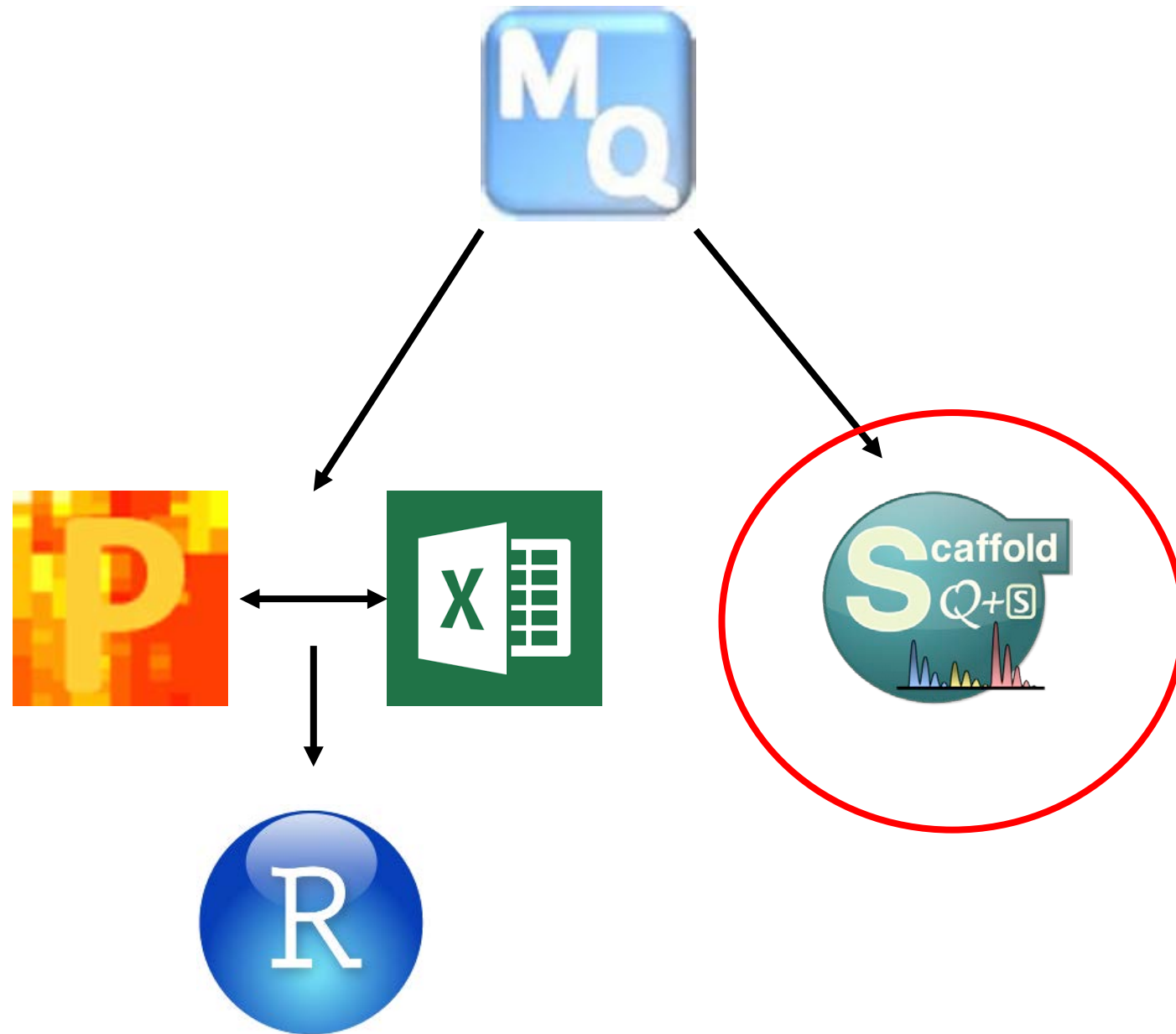
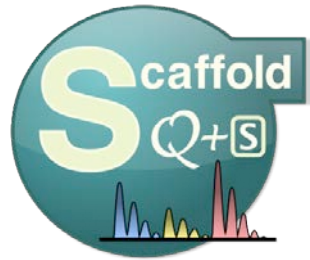


Basics of Using Scaffold and Visualizing Data

Stephanie Byrum, PhD





Scaffold Q+S Quant Module

Experimental Design Wizard

1. Analysis Type
2. **Experiment Type**
3. Edit Sample Names and Categories
4. Organize Quant Samples
5. Approve Settings

Experiment Type

Between-subjects (Independent Groups)

All measurements, including those in the reference category, are taken from independent subjects (e.g., the reference category comprises a control group).

Between-subjects (Common/Pooled Reference)

All measurements are taken from independent subjects, except for the reference, which is common to all BioSamples (e.g., a pooled sample or standardized control sample).

Repeated Measures / Time Course

A within-subjects design in which individual subjects are measured multiple times.

Help Previous Next Done Cancel



Scaffold Q+S Quant Module

Experimental Design Wizard

1. Analysis Type
2. Experiment Type
3. Edit Sample Names and Categories
4. Organize Quant Samples
5. Approve Settings

Click on a cell and type to edit category or sample name. The first category is always the reference category.

Edit Sample Names and Categories

Quant Categories

Number of categories (including the reference): 4

Name	Color
Reference	
Category 1	
Category 2	
Category 3	

Quant Samples

Name	ID	MS Sample	BioSample
Quant 1	TMT-126	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 2	TMT-127N	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 3	TMT-127C	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 4	TMT-128N	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 5	TMT-128C	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 6	TMT-129N	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 7	TMT-129C	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 8	TMT-130N	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 9	TMT-130C	Experiment Workshop_TMT 10plex:...	TMT 1
Quant 10	TMT-131	Experiment Workshop_TMT 10plex:...	TMT 1

Help Previous Next Done Cancel



Scaffold Q+S Quant Module

Experimental Design Wizard

1. Analysis Type
2. Experiment Type
3. Edit Sample Names and Categories
4. Organize Quant Samples
5. Approve Settings

Drag and drop samples or select samples to be moved and right-click on the destination cell.

Create additional Reference Alignment Groups if necessary to associate specific Quant Samples to different references.

Organize Quant Samples

Unorganized Samples

Name	ID	MS Sample	BioSample

Organized Samples

Reference	Category 1	Category 2	Category 3
Quant 10	Quant 1	Quant 14	Quant 17
Quant 20	Quant 11	Quant 15	Quant 18
	Quant 12	Quant 16	Quant 19
	Quant 13	Quant 4	Quant 7
	Quant 2	Quant 5	Quant 8
	Quant 3	Quant 6	Quant 9

Add Reference Alignment Group Remove Reference Alignment Group Clear All

Help Previous Next Done Cancel



Scaffold Q+S Quant Module

Experimental Design Wizard

1. Analysis Type
2. Experiment Type
3. Edit Sample Names and Categories
4. Organize Quant Samples
5. Approve Settings

These are the recommended settings based on what you've told us.

You may change them by clicking the Edit Settings button.

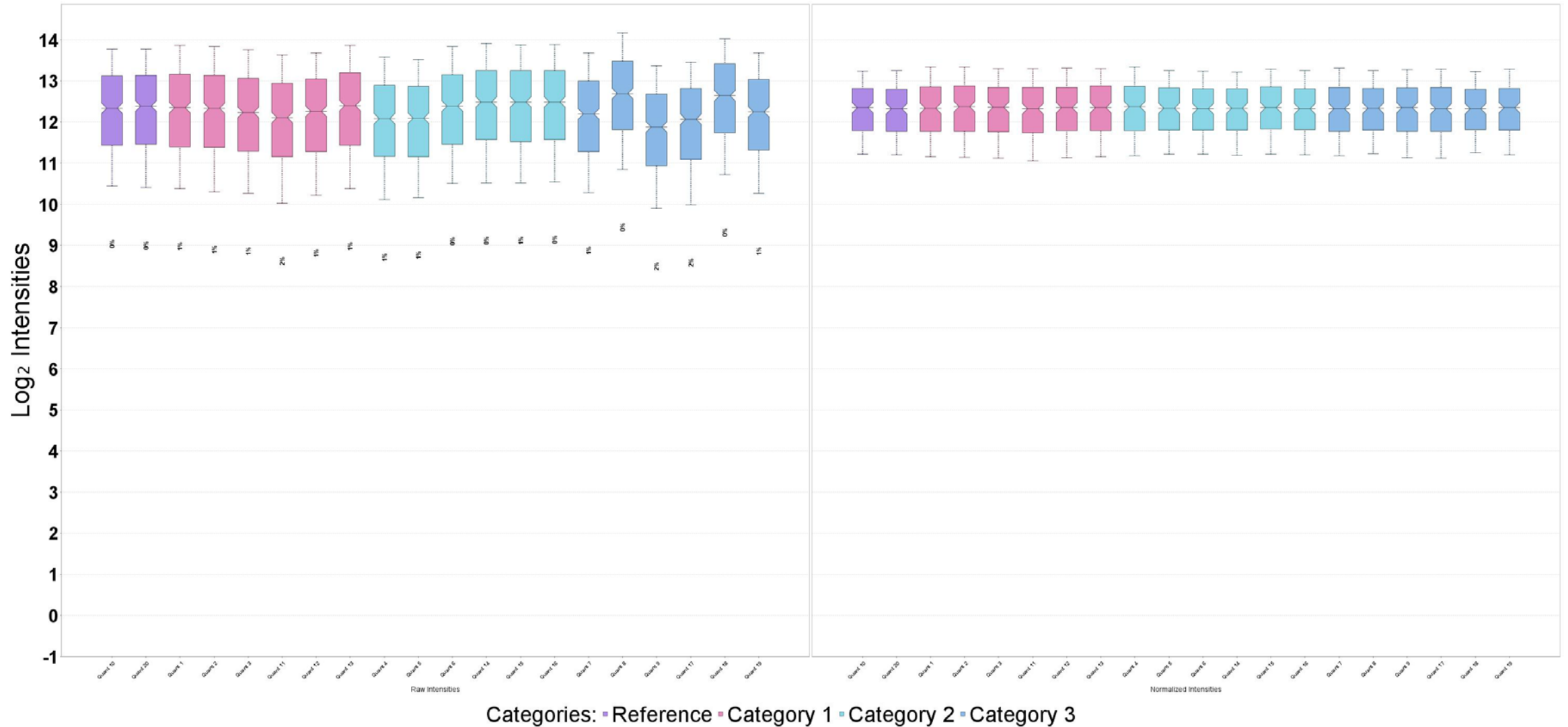
Approve Settings

for Between-subjects (Common/Pooled Reference), Intensity-based Analysis:

<input type="checkbox"/> Experiment:	Workshop_TMT10plex_MQ
<input type="checkbox"/> Quantitation Preferences:	
<input type="checkbox"/> Condenser Preference:	
<input type="checkbox"/> Use Non-Exclusive Peptides:	false
<input type="checkbox"/> Normalization Preference:	
<input type="checkbox"/> Calculation Type:	Median
<input type="checkbox"/> Blocking Level:	Unique Peptides
<input type="checkbox"/> Use Protein Average As Reference:	false
<input type="checkbox"/> Spectrum Quality Filter:	Reference value required



Scaffold Q+S Normalization





Scaffold Q+S Normalization

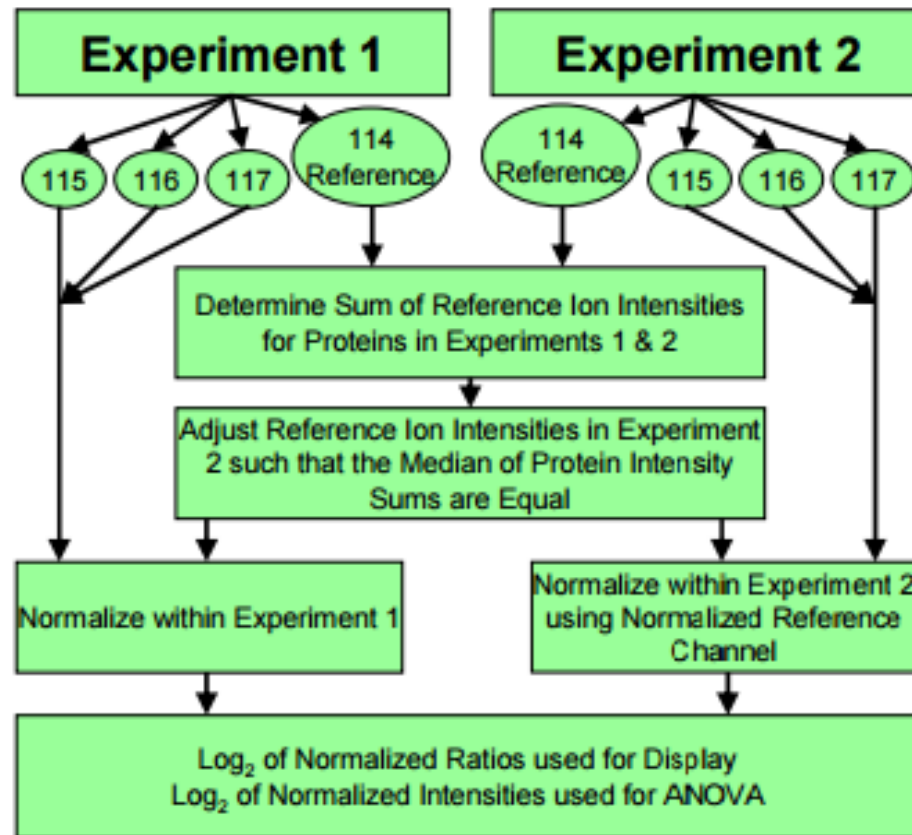
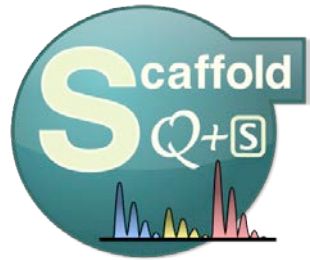


Figure 4: When combining multiple iTRAQ experiments Q+ first normalizes between all experimental reference channels. The normalized references are then used to further normalize ratios within each individual experiment.

Reporter ion channels are typically median normalized with respect to a user-defined reference sample. Employing median ratio normalization ensures that the calculated ratios are no longer dependent on missing or suspect ion intensities because those ratios are not used in the calculation.

An experimental design that daisy chains multiple iTRAQ experiments together using a single reference (but possibly different quantitative samples) can be normalized with an additional step. Before intra-experiment normalization is done, the separate experiments are linked together. The normalization factor is calculated as the median of the summed reporter ion intensities for each protein in the linked reference channels. Normalizing across reference samples in multiple iTRAQ experiments reveals the technical variance between those samples, which can be considered in the final analysis.

\log_2 values of the reporter ion to reference ratios are used to force symmetry between up and down regulated proteins. Statistical test such as the t-Test and ANOVA in Scaffold Q+ are based on \log_2 of the underlying median normalized reporter ion intensities. Testing using the ratios alone ignores variance in the reference channel. Testing using the reporter ion intensities ensures that reference channel variation is considered, while log normalization forces the tests to be magnitude independent.



Scaffold Q+S Normalization

Appendix B. Normalization

Quantitative Raw data imported into the Q+ Quantitation Module are processed for Bias removal with two possible different normalizations methods:

- [Intensity-Based Normalization](#)
- [Ratio-Based Normalization](#)

Intensity-Based Normalization

The method used is a non-parametric version of a statistical model proposed by Ann Oberg and others for analyzing iTRAQ data¹. A technical description of the different steps included in the normalization algorithm developed in Scaffold Q+ or Scaffold Q+S is provided below.

Pre-Processing

1. The raw intensity data is acquired from the spectra and purity-corrected as appropriate.
2. Spectra not assigned exclusively to one protein are discarded (unless the user has indicated otherwise).
3. The data is transformed by applying a logarithm (base 2).
4. Within each MS-Sample a “missing value” is assigned as the larger between (a) the minimum logged intensity acquired and (b) the value whose z-score is -4 for the distribution of all logged values in the MS-Sample. These missing values are applied to all intensities which either had raw value zero, or fall below the “Minimum Dynamic Range.”
5. If the option is selected, spectra with missing values in the reference channel are removed.



Scaffold Q+S Normalization

Iterative Median Polish

A version of Tukey's median polish is applied iteratively to normalized the data. In what follows, all computations are on the logged data, and "average" denotes either median or mean depending on the mode the user has selected. (Median is the default mode.) The normalization has four steps:

1. **Inter-Sample Normalization:** A normalization factor consisting of the global average minus the within-MS-Sample average is added to each data point.
2. **Intra-Sample Normalization:** A normalization factor consisting of the within-MS-Sample average minus the within-channel average is added to each data point.
3. **Peptide/Spectrum Normalization:** For each protein, the averages across the values in each spectrum are brought into alignment by adding a per-spectrum normalization factor (average of these averages minus the particular spectrum's average).
4. **Intensity Weighting:** A weight is assigned to each spectrum based on a t-statistic derived from percent deviations from channel averages.

Steps (1)-(4) are repeated three times.

If the user has indicated that their reference consists of a pooled or common sample, a protein reference normalization step is applied to bring these reference samples into alignment.

Post-Processing

A standard deviation estimate is derived for each spectrum, based on smoothed within-protein deviations of spectral values from averages binned by total intensity.

The weight of each spectrum is divided by the total number of spectra matched to its peptide, providing a form of intermediate peptide-level averaging when subsequently computing protein-level values.

1. Elizabeth G. Hill, John H. Schwacke, Susana Comte-Walters, Elizabeth H. Slate, Ann L. Oberg, Jeanette E. Eckel-Passow, Terry M. Thomeau, and Kevin L. Schey, A STATISTICAL MODEL FOR ITRAQ DATA ANALYSIS, *J Proteome Res.* 2008 Aug; 7(8): 3091–3101. Published online 2008 Jun 26. doi: 10.1021/pr070520u