Visualize

The Multitool for Proteomics!

File

Open

Opens an .ez2 file to be examined.

Import from TPP

Imports data from files created by Trans Proteomic Pipeline. User chooses mzXML, pepXML and FASTA files and Visualize creates corresponding .ez2 file.

Save

Saves the file to a new file name. Option is available after a change has been made the alters the data in the file, such as removing proteins from the file.

Save selection

Saves a selection in the current file to a new file. The selection can be made manually by control or shift clicking, or the selection can be made using the search function. In this way, a new file with only probability scores above a cutoff (manual selection), a specific gene (search selection) or a specific species (search selection) can be saved to a separate file.

Exit

Experiment

Overview

Sub menu of functions to give the user overall information about the data in the file

Experimental Dashboard

Provides an overview of the number of proteins, peptides and scans are contained in the current data set. Data is presented as a series of histograms describing different parameters of the run(see below). Histograms shown vary with the data in the file and the search engine used.

Experimental Summary

Provides detailed data for all of the proteins found, including name, accession number, description, probability score, score (score = (maximum xcorr*peptide count*scan count)/10), peptide count, protein coverage, run count (for combined datasets), spectra or scan count, total Xcorr, maximum Xcorr, total TIC, molecular weight and pI values for the protein. Data is presented in a text window (see below).

Journal Report

Provides the information requested for the publication of a proteomics experiment by many proteomics journals.

Spectra Summary

Provides detailed data for all of the spectra assigned to a protein, including scan name, matched protein , matched peptide sequence, peptide probability score, discriminant score, delta Cn value, delta mass, Xcorr, charge, mass, TIC, ion match (matched: total) and any additional matched proteins. Data is presented in a text window (see below).

Protein Groups

Shows a list of groups of proteins that share identical peptides/scans. To the left of each protein, the number of scans currently assigned to a protein. The radiobox selection within each group indicates which protein will be used to represent the group if it is collapsed. The checkbox adjacent to each group indicates that the group should be collapsed. Collapsing groups removes the redundant proteins/scans leaving only the selected protein. Collapsing all groups ensures that each scan is assigned to only one protein.

Redundant Proteins

Describes the relationship between all proteins that share peptide in the sample, indicating which are twins, children or overlapping proteins.

Analysis

Sub menu of functions to analyze the data in the file as a whole.

Species/Gene Summary

Provides a list of the species and genes found in the dataset. This option is designed to be used with UniProt based databases with references in the format gene species. The command also generates Species and Gene tag clouds. Left clicking on the terms in the tag clouds opens the corresponding page from the UniProt database.

Show Serum Proteins

Identifies the members of the most abundant serum proteins in the file and calculates the percent of serum proteins present.

Show Selected Proteins

Allows the user to use a file containing a list of protein references or accession numbers to analyze the current set of proteins. Output is in the form of an Excel spreadsheet. The user can easily generate the file containing protein identifiers from the UniProt database, other online sources, or by hand from the literature.

Cleavage Site Analysis

Analyzes the terminal amino acids from each of the peptide matches found in the file. The count and percent of each amino acid is reported. For example, this function can be

used to determine the frequency of non-tryptic cleavages in a dataset. Data is presented in a text window (see below).

Amino Acid Analysis

Analyzes the amino acid composition, molecular weight, pI and GRAVY score (hydorphobicity) of each protein. At the end of the list, the overall, spectral count weighted composition of all proteins observed is also calculated. Histograms of the pI values and GRAVY scores observed in the data set are also presented. Data is presented in a text window (see below).

Modification Analysis

Analyzes the current set of data and identifies modified peptides.

Variant Peptide Analysis

Analyzes and displays variant peptides for searches done with a variant peptide database.

David

Uploads the list of protein accession numbers in the current dataset to the DAVID functional annotation tool hosted by NIAID/NIH (http://david.abcc.ncifcrf.gov). Currently works best with UniProt based databases. Opens external link in default browser.

Reactiome

Uploads the list of protein accession numbers in the current dataset to the DAVID functional annotation tool hosted by NIAID/NIH (http://david.abcc.ncifcrf.gov). Currently works best with UniProt based databases. Opens external link in default browser.

iHOP Tag cloud

Download interaction tags from iHOP and makes a tag cloud from the tags

KEGG Summary

This command opens two windows. One contains the summary of all of the proteins matching the KEGG pathway data for the current species, ranked from highest to lowest. Proteins with no matches to the KEGG pathway data are shown at the end. Data is presented in a text window (see below). The second window is a multi-pane window corresponding to the levels in the KEGG ontology. Clicking on one of the terms opens the child terms in the window to the left. For second and third level terms the term is followed by the number of protein matches in brackets. Below the third level term is listed the proteins that matched the term. Right clicking on a third level term with protein matches exports the list of matching proteins to the KEGG website and opens an external webpage in IE displaying the KEGG pathway for the term with the proteins matched highlighted in red. In order for the KEGG summary command to function, the KEGG map file, rodent.kegg for example, must be present in the same directory as the Visualize program.

GO Protein List

Quantitates the GO terms by protein count. Requires the file uniprot_go.dbm be present in the same directory as the Visualize program.

GO Scan List

Quantitates the GO terms by scan count. Requires the file uniprot_go.dbm be present in the same directory as the Visualize program.

Display

Sub menu of functions that display data about the data in the file as a whole

Display Peptide Probability Model

Displays a graph of the peptide probability model used in analyzing the data. Discriminant score is plotted vs. peptide probability. Graphs can be saved in PostScript (.ps) form. These files can be used as .eps files publications or presentations or for better results, converted to .pdf, .eps, .tiff or other formats using other programs such as Adobe Acrobat or GhostScript.

Display Discriminant Score Histogram

Displays a graph of the histogram of discriminant scores for all top level hits for all spectra. Discriminant score is plotted vs. peptide probability. Datasets include total, predicted correct and random hits.

Display Probability vs. False Discovery Rate (FDR) Graph

Displays a graph of. the peptide probability vs. the peptide FDR for the peptide probability model used in analyzing the data. Peptide probability is plotted vs. peptide FDR.

Display FDR from decoy run

Displays the peptide probability vs. peptide FDR for current run based on DECOY- tags on proteins.

Display Chromatogram

Displays the text data and graph for the chromatographic profile of the current run. Graph can be resized by selection section of the graph or by changing the values for the axes from the menu.

Display Scanmap

Displays the mass vs time for each parent ion. Mousing over the graphic displays the protein identification results for each scan. Right clicking on a point opens the spectral display for that scan. Left clicking on a protein name highlights all scans associated with that protein in blue. Right clicking on a protein name opens the peptide graph for that protein. In the peptide graph, mousing over a peak in the graph highlights the peptide sequence in red.

Calculated vs Observed RT

Displays the mass vs time for each parent ion. Mousing over the graphic displays the protein identification results for each scan. Right clicking on a point opens the spectral display for that scan. Left clicking on a protein name highlights all scans associated with that protein in blue. Right clicking on a protein name opens the peptide graph for that protein. In the peptide graph, mousing over a peak in the graph highlights the peptide sequence in red.

1D gel Simulation

Displays the mass vs time for each parent ion. Mousing over the graphic displays the protein identification results for each scan. Right clicking on a point opens the spectral display for that scan. Left clicking on a protein name highlights all scans associated with that protein in blue. Right clicking on a protein name opens the peptide graph for that protein. In the peptide graph, mousing over a peak in the graph highlights the peptide sequence in red.

2D Gel Simulation

Displays the mass vs time for each parent ion. Mousing over the graphic displays the protein identification results for each scan. Right clicking on a point opens the spectral display for that scan. Left clicking on a protein name highlights all scans associated with that protein in blue. Right clicking on a protein name opens the peptide graph for that protein. In the peptide graph, mousing over a peak in the graph highlights the peptide sequence in red.

Extract Results

Displays the mass vs time for each parent ion. Mousing over the graphic displays the protein identification results for each scan. Right clicking on a point opens the spectral display for that scan. Left clicking on a protein name highlights all scans associated with that protein in blue. Right clicking on a protein name opens the peptide graph for that protein. In the peptide graph, mousing over a peak in the graph highlights the peptide sequence in red.

Maintainance

Sub menu of functions to allow the user to examine the contents of the file and add data to the file

Log files

Displays the contents of any Sequest or other .log files in the file. Data is presented in a text window (see below).

Header files

Displays the contents of the Header.txt file in the file. Data is presented in a text window (see below).

Parameter files

If the file is an uncombined file, the contents of the .params file is displayed. If the file is a combined file, the contents of the combined parameter hash is displayed. Displays the contents of the Header.txt file in the file. Data is presented in a text window (see below).

File contents

Lists the directory of the file in a text window (see below).

Update from RAW

Updates header and spectra information from RAW file. Works only on Thermo RAW files and requires both Windows and BioWorks to be installed.

Update from mzXML file

Updates header and spectra information from a mzXML file. .

Update from indexed database

Updates protein sequence and descriptions from indexed database file created by Epitomize.

Update from FASTA file

Updates protein sequence and descriptions from a FASTA format file.

Update from GOA file

Updates GO annotations for proteins from a .goa file.

Add UniProt data

Adds data from UniProt flat files

Update Run Counts

For combined files, updates the run counts for each protein.

Encrypt Data

Encrypts the protein, scan, fasta, dta and out data in the current file using User entered key and Blowfish algorithm. Overwrites data in current file and if the key is lost, there is no way to recover the data. .

Import Files

Imports external files into the .ez2 file. This allows the addition of protocols and other notes, as well as image file (ie. a 2D gel image). Text and image files can be viewed using File contents.

Import additional data

Allows for the import of additional data for the proteins. A dialog box allows the selection of a tab separated values (.tsv) text file in the following format: Column 1 - ,

reference for protein (for UniProt this is the gene_species name), Column 2 and beyond – additional data to be attached to the protein referenced in column 1. The first line of the file should be a header line containing the title of the additional data columns. This file can be generated in Excel and saved as a .tsv file.

Set Preferred Species

Opens a dialog box allowing the user to set the preferred species if this was not done by Epitomize. Note that this does not reorder the matched proteins as is done by Epitomize so it is preferable to set the preferred species when creating the .ez2 file.

Experimental Notes

Opens or adds a time and date stamped notes file inside the .ez2 file.

Protein

The Protein menu is only available after a file has been opened and a protein has been selected. Note that not all databases contain information for all species or proteins, which will sometimes lead to errors on the linked websites.

Protein coverage

Displays the coverage information for the selected protein. In this view, the full sequence is displayed and individual residues are color coded based on the number of times that they were observed. Data is presented in a text window (see below).

Protein Links

Sub menu of functions that allow the user to examine external information for individual proteins

Protein information from Apropos

Exports the accession number of the currently selected protein to the Apropos website (apropos.mcw.edu) and opens the individual IPI record for the protein in an external browser window.

UniProt Protein information

Exports the accession number of the currently selected protein to the UniProt website (www.pir.uniprot.org) and opens the individual UniProt record for the protein in an external browser window.

IPI Protein information

Exports the accession number of the currently selected protein to the EBI website (srs.ebi.ac.uk) and opens the individual IPI record for the protein in an external browser window.

iHOP Protein information

Exports the accession number of the currently selected protein to the iHOP website (www.ihop-net.org) and opens the webpage showing literature references for the protein in an external browser window.

ArrayExpress information

Exports the accession number of the currently selected protein to the EBI ArrayExpress website (www.ebi.ac.uk/microarray-as) and shows expression information for the protein in an external browser window.

GeneCard information

Exports the gene name of the currently selected protein to the GeneCard website (www.genecards.or) and opens the individual gene record for the protein in an external browser window.

String Protein information

Displays protein – protein interactions using the STRING tool. .

BioGPS Protein information

Displays protein information using the BioGPS tool from GNF.

Reactome Protein information

Exports the accession number of the currently selected protein to the UniProt website (www.pir.uniprot.org) and opens the individual UniProt record for the protein in an external browser window.

Secreted Protein information [Not working]

Exports the protein reference to the Secreted Protein Database and opens the record for the protein in an external browser window.

GenPept Graph

Displays the gene encoding the protein in the NCBI GenPept view.

SNP for gene encoding protein

Displays SNPs for gene encoding protein from NCBI database.

Search Google Scholar

Searches for literature references to the protein using Google Scholar..

KEGG information

Information about the currently selected protein is downloaded from the KEGG database and displayed in a text window (see below).

Peptide

Blast Peptide Sequence

Performs a BLAST search for the peptide sequence against the NCBI nr protein database using parameters optimized for short peptide sequences.

Average Ion Match

Calculates an average spectrum from all of the spectra assigned to a peptide and compares it to the theoretical spectra for the peptide sequence. This leads to less noise and background peaks.

Average Spectra

Calculates an average spectrum from all of the spectra assigned to a peptide displays it along with the theoretical spectra for the peptide sequence. Matched peaks are indicated in black, non-matched peaks are grey.

Export Average Spectra

Calculates an average spectrum from all of the spectra assigned to a peptide displays it along with the theoretical spectra, ion match table and match statistics and creates and saves a one page pdf document suitable for use in publications or presentations.

Export Average Spectra dta file

Calculates an average spectrum from all of the spectra assigned to a peptide and exports it as a .dta file.

Scan

The Scan menu is only available after a file has been opened and an individual scan has been selected.

.out file

Displays the search results data from the original .out file in a text window (see below). This is not supported always supported for non-Sequest results.

.dta file

Displays the centroided spectral data from the original .dta file in numerical form in a text window (see below). This is not supported always supported for non-Sequest results.

Ion Match

Displays the table of experimental ions that match the theoretical positions of the b and y ions deduced from the matched peptide sequence.

Spectra

Displays the spectral graph with the theoretical positions of the b and y ions deduced from the matched peptide sequence indicated.

Spectra with different peptide

Allows user to see the current spectra displayed with a different peptide sequence. This is valuable to examine the effects of different sequence variations or modifications.

Export Current Spectra

Displays the experimental spectra along with the theoretical spectra, ion match table and match statistics and creates and saves a one page pdf document suitable for use in publications or presentations.

Export

Extract Protein IDs

Exports the UniProt protein IDs to a text file.

Extract Accession Numbers

Exports the accession numbers of the proteins in the current dataset to a .fasta format file. This text file can be used to upload data to external annotation tools such as Apropos.

Extract Fasta sequences

Exports the sequences of the proteins in the current dataset to a .fasta format file. This fasta file can be used to create new databases for serial searching.

Excel file

Exports the protein and scan information to an Excel format file. Added data is also exported as well as worksheet with amino acid composition information for the identified proteins. If the file was a DECOY search, then FDR is calculated.

Map IDs

Uses the UniProt mapping service to convert the Protein IDs or Accession numbers for the proteins in the file into a wide range of different identifiers and opens the new list in a text window that can be copied, saved or printed.

Excel file

Exports the filtered data into an Excel file. Links are included to the UniProt database. This is a useful file to share with collaborators that do not use Visualize.

Peptidome Excel file

Exports the filtered data into an Excel file formatted for submission to the NCBI Peptidome Repository. (http://www.ncbi.nlm.nih.gov/peptidome)

Amino Acid Excel file

Performs amino acid analysis on the proteins in the file and produces an Excel workbook file with the data as a series of worksheets. Data includes composition, pI and GRAVY

values for all proteins as well as weighted composite composition, pI and GRAVY values for the entire set of proteins. Weighting is based on scan count for each protein.

Amino Acid Excel File

Exports the results of the amino acid analysis for a single file as an Excel workbook

TreeMap

Annotates the protein data with GO ontology terms and formats as a .tm3 file for display with the U. Maryland TreeMap Java application.

Export Spectra

Allows the user to select and export annotated spectra to a single pdf file. This can be used to comply with journal requirements for spectra of single and modified peptide hits

Search

Search function produces a list of all protein that match a criteria (all text, reference, accession, or description). The list can be used to go directly to each of the hits or to select the set of hits and save as a new file. For other searches (peptide sequence, .out file text, or file name) a text window with the search results is produced.

All text

Search for a word, phrase or regular expression (REGEX) in all name, accession or description of all proteins.

Reference

Search for a word, phrase or regular expression (REGEX) in protein names. Can be used to select protein from a given species (_RAT) or corresponding to a specific gene (ACT).

Accession Number

Search for a word, phrase or regular expression (REGEX) in all accession numbers.

Description

Search for a word, phrase or regular expression (REGEX) in protein descriptions.

Peptide Sequence

Search for a sequence or modification in peptide sequences. Use of REGEX searches are particularly useful in this context.

.out file text

Search for a word, phrase or regular expression (REGEX) in all text of all .out files. Caution: This search can be time consuming and should be used only if one of the specific searches above is insufficient.

Scan name

Search for and jump to specific scan name or number.

File name

Search for parts of .dta/.out file names. Can be used to locate data from particular scan or files in combined data.

Remove serum proteins

Searches for members of the most abundant serum proteins group and allows the user to remove them from the file.

Keep only serum proteins

Removes all proteins that do not match the serum proteins group.

Remove common contaminants

Selects proteins flagged in the database by CONTAM_ and allows the user to remove all or some of them.

Remove selected proteins

Allows the user to specify up to six search terms to select proteins or a file of proteins to be removed from the file. The reference, accession, and description fields can be searched using a normal or regex search. Can be combined with the Export Protein IDs option ro subtract proteins found in one file from a second file.

Keep only selected proteins

Similar to Remove selected proteins except it only keeps proteins that match the search.

Subtract .ez2 file

Remove the proteins found in a different .ez2 file from the current file. This command is useful for removing non-specific background proteins from affinity purifications.

Jump to scan

Change current scan to scan number entered. Updates current protein and peptide to reflect new scan.

Combine

Experiment - Complete

Combines two or more .ez2 files originating from independent runs. Protein probability scores are combined such that proteins observed in multiple files are considered to be independent events and the score is improved ($P = 1- (1-P1)*(1-P2)*\ldots(1-Pn)$). Transfers all data including .dta and .out files.

Experiment – Express

Similar to complete but does not transfer the .dta and .out files, increasing speed and decreasing file size. This option is useful for creating files that are to be used for comparisons at the level of groups of data rather than examining data at the individual spectra level.

Search

Combines two or more .ez2 files originating from repeated searches of the same original raw data. For each spectra, the highest scoring spectra between the runs is used. Protein probabilities are recalculated based on the collection of highest scoring spectra.

Aggregate Results

Allows for the combination of results from multiple files without calculation of probability values. A user can aggregate the data from both .ez2 and older .ezf files.
Closes open file and exits the program.
The Experiment menu is only available after a file has been opened. The menu is divided into 5 sections

Quantitate

Spectral Counting

Sub menu of functions to compare files based on spectral counting approaches

Control values

The user selects one .ez2 file to represent the control data and then a group of .ez2 files to compare it to. Comparisons are made to all proteins that are present in the control sample.

Max values

Using a group of user selected .ez2 files, a synthetic standard is generated based on the maximum values for observed for all observed proteins. In this case, proteins do not have to be observed in all samples.

Total values

Using a group of user selected .ez2 files, a synthetic standard is generated based on the total values for observed for all observed proteins. In this case, proteins do not have to be observed in all samples.

Sample groups

Multiple samples in multiple groups are selected. The mean and standard deviation for scan counts for each protein is determined for each sample in a set and the sets compared.

Multigroup lists

For comparison between groups of files that represent for example biological samples. The user constructs a text file in which the name of the group is given on a line beginning with the $ character, the path to each of the files belonging to the group follows with one file path per line. More than two groups can be compared, but for two groups comparisons, ratios and p values will be calculated. The output is in the form of a multi-sheet Excel workbook. For series set of experiments, including a value preceded by a tab will allow the program to do a linear regression of the values and determine the patter of protein abundance.

GO Terms

Compare two samples based on GO terms by spectral counting.

Selected Proteins by Group

Applies the same analysis as selected proteins across multiple lists of proteins and multiple sample searches. Generates values for mean and standard deviation for individual proteins and whole groups of proteins. Output is placed into a multi-sheet Excel workbook.

Selected proteins by group

Using a preselected set of proteins, different sample groups can be analyzed for the abundance of proteins in the set. For example, the membrane proteins in two biological samples can be compared.

Selected proteins by list

Compare proteins between samples but only consider proteins on list. For example, can be used to restrict comparison to membrane or mitochondrial proteins.

Compare file of groups

Combine the use of lists of samples and proteins.

Matched Pairs

Compares groups of data from matched pairs of samples.

SPeCtRA

Sub menu of functions to compare files based on SPeCtRA approach that combines SILAC with spectral counting

Compare SILAC samples

Using a group of user selected .ez2 files, a synthetic standard is generated based on the maximum values for observed for all observed proteins. Peptides that do not include the labeled amino acid are excluded from the comparison. (see Set SILAC label below).

Compare SILAC groups

Allows the user to define groups of rums to be compared using the SPeCtRA method

Compare SILAC groups by average

Similar to compare sample groups but using the SPeCtRA method.

Set SILAC label

Sets the amino acid that peptides must contain to be included in the SILAC ratio calculations

Compare SILAC Peptides

Compares SILAC (heavy and light) searches of the same data files at the protein and peptide level. Opens a new window showing the proteins with the H:L ratio indicated, individual peptides with the H: L ratio, and individual scan results for both searches.

MS/MS mass tags

Quantitate sample containing sample dependent isobaric MS/MS mass tags (iTRAQ and TMT methods). Masses are predefined for iTRAQ but can be up to 8 different tag masses.

Compare Amino Acid Profiles

User selects an Excel workbook for output and then two .ez2 files for comparison. The physiochemical properties of the amino acids and proteins in the sample are saved as multiple spreadsheets in the workbook.

Compare Spectra

Compare two search results from the same set of spectra on a spectrum by spectrum basis. Includes options to incorporate a set of known correct proteins and a tag for decoy or other special sequences. This option is useful for comparing search algorithms or parameter sets.

Create Comparison Gel

Allows the user to create a simulated 1D SDD-PAGE image with up to 10 lanes. Each lane can contain data from an .ez2 file, a standard set of markers or be left blank. The mass range of the gel is variable and normalization can be applied for each individual lane or across the entire gel.

Set Comparison parameters

Sub menu to set values for comparisons

Get protein list for comparison

The user selects a list of protein names (gene_species for UniProt). Future comparisons will be limited to this list of proteins. This allows the user to focus on a specific group of proteins, e.g. Only mitochondrial proteins.

Set normalization protein

For compare to max and compare to total, values are corrected with respect to the abundance of a ‘normalization’ protein, that can either be an exogenously added protein standard or a protein that is expected to remain biologically consistent between samples.

Help

About

Displays the splash screen.

Help
Opens the online help.

Check version
Checks if the version of the program is current. Requires internet access.

License
Displays the license text.

Changelog
Displays the changelog detailing changes made between program versions.

Text Window Menus
Text output window have menus of additional functions:

File

Save
Save the contents of the window as a text file.

Print
Select print options and print the contents of the window. (Windows only).

Close
Closes the window.

Edit

Copy
Copies the contents of the window to the clipboard (Windows only).

Find
Opens Find dialog box. Find highlights and moves to the next occurrence of the search string. Find all highlights all instances of the string. Direction and case sensitivity can be set and regular expressions can be used in the search.

Graph Window Menus
Graphic output windows have menus of additional functions:

File

Save
Saves the graphic in one of several different graphics formats.

Close
Closes the window.

Compare Menus
Comparison commands open a window with menus of additional functions.

File

Save as text file
Saves the results of the comparison as a text file.

Save as TDMS file
Saves the results of the comparison as a .tdms file compatible with microarray tools such as the TMD suite from TIGR.

Save as Excel File
Saves the results of the comparison as an Excel file.

Save as Ratios
Saves ratios and statistics for comparison as an Excel file.

Save as Ingenuity B file
Saves a file formatted for use with the Ingenuity Pathway Analysis tools.

Save as TreeMap file
Annotates proteins with GO information and formats results for visualization in the U. Maryland TreeMap program.

Edit

Copy
Copy text version of results to clipboard (Windows only).

Find
Provides a find function within the text of the window.

View

TIC
Shows ratios based on total TIC values.

Scans
Shows ratios based on scan count values.

Peptides
Shows ratios based on unique peptide count values.

Coverage
Shows ratios based on protein coverage values.

Venn Digram
Draws a Venn Diagram for up to 3 samples. Right clicking on values displays the list of proteins in each group.

Scatter plot
Creates a scatter plot for pairs of samples. Mousing over the points identifies the protein and displays the values. Right clicking labels the point with the protein name. Labels can be moved by dragging or removed by right clicking. Using the scale menu the user can zoom in or zoom out. The plot can be saved as a postscript file that can easily be converted to pdf or imported into drawing or presentation programs.

GO
Shows ratios of GO term annotations based on protein abundance.

Sort

Alphabetic
Sorts proteins alphabetically

Intensity
Sorts proteins by total TIC.

Probability
Sorts proteins by protein probability.

Peptide Count
Sorts proteins number of unique peptides.

Scan Count
Sorts proteins by number of scans.