

September 2009

Rotor-Gene[®] ScreenClust HRM[®] Software User Guide

For use with Rotor-Gene cyclers



Sample & Assay Technologies

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

Product Use Limitations	4
Product Warranty and Satisfaction Guarantee	4
Technical Assistance	4
Introduction	6
About this user guide	6
Rotor-Gene ScreenClust HRM Software	7
Steps in data analysis	7
Rotor-Gene operating software and Rotor-Gene ScreenClust HRM Software	9
Installing Rotor-Gene ScreenClust HRM Software	13
Using Rotor-Gene ScreenClust HRM Software	14
Starting a new analysis	14
Unsupervised mode	17
Supervised mode	22
Modifying analysis	26
Saving analysis	27
Post-analysis features	27
Viewing previous analysis files	28
User interface options	30
Other options and help	32
Guidelines	34
Troubleshooting Guide	36
Appendix: Steps in Data Analysis	39
Normalization	39
Generation of a residual plot	40
Principal component analysis	40
Clustering	41
Supervised mode	42
Unsupervised mode	43
References	46

Product Use Limitations

The Rotor-Gene ScreenClust HRM Software is designed to be used to process High Resolution Melt (HRM) data files generated on the Rotor-Gene Q instrument.* Its purpose is to allow for automated calling of genotypes for supervised and unsupervised data sets. Statistical measures for each sample are also provided.

The Rotor-Gene ScreenClust HRM Software is intended for molecular biology applications. This software is neither intended for the diagnosis, prevention, or treatment of a disease, nor has it been validated for such use either alone or in combination with other products. Therefore, the performance characteristics of the product for clinical use (i.e., diagnostic, prognostic, therapeutic, of blood banking) are unknown.

The Rotor-Gene ScreenClust HRM Software can analyze data sets generated by the Rotor-Gene Q system only.*

The Rotor-Gene ScreenClust HRM Software is intended for use by professional users, such as technicians and scientists trained in molecular biological techniques and the operation of the Rotor-Gene Q system.

Product Warranty and Satisfaction Guarantee

QIAGEN guarantees the performance of all products in the manner described in our product literature. The purchaser must determine the suitability of the product for its particular use. Should any product fail to perform satisfactorily due to any reason other than misuse, QIAGEN will replace it free of charge or refund the purchase price. We reserve the right to change, alter, or modify any product to enhance its performance and design. If a QIAGEN product does not meet your expectations, simply call your local Technical Service Department or distributor. We will credit your account or exchange the product — as you wish. Separate conditions apply to QIAGEN scientific instruments, service products, and to products shipped on dry ice. Please inquire for more information.

A copy of QIAGEN terms and conditions can be obtained on request, and is also provided on the back of our invoices. If you have questions about product specifications or performance, please call QIAGEN Technical Services or your local distributor (see back cover or visit www.qiagen.com).

Technical Assistance

At QIAGEN, we pride ourselves on the quality and availability of our technical support. Our Technical Service Departments are staffed by experienced scientists with extensive practical and theoretical expertise in sample and assay

* Compatible also with Rotor-Gene 6000 instruments using Rotor-Gene Software version 1.7 and higher.

technologies and the use of QIAGEN products. If you have any questions or experience any difficulties regarding Rotor-Gene ScreenClust HRM Software or QIAGEN products in general, please do not hesitate to contact us.

QIAGEN customers are a major source of information regarding advanced or specialized uses of our products. This information is helpful to other scientists as well as to the researchers at QIAGEN. We therefore encourage you to contact us if you have any suggestions about product performance or new applications and techniques.

For technical assistance and more information, please see our Technical Support Center at www.qiagen.com/Support or call one of the QIAGEN Technical Service Departments or local distributors (see back cover or visit www.qiagen.com).

Introduction

About this user guide

This user guide provides information about the functions and features of Rotor-Gene ScreenClust HRM Software. Please refer to the *Rotor-Gene Q User Manual* for complete information about the proper care, maintenance, and use of the Rotor-Gene Q cycler.

Information about Rotor-Gene ScreenClust HRM Software is provided in the following sections:

- Introduction
- Rotor-Gene ScreenClust HRM Software
- Installing Rotor-Gene ScreenClust HRM Software
- Using Rotor-Gene ScreenClust HRM Software
- Troubleshooting Guide
- Appendix: Steps in Data Analysis
- References

Rotor-Gene ScreenClust HRM Software

Rotor-Gene ScreenClust HRM Software is a powerful tool for analysis of high-resolution melting (HRM) data from the Rotor-Gene Q or Rotor-Gene 6000 cyclers. By grouping samples into clusters based on their dissociation (melting) curve characteristics, Rotor-Gene ScreenClust HRM Software enables applications such as genotyping and mutation scanning. The number of clusters can either be defined by the user, if they have known controls for each genotype (supervised mode) or the software can aid the user in determining the number of clusters in a sample set (unsupervised mode).

Rotor-Gene ScreenClust HRM Software provides:

- Innovative mathematical approach to HRM analysis
- Highly accurate identification of genotypes in supervised mode
- Automatic detection of new mutations in unsupervised mode
- Robust statistics for classifying and interpreting HRM data
- Minimal effort and standardized processes for data interpretation

Steps in data analysis

HRM analysis on a Rotor-Gene cycler produces raw data that can be further analyzed using Rotor-Gene ScreenClust HRM Software. Rotor-Gene ScreenClust HRM Software analyzes HRM data in 4 steps (see below and Figure 1 for brief description, see page 39 for more detail).

- 1. Normalization**
- 2. Generation of a residual plot**
- 3. Principal component analysis**
- 4. Clustering**

HRM curves can have different starting points, therefore the scale of each melt is different (Figure 1A). Rotor-Gene ScreenClust HRM Software only compares samples that are on the same scale, which is achieved by normalization. Raw data are normalized by applying curve scaling to a line of best fit so that the highest fluorescence value is equal to 100 and the lowest is equal to zero (Figure 1B). Next, the curves are differentiated and a composite median curve is constructed using the median fluorescence of all samples. The melt traces for each sample are subtracted from this composite median curve to draw a residual plot (Figure 1C). Consecutively, the individual sample characteristics are extracted by principal component analysis from the residual plot. Principal component analysis is a well-established method of data analysis. However, Rotor-Gene ScreenClust HRM Software is the first software application to apply principal component analysis to HRM data. Principal component analysis

highlights similarities and differences in the data and is used to create a cluster plot (Figure 1D).

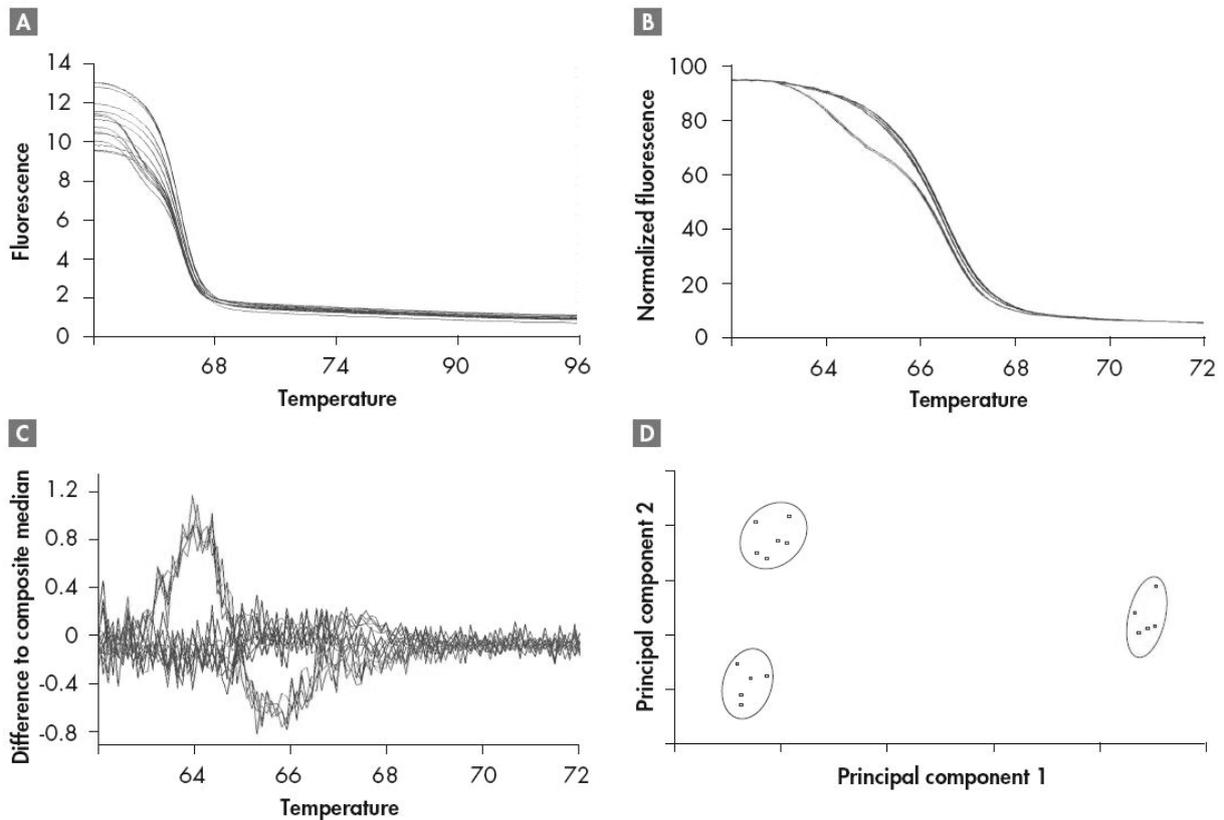


Figure 1. Powerful data analysis enables identification of a class IV SNP. A human A/T SNP in the AHRR gene was analyzed using genomic DNA from wild-type, homozygous mutant, and heterozygous samples. Experiments were performed using the Type-it[®] HRM PCR Kit and a Rotor-Gene Q cyclor with a HRM channel. Data analysis was performed with the unsupervised mode of Rotor-Gene ScreenClust HRM Software. A/T polymorphisms (class IV SNPs) are most difficult to discriminate due to minute differences between homozygote alleles (in this example, less than 0.1°C). **A** HRM raw data, **B** the normalized melting curve, **C** the residual plot, and **D** the cluster plot are shown. All pseudo-unknowns were correctly clustered according to genotype.

Rotor-Gene ScreenClust HRM Software performs clustering (grouping) of data according to allele in either supervised or unsupervised mode. Supervised mode is often used for SNP genotyping, where the genotypes are known. In supervised mode, the user assigns one or more control samples for each cluster and the software classifies (autocalls) all unknown samples to clusters according to their characteristics.

The unsupervised mode is used when there is no or only partial prior knowledge of the genotypes present in the samples. In unsupervised mode, the software calculates the optimal number of clusters by itself. This feature is an excellent tool for the discovery of new polymorphisms.

In addition to the easy-to-interpret cluster plot, Rotor-Gene ScreenClust HRM Software provides statistical probabilities and typicalities in a results table to allow easy comparison of results from different experiments.

The software guides the user through all the steps, giving information about any choices that can be made at each step. Each step is described in more detail in the Appendix (page 39). For more information on the applied statistical methods, refer to the references on page 46.

Rotor-Gene operating software and Rotor-Gene ScreenClust HRM Software

Rotor-Gene operating software is supplied with the Rotor-Gene cyclers. The Rotor-Gene operating software is involved in the running of the cycler and also includes a basic HRM analysis option. Rotor-Gene ScreenClust HRM Software is an extension to the Rotor-Gene operating software. Rotor-Gene ScreenClust HRM Software is not involved in the running of the cycler and is used to analyze raw data which comes from a HRM run. Rotor-Gene ScreenClust HRM Software works with the Rotor-Gene operating software to provide increased functionality and more sophisticated analysis of HRM data. The features provided by the Rotor-Gene operating software and the Rotor-Gene ScreenClust HRM Software are summarized in the table, below.

Software features

Feature	Rotor-Gene ScreenClust HRM Software	Rotor-Gene operating software
Acquires HRM melt curve		✓
Displays raw HRM data	✓	✓
Normalizes melt curve	✓	✓
Provides difference plot using one reference sample		✓
Provides a residual plot using a composite median curve reference	✓	
Provides principal component and cluster plots	✓	
Autocalls genotypes based on one control	✓	✓
Autocalls genotypes based on a group of controls	✓	
Automatically groups genotypes without controls	✓	
Provides probabilities of belonging to each cluster	✓	
Provides typicalities of samples within a cluster	✓	
Generates an analysis report	✓	✓

Both the HRM analysis option of the Rotor-Gene operating software and the Rotor-Gene ScreenClust HRM Software can be used to perform normalization of HRM data (page 39).

The HRM analysis option of the Rotor-Gene operating software uses difference plots for HRM analysis. Difference plots can cause unreliable, difficult-to-

interpret results, and time-consuming manual interpretation may be necessary, particularly when analyzing a large group of samples. In contrast, Rotor-Gene ScreenClust HRM Software performs analysis using residual plots (page 40). Residual plots maximize amount of information that can be extracted from the data. This method also avoids the bias of only using one control sample to achieve a difference plot. However, the main difference between the Rotor-Gene operating software and Rotor-Gene ScreenClust HRM Software is that when using Rotor-Gene ScreenClust HRM Software the residual plot is only an intermediate step in whole data analysis workflow, whereas when using the Rotor-Gene operating software the difference plots represent the final basis for autocalling.

Rotor-Gene ScreenClust HRM Software achieves a more detailed characterization of the HRM melting curves by extracting information from the residual plot using principle component analysis (PCA) (page 40). PCA is a common statistical tool used for characterization of multidimensional data in applications such as image analysis or data mining of large-scale gene expression studies, including microarrays.

Both the HRM analysis option of the Rotor-Gene operating software and the Rotor-Gene ScreenClust HRM Software perform autocalling of genotypes. Autocalling using Rotor-Gene ScreenClust HRM Software provides the advantages of a stringent statistical approach and autocalling based on multiple control samples (page 42). In addition, unlike using difference plots, this approach enables the user to measure the relationship of the sample to other clusters.

Analysis using the unsupervised mode is an important benefit of Rotor-Gene ScreenClust HRM Software. In unsupervised mode, unknown samples in a large data set can be classified without using any control samples (page 43). This provides a powerful tool for the identification of unknown mutations.

Rotor-Gene ScreenClust HRM Software uses established methods of statistical analysis for each sample, indicating the probability that a sample belongs to a particular cluster and how well it belongs in the cluster (typicality). Reliable statistical analysis is critical for correct interpretation of results and is vital for analysis of groups of samples.

Due to the features described, the Rotor-Gene operating software and the Rotor-Gene ScreenClust HRM Software are suited to different applications (see table below).

Software suitable for different applications

Application	Rotor-Gene ScreenClust HRM Software	Rotor-Gene operating software
Genotyping (e.g., SNPs)	++	+
Mutation scanning	++	+
Species identification	++	+
Quantitative analysis (methylation, allelic ratios)	○	+

++ strongly recommended; + suitable; ○ less suitable

Rotor-Gene ScreenClust HRM Software is highly suitable for screening approaches in SNP genotyping, mutation scanning, or pathogen typing because it provides statistics for precise, large-scale curve characterization and classification.

Rotor-Gene ScreenClust HRM Software cannot be used for quantification of allelic ratios (e.g., in the analysis of somatic mutations or of levels of DNA methylation). Such applications require the generation of standard curves and quantification using a calibration formula. Rotor-Gene ScreenClust HRM Software is designed to classify samples into clusters, which is a qualitative and highly specific analysis, but not a quantitative approach. Rotor-Gene ScreenClust HRM Software can group well-defined levels of allele ratios or DNA methylation amounts into corresponding clusters. However, for quantifying unknown levels of methylation, the Rotor-Gene operating software is better suited, in combination with software such as Microsoft® Excel® for the calculation of standard curves.

Installing Rotor-Gene ScreenClust HRM Software

Minimum PC requirements

The table below lists the minimum PC requirements for Rotor-Gene ScreenClust HRM Software.

PC feature	Requirements
Supported operating systems	Windows® XP: WinXP/Pro; Windows Vista®: Vista Business edition
Processor	Pentium® IV 2.6 GHz (desktop) or Pentium M 1.6 GHz (laptop) (Windows XP); Intel® Core™ 2 Duo processor (Windows Vista)
Disk space	80 GB of available hard drive space
Memory	512 MB of RAM (Windows XP); 1 GB of RAM (Windows Vista)
Monitor/graphics card	1024 x 768 monitor (minimum), 1280 x 1024 (preferred); DirectX 9 capable (Windows Vista)
CD-ROM drive	DVD/CD-ROM drive

Installation Procedure

A detailed description of the installation procedure can be found in the installation guide provided with the software.

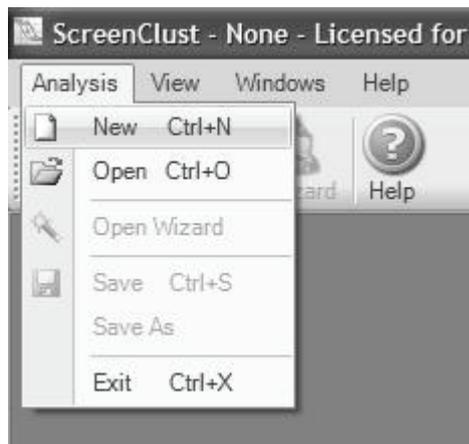
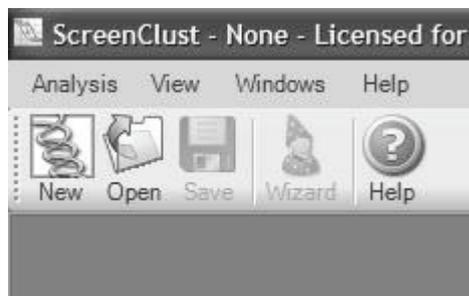
Using Rotor-Gene ScreenClust HRM Software

Starting a new analysis

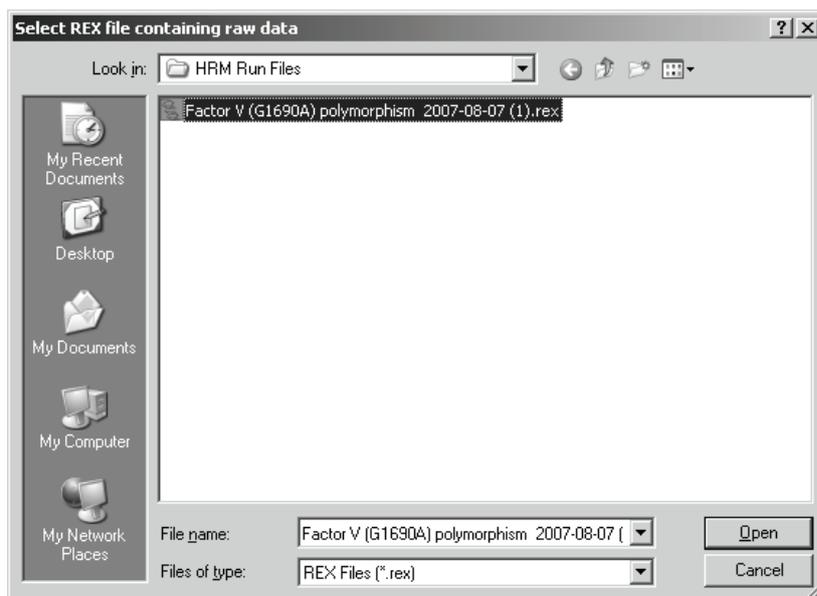
1. Open Rotor-Gene ScreenClust HRM Software by clicking the desktop icon or via the Windows Start menu.



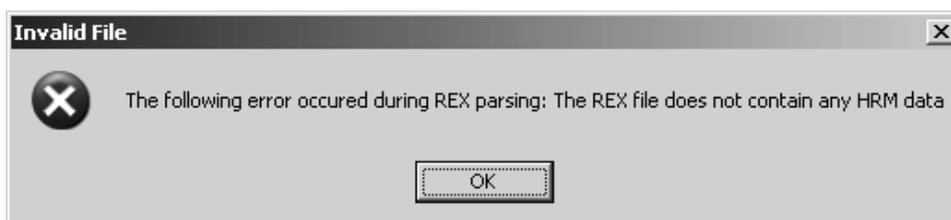
2. Click the "New" icon from the tool bar or select "New" from the "Analysis" menu.



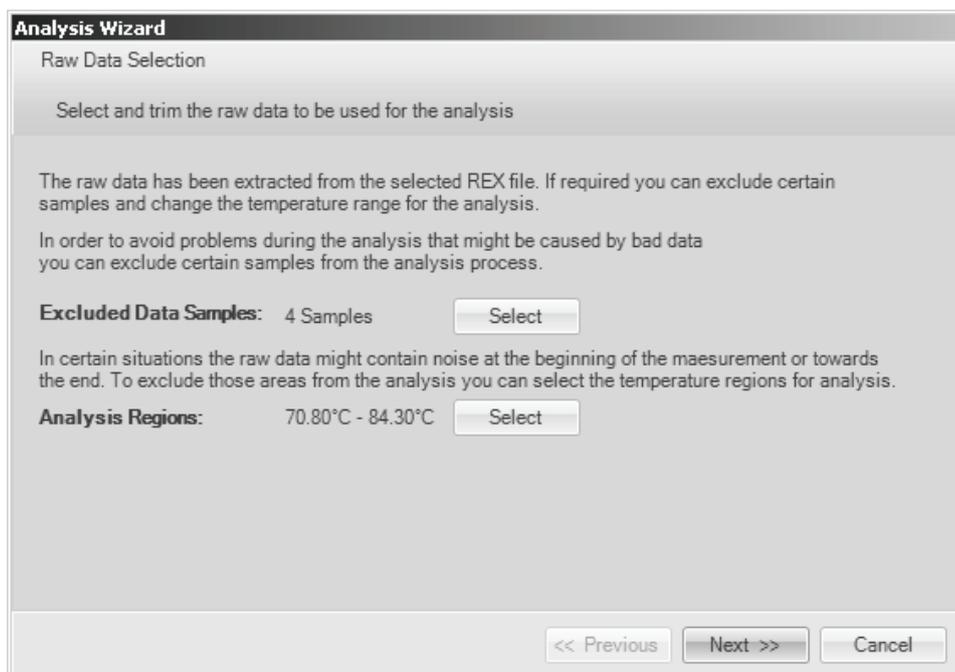
3. Select the *.rex file for analysis from the "File" menu and click "Open".



Ensure that the *.rex file contains HRM data. This is the only file type that can be analyzed using Rotor-Gene ScreenClust HRM Software. A warning message will appear if there is no HRM data in the *.rex file.



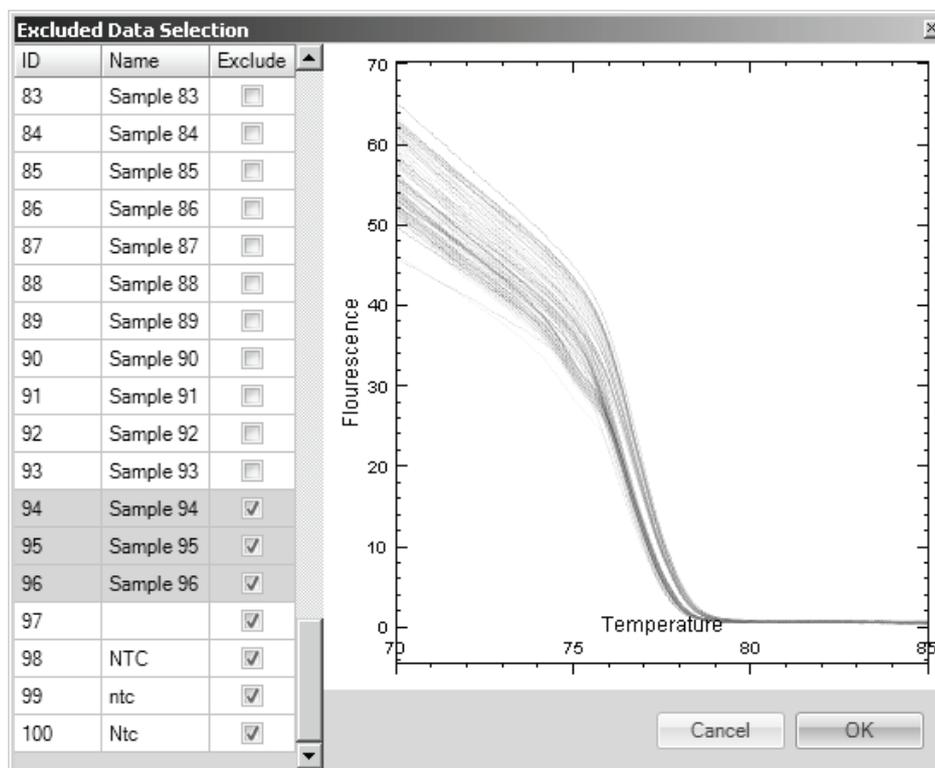
4. The Analysis Wizard automatically begins Raw Data Selection.



5. Exclude any samples that should not be analyzed. Multiple samples can be highlighted and excluded by holding the "Ctrl" key while

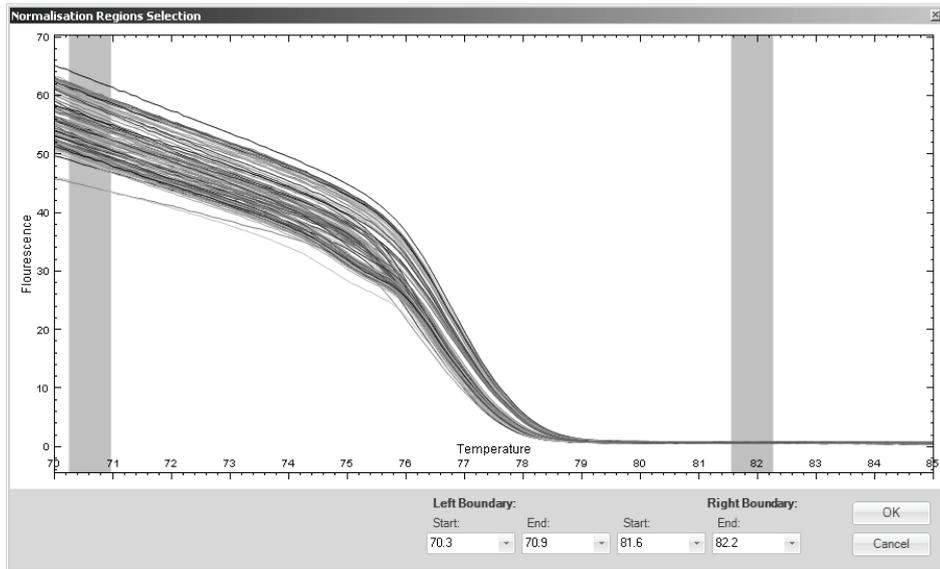
selecting samples with the mouse or by clicking and dragging the mouse over the samples.

Ensure that samples that will be analyzed have melt curve characteristics. Samples with poor melt transitions will negatively affect the result. Samples with the names "NTC", "ntc", or "Ntc" (no template controls), or positions with no names are automatically excluded.



- 6. Refine the Normalization Regions for melt curve analysis if desired. Exclude data by moving the blue shaded areas using the mouse or enter the exact values into the start and end temperature boundary fields. Once the temperature range has been defined, click "OK".**

This option should only be used to remove significant variations in the melt curve such as inflection points not associated with the actual melt product(s) of interest. It is recommended that as much background fluorescence at the beginning of the melt is retained to improve analysis. To further improve analysis, remove as much as half of the background fluorescence at the end of the melt. The background fluorescence at the end of the melt contains very little information and can influence the principal component analysis if too much is present.



7. Click “Next” to proceed. Two options are available: supervised mode and unsupervised mode.

Supervised mode: genotypes are known and controls are available for each one.

Unsupervised mode: genotypes may not be known and/or controls are not available for all genotypes.

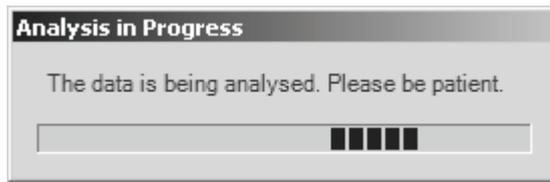
The figure shows an 'Analysis Wizard' window. The title is 'Analysis Wizard'. Below the title is the section 'Analysis Mode' with the instruction 'Select the main analysis mode'. A paragraph explains the two modes: supervised and unsupervised. Below this, the 'Analysis Mode:' section has two radio buttons: 'Supervised' (unselected) and 'Unsupervised' (selected). Underneath is the 'Known Clusters:' section with a table:

	Name	Controls
*		

At the bottom of the window are three buttons: '<< Previous', 'Next >>' (which is highlighted), and 'Cancel'.

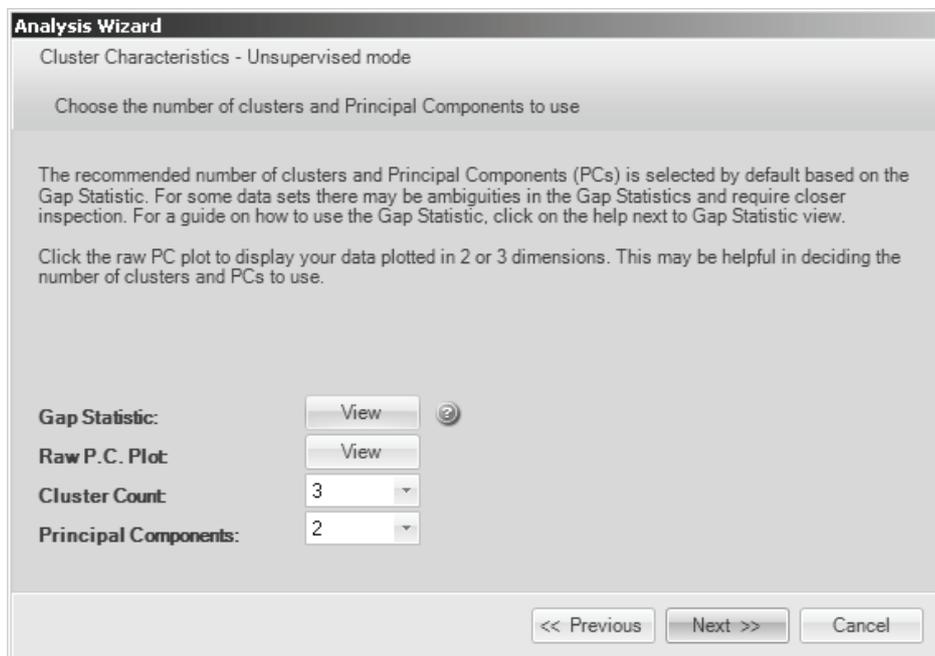
Unsupervised mode

- 1. Select “Unsupervised” and then “Next”. A window appears showing that the analysis is in progress.**



- In unsupervised mode, the number of genotypes is assumed to be unknown. The software calculates a default optimal number of genotypes (clusters) based on the gap statistic (see page 43). The Principal Component dimensionality is also automatically calculated as a default.**

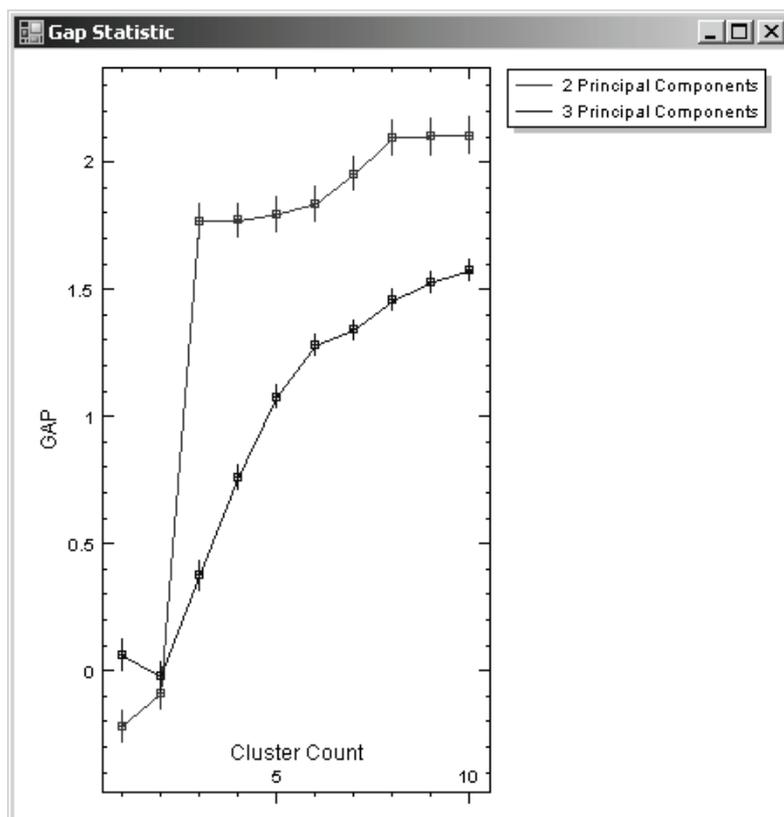
In most cases, the number of clusters calculated matches the number of genotypes present in the data. However, the cluster number can also be changed manually, if desired (see the following steps).



- The gap statistic graph can be viewed by clicking the "View" button to the right of "Gap Statistic". Viewing the gap statistic graph is an advanced feature that can be used to view the data if desired (e.g., to determine whether to manually change the cluster count).**

In the gap statistic graph, Cluster Count (x-axis) is plotted against the gap score (y-axis). The point where the curve plateaus represents the optimal number of clusters (see page 43). The red curve represents the data as observed with 2 principal components; the blue curve represents data observed with 3 principal components. The optimal number of principal components is selected as the curve with the greatest gap score (y-axis).

Further observation of the gap statistic is recommended, as it may indicate the presence of a higher or lower number of clusters. A maximum of 10 clusters is detectable.

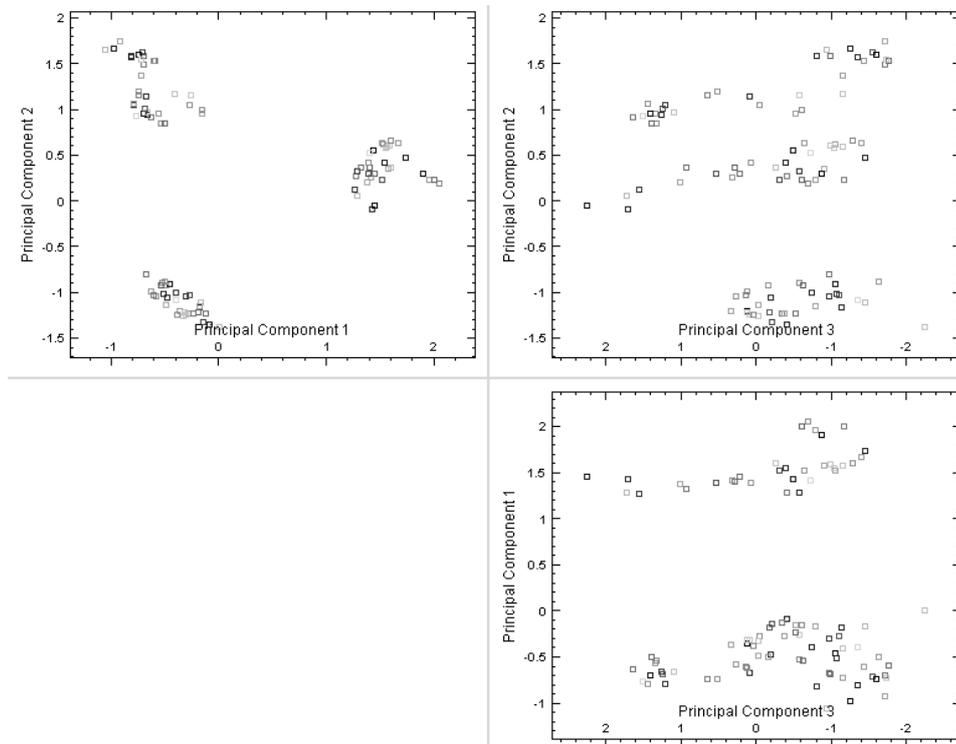


4. **The raw principal components plot can be viewed by clicking the “View” button to the right of “Raw P.C. Plot”. Viewing the raw principal components plot is an advanced feature that can be used to view the data if desired. Together with the gap statistic graph, it enables a decision on the optimal cluster number or PC dimensionality in case of ambiguities (see Appendix, page 39). Depending on the PC dimensionality chosen, the raw principal components plot shows either 2 or all 3 combinations of principal components.**

Generally, for high-quality data sets with small numbers of clusters (usually 2–3 clusters), 2 dimensions are sufficient. Using 3 PC dimensions in these circumstances will often produce poorer results. This is because PC1 and PC2 are usually sufficient to represent most or all the real differences between samples. PC3 may only represent noise, such as the minute differences between replicate samples. Inclusion of PC3 will lower typicality scores and may reduce the accuracy of the assignment of samples into clusters.

For data sets with complex melt curves and/or larger number of clusters, using all 3 PC dimensions provides better separation of samples. In some cases, PC3 represents real differences between samples and an additional dimension provides better separation. Note that using 3 PC dimensions may also improve the separation of poor data, but this is highly dependent on individual data sets.

The default calculation of the PC dimensionality in Rotor-Gene ScreenClust HRM Software takes all these factors into account to make a recommendation of the optimal number of PC dimensions to use.



5. Click "Next" to complete the analysis or "Previous" to make further changes to the analysis criteria.
6. Results are shown as a cluster plot and a table of results.

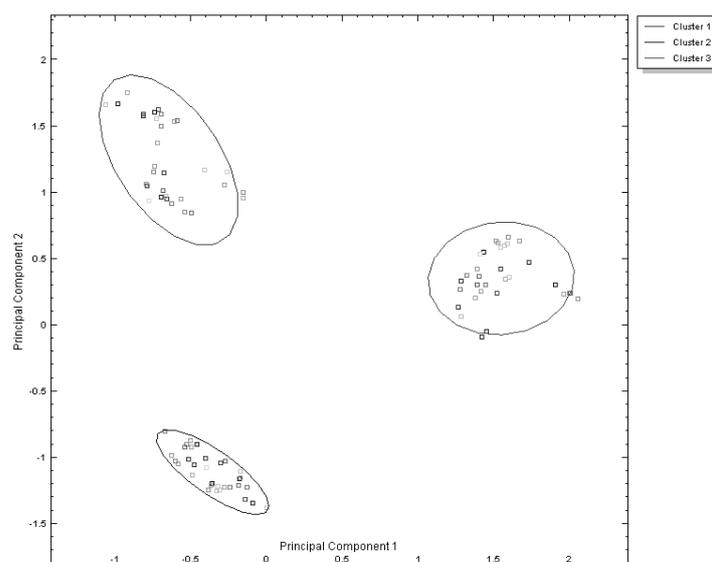
The numbering of each cluster is random and does not represent any particular order. A new analysis of the same data set may generate a different cluster number for the same group of samples. The probabilities and typicalities remain the same if none of the analysis parameters are changed. The entries in the table of results are described below.

Table of results

Column heading	Description
ID	Sample sequence number from the *.rex file
Name	Sample name
Cluster	The sample has been classified to this cluster based on the highest probability value
Typicality	Indicates how well the sample fits into the cluster
Prob. -Cluster #	Probability that the sample belongs to that cluster (the probability that the sample belongs to each of the defined clusters is shown)

ID	Name	Cluster	Typicality	Prob. - Cluster 1	Prob. - Cluster 2	Prob. - Cluster 3
1	Sample 1	Cluster 1	0.50579779	1.00000000	0.00000000	0.00000000
2	Sample 2	Cluster 1	0.13921212	1.00000000	0.00000000	0.00000000
3	Sample 3	Cluster 1	0.23728311	1.00000000	0.00000000	0.00000000
4	Sample 4	Cluster 1	0.53413622	1.00000000	0.00000000	0.00000000
5	Sample 5	Cluster 1	0.21509718	1.00000000	0.00000000	0.00000000
6	Sample 6	Cluster 1	0.45688176	1.00000000	0.00000000	0.00000000

7. The cluster plot is drawn either in 2 principal component dimensions or 3 principal component dimensions (i.e., PC 1 vs. PC 2, PC 2 vs. PC 3, and PC 1 vs. PC 3). Ellipses represent each cluster. The center point of the cluster is calculated using k-means but is not shown (for more information, see Appendix, page 43).



Supervised mode

1. Select "Supervised". The "Known Clusters" table appears.

Analysis Wizard

Analysis Mode

Select the main analysis mode

There are two possible modes for the cluster analysis: supervised or unsupervised. The supervised mode is used if the set of raw data samples contains known control samples and all other samples should be categorized into those known clusters. If the raw data does not contain any controls, the unsupervised mode will categorize all data into a number of clusters which best fit the whole data set.

Analysis Mode: Supervised Unsupervised

Known Clusters:

	Name	Controls
▶*		

<< Previous Next >> Cancel

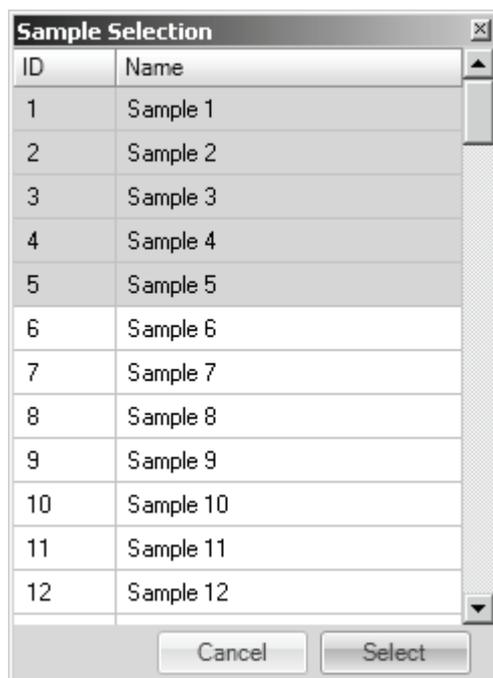
2. Enter the genotype name.

Known Clusters:

	Name	Controls
▶*	Wild Type	
*		

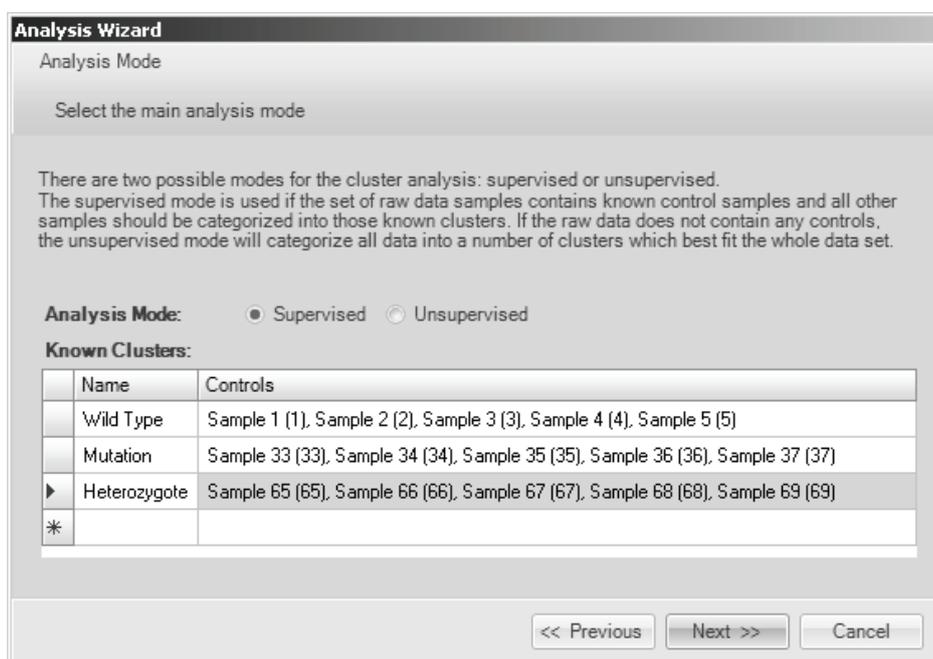
3. Select controls for the genotype from the "Sample Selection" window. Samples that are selected as controls for a genotype will not be available for selection from the "Sample Selection" window for subsequent genotypes. A list of samples selected appears in the "Known Clusters" table with their run position numbers.

Several control samples can be selected by dragging the cursor over the sample names or by selecting individual samples while keeping the CTRL key pressed. In supervised mode, at least one control sample must be selected. If a control sample or a "Known Cluster" name is missing, a red exclamation mark appears.



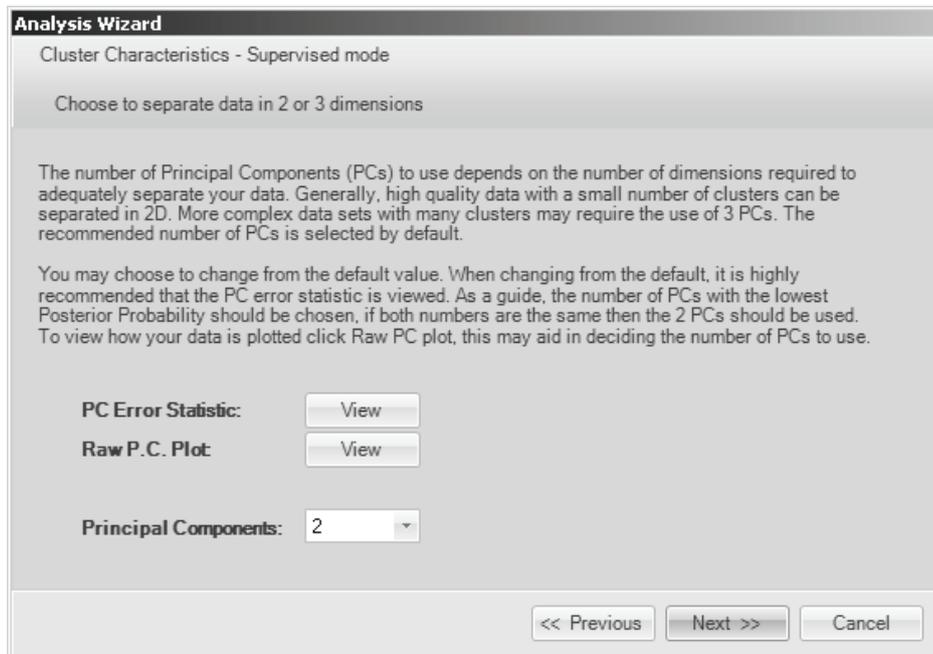
4. When all the genotypes and controls have been named and allocated, click “Next”.

To delete a whole row of information, click the triangle symbol to the left of the row and press the “delete” key on the keyboard.



5. In this window, the number of principal components required is calculated and displayed.

The software automatically calculates a default optimal number for PC dimensionality. In most cases, this correctly reflects the complexity of the sample set. However, the number of PC dimensions can be changed manually if desired (see steps below).



6. To view the “PC Error Statistic” window, click the “View” button to the right of “PC Error Statistic”. The “PC Error Statistic” window appears. Viewing the PC Error Statistic is an advanced feature that can be used to decide on the optimal number of PC dimensions, if desired. The PC dimensions calculation (cross-validation) can be viewed in the “PC Error Statistic” window. The PC dimensionality selected should have the lowest error value. If both error values are equal, the PC dimensionality with the lowest posterior value should be selected. If both posterior values are zero, the lowest PC dimension should be selected.

Principal Components	Error	Posterior
2	0.0000000000	0.0000000000
3	0.0000000000	0.0000000000

7. Select “Next” to complete the analysis or “Previous” to make further changes to the analysis criteria.

8. Results are shown as a cluster plot and a table of results.

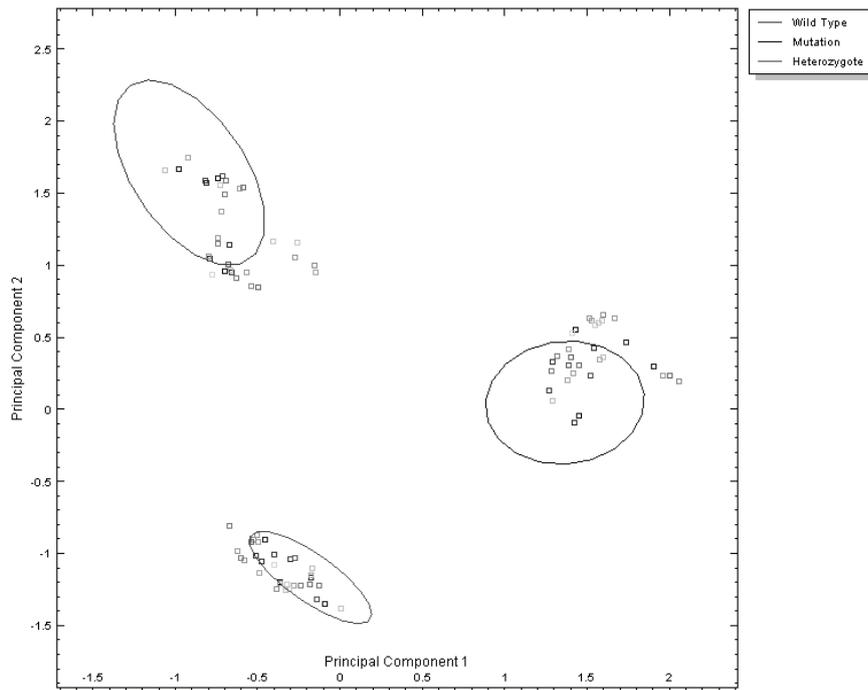
The entries in the table of results are described below.

Table of results

Column heading	Description
ID	Sample sequence number from the *.rex file
Name	Sample name
Cluster	The sample has been classified to this cluster based on the highest probability value
Typicality	Indicates how well the sample fits into the cluster
Prob. -Cluster #	Probability that the sample belongs to that cluster (the probability that the sample belongs to each of the defined clusters is shown)

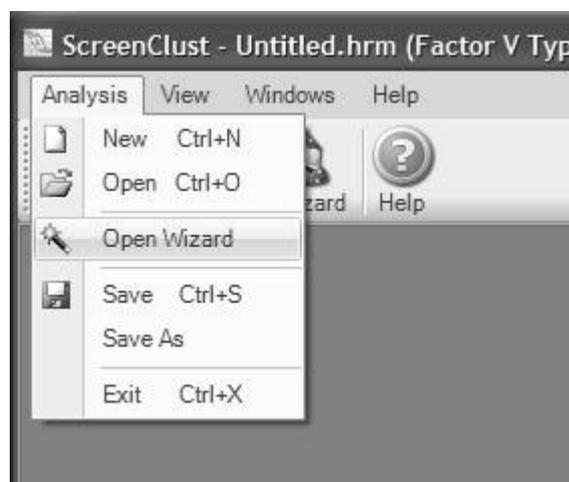
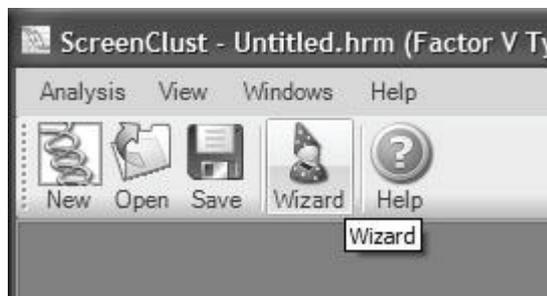
ID	Name	Cluster	Typicality	Prob. - Wild Type	Prob. - Mutation	Prob. - Heterozy...
1	Sample 1	Wild Type	0.51018410	1.00000000	0.00000000	0.00000000
2	Sample 2	Wild Type	0.13698268	1.00000000	0.00000000	0.00000000
3	Sample 3	Wild Type	0.74975411	1.00000000	0.00000000	0.00000000
4	Sample 4	Wild Type	0.50040615	1.00000000	0.00000000	0.00000000
5	Sample 5	Wild Type	0.45418799	1.00000000	0.00000000	0.00000000
6	Sample 6	Wild Type	0.01705566	1.00000000	0.00000000	0.00000000
7	Sample 7	Wild Type	0.00030817	1.00000000	0.00000000	0.00000000
8	Sample 8	Wild Type	0.00100477	1.00000000	0.00000000	0.00000000

9. The cluster plot is drawn either in 2 principal component dimensions or 3 principal component dimensions (i.e., PC 1 vs. PC 2, PC 2 vs. PC 3, and PC 1 vs. PC 3). Ellipses represent each cluster. The center point of the cluster is calculated as the mean point between the control samples. If only one control is used per genotype, then that sample is used as the center point to allocate unknown samples to clusters (for more information, see Appendix, page 42).



Modifying analysis

1. To modify the variables and samples analyzed, click the “Wizard” icon on the tool bar or select “Open Wizard” from the “Analysis” menu.

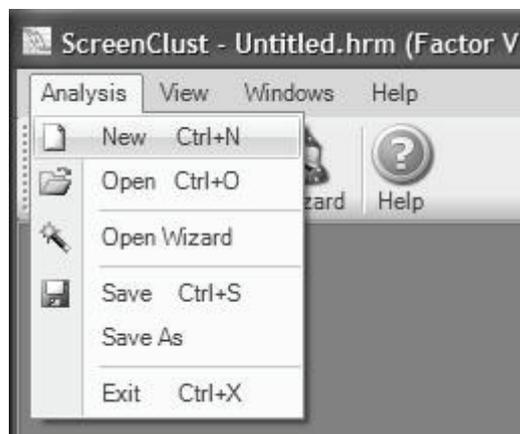
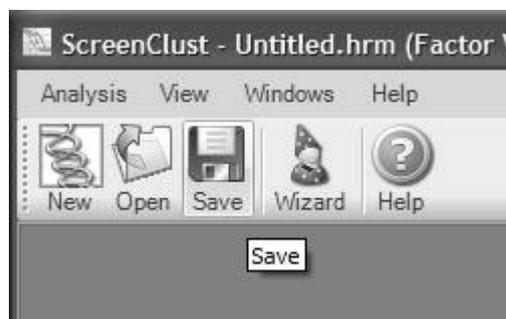


2. **The analysis can be changed from supervised mode to unsupervised mode if desired. All genotypes and controls entered in the “Known Clusters” table are saved. Any modification to the samples, temperature range, or controls will affect the results of the analysis.**

Removal of samples from the analysis will alter the principal components and the principal component plot. Therefore, the clustering and probabilities and typicalities will be altered. In some cases, samples may move to a different cluster.

Saving analysis

To save the final analysis as a Rotor-Gene ScreenClust HRM Analysis File (*.hrm), click the “Save” icon on the tool bar or select “Save” from the “Analysis” menu.

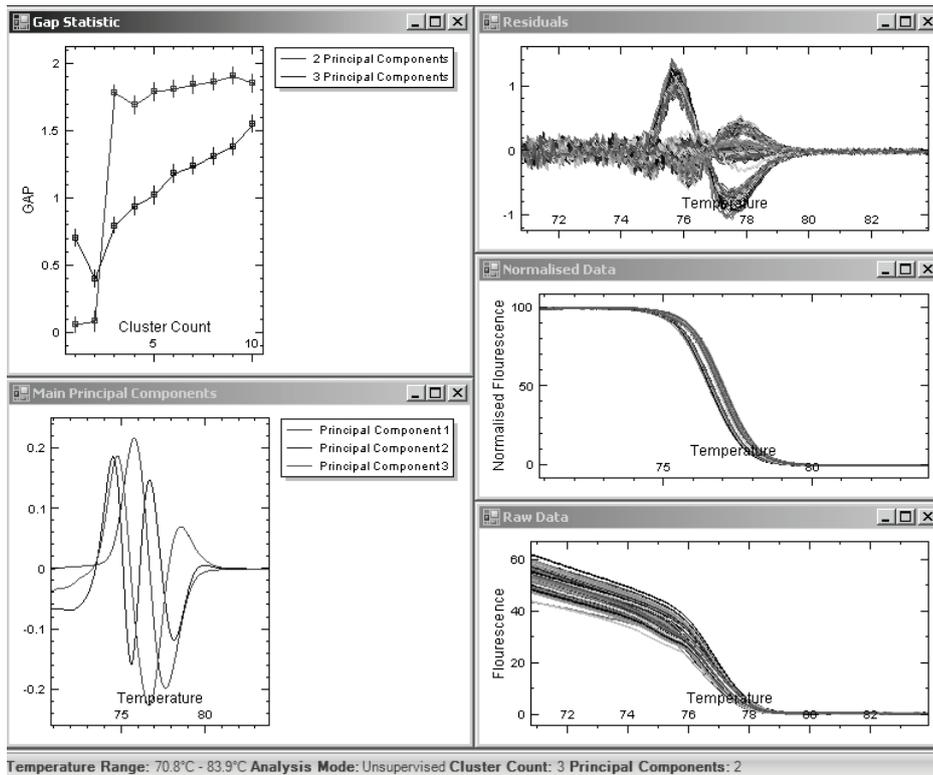


Post-analysis features

To view various parts of the analysis, select from the options in the “View” menu. Raw data, normalized data, residuals, principal components, main principal components plots in 2 or 3 dimensions, the gap statistic (unsupervised mode), PC error statistic (supervised mode), cluster plot, and the table of results can be viewed. If desired, the user can make changes to the analysis options based on observations of the analysis.

The status bar at the bottom of the main window indicates the analysis options

selected. This bar can be removed by deselecting “Status Bar” in the “View” menu. A tick next to “Status Bar” indicates that the bar is visible.

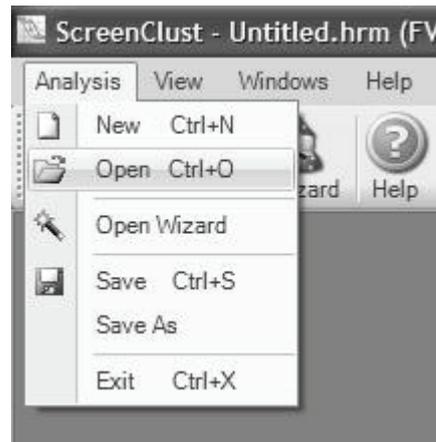
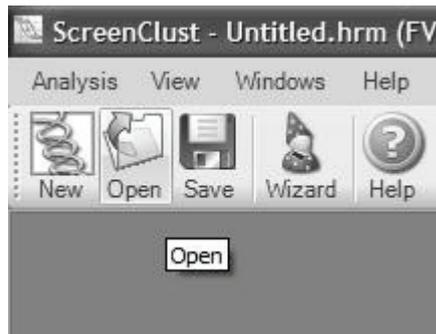


To view a report, select “Report” from the “View” menu. The report shows the table of results and images of the raw data graph and the cluster plot. The report can be printed directly from the software or exported to an Excel® (*.xls) or a *.pdf file.

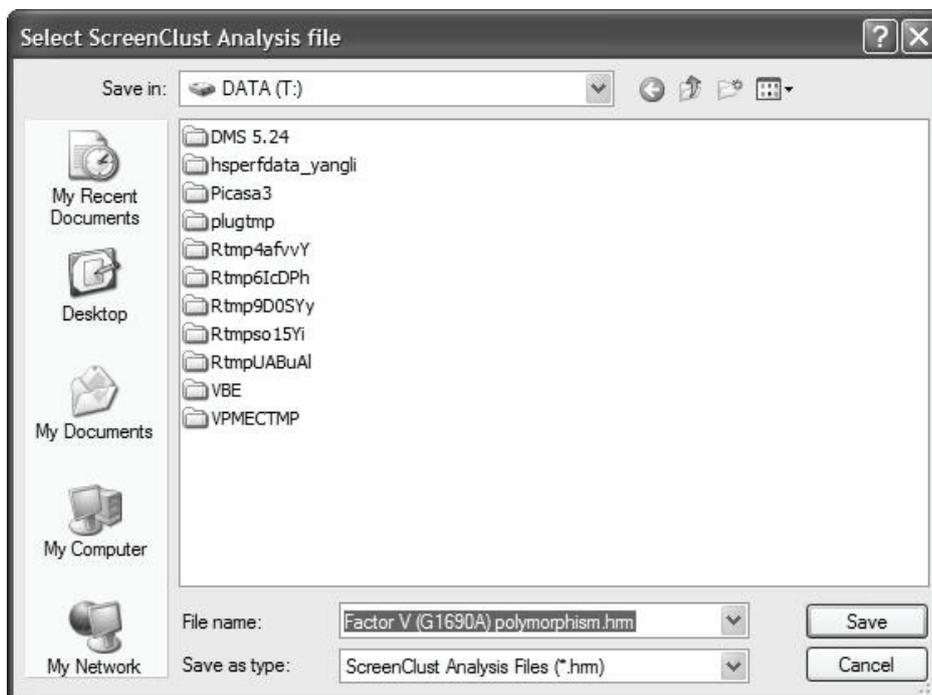
ID	Name	Cluster	Typicality	Wild-type	homozygote	heterozygote
1	Sample 1	Wild Type	0.51018410	1.00000000	0.00000000	0.00000000
2	Sample 2	Wild Type	0.13698268	1.00000000	0.00000000	0.00000000
3	Sample 3	Wild Type	0.74975411	1.00000000	0.00000000	0.00000000
4	Sample 4	Wild Type	0.50040615	1.00000000	0.00000000	0.00000000
5	Sample 5	Wild Type	0.45418799	1.00000000	0.00000000	0.00000000
6	Sample 6	Wild Type	0.01705566	1.00000000	0.00000000	0.00000000
7	Sample 7	Wild Type	0.00030817	1.00000000	0.00000000	0.00000000
8	Sample 8	Wild Type	0.00100477	1.00000000	0.00000000	0.00000000
9	Sample 9	Wild Type	0.04629011	1.00000000	0.00000000	0.00000000
10	Sample 10	Wild Type	0.02051007	1.00000000	0.00000000	0.00000000
11	Sample 11	Wild Type	0.07613923	1.00000000	0.00000000	0.00000000

Viewing previous analysis files

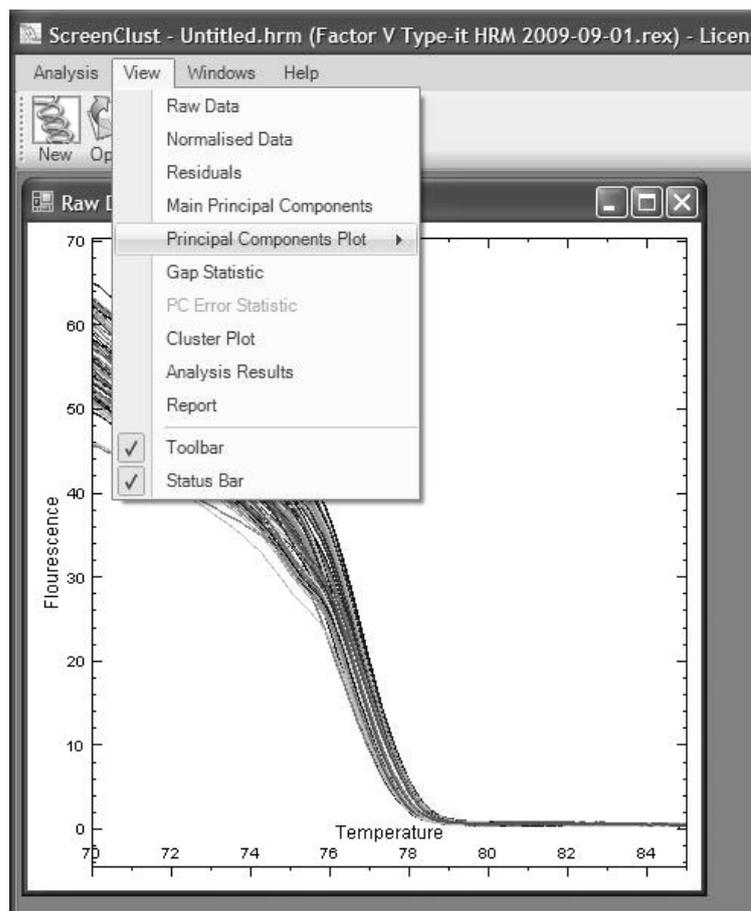
1. To view a previous Rotor-Gene ScreenClust HRM Analysis File (*.hrm), click the “Open” icon on the tool bar or select “Open” from the “Analysis” menu.



2. Select the *.hrm file from the file menu and click "Open".



3. The raw data graph will appear on the screen. To view the cluster plot or table of results, select "Cluster Plot" or "Report" from the "View" menu.

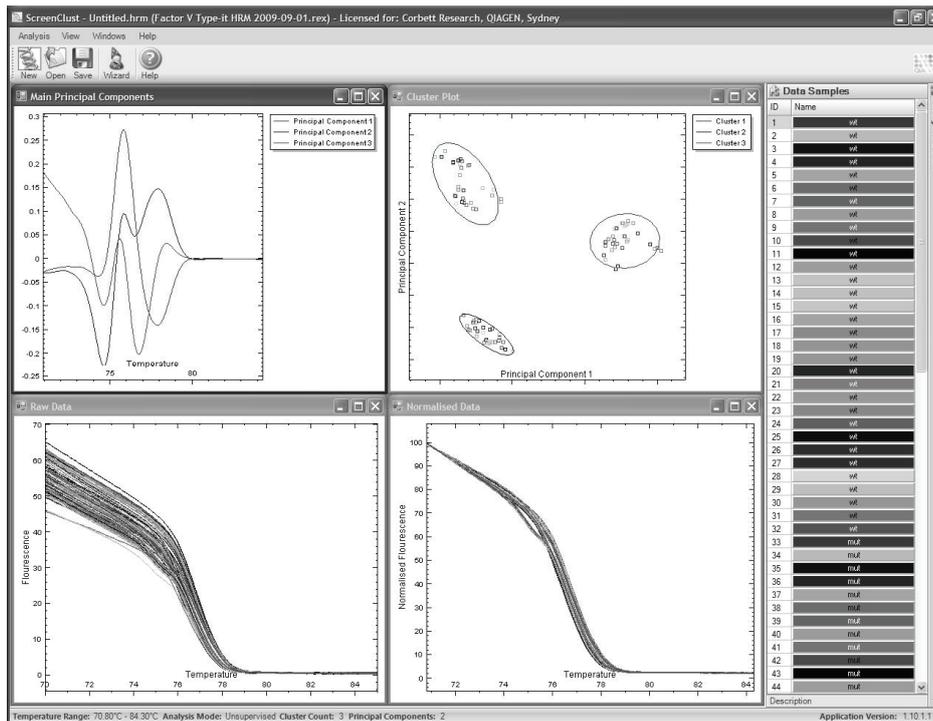


4. If desired, the analysis file can be reanalyzed. Click the “Wizard” icon on the tool bar or select “Open Wizard” from the “Analysis” menu.

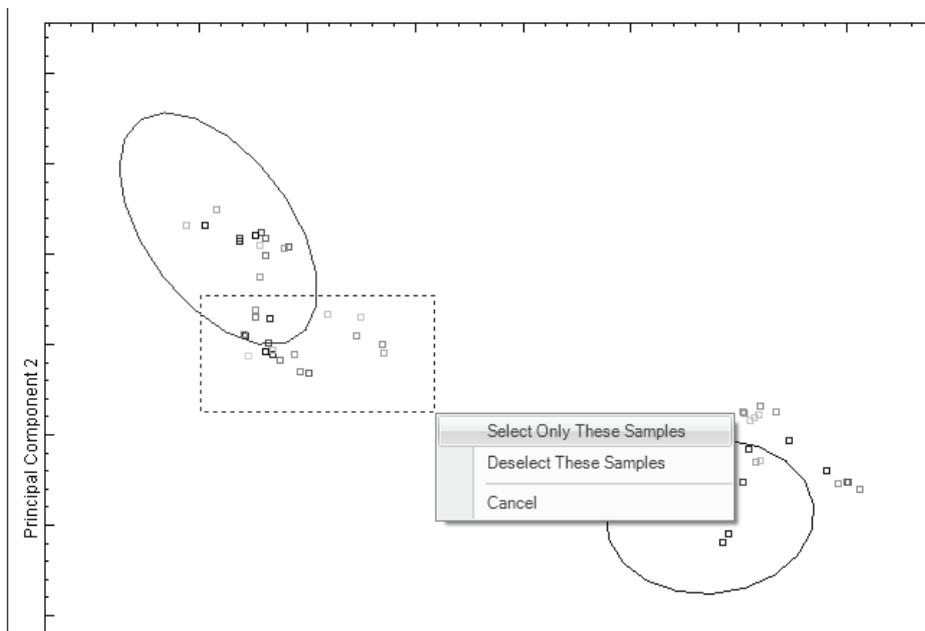
User interface options

Sample selection

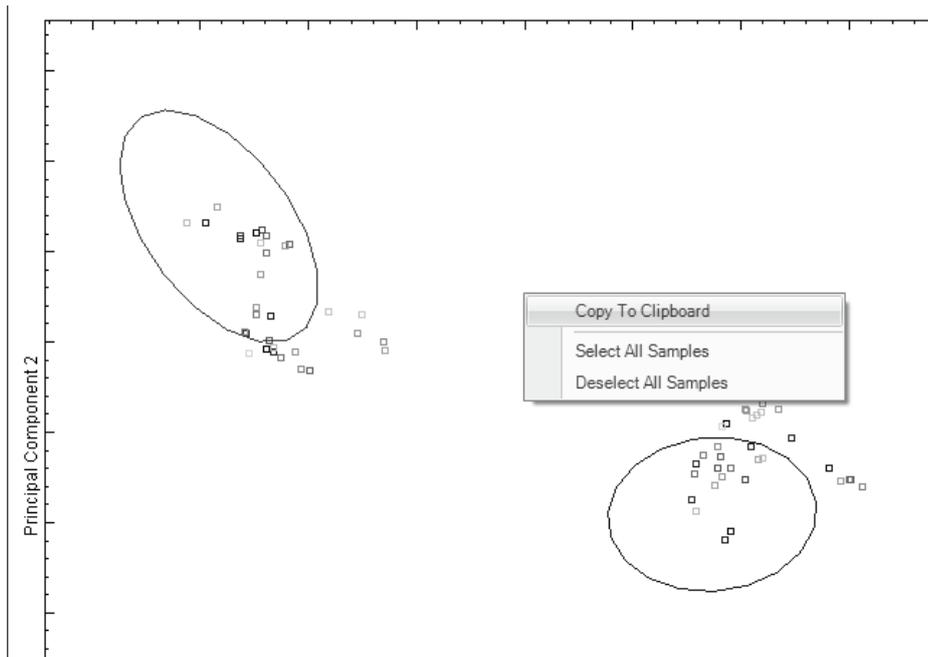
To include or exclude samples from view, use the “Sample Selection” tool bar. Samples that are deselected appear lightly shaded. The “Sample Selection” tool bar can be used for any of the analysis plots including raw data, normalized data, residuals, and cluster plots. To select multiple samples, drag the cursor over the sample name boxes in the tool bar.



Samples can also be selected by holding down the left mouse button and dragging the cursor over the samples. Then choose “Select Only These Samples” or “Deselect These Samples” in the menu that appears.

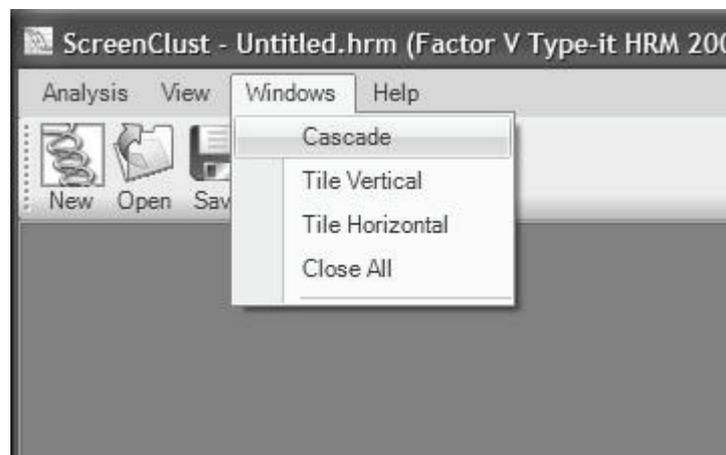


To copy a graph to the Clipboard, place the cursor over the graph and click the right mouse button. Select “Copy To Clipboard” in the menu that appears. The copied image can then be pasted into third-party publishing software if desired.



Windows selection

Each graph or table can be viewed on one page using the “Windows” menu. Individual analysis pages can be aligned by selecting “Cascade”, “Tile Vertical”, or “Tile Horizontal”. To remove all the pages from view, select “Close All”.

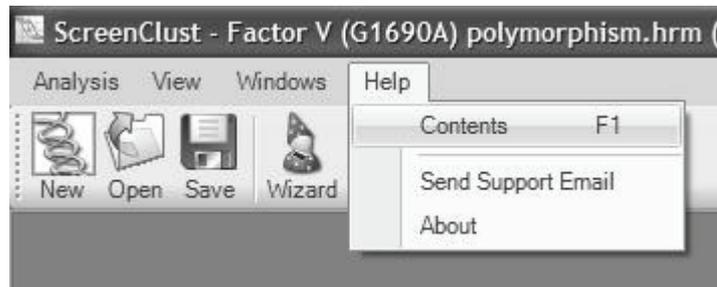


Other options and help

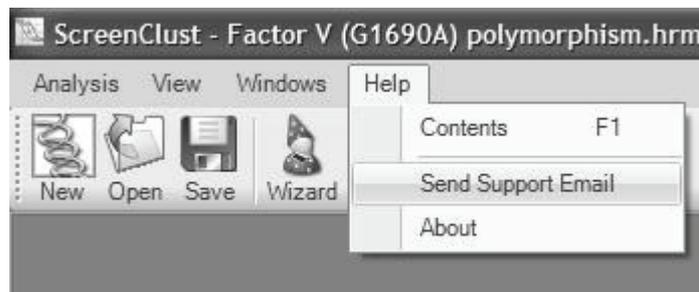
Opened Rotor-Gene ScreenClust HRM Analysis Files can be saved under a new name and to a new location by selecting “Save As” from the “Analysis” menu.

To exit Rotor-Gene ScreenClust HRM Software, select “Exit” from the “Analysis” menu.

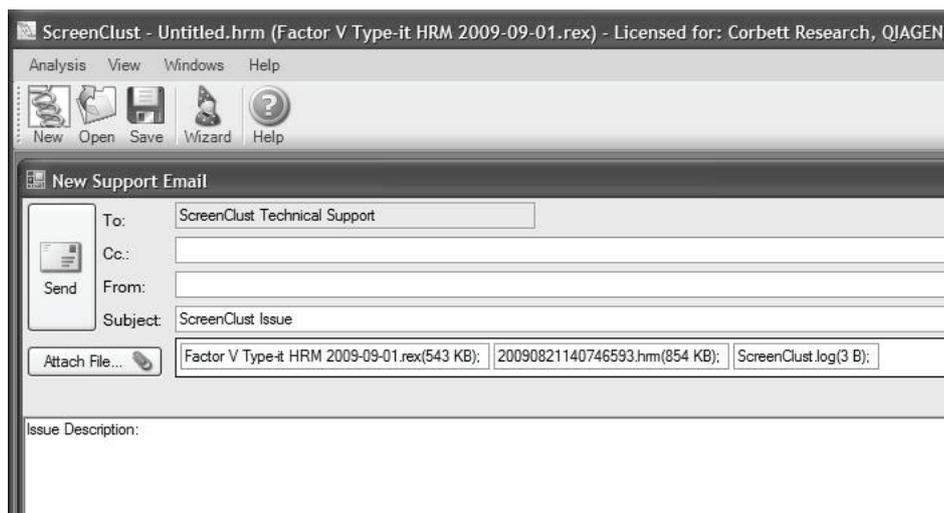
To access the user guide via the software, click the “Help” icon in the tool bar or select “Contents” from the “Help” menu.



To send a support email, select “Send Support Email” from the “Help” menu.



The Rotor-Gene ScreenClust HRM Software automatically attaches the ***.rex** and ***.hrm** files to the e-mail. Simply type in your email address, give a brief description of the issue, and click “Send”.



The e-mail is sent to our Technical Service team, who will analyze your files and respond to your query.

Guidelines

The following guidelines may be useful for optimization of results using Rotor-Gene ScreenClust HRM Software.

- Curve shape can affect normalization and therefore clustering. Ensure that data contains appropriate melt curve features. If some samples do not have melt curve features, remove them from the analysis.
- Samples named "NTC" will automatically be excluded from the analysis. Ensure that all negative controls and no-template controls are not included in the analysis.
- Ensure that all samples have names, as the software will automatically remove samples without names.
- Try to avoid removing samples, especially controls, from the analysis as they represent the variation in your data. Avoid removing controls if reanalyzing.
- Ensure that the temperature range includes as much background at the beginning of the melt as possible to improve analysis. Normalization can be affected if too few points are present at the beginning of the melt, which can influence the line of best fit.
- Remove as much as half of the background at the end of the run. Background at the end of the run influences the principal component extraction and information can be lost in the area of interest if too much background is present.
- Only use supervised mode if there is a control for each genotype. In supervised mode, the software attempts to cluster all samples into given genotypes. Non-related samples will typically sit well outside the cluster.
- If you have only partial information on the genotypes present in the sample, use the unsupervised mode for all unknown and control samples available. All samples will be considered as unknowns and clustered according to the distribution of samples in the principal component plot. The control samples should then be grouped into separate clusters, whereas genotypes without controls should form separate clusters.
- We recommend using at least 3 controls per cluster. Covariance cannot be calculated with less than 4 samples and in this case, the software will have to use the average covariance of the data set. This will reduce the power of the probability and typicality calculated for each sample.

- Typicalities cannot be calculated accurately for clusters with less than 4 samples. For samples in these clusters, the following estimated typicalities are provided:

Number of samples	Estimated typicality (2D)	Estimated typicality (3D)
1	1.00000000	1.00000000
2	0.77880078	0.91889141
3	0.51341712	0.72123337

- Avoid using data sets with less than 5 samples. All statistical calculations used by the software require the use of at least 5 samples to provide optimal results.
- Check the gap statistic graph in unsupervised mode if you suspect that the optimal cluster number is incorrect. There may be another substantial increase in the gap score indicating a higher or lower cluster number (page 43).
- Switching between supervised and unsupervised modes is not recommended. Results will differ as different algorithms are used to calculate clusters and classify unknowns.
- Removing a sample before reanalysis will alter the results. The combination of samples used influences the principal components extraction and therefore the clustering.

Troubleshooting Guide

This troubleshooting guide may be helpful in solving any problems that may arise. For more information, see also the Frequently Asked Questions page at our Technical Support Center: www.qiagen.com/FAQ/FAQList.aspx. The scientists in QIAGEN Technical Services are always happy to answer any questions you may have about either the information and protocols in this handbook or sample and assay technologies (for contact information, see back cover or visit www.qiagen.com).

Comments and suggestions

Cannot initiate analysis

- | | |
|--|--|
| a) ScreenClust is not installed correctly | Ensure that the software is correctly installed. |
| b) Microsoft .NET framework is not correctly installed | Reinstall Microsoft .NET framework 2.0 using the NetFx20SP2_x86.exe installation file included in the Rotor-Gene ScreenClust HRM Software CD. |
| c) R statistical package not correctly installed | Manually install the R statistical package using the R-2.7.1-win32.exe installation file included in the Rotor-Gene ScreenClust HRM Software CD. |
| d) File does not contain HRM data | Ensure that the *. rex file used for analysis contains valid HRM data. Rotor-Gene ScreenClust HRM Software cannot analyze standard melt data. |

Some samples are missing in my analysis

- | | |
|--|--|
| Samples have been automatically excluded | Rotor-Gene ScreenClust HRM Software automatically excludes samples labeled as "NTC" and unlabeled samples. To include these samples, open the "Exclude Data Selection" window and uncheck any excluded samples that are required for analysis. |
|--|--|

The cluster/dimensionality choice in unsupervised mode appears incorrect

- | | |
|---|--|
| More than one ideal cluster/dimensionality choice | Rotor-Gene ScreenClust HRM Software provides a suggested optimal cluster/dimensionality choice using the gap statistic. It is possible that another combination may be a more suitable choice. Users should examine the gap statistic and principal component plot before deciding on the final cluster/dimensionality choice. |
|---|--|

Comments and suggestions

Samples are outside of the cluster

- a) Ellipses do not cover all samples
- The ellipses are a visual aide to give users an idea of how clusters are orientated. They do not cover the entire area of the cluster, so some samples will be positioned outside of the ellipses. Ellipses are computed based on the distribution of samples within each cluster and are designed to cover most samples. Samples lying outside the ellipses tend to have low typicalities.
- b) Controls are not representative of samples in the cluster
- In supervised mode, clusters are centered on the controls. If the controls lie at the edge of the majority of samples from the cluster, the position of the cluster appears shifted away from the samples. Minimize variability between controls and unknown samples in the experimental setup. Differences such as purification method, template concentration/form/treatment, and reaction mixture can all result in different HRM behavior, even for the same amplicon/target sequence.

A control sample is assigned to an incorrect cluster

- If a control is closer to a cluster other than the one to which it was assigned, it is assigned to the closest cluster
- Check that the correct sample was selected as the control. Check that both clusters represent unique melting species. Examine the raw and normalized melt curves of the sample for abnormalities such as unexpected bumps or inflection points.

Typicalities values are the same for many samples

- Clusters have fewer than 4 samples
- For small clusters it is not possible to accurately calculate typicality values. Instead, fixed estimated values are provided based on the number of samples.

Comments and suggestions

Cluster orientation is completely different for 2 different analysis runs of the same assay

Each analysis is unique The principal components in each analysis are derived from the samples included for that analysis. Therefore, the principal component plot is unique for each analysis. Even though the same types of samples may be present, factors such as the number of samples, ratio between sample types, and even minor differences in melt curve, such as the variability between replicates, may result in different principal components and hence different cluster plots.

All posterior probability values equal "Not Available"

Posterior probability error To calculate posterior probability, the covariance matrix is used. For the plotted cluster data, the covariance matrix has not been properly calculated due to the number of principal components used or clusters selected. Change the number of principal components used. If this fails, try to reduce the number of clusters selected.

Appendix: Steps in Data Analysis

This appendix provides details on the algorithm and statistics used in analysis with Rotor-Gene ScreenClust HRM Software. For more detailed background information on the statistical methods, refer to the references on page 46.

Normalization

Raw HRM curves are normalized by applying curve scaling to a line of best fit so that the highest fluorescence value is equal to 100 and the lowest equal to zero. Two regions of the melt data are selected: a region before and a region after melt curve transition. The average fluorescence and slope of the line of best fit are calculated and applied in the normalization.

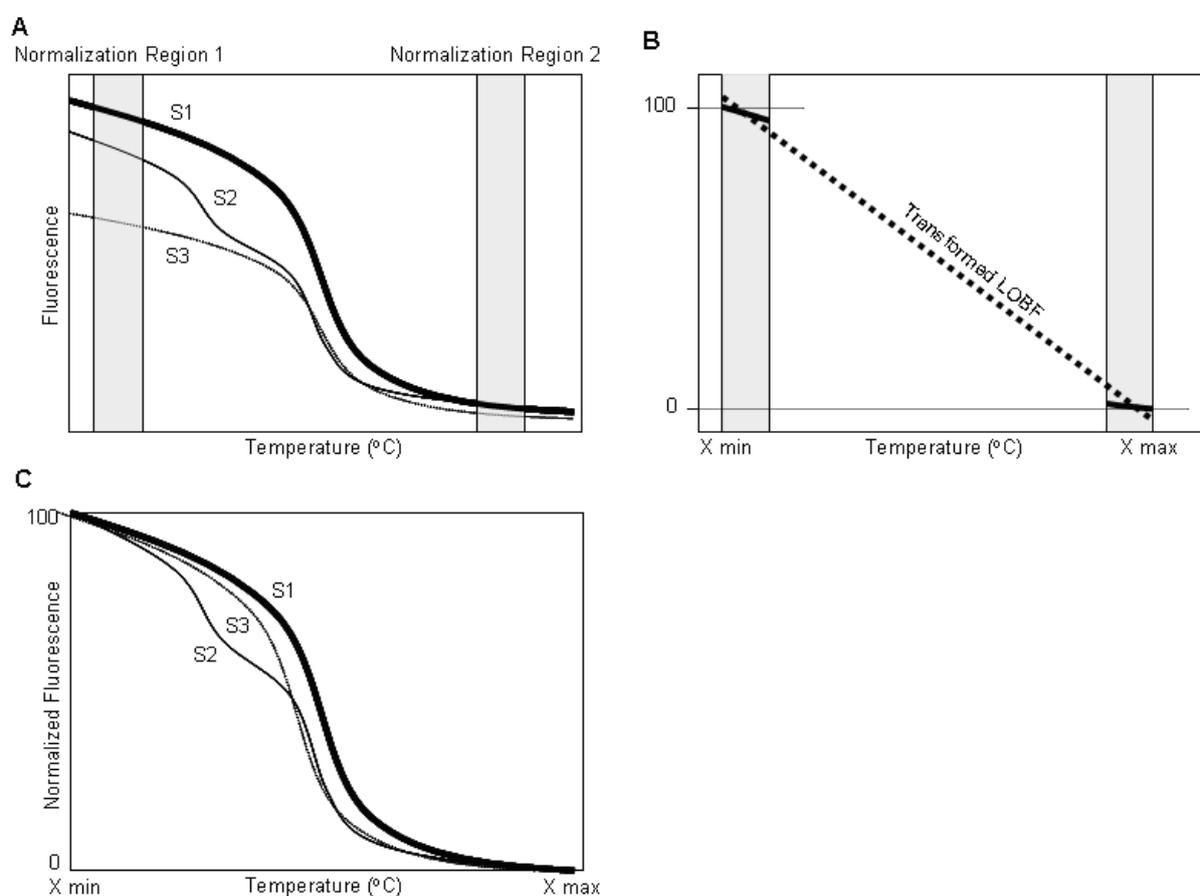


Figure 2. Normalization. **A** In the first step in HRM curve normalization, 2 normalization regions are selected, avoiding areas that are specific to the amplicon. The software focuses on points within the normalization regions. **B** A line of best fit (LOBF) is determined for each HRM curve between points X min and X max. Each individual HRM curve is forced to have the same LOBF from a fluorescence range of 0 to 100. **C** The software transforms each curve to have this LOBF using the following formula: $y = (m2/m1) \times F(x) + (m1 \times b2 - m2 \times b1) / m1$; where m1 is the gradient of the individual HRM curve, m2 is the gradient of the required transformed LOBF, b1 is the intercept for the individual HRM curve, b2 is the intercept for the required transformed LOBF, and F(x) is fluorescence at temperature X.

Generation of a residual plot

The normalized melt curves are differentiated to accentuate the rate of change in melting. To retrieve more information from the data, a composite median curve is constructed by taking the median fluorescence of all the curves for each coordinate along the temperature axis. Residual plots from the composite median curve are drawn for each sample being analyzed. Characteristics of these residual plots are used in principal component analysis.

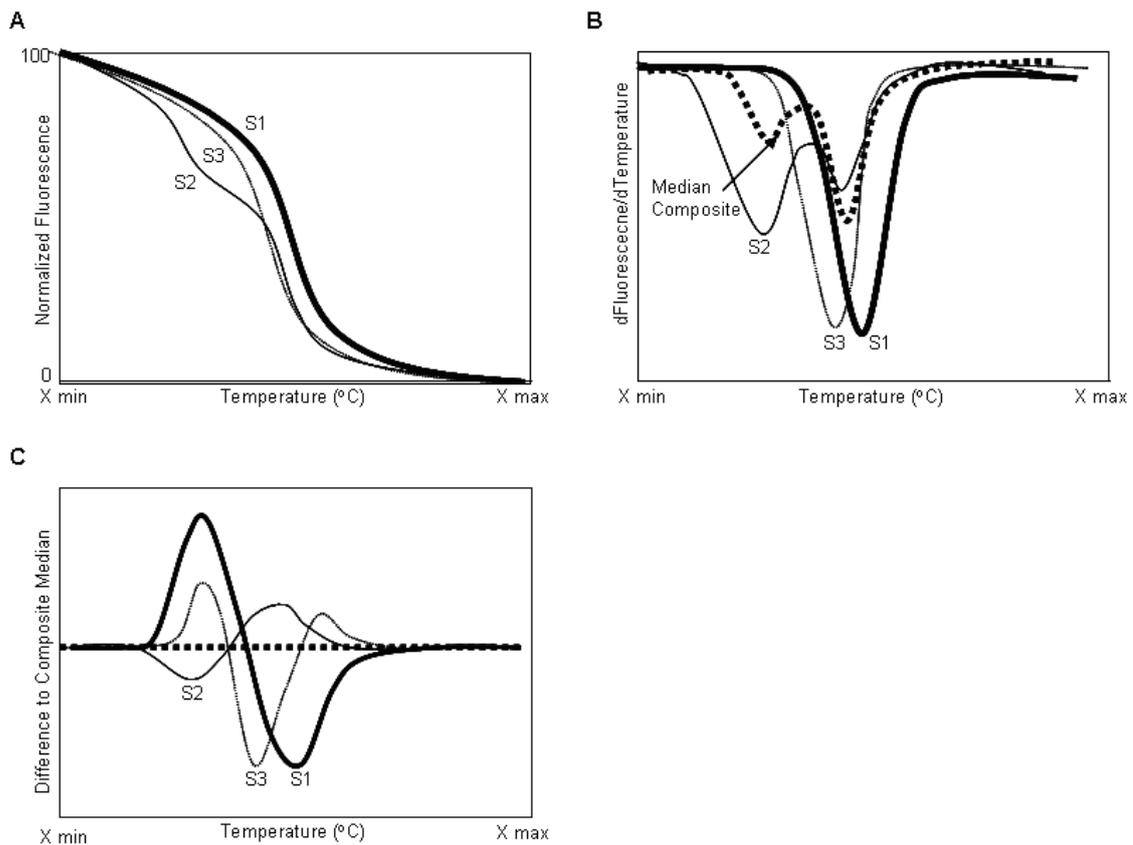


Figure 3. Drawing residual plots. **A** Normalized samples are shown. **B** The normalized samples are differentiated to accentuate the rate of change in the HRM curves. A composite median curve is then determined from all of the differentiated samples present. **C** The composite median curve is then subtracted from each of the individual normalized, differentiated samples to give a residual plot. Principal components are then extracted from the residual plot.

Principal component analysis

Principal component analysis is a well-established method of data analysis. However, Rotor-Gene ScreenClust HRM Software is the first software application to apply principal component analysis to HRM data.

Principal components are determined from the residual plots. Principal component analysis (PCA) enables determination of patterns in multi-

dimensional data, and draws the data so that similarities and differences are highlighted.

The first principal component (PC1) is the new linear combination of data vectors that represent the most variation among samples (1). The second principal component (PC2) further characterizes the remaining variability not accounted for by PC1. Finally, the third principal component characterizes the remaining variability not accounted for by PC1 and PC2. This analysis enables maximum separation of samples. In mathematical terms, PCA is an orthogonal linear transformation that transforms the data to a new coordinate system.

Individual samples are scored against each principal component. The relative scores for each principal component are plotted in 2 or 3 dimensions depending on the number of principal components chosen for the analysis. Typically with any HRM data set, only the first 3 principal components are required. Subsequent principal components are usually background noise that provides very little information about the features of the data.

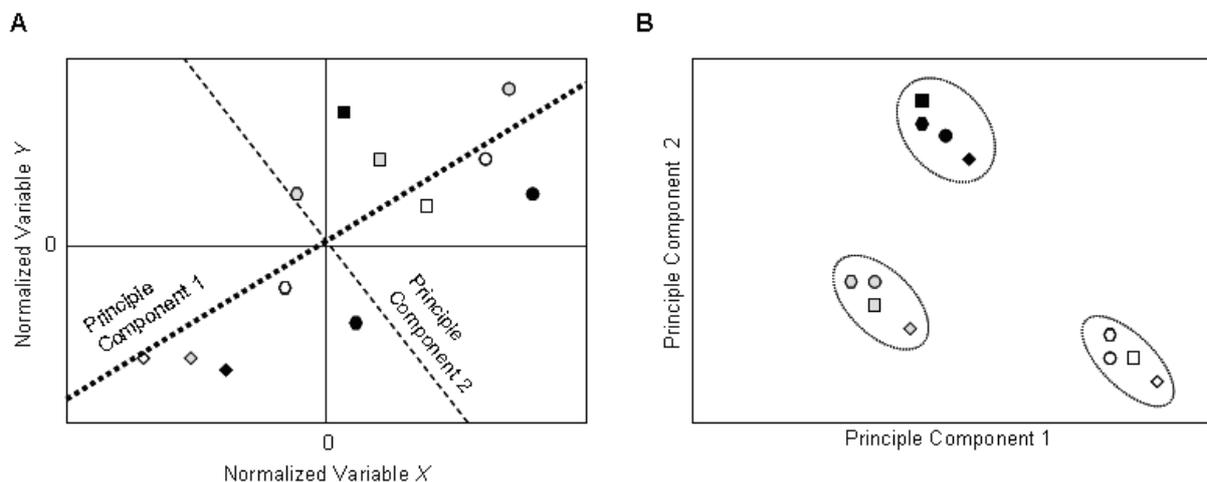


Figure 4. Simplified example of principal component analysis. **A** Detect the vector that best represents the greatest variability across the data, which is the first principal component (PC1). Any remaining variability, perpendicular to PC1 is determined as the second principal component (PC2). In this example, only 2 principal components were required. **B** Once the PCs are determined, the individual samples are scored against each PC and drawn as a principal component plot. Cluster analysis can be performed from the PC plot.

Clustering

Following principal component analysis the samples are clustered into groups. Grouping or clustering can be performed using Rotor-Gene ScreenClust HRM Software in either supervised or unsupervised mode.

In supervised mode, the clusters are known and controls are available for each cluster. Supervised mode is often used for SNP genotyping, where the genotypes are known. In supervised mode, the user assigns control samples for each cluster and the software classifies all unknown samples to clusters according to their characteristics.

In unsupervised mode, the clusters are unknown or controls are not available for every cluster. The unsupervised mode is used to find new mutations in the data, and when there is no prior knowledge of the genotypes present in the samples.

Supervised mode

In supervised mode, the clusters are known and controls are available for each cluster. Supervised mode uses Linear Discriminant Analysis (LDA) as a classification method (2). Using the control samples, LDA calculates a cluster distribution. The centre of the cluster is set as the mean of the controls. Unknown samples are allocated to a cluster based on their proximity to the mean points of the controls. When all the unknown samples are classified, the covariance of each cluster is determined and used to calculate posterior probabilities and typicalities (see below). Although a single control can be used for each cluster (see "Nearest neighbor", page 43), the use of at least 4 samples provides more statistically reliable results.

Following LDA, if less than 4 samples are classified to one cluster, then the average covariance of all the other clusters is used. This enables use of the variance/covariance matrix in calculating probabilities, typicalities, and drawing cluster ellipses.

To determine the correct number of principal components to use for supervised analysis, a cross-validation procedure is performed. Cross-validation involves excluding a known sample, and using the remaining samples to build a classifier. The excluded sample is then classified and any errors made in classification are counted. This process is repeated for the dimensions of principal components and the dimensionality with the lowest error rate is chosen. However, the error rate can be discrete, and in this case the posterior probabilities of the true classes for the known samples can be used. By adding the logs of these probabilities for all the known samples, an overall measure of the quality of the classification is determined. The principal component dimensionality that returns the lowest posterior probabilities value is chosen. If both values are equal then the lower principal component dimensionality is chosen.

Posterior probabilities

The probability of each sample fitting into a particular cluster is calculated and given as a value from zero to one. The sum of all probability values for a single sample is one. Each sample is called into the cluster with the highest probability. Samples with a probability of less than 0.7 of belonging to a particular cluster should be treated with caution.

Typicalities

The typicality measures how well a sample falls within the cluster distribution in which it has been classified. Samples with typicalities below 0.05 should be treated with caution.

Nearest neighbor

If only one control is provided per cluster, the nearest neighbor calculation is used. Unknown samples are allocated to a cluster based on their proximity to the control. Once samples are allocated to a cluster, the posterior probability and typicality values are recalculated based on the newly defined clusters and covariance for each cluster. Clusters with less than 4 classified samples use the average covariance of all the other clusters.

Unsupervised mode

In unsupervised mode, the software calculates the optimal number of clusters automatically. This feature is an excellent tool for the discovery of new polymorphisms.

The unsupervised mode uses the k-means calculation for cluster analysis (3). K-means randomly assigns cluster centers multiple times. The optimal cluster distribution in the spatial principal component plot is determined as the cluster distribution from the many random assignments with the least sum of squares. The optimal distribution is the one where the sum of distances between all samples and their respective cluster center is the lowest.

Combining k-means with the gap statistic allows for the determination of the optimal number of clusters (4). The gap statistic runs a within-cluster sum of squares analysis as a cluster quality measure for each cluster number determined by the k-means algorithm (up to a maximum of 9 clusters) and plots the gap score for each cluster number as well as the standard deviation.

The gap score and its error (y-axis) are plotted against cluster count (x-axis), and the first cluster count where the gap score is greater than or equal the gap score of the following cluster count minus its error is determined to be the optimal cluster number. Two optimal cluster counts are calculated, one for 2 and one for 3 principal components. The number of principal components to use is determined as the optimal cluster count with the highest gap score.

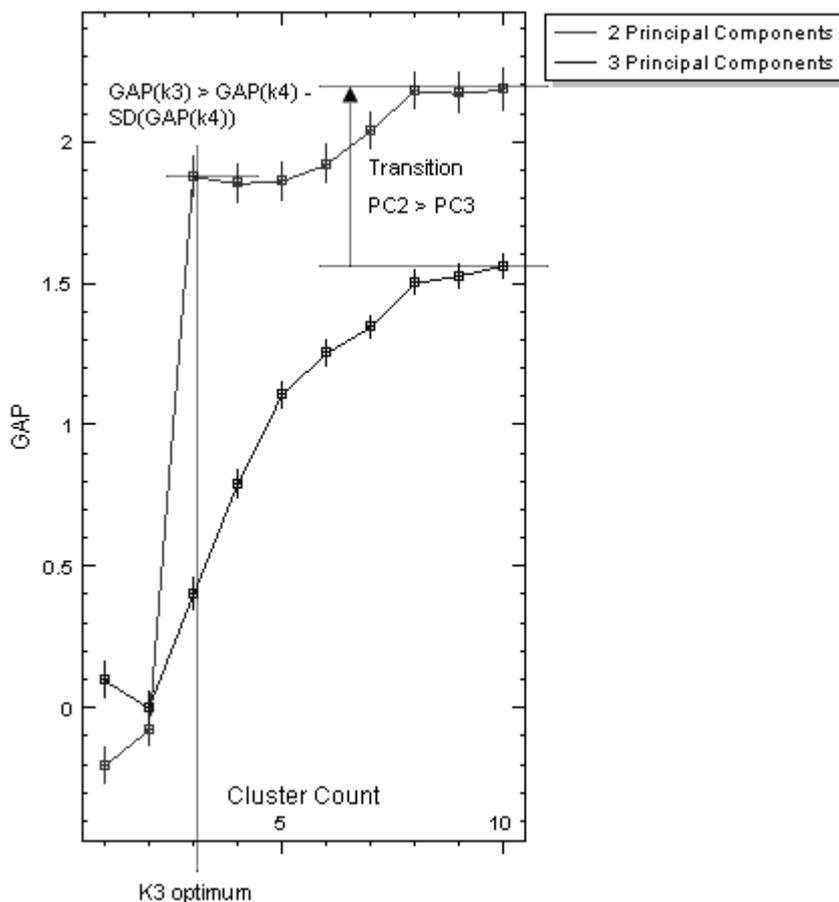


Figure 5. Gap statistic. The gap score along with its error (y-axis) is plotted against cluster count (x-axis) for both 2 and 3 principal components. The cluster count where the mean gap score is greater than or equal to the following gap score minus its error error (i.e., $GAP(k3) > GAP(k4) - SD(GAP(k4))$ in the graph above) is considered to be the optimal cluster number, and the principal component with the highest gap score is chosen.

Although the software automatically determines the optimal number of clusters, further analysis of the gap statistic is recommended. For example, there may be a substantial secondary increase in the gap score following the first plateau, indicating a potential for more clusters within the data set (Figure 6). The user should further investigate the possibility of extra clusters. This can be achieved by simply running more replicates or by sequencing samples within the clusters of interest. Once samples are classified into clusters, the posterior probability and typicality values are calculated as described above.

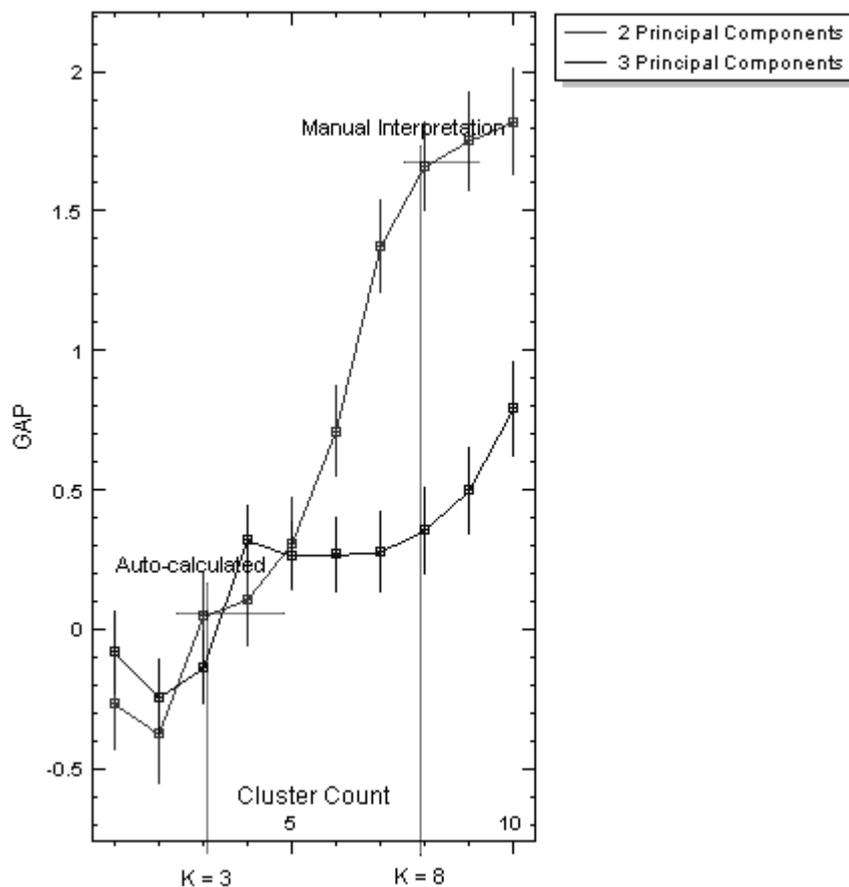


Figure 6. Manual gap statistic interpretation. In this example, the mean gap score for a cluster count of 3 is greater than the mean gap score for a cluster count of 4 minus its error. However, a second substantial increase in the gap transition indicates that another ideal cluster count could be present at 8. This interpretation of the data should be combined with observation of the raw cluster plots and existing knowledge of the experiment.

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.

1. Jolliffe, I. Principal component analysis. Springer. 2nd Edition.
2. Martinez, A.M. and Kak, A.C. (2001) PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence. **23**, 228.
3. Hartigan, J.A. and Wong, M.A. (1979) Algorithm AS 136: A k-means clustering algorithm. Appl. Statist. **28**, 1000.
4. Tibshirani, R., Walther, G., and Hastie, T. (2001) Estimating the number of clusters in a data set via the gap statistic. J. R. Statist. Soc. **63**, 411.
5. Reja, V., Kwok, A., Stone, G., Yang, L., and Bassam, B. (2009) ScreenClust: Advanced Statistical Software for Supervised and Unsupervised High Resolution Melt (HRM) Analysis. Clin. Chem. Submitted for publication.

Trademarks: QIAGEN®, Type-it® (QIAGEN Group); Rotor-Gene®, HRM® (Corbett Research Pty Ltd); Microsoft®, Excel®, Windows®, Vista® (Microsoft Corporation), Intel®, Pentium®, Intel Core™ (Intel Corporation).

The Rotor-Gene ScreenClust HRM Software is protected by copyright law and international treaties. Unauthorized reproduction or distribution of this program, or any portion of it, may result in severe civil and criminal penalties, and will be prosecuted to the utmost extent possible under the law. Copyright © 2009 QIAGEN GmbH. All Rights Reserved.

R/Scilab (D)COM Server 2.50 Copyright © by statconn Thomas Baier.

The software CD contains also a copy of R for Windows V2.7.1.(Copyright © 1998-2006 R Core Team), which is distributed under the GNU General Public License version 2, June 1991. Find the text of the GNU General Public License or find it in the folder "GPL License" on the CD-ROM. Please be advised that R is distributed WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details

This software is intended solely for use with the QIAGEN "Rotor-Gene Q" Real-Time-Cycler. For licensing information see below. The use of this software with non-licensed Real-Time-Cyclers may infringe any number of third party patents.

QIAGEN does not authorise such use and does not make any, and expressly disclaims all, express or implied, warranties or representations of any kind with respect to such use. In no event shall QIAGEN be liable for any special, punitive, incidental or consequential damages (including, without limitation lost profits) arising out of, based on or relating to such unauthorised use.

The purchase of the Rotor-Gene Q includes a limited, non-transferable license to one or more of US Patents Nos 6,787,338; 7,238,321; 7,081,226; 6,174,670; 6,245,514;6,569,627; 6,303,305; 6,503,720; 5,871,908; 6,691,041; 7,387,887; and U.S. Patent Applications Nos. 2003-0224434 and 2006-0019253 and all continuations and divisionals, and corresponding claims in patents and patent applications outside the United States, owned by the University of Utah Research Foundation, Idaho Technology, Inc., and/or Roche Diagnostics GmbH, for internal research use or for non-in vitro diagnostics applications. No right is conveyed, expressly, by implication or estoppel, for any reagent or kit, or under any other patent or patent claims owned by the University of Utah Research Foundation, Idaho Technology, Inc., and/or Roche Diagnostics GmbH, or by any other Party. For information on purchasing licences for in-vitro diagnostics applications or reagents, contact Roche Molecular Systems, 4300 Hacienda Drive, Pleasanton, CA 94588, USA.

This real-time thermal cycler (Rotor-Gene Q) is licensed under U.S. Patent No. 6,814,934 and corresponding claims in any foreign counterpart patent thereof owned by Applied Biosystems LLC, in all fields, including research and development, all applied fields, and human and animal in-vitro diagnostics. No rights are conveyed expressly, by implication or estoppel to any patents on real-time methods, including but not limited to 5' nuclease assays, or to any patent claiming a reagent or kit. For further information on purchasing additional rights, contact the Director of Licensing at Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California, 94404, USA.

License Agreement for the Rotor-Gene ScreenClust HRM Software

1. In the following "Qiagen" refers to Qiagen GmbH and its affiliated companies and "Software" means the Rotor-Gene ScreenClust HRM Software and data supplied on this physical medium (eg. CD-ROM) or over the Internet with these conditions. (If you are unsure of any aspect of this agreement or have any questions they should be emailed to support@qiagen.com.) The Software and any accompanying documentation have been developed entirely at private expense. They are delivered and licensed as "commercial computer software". Please note that this License Agreement does not apply to any other additional programs contained in this medium.

2. License

Your license confers no title or ownership in the Software and is not a sale of any rights in the Software. Qiagen grants you a non-transferable non-exclusive license as follows:

2.1 You use any number of copies of the Software within your organisation, provided that the software is accessible for use only by employees of the organisation. Making this software available for use outside your organisation constitutes a violation of this agreement.

2.2 You may only make copies of the Software only as necessary for backup purposes or when copying is an essential step in the authorised use of the Software. You must reproduce all copyright notices in the original Software on all copies. Under no circumstances may you copy the Software onto any bulletin board, Internet web site or similar public or private distribution system.

2.3 You may not make the Software available to any third party by way of gift or loan or hire.

2.4 You may not incorporate the Software or any part of the Software into programs or computer systems developed or used by you.

2.5 You may not use or otherwise construct data files or other files processed by the Software (save as occurs during the Software's normal operation).

2.6 You may not disassemble, reverse engineer, reverse compile, unlock or translate any part of the Software or make any attempt to discover the source code or underlying algorithms of the Software. You may not alter any of the data files or other files that comprise the Software (save as occurs during the Software's normal operation).

2.7 If this is a demonstration or trial version of the Software you are only licensed to use it for evaluation purposes and within the described restrictions (such as a time limit or limited runs or other limits). The Software may or may not attempt to enforce said restrictions and the Software's failure to enforce said restrictions does not constitute a license for you to go beyond said restrictions.

2.8 You agree to obtain any necessary registration/license key only from Qiagen or an authorised distributor and to keep said key strictly confidential from all third parties.

3. Termination

3.1 If you fail to comply with the terms and conditions of this license then without prejudice to any other rights Qiagen may terminate this license.

3.2 Within 7 days of termination of this license, you will furnish to Qiagen a letter attesting to the destruction of the original and any copies of the Software and the destruction of all copies of any registration/license key. You may terminate this license at any time by furnishing such confirmation.

4. Limited Warranty/Liability

4.1 Qiagen warrants to you only that:

a) If the software is supplied on CD-ROM the CD-ROM is free from defects in materials and workmanship under normal use for a period of ninety days from the date of purchase. (We will replace any defective CD-ROM free of charge.)

b) If used properly the Software will conform substantially in accordance with the documentation supplied with the Software or other specification published by Qiagen for a period of ninety days from the date of purchase.

4.2 Qiagen's entire liability and your exclusive remedy shall be at Qiagen's option either compensation to the value of two hundred and fifty US dollars (US\$250) or replacement of the Software that does not meet the limited warranty.

4.3 WITH THE EXCEPTION OF THE WARRANTIES GIVEN IN SECTION 4.1

ABOVE AND TO THE MAXIMUM EXTENT PERMITTED BY LAW QIAGEN GIVES NO OTHER WARRANTIES IN RESPECT OF THE SOFTWARE.

4.4 TO THE MAXIMUM EXTENT PERMITTED BY LAW AND UNDER NO CIRCUMSTANCES AND UNDER NO LEGAL THEORY, TORT, CONTRACT, OR OTHERWISE WILL QIAGEN BE LIABLE TO YOU OR ANY OTHER PERSON FOR ANY INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES OF ANY CHARACTER INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF GOODWILL, WORK STOPPAGE, COMPUTER FAILURE OR MALFUNCTION, OR ANY AND ALL OTHER COMMERCIAL DAMAGES OR LOSSES, EVEN IF QIAGEN SHALL HAVE BEEN INFORMED OF THE POSSIBILITY OF SUCH DAMAGES. IN ANY CASE QIAGEN'S ENTIRE LIABILITY UNDER THIS AGREEMENT WILL BE LIMITED TO THE LICENSE FEE PAID BY YOU FOR THE SOFTWARE. THIS LIMITATION OF LIABILITY SHALL NOT APPLY TO LIABILITY FOR DEATH OR PERSONAL INJURY TO THE EXTENT APPLICABLE LAW PROHIBITS SUCH LIMITATION.

For updated license terms, see www.qiagen.com. © 2009 QIAGEN, all rights reserved.

www.qiagen.com

Australia ■ Orders 03-9840-9800 ■ Fax 03-9840-9888 ■ Technical 1-800-243-066

Austria ■ Orders 0800/28-10-10 ■ Fax 0800/28-10-19 ■ Technical 0800/28-10-11

Belgium ■ Orders 0800-79612 ■ Fax 0800-79611 ■ Technical 0800-79556

Brazil ■ Orders 0800-557779 ■ Fax 55-11-5079-4001 ■ Technical 0800-557779

Canada ■ Orders 800-572-9613 ■ Fax 800-713-5951 ■ Technical 800-DNA-PREP (800-362-7737)

China ■ Orders 021-3865-3865 ■ Fax 021-3865-3965 ■ Technical 800-988-0325

Denmark ■ Orders 80-885945 ■ Fax 80-885944 ■ Technical 80-885942

Finland ■ Orders 0800-914416 ■ Fax 0800-914415 ■ Technical 0800-914413

France ■ Orders 01-60-920-926 ■ Fax 01-60-920-925 ■ Technical 01-60-920-930 ■ Offers 01-60-920-928

Germany ■ Orders 02103-29-12000 ■ Fax 02103-29-22000 ■ Technical 02103-29-12400

Hong Kong ■ Orders 800 933 965 ■ Fax 800 930 439 ■ Technical 800 930 425

Ireland ■ Orders 1800 555 049 ■ Fax 1800 555 048 ■ Technical 1800 555 061

Italy ■ Orders 02-33430-420 ■ Fax 02-33430-426 ■ Technical 800-787980

Japan ■ Telephone 03-6890-7300 ■ Fax 03-5547-0818 ■ Technical 03-6890-7300

Korea (South) ■ Orders 1544 7145 ■ Fax 1544 7146 ■ Technical 1544 7145

Luxembourg ■ Orders 8002-2076 ■ Fax 8002-2073 ■ Technical 8002-2067

Mexico ■ Orders 01-800-7742-639 ■ Fax 01-800-1122-330 ■ Technical 01-800-7742-639

The Netherlands ■ Orders 0800-0229592 ■ Fax 0800-0229593 ■ Technical 0800-0229602

Norway ■ Orders 800-18859 ■ Fax 800-18817 ■ Technical 800-18712

Singapore ■ Orders 65-67775366 ■ Fax 65-67785177 ■ Technical 65-67775366

Spain ■ Orders 91-630-7050 ■ Fax 91-630-5145 ■ Technical 91-630-7050

Sweden ■ Orders 020-790282 ■ Fax 020-790582 ■ Technical 020-798328

Switzerland ■ Orders 055-254-22-11 ■ Fax 055-254-22-13 ■ Technical 055-254-22-12

UK ■ Orders 01293-422-911 ■ Fax 01293-422-922 ■ Technical 01293-422-999

USA ■ Orders 800-426-8157 ■ Fax 800-718-2056 ■ Technical 800-DNA-PREP (800-362-7737)

