Unlocking the MIQE secrets to successful RT-qPCR gene expression analysis

Jo Vandesompele
professor, Ghent University
co-founder and CEO, Biogazelle

Bio-Rad seminar tour China
May 28-31, 2012
Do you sometimes feel frustrated with scientific literature?

- important information is missing
  - primer sequences
  - DNAse treatment (yes / no)
  - RNA quality control
  - experiment design and run layout
  - ...

- lack of best practice
  - only 1 (non-validated) reference gene
  - no assay validation (specificity, efficiency)
  - statistics on linear data instead of logarithmic transformed data
  - ...

The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments

Stephen A. Bustin, Vladimir Benes, Jeremy A. Garson, Jan Hellemans, Jim Huggett, Mikael Kubista, Reinhold Mueller, Tania Nolan, Michael W. Pfaffl, Gregory L. Shipley, Jo Vandesompele, and Carl T. Wittwer

BACKGROUND: Currently, a lack of consensus exists on how best to perform and interpret quantitative real-time PCR (qPCR) experiments. The problem is exacerbated by a lack of sufficient experimental detail in many publications, which impedes a reader’s ability to evaluate critically the quality of the results presented or to repeat the experiments.

SUMMARY: Following these guidelines will encourage better experimental practice, allowing more reliable and unequivocal interpretation of qPCR results.

© 2009 American Association for Clinical Chemistry

The fluorescence-based quantitative real-time PCR (qPCR)\textsuperscript{15} (1–3), with its capacity to detect and mea-
The MIQE guidelines are summarized in a checklist

<table>
<thead>
<tr>
<th>Item to check</th>
<th>Importance</th>
<th>Item to check</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental design</td>
<td>qPCR oligonucleotides</td>
<td>qPCR protocol</td>
<td>qPCR validation</td>
</tr>
<tr>
<td>Definition of experimental and control groups</td>
<td>E Primer sequences</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Number within each group</td>
<td>E RTPrimerID identification number</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Assay carried out by the core or investigator’s laboratory?</td>
<td>D Probe sequences</td>
<td>DqPCR target information</td>
<td>D</td>
</tr>
<tr>
<td>Acknowledgment of authors’ contributions</td>
<td>D Location and identity of any modifications</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Sample</td>
<td>Manufacturer of oligonucleotides</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Description</td>
<td>E Purification method</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Volume/number of sample processed</td>
<td>D qPCR protocol</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Microcalcification or macrocalcification</td>
<td>E Complete reaction conditions</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processing procedure</td>
<td>E Reaction volume and amount of cDNA/DNA</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If frozen, how and how quickly?</td>
<td>E Primer, (probe), Mg²⁺, and dNTP concentrations</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If fixed, with what and how quickly?</td>
<td>E Polymerase identity and concentration</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Sample storage conditions and duration (especially for FFPE samples)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Details of diffusion or fixation treatment</td>
<td>E Reaction setup (manual/robotic)</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contamination assessment (DNA or RNA)</td>
<td>E Manufacturer of qPCR instrument</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nucleic acid quantification</td>
<td>E qPCR validation</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Instrument and method</td>
<td>E Evidence of optimization (from gradients)</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Purify (A₂₆₀/A₂₈₀)</td>
<td>D Specificity (gel, sequence, melt, or digested)</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Yield</td>
<td>E For SYBR Green Cq of the NSC</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>RNA integrity methodinstrumentation</td>
<td>E Calibration curves with slope and y-intercept</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>RNA/DNA or Cq of 3' and 5' transcripts</td>
<td>E PCR efficiency calculated from slope</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Electrophoresis traces</td>
<td>D Cq for PCR efficiency or 15x</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Correlation testing Cq (dilutions, spikes, or other)</td>
<td>E Slope of calibration curve</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Reverse transcription</td>
<td>E Linear dynamic range</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Complete reaction conditions</td>
<td>E Cq variation at LOD</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Amount of RNA and reaction volume</td>
<td>E Cq throughout range</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Priming oligonucleotide (if using GSP) and concentration</td>
<td>E Evidence for LOD</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Reverse transcription and concentration</td>
<td>E If multiple, efficiency and LOD of each assay</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature and time</td>
<td>E Data analysis</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Manufacturer of reagents and catalogue numbers</td>
<td>D qPCR analysis program (source, version)</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>qPCR target information</td>
<td>Results for NTCs</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene symbol</td>
<td>E Identification of number and choice of reference genes</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Sequence accession number</td>
<td>E Description of normalization method</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Location of amplicon</td>
<td>D Number and concordance of biological replicates</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Amplification length</td>
<td>E Number and stage (reverse transcription or qPCR) of technical replicates</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>In silico specificity screen (BLAST, and so on)</td>
<td>E Repeatability (intrasample variation)</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Gene symbols, retroelements, or other homologs?</td>
<td>D Reproducibility (interassay variation, CV)</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Sequence alignment</td>
<td>E Primer analysis</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Secondary structure analysis of amplicon</td>
<td>E Statistical methods for results significance</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>Location of each primer by exon or intron (if applicable)</td>
<td>E Software (source, version)</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>What splice variants are targeted?</td>
<td>E Cq or raw data submission with RDML</td>
<td>D</td>
<td>D</td>
</tr>
</tbody>
</table>

* All essential information (E) must be submitted with the manuscript. Desirable information (D) should be submitted if available. If primers are from RTPrimerID, information on qPCR target, oligonucleotides, protocols, and validation is available from that source.

** FFPE, formalin-fixed, paraffin-embedded; RIN, RNA integrity number; RQI, RNA quality indicator; GSP, gene-specific priming; dNTP, deoxynucleoside triphosphate.

*** Assessing the absence of DNA with a no-reverse transcription assay is essential when first extracting RNA. Once the sample has been validated as DNA free, inclusion of a no-reverse transcription control is desirable but no longer essential.

** Disclosure of the probe sequence is highly desirable and strongly encouraged; however, because not all vendors of commercial predesigned assays provide this information, it cannot be an essential requirement. Use of such assays is discouraged.
MIQE is there to help you!

Look out! The slope is steep and the road is slippery.
Critical factors contributing to successful RT-qPCR results

Derveaux et al., Methods, 2010
Full text papers available on Biogazelle website

http://www.biogazelle.com
RNA sample quality is determined by its integrity & purity

- various RNA quality measures are available
  - microfluidic electrophoresis (e.g. Bio-Rad’s Experion)
  - 5’-3’ mRNA ratio
  - SPUD inhibition assay
  - UV-VIS absorption spectrum (260:280 > insensitive + stable pH!)
RNA purity assessment using the SPUD assay

- spiking of synthetic sequence lacking homology with any known human sequence into RNA

\[
\begin{align*}
&\text{SPUD} + H_2O \\
&\text{SPUD} + \text{heparin} \\
&\text{SPUD} + \text{RNA1} \\
&\text{SPUD} + \text{RNA2} \\
&\text{SPUD} + \text{RNA3} \\
\end{align*}
\]

\[\Delta C_q > 1: \text{presence of inhibitors}\]

Nolan et al., Anal Biochem, 2006
RNA integrity assessment using 5′-3′ mRNA assay

- universally expressed low abundant reference (HPRT1)
- anchored oligo(dT) reverse transcription (iScript select)

![Graph showing increasing delta-Cq values upon artificial RNA degradation](image-url)
RNA integrity assessment using 5’-3’ mRNA assay

- universally expressed low abundant reference (HPRT1)
- anchored oligo(dT) reverse transcription (iScript select)

5’

AAAAAAA

3’

Cq 5’

Cq 3’

decreasing RQI values upon artificial RNA degradation

RQI 10.0
RQI 9.0
RQI 6.8
RNA quality has impact on reference gene stability ranking

<table>
<thead>
<tr>
<th>Step*</th>
<th>Degraded RNA (CRS samples)</th>
<th>Intact RNA (CRS samples)</th>
<th>Degraded RNA (NP samples)</th>
<th>Intact RNA (NP samples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HPRT1</td>
<td>GAPD</td>
<td>HPRT1</td>
<td>YWHAZ</td>
</tr>
<tr>
<td>2</td>
<td>YWHAZ</td>
<td>YWHAZ</td>
<td>ACTB</td>
<td>B2M</td>
</tr>
<tr>
<td>3</td>
<td>B2M</td>
<td>RPL31A</td>
<td>RPL31A</td>
<td>RPL31A</td>
</tr>
<tr>
<td>4</td>
<td>TBP</td>
<td>B2M</td>
<td>GAPD</td>
<td>UBC</td>
</tr>
<tr>
<td>5</td>
<td>RPL31A</td>
<td>UBC</td>
<td>TBP</td>
<td>GAPD</td>
</tr>
<tr>
<td>6</td>
<td>UBC</td>
<td>HPRT1</td>
<td>YWHAZ</td>
<td>HMBS</td>
</tr>
<tr>
<td>7</td>
<td>ACTB</td>
<td>TBP</td>
<td>HMBS</td>
<td>HPRT1</td>
</tr>
<tr>
<td>8</td>
<td>GAPD</td>
<td>ACTB</td>
<td>SDHA</td>
<td>SDHA</td>
</tr>
<tr>
<td>9</td>
<td>HMBS- SDHA</td>
<td>HMBS- SDHA</td>
<td>B2M- UBe</td>
<td>ACTB- TBP</td>
</tr>
</tbody>
</table>

Perez-Novo et al., Biotechniques, 2005
Degraded RNA is characterized by more expression noise

Vermeulen et al., Nucleic Acids Research, 2011
Degraded RNA is characterized by more expression noise

Vermeulen et al., Nucleic Acids Research, 2011
Impact of RNA quality on results

Sample group size

Significance (-10 log p-value)

Low quality

High quality

Intact RNA delivers more significant results

Vermeulen et al., Nucleic Acids Research, 2011
Determination of RNA quality control cut-off requires work

- cut-off depends on various factors
  - method (RT-qPCR, microarray, RNA-sequencing)
  - expression difference of the target(s)
  - abundance of the target(s)
  - stability of the target(s)
  - sample material (fresh frozen vs. FFPE)
  - application: single gene study vs. multi-gene classification signature

- pilot experiment is needed with positive and negative controls to establish minimal acceptable RNA quality
Good experiment design is key

- very important to achieve biologically meaningful and statistically significant results
- often underestimated or neglected part of the workflow

- power analysis
  - define number and type of replicates (technical vs. biological), control and reference samples, and genes
- choose run layout
  - sample maximization
  - gene maximization
  - perform inter-run calibration if needed
Run lay-out definition: sample vs. gene maximization

How to set-up an experiment with 3 genes of interest (GOI), 3 reference genes (REF), 11 samples (S) and 1 no template control (NTC)?
Run lay-out definition: sample vs. gene maximization

- Sample maximization is to be preferred
  - No increase in variation due to absence of inter-run variation
  - More cost-effective (fewer reactions)
  - Easier to set-up (make master mix only once)
  - Suitable for retrospective studies and controlled experiments

- Gene maximization
  - Introduces (under-estimated) inter-run variation
  - Applicable for prospective studies or large studies in which the number of samples do not fit on the plate anymore
  - Inter-run variation can be measured and corrected for using inter-run calibrators (IRC)
Inter-run calibration corrects inter-run differences

- possible on two levels
  - Cq values
  - normalized relative quantities
    - greater flexibility
    - fewer calculations

- rules for inter-run calibration
  - the more inter-run calibrators, the better
  - use the same set of inter-run calibrators in all runs
  - measure all genes in the inter-run calibrators

- specialised software is needed (e.g. qbase$PLUS$)
Why do we need normalization?

- 2 sources of variation in gene expression results
  - biological variation (true fold changes)
  - experimentally induced variation (noise and bias)

- purpose of normalization is removal or reduction of the experimental variation
  - input quantity: RNA quantity, cDNA synthesis efficiency, ...
  - (input quality: RNA integrity, RNA purity, ...)
Various normalisation strategies have been proposed.

The use of reference genes (housekeeping genes) is the most universal and most appropriate method for normalization.

Huggett et al., Genes and Immunity, 2005
Candidate reference gene expression levels fluctuate

- RT-qPCR analysis of 5 reference genes (belonging to different functional and abundance classes) on 7 normal human blood samples

15 fold difference between A and B if normalized by a single non-validated reference gene (\textit{ACTB} or \textit{HMBS})
the geNorm solution to the normalisation problem

- framework for qPCR gene expression normalisation using the reference gene concept
  - quantified errors related to the use of a single reference gene
    (> 3 fold in 25% of the cases; > 6 fold in 10% of the cases)
  - developed a robust algorithm for assessment of expression stability of candidate reference genes
  - proposed the geometric mean of multiple reference genes for accurate and reliable normalisation

- Vandesompele et al., Genome Biology, 2002

Research

**Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes**

Jo Vandesompele, Katleen De Preter, Filip Pattyn, Bruce Poppe, Nadine Van Roy, Anne De Paepe and Frank Speleman
geNorm M value is an expression stability parameter

- pairwise variation V (between any 2 candidate reference genes)

<table>
<thead>
<tr>
<th>gene A</th>
<th>gene B</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample 1</td>
<td>a1</td>
</tr>
<tr>
<td>sample 2</td>
<td>a2</td>
</tr>
<tr>
<td>sample 3</td>
<td>a3</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>sample n</td>
<td>an</td>
</tr>
</tbody>
</table>

standard deviation = V

- gene stability measure M
  average pairwise variation V of a given reference gene with all other candidate reference genes

- iterative procedure of removing the worst reference gene followed by recalculation of M-values
geNorm analysis is straightforward and helps you on 2 levels

- ranking of candidate reference genes according to their stability
- determination of how many genes are required for reliable normalization
Calculation of a sample specific normalization factor as the geometric mean of the reference genes

geometric mean = \((a \times b \times c)^{1/3}\)

arithmetic mean = \(\frac{a + b + c}{3}\)

- controls for outliers
- compensates for differences in expression level between the reference genes
Geometric averaging is robust (insensitive to outliers)
The use of multiple reference genes for normalization gives statistically more significant results.

Kaplan-Meier cancer patient survival curve

log rank statistics

NF4

NF1

Hoebeeck et al., Int J Cancer, 2006
The use of multiple reference genes for normalization enables accurate measurement of small expression differences.

LEMD3 levels drop to 50% in patient with NMD.

Hellemans et al., Nature Genetics, 2004
geNorm is the *de facto* standard for reference gene validation and normalization

- > 4,500 citations of the geNorm technology
- > 15,000 geNorm software downloads in 100 countries
Large and active geNorm discussion community

> 1000 members, almost 2000 posts

http://tech.groups.yahoo.com/group/genorm/
genorm\textsuperscript{PLUS} is a much improved version built into qbase\textsuperscript{PLUS}

<table>
<thead>
<tr>
<th>Feature</th>
<th>classic geNorm</th>
<th>improved geNorm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Excel Windows</td>
<td>qbase\textsuperscript{PLUS} Win, Mac, Linux</td>
</tr>
<tr>
<td>Speed</td>
<td>1x</td>
<td>20x</td>
</tr>
<tr>
<td>Expert interpretation + report</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Ranking best 2 genes</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Handling missing data</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Raw data (Cq) as input</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
genorm\textsuperscript{PLUS} result interpretation is easy

- expert report without need to understand formulas
- time saver
- higher confidence in the results

\begin{table}
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{Data} & \\
\hline
\textbf{Input} & geNorm analysis was initiated on \textit{10 samples and 10 reference targets}. & \\
\textbf{Missing values} & \textit{none} & \\
\textbf{Experiment design} & \textit{The run lay-out is perfect. All samples are measured in the same run for a given reference target (i.e. sample maximization strategy according to Hellemans et al., Genome Biology, 2007).} & \\
\hline
\textbf{Results} & & \\
\textbf{Optimal reference target selection} & The optimal number of reference targets in this experimental situation is 2 (geNorm V < 0.15 when comparing a normalization factor based on the 2 or 3 most stable targets). As such, the optimal normalization factor can be calculated as the geometric mean of reference targets \textit{HMBS and SDHA}. & \\
\textbf{Reference target stability} & \textit{High reference target stability} (average geNorm M ≤ 0.5). This is typically seen when evaluating candidate reference targets on a homogeneous set of samples (e.g. untreated cultured cells, or blood from normal individuals). More reference values in Table 1 in Hellemans et al., Genome Biology, 2007. & \\
\hline
\end{tabular}
\end{table}
Tips and tricks

- expression levels of reference gene and gene of interest do not need to be similar
- no need to measure reference genes in the same run (plate) as genes of interest
- geNorm pilot experiment can fit in a single 96-well plates
  - 10 representative samples x 8 candidate reference genes (no replicates, no controls, no standard curve)
  - results that can be used in the future if experimental conditions do not change
- at least 2 reference genes are needed
  - quality control of the expression stability and normalization factor
  - more accurate results
Global mean normalization method is preferred when many genes are simultaneously measured

- novel and powerful normalization strategy
- only useful when measuring a large and unbiased set of genes
  - e.g. microRNA gene expression profiling of 755 human miRs
  - e.g. measuring panels of pathway genes
- maximal reduction of technical noise
- improved identification of differentially expressed genes
  - Mestdagh et al., Genome Biology, 2009 (original global mean)
  - D’haene et al., Methods Mol Biol, 2012 (improved global mean)
  - integrated in qbasePLUS

Method

A novel and universal method for microRNA RT-qPCR data normalization
Pieter Mestdagh*, Pieter Van Vlierberghe*, An De Weer*, Daniel Muth†, Frank Westermann†, Frank Speleman* and Jo Vandesompele*

Addresses: *Center for Medical Genetics, Ghent University Hospital, De Pintelaan 185, Ghent, Belgium. †Department of Tumour Genetics, German Cancer Center, Im Neuenheimer Feld 280, Heidelberg, Germany.

Correspondence: Jo Vandesompele. Email: Joke.Vandesompele@UGent.be
How can I extract meaningful biological information from qPCR data?
How can I extract meaningful biological information from qPCR data?

- error bars that make sense and provide confidence in the results
- differences that are meaningful and explain the phenotype or mechanism under investigation
qPCR data-analysis witnessed 3 generations of quantification models

- Livak and Schmittgen (2001)
  - 100% PCR efficiency
  - 1 reference gene

- Pfaffl (2001)
  - adjusted PCR efficiency
  - 1 reference gene

- Hellemans et al. (2007) - qbasePLUS
  - adjusted PCR efficiency
  - multiple reference genes
  - error propagation rules

\[
NRQ = 2^{\Delta \Delta C_q}
\]

\[
NRQ = \frac{E_{g oi}^{\Delta C_q, goi}}{E_{r ef}^{\Delta C_q, ref}}
\]

\[
NRQ = \sqrt[n]{\prod_{i=1}^{n} E_{r ef_i}^{\Delta C_q, ref_i}}
\]
qBase is a state-of-the-art quantification model with error propagation

Step 1
Calculation of the normalization factor \( NF \) for sample \( k \) based on the \( RQs \) of the reference genes.

\[
NF_k = \left( \frac{\sum_{i=1}^{n} Q_{i,k, measured} - Q_{i,k, predicted}}{V_{i,k}} \right)^{\lambda - 2}
\]

\( \lambda \) is a parameter that can be optimized to minimize the error.

Step 2
Calculation of the coefficient of variation \( CV \) of a given reference gene \( p \) across all samples.

\[
CV_p = \sum_{k=1}^{n} \left( \frac{NRQ_{k,p} - \bar{NRQ}_p}{\bar{NRQ}_p} \right)^{2}
\]

\( \bar{NRQ}_p \) is the mean NRQ of the reference gene \( p \) across all samples.

Step 3
Calculation of the mean coefficient of variation for all reference genes.

\[
\sum_{p=1}^{P} CV_p
\]

Step 4
Calculation of the arithmetic mean \( M_{p}^{\text{pairwise}} \) of all pairwise variations \( V_{p,k} \) of a given reference gene \( p \) with all other tested reference genes \( p' \).

\[
M_{p}^{\text{pairwise}} = \frac{1}{P-1} \sum_{k=1}^{n} V_{p,k}^{\text{pairwise}}
\]

Step 5
Calculation of the arithmetic mean \( M_{p}^{\text{overall}} \) of all pairwise variations \( V_{p,k} \) of a given reference gene \( p \) with all other tested reference genes \( p' \), for the same gene \( j \).

\[
M_{p}^{\text{overall}} = \frac{1}{P} \sum_{k=1}^{n} V_{p,k}^{\text{overall}}
\]

Reference gene and IRC stability parameter \( M \)

Since normalization and inter-run calibration are highly analogous, quality evaluation using the stability parameter \( M \) is similar as well. Therefore, both methods are explained in parallel.

Step 1
Calculation of the \( s \times 1 \) matrix \( A_{p,k}^{\text{NRO}} \) in which the \( k \)-th element is the log, transformed ratio between the \( RQs \) of two IRCs \( m \) and \( m' \) for the same gene \( j \) within a run \( k \); matrix \( A_{p,k}^{\text{NRO}} \) is calculated in an analogous manner.

\[
A_{p,k}^{\text{NRO}} = \left[ \frac{NRQ_{k,p}^{m} - NRQ_{k,p}^{m'}}{NRQ_{k,p}^{m}} \right]
\]

Step 2
Calculation of the mean \( RQ \) for all samples \( k \) and a given reference gene \( p \):

\[
\bar{NRQ}_p = \frac{1}{s} \sum_{k=1}^{s} NRQ_{k,p}
\]

Normalization and inter-run calibration

The procedures for normalization and inter-run calibration are highly analogous and are therefore described in parallel.

\[
\text{Conversion of Cq values into relative quantities}
\]

Step 1
Calculation of the average Cq value for all replicates of the same gene/sample combination \( jk \) within a given run \( f \):

\[
Cq_{jk} = \frac{1}{n} \sum_{i=1}^{n} Cq_{i,jk}
\]

Step 2
Transformation of mean Cq value into RQ using the gene specific PCR efficiency \( \Delta_{e} \), with minimization of the overall error:

\[
\Delta_{Q}\text{calc} = Cq_{jk} - \Delta_{Q}\text{ref}
\]

Step 3
Calculation of the mean of RQs:

\[
\bar{RQ}_p = \frac{1}{s} \sum_{k=1}^{s} RQ_{k,p}
\]

Step 4
Calculation of the mean of CNBCs:

\[
\bar{CNBC}_p = \frac{1}{s} \sum_{k=1}^{s} CNBC_{k,p}
\]

Step 5
Calculation of the mean of NRQs:

\[
\bar{NRQ}_p = \frac{1}{s} \sum_{k=1}^{s} NRQ_{k,p}
\]

Step 6
Calculation of the mean of NRQs:

\[
\bar{NRQ}_p = \frac{1}{s} \sum_{k=1}^{s} NRQ_{k,p}
\]

Reference gene and IRC stability parameter \( M \)

Since normalization and inter-run calibration are highly analogous, quality evaluation using the stability parameter \( M \) is similar as well. Therefore, both methods are explained in parallel.

Step 1
Calculation of the \( s \times 1 \) matrix \( A_{p,k}^{\text{NRO}} \) in which the \( k \)-th element is the log, transformed ratio between the \( RQs \) of two IRCs \( m \) and \( m' \) for the same gene \( j \) within a run \( k \); matrix \( A_{p,k}^{\text{NRO}} \) is calculated in an analogous manner.

\[
A_{p,k}^{\text{NRO}} = \left[ \frac{NRQ_{k,p}^{m} - NRQ_{k,p}^{m'}}{NRQ_{k,p}^{m}} \right]
\]

Step 2
Calculation of the mean of NRQs:

\[
\bar{NRQ}_p = \frac{1}{s} \sum_{k=1}^{s} NRQ_{k,p}
\]

Step 3
Calculation of the mean of CNBCs:

\[
\bar{CNBC}_p = \frac{1}{s} \sum_{k=1}^{s} CNBC_{k,p}
\]

Step 4
Calculation of the mean of NRQs:

\[
\bar{NRQ}_p = \frac{1}{s} \sum_{k=1}^{s} NRQ_{k,p}
\]
The qBase paper is cited more than 500 times

Method

qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data

Jan Hellemans, Geert Mortier, Anne De Paepe, Frank Speleman and Jo Vandesompele

Address: Center for Medical Genetics, Ghent University Hospital, De Pintelaan, B-9000 Ghent, Belgium.

Correspondence: Jo Vandesompele. Email: Joke.Vandesompele@UGent.be

Published: 9 February 2007


The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2007/8/2/R19

Received: 31 August 2006

Revised: 7 December 2006

Accepted: 9 February 2007

© 2007 Hellemans et al.; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
qbasePLUS is most powerful, flexible and user-friendly real-time PCR data-analysis software

- based on Ghent University’s geNorm and qBase technology
- up to hundred 384-well plates
- multiple reference genes for accurate normalization
- detection and correction of inter-run variation
- error propagation
- automated analysis
  - genormPLUS
  - biostatistics wizard

http://www.qbaseplus.com
qbasePLUS is the only 3rd party qPCR data-analysis software that is MIQE and RDML compliant

- RDML universal file format compatible
- peer-reviewed data-analysis algorithms
  - multiple reference gene normalization
  - global mean normalization
  - inter-run calibration
  - error propagation
- assay information
- data quality control
  - reference gene stability analysis
  - NTC analysis
  - PCR efficiency calculation
  - replicate variability assessment
qbasePLUS version 2 is a major new release

- copy number analysis
  - automatic calling of copy number status

- sample property annotation
  - sample property table (edit/import)
  - result visualization and rescaling (e.g. control group)

- statistics
  - intuitive wizard
  - 4 easy steps to select proper test

- high-throughput analysis
  - global mean normalization
  - auto exclusion of spurious data points
  - new calculation engine with recalculate button
How to get qbasePLUS?

- basic license
  - no high-throughput features
  - no biostatistics
  - expires after 1 year
- premium license
  - powerful high-throughput features
  - copy number variant analysis
  - biostatistics
  - everlasting license

- new CFX customers get premium qbasePLUS for free
- promotion until the end of June 2012
  - premium license at 35% discount
  - promotion website

http://tinyurl.com/co5kgaf
Summary

- Take home message: successful gene expression analysis relies on validation at each step of the RT-qPCR workflow

- Critical MIQE factors contributing to successful RT-qPCR results
  - RNA quality
  - Reference gene normalization
  - Experiment design
  - Data-analysis
  - Standardized reporting (MIQE & RDML)

- Derveaux et al., Methods, 2010
  *How to do successful gene expression analysis using real-time PCR*
Acknowledgments 谢谢

- UGent
  - Joëlle Vermeulen
  - Stefaan Derveaux
  - Pieter Mestdagh
  - Filip Pattyn
  - Steve Lefever
  - Katleen De Preter
  - Frank Speleman

- Biogazelle
  - Barbara D’haene
  - Jan Hellemans
  - Gaëlle Van Severen

- RDML consortium
- MIQE consortium