RT^2 Profiler PCR Array Data Analysis v3.5 Handbook

For analyzing gene expression data from RT^2 Profiler PCR Arrays

QIAGEN®

QIAGEN Sample and Assay Technologies

QIAGEN is the leading provider of innovative sample and assay technologies, enabling the isolation and detection of contents of any biological sample. Our advanced, high-quality products and services ensure success from sample to result.

QIAGEN sets standards in:

■ Purification of DNA, RNA, and proteins

■ Nucleic acid and protein assays

■ microRNA research and RNAi

■ Automation of sample and assay technologies

Our mission is to enable you to achieve outstanding success and breakthroughs. For more information, visit www.qiagen.com.

Contents

Intended Use

The RT^2 Profiler PCR Array Data Analysis Webportal is for molecular biology applications. This product is not intended for the diagnosis, prevention, or treatment of a disease.

All due care and attention should be exercised in the handling of the products. We recommend all users of QIAGEN products to adhere to the NIH guidelines that have been developed for recombinant DNA experiments, or to other applicable guidelines.

Safety Information

When working with chemicals, always wear a suitable lab coat, disposable gloves, and protective goggles. For more information, please consult the appropriate safety data sheets (SDSs). These are available online in convenient and compact PDF format at www.qiagen.com/safety where you can find, view, and print the SDS for each QIAGEN kit and kit component.

24-hour emergency information

Chemical emergency or accident assistance is available 24 hours a day from:

CHEMTREC

USA & Canada ▪ Tel: 1-800-424-9300

Outside USA & Canada ▪ Tel: +1-703-527-3887 (collect calls accepted)

Introduction

This document contains a quick reference guide for experienced users (page 6) and a detailed guide (page 11) for use of the RT^2 Profiler PCR Array Data Analysis Webportal. The detailed guide contains thorough explanations of the formulas used to generate the results and figures within the RT^2 Profiler Data Analysis Webportal. This document does not provide bench protocols. To obtain bench protocols, please refer to the appropriate user manual or handbook. If additional assistance is required, please contact a local QIAGEN technical representative.

Important Notes

■ Please note that all work must be completed with the RT^2 Profiler PCR Array Data Analysis Webportal in the same session. The data is not stored on a server, so all work is lost once the session, or the web browser, is closed. Be sure to export all processed data and results to a Microsoft® Excel® file saved on a local computer or server.

■ Please set the screen resolution to 1024×768 or greater, if possible.

■ Turn off any window pop-up blockers. Some features of the software will launch separate windows.

■ Data to be uploaded must be saved as an Excel 1997–2003 file. Files with *.xlsx are not compatible with the tool at this time.

Quick Reference Guide

This guide is intended for experienced users of the RT^2 Profiler PCR Array Data Analysis Webportal only. New users should reference the Detailed Reference Guide on page 11.

Procedure: Setting baseline and threshold on the real-time thermocycler

The RT^2 Profiler PCR Array Data Analysis Webportal analyzes threshold cycle (C_T) values to calculate changes in gene expression. The C_T values must be exported from the qPCR instrument and formatted into an Excel spreadsheet. All of the plates/wells must be copied into a single worksheet.

1. **Calculate the C_T value for each well using the real-time cycler software, as described in the following steps.**
    Note: If using the Roche® LightCycler® 480, there are 2 options for data analysis: using the second derivative maximum setting (in this case, there is no need to calculate the C_P) or using "Fit Points" (in this case, the C_P should be defined manually as described in step 3).

2. **Define the baseline by choosing the automated baseline option if the cycler has the adaptive baseline function. If the cycler does not have the adaptive baseline function, set the baseline manually.**
    To set the baseline manually, use the linear view of the amplification plots to determine the earliest visible amplification. Set the cycler to use the readings from cycle number 2 through 2 cycles before the earliest visible amplification, but not past cycle 15. The earliest amplification will usually be visible between cycles 14 and 18.

3. **Manually define the threshold by using the log view of the amplification plots.**
    Choose a threshold value above the background signal but within the lower one-third to the lower one-half of the linear phase of the amplification plot. Ensure that the threshold values are the same across all RT^2 Profiler PCR Array runs in the same analysis and yield C_T^{PPC} values in the range of 19 ± 3 for all arrays or samples.

4. **Export or copy the C_T values from each experimental run into an Excel spreadsheet. C_T values for each sample should be in a single column.**

Procedure: Formatting data for the RT^2 Profiler PCR Array Data Analysis Webportal

The RT^2 Profiler PCR Array Data Analysis Webportal analyzes C_T values to calculate changes in gene expression. For the RT^2 Profiler PCR Array Data Analysis Webportal to correctly read and import data, the C_T values must be exported from the qPCR instrument and then formatted into a new Excel spreadsheet. The RT^2 Profiler PCR Array Data Analysis Webportal only accepts Excel spreadsheets for analysis. All of the plates and wells must be copied into a single worksheet. Please use the Excel files available for download from the "Upload Data" page of the Webportal as guides.

If using a cataloged PCR array where 1 sample was tested on a 96-well plate, 100-well ring, or 384-well plate (the 1x384 layout):

The raw C_T values should be in a single column that lines up with the correct well location on the PCR array. The first two rows should contain sample names (required) and group names (optional). A sample spreadsheet is available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal.

If using a 384-well plate in the 4x96 layout:

Separate each sample's data into a different column by using the "Universal Custom PCR Array Patch" available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal.

If using a custom PCR array:

The raw C_T values for each sample should be in a single column that lines up with the correct well location (column A) on the PCR array. The gene tested in each well and its assay catalog number must be supplied by the user in columns B and C, respectively. The first two rows should contain sample names (required) and group names (optional). A sample spreadsheet is available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal.

A patch is required to format the raw C_T values from Custom RT^2 Profiler PCR Arrays that characterize multiple samples with shorter gene lists on the same plate. A "Universal Custom PCR Array Patch" is also available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal.

If analyzing data from individual qPCR gene expression assays:

The raw C_T values for each sample should be in a single column that lines up with the correct position (column A). The gene tested in each well and its assay catalog number must be supplied by the user in columns B and C, respectively. The first two rows should contain sample names (required) and group names (optional). A sample spreadsheet is available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal.

Procedure: Upload data

Before arriving at the RT^2 Profiler PCR Array Data Analysis Webportal, the QIAGEN product format (Standard/Cataloged RT^2 Profiler PCR Array, Custom RT^2 Profiler PCR Array, or Single or Multi-Gene RT^2 qPCR Primer Assays) and the plate format (96- or 384-well plate format or Rotor-Gene Q) used to perform the experiment should already have been defined. If Standard/Cataloged or Custom RT^2 Profiler PCR Arrays were used, the catalog number of the array should also have been entered.

1. **Once in the RT^2 Profiler PCR Array Data Analysis Webportal at the "Upload Data" page, click the "Browse" button to select the MS Excel file prepared in the above section ("Procedure: Formatting data for the Data Analysis Webportal") that contains all PCR data with a maximum number of 100 samples.**
2. **Click the "Upload" button.**
   Note: Data to be uploaded **must** be saved as Excel version 1997–2003. Data cannot be saved as an "*.xlsx" file.

Procedure: Analysis setup: organization and quality-check of uploaded data

1. **The "Uploaded Data" page displays the raw C$_T$ values, sample name, and defined groups. Confirm that data was uploaded correctly.**
2. **On the "Sample Manager" page, assign "Samples" to their different "Groups," with at least one "Control Group" and at least one "Group 1," or choose to "Exclude Sample" using the dropdown menu for each sample. If RT^2 PreAMP Primer mix was used, click appropriately labeled checkbox. Click the "Update" button when finished.**
3. **Review the "Data QC" page to ensure that each sample has passed the PCR Array Reproducibility, RT (reverse transcription) Efficiency, and Genomic DNA Contamination quality controls.**
4. **On the "Select Normalization Factor" page, define the desired normalization method and housekeeping genes to be used for data normalization; use the appropriate dropdown menu and checkbox selections, and click the "Perform Normalization" button.** This must be performed before proceeding to "Analysis."
5. **Review the "Data Overview" page to see each group's distribution of threshold cycle values and the average raw data in each group.**

Procedure: "Analysis" and "Plots & charts"

1. **Review the "Fold Regulation," "Fold Change," "Average ΔC$_T$," and "2^– ΔC$_T$" pages for the final results and intermediate calculations processed by the software from the inputted data.**

Procedure: Plots and charts

After modifying any parameters on any plot or chart, click the "Update" button to see the new image.

For scatter plots:
1. Define "Groups" to be compared. Choose the fold-change boundary (threshold) of interest.
2. Mouse over the right-hand table entries to highlight symbols on the plot.
3. Click check boxes to remove genes from or add genes to plot.
4. Click the "Update" button when all parameters set to refresh the plot.
5. Click the "Export Data" button to download the results as an Excel file.
6. To save the figure, mouse over and click on the options available from the top right icon.

For volcano plots:
1. Define "Groups" to be compared. Choose the fold-change boundary and p-value thresholds.
2. Mouse over the right-hand table entries to highlight symbols on the plot.
3. Click check boxes to remove genes from or add genes to plot.
4. Click the "Update" button when all parameters set to refresh the plot.
5. Click the "Export Data" button to download the results as an Excel file.
6. To save the figure, mouse over and click the options available from the top right icon.

For clustergrams:
    Note: All default settings are the recommended settings.
1. Choose to cluster "Sample" by "Array" or "Group."
2. Select the "1-D" or "2-D" "Dimension" option to cluster by genes only or by genes and samples, respectively.
3. Select "Join Type" to define the linkage method.
4. Select whether the graph should be "Color coded" by using "Genes," "Samples," or the "Entire Dataset" as the average.
5. Choose whether or not the gene symbols and array or group names are "Item names visible."
6. Click check boxes to remove genes from or add genes to plot.
7. Click the "Update" button when all parameters set to refresh the plot.
8. To save the figure:
    a. Using Windows: Right-click on the figure and save the image/picture.
    b. Using OS X: Hold down the Control key and click, and save the image/picture.

Procedure: Export data

1. **Select the desired data tables using the check boxes and click the "Export" button to download an Excel file with each data table in a different worksheet.**

    Note: The software does not automatically save any results. Please be sure to export and save any figures or tables from each session.

What's next

Follow the instructions on each of the available modules to identify genes and reagents needed for the next best experiment to follow up the analyzed gene expression analysis, with the help of the latest bioinformatics tools. Please note that not all "What's next" modules are available for all species. The following modules are available:

1. **Gene Expression:** Screen or verify the target differentially expressed genes in more samples with individual RT^2 qPCR Primer Assays or a Custom RT^2 Profiler PCR Array.

2. **miRNA Regulation:** Analyze the expression or function of miRNAs predicted to regulate the target differentially expressed genes with miScript$^®$ miRNA qPCR Assays, Custom PCR Arrays, Mimics, or Inhibitors.

3. **Methylation:** Determine if promoter DNA methylation might regulate the expression of the target differentially expressed genes with EpiTect$^®$ Methyl II qPCR Assays or Custom qPCR Arrays.

4. **Transcription Factor / Histone:** Determine if differential transcription factor or modified histone binding might regulate the expression of the target differentially expressed genes with EpiTect ChIP qPCR Assays or Custom qPCR Arrays.

5. **siRNA:** Analyze the function of the target differentially expressed genes with FlexiTube or FlexiPlate siRNA.

6. **Protein Detection:** Determine if cytokine and chemokine gene expression changes results in protein expression level changes with Single Analyte ELISArrays.

7. **Somatic Mutation**: Determine if the target differentially expressed genes contain mutations that might affect their function with qBiomarker Somatic Mutation qPCR Assays or Custom PCR Arrays.

Detailed Reference Guide
Setting baseline and threshold on the real-time thermocycler

For all instruments except the Roche LightCycler:

1. **Set the baseline.** During initial cycles of PCR (e.g. cycles 3–15) the level of background fluorescence is gathered and used to subtract from and correct the fluorescent signal resulting from the increase in PCR product. The baseline may be collected manually and applied uniformly to all wells in the plate or collected and applied to each individual well ("Adaptive Baseline").

   a. **Use "Automated Baseline" if the instrument has an "Adaptive Baseline" function.**
   b. **Use "Manual Baseline" if there is no "Adaptive Baseline" function.**
       i. **Use "Linear View."**
       ii. **Set "Start Cycle" to 2 or 3.**
       iii. **Set "End Cycle" to up to 1 or 2 cycles before earliest amplification (with highest cycle being cycle 15).** A general suitable baseline range is from cycle 3 to 15.

2. **Set the threshold.** The C_T value is defined as the cycle where a statistically significant increase in fluorescence above the background signal is detected. The C_T value is inversely proportional to the relative amount of template. Specifically, the higher the C_T value, the lower the relative amount of template. The threshold value should be set above background signal, but within the lower one third to one half of the linear phase of the amplification plot (see Figure 1). It is imperative that the threshold value is set to the exact same value for each array plate.
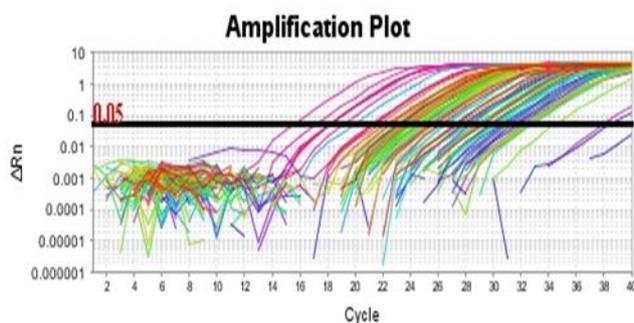

Figure 1. An example of proper threshold value placement. This absolute value will vary based on the qPCR instrument.

Note: Ensure that the threshold values are the same across all RT^2 Profiler PCR Array runs in the same analysis. The absolute position of the threshold is less important than its consistent position across arrays. If the RNA sample is of sufficient quality, the cycling program has been carried out correctly, and threshold values have been defined correctly, the value of C_T^{PPC} should be in the range of 19±3 for all arrays or samples.

3. **Export results.** After the baseline and threshold are set, export data into an Excel-compatible format and save the file.

   Note: Prior to exporting, check the exporting options available with the instrument. **De-select** the "Group by Replicates" option, if this is available. Alternatively, be sure to select the option that will export the data in order of the well numbers; e.g. 1–96 (A1, A2, A3, A4…H12), or 1–384 (A1, A2, A3, A4….P24).

For Roche LightCycler:
For data analysis using the LightCycler 480 software, the use of the **"Second Derivative Maximum"** analysis method available in this software is highly recommended. This algorithm is based on the kinetics of PCR. This method identifies the **crossing point (C$_P$)** of PCR as the point where the reaction's fluorescence reaches the maximum of the second derivative of the amplification curve, which corresponds to the point where the acceleration of the fluorescence signal is at its maximum. Hence, this crossing point should always be located in the middle of the log-linear portion of the PCR amplification plot. The advantage of this analysis method is that it requires little user input.

1. **Obtaining the "Crossing Point (C$_P$)" data values from the Roche LightCycler:**
   a. In the **Navigator** window, find and open the run file to be analyzed.
   b. Click the **"Analysis"** button on the left module bar. The **Analysis Overview** window will open.
   c. Click **"Abs Quant/2nd Derivative Max"** in the **Create new analysis** list. The **Create new analysis** dialog box will then open. Make sure the **"Analysis Type"** is **"Abs Quant/2nd Derivative Max."** Click to confirm the selection.
   d. The **Analysis Window** with the amplification curves will appear. To proceed, click the down arrow button on the lower right corner of the screen to set the analysis to "**High Confidence.**"
   e. Click the "**Calculate**" button to generate results. The calculation may take a few minutes.

2. **Exporting the C_P values from Roche LightCycler:**
   a. Once the C_P data is generated, move the mouse over to the result window and right-click to highlight the result table.
   b. Click the **"Export"** button to export the table data. A **Save table data** dialog box will appear.
   c. Enter a filename to save the data as a text file.
   d. Click the button on the right panel to save the run file now containing the C_P data.

3. **Melt curve analysis on the Roche LightCycler:**
   a. Proceed with the melt curve analysis by clicking the button located on the right of the **Analyses** line to add a new analysis. The **Create new analysis** dialog box will open.
   b. Select **"Tm calling"** for the **Analysis Type**. Then click to confirm the selection.
   c. The **Analysis** window with the raw melting curves will appear. To proceed, click the down arrow button on the lower right corner of the screen to set the analysis to **"SYBR Green I Format."**
   d. Click the **"Calculate"** button to generate results. The calculation may take several minutes.
   e. Once the calculated derivatives of the melting curves appear, check the specificity of each assay by ensuring the presence of a single melt peak in each well.
   f. Click the button on the right panel to save the run file now also containing the melt curve analysis.

Formatting data for the RT² Profiler PCR Array Data Analysis Webportal

The RT² Profiler PCR Array Data Analysis Webportal analyzes C_T values to calculate changes in gene expression. For the RT² Profiler PCR Array Data Analysis Webportal to correctly import and read data, the C_T values must be exported from the qPCR instrument and then formatted into a new Excel spreadsheet, version 1997–2003. The RT² Profiler PCR Array Data Analysis Webportal only accepts Excel spreadsheets for analysis. The data from all plates and wells to be analyzed at the same time must be copied into a single worksheet. Please use the templates described below as guides.

For all plate formats, start by exporting the data from the real-time instrument into an Excel-compatible file. After exporting, check the exported file to ensure that the C_T values are arranged in the correct order as shown in Figure 2, below.

Correct			Incorrect	
Well	**C_T**		**Well**	**C_T**
A01	33.4		A1	33.4
A02	21.2		A10	27.2
A03	26.7		A11	23.5
…			…	
H12	22.0		H9	34.4
…			…	
P24	19.2		P9	37.2

Figure 2. Correct versus incorrect formatting for C_T values in the exported data file.

For cataloged PCR arrays, 1x96 96-well plate, 100-well ring, or 384HC

If using a cataloged PCR array where 1 sample was tested on a 96-well plate, 100-well ring, or 384-well plate (HC content: 1x384 layout), the raw C_T values from one plate or array (each sample) must be in a single column lined up with the correct well location on the PCR array. A sample spreadsheet is available on the "Upload Data" page of the RT[2] Profiler PCR Array Data Analysis Webportal. Download this template and save to a local hard drive.

1. **Delete example data from the template file. Do not delete well numbers.**
2. **Copy and paste the data from the file exported from the instrument into the template.**
   Each column of data represents one sample from one array.
3. **Above the raw data, add Sample Names in first row (Row A). Optionally, add Group designations in second row (Row B).**
   The Group names must precisely read "Control Group" ,"Test Group 1", "Test Group 2", etc. If defining Groups in the software after uploading the data is preferred, simply leave the second row (Row B) blank, and see below. Delete any example Sample and Group names and data (extra columns) that do not apply to the experiment.
4. **Save the file as Excel version 1997–-2003 and use this file to upload data to PCR Array Web-based Data Analysis Portal.**
   Note: See Figure 3 below for an example of a correctly formatted upload file for a cataloged PCR array (based on 96-gene content).

	Resting 6 h	6 h Stimulation	Resting 24 h	24 h Stimulation
	Control Group	**Test Group 1**	**Test Group 2**	**Test Group 3**
A01	29.08	29.89	28.86	27.39
A02	32.02	31.15	33.33	31.89
A03	33.83	31.57	35.7	37.93
A04	33.95	31.3	33.05	29.21
A05	Undetermined	Undetermined	Undetermined	Undetermined
…	…	…	…	…
H12	19.9	20.65	20.09	19.85

Figure 3. An example of a correctly formatted upload file for a cataloged PCR array.

For cataloged PCR arrays, 4x96 on a 384-well plate

If using a cataloged array with 96 assays in a 384-well plate (4x96 layout), each sample's data will need to be reorganized into 4 individual columns and pasted into the data analysis upload template. This operation may be accomplished using the "Universal Custom PCR Array Patch" available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal. The patch will organize each plate into 4 separate samples, as shown in Figure 4.

384-Well Format	Custom Gene List			96-Assay Format Equivalent			
C$_T$ Values	Assay #	Gene Symbol	Assay Catalog #	Sample 1	Sample 2	Sample 3	Sample 4
A01	1			A01	A02	B01	B02
A02	2			A03	A04	B03	B04
A03	3			A05	A06	B05	B06
A04	4			A07	A08	B07	B08
…	…	…	…	…	…	…	…
D24	96			O23	O24	P23	P24
…							
P24							

Figure 4. The "384.96x4" worksheet of the Universal Custom PCR Array Patch Tool for 4x96 384-well.

1. **Copy the C$_T$ values exported from the real-time PCR instrument.**
2. **Paste the values into the yellow cells starting at cell A05 of the patch.**
   The patch will automatically convert the 384-well format of the data into 4 separate columns representing each sample.
3. **For Web-based Data Analysis, copy cells E9 through H100, and paste (using "Paste Special" and choose "Values") to a cataloged PCR array upload template.**
   Proceed as described above for the cataloged PCR array upload template.
4. **Save the file as Excel version 1997–2003 and use this file to upload data to the RT^2 Profiler PCR Array Web-based Data Analysis Portal.**
   Note: See Figure 5 below for an example of a correctly formatted upload file for a cataloged PCR array.

	Resting 6 h	6 h Stimulation	Resting 24 h	24 h Stimulation
	Control Group	Test Group 1	Test Group 2	Test Group 3
A01	29.08	29.89	28.86	27.39
A02	32.02	31.15	33.33	31.89
A03	33.83	31.57	35.7	37.93
A04	33.95	31.3	33.05	29.21
A05	Undetermined	Undetermined	Undetermined	Undetermined
…	…	…	…	…
H12	19.9	20.65	20.09	19.85

Figure 5. An example of a correctly formatted upload file for a cataloged PCR array.

For custom PCR arrays

For custom PCR arrays, the raw C_T values for each sample should be in a single column that lines up with the correct well location (column A) on the PCR array. The gene tested in each well must be supplied by the user in column B, and the RT^2 qPCR Primer Assay catalog number for each gene must be supplied by the user in column C. All information of this information may be obtained from the custom PCR array sign-off form used during ordering. A sample spreadsheet is available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal.

Additionally, if using a custom array (96- or 384-well formats) that analyzes shorter lists of genes in multiple samples on the same plate, each sample's data will need to be sorted into individual columns and pasted into the data analysis upload template. This operation may be accomplished using the "Universal Custom PCR Array Patch" available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal.

1. **Copy the C_T values exported from the real-time PCR instrument.**
2. **Paste the values into the yellow cells starting at cell A05 of the patch.**
   The patch will automatically sort the separate columns representing each sample.
3. **For Web-based Data Analysis, copy the appropriate cells and paste (using "Paste Special" and choose "Values") to the upload template.**
4. **Save the file as <u>Excel version 1997–2003</u> and use this file to upload data to the PCR Array Web-based Data Analysis Portal.**
   Note: See Figure 6 below for an example of a correctly formatted upload template for a custom array (based on 96-gene content).

Position	Gene Symbol	Assay Catalog #	Resting 6 h Control Group	6-h Stimulation Test Group 1	Resting 24 h Test Group 2	24-h Stimulation Test Group 3
A01	BMP1	PPH00515B	29.08	29.89	28.86	27.39
A02	BMP2	PPH00549C	32.02	31.15	33.33	31.89
A03	BMP3	PPH00513F	33.83	31.57	35.7	37.93
A04	BMP4	PPH00546A	33.95	31.3	33.05	29.21
A05	BMP5	PPH00509A	Undetermined	Undetermined	Undetermined	Undetermined
…	…	…	…	…	…	…
H12	PPC		19.9	20.65	20.09	19.85

Figure 6. An example of a correctly formatted upload template for a custom PCR array.

For single or multi-gene qPCR assays

If analyzing data from one or a set of qPCR primers, the raw C_T values for each sample should be in a single column that corresponds to the correct well location (column A) on the plate used. The gene tested in each well must be supplied by the user in column B, and the assay catalog number for each gene must be supplied by the user in column C. A sample spreadsheet is available on the "Upload Data" page of the RT^2 Profiler PCR Array Data Analysis Webportal.

1. **Copy the C_T values exported from the real-time PCR instrument into the Individual Assay upload template.**
2. **Save the file as Excel version 1997–2003 and use this file to upload data to the RT^2 Profiler PCR Array Web-based Data Analysis Portal.**
   Note: See below for an example of a correctly formatted template for analyzing individual assays.

Gene #	Gene Symbol	Assay Catalog #	Resting 6 h Control Group	6-h Stimulation Test Group 1	Resting 24 h Test Group 2	24-h Stimulation Test Group 3
1	BMP1	PPH00515B	29.08	29.89	28.86	27.39
2	BMP2	PPH00549C	32.02	31.15	33.33	31.89
3	BMP3	PPH00513F	33.83	31.57	35.7	37.93
4	BMP4	PPH00546A	33.95	31.3	33.05	29.21
5	BMP5	PPH00509A	Undetermined	Undetermined	Undetermined	Undetermined
6	BMP6	PPH00542F	29	26.67	29.38	26.32
7	BMP7	PPH00527A	Undetermined	37.28	Undetermined	34.3
8	BMP8B	PPH00559C	27.33	29.54	28.24	31
9	CSF1	PPH00124B	25.52	21.26	23.91	20.57
10	GAPDH	PPH00150F	17.89	18.2	17.52	16.86

Figure 7. An example of a correctly formatted template for individual PCR assays.

Upload data

Before arriving at the RT^2 Profiler PCR Array Data Analysis Webportal, the QIAGEN product format (Standard/Cataloged RT^2 Profiler PCR Array, Custom RT^2 Profiler PCR Array, or Single or Multi-Gene RT^2 qPCR Primer Assays) and the plate format (96- or 384-well plate format or Rotor-Gene Q) used to perform the experiment should have been defined. If Standard/Cataloged or Custom RT^2 Profiler PCR Arrays were used, the catalog number of the array should also have been provided.

1. **In the RT^2 Profiler PCR Array Data Analysis Webportal, click the "Browse" button to select the MS Excel file prepared in the above section ("Formatting data for the Data Analysis Webportal") that contains all PCR data with a maximum number of 100 samples.**
2. **Click the "Upload" button.**
   Note: Data to be uploaded **must** be saved as Excel Version 1997–2003. Data cannot be saved as an "*.xlsx" file.

Analysis setup

After uploading the data, the software will move to the "Analysis setup" tab that includes the following submenu items: "Uploaded Data," "Sample Manager," "Data QC," "Select Normalization Method," and "Data Overview."

Uploaded Data

This page will be displayed after the Excel 97-2003 spreadsheet (*.xls) has been uploaded. Ensure that the data, sample names, and group definitions (if used) were uploaded as intended. Any red C_T values or indications of undetectable amplification (e.g. Undetermined, ND, N/A, 0, etc.) will be converted to the "C_T Cut-off" referred to in, and defined by, the "Sample Manager" page.

Sample Manager

1. **Assign samples to groups.** If Samples were already assigned to Groups in the uploaded Excel file, then those assignments will already be defined in the software and displayed here. If not or if changes need to be made, assign which samples were obtained under "Control" and various "Test" conditions and place replicates to the same group using the options available from the dropdown menus next to each "Sample Name". The software supports the analysis of up to 12 test groups; however, each will use the same "Control Group" as its reference Group for the calculations of Fold Change and Fold Regulation reported by the software. To exclude a sample from the data analysis, select "Exclude Sample" from the dropdown menu.

   Note: Only one set of samples may be used as the "Control Group". If multiple control groups are to be used, each analysis with a different control group will need to be performed separately, which can be done the same uploaded data after saving the results of the initial comparison.

2. **Factor in whether RT² PreAMP Primer Mixws used.** If an RT² PreAMP Primer Mix kit was used during the wet-bench protocol, simply select the checkbox reading "Click if a PreAMP Primer Mix Kit was used".

3. **Define housekeeping (reference) genes (for Custom RT^2 Profiler PCR Array and individual RT^2 qPCR Primer Assays only).** When analyzing data obtained from Custom RT^2 Profiler PCR Arrays or individual assays, the end-user must define the genes to be used for manual or automatic reference genes selection for the data normalization described below. Scroll down the gene list in the "Select Housekeeping Genes" table to find the desired housekeeping/reference genes. Select a gene by clicking it and then click the "Add" button to include the gene in the "Selected Housekeeping Genes" list that appears when the first gene/genes are added. Hold down the CTRL key and click on multiple genes to select and add them all at the same time. To remove a gene from the "Selected Housekeeping Genes" list, click on it and then the "Remove" button.

4. **Lower Limit of Detection: Set C$_T$ Cut-off.** For most applications, the "Set C$_T$ Cut-off" value will not need to be changed. Only reset this value if specifically recommended by Technical Support. If a 96- or 384-well plate format was selected, the value of 35 will be pre-populated in the dropdown menu. If the Rotor-Gene Q format was selected, the value of 33 will be pre-populated in the dropdown menu. The "Set C$_T$ Cut-Off" function defines the upper limit of useful C$_T$ values in calculating fold change results. All raw undetermined C$_T$ values (e.g., Undetermined, ND, N/A, 0, etc.) or C$_T$ values greater than the "Set C$_T$ Cut-Off" value are changed to that value. Under special circumstances, the "Set C$_T$ Cut-off" value may be adjusted to C$_T$ values ranging from 30 to 40. Please contact a local QIAGEN technical representative.

    Note: Because of the assigned "Set C$_T$ Cut-Off" value, apparent fold-change may be returned for a gene that is not detected. Therefore, when interpreting the data, it is crucial to confirm the context of the absolute C$_T$ values. See the "Analysis" section.

5. **Click the "Update" button after all changes have been made and before moving to the next tab.**

6. **If needed, click the "Undo All Changes" button to reset all of the settings to the last time that the "Update" button was clicked.**

Data QC

A key benefit of the RT^2 Profiler PCR Arrays is the proprietary quality control assays included on all cataloged arrays (and highly recommended for custom arrays) to provide information on the technical quality of the results. These controls are the Genomic DNA Control (GDC), which measures the potential contribution of genomic DNA to the raw C_T values; the Reverse Transcription Control (RTC), which measures the relative efficiency of the reverse transcription; and the Positive PCR Control (PPC), which measures how well the overall PCR performed. This section of the Analysis setup compares the specific results of the control wells to established parameters and provides information on whether or not the controls passed QC.

If all samples passed all three quality controls and their parameters, then only a table reporting this fact will be displayed. If any sample did not pass any of the three quality controls and their parameters, the table will display which sample(s) did not pass which control(s). Another table will then display the results of all controls from all samples that did not pass at least one control and the parameters that they were supposed to meet. Contact a local QIAGEN technical representative for further troubleshooting support.

Control descriptions

Genomic DNA Control
Within each species, a unique non-coding region within the genome that is far removed from any transcriptional start sites has been identified. The GDC is a very sensitive assay for that region to replace the need to set up a separate reaction to run a traditional "No Reverse Transcription" Control. If the GDC C_T value is lower than the established cutoff, then re-purification of the RNA with genomic DNA removal may be required.

Reverse Transcription Control (RTC)
An artificial mRNA with a poly-A tail that is not homologous to any mammalian or bacterial sequence is pre-loaded into the primer buffer of the RT^2 First Strand cDNA synthesis kit and reverse transcribed along with the messages in the samples. The RTC detecting this sequence determines if the reverse transcription efficiency was similar for all samples; that is, it helps determine whether or not the data from all samples may be used for comparison. RNA quality influences reverse transcription efficiency the most. If the RTC is not within the established parameters when compared to the PPC, then the most likely next step is to re-purify the RNA.

Positive PCR Control (PPC)
The PPC wells contain a small amount of DNA with another artificial sequence (not homologous to the RTC) and primers designed to amplify this sequence. Testing and verification of the PPC defined that the PPC C_T range should always yield C_T values within a specific range and be consistent within an array and between arrays. If the PPC C_T values are not within this range, then the PCR itself was likely negatively impacted.

Control Parameters

PCR Array Reproducibility
This metric measures the technical variability of the PPC wells. If the Average PPC C_T of the three replicates within an array is 20 ± 2 and the Average PPC C_T values of any two arrays is not more than 2 different from one another, then the sample receives a "Pass" from the software; otherwise, it reports "Inquiry".

RT Efficiency
If ΔC_T (AVG RTC – AVG PPC) ≤ 5, RT Efficiency reports "Pass"; otherwise, it reports "Inquiry."

Genomic DNA Contamination
If C_T(GDC)$\geq$35, then the GDC QC reports "Pass"; otherwise it reports "Inquiry."

As a guideline, if the GDC C_T value <30, genomic DNA is likely contributing signal to most if not all the genes of interest on the array. The recommended course of action is to further treat the RNA sample with DNase, either directly on a spin-column or in solution followed by spin-column purification. If the GDC C_T value is >33, then genomic DNA is not likely contributing signal and what signal it may be contributing is not likely to impact the fold change results. If the GDC lies between 30 and 33, it is likely that genomic DNA is contributing to the signal.

If any samples return "Inquiry" for any of these three controls, contact a local QIAGEN technical representative for further troubleshooting support.

Selecting Normalization Method (housekeeping / internal control / normalization gene selection)

Housekeeping genes or reference genes are expressed at relatively constant levels under most experimental conditions. This constant expression level makes these genes ideal for normalization and use as a reference. Cataloged RT^2 Profiler PCR Arrays all possess a standard set of five species-specific housekeeping or reference genes. The housekeeping or reference genes on Custom RT^2 Profiler PCR Arrays or detected by individual assays can be defined in the Sample Manager described above. Unless specific housekeeping genes to be used for normalization have been previously determined, choosing one of the two automated HKG selection methods described below is highly recommended.

Three methods for data normalization are available: Manual Selection, Automatic Selection from HKG Panel, and Automatic Selection from Full Plate.

Manual Selection Procedure
1. **Select the HKG gene(s) for the manual analysis.** For Cataloged PCR Arrays, all five (5) of the standard HKG genes are displayed and selected by default. For Custom PCR Arrays and Individual qPCR Assays, the housekeeping genes selected in the Sample Manager will be automatically displayed and selected. If housekeeping genes were not selected in the Sample Manager for Custom PCR Arrays or Individual qPCR Assays, then an error message will be displayed saying that they must defined in order to use the feature.
2. **Review the arithmetic and geometric means for each group.** The recommended maximum amount of variation in the normalization factor if 1.5 cycles across all of the samples.
3. **If one of the means meets the above criterion, select the appropriate arithmetic or geometric mean as preferred from the dropdown menu. Click the "Perform Normalization" button.**
   Note: Because real time PCR is a non-linear process, the recommended geometric mean is selected as the default setting. However, unless there is a great deal of variability, the arithmetic mean and geometric mean values are not likely to greatly differ, or the use of one of the other is not likely to greatly affect the final results.
4. **If neither mean meets the criterion, then remove (or add back) genes from the standard HKG gene set from the calculation by clicking the appropriate checkboxes. Click the "Perform Normalization" button to recalculate the means in the table. If the new chosen mean still does not meet the criterion, then select the appropriate mean and/or change the chosen HKG panel and click the "Perform Normalization" button again to re-evaluate. Otherwise, proceed with the "Analysis".**

Automatic Selection from HKG Panel Procedure (recommended method)

1. **The most stable gene(s) are automatically identified from within the 5 default HKG set on Cataloged RT^2 Profiler PCR Arrays or chosen genes for Custom RT^2 Profiler PCR Arrays and Individual RT^2 qPCR Primer Assays.**

2. **Automatic and progressive selection is utilized to only return those genes within the group of genes exhibiting minimal changes in raw C_T values.** The maximum amount of variation allowed for inclusion in the returned list is 1.5 cycles across all of the samples.

3. **The C_T values for these genes are automatically geometrically averaged and used for the $\Delta\Delta C_T$ calculations.**

4. **If satisfied with the automatically chosen HKG, click the "Perform Normalization" button and proceed with the "Analysis."**

Automatic Selection from Full Plate Procedure

1. **The most stable genes are automatically identified from the entire gene panel (excluding GDC, RTC and PPC) by the same method as above.**

2. **The C_T values for these genes are automatically geometrically averaged and used for the $\Delta\Delta C_T$ calculations.**

3. **If satisfied with the automatically chosen HKG, click the "Perform Normalization" button and proceed with the "Analysis".**
   Note: The Automatic Selection from Full Plate method is only recommended if the Automatic Selection from the HKG Panel method fails to identify a suitable HKG.

Note: Be sure to click "Perform Normalization Method." The software will not allow movement to "Analysis" if this is not clicked.
Note: If comparing fold-change results between different experiments, ensure that the same HKGs are used for each analysis.
Note: Always record the reference genes that were used.

Data Overview
The "Data Overview" section provides simple tabulation of each group's distribution of C_T values and the average of the raw data in each group. Use the "Select Group" dropdown menu to choose which summary to view. When finished, click on the "Analysis" tab to continue.

Analysis
As described earlier, fold-change calculations or gene expression ratios are calculated using the classic, well-established, and widely adopted $\Delta\Delta C_T$ method. The Analysis page has several submenus: Fold Regulation (default view), Fold Change, Average ΔC_T, and $2^{\wedge} - \Delta C_T$.

Fold Change

Fold Change
Fold change is calculated by using the $\Delta\Delta C_T$ method originally published in Livak K.J. and Schmittgen T.D. (2001) Methods **25**, 402. It is the ratio of the relative gene expression between the Control Group and Test Group. Numbers greater than 1 indicate upregulated or increased gene expression, numbers between 0 and 1 indicate downregulated or decreased gene expression, and a fold change of 1 indicates no change.

For example, a fold-change value of 0.25 occurs when the normalized expression of a gene in the test sample is four times lower than the normalized gene expression in the control sample.

First, ΔC_T, also called the "normalized raw data," for each gene is calculated by subtracting the selected normalization factor from the C_T value of each gene of interest. If one housekeeping gene is selected, then its C_T value is used as the normalization factor. If multiple housekeeping genes are selected, then either the arithmetic or geometric mean of their C_T values is used as the normalization factor based on the selection made in the "Select Normalization Method" page of "Analysis Setup".

$\Delta C_T = C_{T\,(\text{Gene of Interest})} - C_{T\,(\text{Housekeeping / Reference Gene})}$
or
$\Delta C_T = C_{T\,(\text{Gene of Interest})} - \text{Average}\,(C_{T\,(\text{Housekeeping / Reference Genes})})$

Next, gene-specific ΔC_T values in Samples in the same Group will be averaged. If there is only 1 sample per group, this step is skipped.

Average $\Delta C_T = (\Delta C_T\,(\text{Sample1}) + \Delta C_T\,(\text{Sample2}) + \Delta C_T\,(\text{Sample3}) + \ldots \Delta C_T\,(\text{Sample } n)) / n$ Samples

The $\Delta\Delta C_T$ value is calculated by subtracting the ΔC_T value of each Test Group from the Control Group.

$\Delta\Delta C_T = \Delta C_T\,(\text{Test Group } n) - \Delta C_T\,(\text{Control Group})$

The fold-change value is then calculated by converting the $\Delta\Delta C_T$ from a log2 scale to a linear scale using the following equation:

Fold change $= 2^{(-\Delta\Delta C_T)}$

P-value

The p-values are calculated using a Student's t-test (two-tail distribution and equal variances between the two samples) on the replicate $2^{-\Delta C_T}$ values for each gene in each treatment group compared to the control group. The p-values less than 0.05 are indicated in red. Each group (including the control) must contain at least 3 samples for the software to calculate p-values.

> Note: To perform higher order statistics, e.g. Anova, etc., calculate the $2^{-\Delta C_T}$ for each individual gene in each sample and import those into the desired statistical analysis software.

Comments

The Comments provide disclaimers on fold-change (or fold-regulation) results based on the original average C_T values in the Groups compared. Given the higher degree of variability that may be observed with lower copy number templates (e.g. $C_T > 30$), care needs to be taken when interpreting such results and deciding on next steps such as verification, confirmation, functional studies, etc. The "Comments" provide insight into the relative abundance of the transcripts based on their C_T values used in the fold change calculation. An example is shown below.

Table 1. Examples of comments and explanations.

Comment	Explanation
[No Comment displayed]	The gene's expression is relatively high in both the test and control group. The threshold cycle (C_T) is <30.
"A"	The gene's expression is relatively low or not detected in either the test or control group. Low expression means that the average threshold cycle value is high ($C_T > 30$) for that group. The expression in the other group is then relatively high ($C_T < 30$). Because the detection of gene expression can be limited by the original amount of mRNA, the real fold-change value is at least as large as reported. This result may also have variations in p-value (>0.05). Therefore, it is important to have a sufficient number of biological replicates to verify the results for this gene.
"B"	The gene's relative expression level is low in both samples. Low expression means having a high average threshold cycle ($C_T > 30$). In addition, the p-value is either not available or relatively high (>0.05). The fold change or fold regulation result may be more variable due to the low levels of expression. Therefore, it is important to have a sufficient number of biological replicates to verify the results for this gene.
"C"	This gene is not detected in either sample. The gene's average threshold cycle (C_T) is either not determined or greater than the user defined cut-off value. (See "Lower Limit of Detection: Set C_T Cut-Off" section of the "Sample Manger" page in the "Analysis setup" tab). While the software tries to calculate a fold-change or fold-regulation value after normalization, the fold change may be erroneous.

Fold Regulation
Fold regulation provides the same information as fold change (including p-value and comments), but fold regulation displays the data in a format that is easier to read and interpret.

For all fold-change values greater than 1, the fold-regulation and fold-change values are the same. For all fold-change values (X) less than 1, the fold regulation is the negative inverse of the fold change ($-\frac{1}{X}$). For example, a fold-change value of 0.25 corresponds with a fold regulation of −4.0.

Fold-regulation values greater than 1 indicate upregulated or increased gene expression; fold-regulation values less than zero indicate downregulated or decreased gene expression.

For example, a fold-regulation value of −4.0 occurs when the normalized expression of a gene in the test sample is four times lower than the normalized gene expression in the control sample.

With fold change and p-values calculated and comments reviewed, continue to generate **"Plots and charts"** by clicking on the appropriate tab.

Average ΔC_T page
This page provides, in table format, the average ΔC_T within the "Control Group" and "Test Group" for each gene.

$2^{-\Delta C_T}$
This page provides in table format the average $2^{-\Delta C_T}$ within the "Control Group" and "Test Group" for each gene.

Plots and charts

The "Plot" home page shows a sample of each type of figure that the software can generate. Simply click on the type of figure or the sub-menu item for the desired graph.

Scatter plot

The scatter plot compares the normalized expression of each gene on a PCR array between 2 groups. Every gene-specific $2^{-\Delta C_T}$ value, or normalized expression level, in one group is plotted on one axis against the corresponding value in the other group on the other axis on a log base 10 scale to neatly visualize gene expression changes. Boundary lines are used to enable clear visualization of gene expression changes of interest for upregulated and downregulated genes above and below a particular selected fold-change value. The central line indicates unchanged gene expression. The upper left section of the scatter plot (above the fold-change boundary lines) contains genes upregulated in the y-axis group as compared to the x-axis group. The lower right section of the scatter plot (below the fold-change boundary lines) contains genes downregulated in the y-axis group as compared to the x-axis group.

A default scatter plot comparing the "Control Group" to "Test Group 1" with a fold-change threshold of 2 and a blue-yellow color scheme (for downregulated and upregulated genes, respectively) is the default display.

Procedure: Modifying the default scatter plot
1. **Using the dropdown menus at the top of the window, select the groups to be plotted on the "Y Axis" and "X Axis."**
2. **Change the boundary lines or desired threshold for fold-change by entering a new value in the "Boundary" dialog box.** The boundary line settings must be determined by the researcher and set appropriately for the experiment and model system used.
3. **Select the desired "Color Scheme" from the dropdown menu: blue-yellow, blue-red, or green-red for downregulated and upregulated genes, respectively.**
4. **After modifying all of these parameters as desired, click the "Update" button to view the new scatter plot.**
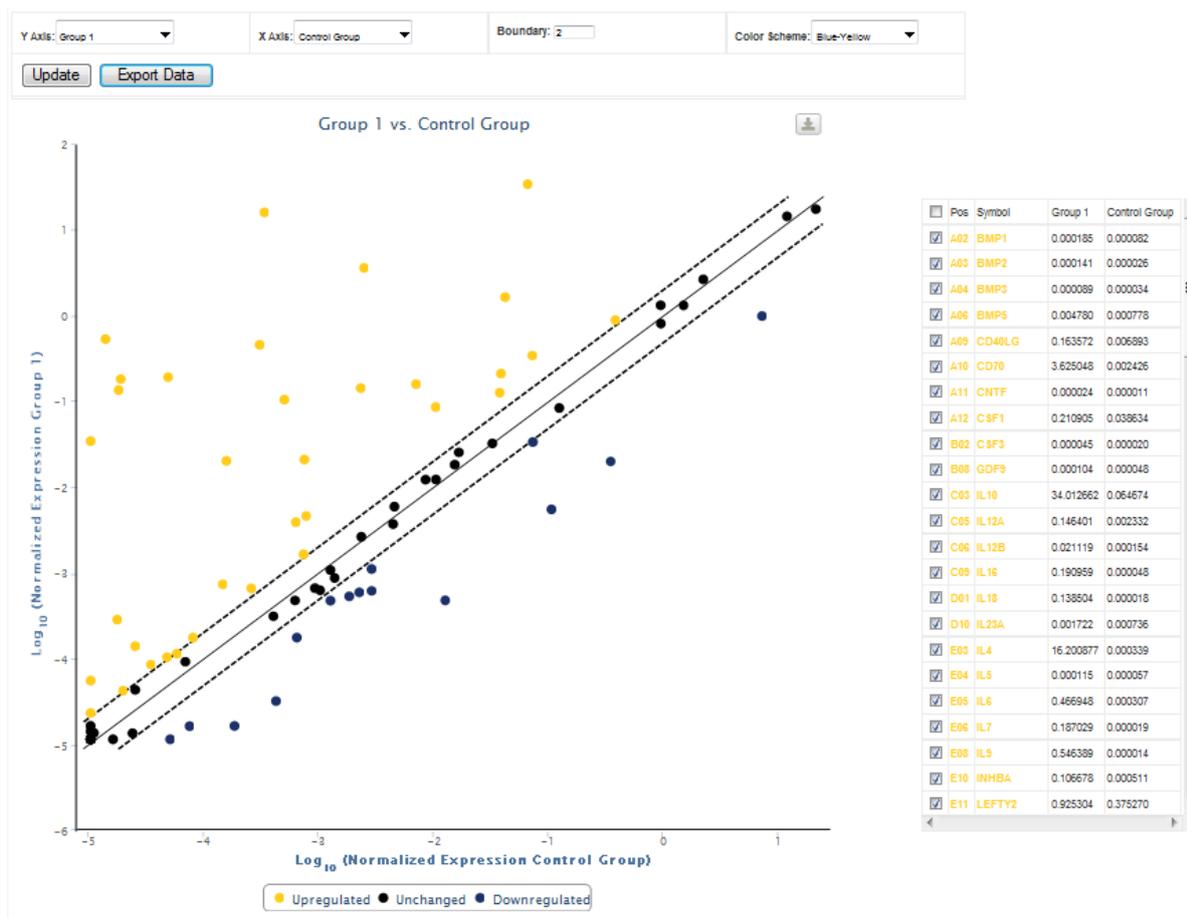
Figure 8. A scatter plot created in the RT² Profiler PCR Array Data Analysis Webportal.

Procedure: Identifying genes on the scatter plot
1. The right-hand table lists the array position ("POS"), "Symbol," and normalized gene expression levels in the two "Groups" used to generate the scatter plot.
2. Highlight and identify points on the plot by moving the mouse over the gene entries in the right-hand table.

Procedure: Excluding genes from the scatter plot
1. Using the table on the right hand side of the window, deselect the check box of any gene to be removed from the scatter plot. (Conversely, reselect any check box of a gene to add it back to the scatter plot.) Click the "Update" button to view the new scatter plot.
2. By default, the GDC, RTC, and PPC control wells are not displayed in the scatter plot.

	Pos	Symbol	Group 1	Control Group
☑	A02	BMP1	0.000185	0.000082
☑	A03	BMP2	0.000141	0.000026
☑	A04	BMP3	0.000089	0.000034
☑	A06	BMP5	0.004780	0.000778
☑	A09	CD40LG	0.163572	0.006893
☑	A10	CD70	3.625048	0.002426
☑	A11	CNTF	0.000024	0.000011
☑	A12	CSF1	0.210905	0.038634
☑	B02	CSF3	0.000045	0.000020
☑	B08	GDF9	0.000104	0.000048
☑	C03	IL10	34.012662	0.064674
☑	C05	IL12A	0.146401	0.002332
☑	C06	IL12B	0.021119	0.000154
☑	C09	IL16	0.190959	0.000048
☑	D01	IL18	0.138504	0.000018
☑	D10	IL23A	0.001722	0.000736
☑	E03	IL4	16.200877	0.000339
☑	E04	IL5	0.000115	0.000057
☑	E05	IL6	0.466948	0.000307
☑	E06	IL7	0.187029	0.000019
☑	E08	IL9	0.546389	0.000014
☑	E10	INHBA	0.106678	0.000511
☑	E11	LEFTY2	0.925304	0.375270

Figure 9. Excluding genes from the scatter plot.

Procedure: Exporting associated data table and scatter plot image
1. Click the "Export Data" button to generate and download an Excel file containing the data used to generate the scatter plot. Save the file on a local computer or server.
2. Mouse over and click on the icon in the upper right hand corner of the scatter plot. Mouse over to the desired file format and click it to save the image of the scatter plot on a local computer or server.

Figure 10. Saving the scatter plot.

Volcano plot

The volcano plot helps evaluate gene expression changes in the context of their statistical significance. The graph plots the log base 2 of each gene's fold-change value on the x-axis and the negative log base 10 of the gene's p-value on the y-axis. The default plot compares the **"Control Group"** with **"Test Group 1."** The p-values are calculated as described in the **Analysis** section using the student's t-test of the replicate raw C_T data, a method widely accepted in scientific literature. The solid vertical represents no change in gene expression ($\log_2(1) = 0$). The data points to the right of the solid vertical line indicate upregulated genes, while the data points left of the solid vertical line indicate downregulated genes. The dotted lines represent a selected threshold or boundary for fold change. The default setting is 2. The solid vertical line represents a selected threshold for the p-value and statistical significance. The default setting is 0.05. Data points above the solid vertical line are statistically significant, while data points below the line are not.

Note: Because the volcano plot utilizes p-values, the volcano plot function is only active if there are 3 or more replicates in each of the groups that are being compared. An error message is returned if trying to compare two groups where either one or both has less than 3 samples or replicates.

Procedure: Modifying the default volcano plot
1. **Select the "Test" and "Control" groups to be compared in the plot using the dropdown menus.**
2. **Input the desired fold-change threshold in the "Boundary" dialog box.**
3. **Input the desired p-value threshold for statistical significance in the "P-value" dialog box.**
4. **Select the desired "Color Scheme" from the dropdown menu: blue-yellow, blue-red, or green-red for downregulated and upregulated genes, respectively.**
5. **After modifying all of these parameters as desired, click the "Update" button to view the new volcano plot.**

Figure 11. A volcano plot generated by the RT² Profiler PCR Array Data Analysis Webportal.

Procedure: Identifying genes on the volcano plot
1. The right-hand table lists the array position ("POS"), "Symbol," and the p-value and the fold-regulation values used to generate the volcano plot.
2. Highlight and identify points on the plot by moving the mouse over the gene entries in the right-hand table.

Procedure: Excluding genes from the volcano plot
1. Using the table on the right-hand side of the window, deselect the check box of any gene to be removed from the volcano plot. (Conversely, reselect any check box of a gene to add back to the volcano plot.)
2. Click the "Update" button to view the new volcano plot.
   Note: By default, the GDC, RTC, and PPC control wells are not displayed in the volcano plot.

Procedure: Exporting the associated data table and volcano plot image
1. Click the "Export Data" button to generate and download an Excel file containing the data used to generate the volcano plot. Save the file on a local computer or server.
2. Mouse over and click on the icon in the upper right-hand corner of the volcano plot. Mouse over to the desired file format and click it to save the image of the volcano plot on a local computer or server.
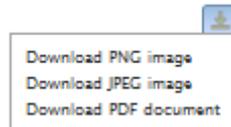

Figure 12. Saving the volcano plot.

Clustergram

The clustergram performs non-supervised hierarchical clustering of the entire dataset to display a heat map with dendrograms indicating co-regulated genes across groups or individual samples. With PCR array data, the clustergram provides an overview of what genes may be co-regulated by a common factor, e.g. a transcription factor, miRNA, methylation status, etc.

Clustering algorithms divide a set of objects (genes or samples) into clusters so that gene expression patterns within a cluster are more similar than patterns across clusters. Cluster analysis includes hierarchical and non-hierarchical clustering methods. Hierarchical clustering is used to understand the relationship of each gene or array to every other gene or array, respectively. Hierarchical clustering is also used when little or no prior knowledge of how the data will be clustered is known, because hierarchical clustering does not set the number of clusters to form before the analysis begins. Hierarchical clustering analysis produces dendrograms, which are treelike figures used to graphically represent the hierarchy.

Note: Clustergrams **do not** cluster genes according to known functional or sequence homology.

The clustergram is based on a hierarchical clustering method that:

- Assigns each gene and each array or group to its own cluster (agglomerative)
- Joins the nearest clusters
- Re-estimates the distance between clusters
- Repeats for each gene and each array or each group

Objects that are more similar are joined at lower heights in the dendrogram and those that are less similar are joined at higher heights. To create a hierarchical cluster, a distance/dissimilarity metric (here based on the magnitude of gene expression) and linkage method must be specified. The magnitude of gene expression is determined by calculating the $2^{-\Delta C_T}$ for each individual gene and normalizing to the average $2^{-\Delta C_T}$ of all genes across all arrays. The distance/dissimilarity metric is used to convert data points into clusters, while the linkage method is used to join the clusters to form a tree.

The software utilizes a correlation coefficient to express similarity of all genes across all arrays calculated based on the following equation for the correlation coefficient between two dimensional profiles:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \, \sigma_Y} = \frac{\sum_{i=1}^{N}(X_i - \mu_X)(Y_i - \mu_Y)}{\sqrt{\sum_{i=1}^{N}(X_i - \mu_X)^2 \cdot \sum_{i=1}^{N}(Y_i - \mu_Y)^2}}$$

The **Clustergram** page provides optional features to change the appearance of the display by selecting new options with the dropdown menus; however, the default settings provide the most easily interpreted representation.

Procedure: Modifying the default clustergram
1. **Select "Array" or "Group" for "Sample."** Selecting **"Array"** (the default setting) uses the gene-specific magnitude of gene expression from each individual PCR array as the distance/dissimilarity metric in the analysis and displays the results from each PCR array as a separate column of color-coded data. Selecting **"Group"** uses the average gene-specific magnitude of gene expression across PCR arrays in the same group (**"Control Group," "Test Group 1,"** etc.) as the distance/dissimilarity metric in the analysis and displays the results from each **"Group"** as a separate column of color-coded data.
2. **Select "Dimension."** Selecting **"2-D"** (the default setting) displays dendrograms on both the x-axis (**"Array"** or **"Groups"**) and y-axis (**"Genes"**). Selecting **"1-D"** only draws a dendrogram for the y-axis.
3. **Select "Join Type."** See the figure below the following descriptions for a visual demonstration of the difference between the three join types.
   a. **Selecting "Maximum" or "complete linkage" defines the distance between clusters using the pair of genes with the greatest difference in their distance/dissimilarity metric.** This method tends to produce dendrograms with shorter branches that clump together.
   b. **Selecting "Minimum" or "single linkage" defines the distance between clusters using the pair of genes with the smallest**

difference in their distance/dissimilarity metric. This method tends to produce dendrograms with longer branches.

    c. **Selecting "Average" linkage defines the distance between clusters using the average of the differences in the distance/dissimilarity metric between all pairs of genes in the two clusters.**

Note: Only one individual gene from each cluster is considered when using "Maximum" or "Minimum," while using "Average" considers all genes in each cluster.
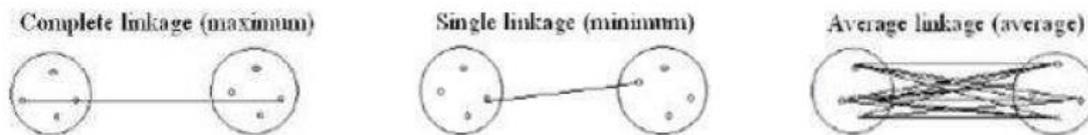

Figure 13. Differences between the three join types.

4. **Select color coding.** The colors in the cells of the clustergram represent the relative magnitude of gene expression based on a selected method of calculating an average value. The black color represents the average magnitude of gene expression. The brightest green represents the smallest value, and the brightest red represents the highest value.

    a. **Selecting "Genes" (the default setting) sets a different average value for each gene on the y-axis.**

    b. **Selecting "Samples" sets a different average for each "Sample" on the x-axis, whether by "Array" or "Group."**

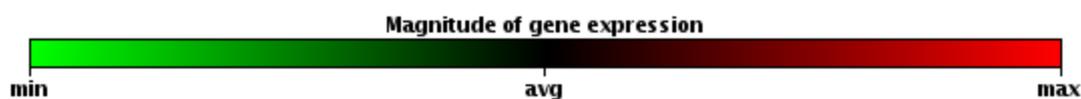   c. **Selecting "Entire Dataset" sets one average for all data used to generate the clustergram.**

Figure 14. Clustergram color-coding.

5. **Select "Item names visible."** Selecting **"Yes"** (the default setting) labels the names of the **"Arrays"** or **"Groups"** and **"Genes"** on the clustergram. Selecting **"No"** hides the names of the **"Arrays"** or **"Groups"** and **"Genes"** on the clustergram.

6. **Once all desired changes have been made to these settings, click the "Update" button to visualize the new clustergram.**

7. **Save the clustergram image.**

    a. **If using Windows, right-click on the figure, select "Save image/picture as …", and follow the prompts.**

    b. **If using OS X, hold down the Control key and then click the image/picture to save it.**

Heat Map

The heat map provides a graphical and color-coded representation of fold regulation expression data between two groups overlaid onto the PCR array plate layout. The black color represents the average magnitude of gene expression. The brightest green represents the smallest value, and the brightest red represents the highest value.

The table below the heat map contains the fold regulation data used to generate the plot for each of the annotated gene symbols in its well position. If any fold-regulation results were annotated with **Comments** in the **Fold Regulation** page of the **Analysis** tab, then those comments are also annotated in the heat map table, and their definitions are re-listed under the table.

Procedure: Modifying the default heat map

1. **Select the "Treated Group" and the "Control Group" to compare in the display.**
2. **Selecting "Log$_2$ transform fold change" checkbox (the default setting) uses the log base 2 of the fold regulation values to determine the color-coding scheme. Deselecting the checkbox uses the unmodified fold-regulation values to determine the color-coding scheme.**
3. **Once all desired changes have been made to these settings, click the "Update" button to display the heat map.**

Procedure: Exporting the associated data table and heat map image

1. **Click the "Export Data" button to generate and download an Excel file containing the data used to generate the heat map. Save the file on a local computer or server.**
2. **To save the heat map image:**
   a. **Using Windows, right-click on the figure, select "Save image/picture as …", and follow the prompts.**
   b. **Using OS X, hold down the Control key and click the image/picture to save it.**

Multigroup plot

The multigroup plot allows the visualization of a few data points across several treatment groups. Data may be plotted for selected genes of interest with either lines or bars with or without error bars using chosen definitions of uncertainty. This plot is best suited for following the trend of expression changes for a few genes across the various treatments in the experiment, such as time-course or dose-response experiments to monitor changes over time or drug concentration.

The x-axis plots all **"Groups,"** and the y-axis may plot either **"Average delta (C_T)," "2^-delta (C_T),"** or **"Fold Change."**

Procedure: Modifying the default multigroup plot
1. **Select the data to be plotted on the y-axis using the appropriate "Y Represent" radial button. Selections include "Average delta (C_T)," "2^-delta (C_T)," and "Fold Change."**
2. **Select the desired "Chart Type" or method to visualize the graphed data, either a "Line Chart" or "Column Chart," using the dropdown menu.**
3. **Click "Show Uncertainty" checkbox to include error bars on the multigroup plot. The error bars are defined by the following parameters depending on the chosen "Y represent" option above:**

Table 2. Parameters for error bars in the multigroup plot

Y-axis	Uncertainty
Average ΔC_T	standard deviation of ΔC_T
$2\char`^-\Delta C_T$	standard deviation of $2^{-\Delta CT}$
Fold Change	$\sigma_{FC} = FC * \ln2 * \sqrt{\sigma_x^2 / n_x + \sigma_y^2 / n_y}$

4. **Define the genes of interest to be plotted by selecting or deselecting the checkboxes in the table on the left side of the multigroup plot. Limiting the selection to 1 to 5 genes of interest maintains figure clarity. The maximum allowable number of selected genes is 10.**
5. **Once all desired changes have been made to these settings, click the "Update" button to visualize the new multigroup plot.**
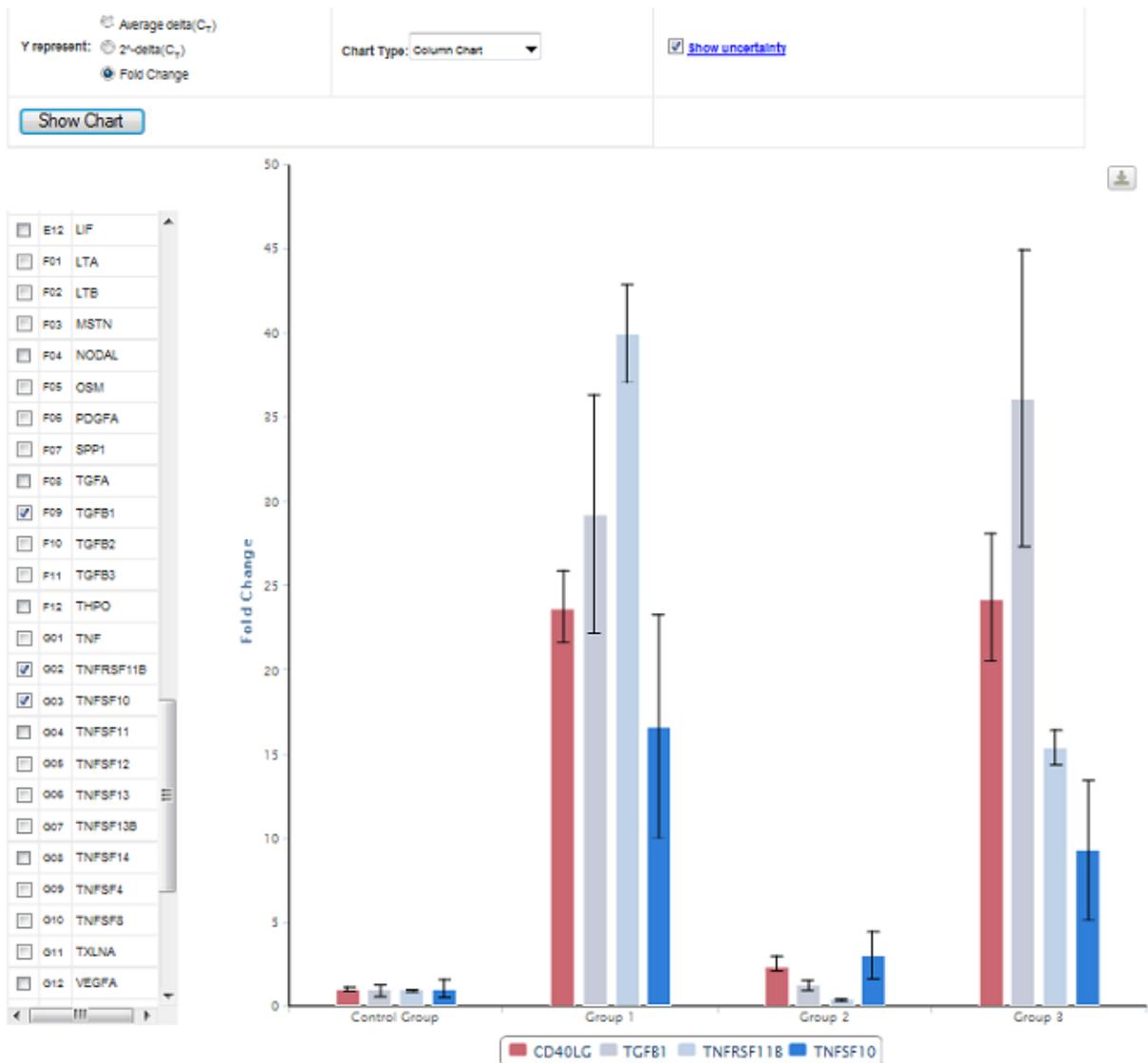
Figure 15. Multigroup plot generated by the RT² Profiler PCR Array Data Analysis Webportal.

Procedure: Exporting the multigroup image

1. Mouse over and click on the icon in the upper right-hand corner of the multigroup plot.
2. Mouse over to the desired file format and click it to save the image of the multigroup plot on a local computer or server.
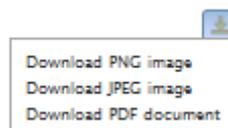

Download PNG image
Download JPEG image
Download PDF document

Figure 16. Saving the multigroup plot.

Export data

1. **Select the data tables to export using the corresponding check boxes and click "Export" to download an Excel file with each data table in a different worksheet.**

PCR Data Analysis Result:

Gene Table	☑
Average C_T	☑
Average Delta(C_T)	☑
2^(-Avg.(Delta(C_T))	☑
Fold Change	☑
P-Value	☑
Fold Regulation	☑
submit	

Figure 17. The export data screen.

2. **After saving the Excel to a local computer or server, click on the "What's next" tab and investigate how QIAGEN can help determine the next best experiment to follow up on the RT² Profiler PCR Array gene expression results.**

    Note: The RT² Profiler PCR Array Data Analysis Webportal does not save any results. Please ensure that any figures or tables from each session have been exported or copied and saved on a local computer or server.

What's next

The data analysis software delivers a list of expression changes in the samples from the supplied data. However, this result often only starts an investigation into the underlying mechanisms at work. In order to assist in further analysis, the software now utilizes the latest bioinformatics tools to analyze the data and suggest regulatory mechanisms and future experiments. Explore each sub-menu item to examine their results. Please note that not all "What's next" modules are available for all species.

Note: All "What's next" modules are available for human and mouse PCR arrays and qPCR assays. The "miRNA Regulation" module cannot support rat PCR arrays and qPCR assays. All other species can only be supported by the "Gene Expression" module.

All "What's next" modules first prompt for the definition of search criteria to be used in the analysis.

Procedure

1. **Step 1: Define the "Control Group" and "Experimental Group" to be compared.**
2. **Step 2: Specify whether "Decreased" or "Increased" expression changes, or "Both," are to be analyzed.**
3. **Step 3: Define the fold regulation and p-value thresholds to be used in the analysis.**
4. **Once the parameters have been defined, click "Show Selected Genes" in Step 4.** A table will be displayed listing each **"Gene Symbol"** meeting the search criteria and its well **"Position," "Fold Regulation,"** and **"p value"** (if applicable).

Note: If either of the two groups in the comparison have less than 3 samples and a p-value threshold is used, an error message is displayed indicating that p-values could not be calculated and that the p-value threshold will be ignored.

Follow the different prompts available in each specific module to eventually save the data as Excel files on a local computer or server.

Gene Expression

This tool will help define a panel of genes based off this experiment's results. This panel may represent a putative biomarker set, a target gene set, or simply a collection of genes. The tool is designed to deliver a list of gene expression assays that enable the user to follow up on the results of the analyzed experiment.

Procedure (following Steps 1–4)

1. **Step 5: Select the reagent to be used in the next experiment by clicking the appropriate radial button.**

2. **Step 6: Export data.**
   a. **For the "Individual Assays" option, click the "Export Data" button to generate and download an Excel file containing the table output from Step 4, annotated with RT^2 qPCR Primer Assay catalog numbers hyperlinked to their product webpages on GeneGlobe. Save to a local computer or server.**
   b. **For the "Custom PCR Array" option, first select the number of genes from the table to include in the Custom RT^2 Profiler PCR Array gene list. A limited number of options (8, 12, 16, 24, 32, 48, and 96) are available that match the offered Custom RT^2 Profiler PCR Array layouts. The software automatically chooses the top-ranked genes based on the absolute values of their fold regulation results. Click the "Export Data" button to generate and download an Excel file containing the recommended Custom RT^2 Profiler PCR Array layout and necessary assay catalog numbers. Save to a local computer or server.**

miRNA Regulation

This tool will identify candidate miRNA regulators in this experiment's results. The tool is designed to deliver a list of miRNAs that could be targeting the genes that had observed changes in expression in the selected samples. If genes only changing expression in one direction or the other are displayed in the table output from Step 4, then only the corresponding table will be displayed. The tables list the miRNAs predicted to be regulated the gene that met the search criteria based on the TargetScan algorithm and the latest version of the miRNome from miRBase.

For each miRNA, the table provides the total number and identity of the predicted target genes, the total number of miRNA binding sites across those genes, and the range of the strength scores (TargetScan Z score) across those binding sites. The table also lists the catalog numbers for miScript miRNA Assay, miScript miRNA Mimic, and miScript miRNA Inhibitor for each miRNA hyperlinked to the products' webpage on GeneGlobe.

Procedure (following Steps 1–4)

1. **Step 5: Click the "Search for miRNA regulators" button. Tables are displayed listing the "miRNA Regulators for Genes with Decreased Expression" and the "miRNA Regulators for Genes with Increased Expression." Click on the column headers to read an explanation of their contents in popup windows.**

2. **Click any "Export list of miRNAs" button to generate and download an Excel file containing all of the information and hyperlinks in all tables displayed upon clicking the "Search for miRNA regulators" button. Save to a local computer or server.**

Methylation

This tool will help define a panel of differentially expressed genes based on the experiment's results. Altered methylation patterns on the genes' promoters may be responsible for the observed gene expression changes. This tool is designed to deliver a list of available DNA methylation assays for those differentially expressed genes that enable the user to follow up their gene expression experiment with epigenetic analysis.

Procedure (following Steps 1–3)

1. **Step 4: The output table includes an additional column listing the catalog number of the "Recommended EpiTect Methyl II qPCR Assay" for each differentially expressed gene that met the search criteria. The catalog number text hyperlinks to their product webpages on GeneGlobe. If an EpiTect Methyl II qPCR Assay is not available for a gene, the table instead returns the message "Not likely regulated by methylation."**

2. **Step 5: Select the reagent to be used in the next experiment by clicking the appropriate radial button.**

3. **Step 6: Export data.**
   a. **For "Individual Assays", click the "Export Data" button to generate and download an Excel file containing the table output from Step 4 complete with EpiTect Methyl II qPCR Assay catalog numbers and hyperlinks. Save to a local computer or server.**
   b. **For "Custom PCR Array": First, select the number of genes from the table to include in the custom PCR array gene list. A limited number of options (6, 8, 12, and 24) are available that match the offered custom PCR array layouts. The software automatically chooses the top-ranked genes with EpiTect Methyl II qPCR Assay catalog numbers based on the absolute values of their fold regulation results. Click the "Export Data" button to generate and download an Excel file containing the recommended custom PCR array layout and needed assay catalog numbers. Save to a local computer or server.**

Transcription Factor / Histone

This tool will help define a panel of differentially expressed genes based on the experiment's results. Altered transcription factor binding activity or histone modification changes on the genes' promoters may be responsible for the observed gene expression changes. This tool is designed to deliver a list of the transcription factors that might regulate the selected differentially expressed genes as well as the available respective gene-specific real-time PCR assays for DNA from anti-transcription factor or anti-histone chromatin immunoprecipitation. These assays enable the user to follow up their gene expression experiment with epigenetic analysis.

Procedure (following Steps 1–3)

1. **Step 4: The output table includes an additional column listing the "Transcription Factors" with predicted binding sites in the 1-kb region immediately upstream of each gene's transcription start sites. Another additional column lists the EpiTect ChIP qPCR Assay catalog number for that region of genomic DNA with a hyperlink to its product webpages on GeneGlobe.**
2. **Step 5: Select the reagent to be used in the next experiment by clicking the appropriate radial button.**
3. **Step 6**
   a. **For "Individual Assays", click the "Export Data" button to generate and download an Excel file containing the table output from Step 4 complete with EpiTect ChIP qPCR Assay catalog numbers and hyperlinks to their product webpages on GeneGlobe. Save to a local computer or server.**
   b. **For "Custom PCR Array": First, select the number of genes from the table to include in the custom PCR array gene list. A limited number of options (8, 12, 16, 24, 32, 48, and 96) are available that match the offered custom PCR array layouts. The software automatically chooses the top-ranked genes based on the absolute values of their fold regulation results. Click the "Export Data" button to generate and download an Excel file containing the recommended custom PCR array layout and needed assay catalog numbers. Save to a local computer or server.**

siRNA

This tool will help define a panel of differentially expressed genes based on the experiment's results. These differentially expressed genes may contribute to the observed differences between the tested sample groups. This tool is designed to deliver a list of differentially expressed genes and corresponding siRNA reagents to test the contribution of each differentially-expressed gene in the experiment. These siRNAs enable the user to follow up their gene expression experiment with a functional analysis.

Procedure (following Steps 1–3)

1. **Step 4: The output table includes an additional column containing a hyperlink to a search result page on GeneGlobe for FlexiTube siRNA products for each gene matching the search criteria.**
2. **Click the "Export Data" button to generate and download an Excel file containing all of the data and hyperlinks in the table. Save to a local computer or server.**
3. **Click the "Batch Search GeneGlobe" button at the bottom of the table to navigate to the "Batch Configurator" to design a Custom FlexiPlate using the list in the Excel file exported above.**

Protein Detection

This tool will help define a panel of cytokines or chemokines based on the experiment's results. This panel may represent a putative biomarker set, a target gene set or simply a collection of genes. The tool is designed to deliver a list of ELISAs that enable the user to follow up the results of the analyzed experiment.

Procedure (following Steps 1–3)

1. **Step 4: The output table includes an additional column containing the "Single-Analyte ELISArray" catalog number (if available) for products of the genes matching the search criteria.** The catalog numbers hyperlink to the products' webpages on GeneGlobe. If a Single Analyte ELISArray is not available for a gene product, then the text "Inquire" is returned instead.
2. **Click the "Export Data" button to generate and download an Excel file containing all of the information and hyperlinks in the table. Save to a local computer or server.**

Somatic Mutation

This tool will help define a panel of genes based on the experiment's results. Mutations in these genes may affect whether their expression changes have any effect or different effects on the experimental model system. The tool is designed to deliver a list of available somatic mutation assays for these differentially expressed genes that enable the user to follow up the results of the analyzed experiment.

Procedure (following Steps 1–3)

1. **Step 4: The output table includes an additional column containing the "Somatic Mutation Assay" catalog number (if available) for the genes matching the search criteria. The catalog numbers hyperlink to the products' webpages on GeneGlobe. If a qBiomarker Somatic Mutation qPCR Assay is not available for a gene product, then the text "Inquire" is returned instead.**

2. **Step 5: Export data. Click the "Export Data" button to generate and download an Excel file containing all of the information and hyperlinks in the table. Save to a local computer or server. For custom PCR arrays, click the link for more information about how to set up and order the array based on the results exported above.**

References

QIAGEN maintains a large, up-to-date online database of scientific publications utilizing QIAGEN products. Comprehensive search options allow you to find the articles you need, either by a simple keyword search or by specifying the application, research area, title, etc.

For a complete list of references, visit the QIAGEN Reference Database online at www.qiagen.com/RefDB/search.asp or contact QIAGEN Technical Services or your local distributor.

Ordering Information

Product	Contents	Cat. no.
RT^2 Profiler PCR Arrays	Arrays of assays for disease, pathway, or functionally related genes; available in 96-well, 384-well, and Rotor-Disc 100 formats	330231
Custom RT^2 Profiler PCR Arrays	Arrays of assays for a user-defined set of genes; available in 96-well, 384-well, and Rotor-Disc 100 formats	330131
RT^2 qPCR Primer Assays (200)	For 200 reactions; Mix of 2 gene-specific primers provided in solution (200 μl); 10 μM each primer	Varies

Notes

Limited License Agreement for RT^2 Profiler PCR Array Data Analysis Webportal.

Use of this product signifies the agreement of any purchaser or user of the product to the following terms:

C. The product may be used solely in accordance with the protocols provided with the product and this handbook and for use with components contained in the kit only. QIAGEN grants no license under any of its intellectual property to use or incorporate the enclosed components of this kit with any components not included within this kit except as described in the protocols provided with the product, this handbook, and additional protocols available at www.qiagen.com. Some of these additional protocols have been provided by QIAGEN users for QIAGEN users. These protocols have not been thoroughly tested or optimized by QIAGEN. QIAGEN neither guarantees them nor warrants that they do not infringe the rights of third-parties.

D. Other than expressly stated licenses, QIAGEN makes no warranty that this kit and/or its use(s) do not infringe the rights of third-parties.

E. This kit and its components are licensed for one-time use and may not be reused, refurbished, or resold.

F. QIAGEN specifically disclaims any other licenses, expressed or implied other than those expressly stated.

G. The purchaser and user of the kit agree not to take or permit anyone else to take any steps that could lead to or facilitate any acts prohibited above. QIAGEN may enforce the prohibitions of this Limited License Agreement in any Court, and shall recover all its investigative and Court costs, including attorney fees, in any action to enforce this Limited License Agreement or any of its intellectual property rights relating to the kit and/or its components.

For updated license terms, see www.qiagen.com.

www.qiagen.com

Australia ▪ techservice-au@qiagen.com

Austria ▪ techservice-at@qiagen.com

Belgium ▪ techservice-bnl@qiagen.com

Brazil ▪ suportetecnico.brasil@qiagen.com

Canada ▪ techservice-ca@qiagen.com

China ▪ techservice-cn@qiagen.com

Denmark ▪ techservice-nordic@qiagen.com

Finland ▪ techservice-nordic@qiagen.com

France ▪ techservice-fr@qiagen.com

Germany ▪ techservice-de@qiagen.com

Hong Kong ▪ techservice-hk@qiagen.com

India ▪ techservice-india@qiagen.com

Ireland ▪ techservice-uk@qiagen.com

Italy ▪ techservice-it@qiagen.com

Japan ▪ techservice-jp@qiagen.com

Korea (South) ▪ techservice-kr@qiagen.com

Luxembourg ▪ techservice-bnl@qiagen.com

Mexico ▪ techservice-mx@qiagen.com

The Netherlands ▪ techservice-bnl@qiagen.com

Norway ▪ techservice-nordic@qiagen.com

Singapore ▪ techservice-sg@qiagen.com

Sweden ▪ techservice-nordic@qiagen.com

Switzerland ▪ techservice-ch@qiagen.com

UK ▪ techservice-uk@qiagen.com

USA ▪ techservice-us@qiagen.com

QIAGEN

Sample & Assay