Methods 50 (2010) 227–230

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

How to do successful gene expression analysis using real-time PCR

Stefaan Derveaux, Jo Vandesompele *, Jan Hellemans

Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium
Biogazelle, Ghent, Belgium

ARTICLE INFO

Article history:
Accepted 25 November 2009
Available online 5 December 2009

Keywords:
RT-qPCR
Quality control
Experiment design
Assay design
Reference gene validation

ABSTRACT

Reverse transcription quantitative PCR (RT-qPCR) is considered today as the gold standard for accurate, sensitive and fast measurement of gene expression. Unfortunately, what many users fail to appreciate is that numerous critical issues in the workflow need to be addressed before biologically meaningful and trustworthy conclusions can be drawn. Here, we review the entire workflow from the planning and preparation phase, over the actual real-time PCR cycling experiments to data-analysis and reporting steps. This process can be captured with the appropriate acronym PCR: plan/prepare, cycle and report. The key message is that quality assurance and quality control are essential throughout the entire RT-qPCR workflow; from living cells, over extraction of nucleic acids, storage, various enzymatic steps such as DNase treatment, reverse transcription and PCR amplification, to data-analysis and finally reporting.

1. Chance favors the prepared mind

Reverse transcription quantitative PCR (RT-qPCR) distinguishes itself from other methods available for gene expression in terms of accuracy, sensitivity, and fast results. Because of this, the technology has established itself as the golden standard for medium throughput gene expression analysis. Due to its apparent simplicity, also inexperienced users can rapidly produce results; however, care should be taken when performing RT-qPCR as numerous critical quality issues may arise throughout the entire workflow influencing the accuracy of the results and the reliability of the conclusions. Intensive quality control is an important and necessary part that can be captured with the appropriate acronym PCR: plan/prepare, cycle and report (Fig. 1). In this first section, we will illustrate four important preparative steps in the RT-qPCR workflow, prior to starting the actual real-time PCR quantifications. Spending careful attention to experiment design, sample and assay quality control, and selection of proper reference genes for normalization will significantly increase the chance of successful results; preparation is everything.

1.1. Experiment design

One of the most neglected points in setting up an RT-qPCR study is experimental design. Nevertheless, proper set-up of the experiment saves time, cuts down on reagent cost and increases the accuracy and precision of the results. Experiment design involves 3 important aspects. First, power analysis should enable careful assessment of the number of biological samples needed to draw meaningful and statistically significant results. Often, researchers are confronted with too few samples to demonstrate significance for the observed differential expression in their experiment. Second, the proper run layout strategy should be selected. Two different experimental set-ups can be followed in an RT-qPCR study [1]. According to the preferred sample maximization method, as many samples as possible are analyzed in the same run. This means that different genes should be analyzed in different runs if not enough free wells are available to analyze the different genes in the same run. In contrast, the gene maximization set-up analyzes multiple genes in the same run, and spreads samples across runs if required. The latter approach is often used in commercial kits or in prospective studies. It is important to realize that in a relative quantification study, the experimenter is usually interested in comparing the expression level of a particular gene between different samples. Therefore, the sample maximization method is highly recommended because it does not suffer from (often underestimated) technical, run-to-run variation between the samples. However, irrespective of the set-up, inter-run calibration is required to correct for possible run-to-run variation whenever not all samples are or can be analyzed in the same run. For this purpose, the experimenter needs to analyze so-called inter-run calibrators (IRC); these are identical samples that are tested in both runs. By measuring the difference in quantification cycle or normalized relative quantity between the IRCs in both runs, it is possible to calculate a correction or calibration factor to remove the run-to-run difference, and proceed as if all samples were analyzed in the same run. The qBase quantification model incorporates state of the art inter-run calibration schemes [1].

* Corresponding author. Address: Center for Medical Genetics Ghent, Ghent University Hospital, De Pintelaan 185, B-9000 Ghent, Belgium. Fax: +32 9 3326549.
E-mail address: Joke.Vandesompele@Ugent.be (J. Vandesompele).

1046-2023/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2009.11.001

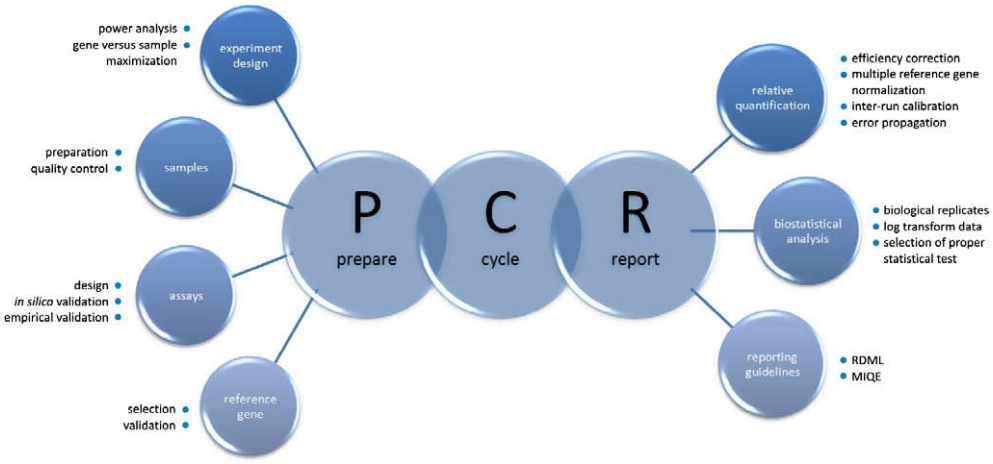

Fig. 1. Overview of all quality control tools throughout the entire qPCR workflow.

Apart from the theoretical considerations of gene versus sample maximization there are also practical reasons to think about run layout. As the cost of microtiter plates is negligible compared to that of samples and reagents, there is no need to force reactions in a single plate until it is completely full. Rather, it may be more appropriate to leave some wells empty in order to have a more logical or practical pipetting scheme. For example, in a 96-well plate 9 genes can be measured in 10 samples (in this example no replicated reactions); however, analyzing only 8 genes allows putting 1 sample or gene in a single column or row, respectively. This not only facilitates run annotation, but also allows easy pipetting with a multi-channel pipette or dispensing robot and prevents mistakes during reaction set-up. In line with this, some experiments are better served with a 384-well instrument for its lower consumable cost, faster results and improved data quality as technical run-to-run variation can be more easily avoided.

1.2. Sample extraction and quality assurance

The need for high quality nucleic acids has been recognized for many years in the microarray field; the cost of such experiments pushed people to assess the purity and integrity of the input RNA molecules. In contrast, RT-qPCR is relatively cheap and – more worrying – almost always result in beautiful sigmoidal amplification curves, even if the template is highly degraded as is the case for e.g. formalin fixed paraffin embedded tissues. An uninformed user is thus easily mislead, and might not appreciate the impact of RNA quality on the results. A well known phrase in computer science 'garbage in, garbage out' also applies to real-time PCR. As RT-qPCR performance is affected by the RNA integrity, Fleige and Pfaffl recommend an RNA quality score (RIN or RQI) higher than five as good total RNA quality and higher than eight as perfect total RNA for downstream applications [2]. A study on the impact of RNA quality on the expression stability of reference genes indicated that it is inappropriate to compare degraded and intact samples, necessitating sample quality control prior to RT-qPCR measurements [3]. Our own recent data indicate that RNA quality has a profound impact on the results, in terms of the significance of differential expression, variability of reference genes and classification performance of a multi-gene signature (Vermeulen et al., submitted for publication). In addition or as an alternative to the use of capillary gel electrophoresis methods that assess the integrity of the ribosomal RNA molecules (as discussed in [2,3]), PCR based tests are also frequently used to determine mRNA integrity. In one such a test, the ratio between the 5' and 3' end of a universally expressed gene is measured upon anchored oligo-dT cDNA synthesis, reflecting integrity of that particular poly-adenylated transcript [4]. Finally, another PCR based assay is often used in clinical diagnostics to determine sample purity. By comparing the Cq value of a known concentration of a spiked DNA or RNA molecule in both a negative water control and in the sample of unknown quality, enzymatic inhibition can be determined [5].

Apart from purity and integrity, there is a third important requirement related to the input RNA material; it should be free of contaminating DNA. While it is sometimes possible to design an assay that spans a large intron such that residual DNA is not co-amplified, this strategy is not recommended. First, it makes the design process more cumbersome, and secondly, for up to 20% of the human genes, it will not work as these genes are either single exon genes or have one or more processed pseudogenes (retropseudogene or intronless copy) in the genome. The most efficient strategy in our hands is to do a proper DNase treatment, followed by a careful check for absence of DNA through qPCR analysis of a DNA target on the crude RNA [6]. When no signal is observed, the RNA was free of DNA.

A last point of attention with respect to RNA samples is the increased use of sample pre-amplification protocols. Indeed, many biological specimens are valuable and often the amount of RNA extracted is limiting large-scale gene-expression studies. RNA pre-amplification methods address this issue by producing micrograms of cDNA starting from a few nanograms of total RNA. Prior to using one of the different amplification protocols that are available, there is only one important criterion that has to be evaluated: fold changes in expression between two samples should be preserved before and after amplification. As different targets might be pre-amplified with a different efficiency, it is irrelevant to compare genes before and after pre-amplification. As such, transcript variant analysis is only possible before pre-amplification.

1.3. Assay design and quality control

Another important preparative step is the design and empirical validation of a real-time PCR assay. Assays are available from multiple sources. You can design by yourself using free or commercially available design software, or you can buy assays from specialized vendors that provide assays of the shelf or design on demand. A third alternative is to look for published and experimentally validated assays in public databases such as rtprimerdb.org [7]. The latest version contains almost 8000 assays from mainly five organisms (human, mouse, rat, rice and *Arabidopsis thaliana*). Irrespective the

source of the primer and probe sequences, it is of great importance to validate the assay, both *in silico* and empirically.

People typically perform specificity analysis of an RT-qPCR primer pair by doing a BLAST or BiSearch query [8]. However, there is more quality control required than only specificity assessment. One should also inspect the presence of SNPs in the primer annealing regions (which may hamper efficient annealing of the primer or prevent amplification of the variant allele at all), and model the secondary structure of the amplicon (using e.g. UNAFold software). It is well known that hairpins overlapping primer annealing sites significantly hamper efficient annealing and negatively impact the PCR efficiency [9]. All these tools are integrated in an automated *in silico* assay evaluation pipeline available from rtprimerdb.org (http://www.rtprimerdb.org).

After *in silico* assay evaluation, the primers need to be empirically validated by doing an actual RT-qPCR experiment and inspecting the length using gel electrophoresis (once), and inspecting the melt curve when using SYBR Green I. In addition, a standard curve needs to be run in order to estimate the PCR efficiency. Important to realize is that the more dilution points and the wider the range (dilution factor), the more precise PCR efficiency can be determined [1]. Best practice is to use a mixture of representative samples as input material for the dilution series.

An important aspect of assay design is to select the correct target sequence, which in many cases is not easy to do, especially for the large proportion of genes that have alternatively spliced isoforms. If no prior knowledge is available on the function of these transcript variants, one may decide to design an assay for a part of the transcript that is expressed in most or all of the variants. Dedicated splice variant quantification requires a specific work flow and its own controls. A straightforward and reliable strategy was published by Vandenbroucke and colleagues [10]. The key message of that study is that so-called absolute or equimolar (same molarity for the various splice variant targets) standard curves are required for accurate assessment of splice variants, and that the primers should be the driving force behind the specificity of the detection (instead of relying on a probe). Absolute or equimolar standard curves are curves in which the number of molecules is known; an efficient way to produce them is to purify a PCR product or synthesize a single stranded long oligonucleotide template, both diluted in carrier DNA (10–100 ng of *E. coli* or yeast tRNA).

1.4. Reference gene validation

A final step in the preparation phase is the selection of proper reference genes for data normalization. Important to realize is that any gene expression quantification result is finally composed of two sources of variation; on the one side there is inherent technical or experimentally induced variation and on the other side, there is the true biological variation that underlies the phenomenon under investigation. The very purpose of any normalization strategy is to remove the technical variation as much as possible, ending up with the true biological changes. At the 3rd London qPCR Symposium organized by Professor Stephen Bustin (April 2005), normalization against 3 or more validated reference genes was considered as the most appropriate and universally applicable method. While many algorithms have been reported to date to evaluate candidate reference genes in terms of expression stability (or suitability as normalizing gene) [11], the geNorm method [12] was the first (http://medgen.ugent.be/genorm) and has established itself as the *de facto* standard with more than 2000 citations (Google Scholar, September 2009). A typical pilot experiment for evaluation of reference gene expression stability measures around 10 candidate reference genes in 10 representative samples from each tissue or sample group. The genes are selected such that they represent different biological pathways and expression abundance levels. The relative gene expression values are subsequently imported in the geNorm program and ranked according to their expression stability. In a subsequent analysis, the software is able to indicate how many reference genes are optimally required to remove most of the technical variation (which depends on the expression stability of the tested genes and on the heterogeneity of the samples under investigation). Typically, between 3 and 5 genes are required for accurate normalization. It is clear that any report documenting small expression changes without validating reference genes is unacceptable as it has been shown that the use of a single non-validated reference gene results in a significant bias (ranging from more than 3-fold in 25% of the results up to 6-fold in 10% of the results) [12].

Besides selection of stably expressed reference genes in a pilot experiment with representative samples, it remains important to assess their expression stability in the final experiment. In the qbase^{PLUS} software, expression stability calculations are automatically performed if more than one reference gene is used for normalization. This approach provides the required reference gene quality control in each experiment.

2. Action and reaction

There are two basic guidelines in PCR set-up to assure successful reactions: maximize precision through the proper use of calibrated pipettes or automated liquid handling systems and minimize contamination by using filter tips and gloves and by keeping amplified products away from the PCR set-up area. RT and PCR replicates can be included to evaluate variation in reverse transcription efficiency and qPCR precision while no-template controls allow verification of the absence of contamination in your qPCR reactions.

Before setting up the qPCR reaction one needs to decide on how to monitor the increase in PCR products as the reaction proceeds. One can choose between intercalating dyes such as SYBR green I or one of the many probe technologies that are available. Both detection methods can deliver excellent results and each has its advantages and disadvantages: probes enable multiplex quantification whereas intercalating dyes are the most cost-effective choice, especially when a limited number of samples are quantified.

A last consideration with significant consequences is the volume of the PCR reaction. In 96- or 384-well plates, PCR reaction volumes can be easily scaled down to 10 μl and 5 μl, respectively. This will not only save on reagents cost, but also allows the use of less material from your precious samples (as only the final concentration of the template impacts the Cq value). Below the proposed reaction volumes, one is likely to encounter evaporation problems and loss of precision. High-throughput systems (now available from Roche, Biotrove and Fluidigm) enable fast screening of thousands of reactions in small (sub)microliter volumes. Of note, there can be a reduction in sensitivity due to the limited amount of sample that can be added to the reaction. A sample pre-amplification step might solve this issue.

3. Trust, but verify

3.1. Relative quantification

The polymerase chain reaction is an exponential process whereby the specifically amplified product ideally doubles each cycle. As such, the measured Cq value (standard name for threshold cycle or crossing point value according to the RDML guideless, http://www.rdml.org [13]) is a logarithmic value that needs to be converted into a linear relative quantity using the following exponential function: $RQ_{unkn} = 2\hat{}(Cq_{cal} - Cq_{unkn})$ (Cq_{cal} is the Cq value of the calibrator sample (or reference sample, e.g. untreated control),

Cq_{unkn} the Cq value of the unknown sample, RQ_{unkn} the quantity of the unknown sample relative to the calibrator, and 2 the base of the exponential function, denoting 100% PCR efficiency). Pfaffl was the first to realize that the use of a gene specific PCR efficiency (when measured accurately) improves the accuracy of the results [14]. However, in his quantification model, only one reference gene can be inserted in the equation for normalization. As already indicated, normalization is a crucial step in the calculation workflow in which sample related technical variation is cancelled out. We and others have shown that the use of multiple stably expressed reference genes results in more accurate data and statistically more significant results, and allows reliable quantification of small expression differences [15,12]. The use of efficiency corrected multiple reference genes is enabled in the universally applicable qBase quantification model [1], forming the basis of the qbase^{PLUS} software (http://www.qbaseplus.com) that also comes in a free version. Such a quantification model not only provides the required flexibility (in terms of selection of one or more reference genes, either or not with correction for gene specific PCR efficiency) but also employs state of the art error propagation rules (providing confidence measures to the final quantification result) and inter-run calibration schemes (see also higher). The latter makes it possible to perform large multi-plate experiments in which many more samples are studied than actually fit in one physical plate or PCR run. This feature not only enables large multi-centric studies but also prospective evaluation of patients in the clinic [16].

3.2. Biostatistical analysis

It is beyond the scope of this article to review all statistical tests to determine significance of a difference in gene expression between 2 or more groups, to identify a diagnostic or prognostic RNA marker with high confidence, to find correlations between gene expression patterns or samples, or to identify relevant pathways or new sample subgroups. However, a few important messages can be conveyed. First, it is good practice to log transform the final gene expression results (i.e. the normalized relative quantities), in order to make the data distribution more symmetrical (as gene expression data is often log normally distributed [17]). Together with the Central Limit Theorem, this allows the use of parametric statistical tests and calculations that rely on a distribution that resembles a normal distribution (e.g. classic t-test, confidence intervals, Analysis of Variance) [18]. Secondly, independent biological replicates are required to draw meaningful and reliable conclusions. The minimum number depends on the statistical test and on the power one wants to achieve (e.g. for confidence interval analysis, at least three replicates are needed, for a non-parametric paired test (Wilcoxon signed-rank test), at least six pairs are needed). It must be clear that statistics on repeated measurements (e.g. PCR replicates) are absolutely nonsense, as only technical variation is measured. Third, the statistical test should be selected prior to doing the actual experiment, whereby the choice is based on the question that needs to be addressed, the number of data points, and the distribution of the data. If in doubt, a (bio)statistician should be consulted. Fourth, if the data speak for themselves, don't interrupt. This means that sometimes, the difference is so striking or obvious that there is no need to perform a statistical test.

3.3. Reporting guidelines – the tower of Babel

Having generated high quality data, the next step is to communicate results to colleagues and collaborators and possibly to submit for publication. One major limitation today is that the real-time PCR instruments' software and third party data analysis software speak a different language and as such create data files that are difficult if not impossible to understand by people not having the same software. To address this issue, the international Real-time PCR Data Markup Language (RDML) consortium was founded (http://www.rdml.org) [13] with one major goal, i.e. the development of an XML-based Real-Time PCR Data Markup Language (RDML) standard. RDML enables straightforward exchange of qPCR data and related information between qPCR instruments and third party data analysis software, between colleagues and collaborators and between experimenters and journals or public repositories.

RDML has recently become part of the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines and accompanying checklist to guarantee inclusion of key data information when reporting experimental results [19] (http://www.rdml.org/miqe/). MIQE describes the minimum information necessary for evaluation of RT-qPCR experiments under the form of a checklist to accompany the initial submission of a manuscript to the publisher. By providing all relevant experimental conditions and assay characteristics, reviewers can assess the validity of the protocols used. Full disclosure of all reagents, sequences, and analysis methods is necessary to enable other investigators to reproduce results. Following these guidelines will encourage better experimental practice, allowing more reliable and unequivocal interpretation of quantitative PCR results.

4. Conclusions

Real-time quantitative RT-PCR is a wonderful method for fast, accurate, sensitive and cost-effective gene expression analysis. However, the simplicity of the technology itself makes it vulnerable for abuse in experiments in which the operator does not perform the required quality control throughout the entire procedure. In this review, we outlined the different steps in the work flow and indicated point by point where and how critical issues can be resolved. Following the advice in this paper, any user should be able to do (more) successful gene expression profiling using the RT-qPCR technology.

References

[1] J. Hellemans, G. Mortier, P.A. De, F. Speleman, J. Vandesompele, Genome Biol. 8 (2007) R19.
[2] S. Fleige, M.W. Pfaffl, Mol. Aspects Med. 27 (2006) 126–139.
[3] C.A. Perez-Novo, C. Claeys, F. Speleman, C.P. Van, C. Bachert, J. Vandesompele, Biotechniques 39 (2005) (pp. 52, 54, 56).
[4] T. Nolan, R.E. Hands, S.A. Bustin, Nat. Protoc. 1 (2006) 1559–1582.
[5] T. Nolan, R.E. Hands, W. Ogunkolade, S.A. Bustin, Anal. Biochem. 351 (2006) 308–310.
[6] J. Vandesompele, P.A. De, F. Speleman, Anal. Biochem. 303 (2002) 95–98.
[7] S. Lefever, J. Vandesompele, F. Speleman, F. Pattyn, Nucleic Acids Res. 37 (2009) D942–D945.
[8] T. Aranyi, G.E. Tusnady, Methods Mol. Biol. 402 (2007) 385–402.
[9] J. Hoebeeck, L.R. van der, B. Poppe, S.E. De, N. Yigit, K. Claes, R. Zewald, G.J. de Jong, P.A. De, F. Speleman, J. Lab. Invest. 85 (2005) 24–33.
[10] I.I. Vandenbroucke, J. Vandesompele, A.D. Paepe, L. Messiaen, Nucleic Acids Res. 29 (2001) E68.
[11] J. Vandesompele, M. Kubista, M.W. Pfaffl, in: J. Logan, K. Edwards, N. Saunders (Eds.), Real-Time PCR Current Technology and Applications, Caister Academic Press, Norfolk, 2009.
[12] J. Vandesompele, P.K. De, F. Pattyn, B. Poppe, R.N. Van, P.A. De, F. Speleman, Genome Biol. 3 (2002).
[13] S. Lefever, J. Hellemans, F. Pattyn, D.R. Przybylski, C. Taylor, R. Geurts, A. Untergasser, J. Vandesompele, Nucleic Acids Res. 37 (2009) 2065–2069.
[14] M.W. Pfaffl, Nucleic Acids Res. 29 (2001) e45.
[15] J. Hellemans, O. Preobrazhenska, A. Willaert, P. Debeer, P.C. Verdonk, T. Costa, K. Janssens, B. Menten, R.N. Van, S.J. Vermeulen, R. Savarirayan, H.W. Van, F. Vanhoenacker, D. Huylebroeck, P.A. De, J.M. Naeyaert, J. Vandesompele, F. Speleman, K. Verschueren, P.J. Coucke, G.R. Mortier, Nat. Genet. 36 (2004) 1213–1218.
[16] J. Vermeulen, F. Pattyn, P.K. De, L. Vercruysse, S. Derveaux, P. Mestdagh, S. Lefever, J. Hellemans, F. Speleman, J. Vandesompele, Nucleic Acids Res. (2009).
[17] M. Bengtsson, A. Stahlberg, P. Rorsman, M. Kubista, Genome Res. 15 (2005) 1388–1392.
[18] H. Motulski, Intuitive Biostatistics, Oxford University Press, 1995.
[19] S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, C.T. Wittwer, Clin. Chem. 55 (2009) 611–622.