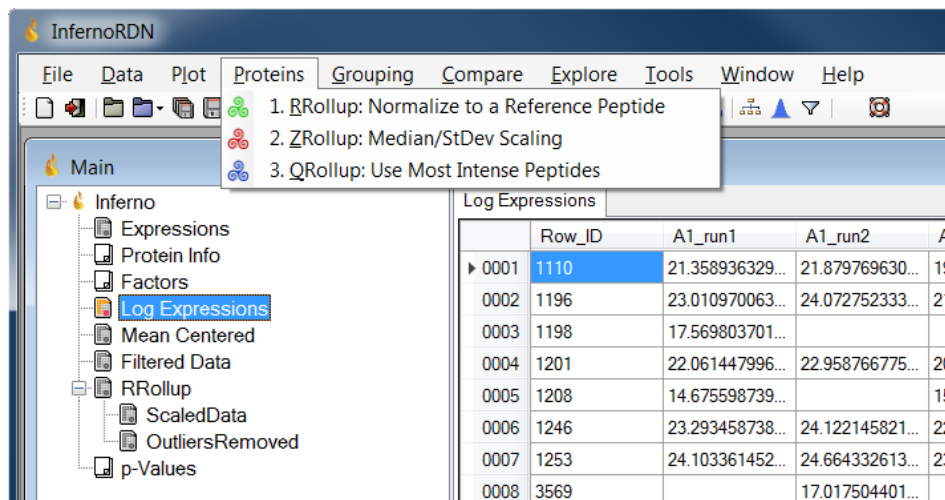


Rollup Methods



InfernoRDN has three methods for estimating protein abundances based on observed peptides, aka protein rollup.

- RRollup: Reference peptide based scaling
 - Identifies a reference peptide, then computes peptide ratios relative to the reference
- ZRollup: Median / standard deviation based scaling
 - Peptides are median centered first and then scaled by the row standard deviation. Protein abundance is obtained as the median of the abundances of the peptides in the group.
- QRollup
 - Takes the top user selected percentage of peptides for each protein then computes the average (or median). In other words, compute a trimmed mean (or trimmed median).

The three methods are described in detail on the following pages.

RRollup - Reference Peptide Based Scaling

Note that this method correctly works only with log transformed data.

A reference peptide which has the most presence across all the datasets, is chosen from the group of peptides that belong to a protein. If there are multiple candidates, the most abundant one is chosen. Then the ratios of peptide abundances with respect to the reference are computed (since the data is assumed to be in log scale, the differences are used) and their median is used as a scaling factor. Protein abundance is obtained as the median of the resulting peptide abundances.

RRollup Options

RRollup - Reference Peptide Based Scaling, Rollup

Data Source: **Log Expressions** This method assumes that the data is in log scale.

Select Options for Peptide Scaling

Minimum Presence of at least one Peptide for a Protein (%) Minimum Number of Peptides required for Grubbs' Test

Exclude peptides from scaling if they are at least not present in this many datasets p-value Cutoff for Grubbs' Test

Include One-Hit-Wonders ☐ Rollup as Mean - default Median ☐

☐ Mean Center Peptides to Zero Mean

☐ Plot each Protein/Peptide profile to a folder (WARNING: Could be very slow)

...

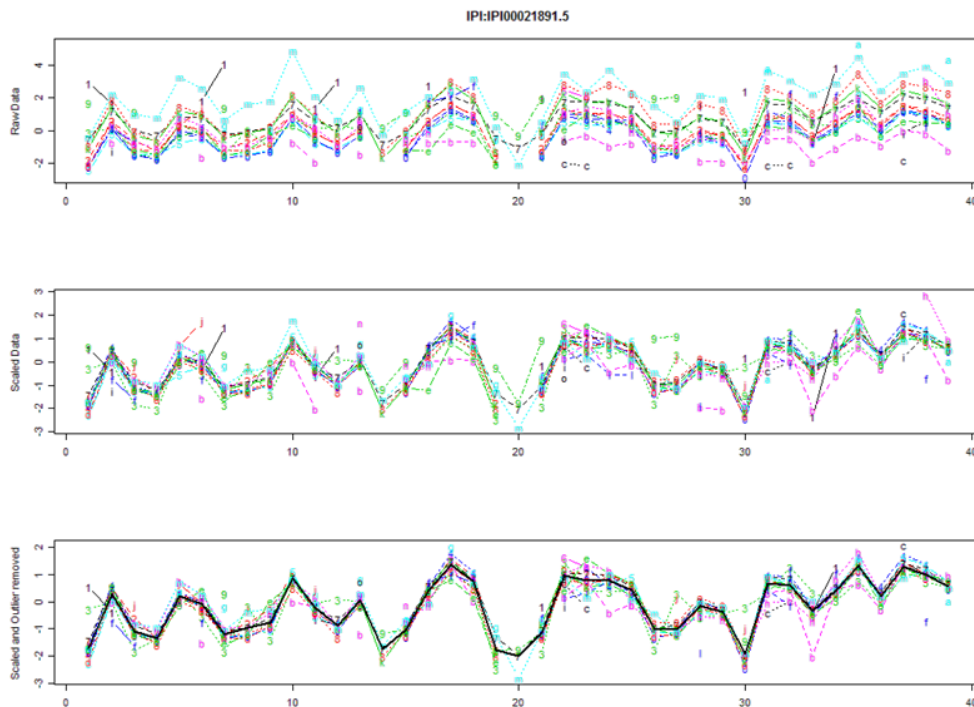
OK Defaults Cancel

Minimum Presence of at least one Peptide for a Protein: Peptides that have too many missing values below this percentage are dropped.

Exclude peptides from scaling if they are at least not present in this many datasets: Within a group of peptides for a specific protein, the ones that do not overlap well (controlled by this value) are not scaled but they are kept to calculate the final protein abundance.

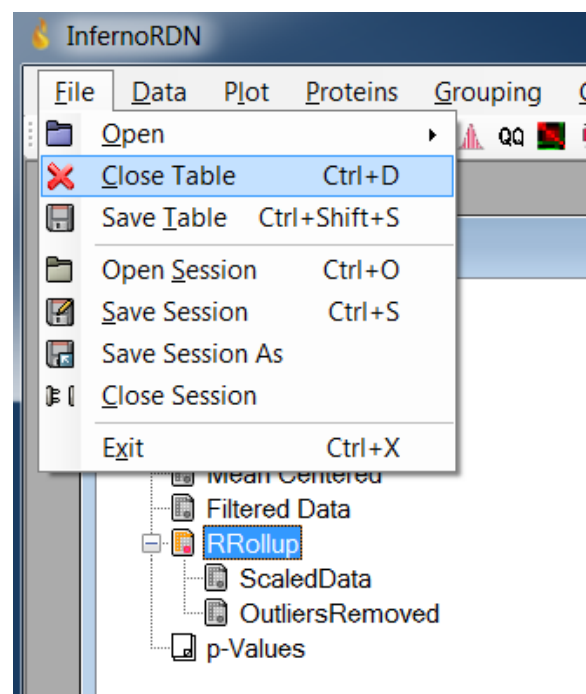
Include Single peptide/protein matches (i.e. 'One-Hit-Wonders'): Protein with only one observed peptide will be included in the final list of proteins. The rationale behind this is that if a particular protein may have only one peptide but it may be quite abundant and present throughout giving some strong confidence on the presence of the protein.

If the plotting is enabled using the checkbox, the scaling results will be plotted for each peptide group as follows:



Note: this option worked in DAnTE but appears broken in InfernoRDN. As an alternative, use menu option Plot, then Protein Rollup as describe on the final page of this file.

To redo RRollup, you must first delete the RRollup table. Select RRollup in the data table list, then choose File, Close Table. Alternatively, use Ctrl+D



ZRollup - Median / Standard Deviation Based Scaling

Peptides are median centered first and then scaled by the row standard deviation. Protein abundance is obtained as the median of the abundances of the peptides in the group.

ZRollup Options

ZRollup - Median / St.Dev. based Scaling and Rollup

Data Source: **Expressions**

Select Options for Peptide Scaling

Minimum Presence of at least one Peptide for a Protein (%) Minimum Number of Peptides required for Grubbs' Test

Include One-Hit-Wonders ☐

Rollup as Mean - default Median ☐

p-value Cutoff for Grubbs' Test

☐ Plot each Protein/Peptide profile to a folder (WARNING: Could be very slow)

F:\My Documents\Projects_CommunityApplications\InfernoRDN\Docume ...

OK Defaults Cancel

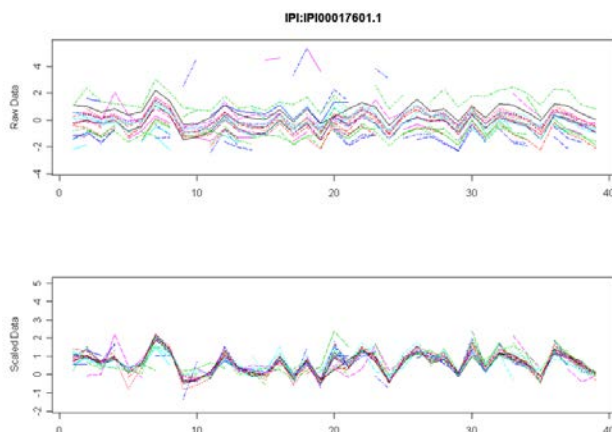
Minimum Number of Peptides required for Grubb's Test: Grubb's test for outliers is performed for peptide groups in each dataset. This value controls the minimum number required to perform the test.

p-value Cutoff for Grubb's Test: This controls the cutoff value for removing outliers determined using Grubb's test.

Include Single peptide/protein matches (i.e. 'One-Hit-Wonders'): Same as in the first method.

If the plotting is enabled using the checkbox, the scaling results will be plotted for each peptide group. As mentioned above, this may not work and you should instead try

Note: this functionality appears broken. As an alternative, use menu option Plot, then Protein Rollup as describe on the final page of this file.



QRollup method

QRollup method simply takes the top user selected percentage of peptides and averages (either mean or median) to obtain the protein abundance. The parameters to control are:

QRollup Options

QRollup

Data Source: **Expressions**

Method

☐ Mean

☒ Median

Top Peptides

☒ Percentage (%) 33

☐ Number 3

Percentage will select the specified top percentage of peptides at each dataset. Top percentage is calculated for each dataset separately.

If top number of peptides to use is specified, the top peptides are selected after summing the abundances of peptides across all the datasets.

Minimum Presence (%) 50 ☐ Include One-Hit-Wonders

OK Cancel

Threshold: All peptides that have abundance values above this percentage at each dataset are used for protein abundance calculation.

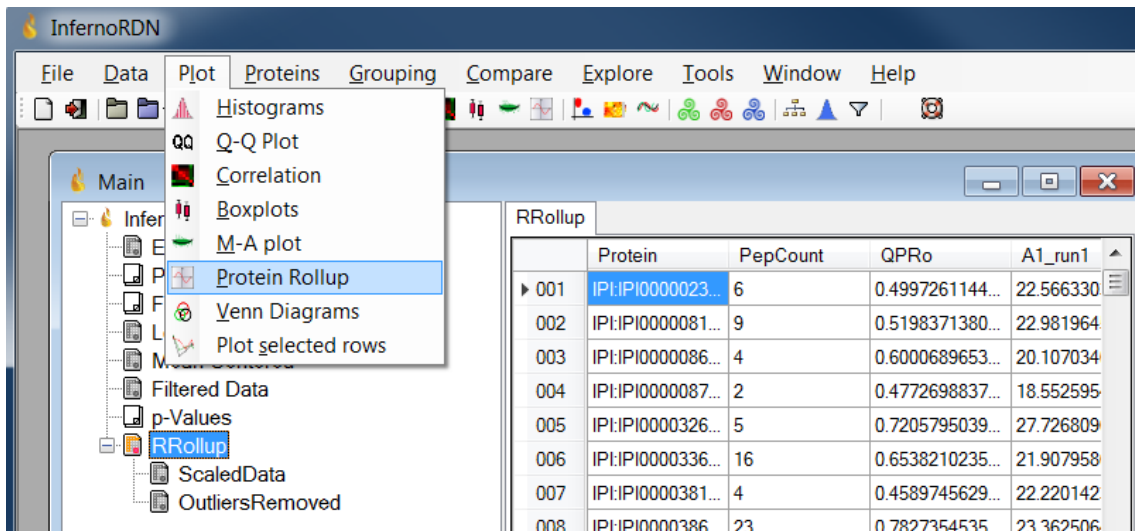
Minimum Presence: Minimum presence of the most present peptide in a peptide group.

Method: Method to calculate Protein abundance from peptides. Either mean or median of the peptides in the group.

Include Single peptide/protein matches (i.e. 'One-Hit-Wonders'): Same as in first method.

Protein Rollup Plots

To visualize protein abundances across samples, select a specific rollup table, like RRollup, then choose Protein Rollup under the Plot menu



You will see a list of protein names. Select one, then click Plot and a new plot will be created.

