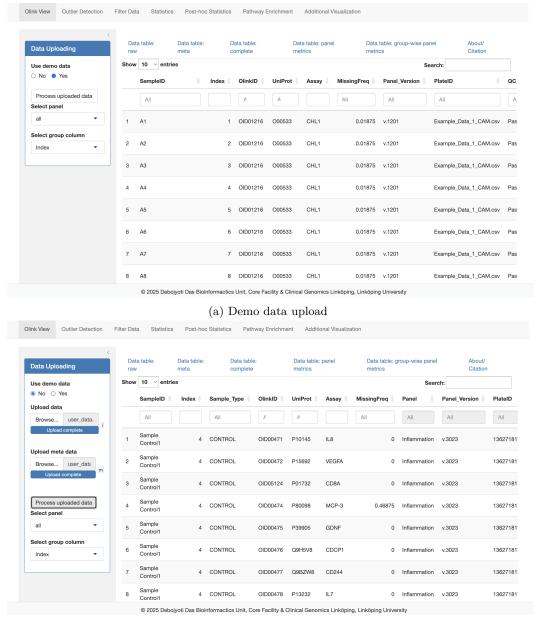
completeOlink - Manual

Here is an easy, step-wise guide to the use of completeOlink webserver for **OLINK** proteomics data analysis A. Data upload:



(b) User data upload

Figure 1: Olink View tab

The **Olink View** tab allows data upload. User can optionally use the demo data by selecting "Yes" for "Use demo data". Figure 1 depicts both scenarios.

User needs to upload the csv file (the CSV file is characterized by the presence of ";" as a delimiter instead of comma) and a sample manifest file (regular CSV file) where information about each unique "SampleID" in the npx raw data file is available. SampleID (s) are required to be character strings.

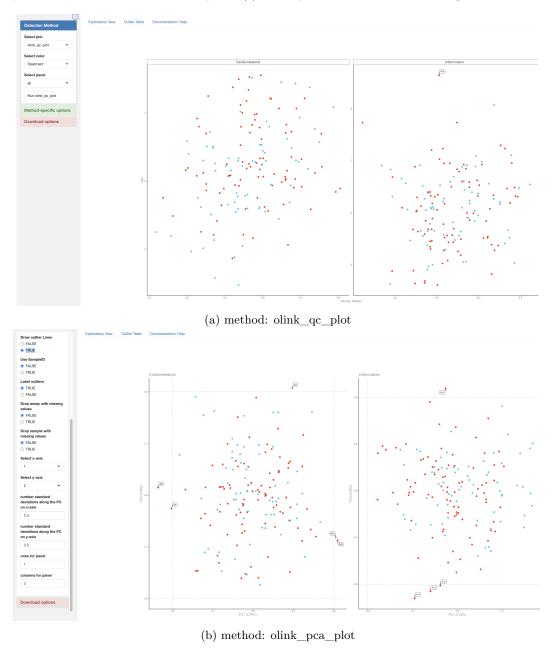


Figure 2: Outlier Detection tab (demo data; color: Treatment)

B. Outlier detection: Three outlier detection methods are available with control over parameters (optional) for accurate detection. The user can choose any of the three methods to find outliers or this step can be entirely skipped. See Figure 2. Note: If this step is used the outliers that are found in the most recent run/execution, are propagated to the next analysis step. Although, the user can choose to use the original data instead of the filtered data at every subsequent step. Results of olink_umap_plot can change in each run because of the inherent stochasticity of UMAP method.

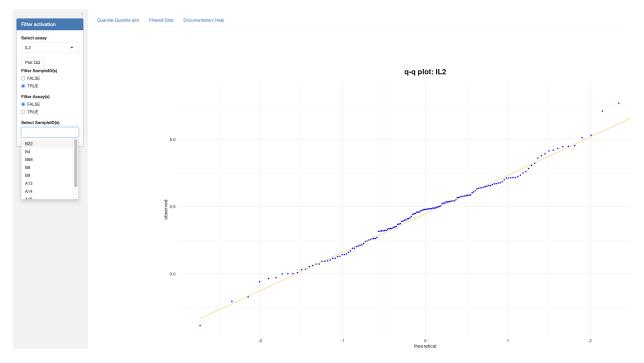


Figure 3: Filter Data tab (demo data)

- C. Filter Data: While the filtering step is optional, the user has the choice of removing samples (outliers from outlier detection step B) as well as assays (having non-parametric or non-standard distribution) before further analysis. See Figure 3. Qualitative feature of each assay is captured using quartile-quartile plot, a standard way to detect if there are departures from normal distribution (non-parametric distribution).
- D. Statistics: Statistical analysis methods offered in the application can be broadly grouped as "parametric" and "non-parametric" statistics. Parametric statistical tests can be further grouped as tests where "Outcome factor with exactly two levels" and "Outcome factor with greater than two levels". The former comprises of t-test (olink_ttest) [See Figure 4.] whereas the latter includes analysis of variance (anova; olink_anova) [See Figure 5.] and linear mixed effects model (lmer; olink_lmer).

The corresponding "non-parametric" test for "Outcome factor with exactly two levels" is the Mann-Whitney U test (olink_wilcox), while for "Outcome factor with greater than two levels" the tests are Kruskal-Wallis test/ Friedman test (olink_one_non_parametric) and two-way ordinal analysis of variance (olink_ordinalRegression). Both olink_ttest and olink_wilcox results can be used to generate volcano plots corresponding to the test outcome. In addition to statistical tests, this section allows user to generate heatmaps for a choice of grouping variable. Clustering can be invoked both on rows and/or columns, and the user can choose to use only significant assays obtained in the most recently run/executed statistical test.

- E. Post-hoc Statistics: All statistical tests in the group "Outcome factor with greater than two levels" [See Figure 6] irrespective of distribution, require post hoc analysis before meaningful results are generated. All active (chosen) functions have their corresponding usage information available in a tab labelled "Documentation/ Help".
- F. Pathway Enrichment: Following t-test or Mann-Whitney U test or the appropriate post-hoc analysis, the user has the option of performing pathway enrichment using over-representation analysis (ORA) or gene set enrichment analysis (GSEA)[See Figure 8].
- G. Additional Visualization: Lastly, significant assays can be visualized using boxplots and the p-value significance added to the plot in case t-test were previously performed. Download handlers have been implemented for easy download of all result tables. Figures generated during analysis can be customized for size and downloaded in three file formats.

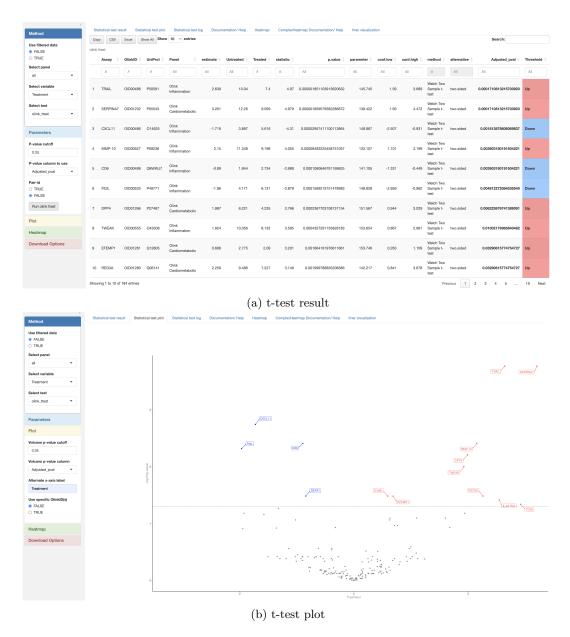
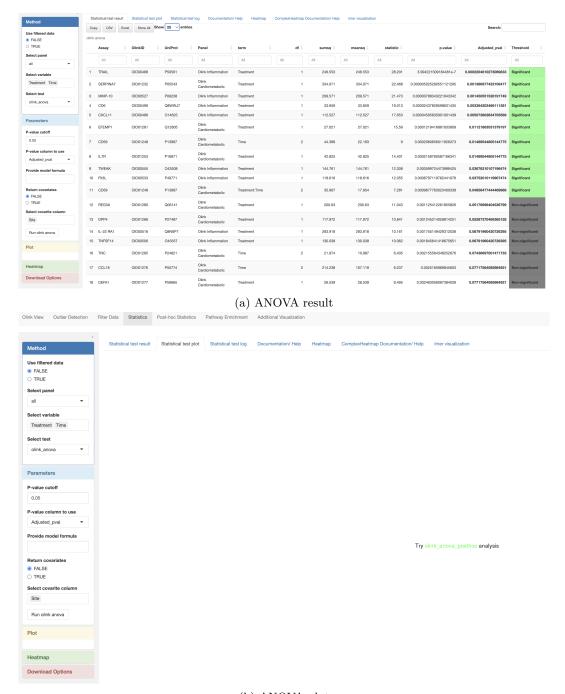
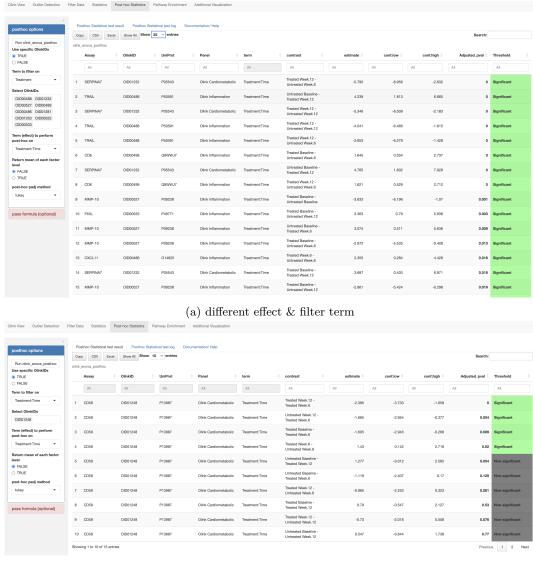


Figure 4: Statistics tab (two-level variable)



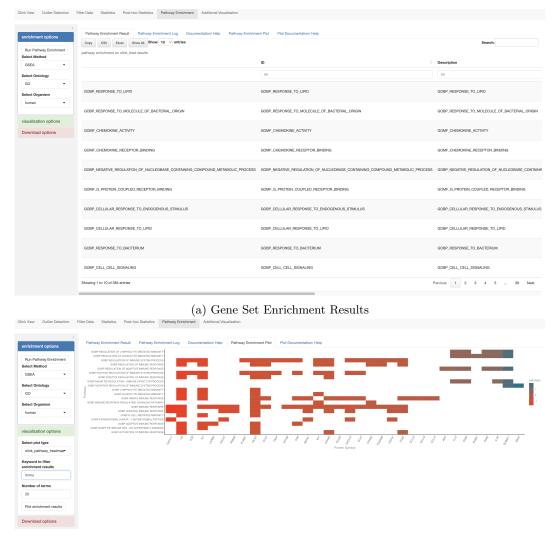
(b) ANOVA plot

Figure 5: Statistics tab (>two-level variable)



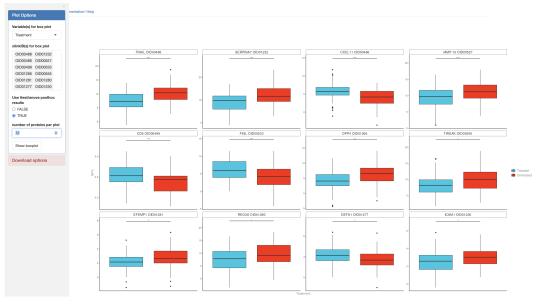
(b) same effect & filter term

Figure 6: Post-hoc Statistics tab (>two-level variable)

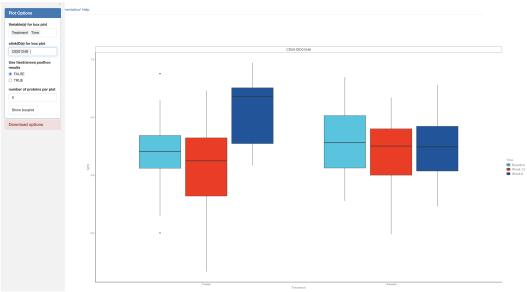


(b) Gene Set Enrichment Heatmap

Figure 7: Pathway Enrichment tab (demo data t-test results)



(a) Boxplot with t-test results



(b) Boxplot with ANOVA results

Figure 8: Additional Visualization tab (demo data)