

Report for PEP Section in mzTab File

example_5

The PEP section of the **mzTab** file contains 26,794 quantified peptide features measured in 54 samples.

	number of peptides	
quantified	26,794	100%
quantified (any zero)	25,270	94.31%
quantified (any NaN)	0	0%
identified (total)	26,794	100%
identified (unique modified)	21,658	80.83%
identified (unique stripped)	19,580	73.08%

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
Oxidation	M	4942
Methylthio	C	4473
Dioxidation	M	112
Label:13C(6)15N(2)	K	26
Label:13C(6)15N(4)	R	17

Table 2: Statistics of modifications.

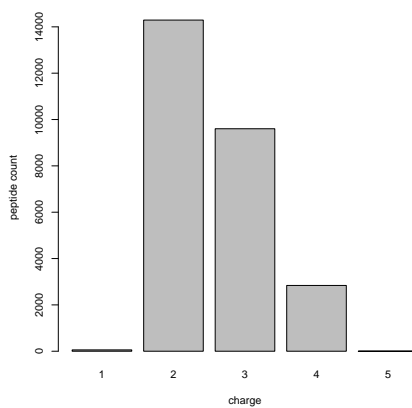


Figure 1: Charge distribution of peptide quantifications.

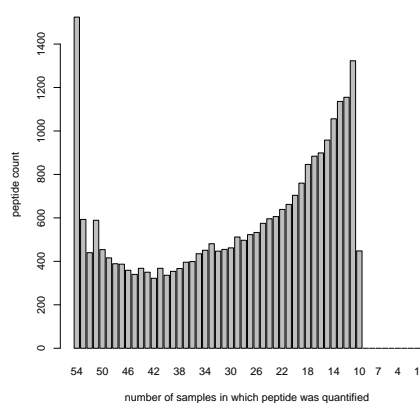


Figure 2: Frequency plot of peptide quantifications.

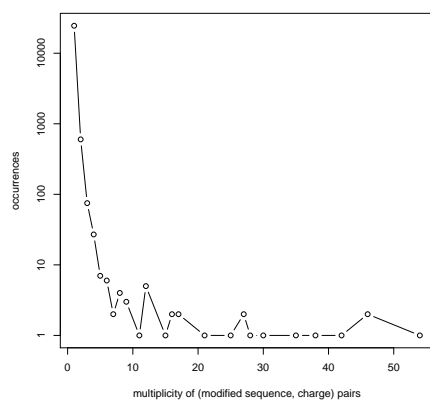


Figure 3: (modified sequence, charge) pair multiplicity vs frequency plot. Each peptide feature (characterised by a (possibly) modified peptide sequence and a charge state) should ideally occur only once in the analysis. In other words, peptides of multiplicity 1 should have a very high frequency. The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

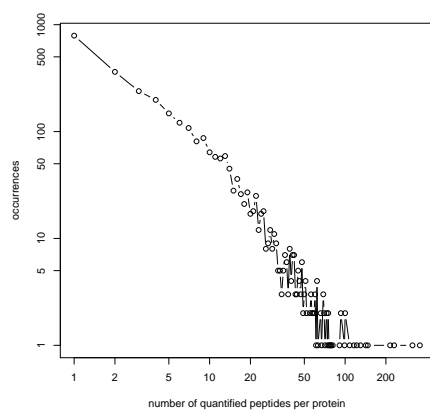


Figure 4: Number of quantified peptides per protein.

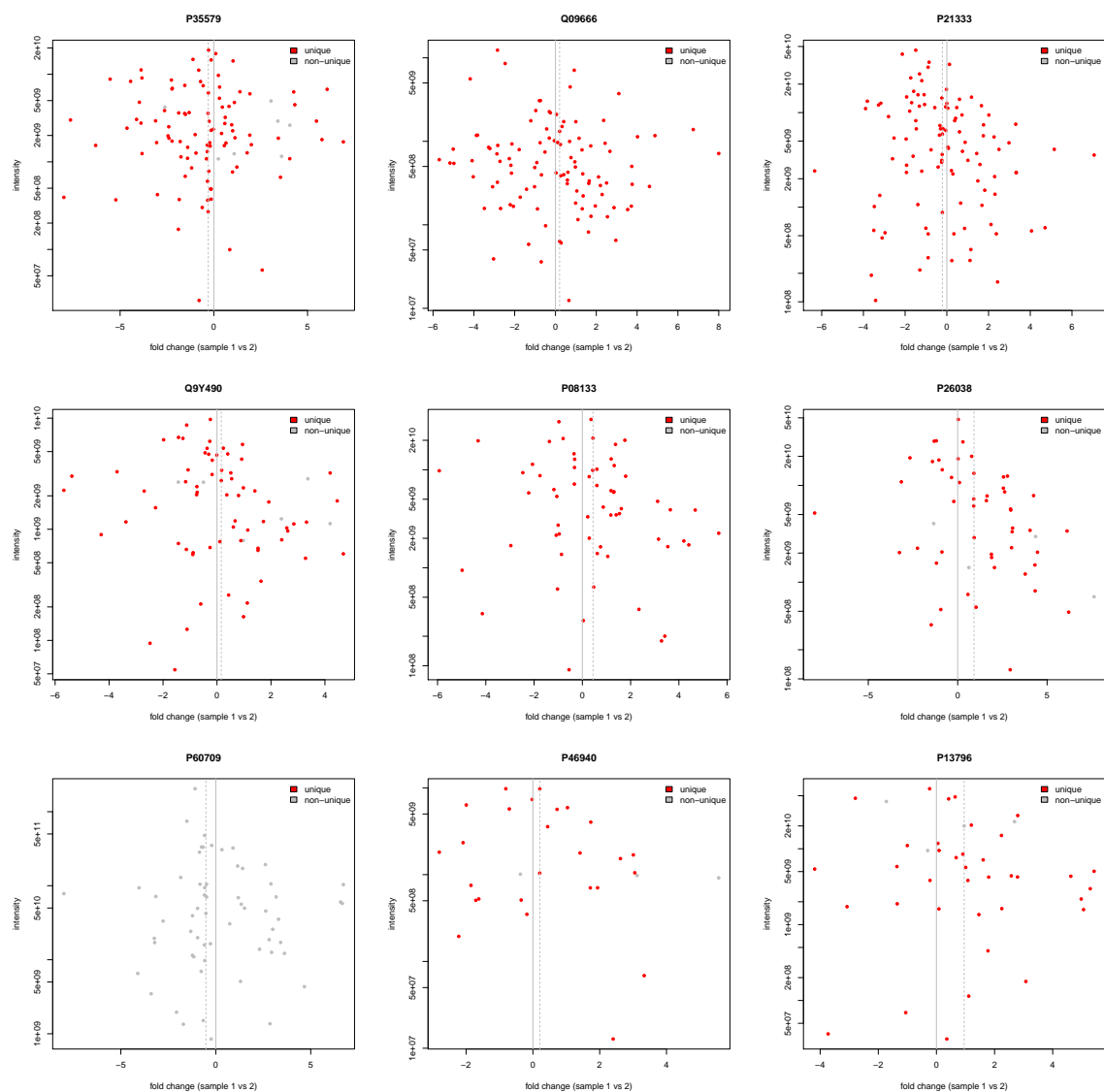
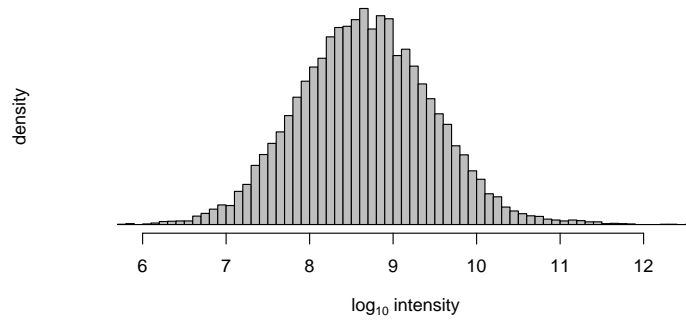
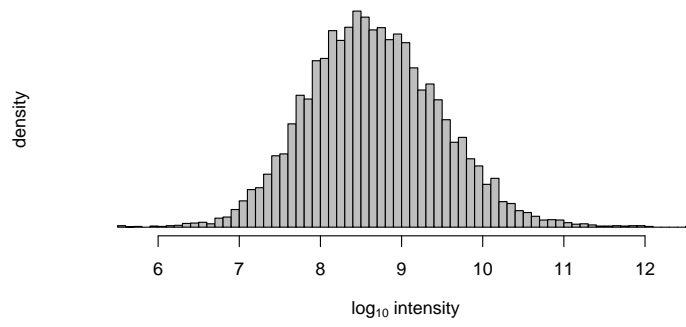


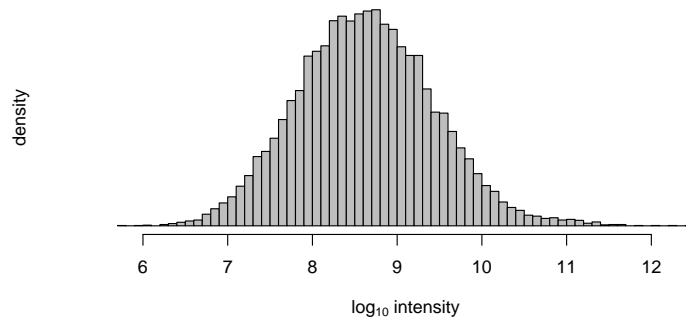
Figure 5: Fold changes of peptide abundances 1 and 2. For proteins with the largest number of quantified peptides.



(a) peptide abundances 1, $\text{median}(\text{intensity}) = 455,025,504$



(b) peptide abundances 2, $\text{median}(\text{intensity}) = 424,578,000$



(c) peptide abundances 3, $\text{median}(\text{intensity}) = 412,578,512$

Figure 6: peptide abundance distributions.

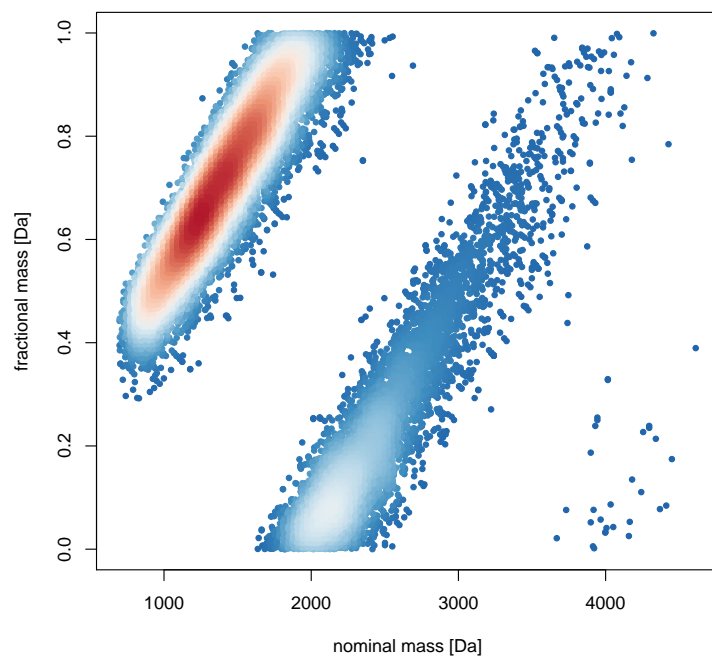


Figure 7: Kendrick nominal fractional mass plot

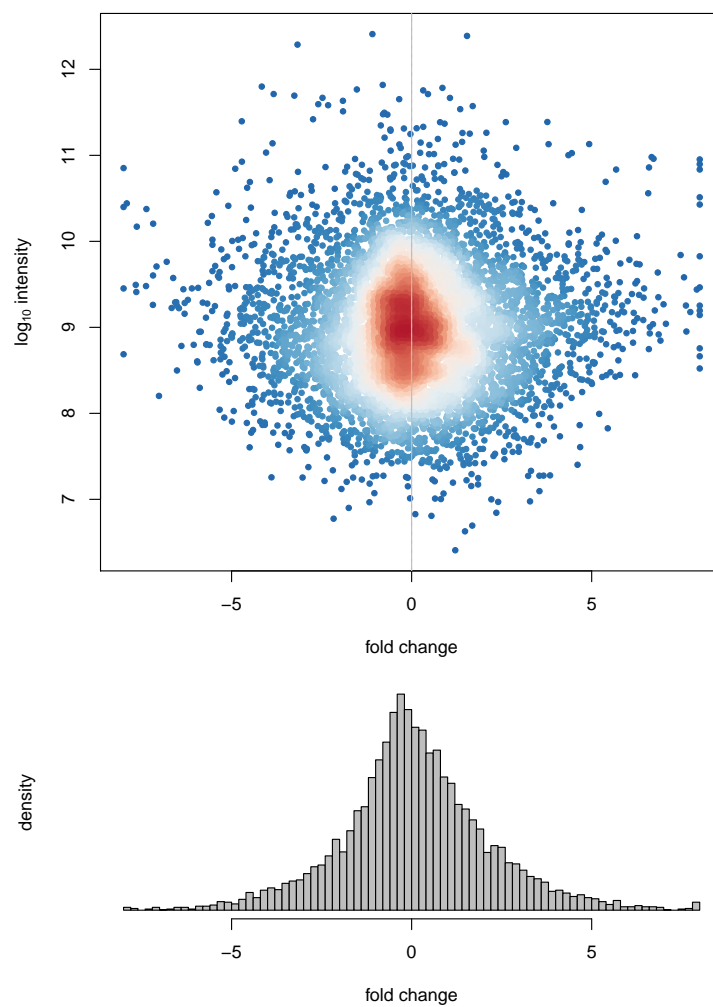


Figure 8: Fold changes of peptide abundances 1 and 2.
 $\text{median}(\text{fc}) = -0.0026$ $\text{sd}(\text{fc}) = 2.0776$

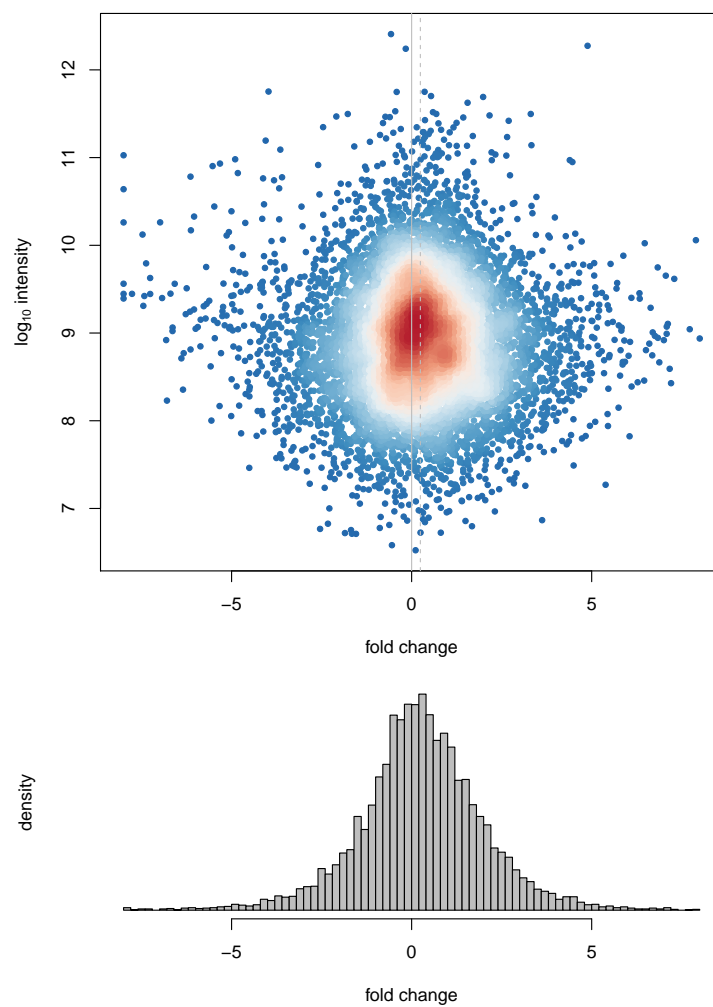


Figure 9: Fold changes of peptide abundances 1 and 3.
 $\text{median}(\text{fc}) = 0.2421$ $\text{sd}(\text{fc}) = 1.7661$

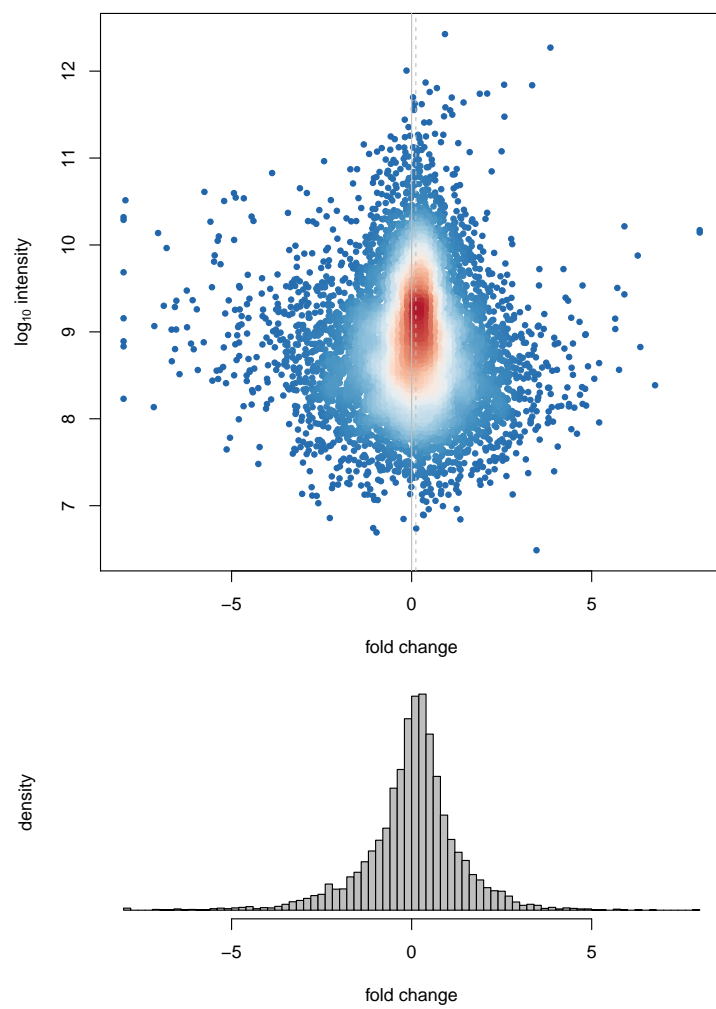


Figure 10: Fold changes of peptide abundances 2 and 3.
 $\text{median}(\text{fc}) = 0.1175$ $\text{sd}(\text{fc}) = 1.3543$

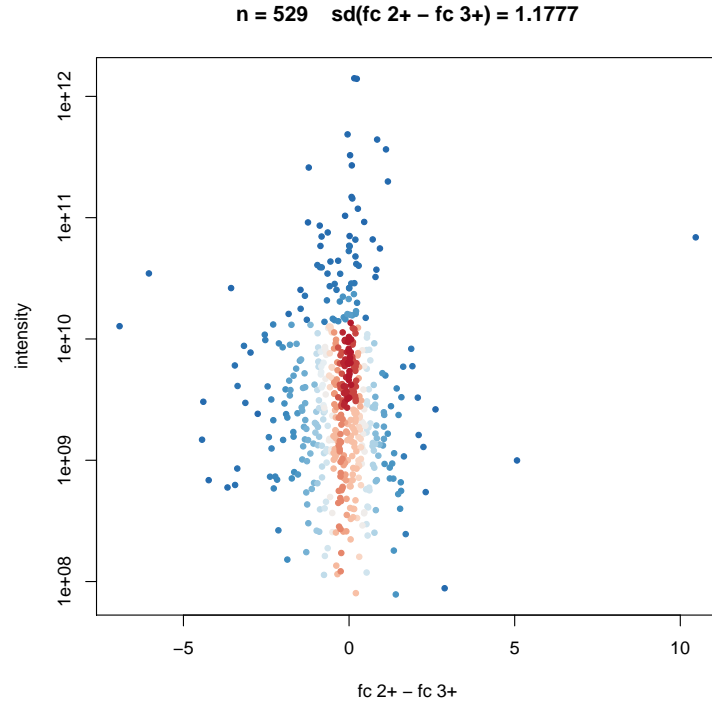


Figure 11: Fold changes of the same peptide in charge 2+ and 3+ are expected to be identical. Here we plot the difference of the 2+ and 3+ fold changes of sample 1 vs. sample 2 of all peptides which were identified and quantified in both charge states.

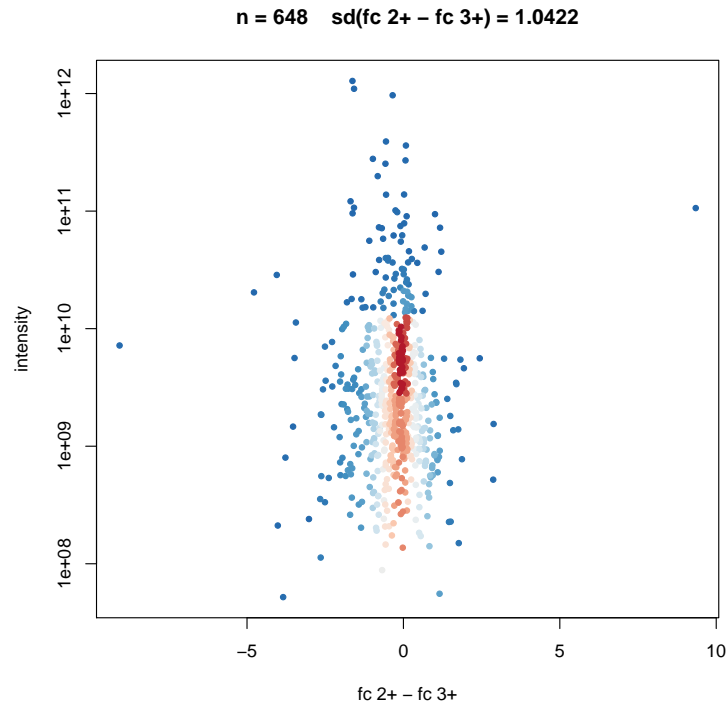


Figure 12: Fold changes of the same peptide in charge 2+ and 3+ are expected to be identical. Here we plot the difference of the 2+ and 3+ fold changes of sample 1 vs. sample 3 of all peptides which were identified and quantified in both charge states.

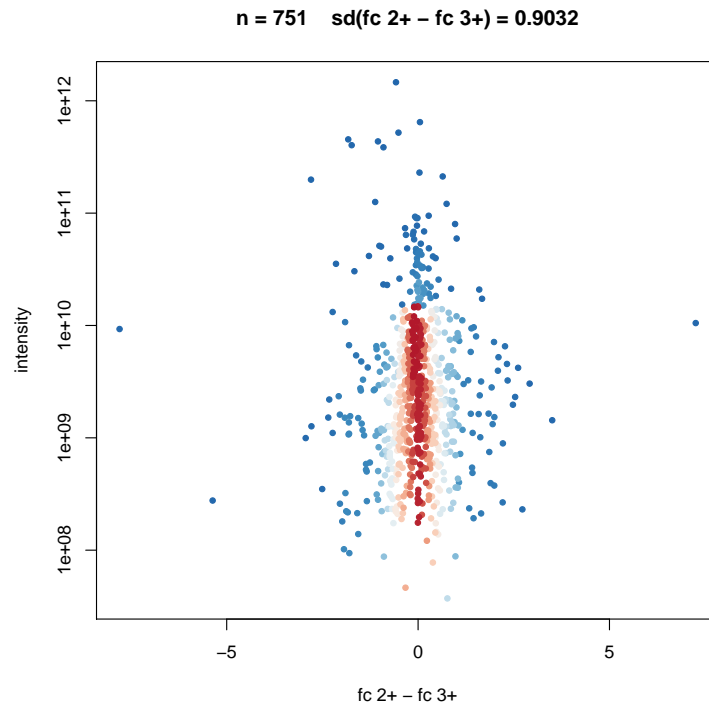


Figure 13: Fold changes of the same peptide in charge 2+ and 3+ are expected to be identical. Here we plot the difference of the 2+ and 3+ fold changes of sample 2 vs. sample 3 of all peptides which were identified and quantified in both charge states.

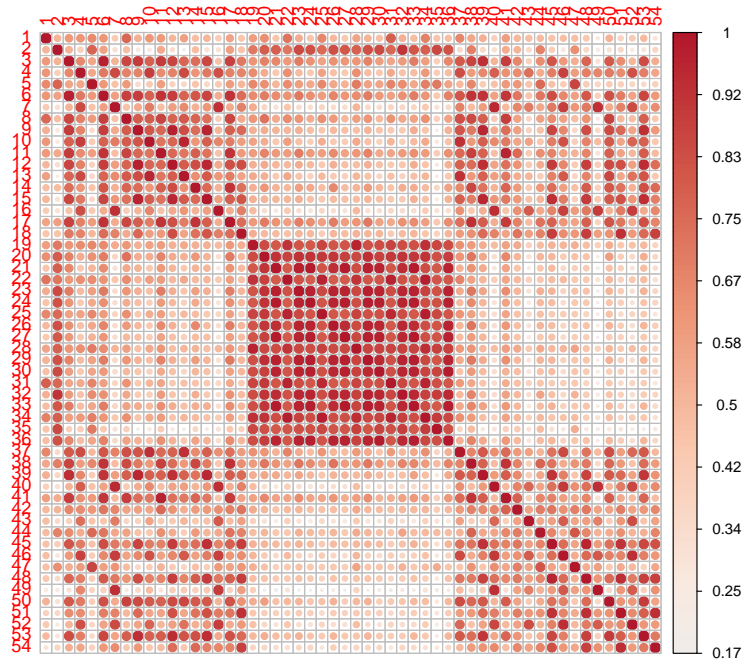


Figure 14: Pearson correlation of all peptide abundances. (min correlation = 0.1688, median correlation = 0.5843, max correlation = 1)

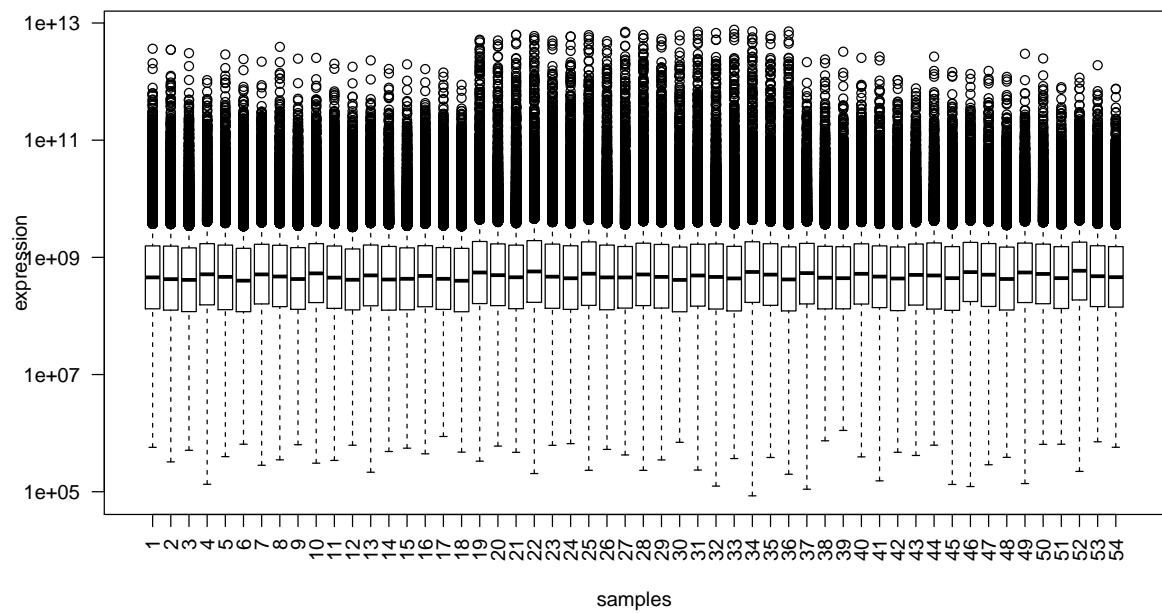


Figure 15: Boxplot of all peptide abundances.

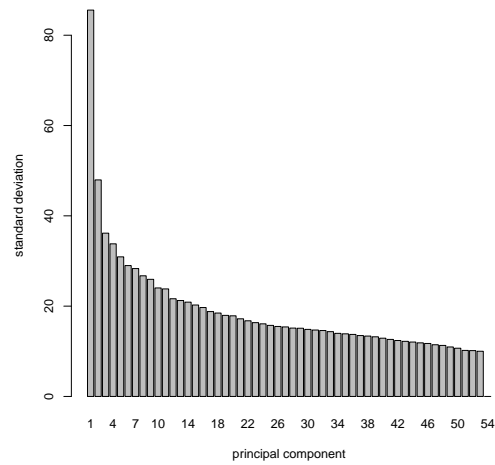


Figure 16: Standard deviation of all principal components.

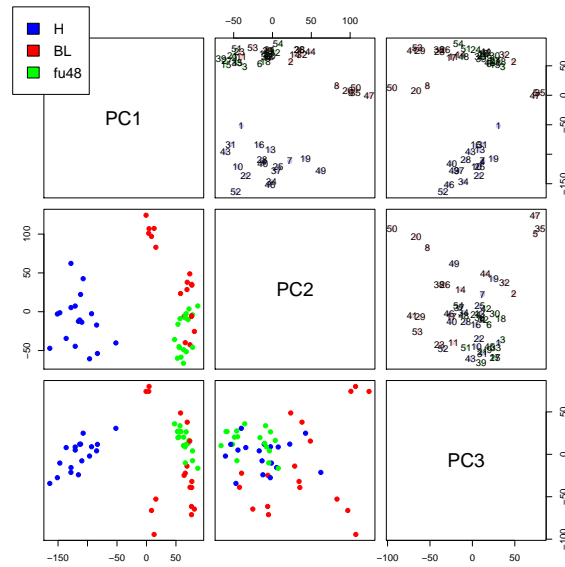


Figure 17: Principal Component Analysis of all peptides with complete quantifications. Any peptides with one or more missing values are ignored. The numbers in the upper right panels correspond to the sample IDs.

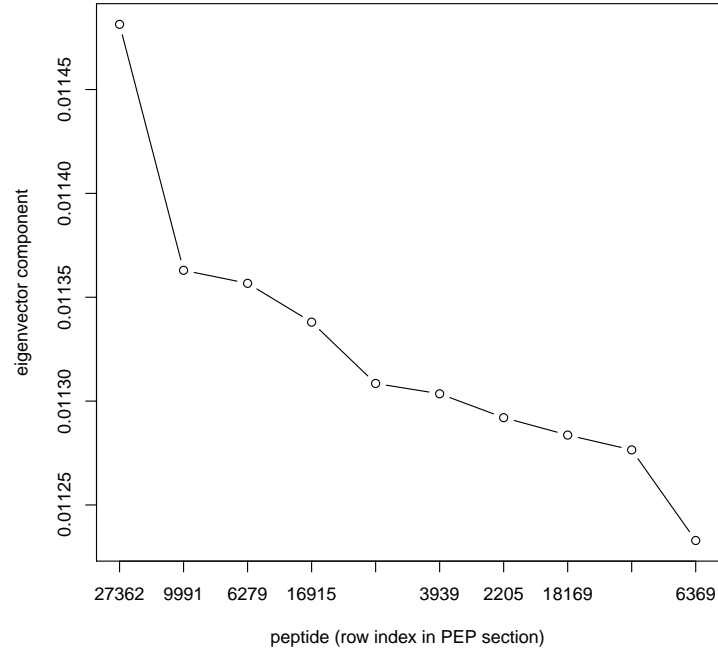


Figure 18: Most important contributions to the first principal component.

row index	modified sequence	accession	charge	retention time	m/z
27362	VLEDGKQQVQVVGGLQER	O00139	3	4606.90	642.35
9991	HLQLAIRNDEELNKLLGK	Q6FI13	4	7205.10	526.80
6279	IITLEEGDIILTGTTPK	Q6P587	2	9379.14	856.99
16915	DLM(Oxidation)TDLKSEISGDL...	P08133	2	10781.67	890.44
24319	AERELSEQIQR	P15311	3	2174.66	453.57
3939	MNVLADALK	P62244	2	6292.91	487.77
2205	VKEGM(Oxidation)NIVEAMER	P62937	3	3924.11	507.92
18169	IISDRDLLAVVFGTEKDK	P12956	4	8899.00	568.06
23075	FTPGTFTNQIAAFREPR	P08865	2	8161.59	1041.03
6369	EVVEAHVDQK	P30042	2	1625.03	577.30

Table 3: Most important contributions to the first principal component.

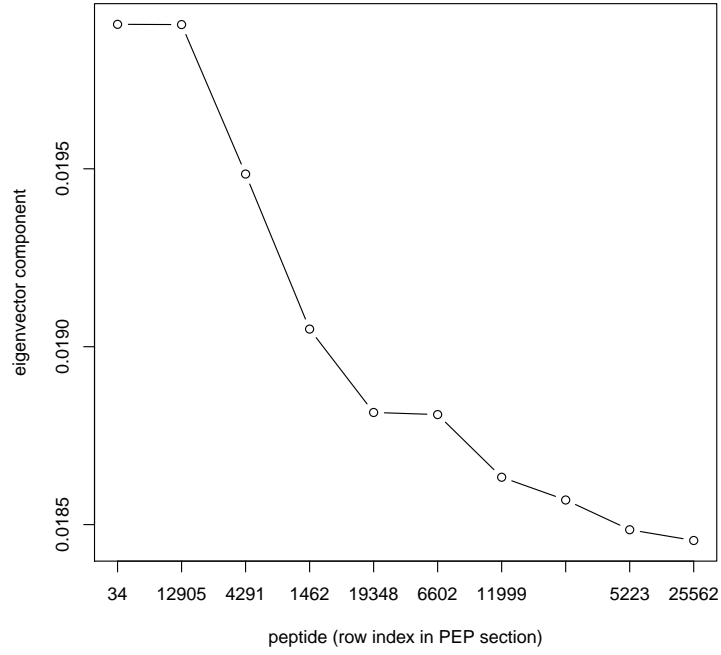


Figure 19: Most important contributions to the second principal component.

row index	modified sequence	accession	charge	retention time	m/z
34	VKASGPGLNTTGVPASLPVEFTIDA...	P21333	3	8675.02	857.13
12905	YEM(Oxidation)FAQTLQQR	P55072	2	4603.57	759.36
4291	LVDQNIFSFYLSRDPDAQPGGELM(...	P07339	3	10656.05	1129.55
1462	GNFGGSFAGSFAGGAGGHAPGVAR	P52272	3	5570.46	678.99
19348	GEGPDVDVNLPK	Q09666	2	5316.29	620.31
6602	GYTSIFNM(Oxidation)ETQQR	O43815	2	5502.98	795.86
11999	KM(Oxidation)QELLQTQDFSK	Q9H4M9	2	3666.16	806.41
20305	DGYAQILR	Q99798	2	5146.79	468.25
5223	YGDFFIR	Q9Y262	2	6543.81	459.23
25562	IYKVPSTETEALASNLMGMFEK	P31150	3	10908.34	820.41

Table 4: Most important contributions to the second principal component.

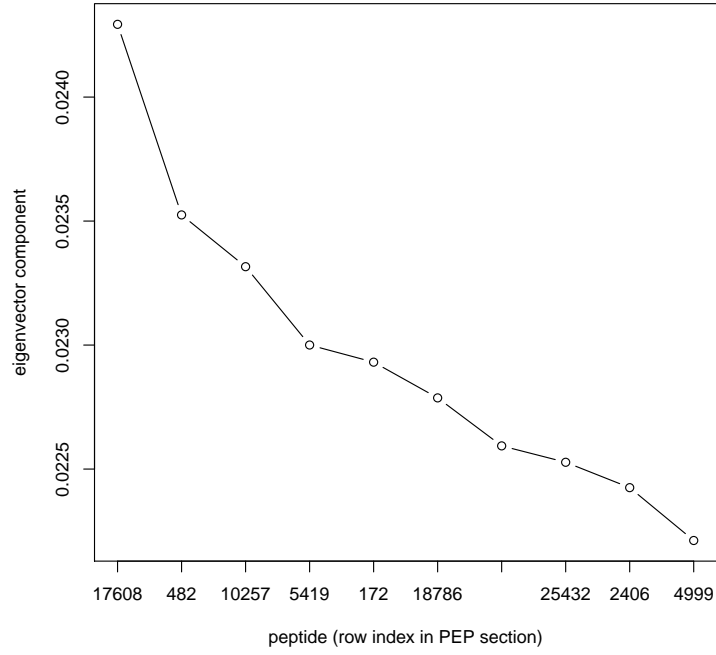


Figure 20: Most important contributions to the third principal component.

row index	modified sequence	accession	charge	retention time	m/z
17608	SKPDLESSLESLVFDLR	Q99549	3	11841.03	645.67
482	SQPAILLLTAAR	Q5JTV8	2	8038.47	627.38
10257	SAYPPPAPAYALSSPR	O14497	2	5841.52	822.92
5419	FLSGLELVK	Q96S44	2	7544.70	503.30
172	EAINVEQAFQTIAR	P51149	2	9126.45	795.42
18786	EGLELPEDEEEKKKQEEKK	Q58FG0	3	2011.00	772.39
14998	NFAELAR	Q9Y3C6	2	3437.92	410.72
25432	LSEHATAPTR	P33316	2	1547.58	541.78
2406	NREPVQLETLSIR	P62314	2	5298.33	777.93
4999	HGVYNPNKIFGVTTLDIVR	P40926	3	7526.08	715.06

Table 5: Most important contributions to the third principal component.

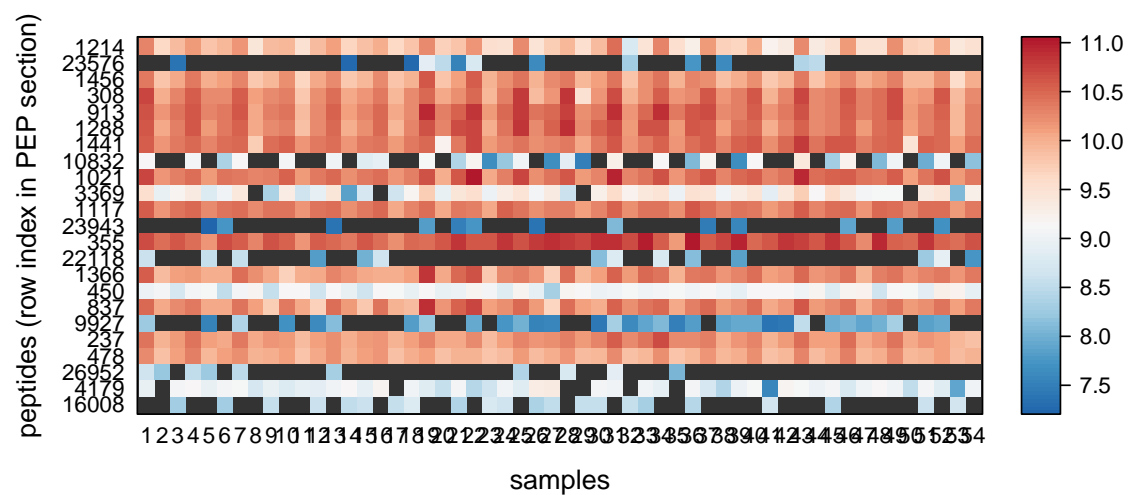


Figure 21: Logarithmic peptide abundances for all peptides of interest.

row index	modified sequence	accession	charge	retention time	m/z
1214	SSAAPPPPPR(Label:13C(6)15...	STD_01	2	1659.92	493.77
23576	SSAAPPPPPR(Label:13C(6)15...	STD_01	2	1605.04	493.76
1456	GISNEGQNASIK(Label:13C(6)...	STD_02	2	2041.39	613.32
308	HVLTSIGEK(Label:13C(6)15N...	STD_03	2	2127.71	496.29
913	DIPVPK(Label:13C(6)15N(2)...	STD_04	2	2653.71	451.28
1288	IGDYAGIK(Label:13C(6)15N(...	STD_05	2	3096.71	422.74
1441	TASEFDSAIAQDK(Label:13C(6...	STD_06	2	4266.53	695.83
10832	TASEFDSAIAQDK	STD_06	2	4271.48	691.83
1021	SAAGAFGPESLR(Label:13C(6)...	STD_07	2	4457.27	586.80
3369	SAAGAFGPESLR(Label:13C(6)...	STD_07	2	4650.67	586.80
1117	ELGQSGVDTYLQTK(Label:13C(...	STD_08	2	5741.14	773.90
23943	ELGQSGVDTYLQTK(Label:13C(...	STD_08	2	5880.07	773.90
355	GLILVGGYGTR(Label:13C(6)1...	STD_09	2	6431.53	558.33
22118	GLILVGGYGTR(Label:13C(6)1...	STD_09	2	4427.35	558.33
1366	GILFVGSGVSGGEEGAR(Label:1...	P52209	2	6781.34	801.41
450	GILFVGSGVSGGEEGAR	P52209	2	6780.92	796.41

837	SFANQPLEVVYSK(Label:13C(6...	STD_11	2	6787.30	745.39
9927	SFANQPLEVVYSK(Label:13C(6...	STD_11	2	6606.62	745.39
237	LTILEELR(Label:13C(6)15N(...	STD_12	2	7538.66	498.80
478	ELASGLSFPVGFK(Label:13C(6...	STD_14	2	9083.08	680.37
26952	ELASGLSFPVGFK(Label:13C(6...	STD_14	2	8764.02	680.37
4179	LSSEAPALFQFDLK(Label:13C(...	STD_15	2	9657.05	787.42
16008	AKGILFVGSGVSGGEEGAR	P52209	3	4941.74	597.65

Table 6: 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
2584	KAEAGAGSATEFQFR	P46783	3	3518.31	523.93
8795	SAVPPGADKKAEAGAGSATEFQFR	P46783	4	4302.34	598.80
5	DYLHLPPEIVPATLRR	P46783	3	8103.29	630.69
3119	IAIYELLFK	P46783	2	10732.68	555.33
2475	AEAGAGSATEFQFR	P46783	2	5036.50	721.34
1721	KAEAGAGSATEFQFR	P46783	2	3516.88	785.39
5284	DYLHLPPEIVPATLR	P46783	3	9558.62	578.66
9684	SAVPPGADKKAEAGAGSATEFQFR	P46783	3	4302.30	798.07
10741	GYVKEQFAWR	P46783	3	4253.53	428.56
15766	GYVKEQFAWR	P46783	2	4257.40	642.33
20490	HPELADKNVPNLHVM(Oxidation...	P46783	4	2806.75	465.25
26010	SRPETGRPRPK	P46783	3	1112.97	427.58
11596	IAIYELLFKEGVM(Oxidation)V...	P46783	3	10197.44	614.01
25208	SRPETGRPRPK	P46783	2	1112.69	640.86
3411	DYLHLPPEIVPATLRR	P46783	4	8100.98	473.27
19976	DYLHLPPEIVPATLR	P46783	2	9548.91	867.48
15136	HFYWYLTNEGIQYLR	P46783	3	9290.39	668.33
13171	LENEVEQR	P12270	2	1693.68	508.75
5886	ISTQLDFASK	P12270	2	4987.08	555.30
4982	NLQEQTVQLQSELSR	P12270	2	6830.64	886.96
11949	FKVESEQQYFEIEKR	P12270	3	4954.28	654.00
5219	TLSSVQNEVQEALQR	P12270	2	7739.83	851.44
11273	GIASDPTANIKPTPVVSTPSK	P12270	3	4902.51	789.09
22013	EKGNEILELK	P12270	2	3452.76	586.83
13646	LESALTELEQLRK	P12270	3	6908.25	510.62
15566	LLSEKEVHTK	P12270	2	1506.47	592.33
22088	LQEQVTDLR	P12270	2	3580.96	551.30
11076	LSQELEYLTEDVKR	P12270	3	6542.97	574.97
14084	LSSQIEKLEHEISHLK	P12270	4	6322.55	473.51
15882	EKEIAETRF EVAQVESLR	P12270	3	5813.97	712.04
13013	RPSTSQTVSTPAPVPVIESTE AIEA...	P12270	3	6727.12	899.14
14775	NLDVQLLDTK	P12270	2	6805.16	579.82
2595	LTIHAPPQELGPPVQR	P12270	3	5244.98	585.00
18730	FLADQQSEIDGLKGRHEK	P12270	4	3196.05	518.52
14534	GQNLLLTNLQTIQGILER	P12270	2	12445.40	1012.58
20518	KLELDILPLQEANAELSEK	P12270	3	9473.26	718.40
17342	ILLSQTTGVAIPLHASSLDDVSLAS...	P12270	3	9456.43	945.52
13673	YLDEIVKEVEAK	P12270	3	7196.92	479.26
26106	SQEQILEILR	P12270	2	7860.87	614.85
26635	KLENEVEQR	P12270	2	1635.88	572.80
26421	SAADDSEAKSNELTR	P12270	2	1834.45	797.37
24602	SAADDSEAKSNELTR	P12270	3	1865.12	531.92
23121	TKEELEAEKR	P12270	2	1216.40	616.83

23610	EGVQGPLNVSLSEEGKSSEQILEIL...	P12270	3	10181.28	951.51
26244	AADSQNSGEGNTGAAESSFSQEVSR	P12270	2	4519.04	1243.53
26016	AALKQLQEIFENYKK	P12270	4	6949.96	456.51
25260	ITELQLKLESALTELEQLRK	P12270	4	10747.62	589.59

Table 7: Proteins of interest.

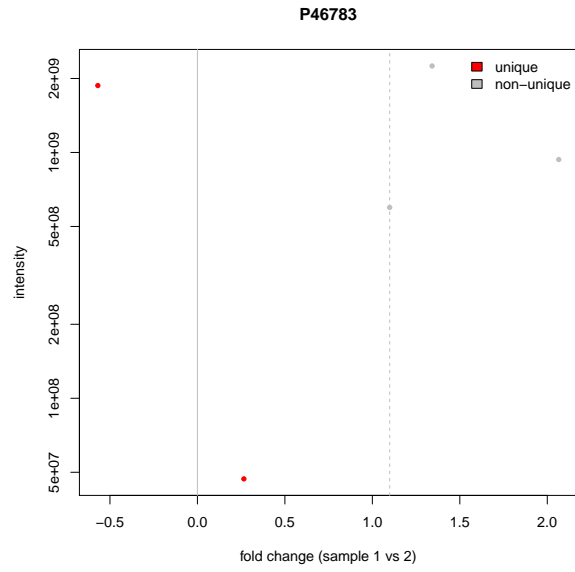


Figure 23: Fold changes of peptide abundances 1 and 2 for first protein of interest.

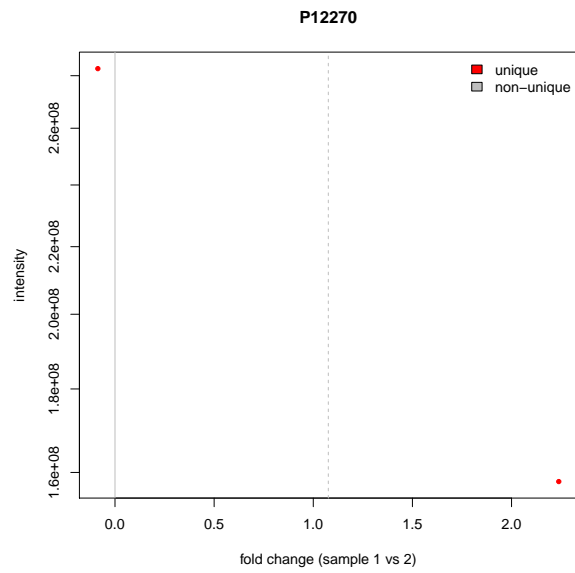


Figure 24: Fold changes of peptide abundances 1 and 2 for second protein of interest.