Identification of dominant housekeeping genes to normalize mRNA level in dioecious Palmyra palm

Suchilkumar Soorya[1]*, Thacheril Sukumaran Swapna[1], Senan Siju[2] and Sabu Kalluvettankuzhy Krishnan Nair[2]
1. Department of Botany, University of Kerala, Kariavattom, Thiruvananthapuram 695581, INDIA
2. Biotechnology and Bioinformatics Division, Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Palode, Thiruvananthapuram 695562, INDIA
*sooryas326@gmail.com

Abstract

A stable housekeeping gene is required for precise and consistent qPCR data analysis. Palmyra palm is an important tropical dioecious tree crop with diverse uses. However, no reliable gene expression studies or housekeeping gene standardization have been documented, which limit crop improvement and utilization. In this work, four frequently used statistical tools such as RefFinder, NormFinder, Delta CT method and BestKeeper were applied to evaluate the constancy and fluctuation in expression of a chosen 10 internal control genes such as NADH dehydrogenase subunit 5-like (NAD5), β-actin, ACTIN CRD, alpha-tubulin, UBCE, EIF, EF2, AP1, GAPDH and UBQ2 in both vegetative and reproductive tissues of dioecious Palmyra palm. Six flower tissues and two leaf tissues from male and female palms were used in this study.

The findings indicated that NAD5-2 is the most reliable and constant housekeeping gene expressed during Palmyra palm flower development. The AGAMOUS and SEEDSTICK genes expression pattern in the flower development were used to validate the normalized expression of the four housekeeping genes that were chosen in order to further support the stability of those genes. Our results will be useful for further research on gene transcription level in Palmyra palms and also the above mentioned tool is very useful for the normalization of housekeeping genes of biological samples with different experimental conditions.

Keywords: *Borassus flabellifer*, gene expression, reference gene, qRT PCR, standardization.

Introduction

Borassus flabellifer L. commonly known as toddy palm or Palmyra palm comes under the family arecaceae. Palmyra is a crop that has been found in India since ancient times. It can be grown with very low investment and yields continuously for 100 years[14]. Palmyra is an asset for the next generations but farmers are hesitant to grow Palmyra palm due to its dioecious nature and they switch to different crops and cut down any trees that are still there in the fields. Palmyra is a symbol of the traditional occupation of the country. The State Government of Tamil Nadu declared it as a State tree. It is known as karpaha, "nungu" "celestial tree" in Tamil culture and is greatly revered since all of its components may be utilized[24]. Palmyra was a boon to nature due to its climate. This is because they can withstand any weather, heat or drought climatic changes and get more yield. As a result, very few researches can be done to identify the essential genes and to clarify the molecular mechanisms associated with yield for the crop improvement of Palmyra palm.

Gene expression analysis in the palm is of interest. However, there is currently limited molecular information known about the Palmyra palm. Only the chloroplast genome sequence and transcripts of male and female inflorescences have been reported recently and the whole genome sequence is not yet available[31,33]. Transcriptome analysis has led to the discovery of numerous important genes involved in the network underlying the genetically complex traits of the Palmyra palm[31]. For the proper utilization of this crop, additional research on the expression of specific genes in various tissues of interest and at various phases of tissue development would be required. In recent years, one of the most used techniques for determining changes in gene expression is quantitative real-time polymerase chain reaction (qRT-PCR)[37].

Compared to more established techniques like northern blotting and ribonuclease protection test, it is increasingly used for messenger RNA (mRNA) identification and quantification[18]. The qRT-PCR result was influenced by various experimental variables which include the quality and amount of template, specificity of primer and efficiency of reverse transcription PCR and pipetting errors[8]. Several methods have been devised to reduce the technical inaccuracy associated with qRT-PCR, the most accurate of which is normalization of the housekeeping genes.

There are a lot of studies on the selection of stable housekeeping genes in various experimental conditions in plants such as oil palm[44], date palm[27], coconut palm[27], rice[13], wheat[26], celery[20], amaranthus[39], arabidopsis[16] etc. Due to their involvement in essential cellular processes such as glycolytic pathway, cytoskeleton components, protein folding, ribosome subunit synthesis, protein degradation and so on, housekeeping genes are frequently used as candidate reference genes, regardless of physiological or biological conditions[2].

An ideal housekeeping gene should exhibit constant expression across tissues and experimental setups. However, no universal stable housekeeping gene in plants or animals has yet been discovered[5]. The most commonly used

housekeeping genes are GAPDH, NAD5, TUBULIN, UBIQUITIN, ACTIN, polyubiquitin (UBQ), translation elongation factor (TEF), 18S (18S rRNA) and 25S ribosomal RNA (25S rRNA), cyclophylin(CYP), elongation factor1-A (EF1A) etc.[10] Thus, before using candidate housekeeping genes in RT-qPCR normalization, it is crucial to evaluate their stability for each experimental situation. Meanwhile, various user-friendly web-based programs have been developed such as geNorm[49], NormFinder[49], BestKeeper[28] and RefFinder[42] to validate the most stable housekeeping gene from a list of possible genes.

The standardization of reference genes in Palmyra palm for qPCR quantification of gene expression studies has not been published. However, detailed research and data analysis of appropriate reference gene expression during different tissues of male and female Palmyra palm leaves and flower have yet to be identified. The goal of this study was to determine the most stable housekeeping genes for precise qRT-PCR standardization of vegetative and reproductive organs during Palmyra palm flower development. Ten potential housekeeping genes and their expression stability were thoroughly evaluated using four different statistical techniques.

Material and Methods

Plant materials: Mature leaves and young and mature inflorescence of male and female Palmyra palm were taken from parassala for RNA isolation as in fig. 1. These samples were taken from 20-year-old Palmyra palms. It is very difficult to collect different stages of male and female inflorescences for this study. Because the primordial of the inflorescence is covered in several thick and very stiff leaf sheets, we cannot distinguish leaf primordial or an inflorescence primordial. After the tip is removed, the plant typically dies. Another challenge was that we could not climb the tree and see how the inflorescences had developed, so we collected the young inflorescences from the leaf sheet just opened and mature tissues collected from the fully open stages. All these samples were stored at -80°C prior to RNA extraction.

Isolation of total RNA, Quality Checking and cDNA Synthesis: In order to isolate RNA, stored leaves and inflorescence tissues were separated according to our study's requirements. The following tissues were used for RNA isolation and labeled as: FL (female leaf), ML (male leaf), IP (immature petal), MP (mature petal), IO (immature ovule), MO (mature ovule), YMF (young male flower) and MMF (mature male flower). RNA was initially obtained using a modified approach due to the hardness of palm leaf and flower materials and the difficulty of directly extracting high-purity RNA. Total RNA was extracted from about 100mg of different tissue samples using modified protocol of RNasy Plant Mini Kit (Qiagen) and CTAB method. By measuring the absorbance of the RNA at wavelengths of 230, 260 and 280 nm, respectively, the concentration and purity of the RNA were detected. (IMPLEN).

The RNA samples' integrity was checked through 1.2 percent agarose gel electrophoresis. Only RNA samples with 260/280 ratios between 1.8 and 2.1 and higher than 2.0 as well as bands for both the 28S and 18S ribosomal RNAs with a density ratio of about 2:1 were used for further analysis. 1 µg of RNA was reverse-transcribed using the PrimeScript RT-reagent kit II in accordance with the manufacturer's instructions to generate first-strand cDNA (TaKaRa, Japan). The reverse transcriptase reaction was carried out at 37°C for 15 minutes before the first strand of cDNA synthesis started. The reaction mixture was then heated for 5 minutes at 85°C in order to inactivate PrimeScript Reverse Transcriptase. Prior to using each cDNA sample as a template in an RT-qPCR reaction, it was diluted 1:10 with RNase-free water and stored at -20°C.

Figure 1: Palmyra palm and different floral parts. A: Palmyra palm tree, B-C: Male and female inflorescence, D: Mature female flower, E: Mature male flower

Housekeeping gene selection and primer designing: For normalization of housekeeping genes consistently expressed in selected Palmyra palm samples, we chose ten candidate genes such as NAD5-1 (NADH dehydrogenase subunit 5-like 1), NAD5-2 (NADH dehydrogenase subunit 5-like 2), ACTB (Beta- actin), ACTB CRD (Beta actin taken from cardomomum), α-tubulin, UBCE (Ubiquitin conjugating enzyme), EIF (Eukaryotic translation initiation factor), EF2 (Elongation factor 2-alpha), GAPDH (Glyceraldehyde-3-phosphate dehydrogenase) and UBQ2 (Ubiquitin) . The NAD5 primer pairs were adopted from stable housekeeping genes of oil palm[4]. All other primer pairs were designed by using IDT's PrimerQuest incorporates Primer3 software (version 2.2.3)[11].

Primers with melting temperatures of 55–64 °C, primer lengths of 18–24 bp, GC content of 45–55 percent and amplicon lengths of 90–150 bp were designed. To predict primer specificity, the NCBI Basic Local Alignment Search Tool (BLAST) was used to compare primers to the available Palmyra palm transcriptome sequence. The primer hybridization related parameters were determined by using Net Primer software available at https://premierbiosoft.com/ netprimer/netprlaunch/netprlaunch.html.

All the primer pairs were synthesized from a commercial supplier Eurofins Genomics India Pvt., Ltd. All primer pairs were checked by RT-qPCR to ensure for the specificity of the amplicon by the melting-curve after amplification with RT-qPCR analysis prior to the routine gene expression study with RT-qPCR. Only the primer pairings that produced single product and no product in the no template control (NTC) were chosen for further analysis. For a single product and the predicted size, 1.5 percent agarose gel electrophoresis and ethidium bromide staining were used to determine primer specificities.

RT-qPCR analysis PCR: For each reference gene, RT-qPCR was performed in an Applied Biosystems Step One Real Time PCR System using specified qPCR primer sets. In a volume of 20 µl, the reaction mixture contained 10 µl SYBR® Premix Ex TaqTM II (Takara), 2 µl cDNA (after dilution), 1 µl each of the forward and reverse primers (5 M) and 6 µl RNase-free water with two-step amplification conditions of 95 °C for 1 min, 40 cycles of 95 °C for 15 s and 60 °C for 34 s (data collection). Three biological samples and three technical replicates for each sample were used in each experiment.

Evaluation of expression stability of candidate genes: The expression of the selected housekeeping genes was analyzed by NormFinder[6], BestKeeper29 and RefFinder[27,49]. NormFinder is a program that finds the best reference gene from a list of candidate genes. It evaluates potential candidate genes based on their expression stability in a particular sample set and experimental strategy[3]. BestKeeper is an Excel-based application that uses raw Cq values to estimate inter-gene relationships of probable reference gene pairs. All genes are included in the calculation of the BestKeeper index, which is used to rank the best reference genes because the stable reference genes have a significant relationship with it[27].

Subsequently we analyzed the findings of all reference genes by using RefFinder (http://150.216.56.64/ referencegene. php) which is a user-friendly, comprehensive, web-based application that compares and ranks the evaluated candidate reference genes using the primary computational methods that are now accessible (geNorm, NormFinder, BestKeeper and the comparative Ct technique).

Validation of reference gene stability: Recent research has discovered that AGAMOUS (AG) and SEEDSTICK (STK) genes are kind of plant transcriptional regulators play an important role in floral whorl identity, particularly in the formation of stamen and ovule during flower initiation. In order to confirm the potential application of the preferred candidate housekeeping gene, the relative expression profiles of the AG and STK genes were assessed in the aforementioned samples using RT-qPCR and normalization was performed using the two most stable and two least stable candidates as ranked by RefFinder. The selected genes sequences were identified and deposited to NCBI GenBank. The gene specific primer was designed by using this sequence through IDT primer quest tool. The expression data were statistically validated using ANOVA Duncan test.

Results

Primers Specificity: PCR was used to test ten primer pairs prepared for putative housekeeping genes against eight cDNA of leaves and inflorescence from male and female Palmyra palm. A complete single band product could only be produced by nine out of the ten housekeeping genes: ACTB, ACT CRD, TUBULIN, NAD5-1, NAD5-2, EIF, EF2, UBCE, UBQ2 and GAPDH as in fig. 2. Presence of a single band indicates that these primers were successful in amplifying the target regions from the chosen cDNA samples. In the qPCR experiment, only DNA primers that amplified complete single bands from chosen samples were employed.

The melting curves of these selected housekeeping genes generated by qRT-PCR were all single peaks as in fig. 3. The amplification efficiency (E%) and correlation coefficients (R2) of the nine potential reference genes were calculated by plotting the slopes of standard curves using 10-fold serial dilutions of template cDNA. The ten primer pair amplification efficiencies varied from 97.80 to 103.1 percent and their correlation coefficients in between 0.96 and 0.999 as in table 1. This finding revealed that the selected primers were more specific and reliable for effective RT-qPCR reaction.

Candidate reference gene expression profiling in real time PCR: The Ct values derived from the amplification curve were used to calculate the expression levels of putative

reference genes. As a result, there was a broad range in the mean Ct values between 18 and 30 cycles across the chosen samples. All eight samples had strongly expressed transcripts for NAD5 and elf, with mean Ct values ranging from 18 to 24 cycles. Since gene expression levels and Ct values are negatively correlated, the gene NAD5 displayed the highest expression abundance and the lowest mean Ct values. While GAPDH, AP2 and UBQ2 displayed relatively low expression (mean Ct = 27.196 and 35.132), TUBULIN, ACT CRD and EIF also displayed very high expression with mean Ct values of 19.097 and 25.374 respectively. Other genes had expression levels that were average.

Computational evaluation of candidate reference genes expression stability: Four statistical programme were used to analyse the expression stability of potential reference genes including RefFinder[7,34], BestKeeper[16,29] and NormFinder[3,15] as well as Microsoft Excel's delta Ct technique.

Table 1
PCR Primers used for the qRT PCR analysis

Gene	Gene name	Primer sequences		Expected amplicon length(bp)	Temp. ◦C
		5' forward 3'	5' reverse3'		
UBCE	Ubiquitin conjugating enzyme	CCAGTGTTTGGCAATGTTCTC	GAGTCCAGCCCTCCAAATC	100	58
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	GCCACGTATGCCGATATTAAGG	GGACCTGCTGTCACCAATAAAG	105	60
NAD5-1	NADH dehydrogenase subunit 5-like gene-1	CATTTCTGGTTCACACGACTTCAG	AGAGAGTAAAACGACCCGAAATCC	112	60
NAD5-2	NADH dehydrogenase subunit 5-like gene-2	CAAGCAAAGATCATATAGC	ATTCCTTGCGGCAACCAC	105	58
α-TUBULIN	alpha-tubulin	CCTTATGTATCGTGGTGATG	AGTAGGACACCAGTCAACAA	120	60
EIF	Eukaryotic initiation factor	CTTGTACCAGTCCAGGTTGG	GGTTACAACCCAGAGAAGATTCC	110	60
ACTB	Beta-actin	TGGATCCTCCGATCCAGAC	AAGGAGATCACTGCCCTTG		62
EF2	Elongation factor 2-alpha	ACCAAAGGCAAAGAACCTACC	TGCCATCAGAAACTGTGATCC	100	58
ACTB CRD	Beta -actin	TGGATCCTCCGATCCAGAC	AAGGAGTCACTGCCCTTG	115	60
UBQ2	Polyubiquitin	ATCTTCGTCAAGACCCTCA	CGGCGGAATACCTTCCYYATC	123	60

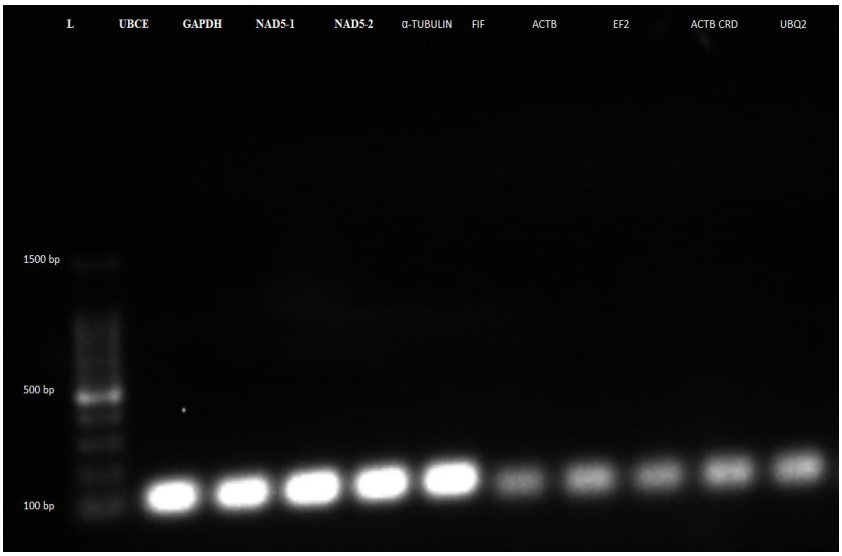


Figure 2: Fragments amplified from the ten housekeeping genes. The specific PCR product the predicted size for each gene is seen on an agarose gel (3%) electrophoresis

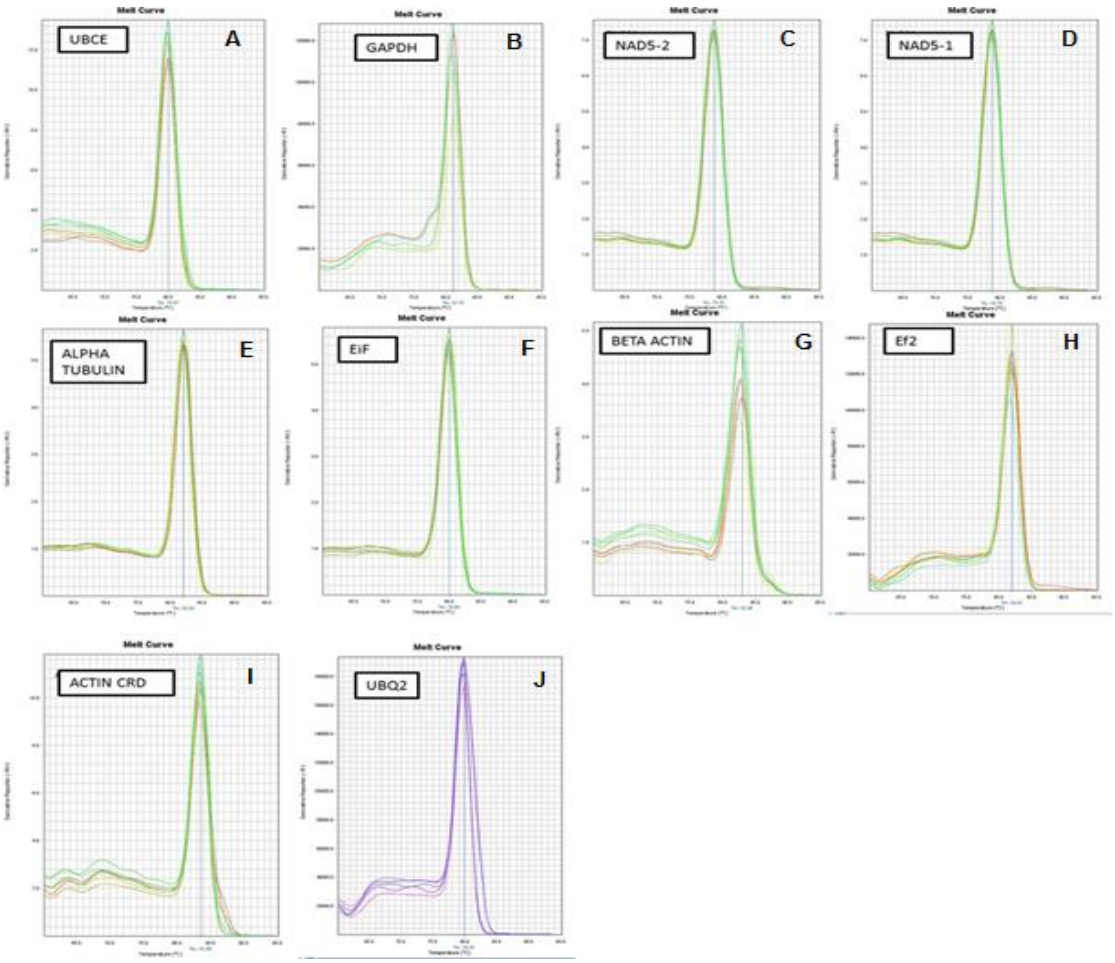


Figure 3: Melt curves analysis generated for amplicons of each of the 10 housekeeping genes.

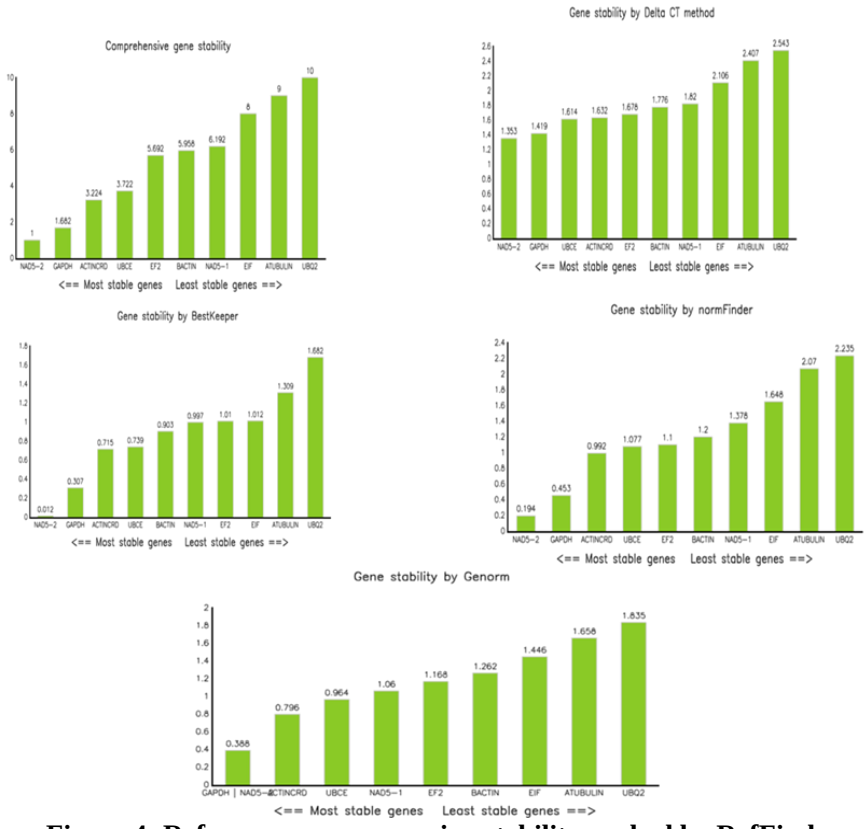


Figure 4: Reference gene expression stability ranked by RefFinder

RefFinder: RefFinder is a thorough and user-friendly web-based tool made for evaluating and screening reference genes from huge experimental datasets. In order to rank and compare the tested potential reference genes, it incorporates the four most important computational tools now in use: geNorm, NormFinder, BestKeeper and the comparative Ct technique http://blooge.cn/RefFinder/. Based on the rankings from each programme, it assigns each gene the proper weight before computing the geometric mean of those weights to arrive at the final overall rating[21]. For each reference gene, the Ct values from all of the samples under various tissues were determined in order to assess the gene stability using RefFinder. According to the RefFinder results, the computer determined that numerous reference genes were among the most stable genes for the samples using a variety of statistical techniques.

The expression stability of selected reference genes followed in the order NAD5-2> GAPDH> ACT CRD> UBCE> EF2> ACTB>NAD5-1>EIF>TUBULIN>UBQ2. Based on the results of the RefFinder study, comprehensive raking recommended that NAD5-2 reference gene expression was constant across all samples of different tissues of Palmyra palm as in table 2. The least stable reference genes are UBQ2 and GAPDH, however GAPDH came second since the CT values were nearly identical across all samples. However, the qPCR procedure employed in this work used 40 cycles of amplification, hence CT values above 30 are regarded to be late amplification and as a result, such a finding was unreliable[19,22].

Quantitative data's precision will vary depending on whether a reference gene is expressed highly or weakly. The ideal reference gene should express itself moderately with a Ct value between 15 and 30[19]. Based on the RefFinder gene stability rating, we can confirm that NAD5 2 is the most stable potential reference gene with a value of 0.194 as in fig. 4.

NormFinder: A method called NormFinder is used to choose the best potential reference gene from a list of candidates. It ranks the group of potential reference genes based on how consistently they express themselves in a specific sample set and under experimental conditions[3]. According to the NormFinder analysis results in this work, the ten candidate genes' expression stability value under the various sample sets followed the order NAD5-2< GAPDH <ACTCRD <UBCE <EF2 <ACTB<NAD51 <EIF <UBQ2 <TUBULIN respectively. Low stability value, according to NormFinder, indicates a gene's high degree of expression stability for the given sample sets. Here are the NormFinder results, which showed that seven genes' stability values were nearly equal to those obtained by the RefFinder tool.

Accordingly, NAD5 2 having the lowest stability value 0.006 is the gene that is most consistently expressed in both the vegetative and reproductive tissues of the Palmyra palm. In addition, UBQ2 and TUBULIN showed highest stability value considered as least stable reference genes with the stability value as in fig. 5.

BestKeeper: Then, we reanalyzed the expression stability of the selected candidate housekeeping genes using another algorithm called BestKeeper. All of the Cq values obtained from the tissue specimens were plotted in an Excel table by the BestKeeper programme, an Excel-based program. According to the pairwise correlation analysis to the BestKeeper Index, the programme estimated the pairwise correlation coefficient for each possible gene and suggested the best reference genes. The arithmetic mean of the Cq values for each examined candidate gene was used to calculate the BestKeeper index.

Table 2
Housekeeping gene expression stability ranked by RefFinder

Ranking Order (Better--Good--Average)										
Method	1	2	3	4	5	6	7	8	9	10
Delta CT	NAD5-2	GAPDH	UBCE	ACTINCRD	EF2	BETA ACTIN	NAD 5-1	EIF	ALPHA TUBULIN	UBQ2
BestKeeper	NAD5-2	GAPDH	ACTINCRD	UBCE	BETA ACTIN	NAD5-1	EF2	EIF	ALPHA TUBULIN	UBQ2
NormFinder	NAD5-2	GAPDH	ACTINCRD	UBCE	EF2	BETA ACTIN	NAD 5-1	EIF	ALPHA TUBULIN	UBQ2
GeNorm	GAPDH | NAD 5-2		ACTINCRD	UBCE	NAD 5-1	EF2	BETA ACTIN	EIF	ALPHA TUBULIN	UBQ2
Recommended comprehensive ranking	**NAD5-2**	**GAPDH**	**ACTINCRD**	**UBCE**	**EF2**	**BETA ACTIN**	**NAD 5-1**	**EIF**	**ALPHA TUBULIN**	**UBQ2**

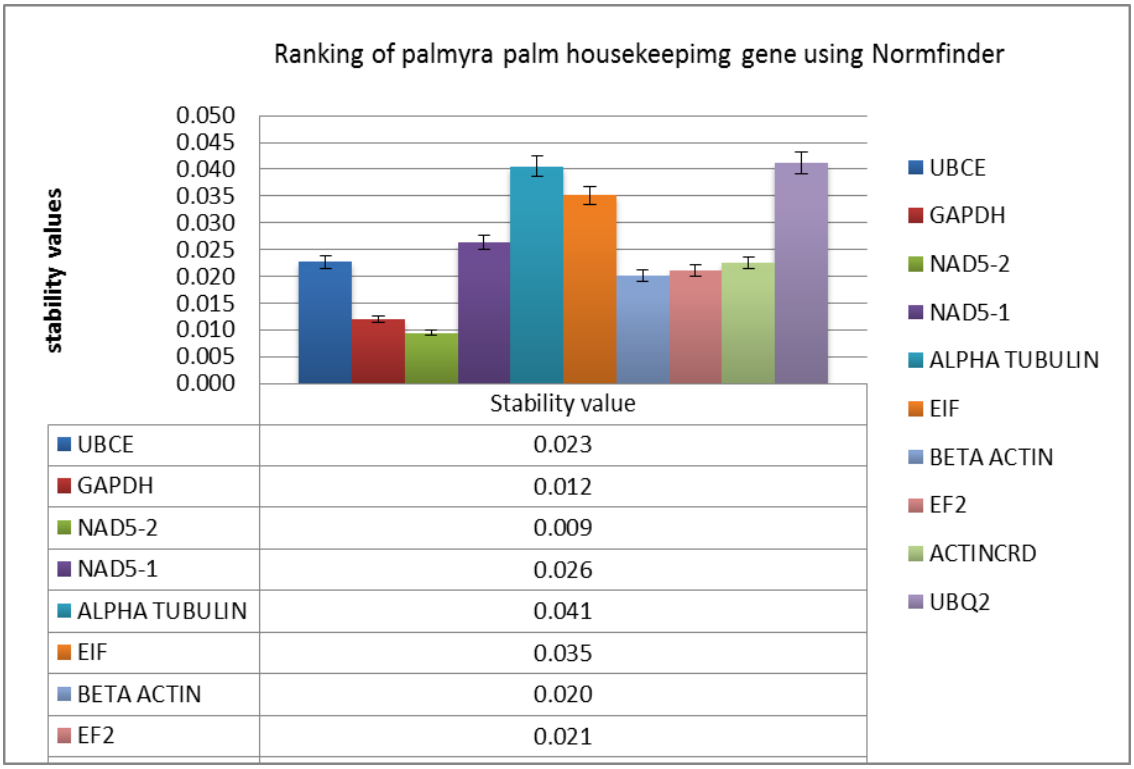


Figure 5: Ranking of stability and transcript accumulation of candidate housekeeping genes. NormFinder ranking of ten candidate reference genes based on quantitative polymerase chain reaction data. According to the NormFinder the lower stability value indicating the higher stability (NAD5-2= 0.009)

Table 3
Housekeeping genes expression stability ranked by BestKeeper at different tissue samples

	UBCE	**GAPDH**	**NAD 5-2**	**NAD 5-1**	**ALPHA TUBULIN**	**EIF**	**BETA ACTIN**	**EF2**	**ACTINCRD**	**UBQ2**
	HKG 1	**HKG 2**	**HKG 3**	**HKG 4**	**HKG 5**	**HKG 6**	**HKG 7**	**HKG 8**	**HKG 9**	**HKG 10**
n	9	9	9	9	8	8	8	8	8	8
geo Mean [CP]	26.61	34.33	18.64	19.73	21.21	21.54	26.04	24.99	27.40	30.33
ar Mean [CP]	26.70	34.34	18.67	19.77	21.28	21.60	26.07	25.02	27.42	30.40
min [CP]	21.29	32.43	18.26	18.31	19.22	19.73	24.63	23.31	25.48	26.95
max [CP]	28.81	35.13	21.84	22.04	25.31	25.37	29.20	27.24	28.59	33.07
std dev [± CP]	1.54	0.51	0.70	1.10	1.31	1.01	0.90	1.01	0.72	1.68
CV [% CP]	5.77	1.49	3.77	5.57	6.15	4.68	3.46	4.04	2.61	5.53
min [x-fold]	-40.07	-3.73	-1.30	-2.68	-3.97	-3.52	-2.66	-3.20	-3.79	-10.41
max [x-fold]	4.58	1.74	9.17	4.96	17.18	14.23	8.92	4.76	2.27	6.69
std dev [± x-fold]	2.91	1.43	1.63	2.14	2.48	2.02	1.87	2.01	1.64	3.21

Additionally, this programme provided the standard deviation between the Cq values of each gene and the entire data set. A strong candidate for a reference gene in qPCR analysis is the gene with the highest coefficient of correlation since it is believed to have the best stability[28].

Based on the result from BestKeeper, NAD52 most stable reference gene shows least standard deviation (SD) and coefficient of correlation (r) values of 0.01 and -0.09. It is consistent with RefFinder and NormFinder. The results of RefFinder and NormFinder supported Best Keeper's identification of the most unstable genes: UBQ2 and TUBULIN had the greatest SD and r value as in table 3.

Validation of reference gene: Wild type flower's reproductive organs such as stamen and ovule are developed by the influence of C-class function floral organ identity gene AGAMOUS and D class function ovule identity gene SEEDSTICK which are also known as AG and STK[12]. In this study, we identified and sequenced AG (GenBank OP629949) and STK (GenBank OQ129372) genes. These genes are used for the validation of selected stable

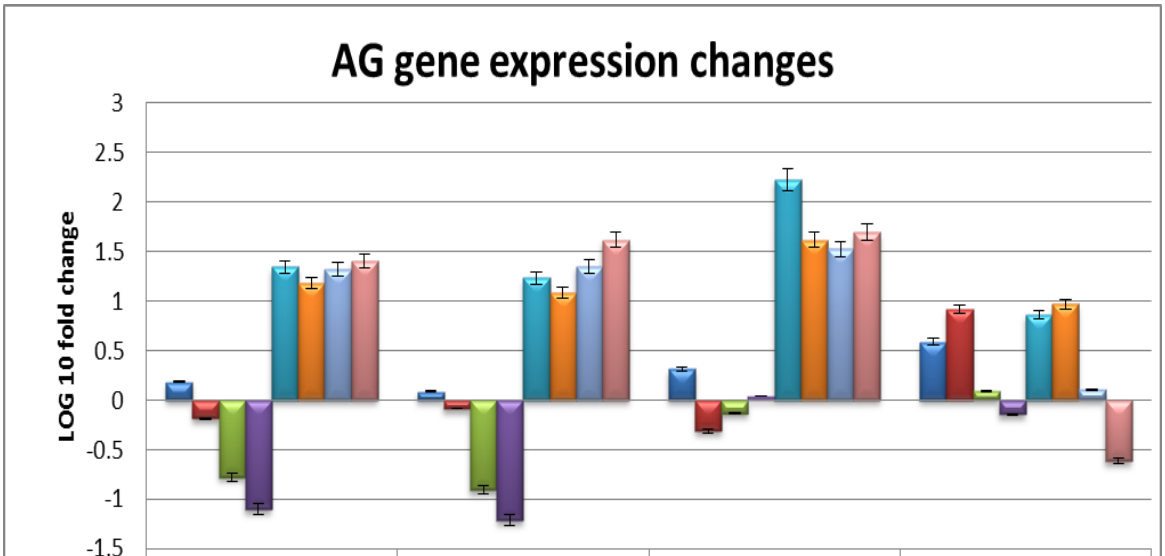
housekeeping genes. Four housekeeping genes were chosen as the calibrator in order to ensure that the those genes were adequate including the two most stable reference genes NAD5-2 and NAD5-1 as well as the two moderately stable reference genes UBCE and EIF. The delta CT method was used to determine the relative expression levels of the AG and STK genes in relation to the chosen reference genes. According to ABCDE model of floral organ development, AG gene involved in the stamen and carpel development and STK specific for ovule development[30,36,43,46].

Based on these information, figure 6 and figure 7 illustrated how various reference genes caused the expression pattern of AG and STK genes to vary more in various tissues. The AG and STK gene expression patterns were consistent using the most reliable housekeeping genes NAD5-2. The expression level AG gene in the mature male flower was upregulated and STK gene were down regulated in male flowers and up regulated in mature and immature ovules with respect to NAD5-2. But in the case of other three reference genes, the target expression shows drastic changes.

However, when the moderately stable housekeeping gene UBCE was employed, AG gene expression level shows a considerable bias when compared to other stable genes. Although NAD5-1 and EIF exhibit more variation in STK gene expression level, this finding showed that accurate normalization of the target gene's expression required the more stable housekeeping gene. Four housekeeping genes were chosen for validation based on the results of various software tools and a one-way ANOVA Duncan test using their CT values in triplicates was run. Consequently, it was revealed that NAD5 -2 has the lowest SE and SD and is more stable as in table 4.

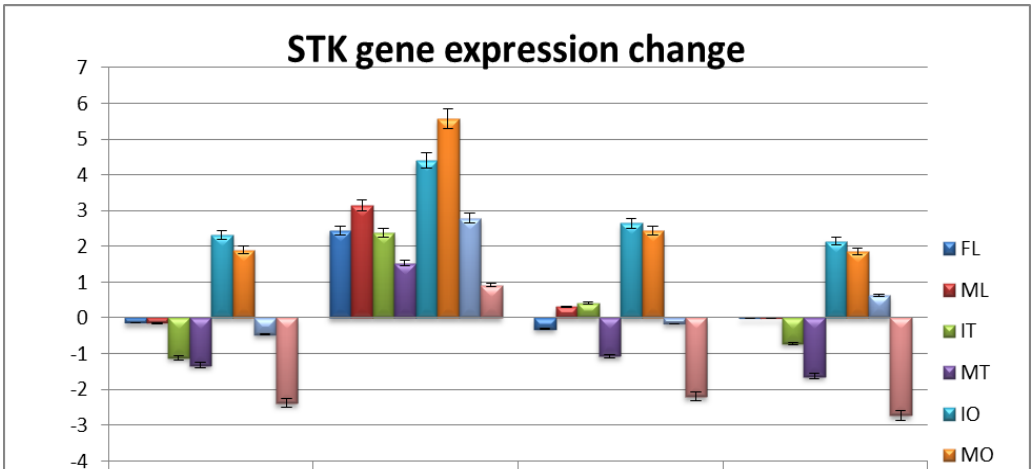
Discussion

Real-time PCR/qPCR assays have emerged as the preferred method for the quick identification and quantification of nucleic acids in a different biological samples with numerous applications including gene expression studies. The most stable reference genes must be identified in order to produce accurate qPCR results[38]. When target gene expression levels fluctuate due to technical issues with total RNA or cDNA synthesis quantity or quality, the expression of reference genes is employed to remedy the problem[5]. However, recent studies have found that these reference genes may display instability in various plant species or genotypes. organs and various experimental setups[40]. The reference genes UBQ and HSP are stable in date palm drought-stressed leaf samples whereas GAPDH is stable in drought-stressed roots to illustrate the aforementioned fact[20].



	NAD52	UBCE	EIF	NAD51
FL	0.183929327	0.086245094	0.313071195	0.590018792
ML	-0.183929327	-0.086245094	-0.313071195	0.918141487
IP	-0.778463569	-0.906852862	-0.132453198	0.093921359
MP	-1.101167724	-1.212398308	0.041241109	-0.143892338
IO	1.344700991	1.231965257	2.233642568	0.864859178
MO	1.181542733	1.084159529	1.621347557	0.971423796
YMF	1.324833011	1.352076226	1.527426198	0.108671828
MMF	1.40340184	1.619691892	1.694798876	-0.614101191

Figure 6: AGAMOUS expression changes in relation with four housekeeping genes of most stable, moderately stable and least stable housekeeping genes by using Delta CT method.



	NAD5-2	NAD5-1	UBCE	EIF
FL	-0.140861703	2.443258292	-0.316142795	-0.022093572
ML	-0.153477332	3.139532142	0.316180099	-0.022093572
IT	-1.115204636	2.370511239	0.410608543	-0.715118285
MT	-1.319664487	1.531137734	-1.069050969	-1.634512015
IO	2.321693687	4.397222031	2.641637684	2.143033546
MO	1.900071489	5.572757588	2.428008376	1.854846296
YMF	-0.463062977	2.780111279	-0.166215625	0.640580806
MMF	-2.387216143	0.909235003	-2.193820026	-2.721246399

Figure 7: SEEDSTICK expression changes in relation with four housekeeping genes by using Delta CT method

Table 4
Means within a column followed by same letters are not significantly (p<0.05) different as determined by Duncan's multiple range test

…….	NAD5-2 (Mean ± SE)	NAD5-1 (Mean ± SE)	UBCE (Mean ± SE)	EIF (Mean ± SE)
FL	20.25±0.01^{c}	18.30±0.01^{f}	26.06±0.09^{e}	21.25±0.14^{bc}
ML	21.09±0.09^{a}	20.42±0.01^{cd}	27.98±0.09^{c}	21.21±0.37^{bc}
IT	20.19±0.03^{c}	21.18±0.03^{abc}	28.99±0.19^{a}	22.08±0.74^{b}
MT	20.22±0.01^{c}	19.11±0.16^{ef}	27.40±0.11^{d}	19.73±0.04^{c}
IO	20.24±0.02^{c}	19.66±0.39^{de}	27.65±0.11^{d}	20.77±0.36^{bc}
MO	20.26±0.01^{c}	21.88±0.16^{a}	28.37±0.08^{b}	20.67±0.29^{bc}
YMF	20.68±0.14^{b}	20.87±0.58^{bc}	28.01±0.07^{c}	24.90±0.55^{a}
MMF	21.05±0.24^{a}	21.43±0.41^{ab}	28.05±0.09^{bc}	20.55±0.65^{c}
Df (n-1) = 7				
F value	14.25***	17.29***	60.48***	11.95***

***significant F value at p<0.001

Another significance of this study is that Palmyra palm is a dioecious plant and the reference gene standardization and floral organ identity gene expression analysis have been performed. Nine candidate reference genes were used in this study. Primer pairs for these genes were carefully designed to enable the amplification of particular PCR products from reverse transcribed cDNA. The appearance of a single amplicon peak in the melting curve analysis provided additional evidence that a single PCR product had been amplified[19].

Under the vegetative and reproductive tissues of the dioecious palm, the expression stability and validity of these selected reference genes were investigated. For gene expression profiling in various plants, these reference genes were also applied. According to prior research, multiple analytic programs are needed for a precise conclusion. Contradictory conclusions may be drawn from the same data because various statistical programs use different calculating principles[47].

In contrast to what has previously been reported, the expression stability assessed by the geNorm, NormFinder and BestKeeper was not constant[1,40]. In order to thoroughly rate the reference genes and evaluate the outcomes of other tools, we employed the RefFinder programme[7,17,21,23,34,35,41,45,48,49]. To cross-validate the expression stability of the potential reference genes in earlier

investigations in the oil palm, date palm and coconut palm, the geNorm, RefFinder, NormFinder and BestKeeper algorithms were employed and get good result[23,27,32]. As a result, we had a satisfactory result from the tools we used.

Expression stability is a key attribute of a reliable reference gene that may be utilized as an internal control. Ideally, tissue types, developmental stages, physiological conditions and experimental settings have no impact on the stability of expression. Our results show that in the leaf and flower whorls of male and female Palmyra palm, NAD5-2 and tubulin are the most stable and least stable reference genes respectively. Related research in the oil palm revealed that NAD5 and UBIQUITIN are abundantly expressed in their chosen sample types but exhibit the lowest expression stability[4]. Reference gene standardization tests in a Salix dioecious plant revealed that ACT and DnaJ were the most stable genes for both female and male flowers at various developmental stages[48].

The target gene's expression levels should ideally be comparable to those of the reference genes[6]. In our study, RefFinder comprehensive ranking displayed that NAD5-1 and NAD5-2 showed highest expression and lowest CT value. The lowest expression and greatest Ct values were seen in UBQ, in contrast. The expression of EIF, UBCE and EF2 was moderate. Based on the level of the target gene expression as well as the particular experimental condition, we recommend to select a stable reference gene. The process of plant development involves a huge number of genes. In our studies, the expression of the AGAMOUS and SEEDSTICK target genes were used to evaluate the stability of the four reference genes that, out of ten, were the most and moderately stable in a variety of tissues of the Palmyra palm. Based on the relative target gene expression data, NAD5-2 is most accurate and consistent reference gene in Palmyra palm.

NAD52 is the most stable reference gene in male and female leaves as well as various developmental stages of floral whorls of the Palmyra palm, according to the different algorithms we applied and thorough examination of RefFinder. For the NADH dehydrogenase subunits on the mitochondrial genome, NAD5 is one of the classic housekeeping genes[25]. Almost no changes have been seen in the amino acid sequence of the protein NAD5, which has been present throughout plant evolution[9]. NAD5 is highly and consistently expressed in a variety of tissues in the current investigation. The proper development of plants has been made possible mainly by the coordinated control of a huge number of genes associated with development. A good foundation for future study on Palmyra palm development is provided by the selection of appropriate reference genes.

Conclusion

In this study, four software tools delta-Ct, BestKeeper, RefFinder and NormFinder were used to examine the expression stability of nine reference genes under various plant tissues of dioecious palm. The results showed that NAD5-2 was an ideal reference gene in all of the chosen samples, suggesting that it would be a good reference gene for RT-qPCR experiments for *Borassus flabellifer*.

Acknowledgement

We gratefully acknowledge the facilities provided by the Department of Botany at the University of Kerala in Kariavattom and the Biotechnology and Bioinformatics Division at the Jawaharlal Nehru Tropical Botanic Garden and Research Institute in Palode. We also thank the lab staff for their discussions and manuscript feedback. We are also thankful for the research fellowship that Suchilkumar Soorya received from the UGC.

References

1. Abbas A., Yu H., Li X., Cui H., Chen J. and Huang P., Selection and validation of reference genes for rt-qpcr analysis in aegilops tauschii (Coss.) under different abiotic stresses, *International Journal of Molecular Sciences*, **22(20)**, https://doi.org/10.3390/ijms222011017 **(2021)**

2. Bustin S.A., Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems, *Journal of Molecular Endocrinology*, **29(1)**, 23–39, https://doi.org/10.1677/jme.0.0290023 **(2002)**

3. Cancer C. and Sets D., Documentation for running Normfinder in R The Normfinder algorithm, 1–7 **(2013)**

4. Chan P.L. et al, Evaluation of reference genes for quantitative real-time PCR in oil palm elite planting materials propagated by tissue culture, *PLoS ONE*, **9(6)**, https://doi.org/10.1371/journal.pone.0099774 **(2014)**

5. Derveaux S., Vandesompele J. and Hellemans J., How to do successful gene expression analysis using real-time PCR, *Methods*, **50(4)**, 227–230, https://doi.org/10.1016/j.ymeth.2009.11.001 **(2010)**

6. Dheda K., Huggett J.F., Bustin S.A., Johnson M.A., Rook G. and Zumla A., Validation of housekeeping genes for normalizing RNA expression in real-time PCR, *BioTechniques*, **37(1)**, 112–119, https://doi.org/10.2144/04371rr03 **(2004)**

7. Gao P., Wang J. and Wen J., Selection of reference genes for tissue/organ samples of adults of Eucryptorrhynchus scrobiculatus, *PLoS ONE*, **15(2)**, 1–15, https://doi.org/10.1371/journal.pone.0228308 **(2020)**

8. Ginzinger D.G., Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream, *Experimental Hematology*, **30(6)**, 503–512, https://doi.org/10.1016/S0301-472X(02)00806-8 **(2002)**

9. Groth-Malonek M., Rein T., Wilson R., Croth H., Heinrichs J. and Knoop V., Different Fates of Two Mitochondrial Gene Spacers in Early Land Plant Evolution, *International Journal of Plant Sciences*, **168(5)**, 709–717, https://doi.org/10.1086/513472 **(2015)**

10. Huggett J., Dheda K., Bustin S. and Zumla A., Real-time RT-PCR normalisation; strategies and considerations, *Genes and*

Immunity, **6(4)**, 279–284, https://doi.org/10.1038/sj.gene.6364190 **(2005)**

11. IDT, IDT Primerquest Tool **(2017)**

12. Irish V., The ABC model of floral development, *Current Biology*, **27(17)**, R887–R890, https://doi.org/10.1016/j.cub.2017.03.045 **(2017)**

13. Fjain M., Nijhawan A., Tyagi A.K. and Khurana J.P., Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR, *Biochemical and Biophysical Research Communications*, **345(2)**, 646–651, https://doi.org/10.1016/j.bbrc.2006.04.140 **(2006)**

14. Jana H. and Jana S., Palmyra palm: Importance in Indian agriculture, *Rashtriya Krishi*, **12(2)**, 2–35 **(2017)**

15. Jia Y., Liu S.C., Zhao J., Chen X.X., Sun L.X., Yu X.F. and Li X., Reference gene selection and validation by qRT-PCR during flower development and in different organs of Primula forbesii, *Journal of Horticultural Science and Biotechnology*, **95(3)**, 383–394, https://doi.org/10.1080/14620316.2019.1681909 **(2020)**

16. Jin Y., Liu F., Huang W., Sun Q. and Huang X., Identification of reliable reference genes for qRT-PCR in the ephemeral plant Arabidopsis pumila based on full-length transcriptome data, *Scientific Reports*, **9(1)**, 1–11, https://doi.org/10.1038/s41598-019-44849-1 **(2019)**

17. Karuppaiya P., Yan X.X., Liao W., Wu J., Chen F. and Tang L., Identification and validation of superior reference gene for gene expression normalization via RT-qPCR in staminate and pistillate flowers of Jatropha curcas - A biodiesel plant, *PLoS ONE*, **12(2)**, 1–17, https://doi.org/10.1371/journal.pone.0172460 **(2017)**

18. Kwan Y.M., Meon S., Ho C.L. and Wong M.Y., Selection of reference genes for quatitative real-time PCR normalization in Ganoderma-infected oil palm (Elaeis guineensis) seedlings, *Australasian Plant Pathology*, **45(3)**, 261–268, https://doi.org/10.1007/s13313-016-0417-4 **(2016)**

19. Li J., Zhang Z., Xu C., Wang D., Lv M. and Xie H., Identification and validation of reference genes for real-time RT-PCR in Aphelenchoides besseyi, *Molecular Biology Reports*, **47(6)**, 4485–4494, https://doi.org/10.1007/s11033-020-05547-8 **(2020)**

20. Li M.Y., Wang F., Jiang Q., Wang G.L., Tian C. and Xiong A.S., Validation and comparison of reference genes for qPCR normalization of celery (Apium graveolens) at different development stages, *Frontiers in Plant Science*, **7**, 1–12, https://doi.org/10.3389/fpls.2016.00313 **(2016)**

21. Li W., Zhang L., Zhang Y., Wang G., Song D. and Zhang Y., Selection and validation of appropriate reference genes for quantitative real-time PCR normalization in staminate and perfect flowers of andromonoecious Taihangia rupestris, *Frontiers in Plant Science*, **8**, 1–13, https://doi.org/10.3389/fpls.2017.00729 **(2017)**

22. Lilly S.T., Drummond R.S.M., Pearson M.N. and MacDiarmid R.M., Identification and validation of reference genes for normalization of transcripts from virus-infected Arabidopsis thaliana, *Molecular Plant-Microbe Interactions*, **24(3)**, 294–304, https://doi.org/10.1094/MPMI-10-10-0236 **(2011)**

23. Mohd Zuhar L., Determination of reference genes for normalisation of gene expression study of Ganoderma-infected oil palms, *Journal of Oil Palm Research*, **31**, 550–560, https://doi.org/10.21894/jopr.2019.0051 **(2019)**

24. New P., Palmyra palm is a genus of five species of fan palms, native to tropical regions of Africa, Asia and Papua New Guinea **(2017)**

25. Nozato N. et al, Cotranscriptional expression of mitochondrial genes for subunits of NADH dehydrogenase, nad5, nad4, nad2, in Marchantia polymorpha, *Molecular & General Genetics: MGG*, **237(3)**, 343–350, https://doi.org/10.1007/BF00279437 **(1993)**

26. Paolacci A.R., Tanzarella O.A., Porceddu E. and Ciaffi M., Identification and validation of reference genes for quantitative RT-PCR normalization in wheat, *BMC Molecular Biology*, **10**, 1–27, https://doi.org/10.1186/1471-2199-10-11 **(2009)**

27. Patankar H.V., Assaha D.V.M., Al-Yahyai R., Sunkar R. and Yaish M.W., Identification of reference genes for quantitative real-time PCR in date palm (Phoenix dactylifera L.) subjected to drought and salinity, *PLoS ONE*, **11(11)**, https://doi.org/10.1371/journal.pone.0166216 **(2016)**

28. Pavithra M.K.S., Procedure for chromatography using mobile phase with high elution strength to recover tacrolimus from the fermentation broth, *Res. J. Chem. Environ.*, **27(1)**, 71-77 **(2023)**

29. Pfaffl M.W., Tichopad A., Prgomet C. and Neuvians T.P., Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper - Excel-based tool using pair-wise correlations, *Biotechnology Letters*, **26(6)**, 509–515, https://doi.org/10.1023/B:BILE.0000019559.84305.47 **(2004b)**

30. Pinyopich A., Ditta G.S., Savidge B., Liljegren S.J., Baumann E., Wisman E. and Yanofsky M.F., Assessing the redundancy of MADS-box genes during carpel and ovule development, *Nature*, **424(6944)**, 85–88, https://doi.org/10.1038/nature01741 **(2003)**

31. Pipatchartlearnwong K., Juntawong P., Wonnapinij P., Apisitwanich S. and Vuttipongchaikij S., Towards sex identification of Asian Palmyra palm (Borassus flabellifer L.) by DNA fingerprinting, suppression subtractive hybridization and de novo transcriptome sequencing, *Peer J*, **2019(7)**, https://doi.org/10.7717/peerj.7268 **(2019)**

32. Rachana K.E. and Rajesh M.K., Selection and validation of reference genes for gene expression normalization in coconut (Cocos nucifera L.) under biotic stress and hormone stimuli, *Plant Gene*, **19**, https://doi.org/10.1016/j.plgene.2019.100184 **(2019)**

33. Sakulsathaporn A., Wonnapinij P., Vuttipongchaikij S. and Apisitwanich S., The complete chloroplast genome sequence of Asian Palmyra palm (Borassus flabellifer), *BMC Research Notes*, **10(1)**, 1–7, https://doi.org/10.1186/s13104-017-3077-8 **(2017)**

34. Samarth and Jameson P.E., Selection of reference genes for flowering pathway analysis in the masting plants, Celmisia lyallii and Chionochloa pallens, under variable environmental conditions, *Scientific Reports*, **9(1)**, 1–16, https://doi.org/10.1038/s41598-019-45780-1 **(2019)**

35. Sarwar M.B. et al, Identification and validation of superior housekeeping gene(s) for qRT-PCR data normalization in Agave sisalana (a CAM-plant) under abiotic stresses, *Physiology and Molecular Biology of Plants*, **26(3)**, 567–584, https://doi.org/10.1007/s12298-020-00760-y **(2020)**

36. Tani E. et al, Characterization and expression analysis of AGAMOUS-like, SEEDSTICK-like and SEPALLATA-like MADS-box genes in peach (Prunus persica) fruit, *Plant Physiology and Biochemistry*, **47(8)**, 690–700, https://doi.org/10.1016/j.plaphy.2009.03.013 **(2009)**

37. Tong Z., Gao Z., Wang F., Zhou J. and Zhang Z., Selection of reliable reference genes for gene expression studies in peach using real-time PCR, *BMC Molecular Biology*, **10**, 1–13, https://doi.org/10.1186/1471-2199-10-71 **(2009)**

38. Udvardi M.K., Czechowski T. and Scheible W.R., Eleven golden rules of quantitative RT-PCR, *Plant Cell*, **20(7)**, 1736–1737, https://doi.org/10.1105/tpc.108.061143 **(2008)**

39. Vera Hernández F.P., Martínez Núñez M., Ruiz Rivas M., Vázquez Portillo R.E., Bibbins Martínez M.D., Luna Suárez S. and Rosas Cárdenas F. de F., Reference genes for RT-qPCR normalisation in different tissues, developmental stages and stress conditions of amaranth, *Plant Biology*, **20(4)**, 713–721, https://doi.org/10.1111/plb.12725 **(2018)**

40. Wang H., Wang J., Jiang J., Chen S., Guan Z., Liao Y. and Chen F., Reference genes for normalizing transcription in diploid and tetraploid arabidopsis, *Scientific Reports*, **4**, 1–8, https://doi.org/10.1038/srep06781 **(2014)**

41. Wang M. and Lu S., Validation of suitable reference genes for quantitative gene expression analysis in Panax ginseng, *Frontiers in Plant Science*, **6**, https://doi.org/10.3389/fpls.2015.01259 **(2016)**

42. Xie F., Xiao P., Chen D., Xu L. and Zhang B., miRDeepFinder: A miRNA analysis tool for deep sequencing of plant small RNAs, *Plant Molecular Biology*, **80(1)**, 75–84, https://doi.org/10.1007/s11103-012-9885-2 **(2012)**

43. Yamada K., Saraike T., Shitsukawa N., Hirabayashi C., Takumi S. and Murai K., Class D and Bsister MADS-box genes are associated with ectopic ovule formation in the pistil-like stamens of alloplasmic wheat (Triticum aestivum L.), *Plant Molecular Biology*, **71(1–2)**, 1–14, https://doi.org/10.1007/s11103-009-9504-z **(2009)**

44. Yeap W.C., Loo J.M., Wong Y.C. and Kulaveerasingam H., Evaluation of suitable reference genes for qRT-PCR gene expression normalization in reproductive, vegetative tissues and during fruit development in oil palm, *Plant Cell, Tissue and Organ Culture*, **116(1)**, 55–66, https://doi.org/10.1007/s11240-013-0382-3 **(2014)**

45. Yu Y. et al, Selection of Reference Genes for qPCR Analyses of Gene Expression in Ramie Leaves and Roots across Eleven Abiotic/Biotic Treatments, *Scientific Reports*, **9(1)**, 1–13, https://doi.org/10.1038/s41598-019-56640-3 **(2019)**

46. Zhang X. et al, Identification and characterization of the mads-box genes and their contribution to flower organ in carnation (Dianthus Caryophyllus L.), *Genes*, **9(4)**, 1–21, https://doi.org/10.3390/genes9040193 **(2018)**

47. Zheng T., Chen Z., Ju Y., Zhang H., Cai M., Pan H. and Zhang Q., Reference gene selection for qRT-PCR analysis of flower development in Lagerstroemia indica and L. speciosa, *PLoS ONE*, **13(3)**, 1–14, https://doi.org/10.1371/journal.pone.0195004 **(2018)**

48. Zhou F., Chen Y., Wu H. and Yin T., A Selection of Reliable Reference Genes for Gene Expression Analysis in the Female and Male Flowers of Salix suchowensis, *Plants*, **11(5)**, https://doi.org/10.3390/plants11050647 **(2022)**

49. Zhu X., Li X., Chen W., Chen J., Lu W., Chen L. and Fu D., Evaluation of New Reference Genes in Papaya for Accurate Transcript Normalization under Different Experimental Conditions, *PLoS ONE*, **7(8)**, https://doi.org/10.1371/journal.pone.0044405 **(2012)**.

(Received 23rd September 2023, accepted 27th October 2023)