A Practical Approach to RT-qPCR, Publishing Data That Conforms to the MIQE Guidelines

Sean Taylor, Michael Wakem, Greg Dijkman, Marwan Alsarraj, and Marie Nguyen, Bio-Rad Laboratories, Inc., Hercules, CA 94547

**Introduction**

Real-time quantitative PCR (qPCR) has become a definitive technique for quantitating differences in gene expression levels between samples. Over the past 10 years, the popularity of this method has grown exponentially, with the publication of well over 25,000 papers from diverse fields of science, including agricultural, environmental, industrial, and medical research, making reference to RT (reverse transcription)-qPCR data.

Apart from the broad applicability of the technique, one of the central factors that has stimulated this impressive growth is the increased demand from journal review panels for the use of RT-qPCR to support phenotypic observations with quantitative, molecular data. Furthermore, gene expression analysis is now being used to support protein expression data from proteomics-based assays. The biotechnology industry has responded to the rapid adoption of this technique with the concomitant development of reagents and instruments to perform RT-qPCR experiments. However, since no strict guidelines have been established, researchers have generally designed their experiments based on information gathered from disparate sources, which has resulted in data of variable quality.

A number of technical deficiencies can affect RT-qPCR assay performance, including improper experimental design, inadequate controls and replicates, lack of well-defined experimental conditions and sample handling techniques, poor quality of the RNA sample, suboptimal choice of primers for reverse-transcription, and qPCR reactions and inappropriate methods of data analysis. In an effort to assist the scientific community in producing consistent, high quality data from qPCR experiments, the minimum information for publication of quantitative real-time PCR experiments (MIQE) guidelines have been recently published (Bustin et al. 2009). This has been followed by the development of an XML-based real-time PCR data markup language (RDML) for the consistent reporting of real-time PCR data by the RDML consortium (Lefever et al. 2009). This consortium is active in the development of appropriate and standardized terminology, guidelines on minimum information for biological and biomedical investigations, and a flexible and universal data file structure with tools to create, process, and validate RDML files. All relevant information about the RDML project is available at [www.rdml.org/](http://www.rdml.org/).

The ultimate goal of RDML and MIQE is to establish a clear framework within which to conduct RT-qPCR experiments and to provide guidelines for reviewers and editors to use in the evaluation of the technical quality of submitted manuscripts against an established yardstick. As a consequence, investigations that use this widely applied methodology will produce data that are more consistent, more comparable, and ultimately more reliable.

Given the highly dynamic nature of mRNA transcription and the potential variables introduced in sample handling and in the downstream processing steps (Garson et al. 2009), a standardized approach to each step of the RT-qPCR workflow is critical for reliable and reproducible results. The MIQE provides this approach with a checklist that contains 85 parameters to assure quality results that will meet the acceptance criteria of any journal (Bustin et al. 2009).

In this paper we demonstrate how to apply the MIQE guidelines ([http://medgen.ugent.be/rdml/guidelines.php](http://medgen.ugent.be/rdml/guidelines.php)) to establish a solid experimental approach.
1. Experimental Design
Proper experimental design is the key to any gene expression study. Since mRNA transcription can be sensitive to external stimuli that are unrelated to the processes studied, it is important to work under tightly controlled and well-defined conditions. Taking the time to define experimental procedures, control groups, type and number of replicates, experimental conditions, and sample handling methods within each group is essential to minimize variability (Table 1). Each of these parameters should be carefully recorded prior to conducting gene expression experiments to assure good biological reproducibility for published data.

Table 1. RT-qPCR experimental design and sample management.
This table summarizes the workflow of a typical RT-qPCR experiment from experimental design to defining control groups, replicates, and experimental conditions to the detailed procedures for sample handling. This ultimately assures that the key steps in RT-qPCR data production lead to high quality, reproducible, and publishable data.

<table>
<thead>
<tr>
<th>Experiment Design</th>
<th>Control Groups</th>
<th>Replicates</th>
<th>Experiment Conditions</th>
<th>Sample Handling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease or treatment groups</td>
<td>Time course study (i.e. t = 0)</td>
<td>Biological (different sample per well)</td>
<td>Growth conditions (media and time or OD)</td>
<td>Precise time to harvest cells or tissue</td>
</tr>
<tr>
<td>Target genes implicated</td>
<td>Normal vs. disease (i.e. Normal)</td>
<td>Technical (same sample per well)</td>
<td>Days of embryonic development</td>
<td>Sample extraction method</td>
</tr>
<tr>
<td>Potential reference genes</td>
<td>Untreated vs. drug treated (i.e. untreated)</td>
<td>—</td>
<td>Amount per mass of drug or compound</td>
<td>Preservation method and time</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Sex, phenotype</td>
<td>Thaw and homogenization procedure</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Incubation time</td>
<td>Total RNA extraction procedure</td>
</tr>
</tbody>
</table>

2. RNA Extraction
If samples must be collected over a period of time or in too large a number to process immediately, they should be stored in appropriate conditions (frozen at ~80°C and/or in RNA storage solution) until use. To minimize handling time during the RNA extraction procedure, it is recommended that samples be processed in relatively small batches of 10 to 20. The RNA extraction procedure should include a DNase I treatment step to remove any contaminating genomic DNA. Many commercially available kits such as the Aurum™ RNA extraction kits (Bio-Rad Laboratories, Inc.) can be used for virtually any sample type from plant to animal tissue and include an on-column DNase I treatment of the samples.

3. RNA Quality Control
Ensuring that only RNA of high purity (no contaminants) and high integrity (not degraded) is used is one of the most critical points in the RT-qPCR experimental workflow. Impurities in the RNA sample may lead to inhibition of the PCR reaction, yielding biased data. Using partially degraded RNA can lead to varying and incorrect quantification results (Feige and Pfaffl 2006, Gingrich et al. 2006). Since sample purity and integrity are not related, both should be assessed to ascertain that the RNA sample meets minimal acceptance criteria for the downstream workflow.

Purity of the sample with respect to protein contamination can be assessed spectrophotometrically by measuring the OD [260/280] ratio. An OD [260/280] ratio of 1.8 to 2.0 indicates good quality RNA that is devoid of protein and phenol contamination (Figure 1B). However, no RNA integrity information can be obtained from a spectrophotometric reading. RNA integrity can be assessed using several methods. The traditional method is visual inspection after electrophoresis on a formaldehyde agarose gel in the presence of a fluorescent dye such as ethidium bromide. Observation of two sharp bands for the large and the small subunit ribosomal RNAs (rRNA) with the intensity of the larger band being about twice that of the smaller band is indicative of intact RNA. While this method is inexpensive, the interpretation of the data is mostly subjective and requires approximately 200 ng of total RNA, which may be difficult if the sample is only available in limited quantities. This method can be refined by quantifying the intensity of the rRNA bands using an imager with densitometry scanning. While the values of the 28S/18S ratio can vary between different tissues or cell types, a ratio between 1 and 2 is indicative of an intact RNA sample.

A major improvement in RNA integrity analysis came with the introduction of microfluidics-based electrophoresis systems (Imbeaud et al. 2005) such as the Experion™ automated electrophoresis system (Bio-Rad) that combines integrity and concentration quantification of RNA in a single step from nanogram or picogram quantities (Figure 1A and 1C). In addition to generating a virtual gel, an electropherogram, and calculating the 28S/18S ratio, the Experion system’s software automatically calculates and reports an RNA quality indicator (RQI) value (Figure 1C) that reflects the integrity of the RNA sample based on several criteria (Denisov et al. 2008) of the electropherogram. The RQI value ranging from 1 (degraded) to 10 (intact) provides an objective assessment of the integrity of the RNA sample and it can be used to screen out degraded samples.

By assuring consistency in both purity and quality across all RNA samples, variability between biological replicates can be reduced (Figure 1D) (Imbeaud et al. 2005). When a batch of RNA samples has successfully met the standards of quality control, we recommend their immediate use in RT-qPCR experiments, or their conversion into much more stable cDNA by reverse transcription to preserve RNA integrity after the quality check.
4. Reverse Transcription

Given the prevalence of RNase in the environment, we recommend performing the reverse transcription of total RNA samples to cDNA immediately following the quality control assessment. This will avoid the risk of RNA samples degradation from multiple freeze/thaws before conversion to cDNA. For the RT step, the key is to assure consistent and complete coverage of the transcribed genome in the extracted RNA sample. Some genes are very long, but the associated RT products for these sequences cannot be as large, especially if the RT primers anneal only at the ends of each mRNA. By annealing primers at the end of the mRNA plus at random points within each mRNA sequence a good sampling of the population of each gene is obtained. This method, which is more representative than just annealing at the ends or at random sites within each transcript, provides better coverage of the transcribed genome.

A reverse transcription buffer should contain a mix of primers that are random in sequence allowing for a better sampling of the mRNA, RNase H the enzyme that specifically degrades RNA in DNA/RNA duplexes, a specific and sensitive reverse transcriptase with broad dynamic range for RNA amounts from 1µg to 1 pg and, a simple and fast protocol. The iScript™ (Bio-Rad) cDNA synthesis kits meet all of these criteria and are well suited for reverse transcription of any total RNA sample. We recommend that the same amount of total RNA be used and that reaction time be maintained for reverse transcription for all experimental samples to minimize variability between biological replicates. Reverse transcribed RNA can be stored frozen at −20°C or −80°C until use.

5. Primer and Amplicon Design

Both primer design and careful choice of target sequence are essential to ensure specific and efficient amplification of the products. Target sequences should be unique, 75–300 bp long with a GC content between 50–60%, and should not contain secondary structures. It is recommended that primers should have a GC content of 50–60% and a melting temperature of 55–65°C. Long G or C stretches in the primer should be avoided, but it is recommended to have G or C at the end of the primers.

A number of programs are available to help design primer pairs and pick target sequences. We recommend designing oligonucleotides using Primer-Blast (www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHomeAd), a program developed by NCBI that uses the algorithm Primer3 (Rozen and Skaletsky 2000). Primer sequences are compared (blasted) to the user-selected databases to ensure they are unique and specific for the gene of interest. The program MFOLD (http://mfold.bioinfo.rpi.edu/cgi-bin/dna-form1.cgi) can then be used to analyze the amplicon
for potential secondary structures that may prevent efficient amplification (Zuker 2003) (Figure 2). At least two sets of oligonucleotides should then be ordered and tested for their performance in a qPCR reaction.

6. qPCR Validation
A validated qPCR assay is one that has been assessed for the optimal range of primer annealing temperatures, reaction efficiency, and specificity using a standard set of samples (Bustin et al. 2009). This will assure that the reaction conditions, buffers, and primers have been optimized and that the cDNA samples are not contaminated with inhibitors of Taq polymerase. Bio-Rad has created a practical web resource (www.bio-rad.com/genomics/pcrsupport) for qPCR design and validation. The major points in assay validation are summarized below.

Determination of Annealing Temperature, Melt Curve Analysis, Gel Analysis of Amplicon, and No Template Control
A critical step in a PCR reaction is the annealing of the primers to their target sequences. It has to be performed at the right temperature for the primers to anneal efficiently to their targets, while preventing nonspecific annealing and primer-dimer formation. The fastest way to determine optimal annealing temperature is to use a thermal cycler equipped with a temperature gradient feature. All Bio-Rad thermal cyclers and real-time PCR instruments are available with a thermal gradient block option. A range of temperatures around the calculated Tm of the primers should be tested (Figure 3A).

It is important to check the specificity of the reaction by analyzing the PCR product. A melt curve analysis, performed at the end of the PCR cycles, will confirm specificity of primer annealing. The melt curve should display a single sharp peak (Figure 3B). In addition, at least one sample per primer pair should be run on an appropriate gel (agarose gel or polyacrylamide gel to resolve smaller DNA fragments) to confirm that the amplicon is of the expected size (Figure 4). Alternatively, an automated electrophoresis system such as the Experion automated electrophoresis system (using the Experion DNA 1K analysis kit) can be used to rapidly and precisely confirm the size of the amplicon.

Since most qPCR experiments involve the study of multiple genes in a single protocol it is important to design primers with similar annealing temperatures and to run the reactions at a temperature that is adequate for all the primer pairs to be studied during the same run. A duplicate, no template control (NTC) reaction should be included in every run for each primer pair to test buffers and solutions for DNA contamination and to assess for primer-dimers.

Establishment of a Standard Curve (to Evaluate PCR Efficiency)
The efficiency of a PCR reaction is a measure of the rate at which the polymerase converts the reagents (dNTPs, oligonucleotides, and template cDNA) to amplicon. The maximum increase of amplicon per cycle is two-fold representing a reaction that is 100% efficient and it is important to measure reaction efficiency as it is indicative of problems with the qPCR reaction that can cause artifactual results. Low-efficiency reactions (< 90%) may be caused by contaminating Taq inhibitors, high or suboptimal annealing temperature, old or inactive Taq, poorly designed primers, or amplicons with secondary structures. High reaction efficiency (> 105%) is generally the result of primer-dimers or nonspecific amplicons. The most common causes of both high and low reaction efficiencies are poorly calibrated pipettes or poor pipetting technique.

A standard curve is generally used to determine the reaction efficiency for any qPCR reaction. The template for this typically is a sample of cDNA or spiked plasmid cDNA in a sample extract. We recommend initially producing a 10-fold dilution series over eight points starting from the

Fig. 2. Effect of annealing temperature on amplicon secondary structure. Analysis of the amplicon sequence using MFOLD predicts a high secondary structure at 60°C (top) and low secondary structure at 65°C (bottom).
most concentrated cDNA sample, to ensure the standard curve covers all potential template concentrations that may be encountered during the study (i.e. broad dynamic range). For each dilution, a standard qPCR protocol is performed in triplicate for all the primer pairs to be used in the experiment and C_t values are determined. The standard curve is constructed by plotting the log of the starting quantity of the template against the C_t values obtained. The equation of the linear regression line, along with Pearson’s correlation coefficient (r) or the coefficient of determination (R^2), can then be used to evaluate whether the qPCR assay is optimized.

Ideally, the dilution series will produce amplification curves with tight technical replicates that are evenly spaced (Figure 5A). If perfect doubling occurs with each amplification cycle, the spacing of the fluorescence curves will be determined by the equation 2^n = dilution factor, where n is the number of cycles between curves at the fluorescence threshold (in other words, the difference between the C_t values of the curves). For example, with a 10-fold serial dilution of DNA, 2^n = 10. Therefore, n = 3.32, and the C_t values should be separated by 3.32 cycles. Evenly spaced amplification curves (Figure 5A) will produce a linear standard curve (Figure 5B) with the goal of 90% to 105% reaction efficiency.

The r or R^2 value of a standard curve represents how well the experimental data fit the regression line, that is, how linear the data are. Linearity, in turn, gives a measure of the variability across assay replicates and whether the amplification efficiency is the same for different starting template copy numbers. A significant difference in observed C_t values between replicates will lower the r or R^2 value. An r with an absolute value >0.990 or an R^2 value >0.980 is desirable for RT-qPCR reactions. Deletion of points at both ends of the standard curve may be required to obtain an acceptable slope (efficiency) and R^2 value. This will ultimately define the dynamic range of cDNA concentration for each primer pair with respect to sample dilution.

**qPCR Reagents, Instrument, and Analysis Program**

There are a wide variety of commercial qPCR reagent kits available. We recommend SsoFast™ EvaGreen® supermix from Bio-Rad, a mix that contains EvaGreen, which is a saturating dye that provides much better sensitivity than SYBR Green. Moreover, the Sso7d fusion polymerase and its optimized buffer are highly resistant to sample contaminants that may inhibit Taq to give good reproducibility.

A large number of qPCR detection instruments are commercially available and the following guidelines can be employed for selecting an appropriate instrument:

1. Sample volumes with a range of 10 to 50 µl in standard 96-well format, 0.2 ml, low-profile thermal cycler plates, strips, or tubes from any manufacturer.

2. Gradient-enabled thermal block with fast ramping.

3. No requirement for proprietary reagents, consumables, or fluorophores.
The CFX96™ and CFX384™ real-time PCR detection systems from Bio-Rad meet all of these criteria.

All qPCR instruments are packaged with data analysis software and we recommend that a good software package include the following features:

1. Flexibility to enter plate setup information such that well identifiers can be loaded and edited before, during, or after the run.
2. Ability to group wells from multiple experiments on one plate.
3. Built-in gene expression analysis that can normalize data to multiple reference genes (Vandesompele et al. 2002) and individual reaction efficiencies (Pfaffl 2001).
4. The ability to combine multiple plates of experimental data into a large gene study.

The CFX Manager™ software from Bio-Rad meets all of these criteria and comes packaged with the CFX96 and CFX384 real-time PCR detection systems.

7. Choice of Reference Genes

In RT-qPCR experiments, reference genes are used as controls to normalize the data by correcting for differences in quantities of cDNA used as template (Gutierrez et al. 2008, Huggett et al. 2005, Vandesompele et al. 2002). A perfect reference gene is therefore one that does not exhibit changes in expression between samples from various experimental conditions or time points. Several genes such as GADPH, β-actin, or rRNA are often used as reference genes. However, a number of studies have indicated that the expression of these genes may vary considerably between tissues or between treatments, which may make them unsuitable for use as reference genes. Reference genes must therefore be carefully selected based on experimental data and we recommend the following protocol:

1. Extract the total RNA from at least one or two samples from each experimental condition or time point and confirm their purity and quality (see Steps 2 and 3 above).
2. Normalize the sample concentration and perform reverse transcription PCR (see Step 4) from the same volume of each sample.
3. Perform the qPCR experiment using the same volume of each cDNA sample as a template.
4. Compare C_{t} for each gene between the different conditions.

The C_{t} differences (ΔC_{t}) between all the conditions or time points should be less than 0.5. Since any given gene may be affected by the experimental conditions, we recommend simultaneously testing at least three or four putative reference genes that belong to different metabolic pathways and to use the geometric average of the one(s) with minimal expression differences to achieve accurate normalization (Vandesompele et al. 2002).
8. Experimental Reproducibility
There are two sources of variability in a gene expression experiment that may affect the results:

1. Biological variability which is due to inherent differences in gene expression levels between individual organisms, tissues, or cell culture samples.
2. Technical variability in the experimental process itself which is typically associated with pipetting and poorly calibrated pipettes.

To mitigate the effect of biological and technical variability, it is generally accepted that at least three biological and two technical replicates per biological replicate be performed for each experiment (Figure 6). If the experiment compares gene expression levels between control and treated samples, the three biological replicates should be samples that were treated in separate and independent experiments.

**Conclusions**
RT-qPCR is the method of choice for gene expression analysis because of its high sensitivity from samples of very low RNA concentrations. However, in order to assure accurate and reproducible, quantitative data, strict standard operating procedures should be followed. All experimental details and controls should be accurately reported when publishing gene expression experiments. This will allow proper assessment of the data by the scientific community, and enable informed comparison of expression data between labs and between experiments.

In summary the key steps for most RT-qPCR experiments include:

1. Experimental design with appropriate number of biological replicates and proper control samples.
2. Sample procurement which requires adherence to strict experimental protocols for acquisition, processing, and storage to assure reproducibility and minimize standard deviations between replicates.
3. Quality control of RNA for purity and integrity.
4. Reverse transcription to convert the total RNA to cDNA.
5. qPCR experiments with proper primer design, choice of target sequence and reference genes, and technical replicates.

The MIQE guidelines were written to provide all the parameters that should be met to publish acceptable results from qPCR experiments. Here we have shown how the most important items in the MIQE checklist can be addressed in practice to assure high quality results.

**Fig. 6. Experimental replicates.** All experiments should be designed with a combination of biological and technical replicates. This illustrates a simple experiment with triplicate biological samples from control and treatment/experimental conditions. For each biological sample, three technical replicates are recommended for the gene of interest as well as for the reference gene(s). This results in a total of at least 36 samples plus the duplicate NTC for a total of >40 wells.
References


Bio-Rad’s real-time thermal cyclers are licensed real-time thermal cyclers under Applera’s United States Patent 6,814,934 B1 for use in research, human in vitro diagnostics, and all other fields except veterinary diagnostics. This product is covered by one or more of the following U.S. patents or their foreign counterparts owned by Eppendorf AG: U.S. Patent Nos. 6,767,512 and 7,074,367.

EvaGreen is a trademark of Biotium, Inc. Bio-Rad Laboratories, Inc. is licensed by Biotium, Inc. to sell reagents containing EvaGreen dye for use in real-time PCR, for research purposes only.

Practice of the patented 5’ Nuclease Process requires a license from Applied Biosystems. The purchase of these products includes an immunity from suit under patents specified in the product insert to use only the amount purchased for the purchaser’s own internal research when used with the separate purchase of Licensed Probe. No other patent rights are conveyed expressly, by implication, or by estoppel. Further information on purchasing licenses may be obtained from the Director of Licensing, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94440, USA.

CFX Manager software is for use with authorized real-time instrument or thermal cycler. Please see user manual for complete licensing statement.

Nanodrop is a trademark of Thermo Fisher Scientific, Inc.

Information in this tech note was current as of the date of writing (2009) and not necessarily the date this version (rev A, 2009) was published.