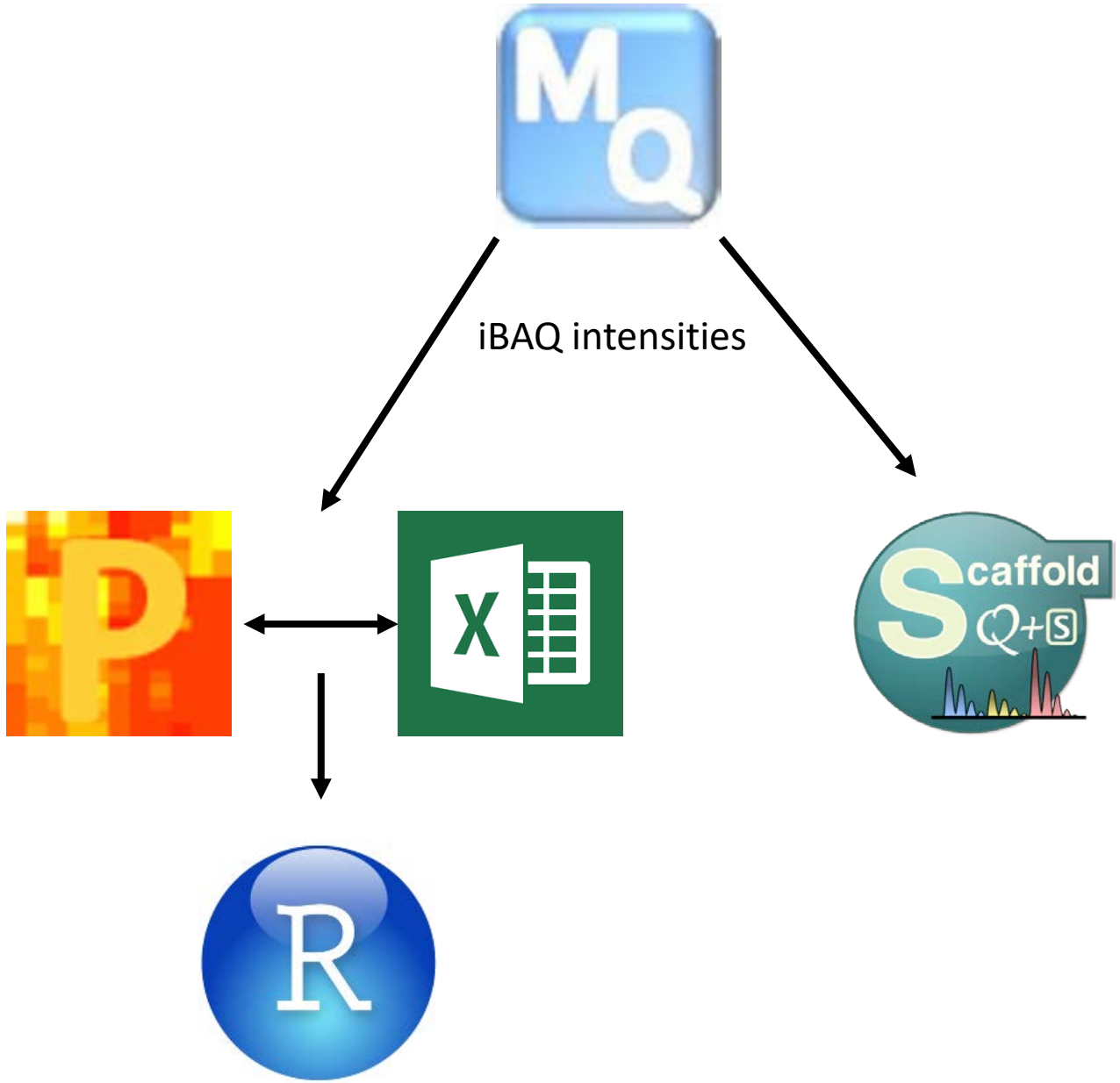
Label-free Protein Intensity

Stephanie Byrum, PhD


iBAQ intensities
Scaffold
Q+S

MaxQuant References

Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ*

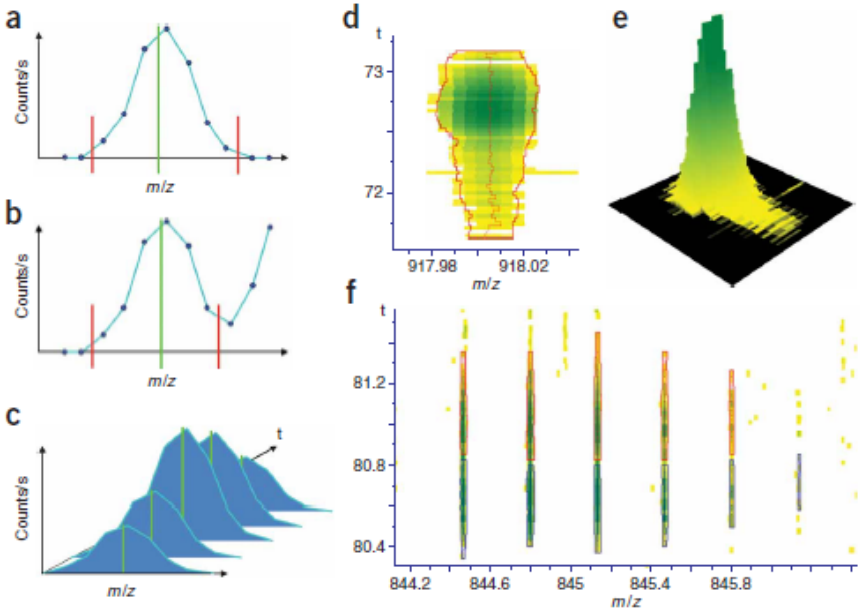
Jürgen Cox‡§, Marco Y. Hein‡, Christian A. Luber‡, Igor Paron‡, Nagarjuna Nagaraj‡, and Matthias Mann‡§

Molecular & Cellular Proteomics 13.9

MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification

Jürgen Cox & Matthias Mann

NATURE BIOTECHNOLOGY VOLUME 26 NUMBER 12 DECEMBER 2008



Global quantification of mammalian gene expression control

Björn Schwanhäusser[1], Dorothea Busse[1], Na Li[1], Gunnar Dittmar[1], Johannes Schuchhardt[2], Jana Wolf[1], Wei Chen[1] & Matthias Selbach[1]

19 MAY 2011 | VOL 473 | NATURE | 337

$$riBAQ = \frac{iBAQ}{\sum iBAQ}$$

Accurate Label-Free Protein Quantitation with High- and Low-Resolution Mass Spectrometers

Jocelyn F. Krey[#1], Phillip A. Wilmarth[#2], Jung-Bum Shin[1], John Klimek[3], Nicholas E. Sherman[4], Erin D. Jeffery[4], Dongseok Choi[5], Larry L. David[2,3], and Peter G. Barr-Gillespie[1,*]

J Proteome Res. 2014 February 7; 13(2): 1034–1044. doi:10.1021/pr401017h.

Popular MS1 methods include the iBAQ algorithm, where a protein's total intensity is divided by the number of tryptic peptides between 6 and 30 amino acids in length, and the “top three” method, where the three peptides with highest intensity are used for quantitation.

iBAQ- intensity based absolute quantification

- Measure for Protein Abundance
- Should have at least 3 peptides per protein
- Sum all identified peptide intensities
  - Maximum detector peak intensities of the peptide elution profile
- Divide by the theoretically observed peptides

- The iBAQ of a protein/protein group is the sum of peak intensities of all peptides divided by the number of theoretically observable peptides.
- iBAQ values are approximately proportionate to the number of moles of protein present and thus $iBAQ_i/\Sigma iBAQ_j$ is the relative molar amount of protein *i*.

http://europepmc.org/articles/PMC4820968;jsessionid=rYVxaiCcNMQmRJht86RG.0

http://greproteomics.lifesci.dundee.ac.uk/webpage%20front%20page/dreamweaver%20webpage/maxquant%20how%20to.pdf

iBAQ- intensity based absolute quantification

[**Theoretical** pI: 8.38 / Mw (average mass): 34081.43 / Mw (monoisotopic mass): 34059.87]
90.2% of sequence covered (you may modify the input parameters to display also peptides < 500 Da or > 100000000000 Da):

10 20 30 40 50 60
MEDYTKiekI GEGTYGVVYK grhkTTGQVV AMKkirLESE EEGVPSTAIR EISLLKelrH

70 80 90 100 110 120
PNIVSLQDVL MQDSRLYLIF EFLSMDLKkY LDSIPPGQYM DSSLVKSYLY QILQGIVFCH

130 140 150 160 170 180
SRrVLHRDLK PQNLLIDDKg tikLADFGLA RAFGIPIRVY THEAITLWYR SPEVLLGSAR

190 200 210 220 230 240
YSTPVDIWSI GTIFAELATK KPLFHGDSEI DQLFRifrAL GTPNNEVWPE VESLQDYKNT

250 260 270 280 290
FPKWKPGSLA SHVKNLDENG LDLLSKMLIY DPAKrisgkM ALNHPYFNDL DNQIKkm

	CDC2	iBAQ value
sum intensity	5097200	268274
theoretical peptides	19	

iBAQ = 5097200/19 = 268274

mass	position	peptide sequence
2289.1033	219-238	ALGTPNNEVWPEVESLQDYK
2212.1536	181-200	YSTPVDIWSIGTIFAELATK
1932.9272	280-295	MALNHPYFNDLDNQIK
1926.9894	107-122	SYLYQILQGIVFCHSR
1912.9361	90-106	YLDSIPPGQYMDSSLVK
1851.9381	60-75	HPNIVSLQDVLMQDSR
1801.9231	201-215	KPLFHGDSEIDQLFR
1631.8753	76-88	LYLIFEFLSMDLK
1551.7954	159-170	VYTHEAITLWYR
1516.7489	37-50	LESEEEGVPSTAIR
1411.7791	128-139	DLKPQNLLIDDK
1330.6848	255-266	NLDENGLDLLSK
1209.6738	244-254	WKPGSLASHVK
1185.6150	10-20	IGEGTYGVVYK
1028.5734	171-180	SPEVLLGSAR
950.5015	267-274	MLIYDPAK
934.5026	25-33	TTGQVVAMK
862.4781	144-151	LADFGLAR
786.3338	1-6	MEDYTK
773.4668	152-158	AFGIPIR
702.4396	51-56	EISLLK
606.3246	239-243	NTFPK
524.3303	124-127	VLHR

iBAQ- intensity based absolute quantification

Data Normalization with iBAQ intensities from MaxQuant

1) Calculate the relative iBAQ (iBAQ/∑iBAQ
2) Use Perseus to Log2 transform and impute missing values based on normal distribution
3) Load the data into R studio
4) Calculate a Normalization Factor based on the median to normalize total protein in each sample
5) Proceed with Statistical Analysis
6) Pathway Analysis

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. Therneau, 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