Reference genes in real-time PCR

Bartłomiej Kozera · Marcin Rapacz

Abstract This paper aims to discuss various aspects of the use of reference genes in qPCR technique used in the thousands of present studies. Most frequently, these are house-keeping genes and they must meet several criteria so that they can lay claim to the name. Lots of papers report that in different conditions, for different organisms and even tissues the basic assumption—the constant level of the expression is not maintained for many genes that seem to be perfect candidates. Moreover, their transcription can not be affected by experimental factors. Sounds simple and clear but a great number of designed protocols and lack of consistency among them brings confusion on how to perform experiment properly. Since during selection of the most stable normalizing gene we can not use any reference gene, different ways and algorithms for their selection were developed. Such methods, including examples of best normalizing genes in some specific cases and possible mistakes are presented based on available sources. Numerous examples of reference genes applications, which are usually in too few numbers in relevant articles not allowing to make a solid fundament for a reader, will be shown along with instructive compilations to make an evidence for presented statements and an arrangement of future qPCR experiments. To include all the pitfalls and problems associated with the normalization methods there is no way not to begin from sample preparation and its storage going through candidate gene selection, primer design and statistical analysis. This is important because numerous short reviews available cover the topic only in lesser extent at the same time giving the reader false conviction of complete topic recognition.

Introduction

The study of gene expression profiles is commonly performed by relying on such techniques like Northern or cDNA micro-arrays and it is mostly thanks to the possibility of analyzing many genes simultaneously and economic aspects (Mallona et al. 2010). However real-time PCR (qPCR) technique is considered to be the most accurate and most reliable for what often serves to validate data obtained by other methods. Undoubtedly, its advantages are sensitivity, real time detection of reaction progress, speed of analysis and precise measurement of the examined material in the sample (Gachon et al. 2004). Moreover expression level for some genes is often so small that qPCR becomes the only technique that can detect such a small number of mRNA copies. But if real-time PCR is about to reach its maximum analytical potential it is necessary to introduce appropriate normalization methods and to validate the results. It is relentlessly stressed that many qPCR experiments lack authors critical evaluation, are wrongly designed and difficult to repeat due to insufficient data quality (Bustin et al. 2009). This appears to be of greater issue for studies where qPCR serves as a supplementary method among others. For example, a common problem is the difference in the extraction of mRNA between samples and performance of reverse transcription and PCR itself (Huggett et al. 2005). To avoid the influence of these factors, normalizing gene is applied against which the level of expression will be determined.
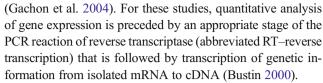
It remains a question of a different matter what makes a process of reference genes normalization, a recurring problem that is being addressed by scientists in recent years (Huggett and Bustin 2011). There is an enormous range of protocols, various methodologies and data available somehow affecting the integrity of scientific literature. At the same time those papers taken together can be misleading: incongruously the concise publication manner makes information about protocol

B. Kozera (✉)
Polish Academy of Sciences, The Franciszek Górski Institute of Plant Physiology, Podłużna 3, 30-239 Kraków, Poland
e-mail: bkozera@ifr-pan.krakow.pl

M. Rapacz
Department of Plant Physiology, University of Agriculture in Krakow, Podłużna 3, 30-239 Kraków, Poland

insufficient to assess the assay and perform validation process. Such terms as "qPCR gold standard", "stable expression and uniform efficiency", "validated in previous publications" are being used over and over again. In other words authors keep shuffling with assumptions as if referring to them as well known facts was enough to provide high-quality data. A good awareness rising example is a report of "catalogue of mistakes, inaccuracies and inappropriate analysis methods as well as contamination and poor assay performance" carried out by Bustin (2008) acting as an expert witness in court for specific case where he was undermining conclusion of a link between authism and entheropathy as previously stated by Uhlmann et al. (2002). It is frequently raised that a consensus on how to perform and evaluate qPCR experiment does not exist which implicates a kind of conservatism in citing references. Review article by Thellin et al. (1999) was cited 643 times (Scopus–August 2013) with 95 citations since 2012 (Scopus–August 2013). Yet, it seems advisable to use such reference as a base for further discussion by anchoring to the most familiar articles to scientists.

The aim of the present review is to provide a complete handbook of reference genes issues in qPCR which includes all the pitfalls and problems associated with the method beginning from sample preparation and its storage, candidate gene selection, primer design and robust statistical analysis. This is important because numerous short reviews available cover the topic only in lesser extent at the same time giving the reader false conviction of complete topic recognition. Numerous examples of reference genes applications, which are usually in too few numbers in articles not allowing to make a solid fundament for readers, will be shown along with instructive compilations to make an evidence for presented statements and enabling planning future qPCR experiments. Another aspect that needs to be raised is discussion about available qPCR software, where comparison of different validation algorithms along with reference genes selection methods, which are equally important, tend to be treated with smaller care.

Problems associated with the method

Polymerase chain reaction was developed in 1983 by Kary Mullis and colleagues (Saiki et al. 1988). It involves logarithmic amplification of genetic material based on the matrix and designed primers that bind to it. The reaction proceeds through three cyclically repeated reactions in their respective temperatures. Those are: matrix denaturation, primer hybridization and elongation. In theory, basal material is duplicated twice in each cycle assuming 100 % efficiency. In today's diagnostic tests, including those dealing with gene expression, real-time PCR is considered as the most reliable and most accurate method which is a modification of the classical PCR

(Gachon et al. 2004). For these studies, quantitative analysis of gene expression is preceded by an appropriate stage of the PCR reaction of reverse transcriptase (abbreviated RT–reverse transcription) that is followed by transcription of genetic information from isolated mRNA to cDNA (Bustin 2000).

Real-time imaging is possible through the use of special fluorescent dyes, including SYBR Green I, which binds to double-stranded DNA and is observed by nearly 1000-fold increase in fluorescence intensity (Huggett and Bustin 2011). More advanced is the use of fluorescent-labeled oligonucleotides where signal is generated only when amplicon-specific probe hybridizes to complementary region. Such chemistry like hydrolysis probes "TaqMan" or structured probes "molecular beacons" can greatly increase reaction specificity.

In the first phase of qPCR DNA replication is the slowest, since the quantity of template is relatively small compared to the later stages, and starters need more time to find complementary regions (Studzińska et al. 2008). The faster the reaction reaches the next phase, called the exponential, the more template of the target sequence is in the sample. The cycle in which fluorescence begins to exceed the background level is named cycle threshold (Ct) and is the beginning of the following logarithmic phase.

The analysis of the amount of gene transcripts may be based on absolute or relative calculations (Tyburski et al. 2008). In the absolute method the determination of the template copy number in the sample is based on a standard curve prepared with serial dilutions of known concentration solution of the test sequence. In the relative method, the curve is only used to calculate the reaction efficiency, and the result is given in relation to the calibrator. Commonly the number of transcripts is given in relation to the amount of pre-defined reference genes. Such normalization can effectively correct the differences between the compared samples.

Achieving reliable results is only possible after application of an appropriate normalization method. It is an absolute necessity because the technique of real-time RT-PCR poses problems at various stages of sample preparation and processing. The most commonly mentioned problems are: RNA extraction procedure along with sample storage and its quality, the process of reverse transcription with cDNA synthesis including poorly selected primers and inappropriate statistical analysis (Bustin et al. 2009; Mallona et al. 2010). Already during obtaining material from tissues there is no assurance that in spite of achieving uniform size and weight of the samples they will contain the same amount of matrix for the reaction. To set an example: same volume of human blood with HIV was acquired but patients with less advanced immunosuppression have fewer cells in 1 ml of blood, and therefore less RNA than was extracted from the same volume of blood of patients who were in a more advanced stage (Huggett et al. 2005). Taking under consideration that a cell under the influence of internal factors and external environment is a dynamic

and variable living form, it is obvious that in various samples high variability is observed, both in absolute amounts and relative content of each mRNA (Bustin et al. 2005). Errors in pipetting and transferring of the material may be one of the reasons of such variability. The process of extraction and purification of RNA may proceed with varying efficiency—it is related to instability of nucleic acids. In addition, biological material often contains many low-specific nucleic acids that are a part of the background not to mention native RT and PCR polymerase inhibitors. To overcome these problems RNA's purity can be checked through absorbance ratios and integrity by SPUD-assay but there is no gold standard for such examination (Nolan et al. 2006; Bustin et al. 2010; De Keyser et al. 2013). Lack of inhibitors can be shown even when absorbance ratio is not perfect allowing RNAs further use. At the same time, RNAs of different quality should never be quantitatively compared. The popular method of determining RIN/RQI (RNA integrity value/RNA quality value) values only takes into account complete electropherogram while software was originally trained by mammalian tissues. This assumes perfect ratio of 2 for 28S/18S RNA which may not be appropriate for plants (not to mention discrepancies between cells and tissues). In such approach applied for plants precisely one must remember that there will not be 28S RNA but 25S RNA or even additional two peaks: 16S RNA and 23S RNA if it is a chloroplast-containing total RNA. This will be identified as degradation peaks resulting in misleading values (Taylor et al. 2010; De Keyser et al. 2013).

In the isolated RNA (referred to as "total RNA") amounts of ribosomal RNA are present, which are subjected to the reverse transcription while the right target is mRNA (Hendriks-Balk et al. 2007). A serious problem in the normalization of total RNA is frequent differences in the rRNA:mRNA ratios which do not guarantee to obtain robust data.

Assuming all the problems the normalization step is essential. Most authors agree that the use of reference genes is the most effective method and is likely to be one of the easiest one to correct errors of the whole research (Huggett et al. 2005). Its use, however, requires a wider look at a variety of relevant genes and at the need for their validation.

Reference genes

qPCR was introduced in 1992 by Higuchi and co-workers (Higuchi et al. 1992) but it was a few years later when a matter of greater importance was put in reference genes. At the same time qPCR was still a novel and developing technique used only in approximately 8 % of mRNA quantifying studies (Thellin et al. 2009). Reference genes are an internal reaction control that have sequences different than the target. For a gene to be regarded as a reliable reference it must meet several important criteria (Chervoneva et al. 2010). The most

important is its expression level unaffected by experimental factors. Also, it should show minimal variability in its expression between tissues and physiological states of the organism. It is desirable to pick such reference that would show a similar threshold cycle with gene of interest. Reference gene must in turn demonstrate the variability resulting from imperfections of the technology used and preparatory procedures—this ensures that any variation in the amount of genetic material will relate to the same extent as the object of research and control. It seems that the perfect fulfillment of these conditions are the basic metabolism genes (called Housekeeping Genes–HKGs) which, by definition, being involved in processes essential for the survival of cells, must be expressed in a stable and non-regulated constant level and in fact they were first to be examined as reference genes (Thellin et al. 1999). This was always questioned, even at the time of forming the assumption since many of them participate not only in basic metabolic processes and what seems to be perfect in one experiment does not guarantee its functionality in another. What sounds simple and obvious forces in turn an individual and complex approach to each experiment and the necessity of their careful choice, along with validation which, to some extent was and still is being deficient in many papers.

There were proposed different statistical approaches and algorithms for the optimal choice of a couple or more reference genes. One ranks genes by the stability measure M, which value is calculated by the average pairwise variation of a single gene with all other candidate reference genes (Vandesompele et al. 2002). It is recommended to always use at least two reference genes, since the use of only one may lead to relatively large errors (Nicot et al. 2005). Another advantage that comes along with using more of them is usually increase of resolution and greater accuracy of the results. In exceptional circumstances, use of a single reference gene is acceptable, if it was previously tested in similar experimental conditions and properly validated (Thellin et al. 1999), but it was lately reported that it still may result in significant bias like three-fold in 25 % of the analyzed results, or even six-fold in a narrower range of 10 % (Derveaux et al. 2010).

To determine the concentration of the studied sequence, a relative method can be used, where one sample (usually the same gene that is not exposed to experimental factor) is a calibrator, i.e., the sample against to which change is given (Tyburski et al. 2008). For each reaction there must be determined efficiency (E), usually close to 2 (but in practice never equal) and using the Ct values-based differences between genes can be considered. Efficiency can be disturbed by aberrant product synthesis as a result of side enzymatic inhibitors or undesired secondary structures (Pettengill et al. 2012) although it is equipment and pipetting (pipette's calibration!) that are the source of most errors (Taylor et al. 2010). Provided there are similar reactions efficiency for each gene comparative method ($\Delta\Delta$Ct) can be used (Livak and Schmittgen

2001) or Pfaffl model (Pfaffl 2001), if larger differences in efficiency are observed. First, for each sample, difference between ΔCt of studied gene and control gene is calculated, then subtract between (so the value of the "$\Delta\Delta Ct$") ΔCt of sample with unknown concentration and ΔCt of the calibrator. Normalized value of the expression level relative to the calibrator is determined by the formula:

$$R = 2^{-\Delta\Delta Ct}$$

The final result will be a multiple of the calibrator concentration where one means no relative change against the calibrator (Livak and Schmittgen 2001). However, it is rarely managed to achieve the same performance of PCR reaction for the studied and the control gene, therefore a more appropriate approach is often to use one of the models that take into account the correction for this difference, for example, Pfaffl model (Pfaffl 2001). It becomes necessary to designate the efficiency of replication for all reactions. The proposed mathematical model, which includes the normalization with the reference gene is expressed as follows:

$$R = \frac{\left(E_{studied\ gene}\right)^{\Delta Ct\ studied\ gene(control-sample)}}{\left(E_{control\ gene}\right)^{\Delta Ct\ control\ gene(control-sample)}}$$

where:

E   amplification efficiency
ΔCt subtract of threshold cycle designated respectively for studied or control gene in the calibration and in the samples (Pfaffl 2001).

The result is also obtained as a multiple level of transcript in the sample against calibration.

Past use of reference genes as an important lesson

qPCR is one of the most rapidly incorporated technique in scientific studies in the last decade growing from ~8 % use to ~73–88 % application in mRNA quantification (Thellin et al. 2009). This leaves the question if the method used so broadly as a supplementary experiment was implemented properly and if all the authors despite the obvious need for normalization with reference genes approached it correctly. The genes used arouse the most controversy—a single gene often served as the only reference without verification of its stability under experimental conditions. That this is the wrong approach is confirmed by practice for several years now, but the beginnings were tough as presented by Suzuki et al. (2000); it was the case for more than 90 % of the articles concerning the analysis of RNA transcript that were published in 1999 in high impact journals. Moreover, the selection of HKGs was based on their precarious belonging to this group, because the classification was carried out mainly with qualitative methods (e.g., histochemical analysis, Northern analysis technique)—insufficient in the case of real-time RT-PCR, which is expected to obtain credible quantitative results (Gutierrez et al. 2008b). Another issue in this case is a common lack of an adequate validation of these genes by molecular biologists. The same authors emphasize that within 6 months, from July to December 2007 in three leading research journals in plant biology ("The Plant Cell", "Plant Physiology", "The Plant Journal") among 188 real-time RT-PCR analyses presented there, only 3.2 % of them were conducted with the appropriate validation of reference genes. Their choice was frequently based mainly on data acquired from earlier publications. A serious mistake in this approach is not to take into account that samples on which the stability of the expression was obtained and presented in these articles were collected in certain experimental conditions and are relevant only to them (Guénin et al. 2009). Although the problem was highlighted more than a decade ago, it is still a common methodological error made by researchers. Referring to data from reviewing articles and observations for which ones are most often subjected to validation, some examples of commonly used reference genes can be given (Thellin et al. 1999; Jin et al. 2004; Radonić et al. 2004; Huggett et al. 2005; Nicot et al. 2005; Hendriks-Balk et al. 2007; Guénin et al. 2009; Paolacci et al. 2009). Those might be: *18S rRNA* (18S ribosomal RNA), *28S rRNA* (28S ribosomal RNA), *TUBA* (α-tubulin), *ACTB* (β-actin), *β2M* (β2-microglobulin), *ALB* (albumin), *RPL32* (ribosomal protein L32), *TBP* (TATA sequence binding protein), *CYCC* (cyclophilin C), *EF1A* (elongation factor 1α), *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase), *HPRT* (hypoxanthine phosphoribosyl transferase), *RPII* (RNA polymerase II).

Taking a closer look at some research made on a certain topic and specified organism, a large variety of more or less relevant reference genes and methods applied for their selection can be observed. Confirmation of this may be studies using real-time RT-PCR that were carried out on barley published in the period of January 1996–March 2008 (Paolacci et al. 2009). Among the 26 reports examined, the authors found that there were used 16 different reference genes: most frequently it was *18S RNA* (eight times), *ACT* (seven times) and *TUBA* (five times). The most striking is that only four studies include the use of several genes and as many as 15 reports present the use of a single gene without proper validation. This illustrates how often in the past the issue of reference genes was not paid enough attention but for a scientist an error always should lead to an outcome of enlightening conclusions and not to be made again.

Some breakthrough moments for qPCR and reference genes were the 1st International qPCR Symposium in Germany (March 2004), continued further on in different countries as lead by Prof. Stephen Bustin, and the publication of MIQE Guidelines which goes for Minimum Information for Publication of Quantitative Real-Time PCR Experiments

(Bustin et al. 2009; Bustin et al. 2010; Derveaux et al. 2010). At the same time, implementation of MIQE goes with varying effectiveness, for example the report by Bandelj et al. (2013) was actually the first study concerning *Clostridium difficile* applying those guidelines and improving previous protocol from 2005. They achieved much greater precision and sensitivity to a large extent by testing various TaqMan universal PCR master mixes as suggested by MIQE. The guideline also encourages the use of validated assays like those available from RTPrimerDB as it helps with standardization (Bustin et al. 2010). Different RT chemistry and their influence on various starter sets was compared with MIQE standards by Jacob et al. (2013) which also showed to be not without significance.

GAPDH—one of the most commonly used reference genes

GAPDH is one of the most commonly used reference genes and a great majority of the most important scientific journals concerns its use through what is often referred as "classical" (de Jonge et al. 2007). The use of *GAPDH* in many studies brings good results, in others it is not recommended due to variability of expression caused by exhibition to the specified experimental factors. This casts into question its classification as a HKG, because this may suggest that it is involved not only in the fundamental processes of cells but might also be significantly influenced by other processes induced during the experiment.

The flagship example of how the level of *GAPDH* may vary within one type of organism were large scale studies conducted on 72 types of human tissues −1595 samples were collected from over 600 donors and a total of 5154 measurements were made (Barber et al. 2005). One of the biggest difference was observed for the skeletal muscle tissue and breast cells which was 14-fold. An example of similar analogy to human can be medaka (*Oryzias latipes*) set of tissues which showed organ-dependent *GADPH* transcript quantity; for example, the difference in expression between intestine (top) and muscle (the lowest) was approximately 5.45-fold (Zhang and Hu 2007). However, not any result can exclude the usefulness of *GAPDH* as a reference within the same tissue for sure but they underline the need of validation every time. The expression level of certain genes can vary in diseases, such as when you suffer from tuberculosis (Dheda et al. 2004; Dheda et al. 2005). Samples in this case were collected from blood, and the aim was to examine changes in expression of *IL-4* (in this case, a mediator of infection) after a 6-month treatment. It turned out the results obtained performing normalization with *GAPDH* were simply wrong, i.e., the number of transcripts of *IL-4* would remain unchanged (a false negative result for tuberculosis) for both healthy subjects and patients before treatment while after treatment their concentration would

increase, suggesting the aggravation of disease (false positive). The absolute need for validation of *GAPDH* as a reference gene can attest studies on adipose-derived stem cells (Fink et al. 2008). There were previous reports suggesting the usefulness of this gene in preadipocytes so the authors decided to include it as a candidate for the reference gene. Such an approach, which considers an existing successful reference gene application in a somewhat similar study as sufficient evidence is still a common practice. In fact, *GADPH* showed no changes in the amount of transcript in subsequent passages, but this was not the only criterion taken into account. Some cells were also grown under conditions of hypoxia (which ultimately was the object of research), while in others there were induced chondrogenesis, osteogenesis and adipogenesis. In such cases it was shown that during chondrogenesis and hypoxia a positive regulation of expression was observed, which eliminated the use of *GAPDH* as a reference gene in these conditions.

A dangerous way to include/exclude *GADPH* gene as a reference in a study is its categorization in a species-related manner. Exaggerating this statement this may even lead to a series of wrongly conducted studies of bad data migration and spread. To extend the above statement a quick look on the scientists, whose work domain are plants will be made: sometimes *GADPH* is chosen and often with very good application. Such satisfying results were obtained for *Coffea arabica*, which was used to test five different samples (Barsalobres-Cavallari et al. 2009). *GAPDH* proved to be the most stable among the evaluated genes, and showed no variation between tissues. Its equally high utility has been demonstrated for flax (*Linum usitatissimum* L.), where in addition to various tissues, the different stages of development of leaves and flowers were also taken into account (Huis et al. 2010). But the matter becomes more complicated when you compare different cultivars of the same plant species. In six cultivars of rice, *GAPDH* showed up to two-fold variation between samples (Kim et al. 2003), while among two cvs. of petunia (*Petunia hybrida* L.), the difference in stability between them was four-fold (Mallona et al. 2010). In the above-mentioned coffee, when taking into account the different leaf tissue of different cultivars *GAPDH* can still be considered as an optimal reference gene (Cruz et al. 2009).

Yet, there is another thing that almost always accompanies plant studies or sample acquisition: different environmental conditions, sometimes extreme, known as abiotic stress to which plant adapts in different ways and degrees. This may also have some influence on the expression of certain genes, including ubiquitous *GAPDH* not directly associated with such response. In the case of *Lolium temulentum* where a set of stress factors was applied *GAPDH* expression level increased under treatment of heat stress, or when exposed to UV light (Dombrowski and Martin 2009). On the other hand, exposing tomato to light stress treatment had no significant

effect on the expression of *GAPDH*, but low temperatures and lack of nitrogen were the source of such change (Løvdal and Lillo 2009).

A number of factors affecting gene expression can be multiplied, however it is possible to include as many conditions so that a single stress response affecting *GAPDH* can be a minor part of a whole not excluding its use, or vice versa: narrow application rules gene out. Lately a study of papaya samples set from 13 different conditions including even storage temperatures, postharvest ripening and atmosphere packaging showed that not only was *GAPDH* bad as a reference but it was one of the worst among 21 validated genes although limited scope of its use could be found (Zhu et al. 2012). All these examples illustrate how the usefulness of *GAPDH* as a reference gene can vary depending on the organism, tissue, diseases, and many other factors. Despite the prevailing opinion of this gene belonging to a group of HKGs, its expression does not have to be constant. The observed extremes of applications, from very successful to completely unrecommended, indicate the need for validation in each case separately. At the same time it seems as if *GAPDH* was as good as any other reference to be evaluated, why is it then it is being picked in most of the publications? This may be due to the fact its sequence is relatively conservative hence, it can be isolated as such even in non-model plants, like azalea, while other candidates may require degenerate PCR (De Keyser et al. 2013).

Gene encoding ribosome 18S rRNA subunit as a reference gene

Ribosomal subunits, with 18S as the most common among them, are widely used as normalization genes. It needs to be stressed that there are several difficulties associated with their use by which they have not supplanted other reference genes:

– Typically, the expression of rRNA is much higher than the target gene and its degradation is reduced compared with the mRNA (Paolacci et al. 2009). As a result, there is much more of it than transcripts of studied gene, which contradicts the theoretical assumptions for the reference genes.

– rRNA transcription is linked with the RNA polymerase I, while the mRNA uses RNA polymerase II (Radonić et al. 2004). This makes the control of synthesis of both types of RNA independent from each other.

– rRNA is absent in the purified mRNA fraction, so the total RNA must be used, where it stands a majority (Vandesompele et al. 2002). This involves a threat associated with the different rRNA:mRNA ratio between groups and samples. This was demonstrated by showing the differences in this ratio ranging up to 7.5 % in the case of mammary adenocarcinoma.

– Ribosomal subunits are not polyadenylated and will not take part in reverse transcription reaction if the cDNA has been isolated with the use of oligo(dT) primers (Radonić et al. 2004). In such cases it requires the use of random primers specific for rRNA by which other unwanted sequences can also be replicated.

– Normalization with their use does not include the efficiency of enzymatic reactions (de Jonge et al. 2007).

– It is indicated that their transcription may be regulated by some biological and chemical substances (Nicot et al. 2005).

– rRNA's of the same RIN/RQI integrity and quality values can contain mRNA's differing significantly in their integrity (Bustin et al. 2010).

To apply the use of *18S rRNA*, an endogenous standard with the predictable efficiency in the experiments that require oligo(dT) primers (which provide a single product as opposed to random oligomers), a method called coapplication reverse transcription (Co-RT) was developed (Zhu and Altman 2005; Kuchipudi et al. 2012). In this method, two primer sets are mixed in the initial reverse transcription reaction; these are oligo(dT) and *18S rRNA* sequence specific primers. Running two reactions in the same tube guarantees the same conditions (efficiency) and same enzyme (reverse transcriptase) for both so normalization is made properly. Zhu and Altman (2005) provided data that shows no cross-interference between reactions in the Co-RT approach.

Ubiquitous abundance of *18S rRNA* may lead to the situation similar to that where its transcript number was 5700-fold more than dataset's average (Tong et al. 2009). This requires dilutions that eventually may lead to greater errors. Thus this transcript is recommended to be used in studies where mRNA abundance can be high due to experiment factors. Interesting conclusions can also be reached by analyzing the two studies made on rice cultivars. Six cultivars (Dasan, Anda, Odae, Ilpum, Jukjinju, Hukjinju) were taken into consideration where *18S rRNA* proved to be the best reference gene (Kim et al. 2003) but a few years later for two other cultivars (9311 and Pai'ei 64S) gene validation has been made for developing seeds where it showed very high expression variability both among cultivars and their seeds (Li et al. 2010). In this case some uncertainty can arise if different conclusions are because of experimental design (cultivars), a small number of other candidates or experiment protocol lacking the approach of validation in older papers. It is therefore necessary to re-evaluate the usefulness of this gene whenever new cultivars are investigated, even if it had not previously appeared to have variable stability. The fact that *18S rRNA* is not the best choice for plants if studying their embryonic development also seems to confirm the research made for longan tree (*Dimocarpus longan* Lour.) during somatic embryogenesis (Lin and Lai 2010). This gene showed a high variability which is associated

with cell differentiation, forming organs and their maturation. Examples could be multiplied, not bringing too much confusion we will remain in the plant kingdom where this gene was recommended as reference for example for rice cultivars (Kim et al. 2003) and potato (Nicot et al. 2005) while going further in time (as it seems) it was generally not recommended for plants like peach (Tong et al. 2009), rice (Li et al. 2010), cucumber (Wan et al. 2010) or papaya (Zhu et al. 2012). Some statements can be risked that over time *18S rRNA* looses the advantage of first choice as a reference over other validated genes which prove themselves to be much more stable. This does not necessarily mean it is bad as reference but often there are better choices that can be made avoiding the problems of, e.g., high transcript abundance and others mentioned earlier.

The variability of gene expression levels under different factors

The issue previously signaled for two genes of HKGs group: *GAPDH* and *18S rRNA* is obviously not specific only to them. A number of further discussed factors may have different impacts on the expression and apply in varying degrees for different genes. This information is particularly important because it results in the absolute need in every study to analyze the factors which may affect the reference gene expression and, consequently, the result of the measurement obtained with the real-time RT-PCR.

Type of tissue

There is no need to recall previously quoted examples for inter tissue HKG's variable expression patterns but what may not seem so obvious is a fact that variability is possible via intra tissue. Such differences can be observed within the same tissue performing its physiological functions, as is evidenced in a muscle during physical exercise (Jemiolo and Trappe 2004). In isolated muscle fibers expression levels of *β2M*, *GAPDH*, *ACTB* and *18S rRNA* were tested. It turned out that discrepancies in the amount of transcripts can be up to 52-fold for *β2M* normalized with *18S rRNA* in relation to the tissue before exercise and 4 h afterwards. The most stable gene in this case was *GAPDH* and it could only show a significant variation when normalized to *β2M*. Indeed, as it is recently stressed genetically identical cells exposed to uniform environment show significant variation in mRNA quantity due to different physical localization of mRNA (Bustin 2010). Hence, the most important is to obtain such a sample that will be representative in certain conditions.

Developmental stage

Development of the organisms, differentiation and cell growth induces a change in the number of transcripts of many genes. Some of the developmental processes have such an influence on the phenotype that it must go along with major impact on HKGs expression. To name some examples for plant studies we may refer to somatic embryogenesis on a model plant, longan tree which is interesting because it included fluctuations in temperature (Lin and Lai 2010). When *18S rRNA* gene was taken into consideration as the reference, the differences between the samples were significant but embryo stages for which there was up-regulation and for which down-regulation of expression could not even be clearly defined, since the amount of the transcript was also highly dependent on the temperature. While at 25 °C, the first and the last stages were characterized by the lowest expression of this gene, at 20 °C and 30 °C the results were opposite. Other cross-section development studies may concern days following the flowering (Li et al. 2010) and flower and leaf development (Mallona et al. 2010). The development goes on also after harvesting affecting HKG expression as reported for storing peach and postharvest ripening papaya, respectively (Tong et al. 2009; Zhu et al. 2012).

Growing conditions not only determine future phenotype but dynamically affect HKG's. Rushing with an explanation we will use *Saccharomyces cerevisiae* as an example. It should be noticed that it must constantly adapt to changes in the environment induced by a progressive increase in cell number and nutrient depletion (Teste et al. 2009). These yeast show different expression levels of certain genes, depending on the current growth phase and their physiological state. Although gene expression may be stable in a certain measure point there is no clear pattern of how it will adjust in the next stage of growth as well as there is no evidence to predict if those changes will be positive or negative in general for examined housekeeping genes. In this case a different stage of yeast growth (i.e., early exponential growth stage) can be set as the calibrator to trace such expression changes. To set an example: *PDA1* (pyruvate dehydrogenase subunit E1 alfa) showed a several fold increase in transcript number when the organism consumed glucose and was switching to metabolism of another sugar. Then, having entered into the next phase expression level was down-regulated, until stationary phase, where the amount of mRNA of this HKG was more than ten-fold lower than for the calibration sample. Such regulatory scheme: first up-regulated, then down-regulated is obviously not the same for each yeast gene. To contradict, *IPP1* (Inorganic pyrophosphatase) shows only a negative adjustment in relation to the calibrator when *GPH1* (glycogen phosphorylase) is always positive. To address the issue away from microbial *in vitro* cultures we refer to study where genes were validated in *Heterobasidion annosum* grown on three

different substrates: pine bark, pine heartwood, and pine sap-wood (Raffaello and Asiegbu 2013), plants treated with different hormones: salicylic acid, methyl jasmonic acid, abscisic acid (Wan et al. 2010) and the influence of auxin over time after application (de Almeida et al. 2010).

The need for separate validation of reference genes due to different developmental stage and specific set of tissues is perfectly highlighted by the example of two separate experiments designed for the same group of 14 genes in *Arabidopsis thaliana*, but with a different set of tissues including both vegetative and generative organs (Gutierrez et al. 2008a). They consist of such samples as: floral buds, inflorescence, open flowers, siliques at different times after flowering as well as older and younger leaves and root tissues. Some HKGs expression patterns were different in a second study involving the same set of genes, the same plant, but a different set of tissues, which excluded siliques at different times after flowering (Guénin et al. 2009).

A very important issue is *in vitro* vs. *in vivo* studies which should reflect actual development processes occurring in the natural environment. Such comparison was carried out for grass *Lolium perenne* L. by comparing 13 plants grown *in vitro* with 422 leaf samples obtained in April, July and October with the laboratory results consisting of four sets of samples collected at the peak of each season (Lee et al. 2010). As expected, HKGs stability rankings were very different for each case. Greenhouse vegetation compiled with *in vitro* cultures also bring discrepancies (Podevin et al. 2012).

Related species

To extend previously mentioned variability in plant cultivars from different points of the view and applications we will make a combination of two corresponding rodent studies. For such non-model animals validation of reference genes is often based on previously obtained databases for species that are close relatives to them, because a high degree of similarity between genomes gives the chance that the expression may be the same or on a similar level. This approach was used in the two research works concerning rodents. For wild yellow-necked mouse (*Apodemus flavicollis*), six sets of primers from RTPrimer database originally developed for lab mice (*Mus musculus*) were used (Axtner and Sommer 2009). It turned out that none of them gave amplification products. However, they appeared when primers were designed based on the Mouse Normalization Gene Panel database for eight other reference genes. Despite the heterogeneity of wild specimens in terms of age or physiological condition, five of the genes originally used for laboratory mice (*Mus musculus*) were characterized by stability of expression and could provide the correct normalization (those were: *Rps18*, *Sdha*, *Canx*, *Pgk1*, *Actg1*). Genes such as *UBC*, *Rp113* and *Actb* were not applicable here because of too large a variability. For the endemic Brazilian rodent *Delomys sublineatus* a study was also designed based on the selection of primer pairs among the second database, and for six out of nine genes the expected reaction products of RT-PCR were achieved (Weyrich et al. 2010). It was a similar set of genes as in the previous species: *Rps18*, *Actg1*, *Sdha*, *Actb*, *Pgk1* and *Canx*, but the first one showed presence of nonspecific products during melting. All these genes prove to be a good reference with the exception of *Actg1*. This example indicates that it is possible to find common reference genes for relatively close species of rodents. For such universal (for those cases) normalizing gene could stand *Sdha*, classified as the best in the first publication and as the second best in the other.

Abiotic stress

Due to the specificity of their growth in changing environment, plants had to develop their own responses to the periodically occurring extreme adverse conditions. This involves changes in complex regulatory networks and stress regulated genes for which changes in the transcript level are identified with the use of real-time RT-PCR. It is important that the reference gene is not subjected to the same mechanisms of regulation, which is induced by environmental stress in the case of resistance genes, since it will be impossible to determine the actual increase (or decrease) of transcripts quantity (Paolacci et al. 2009). This is not always easy to determine, since seemingly not being directly related to plant stress response, HKGs may show co-regulation with transcription factors which affects often a wide range of genes. It seems that the catch lies in stress definition itself because, for example the temperature does not necessarily need to reach extreme values, and even so it will significantly affect reference gene expression, as was discussed in the study of somatic embryo-genesis (Lin and Lai 2010). It is rarely seen to acknowledge information if samples were acquired in the same temperature, soil water content (so if watering and evaporation had same effect in the specified moment), light exposition etc.—in other words if it was precisely in the same set of stress factors (unless they are included as the validated stress factor for reference gene sets).

It is remarkable how *ACT* gene found its place in the stress-related studies. At the same time by following previous papers some general trends can be drawn that may decide if to include specified gene or not. Focusing on salinity stress and difficulties in water extraction from the soil—for three different plants: potato (Nicot et al. 2005), *Lolium temulentum* L. (Dombrowski and Martin 2009) and cucumber (Wan et al. 2010), *ACT* gene was ranked in the last place in the ranking of stability every time, which, to some extent, may indicate the relationship between the plant response to such stress and the expression of *ACT*. Careful validation of *ACT* under drought

treatment in barley was proved to be stable in control and stressed plants but at the same time highlights that expression pattern was dependent on developmental stage and *ACT* was recommended as the best reference gene only during heading (Rapacz et al. 2012). It did not cross out its previous use for drought stressed barley at seedling stage due to less substantial water deficit in leaves than in previously quoted study and still was characterized by relatively high stability (Rapacz et al. 2010; Wójcik-Jagła et al. 2012). *ACT* was also successfully adapted as reference when investigating effects of cold, light and time of day during low temperature shift and other abiotic stresses in *Festuca pratensis* (Jurczyk et al. 2012; Pawłowicz et al. 2012) or cold acclimation of four genotypes of barley (Rapacz et al. 2008).

Changes in gene expression can even be caused by plant wounding, therefore, collected samples should be immediately frozen. In *L. tumulentum* Ct value for *CAP* gene was reached about two cycles earlier when samples were analyzed 11 h after wounding than in unwounded tissues, but this difference decreased after 24 h (Dombrowski and Martin 2009). The same stress can also cause different changes in various tissues, as observed in coffee exposed to drought (Cruz et al. 2009). Also nutrient deficiency, e.g., nitrogen shortage may have an effect on the expression of potential reference genes and their stability which was presented in tomato for *TUB* gene (Løvdal and Lillo 2009).

Diseases and infections

An example when defense response and burden of illness have no effect on any validated HKG's expression is impossible to find. Patients are classified in the same group as healthy or ill, while they differ on the gene expression level as a result of health deterioration. Nowadays, a very common physical disorder is obesity and the metabolic syndrome (Mehta et al. 2010). Symptoms can have somewhat smooth margins that do not guarantee a strict threshold above which certain influence on gene expression is observed. To examine how they affect the number of transcripts of some genes, adipocyte cells were taken from healthy and morbidly obese patients. It turned out that significant differences were noticed in the expression of many HKG's including the *18S rRNA*. Concluding the fact that tissue cells can differ from one another just like individual patients it appears that the most appropriate term to use would be "reflects a snapshot of mRNA" which is hopefully most representative for certain cases (Bustin 2010; De Keyser et al. 2013). Still such may not be relevant to some rare unique genotypes. Genotypes that not only refer to patient's sample but disease source itself which may be influenza viruses showing broad diversity in their nature. Kuchipudi et al. (2012) searched for best performing reference genes for five subtypes of influenza A virus where depending on subtype stability of a single gene was varying. Having in mind how

dynamic in evolving viruses are it casts into question if reference gene evaluation should not be performed prophylactically from time to time even when working on the same matter for a longer period.

Studies in plants allow to answer if a common infection influences HKG's expression in the same way in relatively closely related species. BYDV (barley yellow dwarf virus) affects the expression stability of genes previously used as reference in wheat *Triticum aestivum* L., barley *Hordeum vulgare* L. and oat *Avena sativa* L. (Jarošová and Kundu 2010). It turned out that these differences were dependent on the species; in the case of *TUB* gene, the greatest instability was observed in wheat, while in the other two species, this gene proved to be reasonably stable. In the case of other genes there were probably inter-specific similarities in regulation, and similar acceptable stability was determined for *GAPDH* and *18S rRNA*. So a set of evaluated HKGs is not the best way to predict to what extent any other gene expression will be affected.

Tumors

Cancer studies stress a different problem: it is crucial to make biopsy representative for malignant tissue while it will still remain unclear if we can predict tumor's behavior based on the results obtained this way (Bustin 2010). Also some objections can arise toward inconsistent use of controls. There can be made compilations of different and similar cancer studies, where expression patterns of HKGs can be comparable among the organs or totally different even in the same type of carcinoma (Andersen et al. 2004; Rubie et al. 2005; Chari et al. 2010; Chervoneva et al. 2010). This illustrates the difference in the metabolism of cancer cells when different cases (even concerning the same organ) are compared. Differences in the results of reference gene evaluation are also associated with distinct validation methods used by the authors. Just to mention that Andersen et al. (2004) using the method of comparison of gene pairs and model that takes into account inter- and intra-group variations come to different conclusions. Applying latest qPCR validation standards may even extend differences. Based on the analysis of colon cancer samples, it was concluded that reducing or increasing gene expression is closely associated with changes in the structure of chromosomes (Tsafrir et al. 2006). This may explain for same cases differences in the amount of transcripts of certain genes in the same type of cancer, because changes associated with the loss or the appearance of extra copies of chromosomal DNA occurred in varying degrees in tested samples and their severity was associated with the type of cancer and its progression in the organism. It seems that cancer-related studies are most case-specific.
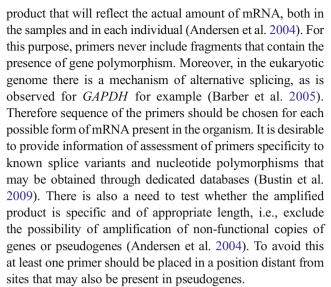
Alternative splicing

Many genes of multicellular organisms are alternatively spliced producing various protein isoforms which may be specific for certain tissue or developmental stage. To be precise it is often more advisable to say that expression level is measured for 1–2 exons instead of an entire gene (Bustin et al. 2009). To address the issue, it was recently proved that alternatively spliced domain of *NDC1* (NAD(P)H dehydrogenase) gene, whose expression is relatively unresponsive to stress treatments, strongly affects the expression of the *ACTIN2* reference gene in *A. thaliana* (Wallström et al. 2012). *NDC1* expression is induced by light and is present in all plant organs but ratio of two alternatively spliced mRNAs (*NDC1-1* and *NDC 1-2*) is different in various tissues. *NDC1-2* consists of early stop codon because of additional nucleotides from intron 5 resulting in frameshift but it does not cause a major degradation of the transcript. T-DNA insertions in intron 5 which disrupted the reading frame of *NDC1-2* proved that such change increases *ACT2* expression by three to four-fold. This suggests an effect on the signaling paths for *ACT2* gene expression and interactions between genes from separate functional domains. This example strongly emphasizes the complexity of signaling paths and underlines a caution approach to every studied genotype and reference gene, even such as *ACT2* which is commonly used due to pair of primers commercially sold.

Validation of reference genes

The absolute need of reference genes validation was underlined several times in this paper as was supported with numerous examples. This is necessary for proper normalization whose task is to compensate for the intra and inter-kinetic real-time RT-PCR variations that result from basic difficulties connected with the method (Chervoneva et al. 2010). The fundamental idea is to find a gene that will be characterized with undoubted stable expression, i.e., one that will be the same in every tissue and individual sample and will meet the criteria described previously for the reference genes. Several statistical approaches and algorithms were developed that enable identification of the most stably expressed gene. Validation is of course impossible for each individual gene present in the genome and can be carried out only for a few to several genes previously selected. That is why an important step is the selection of genes whose expression stability will be verified.

Selection of genes

First, structure and sequence of candidate genes should be recognized so the designed primers would always amplify the product that will reflect the actual amount of mRNA, both in the samples and in each individual (Andersen et al. 2004). For this purpose, primers never include fragments that contain the presence of gene polymorphism. Moreover, in the eukaryotic genome there is a mechanism of alternative splicing, as is observed for *GAPDH* for example (Barber et al. 2005). Therefore sequence of the primers should be chosen for each possible form of mRNA present in the organism. It is desirable to provide information of assessment of primers specificity to known splice variants and nucleotide polymorphisms that may be obtained through dedicated databases (Bustin et al. 2009). There is also a need to test whether the amplified product is specific and of appropriate length, i.e., exclude the possibility of amplification of non-functional copies of genes or pseudogenes (Andersen et al. 2004). To avoid this at least one primer should be placed in a position distant from sites that may also be present in pseudogenes.

Good output to work on can provide web tools for primer design: PrimerQuest software (Integrated DNA Technologies, http://eu.idtdna.com/scitools/applications/primerquest/default.aspx), Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) or Primer-Blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/), a program developed by NCBI that uses the algorithm Primer3 (D'haene et al. 2010; Taylor et al. 2010). It is desirable to keep GC content in primers close to 50 % with primer length about 20 bp as the most optimal length. A huge database of previously used qPCR primers can be found in RTPrimerDB (http://www.rtprimerdb.org, Lefever et al. 2009). Therefore, primers should be validated *in silico* for possible secondary structures (e.g., MFOLD program: http://mfold.rit.albany.edu/), amplicon length, single nucleotide polymorphisms (SNP) etc., which to some extent can provide NCBI BLAST tool.

If the studied material is a cancerous tissue, it is possible that additional or missing chromosomes may be present in the genome, so it is advisable to examine also the genotype of tissues to exclude the presence of additional copies of the gene (Tsafrir et al. 2006). It was also shown that use of random hexameric primer sequences for reverse transcription may overstate the actual amount of mRNA up to 19 times, therefore more accurate results might be obtained by using primers with sequences specific to a given area (Bustin et al. 2005).

Ideally, the reference gene should have an identical expression level as the tested gene, but it is not always possible so the general guidelines do not recommend that it would be very low (Ct>30) or very high (Ct<15) (Wan et al. 2010). Otherwise, there may occur a situation where starters are more likely hybridized to a sequence representing a larger number of gene copies (Bustin 2000). Some of those rules might have been omitted in the past, just like Pettengill et al. (2012) indicates that a report by Brunner et al. (2004) lacks information about PCR efficiencies for primers and amplification product length.

A set of genes for validation is often selected with the use of tools that enable global analysis of transcripts, such as chips and microarrays (Czechowski et al. 2004). They are only able to perform a qualitative analysis, since there is no strict relationship between signal intensity and the amount of transcript. However, if gene had a high level of expression in the microarray analysis it might be similar for the real-time PCR. Such a study should also include experimental conditions, since all major changes in gene expression are shown in these methods and they mainly concern genes that have a large number of transcripts in the cells. Analysis of potential *Arabidopsis* reference genes was based on Affymetrix chips, they were selected for further validation by meeting preset criteria: detection signal strong enough and small standard deviation for the measurement of expression. Because microarrays are based on hybridization and are highly qualitative, an alternative for them can be a SAGE technique (Serial Analysis of Gene Expression), that was applied in research of *Lolium perenne* L. (Lee et al. 2010). SAGE library made of concatamers constructed from short genome fragments of plants grown in the field, allowed the identification of genes among the tags with expression moderately stable between seasons and a similar expression level.

Studying non-model organisms, the gene database of the model plant, such as that published for *Arabidopsis*, can be a good starting point (Gutierrez et al. 2008b). This approach was applied in the study of *Brachiaria brizantha*, where BLAST tool was used to search for genes and sequences with high complementarity to the previously used reference genes of other organisms (Silveira et al. 2009). BLAST is obviously one among many sources where search of expressed sequence tag (EST) and gene sequences can be performed.

Many plants and animals have their own dedicated databases, such as wheat (Paolacci et al. 2009). In this case, the database searched was dbEST (NCBI), *Triticum aestivum* UniGene and TIGR. In designed *in silico* method, for 41.256 found Unigene clusters there was estimated a number of transcripts using program "The ProfileViewer" which compared nested ESTs to their total number derived from cDNA libraries from different tissues (ten in this case) and calculated the value of transcripts per million (TPM). Then the selection was performed based on three criteria: each Unigene cluster having at least 60 single ESTs, had to show expression in each tissue and TPM value should not be less than 40 % for the six most representative tissues. For the 177 Unigene clusters selected using the TPM values, there were calculated mean values, variance, standard deviation and coefficient of variation for the six tissues. Next, the selected clusters were introduced to Unigene BLAST tool that searched TIGR wheat gene database and final work was performed. The next step after selecting a group of candidate genes is to subject them to detailed statistical analysis, in which the most stable ones will emerge from the tested reference genes.

Validation

The basic assumption of reference genes is that they should be characterized by a permanent and unchanging expression in each of the samples tested, despite the impact of experimental factors. However, in practice there is always observed variability (or variance) of the obtained Ct values data and the key in this case is the selection of genes with the most stable expression, or showing the least deviation from the mean (Vandesompele et al. 2002). As in the case of the reference genes we have no reference point, so Vandesompele et al. (2002) suggested that the stability index would be calculated by the average variation among pairs of genes by comparing control gene with other validated candidates. In this case gene-stability measure, M is defined as the standard deviation of the logarithmically transformed expression values of the compared genes. This stability index together with its decreasing value gives rise to the rank and stability of gene. The authors have developed a tool based on Microsoft Excel, geNorm, which performs the appropriate calculations and is freely available. In the first stage the algorithm selects a pair of genes, which have the lowest value of M, and further genes are classified based on the highest degree of compatibility with the other and with a geometric mean of the first pair.

Usually one gene is not enough to complete the normalization of satisfactory assurance. Therefore the same authors recommended that the next step should include the calculation of the variation between successive compared pairs, that is normalization factor (NF), a degree determining the influence of the subsequent reference gene added to the others. Such comparisons are made by confronting an NF_n for n number of genes with an NF_{n+1} containing the same set of genes with one additional that was next after them in the stability ranking. When variation $V_{n/n+1}$ between two NF factors is high, this means that another gene added has a significant impact on the quality of normalization. This value is set by the researcher and in many studies the number of genes is most often determined by achieving variation below 0.15. Currently geNorm's use makes it a standard program and according to Google Scholar more than 6300 papers cited this method as was provided by the official site (http://medgen.ugent.be/~jvdesomp/genorm/—August 2013). On the other hand it seems that a better approach is to select genes based on linear models describing the geometric mean of expression as a complex set of factors, such as gene expression in a group, the amount of mRNA in the sample and the random variation caused by biological and experimental factors (Andersen et al. 2004). Such additive models allow to apply different statistical approaches like classical ANOVA model and other methods for mixed linear models consisting of observations and independent variables. Taking into account the average influence of gene within tissue (group) and individual impact of a gene, the authors calculated intra and intergroup variations, which
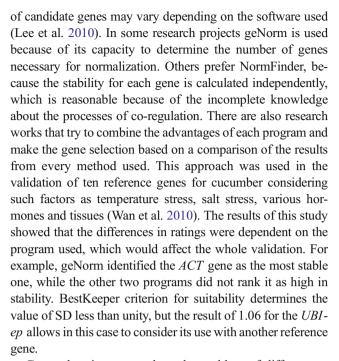
were combined into a stability value. This algorithm was used in NormFinder software. In order to use ANOVA model and the F-test assumptions concerning homogeneity of variance and normality of data must be fulfilled (Brunner et al. 2004). Yet, today's advanced statistical programs supplanted such approach as pointed by Pettengill et al. (2012) who also studied *Populus* implementing MIQE standards. Various models have also been developed by Szabo et al. (2004), including one which considers heteroscedasticity, i.e., this model is in contrast to the classical ANOVA model and takes into account the case where random variables have different variances. The authors used a statistical calculation program statistical analysis software (SAS), and received a similar stability ranking as was obtained using geNorm program.

Chervoneva et al. (2010) constructed their model in a way which enabled a choice between certain sets of genes in terms of different criteria: reduced variability of NF, the use of the smallest possible number of genes meeting the limits of variability of NF or minimized average variance of the NF. The approach is based on estimating the covariance matrix of all validated genes of which on the basis there is calculated the variances of log NF relating to all possible sets of genes. A popular tool for validation of reference genes is also BestKeeper, a software developed by Pfaffl et al. (2004). Evaluation is achieved based on the Ct values and calculated for them: the geometric mean, arithmetic mean, minimum and maximum value, standard deviation and coefficient of variation. There is also defined the x-fold over- or under-expression toward the geometric mean of measurement (Ct) that takes into account the PCR efficiency (computed either as sample specific or as factor specific). On the basis of the calculated variation (SD and CV) genes are ranged, where the most stable is characterized by the lowest result, and as a reference can be used only those that achieved a value of SD below one. Variability between genes is estimated based on comparisons of gene pairs for which there are determined the Pearson coefficient of correlation (r) and probability value. For genes indexed in this way there is calculated the correlation between the candidate gene (i.e., the geometric mean of the obtained Ct values) and calculated index (BestKeeper), whose relation is expressed, beyond the previous two determinants, also by the determination coefficient (r^2). However, beyond the evaluation of reference genes the program does not allow necessary estimation of the minimum number of genes for normalization. The described procedure is also carried out in the further analysis of target genes. Although there is no clear evidence to point which approach is appropriate, geNorm has been used a couple times more than other programs (Pettengill et al. 2012).

Comparison of validation programs

After reviewing the algorithms for programs such as geNorm, NormFinder or BestKeeper it is not surprising that the rankings

of candidate genes may vary depending on the software used (Lee et al. 2010). In some research projects geNorm is used because of its capacity to determine the number of genes necessary for normalization. Others prefer NormFinder, because the stability for each gene is calculated independently, which is reasonable because of the incomplete knowledge about the processes of co-regulation. There are also research works that try to combine the advantages of each program and make the gene selection based on a comparison of the results from every method used. This approach was used in the validation of ten reference genes for cucumber considering such factors as temperature stress, salt stress, various hormones and tissues (Wan et al. 2010). The results of this study showed that the differences in ratings were dependent on the program used, which would affect the whole validation. For example, geNorm identified the *ACT* gene as the most stable one, while the other two programs did not rank it as high in stability. BestKeeper criterion for suitability determines the value of SD less than unity, but the result of 1.06 for the *UBI-ep* allows in this case to consider its use with another reference gene.

Comprehensive approach to the problem of different results being dependent on the algorithm was presented by Mallona et al. (2010), who used four available validation programs (NormFinder, BestKeeper, qBasePlus, geNorm). For each the required values and batch data were calculated and the results obtained with these different statistical approaches were combined to obtain an aggregated stability list with the use of Monte Carlo algorithm together with calculated Spearman footrule distances conducted by RankAgregg program. Concerning the influence of experimental conditions, most of the programs identified *GAPDH* as the least stable gene, regardless of the analyzed petunia line. Unified ranks summarized together in the graph with the results of different statistical approaches, i.e., stability on the basis of the M value of program geNorm, the stability calculated by the NormFinder, stability expressed in the M value calculated with qBasePlus, CV values obtained from calculation of qBasePlus and determination coefficient (r^2) obtained with BestKeeper, clearly illustrate how they can vary greatly in comparison to the individual rankings. For example, *CYP* gene was located on the last position in the V30 petunia line with NormFinder and on the first when combining the results from each program. Analyzing the differences in gene evaluation with each program it is difficult to identify the most reliable one. Certainly, the results of standard PCR experiment would be significantly different if, for example, one of the researchers guided by the results of BestKeeper would apply *CYP* and *EF1A* genes for the V30 line (being on the last and seventh position, respectively in the ranking received with NormFinder) and the other chosen genes *UBQ* and *RAN1* which were suggested by the NormFinder (the fifth and the ninth location in BestKeeper ranking, respectively).

A similar study was performed for drought-stressed barley at different developmental stages and leaf age (Rapacz et al. 2012). The stability was analyzed using qBase^PLUS-geNorm^PLUS, NormFinder and BestKeeper. Obtained rankings showed again how different models can affect reference gene order where, for example, *L-TUB* was bottom ranked in stressed plants at the seedling stage according to qBase and NormFinder, while BestKeeper put it in second place. Another study in maize showed satisfactory and similar performance among different softwares (in this case: geNorm^PLUS, NormFinder and BestKeeper) in evaluating suitability of reference genes (Manoli et al. 2012). All three algorithms identified the same top five reference genes (although with mismatched order) for different maize tissues and samples collected during various experimental conditions. Yet, a high correlation between programs is a dangerous shortcoming. Jacob et al. (2013) pointed out that in their study correlation between geNorm and NormFinder was high ($r = 0.9$) while ranking was identical only for five out of 12 genes.

Following different results of different algorithms, when it is not certain if the right reference gene is used, it is advisable to use at least the pair of genes responsible for distant functions, because of a very little chance for a common regulation of their expression. In order to achieve best results at least three reference genes should be used (Bustin et al. 2010; Derveaux et al. 2010), three different validation programs (Jacob et al. 2013) and take three samples (biological replicates) with three repeats for each genotype (Pawłowicz et al. 2012; Rapacz et al. 2012). Altogether those can be placed under the common rule of "Best 3".

Conclusions

Expectations for real-time RT-PCR technique impose an extremely carefully made research, which may be a reference for many other techniques. Thousands of studies are base on qPCR data while in fact only a small number of *de novo* studies are being developed. To obtain reliable results it is necessary to carry out the process of normalization with reference genes and to interpret them rationally. It is always a complicated matter to point out if a two-fold variation in gene expression is already of biological importance because the genetic variability discussed above may trigger intrinsic stochastic kinetic noise of biochemical reactions. A single cell is always unique, so a snapshot of mRNA in current experimental conditions is what we get—it is crucial to relate to proper experimental context and studied object(s).

In the past, validation process was often avoided and HKGs were used due to a common belief that they are characterized by constant expression level regardless of the conditions and origins. As awareness of the complex expression regulation networks in the cell function grew, this statement begun to be undermined and experimental confirmation of the stability of candidate genes is now a standard requirement. This has been repeatedly shown, among examples for *GAPDH* and *18S rRNA* genes. At the same time it is highly desirable to support results with regulatory RNAs studies, protein levels and their activities. In many reports can be found a fundamental rule that there is no universal reference gene and when analyzing dozens of cited examples for expression variability between the tissues, caused by stress factors or tumors and diseases, it is difficult to disagree with this statement. What is more, living organisms as dynamic creations and constantly adapting to changing conditions, exhibit different expression profiles of HKGs in laboratory and field conditions. Therefore, there is emphasized an absolute necessity of a unique approach to each experiment and validation of candidate reference genes, even their careful selection, since it is not possible to carry out this process for all genes in the cell. It is obligatory for every study, even those similar to the previously performed, because it may vary, for example in an additional type of tissue acquired for research or occurrence of some additional factors. Even differences in chemistry used for RNA isolation and RNA purification method are sufficient enough to reevaluate the whole protocol.

Previous studies can be a perfect starting point for selection of reference gene candidates. The knowledge of transcription regulation is still not enough to develop a reliable statistical model to analyze the expression of candidate reference genes and therefore many different approaches are in use. Some authors rightly point out the lack of benchmark for the reference genes and they developed models that examine the variability of all possible gene pairs or those that are based on a linear additive model consisting of observations and independent variables. Gene expression stability rankings obtained by means of these models are often very different and it should always be a priority to find such a reference gene that is characterized by a high stability not only for a single algorithm. Despite a certain amount of uncertainty associated with the use of genes for the normalization, real-time PCR still remains one of the most accurate methods for transcript quantification. Yet, during review of the literature the feeling accompanies that most of the research works make their protocols only on the basis of previous achievements trusting the statistics of successful application in the majority of cases. Paraphrasing a joke about a man who drown in a river that had in average only 1 m depth, without understanding of validation algorithms and reference genes diving in scientific literature about qPCR can end up in wrongly designed experiments in spite of applying general rules and following well worn protocols. So it would be somewhat right to say that qPCR results are most dependent on user following the rule: "garbage in—garbage out".

References

Andersen CL, Jensen JL, Ørntoft 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 Res 64: 5245–5250

Axtner J, Sommer S (2009) Validation of internal reference genes for quantitative real-time PCR in a non-model organism, the yellow-necked mouse, *Apodemus flavicollis*. BMC Res Notes 2:264

Bandelj P, Logar K, Usenik AM, Vengust M, Ocepek M (2013) An improved qPCR protocol for rapid detection and quantification of *Clostridium difficile* in cattle feces. FEMS Microbiol Lett 341:115–121

Barber RD, Harmer DW, Coleman RA, Clark BJ (2005) *GAPDH* as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. Physiol Genomics 21:389–395

Barsalobres-Cavallari CF, Severino FE, Maluf MP, Maia IG (2009) Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. BMC Mol Biol 10:1

Brunner AM, Yakovlev IA, Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. BMC Plant Biol 4:14

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

Bustin SA (2008) RT-qPCR and molecular diagnostics: no evidence for measles virus in the GI tract of autistic children. Eur Pharm Rev Dig 1:11–16

Bustin SA (2010) Why the need for qPCR publication guidelines? The case for MIQE. Methods 50:217–226

Bustin SA, Benes V, Nolan T, Pfaffl MW (2005) Quantitative real-time RT-PCR—a perspective. J Mol Endocrinol 34:597–601

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622

Bustin SA, Beaulieu JF, Huggett J, Jaggi R, Kibenge FS, Olsvik PA, Penning LC, Toegel S (2010) MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. BMC Mol Biol 11:74

Chari R, Lonergan KM, Pikor LA, Coe BP, Zhu CQ, Chan THW, MacAulay CE, Tsao MS, Lam S, Ng RT, Lam WL (2010) A sequence-based approach to identify reference genes for gene expression analysis. BMC Med Genomics 3:32

Chervoneva I, Li Y, Schulz S, Croker S, Wilson C, Waldman SA, Hyslop T (2010) Selection of optimal reference genes for normalization in quantitative RT-PCR. BMC Bioinforma 11:253

Cruz F, Kalaoun S, Nobile P, Colombo C, Almeida J, Barros L, Romano E, Grossi-de-Sa MF, Vaslin M, Alves-Ferreira M (2009) Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. Mol Breeding 23:607–616

Czechowski T, Bari RP, Stitt M, Scheible WR, Udvardi MK (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

D'haene B, Vandesompele J, Hellemans J (2010) Accurate and objective copy number profiling using real-time quantitative PCR. Methods 50:262–270

de Almeida MR, Ruedell CM, Ricachenevsky FK, Sperotto RA, Pasquali G, Fett-Neto AG (2010) Reference gene selection for quantitative reverse transcription-polymerase chain reaction normalization during in vitro adventitious rooting in *Eucalyptus globulus* Labill. BMC Mol Biol 11:73

de Jonge HJM, Fehrmann RSN, De Bont ESJM, Hofstra RMW, Gerbens F, Kamps WA, de Vries EGE, van der Zee AGJ, te Meerman GJ, ter Elst A (2007) Evidence based selection of housekeeping genes. PLoS One 2(9):e898

De Keyser E, Desmet L, Van Bockstaele E, De Riek J (2013) How to perform RT-qPCR accurately in plant species? A case study on flower colour gene expression in an azalea (*Rhododendron simsii* hybrids) mapping population. BMC Mol Biol 14(1):13

Derveaux S, Vandesompele J, Hellemans J (2010) How to do successful gene expression analysis using real-time PCR. Methods 50:227–230

Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A (2004) Validation of housekeeping genes for normalizing RNA expression in real-time PCR. BioTechniques 37:112–119

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. Anal Biochem 344:141–143

Dombrowski JE, Martin RC (2009) Evaluation of reference genes for quantitative RT-PCR in *Lolium temulentum* under abiotic stress. Plant Sci 176:390–396

Fink T, Lund P, Pilgaard L, Rasmussen JG, Duroux M, Zachar V (2008) Instability of standard PCR reference genes in adipose-derived stem cells during propagation, differentiation and hypoxic exposure. BMC Mol Biol 9:98

Gachon C, Mingam A, Charrier B (2004) Real-time PCR: what relevance to plant studies? J Exp Bot 55:1445–1454

Guénin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L (2009) Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions specific, validation of references. J Exp Bot 60(2):487–493

Gutierrez L, Mauriat M, Guénin S, Pelloux J, Lefebvre JF, Louvet R, Rusterucci C, Moritz T, Guerineau F, Bellini C, Van Wuytswinkel O (2008a) The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. Plant Biotechnol J 6:609–618

Gutierrez L, Mauriat M, Pelloux J, Bellini C, Van Wuytswinkel O (2008b) Towards a systematic validation of references in Real-Time RT-PCR. Plant Cell 20:1734–1735

Hendriks-Balk MC, Michel MC, Alewijnse AE (2007) Pitfalls in the normalization of real-time polymerase chain reaction. Basic Res Cardiol 102:195–197

Higuchi R, Dollinger G, Walsh PS, Griffith R (1992) Simultaneous amplification and detection of specific DNA sequences. Biotechnology (NY) 10:413–417

Huggett J, Bustin S (2011) Standardisation and reporting for nucleic acid quantification. Accred Qual Assur 16:399–405

Huggett J, Dheda K, Bustin S, Zumla A (2005) Real-time RT-PCR normalisation; strategies and considerations. Genes Immun 6:279–284

Huis R, Hawkins S, Neutelings G (2010) Selection of reference genes for quantitative gene expression normalization in flax (*Linum usitatissimum* L.). BMC Plant Biol 10:71

Jacob F, Guertler R, Naim S, Nixdorf S, Fedier A, Hacker NF, Heinzelmann-Schwarz V (2013) Careful selection of reference genes is required for reliable performance of RT-qPCR in human normal and cancer cell lines. PLoS One 8(3):e59180

Jarošová J, Kundu JK (2010) Validation of reference genes as internal control for studying viral infections in cereals by quantitative real-time RT-PCR. BMC Plant Biol 10:146

Jemiolo B, Trappe S (2004) Single muscle fiber gene expression in human skeletal muscle: validation of internal control with exercise. Bioch Biophys Res Commun 320:1043–1050

Jin P, Zhao Y, Ngalame Y, Panelli MC, Nagorsen D, Monsurro V, Smith K, Hu N, Su H, Taylor PR, Marincola FM, Wang E (2004) Selection and validation of endogenous reference genes using a high throughput approach. BMC Genomics 5:55

Jurczyk B, Rapacz M, Budzisz K, Barcik W, Sasal M (2012) The effects of cold, light and time of day during low-temperature shift on the expression of *CBF6*, *FpCor14b* and *LOS2* in *Festuca pratensis*. Plant Sci 183:143–148

Kim BR, Nam HY, Kim SU, Kim SI, Chang YJ (2003) Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. Biotechnol Lett 25(21):1869–1872

Kuchipudi SV, Tellabati M, Nelli RK, White GA, Perez BB, Sebastian S, Slomka MJ, Brookes SM, Brown IH, Dunham SP, Chang KC (2012) 18S rRNA is a reliable normalisation gene for real time PCR based on influenza virus infected cells. Virol J 9:230

Lee JM, Roche JR, Donaghy DJ, Thrush A, Sathish P (2010) Validation of reference genes for quantitative RT-PCR studies of gene expression in perennial ryegrass (*Lolium perenne* L.). BMC Mol Biol 11:8

Lefever S, Vandesompele J, Speleman F, Pattyn F (2009) RTPrimerDB: the portal for real-time PCR primers and probes. Nucleic Acids Res 37(Database issue):D942–D945

Li QF, Sun SSM, Yuan DY, Yu HX, Gu MH, Liu QQ (2010) Validation of candidate reference genes for the accurate normalization of real-time quantitative RT-PCR data in rice during seed development. Plant Mol Biol Rep 28:49–57

Lin YL, Lai ZX (2010) Reference gene selection for qPCR analysis during somatic embryogenesis in longan tree. Plant Sci 178:359–365

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25(4):402–408

Løvdal T, Lillo C (2009) Reference gene selection for quantitative real-time PCR normalization in tomato subjected to nitrogen, cold, and light stress. Anal Biochem 387(2):238–242

Mallona I, Lischewski S, Weiss J, Hause B, Egea-Cortines M (2010) Validation of reference genes for quantitative real-time PCR during leaf and flower development in *Petunia hybrida*. BMC Plant Biol 10:4

Manoli A, Sturaro A, Trevisan S, Quaggiotti S, Nonis A (2012) Evaluation of candidate reference genes for qPCR in maize. J Plant Physiol 169(8):807–815

Mehta R, Birerdinc A, Hossain N, Afendy A, Chandhoke V, Younossi Z, Baranova A (2010) Validation of endogenous reference genes for qRT-PCR analysis of human visceral adipose samples. BMC Mol Biol 11:39

Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. J Exp Bot 56(421):2907–2914

Nolan T, Hands RE, Ogunkolade W, Bustin SA (2006) SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. Anal Biochem 351(2):308–310

Paolacci AR, Tanzarella OA, Porceddu E, Ciaffi M (2009) Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. BMC Mol Biol 10:11

Pawłowicz I, Kosmala A, Rapacz M (2012) Expression pattern of the *psbO* gene and its involvement in acclimation of the photosynthetic apparatus during abiotic stresses in *Festuca arundinacea* and *F. pratensis*. Acta Physiol Plant 34:1915–1924

Pettengill EA, Parmentier-Line C, Coleman GD (2012) Evaluation of qPCR reference genes in two genotypes of Populus for use in photoperiod and low-temperature studies. BMC Res Notes 5:366

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT–PCR. Nucleic Acids Res 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. Biotechnol Let 26(6):509–515

Podevin N, Krauss A, Henry I, Swennen R, Remy S (2012) Selection and validation of reference genes for quantitative RT-PCR expression studies of the non-model crop *Musa*. Mol Breed 30(3):1237–1252

Radonić A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A (2004) Guideline to reference gene selection for quantitative real-time PCR. Biochem Biophys Res Commun 313(4):856–862

Raffaello T, Asiegbu FO (2013) Evaluation of potential reference genes for use in gene expression studies in the conifer pathogen (*Heterobasidion annosum*). Mol Biol Rep 40(7):4605–4611

Rapacz M, Wolanin B, Hura K, Tyrka M (2008) The effects of cold acclimation on photosynthetic apparatus and the expression of *COR14b* in four genotypes of barley (*Hordeum vulgare*) contrasting in their tolerance to freezing and high-light treatment in cold conditions. Ann Bot 101:689–699

Rapacz M, Kościelniak J, Jurczyk B, Adamska A, Wójcik M (2010) Different patterns of physiological and molecular response to drought in seedlings of malt- and feed-type barleys (*Hordeum vulgare*). J Agron Crop Sci 196:9–19

Rapacz M, Stępień A, Skorupa K (2012) Internal standards for quantitative RT-PCR studies of gene expression under drought treatment in barley (*Hordeum vulgare* L.): the effects of developmental stage and leaf age. Acta Physiol Plant 34:1723–1733

Rubie C, Kempf K, Hans J, Su T, Tilton B, Georg T, Brittner B, Ludwig B, Schilling M (2005) Housekeeping gene variability in normal and cancerous colorectal, pancreatic, esophageal, gastric and hepatic tissues. Mol Cell Probes 19:101–109

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Silveira ED, Alves-Ferreira M, Guimarães LA, da Silva FR, de Campos Carneiro VT (2009) Selection of reference genes for quantitative real-time PCR expression studies in the apomictic and sexual grass *Brachiaria brizantha*. BMC Plant Biol 9:84

Studzińska A, Tyburski J, Daca P, Tretyn A (2008) PCR in real time: the idea of the method and strategies of reaction monitoring. Biotechnologia 1(80):71–85 (in Polish with English abstract)

Suzuki T, Higgins PJ, Crawford DR (2000) Control selection for RNA quantitation. Biotechniques 29(2):332–337

Szabo A, Perou CM, Karaca M, Perreard L, Quackenbush JF, Bernard PS (2004) Statistical modeling for selecting housekeeper genes. Genome Biol 5(8):R59

Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M (2010) A practical approach to RT-qPCR—Publishing data that conform to the MIQE guidelines. Methods 50:S1–S5

Teste MA, Duquenne M, François JM, Parrou JL (2009) Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. BMC Mol Biol 10:99

Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A, Heinen E (1999) Housekeeping genes as internal standards: Use and limits. J Biotechnol 75:291–295

Thellin O, ElMoualij B, Heinen E, Zorzi W (2009) A decade of improvements in quantification of gene expression and internal standard selection. Biotechnol Adv 27:323–333

Tong Z, Gao Z, Wang F, Zhou J, Zhang Z (2009) Selection of reliable reference genes for gene expression studies in peach using real-time PCR. BMC Mol Biol 10:71

Tsafrir D, Bacolod M, Selvanayagam Z, Tsafrir I, Shia J, Zeng Z, Liu H, Krier C, Stengel RF, Barany F, Gerald WL, Paty PB, Domany E, Notterman DA (2006) Relationship of gene expression and chromosomal abnormalities in colorectal cancer. Cancer Res 66:2129–2137

Tyburski J, Studzińska A, Daca P, Tretyn A (2008) PCR in real time: methods of data analysis. Biotechnologia 1(80):86–96 (in Polish with English abstract)

Uhlmann V, Martin CM, Sheils O, Pilkington L, Silva I, Killalea A, Murch SB, Walker-Smith J, Thomson M, Wakefield AJ, O'Leary

JJ (2002) Potential viral pathogenic mechanism for new variant inflammatory bowel disease. Mol Pathol 55:84–90

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):research0034.1–0034.11

Wallström SV, Aidemark M, Escobar MA, Rasmusson AG (2012) An alternatively spliced domain of the *NDC1* NAD(P)H dehydrogenase gene strongly influences the expression of the *ACTIN2* reference gene in *Arabidopsis thaliana*. Plant Sci 183:190–196

Wan H, Zhao Z, Qian C, Sui Y, Malik AA, Chen J (2010) Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. Anal Biochem 399:257–261

Weyrich A, Axtner J, Sommer S (2010) Selection and validation of reference genes for realtime RT-PCR studies in the non-model species *Delomys sublineatus*, an endemic Brazilian rodent. Biochem Biophys Res Commun 392:145–149

Wójcik-Jagła M, Rapacz M, Barcik W, Janowiak F (2012) Differential regulation of barley (*Hordeum distichon*) *HVA1* and *SRG6* transcript accumulation during the induction of soil and leaf water deficit. Acta Physiol Plant 34:2069–2078

Zhang Z, Hu J (2007) Development and validation of endogenous reference genes for expression profiling of medaka (*Oryzias latipes*) exposed to endocrine disrupting chemicals by quantitative real-time RT-PCR. Toxicol Sci 95:356–368

Zhu L, Altman SW (2005) mRNA and 18S-RNA coapplication-reverse transcription for quantitative gene expression analysis. Anal Biochem 345:102–109

Zhu X, Li X, Chen W, Chen J, Lu W, Chen L, Fu D (2012) Evaluation of new reference genes in papaya for accurate transcript normalization under different experimental conditions. PLoS One 7(8):e44405