MS-DAP: Mass Spectrometry Downstream Analysis Pipeline

version: beta 0.2.7.1 https://github.com/ftwkoopmans/msdap/

Contents

1 Quality control

The quality control figures in this section enable you to investigate reproducibility and global clustering of samples by visualizing:

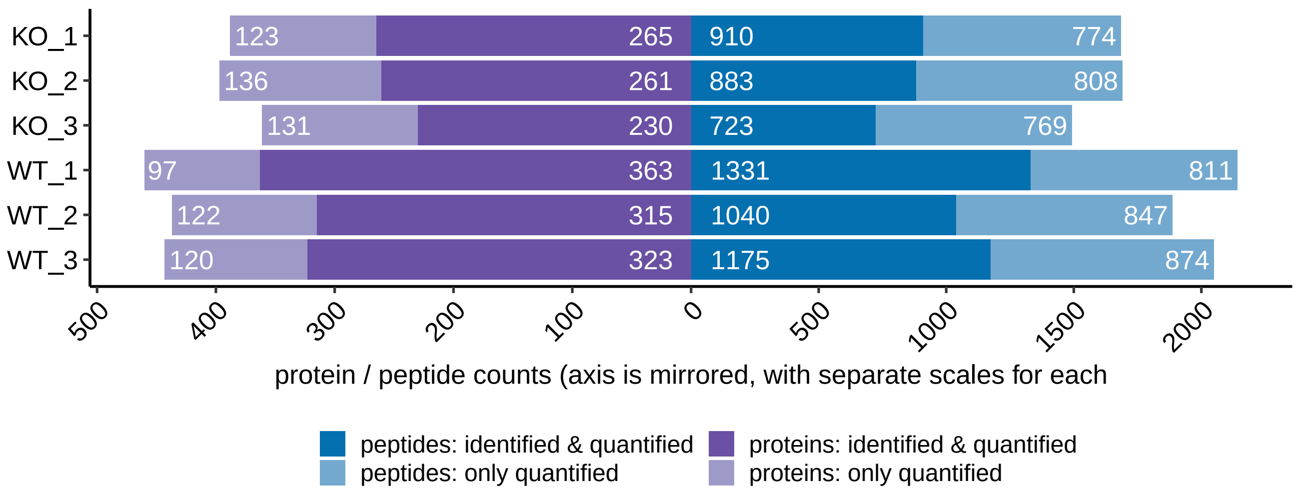
- number of peptides/proteins detected in each sample
- dataset completeness
- local effects in HPLC peptide retention time per sample
- reproducibility of peptide quantification among replicates
- PCA of all samples to visualize clustering

The first set of quality control figures describes individual samples, thereafter group-level quality metrics are described and finally sample clustering is used to highlight structure in the entire dataset.

1.1 number of peptides and proteins

These plots show the number of (target) peptides that are 'detected' per sample. For DDA, 'detected' implies the peptide has a MS/MS identification. Peptides quantified through match-between-runs (MBR) are quantified but not detected/identified. In case of DDA, we also show the number of peptides quantified through MBR. For DIA, we refer to a peptide as 'detected' if the confidence score (for identification) is <= 0.01.

Samples in this plot are sorted by their experimental group, and then ordered and by their name within each group. This data is also available in the output table 'samples.xlsx'.

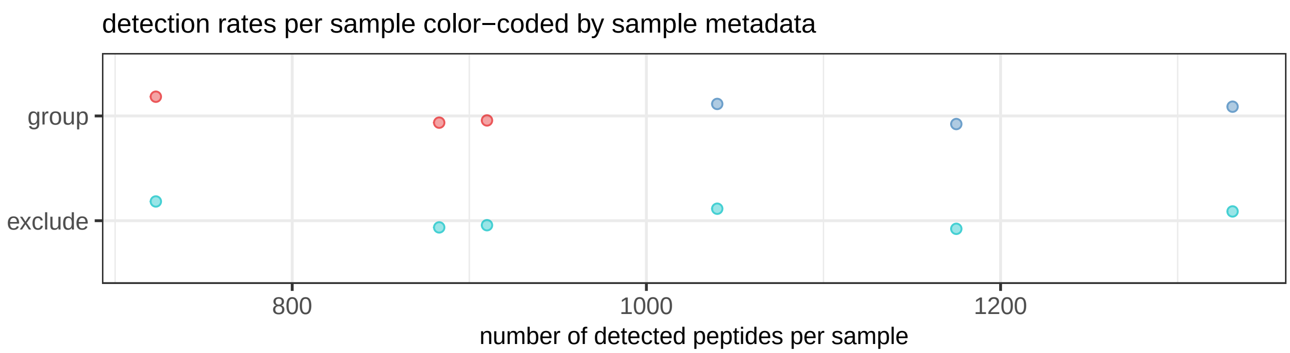


1.1.1 color-coding sample metadata

The number of detected peptides in a sample, as compared to other samples within a dataset, can be used as a measure for sample quality. Color-coding individual samples for metadata that you provided as input (e.g. experiment batch, sample handling order, gel lanes, etc.) allows visual inspection as to whether these relate to the rate of successful peptide detection.

The figure below aims to provide an overview of all sample metadata at a first glance. The y-axis shows all sample metadata. On each row all samples in the dataset are shown as a dot, each color-coded by the respective property shown on the y-axis (with minor vertical jitter for visual clarity). If there is a major effect of any experimental condition on the number of detected peptides, this is easily spotted as all outlier samples (extremes in number peptide detection, x-axis) will have the same color. The following figures expand this overview into a detailed figure for each sample property, those visualization are designed to further dive into each experimental condition and provide respective color legends.

Note that the visualization of sample metadata in this report depends on user-provided input; each column in the metadata input table (besides sample names) that contains more than 1 unique value is automatically used as a factor for color-coding all figures in this section. So the data used for color-coding is also available in the output table 'samples.xlsx'.



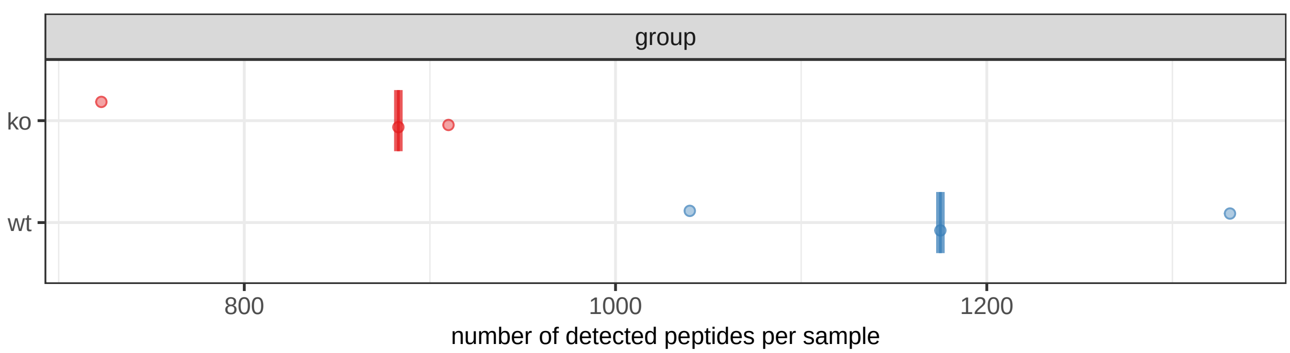
color-coding sample metadata, expanded

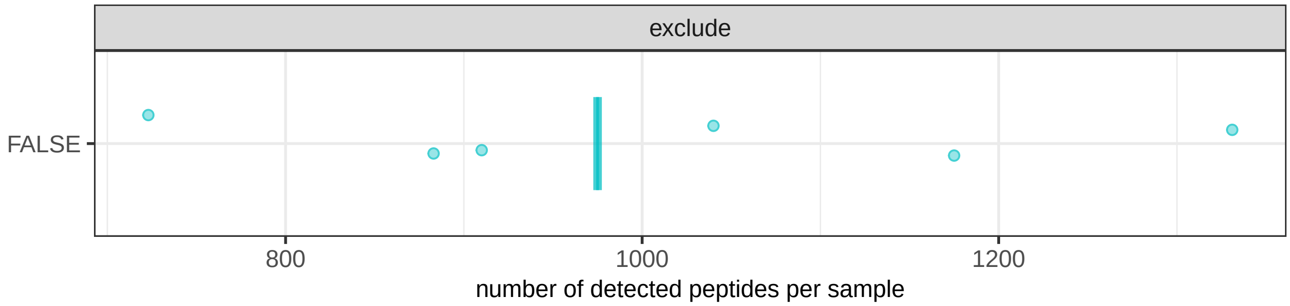
To further detail each sample property, each row in the above figure is now split into separate plots. Thus, a figure is generated for each experimental condition from the user-provided metadata (column in the samples table, it's name shown in the plot title).

The y-axis shows all unique variables, the x-axis the number of detected peptides and each dot is a sample. The vertical line indicates the median value for each row. Colors are consistent with the above plot.

For example: the first plot shows color-coding by the 'group' property, so each row represents a sample group. If samples in a particular group systematically yield fewer peptides than another group, a clear pattern will be visible.

Note that the *exclude* metadata is a user-provided indication of 'outlier samples' in the dataset that are excluded from downstream statistical analysis.

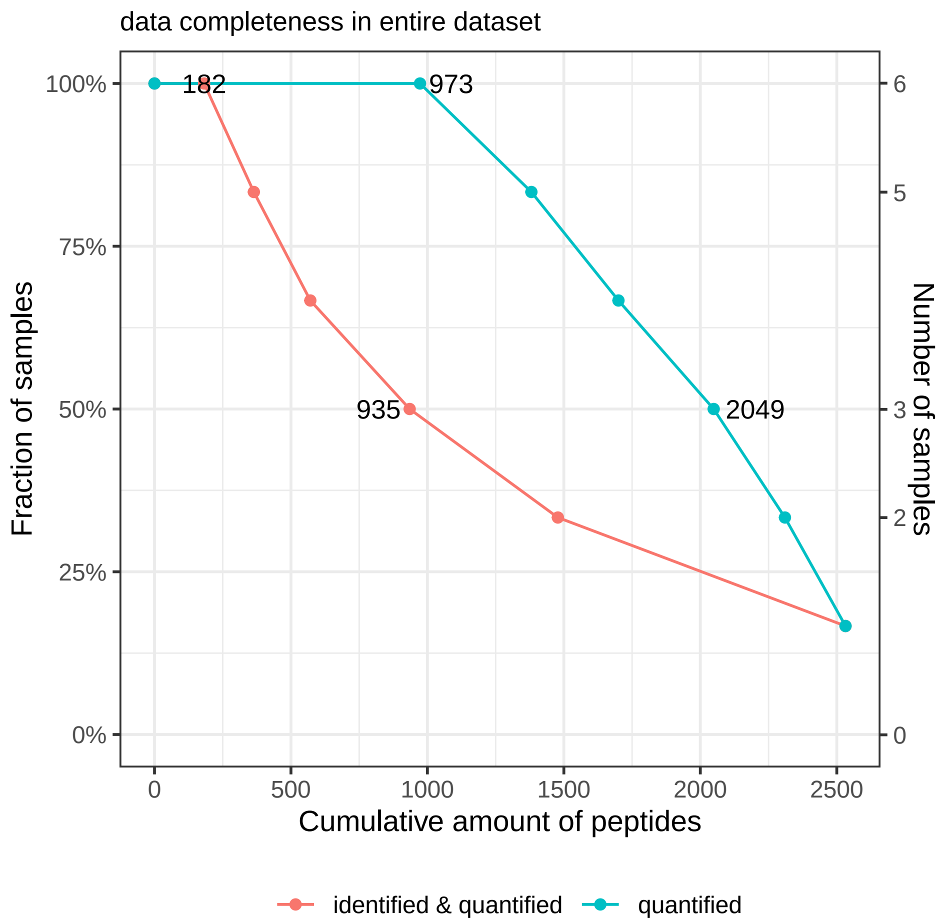



exclude
FALSE
800
1000
1200
number of detected peptides per sample

1.2 data completeness

To visualize how many peptides are consistently identified in multiple samples, the first figure summarizes how common missing values are in the entire dataset. Optimally, most peptides are identified in 100% of samples and this curve slowly falls of. The following figure shows for each sample whether its peptides are also present in other samples in the dataset or whether these are unique to a (minor) subset of samples. You can use this mark of experimental consistency to compare datasets generated by similar protocols and mass-spec acquisition.

1.2.1 cumulative distribution

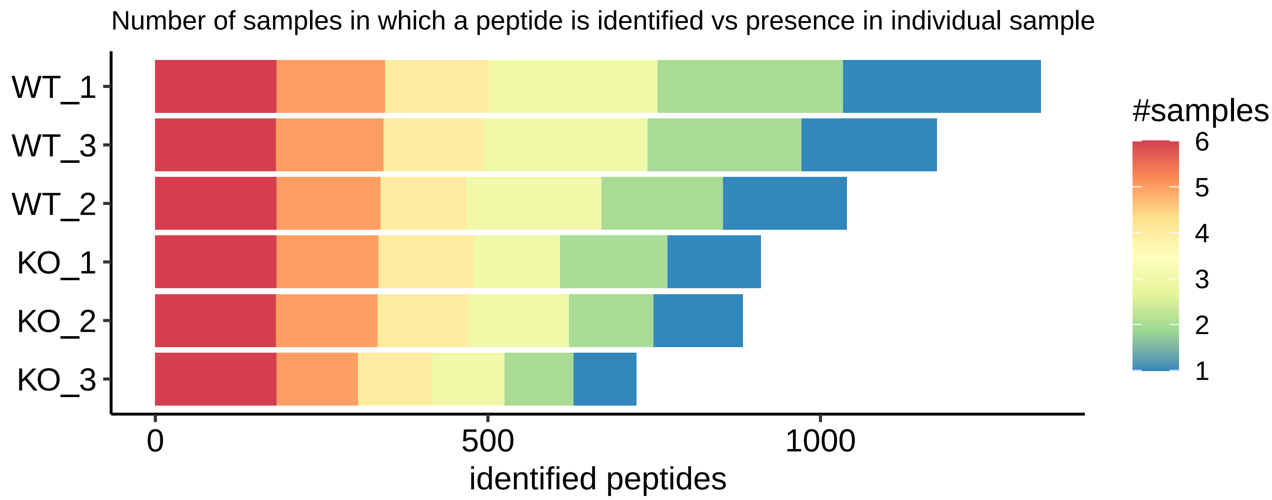


Samples flagged as 'exclude' (by user) are not taken into account in this figure. Exact values are shown for data points matching 90% and 50% of samples to convenience comparison between analyses (e.g. before/after configuring 'exclude' samples, or comparing between experiments of similar protocol).

1.2.2 peptide detection frequency

Each identified peptide in a sample is classified and color-coded by the number of other samples where the same peptide is present. Visualization of the amount of peptides that overlap with other samples in the dataset, from peptides identified in most samples (red) to one-hit-wonders (blue), helps identify uncommon samples (more blue/green than other samples).

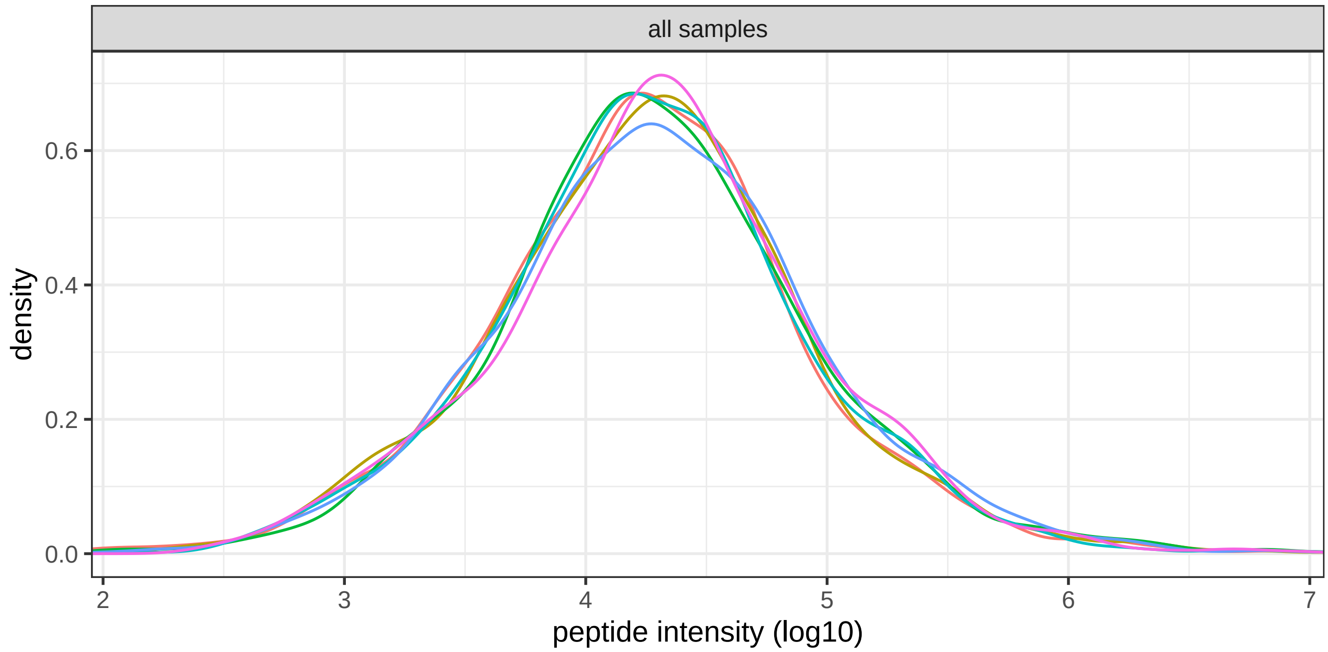
Optimally, the majority of peptides in each sample are red~orange with relatively few uniquely identified peptides (blue~green). Samples are sorted by the total amount of detected peptides.

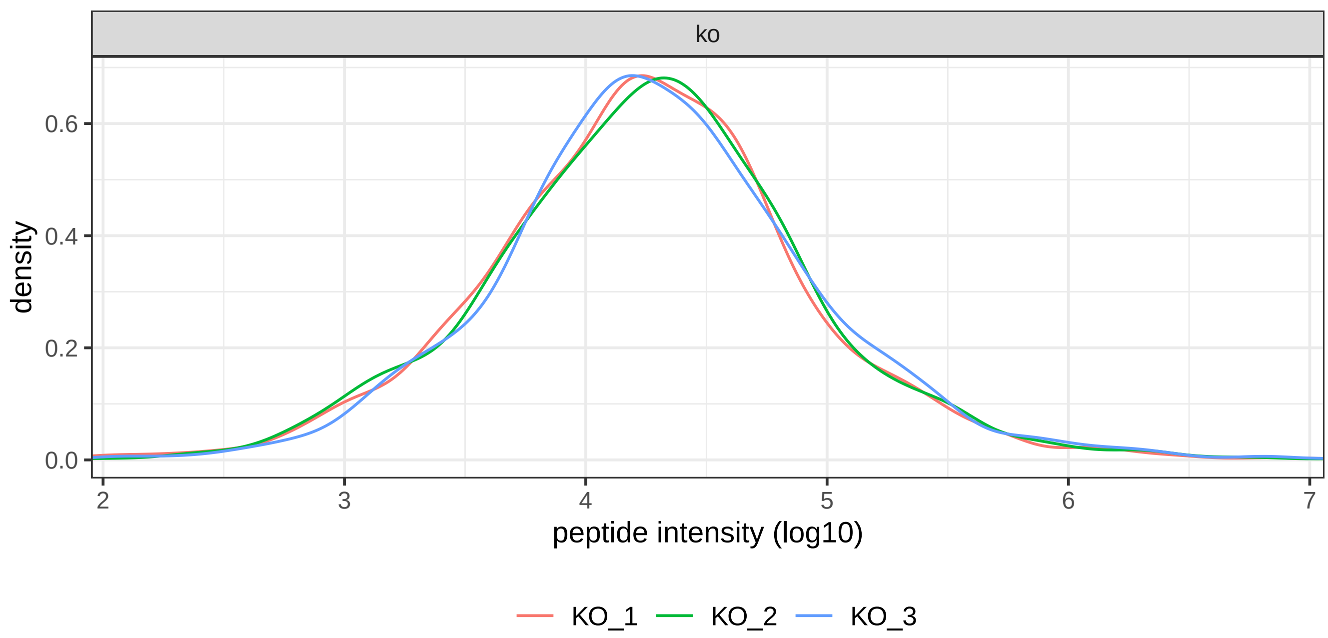


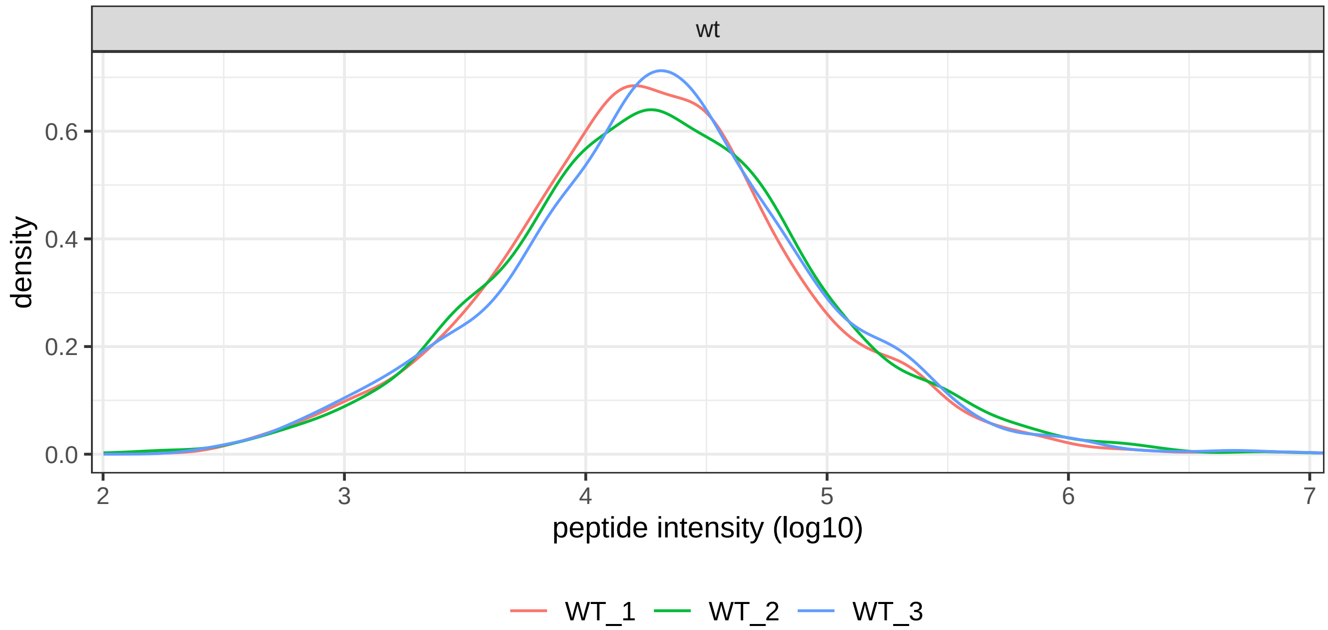
Number of samples in which a peptide is identified vs presence in individual sample
WT_1
WT_3
WT_2
KO_1
KO_2
KO_3
0
500
1000
identified peptides
#samples
6
5
4
3
2
1

1.3 abundance distributions

The figures in this subsection are used to identify unexpected mass-spec sensitivity or sample loading differences. Peptide data is shown as provided in input files, so peptide filtering nor intensity normalization has been applied yet (for proper QC, make sure the software that generated the input data did not apply normalization prior). If the dataset is DDA, match-between-runs (MBR) peptides are included in these distributions whereas for DIA only 'detected' peptides (based on confidence score threshold) are included.





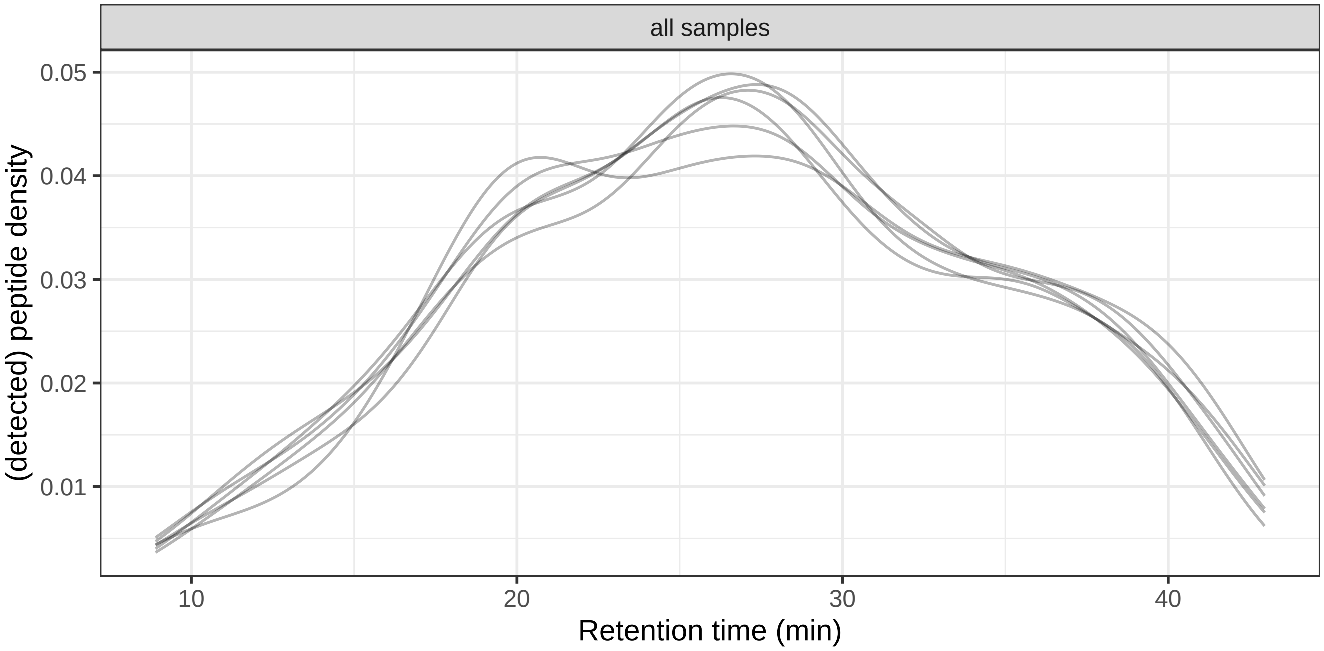

wt
density
0.6
0.4
0.2
0.0
2
3
4
5
6
7
peptide intensity (log10)
WT_1
WT_2
WT_3

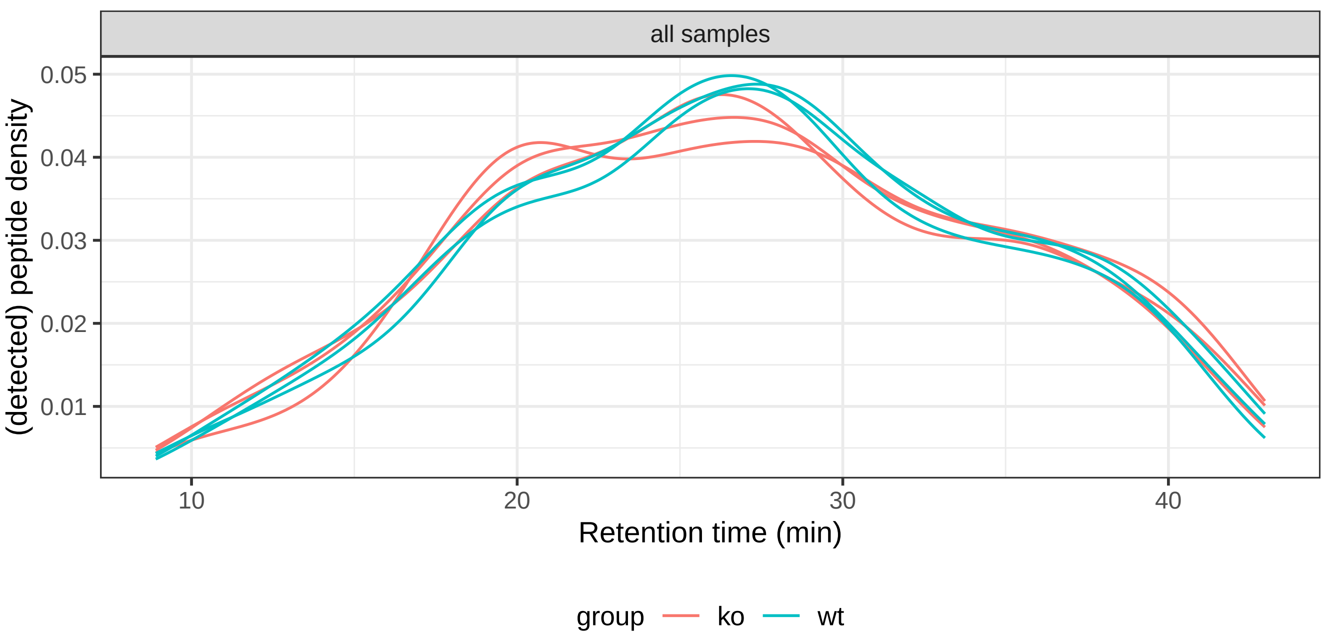
1.4 retention time

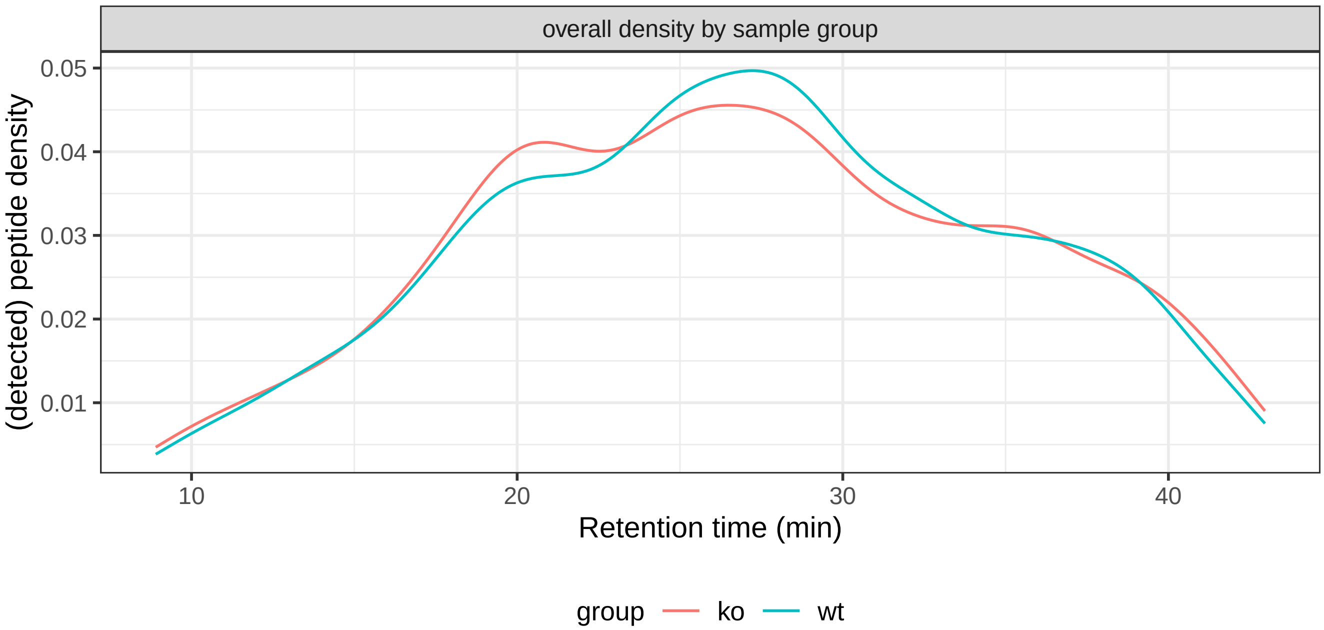
The figures in this section allow you to identify potential problems during HPLC elution, such as a temporarily blocking column, failing ionization spray or decreasing sensitivity over time. For each sample, all peptides that are also observed in a replicate (such that there is a point of reference available) are visualized.

1.4.1 retention time distributions

The density of the number of peptides eluting at each point in time. The figure below presents an overview of all samples that allows for the identification of outlier samples that follow distinct elution patterns. The following section shows details for each sample. Samples marked as 'exclude' in the provided sample metadata table are visualized as dashed lines.





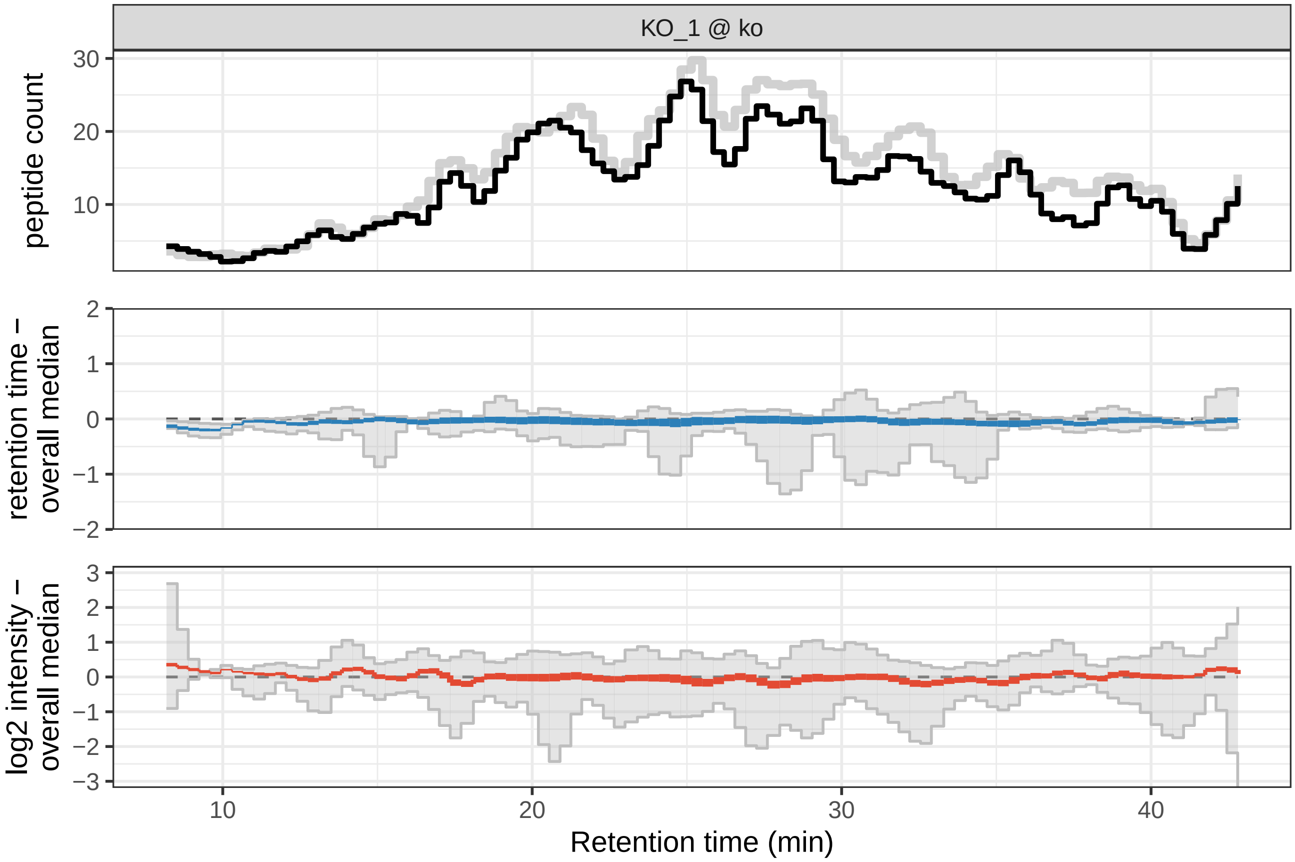


1.4.2 retention time local effects

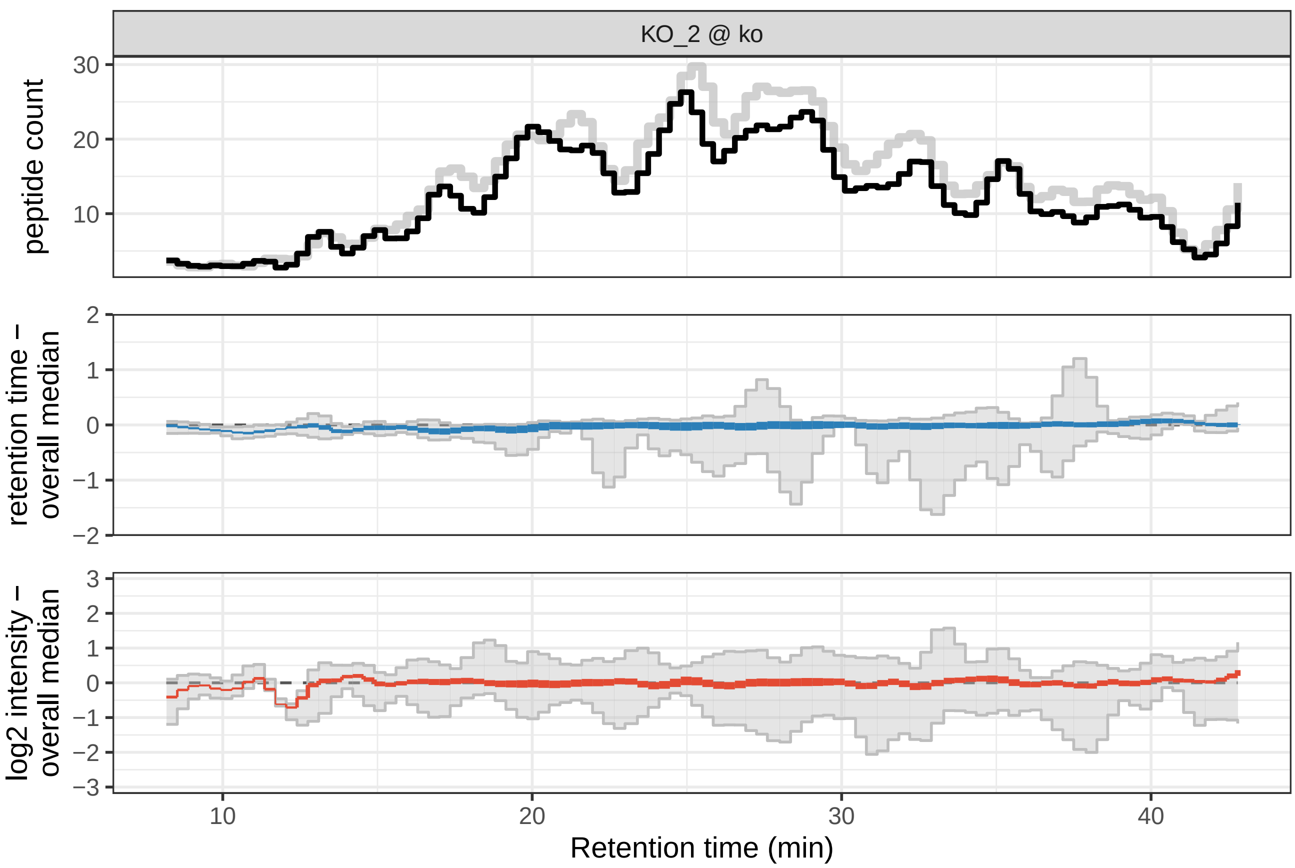
To investigate how each measurement differs from others, we visualize each sample as a 3 panel figure. First, the data is binned across the retention time dimension (x-axis). If a samples was marked as 'exclude' in the provided sample metadata, this is indicated in the plot title.

The top panel shows the number of peptides in the input data, e.g. as recognized by the software that generated input for this pipeline, over time (black line). For reference, the grey line shows the median amount over all samples (note; if this is the exact same in all samples, the grey line may not be visible as it falls behind the black line).

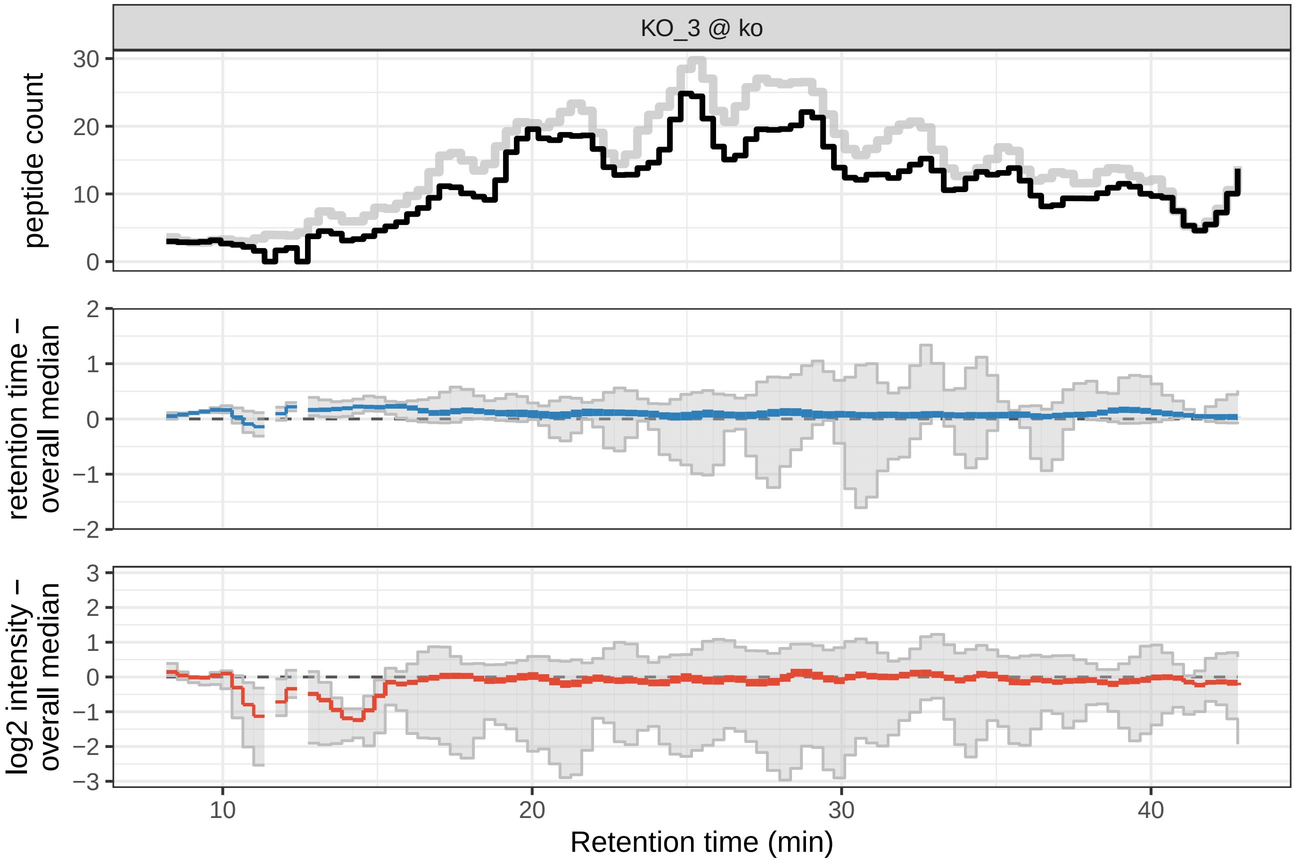
The middle panel indicates whether peptide retention times deviate from their median over all samples (blue line). The grey area depicts the 5% and 95% quantiles, respectively. The line width corresponds to the number of peptides eluting at that time (data from first panel). Analogously, the bottom panel shows the deviation in peptide abundance as compared to the median over all samples (red line).

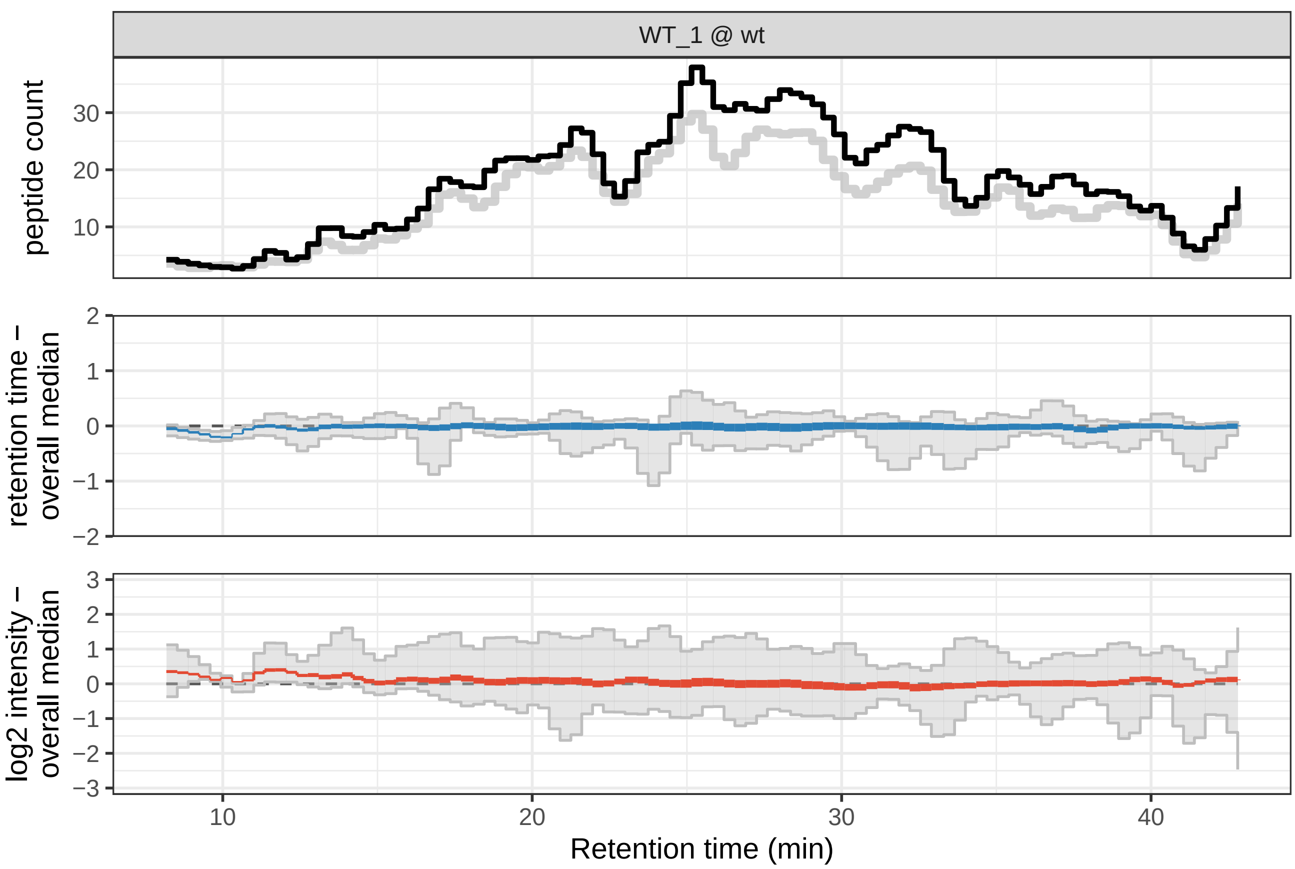


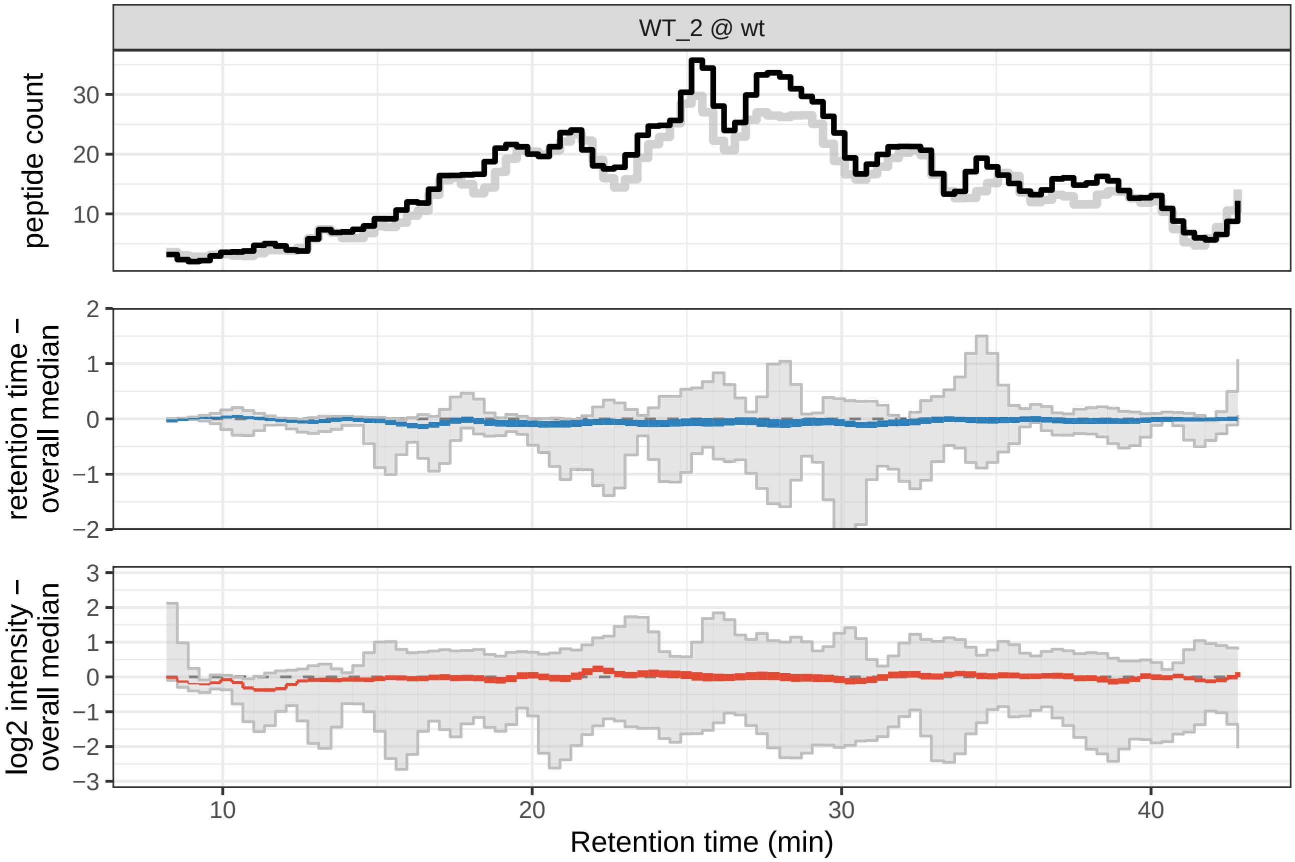
KO_1 @ ko
peptide count
retention time – overall median
log2 intensity – overall median
Retention time (min)

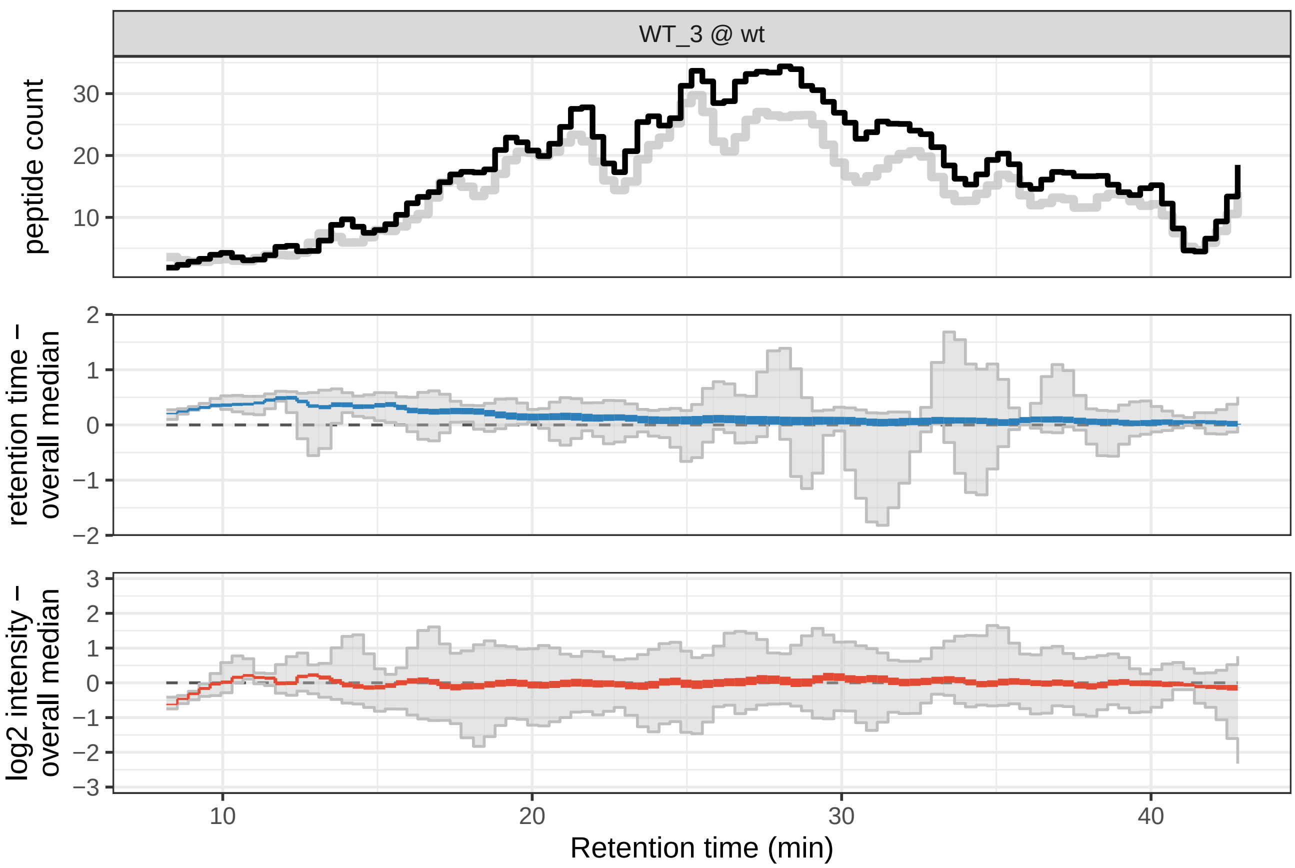


KO_2 @ ko
peptide count
retention time – overall median
log2 intensity – overall median
Retention time (min)


KO_3 @ ko
peptide count
retention time – overall median
log2 intensity – overall median
Retention time (min)


WT_1 @ wt
peptide count
retention time – overall median
log2 intensity – overall median
Retention time (min)


WT_2 @ wt
peptide count
retention time – overall median
log2 intensity – overall median
Retention time (min)

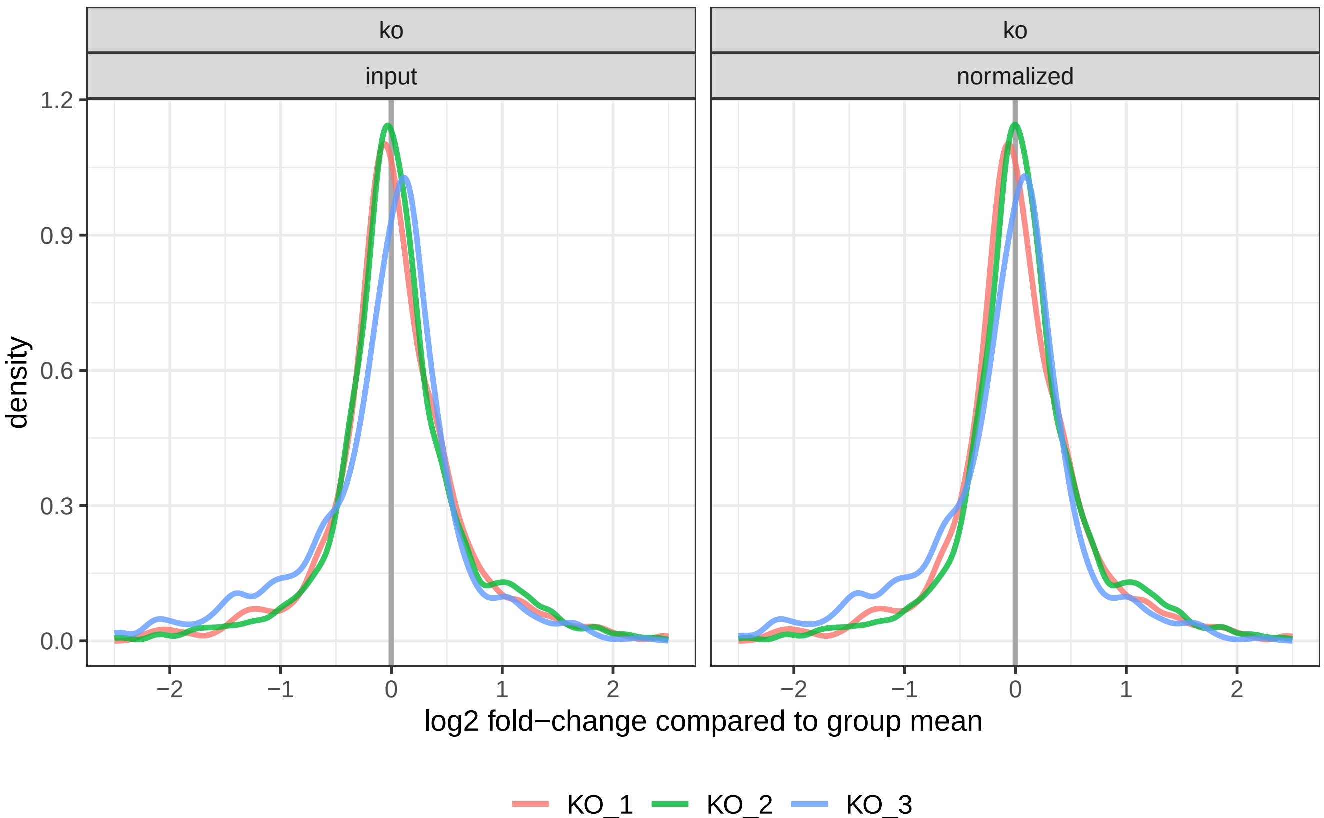

WT_3 @ wt
peptide count
retention time – overall median
log2 intensity – overall median
Retention time (min)

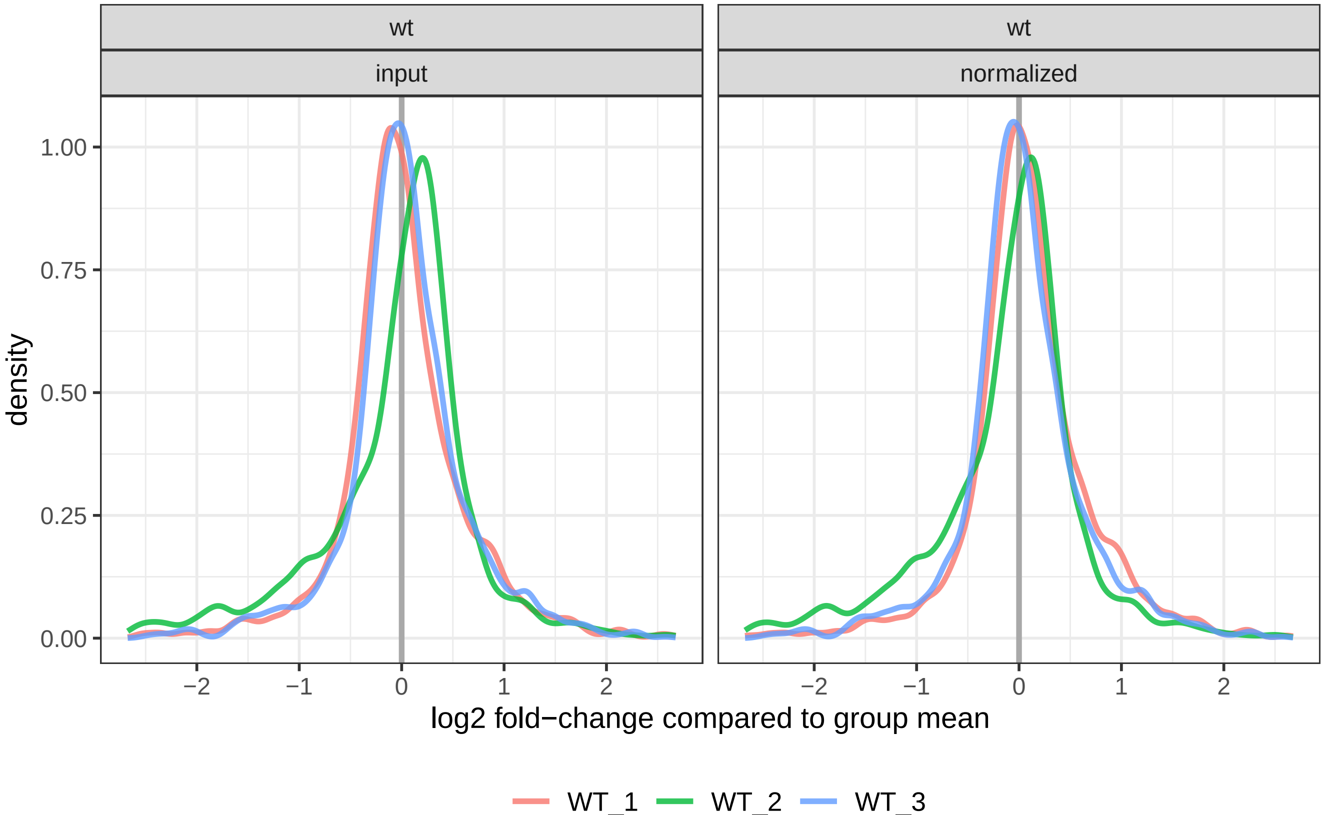
1.5 variation among replicates

The reproducibility of replicate measurements is expressed in three different analyses. First, the difference between peptide intensities in each sample are compared to the mean value among all replicates (foldchange distributions). Next, the Coefficient of Variation (CoV) is used as a metric for reproducibility to explore how much the CoV within a sample group can be improved by removing a single sample (eg; if CoV strongly improved after removing sample s, it could be regarded as an outlier). Finally, the CoV within each sample group is visualized as a boxplot and a violin plot, figures commonly seen in proteomics literature and useful for comparing across experiments (of similar protocol).

1.5.1 within-group foldchange distributions

The foldchange of all peptides in a sample is compared to their respective mean value over all samples in the group. This visualizes how strongly each sample deviates from other samples in the same group which helps identify outlier samples. The same data was used as detailed in the “retention time” section above. Samples marked as ‘exclude’ in the provided sample metadata table are visualized as dashed lines.




wt
input
wt
normalized
density
1.00
0.75
0.50
0.25
0.00
−2
−1
0
1
2
log2 fold−change compared to group mean
WT_1
WT_2
WT_3

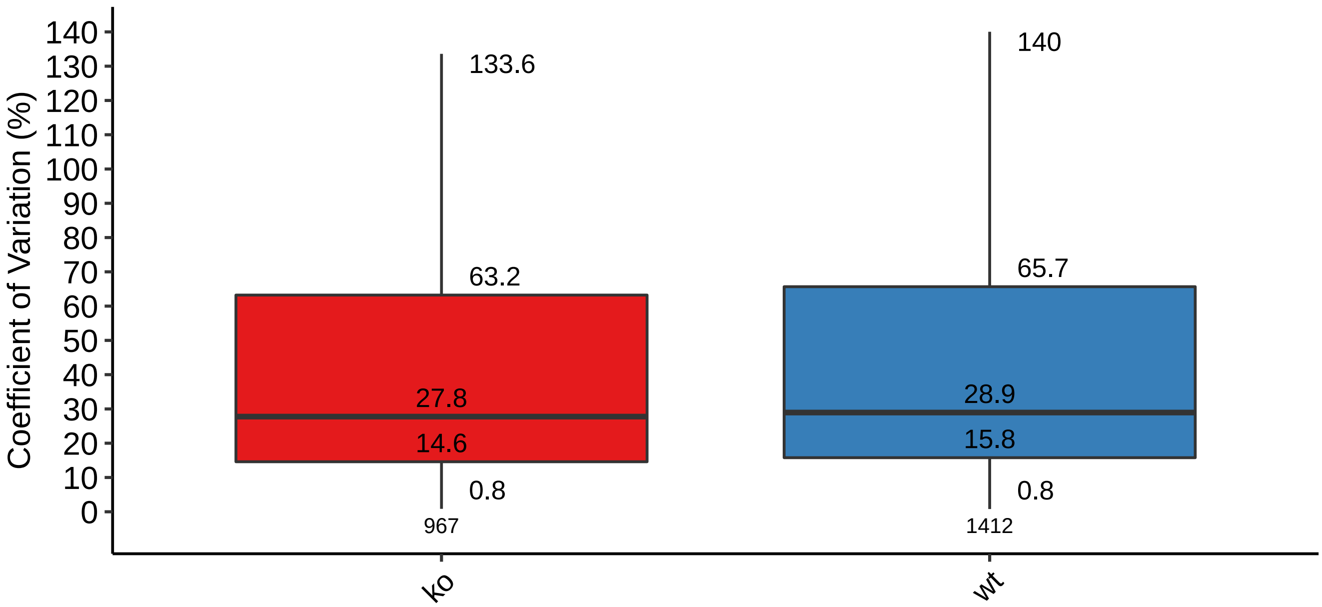
1.5.2 CoV, leave-one-out

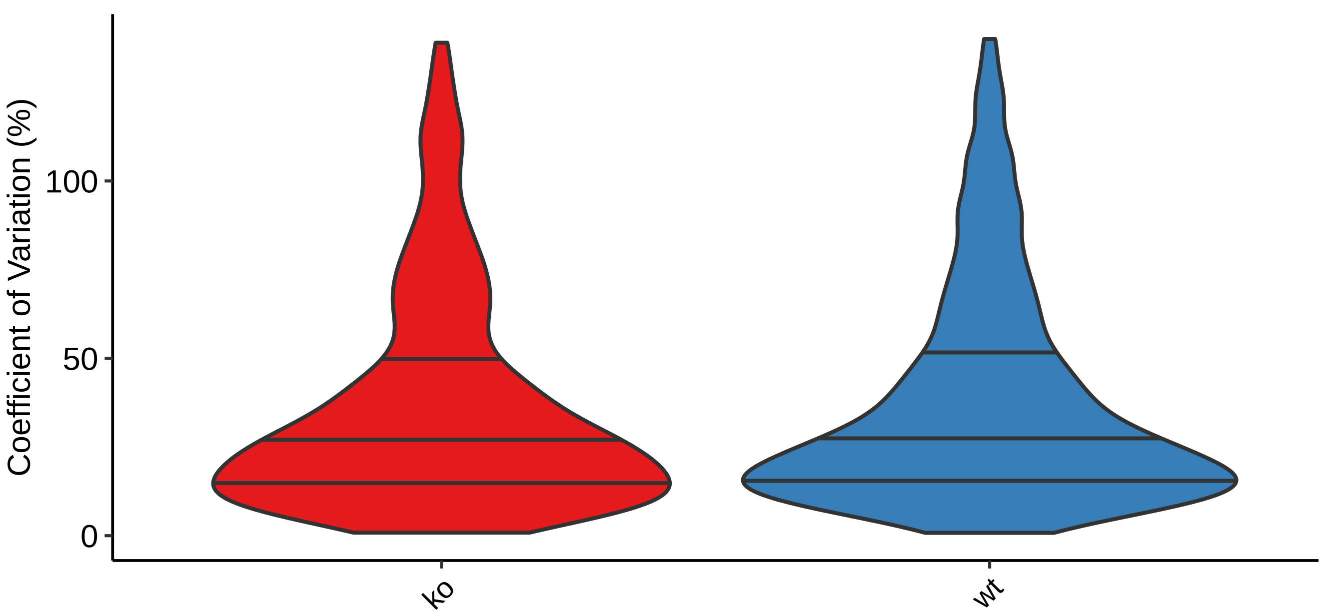
No CoV leave-one-out computations could be made, the dataset lacks sample groups with at least 4 replicate samples.

1.5.3 Coefficient of Variation

The Coefficient of Variation (CoV) is a quality metric for the reproducibility of replicate measurements, here visualized using box- and violin-plots.

The same data was used as detailed in the "retention time" section above, with only peptides quantified in at least 3 replicates used for CoV computation in a group. Only samples that are NOT marked 'exclude' in the provided sample metadata and are in a sample group among at least 3 replicates are used for these figures.





1.6 PCA

A visualization of the first three PCA dimensions illustrates sample clustering. The goal of these figures is to detect global effects from a quality control perspective, such as samples from the same experiment batch clustering together, not to be sensitive to a minor subset of differentially abundant proteins (for which specialized statistical models can be applied downstream).

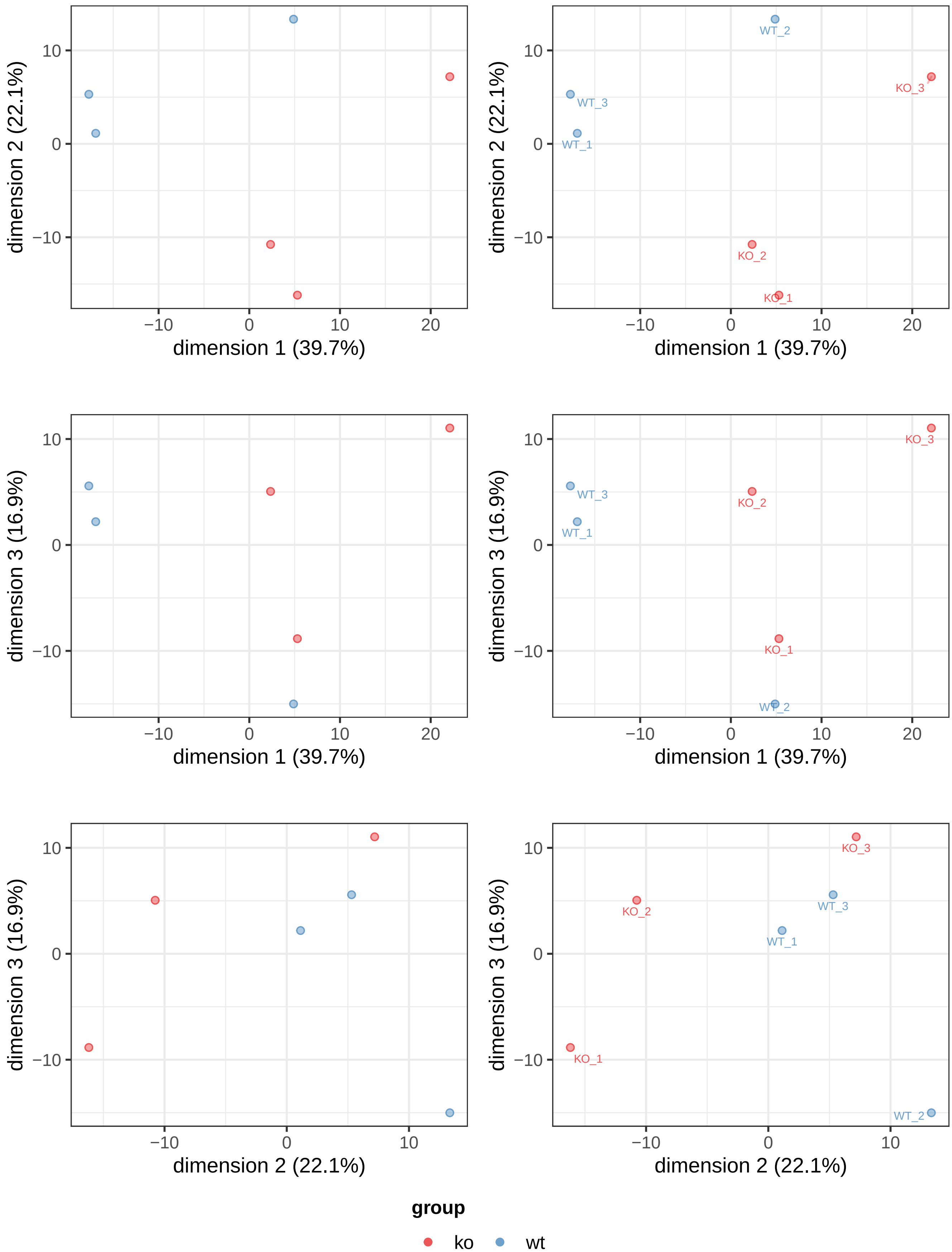
If additional sample metadata was provided, such as experiment batch, sample-prep dates, gel, etc., multiple PCA figures will be generated with respective color-codings. Users are encouraged to provide relevant experiment information as sample metadata and use these figures to search for unexpected batch effects.

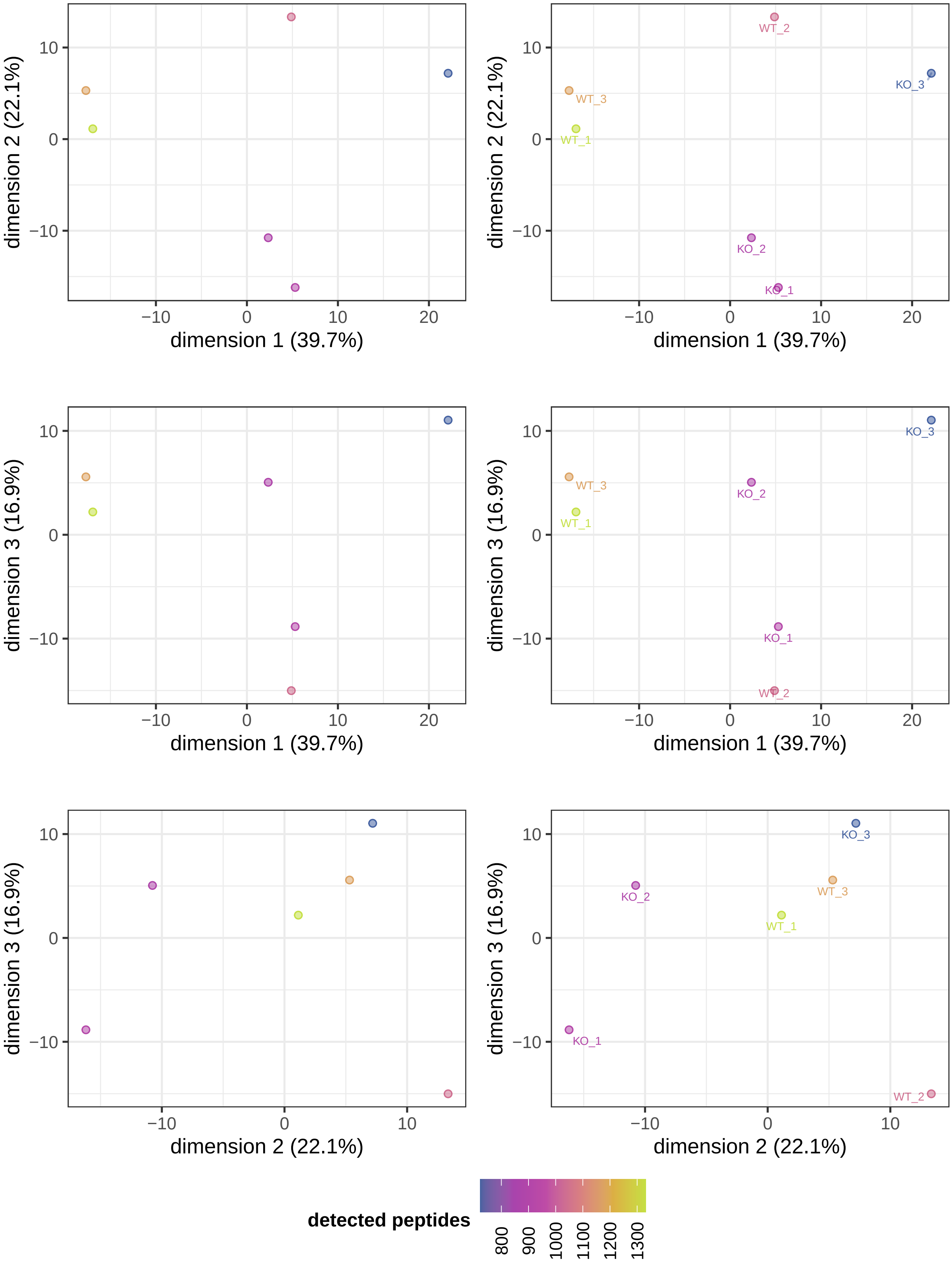
The pcaMethods R package is used here to perform the Probabilistic PCA (PPCA). The set of peptides used for this analysis consists of those peptides that pass your filter criteria in every sample group. If any samples are marked as ‘exclude’ in the provided sample metadata, an additional PCA plot is generated with these samples included (depicting the ‘exclude’ samples as square symbols).

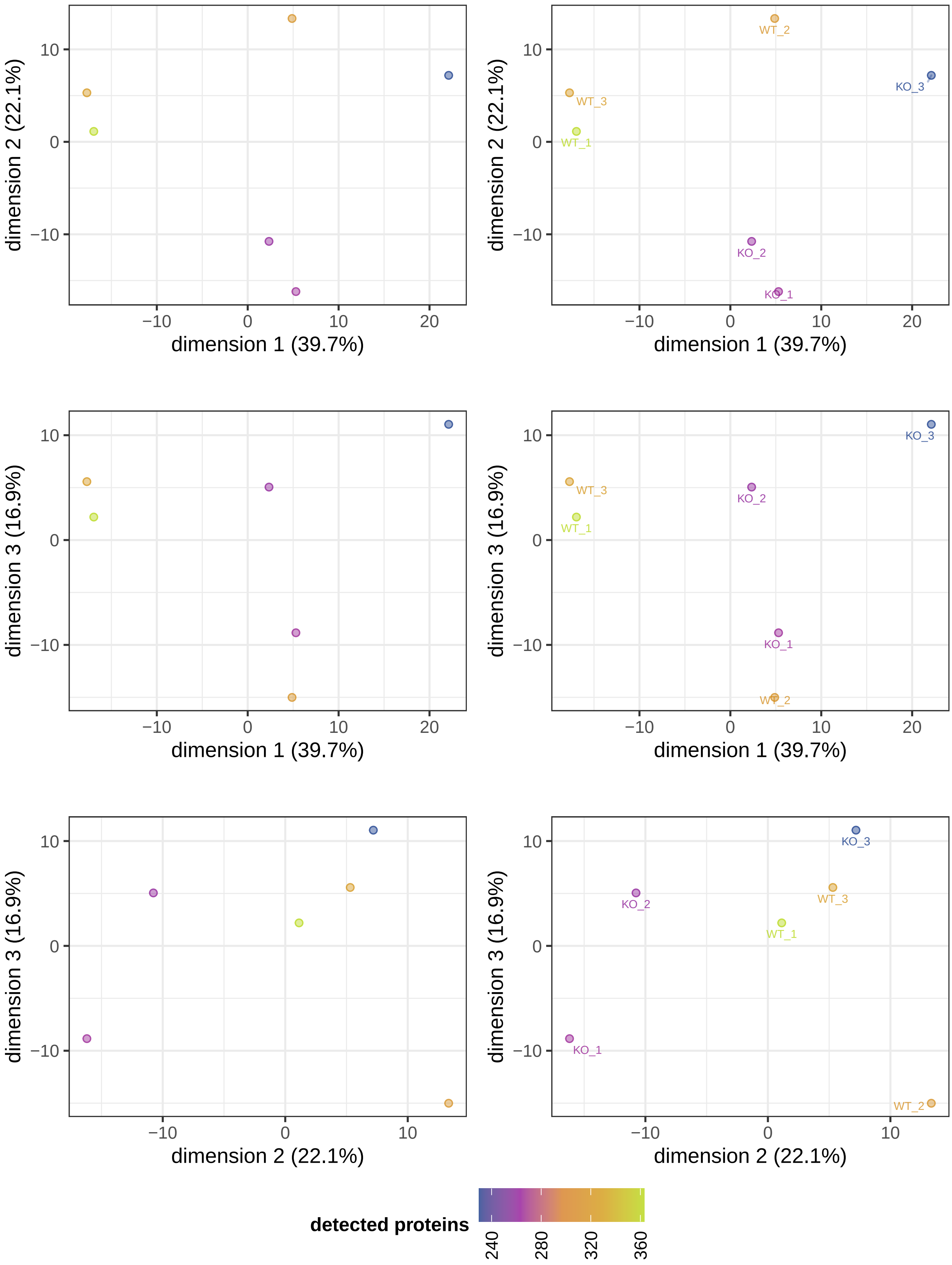
Rationale behind data filter
As mentioned above, the aim of the PCA figures is to identify global effects. To achieve this, we compute sample distances on the subset of peptides identified in each group which prevents rarely detected peptides/proteins from having a disproportionate effect on sample clustering. This pertains not only to ‘randomly detected contaminant proteins’ but also to proteins with abundance levels near the detection limit, which may be detected in only a subset of samples (eg; some measurements will be more successful/sensitive than others).

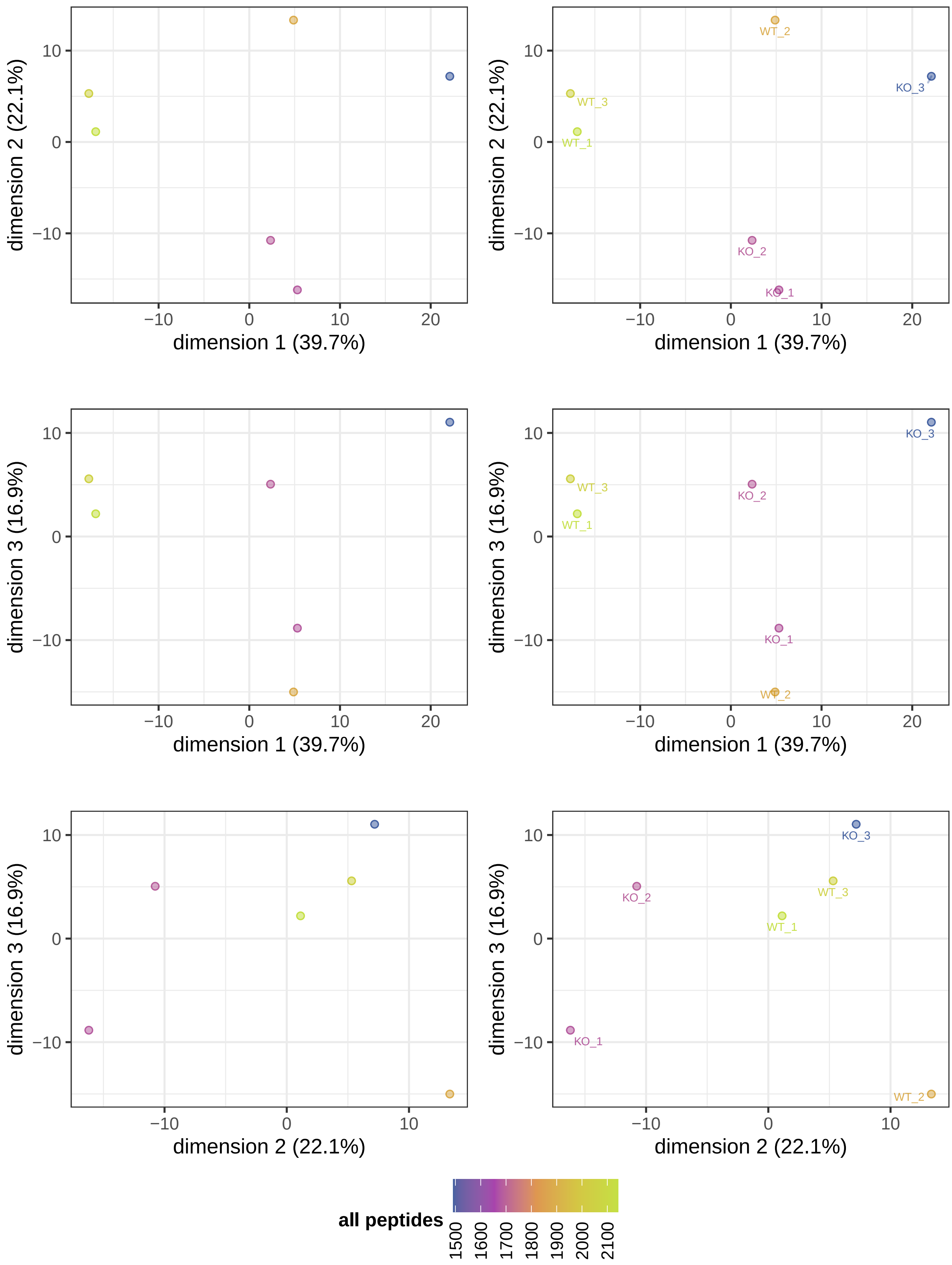
Figure legends
The first 3 principle components compared visually (1 *vs* 2, 1 *vs* 3, 2 *vs* 3) on the rows. Left- and right-side panels on each row represent the same figure without and with sample labels. The principle components are shown on the axis labels together with their respective percentage of variance explained. Samples marked as ‘exclude’ in the provided sample metadata, if any, are visualized as square shapes.

PCA only on samples not flagged as 'exclude', using 716 peptides

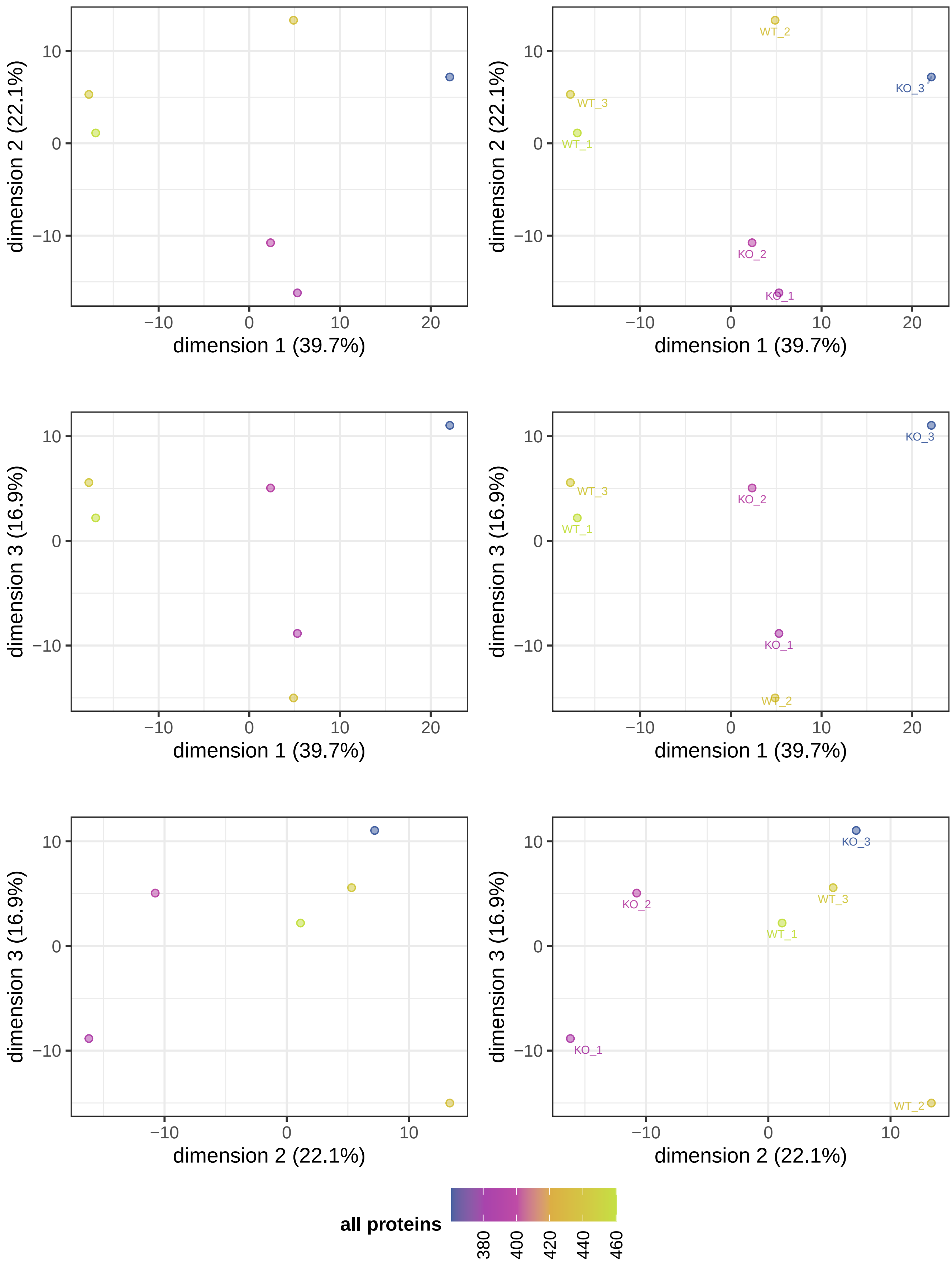



dimension 1 (39.7%)
dimension 2 (22.1%)
dimension 3 (16.9%)
WT_1
WT_2
WT_3
KO_1
KO_2
KO_3
detected peptides
800
900
1000
1100
1200
1300


dimension 1 (39.7%)
dimension 2 (22.1%)
dimension 3 (16.9%)
WT_1
WT_2
WT_3
KO_1
KO_2
KO_3
detected proteins
240
280
320
360



dimension 1 (39.7%)
dimension 2 (22.1%)
dimension 3 (16.9%)
WT_1
WT_2
WT_3
KO_1
KO_2
KO_3
all peptides
1500
1600
1700
1800
1900
2000
2100


dimension 1 (39.7%)
dimension 2 (22.1%)
dimension 3 (16.9%)
WT_1
WT_2
WT_3
KO_1
KO_2
KO_3
all proteins
380
400
420
440
460

2 Differential abundance analysis

goal: maximize reliable features for quantification

In a pairwise analysis of two groups of samples, only peptides with N data-points in both groups are used for quantitative analysis (where N = defined by user settings). For example; if peptide *p* is consistently quantified in sample groups A and B but not in C/D/E, it can be used when comparing group A *versus* group B but should not be used in any other group comparisons. This approach is particularly suited to maximize the number of peptides used for statistical analysis in experimental designs with many sample groups.

A common alternative strategy is a global filtering approach where peptides are selected based on their properties in the overall dataset (eg; present in x% of samples or x% of replicates in all groups) and subsequentially the resulting data matrix is used for all downstream statistical analyses. In the example above where peptide *p* is present in a subset of sample groups, *p* would either be left out (not present in majority of samples in entire dataset) or erronously used when applying t-statistics to groups B and C (since *p* is not present in group C, it may differentially detected but there are no features available for quantitative analysis)

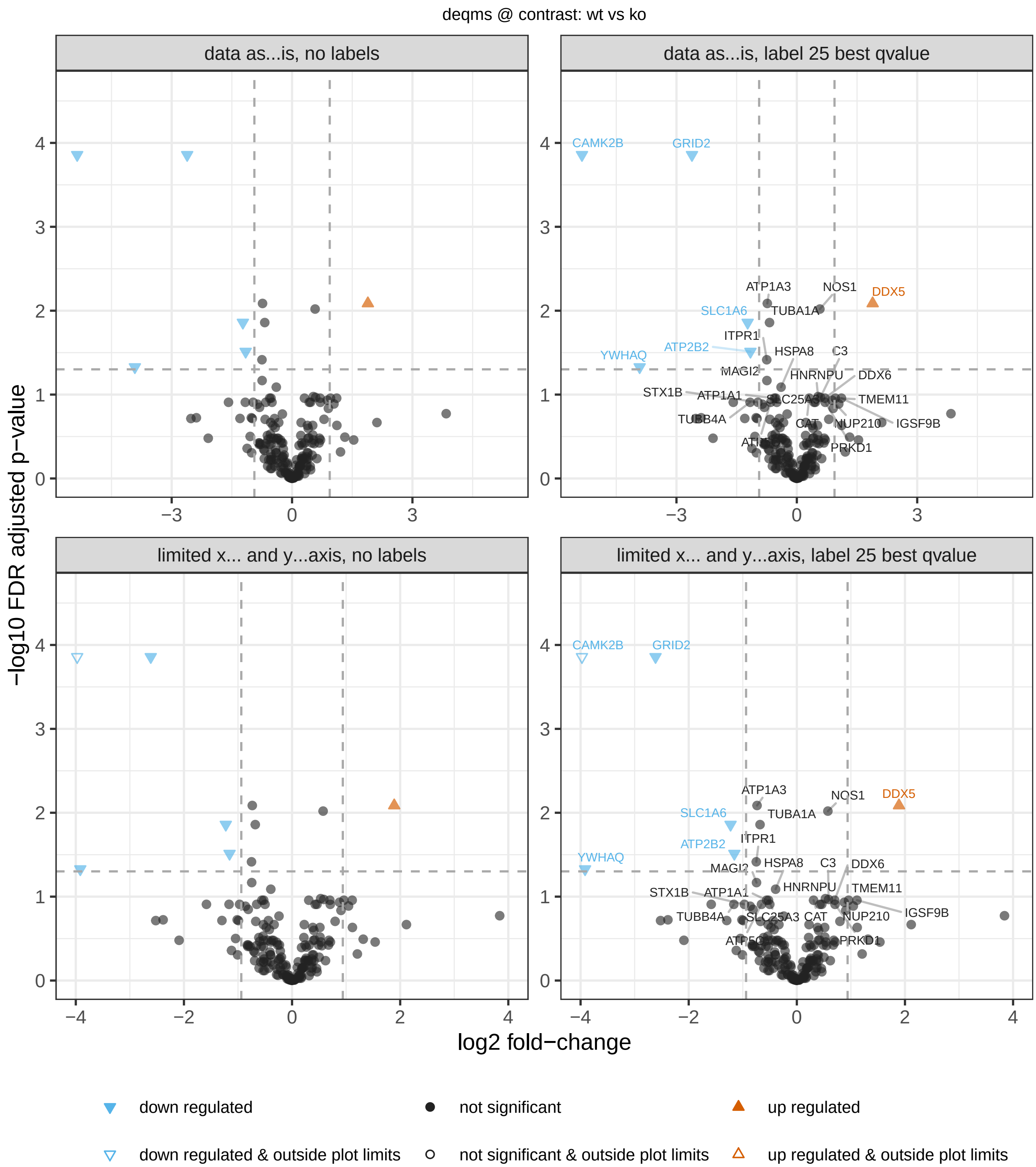
2.1 wt vs ko

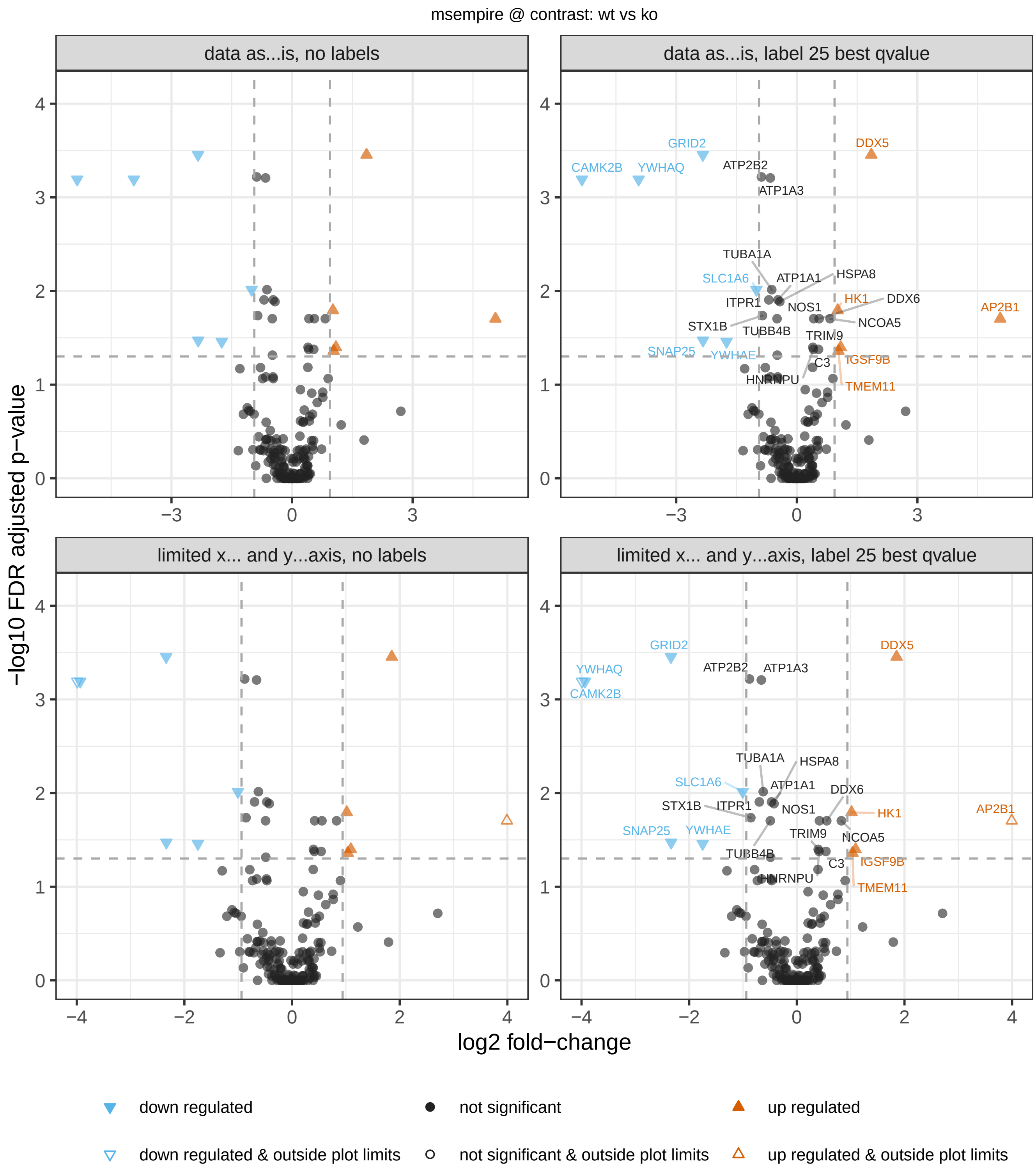
- **user setting:** using 'filter by contrast' peptide filtering approach
- 716 peptides in 193 proteins remain in the current contrast after peptide filters and are used for the statistical analysis in this section
- qvalue threshold: 0.05
- log2 foldchange threshold: 0.939

2.1.1 volcano

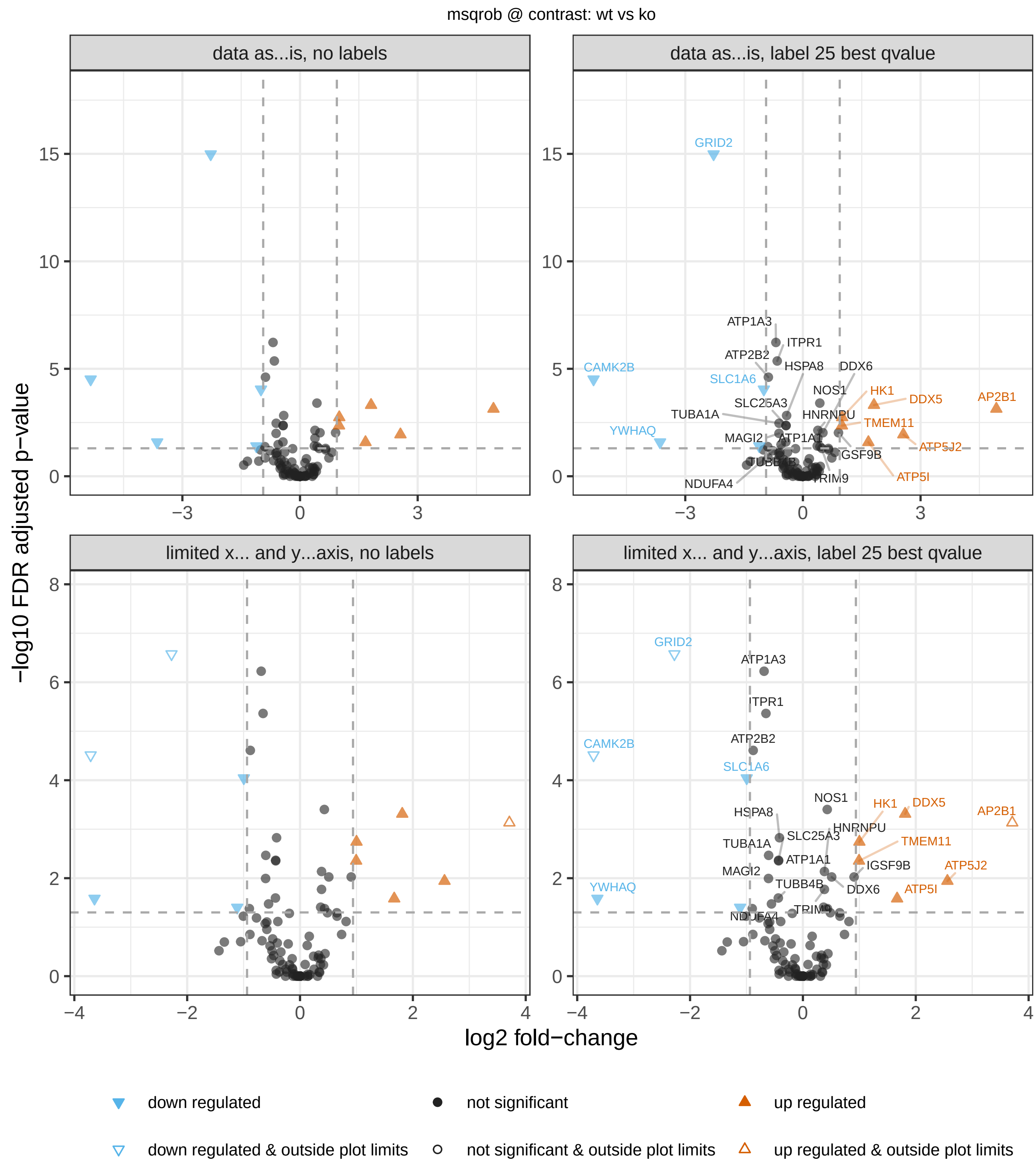
The plot title shows the statistical model and contrast (sample groups in the comparison). Left- and right-side figure panels on each row represent the same figure without and with labels for the 25 proteins with lowest p-value.

Bottom figure panels have limited x- and y-axis. For datasets with a small number of strong outliers in p-value or fold-change, which may have a profound effect on the plot scales, this allows inspection of the remainder of the volcano plot without disproportionate influence by 'extreme' values.


deqms @ contrast: wt vs ko
data as...is, no labels
data as...is, label 25 best qvalue
limited x... and y...axis, no labels
limited x... and y...axis, label 25 best qvalue
−log10 FDR adjusted p−value
log2 fold−change
CAMK2B
GRID2
SLC1A6
ATP2B2
YWHAQ
ATP1A3
NOS1
DDX5
TUBA1A
ITPR1
HSPA8
C3
MAGI2
HNRNPU
DDX6
STX1B
ATP1A1
SLC25A3
TMEM11
TUBB4A
CAT
NUP210
IGSF9B
PRKD1
down regulated
not significant
up regulated
down regulated & outside plot limits
not significant & outside plot limits
up regulated & outside plot limits



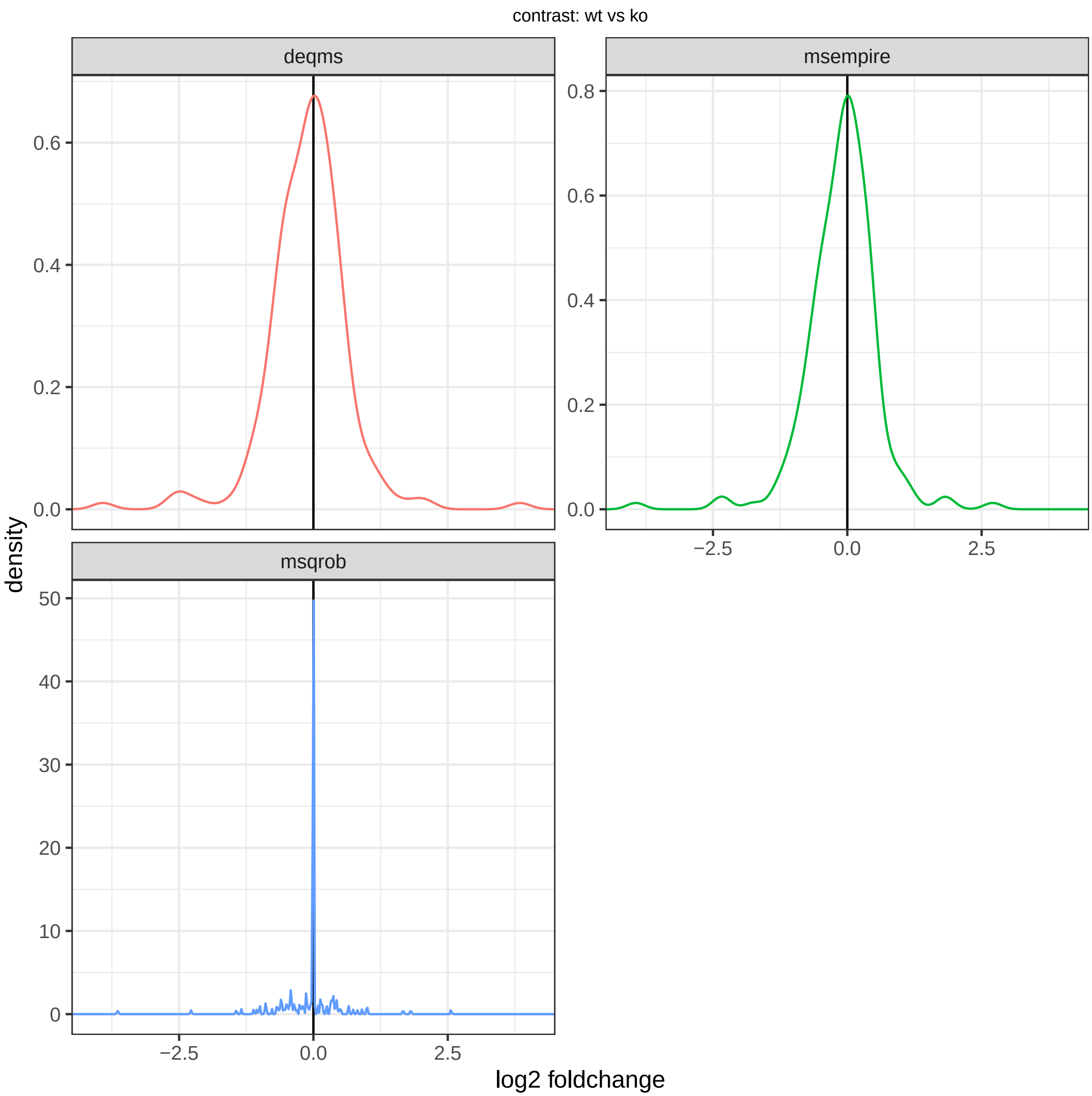
msempire @ contrast: wt vs ko
data as...is, no labels
data as...is, label 25 best qvalue
limited x... and y...axis, no labels
limited x... and y...axis, label 25 best qvalue
−log10 FDR adjusted p−value
log2 fold−change
GRID2
CAMK2B
YWHAQ
ATP2B2
ATP1A3
DDX5
TUBA1A
SLC1A6
ATP1A1
HSPA8
ITPR1
NOS1
HK1
DDX6
AP2B1
STX1B
TUBB4B
NCOA5
TRIM9
SNAP25
YWHAE
C3
IGSF9B
HNRNPU
TMEM11
down regulated
not significant
up regulated
down regulated & outside plot limits
not significant & outside plot limits
up regulated & outside plot limits


msqrob @ contrast: wt vs ko
data as...is, no labels
data as...is, label 25 best qvalue
limited x... and y...axis, no labels
limited x... and y...axis, label 25 best qvalue
−log10 FDR adjusted p−value
log2 fold−change
GRID2
CAMK2B
SLC1A6
YWHAQ
ATP1A3
ITPR1
ATP2B2
HSPA8
DDX6
NOS1
HK1
DDX5
AP2B1
SLC25A3
HNRNPU
TUBA1A
TMEM11
MAGI2
ATP1A1
ATP5J2
IGSF9B
TUBB4B
NDUFA4
TRIM9
ATP5I
down regulated
not significant
up regulated
down regulated & outside plot limits
not significant & outside plot limits
up regulated & outside plot limits

2.1.2 foldchange distribution

Distributions of estimated foldchanges produced by the statistical models. If the mode is far from 0, consider alternative normalization strategies. Do note the scale on the x-axis, for some experiments the foldchanges are very low which in turn may exaggerate this figure.

note; the MSqRob model tends to assign zero (log)foldchange for proteins with minor difference between conditions where the model is very sure the null hypothesis cannot be rejected (shrinkage by the ridge regression model). As a result, many foldchanges will be zero and the density plot for MSqRob may look like a spike instead of the expected Gaussian shape observed in other models



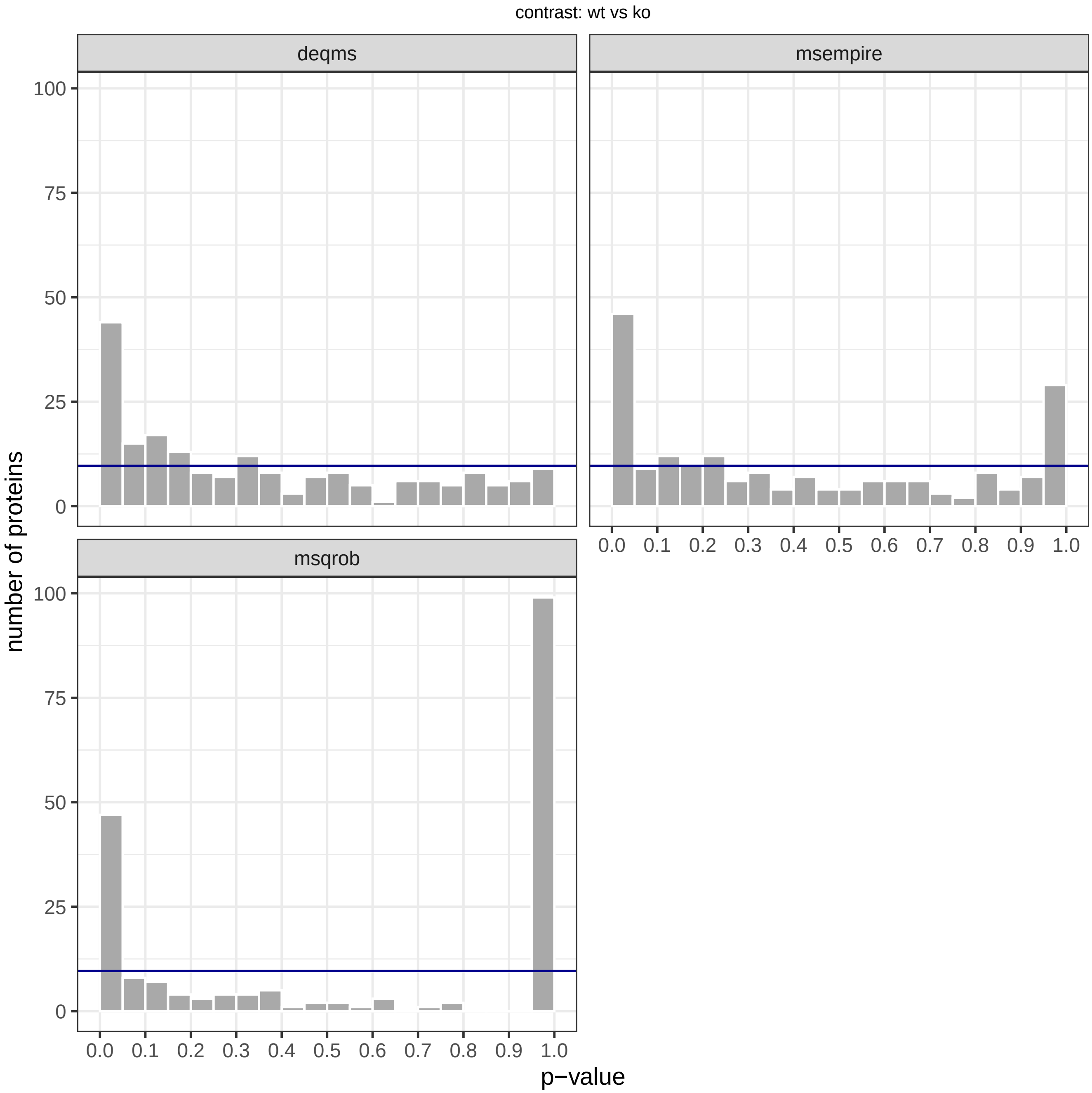
2.1.3 p-value distribution

Histogram of p-values computed by differential expression analysis algorithms, as-is, for quality-control inspection. The horizontal line indicates the expected counts assuming a uniform distribution (total number of p-values divided by number of histogram bins)

See further: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6164648/

See further: http://varianceexplained.org/statistics/interpreting-pvalue-histogram/

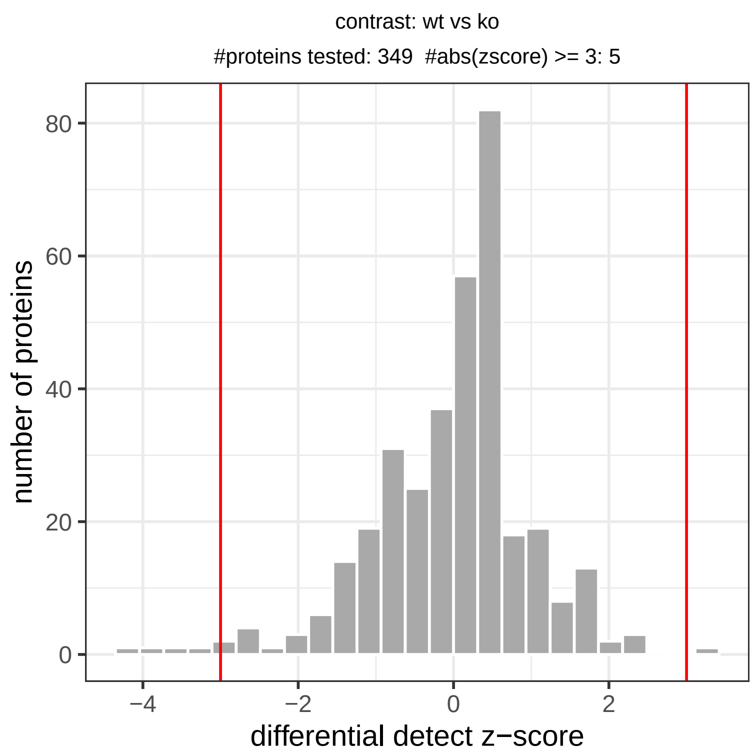
note; the MSqRob and MS-EmpiRe models often yield p-value distributions that show a large peak at p-value 1, these are typically proteins with estimated log foldchanges at/near zero where these models are very sure the null hypothesis cannot be rejected



2.1.4 differential detect

Some proteins may not have peptides with sufficient data points over samples to be used for differential expression analysis (depending on the user-defined filtering criteria in how many replicates peptides should be observed), but do show a strong difference in the number of detected peptides between sample groups. In some proteomics experimental designs, for example a wildtype-knockout APMS study, those are interesting proteins. The DEA based on peptide abundance values (volcano plots above) are the main result for differential testing in MS-DAP but as a situationally useful tool MS-DAP also includes a 'protein detection' z-score, based on the number of times a peptide for each protein was detected per sample group (/experimental condition), as an alternative means of differential testing.

Below figure shows the distribution of these scores with thresholds at 3 std. Both the z-scores and the counts these are based upon are available in the statistical result Excel table.



3 Summary of differential testing

Differential Expression Analysis: number of proteins found statistically significant.

contrast	algorithm	#test	#hits	top10 significant
wt vs ko	deqms	193	6	grid2, camk2b, ddx5, slc1a6, atp2b2, ywhaq
wt vs ko	msempire	193	11	grid2, ddx5, ywhaq, camk2b, slc1a6, hk1, ap2b1, snap25, ywhae, igsf9b
wt vs ko	msqrob	193	11	grid2, camk2b, slc1a6, ddx5, ap2b1, hk1, tmem11, atp5j2, ywhaq, atp5i

Differential Detection: prioritize proteins with more peptide detections in some group. A simple metric to complement results from DEA, which is the main result (eg; consider proteins with too few data points for DEA).

contrast	#proteins tested	#abs(zscore) >= 3	top10
wt vs ko	349	5	gria2, shisa6, gria3, acap3, gria1, camk2g, camk2b, camk2a, nedd4l, camk2d

4 log

[info] reading MetaMorpheus search task...

[info] Parsing MetaMorpheus protein-groups, all non-unique peptides were assigned to their largest protein-group (by peptide count)

[info] 2850 target precursors, 2798 (plain)sequences, 682 proteins

[info] 1131/1131 protein accessions and 682/682 protein groups were mapped to provided fasta file(s)

[progress] filtered proteins: >tr|A0A075B5J6|A0A075B5J6_MOUSE Predicted gene, 20730 OS=Mus musculus OX=10090 GN=Gm20730 PE=4 SV=7;>sp|P01629|KV2A4_MOUS ... >tr|A0A075B5J9|A0A075B5J9_MOUSE Isoform of A0A0G2JDN5, Immunoglobulin kappa variable 17-127 OS=Mus musculus OX=10090 GN=I ... >tr|A0A075B5K0|A0A075B5K0_MOUSE Immunoglobulin kappa variable 14-126 (Fragment) OS=Mus musculus OX=10090 GN=Igkv14-126 PE ... >tr|A0A075B5K1|A0A075B5K1_MOUSE Isoform of P01650, Immunoglobulin kappa variable 11-125 OS=Mus musculus OX=10090 GN=Igkv1 ... >tr|A0A075B5K2|A0A075B5K2_MOUSE Isoform of A0A0G2JDR3, Immunoglobulin kappa chain variable 9-124 OS=Mus musculus OX=10090 ... >tr|A0A075B5K5|A0A075B5K5_MOUSE Isoform of P01627, Immunoglobulin kappa variable 2-112 OS=Mus musculus OX=10090 GN=Igkv2- ... >tr|A0A075B5K6|A0A075B5K6_MOUSE Immunoglobulin kappa variable 2-109 (Fragment) OS=Mus musculus OX=10090 GN=Igkv2-109 PE=4 ... >tr|A0A075B5K7|A0A075B5K7_MOUSE Isoform of A0A0G2JG36, Immunoglobulin kappa chain variable 14-100 OS=Mus musculus OX=1009 ... >tr|A0A075B5K9|A0A075B5K9_MOUSE Isoform of A0A0G2JGF9, Immunoglobulin kappa variable 12-98 OS=Mus musculus OX=10090 GN=Ig ... >tr|A0A075B5L1|A0A075B5L1_MOUSE Immunoglobulin kappa variable 10-94 (Fragment) OS=Mus musculus OX=10090 GN=Igkv10-94 PE=4 ... >tr|A0A075B5L7|A0A075B5L7_MOUSE Immunoglobulin kappa variable 4-80 (Fragment) OS=Mus musculus OX=10090 GN=Igkv4-80 PE=4 SV=2 >tr|A0A075B5M1|A0A075B5M1_MOUSE Isoform of A0A0G2JFU6, Immunoglobulin kappa variable 4-63 OS=Mus musculus OX=10090 GN=Igk ... >tr|A0A075B5M3|A0A075B5M3_MOUSE Isoform of A0A0G2JGI6, Immunoglobulin kappa variable 4-58 OS=Mus musculus OX=10090 GN=Igk ... >tr|A0A075B5M7|A0A075B5M7_MOUSE Isoform of A0A0G2JDV4, Immunoglobulin kappa variable 5-39 OS=Mus musculus OX=10090 GN=Igk ... >tr|A0A075B5N2|A0A075B5N2_MOUSE Isoform of A0A0G2JEI6, Immunoglobulin kappa chain variable 6-29 OS=Mus musculus OX=10090 ... >tr|A0A075B5N3|A0A075B5N3_MOUSE Isoform of A0A0G2JE47, Immunoglobulin kappa variable 8-28 OS=Mus musculus OX=10090 GN=Igk ... >tr|A0A075B5N4|A0A075B5N4_MOUSE Isoform of A0A0G2JF28, Immunoglobulin kappa chain variable 8-27 OS=Mus musculus OX=10090 ... >tr|A0A075B5N5|A0A075B5N5_MOUSE Immunoglobulin kappa variable 8-26 (Fragment) OS=Mus musculus OX=10090 GN=Igkv8-26 PE=4 SV=7 >tr|A0A075B5N6|A0A075B5N6_MOUSE Isoform of A0A0G2JDG9, Immunoglobulin kappa variable 8-16 OS=Mus musculus OX=10090 GN=Igk ... >tr|A0A075B5N7|A0A075B5N7_MOUSE Isoform of A0A0G2JGY3, Immunoglobulin kappa variable 6-13 OS=Mus musculus OX=10090 GN=Igk ... >tr|A0A075B5N9|A0A075B5N9_MOUSE Isoform of A0A0G2JE31, Immunoglobulin kappa variable 3-7 OS=Mus musculus OX=10090 GN=Igkv ... >tr|A0A075B5P1|A0A075B5P1_MOUSE Immunoglobulin kappa variable 3-1 (Fragment) OS=Mus musculus OX=10090 GN=Igkv3-1 PE=1 SV=7 >tr|A0A075B5P3|A0A075B5P3_MOUSE Isoform of P01867, Immunoglobulin heavy constant gamma 2B (Fragment) OS=Mus musculus OX=1 ... >tr|A0A075B5P4|A0A075B5P4_MOUSE Isoform of P01869, Ig gamma-1 chain C region secreted form (Fragment) OS=Mus musculus OX= ... >tr|A0A075B5P5|A0A075B5P5_MOUSE Isoform of A0A1Y7VJN6, Immunoglobulin heavy constant gamma 3 (Fragment) OS=Mus musculus O ... >tr|A0A075B5P9|A0A075B5P9_MOUSE Immunoglobulin heavy variable 5-4 (Fragment) OS=Mus musculus OX=10090 GN=Ighv5-4 PE=4 SV=1 >tr|A0A075B5Q2|A0A075B5Q2_MOUSE Immunoglobulin heavy variable 5-9 (Fragment) OS=Mus musculus OX=10090 GN=Ighv5-9 PE=4 SV=1 >tr|A0A075B5Q3|A0A075B5Q3_MOUSE Isoform of A0A0A6YY69, Immunoglobulin heavy variable 2-5 OS=Mus musculus OX=10090 GN=Ighv ... >tr|A0A075B5Q4|A0A075B5Q4_MOUSE Immunoglobulin heavy variable 5-12 (Fragment) OS=Mus musculus OX=10090 GN=Ighv5-12 PE=4 SV=1 >tr|A0A075B5Q6|A0A075B5Q6_MOUSE Isoform of A0A0A6YVS4, Immunoglobulin heavy variable 5-9-1 OS=Mus musculus OX=10090 GN=Ig ... >tr|A0A075B5Q9|A0A075B5Q9_MOUSE Isoform of A0A0A6YWC7, Immunoglobulin heavy variable 5-15 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B5R0|A0A075B5R0_MOUSE Isoform of A0A0B4J1P4, Immunoglobulin heavy variable 5-16 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B5R1|A0A075B5R1_MOUSE Immunoglobulin heavy variable 5-17 (Fragment) OS=Mus musculus OX=10090 GN=Ighv5-17 PE=4 SV=1 >tr|A0A075B5R2|A0A075B5R2_MOUSE Immunoglobulin heavy variable 7-3 (Fragment) OS=Mus musculus OX=10090 GN=Ighv7-3 PE=4 SV=2 >tr|A0A075B5R4|A0A075B5R4_MOUSE Immunoglobulin heavy variable 14-1 (Fragment) OS=Mus musculus OX=10090 GN=Ighv14-1 PE=4 SV=1 >tr|A0A075B5R5|A0A075B5R5_MOUSE Immunoglobulin heavy variable 4-1 (Fragment) OS=Mus musculus OX=10090 GN=Ighv4-1 PE=1 SV=7 >tr|A0A075B5R7|A0A075B5R7_MOUSE Immunoglobulin heavy variable 14-2 (Fragment) OS=Mus musculus OX=10090 GN=Ighv14-2 PE=4 SV=1 >tr|A0A075B5S2|A0A075B5S2_MOUSE Immunoglobulin heavy variable 7-1 (Fragment) OS=Mus musculus OX=10090 GN=Ighv7-1 PE=1 SV=2 >tr|A0A075B5S5|A0A075B5S5_MOUSE Isoform of A0A0A6YXL5, Immunoglobulin heavy variable 7-4 OS=Mus musculus OX=10090 GN=Ighv ... >tr|A0A075B5S9|A0A075B5S9_MOUSE Isoform of A0A0A6YWH6, Immunoglobulin heavy variable 9-4 OS=Mus musculus OX=10090 GN=Ighv ... >tr|A0A075B5T2|A0A075B5T2_MOUSE Immunoglobulin heavy variable 6-3 (Fragment) OS=Mus musculus OX=10090 GN=Ighv6-3 PE=4 SV=7 >tr|A0A075B5T3|A0A075B5T3_MOUSE Immunoglobulin heavy variable 6-6 (Fragment) OS=Mus musculus OX=10090 GN=Ighv6-6 PE=4 SV=1 >tr|A0A075B5T4|A0A075B5T4_MOUSE Immunoglobulin heavy variable 1-4 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-4 PE=4 SV=1 >tr|A0A075B5T5|A0A075B5T5_MOUSE Isoform of A0A0A6YWG2, Immunoglobulin heavy variable V1-5 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B5T6|A0A075B5T6_MOUSE Immunoglobulin heavy variable V10-3 (Fragment) OS=Mus musculus OX=10090 GN=Ighv10-3 PE=4 SV=1 >tr|A0A075B5T9|A0A075B5T9_MOUSE Isoform of A0A0A6YW37, Immunoglobulin heavy variable V1-9 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B5U6|A0A075B5U6_MOUSE Isoform of A0A0A6YX66, Immunoglobulin heavy variable V1-20 OS=Mus musculus OX=10090 GN=Ig ... >tr|A0A075B5U7|A0A075B5U7_MOUSE Immunoglobulin heavy variable 1-22 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-22 PE=1 SV=1 >tr|A0A075B5V5|A0A075B5V5_MOUSE Immunoglobulin heavy variable 1-39 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-39 PE=4 SV=1 >tr|A0A075B5V6|A0A075B5V6_MOUSE Isoform of A0A0A6YY38, Immunoglobulin heavy variable V1-42 OS=Mus musculus OX=10090 GN=Ig ... >tr|A0A075B5V7|A0A075B5V7_MOUSE Isoform of A0A0A6YXN5, Immunoglobulin heavy variable V1-43 OS=Mus musculus OX=10090 GN=Ig ... >tr|A0A075B5V8|A0A075B5V8_MOUSE Isoform of A0A0A6YY41, Immunoglobulin heavy variable 1-47 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B5W2|A0A075B5W2_MOUSE Isoform of A0A0A6YXC3, Immunoglobulin heavy variable 1-52 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B5W3|A0A075B5W3_MOUSE Immunoglobulin heavy variable 1-53 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-53 PE=4 SV=1 >tr|A0A075B5W4|A0A075B5W4_MOUSE Isoform of A0A0A6YXA3, Immunoglobulin heavy variable V8-6 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B5W5|A0A075B5W5_MOUSE Immunoglobulin heavy variable V1-54 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-54 PE=4 SV=1 >tr|A0A075B5W6|A0A075B5W6_MOUSE Immunoglobulin heavy variable 1-55 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-55 PE=1 SV=1 >tr|A0A075B5W9|A0A075B5W9_MOUSE Immunoglobulin heavy variable 1-58 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-58 PE=4 SV=1 >tr|A0A075B5X3|A0A075B5X3_MOUSE Immunoglobulin heavy variable 1-64 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-64 PE=4 SV=1 >tr|A0A075B5X7|A0A075B5X7_MOUSE Isoform of A0A0G2JFS5, Immunoglobulin heavy variable 1-69 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B5Y1|A0A075B5Y1_MOUSE Isoform of A0A0G2JGK2, Immunoglobulin heavy variable V1-74 OS=Mus musculus OX=10090 GN=Ig ... >tr|A0A075B5Y2|A0A075B5Y2_MOUSE Immunoglobulin heavy variable 1-75 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-75 PE=4 SV=1 >tr|A0A075B5Y3|A0A075B5Y3_MOUSE Immunoglobulin heavy variable 1-80 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-80 PE=1 SV=1 >tr|A0A075B5Y4|A0A075B5Y4_MOUSE Immunoglobulin heavy variable 1-81 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-81 PE=4 SV=1 >tr|A0A075B5Y5|A0A075B5Y5_MOUSE Isoform of P06327, Immunoglobulin heavy variable 1-84 (Fragment) OS=Mus musculus OX=10090 ... >tr|A0A075B5Y6|A0A075B5Y6_MOUSE Immunoglobulin heavy variable 1-85 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-85 PE=4 SV=1 >tr|A0A075B666|A0A075B666_MOUSE Immunoglobulin kappa chain variable 13-85 (Fragment) OS=Mus musculus OX=10090 GN=Igkv13-8 ... >tr|A0A075B674|A0A075B674_MOUSE Isoform of A0A0G2JGN3, Immunoglobulin heavy variable 1-78 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A075B677|A0A075B677_MOUSE Isoform of A0A0G2JFC6, Immunoglobulin kappa variable 4-53 OS=Mus musculus OX=10090 GN=Igk ... >tr|A0A075B680|A0A075B680_MOUSE Isoform of A0A0A6YXZ4, Immunoglobulin heavy variable 1-62-2 OS=Mus musculus OX=10090 GN=I ... >tr|A0A075B684|A0A075B684_MOUSE Isoform of A0A075B5U3, Immunoglobulin heavy variable 1-62-1 OS=Mus musculus OX=10090 GN=I ... >tr|A0A075B696|A0A075B696_MOUSE Immunoglobulin heavy variable 2-3 (Fragment) OS=Mus musculus OX=10090 GN=Ighv2-3 PE=4 SV=3 >tr|A0A075B6D5|A0A075B6D5_MOUSE Isoform of A0A0G2JFZ3, Immunoglobulin kappa chain variable 19-93 OS=Mus musculus OX=10090 ... >tr|A0A087WPN7|A0A087WPN7_MOUSE Isoform of A0A0A6YWN3, Immunoglobulin heavy variable 13-2 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A0A0MQC1|A0A0A0MQC1_MOUSE Isoform of P18533, Immunoglobulin heavy variable 3-5 OS=Mus musculus OX=10090 GN=Ighv3-5 ... >tr|A0A0A6YVW3|A0A0A6YVW3_MOUSE Immunoglobulin heavy variable V1-23 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-23 PE=4 ... >tr|A0A0A6YWI9|A0A0A6YWI9_MOUSE Immunoglobulin heavy variable V1-11 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-11 PE=4 SV=1 >tr|A0A0A6YXN4|A0A0A6YXN4_MOUSE Immunoglobulin heavy variable V1-18 (Fragment) OS=Mus musculus OX=10090 GN=Ighv1-18 PE=1 SV=1 >tr|A0A0A6YXQ0|A0A0A6YXQ0_MOUSE Immunoglobulin heavy variable 8-8 (Fragment) OS=Mus musculus OX=10090 GN=Ighv8-8 PE=4 SV=1 >tr|A0A0A6YY53|A0A0A6YY53_MOUSE Isoform of F6TQW2, Immunoglobulin heavy constant gamma 2C (Fragment) OS=Mus musculus OX=1 ... >tr|A0A0A6YYE7|A0A0A6YYE7_MOUSE Immunoglobulin kappa variable 4-57 (Fragment) OS=Mus musculus OX=10090 GN=Igkv4-57 PE=4 SV=2 >tr|A0A0B4J1H6|A0A0B4J1H6_MOUSE Immunoglobulin kappa chain variable 2-137 (Fragment) OS=Mus musculus OX=10090 GN=Igkv2-13 ... >tr|A0A0B4J1H7|A0A0B4J1H7_MOUSE Immunoglobulin kappa variable 1-135 (Fragment) OS=Mus musculus OX=10090 GN=Igkv1-135 PE=4 ... >tr|A0A0B4J1H9|A0A0B4J1H9_MOUSE Immunoglobulin kappa variable 1-132 (Fragment) OS=Mus musculus OX=10090 GN=Igkv1-132 PE=4 ... >tr|A0A0B4J1I1|A0A0B4J1I1_MOUSE Immunoglobulin kappa variable 16-104 (Fragment) OS=Mus musculus OX=10090 GN=Igkv16-104 PE ... >tr|A0A0B4J1I4|A0A0B4J1I4_MOUSE Immunoglobulin kappa chain variable 4-72 (Fragment) OS=Mus musculus OX=10090 GN=Igkv4-72 ... >tr|A0A0B4J1I7|A0A0B4J1I7_MOUSE Immunoglobulin kappa variable 4-68 (Fragment) OS=Mus musculus OX=10090 GN=Igkv4-68 PE=4 SV=1 >tr|A0A0B4J1I8|A0A0B4J1I8_MOUSE Immunoglobulin kappa variable 4-59 (Fragment) OS=Mus musculus OX=10090 GN=Igkv4-59 PE=4 S ... >tr|A0A0B4J1I9|A0A0B4J1I9_MOUSE Isoform of P04945, Immunoglobulin kappa variable 4-55 (Fragment) OS=Mus musculus OX=10090 ... >tr|A0A0B4J1J0|A0A0B4J1J0_MOUSE Immunoglobulin kappa variable 4-50 (Fragment) OS=Mus musculus OX=10090 GN=Igkv4-50 PE=4 SV=1 >tr|A0A0B4J1J2|A0A0B4J1J2_MOUSE Immunoglobulin kappa chain variable 5-43 (Fragment) OS=Mus musculus OX=10090 GN=Igkv5-43 ... >tr|A0A0B4J1J4|A0A0B4J1J4_MOUSE Immunoglobulin heavy variable V9-2 (Fragment) OS=Mus musculus OX=10090 GN=Ighv9-2 PE=4 SV=1 >tr|A0A0B4J1J5|A0A0B4J1J5_MOUSE Immunoglobulin heavy variable V9-3 (Fragment) OS=Mus musculus OX=10090 GN=Ighv9-3 PE=1 SV=1 >tr|A0A0B4J1J6|A0A0B4J1J6_MOUSE Immunoglobulin heavy variable 10-1 (Fragment) OS=Mus musculus OX=10090 GN=Ighv10-1 PE=4 SV=1 >tr|A0A0B4J1J7|A0A0B4J1J7_MOUSE Isoform of A0A0G2JEU7, Immunoglobulin heavy variable 1-82 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A0B4J1M0|A0A0B4J1M0_MOUSE Isoform of A0A0G2JGS9, Immunoglobulin heavy variable 1-77 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A0B4J1N0|A0A0B4J1N0_MOUSE Isoform of A0A0G2JFE9, Immunoglobulin heavy variable 1-76 OS=Mus musculus OX=10090 GN=Igh ... >tr|A0A0G2JDE1|A0A0G2JDE1_MOUSE Immunoglobulin heavy variable V8-12 (Fragment) OS=Mus musculus OX=10090 GN=Ighv8-12 PE=1 SV=1 >tr|A0A0G2JE99|A0A0G2JE99_MOUSE Immunoglobulin lambda constant 1 (Fragment) OS=Mus musculus OX=10090 GN=Iglc1 PE=4 SV=1;> ... >tr|A0A0G2JGQ8|A0A0G2JGQ8_MOUSE Isoform of P01845, Ig lambda-3 chain C region (Fragment) OS=Mus musculus OX=10090 GN=Iglc ... >tr|A0A140T8M0|A0A140T8M0_MOUSE Immunoglobulin kappa variable 1-117 (Fragment) OS=Mus musculus OX=10090 GN=Igkv1-117 PE=4 ... >tr|A0A140T8M2|A0A140T8M2_MOUSE Immunoglobulin kappa variable 12-44 (Fragment) OS=Mus musculus OX=10090 GN=Igkv12-44 PE=4 ... >tr|A0A140T8M3|A0A140T8M3_MOUSE Immunoglobulin kappa chain variable 8-30 (Fragment) OS=Mus musculus OX=10090 GN=Igkv8-30 ... >tr|A0A140T8M4|A0A140T8M4_MOUSE Isoform of A0A0G2JFN8, Immunoglobulin kappa variable 8-19 OS=Mus musculus OX=10090 GN=Igk ... >tr|A0A140T8M5|A0A140T8M5_MOUSE Immunoglobulin kappa variable 6-15 (Fragment) OS=Mus musculus OX=10090 GN=Igkv6-15 PE=1 SV=2 >tr|A0A140T8N0|A0A140T8N0_MOUSE Isoform of P01639, Immunoglobulin kappa chain variable 9-120 (Fragment) OS=Mus musculus O ... >tr|A0A140T8N2|A0A140T8N2_MOUSE Isoform of P01642, Protein Igkv5-48 (Fragment) OS=Mus musculus OX=10090 GN=Igkv5-48 PE=1 ... >tr|A0A140T8N3|A0A140T8N3_MOUSE Immunoglobulin kappa chain variable 13-84 (Fragment) OS=Mus musculus OX=10090 GN=Igkv13-8 ... >tr|A0A140T8N5|A0A140T8N5_MOUSE Immunoglobulin kappa variable 6-23 (Fragment) OS=Mus musculus OX=10090 GN=Igkv6-23 PE=4 SV=2 >tr|A0A140T8N7|A0A140T8N7_MOUSE Immunoglobulin kappa chain variable 6-25 (Fragment) OS=Mus musculus OX=10090 GN=Igkv6-25 ... >tr|A0A140T8N8|A0A140T8N8_MOUSE Immunoglobulin kappa variable 9-123 (Fragment) OS=Mus musculus OX=10090 GN=Igkv9-123 PE=4 ... >tr|A0A140T8N9|A0A140T8N9_MOUSE Immunoglobulin kappa variable 6-32 (Fragment) OS=Mus musculus OX=10090 GN=Igkv6-32 PE=4 SV=2 >tr|A0A140T8P1|A0A140T8P1_MOUSE Immunoglobulin kappa variable 6-14 (Fragment) OS=Mus musculus OX=10090 GN=Igkv6-14 PE=4 SV=2 >tr|A0A140T8P3|A0A140T8P3_MOUSE Immunoglobulin kappa chain variable 15-103 (Fragment) OS=Mus musculus OX=10090 GN=Igkv15- ... >tr|A0A140T8P5|A0A140T8P5_MOUSE Immunoglobulin kappa chain variable 8-24 (Fragment) OS=Mus musculus OX=10090 GN=Igkv8-24 ... >tr|A0A140T8P6|A0A140T8P6_MOUSE Immunoglobulin kappa variable 12-46 (Fragment) OS=Mus musculus OX=10090 GN=Igkv12-46 PE=1 ... >tr|A0A140T8P7|A0A140T8P7_MOUSE Immunoglobulin kappa variable 8-21 (Fragment) OS=Mus musculus OX=10090 GN=Igkv8-21 PE=1 SV=2 >tr|A0A140T8P9|A0A140T8P9_MOUSE Immunoglobulin kappa variable 14-111 (Fragment) OS=Mus musculus OX=10090 GN=Igkv14-111 PE ... >tr|A0A140T8Q1|A0A140T8Q1_MOUSE Isoform of P01635, Protein Igkv12-41 (Fragment) OS=Mus musculus OX=10090 GN=Igkv12-41 PE= ... >sp|P01631|KV2A7_MOUSE Ig kappa chain V-II region 26-10 OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P01633|KV5A1_MOUSE Ig kappa chain V19-17 OS=Mus musculus OX=10090 GN=Igk-V19-17 PE=1 SV=1 >sp|P01636|KV5A4_MOUSE Ig kappa chain V-V region MOPC 149 OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P01644|KV5AB_MOUSE Ig kappa chain V-V region HP R16.7 OS=Mus musculus OX=10090 PE=1 SV=1;>sp|P01645|KV5AC_MOUSE Ig ka ... >sp|P01654|KV3A1_MOUSE Ig kappa chain V-III region PC 2880/PC 1229 OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P01660|KV3A8_MOUSE Ig kappa chain V-III region PC 3741/TEPC 111 OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P01665|KV3AD_MOUSE Ig kappa chain V-III region PC 7043 OS=Mus musculus OX=10090 PE=1 SV=1;>sp|P01666|KV3AE_MOUSE Ig k ... >sp|P01671|KV3AJ_MOUSE Ig kappa chain V-III region PC 7175 OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P01675|KV6A1_MOUSE Ig kappa chain V-VI region XRPC 44 OS=Mus musculus OX=10090 PE=1 SV=1;>sp|P01676|KV6A2_MOUSE Ig ka ... >sp|P01837|IGKC_MOUSE Immunoglobulin kappa constant OS=Mus musculus OX=10090 GN=Igkc PE=1 SV=2 >sp|P01844|LAC2_MOUSE Ig lambda-2 chain C region OS=Mus musculus OX=10090 GN=Iglc2 PE=1 SV=1;>tr|Q99JC1|Q99JC1_MOUSE Isof ... >sp|P01863|GCAA_MOUSE Ig gamma-2A chain C region, A allele OS=Mus musculus OX=10090 GN=Ighg PE=1 SV=1;>sp|P01865|GCAM_MOU ... >sp|P01864|GCAB_MOUSE Ig gamma-2A chain C region secreted form OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P03977|KV3A4_MOUSE Ig kappa chain V-III region 50S10.1 OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P06330|HVM51_MOUSE Ig heavy chain V region AC38 205.12 OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P18528|HVM57_MOUSE Ig heavy chain V region 6.96 OS=Mus musculus OX=10090 PE=4 SV=1 >sp|P18529|HVM58_MOUSE Ig heavy chain V region 5-76 OS=Mus musculus OX=10090 PE=1 SV=1 >sp|P18531|HVM60_MOUSE Ig heavy chain V region 3-6 OS=Mus musculus OX=10090 GN=Ighv3-6 PE=1 SV=1 >sp|P84750|KVM5_MOUSE Ig kappa chain V region Mem5 (Fragment) OS=Mus musculus OX=10090 PE=1 SV=1 >sp|Q3UV17|K22O_MOUSE Keratin, type II cytoskeletal 2 oral OS=Mus musculus OX=10090 GN=Krt76 PE=1 SV=1 >tr|Q792Z1|Q792Z1_MOUSE MCG140784 OS=Mus musculus OX=10090 GN=Try10 PE=1 SV=1 >sp|Q9Z320|K1C27_MOUSE Keratin, type I cytoskeletal 27 OS=Mus musculus OX=10090 GN=Krt27 PE=1 SV=1

[info] 141/682 proteins matching your filters were removed from the dataset

[info] contrast: wt vs ko

[info] output directory does not exist yet, creating; /data/dataset_Klaassen2018_pmid26931375

[info] sample metadata contains samples with multiple fractions, 'sample_id' with the same 'shortname' are now merged by summation of their respective peptide intensities

[info] using 23 threads for multiprocessing

[progress] caching filter data took 1 seconds

[progress] peptide filtering and normalization took 1 seconds

[info] peptide to protein rollup strategy: maxlfq

[info] differential abundance analysis for contrast: wt vs ko

[info] using data from peptide filter: filter by contrast

[progress] peptide to protein rollup with MaxLFQ took 1 seconds

[info] log2 foldchange threshold estimated by bootstrap analysis: 0.939

[progress] DEqMS took 1 seconds

[progress] MS-EmpiRe took 2 seconds

[info] msqrob linear regression formulas (these are prioritized. eg; if a model fit fails due to lack of data, the next formula is used); expression ~ (1 | condition) + (1 | sample_id) + (1 | peptide_id) , expression ~ (1 | condition)

[progress] msqrob took 16 seconds

[info] differential detection analysis: min_samples_observed=2

[progress] creating PDF report…

[progress] report: constructing plots specific for each contrast

[progress] report: rendering report (this may take a while depending on dataset size)

[progress] RT plots: preparing data took 1 seconds

[progress] RT plots: creating plots took 1 seconds

[warning] No data available for CoV leave-one-out computation in sample group 'ko', skipping plots

[warning] No data available for CoV leave-one-out computation in sample group 'wt', skipping plots

5 R command history

This shows the history commands from your R script that starts this pipeline, thereby automatically documenting the parameters/settings used. All lines of executed code since (last) importing data using this R package are shown.

Using this feature

Do not use RStudio's `source` option to execute our pipeline since it will only write `source(...yourscript.R)` to the session history, and consequentially that is all you see in this 'code log'. Instead, select all lines in your script (`control + A`) and then "run" the selected code (either click the run button in RStudio, or use `control + enter`). All lines shown in this section are the same as shown in the RStudio 'History' pane (a tab on the top-right of its UI).

```
dataset = import_dataset_metamorpheus(
  path = getwd(),
  protein_qval_threshold = 0.05
)
dataset = import_fasta(dataset,
  files = c(
    "UP000000589_10090.fasta",
    "UP000000589_10090_additional.fasta"
  )
)
dataset = remove_proteins_by_name(dataset,
  regular_expression = "ig \\S+ chain|keratin|GN=(krt|try|igk|igg|igkv|ighv|ighg)"
)
dataset = import_sample_metadata(dataset,
  filename = "sample_metadata.xlsx"
)
dataset = setup_contrasts(dataset,
  contrast_list = list(c(
    "wt",
    "ko"
  ))
)
dataset = analysis_quickstart(dataset,
  filter_min_detect = 1,
  filter_min_quant = 3,
  filter_by_contrast = TRUE,
  norm_algorithm = c(
    "vwmb",
    "modebetween_protein"
  ),
  dea_algorithm = c(
    "deqms",
    "msempire", "msqrob"
  ),
  dea_qvalue_threshold = 0.05,
  dea_log2foldchange_threshold = NA,
  diffdetect_min_samples_observed = 2,
  output_qc_report = TRUE,
  output_dir = "/data/dataset_Klaassen2018_pmid26931375",
  output_within_timestamped_subdirectory = TRUE
)
```

6 R session info

The computer system and versioning of all R packages used to run this analysis are shown below to facilitate, in combination with the previous section, reproducibility.

setting	value
version	R version 3.6.3 (2020-02-29)
os	Debian GNU/Linux 10 (buster)
system	x86_64, linux-gnu
ui	RStudio
language	(EN)
collate	en_US.UTF-8
ctype	en_US.UTF-8
tz	Etc/UTC
date	2020-12-13

System

package	loadedversion	source
dplyr	0.8.5	CRAN (R 3.6.3)
ggplot2	3.3.0	CRAN (R 3.6.3)
msdap	0.2.7.1	Github (ftwkoopmans/msdap@67ec44e)
rlang	0.4.5	CRAN (R 3.6.3)
tibble	3.0.1	CRAN (R 3.6.3)
tidyr	1.0.2	CRAN (R 3.6.3)

Attached packages

package	loadedversion	source
affy	1.64.0	Bioconductor
affyio	1.56.0	Bioconductor
arrangements	1.1.8	CRAN (R 3.6.3)
askpass	1.1	CRAN (R 3.6.3)
assertthat	0.2.1	CRAN (R 3.6.3)
backports	1.1.6	CRAN (R 3.6.3)
Biobase	2.46.0	Bioconductor
BiocGenerics	0.32.0	Bioconductor
BiocManager	1.30.10	CRAN (R 3.6.3)
BiocParallel	1.20.1	Bioconductor
bit	1.1-15.2	CRAN (R 3.6.3)
bit64	0.9-7	CRAN (R 3.6.3)
blob	1.2.1	CRAN (R 3.6.3)
boot	1.3-24	CRAN (R 3.6.3)
callr	3.4.3	CRAN (R 3.6.3)
cli	2.0.2	CRAN (R 3.6.3)
codetools	0.2-16	CRAN (R 3.6.3)
colorspace	1.4-1	CRAN (R 3.6.3)
cowplot	1.0.0	CRAN (R 3.6.3)
crayon	1.3.4	CRAN (R 3.6.3)
data.table	1.12.8	CRAN (R 3.6.3)
DBI	1.1.0	CRAN (R 3.6.3)
DEqMS	1.4.0	Bioconductor
desc	1.2.0	CRAN (R 3.6.3)
devtools	2.3.0	CRAN (R 3.6.3)
diann	1.0.1	Github (vdemichev/diann-rpackage@af538f6)
digest	0.6.25	CRAN (R 3.6.3)
doParallel	1.0.16	CRAN (R 3.6.3)
ellipsis	0.3.0	CRAN (R 3.6.3)
evaluate	0.14	CRAN (R 3.6.3)
fansi	0.4.1	CRAN (R 3.6.3)
farver	2.0.3	CRAN (R 3.6.3)
foreach	1.5.1	CRAN (R 3.6.3)
formatR	1.7	CRAN (R 3.6.3)
fs	1.4.1	CRAN (R 3.6.3)
ggpubr	0.2.5	CRAN (R 3.6.3)
ggrepel	0.8.2	CRAN (R 3.6.3)
ggsignif	0.6.0	CRAN (R 3.6.3)
glue	1.4.0	CRAN (R 3.6.3)
gmp	0.5-13.6	CRAN (R 3.6.3)
gridExtra	2.3	CRAN (R 3.6.3)
gtable	0.3.0	CRAN (R 3.6.3)
gtools	3.8.2	CRAN (R 3.6.3)
hms	0.5.3	CRAN (R 3.6.3)
htmltools	0.4.0	CRAN (R 3.6.3)
impute	1.60.0	Bioconductor
IRanges	2.20.2	Bioconductor
iterators	1.0.13	CRAN (R 3.6.3)
knitr	1.28	CRAN (R 3.6.3)
labeling	0.3	CRAN (R 3.6.3)

package	loadedversion	source
lattice	0.20-38	CRAN (R 3.6.3)
lifecycle	0.2.0	CRAN (R 3.6.3)
limma	3.42.2	Bioconductor
lme4	1.1-23	CRAN (R 3.6.3)
magrittr	1.5	CRAN (R 3.6.3)
MALDIquant	1.19.3	CRAN (R 3.6.3)
MASS	7.3-51.5	CRAN (R 3.6.3)
Matrix	1.2-18	CRAN (R 3.6.3)
matrixStats	0.56.0	CRAN (R 3.6.3)
memoise	1.1.0	CRAN (R 3.6.3)
minqa	1.2.4	CRAN (R 3.6.3)
msEmpiRe	0.1.0	Github (zimmerlab/MS-EmpiRe@8a85757)
MSnbase	2.12.0	Bioconductor
munsell	0.5.0	CRAN (R 3.6.3)
mzID	1.24.0	Bioconductor
mzR	2.20.0	Bioconductor
ncdf4	1.17	CRAN (R 3.6.3)
nlme	3.1-144	CRAN (R 3.6.3)
nloptr	1.2.2.1	CRAN (R 3.6.3)
openxlsx	4.1.4	CRAN (R 3.6.3)
patchwork	1.0.0	CRAN (R 3.6.3)
pcaMethods	1.78.0	Bioconductor
pdftools	2.3	CRAN (R 3.6.3)
pillar	1.4.3	CRAN (R 3.6.3)
pkgbuild	1.0.6	CRAN (R 3.6.3)
pkgconfig	2.0.3	CRAN (R 3.6.3)
pkgload	1.0.2	CRAN (R 3.6.3)
plyr	1.8.6	CRAN (R 3.6.3)
preprocessCore	1.48.0	Bioconductor
prettyunits	1.1.1	CRAN (R 3.6.3)
pROC	1.16.2	CRAN (R 3.6.3)
processx	3.4.2	CRAN (R 3.6.3)
ProtGenerics	1.18.0	Bioconductor
ps	1.3.2	CRAN (R 3.6.3)
purrr	0.3.4	CRAN (R 3.6.3)
qpdf	1.1	CRAN (R 3.6.3)
R6	2.4.1	CRAN (R 3.6.3)
RColorBrewer	1.1-2	CRAN (R 3.6.3)
Rcpp	1.0.4.6	CRAN (R 3.6.3)
RcppEigen	0.3.3.7.0	CRAN (R 3.6.3)
readr	1.3.1	CRAN (R 3.6.3)
remotes	2.1.1	CRAN (R 3.6.3)
rmarkdown	2.1	CRAN (R 3.6.3)
rprojroot	1.3-2	CRAN (R 3.6.3)
RSQLite	2.2.0	CRAN (R 3.6.3)
rstudioapi	0.11	CRAN (R 3.6.3)
S4Vectors	0.24.4	Bioconductor
scales	1.1.0	CRAN (R 3.6.3)
sessioninfo	1.1.1	CRAN (R 3.6.3)
statmod	1.4.34	CRAN (R 3.6.3)

package	loadedversion	source
stringi	1.4.6	CRAN (R 3.6.3)
stringr	1.4.0	CRAN (R 3.6.3)
styler	1.3.2	CRAN (R 3.6.3)
testthat	2.3.2	CRAN (R 3.6.3)
tidyselect	1.0.0	CRAN (R 3.6.3)
usethis	1.6.0	CRAN (R 3.6.3)
vctrs	0.2.4	CRAN (R 3.6.3)
viridis	0.5.1	CRAN (R 3.6.3)
viridisLite	0.3.0	CRAN (R 3.6.3)
vsn	3.54.0	Bioconductor
withr	2.2.0	CRAN (R 3.6.3)
xfun	0.13	CRAN (R 3.6.3)
XML	3.99-0.3	CRAN (R 3.6.3)
xtable	1.8-4	CRAN (R 3.6.3)
yaml	2.2.1	CRAN (R 3.6.3)
zip	2.0.4	CRAN (R 3.6.3)
zlibbioc	1.32.0	Bioconductor

Packages that are not attached