Probabilistic Validation for Targeted Proteomics Using Parallel Reaction Monitoring

Alex W. Joyce[1-3], Ariana E. Shannon[3,4], Yemeng Wu[5], and Brian C. Searle[1-4]

1-Biomedical Graduate Studies Program, 2-Department of Biomedical Informatics, 3 – Pelotonia Institute for Immuno-Oncology, 4 –Chemistry and Biochemistry, 5- Department of Biomedical Education and Anatomy

THE OHIO STATE UNIVERSITY
WEXNER MEDICAL CENTER

Searle Lab

Searle

Introduction

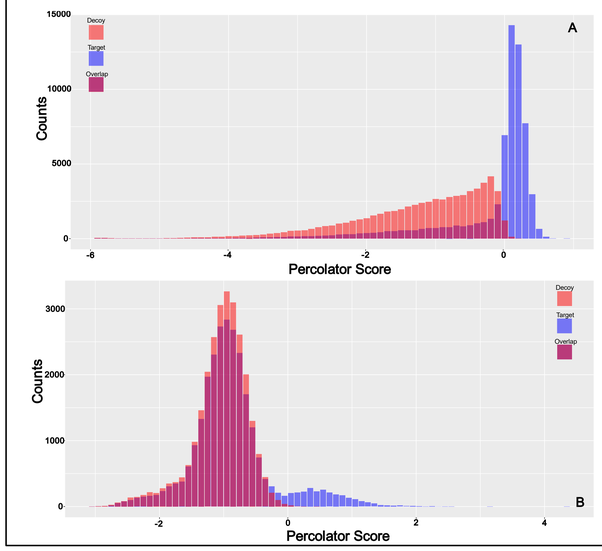
Due to the fact that targeted proteomic methods revolve around the identification of specific peptides, it is paramount to develop reliable ways to ensure confidence in the detection of targets. Unlike selected reaction monitoring (SRM) [1], parallel reaction monitoring (PRM)[2] can identify all peaks in a given spectra along with co-eluting peptides alongside the targets.

Here, we demonstrate that these features of PRM closely resemble those seen in top-down approaches such as data-independent acquisition (DIA)[4], allowing for the use of software tools used for top-down proteomics with targeted data. This is demonstrated by spectral library searching through the EncyclopeDIA platform[5] against a set of raw PRM data. The underlying framework of this process is the Percolator algorithm[6] which gives a set of scores as well as posterior error probabilities for each peptide-spectrum match (PSM).

Hypothesis: We can build probability distributions that resemble those used in DIA verification and structure them against targeted searches, giving a clear illustration of confidence in targeted methods.

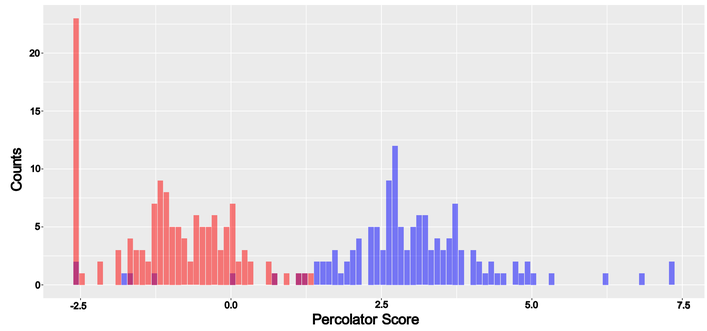
Global searches provide scoring frameworks for PRM experiments

Figure 1: Comparison of DIA and PRM score distributions. A). Distribution of scores from a DIA Chromatogram library searched against a DIA raw file. B). Distribution of scores from a DIA chromatogram library searched against a PRM raw file.



Local searches show discrimination between targets and decoys

Figure 2: Distribution of scores from a PRM targeted library searched against a PRM raw file using the Percolator model derived from the search shown in figure 1B.



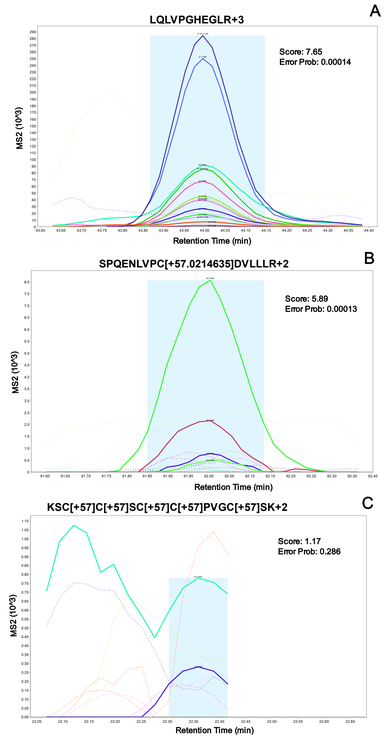
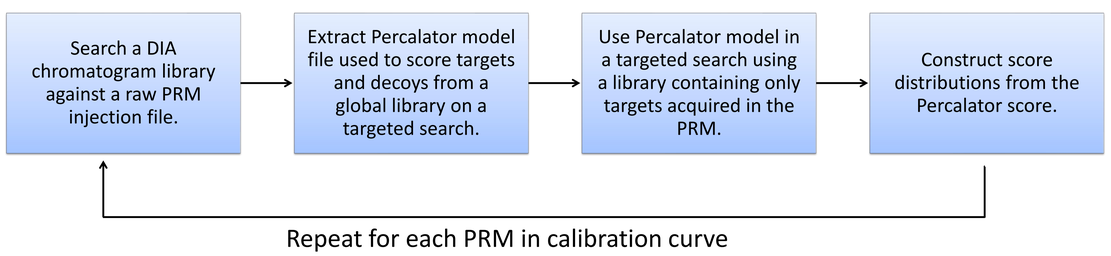


Figure 3 : Representative chromatograms were selected to show peptides and corresponding scores from the distribution featured in figure 2. A). example of a high-scoring chromatogram that can be clearly resolved. B). example of a medium-scoring chromatogram that can be resolved with difficulty. C). example of a low-scoring chromatogram that cannot be resolved.

Methods

A 'global' search was first conducted using a DIA library (.elib) against a PRM raw file (.mzml) using the mouse genome FASTA as a background file. The resultant Percolator model file produced from this search was used as a set of scoring parameters for a 'local' search where the DIA library was substituted for a PRM target library.



PEP correlates with LOQ

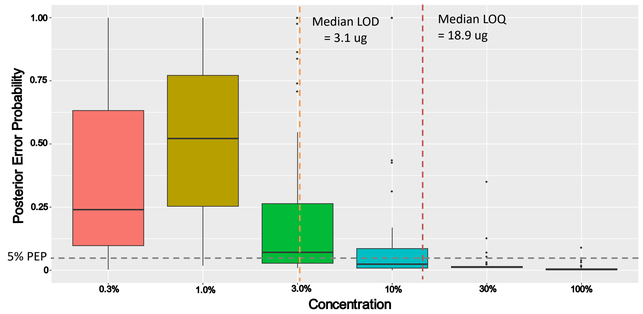


Figure 4: Boxplot displaying the PEP for peptides in each dilution. The majority of peptides above the LOD are below the 5% PEP.

Conclusion and Future Directions

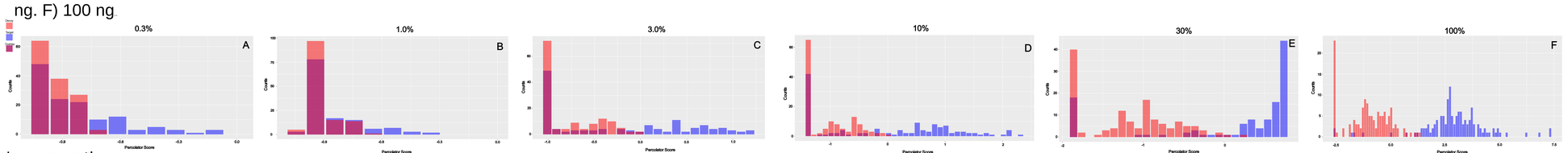
- PEP below 5% correlates with peptides above the limit of detection.
- PEP in conjunction with Percolator can be used to differentiate targets and decoys in assigning confidence to targeted proteomics matches.
- We have also demonstrated that sample concentration plays a significant role in both confidence and information content of the data.
- In the future, we would like to see if using background peptides present in small isolation windows can be used in addition to decoys to increase confidence in detecting targets.

References

1. Picotti, P., & Aebersold, R. (2012). Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nature methods, 9(6), 555–566.
2. Rauniyar N. Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. Int J Mol Sci. 2015 Dec 2;16(12):28566-81.
3. Krasny, L., & Huang, P. H. (2021). Data-independent acquisition mass spectrometry (DIA-MS) for proteomic applications in oncology. Molecular omics, 17(1), 29–42
4. Searle, B. C., Pino, L. K., Egertson, J. D., Ting, Y. S., Lawrence, R. T., MacLean, B. X., Villén, J., & MacCoss, M. J. (2018). Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. Nature communications, 9(1), 5128.
5. The M, MacCoss MJ, Noble WS, Käll L. Fast and Accurate Protein False Discovery Rates on Large-Scale Proteomics Data Sets with Percolator 3.0. J Am Soc Mass Spectrom. 2016 Nov;27(11):1719-1727.

Discrimination between target-decoy increases with concentration

Figure 5: The effect of protein concentration on targeted analysis. Using the same method discussed in figures 1 and 2, searches were performed against a set of raw PRM files obtained from experiments using titrating peptide amounts in a matrix-matched background. Dimethyl labelled peptides served as the matrix matched background in the calibration curve. A). 0.3 ng. B). 1 ng. C). 3 ng. D). 10 ng E). 30 ng. F) 100 ng.



Less separation between target-decoy ⟷ Less separation between target-decoy