QLUCORE

How to import data and annotations

Contents

1. INTRODUCTION

This document covers the import functionality of the Qlucore Omics Explorer Base module and covers both import of data and annotations.

2. EXAMPLES ON DATA SOURCES AND APPLICATION AREAS

In this section we outline the data import options for common application areas.

2.1. RNA-SEQ DATA (GENE EXPRESSION)

For gene expression RNA-seq data there are several import options.

a) Aligned BAM files. One for each sample/subject. Import details are available in section 4.
b) A raw count matrix in a .tsv, .txt or .csv file. Import details are available in section 11.

2.2. PROTEOMICS

You will need quantified data in a file or in multiple files. Data can be either normalized or not. The program supports several normalization options. The file extension shall be .txt, .csv or .tsv. These can then be imported using the Wizard, see section 10. Data can be organized in several ways. Both as a matrix and as vectors where data for all samples are presented for each variable after each other.

The data in the .csv file may be organized as vectors (columns) or as a matrix, the Wizard can handle both.

You may also have protein panels. One example is Olink® panels. You can then use the Olink® NPX Signature software to export normalized data from the Olink® panels. Export can be done to Microsoft Excel (.xlsx) files or text (.csv) files. Before importing the .csv into Qlucore Omics Explorer using the Wizard you may need to transpose the data in the .csv file. For more information, please read the document "How to analyze Olink data A.pdf" .

Note: The Wizard can handle import of multiple files and it will work even if the files contain different variables.

If the data has not been log transformed as part of the normalization steps done before importing, it can be done from the Data tab.

The program supports various methods for data preparation, prefiltering and management of missing values. See the reference manual or dedicated documentation.

2.3. SINGLE CELL RNA-SEQ
We support several options for import of single cell data:

a) As aligned BAM files as described in section 4.
b) As a matrix of data in a .txt, .tsv or .csv file (see section 10).
c) For 10X Genomics data using the dedicated templates. Data that has been process with the Cell Ranger. Launch the Template browser and select the template and follow the steps. See section 7.

2.4. QPCR DATA
qPCR data can be imported both as normalized or un-normalized data. If your data is not normalized, then you can normalize the data using Housekeeping normalization The extension shall be .txt, .csv or .tsv. These can then be imported using the Wizard, see section 10.

The program supports various methods for data preparation, prefiltering and management of missing values. See the reference manual or dedicated documentation.

If the data has not been log transformed as part of the normalization steps done before importing, it can be done from the Data tab.Metabolomics and Mass spec Data

You will need quantified and normalized data (concentrations, levels,…) in a file or in multiple files. The extensions shall be .txt, .csv or .tsv.

The data can then be imported using the Wizard, see section 9.

Note: The Wizard can handle import of multiple files and it will work even if the files contain different variables.

If the data has not been log transformed as part of the normalization steps done before importing, it can be done from the Data tab.

2.5. FLOW CYTOMETRY
The easiest way to export features (variables) describing each sample from gating software to QOE is by using .csv files. Counts of different populations is often a natural choice of features to export. However, any feature varying in a relevant way between samples can be used, for examples relative frequencies, mean fluorescence intensities (MFI's) or MFI ratios. You need to ensure that either each column or each row corresponds to a sample.

If you use FlowJo (http://www.flowjo.com) for manual gating you can use a synchronized group to do batch analysis of all samples (http://docs.flowjo.com/d2/workspaces-and-samples/ws-groups/ws-gatecopying/). Then use the Table Editor

(http://docs.flowjo.com/d2/tabular-reports/) to define interesting features for export to a .csv file. To set the gates, you can either use one typical sample as a template, or you can first concatenate all the sample files into a new .fcs file and do the gating on the concatenated file. Using a concatenated file allows you to consider the data in all samples simultaneously.

The data can then be imported using the Wizard, see section 9.

After loading the data, you need to decide which normalization to use and if you should log-transform your data. These decisions should be based on which features you use and what type of patterns/deviations you want to detect in your data. General guidelines are given in the "How to work with flow cytometry data" that can be found on www.glucore.com/documentation.

2.6. DNA AND RNA SEQUENCING (NGS MODULE)
For DNA-seq (and RNA-seq) data the NGS module is required for browsing the genome and detecting variants. DNA-seq data is imported by using the project manager and a set of different files are required to define a NGS project. The minimum requirements are:

- Aligned BAM files to define the sample(s)
- A reference genome (FAST file)
- A GTF file defining the transcripts

Details are presented in the Documentation and Help Manager that is shipped with the program.

2.7. GENE FUSION DATA (NGS MODULE)
From a range of gene fusion callers. Examples are: Gene fusion output files from Arriba, STAR-Fusion, Fusion Catcher or Manta (TSO500, Illumina).

2.8. CHIP-SEQ AND ATAC-SEQ DATA
The peak analysis support in Qlucore Omics Explorer allows for comprehensive analysis of peak data such as ChIP-seq and ATAC-seq. The main components of the peak analysis processing are peak detection, consensus peaks and count matrix generation.

Peak analysis is an NGS feature and requires the NGS module to work. Peak analysis is always performed in the context of a Qlucore NGS project created in the NGS Project Manager.

Qlucore Omics Explorer supports Peak detection with the algorithm Macs2. This will generate one bed type file (narrowPeak or broadPeak) per peak bam file with each feature in it being a peak detected by the Macs2 algorithm.

Qlucore Omics Explorer supports adding a group of input bam-files that can be used by the peak detection to get more accurate results. Although not strictly necessary, it is advisable in most cases to generate input data for ChIP-seq studies.

Details are presented in the Documentation and Help Manager that is shipped with the program.

2.9. TCGA DATA

From the Template browser you can access and start a template that downloads TCGA mRNA gene expression data from GDAC (Broad Institute). The template is called "TCGA RSEM". Start the template and follow the steps.

3. SUPPORTED DATA FORMATS/ARRAYS

This section applies to experiment data. Other sections will cover import of annotations (clinical information) and system biology related information such as gene sets and pathways.

The base module supports direct import of experiment data including normalization for:

- Aligned BAM files with RNA-seq data
- Thermo Fisher (Affymetrix) 3" and WT arrays, including Clariom arrays.
- 10X Genomics files from Cell Ranger pipeline
- Agilent mRNA arrays and Agilent microRNA arrays

For Illumina array data the recommended workflow is to normalize data with the GenomeStudio or BeadStudio software and then use the Wizard in Qlucore Omics Explorer.

For data generated with other arrays/instruments or resulting from other type of sources, Qlucore offers a wide range of import alternatives. If data is stored in a .txt, .csv or .tsv file the Import Wizard, see section 10, is a useful tool for efficient data import. **If you are uncertain of how your data is structured the Wizard is normally the best way to import data.**

In total 11 different data set file formats are supported in the base module:

- BAM files (.bam)
- Affymetrix CEL files (.cel)
- Affymetrix Probe Set Files (.chp)
- Agilent Text Files (.txt)
- Agilent Gene View files (.txt)
- Simple Data Files (.txt)
- Qlucore Data Files (.gedata)
- GEO Data Sets (.soft and .soft.gz) [1]
- GEO Series Matrix (.txt and .txt.gz) [2]
- Compact Text Files (.txt; .csv)
- BioArray Software Environment Files (.base)
- 10X Genomics files from Cell Ranger pipeline ("count" and "aggr")

---

[1], [2] Note that it is only Data Sets from Gene Expression Omnibus in these formats that are supported. Not just any *.soft or *.soft.gz file.

4. HOW TO LOAD AND OPEN FILES

Files are loaded by using the File->Open menu. Then there are different filters that you can select. See the picture below. *Note that the picture might vary slightly depending on your operating system.*
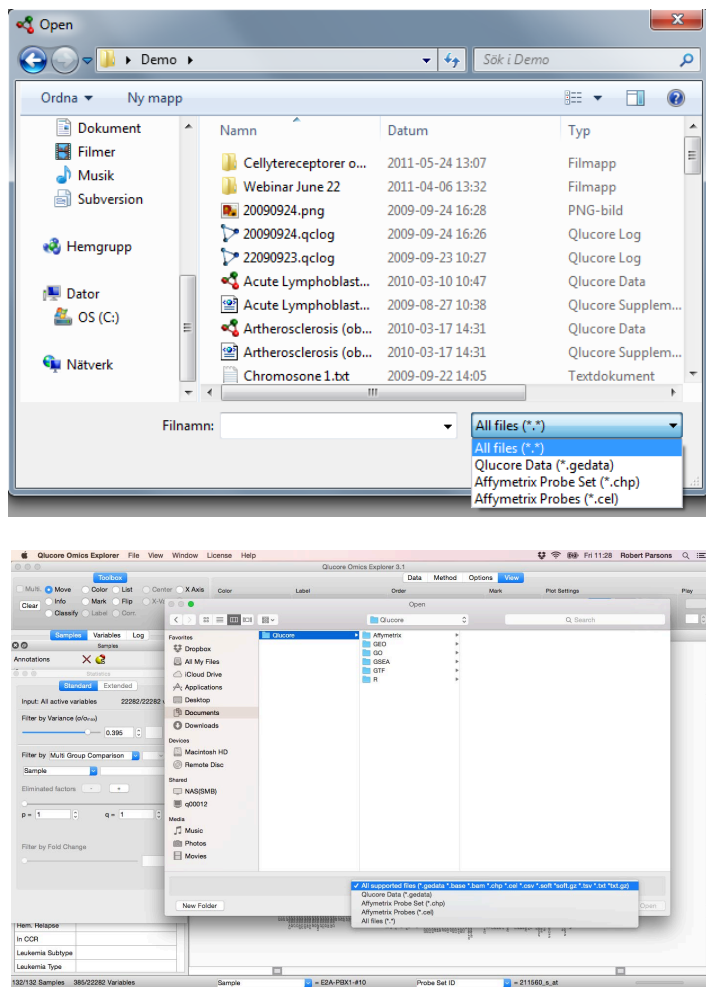
It is also possible to load both variable and sample annotation files, see section 15.

5. ALIGNED BAM FILES

You can import and normalize RNA-seq data from aligned BAM file(s).

Select the **File → Open BAM** files menu item.

Use the Add button to select individual files. Or use the Add Folder button to select all BAM files in a folder.

Press OK to start loading the selected files. This can take a while.

A GTF (Gene Transfer Format) file is needed to describe where the genes are located in the reference genome. file. GTF files can be downloaded from ftp://ftp.ensembl.org/pub/current_gtf/.

Note: The GTF shall based on the same reference genome as the one you have used for alignment.

The Normalization options are controlled from the Data tab after import and are described in the Reference manual.

6. 10X GENOMICS FILES

10X Genomics data is imported with the "Load 10x data" template that is reached from the Template Browser.

Any dataset containing results from the Cell Ranger "count" or "aggr" pipelines work.

You need to provide the path to the directory where data is stored. The directory **must** contain the following 3 files to make the template work:

- barcodes.tsv.gz
- features.tsv.gz
- matrix.mtx.gz

7. H5 FILES

H5 files can be imported with the "10X H5 data import" template. This template is reached from the Template browser. The template imports H5 (HDF5) data processed by 10X Genomics Cell Ranger pipelines 'count', 'aggr', or 'multi'. Both normalized data and raw data can be imported. If data contains integers, it will be assumed to be raw data. Raw data will be filtered and normalized using LogRPM. If the number of cells is 200 000 and above, you get a warning. You can proceed by pressing OK. The H5 data importer can import H5 files generated by Cell Ranger version 2.0 to 7.0, and possibly also higher versions (valid in October 2024).

8. IMPORT WIZARD FOR .TXT, .CSV AND .TSV FILES

The Wizard supports import of data in many different formats. The starting point is that data is stored in a file with extension .txt, .csv or .tsv.

Note: The main benefit of the Wizard is that it helps you to define the parts of a data file that you would like to import. The Wizard accepts data structured and organized in many ways, see examples below. Data can be organized both as a matrix or as vectors where data for all samples are presented for each variable after each other.

	Variable ID	Var 1	Var 2	Var 3	Var 4	Var 5
	Weight	1	2	3	4	5
	Concentratio	200	300	100	500	66
Sample id	Treatment					
ID_1	Control	219.49	423.8	11925	3271.9	5686.5
ID_2	Control	233.23	390.23	10999	3654.3	5760.7
ID_3	Treated	144.9	673.68	9137.7	2876.1	3520.4

		Sample id	ID_1	ID_2	ID_3
		Treatment	Control	Control	Treated
PROBE_ID	SYMBOL				
ILMN_1762337	7A5		219.49	233.23	144.9
ILMN_2055271	A1BG		423.8	390.23	673.68
ILMN_2383229	A1CF		11925	10999	9137.7
ILMN_1806310	A1CF		3271.9	3654.3	2876.1
ILMN_1779670	A1CF		5686.5	5760.7	3520.4

		Sample id	6116725068_A.AVG_Signal	6116725068_A.FC	6116725068_C.AVG_Signal	6116725068_C.FC	
		Treatment	Control		Control		
PROBE_ID	SYMBOL						
ILMN_176	7A5		219.49	1,97	144.9	2,43	
ILMN_205	A1BG		423.8	4,16	673.68	0,87	
ILMN_238	A1CF		11925.00	2,00	9137.7	0,66	
ILMN_180	A1CF		3271.9	0,50	2876.1	2,57	
ILMN_177	A1CF		5686.5	2,33	3520.4	3,67	
ILMN_171	A26C3		152.1	0,47	134.27	1,85	
ILMN_181	A2BP1		192.01	1,23	136.38	1,47	

Examples of different structures and layouts that are accepted. Transposed or not, with extra sample columns or not

	Instrument	Machine special # 3					
	Date	2019-01-06					
				Sample id	ID_1	ID_2	ID_3
				Treatment	Control	Control	Treated
Variable ID	Weight	Concentration	Link				
Var 1	1	200	www.ggg.org		219.49	233.23	144.9
Var 2	2	300	Blank		423.8	390.23	673.68
Var 3	3	100	Blank		11925	10999	9137.7
Var 4	4	500	Blank		3271.9	3654.3	2876.1
Var 5	5	66	Blank		5686.5	5760.7	3520.4

Example on different structures and layouts that are accepted, with header.

Sample ID	OlinkID	LOD	NPX
Sample_01	OID20838	0,2178	-0.1219
Sample_02	OID20838	0,2178	-0.0869
Sample_01	OID20839	0.9848	0.1711
Sample_02	OID20839	0.9848	-0.1490
...			

Example of vector layout

The samples can be in one or multiple files. If the samples are in multiple files, the file format for all samples must be identical. You reach the Wizard by the menu **File->Open with Wizard.**

Note: After import data can be pre-processed and normalized in various ways to fit your type of data.

The Wizard guides you through a number of steps where you define how the layout of the data shall be interpreted and imported. The Wizard is very flexible and can for example handle

- Samples as columns and variables as rows, or the transposed version
- Samples in multiple files, with or without the exact same number of variables in each file.
- Count and continuous data
- Various data separators [",", ";", "tab" or other]
- Data on every ith column or every ith row.
- File header or not
- Data that shall not be imported
- Various indications of missing values [empty cell, n/a, NA,…]
- Annotations

If you are uncertain of how your data is structured, then the Wizard is the best option for data import.
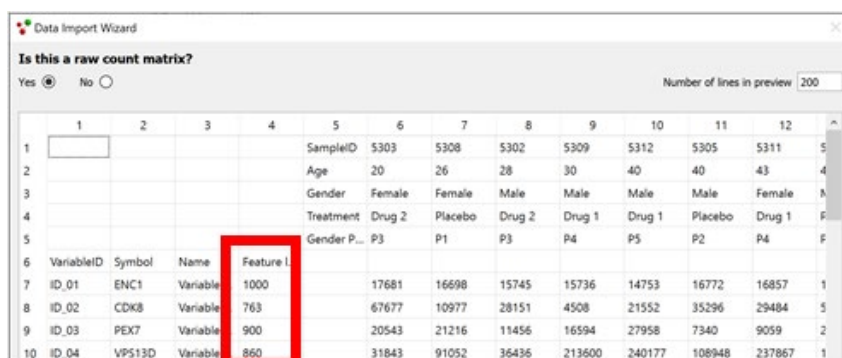
Note: The Import Wizard is an automatic import method for very many different types of files. Make sure to verify that the import results in the expected number of samples and variables.

9. IMPORT WIZARD FOR RAW COUNT DATA

The starting point is that data is stored in a matrix in a file with extension .txt, .csv or .tsv.

In the second step in the Wizard you define that you have count data. There are several Normalization methods, these are from version 3.10 selected dynamically in the Data tab. This is done after the data is imported. After the data is imported you also need to provide the length of each feature (gene lengths in base pairs for gene data). That is also done in the Data tab.

Note: When you select variable annotations to import the feature lengths annotation shall be included if you would like to use it later.

Example of count matrix - the column marked in red contains the gene lengths.

10. QLUCORE DATA FILES (.GEDATA)

The Qlucore Data file format has been designed to make it easy to create a file by copying and pasting in a spread sheet application. Data resulting from other tools and applications can easily be formatted and imported into Qlucore Omics Explorer using this file format. A Qlucore Data File contains a complete data frame, including sample and variable annotations. [3]

There are three important things to note when a file shall be created

1. It is a tab-separated text file.
2. The file shall have the file name extension .gedata.
3. Data shall be formatted as described below

11. R USERS

The program offers the possibility both to import R data files as data sets, and to export an open data set as an R data file. The R file protocol used is a version of the .rds format. To import an R data file go to File->Open R Data Frame. Note: The .rds file must contain a particular set of data frames as described in the reference manual.

Additionally two R-scripts are shipped with Qlucore Omics Explorer. They make it easy to convert data back and forth from R.

The script function definition files are located in the user's Documents folder. The path is Documents/Qlucore/R-import/, where the two files read.gedata.1.1.R and write.gedata.1.1.R can be found.

For detailed description on how to use the scripts see the "R to .gedata" section in the Documentation and Help system.

12. THERMOFISHER (AFFYMETRIX) FILES

This applies to mRNA array based data. For miRNA data, only normalized data (as a .txt, .tsv or .csv file) is accepted, and the suggested workflow is to import the normalized data using the Wizard, see section 10.

12.1. THERMOFISHER (AFFYMETRIX) CEL FILES (.CEL)

The import and normalization will normally include four steps

1. Select files
2. Select normalization method
3. Select which annotations to import
4. Inspect the QC-report

---

[3] A Qlucore data file does not contain information about the colors used to represent the values of a sample attribute in a plot, or the default variable annotation used when matching variable name lists with variable annotations in the data frame. That information is stored in supplementary data files.

To start the import: Select File->Open and select the files you would like to open. To select multiple files, use the shift key on the keyboard.

The data import process includes downloading information about the array from a ThermoFisher server. You need to be connected to Internet if you do not have this data stored locally.

After a short while you are asked to select normalization method. Three methods are provided

- RMA
- RMA Sketch
- Plier

The next step is to download annotations from the Thermo Fisher NetAffx server [4] . You can select which annotations to include in your data set.

Finally, the QC-report is displayed.

In OE, normalization and summarization will be performed on gene (transcript level) both for 3" arrays and WT arrays.

A wide range of different arrays for a wide range of species are supported, this includes for instance the newer Clariom arrays as well as older U133 arrays of various generations.

12.2. THERMO FISHER (AFFYMETRIX) PROBE SET FILES (.CHP)

Select **File->Open** and select the files you would like to open. To select multiple files, use the shift key on the keyboard.

You will be asked if you would like to download annotation data from Affymetrix NetAffx server automatically [5]. When loading probe set annotations from Affymetrix Whole Transcript arrays, such as Human Gene 1.0 ST, Omics Explorer automatically creates simplified annotations (Gene Symbol, Gene Title,…) in addition to the annotations in the Affymetrix annotation file. This makes it easier to interpret the results and to search for interesting genes.

13. AGILENT FILES

13.1. AGILENT TEXT FILES

Omics Explorer can import text files created with Agilent Feature Extraction Software. These files contain one array (sample) per file.

How to open Agilent text files

[4,4] This requires that the user is registered at the Thermo Fisher(Affymetrix) website. See section 15.3 for more information about downloading this website.

Select the **File→Open** menu item.

Select the Agilent Text File (*.txt) file filter.

Select the files and press Open. (Normally you select several files.)

In the Normalization Dialog, select a preprocessing scheme by pressing one of the four quick setting buttons.

- One-color mRNA
- One-color miRNA
- Two-color mRNA
- Two-color miRNA

Instead of using the above quick settings, you can select your own customized preprocessing. See the OE reference manual, section "Agilent .txt files", for details.

13.2.   AGILENT GENEVIEW FILES (.TXT)

Qlucore Omics Explorer can open Agilent GeneView files. These files contain miRNA expression levels. They are created with the Agilent Feature Extraction Software. Each file contains data from a single array. To select multiple files use the shift key on the keyboard.

Normalization

The data is normalized as recommended by Agilent:

- thresholding
- 2-logarithm
- percentile shift (optional, but recommended)

You can select the threshold value and the percentile. Recommended values are threshold = 1.0 and percentile = 75.

13.3.   HOW TO CREATE A QLUCORE DATA FILE

There are several ways to start the creation of the data file (.gedata).

a) If you have installed OE on your computer you can create a new data file template by right clicking on your desktop and selecting New and Qlucore Omics Explorer. This will generate a new file on your desktop. If your right click on the new file you can open the file with Excel or another spreadsheet application. Then you need to copy and paste data from your original data file into the new file. A guide on how to do this follows below in the section Formatting.

b) Open your original data file and include rows and columns as described in the section Formatting.

13.4.   FORMATTING

Copy and paste the data into the spreadsheet, following the layout below. You can use any spreadsheet application, such as Excel or OpenOffice. When you are done and have formatted data as below, go through the following three steps.

1. Select 'Save as' in the spread sheet application
2. Save the spread sheet as a tab delimited text file.
3. Rename the file by changing the file name extension from the default .txt to .gedata.

13.5. LAYOUT

The layout of the spreadsheet is best explained through an example. The data below contains

1. The measured data (cells are blue)
2. Sample annotation IDs: Array ID, Age, Sex (cells are pink)
3. Sample annotations (cells are yellow)
4. Variable annotation IDs: VarID and Symbol (cells are light blue).
5. Variable annotations (cells are green)

In a spread sheet application, this data frame should be arranged as follows:

Qlucore	gedata	version 1.0						
4	samples	with	3	attributes				
5	variables	with	2	annotations				
		Array ID	5301	5302	5303	5304		
		Age	34	22	47	41		
		Sex	Female	Male	Male	Male		
varID	Symbol							
3140	MR1		2.28	-1.23		0.45		
622	BDH1		1.04	0	-0.03	0		
7551	ZNF3		-0.67	3.14	2.18	0.53		
1537	CYC1		-1.34	2.34	-0.3	2.73		
961	CD47		2.73	1.07	0.83	-1.52		

Always included

Note: The start of the file shall include the text and information as indicated by row 1 to 5. For your own data you need to adjust the number of samples, variables and annotations.

13.6. TRANSPOSED DATA

The example given above has samples as columns and variables as rows. For some data it is more natural to have the opposite, i.e. samples as rows and variables as columns.

Qlucore OE is capable to import .gedata files with data organized in this way if the word "transposed" is added to cell D1, see the table below. Note that the table is not complete.

Qlucore	gedata	version 1.0	transposed			
5	samples	with	2	attributes		
4	variables	with	3	annotations		
		Variable ID	101	102	103	104

14. SIMPLE DATA FILE

The simple data file is a tab separated text file with the extension .txt. The table below shows how data and identifiers should be organized. The rows are variables and the columns are samples. The first column should include a unique variable identifier (green cells) and the first row a unique sample identifier (yellow cells). Data is stored in the blue cells.

	5301	5302	5303	5304	
3140	2.28	-1.23		0.45	
622	1.04	0	-0.03	0	
7551	-0.67	3.14	2.18	0.53	
1537	-1.34	2.34	-0.3	2.73	
961	2.73	1.07	0.83	-1.52	

If there is a text in the upper left cell, that information will be used as the variable annotation header, otherwise the name will be *Variable ID*. The sample annotation header, corresponding to the sample annotation in the first row, will be the *Sample ID.*

Normally you would like to complement your data import with annotations, see chapter 10.

15. COMPACT TEXT FILE (.CSV)

Omics Explorer supports compact text files. These contain data, sample annotations and variable annotations for multiple samples in a single file. If compared to a ".gedata" file, see chapter 7, the compact text file is lacking a file header and has samples as rows. The layout of the files is described below. The last variable annotation (in the example below the second

row) will be used as variable id. Note that the name of this annotation is not included in the file; after loading the file the annotation will be called Variable Id.

The file can have either the extension ".txt" or ".csv", both options work.

		Symbol	ENC1	CDK8	PEX7	SNN	DLX6
Array	Age	Gender	201314_at	204831_at	205420_at	218032_at	221289_at
5301	34	Female	2.28	1.04	-0.67	-1.34	2.73
5302	22	Male	-1.23	0	3.14	2.34	1.07
5303	47	Male	1.45	-0.03	2.18	-0.3	0.83
5304	41	Female	0.45	0	0.53	2.73	-1.52

How to open a compact text file

Select the **File → Open** menu item.

Select the file and press Open.

If the file contains only one variable annotation you will be asked to enter the index of the first data column (in the example above it is column 4).

16. IMPORT SAMPLE AND VARIABLE ANNOTATIONS

Annotations are information about your samples and/or variables. It can for instance be clinical information about patients. The description below is a summary of options. More details are provided in the section with the same name in the Documentation and Help manager in the program itself. In addition to importing sample annotations they can be added through manual input in the program. See the "How to add sample annotations" document.

Note 1. Many times, annotations are imported with the data. This section covers the import of additional annotations.

Note 2. To import annotations a data set first should be loaded first.

16.1. FROM A .TXT OR .CSV FILE

There are two options. Using the Annotation Wizard or importing a file of the format specified below.

Using the Annotation Wizard

Select the **File → Import → Sample Annotations via Wizard** or the **File → Import → Variable Annotations via Wizard** menu item.

The Annotation Wizard is launched and by following the instructions it is possible to import data from *.txt, '.csv or *.tsv files. The Wizard will assist you to pick out the data you need.

Direct From a .txt or .csv file

You can add sample and variable annotations to an existing data set by importing an annotation text file. These files are tab- or comma-delimited text files. The file name extension is .txt if the file is tab-delimited and .csv if the file is comma-delimited.

How to import a sample/variable annotation text file

> Select the **File → Import → Sample Annotations** or the **File → Import → Variable Annotations** menu item.
>
> Select the Tab-separated Text (*.txt) or the Comma-separated Text (*.csv) file filter.
>
> Select the file and press Open.
>
> Select the annotations you want to import and press OK.

When importing an annotation text file, the samples/variables in the file will be matched with the samples/variables in the data frame using the annotation in the first column of the annotation file and the sample/variable ID annotation in the data set (This can be changed in the Data tab). In this way the ordering of the rows in the annotation file and the data frame does not matter.

16.2. CREATE A SAMPLE/VARIABLE ANNOTATION TEXT FILE

Place the data in a spreadsheet, following the layout below. You can use any spreadsheet application, such as Excel or OpenOffice. Save the spreadsheet as a tab- or comma-delimited text file. Use the file name extension .txt if the file is tab-delimited and .csv if the file is comma-delimited. We recommend using tab-delimited text files.

The following layout should be used:

Probe Set	Transcript	Gene Symbol	Chromosome	Entrez Gene
1007_s_at	U48705	DDR1	chr6p21.3	780
1053_at	M87338	RFC2	chr7q11.23	5982
117_at	X51757	HSPA6	chr1q23	3310
121_at	X69699	PAX8	chr2q12-q14	7849
1255_g_at	L36861	GUCA1A	chr6p21.1	2978
1294_g_at	L13852	UBA7	chr3p21	7318

The first row contains the name of each annotation. Each of the remaining rows contains the value of each annotation for one sample/variable. The first column should contain an annotation that matches the sample/variable ID in the data set.

This table should be saved as a tab-separated text file with file name extension .txt or as a comma-separated text file with file name extension .csv.

16.3. FROM NETAFFX

You can download variable annotations for Affymetrix arrays from the Affymetrix NetAffx server and add them to the data set.

The downloaded annotation files will be stored on your computer, and can be used without having to download them from the server again. When you select to download annotations, you will be notified if updated annotations are available on the Affymetrix server.

When downloading probe set annotations for the Affymetrix whole transcript arrays, Omics Explorer will create simplified annotations (Gene Symbol, Gene Title,...) in addition to the annotations in the downloaded annotation file.

How to download and import annotations from the Affymetrix NetAffx server [6]

Select the **File → Download→ Affymetrix Probe Set Annotations** menu item.

Select the array type and press OK .

Select the annotations you want to import and press OK.

17. TEMPLATES

Templates are scripts that enables a pre-configured set of operations to be executed in the program. Templates are based on Python.

Templates can be opened from **File → Execute Template.**

The Template Browser (**File → Template Browser**) shows an overview of all templates available at the location(s) specified in the Template section of the Preferences.

Current standard templates include import of TCGA data and 10X Genomics Single cell data.

---

[6] The first time you use this function, you will also be asked to provide: your *Affymetrix user name* and your Thermo Fisher (*Affymetrix) password* and the folder where the downloaded annotation files will be stored. If you do not have an Affymetrix user name and password, you can get one by pressing the Register button. This will take you to the registration page on the Affymetrix web site.

TERMINOLOGY

Samples: We use samples to describe units such as patients, persons, animals, treatments, dates,…

Variables: We use variables to describe quantities that have been measured for each samples, such as Gene Expression Levels, Protein concentrations, answers to a question in a questionnaire,…

Annotation: A description of a sample or a variable. One sample or variable can be described by one or many annotations.

DECIMAL POINT AND DECIMAL COMMA

Both separators are supported in all tab separated .txt files and .csv files.

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 only intended for research purposes.

TRADEMARK LIST

Excel, Windows 10 and Windows 11 are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

Olink and NetAffx is the trademark of Thermo Fisher Scientific (Affymetrix).

GenomeStudio and BeadStudio are trademarks of Illumina

10X, 10X GENOMICS and CELL RANGER are trademarks of 10X Genomics, Inc.