

Received:
23 September 2016
Revised:
16 December 2016
Accepted:
9 January 2017

Heliyon 3 (2017) e00233



Analysis of variance, normal quantile-quantile correlation and effective expression support of pooled expression ratio of reference genes for defining expression stability

Himanshu Priyadarshi^{a,*}, Rekha Das^b, Shivendra Kumar^c, Pankaj Kishore^d,
Sujit Kumar^e

^a College of Fisheries (CAU-Imphal), Lembucherra, Tripura 799210, India

^b ICAR RC for NEH Region, Tripura 799210, India

^c Dr. Rajendra Prasad Central Agricultural University, Samastipur 848125, India

^d ICAR-Central Institute of Fisheries Technology, Cochin 682029, India

^e ICAR-Central Institute of Fisheries Education, Mumbai 400061, India

*Corresponding author at: Department of Fish Genetics and Reproduction, College of Fisheries (CAU-Imphal), Lembucherra, Tripura 799210, India.

E-mail address: priyadarshimanshu@gmail.com (H. Priyadarshi).

Abstract

Identification of a reference gene unaffected by the experimental conditions is obligatory for accurate measurement of gene expression through relative quantification. Most existing methods directly analyze variability in crossing point (Cp) values of reference genes and fail to account for template-independent factors that affect Cp values in their estimates. We describe the use of three simple statistical methods namely analysis of variance (ANOVA), normal quantile-quantile correlation (NQQC) and effective expression support (EES), on pooled expression ratios of reference genes in a panel to overcome this issue. The pooling of expression ratios across the genes in the panel nullify the sample specific effects uniformly affecting all genes that are falsely reflected as instability. Our methods also offer the flexibility to

include sample specific PCR efficiencies in estimations, when available, for improved accuracy. Additionally, we describe a correction factor from the ANOVA method to correct the relative fold change of a target gene if no truly stable reference gene could be found in the analyzed panel. The analysis is described on a synthetic data set to simplify the explanation of the statistical treatment of data.

Keywords: Bioinformatics, Mathematical biosciences

1. Introduction

Relative quantification of gene expression has been accepted as the gold standard method in several modern techniques such as quantitative real-time PCR, microarray, and high-throughput sequencing (Grozinger et al., 2007; Wang et al., 2012a; Wang et al., 2012b; Paria et al., 2013; Das et al., 2015; Van Hoeck et al., 2015) whose accuracy depends largely on the stability of the reference gene employed for normalization (Bustin, 2002; Leelatanawit et al., 2012). Ideally, to qualify as stable, a reference gene should express at a constant level in each sample across the groups/categories (treatments, tissues, developmental stages, etc). Unfortunately, such ideal reference genes are rare (Vandesompele et al., 2002; Pfaffl, 2004; Chandna et al., 2012). Additionally, factors like RNA isolation, quantification, integrity, efficiency of cDNA synthesis etc. varies across samples, batches and tissues leading to a difference between machine-read expression level of a reference gene and its original expression level (Livak and Schmittgen, 2001; Bustin, 2002). Alternately this also becomes a reason for variation in expression level of a reference gene between different samples/groups/categories which is then reflected as its instability. Instability due to this reason is not a true instability, because this is equally likely to affect both reference and target genes expression read from the same sample and will nullify in relative quantification procedure (Pfaffl et al., 2004), provided both reference and target genes have same PCR efficiency (Livak and Schmittgen, 2001). For the sake of brevity, hereafter we shall refer to the variation in expression not arising as a result of the experimental treatment as “false instability”. Hence, to achieve high precision in the measurements of a gene expression, an in-depth analysis of reference gene stability pertaining to a specific experiment is required.

Variation in the Cp value between different samples for a reference gene has directly been used to evaluate its expression stability (Pfaffl et al., 2004). However, there are fair chances that a reference gene, which is actually up/down-regulated in a particular group/category/sample in an experiment, reflect Cp value equal to rest of the group/category/sample, or the other way round, because of false instability. To circumvent such problems, methods like ΔC_t (Silver et al., 2006) and geNorm (Vandesompele et al., 2002) have been developed, which compare each reference gene against the rest of the reference genes in a panel. These methods work quite

well when dealing with a panel comprising fairly large number of reference genes. However their precision reduces in a small panel comprised of reference genes of divergent nature (Silver et al., 2006). It is always advised to evaluate a large number of reference genes and use at least two for the normalization of a target gene (Vandesompele et al., 2002; Andersen et al., 2004; Silver et al., 2006; Chandna et al., 2012; Leelatanawit et al., 2012; Xie et al., 2012; Ling et al., 2014; Taki and Zhang, 2013; Guo et al., 2014; Hildyard and Wells, 2014). Selecting and employing more than one reference gene overburdens a researcher with limited funds and reference gene sequence information. In this condition the researchers are practically forced to normalize a target gene with any one of the popularly used reference genes. This is very much apparent from most of the literature (Paria et al., 2013; Das et al., 2015; Sharabi et al., 2015). Hence, there is a desperate need for a method, with good sensitivity in both large and small panels to identify a stable reference gene and provide possible correction measures, if no reference gene in the panel exhibits stable expression.

PCR efficiency (E) is another factor that affect Cp value and due weightage is given to this parameter during normalization of a target gene with a reference gene (Pfaffl, 2004). However differential PCR efficiency that affects Cp value within a reference gene itself is often overlooked leading to erroneous evaluation of stability. A reference gene with PCR efficiency less than 2 will have higher variation in its Cp values than a gene with PCR efficiency of 2 (Pfaffl, 2004). This variability in Cp value is another source of false instability. Additionally, for each reference gene PCR efficiency may differ from sample to sample due to variation in impurity or PCR inhibitory compound levels (Wilson, 1997; Guescini et al., 2008), which certainly will affect Cp value and the stability measure of a reference gene. A method to measure PCR efficiency of each sample during quantitative real-time PCR has been developed (Tichopad et al., 2003). Hence, a kind of stability evaluation method is needed, which can account for PCR efficiency and nullify the possible variation arising due to the differences in PCR efficiency.

Statistical methods developed so far for evaluation of the reference gene stability are based on real experimental data. This certainly is advantageous; however, the complexity of real life data does not allow identification of the limitation of a statistical method i.e. to match the stability value estimated and the expression profile of a reference gene. Hence, it would be classical to test the developed statistical method over a predetermined synthetic data set, for a direct comparison of estimated stability value with the expression profile of a reference gene.

We devised three simple statistical methods, namely ANOVA, NQQC and EES, and validated them over a synthetic data set of which, ANOVA and NQQC, select a reference gene with stable expression profile in both large and small panels and provide an absolute stability value. While the method EES also works in both small

and large panel, it provides the highest stability value to more than one reference gene having same nature in a panel. These methods also account for sample specific PCR efficiency of each reference gene to estimate more precise stability value. For comparison purpose, we evaluated the same data sets with other widely used methods viz., geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), BestKeeper (Pfaffl et al., 2004) and Δ Ct (Silver et al., 2006).

2. Methods

2.1. Synthetic data

A panel of nine reference genes were synthesized in a manner that it comprised of all kinds of reference genes to make at par of the possibility in a natural/real world data i.e. up-regulated, down-regulated and consistent. Total nine reference genes (A, B, C, D, E, F, G, H and I) of different expression profiles were synthesized manually. Samples were divided into three groups (or categories) namely 1, 2 and 3. In each group, there were eight samples. The Cp values of three genes (A, C and F) were kept almost same in all the three groups, while the Cp values of three genes (B, D and I) were adjusted in a manner to achieve a downregulation in the group 2 and 3, in comparison to group 1. Whereas for the rest three genes (E, G and H) Cp values were fixed in a manner to get up-regulation in the group 2 and 3 in comparison to group 1. For stability value determination, five panels namely P-I, P-II, P-III, P-IV and P-V comprising different nature of reference genes were formed (Table 1). Further, we kept three different PCR efficiency (E) conditions for different data sets: 1) sample specific PCR efficiency for each reference gene, 2) gene specific PCR efficiency, and 3) global PCR efficiency ($E = 2$) for each reference gene in each sample. (See supplementary file 1: Data sets for analysis).

Table 1. Reference gene composition of different panel with their nature and number.

Panel	Number of reference genes	Genes in the Panel		
		Stable	Upregulated	Downregulated
P-I	9	A, C, F	E, G, H	B, D, I
P-II	4	F, A	G	I
P-III	4	F	G	I, D
P-IV	4	F	G, H	I
P-V	3	F	G	I

2.2. Stability evaluation

2.2.1. Common Terms/Formula used in this study (adopted from Pfaffl, 2004)

i) Reference sample = a sample whose Cp value was used to get fold change in all the samples of the respective reference gene (It can be any sample from any group, but must be the same for all the reference genes in the panel; Here we kept S1 i.e. Sample 1 of group 1 for each reference gene);

ii) Transcript number of each reference gene in each sample were calculated using the formula E^{-Cp} . However the exact transcript number of any reference gene in a sample can only be estimated from gene-specific standard curve. Hence, for each reference gene, we have normalized the transcript number in each sample to the reference sample. So, a factor missed due to lack of standard curve were neutralized and further statistical procedure was unbiased;

iii) Intra-gene transcript expression fold ratio (*intra-GTEFR*) = fold ratio of the respective reference gene transcript in a sample normalized to reference sample $\{[E^{-Cp \text{ sample } X (X = 1 \text{ to } n)}] \div [E^{-Cp \text{ reference sample (S1)}] \text{ or } [E^{Cp \text{ reference sample (S1)} - Cp \text{ sample } X (X= 1 \text{ to } n)}]\}$;

iv) Inter gene (relative) transcript expression fold ratio (*inter-GTEFR*) = expression ratio of two reference genes in a sample i.e. ratio of two *intra-GTEFR* in a sample $\{[E_T^{Cp \text{ of } TS1}] * [E_R^{Cp \text{ of } RSX}] \div [E_T^{Cp \text{ of } TSX}] * [E_R^{Cp \text{ of } RS1}]\}$ where TS1: indicates target reference gene (T) in the reference sample; RSX: indicates reference gene (R) in sample X; TSX: indicate target reference gene (T) in sample X, RS1: indicates reference gene (R) in the reference sample; X indicates any sample from 1st to last. The notation T is used in this equation in place of T_{RG} to reduce visual complexity.

v) T_{RG} = target reference gene (a reference gene in a panel whose stability is being evaluated); the other genes in the panel will be simply referred to as reference genes (R).

2.2.2. Method 1: Effective Expression Support (EES) from correlation coefficients

It is the resultant number of reference genes (R) in a panel that have significant positive correlation with T_{RG} for *intra-GTEFR*. It was calculated by subtracting the number of reference genes showing a significant negative correlation with T_{RG} from the number of reference genes showing a significant positive correlation with T_{RG}. The reference genes showing insignificant positive/negative correlation with T_{RG} was dropped out from calculation. Pearson correlation coefficient between two

reference genes for log₂ transformed *intra-GTEFR* were calculated using Statistical Package for Social Science (SPSS) V. 16 at 5% level of significance.

2.2.3. Method 2: Normal Quantile-Quantile Correlation (NQQC)

It measures the Pearson correlation coefficient between the distribution of pooled *inter-GTEFR* of a T_{RG} and theoretical standard normal distribution of same sample size. Transcript number followed by *intra-GTEFR* for each sample of each reference gene were estimated. After that, one by one each reference gene was used as T_{RG} and its *inter-GTEFR* in each sample were calculated against the rest of all the reference genes in that panel. Hence, for each T_{RG} there was a total of $n(N-1)$ number of pooled *inter-GTEFR* value, where 'n' is the total number of samples and 'N' is the number of reference genes in the panel. Pooled *inter-GTEFR* data was log₂ transformed, followed by drawing a normal quantile-quantile plot (NQQP)

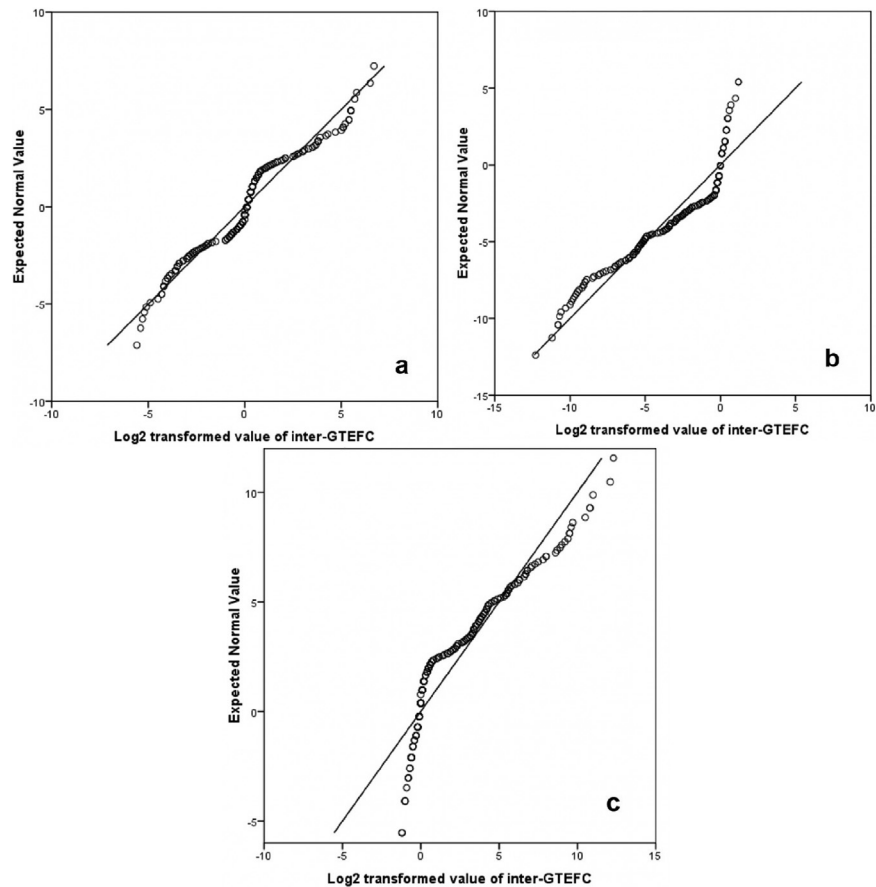


Fig. 1. Normal Q-Q plot of reference genes A, B and E from panel-I at global PCR efficiency ($E = 2$). The plots a, b and c shows the Log₂ transformed value of inter-GTEFR of the reference genes A, B and E respectively. Gene A was evaluated as the most stable by NQQC method as it has the highest quantile-quantile correlation coefficient. Note that the Q-Q plot of gene A is closer to the standard normal curve than B or E.

against theoretical standard normally distributed data point of same sample size using SPSS V. 16 (Fig. 1). In NQQ plot Log2 transformed pooled *inter-GTEFR* produces a minimum deviation from normally distributed data for the most stable reference gene in a panel, whereas an unstable reference gene (up/down-regulated) produces a large deviation. To quantify deviation, correlation coefficient was calculated between Log2 transformed pooled *inter-GTEFR* of each reference gene sorted in ascending order and quantile values of theoretical standard normally distributed data for same sample size. In a panel with equal representation of reference genes of all kinds (stable, up-regulated and down-regulated), Log2 transformed pooled *inter-GTEFR* of the most stable reference gene will produce a distribution closest to normal with maximum normal quantile-quantile correlation/NQQC (Tsai and Yang, 2005; Whiting and Arriaga, 2007).

2.2.4. Method 3: Analysis of variance

In this method, we subjected the categorized log2 transformed pooled *inter-GTEFR* of each T_{RG} to one-way ANOVA in SPSS V16. In general terms, those genes from the panel shall be considered stable which have a low F ratio value, indicating that the expression does not vary with the treatments (Larson, 2008; Kim, 2014). Among those, the gene with the lowest F value will be evaluated as the most stable, as it shows minimal variation among all. When all genes in the panel show high F value (high between-treatment variation), the researcher could opt the gene with the lowest F value and use a correction factor as explained in the following section, if they do not want to enlarge their panel and continue evaluation. Prior to employing the three methods, the Inter gene (relative) transcript expression fold ratio (*inter-GTEFR*) were log2 transformed to achieve normality in data (Osborne, 2002).

2.2.5. Correction factor from ANOVA

The corrected relative fold change of a target gene in a category (treatment/tissue type) is the product of its uncorrected relative fold change and the correction factor. Correction factor is estimated as $2^{(Y-W)}$ where Y and W are the averages of log2 transformed *inter-GTEFR* of a T_{RG} in treatment group and control group respectively, selected for and employed for the normalization. The base of 2 is employed to back transform the data. The Correction factor can be easily obtained from ANOVA output of Statistical Package for the Social Sciences (SPSS) performed over *inter-GTEFR*.

The different panels used in our studies were also analyzed by geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), BestKeeper (Pfaffl et al., 2004) and ΔCt (Silver et al., 2006) to see the validity of our statistical assumption.

Table 2. EES of different reference genes in the different panels; α = Sample specific PCR efficiency (E); β = Gene specific PCR efficiency (E); γ = Global PCR efficiency (E); Superscript * indicates the genes having highest EES in panel.

Panel	Gene→ $E\downarrow$	EES Value								
		A	B	C	D	E	F	G	H	I
P-I	α	0	0	0	-2	0	0	0	0	1*
	β	2*	2*	1	2*	-4	2*	-4	-4	1
	γ	2*	2*	1	2*	-4	2*	-4	-4	1
P-II	α	0*	-	-	-	-	0*	0*	-	0*
	β	1*	-	-	-	-	1*	-3	-	1*
	γ	1*	-	-	-	-	1*	-3	-	1*
P-III	α	-	-	-	-1	-	0*	-1	-	0*
	β	-	-	-	1*	-	1*	-3	-	1*
	γ	-	-	-	1*	-	1*	-3	-	1*
P-IV	α	-	-	-	-	-	0	1*	1*	0
	β	-	-	-	-	-	-1*	-1*	-1*	-1*
	γ	-	-	-	-	-	-1*	-1*	-1*	-1*
P-V	α	-	-	-	-	-	0	1*	-	1*
	β	-	-	-	-	-	0*	-2	-	0*
	γ	-	-	-	-	-	0*	-2	-	0*

3. Results

3.1. Effective Expression Support (EES)

EES of all the reference genes in the respective panels was estimated and is detailed in Table 2.

3.2. NQQC test

It can be seen that the ranking of the genes by the NQQC method is significantly affected by the PCR efficiency (E) value employed in the calculation (Table 3). Genes with stable Cp values were adjudged as the best in all but one panel (panel-II) when gene specific and global PCR efficiency values were used.

3.3. ANOVA test

The results of the ANOVA test for stability evaluation is summarized in Table 4. While the gene with stable Cp values was selected in most of the cases (A and F), it

Table 3. Stability values of different reference genes in the different panels from NQQC test; α = Sample specific E ; β = Gene specific E ; γ = Global E ; gene having highest stability has been ranked 1st followed by subsequent stable gene in panel.

Panel	Stability Rank→ E↓	1 st	2 nd	3 rd	4 th	5 th	6 th	7 th	8 th	9 th
P-I	α	B(0.872)	D(0.862)	A(0.831)	F(0.828)	E(0.823)	H(0.823)	I(0.820)	G(0.819)	C(0.701)
	β	A(0.980)	C(0.978)	F(0.978)	G(0.964)	I(0.961)	D(0.952)	B(0.951)	H(0.951)	E(0.950)
	γ	A(0.977)	F(0.976)	C(0.975)	I(0.973)	D(0.951)	B(0.949)	G(0.947)	E(0.942)	H(0.942)
P-II	α	A(0.873)	I(0.865)	G(0.851)	F(0.789)	–	–	–	–	–
	β	G(0.975)	A(0.973)	F(0.967)	I(0.961)	–	–	–	–	–
	γ	G(0.977)	A(0.967)	F(0.958)	I(0.950)	–	–	–	–	–
P-III	α	D(0.971)	G(0.907)	I(0.902)	F(0.855)	–	–	–	–	–
	β	F(0.989)	G(0.969)	D(0.950)	I(0.957)	–	–	–	–	–
	γ	F(0.989)	I(0.974)	G(0.968)	D(0.963)	–	–	–	–	–
P-IV	α	I(0.921)	G(0.866)	H(0.842)	F(0.770)	–	–	–	–	–
	β	F(0.984)	G(0.968)	H(0.959)	I(0.956)	–	–	–	–	–
	γ	F(0.977)	I(0.953)	G(0.947)	H(0.945)	–	–	–	–	–
P-V	α	I(0.851)	G(0.815)	F(0.739)	–	–	–	–	–	–
	β	F(0.986)	G(0.976)	I(0.959)	–	–	–	–	–	–
	γ	F(0.978)	G(0.976)	I(0.950)	–	–	–	–	–	–

can be seen that in three instances an apparently down regulated gene (I) was selected when sample specific PCR efficiency was employed. In panel P-IV, none of the genes qualified to be deemed stable, irrespective of whether classified as stable or downregulated as per Cp values.

3.4. Comparative analysis with the other reported methods

Table 5 presents the comparative analysis of all the methods viz, NQQC, ANOVA, Δ Ct, BestKeeper, NormFinder and geNorm. It can be seen from Table 5 that our techniques produce a result different from the existing methods only in instances where gene specific and sample specific PCR efficiencies are used for calculation.

4. Discussion

Most of the statistical methods described to evaluate the stability of reference genes from a panel of several genes rely on assessing the variability in the Cp value data (Pfaffl et al., 2004; Paolacci et al., 2009). The principle behind these analyses are that Cp values are the direct numerical representation of gene expression. The source of variation in Cp values is two fold- process/sample induced (false instability) and treatment induced variation in the gene expression (true instability).

Table 4. Stability ranking (based on F-value) of different reference genes in the different panels from ANOVA test; α = Sample specific E ; β = Gene specific E ; γ = Global E . Either of the superscript * or # indicated the genes having lowest F-value and highest stability in panel with respective PCR efficiency in panel. Superscript * shows genes that clearly did not have a significant difference among treatment groups. Superscript # indicates gene was selected despite showing significant difference among the treatment groups ($p < 0.05$). The Cp values of this gene would have to be subjected to correction prior to analysis.

		Nature of the genes included in the panel								
Panel	PCR efficiency (E) value employed	Stable			Upregulated			Downregulated		
		A	C	F	E	G	H	B	D	I
P-I	α	0.415*	8.27	2.25	20.28	11.96	15.7	28.02	17.35	1.72
	β	0.137*	0.592	0.674	81.86	30.33	87.63	119.06	63.76	27.13
	γ	0.184*	0.624	0.6	74.55	39.73	73.14	106.84	92.36	20
P-II	α	0.26*	–	3.38	–	10.84	–	–	–	1.78
	β	0.163*	–	0.872	–	82.07	–	–	–	60.91
	γ	0.138*	–	1.27	–	98.83	–	–	–	47.34
P- III	α	–	–	1.185	–	17.34	–	–	13.45	0.57*
	β	–	–	1.93*	–	90.34	–	–	35.11	7.99
	γ	–	–	1.62*	–	8.041	–	–	49.63	3.91
P- IV	α	–	–	6.38	–	5.28	9.12	–	–	3.18 [#]
	β	–	–	4.63 [#]	–	7.12	45.12	–	–	60.23
	γ	–	–	5.75 [#]	–	11.16	33.98	–	–	58.74
P- V	α	–	–	2.95	–	7.94	–	–	–	0.935*
	β	–	–	0.576*	–	64.51	–	–	–	40.79
	γ	–	–	0.871*	–	75.81	–	–	–	32.63

In other words, Cp values are highly influenced by sample differences arising out of procedural errors during dissection, storage, RNA isolation etc. The Cp variability arising out of these reasons are not easily identified in these methods leading to low accuracy of stability estimations. Here we report three methods that effectively neutralize the false instability and accounts only true instability.

The simplest of the three methods described here is the EES. It functions on the basic assumption that functionally unrelated reference genes have a low probability of co-regulation (Vandesompele et al., 2002; Andersen et al., 2004) for any given experimental condition. This should follow that there should not be a negative correlation in the expression of any of the genes in the chosen panel; any negative correlation between two given genes indicates instability of expression of either one or both of them. In a cumulative manner, thus the stability indication of a given reference gene would be the effective number of genes that show a positive

Table 5. Comparative results of stability analysis by the different methods at PCR efficiency of 2 (global PCR efficiency); The values in the parenthesis are the stability indices as produced by each method. For NQQC and ANOVA, superscript * indicate gene evaluated as most stable at global PCR efficiency and # indicate gene evaluated as most stable at sample specific and/or gene specific PCR efficiency.

METHOD	P-I	P-II	P-III	P-IV	P-V
NQQC	A*/B#	G*/A#	F*/D#	F*/I#	F*/I#
ANOVA	A*/A#	A*/A#	F*/I#	F*/I#	F*/I#
ΔC_t	A(1.52)	A(1.07)	F(1.70)	G(1.66)	F(1.46)
BestKeeper	A(0.18)	A(0.18)	F(0.35)	F(0.35)	F(0.35)
NormFinder	A(0.203)	A(0.179)	F(0.517)	F(0.872)	F(0.394)
geNorm	AIC(0.207)	AIF(0.346)	FII(1.154)	GIH(0.511)	FII(1.154)

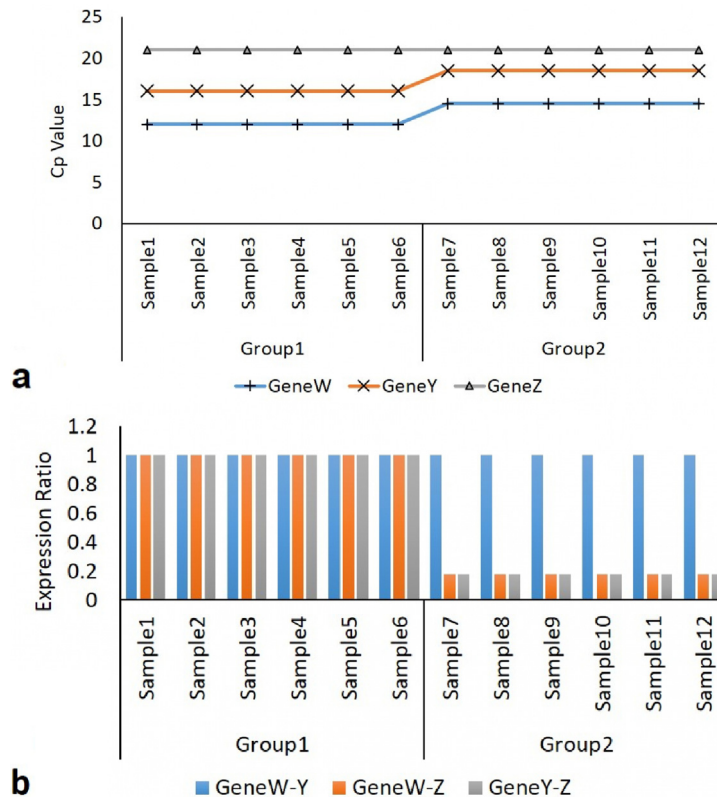


Fig. 2. Concept of EES: a) Plot of Cp values of genes W, Y and Z in each sample. There is an apparent downregulation (upward trend of Cp values) in W and Y but not in Z. b) Plot of expression ratios of the three genes calculated sequentially to one another. Note that downregulation of expression ratios between the groups occur only when gene Z is involved in the calculation. When the expression ratio between W and Y is estimated, which are apparently downregulated as per Cp data, expression ratio is consistent between the treatment groups. In other words the apparent stability of gene Z as reflected by Cp values could be the result of false instability elaborated in the introduction section.

correlation to it (Fig. 2). However, a major disadvantage of EES is that it does not accommodate variation (stability or instability) conferred from those reference genes in the panel showing insignificant correlation with T_{RG} . Additionally, reference genes belonging to different functional classes never warrant that they will not show co-regulation in a biological phenomenon or experiment and it would be an injustice to assume constant expression ratio or high EES as stability index. EES only provide the prima facie stability measure of a reference gene and hence to estimate stability in absolute terms, two additional methods, namely NQQC and ANOVA were adopted.

For ANOVA and NQQC, first of all transcript expression fold ratio of a concerned reference gene (T_{RG}) in comparison to the rest of all the reference genes in a panel was estimated. Thereafter, the stability value of a concerned T_{RG} was estimated as: i) correlation coefficient between the quantiles of transcript expression fold ratio and the quantile of theoretically standard normally distributed data for same sample size and ii) the between-group and the within-group variance (ANOVA) for categorized pooled transcript expression fold ratio.

The basic assumption of the NQQC method is that a stable T_{RG} exhibits a stable expression (i.e. constant *inter-GTEFR*) in different categories (tissue type/ treatments/stage etc.) when normalized with another stable reference gene from the panel. However, it exhibits an up-regulation when normalized with a down-regulated reference gene and vice versa. If a panel consists of all kinds of reference genes, the most stable reference gene exhibits the most homogeneous variation in its pooled *inter-GTEFR* distribution and will be closest to the normal. On the contrary, an up-regulated T_{RG} will exhibit a stable distribution across categories when normalized with an up-regulated reference gene, but its distribution becomes up-regulated upon normalization with a stable gene and increasingly so with a down-regulated reference gene. The condition will be similar but in opposite direction for a downregulated T_{RG} . To sum up, they will have a non-homogenous variation in its pooled *inter-GTEFR*, skewed towards one side.

In order to identify instability in reference genes in the panel by ANOVA, the direct observation would be the between-treatment variation in expression (C_p values) of a particular gene (Paolacci et al., 2009). This would mean that the treatment means of C_p values for a given gene 'A' would be compared independent of other genes in the panel. However, the C_p variation in gene 'A' might have arisen as a result of sample infidelities, which is probably reflected equally in the C_p data of other genes in the panel, though we chose not to consider it. In other words, ANOVA of C_p values fail to separate true and false instability, simply because it functions on the assumption of sample uniformity. In our method, we remove this assumption of sample uniformity. In place of C_p values of gene 'A' a given sample S_1 would have the expression ratios of 'A' normalized

iteratively to each of the rest of the genes in the panel. In doing so, we bring in the variation imparted to all the genes in the panel by the sample S_1 . This follows that under the treatment T, for gene A, there are now a total of $(S*N)$ expression ratio values, where S is the number of samples studied under T, N is the number of genes in the panel. The error factor of ANOVA now arises as a result of sample imparted variations that are highly specific to gene 'A'; sample imparted variations that affect all genes in general have been normalized within the expression ratio values. One of the major causes of this kind of error is the contamination of sample tissue by the surrounding tissue types during dissection. The expression of a reference gene in the panel may or may not be differentially expressed in response to the treatment conditions in the contaminating tissue, giving rise to the error.

Thus arises the need for a word of caution in judging the true instability of a gene on the sole basis of low F ratio. Since the F value is the ratio of between-group variance and within-group variance, there could be the possibility of high within group variance contributing to the lowered F value. To rule out this issue, albeit of low probability, the researcher is advised to look out for unreasonably high error sum of square values in the ANOVA (Larson, 2008; Kim, 2014). Student's t-test of the relative quantity of a reference gene transcript (*intra-GTEFR*) of each reference gene (Lardizábal et al., 2012) have also been proposed to test the expression stability of a reference gene. However, unlike our method, this method does not account for variation from all the reference genes in the panel.

The output of NQQC and ANOVA disagreed with that of ΔCt and geNorm in panel-IV. The authors of ΔCt method caution that the results get affected in a small panel if the panel comprises of reference genes of divergent nature (Silver et al., 2006) as was the case here. The method geNorm, remove a least stable reference gene (poorest for constant expression ratio with the majority of the reference gene in the panel) from the panel in a stepwise manner and conclude on two reference genes in the panel. GeNorm takes separate reference point for each gene (minimal Cp value of the respective reference gene). Hence, there is every possibility that many reference genes will have minimal Cp value in the different samples. In such case the estimated expression ratio will be biased. To avoid this, in NQQC and ANOVA method, we consider Cp value of the respective reference gene from a single sample as reference Cp.

BestKeeper accounts for variation in Cp value of the individual reference gene (T_{RC}) and considers the gene with minimal variation in Cp value as the most stable (Pfaffl et al., 2004). The output of BestKeeper was as expected, since we had taken a synthetic predefined data set in which the stable reference gene has minimum Cp variation. NQQC and ANOVA produced the same output as BestKeeper and NormFinder in most instances for this data set. When validated with a real life data set (data not shown), there were disagreements with BestKeeper, but not with

NormFinder. In NormFinder the intra-group and inter-group variation in expression of the individual T_{RG} are measured (Andersen et al., 2004), partially explaining this observation. On the other hand, BestKeeper relies on the raw Cp value variations within one gene, while our methods account for variation from all the reference genes in the panel to produce a more accurate stability measure for a T_{RG} .

Of the four established methods, none facilitate the use of sample specific PCR efficiencies in computations. It is interesting to note that any variation in the results between our methods and the other methods (especially NormFinder) occurred when sample specific PCR efficiencies were employed in the calculations. The chosen best reference genes in these cases were not necessarily ones with a stable Cp data. For example, with sample specific PCR efficiency, in panel-IV gene F appears worst, however with gene specific and global PCR efficiency, the same gene appears the best, which also comply with the result of NormFinder and geNorm, because these two methods consider only the global PCR efficiency ($E = 2$). More importantly the occurrence of such deviations in the results were not influenced by the size of the panel as a result of pooling all possible sources of Cp value variations, indicating the sensitivity and robustness of our methods. The significance of PCR efficiency in the estimation of transcript copy number from the Cp values is well understood (Livak and Schmittgen, 2001; Pfaffl, 2004). Considering that Cp values could vary due to sample specific conditions, inclusion of sample specific PCR efficiency in computations become indispensable for accuracy. Both the methods described in this paper facilitate the option of including this important factor in analyses. The observations were consistent for any sub panel drawn from this data set.

Though the results of NQQC and ANOVA estimated from gene-specific PCR and global PCR efficiencies are similar in both large and small panels, the results contradict in the large panel (P-I) when sample specific PCR efficiency is used. (The similarity and dissimilarity of results are discussed with respect to the nature of the gene chosen, but not the specific gene per say). NQQC-ANOVA result contradictions were noted when panels had an unequal representation of gene types in the panel, which is one of the basic assumptions of NQQC test. In all such cases ANOVA results are more accurate as ANOVA does not hold onto such an assumption. However a combination of NQQC and EES becomes significant when data is uncategorized and cannot be analyzed by ANOVA.

Among the four established methods, BestKeeper (Pfaffl et al., 2004) and ΔC_t (Silver et al., 2006) methods simply rank the stability value of the reference genes in a panel and does not provide any correction factor in case of unavailability of a truly stable reference gene. The other two methods, geNorm (Vandesompele et al., 2002) and NormFinder (Andersen et al., 2004) provide a correction factor,

but since they accommodate the variation from only selected reference genes from the panel, this would be unjustifiable as explained earlier. ANOVA method, developed in this study, produces only one most stable reference gene as result with the correction factor which accounts for variation from all the reference genes in a panel i.e. giving equal weightage to all the reference genes.

5. Conclusion

It is evident from the results that with a change in PCR efficiency, ANOVA, NQQC and EES test, change the stability value of reference genes in a panel on the ground of a possible change in transcript expression ratio. While EES provide a prima facie indication of the stability of the gene, NQQC and ANOVA effectively determines a stable reference gene from a panel for non-structured and structured data respectively. The NQQC and ANOVA test account for all the variation within a reference gene panel and analyze category wise pooled expression ratio to assign a stability value. This makes this method more sensitive and robust in both large and small panels. Further, we developed a Virtual basic application (VBA) in Microsoft excel, which directly produces the result for NQQC test and category-wise pool of log₂ transformed expression ratio (*inter-GTEFR*) for performing ANOVA in SPSS. The developed NQQC-ANOVA VBA can handle a panel of virtually any number of genes and samples. See Supplementary file 2 for NQQC-ANOVA MS-Excel VBA and Supplementary file 3 for procedure to operate NQQC-ANOVA MS-Excel VBA.

Declarations

Author contribution statement

Himanshu Priyadarshi: Conceived and designed the experiments.

Rekha Das: Wrote the paper.

Shivendra Kumar: Analyzed and interpreted the data.

Pankaj Kishore: Performed the experiments.

Sujit Kumar: Contributed reagents, materials, analysis tools or data.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interest statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at <http://dx.doi.org/10.1016/j.heliyon.2017.e00233>

Acknowledgment

The authors are thankful to Mr. Amit Dutt Tripathi (email: amittripathi4@gmail.com), Senior software engineer in Codeit Software Ltd., New Delhi, India for developing VBA in MS-Excel.

References

- Andersen, C.L., Jensen, J.L., Ørntoft, T.F., 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 (15), 5245–5250.
- Bustin, S.A., 2002. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J. Mol. Endocrinol.* 29 (1), 23–39.
- Chandna, R., Augustine, R., Bisht, N.C., 2012. Evaluation of candidate reference genes for gene expression normalization in *Brassica juncea* using real time quantitative RT-PCR. *PLoS one* 7 (5) e36918.
- Das, R., Krishna, G., Priyadarshi, H., Gireesh-Babu, P., Pavan-Kumar, A., Rajendran, K.V., Reddy, A.K., Makesh, M., Chaudhari, A., 2015. Captive maturation studies in *Penaeus monodon* by GIH silencing using constitutively expressed long hairpin RNA. *Aquaculture* 448, 512–520.
- Grozinger, C.M., Fan, Y., Hoover, S.E., Winston, M.L., 2007. Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Mol. Ecol.* 16 (22), 4837–4848.
- Guescini, M., Sisti, D., Rocchi, M.B., Stocchi, L., Stocchi, V., 2008. A new real-time PCR method to overcome significant quantitative inaccuracy due to slight amplification inhibition. *BMC Bioinform.* 9 (1), 1.
- Guo, J., Ling, H., Wu, Q., Xu, L., Que, Y., 2014. The choice of reference genes for assessing gene expression in sugarcane under salinity and drought stresses. *Sci. Rep.* 4, 7042.
- Hildyard, J.C., Wells, D.J., 2014. Identification and validation of quantitative PCR reference genes suitable for normalizing expression in normal and dystrophic cell culture models of myogenesis. *PLOS Curr.: Muscular Dystrophy*.

- Kim, H.Y., 2014. Analysis of variance (ANOVA) comparing means of more than two groups. *Restor. Dent. Endod.* 39 (1), 74–77.
- Lardizábal, M.N., Nocito, A.L., Daniele, S.M., Ornella, L.A., Palatnik, J.F., Veggi, L.M., 2012. Reference genes for real-time PCR quantification of microRNAs and messenger RNAs in rat models of hepatotoxicity. *PloS one* 7 (5) e36323.
- Larson, M.G., 2008. Analysis of variance. *Circulation* 117 (1), 115–121.
- Leelatanawit, R., Klanchui, A., Uawisetwathana, U., Karoonuthaisiri, N., 2012. Validation of reference genes for real-time PCR of reproductive system in the black tiger shrimp. *PloS one* 7 (12) e52677.
- Ling, H., Wu, Q., Guo, J., Xu, L., Que, Y., 2014. Comprehensive selection of reference genes for gene expression normalization in sugarcane by real time quantitative RT-PCR. *PloS one* 9 (5) e97469.
- Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* 25 (4), 402–408.
- Osborne, J., 2002. Notes on the use of data transformations. *Practical Assessment, Research and Evaluations* 8 Retrieved May 10, 2003.
- Paolacci, A.R., Tanzarella, O.A., Porceddu, E., Ciaffi, M., 2009. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Mol. Biol.* 10 (1), 1.
- Paria, A., Greeshma, S.S., Chaudhari, A., Makesh, M., Purushothaman, C.S., Rajendran, K.V., 2013. Nonspecific Effect of Double-Stranded (ds) RNA on Prophenoloxidase (proPO) Expression in *Penaeus monodon*. *Appl. Biochem. Biotechnol.* 169 (1), 281–289.
- Pfaffl, M.W., 2004. Quantification strategies in real-time PCR. *AZ of quantitative PCR* 1, 89–113.
- Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P., 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* 26 (6), 509–515.
- Sharabi, O., Manor, R., Weil, S., Aflalo, E.D., Lezer, Y., Levy, T., Aizen, J., Ventura, T., Mather, P.B., Khalaila, I., Sagi, A., 2015. Identification and characterization of an insulin-like receptor involved in crustacean reproduction. *Endocrinology* 157 (2), 928–941.

- Silver, N., Best, S., Jiang, J., Thein, S.L., 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* 7 (1), 1.
- Taki, F.A., Zhang, B., 2013. Determination of reliable reference genes for multi-generational gene expression analysis on *C. elegans* exposed to abused drug nicotine. *Psychopharmacology* 230 (1), 77–88.
- Tichopad, A., Dilger, M., Schwarz, G., Pfaffl, M.W., 2003. Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucleic Acids Res.* 31 (20), e122.
- Tsai, D.M., Yang, C.H., 2005. A quantile–quantile plot based pattern matching for defect detection. *Pattern Recogn. Lett.* 26 (13), 1948–1962.
- 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 Biol.* 3 (7), 1.
- Van Hoeck, V., Scolari, S.C., Pugliesi, G., Gonella-Diaza, A.M., Andrade, S.C., Gasparin, G.R., Coutinho, L.L., Binelli, M., 2015. Gene expression profiling by high throughput sequencing to determine signatures for the bovine receptive uterus at early gestation. *Genom. Data* 5, 94–96.
- Wang, X., Gong, Y., Wang, D., Xie, Q., Zheng, M., Zhou, Y., Li, Q., Yang, Z., Tang, H., Li, Y., Hu, R., 2012a. Analysis of gene expression profiling in meningioma: deregulated signaling pathways associated with meningioma and EGFL6 overexpression in benign meningioma tissue and serum. *PLoS one* 7 (12), e52707.
- Wang, Z.L., Liu, T.T., Huang, Z.Y., Wu, X.B., Yan, W.Y., Zeng, Z.J., 2012b. Transcriptome analysis of the Asian honey bee *Apis cerana cerana*. *Plos one* 7 (10), e47954.
- Whiting, C.E., Arriaga, E.A., 2007. Evaluation of individual particle capillary electrophoresis experiments via quantile analysis. *J. Chromatogr. A* 1157 (1), 446–453.
- Wilson, I.G., 1997. Inhibition and facilitation of nucleic acid amplification. *Appl. Environ. Microbiol.* 63 (10), 3741.
- Xie, F., Xiao, P., Chen, D., Xu, L., Zhang, B., 2012. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol. Biol.* 80 (1), 75–84.