Report for `PEP` Section in `mzTab` File
example_6

The `PEP` section of the `mzTab` file contains $22,310$ quantified peptide features measured in 2 samples.

	number of peptides	
quantified	22,310	100%
quantified (any zero)	5,772	25.87%
quantified (any NaN)	0	0%
identified (total)	22,310	100%
identified (unique modified)	9,807	43.96%
identified (unique stripped)	8,214	36.82%

Table 1: Total number of quantified and identified peptides. (any zero) corresponds to peptides which are absent in one or more samples. (any NaN) corresponds to peptides which could not be quantified due to overlapping peptide features.

mod	specificity	number
Carbamidomethyl	C	4611
Label:13C(6)	K	3957
Label:13C(6)	R	2625

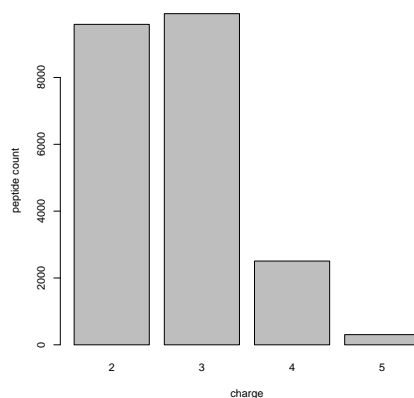Table 2: Statistics of modifications.

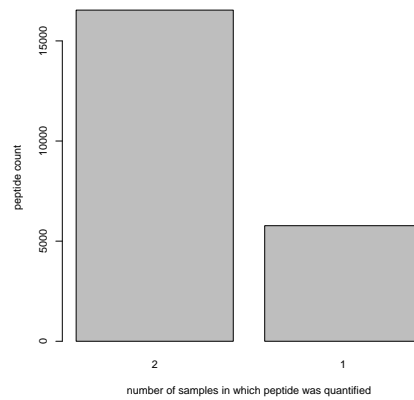Figure 1: Charge distribution of peptide quantifications.
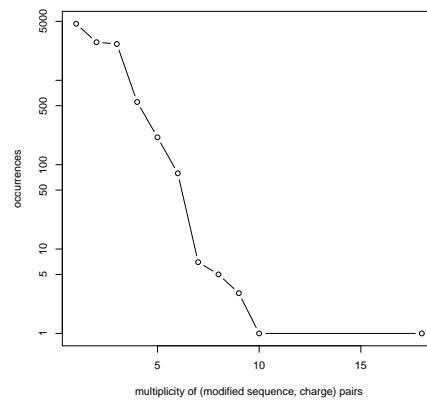
Figure 2: Frequency plot of peptide quantifications.

Figure 3: (modified sequence, charge) pair multiplicity vs frequency plot. Each peptide feature (characterised by a (possibly) modified peptide sequence and a charge state) should ideally occur only once in the analysis. In other words, peptides of multiplicity 1 should have a very high frequency. The plot below should show a significant spike on the left and can be used as QC of the analysis.
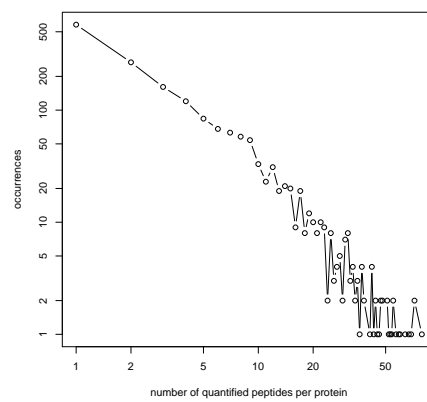

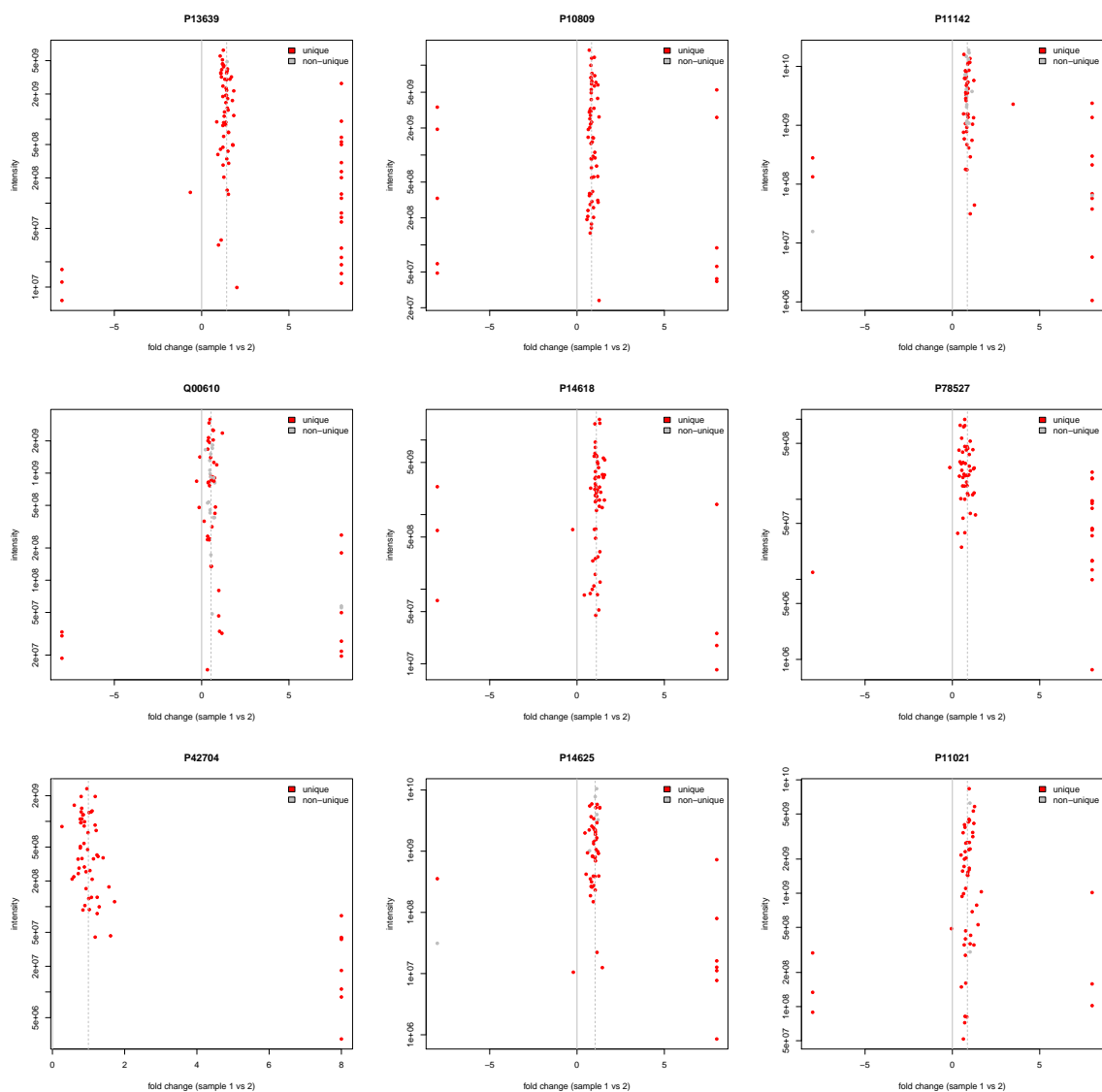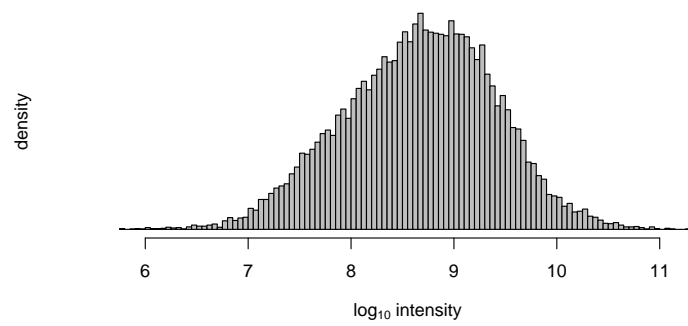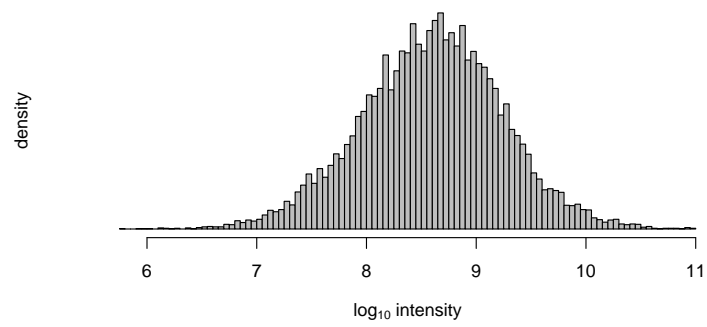Figure 4: Number of quantified peptides per protein.

Figure 5: Fold changes of peptide abundances 1 and 2. For proteins with the largest number of quantified peptides.

(a) peptide abundances 1, median(intensity) = $454, 615, 008$

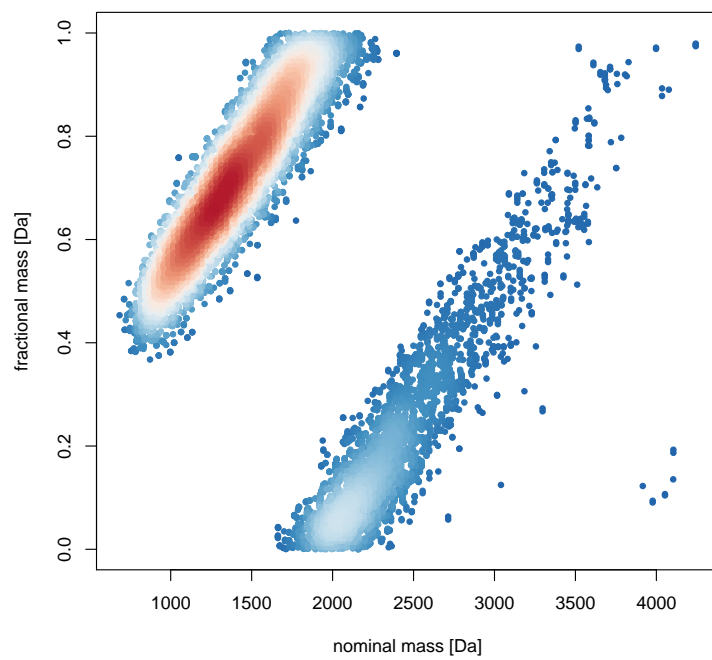(b) peptide abundances 2, median(intensity) = $243, 896, 504$

Figure 6: peptide abundance distributions.
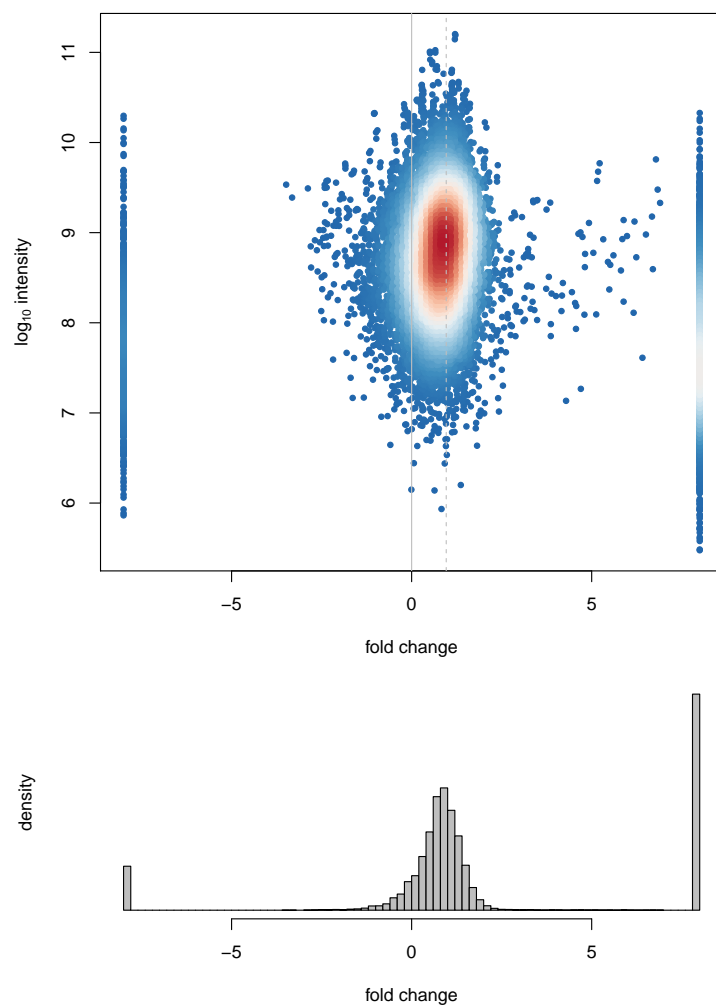
Figure 7: Kendrick nominal fractional mass plot

Figure 8: Fold changes of peptide abundances 1 and 2.
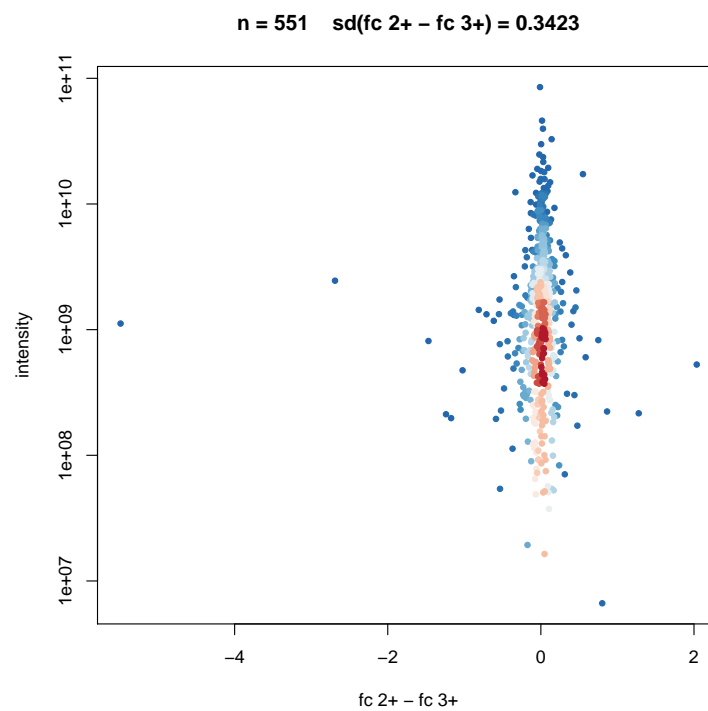median(fc) = 0.9609   sd(fc) = 3.6836

Figure 9: Fold changes of the same peptide in charge 2+ and 3+ are expected to be identical. Here we plot the difference of the 2+ and 3+ fold changes of sample 1 vs. sample 2 of all peptides which were identified and quantified in both charge states.
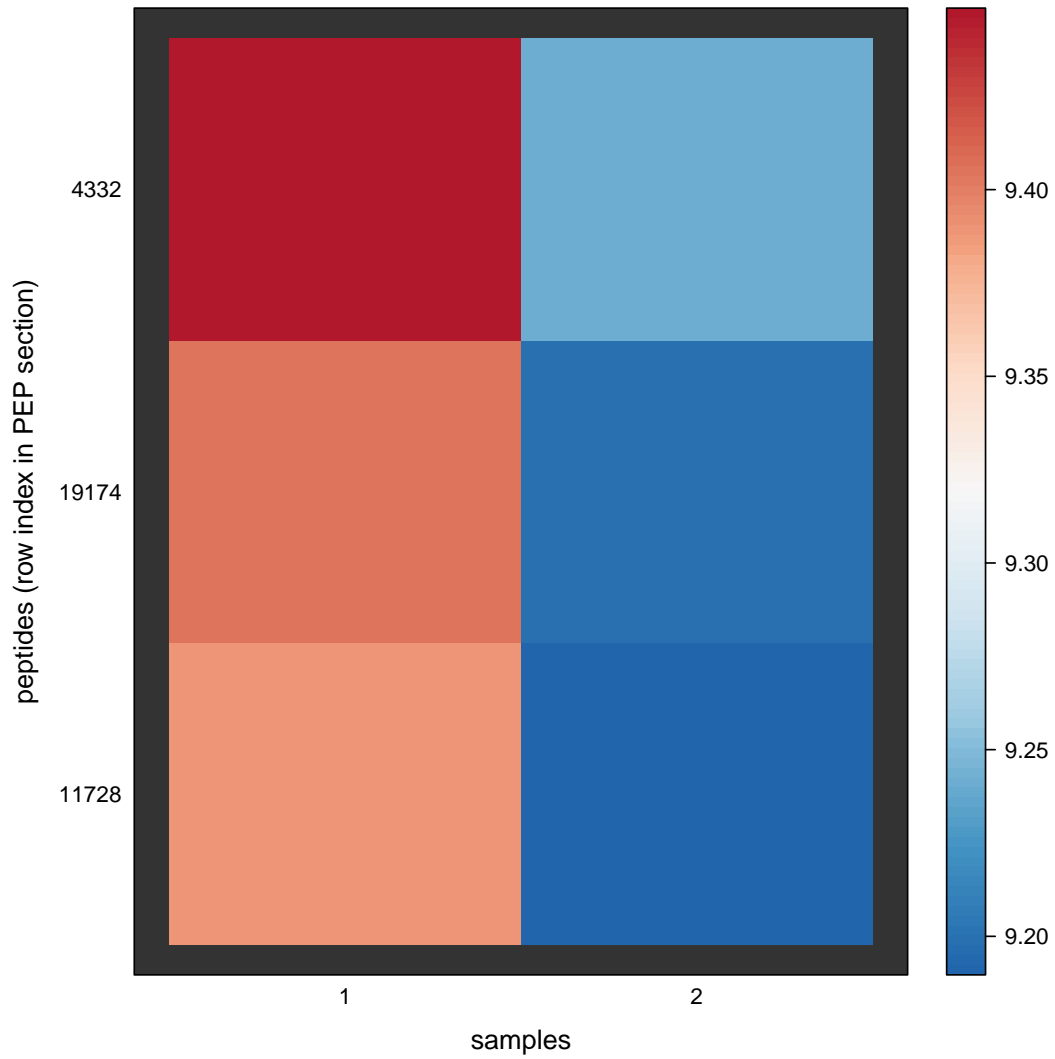
Figure 10: Logarithmic peptide abundances for all peptides of interest.

row index	modified sequence	accession	charge	retention time	m/z
4332	GILFVGSGVSGGEEGAR	P52209	2	2334.78	796.41
19174	GILFVGSGVSGGEEGAR	P52209	2	2335.85	796.41
11728	GILFVGSGVSGGEEGAR	P52209	2	2335.49	796.41

Table 3: Peptides of interest. Please note that the script requires a vector of *stripped* peptides sequences, but in the above table we list the *modified* peptide sequences.

row index	modified sequence	accession	charge	retention time	m/z
14127	IAIYELLFK	P46783	2	3767.59	555.33
6755	IAIYELLFK	P46783	2	3772.74	555.33
21653	IAIYELLFK	P46783	2	3770.19	555.33
9363	HPELADKNVPNLHVMK	P46783	4	1355.06	461.25
2002	KAEAGAGSATEFQFR	P46783	3	1385.40	523.93
16849	KAEAGAGSATEFQFR	P46783	3	1391.61	523.93

| 9428	KAEAGAGSATEFQFR	P46783	3	1387.19	523.93
16806	HPELADKNVPNLHVMK	P46783	4	1369.56	461.25
1938	HPELADKNVPNLHVMK	P46783	4	1356.00	461.25
13450	DYLHLPPEIVPATLR	P46783	3	3288.28	578.66
16052	NVPNLHVMK	P46783	3	1030.46	351.20
20332	DYLHLPPEIVPATLRR	P46783	4	2891.25	473.27
1999	KAEAGAGSATEFQFR	P46783	2	1384.54	785.39
8586	NVPNLHVMK	P46783	3	1021.79	351.20
12851	DYLHLPPEIVPATLRR	P46783	4	2888.02	473.27
9422	KAEAGAGSATEFQFR	P46783	3	1382.02	523.93
9426	KAEAGAGSATEFQFR	P46783	2	1386.17	785.39
16851	KAEAGAGSATEFQFR	P46783	2	1392.91	785.39
14761	GQNLLLTNLQTIQGILER	P12270	3	4285.16	675.39
22267	GQNLLLTNLQTIQGILER	P12270	3	4285.98	675.39
7378	GQNLLLTNLQTIQGILER	P12270	3	4293.51	675.39
303	LSDK(Label:13C(6))VVASVK(...	P12270	3	650.82	349.21
7726	LSDKVVASVK	P12270	3	645.78	349.21
945	NIEELQQQNQR	P12270	2	929.09	700.36
3157	NLDVQLLDTKR	P12270	3	1847.85	438.92

Table 4: Proteins of interest.
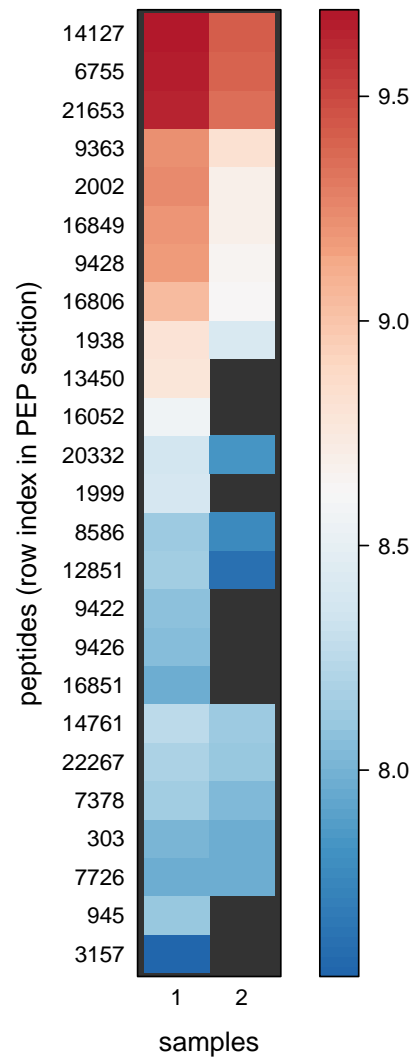
Figure 11: Logarithmic peptide abundances for all proteins of interest.
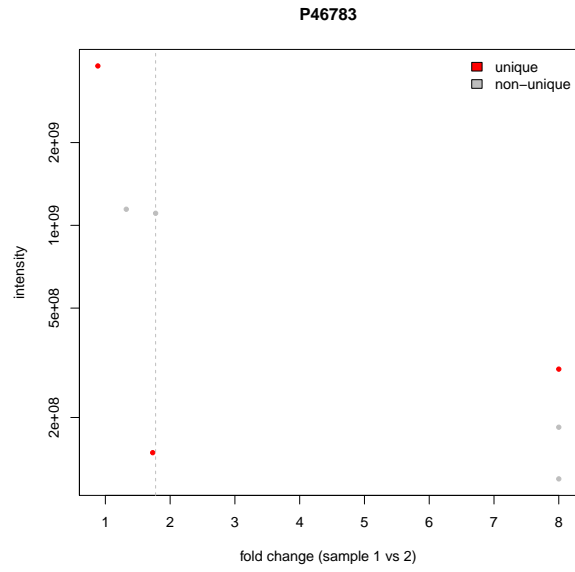
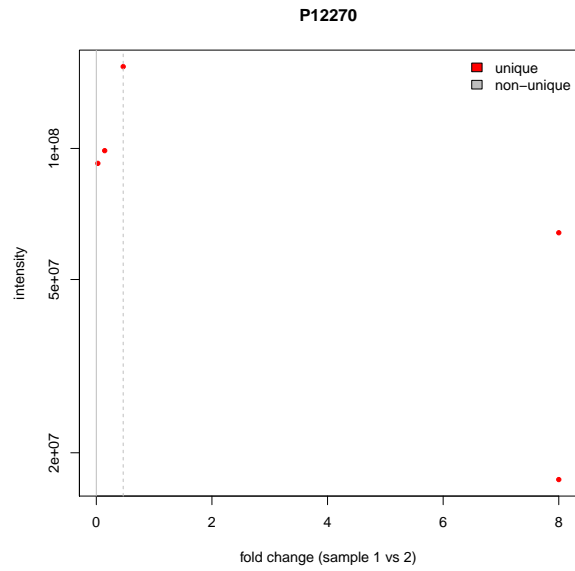Figure 12: Fold changes of peptide abundances 1 and 2 for first protein of interest.

Figure 13: Fold changes of peptide abundances 1 and 2 for second protein of interest.