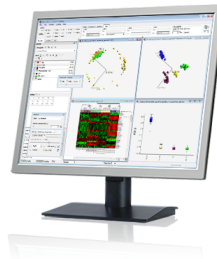


Qlucore Omics Explorer



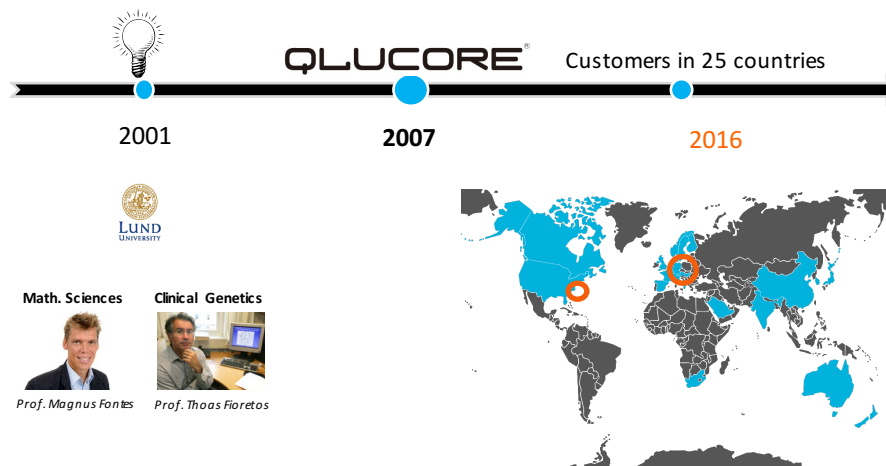
Basic Training

Sara Strandberg
Qlucore

Objectives

- Introduction and terminology.
- Get acquainted with the user interface
- Import data and annotations
- Identify discriminating variables using basic statistical tests
- Use visualization to enhance analysis and interpret results
- Presentation of data in different plot types
- Export of variable lists and images
- Biological interpretation of results
- Data Exploration - find structure in large data sets

Background



QLUCORE®

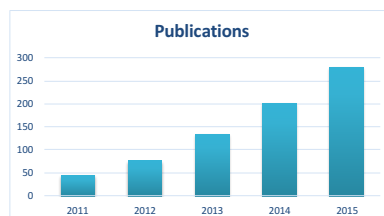
Course material – Basic training. © Qlucore 2016

References and Front Covers

The screenshot shows the Qlucore Omics Explorer website. The header includes the Qlucore logo and navigation links: HOME, COMPANY, PRODUCTS, CONTACT, SUPPORT, DOWNLOADS, and LOGIN. Below the header, there's a section titled "REFERENCES" with a list of scientific papers. To the right of the references, there are buttons for "VERSION 2.10 NEWS", "HOW TO USE", and "GET THE INFORMATION GUIDE". Below these buttons, there are sections for "DATA" (listing Qlucore Omics Explorer, PROTEIN ARRAY, RNA, and DNA METHYLATION), "DISEASE AREA" (listing CANCER, OBESITY, and DIABETES), "PERFORMANCE" (with a bar chart), and "SCREENSHOTS" (with three small images).

85% renewal rate

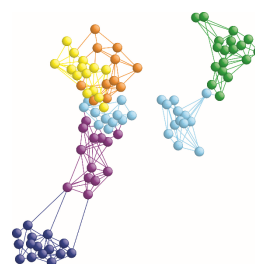
QLUCORE®



Course material – Basic training. © Qlucore 2016

Qlucore Omics Explorer

A very fast and easy to use tool to analyze and explore data without being a statistical expert



QLUCORE®

QLUCORE®

Data types

Examples

- Gene Expression Data (array and NGS)
- DNA Methylation
- Proteomics
- Metabolomics
- Protein array data
- miRNA data
- qPCR data
- Flow cytometry data
- Multiplex

Supported File formats

- Affymetrix GeneChip compatible (.cel and .chp files)
- Agilent txt files
- BAM files (aligned for **RNA-seq**)
- GEO soft files
- Wizard (*.txt, *.csv)

Any multivariate data

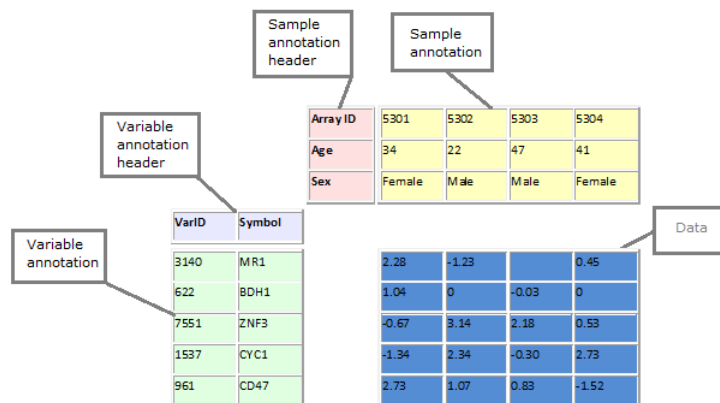


GEO
Gene Expression Omnibus

QLUCORE®

Course material – Basic training. © Qlucore 2016

Terminology



QLUCORE®

Course material – Basic training. © Qlucore 2016

Work modes

Visualize and Explore

- Check quality
- Identify structures in the data
- Generate new hypotheses

Statistical Analysis

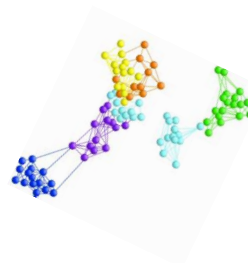
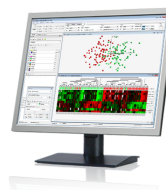
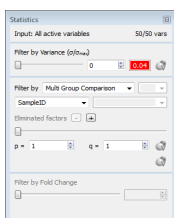
- t-test, ANOVA, Regression analysis.
- Open API to R
- Easy generation and export of documentation and plots

Achieve Biological Insight

- Explore annotations
- GO Browser
- GSEA – Gene Set Enrichment Analysis

Classify and Predict

- Build classifiers
- kNN, SVM, RT
- Use classifiers for prediction



QLUCORE®

Course material – Basic training. © Qlucore 2016

Statistical tests

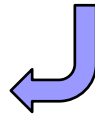
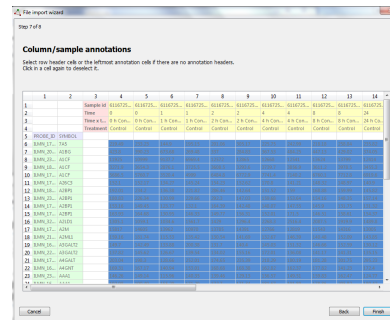
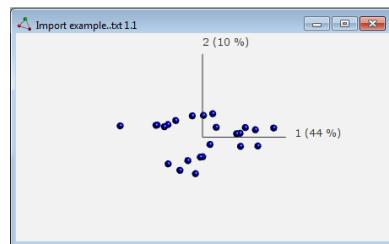
- **t-test (two group comparison)**
find variables that are statistically different between 2 groups.
- **F-test (ANOVA - Multigroup comparison)**
find variables that are statistically different between 2 or more groups.
- **Regression Analysis**
(used for numerical annotations e.g. time, dose, age etc)
find variables that increase or decrease over time, dose etc.
- **Open API to R**
(Welch, Limma, Mann Whitney - included)

Exercises

1. Import data with Wizard
2. ANOVA (F-test) – Export Variable list
3. t-test – Heatmap
4. Box plot
5. Linear Regression
6. Rank Regression
7. Bar Plot
8. Compare lists – Set operations
9. GEO data – additional data set for cross comparison
10. Exploratory Data Analysis – Exploration
11. Clustering
12. Import data and add annotations

Exercise 1 – Import data with the Wizard


1. Select File/Open Wizard
2. Select the file Import example.txt
3. Select Wizard
4. Step through the Wizard
5. The dataset is loaded and a Sample PCA plot is displayed

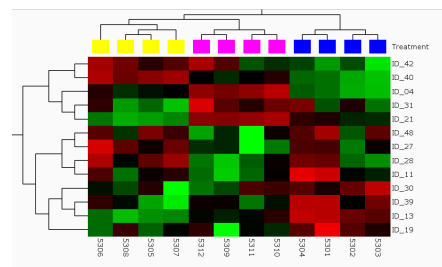


GLUCORE®

Course material – Basic training. © Glucore 2016

Exercise 2 - ANOVA

- Perform ANOVA - F-test (Multigroup comparison)
 - Heat map
 - Color samples according to annotations
 - Find discriminating variables
 - Save list, Export list
- 



GLUCORE®

Course material – Basic training. © Qlucore 2016

Techniques 2

Select heat map

Color samples

ANOVA (Multi group comparison) p-value slider

Order hierarchical clustering

Copy the list

Add more info to the list

Save the list

Course material – Basic training. © Qlucore 2016

Steps 2

1. Load the Qlucore Test Data set from Help/Example
 2. Select Heat Map in the Method tab
 3. Select the View Tab and change the y-axis label to Symbol
 4. Go to the View Tab and change Order to Hierarchical clustering for both the samples and the variables
 5. Colour the samples according to the Treatment annotation.
 6. Select Multi group comparison in the Statistics dialog. Select the annotation Treatment.
 7. Drag the p-value slider in the Statistics dialog. Monitor the p-values and q-values of the remaining variables.
- Continue on next page.....

View

Order
Sam. H. clustering
Var. H. clustering

Statistics

Input: All active variables 50/50 vars

Filter by Variance (σ^2/σ_{max})

Filter by Multi Group Comparison

Treatment

p = 1 q = 1

F_{2,9} ≥ 0 R² ≥ 0

Filter by Fold Change

Steps 2 continued

8. Create a list of the variables by making a copy of the active variable list (Qlucore Test Data set) in the Variables tab.
9. Double click on the list to rename it.
10. Add more information to the list by clicking on the Play symbol. Add a comment.
11. Export the list to your PC (select the save icon) (include annotations)
12. Open the list and look at the content

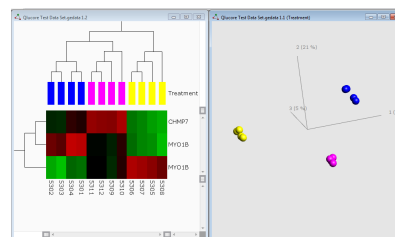


QLUCORE®

Course material – Basic training. © Qlucore 2016

Exercise 3 – t-test

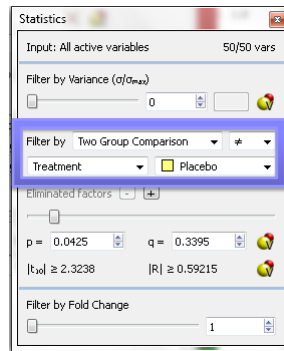
- Perform a t-test (Two Group comparison)
- Find discriminating variables
- Show the samples and variables in synchronized plots - heat map and PCA
- Create a variable list from the heat map



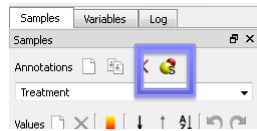
QLUCORE®

Course material – Basic training. © Qlucore 2016

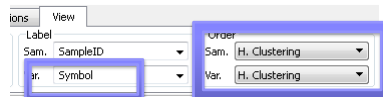
Techniques 3



Two group comparison (t-test)

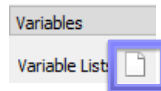


Colour the samples

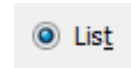


Change label

Hierarchical clustering



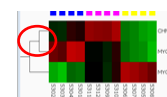
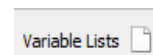
Create new variable list



The mouse List tool

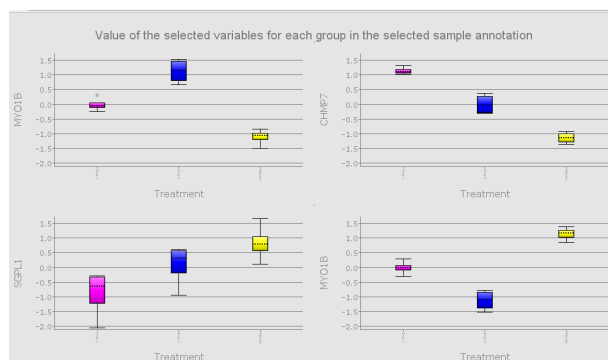
Steps 3

1. Select Two Group comparison in the statistics dialogue.
2. Compare the Placebo group with both treated groups (Select Placebo)
3. Drag the statistical slider to find discriminating variables
4. Select Window/New synchronized plot
5. Tile the Plots using CTRL+T or Cmd+T
6. Colour the new Sample PCA plot according to Treatment
7. Observe that the same clusters can be viewed in both the sample PCA and the heat map
8. Create New Variable list by clicking on the icon in the variables tab.
9. Change the mouse tool to List
10. Select the down regulated genes for the placebo group with the list tool (select the node in the heat map)



Exercise 4 – Box plot

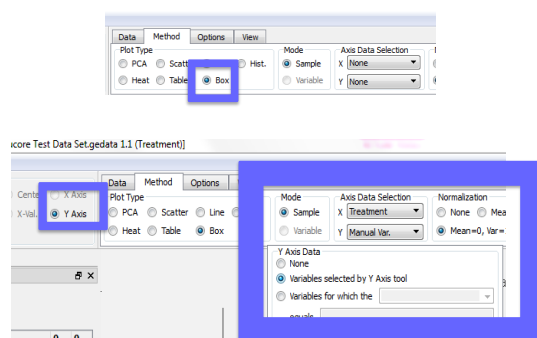
- Display discriminating variables in Box plots
- Export plots



QLUCORE®

Course material – Basic training. © Qlucore 2016

Techniques 4



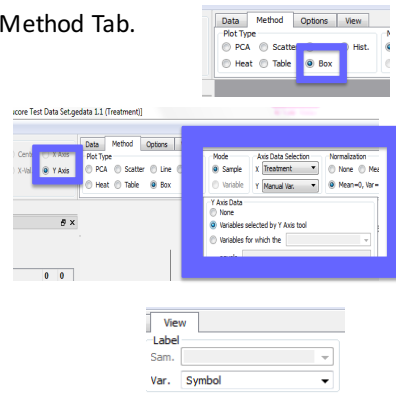
X-axis Sample annotation: Treatment
Y axis Variables selected by Y axis tool

QLUCORE®

Course material – Basic training. © Qlucore 2016

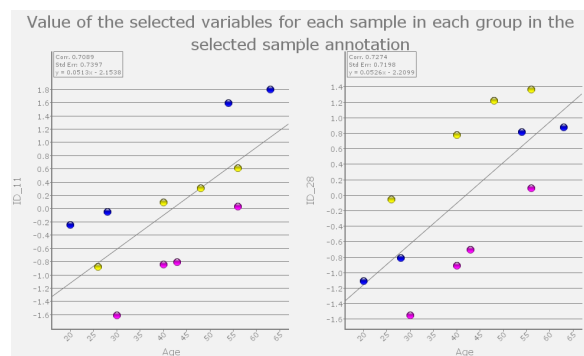
Steps 4

1. Change the Heat map to Box Plot in the Method Tab.
2. Select **Treatment** on the x-axis.
3. Select **Variables selected by y-axis tool** on y-axis and click on the active list.
4. Change the label to Symbol
5. Export the image – File/Export/Image
6. Open the exported file.



Exercise 5 – Linear regression

- Perform Linear regression
- Find discriminating variables
- Show the variables in a scatter plot



Steps 5

1. Select Linear Regression in the Statistics dialogue
2. Select the annotation Age
3. Drag the statistical slider to find discriminating variables
4. Show the variables in Scatter plots

Filter by Linear Regression
Age

- Select Scatter Plot in the Method Tab
- Select the dropdown menu for the x-axis in the Axes Data Selection in the Method tab
- Select Sample annotation Age and Numerical scale
- Select the dropdown menu for the y-axis in the Axes Data Selection in the Method tab
- Select Variables selected by the Y-axis tool
- Click on your list of variables in the variables tab or select individual variables in the variable list

Scatter

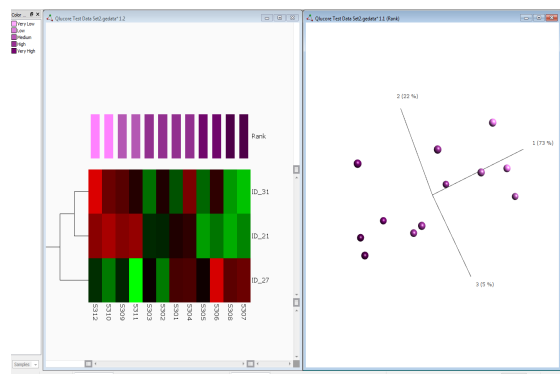
Axis Data Selection
X Age (num)
Y Manual Var.

QLUCORE®

Course material – Basic training. © Qlucore 2016

Exercise 6 – Rank regression

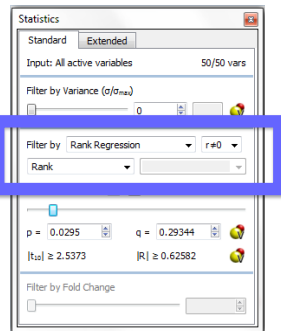
- Perform Rank regression
- Work with Color gradient



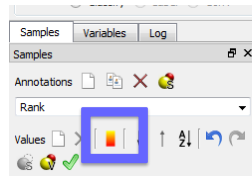
QLUCORE®

Course material – Basic training. © Qlucore 2016

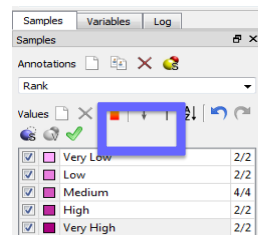
Techniques 6



Rank regression



Color gradient



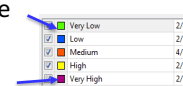
Make sure values are in the ranked order

QLUCORE®

Course material – Basic training. © Qlucore 2016

Steps 6

1. Load the Qlucore Test Data Set2.gedata from the training folder. File/open.
2. Open a new synchronized plot. Make it a heat map. Tile the plots
3. Select the Rank annotation (Sample tab). Sort the rows in rank order if they are not correctly sorted. Use the up and down arrows.
4. Create a color gradient for the Rank annotation:
 - Double click the colored box next to the "very low" value to change color to a light color
 - Do the same thing on the "very high" value but select a dark color. Pick the colors you like.
 - Now click the gradient icon and follow the instructions.

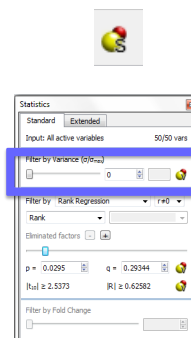
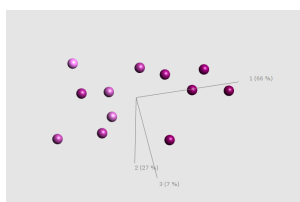


QLUCORE®

Course material – Basic training. © Qlucore 2016

Steps 6 – cont.

5. Color your PCA plot and your heatmap according to the rank annotation.
6. In the Statistics box select Rank regression and select the annotation Rank.
7. Drag the statistical slider until you have around 3-4 variables left.
8. Notice that the samples line up in the rank order

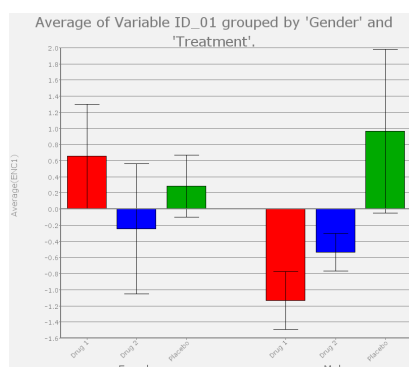
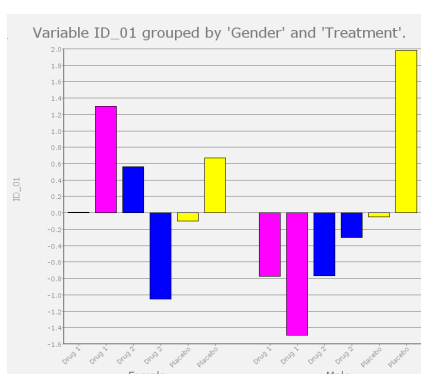


QlUCORE®

Course material – Basic training. © Qlucore 2016

Exercise 7– Bar plot

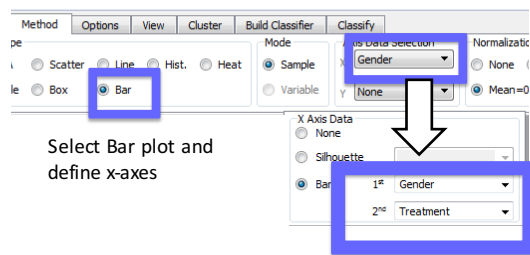
- Display results in Bar plots



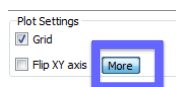
QlUCORE®

Course material – Basic training. © Qlucore 2016

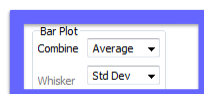
Techniques 7



Select Bar plot and
define x-axes



View tab – More plot settings



Options tab – for average bars

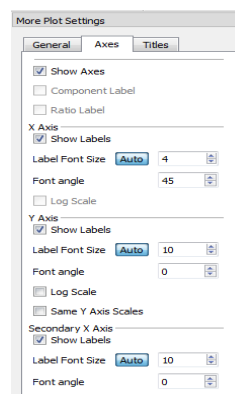
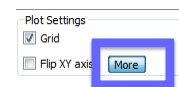
Steps 7

1. Load the Qlucore Test Data Set from the Help menu.
 2. Select Bar plot and Select Gender and Treatment for the x-axes
 3. Select one (or several) variable(s) for the y-axis
 4. Colour the bars according to Treatment
 5. Create average bars by selecting Average from Bar options – to the far right in the Options tab
 6. Select Whiskers: Std Dev
 7. Change the y-axis label to Symbol
- Continue on next page....



Steps 7 - continued

8. Click the More button in the View tab and open the More plot settings Window
9. Set the size of the 2nd x-axis label to 15
10. Change the angle of the first x-axis label

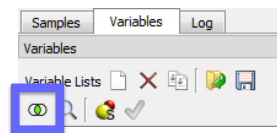


Exercise 8 – Set operations

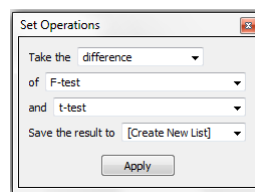
- Compare lists
- Create intersection, unions, differences

Variables			
Variable Lists			
<input type="checkbox"/>	Search	0	0
<input checked="" type="checkbox"/>	Qlucore Test Data Set.gedata 1	50	0
<input checked="" type="checkbox"/>	Acute Lymphoblastic Leukemia.g...	501	501
<input type="checkbox"/>	F-test	502	502
<input type="checkbox"/>	t-test	501	501
<input type="checkbox"/>	Intersection	211	211
<input type="checkbox"/>	Diff	291	291

Techniques 8



Set operations



QLUCORE[®]

Course material – Basic training. © Qlucore 2016

Steps 8

1. Load the Leukemia dataset from the Help menu
2. Perform a Multigroup comparison on Leukemia Subtype – use the p-value slider and filter to approx 500 variables
3. Copy the list and call it "F-test"
4. Perform a two group comparison on Leukemia subtype T-ALL and filter to approx 500 variables
5. Copy the list and call it "t-test"
6. Use the Set Operations tool and make an intersection of the lists – rename the list "Intersection"
7. Use the Set Operations tool and make a difference between F-test and t-test. Rename the list Diff.

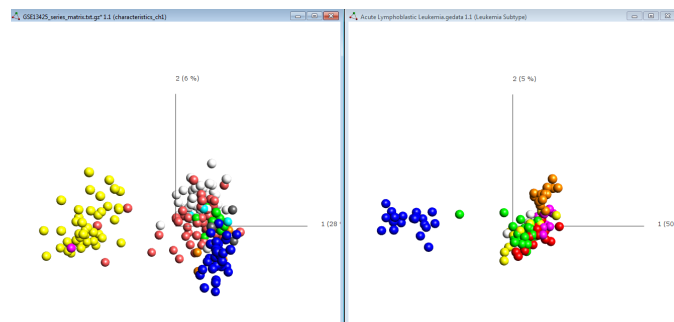
Variable Lists	Count	Count
Search	0	0
Qlucore Test Data Setgedata 1	50	0
Acute Lymphoblastic Leukemia.g...	501	501
F-test	502	502
t-test	501	501
Intersection	211	211
Diff	291	291

QLUCORE[®]

Course material – Basic training. © Qlucore 2016

Exercise 9 – GEO data set

- Import data set from Gene Expression Omnibus (GEO)
- Use a list generated from one data set and cross compare in a second data set.



QLUCORE[®]



Course material – Basic training. © Qlucore 2016

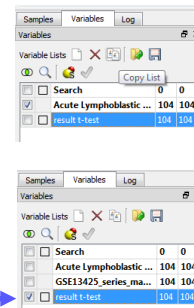
Techniques 9

QLUCORE[®]

Course material – Ba

Steps 9

1. Load the Acute Lymphoblastic Leukemia.gedata data set from the Help menu.
2. Color the data according to the annotation Leukemia Subtype.
3. Perform a t-test on Leukemia subtype, select TEL/AML1 in the statistics box.
4. Filter down to about 100 genes, click the active list and press  to copy the list and rename it to “**result t-test**”.
5. Import a data set from GEO. Search for GSE13425 (childhood ALLs). Tile the plots using CTRL+T or CMD+T. (If no internet connection find the data set in the training folder).
6. Use the generated list **result t-test**, which separated the TEL/AML1 subtype, as input in the GSE13425 data set. 
7. Notice the change and color the GSE13425 data set according to the “characteristics_ch1” annotation.

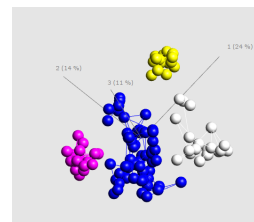


QLUCORE[®]

Course material – Basic training. © Qlucore 2016

Exercise 10 – Exploratory Data Analysis

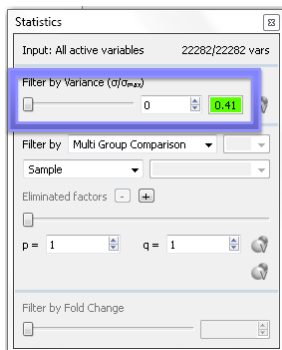
- Use the statistical and visual support in Qlucore Omics Explorer for exploratory analysis and hypothesis generation.
- Find new, previously unknown structures in data.
- Quality control – detect anomalies.
- Use the powerful pattern recognizing capabilities of the human brain.



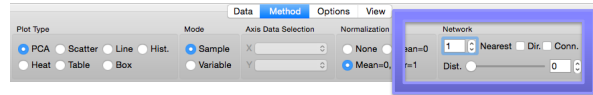
QLUCORE[®]

Course material – Basic training. © Qlucore 2016

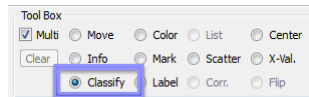
Techniques 10



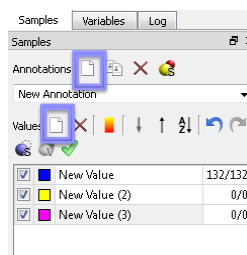
Variance filtering



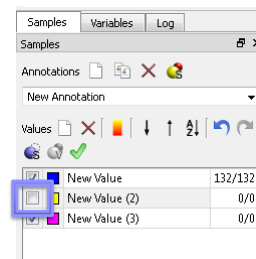
Network



Classify



New annotation and values



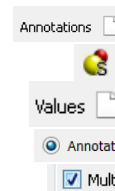
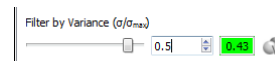
Deselect group

QLUCORE®

Course material – Basic training. © Qlucore 2016

Steps 10

1. Load the Acute Lymphoblastic Leukemia.gedata data set from the Help menu.
2. Look at the data. Can you see any patterns?
3. Remove variables with low variance by moving the variance slider
4. Find the maximum Projection Score, 0.43.
5. Use networks to connect each sample to its closest neighbor.
6. Create a new sample annotation.
7. Color the data set according to the new annotation.
8. Create a new annotation value.
9. Classify the found group.
10. Deselect the found group.
11. Repeat steps 5-10.

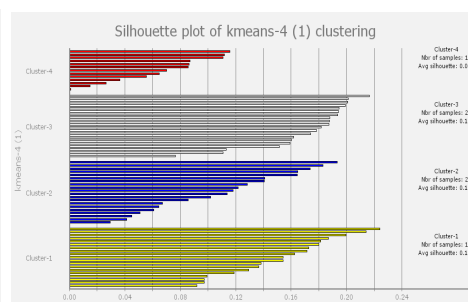
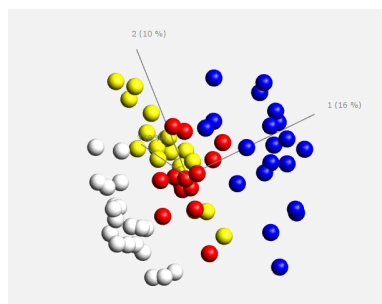


QLUCORE®

Course material – Basic training. © Qlucore 2016

Exercise 11 - Clustering

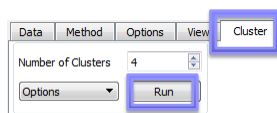
- Use k-means clustering to find structures
- Use Silhouette plots to evaluate clusters



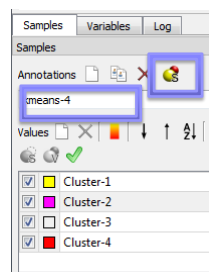
QLUCORE®

Course material – Basic training. © Qlucore 2016

Techniques 11

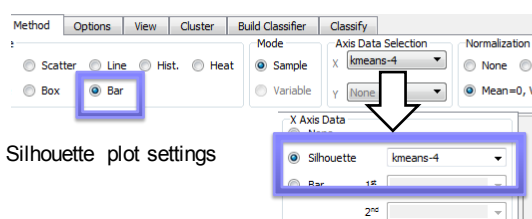


Cluster



Colour the clusters

QLUCORE®



Silhouette plot settings

Course material – Basic training. © Qlucore 2016

Steps 11

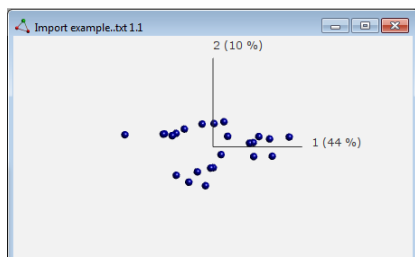
1. Load the Leukemia Example Clustering dataset
2. Perform variance filtering so that you keep 10% of the variables – e.g approx 2000
3. Go to the Cluster tab, set no of Clusters to 3 and press Run
4. Colour the samples according to the new clustering annotation (k-means-3)
5. Open a New synchronized plot and make it a Silhouette plot:
Select Bar plot in method tab and for X-axis select Silhouette and k-means 3.
6. Inspect the clusters in the silhouette plot
Use the mouse tool Mark to mark the sample with a negative Silhouette value
7. Repeat step 3-6 with 4 clusters. Which cluster option seems most relevant?

Data Import

- .txt files – open with Wizard
- .gedata files – File/Open
- GEO data
- Affymetrix cel files
- Agilent txt files
- Aligned RNA-seq BAM files

Exercise 12a – The Wizard

1. Select File/Open Wizard
2. Select the file **Import example without annotations.txt**
3. Select Wizard
4. Step through the Wizard
5. The dataset is loaded and a Sample PCA plot is displayed



File Import Wizard

Step 7 of 8

Column/sample annotations

Select row header cells or the leftmost annotation cells if there are no annotation headers. Click in a cell again to deselect it.

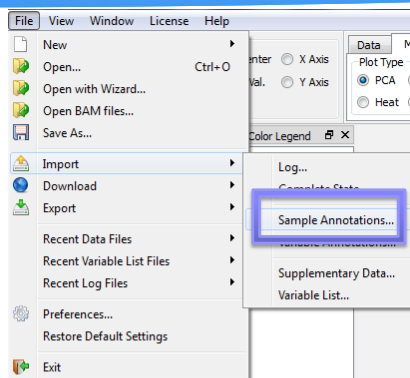
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Sample id	6116725068_A.AVG_Signal	6116725068_B.AVG_Signal	6116725068_C.AVG_Signal	6116725068_D.AVG_Signal	6116725068_E.AVG_Signal	6116725068_F.AVG_Signal	6116725068_G.AVG_Signal	6116725068_H.AVG_Signal	6116725068_I.AVG_Signal	6116725068_J.AVG_Signal	6116725068_K.AVG_Signal	6116725068_L.AVG_Signal	6116725068_L.AVG_Signal
2	Treatment	Control	Control	Control	Control	Control	Control	Control	Control	Control	Control	Control	Control	Control
3	Time	0	0	1	1	2	2	4	4	8	8	24	24	24
4	Time x treatment	0 h Control	0 h Control	1 h Control	1 h Control	2 h Control	2 h Control	4 h Control	4 h Control	8 h Control	8 h Control	24 h Control	24 h Control	24 h Control

QLUCORE®

Course material – Basic training. © Qlucore 2016

Exercise 12b – Add more annotations

Sample id	Treatment	Time	Time x treatment
6116725068_A.AVG_Signal	Control	0	0 h Control
6116725068_B.AVG_Signal	Control	0	0 h Control
6116725068_C.AVG_Signal	Control	1	1 h Control
6116725068_D.AVG_Signal	Control	1	1 h Control
6116725068_E.AVG_Signal	Control	2	2 h Control
6116725068_F.AVG_Signal	Control	2	2 h Control
6116725068_G.AVG_Signal	Control	4	4 h Control
6116725068_H.AVG_Signal	Control	4	4 h Control
6116725068_I.AVG_Signal	Control	8	8 h Control
6116725068_J.AVG_Signal	Control	8	8 h Control
6116725068_K.AVG_Signal	Control	24	24 h Control
6116725068_L.AVG_Signal	Control	24	24 h Control



- The unique sample/variable id in the first column must match the id of the imported dataset.

QLUCORE®

Course material – Basic training. © Qlucore 2016

Steps 12b

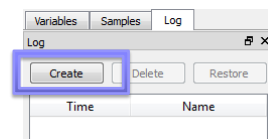
1. Select File/Open with Wizard
2. Select the file **Import example without annotations.txt**
3. Check the sample annotations in the Sample tab and see that you only have one annotation – Sample id
4. Select File/Import/Sample annotations
5. Select the file **Annotations.txt**
6. Import all the annotations
7. Check the sample annotations in the Sample tab and see that you now have all 4 annotations

QLUCORE®

Course material – Basic training. © Qlucore 2016

Save dataset

- Once a data set is loaded you can save it as a gedata.file, which will contain all the imported data including annotations.
(File/Save as)
- Save logpoint to save a snap shot of your analysis at a specific point in time.

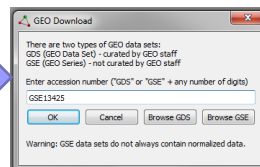
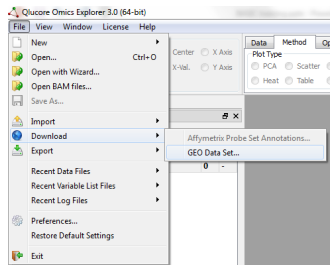


QLUCORE®

Course material – Basic training. © Qlucore 2016

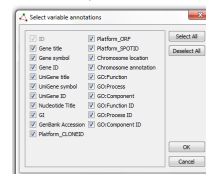
Import a Data Set From GEO

1. File/Download/GEO Data Set



2. Enter Accession Number GSE13425 and Press OK

3. Select Annotations to be included and Press OK



Import of cel files

Preparations:

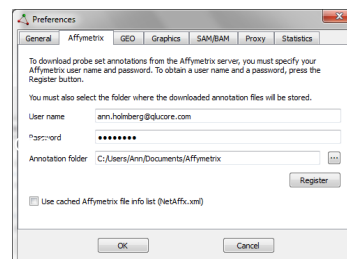
- Register on Affymetrix NetAffx
- Enter you credentials in File/Preferences/Affymetrix

Import:

1. Select File Open, select your files, press OK, select normalization method, press OK

The data will be normalized, log transformed and opened up as a PCA plot

2. Import annotations
see Exercise 12b.



Note: The Qlucore software must be allowed to download information from Affymetrix- if this is prevented by firewall settings etc please contact your IT support.

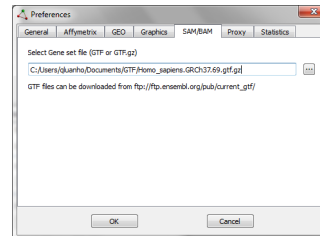
Import of RNA-seq BAM files

Preparations

- Enter the path to the Reference Genome (gtf) in File/Preferences

Import

- Select File/Open BAM files
- Select the files
- Press OK
- The BAM files will be counted, normalized using TMM, log transformed and the data will be opened up in a PCA plot

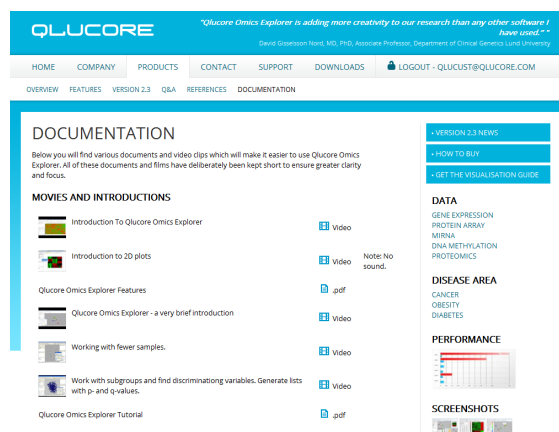


Step2: Import annotations – see Exercise 12b.

Note: The BAM files need to be aligned and sorted on coordinate.
The reference genome must be the same used for the alignment.

More info – www.qlucore.com

- Qlucore Omics Explorer
 - Tutorial
 - Reference manual
- Homepage
 - Films
 - How to documents
 - FAQ
- Monthly webinars
 - Register on the homepage



Contact/Support

- Please feel free to contact us if you have questions.

sara.strandberg@glucore.com

support@glucore.com