

# whole genome miRNA profiling

qbase<sup>PLUS</sup>



Easy . Fast . Reliable .

## Subject

### Step-by-step

Step 1	Stop automated calculations	p. 2
Step 2	Create a new experiment	p. 2
Step 3	Import run data	p. 3
Step 4	Add custom sample properties (optional)	p. 4
Step 5	Define calculation parameters	p. 5
Step 6	Define quality control setting	p. 6
Step 7	Browse results	p. 7
Step 8	Export results or perform bio-statistical analysis	p. 7

### References

## Subject

This chapter describes how to do whole genome miRNA profiling analysis using the global mean normalization strategies implemented in qbase<sup>PLUS</sup> as of version 2.0 [premium license only].

In 2009, Mestdagh et al. introduced a global mean normalization method to normalize data from RT-qPCR miRNA profiling studies in which a large number of miRNAs are tested per sample (e.g. whole miRNome). This method outperformed other normalization strategies commonly used at the time (e.g. multiple target reference normalization using endogenous small RNA controls) and it has been considered the gold standard method since. Biogazelle has perfected this method and integrated two modified global mean normalization strategies into its qbase<sup>PLUS</sup> software [D'haene et al., 2011].

The modified global mean normalization strategy conveniently generates normalized relative quantities (NRQ values) in a linear scale. These values can be obtained using four simple steps: (1) auto-exclusion of miRNAs below a certain expression level, (2) conversion of Cq values into relative quantities (RQs), (3) calculation of sample specific normalization factor (NF) as the geometric mean of the RQs of the targets, and (4) conversion of RQs into normalized RQs (NRQs) by dividing the RQs with the sample specific NF. Steps 2-4 are simultaneously performed if the user selects one of the two available global mean normalization strategies. The NF can be calculated based on the RQs of all expressed targets in the sample for which the normalization is determined without taking into account the other samples ('modified global mean normalization'). Alternatively, the normalization factor is calculated based on the RQs of the targets that are expressed in all samples ('modified global mean normalization on common targets').

The analysis of a published dataset [Mestdagh et al., 2009] that includes expression profiles for 430 miRNAs and 18 controls in 61 neuroblastoma (NB) tumor samples is used as an example. Two sample groups can be distinguished; samples with MYCN amplification (MNA) in comparison to samples with a normal MYCN copy number (MNSC).

# Step-by-step

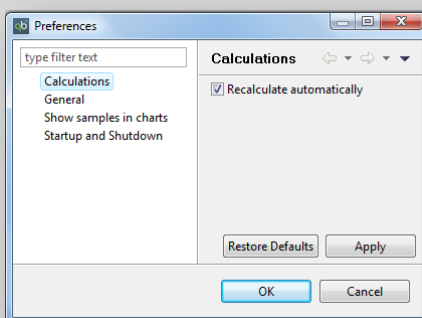
## Step 1 - Stop automatic calculations

By default, qbase<sup>PLUS</sup> recalculates everything immediately (intermediate and final results, quality controls, and specific analyses) as soon as something changes in the data or settings. For large datasets such as whole genome miRNA profiles, it is recommended to stop the automatic recalculations in order to avoid losing several seconds for every minor change propagated in a large experiment. The option to stop automatic recalculations is accessible in the Preferences window (Figure 1). Without automatic calculations, a recalculate button will appear in the qbase<sup>PLUS</sup> toolbar, which needs to be clicked manually to execute the calculations (Figure 2).

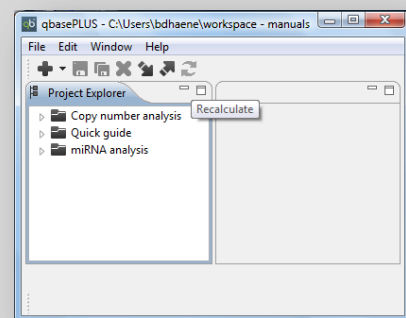
## Step 2 - Create a new experiment

In qbase<sup>PLUS</sup> qPCR run data are organized in experiments. To start data analysis, a new empty experiment needs to be created. Right click on the project (📁) in the qbase<sup>PLUS</sup> project explorer tree in which you want to start a new experiment and click *New experiment*. If needed, create a new project first in a similar way (*New > Project*). An experiment name and description can be provided in the *Experiment properties* window.



▼ Figure 1



▼ Figure 2



### Step 3 - Import run data

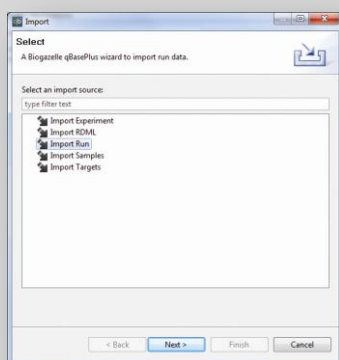
qbase<sup>PLUS</sup> experiments contain data from one or more runs. Run data can be imported by clicking the downward pointing arrow (  ) in the qbase<sup>PLUS</sup> toolbar followed by the selection of the Import Run option (Figure 3). Alternatively, right click on the project explorer Run element (  ) of the experiment in which you want to import the run and select Import Runs ... .

#### Make the appropriate settings and selections in the next window

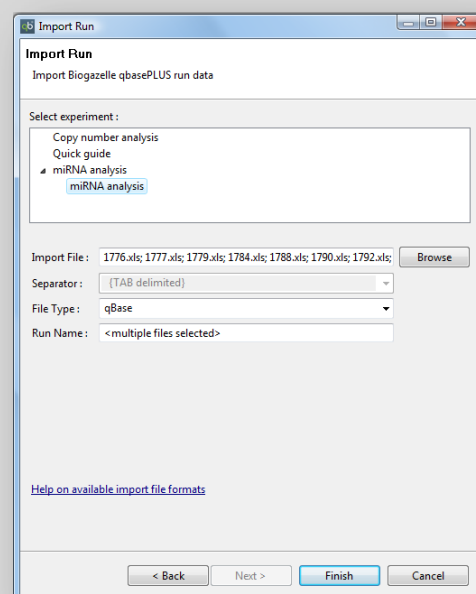
1. Select the experiment in which you want to import the run data (miRNA analysis in this example).
2. Browse for your run files. If multiple runs need to be imported in the selected experiment and all runs have the same file type, i.e. are derived from the same instrument using the same data collection software, you can make use of the batch import function (CTRL + click in Windows, command + click in MacOSX). In this miRNA analysis experiment, we import 61 qBase files (one for each sample) (Figure 4).
3. Select the file type that corresponds to the instrument and data collection software version that has been used for the creation of your run files ('qBase' for this example, one for each sample).

Click Finish to start the run import procedure. If needed, well annotation such as sample and target name can be added or modified in the run editor window that can be opened by double clicking the run name in the Project explorer tree. In this example, a fully annotated run file (containing a target and sample name for each well) is used, which does not require further editing.

▼ Figure 3



▼ Figure 4



## Step 4 - Add custom sample properties (optional)

If multiple sample groups exist within one experiment (e.g. MNA and MNSC), it is possible to define these groups by adding custom sample properties.

There are two ways to enter custom sample properties and its corresponding values in qbase<sup>PLUS</sup>.

- Custom properties can be manually entered in the Sample properties window (Figure 1). To access the Sample properties window, double click on the Annotations icon (🔗) and then on the Sample properties icon (🔗). Custom properties are used for visualization (sample grouping and rescaling) and statistics. To create a custom property click Add property at the bottom and enter a name for the property. For each sample, a property value may be entered for the newly created property.

Manual editing of custom property values is most useful for small corrections or in experiments with a limited number of samples

- Importing sample properties from additional columns in a sample import file (Figure 5) is a faster alternative to the individual creation of new samples or properties. The format of these import files, as well as example files can be found at <http://www.biogazelle.com/support/qbaseplus/formats>.

In addition to the complete overview of sample properties in the Sample properties window, properties for a given sample can be found in the Properties tab of the Sample window.

In this miRNA analysis example, we use a sample import file (Sample properties miRNA analysis.xls). Click the downward pointing arrow (⌵) in the qbase<sup>PLUS</sup> toolbar and choose the Import Samples option. Select the properties to be copied from the sample list (Sample name and type are mandatory, other properties are optional) and click the Finish button to proceed.

More information can be found in the chapter on 'Run annotation' (<http://www.biogazelle.com/support/qBasePlus/manual>).

▼ Figure 5 – Sample import file

	A	B	C	D	E	F	G
1	name	type	description	quantity	normalization factor	positive control quantity	Treatment
2	sample01	UNKN					No
3	sample02	UNKN					Yes
4	sample03	UNKN					No
5	sample04	UNKN					Yes
6	sample05	UNKN					No
7	sample06	UNKN					Yes
8	sample07	UNKN					No
9	sample08	UNKN					Yes
10	sample09	UNKN					No
11	sample10	UNKN					Yes
12	sample11	UNKN					No
13	sample12	UNKN					Yes
14	sample13	UNKN					No
15	sample14	UNKN					Yes
16	sample15	UNKN					No
17	sample16	UNKN					Yes
18	water	NTC					
19	standard1	STANDARD					
20	standard2	STANDARD					
21	standard3	STANDARD					
22	standard4	STANDARD					
23	standard5	STANDARD					

## Step 5 - Define calculation parameters

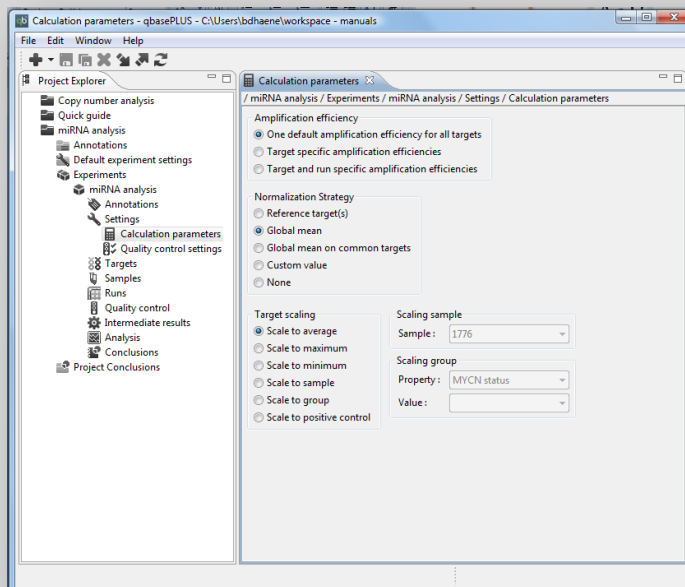
The parameters for calculations can be defined in the Calculation parameters window (Project Explorer > Settings > Calculation parameters). This is the place to select the appropriate normalization strategy and target scaling option (Figure 6).

Please note that the multiple reference target normalization strategy is used by default. Users with a premium license have the option to choose different normalization procedures to suit their specific needs in different types of experiments.


The Global mean or the Global mean on common targets normalization strategy are useful for any experiment in which a sufficiently large set of unbiased genes/targets is quantified.

- *Global mean* [premium license only]  
The global mean normalization procedure was initially developed for the normalization of extended miRNA screening experiments where it was shown to be a superior alternative to the commonly used small nuclear and small nucleolar RNAs as reference targets [Mestdagh et al., 2009]. It is however useful for any experiment in which a sufficiently large set of unbiased genes is quantified. It is based on the same principles commonly used for microarray normalization. The results for a given sample will be normalized by the geometric mean of the relative quantities of all the targets that are expressed in that sample. This normalization approach is best combined with the option to exclude too high Cq values (see auto exclusion).
- *Global mean on common targets* [premium license only]  
Similar to the 'Global mean' normalization procedure with the exception that the normalization factor will only be based on the targets that are expressed in all samples.

▼ Figure 6 - Calculation parameters window



## Step 6 - Define quality control parameters

The parameters for quality control can be defined in the Quality control setting window (Project Explorer > Settings > Quality control settings ).

The auto exclusion criteria allow data points to be automatically excluded in three different situations (Figure 7):

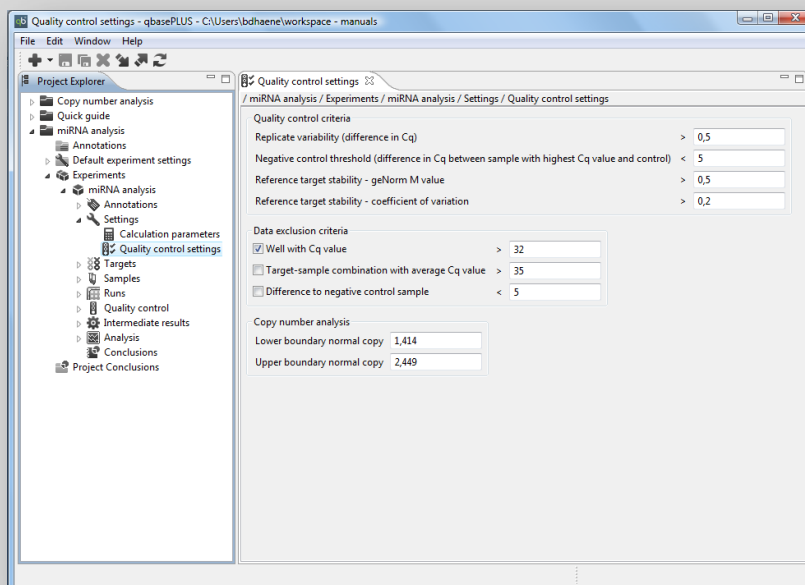
- 'Difference to negative control sample < ...'  
used to automatically exclude data points that could be significantly impacted by the signal found in the negative control
- 'Well with Cq value > ...'  
used to automatically exclude data points in a Cq range with inaccurate results
- 'Target-sample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

The latter two options are particularly relevant when analyzing a miRNA screening with the global mean normalization strategy. In this miRNA analysis example, all Cq values above 32 are considered noise and are excluded from further analysis.

Data points that have been automatically excluded are grayed out in the replicate quality control. 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 reincluded in the replicate quality control screen or the run editor, ie they are strictly linked to the auto exclusion settings.

More information can be found in the manual chapter on quality control (<http://www.biogazelle.com/support/qBasePlus/manual>).

▼ Figure 7 - Quality control parameters window



## Step 7 - Browse results

Click the Recalculate button in the qbase<sup>PLUS</sup> toolbar to initiate calculations.

Double click a target to visualize the fold changes between samples in the Target bar chart (Project Explorer > Targets).

If custom properties have been entered (see Step 4), it is possible to group results in the Target bar chart. This functionality facilitates interpretation of results when multiple groups are to be compared (Figure 8).

Open the Property 1 drop down menu in the Grouping box at the bottom of the target bar chart and select the custom property. This action will result in a grouping of the results according to the selected custom property (Figure 8).

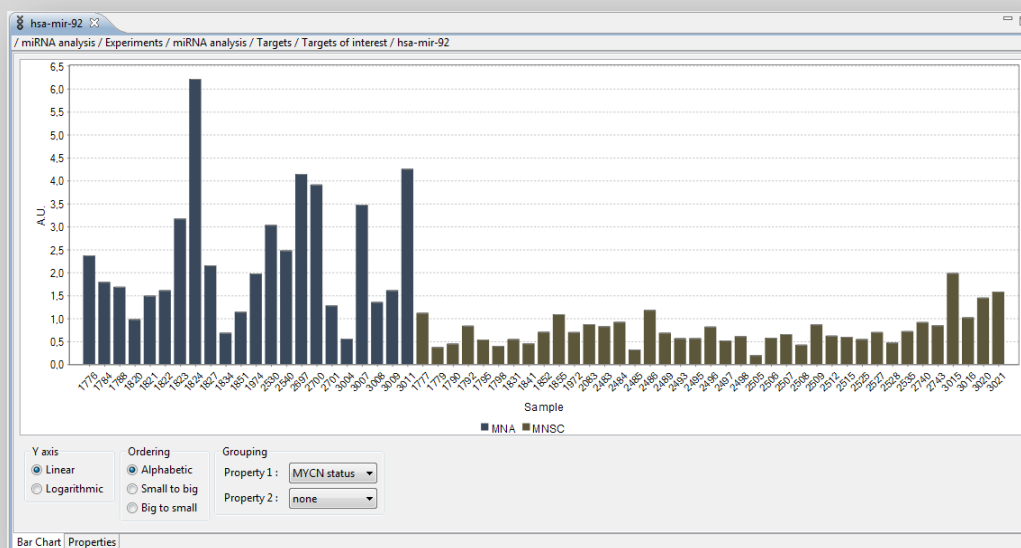
In this miRNA analysis experiment, we select 'MYCN status' as Property 1. This action results in a grouping of the MNA samples at the left-hand side and the MNSC samples at the right-hand side.

## Step 8 - Export results or proceed with statistical analysis

The results can be exported for reporting or further processing using different formats. Individual charts can be printed and exported as figures, or the entire result set can be exported in tabular format by clicking the upward pointing arrow (↕) in the qbase<sup>PLUS</sup> toolbar and selecting Export result table. You will be given the choice to export results (normalized relative quantities) only or to include the errors (standard error of the mean) as well.

As of version 2.0, qbase<sup>PLUS</sup> has been enhanced with tools to perform commonly used statistical tests on the results that have been generated [premium license only]. This new module is tailored towards the typical needs of biologists performing qPCR analysis rather than to needs and habits of statisticians. This philosophy is reflected in a wizard based approach that minimizes the use of statistical lingo and provides the needed information to guide the user to appropriate statistical test and settings. More information on this topic can be found in the statistics manual.

▼ Figure 8 - Target bar chart after sample grouping



## References

Mestdagh P, Van Vlierberghe P, De Weer A, Muth D, Westermann F, Speleman F, Vandesompele J

A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biol.* 2009 16;10(6):R64.

D'haene B, Mestdagh P, Hellemans J, Vandesompele J

miRNA expression profiling - from reference genes to global mean normalization. Second edition of *MicroRNA Expression Profiling: Methods and Protocols*. In press.

