

OPEN ACCESS

Selection of suitable housekeeping genes for gene expression study in caraway (Bunium persicum)

Davod Darvishi Zeidabadi¹, Mokhtar Jalali Javaran^{2*}, Hamid Dehghani², Sajad Rashidi Monfared², Amin Baghizadeh³

¹Department of Plant Breeding, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran ²Department of Plant Breeding, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran ³Department of Biotechnology, Graduate University of Advanced Technology, Kerman, Iran

Article published on September 17, 2015

Key words: Housekeeping gene, Gene expression, *Bunium persicum*, Data normalization, Quantitative PCR technique.

Abstract

One of the most important issues in genetic engineering and molecular genetics is to study target gene expression that is aimed to quantify data resulted from measuring mRNA levels. Studying amount of gene expression by Real time PCR technique should normalize and compare the amount of mRNA in target gene to amount of mRNA in a housekeeping gene. The level of housekeeping gene expression is stable in different tissues and in all growth stages of an organism and it is not affected by experimental treatments. In each test of gene expression, several housekeeping genes should be studied in order to choose the best housekeeping gene for normalization of data resulted from the target gene expression. In the present study, stability of three housekeeping gene expressions, including actin-beta (*ACT*), ubiquitin (*UBC*) and elongation factor 1-alpha (*EF1A*) has been evaluated in germination, multi-leaf and flowering stages using real time PCR technique. Results obtained from data analysis via geNorm and Bestkeeper-1 software tools showed that EF1A gene is expressed more stable than two other genes and standard error (SE) is higher in the *ACT* and *UBC* genes and they are less expressed than *EF1A*. Concerning studies done on stability, uniformity and stability of *EF1A* gene expression in different ecotypes of*B*. *persicum* is recommended to normalize data of the target gene expression.

*Corresponding Author: Mokhtar Jalali Javaran 🖂 m_jalali@modares.ac.ir

Introduction

Kermani Caraway (Bunium persicum) is one of important pharmaceutical plants with high economic value in Iran and throughout the world (Sofi et al., 2009; Jalilzadeh-Amin et al., 2011). This diploid plant belongs to Apiaceae or Umbeliferae family and naturally it is found in moderate and dry regions of Iran and some countries such as Afghanistan and north of Pakistan (Jalilzadeh-Amin et al., 2011). Oroojalian et al. (2010) indicated that terpene materials are the most common compounds present in b. persicum among which gamma terpinene has the highest amount. Today, quantitative PCR technique is used to measure and compare RNA levels due to its high accuracy, repeatability, specialization, sensitivity and operational power (Infante et al., 2008; McCurley and Callard, 2008; Small et al., 2008; Zhong et al., 2008; Mitter et al., 2009). Real time PCR method has some advantages compared to conventional PCR methods and the most important advantages are as follows: amount of the product can be tracked in each cycle while the product of traditional methods is determined at the end of electrophoretic reaction. It is possible to study and analyze several transcripts in one tube. Its dynamic magnitude and sensitivity are 1000 times of traditional RT-PCR. By using this technique, quantitative valuation can be done and amount of the primary pattern can be estimated. Despite its advantages, this technique is unable to estimate size of proliferated product and it is expensive (Pfaffl, 2001). In this technique, amount of mRNA in the target gene compared to that in internal housekeeping gene with a relative stability and uniformity can be measured to quantify data resulted from measuring mRNA levels (Bustin, 2000). Normalization of data resulted from quantitative PCR is necessary to control and remove the error resulted from application of samples in different time stages within experimental process (Dheda et al., 2005). Level of gene expression of internal control should be constant and stable among different tissues of an organism and in all growth stages and its expression should not be affected by experimental treatments

(Bustin, 2000). Shamsifard et al. (2014) indicated that transcript expression of GAPGH housekeeping gene is nearly equal in tissues of root, stem, leaf, capsule and seed of two early maturing and late maturing genotypes of B. persicum (Shamsi-Fard et al., 2014). During testing gene expression, stability of several housekeeping gene expressions should be studied and evaluated in order to choose the best housekeeping gene for normalization of data resulted from target gene expression (Tang et al., 2007). In order to choose and introduce a housekeeping gene, a comprehensive research is essential to ensure that no adjustment is performed on them. Stability of the housekeeping gene expression is the first condition for internal standard of data related to the target gene expression and many housekeeping genes with assumed stable expression may undergo adjustment of up or down-stream genes (Pfaffl, 2004). Software used for studying housekeeping genes shows some differences in results because they have no equal calculative basis. Data can be analyzed by several software tools and results can be compared. Therefore, software tools such as Bestkeeper-1 (Bfaffl et al., 2004), geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), qBase (Hellemans et al., 2007) and GenEx (MultiD Analysis AB, Sweden, version GenEx 5 Enterprise) can be used. It is noteworthy that no study has been done on gene expression of B. persicum. The first step in studying gene expression in every organism is to identify a suitable housekeeping gene and the present study is the first step taken for investigation of gene expressions in B. persicum.

Methods and materials

Plant Materials

No information related to housekeeping genes of *B. persicum* has been recorded yet in the gene bank. Therefore, in order to identify genes involved in biosynthetic pathway of secondary metabolites, similar information related to protected regions for other plants, especially co-family plants should be used. First, the seed of five ecotypes of *B. persicum* was gathered from provenances of four provinces.

These five ecotypes are from Esfahan, Kerman1 (Khabr), Semnan, Kerman2 (Sirch) and Hormozgan. A sample from each ecotype was washed and disinfected individually and after treating with hormonal compounds (Giberellic acid and Tidiazuzon), they were refrigerated for three weeks in 4 °C. By starting germination, the seeds were placed in sterile petri dishes in 15 °C within germinator. One week after germination, seedlings were sampled. Weight of each sample was 100 mg and the samples were placed in refrigerator in -80 °C until RNA extraction. In flowering and multi-leaf stages, scrubs cultivated in pots were sampled. Complete random blocks with three replications have been used and the samples were mixed with each other and used to extract RNA.

Housekeeping Genes

Among housekeeping genes (*UBC*, *EF1A*, *GAPDH1*, *GAPDH2*, *ACT*, *ACT2*, *PPa2*), three genes so called Ubiquitin (*UBC*) (AEQ75497.1), Actin (*ATC*) (Yousefzadei ,2012) and *Elongation factor1 alfa* (*EF1A*) (JN399225.1) were selected (Table 1).

These genes had better proliferation in primary PCR that was performed for testing primers and the length of the part proliferated by them was proper as well and suitable primers were designed and prepared for them regarding length and required features. Some features of primers such as annealing, GC percentage, hairpin, self-dimer and hetero-dimer were studied using Integrated DNA OligoAnalyser Technologies, Inc. extraction of total RNA from different tissues of B. persicum was done by Ribospin RNA plant kit the manufacturer according to (GeneAll). Concentration of extracted RNA and ultraviolet absorption ratios of 260/280 nm and 260/230 were measured using 2µl of RNA via Nanodrop system, Model Biotec. cDNA was fabricated by RevertAid Ferst strand cDNA synthesis kit from Fermentase company. Housekeeping genes chosen in a common PCR had good performance and they were proliferated with cDNA of Bunium persicum.

Real Time PCR

In this stage, three samples were separated from five ecotypes of fresh tissues of the plant during germination, multi-leaf and flowering stages of B. persicum and total RNA was extracted after mixing cDNA three samples. Then, with certain concentration was prepared from different tissues of seedlings different in growth stages and simultaneously primers of these three genes with three replications were used in Real Time PCR program. Real time PCR conditions in BioRad system include temperature of 95 °C for 15 minutes for denaturation or primary activation (because DNA polymerase requires high temperature in a long time to be activated), 40 cycles, each cycle with 95°C for 20 seconds, optimal temperature of primer annealing (57 °C) for 20 seconds and 72 °C for 30 seconds and the final cycle with 60-95 °C for 5 minutes for obtaining melt curve. Levels of gene expression are recorded as CT values which mean a cycle of PCR where fluorescent signal is beyond assumed threshold line. The base line and assumed threshold line were adjusted automatically using CFX software (CFX manager, 1999). The line is drawn where the curve related to genes is ascending and it is plotted on the best point automatically by the software but it can be changed. Negative control in the reaction including application of all reactive compounds except cDNA was prepared in order to study possibility of external contamination. Since Real time PCR using Eva Green (that is non-specialized fluorcent color) is able to bind the double stranded DNA and to form dimers and the results are overestimated, it is necessary to ensure the least proliferation of non-specialized DNA of dimers by CFX software. Results obtained from data related to different growth stages of b. persicum were analyzed to study stability of housekeeping gene expressions and to select the best gene using Bestkeeper-1 software (Pfaffl et al., 2004). It should be noted that in order to analyze information via this software, information related to specialized genes and housekeeping genes can be studied simultaneously. This software is able to analyze simultaneously information related to ten specialized genes and ten

housekeeping genes with more than 100 samples (Pfaffl *et al.*, 2004). The study of the best housekeeping gene was done by geNorm software (Vandesompele *et al.*, 2002) and obtained results are based on calculation of stability Eigen value of gene expression (M-value) and the less this value, the higher the gene expression.

Results and discussion

Housekeeping genes chosen in conventional PCR had good performance and cDNA of *B. persicum* was proliferated well and among seven housekeeping genes (*UBC, EF1A, GAPDH1, GAPDH2, ACT, ACT2* and *PPa2*), three genes of *UBC, EF1A* and *ACT* were more optimal so they were selected (fig.1). In order to measure transcript level of RNA in housekeeping genes, Real Time PCR was done. After testing and plotting melt curve (Melt curve analysis) for the three housekeeping genes (*ACT, EF1A, UBC*), non-specialized DNA proliferation and dimers' proliferations should be ensured and it was determined that melt curve has only one peak point for each gene (fig.2).

Results of data analyzed by Bestkeeper-1 software ranked stability of housekeeping genes based on raw values of CT as follows: UBC > EF1A > ACT and correlation of each gene with Bestkeeper-1 index is 0.97, 0.93 and 0.77 for *UBC*, *Ef1A* and *ACT* respectively (Table 2).

Table 1. Primer sequence characteristics used as Housekeeping Genes in Caraway.

Primer	5' 3´ ->	GC%	Length of fragment	Tm(°C)
name	Primer sequence			
F-UBC	CCA AAG GTT GCA TTC AGG AC	50	241	54.3
R-UBC	ACT TCT GGG TCC AGC TCC TT	55		58
F-EF1A	CTG GTG GTT TTG AAG CTG GT	50	243	55.8
R-EF1A	TGT TGT CAC CCT CGA ATC CA	50		56.2
F-ACT	GCA GGG ATC CAC GAG ACC ACC	66.7	93	62.3
R-ACT	CCC ACC ACT GAG CAC AAT GTT CC	56.5		60.5

Table 2. The results of analysis different stage data by Bestkeeper-1 software for selection of the best housekeeping gene of five Caraway (*B. persicum*) ecotypes.

Correlation Analysis	UBC	EF1A	ACT
Coefficient of Correlation	0.97	0.93	0.78
Coefficient of determination	0.95	0.87	0.6
Intercept	-21.85	7.40	12.32
Slope	1.98	0.52	0.63
CT means	26.20	19.93	27.52
SE [CT]	0.80±	±0.342	± 0.865
SD[CT]	2.62	0.70	1.01
p-value	0.001	0.001	0.002
Power of HKG [x-fold]	3.95	1.43	1.54
CV%	9.94	3.49	3.68

geNorm software ranked value of gene stability for each gene as follows: ACT = EF1A = 0.057 and UBC = 0.11. These results showed that primer of *UBC* gene although has high correlation coefficient, is not suitable as a housekeeping gene due to low expression (high average of CT), high standard error. Therefore, between primers of two genes of *EF1A* and ACT, *EF1A* primers has been better than *ACT* and *UBC* primers due to good correlation coefficient (0.93) with Bestkeeper-1 index, proper explanatory coefficient, low SE and high expression. Based on results of geNorm software, the best housekeeping gene or internal control in *B. persicum* is *EF1A* that its stability Eigen value of gene expression is 0.057 that is lower than values of other genes (fig.3).

Plotting Heatmap based on CT (threshold cycle) values related to each gene by GenEx software (MultiD Analyses. 2015) confirms results analyzed by Bestkeeper-1 and geNorm software. According to this map, *EF1A* is expressed higher and more uniform than two other genes (fig. 4).

In cucumber, gene expression of *EF1A* was the most stable in biotic and abiotic stresses and in different tissues (Wan *et al.*, 2010). *EF1A* gene was the most stable in condition of nitrogen shortage in tomato (Lovdal and Lillo, 2009).



Fig. 1. Investigation of PCR product by agarose gel electrophoresis 1% for comparison of seven housekeeping gene in Caraway: M= Fermentase 1kb, 1=*UBC*, 2=*EF1A*, 3=*GAPDH1*, 4=*GAPDH2*, 5=*ACT*, 6=*ACT2*, 7=*PPa2*, C⁺=*IPI* and C^{-.}

This gene is good for normalization of many tissues in soybean and citruses or flower organs of cotton (Libault *et al.*, 2008; Artico, 2010; Mafra *et al.*, 2012). Results of previous researches indicated that among 12 genes under study, four genes of *CACS*, *TIP41*, *EF1A* and F-box are the most stable housekeeping genes expressed in most samples and subgroups of cucumber. TIP41 and EF1A were the most stable genes in roots whereas EF1A and CACS expressed equally in leaves of cucumber (Warzybok and Migocka, 2013). These results confirmed previous studies on housekeeping genes of cucumber that indicated genes of CACS, TIP14, F-box and EF1A have the highest stability of transcription in heavy metals stress, salt, oxidative, osmotic conditions and using growth regulators (Migocka and Papierniak, 2011). After studying 11 housekeeping genes in Jatropha curcas, it was found that three genes of ACT, EF1A and GAPDH were the most stable among genes under consideration in two growth stages of the plant. In contrast, two genes of EF1A and GAPDH were ranked differently (Zhang et al., 2013). In vegetative and reproductive stages, it was determined bv NormFinder software that ACT and GAPDH genes or EF1A and UEP genes had the most stable expressions whereas ACT and TUB8 genes or EF1A and GAPDH genes were identified by qBase as the most stable genes (Zhang et al., 2013). Previous results of Jatropha showed that there are more than one suitable housekeeping and different gene housekeeping genes should be selected based on type of samples.



Fig. 2. Melting Curve of three housekeeping genes in multi-leaf stage of Caraway (*B. persicum*), the negative derivative plat of relative fluorescent signals per derivative temperature versus temperature.

54 | Zeidabadi et al.



Fig. 3. Diagram of M-Value calculated by geNorm software for selection of the best housekeeping gene of five Caraway (*B. persicum*) ecotypes by use different stages data.

According to obtained results, the best housekeeping genes for normalization of gene expression of Jatropha are as follows: EF1A + GAPDH + ACT for all growth stages, TUB8 + ACT for vegetative stage, GAPDH + EF1A for reproductive stage, TUB5 + GAPDH + ACT for treatments of drought and cold stresses, TUB8 + TUB5 for treatment of drought stress and ACT + GAPDH for treatment of cold stress. These results contain precise information about gene expression of Jatropha based on qRT-PCR (Zhang et al., 2013). As seen in the results of housekeeping genes in Jatropha, EF1A is one of housekeeping genes that can be used in all growth stages of the plant. Another point is that software used for studying housekeeping genes shows some differences in results because they have no equal calculative basics. Data can be analyzed by several software tools, then the results can be compared. Since strategy of data normalization of the target gene expression based on only one housekeeping gene can lead to error in data normalization, it is recommended that one normalization factor is used based on mean geometry of the best housekeeping genes with highest stability and expression in order to normalize data of target gene expression (Strube et al., 2008; Vandesompele et al., 2002), in order to ensure stability of the housekeeping gene, analysis with different software tools is suggested in order to get a final conclusion for selection of the most suitable housekeeping gene.



Fig. 4. Heatmap of Housekeeping genes (UBC, EF1A and ACT) by CT.

Also it is recommended that biological and functional features of housekeeping genes are paid attention while selecting them for data normalization of target gene expression and housekeeping genes with different functions in cells are used if possible (Saele *et al.*, 2009). Therefore, genes selected in present research had good and various performances in the plant and this helps easy selection of housekeeping

gene. Totally after different analyses with several software tools, among three housekeeping genes, *EF1A* is recommended for analyses of gene expression in *Bunium persicum*.

Conclusion

Housekeeping genes chosen in conventional PCR had good performance and cDNA of B. persicum was proliferated well and three genes of UBC, EF1A and ACT were more optimal so they were selected. In order to measure transcript level of RNA in housekeeping genes, Real Time PCR was done. Results of data analyzed by Bestkeeper-1 software ranked stability of housekeeping genes based on raw values of CT. Based on results of geNorm software, the best housekeeping gene or internal control in B. persicum is EF1A that its stability Eigen value of gene expression is 0.057 that is lower than values of other genes. Plotting Heatmap based on CT (threshold cycle) values related to each gene by GenEx software confirms results analyzed by Bestkeeper-1 and geNorm software. According to this map, EF1A is expressed higher and more uniform than two other genes.

References

Andersen CL, Jensen JL, Orntoft TF. 2004. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Research **64**, 5245–5250.

Artico S, Nardeli SM, Brilhante O, Grossi-de-Sa MF, Alves-Ferreira M. 2010. Identification and evaluation of new reference genes in Gossypium hirsutum for accurate normalization of real-time quantitative RT-PCR data. BMC Plant Biology **10**, 49.

Bustin SA. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. Journal of Molecular Endocrinology **25**, 169-193.

Manager CFX. 1999. University of Chicago. All rights reserved.

Dheda K, Huggett JF, Chang JS, Kim LU, Bustin SA, Johnson MA, Rook GAW, Zumla A. 2005. The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. Analytical Biochemistry 344, 141–143.

Hellemans J, Mortier G, de Paepe A, Speleman F, Vandesompele J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biology **8**, R19.

Infante C, Matsuoka MP, Asensio E, Canavate JP, Reith M, Manchado M. 2008. Selection of housekeeping genes for gene expression studies in larvae from flatfish using real-time PCR. BMC Molecular Biology **9**, 28.

Integrated DNA Technologies, Inc. 2014. Privacy 1710 Commercial Park Coralville, IA 52241 USA.

Jalilzadeh-Amin G, Maham M, Dalir-Naghadeh B, Kheiri F. 2011. Effects of *Bunium persicum* (Boiss.) Essential oil on the contractile responses of smooth muscle (An in vitro Study). Veterinary Research Forum 2, 87-96.

Libault M, Thibivilliers S, Bilgin DD, Radwan O, Benitez M. 2008. Identification of four soybean reference genes for gene expression normalization. Plant Genome **1**, 44–54.

Lovdal T, Lillo C. 2009. Reference gene selection for quantitative real-time PCR normalization in tomato subjected to nitrogen, cold, and light stress. Anal Biochemistry **387**, 238–242.

Mafra V, Kubo KS, Alves-Ferreira M, Ribeiro-Alves M, Stuart RM. 2012. Reference genes for **McCurley AT, Callard GV.** 2008. Characterization of housekeeping genes in zebrafish: male-female differences and effects of tissue type, developmental stage and chemical treatment. BMC Molecular Biology **9**, 102.

Migocka M, Papierniak A. 2011. Identification of suitable reference genes for studying gene expression in cucumber plants subjected to abiotic stress and growth regulators. Molecular Breeding **28**, 343–357.

Mitter K, Kotoulas G, Magoulas A, Mulero V, Sepulcre P, Figueras A, Novoa B, Sarropoulou E. 2009. Evaluation of candidate reference genes for QPCR during ontogenesis and of immune-relevant tissues of European seabass (*Dicentrarchus labrax*). Comparative Biochemistry and Physiology - Part B: Biochemistry and Molecular Biology **153**, 340-347.

Multid Analyses. 2015. .se. Accessed on 18 Feb 2015. http://www.multid

Oroojalian F, Kasra-Kermanshahi R, Azizi M, Bassami MR. 2010. Synergistic antibacterial activity of the essential oils from three medicinal plants against some important food-borne pathogens by microdilution method. Iranian Journal of Medicinal and Aromatic Plants **26(2)**, 133-146.

Pfaffl MW. 2001. A new mathematical model for relative quantification in realtime RT-PCR. Nucleic Acids Research **29(9)**, e45.

Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper - Excel-based tool using pair-wise correlations. Biotechnology Letters **26**, 509-515. Sæle O, Nordgreen A, Hamre K, Olsvik PA. 2009. Evaluation of candidate reference genes in Q-PCR studies of Atlantic cod (*Gadus morhua*) ontogeny, with emphasis on the gastrointestinal tract. Comparative Biochemistry and Physiology - Part B: Biochemistry and Molecular Biology **152**, 94–101.

Shamsi-Fard MH, Mirzaghaderi G, Majdi M. 2014. Transcript expression analysis of geranyl diphosphate synthase gene in different tissues of black cumin (*Nigella sativa* L.). Iranian Journal of Rangelands and Forests Plant Breeding and Genetic Research **22(2)**, 143-155.

Small BC, Murdock CA, Bilodeau-Bourgeois AL, Peterson BC, Waldbieser GC. 2008. Stability of reference genes for real-time PCR analyses in channel catfish (*Ictalurus punctatus*) tissues under varying physiological conditions. Comparative Biochemistry and Physiology - Part B: Biochemistry and Molecular Biology **151**, 296-304.

Sofi PA, Zeerak NA, Singh P. 2009. Kala zeera (*Bunium persicum* Bioss.): a Kashmirian high value crop. Turkish Journal of Biology **33**, 249-258.

Strube C, Buschbaum S, Wolken S, Schnieder T. 2008. Evaluation of reference genes for quantitative real-time PCR to investigate protein disulfide isomerase transcription pattern in the bovine lungworm *Dictyocaulus viviparus*. Gene **425**, 36-43.

Tang RY, Dodd A, Lai D, McNabb WC, Love DR. 2007. Validation of zebrafish (*Danio rerio*) reference genes for quantitative real-time RT-PCR normalization. Acta Biochimica et Biophysica Sinica **39**, 384-390.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology **3**, 1-12. Wan H, Zhao Z, Qian C, Sui Y, Malik AA. 2010. Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. Anal Biochemistry **399**, 257–261.

Warzybok A, Migocka M. 2013. Reliable reference renes for normalization of gene expression in cucumber grown under different nitrogen nutrition. Plos One **8(9)**, 1-11.

Yousefzadi M, Sharifi M, Behmanesh M, Ghasempour A, Moyano E, Palazon J. 2012. The effect of light on gene expression and podophyllotoxin biosynthesis in Linum album cell culture. Plant Physiology Biochemistry **56**, 41-46.

Zhang L, He LL, Fu QT, Xu ZF. 2013. Selection of Reliable Reference Genes for Gene Expression Studies in the Biofuel Plant *Jatropha curcas* Using Real-Time Quantitative PCR. International Journal of Molecular Sciences **14**, 24338-24354.

Zhong Q, Zhang Q, Wang Z, Qi J, Chen Y, Li S, Sun Y, Li C, Lan X. 2008. Expression profiling and validation of potential reference genes during *Paralichthys olivaceus* embryogenesis. Marine Biotechnology **10**, 310–318.