Iran. J. Biotechnol. April 2023;21(2): e3357 DOI: 10.30498/ijb.2023.338375.3357

Research Article

Evaluating and Validating Sunflower Reference Genes for Q-PCR Studies Under High Temperature Condition

Masood Soltani Najafabadi, Nazanin Amirbakhtiar

National Plant Genebank, Seed and Plant Improvement Institute, Agricultural Research, Education, and Extension Organization, Karaj, Iran

*Corresponding author: Masood Soltani Najafabadi, National Plant Genebank, Seed and Plant Improvement Institute, Agricultural Research, Education, and Extension Organization, Karaj, Iran. Tel: +98-2632701260, Fax: +98-2632709405, E-mail: soltanin@areeo.ac.ir

Received: 2022/04/19; **Accepted:** 2023/01/ 07

Background: Q-PCR is the method of choice for PCR- based transcriptomics and validating microarray-based and RNA-seq results. Proper application of this technology requires proper normalization to correct as much as possible errors propagating during RNA extraction and cDNA synthesis

Objectives: The investigation was performed to find stable reference genes in sunflower under shifting in ambient temperature.

Materials and Methods: Sequences of five well-known reference genes of Arabidopsis (*Actin*, *Ubiquitin*, *Elongation factor-1*, *GAPDH*, and *SAND*) and one well-known reference gene in human, *Importin*, were subjected to BLASTX against sunflower databases and the relevant genes were subjected to primer designing for q-PCR. Two sunflower inbred lines were cultivated at two dates so that anthesis occurred at nearly 30 °C and 40 °C (heat stress). The experiment was repeated for two years. Q-PCR was run on samples taken for two planting date separately at the beginning of anthesis for each genotype from leaf, taproots, receptacle base, immature and mature disc flowers and on pooled samples comprising of the tissues for each genotype, planting dates and also all tissues for both genotypes and both planting dates. Basic statistical properties of each candidate gene across all the samples were calculated. Furthermore, gene expression stability analysis was done for six candidate reference genes on Cq mean of two years using three independent algorithms, geNorm, Bestkeeper, and Refinder.

Results: Designed primers for *Actin2*, *SAND*, *GAPDH*, *Ubiquitin*, *EF-1a*, and *Importin* yielded a single peak in melting curve analysis indicating specificity of the PCR reaction. Basic statistical analysis showed that *Actin2* and *EF-1a* had the highest and lowest expression levels across all the samples, respectively. *Actin2* appeared to be the most stable reference gene across all the samples based on the three used algorithms. Pairwise variation analysis revealed that for samples taken under ambient temperature of 30 °C, *Actin2*, *EF-1a*, *SAND* and for those taken under ambient temperature of 40 °C, *Actin2*, *EF-1a*, *Importin* and *SAND* have to be used for normalization in q-PCR studies. Moreover, it is suggested that normalization to be based on *Actin2*, *SAND* and *EF-1a* for vegetative tissues and *Actin2*, *EF-1a*, *SAND* and *Importin* for reproductive tissues.

Conclusions: In the present research, proper reference genes for normalization of gene expression studies under heat stress conditions were introduced. Moreover, the presence of genotype-by- planting date interaction effects and tissue specific gene expression pattern on the behavior of the most three stable reference genes was indicated

Keywords: Heat stress, Reference genes, Sunflower

1. Background

Climate change significantly affects crop physiology, especially by elevating ambient temperature (1). Indeed, it has caused a violation in phenological-developmental stages; for instance, flowering time faces warmer temperatures compared with before. The molecular bases of these changes in plant biology need to be elucidated through molecular tools such as q-PCR.

Gene expression analysis is the backbone of transcriptomics and the related sciences. Starting from Northern blotting (2) gene expression analysis methods subjected to evolutions towards high throughput platforms, such as RNASeq (3). In this evolution of technology and concept, quantitative Polymerase Chain Reaction (q-PCR) has been introduced as the most sensitive, accurate, and reliable methods (4) to validate data obtained by other methods (5).

Factors such as variation in the amount of input total RNA, the efficiency of complementary DNA (cDNA) synthesis, and (messenger RNA) mRNA degradation severely influence gene expression measurements via q-PCR. Thus, selecting and using appropriate reference genes are crucial steps in gene expression analysis (6).

Reference genes have traditionally been selected from constitutively and stably expressed genes. To meet these criteria, genes involved in basic cellular functions and structures (7) have been evaluated in many plants. For example, stable reference genes have been identified and evaluated in gene expression analysis in response to abiotic and biotic stresses in *Hordeum brevisubulatum* (8), cowpea (9), and wheat (10). Moreover, stably expressed reference genes at various developmental stages have been evaluated and reported in different plants such as soybean (11), *Brassica juncea* (12), and sunflower (13). Ochogavía et al. (13) evaluated stability of five commonly used reference genes *Actin*, *Elongation factor1*, *Plastid-encode RNA polymerase*, *Tubulin*, and *Ubiquitin* in addition to five new candidates in different tissues at various developmental stages.

Sunflower is a staple oil seed crop with kernel oil content of about 55%. In terms of oil production, sunflower is ranked as the fourth most important oily crop in the world (14). The crop shows wide adaptability to various climate conditions, attracting much attentions for breeding with the aim of yield improvement (15).

Under high ambient temperature, and exceeding a threshold (defined as the temperature above which growth and reproduction of a species is negatively affected (16), plants experience heat stress. The intensity of the effect of environmental stresses on the plants depends on the stress severity, plant developmental stages, and duration of stress conditions (17). In sunflower, when heat stress occurs during anthesis, pollen and ovule sterility may occur (18, 19); thus, anthesis is the sensitive stage influenced by heat stress (20). There is variation in the threshold temperature among genotypes (21) and different growth stages (20). It has been indicated that temperatures above 27 °C can result in dramatic problems in the growth rate and embryo survival of sunflowers (19). Based on the previous studies, high temperature affects leaf growth. For example, Rawson and Hindmarsh (22) reported that temperatures above 36 °C decreased leaf growth period by 1.04 days $°C^{-1}$ in sunflowers. Moreover, heat stress can accelerate leaf senescence in many plant species (23).

Selection and validation of stable reference genes under heat stress situation has been performed in *Hordeum brevisubulatum* (8), *Diospyros kaki* (24), and *Cajanus cajan* (25). Sinha et al. (25) evaluated the expression stability of 10 commonly used housekeeping genes (*EF1α, UBQ10, GAPDH, 18Sr RNA, 25Sr RNA, TUB6, ACT1, IF4α, UBC, and HSP90*) in root, stem, and leaves tissues of *Cajanus cajan* under heat stress. They found *UBC, HSP90*, and *GAPDH* were the most stable reference genes under heat stress. In wild barley *Hordeum brevisubulatum*, 11 candidate reference genes, including *Actin*, *ADP-ribosylation factor 1*, *Cyclophilin 2*, *Elongation factor 1-a*, *Glyceraldehyde 3-phosphate dehydrogenase*, *Heat shock protein 90*, *Alpha-tubulin*, *Beta-tubulin 6*, *Ubiquitin*, *18SrRNA-1* and *18SrRNA-3* were subjected to the expression stabilities analysis in shoots and roots under various stress conditions including heat stress. There are a few studies on the expression stability of candidate reference genes in sunflower (13, 26), but no study was conducted on candidate genes stably expressed under different ambient temperatures.

2. Objectives

Investigation on the gene expression under high ambient temperature is demanding especially under climate change conditions. As far as the authors know, no reference genes have been introduced in sunflower in high temperatures. The aim of this study was to find the most stable reference genes for gene expression normalization

in q-PCR studies under heat stress in sunflowers.

3. Material and Methods

The methodology, annotations, and terms used in this paper are in accordance with the MIQE guidelines (27).

3.1. Plant Genetic Materials and Treatments

Sunflower (Helianthus annuus L.) inbred lines of B-line19 and Bline-1221, developed at the Oil Seed Crop Research Department, Seed and Plant Improvement Institute (SPII), Karaj, Iran, were used in this investigation. Seeds of the inbred lines were planted 25 cm apart on 4 m length rows, with row spacing of 60 cm. The experiment was conducted in two years, 2017 and 2018, in a complete randomized block design with three replicates. Inbred line (in two levels) and planting date (in two levels) were considered as factors. Planting dates were chosen such that the onset of anthesis occurs when the ambient temperature were 32±2 °C and 40 ±1 °C. The first planting date was in mid-April and a 10-day interval was used for the next three planting dates. NPK fertilizers were supplied based on soil analysis. Weeds were controlled manually. Plots were irrigated every 14 days before anthesis.

Sampling was performed at the beginning of anthesis in two replicates (out of three replicates). The leaf sample consisted of leaf the blade and its stalk. Roots were taken out from the wet soil by shovel based on the method described by Trachsel *et al*. (28). Sampling was also performed from a similar region of the receptacle base in all plants. For flower sampling, all immature and mature disk flowers were collected by a razor. The specimen was flash frozen in liquid nitrogen and kept under -80 °C for further analysis.

3.2. Selection of Candidate Reference Genes and Design of Specific Primers

Sequences of five well-known reference genes of *Arabidopsis* (*Actin*, *Ubiquitin*, *Elongation factor-1*, *GAPDH*, and *SAND*) (29), and one well-known reference gene in humans, *Importin* (30) were subjected to BLASTX against *Helianthus annuus* databases in National Center for Biotechnology Information (NCBI) and the orthologous genes were identified in sunflower (**Table 1**). Primers of sunflower candidate genes were designed using Oligo 7 (31). The criteria for primer designing were as follows (32):

An amplicon size of 100-150 base pairs (bp), annealing temperature of 60 ± 2 °C, GC content of 40% to 60% and primer length of 18-24 bp. All primers were synthesized by SinaClon Company (SinaClon, Iran).

3.3. Gene Expression Experiment

Total RNA was extracted from 100mg of the frozen samples using TransZol Up Plus RNA Extraction kit (Beijing, China) according to the manufacturer's instruction. Quantity and quality of the RNA was evaluated by Nanodrop (ND-1000 Spectrophotometer, Thermo Scientific, Wilmington DC, USA) and agarose gel electrophoresis. After treating the total RNA with RNase-free DNaseI (Roche, Mannheim, Germany), q-PCR was done using *Actin2* primer pairs (**Table 1**) to

Table 1. Description of the primers used in this work.

Gene name	Genebank ID	Forward (5' to 3')	Reverse (5' to 3')	Amplicon length	Orthologue in *Arabidopsis*
GAPDG	110920787	GATGACATGGTGAAGGTGATTG	AATTAAATGACTGATGACTGAAGC	128	*AT1G13440
Importin	KAF5807756	GCAACAACAATGAAGGATATAGAG	AAGTGAAGAGCAACCATTAACC	100	**AT5G19820
Actin2	110909803	AACTATTATGTAAGACTGGCAGAC	TCACAACCACTCTCCAACTAC	130	*AT3G18780
SAND	110904297	GGCAGGAACCCCTTTCAC	TGGGCTGGGCTTTGTTTG	150	*AT2G28390
EF-1a	110868526	GTTTTGTTATTTGATCCTTGAAGC	TTAAACCACCAACTAGCAACTC	110	*AT5G60390
Ubiquitin	110910892	GCTTGTCTTATGTTTGAACTCTTG	TTGAAACCTGAATGAACTTGGG	141	*AT4G05320
Pectin acetylesterase	CD850746§	AGTTTGGCAGTGGGAATAGCAA	TGGATTCGGAAACTCTCCTTCA	210	-
Ferredoxin	CD849228§	GCAGTCGTGCTCAGACTTTCC	GACCATGATGGCCACCTACAA	175	-

* reported by (28) Wang *et al*., 2014, ** extracted from www. Arabidopsis.org, § reported by Hewezi et al., 2008.

be sure of the lack of genomic DNA. Integrity of RNA was checked on a 1% agarose gel prior to, and after *DNase*I treatment.

First strand cDNA synthesis was carried out using 2µg of the *DNase* I-treated RNA using SuperScript III reverse transcriptase kit (Invitrogen, Karlsruhe, Germany) according to the manufacturer's instruction.

Q-PCR reactions were conducted using a BioRad MJ MiniTM Thermal Cycler (California, USA). Reactions with a final volume of 20µL contained 4µL of template (cDNA or total RNA), 200mM of each primer (1µL of mixed 0.5mM forward and reverse primers, **Table 1**), and 10µL of SYBR Green Realq -plus 2X master mix (Ampliqon). Thermal profile used was 95 ºC for 10 min; 40 cycles of 95 °C for 15s, and 60 °C for 1 min. After 40 cycles, the specificity of the amplifications was checked by heating from 60 °C to 95 °C with a ramp speed of 1.9 °C.min^{-1}, and producing the melting curves.

Two independent biological replicates were used. Reactions with no template were also included as a negative control. The quantification cycle (Cq) values used for analysis, was calculated by averaging the two biological replicates and the two years. Specificity of amplification was verified for the lack of primer dimers or non-specific amplicons through melting curve analysis.

To validate the selected reference genes for q-PCR normalization, the relative expression of two heat responsive genes in sunflower (**Table 1**) leaves was measured in the second planting date (as heat stress condition) compared to the first planting date (as normal ambient temperature). The normalization was done using the most and least stable and also the two most stable genes according to the $2^{-\Delta\Delta Ct}$ formula (33).

3.4. Pooling Samples and Cq's

For each year and each biological replicate, 23 samples were subjected to q-PCR analysis (**Table 2**). For pooled samples, equal amounts of cDNA from the relevant samples were mixed and then subjected to q-PCR analysis. Finally, for 10 cases, the Cq of samples was pooled, and the pooled data were used for stability analysis (**Table 2**). Pooled samples comprised of the tissues for each genotype, planting dates and also all tissues for both genotypes and both planting dates.

3.5. Evaluating Candidate Reference Genes Expression and Stability

Mean, median, coefficient of variation (CV) and Box-plot of Cq values in different tissues, planting dates and pooled samples were considered basic statistical properties of each candidate gene and calculated by Mini tab v18.

Gene expression stability was evaluated via three independent algorithms. The geNorm algorithm (34) was run using geNorm v3 software. To do this, first the minimum Cq in all the candidate reference genes and samples was used as a base for calculating relative expression according to the geNorm manual. The relative expressions were then used as input for geNorm software. For the rest of the two algorithms, Cq values were directly used as input. A gene stability analysis by the NormFinder (35) was performed in the NormFinder Excel program. Bestkeeper (36) was used as the third gene expression stability algorithm.

4. Results

4.1. Identification of Sunflower Candidate Reference Genes

Among the investigated genes, four were protein-coding genes (*Actin2*, *EF-1a*, *GAPDH*, and *Ubiquitin*) frequently used in expression analysis and recognized as good candidates based on previously published studies (26, 32). Moreover, one newly used protein-coding reference gene in planta (*SAND)* (37) and one reference gene in animals (*Importin*) (30), which have been reported as highly stable reference genes in various cases and conditions, were also evaluated. These genes were subjected to q-PCR and the melting temperature analysis, subsequently. For all the investigated genes, a single pick was observed in their melting curve (**Fig. S1**) indicating of specificity of the primer's pairs and lack of primer self- and cross- dimers.

4.2. Evaluating the Expression Levels of Reference Candidate Genes

The Cq values were used to provide an overview of the expression levels of the candidate genes for all tissues, planting dates and pooled samples. The mean Cq values of the reference genes, averaged over the two years, ranged from 16.9 to 23.8 for different samples as shown in **Figure 1**. The *Actin2* was found to be the most expressed gene with the lowest mean Cq value (16.9±0.23) in leaf followed by the same gene (Cq=17.3± 1.1) in root, while *EF-1a* was expressed at the lowest level (23.8± 0.39) in flower tissues.

Table 2. Composition of samples and data subjected to q-PCR and stability analysis.

Item for analysis	Composition of items		Type of analysis
Individual samples			
	B-line 19/flower/Planting date 1		Q-PCR
	B-line 19/flower/Planting date 2		Q-PCR
	B-line1221/flower/Planting date 1		Q-PCR
	B-line1221/flower/Planting date 2		Q-PCR
	B-line1221/Receptacle base/Planting date 1		Q-PCR
	B-line1221/Receptacle base/Planting date 2		Q-PCR
	B-line 19/Receptacle base/Planting date 1		Q-PCR
	B-line 19/Receptacle base/Planting date 2		Q-PCR
	B-line1221/Leaf/Planting date 1		Q-PCR
	B-line1221/Leaf/Planting date 2		Q-PCR
	B-line 19/Leaf/Planting date 1		Q-PCR
	B-line 19/Leaf/Planting date 2		Q-PCR
	B-line 19/Root/Planting date 1		Q-PCR
	B-line 19/Root/Planting date 2		Q-PCR
	B-line1221/Root/Planting date 1		Q-PCR
	B-line1221/Root/Planting date 2		Q-PCR
Pooled samples			
	B-line 19 at Planting date 1 composed of all tissues		Q-PCR
	B-line 19 at Planting date 2 composed of all tissues		Q-PCR
	B-line1221 at Planting date 1 composed of all tissues		Q-PCR
	B-line1221 at Planting date2 composed of all tissues		Q-PCR
	First planting date composed of all tissues for both genotypes		Q-PCR
	Second planting date composed of all tissues for both genotypes		Q-PCR
	All samples composed of all tissues for both genotypes and both planting dates		Q-PCR
Pooled Cq			
	Reproductive tissues	two genotypes + two planting dates	Stability analysis
	Vegetative tissues	two genotypes + two planting dates	Stability analysis
	B-line 19	two planting dates + all tissues	Stability analysis
	B-line1221	two planting dates + all tissues	Stability analysis
	Flower	two planting dates for both genotypes	Stability analysis
	Receptacle base	two planting dates for both genotypes	Stability analysis
	Root	two planting dates for both genotypes	Stability analysis
	Leaf	two planting dates for both genotypes	Stability analysis
	Reproductive tissues	flower and receptacle base of two planting dates for both genotypes	Stability analysis
	Vegetative tissues	leaf and root two planting dates for both genotypes	Stability analysis

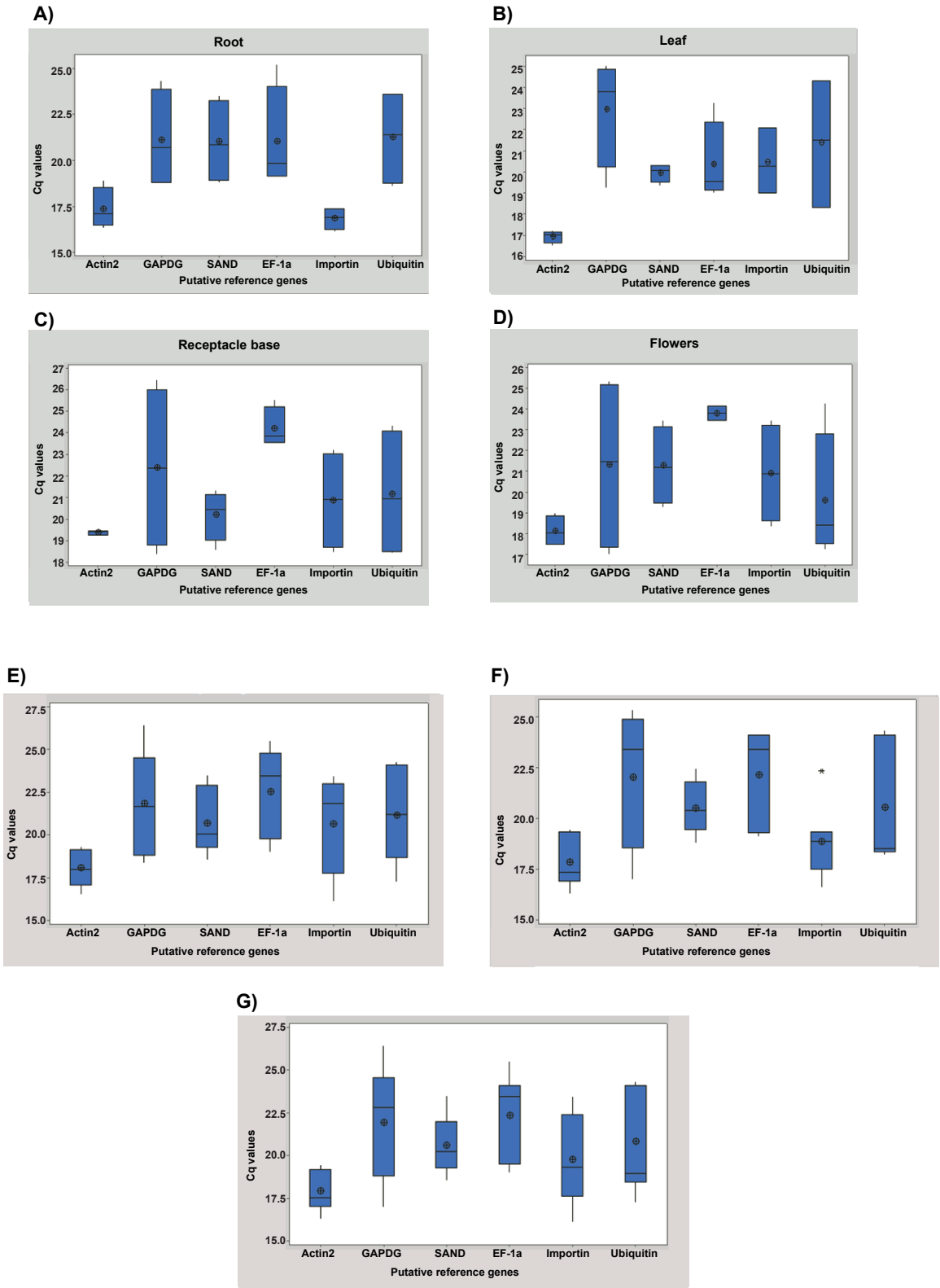


Figure 1. **Expression levels of six candidate reference genes in different samples**. Pooled sample is comprised of all tissues, planting dates of all the genotypes. Lines crossing the boxes depict the medians. Crossed circles indicate means. Stars indicate outlier data; Boxes indicate the interquartile range. Whiskers represent 95% confidence intervals.

The other putative reference genes used in the investigation were moderately expressed, with mean Cq values ranging from 17.8 to 23.8. Moreover, *Actin2* showed the least coefficient of variation (CV) of 0.69% in its transcript level across all samples and *EF-1a* showed the second least variation in the gene expression with a CV of 1.6%. *GAPDG* was the most variable reference gene (CV=20.0%) in flower tissues (**Fig. 1**). These results clearly indicate that the expression of internal control genes varied across the tissues and validating reference genes for use in q-PCR normalization is therefore necessary.

4.3. Stability of the Putative Reference Genes

Three famous algorithms used for evaluating gene expression stability, *i e*. geNorm, NormFinder, and Bestkeeper were employed on the data sets (**Table 2**) and their results are presented here.

4.4. GeNorm Analysis

The geNorm algorithm was applied to the expression data of all the candidate genes gathered on genotypes, planting dates, and sunflower tissues and organs. For both planting dates, flower, receptacle base, reproductive tissues (pool of Cq values of flower and receptacle base), maximum stability index, M, were observed for *Actin2* and *EF-1a* (**Fig. 2D-2G** and **Table S1**). Moreover, *Actin2*, followed by *SAND* and *EF-1a*, were the most stable candidates in vegetative tissues (pool of Cq values of leaf and root) (**Table S1**). Stability analysis for the two genotypes led to different results in such a way that *Actin2* and *EF-1a* were identified as the most stable reference genes in B-line 19, while *SAND*, followed by *Actin2*, was recognized as the most stable one in BF-line 1221 (**Fig. 2H-2I**). Finally, when all the samples were pooled into a single body, *Actin2* and *EF-1a* were commonly identified as the most stable candidates (**Table S1** and **Fig. 2A**).

4.5. The Effect of Planting Date ×Genotype Interaction on Behavior Stability of Candidates

According to the geNorm algorithm results (**Table S1** and **Fig. 2**), the most stable reference gene in the four pooled samples, planting dates and genotypes was *Actin2*. Nevertheless, when the geNorm stability index for the second most stable and the most unstable reference genes were graphed for the mentioned samples, shifts in their behaviors were observed (**Fig. 3**). For the second most stable genes, the stability index of the first planting date in the Bline-1221 genotype was lower than that for the second planting date. Nevertheless, their ranking order was shifted in Bline-19. The same type of shits was observed for the most unstable reference gene. The shift in position of different levels of one factor (*i e.* planting dates) across the levels of other factor is called interaction effect (38).

4.6. Determining Optimum Number of Reference Genes

It has been advised to use the geometric mean of expression of at least three reference genes for gene expression normalization (34). In order to find the minimum number of reference genes, we calculated pairwise variations ($Vn/Vn+1$) by the geNorm software. Vandesompele *et al*. (34) suggested a cut of value of Vn/Vn+1<0.15 means that the addition of more reference gene would have no significant contribution to normalization in q-PCR data analysis. As shown in **Figure 4**, a reduction in the variation was observed from V2/3 to V3/4 for B-line1221, the first planting date, root, and leaf tissues. The reduction was observed from V3/4 to V4/5 for the rest of the samples and the all samples pooled. These indicate that for data normalization in the first planting date, root, and leaf, three reference genes and for reproductive-related tissues and second planting date four reference genes are sufficient for gene expression normalization in the sunflower.

4.7. Norm Finder Analysis

The Norm Finder analysis was used as another algorithm to evaluate the gene expression stability of the candidates. The analysis indicated that *Actin2* was the most stable candidate gene for seven out of 11 samples. However, *EF-1a* was identified as the most stable candidate for B-line19, receptacle base, and reproductive tissues. *Importin* was identified as the most stable one for root (**Table S1**).

4.8. Best Keeper Analysis

The results obtained for stability analysis using the Best Keeper algorithm revealed that *Actin2* had the most expression stability in B-line1221, planting date 1, receptacle base, and leaf and vegetative tissues. *EF-1a* were the most stable candidate for B-line19, flower tissue and reproductive tissues. Moreover, *SAND* and *Importin* identified to be the stable candidate for planting date 2 and root, respectively (**Table S1**).

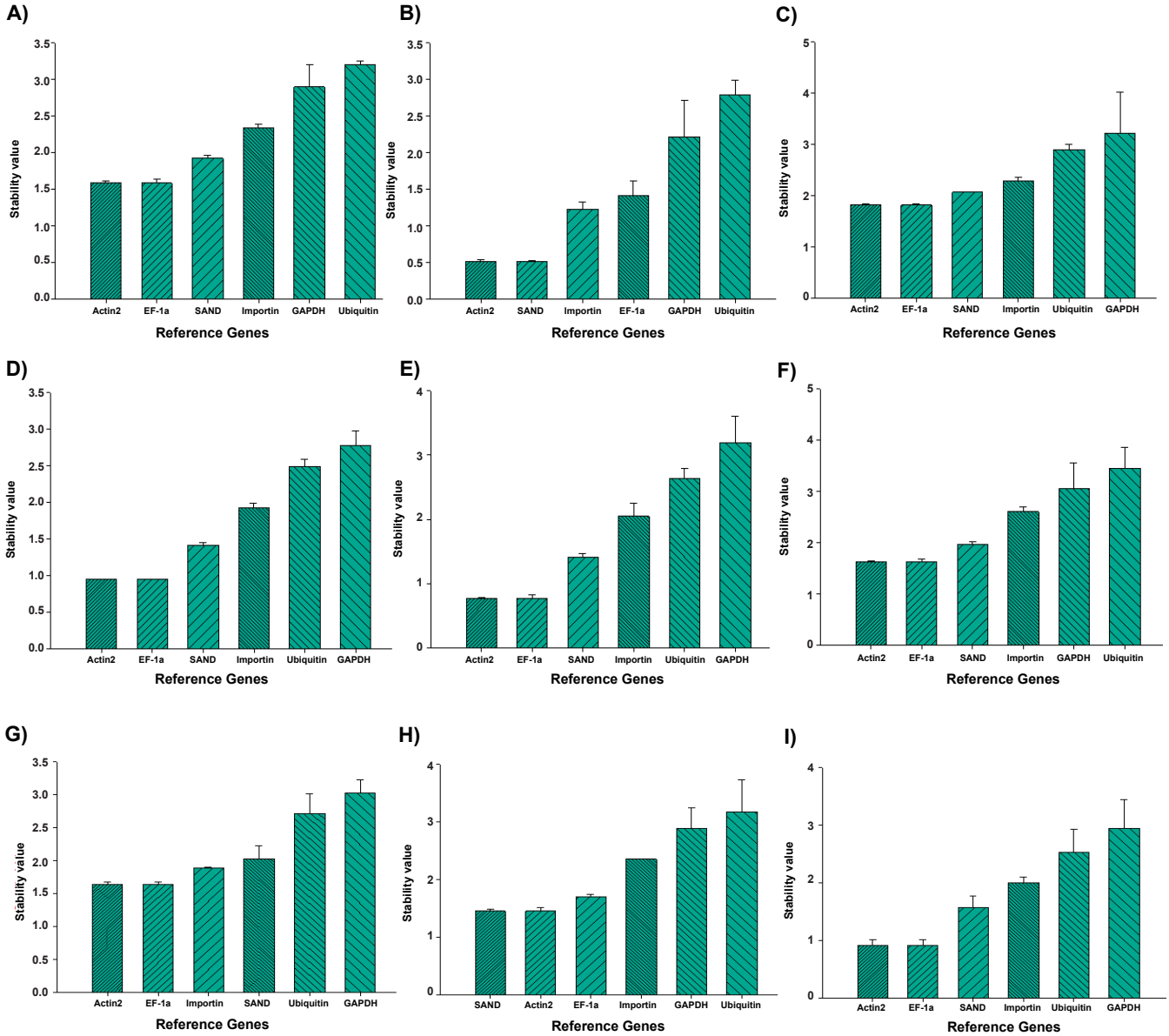


Figure 2. Stability analysis using GeNorm algorithm. A)The analysis was performed for pooled samples, two vegetative tissues: **B**) leaf and **C**) root, two reproductive tissues: **D**) receptacle base and **E**) flower, two planting dates: **F**) date 1 and **G**) date 2, two sunflower inbred lines: **H**) B-line 1221 and **I**) B-line 19. Vertical axis shows average expression stability values (**M**) and the genes are arranged on the horizontal axis in each panel from the most stable (left side) to the least stable (right side) ones.

4.9. Validation of the Selected Reference Genes

To validate the effectiveness of the selected reference genes, the expression patterns of two heat responsive genes, *CD850746* and *CD849228* were analyzed. The genes have been reported to be upregulated upon heat stress in the sunflower leaves (39). Q-PCR analysis of leaves samples under heat stress showed that the relative expressions were 4.8 and 2.85 for *CD850746* and *CD849228*, respectively (39). In the present investigation, the expression patterns of the two genes were calculated in heat stress condition (ambient temperature around 40 °C) relative to the normal temperature (ambient temperature around 30 °C).

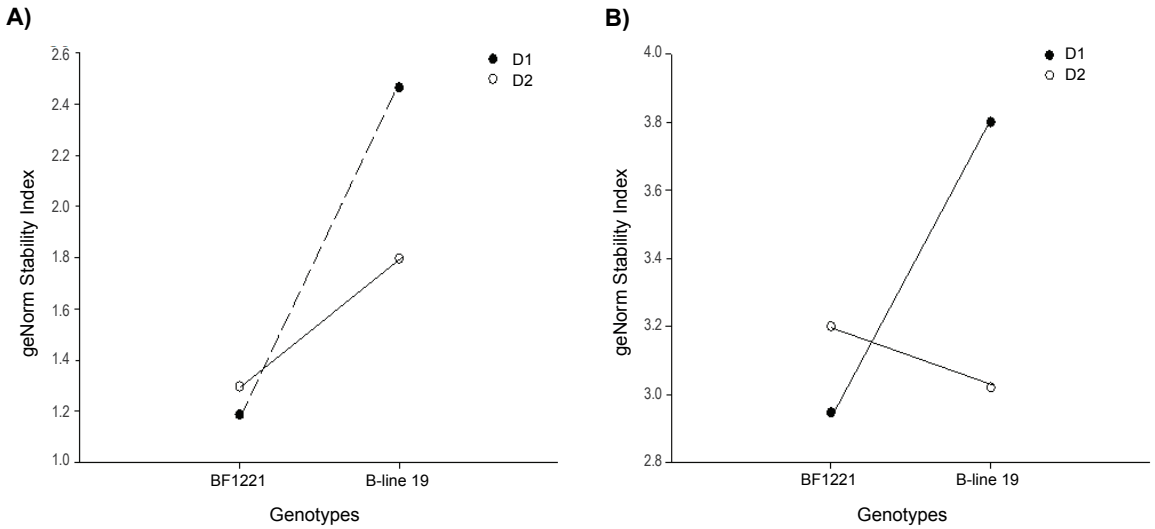
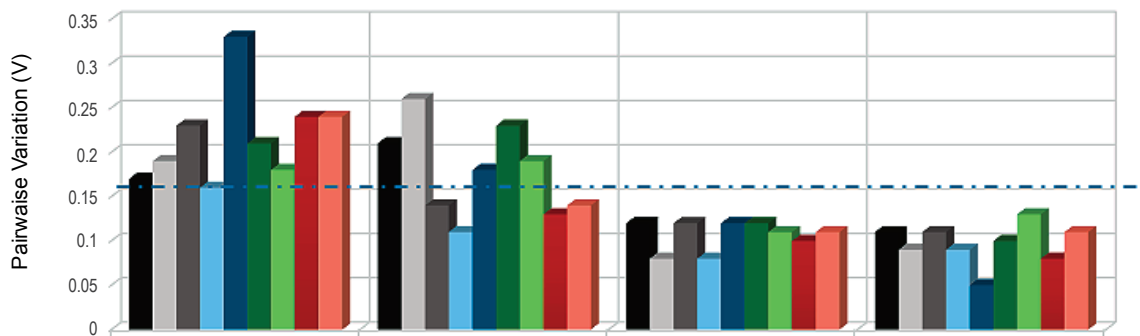


Figure 3. **The effect of planting date ×genotype interaction on behavior stability of six candidate reference genes according to geNorm algorithm.** The geNorm stability index of the second most stable **A**) and the most unstable **B**) reference genes were plotted for BF1221 and B-line19. Crossing of the line connected planting dates indicate interactions between the genotypes and the planting dates.



	V2/3	V3/4	V4/5	V5/6
All data	0.17	0.21	0.12	0.11
B-Line 19	0.19	0.26	0.08	0.09
B-Line1221	0.23	0.14	0.12	0.11
D1	0.16	0.11	0.08	0.09
D2	0.33	0.18	0.12	0.05
Flower	0.21	0.23	0.12	0.1
Receptacle base	0.18	0.19	0.11	0.13
Leaf	0.24	0.13	0.1	0.08
Root	0.24	0.14	0.11	0.11

Figure 4. **Pairwise variation *(Vn/Vn*+1*)* was analyzed to find optimal number of reference genes for accurate gene expression normalization for q-PCR data.** The analysis was performed on 6 candidate reference genes by geNorm software. Dashed line shows the cut of value so that Vn/Vn+1<0.15 means that the addition of more reference gene would have no significant contribution to normalization in q-PCR data analysis. D1 and D2 stands for the first and second planting date, respectively. All data refers to pooling all tissues from the two genotypes and two planting dates into a unit sample.

The transcript abundances of these genes were normalized based on the most stable, the most two stables and the most unstable reference genes under normal temperature (*Actin2*, *Actin2* + *EF-1a*, and *Ubiquitin*, respectively) and those under heat stress condition (*Actin2*, *Actin2* + *EF-1a*, and *GAPDG*). Fluctuation in the relative transcript abundance is indicated by the standard error bars (**Fig. 5**). The relative expression analysis for both heat responsive genes showed high standard error bars when the most unstable reference genes was used. The fluctuation in the relative gene expression was reduced when the most stable reference genes was used as normalizer. Interestingly, for both tested heat stress responsive genes, when the two most stable reference genes used for normalization, the least fluctuations in the relative gene expression were observed (**Fig. 5**).

5. Discussion

The normalization procedure based on reference genes is a crucial step in transcriptomics-related experiments (40-42). Genes proposed to be as references should meet at least two criteria: having moderate to high expression level (43) and stability in expression over developmental stages (44) and environmental conditions (25).

Although it is simple to report and use a single unique gene for performing expression analysis in a species, the normalization results are far from reality. Usually, the reference genes are chosen from housekeeping genes such as *Actin*, *Ubiquitin*, tubulin, *importin*, *GAPDH*, *Elongation factor (EF) subunits*, and *Ribosomal* genes, involved in basic cellular functions and structures. These genes are supposed to be constitutively and stably expressed in varying physiological and experimental conditions (7). Nevertheless, environmental conditions (25) and developmental stages (44) usually shift the housekeeping expression patterns. This necessitates performing precise expression analysis under various environmental conditions and developmental stages for each species to find the most stable housekeeping genes to be used as a reference in gene normalization.

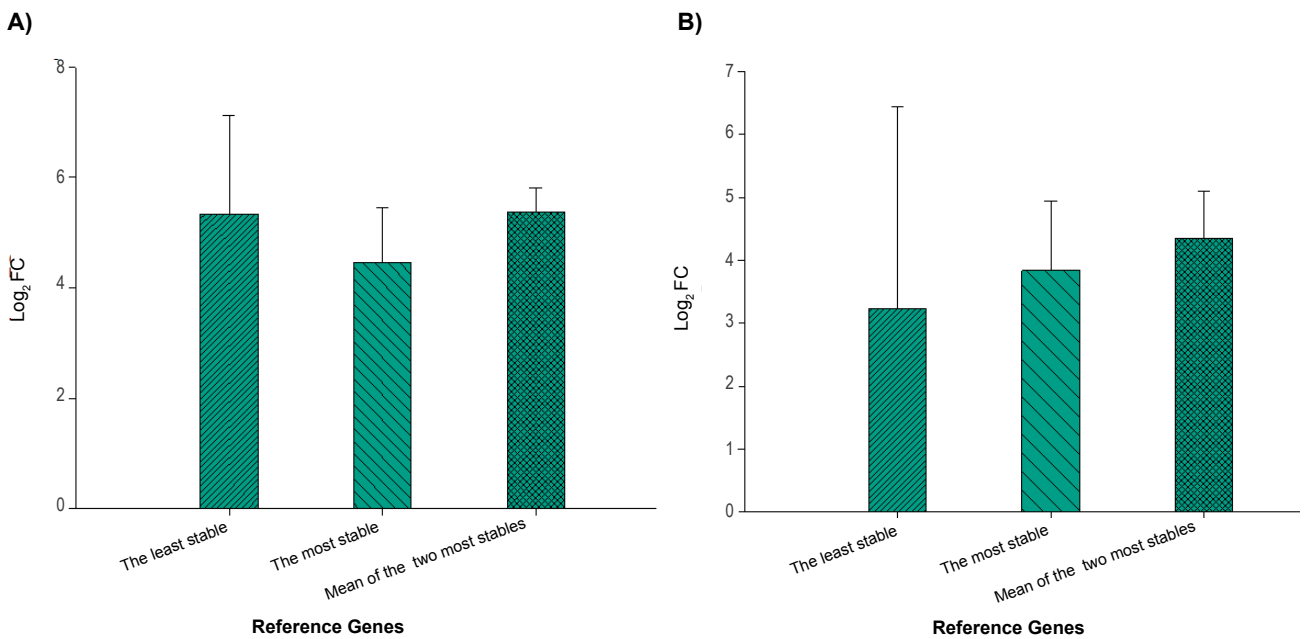


Figure 5. Validation of selected reference genes under heat stress condition. Relative expression of **A**) *CD850746* and **B**) *CD849228* in the sunflower leave at heat stress condition compared to normal ambient temperature was calculated after normalization based on the most stable, the two most stable and the most unstable reference genes. Under normal temperature *Actin2*, *Actin2* + *EF-1a*, and *Ubiquitin* were the most stable, the two most stable and the most unstable reference genes, respectively. Under heat stress condition, *Actin2*, *Actin2*+ *EF-1a*, and *GAPDG* were the most stable, the two most stable and the most unstable reference genes, respectively. Relative gene expression analysis was performed according to the ……Bars indicate the standard errors (n=3).

Factors such as variation in the amount of input total RNA, the efficiency of cDNA synthesis, and mRNA degradation significantly affect gene expression measurement via q-PCR. Selecting and using appropriate reference genes have been proposed to reduce theses effects during gene expression analysis in q-PCR (45). In the present paper, the expression stability of six candidate reference genes over different organs and environmental conditions was investigated in the sunflower.

The experiments were performed under field conditions and repeated for two years. Thus, our results seem to be closer to the real situation compared with most of the experiments performed under controlled conditions (**Table S2**). The results obtained under field conditions increase the levels of complexity for selecting appropriate reference genes (46); nevertheless, repeating the experiments in different years attenuates environmentally inevitable variations and makes the results more confident (38).

To the best of our knowledge, all the expression analysis experiments performed on sunflower vegetation and reproductive tissues under various developmental stages or environmental conditions have employed almost one reference gene, *Actin*, for normalization (**Table S2**). The stability of some reference genes was reported to be affected by experimental conditions (47). In expression analysis platforms (such as array-based methods (32) and RNASeq analysis studies), q-PCR on a subset of differentially expressed genes (DEGs) is usually used to confirm the expression analysis results (**Table S2**).

Whole plant expression analysis (done by pooling all plant tissues and organs and measuring the expression of many genes) may be used in the near future as a tool in biological systems (48). To this end, having reference genes with general high stability in expression will be highly demanded. In our experiment, we mixed all the samples from both two genotypes, planting dates, and their relevant tissues and investigated the expression stability of the six candidate reference genes. All the three methods used for gene expression stability indicated *Actin2* as the most stable candidate. Ochogavía *et al*. (13) evaluated the stability of 10 new and conventional reference genes and found that *Actin*, an unknown protein, and *EF-1a* were the most stable genes in all the samples. Nevertheless, when the validating the expression of many genes through q-PCR is the mater of interest (as proposed by Vandesompele *et al*. (34), normalization based on more than one reference gene is demanding. We recommend that geometric mean of Cq of four reference genes of *Actin2*, *EF-1a*, *SAND*, and *Importin*, as the four most stable candidates, be used to normalize gene expression in mixed samples.

In the present experiment, we mimicked the temperature changes by managing planting dates. The plants cultivated at the first and second planting dates were exposed to ambient temperatures of 32 ± 2 °C and 40 ± 1 °C at the beginning of anthesis, respectively. The effect of ambient temperatures on the expression stability of candidate genes is shown in **Figure 2**. Based on the geNorm algorithm, *Actin2*, *EF-1a*, and *SAND* at 32±2 °C and *Actin2*, *EF-1a*, and *Importin* at 40±1 °C were identified as the three most stable candidates (**Table S1**). Thus, based on the recommendation of Vandesompele *et al*. (34), it is proposed to use the geometric mean of *Actin2*, *EF-1a*, and *SAND* Cq's for the expression normalization in normal temperature (32±2 °C) and *Actin2*, *EF-1a*, and *Importin* for normalization in high temperature (40±1 °C) at the beginning of anthesis. However, as *Actin2* was nominated by the three algorithms as the most stable reference gene, this gene is offered as an expression normalizer when a limited number of genes are subjected to expression analysis at various ambient temperatures. *Actin* and *EF1* have been frequently employed as reference genes in sunflower (49-52).

In most of the investigations performed on reference genes identifications, usually one genotype (53, 54) or pools of different genotypes (43) were subjected to various treatments or sampled in various developmental stages (55, 56). This has created ambiguity in reporting a set of reference genes for the investigated materials. We observed presence of genotype × planting date interactions on the second most stable and the most unstable reference genes. These observations elucidate presence of the effect of different genetic backgrounds on the behavior stability of reference genes. As the ranking of the three most stable candidates changed over the two genotypes (**Table S1**), for any number of genes under study, it is highly recommended to utilize the geometric mean of *Actin2*, *EF-1a*, and *SAND* Cq's for normalization when different genotypes are under investigations.

Tissue-specific gene expression was observed for the reference genes. While the most three stable reference genes for flower tissues were *Actin2*, *EF-1a*, and *SAND*,

those for leaf tissues were *Actin2*, *SAND*, and *Importin*. It has been reported in sunflower that while *ETIF5* was one of the most stable genes in vegetative tissues, it exhibited the worst stability among reproductive tissues (13).

Ochogavía *et al*. (13) evaluated the stability in the expression of 10 reference genes in vegetative and reproductive tissues of the sunflower and found that the three and five most stable reference genes should be used for normalizations for sunflower vegetative and reproductive tissues, respectively. However, in the present investigation, we suggest that the three and four most stable reference genes to be employed for normalization. The genes are *Actin2*, *EF-1a*, *Importin*, and *SAND* for reproductive tissues, and *Actin2*, *SAND*, and *EF-1a* for vegetative tissues.

Suitability of the evaluated reference genes was assessed through expression analysis of two heat stress-responsive genes. The relative expression analysis of these two genes in the sunflower leaves in planting date 2 compared to that in planting date 1 were normalized according to the most stable, the two most stable and the most unstable reference genes. We observed that the relative transcript abundance of both genes showed high vitiation when normalization was done based on the most unstable reference genes. The variation in the relative gene expression patterns was reduced when stable reference genes used for the gene expression normalization. These results indicates that the identified reference genes are suitable for transcript normalization in the sunflower under the heat stress condition.

We used three different algorithms for evaluating the gene expression stabilities. These algorithms use different mathematical and statistical models. Usually, these different algorithms differently rank the reference genes according to their analytical principles. As there are no definite criteria to select the best algorithm, the three algorithms have been utilized in most all the researches performed in the field of reference gene stability analysis. Nevertheless, because of its dual function in the gene expression stability analysis and determining the optimal number of reference genes, researchers usually refer to the geNorm results when discrepancies are observed among the algorithms results.

6. Conclusions

The expression stability of six candidate reference genes was surveyed in two sunflower genotypes, two planting dates, and vegetative and reproductive tissues using three different algorithms. For the first time, *Importin* was evaluated as a reference candidate in planta and proposed as one of the three most stable reference genes for heat stress experiments (*i e*. planting date 2) and leaf tissue. Although *Actin2* was identified as the most stable gene in all samples based on at least one algorithm, for reproductive tissues, geometric mean of *Actin2*, *EF-1a*, *SAND*, and *Importin*, and for vegetative tissues, that of *Actin2*, *SAND*, and *EF-1a* is proposed for normalization. At planting date 1 and 2, respectively, sets of three genes of *Actin2*, *SAND*, and *EF-1a* and four genes of *Actin2*, *EF-1a*, *Importin*, and *SAND* are advised to be used to calculate normalization factors. Genotype × planting date interaction and tissue specific gene expression was observed for the reference genes.

In the present research, proper reference genes for normalization of gene expression studies under heat stress conditions were introduced. Moreover, the presence of genotype-by-planting date interaction effects and tissue specific gene expression pattern on the behavior of the most three stable reference genes was indicated.

Acknowledgement

This research was conducted in Seed and Plant Improvement Institute (SPII). The research was financially supported by Iran National Science Foundation Grant No. 92017948 which is highly appreciated.

References

1. Martins MQ, Fortunato AS, Rodrigues WP, Partelli FL, Campostrini E, Lidon FC, *et al*. Selection and validation of reference genes for accurate RT-qPCR data normalization in Coffea spp. under a climate changes context of interacting elevated (CO_2) and temperature. *Front Plant Sci*. 2017;**8**:307. doi: 10.3389/fpls.2017.00307
2. Goldsworthy S, Goldsworthy T, Sprankle C, Butterworth B. Variation in expression of genes used for normalization of Northern blots after induction of cell proliferation. *Cell Prolif*. 1993;**26**(6):511-517. doi: 10.1111/j.1365-2184.1993.tb00029.x
3. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;**10**(1):57-63. doi: 10.1038/nrg2484
4. Gachon C, Mingam A, Charrier B. Real-time PCR: what relevance to plant studies? *J Exp Bot*. 2004;**55**(402):1445-1454. doi: 10.1093/jxb/erh181
5. Wu AR, Neff NF, Kalisky T, Dalerba P, Treutlein B, Rothenberg ME, *et al*. Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods*. 2014;**11**(1):41-46. doi: 10.1038/nmeth.2694
6. Bustin S, Benes V, Nolan T, Pfaffl M. Quantitative real-time

RT-PCR–a perspective. *J Mol Endocrinol.* 2005;**34**(3):597-601. doi: 10.1677/jme.1.01755
7. Panina Y, Germond A, Masui S, Watanabe TM. Validation of common housekeeping genes as reference for qPCR gene expression analysis during iPS reprogramming process. *Sci Rep*. 2018;**8**(1):1-8. doi: 10.1038/s41598-018-26707-8
8. Zhang L, Zhang Q, Jiang Y, Li Y, Zhang H, Li R. Reference genes identification for normalization of qPCR under multiple stresses in Hordeum brevisubulatum. *Plant Methods*. 2018;**14**(1):1-14. doi: 10.1186/s13007-018-0379-3
9. Amorim LLB, Ferreira-Neto JRC, Bezerra-Neto JP, Pandolfi V, de Araújo FT, da Silva Matos MK, *et al*. Cowpea and abiotic stresses: identification of reference genes for transcriptional profiling by qPCR. *Plant Methods*. 2018;**14**(1):1-17. doi: 10.1186/s13007-018-0354-z
10. Paolacci AR, Tanzarella OA, Porceddu E, Ciaffi M. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Mol Biol*. 2009;**10**(1):1-27. doi: 10.1186/1471-2199-10-11
11. Hu R, Fan C, Li H, Zhang Q, Fu Y-F. Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Mol Biol*. 2009;**10**(1):1-12. doi: 10.1186/1471-2199-10-93
12. Chandna R, Augustine R, Bisht NC. Evaluation of candidate reference genes for gene expression normalization in Brassica juncea using real time quantitative RT-PCR. *PloS One*. 2012;**7**(5):e36918. doi: 10.1371/journal.pone.0036918
13. Ochogavía AC, Novello MA, Picardi LA, Nestares GM. Identification of suitable reference genes by quantitative real-time PCR for gene expression normalization in sunflower. *Plant Omics*. 2017;**10**(4). doi: 10.21475/poj.10.04.17.pne831
14. Grompone MA. Sunflower and high-oleic sunflower oils. *Bailey's industrial oil and fat products*. 2005:1-54. doi: 10.1002/047167849X.bio017.pub2
15. Vear F. Changes in sunflower breeding over the last fifty years. *OCL - Oilseeds Fats Crops Lipids.* 2016;**23**(2):1-8. doi : 10.1051/ocl/2016006
16. Kalyar T, Rauf S, Teixeira Da Silva JA, Shahzad M. Handling sunflower (*Helianthus annuus* L.) populations under heat stress. *Arch. Acker Pflanzenbau Bodenkd.* 2014;**60**(5):655-672. doi: 10.1080/03650340.2013.799276
17. Chaves MM, Pereira JS, Maroco J, Rodrigues ML, Ricardo CP, Osório ML, *et al*. How plants cope with water stress in the field? Photosynthesis and growth. *Ann Bot.* 2002;**89**(7):907-916. doi: 10.1093/aob/mcf105
18. Chimenti C, Hall A. Grain number responses to temperature during floret differentiation in sunflower. *Field Crops Res*. 2001;**72**(3):177-184. doi: 10.1016/S0378-4290(01)00175-7
19. Chimenti C, Hall A, Lopez MS. Embryo-growth rate and duration in sunflower as affected by temperature. *Field Crops Res.* 2001;**69**(1):81-88. doi: 10.1016/S0378-4290(00)00135-0
20. Moriondo M, Giannakopoulos C, Bindi M. Climate change impact assessment: the role of climate extremes in crop yield simulation. *Clim Change*. 2011;**104**(3):679-701. doi: 10.1007/s10584-010-9871-0
21. Kalyar T, Rauf S, Teixeira da Silva J, Haidar S, Iqbal Z. Utilization of leaf temperature for the selection of leaf gas-exchange traits to induce heat resistance in sunflower (*Helianthus annuus* L.). *Photosynthetica*. 2013;**51**(3):419-428. doi: 10.1007/s11099-013-0038-x
22. Rawson H, Hindmarsh J. Effects of temperature on leaf expansion in sunflower. *Funct Plant Biol.* 1982;**9**(2):209-219. doi: 10.1071/PP9820209
23. Fahad S, Bajwa AA, Nazir U, Anjum SA, Farooq A, Zohaib A, *et al*. Crop production under drought and heat stress: plant responses and management options. *Front Plant Sci.* 2017;**1147**. doi: 10.3389/fpls.2017.01147
24. Wang P, Xiong A, Gao Z, Yu X, Li M, Hou Y, *et al*. Selection of suitable reference genes for RT-qPCR normalization under abiotic stresses and hormone stimulation in Persimmon (*Diospyros kaki* Thunb). *PloS One*. 2016;**11**(8):e0160885. doi: 10.1371/journal.pone.0160885
25. Sinha P, Saxena RK, Singh VK, Krishnamurthy L, Varshney RK. Selection and validation of housekeeping genes as reference for gene expression studies in pigeonpea (*Cajanus cajan*) under heat and salt stress conditions. *Frontiers in Plant Sci*. 2015;**6**:1071. doi: 10.1371/journal.pone.0122847
26. Fernandez P, Di Rienzo JA, Moschen S, Dosio GA, Aguirrezábal LA, Hopp HE, *et al*. Comparison of predictive methods and biological validation for qPCR reference genes in sunflower leaf senescence transcript analysis. *Plant Cell Rep*. 2011;**30**(1):63-74. doi: 10.1007/s00299-010-0944-3
27. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, *et al*. The MIQE guidelines: Minimum information for publication of quantitative real-rime PCR experiments. *Oxford University Press*; 2009. doi: 10.1373/clinchem.2008.112797
28. Trachsel S, Kaeppler SM, Brown KM, Lynch JP. Shovelomics: high throughput phenotyping of maize (*Zea mays* L.) root architecture in the field. *Plant Soil*. 2011;**341**(1):75-87. doi: 10.1007/s11104-010-0623-8
29. Wang H, Wang J, Jiang J, Chen S, Guan Z, Liao Y, *et al*. Reference genes for normalizing transcription in diploid and tetraploid Arabidopsis. *Sci Rep.* 2014;**4**(1):1-8. doi: 10.1038/srep06781
30. Ayakannu T, Taylor AH, Konje JC. Selection of endogenous control reference genes for studies on type 1 or type 2 endometrial cancer. *Sci Rep.* 2020;**10**(1):1-10. doi: 10.1186/s13326-017-0146-9
31. Rychlik W. OLIGO 7 primer analysis software. In: Yuryev A. (eds) PCR Primer Design. *Methods Mol Biol***.** 2007;**402**:35-59. doi: 10.1007/978-1-59745-528-2_2
32. Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible W-R. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol.* 2005;**139**(1):5-17. doi: 10.1104/pp.105.063743
33. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*. 2001;**25**(4):402-408. doi: 10.1006/meth.2001.1262
34. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, *et al*. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol***.** 2002;**3**(7):1-12. doi: 10.1186/gb-2002-3-7-research0034
35. Andersen CL, Jensen JL, Ørntoft TF. 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.* 2004;**64**(15):5245-5250. doi: 10.1158/0008-5472.CAN-04-0496.
36. Pfaff MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulatedtarget genes and sample integrity: BestKeeper -Excel-based tool

usingpair-wise correlations. *Biotechnology Letter*. 2004;**26**:509-515. doi: 10.1023/B:BILE.0000019559.84305.47

37. Zhao D, Wang X, Chen J, Huang Z, Huo H, Jiang C, *et al*. Selection of reference genes for qPCR normalization in buffalobur (S*olanum rostratum* Dunal). *Sci Rep.* 2019;**9**(1):1-10. doi: 10.1038/s41598-019-43438-6
38. Rohlf FJ, Sokal RR. Biometry: the principles and practice of statistics in biological research. *J Am Stat Assoc.* 1982;**77** (380):946-947. doi: 10.2307/2287349
39. Hewezi T, Ger ML, Gentzbittel L. A comprehensive analysis of the combined effects of high light and high temperature stresses on genee expression in sunflower. *Ann Bot.*2008;**102**:127-140. doi: 10.1093/aob/mcn071
40. Bhuva DD, Cursons J, Davis MJ. Stable gene expression for normalisation and single-sample scoring. *Nucleic Acids Res Spec Publ.* 2020;**48**(19):e113-e. doi: 10.1093/nar/gkaa802
41. Abbas-Aghababazadeh F, Li Q, Fridley BL. Comparison of normalization approaches for gene expression studies completed with high-throughput sequencing. *PloS One*. 2018;**13**(10):e0206312. doi: 10.1371/journal.pone.0206312
42. Liu X, Li N, Liu S, Wang J, Zhang N, Zheng X, *et al*. Normalization methods for the analysis of unbalanced transcriptome data: a review. *Front Bioeng Biotechnol*. 2019;**7**:358. doi: 10.3389/fbioe.2019.00358
43. Caldana C, Scheible W-R, Mueller-Roeber B, Ruzicic S. A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. *Plant Methods*. 2007;**3**(1):1-9. doi: 10.1186/1746-4811-3-7
44. Fischer M, Skowron M, Berthold F. Reliable transcript quantification by real-time reverse transcriptase-polymerase chain reaction in primary neuroblastoma using normalization to averaged expression levels of the control genes *HPRT1* and *SDHA*. *J Mol Diagn*. 2005;**7**(1):89-96. doi:10.1016/S1525-1578(10)60013-X
45. Kozera B, Rapacz M. Reference genes in real-time PCR. *J Appl Genet*. 2013;**54**(4):391-406. doi: 10.1007/s13353-013-0173-x
46. Jameson PE. Selection of reference genes for flowering pathway analysis in the masting plants, *Celmisia lyallii* and *Chionochloa pallens*, under variable environmental conditions. *Sci Rep.* 2019;**9**(1):1-16. doi: 10.1038/s41598-019-45780-1
47. Xu Z, Dai J, Su W, Wu H, Shah K, Xing L, *et al*. Selection and validation of reliable reference genes for gene expression studies in different genotypes and TRV-infected fruits of peach (*Prunus persica* L. Batsch) during Ripening. *Genes*. 2022;**13**(1):160. doi: 10.3390/genes13010160
48. Libault M, Pingault L, Zogli P, Schiefelbein J. Plant systems biology at the single-cell level. *Trends Plant Sci.* 2017;**22**(11):949-960. doi: 10.1016/j.tplants.2017.08.006
49. Breccia G, Vega T, Felitti SA, Picardi L, Nestares G. Differential expression of acetohydroxyacid synthase genes in sunflower plantlets and its response to imazapyr herbicide. *Plant Sci*. 2013;**208**:28-33. doi: 10.1016/j.plantsci.2013.03.008
50. Moschen S, Bengoa Luoni S, Paniego NB, Hopp HE, Dosio GA, Fernandez P, *et al*. Identification of candidate genes associated with leaf senescence in cultivated sunflower (*Helianthus annuus* L.). *PloS One*. 2014;**9**(8):e104379. doi: 10.1371/journal.pone.0104379
51. Ochogavía AC, Breccia G, Vega T, Felitti SA, Picardi LA, Nestares G. Acetohydroxyacid synthase activity and transcripts profiling reveal tissue-specific regulation of ahas genes in sunflower. *Plant Sci*. 2014;**224**:144-150. doi:10.1016/j.plantsci.2014.04.018
52. Ungerer MC, Kawakami T. Transcriptional dynamics of LTR retrotransposons in early generation and ancient sunflower hybrids. *Genome Biol Evol*. 2013;**5**(2):329-337. doi:10.1093/gbe/evt006
53. Guo S, Zuo Y, Zhang Y, Wu C, Su W, Jin W, *et al*. Large-scale transcriptome comparison of sunflower genes responsive to *Verticillium dahliae*. *BMC Genom*. 2017;**18**(1):1-13. doi: 10.1186/s12864-016-3386-7
54. Moschen S, Marino J, Nicosia S, Higgins J, Alseekh S, Astigueta F, *et al*. Exploring gene networks in two sunflower lines with contrasting leaf senescence phenotype using a system biology approach. *BMC Plant Biol*. 2019;**19**(1):1-15. doi:10.1186/s12870-019-2021-6
55. Hewezi T, Léger M, El Kayal W, Gentzbittel L. Transcriptional profiling of sunflower plants growing under low temperatures reveals an extensive down-regulation of gene expression associated with chilling sensitivity. *J Exp Bot.* 2006;**57**(12):3109-3122. doi: 10.1093/jxb/erl080
56. Lim AR, Kong Q, Singh SK, Guo L, Yuan L, Ma W. Sunflower *WRINKLED1* plays a key role in transcriptional regulation of oil biosynthesis. *Int J Mol Sci.* 2022;**23**(6):3054. doi: 10.3390/ijms23063054