

MS-DAP: Mass Spectrometry Downstream Analysis Pipeline

version: 1.0.3 <https://github.com/ftwkoopmans/msdap/>

Contents

1 Quality control	2
1.1 DIA confidence score distributions	2
1.2 number of peptides and proteins	26
1.3 data completeness	34
1.4 abundance distributions	36
1.5 retention time	45
1.6 variation among replicates	152
1.7 PCA	178
2 Differential abundance analysis	194
2.1 control_Berlin,control_Magdeburg... vs AD_Berlin,AD_Magdeburg,AD_Sweden	194
3 Summary of differential testing	200
4 log	201
5 R command history	202
6 R session info	203

1 Quality control

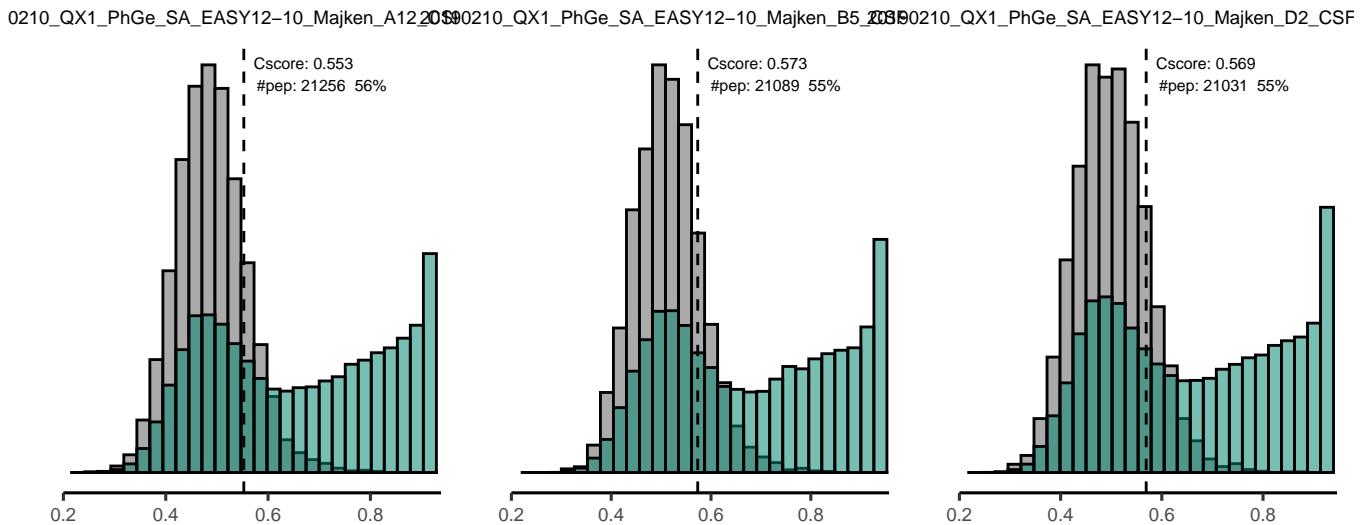
The quality control figures in this section enable you to investigate reproducibility and global clustering of samples by visualizing:

- 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

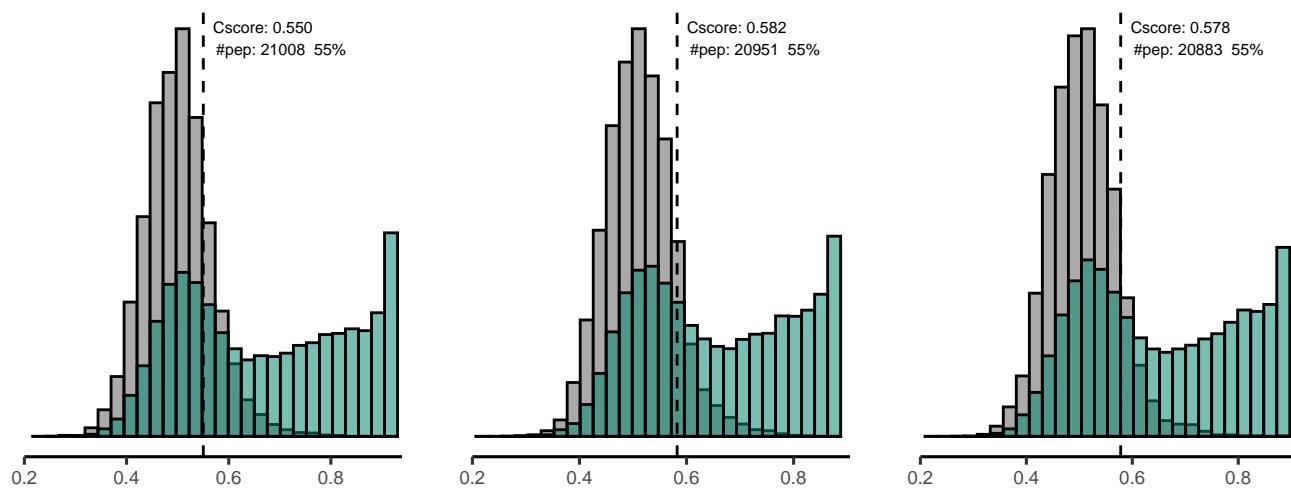
The first set of quality control figures describes individual samples, thereafter group-level quality metrics are described and finally sample clustering is used to highlight structure in the entire dataset.

1.1 DIA confidence score distributions

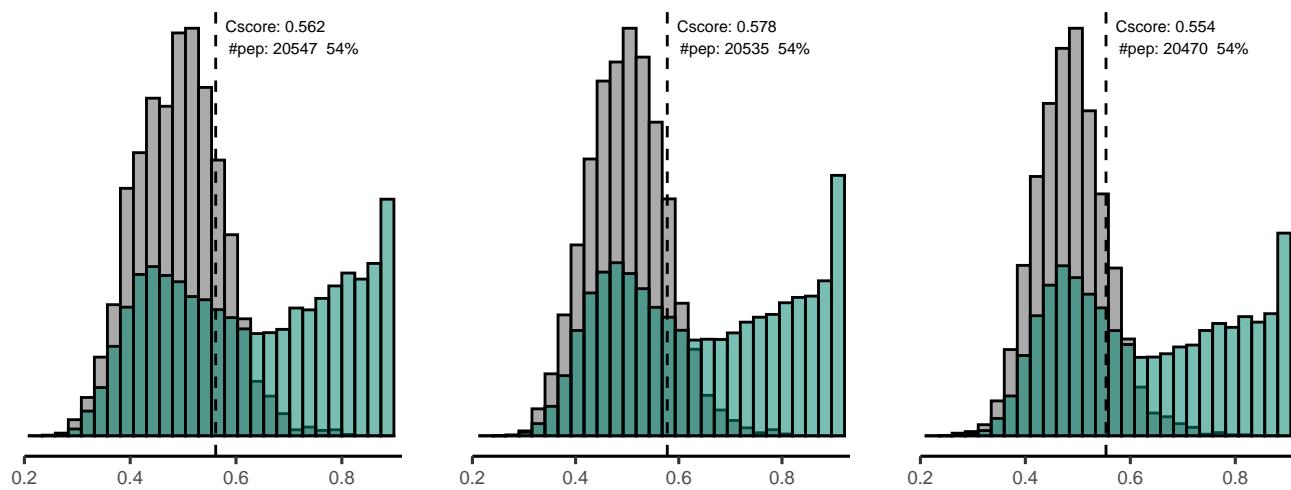
DIA data was used as input for this analysis. Below histograms visualize both target (green) and decoy (grey) cscores in each sample, indicating how confident the input software was in the identification of peptides from the spectral library in the raw DIA data. Samples are ordered by the number of precursors quantified at q-value confidence threshold 0.01. At this threshold, the respective cscore and number of peptides is shown.



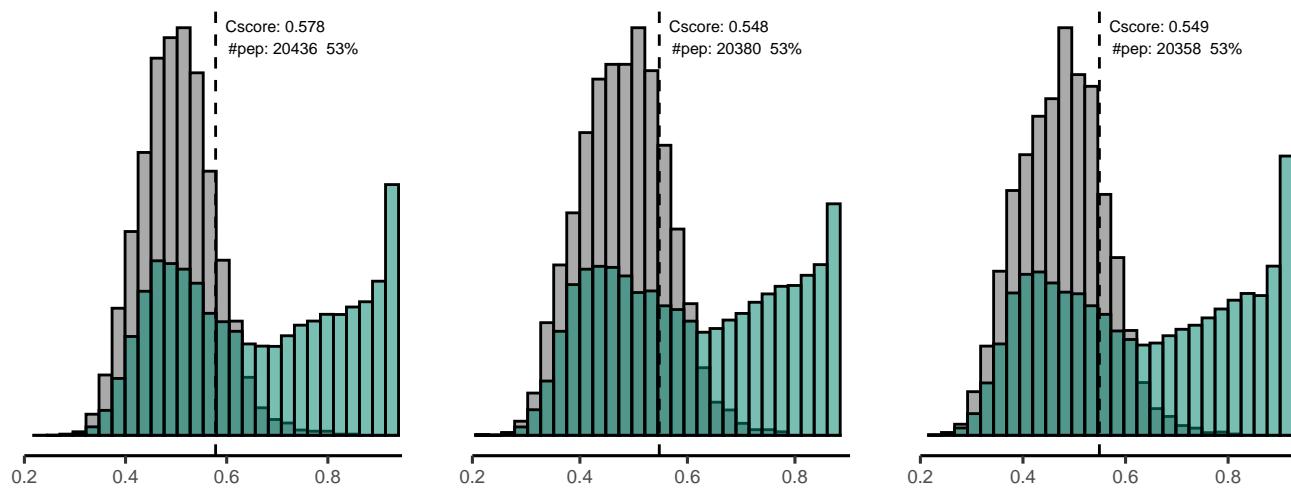
_QX1_PhGe_SA_EASY12-10_Deigendesch_D6_129_QX1_PhGe_SA_EASY12-10_JaBa_SA_LC12_5_CSF1_1_8-1xD12910618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



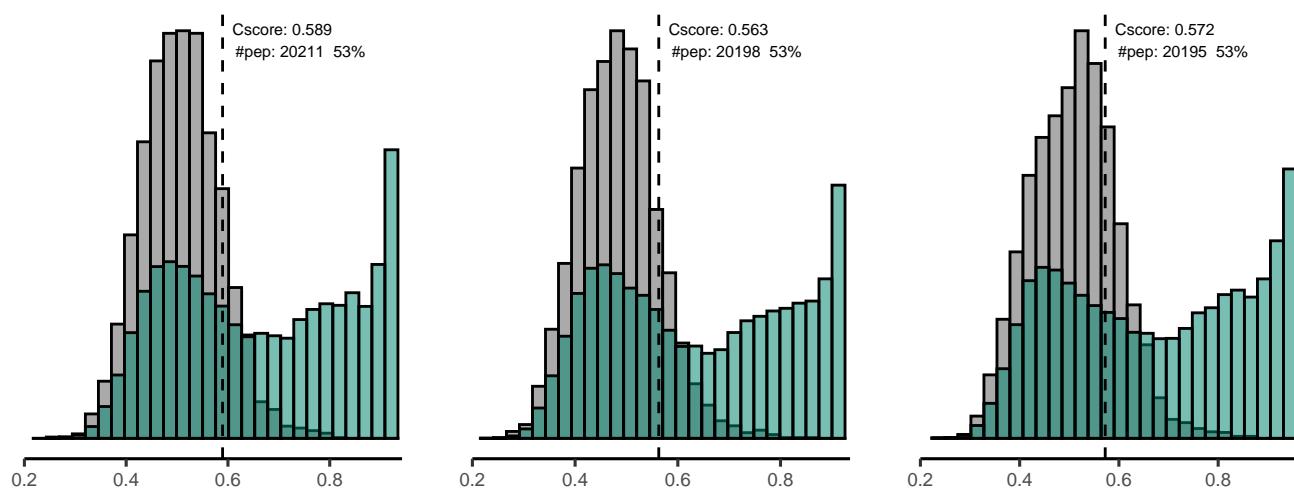
)20210_QX1_PhGe_SA_EASY12-10_Majken_A2_20590210_QX1_PhGe_SA_EASY12-10_Majken20590210_QX1_PhGe_SA_EASY12-10_Deigendesch_D6_129



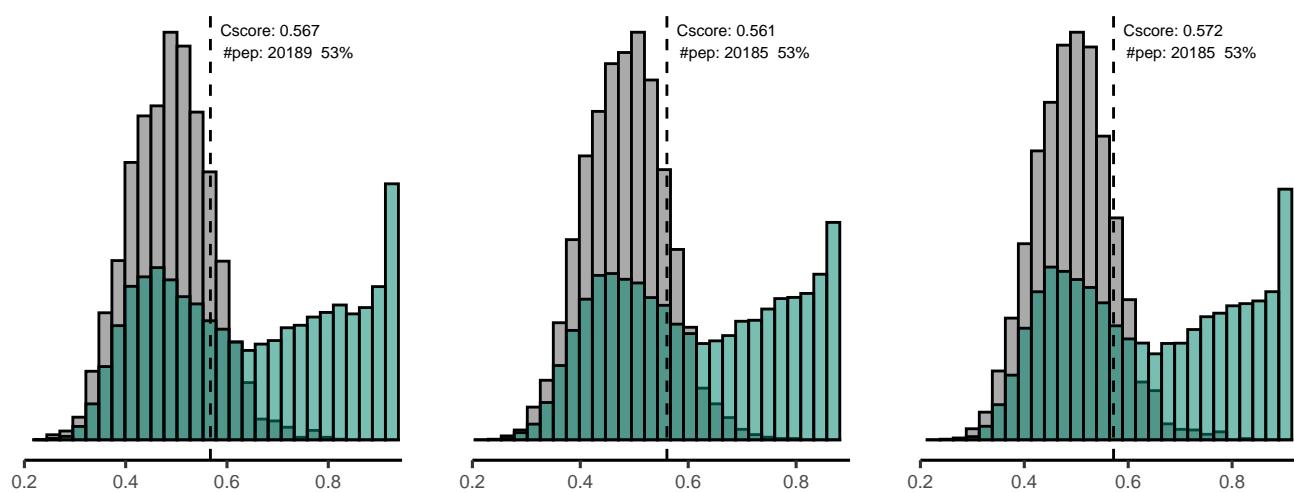
)20210_QX1_PhGe_SA_EASY12-10_Majken_C10_20590210_QX1_PhGe_SA_EASY12-10_Majken_B7_20590210_QX1_PhGe_SA_EASY12-10_Majken_C2_CSF.



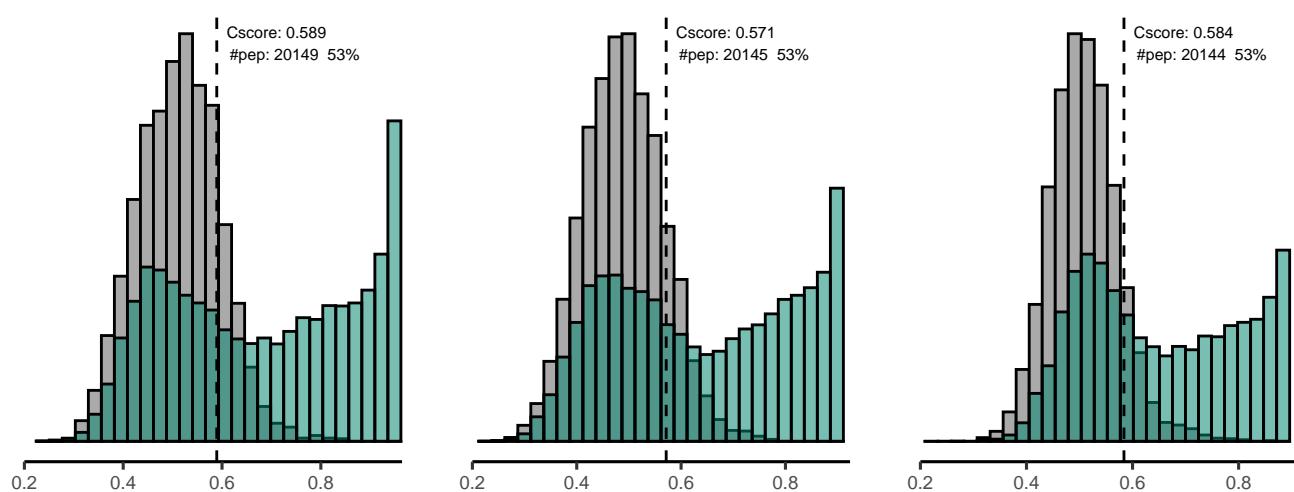
_QX1_PhGe_SA_EASY12-10_Deigendesch120190210_QX1_PhGe_SA_EASY12-10_Deigendesch_G120190210_QX1_PhGe_SA_EASY12-10_Majken_D1_CSF



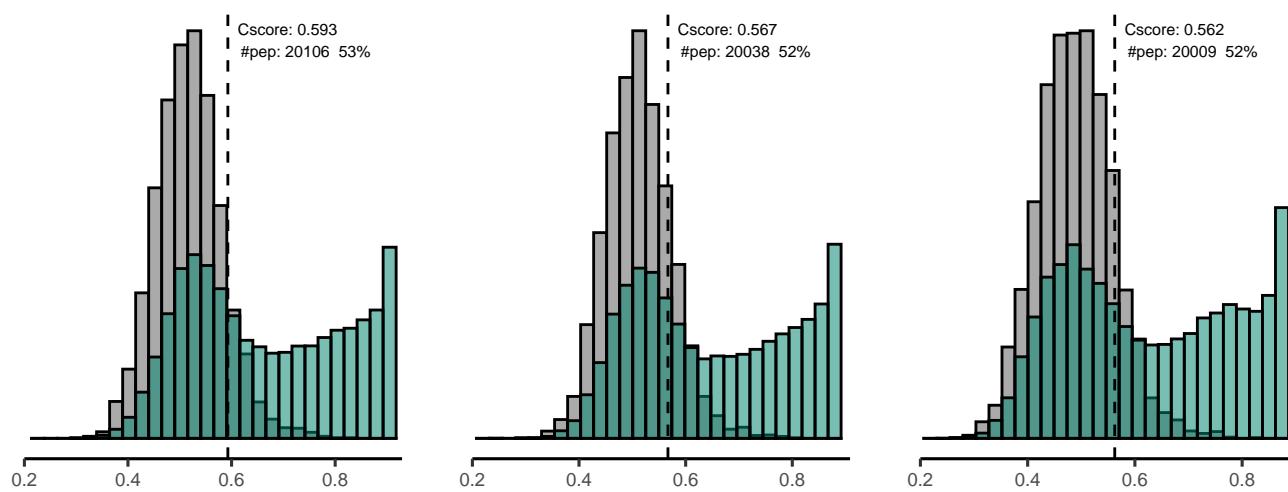
_QX1_PhGe_SA_EASY12-10_Majken_B8_20190210_QX1_PhGe_SA_EASY12-10_Majken_20190210_QX1_PhGe_SA_EASY12-10_Deigendesch_D11_129



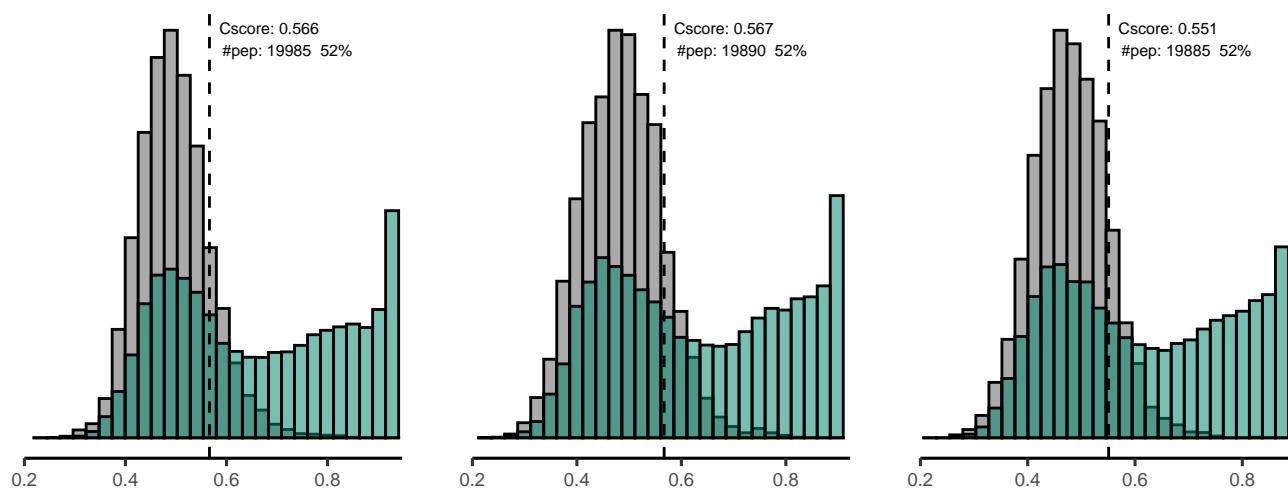
_QX1_PhGe_SA_EASY12-10_Majken_A3_20190210_QX1_PhGe_SA_EASY12-10_Majken_A5_20190218_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fm1.



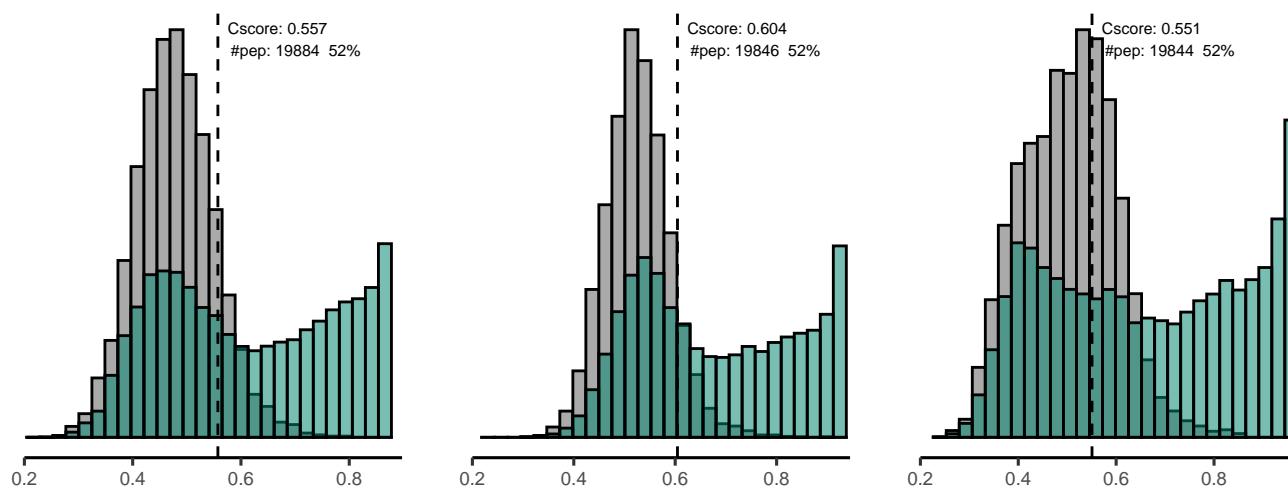
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1028_QX1_PhGe_SA_EASY12-10_Deigendesch_C01028_QX1_PhGe_SA_EASY12-10_Deigendesch_D4_



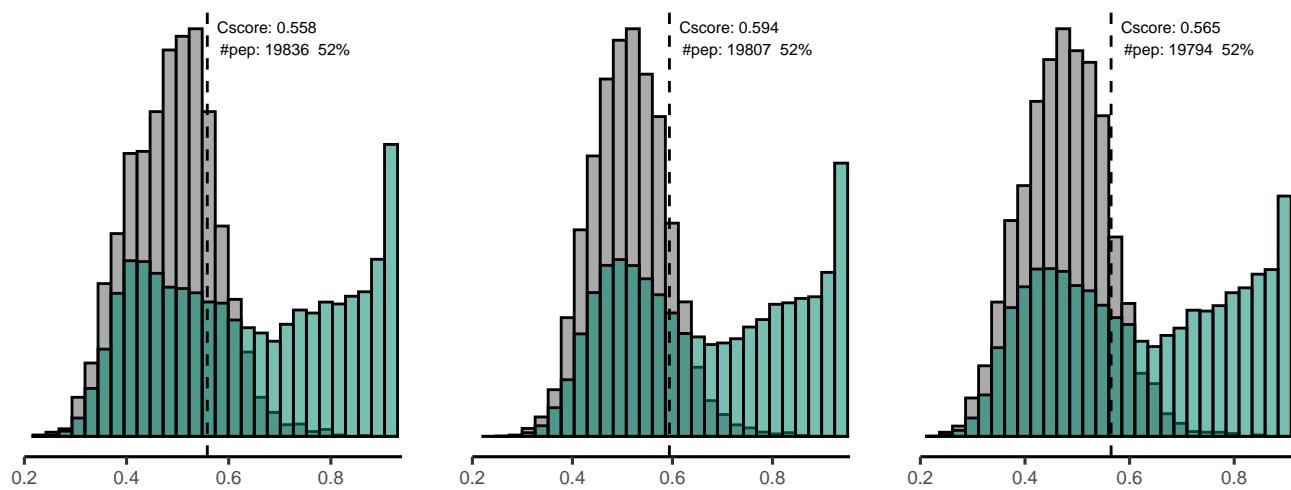
_QX1_PhGe_SA_EASY12-10_Deigendesch_201028_QX1_PhGe_SA_EASY12-10_Deigendesch_A82090210_QX1_PhGe_SA_EASY12-10_Majken_B4_CSF



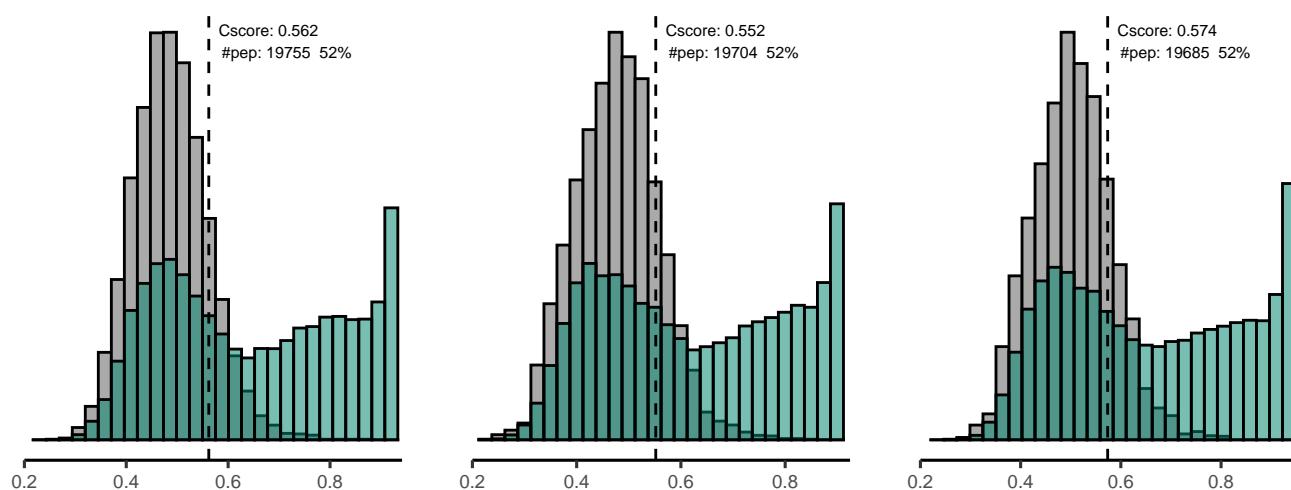
_0210_QX1_PhGe_SA_EASY12-10_Majken_A9168718_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1028_QX1_PhGe_SA_EASY12-10_Majken_D6_CSF



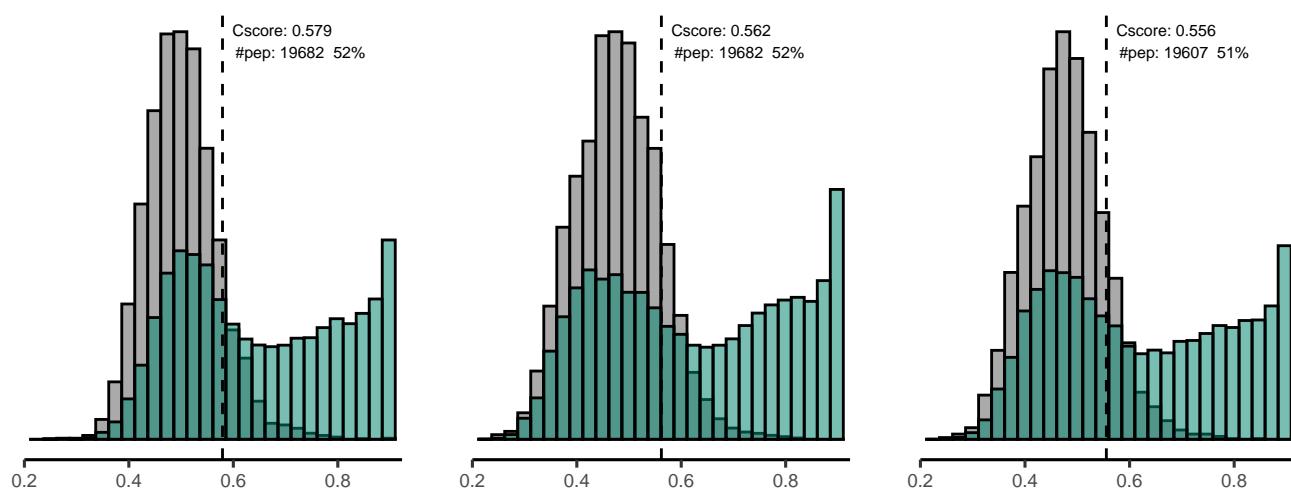
1228_QX1_PhGe_SA_EASY12-10_Deigendesch_20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_20190210_QX1_PhGe_SA_EASY12-10_Majken_D3_CSF



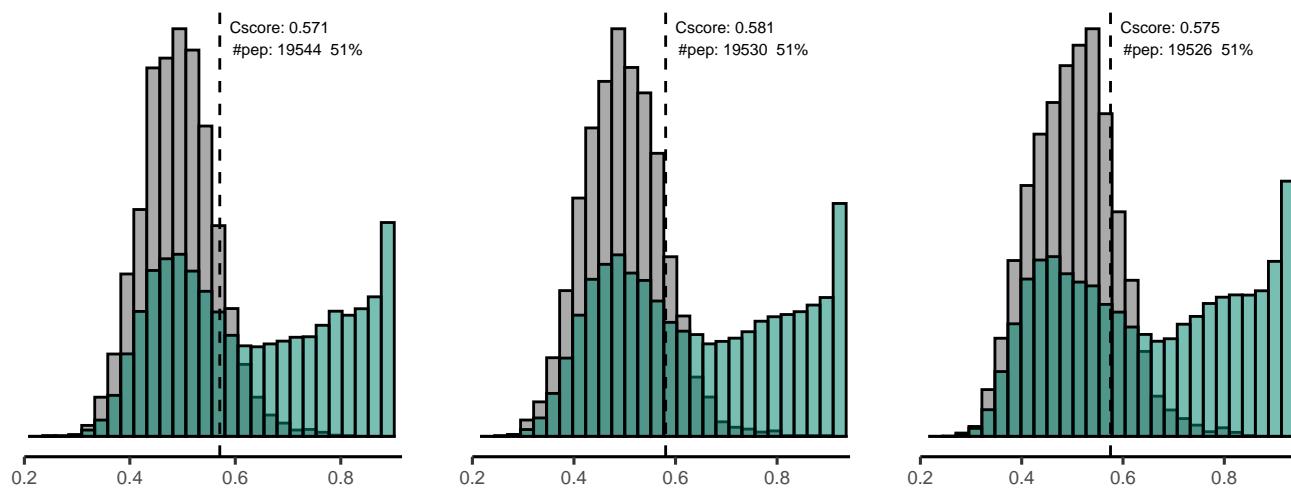
3_QX1_PhGe_SA_EASY12-10_Deigendesch_D920190210_QX1_PhGe_SA_EASY12-10_Majken_B12_QX1_PhGe_SA_EASY12-10_Deigendesch_B3



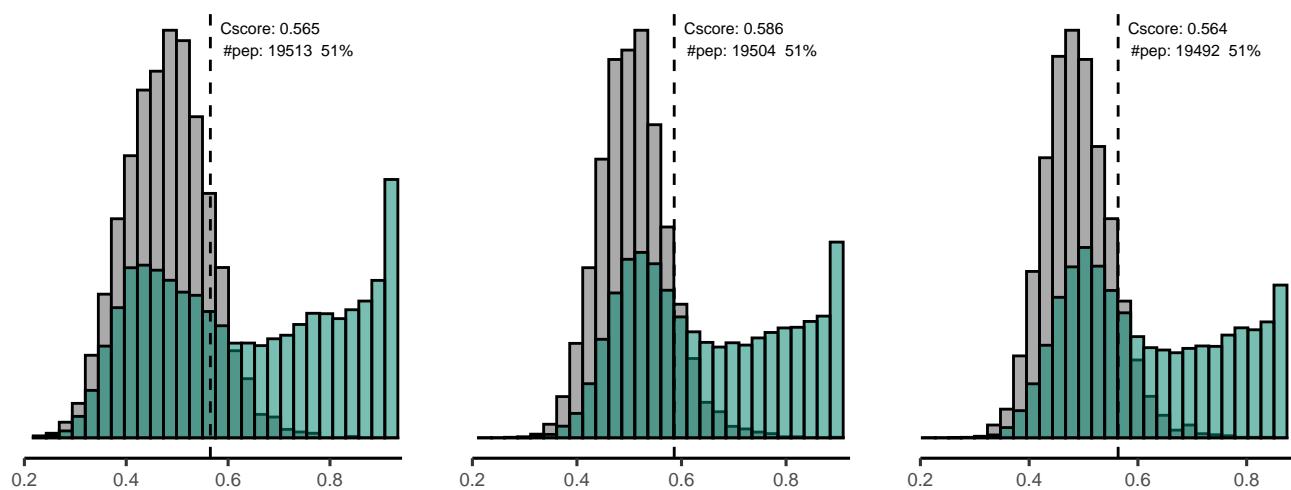
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1x20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_E5_20190210_QX1_PhGe_SA_EASY12-10_Majken_C1_CSF



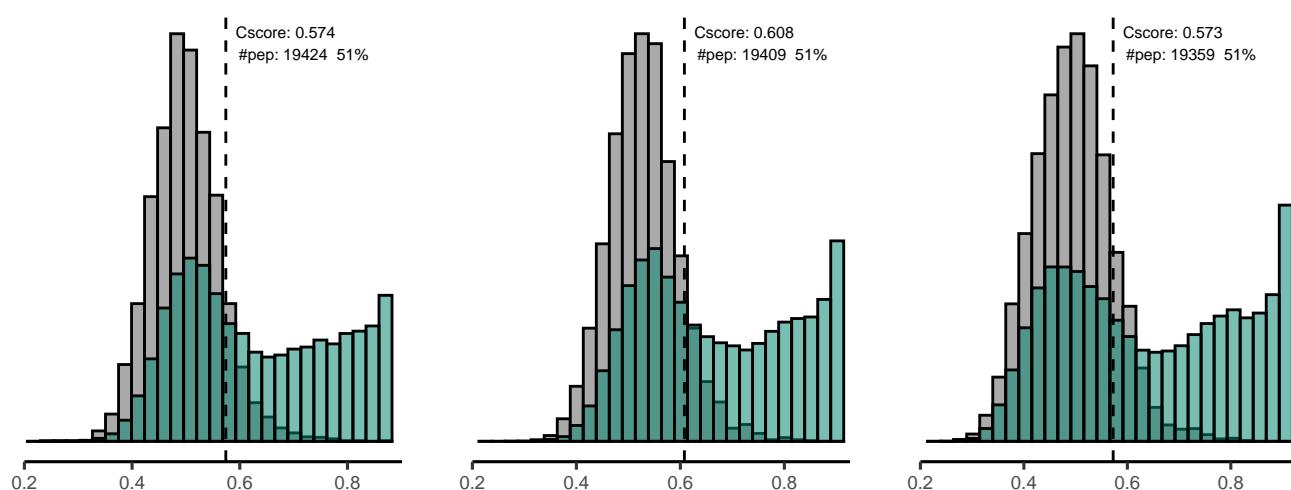
228_QX1_PhGe_SA_EASY12-10_Deigendesch_F20190210_QX1_PhGe_SA_EASY12-10_20190210_QX1_PhGe_SA_EASY12-10_Majken_A7_CSF_531-12_c



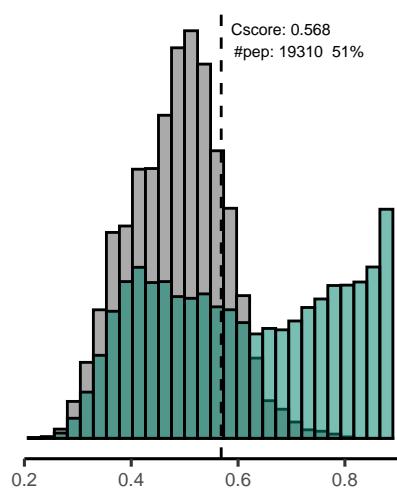
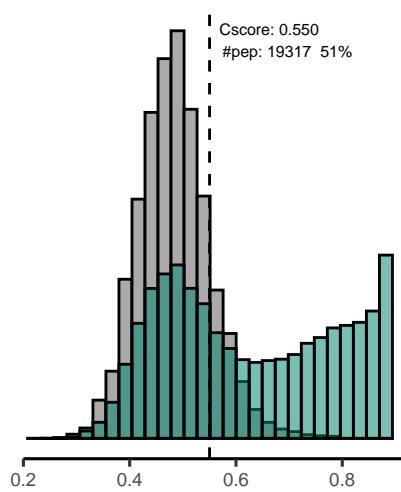
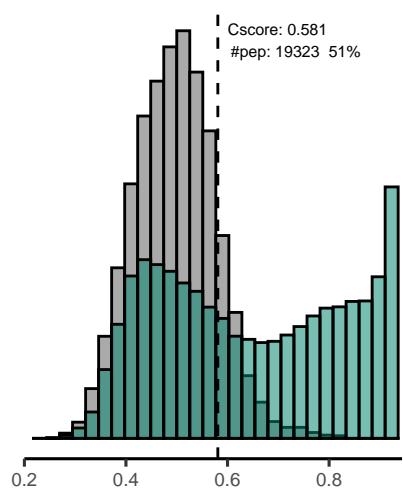
190210_QX1_PhGe_SA_EASY12-10_Majken_20190210_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



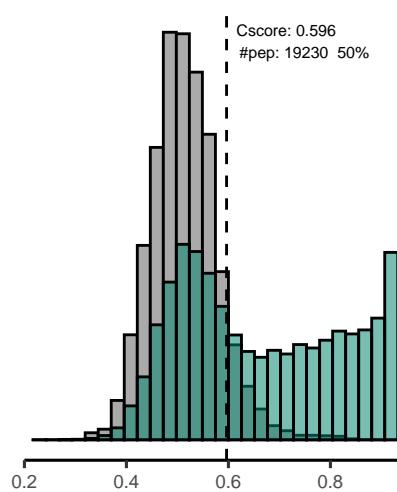
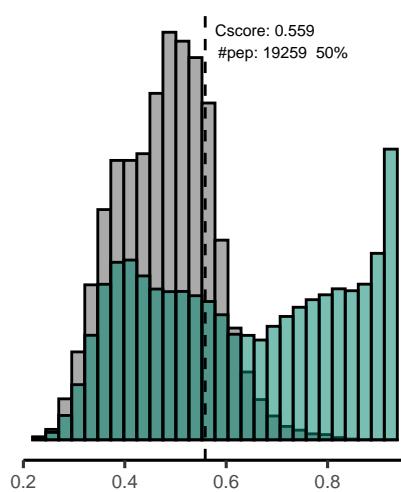
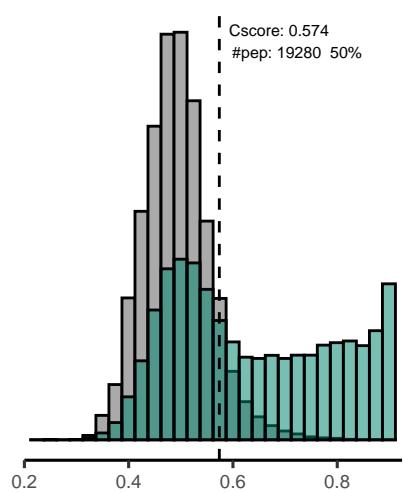
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1_20190218_QX1_PhGe_SA_EASY12-10_Deigendesch_F8_190:



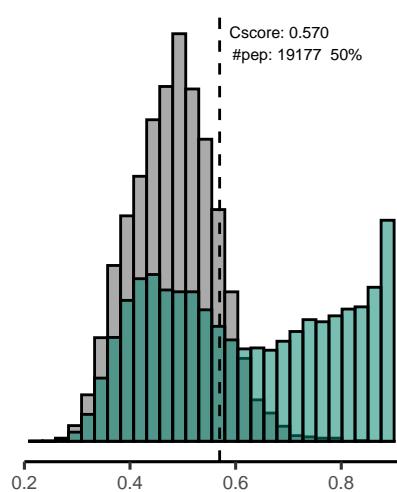
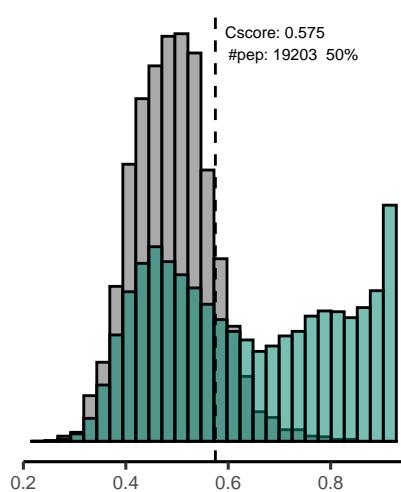
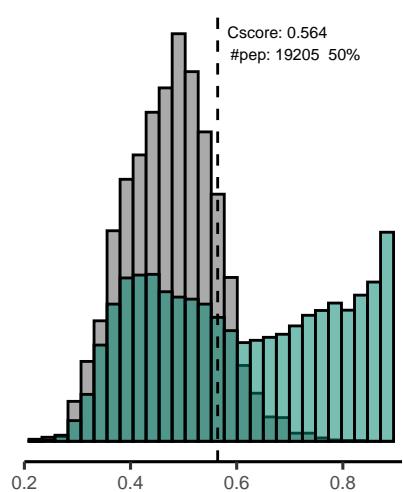
0210_QX1_PhGe_SA_EASY12-10_Majken_C2000



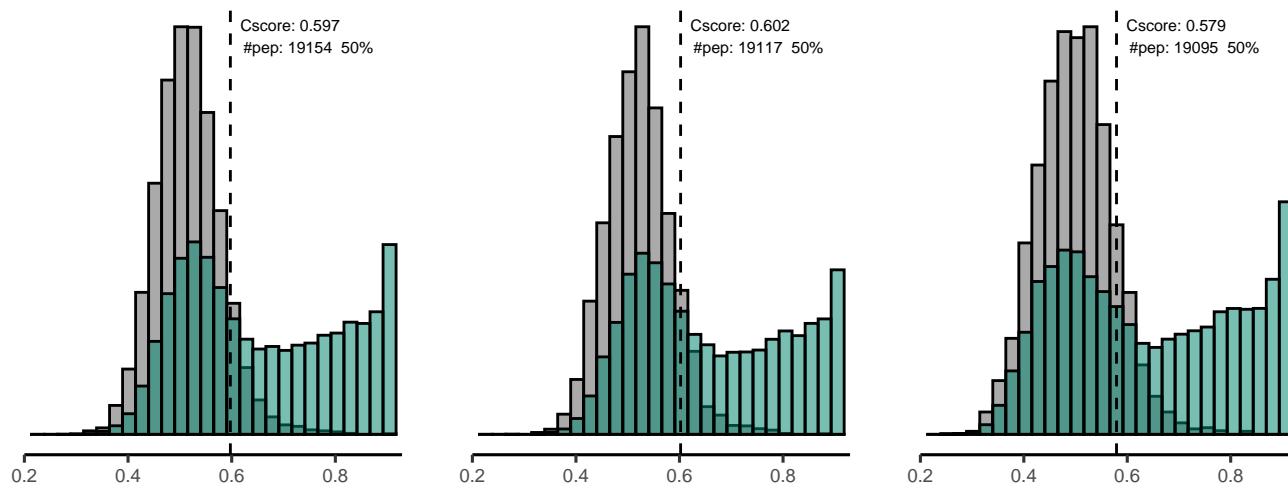
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1_5



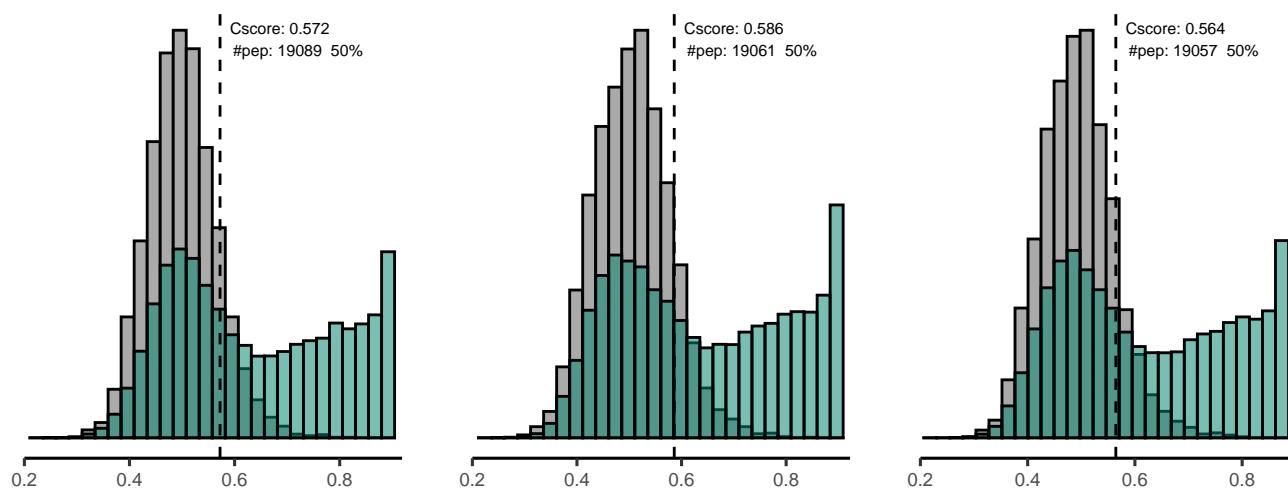
0210_QX1_PhGe_SA_EASY12-10_Majken_C62000



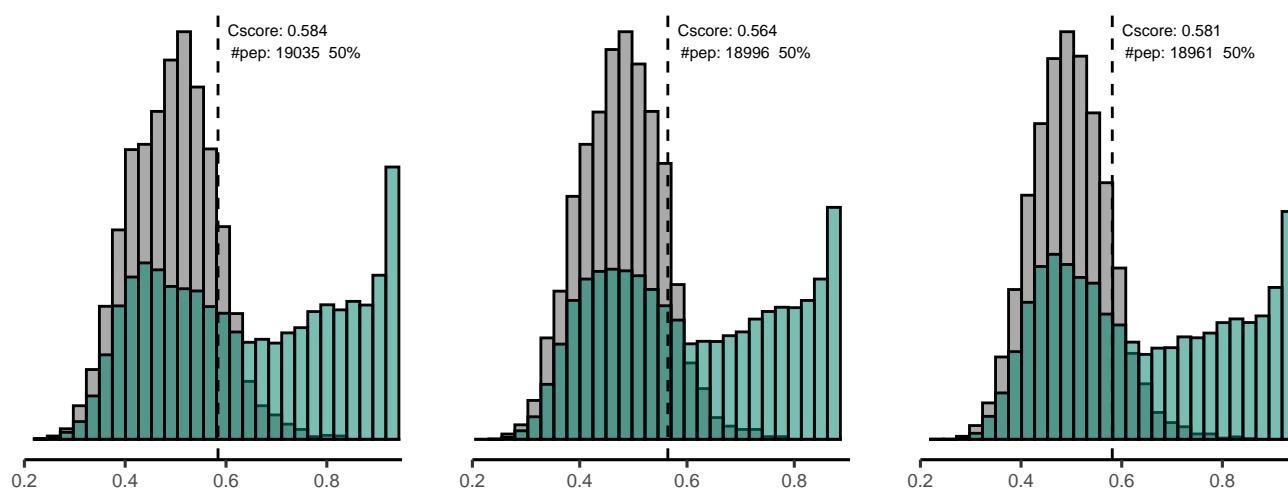
3_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1x**20100618**_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1x**2010228**_QX1_PhGe_SA_EASY12-10_Deigendesch_A11_



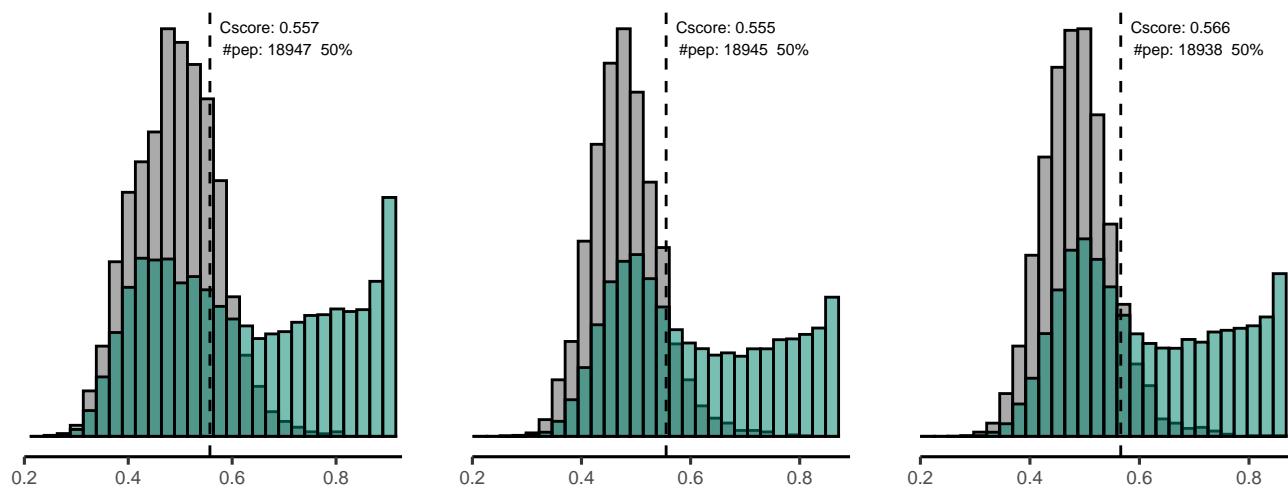
?28_QX1_PhGe_SA_EASY12-10_Deigendesch**2010228**_QX1_PhGe_SA_EASY12-10_Deigendesch_A**2010228**_QX1_PhGe_SA_EASY12-10_Deigendesch_H9_



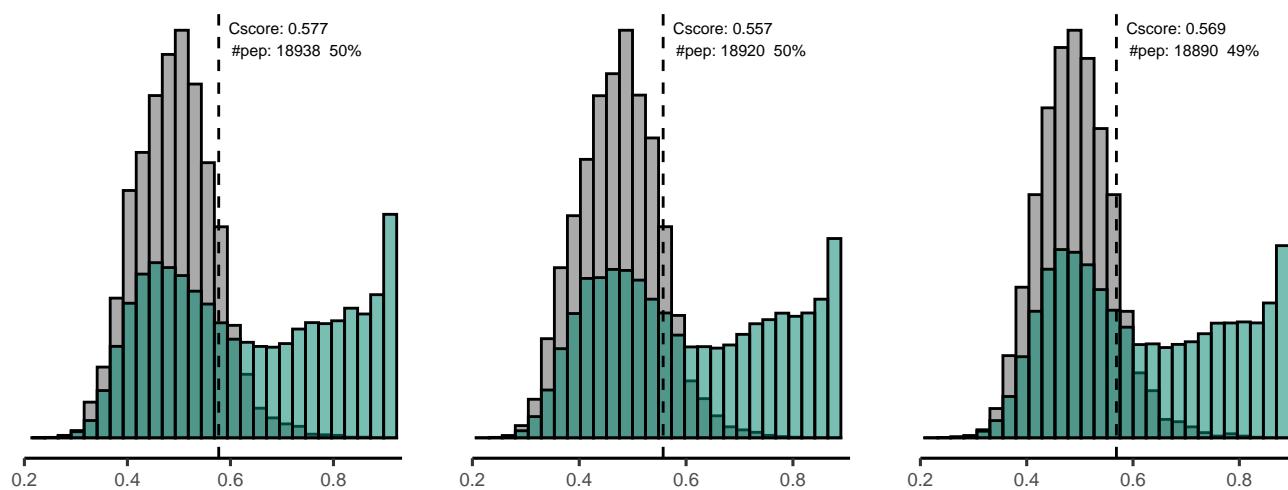
228_QX1_PhGe_SA_EASY12-10_Deigendesch**2010228**_QX1_PhGe_SA_EASY12-10_Deigendesch**2010210**_QX1_PhGe_SA_EASY12-10_Majken_B2_CSF



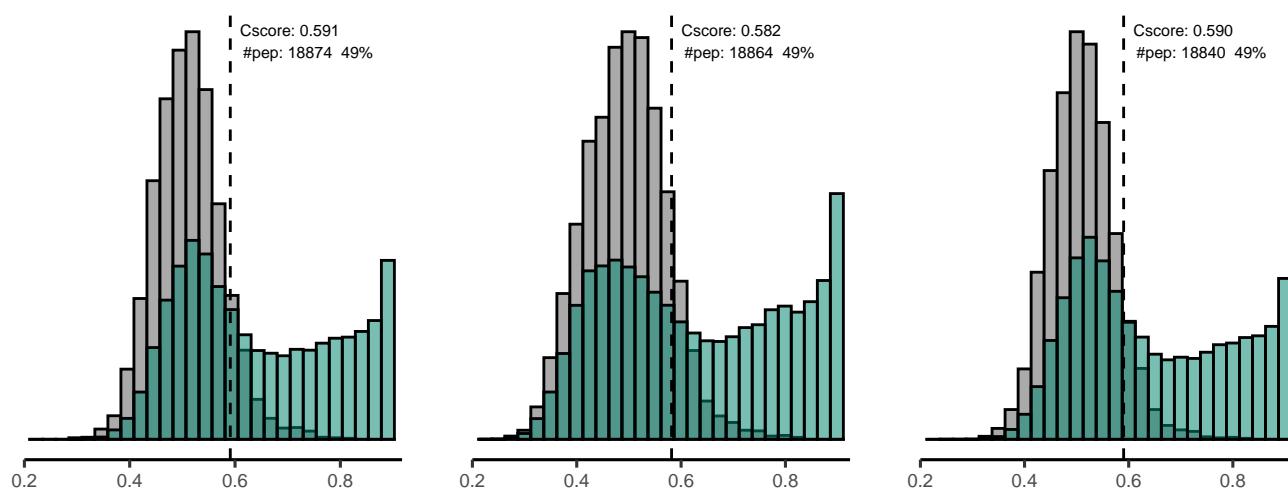
!28_QX1_PhGe_SA_EASY12-10_Deigendes20100618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fm1.



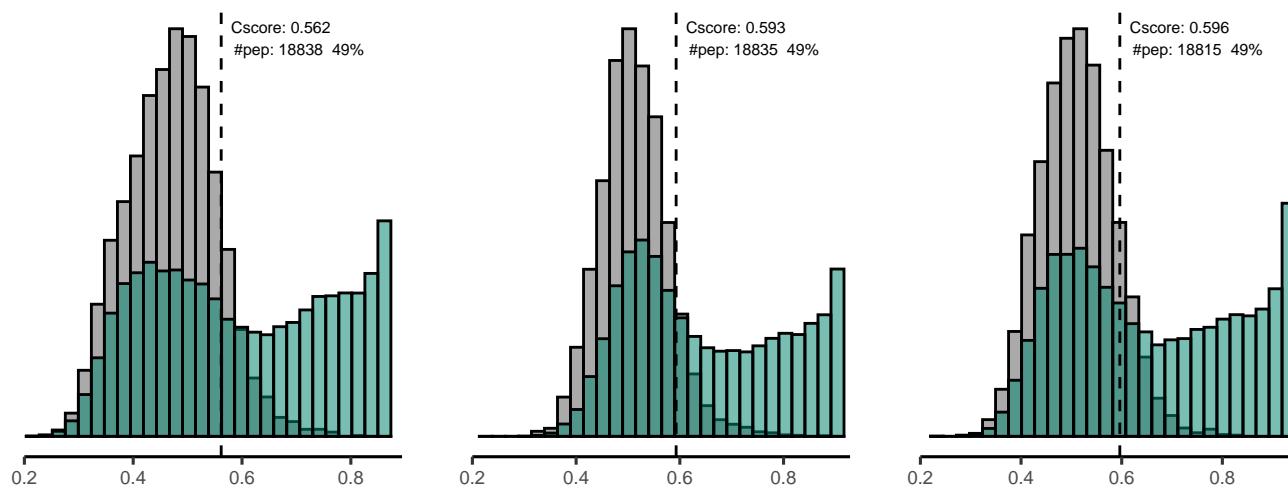
_QX1_PhGe_SA_EASY12-10_Deigendesch_D2010228_QX1_PhGe_SA_EASY12-10_Deigendesch2010228_QX1_PhGe_SA_EASY12-10_Deigendesch_G12_



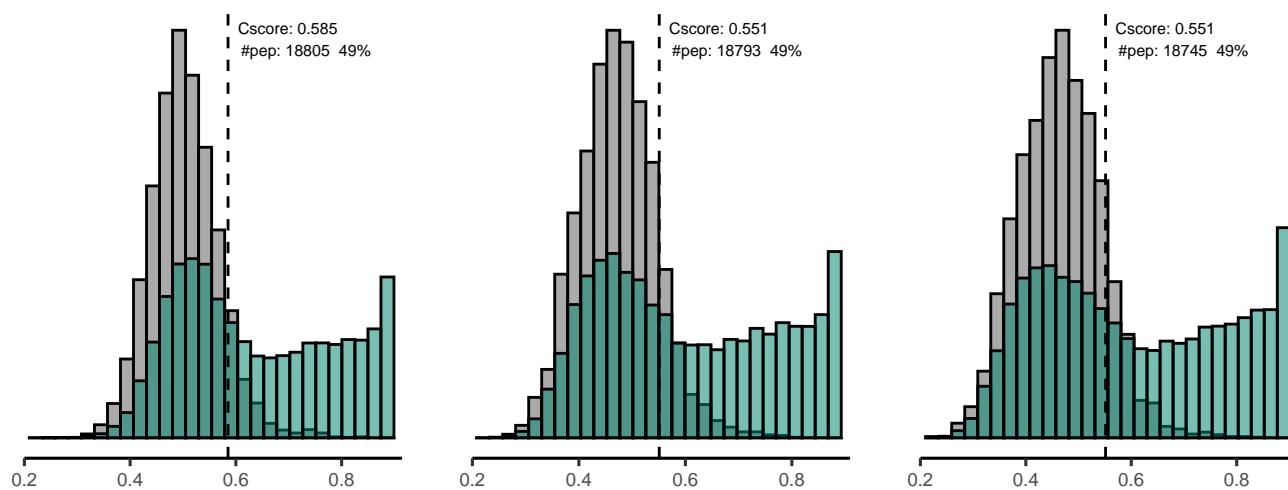
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fm1.20100228_QX1_PhGe_SA_EASY12-10_Deigendesch20100618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fm1.



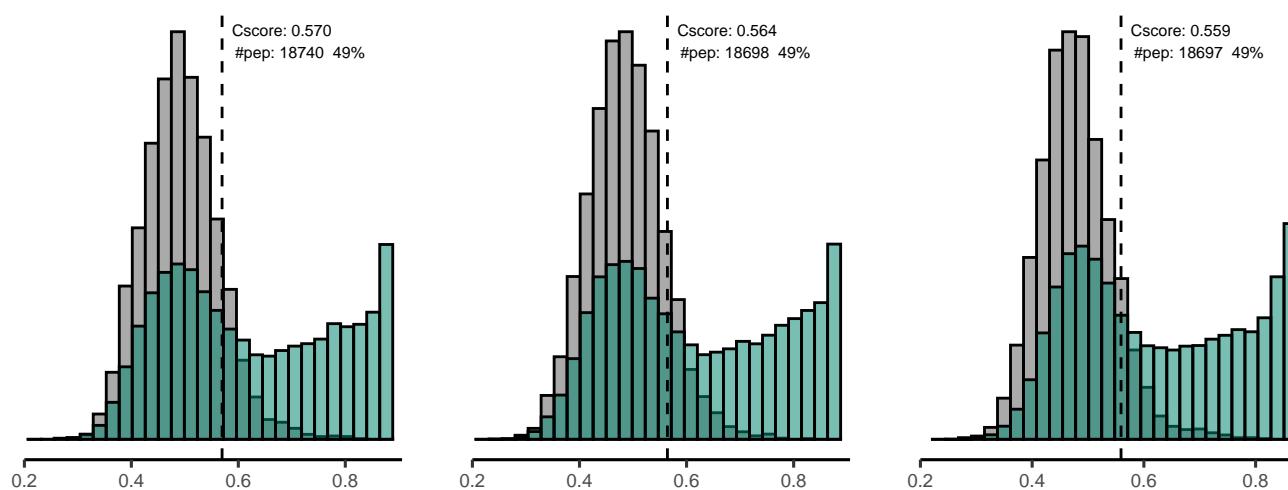
0210_QX1_PhGe_SA_EASY12-10_Majken_C6168718_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fm1 0228_QX1_PhGe_SA_EASY12-10_Deigendesch_B2_



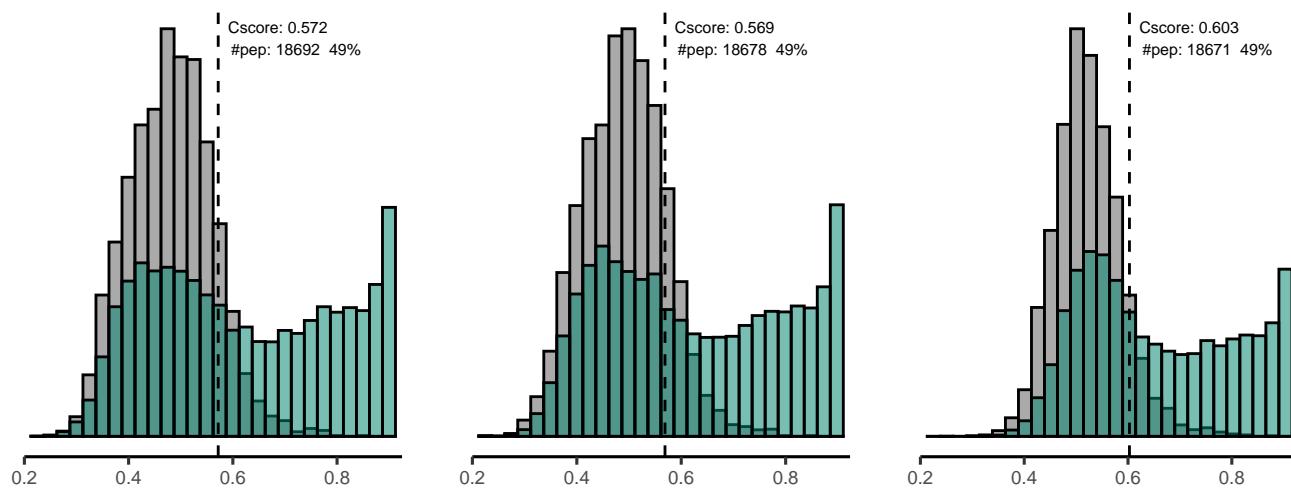
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fm1 0228_QX1_PhGe_SA_EASY12-10_Deigendesch_A7_20190210_QX1_PhGe_SA_EASY12-10_Majken_E12_C



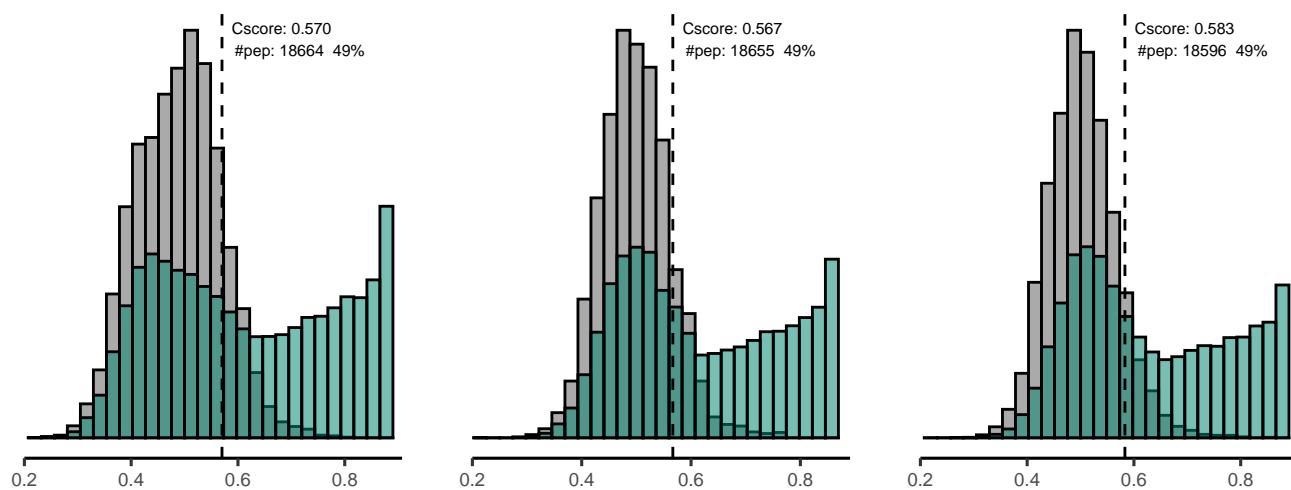
228_QX1_PhGe_SA_EASY12-10_Deigendesch_A4 20190210_QX1_PhGe_SA_EASY12-10_Majken_C6168718_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fm1.



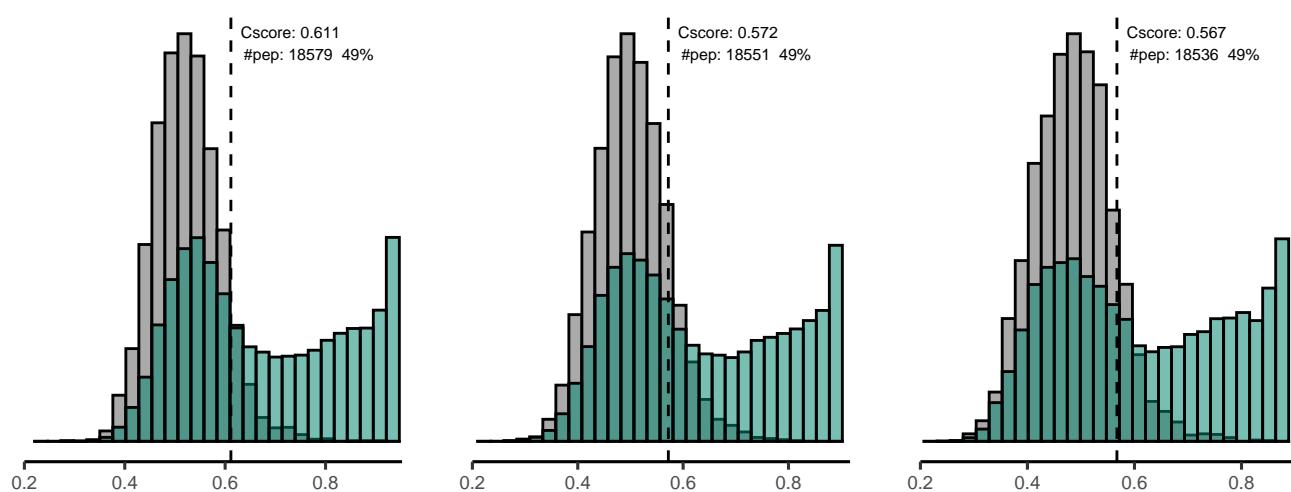
1228_QX1_PhGe_SA_EASY12-10_Deigendes20106228_QX1_PhGe_SA_EASY12-10_Deigendes2010618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



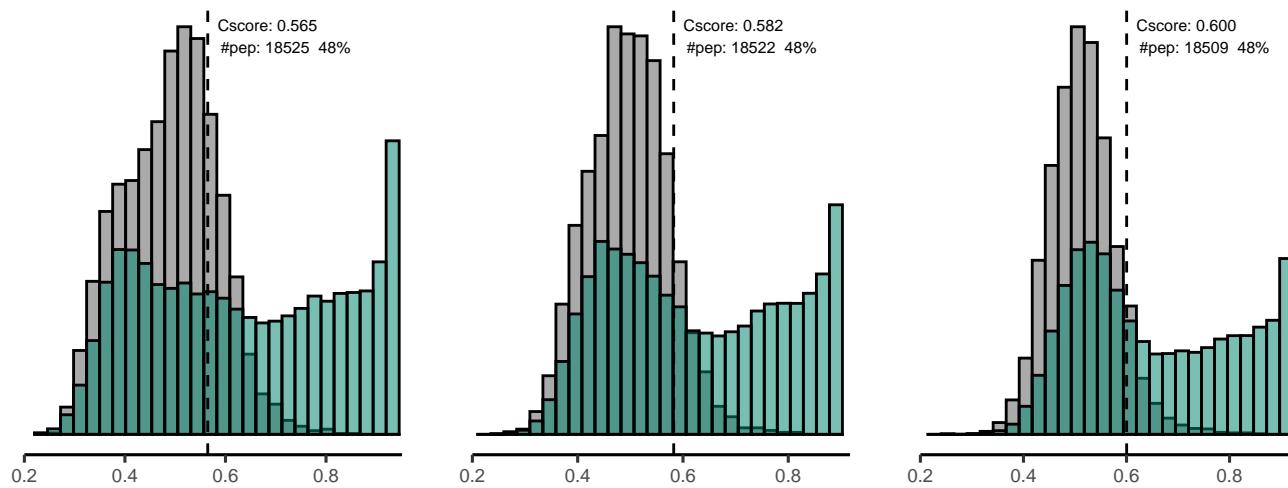
1228_QX1_PhGe_SA_EASY12-10_Deigendes20106228_QX1_PhGe_SA_EASY12-10_Deigendes2010618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



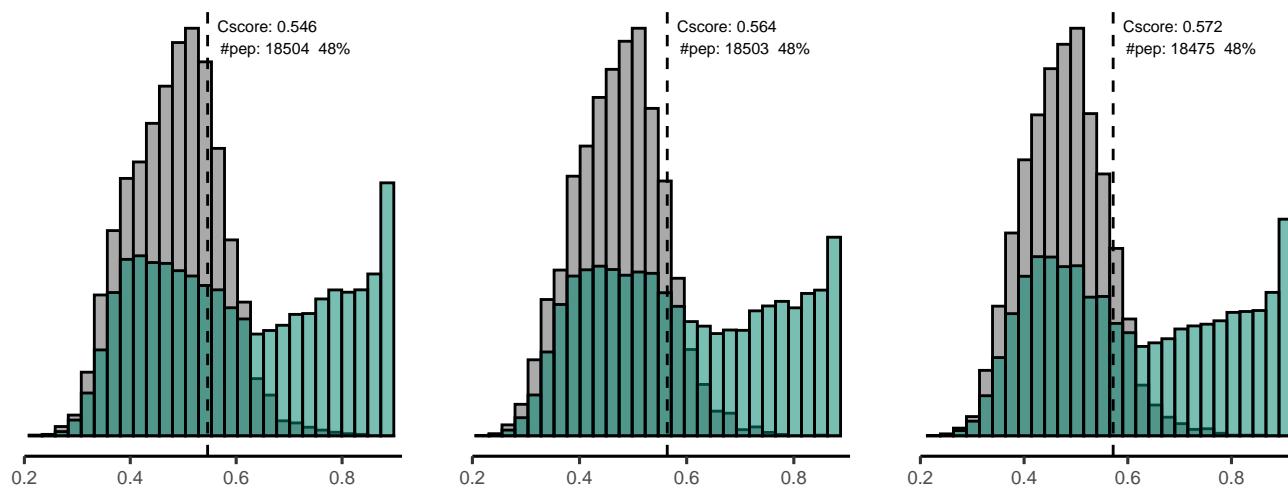
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1 20106228_QX1_PhGe_SA_EASY12-10_Deigendes2010618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



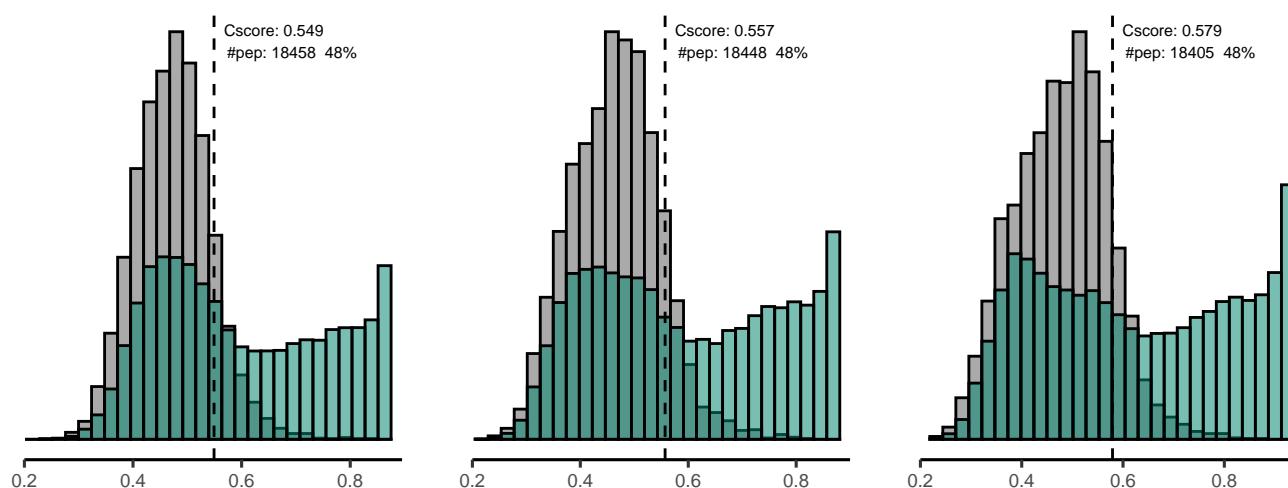
0210_QX1_PhGe_SA_EASY12-10_Majken_B9_CSF



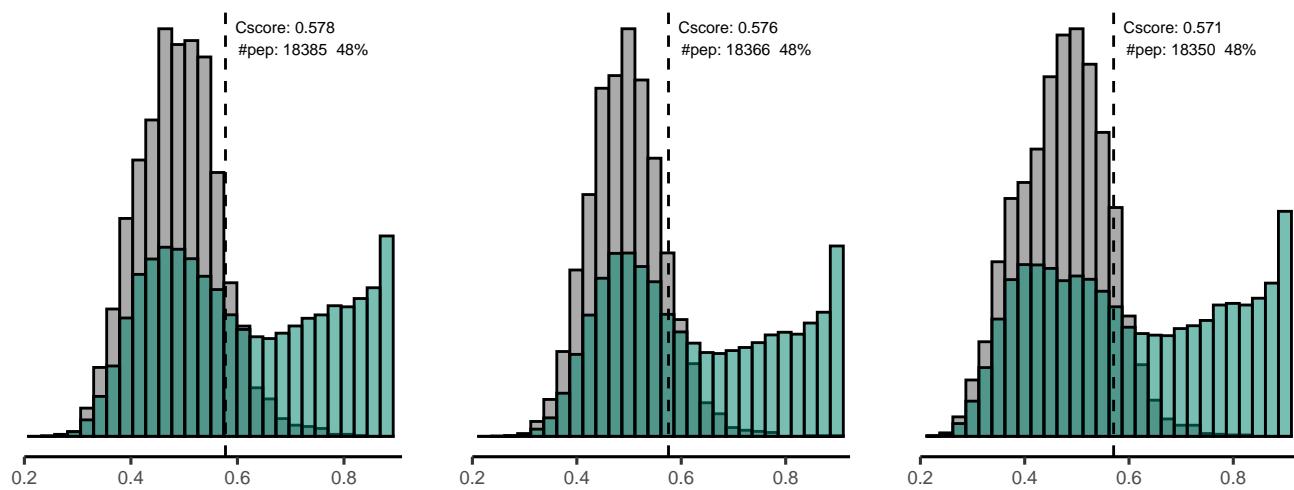
)228_QX1_PhGe_SA_EASY12-10_Deigendesch_E11_828720390210_QX1_PhGe_SA_EASY12-10_Majken_B9_CSF



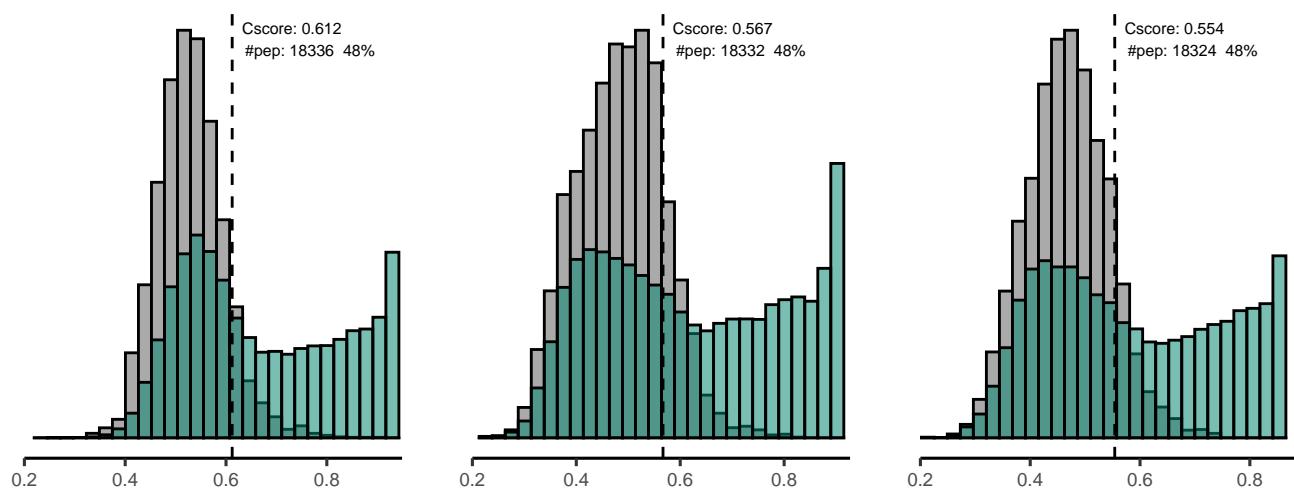
228_QX1_PhGe_SA_EASY12-10_Deigendesch_E320290210_QX1_PhGe_SA_EASY12-10_Majken_C7_CSF



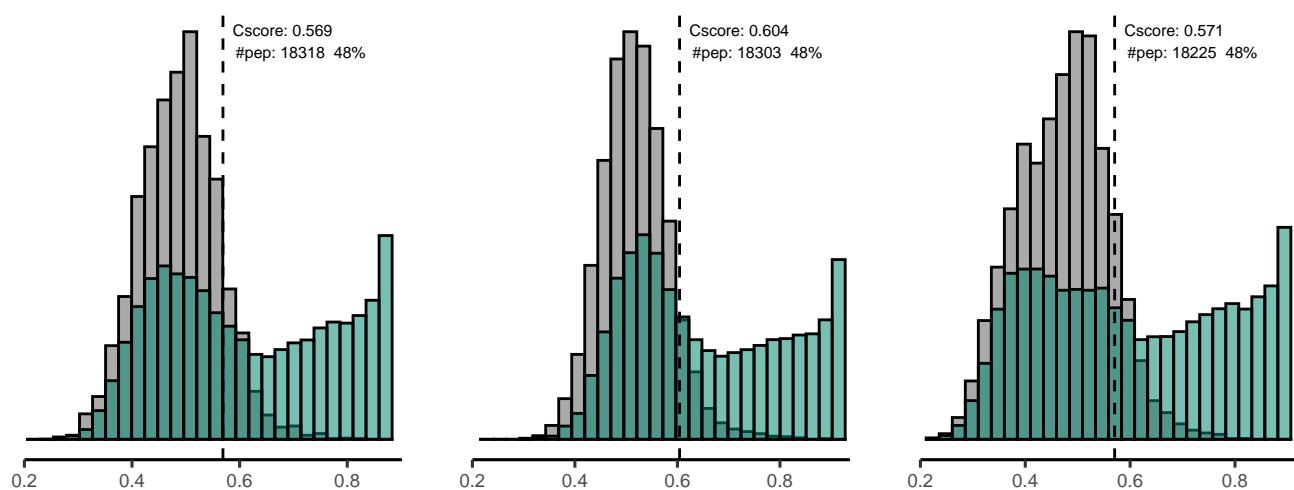
190210_QX1_PhGe_SA_EASY12-10_Majken_E5_C0190210_QX1_PhGe_SA_EASY12-10_Majken_E20C0210_QX1_PhGe_SA_EASY12-10_Majken_D4_CSF



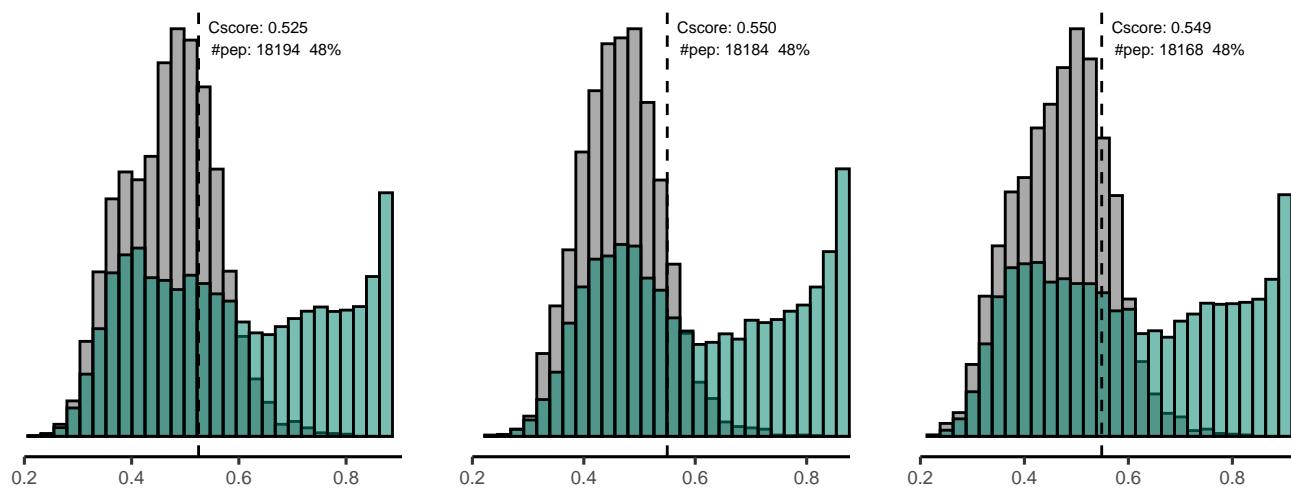
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS01M0228_QX1_PhGe_SA_EASY12-10_Deigendesch20390210_QX1_PhGe_SA_EASY12-10_Majken_B3_CSF



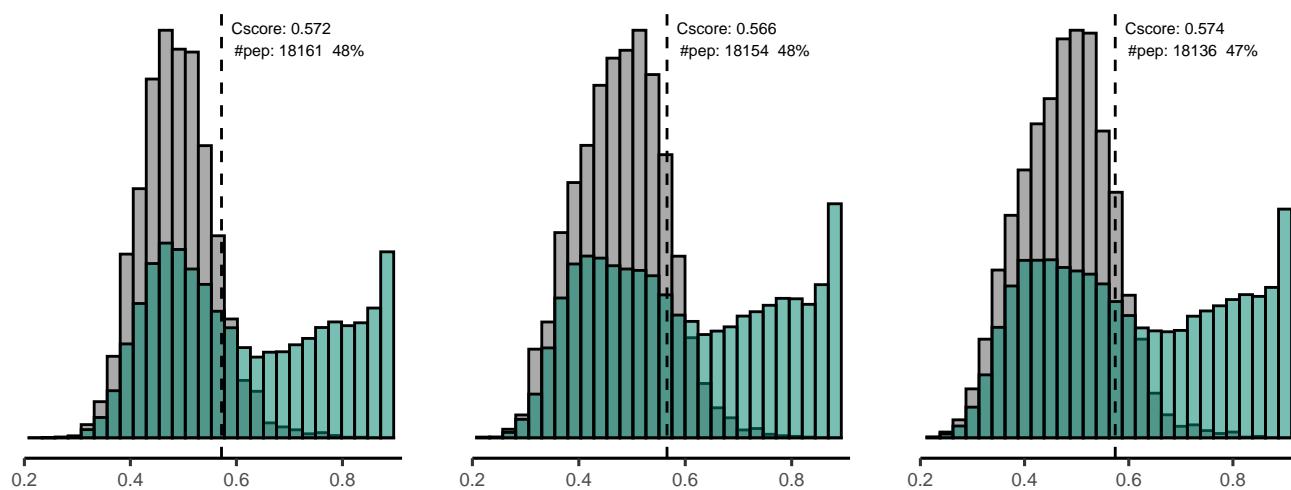
228_QX1_PhGe_SA_EASY12-10_Deigendesch2016618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1M190210_QX1_PhGe_SA_EASY12-10_Majken_D5_CSF



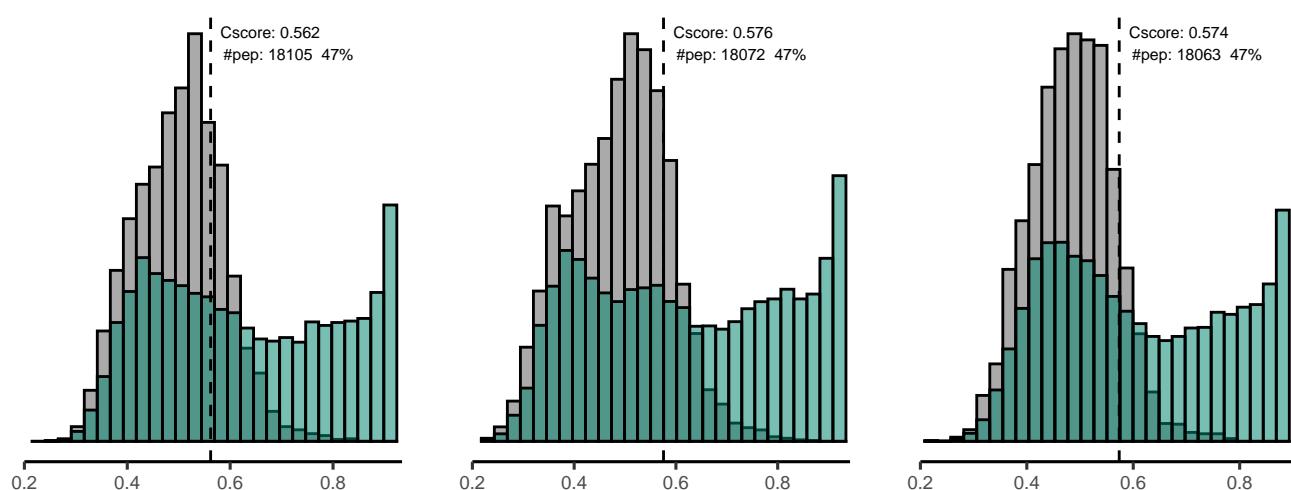
228_QX1_PhGe_SA_EASY12-10_Deigendesch_20190210_QX1_PhGe_SA_EASY12-10_Majken_A1020190210_QX1_PhGe_SA_EASY12-10_Majken_C9_CSF



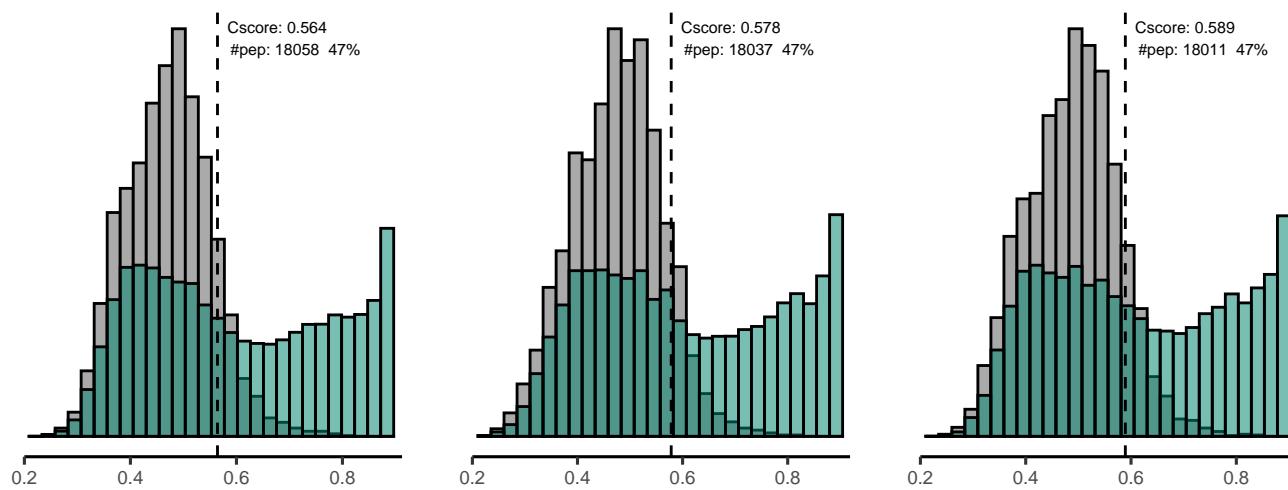
_QX1_PhGe_SA_EASY12-10_Deigendesch_F20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_G20190210_QX1_PhGe_SA_EASY12-10_Majken_F8_CS



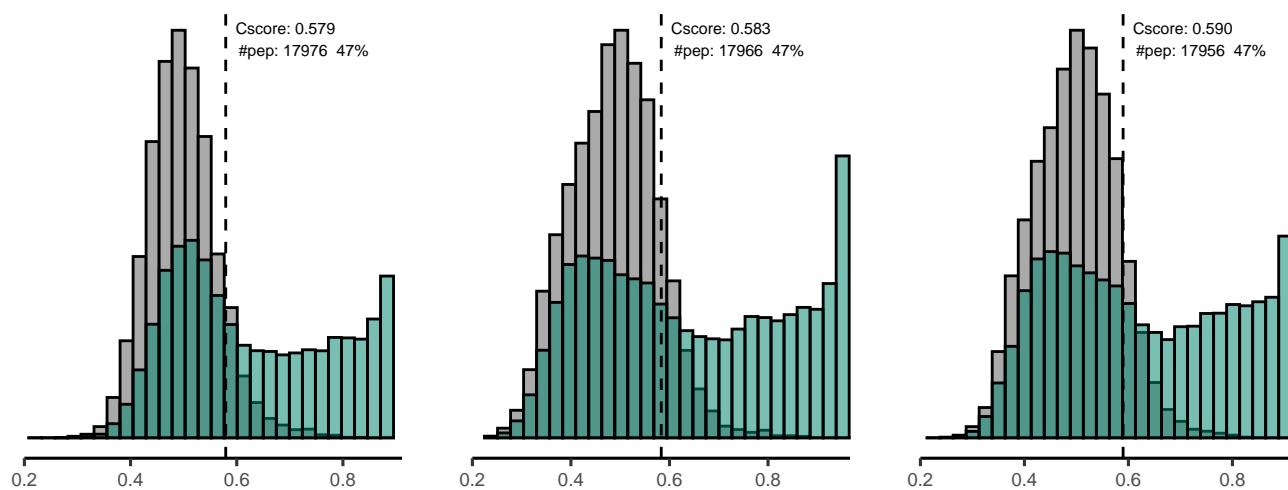
)228_QX1_PhGe_SA_EASY12-10_Deigendesch_20190210_QX1_PhGe_SA_EASY12-10_Majken_20190224_QX1_PhGe_SA_EASY12-10_Deigendesch_E8_1291885



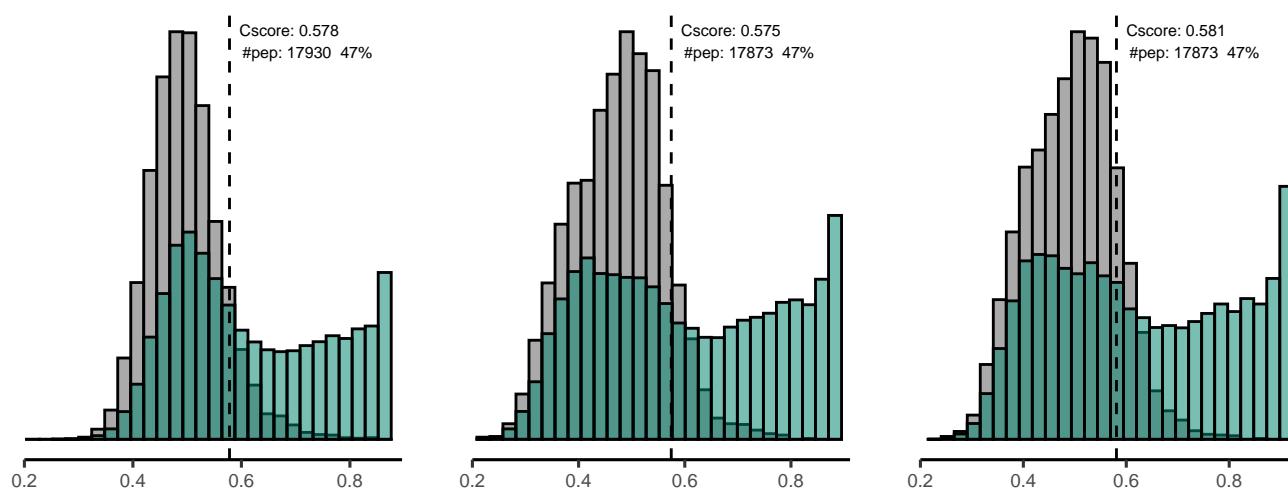
J228_QX1_PhGe_SA_EASY12-10_Deigendesch_A@20190210_QX1_PhGe_SA_EASY12-10_Majken_F4_20190210_QX1_PhGe_SA_EASY12-10_Majken_E2_C5



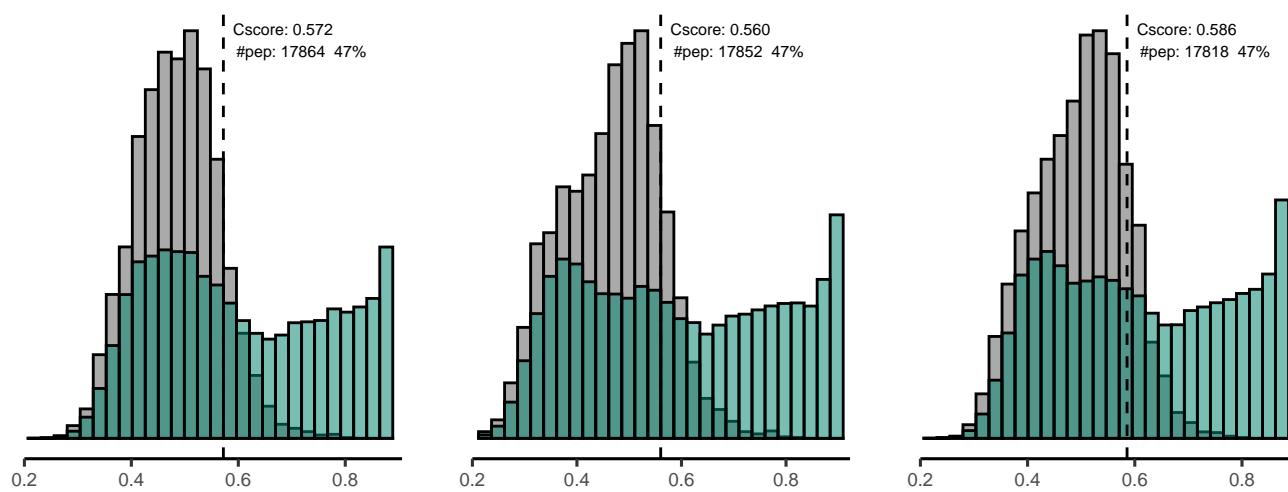
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1D1@20190210_QX1_PhGe_SA_EASY12-10_Majken_A12@20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_B5_



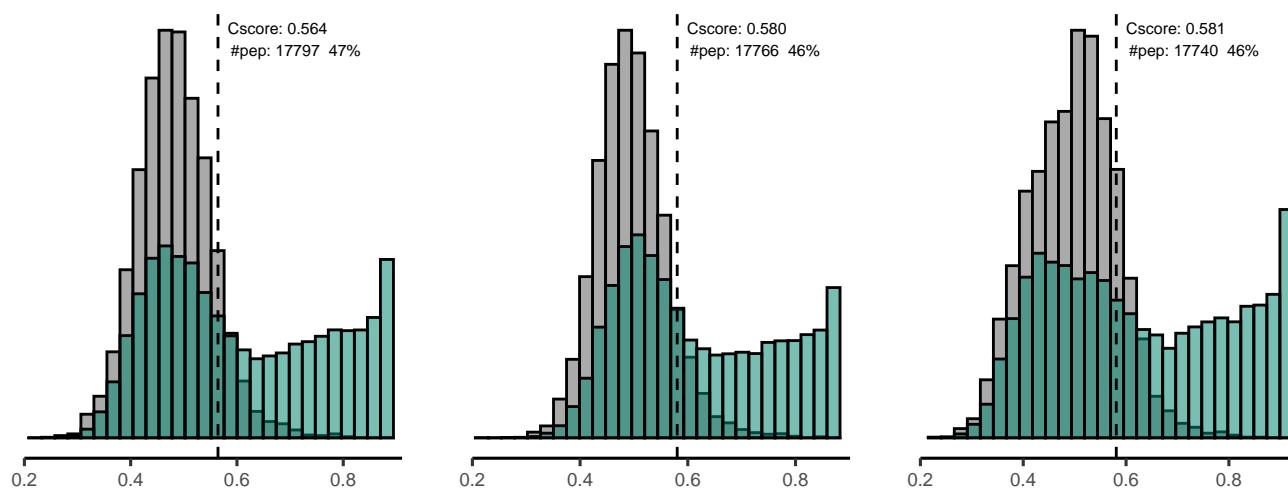
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1D1@20190210_QX1_PhGe_SA_EASY12-10_Majken_20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_E9_828



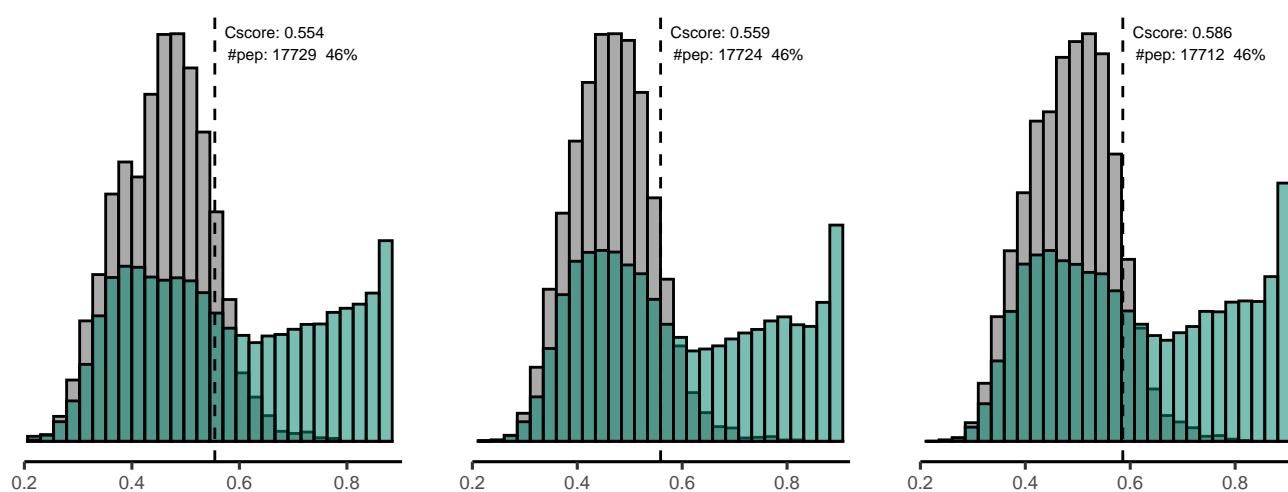
?228_QX1_PhGe_SA_EASY12-10_Deigendesch_B10



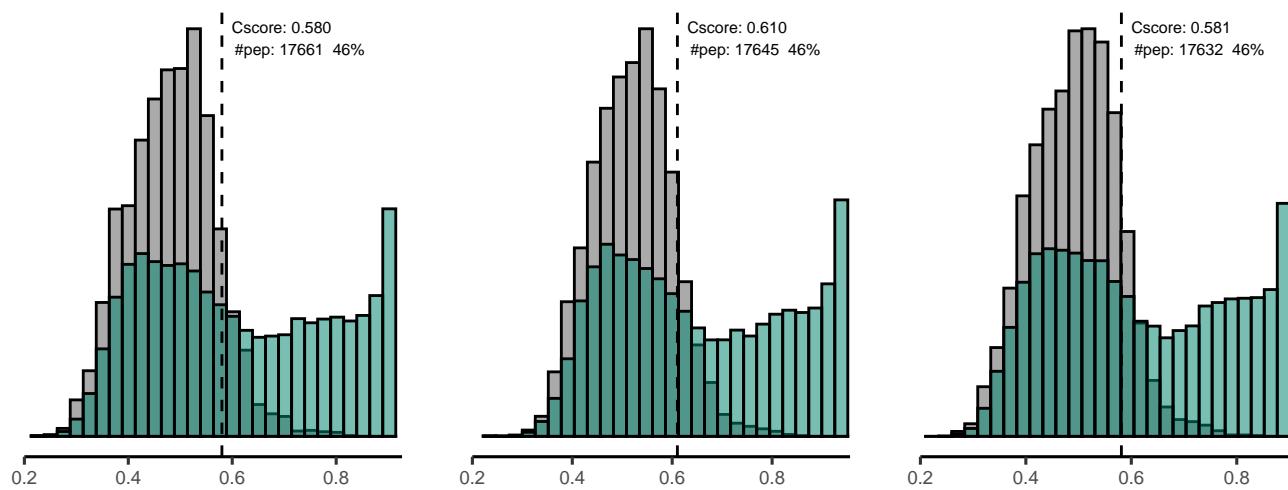
?28_QX1_PhGe_SA_EASY12-10_Deigendesch_B10



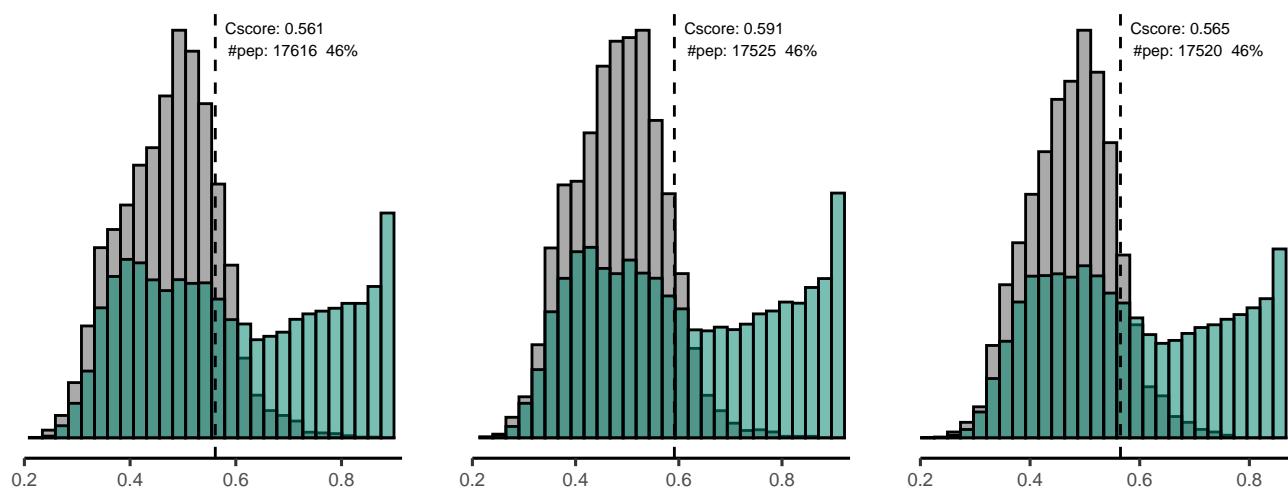
190210_QX1_PhGe_SA_EASY12-10_Majken_B10



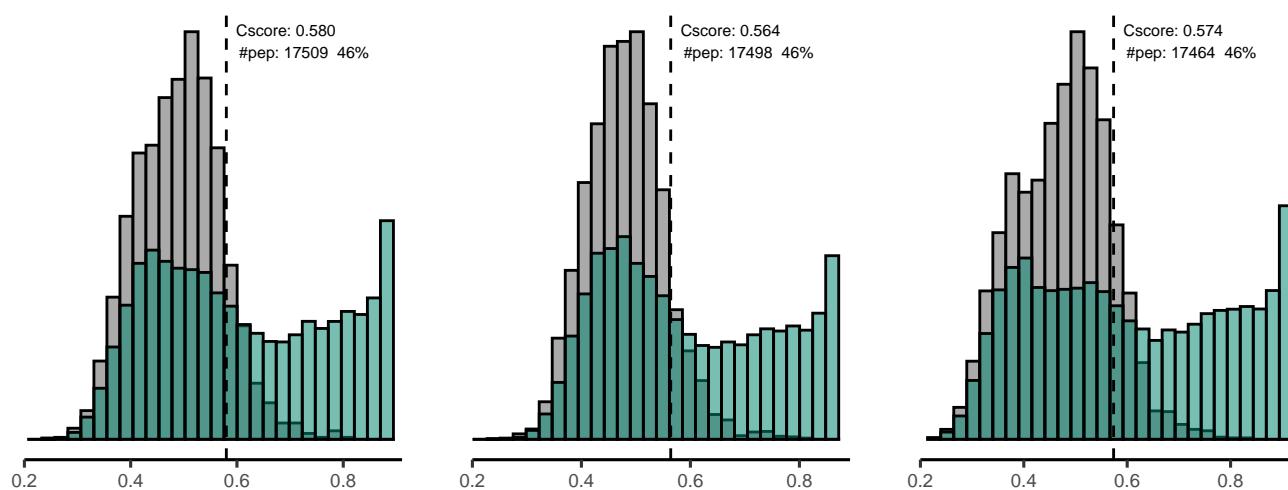
!228_QX1_PhGe_SA_EASY12-10_Deigendesch_B20190210_QX1_PhGe_SA_EASY12-10_Deigendesch_D120190228_QX1_PhGe_SA_EASY12-10_Deigendesch_A2_



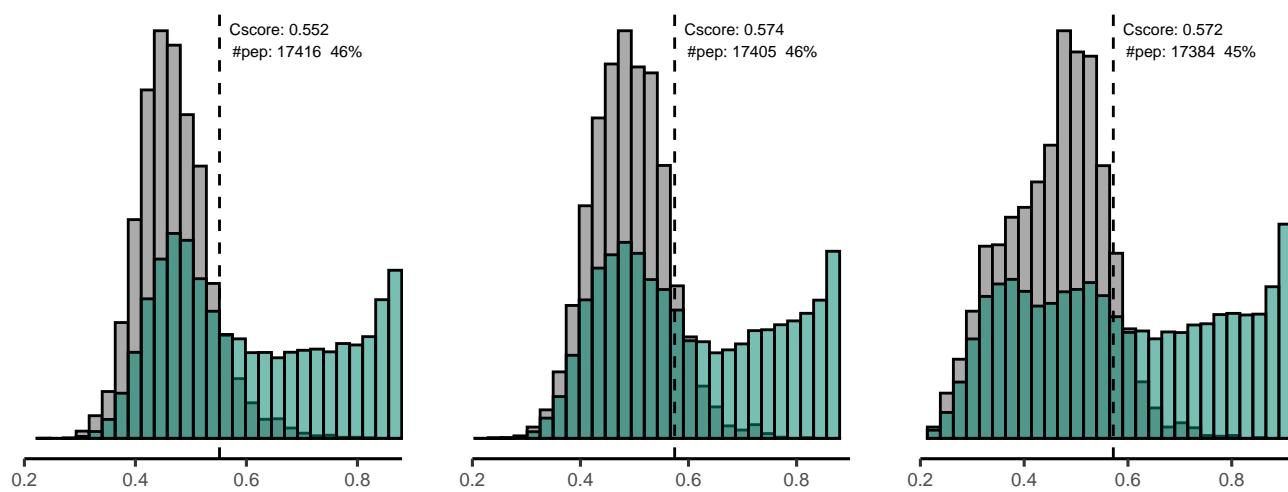
?28_QX1_PhGe_SA_EASY12-10_Deigendesch_B20190210_QX1_PhGe_SA_EASY12-10_Majken_E20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_D1_



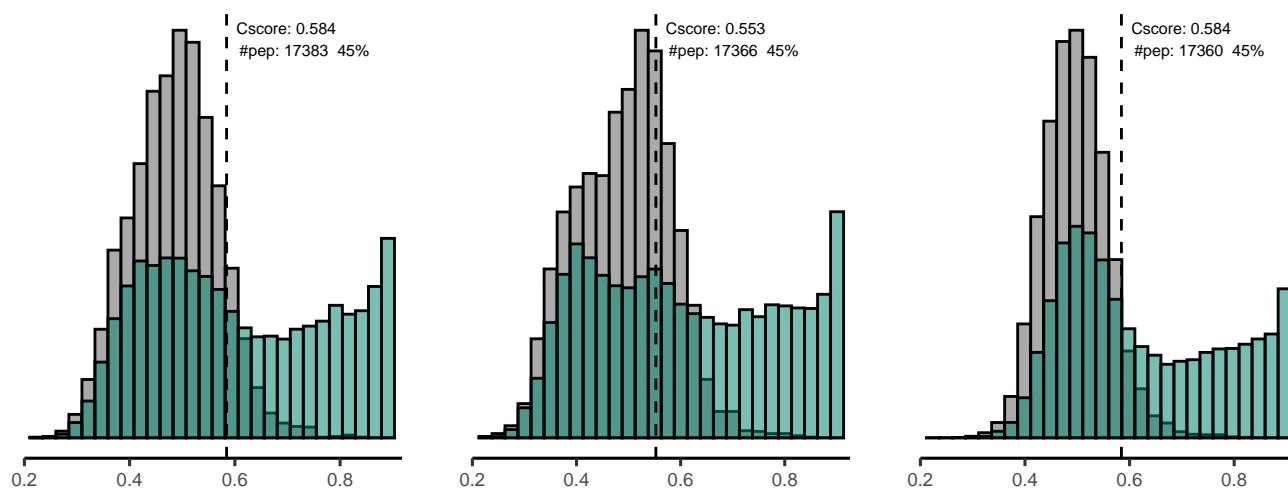
_QX1_PhGe_SA_EASY12-10_Deigendesch_A20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_D20190210_QX1_PhGe_SA_EASY12-10_Majken_A8_CSF



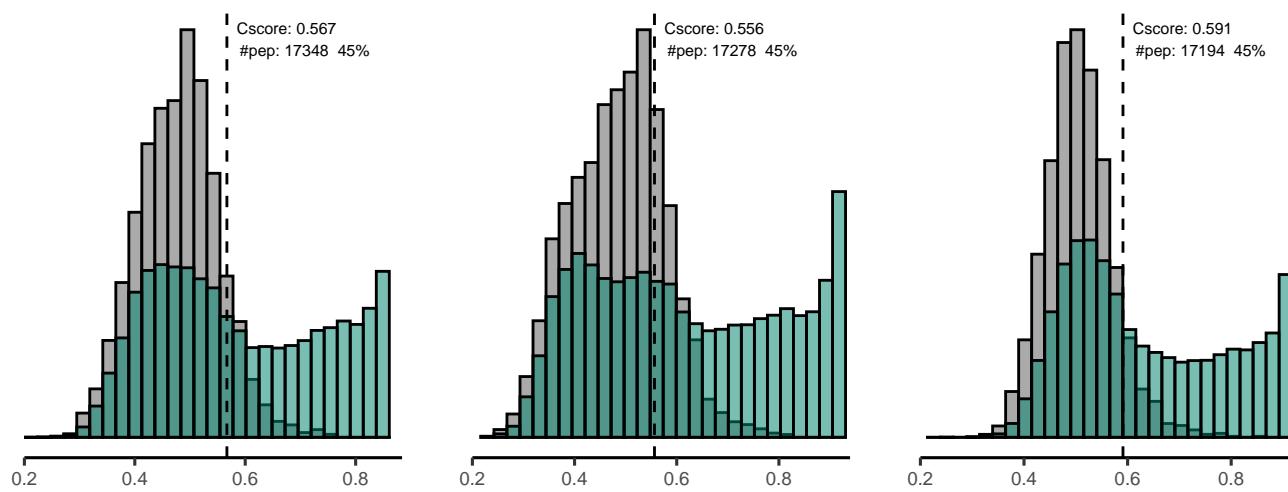
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1 20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_20180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1



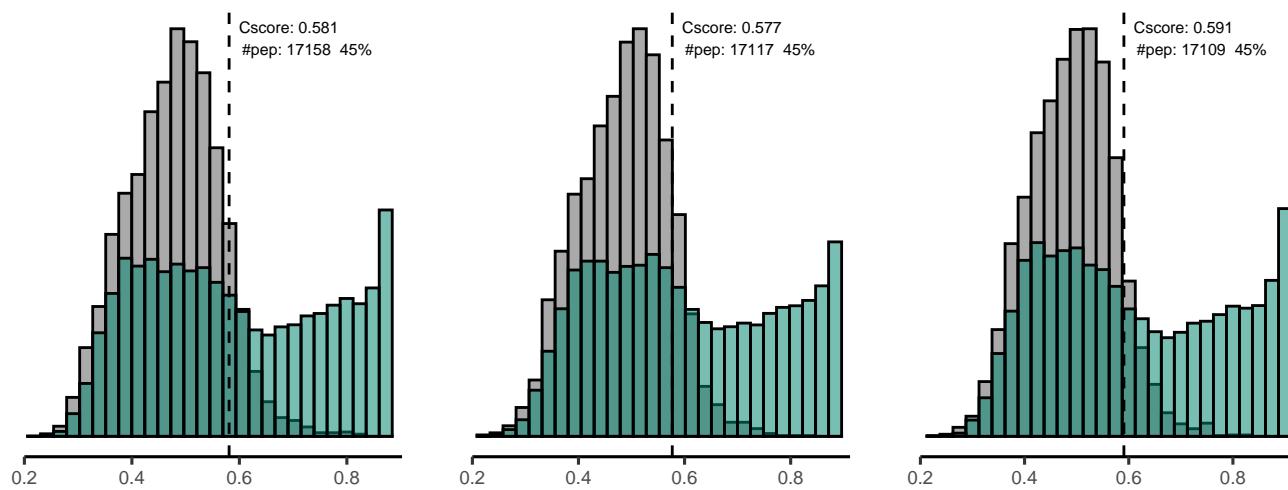
190210_QX1_PhGe_SA_EASY12-10_Majken_20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_20180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1



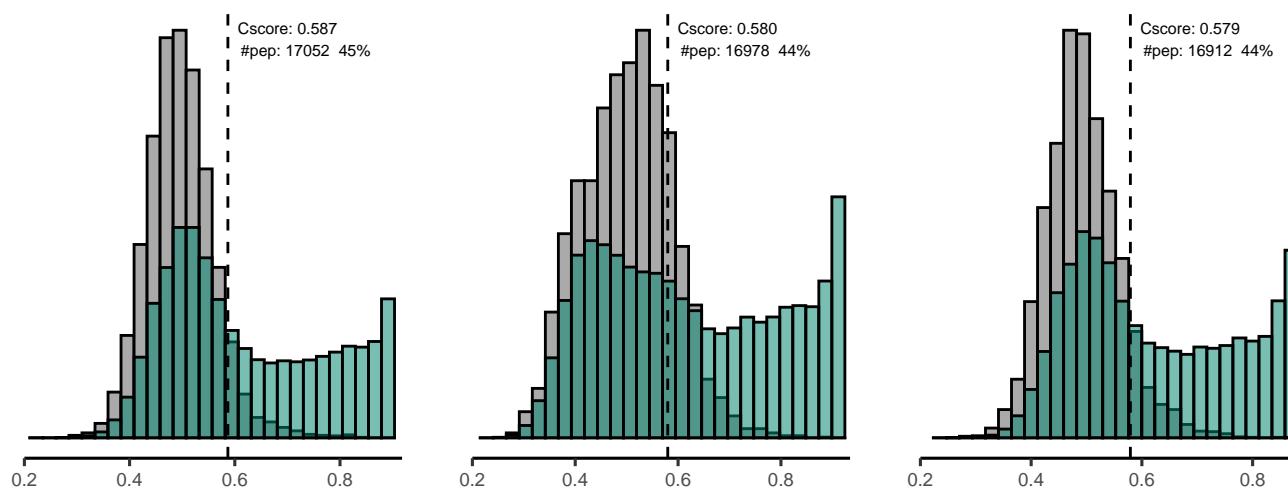
228_QX1_PhGe_SA_EASY12-10_Deigendesch_20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_20180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1



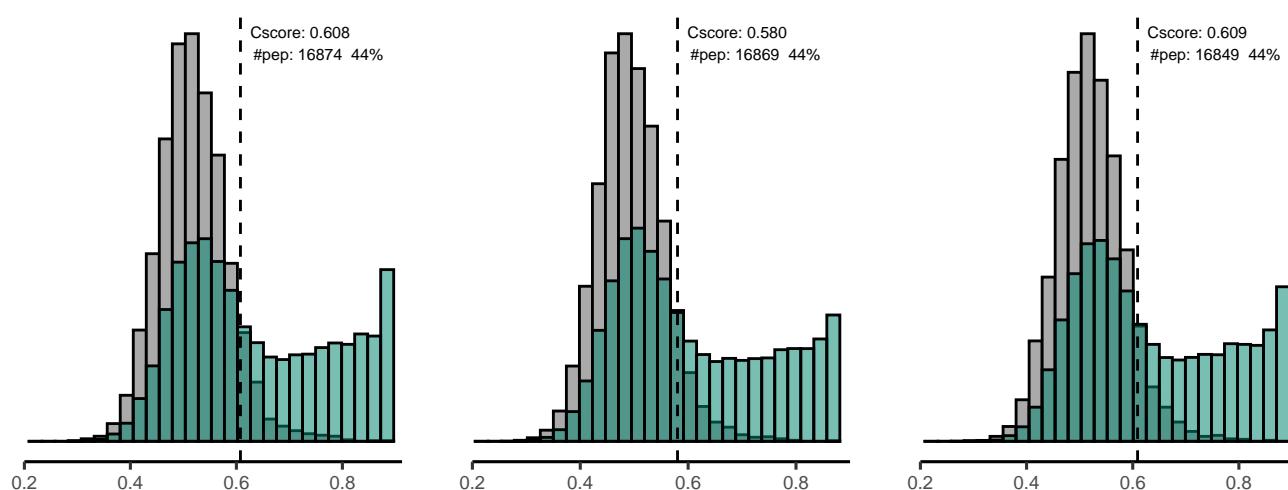
1228_QX1_PhGe_SA_EASY12-10_Deigendesch_B9_



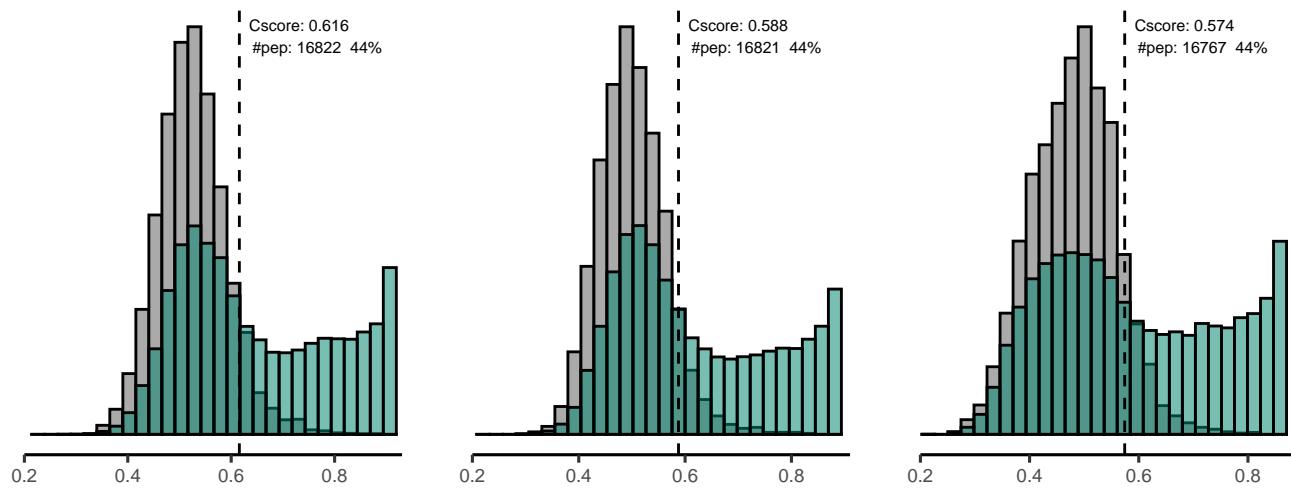
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1_



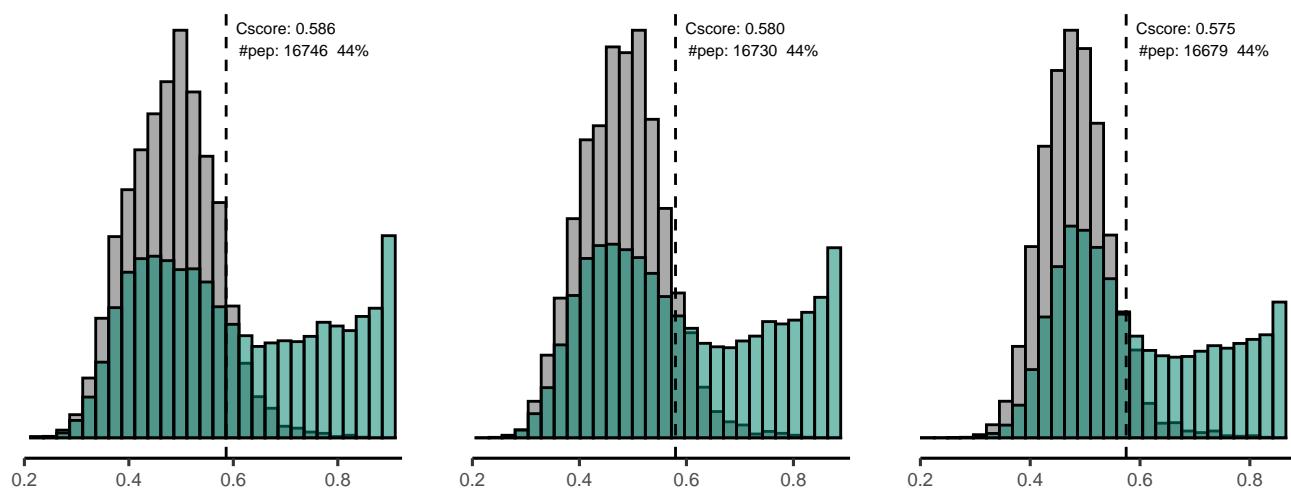
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1_



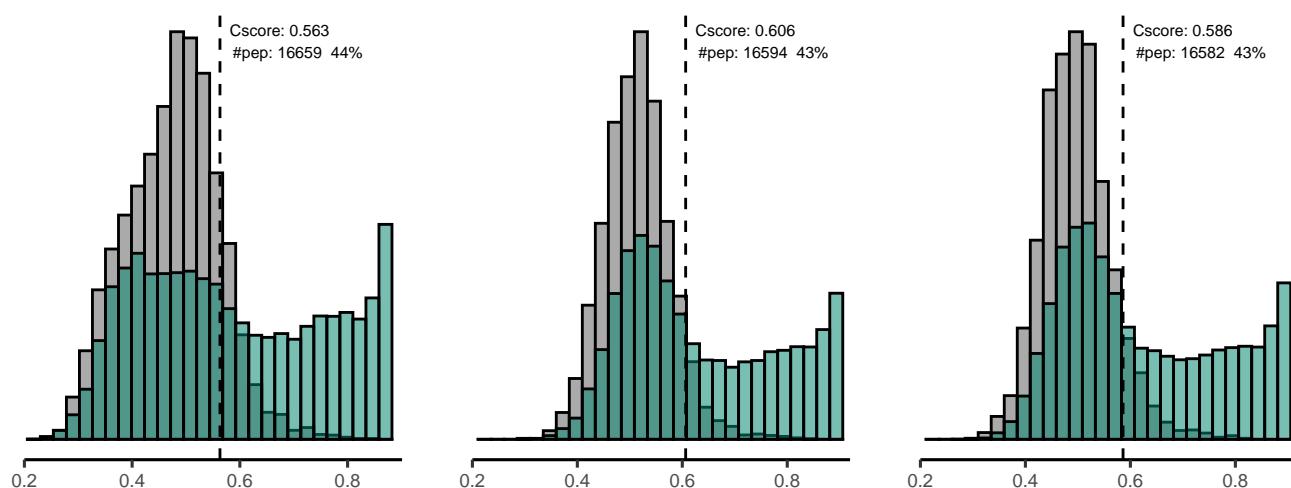
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1 2019072618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1 2019072618_QX1_PhGe_SA_EASY12-10_Deigendesch_B8_



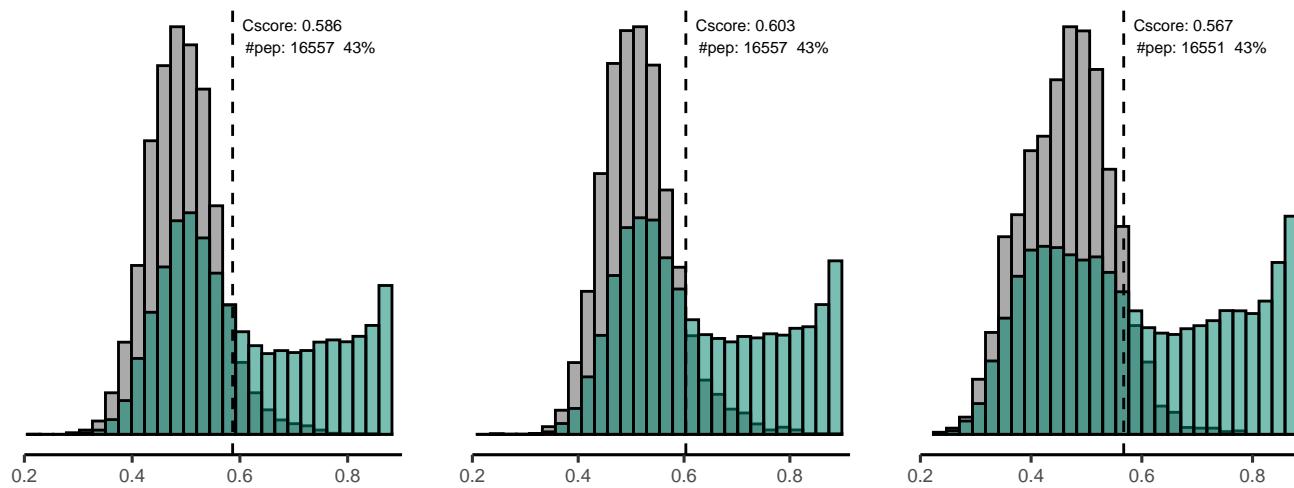
190210_QX1_PhGe_SA_EASY12-10_Maj 2019072618_QX1_PhGe_SA_EASY12-10_Deigendesch_F2019072618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1



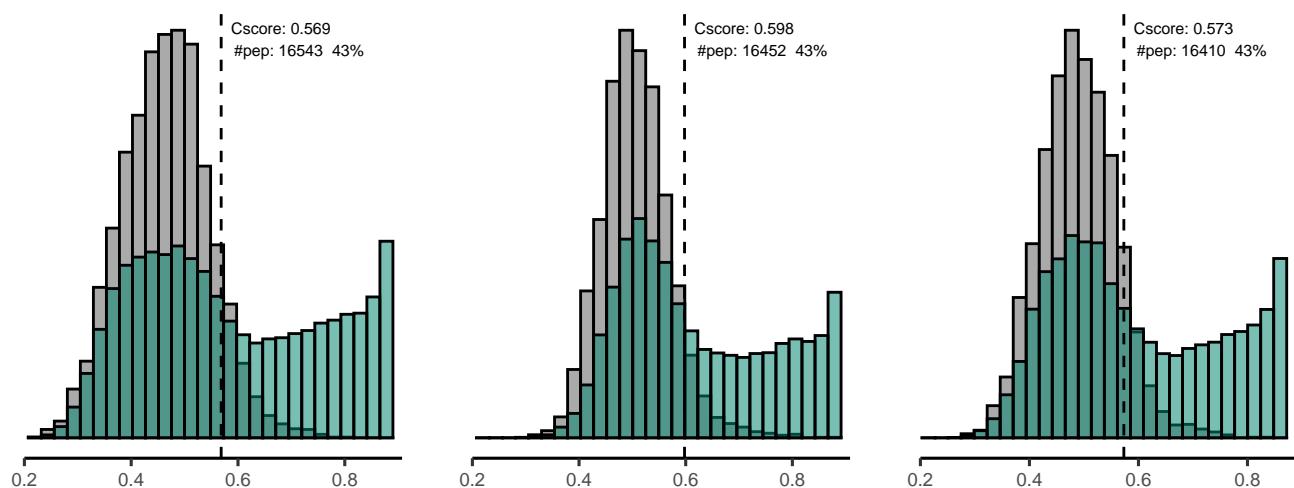
228_QX1_PhGe_SA_EASY12-10_Deigendesch 20190618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1 20190618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1



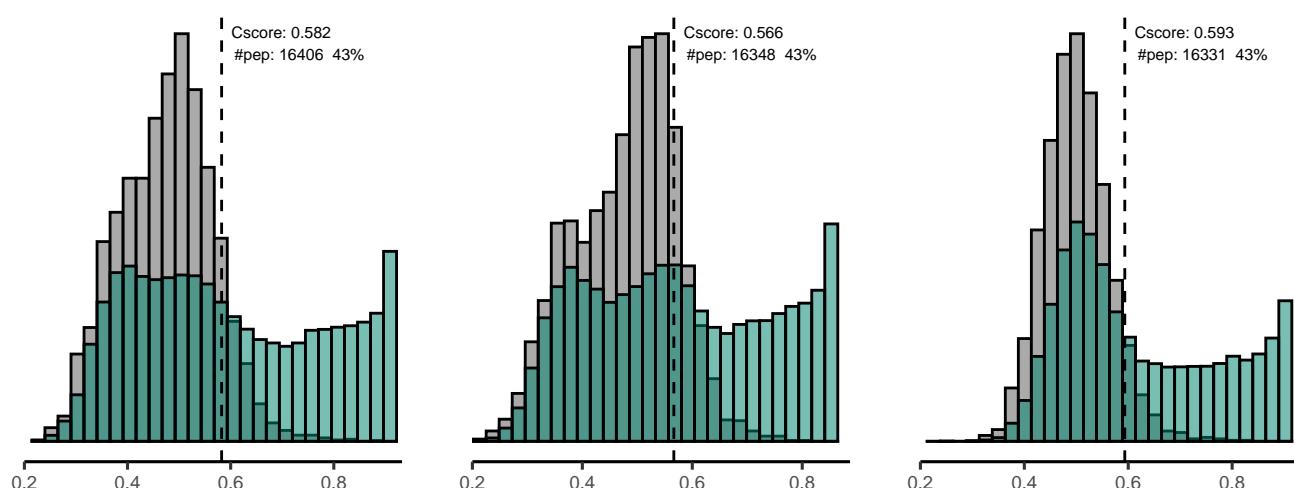
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1 18_QX0_JaBa_SA_LC12_5_CSF1_1_8-2xD1xS1fM1 18_QX0_JaBa_SA_EASY12-10_Deigendesch_F9_828801



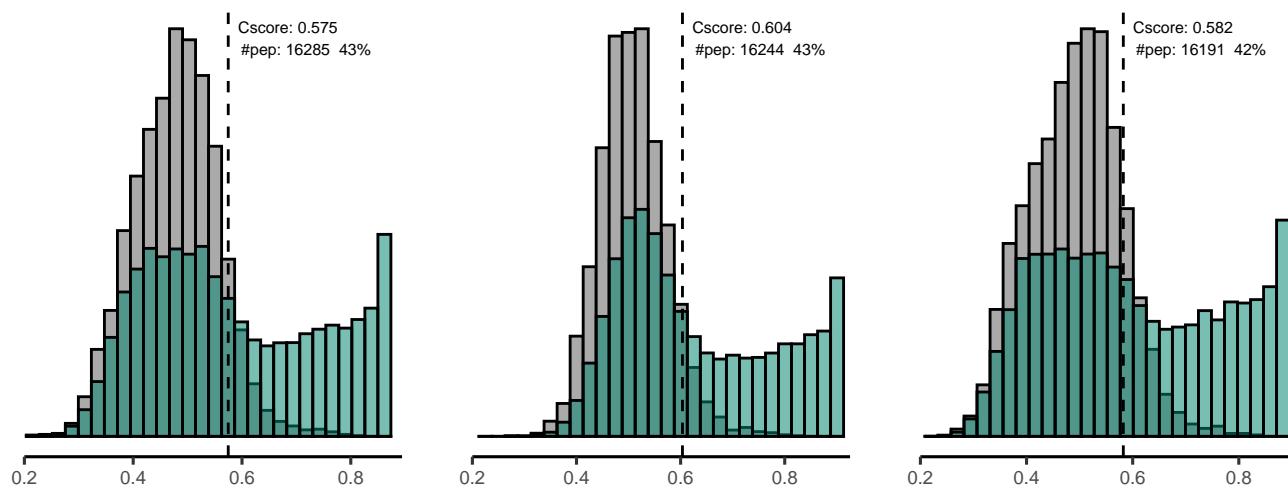
190210_QX1_PhGe_SA_EASY12-10_Majken 2018_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1 2018_QX1_PhGe_SA_EASY12-10_Deigendesch_C5_



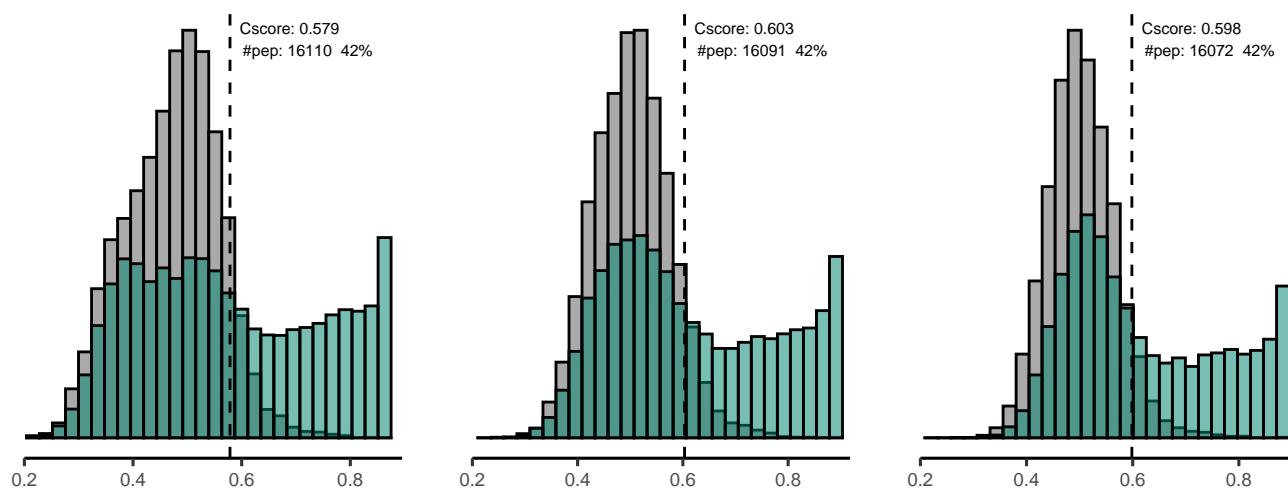
190210_QX1_PhGe_SA_EASY12-10_Majken 2018_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1 2018_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1



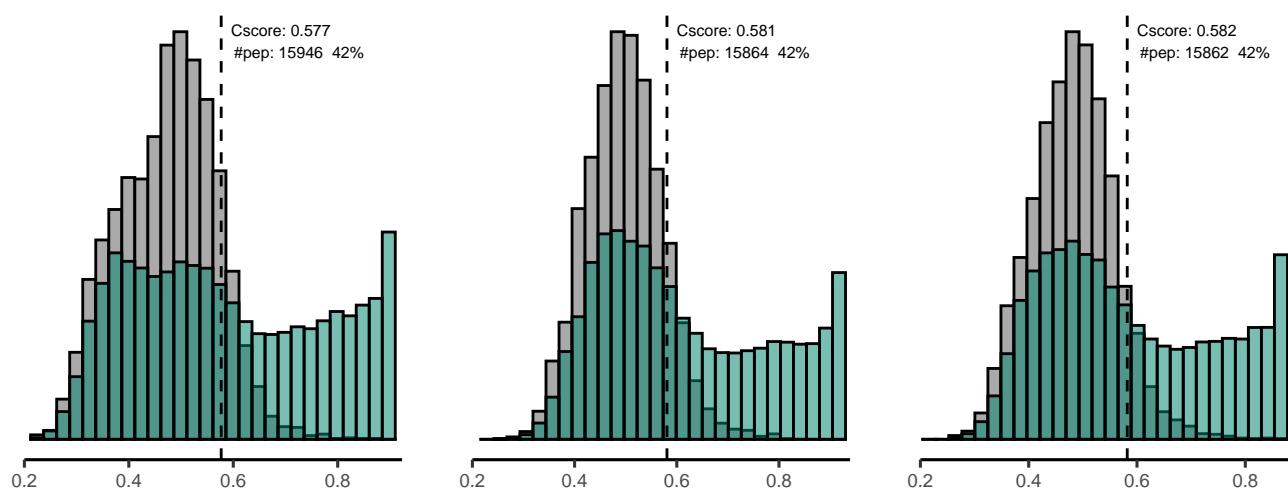
_QX1_PhGe_SA_EASY12-10_Deigendesch_Z0118_QX0_JaBa_SA_LC12_5_CSF1_1_8201x01281\QX1_PhGe_SA_EASY12-10_Deigendesch_F6_1291899



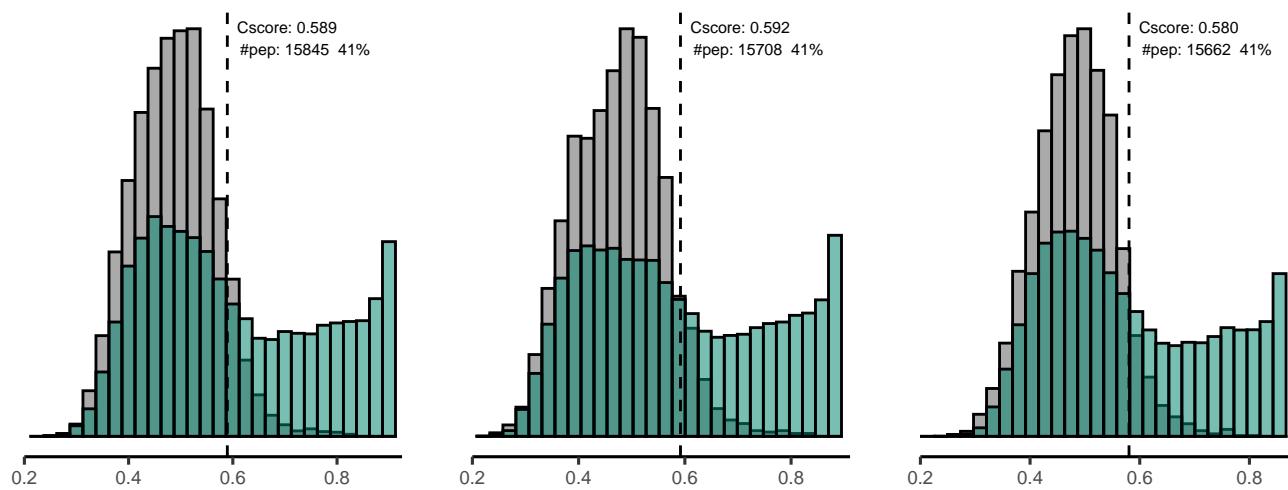
_QX1_PhGe_SA_EASY12-10_Deigendesch_Z0118_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



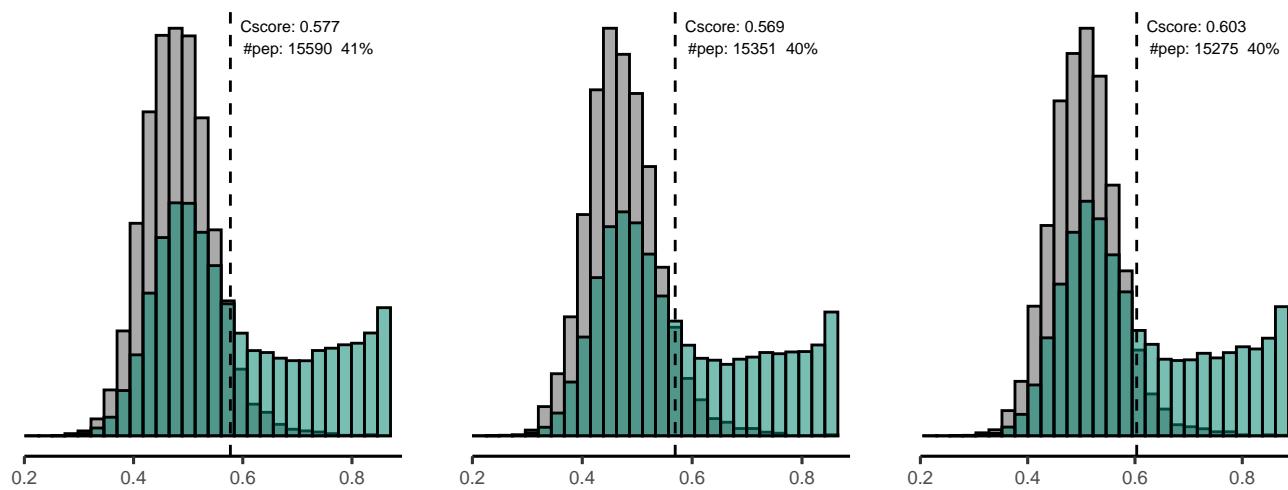
190210_QX1_PhGe_SA_EASY12-10_Maj_20190228\QX1_PhGe_SA_EASY12-10_Deigendesch_E\20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_A12_



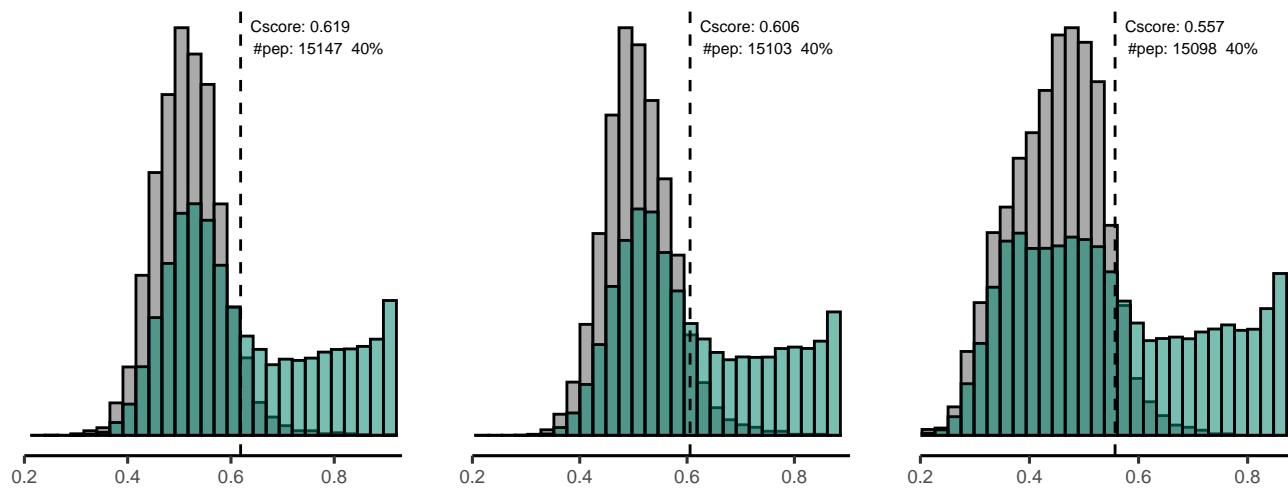
1228_QX1_PhGe_SA_EASY12-10_Deigendesch_D20190210_QX1_PhGe_SA_EASY12-10_Majken_20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_E10_



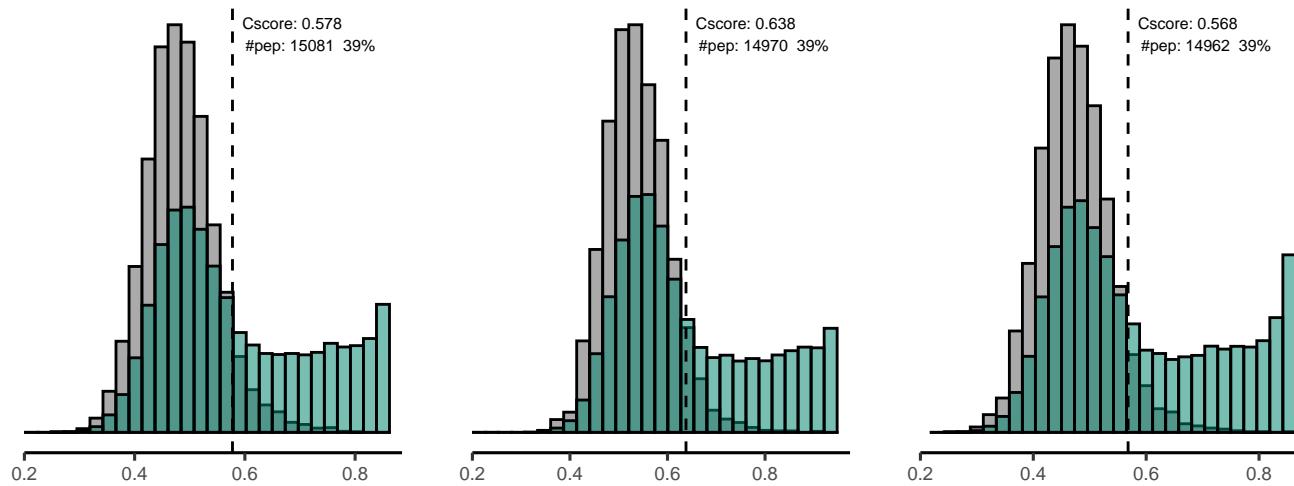
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



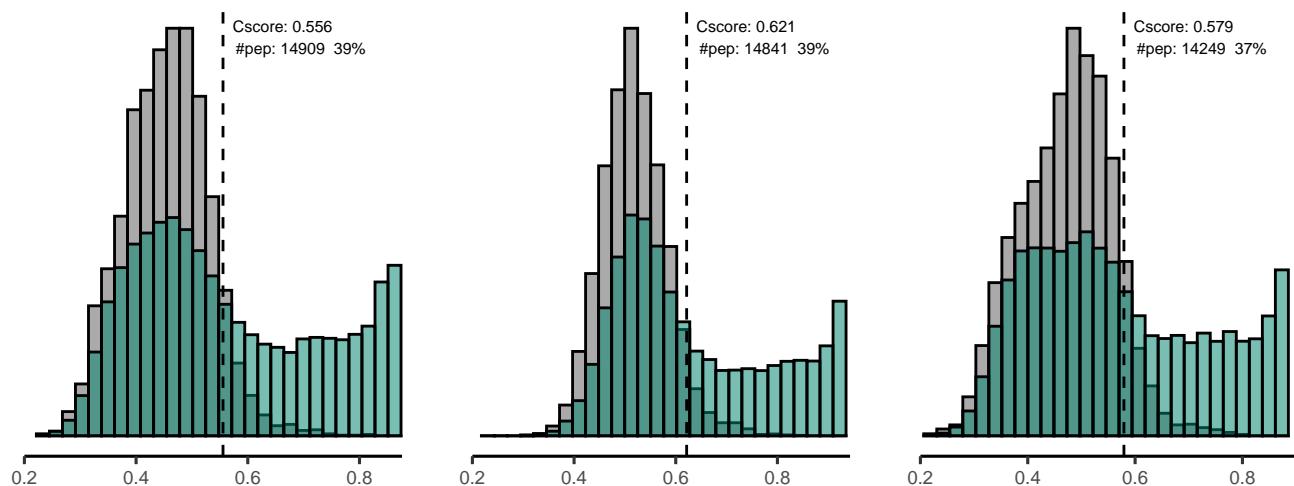
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1_20190228_QX1_PhGe_SA_EASY12-10_Deigendesch_G7_



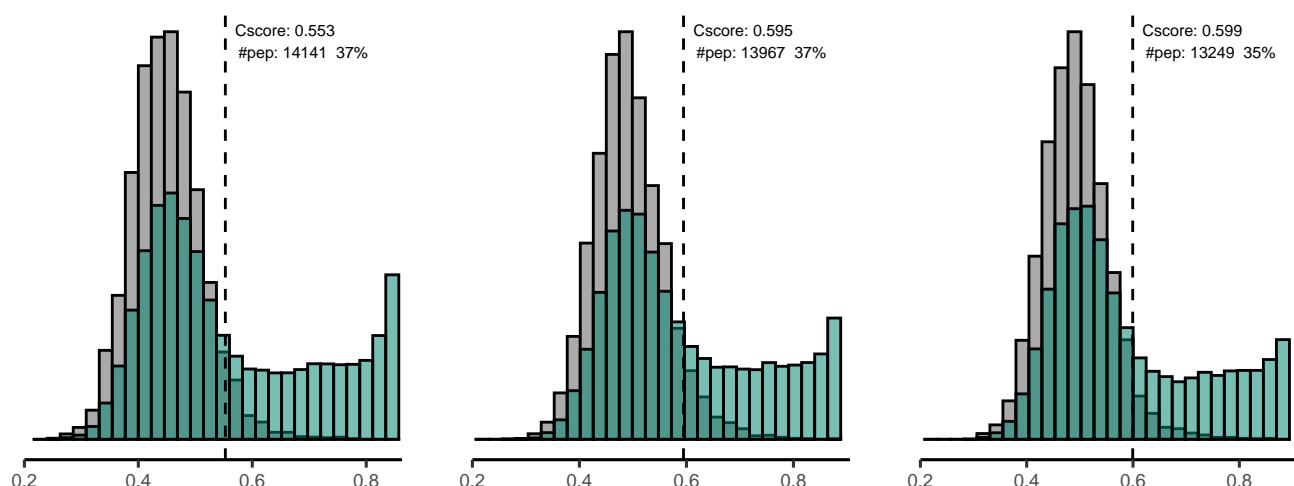
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



228_QX1_PhGe_SA_EASY12-10_Deigendes20180618_QX0_JaBa_SA_LC12_5_CSF1_1_8201D028810X1_PhGe_SA_EASY12-10_Deigendesch_G9_828799



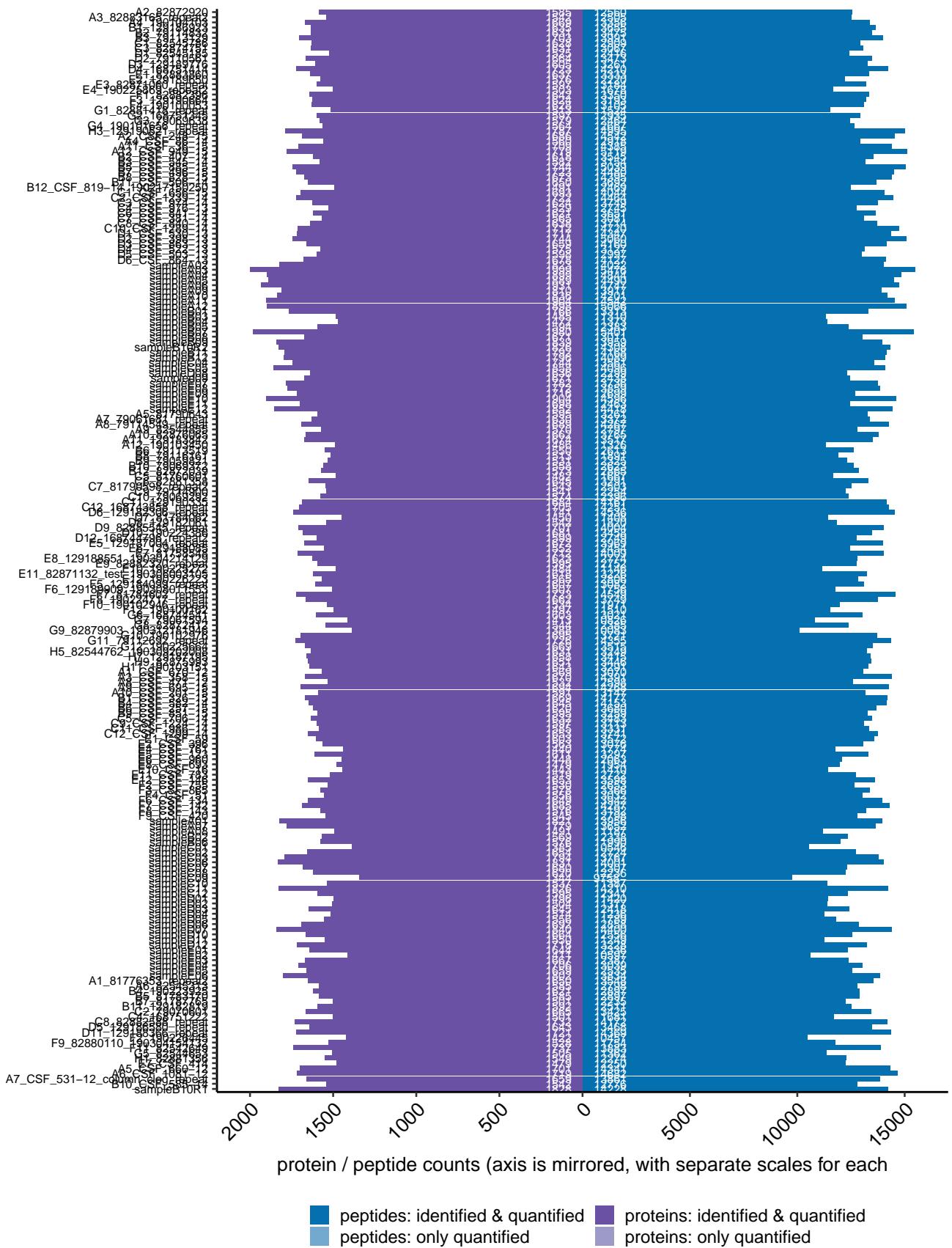
18_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD128180618_QX0_JaBa_SA_LC12_5_CSF1_1_8-1xD1xS1fM1.



1.2 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’.

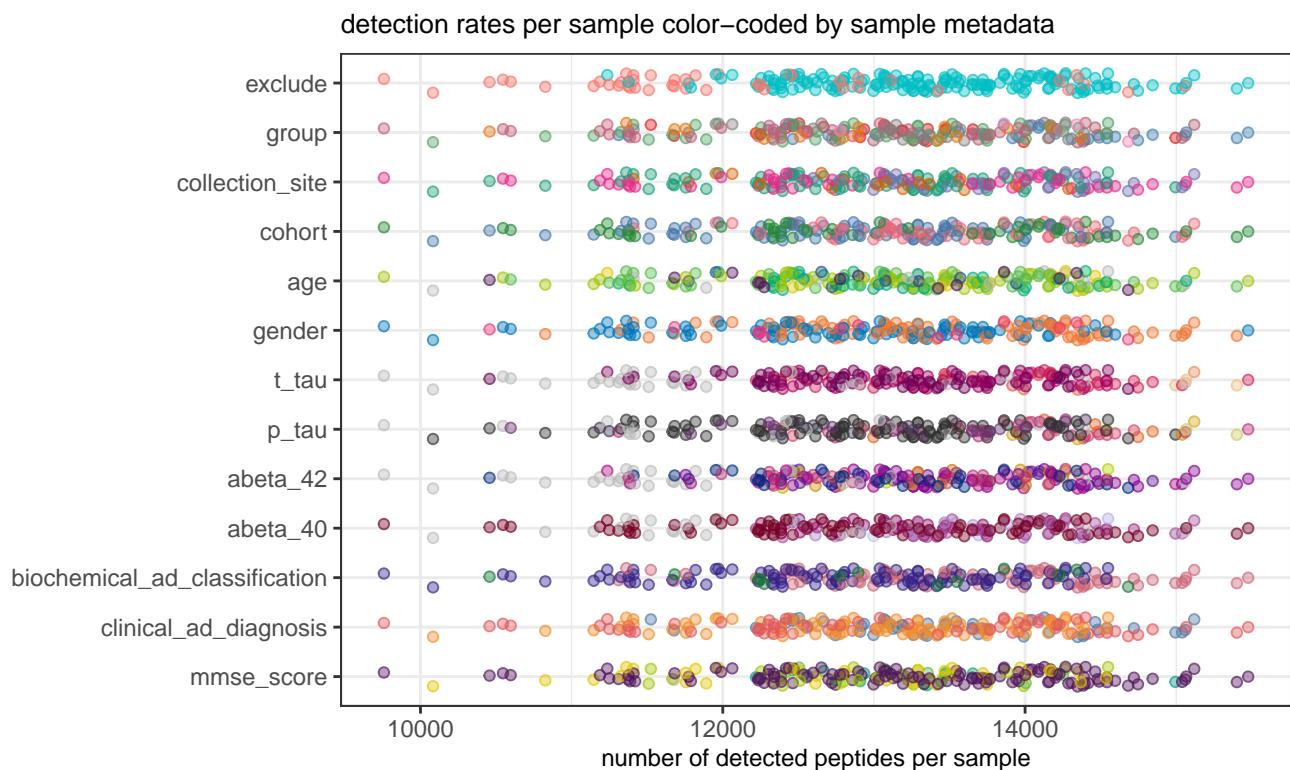


1.2.1 color-coding sample metadata

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 provides an overview of all sample metadata at a first glance. On each row all samples in the dataset are shown as a data point, each color-coded by the respective property shown on the y-axis (with minor vertical jitter for visual clarity). If any of these metadata coincide with a major effect on the number of detected peptides, this should become apparent by a clustering of samples by color-code. Hereafter, an additional set of figures will further expand this overview into detailed figures for each sample property.

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. All information shown in these figures is also available in the output table ‘samples.xlsx’.



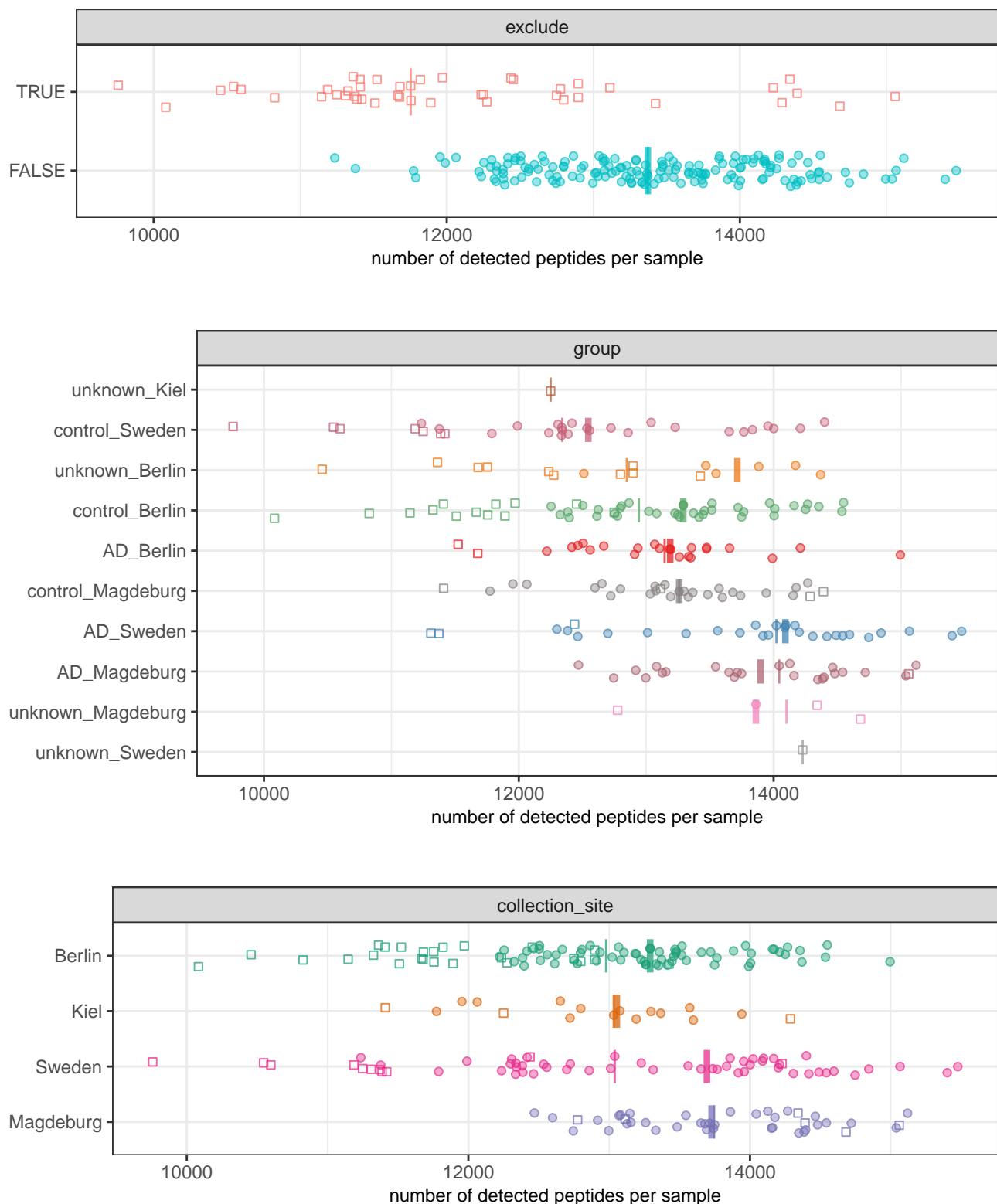
color-coding sample metadata, expanded

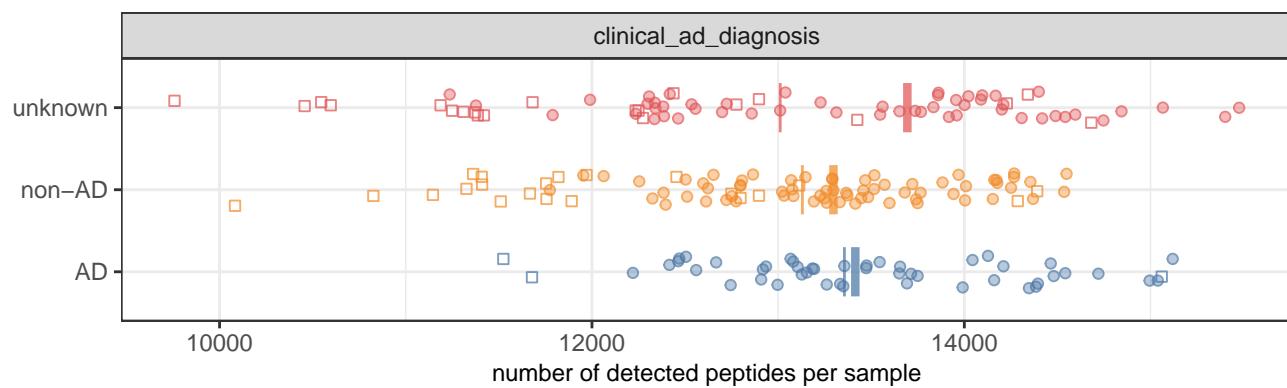
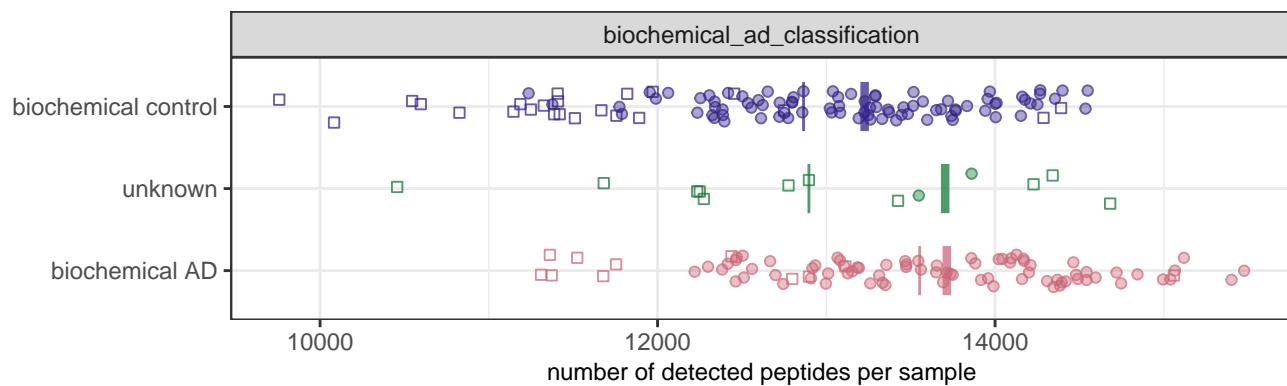
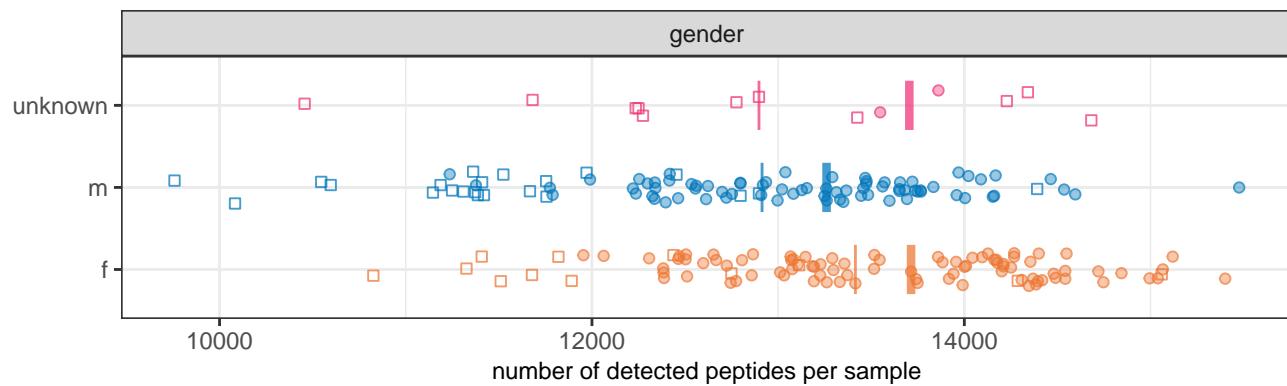
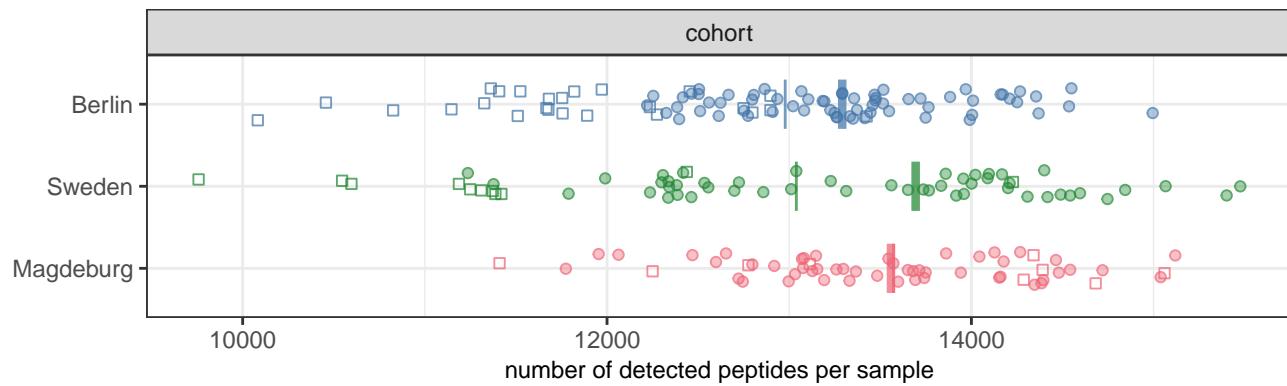
To further detail each sample property, each row in the above figure is now split into separate plots. Thus, a figure is generated for each property in the user-provided metadata (column in the samples table, its name shown in the plot title).

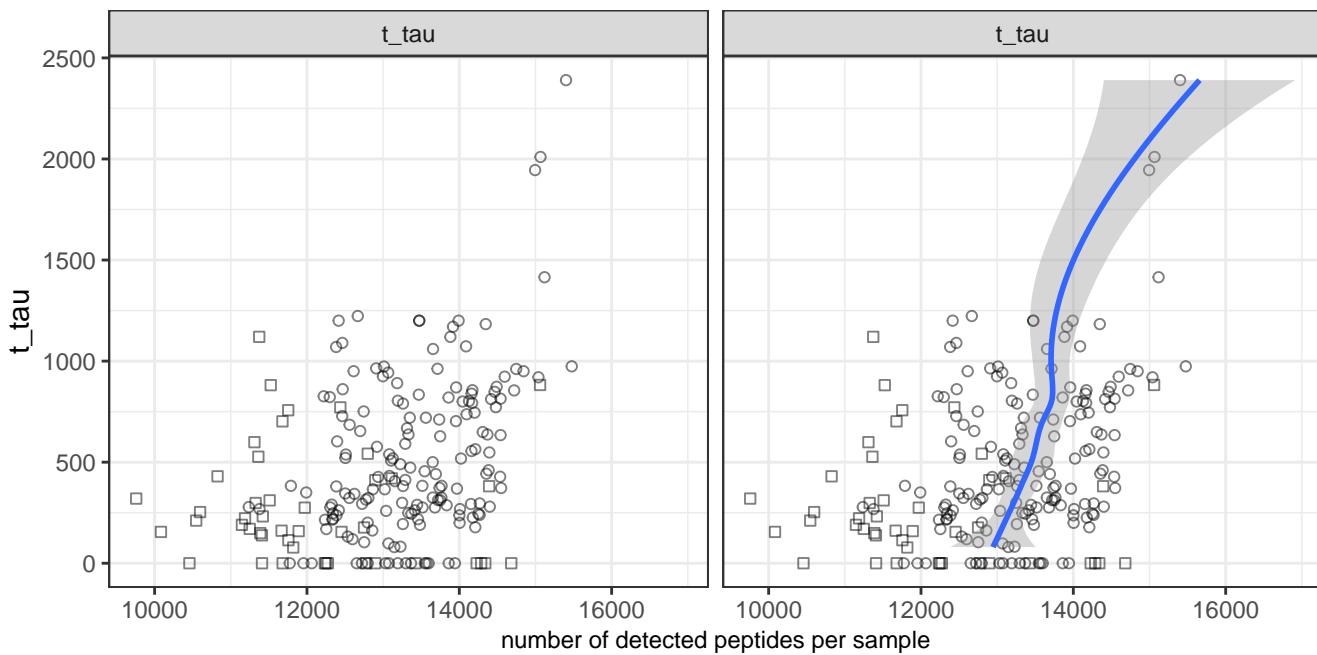
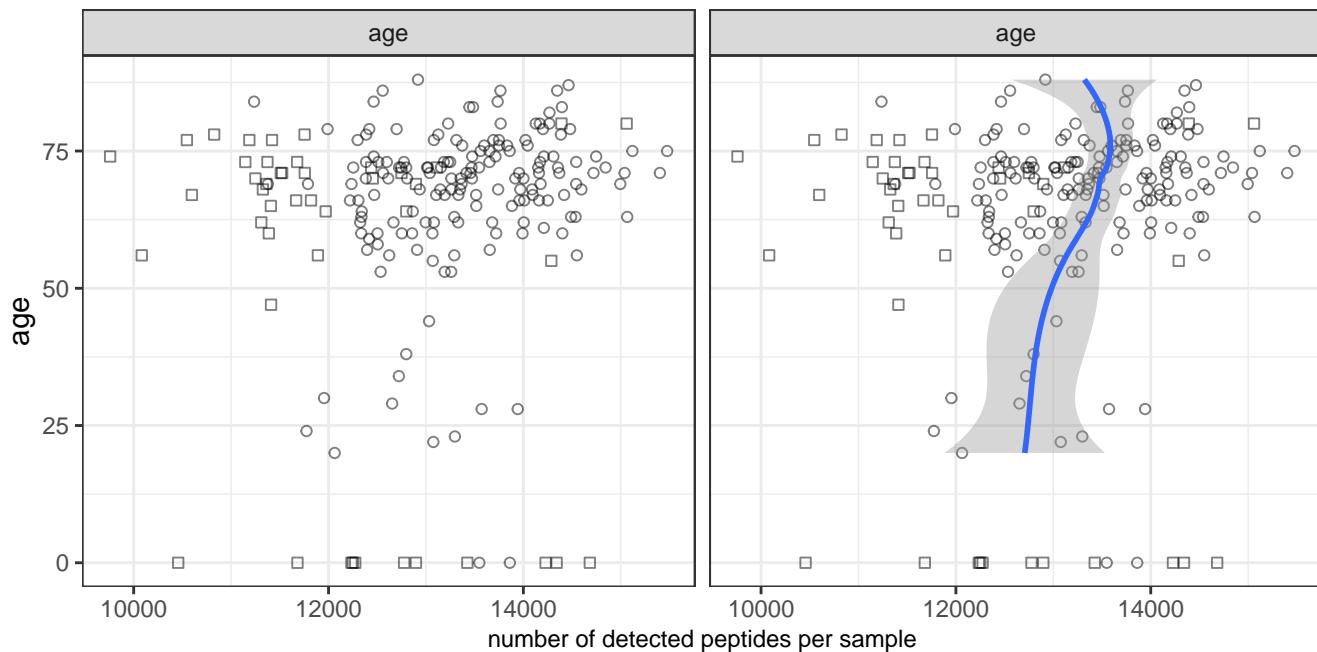
For categorical variables, a scatterplot shows on the y-axis all unique variables while the x-axis depicts the number of detected peptides. Colors are consistent with the above plot. *exclude* samples, if any, are depicted as squares. The median value is shown as a vertical line (thin line = median over all samples, wider line = median while discarding *exclude* samples). For continuous variables, a scatterplot without (left panel) and with Loess fit is shown (right panel, visualized as blue line if data was successfully fitted).

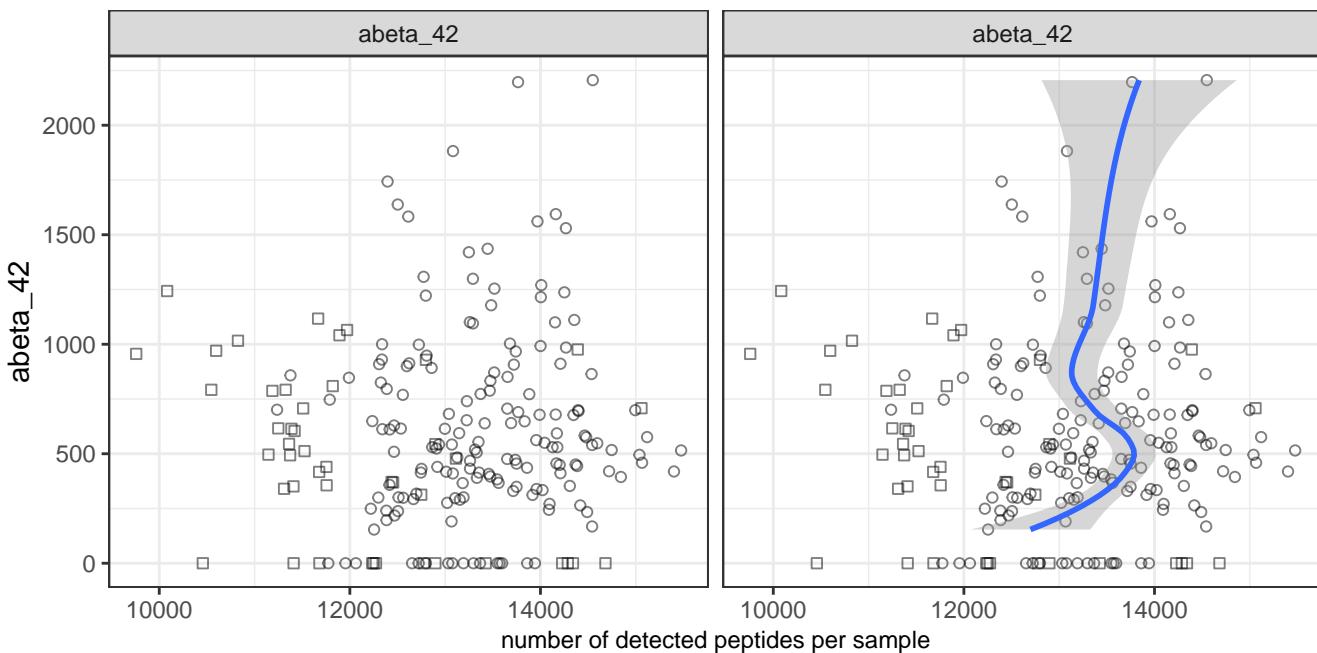
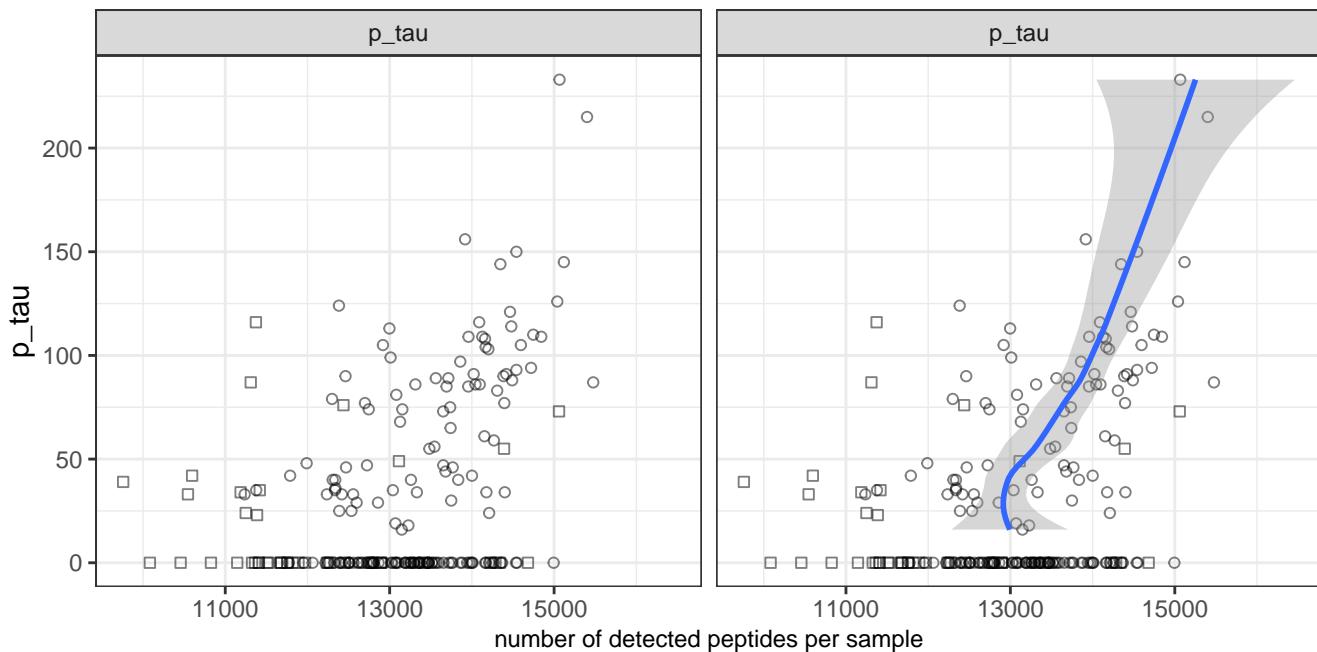
Note that samples flagged as *exclude* are user-provided in the sample metadata table. These are included in data visualizations but excluded from downstream statistical analysis (later part of the report).

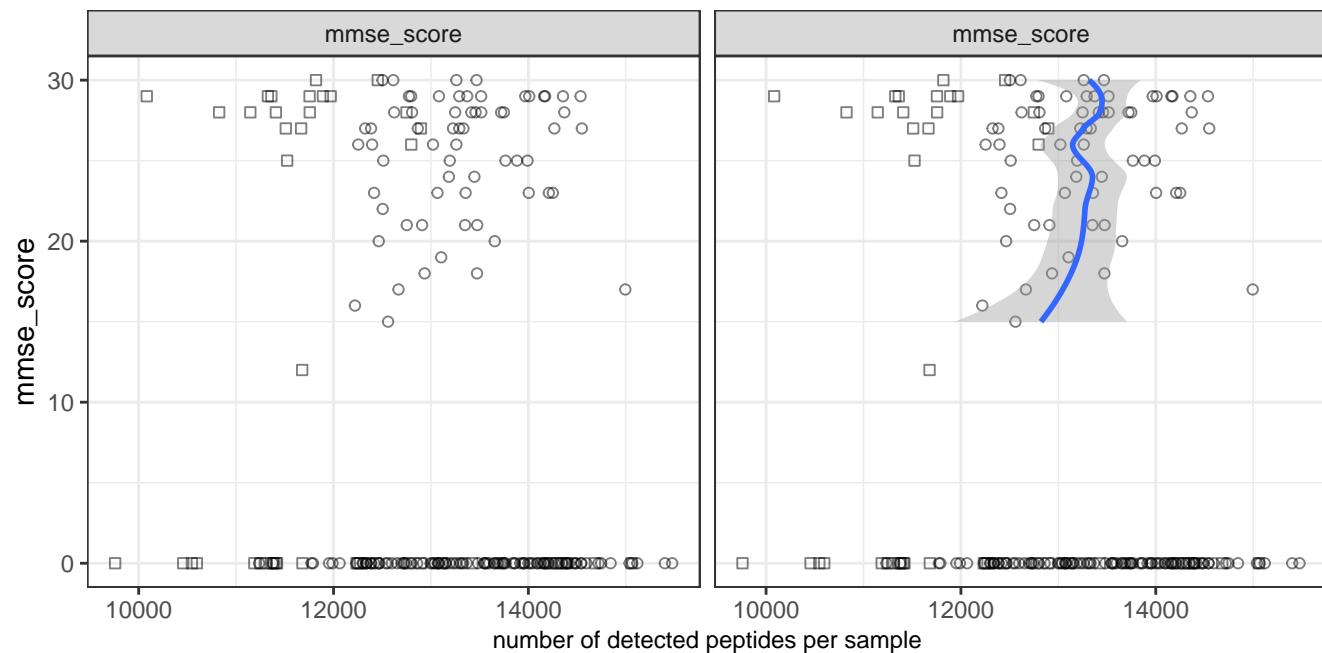
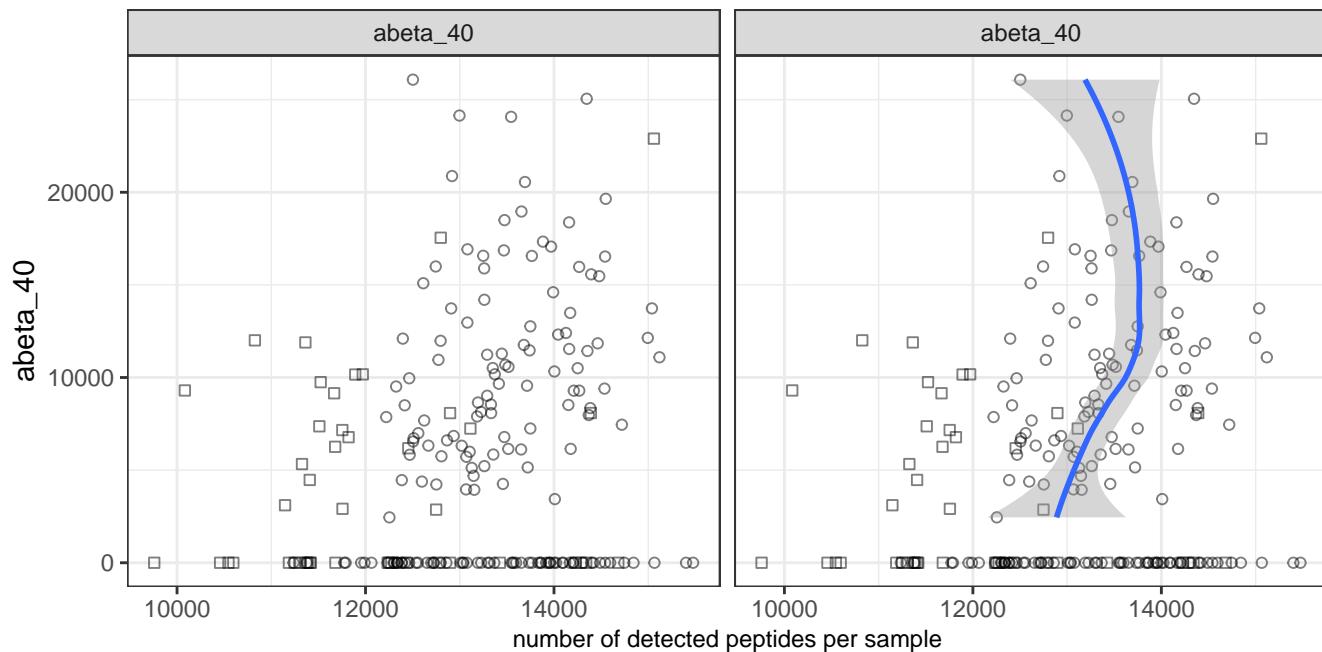
For example: the first plot shows color-coding by the ‘group’ property, so each row represents a sample group. If samples in a particular group systematically yield fewer peptides than another group, a clear pattern will be visible.







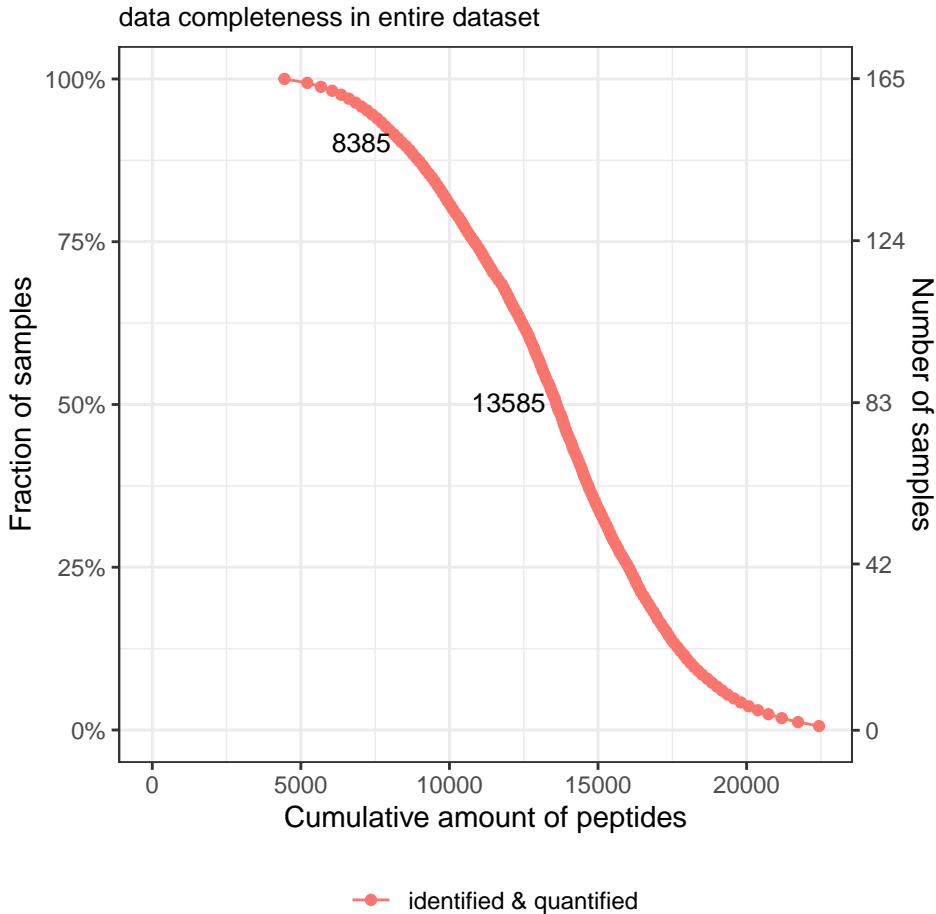




1.3 data completeness

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 off. 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.

1.3.1 cumulative distribution



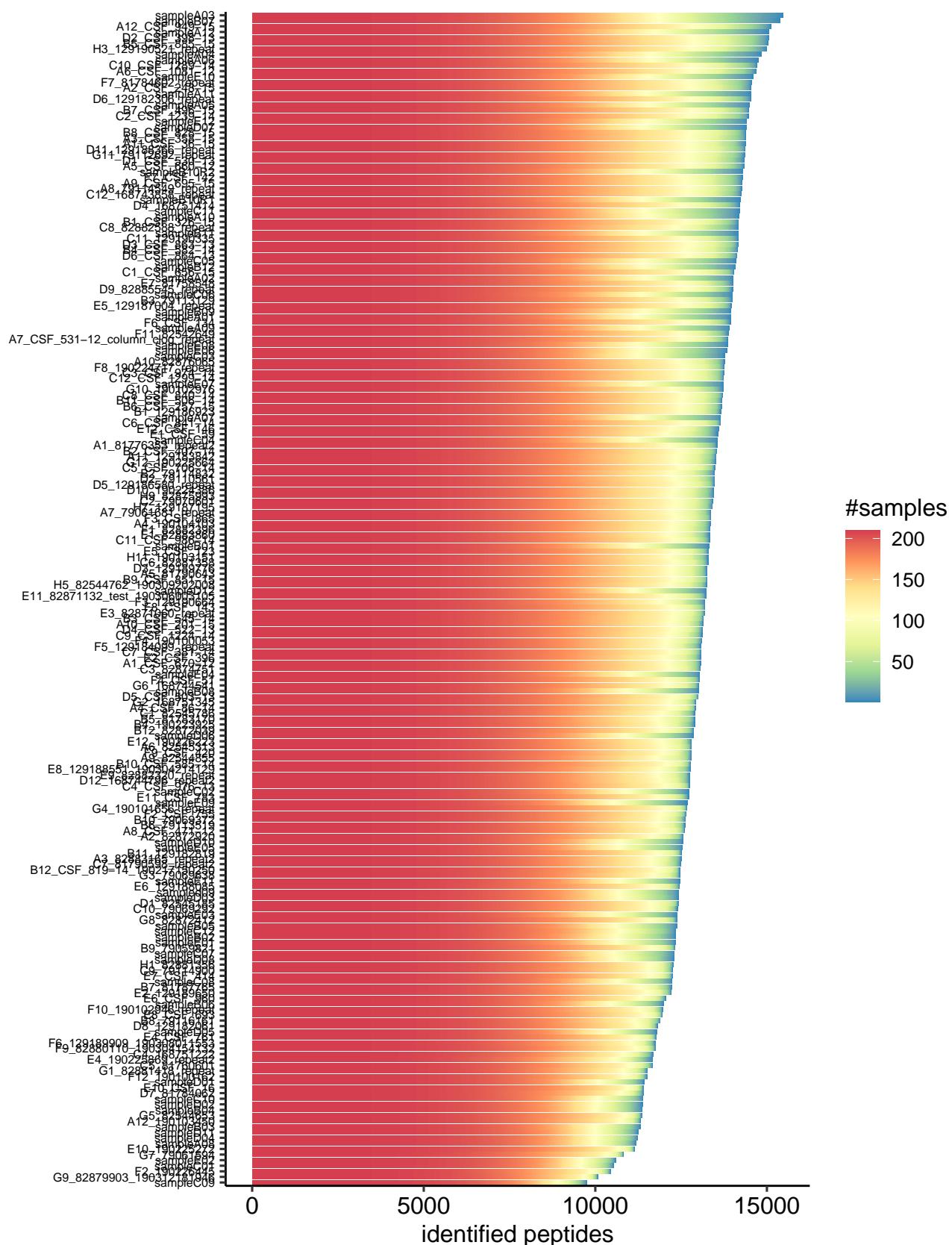
Samples flagged as ‘exclude’ (by user) are not taken into account in this figure. Exact values are shown for data points matching 90% and 50% of samples to convenience comparison between analyses (e.g. before/after configuring ‘exclude’ samples, or comparing between experiments of similar protocol).

1.3.2 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).

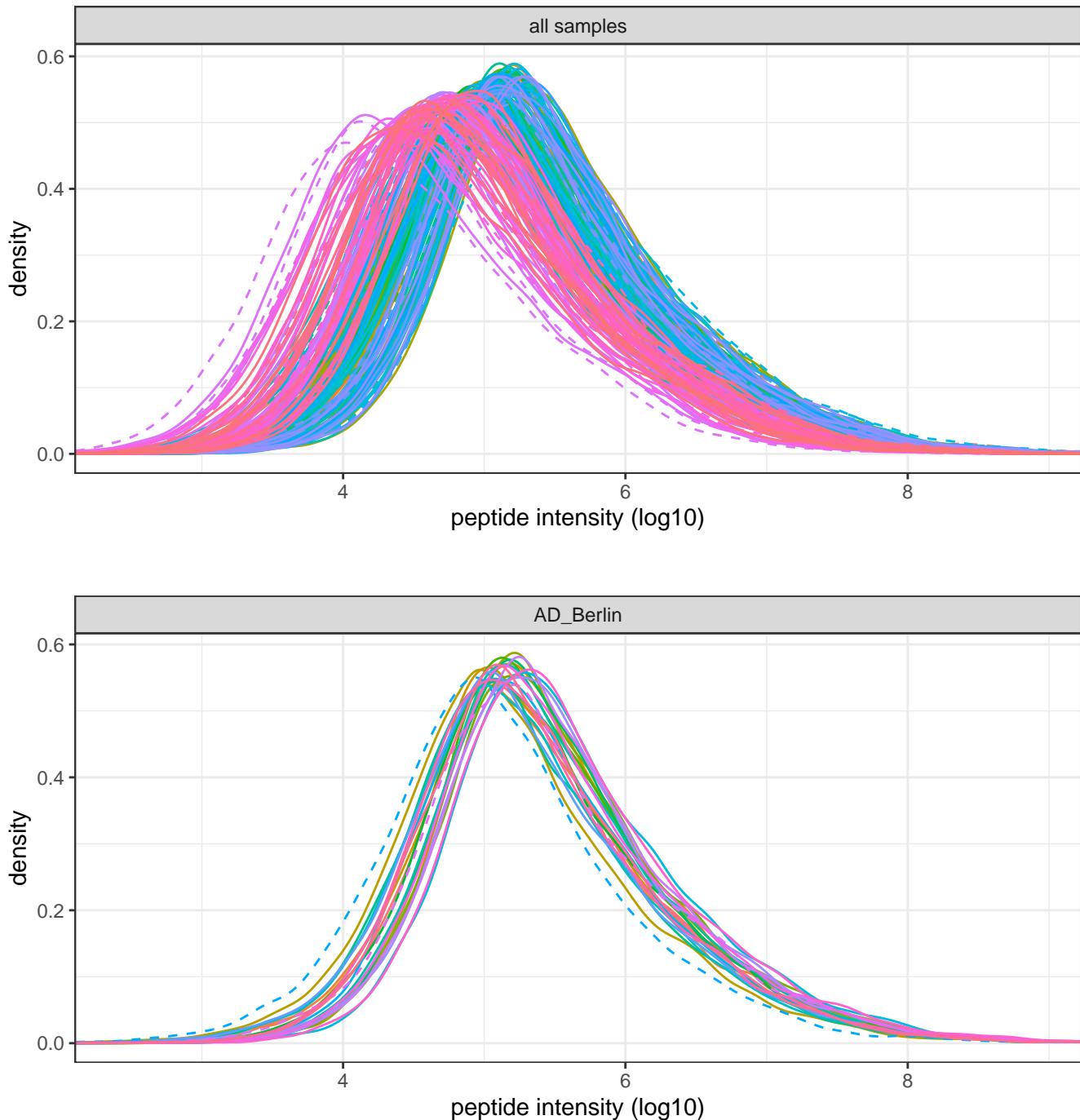
Optimally, the majority of peptides in each sample are red~orange with relatively few uniquely identified peptides (blue~green). Samples are sorted by the total amount of detected peptides.

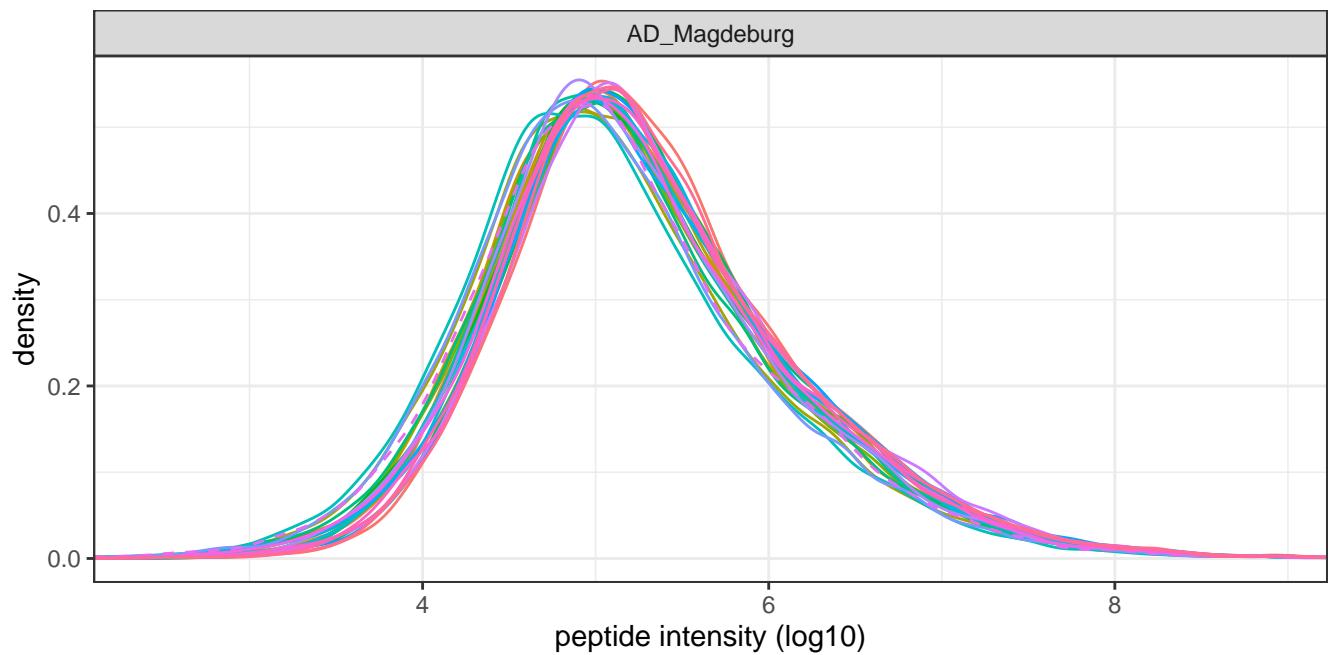
Number of samples in which a peptide is identified vs presence in individual sample

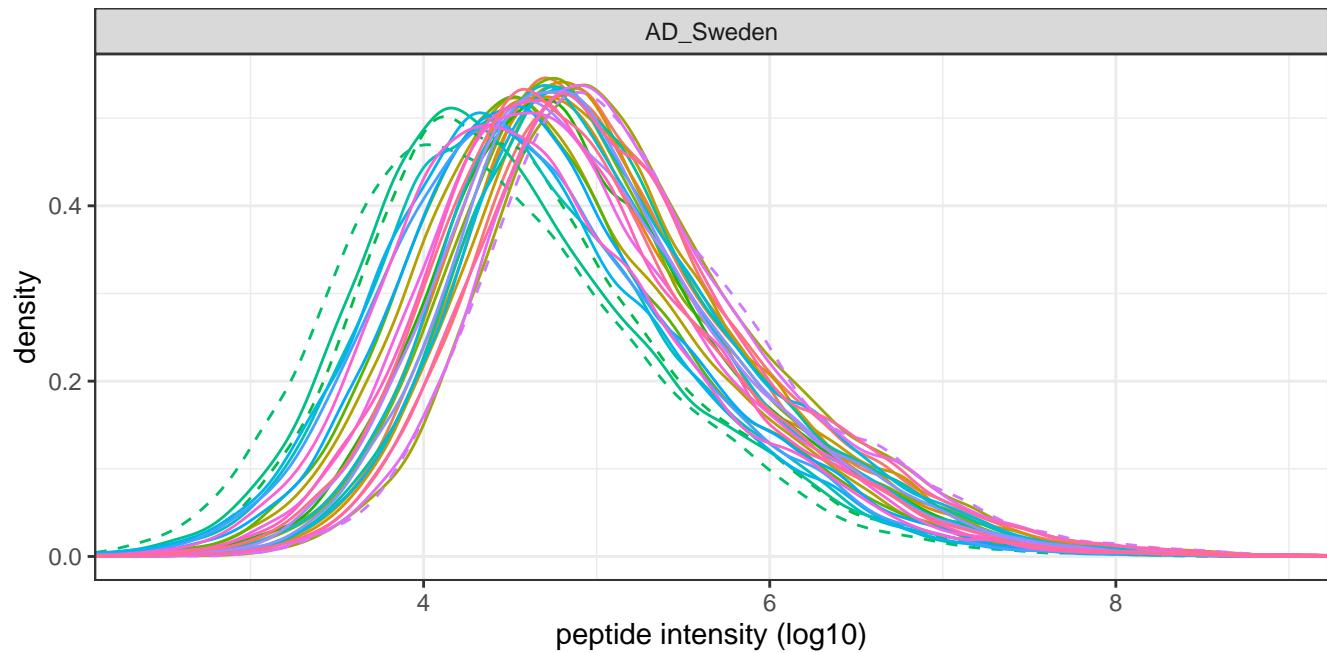
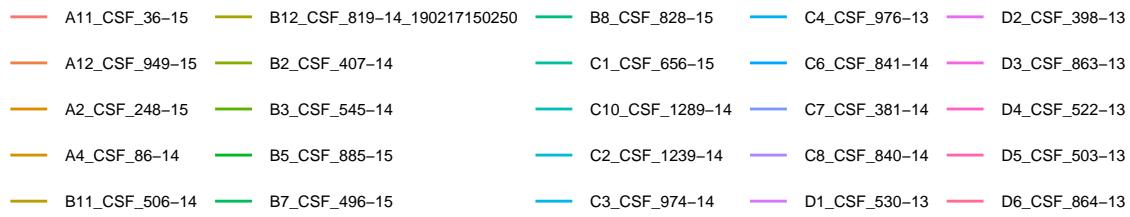


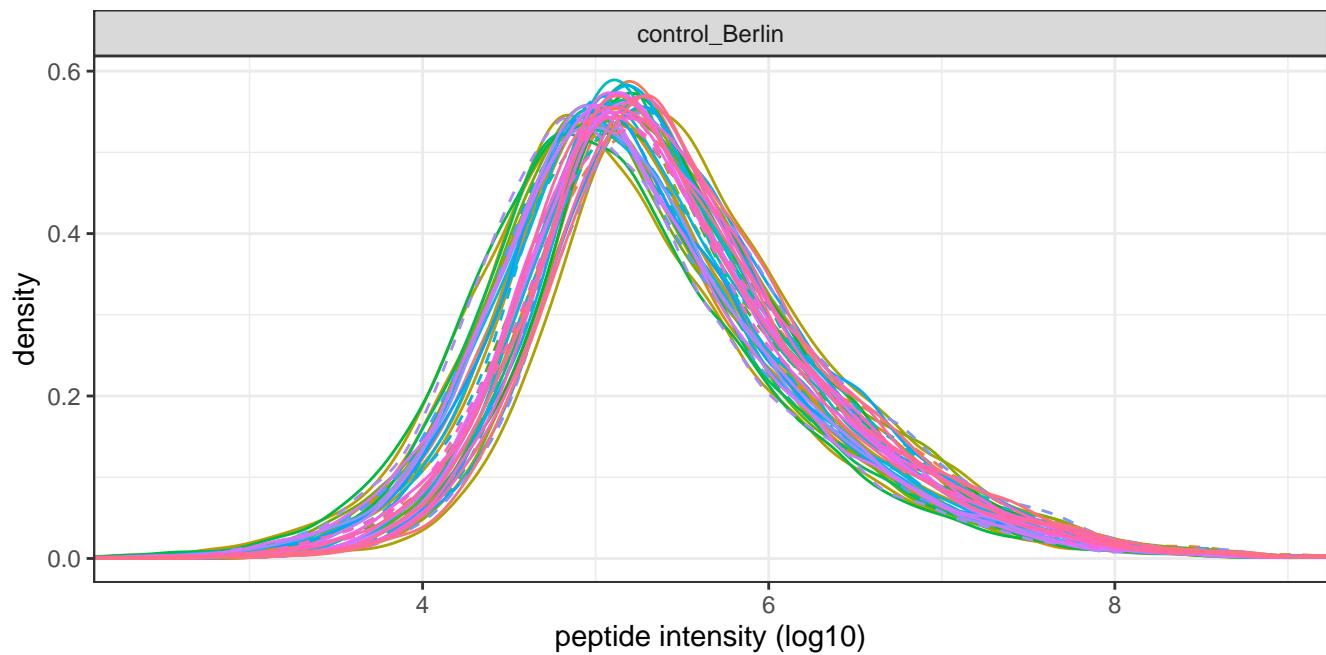
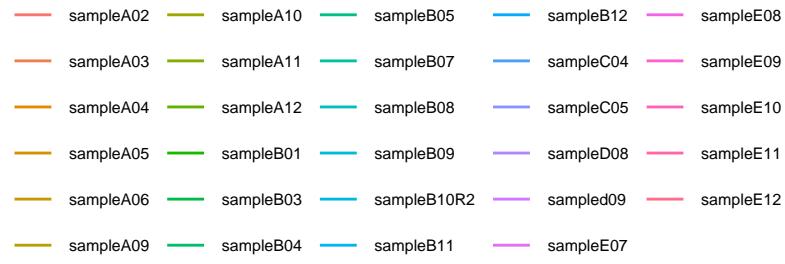
1.4 abundance distributions

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.

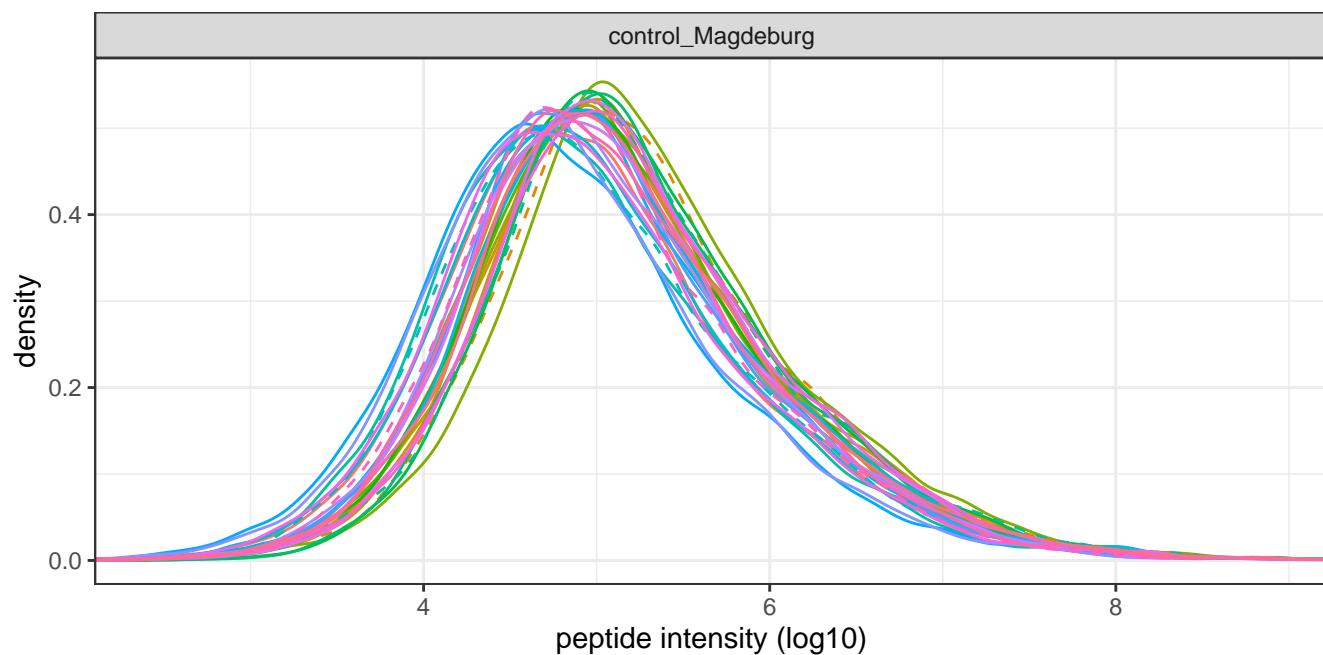


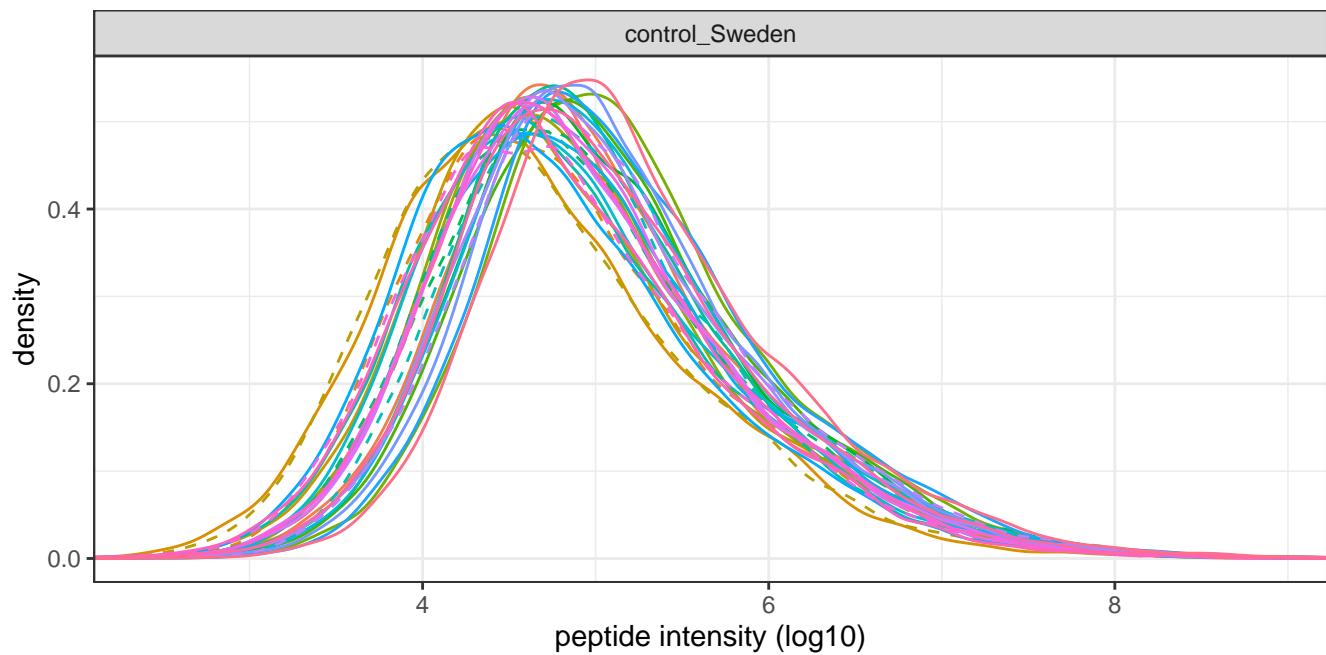
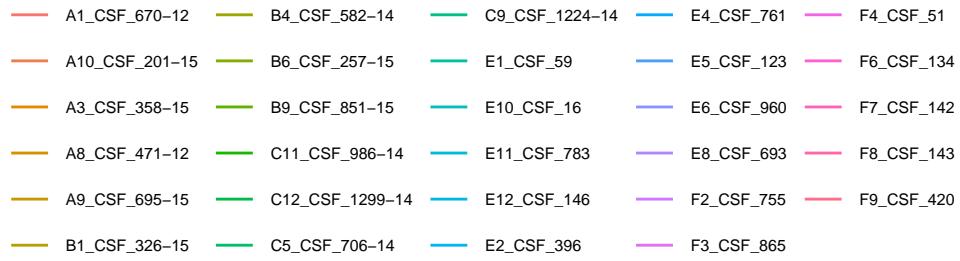


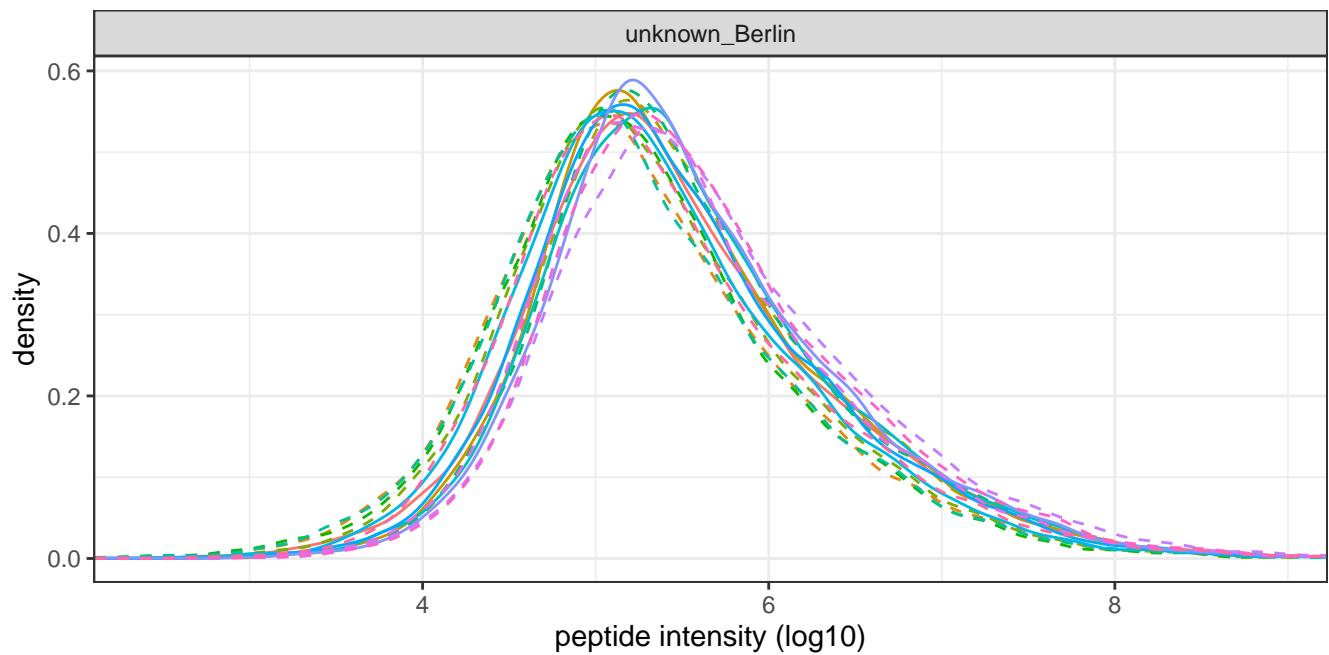




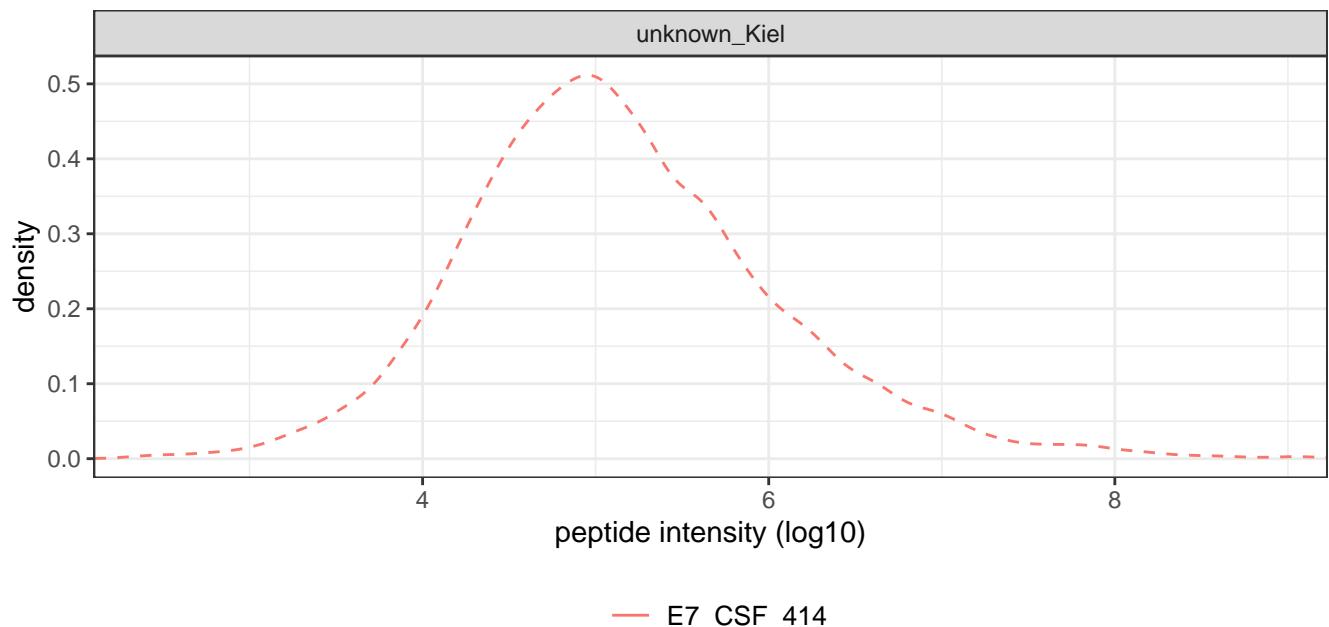
A10_82876065	B8_79116161	D12_168744796_repeat2	E7_81758548	G11_79112692_repeat
A11_129183842	B9_79059821	D6_129182306_repeat	E8_129188551_190304214129	G12_190225664
A12_190103450	C10_79069292	D7_81784062	E9_82882320_repeat	G6_168744541
A5_81790643	C11_129190335	D8_129182061	F10_190102946_repeat	G7_79061594
A7_79061681_repeat	C12_168743858_repeat	D9_82885545_repeat	F12_190100162	G8_82872412
A8_79114549_repeat	C5_81760601	E10_190225272	F5_129184099_repeat	G9_82879903_1903121815
A9_82544855	C6_82881358	E11_82871132_test_190306003105	F6_129189909_190308011553	H11_190103151
B10_79069372	C7_81790598_repeat2	E12_190226223	F7_81784602_repeat	H5_82544762_190309202C
B12_82872039	C9_79114900	E5_129187004_repeat	F8_190224717_repeat	H7_129187195
B6_79113519	D10_190224386	E6_129188085	G10_190102976	H9_82875993

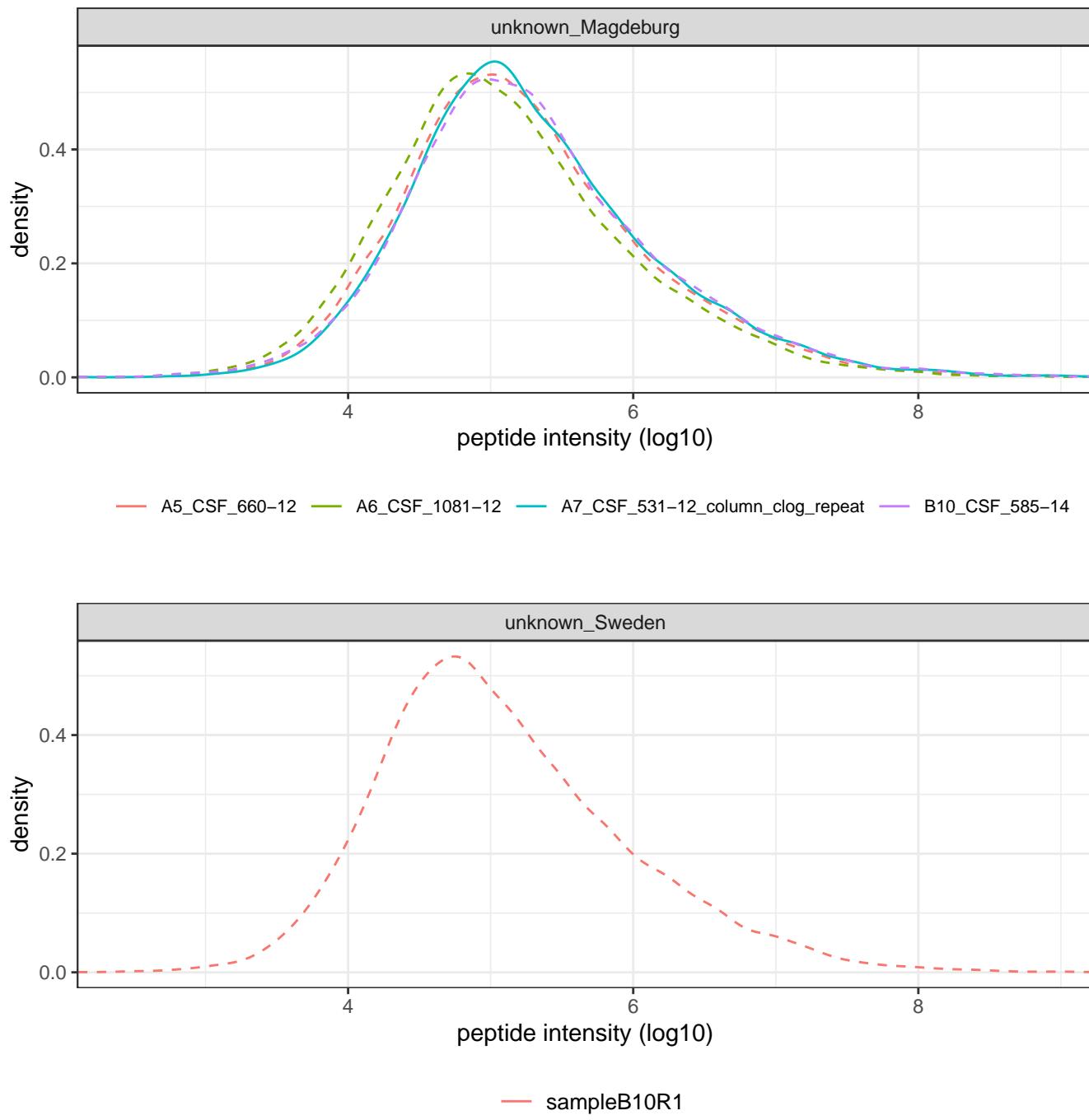






— A1_81776353_repeat2 — B5_81783170 — C8_82882588_repeat — F2_190226445
— A6_82545313 — B7_81787765 — D11_129188366_repeat — F9_82880110_190304154132
— B11_129182819 — C2_79070601 — D5_129186580_repeat — G5_82544653
— B4_190223925 — C4_168751222 — F11_82542649 — H1_82881356



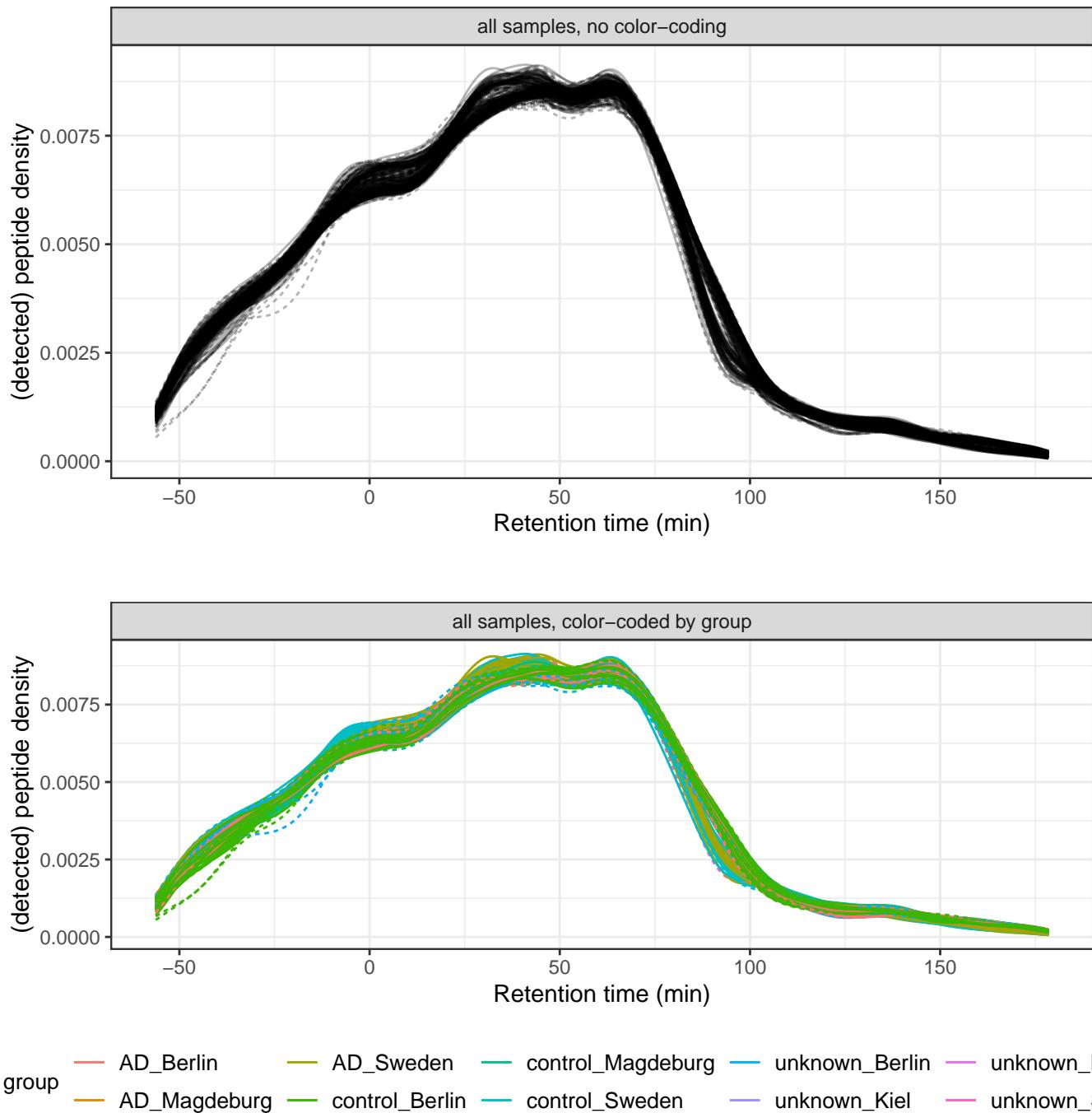


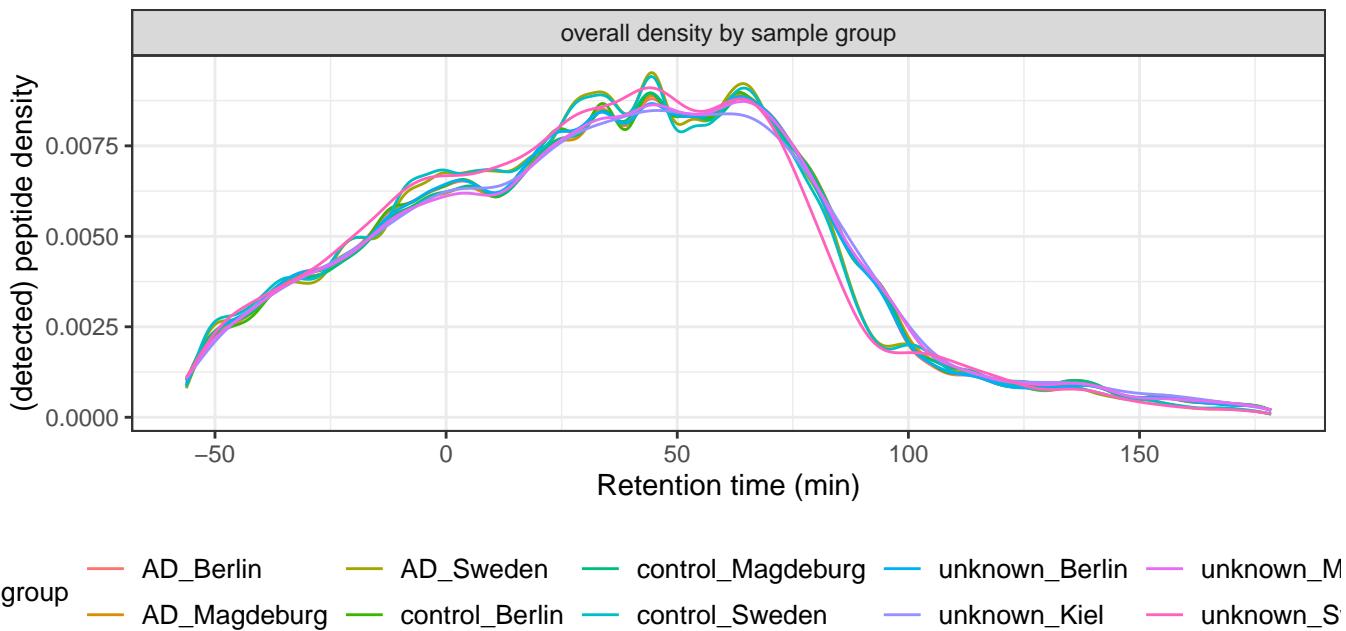
1.5 retention time

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.

1.5.1 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.



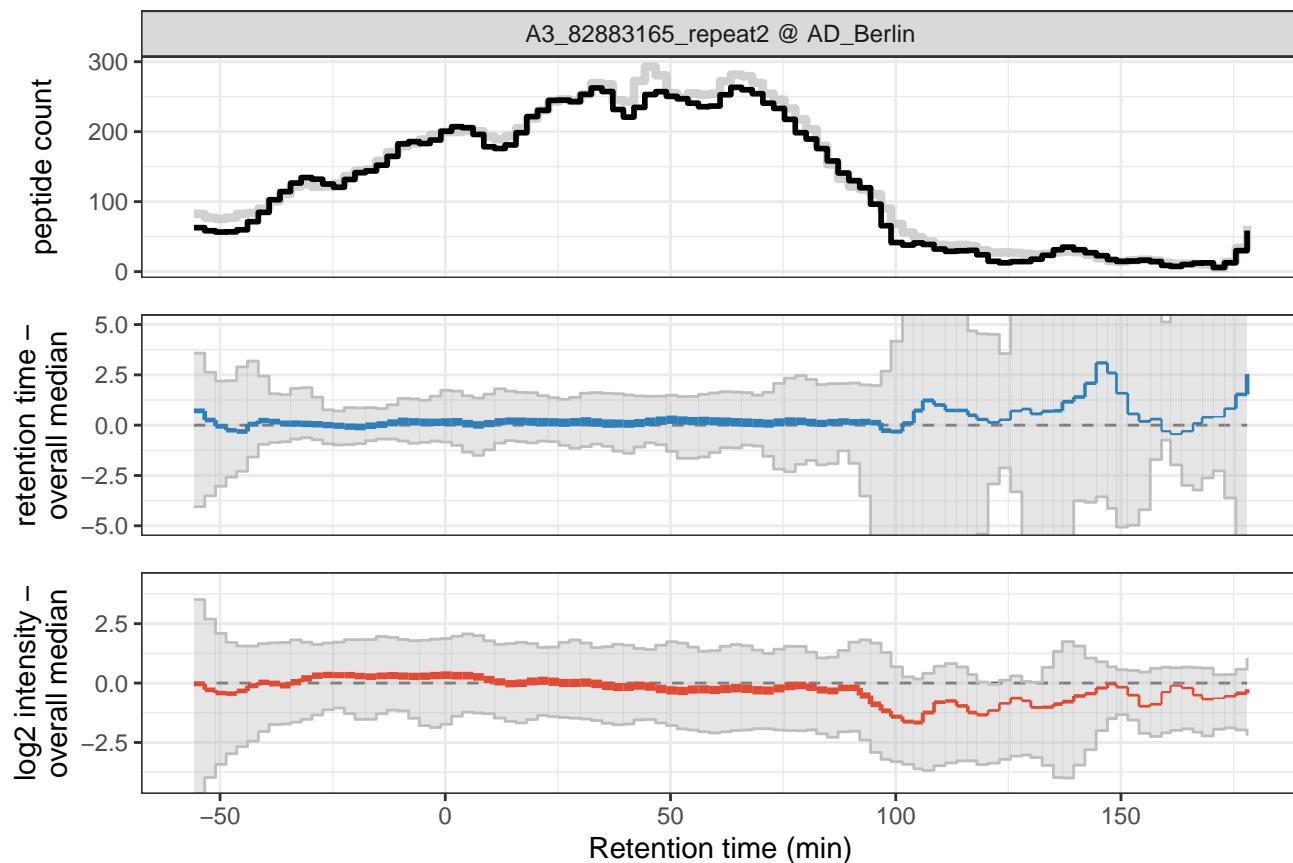
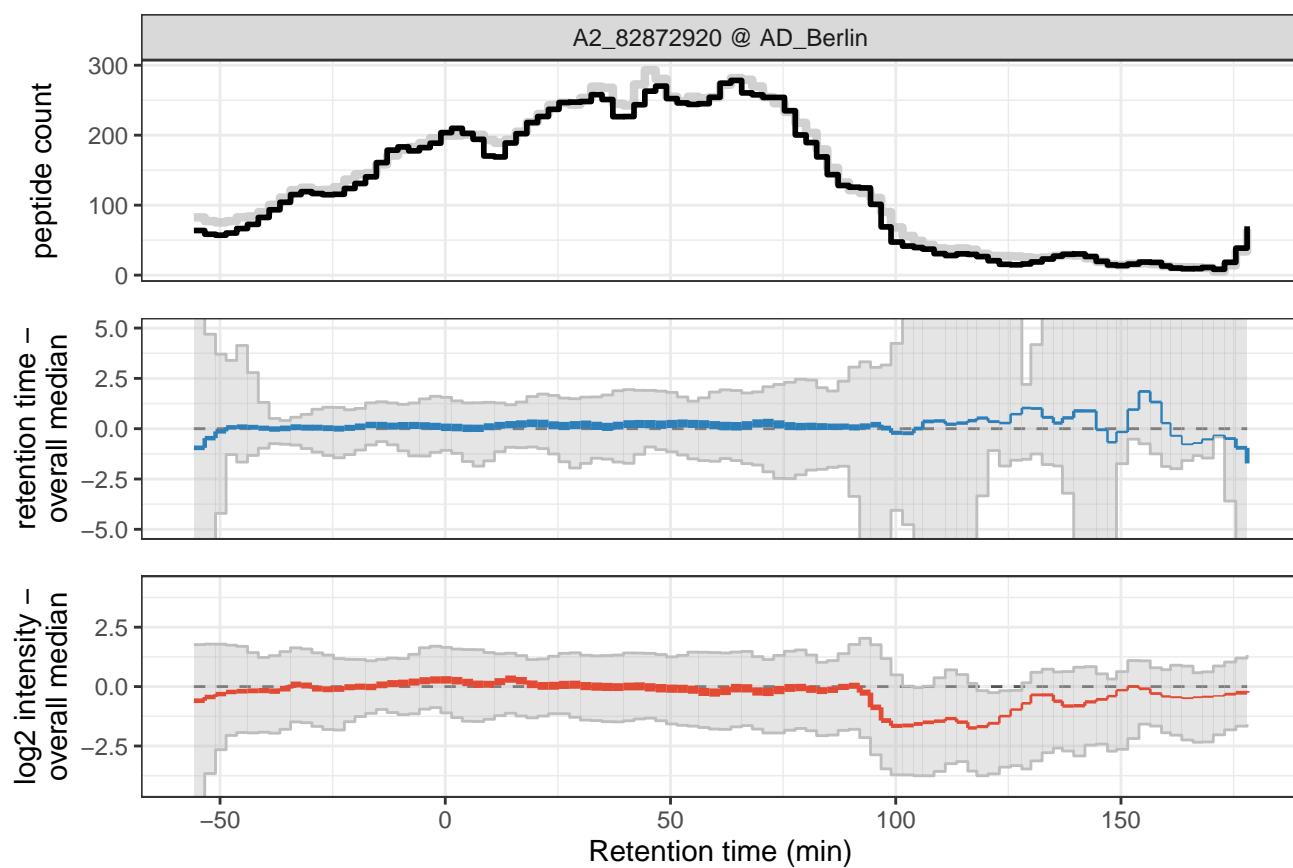


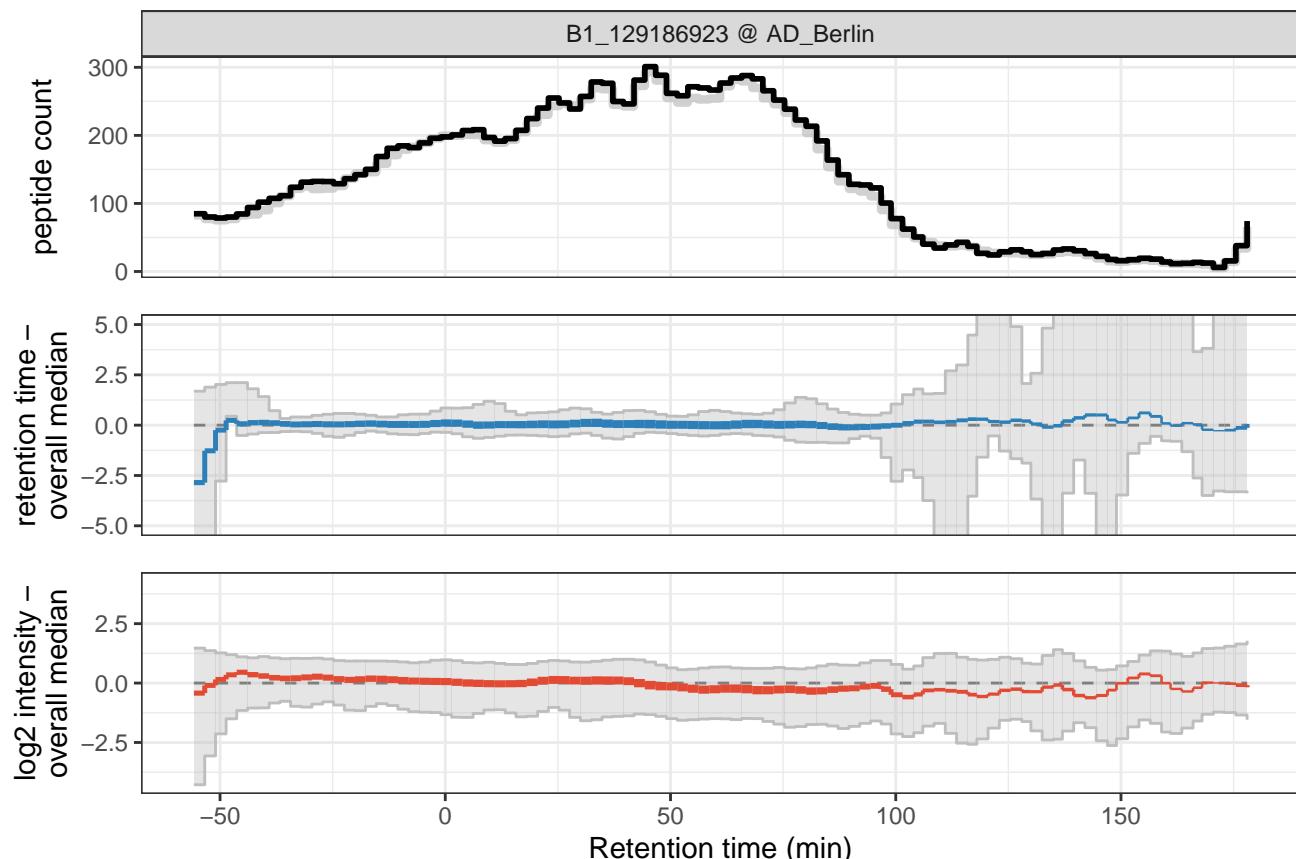
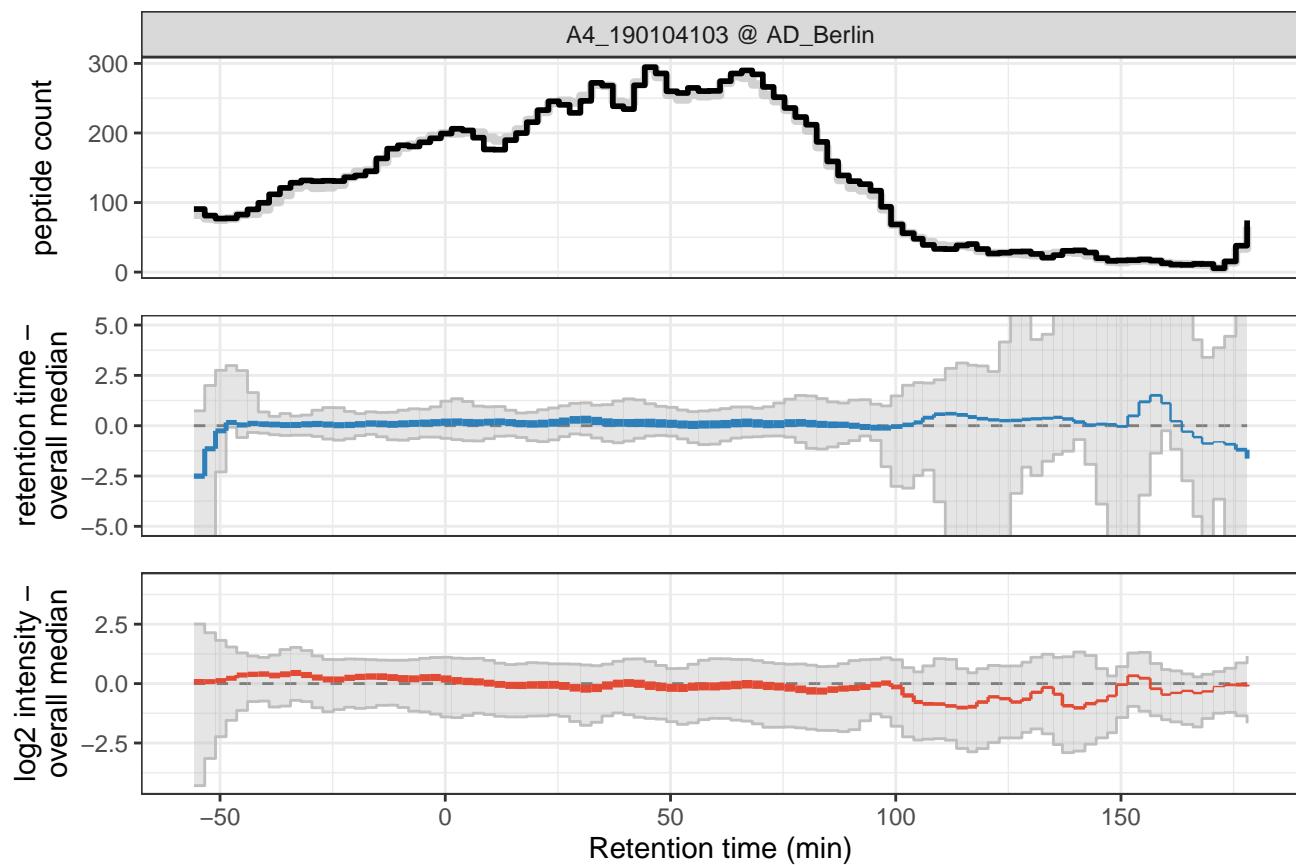
1.5.2 retention time local effects

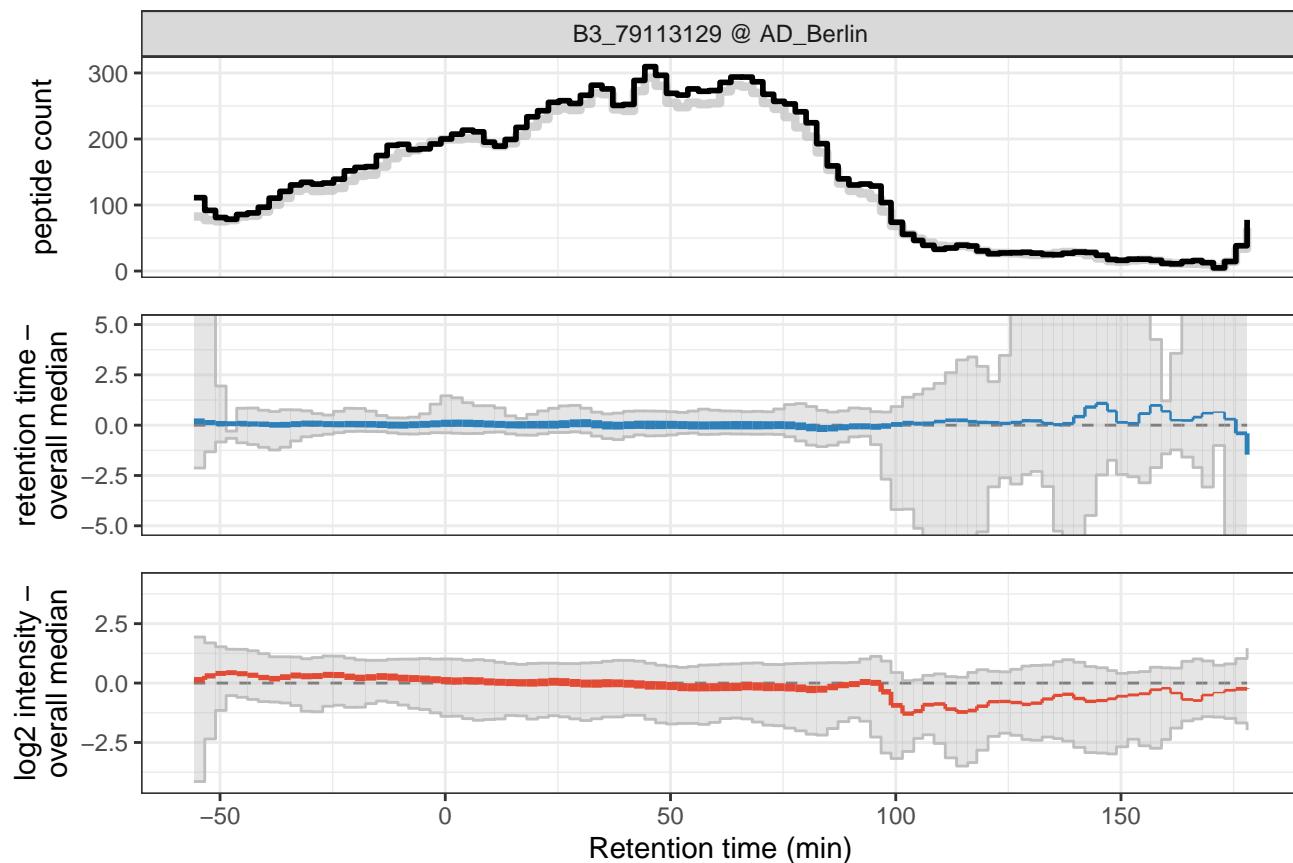
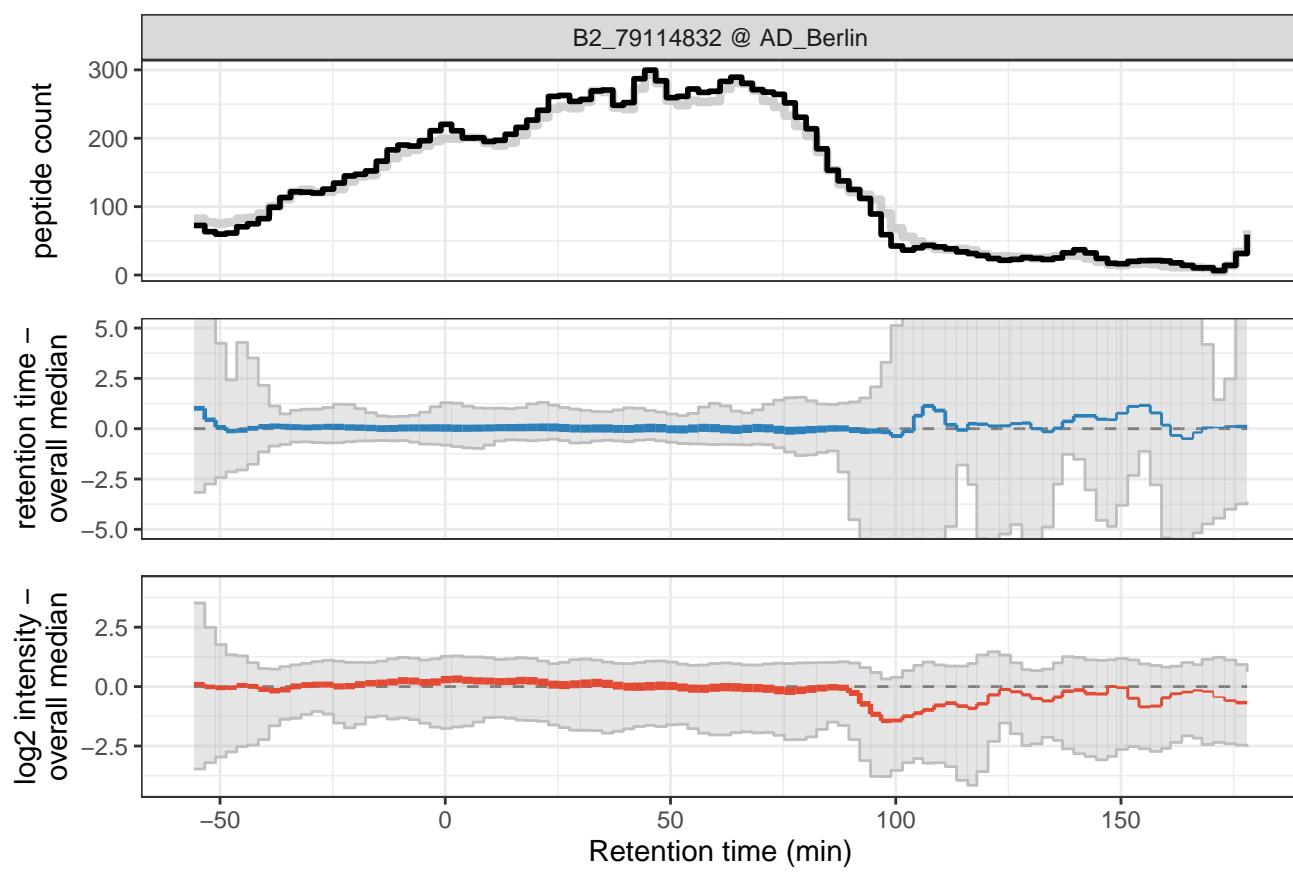
To investigate how each measurement differs from others, we visualize each sample as a 3 panel figure. First, the data is binned across the retention time dimension (x-axis). If a samples was marked as ‘exclude’ in the provided sample metadata, this is indicated in the plot title.

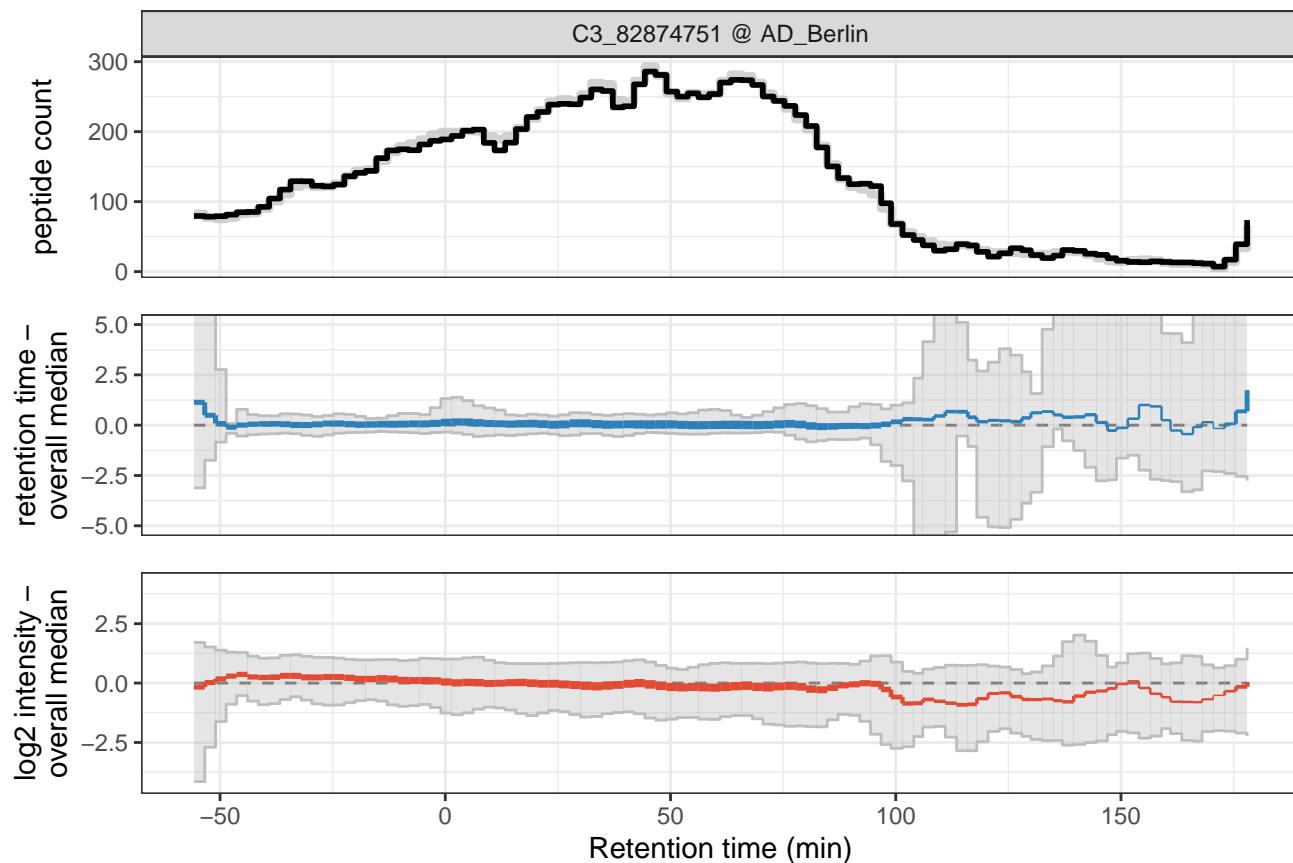
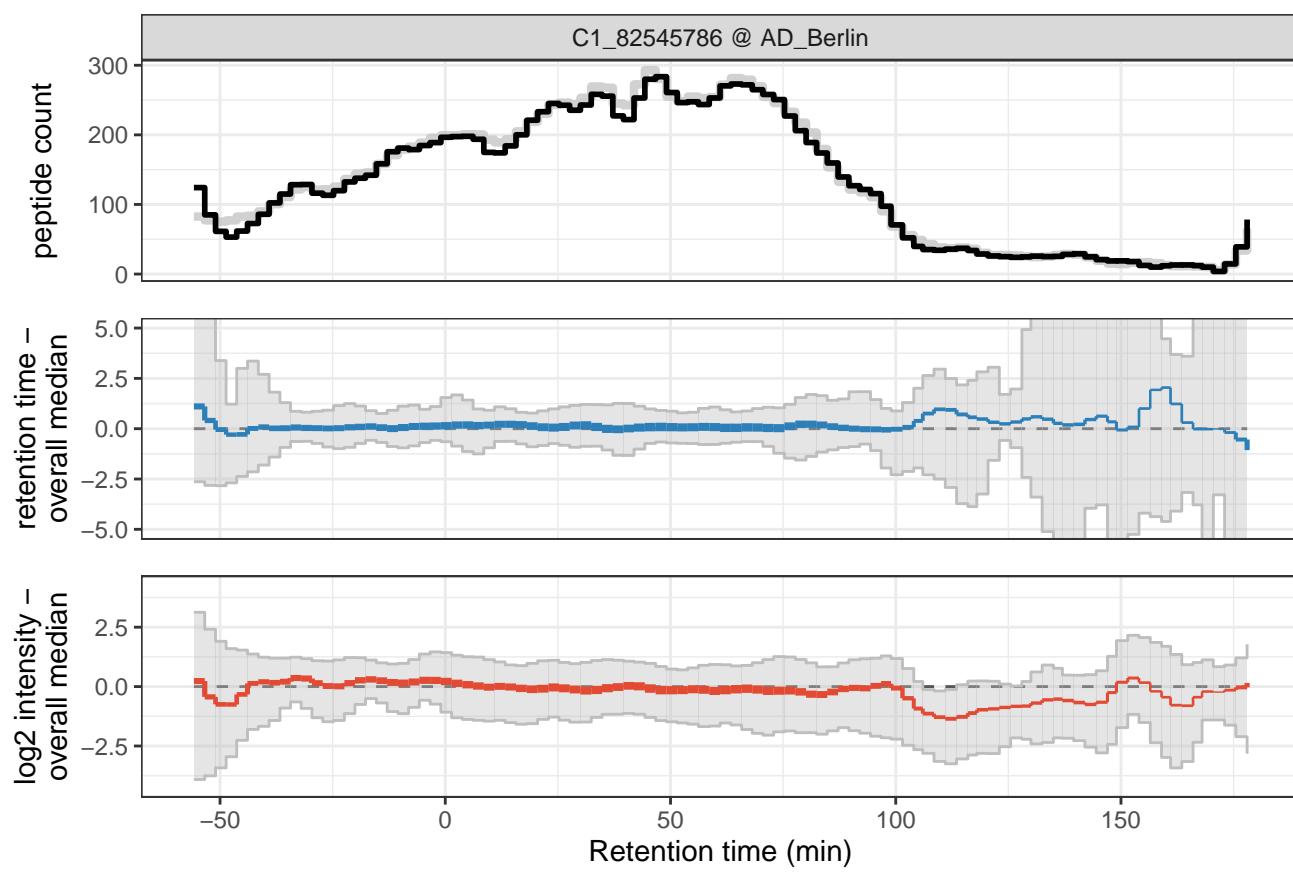
The top panel shows the number of peptides in the input data, e.g. as recognized by the software that generated input for this pipeline, over time (black line). For reference, the grey line shows the median amount over all samples (note; if this is the exact same in all samples, the grey line may not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

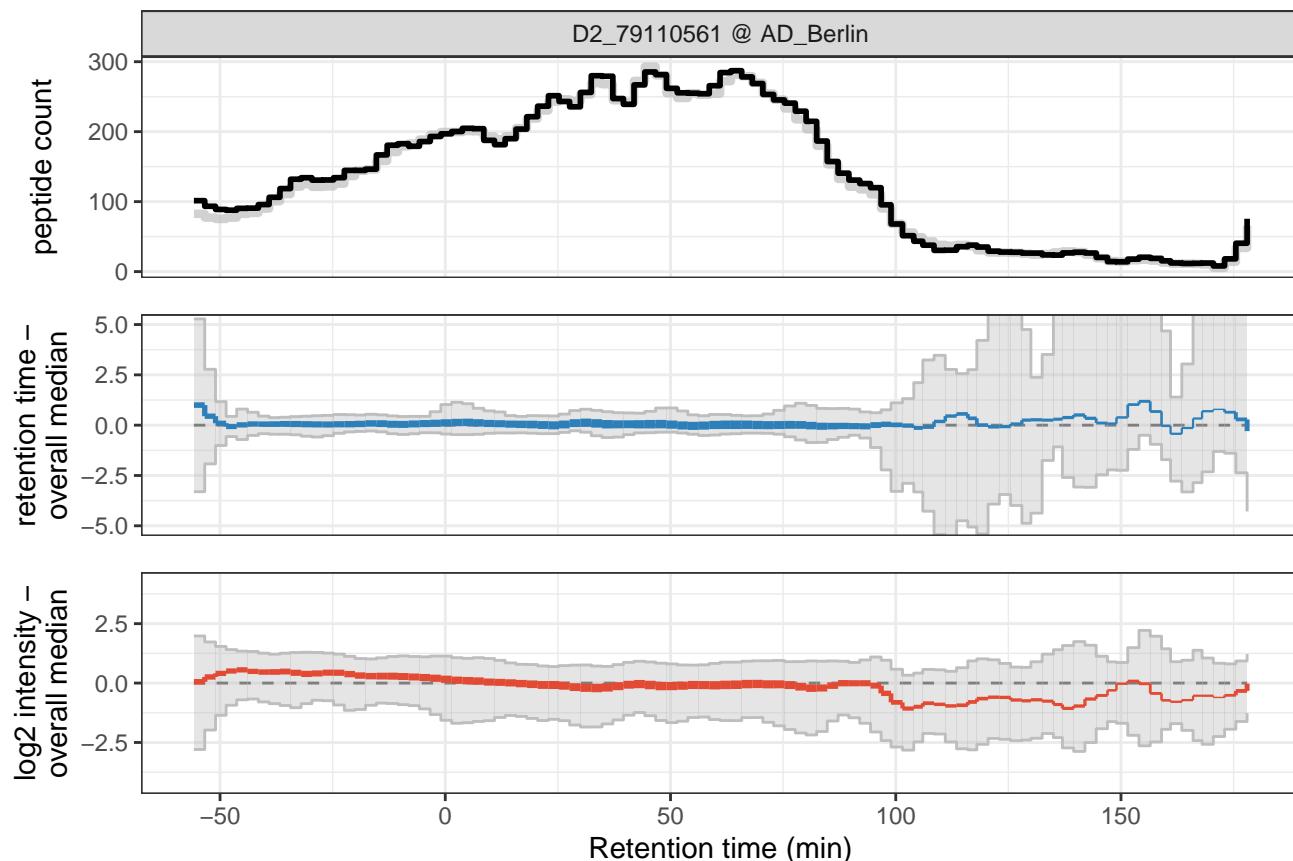
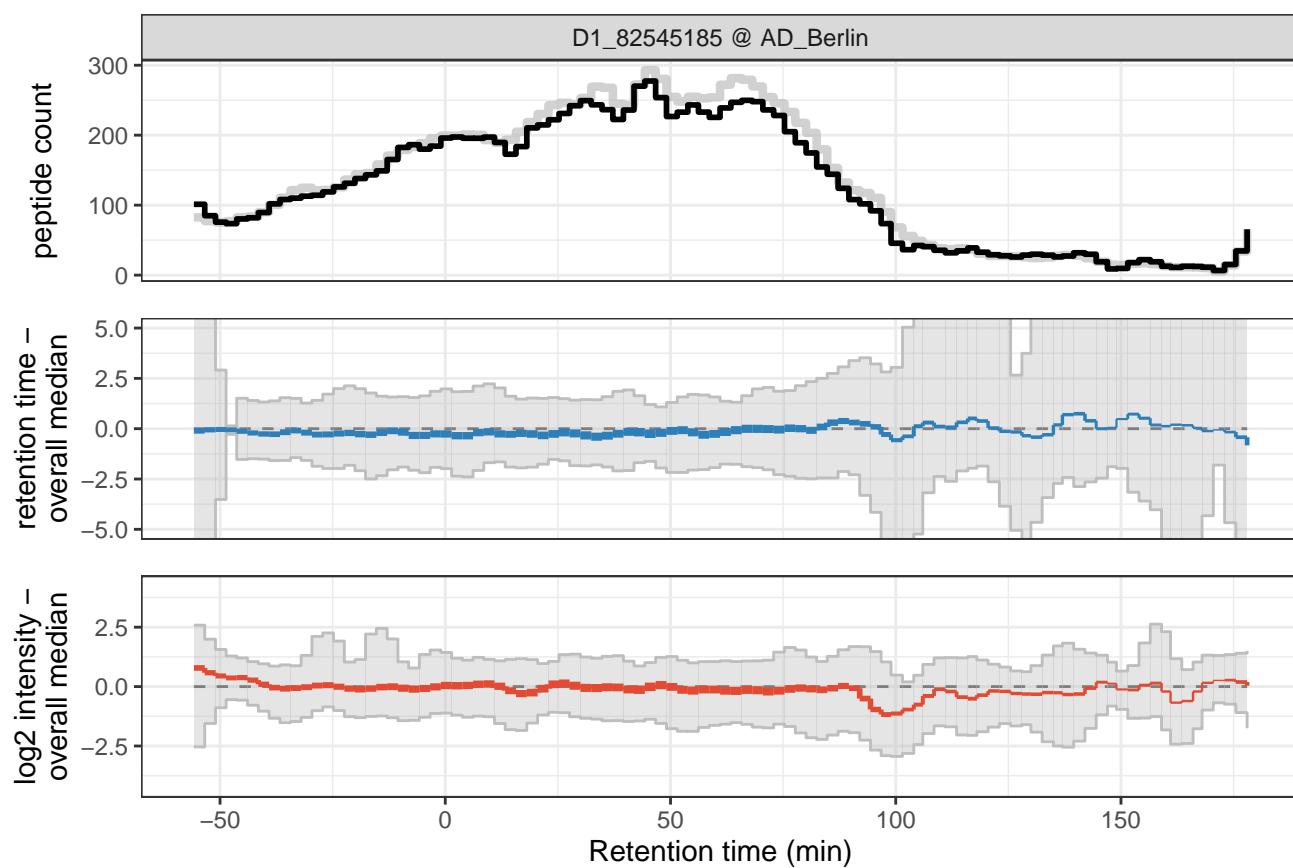
The middle panel indicates whether peptide retention times deviate from their median over all samples (blue line). The grey area depicts the 5% and 95% quantiles, respectively. The line width corresponds to the number of peptides eluting at that time (data from first panel). Analogously, the bottom panel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

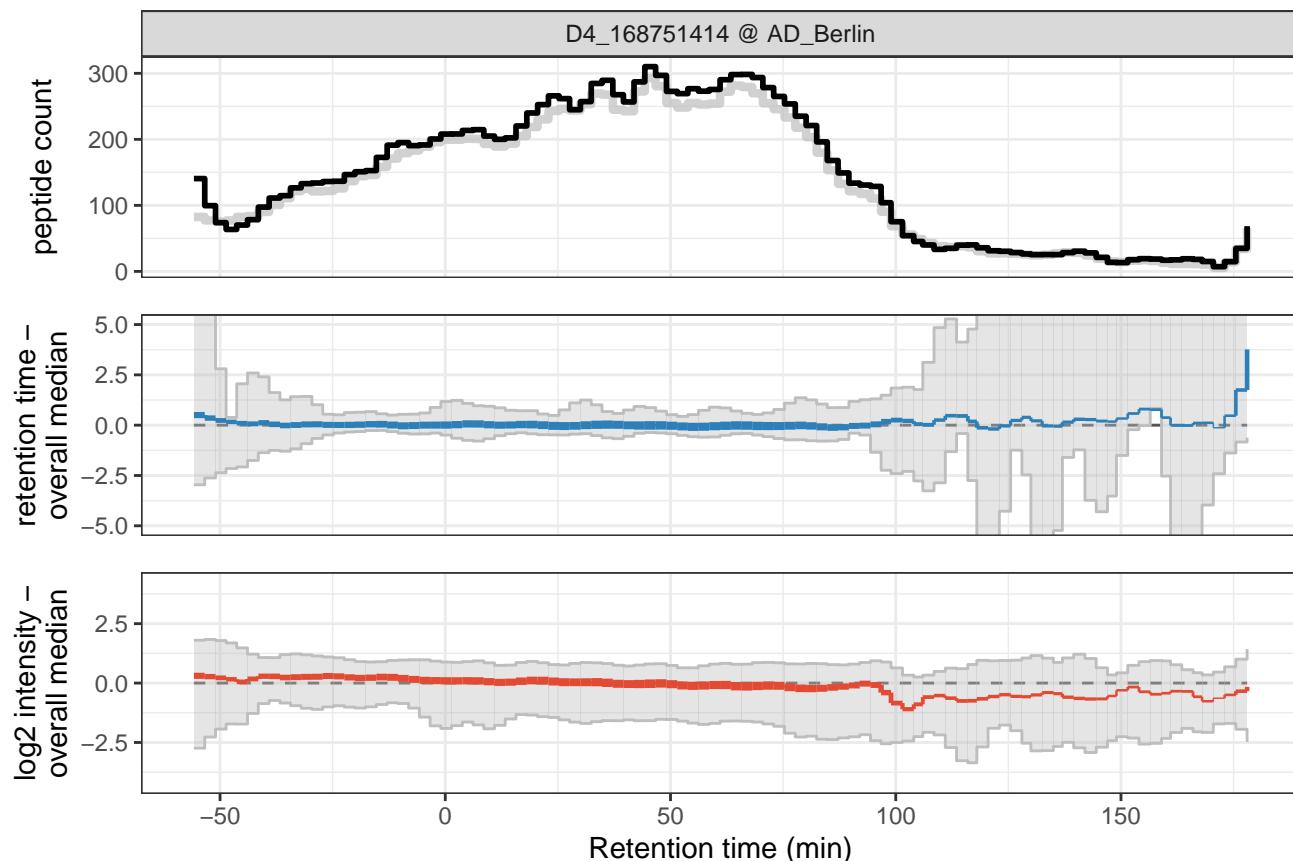
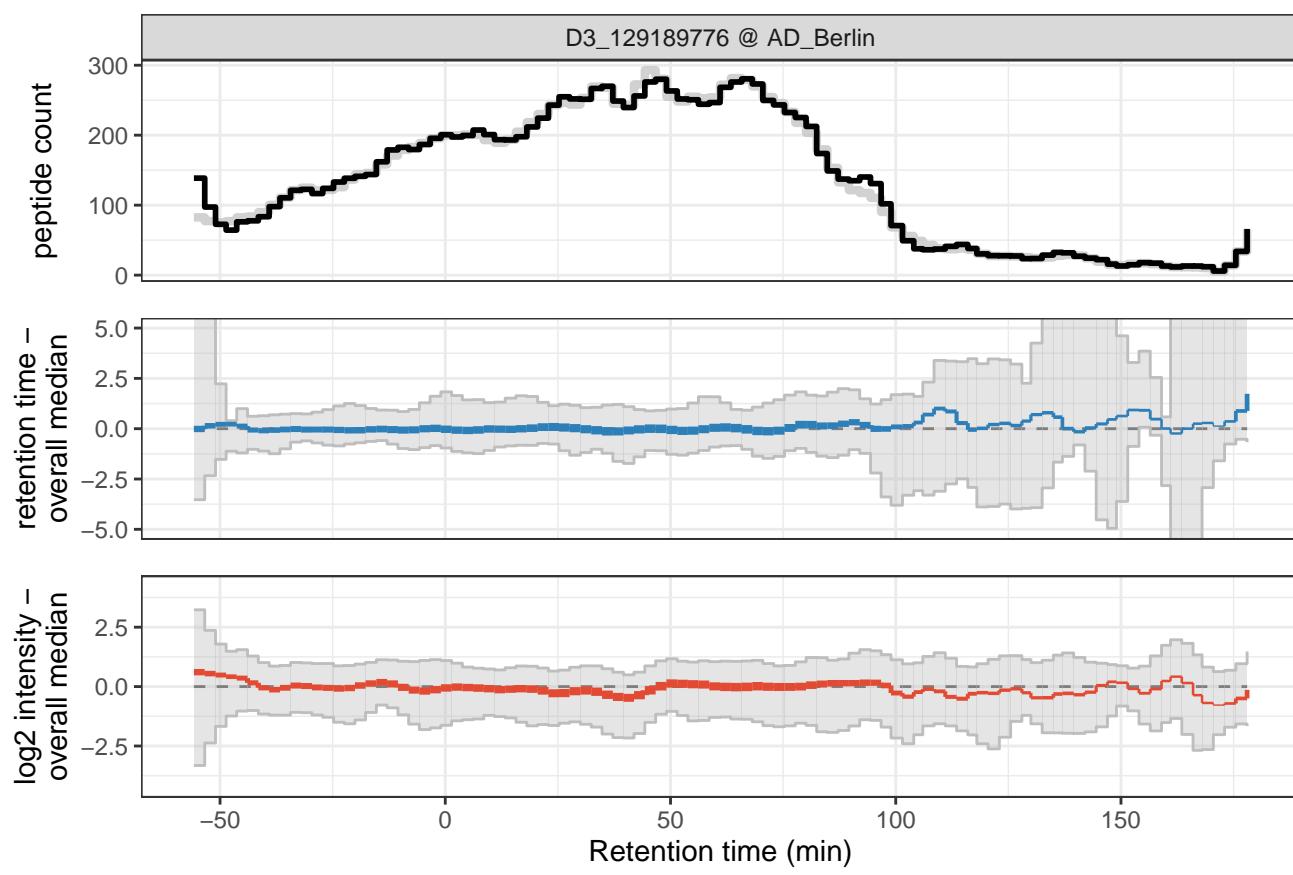


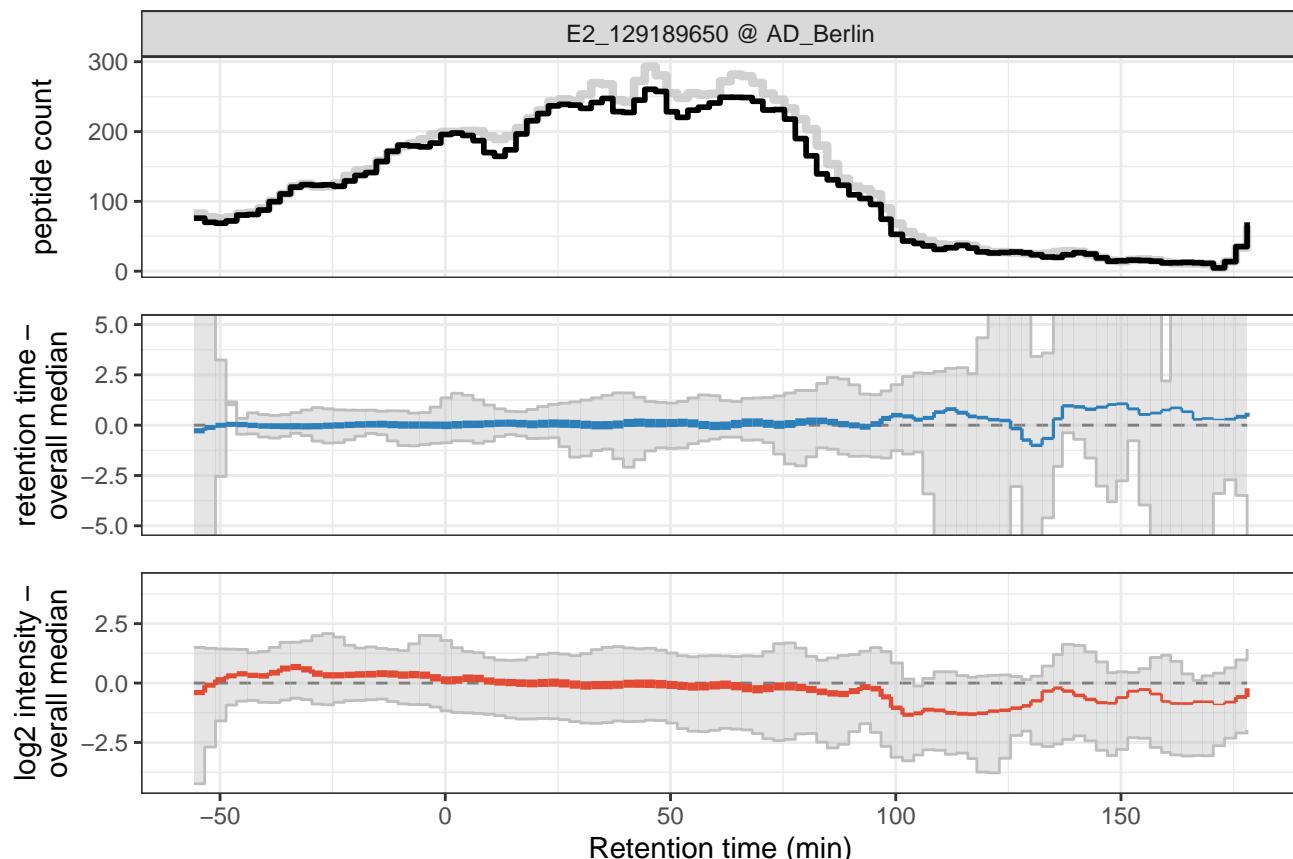
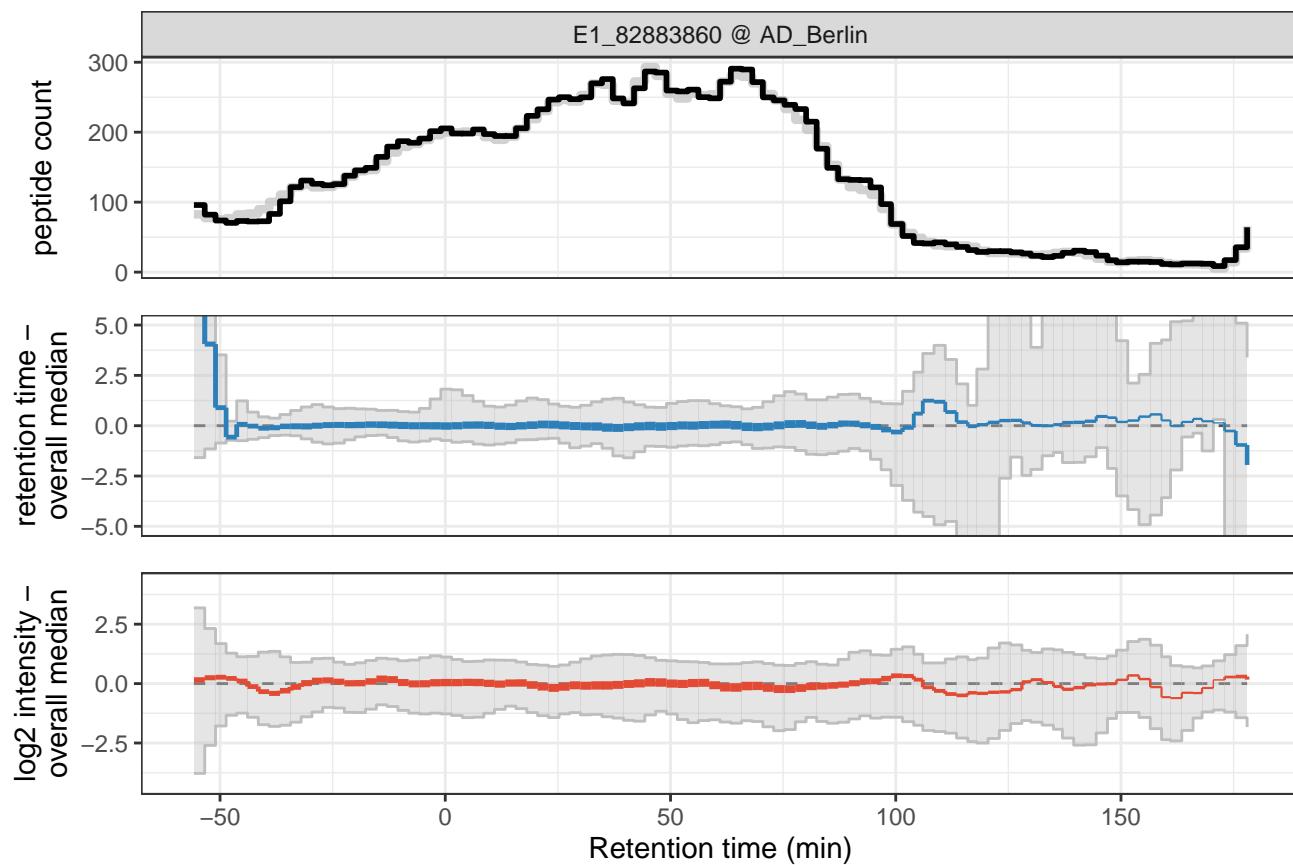


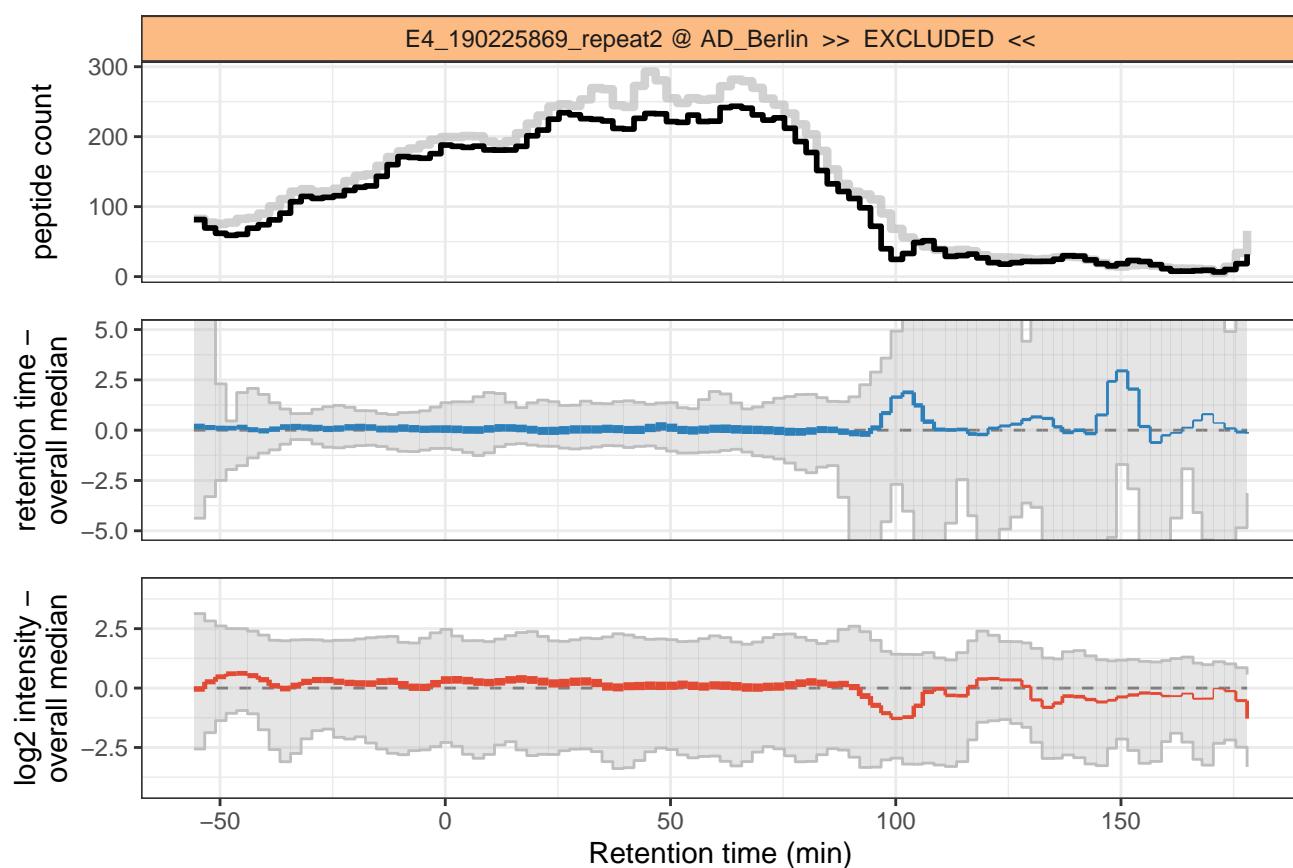
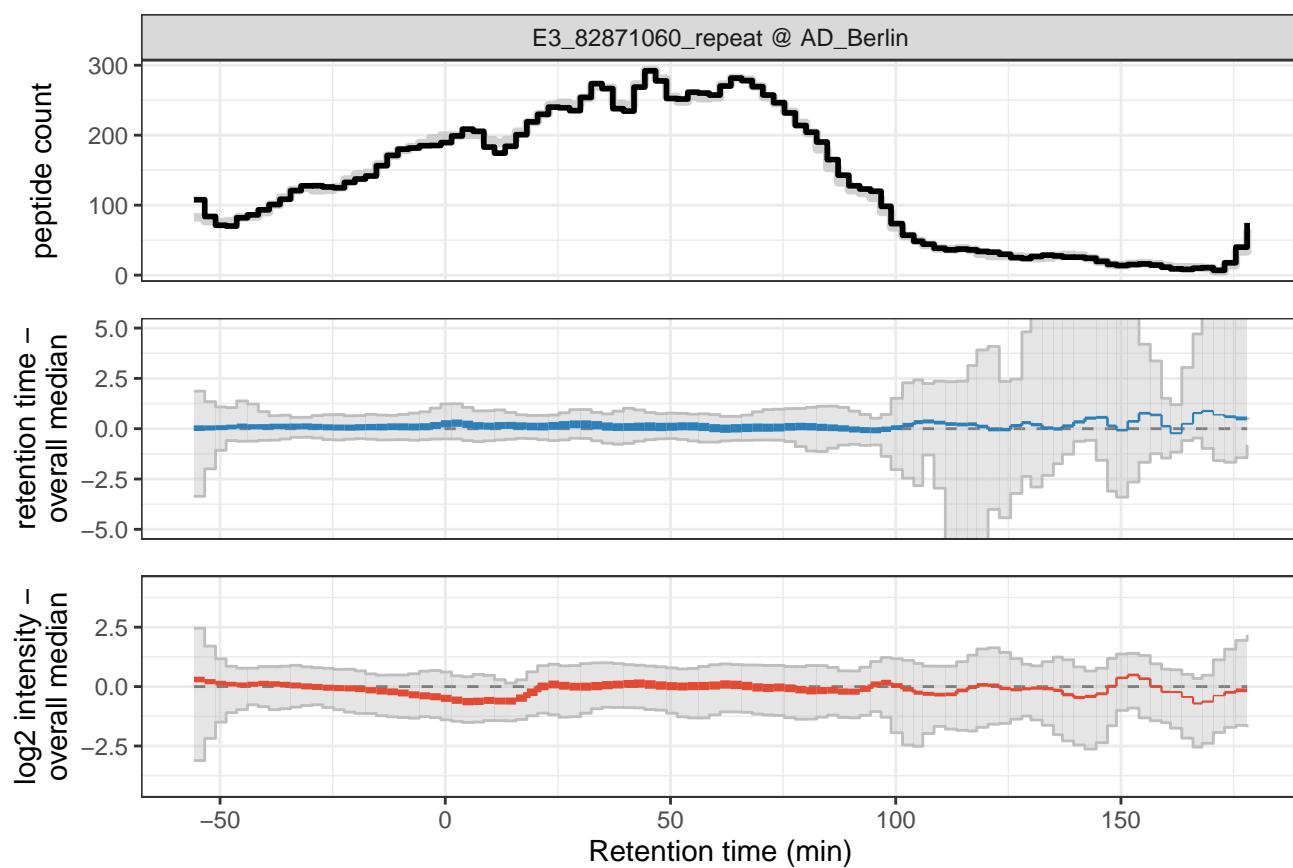


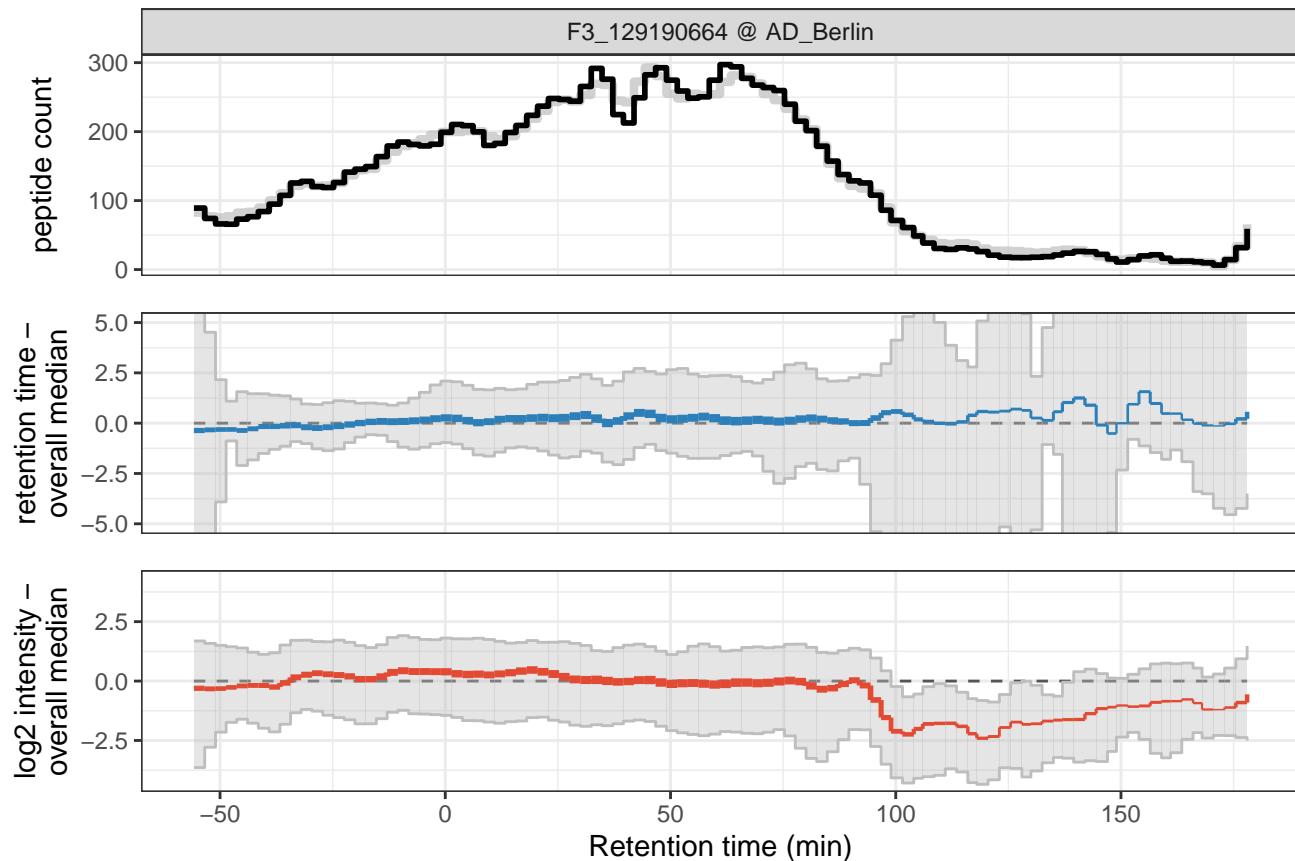
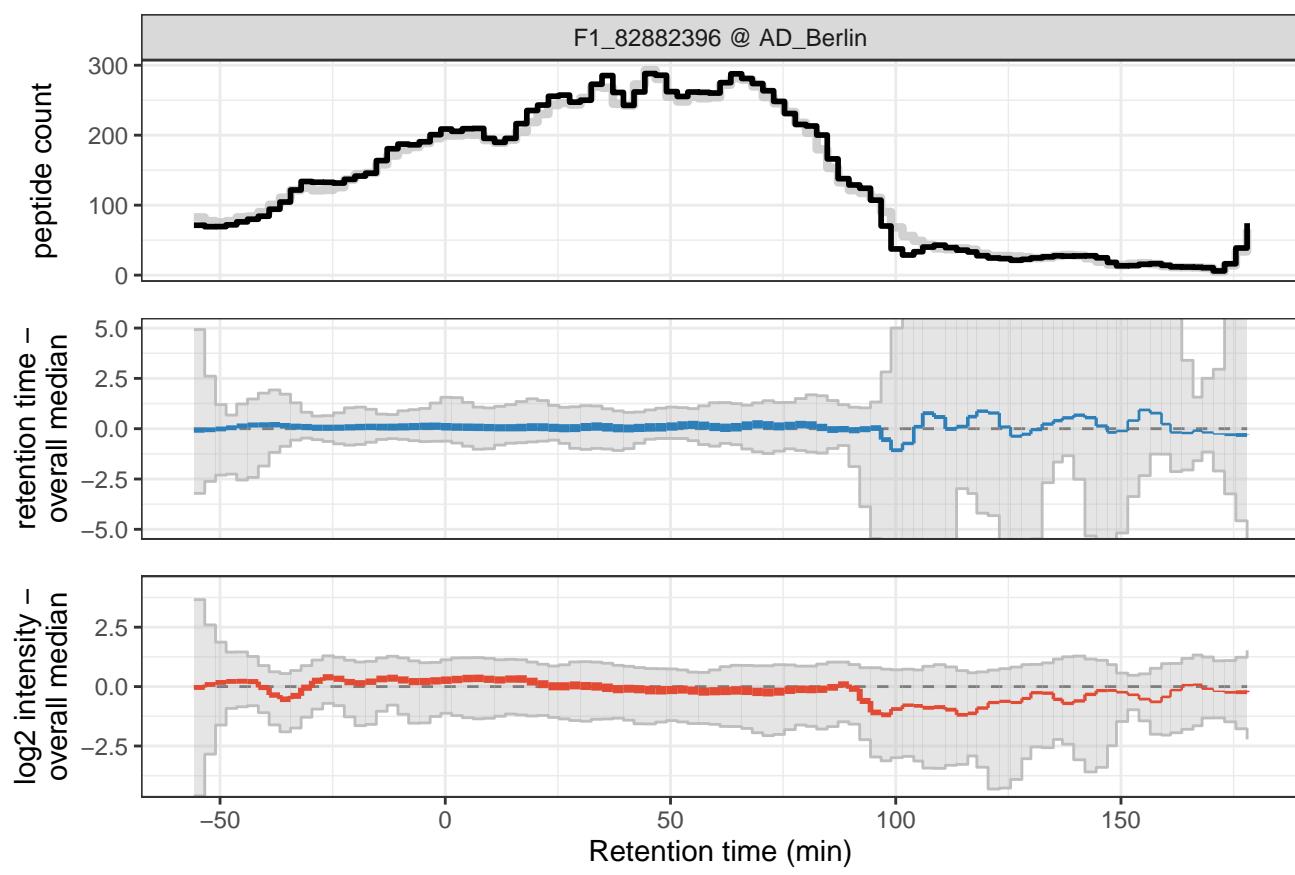


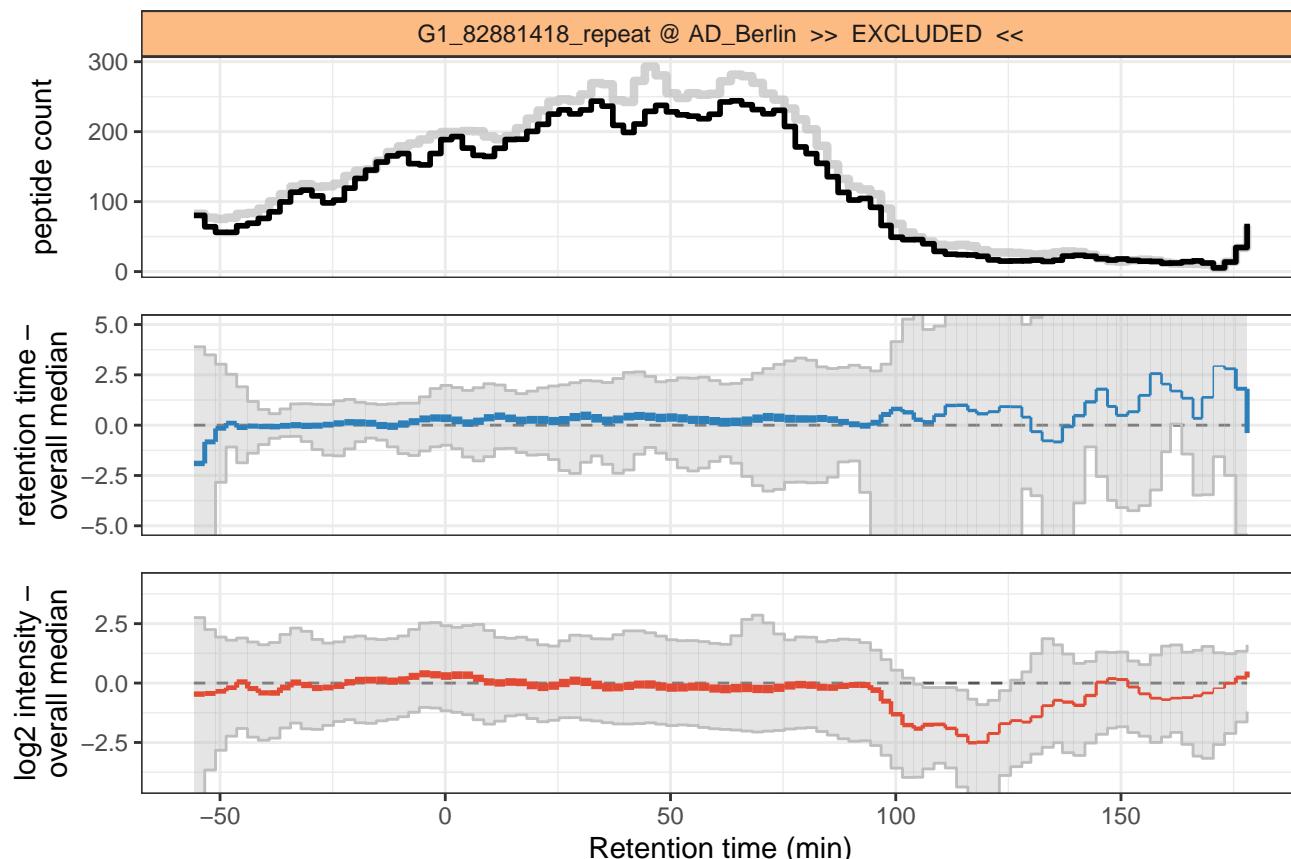
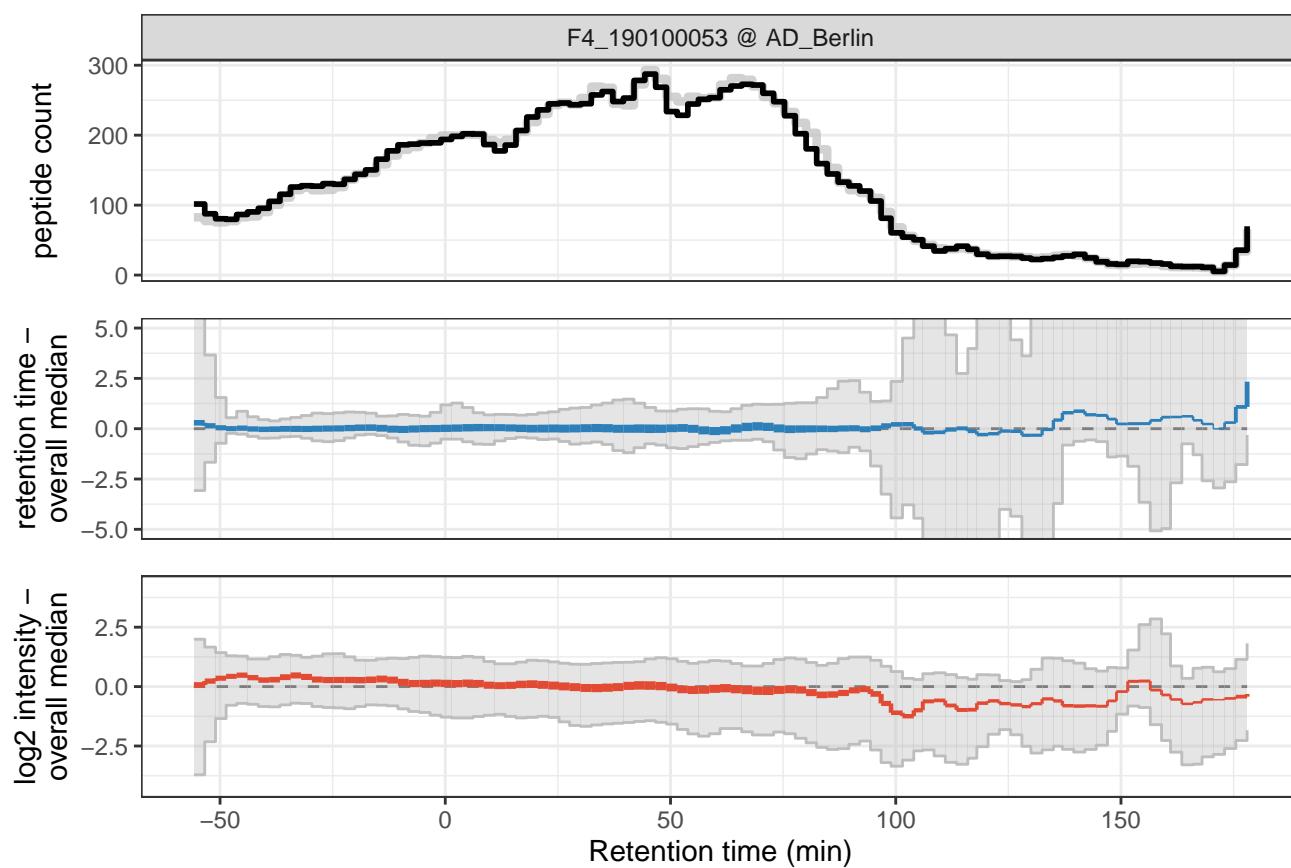


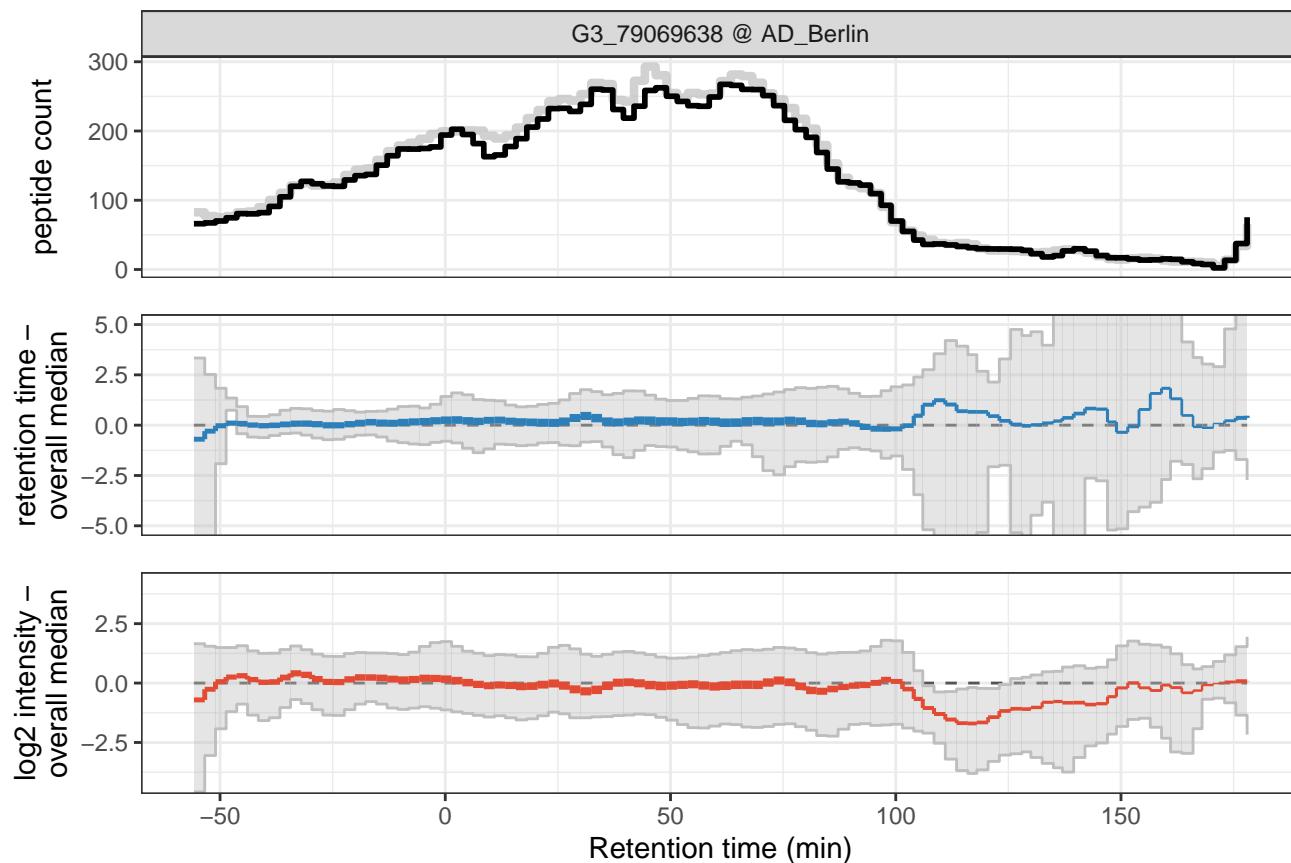
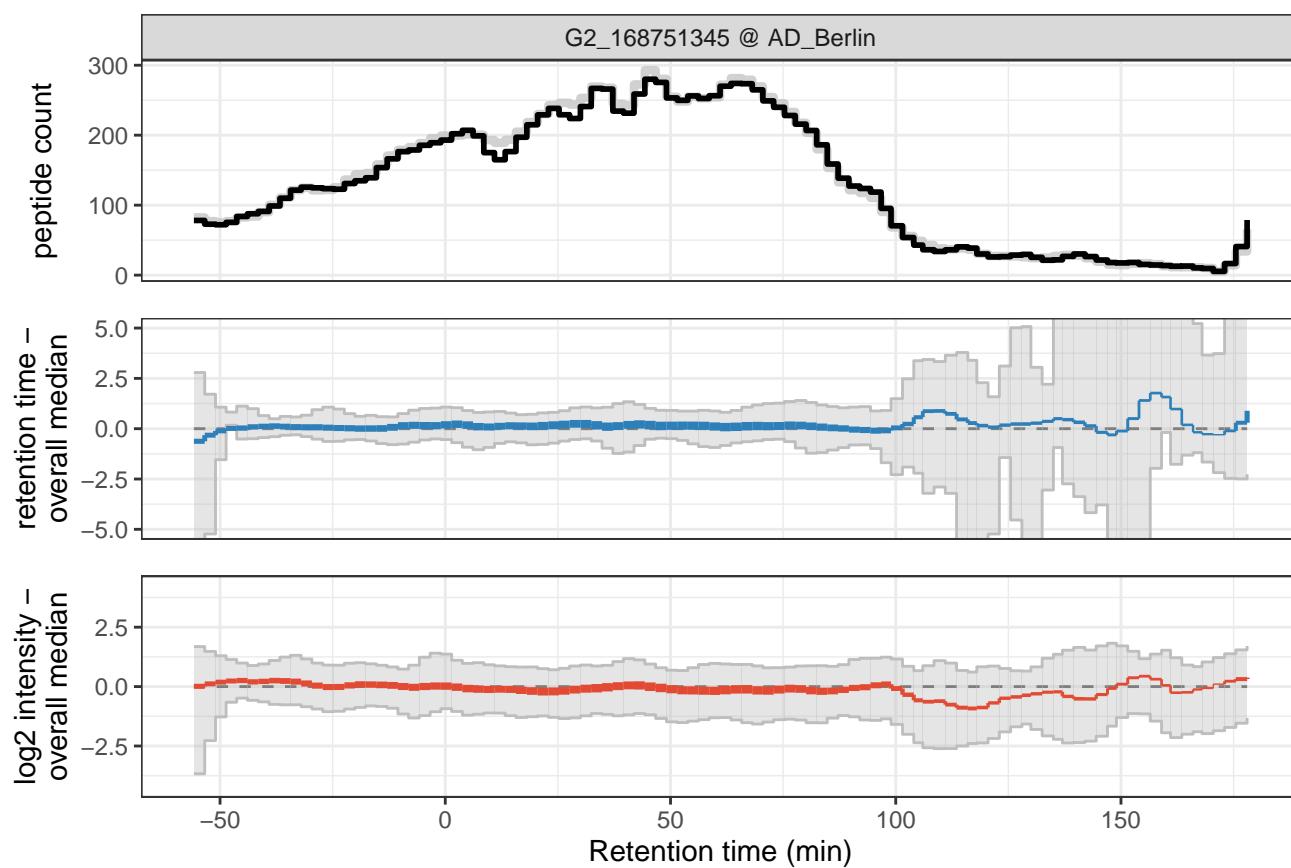


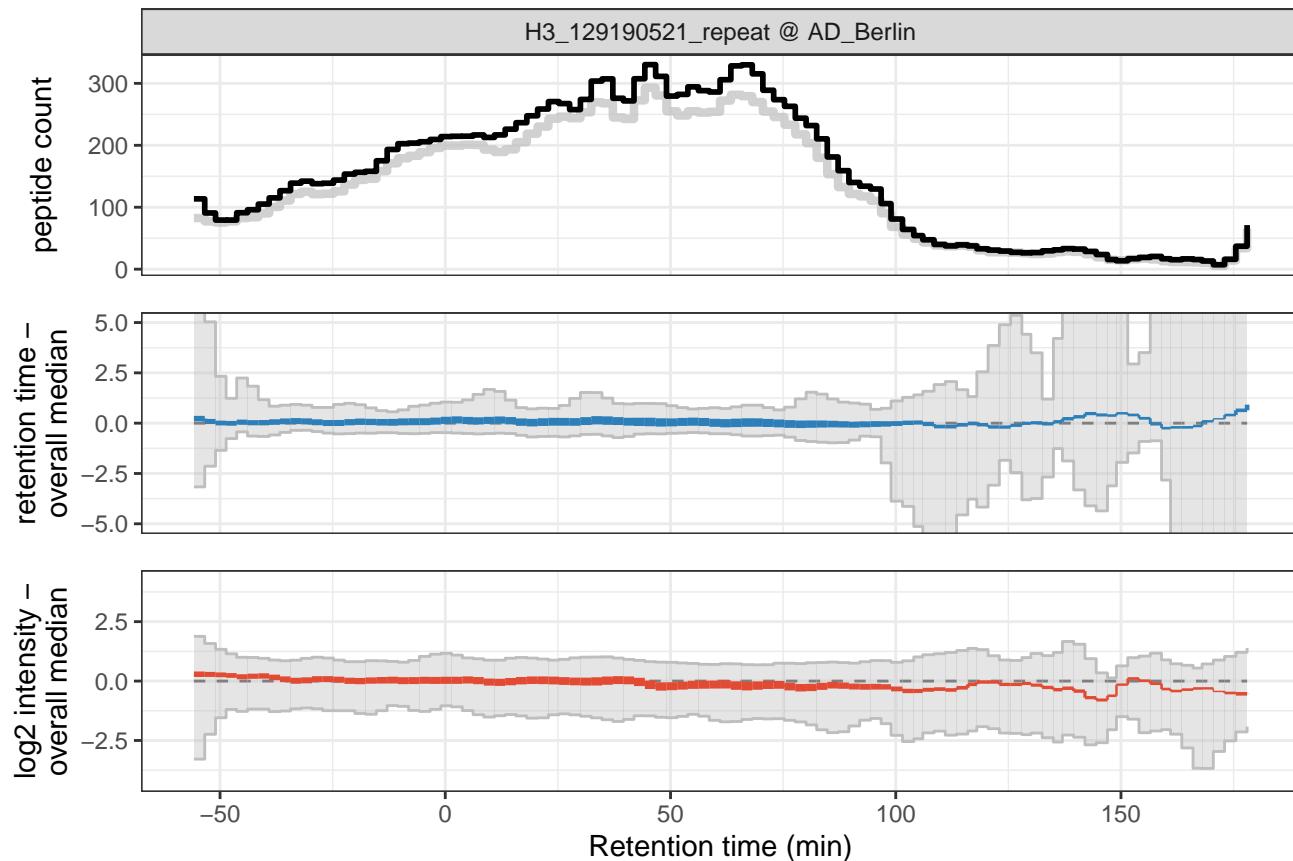
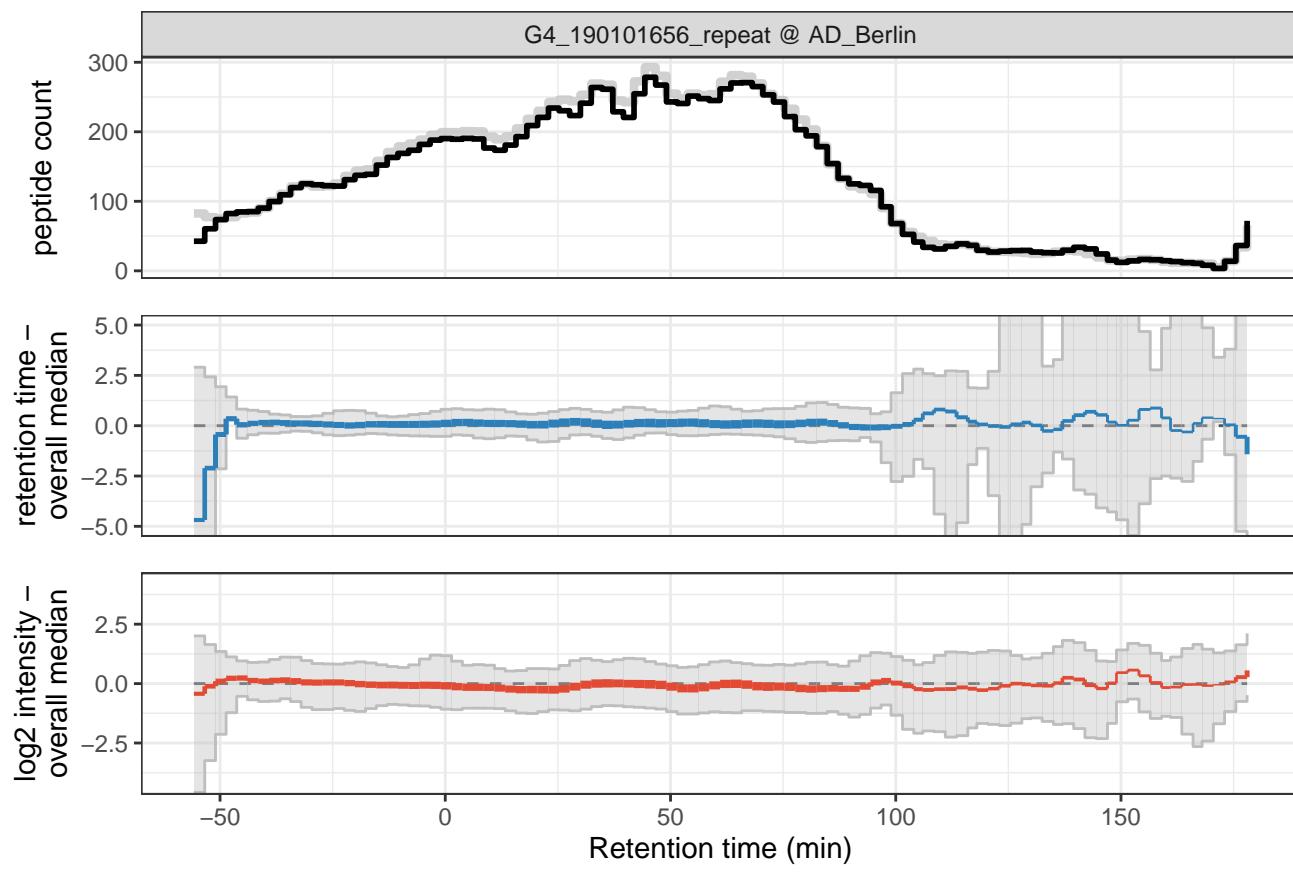


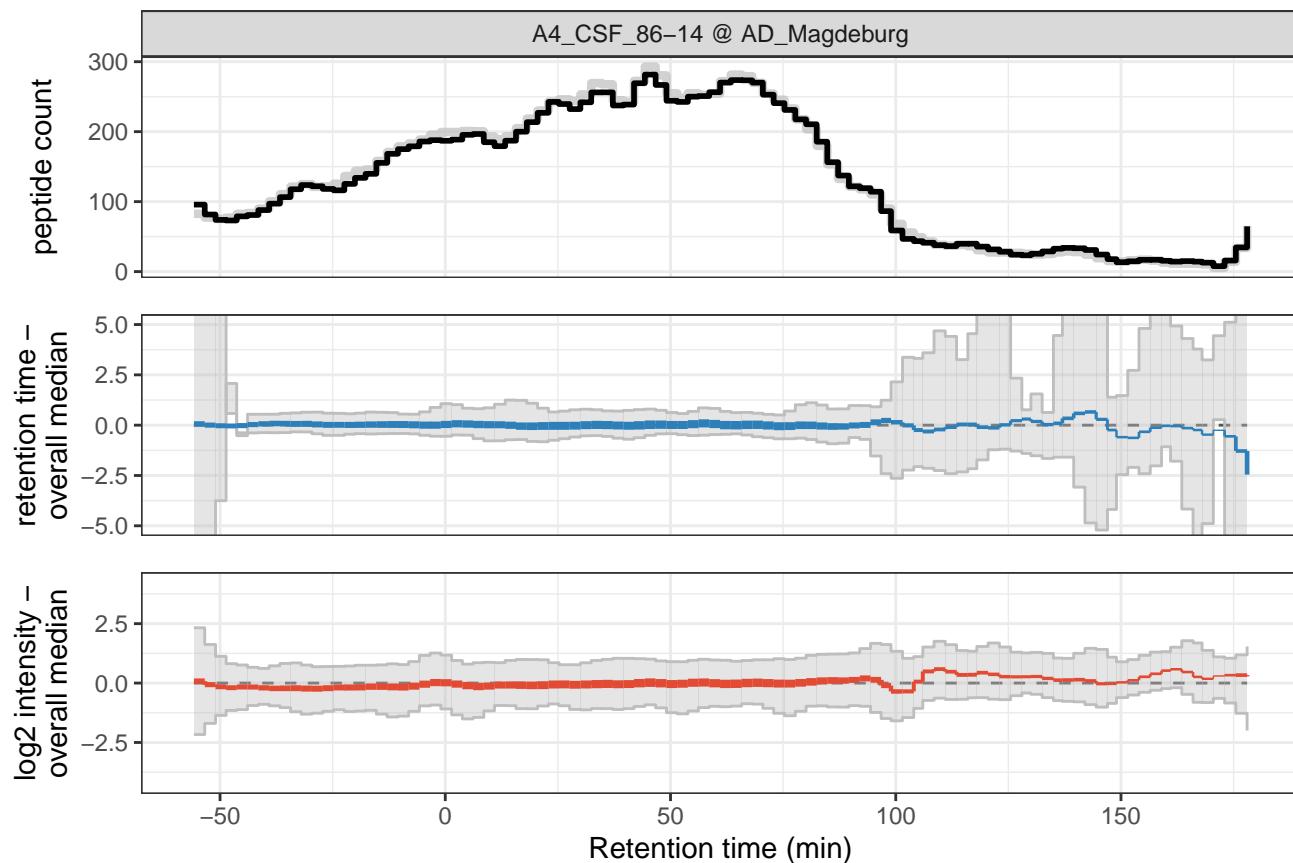
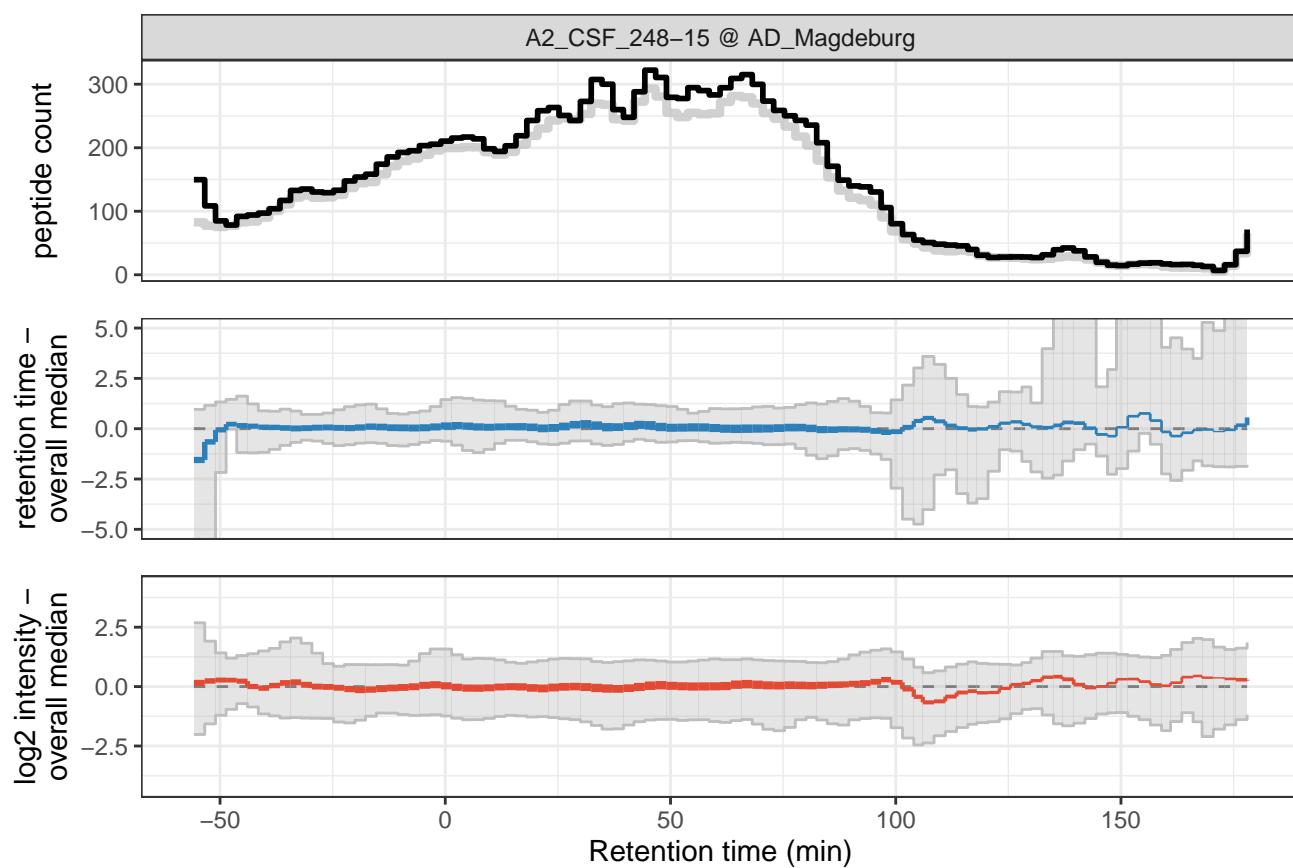


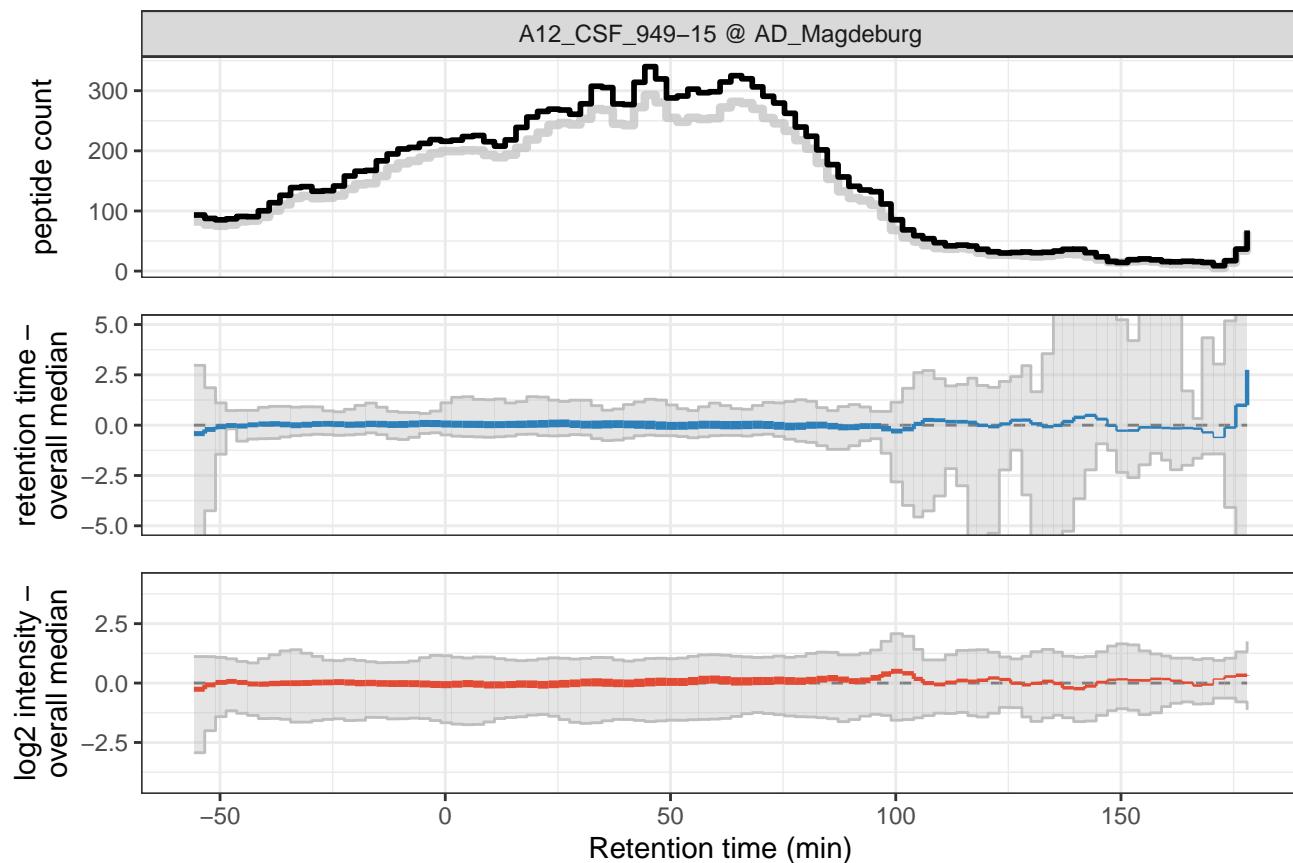
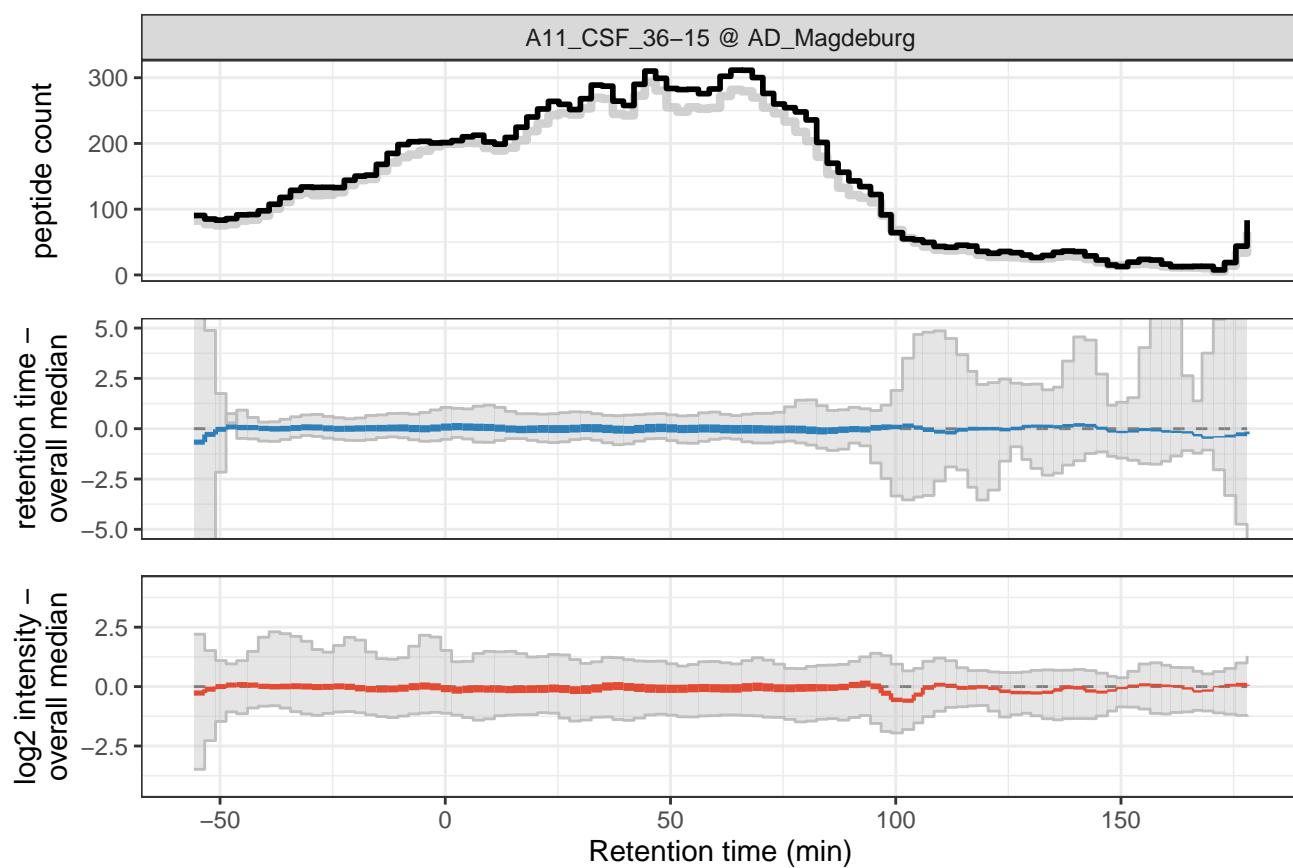


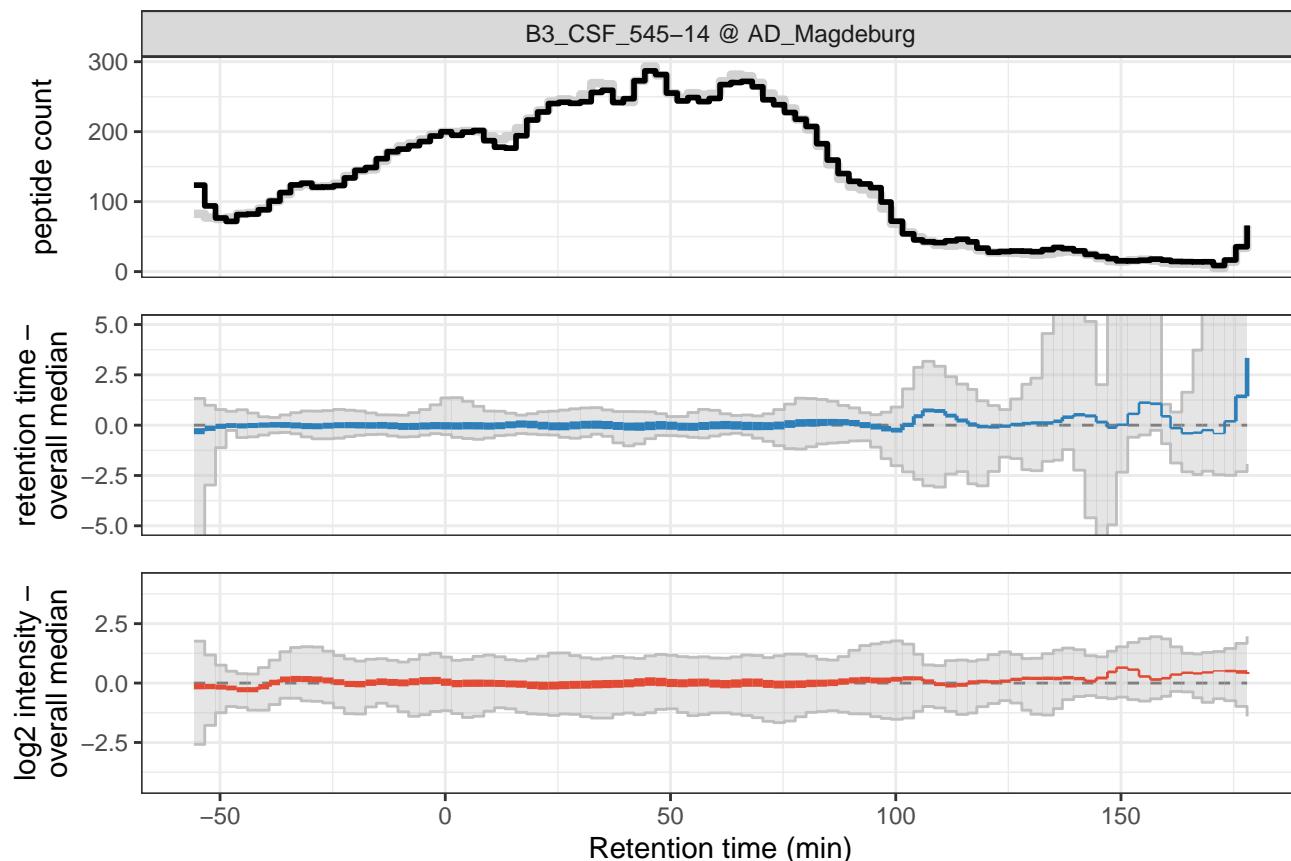
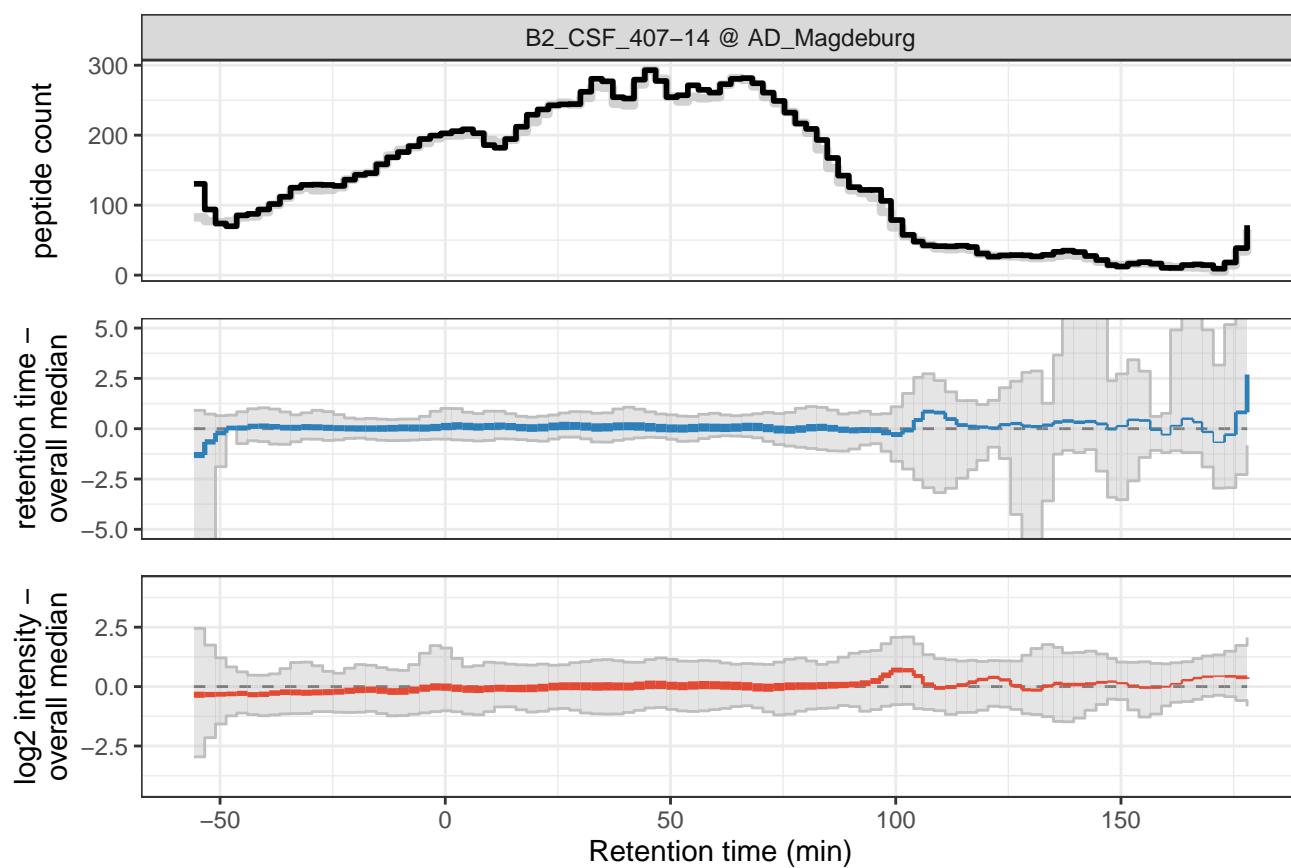


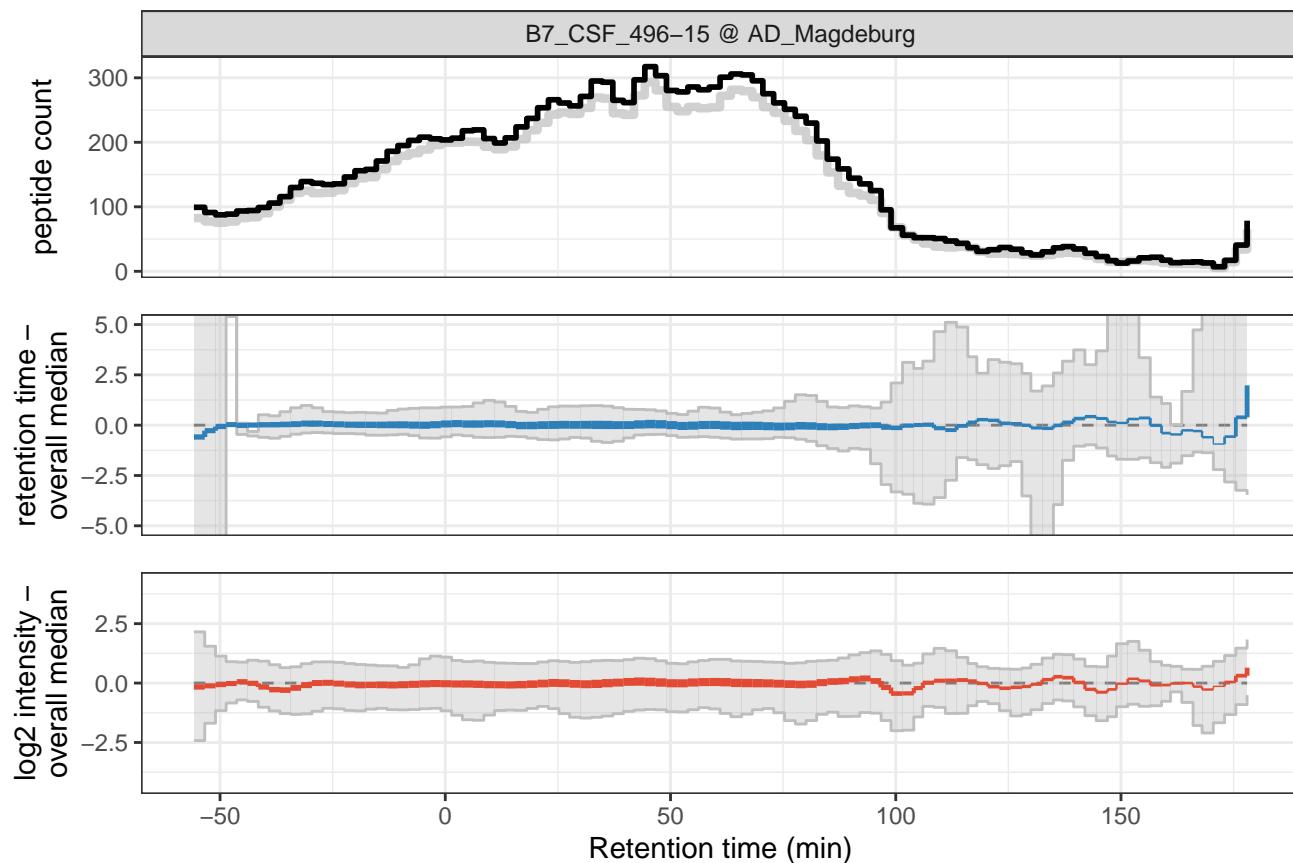
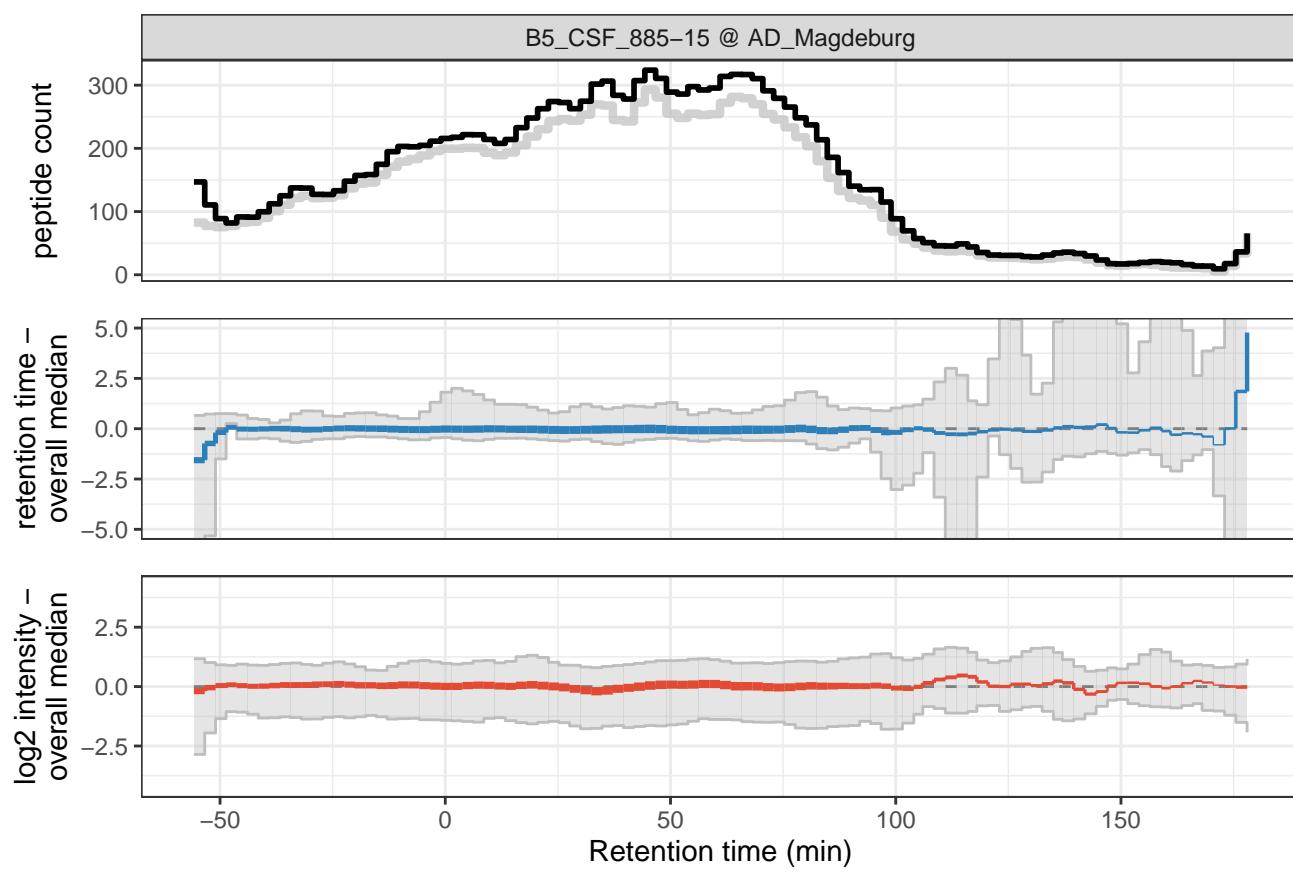


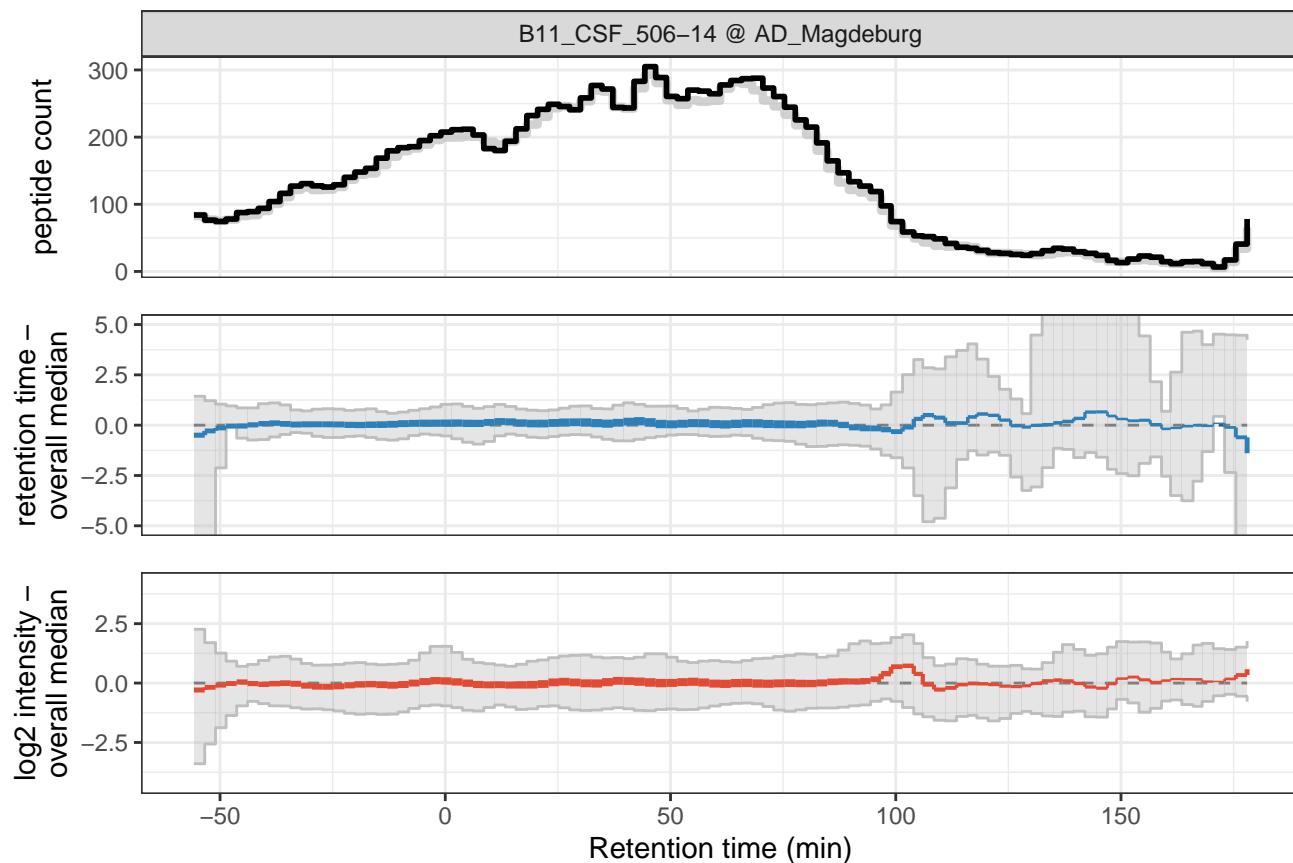
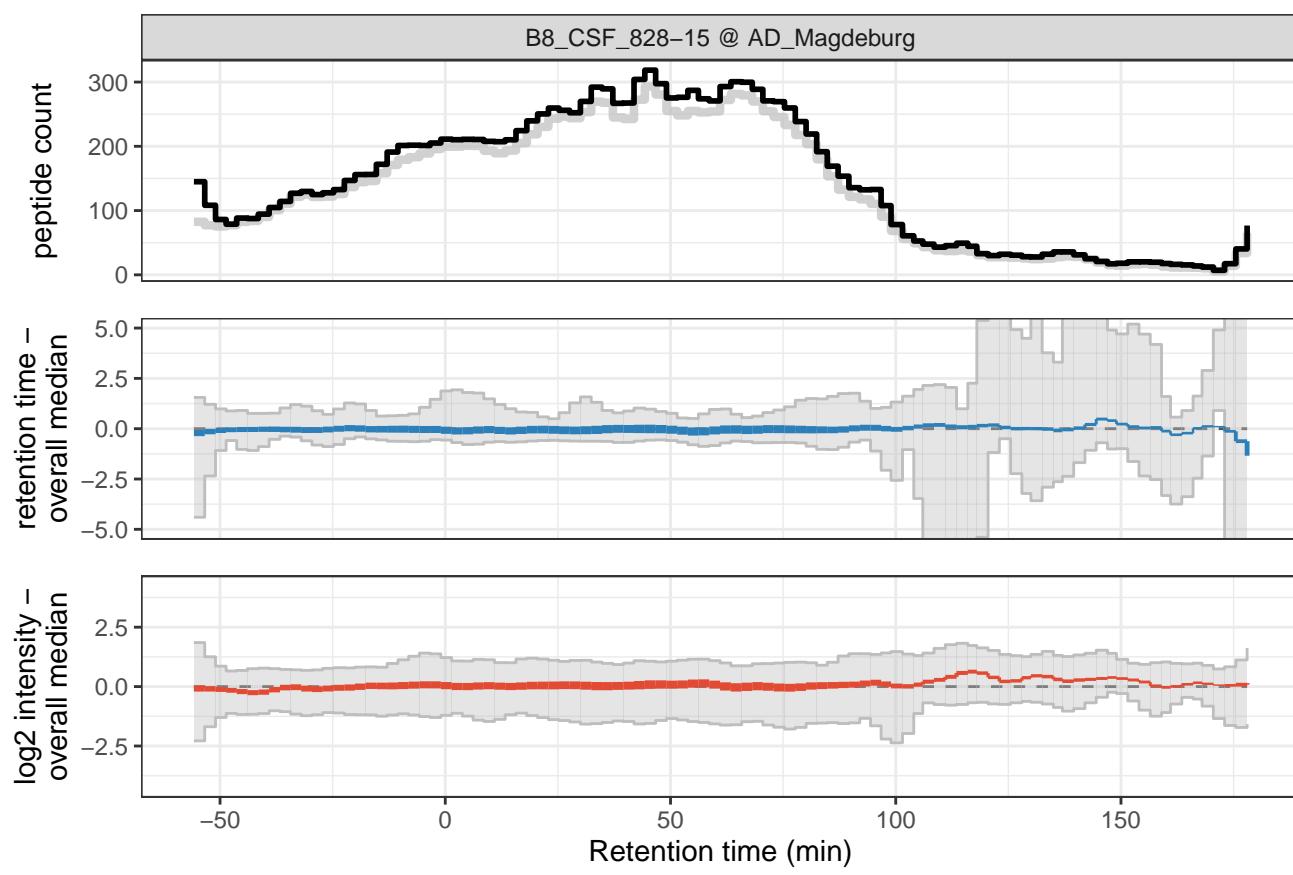


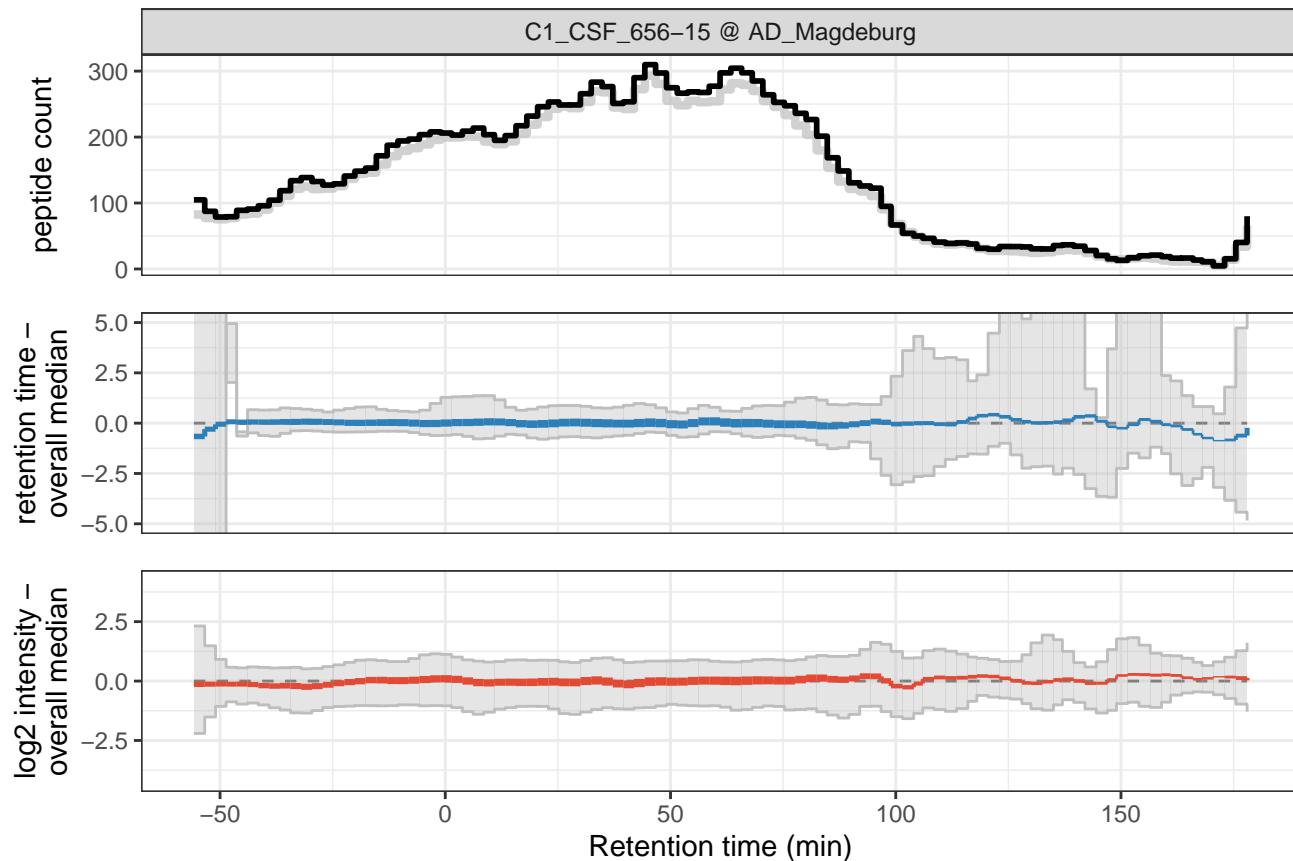
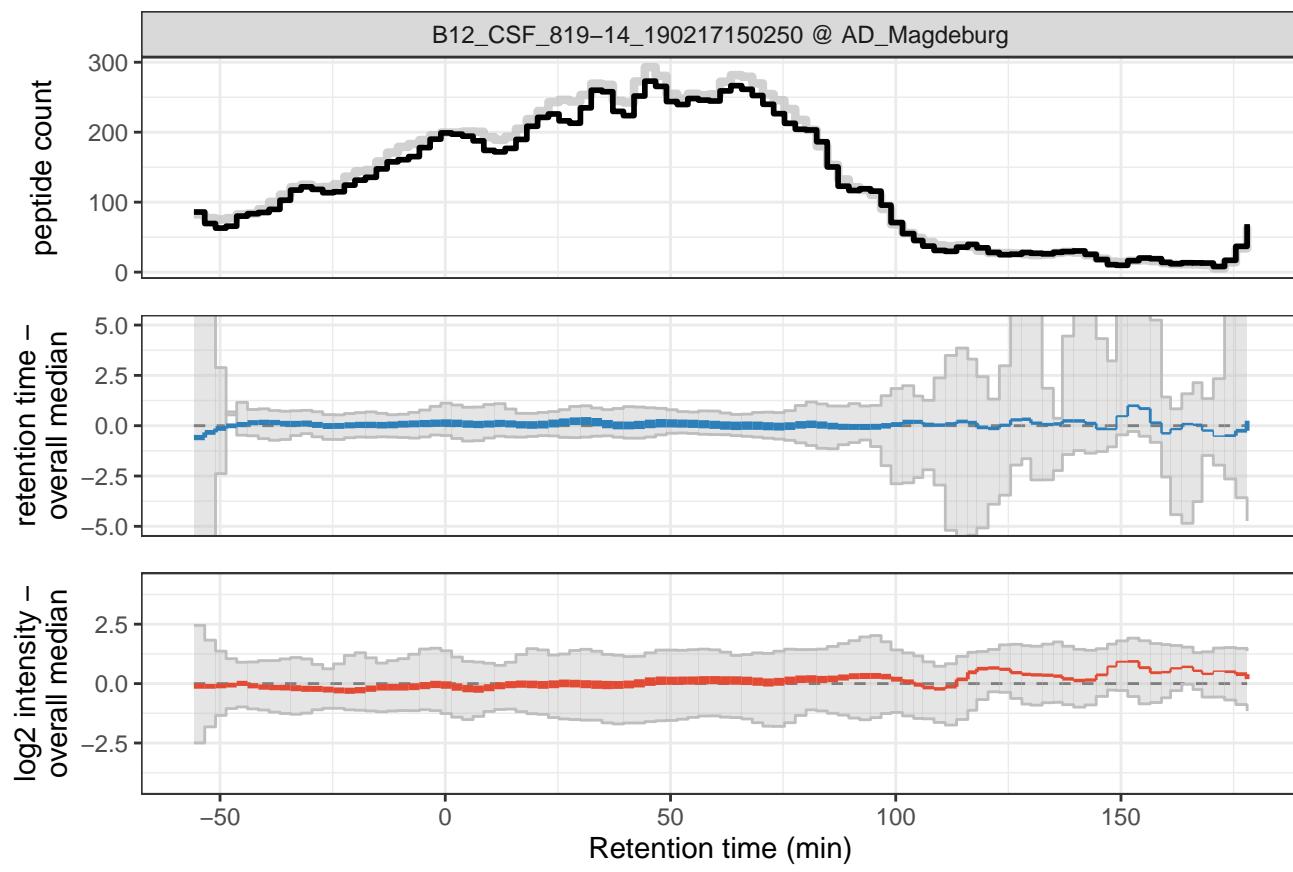


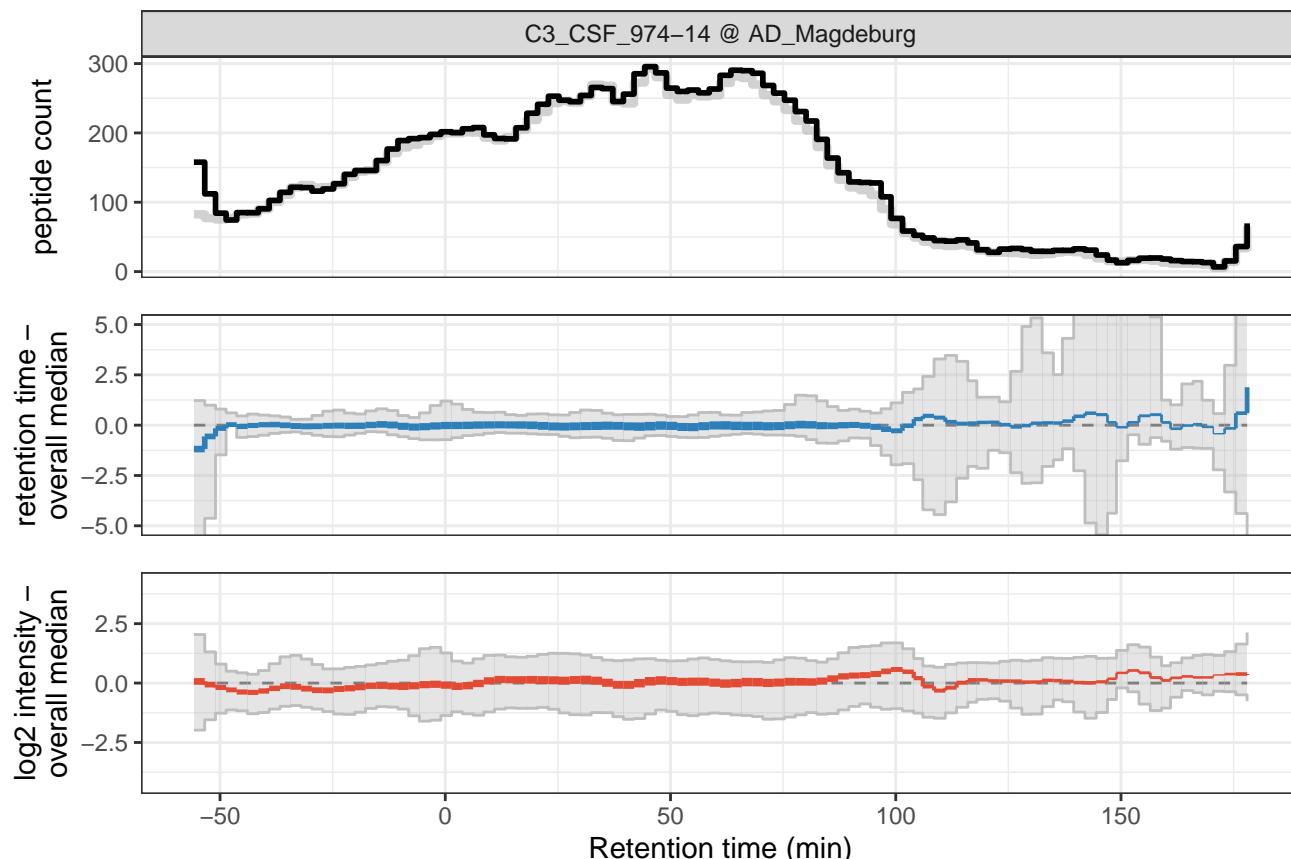
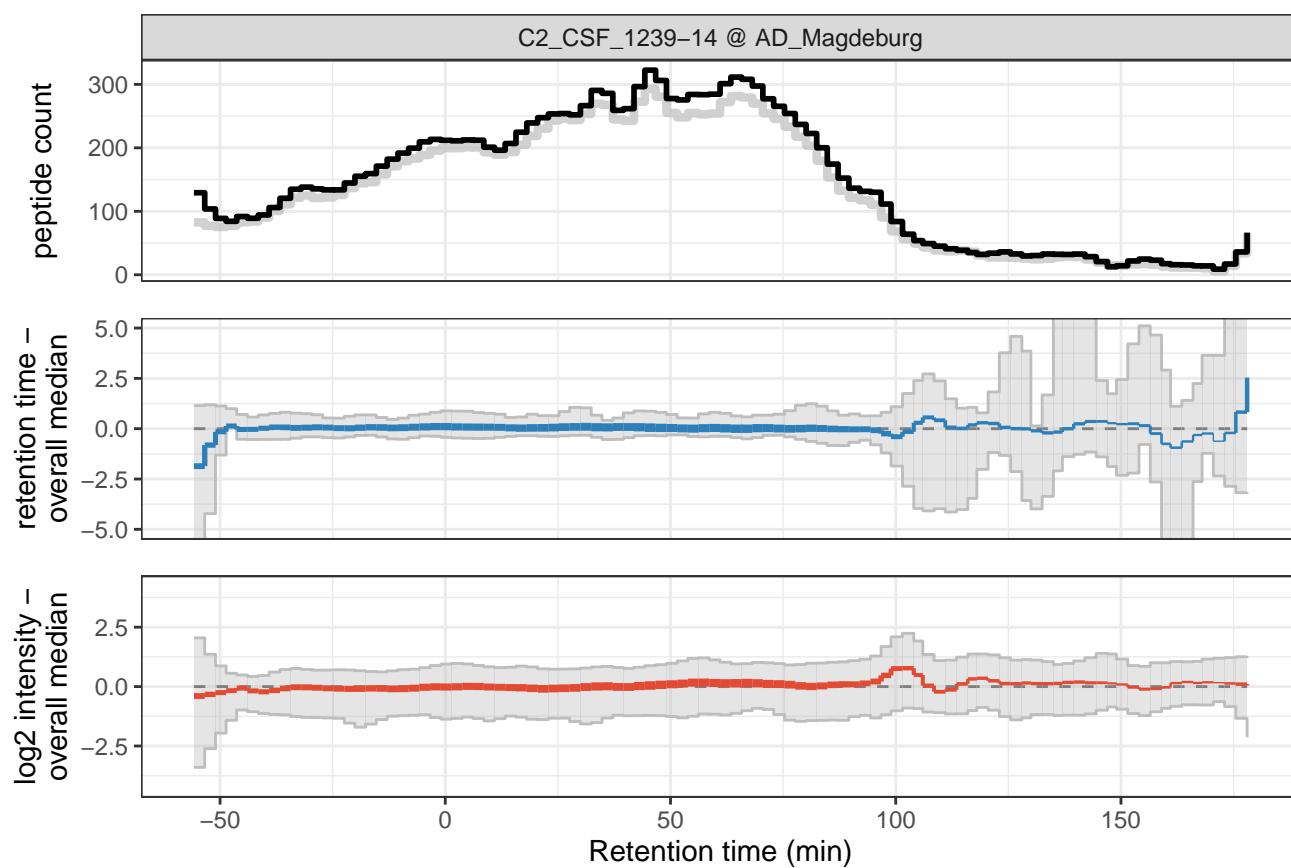


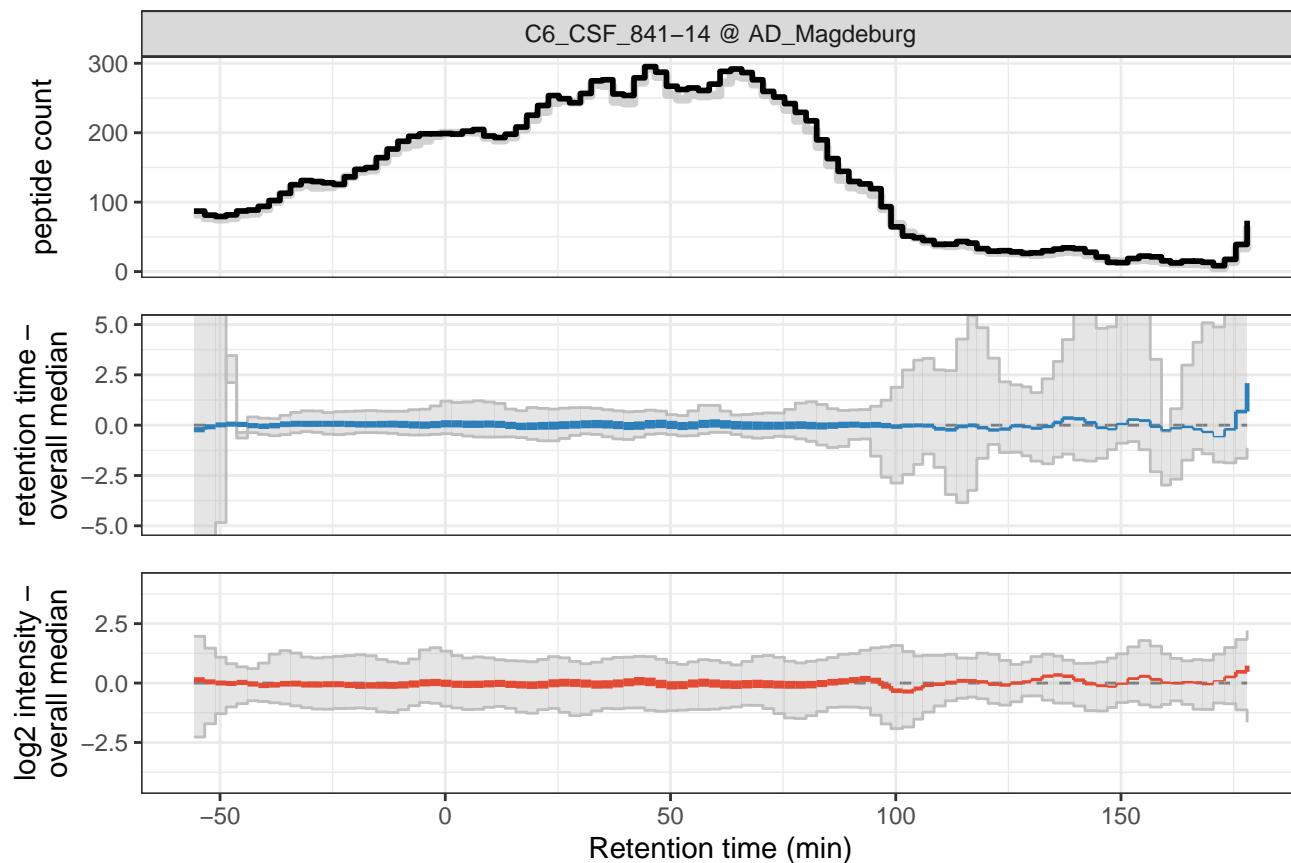
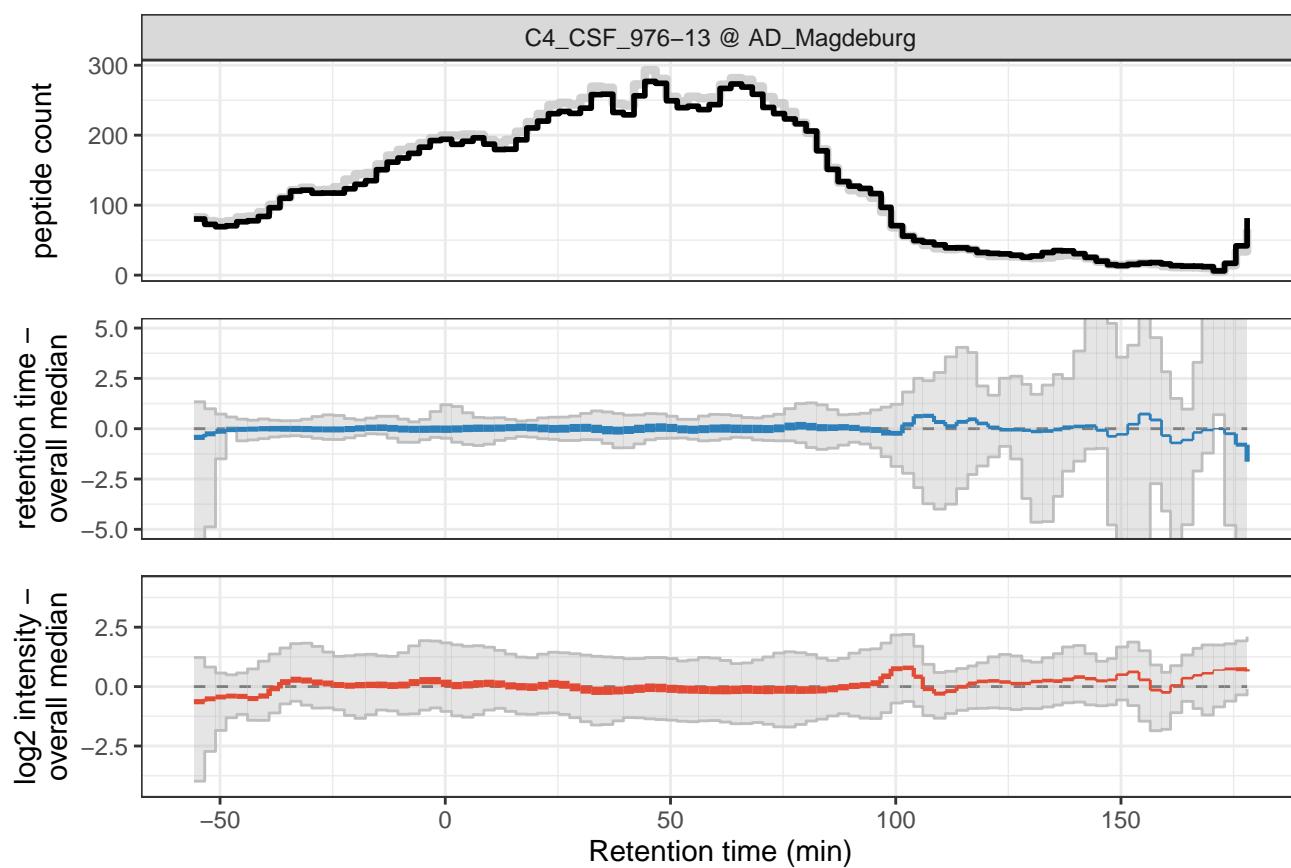


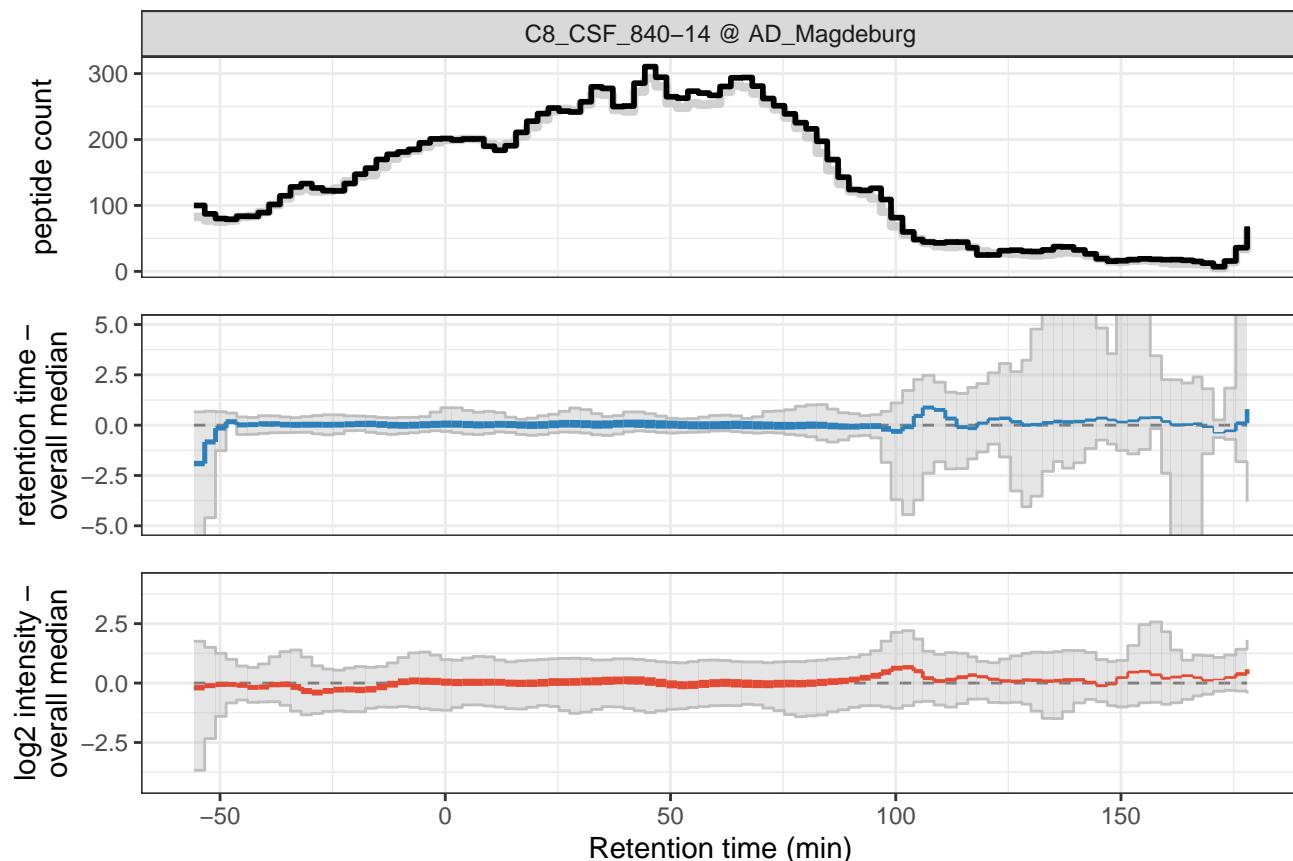
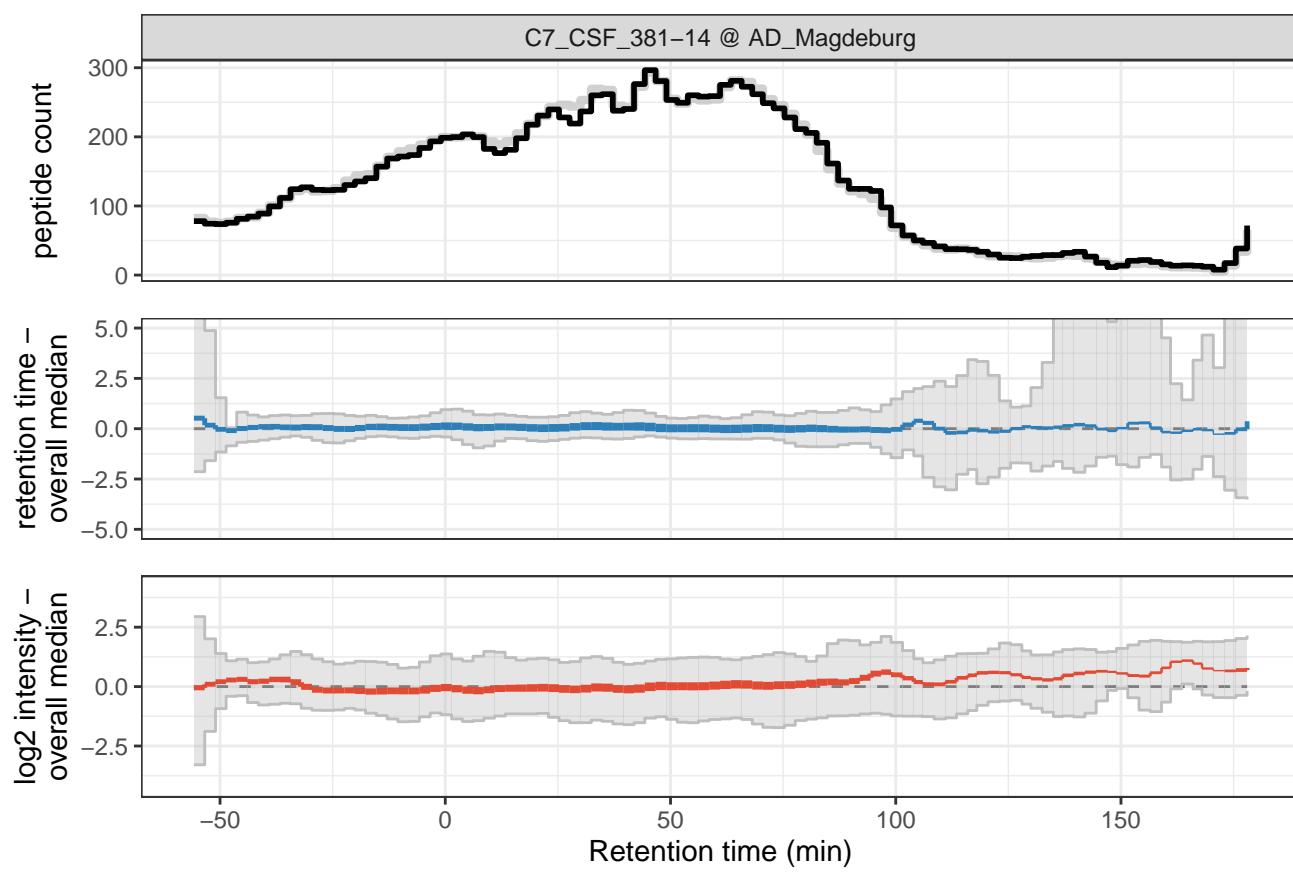


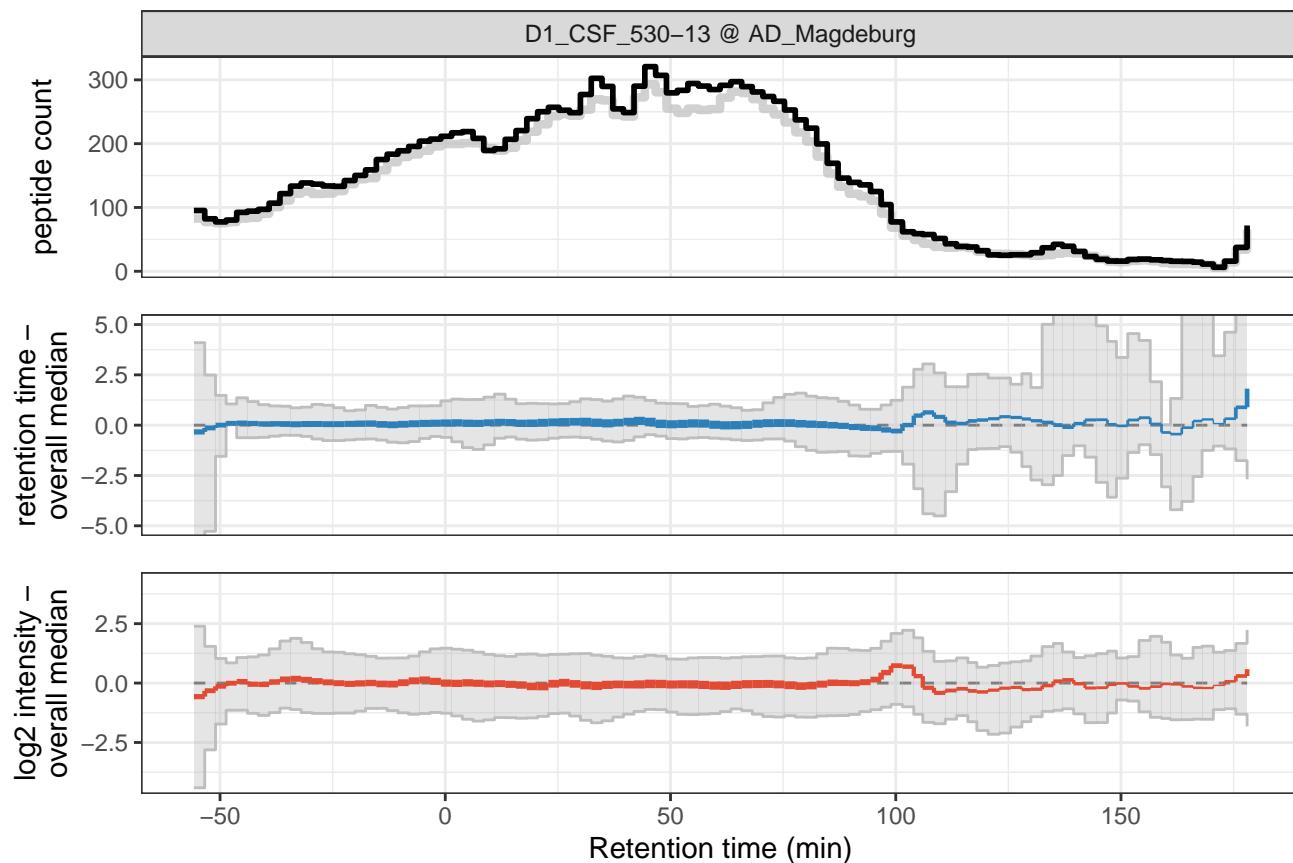
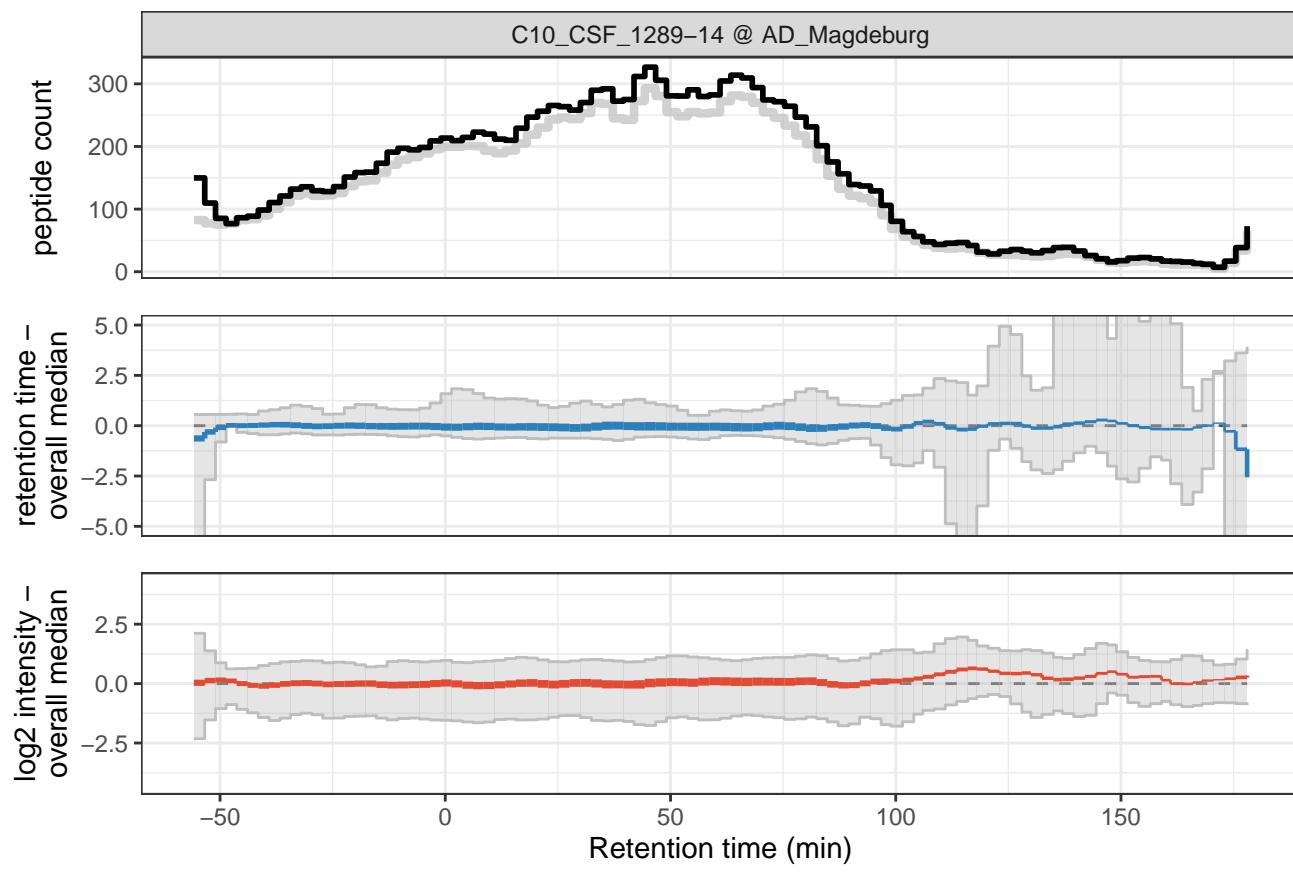


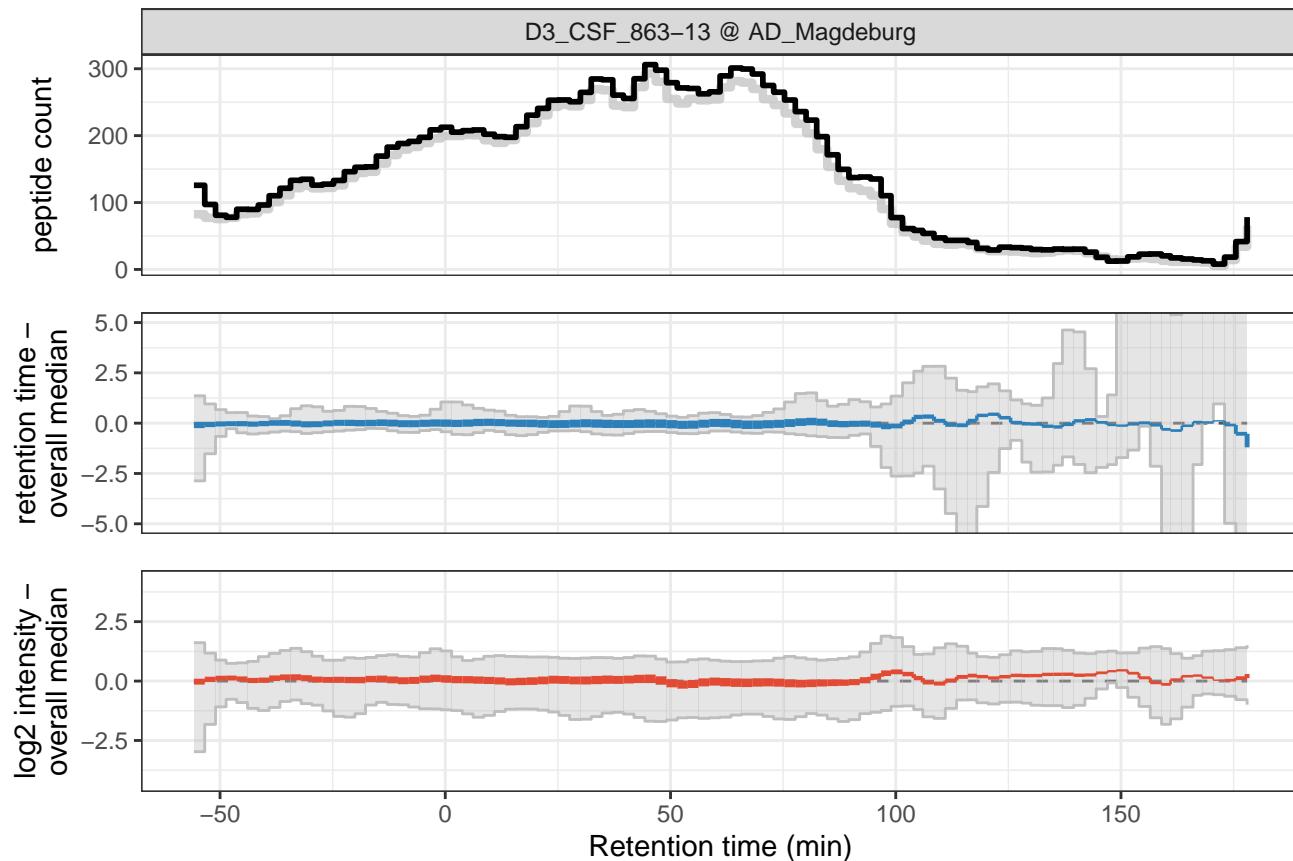
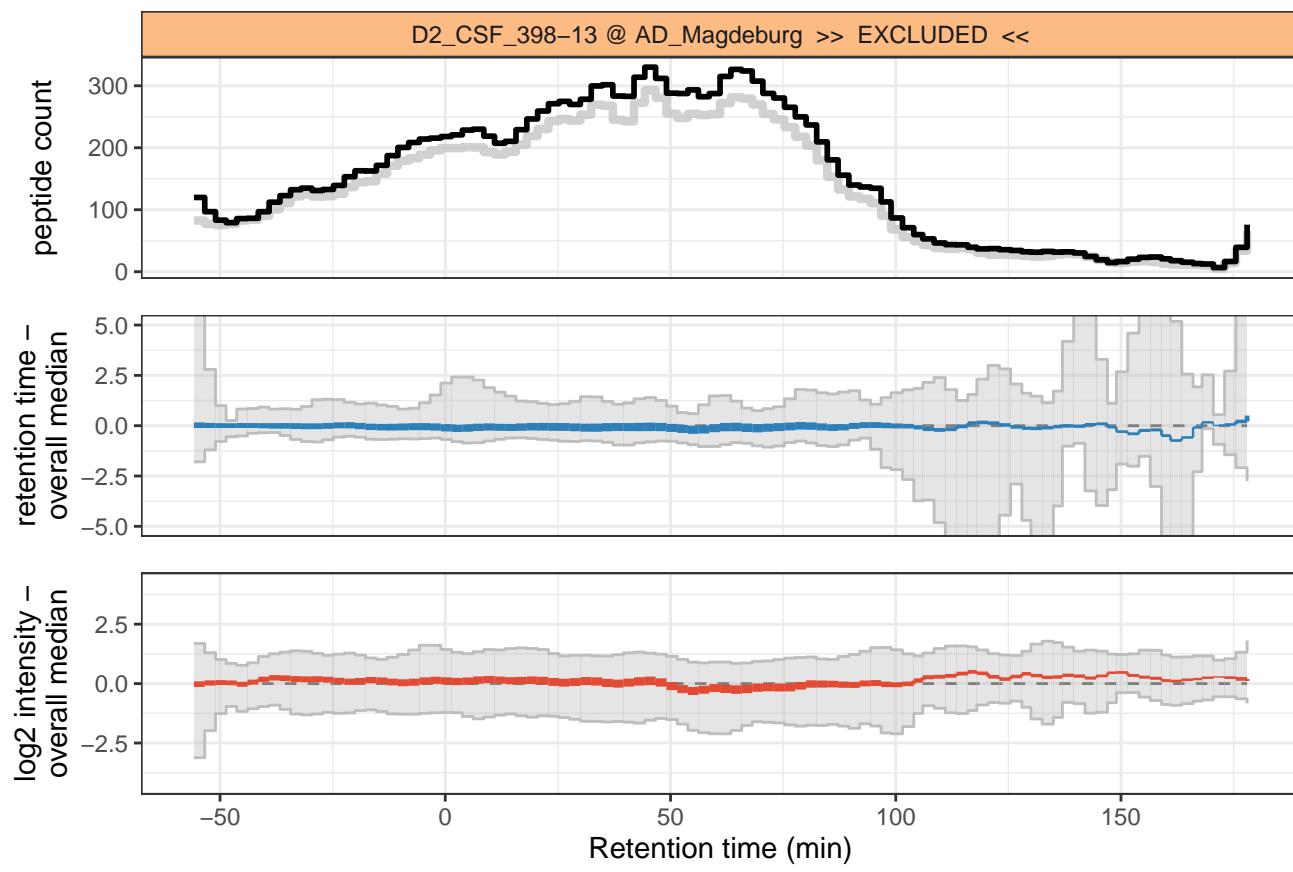


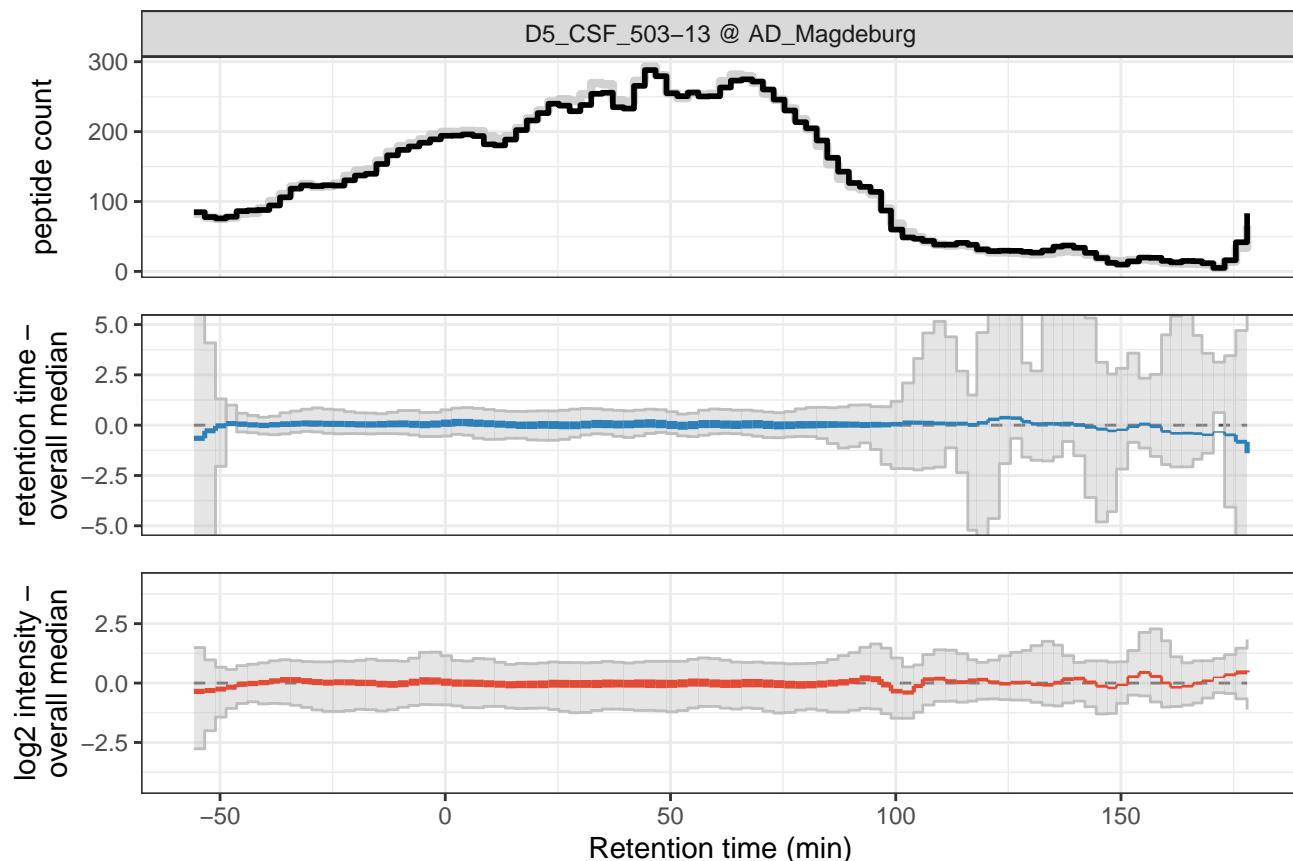
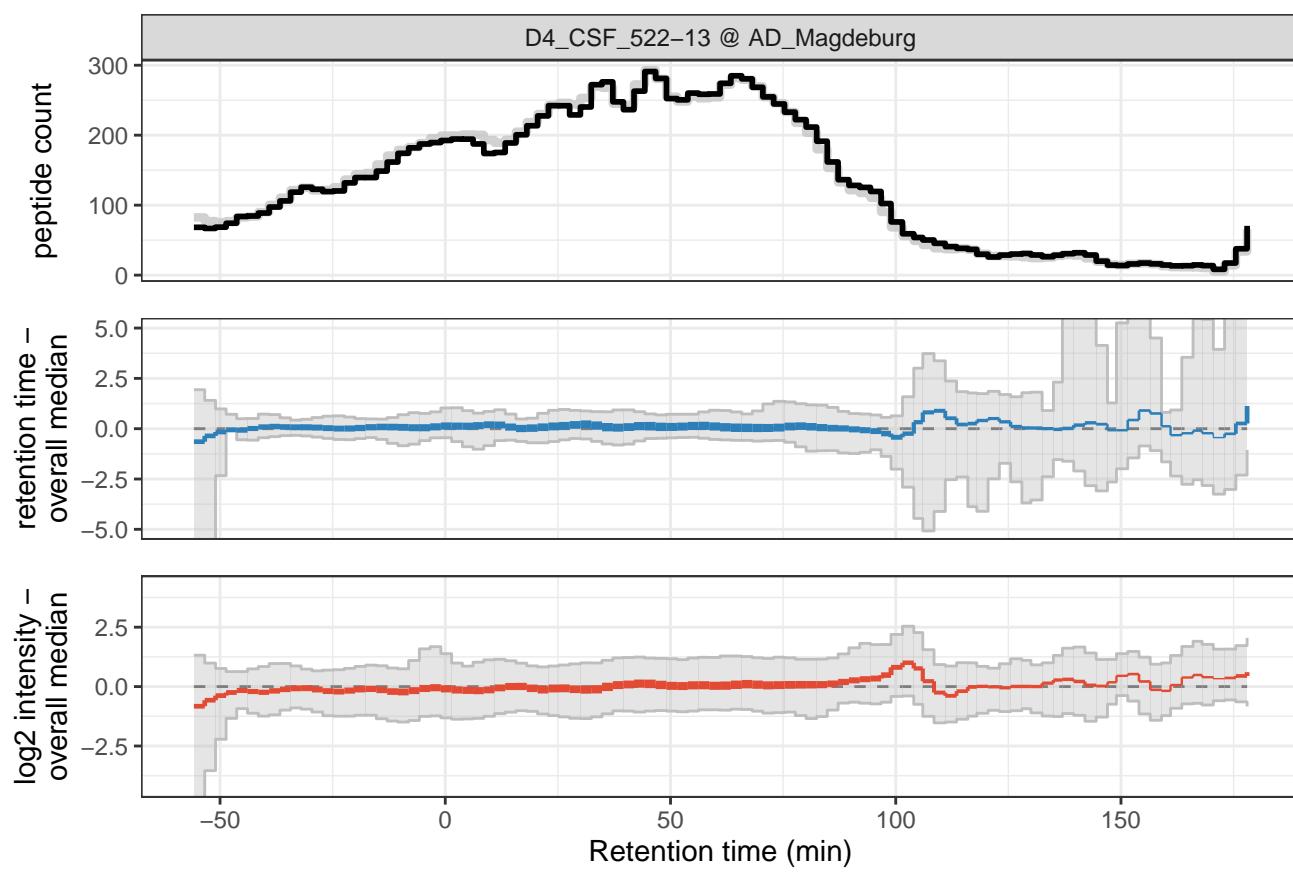


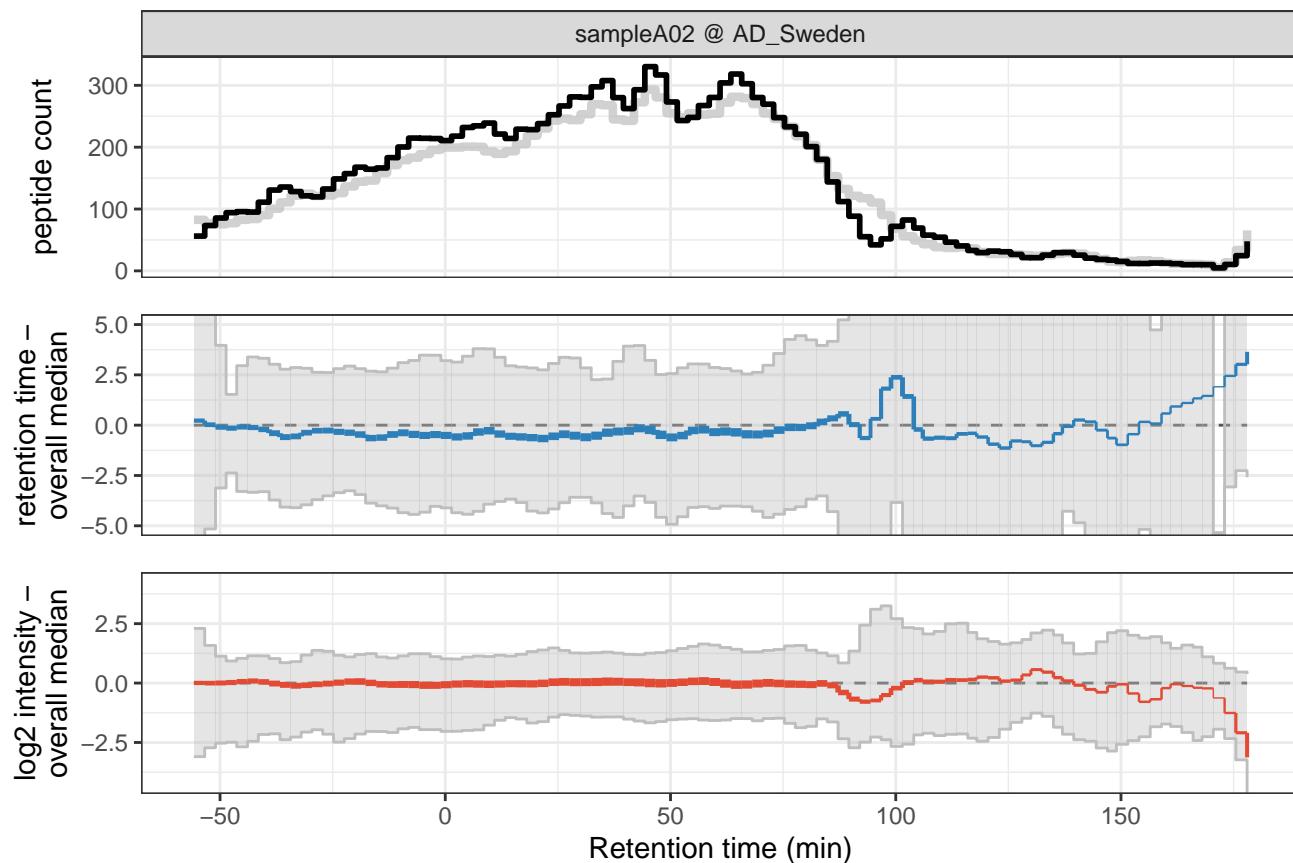
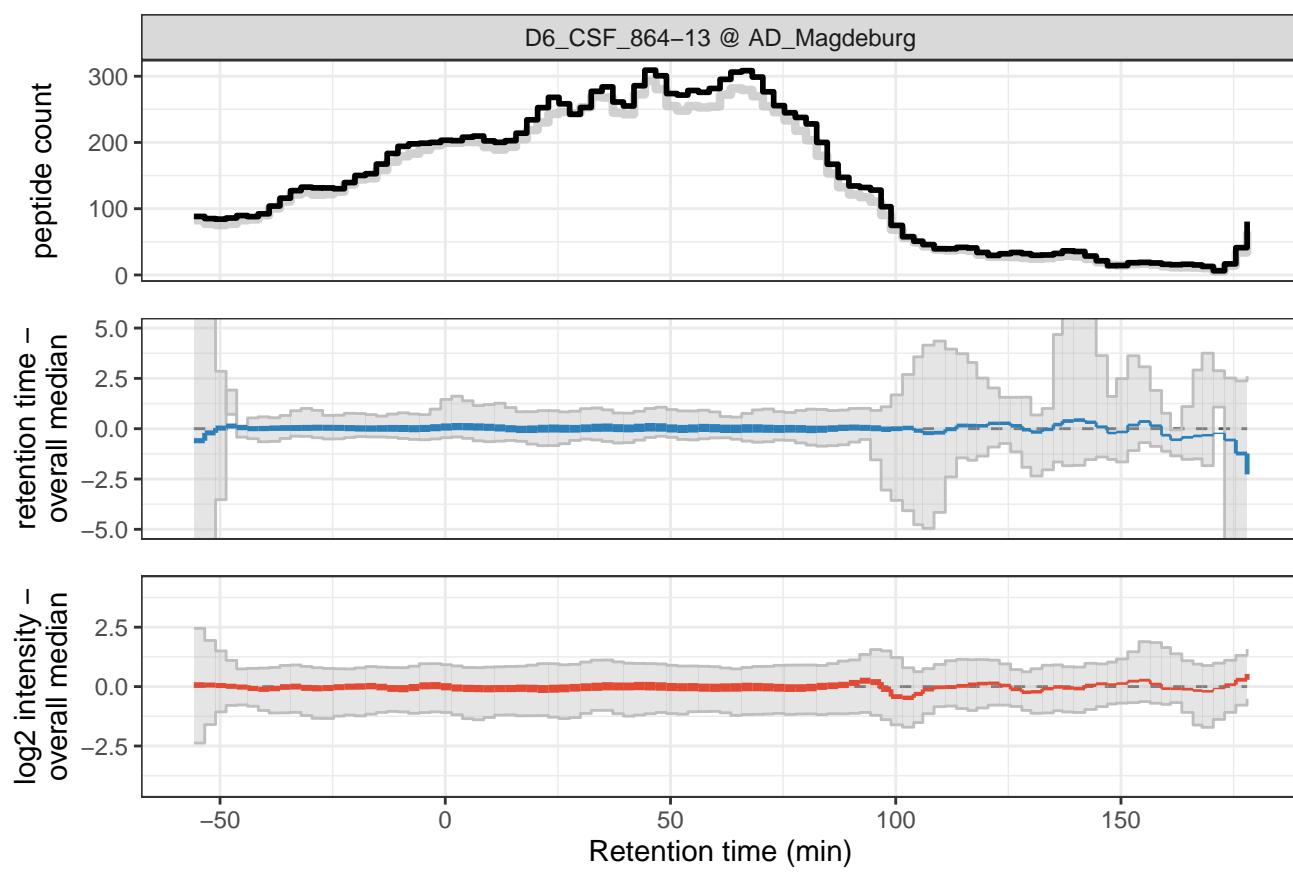


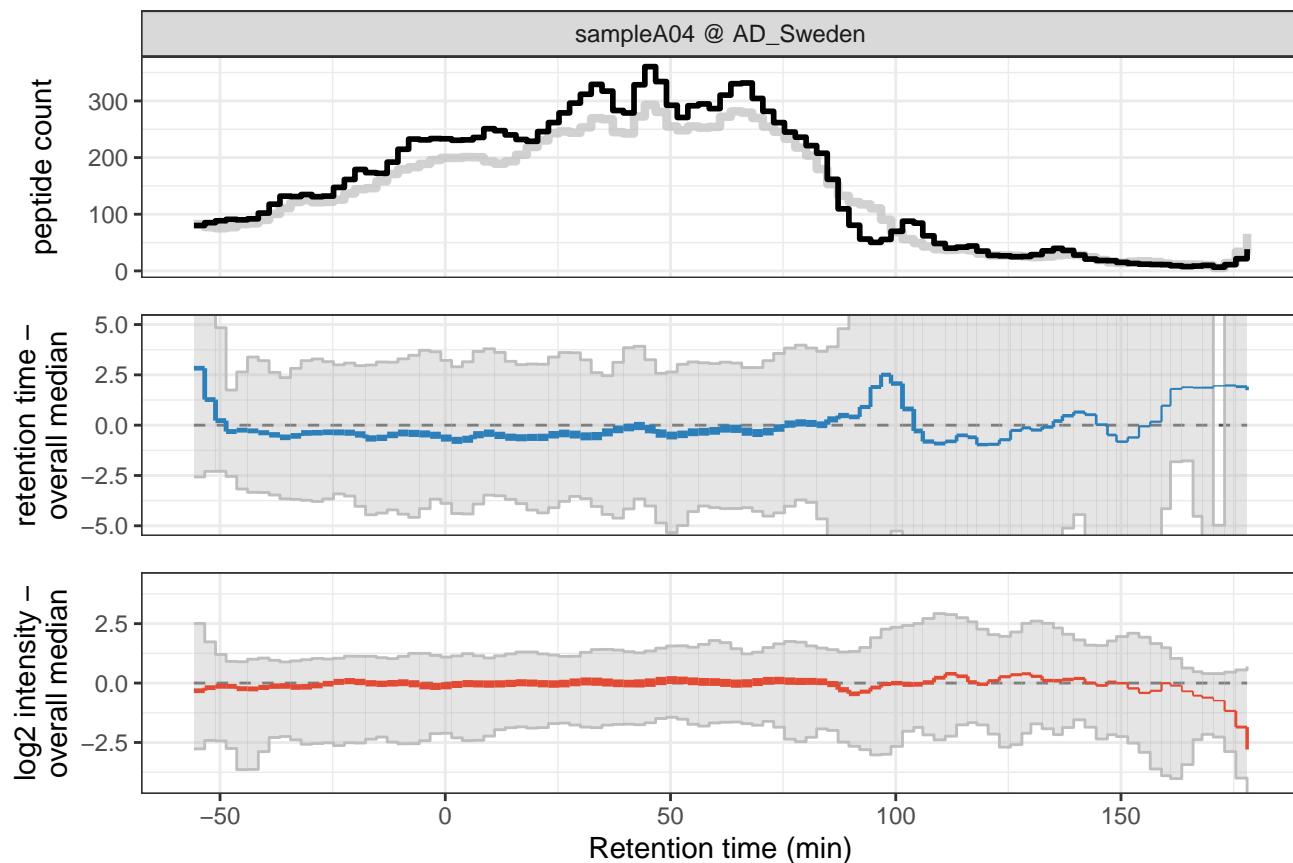
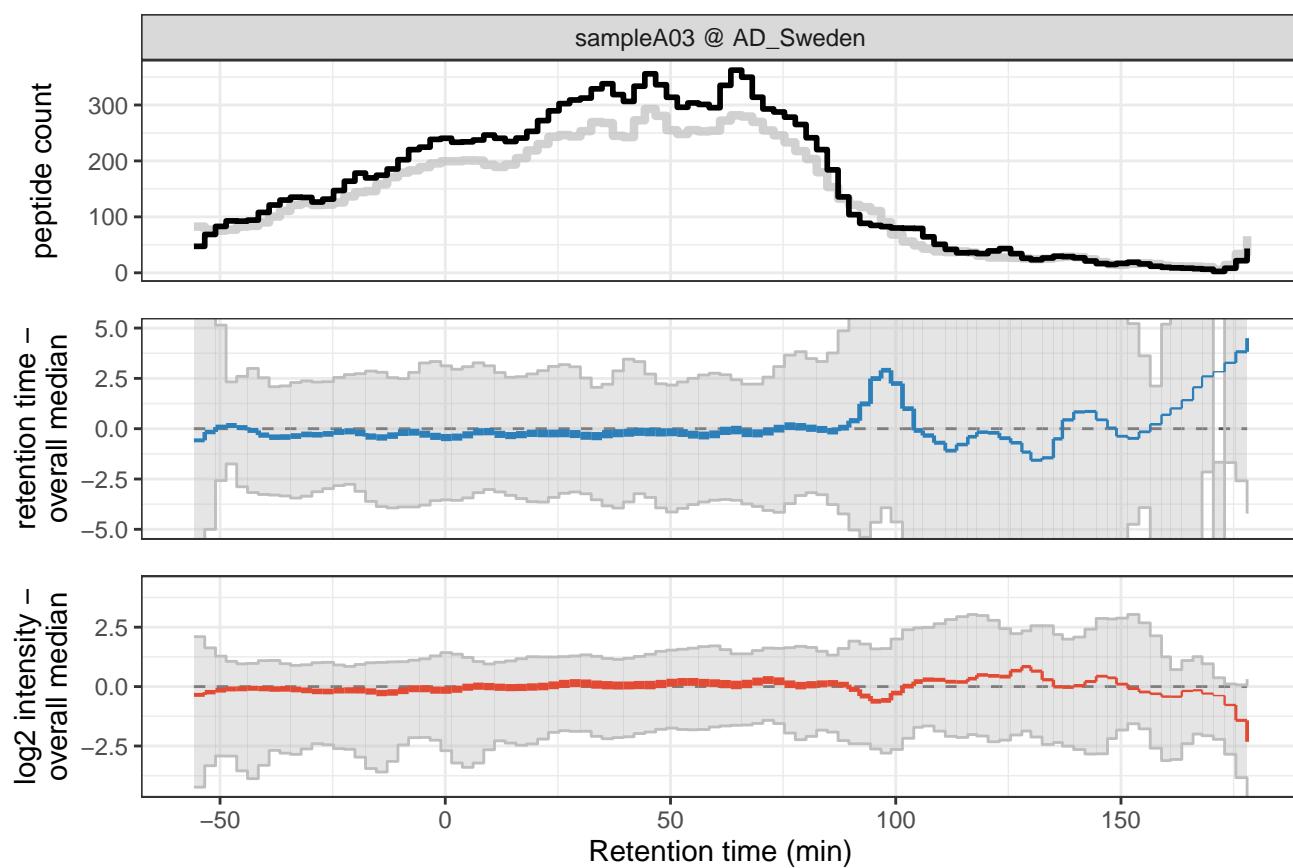


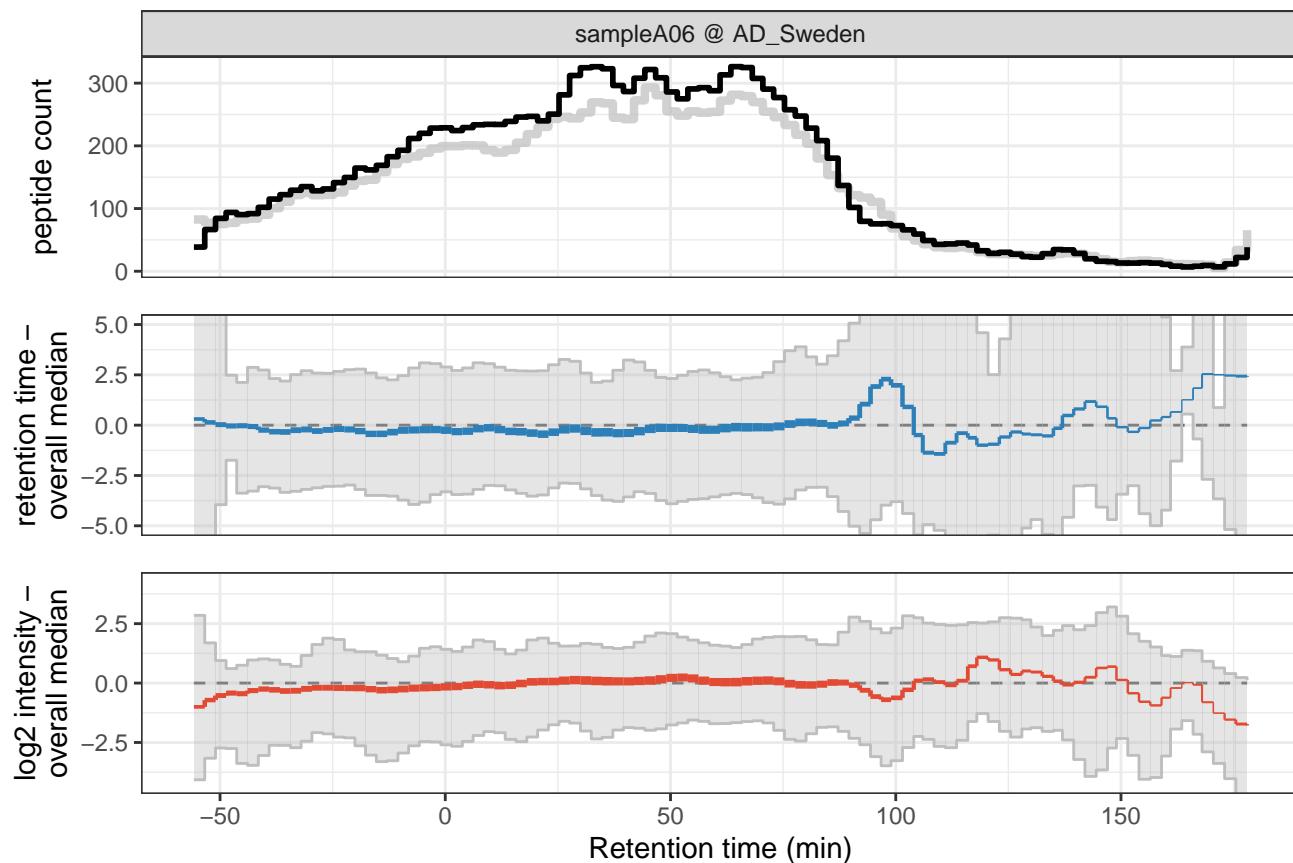
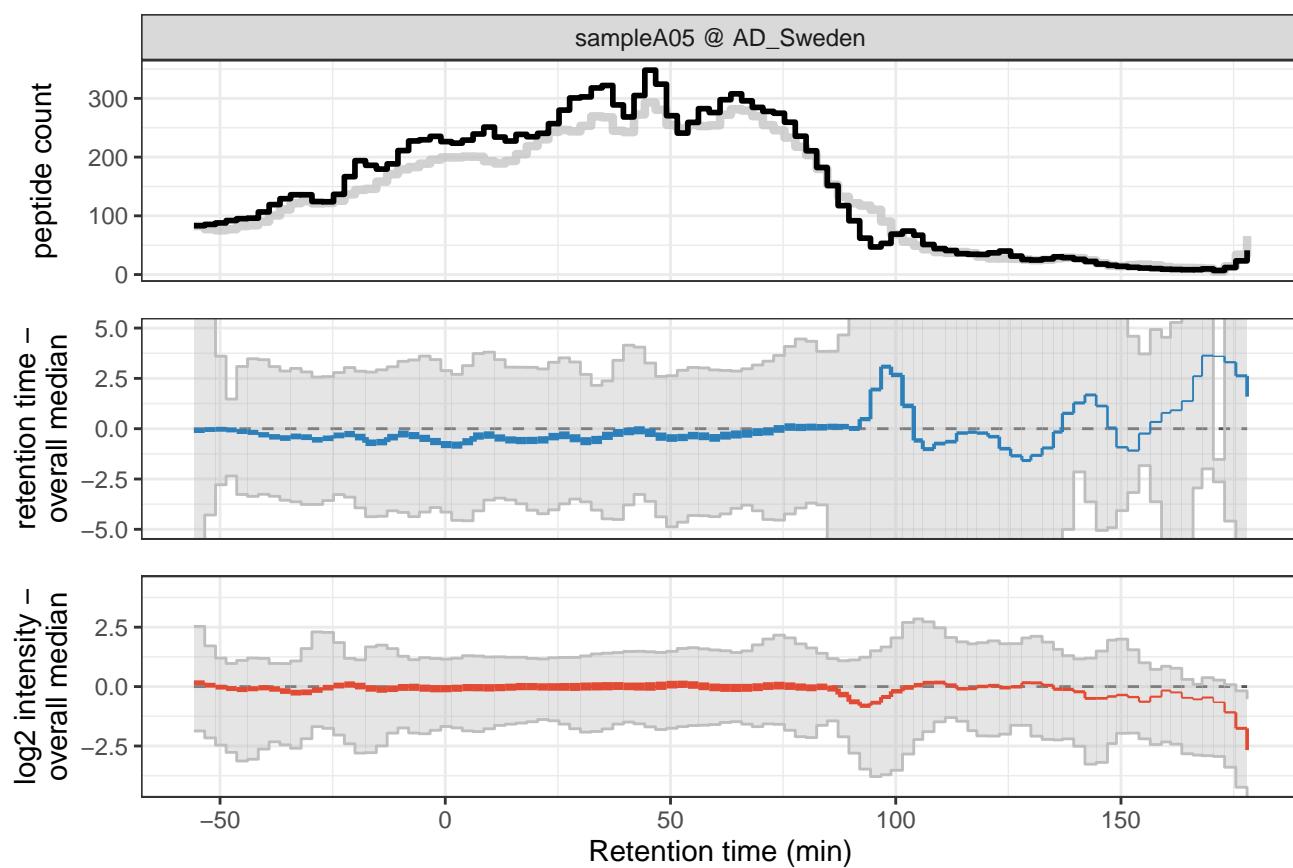


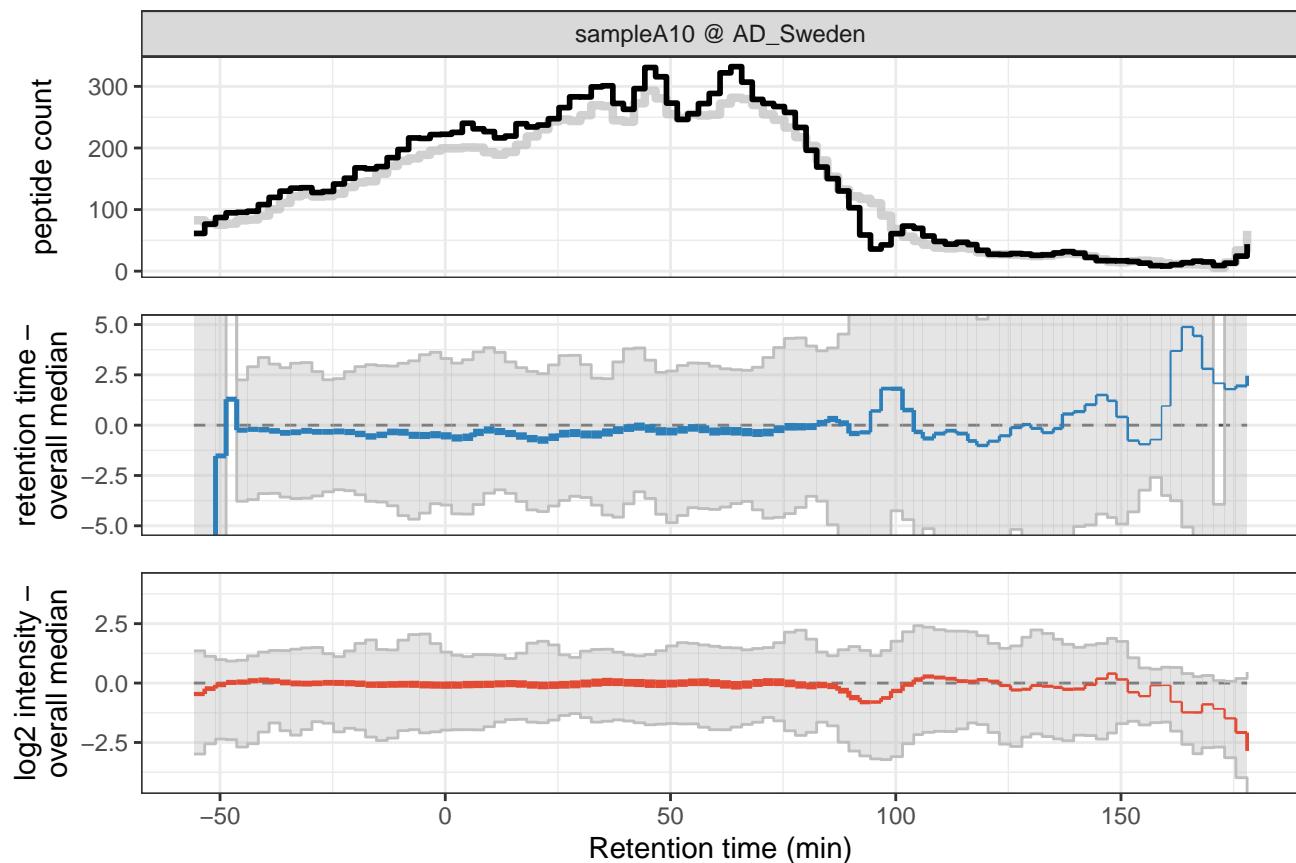
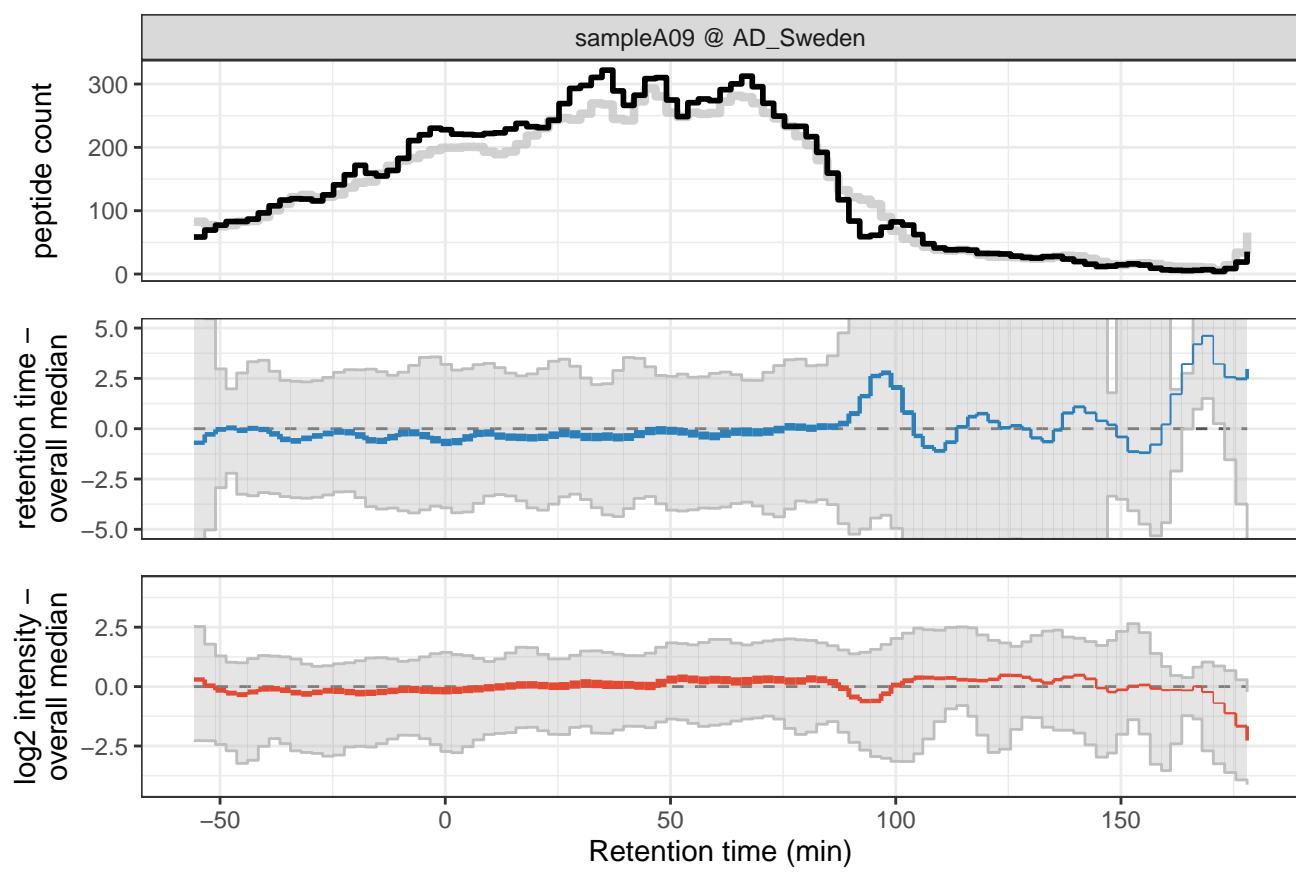


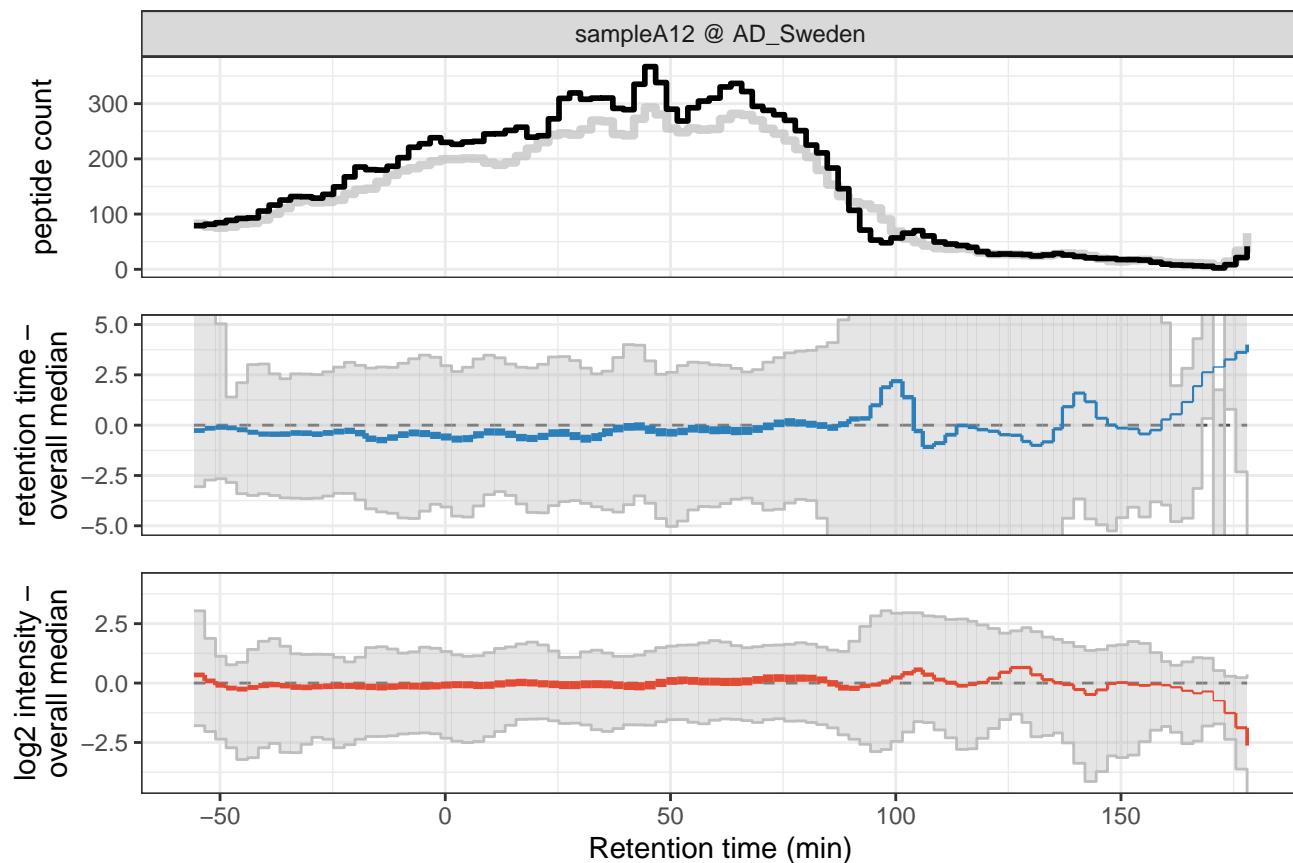
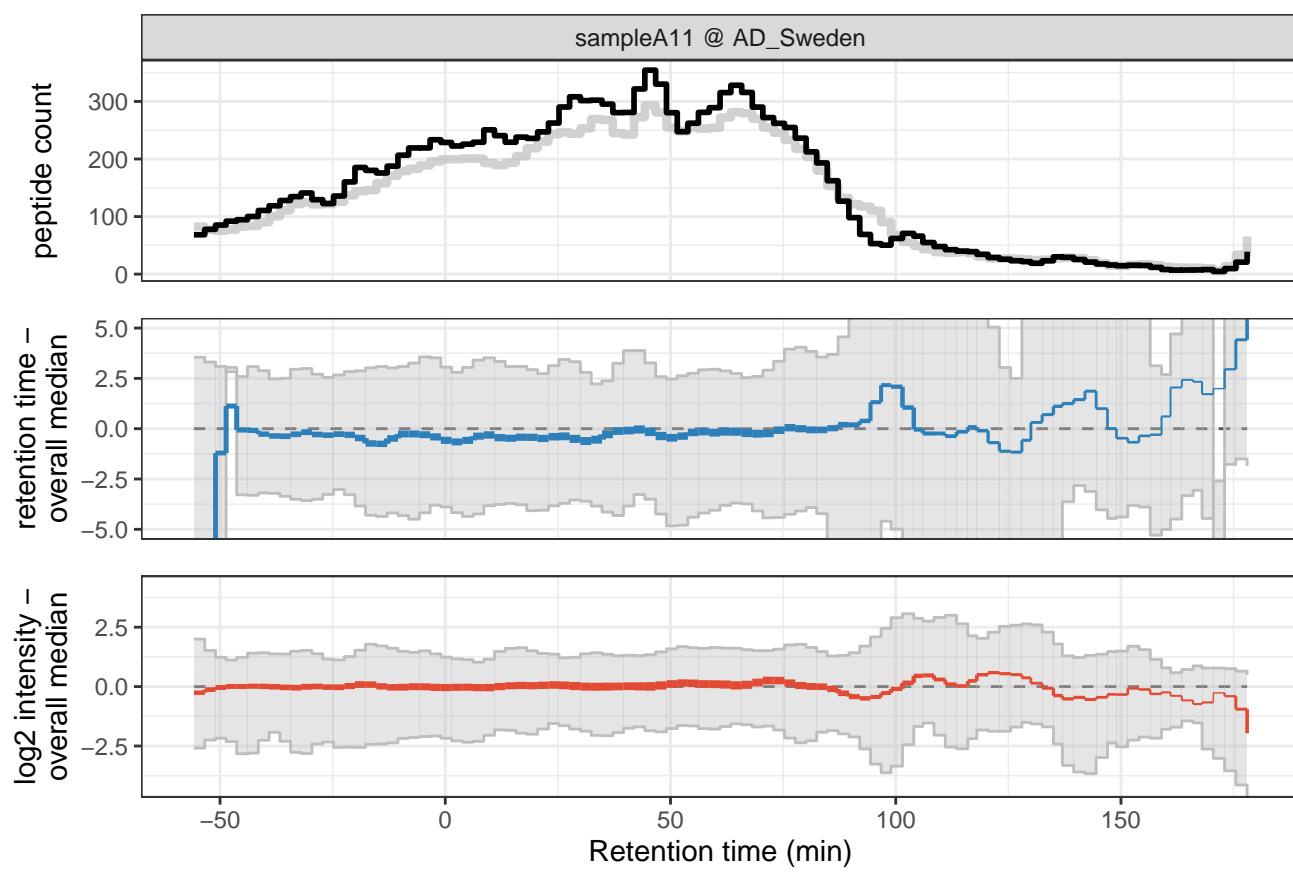


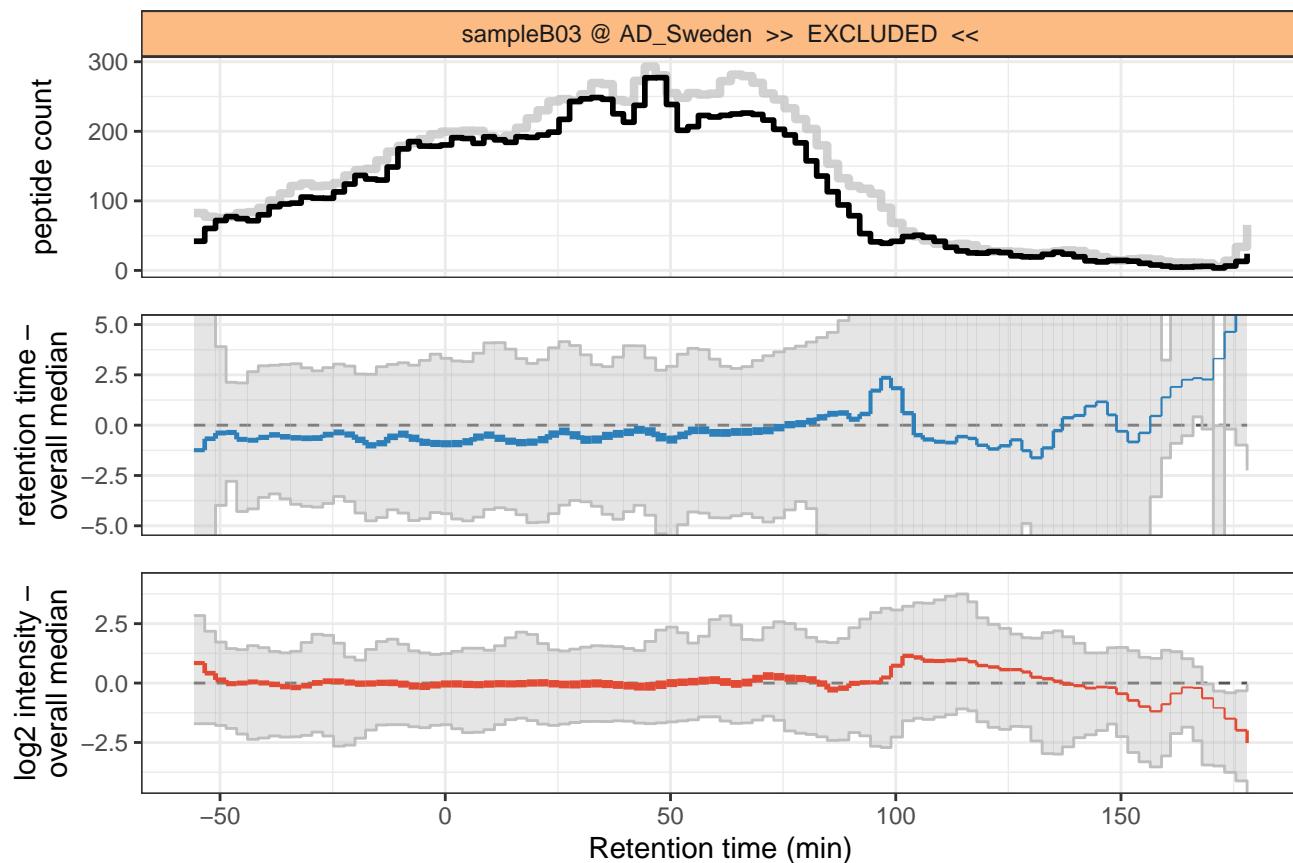
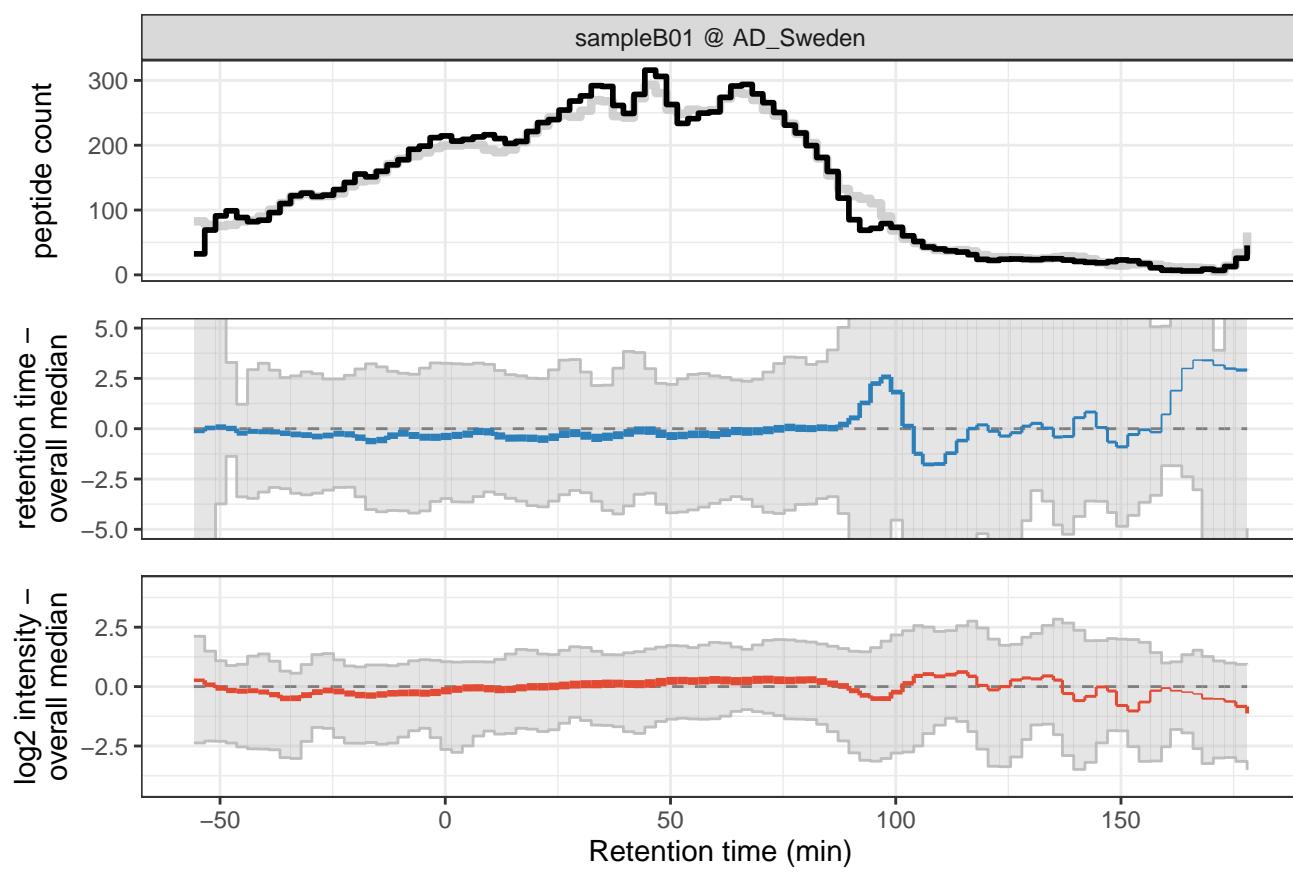


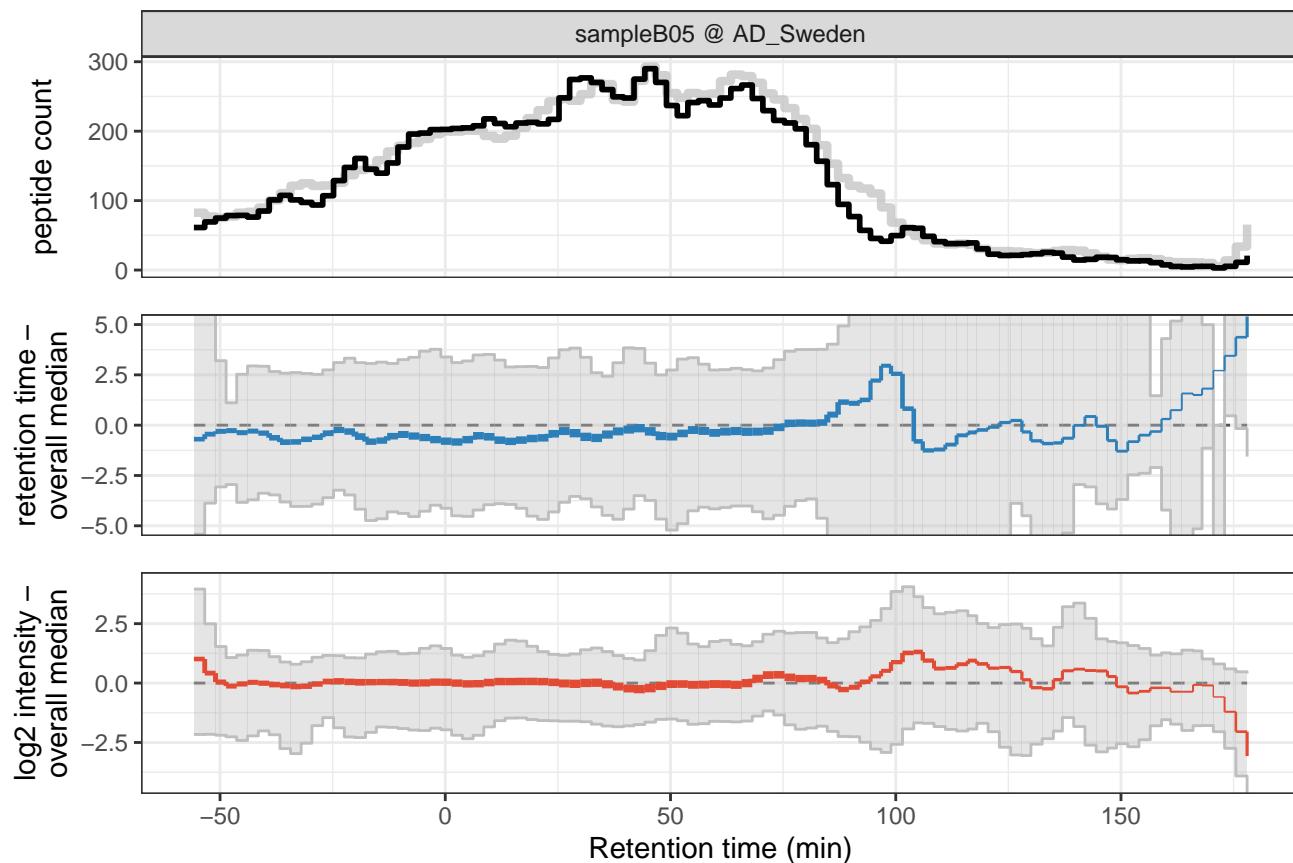
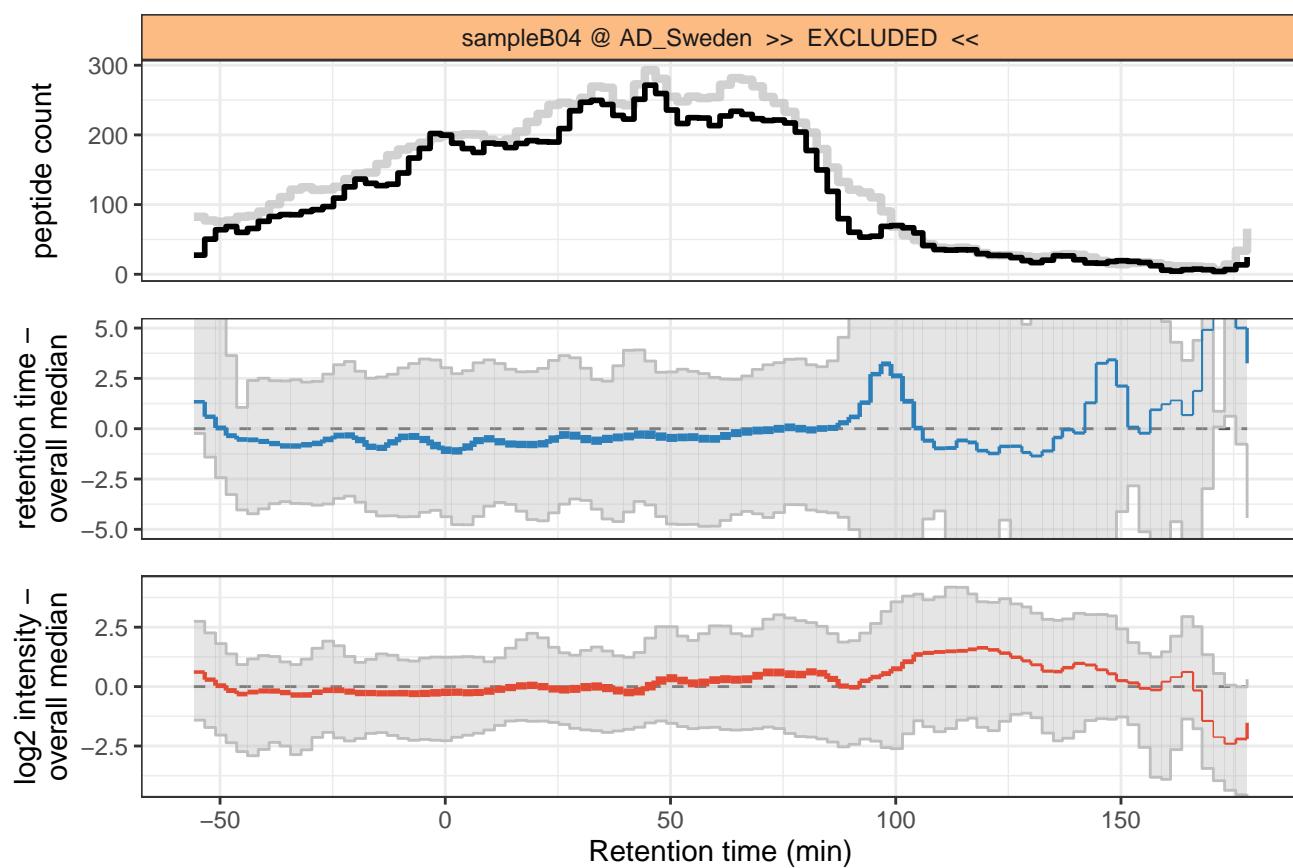


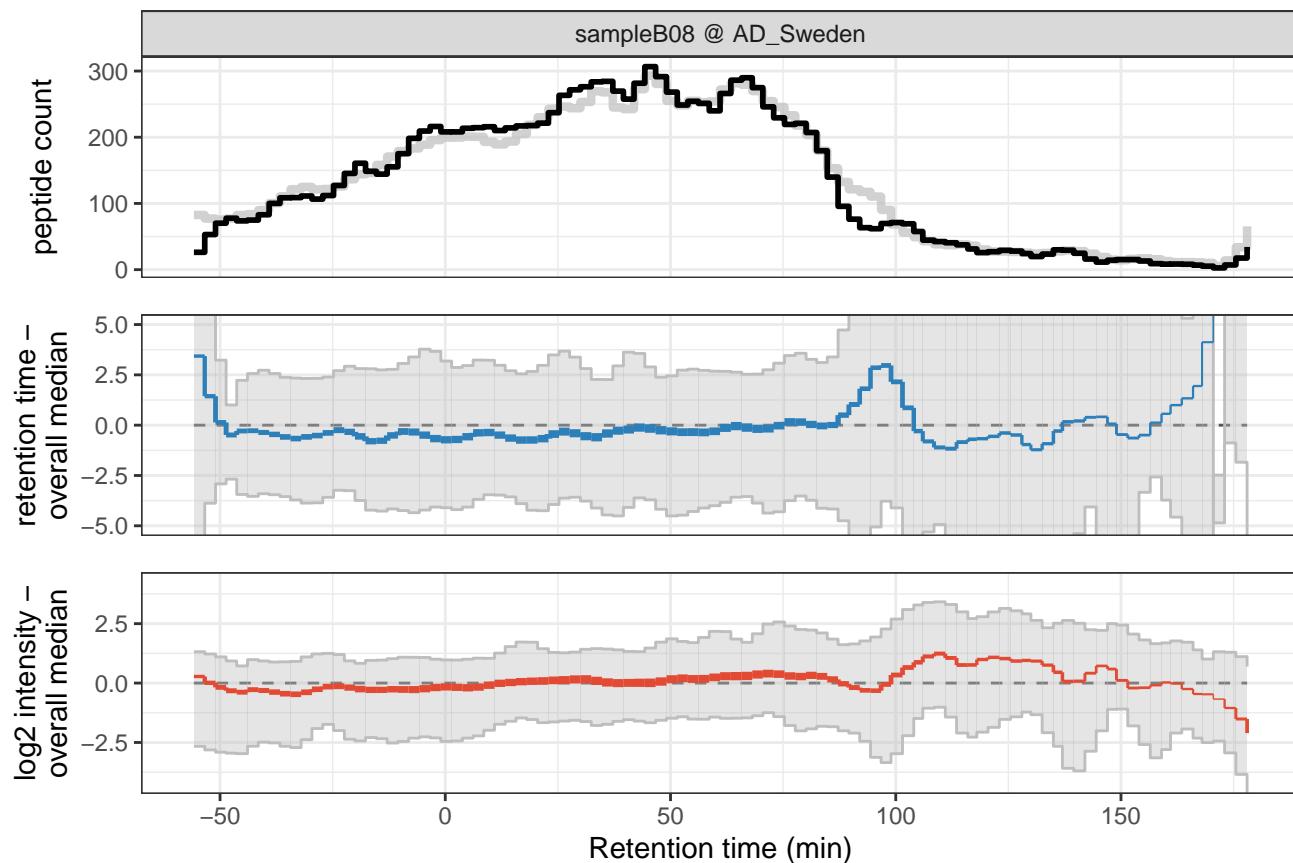
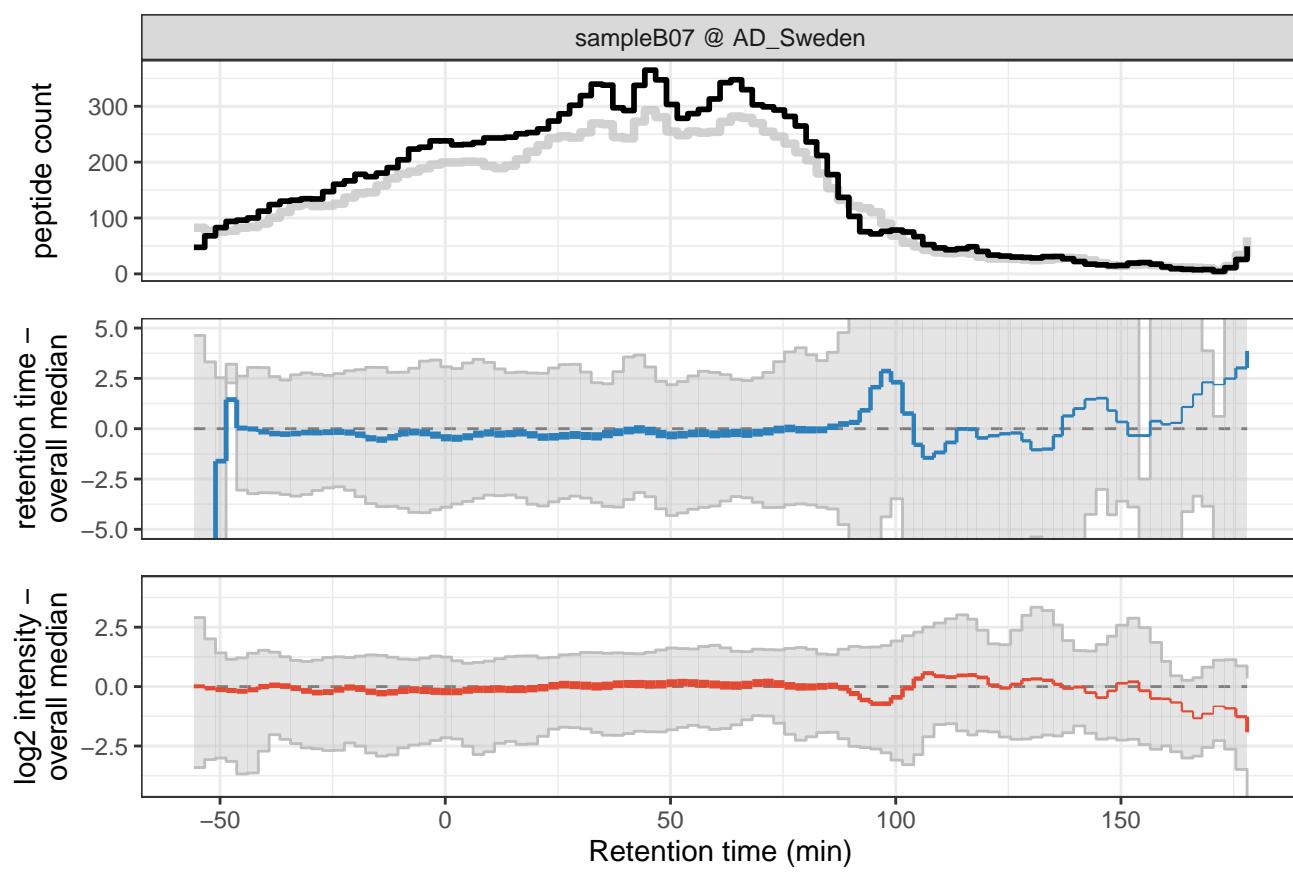


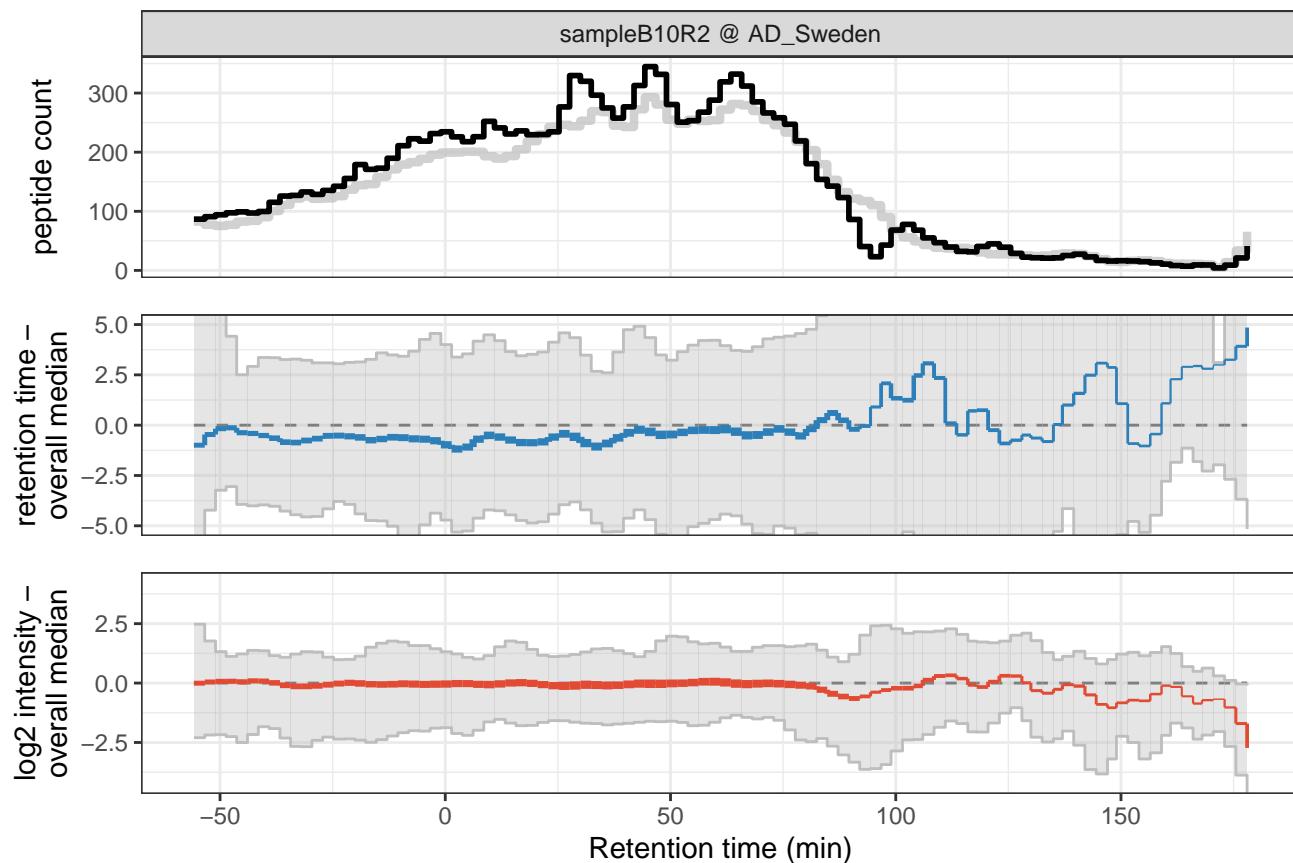
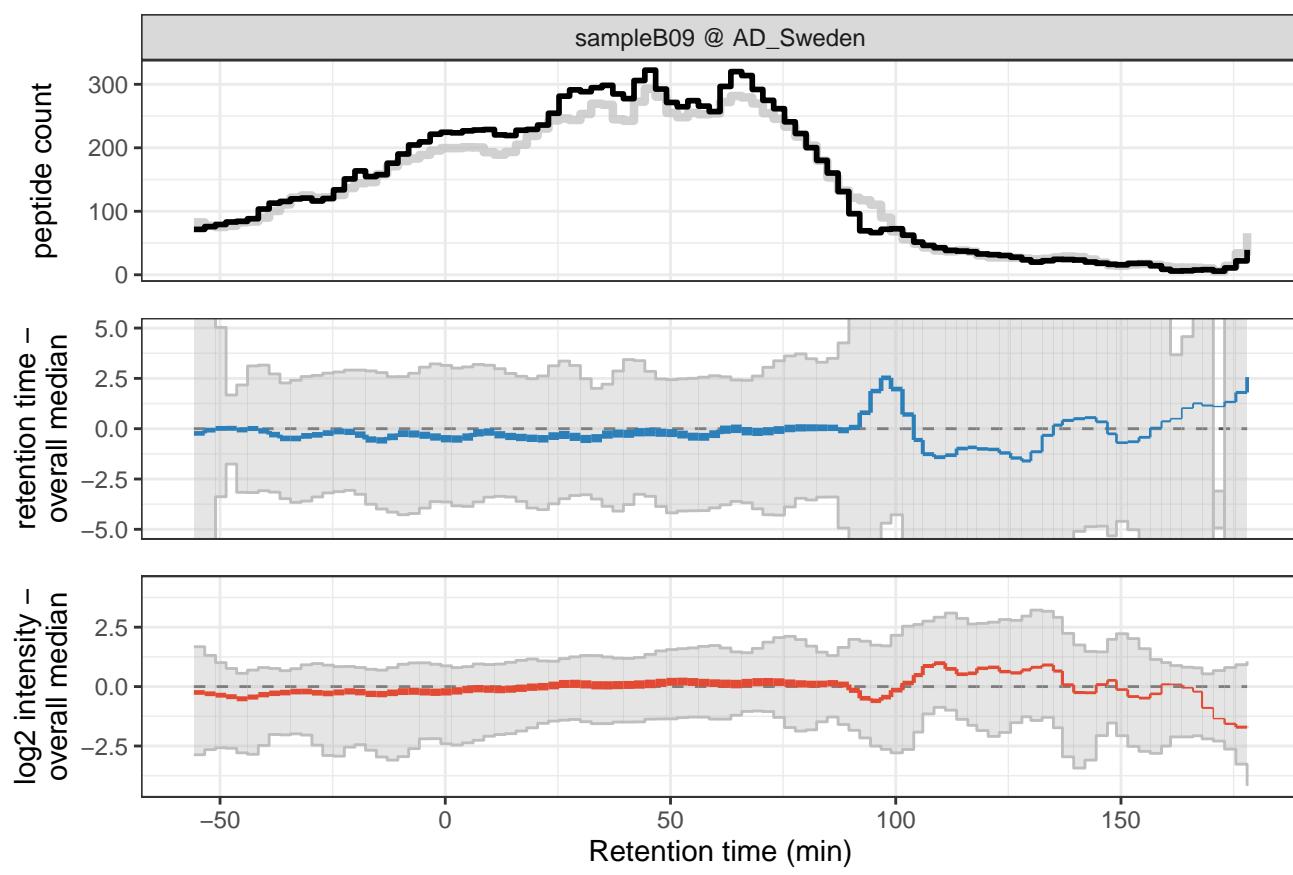


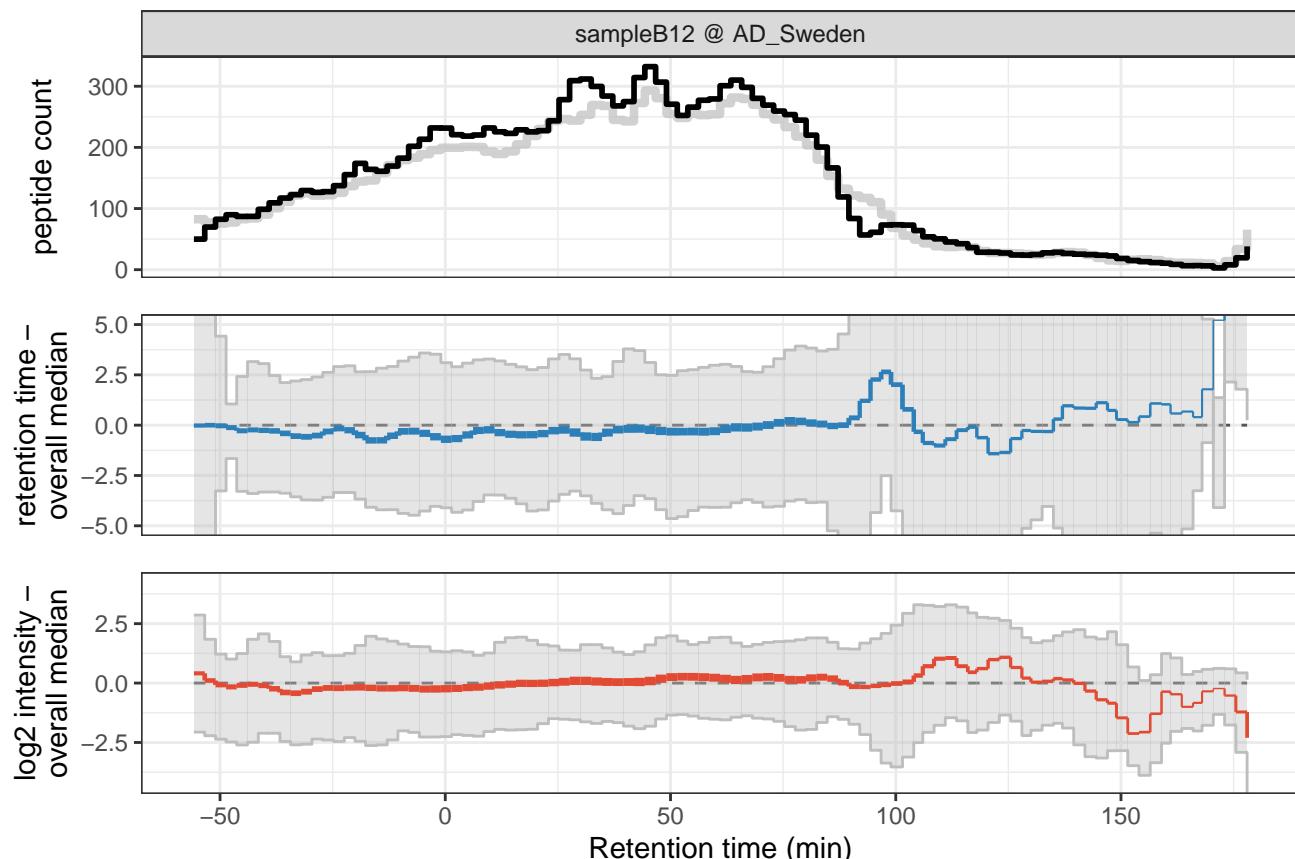
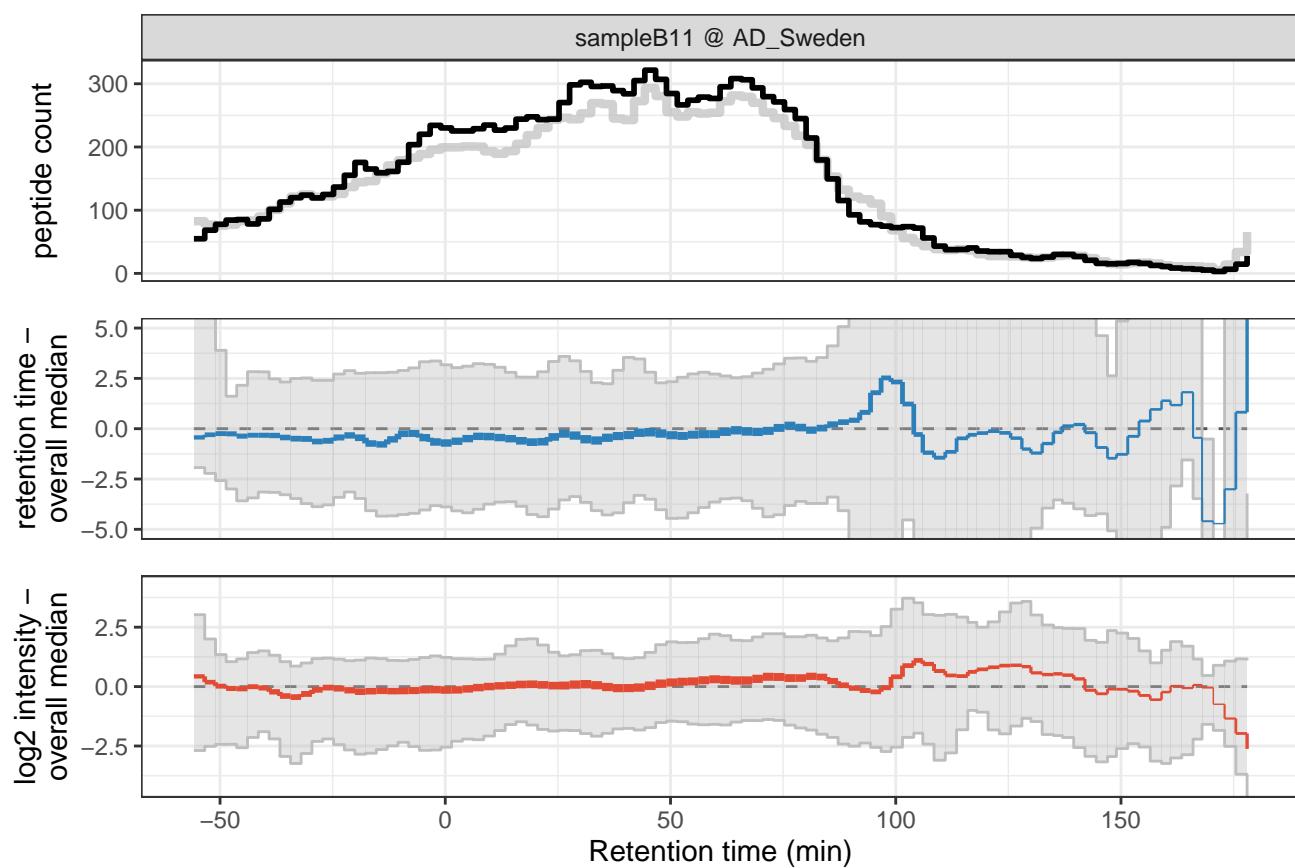


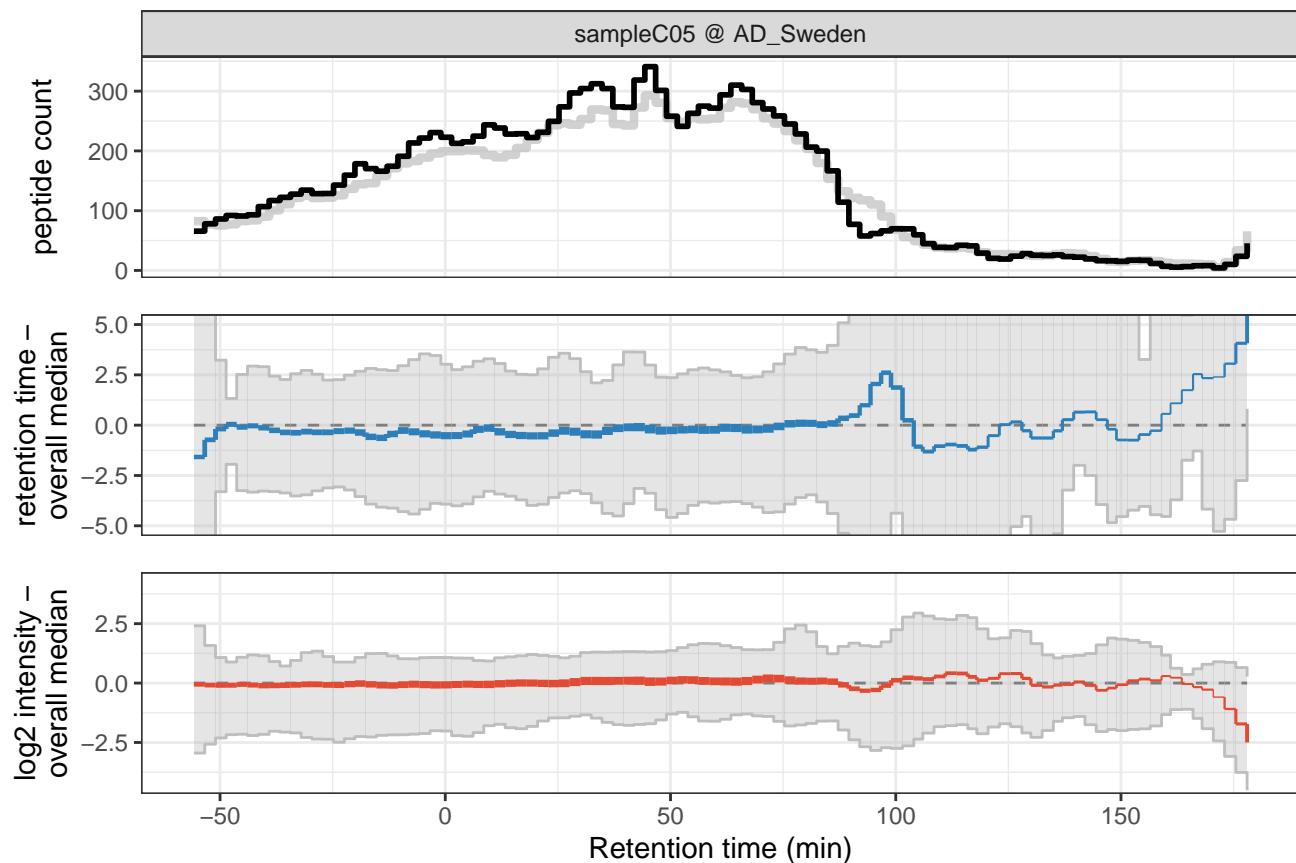
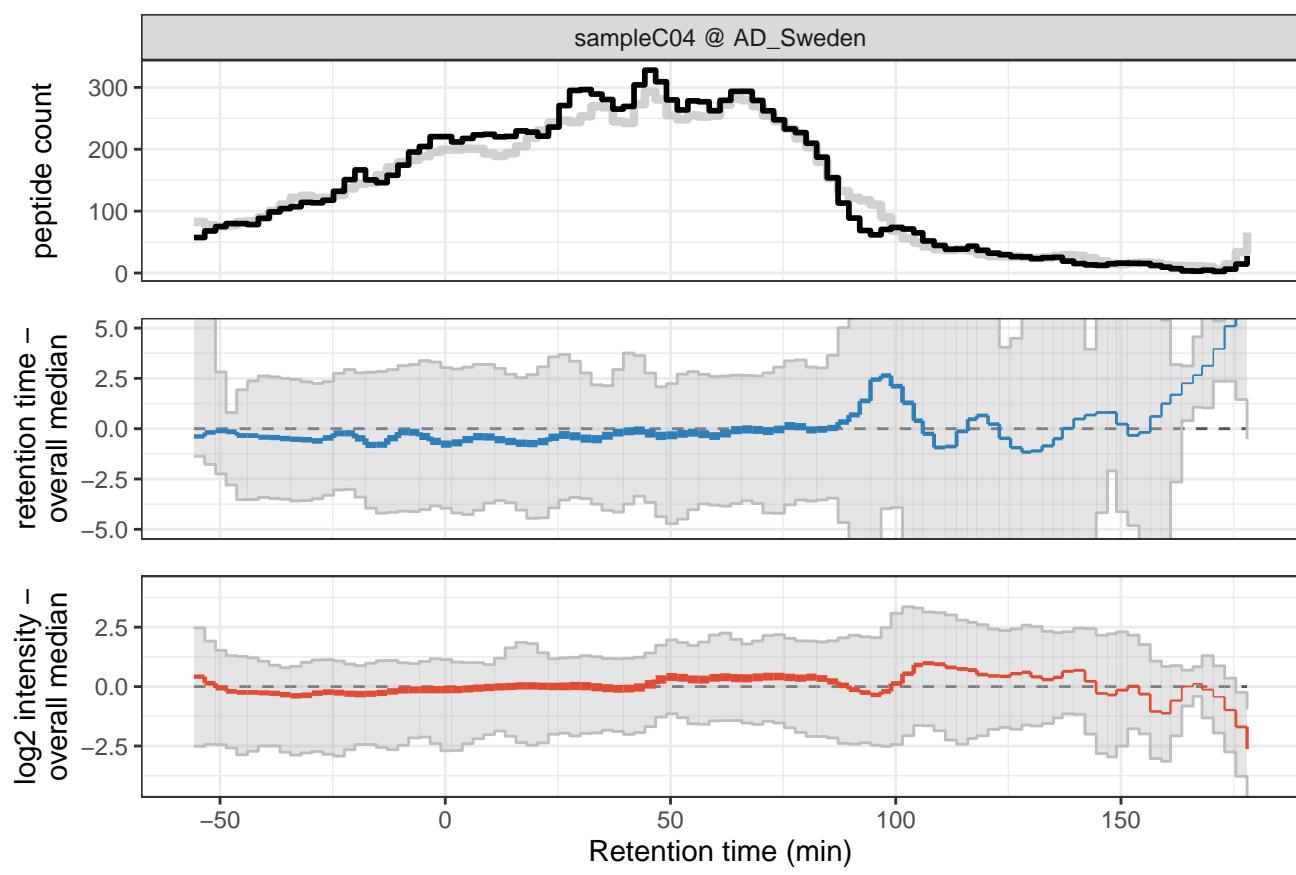


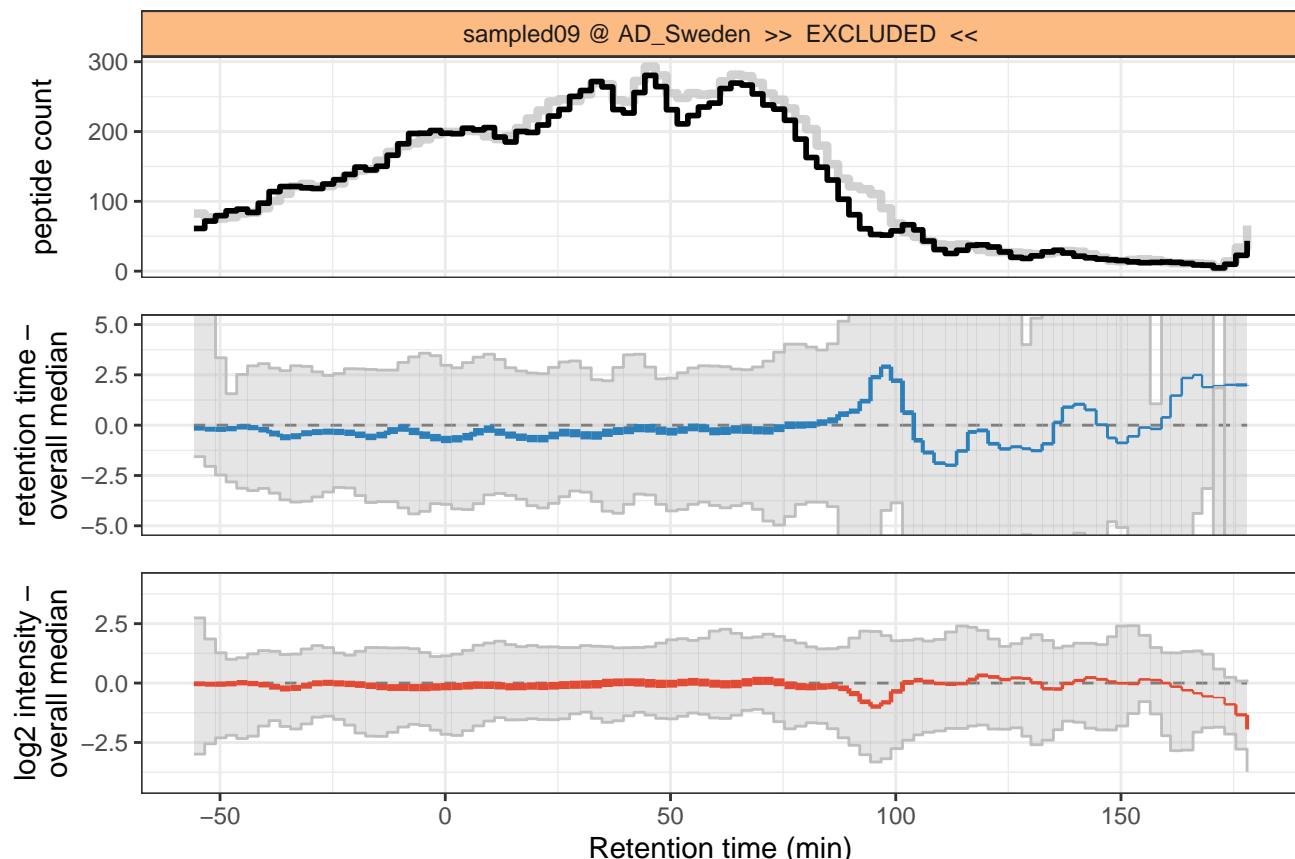
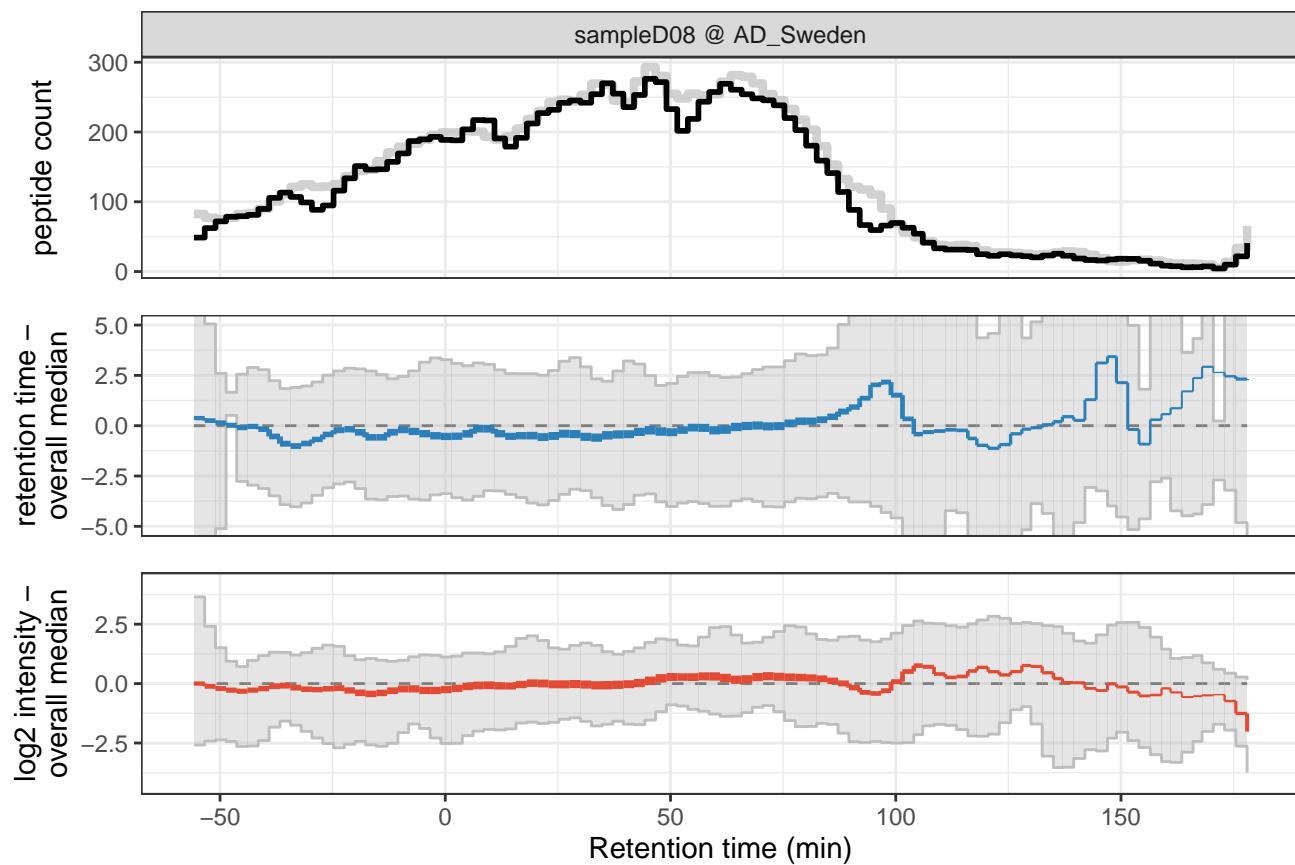


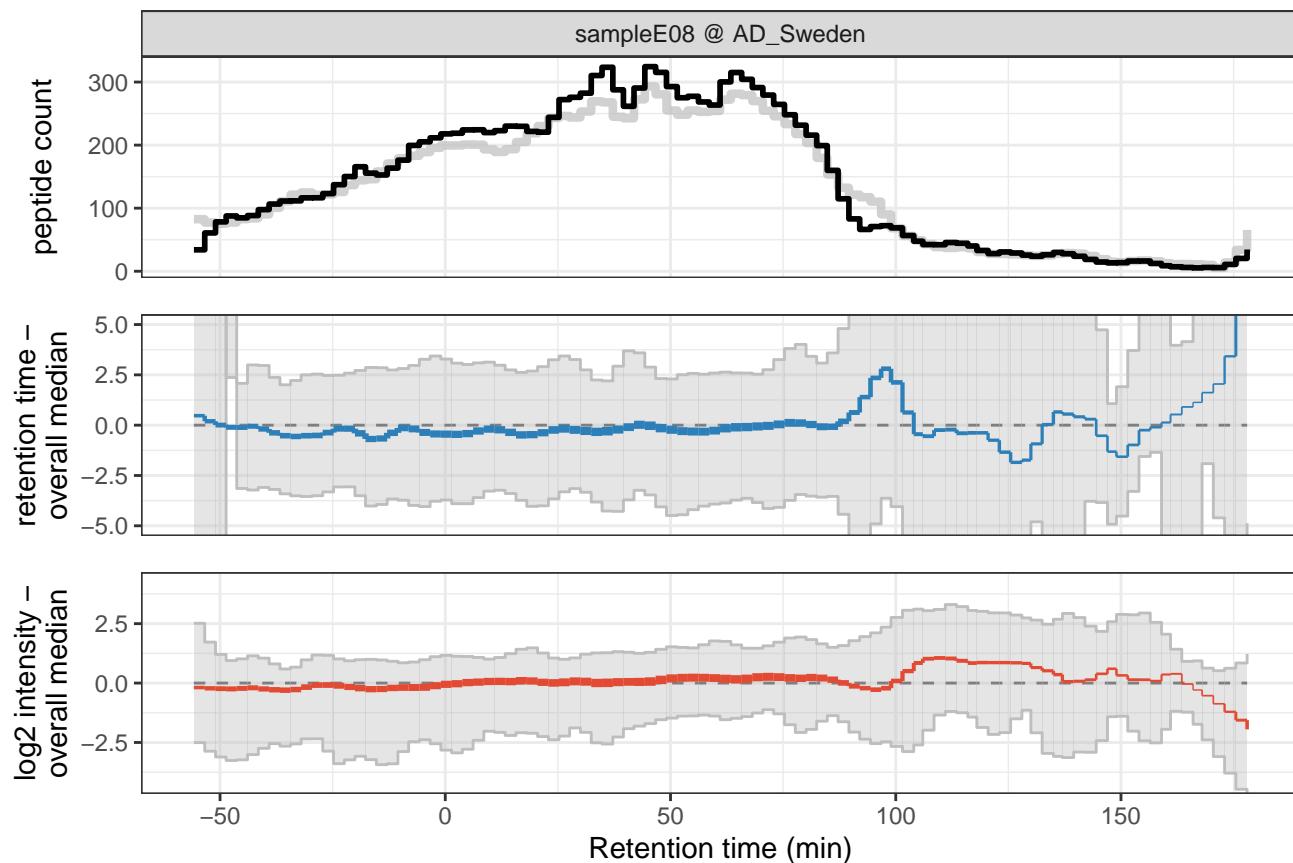
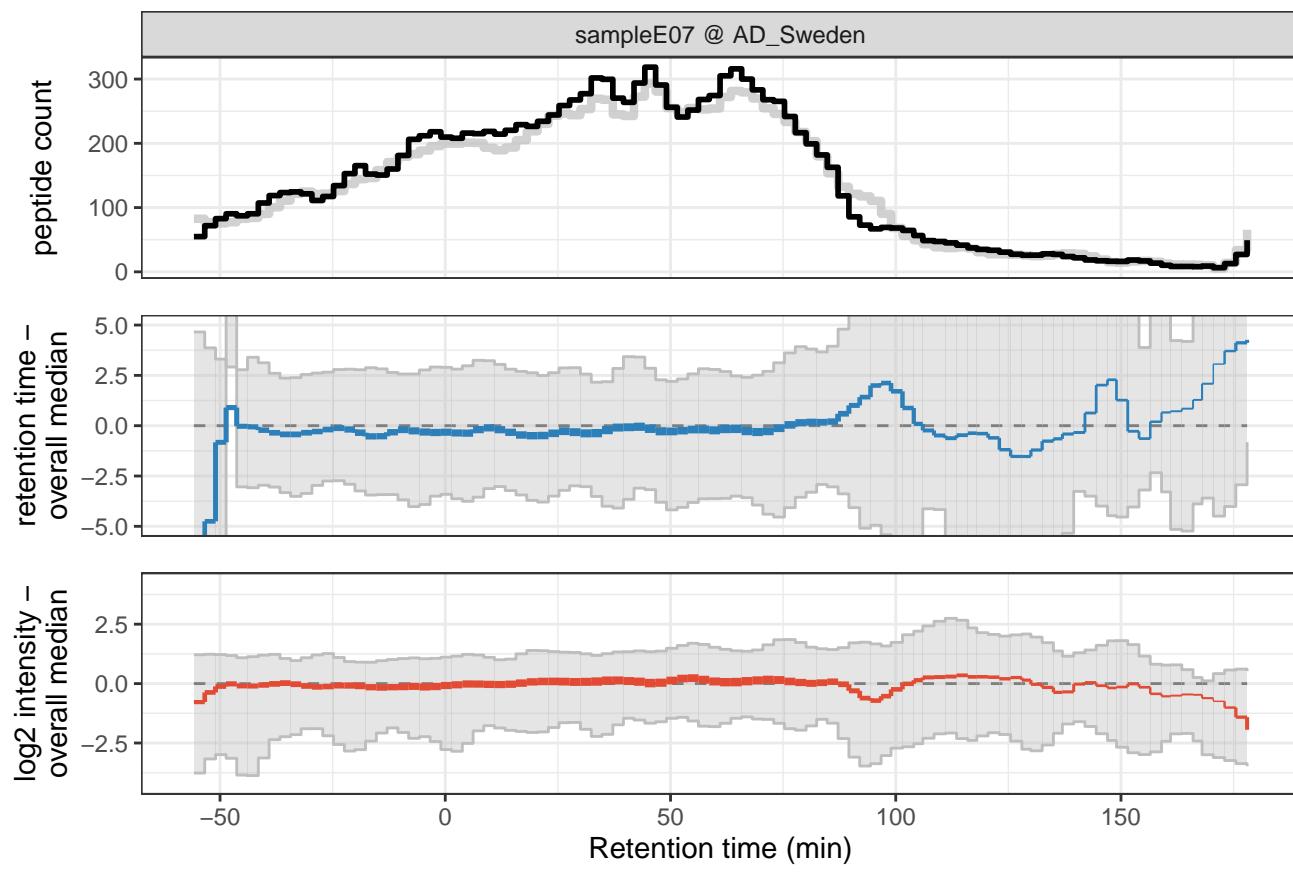


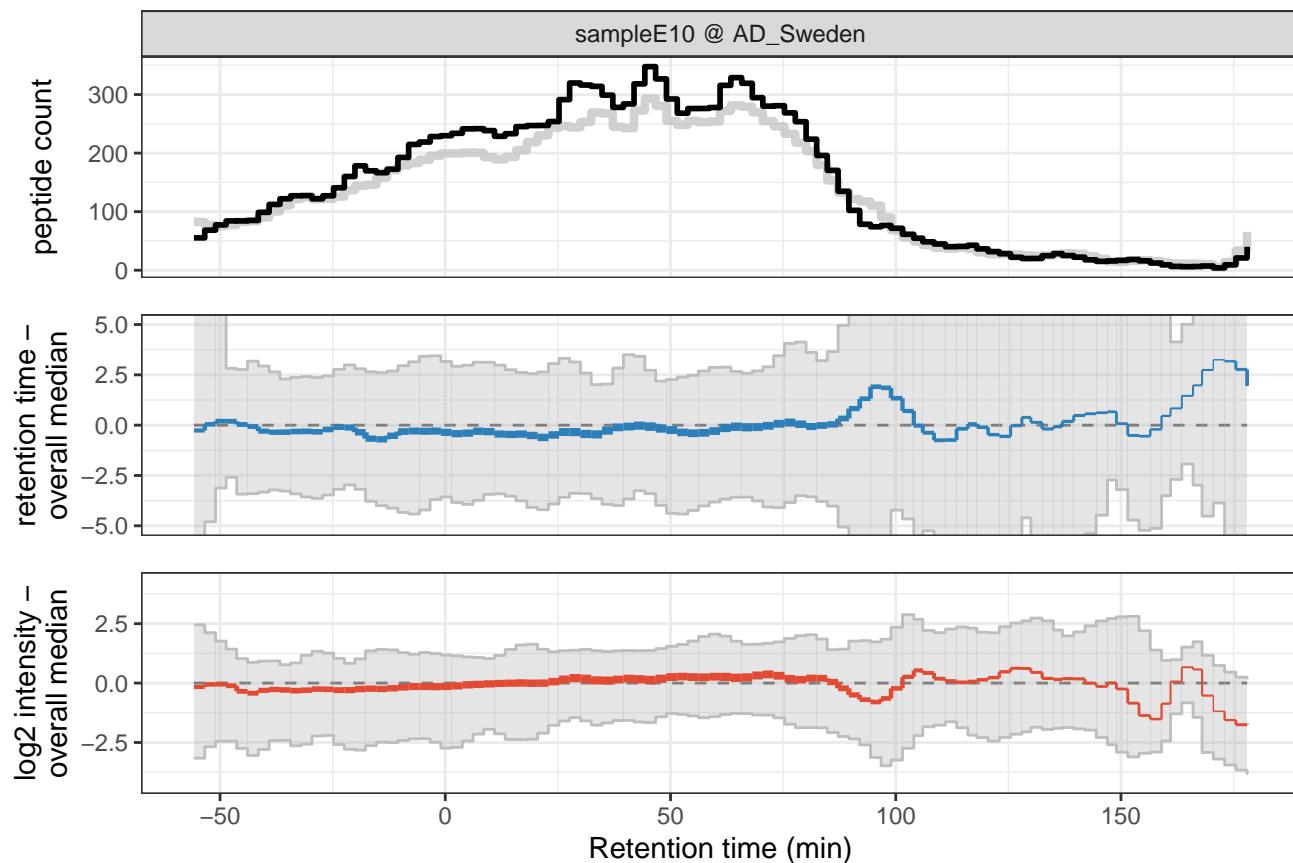
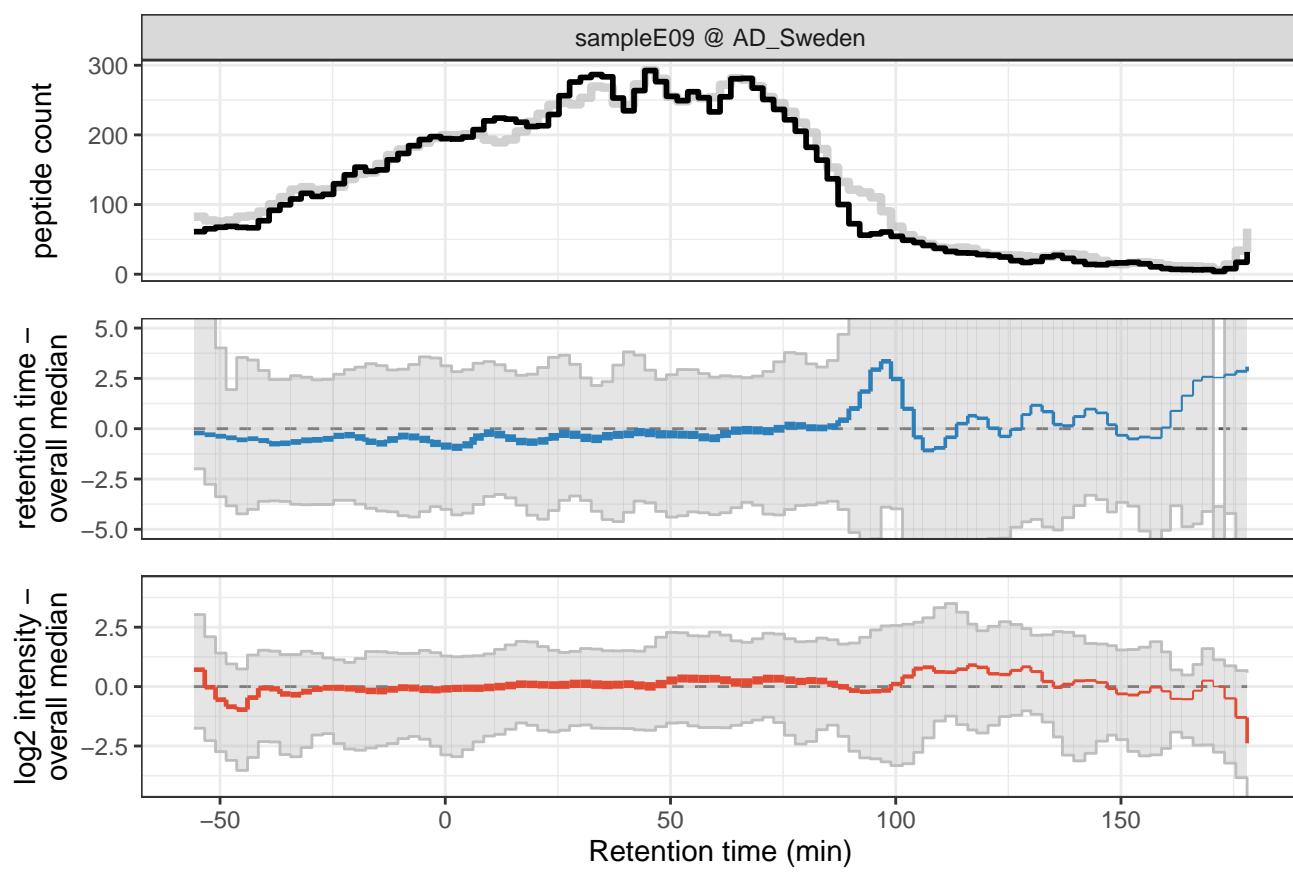


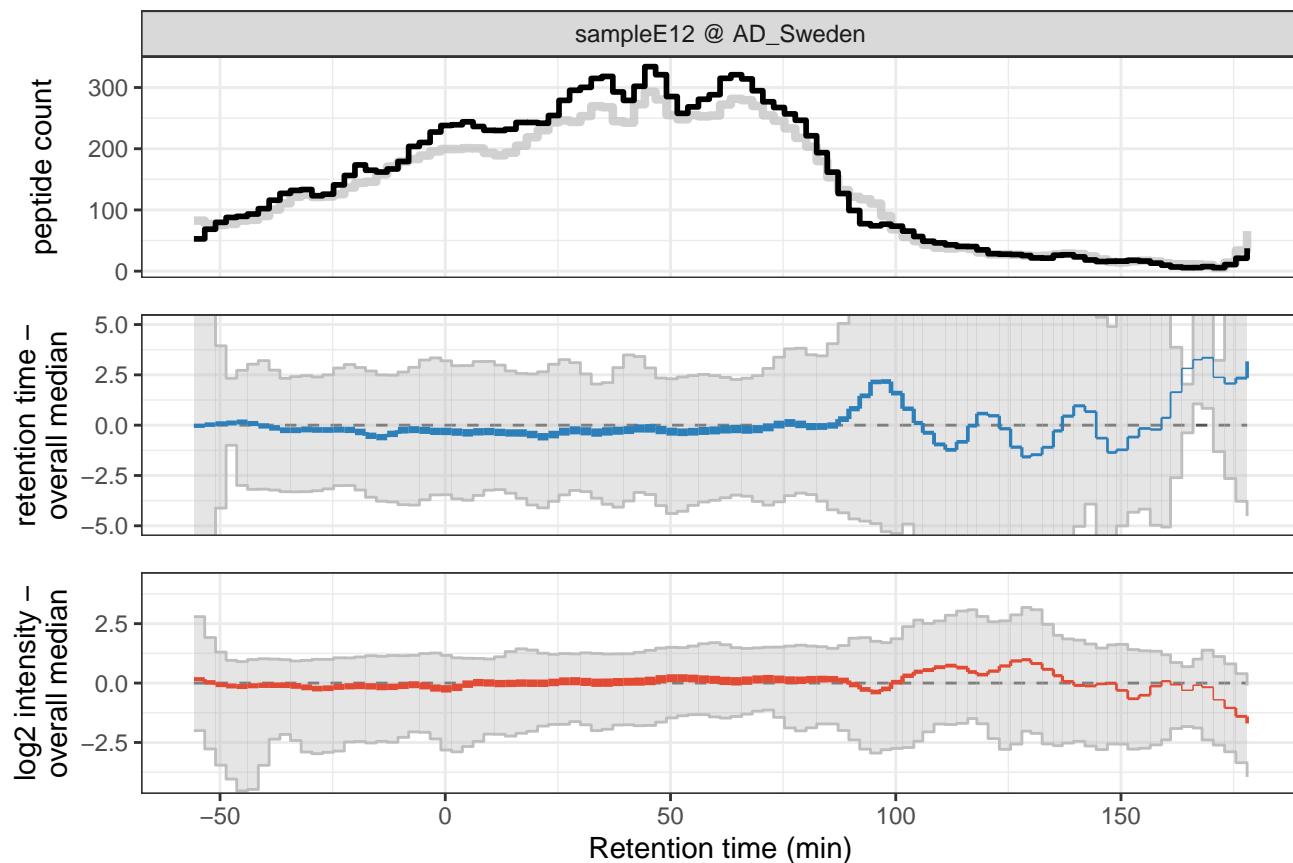
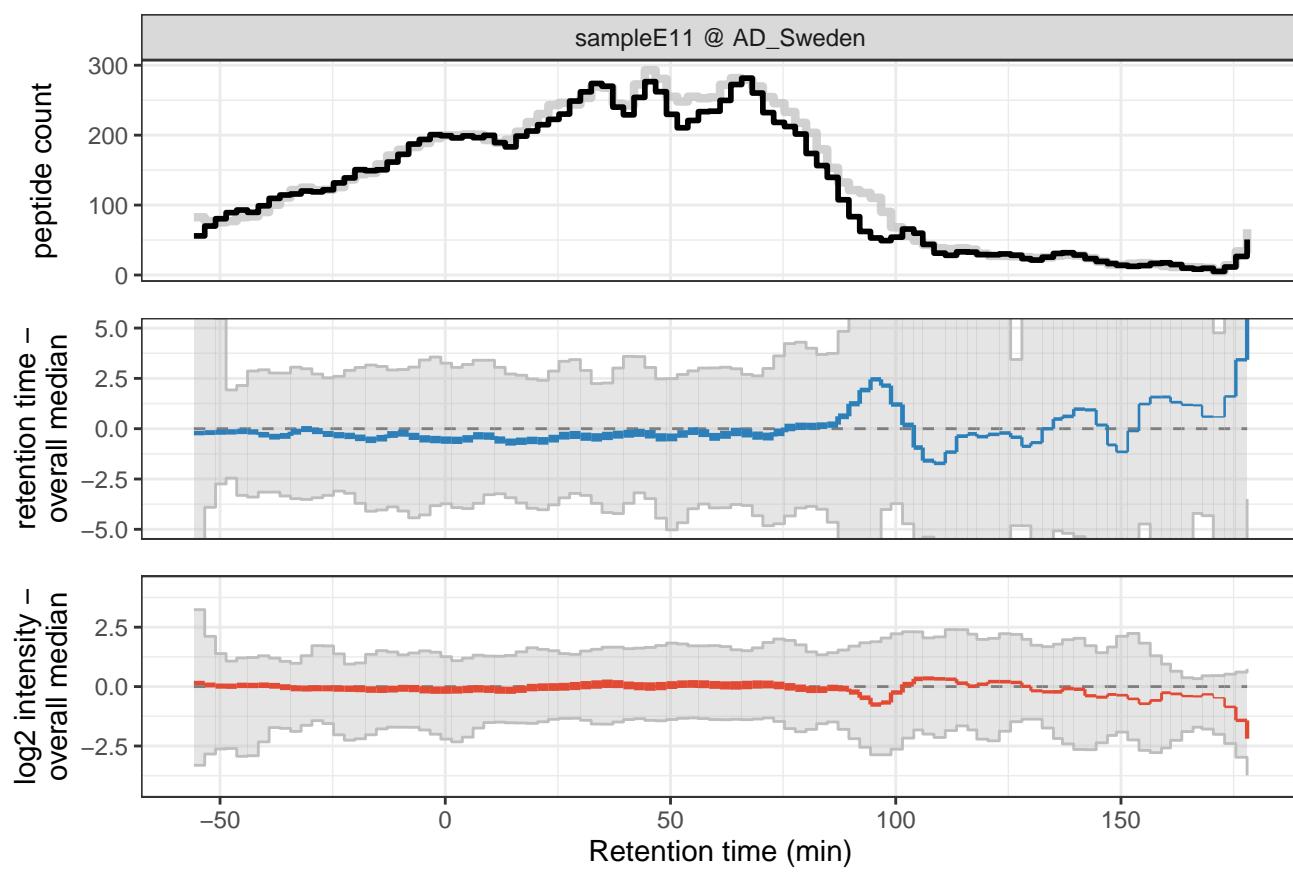


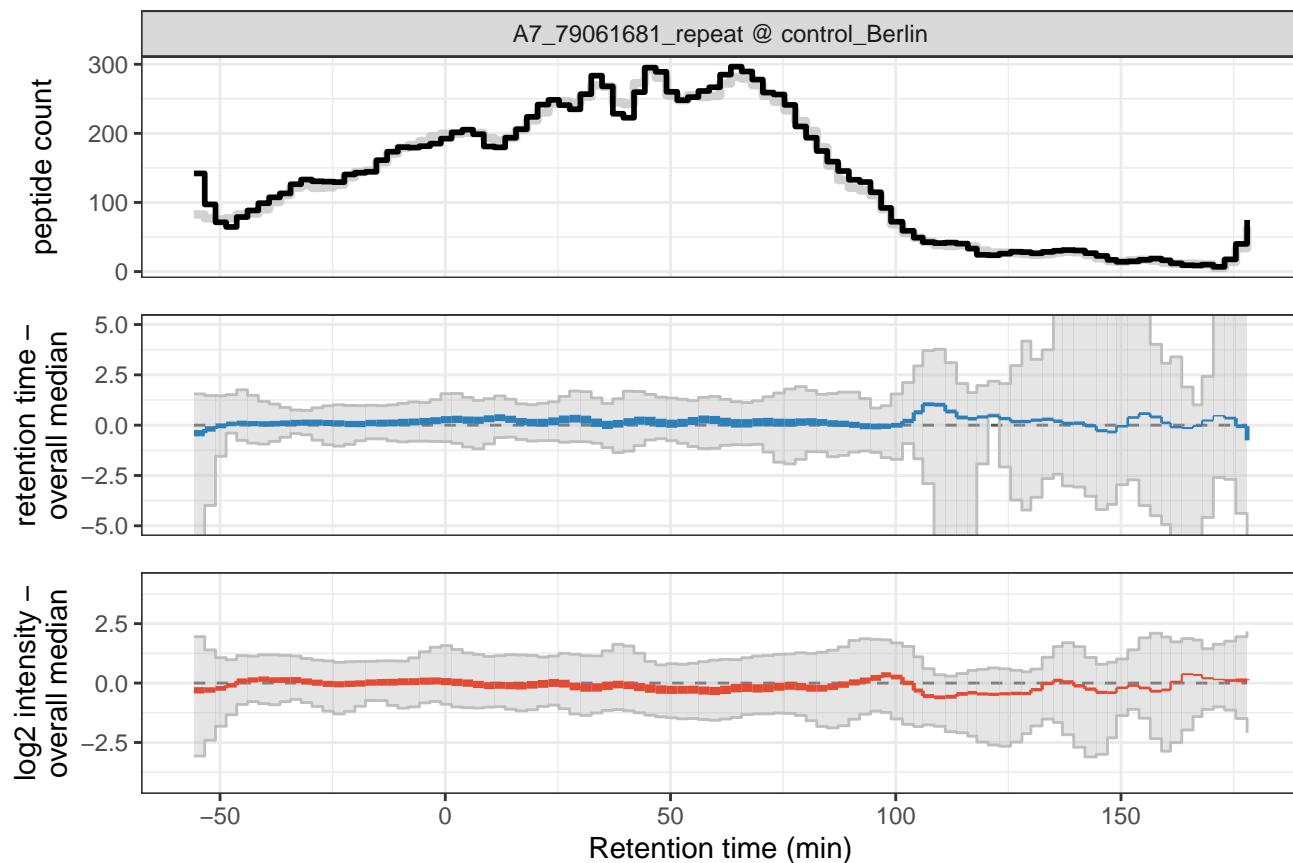
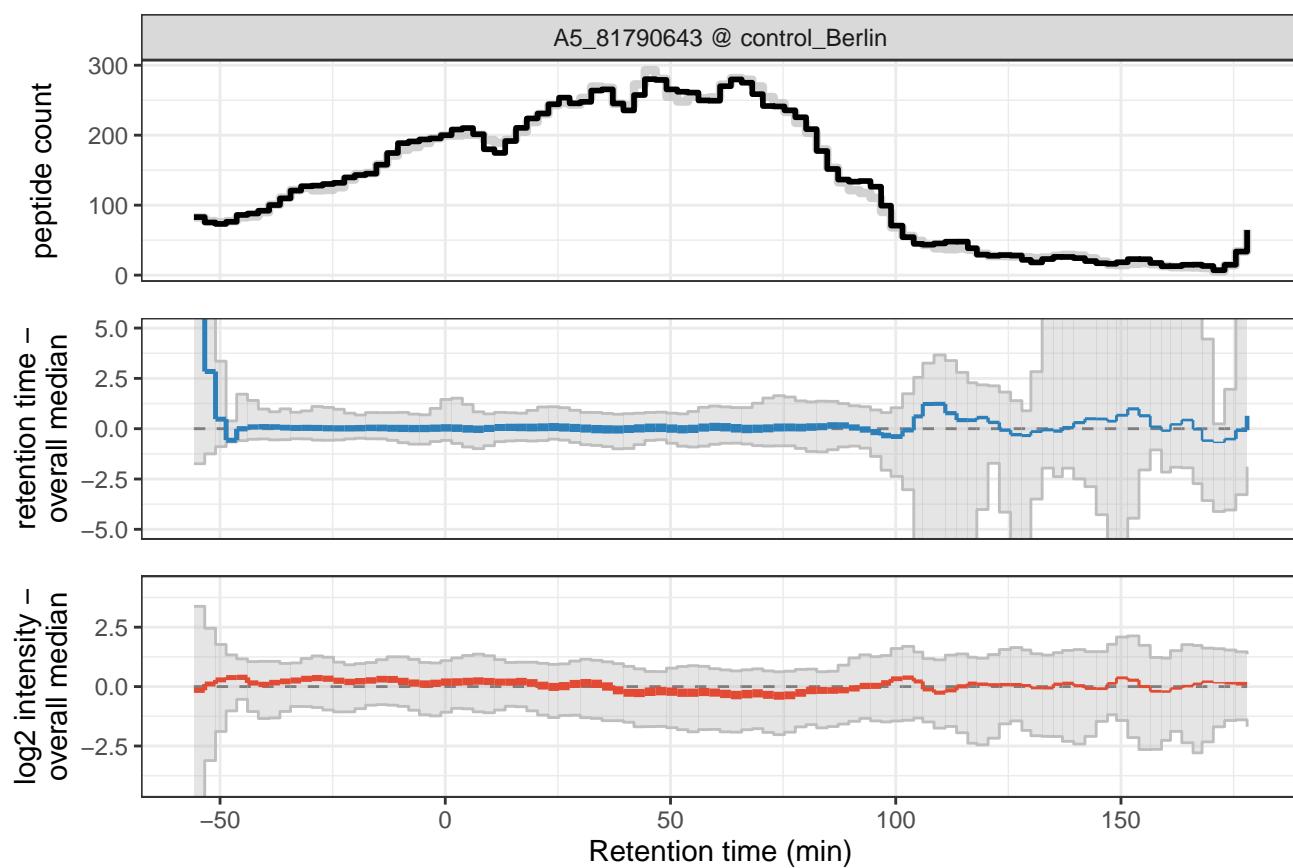


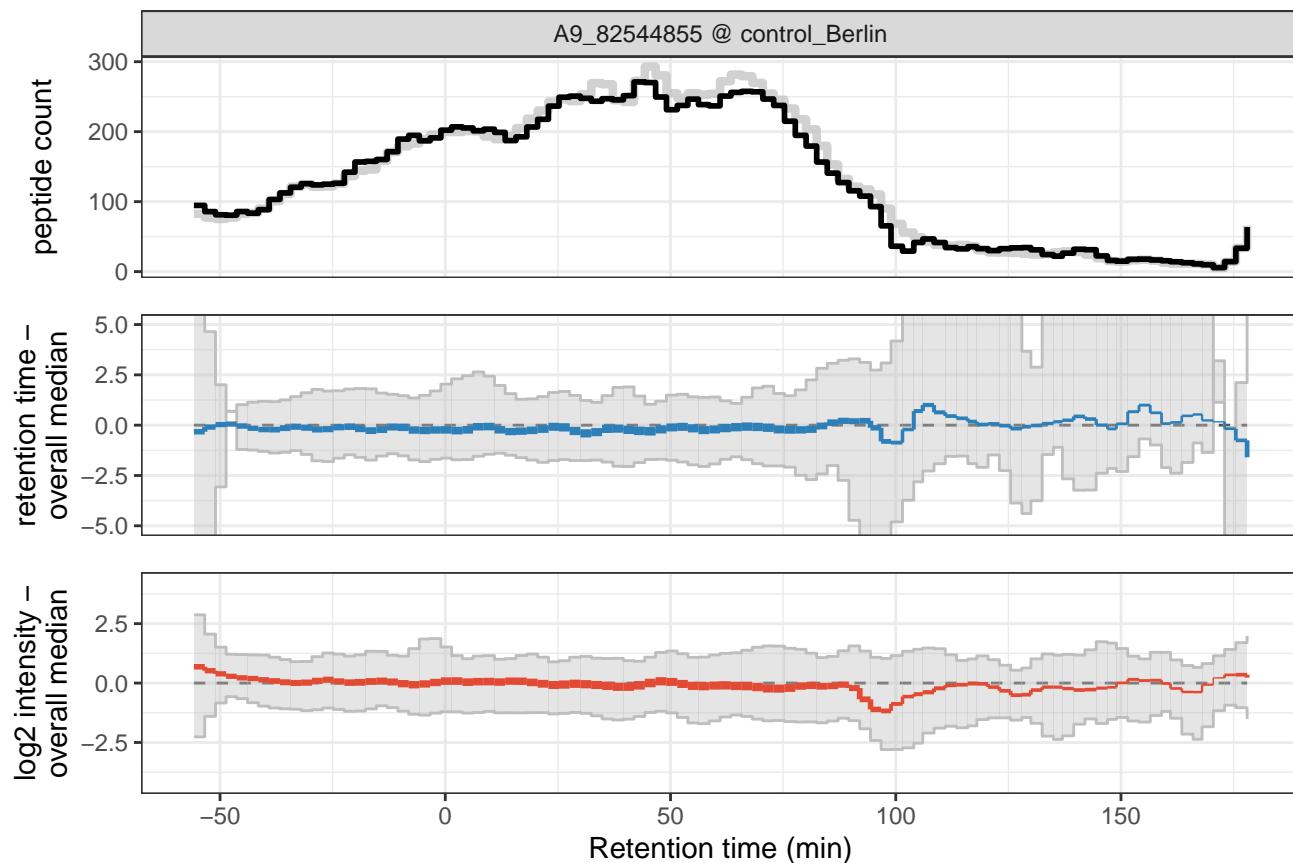
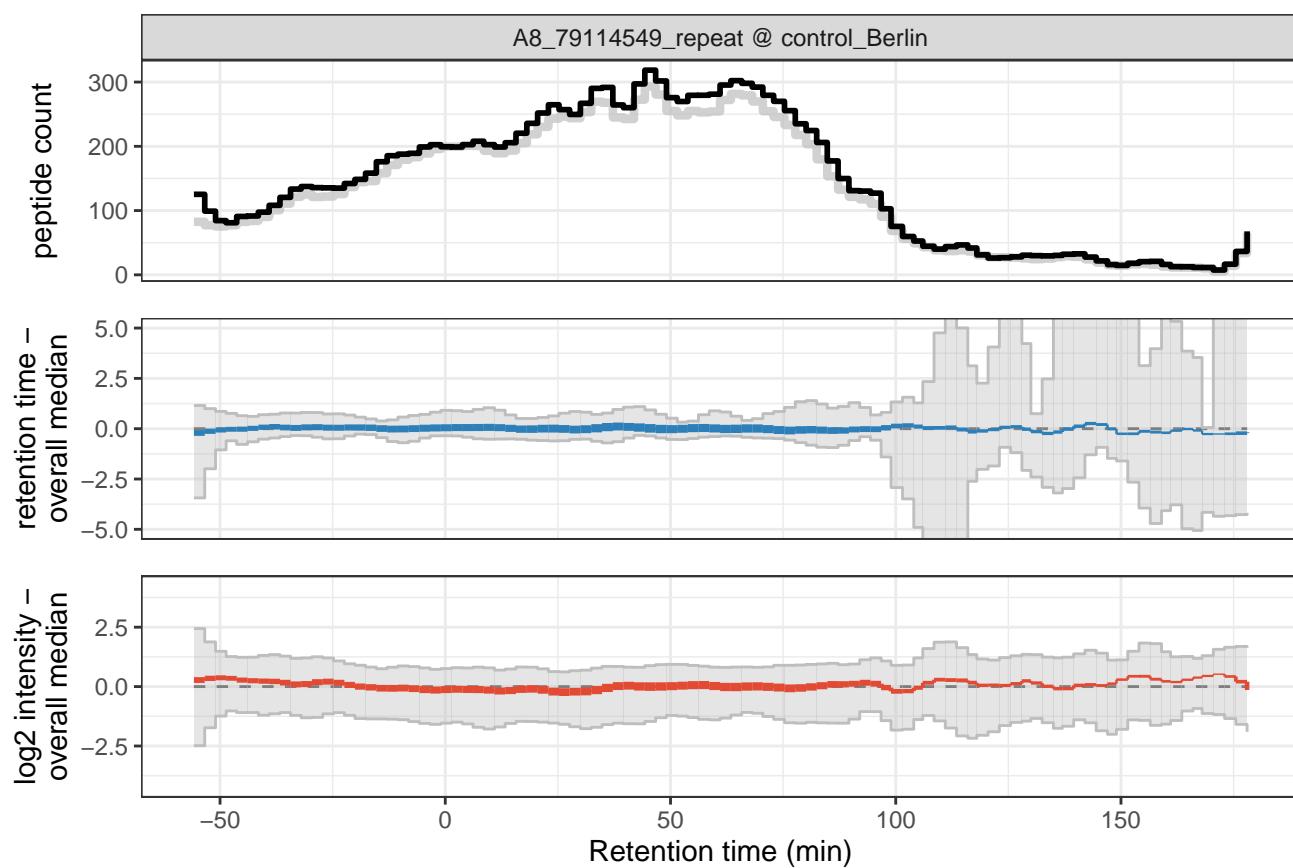


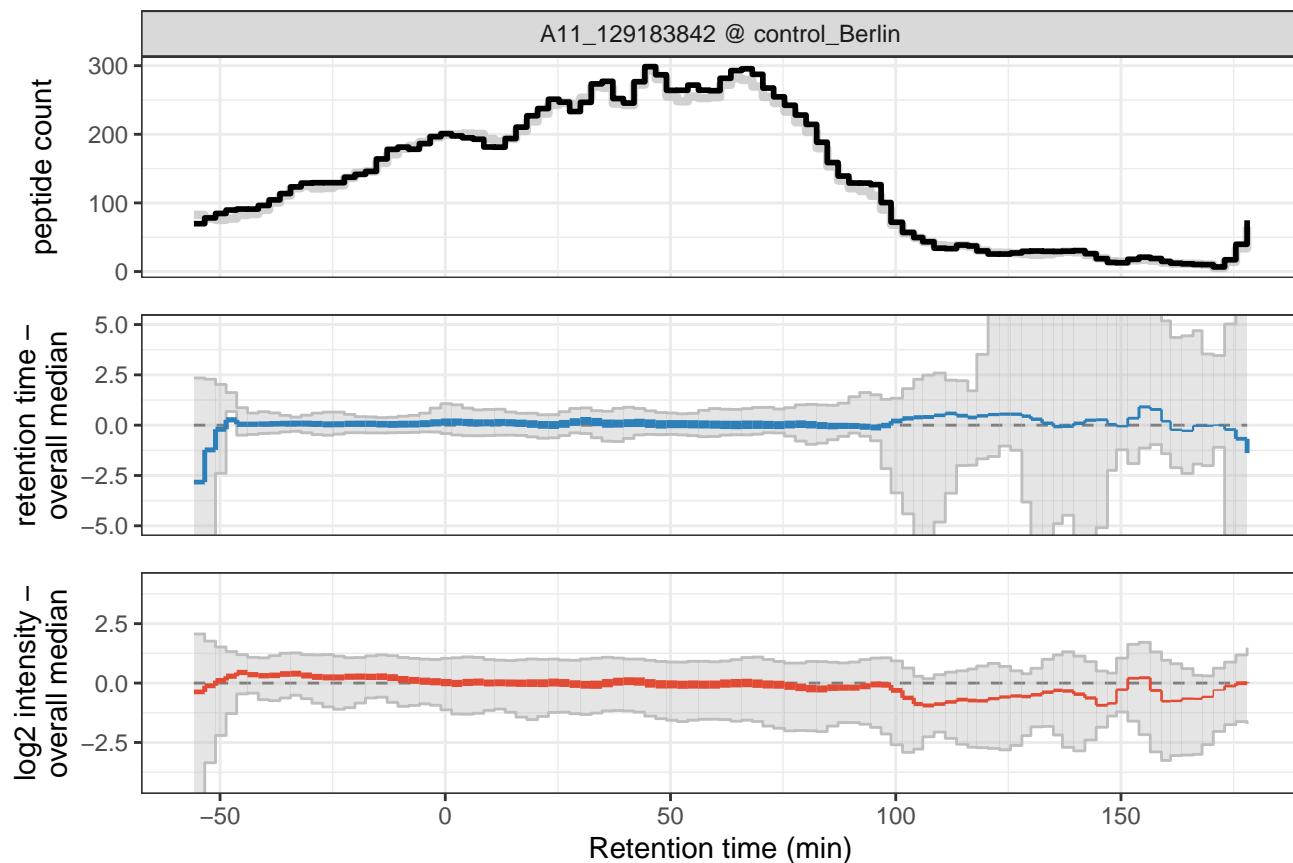
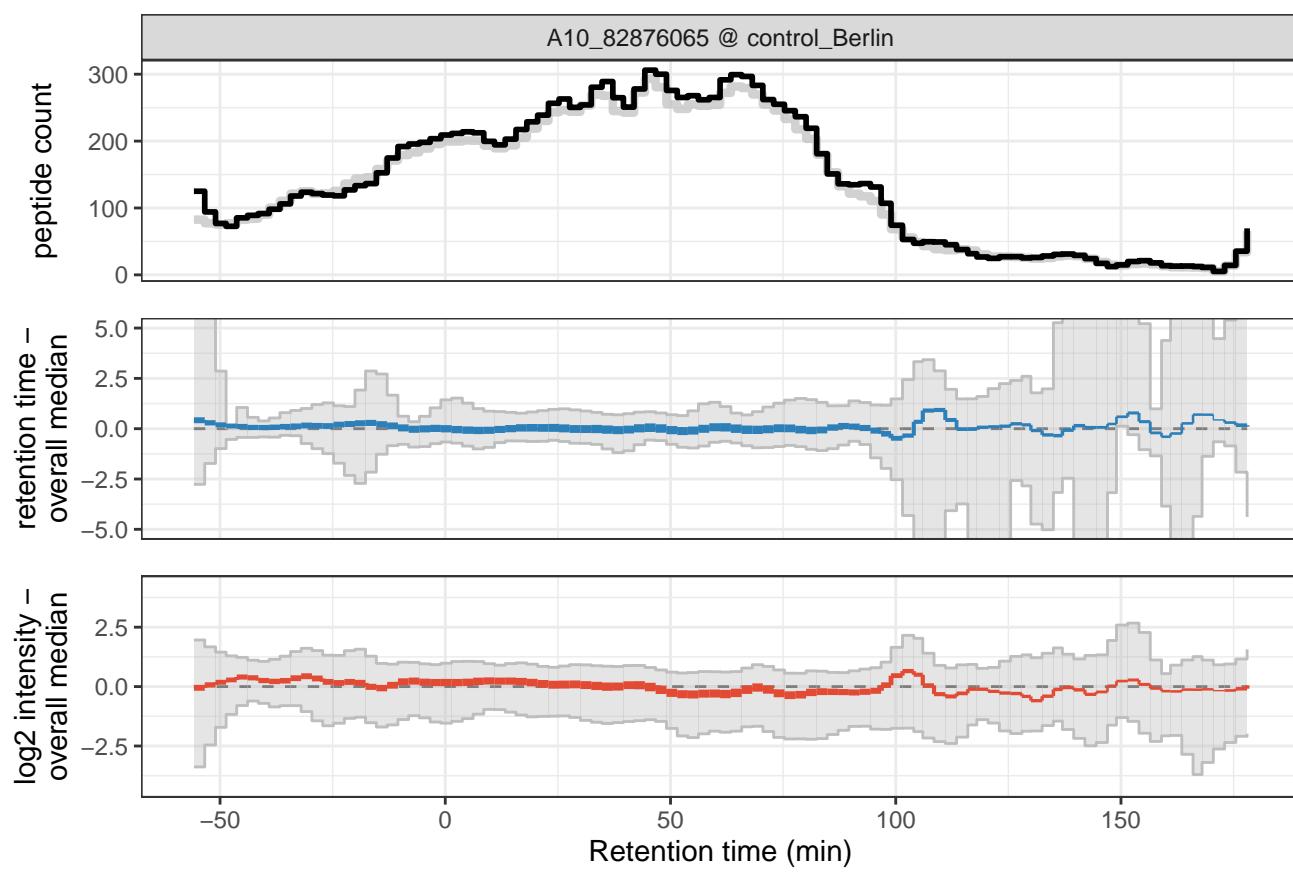


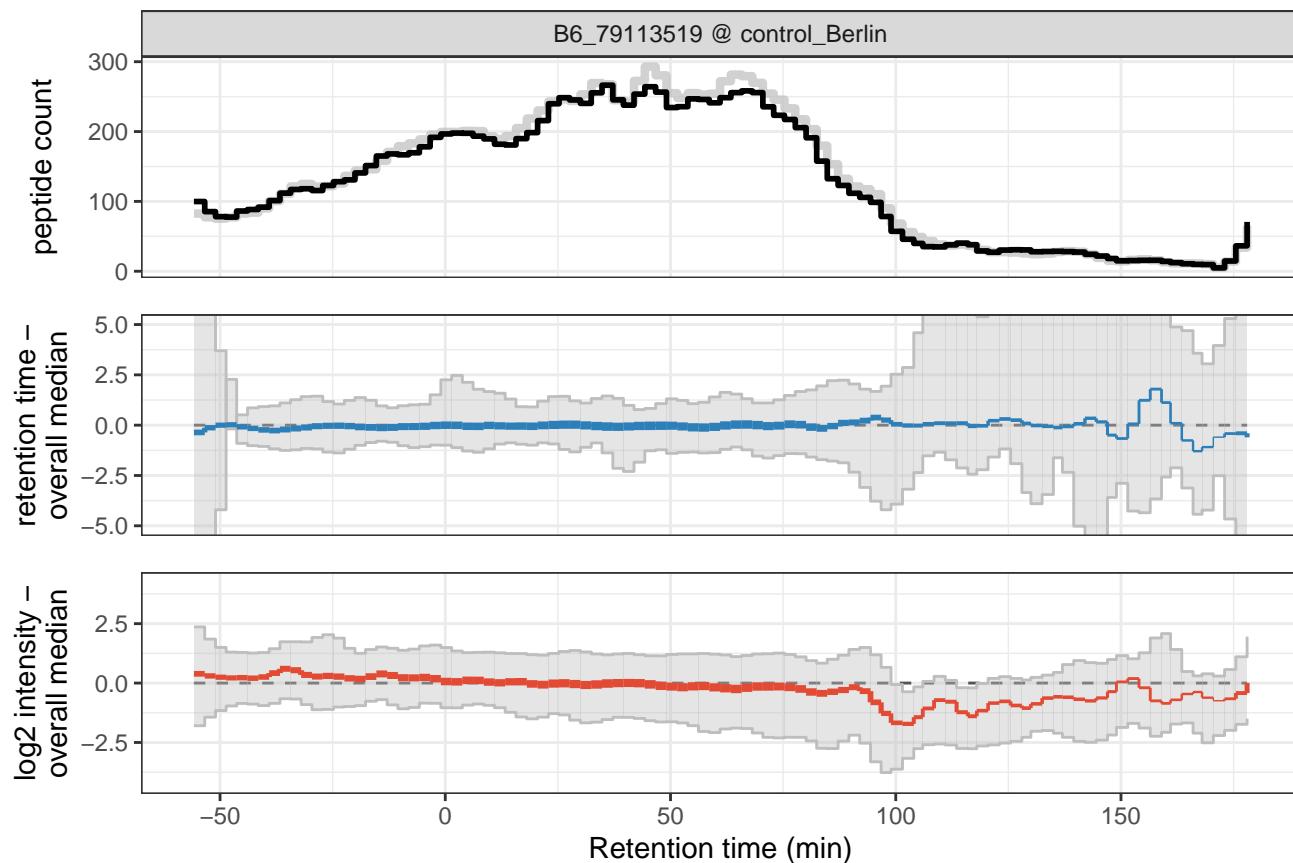
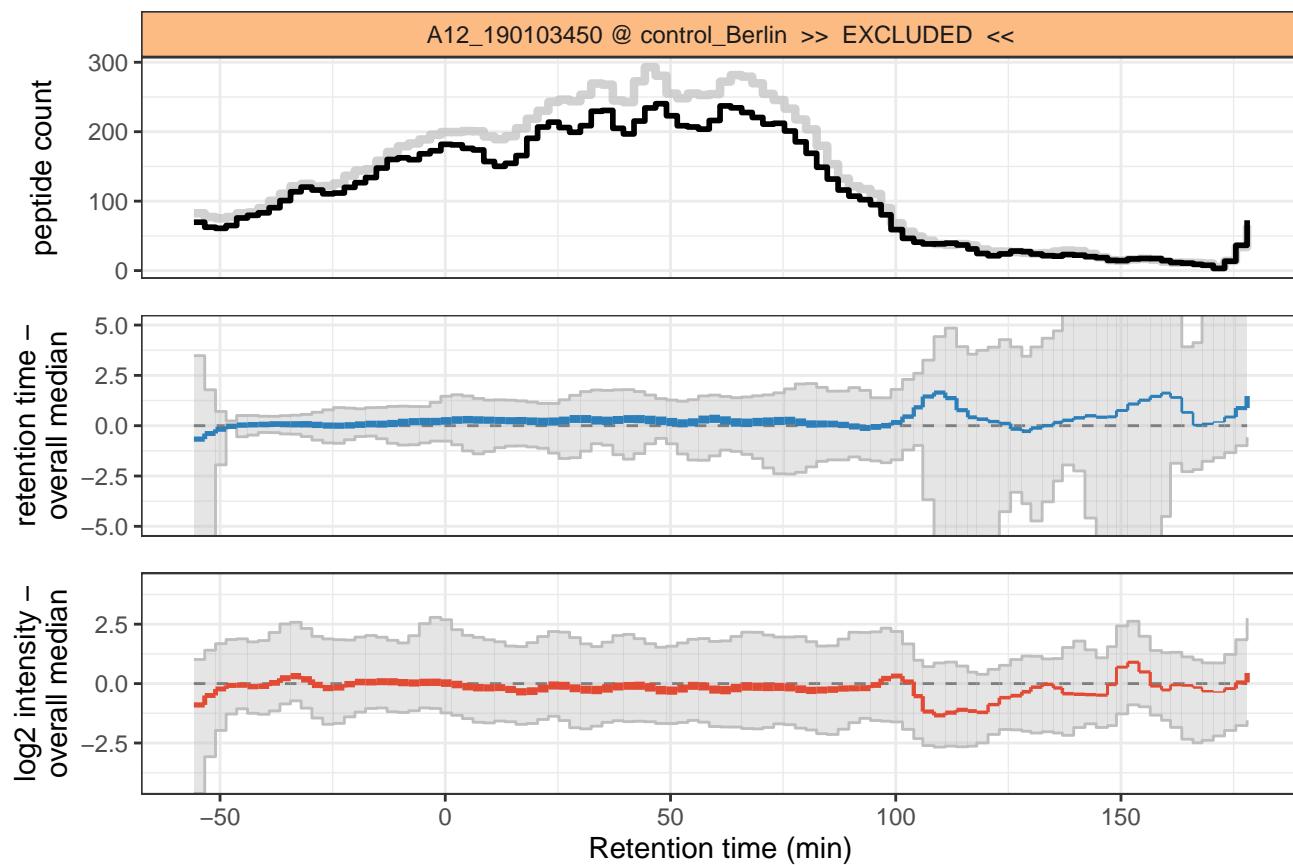


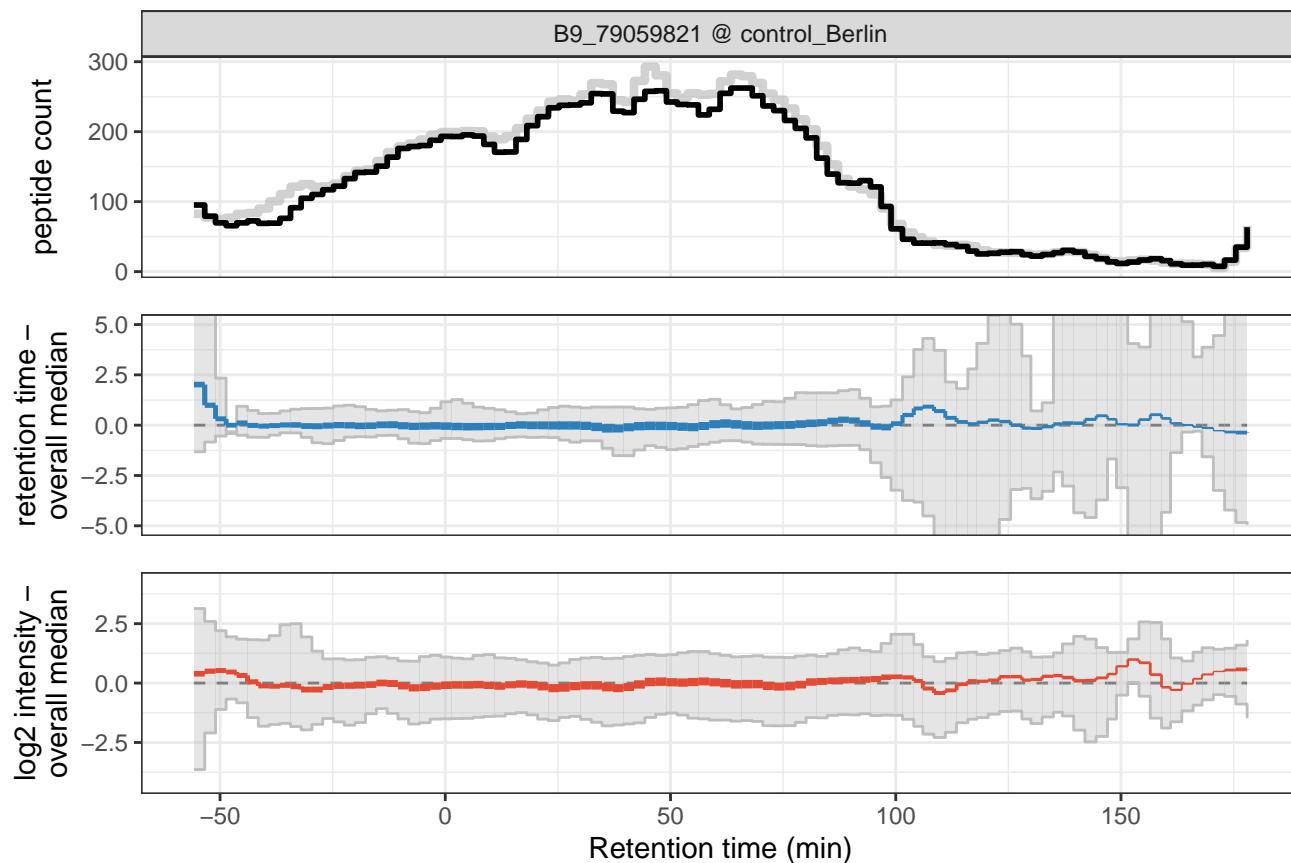
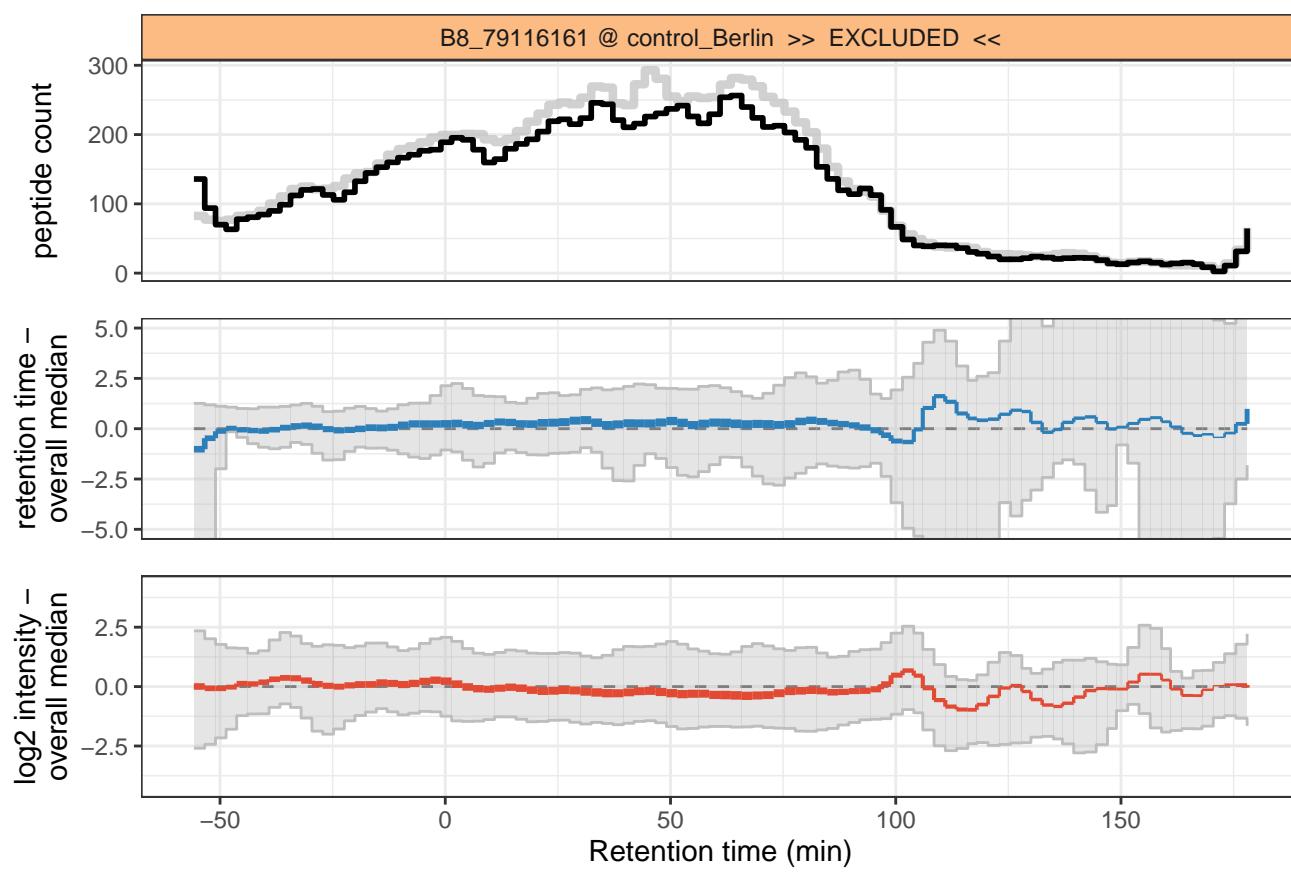


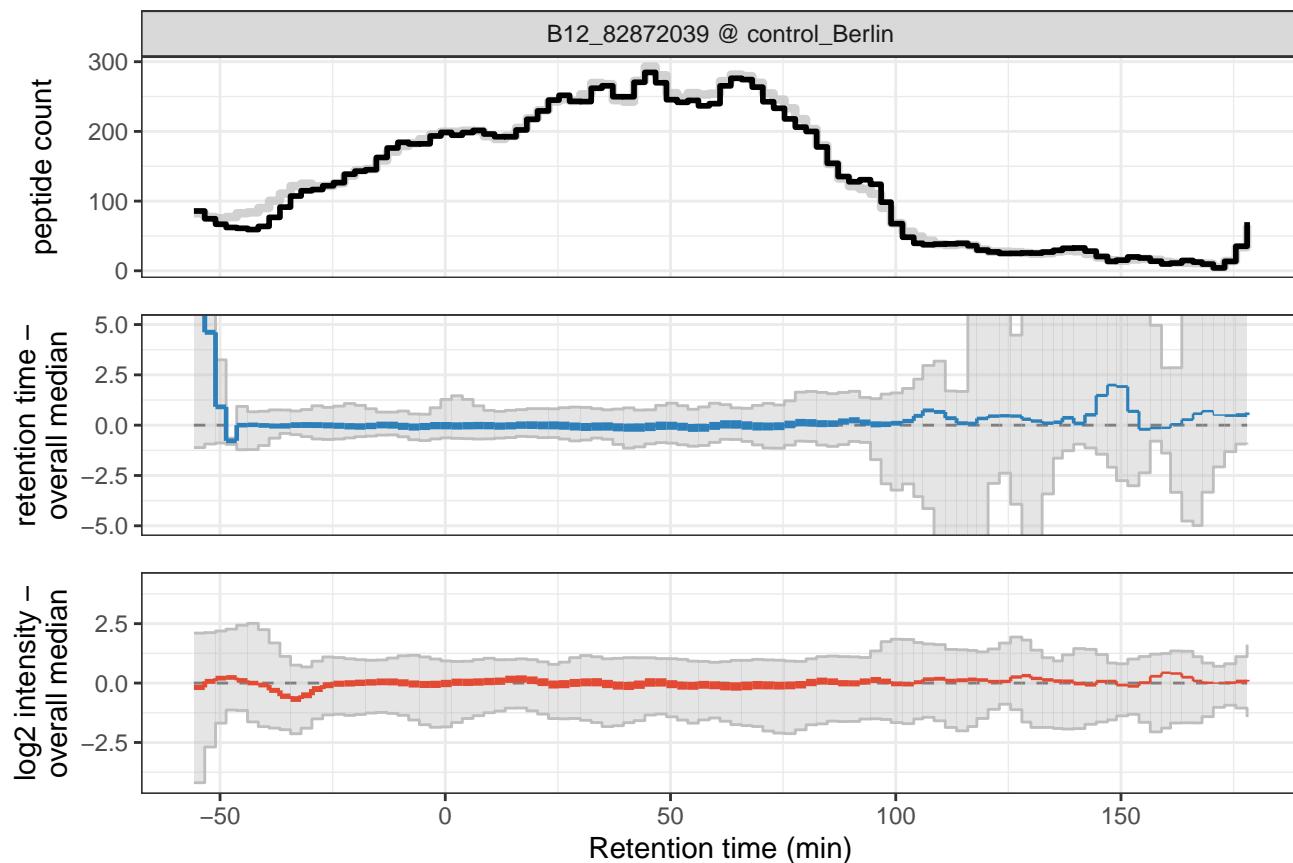
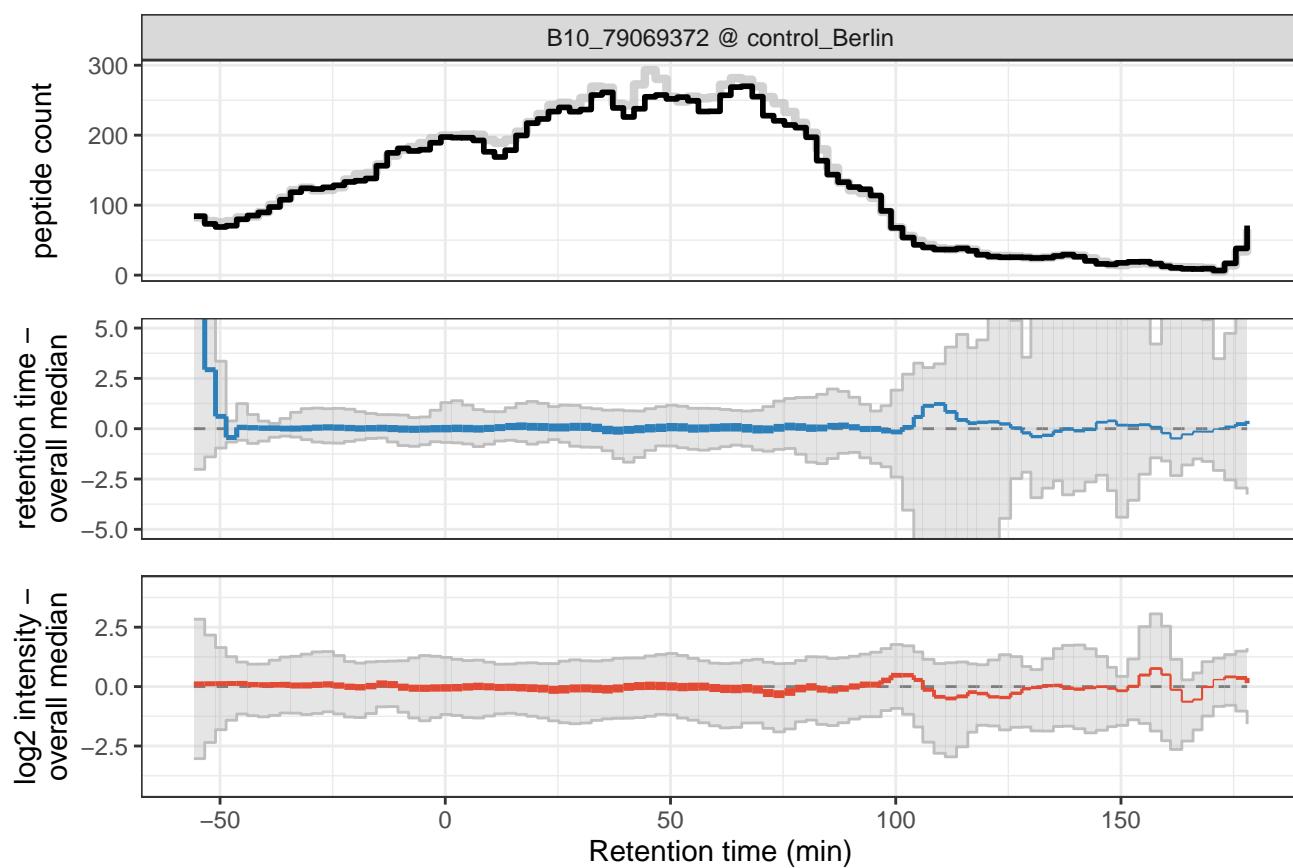


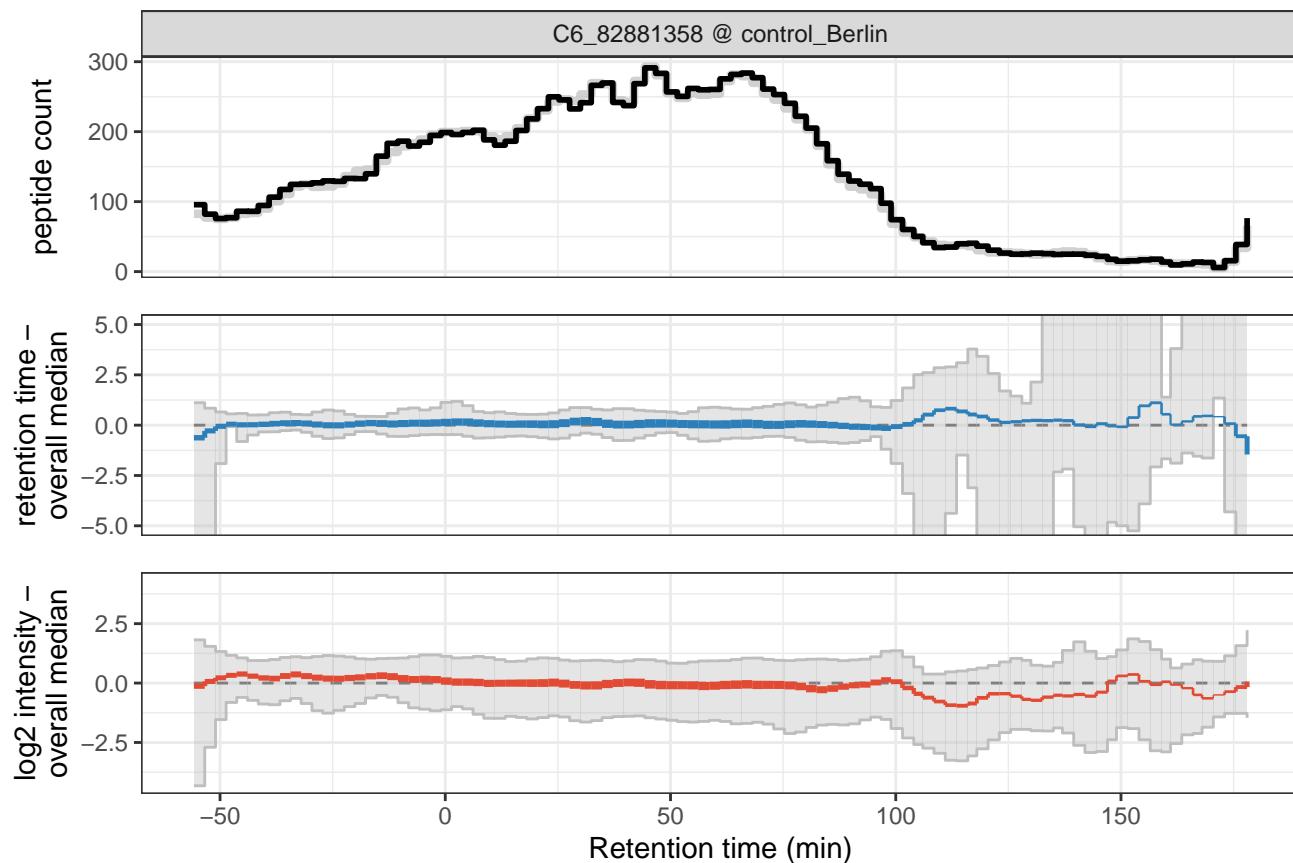
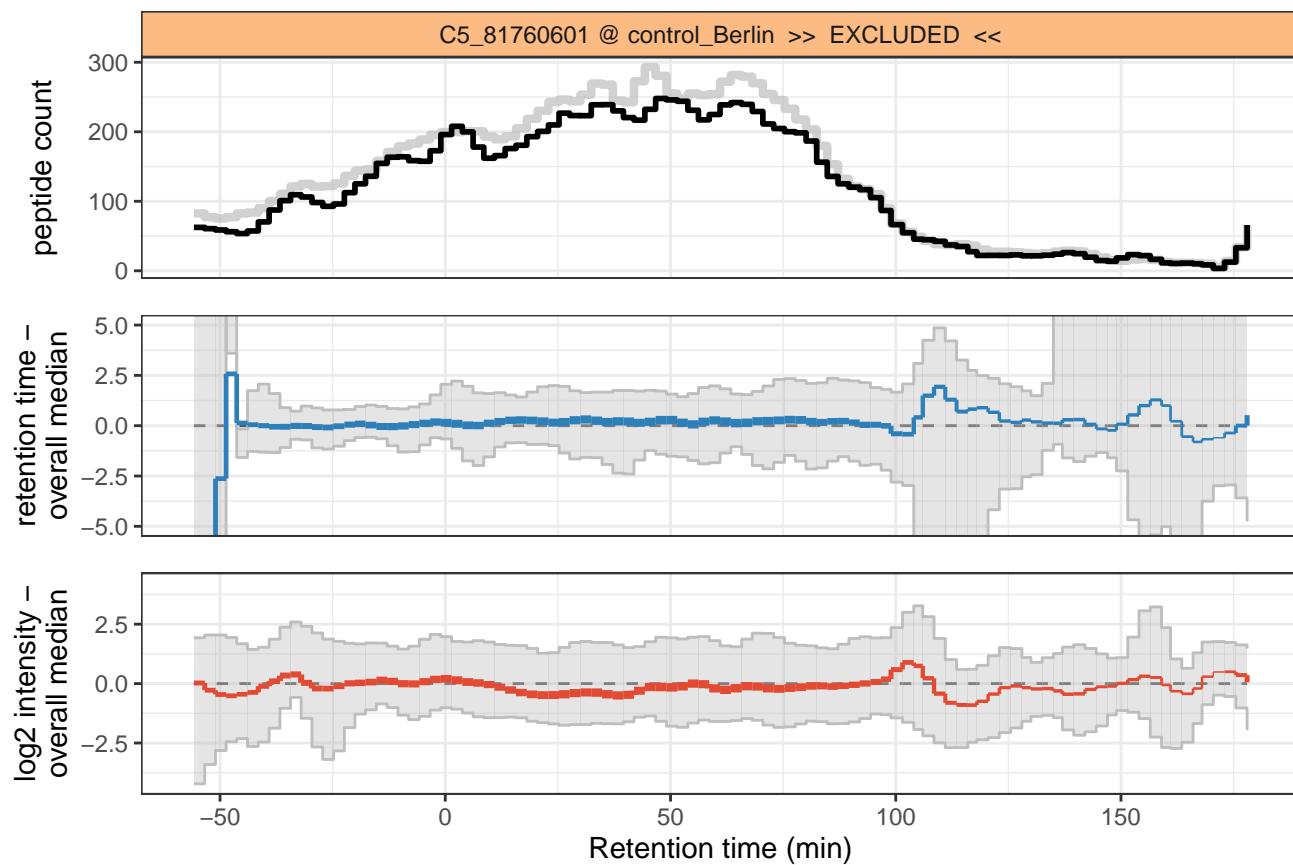


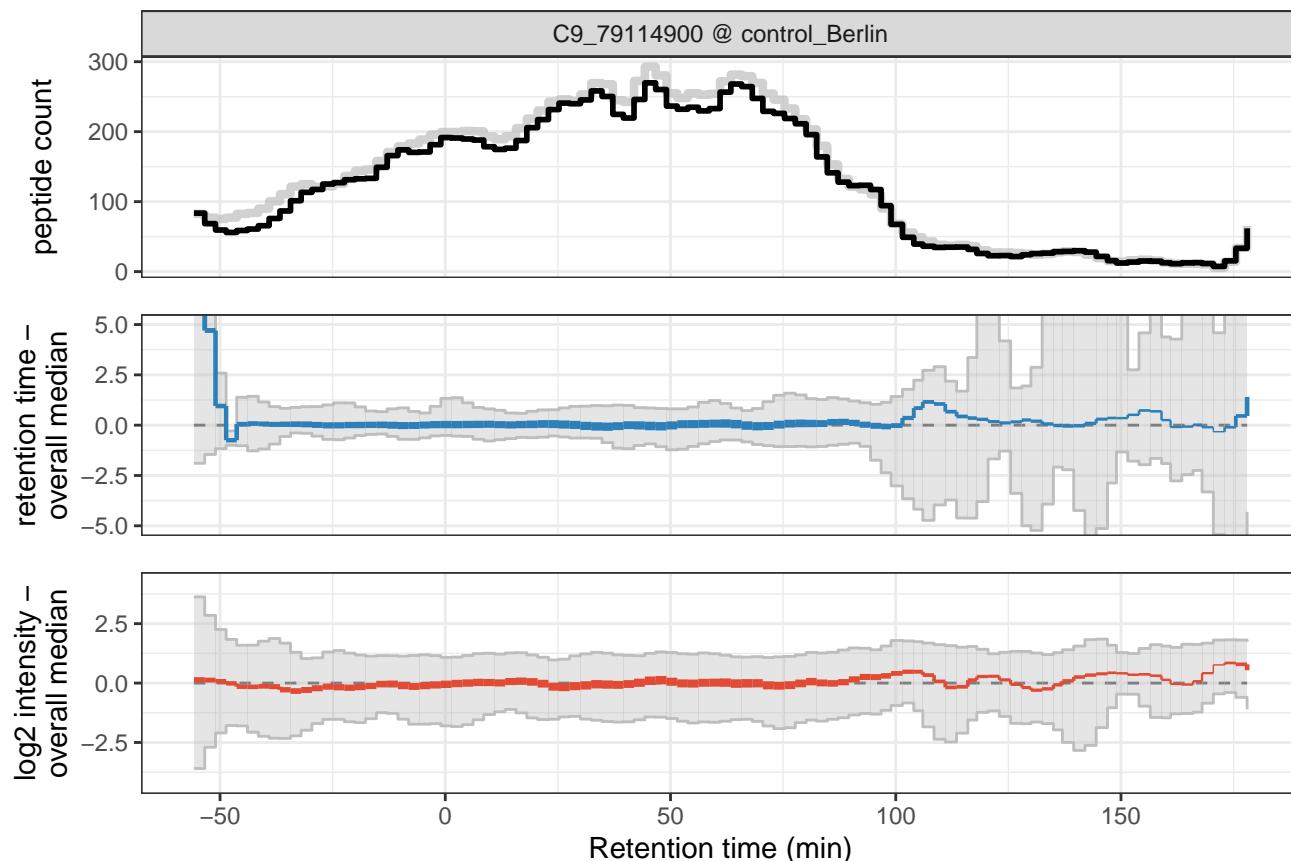
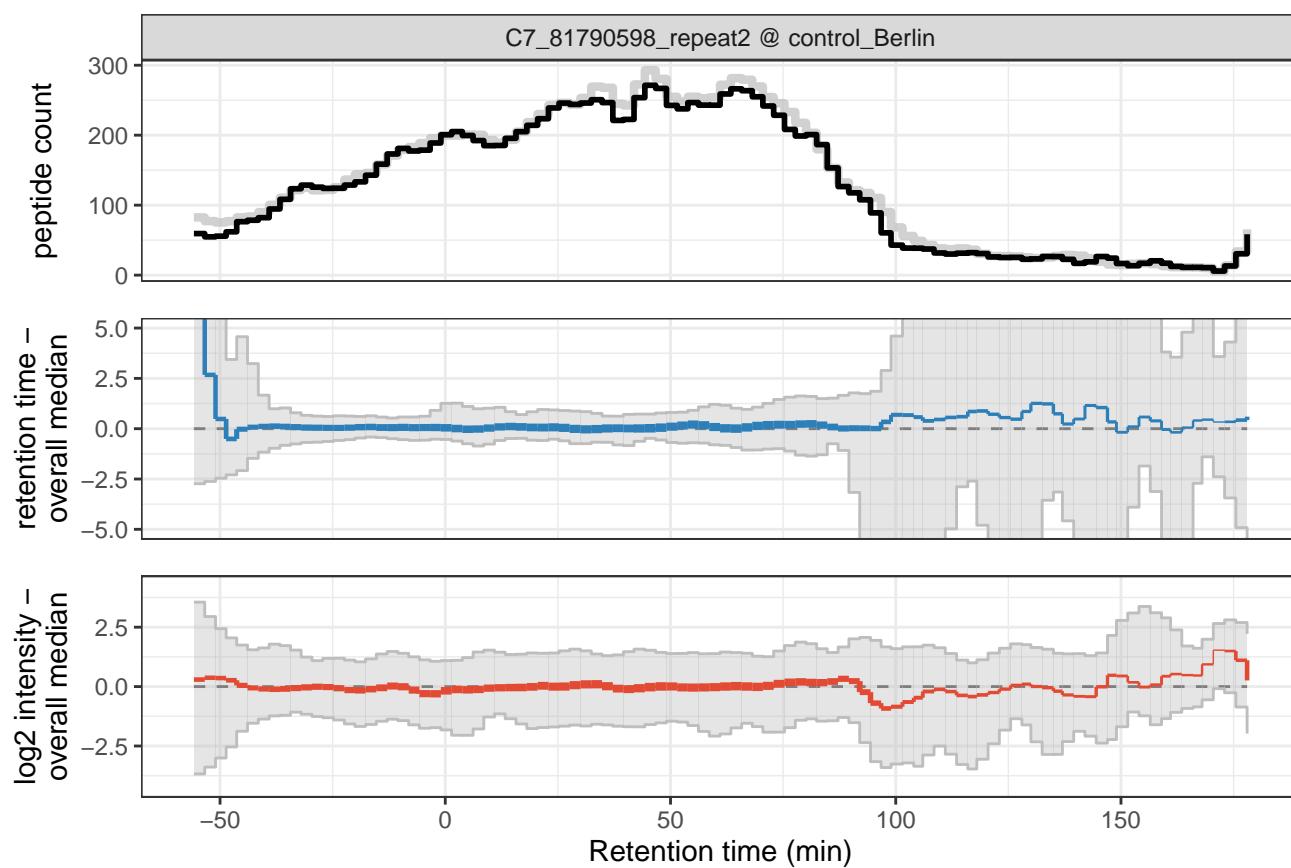


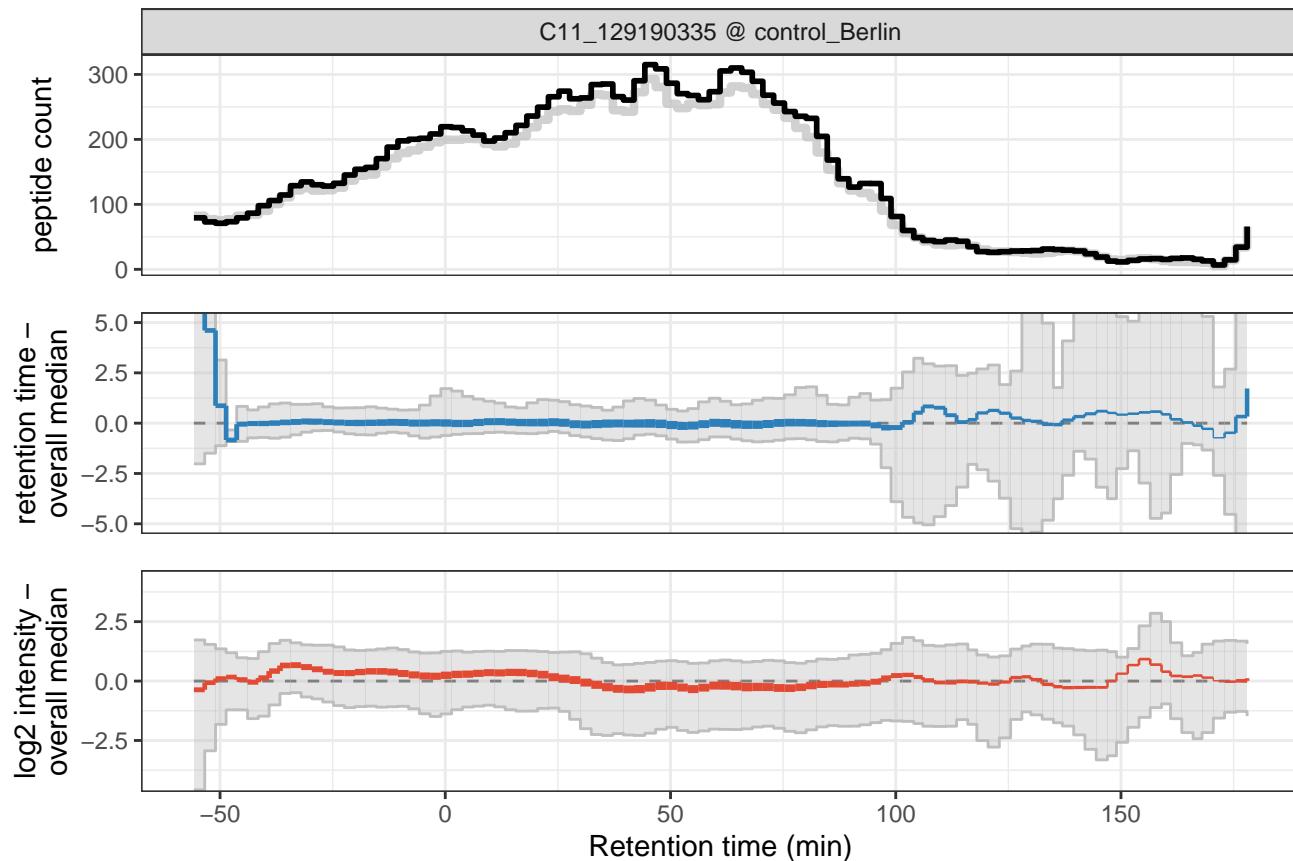
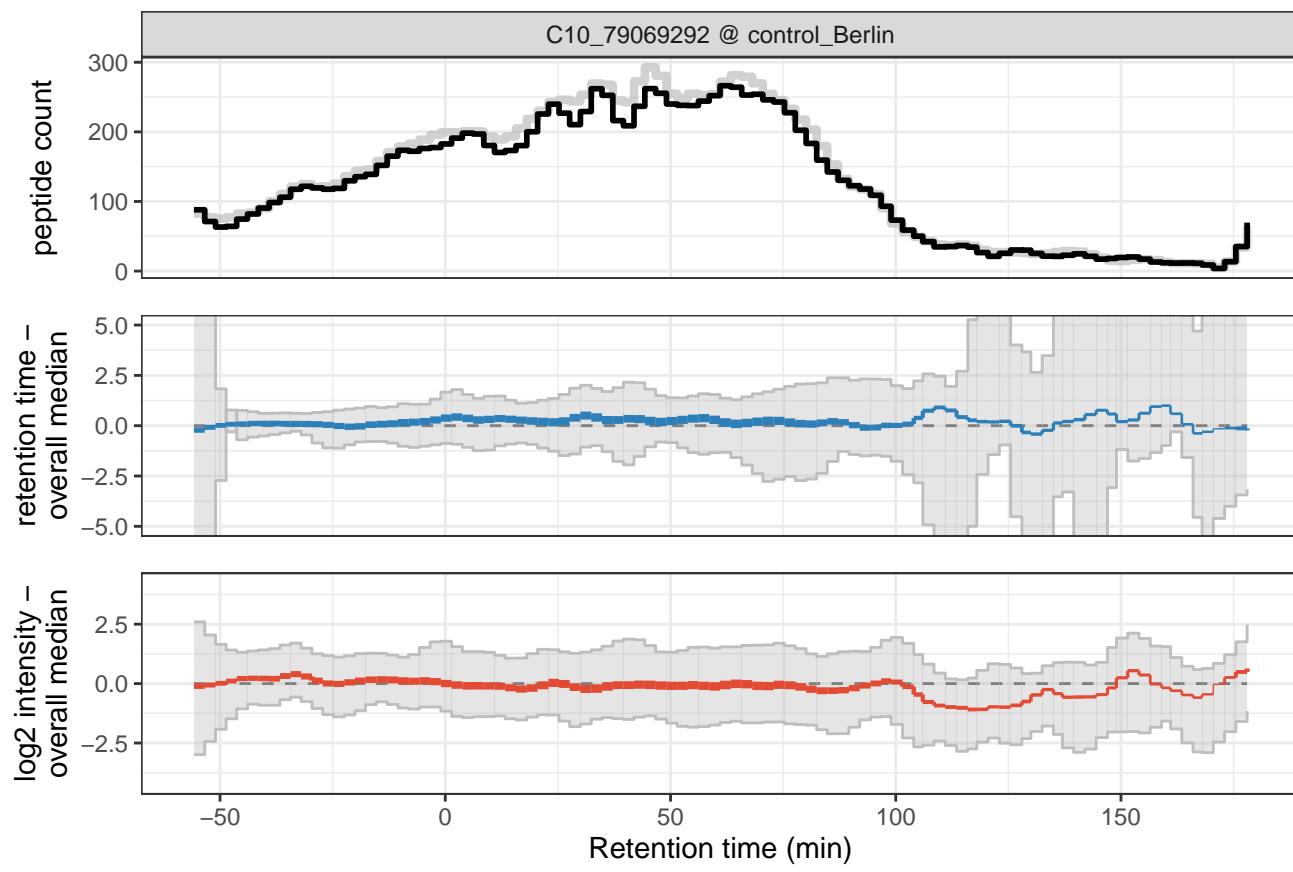


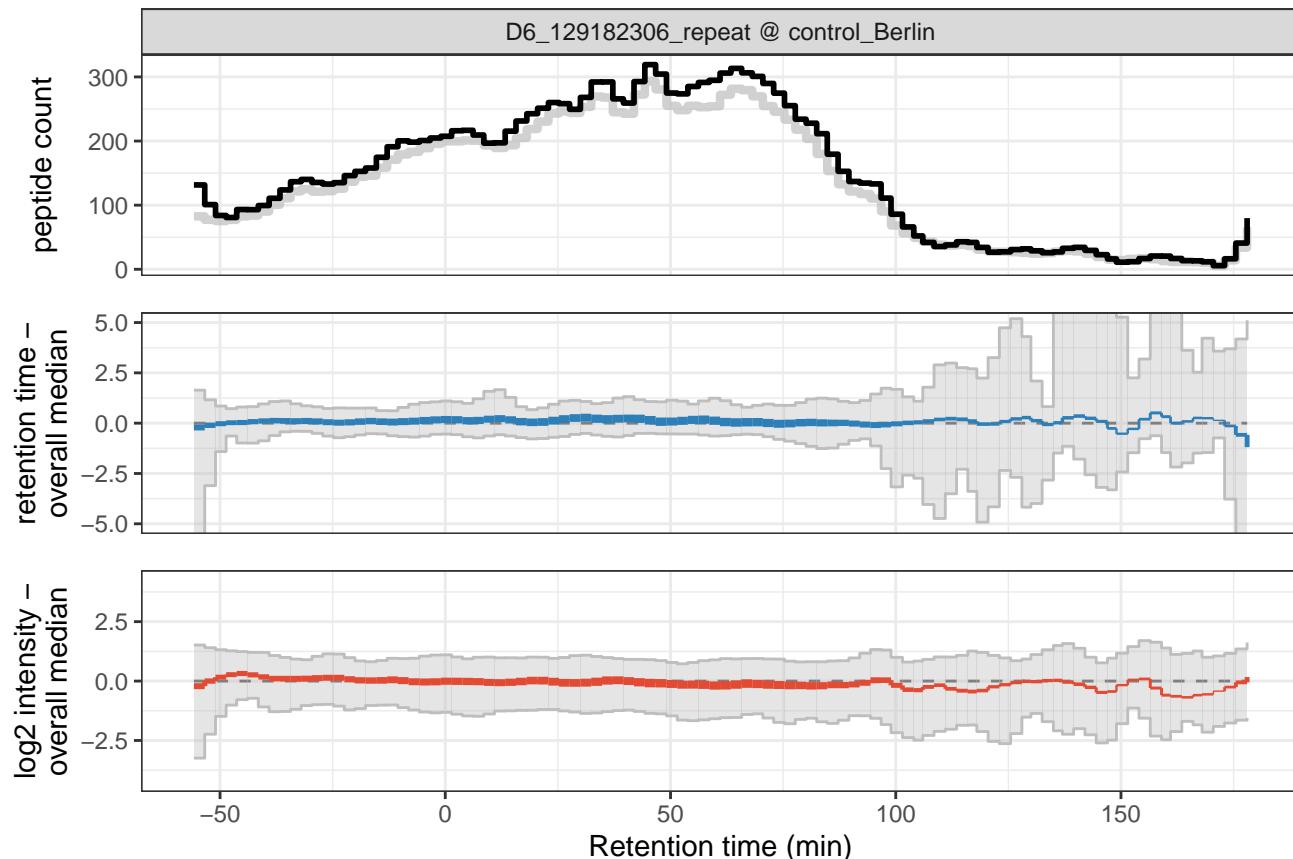
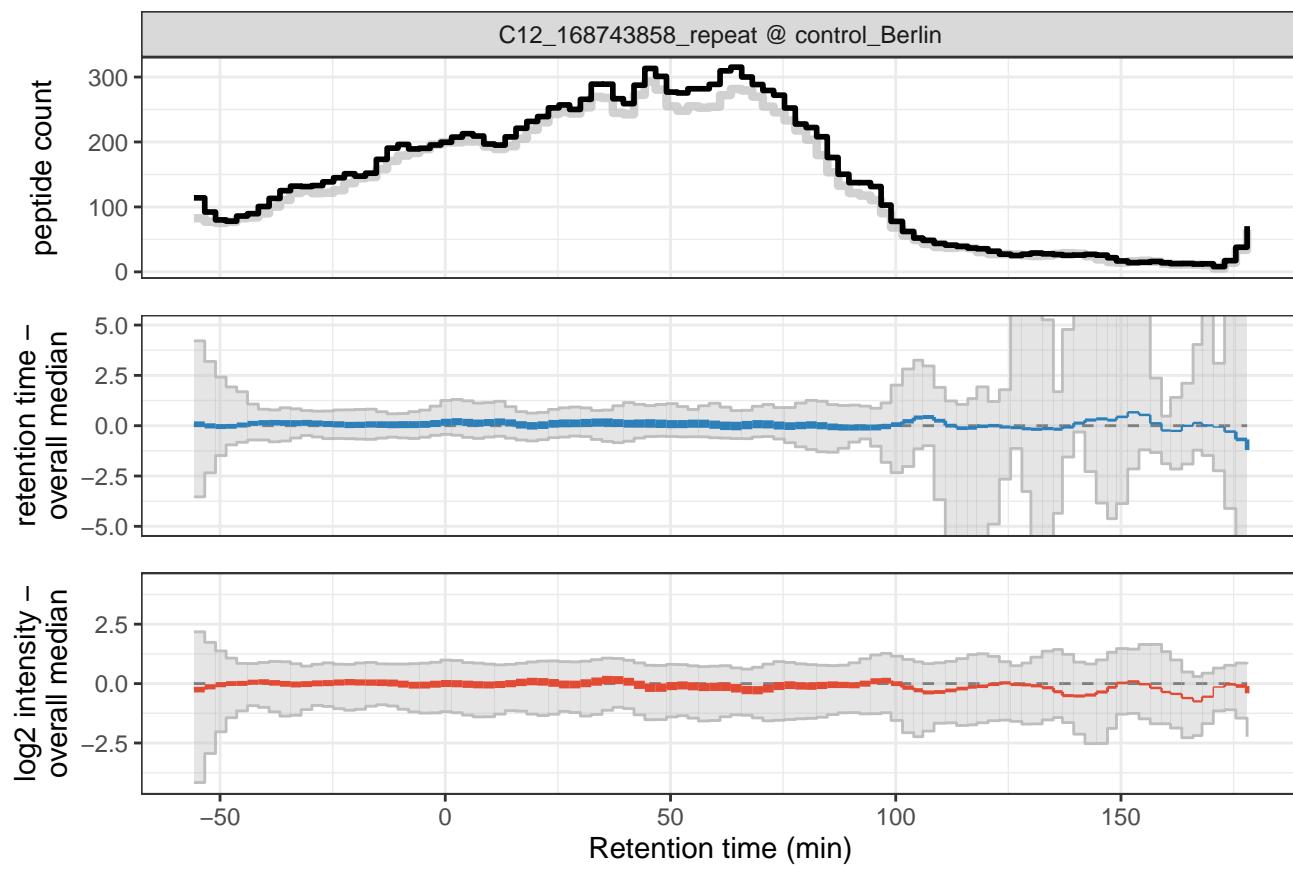


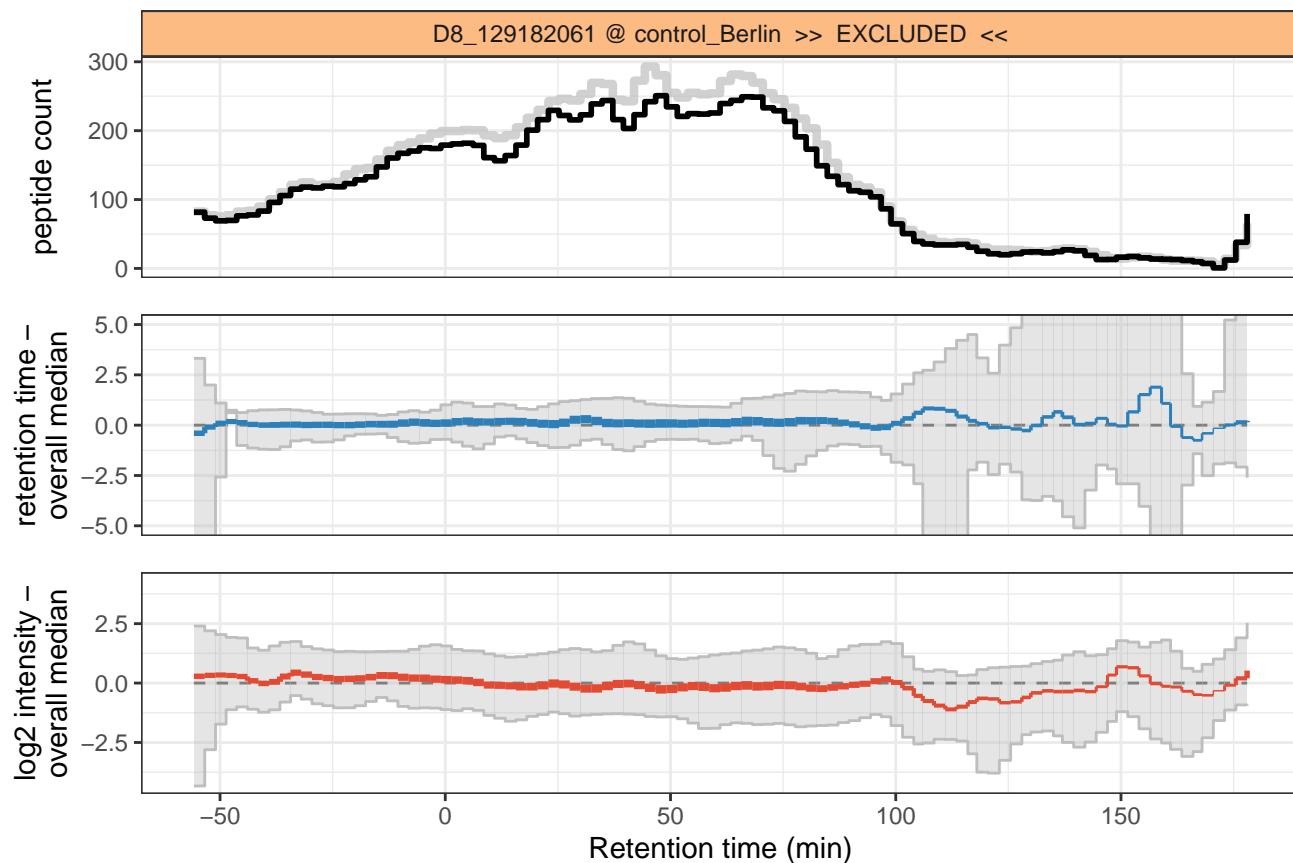
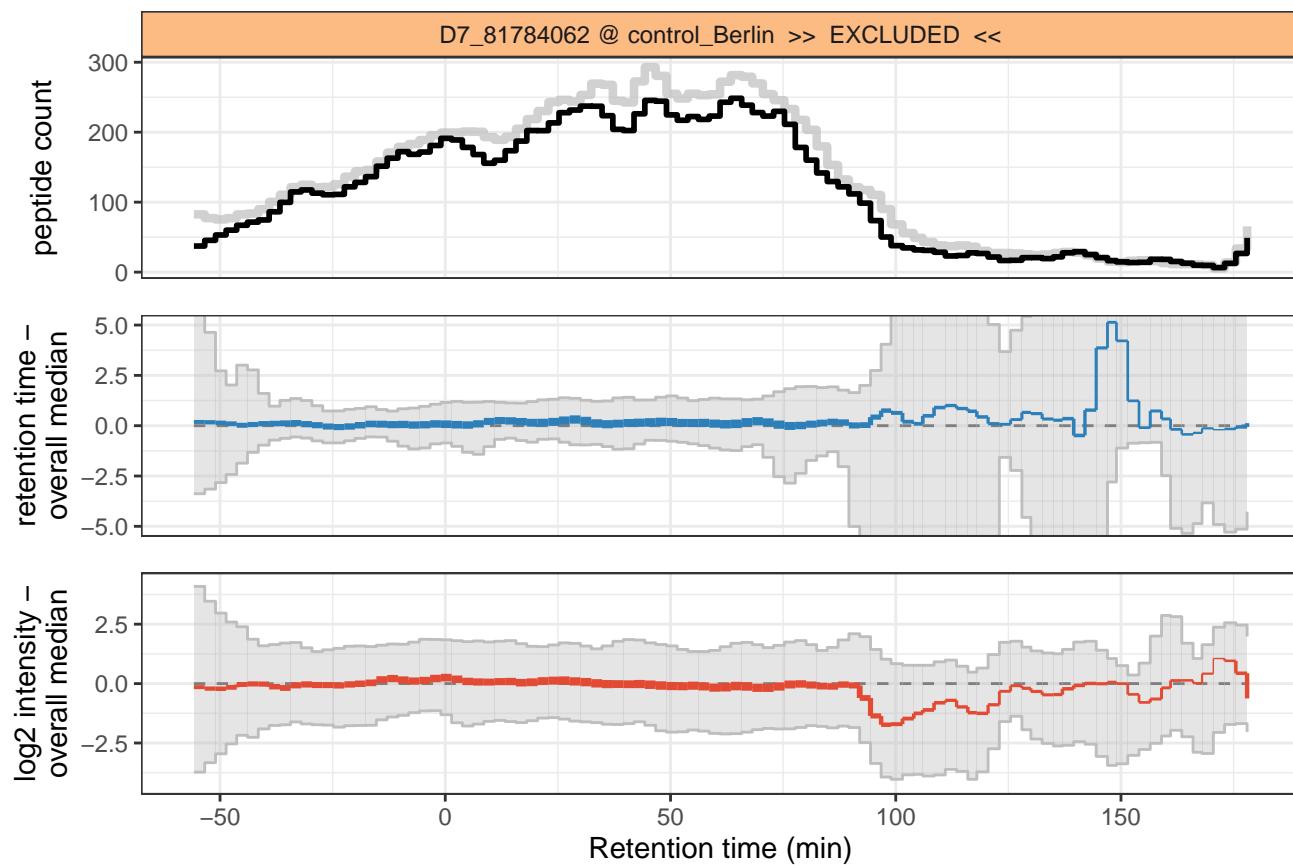


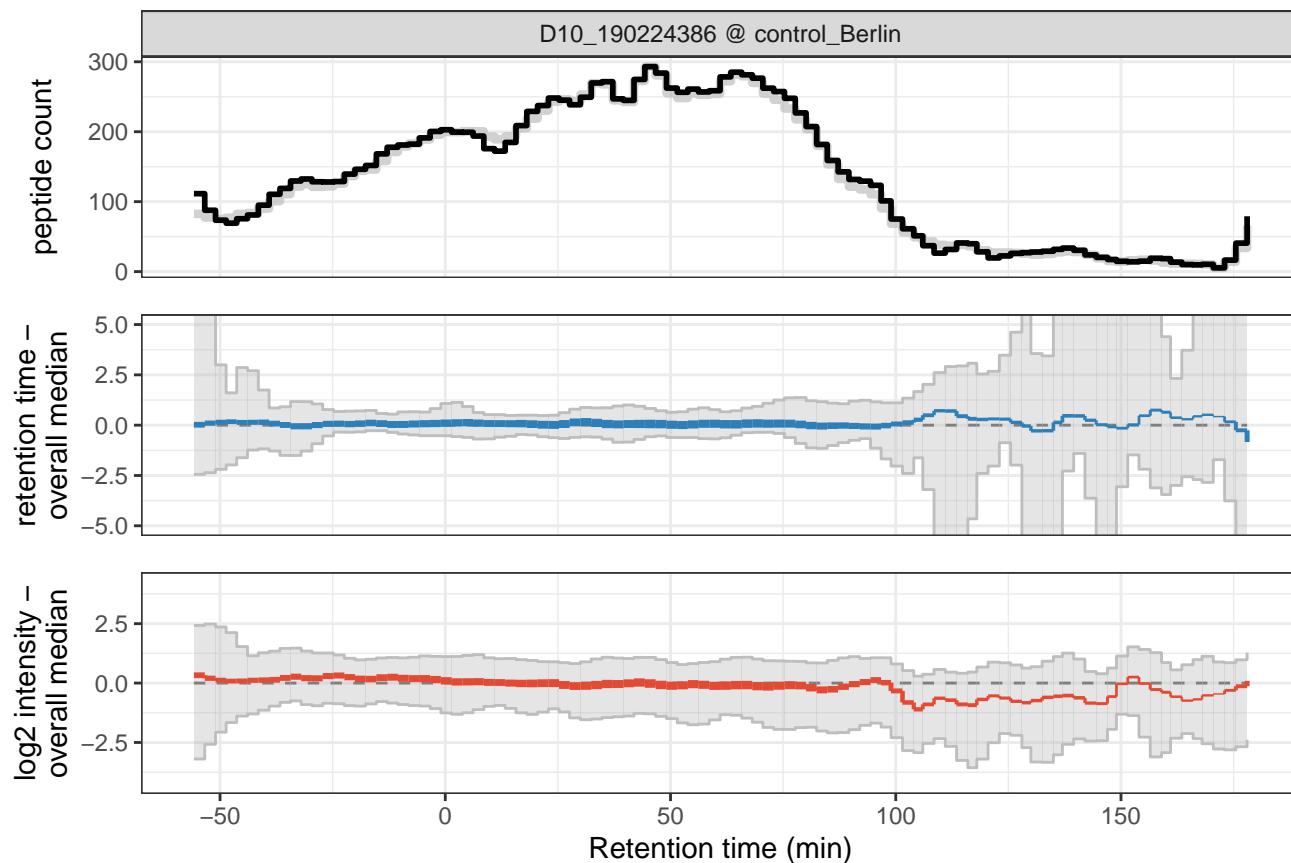
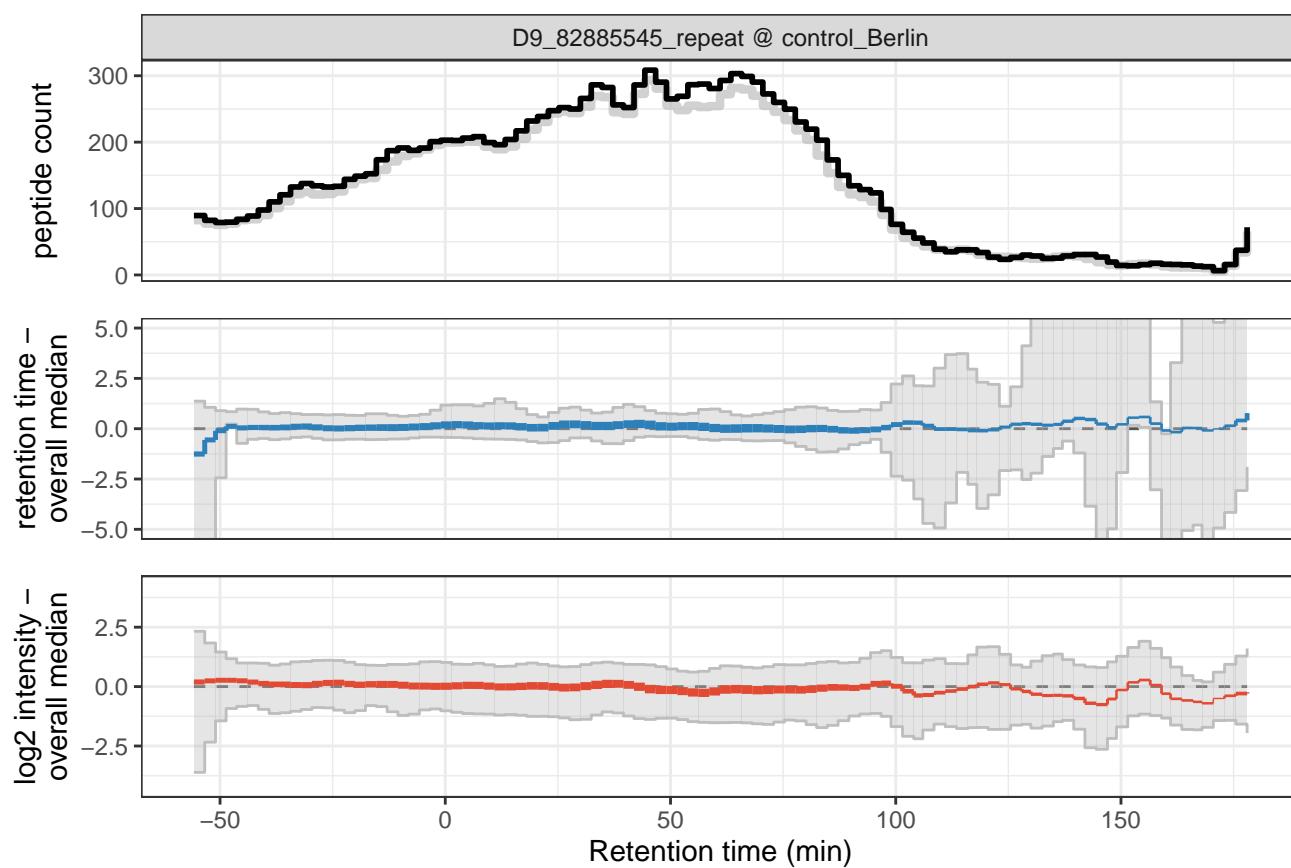


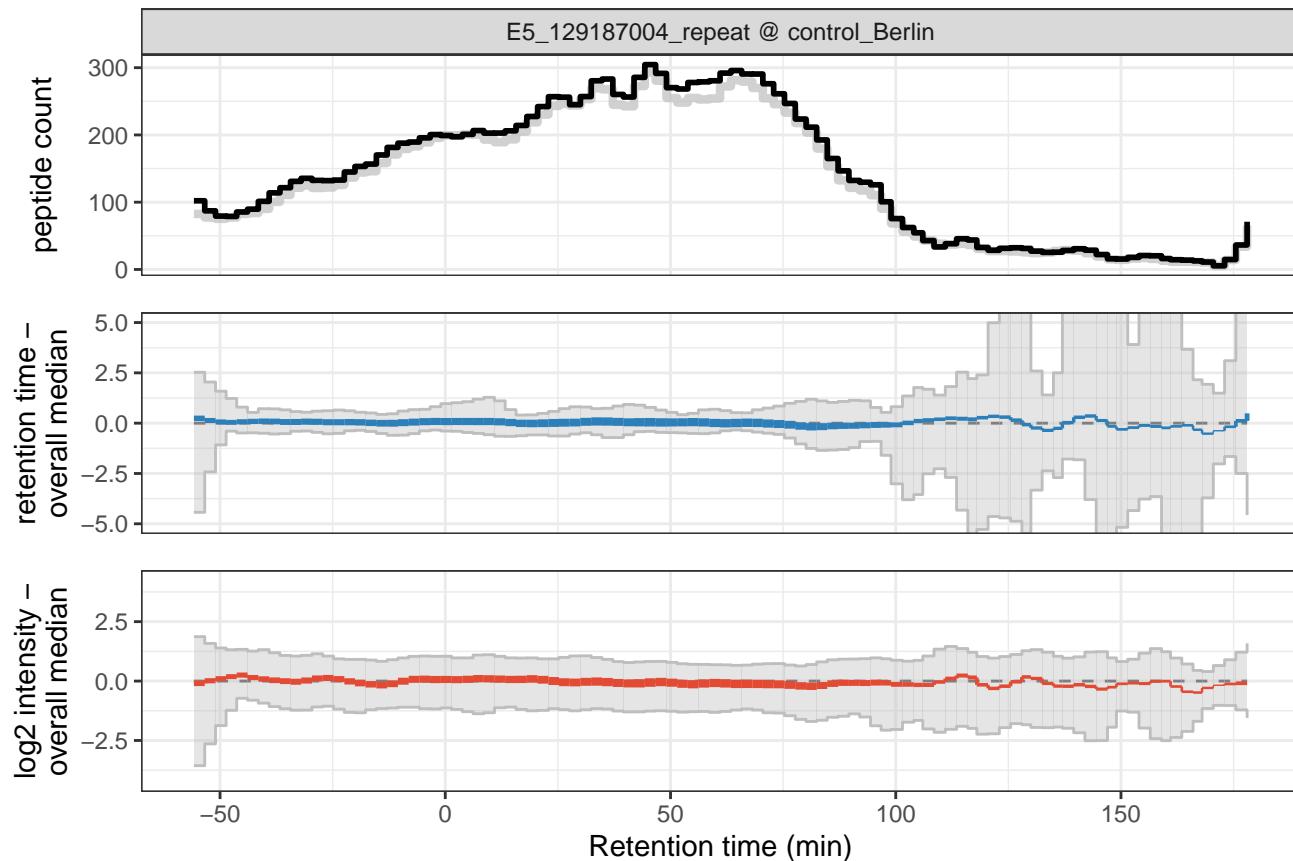
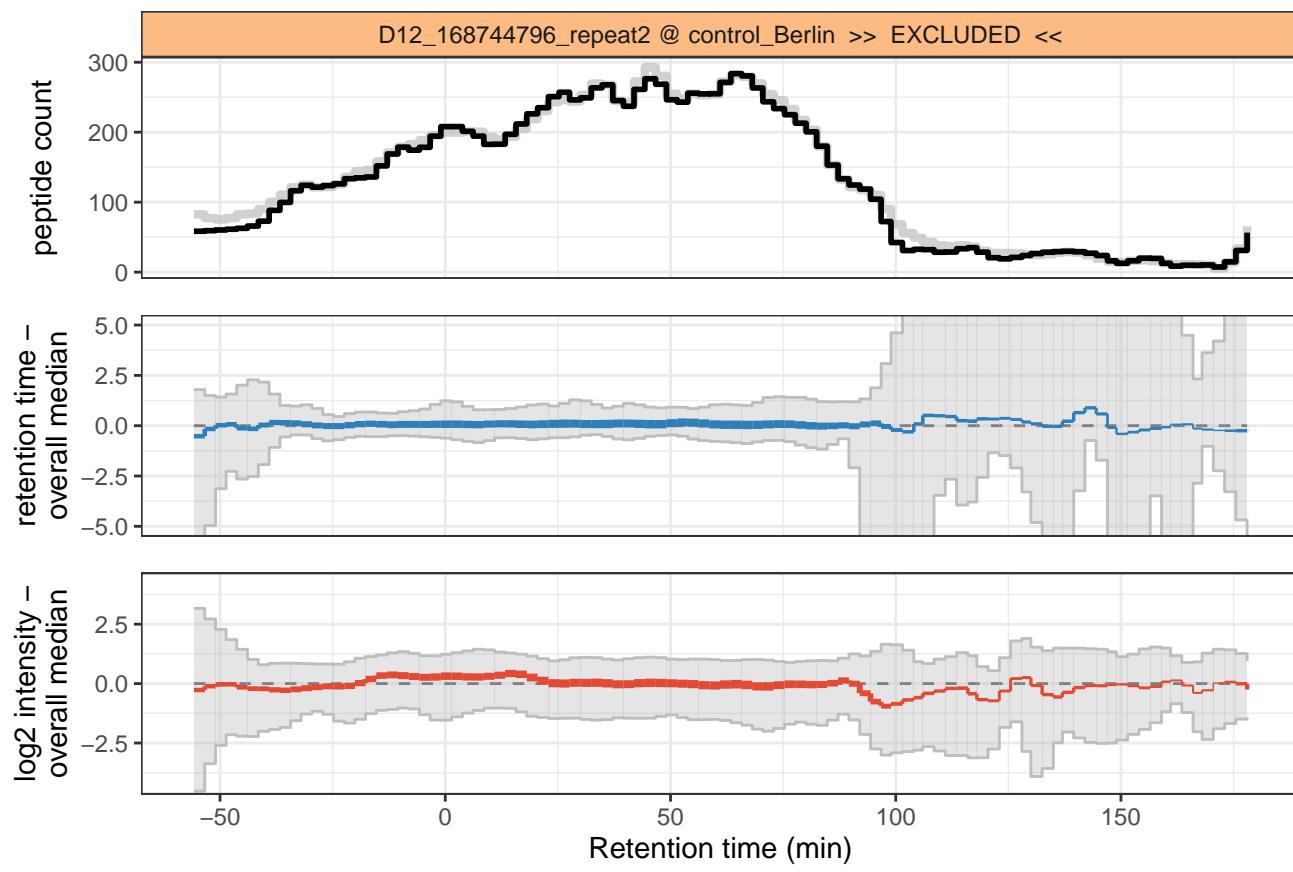


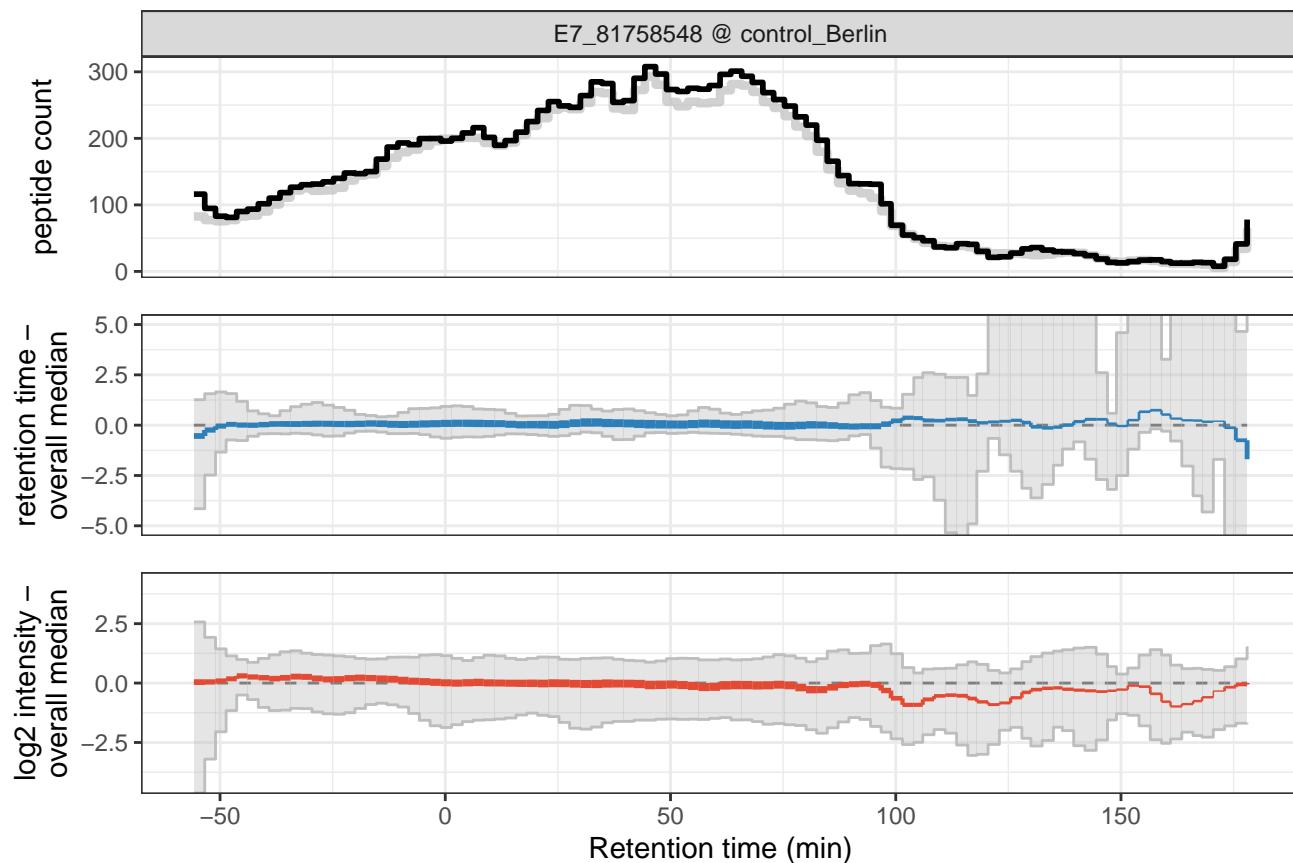
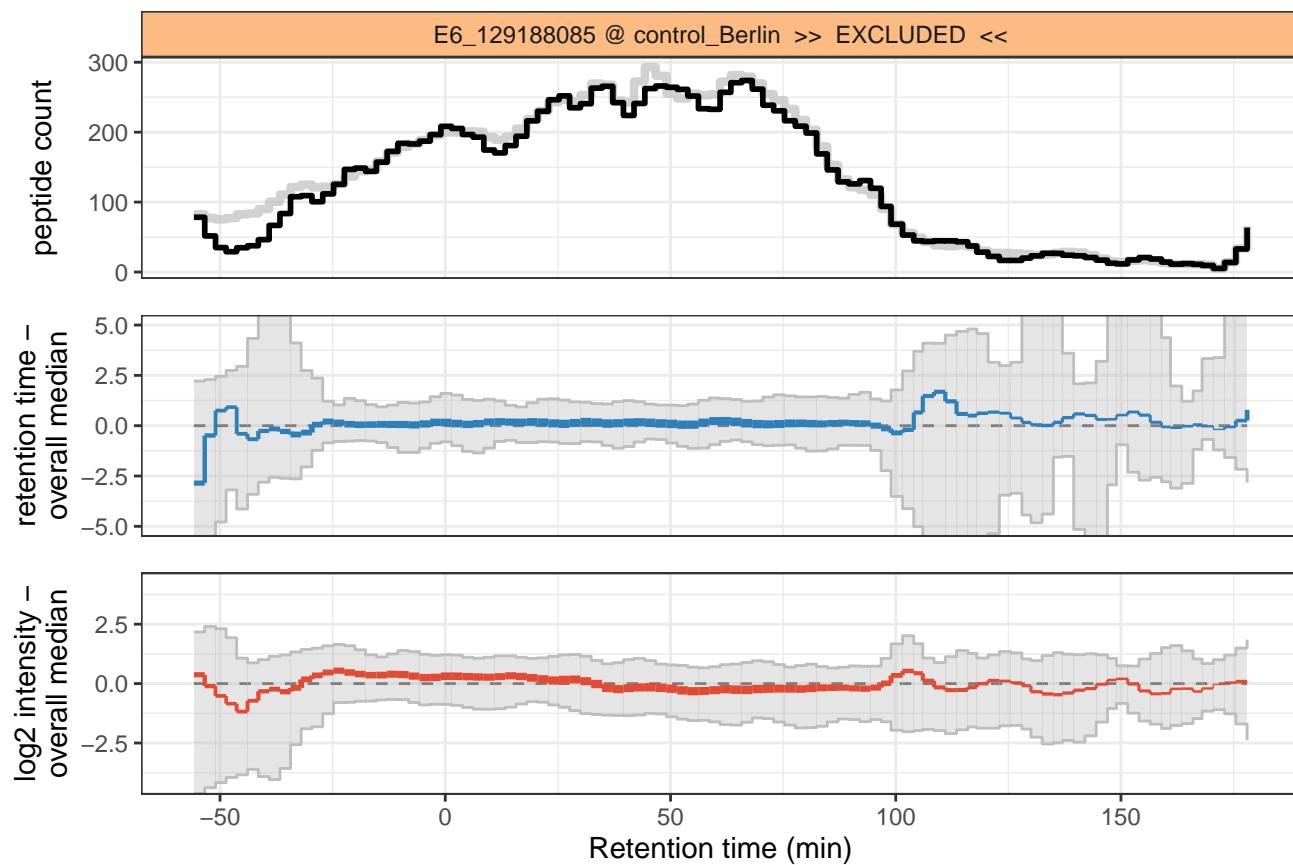


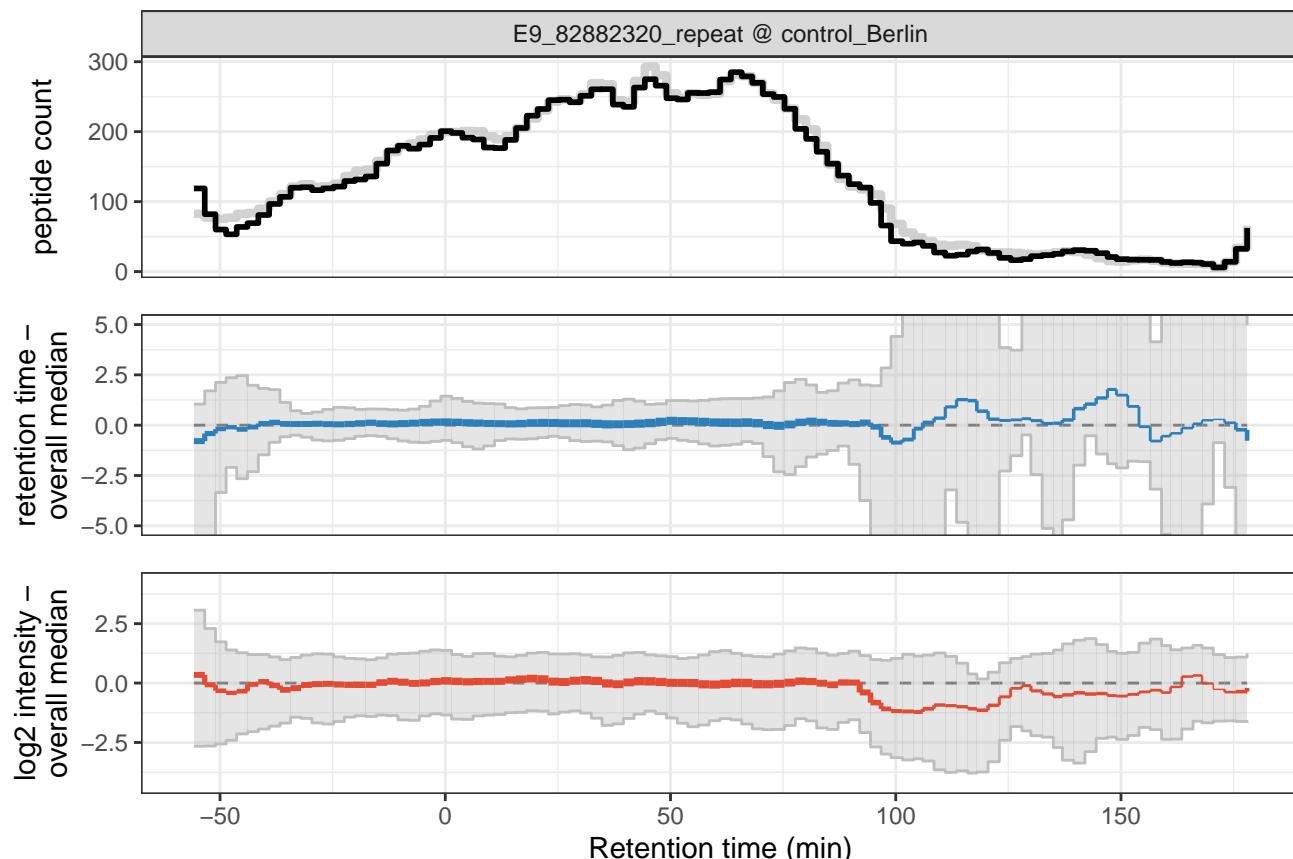
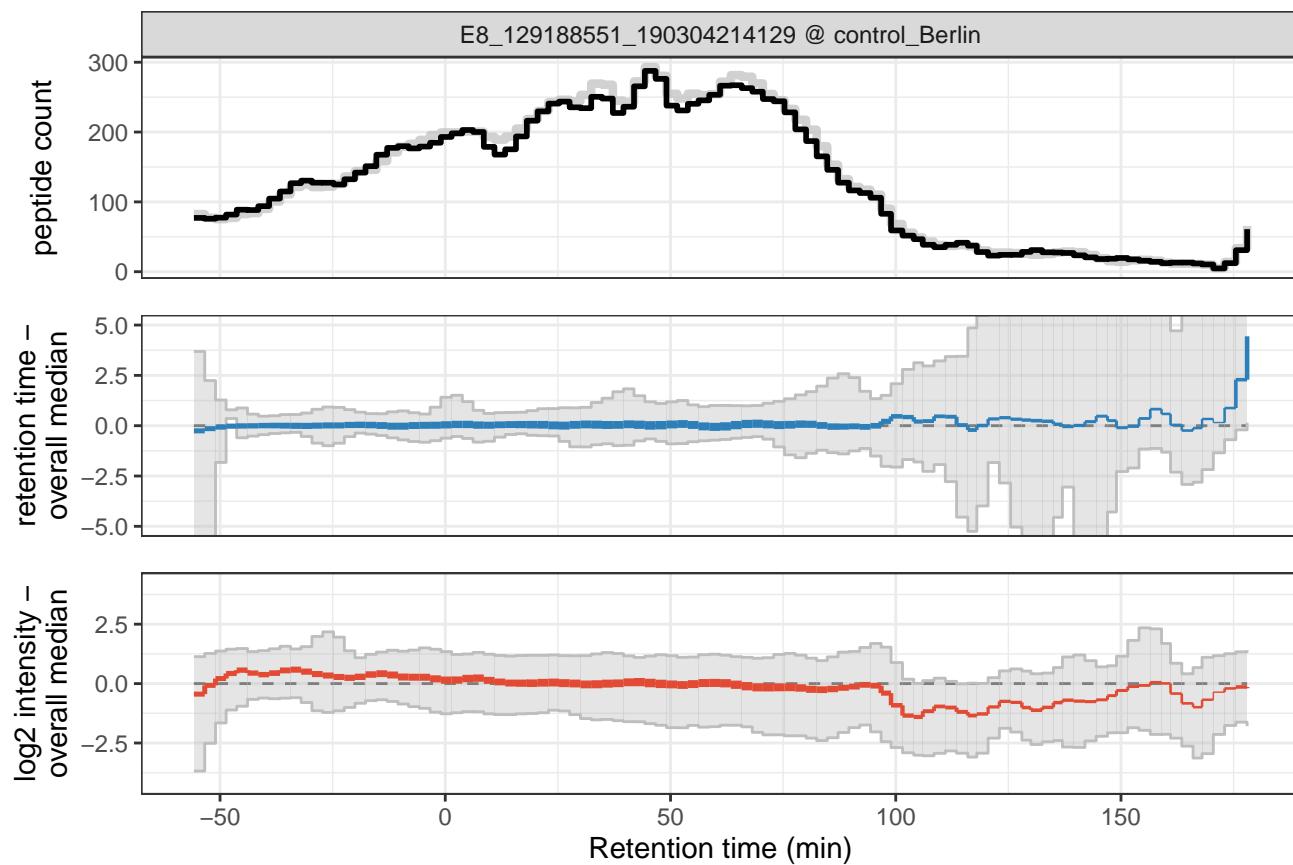


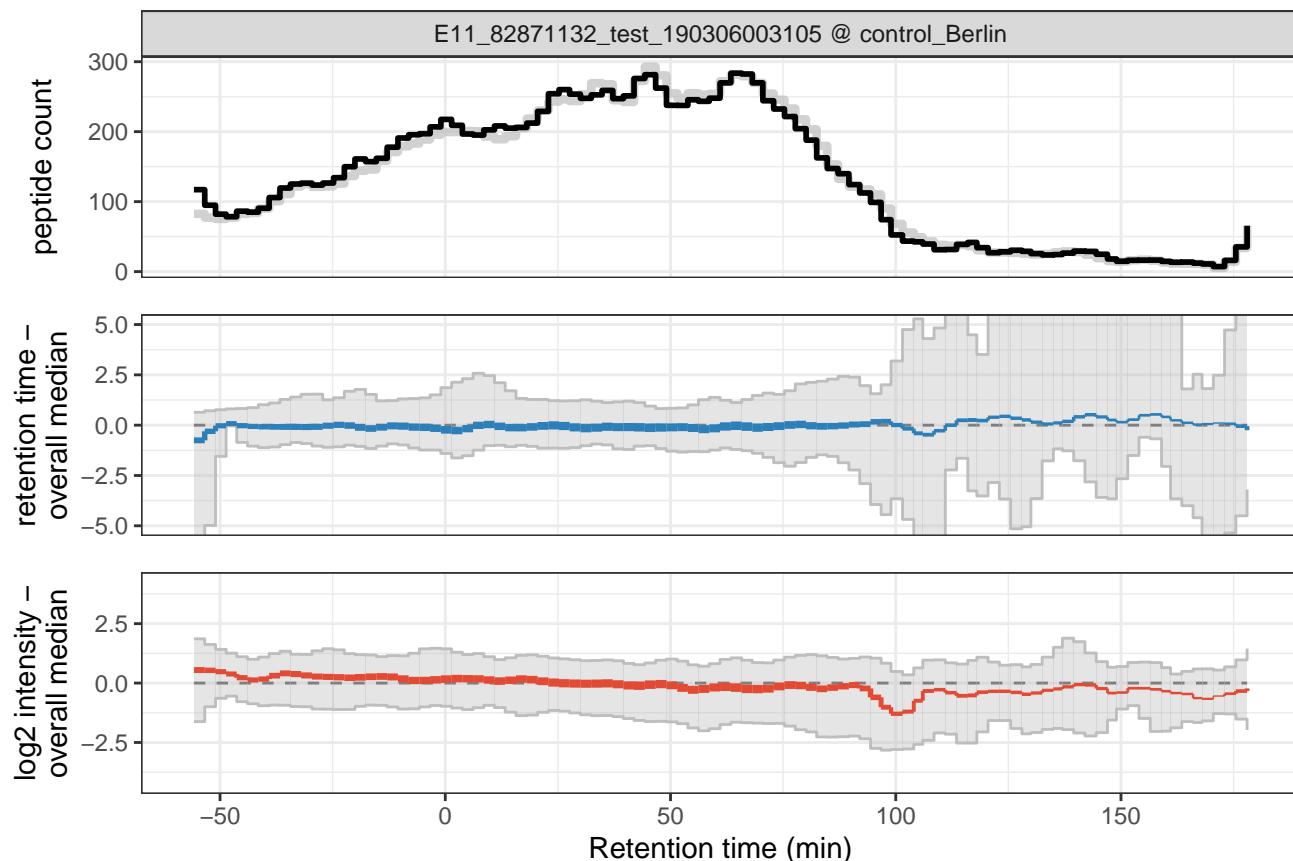
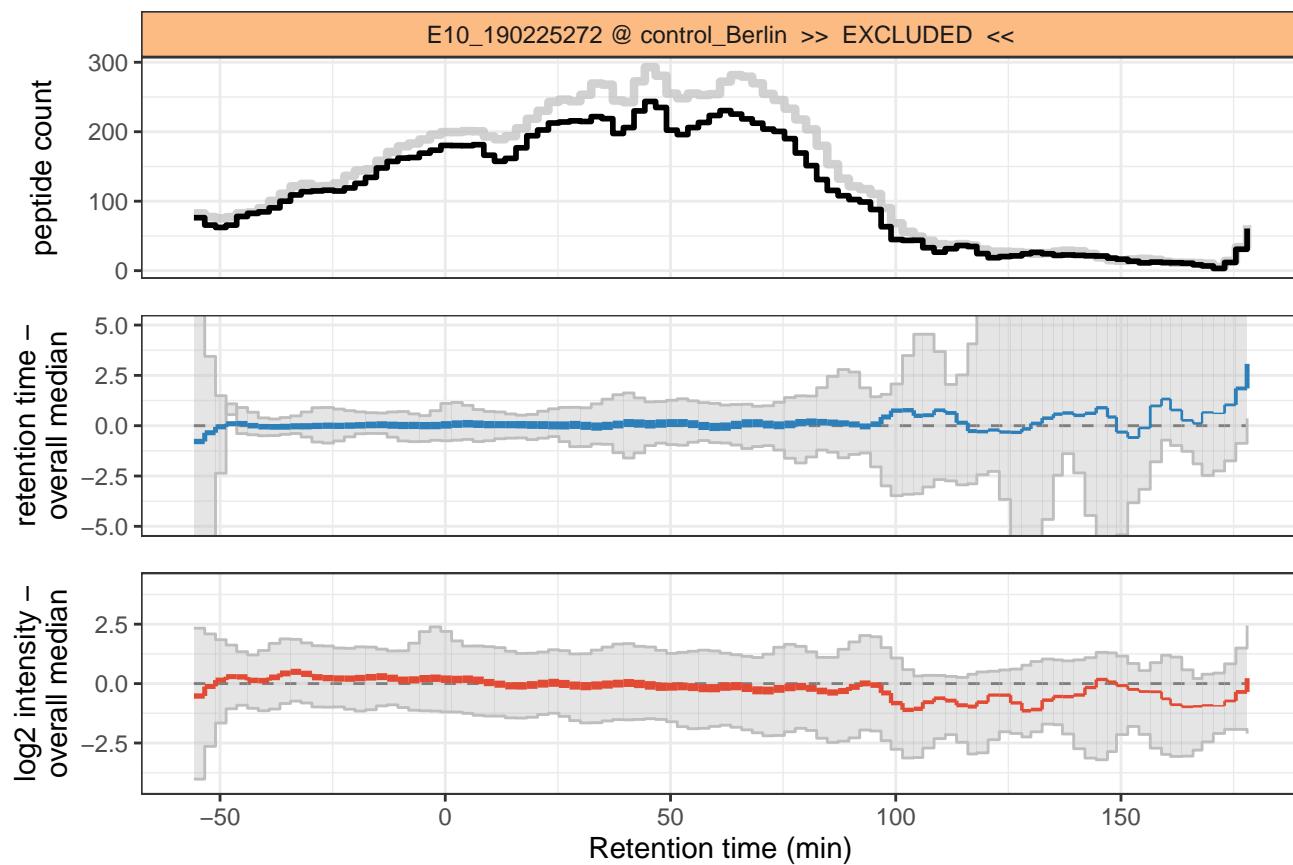


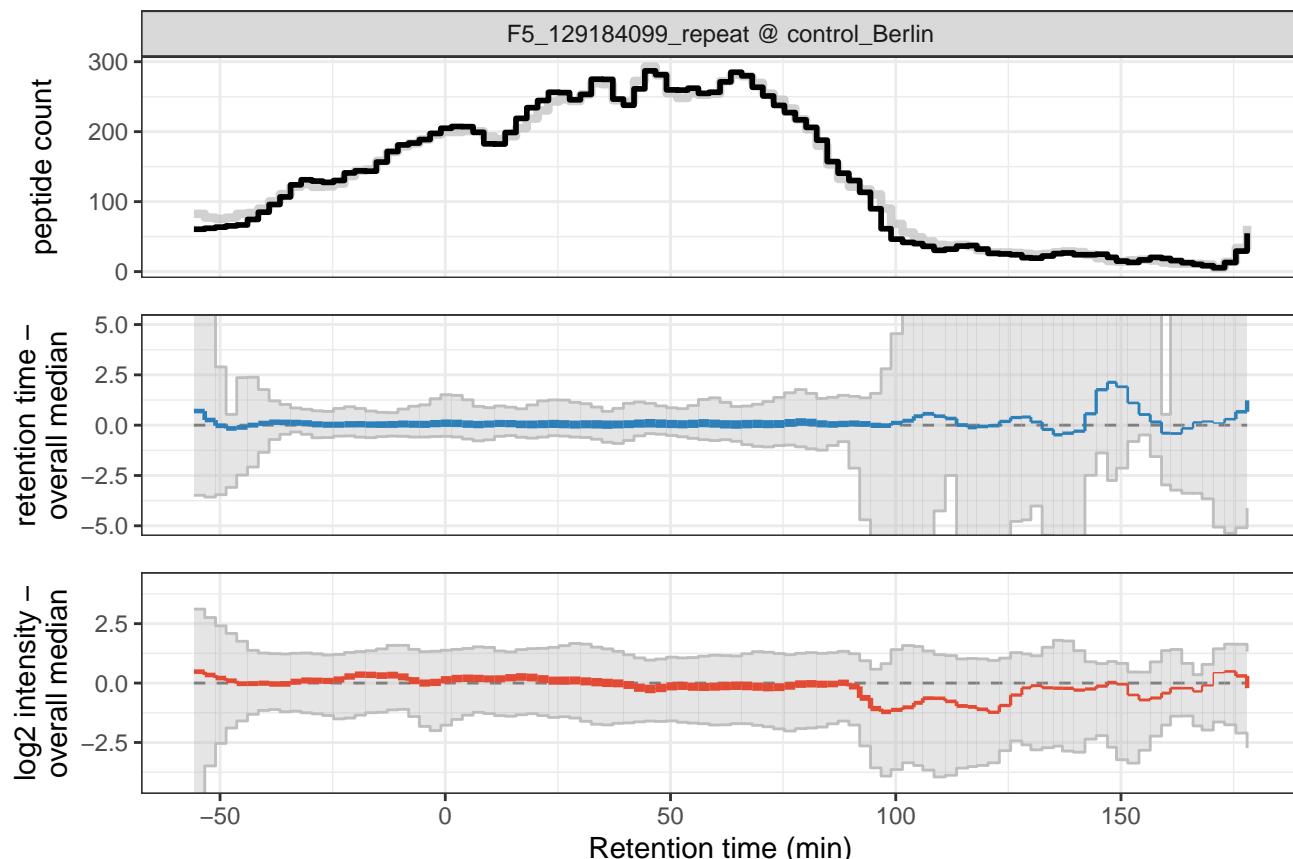
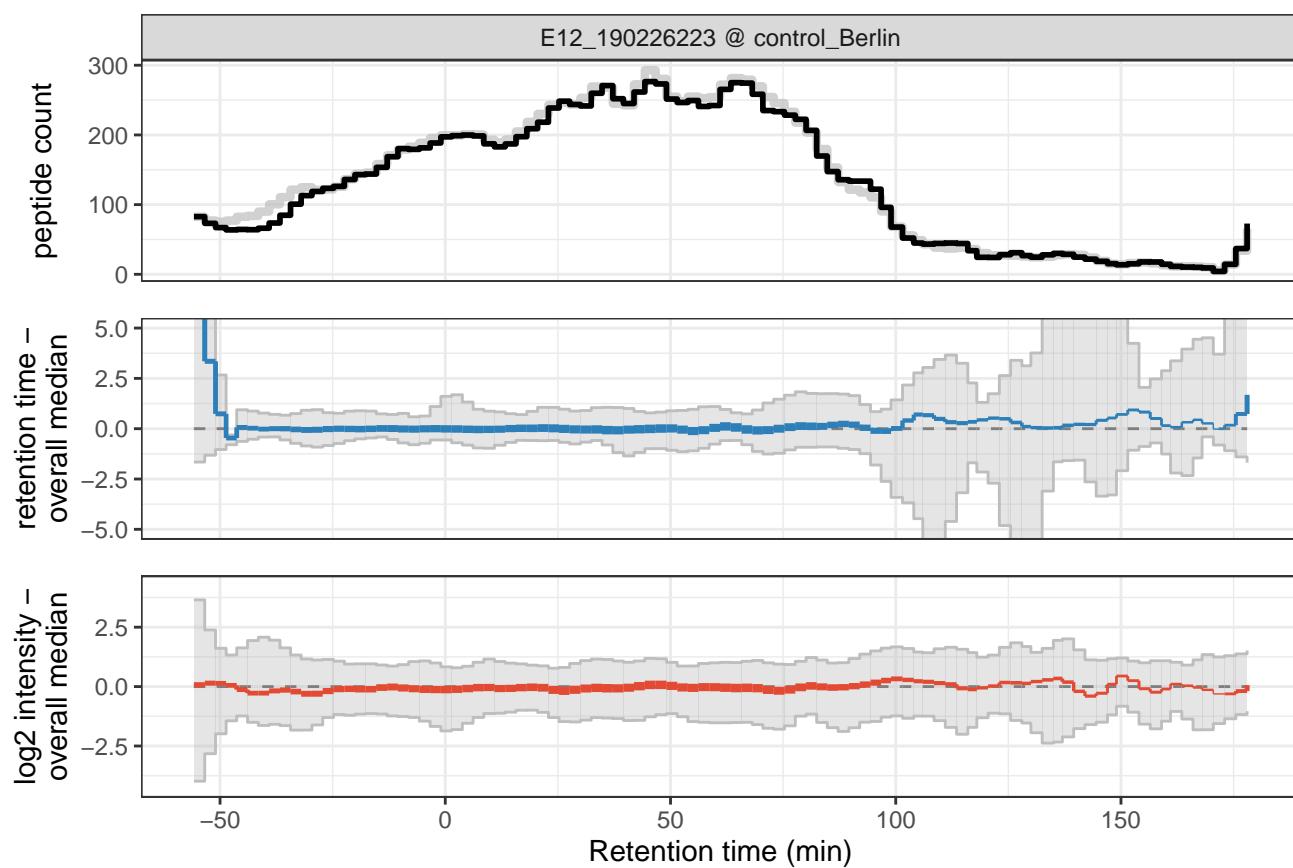


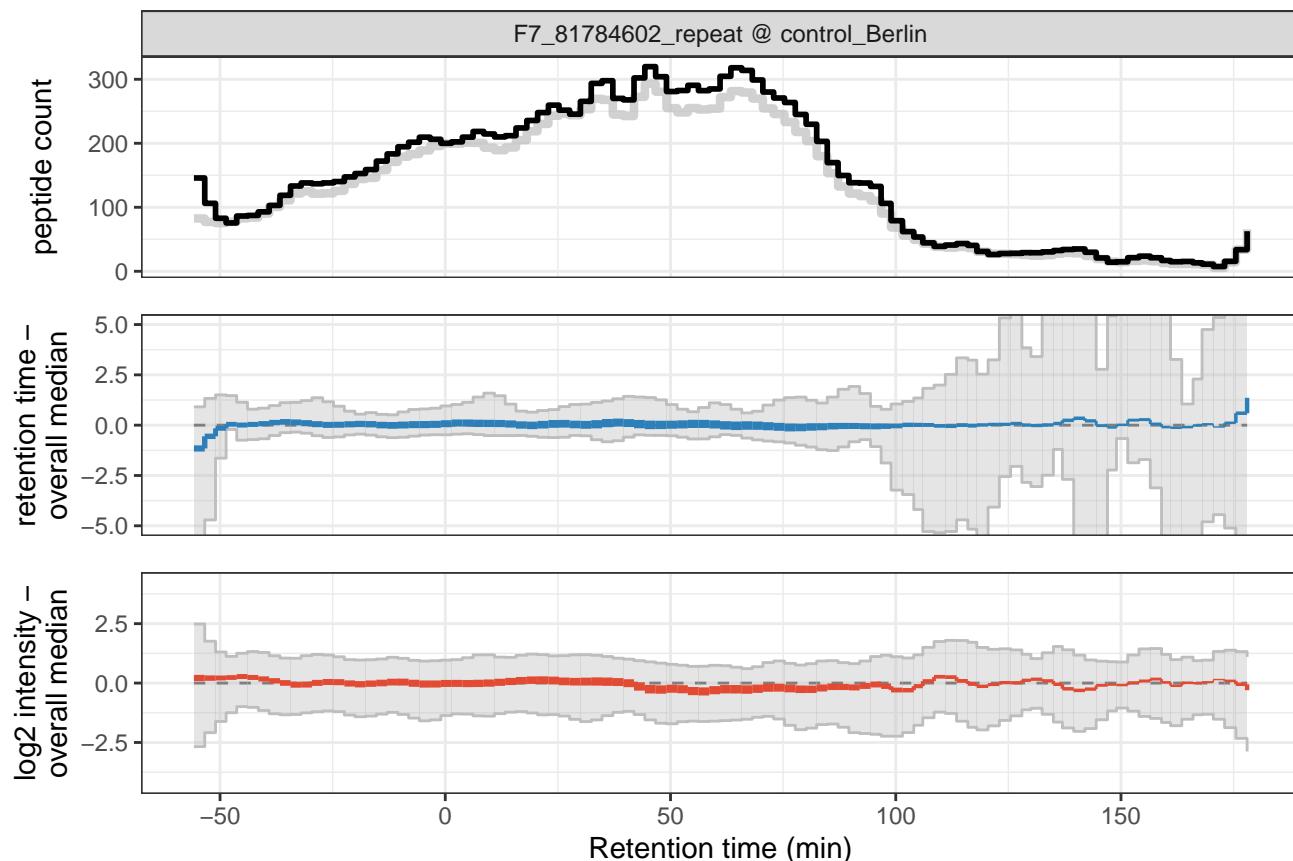
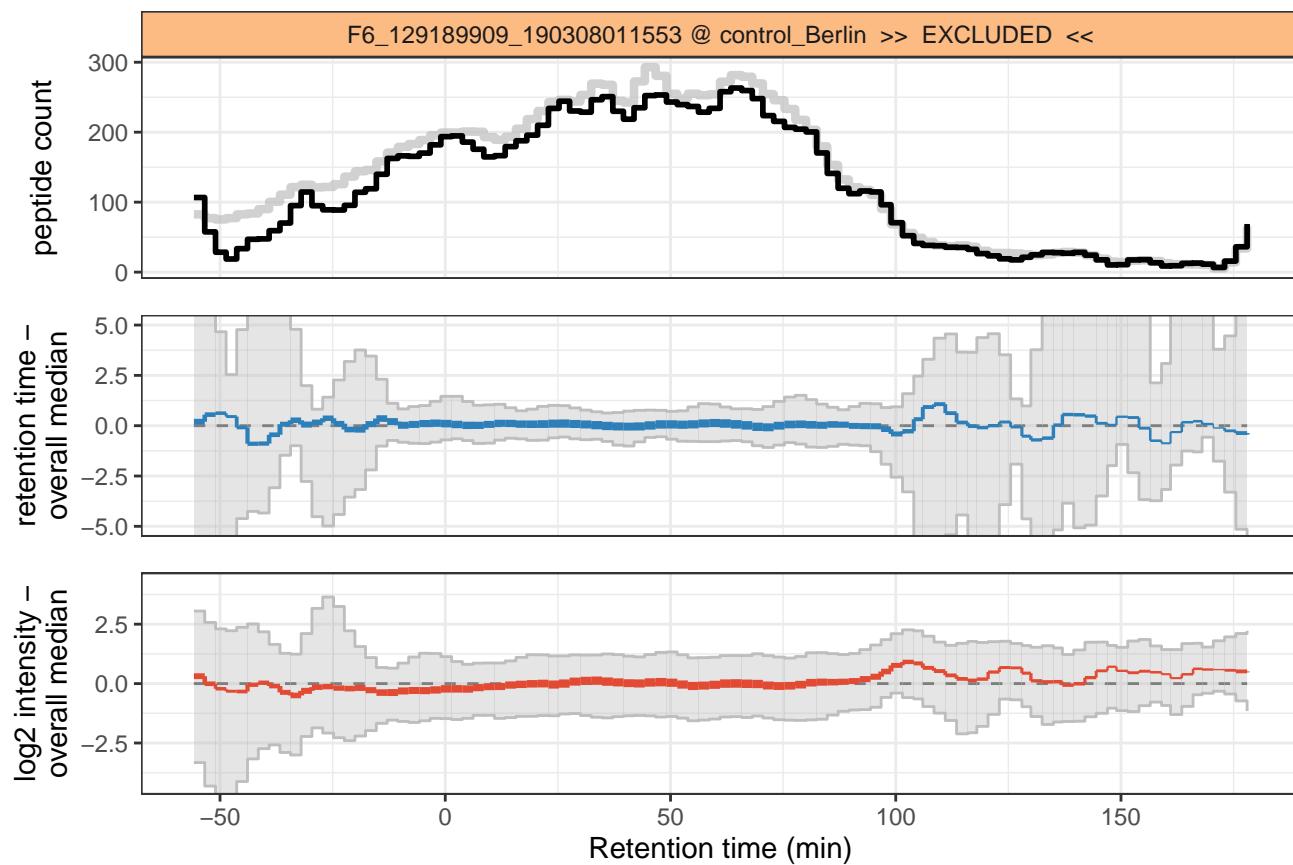


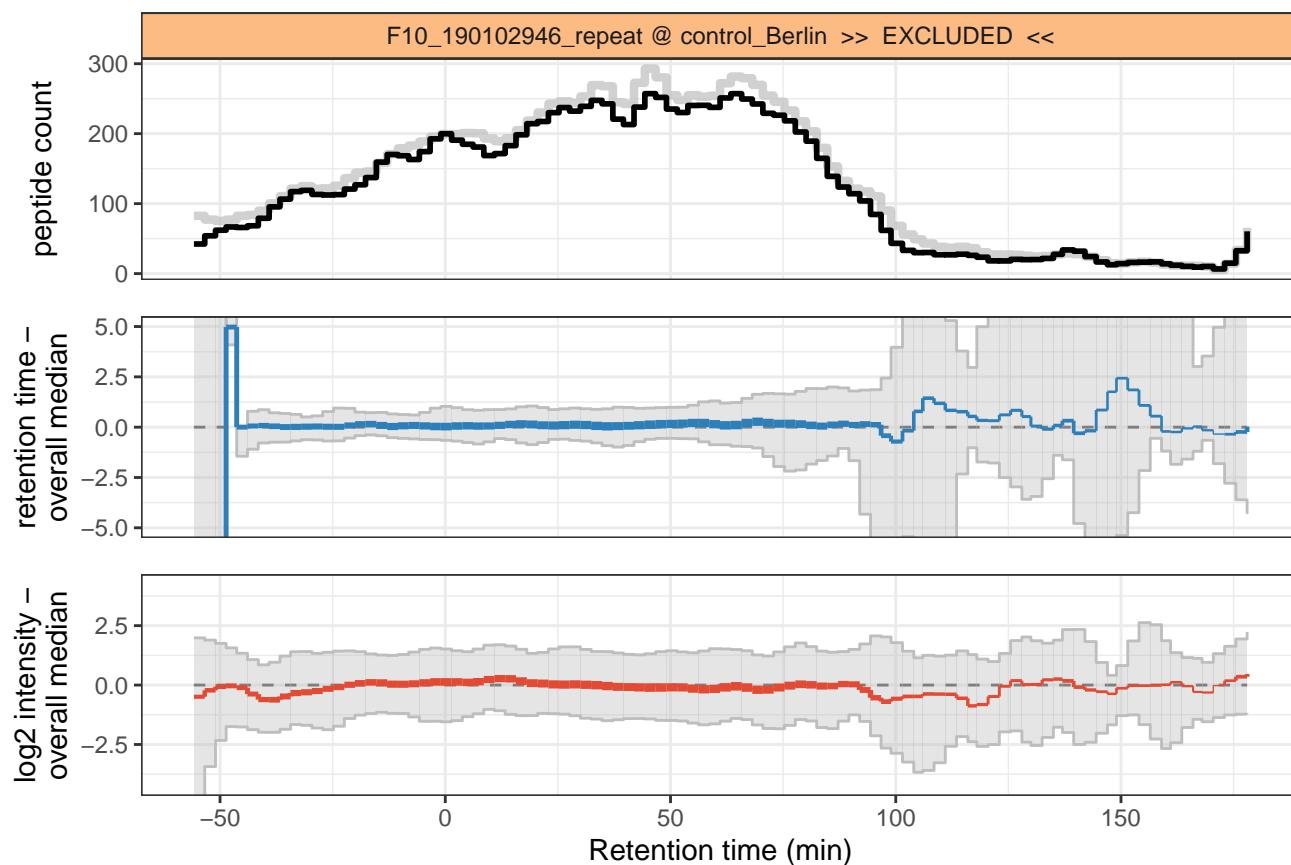
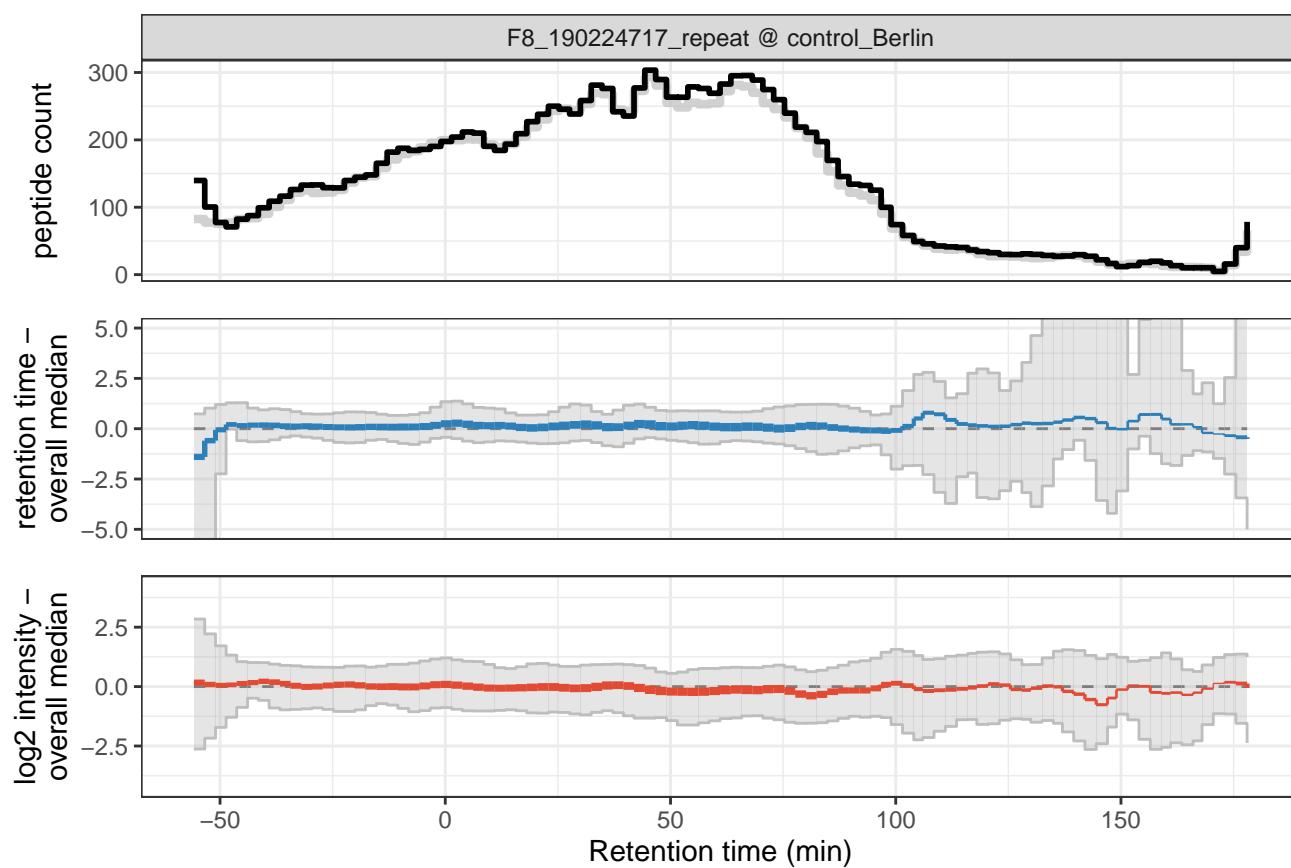


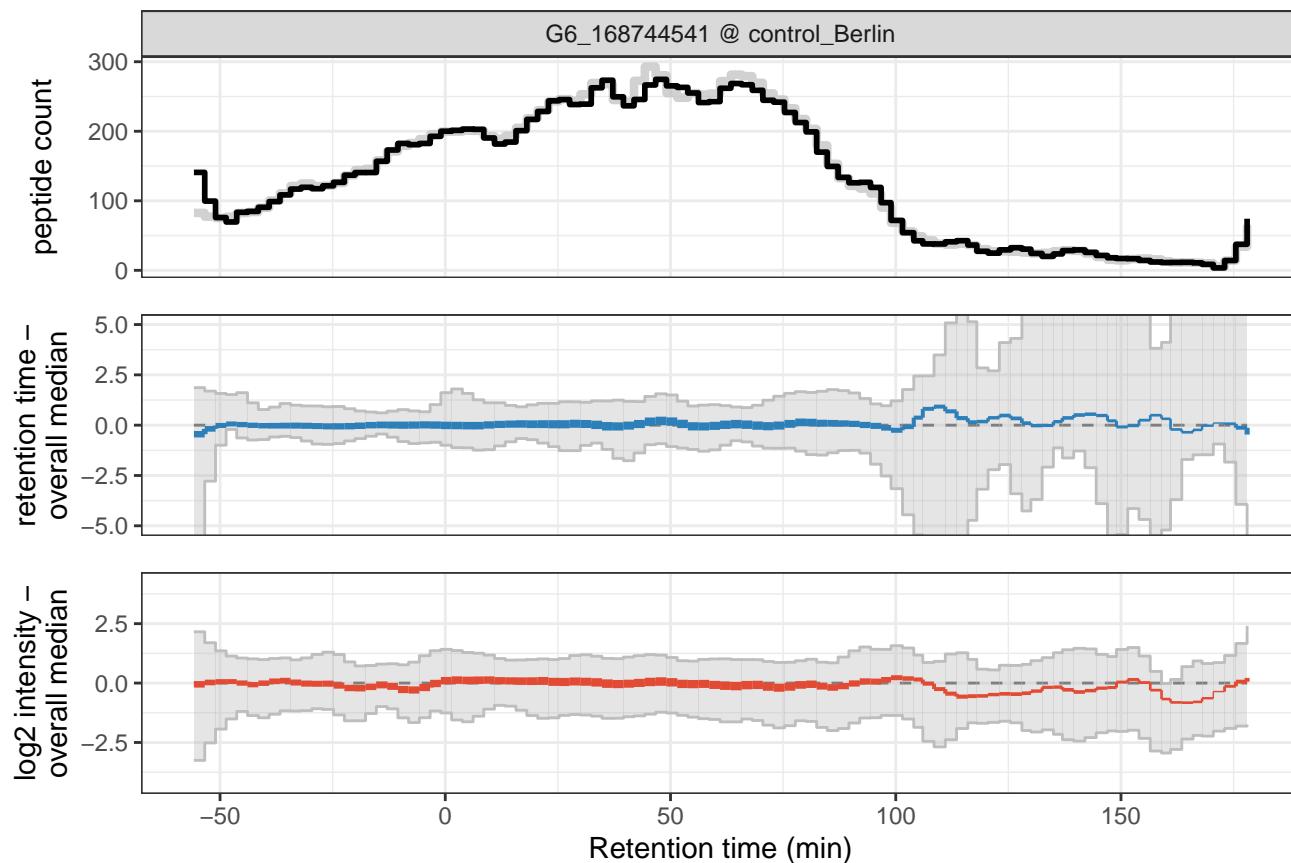
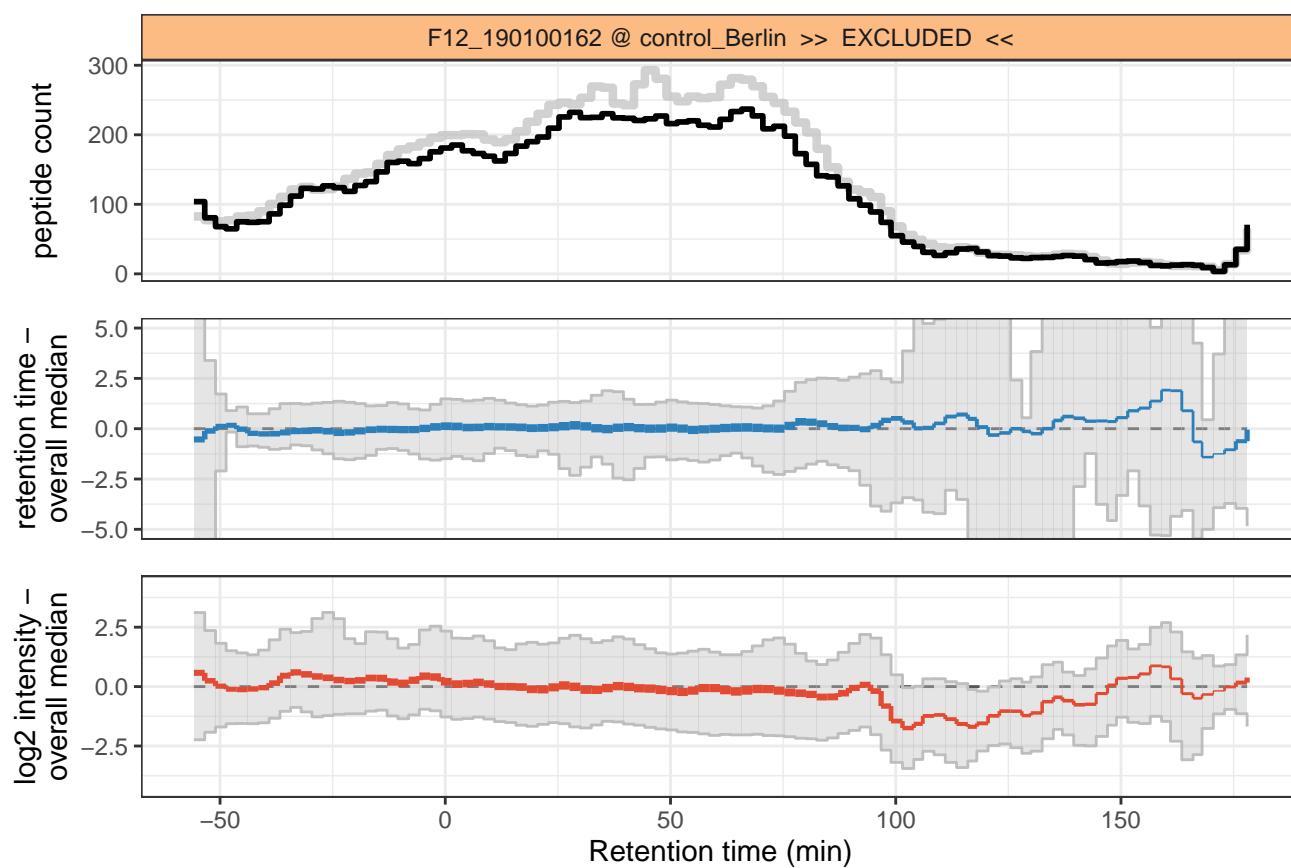


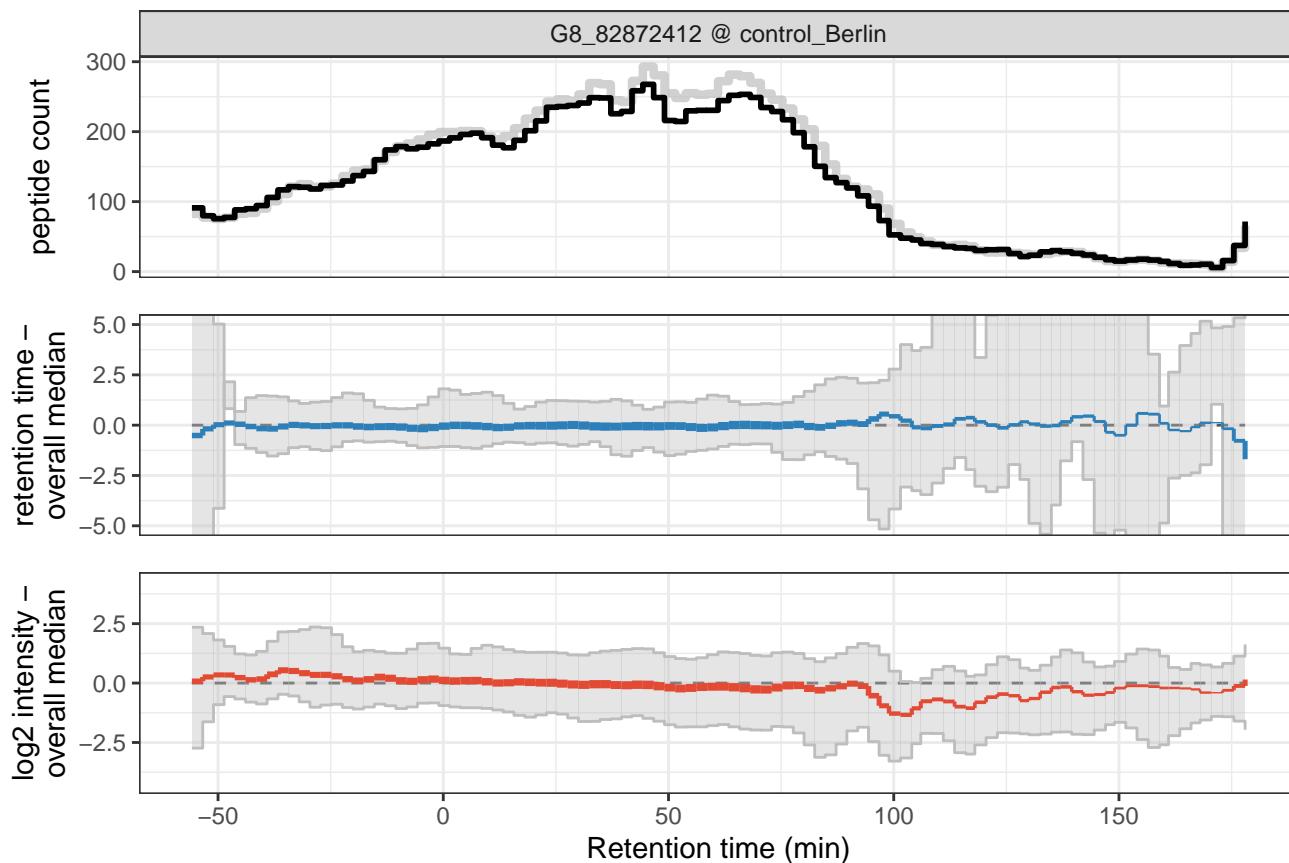
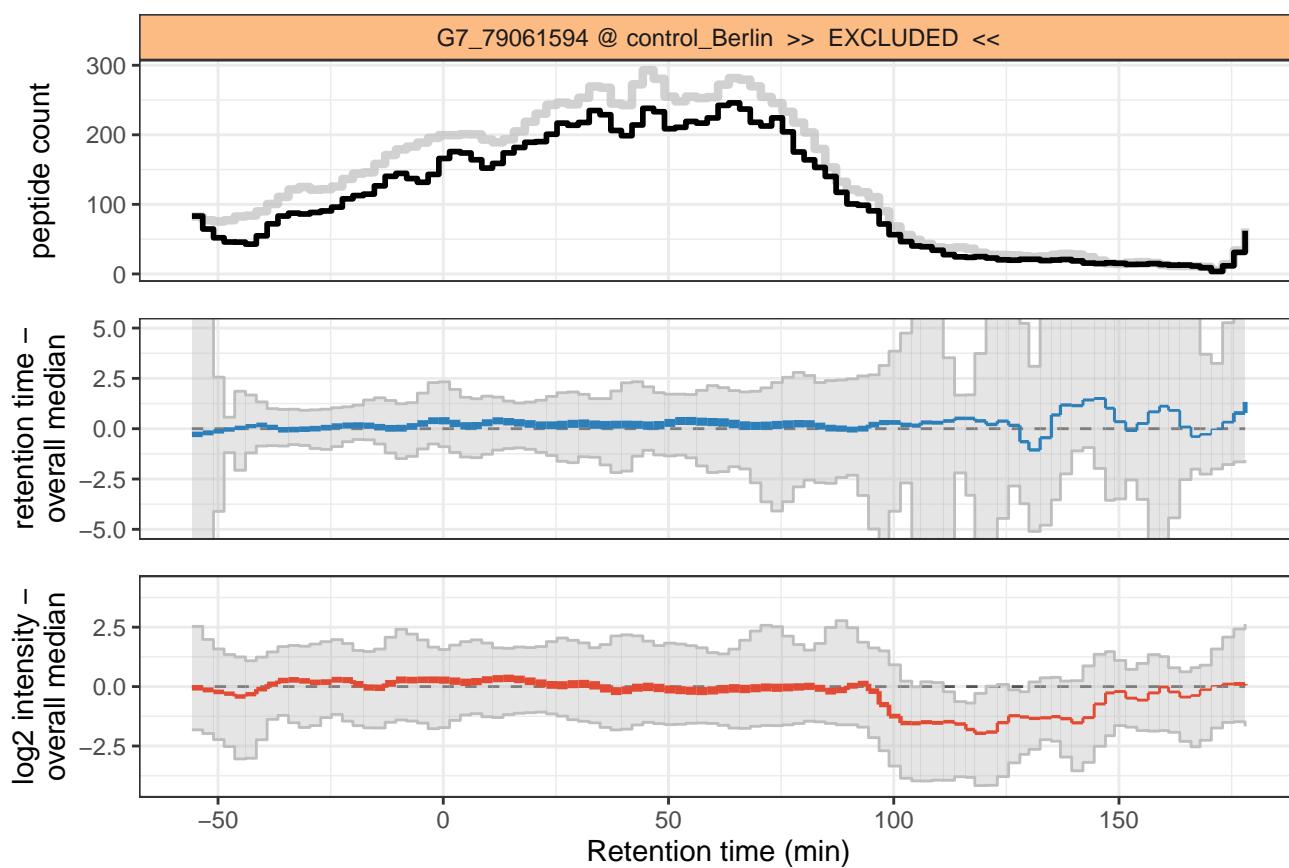


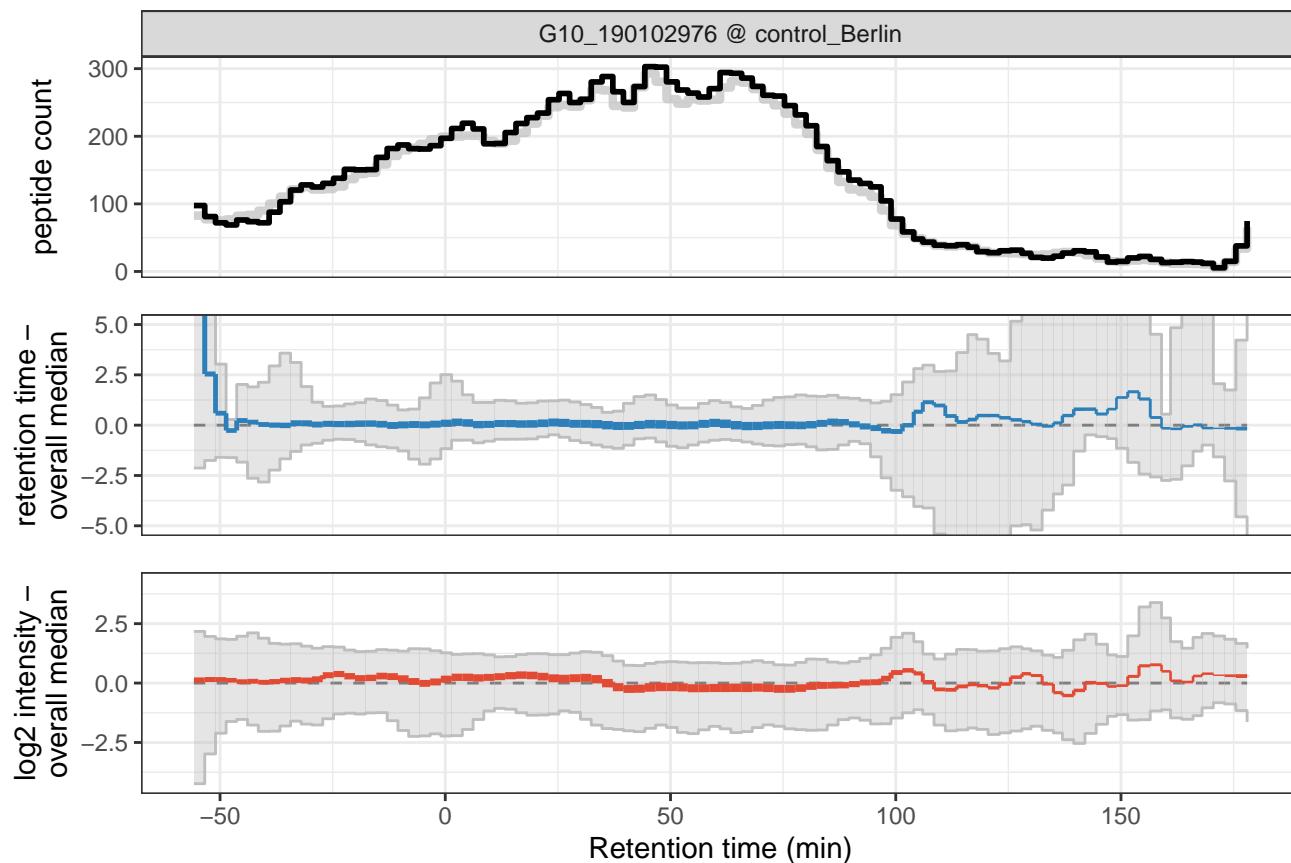
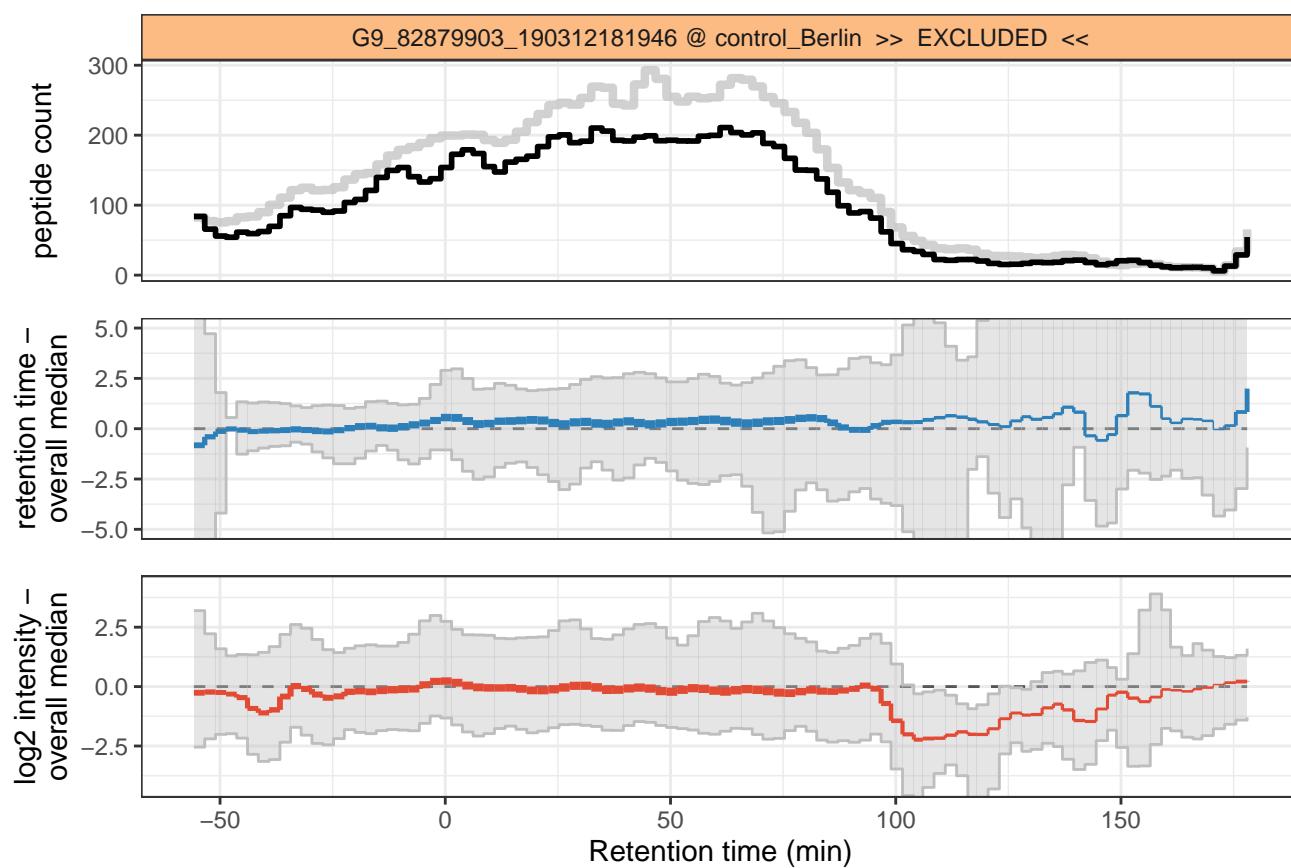


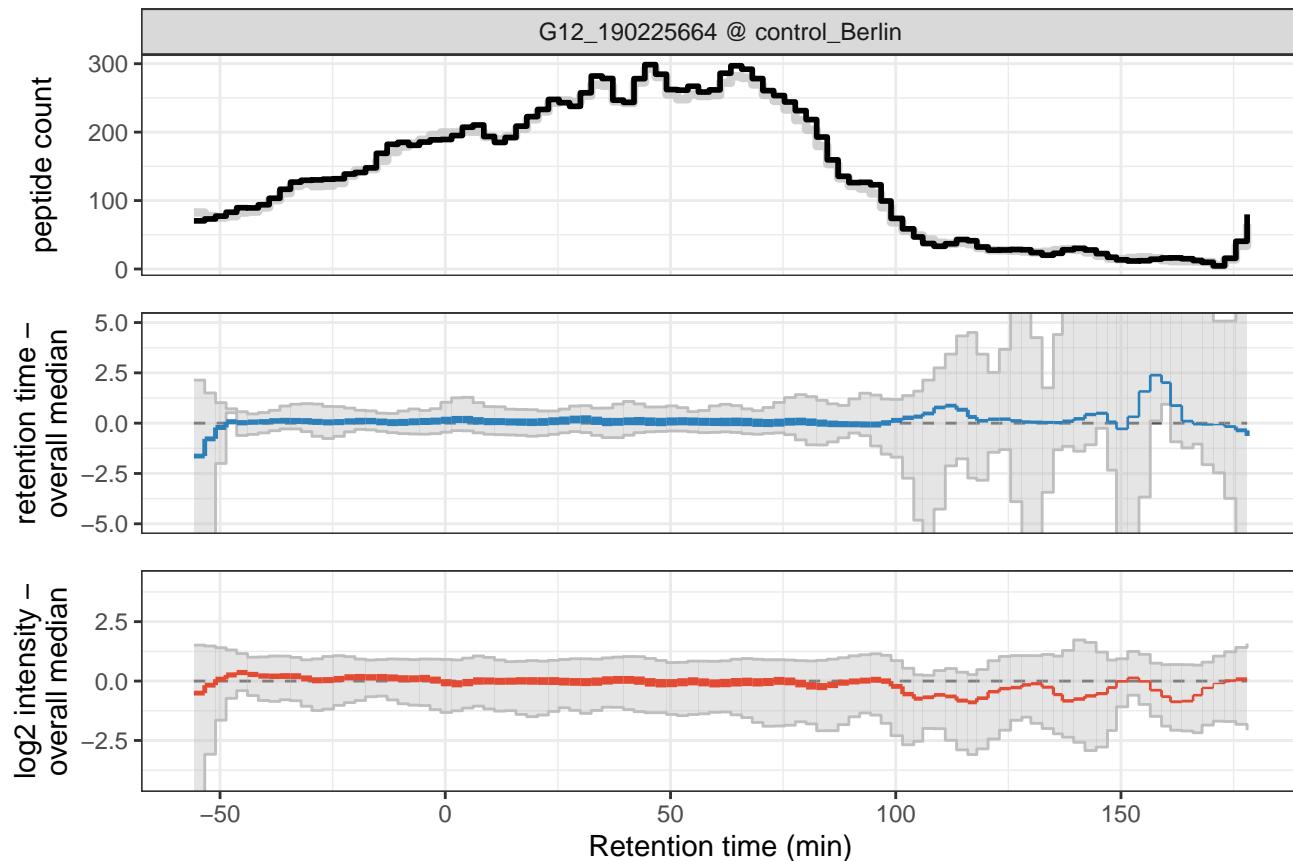
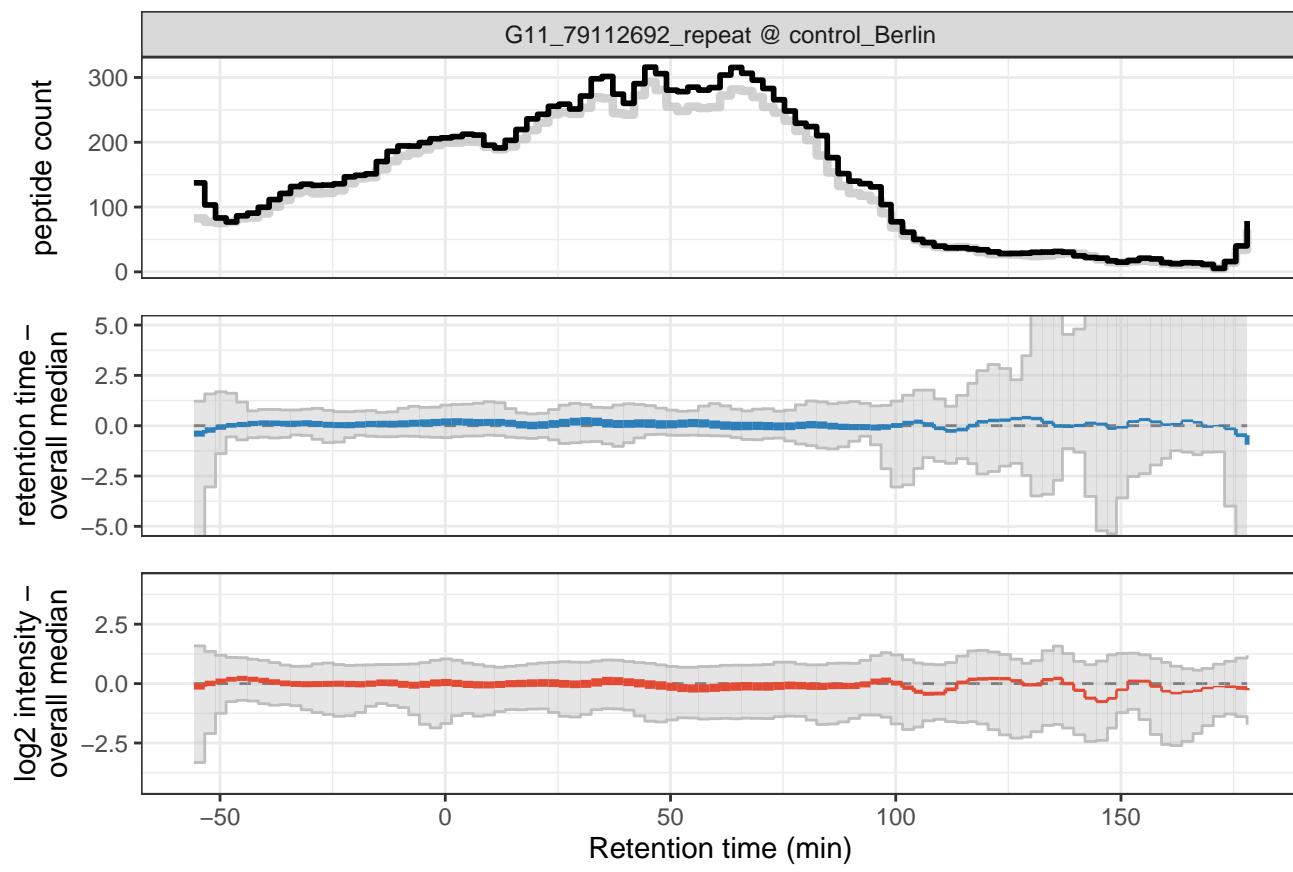


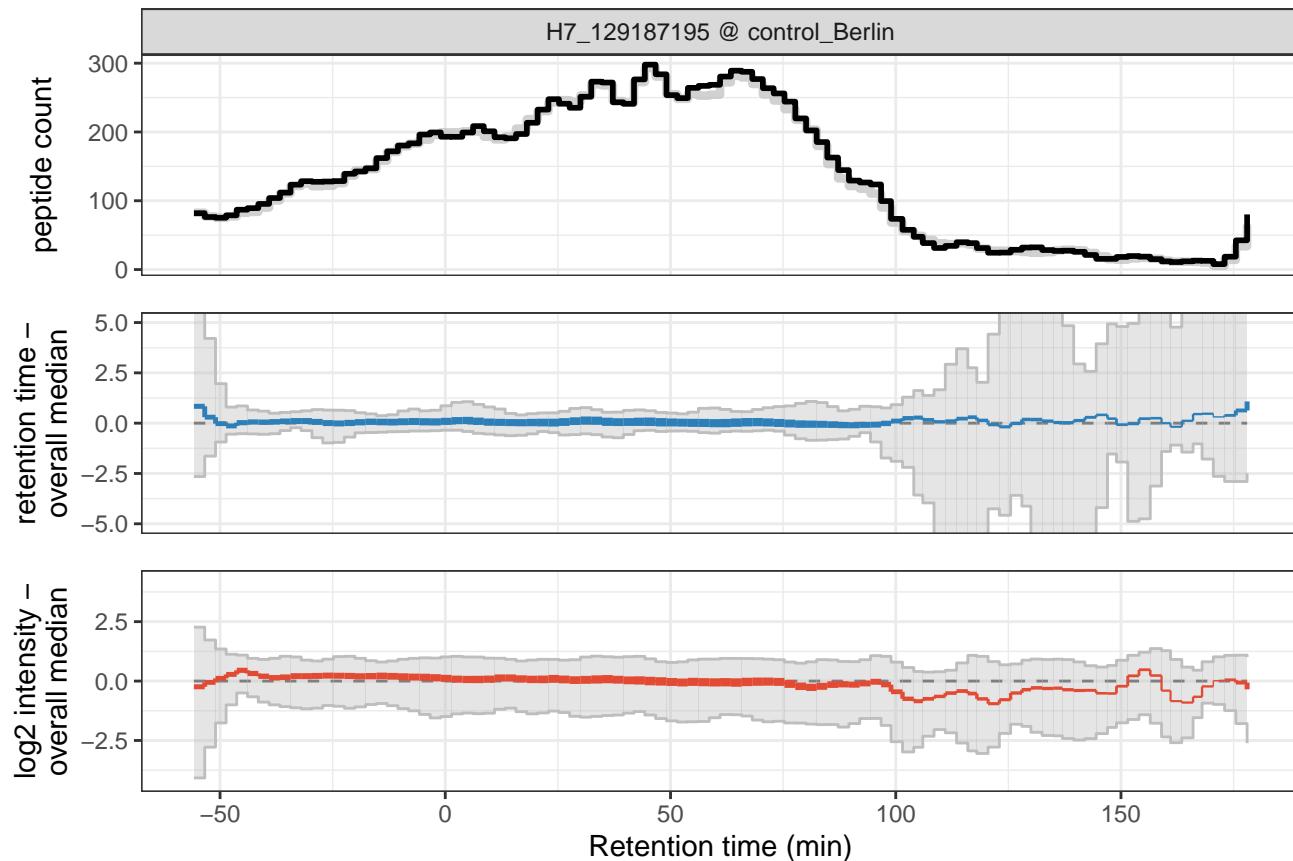
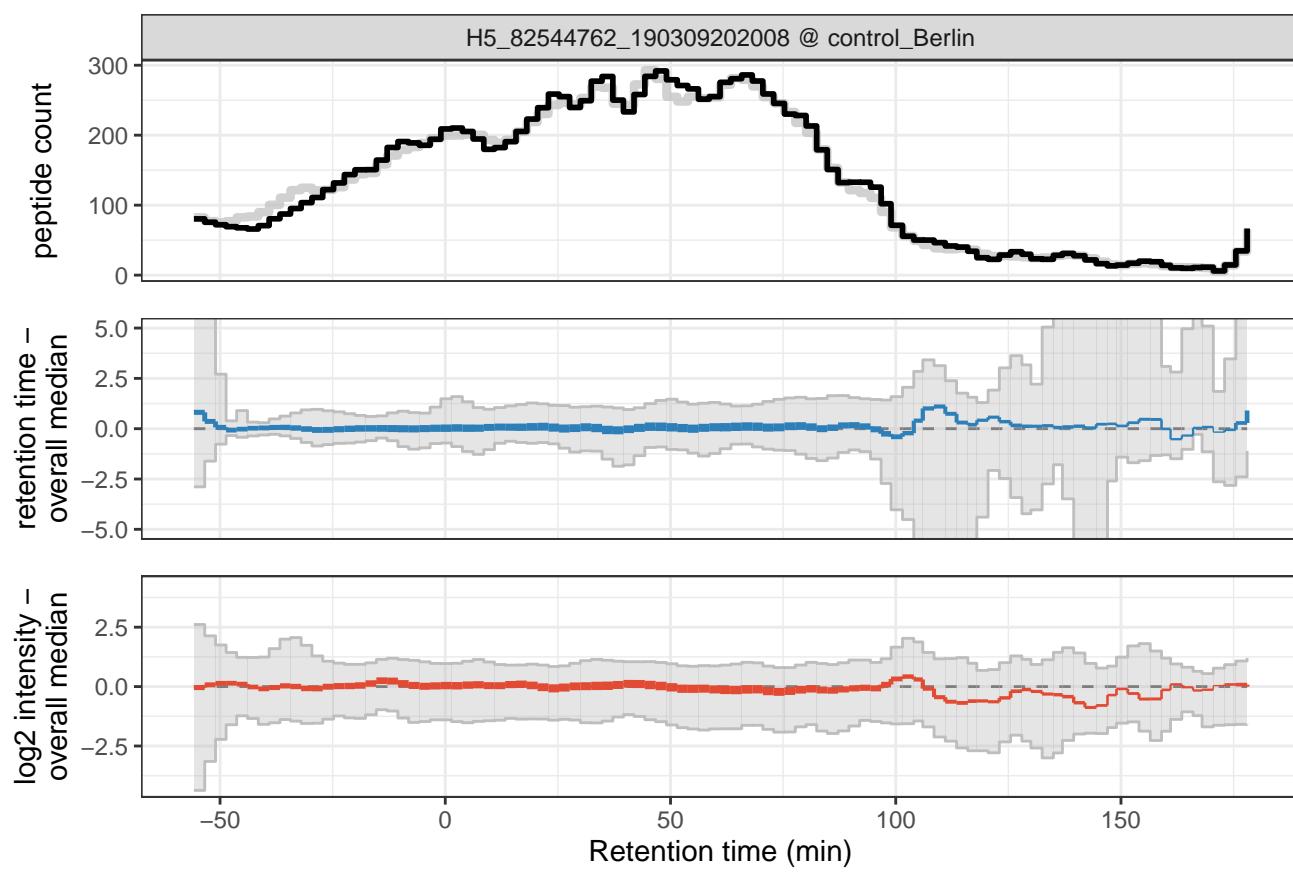


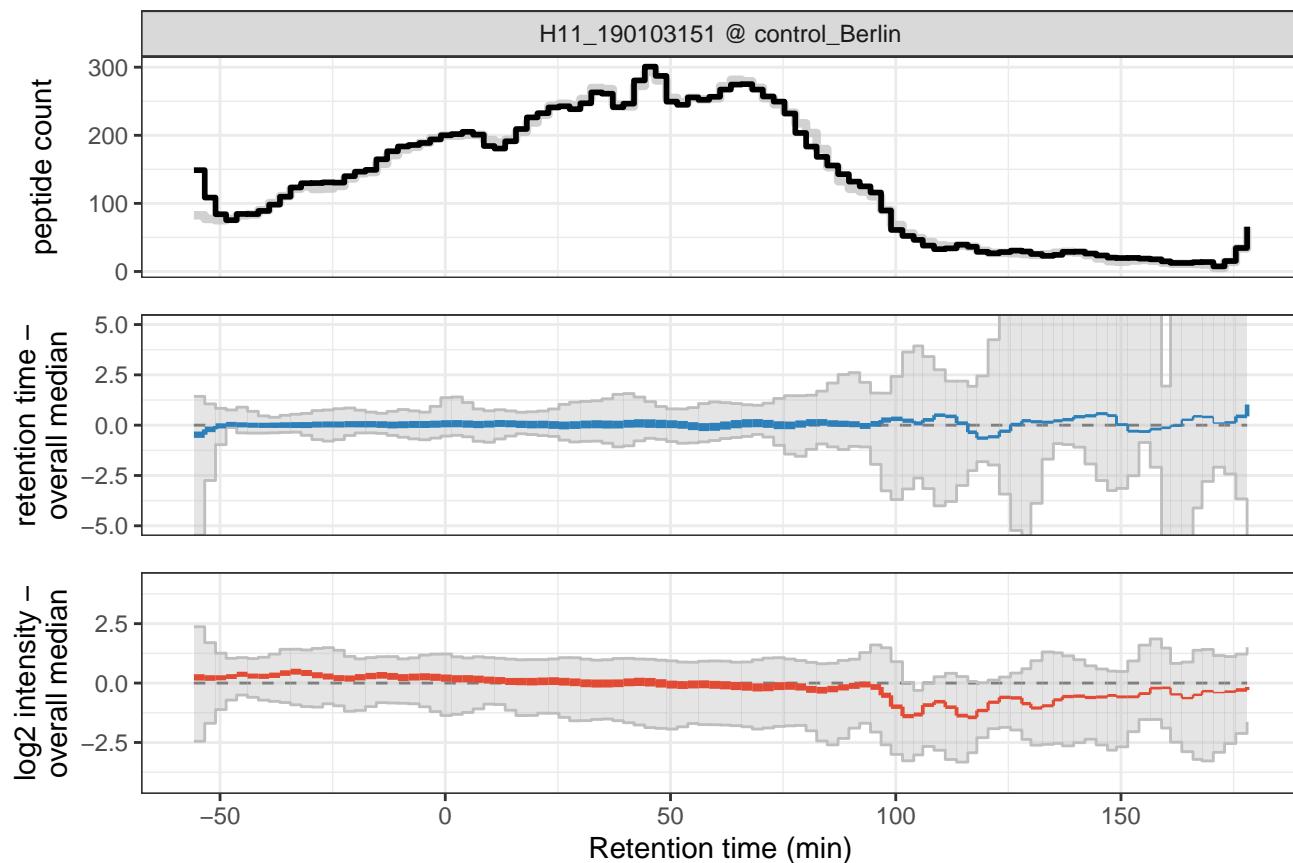
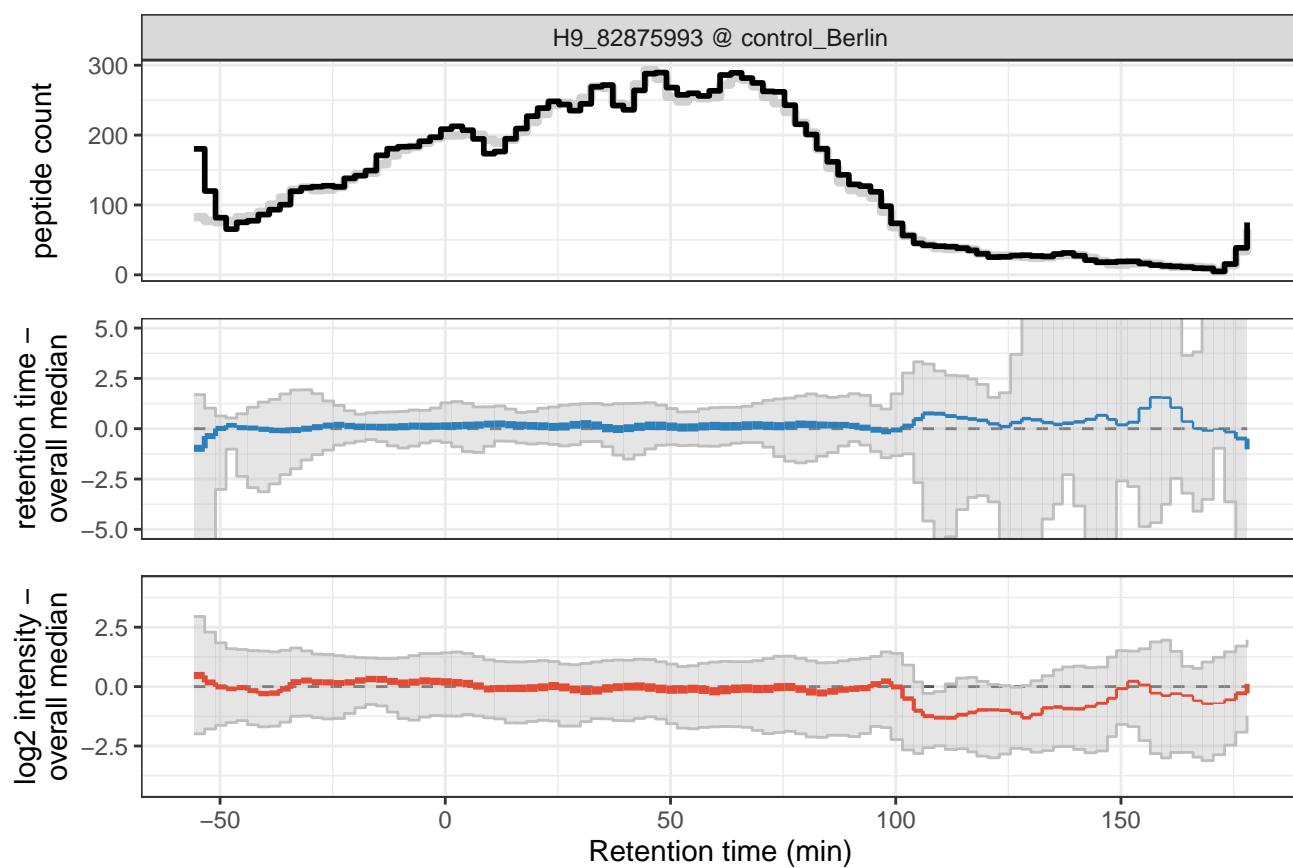


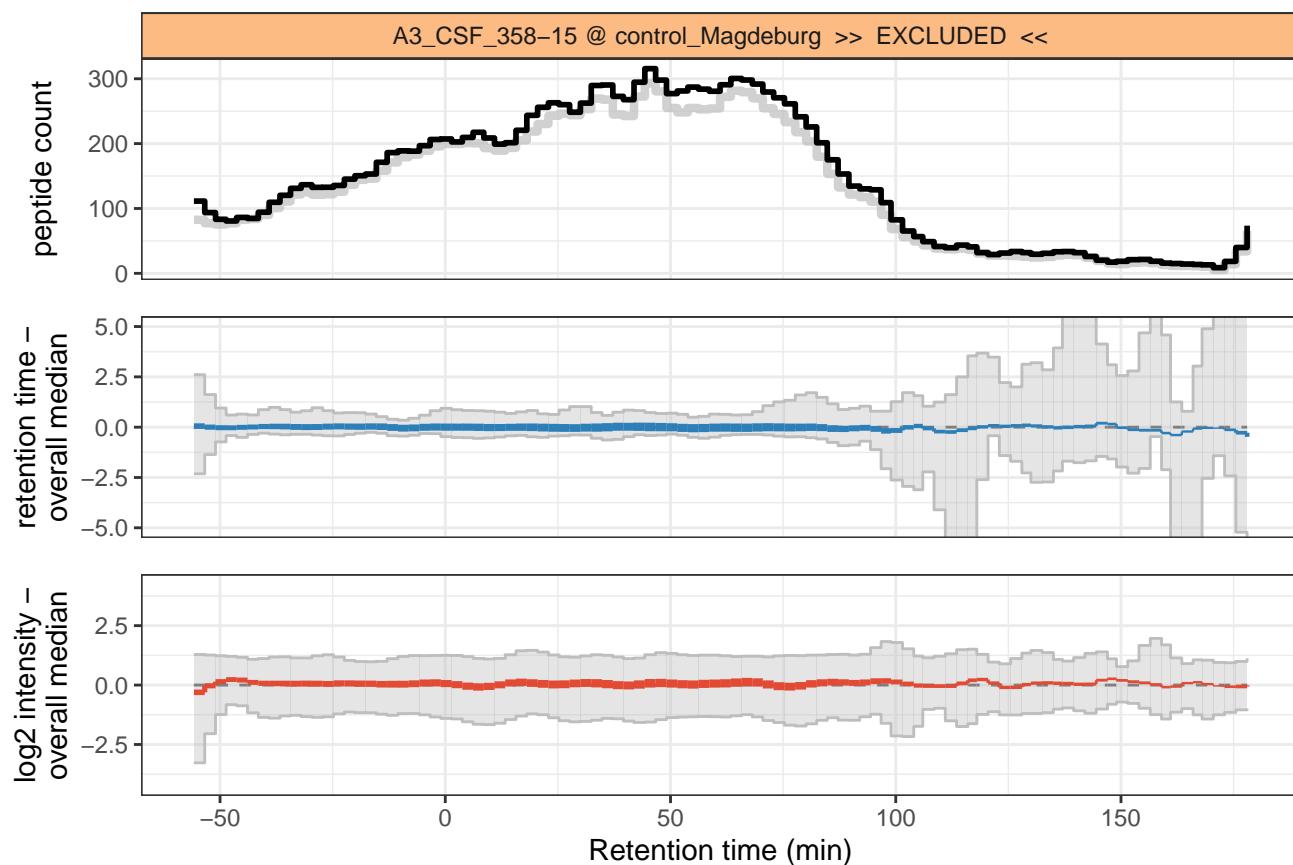
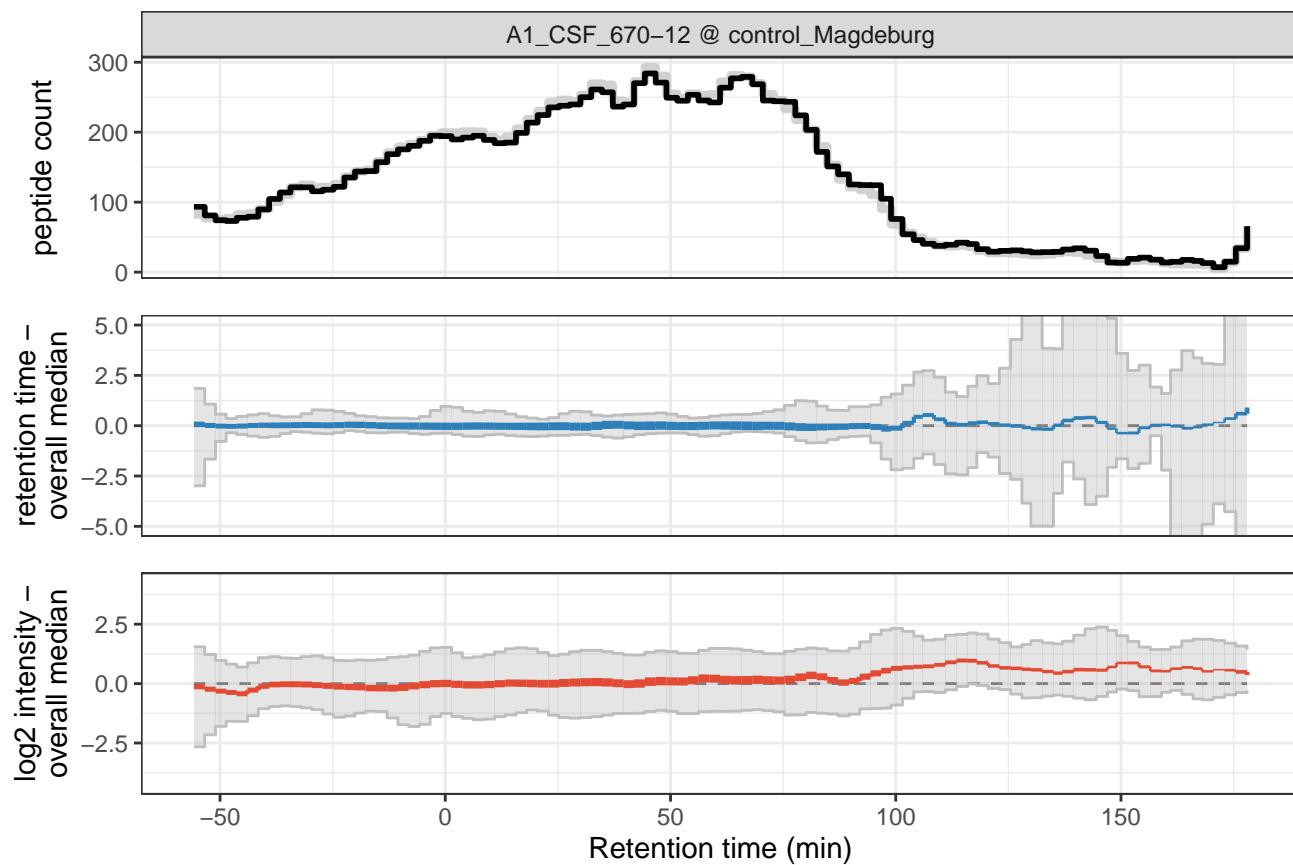


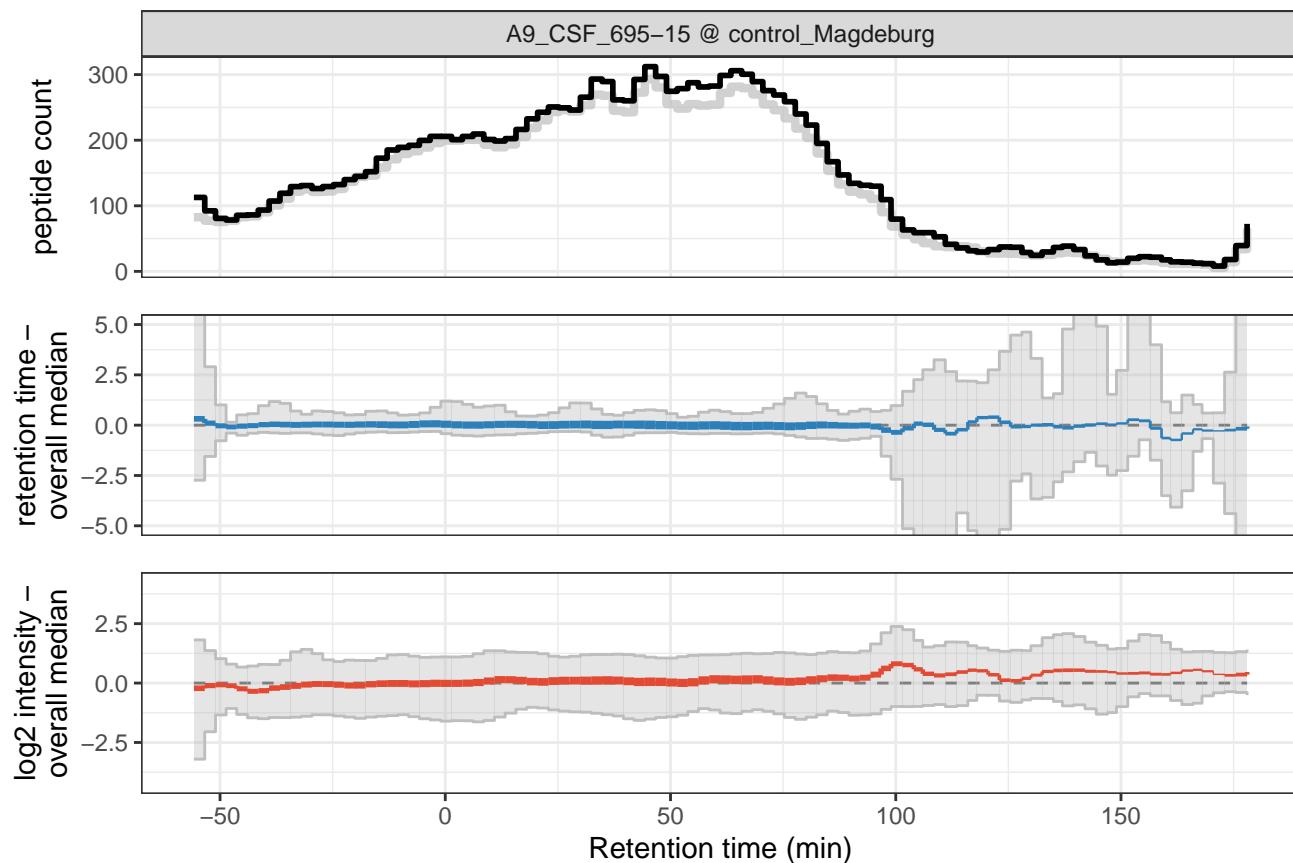
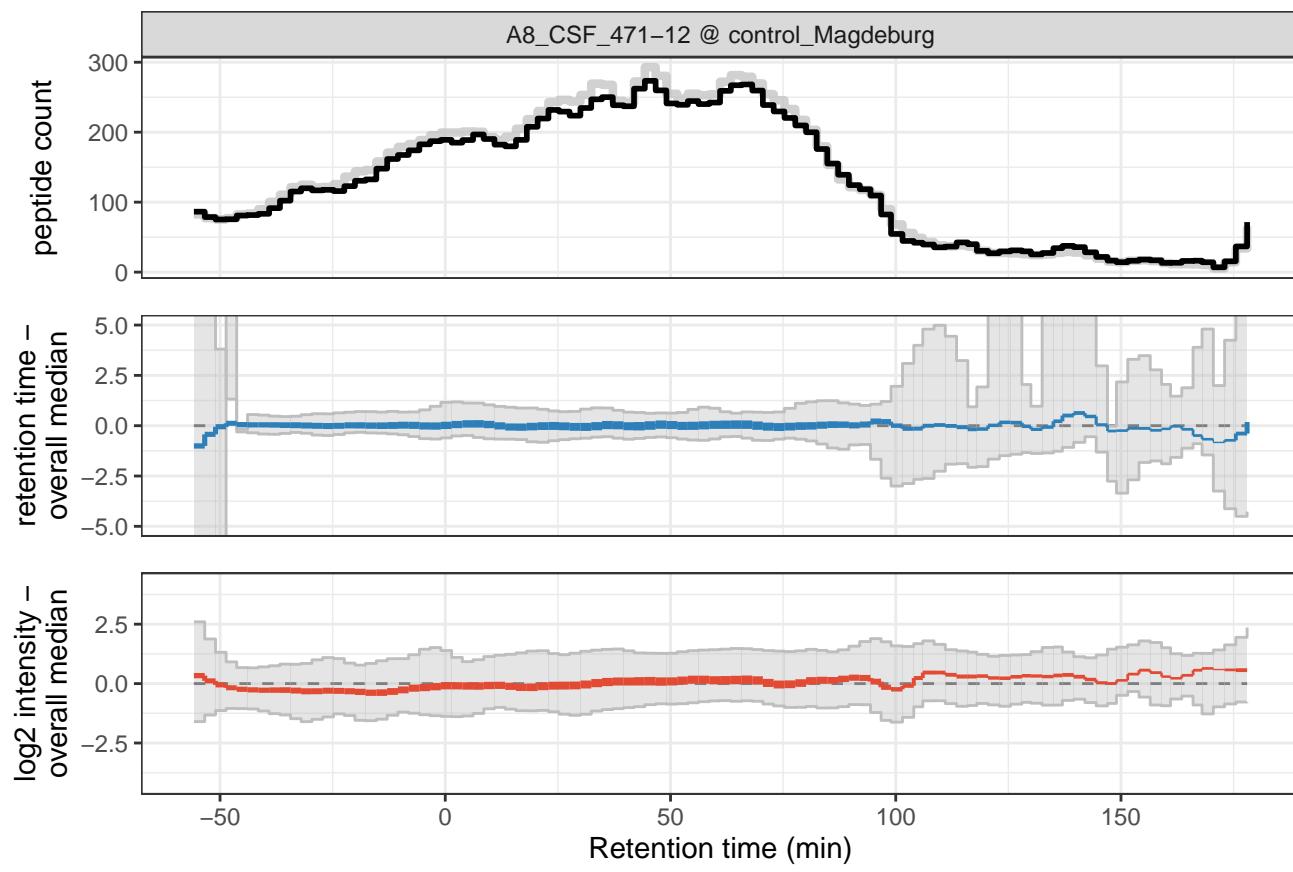


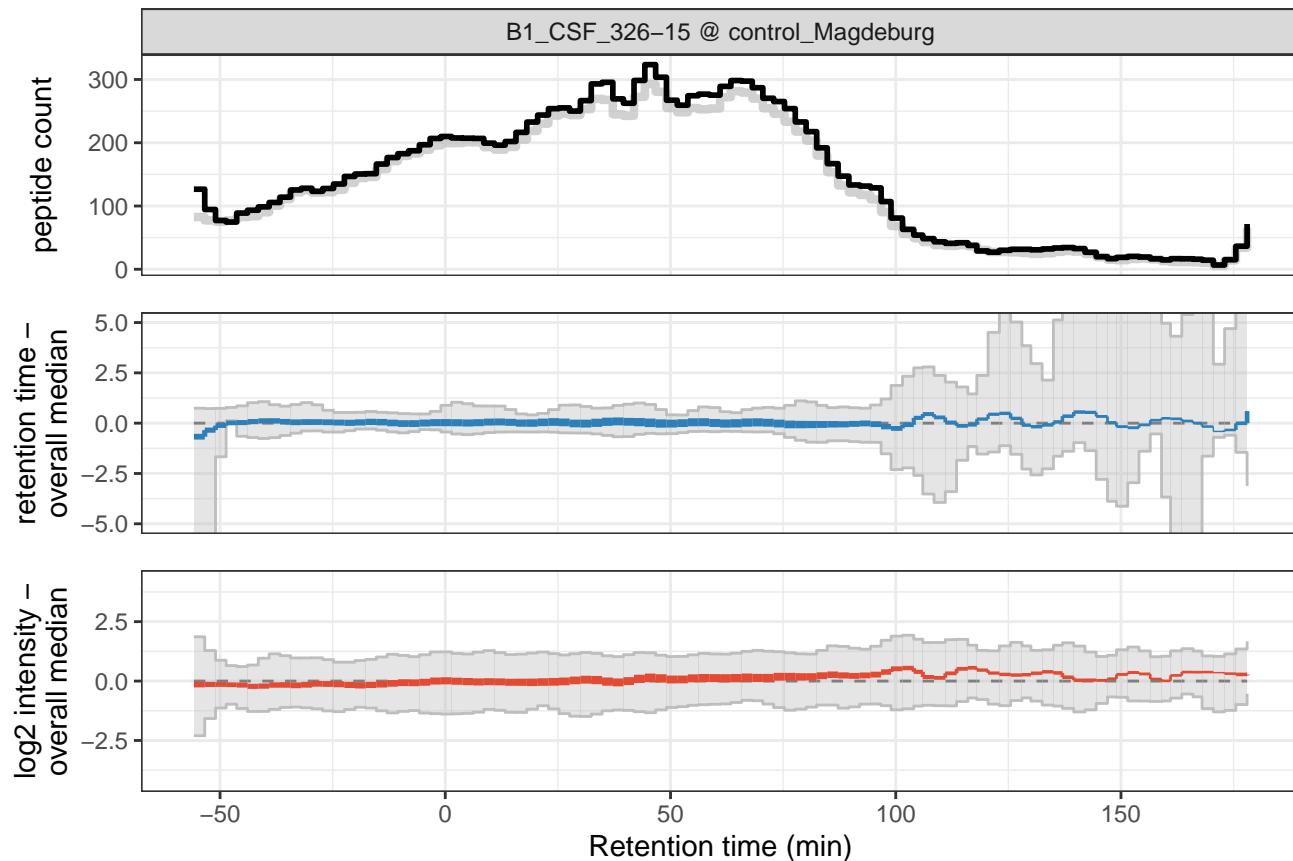
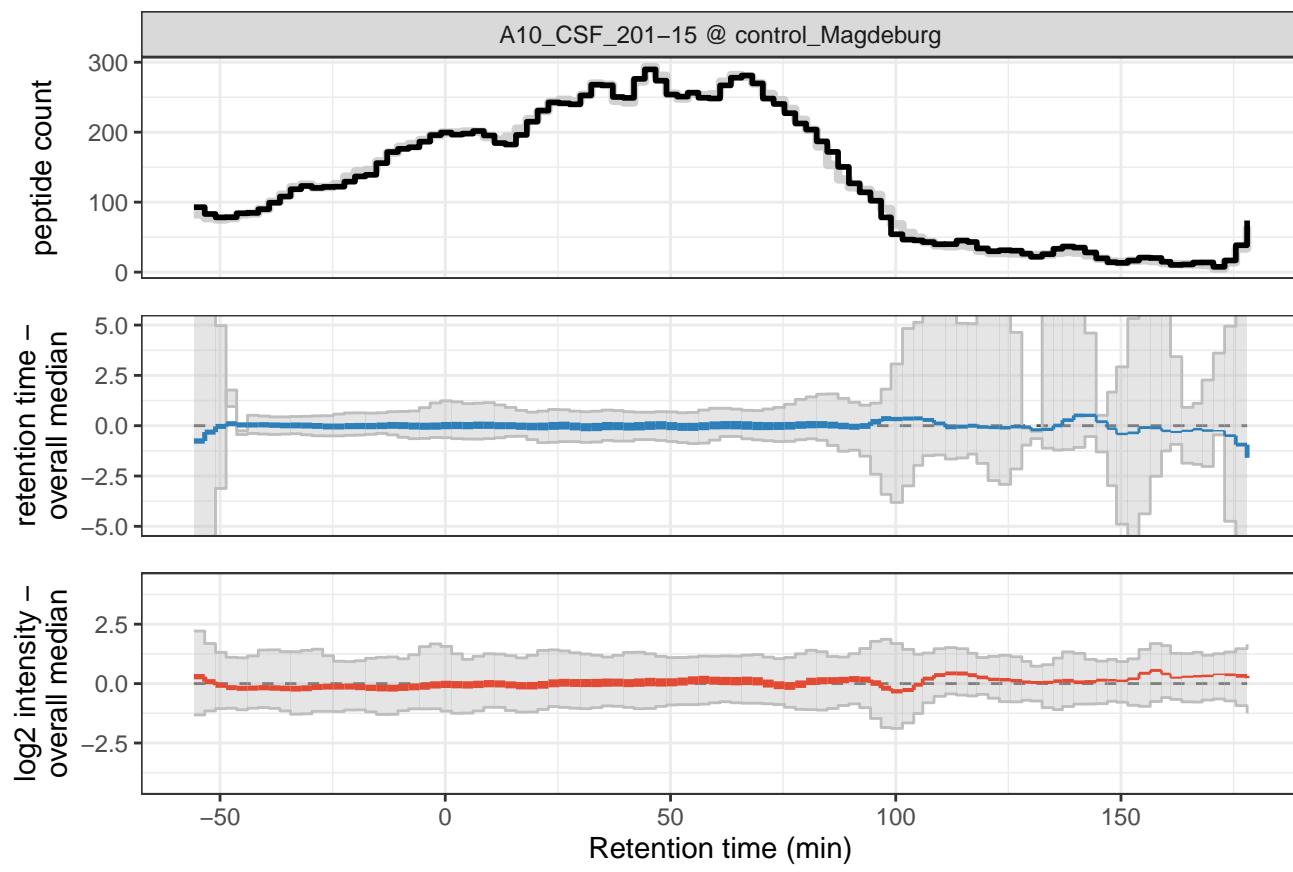


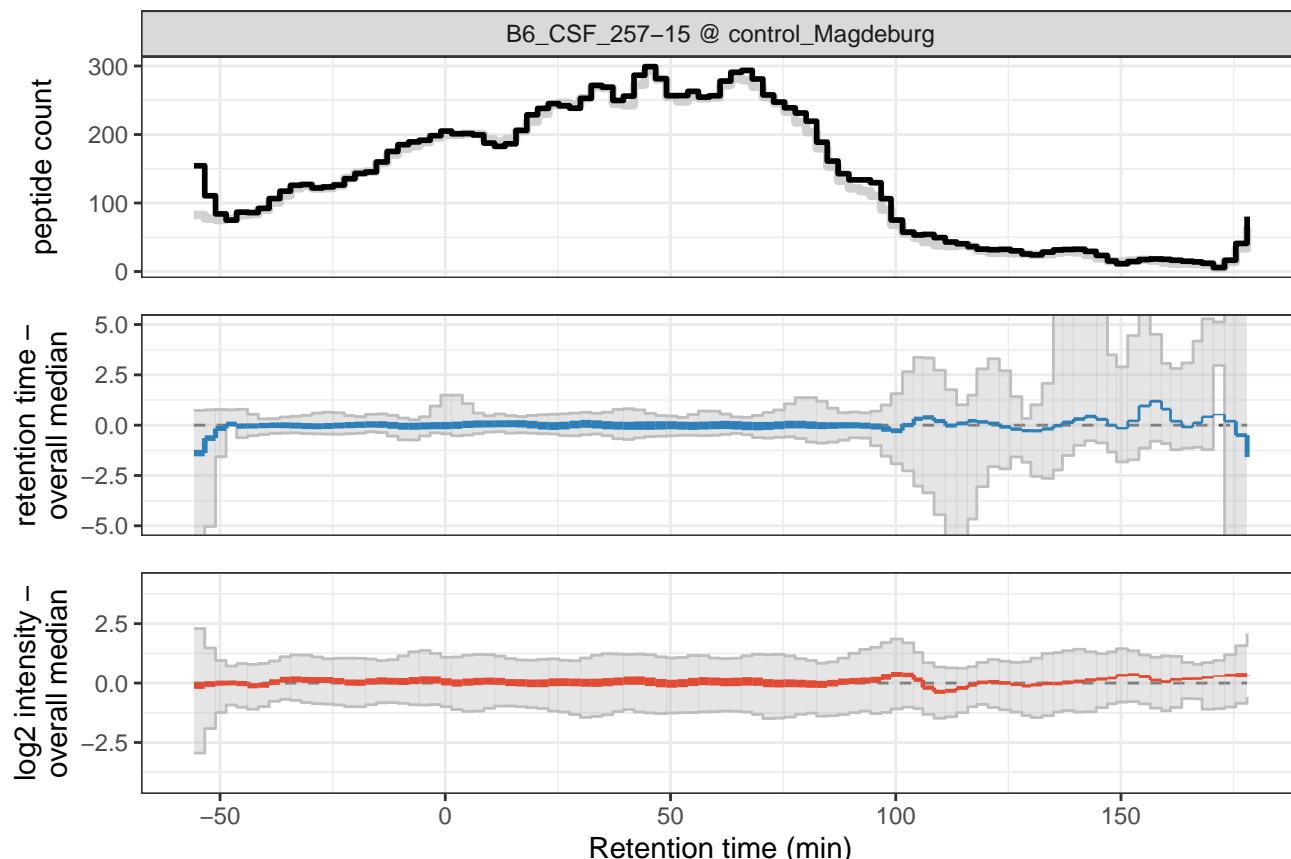
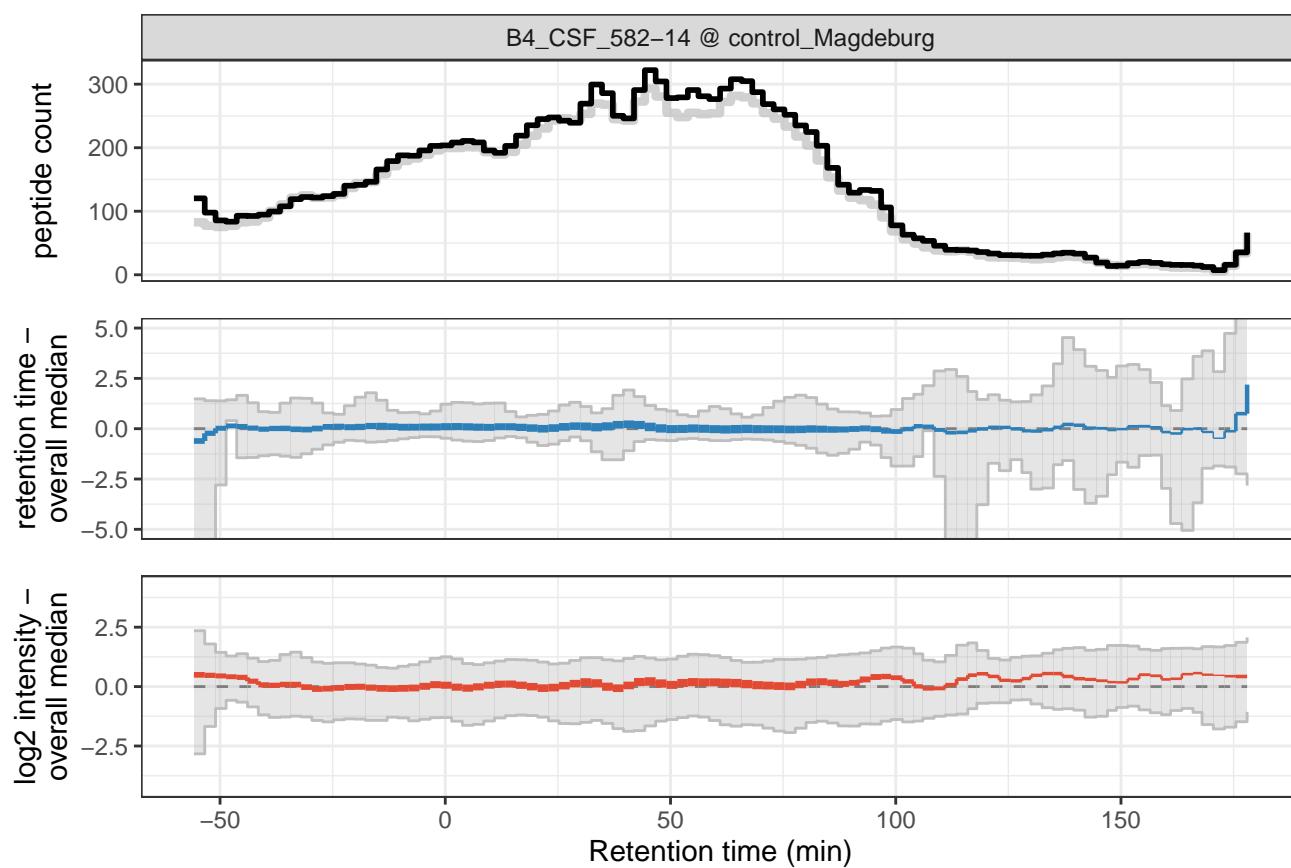


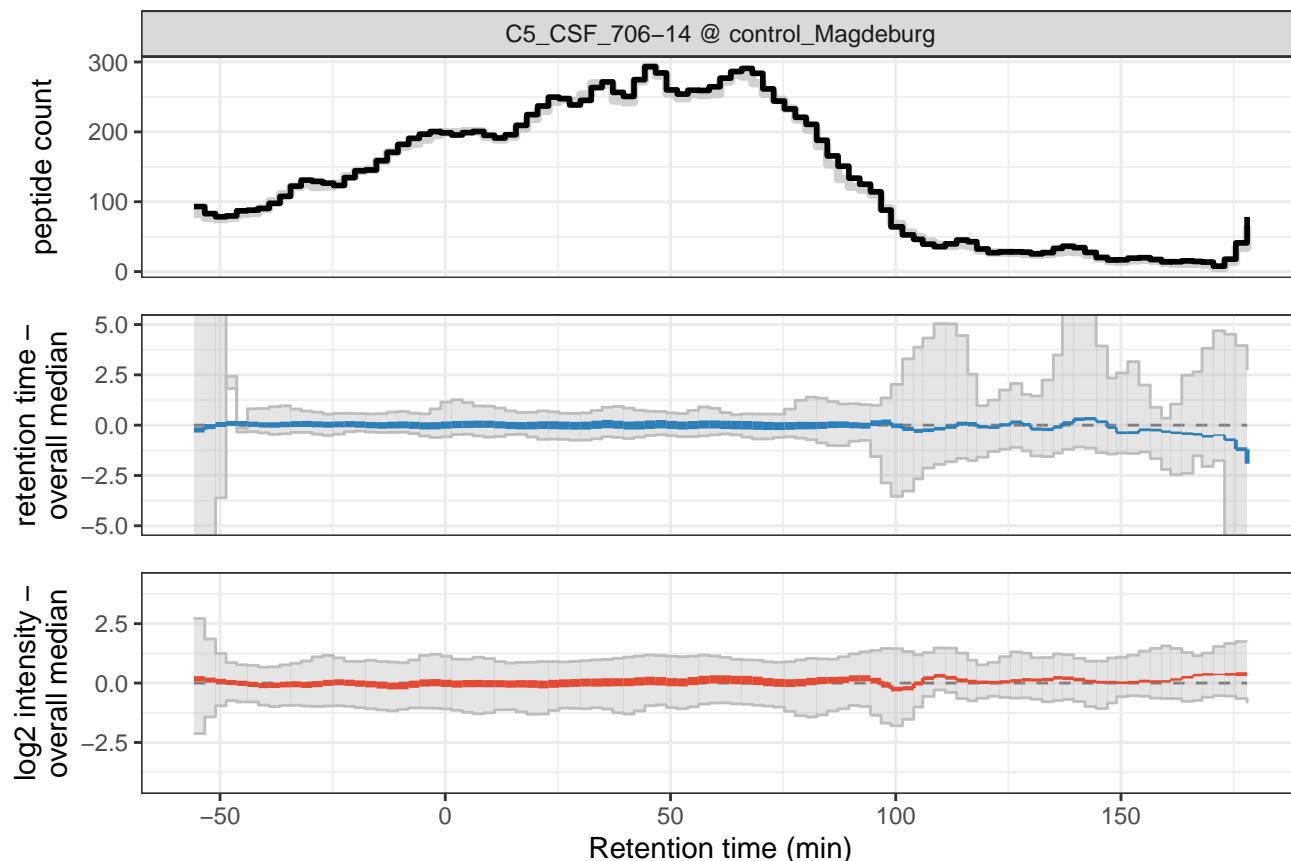
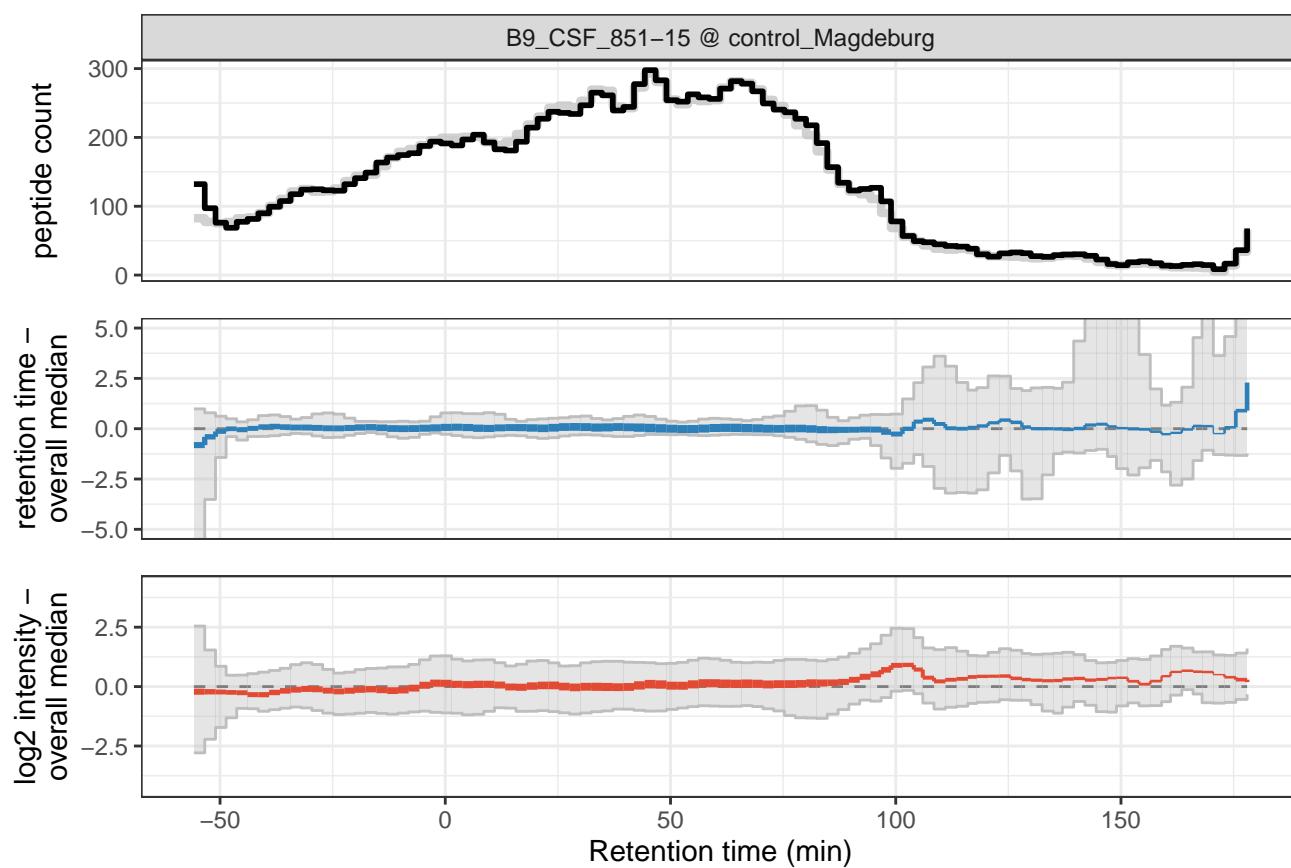


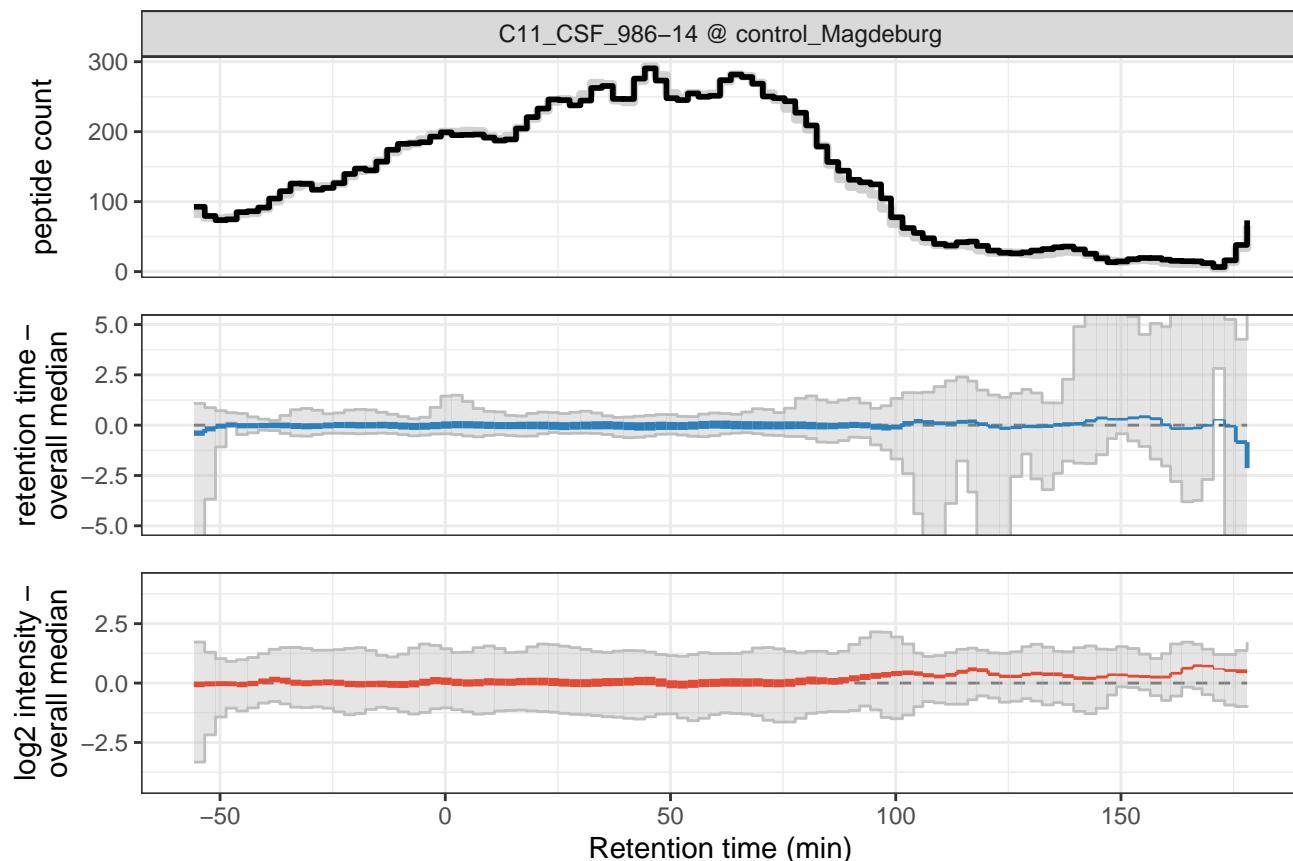
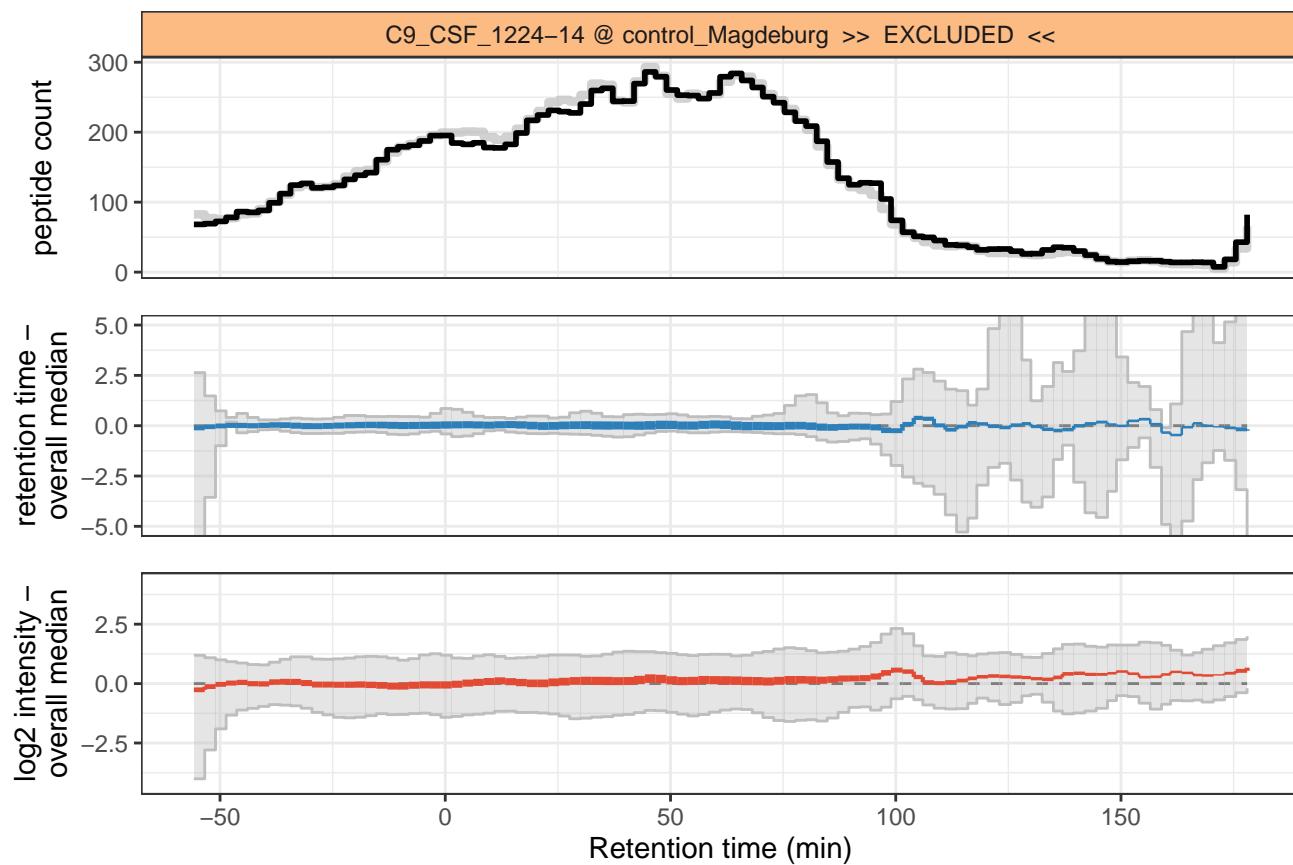


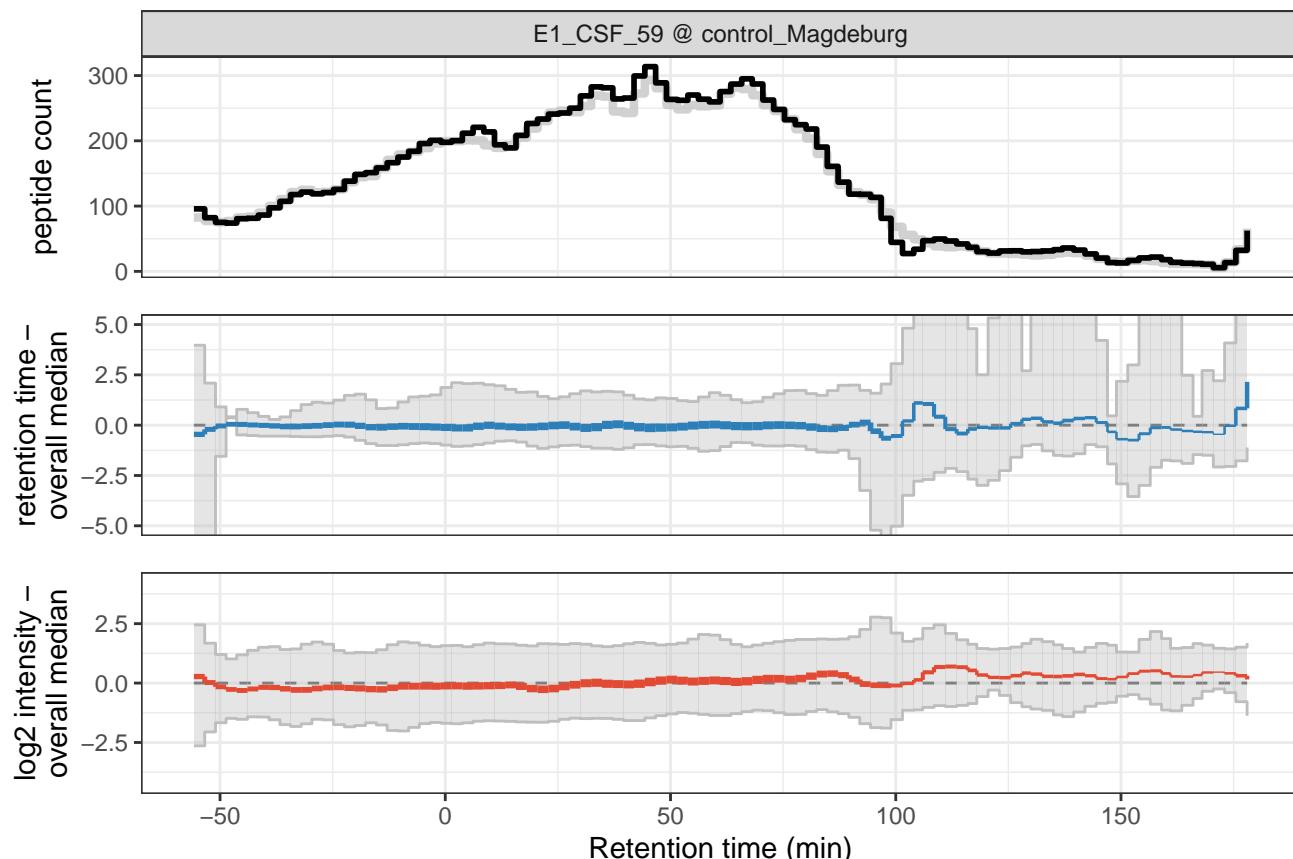
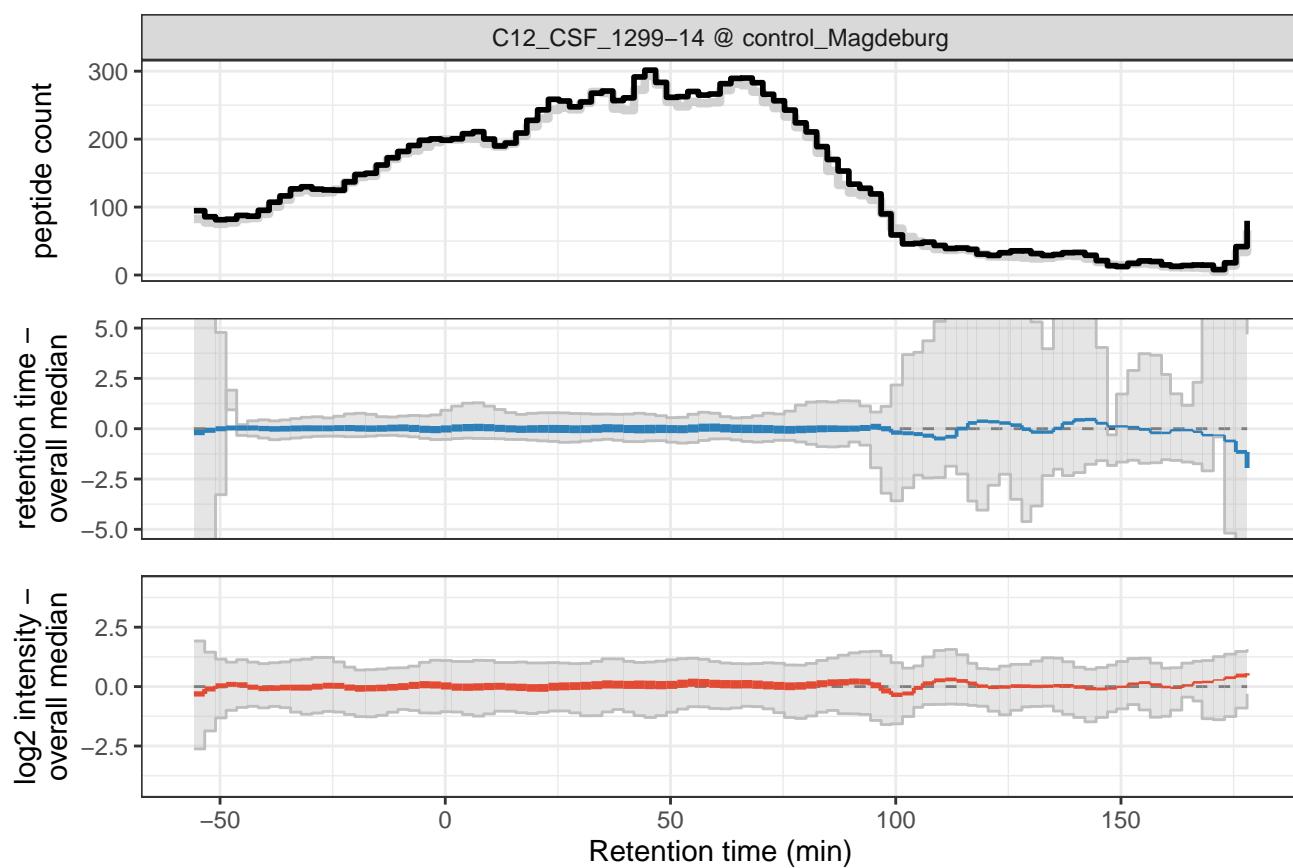


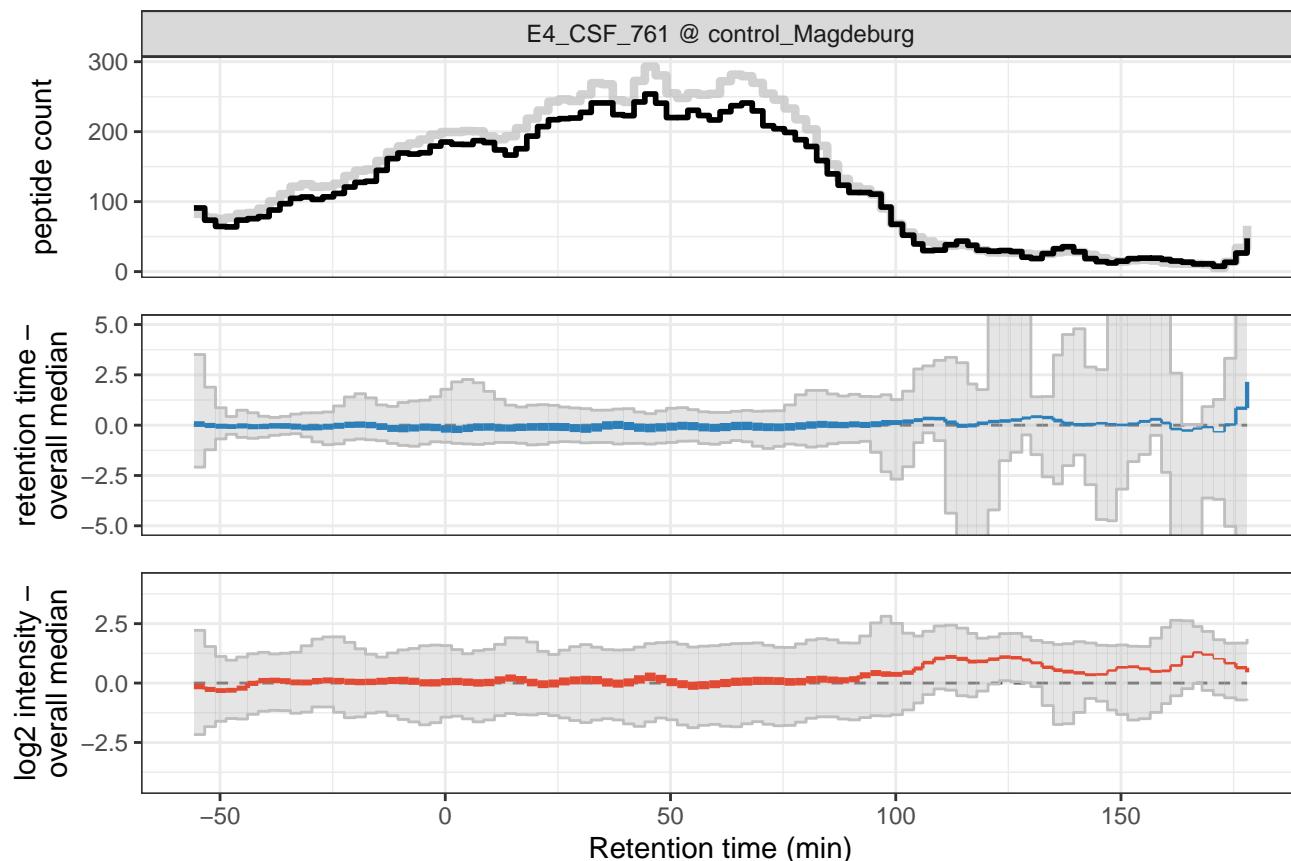
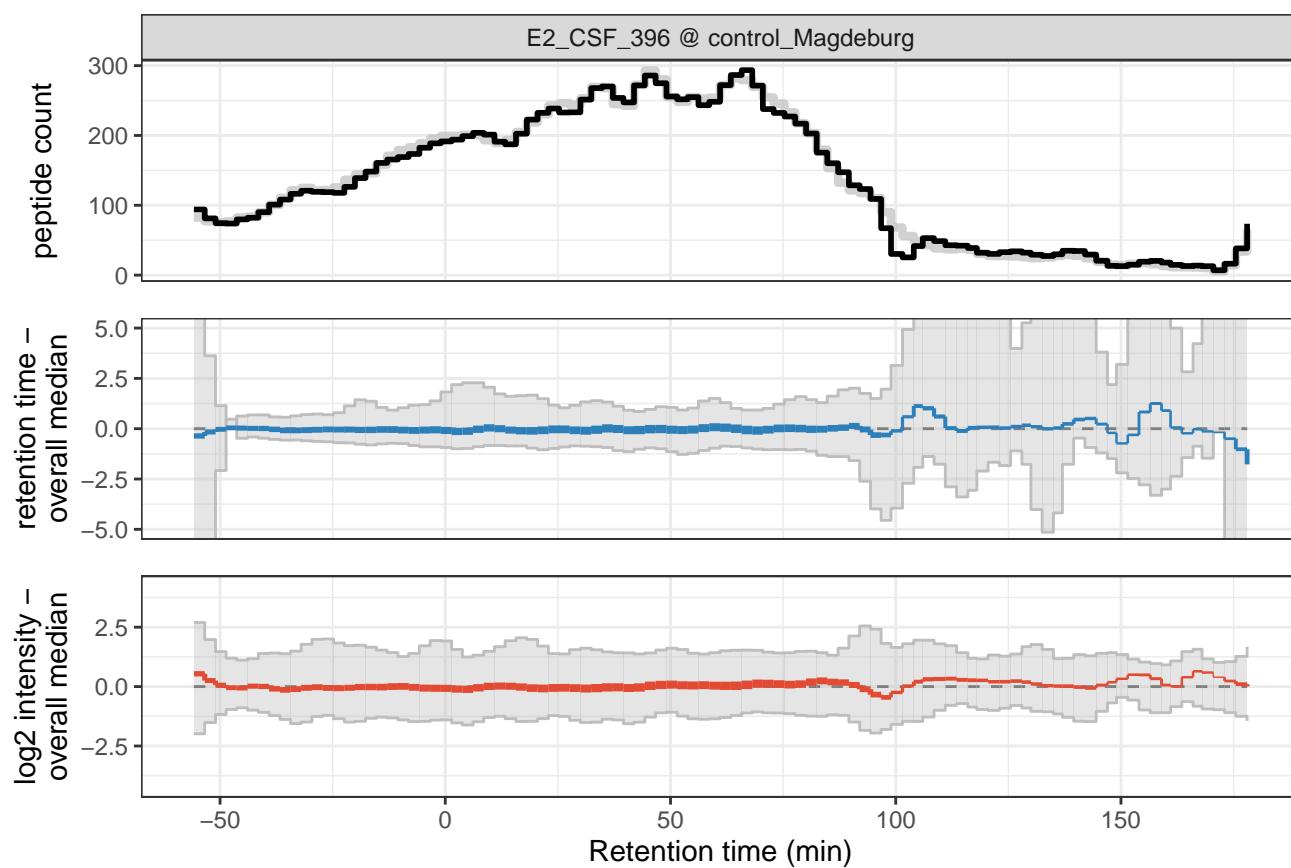


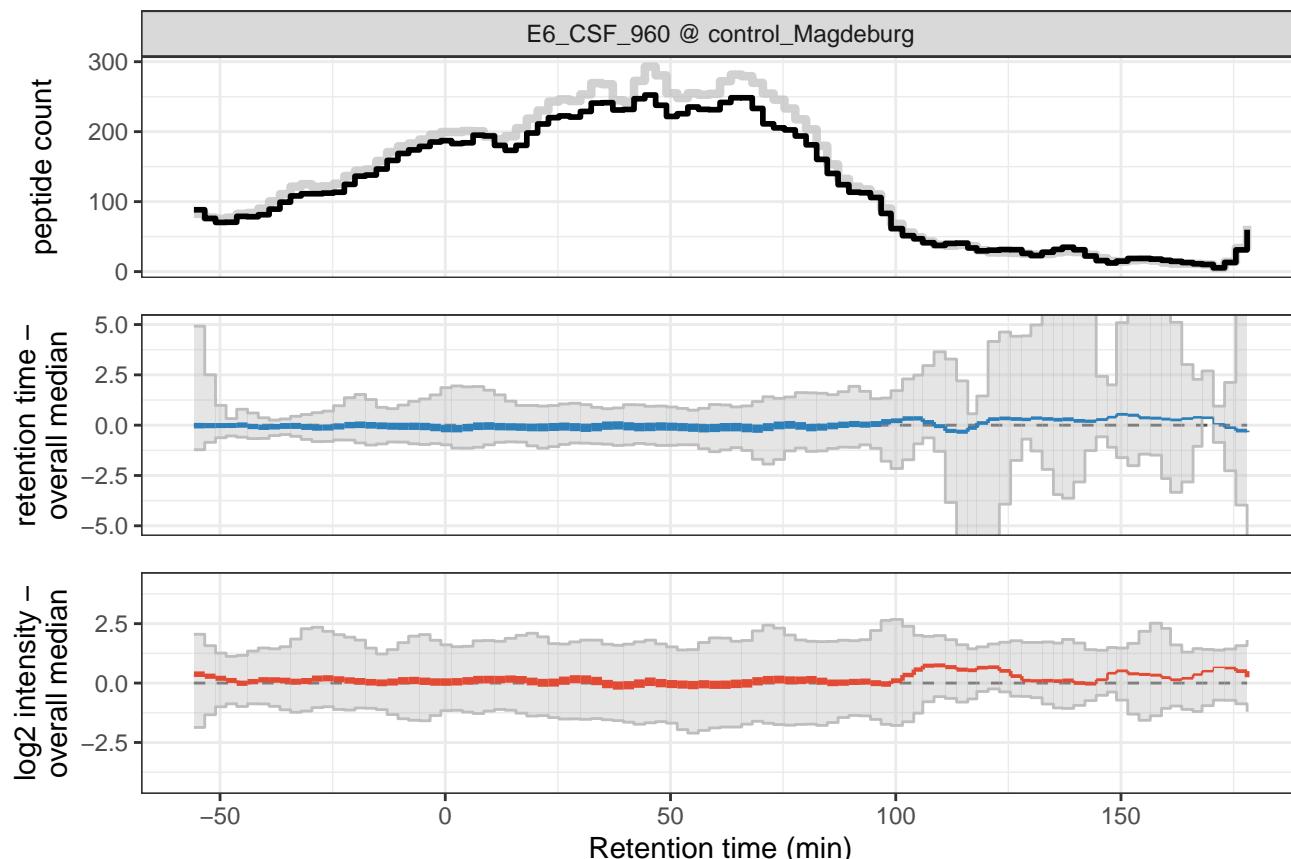
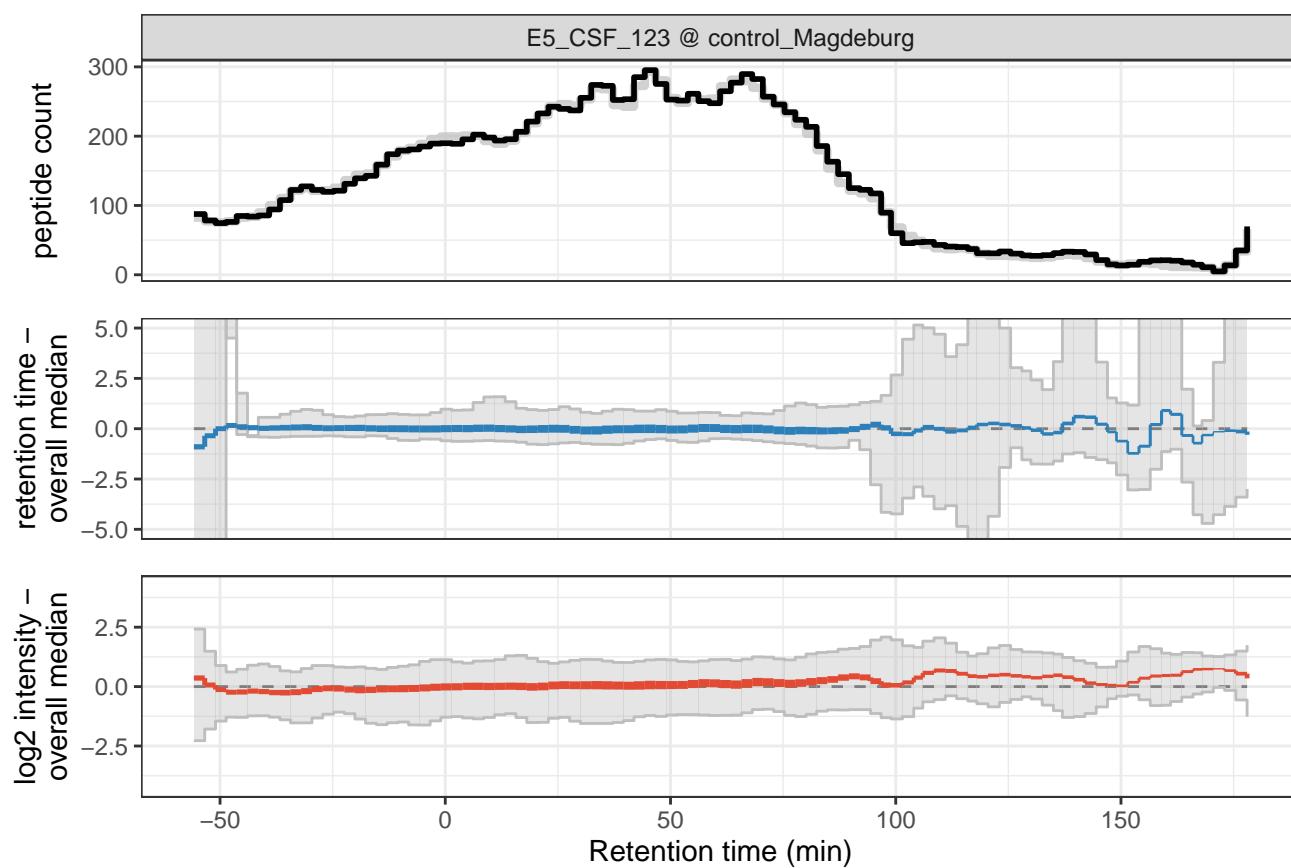


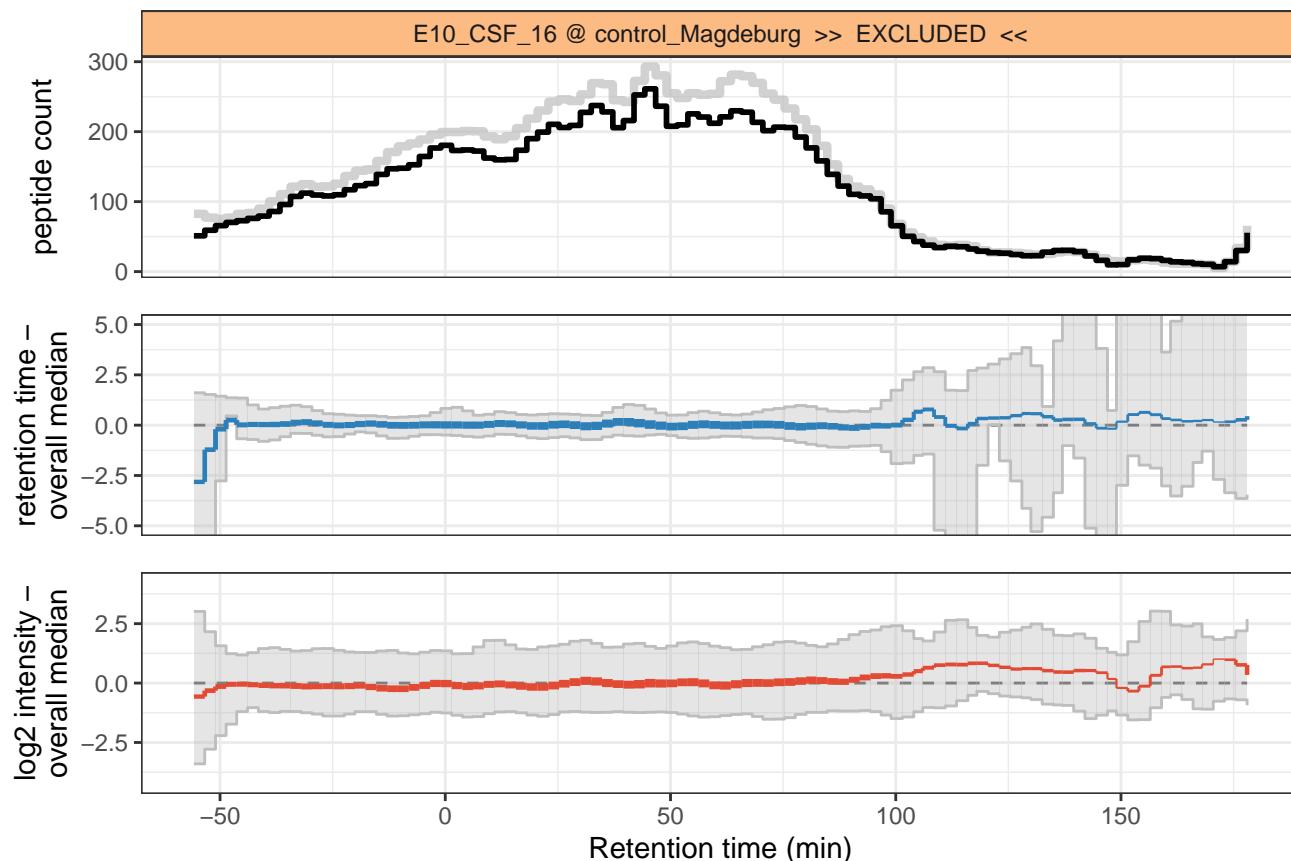
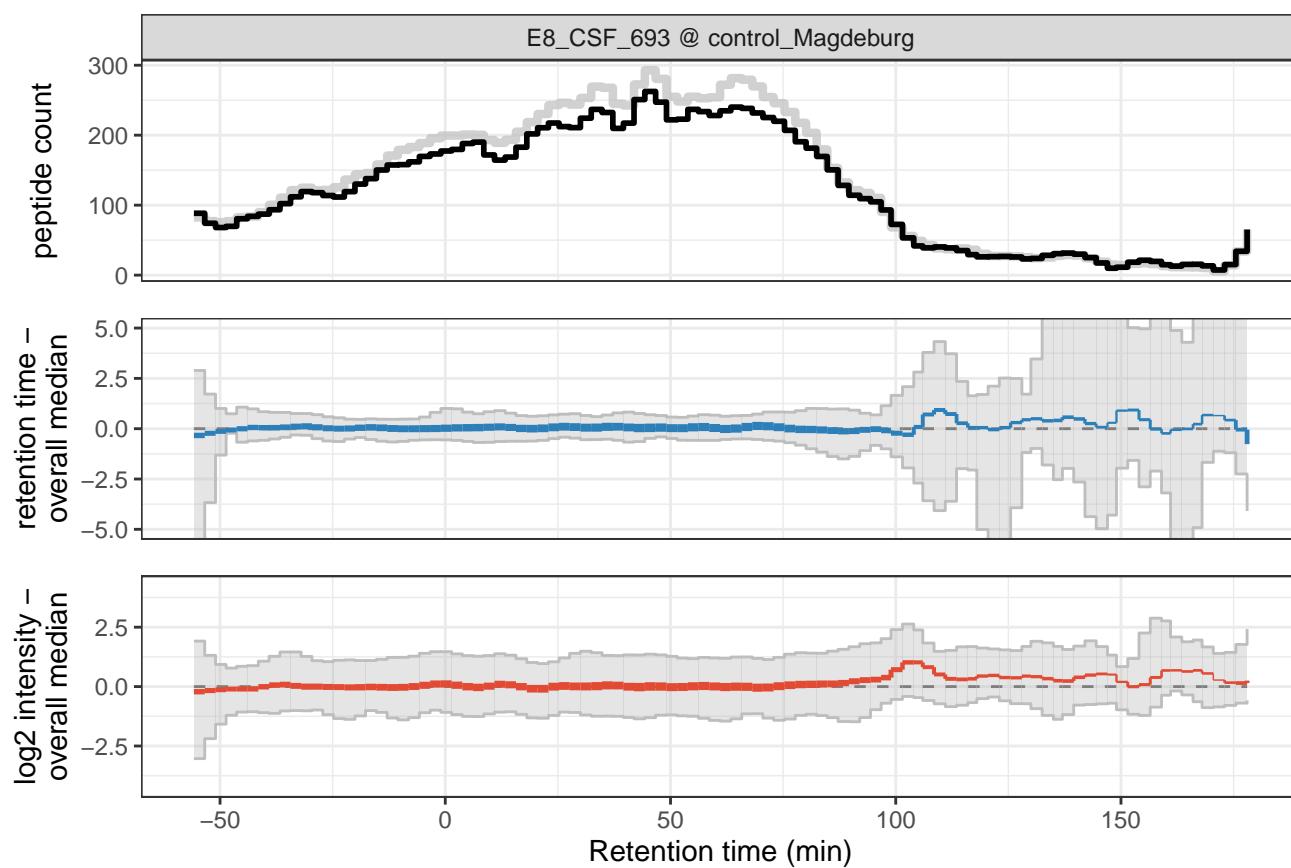


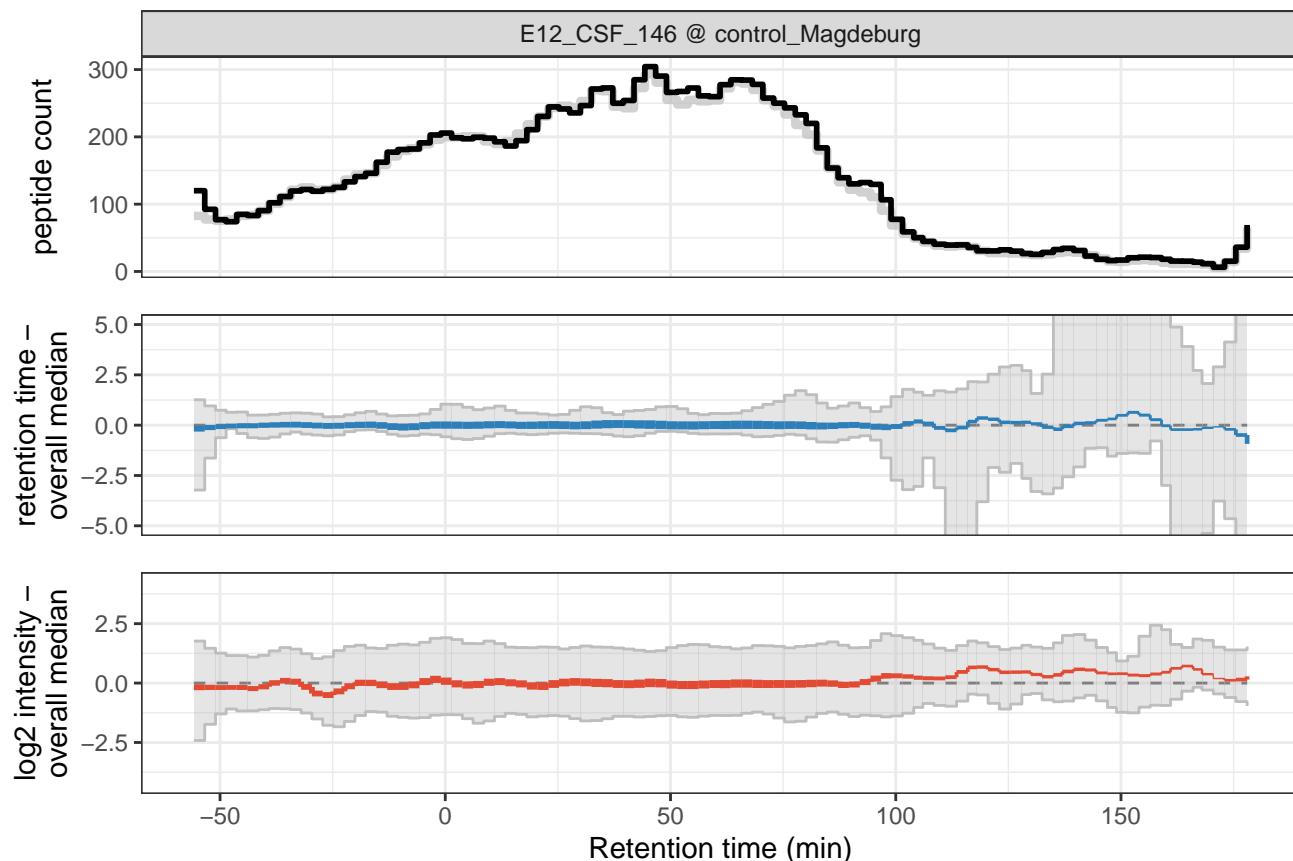
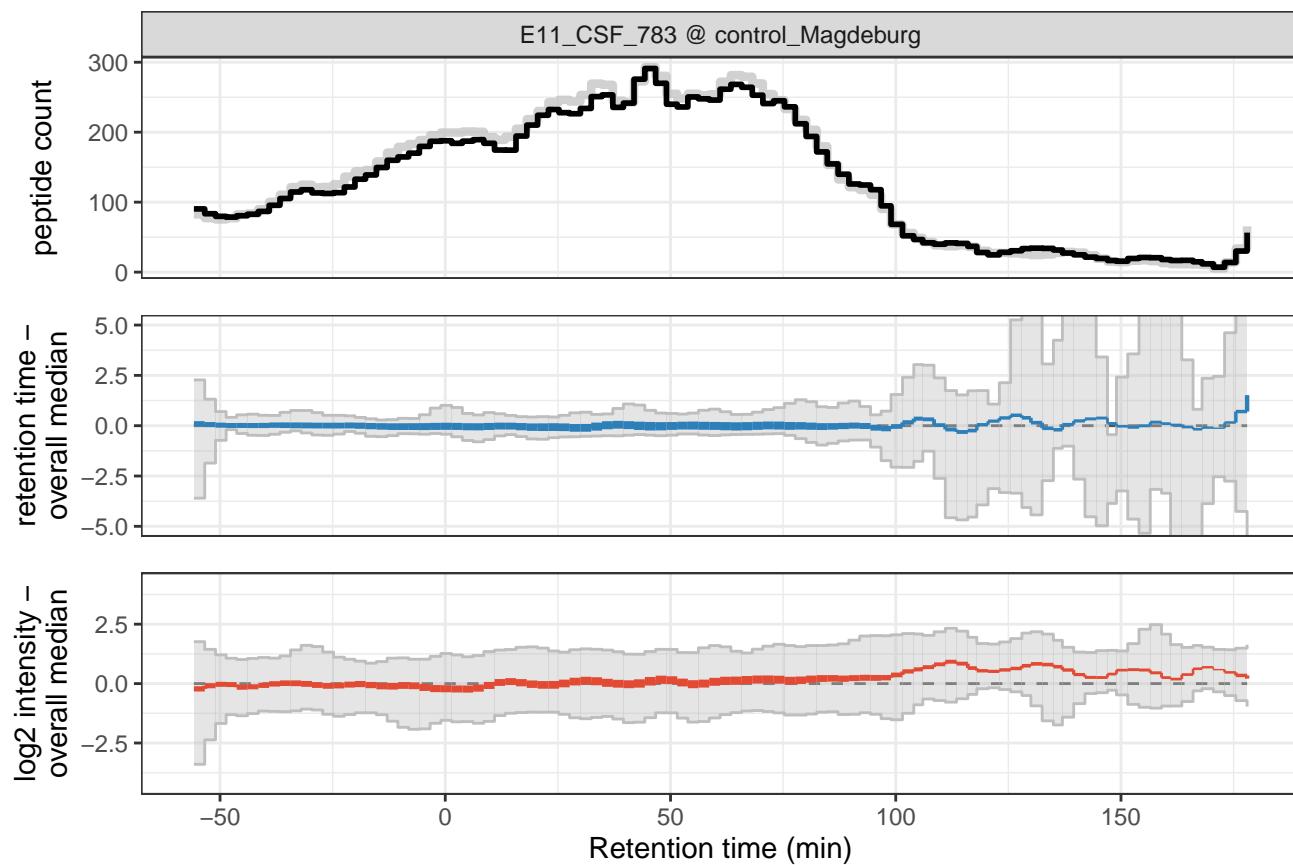


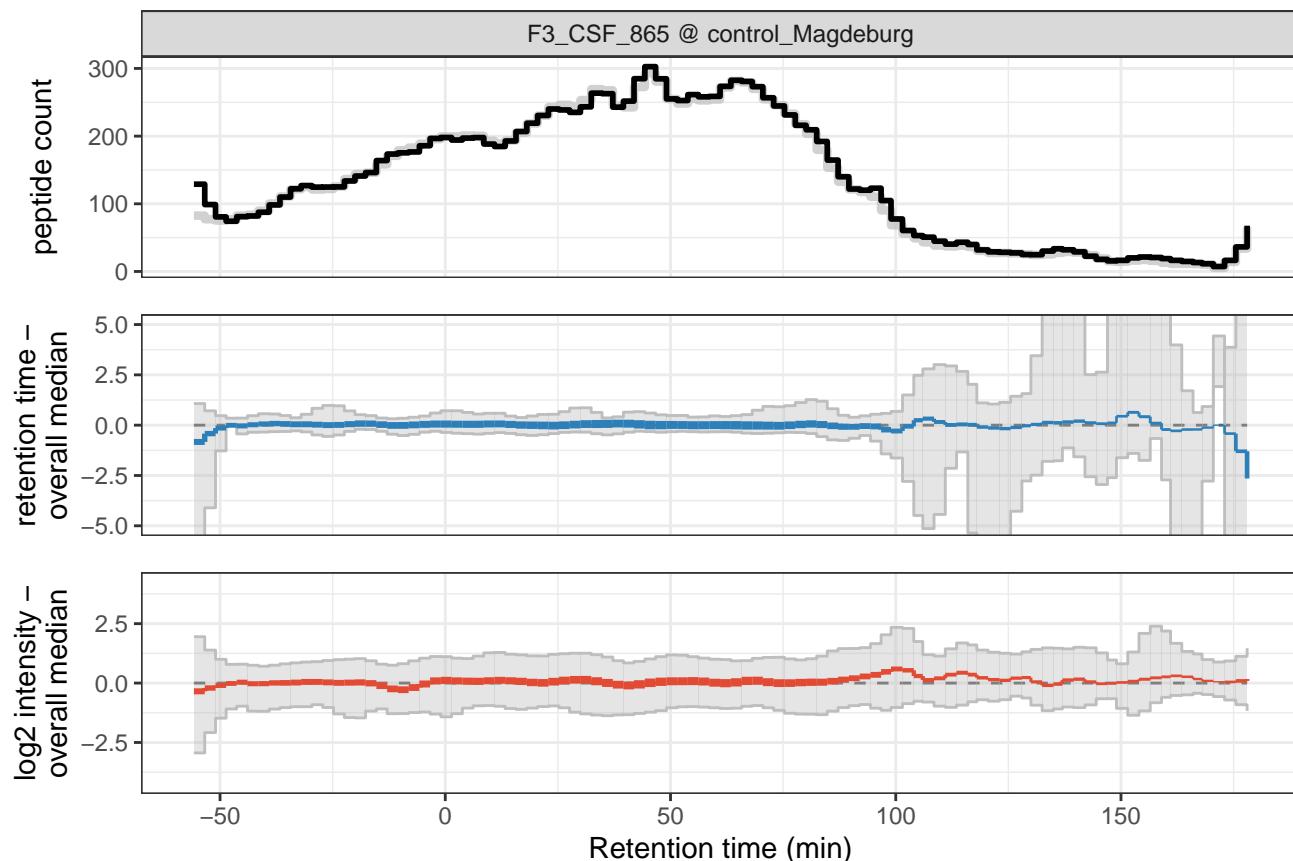
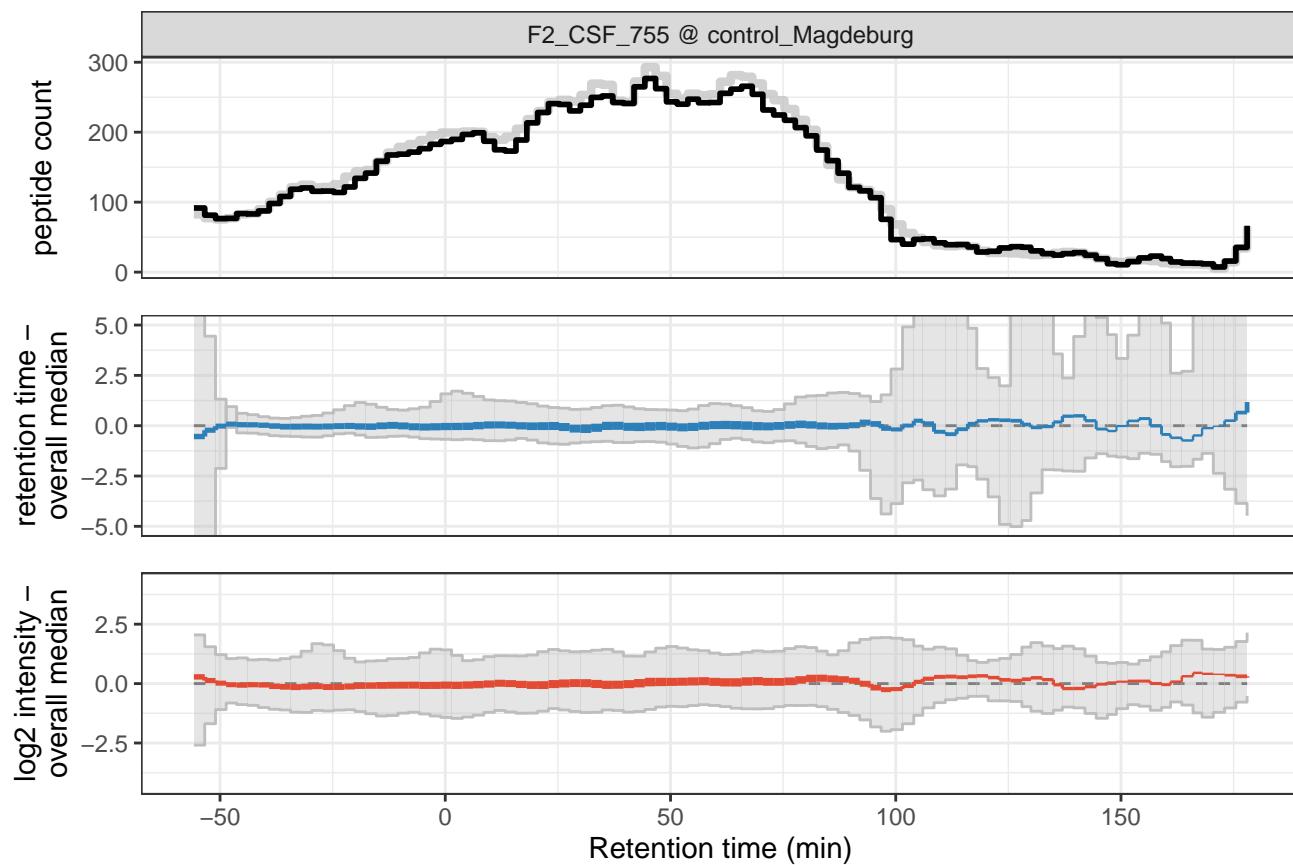


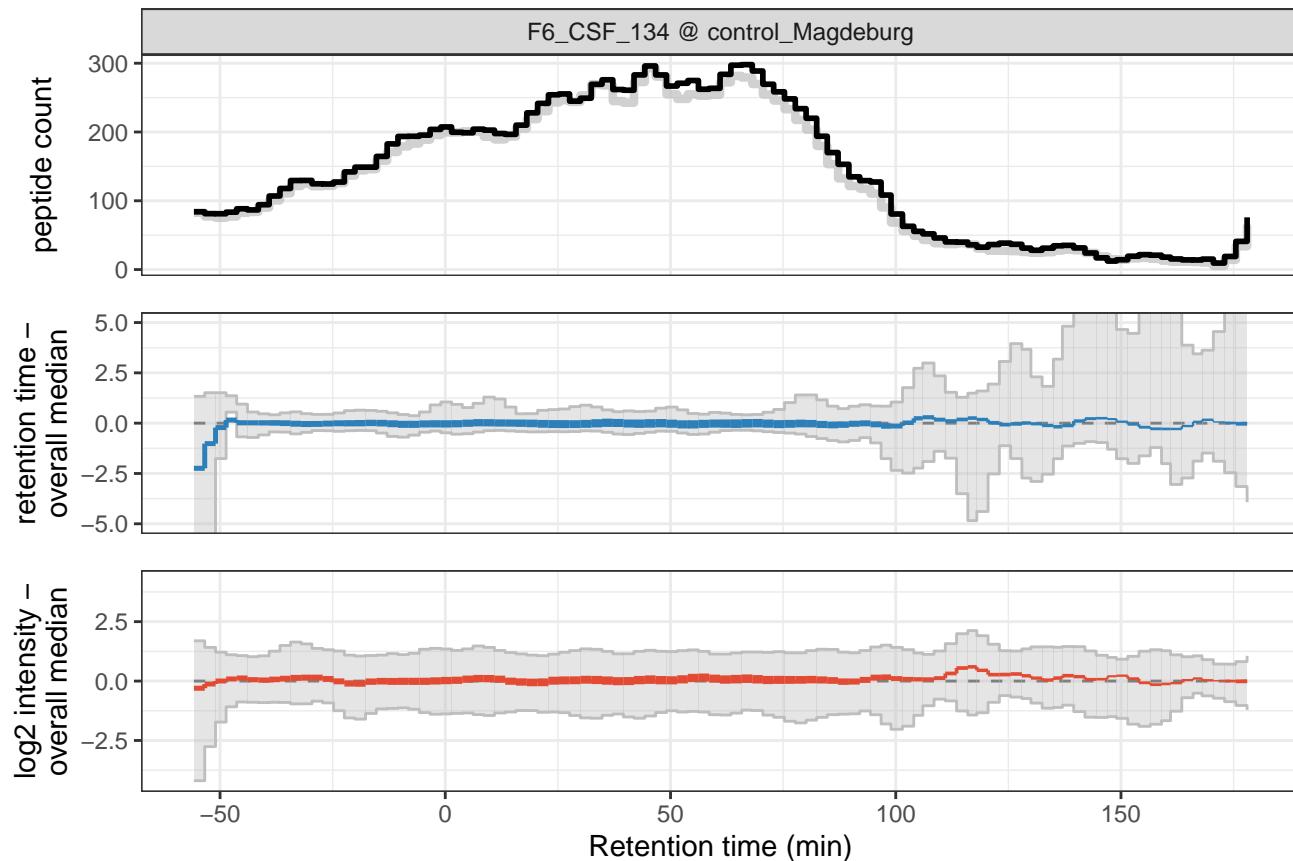
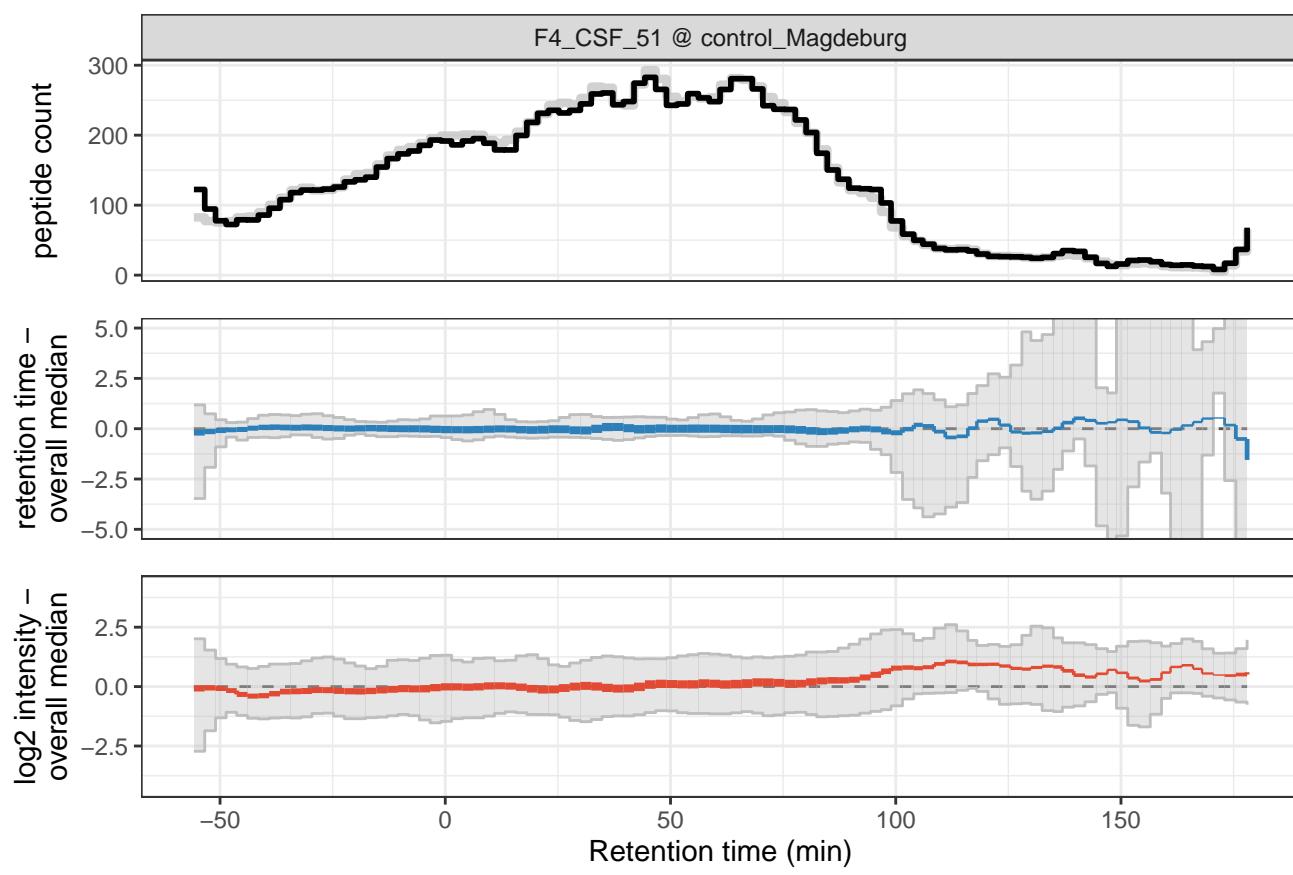


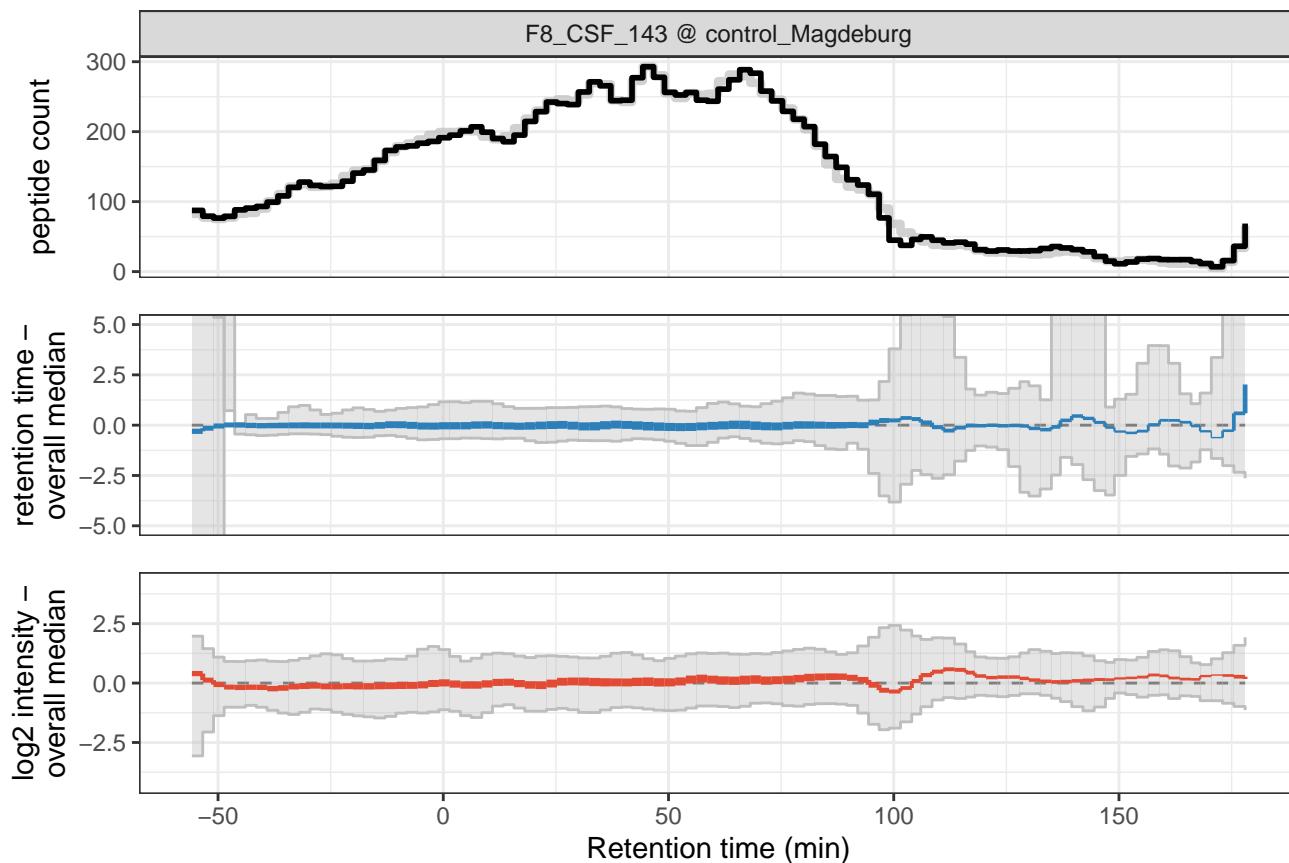
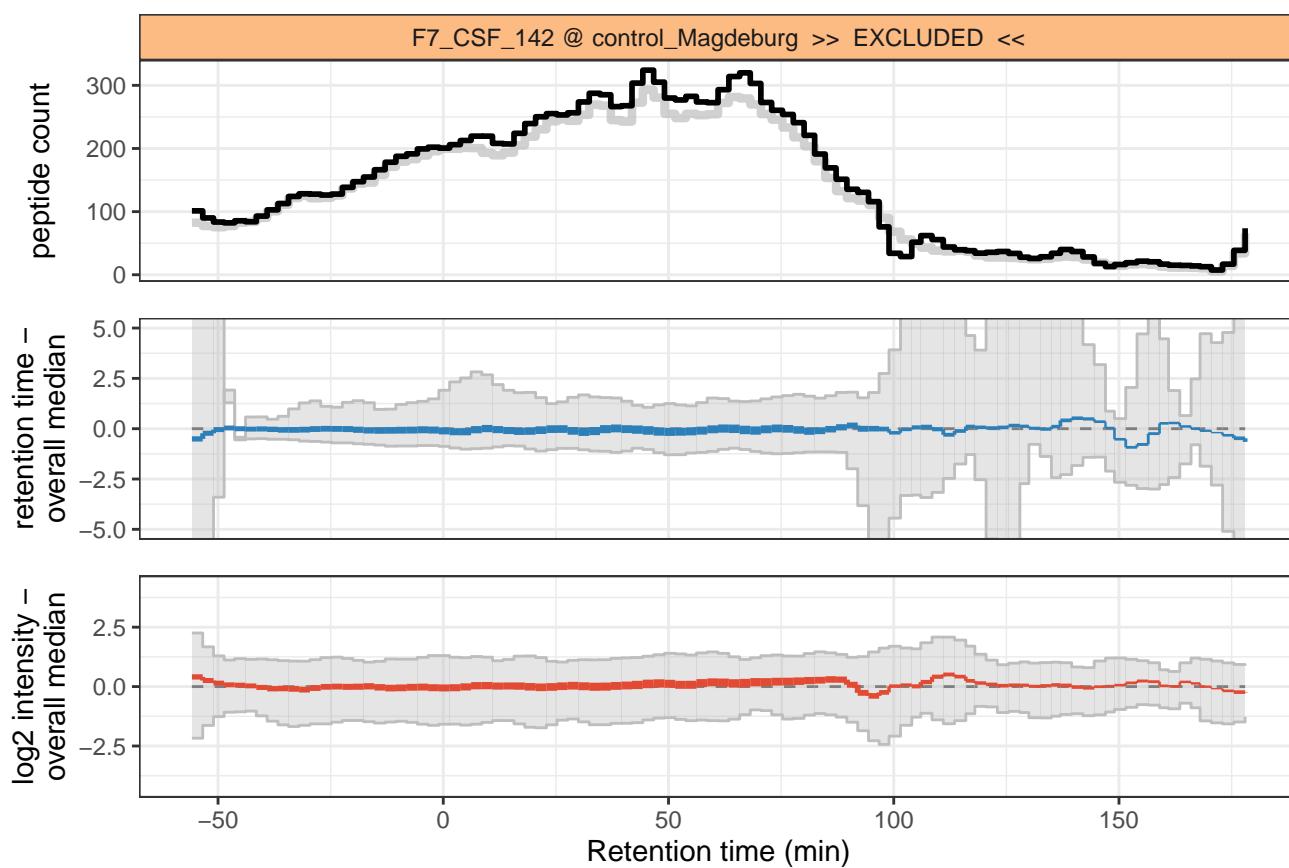


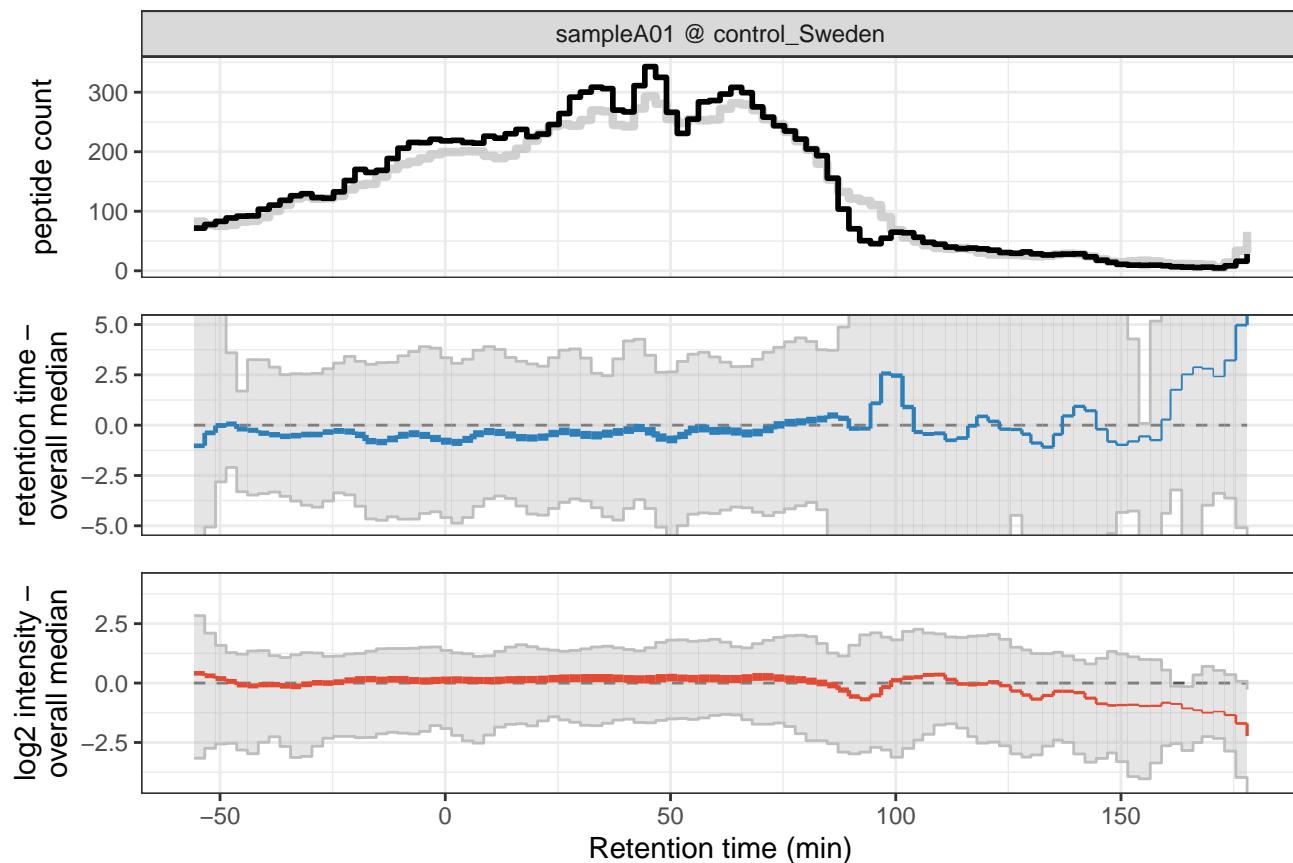
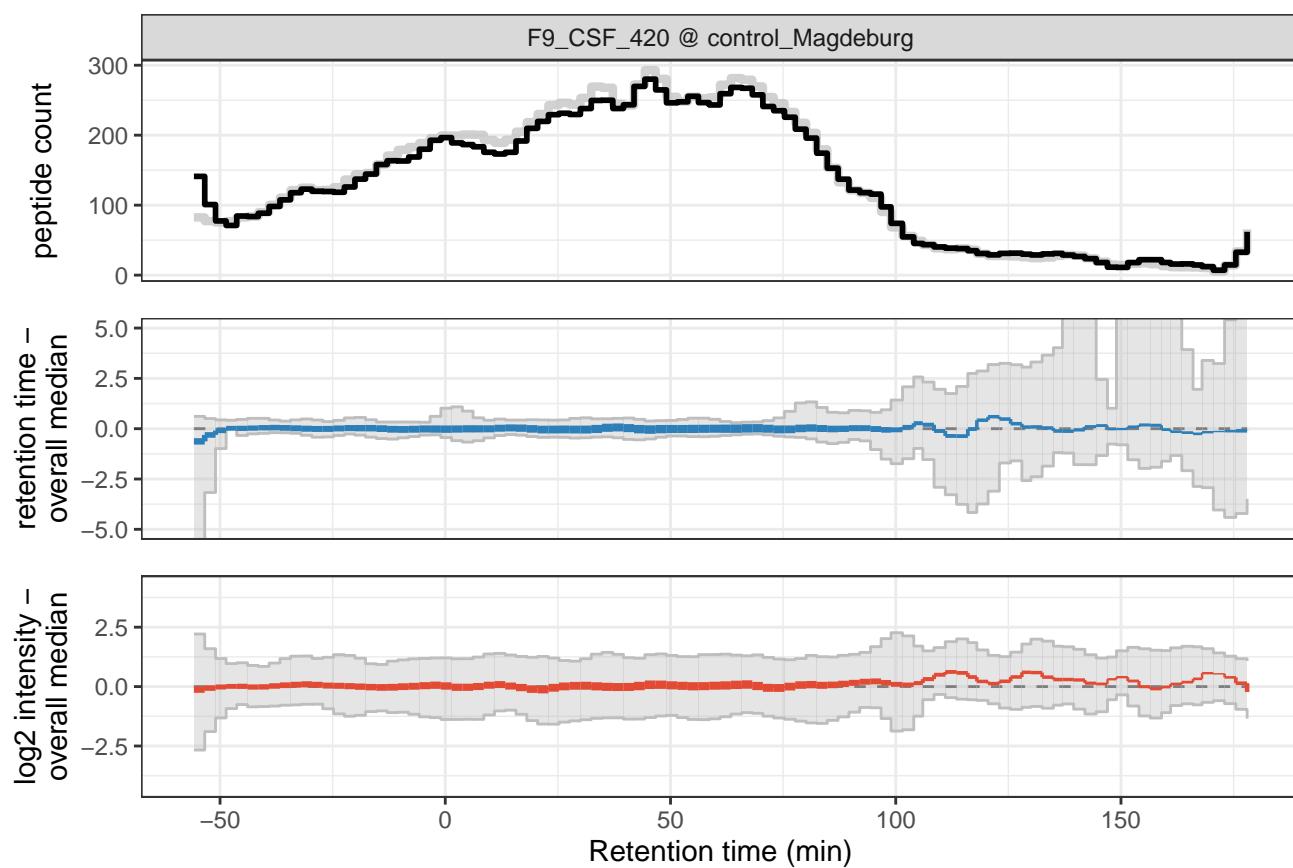


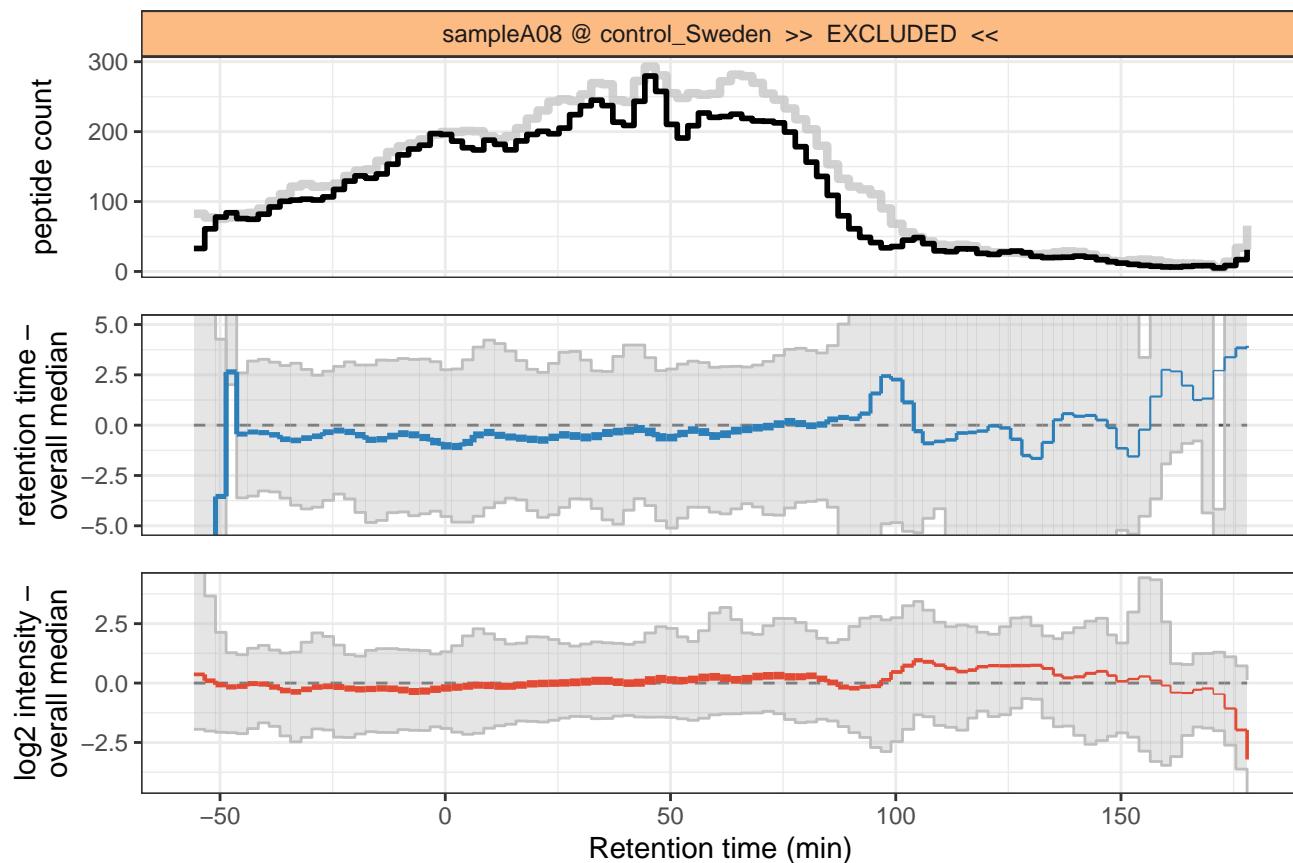
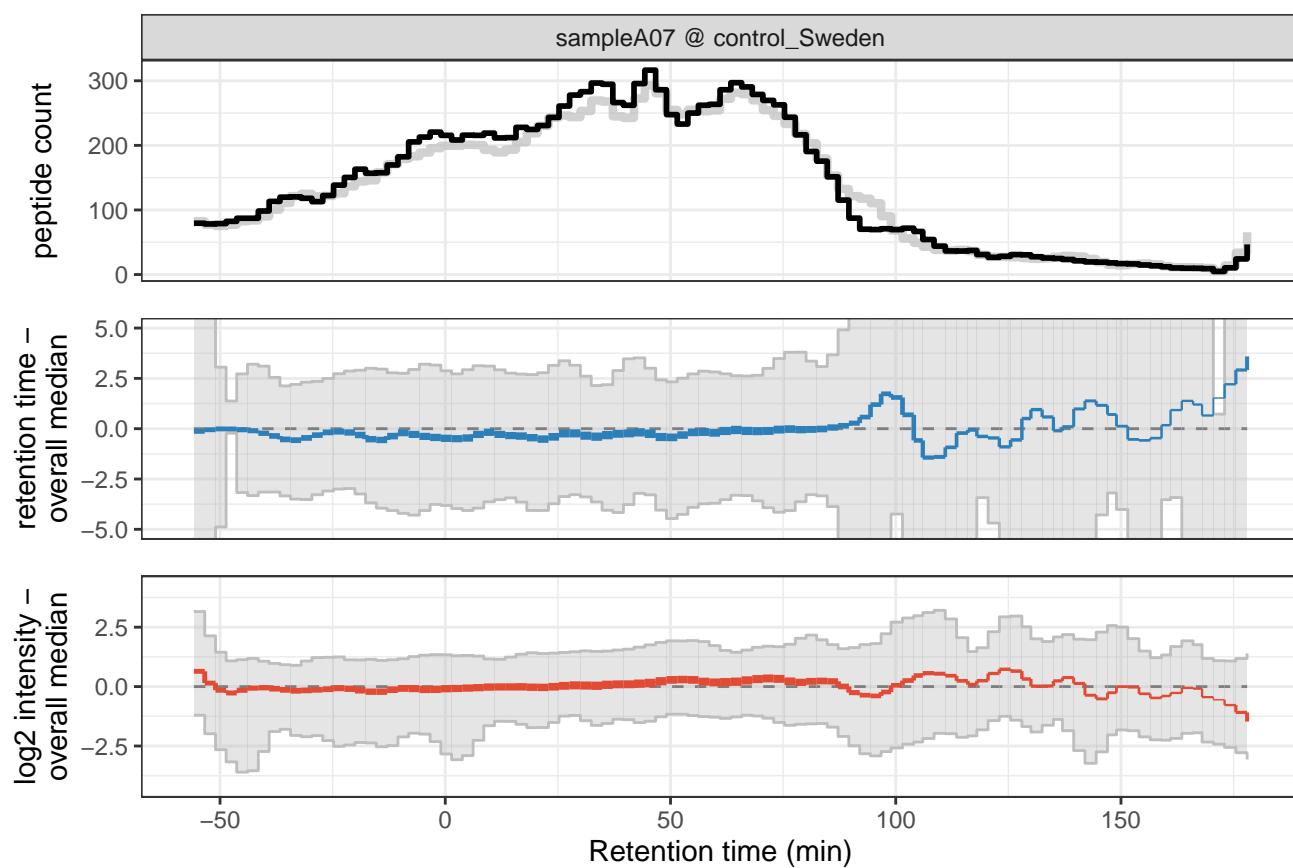


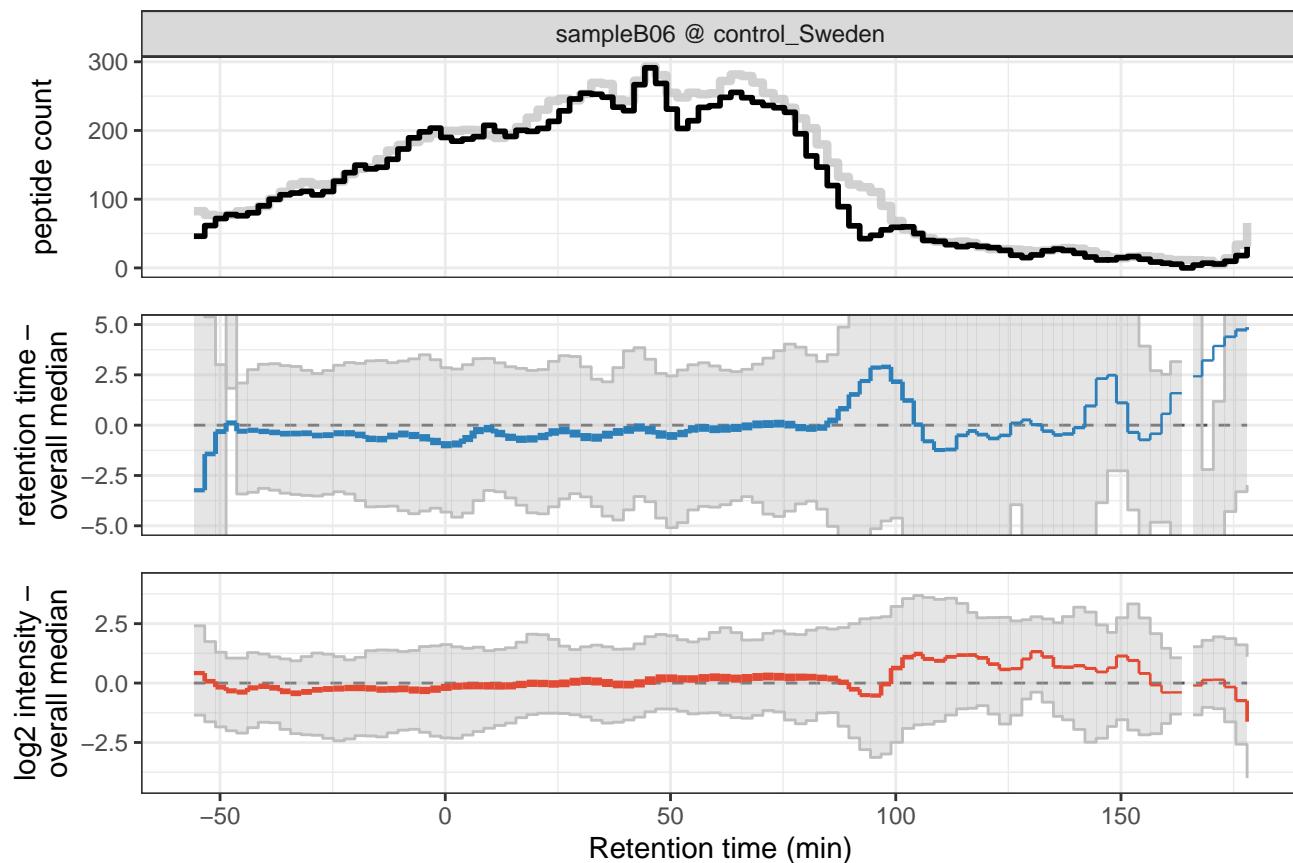
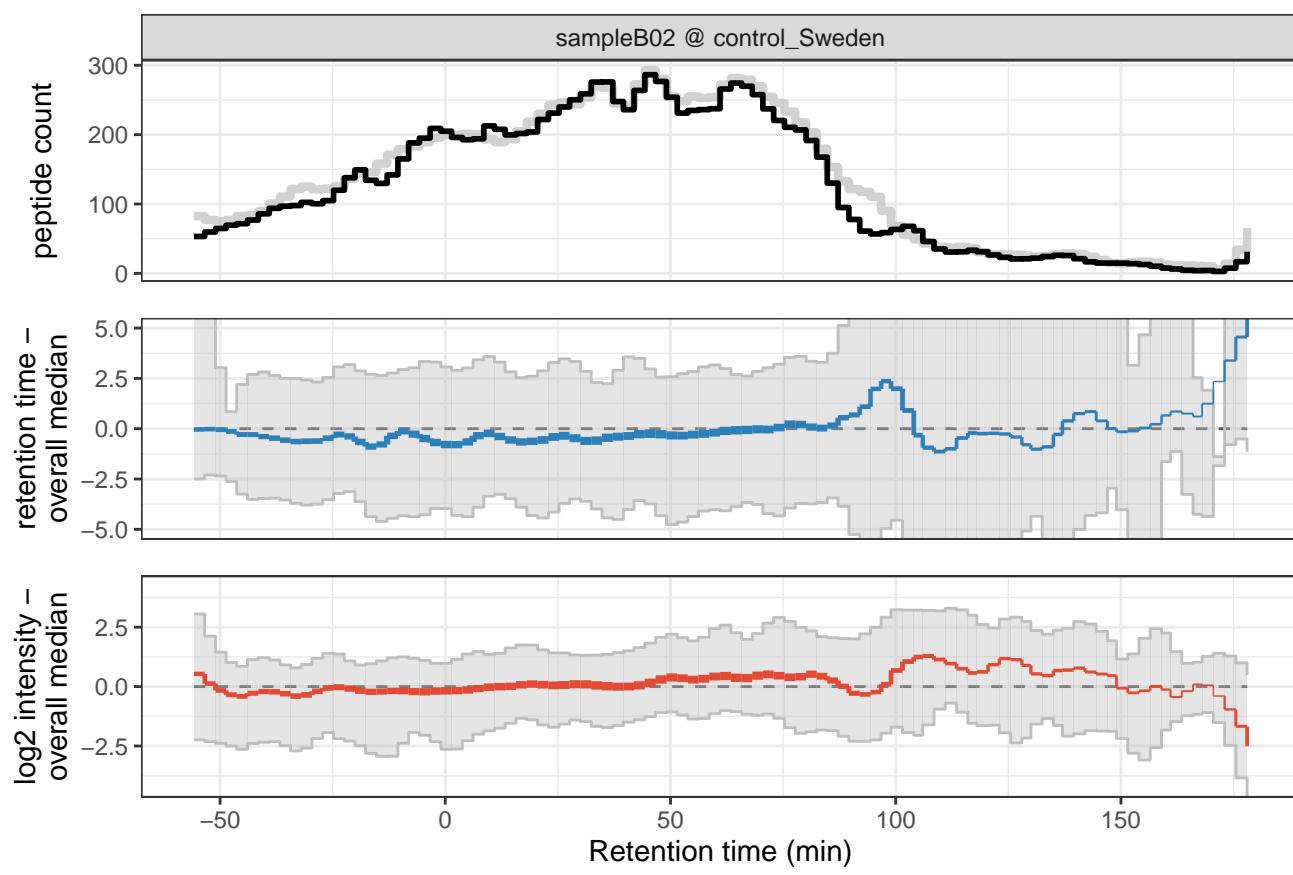


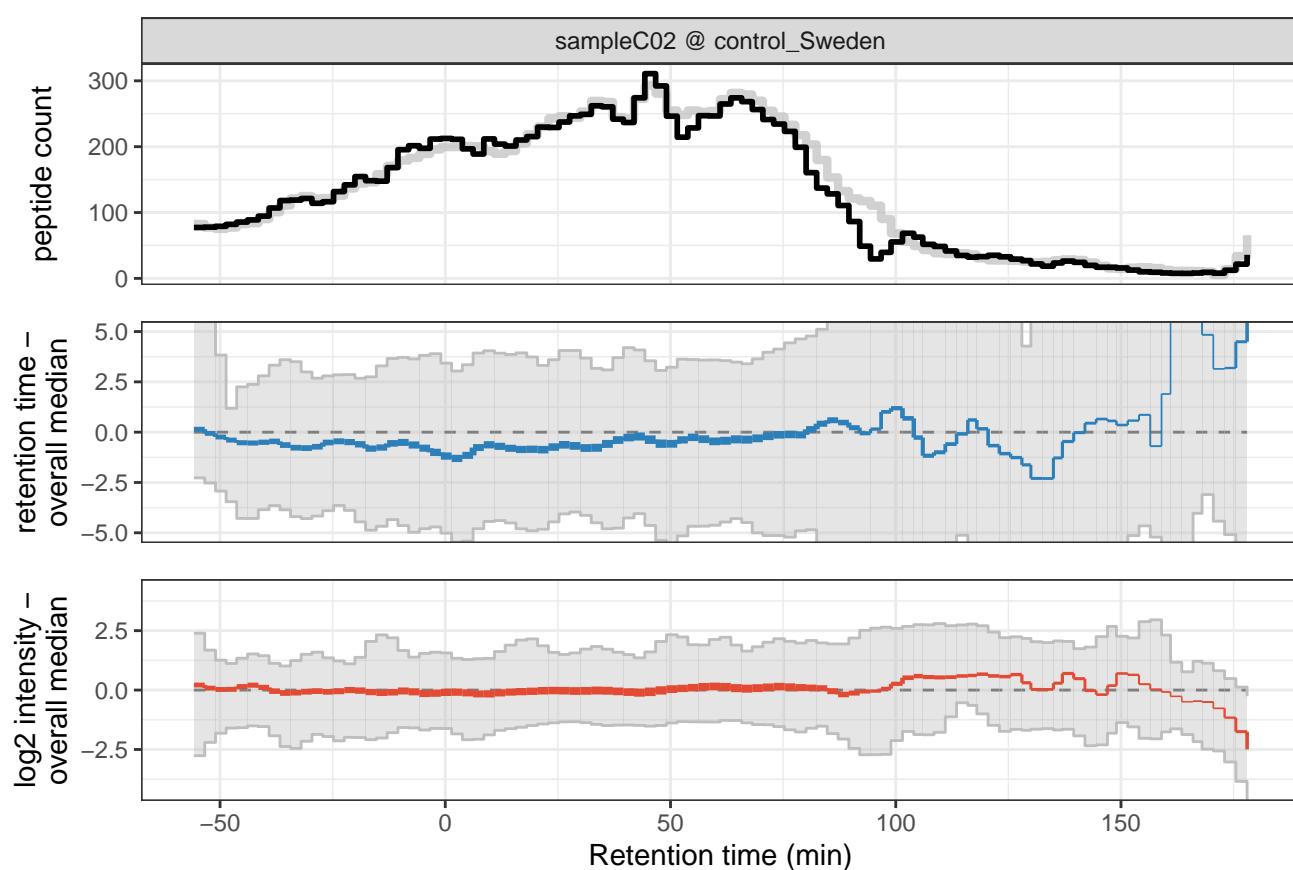
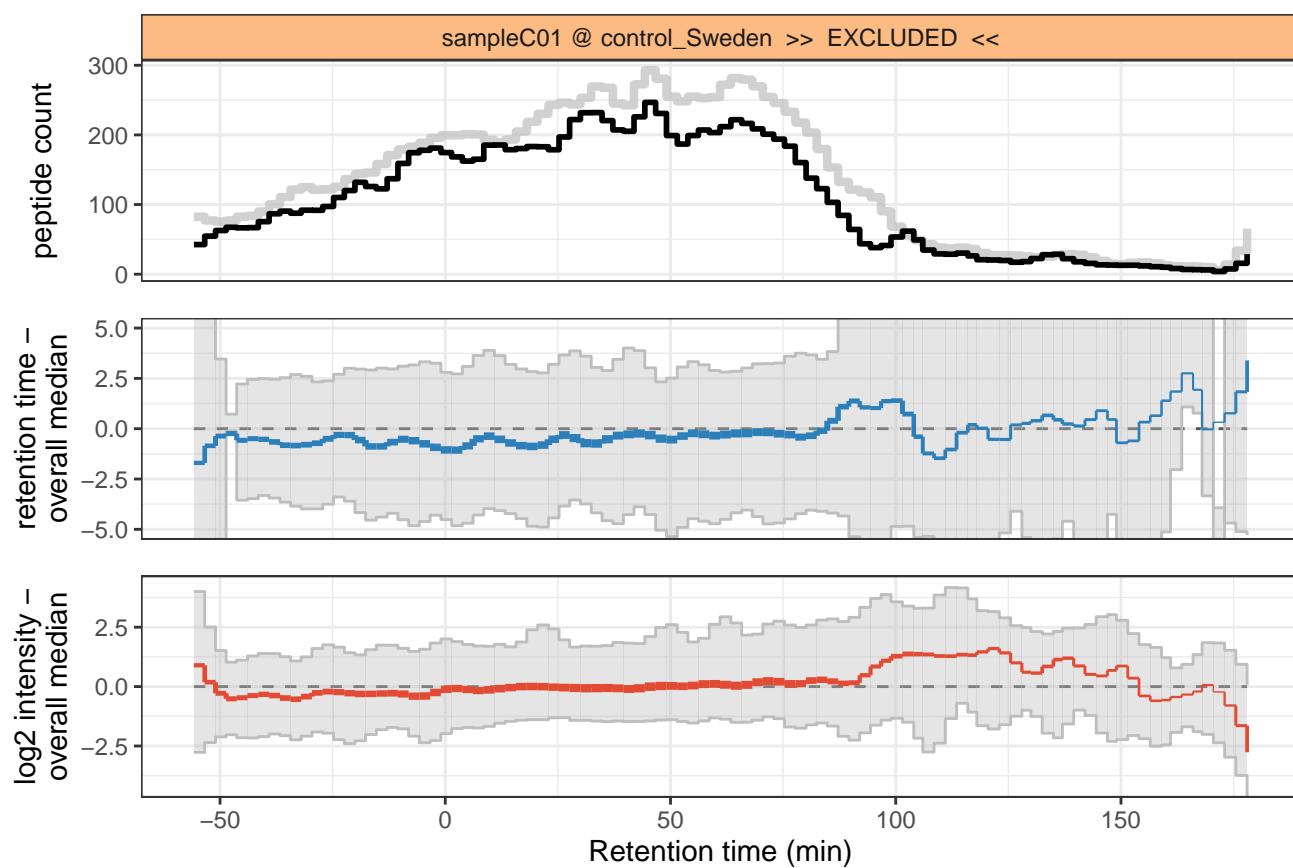


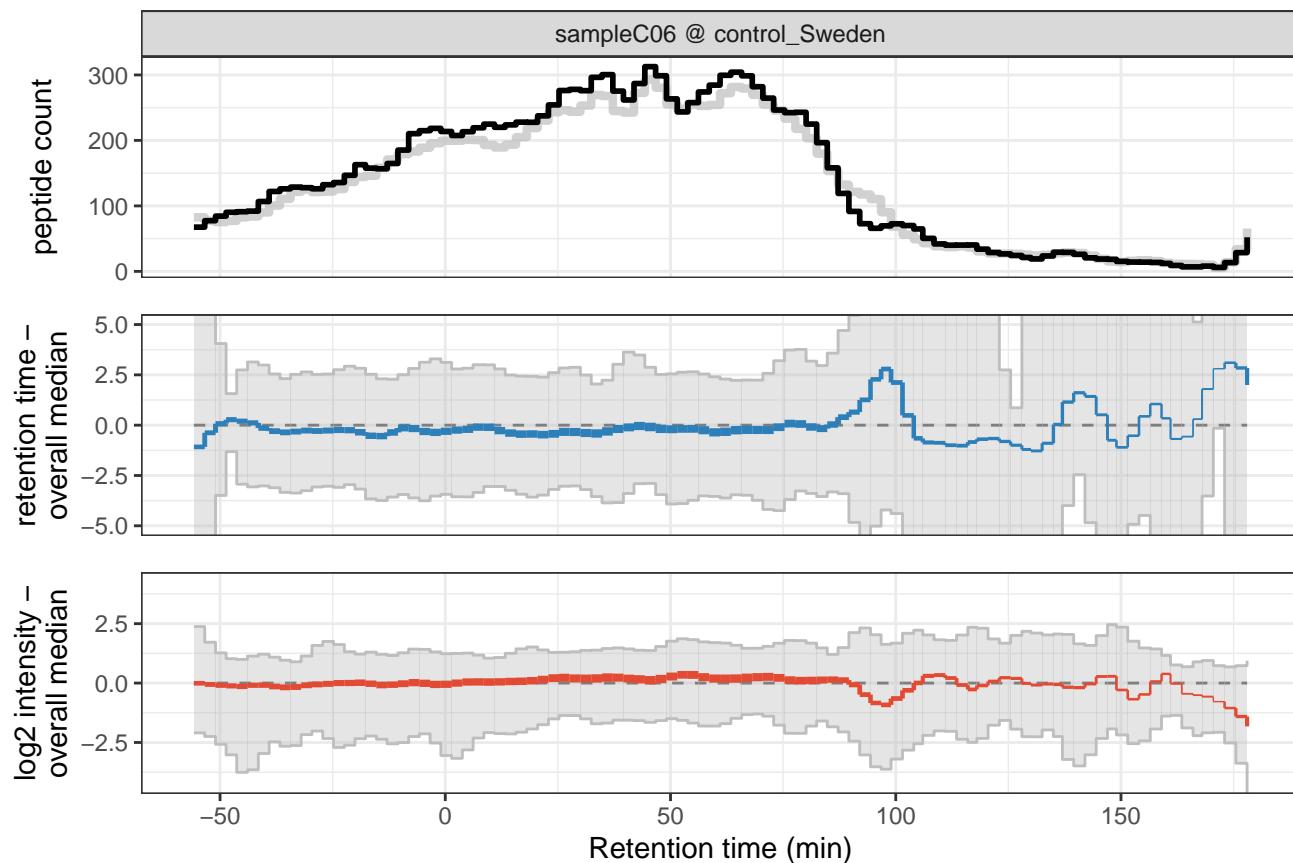
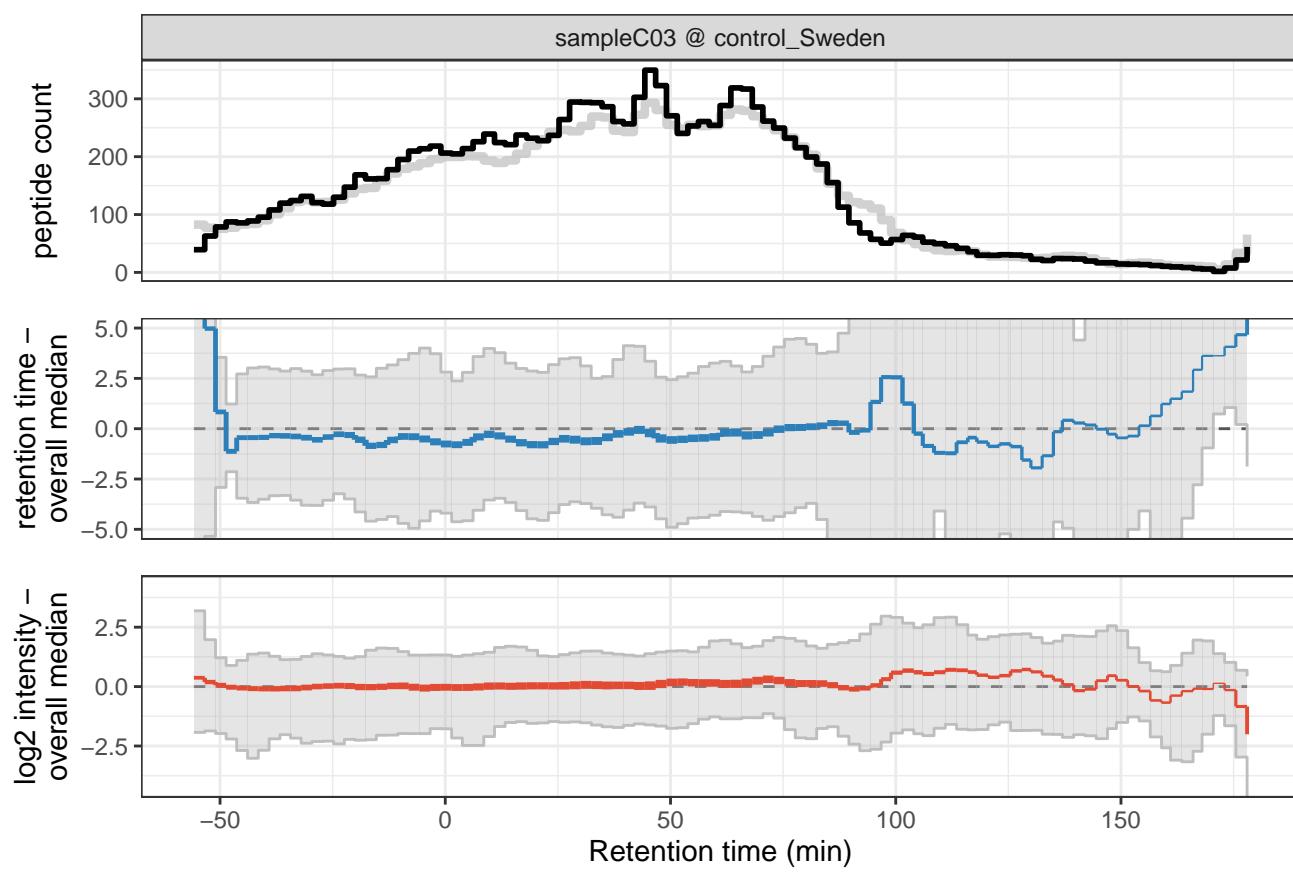


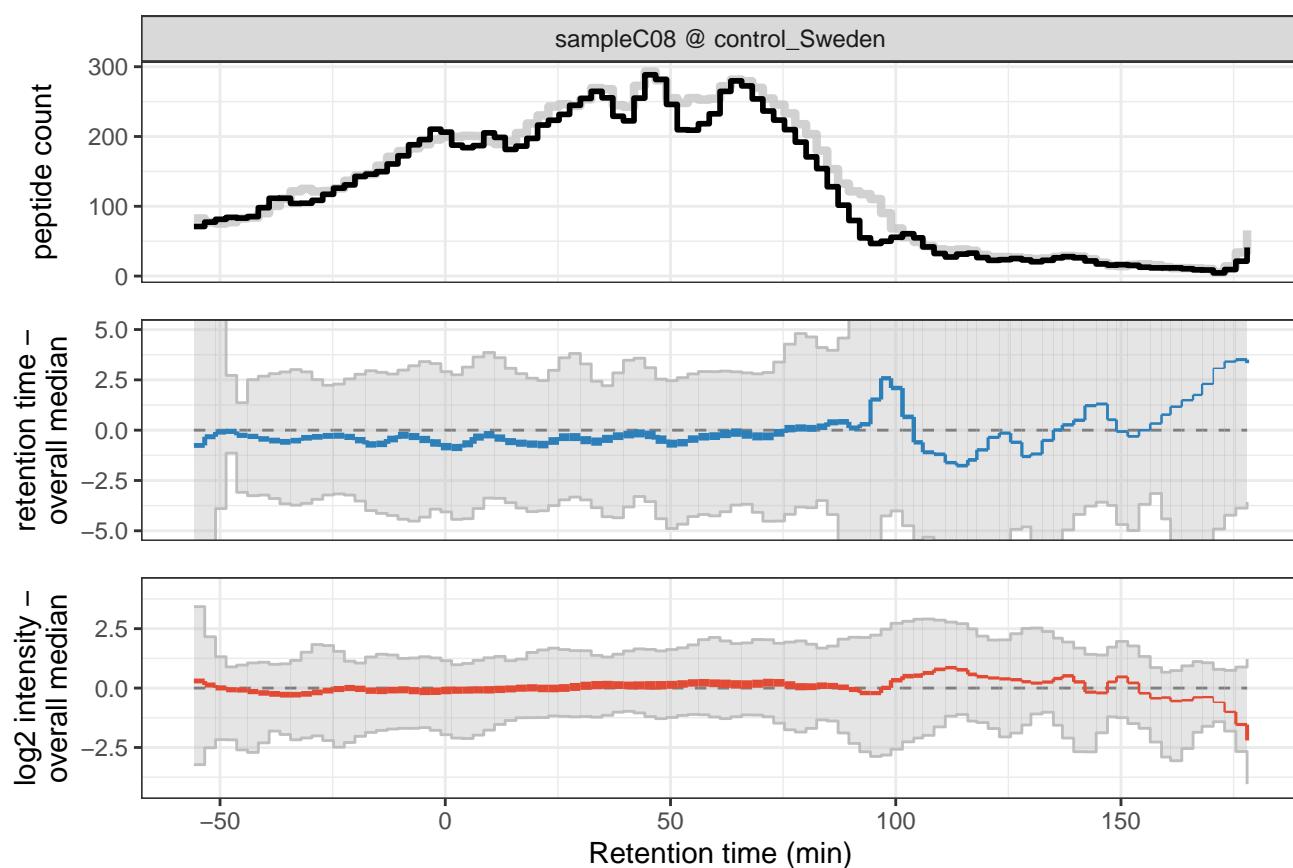
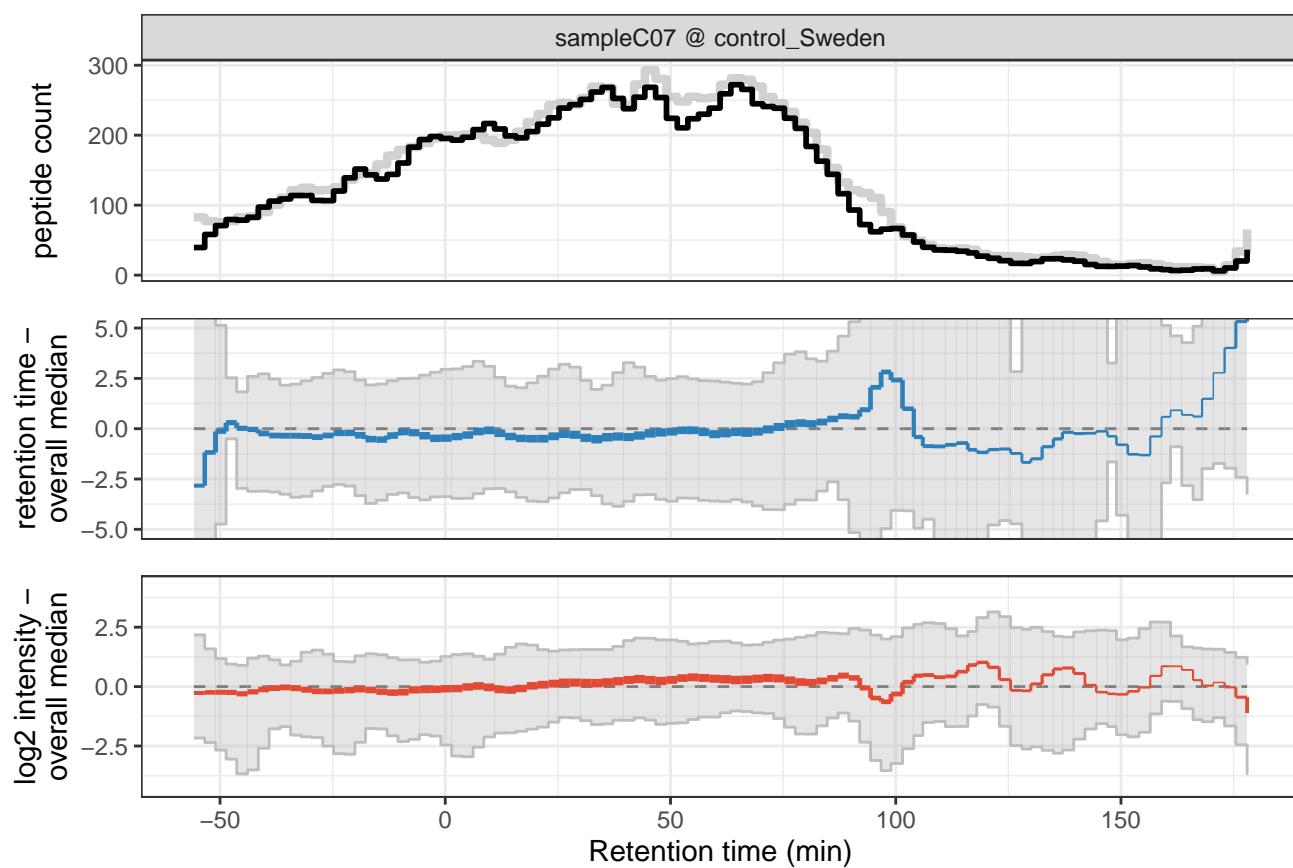


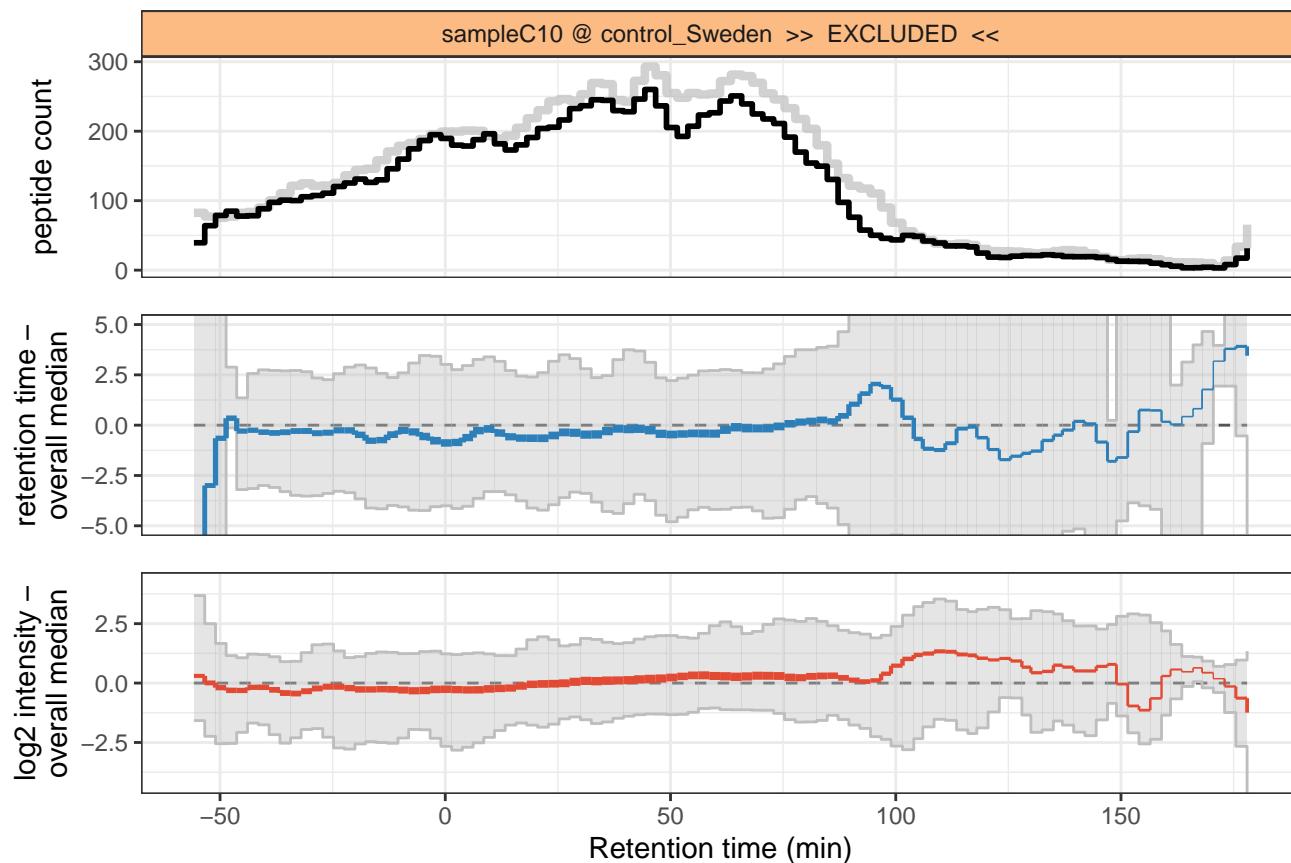
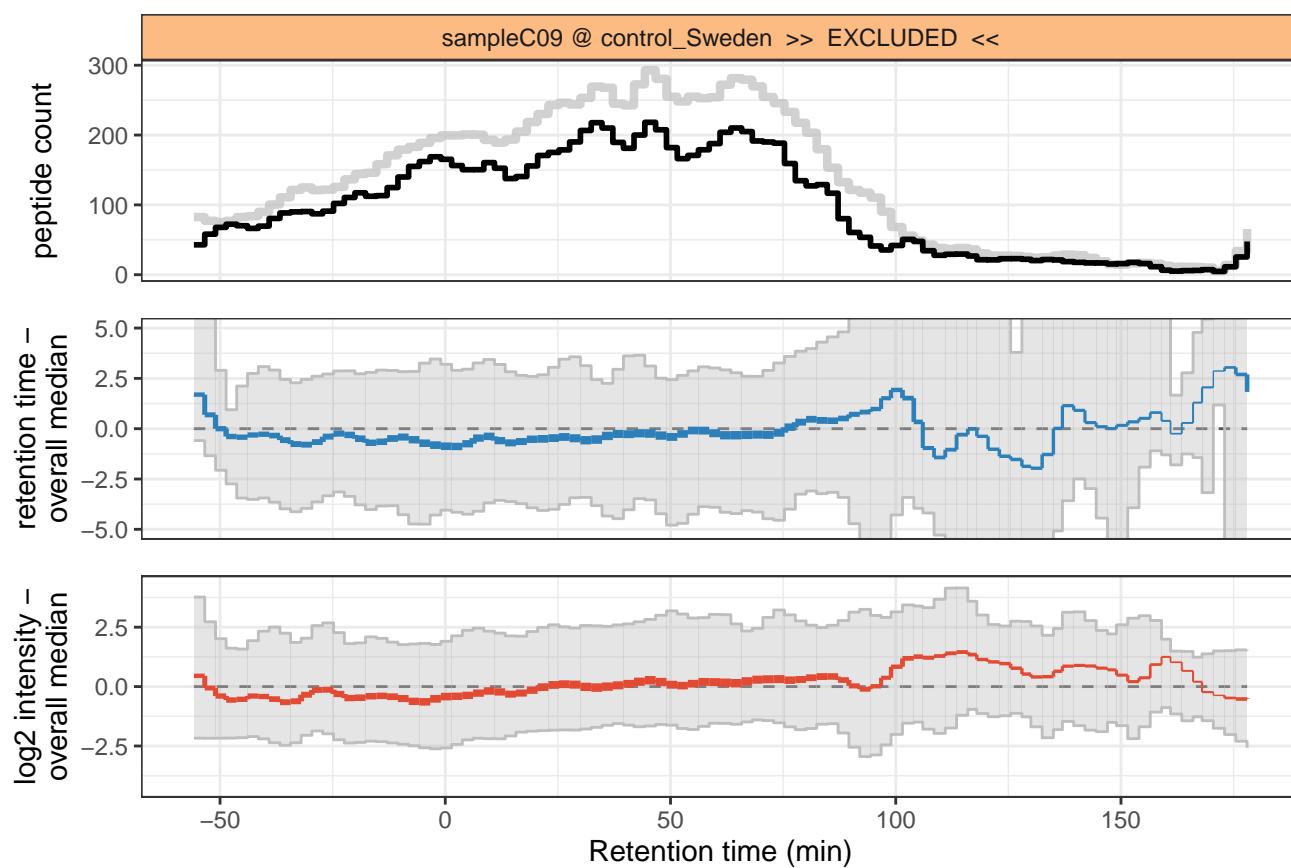


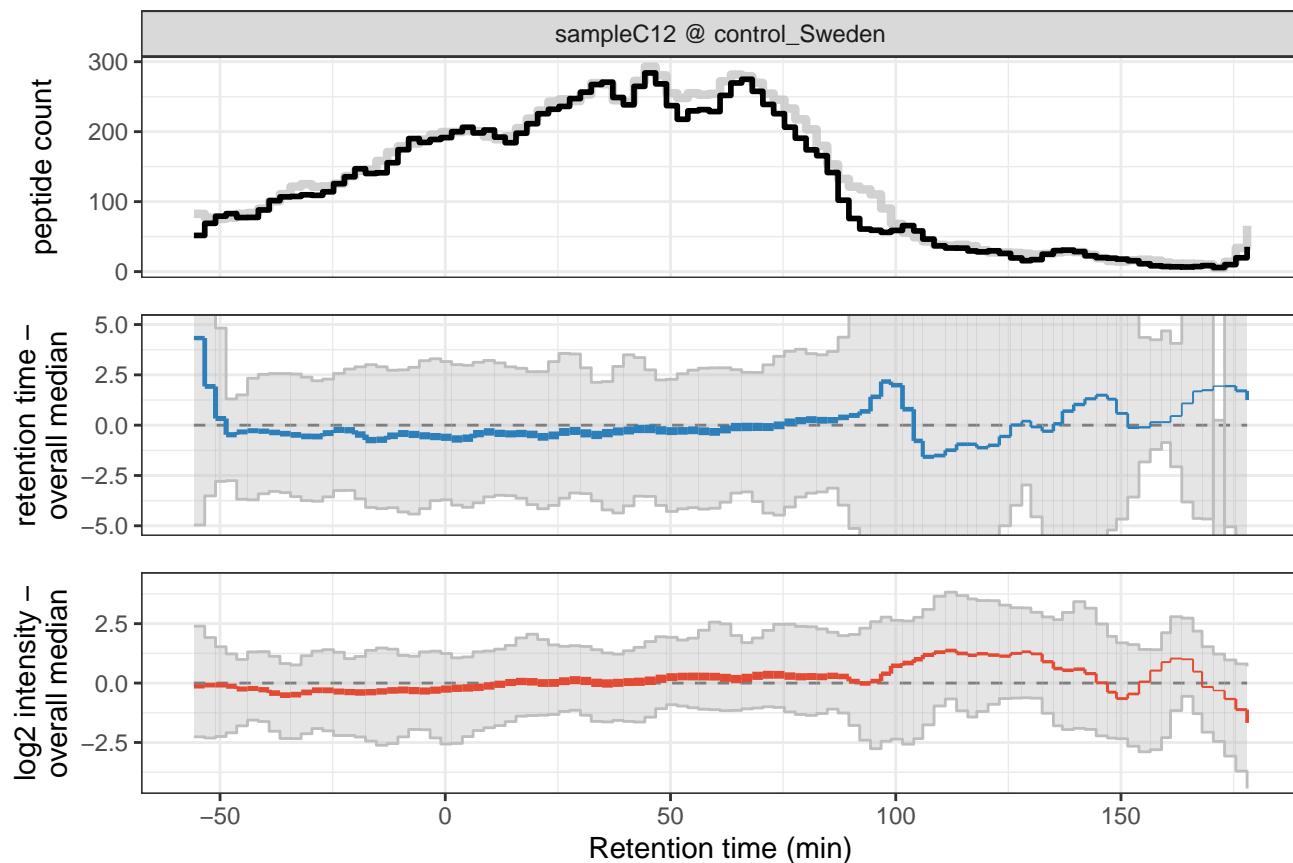
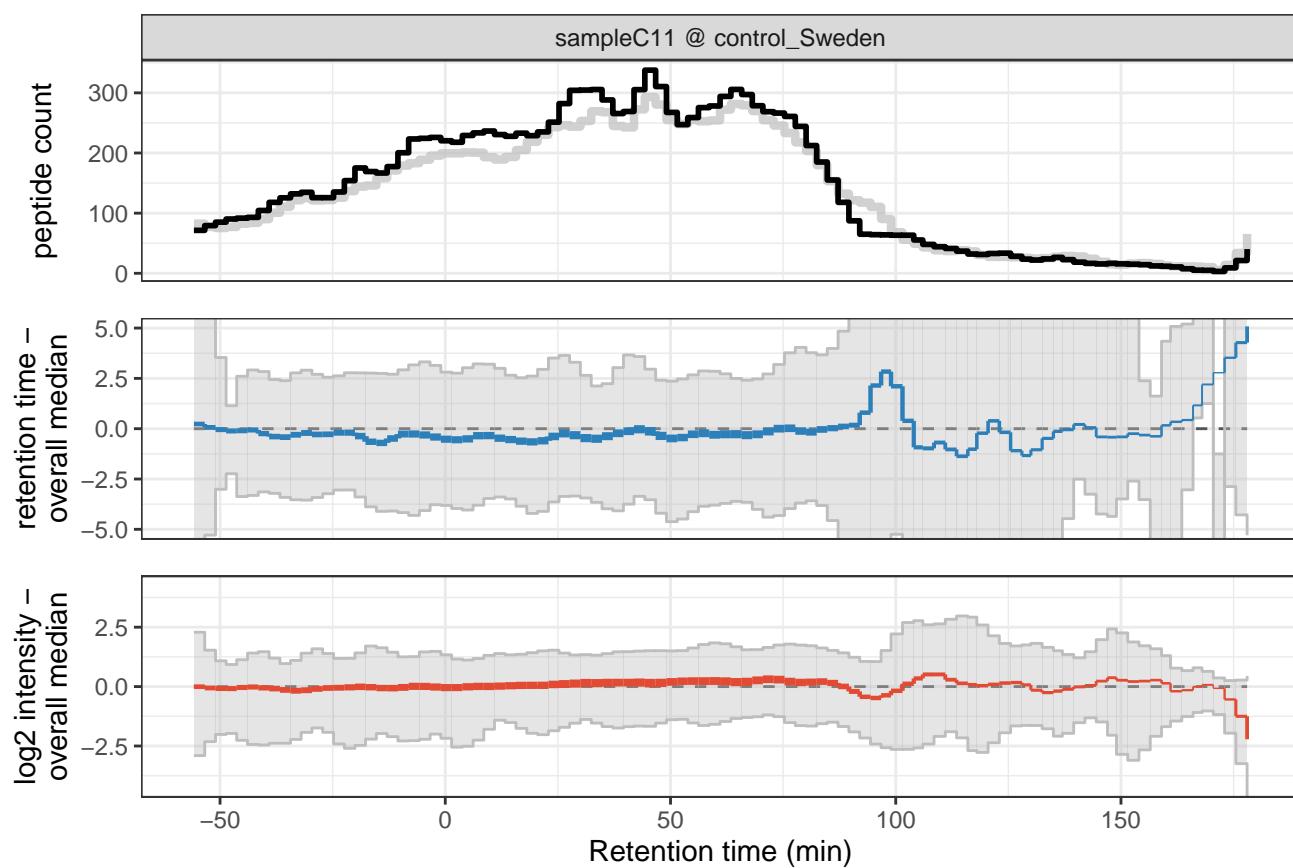


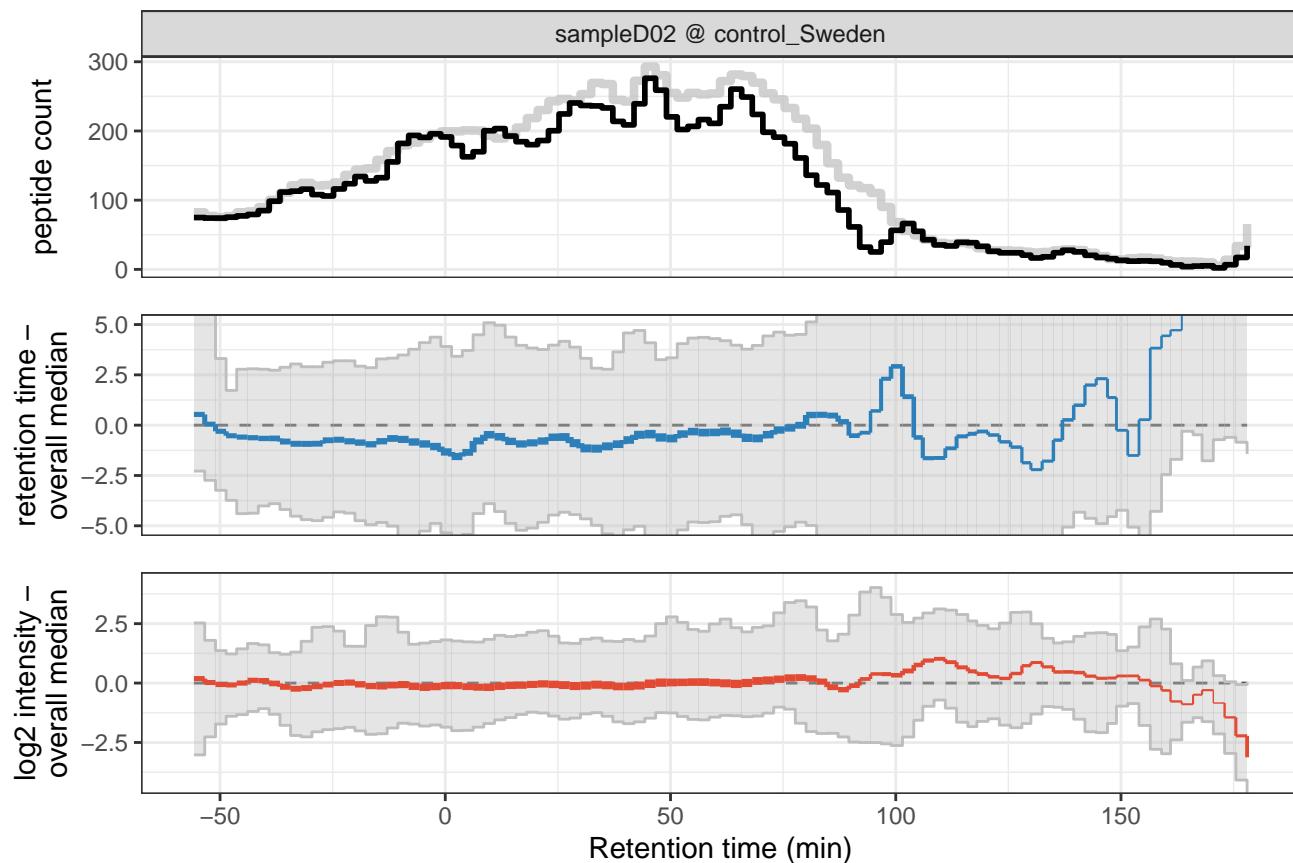
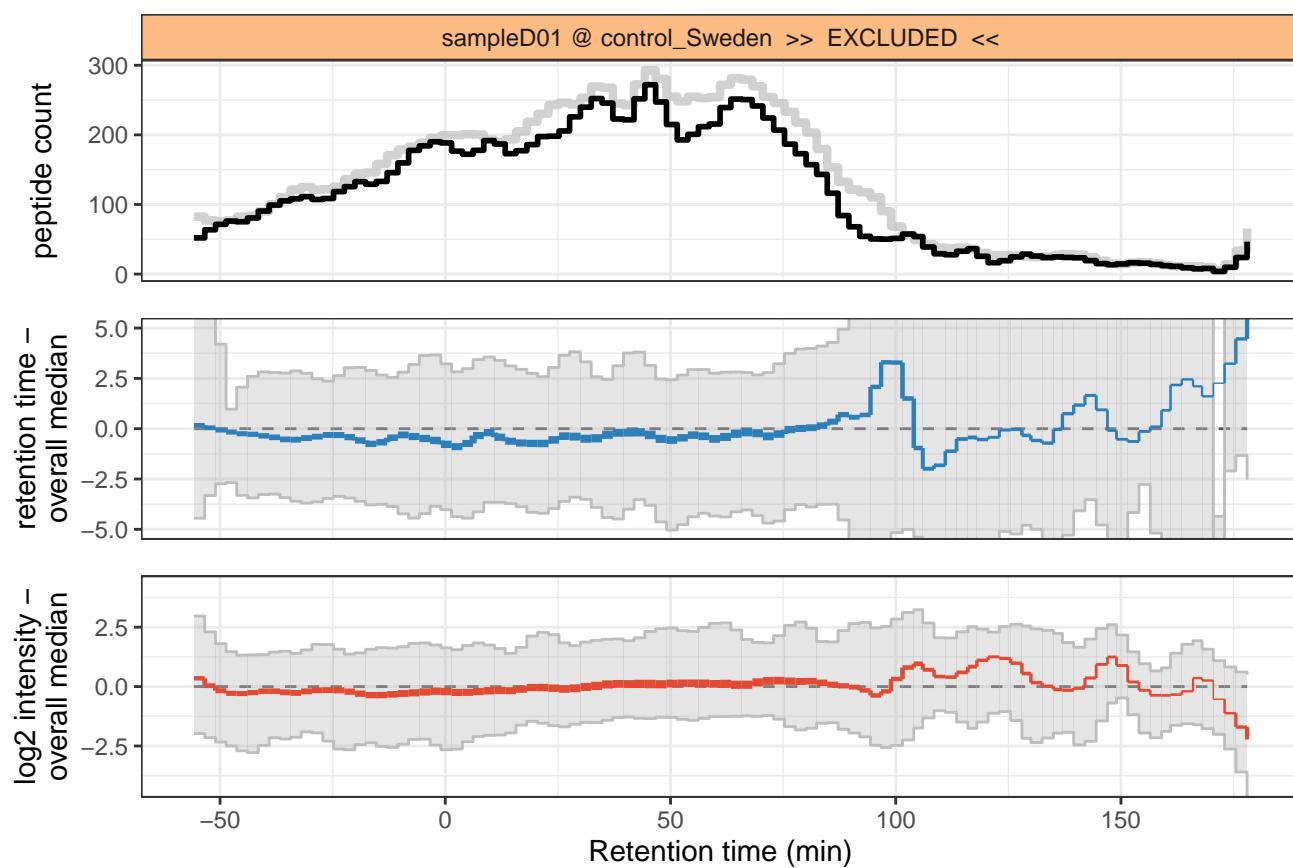


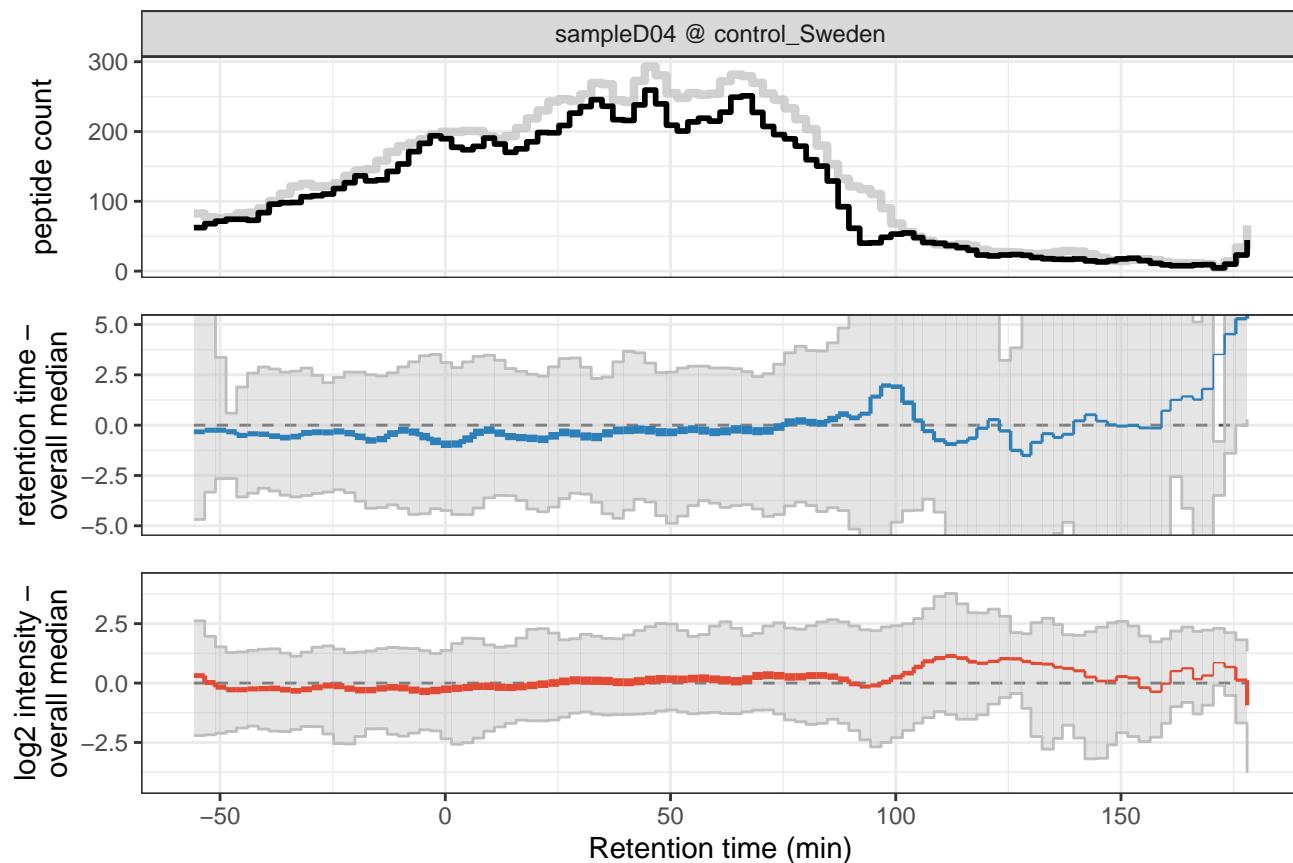
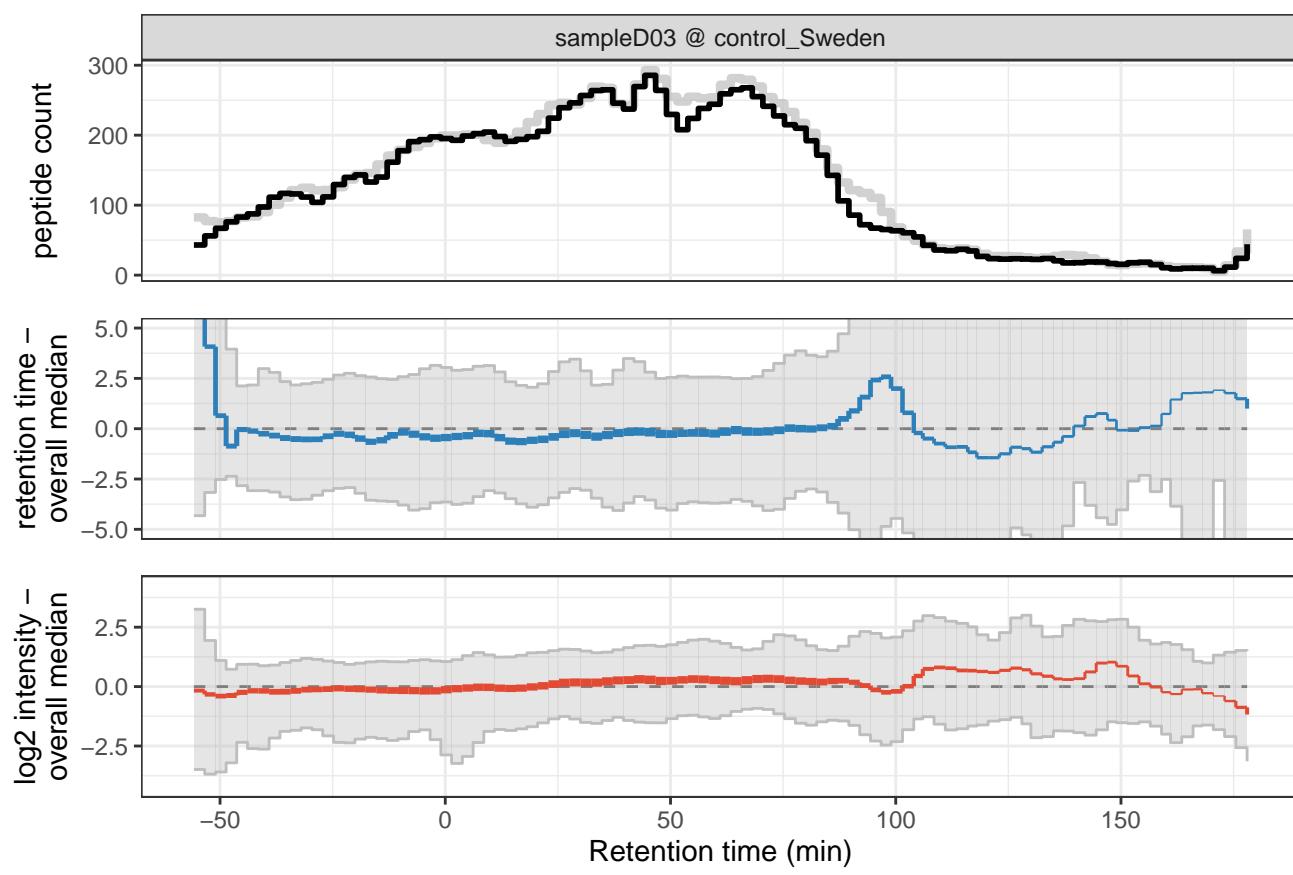


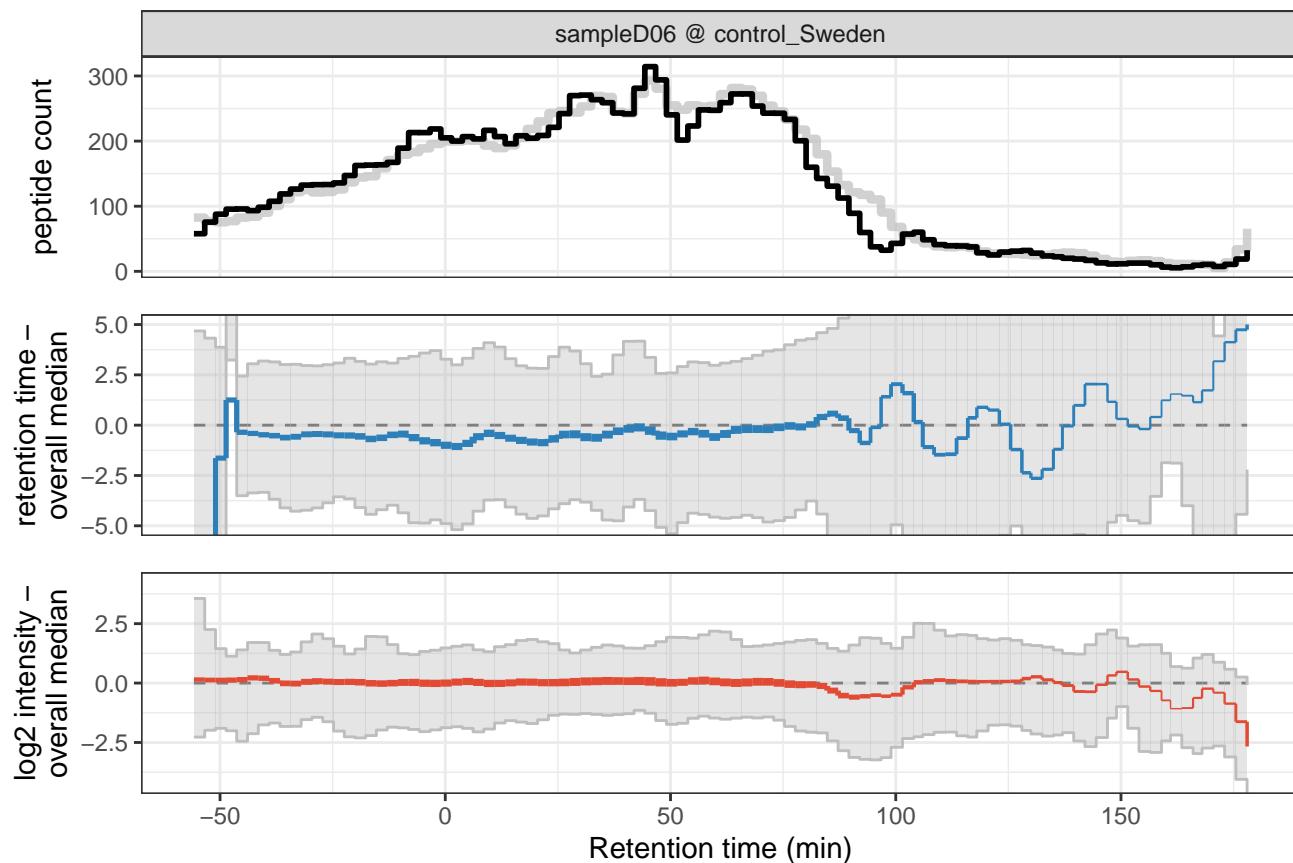
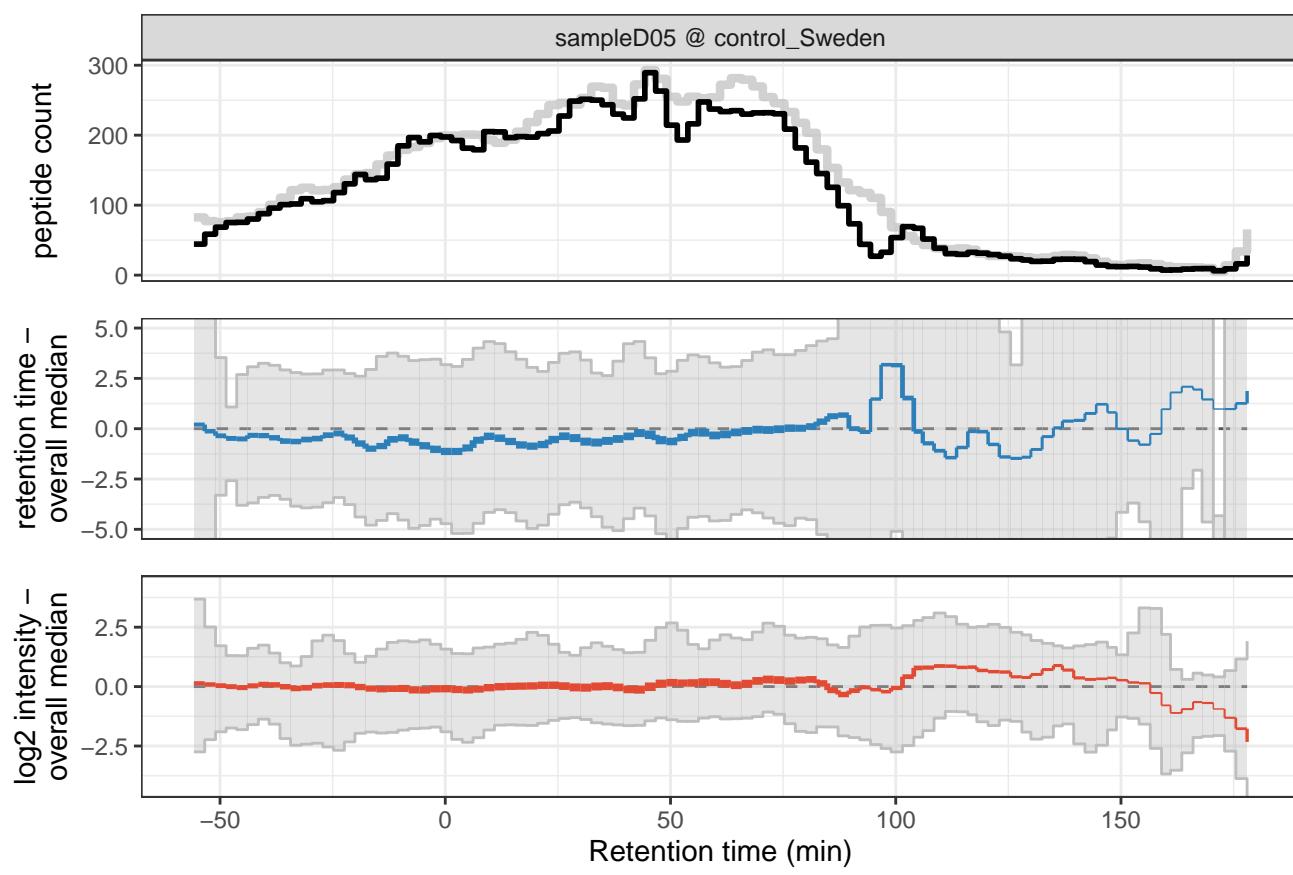


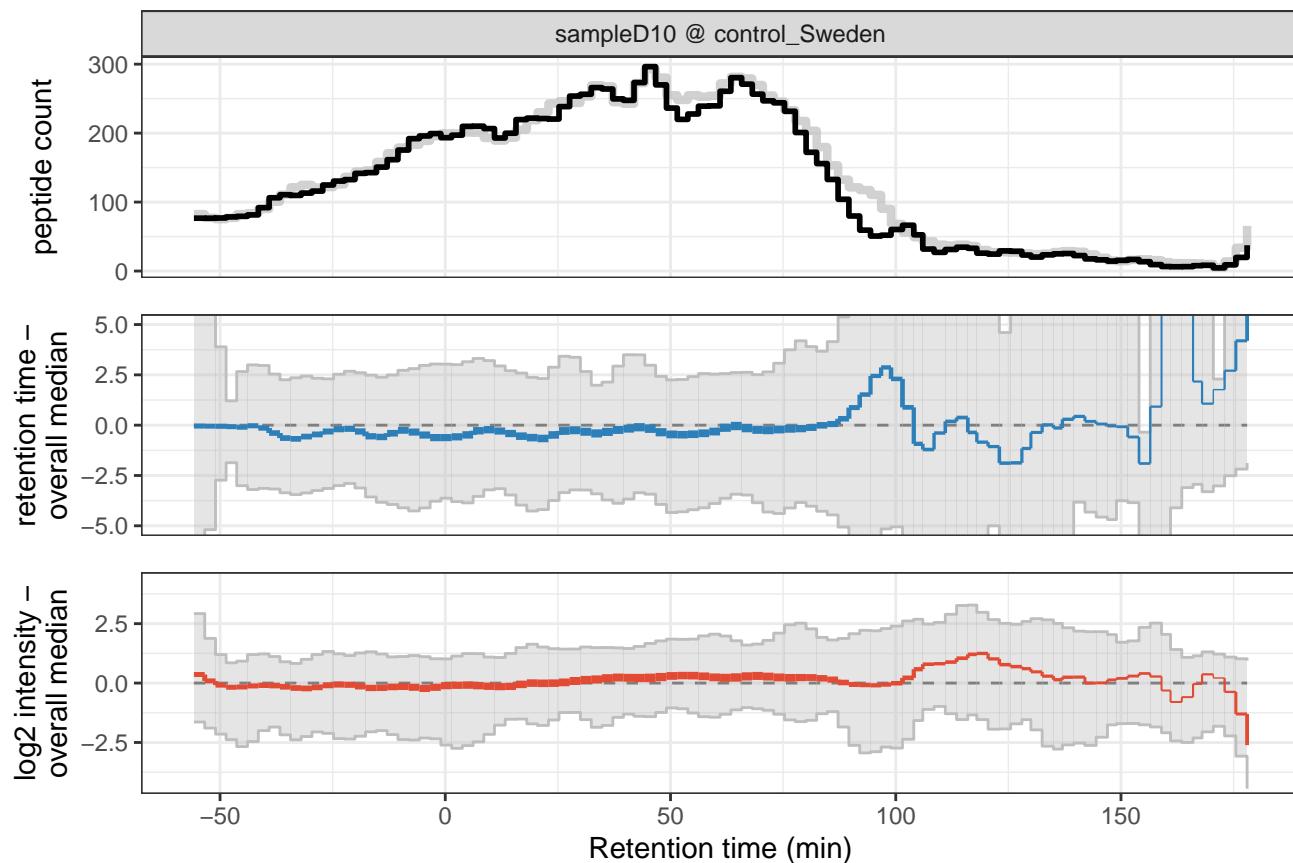
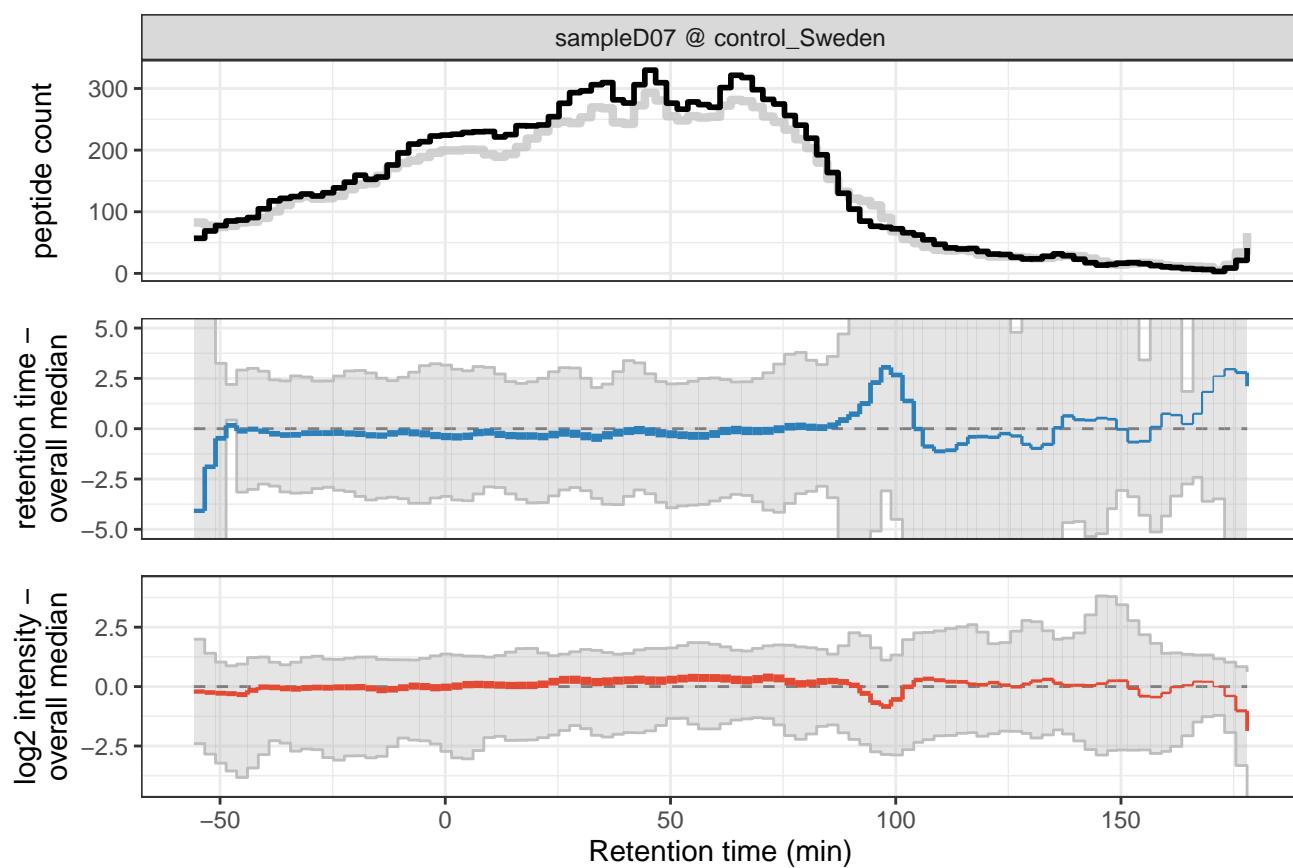


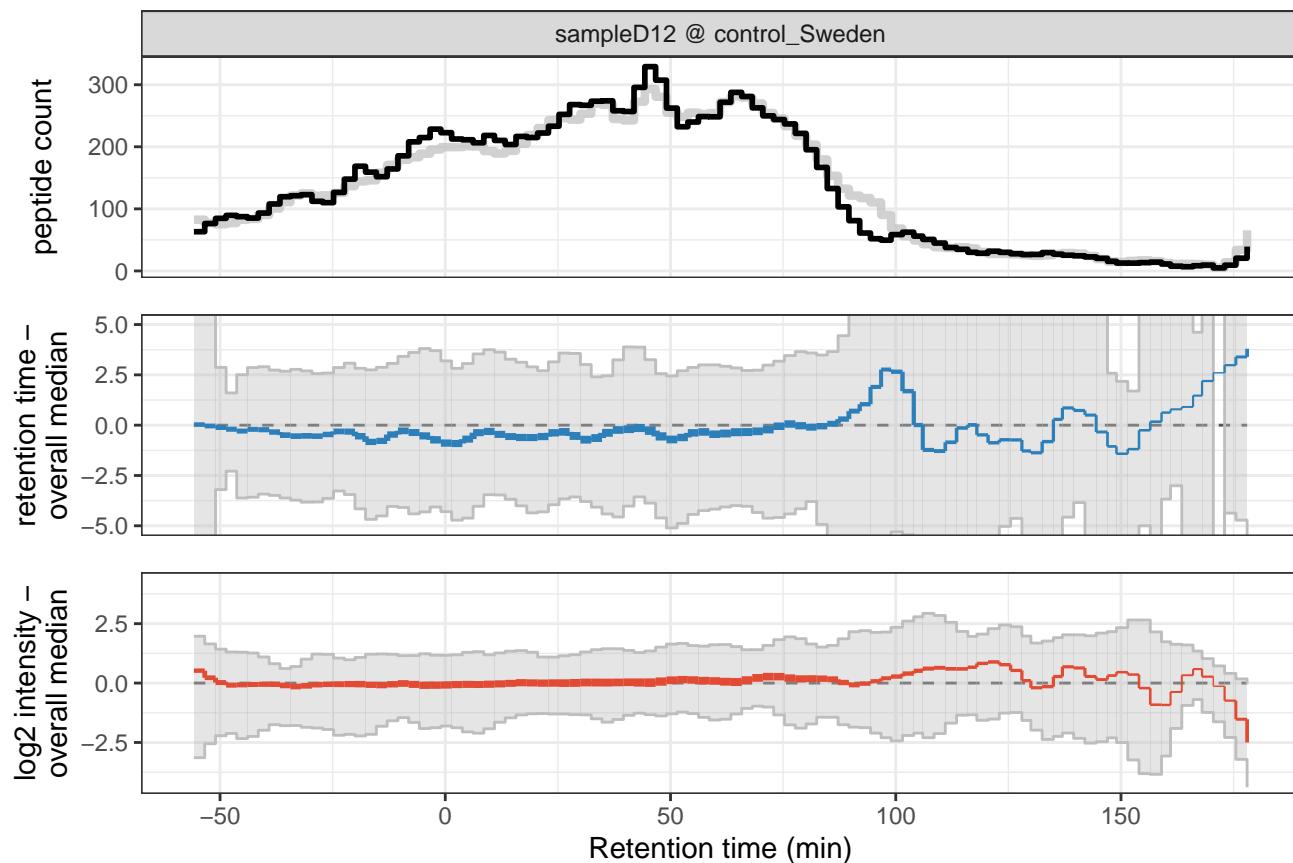
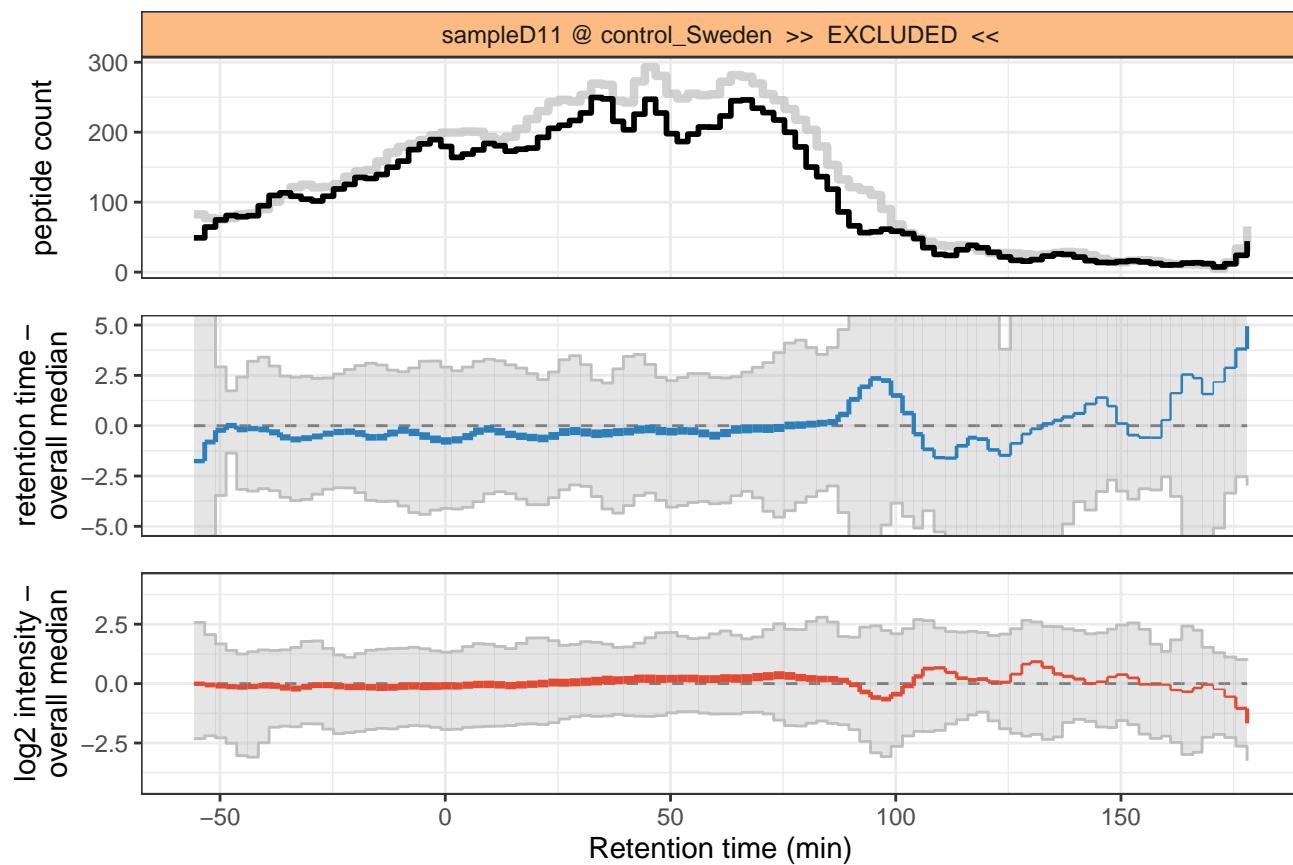


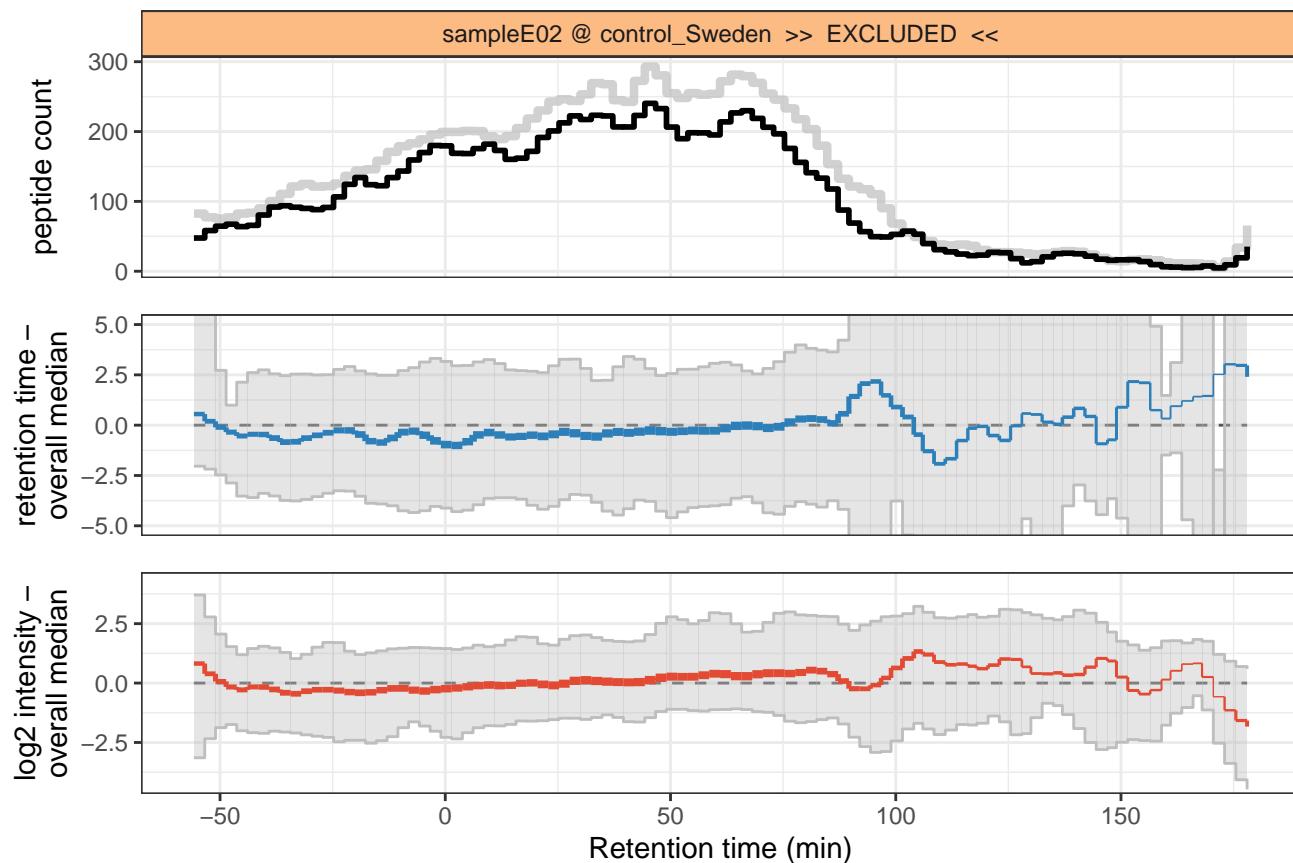
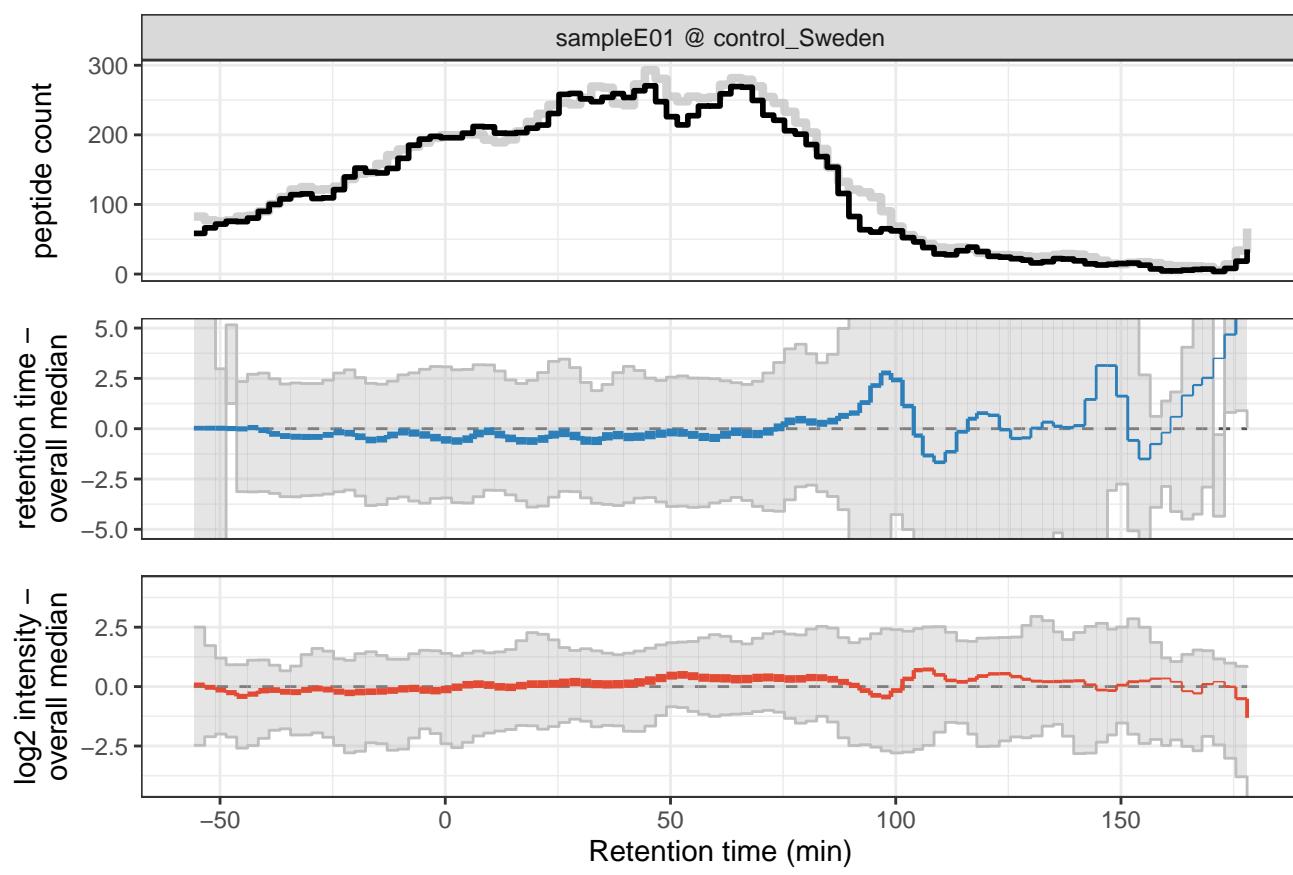


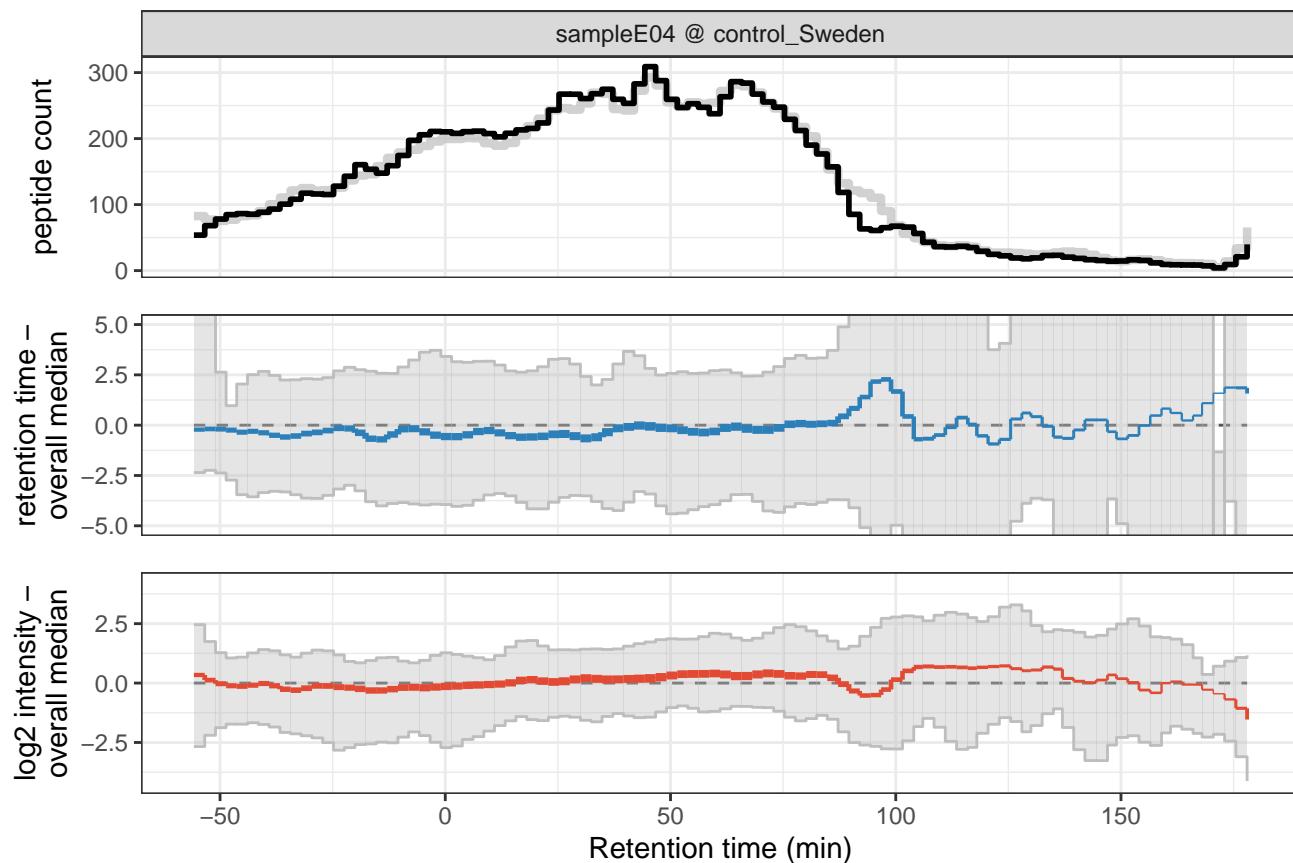
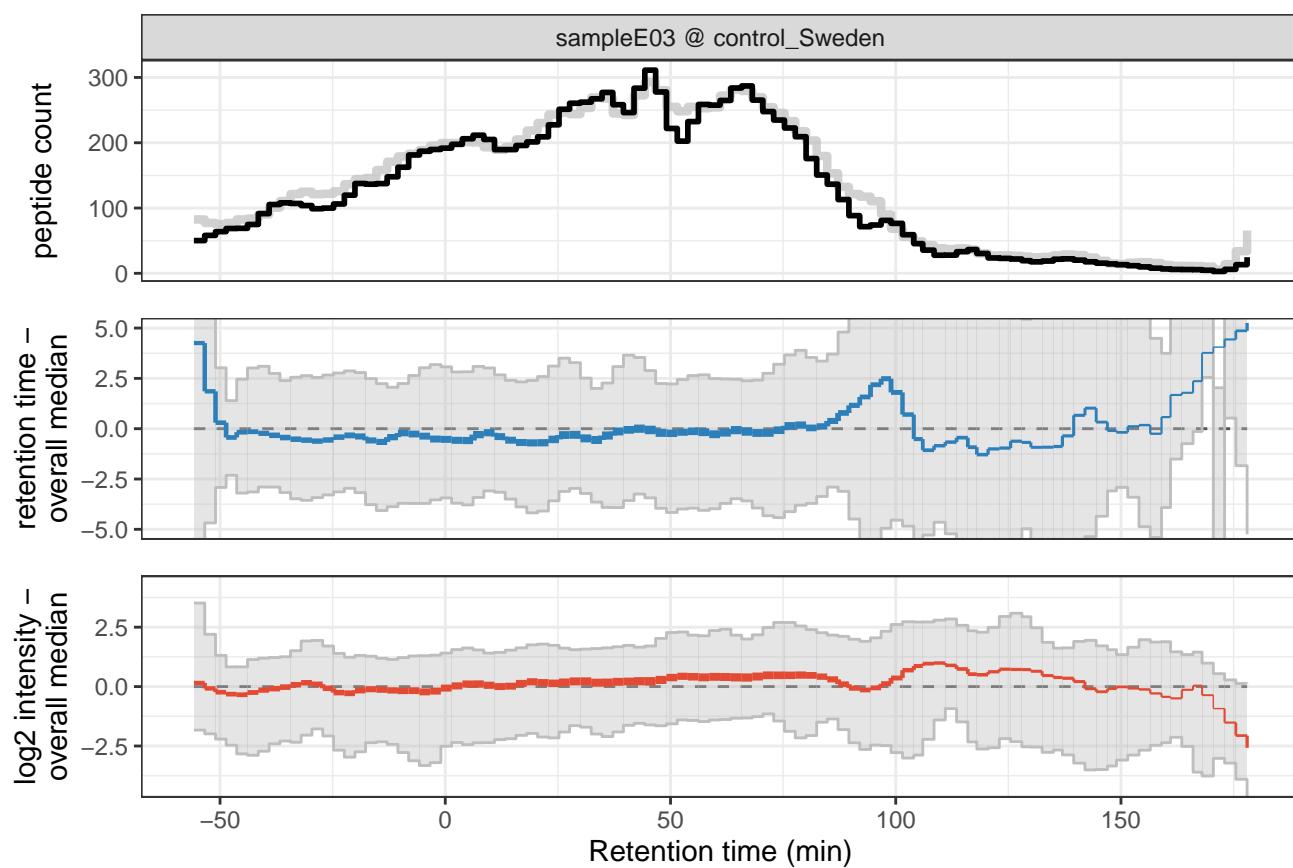


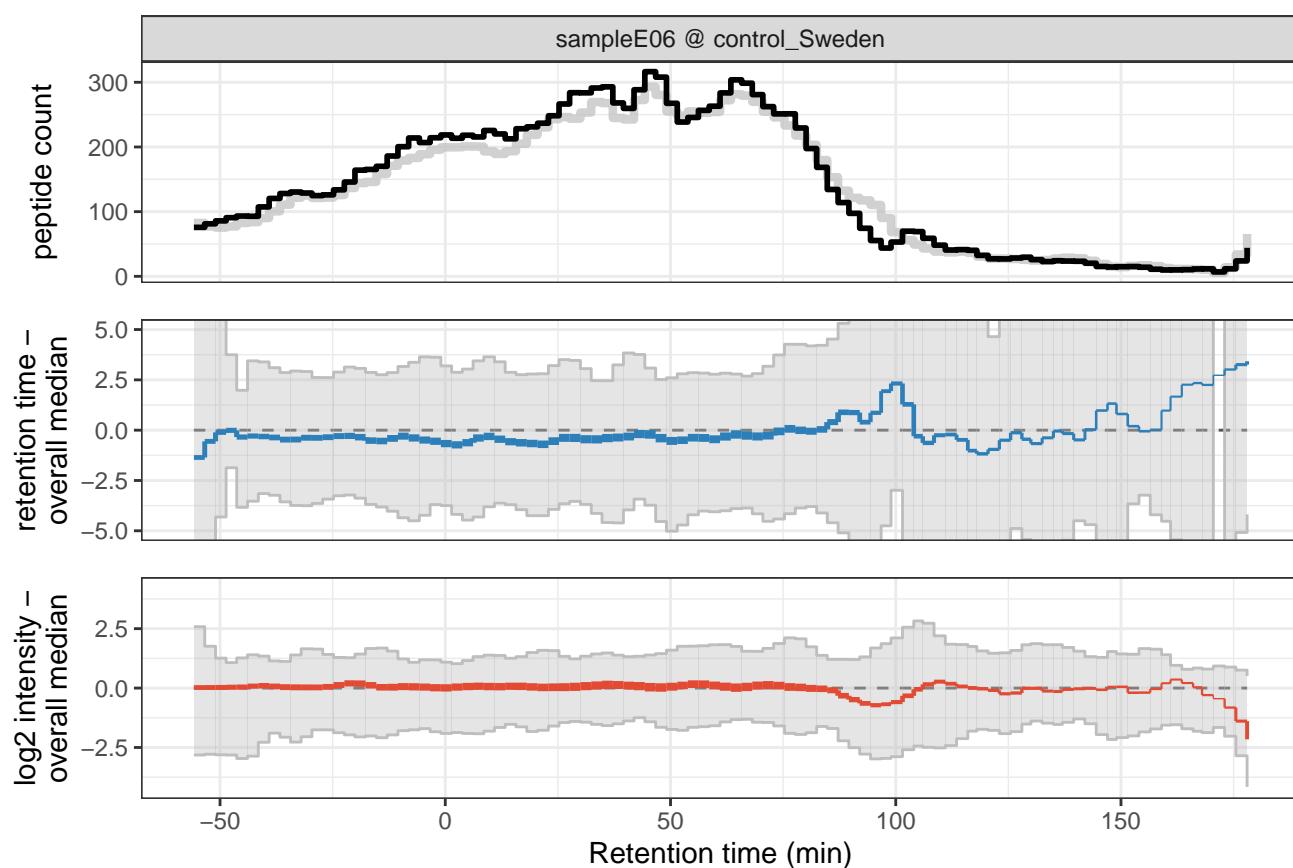
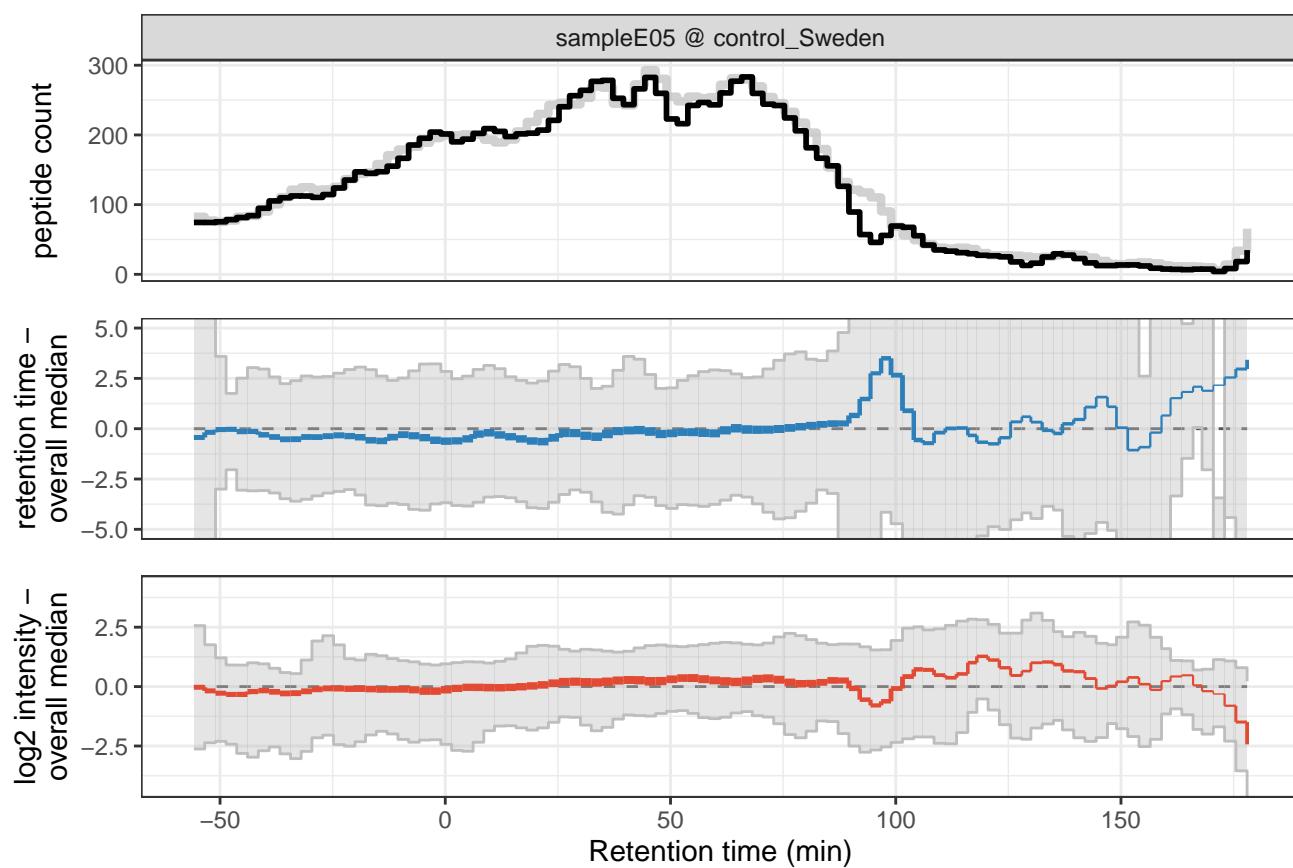


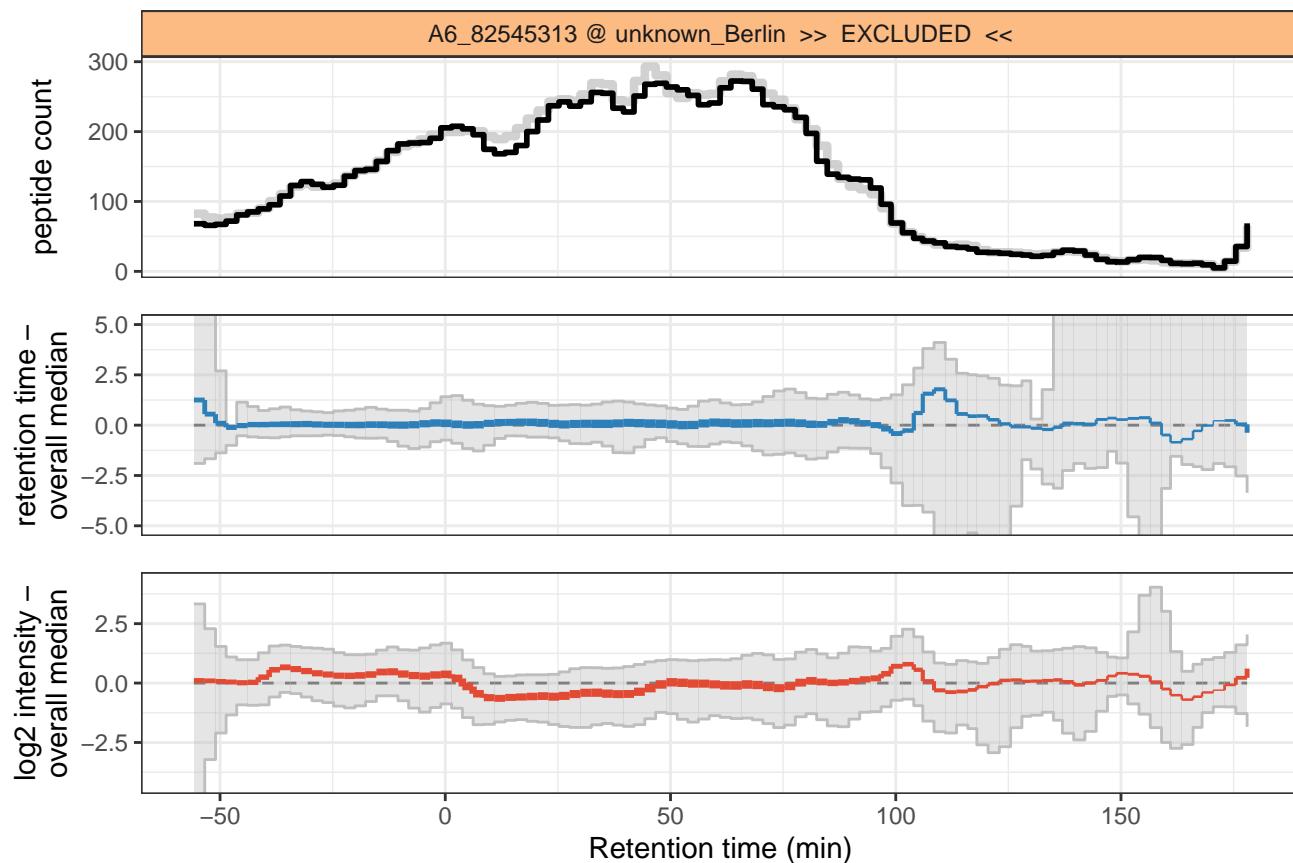
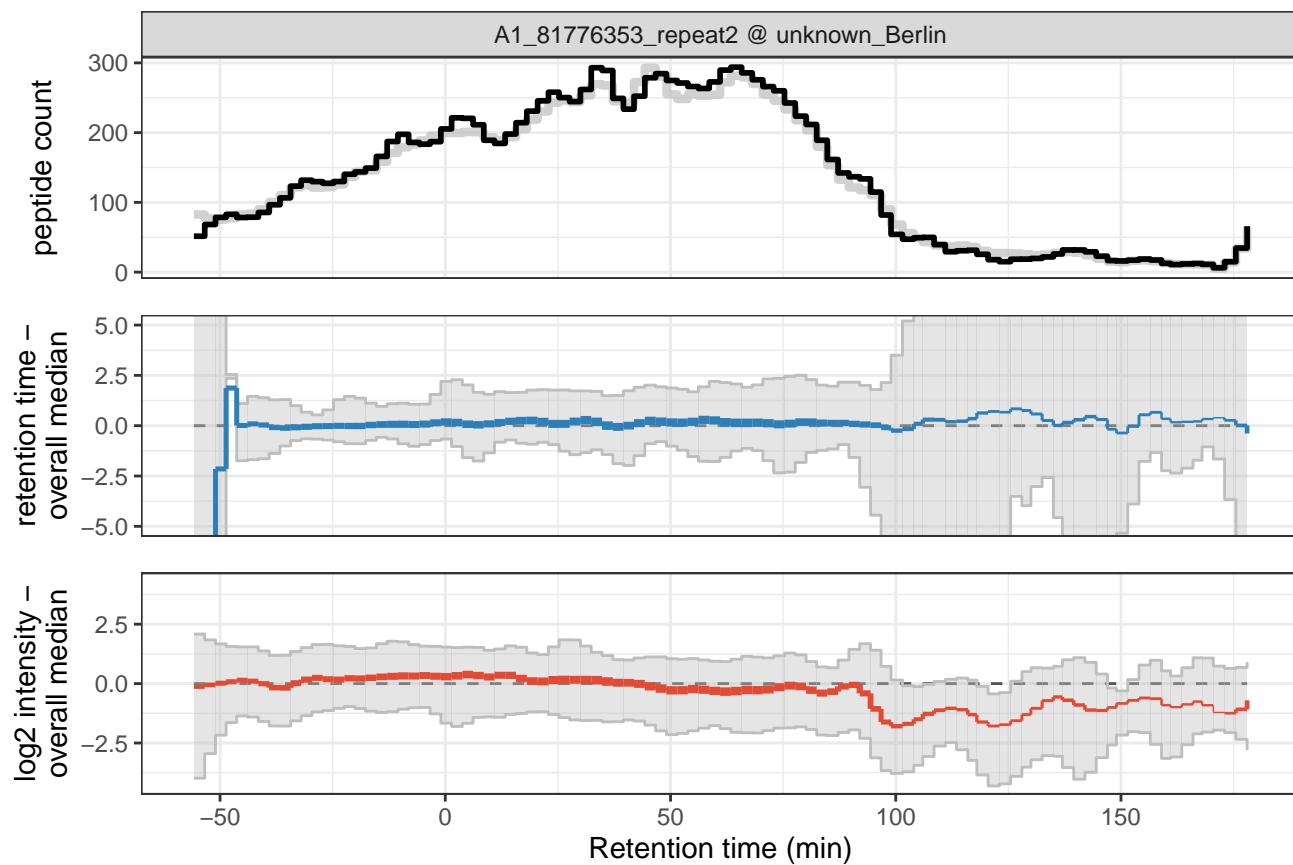


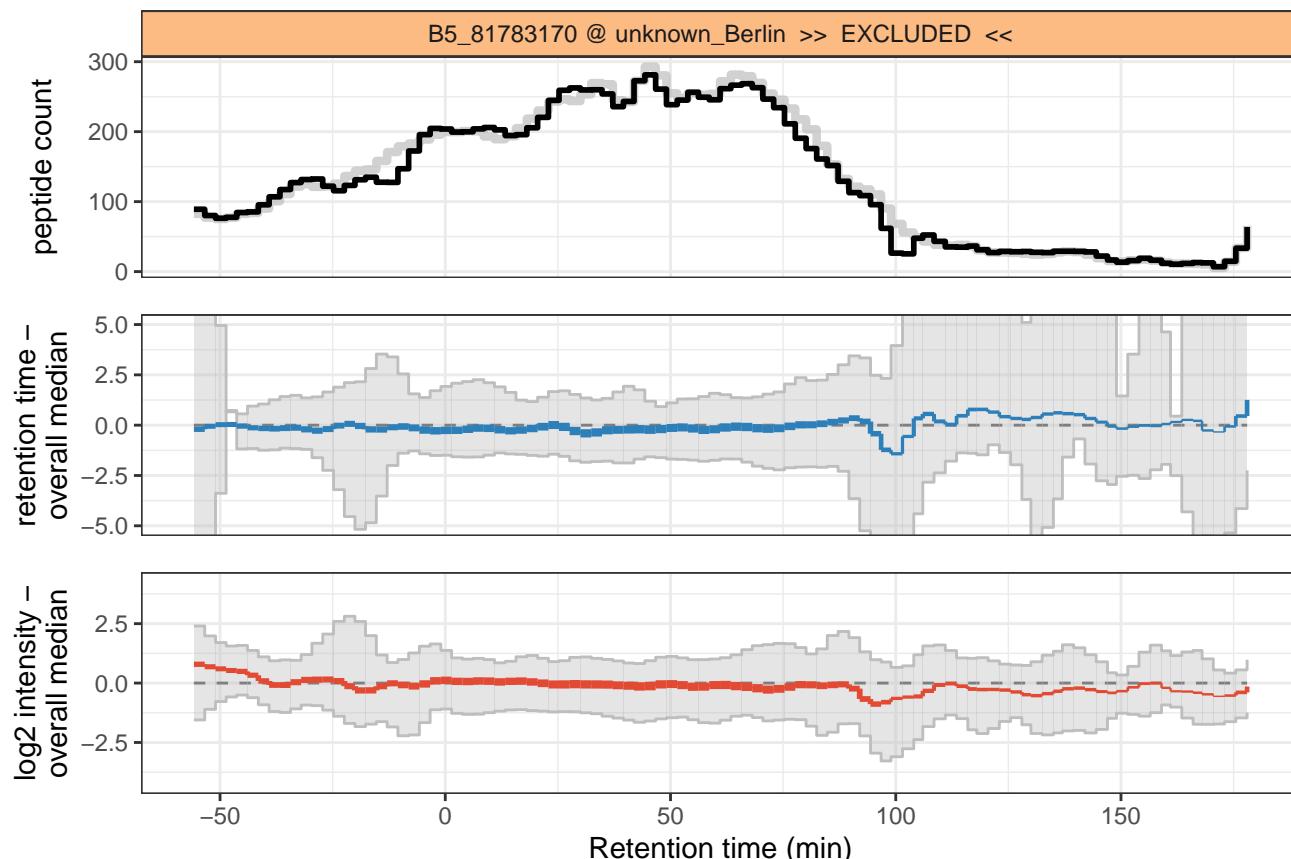
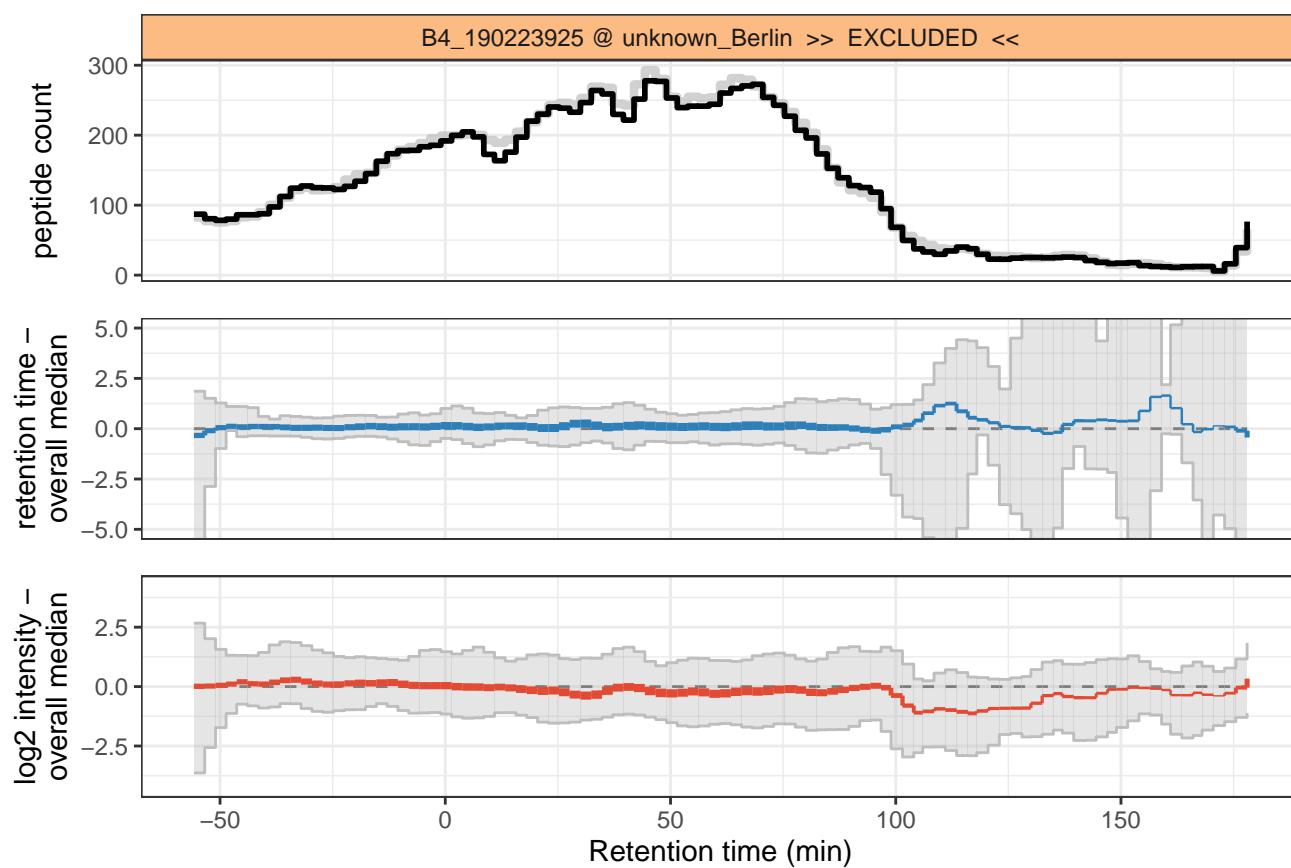


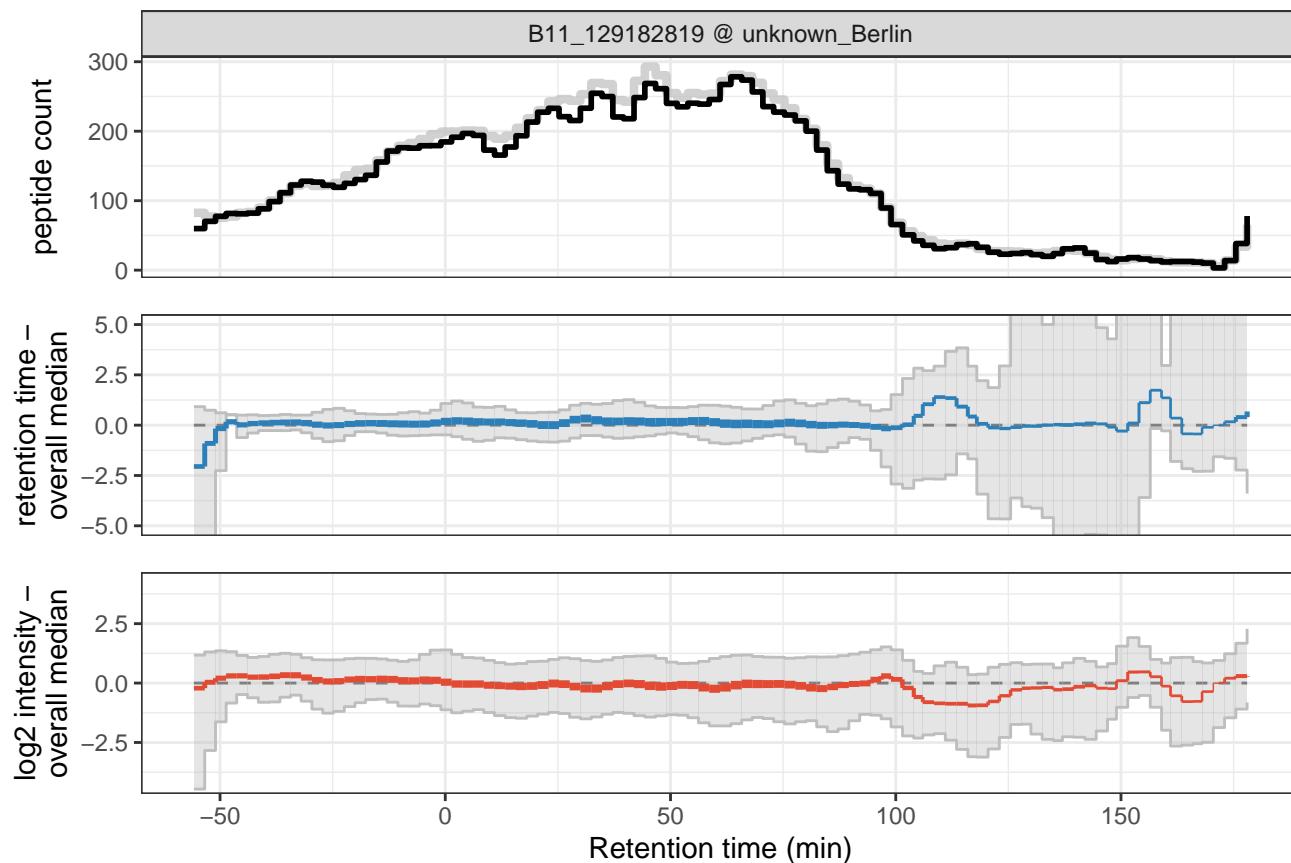
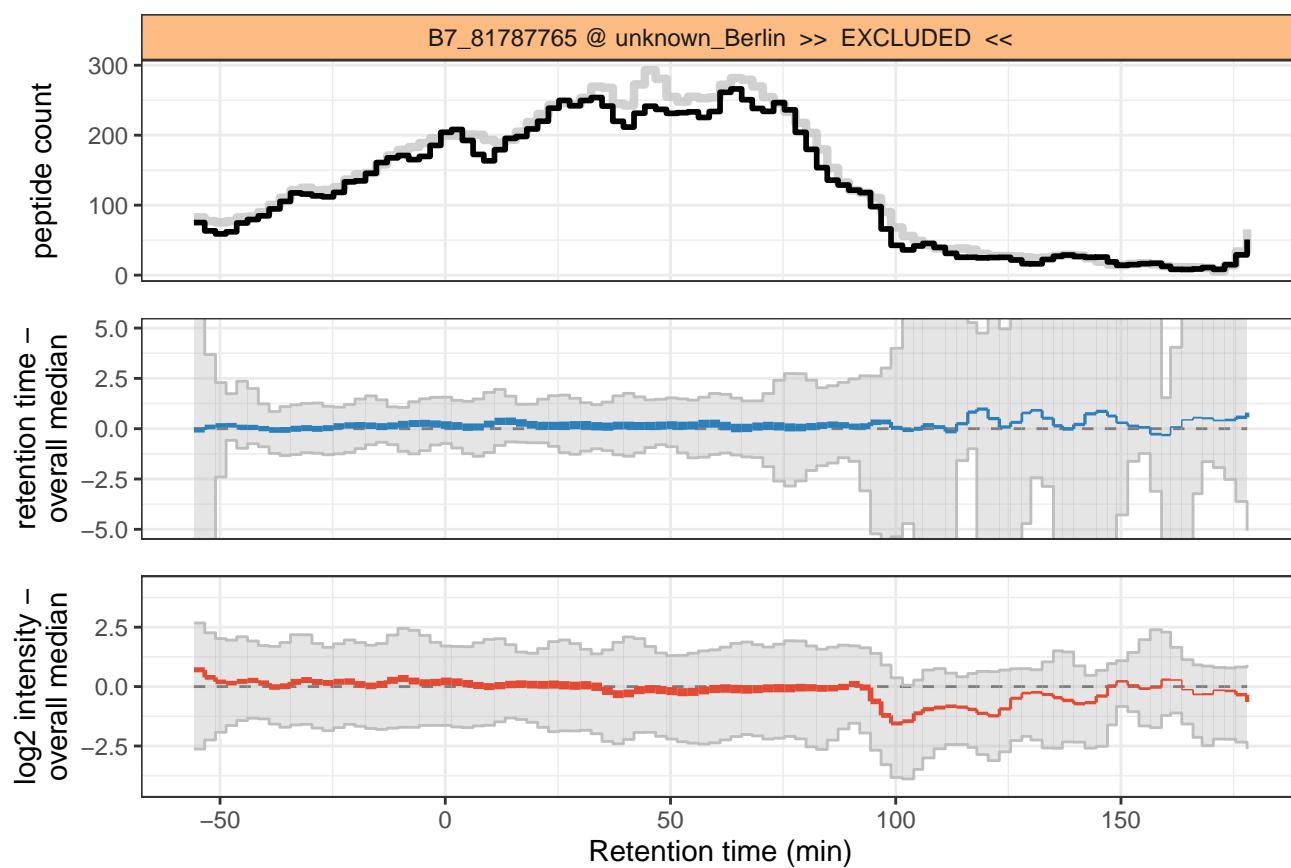


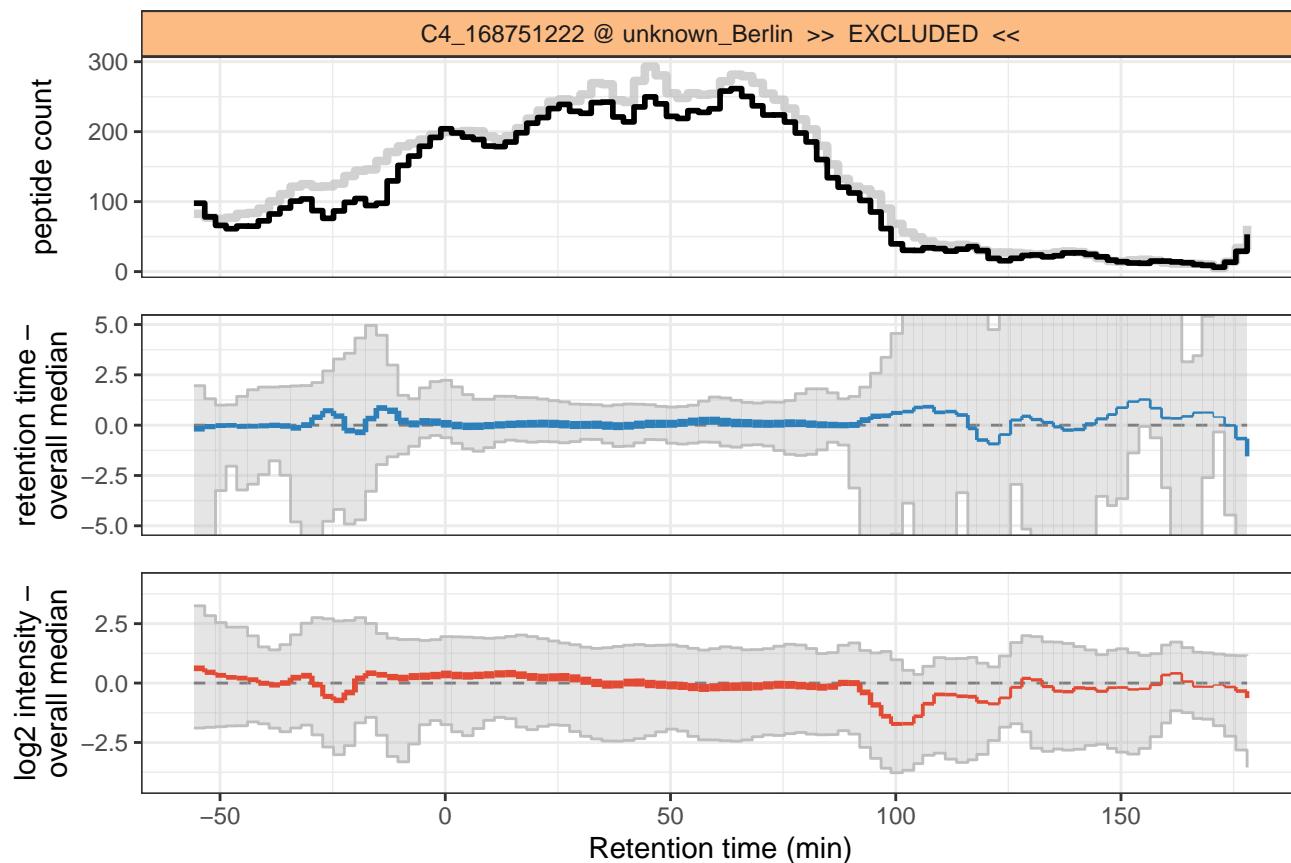
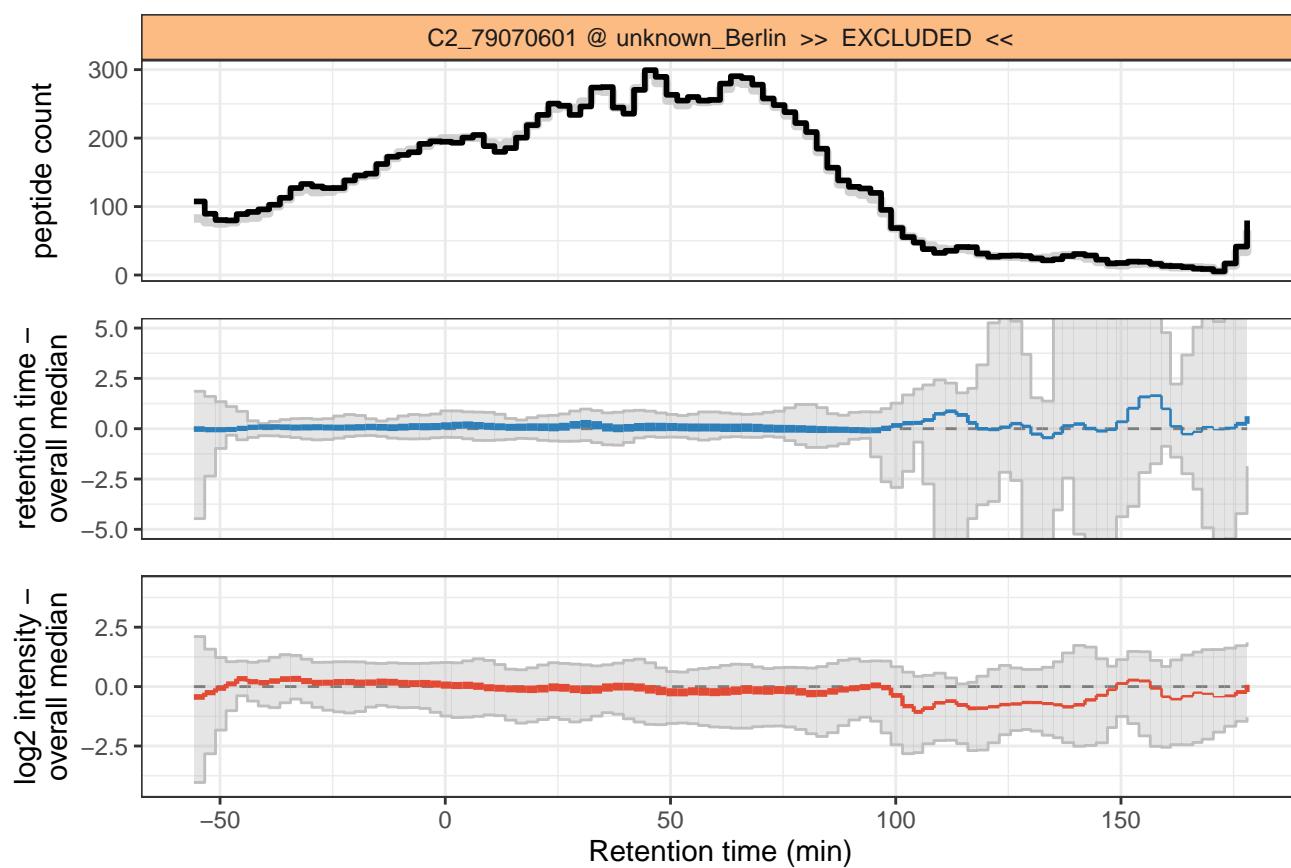


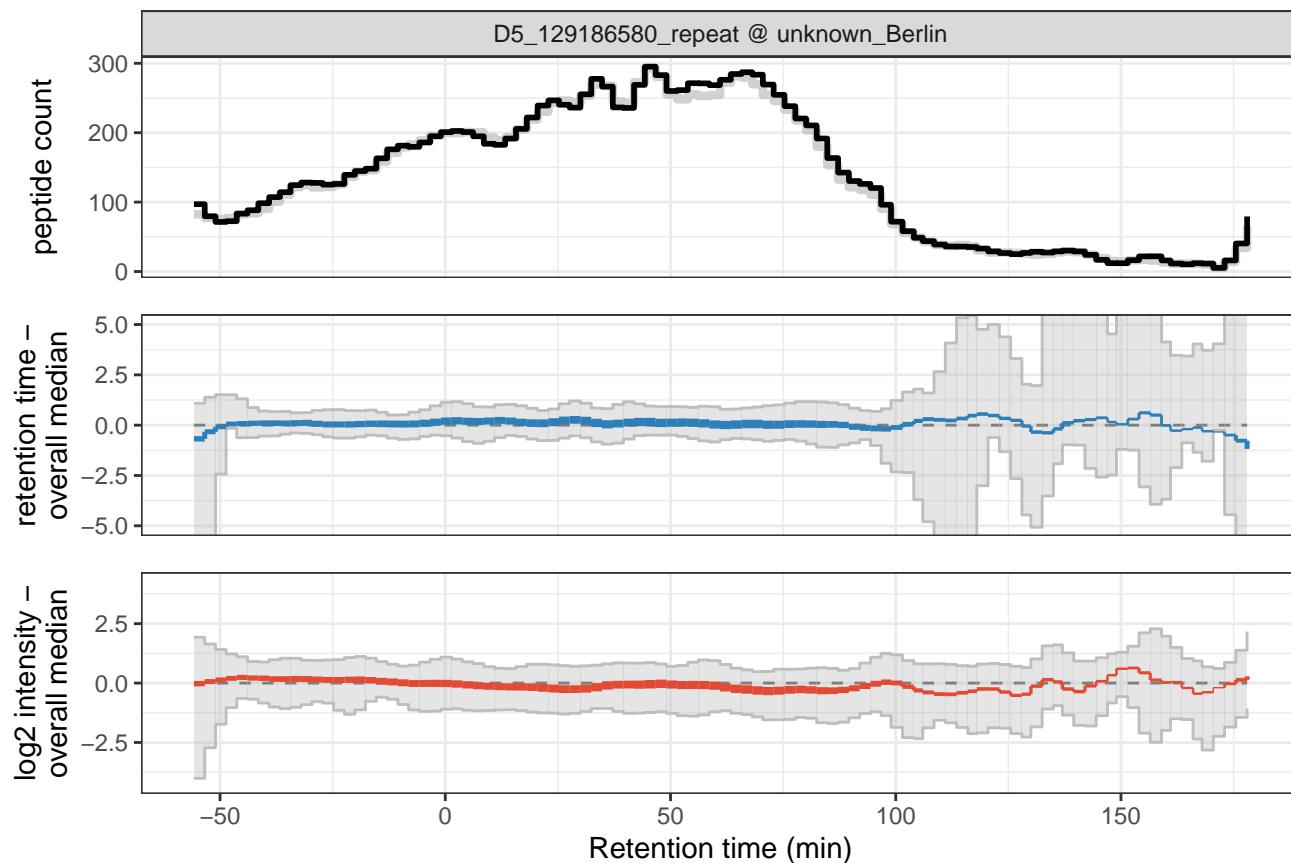
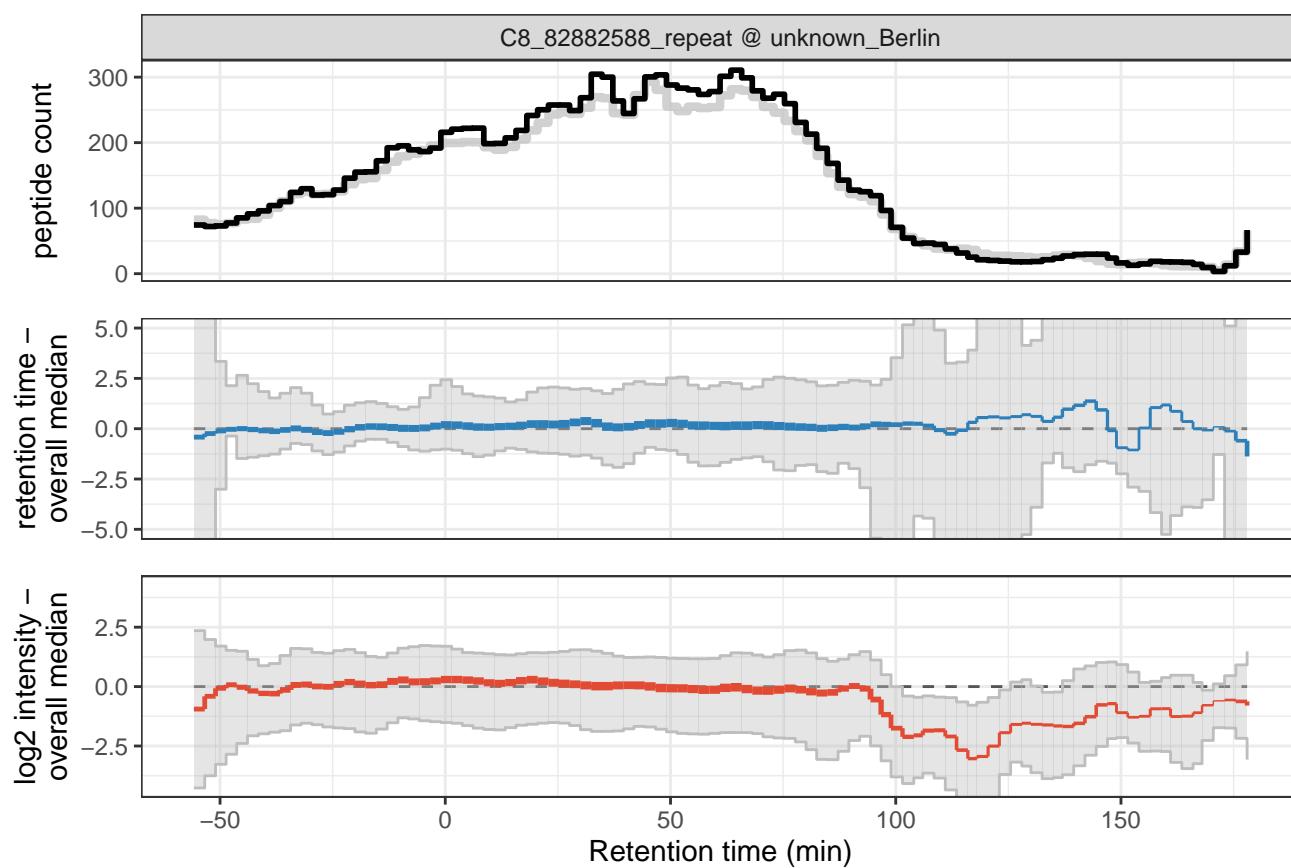


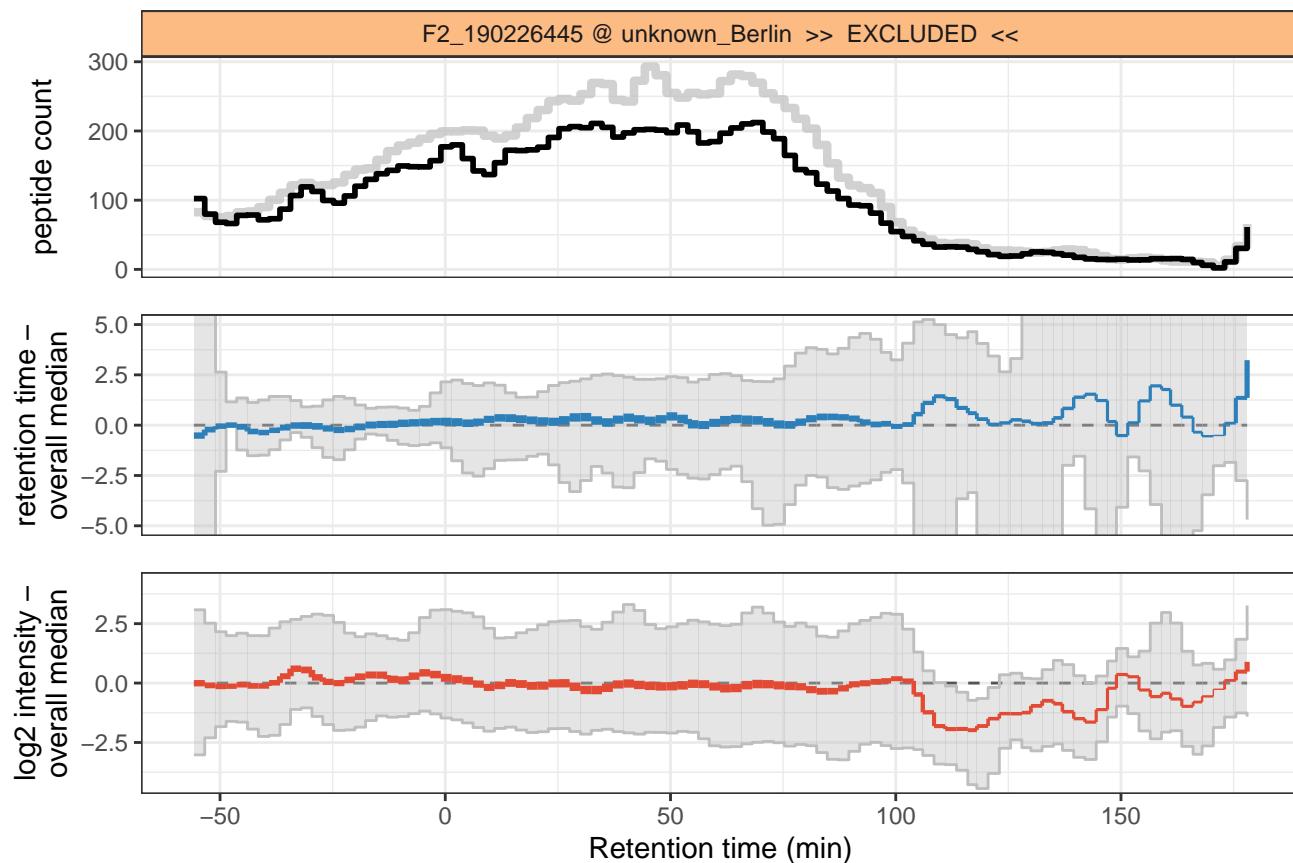
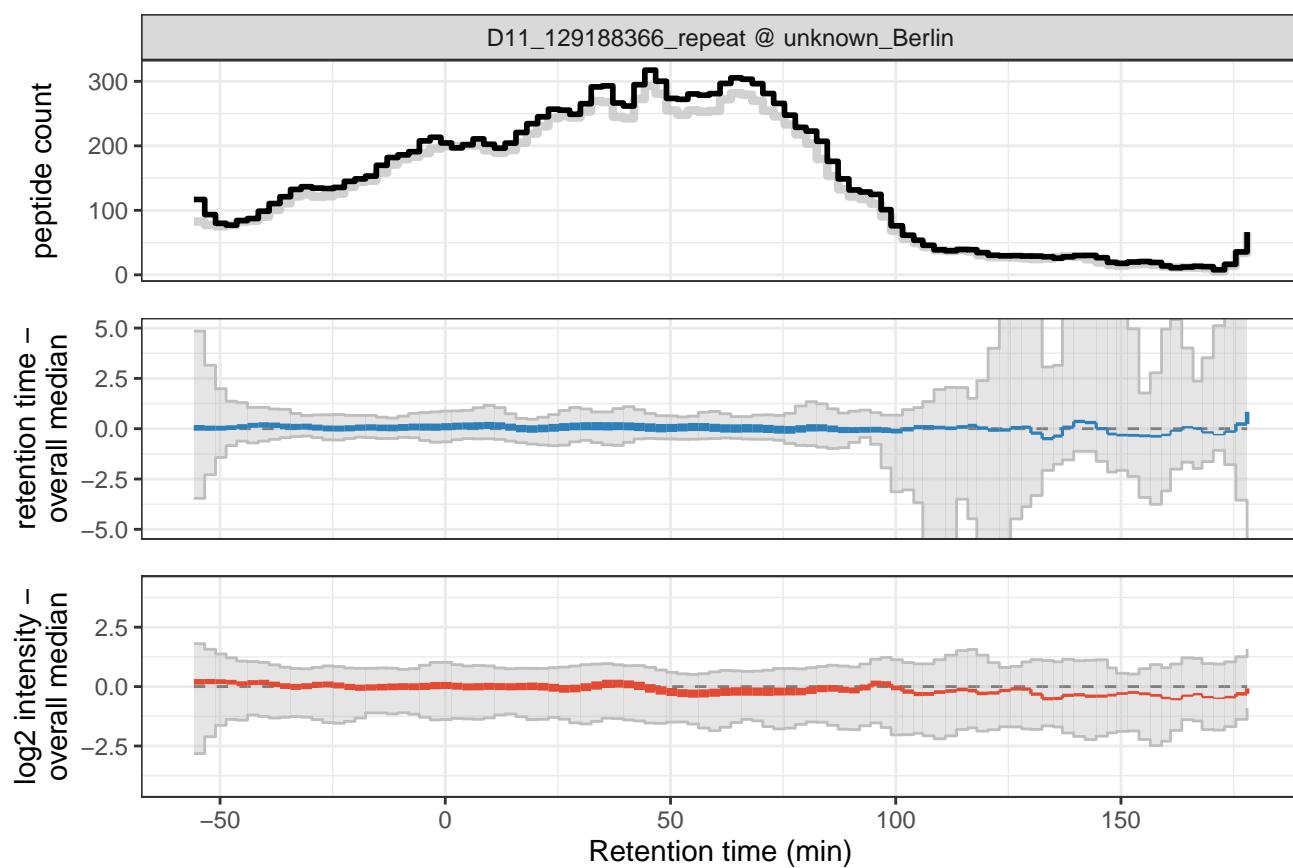


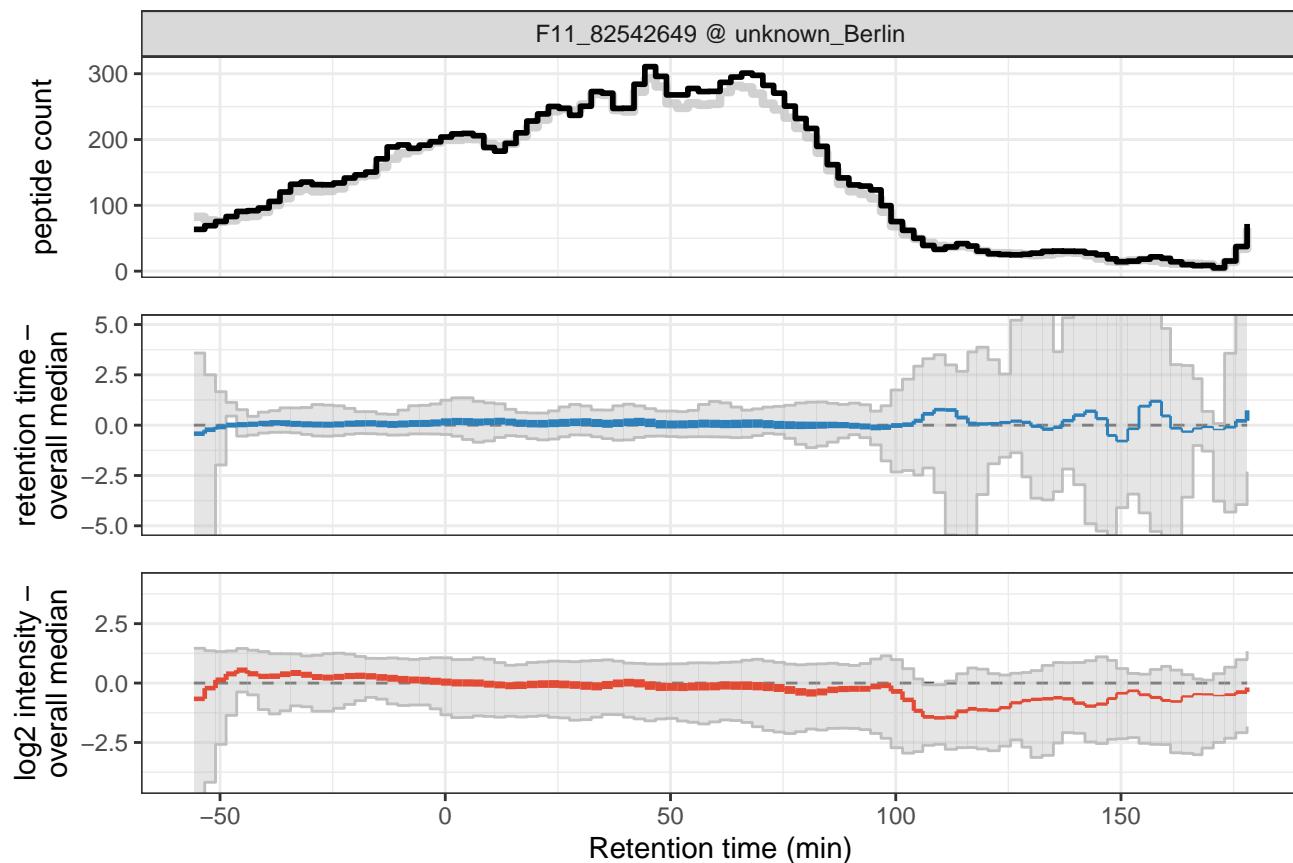
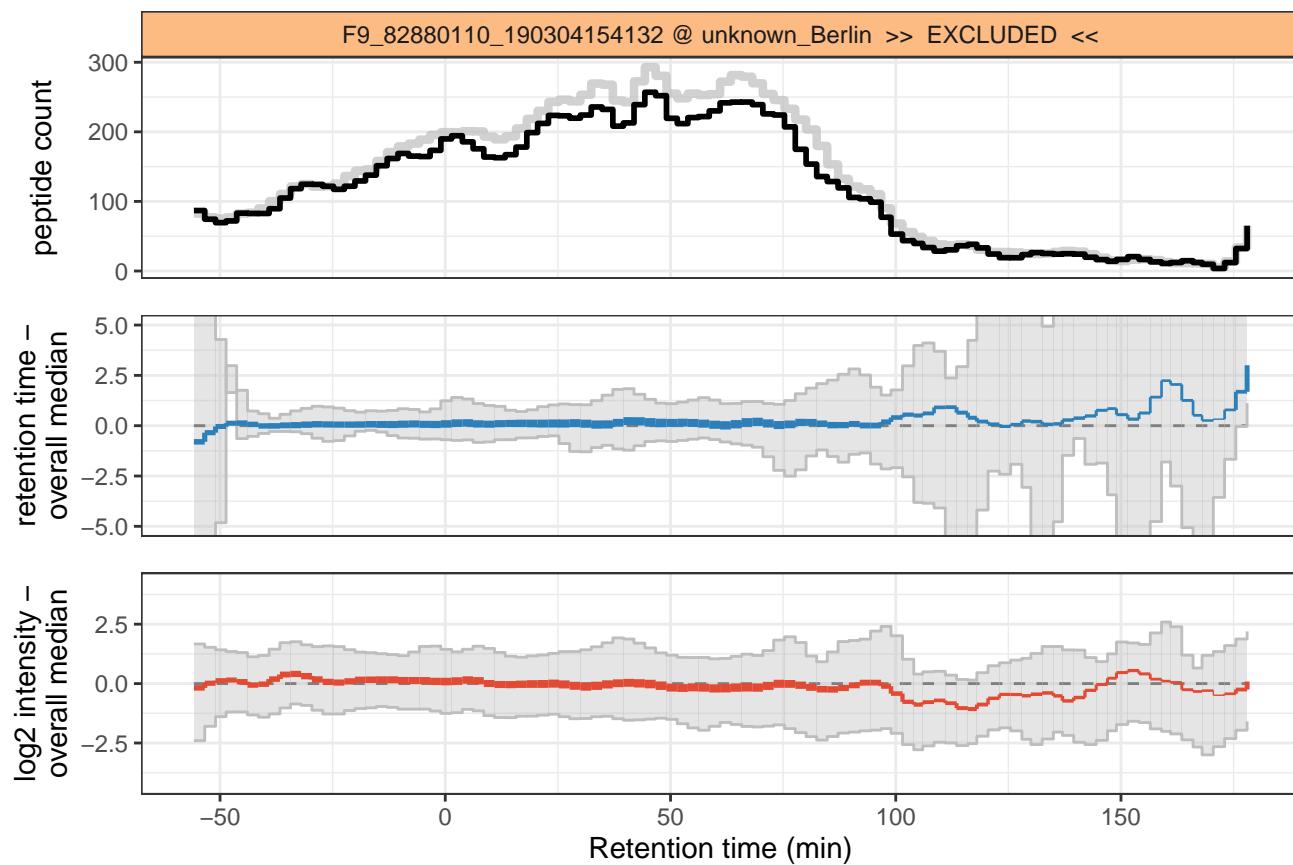


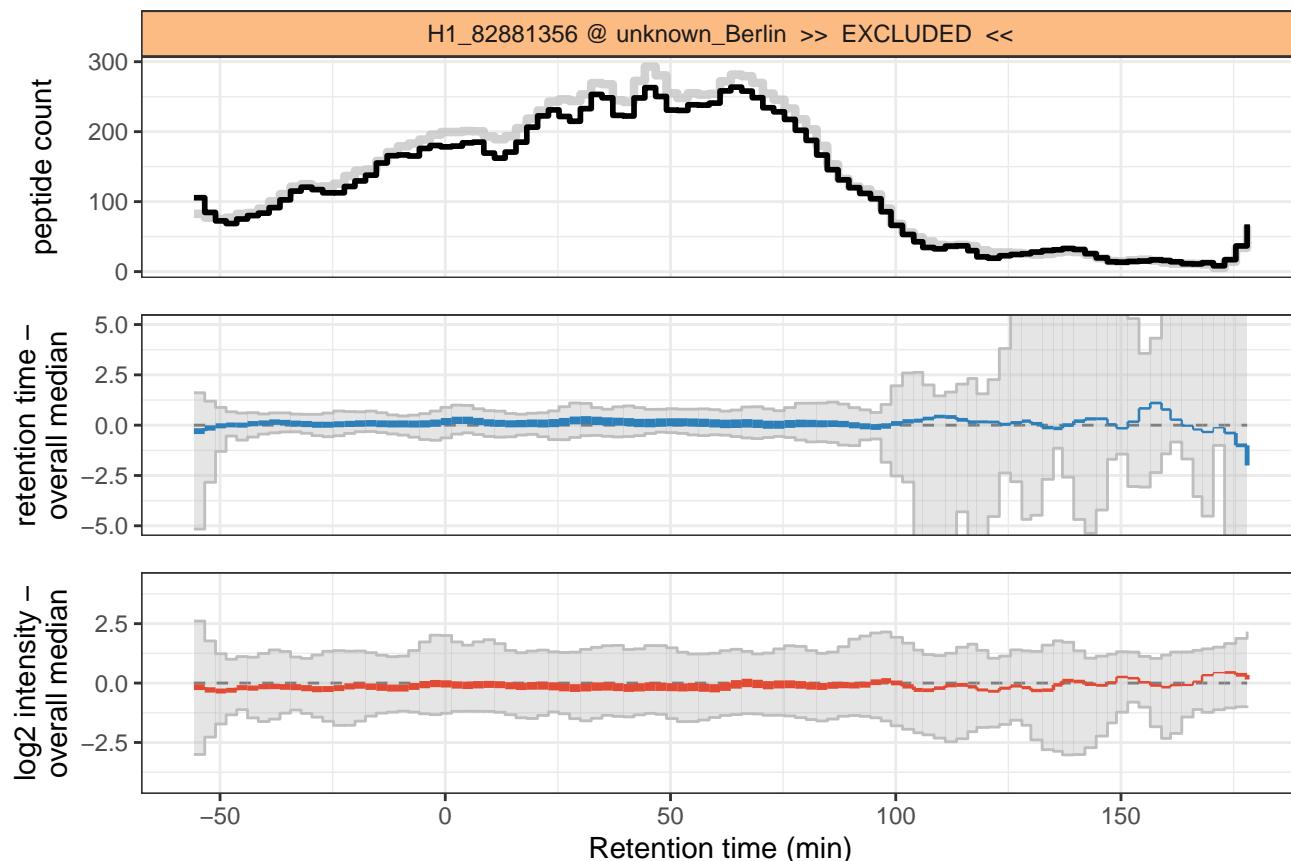
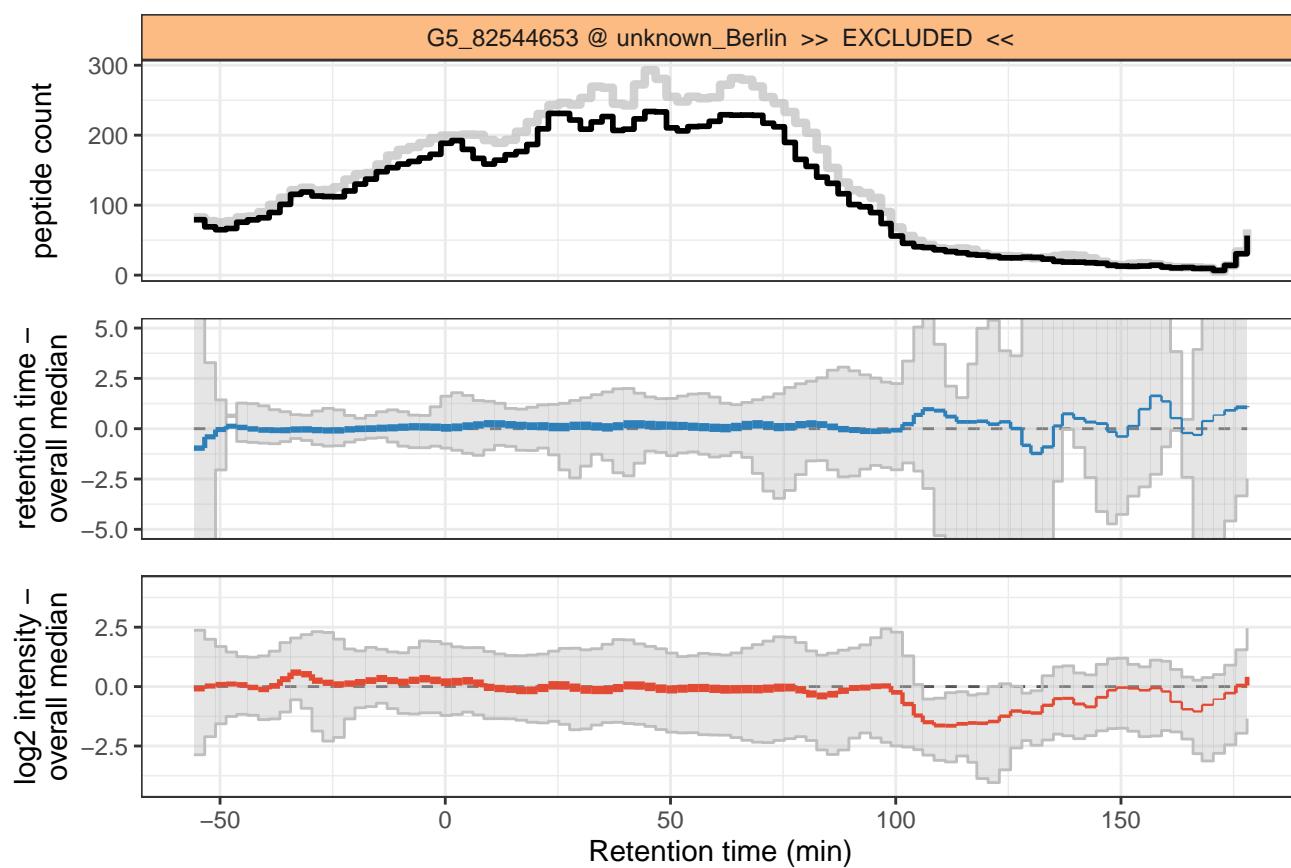


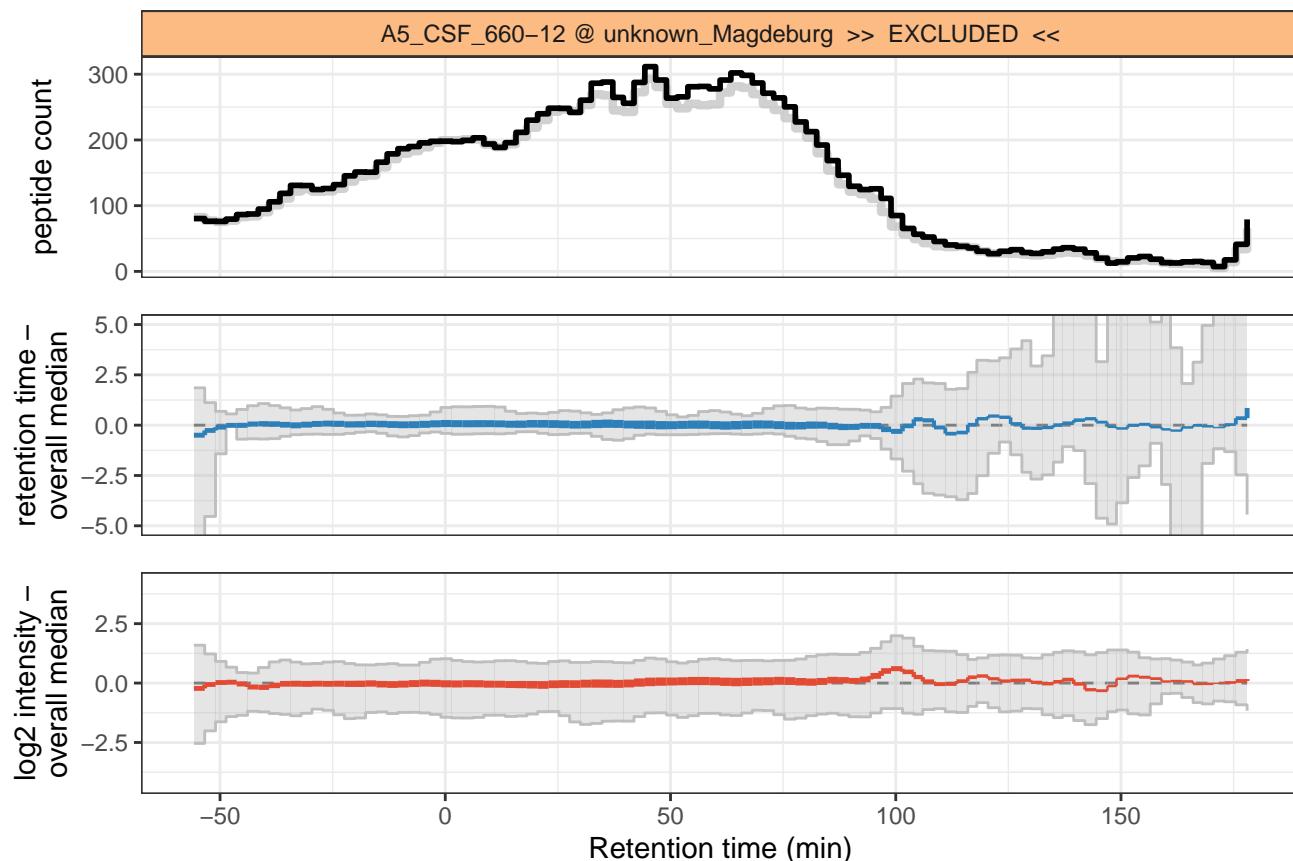
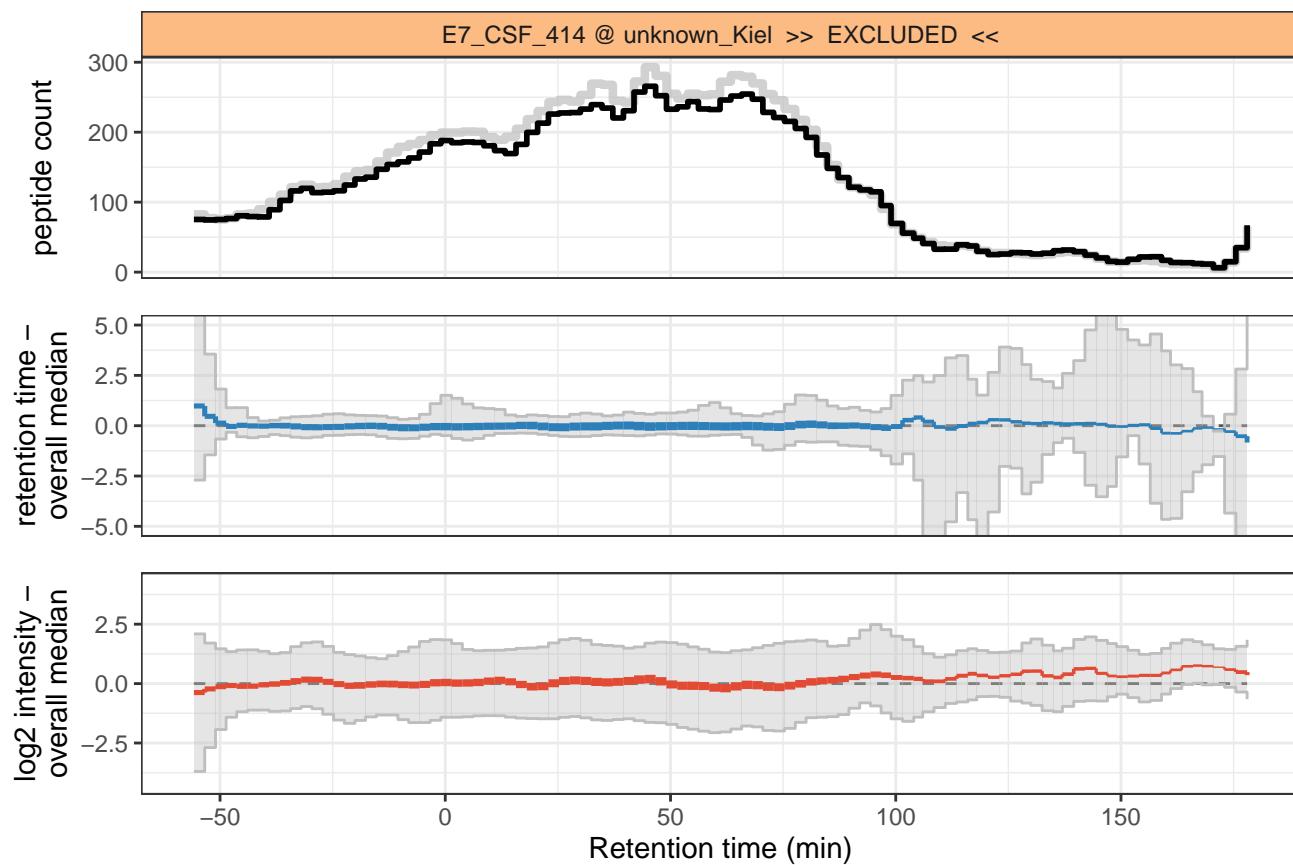


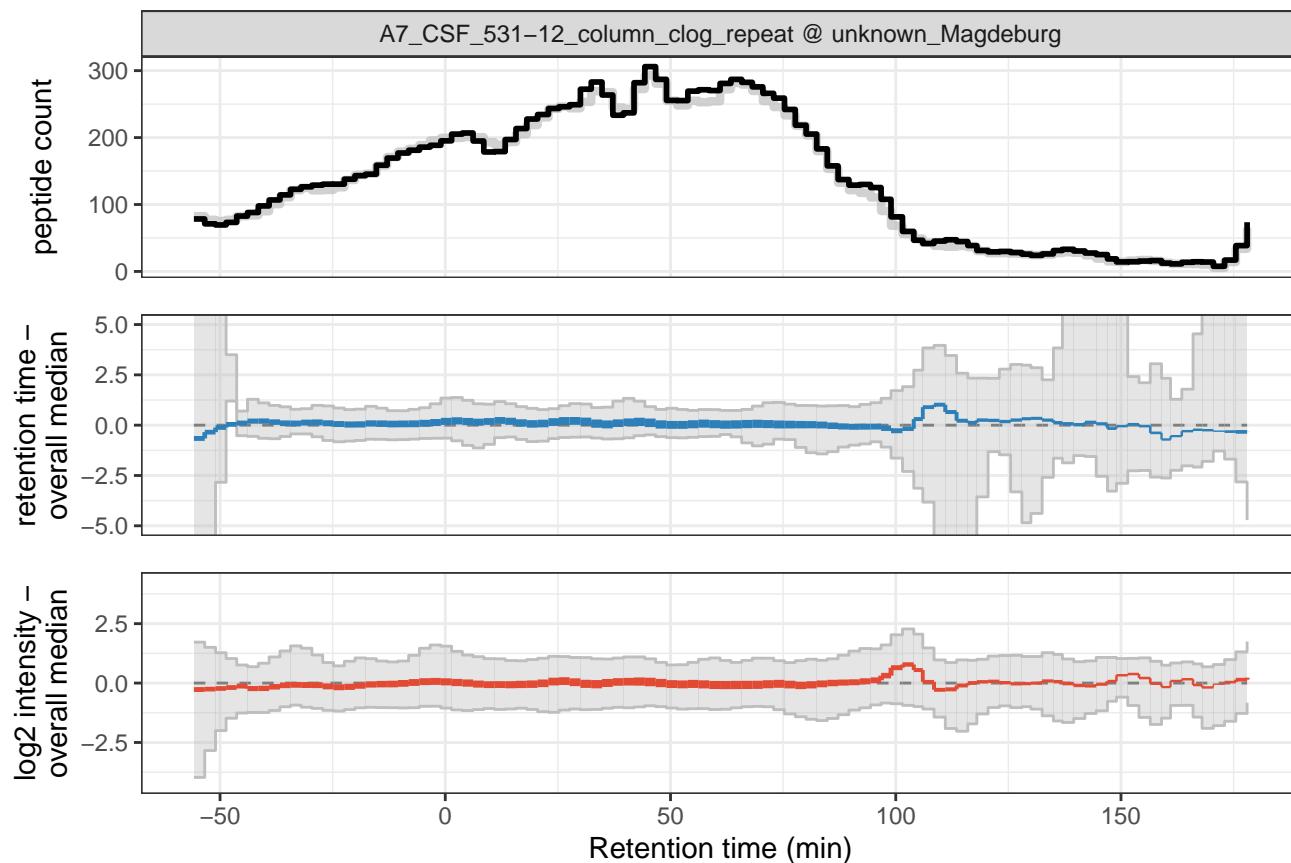
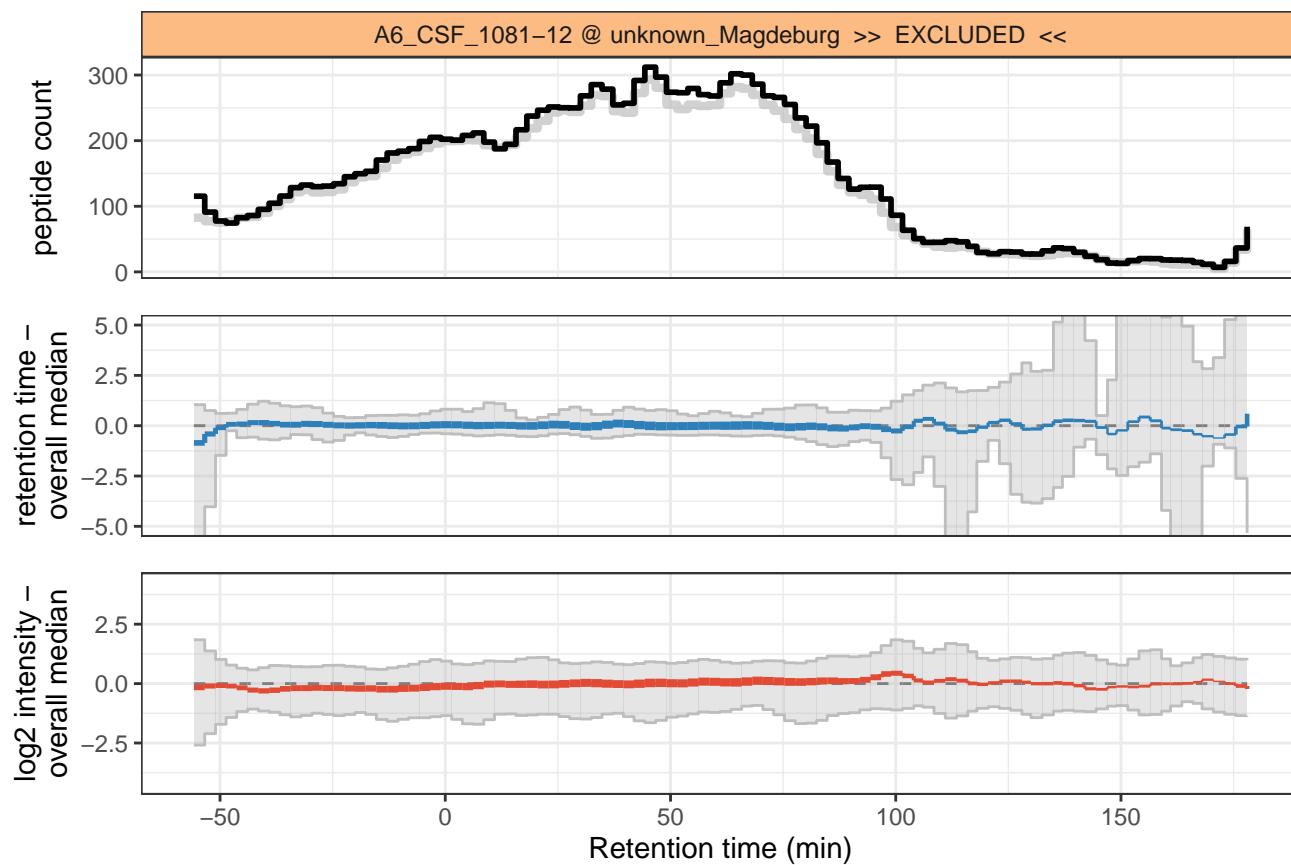


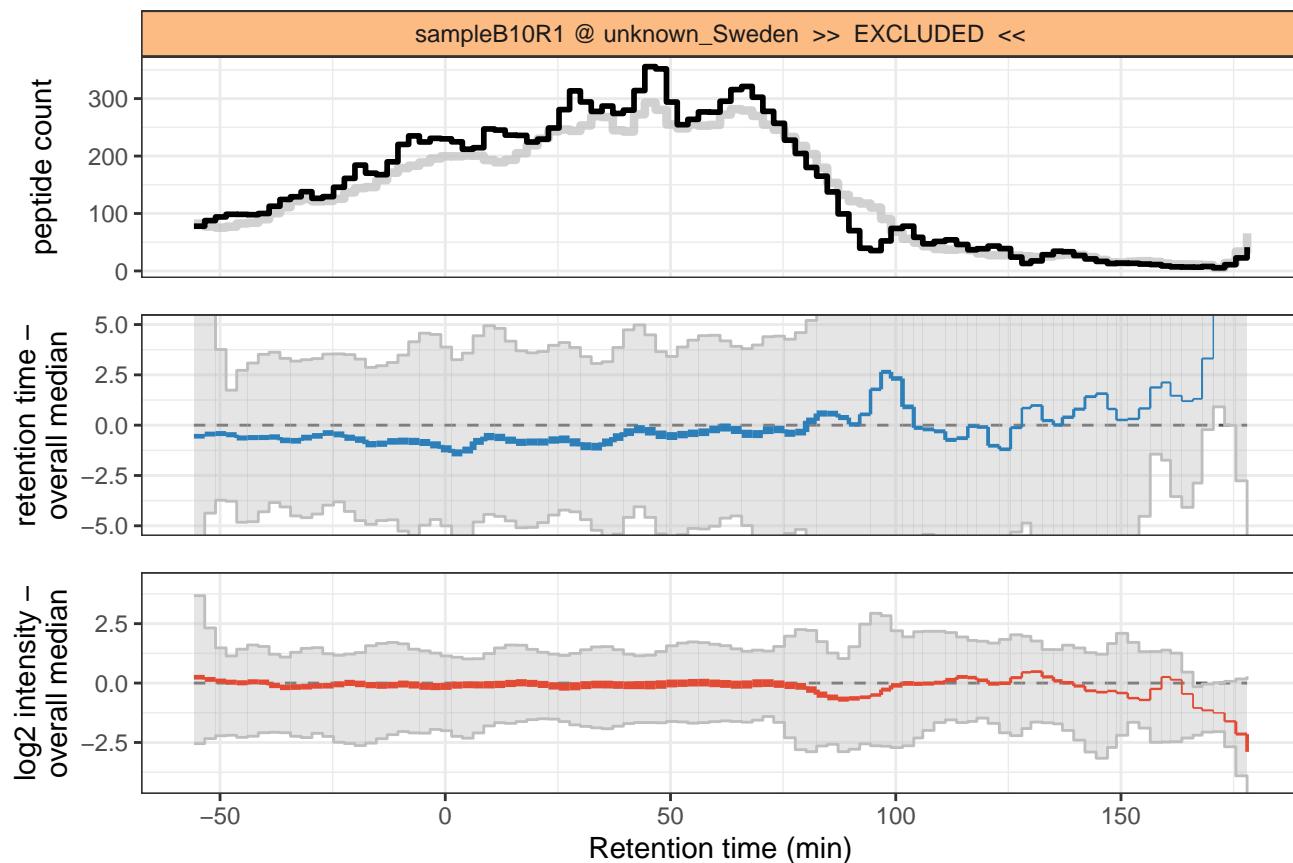
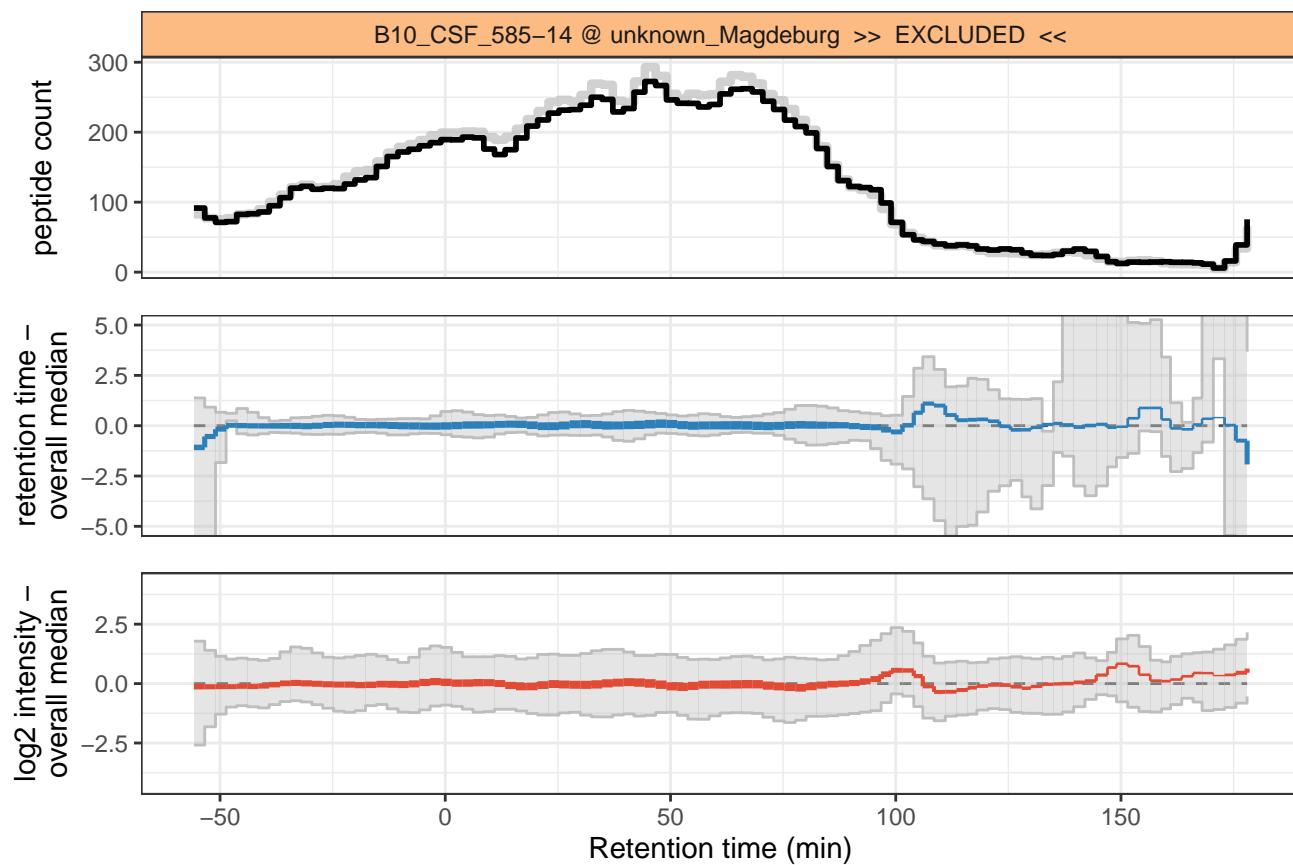












1.6 variation among replicates

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).

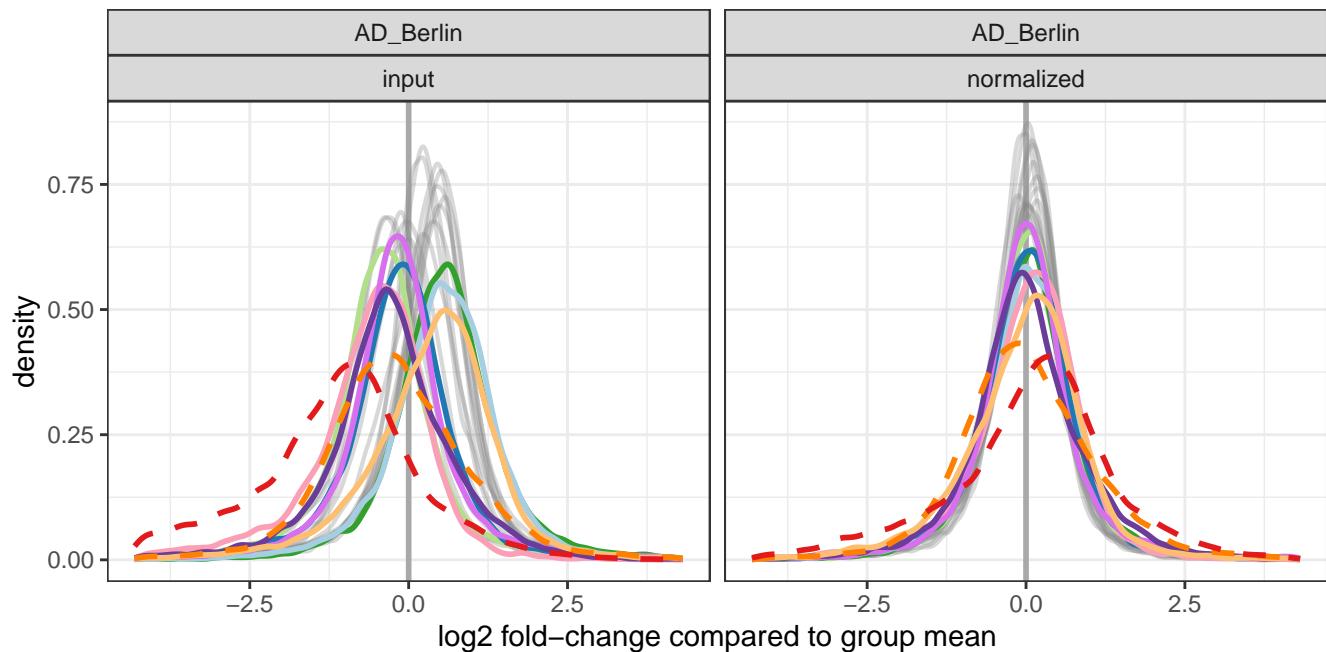
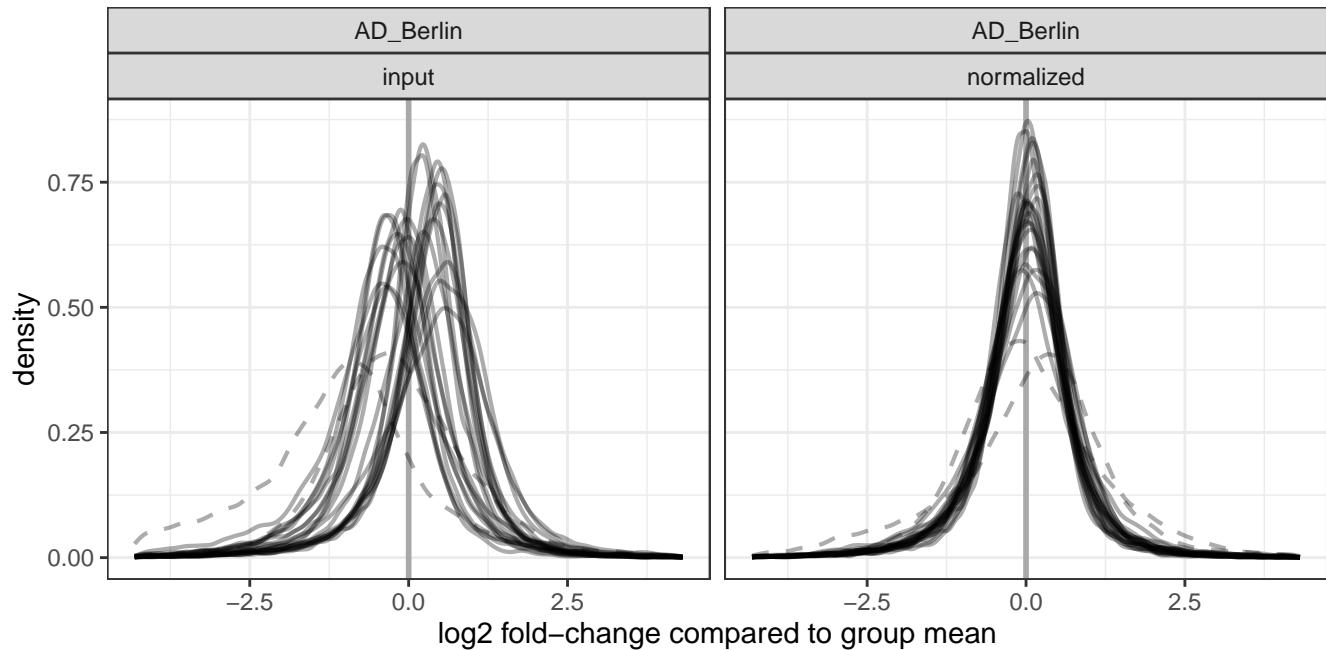
1.6.1 within-group foldchange distributions

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.

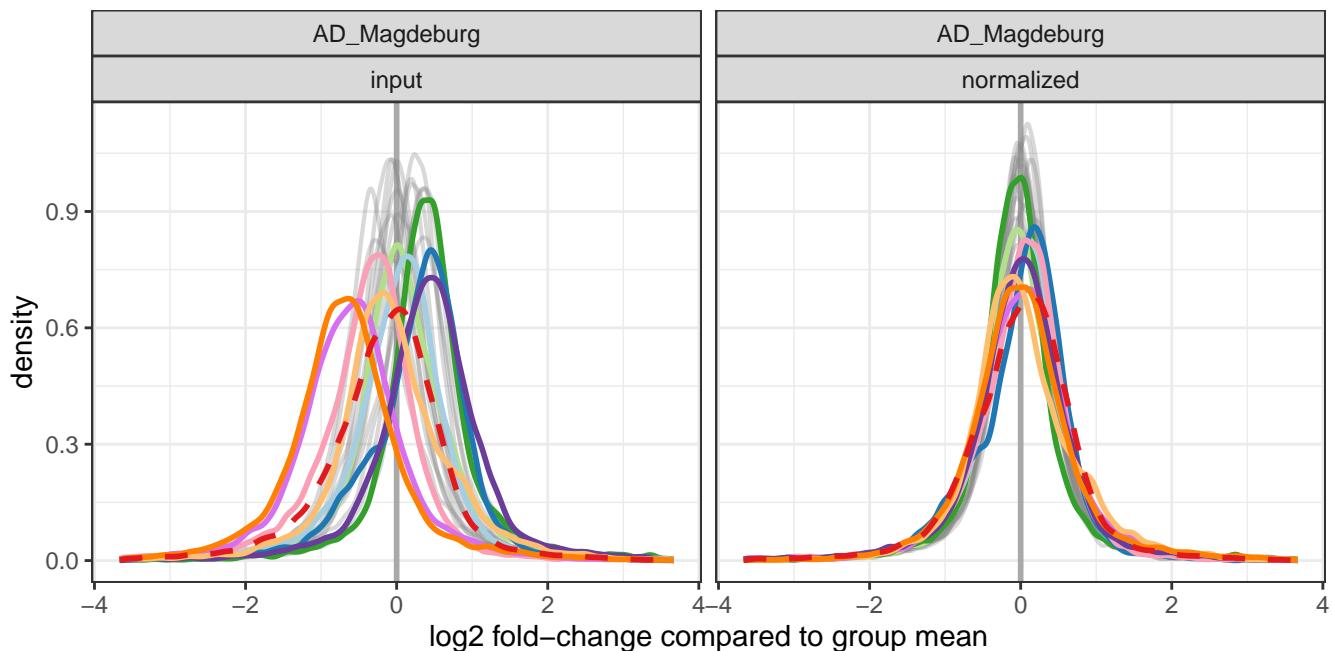
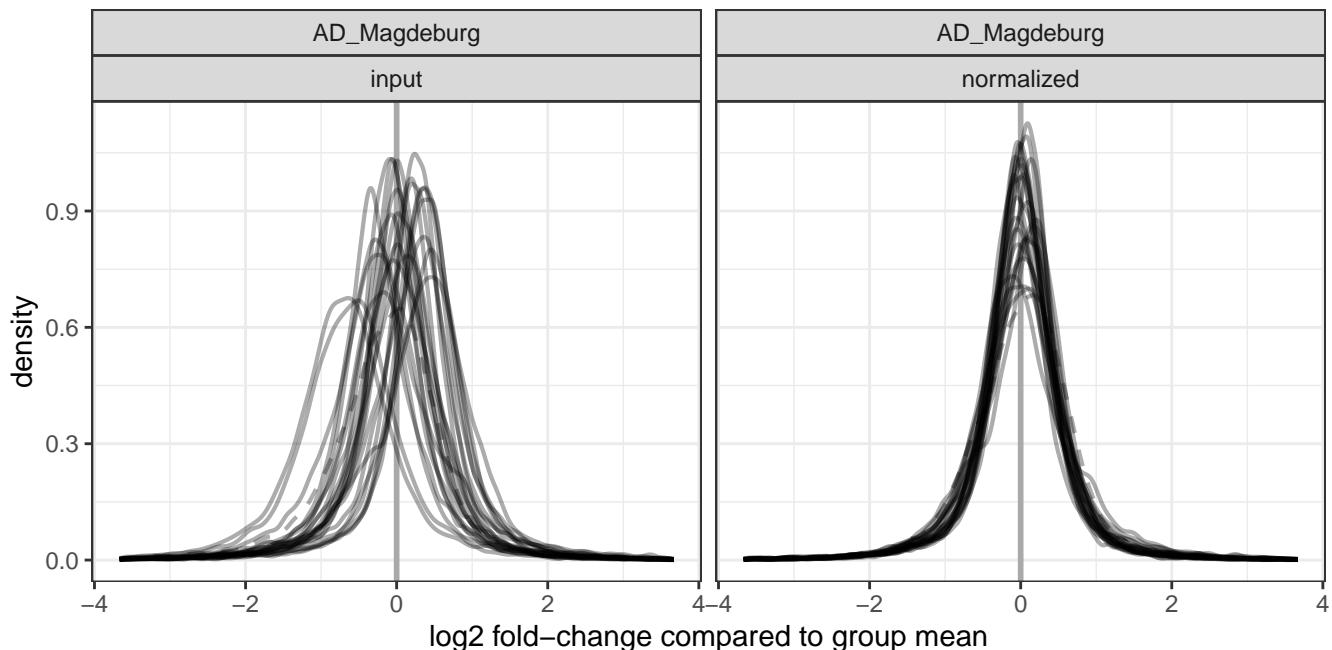
For each sample group, two plots are shown: 1) a basic monochrome plot and 2) a variant that color-codes the top10 ‘worst’ samples based on the standard deviation (sd) of their respective distributions. Note; per sample, the 0.5% quantiles of both top- and bottom-most outliers are disregarded in sd computation. The legend is sorted in column-first descending order (sample with highest sd in column 1 row 1, sample with second highest sd in column 1 row 2, etc.). If there are 10 or fewer samples, all are color-coded. On the pages after the plots, these sd values are shown as tables.

The ‘input’ panel is based on the peptide intensities as-is (i.e. the user-provided input data from upstream software), the ‘normalized’ panel shows the exact same samples and peptides after normalization (as specified by user). Samples marked as ‘exclude’ in the provided sample metadata table are visualized as dashed lines.

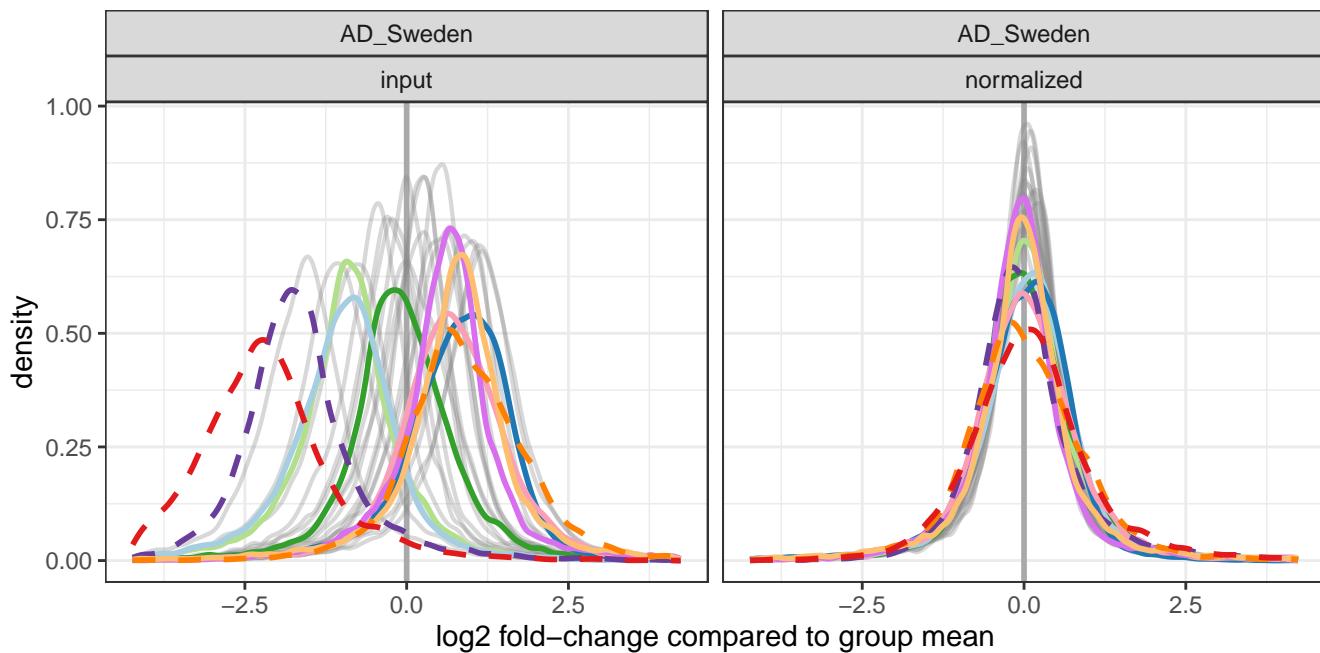
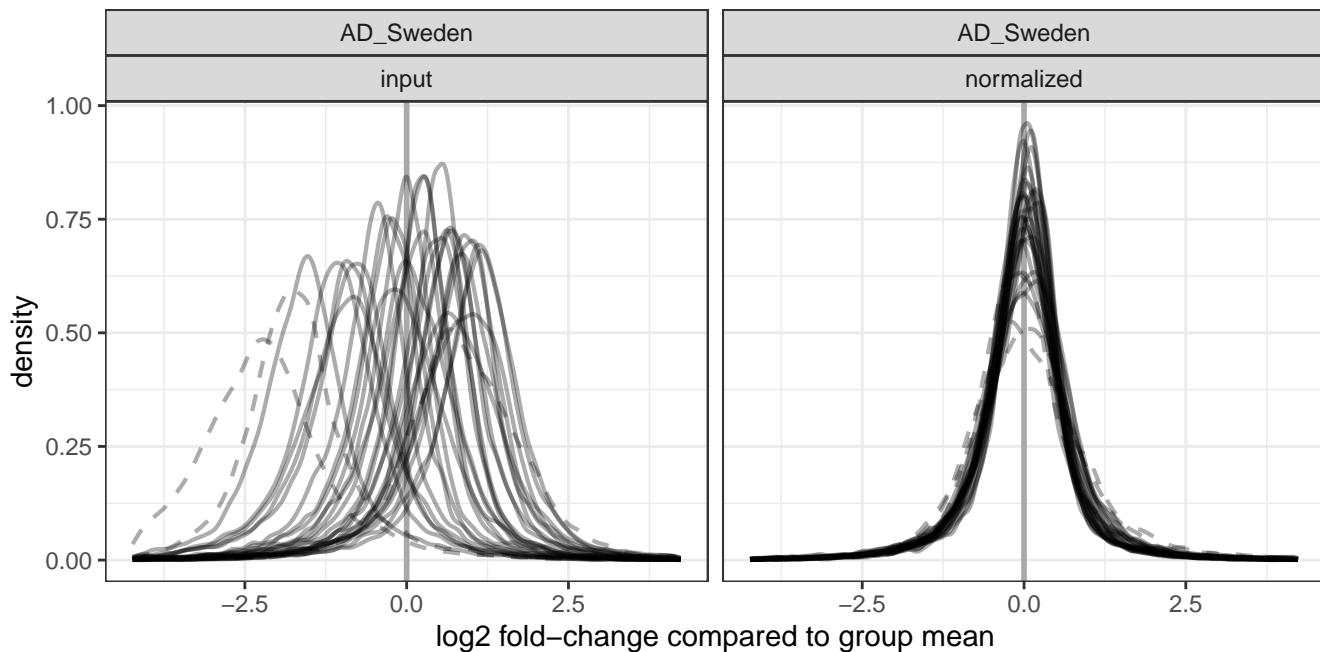
info: skipping within-group foldchange plots for sample group 'unknown_Kiel', require at least 2 replicates info:
info: skipping within-group foldchange plots for sample group 'unknown_Sweden', require at least 2 replicates



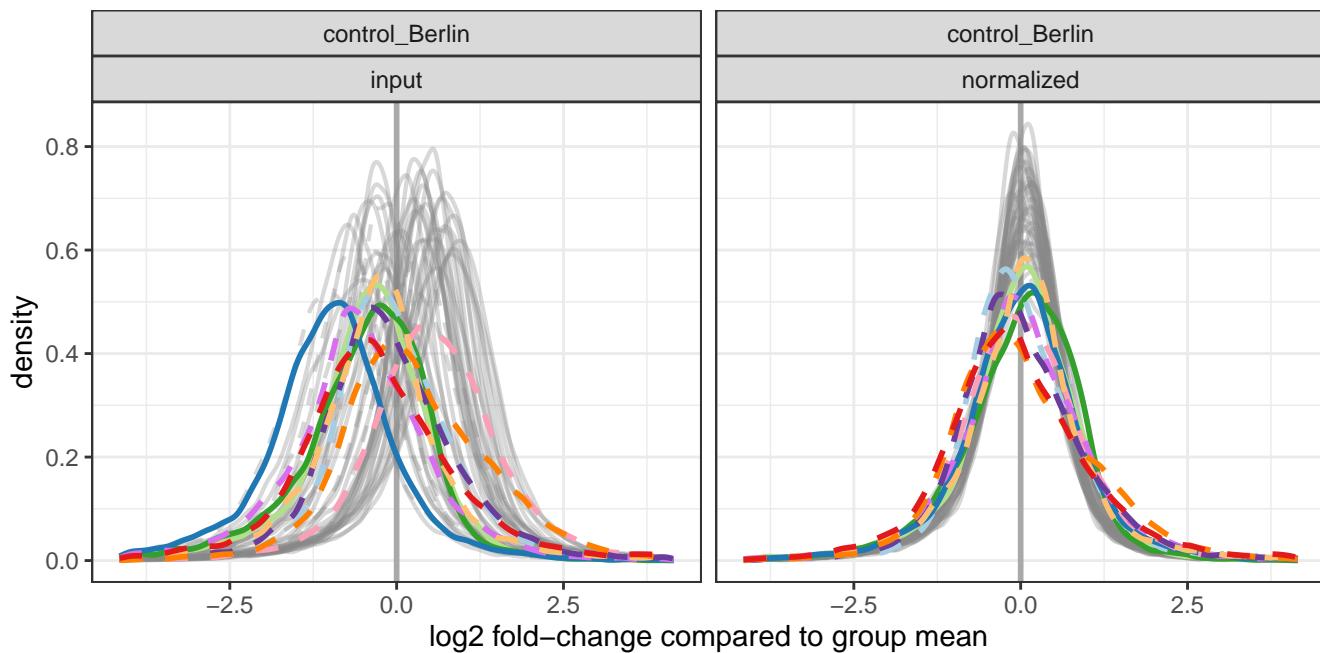
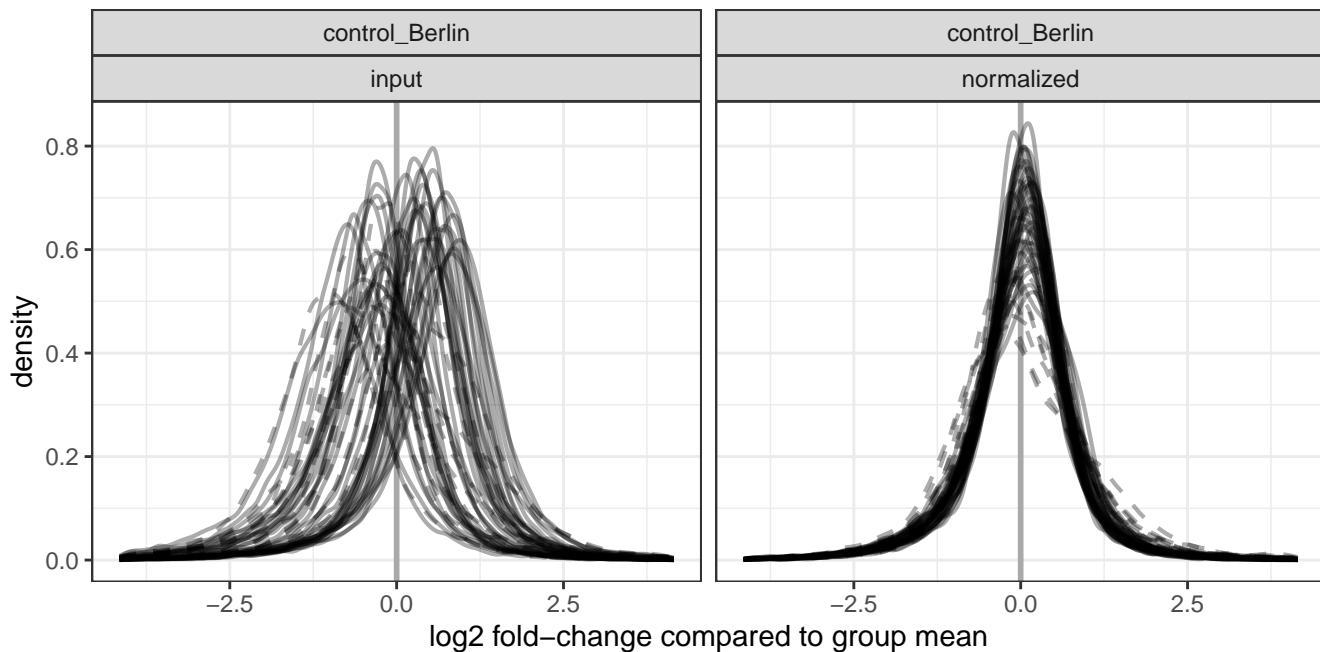
- | | |
|------------------------|----------------|
| — E4_190225869_repeat2 | — B2_79114832 |
| — G1_82881418_repeat | — A2_82872920 |
| — F3_129190664 | — E2_129189650 |
| — A3_82883165_repeat2 | — G3_79069638 |
| — D1_82545185 | — D3_129189776 |



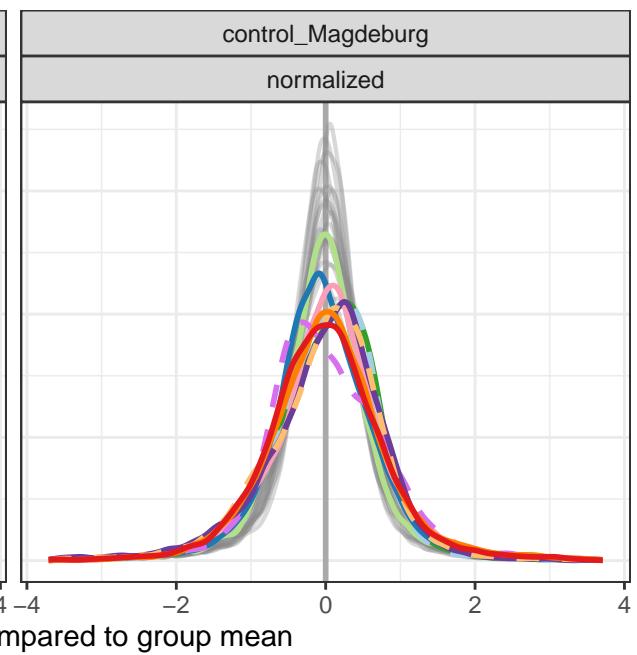
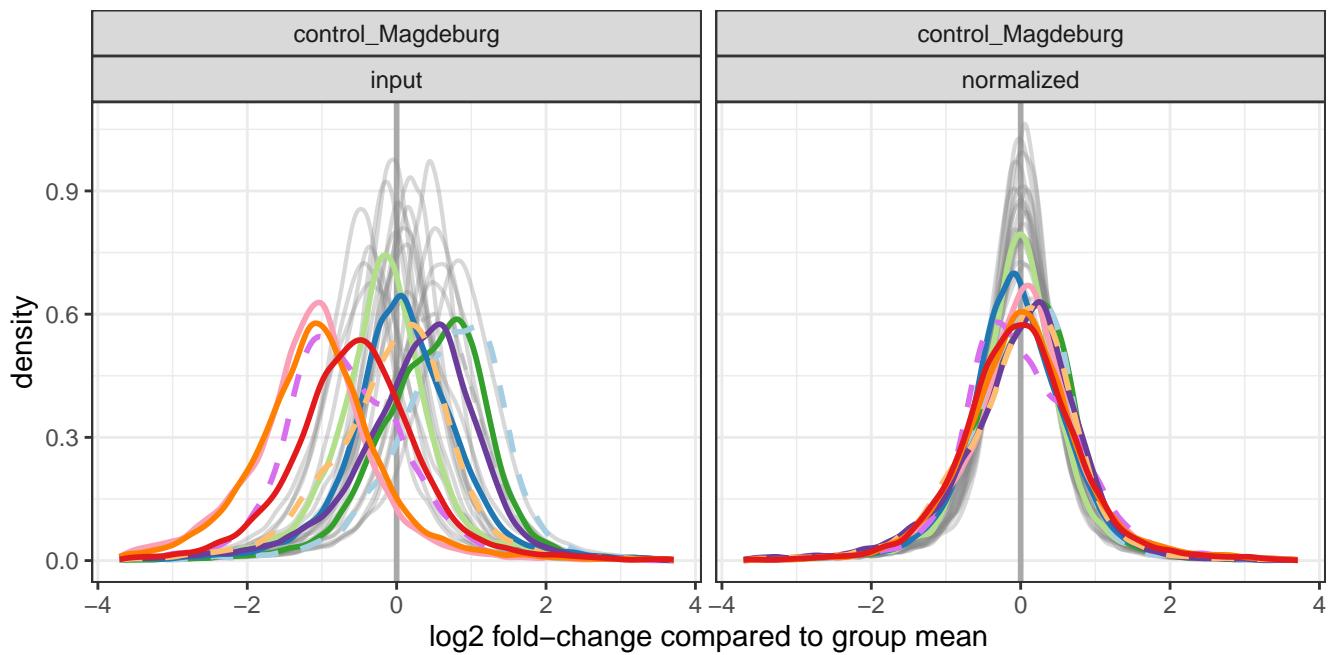
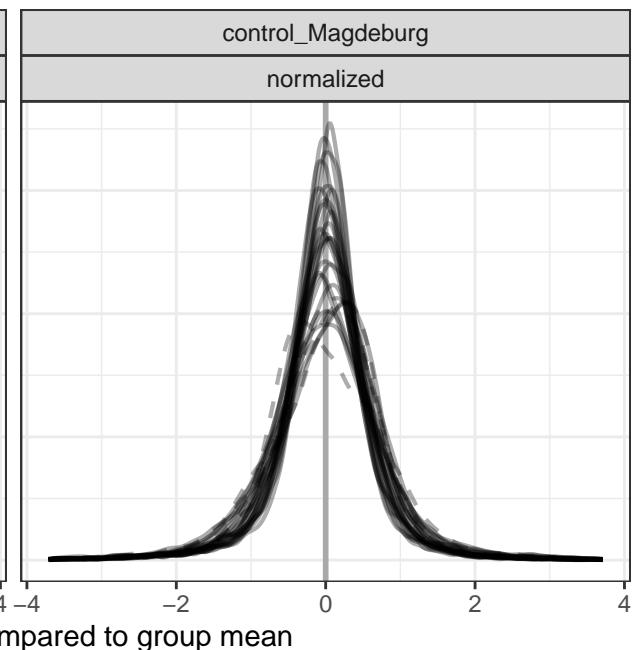
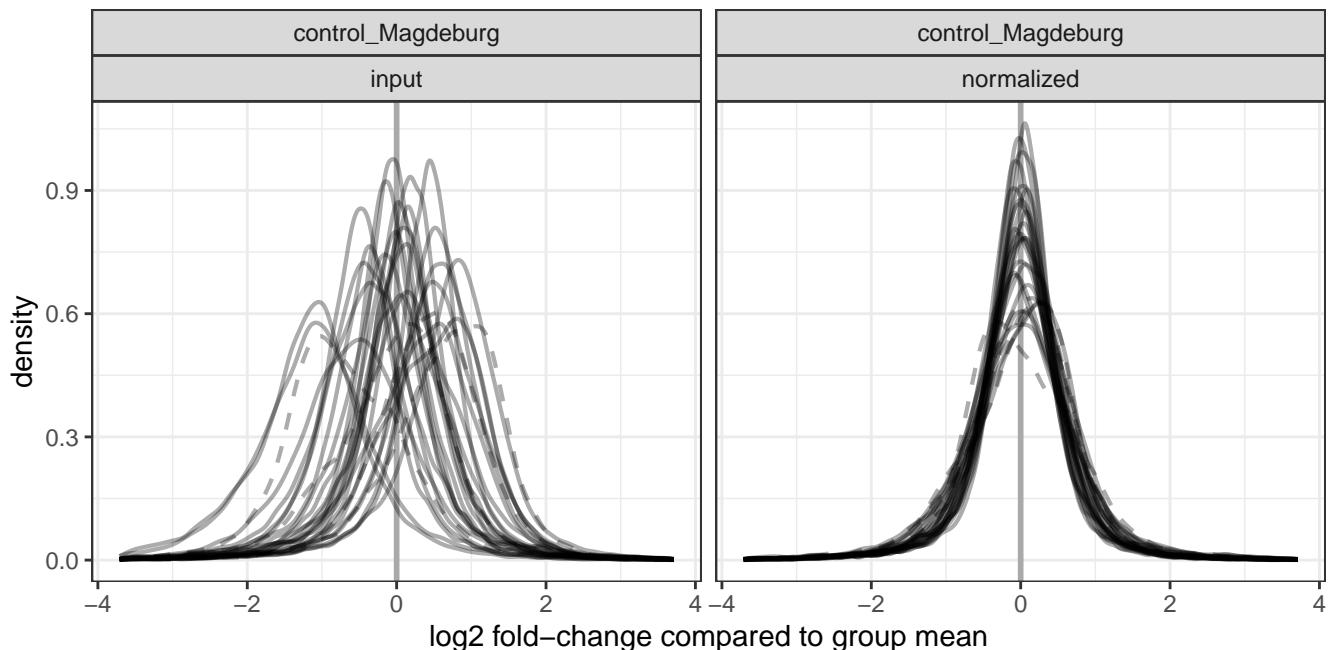
- D2_CSF_398-13
- C10_CSF_1289-14
- B12_CSF_819-14_190217150250
- A12_CSF_949-15
- C4_CSF_976-13
- A2_CSF_248-15
- D1_CSF_530-13
- A11_CSF_36-15
- C7_CSF_381-14
- D4_CSF_522-13



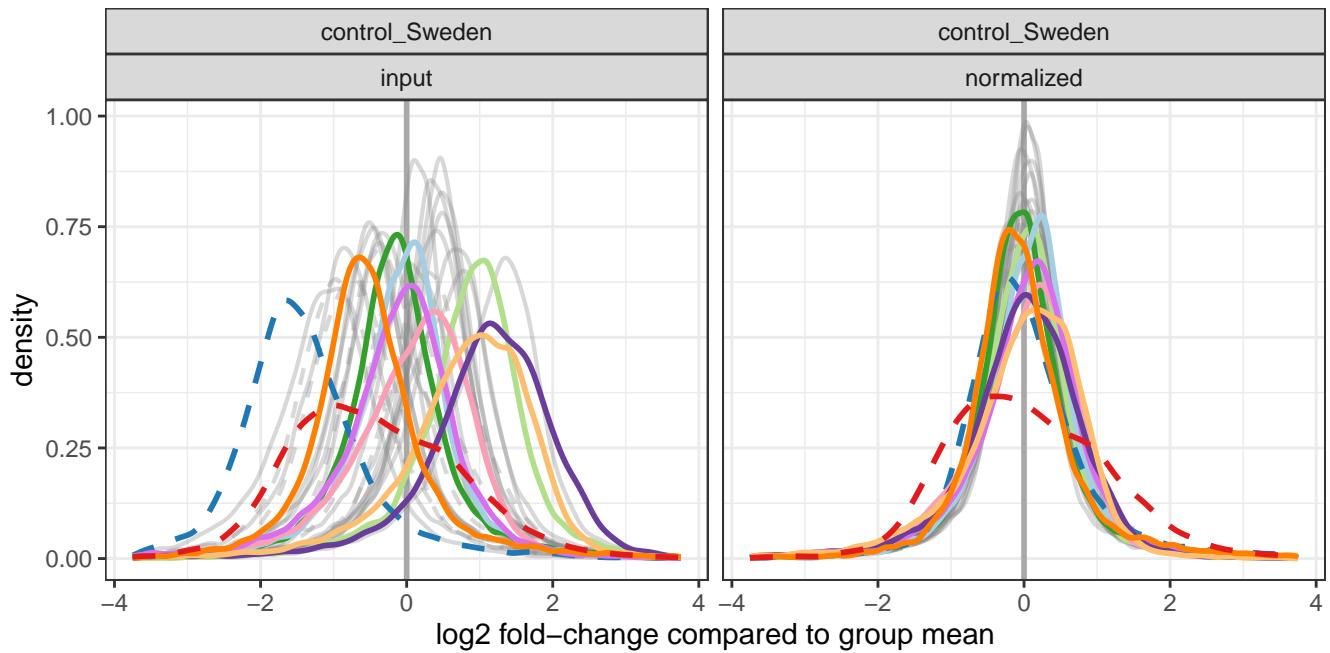
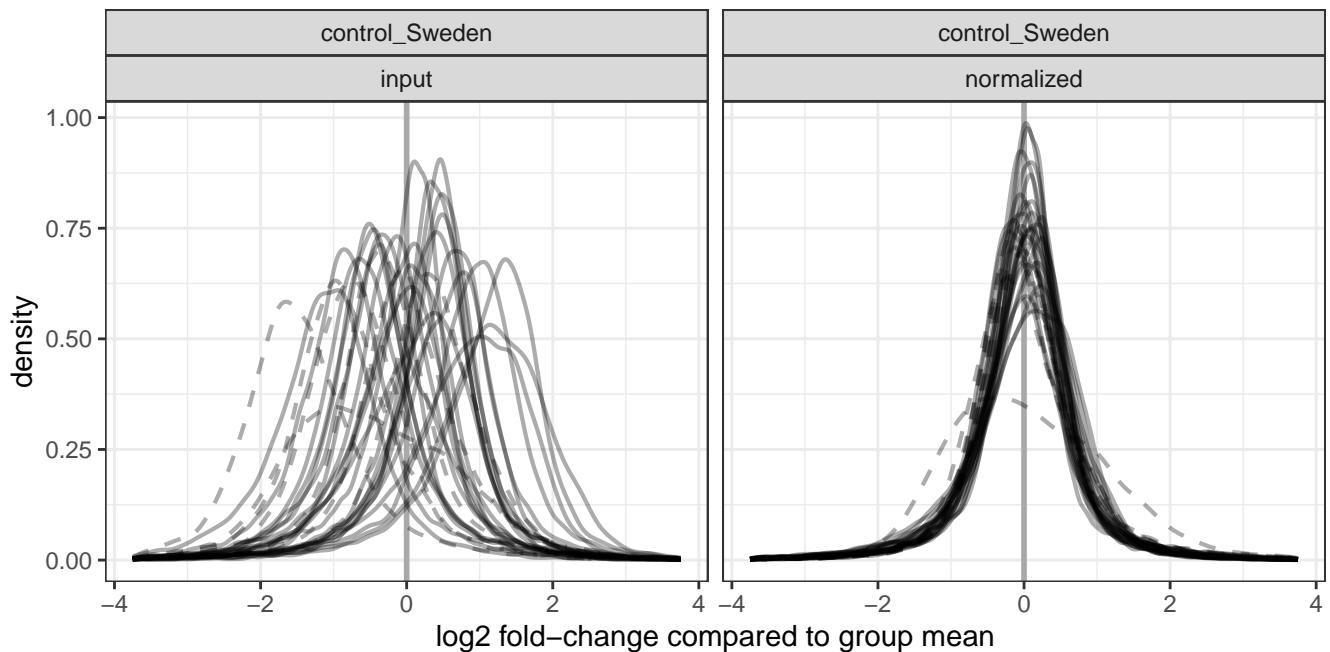
- sampleB04 — sampleE11
- sampleD09 — sampleB07
- sampleA06 — sampleB11
- sampleB03 — sampleD08
- sampleB10R2 — sampleE09



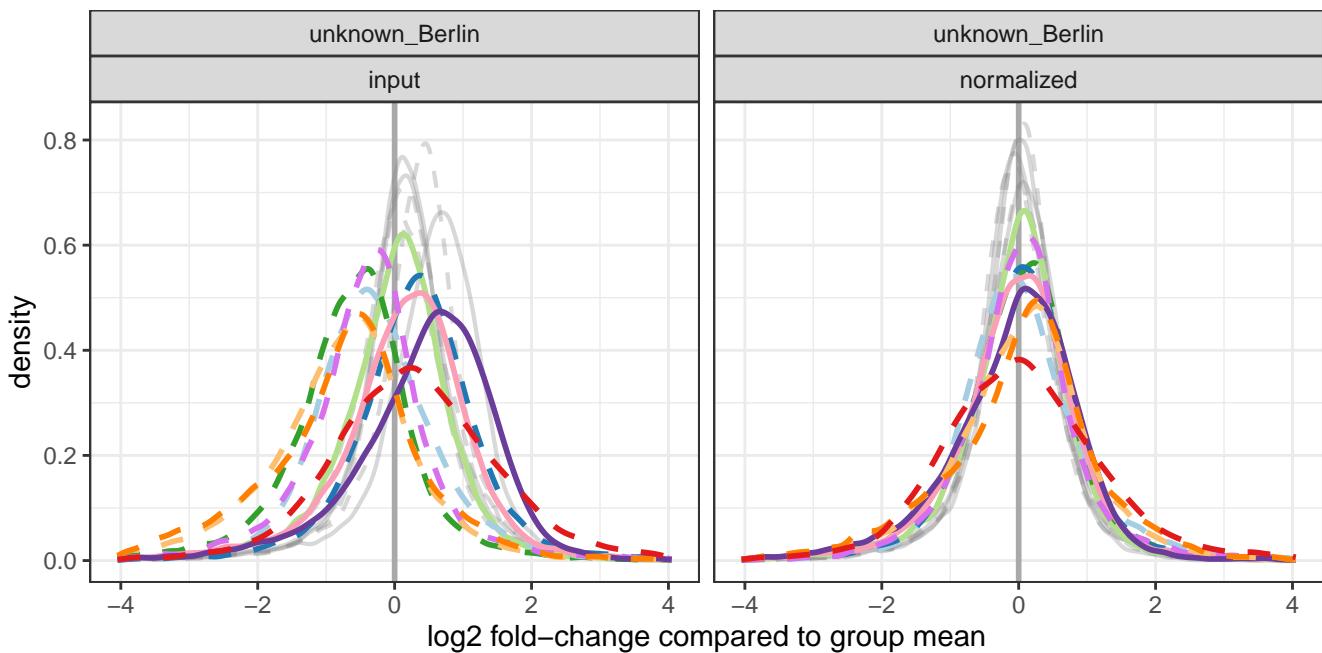
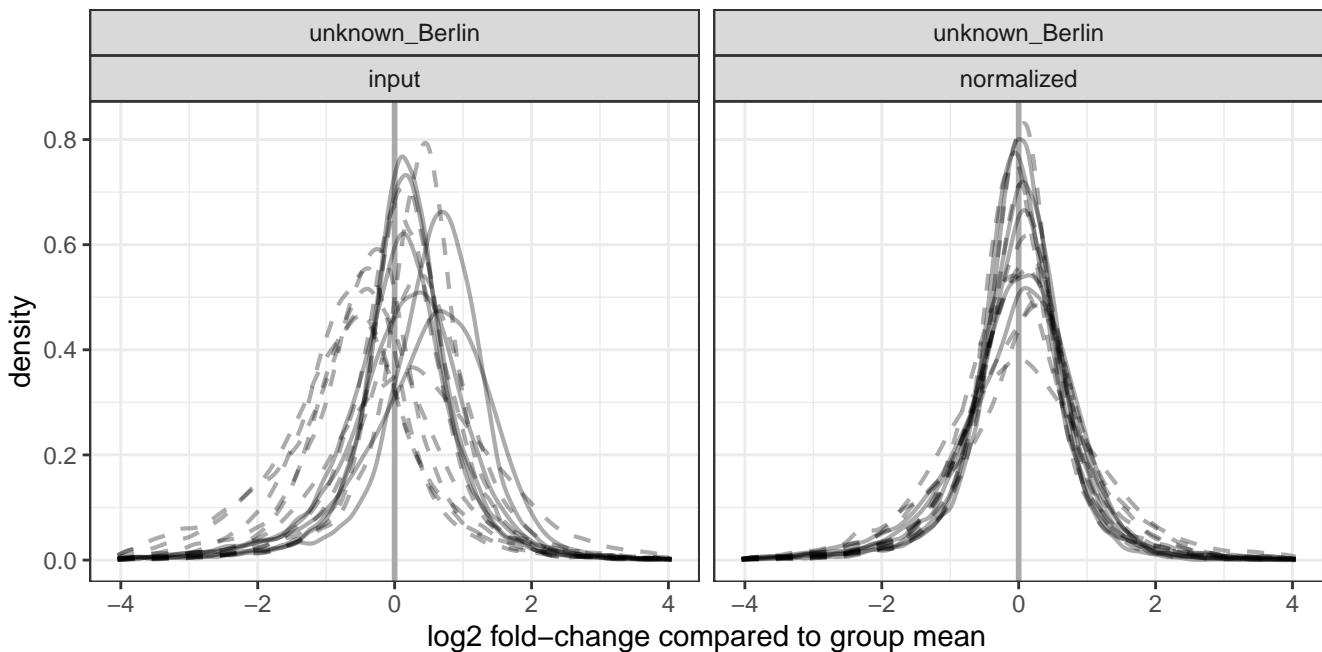
— G9_82879903_190312181946 — A12_190103450 — G7_79061594 — C5_81760601 — D7_81784062	— F12_190100162 — C7_81790598_repeat — B8_79116161 — C11_129190335 — F5_129184099_repeat
--	--



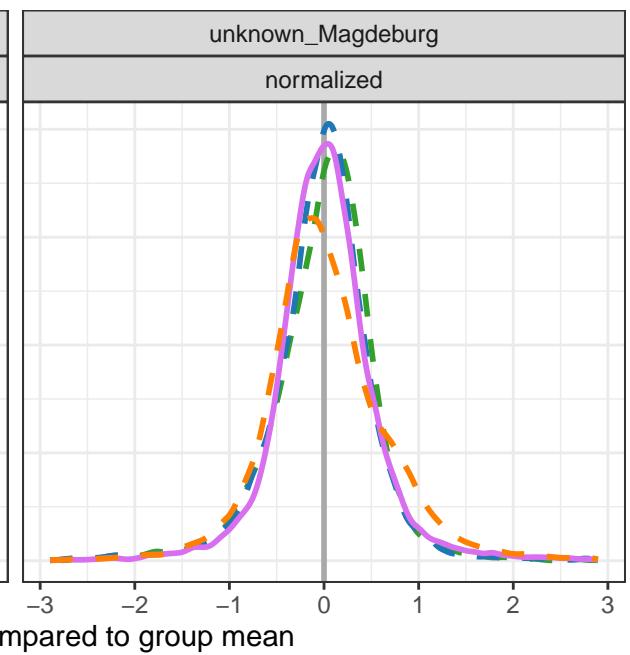
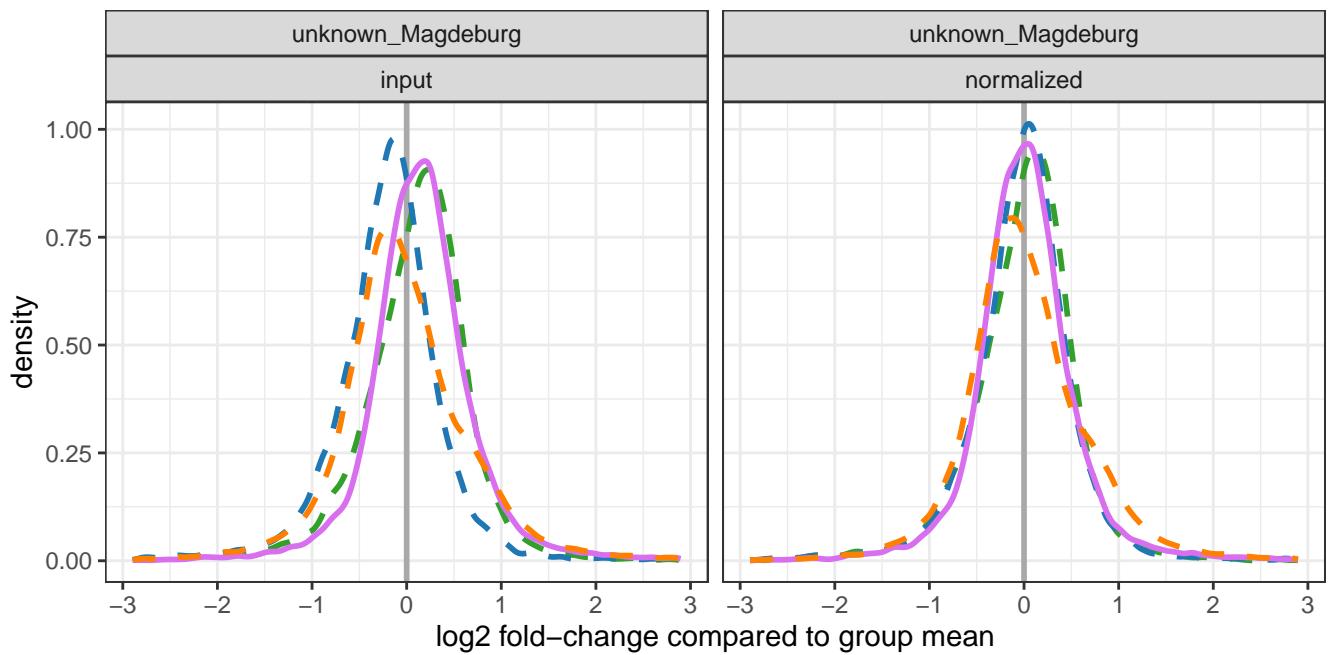
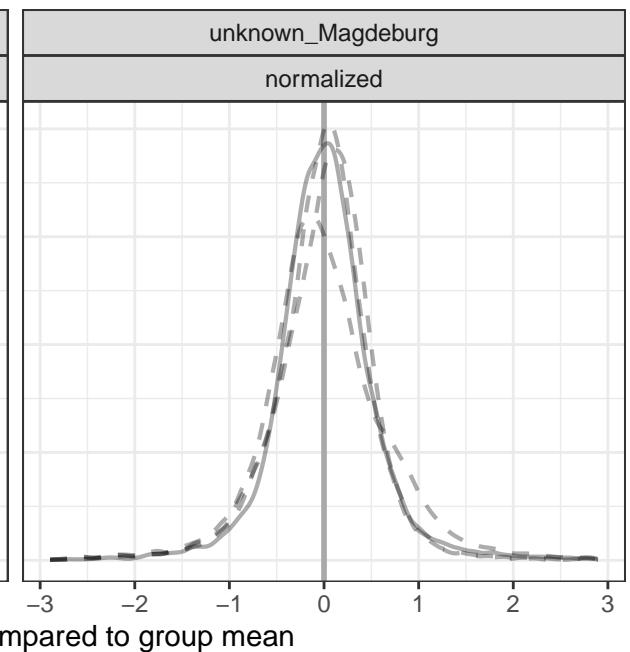
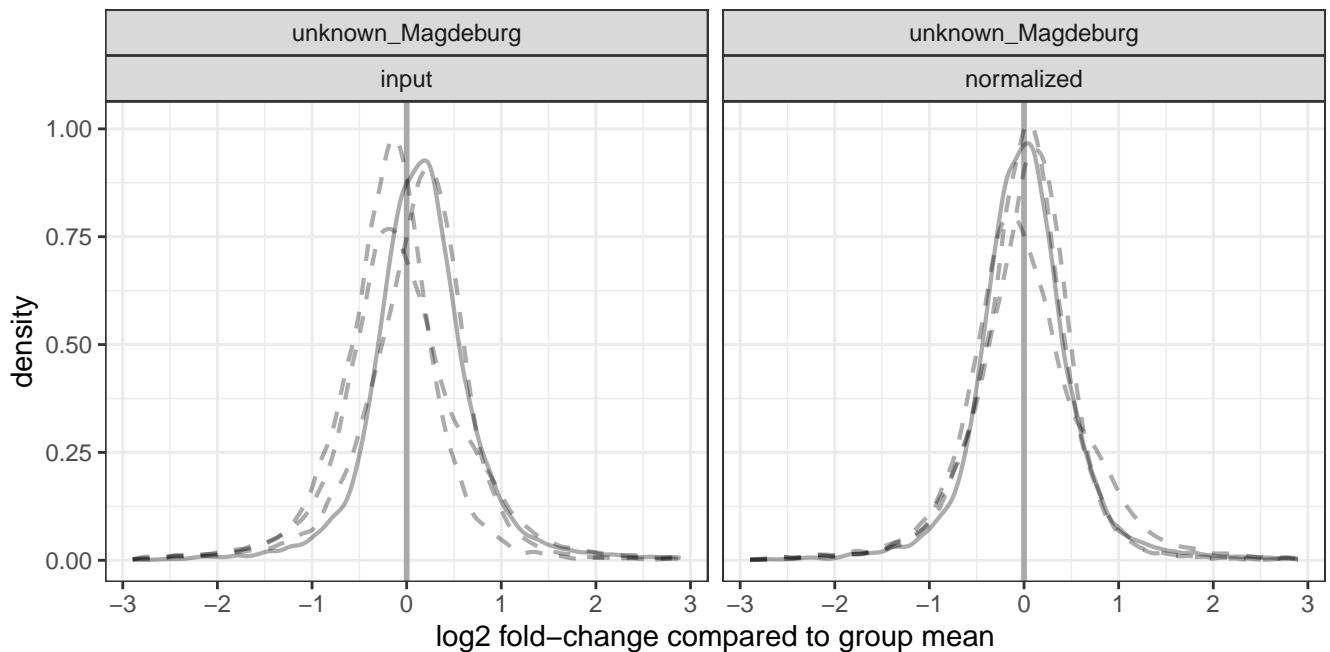
- | | |
|---------------|---------------|
| E1_CSF_59 | E4_CSF_761 |
| E6_CSF_960 | E12_CSF_146 |
| F7_CSF_142 | A3_CSF_358-15 |
| B4_CSF_582-14 | A9_CSF_695-15 |
| E10_CSF_16 | E2_CSF_396 |



—	sampleC09	—	sampleA01
—	sampleD02	—	sampleC01
—	sampleD07	—	sampleE01
—	sampleC06	—	sampleC02
—	sampleE03	—	sampleD06



— F2_190226445 — C4_168751222 — B7_81787765 — C8_82882588_repeat — B5_81783170	— A1_81776353_repeat2 — G5_82544653 — H1_82881356 — A6_82545313 — D11_129188366_repeat
---	---



— B10_CSF_585-14 — A7_CSF_531-12_column_clog_repeat	— A6_CSF_1081-12 — A5_CSF_660-12
---	--

top10 outliers per sample group

Group	Sample	Standard Deviation
AD_Berlin	F4_190100053	0.729
AD_Berlin	D4_168751414	0.728
AD_Berlin	E3_82871060_repeat	0.727
AD_Berlin	B1_129186923	0.703
AD_Berlin	G4_190101656_repeat	0.695
AD_Berlin	D2_79110561	0.687
AD_Berlin	A4_190104103	0.668
AD_Berlin	B3_79113129	0.657
AD_Berlin	G2_168751345	0.656
AD_Berlin	C3_82874751	0.635

Group	Sample	Standard Deviation
AD_Magdeburg	B11_CSF_506-14	0.589
AD_Magdeburg	D5_CSF_503-13	0.578
AD_Magdeburg	C3_CSF_974-14	0.570
AD_Magdeburg	B7_CSF_496-15	0.570
AD_Magdeburg	C6_CSF_841-14	0.567
AD_Magdeburg	C1_CSF_656-15	0.564
AD_Magdeburg	D3_CSF_863-13	0.563
AD_Magdeburg	C2_CSF_1239-14	0.556
AD_Magdeburg	D6_CSF_864-13	0.553
AD_Magdeburg	C8_CSF_840-14	0.535

Group	Sample	Standard Deviation
AD_Sweden	sampleA02	0.715
AD_Sweden	sampleA12	0.705
AD_Sweden	sampleE10	0.699
AD_Sweden	sampleA11	0.690
AD_Sweden	sampleA04	0.680
AD_Sweden	sampleE08	0.674
AD_Sweden	sampleE12	0.674
AD_Sweden	sampleB09	0.673
AD_Sweden	sampleA09	0.659
AD_Sweden	sampleC05	0.628

Group	Sample	Standard Deviation
control_Berlin	C12_168743858_repeat	0.727
control_Berlin	H5_82544762_190309202008	0.723
control_Berlin	D10_190224386	0.718
control_Berlin	G6_168744541	0.715
control_Berlin	A11_129183842	0.710
control_Berlin	C6_82881358	0.710
control_Berlin	D9_82885545_repeat	0.703
control_Berlin	B10_79069372	0.703
control_Berlin	G12_190225664	0.687
control_Berlin	F8_190224717_repeat	0.680

Group	Sample	Standard Deviation
control_Magdeburg	B1_CSF_326-15	0.640
control_Magdeburg	C12_CSF_1299-14	0.624
control_Magdeburg	A10_CSF_201-15	0.614
control_Magdeburg	F2_CSF_755	0.612
control_Magdeburg	F9_CSF_420	0.606
control_Magdeburg	C5_CSF_706-14	0.603
control_Magdeburg	F4_CSF_51	0.585
control_Magdeburg	F3_CSF_865	0.574
control_Magdeburg	F8_CSF_143	0.564
control_Magdeburg	B9_CSF_851-15	0.562

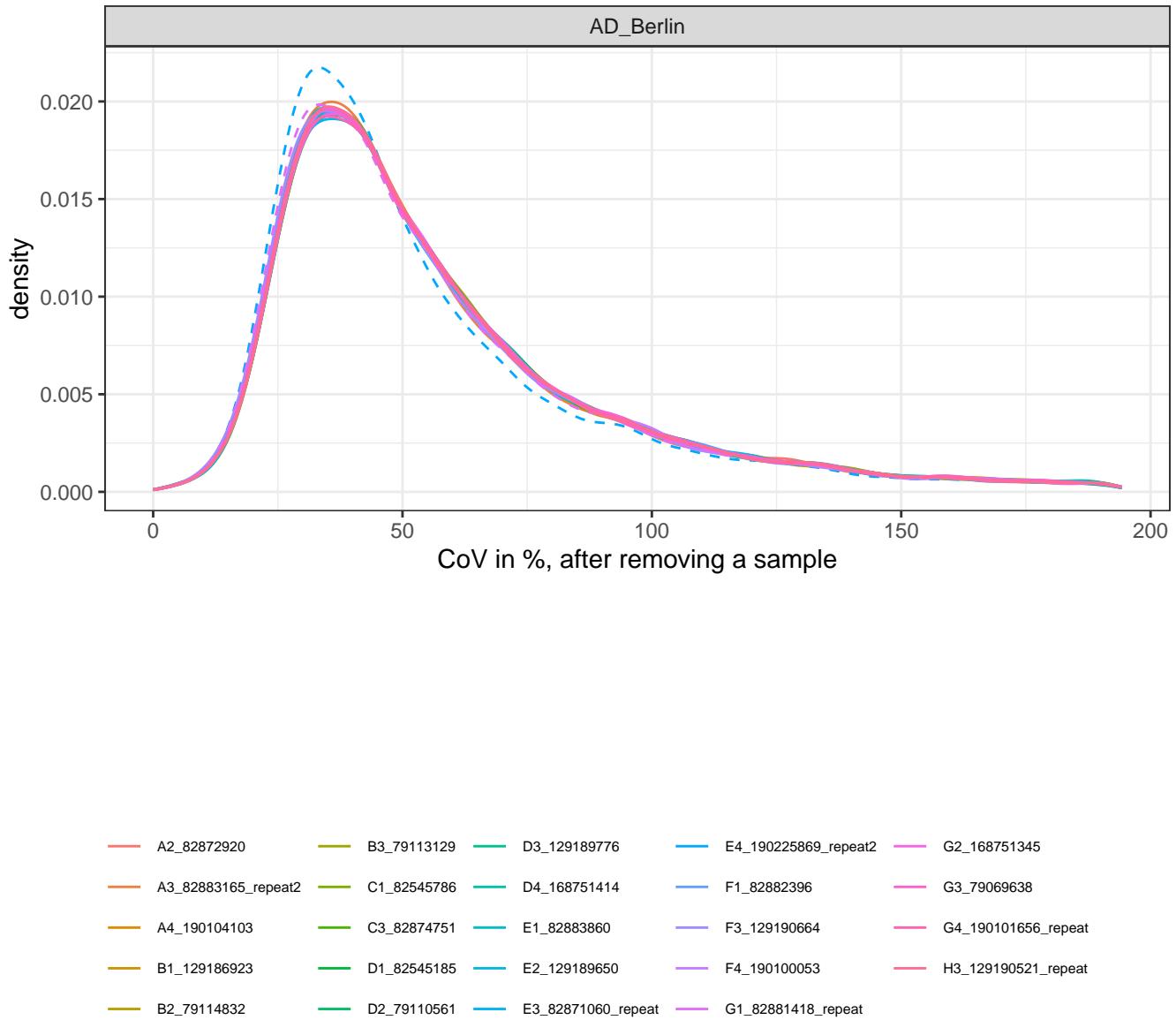
Group	Sample	Standard Deviation
control_Sweden	sampleD04	0.715
control_Sweden	sampleC12	0.712
control_Sweden	sampleD03	0.703
control_Sweden	sampleB06	0.696
control_Sweden	sampleC10	0.689
control_Sweden	sampleE04	0.686
control_Sweden	sampleE05	0.667
control_Sweden	sampleD10	0.662
control_Sweden	sampleC08	0.629
control_Sweden	sampleD12	0.623

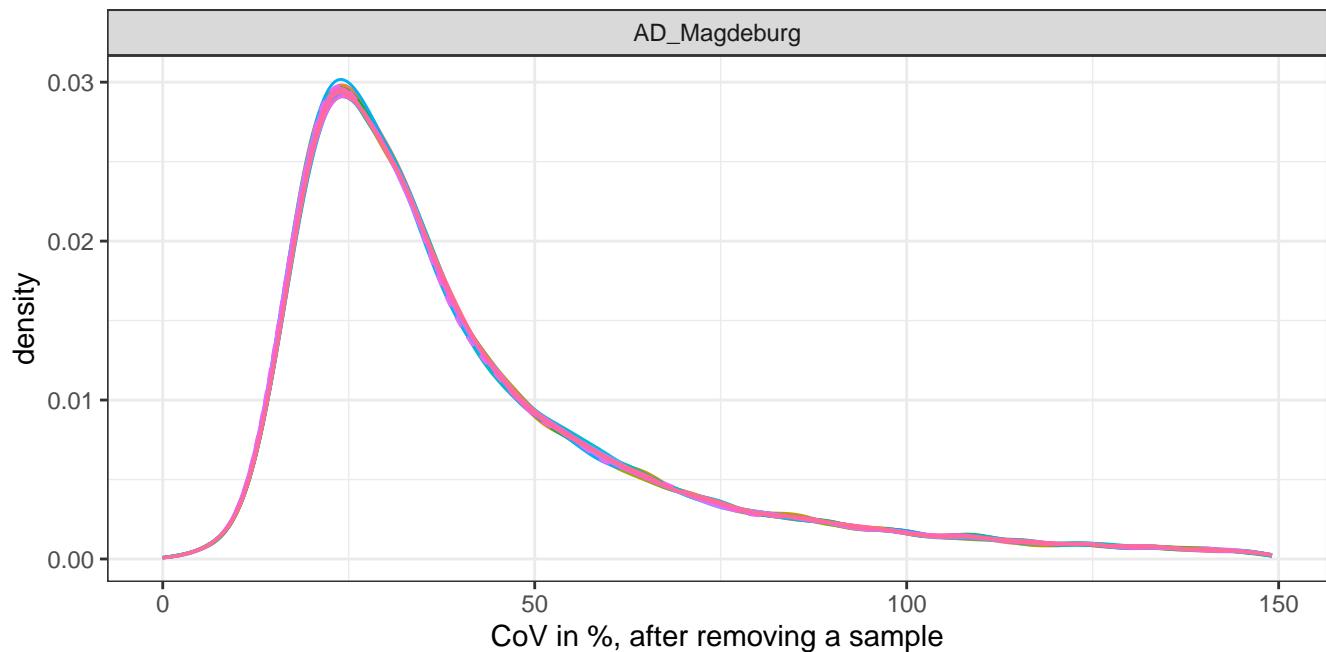
Group	Sample	Standard Deviation
unknown_Berlin	G5_82544653	0.867
unknown_Berlin	H1_82881356	0.865
unknown_Berlin	A6_82545313	0.840
unknown_Berlin	D11_129188366_repeat	0.800
unknown_Berlin	F9_82880110_190304154132	0.747
unknown_Berlin	B4_190223925	0.713
unknown_Berlin	F11_82542649	0.710
unknown_Berlin	D5_129186580_repeat	0.693
unknown_Berlin	B11_129182819	0.662
unknown_Berlin	C2_79070601	0.662

Group	Sample	Standard Deviation
unknown_Magdeburg	B10_CSF_585-14	0.632
unknown_Magdeburg	A7_CSF_531-12_column_clog_repeat	0.522
unknown_Magdeburg	A6_CSF_1081-12	0.519
unknown_Magdeburg	A5_CSF_660-12	0.519

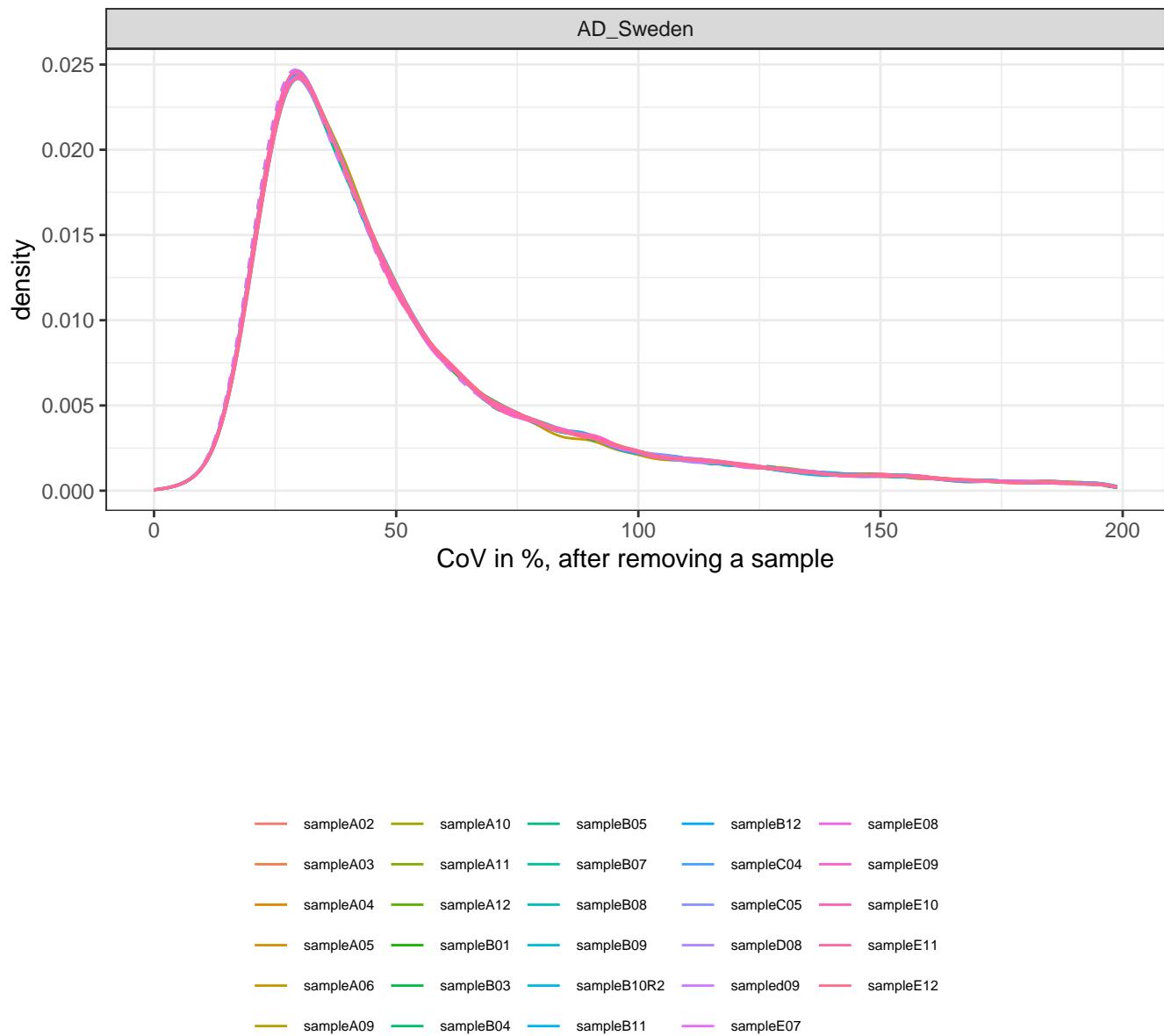
1.6.2 CoV, leave-one-out

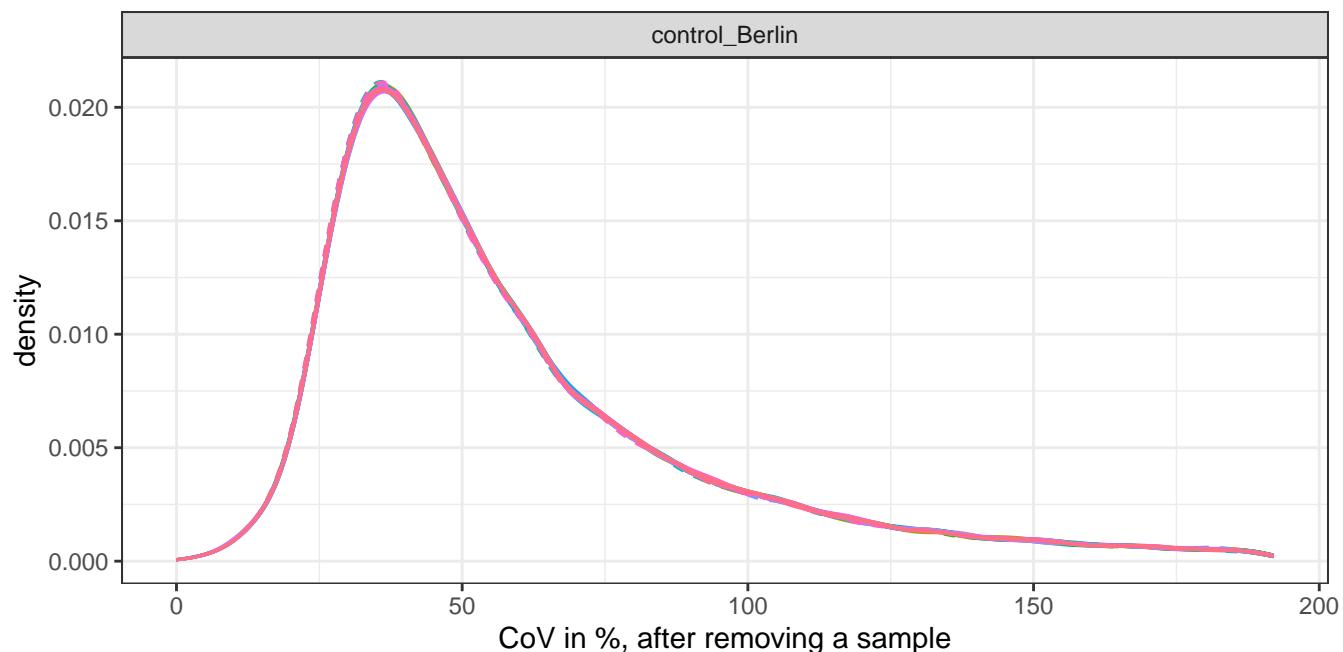
The figures below describe the effect of removing a particular sample prior to within-group Coefficient of Variation (CoV) computation. The lower the CoV distribution is for a sample, the better reproducibility we get by excluding it. Only sample groups with at least 4 replicates can be used for this analysis, so 3 samples remain after leaving one out. Samples marked as ‘exclude’ in the provided sample metadata are included in these analyses (shown as dashed lines), and only peptides with at least 3 data points across replicate samples (after leave-one-out) are used for each CoV computation.



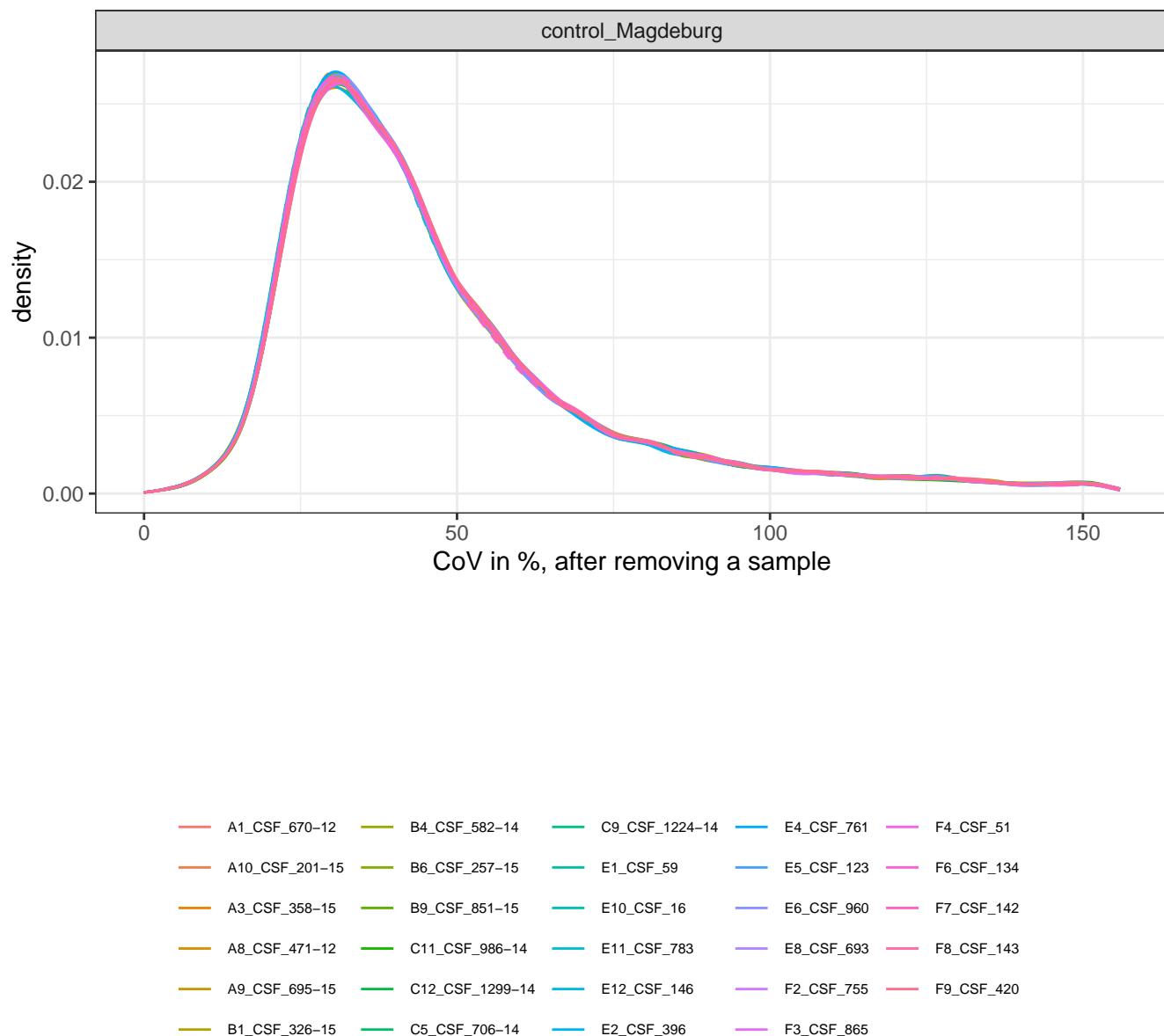


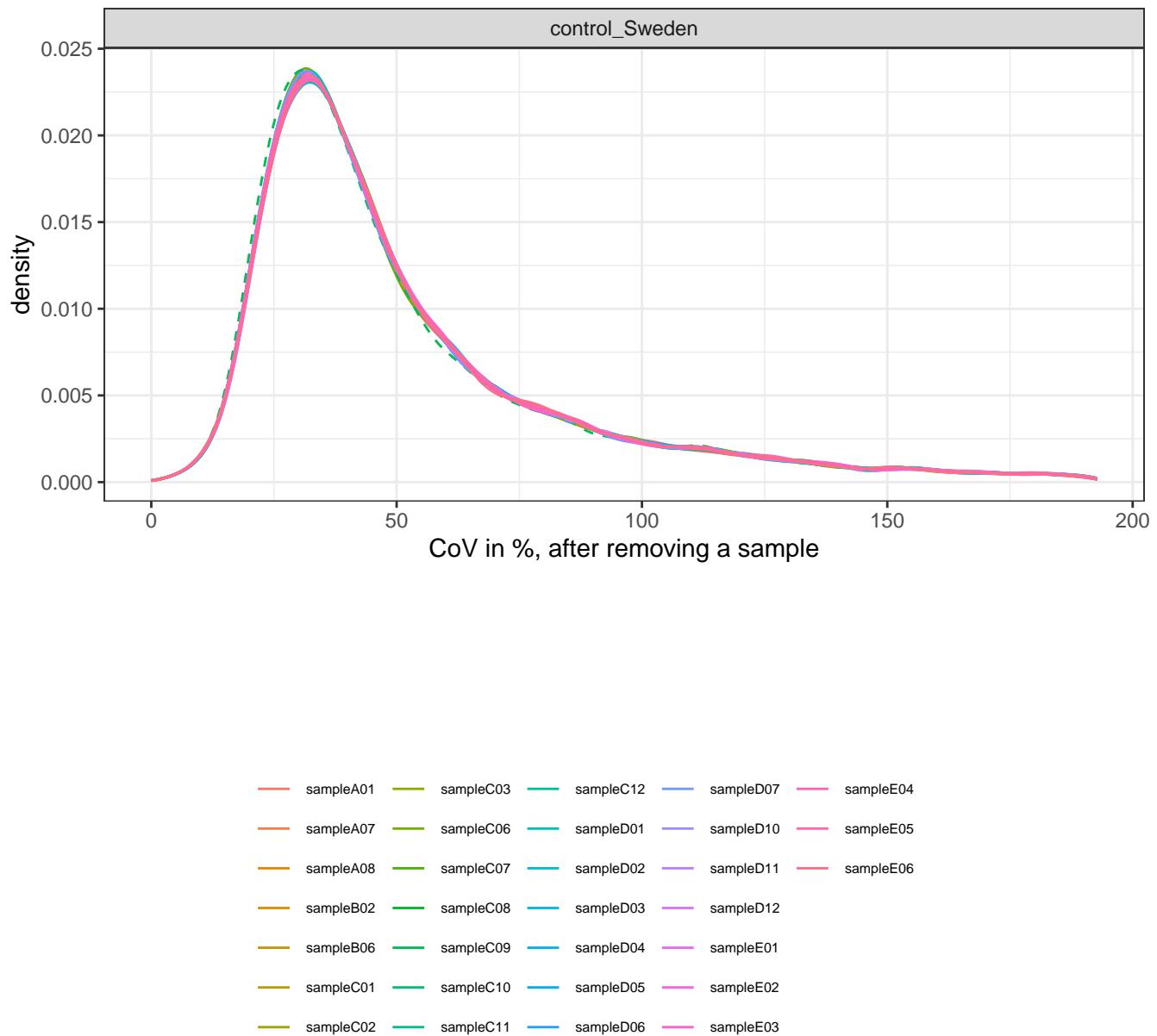
A11_CSF_36-15	B12_CSF_819-14_190217150250	B8_CSF_828-15	C4_CSF_976-13	D2_CSF_398-13
A12_CSF_949-15	B2_CSF_407-14	C1_CSF_656-15	C6_CSF_841-14	D3_CSF_863-13
A2_CSF_248-15	B3_CSF_545-14	C10_CSF_1289-14	C7_CSF_381-14	D4_CSF_522-13
A4_CSF_86-14	B5_CSF_885-15	C2_CSF_1239-14	C8_CSF_840-14	D5_CSF_503-13
B11_CSF_506-14	B7_CSF_496-15	C3_CSF_974-14	D1_CSF_530-13	D6_CSF_864-13

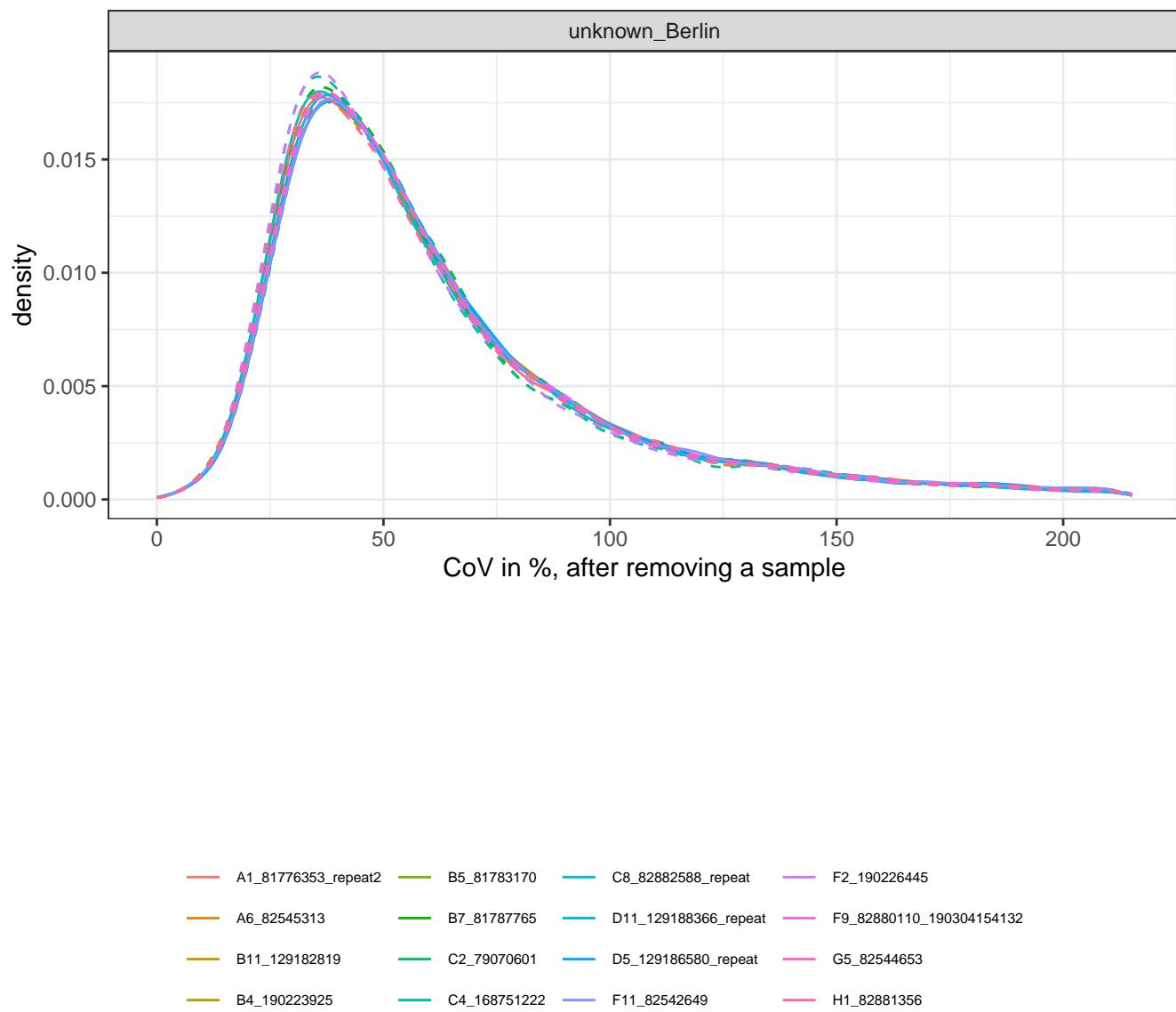


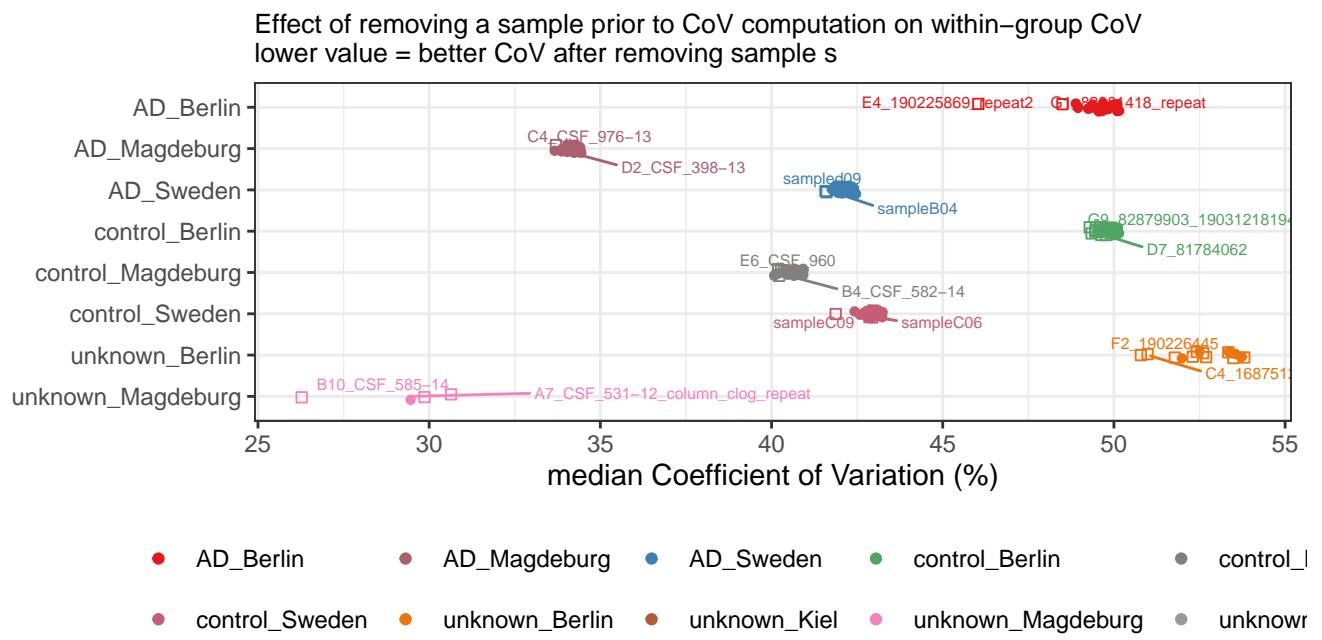
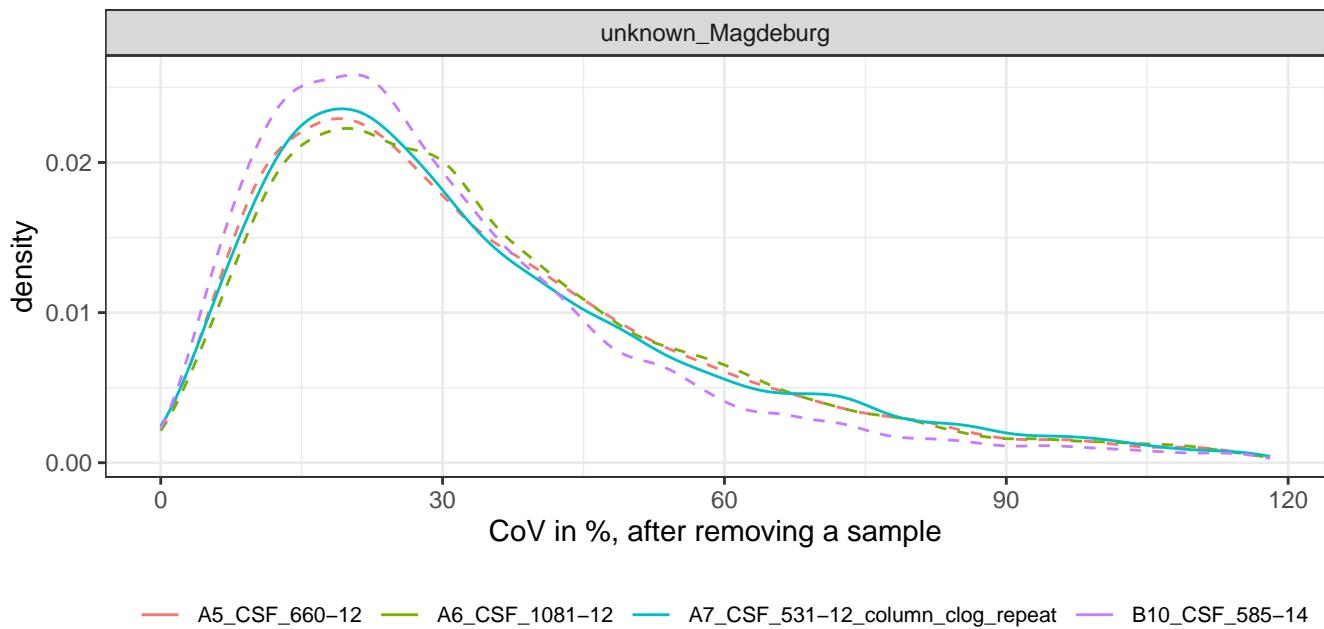


A10_82876065	B8_79116161	D12_168744796_repeat2	E7_81758548	G11_79112692_repeat
A11_129183842	B9_79059821	D6_129182306_repeat	E8_129188551_190304214129	G12_190225664
A12_190103450	C10_79069292	D7_81784062	E9_82882320_repeat	G6_168744541
A5_81790643	C11_129190335	D8_129182061	F10_190102946_repeat	G7_79061594
A7_79061681_repeat	C12_168743858_repeat	D9_82885545_repeat	F12_190100162	G8_82872412
A8_79114549_repeat	C5_81760601	E10_190225272	F5_129184099_repeat	G9_82879903_190312181
A9_82544855	C6_82881358	E11_82871132_test_190306003105	F6_129189909_190308011553	H11_190103151
B10_79069372	C7_81790598_repeat2	E12_190226223	F7_81784602_repeat	H5_82544762_190309202
B12_82872039	C9_79114900	E5_129187004_repeat	F8_190224717_repeat	H7_129187195
B6_79113519	D10_190224386	E6_129188085	G10_190102976	H9_82875993









shortname	group	exclude	median	CoV
E4_190225869_repeat2	AD_Berlin	TRUE	46.0	
G1_82881418_repeat	AD_Berlin	TRUE	48.5	
F3_129190664	AD_Berlin	FALSE	48.9	
A3_82883165_repeat2	AD_Berlin	FALSE	49.0	
E2_129189650	AD_Berlin	FALSE	49.2	
D1_82545185	AD_Berlin	FALSE	49.3	
B2_79114832	AD_Berlin	FALSE	49.3	
A2_82872920	AD_Berlin	FALSE	49.4	
G3_79069638	AD_Berlin	FALSE	49.5	
D3_129189776	AD_Berlin	FALSE	49.5	
H3_129190521_repeat	AD_Berlin	FALSE	49.6	
E1_82883860	AD_Berlin	FALSE	49.7	
C1_82545786	AD_Berlin	FALSE	49.7	
F4_190100053	AD_Berlin	FALSE	49.7	
E3_82871060_repeat	AD_Berlin	FALSE	49.7	
F1_82882396	AD_Berlin	FALSE	49.8	
D4_168751414	AD_Berlin	FALSE	49.8	
B1_129186923	AD_Berlin	FALSE	49.8	
G4_190101656_repeat	AD_Berlin	FALSE	49.9	
A4_190104103	AD_Berlin	FALSE	49.9	
D2_79110561	AD_Berlin	FALSE	50.0	
B3_79113129	AD_Berlin	FALSE	50.1	
G2_168751345	AD_Berlin	FALSE	50.1	
C3_82874751	AD_Berlin	FALSE	50.2	
C4_CSF_976-13	AD_Magdeburg	FALSE	33.7	
D2_CSF_398-13	AD_Magdeburg	TRUE	33.7	
B12_CSF_819-14_190217150250	AD_Magdeburg	FALSE	33.9	
C7_CSF_381-14	AD_Magdeburg	FALSE	33.9	
A12_CSF_949-15	AD_Magdeburg	FALSE	33.9	
A2_CSF_248-15	AD_Magdeburg	FALSE	34.0	
C10_CSF_1289-14	AD_Magdeburg	FALSE	34.0	
A11_CSF_36-15	AD_Magdeburg	FALSE	34.0	
D4_CSF_522-13	AD_Magdeburg	FALSE	34.0	
D1_CSF_530-13	AD_Magdeburg	FALSE	34.0	
B3_CSF_545-14	AD_Magdeburg	FALSE	34.1	
B5_CSF_885-15	AD_Magdeburg	FALSE	34.2	
B2_CSF_407-14	AD_Magdeburg	FALSE	34.2	
B8_CSF_828-15	AD_Magdeburg	FALSE	34.2	
A4_CSF_86-14	AD_Magdeburg	FALSE	34.2	
B11_CSF_506-14	AD_Magdeburg	FALSE	34.3	
C6_CSF_841-14	AD_Magdeburg	FALSE	34.3	
C3_CSF_974-14	AD_Magdeburg	FALSE	34.3	
C8_CSF_840-14	AD_Magdeburg	FALSE	34.3	
D5_CSF_503-13	AD_Magdeburg	FALSE	34.3	
C1_CSF_656-15	AD_Magdeburg	FALSE	34.4	
B7_CSF_496-15	AD_Magdeburg	FALSE	34.4	
D3_CSF_863-13	AD_Magdeburg	FALSE	34.4	
C2_CSF_1239-14	AD_Magdeburg	FALSE	34.4	
D6_CSF_864-13	AD_Magdeburg	FALSE	34.4	
sampled09	AD_Sweden	TRUE	41.6	

shortname	group	exclude	median	CoV
sampleB04	AD_Sweden	TRUE	41.6	
sampleA06	AD_Sweden	FALSE	41.8	
sampleE11	AD_Sweden	FALSE	41.8	
sampleE09	AD_Sweden	FALSE	41.9	
sampleB11	AD_Sweden	FALSE	41.9	
sampleD08	AD_Sweden	FALSE	41.9	
sampleB10R2	AD_Sweden	FALSE	42.0	
sampleB03	AD_Sweden	TRUE	42.0	
sampleB01	AD_Sweden	FALSE	42.0	
sampleB07	AD_Sweden	FALSE	42.0	
sampleA05	AD_Sweden	FALSE	42.1	
sampleB08	AD_Sweden	FALSE	42.1	
sampleB05	AD_Sweden	FALSE	42.1	
sampleC04	AD_Sweden	FALSE	42.2	
sampleA10	AD_Sweden	FALSE	42.2	
sampleE07	AD_Sweden	FALSE	42.2	
sampleA03	AD_Sweden	FALSE	42.2	
sampleB12	AD_Sweden	FALSE	42.2	
sampleA12	AD_Sweden	FALSE	42.3	
sampleE10	AD_Sweden	FALSE	42.3	
sampleA11	AD_Sweden	FALSE	42.3	
sampleA02	AD_Sweden	FALSE	42.3	
sampleA04	AD_Sweden	FALSE	42.3	
sampleE08	AD_Sweden	FALSE	42.3	
sampleA09	AD_Sweden	FALSE	42.4	
sampleB09	AD_Sweden	FALSE	42.4	
sampleC05	AD_Sweden	FALSE	42.4	
sampleE12	AD_Sweden	FALSE	42.5	
G9_82879903_190312181946	control_Berlin	TRUE	49.3	
D7_81784062	control_Berlin	TRUE	49.3	
G7_79061594	control_Berlin	TRUE	49.5	
A12_190103450	control_Berlin	TRUE	49.5	
C7_81790598_repeat2	control_Berlin	FALSE	49.5	
C5_81760601	control_Berlin	TRUE	49.6	
F12_190100162	control_Berlin	TRUE	49.6	
F6_129189909_190308011553	control_Berlin	TRUE	49.6	
C11_129190335	control_Berlin	FALSE	49.7	
F10_190102946_repeat	control_Berlin	TRUE	49.7	
F7_81784602_repeat	control_Berlin	FALSE	49.7	
B8_79116161	control_Berlin	TRUE	49.7	
G10_190102976	control_Berlin	FALSE	49.7	
F5_129184099_repeat	control_Berlin	FALSE	49.7	
B6_79113519	control_Berlin	FALSE	49.8	
C9_79114900	control_Berlin	FALSE	49.8	
E10_190225272	control_Berlin	TRUE	49.8	
G8_82872412	control_Berlin	FALSE	49.8	
A8_79114549_repeat	control_Berlin	FALSE	49.8	
E8_129188551_190304214129	control_Berlin	FALSE	49.8	
C10_79069292	control_Berlin	FALSE	49.8	
E11_82871132_test_190306003105	control_Berlin	FALSE	49.8	

shortname	group	exclude	median CoV
D12_168744796_repeat2	control_Berlin	TRUE	49.9
E7_81758548	control_Berlin	FALSE	49.9
A9_82544855	control_Berlin	FALSE	49.9
B12_82872039	control_Berlin	FALSE	49.9
D8_129182061	control_Berlin	TRUE	49.9
H11_190103151	control_Berlin	FALSE	49.9
E12_190226223	control_Berlin	FALSE	49.9
E6_129188085	control_Berlin	TRUE	49.9
C6_82881358	control_Berlin	FALSE	49.9
E9_82882320_repeat	control_Berlin	FALSE	49.9
H9_82875993	control_Berlin	FALSE	49.9
A7_79061681_repeat	control_Berlin	FALSE	50.0
A5_81790643	control_Berlin	FALSE	50.0
B9_79059821	control_Berlin	FALSE	50.0
H7_129187195	control_Berlin	FALSE	50.0
A10_82876065	control_Berlin	FALSE	50.0
D10_190224386	control_Berlin	FALSE	50.0
E5_129187004_repeat	control_Berlin	FALSE	50.0
D6_129182306_repeat	control_Berlin	FALSE	50.0
B10_79069372	control_Berlin	FALSE	50.0
G6_168744541	control_Berlin	FALSE	50.0
H5_82544762_190309202008	control_Berlin	FALSE	50.1
C12_168743858_repeat	control_Berlin	FALSE	50.1
A11_129183842	control_Berlin	FALSE	50.1
F8_190224717_repeat	control_Berlin	FALSE	50.1
G11_79112692_repeat	control_Berlin	FALSE	50.1
G12_190225664	control_Berlin	FALSE	50.1
D9_82885545_repeat	control_Berlin	FALSE	50.1
E6_CSF_960	control_Magdeburg	FALSE	40.1
B4_CSF_582-14	control_Magdeburg	FALSE	40.1
E1_CSF_59	control_Magdeburg	FALSE	40.2
F7_CSF_142	control_Magdeburg	TRUE	40.2
E10_CSF_16	control_Magdeburg	TRUE	40.2
A3_CSF_358-15	control_Magdeburg	TRUE	40.3
A9_CSF_695-15	control_Magdeburg	FALSE	40.3
E4_CSF_761	control_Magdeburg	FALSE	40.4
C9_CSF_1224-14	control_Magdeburg	TRUE	40.4
E12_CSF_146	control_Magdeburg	FALSE	40.4
E8_CSF_693	control_Magdeburg	FALSE	40.5
E2_CSF_396	control_Magdeburg	FALSE	40.5
F6_CSF_134	control_Magdeburg	FALSE	40.5
B6_CSF_257-15	control_Magdeburg	FALSE	40.5
A8_CSF_471-12	control_Magdeburg	FALSE	40.6
E11_CSF_783	control_Magdeburg	FALSE	40.6
E5_CSF_123	control_Magdeburg	FALSE	40.6
C11_CSF_986-14	control_Magdeburg	FALSE	40.6
A1_CSF_670-12	control_Magdeburg	FALSE	40.7
C12_CSF_1299-14	control_Magdeburg	FALSE	40.7
C5_CSF_706-14	control_Magdeburg	FALSE	40.7
F9_CSF_420	control_Magdeburg	FALSE	40.7

shortname	group	exclude	median CoV
F4_CSF_51	control_Magdeburg	FALSE	40.8
F2_CSF_755	control_Magdeburg	FALSE	40.8
A10_CSF_201-15	control_Magdeburg	FALSE	40.8
B1_CSF_326-15	control_Magdeburg	FALSE	40.8
F3_CSF_865	control_Magdeburg	FALSE	40.9
B9_CSF_851-15	control_Magdeburg	FALSE	40.9
F8_CSF_143	control_Magdeburg	FALSE	40.9
sampleC09	control_Sweden	TRUE	41.9
sampleC06	control_Sweden	FALSE	42.4
sampleD07	control_Sweden	FALSE	42.6
sampleA01	control_Sweden	FALSE	42.7
sampleE03	control_Sweden	FALSE	42.7
sampleB02	control_Sweden	FALSE	42.8
sampleD02	control_Sweden	FALSE	42.8
sampleE02	control_Sweden	TRUE	42.9
sampleC01	control_Sweden	TRUE	42.9
sampleC07	control_Sweden	FALSE	42.9
sampleE01	control_Sweden	FALSE	42.9
sampleC11	control_Sweden	FALSE	42.9
sampleD06	control_Sweden	FALSE	42.9
sampleC02	control_Sweden	FALSE	42.9
sampleD05	control_Sweden	FALSE	42.9
sampleD11	control_Sweden	TRUE	42.9
sampleE06	control_Sweden	FALSE	43.0
sampleD01	control_Sweden	TRUE	43.0
sampleC10	control_Sweden	TRUE	43.0
sampleD04	control_Sweden	FALSE	43.0
sampleB06	control_Sweden	FALSE	43.0
sampleA08	control_Sweden	TRUE	43.0
sampleA07	control_Sweden	FALSE	43.0
sampleE04	control_Sweden	FALSE	43.1
sampleC03	control_Sweden	FALSE	43.1
sampleE05	control_Sweden	FALSE	43.1
sampleC12	control_Sweden	FALSE	43.1
sampleD03	control_Sweden	FALSE	43.1
sampleD10	control_Sweden	FALSE	43.2
sampleD12	control_Sweden	FALSE	43.2
sampleC08	control_Sweden	FALSE	43.2
F2_190226445	unknown_Berlin	TRUE	50.8
C4_168751222	unknown_Berlin	TRUE	51.0
B7_81787765	unknown_Berlin	TRUE	51.8
C8_82882588_repeat	unknown_Berlin	FALSE	52.0
H1_82881356	unknown_Berlin	TRUE	52.3
B5_81783170	unknown_Berlin	TRUE	52.4
A1_81776353_repeat2	unknown_Berlin	FALSE	52.5
A6_82545313	unknown_Berlin	TRUE	52.6
G5_82544653	unknown_Berlin	TRUE	52.7
F9_82880110_190304154132	unknown_Berlin	TRUE	53.3
D11_129188366_repeat	unknown_Berlin	FALSE	53.4
B4_190223925	unknown_Berlin	TRUE	53.5

shortname	group	exclude	median CoV
F11_82542649	unknown_Berlin	FALSE	53.5
B11_129182819	unknown_Berlin	FALSE	53.7
D5_129186580_repeat	unknown_Berlin	FALSE	53.7
C2_79070601	unknown_Berlin	TRUE	53.8
B10_CSF_585-14	unknown_Magdeburg	TRUE	26.3
A7_CSF_531-12_column_clog_repeat	unknown_Magdeburg	FALSE	29.5
A5_CSF_660-12	unknown_Magdeburg	TRUE	29.9
A6_CSF_1081-12	unknown_Magdeburg	TRUE	30.6

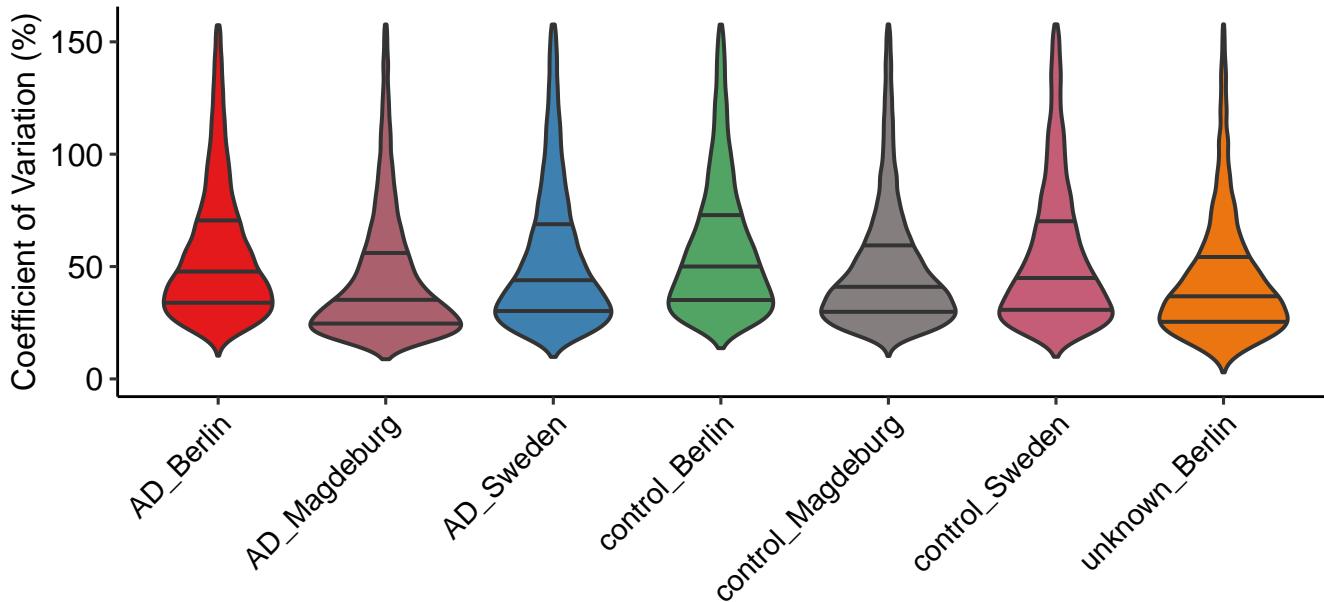
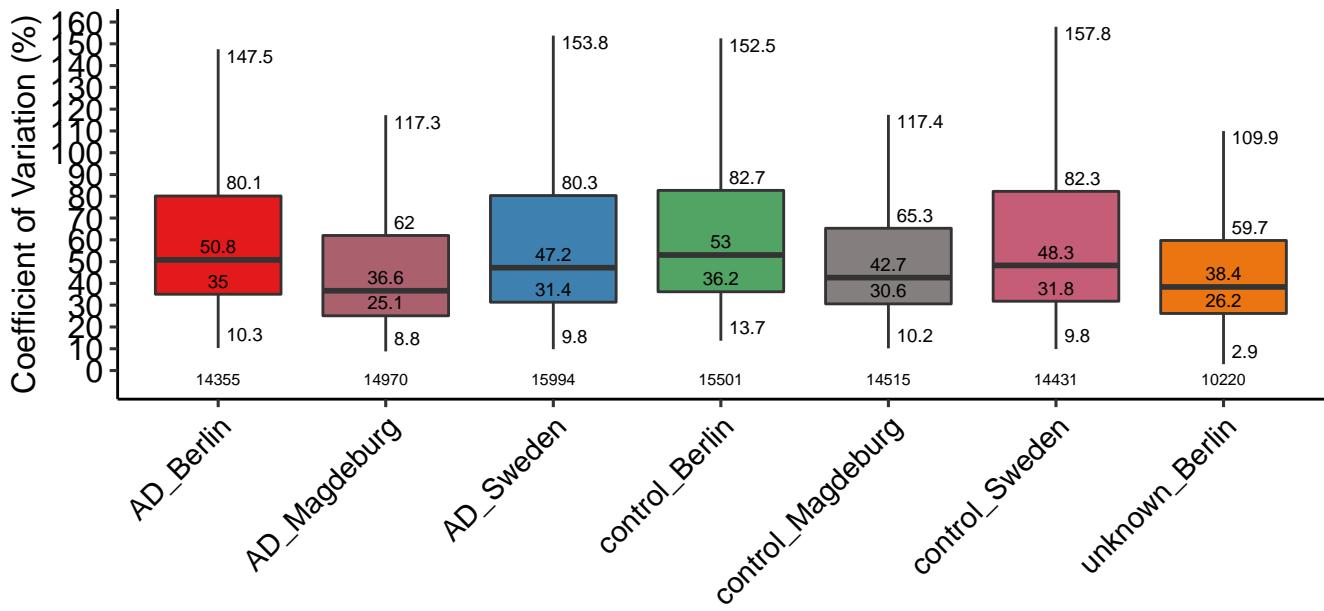
Leave-one-out impact on within-group CoV (%)

1.6.3 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.

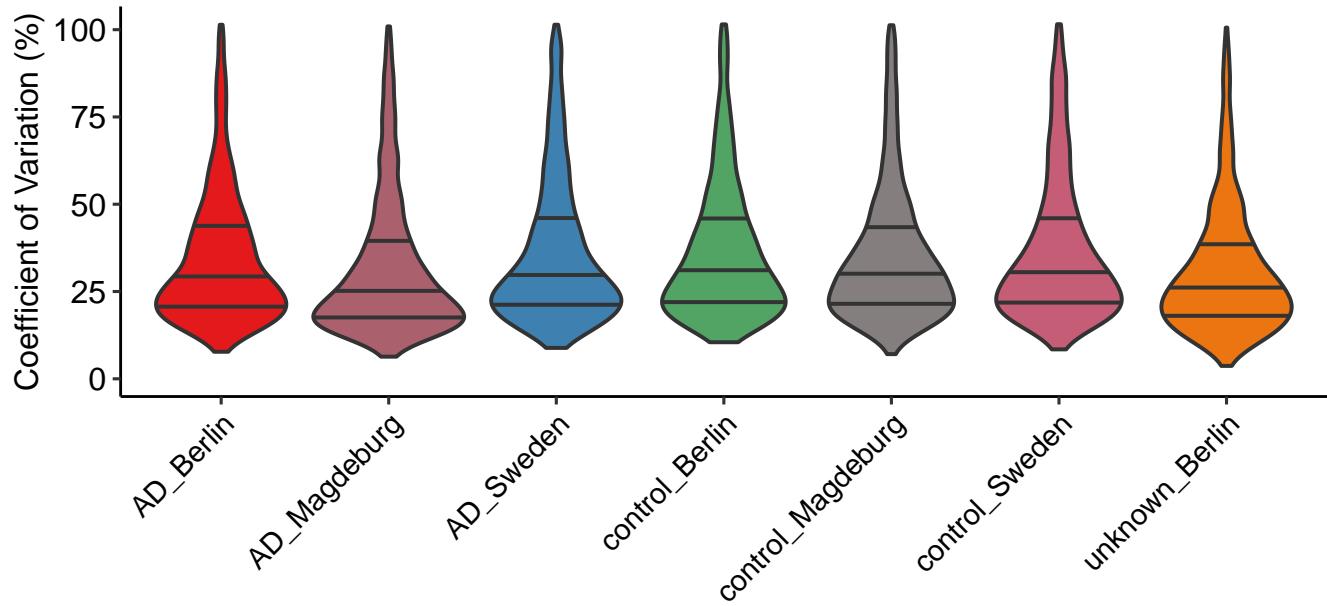
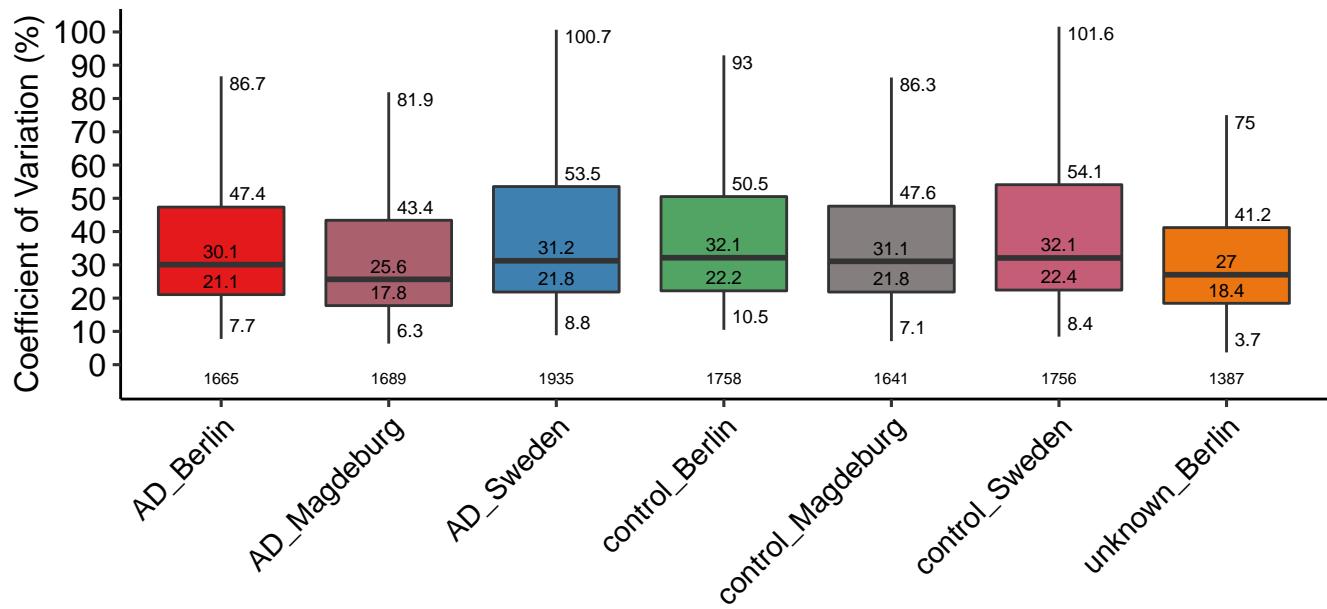
Only samples that are NOT marked ‘exclude’ in the provided sample metadata and are in a sample group among at least 3 replicates are used for these figures. The user-specified filtering rules (eg; filter_min_detect, filter_min_peptide_per_prot, etc.) were applied within each sample group independently and remaining peptides were subsequently normalized using the ”vsn&modebetween_protein” algorithm (an important parameter to keep in mind when comparing across datasets/analyses). Only peptides with at least 3 data points across replicate samples are used for each CoV computation.

Peptide-level CoV:



Protein-level CoV:

(analogous to peptide CoV's, but with additional rollup to protein abundances using 'maxlfq' algorithm)



1.7 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

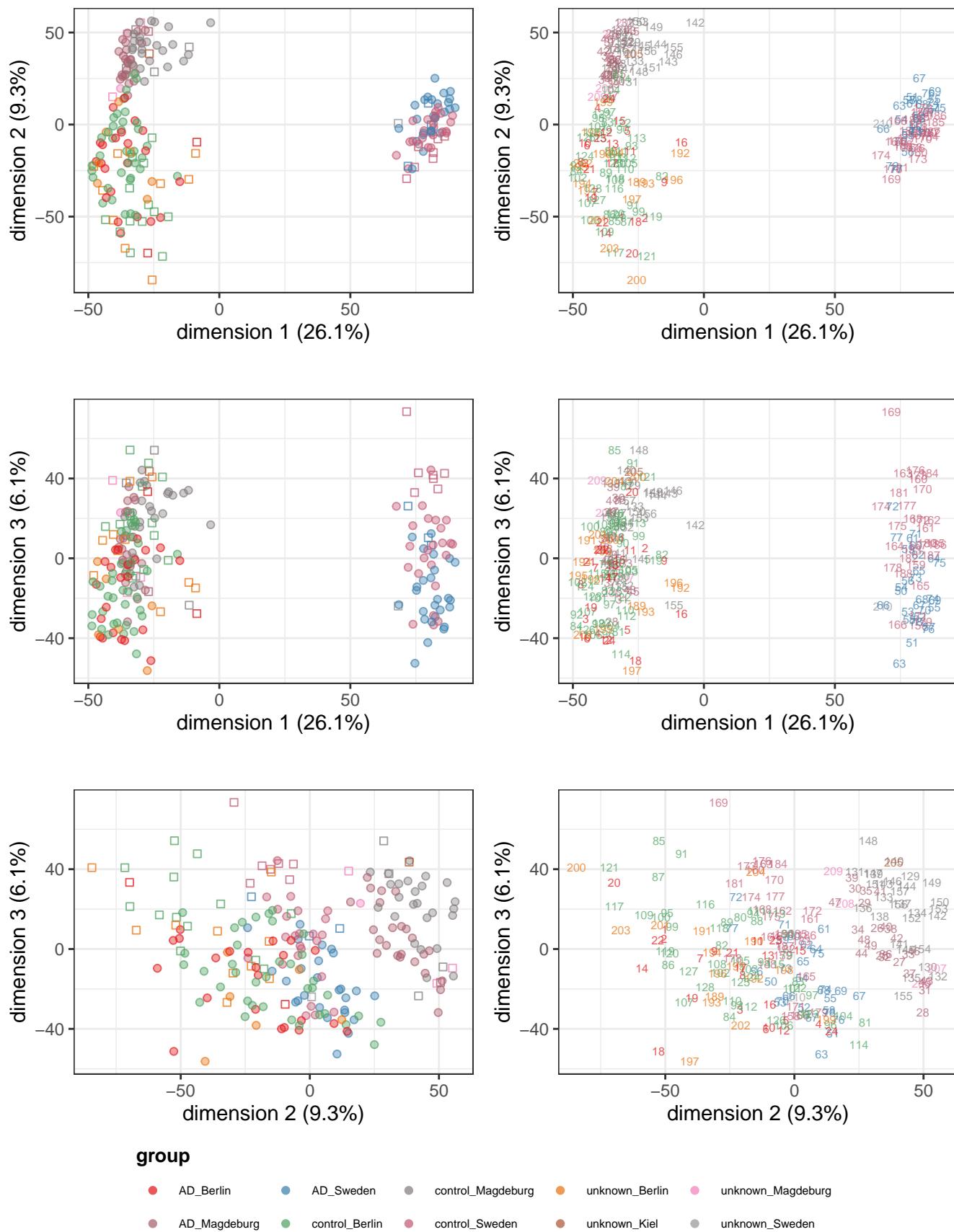
As mentioned above, the aim of the PCA figures is to identify global effects. To achieve this, we compute sample distances on the subset of peptides identified in each group which prevents rarely detected peptides/proteins from having a disproportionate effect on sample clustering. This pertains not 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.

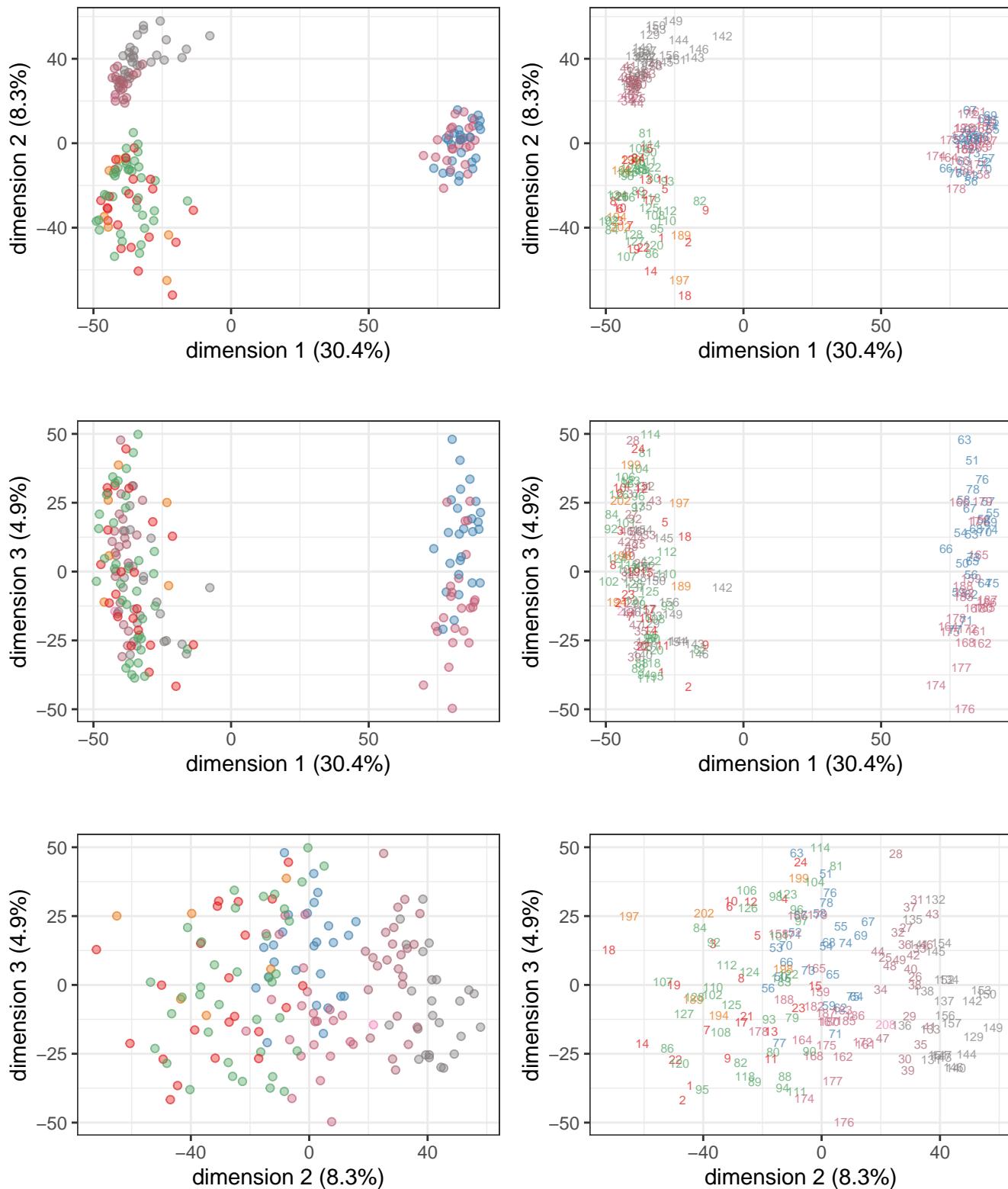
PCA of all samples, including those flagged as 'exclude', using 8892 peptides

samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)

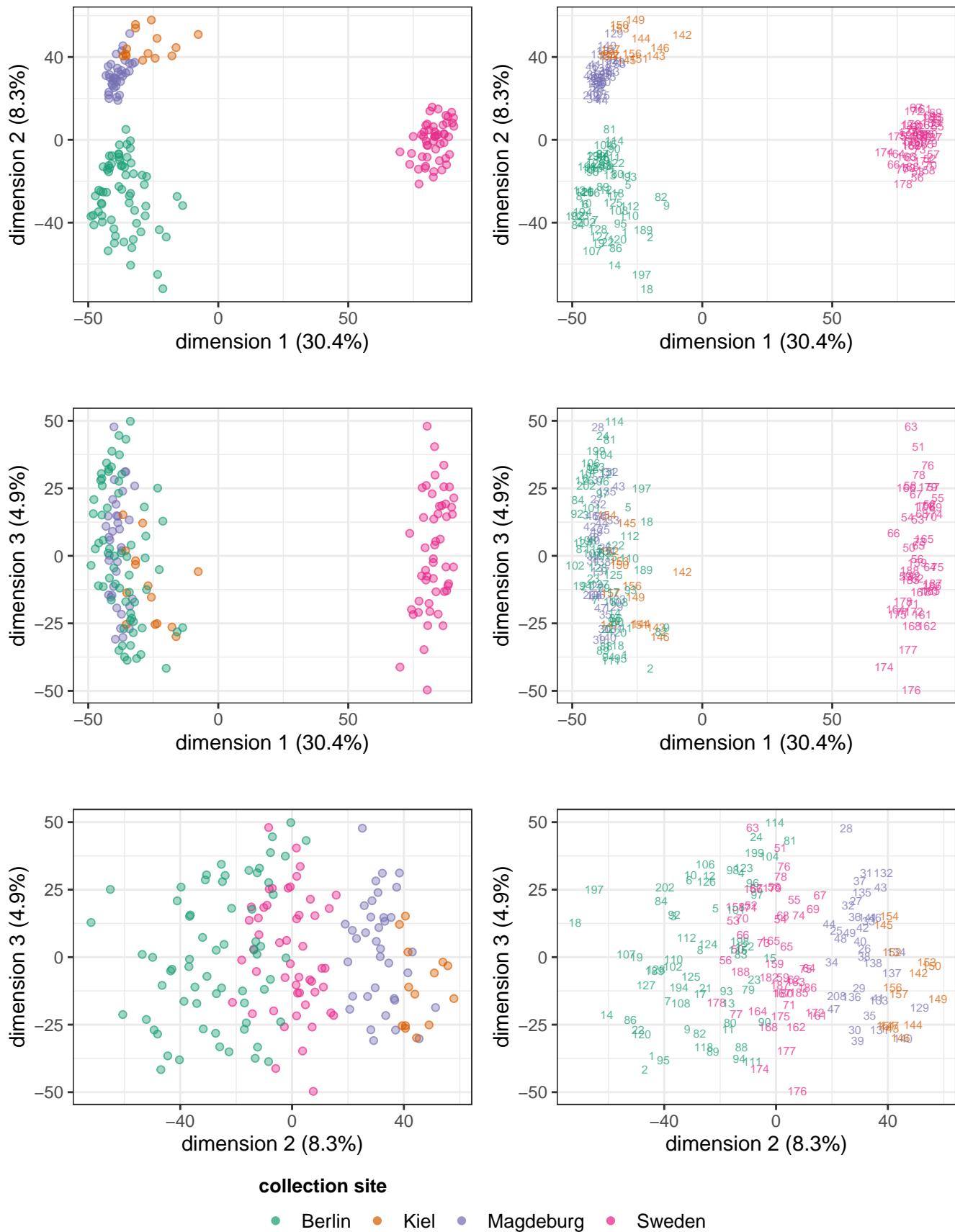


PCA only on samples not flagged as ‘exclude’, using 9164 peptides

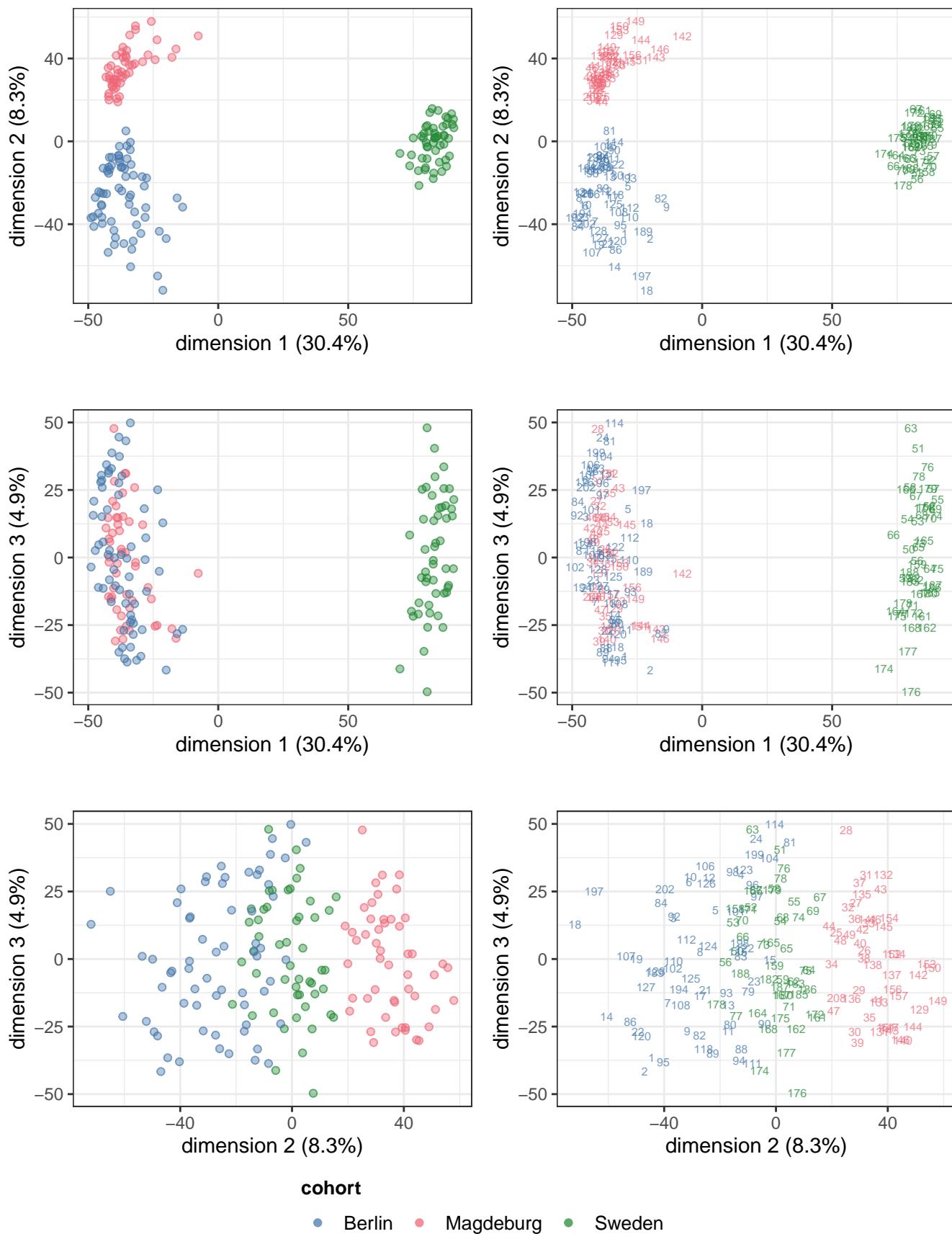
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



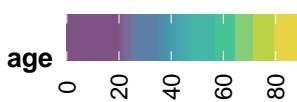
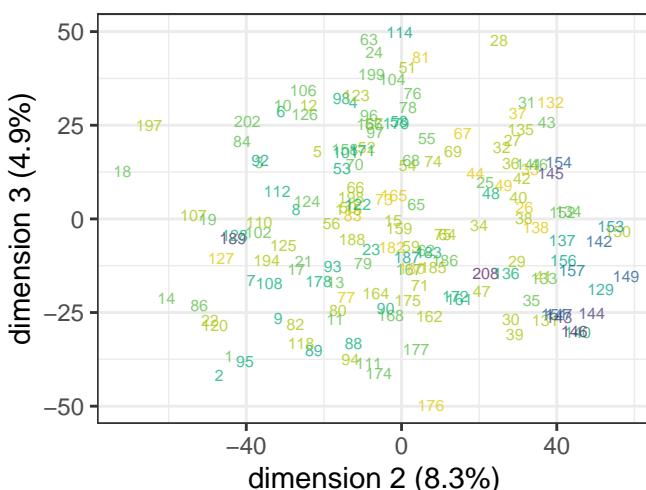
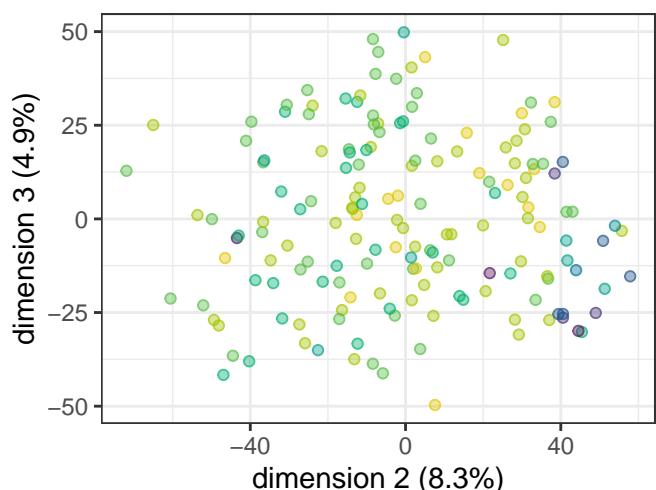
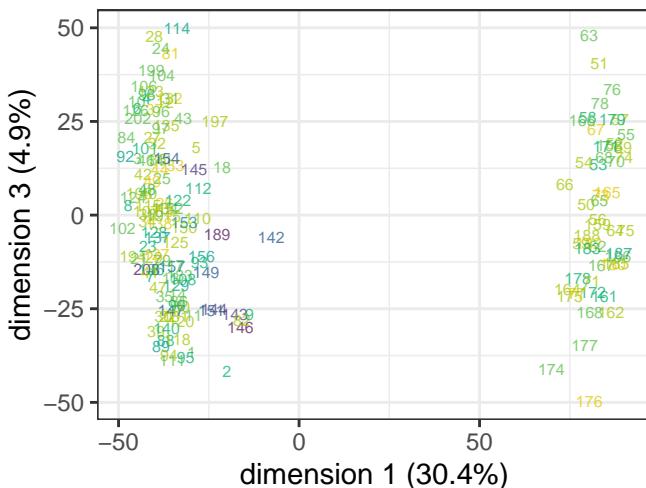
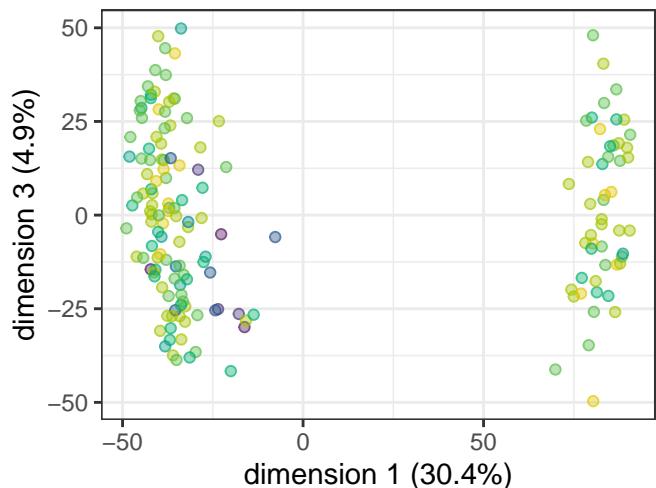
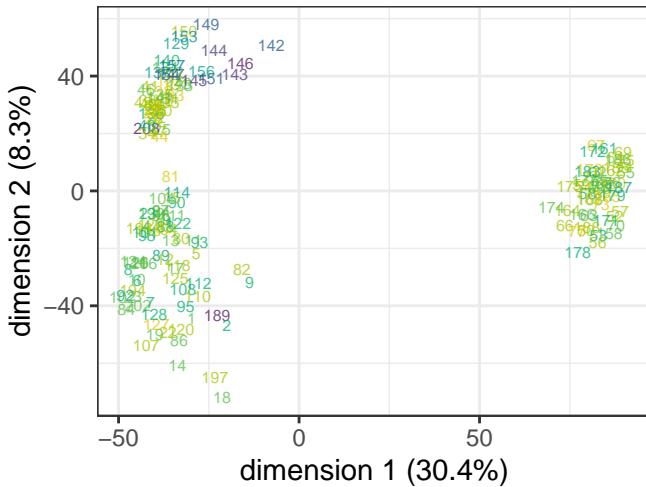
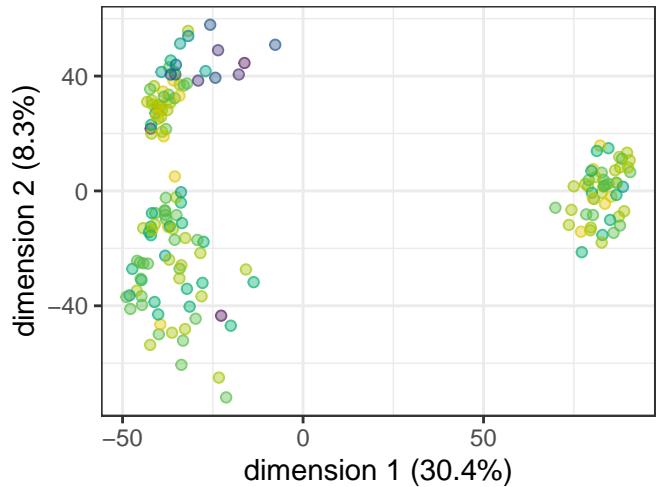
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



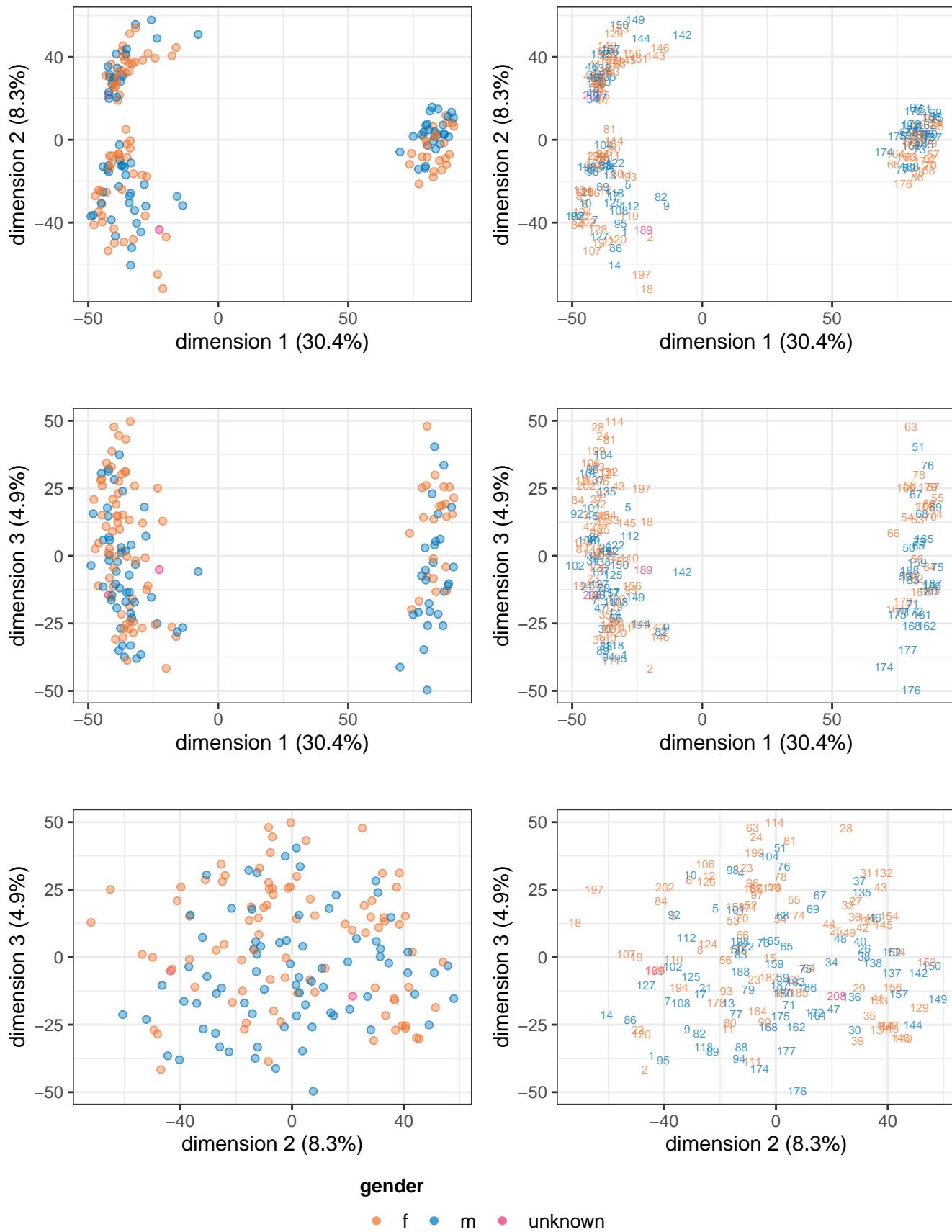
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



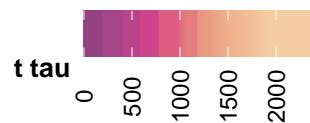
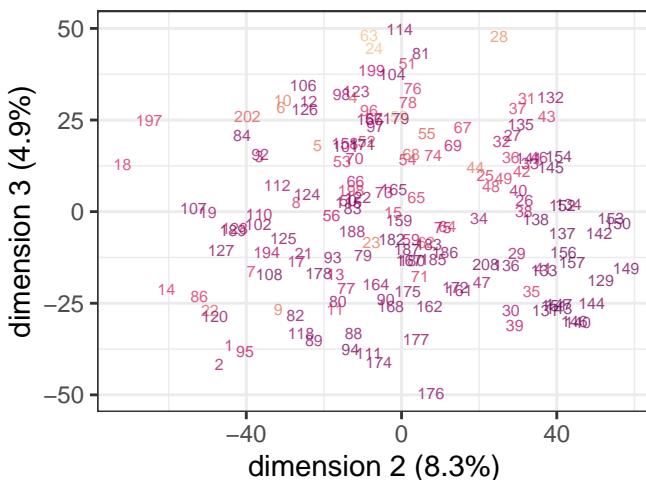
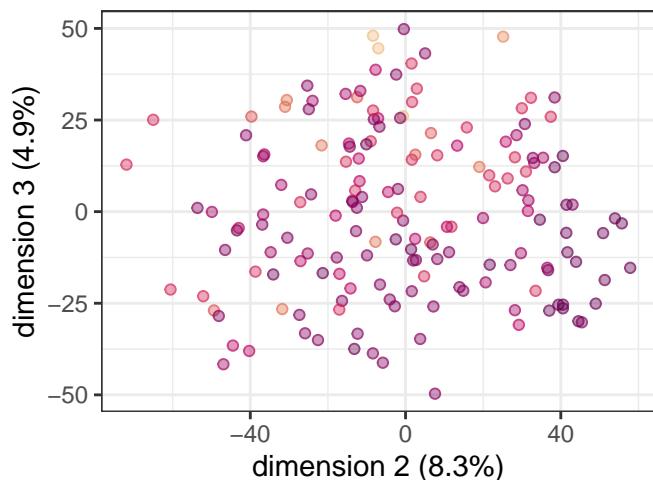
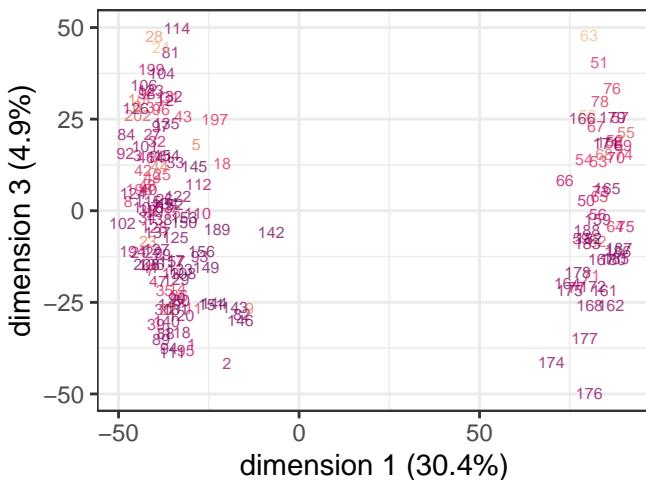
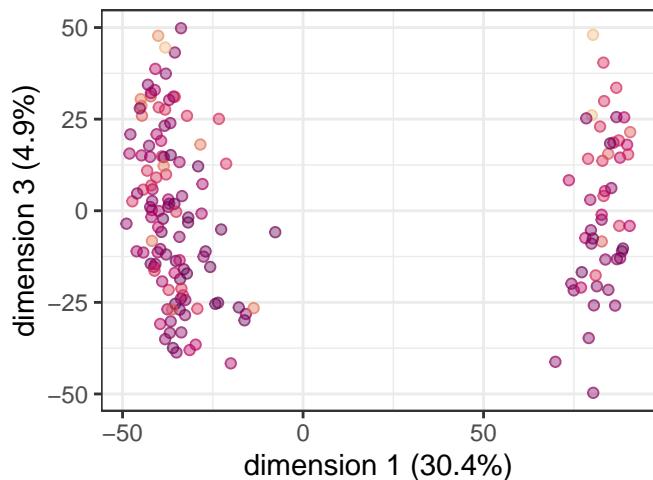
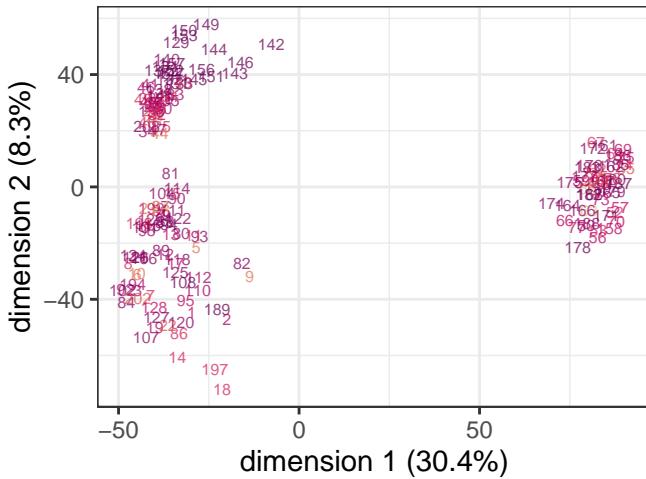
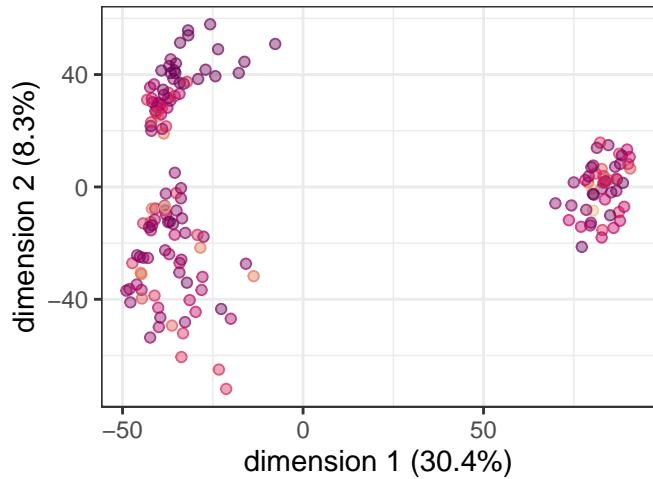
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



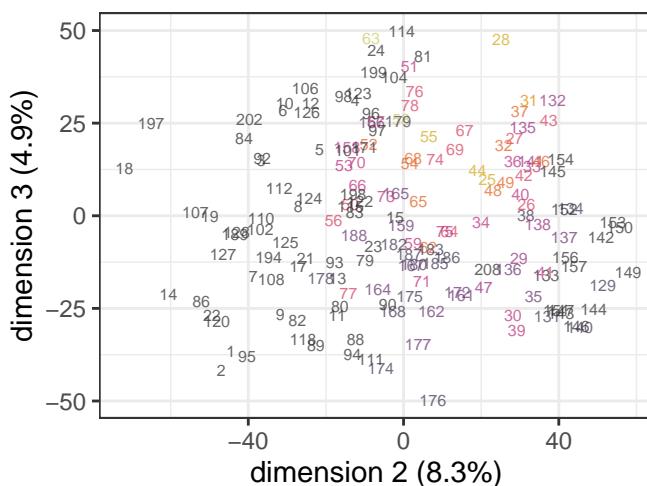
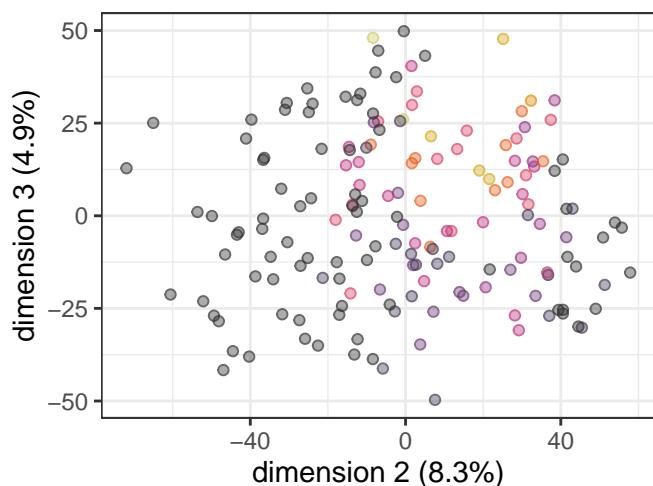
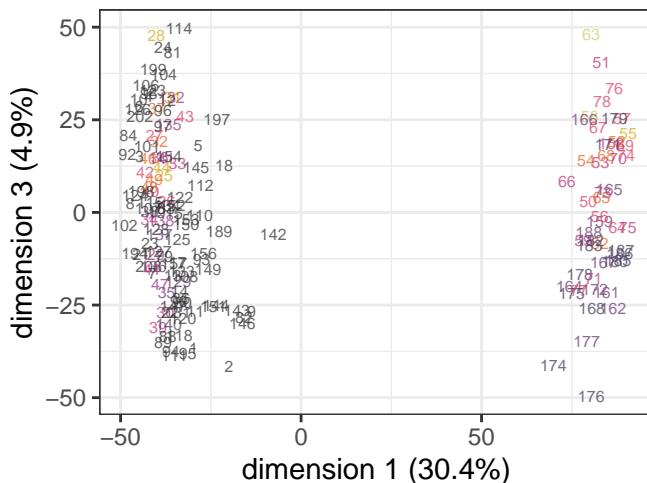
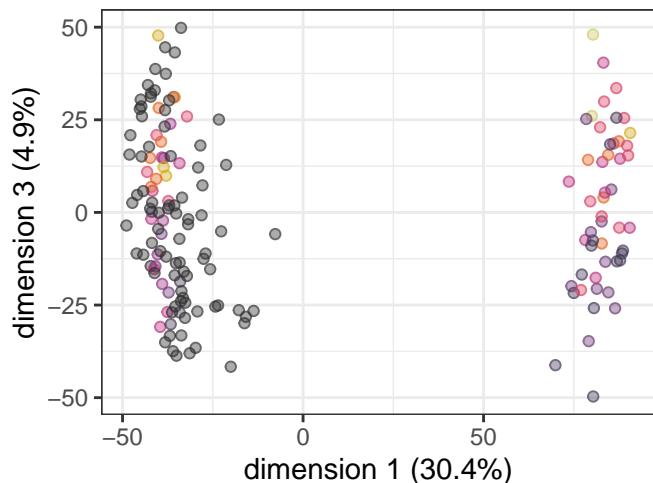
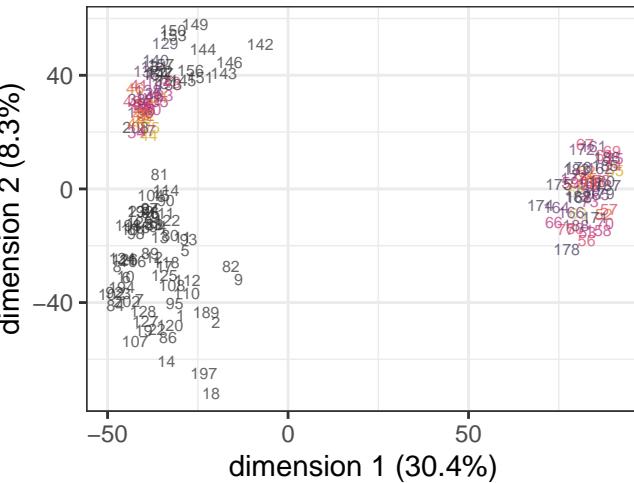
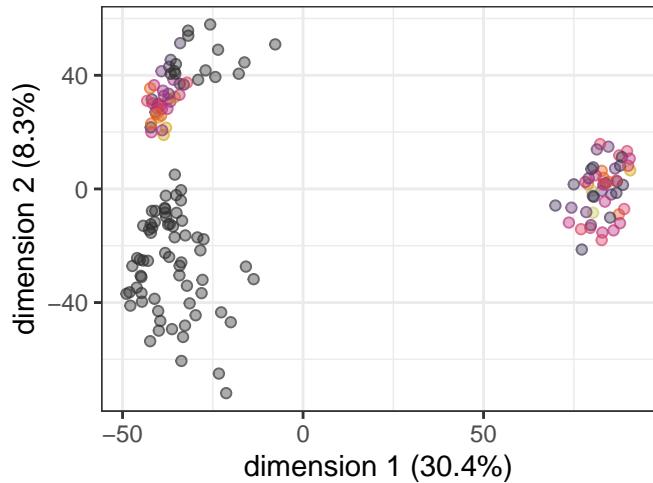
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



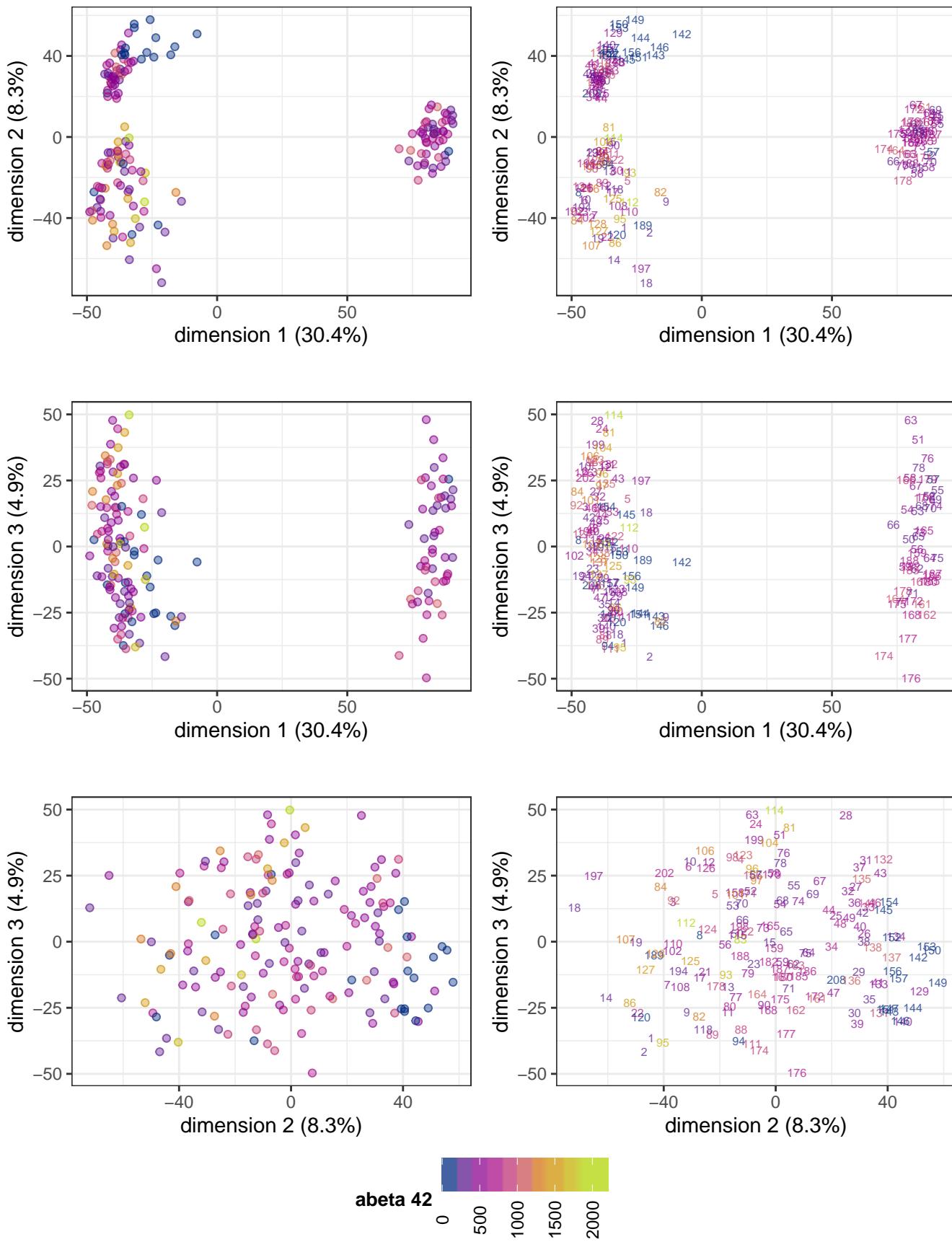
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



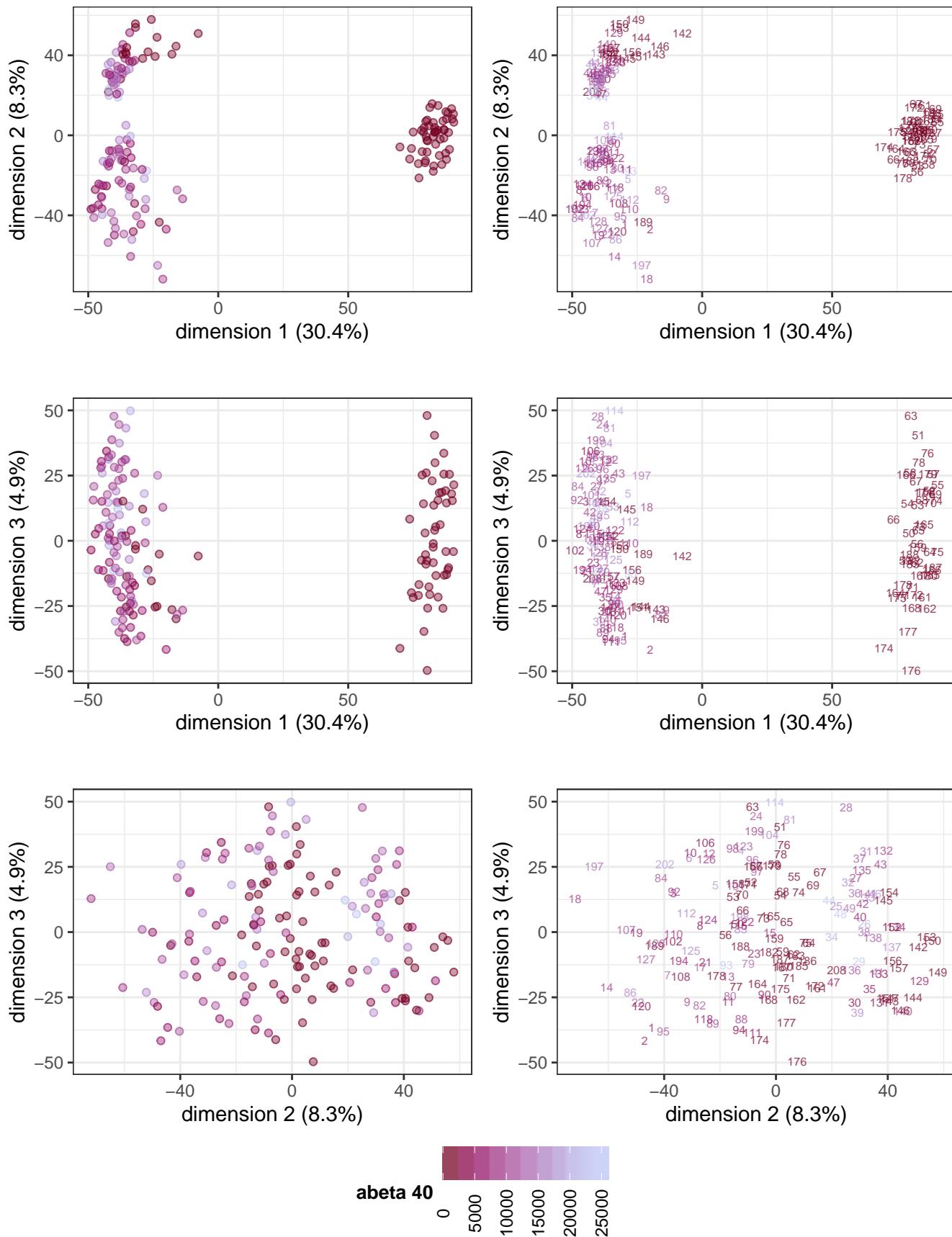
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



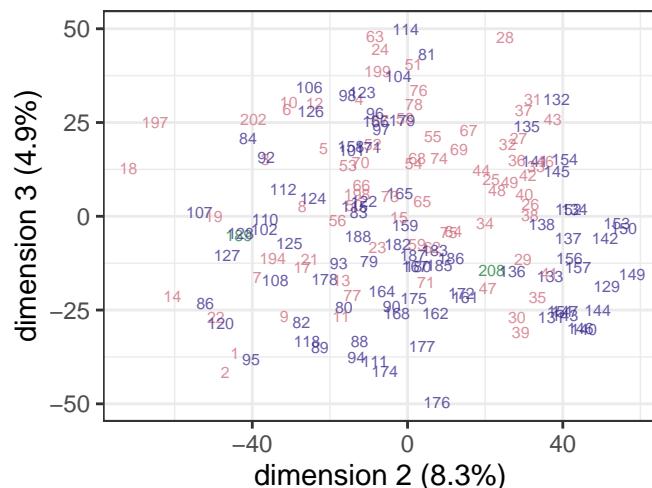
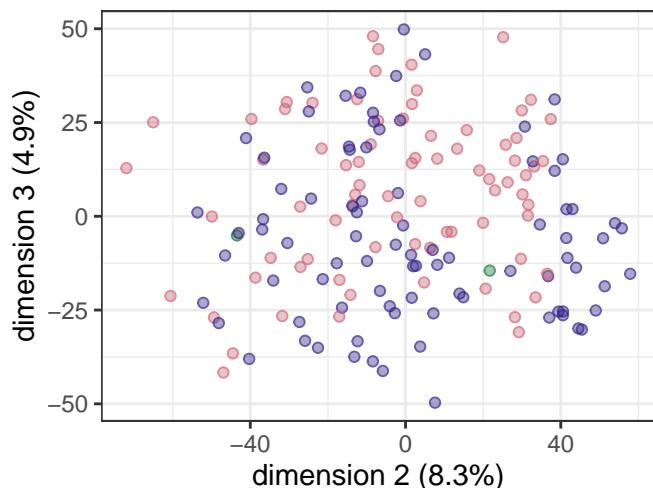
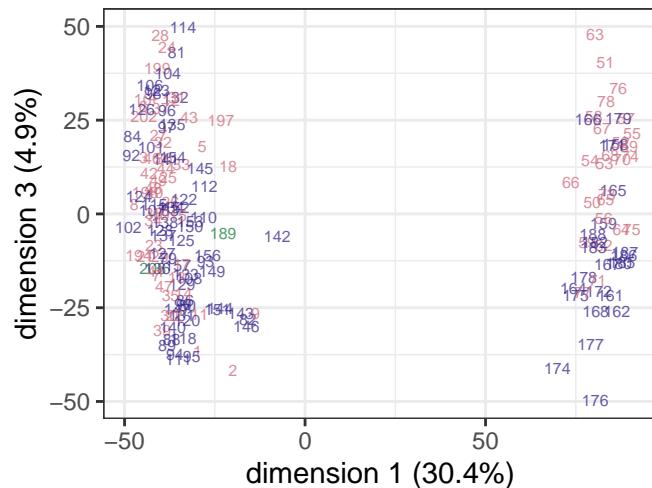
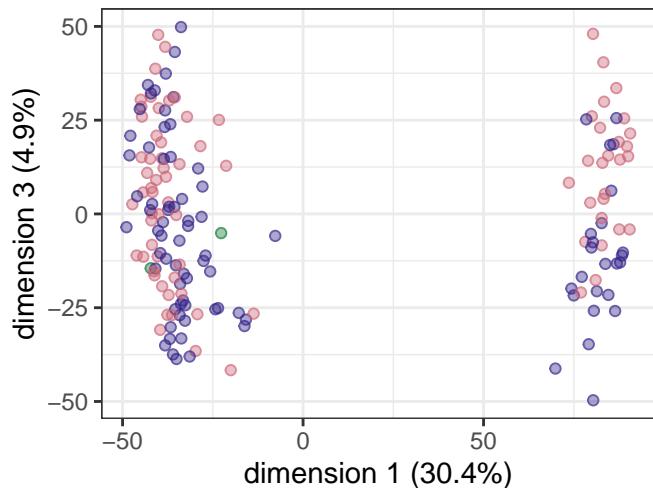
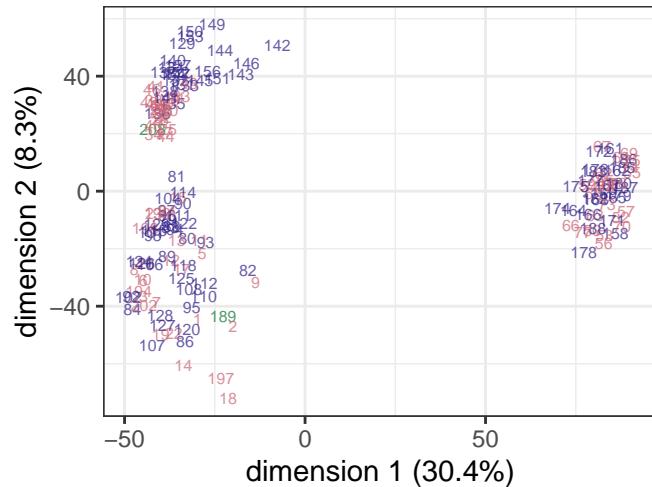
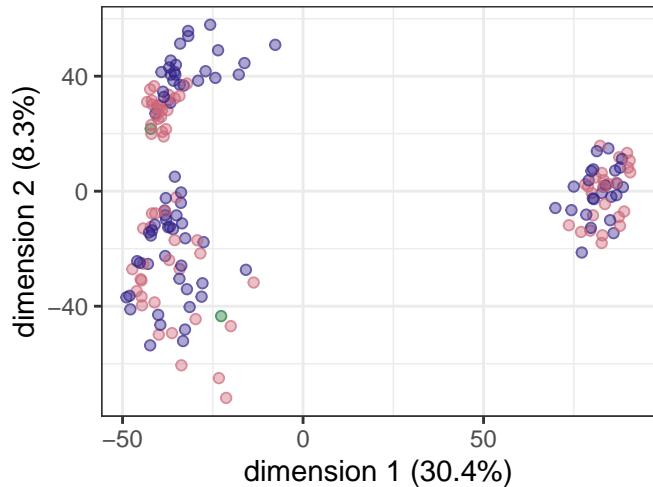
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



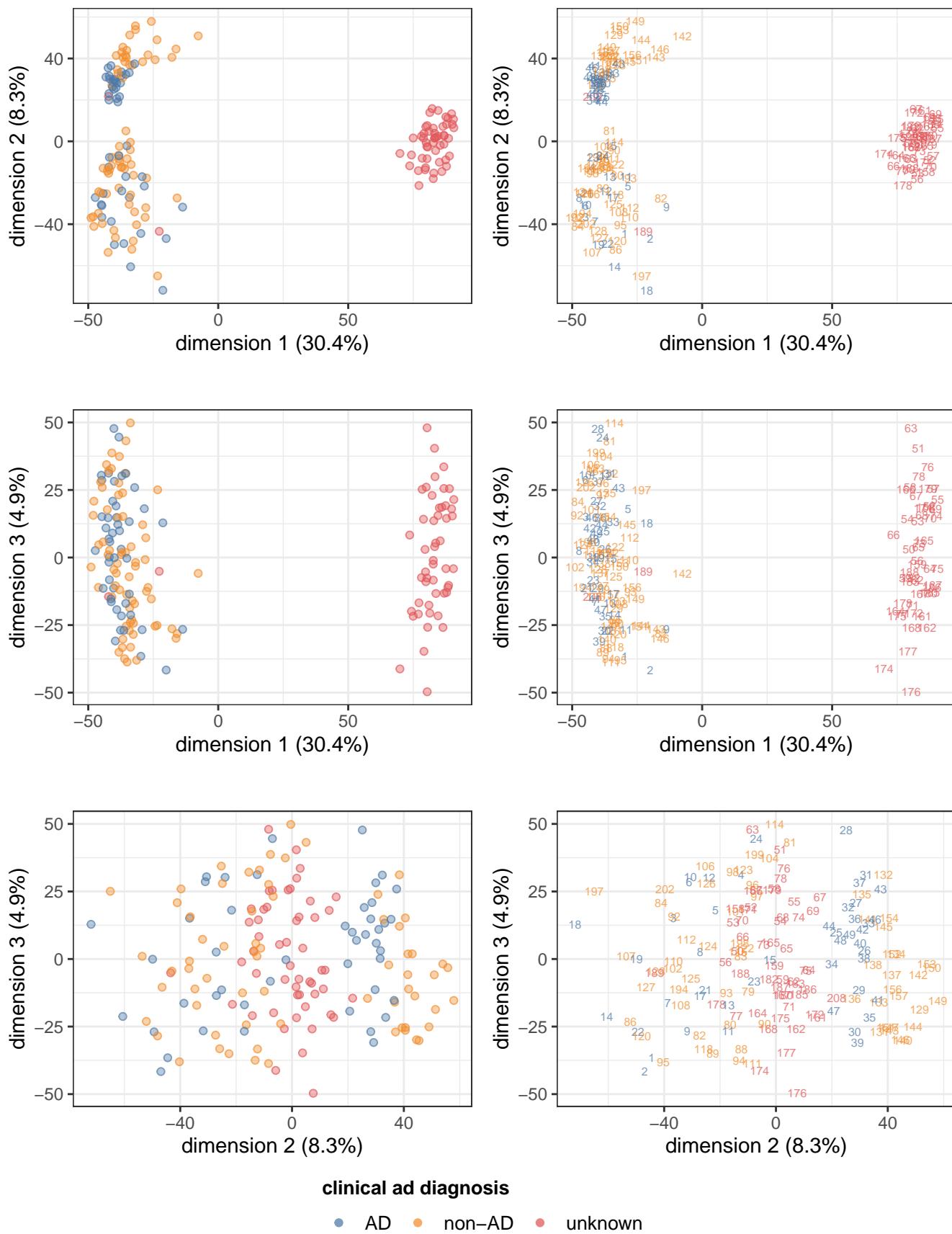
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



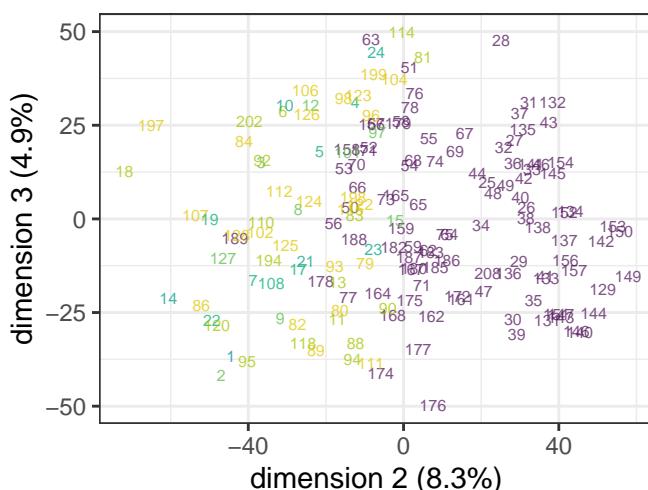
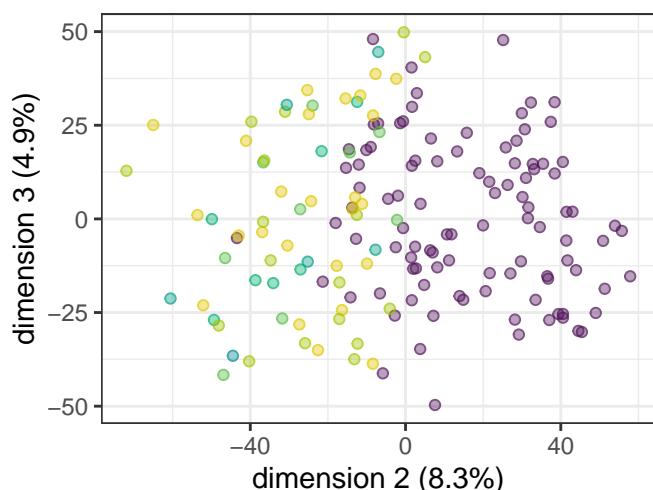
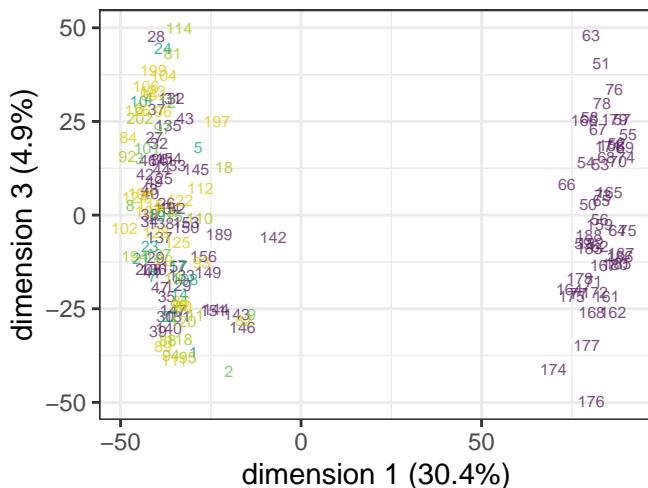
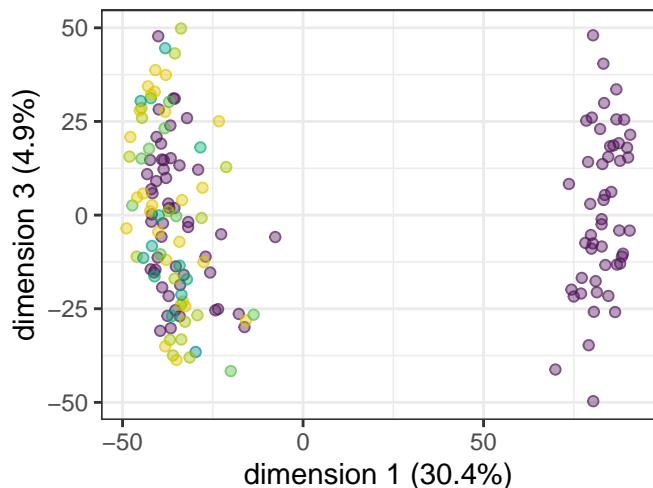
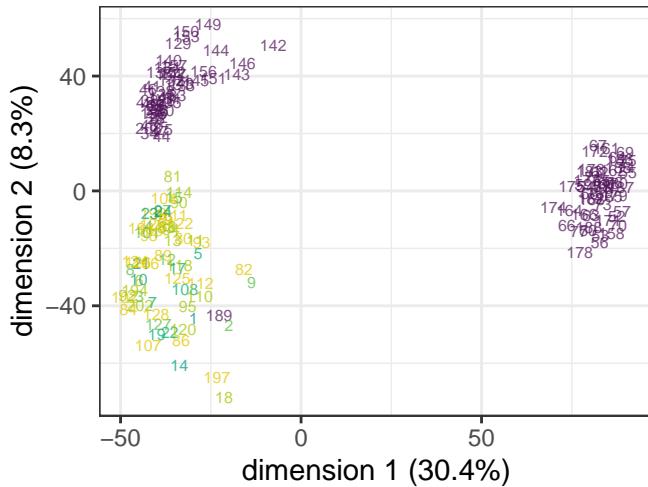
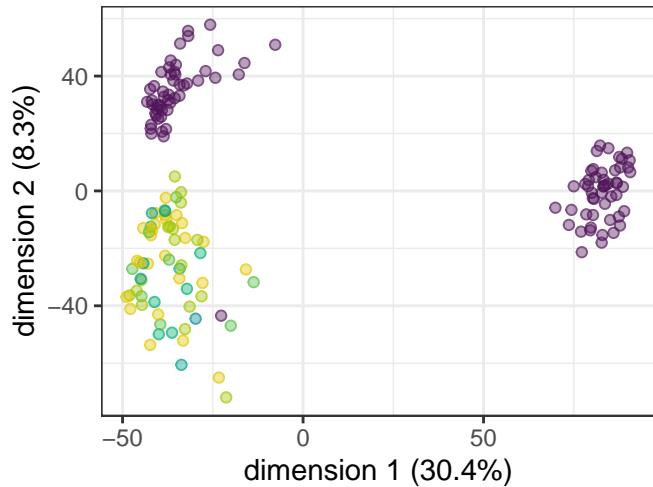
biochemical ad classification

- biochemical AD ● biochemical control ● unknown

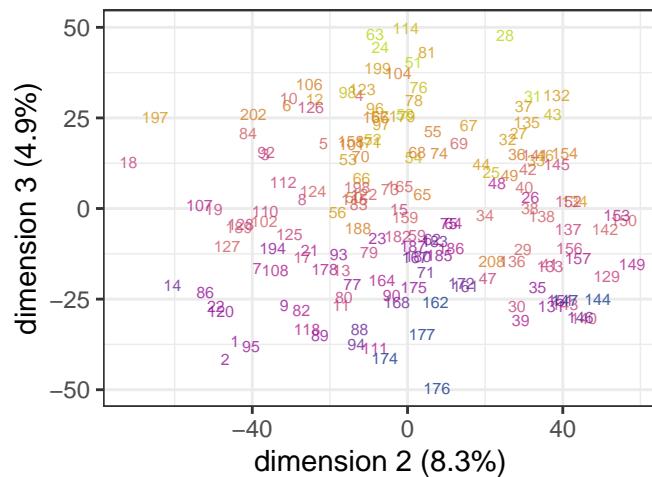
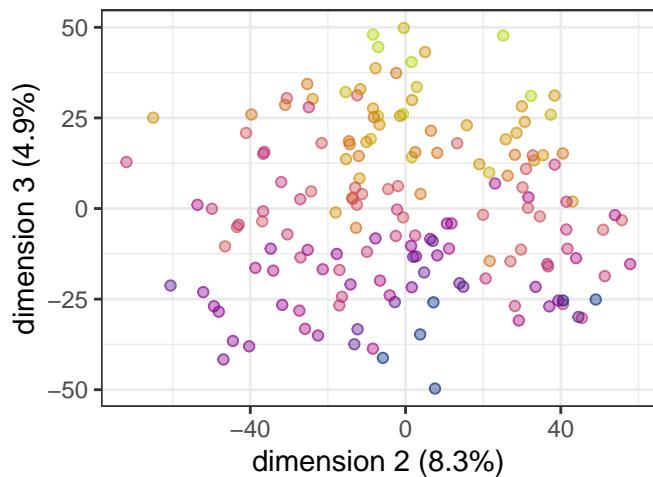
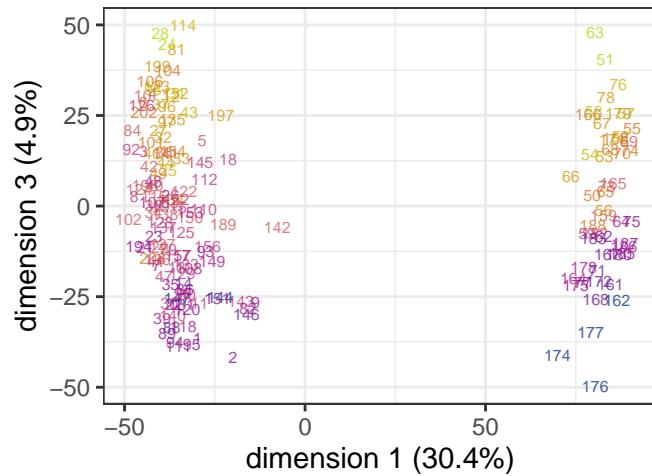
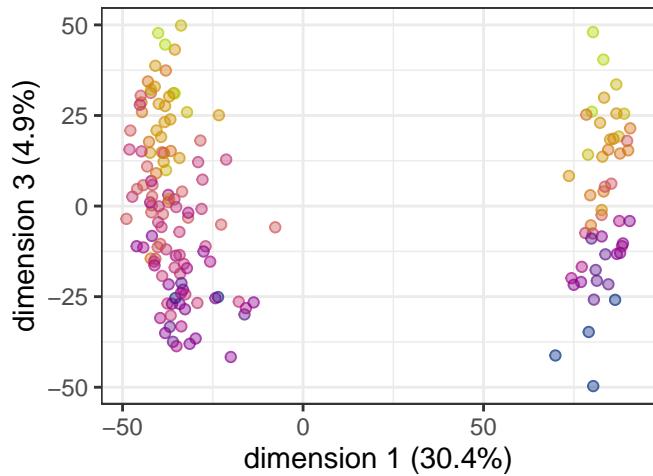
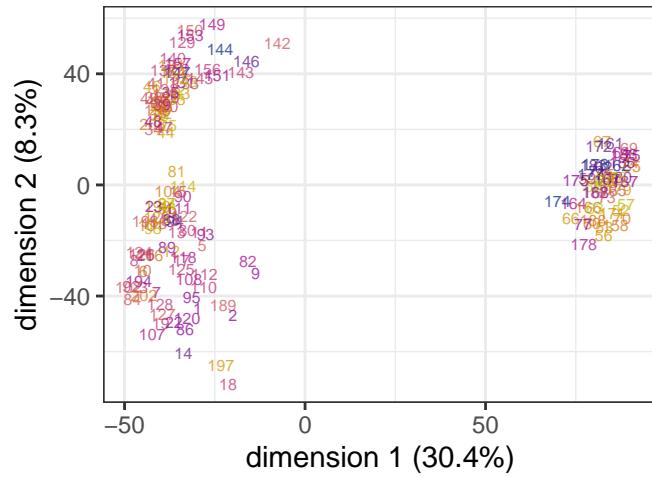
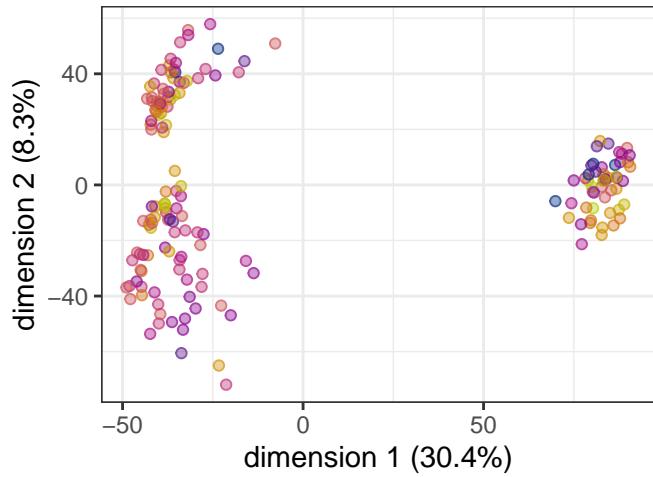
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



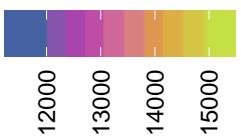
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



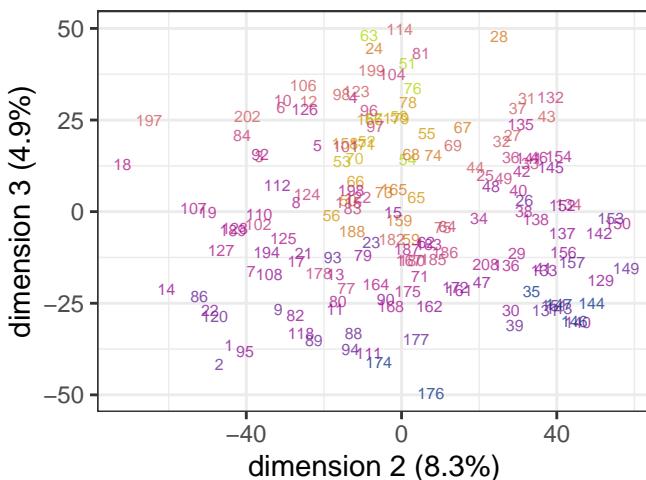
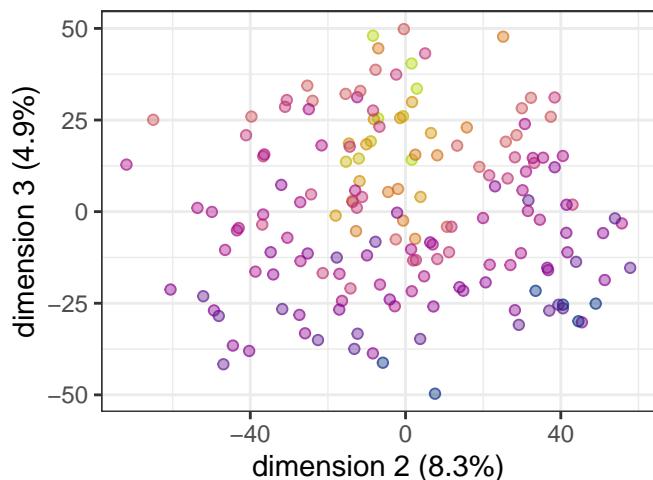
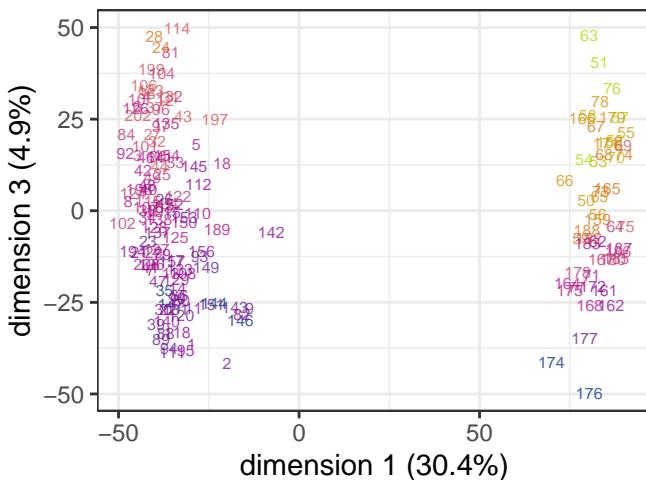
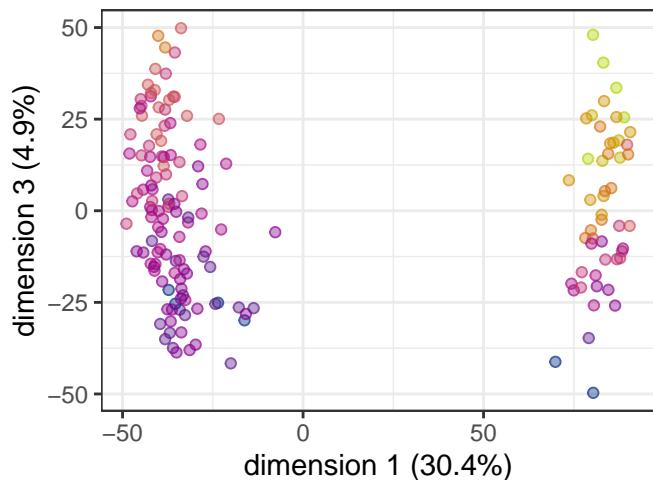
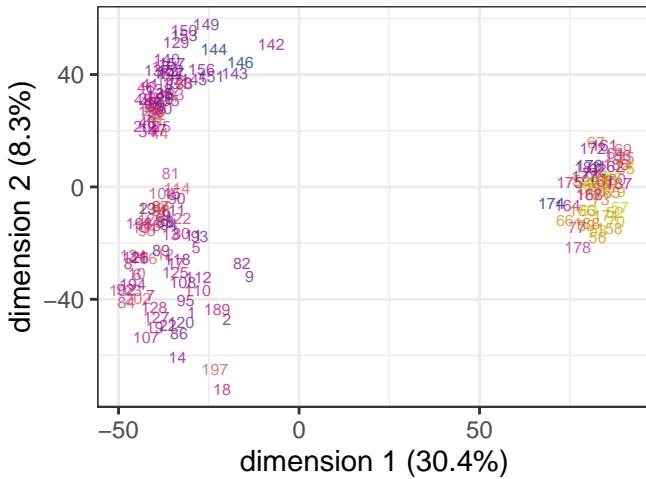
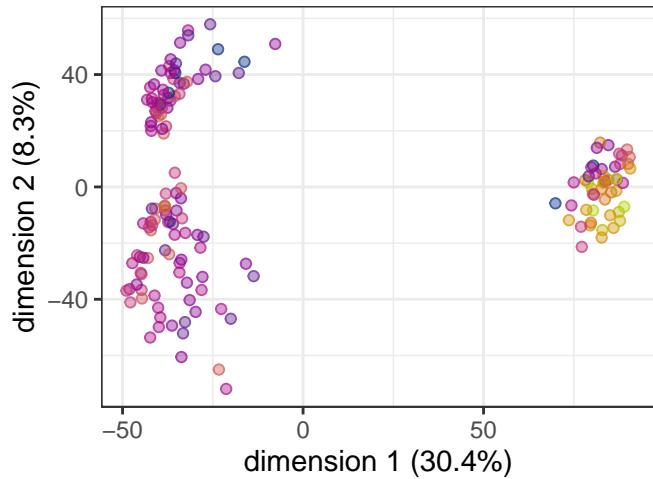
samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



detected peptides



samples are labeled by "sample_index" (described in samples.xlsx included with MS-DAP output)



detected proteins



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. In MS-DAP this is referred to as “by contrast” filtering.

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 subsequently 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 erroneously 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)

contrasts and foldchanges

Note that a MS-DAP contrast for “A vs B” returns foldchanges for B/A. For example, for the contrast “control vs disease” a positive log2 foldchange implies protein abundances are higher in the “disease” sample group.

2.1 control_Berlin,control_Magdeburg... vs AD_Berlin,AD_Magdeburg,AD_Sweden

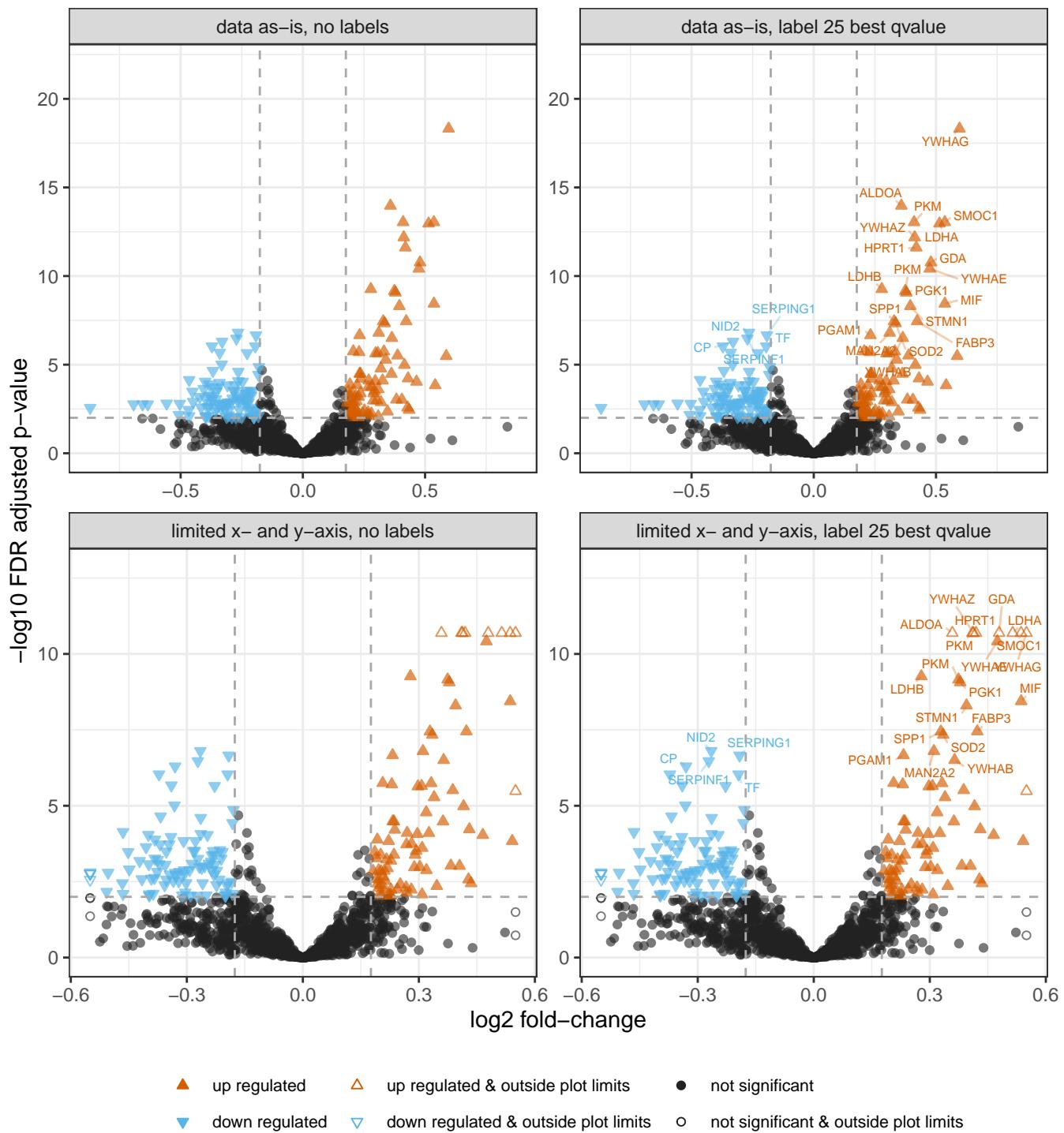
- **user setting:** using ‘filter by contrast’ peptide filtering approach
- 11824 peptides in 1445 proteins remain in the current contrast after peptide filters and are used for the statistical analysis in this section
- qvalue threshold: 0.01
- log2 foldchange threshold: 0.176

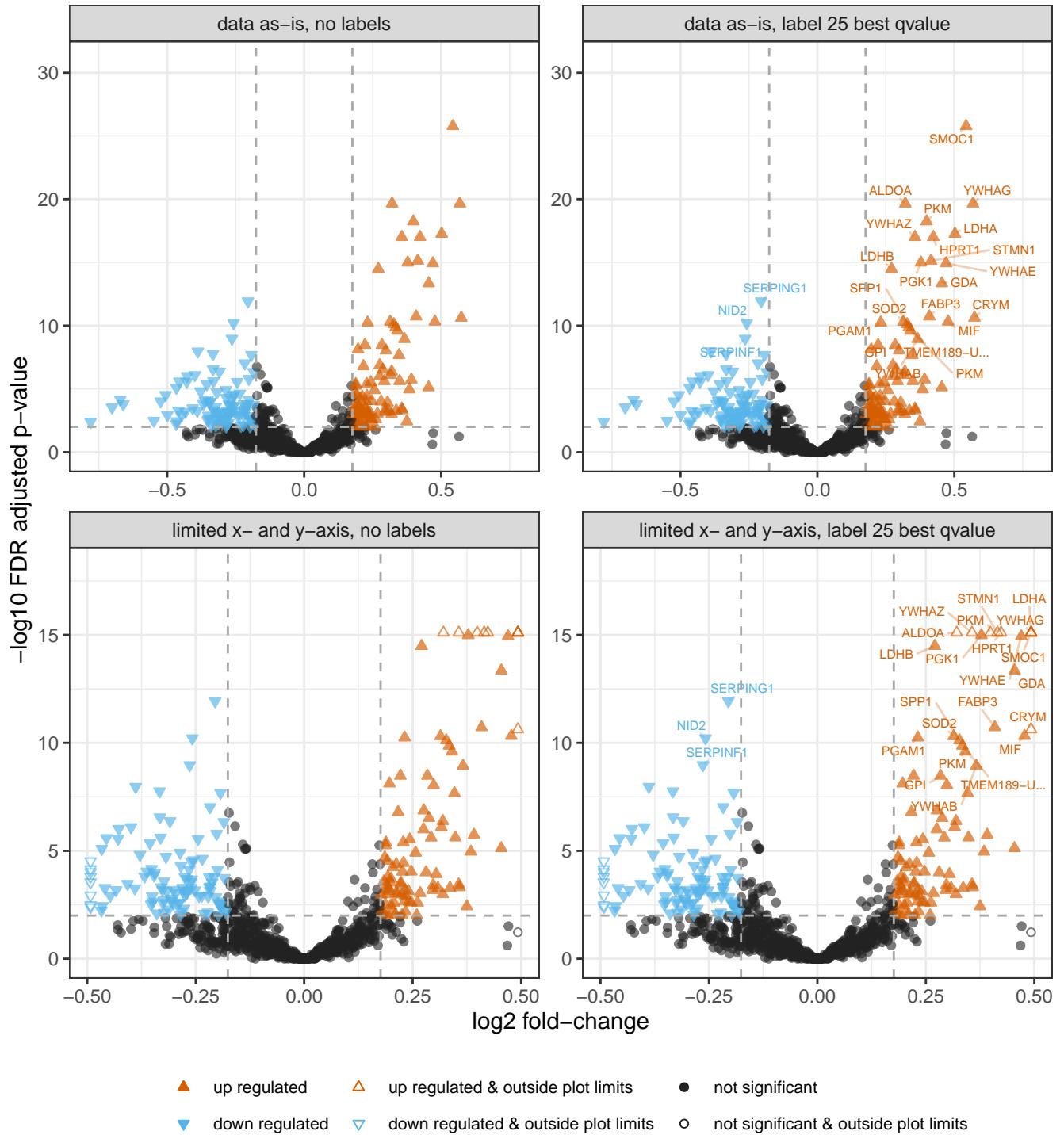
2.1.1 volcano

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.

Labels for proteins that are more than 12 characters long are truncated for visual clarity (indicated by trailing ...). For protein identifiers that are ambiguous, e.g. a protein-group with assigned genes “gene1a;gene1b”, only the first label/ID is shown for visual clarity (indicated by trailing *).

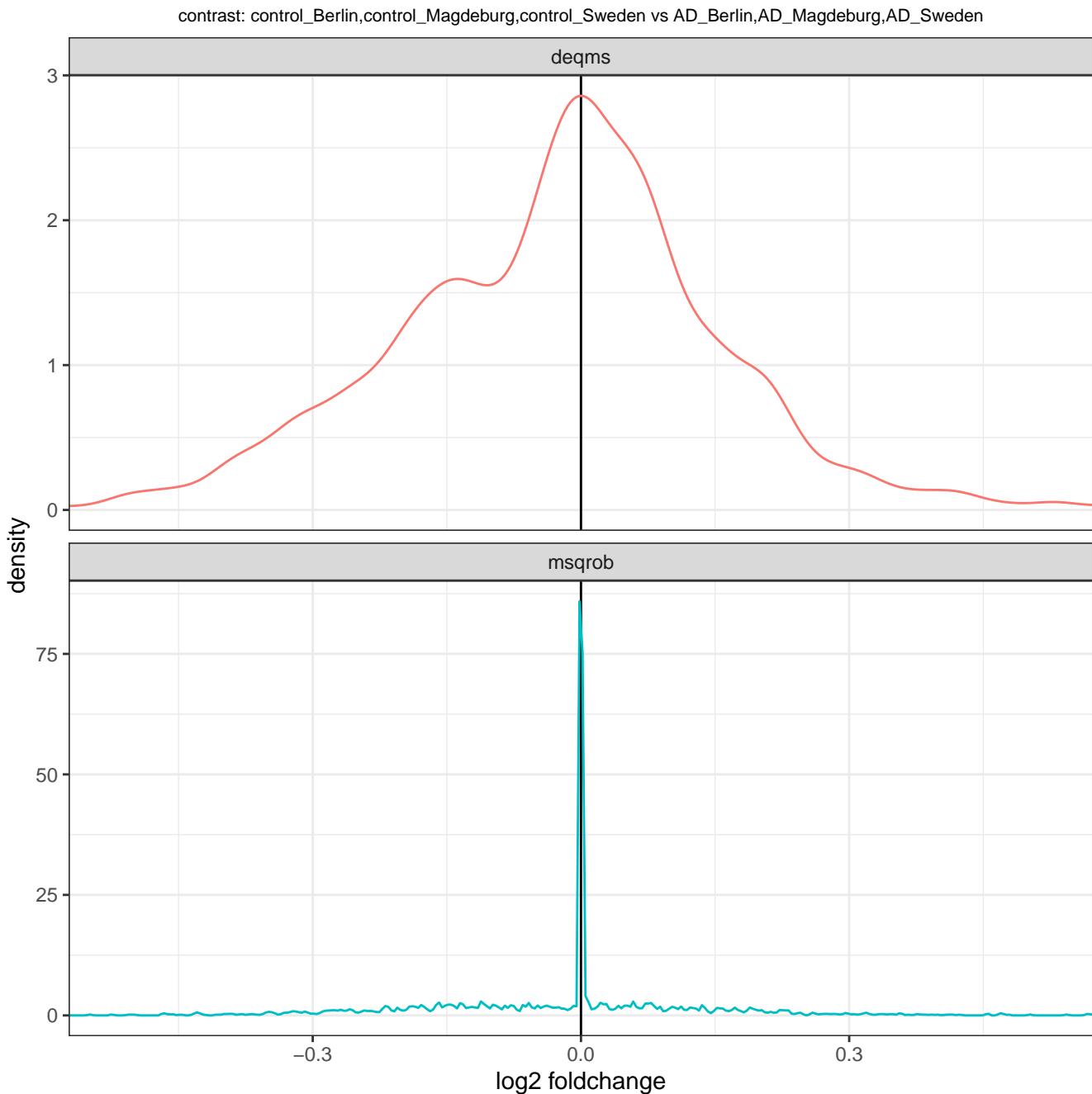




2.1.2 foldchange distribution

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



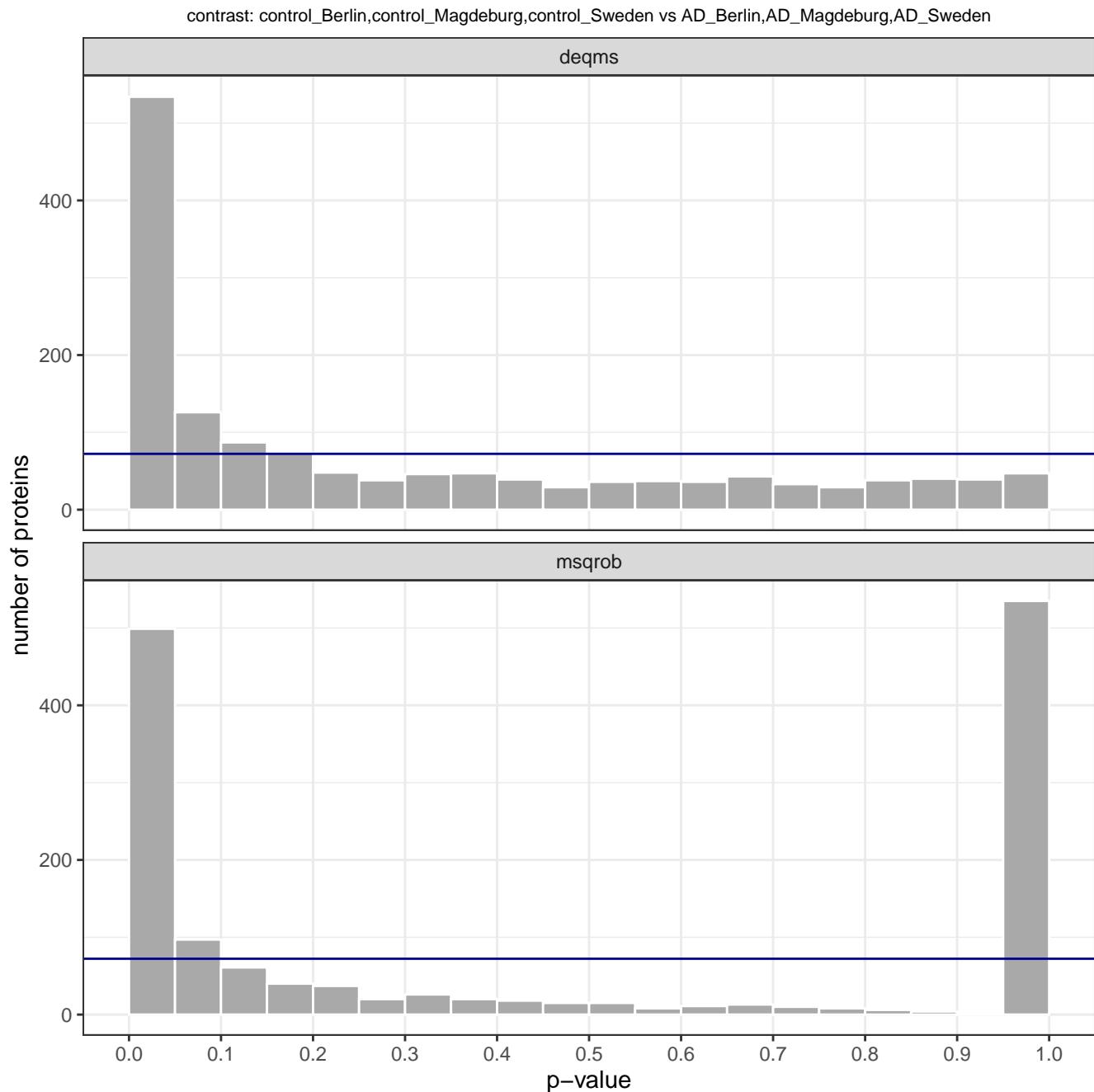
2.1.3 p-value distribution

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/>

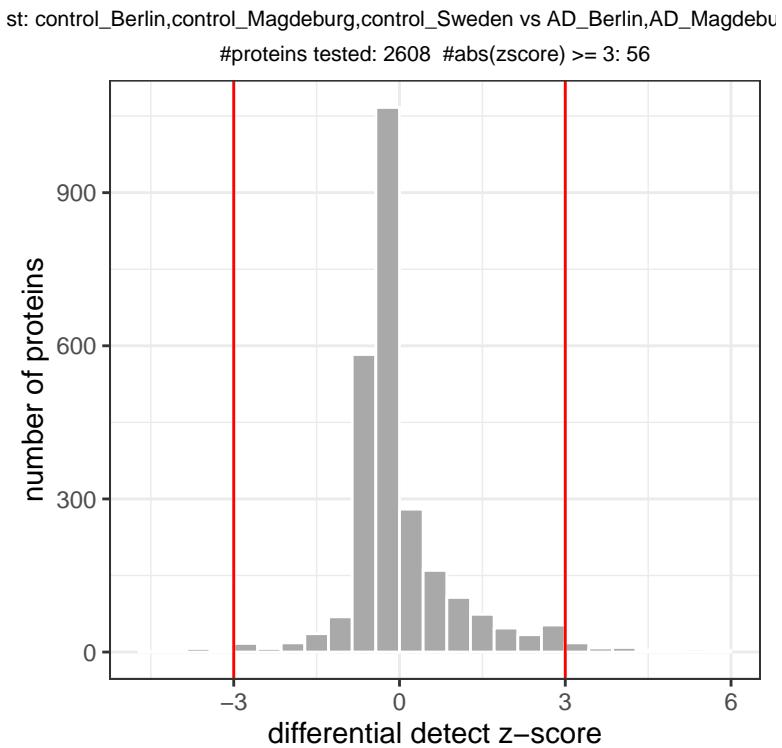
note; the MSqRob and MS-Empire models often yield p-value distributions that show a large peak at p-value 1, these are typically proteins with estimated log foldchanges at/near zero where these models are very sure the null hypothesis cannot be rejected



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
control_Berlin,...	vs	deqms	1445	180	ywhag, aldoa, smoc1, pkm, ldha, ywhaz, hpert1, gda, ywphae, ldhb
AD_Berlin,AD_Ma...					
control_Berlin,...	vs	msqrob	1445	197	smoc1, ywhag, aldoa, pkm, ldha, hpert1, yw-
AD_Berlin,AD_Ma...					haz, stmn1, pgk1, ywphae

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
control_Berlin,...	vs	2608	56	tubb, park7, plppr4, smarce1, cdon, dlat, glce, gsta4, rbmxl1;..., sqstm1
AD_Berlin,AD_Ma...				

4 log

```
[info] reading Spectronaut report...
[info] 22557/31469 precursors remain after selecting the 'best' precursor for each modified sequence
[info] 6882/6882 protein accessions and 2771/2771 protein groups were mapped to provided fasta file(s)
[info] contrast: control_Berlin,control_Magdeburg,control_Sweden vs AD_Berlin,AD_Magdeburg,AD_Sweden
[info] using 23 threads for multiprocessing
[progress] caching filter data took 3.2 minutes
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 2 seconds
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 1 seconds
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 1 seconds
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 2 seconds
[info] filter dataset with settings: min_detect = 8; norm_algorithm = 'vsn&modebetween_protein'; rollup_algorithm = 'maxlfq'
9164/22557 peptides were retained after filtering over all groups
11824/22557 peptides were retained after filtering within contrast: control_Berlin,control_Magdeburg,control_Sweden vs AD_Berlin,AD_Magdeburg,AD_Sweden
18265/22557 peptides were retained after filtering within each group independently ("by group")
[progress] peptide filtering and normalization took 4.6 minutes
[info] differential expression analysis for contrast: control_Berlin,control_Magdeburg,control_Sweden vs AD_Berlin,AD_Magdeburg,AD_Sweden
[info] using data from peptide filter: filter by contrast
[info] random variables used in current contrast: cohort
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 1 seconds
[info] log2 foldchange threshold estimated by bootstrap analysis: 0.176
[progress] DEqMS took 1 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) + (1 | cohort), expression ~ (1 | condition) + (1 | cohort), expression ~ (1 | condition)
[progress] msqrob took 1.1 minutes
[info] differential detection analysis: min_samples_observed=2
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 2 seconds
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 1 seconds
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 1 seconds
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 1 seconds
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 1 seconds
[progress] creating PDF report...
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 2 seconds
[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 26 seconds
[progress] RT plots: creating plots took 7 seconds
[info] skipping within-group foldchange plots for sample group 'unknown_Kiel', require at least 2 replicates
[info] skipping within-group foldchange plots for sample group 'unknown_Sweden', require at least 2 replicates
[info] No data available for CoV leave-one-out computation in sample group 'unknown_Kiel', skipping plots
[info] No data available for CoV leave-one-out computation in sample group 'unknown_Sweden', skipping plots
[progress] leave-one-out CoV plot computations took 11 seconds
[info] no CoV computation for sample group 'unknown_Kiel', require at least 3 replicates
[info] no CoV computation for sample group 'unknown_Magdeburg', require at least 3 replicates
[info] no CoV computation for sample group 'unknown_Sweden', require at least 3 replicates
[progress] peptide to protein rollup with MaxLFQ (implementation: iq) took 1 seconds
[info] no CoV computation for sample group 'unknown_Kiel', require at least 3 replicates
[info] no CoV computation for sample group 'unknown_Magdeburg', require at least 3 replicates
[info] no CoV computation for sample group 'unknown_Sweden', require at least 3 replicates
```

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_spectronaut(
  filename = "E:/DATA/PMID32485097/20200914_185722_SNA-fSN1.1_CLib1.1 with deamidation_SNA-SER1.1 BGS default_all final raw files Zet Ma
  confidence_threshold = 0.05,
  use_normalized_intensities = FALSE,
  use_irt = TRUE, do_plot = TRUE
)
dataset = import_fasta(dataset,
  files = c(
    "C:/VU.fasta/UniProt_2018-05/UP000005640_9606.fasta",
    "C:/VU.fasta/UniProt_2018-05/UP000005640_9606_additional.fasta",
    "C:/VU.fasta/UniProt_2017-06/UP000005640_9606.fasta",
    "C:/VU.fasta/UniProt_2017-06/UP000005640_9606_additional.fasta"
  )
)
dataset = import_sample_metadata(dataset,
  filename = "E:/DATA/PMID32485097/sample_metadata.xlsx"
)
dataset = setup_contrasts(dataset,
  contrast_list = list(list(
    c(
      "control_Berlin",
      "control_Magdeburg",
      "control_Sweden"
    ),
    c(
      "AD_Berlin",
      "AD_Magdeburg",
      "AD_Sweden"
    )
  )),
  random_variables = "cohort"
)
dataset = analysis_quickstart(dataset,
  filter_min_detect = 8,
  filter_by_contrast = TRUE,
  filter_topn_peptides = 0,
  filter_min_peptide_per_prot = 1,
  norm_algorithm = c(
    "vsn",
    "modebetween_protein"
  ),
  dea_algorithm = c(
    "deqms",
    "msqrob"
  ), dea_qvalue_threshold = 0.01,
  dea_log2foldchange_threshold = NA,
  diffdetect_min_samples_observed = 2,
  output_qc_report = TRUE,
  output_abundance_tables = TRUE,
  output_dir = "E:/DATA/PMID32485097/",
  output_within_timestamped_subdirectory = TRUE,
  dump_all_data = 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 4.2.1 (2022-06-23 ucrt)
os	Windows 10 x64 (build 19043)
system	x86_64, mingw32
ui	RStudio
language	(EN)
collate	English_United States.utf8
ctype	English_United States.utf8
tz	Europe/Berlin
date	2022-10-30
rstudio	2022.07.0+548 Spotted Wakerobin (desktop)
pandoc	2.18 @ C:/Program Files/RStudio/bin/quarto/bin/tools/ (via rmarkdown)

System

package	loadedversion	source
dplyr	1.0.9	CRAN (R 4.2.1)
ggplot2	3.3.6	CRAN (R 4.2.1)
msdap	1.0.3	
rlang	1.0.3	CRAN (R 4.2.1)
testthat	3.1.4	CRAN (R 4.2.1)
tibble	3.1.7	CRAN (R 4.2.1)
tidyverse	1.2.0	CRAN (R 4.2.1)

Attached packages

package	loadedversion	source
abind	1.4-5	CRAN (R 4.2.0)
affy	1.74.0	Bioconductor
affyio	1.66.0	Bioconductor
aod	1.3.2	CRAN (R 4.2.1)
askpass	1.1	CRAN (R 4.2.1)
assertthat	0.2.1	CRAN (R 4.2.1)
backports	1.4.1	CRAN (R 4.2.0)
Biobase	2.56.0	Bioconductor
BiocGenerics	0.42.0	Bioconductor
BiocManager	1.30.18	CRAN (R 4.2.1)
BiocParallel	1.30.3	Bioconductor
bit	4.0.4	CRAN (R 4.2.1)
bit64	4.0.5	CRAN (R 4.2.1)
bitops	1.0-7	CRAN (R 4.2.0)
blob	1.2.3	CRAN (R 4.2.1)
boot	1.3-28	CRAN (R 4.2.1)
brio	1.1.3	CRAN (R 4.2.1)
broom	1.0.0	CRAN (R 4.2.1)
cachem	1.0.6	CRAN (R 4.2.1)
callr	3.7.0	CRAN (R 4.2.1)
car	3.1-0	CRAN (R 4.2.1)
carData	3.0-5	CRAN (R 4.2.1)
caTools	1.18.2	CRAN (R 4.2.1)
cli	3.3.0	CRAN (R 4.2.1)
clue	0.3-61	CRAN (R 4.2.1)
cluster	2.1.3	CRAN (R 4.2.1)
codetools	0.2-18	CRAN (R 4.2.1)
colorspace	2.0-3	CRAN (R 4.2.1)
cowplot	1.1.1	CRAN (R 4.2.1)
crayon	1.5.1	CRAN (R 4.2.1)
data.table	1.14.2	CRAN (R 4.2.1)
DBI	1.1.3	CRAN (R 4.2.1)
DEqMS	1.14.0	bioc_xgit (@a30da3599ce70afb20527ef3c001d4c15a51a662)
desc	1.4.1	CRAN (R 4.2.1)
devtools	2.4.3	CRAN (R 4.2.1)
diann	1.0.1	Github (vdemichev/diann-rpackage@af538f6e2cd5ab715e1381632e17cb8f234ebf53)
digest	0.6.29	CRAN (R 4.2.1)
doParallel	1.0.17	CRAN (R 4.2.1)
doRNG	1.8.2	CRAN (R 4.2.1)
ellipsis	0.3.2	CRAN (R 4.2.1)
evaluate	0.15	CRAN (R 4.2.1)
fansi	1.0.3	CRAN (R 4.2.1)
farver	2.1.1	CRAN (R 4.2.1)
fastmap	1.1.0	CRAN (R 4.2.1)
forcats	0.5.1	CRAN (R 4.2.1)
foreach	1.5.2	CRAN (R 4.2.1)
formatR	1.12	CRAN (R 4.2.1)
fs	1.5.2	CRAN (R 4.2.1)
generics	0.1.3	CRAN (R 4.2.1)
ggpibr	0.4.0	CRAN (R 4.2.1)

package	loadedversion	source
ggrepel	0.9.1	CRAN (R 4.2.1)
ggsignif	0.6.3	CRAN (R 4.2.1)
glue	1.6.2	CRAN (R 4.2.1)
gplots	3.1.3	CRAN (R 4.2.1)
gridExtra	2.3	CRAN (R 4.2.1)
gtable	0.3.0	CRAN (R 4.2.1)
gtools	3.9.2.2	CRAN (R 4.2.1)
hms	1.1.1	CRAN (R 4.2.1)
htmltools	0.5.2	CRAN (R 4.2.1)
impute	1.70.0	Bioconductor
iq	1.9.6	CRAN (R 4.2.1)
IRanges	2.30.0	Bioconductor
iterators	1.0.14	CRAN (R 4.2.1)
itertools	0.1-3	CRAN (R 4.2.1)
KernSmooth	2.23-20	CRAN (R 4.2.1)
knitr	1.39	CRAN (R 4.2.1)
labeling	0.4.2	CRAN (R 4.2.0)
lattice	0.20-45	CRAN (R 4.2.1)
lifecycle	1.0.1	CRAN (R 4.2.1)
limma	3.52.2	Bioconductor
lme4	1.1-30	CRAN (R 4.2.1)
magrittr	2.0.3	CRAN (R 4.2.1)
MALDIquant	1.21	CRAN (R 4.2.1)
MASS	7.3-57	CRAN (R 4.2.1)
Matrix	1.5-1	CRAN (R 4.2.1)
matrixStats	0.62.0	CRAN (R 4.2.1)
memoise	2.0.1	CRAN (R 4.2.1)
mgcv	1.8-40	CRAN (R 4.2.1)
minqa	1.2.4	CRAN (R 4.2.1)
missForest	1.5	CRAN (R 4.2.1)
MsCoreUtils	1.8.0	Bioconductor
msEmpiRe	0.1.0	Github (zimmerlab/MS-EmpiRe@8a85757c8d604014130ee9d379aa8cfcb05e3855)
MSnbase	2.22.0	Bioconductor
munsell	0.5.0	CRAN (R 4.2.1)
mzID	1.34.0	Bioconductor
mzR	2.30.0	Bioconductor
ncdf4	1.19	CRAN (R 4.2.0)
nlme	3.1-157	CRAN (R 4.2.1)
nloptr	2.0.3	CRAN (R 4.2.1)
openssl	2.0.2	CRAN (R 4.2.1)
openxlsx	4.2.5	CRAN (R 4.2.1)
patchwork	1.1.1	CRAN (R 4.2.1)
pbkrtest	0.5.1	CRAN (R 4.2.1)
pcaMethods	1.88.0	Bioconductor
pdftools	3.3.0	CRAN (R 4.2.1)
Peptides	2.4.4	CRAN (R 4.2.1)
pillar	1.7.0	CRAN (R 4.2.1)
pkgbuild	1.3.1	CRAN (R 4.2.1)
pkgconfig	2.0.3	CRAN (R 4.2.1)
pkgload	1.3.0	CRAN (R 4.2.1)

package	loadedversion	source
plyr	1.8.7	CRAN (R 4.2.1)
preprocessCore	1.58.0	Bioconductor
prettyunits	1.1.1	CRAN (R 4.2.1)
pROC	1.18.0	CRAN (R 4.2.1)
processx	3.7.0	CRAN (R 4.2.1)
progress	1.2.2	CRAN (R 4.2.1)
ProtGenerics	1.28.0	Bioconductor
ps	1.7.1	CRAN (R 4.2.1)
purrr	0.3.4	CRAN (R 4.2.1)
qpdf	1.2.0	CRAN (R 4.2.1)
R.cache	0.15.0	CRAN (R 4.2.1)
R.methodsS3	1.8.2	CRAN (R 4.2.0)
R.oo	1.25.0	CRAN (R 4.2.0)
R.utils	2.12.0	CRAN (R 4.2.1)
R6	2.5.1	CRAN (R 4.2.1)
randomForest	4.7-1.1	CRAN (R 4.2.1)
rbibutils	2.2.8	CRAN (R 4.2.1)
RColorBrewer	1.1-3	CRAN (R 4.2.0)
Rcpp	1.0.9	CRAN (R 4.2.1)
RcppEigen	0.3.3.9.2	CRAN (R 4.2.1)
Rdpack	2.3.1	CRAN (R 4.2.1)
readr	2.1.2	CRAN (R 4.2.1)
remotes	2.4.2	CRAN (R 4.2.1)
reshape2	1.4.4	CRAN (R 4.2.1)
RhpcBLASctl	0.21-247.1	CRAN (R 4.2.0)
rmarkdown	2.14	CRAN (R 4.2.1)
rngtools	1.5.2	CRAN (R 4.2.1)
rprojroot	2.0.3	CRAN (R 4.2.1)
RSQLite	2.2.14	CRAN (R 4.2.1)
rstatix	0.7.0	CRAN (R 4.2.1)
rstudioapi	0.13	CRAN (R 4.2.1)
S4Vectors	0.34.0	Bioconductor
scales	1.2.0	CRAN (R 4.2.1)
sessioninfo	1.2.2	CRAN (R 4.2.1)
stringi	1.7.6	CRAN (R 4.2.0)
stringr	1.4.0	CRAN (R 4.2.1)
styler	1.7.0	CRAN (R 4.2.1)
tidyselect	1.1.2	CRAN (R 4.2.1)
tinytex	0.40	CRAN (R 4.2.1)
tzdb	0.3.0	CRAN (R 4.2.1)
useThis	2.1.6	CRAN (R 4.2.1)
utf8	1.2.2	CRAN (R 4.2.1)
variancePartition	1.26.0	bioc_xgit (@b1731297b2f7335ef9155c6ac7df0e246787d9d6)
vctrs	0.4.1	CRAN (R 4.2.1)
viridis	0.6.2	CRAN (R 4.2.1)
viridisLite	0.4.0	CRAN (R 4.2.1)
vroom	1.5.7	CRAN (R 4.2.1)
vsn	3.64.0	Bioconductor
withr	2.5.0	CRAN (R 4.2.1)
xfun	0.31	CRAN (R 4.2.1)

package	loadedversion	source
XML	3.99-0.10	CRAN (R 4.2.0)
xtable	1.8-4	CRAN (R 4.2.1)
yaml	2.3.5	CRAN (R 4.2.0)
zip	2.2.0	CRAN (R 4.2.1)
zlibbioc	1.42.0	Bioconductor

Packages that are not attached