Visual and Intuitive Access to Repository Data

Grant M Fujimoto, Michael Degan, Christopher O Overall, Lillian Ryadinskiy, Christopher Wilkins, Joon-Yong Lee, Richard D Smith, Samuel H Payne

Overview
- Release of the PNNL Biodiversity Plugin for Skyline which serves as a visual access point to PNNL’s mass spec data repository.
- Announce publication and deposition in open repositories of extensive proteomics data for 112 organisms including Human which has 30 separate tissues.
- Proteome coverage per organism averages 80% of functionally annotated genes in KEGG pathways.

Introduction
- A diverse set of organisms has been run by a variety of groups around the world, but they were not analyzed or processed in a standard way.
- Publically available data needs to be presented visually for different uses by various communities.
- Few tools exist for viewing data within the familiar context of biological processes and pathways.

Library Data
- Human proteomics data (Kim et al. Nature 2015)
  - 30 adult tissues
- Global proteomics data for 112 bacteria and archaea
  - Raw mass spectra (vendor and mzML formats)
  - Peptide identifications from MS-GF+ (mzid)
  - Condensed spectral library (bibliospec)
  - Proteins mapped to KEGG pathways
- On average, 2100 proteins per organism
- Data uploaded to ProteomeXchange via MassIVE
  - MSV000079053
  - PXD001860
- 13 TB of data

Library Visualization
- Identified peptides are linked to their highest scoring MS/MS scan with relevant peaks annotated.
- Enables higher quality SRM/MRM design.

Conclusion
- The PNNL Biodiversity Plugin is a useful tool for streamlining and simplifying the design of SRM experiments.
- Freely available at omics.pnl.gov/software or at the Skyline Tool Store.