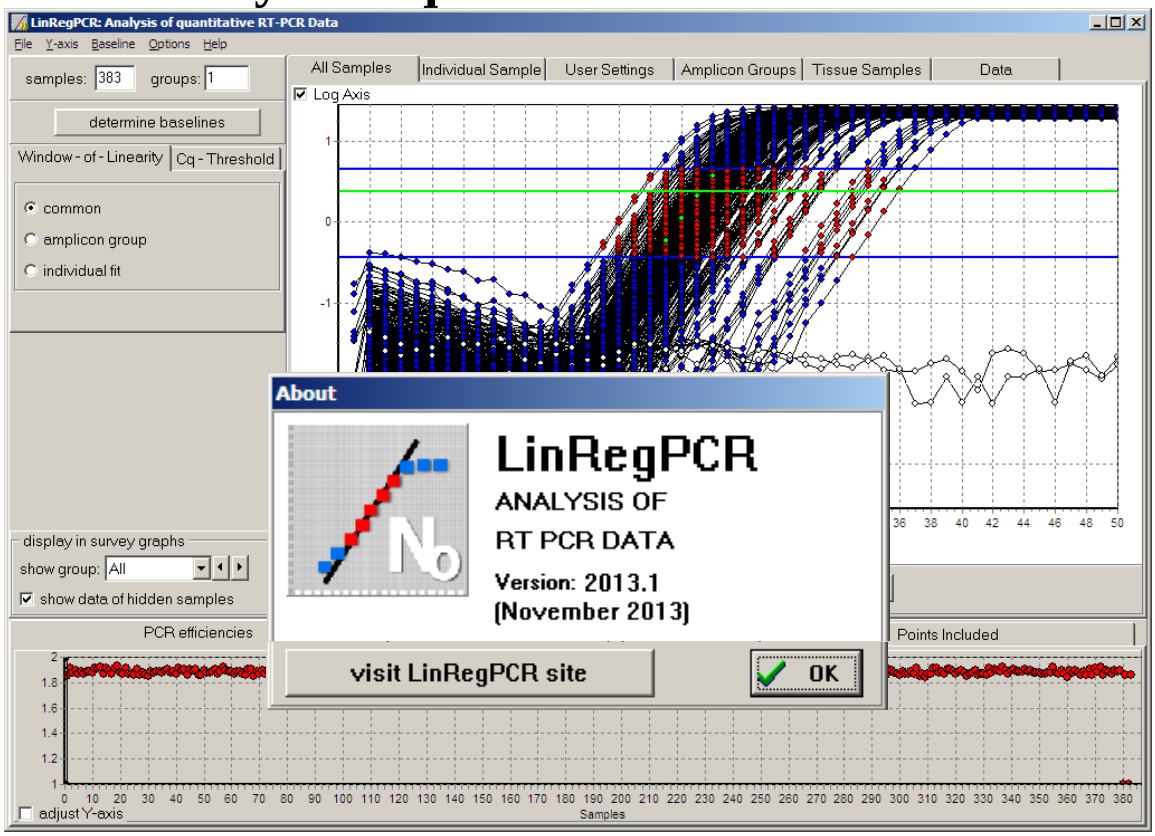
LinRegPCR (2014.x)

Analysis of quantitative RT-PCR data



J. M. Ruijter
A. Ilgun
Q.D. Gunst

Department of Anatomy, Embryology & Physiology
Academic Medical Center
Amsterdam, the Netherlands

Version: January 2014

Preferred way of referring to LinRegPCR

When you refer to LinRegPCR in your paper, we prefer that you use the reference **Ruijter et al. (2009)** Nucleic Acids Research 37: e45. When you are using hydrolysis probes, you can also use the reference **Tuomi et al. (2010)** Methods 50: 313–322. The handling of other chemistries is described in **Ruijter et al. (2014)** Microchimica Acta). A comparison of the different amplification curve analysis approaches can be found in **Ruijter et al. (2013)** Methods 59: 32-46.

Please do not use the abbreviation LinReg when you refer to LinRegPCR.

Known bugs

Opening a second data set after analyzing the first sometimes gives an error, especially when the second set is smaller than the first. Therefore, it is better to close the program and start again when you have more datasets to analyze. Note that you can analyze different datasets in parallel by starting several instances of LinRegPCR simultaneously.

Acknowledgements

We wish to thank Prof. Antoon Moorman, Prof. Vincent Christoffels, Dr Maurice van den Hoff (Department of Anatomy, Embryology & Physiology, AMC, Amsterdam) and Dr. Onno Bakker (Department Endocrinology and Metabolism, AMC, Amsterdam) for their advice and support. All qPCR users in the Heart Failure Research Center and the department of Neurogenetics (AMC, Amsterdam) are acknowledged for their patience, critical comments and suggestions when using the different beta-versions of this program. Darren Scicluna designed the LinRegPCR logo.

Contacts

Frequently asked questions: http://LinRegPCR.nl
Please direct questions and remarks to LinRegPCR@amc.uva.nl . This mailbox will be checked once per week.

CONTENTS

Preferred way of referring to LinRegPCR 2
Known bugs 2
Acknowledgements 2
Contacts 2
1. About LinRegPCR. 5
1.1 Disclaimer 5
1.2 Update history of LinRegPCR 5
1.3 Frequently asked questions 5
1.4 Installation and requirements of LinRegPCR 5
1.5 Note for system managers 6
1.6 LinRegPCR references 6
2. Introduction to LinRegPCR 7
2.1 Definitions of terms used in this manual and in the LinRegPCR program 7
2.2 Theory behind the LinRegPCR program. 8
2.3. Interface components 9
2.3.1 Menu and Submenus 9
2.3.2 Tab pages 11
2.3.3 Survey graphs 13
2.4 Analysis principle and basic procedure 14
3. PCR data analysis with LinRegPCR 15
3.1 Exporting raw data from the PCR apparatus 15
3.1.1. Opening a text file in Excel 15
3.2 Reading data into LinRegPCR 15
3.2.1 Read data from Excel 15
3.2.2 Read data from RDML 17
3.3 Baseline Estimation and flagging of deviating samples 20
3.3.1 Relaxing baseline estimation criteria 21
3.3.2 Manual setting of the fluorescence baseline 22
3.4 Amplicon groups 23
3.4.1 Groups Statistics 25
3.5 Check individual samples 26
Survey graphs 26
3.6 Mean efficiency 28
3.7 Window-of-Linearity and Nq threshold 28
3.7.1 Automatic setting of W-o-L and N_q threshold 29
3.7.2 Manual setting of the Window-of-Linearity 30
3.7.3 Manual setting of the N_q threshold 31
3.8 Handling noisy datasets 31
3.8.1 Avoid noisy samples in the lab 31
3.8.2 Rescue noisy samples during data analysis 32
3.9 Probes-based monitoring chemistries and cumulative fluorescence 32
3.10 User Settings 34
3.11 Tissue Samples 35
3.12 Save (data and) results 36
3.12.1 Save to Excel 36
Save options 36
3.12.2 Save to RDML 38
3.13 Quality control output 39
Appendix. Implemented formats for Excel 41

1. About LinRegPCR.

The first two chapters of this manual contain the information on the installation of LinRegPCR and the basic principles of the program. Chapter 2 also describes the menus and different interface pages. Detailed descriptions of the interactive procedures of LinRegPCR are given in the chapter 3. The LinRegPCR web page (http://LinRegPCR.nl) gives answers to some frequently asked questions.

1.1 Disclaimer

LinRegPCR is based on the procedures described in the paper by Ruijter et al., Nucleic Acids Research, 2009. The procedures in this program can be used for fluorescence data resulting from monitoring the PCR reaction. You should preferably use raw fluorescence data and thus you should check which data are exported by your PCR system. Negative values should not occur. The fact that LinRegPCR can read the data should not be seen as proof that you have the correct data.
By using this software you acknowledge that you have read the above paper, understand it, and agree with its conclusions. Therefore, you assume all responsibility and liability for the selection of this program to achieve your intended goals, and for the conclusions you draw from the results. The authors cannot be held responsible for any consequences of the use of this program.

1.2 Update history of LinRegPCR

The recent update history of LinRegPCR can be found at: http:/LinRegPCR.nl . The complete update history is in the UpdateHistory.pdf in the LinRegPCR.zip. Please register yourself as LinRegPCR user on the website. When you do, you enable us to send you information on important updates and bug fixes.

1.3 Frequently asked questions

Please regularly use the button on the **About** box and in the **Help** of the program to access the LinRegPCR website and look at the FAQ's (http://LinRegPCR.nl).

1.4 Installation and requirements of LinRegPCR

The version of LinRegPCR that you received has been zipped. In the LinRegPCR.zip file you will find:

- LinRegPCR.exe: the executable program.
- LinRegPCR_help.pdf: this help manual in Adobe Acrobat readable format.
- UpdateHistory.pdf.

Unzip the files into a directory and you are ready to run. No special configuration is needed. For easy use it is handy to put shortcuts to the program and the help file in your **Start** menu or on your Desktop.

LinRegPCR stores some of the choices you make during the use of the program in the Windows Registry (Key: HKEY_CURRENT_USER\Software\AEL Research\LinRegPCR). This enables the program to select these choices as default the next time you use the program.

The program requires Windows 95 or later and Microsoft Excel 97 or later. When your input data are in Excel, LinRegPCR program expects Excel to use the decimal separator that is defined in Windows. To make sure of this, in Excel go to ***Tools - Options*** and in the ***Options*** dialog, on the ***International*** Tab make sure that ***Use system separator*** is checked. An Excel version that expects decimal points will read numbers with decimal commas like 3,9; 3,87; 3,868 as 39; 387; 3686, respectively. Thus creating a 100 fold difference where there was only difference in number of decimal places.

1.5 Note for system managers

LinRegPCR writes the user choices and its version number to the Windows Registry (Key: HKEY_CURRENT_USER\Software\AEL Research\LinRegPCR). At start-up LinRegPCR compares its version number with the one stored in the Registry. When they differ, the user will be notified that you installed a new version. This way, the users are automatically notified when you have installed a new version of the program on an application server. The opening of the user manual from the program requires the LinRegPCR_help.pdf to be present in the program's directory.

1.6 LinRegPCR references

Ruijter JM, Lorenz P, Tuomi JM, Hecker M, van den Hoff MJB (2014) Fluorescent increase kinetics of different fluorescent reporters used for qPCR depends on monitoring chemistry, targeted sequence, type of DNA input and PCR efficiency. Microchimica Acta. *This paper is an extension of Tuomi et al. (2010) and describes the processing of qPCR data resulting from 6 different categories of monitoring chemistries.*

Ruijter JM, Pfaffl MW, Zhao S, Spiess AN, Boggy G, Blom J, Rutledge RG, Sisti D, Lievens A, De Preter K, Derveaux S, Hellemans J, Vandesompele J. (2013) Evaluation of qPCR curve analysis methods for reliable biomarker discovery: bias, resolution, precision, and implications. Methods 59: 32-46. *This paper describes the comparison of the performance, in terms of precision, linearity, and sensitivity of all available curve qPCR analysis methods using the same biomarker data sets.*

Tuomi JM, Voorbraak F, Jones DL, Ruijter JM. (2010) Bias in the Cq value observed with hydrolysis probe based quantitative PCR can be corrected with the estimated PCR efficiency value. Methods 50, 313–322.
This paper describes the use of LinRegPCR in the analysis of cumulative fluorescence datasets obtained with hydrolysis probes like Taqman® probes.

Ruijter JM, Ramakers C, Hoogaars W, Bakker O, van den Hoff MJB, Karlen Y, Moorman AFM. (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research 37: e45.
This paper describes the basics of the LinRegPCR program and gives details on the fluorescence baseline correction and setting of the window-of-linearity.

Ramakers C, Ruijter JM, Lekanne Deprez RH, Moorman AFM. (2003) Assumption-free analysis of quantitative real-time PCR data. Neurosci Letters 339: 62-66.
This paper describes the estimation of PCR efficiencies for individual samples from a subset of data points in the log-linear phase of the amplification curves. Such handling of individual samples is no longer recommended (JMR 2013).

The linear regression procedure used in this program can be found in every statistics textbook. **Windows** and **Excel** are trademarks of Microsoft Inc. The Real-time PCR Data Markup Language, **RDML**, is developed and maintained by the international RDML consortium (http://www.rdml.org) to provide the qPCR community with a common data exchange format.

2. Introduction to LinRegPCR

LinRegPCR is a program for the analysis of real time RT-PCR Data, also called quantitative PCR (qPCR) data. The program imports non-baseline corrected data, performs a baseline correction on each sample separately, determines a window-of-linearity and then uses linear regression analysis to determine the PCR efficiency per sample from the slope of the regression line. The mean PCR efficiency per amplicon and the Cq value per sample are used to calculate a starting concentration per sample, expressed in arbitrary fluorescence units. Data input and output are from/to an Excel spreadsheet or from/to an RDML file.

2.1 Definitions of terms used in this manual and in the LinRegPCR program

Amplicon group: A set of samples in which the same pair of primers is used to amplify the DNA-of-interest (=amplicon).

Baseline: Measured fluorescence when no amplification-specific fluorescence can yet be determined. This fluorescence includes fluorescence from cDNA, primers and unbound SYBR Green or inefficiently quenched fluorescently labeled probe.

Cq or **quantification cycle**: Fractional number of cycles needed to reach the fluorescence threshold. Samples with higher starting concentrations will reach this threshold earlier and will have a low Cq value; Cq is used to calculate N0. Cq is also known as Cp or Ct but the use of those terms is discouraged by the MIQE guidelines (Bustin et al., Clinical Chemistry 2009)

Cumulative fluorescence: Observed fluorescence data obtained when the PCR amplification is monitored with a monitoring systems for which fluorescence accumulates. These reporters become fluorescent during annealing or elongation and the fluorescent accumulates during subsequent cycles.

E or PCR Efficiency: Amplification efficiency expressed as a value between 1 and 2 and is thus defined as “the fold increase per cycle”. This efficiency is calculated from the slope of the amplification curve in the exponential phase. Ideally the PCR efficiency is 2, meaning that in each cycle the amount of amplicon doubles. This is also referred to as a 100% efficient PCR. An efficiency value of 1 means no amplification.

Exponential or **log-linear phase:** The portion of the PCR curve in which the level of generated fluorescence exceeds baseline fluorescence, but reagents have not yet begun to be limiting. In this phase the amplification efficiency is constant and similar across all samples in an amplicon group regardless of their starting concentration.

Fluorescence threshold, quantification threshold or **Nq**: Fixed amount of fluorescence, set to determine the Cq value. By default set at one cycle below the upper limit of the window-of-linearity.

N0: Target quantity or starting concentration per sample, expressed in arbitrary fluorescence units. The N0 value of a target and one or more reference genes can be used to calculate the efficiency corrected gene expression ratio.

Non-cumulative fluorescence: Observed fluorescence data obtained when the PCR amplification is monitored with a DNA binding dye or hybridization probe. The reporter is only fluorescent when bound to dsDNA; fluorescence intensity is proportional to the amount of DNA present at each cycle.

Nq: see **Fluorescence threshold**

Plateau phase: Endpoint of the PCR reaction - the phase in which reaction components become limiting. In the plateau phase the amplification curves of the real-time PCR are no longer exponential and the PCR efficiency drops to 1.

Raw data: Fluorescence data that is not baseline-corrected by the PCR system. Raw data should not contain negative values.

RDML: Real-time PCR Data Markup Language that can be used to exchange data between qPCR systems and data analysis software. An RDML file is a zipped rdml_data.xml file.

Sample: Reaction unit (well in a PCR plate)

Tissue: Biological material in which DNA or RNA levels have to be determined. A tissue can be measured in different wells to measure different genes.

Window-of-Linearity or **W-o-L**: Subset of data points in the exponential phase that is used to determine the PCR amplification efficiency (E) per sample. The Window-of-Linearity is set automatically by the program but can be modified by the user. Default the W-o-L width is set to include 4 amplification cycles.

2.2 Theory behind the LinRegPCR program.

Basic principle

The basic equation for PCR kinetics is: $N_C = N_0 \times E^C$ [Eq. 1]

Where the amount of amplicon after *C* cycles (N_c) is the starting concentration of the amplicon (N_0) times the PCR efficiency (*E*) to the power *C*. The PCR efficiency in this equation is defined as fold increase per cycle. It is therefore expressed as a value between 1 and 2 where an efficiency of 2 represents a perfect doubling of the amplicon in each cycle This is also referred to as a 100% efficient PCR.

$$N_c = N_0 E^c \quad \longrightarrow \quad log(N_c) = log(N_0) + c.log(E)$$

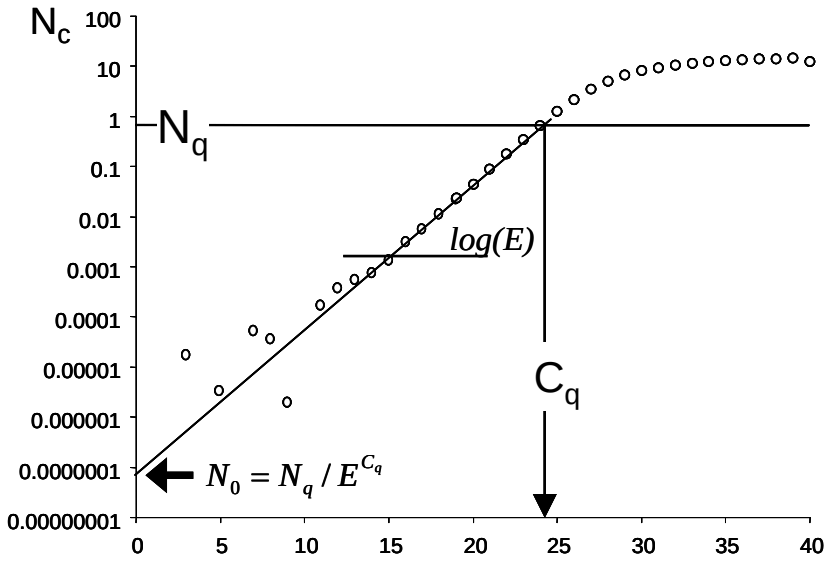


Figure 1: Basic principle of qPCR data analysis. The logarithmic version of the basic kinetics equation (see top row) shows that the data points in the exponential phase, when plotted on a log(N_c) axis, are on a straight line with slope log(E). Calculating the target quantity with Eq. 2 is mathematically equivalent to extending the regression line to its intercept with the log(N_c) axis

LinRegPCR is based on the fact that, when plotted on a log-fluorescence axis, the slope of the linear regression line through the data points in the exponential phase of the amplification curve can be used to estimate the PCR efficiency (*E*) (Ramakers et al 2003).

After setting the quantification threshold Nq and determining the fractional cycle number that was needed to reach this threshold (Fig. 1), the starting concentration or target quantity can be calculated with the inverse of Eq. 1:

$$N_0 = N_q / E^{C_q} \quad \text{[Eq. 2]}$$

The Efficiency value used in Eq. 2 is the mean of the observed efficiencies per amplicon (Ruijter et al 2009).

Different monitoring chemistries

Because their fluorescence is linearly related to the amount of the amplicon, Eq. 1 describes the increase in fluorescence for PCR runs monitored with a DNA binding dye. However, some monitoring chemistries, the fluorescent reporter is released from quenching during elongation and remains fluorescent in subsequent cycles; fluorescence thus accumulates. For other chemistries, the first release of fluorescence occurs only after a number of initial lag-cycles. Some chemistries combine lag with accumulation. The different kinetics lead to a shift in the Cq values of these curves and correction for the observed, earlier or later, Cq values is implemented (Tuomi et al 2010; Ruijter et al., 2014). The required correction depends on the input DNA (ss cDNA or ds DNA) and the chemistry; the targeted sequence in case of fluorescent probes (mRNA or cDNA strand) does not displace the amplification curve observably.

Plotted on a logarithmic fluorescence axis every chemistry shows the amplification curve for each reaction as a straight line, with a slope equal to log(E). This is true for all probe-based chemistries used for monitoring DNA increase in a real-time PCR run.

2.3. Interface components

2.3.1 Menu and Submenus

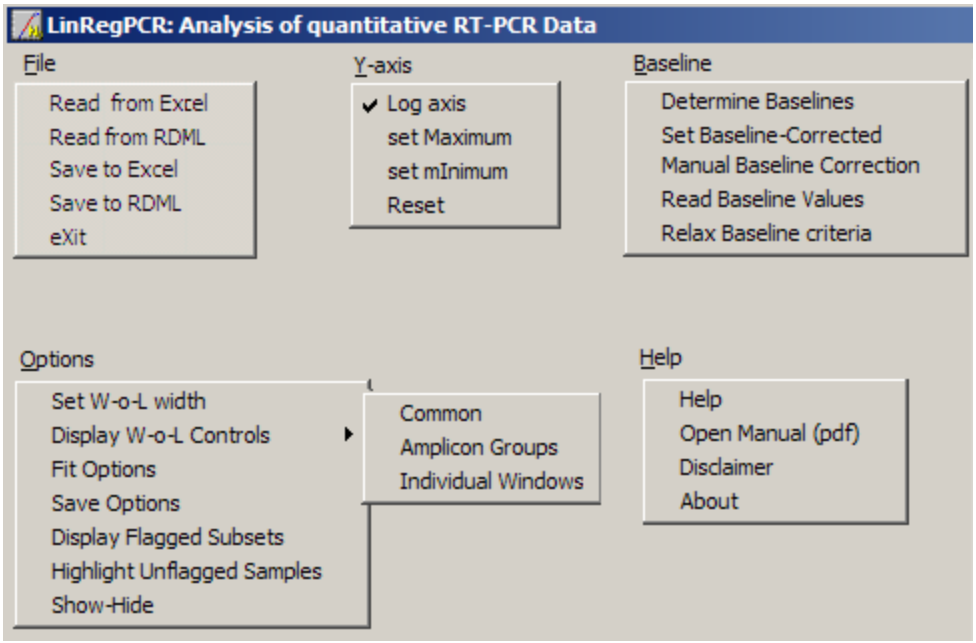


Figure 2: Main menu and sub-menu choices.

File Menu

Read from Excel:

Opens the dialog to read data from Excel (3.2.1)

Read from RDML:

Opens the dialog to read data from an RDML file (3.2.2)

Save to Excel:

Use this option to save results to Excel. Excel must be running and be in 'Ready' status (3.12.1)

Save to RDML:

Use this option to save fluorescence data and results to a new or existing RDML file (3.12.2)

Exit:

Closes the program without saving any results or choices.

Y-axis menu

Log axis:

You can use this option to toggle between a logarithmic Y-axis (default and recommended) and a normal Y-axis. The check boxes on the graphs have the same effect.

Set Maximum / Set Minimum:

These choices will open a dialog window in which you can enter the Y-axis maximum or minimum. Depending on the status of **Log axis** the input has to be a log-value or a normal value.

Reset:

This choice resets the Y-axis to the default values that were determined automatically after the **Read from Excel** and again after the baseline correction.

Baseline Menu

Determine baselines:

Has the same effect as pressing the ***Determine Baseline*** button (3.3).

Set Baseline Corrected:

This option enables you to skip the baseline subtraction procedure when you had forgotten to indicate in the Read-from-Excel dialog that the PCR apparatus had already done a baseline subtraction.

Manual Baseline:

This option opens the manual baseline controls that allow you to manually set or correct the baseline (3.3.2). The aim of a manual baseline setting is to find a baseline value that puts the data points in the exponential phase on a straight line.

Read Baseline Values

In case you reanalyze a dataset, you can read the saved baseline values from the output of the earlier analysis (3.10).

Relax Baseline Criteria

In case of a noisy dataset or low plateau-to-baseline distance you can relax the continuously increasing data points requirement of the baseline estimation procedure (3.3.1).

Options menu

Set W-o-L width:

By default 4 data points are included in the window-of-linearity. The W-o-L limits are indicated by blue horizontal lines in the graphs. You can also change the number of included points on the Fit / Save Options Tab page. Using fewer points is not recommended because it increases the variation in PCR efficiency values.

Display W-o-L Controls:

In the submenu of this option you can choose to open the controls that enable you to manually set the upper and lower limit of the window-of-linearity. You can choose to set the common, amplicon group or individual windows (3.7.2). The aim should be to set a window with minimal variation in efficiency values.

Fit Options:

Choosing this option opens the Fit / Save Options page. On this page you can set the fit options for the individual window-of-linearity. Note that using the

individual windows is not recommended. This option is maintained for backward compatibility.

Save Options:

Choosing this option opens the Fit / Save Options page. On this page you can choose the output columns for the User-Defined output. These choices will be saved to the Windows Registry and be automatically used the next time you use the program (3.12.1).

Display Flagged Subsets:

This option opens a list box (on the middle left of the screen) from which you can choose to display the samples that the program has flagged as: **no-amplification**, **baseline error**, **no-plateau, noisy sample,** or **efficiency outlier**. The first three categories are assigned by the baseline estimation procedure, the "efficiency outliers" are determined while saving the results to Excel or by pressing the button on the **User Settings** page (3.10).

Highlight Unflagged Samples:

Choosing this option displays the unflagged samples (see above) in green and the other samples in red in the survey graph at the bottom of the screen.

Show-Hide:

Opens the Show-Hide page on which you can manually choose which samples to display in the "All Sample" graph. This choice does not affect the calculations or the output and is overruled when you choose to display separate amplicon groups or flagged samples.

Help menu

Help:

Opens the **Help** window.

Open Manual (pdf)

This menu option opens the LinRegPCR_help.pdf file. This file should be present in the directory where the program is stored.

Help on Noisy data sets:

This menu item will only appear when a noisy sample is encountered (3.8)

Disclaimer:

Opens the **Disclaimer** window. This window is also displayed at first use of the program.

About:

Opens the **About** box with the version information and a link to the LinRegPCR web page.

2.3.2 Tab pages

Figure 3: Tab pages of the LinRegPCR program. The **Data** page is opened to display its sub-pages. The **Fit / Save options** page and the **Show / Hide** page are not shown. These pages can be displayed from the **Options** menu.

All Samples

This page shows the amplification curves of all imported samples (see e.g. Fig. 7). The blue lines indicate the upper and lower limit of the W-o-L; the green line is the Nq threshold. Points in the W-o-L are shown in red. When you have defined amplicon groups, which have

different W-o-L and Nq you do not see those lines. The data points in the W-o-L of the sample shown at the ***Individual Sample*** page are displayed in green.

Individual Sample
This page shows the amplification curve of the currently selected sample (see e.g. Fig 13). Lines and colors are as in the ***All Samples*** graph. When the sample is fitted to a common W-o-L the data points in the W-o-L are shown in red, when the W-o-L is set per amplicon group the data points are orange. Yellow data points indicate an individual W-o-L.

User Settings
On this page (3.10; Fig. 23) there are 5 groups of controls:

- Top-left: Setting of the criteria for exclusion of samples from the calculation of the mean PCR efficiency.
- Bottom-left: read baseline values
- Bottom-middle: pairs of radio buttons to enable you to choose what error and bias to accept when the observed Cq value of a sample is too low (see section 3.8, Fig. 9).
- Top-right: radio buttons with which you can relax the criteria that are used for baseline estimation.

The choices in the middle of the page only appear when you are using a cumulative chemistry and input of ss-cDNA.

Amplicon Groups
On this page you can define the amplicon groups (3.4) and view group statistics (3.4.1). The setting of the criteria for exclusion of samples from the calculation of the mean PCR efficiency is done on the User Setting page (3.10).

Tissue Samples
On this page (3.11; Fig. 24) you can define the tissues samples. Doing this is handy for statistical analysis of the results saved to Excel and required for output to RDML.

Data
This page contains 4 sub-pages:

- **input data**
  This Tab page shows the imported data. The only function of this page is to check that all data have been imported.
- **valid data**
  This Tab page shows the baseline corrected data with the negative (below baseline) data removed. You can save these data to Excel. On a later date you can import them again into LinRegPCR without going through the baseline correction again. This is especially handy when you had to do a lot of manual baseline setting. Import them as ABI or Lightcycler480 format and check **Yes** at baseline corrected on the ***Read from Excel*** dialog. Note: baseline values can also be read by using the menu option ***Baseline - Read Baselines*** and the controls on the **User Settings** page
- **sample data**
  This page shows the data of the current sample, which is the sample displayed in the **Individual Sample** graph. The data points labeled ‘Y’ are the ones included in the window-of-linearity and used for the calculation of the individual PCR efficiency.
- **RDML input data**
  This page shows the imported data when data are read from an RDML file (3.2.2; Fig. 10)

Fit and Save Options

- The choices in the top part determine the rules for fitting individual samples. These are no longer used when you set a common W-o-L or a W-o-L per amplicon group. When you choose to fit individual W-o-Ls these fit options are used. This part of the page is just there for backward compatibility. Using individual W-o-Ls is not recommended
- The bottom part of this page contains the checkboxes for selecting of the columns you want to see in the output of the analysis results to Excel (3.12.1).

Show or hide

With the list boxes on this page you can select which samples are shown in the graph on the **All Samples** page and in the **Survey graphs** (2.3.3). This can be handy when you want to show amplification curves of a selection of your data in a presentation. The choices made on this page are overruled when you browse through amplicon groups and when you select marked subsets for display.

2.3.3 Survey graphs

At the bottom of the LinRegPCR screen there is a notebook (Fig. 4) containing three Tab pages displaying a survey of all PCR efficiencies, correlation coefficients and number of included points, respectively. Inspect these Tab pages as part of the quality check before saving results to RDML or Excel (3.5).

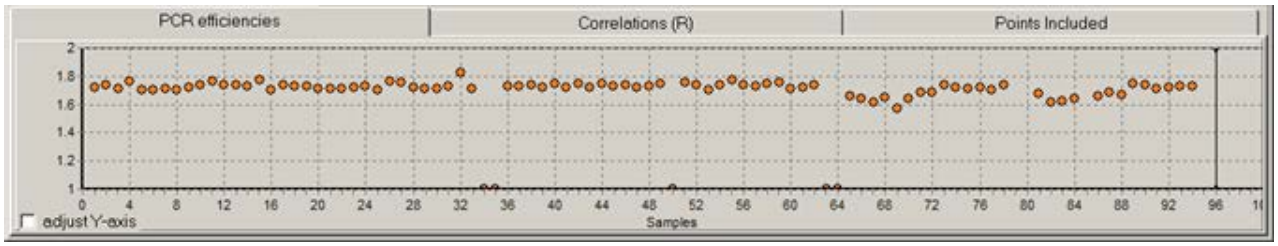


Figure 4: Notebook with Tab pages to display the survey graphs of PCR efficiencies, correlation coefficients and number of included points.

The data points in the survey graphs are red when you are using a common window, orange when you have defined amplicon groups and yellow when a sample is fitted to an individual window.

Exporting graphs

When you click the right mouse button on a graph, the graph is copied to the Windows Clipboard as a bitmap. This bitmap can be pasted into your presentation.

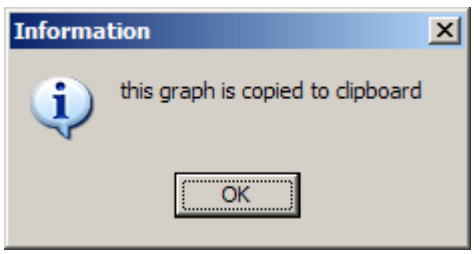


Figure 5: Notification that the graph has been copied to the Windows Clipboard.

2.4 Analysis principle and basic procedure

As described in 2.2 the LinRegPCR program is used for the calculation of the starting concentration (N_0) using the following parameters:
- N_q : the fluorescence threshold,
- C_q : the fractional cycle number that was needed to reach this threshold
- E : the mean PCR Efficiency per amplicon

Therefore, it is essential that C_q and E are determined as accurate as possible. To this end LinRegPCR imports non-baseline corrected data, performs a baseline correction on each sample separately, determines a window-of-linearity per amplicon and then uses linear regression analysis to fit a straight line to the selected PCR data per sample. Subsequently an individual E is calculated from the slope of this line and the mean efficiency per amplicon is calculated. Nq is set and Cq is determined per sample. Then the mean PCR efficiency and Nq per amplicon and the Cq per sample are used to calculate the starting concentration per sample (Eq. 2), expressed in arbitrary fluorescence units.

Overall, the determination of Efficiency, Nq and Cq includes the following minimal steps in operating the LinRegPCR program:

1. Export data that are not baseline-corrected from the qPCR system to Excel or RDML.
2. Read these raw data into LinRegPCR
   You can check the ***Data*** page, sub-page ***(RDML) Input***, to see whether all data are there.
3. Determine the baselines.
   The program will do a baseline correction per sample and will continue with determining a common window-of-linearity for all samples. The program then sets the quantification threshold Nq at one cycle below the upper limit of this window.
4. Define grouping of samples per amplicon (only after input from Excel) and set W-o-L per amplicon.
5. Check the quality of the individual samples; if needed
   - Correct the baseline in individual samples and recalculate the window-of-linearity.
   - Adjust the window-of-linearity for an amplicon group or an individual sample.
   - Export graphs to clipboard (click the right mouse button on a graph).
6. Save results to Excel or RDML.

3. PCR data analysis with LinRegPCR

3.1 Exporting raw data from the PCR apparatus

The LinRegPCR program reads data from an Excel spreadsheet or an RDML file. The latter format has been implemented in some qPCR systems and the export consists of raw data. All real-time PCR apparatus enable the export of the fluorescence data per cycle, although some only export into a text file format (3.2.1). Note that the exported data have to be corrected for the technical background and, if used, the internal reference fluorescence (e.g. ROX) but should NOT be corrected for the fluorescence baseline. LinRegPCR estimates this baseline per sample and does the baseline subtraction.

Please find out for your own PCR apparatus how to export raw data. You are on your own at this point. There are too many different types of machines to keep track of all these procedures. Sorry.

3.1.1. Opening a text file in Excel

A tab-delimited (*.TXT or *.DEL) or a comma-delimited (*.CSV) text file can be opened in Excel.

1. Start Excel
2. Select ***File*** – ***Open*** from the menu
3. In the ***Open-dialog*** set file type to '***Text files'***
4. Choose the file you want to open
5. In the following dialog, check ***'Delimited'*** and press ***Next***
6. Check ***'Tab'***, ***'Comma'*** or give the separator and press ***Finish***
7. ***Save*** the imported data as an .XLS file before you open it in LinRegPCR.

LightCycler 480 users will find that the raw data output results in a text file of one line per cycle and melting curve point per sample. To convert the amplification data in such a text file into an Excel sheet that can be read by LinRegPCR they can use the program LC480Converter that can be obtained by clicking this link: http://HFRC.nl , go to ***Downloads*** and find the **LC480 Conversion** program under ***Applications***.

3.2 Reading data into LinRegPCR

LinRegPCR can import qPCR data from an Excel file or from an RDML file.

3.2.1 Read data from Excel

Make sure the Excel is running, in 'Ready' state, and that the file containing the data is open! To read the data from Excel:

1. Start LinRegPCR
2. Choose ***File - Read data from Excel***. The program will open the ***Read-from-Excel*** dialog (Fig. 6). The active Excel book and sheet will be shown in the ***book*** and ***sheet*** list boxes.

If these are not the book and sheet that you want to use:

3. Choose the **book** and **sheet** from the dropdown lists that will appear when the arrow head next to the boxes is pressed.
   If the book or the sheet does not show up in the list boxes:
   a. Press ***Cancel*** and go to Excel
   b. Save your data file as an Excel workbook (*.XLS)
   c. Go back to LinRegPCR and choose ***File*** – ***Read from Excel*** from the menu
   d. Return to step 3.
4. Select the ***data file format*** or select your qPCR machine.
5. Set the ***column*** letters and ***row*** numbers to define the range of cells that has to be read.

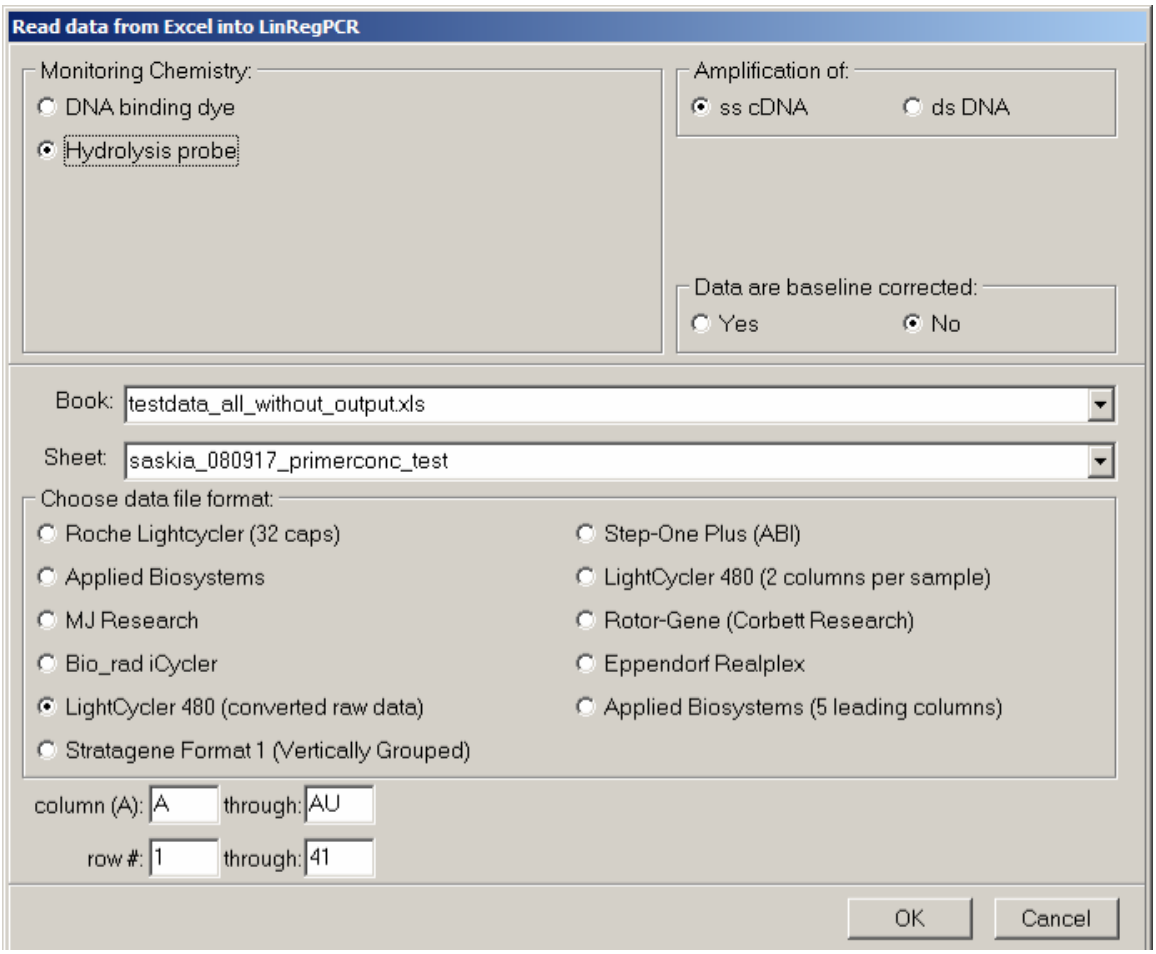


Figure 6: Read-from-Excel dialog. You have to choose the workbook and the worksheet with your data (the currently open workbook and sheet are displayed by default), the format of your input file and the range of cells that you want to read. By default the program expects the data to be not baseline-corrected. If they are, you have to choose ***Yes***. Also your monitoring chemistry and the material that you start the amplification with.

NOTE: As indicated in the description of the formats, each format contains a pre-defined number of leading columns and/or header rows. These header rows and leading columns are used to define samples names or cycle numbers. All described columns and rows have to be given in the range you enter in the columns and rows edit fields of the ***Read-from-Excel*** dialog. When you do not include them, sample names will be based on the fluorescence data which will thus be lost from the analysis. When there are spaces in the entries in the leading columns or header rows that are used to define sample names, those spaces will be deleted. *The format you choose and the column and row values you give will be saved in the Windows Registry. The next time you use the program these values will be displayed as default.*

6. Set ***baseline-corrected*** to ***Yes*** or ***No***. The program expects data that are not baseline-corrected. When your data are baseline corrected by the PCR system you have to indicate this in the ***Read-from-Excel*** dialog. When you forget to do this, you can use the menu option ***Baseline - Set Baseline Corrected*** to skip the baseline correction of LinRegPCR.

Note: you are advised to leave the baseline correction to LinRegPCR. Your qPCR system has used the first cycles to determine a baseline trend which, because of the noise in these early data, is often completely wrong (Ruijter et al 2009).

7. Choose the ***monitoring chemistry*** and the ***input*** material.

These choices are also stored in the Windows Registry and used as default in the next run.

8. Press ***OK*** to import the data.

Do not close Excel! You will need it again to save the results**.**

LinRegPCR recognizes several input formats (Fig. 6). A full description of these formats is given in the Appendix. When your PCR apparatus is not in this list, it will almost certainly have an output format that fits with one of those in the list. If not, try to add rows and columns to make it fit, or try to merge columns to remove excess columns. If you cannot get it right, contact us and we will try to find a solution.

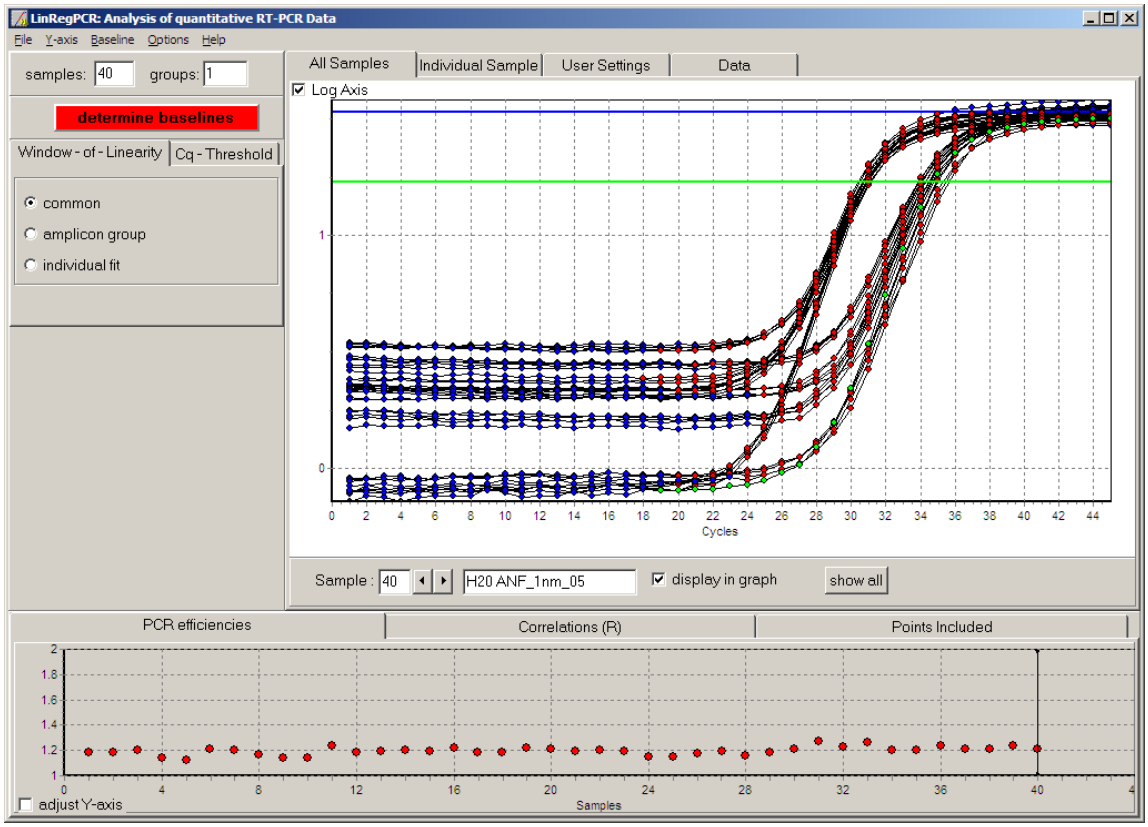


Figure 7: LinRegPCR immediately after importing a not baseline-corrected data set from Excel. Note that the ***Determine Baselines*** button is highlighted to indicate the next step in the analysis.

Trouble shooting:

- Make sure the decimal separator used by your qPCR system, the Windows version on your computer and your Excel program are the same. Excel should be set to ***'use system separators'*** (in Excel go to ***Tools - Options - International***; see 1.4). When your qPCR system exports the wrong separator in a text file you have to do a ***replace*** in a text editor.
- Make sure the status of Excel is **'Ready**' (in the left bottom corner of Excel). When you are editing a cell, importing the data into LinRegPCR will fail. LinRegPCR will then display the message "Call was rejected by callee ".
- Some sheet names are not compatible with LinRegPCR. When a sheet is not displayed in the sheets drop down list: cancel the ***Read-from-Excel*** dialog, go to Excel, rename the sheet and try again.
- Excel Chart sheets are incompatible with LinRegPCR and disrupt the automatic recognition of the data sheets.
- Note that Excel may have 'hidden' sheets that it uses for storage of macros and functions. These sheets may turn up in the sheets-listbox. Choosing one of those sheets will lead to a crash.

3.2.2 Read data from RDML

The RDML file contains the raw fluorescence data that can be read directly into LinRegPCR.

To read data from an RDML file:
1. Start LinRegPCR.
2. Choose ***Read data from RDML*** from the ***File*** menu.

The *Read-from-RDML* dialog (Fig. 8) will appear.

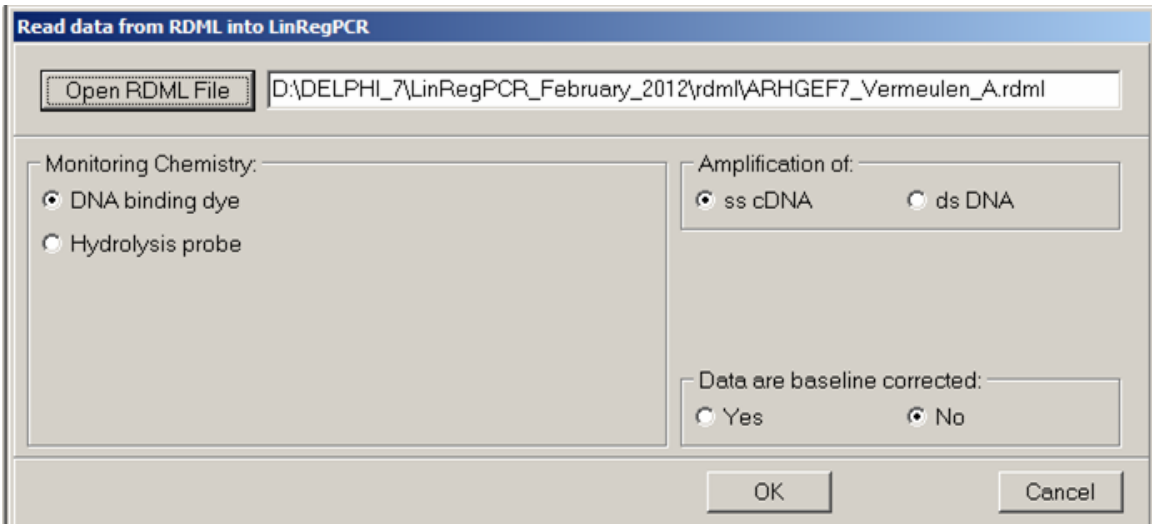


Figure 8: Read-from-RDML dialog. You have to choose the RDML file. By default the program expects the data to be not baseline-corrected. If they are, you have to choose the ***Yes*** radio button. Also choose your monitoring chemistry, and the input material.

3. Press ***Open RDML file*** and choose the RDML file in the ***file browser*** at the top of the dialog (Fig. 9).

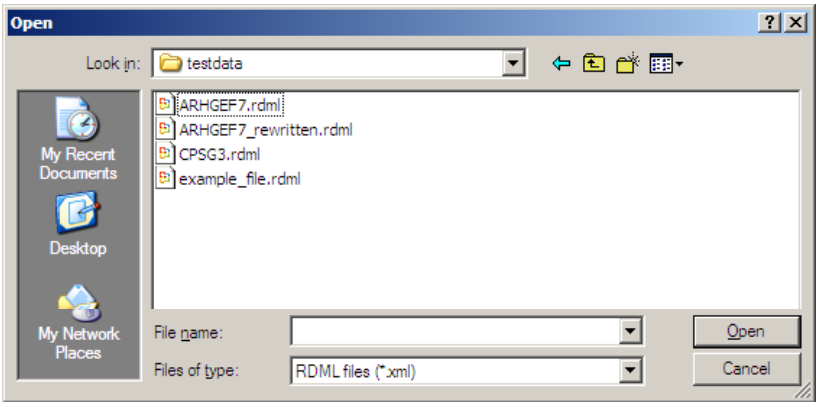


Figure 9: Open dialog that will be displayed to look-up your RDML file. Browse the directory tree, choose the RDML file and press Open.

LinRegPCR can not yet recognize the DNA-monitoring chemistry from the RDML file.
4. Indicate which ***chemistry*** is used to monitor the increasing fluorescence.
5. Choose the ***input*** material (ss cDNA or ds DNA).

When all choices have been made:
6. Press ***OK*** to read the data.

The RDML file will be read and the raw data will be displayed in the ***All Samples*** graph.

The user can inspect the ***Data*** page, sub-page ***RDML Input***, to check whether all data are there (Fig. 10). The ***Amplicon Groups*** Tab and the ***Tissue samples*** Tab will be displayed and will show the information read from the RDML file. The sample names are a concatenation of the reaction number, tissue sample and amplicon, separated by spaces. When the amplicon group information is not present or wrong you can use the controls on the ***Amplicon Group*** page to provide the program with the correct information

Because amplicon information (3.4) is present in the RDML file, LinRegPCR immediately proceeds the setting of a window-of-linearity (3.7) per amplicon after you have pressed the red baseline estimation button (3.3).

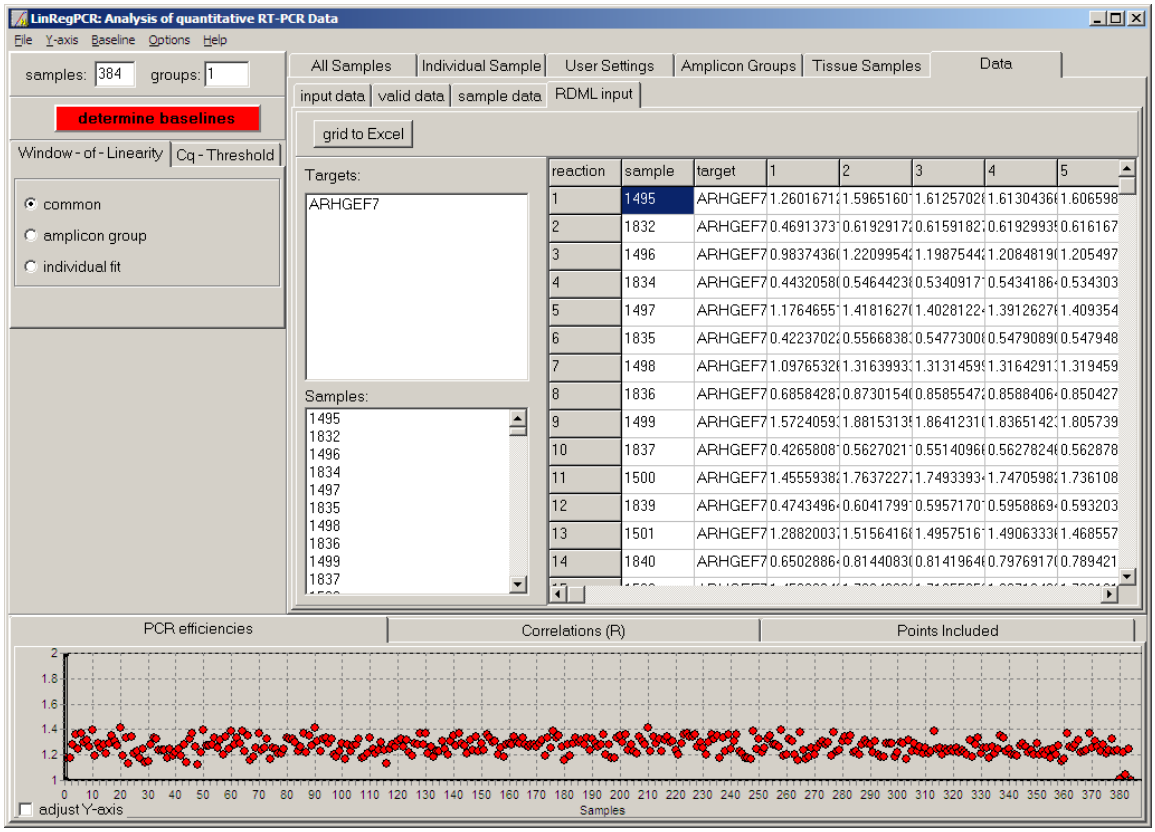


Figure 10: After input from RDML the **RDML Input** tab on the **Data** page will display the data that have been read from the RDML file. The Targets listbox contains the names of the amplicons found in the RDML file; the Samples listbox displays the names of the tissue samples. Reaction numbers are read from the RDML file.

NOTES:

- The RDML output of some qPCR machines is in RDML version 1.0, whereas the latest standard is version 1.1. LinRegPCR can read both versions and you will not notice the difference while importing data. However, LinRegPCR saves the results of the analysis in RDML 1.1 format to maintain compatibility with other software based on RDML import. See 3.12.2 for details on saving to RDML.
- In the definition of RDML a floating point value is supposed to be formatted with a decimal point. However, some qPCR systems place the decimal separator that is used by the computer in the RDML file. LinRegPCR can read RDML files with a decimal comma. However, LinRegPCR saves the results of the analysis in with decimal points to maintain compatibility with other software based on RDML import. See 3.12.2 for details on saving to RDML.
- The RDML file is a zipped rdml_data.xml file. Unfortunately LinRegPCR cannot handle all types of zip-archives. In case the qPCR machine's zip is incompatible with LinRegPCR you will be notified. In that case you will have to re-zip your file with e.g. WinZip:
  1. make a copy of the original RDML file as a backup
  2. rename the AAA.rdml file into AAA.zip
  3. extract the rdml_data.xml from the AAA.zip file
  4. use your Zip program to create a new BBB.zip file without directory information.
  5. rename the file BBB.zip to BBB.rdml

Where AAA and BBB are the original and new names of your RDML file. While renaming the Zip to RDML files and vice versa you can ignore the warnings that this may affect the usability of these files.

- After import of RDML data, the Group Statistics tab on the Amplicon Group Tab will become visible (see 3.4.1).

3.3 Baseline Estimation and flagging of deviating samples

LinRegPCR determines the baseline fluorescence **per sample** by reconstructing the log-linear phase. It determines the cycle at the maximum of the Second Derivative (SDM) of the observed fluorescence values in each sample. The SDM cycle is the start of the plateau phase. Then LinRegPCR does an iterative search for a baseline value that gives the longest straight line of data points down from the plateau phase. For details on the baseline estimation procedure: see Ruijter et al., Nucleic Acids Research, 2009.

As explained in the above paper, an accurate baseline estimation is a requirement for the correct determination of E and to a lesser extend C_q. The implemented baseline estimation requires a sufficiently long log-linear phase and thus optimized qPCR conditions have to be used. However, when the baseline-to-plateau distance is too small the baseline determination can fail. LinRegPCR determines and flags the following deviating samples during the process of baseline estimation. Also the manual correction of the assignment of such deviating samples is described here.

- **no-amplification**: samples with less than 7-times overall increase between the first cycles and the last cycles. These samples will be ignored by the program and are always excluded from the calculation of the mean PCR efficiency and will receive a starting concentration of -999. This assignment can be overruled by un-checking the ***Mark sample as excluded*** checkbox in the lower right of the individual sample window.
- **baseline error**: samples for which no baseline value can be determined because no fluorescence value between the SDM and the minimum observed value results in a straight exponential phase. These samples will be ignored by the program and are always excluded from the calculation of the mean PCR efficiency and will receive a starting concentration of -999. This assignment can be overruled by setting a manual baseline. The menu option ***Baseline - Manual Baseline*** displays the manual baseline controls (3.3.2).
- **noisy sample**: sample in which the fluorescence in cycle *C* is higher than it is in cycle *C+1* in a W-o-L just below the SDM cycle (Fig. 12A). This high in the exponential phase, the fluorescence values should be continuously increasing. In such a case, the baseline estimation will result in a warning that the dataset is **too noisy** to be analyzed automatically. This does not have to mean that the dataset is completely lost. The program will display the ***manual baseline*** and ***window-of-linearity controls***. See 3.8 for help on the handling of noisy and other law quality datasets..
- **no-plateau**: samples that do not reach the plateau phase. This assignment cannot be overruled. The user can decide to include or exclude the **no-plateau** samples from the calculation of the mean PCR efficiency (see 3.6). For these samples a starting concentration will be calculated when they are not also an **efficiency outlier**. The inclusion or exclusion choices are made on the **User Settings** page (3.10, Fig 23). *These choices are saved to the Windows Registry and set automatically next time you use LinRegPCR.*

The deviating and flagged samples can be separately displayed using the menu item: ***Options - Display Flagged subsets***.

You can also use the ***Mark sample as excluded*** checkbox to exclude samples that you do not want to include in the calculation; they will be marked as **User Excluded** in the results.

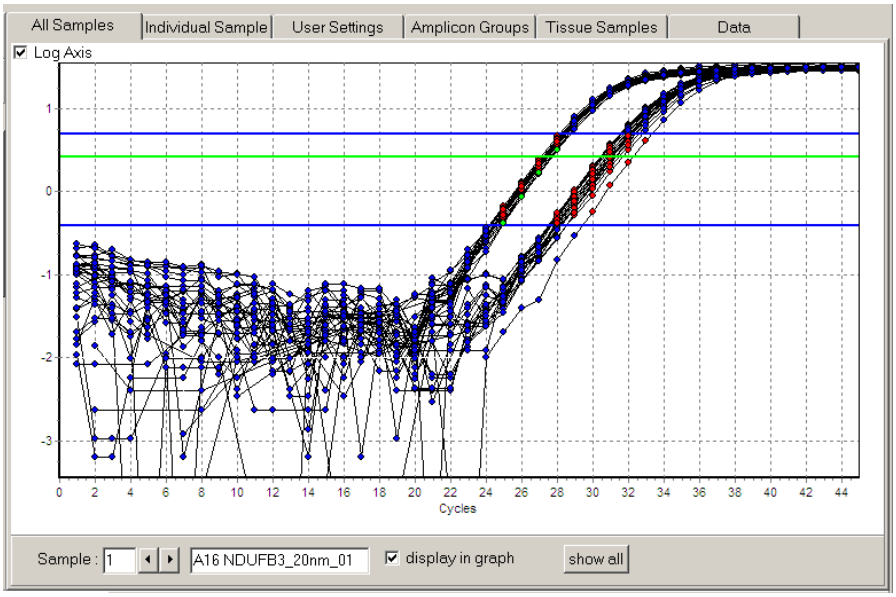


Figure 11:
All Samples graph of the data of Fig. 7 after estimation and subtraction of the fluorescence baseline.

3.3.1 Relaxing baseline estimation criteria

When you encounter a lot of baseline errors and noisy samples it may be that the Log-linear phase in your dataset is too short. This is most probably because the baseline is too high which is limiting the baseline-to-plateau distance. The program then has not enough data points in the log-linear phase to estimate the fluorescence baseline. It may also be that measurement noise leads to discontinuous data in the log-linear phase. Both problems make that the data points downwards from the plateau phase do not form a continuously decreasing set of points. Because LinRegPCR uses only the points that are continuously decreasing from the plateau phase too few points are available for the baseline estimation. In that case the baseline estimate will be wrong, in most cases too high.

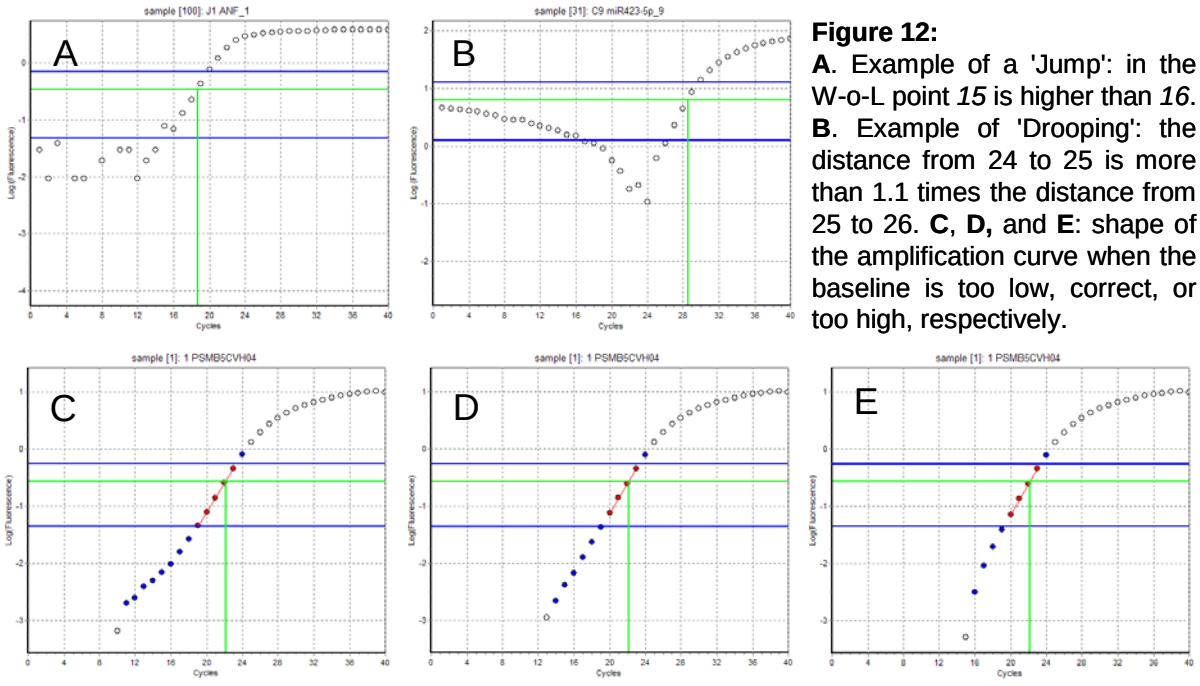


Figure 12:
A. Example of a 'Jump': in the W-o-L point *15* is higher than *16*. **B**. Example of 'Drooping': the distance from 24 to 25 is more than 1.1 times the distance from 25 to 26. **C**, **D,** and **E**: shape of the amplification curve when the baseline is too low, correct, or too high, respectively.

When only a few samples are deviating you can choose ***Manual baseline*** setting to rescue those samples (3.4.2). Otherwise you can choose to relax the continuity criterion that LinRegPCR uses to determine which points to include into the Log-linear phase.

On the **User Settings** page (3.10, Fig. 23), '***strictly continuous Log-linear phase***' is set as the default. The fluorescence values are then required to increase continuously and the lowest point should not 'droop'. A drooping point is defined as a point for which the distance to the next point is more than 1.1 times the distance between the next two points (Fig. 12B).

When you choose '***continuous with jumps***' you allow that, as a result of measurement noise, in the log-linear phase data points C is higher than data point *C+1* (Fig. 12A). You can also choose to '***allow jumps and drooping points***' and thus also allow the lowest point to droop (Fig. 12B). Note that with the 'strict baseline estimation' the data points in and below the W-o-L are <u>on</u> a straight line, in both 'relaxed' cases the included data points are -or should be- <u>around</u> a straight line (Fig 12D). Because you now allow some measurement noise, the variation between individual efficiency values will be increased for 'relaxed' baselines. Check that this not affects the mean efficiency too much.

NOTE*:** For some experiments the ground phase noise and a limited signal-to-noise ratio can lead to a baseline error assignment for samples that show a good straight log-linear phase. When you encounter such samples during the check of the amplification curves (3.5), you can press 'ignore error***' in the ***manual baseline controls*** to include the sample. In the results those samples will be marked as '**manual baseline**' and '**included by user**'. When you want to undo this action you have to press '***reset***' in the ***manual baseline controls***.

3.3.2 Manual setting of the fluorescence baseline

When the '***also show raw data***' check box on the **Individual Sample** graph is checked, the raw data are shown as red diamonds and the baseline is indicated by a red line (Fig. 13). The user can move the baseline by drag-and-drop of the red baseline. When you do so, the ***manual baseline controls*** will appear at the right bottom of the **Individual sample** page.

The baseline value can be entered in the ***edit*** field or changed stepwise by clicking the ***up-down arrows***. The ***step size*** is determined by the choices (0.0001 thru 0.01 times the current baseline). When setting a manual baseline is required, the aim should be to construct a straight line of data points downwards from the plateau phase (Fig. 12 D). When the points slope upwards at lower cycles the baseline is too low; when they suddenly drop downwards the baseline is too high (Fig. 12 C and E, respectively; Ruijter et al, 2009).

When you start to manual threshold setting it is a good idea to start with the minimum observed fluorescence of the sample:

1. Press the ***get Min.Flu***' button to have the value displayed in the edit field
2. Press ***apply baseline*** to subtract this value from the raw data
3. Use the ***up-down*** arrows to adjust the baseline.

***NOTES*:**

- When many samples require a manual baseline setting it may be good to try to relax the baseline criteria first (see section 3.3.1).
- When the current baseline is 0 the ***up-down*** arrows will have no effect. Fill in a small number or press the ***Get Min.Flu*** button and then press the ***apply baseline*** button.
- When a manual baseline is set, the window-of-linearity should be recalculated by pressing the ***Recalculate W-o-L*** button at the bottom right of the page.

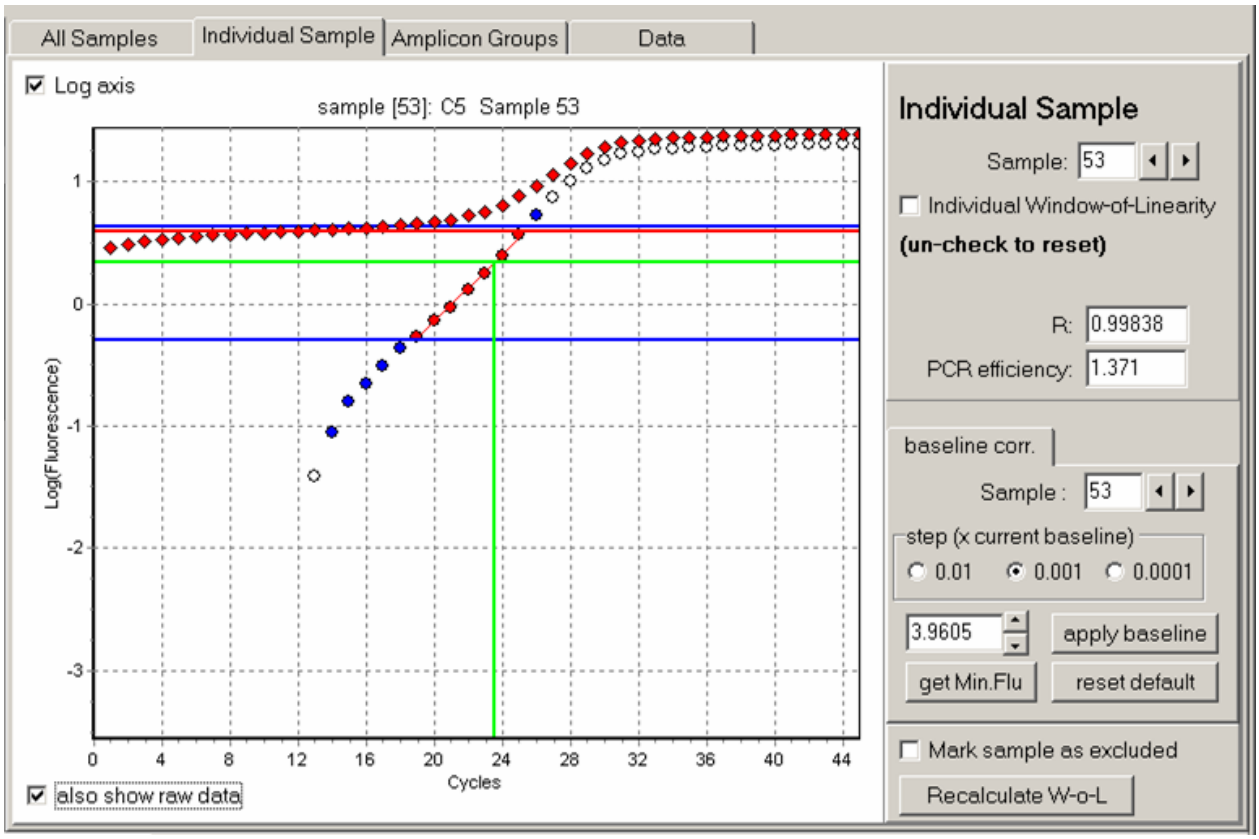


Figure 13: Individual Sample page showing the baseline-corrected sample (white, blue and orange points), the raw data (red diamonds) and the fluorescence baseline (red line). The baseline can be set manually with the controls on the lower right of the screen.

- When you apply a manual baseline to a **baseline error** sample, the sample will be included in all calculations; when you check the ***Mark sample as excluded*** checkbox it will be excluded from the mean efficiency but receive a N0 value.
- Samples with a **manual baseline** will be labeled as such in the output (3.13)

3.4 Amplicon groups

The subset of samples in which the same pair of primers is used is called an amplicon group. Individual samples have slightly variable observed PCR efficiencies but the PCR efficiency per amplicon is assumed to be constant. Therefore, LinRegPCR uses the mean PCR efficiency per amplicon in its calculations (Ruijter et al. 2009). To enable this, the samples have to be assigned to amplicon groups (Fig. 14).

After **RDML** input, the samples are already assigned to their amplicon group and LinRegPCR has set a W-o-L per amplicon immediately after reading the data (3.2.2). You can check and, if necessary, correct this assignment on the ***Amplicon Groups*** page.

After **Excel** input, amplicon groups have to be assigned. To do this, go to the ***Amplicon Groups*** page (Fig. 14) where you find 4 methods to assign amplicon groups.

1) no groups: all samples are one group.

When this is the case you are advised still to give the amplicon name in the ***amplicon*** field and thus replace 'not_named'. This will help you identify the results after you saved them to Excel or RDML. After the baseline correction a common window-of-linearity was already set assuming that in all samples the same amplicon is produced.

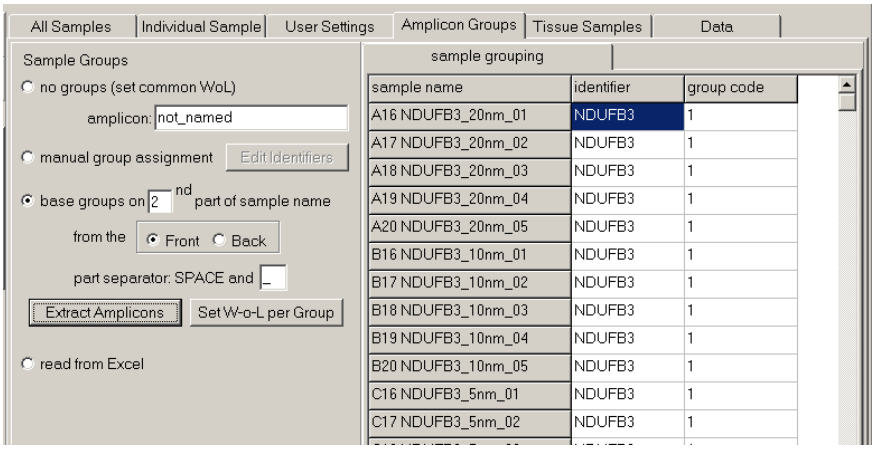


Figure 14: Amplicon Groups Tab after setting the grouping parameters to 'base groups on the second part from the front of the sample name using space and underscore as separators'.

2) manual group assignment:

When the sample names do not contain information on the amplicon, you can assign the grouping manually.

1. choose ***'manual group assignment'***
2. press the ***Edit Identifiers*** button

This button puts the sample group grid in edit mode. You can now enter values into the **identifier** column of this grid. Activate a cell by ***double clicking*** and then enter the identifier. The group code is assigned automatically and cannot be edited.

When you have given all identifiers:

3. Press the ***Apply Groups*** button (which is the same button you just pressed).

The program now sets a window-of-linearity and fluorescence threshold per amplicon group.

3) base groups on part X of the sample name

When the sample name contains the amplicon identifier in a fixed format, you can base the group assignment on the sample name.

1. choose '***base groups on***'
2. enter the ***index*** of the part of the sample name that describes the amplicon
3. indicate whether you are counting from the ***front*** or the ***back*** of the sample name
4. when the ***separator*** is not (only) a space, give the extra separator, such as the underscore (see example Fig. 14)
5. press the ***Extract Identifiers*** button

The program now extracts the amplicon **identifier** and assigns group codes. Check the identifiers and change the above entries when required. If you want you can switch to ***'manual group assignment'*** to correct the identifiers.

When the assignments are correct:

6. press ***Set W-o-L per Group***. The program now sets a window-of-linearity and fluorescence threshold per amplicon group.

4) read from Excel

Choosing this option will display the ***read-from-Excel*** controls (Fig. 15).

1. Use the list box controls to choose the right ***Excel book*** and ***sheet***.
2. Give the ***column*** with the amplicon identifiers.
3. Give the ***first*** and ***last row*** in which the amplicon identifiers are to be found.

Do NOT include a header row in the row specification. The rows should be in the same order as the samples in the fluorescence input.

4. Press ***Read Amplicons***.

The identifiers will appear in the grid (Fig. 14). When you are satisfied:

5. Press ***Set W-o-L per Group***.

The program now sets a window-of-linearity and fluorescence threshold per amplicon group.

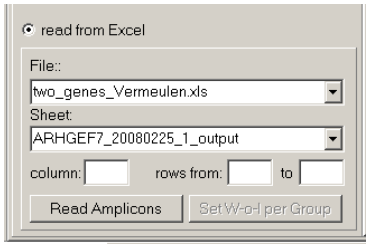


Figure 15:
Read amplicon groups from Excel. Use the list box controls to choose the book and sheet where the amplicon per sample can be read.

You can look at the groups separately by browsing through the groups with the controls that have appeared on the lower left of the screen (Fig. 16). You can select individual samples by clicking on the sample in the PCR efficiencies graph.

***NOTES*:**
- LinRegPCR has now set a window-of-linearity per group. When you browse through the groups you can now also make corrections in this W-o-L setting (see 3.7.2)
- In defining amplicon groups take care of handling the minus-RT samples. When you are sure the amplified product in these samples is your amplicon-of-interest, you can include them in the amplicon group. However, when the product may be something different (e.g. genomic DNA in stead of cDNA) the minus-RT samples may display a different PCR efficiency. In such a case it is better to treat them as a separate amplicon group or to exclude them manually by checking the ***Mark Sample as excluded*** box.

3.4.1 Groups Statistics
When Amplicons have been defined the Groups Statistics Tab becomes visible on the Amplicon Groups Tab (Figure 15a). With the 'Calculate' button on this page, the user can calculate the mean efficiency per amplicon group, together with its standard deviation, SEM and relative standard error (RSE). The displayed can be exported to Excel. Do not forget to rename the sheet added to Excel.

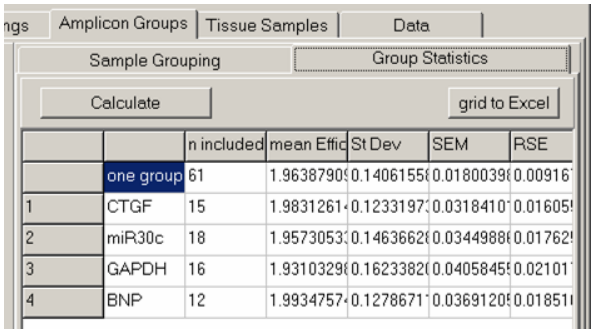


Figure 15a:
Group Statistics can be displayed on the Amplicon Groups Tab and exported to Excel.

***NOTES*:**
- After import of RDML data, which contain amplicon grouping information, the Group Statistics Tab becomes visible.
- After saving the results to Excel or RDML the Groups Statistics will be displayed.
- Use the Calculate button to recalculate the Groups Statistics when samples are excluded or included by the user (see 3.5)

3.5 Check individual samples

Errors can occur during sample preparation and measurement noise will always be present. Therefore, the baseline estimation cannot be made completely fail-proof. The user of LinRegPCR is urged to browse through the groups of samples (inspecting the graph on the All Samples page) or through the individual samples to check deviating samples that were missed by the program or to correct faultily assigned flags.

The user can exclude a sample by checking the ***Mark sample as excluded*** box at the lower right next to the graph. The user can also use the manual baseline controls to try to correct the baseline in a sample when the program failed to find the right baseline because the data were too noisy (3.3.2 and 3.8). When extra samples are marked as excluded, or when the baseline of a sample is manually corrected, the user should press ***Recalculate W-o-L*** to adjust the W-o-L to the new data.

When required, the user can also:

- Change the settings of the common or amplicon group window-of-linearity manual W-o-L settings (see 3.7).
- For individual samples, data points can be included in - or excluded from - the W-o-L. *This will set an individual W-o-L for that sample. Samples with an individual W-o-L will be included in the calculation of the mean PCR efficiency of their amplicon group.*

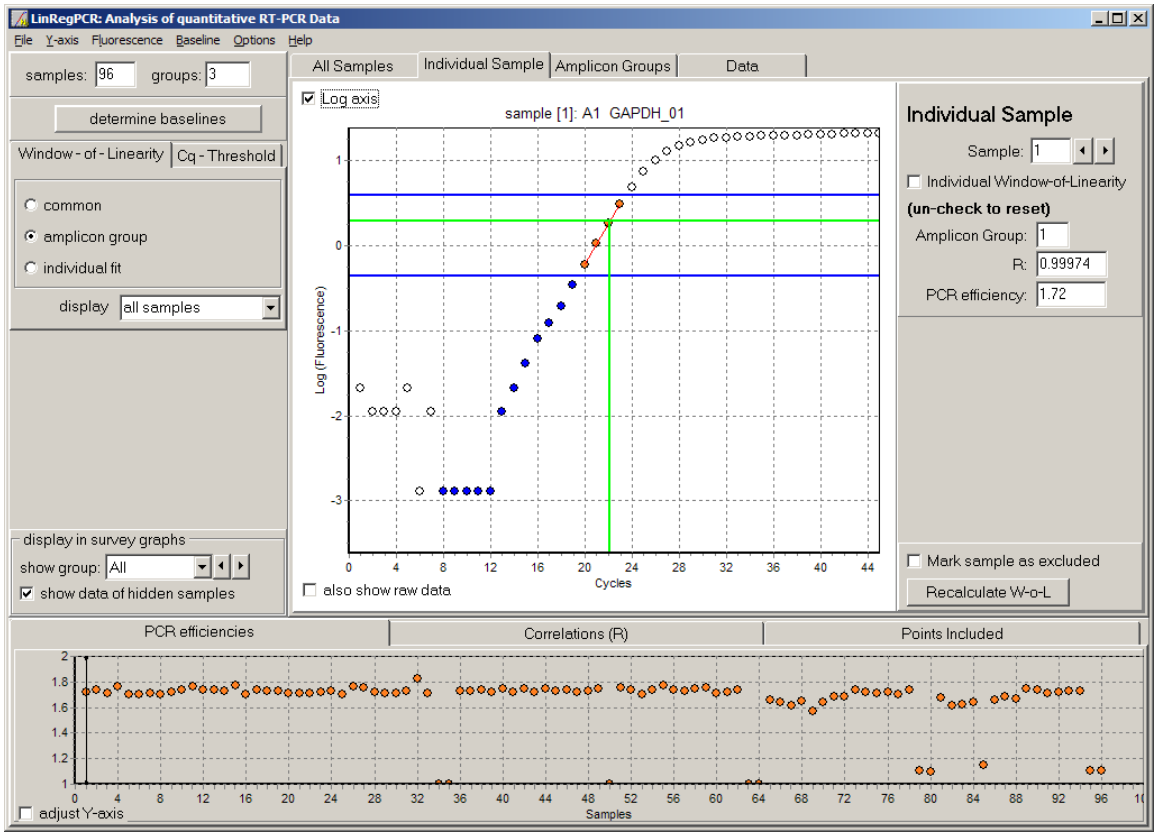


Figure 16: LinRegPCR enables the user to browse through the individual samples and to check the baseline correction. Samples with strongly deviating PCR efficiency, as shown in the bottom panel, require special attention.

Survey graphs

At the bottom of the LinRegPCR screen there is a notebook containing three Tab pages displaying a survey of all PCR efficiency values (Fig. 17A), correlation coefficients of the

linear fit (Fig. 17B) and number of points included in the W-o-L (Fig. 17C).These graphs can be used to pin-point deviating samples:

- PCR efficiencies should be normally distributed per amplicon and should never exceed 2.
- the correlation coefficients should be close to 1; a low R^2 can indicate the inclusion of measurement noise in the W-o-L.
- the number of included points should be 4 although a small number of samples with only 3 points is acceptable when the correlation coefficient is high.

In each of the survey graphs you can click on a sample to highlight this sample in the ***All Samples*** graph **(**it will there be shown as green points; e.g. Fig. 11) or in the ***Individual Sample*** graph. The easiest check of deviating samples is to inspect the chosen sample in the **Individual Sample** graph (Fig. 16).

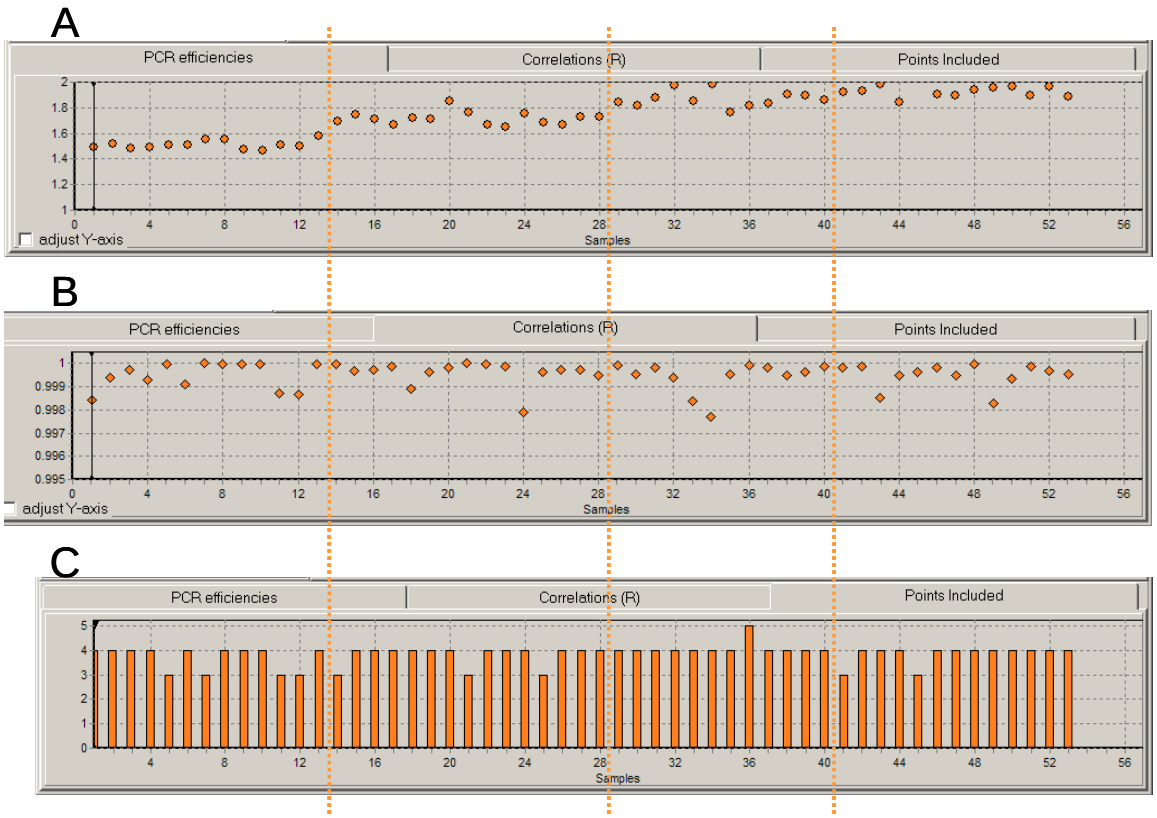


Figure 17: Notebook with Tab pages to display the survey graphs of PCR efficiencies (**A**), correlation coefficients (**B**) and number of included points (**C**) for 4 amplicon groups (indicated by vertical dotted lines) with increasing efficiency value. The Y-axis of the PCR efficiency and correlation coefficient graphs can be adjusted to show all values; by default only allowable values are shown.

The data points in the survey graphs are red when you are using a common window, orange when you have defined amplicon groups and yellow when a sample is fitted to an individual window.

Exporting graphs

When you click the right mouse button on a graph, the graph is copied to the Windows Clipboard as a bitmap (2.3.3; Fig. 5). This bitmap can be pasted into your presentation.

3.6 Mean efficiency

The group of samples in which the same pair of primers is used are called an amplicon group. Amplicon groups can be assigned with the controls on the Amplicon Groups page (see 3.4). The PCR efficiency per amplicon is assumed to be constant; nevertheless individual samples have slightly variable observed PCR efficiencies. Therefore, it is recommended to use the average efficiency of the samples for each amplicon group to calculate the target quantity. (Ruijter et al., Nucleic Acids Research, 2009).

Determining the correct mean efficiency is very important and deviating samples should therefore be excluded. On the User Settings page (see 3.10, Fig. 23) you can decide on to include or exclude deviating samples from the calculation of the mean efficiency.

- **No amplification**: Samples without amplification are always excluded.
- **No Plateau**: Samples that do not reach the plateau phase can be excluded. Excluding these samples may give you a better mean efficiency but, because this mean is then based on fewer samples, it may not be the best mean.
- **Cq too low**: Samples of which the Cq value is too low to determine the correct PCR efficiency (only with cumulative fluorescence; see also 3.9) can be excluded.
- **Efficiency outliers:** Samples with a deviating individual PCR efficiency can be excluded from the calculation of the mean efficiency. You can give the range around the median efficiency that you want to allow. The default range of 5% means that around a median efficiency of 1.83, a range from 1.78 to 1.88 is allowed. This exclusion will only affect the mean efficiency when the distribution of observed individual efficiencies is much skewed.

When you have chosen to exclude **efficiency outliers**, you can display those samples by choosing the menu option ***Options - Display Flagged Subsets*** and then choose the set from the list box that appeared at the left side of the window (Fig. 16).

***NOTES*:**

- Excluded samples are only excluded from the calculation of the mean efficiency per amplicon group. The sample will receive a starting concentration (N0) in the output based on the mean PCR efficiency of the remaining samples.
- The choices you make will be saved to the Windows Registry and will be loaded as default the next time you run the program.

3.7 Window-of-Linearity and Nq threshold

To determine the mean PCR efficiency for an amplicon group, LinRegPCR uses a selection of data points in the log-linear phase to determine the PCR efficiency (E) per sample. This selection of cycles is called the **Window-of-Linearity (W-o-L**). Default the W-o-L width is set to 4 cycles. The W-o-L is set in such a way the variation between individual efficiency values of the included samples is lowest.

In the Individual Sample graph the W-o-L is indicated by blue horizontal lines that represent the upper and lower limit of the W-o-L. The data points used for the estimation of the PCR efficiency are marked in red (or orange when amplicon groups are defined).

The fluorescence threshold Nq, represented by the horizontal green line, is by default set at 1 cycle below the upper limit of the W-o-L; the vertical green line indicates the Cq value.

In LinRegPCR you can distinguish between three window-of-linearity settings:

- **common window**, which you should use when all samples are amplified with the same pair of primers.

– **amplicon group window**, which you should use when different sets of primers are used and different efficiency values per groups are expected (3.4).
– **individual window**. This option is included to provide of backward compatibility. When you choose this option the program will set a W-o-L per sample in the same way it did in LinRegPCR; version 7.4. The PCR efficiency survey graph (Fig. 17A) will show that the variation in PCR efficiency values increases significantly.

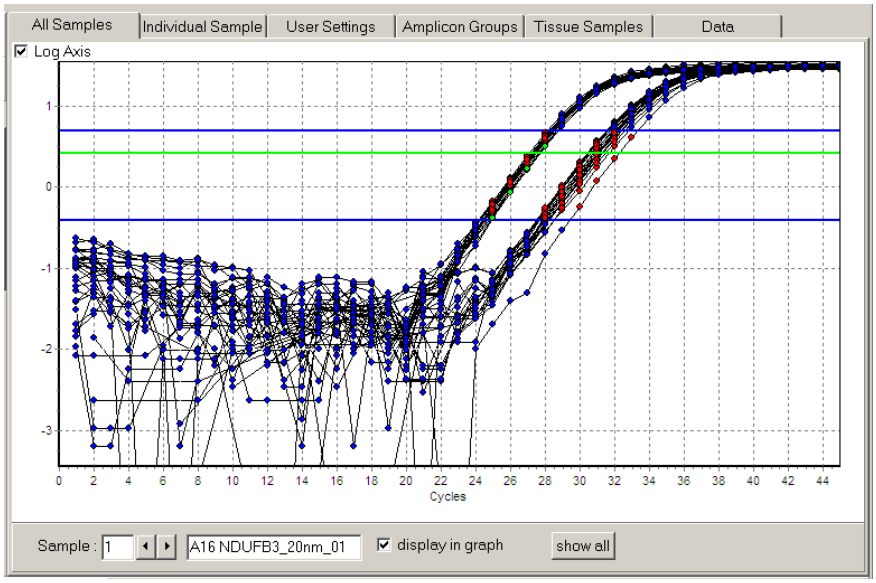


Figure 19: All Samples graph showing a common window -of-linearity (blue lines) and quantification threshold (green line).

3.7.1 Automatic setting of W-o-L and N_q threshold

After the baseline correction a common W-o-L is set. After the definition of amplicon groups LinRegPCR sets a W-o-L per amplicon group. LinRegPCR sets the window-of-linearity by looking for the W-o-L with the smallest coefficient of variation (CV) of the individual PCR efficiencies in the group of samples included. To this end, LinRegPCR moves a W-o-L, starting at the plateau phase, downward till a sample is not longer included or till the CV shows a second minimum. The latter criterion was added after the publication of the method (Ruijter et al 2009) to avoid the situation where the decreasing average efficiency close to the baseline causes the CV to decrease.

All W-o-L's can be adjusted manually using the W-o-L controls that are displayed when you choose the menu option ***Display W-o-L Controls*** from the ***Options*** menu.

The Nq threshold is automatically set at 1 cycle below the upper limit of the W-o-L.
– The common threshold is used when you choose a common window or when you have chosen to use individual windows.
– The threshold per amplicon group is used when you choose amplicon windows.

***NOTES*:**
- In case you want to do your own comparative Cq-like calculations with the exported efficiency and Cq values it is required to use a common threshold for all amplicon groups.
- You can set a manual threshold using the edit field on the ***Cq Tab*** of the notebook on the top left of the screen. When you have defined amplicon groups, the given threshold will be set for the currently displayed group.

The N0 results that LinRegPCR reports are NOT affected by the choice of Nq threshold. However, obviously the Cq values do depend on the fluorescence threshold.

3.7.2 Manual setting of the Window-of-Linearity

The blue lines in the ***All Samples*** graph and the ***Individual Sample*** graphs indicate the upper and lower limit of the window-of-linearity. Data points between these lines are used by the LinRegPCR to calculate the PCR efficiency per sample. The user can change these window settings manually if required.

- Common window.

When a common window is used the user can drag-and-drop the window limits in the ***All Samples*** graph. Click on the blue lines until a vertical ↔ cursor appears, drag the line and click again to drop the line. This action will also open the ***Common Window*** controls (middle left on the screen; Fig. 11) that allow you to use the ***up-down*** arrows to fine tune window settings. You can also fill in values that you know from earlier analyses.

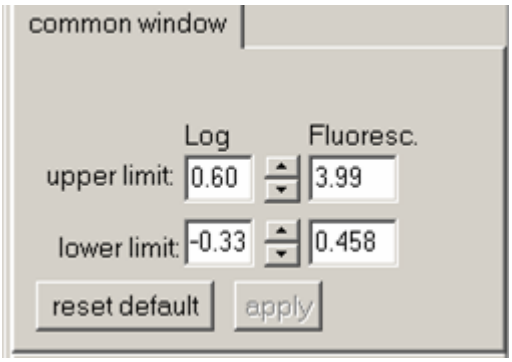


Figure 20: Controls that can be used to set a common window-of-linearity. Data can be entered in the edit fields, or the up-down arrows can be clicked. When the edit fields are used, the values are applied after pressing the ***apply*** button. Similar controls are available for group W-o-L and individual W-o-L setting.

- Amplicon windows.

When amplicon groups have been defined, and you are browsing through the amplicon groups, the ***All Samples*** graph will show the W-o-L per amplicon group. You can drag-and-drop the window limits or use the amplicon window controls on the middle left of the screen to adjust the window.

- Individual windows.

When you drag-and-drop the blue window-of-linearity limits in the ***Individual Sample*** graph, the sample will be marked as having an individual window (data points will be displayed in yellow). This will also open the individual window controls at the middle right of the screen. You can also include or exclude a data point in the individual window by clicking on the point itself. Samples with an individual W-o-L will still be treated as member of the amplicon group.

***NOTES*:**

- The common and amplicon group windows set automatically by LinRegPCR are the windows that give the least variation between individual efficiency values per group. Adjusting the window manually will probably lead to increased variation between samples.
- Setting an individual window will only help you bring the individual efficiency closer to the mean efficiency and thus lead to a better mean efficiency. The individual target quantity depends mainly on the Cq value which is not affected by the individual W-o-L setting.
- Only when all samples are set to an individual window by choosing **individual fit** from the ***Window-of-Linearity*** tab (upper left) the individual efficiencies will be used in the calculation of the N0 values. This setting is equivalent to the operation of LinRegPCR up till version 7.4 and is no longer recommended (Ruijter et al. , 2009).

3.7.3 Manual setting of the N_q threshold

The horizontal green line in the ***All Samples*** graph and the ***Individual Sample*** graph represents the Nq threshold; the vertical green line indicates the Cq value. You can change the Nq threshold by drag-and-drop of the horizontal green line.

When you are in the ***All Samples*** or in the ***Individual Sample*** graph, and you are using a common W-o-L, drag-and-drop of the green Nq threshold will change the common threshold value. However, when you have defined amplicon groups, and are displaying only one amplicon group, drag-and-drop in the ***All Samples*** graph will change the Nq threshold for that amplicon group only. In the ***Individual Sample*** graph the drag-and-drop of the Nq threshold will then also set the threshold for the amplicon group in which the current sample is placed.

3.8 Handling noisy datasets

A noisy sample is defined as a sample in which the data in the window-of-linearity are not continuously increasing. This mostly occurs when the observation noise is too large. When **noisy samples** occur, LinRegPCR cannot accurately determine the fluorescence baseline. So, you probably also have a number of samples that have a **baseline error** assigned to them.

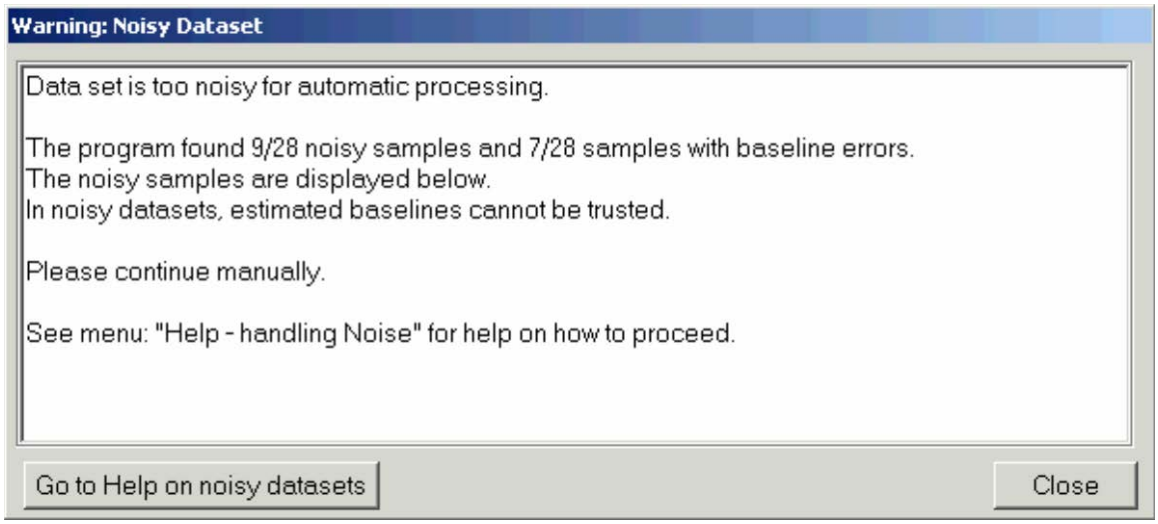


Figure 21: Warning message that pops-up when the program has detected a ***noisy sample*** during the estimation of the fluorescence baselines.

The program checks for noisy samples immediately after the estimation of the fluorescence baseline. When noisy samples occur, the program marks **noisy samples** for exclusion from calculation of the mean efficiency. Then it determines a common window-of-linearity for the remaining samples. After this the program directs you to the tab page on ***Handling noisy data*** in the Help. A short-cut to this page is then added to the ***Help*** menu. The program also shows the ***Warning*** window (Fig. 21)

3.8.1 Avoid noisy samples in the lab

Noisy data might occur when the distance between baseline and plateau phase is not large enough which leaves the program with not enough room to determine the baseline accurately. To solve these issues you have to go back to the lab. You have several ways to improve the data quality in the future.

Firstly look at the exposure-time setting in your PCR apparatus and set this to a fixed exposure time. The automatic exposure time setting of your qPCR apparatus exaggerates noise at low fluorescence intensities. Secondly you can try to increase the baseline - plateau distance by lowering the primer concentration in the PCR reaction. When you are using a fluorescent probes and quenched reporters you have to make sure that the quenching efficiency of the undigested probe is as high as possible (see Tuomi et al 2010 for suggestions).

3.8.2 Rescue noisy samples during data analysis

You can try to RESCUE the current dataset by doing the data analysis manually. To this end you will have to:

1. set the correct baseline for, at least, the 'noisy' and the 'baseline error' samples (3.3.2).
2. set a manual window-of-linearity (3.7.2).

Note that the Cq values are hardly dependent on the baseline correction. So, when using the mean efficiency per amplicon and the individual Cq values, like LinRegPCR does, you may get reliable data even from a noisy data set. But keep in mind that you are trying to rescue low quality data. Baseline errors and Noisy data stay marked in the output.

NOTE

- When many noisy samples require a manual baseline setting it may be best to first look at the effect of relaxing the baseline criteria (3.3.1).
- When you have set the correct baselines for the **noisy samples**, it may help to do the same for the samples that were assigned a **baseline error**. It is advisable to also check the samples that passed the automatic baseline estimation.
- When you have checked and corrected the baselines for all samples you can try to set an automatic window-of-linearity by pressing the ***Recalculate W-o-L*** button.
- The fluorescence threshold for determining Cq values is automatically set at 1 cycle below the upper limit of the W-o-L. You can choose your own threshold by entering the value in the edit field on the **Nq-threshold** tab (top left).

3.9 Probes-based monitoring chemistries and cumulative fluorescence

As described in 2.2, LinRegPCR can be used with all types of amplification monitoring chemistries. DNA binding dyes and hybridization probes increase their fluorescence upon binding with dsDNA and result in fluorescence data that are proportional to the amount of DNA. On the other hand, hydrolysis probes (e.g. TaqMan® probes) contain a fluorescent reporter that is quenched by the proximity of a quencher molecule. After the probe is bound to ssDNA, it is hydrolyzed by the polymerase and the reporter is free to fluoresce. It stays fluorescent in subsequent cycles and therefore this monitoring chemistry results in fluorescence data that are cumulative. The quenching of the reporter in hydrolysis probes is often not very efficient, causing a high fluorescence baseline and a low baseline-to-plateau distance. This distance is much less than the one observed in datasets obtained with DNA binding dyes (Tuomi et al 2010). Therefore, cumulative fluorescence data may cause problems in setting fluorescence baselines and may give more variable PCR efficiency values per amplicon.

On a logarithmic fluorescence axis, after enough cycles, every monitoring chemistry results in an amplification curve that becomes a straight line with a slope that is equal to the slope that would have been found when the fluorescence had not accumulated (Tuomi et al 2010). This means that the PCR efficiency derived from the slope of this curve is a correct estimate of the PCR efficiency used in Eq. 1.

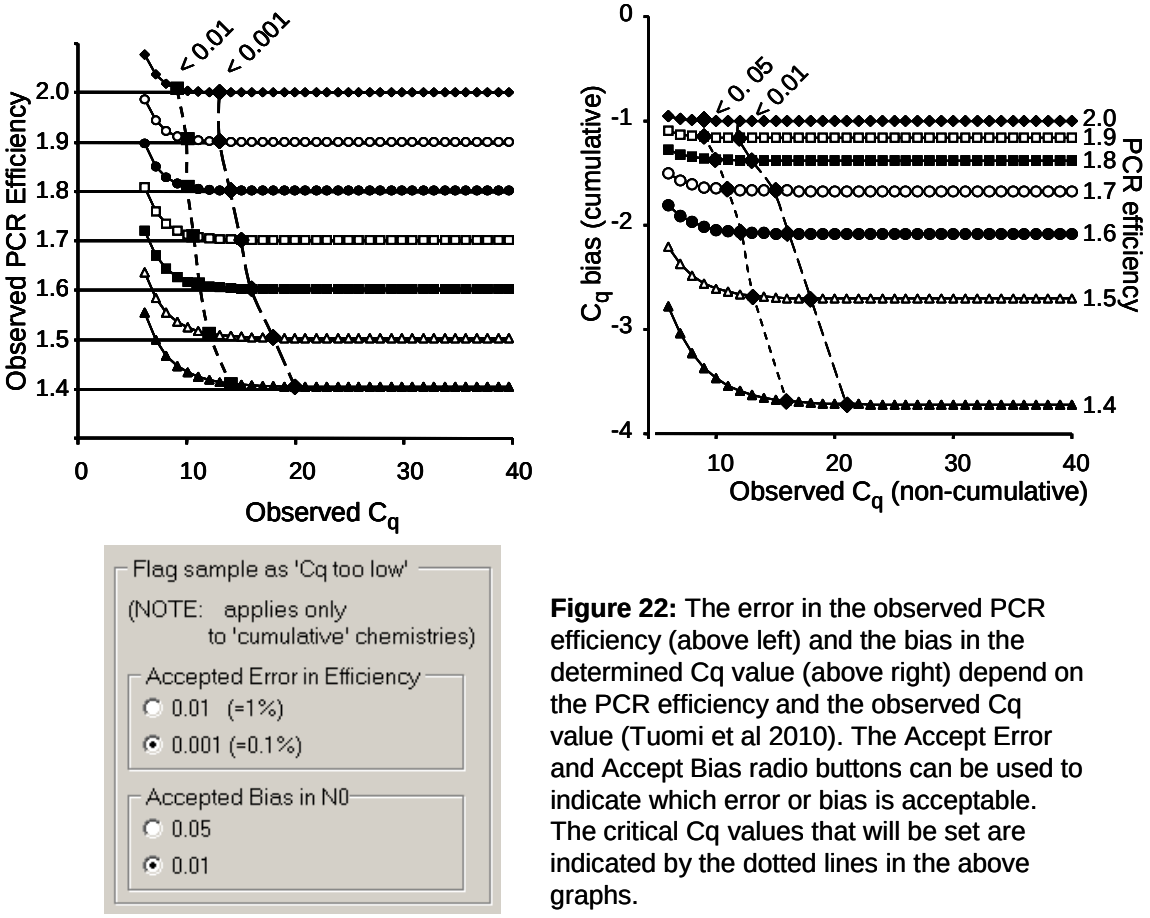


Figure 22: The error in the observed PCR efficiency (above left) and the bias in the determined Cq value (above right) depend on the PCR efficiency and the observed Cq value (Tuomi et al 2010). The Accept Error and Accept Bias radio buttons can be used to indicate which error or bias is acceptable. The critical Cq values that will be set are indicated by the dotted lines in the above graphs.

Because the cumulative data are higher, the Nq threshold is reached at an earlier cycle. This decrease in Cq, dubbed Cshift, is dependent on the PCR efficiency (Tuomi et al, 2010). Other monitoring chemistries depend on the incorporation of fluorescently labeled primers into the amplicon and thus show a lag in the occurrence of fluorescence and an increase in Cq. Both effects can also occur together. For each chemistry, this effect on the observed Cq can be corrected when the PCR efficiency is estimated (Ruijter et al., 2014, Microchimica Acta).

However, both the PCR efficiency estimation and the Cshift correction are only valid when the Cq value of the sample is not too low. Therefore the program checks the Cq value against the critical values shown in figure 22.

When the Cq of a sample is too low to allow a valid PCR efficiency to be estimated you can choose to exclude this sample from the calculation of the mean efficiency by checking the ***Exclude samples with too low Cq value*** box on the ***User Settings*** page (Fig. 22, bottom). Which error in PCR efficiency you are willing to accept because of a low Cq can be indicated on the ***User Settings*** page.

The Cq value can be too low to perform the right Cshift correction which will affect N0. In that case the program will assign a flag to the sample and display this flag in the ***Quality label*** column of the output. On the ***User Settings*** page (3.10) you can indicate which bias in N0 you are willing to accept before this flag is set.

3.10 User Settings

The User Settings page contains controls for exclusion of samples from calculation of the mean efficiency, relaxing the baseline setting requirements, reading baseline values from an earlier LinRegPCR analysis of the same data file and setting criteria for flagging samples when the Cq is too low.

Exclusion of samples from the mean efficiency.

The program offers four reasons to exclude samples from the calculation of the mean efficiency:

1. Samples without amplification: these are always excluded, the check cannot be removed.
2. Samples that do not reach the plateau phase.
3. Samples with a very low Cq value (only for probe based chemistries; 3.9).
4. Samples with a deviating individual PCR efficiency.

The user can choose to exclude samples because of criteria 2 and 4. Which choice is made depends on the number of samples per amplicon group (less samples gives a less valid mean PCR efficiency) and the variation between samples. Criterion 3 is described in section 3.9 (Fig. 22). When the DNA monitoring chemistry is not cumulative, low Cq values are no problem; because of the way LinRegPCR determines the baseline by reconstructing the log-linear phase it does not need early cycles with negligible amplicon presence to determine the fluorescence baseline.

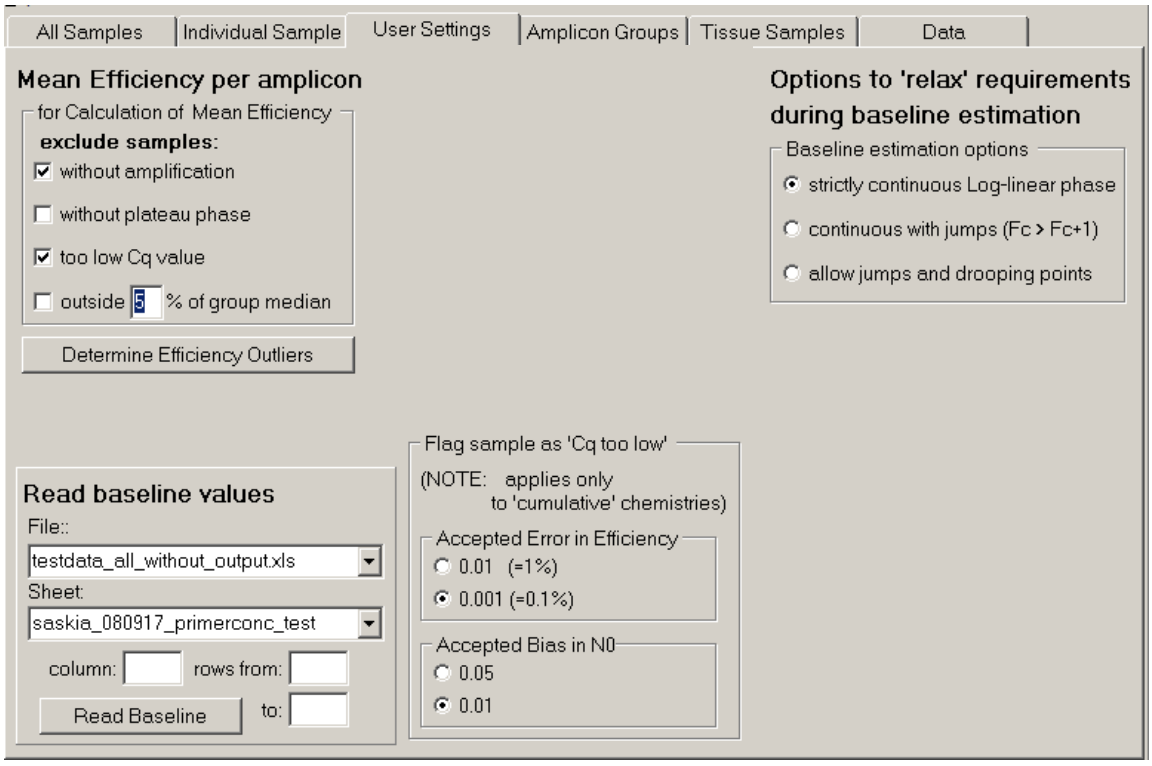


Figure 23: User Settings page with controls for inclusion or exclusion of samples from calculation of the mean efficiency, setting probe target per amplicon group, relaxing the baseline setting requirements, reading baseline values from an earlier LinRegPCR analysis of the same data file and setting criteria for flagging samples when the Cq is too low.

Read baseline values.

After analysis of a qPCR dataset with LinRegPCR, the results of this analysis, including the determined baseline values can be saved to Excel (3.12.2). When you want to reanalyze these data, for instance to define other amplicon groups, you do not have to run the baseline correction again. To read baseline values from Excel:

1. Open the ***Baseline*** menu
2. Choose '***Read Baseline values'***

The program will display the ***Read Baseline values*** control in the left bottom corner of the ***User Settings*** page.

3. Use the list box controls to choose the right ***Excel book*** and ***sheet.***
4. Give the ***column*** in which the baselines were saved.
5. Give the ***first*** and ***last row*** in which the baselines can be found.

Do NOT include a header row in the row specification.

6. Press '***Read Baselines'***.

The program will apply the baselines and proceed with setting a common W-o-L. After this step the situation is the same as after pressing the red '***determine baselines***' button.

Relaxing baseline requirements.
This control (in the right top corner of the page) has been described in the section on baseline estimation (3.3)

Flagging Cq too low.
The use of this control has already been described in section 3.9.

3.11 Tissue Samples

Quantitative PCR data that were read from an RDML file contain information on the tissue samples that were used as input material. Each of those tissue samples can be used as input for different reactions (wells) within a qPCR run. The tissue sample information is also required for the normalization of N0 values based on reference genes and for further statistical analysis. To fully benefit from the RDML format, tissue sample information should also be added to qPCR data read from Excel. The **Tissue Sample** page provides the controls to add this information.

On the **Tissue Sample** page the user can choose the following options (Fig. 24):

1. **one tissue sample**: all reactions are derived from the same tissue.
   When this is the case you are still advised to give the tissue description in the ***edit*** field and thus replace 'OneTissueEdit'. This will help you identify the results after you saved them to Excel or to RDML.
2. **manual identification:**
   When the sample names do not contain information on the tissue, you can assign the tissues manually.
   1. choose the option ***'manual identification'***
   2. press the ***edit tissues*** button

   This button puts the tissue grid in edit mode. You can now enter values into the **tissue** column of this grid. Activate a cell by ***double clicking*** it and then enter the identifier.

3. **extract tissue from part X of the sample name** (Fig. 24)
   When the sample name contains the tissue identifier in a fixed format, you can base the tissue assignment on the sample name.
   1. choose the option ***'extract tissue from***'
   2. enter the ***index*** of the part of the sample name that describes the tissue
   3. indicate whether you are counting from the ***front*** or the ***back*** of the sample name
   4. when the ***separator*** is not (only) a space, give the extra separator
   5. press ***'Extract Tissues'***

The program now extracts the tissue identifier. Browse through the grid to check the assignments and change the above entries when required. If you want you can switch to '***manual identification***' to correct the identifiers.

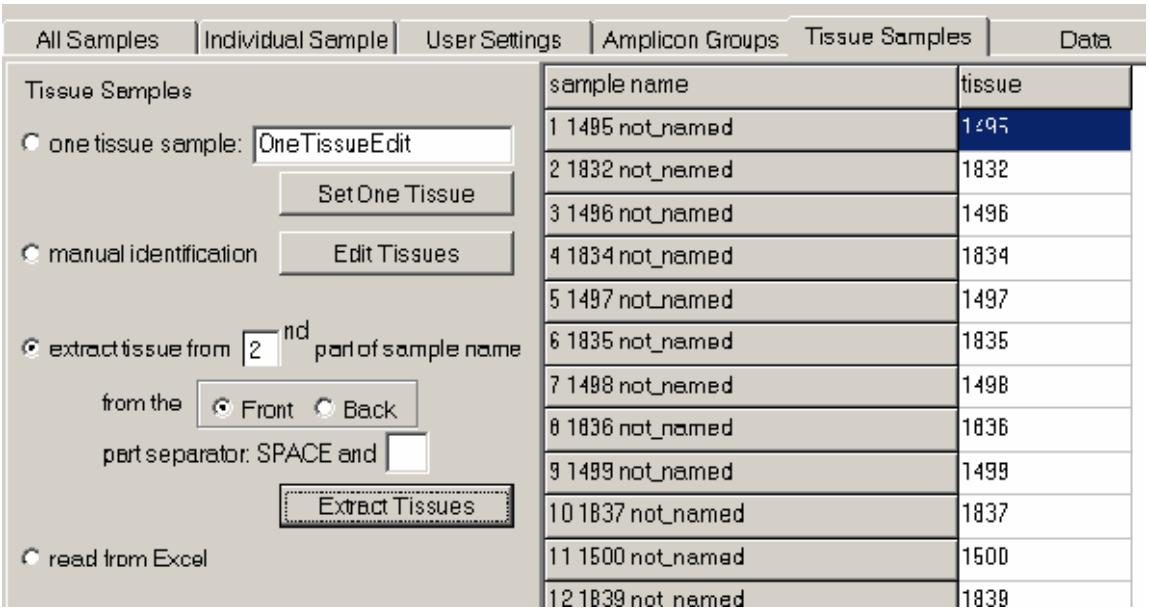


Figure 24: Tissue Sample page. The user can choose between four options to define the tissue sample that forms the input for each of the reactions. The figure shows how the tissue is extracted from the sample name as the second part from the name, separated by spaces.

4. read from Excel

Choosing this option will display the ***Read-from-Excel*** control (conform Fig. 15).

1. Use the list box controls to choose the right ***Excel book*** and ***sheet***.
2. Give the ***column*** with the tissue identifiers.
3. Give the ***first*** and ***last row*** in which the tissue identifiers are to be found.

Do NOT include a header row in the row specification. The row order should be the same as in the input of the fluorescence data.

4. Press '***Read Tissues'***.

The tissues will appear in the grid. Check them and if necessary correct them manually.

3.12 Save (data and) results

3.12.1 Save to Excel

Save options

You can decide for yourself which results you save to Excel on the ***Fit / Save*** Options page that you can open by choosing ***Options - Save Options*** from the main menu. By (un)checking the boxes in front of each parameter you determine which parameters are exported. Checking the ***all columns*** checkbox will give the complete output. Some columns cannot be unchecked because this would turn the output useless.

***NOTES*:**

- The choices you make will be saved to the Windows Registry and will be used as default the next time you use LinRegPCR. This ensures that the format of the output of LinRegPCR will be the same and that you can easily apply the analysis templates you made in Excel.

- The **N0 individual** value is only included for backward compatibility. The use of these results is not recommended.

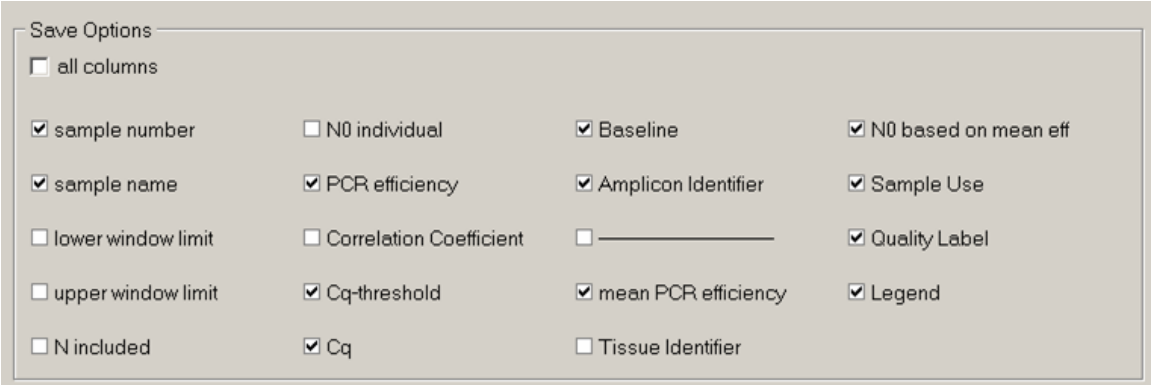


Figure 26: Save Options. By (un)checking the boxes you can decide which parameters have to be saved to Excel. Sample number, sample name, N0 based on mean efficiency, Sample use, Quality label and Legend cannot be unchecked and will always be exported.

Saving to Excel

To save the analysis results to an Excel spreadsheet, choose ***File - Save to Excel*** from the main menu. **Excel has to be running, 'Ready' and a book should be opened!**

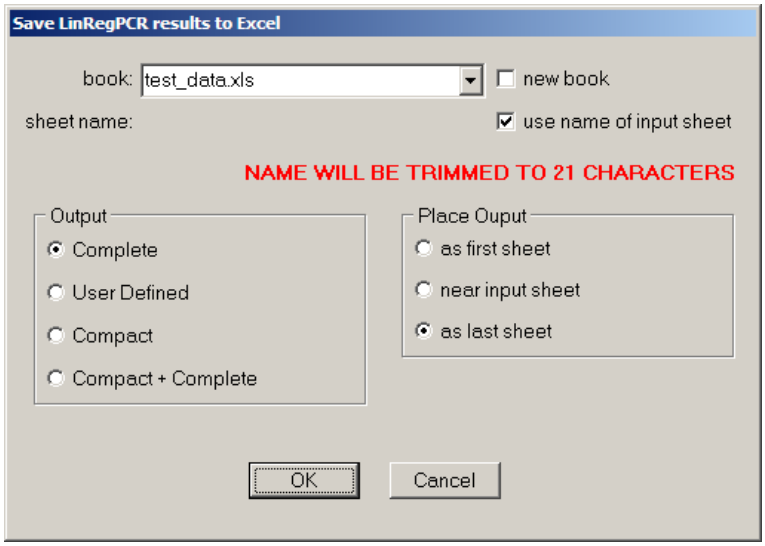


Figure 27: Save to Excel dialog.

In the ***Save to Excel*** dialog (Fig. 27):

1. Choose to save to an existing Excel ***book*** or to a new book. By default LinRegPCR suggests to save to the book you read the data from, which is the active book.
2. Give the name of the ***output sheet***.

As default, LinRegPCR uses the name of the input sheet extended with the kind of output and the date.

3. Choose the kind of output you want to save:
   - **Complete output**: all 19 columns, also the ones that are there for backward compatibility.
   - **User defined output**: you can select the output columns on the Fit / Save Options page.
   - **Compact + Complete output**: this is the recommended option when you start using LinRegPCR. The compact output page gives the columns required to further analysis of your data; the complete data is available for trouble shooting.

- **Compact**: The compact output page gives the columns required to further analysis of your data. You may miss the results required for trouble shooting

4. Choose the ***position*** of the output sheet(s) in the Excel book.
5. Press ***OK*** to save the results.

***NOTES*:**

- When during the analysis you have changed the Excel file that you read the data from, saving the results to this file may fail. In that case, saving to a new book will still be possible.
- You are not allowed to delete sheets from the active book in Excel while the ***Save-to-Excel*** dialog is open.
- When the output sheet name already exists in Excel, '_1' will be added to the name of the new file; '_2' when the '_1' is already present, etcetera.

3.12.2 Save to RDML

When you save results to RDML, LinRegPCR can will also save the raw fluorescence data that were read from an Excel file to this RDML file. When data were read from RDML, only the results of the analysis will be saved to the existing RDML file. The advantage of saving the results to RDML is that it allows normalization and further statistical analysis with RDML compliant software packages, like qBasePlus.

To save data to RDML:

1. From the ***File*** menu, choose '***Save to RDML'***

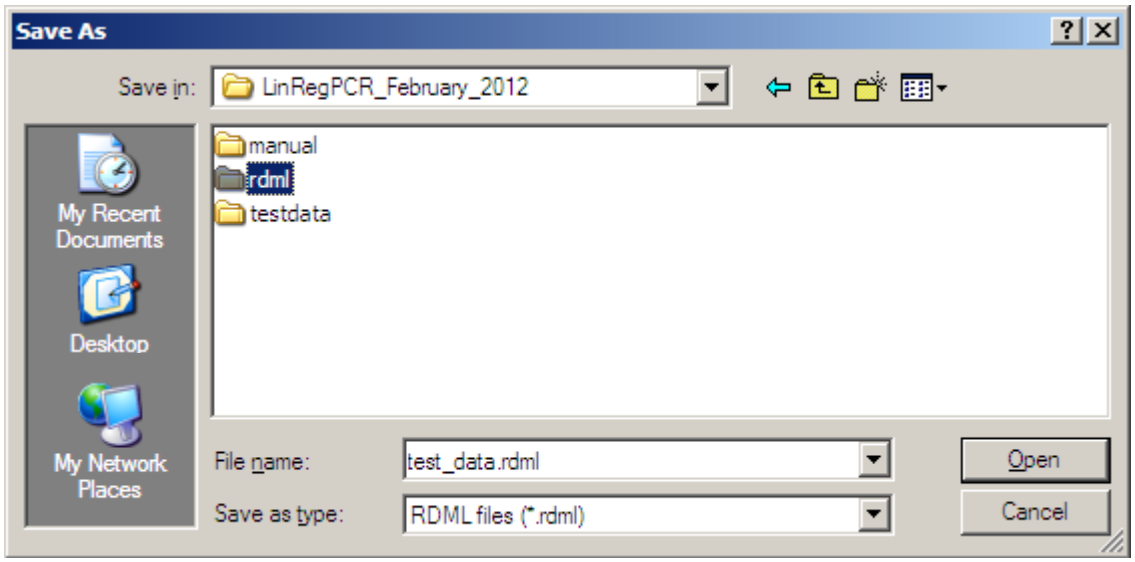


Figure 25: Save As RDML dialog. The user can browse the directory tree to determine the path and the name to save the RDML file.

2. In the ***Save*** dialog, choose or give the name of the RDML file
3. Press Open to save the dta and results to RDML

When you choose an existing RDML file you will be asked permission to overwrite this file.

Saving after RDML input.

When you had read the raw fluorescence data from an RDML file these data will not be overwritten. LinRegPCR only adds the PCR efficiency and quantification threshold per amplicon and the Cq value and baseline value per reaction to the existing RDML file. When

no Cq value could be determined the value -1 will be exported. Note that the quality control information (3.13) of LinRegPCR is not saved to the RDML file. When you want to make use of this information you will also have to save your results to Excel (3.12.1).

NOTES:

- LinRegPCR saves data into RDML version 1.1 format to maintain compatibility with downstream analysis programs. In case the imported RDML file was in RDML version 1.0 format, you will be notified when saving results to RDML. When you decide to continue, the suggested filename for the results file will be the name of the imported RDML file with the prefix 'Results_'. Do not overwrite your original RDML file because this will lead to data loss.
- The RDML format definition expects floating point values to be written with decimal points. When your input contained decimal commas, LinRegPCR will save the analysis results to a new file with decimal points. You will then be notified when saving results to RDML. When you decide to continue, the suggested filename for the results file will be the name of the imported RDML file with the prefix 'Results_'. Do not overwrite your original RDML file because this will lead to data loss.
- When you import an RDML 1.0 file, the result file will contain the Tissues and Target information, the Experiment and Run Identifiers, and the cycle and fluorescence data from the imported file, together with the analysis results as described above. Other data from the original RDML 1.0 file are not transferred to the RDML 1.1 results file.

Saving after Excel input.
When you had started with raw fluorescence values from Excel, you can use LinRegPCR to convert these data into RDML format. LinRegPCR will create a new RDML file containing the raw fluorescence data, as well as the amplicon group (3.4) and tissue sample (3.11) that you assigned during the LinRegPCR analysis of the data. Additionally the analysis results are saved as described above for **Saving after RDML input**.

3.13 Quality control output

The last three columns of the Excel output give information on the quality of the samples, as determined by LinRegPCR (Fig. 28). Other columns that play a role in quality control are the individual PCR efficiency column and the correlation coefficient column. The use of these parameters in quality control has been described in section 3.5.

The ***Sample_Use*** column tells you whether the sample is used for: 1: setting the window-of-linearity, 2: calculating the mean efficiency, and 3: whether a starting concentration (N0) is calculated. When either of these is not applicable the column contains a '0'.

The ***Quality_checks*** column contains 0 when the sample is OK. The other 9 positions can contain values of 1 thru 9 of which the explanation is given in the ***LEGEND*** column under **Quality Checks**. Based on these values you can decide to include or exclude the N0 value from further analysis.

Samples without amplification (**1**), usually also do not reach the plateau (**3**) and are excluded by the program (**6**). Similarly, samples that are assigned **baseline error** (**2**), usually have a deviating PCR efficiency (**5**) and are always excluded by the program (**6**), unless you decide to set a manual baseline (**9**) which results in a N0 value being calculated.
Samples that do not reach the plateau phase (**3**) OR have a deviating PCR efficiency (**5**) are not included in the calculation of the mean efficiency but a N0 value is calculated. However, when both conditions apply (**3 5**) the samples are rejected by the program (**6**) and receive a

N0 of ‘-999’. When you are using a cumulative monitoring chemistry two extra quality flags may appear, indicating that the observed Cq value was too low to determine a valid PCR efficiency (**A**) or to calculate a valid corrected Cq value (**B**).

At the bottom of the ***LEGEND*** column (not shown in figure 28) the choices that the user made with respect to the calculation of the mean PCR efficiency are displayed.

N0 = threshold /(Eff_mean^Cq)

N0	Sample_Use	Quality_checks
-9.99E+02	0 0 0	- 1 - 3 - 5 6 - - - - -
1.10E-10	0 0 3	- - - 3 - - 6 - - - - -
4.29E-10	0 0 3	- - - 3 - - 6 - - - - -
1.55E-09	0 0 3	- - - 3 - - 6 - - - - -
1.95E-09	0 0 3	- - - 3 - - 6 - - - - -
-9.99E+02	0 0 0	- 1 - 3 - 5 6 - - - - -
5.25E-09	1 2 3	0 - - - - - - - - - - -
3.58E-09	1 2 3	0 - - - - - - - - - - -
-9.99E+02	0 0 3	- - 2 - 4 - 6 - - - - -
1.58E-08	1 2 3	- - 2 - 4 - 6 - - 9 - -
2.72E-08	1 2 3	0 - - - - - - - - - - -
2.74E-10	0 0 3	- - - 3 - - 6 - - - - -
8.38E-09	1 2 3	- - - - - 5 - - - - - -
1.11E-08	1 2 3	0 - - - - - - - - - - -
1.84E-10	0 0 3	- - - 3 - - 6 - - - - -
5.46E-09	1 2 3	0 - - - - - - - - - - -
2.79E-10	0 0 3	- - - 3 - - 6 - - - - - - -
9.45E-10	1 0 3	- - - - - 5 6 - - - - - - -
4.82E-10	0 0 3	- - - 3 - - 6 - - - - - - -
-9.99E+022	0 0 0	- 1 - 3 - 5 6 - - - - - - -

LEGEND
Sample Use:
1: used for W-o-L setting
2: contributes to mean PCR efficiency
3: N0 value calculated
0: not used / calculated

Quality Checks:
0: passed all checks
1: no amplification
2: baseline error
3: no plateau
4: noisy sample
5: PCR efficiency outside 5%
6: excluded from mean Eff
7: excluded by user
8: included by user
9: manual baseline

for cumulative fluorescence only
A: Cq too low for Eff
B: Cq too low for N0

Figure 28: Example of the last 4 columns of the output to Excel showing the quality check output

When no starting concentration is calculated the **N0** or **N0_(mean eff)** column will contain '-999'. When you do not agree with this ‘decision of the program’ you can overrule this by calculating the N0 value in Excel. You can calculate the N0 by hand by using the equation above the **N0** column. All required data are present in the compact and the complete output. Make sure you have the **Nq-threshold**, **Cq-value** and **mean PCR efficiency** columns in your user-defined output (3.12.1; Save Options) to be able to calculate N0 manually.

Appendix. Implemented formats for Excel

NOTE:

Each format contains a pre-defined number of leading columns and/or header rows. Some rows and columns are ignored by the program. However, header rows and leading columns are also used to define samples names. These columns and rows have to be given in the range you enter in the columns and rows edit fields of the ***Read-from-Excel*** dialog. When you do not include them, sample names will be based on the fluorescence data which will thus be lost from the analysis. When there are spaces in the entries in columns or rows that are used to define sample names, those spaces will be deleted.

Also note that these formats may already be different from the ones that are generated by the current versions of the software of the PCR apparatus. In that case, try to find a format that fits to the exported data or contact us.

Roche LigthCycler (32 capillaries)

This format consists of 10 header rows that are ignored by LinRegPCR and two columns per sample starting with a two row header followed by cycle number - fluorescence value pairs. The range you have to give in the columns and rows edit fields is inclusive the two header rows per sample. For a 40 cycle PCR and the full 32 samples the range would be: columns A thru BL and rows 11 thru 52. The entries in the left cell of first header row per sample will be used as sample name; the right cell and the second header row will be ignored. The numbers in the first column per sample are used as cycle numbers.

Applied Biosystems

This format consists of one header row (with cycle numbers) and two leading columns around a data block that consists of one sample per row, one cycle per column. The whole range has to be given in the columns and rows edit fields. For a 40 cycle PCR and 72 samples the range would be: columns A thru AP and rows 1 thru 73. The two leading columns will be combined to one sample name, the entries of the columns being separated by a space. The header row will be converted into integer cycle numbers.

MJ Research

This format consists of one header row (with sample identification) and three leading columns around a data block that consists of one sample per column and one cycle per row. For a 40 cycle PCR and 60 samples the range would be: columns A thru BK and rows 1 thru 41. The entries in the header row are used as sample name. The entries in three leading columns are ignored and the row number in the data block is used as integer cycle number.

Bio-Rad iCycler

This format consists of two header rows and one leading column around a block of data that consists of one sample per column and one cycle per row. The leading column contains fractional cycle numbers, the first header row is empty except for the word ‘well’ in cell A1 and the second header row contains the well identification. Both header rows have to be included in the range given in the columns and rows edit fields. For a 45 cycle PCR and 30 samples the range would be: columns A thru AE and rows 1 thru 47. The entries in first header row are ignored by the program; those in the second row are used as sample names. The leading column is converted into fractional cycle numbers.

Roche LightCycler 480 (converted raw data: download: http://LC480Conversion.hfrc.nl). This format is the same as the Applied Biosystems format (see above).

Stratagene Format 1 (Vertically Grouped by Sample)
This format consists of two header rows (‘Amplification Plots’ and an empty row) followed by a data block per sample This block starts with a sample header row with the entries sample name, Cycles and Fluorescence (Rn) followed by one row per cycle. The total input block is therefore 2 + Nsamples * (Ncycles +1) rows long and 3 columns wide. For a 40 cycle PCR with 66 samples the range would be: columns A thru C and rows 1 thru 2708 {=2 + 66 *(40 +1)}. The two header rows are ignored, the entry in column A of the sample header row is used as sample name and the entries in column B are used as integer cycle numbers.

Step-One Plus (ABI)
This format consists of an information block of 6 lines, an empty line followed by one header row above a data block of three columns with one row per cycle per sample. Empty, unused, wells are included in the output but the Rn value is omitted for those wells. When only the well, cycle and Rn columns are included in the output, and a 96-wells plate is used, for a 50 cycle PCR the range to enter in the edit fields would be: columns A thru D and rows 8 thru 4808. The entries in the header row are used to determine the cycle and fluorescence (Rn) columns. Columns labeled ‘well’ and ‘target name’ are used to form the sample names. Other columns are ignored. The entries in the cycle and Rn columns are used as cycles and fluorescence values, respectively. The empty wells are removed. When LinRegPCR cannot find a cycle or Rn column it will ask you to provide the column numbers.

LightCycler 480, 2 columns per well
This format consists of one header rows (with, alternating, ‘X’ and sample ID) and one leading column (with row numbers) around a data block that consists of one well per two columns with cycle - fluorescence value pairs per row. Empty, unused wells are included in the data block. The data range for a 40 cycle PCR run on a 96-wells plate should be given as columns A thru GK and rows 1 thru 41. The leading column is ignored and the sample ID entries in the header row are used as sample name. The default sample ID entry (‘Sample’) is used to remove empty wells. Therefore this default entry should not be used in the sample name. Note that in most versions of Excel a 384 wells plate cannot be exported in this format due to the 256 column restriction in those Excel versions.

Rotor-Gene (Corbett Research)
Rotor-Gene can export data for LinRegPCR. This format consists of an information block of 10 lines, an empty line and a data block consisting of two header rows (with cycle numbers and ‘background’) and two leading columns (the first with sample IDs, the second empty) around a data block that consists of one sample per row, one cycle per column. The range of the data block, including the header rows and leading columns, has to be given in the columns and rows edit fields. For a 40 cycle PCR and 15 samples the range would be: columns A thru AP and rows 12 thru 28. The first leading column is used as the sample name. The header rows are ignored and the cycle number is derived from the column number.

Eppendorf Realplex
The format consists of a header of 3 or 4 lines followed by a header row starting with 6 column headers (Position, Type, Sample, Name, Amount, Target, Dye) followed by the cycle numbers. This row, and the 6 leading columns, should be included in the range of the data block. A PCR run with 45 samples and 40 cycles is given as: columns A thru AT and rows 4 thru 49, when the word ‘Position’ is in cell A4. The 6 leading columns are combined to one sample name (separated by ‘_”).The header row is ignored and the cycle number is derived from the column number.

Applied Biosystems (5 leading columns)
The format consists of a header of 10 lines followed by a header row starting with 6 column headers (Well, Sample, Name, Detector, Task, Reporter, and Rn) followed by the cycle numbers. The header row, and the 6 leading columns, should be included in the range of the data block. A PCR run with 72 samples and 40 cycles is given as: columns A thru AT and rows 11 thru 83, when the word 'Well' is in cell A11. The 6 leading columns are combined to one sample name (separated by '_"). Note that the Rn column is empty but should be included. The header row is ignored and the cycle number is derived from the column number.

Illumina Eco
The "Component data" format of the Eco software consists of a header of 7 lines and an empty line followed by a header row starting with 3 column headers (Well, Exclude, Dye) followed by the cycle numbers. This header row, and the 3 leading columns, should be included in the range of the data block. A PCR run with 48 samples and 40 cycles is given as: columns A thru AQ and rows 9 thru 57, when the word 'Well' is in cell A9. The first and third leading column are combined to the sample name (separated by a space). When a sample is flagged to be excluded in the Eco software, or when a well is empty, no fluorescence data appear in the export and the well will be ignored by LinRegPCR.

Piko Real (Thermo Scientific)
Export the raw data from the Piko Real apparatus, resulting in an Aquired_dat_1 sheet in Excel. The export format consists of a 10 row header, followed by a data block in which each well is in a column and each cycle in a row. In the Read-from-Excel dialog, omit the 10 lines header and read only the range of data, including the header row with well identifiers and the 3 leading columns, which are labeled Channel, Cycle and Temp. A qPCR experiment with 24 reactions and 40 cycles is then loaded by giving the columns A thru AA and rows 11 thru 51.

Trouble shooting:

- Make sure the decimal separator in the export of your qPCR system, the Windows version on your computer and your Excel program are the same. Excel should be set to ***'use system separators'*** (in Excel go to **Tools - Options - International**). When your qPCR system exports the wrong separator in a text file you have to do a ***replace*** in a text editor. An Excel version that expects decimal points will read numbers with decimal commas like 3,9; 3,87; 3,868 as 39; 387; 3686 creating a 100 fold difference where there was only a rounding-off error.
- Excel has to be running, and the book (spreadsheet) with your data should be opened.
- Make sure Excel is in **'Ready'** mode (in the left bottom corner of Excel). When you are editing a cell that has to be read by LinRegPCR, importing the data into LinRegPCR will fail. LinRegPCR will then display the message "Call was rejected by callee ".
- Some sheet names are not compatible with LinRegPCR. When a sheet does not turn up in the drop down list: cancel the ***Read-from-Excel*** dialog, go to Excel, rename the sheet and try again.
- Excel Chart sheets are incompatible with LinRegPCR and disrupt the automatic recognition of the sheets.
- LinRegPCR expects that Excel uses the same decimal separator as Windows. To make sure this is the case: in Excel go to ***Tools - Options***, choose ***International*** and set Excel to ***Use system separators***.

- Sometimes the book or sheet does not show up in the list boxes. In that case check for typing errors in the file extension (*.TXT or *.CSV). You can almost always solve this problem by saving your data file as an Excel-workbook (*.XLS).
- Note that Excel may have ‘hidden’ sheets that it uses for storage of macros and functions. These sheets may turn up in the sheets-listbox. Do not choose one of those sheets. the program will crash.