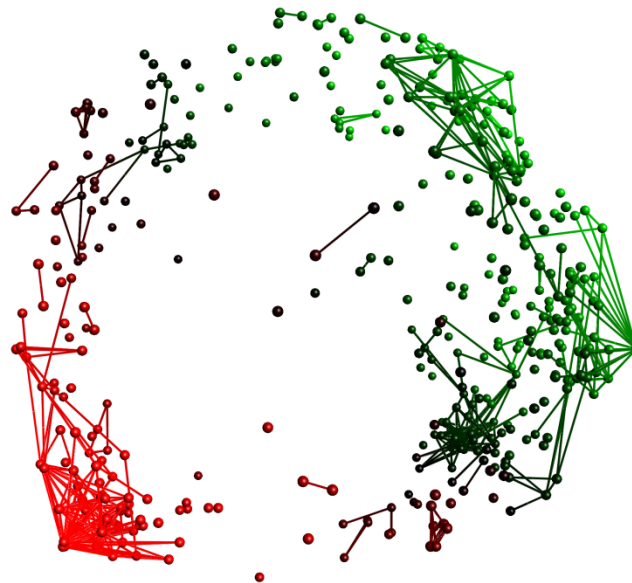# Qlucore Omics Explorer

A statistical test

Analysis in an instant

Preface

Qlucore Omics Explorer is supplying new technology in data analysis and data mining. It is built on state-of-the-art mathematical and statistical methods.

The main features of Qlucore Omics Explorer are the ease and speed with which you will be able to analyze and explore your data sets. You will rapidly and easily visualize and work interactively with your data sets in real time directly on the computer screen. Throughout the analysis you are supported by general statistical methods and measures such as t-tests and F-tests with corresponding p-values and false discovery rates.

This document is meant to enable first-time users to understand and use the basic capabilities regarding performing a statistical test with Qlucore Omics Explorer.

An introduction to the user interface and basic functionality is presented in the Tutorial (**Help Menu**).

A comprehensive description of all functions can be found in the **Reference Manual** that is supplied in the **Help Menu** of Qlucore Omics Explorer. On www.qlucore.com/documentation.aspx is more information provided. There you can also watch instruction movies.

Statistical Analysis with visual feedback

In this section we will continue to work with only one active plot. In this section it will be a heatmap. We will also introduce the statistics functionality. In the Tutorial "Appendix: An introduction to the statistical concepts", you can read about the statistical concepts used.
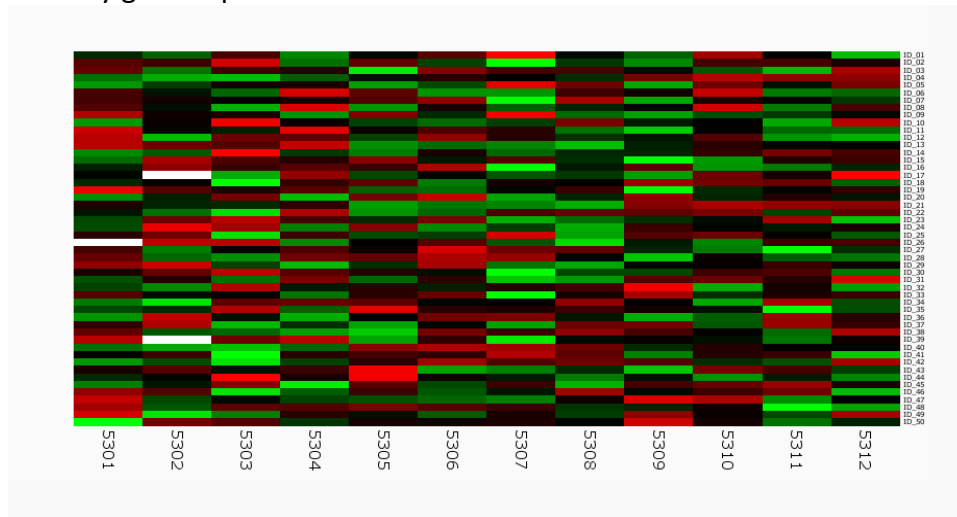
We will do a statistical test and visualize the result simultaneously. We will use the heatmap to generate the visual feedback. We could have worked with any plot type.

Load the Qlucore test data set from the **Help Menu.**

Select **Heat** in the **Method** tab.
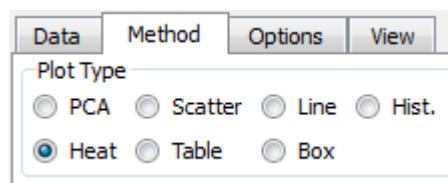
You will directly get the plot below.

*Note 1: The white parts of the heatmap indicate values that have been reconstructed using missing value reconstruction. In the **Data tab** the method for missing value reconstruction can be chosen. Please refer to the **Reference manual** for details about missing values.*
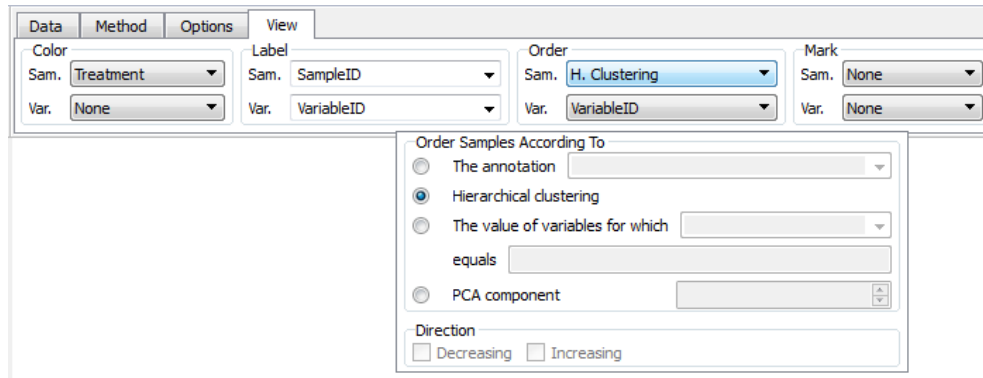
*Note 2: The plot shows normalized values. You can select the normalization in the **Method** tab*

First we enhance the plot by adding additional visual elements.

- Select **Color Samples** in the **View tab** and then color by the annotation *"Treatment"*.
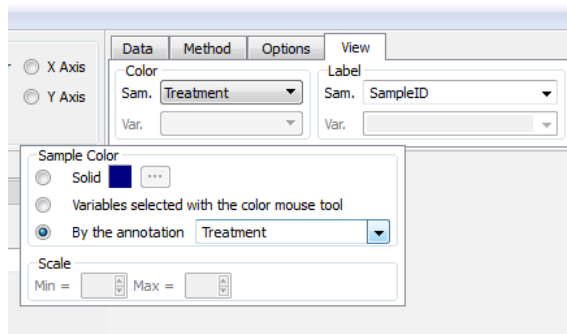
- Then select **Order Samples** in the **View tab** and select **"Hierarchical clustering"**

There are four different algorithms (**Linkage**) that you will find in the **Options tab** for generating the clusters (mean, weighted mean, minimum linkage and maximum linkage). We refer to the reference manual for more information on this and related information on heat maps and clustering.

We are now ready to introduce a statistical test to find the variables that are best on separating "Placebo" from "Drug 1" and "Drug 2".

In the **Statistics Dock Window** you control what type of statistical test you would like to set-up. The tests are:
- Two group comparison (t-test)
- Multi Group Comparison (F-test)
- Linear Regression
- Quadratic Regression
- Rank Regression

Variance slider

Available statistical tests
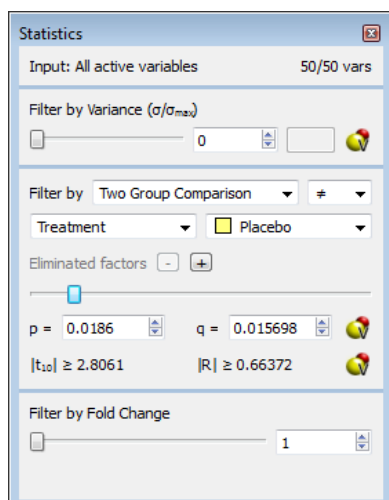
p-value slider
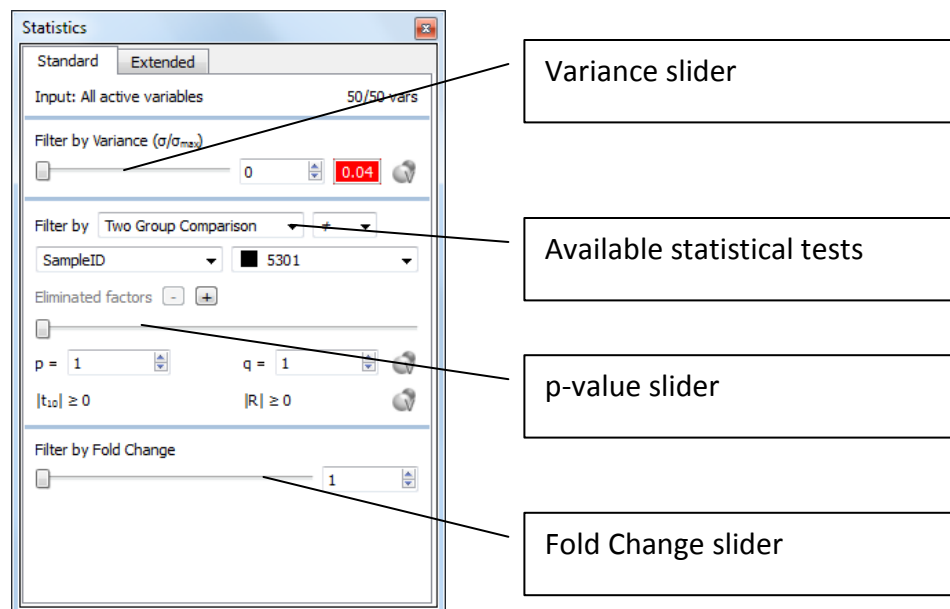
Fold Change slider

The other components of the statistical dialog are the **Variance slider** at the top, the **p-value slider** and the **Fold Change slider**. The use of the variance slider is covered in the section "Basic exploration" in the Tutorial (**Help Menu**). The **p-value slider** is used to select the p-value cut off for the chosen statistical test.

Note: The Extended tab is the interface to statistical scripts written in R. Qlucore Omics Explorer provides this openness and expandability for users you need a wider range of options than the inbuilt tests. Using this functionality requires R to be installed on the computer. You can find more information about this in the reference manual.

To set-up the test that finds the variables that are best on separating "Placebo" from "Drug 1" and "Drug 2", take the following steps.
- Select to Filter by **Two group comparison** (t-test) in the Statistics window.
- Select "Treatment" in the corresponding Combo box
- Verify that "Placebo" is highlighted in the third Combo box

The p-value is presented to the lower left in the **Statistics** window. Adjust the **p-value slider** to a p-value of 0.0265, do it slowly. You can see how the heatmap is updated continuously and the number of variables that fulfill the cut of criteria is decreasing. There are 3 variables that have a p-value of 0.0265 or lower for the selected t-test (Testing "Placebo" against the samples in the groups "Drug 1" and "Drug 2").

The corresponding q-value (0.015698) (which can be interpreted as a false discovery rate threshold) is also displayed in the **Statistics** window and the 3 variables that now are left in the analysis have a high statistical significance. How much you should filter depends on the structure of your data and what significance levels you want to achieve. The results are shown in the plot below.
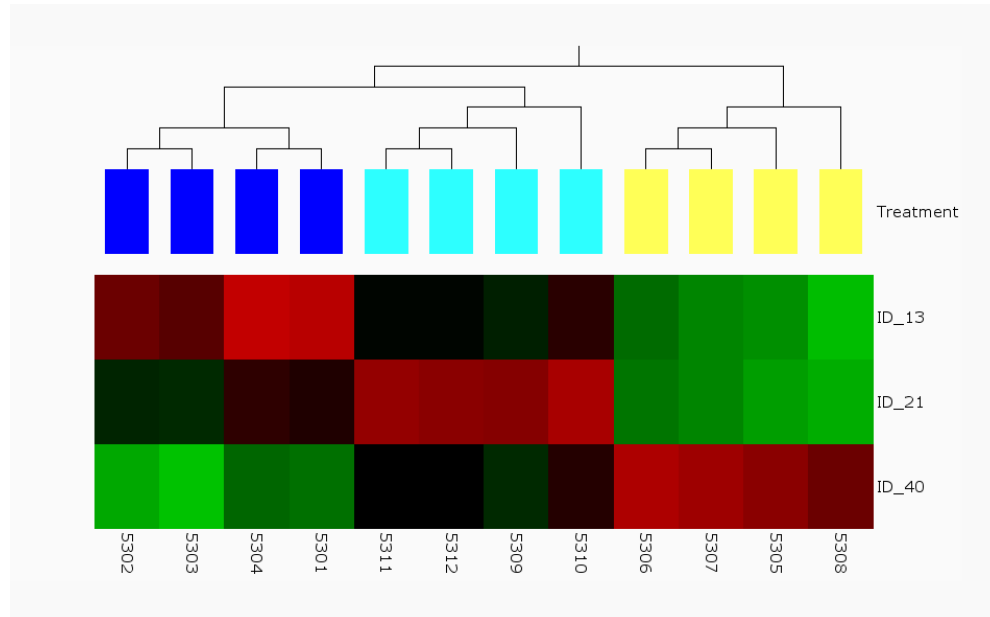
Using the plot we can do multiple observations.

- The three variables (ID_13, ID_21 and ID_40) have the best statistical significance (since they are present)
- The first branch (counting from the top) of the Hierarchical clustering divides the Samples into two clusters; "Placebo" and "Drug 1 and Drug2". The second branch splits the "Drug 1 and Drug 2" cluster into "Drug 1" and "Drug 2". This is expected since we have used the statistical test to identify variables that actually are best at separating "Placebo" from "Drug 1 and Drug2".
- From the regulation we can for example conclude that the variable ID_21 has high values for samples in the group "Drug 1" and that the variable ID_13 has low values for all samples in the "Placebo" group. Red color (default) in the heatmap indicates that a variable has a high value for that sample and green indicates that a variable has a low value for that sample.[1]

One of the key usage models in OE is that it shall be very easy to change the analysis path. To give an example on this select the **Sample window** and select the "Treatment" annotation. Then deselect the "Drug 2" group check box. The updated result is immediately presented, see picture below. The result is an updated statistical test which finds the variables that best separates "Placebo" from "Drug 1". **With the same settings in the statistics dialog (the p-value slider is not moved) more variables (7) are now found that match the earlier selected test criteria.**

---

[1] The color scale can be adjusted. Select Plot settings in the View tab and then the More button.
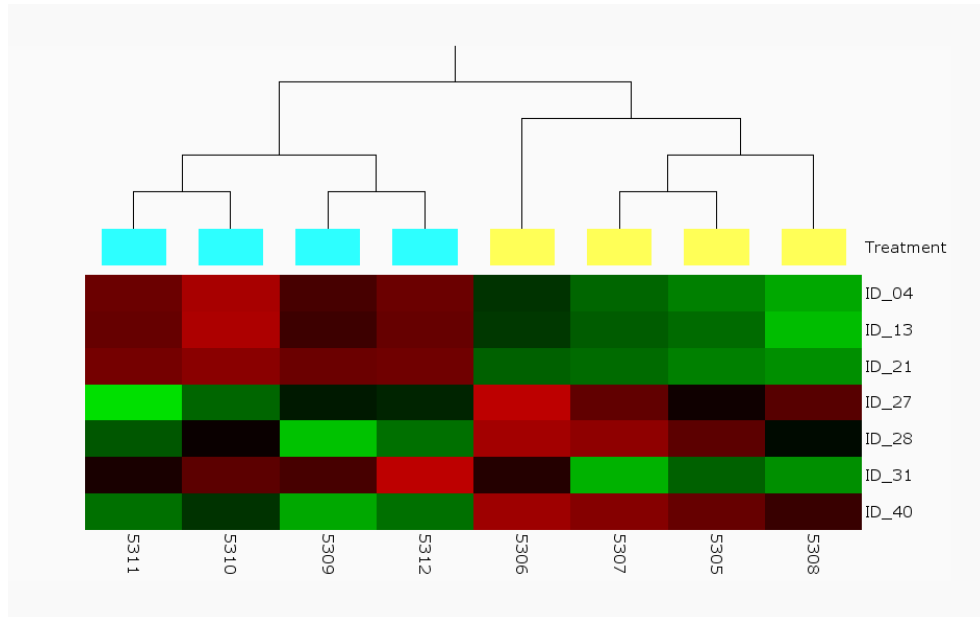
A natural step could now be to out of the 7 most significant variables select the ones with highest Fold Change. Do that by moving the Fold Change slider in the statistics Dialog. *(Note: Fold Change is valid when you have selected two group comparison.)* In the figure below you can see the two variables (ID_21 and ID_28) which have a Fold Change greater than 3.
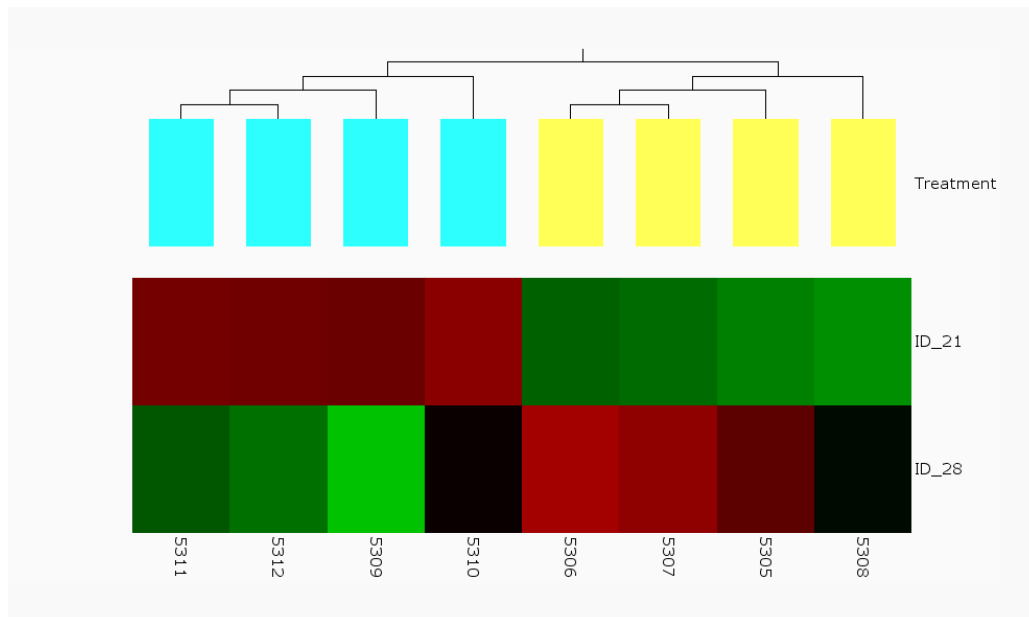
Exporting images, animations and other data

You can at any time during an ongoing analysis export an image or an animation in OE.

You do this by selecting **File > Export > Image** or **File > Export > Video,** and then supplying the name and other characteristics of the exported file.

You can also export principle components, sample distances, annotations and other useful data for downstream analysis, see **File > Export.**

Also note that it is possible, at any point in an analysis, to save the complete current state of OE by using the Log function**.** You can then return to that specific point in the analysis, by opening the corresponding log file in OE, and select the specific log point.

Disclaimer

The contents of this document are subject to revision without notice due to continuous progress in methodology, design, and manufacturing.
Qlucore shall have no liability for any error or damages of any kind resulting from the use of this document.

Qlucore Omics Explorer is intended for research use only.

Trademark List

Windows Vista, Windows 7 and Windows 8 are trademarks of Microsoft.
Affymetrix is a trade mark of Affymetrix
Agilent is a trade mark of Agilent