biogazelle

accelerating your analysis

quality control

Easy.Fast.Reliable.

Subject

Summary

Detailed information

References

Technologiepark 3, B-9052 Zwijnaarde, Belgium
www.biogazelle.com

Subject

qbase^{PLUS} provides data quality control on various levels to improve the accuracy and reliability of your results.

Summary

qbase^{PLUS} contains four levels of quality control: control on the PCR replicate variation, assessment of positive and negative control samples, determination of the reference target expression stability and evaluation of deviating normalization factors. Parameters for quality control of the data can be accessed in the Quality control settings window trough the project explorer (Setting > Quality control settings). The parameters do not affect the results of the analysis, but define the required precision and accuracy of the analysis.

Detailed information

Quality control is an important feature in qbase^{PLUS}. The program contains several types of quality control that can be accessed by double-clicking the Quality control icon in the project explorer:

- quality control on technical replicates
- quality control on positive and negative controls
- quality control on the stability of reference targets

A fourth type of quality control, that on the Normalization factors is available in the Intermediate results section.

The parameters for quality control can be defined in the Quality control setting window (Project Explorer > Settings > Quality control settings). These parameters do not affect the results of the analysis, but define the required precision and accuracy of the analysis.

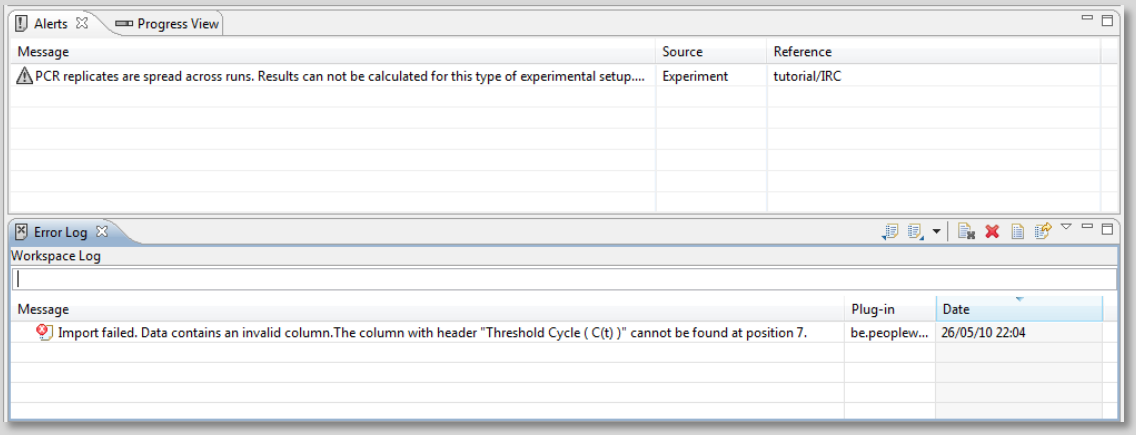
More information about post-PCR quality control can be found in the paper 'The Importance of Quality Control During qPCR Data Analysis' [D'haene and Hellemans, International Drug Discovery, 2010]. This and other papers are available on the Biogazelle website (http://www.biogazelle.com/resources/articles).

Technical replicates

qbase^{PLUS} automatically deals with technical replicates or repeated measurements, which are recognized as different PCR wells with an identical sample and target name. Importantly, **PCR replicates should be measured in the same run**. If PCR replicates are measured in different runs, an alert will be displayed in the Alert window ('PCR replicates are spread across runs. Results cannot be calculated for this type of experimental setup.') (Figure 1).

The Cq values of all (non-empty) wells of a replicate group are averaged at the very beginning of the calculation process. Outliers can be excluded so they do not take part in the calculations.

Figure 1 – Alert for PCR replicates spread across runs



Step 1 - Inspecting replicate quality control

The Replicates window (Project explorer > Quality control) contains an overview of all replicate groups (defined by wells with the same target and sample name) (Figure 2). qbase^{PLUS} can selectively display the failed replicates, which are replicates that do not meet the quality threshold as defined in the Quality control settings (see step 2). This is done by checking the appropriate box in the Well selection box. The summary on top of the list contains the pass rate of the replicates and informs the user about the overall replicate quality of the experiment.

Step 2 - Excluding bad replicates

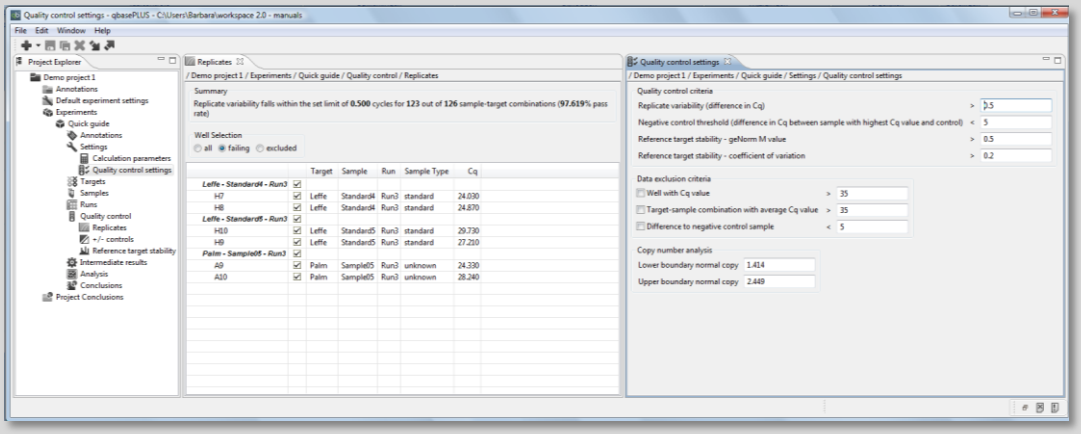
Failing replicates can be excluded from further calculations by removing the tick in front of the well (second column in the table) (Figure 3). The number of failed replicates is determined by the quality of the data in combination with the Quality control settings . Qbase^{PLUS} flags bad replicates based on a user defined maximum allowed difference in Cq value. This value can be defined in the Quality control settings window in the Replicate variability box (Figure 2). By default, the Replicate variability threshold is set at 0.5, which means that the difference in Cq value between the replicate with the highest Cq value and the replicate with the lowest Cq value (ΔCq) must be smaller than 0.5 cycles.

If a more stringent quality control is used (which is equal to a lower threshold value), more replicates will fail the quality control: the list of failed replicates will be longer and the summary will report a lower pass rate.

As of version 2.0, qbase^{PLUS} can also selectively display the excluded replicates.

Important: replicated wells should only be excluded if there is a good reason for it (e.g. abnormal melt curve, no sample added). When in doubt, keep all replicates, as the higher replicate variability will simply result in a larger propagated error on the final result.

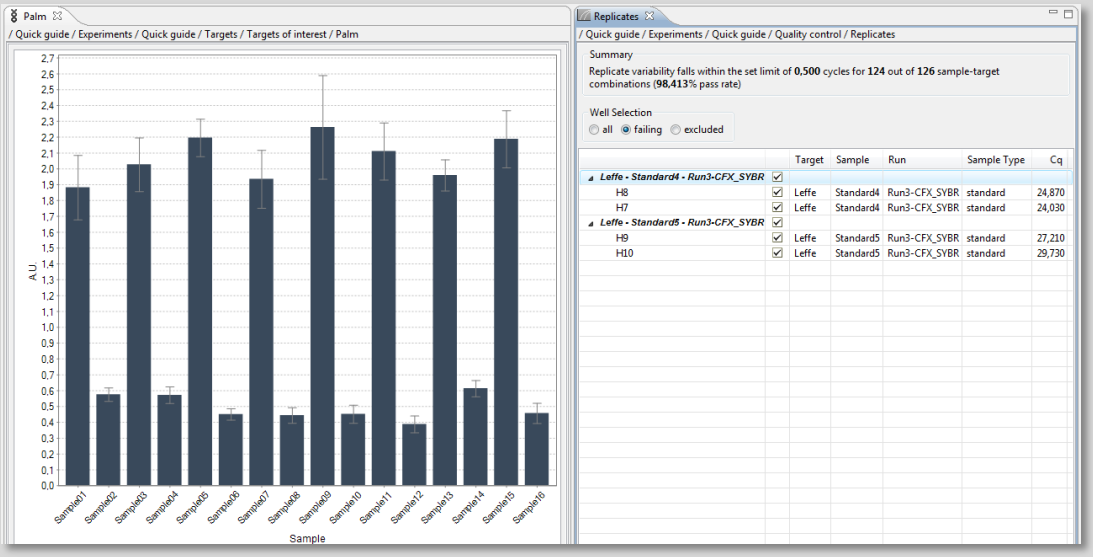
Figure 2 - PCR replicate quality control



PCR replicates (repeated measurements of the same sample in the same run) are useful for a number of reasons: (i) it allows for quality control on the technical reproducibility of the qPCR data, (ii) provides better accuracy, and (iii) allows the generation of results when individual qPCR reactions failed. Strict guidelines on the optimal number of PCR replicates cannot be given since it depends on the purpose of the study (e.g. diagnostics versus research), the quality of the PCR assay, the precision of the qPCR instrument, the Cq determination method, and the pipetting skills of the operator. Ultimately, it all comes down to determine how much confidence is needed for a given data point.
Important to note is that the biological variability is often much larger than the technical variability. Therefore, it is acceptable to omit PCR replicates when the sample size is sufficiently large or when a screening experiment is run (as is often the case for miRNA expression profiling studies).

By placing the Quality control settings window next to the Replicates window, you can see the effect of changing the Replicate variability threshold in real-time (Figure 3). Users interested in flexible use of windows in qbase^{PLUS} should read the 'User interface' chapter.

Figure 3 – Excluding bad PCR replicates



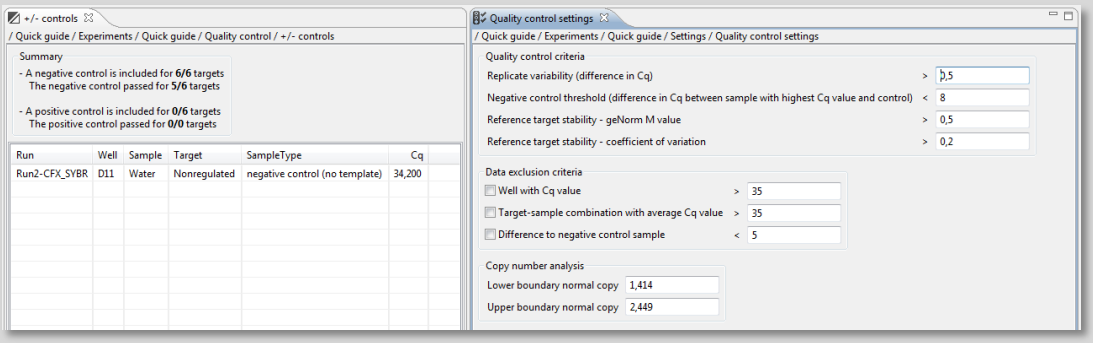
Positive and negative controls

The second type of quality control is evaluation of the positive and negative sample controls. The +/- controls window (Project explorer > Quality Control) shows a summary of the QC results and a list of the samples that failed quality control. qbase^{PLUS} flags suspicious no-template control (NTC) samples based on a user defined threshold, which can be found in the Quality control settings window in the Negative control threshold box. The default value is flagging of non-negative NTC results that are less than 5 cycles away from the sample of interest with the highest Cq value (Figure 4).

In order to get a result for the positive and negative controls, positive and negative control samples should be appropriately labeled in the sample list. If not, no results are shown in the +/- controls window. More information on the annotation of samples can be found in chapter 'Run annotation'.

An amplification signal in the NTC sample indicates a potential contamination issue or formation of primer dimers. Such problems can be ignored as long as the difference in Cq value between the NTC and the sample with the highest Cq value is sufficiently large. For example, a Cq value difference of 5 corresponds to a fold difference of about 32, indicating that approximately 3 % of the signal in your samples may be caused by these unwanted signals (well below the technical error on PCR replicates). Smaller differences between the NTC and the unknown samples should be avoided.

Figure 4 – Positive and negative controls



Stability of reference targets

When using the multiple reference target normalization approach (Project explorer > Settings > Calculation parameters), reference target stability is the third type of quality control. The user is able to choose a minimal acceptable reference target stability by defining a threshold value for two indicators of expression stability of the used reference genes: the geNorm expression stability value of the reference gene (M) and the coefficient of variation of the normalized reference gene relative quantities (CV) [Vandesompele et al., Genome Biology, 2002; Hellemans et al., Genome Biology, 2007]. Both parameters are automatically calculated by qbase^{PLUS} during the analysis. More information on finding the best set of reference genes in a given experimental condition can be found in chapter 'genorm^{PLUS}'.

Step 1 - Selecting reference targets

At least two reference targets are required for the evaluation of their stability. Because qbase^{PLUS} automatically recalculates by default the stability values each time a parameter is changed by the user, quality values are automatically updated as more reference targets are appointed.

Please note that as of version 2.0, there is an option to disable automatic calculations. This option is available in the preferences window.

In order to get a result for the reference target stability, reference targets have to be appointed in the Target list. If no reference targets have been appointed, no values are shown in the Reference target stability window. More information on the annotation of targets can be found in chapter 'Run annotation'.

Step 2 - Reference target quality control settings

The threshold value for M and CV can be set in the Reference target stability boxes (Project explorer > Quality control settings). By default, these thresholds are set at 0.5 for the M value and 0.2 for the CV value.

Figure 5 – Good reference genes

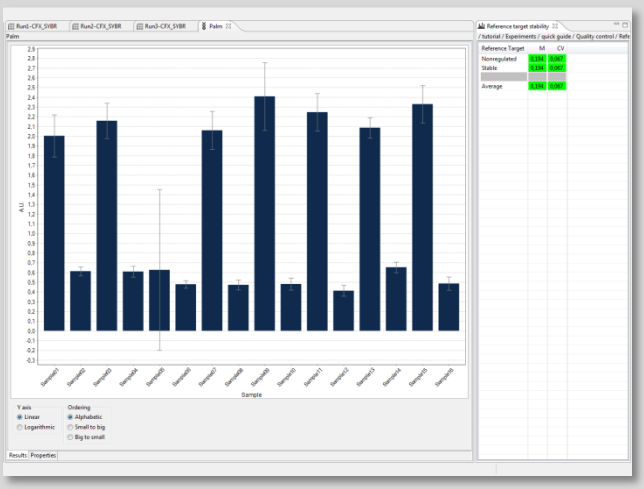
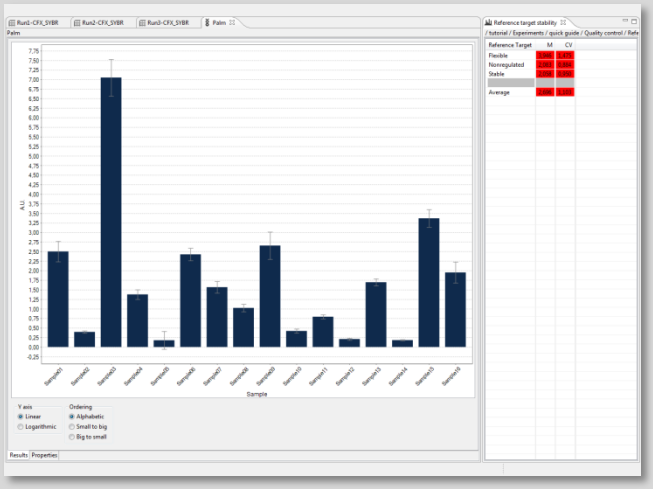


Figure 6 – Low quality reference genes



Step 3 - Inspecting reference target stability

M and CV values passing the QC are highlighted in green (Figure 5) in the Reference target stability window (Project explorer > Quality Control), while failing ones are in red (Figure 6).

Please note that both M and CV values will not be calculated if there are samples that have missing data for one of the reference genes. Either delete the sample(s) or make the reference target a target of interest.

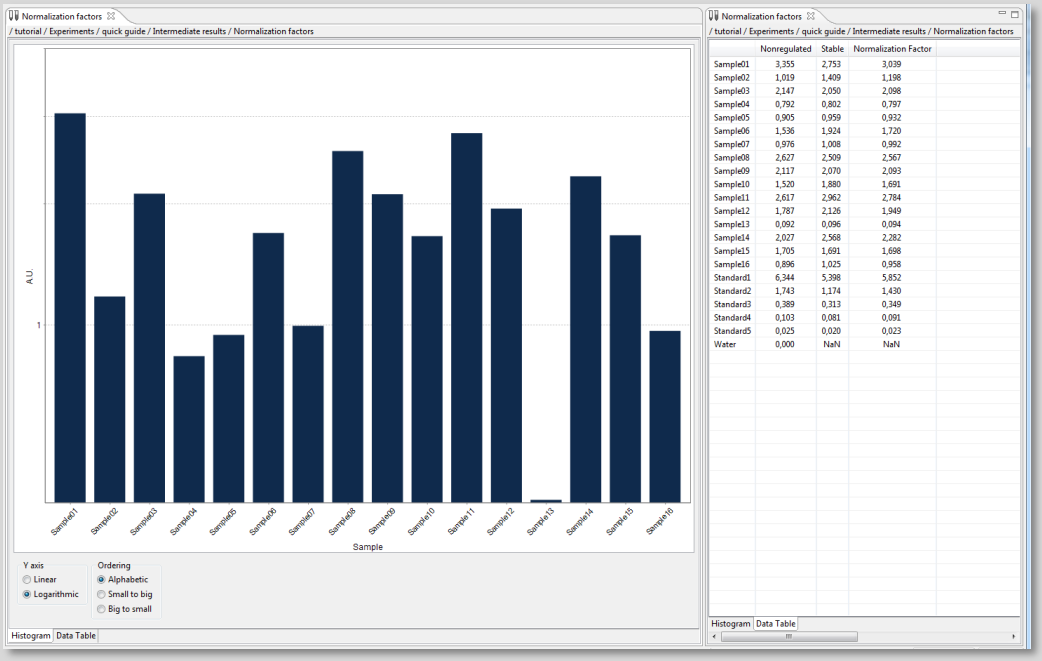
Normalization factors

Inspection of normalization factors allows you to inspect possible experimental problems. qbase^{PLUS} displays the calculated normalization factor (calculation depends on the selected normalization approach) for each sample both in tabular form and in a bar chart. Using approximately equal amounts of equal quality input material and stably expressed reference genes, the normalization factor values should be similar for all samples. High variability of the normalization factors indicates large differences in starting material quantity or quality or it might point out a problem with one of the reference genes (either not stably expressed, or not adequately measured). A variation of 2- to 3-fold is generally acceptable (this is the experimental variation that you want to remove in the normalization process). Any higher variation should be treated with care.

Step 1 - Inspecting the normalization factors window

Go to Intermediate Results in the Project explorer and double click on Normalization factors. By default, the bar chart is shown (Figure 7 – left). Switch to the Data table tab at the bottom of the window to see the normalization factor table (Figure 7 – right). This table can be exported by using the Export normalization factors option in the Normalization factor context menu.

Figure 7 – Normalization factors



Automatic data exclusion

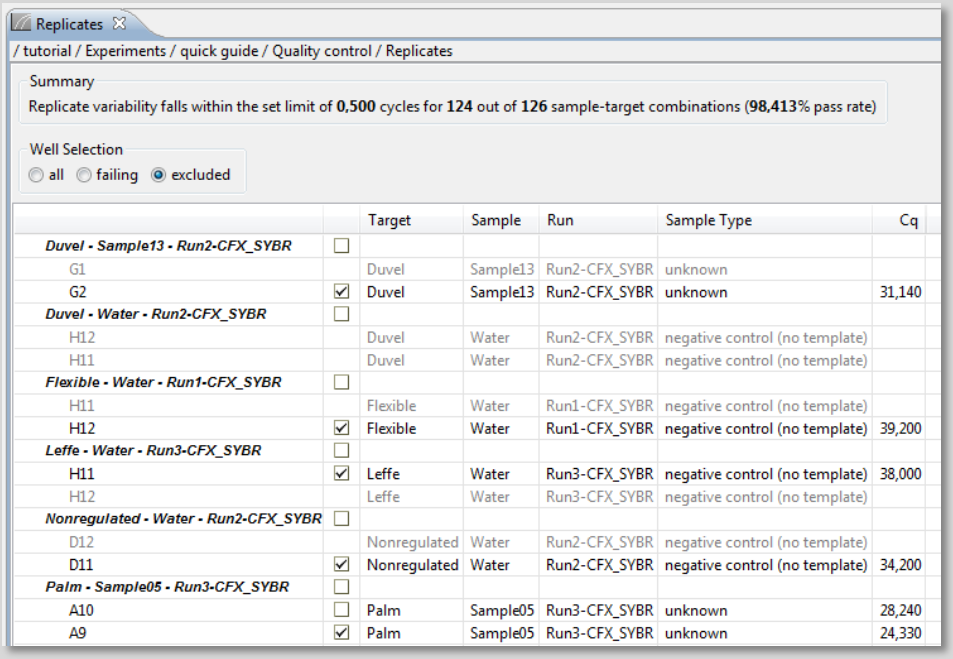
As of version 2.0, it is possible to define auto-exclusion criteria that allow data points to be automatically excluded in three different situations:

- 'Difference to negative control sample < …' – used to automatically exclude data points that could be significantly impacted by the signal found in the negative control
- 'Well with Cq value > …' - used to automatically exclude data points in a Cq range with inaccurate results
- 'Target-sample combination with average Cq value > …' - used to automatically exclude replicates with an average Cq in the range with inaccurate results

The threshold value for auto-exclusion can be set in the Data exclusion criteria box (Project explorer > Quality control settings).

Data points that have been automatically excluded are grayed out in the replicate quality control (Figure 8). These data points, like those that have been manually excluded, will not be used for calculations. In contrast to manually excluded data points they cannot be re-included in the replicate quality control screen or the run editor, i.e. they are strictly linked to the auto exclusion settings.

Figure 8 – Automatically excluded data points



		Target	Sample	Run	Sample Type	Cq
Duvel - Sample13 - Run2-CFX_SYBR	☐					
G1		Duvel	Sample13	Run2-CFX_SYBR	unknown	
G2	☑	Duvel	Sample13	Run2-CFX_SYBR	unknown	31,140
Duvel - Water - Run2-CFX_SYBR	☐					
H12		Duvel	Water	Run2-CFX_SYBR	negative control (no template)	
H11		Duvel	Water	Run2-CFX_SYBR	negative control (no template)	
Flexible - Water - Run1-CFX_SYBR	☐					
H11		Flexible	Water	Run1-CFX_SYBR	negative control (no template)	
H12	☑	Flexible	Water	Run1-CFX_SYBR	negative control (no template)	39,200
Leffe - Water - Run3-CFX_SYBR	☐					
H11	☑	Leffe	Water	Run3-CFX_SYBR	negative control (no template)	38,000
H12		Leffe	Water	Run3-CFX_SYBR	negative control (no template)	
Nonregulated - Water - Run2-CFX_SYBR	☐					
D12		Nonregulated	Water	Run2-CFX_SYBR	negative control (no template)	
D11	☑	Nonregulated	Water	Run2-CFX_SYBR	negative control (no template)	34,200
Palm - Sample05 - Run3-CFX_SYBR	☐					
A10	☐	Palm	Sample05	Run3-CFX_SYBR	unknown	28,240
A9	☑	Palm	Sample05	Run3-CFX_SYBR	unknown	24,330

References

D'haene B and Hellemans J.

The Importance of Quality Control During qPCR Data Analysis. International Drug Discovery August/September 2010

D'haene B, Vandesompele J, Hellemans J.

Accurate and objective copy number profiling using real-time quantitative PCR. Methods. 2010 Apr 50(4):262-70.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F.

Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002 Jun 18;3(7):RESEARCH0034

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J.

qBase: relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. 2007;8(2):R19.