Olink®

Olink® Differential Data Analysis

PROJECT NAME	Example - 91061
CUSTOMER	Customer Organization
	customer email address
BUSINESS DEVELOPMENT MANAGER	BDM/S Name
	BDM/S email address
DATA SCIENCE SERVICE ITEM	Olink® Differential Data Analysis
	91061
	biostat@olink.com
DATE	August 22, 2024

Contents

Product Overview

The Olink Differential Data Analysis product is used to determine differential expression of proteins between experimental groups. The input for this product include:

- Olink data exported from Olink NPX^{TM} software. NPX data is preferred. In the case of data in pg/ml, the data will be log2 transformed prior to analysis.
- A manifest mapping SampleID to the variable group(s) of interest and (in the case of longitudinal studies) mapping SampleID to Subject.

The product's deliverables include:

- a report summarizing the analysis and results
- an excel spreadsheet containing the statistical results
- separate pdfs for the figures within the report
- a boxplot for every significant assay, or for the top assays in the case of no significant assays.

Project Overview

The analysis was performed on 156 samples from project Q-XXXXX. The aim of this analysis was to determine assays that were differentially expressed across groups. A summary of the number of samples per group is shown in Table 1.

Table 1: Groups and number of samples ran in the project Q-XXXXX.

Treatment	No. Samples
A	77
B	79

Preliminary Data Analysis and QC

Interquartile Range (IQR) vs. Sample Median plot (Figures 1 and 2) and Principal Component Analysis (PCA) plot (Figures 3 and 4) along the first two principal components were generated to determine if any samples appeared to be outliers in AS project Q-XXXXX. Each set of plots show samples colored by QC status and Treatment.

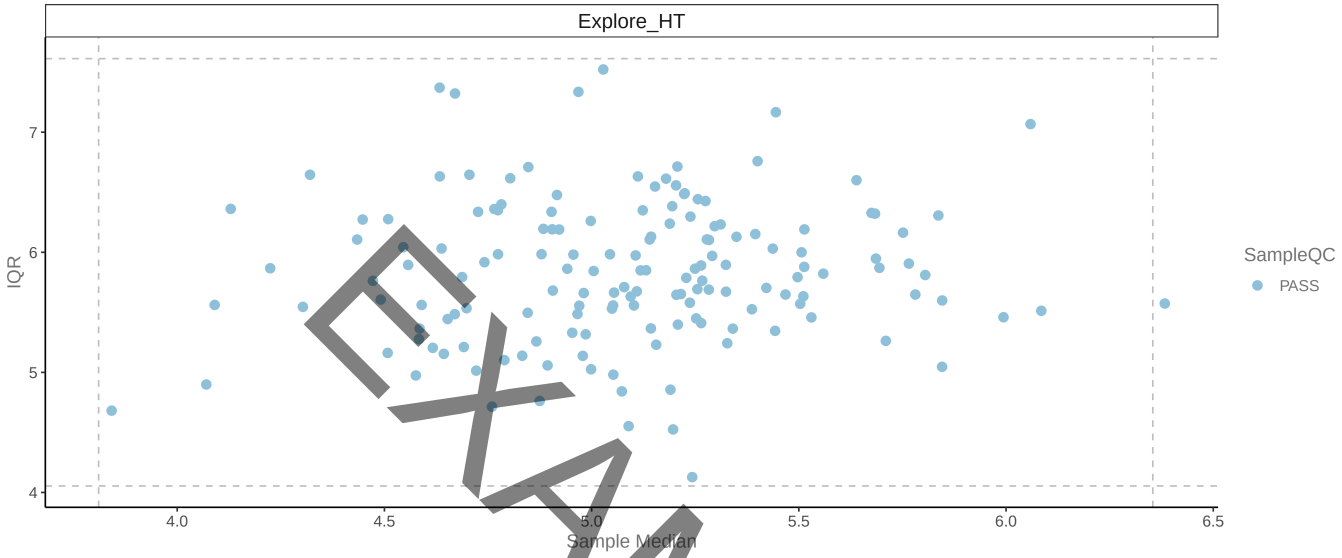

Figure 1: Interquartile range vs. sample median of project Q-XXXXX. Samples are color coded by their SampleQC status. Horizontal dashed lines indicate +/-3 standard deviations from the mean IQR. Vertical dashed lines indicate +/-3 standard deviations from the mean sample median.

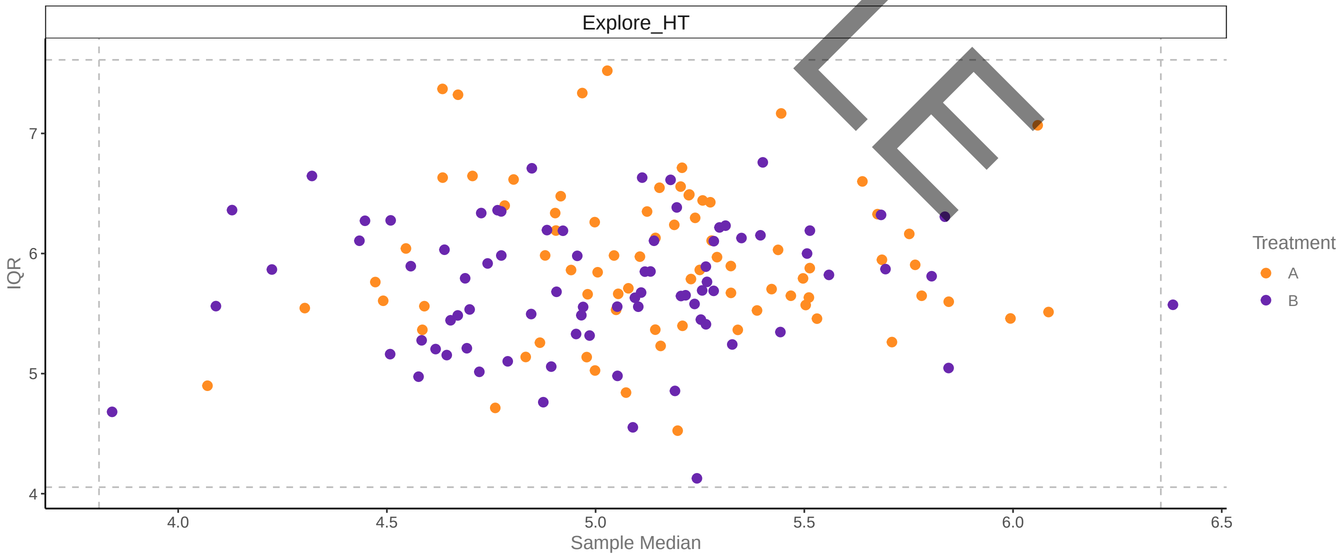

Figure 2: Interquartile range vs. sample median of project Q-XXXXX. Samples are color coded by their Treatment status. Horizontal dashed lines indicate +/-3 standard deviations from the mean IQR. Vertical dashed lines indicate +/-3 standard deviations from the mean sample median.

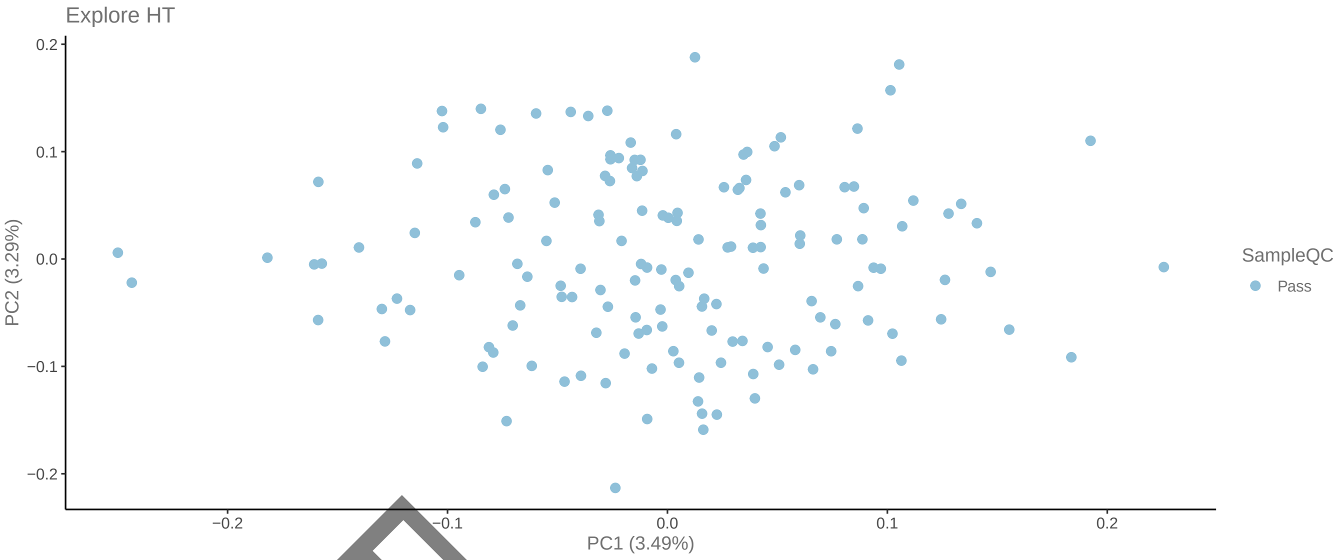

Figure 3: PCA plot along the first two principal components of all samples from project Q-XXXXX. Samples are colored by their SampleQC status

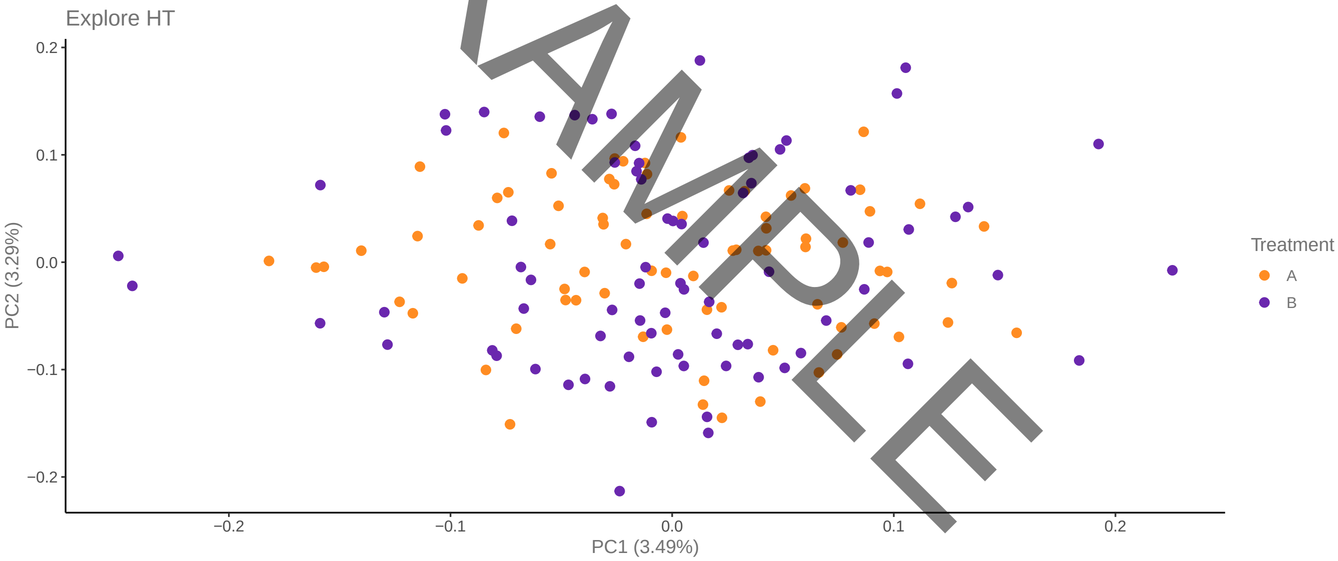

Figure 4: PCA plot along the first two principal components of all samples from project Q-XXXXX. Samples are color coded by their Treatment status.

Per this project, no samples appear to be to be outliers.

Differential Expression Analysis — T-test

The project Q-XXXXX contained samples from two groups. T-test analysis is performed to assess whether the groups have a significant difference in protein expression.

Five assays were found to be significant after multiple testing adjustment.

Differential Expression Visualizations — T-test

Volcano Plot

The volcano plot below visualizes the significant assays from the result of t-test (Figure 5). The estimate, which is the difference in means between the two groups, is plotted along the x-axis, while the nominal p-value (before multiple testing adjustment) is plotted along the y-axis. Assays are colored based on whether or not they are significant after Benjamini-Hochberg multiple testing adjustment, and the dashed line is an unadjusted p-value of 0.05.

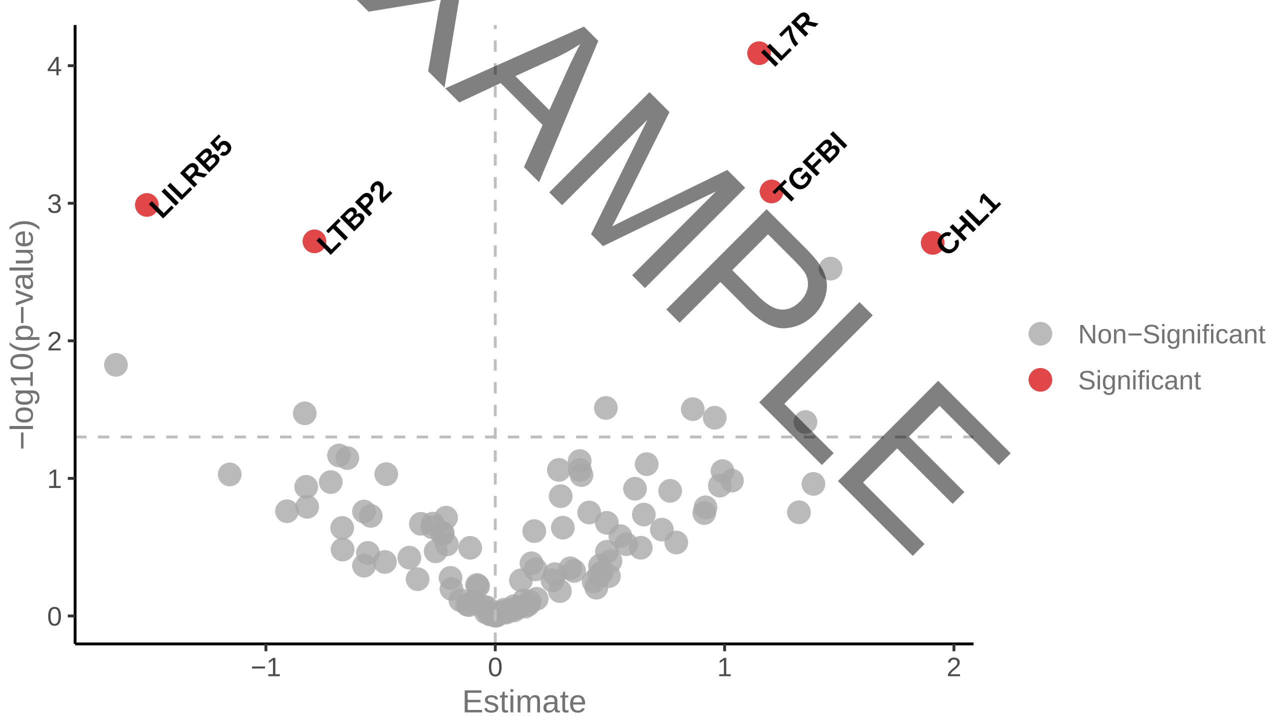

Figure 5: Volcano plot highlighting the assays which were found to be significant after multiple testing adjustment, if any.

Boxplot

Boxplots were generated for the top 12 assays, colored by group (Figure 6). Boxplots of all the statistically significant assays after multiple testing adjustment, if any, are attached to this report.

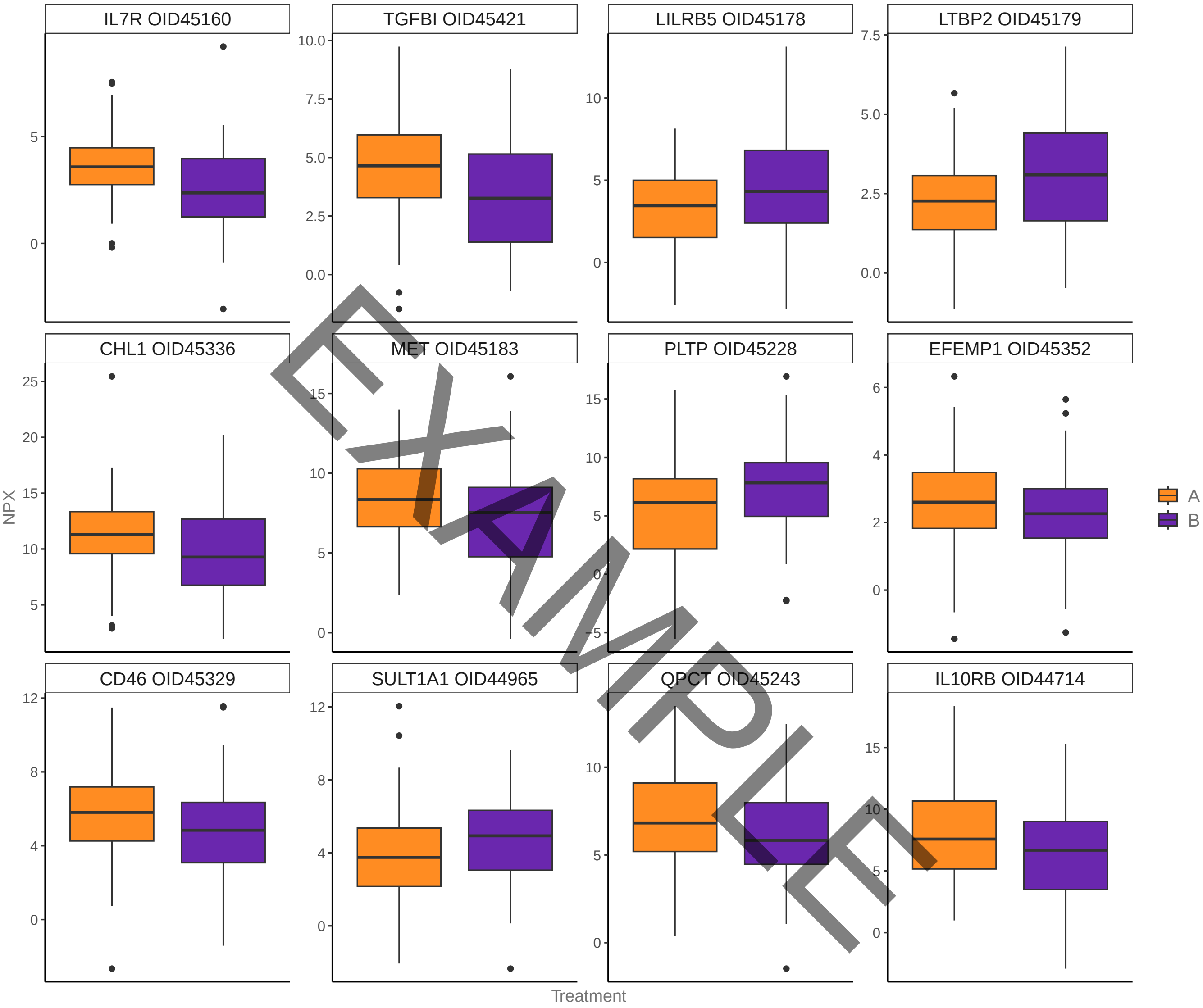

Figure 6: Boxplot of the top 12 assays from which the first 5 found to be significant after multiple testing adjustment.

Heatmap

A heatmap was generated using the top 25 assays after multiple test adjustment (Figure 7). The heatmap is annotated by group in the left-most column

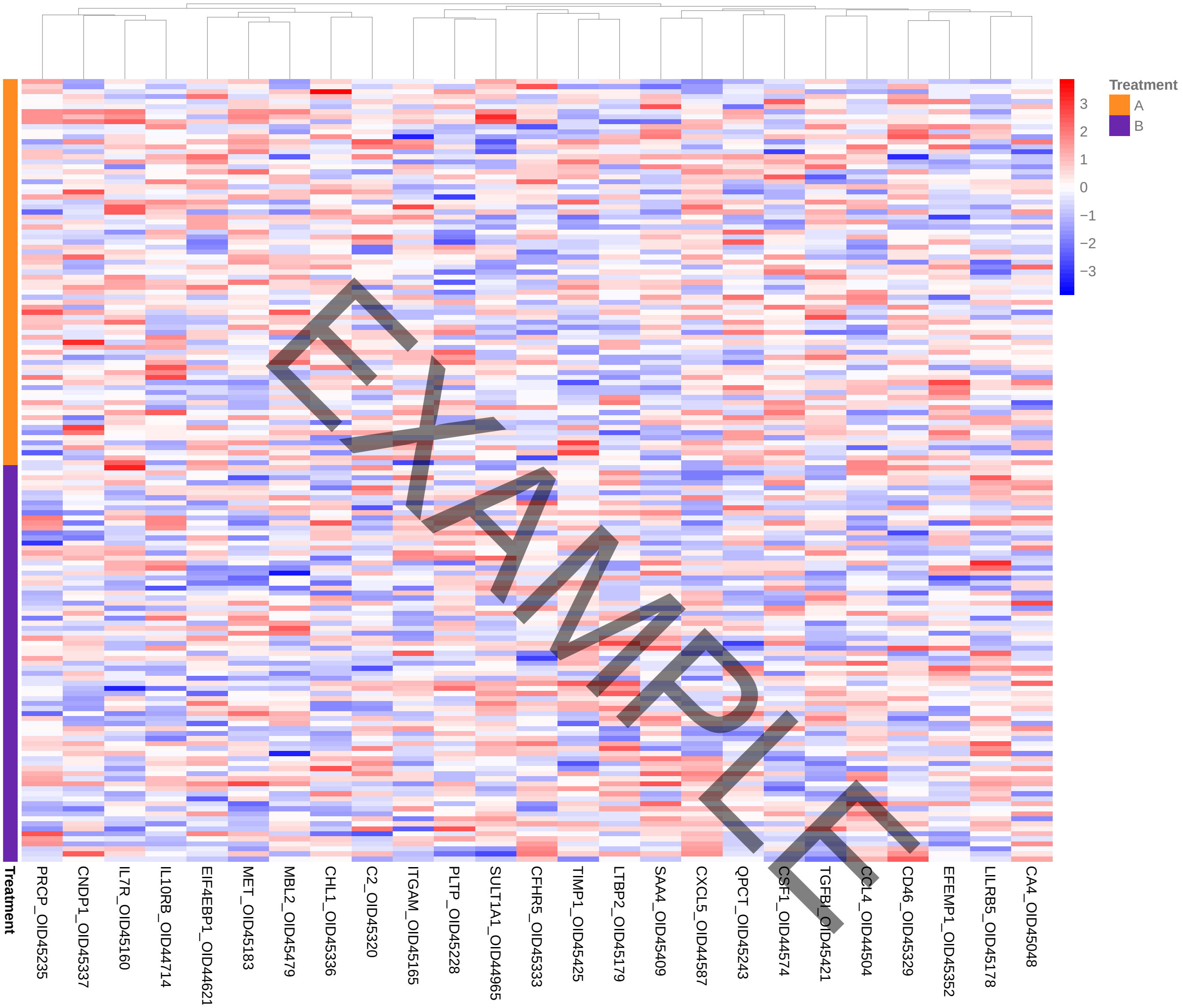

Figure 7: Heatmap using top 25 assays out of which 5 found to be significant after multiple testing adjustment.

Summary

A t-test analysis was performed on Q-XXXXX, which contained 156 samples belonging to two groups, in order to assess whether there were any significant differences in protein expression among the group.

Legal Disclaimer

R Session Information

- R version 4.4.0 (2024-04-24), `x86_64-pc-linux-gnu`
- Locale: `LC_CTYPE=en_US.UTF-8`, `LC_NUMERIC=C`, `LC_TIME=en_US.UTF-8`, `LC_COLLATE=en_US.UTF-8`, `LC_MONETARY=en_US.UTF-8`, `LC_MESSAGES=en_US.UTF-8`, `LC_PAPER=en_US.UTF-8`, `LC_NAME=C`, `LC_ADDRESS=C`, `LC_TELEPHONE=C`, `LC_MEASUREMENT=en_US.UTF-8`, `LC_IDENTIFICATION=C`
- Time zone: `America/New_York`
- TZcode source: `system (glibc)`
- Running under: `Ubuntu 22.04.4 LTS`
- Matrix products: default
- BLAS: `/usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3`
- LAPACK: `/usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblasp-r0.3.20.so`; LAPACK version3.10.0
- Base packages: base, datasets, graphics, grDevices, methods, stats, utils
- Other packages: dplyr 1.1.4, english 1.2-6, extrafont 0.19, forcats 1.0.0, ggplot2 3.5.1, kableExtra 1.4.0, knitr 1.46, lubridate 1.9.3, OlinkAnalyze 3.7.0, purrr 1.0.2, readr 2.1.5, stringr 1.5.1, tibble 3.2.1, tidyr 1.3.1, tidyverse 2.0.0
- Loaded via a namespace (and not attached): abind 1.4-5, arrow 15.0.1, assertthat 0.2.1, backports 1.4.1, bit 4.0.5, bit64 4.0.5, boot 1.3-30, broom 1.0.5, cachem 1.0.8, car 3.1-2, carData 3.0-5, cellranger 1.1.0, cli 3.6.2, codetools 0.2-20, colorspace 2.1-0, compiler 4.4.0, digest 0.6.35, emmeans 1.10.1, estimability 1.5, evaluate 0.23, extrafontdb 1.0, fansi 1.0.6, farver 2.1.1, fastmap 1.1.1, fs 1.6.4, generics 0.1.3, ggplotify 0.1.2, ggrepel 0.9.5, glue 1.7.0, grid 4.4.0, gridGraphics 0.5-1, gtable 0.3.5, hms 1.1.3, htmltools 0.5.8.1, labeling 0.4.3, lattice 0.22-6, lifecycle 1.0.4, lme4 1.1-35.3, lmerTest 3.1-3, magrittr 2.0.3, MASS 7.3-60.2, Matrix 1.7-0, memoise 2.0.1, minqa 1.2.6, multcomp 1.4-26, munsell 0.5.1, mvtnorm 1.2-4, nlme 3.1-164, nloptr 2.0.3, numDeriv 2016.8-1.1, pheatmap 1.0.12, pillar 1.9.0, pkgconfig 2.0.3, R6 2.5.1, RColorBrewer 1.1-3, Rcpp 1.0.12, readxl 1.4.3, rlang 1.1.3, rmarkdown 2.27, rstatix 0.7.2, rstudioapi 0.16.0, Rttf2pt1 1.3.12, sandwich 3.1-0, scales 1.3.0, splines 4.4.0, stringi 1.8.3, survival 3.5-8, svglite 2.1.3, systemfonts 1.0.6, TH.data 1.1-2, tidyselect 1.2.1, timechange 0.3.0, tinytex 0.50, tools 4.4.0, tzdb 0.4.0, utf8 1.2.4, vctrs 0.6.5, viridisLite 0.4.2, withr 3.0.0, xfun 0.43, xml2 1.3.6, xtable 1.8-4, yaml 2.3.8, yulab.utils 0.1.4, zip 2.3.1, zoo 1.8-12