

How to handle different types of experiment designs

The general framework for statistical tests in Qlucore Omics Explorer is flexible enough to handle many experiment designs. In this document we will cover the following examples:

- **Paired data:** Each subject has received two different treatments.
- **Multiple treatments:** Several subjects that each has received several treatments.
- **Multiple treatments with before and after measurements:** Several subjects have received a number of different treatments. For each subject and each treatment there is a sample taken before the treatment and a sample taken after the treatment.

The more extensive document “How to use ANOVA in Qlucore” covers all the above examples as well as additional models.

The built-in statistical methods handle all examples discussed above. It is also possible to enable the Extended statistics functionality using the Open API to R. See the Documentation and Help Manager for more information.

Note that the examples below is constructed using a random data set and hence the q-value is always high.

TERMINOLOGY

We use samples to denote units such as patients, persons, study participants, animals, plants, cells,...

We use variables to denote quantities that have been measured for each sample, such as gene expression levels, miRNA expression levels, protein concentrations, antibody concentrations, methylation levels, answers to questions in a questionnaire, ...

Normally a data set is described by a matrix where the columns represent samples and the rows represent variables.

EXAMPLE 1: PAIRED DATA

A common experiment design is paired data, where each subject has received two different treatments. Thus, there are two samples from each subject.

For paired data one should not do an ordinary Two Group Comparison of the two treatments, because differences between the subjects may obscure the difference between the two treatments. Instead one should do a *paired test*. In Qlucore Omics Explorer this is done by eliminating the contribution from the subject factor and then doing an ordinary Two Group Comparison.

Depending on whether Two Group Comparison or Multi Group Comparison is chosen, Qlucore Omics Explorer will select a t- or F-test with the correct number of degrees of freedom. This procedure is mathematically equivalent to a paired t-test or a repeated measures ANOVA.

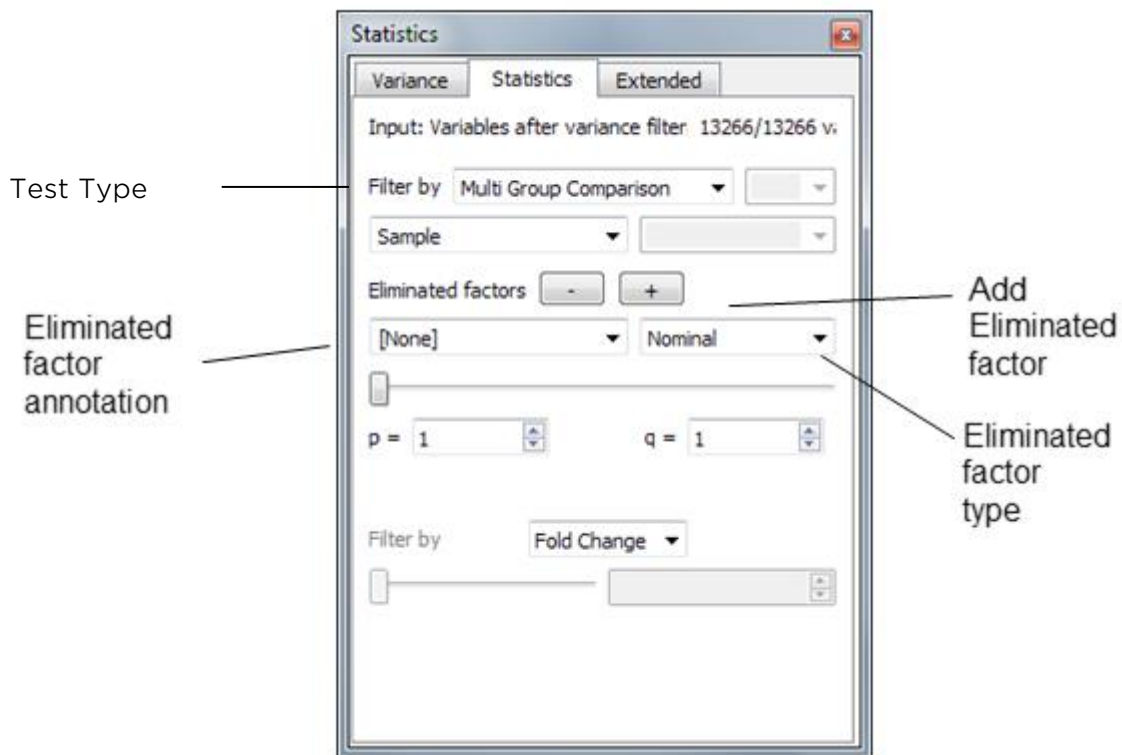


Figure 1

How to analyze paired data

- Identify the relevant sample annotations. There should be one sample annotation that defines the subjects and one that defines the two treatments.
Sometimes you may have to edit the sample annotations to get them in this form. There may for instance be a single sample annotation that defines both subject and treatment.
- Select *Two Group Comparison* in the *Test Combobox*.
- Select a direction (<, > or ≠) in the *Test Direction Combobox*.
- Select the annotation that defines the two treatments in the *Test Annotation Combobox*.
- Select one of the two treatments in the *Group One Combobox*.
- Press the *Add Eliminated Factor Button*.
- Select the annotation that identifies the subjects in the *Eliminated Factor Annotation Combobox*.

- Select *Nominal* in the *Eliminated Factor Type Combobox*.
- Select a significance level or q-value (False Discovery Rate).
- Potentially also select the required Fold Change.

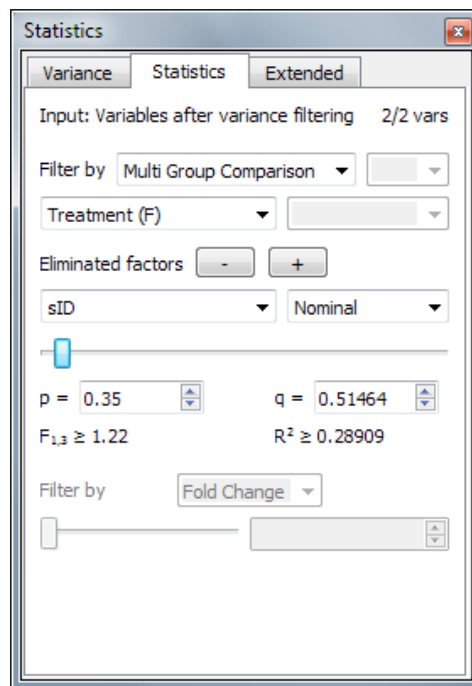
You can also select *Multi Group Comparison*. That has the same effect as selecting *Two Group Comparison* and \neq .

Samples that are not part of a pair (i.e. subjects for which one of the two treatments is missing) will automatically be eliminated from the analysis.

EXAMPLE 2: SEVERAL SUBJECTS THAT EACH HAS RECEIVED SEVERAL TREATMENTS

A number of subjects have received a number of different treatments. There is one sample for each subject and each treatment. We want to find variables that are associated with the treatment factor.

To do this, we proceed as for paired data in example 1 above, but since there are more than two treatments, we select *Multi Group Comparison* instead, see *Figure 1* and test type.



In the example above is the relevant sample annotation "Treatment(F) and the annotation that identifies the subject "sID".

EXAMPLE 3: SEVERAL SUBJECTS BEFORE AND AFTER EACH OF SEVERAL TREATMENT LEVELS

A number of subjects have received a number of different treatments. For each subject and each treatment there is a sample taken before the treatment and a sample taken after the treatment.

Say that there are a number of subjects and 3 treatments, and that the information is provided by three sample annotations as follows:

- *Subject* with values: *A, B, C, ...*,
- *Treatment* with values: *X, Y, Z*.
- *Type* with values: *0* (=before) and *1* (= after).

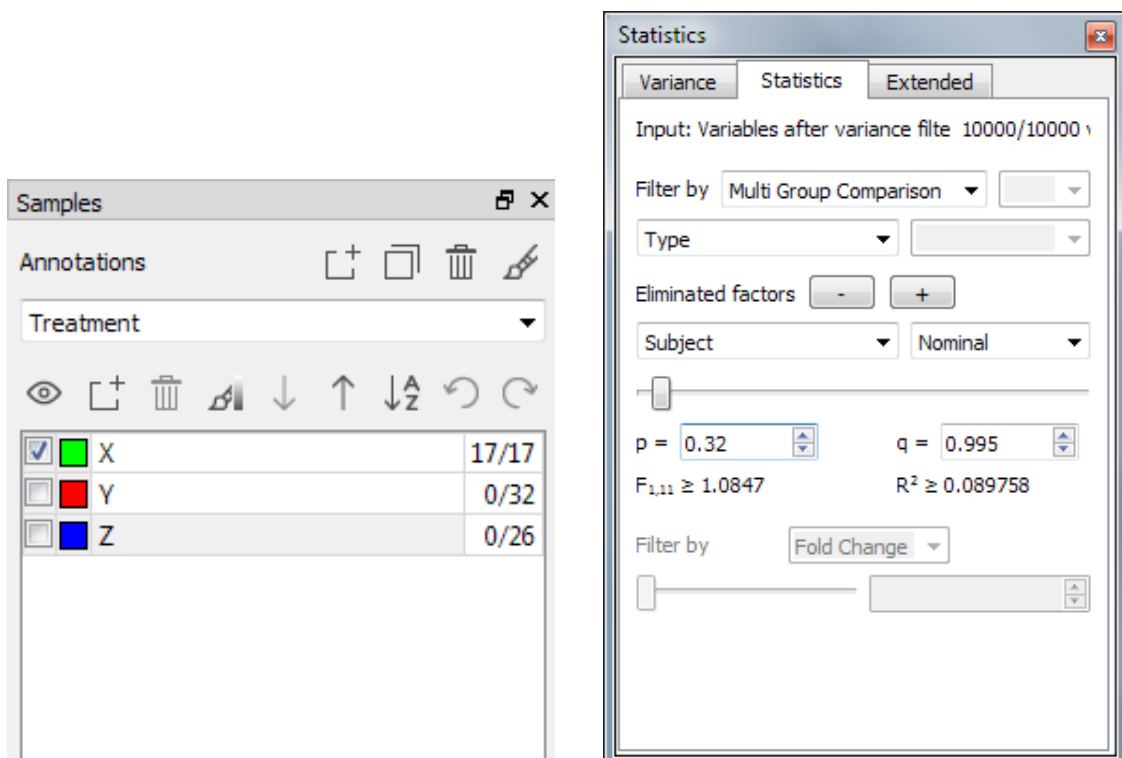
We can then form interaction annotations, such as

- *Subject x Treatment* with values: *AX, AY, AZ, BX, BY, BZ, ...*,
- *Subject x Type* with values: *A0, A1, B0, B1, C0, ...*,
- *Treatment x Type* with values: *X0, X1, Y0, Y1, Z0, Z1*

The data is paired and the pairs are defined by the *Subject x Treatment* annotation. Each pair contains one sample of *Type 0* and one sample of *Type 1*.

3.1. Find variables that respond to treatment X

To answer this question, we let the treatment X samples be the active samples and exclude the other two treatments. Thus, we include both *X0* (= before X) and *X1* (= after X). Then the pairs are defined by the *Subject* annotation. We can now proceed exactly as in Example 2:



The image shows two windows from a software application. The 'Samples' window on the left displays a table of annotations for 'Treatment'. The 'Statistics' window on the right shows analysis results for 'Multi Group Comparison'.

Samples		
Annotations		
Treatment		
<input checked="" type="checkbox"/>	X	17/17
<input type="checkbox"/>	Y	0/32
<input type="checkbox"/>	Z	0/26

Statistics

Variance | Statistics | Extended

Input: Variables after variance filter 10000/10000

Filter by Multi Group Comparison

Type

Eliminated factors - +

Subject Nominal

p = 0.32 q = 0.995

$F_{1,11} \geq 1.0847$ $R^2 \geq 0.089758$

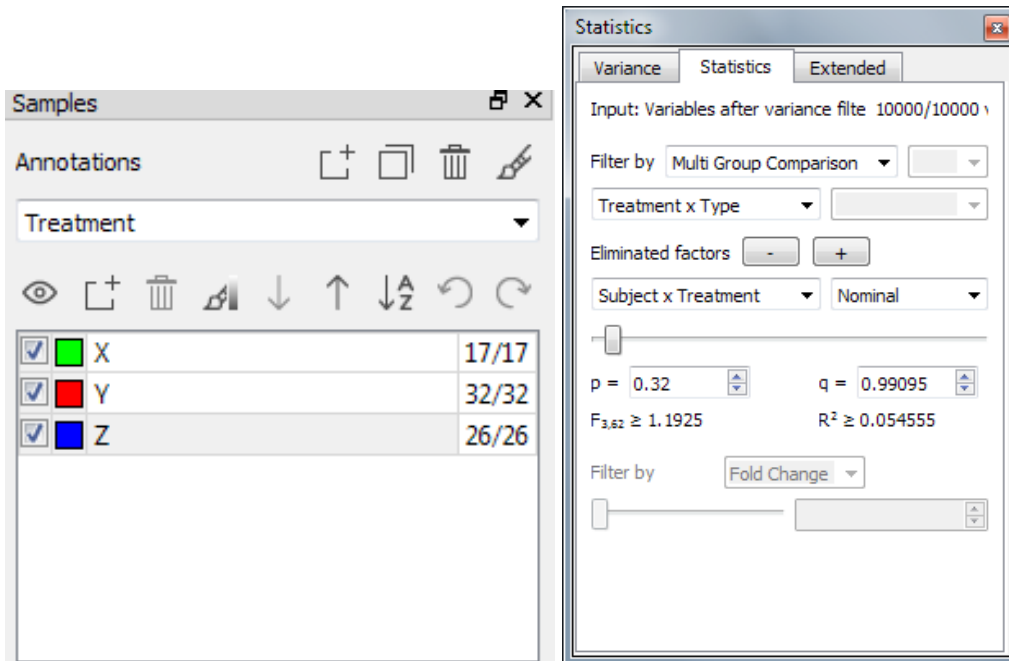
Filter by Fold Change

In this case one could also do a *Two Group Comparison*.

3.2. Find variables that respond to any of the three treatments

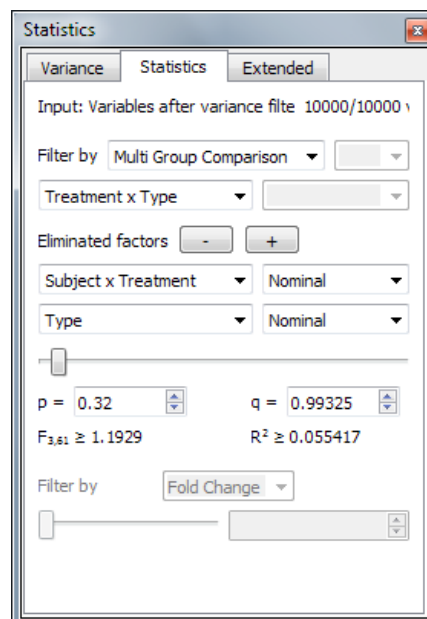
We are now using all samples. Then the pairs are defined by the *Subject x Treatment* annotation. Each pairs consists of one sample of *Type 0* (=before) and *1* (= after). We are

interested in effects given by the *Treatment x Type* annotation. Thus we use the following settings:



3.3. Find variables that do not respond the same way to all three treatments

This is similar to 3.2, but now we want to exclude variables that respond the same way to all three treatments. If a variable responds the same way to all three treatments, that response then can be modelled by the annotation *Type*. Thus we proceed as in 3.2 but we exclude the annotation *Type* as well:



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.