Set in includes 18 samples of pure and mixed blood, semen, and saliva. All samples were collected on the TripleTOF 6600 and searched in Protein Pilot with the Uniprot human database.

Processing:

Matches were filtered for confidence >= 95% (per Protein Pilot), and FDR < 1% estimated with reverse sequence decoy search.

Matches to keratin and trypsin were removed.

Peptide intensities were summed over multiple spectra, and log2 transformed.

Sample descriptions: type and ratio by protein weight of mixture. Numbers of peptides and proteins identified in each sample, after filtering.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| sample | | type | ratio | peptides | | proteins | |
| 20181127\_500ngBloodYoung1 | | blood | NA | 1155 | | 170 | |
| 20181206\_Blood\_WM24 | | blood | NA | 721 | | 125 | |
| 20181206\_Blood\_WM26 | | blood | NA | 1068 | | 171 | |
| 20181127\_500ngSA7 | | saliva | NA | 941 | | 184 | |
| 20181127\_500ngSA8 | | saliva | NA | 744 | | 170 | |
| 20181127\_500ngSA9 | | saliva | NA | 868 | | 156 | |
| 20181206\_SE1005 | | semen | NA | 1082 | | 230 | |
| 20181206\_SE1022 | | semen | NA | 1639 | | 332 | |
| 20181206\_SE1027 | | semen | NA | 1429 | | 322 | |
| 20181206\_B\_SA\_1.1 | | blood/saliva | 1/1 | 952 | | 134 | |
| 20181206\_B\_SA\_1.100 | | blood/saliva | 1/100 | 1111 | | 130 | |
| 20181206\_B\_SA\_100.1 | | blood/saliva | 100/1 | 1015 | | 156 | |
| 20181206\_B\_SE\_1.1 | | blood/semen | 1/1 | 1356 | | 230 | |
| 20181206\_B\_SE\_1.100 | | blood/semen | 1/100 | 1257 | | 248 | |
| 20181206\_B\_SE\_100.1 | | blood/semen | 100/1 | 1002 | | 155 | |
| 20181206\_SE\_SA\_1.1 | | semen/saliva | 1/1 | 1821 | | 315 | |
| 20181206\_SE\_SA\_1.100 | | semen/saliva | 1/100 | 748 | | 130 | |
| 20181206\_SE\_SA\_100.1 | | semen/saliva | 100/1 | 1214 | | 252 | |
|  |  | | | |  | |  | |

6035 total unique peptides were found across all 18 samples.

3844 peptides were found in at least two samples. Peptides that did not meet this criteria were removed.

Three data tables are included: non-normalized (**BF\_mixtures\_nonormalization\_matrix.csv**), quantile normalization (**BF\_mixtures\_quantilenorm\_matrix.csv**), and median centered data (**BF\_mixtures\_mediancenter\_matrix.csv**). I wasn’t sure which would be most appropriate since this set includes samples with very different protein composition.

All tables are log2 intensities for 3844 peptides and 18 samples, rows are peptides, columns samples.

Heatmap of normalized intensities of 3844 peptides in 18 body fluid samples, using median-centered data. Black = missing.

