

# OmicsQ — Documentation and Tutorial

Veit Schwämmle

4/28/25

## Table of contents

Introduction . . . . .	2
Step I: Data Import . . . . .	3
1. File Input . . . . .	3
2. File-Specific Settings . . . . .	4
3. Select and Adjust Your Columns . . . . .	4
4. Data Preview . . . . .	5
5. Proceed to Experimental Design . . . . .	6
Step II: Experimental Design . . . . .	7
1. Automatic Selection of Experimental Groups . . . . .	7
2. Manually Assign Sample Types & Replicate Numbers . . . . .	7
3. Annotating Batches . . . . .	8
4. Experimental Design Table . . . . .	8
5. Proceed to Data Pre-Processing . . . . .	9
💡 Tips & Reminders . . . . .	9
Step III: Data Pre-Processing . . . . .	9
1. Add or Delete Data Columns . . . . .	10
2. Data Manipulation & Adjustment . . . . .	10
3. Summary & Proceed . . . . .	10
4. Visual Feedback . . . . .	11
💡 Tips . . . . .	12
Step IV: Further Integrated Analysis . . . . .	12
1. Sending Data to External Apps . . . . .	12
2. Paired vs. Unpaired Design . . . . .	13
3. Working with Results . . . . .	13
4. Details about the External Apps . . . . .	14



## Introduction

OmicsQ is a modular Shiny application that streamlines the workflow of quantitative omics data analysis *without any need to impute missing values*. This tutorial will guide you through all the major steps — from importing your data, annotating your experimental design, and applying pre-processing methods, to sending your finalized dataset to dedicated external apps for *statistical testing, clustering, or complex behavior analysis*.

## General remarks

*Online documentation:* Further information about the different operations of the user interface are available through the information (i) buttons.

*Recording carried operations:* To ensure reproducibility of the results, all data processing steps are recorded in the **Summary of operations**.

*Brief overview:* Each of the following sections is represented by a different tab in OmicsQ:

### Step I: Data Import

- Upload and verify your files (e.g., Excel or CSV/TSV).
- Adjust file-specific settings such as delimiters, decimal separators, etc.

### Step II: Experimental Design

- Automatically or manually assign sample groups, replicates, and batch labels.
- Ensure your design is correct.

### Step III: Data Pre-Processing

- Apply transformations (e.g., logarithmic), normalization, and summarization.
- Remove features or samples with too many missing values

- Check and correct for batch effects.

#### Step IV: Further Integrated Analysis

- Send processed data to specialized tools, such as [PolySTest](#) for statistical analysis, [VSClust](#) for clustering, or [ComplexBrowser](#) to explore protein complexes.
- Retrieve and incorporate results back into OmicsQ for having a single file with all information.

*Layout:* Each of these sections consist of documentation of the different features and a tutorial part. The tutorial part is meant to guide you through the different steps of the analysis. By the end of this tutorial, you will have a fully processed dataset aligned with your experimental design and enriched with features for further biological interpretation — all managed within a single Shiny application. Whether you are working with proteomics, transcriptomics, or other omics data, OmicsQ’s flexible modules offer an interactive and user-friendly approach.

#### Literature and further materials

The OmicsQ manuscript: *in preparation*

For the publications of the other apps, see the section [About the external Apps](#)

### Step I: Data Import

In this part of OmicsQ, you will upload your data file containing the quantifications of biomolecules across different samples and make any quick, simple adjustments to ensure the input data is ready for further analysis. Below are the features you see on your screen and how to use them:

#### 1. File Input

- *Upload your file:* Click *Browse* to select a file from your computer. Alternatively, you can use the example file. Supported file types typically include Excel (*.xls* or *.xlsx*) and text-based formats such as *CSV* or *TSV*.
- *Run example file:* Click *Run example file* to load a ready-to-use test dataset. This is useful if you want to explore the interface and its features.
- *Feedback and issue reporting:* Find the link to OmicsQ’s open source software repository for raising issues, or to contact the main developer for additional support. There is more information via the *!About* button in the upper right of the user interface.

## Tutorial

Download the example file containing protein abundances of an cell differentiation experiment from the GitHub repository [here](#). You can use the *download raw file* button for direct download. Then upload the file into OmicsQ.

## 2. File-Specific Settings

Once you upload a file, OmicsQ will automatically detect if it is an Excel or a CSV/TSV file and determine the layout. Depending on the file type, different settings appear:

- *For Excel:* You may select the *sheet* if your Excel file has multiple worksheets.
- *For CSV/TSV or other text files:* You will see options to specify:
  - *Delimiter* (e.g., commas, semicolons, tabs)
  - *Decimal separator* (e.g., period . or comma ,)
  - *Skip lines at the beginning* to ignore extra header lines
  - *Indicate if the file has a header row*

If the automatic detection looks correct, you can keep the default settings. If not, simply adjust these fields to properly format your data.

## Tutorial

The file should have been read correctly. Observe that there are many duplicated values which could be of concern.

Normally, we do not expect duplicated values when the instrument provides sufficient precision in its measurements. However, in this case the values have been rounded quite stringently, leading to the duplications.

*Note:* Missing values are not considered when counting duplications.

## 3. Select and Adjust Your Columns

After your file is read, you can *choose which columns are your ID columns* (e.g., gene identifiers, protein IDs) and which columns hold the *quantitative values* (e.g., abundance measurements).

- *Select ID column*: Use the dropdown to pick exactly one column that represents the unique or primary identifier of each row.
- *Select quantitative columns*: Use the dropdown to select one or more columns that contain numeric data you'll want to analyze.
- *Zeroes to missing values*: Convert all zeroes in selected columns to “missing” values.
- *Non-numeric to missing values*: Convert any unexpected strings (e.g., “#DIV/0!”) into missing values.

## **Tutorial**

Now you need to select which columns will be used for identifying the molecules (here proteins given by protein accession numbers) and which columns provide the quantifications.

The **id** column has already been chosen correctly as indicated by a “Type” **id** above the column.

Now press the button below *Select quantitative columns*. You can use the *select all* option and then click on *ProteinId* to get all columns but the one for the identifications. Then click anywhere in the app. Now all quantitative columns should be marked as **quant**.

## **4. Data Preview**

At the bottom of the page, you'll find a table showing the top portion of your uploaded file. This helps you confirm that the data has been loaded and recognized in the correct format. Columns that have been assigned a certain role (**id** or **quant**) will be highlighted accordingly.

OmicsQ: a toolkit for quantitative Omics analysis

Reading data Experimental design Pre-processing Send and retrieve

Summary of operations ? About

**File input**

Data table

Browse... Myo.csv

Upload complete

Run example file

For feedback and bugs, please create an issue in the software repository or write the main developer: [veits@bmb.sdu.dk](mailto:veits@bmb.sdu.dk)

**File specific settings**  
(modify if necessary)

delimiter: auto

decimal separator: point

remove lines at beginning?: 0

☒ Does file have a header?

**Select and adjust**

Select id and data columns:

Select ID column: ProteinID

Select quantitative columns: C1\_1, C1\_2, C1\_3, C2\_1, C2\_2, C2\_3, C3\_1, \*

Simple manipulations and corrections:

☐ Zeros to missing values

☐ Non-numeric to missing values

Proceed to experimental design

Proceed

Show 10 entries

Columns	ProteinID	C1_1	C1_2	C1_3	C2_1	C2_2	C2_3	C3_1	C3_2	C3_3	C4_1	C4_2	C4_3	C5_1	C5_2
Type	id	quant	quant	quant	quant	quant	quant	quant	quant	quant	quant	quant	quant	quant	quant
Duplicated values	0	1836	1837	1828	1808	1795	1835	1879	1833	1834	1831	1839	1826	1841	1826
1	P60709	140000000000	139000000000	134000000000	119000000000	142000000000	130000000000	116000000000	121000000000	114000000000	124000000000	115000000000	106000000000	130000000000	121000000000
2	Q8WZ42	32400000	52600000	39300000	92400000	87200000	83300000	4570000000	5630000000	7440000000	43400000000	39000000000	35200000000	90100000000	83100000000
3	Q14315	274000000000	280000000000	271000000000	342000000000	342000000000	338000000000	506000000000	488000000000	497000000000	663000000000	645000000000	603000000000	756000000000	750000000000
4	P14618	578000000000	624000000000	607000000000	688000000000	651000000000	692000000000	679000000000	651000000000	708000000000	577000000000	593000000000	599000000000	576000000000	549000000000
5	P48681	395000000000	366000000000	387000000000	359000000000	383000000000	353000000000	380000000000	374000000000	342000000000	402000000000	411000000000	359000000000	517000000000	491000000000
6	P17661	116000000000	129000000000	126000000000	201000000000	184000000000	190000000000	307000000000	295000000000	285000000000	356000000000	371000000000	340000000000	489000000000	486000000000
7	P68032	154000000000	160000000000	221000000000	258000000000	431000000000	258000000000	188000000000	168000000000	209000000000	677000000000	218000000000	293000000000	364000000000	427000000000
8	P11055	21000000	42700000	23700000	56600000	26600000	10300000	27000000	27600000	27200000	13300000000	9530000000	10300000000	256000000000	222000000000
9	Q15149	294000000000	286000000000	280000000000	315000000000	315000000000	297000000000	368000000000	360000000000	376000000000	441000000000	447000000000	405000000000	481000000000	457000000000
10	P35579	429000000000	426000000000	434000000000	446000000000	450000000000	413000000000	488000000000	473000000000	461000000000	450000000000	443000000000	387000000000	407000000000	381000000000

Showing 1 to 10 of 3,493 entries

Previous 1 2 3 4 5 ... 350 Next

## 5. Proceed to Experimental Design

Once you have at least one *quantitative* column selected, you can click the *Proceed* button to continue. If you have not selected any numeric columns, you will see a message prompting you to do so before you can move on.

*Note:* This becomes *enabled* only after valid *id* column and *quant* columns were selected.

### Tutorial

Click *Proceed* to take you to the next step, where you define group labels, replicates, and other experimental information.

**Tip:** If anything seems off, return to the upper panel to adjust the file settings or reassign columns. Once everything looks good, use the *Proceed* button to advance to the next step.

## Step II: Experimental Design

After importing your data, you move to the *Experimental Design* tab. Here, you can automatically group and label samples, define replicate numbers, and assign batch information before proceeding to downstream analysis.

### 1. Automatic Selection of Experimental Groups

- *Threshold & string distance*: The slider lets you set a *threshold* that controls how similar or different column names must be to be grouped automatically.
- *String distance type*: the dropdown menu offers different methods for comparing column names (e.g., Levenshtein, Jaro-Winkler).
- *Automatic grouping*: As you adjust the threshold and/or change the distance type, the app clusters your sample names and assign them to different groups. In case of sufficiently well formulated sample names, this procedure should help you to specify the different conditions and replicates in your experiment.

### Tutorial

Play with the rules of the string distance to see how the sample grouping changes. Remember to place it into a similar position as before at the end.

### 2. Manually Assign Sample Types & Replicate Numbers

- *Select columns for sample type*: Adjust the slider to assign them a specific sample-type number (e.g., 1, 2, 3, ...). Use the dropdown to select which samples should be considered the same “type” or condition.
- *Replicate assignment*: Once samples share the same type, their replicate numbers are automatically updated in the table.  
If needed, you can annotate replicate numbers directly in the table at the bottom of the page (double-click on the number).

OmicsQ: a toolkit for quantitative Omics analysis

Reading data Experimental design Pre-processing Send and retrieve

Summary of operations ? About

### Automatic selection of experimental groups

Find most suitable settings. You can manually edit the experimental design in the table below. Replicates with equal number and of the same sample type will be summarized to one replicate.

Threshold for string distance: 0.130

String distance type: Jaro-Winkler

Assign Sample Types and Batch Number

Select columns for setting sample type: Nothing selected

Set to this sample type: 1

Select columns for setting batch number: C4\_1, C4\_2, C4\_3, C5\_1, C5\_2, C5\_3, C6\_1

Set to this batch number: 2

Proceed to data pre-processing

Note: Samples with equal group and replicate number will be merged.

Proceed

Download table

Show 10 entries

	Group	Replicate	Batch
C1_1	1	1	1
C1_2	1	2	1
C1_3	1	3	1
C2_1	2	1	1
C2_2	2	2	1
C2_3	2	3	1
C3_1	3	1	1
C3_2	3	2	1
C3_3	3	3	1
C4_1	4	1	2

Showing 1 to 10 of 18 entries

Previous 1 2 Next

### 3. Annotating Batches

This helps keep track of potential batch effects in subsequent analyses.

- *Specify batch number*: Use the slider to assign a batch ID (e.g., 1, 2, 3, ...).
- *Select columns for batch assignment*: Pick any samples for the specific batch ID that were measured or processed in the same batch.

### Tutorial

Although there are no batches in this experiment, assign a batch number of *two* to the second half of the samples (C4\_1 to C6\_3).

### 4. Experimental Design Table

At the bottom of the page, you'll see a table summarizing your current group assignments, replicate numbers, and batch info:



- *Editing cells*: Double-click a cell to change its value (e.g., alter group labels or replicate numbers).
- *Color-coding*: The table gets color-coded based on assigned groups.
- *Download table*: You can also download the updated experimental design table as a CSV file for your records or offline editing.

## 5. Proceed to Data Pre-Processing

Once you've finished assigning groups, replicates, and batch numbers, click *Proceed*, and you will be taken to the next step in OmicsQ, where you can review and further process your data.

### Tutorial

Press the button to *proceed* to the next tab.

### Tips & Reminders

- Adjust the *string distance threshold* and the *string distance type* carefully to get meaningful groups.
- You can always *manually override* group or replicate information in the table below.
- Remember to *assign batch numbers* if your samples were run in separate experimental batches. You can do this for comparisons, even when you initially do not intend to correct for batch effects.

## Step III: Data Pre-Processing

You can apply various transformations and filters to your data, remove or add sample columns to achieve balanced designs, and explore immediate visual feedback through *Principal Component Analysis* (PCA) and *correlation* plots. This allows you to try different settings and see their effects in real time.

## 1. Add or Delete Data Columns

- *Pick the samples to remove:* Use the dropdown to select sample columns you no longer want in the analysis.
- *Fill with empty columns:* If your experiment is unbalanced (e.g., one group has fewer replicates than another), you can add columns with only missing values to make all groups have the same number of replicates. This helps maintain a more balanced design by introducing blank values.
- *Current state:* A short summary is shown, indicating whether your design is balanced or unbalanced.

## 2. Data Manipulation & Adjustment

- *Log-transformation:* Check if data is already log-transformed. If not, OmicsQ will log2-transform the data. OmicsQ initially estimates whether the data was transformed. Always check to confirm whether the estimation was correct.
- *Max NA per feature:* Allows you to set a maximum threshold for missing values across samples. Any feature (row) exceeding that threshold is removed. This removes features with only spurious measurements.
- *Normalization:*
  - Choose from *none*, *median*, *mean*, or *cyclic LOESS* normalization.
  - Compare how each method affects your data by checking the plots at the bottom of the page.
- *Summarize to ID features:* If your *id* column (e.g., gene/protein IDs) has duplicates, you can aggregate them by **sum**, **mean**, **median**, etc. This reduces multiple rows with the same ID to a single representative row.
- *Batch Correction:* If you assigned batches in the previous step, you can select a correction method, such as **limma** or **Combat**, to adjust for batch effect. When there is only one batch or fewer than two replicates per batch, this option is disabled.

## 3. Summary & Proceed

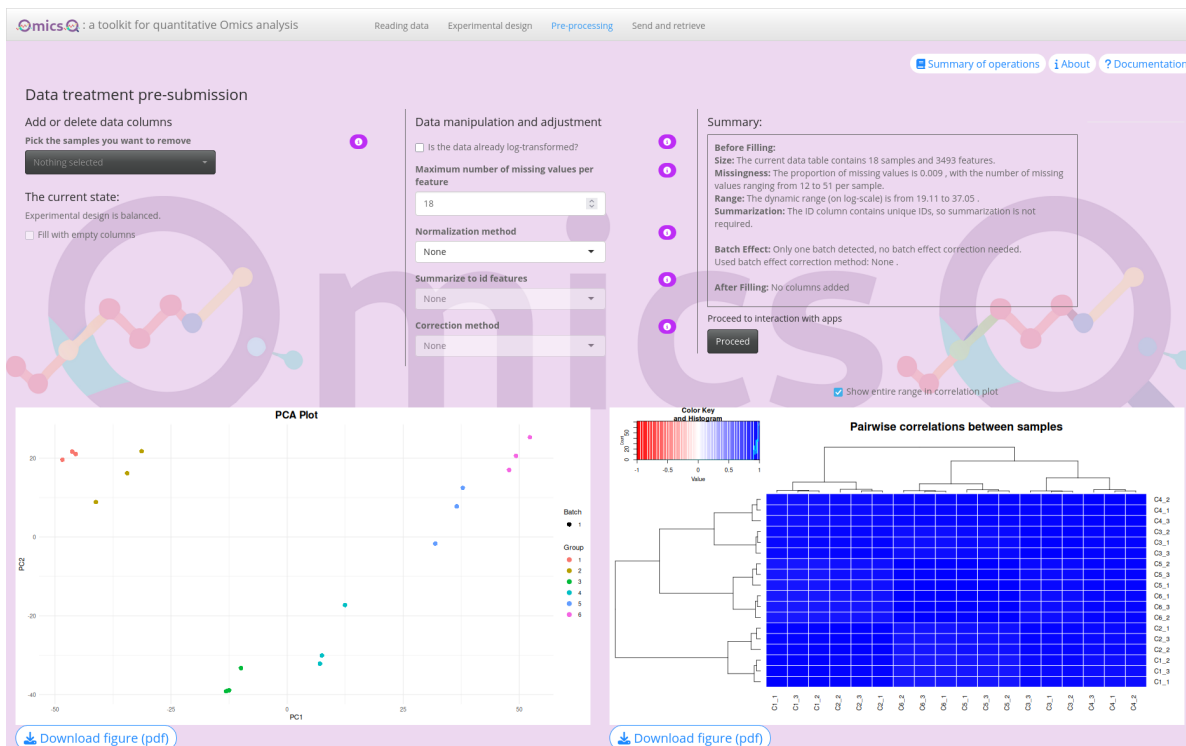
- *Summary box:* Displays key information, including the final number of features and samples, the missing data rate, and whether any features show significant batch effects.

- *Proceed to interaction with apps*: Once you are satisfied with the data adjustments, click *Proceed* to move on to the next part of OmicsQ, which may include more specialized data exploration or analysis apps.

#### 4. Visual Feedback

Beneath the summary panels, you find two data visualizations that let you see the effect of your chosen data transformations:

- *PCA Plot*: Shows the separation of samples by the first two principal components.
- *Correlation Plot*: Displays a heatmap of pairwise sample correlations, highlighting how similar or different they are after processing. This helps identifying potential outliers or technical artifacts that might need addressing. The button *Show entire range in correlation plot* allows you to adjust the coloring range of correlation values.



#### Tutorial

Try different normalization methods and observe the changes in the two plots.

Check whether OmicsQ suggests batch effects. Also try different batch correction methods.

Despite several features apparently being affected by batch effects, their number is rather low and we thus *will omit batch correction*. Actually, these changes are mostly due to sample differences (late differentiation state versus early), and thus show that it can be misleading when using batches that contain sample types that are different for the different batches.

The correlations between the samples are very high. Therefore, click on the button *Show entire range in correlation plot* to get a more distinguishable picture.

Click **Proceed** to continue to the next module of OmicsQ.

---

### 💡 Tips

- Experiment with different *normalization* and *batch correction* methods, then observe changes in the *PCA* and *correlation* plots.
  - Use *max NA* filtering to remove features with extensive missing values, which can simplify downstream analyses.
  - If your design is unbalanced, explore removing columns from the largest groups or adding NA columns to the smallest groups to maintain parity.
- 

## Step IV: Further Integrated Analysis

After you have completed data pre-processing, OmicsQ allows you to send your processed dataset to additional apps for deeper analysis. These tools can be used for statistical testing, clustering, and exploring protein complexes. Below is an overview of each app and how to access them within OmicsQ.

### 1. Sending Data to External Apps

Here, you can send the data to specific web applications. Parameters such as experimental design are automatically configured.

- **PolySTest:**
  - Designed for statistical testing of multi-omics or proteomics data with missing values.

- Once clicked, your processed data is uploaded, and a new window or tab opens where you can run various statistical tests.
- Use *Retrieve results* to import any outcomes back into OmicsQ.
- **VSClust:**
  - A clustering tool that groups features (e.g., proteins, peptides, or genes) based on abundance profiles and their variances, and does not require value imputation.
  - Similar to PolySTest, you click *Send* to push data to VSClust and *Retrieve results* to pull them back.
  - This App identifies co-expression or co-regulation patterns in your data.
- **ComplexBrowser:**
  - Focuses on exploring protein complexes and how their subunits behave under different experimental conditions.
  - Clicking *Send to ComplexBrowser* uploads the data, and you can investigate complex membership, stoichiometry changes, or differential expression of subunits.

## 2. Paired vs. Unpaired Design

- *Paired toggle*: If your design is *paired* (e.g., matched samples before/after treatment), switch this on to ensure apps handle the data appropriately.
- *URL fields*: By default, the apps point to our public server `computproteomics.bmb.sdu.dk`. You can change these URLs to local or private servers if necessary. All apps can be run in isolated Docker containers.

## 3. Working with Results

- *Connection status*: When an app finishes processing, click *Retrieve results* to pull them back into OmicsQ and see whether this worked.
- *Result columns*: Results from each app are usually added to the current table with columns labeled `PolySTest_...`, or `VSClust_...`. You can download the updated table at any point by clicking *Download table*.

OmicsQ: a toolkit for quantitative Omics analysis

Analyze the table with the different apps

Statistical testing

Clustering

Investigate protein complex behavior

Download table

ProteinID	C1_1	C1_2	C1_3	C2_1	C2_2	C2_3	C3_1	C3_2	C3_3	C4_1	C4_2	C4_3	C5_1	C5_2	C5_3	C6_1	C6_2	C6_3
P0709	5.40900841728025	5.40908254714455	9.3537897120946	0.18427502703409	9.36105400404364	0.354728306378984	9.10809112373836	9.15328882097437	9.09748797378754	9.21945324074637	9.11420737394166	9.0505051335877	9.35441201884479	9.322882468884106	9.18474420221276	9.30071963170412	9.39040388122492	9.43300820913797
Q9W242	1.35128978851924	1.93535144599424	1.58918864473058	2.482278769097286	2.50977867880535	2.26127252120264	3.50023016818817	7.81840405635959	3.79348115174048	9.88511337508706	3.48381893382961	5.27164389548007	5.213887571899397	5.07378372424992	5.566418577695037	5.41818118894668	3.807402226108071	
Q14315	7.38157426535855	7.61546815947189	7.57435103464448	7.88092663046251	7.799620370794359	7.871418297459162	8.31691173209938	8.25524274342515	8.5003407395326	7.78897678533969	7.75940095551925	7.710850185509177	8.094101260007149	8.18421397262806	8.051946231467507	8.10474455893879	8.228613194767947	8.22895219154207
P14818	8.12011348991978	8.23218112923562	8.1942459623865	8.3594820828586	8.21681791813653	8.3565553416973	8.300975643867442	8.22963648730248	8.37177945086603	8.13897932621091	8.187248222815041	8.2605239525874	8.201634813150411	8.205404801974067	8.228072474219532	8.374148380752348	8.4078488627604	8.31316261284303
P06081	7.3063679462112	7.6857471605454	7.78238033476031	7.64511205716275	7.67276093182279	7.65097296021648	7.69285914895762	7.6463734268134	7.52828374117012	7.49527636464008	7.48033673037103	7.36439560601572	7.81345464340601	7.812121270984628	7.695778384324127	7.94072040733912	7.52968855214629	8.04955995936124
P17661	6.47783368745209	6.622142464136277	6.59160577732445	7.191153785574509	7.013483784982221	7.14421016088089	7.782380838426314	7.63837314337871	7.61809383462843	6.76945043738715	6.843878752808439	6.7652452811719	7.353191752872677	7.424496177402011	7.25742120429874	7.56132440846156	7.8762648818918	7.8948825959217
P06032	6.8988325705087	6.76220114711895	7.23768138336326	7.45886171334587	7.76047158494719	7.48635283987138	6.9483868872357	6.79233899185465	7.11893972288203	6.461152774630312	6.275641772960788	6.741054418556113	7.04310861028389	7.33548388391527	7.58847778730202	7.6748551388524	7.696788077543493	7.451746714823348
P11055	-1.01240818772863	0.154815938818358	-0.800719193201848	-0.15378079338288	-0.742889345173302	-2.14307132708405	-0.44919081138936	-0.829788051762417	-0.622801777544056	-2.218330231982187	-2.812768019719567	-2.473173479324826	1.3817710574801078	1.271393839382412	1.241996383772319	2.2316340086281	2.188587808338675	2.277215802820236
Q15149	7.440681759873821	7.41450394074521	7.384957388718982	7.34132762885044	7.474017613267869	7.48762280833144	7.71813757853225	7.67233739330379	7.359411525830634	7.44970978837828	7.476117180718752	7.389321393559601	7.638191886227754	7.637227940146505	7.61489909379638	7.709123725347083	7.750642087733016	7.78637634548364
P25579	7.89584523845052	7.89583923132492	7.718848032760344	7.752181978162199	7.692546633102362	7.666741604309134	7.84338389909747	7.78185427848445	7.77231249194382	7.768841237823949	7.75408446228726	7.622988032131268	7.69142866494984	7.670570813059463	7.638674792699412	7.783754861230075	7.754713760370701	7.847320973916074

Showing 1 to 10 of 3,493 entries

Previous 1 2 3 4 5 ... 350 Next

## 4. Details about the External Apps

### PolySTest

- A statistical toolkit for integrated, multi-condition testing in proteomics/omics experiments.
- Contains *Miss Test* for inclusion of missing data in the testing without any assumptions about the missingness.
- Useful for identifying differentially expressed features across multiple groups or time-points.

**Tutorial:** [Link](#) to R package vignette (the tutorial contains the same steps as carried out in the web application).

**Source Code**

**Publication:** [PubMed link]<https://pubmed.ncbi.nlm.nih.gov/32424025/>)

**R Package:**

A tool to determine and visualize differentially regulated features using multiple approaches. For more information, press the question mark button on the upper right.

Data input

Data layout

First column for quantification

2

Replicates are grouped

Number of replicates

3

Number of conditions

6

Statistical testing

Run analysis

Paired tests?

Comparison 1:

C2 vs C1

Comparison 2:

C3 vs C1

Comparison 3:

C4 vs C1

Comparison 4:

C5 vs C1

Comparison 5:

C6 vs C1

Add new comparison

Help (click on the right to see the help page)

Data details

Number of features: 3493

Number of data columns in file: 18

Percentage of missing values: 0.92

Condition 1: C1\_1, C1\_2, C1\_3

Condition 2: C2\_1, C2\_2, C2\_3

Condition 3: C3\_1, C3\_2, C3\_3

Condition 4: C4\_1, C4\_2, C4\_3

Condition 5: C5\_1, C5\_2, C5\_3

Condition 6: C6\_1, C6\_2, C6\_3

Thresholds

To select feature groups and adapt visualization groups below.

FDR threshold

0.01

lower log-ratio threshold

0

upper log-ratio threshold

0

Select all features filtered by FDR and fold-change

Select tests and comparison

Further criteria for feature selection.

Which statistical test(s)?

☒ PolySTest

☐ limma

☐ Miss Test

☐ rank products

☐ permutation test

☐ t-test

Which comparison(s)?

☒ C2\_vs\_C1

☐ C3\_vs\_C1

☐ C4\_vs\_C1

☐ C5\_vs\_C1

☐ C6\_vs\_C1

Select features

Number selected features: 179

Select all shown features

Select all (filtered) features

Clear selected features in table

Download results

Column visibility

Copy

Print

Show 5 entries

Search:

log\_ratios

PolySTest

Feature	C2_vs_C1	C3_vs_C1	C4_vs_C1	C5_vs_C1	C6_vs_C1	C2_vs_C1	C3_vs_C1	C4_vs_C1	C5_vs_C1	C6_vs_C1	C2_vs_C1	C3_vs_C1	C4_vs_C1	C5_vs_C1	C6_vs_C1
P60709	7.3e+9	2.1e+10	2.3e+10	1.7e+10	1.2e+10	1.0	0.0000017	0.0000036	0.000046	0.0013	0.30	0.0063			
Q8W242	-4.6e+7	-5.8e+9	-3.9e+10	-8.4e+10	-1.0e+11	0.88	0.00018	0.0000034	5.5e-9	5.3e-10	0.58	0.14			
Q14315	-6.6e+9	-2.2e+10	-3.6e+10	-4.6e+10	-5.0e+10	0.000099	9.3e-8	3.6e-10	1.5e-11	3.5e-12	0.0046	2.3e-8			
P14618	-7.4e+9	-7.6e+9	1.3e+9	3.9e+9	-6.3e+8	0.00020	0.00018	0.18	0.017	0.35	0.0036	0.00046			
P48681	1.8e+9	1.7e+9	-8.0e+8	-1.0e+10	-1.8e+10	0.48	0.44	0.89	0.00030	0.0000093	0.39	0.22			

Showing 1 to 5 of 3,493 entries

Previous 1 2 3 4 5 ... 699 Next

Details on tests and feature (e.g. proteins) expression changes (max. 30 with lowest unified FDRs shown)

Scale features to mean

Feature expression over conditions

## VSClust

- Clusters features with similar expression/abundance profiles considering their individual variances.
- Can be applied to data with missing values without imputation.
- Helps reveal co-regulated sets of molecules (e.g., proteins, genes) and highlight potential biological modules.

**Tutorial:** [PubMed link](#)

**Publication:** [PubMed link](#)

**Source Code**

**R Package:**



## ComplexBrowser



- Allows inspection of known protein complexes, exploring how their subunits behave across conditions or treatments.
- Facilitates a better understanding of complex stoichiometry or stability changes under different experimental scenarios.

**Tutorial:** [Online tutorial](#)

**Publication:** [PubMed link](#)

**Source Code**



If you encounter connectivity issues, ensure your *browser* and *network settings* allow opening new windows/tabs and sending data to these external URLs. Then try again.

## **Tutorial**

### **PolySTest for differential regulation**

Press the button *Send to PolySTest*. This will open a new tab with PolySTest (you might need to allow this in your browser settings and press the button again). The data should have been loaded into PolySTest which you can see in correctly set sliders for parameters like “Number of replicates” and the window “Data details”. Repeat sending the data from OmicsQ if the data was now uploaded.

Then press the button *Run analysis* to start the statistical tests. This will take a while. Removing some of the comparisons will accelerate the analysis.

When the tests are finished, change the FDR threshold to 0.001, fold-change thresholds of -4 and 4, choose the comparison *C6\_vs\_C1*, and press *Select all features filtered by FDR and fold-change*. This should select 30 proteins that are differentially regulated between the first and the second time point.

Open the next two green fields by clicking on the “+” to see expression profiles and a hierarchical clustering of the differentially regulated proteins.

Then move back to OmicsQ and press *Retrieve results from PolySTest* to add FDR values, fold-changes and other fields to the table in OmicsQ. You can scroll the table and should then see the retrieved columns.

### **VSClust for variance-sensitive clustering**

We will now upload our differentiation data to VSClust. Press the button *Send to VSClust*. This will open another tab with VSClust. The PCA plot in VSClust should show the different samples. Again, if this does not work the first time, please try again.

First, we need to estimate the number of clusters in the data set. Go to the tab *Estimation of cluster number* and press *Estimate parameters*. You can accelerate this example analysis by decreasing the maximum number of clusters to 10.

The resulting figures help to identify suitable numbers of clusters from so-called validity indices such as *Minimum Centroid Distance* and *Xi-Beni Index*. We always compare the results for the variance-sensitive clustering, providing more accurate control of noise levels, with standard fuzzy c-means clustering. The right figure shows the number of features that are assigned to clusters. Features lying between clusters are discarded.

You should now observe that both indices indicate 3 clusters as most suitable selection (black squares). However, we want to see a more fine-grained picture and therefore choose a higher cluster number with a considerable jump of the Minimum Centroid Distance such as 8.

Now, change to the tab *Clustering results / variance-based*, select 8, and run the clustering. You will see the results from the 8 clusters, each of them corresponding to a different temporal pattern.

Go back to OmicsQ and press the button *Retrieve results from VSClust*. This should add the columns with cluster number, presence in cluster, and others, to the final table.

We now want to take the proteins that were increasing during the cell differentiation process. This should be the second or the third cluster showing a continuous increase. Select cluster number 2 in the column `VSClust_clusters` and select `true` in `VSClust_isClusterMember`. You can additionally filter for proteins that are differentially regulated between two (or more) conditions. Then press *Send (filtered) features to StringDB*. This will open a new tab on StringDB. There, select “Homo sapiens” and press *Continue* twice.

You will end with a quite large protein-protein interaction network. Below, you find the enriched pathways in the *Analyze* part. The most dominant terms should be RNA processing and metabolism. You can repeat this procedure for different clusters and different comparisons from the statistical tests.

## ComplexBrowser for Protein Complex Behavior

*Note:* This part is meant for proteomics data only.

Press the button *Send to ComplexBrowser* to open the application. Again, the data and all parameters should automatically be set. You should see several visualizations which can be used to further investigate data properties. We will dismiss this part and move on to the analysis of protein complexes.

Click on the tab *Complex analysis* in the upper left. Select *Complex Portal* in *select database for analysis* and click on *Run analysis* to start calculating protein complex quantifications.

You get a large table with details about the different protein complexes in the data set and their quantitative behavior. Click twice on the column name `NQS`, standing for the number of quantified complex subunits found in the data sets. This will sort the table by the number of subunits, and provide parts of the Spliceosome on the top.

When selecting a complex, you will get details about its behavior in the different samples. Type “MCM” into the search field of `ComplexName` and select the *MCM complex*. You will see that a) the complex quite drastically decreases in its abundance during the differentiation process and b) that its subunits are equally affected exhibiting very similar quantitative changes.

We will now check whether the proteins of this complex also co-regulate that strongly in human cells in general. Therefore, press the button *Submit proteins to CoExpresso* in the lower left

panel *Protein complex visualization*. This will open yet another application, CoExpresso, which will show you the co-expression of the proteins in the MCM complex in different tissues and cell types. You will see that the complex subunits of the MCM complex are usually highly co-regulated in human cells, shown through high correlations and significant p-values. This indicates that the MCM complex is a stable complex in human cells.

You can use ComplexBrowser to investigate other complexes in the same way. We recommend prioritizing the complexes that are most affected by the differentiation process, and thus show highest fold-changes and lowest FDR values for C6/C1.

---

### 💡 Tips

- *Privacy & Security*: If you work with sensitive data, consider hosting these apps on a **secure internal server** and update the URL fields in OmicsQ accordingly.
- *Data Size*: Uploading very large datasets might require additional time or could fail if your internet connection is unstable or if the server imposes size limits.

*Next Step*: After analyzing or clustering your data in these apps, revisit your results in OmicsQ to combine insights or proceed with downstream workflows.