LETTER TO THE EDITOR

Towards a Systematic Validation of References in Real-Time RT-PCR

Real-time RT-PCR (also known as quantitative RT-PCR [qRT-PCR]) is a powerful tool for quantifying gene expression, combining both high sensitivity and specificity with efficient signal detection. It has relatively recently begun to be used to monitor gene expression in plants (reviewed in Gachon et al., 2004) and remains underused, considering its ability to discriminate between the expression of closely related genes and to quantify transcript levels of very weakly expressed genes (Czechowski et al., 2004). This may be due in part to a lack of confidence in real-time RT-PCR, arising from disappointments in the contradictory results sometimes obtained when using this technique without a robust normalization strategy. To account for between-sample variations in the amounts of starting material and the efficiency of the quantification process, determinations of mRNA species in real-time RT-PCR analysis should be normalized according to the total amounts of mRNA present in the samples. For this purpose, the expression levels of target genes are described in terms of the ratios of target mRNA levels to the level of a reference mRNA species, which should be the product of a stably expressed gene whose abundance is strongly correlated to the total amounts of mRNA present in each sample (Huggett et al., 2005).

In December 2006, a detailed discussion (inter alia) of factors affecting the reliability of real-time RT-PCR was published in *Nature Protocols* (Nolan et al., 2006). The authors mentioned that the key measure for ensuring that the technique delivered highly accurate results was to apply a robust normalization strategy (i.e., to normalize using a reference gene that has been shown specifically to be stably expressed under relevant experimental conditions). Owing to the near-infinite diversity of possible experimental conditions and the inability of controlling all experimental parameters, a reference gene should be considered as being valid solely when validated under the same conditions as those used to assess expression of the target gene. This issue has also been addressed in several other articles (Dheda et al., 2005; Huggett et al., 2005), which focused on the merits of using one of several algorithms that have been developed to identify the best reference genes to use under given experimental conditions (Pfaffl et al., 2002, 2004; Vandesompele et al., 2002; Andersen et al., 2004). Use of these algorithms allows for easy systematic validation (i.e., assessment of the expression stability of candidate genes under a variety of conditions to find specific references appropriate for each real-time RT-PCR analysis).

Efforts to adopt such systematic validation procedures are growing, especially in the medical field, as illustrated by the rapid increase during 2007 in the use of geNorm (Vandesompele et al., 2002), one of the most commonly used algorithms. geNorm software has been cited hundreds of times and downloaded thousands of times (http://medgen.ugent.be/~jvdesomp/genorm/).

Unfortunately, however, the increased awareness of the importance of systematic validation has not permeated fully throughout the community of molecular biologists, and although the potentially highly misleading effects of using inappropriate references for normalization are widely known, they are also still widely disregarded. Consequently, real-time RT-PCR data are being normalized alarmingly poorly in many molecular analyses, especially in the plant sciences. For instance, genes that had been validated as references were used in only 3.2% of 188 real-time RT-PCR analyses published during a 6-month period from July through December 2007 in *The Plant Cell*, *Plant Physiology*, and *The Plant Journal* (the three leading primary research journals in plant biology according to the ISI Web of Knowledge), mainly using geNorm software.

In all of the remaining analyses, putative housekeeping genes were used as reference genes. While many, if not most, of the conclusions in these studies likely are still valid, it is impossible to say which are and which are not without data that validates the choice of the reference gene in each study. The choice of such genes as references may be inappropriate since their status as housekeeping genes is generally based on methods known to be mainly qualitative (e.g., RNA gel blotting or histochemical analysis) and is inconsistent with (indeed undermines) the high accuracy associated with real-time RT-PCR. It is generally assumed that housekeeping genes encoding proteins required for basal cell activities, such as central carbon metabolism, protein translation, cytoskeleton maintenance, and protein turnover, are expressed uniformly in different tissues and organs. However, it has been found that the expression of a number of such housekeeping genes, although constant in some experimental conditions, varies considerably in other cases (Volkov et al., 2003; Czechowski et al., 2005; Nicot et al., 2005; Remans et al., 2008). Therefore, their systematic use as reference genes may result in the misinterpretation of results. In addition, such genes are often used as references in experimental conditions that differ from those in which their stability of expression was tested.

We have recently shown that genes commonly used as references may be expressed unstably during the development of *Arabidopsis* plants (Gutierrez et al., 2008). Using validated and nonvalidated references, we illustrated the extent to which the use of arbitrarily chosen genes could affect the quantification of target gene expression levels. Up to 100-fold variations were found in the expression of a target gene, which could be attributed only to variations in the expression of the reference genes, with consequently huge potential

scope for misinterpretation of the results. Therefore, there is an urgent need to regard the systematic validation of reference genes as an essential component of real-time RT-PCR analysis to improve the reliability of published results and retain the accuracy of this powerful technique, which may be lost owing to inappropriate normalization.

An important point for identifying suitable references is the choice of the initial subset of genes used in the validation procedure, which should consist of genes already assessed as being good candidates regarding their expression stability. As suggested in our recent article (Gutierrez et al., 2008), the database published by Czechowski et al. (2005) is a good starting point for choosing reference genes for analyses, not only in *Arabidopsis* but also other model plants, since it provides a list of suitable candidate genes for tests under specific experimental conditions using one of the available algorithms developed for this purpose. Recently, Rieu et al. (2008) performed a real-time RT-PCR analysis of outstanding quality. By testing the stability of candidate genes from the database published by Czechowski et al. (2005), they identified suitable references to be used for each one of their four sets of experimental conditions. Thanks to this robust normalization strategy, they avoided introducing a bias in their analysis and obtained highly accurate real-time RT-PCR results. Notably, the genes found to be the most stably expressed varied between different sets of experiments, illustrating the specificity of the valid references and, thus, the necessity for the validation procedure to be systematic (i.e., to be performed for each experiment being conducted).

Such systematic validation will only be applied consistently and universally if and when all reviewers start to regard its use as essential when evaluating the validity of real-time RT-PCR analyses prior to publication. Given the importance of the issue, we believe it would be advisable to start insisting on the use of appropriate, robust validation in all published transcription analyses as soon as possible.

Laurent Gutierrez
Department of Forest Genetics and Plant Physiology
Umeå Plant Science Centre
Swedish University of Agricultural Sciences
SE-901 87 Umeå, Sweden
laurent.gutierrez@genfys.slu.se

Mélanie Mauriat
Department of Forest Genetics and Plant Physiology
Umeå Plant Science Centre
Swedish University of Agricultural Sciences
SE-901 87 Umeå, Sweden

Jérôme Pelloux
EA3900 BioPI
Faculté des Sciences
Université de Picardie Jules Verne
F-80039 Amiens Cedex, France

Catherine Bellini
Department of Forest Genetics and Plant Physiology
Umeå Plant Science Centre
Swedish University of Agricultural Sciences
SE-901 87 Umeå, Sweden
Laboratoire de Biologie Cellulaire
Institut National de la Recherche Agronomique
78026 Versailles Cedex, France

Olivier Van Wuytswinkel
EA3900 BioPI
Faculté des Sciences
Université de Picardie Jules Verne
F-80039 Amiens Cedex, France

REFERENCES

Andersen, C.L., Jensen, J.L., and Ørntoft, T.F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. **64:** 5245–5250.

Czechowski, T., Bari, R.P., Stitt, M., Scheible, W.R., and Udvardi, M.K. (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.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. **139:** 5–17.

Dheda, K., Huggett, J.F., Chang, J.S., Kim, L.U., Bustin, S.A., Johnson, M.A., Rook, G.A., and Zumla, A. (2005). The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. Anal. Biochem. **344:** 141–143.

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

Gutierrez, L., Mauriat, M., Guénin, S., Pelloux, J., Lefebvre, J.-F., Louver, R., Rusterucci, C., Moritz, T., Guerineau, F., Bellini, C., and Van Wuytswinkel, O. (2008). 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., in press.

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

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

Nolan, T., Hands, R.E., and Bustin, S.A. (2006). Quantification of mRNA using real-time PCR. Nat. Protoc **1:** 1559–1582.

Pfaffl, M.W., Horgan, G.W., and Dempfle, L. (2002). Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res. **30:** e36.

Pfaffl, M.W., Tichopad, A., Prgomet, C., and Neuvians, T.P. (2004). Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. Biotechnol. Lett. **26:** 509–515.

Remans, T., Smeets, K., Opdenakker, K., Mathijsen, D., Vangronsveld, J., and Cuypers, A. (2008). Normalisation of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. Planta **227:** 1343–1349.

Rieu, I., Ruiz-Rivero, O., Fernandez-Garcia, N., Griffiths, J., Powers, S.J., Gong, F., Linhartova, T., Eriksson, S., Nilsson, O., Thomas, S.G., Phillips, A.L., and Hedden, P. (2008). The gibberellin biosynthetic genes AtGA20ox1 and AtGA20ox2 act, partially redundantly, to promote growth and development throughout the Arabidosis life cycle. Plant J. **53:** 488–504.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. **3:** RESEARCH0034.

Volkov, R.A., Panchuk, I.I., and Schoffl, F. (2003). Heat-stress-dependency and developmental modulation of gene expression: The potential of house-keeping genes as internal standards in mRNA expression profiling using real-time RT-PCR. J. Exp. Bot. **54:** 2343–2349.

This information is current as of September 1, 2008

References	This article cites 14 articles, 6 of which you can access for free at: http://www.plantcell.org/cgi/content/full/20/7/1734#BIBL
Permissions	https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&issn=1532298X&WT.mc_id=pd_hw1532298X
eTOCs	Sign up for eTOCs for *THE PLANT CELL* at: http://www.plantcell.org/subscriptions/etoc.shtml
CiteTrack Alerts	Sign up for CiteTrack Alerts for *Plant Cell* at: http://www.plantcell.org/cgi/alerts/ctmain
Subscription Information	Subscription information for *The Plant Cell* and *Plant Physiology* is available at: http://www.aspb.org/publications/subscriptions.cfm