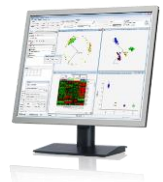
QLUCORE

Qlucore Omics Explorer



Basic Training
&
Data Integration Workshop

NCI June 2nd 2015

Carl Johan Ivarsson
Sara Strandberg
310 945 3047
sara.strandberg@qlucore.com

Agenda June 2nd

- Introduction and Live demonstration — 30 min
- Basic Hands-on Training including Data Integration mRNA/methylation — 2 h

QLUCORE

Objectives: Qlucore Basic training Data Integration

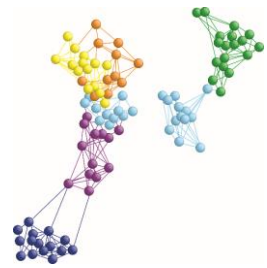
After the training you should be able to do the following types of analyses using Qlucore Omics Explorer:

- Visualize data using PCA
- Identify discriminating variables using basic statistical tests
- Present data with different plot types
- Export variable lists and images
- Integrate data sets (mRNA/Methylation data).

QLUCORE

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
- Multiplex, FACS

Any multivariate data

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)

QLUCORE

GEO Gene Expression Omnibus

Terminology

- Sample
- Variable
- Data set
- Sample annotation
- Variable annotation

QLUCORE

Terminology

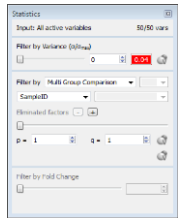
		Array	5301	5302	5303	5304					
		Age	34	22	47	41					
		Gender	Female	Male	Male	Female			Sample annotations		
Probe Set	Symbol								Variable annotations		
201314_at	ENC1		2,28	-1,23		0,45			Data/Results		
204831_at	CDK8		1,04	0	-0,03	0					
205420_at	PEX7		-0,67	3,14	2,18	0,53					
218032_at	SNN		-1,34	2,34	-0,3	2,73					
221289_at	DLX6		2,73	1,07	0,83	-1,52					

QLUCORE

QLUCORE

Visualize and Explore

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

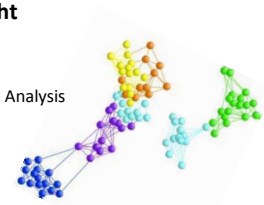


Statistical Analysis

- 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



QLUCORE

Statistical tests

- **F-test (ANOVA - Multigroup comparison)**
  find variables that are statistically different between 2 or more groups.
- **t-test (two group comparison)**
  find variables that are statistically different between 2 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, Wilcoxon included)

QLUCORE

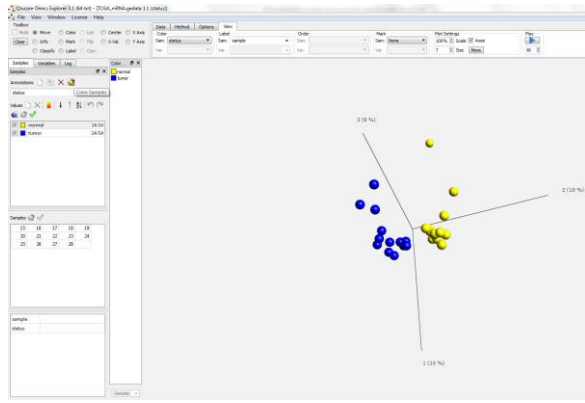
Exercises

1. Visualize (PCA, color according to annotation)
2. Identify discriminating variables (t-test)
3. Create, Save and Export variable Lists and Images
4. Present the result in different plots (heat map and box plot)
5. Data Integration – mRNA/methylation

QLUCORE

Exercise 1 – PCA

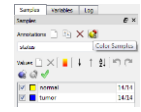
1. Visualize (PCA, color according to annotation)

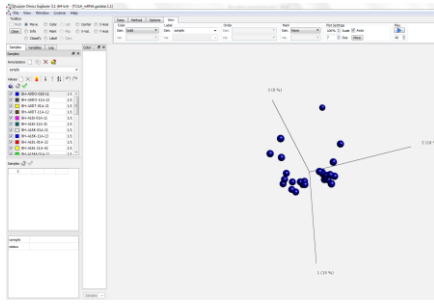


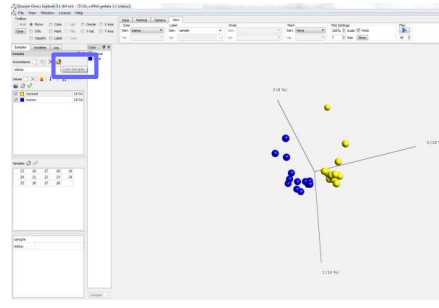
QLUCORE

Steps 1

1. Load mRNA Data set
2. In Sample Tab: Select the Sample Annotation Status
3. Click the Color Sample Button



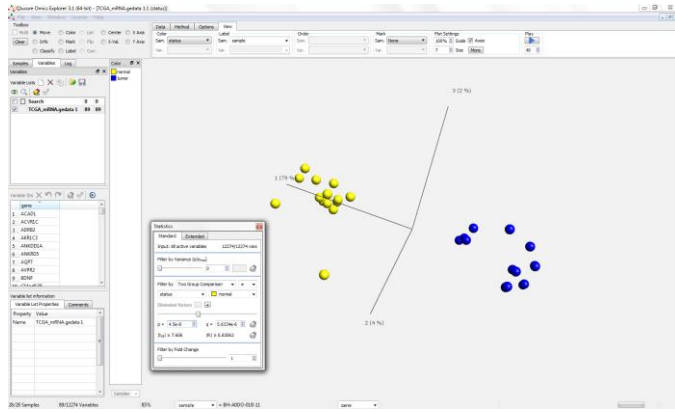




QLUCORE

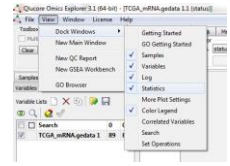
Exercise 2 – t-test

1. Identify discriminating variables (t-test)



Steps 2

1. Start with the Sample PCA plot from exercise 2

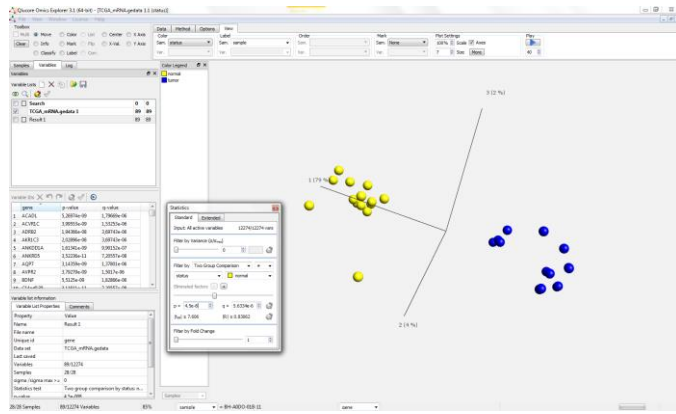


2. Pick up the Statistics dock window from View/Dock Windows

3. Select Two Group comparison in the statistics dialogue.
4. Select Sample annotation Status, compare normal vs tumor
5. Drag the statistical slider to find discriminating variables, stop around 90 variables.

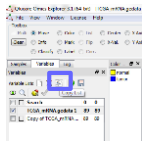
Exercise 3 – Variable list and Images

1. Create and Save and export Variable List.
2. Export image

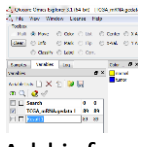


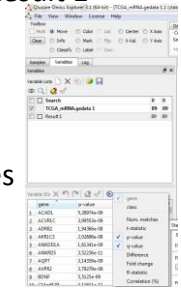
Steps 3

1. In the Variable tab: make a copy of the active list



2. Give it a name by double clicking the list, call it Result 1

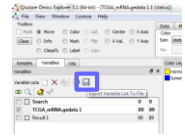




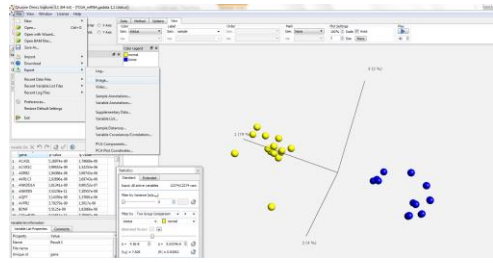
3. Add information to the list, p-values and q-values

Steps 3

4. Export Variable List clicking the Export Icon



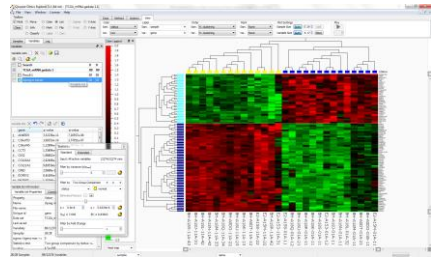
5. Export image File/Export/Image



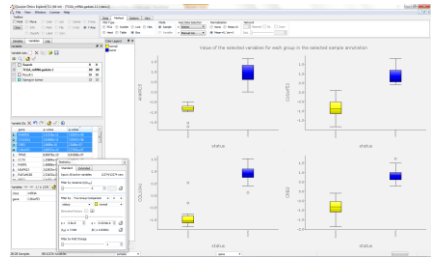
QLUCORE

Exercise 4 – Different plots

1. Generate a Heat map



2. Generate Box plots



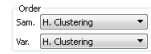
QLUCORE

Steps 4

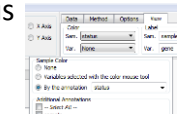
Heat map

1. Start with the Sample PCA plot from Exercise 3

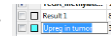
2. Select Heat in the Method tab
3. Select the View Tab and change the order to **hierarchical clustering** for both samples and variables



4. Color the sample according to annotation Status



5. Create a variable list
   - First create a new variable list 
   - Give it a name, by double clicking the list call it *"Upreg in tumor"*



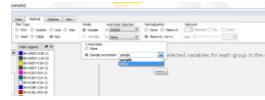
   - Select the List tool in the toolbox 
   - Select upreg variables for tumor sampes by clicking in the heat map.

QLUCORE

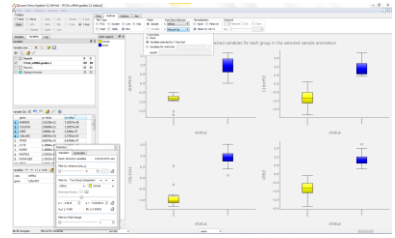
Steps 4

Box Plot

1. Change the plot to Box Plot in the Method tab
2. Set x-axis to Sample Annotation Satus



3. Select variables for y axis with y-axis tool, then hold down shift and click on the top 4 genes in the variables list (make sure the ist is ordered based on q-value by clicking on the header q-value).



QLUCORE

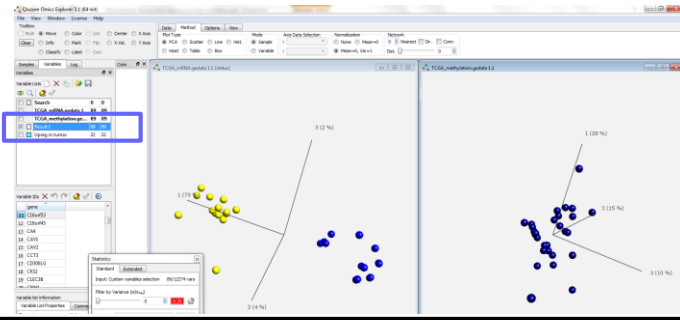
Exercise 5 – Data Integration mRNA/methylation

1. Look for variables from both mRNA and methylation data sets that explains the disease status.
2. Combine mRNA data set with methylation data set.

QLUCORE

Exercise 5 – Data Integration mRNA/methylation

1. mRNA data set: Click on PCA sample in the Method tab
2. **Remove the statistical filtering in the mRNA data → set p-value slider to the left.**
3. Load Methylation data set.
4. Click Ctrl+t (tile the two data sets).
5. Select the checkbox next to the Result 1 list.

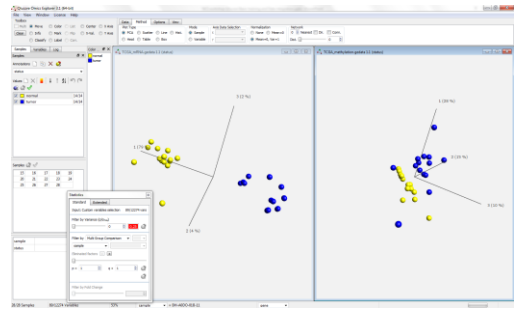


QLUCORE

Exercise 5 – Data Integration mRNA/methylation

5. Color the methylation data set according to Sample annotation Status.

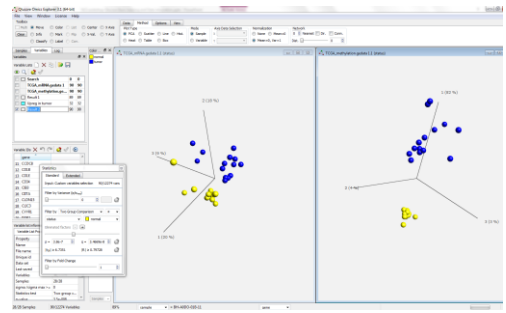
- The methylation PCA plot becomes slightly clearer with **Result 1** as input. The amount of captured variance increases but not that much.
- This is an indication that the identified genes (strongest signal based on t-test) based on the mRNA data do not have methylation levels that are that clear in terms of separating the samples based on the annotation status.



Exercise 5 – Data Integration mRNA/methylation

6. Perform t-test on the DNA methylation data set

- Filter to 90 variables.
- Make a copy of the list and call it Result 2.
- Use the Result 2 list as input by clicking the check box.
- There is a slight pattern in the mRNA data set with the methylation list as the base. We do see a pattern, but not very strong.



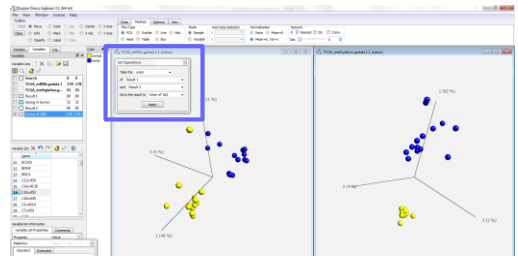
Conclusion: The most significant mRNA and methylation variables do not show corresponding behavior.

Exercise 5 – Data Integration mRNA/methylation

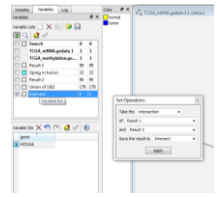
7. Combine generated lists using the Set Operation tool

- Click the set operation tool

- Take the union of the two lists
- Use this list as input. Plots become clearer



- Take the intersect of the two lists.
- Only one gene in common. HOXA4

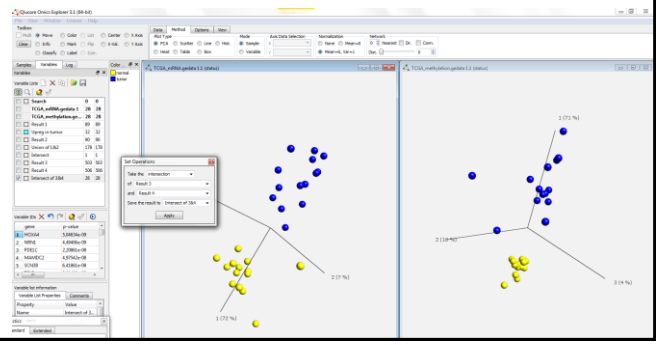


QLUCORE

Exercise 5 – Data Integration mRNA/methylation

Conclusion: The original lists are too short. Let's start with longer lists.

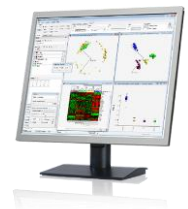
1. Redo the t-tests and create lists with approximately 500 variables long for each data set.
2. Use the set operation tool again select Intersect.
3. This gives a list of 28 candidates. Use the list as input.
4. Shows clear patterns in both data sets.



QLUCORE

Exercise 5 – Data Integration mRNA/methylation

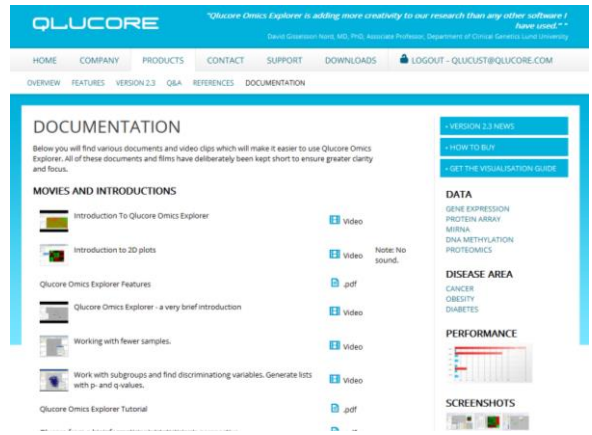
Conclusion: By looking at the experiment using multiple methods more potential candidates that explain the biology emerges.



QLUCORE

More info on www.qlucore.com

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



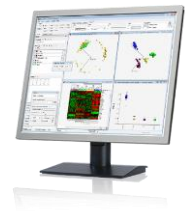
QLUCORE

OSTR Subsidy

https://ostr.cancer.gov/subsidy_program

To obtain access to Qlucore Omics Explorer software, please obtain a quotation from Sara Strandberg. Once a quotation is obtained, please submit a subsidy request for 50% support from OSTR. Upon subsidy approval, the software must be purchased through your Purchasing Agent using your laboratory CAN or Frederick Project Number. Once the order has been placed, the subsidy funds will be transferred to your account.

Bioinformatics Software
Qlucore
Sara Strandberg
Sales Representative
Phone: 310-945-3047
Email: sara.strandberg@qlucore.com
Website: http://www.qlucore.com



QLUCORE

Contact and Support

- Please feel free to contact us with any support questions you might have.

sara.strandberg@qlucore.com

support@qlucore.com

QLUCORE