PVIEW2: A Comprehensive Open Source Software System for Isotope Labeled and Label-Free Protein Quantification

Zia Khan, Leonid Kruglyak, and Mona Singh
zkhan@princeton.edu
Princeton University

ASMS 2010 Annual Conference Thursday Poster 008 10:30am - 2:30pm

web site: http://compbio.cs.princeton.edu/pview

ABSTRACT

The extracted ion chromatogram (XIC) forms the basis of protein and peptide quantification in complex LC-ESI-MS/MS data sets collected using modern, high resolution tandem mass spectrometers. XICs are peaks in MS1 spectra that occur in a narrow m/z range across several scans. The area under these XICs is correlated with relative peptide abundance. PVIEW2 is an easy to use software program for isotope labeled and label-free protein and peptide quantification. The software couples fast algorithms for finding and quantifying XICs, database search, protein grouping, and visualization all in one program. PVIEW2 is open source and freely available for academic use. PVIEW2 runs on MacOS, Windows, and Linux and works with high resolution QTOF, Orbitrap, and FTICR instruments. Source code and Windows 64-bit and Windows 32-bit binaries can be downloaded from http://compbio.cs.princeton.edu/pview.

METHODS

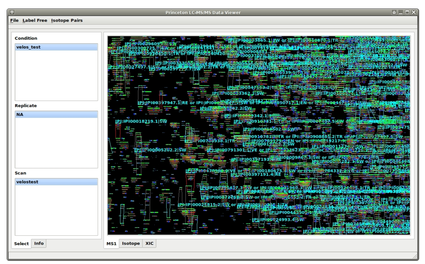
PVIEW2 uses algorithms for finding XICs from [Khan, PNAS, 106(37):15544-15548, 2009]. It uses a space partitioning data structure containing all centroided peaks to process MS1 spectra efficiently. In contrast to our previous work, PVIEW2 couples our algorithms with database search. PVIEW2 applies spectrum filtering, improves precursor mass accuracy using the intensity weighted m/z average over an XIC, includes an integrated database search score that relies on the fragmentation model from OMSSA [Geer et al, JPR, 3, 958-964, 2004], estimates statistical significance by a concatenated reverse decoy database, false discovery rate (FDR), and q-values [Kall JPR, 7, 29-34, 2008], and includes a novel algorithm for forming protein groups that handles shared peptides [Nesvizhskii, MCP, 4, 1419-1440, 2005].

SOFTWARE FEATURE SUMMARY

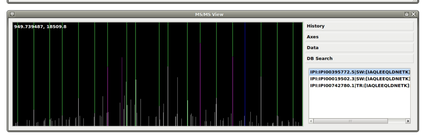
- **Isotope labeled heavy vs. light quantification (e.g. SILAC)**
- **15N heavy vs. light quantification**
- **Integrated MS/MS database search**
- **Integrated FDR and q-value estimation using a reverse concatenated decoy database.**
- **External search engines (e.g. Mascot, SEQUEST, X! Tandem) supported via PepXML format.**
- **Label free nonlinear retention time alignment based quantification.**
- **Label free XIC-based quantification based on MS/MS search results.**
- **Isotope Labeled Heavy vs. Medium vs. Light (e. g. pulsed SILAC) quantification.**
- **Label free retention time alignment based quantification.**
- **Integrated protein grouping algorithm to account for shared peptides between proteins and isoforms.**
- **MacOS, Linux, and Windows support.**
- **Support for high resolution QTOF, FTICR, and Orbitrap Instruments.**
- **Tight integration with the R statistical programming lange.**
- **Completely open source.**

CROSS PLATFORM GUI

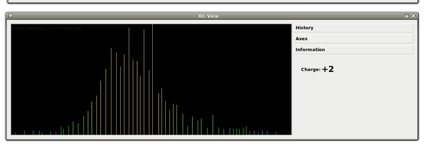
View data in 2-d
x-axis = retention time
y-axis = m/z



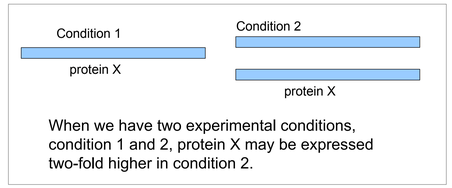
View MS/MS spectra with database search results.



View extracted ion chromatograms (XICs).

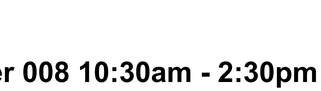


PRELMINARY DATA – Group 1 vs. Group 2 Ratio Plots

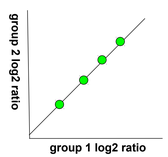


When we have two experimental conditions, condition 1 and 2, protein X may be expressed two-fold higher in condition 2.

A bottom up isotope labeled mass spec experiment, where condition 2 has been labeled with a heavy isotope, measures condition 2 vs. condition 1 ratios from unique (tryptic) fragments from the protein X. Assuming no post-translation modifications of the fragments, we expect the condition 2 vs. condition 1 ratio read from each of these fragments to be 2 to 1.



We can take each fragment ant put them into two groups, group 1 and group 2.



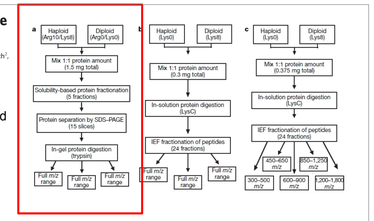
We expect the median ratio of each group to be perfectly correlated if (1) the ratio was measured correctly and (2) if the MS/MS data base search was conducted correctly.
(Thanks to Steve Gygi's lab for coming up with great method for evaluating quantitative mass spec data!!!!!

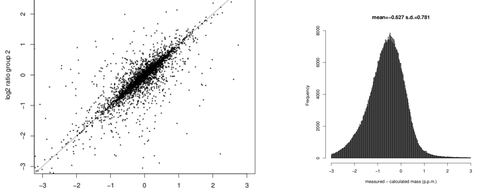
Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast

Lynn M. F. de Godoy[1]*, Jesper V. Olsen[1]*, Jürgen Cox[1]*, Michael L. Nielsen[1]*, Nina C. Hubner[2], Florian Fröhlich[2], Tobias C. Walther[2] & Matthias Mann[1]

Nature Vol 455 | 30 Oct 2008

We analyzed a large 225 LC-MS/MS run isotope labeled set using PVIEW. At 1% FDR PVIEW quantified 4,116 proteins vs. 3,569 quantified by Maxquant.





Furthermore, the precursor mass accuracy was in the p.p.b range.

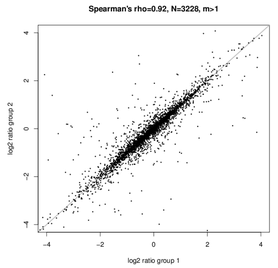
The Spearmans' rank correlation was 0.81 for this data set indicating good quantification and accurate MS/MS database search by PVIEW.

Super-SILAC mix for quantitative proteomics of human tumor tissue

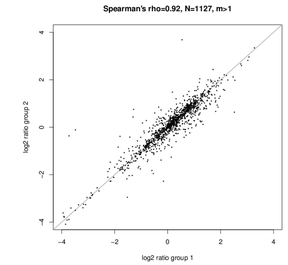
Tamar Geiger, Juergen Cox, Pawel Ostasiewicz, Jacek R Wisniewski & Matthias Mann

Nature Methods Vol 7| 4 Apr 2010

We also evaluated PVIEW on a recent data set that used a mix of heavy labeled cells as a reference for use in quantitative proteomics of human tumor tissue. Below we show the group 1 vs. group 2. fragment plots for an LTQ-Orbitrap and an LTQ-Orbitrap Velos data set bellow.



The Lobular sample vs. super-SILAC cell mix 3 replicates 6 fractions each (18 LTQ-Orbitrap instrument runs).



Unfractionated tumor sample vs. heavy labeled super-SILAC run in tripicate on an LTQ-Orbitrap Velos. Note the high Spearman's rank correlation between tryptic fragment ratio groups.