

Discussion

Quantitative real-time PCR is now commonly used in many laboratories to undertake accurate expression profiling of different genes of interest, as well as for validation of high-throughput transcriptional analysis. However, reliable relative quantification can only be achieved if accurate normalization is performed with suitable reference genes (Bustin 2002). Large scale transcriptional assays enable detection of suitable reference genes directly chosen from those data sets (Czechowski et al., 2005; Narsai et al., 2010; Zhu et al., 2013), but in specific expression studies that do not require application of these massive techniques, selection of candidate genes relies on the use of genes widely described as stable expressed in different systems. Genes involved in basic cellular processes (e.g. UBQ and EF1) or in cellular structure (e.g. ACT and TUB) have been widely used as normalizers in plants (Bomal et al., 2008; Bracha-Drori et al., 2008; Cuevas et al., 2008). Nonetheless, there is increasing evidence that expression of these genes has not been particularly stable in different plant species or under certain experimental conditions (Fernandez et al., 2011; Radonic et al., 2004; Selvey et al., 2001; Thorrez et al., 2008). For example, the gene Subunit 10 of Ubiquitin is very stable in *Arabidopsis* (Czechowski et al., 2005) and flax (Huis et al., 2010), but not in rice (Jain et al., 2006) or soybean (Hu et al., 2009). Thus, it is not always possible to assume that a suitable reference gene identified in one species will also be suitable in another species. Recently published analyses focus on identification of novel genes that keep stable expression among different genotypes, tissues, and experimental conditions. Transcription factors, protein-coding genes unknown function, and miRNAs have been recently proposed as the best normalization genes in soybean, lettuce, peanut, and sesame. Some of these novel genes have offered better results compared with traditionally used reference genes (Borowski et al., 2014; Chi et al., 2012; Hu et al., 2009; Wei et al., 2013).

As in other plant species, TUB, ACT, and EF1 have been the most frequently employed as reference genes in sunflower (Breccia et al., 2013; Moschen et al., 2014; Ochogavia et al., 2014; Ungerer and Kawakami 2013). However, no detailed validation has been carried out to test the suitability of these genes and novel genes in different sunflower tissues.

In this work, we report the evaluation of ten candidate reference genes among ten biological samples from sunflower tissues. Five of the genes were traditionally used, which includes genes that code for proteins involved in essential cellular processes (EF1, PEP and UBQ) and cytoskeleton structure (ACT and TUB). The five novel genes evaluated included a translation factor (ETIF5), a DNA-interacting protein (SKIP), a protein-coding gene of unknown protein (UNK2), and microRNA precursors (MIR156 and MIR171). The expression level of these genes was determined in five tissues (roots, leaves, female dissected flowers, male dissected flowers, and embryos) from two developmental stages.

Four mathematical and statistical models for quantification of gene expression variability were applied: comparisons through the CV of Cq values, M and V values with geNorm software, Sv with NormFinder application, and BKI with BestKeeper. According to CVs, the TUB gene had the lowest variability among all samples (CV= 8.3%). Consistent with previous results, the CV of this gene from a total of twenty different tissues from sunflower developmental stages was under 9% (Ochogavia et al., 2014). However, CV cannot

clearly define the most stably expressed reference gene among a group of genes because it measures the individual variability of each gene. The three methods selected ACT as one of the first three best-ranked genes of all tissues and developmental stages. NormFinder and BestKeeper also included UNK2 in the best three candidate genes, and geNorm and BestKeeper selected EF1 as one of the best three candidate genes. In this work, employing three statistical methods allowed us to obtain clearly conserved trends, but the best and worst ranked individual genes were not exactly the same. Nonetheless, differences between rankings were expected since they are based on divergent statistical algorithms. Here, both traditional reference genes and novel ones were classified as the most stable genes among all samples. Previously, conflicting results were reported on traditional gene expression in several plant species. While the Actin 2/7 gene was identified as the most stably expressed in soybean (Jian et al., 2008) and peanut (Chi et al., 2012), it has been reported to perform poorly as a normalizer in rice, potato, and sugar cane (Iskandar et al., 2004; Kim et al., 2003; Nicot et al., 2005). Likewise, the Actin gene was recognized as an accurate reference gene in sunflower leaves (Fernandez et al., 2011; Ungerer and Kawakami 2013). Elongation Factor-1α gene has been described as highly stable in different tissues of rice (Jain et al., 2006), but unstable in organs of tomato (Exposito-Rodriguez et al., 2008). While certain genes traditionally used as references exhibit stable expression levels in some plant species, they may not be suitable as normalizers in other ones. On the other hand, recent studies with novel candidate genes have reported high stability of these genes in tissues and stages of different plant species. Consistent with our results, UKN2 was found to be one of the most stable genes from a total of 14 examined genes across 116 biological samples of soybean (Hu et al., 2009), as well as in various tissue types, seed developmental stages, and treatments in peanut (Chi et al., 2012) and tomato (Exposito-Rodriguez et al., 2008).

In the present work, two tissue sub-groups were analyzed independently. Different genes were selected as good normalizers in each group, but UNK2 remained at the top of the three rankings. Tissue-specific gene expression of some genes was also detected. While ETIF5 was one of the most stable genes in vegetative tissues, it exhibited the worst stability among reproductive tissues. The opposite behavior was observed for MIR156, which was one of the most stable genes in reproductive tissues and the worst in vegetative ones. Similarly, MIR156 was reported as the most variable microRNA precursor gene from four MIRs evaluated in lettuce leaves exposed to different kinds of abiotic stresses (Borowski et al., 2014). ETIF5 and MIR156 exhibit clear tissue-specific expression in sunflower, but they should be accurate normalizers if used in the respective tissue. Pairwise variation analysis did not allow us to identify the optimal number of reference genes. However, as suggested in the qBASE software manual, the pairwise variation $(Vn/Vn + 1)$ graph should serve as a guide for determination of the number of reference genes and not be a strict cut-off parameter (www.biogazelle.com). Three different best-genes groups were identified when all tissues were analyzed together, as well as when they were separated into vegetative and reproductive sub-groups.

For validation studies, *ahas1* gene expression was determined in young and expanded leaves. In a previous study, we had already measured the expression profile of this gene from different tissues and a wide range of developmental stages. The *ahas1* expression profile confirmed its over-representation in young leaves with respect to older ones (Ochogavia et al., 2014). In this work, *ahas1* was normalized using individual best-ranked genes selected by the three algorithms for vegetative tissues, and with the Normalization Factor of two and three of these common genes (Table 3). Although all normalization strategies allowed detection of similar *ahas1* transcript abundance in leaves tissues, significant differences were obtained when using the best-ranked reference genes.

Materials and Methods

Plant material

The public sunflower inbred line HA89 released by USDA was used. Vegetative tissue was obtained from seedlings grown in pots as described Breccia *et al.* (2013) during 15 days. Leaf and root samples were collected at 8 and 15 days post germination. Furthermore, reproductive tissues were obtained from field grown plants under drip irrigation conditions at the Experimental Field Station of the Universidad Nacional de Rosario (33∘1_S; 60∘53_W). Reproductive tissue samples were collected at different stages: from the first reproductive stage (R5) to physiological maturity (R9) (Schneiter and Miller 1981). Immature and mature disc flowers were collected from the flower bud and male reproductive stages, respectively, as described by Miller (1987). Immature embryos were collected at 9 and 25 days after pollination. All collected tissues were immediately frozen until RNA extraction.

Selection of candidate genes and design of specific primers

A total of ten genes were selected to assess their stability in different tissues and developmental stages. Five of these genes have been traditionally used in expression studies (Actin, Elongation Factor 1α, Plastid-encoded RNA polymerase, Tubulin subunit β, and Ubiquitin, as ACT, EF1, PEP, TUB, and UBQ respectively) (Fernandez et al., 2011; Jain et al., 2006) and the other five have recently been proved as good normalizers in gene expression analysis in another plant species (Eukaryotic translation initiation Factor 5, MicroRNA precursor 171, MicroRNA precursor 156, SKIP1/Ask-Interacting Protein, and Unknown protein 2 as ETIF5, MIR171, MIR156, SKIP, and UNK2 respectively) (Borowski et al., 2014; Hu et al., 2009; Wei et al., 2013). Sunflower protein-coding genes sequences were extracted from GenBank database (http://www.ncbi.nlm.nih.gov/genbank/). The sunflower orthologous gene UNK2 (At4G33380) was obtained by comparing the *Arabidopsis* gene sequence with the sunflower expressed sequences database via the MegaBLAST program (http://www.blast.ncbi.nlm.nih.gov/Blast.cgi). Sunflower precursors of orthologous miRNA (MIR156 and MIR171) were extracted from *Helianthus annuus* miRNA database in miRBASE website (http://www.mirbase.org/). Gene accession numbers are listed in Table 1. Specific primers were designed using the Primer3Plus program (http://www.bioinformatics.nl/cgi-bin / primer3plus / primer3plus.cgi) with options for GC content greater than 60%, Tm of 63 -66 °C, and fragment size from 90 to 255 nucleotides. Primer specificity was achieved by contrasting primer sequences against sunflower non-redundant sequences using the PrimerBlast from NCBI (http://blast.ncbi.nlm.nih.gov/primer-blast.cgi). Primer sequences, fragment sizes, and melting temperatures are detailed in Table 1.

RNA extraction and qPCR analysis

RNA was extracted from 80-180 mg of each sample using PureLink™ RNA Mini Kit (Invitrogen Life Technologies) according to the instructions provided by the manufacturer. Two independent RNA extractions were performed from each tissue. DNAse I treatment was performed using PureLink™ DNAse with column application (Invitrogen Life Technologies). RNA integrity was verified on a 1% agarose gel. RNA was stored at −80∘C until use. First strand cDNA synthesis was performed by the SuperScript® First-Strand Synthesis System for RT-PCR (Invitrogen Life Technologies) according to the manufacturer's instructions. Reverse transcription was initiated using oligo-dT supplied with the kit. cDNA was diluted 1:5 with nuclease-free water. A checklist outlining the RNA to qRT-PCR quality/methodology as described in Bustin et al. (2009) is presented in Supplementary Table 1.

Real-time PCR assays were performed using the Rotor-Gene Q with HRM (Qiagen®) thermal cycler. The reaction medium contained 1× SYBR Green PCR Master Mix (Mezcla Real®, Biodynamics, Argentina), 0.4µM of the forward and reverse primers and 1µL of cDNA in a total volume of 15µL. An initial denaturalization step was included (2 min at 70∘C) followed by 40 cycles that consisted of denaturation for 15s at 95°C, annealing for 30s at 63–65°C, and extension for 40s at 72°C. All reactions were run by triplicate on two independent biological replicates. No template controls were also included. The Cq and efficiency values (E) of each independent qPCR run were obtained from the Comparative Quantitation software supplied by Corbett Research for Rotor Gene. Primer amplification specificity was verified for the lack of primer dimers or non-specific amplicons by the presence of a single peak in qPCR melting curve products and the presence of a single band of the expected size on a 2.5% agarose gel. Electrophoresis was performed in a horizontal cube in 1X TBE buffer maintaining constant amperage at 60mV and was observed on UV light.

Reference gene expression stability determination

The conventional coefficient of variance (CV) statistical test was used to calculate Cq values variability. Expression stability of candidate genes was analyzed with the geNorm function (Vandesompele et al., 2002), and two publicly available software applications, i.e., BestKeeper (Pfaffl et al., 2004) and NormFinder (Andersen et al., 2004).

Validation of reference gene analysis

Expression normalization of the transcript levels of subunit 1 of acetohydroxyacid synthase, *ahas1,* in leaves tissues was performed with the three most stable genes, as well as the less stable gene detected for vegetative tissue. Relative expression was calculated with respect to the expression minimal value of each sample, and differences between them were evaluated using qBASE software (Vandesompele et al., 2002).

Conclusion

As it was demonstrated in the present study, commonly used reference genes should be carefully evaluated for each experimental condition tested, since their expression may be significantly affected depending on organ identity or experimental assay. In general, novel stably expressed genes, like UNK2, provide a promising option as a universally suitable normalizer gene. The present identification and selection of suitable normalizer genes will facilitate future development of transcriptomics studies in sunflower.

References

Andersen CL, Jensen JL and Orntoft TF (2004) Normalization of realtime 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(15):5245–5250.

Bomal C, Bedon F, Caron S, Mansfield S, Levasseur C, Cooke J, Blais S, Tremblay L, Morency M and Pavy N (2008) Involvement of *Pinus taeda* MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis. J Exp Bot. 59(14):3925-3939.

Borowski J, Galli V, Messias R, Perin E, Hugh Buss J, Delmar dos Anjos e Silva S and Rombaldi C (2014) Selection of candidate reference genes for real-time PCR studies in lettuce under abiotic stresses. Planta. 239(6):1187–1200.

Bracha-Drori K, Shichrur K, Lubetzky T and Yalovsky S (2008) Functional analysis of *Arabidopsis* postprenylation CaaX processing enzymes and their function in subcellular protein targeting. Plant Physiol. 148(1):119-131.

Breccia G, Vega G, Felitti S, Picardi L and Nestares G (2013) Differential expression of acetohydroxyacid synthase genes in sunflower plantlets and its response to imazapyr herbicide. Plant Sci. 208(1):28-33.

Bustin SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT -PCR): trends and problems. J Mol Endocrinol. 29(1):23-39.

Cuevas J, López-Cobollo R, Alcázar R, Zarza X, Koncz R, Altabella T, Salinas J, Tiburcio A and Ferrando A (2008) Putrescine is involved in *Arabidopsis* freezing tolerance and cold acclimation by regulating abscisic acid levelsin response to low temperature. Plant Physiol. 148(2):1094-1105.

Czechowski T, Stitt M, Altmann T, Udvardi MK and Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. Plant Physiol. 139(1):5-17.

Chen L, Zhong H, Kuang J, Li JG, Lu WJ and Chen JY (2011) Validation of reference genes for RT-qPCR studies of gene expression in banana fruit under different experimental conditions. Planta. 234(2):377–390.

Chi Y, Hu R, Yang Q, Zhang X, Pan L, Chen N, Chen M, Yang Z, Wang T, He W and Yu S (2012) Validation of reference genes for gene expression studies in peanut by quantitative real-time RT-PCR. Mol Genet Genomics. 287(1):167-176.

Exposito-Rodriguez M, Borges AA, Borges-Perez A and Perez JA (2008) Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol. 8(131)

Fernandez P, Di Rienzo JA, Moschen S, Dosio GA, Aguirrezabal LA, Hopp HE, Paniego N and Heinz RA (2011) Comparison of predictive methods and biological

validation for qPCR reference genes in sunflower leaf senescence transcript analysis. Plant Cell Rep. 30(1):63-74.

Hewezi T, Léger M, El Kayal W and Gentzbittel L (2006) Transcriptional profiling of sunflower plants growing under low temperatures reveals an extensive down-regulation of gene expression associated with chilling sensitivity. J Exp Bot. 57(12):3109-3122.

Hu R, Fan C, Li H, Zhang Y and Fu Y (2009) Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. BMC Molec Biol. 10:93.

Huis R, Hawkins S and Neutelings G (2010) Selection of reference genes for quantitative gene expression normalization in flax (*Linum usitatissimum* L.). BMC Plant Biol. 10:71.

Iskandar HM, Simpson RS, Casu RE, Bonnett GD, MacLean DJ and Manners JM (2004) Comparison of reference genes for quantitative realtime polymerase chain reaction analysis of gene expression in sugarcane. Plant Mol Biol Rep. 22(1):325-337.

Jain M, Nijhawan A, Tyagi AK and Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. Biochem Bioph Res Co. 345(2):646-651.

Jian B, Liu B, Bi Y, Hou W, Wu C and Han T (2008) Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol Biol. 9(59)

Kim BR, Nam HY, Kim SU, Kim SI and Chang YJ (2003) Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. Biotechnol Lett 25(21):1869-1872.

Kolkman JM, Slabaugh MB, Bruniard JM, Berry S, Bushman BS, Olungu C, Maes N, Abratti G, Zambelli A, Miller JF, Leon A and Knapp SJ (2004) Acetohydroxyacid synthase mutations conferring resistance to imidazolinone or sulfonylurea herbicides in sunflower. Theor Appl Genet. 109(6):1147-59.

Libault M, Thibivilliers S, Bilgin D, Radwan O, Benitez M, Clough S and Stacey G (2008) Identification of four soybean reference genes for gene expression normalization. Plant Genome-US. 1(1):44–54.

Miller A (1987) Sunflower. In: Principles of cultivar development. Ed.:Fehr, W. Macmillan Publishing Company:626-668.

Moschen S, Bengoa Luoni S, Paniego N, Hopp E, Dosio G, Fernandez P and Heinz R (2014) Identification of candidate genes associated with leaf senescence in cultivated sunflower (*Helianthus annuus* L.) PLOS ONE. 9(8):e104379.

Narsai R, Ivanova A, Ng S and Whelan J (2010) Defining reference genes in *Oryza sativa* using organ, development, biotic and abiotic transcriptome datasets. BMC Plant Biol 10:56.

Nicot N, Hausman JF, Hoffmann L and Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato duringbiotic and abiotic stress. J Exp Bot. 56(421):2907-2914.

Ochogavia AC, Breccia G, Vega T, Felitti SA, Picardi LA and Nestares G (2014) Acetohydroxyacid synthase activity and transcripts profiling reveal tissue-specific regulation of ahas genes in sunflower. Plant Sci. 224:144-50.

Pfaffl MW, Tichopad A, Prgomet C and Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. Biotechnol Lett. 26(6):509-15.

Radonic A, Thulke S, Mackay IM, Landt O, Siegert W and Nitsche A (2004) Guideline to reference gene selection for quantitative realtime PCR. Biochem Bioph Res Co 313(4):856-862.

Sala CA, Bulos M, Altieri E and Ramos ML (2012) Genetics and breeding of herbicide tolerance in sunflower. Helia. 35(57):57-70.

Schneiter AA andMiller JF (1981) Description of sunflower growth stages. Crop Sci. 21(1):901-903.

Selvey S, Thompson EW, Matthaei K, Lea RA, Irving MG and Griffiths LR (2001) Beta-actin an unsuitable internal control for RT-PCR. Mol Cell Probe. 15(5):307-311.

Thorrez L, Van Deun K, Tranchevent LC, Van Lommel L, Engelen K, Marchal K, Moreau Y, Van Mechelen I and Schuit F (2008) Using ribosomal protein genes as reference: a tale of caution. PLOS ONE. 3(3):e1854.

Ungerer M andKawakami T (2013) Transcriptional dynamics of LTR retrotransposons in early generation and ancient sunflower hybrids. Genome Biol Evol. 5(2):329–337.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3(7):research0034.1.

Wan H, Zhao Z, Qian C, Sui Y, Malik AA and Chen J (2010) Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. Anal Biochem. 399(1):257–261.

Wei L, Miao H, Zhao R, Han X, Zhang T and Zhang H (2013) Identification and testing of reference genes for sesame gene expression analysis by quantitative real-time PCR. Planta. 237(1):873-889.

Zhu J, Li W, Yang W, Qi L and Han SQ (2013) Identification of microRNAs in *Caragana intermedia* by high throughput sequencing and expression analysis of 12 microRNA s and their targets under salt stress. Plant Cell Rep. 32(9):1339–1349.