한작지(Korean J. Crop Sci.), 55(4): 371~378(2010)

Identification of the Housekeeping Genes Using Cross Experiments via *in silico* Analysis

Won Cheol Yim*, Changwon Keum**, Saehwan Kim***, Cheol Seong Jang****, and Byung-Moo Lee*†

*Department of Plant Biotechnology, Dongguk Univ.-Seoul, Seoul, Korea
**Bioinformatics and Molecular Design Research Center (BMDRC), Seoul, Korea
***Macrogen Inc, Seoul, Korea
****Division of Bio-resources Technology, Kangwon National University, Korea

ABSTRACT For sensitive and accurate gene expression analysis, normalization of gene expression data against housekeeping genes is required. There are conventional housekeeping gene (e.g. ACT) that primarily function as an internal control of transcription. In this study, we performed an *in silico* analysis of 278 rice gene expression samples (GSM) in order to identify the gene that is most consistently expressed. Based on this analysis, we identified novel candidate housekeeping genes that displayed improved stability among the cross experimental conditions. Furthermore four of the most conventional housekeeping genes were included in our 30 other housekeeping genes among the most stable genes. Therefore, these 30 genes can be used to normalize transcription results in gene expression studies on rice at a broad range of experimental conditions.

Keywords : housekeeping genes; normalization; GEO; microarray

Knowing the complete genome sequences of model plants including rice (Sequencing Project International Rice Genome, 2005) and *Arabidopsis* (Arabidopsis Genome Initiative, 2000) has paved the way for us to carry out global analyses to gain a more complete understanding of the information in their genome. Indeed, rice microarray dataset using cDNA or oligonucleotide chips has allowed for the simultaneous quantification of tens of thousands of gene transcripts. This wealth of information is now in repository public databases, such as the gene expression omnibus (GEO; Edgar *et al.*, 2002) in NCBI. Advanced molecular techniques, such as quantitative reverse transcription (qRT)-PCR, which is the most sensitive method for the detection of low abundance of mRNA, have been used to measure the expression of a small set of genes (Bustin *et al.*, 2005). Moreover, RT-PCR can be used in different application such as clinical diagnostic (Bustain and Dorudi, 1998) for the analysis of tissue-specific gene expression (Gachon *et al.*, 2004), and for plant studies (Czechowski *et al.*, 2004). With recent advances in microarray technology, real-time RT-PCR can now be used to validate microarray gene expression data (Gachon *et al.*, 2004; Czechowski *et al.*, 2004; Chuaqui *et al.*, 2002). Real-time RT-PCR measures the expression of genes relative to the expression of a reference internal control gene, which need to be constitutively express regardless of the experimental conditions.

Ideally, the conditions of the experiments should not influence the expression of this internal control gene (Schmittgen and Zakrajsek, 2000). The most commonly used reference transcripts for normalization in plants and animals include 18s rRNA, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), elongation factor-1a (EF-1a), polyubiquitin (UBQ), actin (ACT), and a-tubulin and b-tubulin (TUA and TUB, respectively) genes (Goidin *et al.*, 2001; Bustin, 2002; Kim *et al.*, 2003; Andersen *et al.*, 2004; Brunner *et al.*, 2004; Dheda *et al.*, 2004; Radonić *et al.*, 2004). These genes were known or suspected housekeeping genes, which play important roles in basic cellular processes such as those involving cell structure and primary metabolism, in the pre-genomic era. Consequently, they are often referred to as housekeeping genes and have been used as internal controls of transcription in expression studies. Despite their wide spread use as internal controls, a number of studies have shown that their transcription and expression levels are not always stable (Schmittgen and Zakrajsek, 2000;

†Corresponding author: (Phone) +82-2-2260-3307
(E-mail) bmlee@dongguk.edu <Received November 12, 2010>

Thellin *et al.*, 1999; Suzuki *et al.*, 2000; Lee *et al.*, 2002; Zhong *et al.*, 1999)

Rice is the monocot model plant used in genetic and molecular studies. Due to the availability of complete genome (Sequencing Project International Rice Genome, 2005), it has been possible to conduct genome-wide studies. In addition, expression studies have become extremely important in regards to providing a better understanding of the information contained within the rice genome. In this study, we used a large set of expression data from 278 published rice gene expression profiles, to determine the stability of expression of commonly used housekeeping genes and to identify novel reference genes with significantly more stable expression levels. This result was to identify a large set of housekeeping genes that together cover a wide range of absolute expression levels for a wide range of different experimental conditions (e.g. qRT-PCR, etc).

MATERIALS AND METHODS

Rice gene expression datasets

To evaluate global expression profiles of rice genes of interest, we employed the gene expression omnibus (GEO) in NCBI (Edgar *et al.*, 2002). To remove potential errors from cross-platform variation, we only used gene expression datasets obtained with the Affymetrix Rice Genome Array, GEO platform (GPL) ID: GPL2025, targeted for *Oryza sativa*. We filtered out two GEO series (GSE), GSE6737 and GSE15071, since GSE6737 is a redundant super series containing two other GSE, GSE6719 and GSE6720, and GSE15071 is a set of genome variation profiles rather than gene expression profiles. A total of 278 rice gene expression samples (GSM) were selected, resulting in 23 GSE datasets. (Detailed information of all GSE and GSM datasets is described in Table 1).

Table 1. The 23 GSE series of gene expression profile datasets.

Series	Title	PMID
GSE7256	Identification of rice genes differentially expressed upon virulent infection by Magnaporthe grisea	17905473
GSE6893	Expression data for reproductive development in rice	17293439
GSE4471	Expression data from rice varieties Azucena and Bala grown in 0 and 1ppm arsenate	
GSE10054	Expression information of splicing factor OsSKIPa knock-down and overexpressed rice	
GSE14275	Expression data for heat shock in rice seedlings	
GSE6908	Transcript Profiling of the Aerobic and Anoxic Rice Coleoptile	17369434
GSE11025	Comparative transcriptional profiling of two contrasting rice genotypes in response to rice stripe virus infection	
GSE6719	Cytokinin responsive genes in rice	17293362
GSE10872	Molecular characterization and genetic analysis reveal SA0420 as an early senescing rice mutant with pleotropic	
GSE4438	Expression data from rice under salinity stress	17160619
GSE6901	Expression data for stress treatment in rice seedlings	17293439
GSE10857	Gene expression of rice root tips before, at and buckled by a hard layer in two rice varieties	
GSE12097	antiOsLIC collar chip	18953406
GSE7197	Down-regulation of OsSRT1 induces DNA fragmentation and cell death in rice	17468215
GSE3053	Rice salt expression	16183841
GSE12069	Mycroarray analyses reveal that plant mutagenesis may induce more transcriptomic changes than transgene insertion	18303117
GSE8380	Rice Gene Network Inferred from Expression Profiling of Plants Overexpressing OsWRKY13	
GSE7951	Genome-wide gene expression profiling of rice stigma	17556504
GSE6720	Gene expression in OsRR6-overexpression line	17293362
GSE11966	Expression data from rice embryo, endosperm, root, leaf and seedling	
GSE10373	Rice cultivars undergoing a susceptible and resistant interaction with the parasitic plant Striga hermonthica	18507775
GSE9498	Global gene expression profiles of Oryza sativa wild type Zhonghua11 and mutant gif1 in filling stage	18820698

Data preprocessing

Affymetrix GeneChip spot IDs matched to the TIGR locus IDs, which we considered one unit of an individual gene. Each probe, targeted to more than one gene, was filtered out while intensity values of two or more probes, corresponding to one gene, were averaged. A total of 57,381 probe sets were transformed into 34,031 genes. To normalize gene expression samples across the different GEO series, we converted gene expression intensity to gene expression rank by ordering genes according to their raw intensities. Therefore, the highest expressed gene in each sample was assigned a rank value of 1 and the lowest expressed gene was assigned a rank value of 34,031.

Housekeeping gene identification

We first calculated the standard deviation of the gene rank in each GEO dataset to measure the expression variation of a gene in each GEO dataset. Then, we calculated the average standard deviation of each gene across the whole GEO dataset. Finally, genes with the lowest average standard deviations (SD) were selected as candidate rice housekeeping genes.

RESULTS & DISCUSSION

We attempted to indentify Rice housekeeping genes based on public gene expression datasets at various biological conditions and tissue types. Similarly, housekeeping genes were selected in the human genome (Jonge *et al.*, 2007), in which all gene expression samples were normalized for further analysis by quantile normalization. However, gene expression samples from each series displayed their own technical and experimental variations (Rhodes *et al.*, 2004); thus, normalization of all samples from different series could lead to erroneous and ambiguous results. To avoid these potential problems, we normalized the gene expression samples across different data series using a gene rank in each sample rather than the raw intensity.

The top 30 genes that were constitutively expressed across whole datasets are summarized in the list (Table 2). The top 30 ranked housekeeping genes on average had a gene rank ranging from 4.23 to 375.22, with an average rank standard deviation that varied from 1.89 to 65.18 across the datasets. The gene rank of the top ranked housekeeping genes was relatively consistent in each sample and as was the gene rank variation across the dataset compared to the non housekeeping genes (Fig. 1). The top 30 ranked candidate housekeeping genes include the RNA recognition motif containing protein (LOC_Os12g43600, LOC_Os03g46770), coiled-coil domain-containing protein 72 (LOC_Os06g47230), translationally-controlled tumor protein (LOC_Os11g43900), endothelial differentiation-related factor 1 (LOC_Os11g43900), etc. This list clearly reveals novel housekeeping genes that displayed a more stable expression under a broad range of experimental conditions relative to the conventional housekeeping genes. Interestingly, the conventional housekeeping genes, 18s rRNA and a-tubulin and b-tubulin (TUA and TUB, respectively) were not identified as one of the top ranked housekeeping genes. However, elongation factor-1a (EF-1a, LOC_Os02g32030), polyubiquitin (UBQ, LOC_Os06g46770), glyceraldehyde-3-phosphate dehydrogenase (GAPDH, LOC_Os08g03290) and actin (ACT, LOC_Os11g06390) were identified among the top 30 ranked candidates. Two of the RNA recognition motif (RRM) containing proteins (LOC_Os12g43600, LOC_Os03g46770) had the least variation in this study. The function of the RRM protein is still unknown. However, these proteins, which are involved in the most essential processes of post-transcriptional gene regulation, are conserved in plants (Lorkovic *et al.*, 2002). The rice genome encodes 170 RRM-containing proteins based on the TIGR rice annotation release 6 (http://rice.plantbiology.msu.edu/). These genes are involved in post-transcriptional regulation of gene expression and their expression is not tissue specific and stress independent. The translationally-controlled tumor protein (TCTP) (LOC_Os11g43900, 4th) was expressed with minimal variations. The rice genome contains only one TCTP gene, which was originally identified as a tumor protein in mouse ascetic tumor and in mouse erythroleukemia (Yenofsky *et al.*, 1983; Chitpatima *et al.*, 1988). The sequence of the TCTP gene was highly conserved during evolution and has been shown to play an essential role in the development and normal function of various organisms (Thaw *et al.*, 2001; Thiele *et al.*, 2000; Rao *et al.*, 2002). Interestingly, most studies reported that expression of TCTP was highly regulated in response to a wide range of extracellular signals and cellular conditions (Bommer and Thiele 2004). The high degree of homology from plants to man and its expression

Table 2. Top 30 housekeeping genes ordered by average standard deviation.

MSU locus*	Avr.SD	SD of SD	Avr.Rank	Annotation
LOC_Os12g43600	1.89	1.82	4.23	RNA recognition motif containing protein
LOC_Os03g46770	7.11	5.41	22.11	RNA recognition motif containing protein
LOC_Os06g47230	8.96	10.11	24.95	coiled-coil domain-containing protein 72
LOC_Os11g43900	13.87	12.17	48.73	translationally-controlled tumor protein
LOC_Os08g27850	15.01	9.7	45.24	endothelial differentiation-related factor 1
LOC_Os05g41060	15.17	11.89	38.23	ADP-ribosylation factor
LOC_Os06g05880	16.98	16.96	56.98	profilin domain containing protein
LOC_Os08g03290	22.81	19.93	51.88	glyceraldehyde-3-phosphate dehydrogenase
LOC_Os07g34589	27.6	34.74	54.1	translation initiation factor SUI1
LOC_Os06g46770	30.59	30.38	83.19	ubiquitin family protein
LOC_Os11g23854	33.4	30.81	120.43	4F5 protein family protein
LOC_Os12g38000	36.58	33.33	120.6	60S ribosomal protein L8
LOC_Os06g04030	37.09	29.47	108.33	histone H3
LOC_Os02g32030	37.68	31.02	141.62	elongation factor
LOC_Os07g08840	38.96	31.16	131.06	thioredoxin
LOC_Os06g51220	40.37	51.83	79.47	HMG1/2
LOC_Os04g53620	46.25	29.73	161.08	ubiquitin family protein
LOC_Os03g50290	46.78	71.42	117.98	14-3-3 protein
LOC_Os03g60590	46.96	40.56	203.62	actin-depolymerizing factor
LOC_Os10g33800	50.65	51.08	127.56	lactate/malate dehydrogenase
LOC_Os05g49890	51.42	44.23	110.52	ras-related protein
LOC_Os01g60410	51.64	38.4	208.51	ubiquitin-conjugating enzyme
LOC_Os06g41010	51.8	43.59	124.69	zinc finger A20 and AN1 domain-containing stress-associated protein
LOC_Os11g06390	57.64	84.53	169.58	actin
LOC_Os02g02890	58.57	68.3	375.22	peptidyl-prolyl cis-trans isomerase
LOC_Os05g38520	58.79	48.89	206.93	60S ribosomal protein L36-2
LOC_Os09g30412	59.71	41.41	184.24	heat shock protein
LOC_Os05g41900	62.52	44.31	238	translation initiation factor SUI1
LOC_Os07g32800	63.65	53.23	244.04	autophagy-related protein
LOC_Os11g21990	65.18	49.01	300.08	expressed protein

*MSU indicates Michigan State University Rice pseudomolecule locus ID.

in many tissues suggests that TCTP most likely has a cell housekeeping function (Sanchez *et al.*, 1997). Cans *et al.* (2003) have shown that TCTP preferentially stabilizes the GDP form of the translation elongation factor eEF1A, and impairs the GDP exchange reaction promoted by its guanine nucleotide exchange factor eEF1Bbeta, which implicates TCTP in the elongation step of protein synthesis. Hence, the traditional housekeeping gene, Elongation factor (LOC_Os02g 32030, 14th) and TCTP were significantly related due ot their constant expression levels.

Traditionally, housekeeping genes have been selected by considering the gene function in a cellular context. Actin, polyubiquitin and tubulin are considered representative housekeeping genes on account of their roles in fundamental biological processes. However, there is more than one gene that has the same function. In the case of Actin, there are

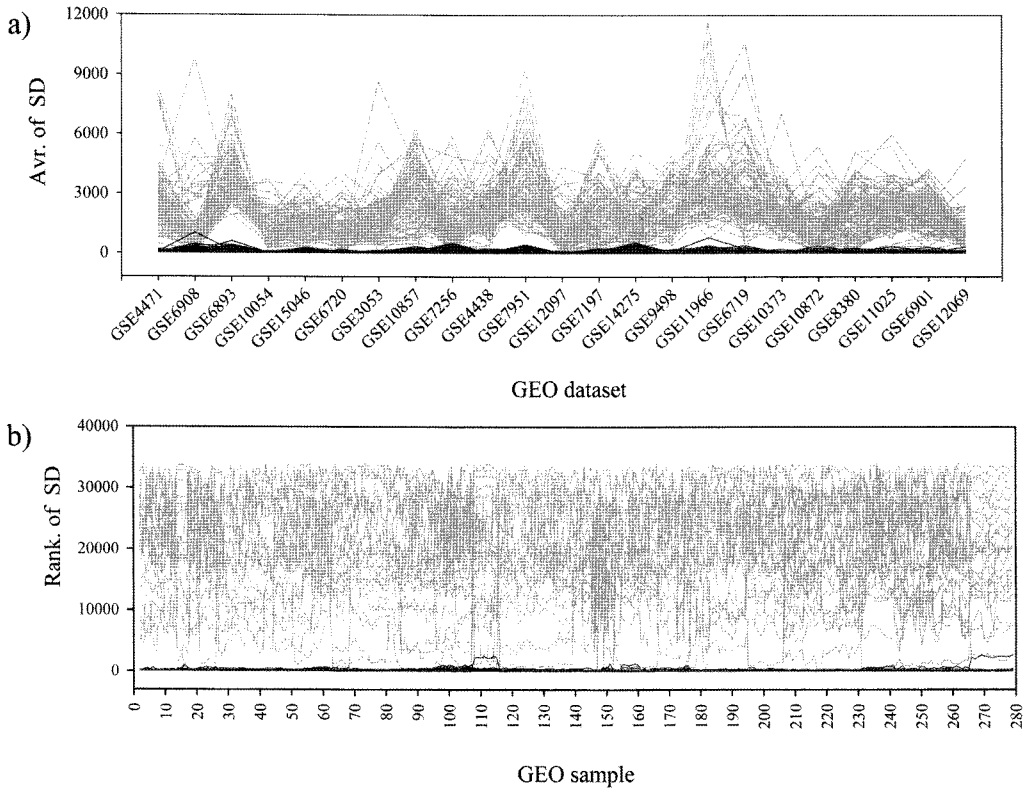

Fig. 1. Plots of the standard deviation within each GEO series. a) top 100 housekeeping genes (red) and 100 non-housekeeping genes (gray, ranked 15,000 to 15,100) in each GSE according to the average standard deviation. b) top 30 housekeeping genes (red) and 100 non-housekeeping genes (gray, ranked 15,000 to 15,030) in each GSM according to the average standard deviation. For x-axis descriptions in detail see the Table 1.

14 genes which have the same function (Table 3). The average SD for these genes ranged from 57.64 to 3170.06, indicating that expression of these genes varied substantially. Actin (LOC_Os03g61970), which was ranked in 1490th, is typically used as an internal control. However, actin (LOC_Os11g06390) was ranked 24th, thus it would be a better housekeeping gene than actin (LOC_Os03g61970). Likewise, the ubiquitin family protein (LOC_Os06g46770), which was ranked 10th, would be a more suited housekeeping gene than the ubiquitin family protein (LOC_Os02g06640). Figure 2 clearly shows that the top ranked genes had smaller perturbations in the average SD than the other genes examined in this expression dataset series. Thus, we believe the results obtained in this study can be used as a basis to select more stable housekeeping genes relative to the genes currently being used as internal controls.

In prior studies, housekeeping genes were selected based on a small number of expression datasets (Zhong *et al.*, 1999; Jain, 2009). Thus, these selected housekeeping genes only displayed consistent expression under certain experimental conditions. Condition specific housekeeping genes showed low expression variation at a given the condition, but the overall variation in expression was high. The rank of the expressed protein (LOC_Os12g41640) in GSE6908, gene expression dataset for Aerobic and Anoxic Rice Coleoptile, based on SD was only 1.5. However, its average rank based on standard deviation across all 23 GEO datasets was 2,135.91 with an average sample rank of 15,852. Since we used all available gene expression datasets at unlimited biological conditions in the selection of housekeeping genes, we could avoid this condition specific housekeeping gene selection problem. The top 3 ranked genes were condition non-specific housekeeping genes since their average standard deviations were lower than 8.96 and they were ranked at most 173th in each gene expression dataset.

We demonstrated that housekeeping genes selected using

Table 3. Examples of commonly used housekeeping genes.

Annotation	MSU locus*	Avr.SD	SD of SD	Avr. Rank	Rank in list
actin	LOC_Os11g06390	57.64	84.53	169.58	24
	LOC_Os05g01600	140.57	147.61	824.74	137
	LOC_Os01g64630	213.31	399.47	534.72	302
	LOC_Os03g56970	385.43	266.58	2,808.03	891
	LOC_Os08g04280	445.88	380.18	2,083.99	1147
	LOC_Os03g61970	515.44	389.6	3,859.18	1490
	LOC_Os10g36650	540.13	501.62	1,357.22	1636
	LOC_Os01g16414	583.18	331.35	10,434.18	1986
	LOC_Os08g28190	671.46	455.65	5,931.17	2748
	LOC_Os01g73310	739.81	817.39	3,361.68	3510
	LOC_Os04g09860	976.9	763.09	11,437.91	6235
	LOC_Os05g36290	1,212.94	1,435.86	3,905.85	8529
	LOC_Os02g38340	1,293.37	617.16	8,959.89	9177
	LOC_Os12g06660	3,170.06	1,709.12	13,455.72	28049
ubiquitin family protein	LOC_Os06g46770	30.59	30.38	83.19	10
	LOC_Os04g53620	46.25	29.73	161.08	17
	LOC_Os02g06640	65.58	66.47	305.93	31
	LOC_Os05g42424	90.38	90.61	392.75	63
	LOC_Os09g25320	289.98	266.09	1,008.45	545
	LOC_Os03g24920	484.01	314.17	3,047.49	1305
	LOC_Os10g39620	557.66	347.88	5,357.25	1779
	LOC_Os10g31790	639.46	392.58	4,986.17	2461
	LOC_Os06g44080	723.7	654.17	3,108.27	3305
	LOC_Os08g19830	752.34	1,815.85	2,616.78	3676
	LOC_Os07g31540	760.02	375.34	8,739.90	3776
	LOC_Os05g28500	1,035.07	884.53	10,373.43	6885
	LOC_Os08g08760	1,772.58	929.96	19,566.84	12571
	LOC_Os01g53100	1,999.41	868.11	18,260.22	14334
	LOC_Os10g33620	2,445.88	1,677.74	14,147.84	18572
	LOC_Os10g34960	2,498.15	1,366.32	27,751.32	19169
	LOC_Os01g45420	2,525.57	1,861.84	29,320.43	19475
	LOC_Os01g67950	2,568.20	2,097.77	8,807.52	19933
	LOC_Os09g31031	2,732.30	1,249.65	10,999.23	21771
	LOC_Os02g38410	2,861.44	2,167.05	15,979.90	23372
	LOC_Os08g08700	3,275.63	1,871.03	21,743.24	29627

*MSU indicates Michigan State University Rice pseudomolecule locus ID.

gene expression datasets acquired at various biological conditions and sample types show more consistent expression than commonly used housekeeping genes or candidate housekeeping genes based on a single gene expression dataset (Jain, 2009). Since the candidate housekeeping genes, identified here, were selected based on gene expression datasets acquired under various laboratory settings and experimental conditions, including various stress conditions and tissue types, they could be used as reference housekeeping genes for various applications in common laboratory experiments such as RT-PCR.

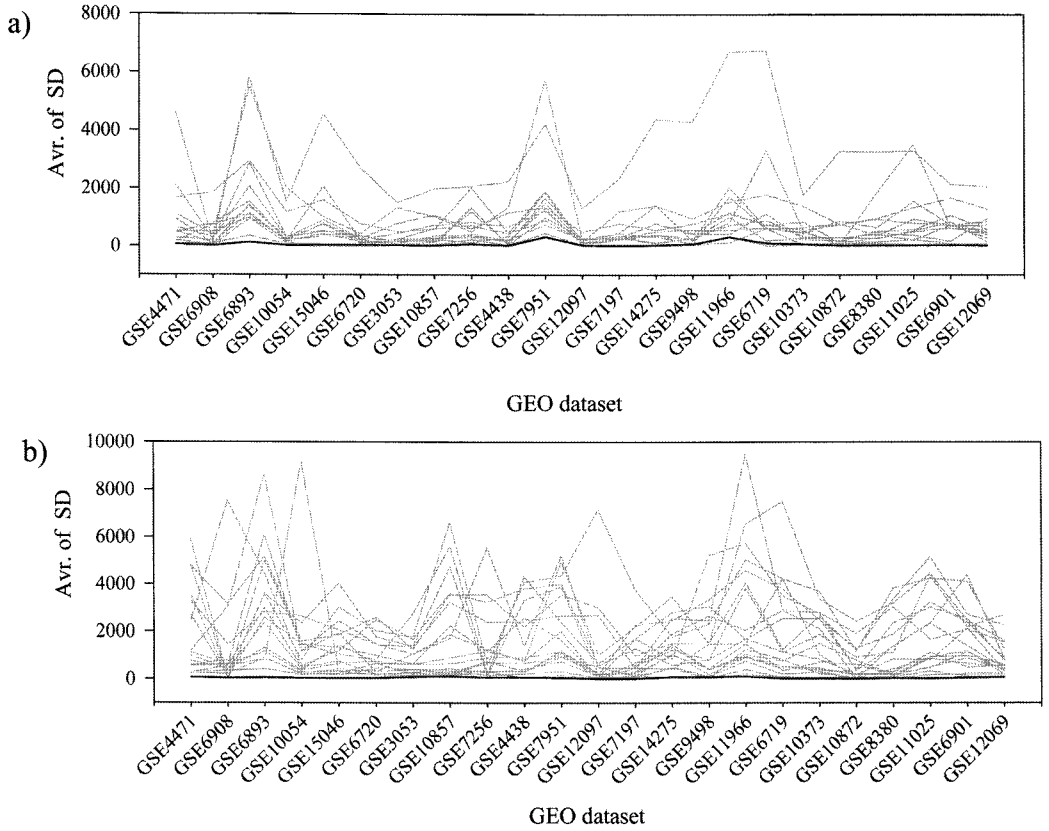

Fig. 2. Plots of the standard deviation within each GEO series. a) actin as housekeeping genes (gray) and the highest ranked actin (red, ranked 24) in each GSE according to the average standard deviation. b) ubiquitin family protein as housekeeping genes (gray) and the highest ranked ubiquitin family protein (red, ranked 10) in each GSE according to the average standard deviation. For x-axis descriptions in detail see the Table 1.

ACKNOWLEDGEMENT

This work was supported by a grant (No. 20070301-034-016-009-01-00) from BioGreen 21 Program, Rural Development Administration, Republic of Korea.

REFERENCES

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408: 796-815.

Andersen, C. L., J.L. Jensen and T.F. Ørntoft 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 Res. 64: 5245-5250.

Bommer, U. and B. Thiele. 2004. The translationally controlled tumour protein (TCTP). The International Journal of Biochemistry & Cell Biology 36: 379-385.

Brunner, A., I. Yakovlev and S. Strauss. 2004. Validating internal controls for quantitative plant gene expression studies. BMC Plant Biology 4: 14.

Bustin, S. A., V. Benes, T. Nolan, and M. W. Pfaffl. 2005. Quantitative real-time RT-PCR--a perspective. J. Mol. Endocrinol 34: 597-601.

Bustin, S. A. and S. Dorudi. 1998. Molecular assessment of tumour stage and disease recurrence using PCR-based assays. Mol. Med. Today 4: 389-396.

Bustin, S. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J. Mol. Endocrinol 25: 169-193.

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

Cans, C., B. J. Passer, V. Shalak, V. Nancy-Portebois, V. Crible, N. Amzallag, D. Allanic, R. Tufino, M. Argentini, D. Moras, G. Fiucci, B. Goud, M. Mirande, R. Amson and A. Telerman. 2003. Translationally controlled tumor protein acts as a guanine nucleotide dissociation inhibitor on the

translation elongation factor eEF1A. Proc. Natl. Acad. Sci. U.S.A 100: 13892-13897.

Chitpatima, S. T., S. Makrides, R. Bandyopadhyay and G. Brawerman, 1988. Nucleotide sequence of a major messenger RNA for a 21 kilodalton polypeptide that is under translational control in mouse tumor cells. Nucleic Acids Res. 16: 2350.

Chuaqui, R. F., R. F. Bonner, C. J. M. Best, J. W. Gillespie, M. J. Flaig, S. M. Hewitt, J. L. Phillips, D. B. Krizman, M. A. Tangrea, M. Ahram, W. M. Linehan, V. Knezevic and M. R. Emmert-Buck. 2002. Post-analysis follow-up and validation of microarray experiments. Nat. Genet. 32: 509-514.

Czechowski, T., R. P. Bari, M. Stitt, W. Scheible, and M. K. Udvardi. 2004. Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. Plant J. 38: 366-379.

Dheda, K., J. F. Huggett, S. A. Bustin, M. A. Johnson, G. Rook and A. Zumla. 2004. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. Bio-Techniques 37: 112-114, 116, 118-119.

Edgar, R., M. Domrachev and A. E. Lash. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30: 207-210.

Gachon, C., A. Mingam and B. Charrier. 2004. Real-time PCR: what relevance to plant studies? Journal of Experimental Botany 55: 1445-1454.

Goidin, D., A. Mamessier, M. J. Staquet, D. Schmitt and O. Berthier-Vergnes. 2001. Ribosomal 18S RNA prevails over glyceraldehyde-3-phosphate dehydrogenase and beta-actin genes as internal standard for quantitative comparison of mRNA levels in invasive and noninvasive human melanoma cell subpopulations. Anal. Biochem. 295: 17-21.

Jain, M. 2009. Genome-wide identification of novel internal control genes for normalization of gene expression during various stages of development in rice. Plant Science 176: 702-706.

de Jonge, H.J.M., R. S. N. Fehrmann, E. S. J. M. de Bont, R. M. W. Hofstra, F. Gerbens, W. A. Kamps, E. G. E. de Vries, A. G. J. van der Zee, G. J. te Meerman and A. ter Elst. 2007. Evidence Based Selection of Housekeeping Genes. PLoS ONE 2: e898.

Kim, B., H. Nam, S. Kim, S. Kim, Y. Chang. 2003. Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. Biotechnol. Lett. 25: 1869-1872.

Lee, P.D., R. Sladek, C. M. Greenwood, T. J. Hudson. 2002. Control Genes and Variability: Absence of Ubiquitous Reference Transcripts in Diverse Mammalian Expression Studies. Genome Research 12: 292-297.

Lorkovic, Z. J. and A. Barta. 2002. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. Nucleic Acids Res 30: 623-635.

Radonic, A., S. Thulke, I. M. Mackay, O. Landt, W. Siegert and A. Nitsche. 2004. Guideline to reference gene selection for quantitative real-time PCR. Biochem. Biophys. Res. Commun 313: 856-862.

Rao, K.V.N., L. Chen, M. Gnanasekar, and K. Ramaswamy. 2002. Cloning and characterization of a calcium-binding, histamine-releasing protein from Schistosoma mansoni. J. Biol. Chem. 277: 31207-31213.

Rhodes, D.R., J. Yu, K. Shanker, N. Deshpande, R. Varambally, D. Ghosh, T. Barrette, A. Pandey and A. M. Chinnaiyan. 2004. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proceedings of the Proc. Natl. Acad. Sci. U.S.A 101: 9309-9314.

Sanchez, J.C., D. Schaller, F. Ravier. O. Golaz, S. Jaccoud, M. Belet, M. R. Wilkins, R. James, J. Deshusses, and D. Hochstrasser. 1997. Translationally controlled tumor protein: a protein identified in several nontumoral cells including erythrocytes. Electrophoresis 18: 150-155.

Schmittgen, T.D. and B. A. Zakrajsek. 2000. Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. Journal of Biochemical and Biophysical Methods 46: 69-81.

Sequencing ProjectInternational Rice Genome. 2005. The map-based sequence of the rice genome. Nature 436: 793-800.

Suzuki, T., P. Higgins and D. Crawford. 2000. Control selection for RNA quantitation. Biotechniques 29: 332-337.

Thaw, P., N. J. Baxter, A. M. Hounslow, C. Price, J. P. Waltho and C. J. Craven, C.J., 2001. Structure of TCTP reveals unexpected relationship with guanine nucleotide-free chaperones. Nat. Struct. Mol. Biol. 8:701-704.

Thellin, O., W. Zorzi, B. Lakaye, B. De Borman, B. Coumans, G. Hennen, T. Grisar, A. Igout, and E. Heinen. 1999. Housekeeping genes as internal standards: use and limits. Journal of Biotechnology 75: 291-295.

Thiele, H., M. Berger, A. Skalweit and B. Thiele. 2003. Expression of the gene and processed pseudogenes encoding the human and rabbit translationally controlled tumour protein (TCTP). European Journal of Biochemistry 267: 5473-5481.

Yenofsky, R., S. Cereghini, A. Krowczynska and G. Brawerman. 1983. Regulation of mRNA utilization in mouse erythroleukemia cells induced to differentiate by exposure to dimethyl sulfoxide. Mol. Cell. Biol 3: 1197-1203.

Zhong, H. and J. W. Simons. 1999. Direct comparison of GAPDH, beta-actin, cyclophilin, and 28S rRNA as internal standards for quantifying RNA levels under hypoxia. Biochem. Biophys. Res. Commun 259: 523-526.