Automated Validation of Results and Removal of Fragment Ion Interferences in Targeted Analysis of Data Independent Acquisition MS using SWATHProphet

Andrew Keller, Samuel L. Bader, David Shteynberg, Leroy Hood and Robert L. Moritz*

Institute for Systems Biology, 401 Terry Avenue North, Seattle, WA 98109, USA

Supplementary Information

Supplementary Information

Choice of Converter for Spectral Library Construction

We used the ProteoWizard 3.0.4624 converter with Vendor peak picking because it is open source and yielded similar numbers of library assays and assay peptides as the AB Sciex MS Data Converter (Fig. S1). We also investigated the use of a newer version of ProteoWizard (ver 3.0,7091) with peak picking either by Vendor or by the new (CantWaiT or CWT) wavelet-based peak-picker and precursor charge determination (Turbocharger) algorithm (1). Vendor peak picking with this version achieved the same result as with the earlier version (ver 3.0.4624), the CWT peak picking yielded lower numbers of assay peptides, and the default algorithm within ProteoWizard msConvert yielded the poorest result.

Control Data Set

A control data set was generated in order to evaluate SWATHProphet targeted analysis of DIA MS data. 1055 synthetic heavy C-terminal labeled *Mtb* peptides were spiked into a neat solvent and human urine background in triplicate, at dilutions of 1:1, 1:4, 1:16, 1:64, and 1:256. Data was collected on an AB Sciex TripleTOF 5600+ mass spectrometer using 32 precursor selection windows of 27.5625Th width and overlap 1Th, ranging from 355.19. Raw files were converted to mzXML using msConvert, and analyzed using SWATHProphet with a spectral library containing the *Mtb* spike-in peptides (1316 precursor assays), 1001 human urine peptides (1092 precursor assays), 2556 assays of *Mtb* peptides not present in the sample and not homologous to any human peptides ('target false positives'), and corresponding decoys (½ reversed, ½ randomized followed by a single amino acid substitution that preserves the precursor selection window). The library also contained predicted isotope peak intensities computed with the Isotope Pattern Calculator (PNNL, http://omics.pnl.gov/software/isotope-pattern-calculator). A time tolerance of ±7.5 minutes and *m/z* tolerance of ±0.05Da was used. Normalized retention times were computed based on a set of 27 abundant *Mtb* and human urine peptides.

Mtb peptides were detected in a 256-fold range in the neat solvent background samples, and in a 64-fold range in the human urine background samples (Table S1). Human

peptides were identified only in the urine background samples. True positives were inferred objectively as peak groups of 1:1 human urine background samples also identified in all three 1:1 and all three 1:4 neat solvent background replicate samples within 1 min normalized retention time and with an observed *ln* XIC ratio within ±1 of the expected values. On average, 663 true positives were inferred in each of the 3 1:1 urine background samples, with estimated 0% FDR based on their numbers of decoys and target false positives.

Evaluation of Decoys and Peak Group Scores

Thirteen peak group scores are computed for each result, and used to generate the discriminant score (linear combination) in each analyzed sample. They include parent scores based on the extracted trace of Q1 in the fragment ion scan, and mz delts and isocorr scores leveraging high mass resolution of the data:

1. Parent mzdelts: Precursor ion peak intensity weighted average of the deviation of predicted and maximum observed intensity *m/z* values, minus the average deviation observed for the run based on the retention time normalization peptides.
2. Parent isocorr: Precursor ion peak intensity weighted average of the correlation of the observed and predicted isotope peak intensities.
3. Delta irt: the absolute value of the difference between observed and library predicted normalized retention times.
4. Mz delts: Fragment ion peak intensity weighted average of the deviation of predicted and maximum observed intensity *m/z* values, minus the average deviation observed for the run based on the retention time normalization peptides. The score is averaged over all fragment ions in the peak group, weighted by peak area.
5. Parent peak shape: The correlation of the parent peak shape with those of the fragment ions. The score is the average correlation with all fragment ions in the peak group.
6. Parent co-elution: The average difference of parent and fragment ion peak retention times relative to the peak group width. The score is the average difference with respect to all fragment ions in the peak group.
7. Fraction contig: The fraction of total fragment b and y ion peaks that are identified contiguously at the peak group apex time.

8. Fraction nonlib y plus b: The fraction of total b and y fragment ion peaks not in the library assay that are identified at the peak group apex time.
9. Isocorr: Fragment ion peak intensity weighted average of the correlation between observed and predicted isotope peak intensities. The score is averaged over all fragment ions in the peak group, weighted by peak area.
10. Co-elution: The average difference of fragment ion peak retention times relative to the peak group width. The score is the average difference with respect to all fragment ions in the peak group.
11. Fraction missing assay ints: The fraction of library assay fragment ion total predicted intensity that is not detected.
12. Intensity correlation with library (Normalized int Euclid): the square root normalized Euclid intensity correlation score between observed and library predicted fragment peak intensities (2).
13. Peak shape: The correlation of the fragment ion peak shapes. The score is the average pairwise correlation of all fragment ions in the peak group.

Decoys are created to simulate false positives in the sample. We chose to use a diverse collection of 50% randomized and 50% reversed sequence decoys, each followed by a single amino acid substitution that changes the precursor *m/z* while preserving the precursor selection window. Decoys that by chance share all their assay fragment ions with a library target precursor ion in the same precursor selection window, within mass tolerance, are excluded.

It is desirable that decoys have peak group score distributions similar to false positives. This was tested using both the Wilks' Lambda test and area under ROC (receiver operating characteristic) curve for the control data set 1:1 human background dilution sample results (Fig. S2A). A high Wilks' Lambda and low area under ROC indicate similarity between the distributions among decoys and target false positives, suggesting the decoys are good representatives of false positives in the sample. Fig. S2B shows that the discriminant score distribution of decoys is very similar to that of target false positives, as are the distributions of randomized and reversed decoys. Using the Kolmogorov-Smirnov test (Fig. S2C), the discriminant score distributions of both the randomized and reversed decoys were

found not to differ significantly from that of the target false positives at the 90% confidence level.

The same criteria were used to evaluate the power of the peak group scores to discriminate between true positives and decoys (Fig. S3). In this case, a low Wilks' Lambda and high area under ROC indicate good discrimination between true positives and decoys. It is evident that the peak shape and intensity correlation with library scores are the most discriminating.

Analysis of SWATHProphet Results in the TPP

Conversion of SWATHProphet results to pepXML format enables their analysis by tools in the TPP in a manner similar to results of shotgun MS data (Fig. S4). The pepXML format, originally designed to store search results of shotgun MS data, was adapted to store scores and computed probabilities of SWATHProphet results. Whereas shotgun results include database peptides assigned to an MS/MS spectrum with a unique name, SWATHProphet results include a library assay precursor assigned to a peak group stored in pepXML as a unique spectrum name with start and stop scans corresponding to the peak group time boundaries. SWATHProphet probabilities are stored in PeptideProphet elements to enable subsequent analysis by iProphet and ProteinProphet. In the future, we plan to further adapt the pepXML format for SWATHProphet results with its own SWATHProphet element, and modify iProphet and ProteinProphet to make use of that format.

The use of pepXML and protXML are internal TPP working formats for representing data that also work well with the growing set of established set of tools that can read, process and export such data. However, the TPP also exports the newer community developed standard of mzIdentML using the idconvert tool from the ProteoWizard toolbox (3).

Calculation of Detected Interference Strengths

The strength of an interference is computed as the sum of contributions to the observed XIC of a peak group originating from the interfering precursor ion(s).

Intra-Library detected interferences

The Intra-Library approach to detect interferences finds pairs of co-eluting peak groups assigned to different library precursors that share one or more assay fragment ion *m/z* values, within mass tolerance. The shared peak area intensity of such fragments is confined to the overlap time range of the two peak groups. Only interferences originating from a peak group with high confidence of being correct (probability 0.9 or greater) are reported.

The strength of an Intra-Library detected interference of peak group A originating from peak group B is computed based on the shared and library predicted fragment ion intensities:

$$XIC^{A}_{obs} = TotalObservedXicPeakGroup_{A}$$

$$XIC^{A}_{true} = TrueXicPeakGroup_{A}$$

$$XIC^{A}_{unshared} = TotalObservedUnsharedFragmentIntensityPeakGroup_{A}$$

$$f^{A}_{unshared} = LibraryIntensityFractionOfUnsharedFragments$$

$$f^{A}_{i} = LibraryIntensityFractionOfSharedFragment_{i}$$

$$S_{i} = SharedIntensityFragment_{i}$$

$$FractionOfShared^{A}_{i} = \frac{\frac{f^{A}_{i} XIC^{A}_{unshared}}{f^{A}_{unshared}}}{\frac{f^{A}_{i} XIC^{A}_{unshared}}{f^{A}_{unshared}} + \frac{f^{B}_{i} XIC^{B}_{unshared}}{f^{B}_{unshared}}}$$

$$InterferenceToFragment^{A}_{i} = (1 - FractionOfShared^{A}_{i})$$

$$XIC^{A}_{obs} = XIC^{A}_{unshared} + \sum_{sharedFragments\_i} S_{i}$$

$$XIC^{A}_{true} = XIC^{A}_{unshared} + \sum_{sharedFragments\_i} FractionOfShared^{A}_{i} * S_{i}$$

$$InterferenceStrength_{A} = \frac{XIC^{A}_{obs} - XIC^{A}_{true}}{XIC^{A}_{obs}} = \frac{\sum_{sharedFragments\_i} InterferenceToFragment^{A}_{i} * S_{i}}{XIC^{A}_{obs}}$$

Score-Based detected interferences

The Score-Based approach to detect interferences relies on identifying valid peak groups with unexpectedly low intensity with correlation score. Its analysis is restricted to peak groups with a probability of 0.95 or greater (high confidence of being correct), as well as those with probabilities that could potentially be increased to that range after adjustment to the intensity correlation score upon *in silico* removal of the interference (see following section). Based on the probability and initial score value of a peak group, the magnitude of score increase required to adjust its probability to 0.95 or greater can be determined. If this magnitude is possible without exceeding the maximum score value of 1, the analysis will proceed. A Score-Based detected interference is only reported if its *in silico* removal does indeed result in a recomputed interference-free probability of 0.95 or greater.

The strength of a Score-Based detected interference of peak group A is computed based on its library predicted fragment ion intensities in a similar manner to that of an Intra-Library detected interference (see above), but with the following differences:

$$XIC_i^A = TotalObservedIntensityFragment_i$$

$$FractionOfShared_i^A = \frac{f_i^A XIC_{unshared}^A}{f_{unshared}^A} * \frac{1}{XIC_i^A}$$

$$InterferenceToFragment_i^A = (1 - FractionOfShared_i^A) * XIC_i^A$$

Adjustment to Peak Group Intensity Correlation Score and Probability upon *In Silico* Removal of Detected Interferences

The intensity correlation score of peak group A with a detected interference is recomputed after removing from the peak area of each affected fragment ion i, its contribution originating from the interfering precursor ion, $InterferenceToFragment_i^A$. The validity of adjustments to the intensity correlation score upon *in silico* removal of detected interferences was assessed by comparing the unadjusted and adjusted scores for the *Mtb* peak groups with interference from a human urine peptide with those of the corresponding peak group in the neat solvent background sample lacking the interference (Fig. S5). It was observed that the adjusted scores agreed more closely with those of the neat solvent background peak groups than did the unadjusted scores, for both Intra-Library (mean

square error 0.09 vs. 0.22 for strengths 0.2 or greater) and Score-Based detected interferences (mean square error 0.01 vs. 0.1 for strengths 0.2 or greater). Removal of interferences *in silico*, particularly Intra-Library detected interferences of peak groups with low initial probabilities, should therefore enable greater sensitivity of precursor detection by increasing their probability values.

Note that currently the intensity correlation score uses the intensities of fragment ion peaks, whereas the interference removal is made at the peak area level. The increase in peak area intensity correlation score following removal of the interference is then applied as if the same magnitude increase was observed at the peak intensity level, in order to translate the removal to a revised probability value. This is likely roughly the case, but could introduce error. In the future, use of the peak area for the primary intensity correlation score will be explored. That will require, however, that the spectral library include predicted fragment peak areas.

Since the intensity correlation score contributes to the linear combination discriminant score with known weight, the increase in correlation score upon *in silico* removal of an interference can be propagated to an increase in discriminant score. And since discriminant scores are converted to probabilities by the mixture model learned from the data set using decoys as known incorrect results, applying this model to the new discriminant score value yields an updated interference-free probability that the peak group is correctly assigned to its library precursor ion.

References

1. French W.R., Zimmerman L.J., Schilling B., Gibson B.W., Miller C.A., Townsend R.R., Sherrod S.D., Goodwin C.R., McLean J.A., Tabb D.L. (2014) Wavelet-Based Peak Detection and a New Charge Inference Procedure for MS/MS Implemented in ProteoWizard's msConvert. *J Proteome Res.* Dec 2. [Epub ahead of print].

2. Toprak U.H., Gillet L.C., Maiolica A., Navarro P., Leitner A., and Aebersold R. (2014) Conserved Peptide fragmentation as a benchmarking tool for mass spectrometers and a discriminating feature for targeted proteomics. *Mol.Cell.Proteomics.* Aug;**13**(8), 2056-2071.

3. Chambers M.C., Maclean B., Burke R., Amodei D., Ruderman D.L., Neumann S., Gatto L., Fischer B., Pratt B., Egertson J., Hoff K., Kessner D., Tasman N., Shulman N., Frewen B., Baker T.A., Brusniak M.Y., Paulse C., Creasy D., Flashner L., Kani K., Moulding C., Seymour S.L., Nuwaysir L.M., Lefebvre B., Kuhlmann F., Roark J., Rainer P., Detlev S., Hemenway T., Huhmer A., Langridge J., Connolly B., Chadick T., Holly K., Eckels J., Deutsch E.W., Moritz R.L., Katz J.E., Agus D.B., MacCoss M., Tabb D.L. and Mallick P. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nat.Biotechnol* **30**, 918-920

Supplementary Information Figures

A

Converter	*Peak Picker*	*File size [Mb]*	*Processing Time*
ProteoWizard 3.0.4624	Vendor	518	30 minutes
AB Sciex MS Data	ProteinPilot	978	30 minutes
ProteoWizard 3.0.7091	CWT	554	40 minutes
ProteoWizard 3.0.4624	msConvert	5311	35 hours

B

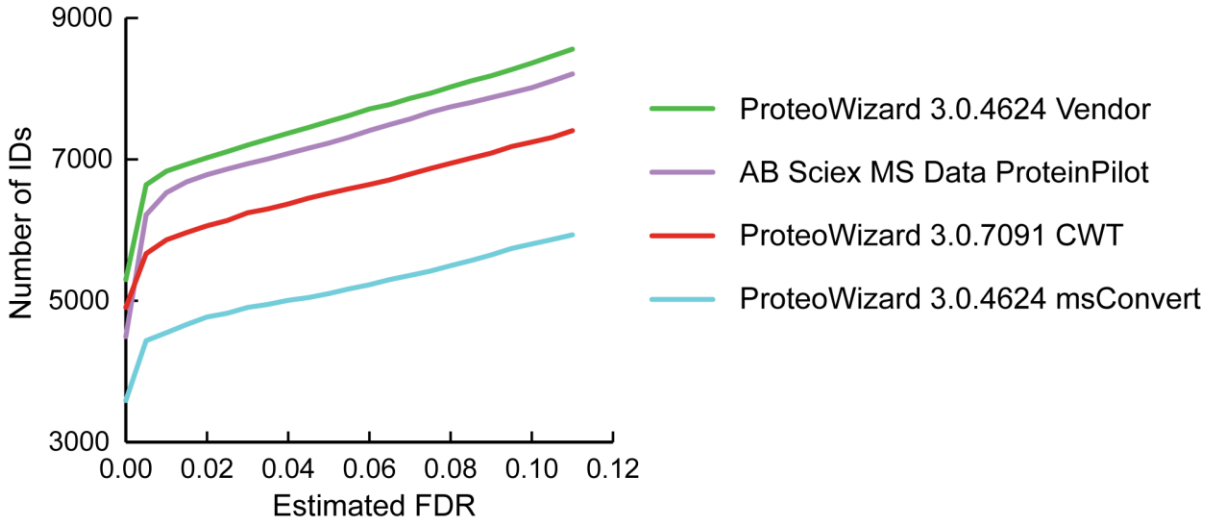
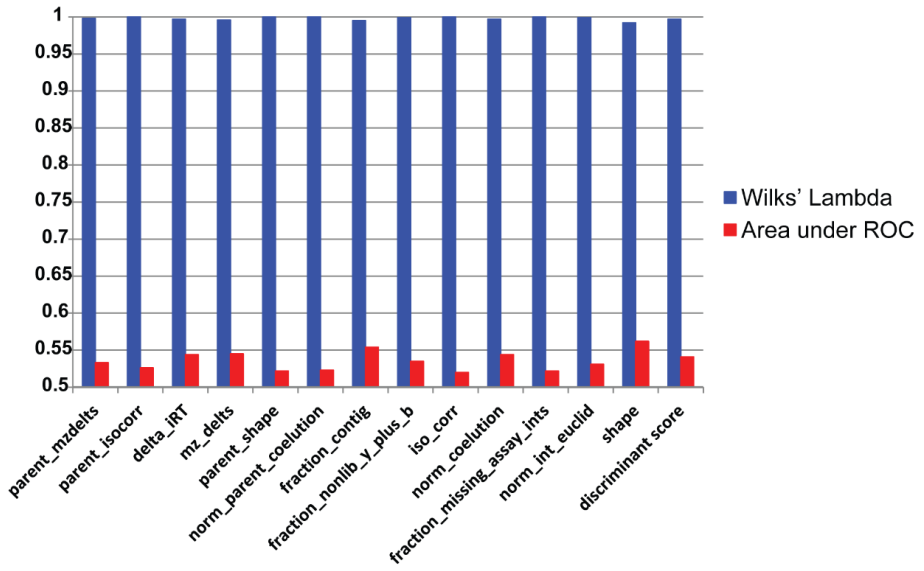
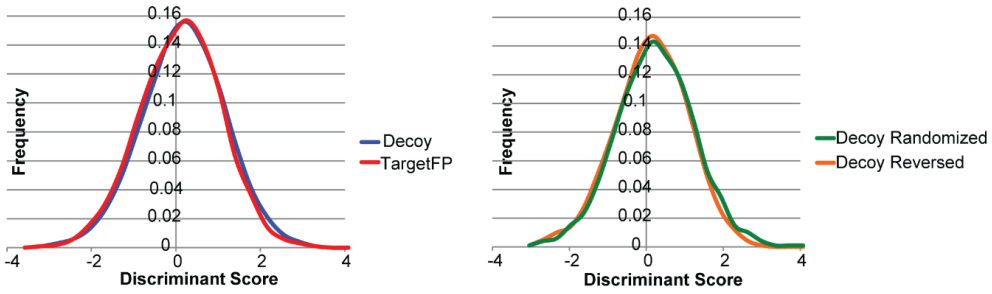


Figure S1. Comparison of ProteoWizard and AB Sciex MS Data converters. Output mzML file size and approximate required processing time (A) and numbers of identified peptides (B) in the human urine shotgun MS data set using indicated converters and peak picking settings. The ProteoWizard converter with Vendor specific peak picking, which yielded the greatest number of identifications, was used in this study.

A



B



C

Replicate	*Randomized*	*Reversed*	*50% Each*
1	0.048 vs 0.054	0.019 vs 0.054	0.041 vs 0.043
2	0.056 vs 0.054	0.063 vs 0.054	0.071 vs 0.043
3	0.041 vs 0.054	0.035 vs 0.054	0.044 vs 0.043

Figure S2. Similarity between decoy and target false positive score distributions. A, Wilks' Lambda and area under ROC for distributions of peak group scores and discriminant score among decoys versus target false positives in the control data set 1:1 dilution samples in the human urine background. B, Discriminant score distributions of decoys (½ reversed, ½ randomized followed by a single amino acid substitution that preserves the precursor selection window) versus target false positives, and of reversed vs. randomized decoys. C, Kolmogorov-Smirnov test of similarity between target false positive and decoy (randomized, reversed, or 50% each) discriminant score distributions in the control data 1:1 human urine background replicates. Show are the test statistic versus the minimum threshold for 90% confidence of a difference in the distributions.

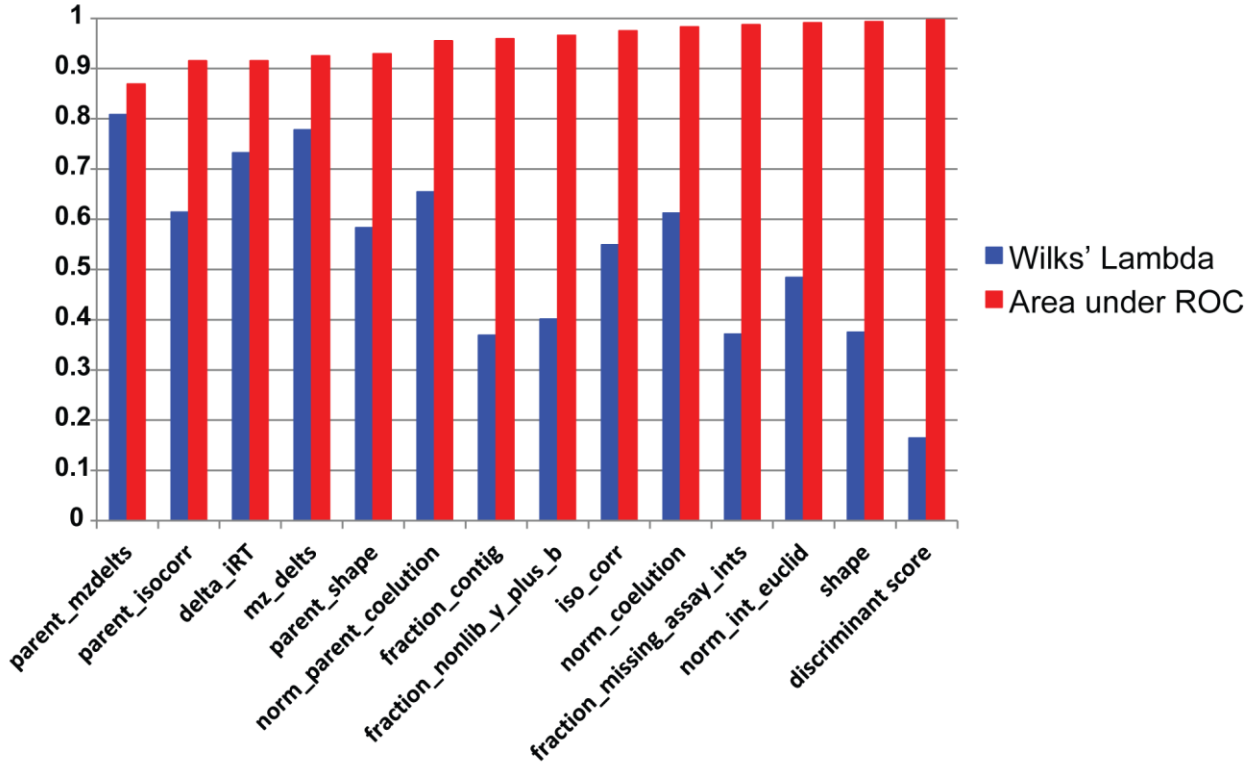


Figure S3. Discriminating power of peak group scores. Wilks' Lambda and area under ROC for distributions of peak group scores and discriminant score among true positives versus decoys in the control data set 1:1 dilution samples in the human urine background.

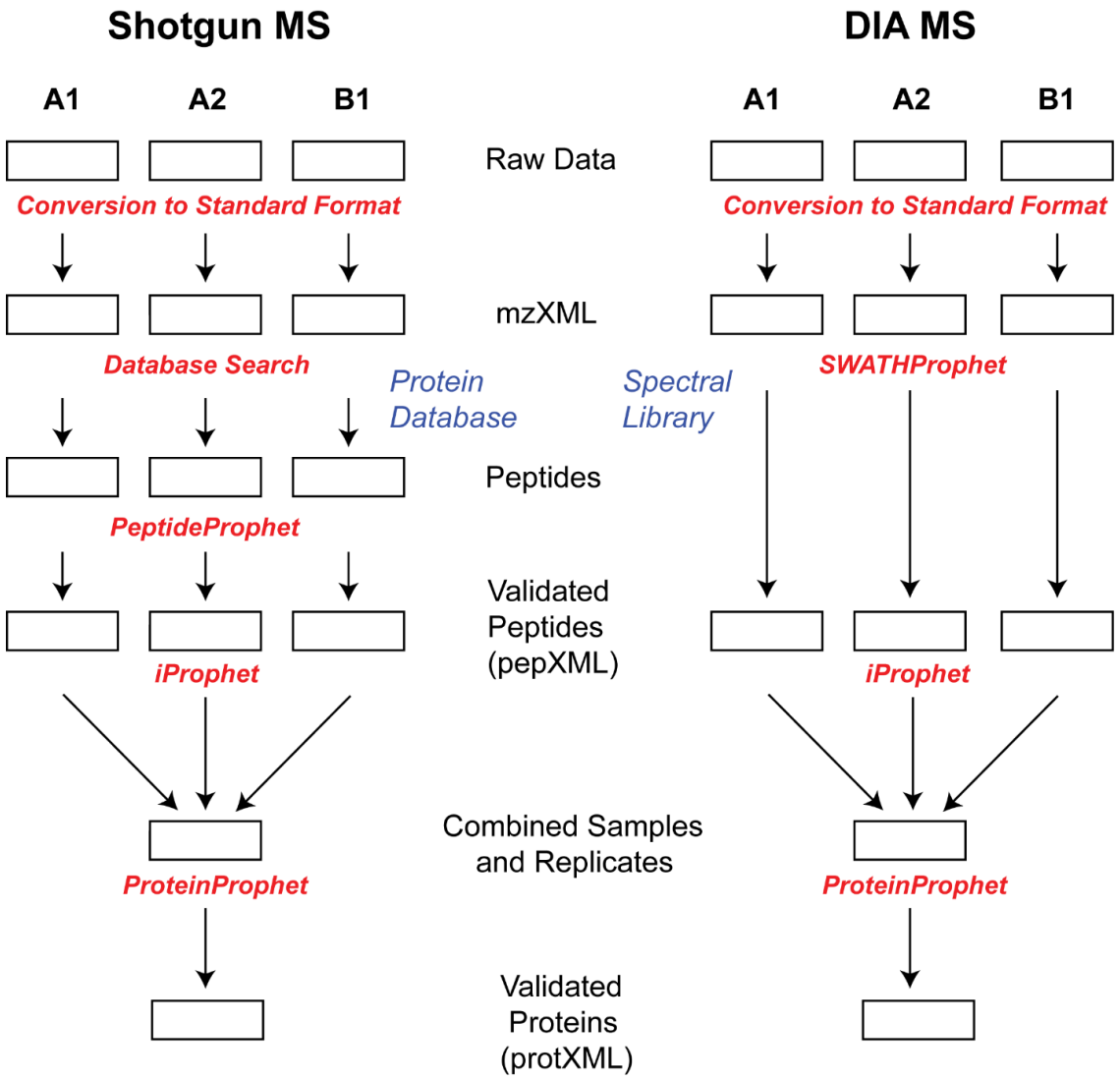


Figure S4. Comparison of shotgun and DIA MS data analysis workflows. Raw data for two replicates of sample A and one replicate of Sample B are converted to mzXML standard file format and analyzed to yield identified peptides with probabilities of being correct in pepXML format. Whereas in shotgun MS analysis, peptides are assigned to MS/MS spectra by database search and validated with probabilities of being correct by PeptideProphet, in DIA MS analysis, extracted peak groups are assigned to library precursor ions and validated with probabilities by SWATHProphet. Subsequent common analysis steps in the TPP include combining together results of multiple replicates and samples with adjusted probabilities by iProphet, and inference of sample proteins by ProteinProphet.

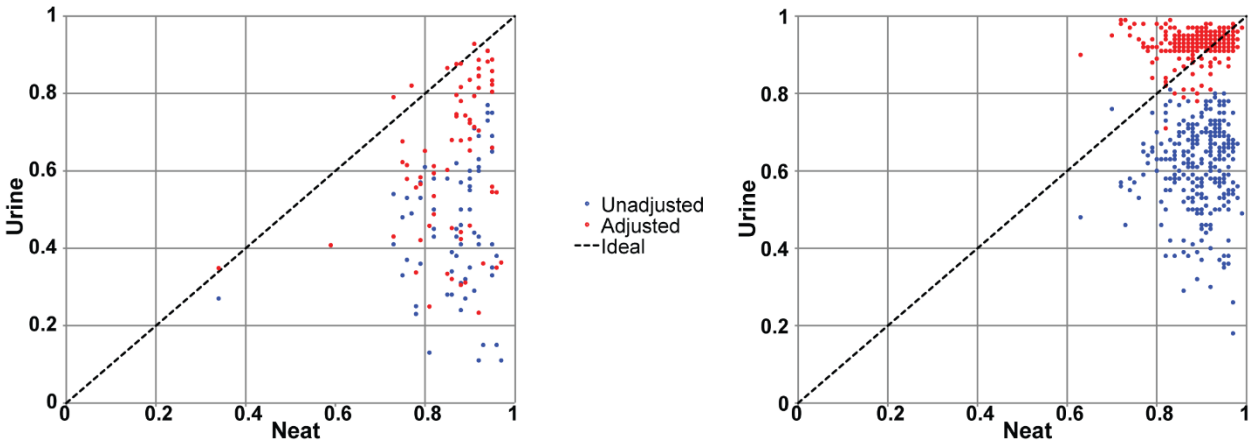


Figure S5. Improvement in intensity correlation score upon *in silico* removal of detected interferences in control data set. Intensity correlation scores of *Mtb* peak groups in human background samples were compared with those of corresponding peak groups in the neat solvent background samples, before and after *in silico* removal of the detected Intra-Library (A) or Score-Based (B) human peptide interference with strength 0.2 or greater.

	Mtb Precursor Ions			Human Urine Precursor Ions		
Sample	*SWATHProphet*	*iProphet*	*iProphet Percent Increase*	*SWATHProphet*	*iProphet*	*iProphet Percent Increase*
urine1:1	1132	1201	6	888	1020	15
urine1:4	984	1047	6	1079	1097	2
urine1:16	508	633	25	1062	1093	3
urine1:64	134	219	63	1046	1085	4
urine1:256	3	14	310	1006	1071	6
neat1:1	1163	1289	11			
neat1:4	1058	1163	10			
neat1:16	809	861	6			
neat1:64	633	659	4			
neat1:256	429	484	13			

Table S1. Increased numbers of precursor identifications after iProphet analysis of combined control data. The average numbers of *Mtb* and human urine precursor ions identified in the 3 replicates of each control data set dilution series sample at 0% estimated FDR based on either SWATHProphet or iProphet adjusted probabilities. The FDR was computed using the frequency of observed target false positive results. It is evident that iProphet is of particular benefit to the identification of *Mtb* precursors in the urine background samples in which they are present at lowest concentration.

Software and Tutorial

An in-depth tutorial and software code for SWATHProphet is included at the website http://tools.proteomecenter.org/software/SWATHProphet/

This website is updated for new versions of the code and deployment of SWATHProphet and modified iProphet for use with SWATHProphet results is included with each new release of the TPP.