

Supplementary Notes

General Statistical framework for quantitative proteomics

by stable isotope labeling

Pedro Navarro ^{#1,2}, Marco Trevisan-Herraz ^{#1,3}, Elena Bonzon-Kulichenko ^{#1,3}, Estefanía Núñez^{1,3}, Pablo Martínez-Acedo^{1,3}, Daniel Pérez-Hernández^{1,3}, Inmaculada Jorge^{1,3}, Raquel Mesa^{1,3}, Enrique Calvo³, Montserrat Carrascal⁴, María Luisa Hernández⁵, Fernando García⁶, José Antonio Bárcena⁷, Keith Ashman^{6,8}, Joaquín Abian⁴, Concha Gil⁵, Juan Miguel Redondo³ and Jesús Vázquez^{1,3}*

Contents

1. Supplementary Materials and Methods

- 1.1 Stable isotope labeling techniques
- 1.2 Targeted label-free relative protein quantification
- 1.3 Processing of quantitative ^{18}O data using UNiquant algorithm.
- 1.4 Variance-stabilizing normalization (VSN) transformation of iTRAQ data
- 1.5 Quantitation of data from SILAC-HR NH experiment using MaxQuant.

2. The WSPP statistical model

- 2.1 The null hypothesis
- 2.2 Estimation of general variances and confidence intervals
- 2.3 Estimation of the proline factor
- 2.4 Integration of data from technical replicates
- 2.5 Identification of outliers and significant protein abundance changes

3. Supplementary Results

3.1 Evidence that three independent sources of variance are necessary in the WSPP model

3.2 Comparison of the WSPP model with other existing algorithms

4. Tables and Figures

Supplementary Table 1: Elements of the WSPP statistical model

Supplementary Table 2: Definition of fitting weights

Supplementary Table 3: Comparison of significant protein abundance changes in the control vs H₂O₂-treatment experiments

Supplementary Table 4: List of proteins and peptides subjected to a label-free Parallel Reaction Monitoring

Supplementary Table 5: Side-by-side comparison of protein abundance changes in the control vs H₂O₂-treatment SILAC-HR experiment determined by MaxQuant and WSPP.

Supplementary Table 6: Comparison of protein abundance changes obtained in the control vs H₂O₂-treatment experiments using WSPP and in the SILAC-HR experiment using MaxQuant.

Supplementary Figure 1: Analysis of local normality

Supplementary Figure 2: Estimation of the weight constant (k_e) for the different SIL experiments

Supplementary Figure 3: Evidence that three separate sources of variance at the spectrum, peptide and protein level are needed to give a correct description of the different NH experiments.

Supplementary Figure 4: Analysis at the spectrum, peptide and protein levels of the controls vs H₂O₂-treatment experiments

Supplementary Figure 5: Analysis of the high-resolution (A,C,E) and low-resolution (B,D,F) ¹⁸O NH experiments using UNiquant algorithm.

Supplementary Figure 6: Analysis of the TOFTOF and PQD iTRAQ NH experiments using the variance-stabilizing normalization (VSN) strategy.

Supplementary Figure 7: Analysis of the SILAC-HR, pseudo-NH experiment using MaxQuant.

Supplementary Figure 8: Performance of the WSSP model in the analysis of spike-in mixtures of known protein concentrations in a complex background.

Supplementary Figure 9: Verification of some protein abundance changes in the control vs treatment experiments by label-free Parallel Reaction Monitoring.

Supplementary Figure 10: Comparison of changing ribosomal, oxidative stress and metabolic proteins in the control vs treatment experiments.

See also additional information in supplementary files:

Supplementary File 1: Complete list of proteins quantified by the different SIL-MS approaches for the null hypothesis

Supplementary File 2: Complete list of proteins quantified by the different SIL-MS approaches for the controls vs treatment experiments

Supplementary File 3: Complete list of protein quantifications integrated from the different SIL-MS approaches for the null hypothesis and for the controls vs treatment experiments

Supplementary File 4: Information about the meaning of the parameters used in QuiXoT

5. References

1. Supplementary Materials and Methods

1.1. Stable isotope labeling techniques

iTRAQ and ^{18}O labeling. For iTRAQ labeling, dried peptides were taken up in 30 μl of iTRAQ dissolution buffer provided with the kit (Applied Biosystems) and labeled by adding 70 μl of the corresponding iTRAQ reagent in ethanol and incubating for 1h at room temperature in 70% ethanol, 180 mM triethylammonium bicarbonate (TEAB), pH 8.53. Sample A was labeled with iTRAQ tags 114 and 116, and sample B with tags 115 and 117. After quenching the reaction with 200 μl 0.1% formic acid for 30 min, samples were brought to dryness to completely stop the labeling reaction. This quenching process was repeated once more to promote TEAB volatilization. The four labeled samples were taken-up in 200 μl 0.1% formic acid and combined into one tube. The mixture was dried down, dissolved in 12 μl 5 mM ammonium formate (pH 3), cleaned up with MCX Oasis cartridges (Waters, Milford; Massachusetts, USA) using as elution solution 1 M ammonium formate pH 3, containing 25% ACN, and dried down. The peptide pool was taken-up in 0.5 ml 0.1% TFA, desalted onto C18 Oasis cartridges using as elution solution 50% ACN in 0.1% TFA, and dried down.

Mass spectrometry analysis. Low-resolution analyses of SILAC and $^{16}\text{O}/^{18}\text{O}$ -labeled samples by LC-MS/MS were performed using a linear ion trap LTQ (Thermo-Finnigan, San Jose, CA, USA), as previously described¹⁻³, in the Cardiovascular Proteomics Laboratory at the Centro de Biología Molecular Severo Ochoa (Madrid). The LTQ was operated in a data-dependent ZoomScan- and MS/MS- switching mode¹. Zoom target parameters, number of

microscans, normalized collision energy, and dynamic exclusion parameters were as previously described¹⁻³.

High-resolution analyses of SILAC and ¹⁶O/¹⁸O-labeled samples were performed using a nano-HPLC Ultimate 3000 (LC-Packings) coupled to an LTQ-Orbitrap XL ETD mass spectrometer (Thermo-Finnigan) in the Proteomics Unit of the Centro Nacional de Investigaciones Cardiovasculares (Madrid). Peptides were loaded onto a C-18 reversed phase (RP) nano-column (100 µm I.D., 12 cm, Teknokroma) and analyzed in a continuous acetonitrile gradient consisting of 0-43% B for 140 min, 50-90% B for 1 min (B=95% acetonitrile, 0.5% acetic acid). Peptides were eluted from the RP nano-column to an emitter nanospray needle at 300 nL/min for real time ionization and peptide MS analysis. An enhanced FT-resolution survey scan (resolution=60,000) in the mass range of m/z 300–2,000 was followed by a data-dependent MS/MS of the five most intense ions in the LTQ. The AGC target value in the Orbitrap for the survey scan was set to 1,000,000. Fragmentation in the LTQ was performed by collision-induced dissociation with a target value of 10,000 ions. Full target was 30,000, 1 microscan, 100 ms injection time and 35% normalized collision energy. Dynamic exclusion was set at 0.5 min.

iTRAQ-labeled samples were analyzed using a linear ion trap LTQ (Thermo-Finnigan, San Jose, CA, USA) in the PQD scanning mode at the Proteomics Facility of the IDIBAPS (Barcelona). The LTQ was programmed to perform a data-dependent MS/MS scan on the 15 most intense precursors detected in a full scan from 400 to 1,600 amu (3 µscans, 10 ms injection time, 10,000 ion target parameter). Singly-charged ions were excluded from the MS/MS analysis. Dynamic exclusion was enabled by using the following parameters: 2 repeat counts, 20 s repeat duration, 500 exclusion size list, 30 s exclusion duration and 1.2 amu exclusion mass width. PQD parameters were set at 100 ms injection time, 3 microscans per

scan, 2 amu isolation width, 32% normalized collision energy, 0.6 activation Q and 0.3 ms activation time. For PQD spectra generation 10,000 was used as ion target parameter and automatic gain control was used to prevent over-filling of the ion trap.

Analysis of iTRAQ-labeled samples by MALDI-TOF/TOF was performed at the Proteomics Unit of the Centro Nacional de Investigaciones Oncológicas (Madrid). Peptides from each OFFgel fraction were trapped into a 200 μ m x 5 mm monolithic PS-DBV capillary column previously equilibrated in 0.065% heptafluorobutyric acid (HFBA) in water. Desalted peptides were separated on a 200 μ m x 50 mm monolithic PS-DVB column (PepSwift. LC-Packings, Amsterdam), using a linear gradient from 100 % A (0.055% TFA in water) to 60% B (0.055% TFA in 50% ACN, v/v) for 40 minutes at 2.5 μ l/min at 60°C. A total of 260 six-second fractions were spotted on a stainless steel plate together with 2.5 mg/ml α CHCA in 50% ACN/0.1% TFA (v/v), 5mM ammonium dihydrogen-phosphate, containing 10 fmol/ μ l Glu-Fib (MW=1570.668 Da) as internal standard. Samples were subjected to MALDI MS/MS analysis using a 4800 MALDI-TOF/TOF Proteomics Analyzer equipped with TOF/TOF ion optics (Applied Biosystems/MDS Analytical Technologies, Foster City, CA) and 4000 Series Explorer V.3.5.1. The instrument was operated in positive ion reflector mode and calibrated with 25 fmol Peptide Calibration Mass Standard (Applied Biosystems/MDS Analytical Technologies, Foster City, CA) and the above mentioned Glu-Fib solution as the internal standard. The laser power was set to 3,600 arbitrary units for MS and 4,400 for MS/MS. MS spectra were acquired across the mass range 700–3,600 Da with a total accumulation of 2,000 laser shots and minimum S/N filter of 25 for precursor ion selection. MS/MS spectra were acquired for the 10 or 15 most abundant precursor ions, with a total accumulation of 4,000 laser shots for each precursor.

Peptide identification. Except for TOF/TOF MS/MS files, protein identification was carried out by searching against a joint Human+YeastSwissprot database (Uniprot release 57.3 May 2009, 26885 entries), in order to increase statistical power of peptide identification, which was supplemented with porcine trypsin, and using SEQUEST algorithm (Bioworks 3.2 package, Thermo Finnigan). For ¹⁸O-labeled samples variable (methionine oxidation, lysine and arginine modification of +4 Da) and fixed modifications (cysteine carboxamidomethylation) were used. For iTRAQ, we allowed variable (methionine oxidation) and fixed modifications (cysteine carboxamidomethylation, lysine and N-terminal modification of +144.1020 Da). For SILAC, we used variable (methionine oxidation, lysine and arginine modification of +6 Da) and fixed modifications (cysteine carboxamidomethylation). Precursor mass tolerance was set to 2Da, fragment mass tolerance to 1.2Da and up to two missed cleavage sites for trypsin were allowed. The same collections of MS/MS spectra were also searched against inverted databases constructed from the same target databases. TOF/TOF MS/MS spectra were converted to MGF files by using the 4000 Series Explorer V.3.5.1 internal processor and the following parameters: mass range from precursor -20 Da to 60 Da, peak density of 50 peaks per 200 Da with S/N >1 and minimum area >1 with a maximum number of 100 peaks per precursor. These files were analyzed with the Mascot V2.2.06 search engine using the database and search conditions indicated above.

1.2. Targeted label-free relative protein quantification

The relative quantitative differences in several proteins between control and H₂O₂-treated samples observed by SIL we confirmed by label free Parallel Reaction Monitoring (PRM) on a quadrupole-equipped bench-top Orbitrap instrument⁴. Five proteins changing their abundance (Supplementary Table 3), together with 4 control non-changing proteins detected in our previous SIL experiments were selected for this experiment. For these 9 proteins

(Supplementary Table 4), proteotypic peptides were selected according to the criteria described by⁵. Desalted peptides from whole proteome in-gel digestion⁶ of samples A and B were loaded onto a C-18 reversed phase (RP) Acclaim PepMap 100 nano-viper column (75 μm I.D., 50 cm, 3 μm, 100 Å using a nano-HPLC Easy-nLC 1000 coupled to a QExactive instrument (Thermo-Finnigan, San Jose, CA, USA). Peptides were RP-fractionated using a continuous acetonitrile gradient 0-30% B in 1 hour (B=95% acetonitrile, 0.1% formic acid) at 200 nL/min. Samples A and B were run 6 times each, and 3 blank runs were performed between each sample. A full mass spectrum at 70.000 resolution (AGC target 1.000.000, 250 ms maximum injection time, m/z 300-1300) was followed by 34 MS/MS scans triggered by an inclusion list (Supplementary Table 4) at 17.500 resolution (AGC target 2x105, 120 ms maximum injection time), ±2 Th isolation width and 27% normalized collision energy.

MS/MS data were searched with Mascot (MatrixScience) against a database containing the 9 yeast proteins to be targeted. A decoy database was generated by reverting the amino acid sequences between tryptic cleavage sites and appended to the target database. Precursor mass tolerance was set to 10 ppm, fragment mass tolerance to 10 mmu and no missed cleavages by trypsin were allowed. The search results were used to create a spectral library using Skyline⁷ at 5% FDR. The sum of the peak areas for the extracted ion chromatogram of the precursor and the corresponding fragment ions in the high-resolution MS1 and MS2 scans of the targeted LC-MS/MS runs was calculated automatically for each peptide by the Skyline software. Statistical analysis of the relative protein quantification was performed using SRMstats⁸. Results are presented in Supplementary Figure 9.

1.3. Processing of quantitative ¹⁸O data using UNiquant algorithm.

To avoid bias due to labeling efficiency in the high and low resolution ¹⁸O null-hypothesis (NH) experiments the peak intensities of the heavy and light species in each spectrum were

calculated taking into account the labeling efficiency for each identified peptide³. The relative abundance of each peptide was then calculated as the sum of peak intensities for all the heavy species of the peptide divided by the sum of intensities of the light species of the peptide⁹. Likewise, the relative abundance of each protein was calculated as the ratio between the sum of intensities of all its heavy peptides and the sum of intensities of all its light peptides⁹. Normalized protein quantifications were calculated as $z_{eq} = (x_{eq} - \bar{x})/\sigma$, where \bar{x} was the median of all protein quantifications and σ was estimated by the robust method described in the section 2.2. Protein abundance changes were considered as statistically significant at 5% FDR_{eq} .

1.4. Variance-stabilizing normalization (VSN) transformation of iTRAQ data.

iTRAQ reporter intensities corrected for the isotopic contaminants were subjected to variance-stabilizing normalization transformation using the VSN software¹⁰. To compare the results after this transformation with those obtained with our statistical model, all vsn-transformed log2-ratios were corrected by subtracting the median of all data and fitted to a cumulative distribution with zero mean and a variance estimated by the robust median absolute squared deviation method described in the section 2.2. Spectra quantifications were considered as outliers at $FDR_{eqps} < 5\%$. We then assumed the spectrum variance to be constant after VSN transformation and ignored the presence of variances at the peptide and protein levels. This was done by integrating spectrum data into peptide and protein values using the same formulas used in the model and setting the weight constant k_e and the generalized peptide σ_{Pe}^2 and protein variances σ_{Qe}^2 to zero.

1.5. Quantitation of data from SILAC-HR NH experiment using MaxQuant.

The raw files from the SILAC-HR NH experiment were processed by MaxQuant (version 1.2.2.5)¹¹ using the same protein database and fixed and variable modifications used for the analysis with the WSPP model. The precursor mass tolerance was 20 ppm for the first search and 6 ppm for the main search, whereas the fragment mass tolerance was set to 1.2 Da. In MS/MS spectra a peak density of 10 top peaks per 100 Da were used. The minimum peptide length was 6 amino acids and the FDR for peptide and protein identification was set to 0.01. Up to two missed cleavage sites for trypsin were allowed. To compare MaxQuant results with those obtained using WSPP, the normalized protein ratios obtained from MaxQuant were transformed into log₂-ratios (x_q) and ranked according to summed peptide intensities. The significance of protein abundance change was assessed calculating the standardized variable z_{eq} in sliding windows of 200 data-points using a robust estimation of the variance as described¹¹ and using 5% FDR_{eq} .

2. The WSPP statistical model

2.1 The null hypothesis

The main parameters used by the WSPP model are listed in Supplementary Table 1, together with definitions and some useful comments. The relative concentration of a peptide as measured in a given spectrum is expressed as the 2-base logarithm of the ratios in the control and treated samples, expressed in units of area or peak intensity. The null hypothesis model states that any quantification made in scan s coming from peptide p derived from protein q in experiment e may be expressed as the results of a combination of random errors and fixed effects at three levels:

$$x_{eqps} = \mu_e + \rho_{eq} + \beta_{eqp} + \xi_{eqps} \quad (1)$$

where μ_e and ρ_{eq} measure the systematic and random errors in the concentration of protein q introduced by manipulation of protein extracts in experiment e ; β_{eqp} is the error introduced when peptide p is prepared from protein q in experiment e ; and ξ_{eqps} is the error introduced during the quantification of spectrum s from peptide p coming from protein q in experiment e .

The model assumes that the three random errors are normally distributed:

$$\begin{aligned} \rho_{eq} &\sim N(0, \sigma_{Qe}^2) \\ \beta_{eqp} &\sim N(\phi \cdot r_{eqp}, \sigma_{Pe}^2) \\ \xi_{eqps} &\sim N(0, k_e / v_{eqps} + \sigma_{Se}^2) \end{aligned} \quad (2)$$

where σ_{Qe}^2 and σ_{Pe}^2 are the **general protein and peptide variances**, respectively, ϕ is the **proline factor**, defined as the systematic deviation introduced by each Pro residue as a consequence of Arg to Pro conversion during SILAC labeling, and is a constant characteristic of the cell culture; r_{eqp} is the number of Pro residues in the peptide; v_{eqps} is the **fitting weight** of the spectrum; k_e is the **weight constant**, and σ_{Se}^2 is the **asymptotic spectrum variance**.

The WSPP model assumes that all the proteins in a given experiment are subjected to the same error source; therefore the general protein variance σ_{Qe}^2 is assumed to be constant in experiment e . The same assumption was made for the general variance at the peptide level, so that σ_{Pe}^2 is also constant in experiment e . These assumptions were found to give a satisfactory description of the distribution of these errors, so that no further refinements were considered.

To estimate the error produced when peptides are quantified by mass spectrometry, a fitting weight v_{eqps} was first calculated for each spectrum as defined in Supplementary Table 2. The fitting weights are expressed in arbitrary units and consider quality features, such as peak intensity or goodness of fit, which are characteristic of the MS detector and SIL method used and allow a preliminary ranking of the spectra according to the quality of the quantification. In each experiment e , the variance of the quantification in each spectrum is then modeled as a hyperbolic function of the fitting weight, using the two constant parameters k_e and σ_{Se}^2 , as indicated in Eq. 2. The first parameter acts as a calibration constant that transforms the arbitrary units of v_{eqps} into variance units, and the second parameter is the residual or asymptotic spectrum variance that is always found to be present when the intensity of the MS peaks increase. Taken together, the four parameters describe the distribution of errors in a given experiment. These parameters are estimated from the whole set of quantifications in the experiment using an iterative method, as explained in section 2.2.

In each experiment e , the peptide log₂-ratios, x_{eqp} , are calculated as weighted averages of the corresponding spectrum values; the protein log₂-ratios, x_{eq} , are weighted averages of the peptide values, and the experiment mean, x_e , is the weighted average of protein values, and is a non-biased estimate of the systematic error μ_e .

$$x_{eqp} = \frac{\sum_s w_{eqps} x_{eqps}}{\sum_s w_{eqps}} \quad (3); \quad x_{eq} = \frac{\sum_p w_{eqp} x_{eqp}}{\sum_p w_{eqp}} \quad (4)$$

$$x_e = \frac{\sum_q w_{eq} x_{eq}}{\sum_q w_{eq}} \quad (5);$$

According to classical error propagation theory the statistical weights of each one of the magnitudes in Eqs.3-5 are the inverses of their variances. Furthermore, the variances produced by two independent error sources are additive and the inverse of the variance of an averaged value is the sum of the inverses of variances of the values used to calculate the average. Hence, and according to Eqs.1 and 2, the local variances associated with each one of the spectrum, peptide and protein values are given by:

$$\sigma_{eqps}^2 = \frac{k_e}{v_{eqps}} + \sigma_{Se}^2 \quad (6); \quad \sigma_{eqp}^2 = \frac{1}{\sum_s \frac{1}{\sigma_{eqps}^2}} + \sigma_{Pe}^2 \quad (7); \quad \sigma_{eq}^2 = \frac{1}{\sum_p \frac{1}{\sigma_{eqp}^2}} + \sigma_{Qe}^2 \quad (8)$$

And the statistical weights at each one of the levels may be calculated from the inverses of these values:

$$w_{eqps} = \frac{1}{\frac{k_e}{v_{eqps}} + \sigma_{Se}^2} \quad (9); \quad w_{eqp} = \frac{1}{\frac{1}{\sum_s w_{eqps}} + \sigma_{Pe}^2} \quad (10)$$

$$w_{eq} = \frac{1}{\frac{1}{\sum_p w_{eqp}} + \sigma_{Qe}^2} \quad (11);$$

2.2 Estimation of general variances and confidence intervals

The weight constant k_e , the asymptotic spectrum variance σ_{Se}^2 and the general variances at the protein σ_{Qe}^2 , and peptide levels σ_{Pe}^2 , are estimated on the basis of a robust method which uses the median absolute deviation. For a large, normally distributed population, a robust estimation of the variance is given by $\sigma^2 \approx \text{median}\{(x_i - \bar{x})^2\} / \{\Phi^{-1}(3/4)\}^2$,

where $\Phi^{-1}(3/4) \approx 0.6745$ is the 75th percentile of the normal distribution with $\sigma = 1$. This estimate may also be reformulated as $\text{median}(z_i^2)/\{\Phi^{-1}(3/4)\}^2 \approx 1$, where $z_i = (x_i - \bar{x})/\sigma$ is the standardized variable.

To estimate the weight constant k_e and the asymptotic spectrum variance σ_{Se}^2 , peptides quantified by only one spectrum are ignored and the remaining spectra (*non-unique* spectra) are ranked according to their fitting weights v_{eqps} . A local variance estimate VE_{eqps} is then computed in sliding windows containing 200 spectra by using the robust method:

$$VE_{eqps} = \text{median} \left\{ \left(x_{eqps} - x_{eqp} \right)^2 \cdot \left(\frac{n_{eqp}}{n_{eqp} - 1} \right) \right\} \cdot \frac{1}{\{\Phi^{-1}(3/4)\}^2} \quad , \quad n_{eqp} > 1 \quad (12)$$

where n_{eqp} is the number of spectra per peptide and $n_{eqp}/(n_{eqp} - 1)$ is a local correction for degrees of freedom that ensures that the estimation of local variance is unbiased no matter the value of n_{eqp} and allows the grouping of spectra irrespective of the peptide they come from. The values of VE_{eqps} , which may be considered as estimates of σ_{eqps}^2 , are then plotted against v_{eqps} , and k_e and σ_{Se}^2 are calculated by fitting the curve to Eq.9. The algorithm is run iteratively; peptide averages x_{eqp} in Eq.12 are calculated using initial estimates of k_e and σ_{Se}^2 , and the resulting values are used to re-estimate these parameters until convergence is met. The final values are then used to calculate the local variances σ_{eqps}^2 for all the quantified spectra (unique and non-unique) from their fitting weights v_{eqps} using Eq.9. Representative examples of the curves used to estimate k_e and σ_{Se}^2 are shown in Supplementary Fig.2.

The **standardized variable at the spectrum level** is defined as

$$z_{eqps} = \frac{x_{eqps} - x_{eqp}}{\sqrt{k_e / v_{eqps} + \sigma_{Se}^2}} \cdot \sqrt{\frac{n_{eqp}}{n_{eqp} - 1}} , \quad n_{eqp} > 1 \quad (13)$$

According to the model, if k_e and σ_{Se}^2 are correctly estimated this variable should be normally distributed with a mean equal to zero and a standard deviation of one. In addition, the following condition should also be met:

$$\text{median}(z_{eqps}^2) / \{\Phi^{-1}(3/4)\}^2 \approx 1 \quad (14)$$

Similarly, the **standardized variable at the peptide level** is defined in the population of non-unique peptides (i.e. those belonging to proteins with two or more peptides):

$$z_{eqp} = \frac{x_{eqp} - x_{eq}}{\sqrt{\frac{1}{\sum_s w_{eqps}} + \sigma_{Pe}^2}} \cdot \sqrt{\frac{n_{eq}}{n_{eq} - 1}} , \quad n_{eq} > 1 \quad (15)$$

where n_{eq} is the number of peptides per protein.

The **standardized variable at the protein level** can be defined for the entire population of proteins in the experiment:

$$z_{eq} = \frac{x_{eq} - x_e}{\sqrt{\frac{1}{\sum_p w_{eqp}} + \sigma_{Qe}^2}} \cdot \sqrt{\frac{n_e}{n_e - 1}} , \quad n_e > 1 \quad (16)$$

where n_e is the number of proteins per experiment.

The general peptide and protein variances σ_{Pe}^2 and σ_{Qe}^2 are calculated in a similar manner. These parameters are iteratively varied until the following conditions are met:

$$\text{median}(z_{eqp}^2) / \{\Phi^{-1}(3/4)\}^2 \approx 1 \quad (17)$$

$$\text{median}(z_{eq}^2) / \{\Phi^{-1}(3/4)\}^2 \approx 1 \quad (18)$$

Confidence intervals of the estimated variances are determined by computer simulation using a Monte Carlo approach. For each experiment, 100 virtual experiments are randomly constructed with the same number of quantifications from the spectrum/peptide/protein structure and from the distribution of fitting weights of the real experiment. The variances are then estimated in each of the virtual experiments and the set of estimated variances are used to determine the 95% confidence intervals.

2.3 Estimation of the proline factor

During SILAC labeling some metabolic conversion from Arg to Pro takes place, so that a proportion of Pro residues become labeled; these Pro-labeled species have an increased molecular mass and hence appear at a m/z different to that of the labeled peptide containing non-labeled Pro residues (heavy peptide). As a consequence, when a peptide doublet is quantified the amount (area or intensity) of the heavy species of peptide p coming from protein q in experiment e will diminish in a proportion given by $\theta^{r_{eqp}}$, where θ is the **fraction of Pro residues that remain unlabeled** and r_{eqp} is the number of Pro residues in the peptide sequence. This effect will shift the peptide log₂-ratio by a constant summand given by $r_{eqp}\phi$, where $\phi = -\log_2 \theta$ is the **proline factor**, defined as the systematic deviation introduced by each Pro residue in the peptide value.

The proline factor was estimated by the maximum likelihood method. The likelihood function is the product of Gaussian distributions associated with each one of the spectra, peptides and proteins:

$$L_e(\vec{\theta}|\vec{X}) = L_e(k_e, \sigma_{Se}^2, \sigma_{Pe}^2, \sigma_{Qe}^2, \phi | x_{eqps}, x_{eqp}, w_{eqp}, r_{eqp}) = \prod_q \prod_p \prod_s \frac{\sqrt{w_{eqps}}}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} w_{eqps} (x_{eqps} - x_{eqp})^2\right) \cdot \prod_q \prod_p \frac{\sqrt{w_{eqp}}}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} w_{eqp} (x_{eqp} - r_{eqp}\phi - x_{eq})^2\right) \cdot \prod_q \frac{\sqrt{w_{eq}}}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} w_{eq} (x_{eq} - x_e)^2\right) \quad (19)$$

By taking natural logarithms and applying the maximum likelihood condition to the proline factor, $\partial \ln L_e / \partial \phi = 0$, we get

$$\sum_q \sum_p w_{eqp} (x_{eqp} - r_{eqp}\phi - x_{eq}) \cdot r_{eqp} = 0 \quad (20)$$

and solving for ϕ , the maximum likelihood estimate of the proline factor takes the form:

$$\phi = \frac{\sum_q \sum_p w_{eqp} (x_{eqp} - x_{eq}) \cdot r_{eqp}}{\sum_q \sum_p w_{eqp} \cdot r_{eqp}^2} \quad (21)$$

2.4 Integration of data from technical replicates

To integrate quantitative data obtained for each one of the proteins in different technical replicates, the corresponding \log_2 -ratios must be corrected for the systematic error in each one of the experiments. The **corrected log₂-ratio of protein q in experiment e** is defined as

$$x'_{qe} = x_{eq} - x_e \quad (22)$$

This correction does not change protein variance; however, for consistency in the formulation we will define the **corrected weight and variance of protein q in experiment e** as

$$w'_{qe} = w_{eq} \quad (23) ; \quad {\sigma'}_{qe}^2 = \sigma_{eq}^2 \quad (24)$$

For each protein q , an **averaged protein log₂-ratio** is obtained as a weighted average of the corrected values of q in all the experiments:

$$x'_q = \frac{\sum_e w'_{qe} x'_{qe}}{\sum_e w'_{qe}} \quad (25)$$

The statistical weight and the variance of the averaged protein values are given by:

$$w'_q = \sum_e w'_{qe} \quad (26) ; \quad \sigma'^2_q = \frac{1}{\sum_e w'_{qe}} \quad (27)$$

The variable z_{eq} from the previous section describes the dispersion of protein values around the grand mean in a given experiment (between proteins, within experiments). When a protein is quantified in several experiments (technical replicates), an averaged protein value x'_q is obtained using Eq.25. The dispersion of protein values in different experiments around the protein average (between experiments, within proteins) in the population of non-unique proteins (i.e., proteins quantified in two or more experiments) is described by the standardized variable z'_{qe} :

$$z'_{qe} = \frac{x'_{qe} - x'_q}{\sigma'_{qe}} \cdot \sqrt{\frac{n'_{qe}}{n'_{qe} - 1}} , \quad n'_{qe} > 1 \quad (28)$$

where n'_{qe} is the number of experiments where the protein q is quantified. Again, if the local protein variances are correctly calculated, then z'_{qe} should follow a normal distribution $N(0,1)$. Deviations from this inter-experiment behavior indicate the existence of an experiment yielding a protein value significantly different from the averaged value obtained taking into account all experiments.

Finally, the dispersion of all averaged protein values (unique and non-unique) is described by the standardized variable z'_q :

$$z'_q = \frac{x'_q}{\sigma'_q} \quad (29)$$

Under the null hypothesis, this parameter should also follow a normal distribution $N(0,1)$. Deviations from this behavior correspond to the proteins that undergo a true expression change.

2.5 Identification of outliers and significant protein abundance changes

The standardized variables z_{eqps} , z_{eqp} and z_{eq} defined above may be used to determine the probability that a given spectrum, peptide or protein value is an outlier with respect to the corresponding average. This is done by calculating the two-tailed probability that z deviates from the normal distribution $N(0,1)$, given by

$$P(z) = 1 - \text{erf}\left(\frac{|z|}{\sqrt{2}}\right) \quad (30)$$

Detection of the presence of outliers is made by using multiple hypothesis testing and controlling for the false discovery rate (*FDR*), defined as the proportion of values expected to deviate by chance alone within the population of observed outliers. Since the number of expected random changes is the product of the probability of being an outlier (Eq.30) by the total number of determinations at each level, the *FDR* at each one of the three levels is calculated by

$$FDR_{eqps} = \frac{P(z_{eqps}) \cdot NS_e}{O(z_{eqps})} \quad (31); FDR_{eqp} = \frac{P(z_{eqp}) \cdot NP_e}{O(z_{eqp})} \quad (32); FDR_{eq} = \frac{P(z_{eq}) \cdot n_e}{O(z_{eq})} \quad (33)$$

where NS_e , NP_e and n_e are the total number of non-unique spectra, non-unique peptides and proteins in experiment e , respectively, and $O(z)$ is the observed number of values detected in each population whose standardized variable is in absolute terms higher than z . These *FDRs* allows detection of the presence of outliers in the corresponding population.

Similarly, the standardized variable z'_{qe} can be used to calculate the probability that a protein value in experiment e deviates from the averaged protein value (inter-experiment variation), and the standardized variable z'_q to calculate the probability that an averaged protein value deviates from the expected value of zero (significant expression change at the protein level). The corresponding *FDR* expressions are given by:

$$FDR'_{qe} = \frac{P(z'_{qe}) \cdot NUQ}{O(z'_{qe})} \quad (34); FDR'_q = \frac{P(z'_q) \cdot NQ}{O(z'_q)} \quad (35)$$

where NUQ is the total number of quantifications of non-unique proteins and NQ is the number of different proteins quantified in one or more experiments.

3. Supplementary Results

3.1 Evidence that three independent sources of variance are necessary in the WSPP model

We analyzed whether it was possible to construct a statistical model containing fewer parameters. To that end, we firstly analyzed if the dispersion of the data could be explained in terms of a single source of variance. As shown in Supplementary Figure 3A, a model that assumes the existence of a unique error source at the spectrum level is, in general, unable to

describe the distribution of the data. The result is slightly better when the variance is supposed to generate at the peptide level only, but it still fails to provide a description of general validity (Supplementary Figure 3B). The best agreement with the dispersion of the results are obtained when variance is assumed to arise at the protein level (Supplementary Figure 3C); however, the model fails to describe the dispersion of the data at the extremes, and the number of false protein abundance changes is, in general, unacceptable.

The results improve considerably when the model includes two error sources at the peptide and protein levels (Supplementary Figure 3D) or at the spectrum and protein levels (Supplementary Figure 3E). In the last case the model gives a very satisfactory description of NH data for the iTRAQ and SILAC NH experiments, being the results almost indistinguishable from those produced by the full WSPP model (Supplementary Figure 3F). This finding is in agreement with the nature of SILAC labeling, which is performed at the protein level and therefore is expected not to produce errors at the peptide level. It also suggests that the contribution of peptide variance in the iTRAQ method is small and could be ignored in a first approach. We must note here that a model with two sources of technical variance has never proposed to date to describe SIL experiments, and such a model is just a particular case of the full WSPP model. In spite of these findings, we also observed that inclusion of peptide variance, in addition to spectrum and protein variance was necessary to describe NH data from the ^{18}O experiments with a negligible number of protein outliers. In this last case, the peptide variance is presumably introduced during ^{18}O -labeling, a process that is known to introduce certain variability³.

We conclude that the three-layered WSPP model is therefore necessary to make a general description of quantitative data obtained by SIL techniques. The three-layered model has the added value that it provides a statistical framework to control separately the

experimental errors produced during protein preparation, during peptide preparation and labeling and during mass spectrometry analysis. This makes the model particularly attractive to ensure robustness during experimental application of SIL techniques.

3.2 Comparison of the WSPP model with other existing algorithms

The performance of the WSPP model was compared with that of other representative existing statistical models specifically developed for $^{18}\text{O}^9$, iTRAQ¹² or SILAC¹¹.

We firstly analyzed the accuracy with which each model described the results obtained in the different NH experiments and the number of false protein abundance changes. In the case of the ^{18}O labeling technique, one of the most used algorithms for analyzing this kind of quantitative results is UNiquant⁹. This method uses the sum of spectra intensities to calculate relative peptide abundances, and the sum of peptide abundances to calculate relative protein abundances⁹. To compare the results obtained using this algorithm for data integration with those obtained with the WSPP model, we analyzed the ^{18}O -HR and ^{18}O -LR NH experiments and estimated the local protein variance in sliding windows containing 200 proteins and used these values to calculate a standardized variable at the protein level. As shown in Supplementary Figure 5, a significant number of false protein abundance changes was detected. In the same experiments the WSPP model yielded only 2 and 0 protein outliers, respectively.

The accuracy of peptide quantitation using iTRAQ reporter ions has been recently analyzed, confirming the heterogeneity of variance at the spectrum level¹². In that work a variance-stabilizing normalization (VSN) transformation similar to those used to treat microarray data was used to correct variance heterogeneity¹⁰. To analyze the performance of the VSN method, we subjected our data to the VSN transformation and compared the results with those obtained by the WSPP model. The distributions of log2-ratios (Supplementary

Figure 6A,B) and VSN-ratios (Supplementary Figure 6C,D) show the ability of the VSN transformation at the spectrum level to stabilize the data dispersion across the entire intensity range for high-resolution and low-resolution NH iTRAQ experiments. The cumulative distributions of the VSN-transformed variable (Supplementary Figure 6E,F) and of the standardized protein variable (Supplementary Figure 6G,H), calculated assuming only the spectrum layer, followed very closely a pure Gaussian distribution with zero mean and unit variance for both the TOFTOF and PQD NH experiments. However, a significant number of false protein abundance changes were observed in the case of the PQD experiment (Supplementary Figure 6H), while the WSPP model was able to describe the same set of data with only 1 false positive (Table 1A). These results suggest that the VSN transformation is effective to normalize the data at the spectrum level, but inclusion of variance at the peptide and protein levels is needed for the method to be effective to describe the NH experiments.

The data for the SILAC-HR NH experiment was analyzed using MaxQuant, and the protein variance was calculated using robust estimates in sliding windows of proteins ordered according to the summed peptide intensities, as described¹¹. The results obtained are shown in Supplementary Figure 7. The WSPP model was able to quantify 1459 proteins in the same experiment and detected 18 significant abundance changes, from which 12 had the same behavior as reported by MaxQuant. These results suggested that the two algorithms had a similar performance describing the NH in SILAC-HR experiments. The main difference was that MaxQuant used a local, sliding-window protein variance estimate, while the WSPP model described the NH data using only a set of four variables (k_e , σ_{Se}^2 , σ_{Pe}^2 , σ_{Qe}^2).

The performance of our model was also compared with that of a weighted least-squares method for iTRAQ¹³ in terms of accuracy and FDR using a yeast background with spiked-in

proteins at different concentrations and ratios¹³. This was made by taking from that work iTRAQ 4-plex data obtained on a LTQ-Orbitrap instrument (Study 1) and analyzing the results using the WSPP model. WSPP was able to calibrate satisfactorily the local variances obtained from these data (Supplementary Figure 8A), and to describe with good accuracy the distribution of protein quantifications from experiments IA and IIA (Supplementary Figure 8G,H). A set of figures were then produced that allowed comparison of results obtained with the WSPP model with those published¹³ (Supplementary Figure 8B-F). The comparative suggests that the performance of the WSPP model is similar to the published method, although the WSPP model produces a better discrimination between the true and false positives at low *p*-values.

The performance of the WSPP model to determine protein abundance changes produced by H₂O₂ treatment was also compared with that of MaxQuant. To eliminate the effect of peak quantification and peptide identification algorithms, the raw data from the corresponding SILAC-HR experiment were processed by MaxQuant and the quantitative data at the peak level were then reanalyzed using the WSPP model (Supplementary Table 5). In addition, we also compared the results obtained in the different experiments using our own algorithms with those directly obtained by MaxQuant (Supplementary Table 6). As shown in the two tables, the results from the two algorithms were quite similar, and MaxQuant results were in good agreement with those directly produced by the WSPP model.

Finally, and concerning the SILAC-LR NH and H₂O₂-treatment experiments, we did not find in the bibliography a suitable statistical model to analyze SILAC results from low-resolution ion trap spectra.

4. Tables and Figures

Supplementary Table 1. Elements of the WSPP statistical model

Parameter	Definition	Comments
k_e	weight constant in experiment e	calibration constant that allows determination of local variance from the fitting weight of a spectrum
$\sigma_{S_e}^2$	asymptotic spectrum variance in experiment e	residual error remaining in a “perfect” spectrum, attributed to the MS detector
$\sigma_{P_e}^2$	general peptide variance in experiment e	error introduced during peptide production and manipulation
$\sigma_{Q_e}^2$	general protein variance in experiment e	error introduced during protein extraction and manipulation
x_{eqps}	log ₂ -ratio (LR) of spectrum s of peptide p derived from protein q in experiment e	relative quantification of peptide pairs in the control and treated samples on a base 2 logarithmic scale, expressed in units of area or peak intensity
$v_{eqps}, w_{eqps}, \sigma_{eqps}^2$	fitting weight, statistical weight and local variance of the spectrum	the fitting weight is calculated when a spectrum is quantified and is used to calculate local spectrum variance; the statistical weight is the inverse of variance
x_{eqp}	LR of peptide p derived from protein q in experiment e	determined by weighted averaging of the spectra in which the peptide is detected
$w_{eqp}, \sigma_{eqp}^2, n_{eqp}$	statistical weight, local variance and number of spectra of the peptide	
x_{eq}	LR of protein q in experiment e	determined by weighted averaging of the peptides that belong to the protein

$w_{eq}, \sigma_{eq}^2, n_{eq}$	statistical weight, local variance and number of peptides of the protein	
x_e	grand mean in experiment e	weighted average of protein values in experiment e ; it is an estimate of the systematic error of the experiment
ϕ, r_{eqp}	proline factor; the number of proline residues in peptide p from protein q in experiment e	measures the systematic deviation introduced by each Pro residue in the peptide LR; this deviation is additive
z_{eqps}	standardized variable of spectrum s around peptide p from protein q in experiment e	describes variation of spectrum LR around the corresponding peptide average in units of standard deviation
FDR_{eqps}	False Discovery Rate (FDR) associated with z_{eqps}	false discovery rate of the population of spectra having equal or higher z_{eqps}
z_{eqp}	Standardized variable of peptide p around protein q in experiment e	describes variation of peptide LR around the corresponding protein average in units of standard deviation
FDR_{eqp}	FDR associated with z_{eqp}	false discovery rate of the population of peptides having equal or higher z_{eqp}
z_{eq}	Standardized variable of protein q around experiment e	describes variation of protein LR around the grand mean of the experiment in units of standard deviation
FDR_{eq}	FDR associated with z_{eq}	false discovery rate of the population of proteins having equal or higher z_{eq} ; used to determine significant changes in protein abundance
x'_{qe}	corrected LR of protein q in experiment e	protein LR values corrected for the systematic deviation (grand mean) of the corresponding experiment

$w'_{qe}, \sigma'^2_{qe}, n'_{qe}$	corrected weight, variance of the protein q and number of experiments in which the protein q is quantified	
x'_q	averaged LR of protein q	weighted average of the protein quantified in different experiments
w'_q, σ'^2_q	statistical weight and variance of the averaged LR of the protein	
z'_{qe}	standardized variable of experiment e around protein q	describes variation of protein LR in a given experiment around the averaged protein LR in units of standard deviation
FDR'_{qe}	FDR associated with z'_{qe}	false discovery rate of the population of proteins having equal or higher z'_{qe} ; can be used to determine which experiment values deviate significantly from protein average
z'_q	standardized variable of protein q	describes variation of protein LR around zero in units of standard deviation
FDR'_q	FDR associated with z'_q	false discovery rate of the population of proteins having equal or higher z'_q ; used to determine significant changes in protein abundance

Supplementary Table 2. Definition of fitting weights

Labeling method	MS platform	Fitting weight (v_{eqps})	Comments ⁽¹⁾
¹⁸ O	LTQ (Zoom scan)	$\frac{I_{\max}^2}{SQ_{\text{pep}} + SQ_{\max}}$	SQ_{pep} : in the peptide isotopic envelope; SQ_{\max} : within +/- 2 m/z units from the isotopic envelope depending on the direction of the change.
	Orbitrap (Full scan)	$\frac{I_{\max}^2}{SQ}$	SQ : in the peptide isotopic envelope
SILAC	LTQ (Zoom scan)	$\frac{I_{\max}^2}{SQ_1 + SQ_2}$	SQ_1 : in the light peptide isotopic envelope; SQ_2 : in the heavy peptide isotopic envelope
	Orbitrap (Full scan)	$\frac{I_{\max}^2}{SQ}$	SQ : in the peptide isotopic envelope
iTRAQ	PQD & TOF/TOF (MS/MS)	I_{\max}	No SQ is computed since signals are taken in centroid mode

¹ I_{\max} is the intensity of the most abundant species and SQ the mean squared deviation between the experimental and the theoretical spectra. SQ is calculated as the peptides are quantified from the mass spectra, as explained in Material and Methods.

Supplementary Table 3. Comparison of significant protein abundance changes in the control vs H₂O₂-treatment experiments

Protein Description ^a	z_{eq}							Remarks
	-10 18O HR AvsB	-6 18O LR AvsB	-2 iTRAQ PQD 114-115 AvsB	0 iTRAQ PQD 116-117 AvsB	2 iTRAQ TOFTOF 114-115 AvsB	6 iTRAQ TOFTOF 116-117 AvsB	10 SILAC HR AvsB	
sp P41816 OYE3_NADPH dehydrogenase 3	-14.2						-12.2	
sp Q04120 TSA2_Peroxiredoxin TSA2	-13.9	-13.4	-8.9	-7.8			-16.8	-14.1
sp P38143 GPX2_Glutathione peroxidase 2	-8.7	-9.5	-3.4				-11.1	-10.4
sp Q96VH4 HBN1_Putative nitroreductase HBN1					-6.2	-4.6		oxidative stress response
sp P02365 RS6_40S ribosomal protein S6			-4.4		-5.7	-5.9		oxidative stress response, regulated by Yap1
sp P22803 TRX2_Thioredoxin-2		-5.4	-4.5	-5.1	-5.3	-5.5	-3.6	oxidative stress response, regulated by Skn7
sp P00431 CCPR_Cytchrome c peroxidase	-5.8	-4.9	-4.1	-3.6	-5.0	-5.8	-5.3	oxidative stress response, regulated by Yap1
sp P54114 ALDH3_Aldehyde dehydrogenase [NAD(P)+] 2	-4.7				-3.6			oxidative stress response
sp Q12068 GRE2_NADPH-dependent methylglyoxal reductase	-4.3	-3.7						oxidative stress response
sp P29509 TRXB1_Thioredoxin reductase 1	-4.2	-3.7	-3.5		-3.8			oxidative stress response, regulated by Yap1, Skn7
sp P33315 TKT2_Transketolase 2	-3.7		-3.8					oxidative stress response, regulated by Msn2/4
sp P28007 GAR1_H/ACA ribonucleoprotein complex subunit 1	-5.3				3.9	3.3	5.5	ribosome biogenesis
sp P40825 SYAC_Alanyl-tRNA synthetase					4.1	3.7		protein translation
sp P07262 DHE4_NADP-specific glutamate dehydrogenase 1					3.9	4.1		synthesis of α-Ketoglutarate
sp P07991 OAT_Orotidine aminotransferase					4.9	4.3		arginine catabolic process
sp P10127 ADH4_Alcohol dehydrogenase 4	4.0	5.2						alcohol biosynthetic process

^aThe list includes proteins that present a statistically significant change in abundance at $FDR_{eq} < 0.05$

<0.05 in two or more experiments. The list does not include two proteins (sp|P47912|LCF4_Long-chain-fatty-acid--CoA ligase 4 and sp|Q03558|OYE2_NADPH dehydrogenase 2) that are significantly increased in the two SILAC control vs H₂O₂-treatment experiments and also in the two SILAC pseudo-null hypothesis experiments. Proteins are sorted according to their z_{eq} values. The magnitudes of the standardized variable are shaded according to the color scales at the top.

Supplementary Table 4. List of proteins and peptides subjected to a label-free Parallel Reaction Monitoring

	Proteins	Peptides	Precursor m/z	Precursor charge
candidate changing proteins	ADH4	DLNVAIYDK	525.77694	2
		FTIISNEEK	540.78222	2
		GIDLINESLVAAYK	753.41413	2
	OYE3	AGNYALHPEVVR	663.35167	2
		AGNYALHPEVVR	442.57020	3
		DTNLFEPIK	538.78476	2
		FFISNPDLVYR	685.85880	2
		FFISNPDLVYR	457.57496	3
		TDEYGGTIENR	627.78348	2
	TRX2	FAEQYSDAAFYK	720.32752	2
		LDVDEVSDVAQK	659.33026	2
		SASEYDSALASGDK	700.81244	2
	TSA2	HITINDLSVGR	612.83839	2
		LVEGFQWTDK	611.80877	2
		TAVVDGIFEEISLEK	825.43526	2
		TAVVDGIFEEISLEK	550.62593	3
	GPX2	GFVILGFP <u>C</u> NQFGK	792.40559	2
		GFVILGFP <u>C</u> NQFGK	528.60615	3
		IDVNGSNADSVNYLK	886.42850	2
		IDVNGSNADSVNYLK	591.28809	3
		PSSLDQEIQSLLSK	772.91195	2
		PSSLDQEIQSLLSK	515.61039	3
control non-changing proteins	COPG	EGPSAPNPSLYVR	693.85424	2
		LFQHQNDPLR	634.33074	2
		LFQHQNDPLR	423.22292	3
		NISTYAITTLK	669.38739	2
	HSP7F	AEEWLYDEGFDSIK	851.38577	2
		GIDIVVNEVSNR	657.85424	2
		IIGLDYHHPDFEQESK	643.31101	3
	TIF31	DSLAYGYTESR	631.28840	2
		ISELEIVSR	523.29804	2
		LQTAALDTTPER	658.34625	2
	YJM3	DITNDITITR	631.83297	2
		IVVGETEEVLR	622.34826	2

The candidate changing proteins were selected from the **Supplementary Table 3**. The control non-changing proteins were selected from the integrated SIL experiments. C – carbamidomethylcysteine (+C₂H₃NO).

Supplementary Table 5. Side-by-side comparison of protein abundance changes in the control vs H₂O₂-treatment SILAC-HR experiment determined by MaxQuant and WSPP.

Protein Description	corrected log ₂ -ratio		Fold Change	
	MaxQuant	WSPP	MaxQuant	WSPP
sp Q99394 TR533_YEAST Transport protein particle 33 kDa subunit OS	-1.81	-3.96	3.50 up	15.56 up
sp Q04120 TS22_YEAST Peroxiredoxin TS2A OS	-2.46	-2.39	5.51 up	5.24 up
sp Q02725 VTC3_YEAST Vacuolar transporter chaperone 3 OS	-1.70	-3.19	3.26 up	9.14 up
sp P38143 GPX2_YEAST Glutathione peroxidase 2 OS	-1.79	-1.67	3.45 up	3.18 up
sp P53332 YGST3_YEAST Uncharacterized protein YGR277C OS	-1.28	-1.83	2.43 up	3.55 up
sp Q04067 EIF3G_YEAST Eukaryotic translation initiation factor 3 subunit G OS	-1.72	0.16	3.29 up	1.11 down
sp P08421 YOS51_YEAST Uncharacterized protein YOR051C OS	-1.17	-0.76	2.25 up	1.70 up
sp P53829 CAF40_YEAST Protein CAF40 OS	-1.09	-0.96	2.13 up	1.94 up
sp P00431 CCPR_YEAST Cytochrome c peroxidase, mitochondrial OS	-0.94	-0.77	1.91 up	1.71 up
sp Q96VH4 HBN1_YEAST Putative nitroreductase HBN1 OS	-0.90	-0.76	1.87 up	1.69 up
sp P07279 RIL8_YEAST 60S ribosomal protein L18 OS	-0.30	-1.61	1.23 up	3.05 up
sp P18239 ADOT2_YEAST ADP-ATP carrier protein 2 OS	-0.75	-0.60	1.68 up	1.52 up
sp P32337 IMB3_YEAST Importin subunit beta-3 OS	-0.94	-0.50	1.92 up	1.41 up
sp P11745 RNA1_YEAST Ran GTPase-activating protein 1 OS	-0.67	-0.63	1.59 up	1.54 up
sp P18900 COQ1_YEAST Hexaprenyl pyrophosphate synthetase, mitochondrial OS	-0.94	-0.68	1.92 up	1.60 up
sp P32466 HXT3_YEAST Low-affinity glucose transporter HXT3 OS	-0.54	-0.82	1.45 up	1.77 up
sp P04046 PUR1_YEAST Amidophosphoribosyltransferase OS	-0.48	-1.01	1.39 up	2.02 up
sp P38627 URA8_YEAST CTP synthase 2 OS	-1.20	-0.03	2.30 up	1.02 up
sp P22803 TRX2_YEAST Thioredoxin-2 OS	-0.54	-0.53	1.45 up	1.45 up
sp P36156 GTO2_YEAST Glutathione S-transferase omega-like 2 OS	-0.71	-0.47	1.63 up	1.39 up
sp P53334 SCW4_YEAST Probable family 17 glucosidase SCW4 OS	-0.64	-0.50	1.55 up	1.41 up
sp P54000 SUB1_YEAST RNA polymerase II transcriptional coactivator SUB1 OS	-0.92	-0.19	1.89 up	1.14 up
sp P25451 PSB3_YEAST Proteasome component PUP3 OS	-0.16	-1.18	1.12 up	2.26 up
sp P38737 ECM29_YEAST Proteasome component ECM29 OS	-0.53	-0.75	1.44 up	1.68 up
sp P23641 MPCP_YEAST Mitochondrial phosphate carrier protein OS	-0.45	-0.51	1.37 up	1.43 up
sp Q07896 NOC3_YEAST Nucleolar complex-associated protein 3 OS	-0.68	-0.60	1.60 up	1.51 up
sp P00445 SDOC_YEAST Superoxide dismutase [Cu-Zn] OS	-0.52	-0.39	1.44 up	1.31 up
sp Q06440 CORO_YEAST Coronin-like protein OS	-0.47	-0.53	1.38 up	1.44 up
sp P32897 TIM23_YEAST Mitochondrial import inner membrane translocase subunit TIM23 OS	-0.55	-0.59	1.46 up	1.50 up
sp P38988 YHM1_YEAST Putative mitochondrial carrier protein YHM1/YSHM1 OS	-0.62	-0.48	1.54 up	1.39 up
sp P33315 TKT2_YEAST Thioredoxin-2 OS	-0.46	-0.54	1.38 up	1.46 up
sp P40582 GST1_YEAST Glutathione S-transferase 1 OS	-0.48	-0.50	1.39 up	1.41 up
sp P05150 OTC_YEAST Ornithine carbamoyltransferase OS	-0.58	-0.38	1.49 up	1.30 up
sp P38113 ADH5_YEAST Alcohol dehydrogenase 5 OS	-0.35	-0.57	1.27 up	1.49 up
sp P36013 MA0N_YEAST NAD-dependent malic enzyme, mitochondrial OS	-0.45	-0.56	1.37 up	1.47 up
sp P12512 ZPS1_YEAST Protein ZPS1 OS	-0.55	-0.29	1.47 up	1.22 up
sp P47912 LCF4_YEAST Long-chain-fatty-acid-CoA ligase 4 OS	-0.70	-0.18	1.62 up	1.13 up
sp P20447 DBP3_YEAST ATP-dependent RNA helicase DBP3 OS	-0.40	-0.55	1.32 up	1.47 up
sp P38998 LYS4_YEAST Saccharopine dehydrogenase [NAD+, L-lysine-forming] OS	-0.47	-0.55	1.38 up	1.47 up
sp Q12503 YF21B_YEAST Transposon Ty2-F/Ty2-G2 Gag-Pol polyprotein OS	-0.18	-0.92	1.14 up	1.89 up
sp Q12496 YO098_YEAST Uncharacterized protein YOL098C OS	-0.54	-0.29	1.46 up	1.22 up
sp P22943 HSP12_YEAST 12 kDa heat shock protein OS	-0.49	-0.33	1.40 up	1.25 up
sp P53278 YGA3_YEAST Uncharacterized protein YGR130C OS	-0.51	-0.23	1.43 up	1.17 up
sp P50945 YNK0_YEAST Uncharacterized protein YNL100W OS	-0.53	-0.52	1.44 up	1.43 up
sp P00044 CYC1_YEAST Cytochrome c iso-1 OS	-0.50	-0.50	1.41 up	1.41 up
sp Q04471 YMO4_YEAST Uncharacterized protein YMR114C OS	-0.53	-0.09	1.44 up	1.07 up
sp P40202 CCS1_YEAST Superoxide dismutase 1 copper chaperone OS	0.49	-0.14	1.40 down	1.10 up
sp P19659 MFD15_YEAST Mediator of RNA polymerase II transcription subunit 15 OS	0.03	0.61	1.02 down	1.53 down
sp P21375 OSM1_YEAST Osmotic growth protein 1 OS	0.06	0.63	1.04 down	1.55 down
sp P53040 TAFF6_YEAST Transcription initiation factor TFIID subunit 6 OS	0.14	0.51	1.10 down	1.43 down
sp P38235 YBQ3_YEAST Uncharacterized protein YBR053C OS	0.08	0.49	1.05 down	1.40 down
sp P53172 SDS23_YEAST Protein SDS23 OS	0.07	0.58	1.05 down	1.50 down
sp P47116 PTK2_YEAST Serine/threonine-protein kinase PTK2/STK2 OS	0.67	-0.07	1.59 down	1.05 up
sp P53920 NM111_YEAST Pro-apoptotic serine protease NMA111 OS	0.14	0.65	1.10 down	1.57 down
sp Q06408 ARO10_YEAST Transaminated amino acid decarboxylase OS	0.39	0.54	1.31 down	1.45 down
sp Q12298 YD061_YEAST Uncharacterized ABC transporter ATP-binding protein YDR061W OS	0.39	0.75	1.10 down	1.68 down
sp Q02805 ROD1_YEAST Protein ROD1 OS	0.39	0.55	1.31 down	1.47 down
sp P14680 YAK1_YEAST Dual specificity protein kinase YAK1 OS	0.47	0.57	1.39 down	1.49 down
sp P38276 YBY7_YEAST UPF0303 protein YBR137W OS	0.69	0.27	1.61 down	1.20 down
sp P11154 PYC1_YEAST Pyruvate carboxylase 1 OS	0.13	0.66	1.09 down	1.58 down
sp P40043 YEP7_YEAST Uncharacterized protein YER067W OS	0.42	0.55	1.34 down	1.47 down
sp O3E841 YNO34_YEAST Uncharacterized protein YNR034W-A OS	0.41	0.50	1.33 down	1.41 down
sp P07262 DHE4_YEAST NADP-specific glutamate dehydrogenase 1 OS	0.47	0.51	1.38 down	1.42 down
sp P38009 PUR92_YEAST Bifunctional purine biosynthesis protein ADE17 OS	0.42	0.54	1.34 down	1.46 down
sp P07991 OAT_YEAST Ornithine aminotransferase OS	0.45	0.52	1.36 down	1.43 down
sp Q12265 KPR5_YEAST Ribose-phosphate pyrophosphokinase 5 OS	0.11	2.32	1.08 down	4.98 down
sp P36059 YKP1_YEAST Uncharacterized protein YKL151C OS	0.58	0.61	1.49 down	1.53 down
sp P32485 KRE6_YEAST Beta-glucan synthesis-associated protein KRE6 OS	0.82	0.51	1.76 down	1.42 down
sp P53878 YNS1_YEAST Uncharacterized protein YNL181W OS	0.45	1.61	1.37 down	3.05 down
sp Q03558 OYE2_YEAST NADPH dehydrogenase 2 OS	0.59	0.63	1.50 down	1.55 down
sp P10127 ADH4_YEAST Alcohol dehydrogenase 4 OS	0.59	0.63	1.51 down	1.55 down
sp P08466 NUC1_YEAST Mitochondrial nucleic OS	1.21	0.07	2.32 down	1.05 down
sp P47127 V51_YEAST Uncharacterized protein YIR080C OS	0.57	1.30	1.48 down	2.46 down
sp Q08987 FDH2_YEAST Formate dehydrogenase 2 OS	0.56	0.92	1.47 down	1.90 down
sp P14747 PP2B2_YEAST Serine/threonine-protein phosphatase 2B catalytic subunit A2 OS	0.36	1.71	1.29 down	3.28 down
sp P53734 DBP6_YEAST ATP-dependent RNA helicase DBP6 OS	-0.13	2.52	1.10 up	5.75 down
sp P46970 NWDS_YEAST Nonsense-mediated mRNA decay protein 5 OS	-0.05	2.40	1.04 up	5.27 down
sp P32657 CHD1_YEAST Chromo domain-containing protein 1 OS	0.25	1.74	1.19 down	3.33 down
sp Q3E754 RS21B_YEAST 40S ribosomal protein S21-B OS	-0.10	2.56	1.07 up	5.88 down
sp P38700 APM2_YEAST Adaptin medium chain homolog APM2 OS	1.16	1.39	2.24 down	2.63 down
sp P00427 COX6_YEAST Cytochrome c oxidase subunit 6, mitochondrial OS	1.34	2.09	2.54 down	4.27 down
sp Q12207 ERR1_YEAST Endonuclease-related protein 1/2 OS	1.76	1.56	3.38 down	2.95 down
sp Q03280 TOM1_YEAST E3 ubiquitin-protein ligase TOM1 OS	0.00	3.93	1.00 down	15.20 down
sp P32597 STH1_YEAST Nuclear protein STH1/NPS1 OS	3.04	0.00	8.23 down	1.00 down
sp Q01519 COX12_YEAST Cytochrome c oxidase subunit 6B OS	1.91	4.02	3.76 down	16.22 down

The peptide pairs from the SILAC-HR A vs B experiment were identified and quantified using MaxQuant. The quantitative data at the peak level obtained by MaxQuant were then reanalyzed using the WSPP model. The table compares the results produced at the protein level by the two algorithms. Only proteins changing in abundance >1.4 fold are shown. The magnitudes of the expression change are shaded according to the color scales at the top. The few discrepancies between the two algorithms in protein abundance changes were mainly due to differences in the algorithms used to eliminate peptide outliers.

Supplementary Table 6. Comparison of protein abundance changes obtained in the control vs H₂O₂-treatment experiments using WSPP and in the SILAC-HR experiment using MaxQuant.

Protein Description ^a	WSPP								MaxQuant			
	corrected log ₂ -ratio				x' _q	Fold change	z' _q	FDR' _q	corrected log ₂ -ratio	Fold change	SILAC-HR	
	18O HR	18O LR	TRAQ PCD 114-115	TRAQ PCD 116-117								
Protein Description ^a	0.8	0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8			
sp Q04120 TSA2 Peroxiredoxin TSA2	3.09	-2.53	-2.05	-1.77	-2.42	-2.32	2.38	8.52	up	-31.23	4.E-211	
sp P38143 GPX2 Glutathione peroxidase	2.13	-1.49	-0.59	-0.70	-1.61	-1.79	4.37	up	-19.33	1.E-80	1.79	
sp P41816 OYE3 NADPH dehydrogenase 3	3.43				-1.94	2.30	10.79	up	-18.03	3.E-70	0.18	
sp P00431 CCP Cytchrome c peroxidases	1.26	0.76	-0.63	-0.61	-0.74	-0.97	-0.77	0.78	2.40	up	-13.63	
sp P22803 TRX2 Thioredoxin-2	0.59	-0.75	-0.55	-0.58	-0.74	-0.47	-0.54	0.62	1.69	up	-12.68	
sp P02365 R56 40S ribosomal protein S	-0.25	0.54	-0.52	-0.35	0.46	-0.62	-0.25	-0.28	0.45	1.53	up	-10.20
sp P29509 TRX6 Thioredoxin reductase	0.88	-0.58	-0.48	-0.29	-0.59	-0.47	-0.40	-0.44	0.48	1.84	up	-8.85
sp P54114 ALDH3 Aldehyde dehydrogenases	1.09	-0.60	-0.74	-0.46	-0.26	-0.54	-0.50	-0.35	0.50	2.02	up	-7.88
sp P05743 R126 60S ribosomal protein	-0.12	-0.44	-0.53	-0.35	-0.36	-0.26	-0.23	-0.24	0.35	1.44	up	-7.16
sp P38616 YGP1 Protein YGP1	0.42	-0.42	-0.22	-0.38	-0.41	-0.59	-0.07	-0.38	-0.36	1.50	up	-7.01
sp P02406 R128 60S ribosomal protein	0.56	-0.16	-0.15	-0.38	-0.43	-0.28	-0.30	-0.31	-0.31	1.48	up	-6.34
sp P53829 CAF40 Protein CAF40	0.05				-1.17	0.84	2.25	up	-6.24	3.E-08	1.09	
sp C96VH4 HBN1 Putative nitroreductas	-0.09	-0.19	-0.37	-0.42	-0.35	-0.43	-0.25	-0.29	0.32	1.34	up	-6.17
sp P05756 R151 40S ribosomal protein	-0.24	0.42	-0.24	-0.48	0.30	-0.33	-0.27	-0.26	0.32	1.39	up	-5.93
sp P05753 R54 40S ribosomal protein S	-0.14	-0.25	-0.29	-0.33	-0.22	-0.47	-0.40	-0.44	0.48	1.51	up	-5.43
sp P09938 R122 Ribonucleotide-diphosph	-0.06	-0.33	-0.18	-0.18	-0.31	-0.61	-0.23	-0.25	0.27	1.52	up	-5.28
sp P05150 OTC Ornithine carbamoyltransferase	-0.17	-0.34	-0.71	-0.40	-0.01	-0.83	-0.70	-0.47	1.77	up	-5.21	
sp P38601 R123 60S ribosomal protein	0.26	-0.43	-0.26	-0.39	-0.28	-0.31	-0.32	1.34	up	-5.18	1.E-05	
sp P14126 R13 60S ribosomal protein L	0.25	-0.26	-0.30	-0.34	-0.32	-0.02	-0.32	-0.23	-0.26	1.27	up	-5.00
sp P33315 TK12 Transketolase 2	1.06		-1.14	-0.64	-0.28	-0.57	0.52	2.20	up	-4.99	2.E-05	
sp Q03102 YMN1 Uncharacterized membrane protein	-0.77	-0.46	-1.17	-0.30	-0.37	-0.54	2.25	up	-4.98	2.E-05		
sp Q03104 MSC1 Meliottis sister chromatid exchange protein	0.63	-0.31	-0.13	-0.12	-0.32	-0.42	-0.23	1.55	up	-4.96	3.E-05	
sp P39938 R526A 40S ribosomal protein	-0.36	-0.22	-0.25	-0.03	-0.43	-1.03	-0.31	-0.47	-0.33	2.04	up	-4.91
sp P42846 KR11 Protein KR11	2.35				-0.30	0.81	5.10	up	-4.87	4.E-05		
sp P05738 R194 60S ribosomal protein	-0.03	-0.28	-0.25	-0.60	-0.03	-0.32	-0.26	-0.30	1.52	up	-4.78	
sp P05754 R58 40S ribosomal protein S	0.00	-0.29	-0.23	-0.43	-0.26	-0.35	-0.29	-0.28	1.35	up	-4.77	
sp P06115 CATT Catalase T	0.25	-0.44	-0.52	-0.37	-0.29	-0.42	-0.36	-0.28	0.35	1.44	up	-4.75
sp Q07653 HBT1 Protein HBT1	0.97				-0.50	-0.44	1.95	up	-4.77	5.E-05		
sp Q12068 GRE2 NADPH-dependent methyl transferase	0.90	-0.60	-0.47	0.01	-0.15	-0.37	1.87	up	-4.61	1.E-04	0.07	
sp P22943 HSP12 12 kDa heat shock protein	-0.03	-0.34	-0.14	-0.36	-0.40	-0.43	-0.29	1.35	up	-4.56	1.E-04	
sp P14796 R140 60S ribosomal protein	-0.35	-0.21	-0.29	-0.37	0.35	-0.31	-0.27	-0.26	1.29	up	-4.52	
sp P22768 ASST Argininosuccinate synthetase	0.30	-0.11	-0.17	-0.43	-0.40	-0.45	-0.25	1.34	up	-4.41	3.E-04	
sp P05030 PM11 Plasma membrane ATPase	0.12	-0.13	-0.11	-0.25	-0.19	-0.23	-0.29	-0.25	1.23	up	-4.35	
sp P47137 Y166 Uncharacterized oxidoreductase	0.64		-0.21	0.17	-0.35	-0.53	-0.41	1.56	up	-4.34	3.E-04	
sp P32827 R523 40S ribosomal protein	-0.18	-0.47	-0.28	-0.23	-0.20	-0.26	-0.29	1.39	up	-4.33	3.E-04	
sp P05740 R1174 60S ribosomal protein	-0.44	-0.16	-0.15	-0.27	-0.17	-0.34	-0.21	-0.22	1.36	up	-4.27	
sp P14065 GCY GCY Protein GCY	0.72	-0.43	-0.53	0.10	-0.15	-0.33	-0.27	-0.26	1.65	up	-4.27	
sp P39078 TCPD T-complex protein 1 subunit 1	-0.24	-0.22	-0.66	-1.06	-0.15	-0.39	-0.36	2.09	up	-4.20	6.E-04	
sp Q03246 RT17 37S ribosomal protein	-0.39				-0.69	0.65	1.61	up	-4.14	7.E-04		
sp P00447 SODM Superoxide dismutase [cytosol]	0.68	-0.23	-0.20	-0.20	-0.16	-0.08	-0.29	-0.26	1.60	up	-4.13	
sp P00445 SODM Superoxide dismutase [mitochondrion]	-0.15	-0.40	-0.24	-0.00	-0.16	-0.54	-0.02	-0.03	1.45	up	-4.04	
sp P47771 ALD2 Aldehyde dehydrogenases	0.34	-0.71	-0.53	-0.33	-0.04	-0.19	-0.14	-0.32	1.63	up	-3.99	
sp P34227 PRX1 Mitochondrial peroxiredoxin	0.69	-0.42	-0.17	-0.20	-0.11	-0.03	-0.15	-0.08	-0.25	1.61	up	-3.96
sp P33442 R534 40S ribosomal protein	-0.11	-0.29	-0.19	-0.20	-0.05	-0.29	-0.20	-0.14	-0.19	1.23	up	-3.87
sp P23327 IMB3 Importin subunit beta-1	0.01				-0.59	0.51	1.51	up	-3.84	2.E-03		
sp P23248 R538 40S ribosomal protein	0.60	0.04	-0.35	-0.03	-0.29	-0.34	0.25	1.52	up	-3.83	2.E-03	
sp P39741 R135 60S ribosomal protein	0.20				-0.87	0.69	1.83	up	-3.87	3.E-03		
sp P38962 TVP23 Golgi apparatus membrane protein	-0.13	-0.15	-0.55	-0.63	-0.48	1.55	up	-3.86	4.E-03			
sp P36148 GPT2 Glycerol-3-phosphate O-acyltransferase	0.36				-0.13	-0.55	-0.55	0.52	1.47	up	-3.83	
sp Q07896 NOC3 Nucleolar complex-assoc	0.31	-0.22	-0.14	-0.19	-0.17	-0.34	0.16	0.20	-0.15	1.26	up	-3.82
sp P13803 AH1 Peroxiredoxin type-2 O	0.31	-0.22	-0.14	-0.19	-0.17	-0.34	0.16	0.20	-0.15	1.26	up	-3.82
sp P49334 TOM22 Mitochondrial import protein	-0.18	-0.10	-0.15	-0.15	-0.62	-0.62	1.54	up	-3.80	5.E-03		
sp Q04432 HSP23 Probable chaperone protein	0.61	-0.29	-0.12	-0.14	-0.32	-0.42	-0.24	1.53	up	-3.79	7.E-04	
sp P04449 R126 60S ribosomal protein	-0.15	-0.22	-0.19	-0.23	-0.28	-0.24	-0.22	1.21	up	-3.77	5.E-03	
sp P16547 R545 Mitochondrial outer membrane protein	-0.22	-0.15	-0.46	-0.15	0.08	1.34	-0.32	-0.27	2.53	up	-3.71	
sp P02753 R121A 60S ribosomal protein	0.22	-0.15	-0.18	-0.33	-0.28	-0.33	-0.22	1.26	up	-3.68	7.E-03	
sp P49631 R123 60S ribosomal protein	-0.12	-0.29	-0.02	-0.31	-0.37	-0.35	-0.12	-0.25	1.29	up	-3.66	
sp P53315 SOL4 Probable 6-phosphoglucan branching enzyme	1.07	-0.22	-0.82	-0.82	-0.24	-0.29	-0.29	0.35	2.10	up	-3.62	
sp P06367 R514A 40S ribosomal protein	-0.18	-0.28	-0.04	-0.05	-0.33	-0.11	-0.12	-0.19	-0.17	1.26	up	-3.59
sp P04451 R123 60S ribosomal protein	-0.25	-0.17	-0.42	0.17	0.04	-0.76	-0.21	-0.23	-0.21	1.69	up	-3.59
sp P18544 ARG6 Acetylornithine aminotransferase	-0.04				-0.06	-0.80	-0.36	-0.48	0.35	1.74	up	-3.56
sp Q06405 ATP5 synthase subunit f	0.60	-0.43	-0.33	-0.30	-0.31	-0.19	-0.30	1.52	up	-3.54	1.E-02	
sp P53858 BN1 Protein BN14	0.47		-1.13	-0.53	-0.55	-0.03	-0.24	-0.42	-0.23	0.77	up	-3.52
sp P36139 PET10 Protein PET10	0.02	-0.50	-0.04	-0.39	0.26	-0.76	-0.09	-0.17	0.20	1.70	up	-3.51
sp P32288 GLNA Glutamine synthetase O	0.74		-0.16	-0.11	-0.21	-0.16	-0.28	-0.25	1.63	up	-3.49	
sp Q04304 YMR0 UPF0659 protein YMR090	-0.20	-0.14	-0.12	-0.14	-0.14	-0.35	-0.21	-0.20	-0.17	1.28	up	-3.48
sp P49723 R124 Ribonucleoside-diphosphorylase	-0.14	-0.04	-0.25	-0.01	-0.35	-0.31	-0.11	-0.08	-0.16	1.27	up	-3.47
sp P41805 R110 60S ribosomal protein	-0.01	0.36	0.27	0.84	-0.78	-0.78	-0.35	1.86	up	-3.46	8.E-03	
sp P38113 ADH5 Alcohol dehydrogenase	-0.17	0.02			-0.47	-0.29	-0.28	1.38	up	-3.45	2.E-02	
sp P02217 TRX1 Thioredoxin-1	0.21	0.04	-0.26	-0.24	0.40	-0.31	0.06	0.07	-0.16	1.32	up	-3.44
sp P40515 F151 Mitochondria fission 1	0.27				-0.65	0.49	-0.43	1.40	up	-3.43	3.E-02	
sp P47122 NP43 GTPase NP43	-0.09				-0.65	-0.48	1.57	up	-3.42	3.E-02		
sp Q07163 TYL1B Transposon Tyh3 Gag-P	0.03	0.02	-0.08	-0.21	-0.22	0.66	-0.21	-0.15	1.58	up	-3.42	
sp P53221 R126 60S ribosomal protein	-0.15	-0.23	-0.42	0.02	-0.36	-0.29	-0.25	1.28	up	-3.41	3.E-02	
sp P54861 DNM1 Dynamin-related protein	1.06				-0.29	-0.31	-0.31	1.34	up	-3.40	4.E-02	
sp P48361 ASK10 Protein ASK10	0.63				-0.36	0.56	2.04	up	-3.39	4.E-02		
sp P12000 TM46 Translation machinery	0.63				-0.34	-0.40	1.54	up	-3.38	4.E-02	0.28	

Supplementary Table 6 (continues)

Protein Description ^a	WSPP												MaxQuant				
	corrected log2-ratio						x'_q	Fold change	z'_q	FDR'_q	SILAC-HR		corrected log2-ratio	Fold change			
	18O HR	18O LR	ITRAQ PD114-115	ITRAQ PD116-117	ITRAQ TOF114-115	ITRAQ TOF116-117					SILAC HR	SILAC LR					
sp P29952 MPI Mannose-6-phosphate iso	0.16	0.14	0.14	0.13	0.25	0.13	0.47	0.20	0.20	0.17	1.19	down	3.01	5.E-02	0.13	1.09 down	
sp P23783 YC21A Transposon Ty2-C Gag											3.03			5.E-02	0.22	1.17 down	
sp P32775 [GLGB 1,4-alpha-glucan-branc	0.30		0.55	-0.14			0.34	0.19	0.28	1.46	down		3.04	4.E-02	0.15		
sp P40510 [SER33 D-3-phosphoglycerate	0.45	0.32	0.17	-0.01			0.19			0.23	1.36	down	3.08	4.E-02	0.04	1.03 down	
sp P15019 TAL1 Transaldolase	0.02	0.01	0.04	0.21	0.15	0.24	0.24	0.23	0.13	1.18	down	3.08		4.E-02	0.15	1.11 down	
sp P11154 PYC1 Pyruvate carboxylase 1	0.35	0.20	0.00	0.07			0.25	0.41	0.17	1.33	down	3.09		4.E-02	0.13	1.09 down	
sp P00925 ENO2 Enolase 2	-0.04	0.13	0.09	0.12	0.05	0.22	0.21	0.20	0.12	1.17	down	3.11		4.E-02	0.12	1.09 down	
sp P38934 [BFR1 Nuclear segregation pr	0.07	0.31	0.08	0.21	0.22	0.14	0.15	0.19	0.16	1.26	down	3.11		4.E-02	0.11	1.08 down	
sp Q08224 THI20 Phosphomethylpyrimidi	0.49	0.14					0.38	0.36	0.32	1.40	down	3.20		3.E-02	0.31	1.24 down	
sp Q36841 YNO34 Uncharacterized prote	0.10	0.18	-0.24	0.40	0.39	0.01	0.48			0.23	1.40	down	3.22	3.E-02	0.41	1.33 down	
sp P33754 SEC66 Translocation protein							-0.26	1.16	0.44	2.23	down	3.25		3.E-02			
sp P37291 [GLYC Serine hydroxymethyltr	0.15	0.15	0.09	0.12	0.15	0.16	0.32	0.20	0.16	1.25	down	3.28		2.E-02	0.20	1.15 down	
sp P04807 HXXK8 Hexokinase-2	0.01	0.13	0.07	0.21	0.41	0.38	0.24	0.20	0.18	1.33	down	3.29		2.E-02	0.15	1.11 down	
sp P07149 [FAS1 Fatty acid synthase su	0.08	0.10	0.16	0.08	0.21	0.19	0.13	0.16	0.11	1.16	down	3.29		2.E-02	0.12	1.09 down	
sp P00560 PGK Phosphoglycerate kinase	-0.12	0.11	0.15	0.12	0.05	0.20	0.12	0.16	0.11	1.15	down	3.32		2.E-02	0.02	1.01 down	
sp P39003 HXT6 High-affinity hexose t	0.14	0.13	-0.11				0.74			0.30							
sp P40443 YEP7 Uncharacterized protei	0.11	0.10	-0.22	0.12			0.53	0.45	0.28	1.44	down	3.37		2.E-02	0.42	1.34 down	
sp P11632 NHP6A Non-histone chromosom			0.44	0.21			0.50			0.43	1.42	down	3.46		2.E-02		
sp P07283 PMMP Phosphomannomutase	0.07	0.22	0.20	0.24	0.00	-0.08	0.30	0.31	0.18	1.24	down	3.52		1.E-02	0.23	1.17 down	
sp P46889 ATG27 Autophagy-related pro	-0.19				-0.01	0.69	0.72		0.46		1.85	down	3.54	1.E-02	0.17	1.13 down	
sp P21524 RIR1 Ribonucleoside-diphosp	0.54	0.19	0.24	0.20			0.41			0.33	1.45	down	3.55	1.E-02	0.30	1.23 down	
sp Q03497 STE20 Serine/threonine-prot			0.76				0.37			0.51	1.62	down	3.59	1.E-02	0.02	1.01 down	
sp P19358 METK2 S-adenosylmethionine	0.26	0.15	0.15	0.18	0.19	0.09	0.31	0.26	0.20	1.26	down	3.62		1.E-02	0.27	1.20 down	
sp P46367 ALDH4 Potassium-activated a	0.05	0.06	0.03	0.06	0.25	0.40	0.44	0.38	0.18	1.36	down	3.68		8.E-03	0.35	1.27 down	
sp P47120 DOH4 Deoxyphosphovaleric acid	0.27	0.22	0.49	1.29	-0.28	0.28	0.76	0.37	0.29	2.29	down	3.70		8.E-03	0.17	1.12 down	
sp P04046 PUR1 Amidophosphoribosyltria	-0.01				0.29	0.49	1.19	0.64	0.17	1.49	down	4.62		2.E-04	0.48	1.39 up	
sp P54115 ALDH6 Magnesium-activated a	0.13	0.13	0.17	0.23	0.22	0.24	0.17	0.08	0.17	1.18	down	3.80		6.E-03	0.08	1.06 down	
sp P01095 IPB2 Protease B inhibitors	0.24	-0.02	0.24	0.51	0.24	0.21	0.26	0.29	0.26	1.42	down	3.83		6.E-03	0.20	1.15 down	
sp P006408 ARO10 Transaminated amino a	0.52						0.59			0.57	1.50	down	3.91	4.E-03	0.39	1.31 down	
sp P00812 ARG1 Arginase	0.55		-0.07	0.32			0.77			0.51	1.70	down	4.08	2.E-03	0.09	1.06 down	
sp P39876 LDL3 3-lactate dehydrogenas	0.14	0.20	0.30	0.31	0.22	0.05	0.36	0.39	0.26	1.31	down	4.19		1.E-03	0.22	1.17 down	
sp Q04371 YMR7 UPP0364 protein YMR027	0.53		1.35	-0.17			0.04	0.57	0.39	2.56	down	4.50		4.E-04	0.01	1.01 down	
sp Q08911 FDH1 Formate dehydrogenase	0.24		0.40	0.52			0.57			0.47							
sp P43616 CPGL Glutamate carboxypepti	0.02	0.38	0.23	0.35	-0.21	0.99	0.31	0.36	0.30	1.99	down	4.88		7.E-05	0.22	1.16 down	
sp P00924 ENO1 Enolase 1	-0.04	0.13	0.17	0.27	0.13	0.25	0.30	0.32	0.19	1.25	down	5.22		1.E-05	0.25	1.19 down	
sp P07991 OAT Ornithine aminotransfer	0.07	0.10	0.00	-0.03	0.42	0.83	0.55	0.64	0.36	2.68	down	5.25		1.E-05	0.45	1.36 down	
sp P16474 GRP78 78 kDa glucose-regula	0.06	0.27	0.10	0.27	0.29	0.22	0.30	0.26	0.22	1.23	down	5.31		9.E-06	0.24	1.18 down	
sp P38009 PUR92 Bifunctional purine b	0.56	0.15	0.53	0.21			0.50	0.46	0.38	1.47	down	5.36		8.E-06	0.42	1.34 down	
sp P17967 PDI Protein disulfide-isom	0.08	0.16	0.29	0.25	0.23	0.73	0.26	0.25	0.26	1.16	down	5.49		4.E-06	0.19	1.14 down	
sp P03529 SCS7 Inositolphosphorylcera			0.69	-0.14	0.36		0.88			0.69							
sp P11986 NOL1 Inositol-3-phosphate s	0.27	0.24	0.27	0.15	0.35	0.12				0.25	1.27	down	5.79	1.E-06	0.45	1.36 down	
sp P07262 DHE4 NADP-specific glutamat	0.11	0.16	0.31	0.28	0.13	0.13	0.56	0.55	0.25	1.48	down	6.19		1.E-07	0.47	1.38 down	
sp P40893 FKS2 1,3-beta-glucan synth			1.55	-0.14			0.25			1.04							
sp P40825 SYAC1 Alanyl-tRNA synthetase	0.41	0.32	0.23	0.43	0.40	0.65	0.46	0.53	0.40	1.57	down	8.16		0.E+00	0.36	1.28 down	
sp P10127 ADH4 Alcohol dehydrogenase	0.39	0.42	0.38	0.54	0.56	0.47	0.69	0.65	0.51	1.61	down	10.99		0.E+00	0.59	1.51 down	

^aProteins changing in abundance at $FDR'_q < 0.05$ after data integration and detected in two or

more experiments. The list does not include the protein sp|P47912|LCF4_YEAST Long-chain-

fatty-acid--CoA ligase 4, which is significantly increased after data integration, but is also

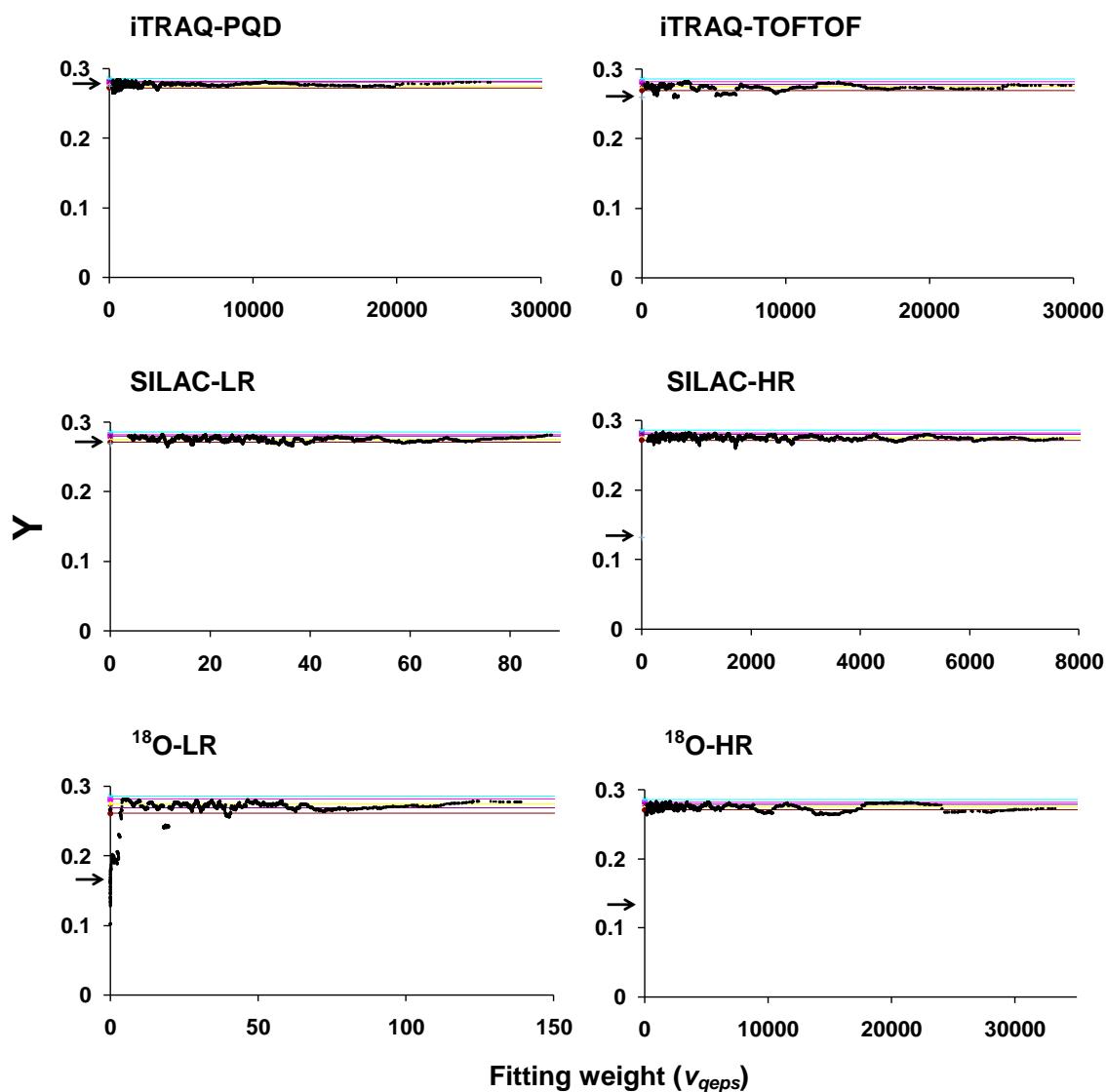
increased in the SILAC-HR and SILAC-LR null hypothesis experiments. Proteins are sorted by

z'_q . The magnitudes of the expression change and of the standardized variable are shaded

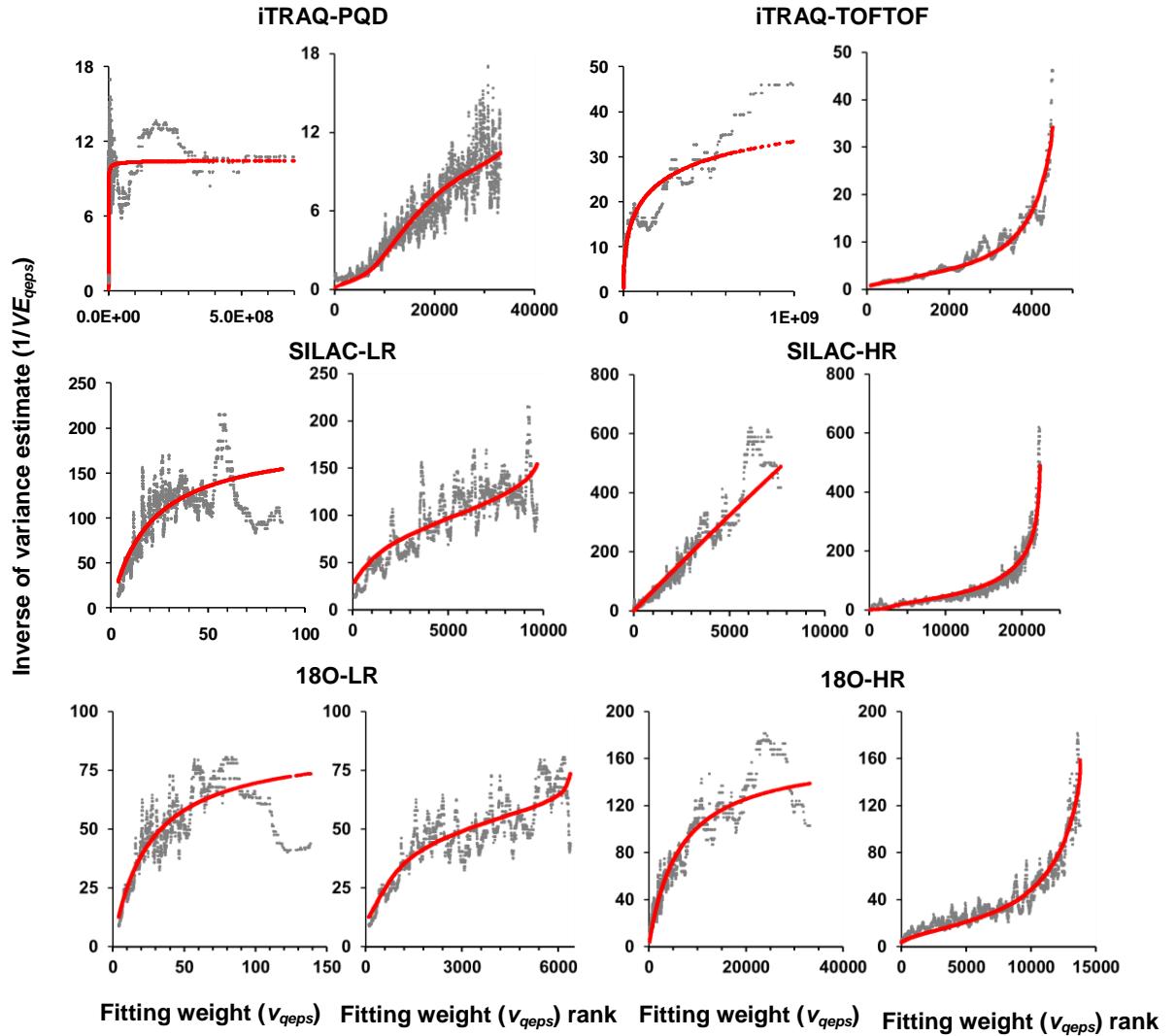
according to the color scales at the top. Proteins in the two columns at the right correspond to the

SILAC-HR A vs B experiment and were identified and quantified using MaxQuant.

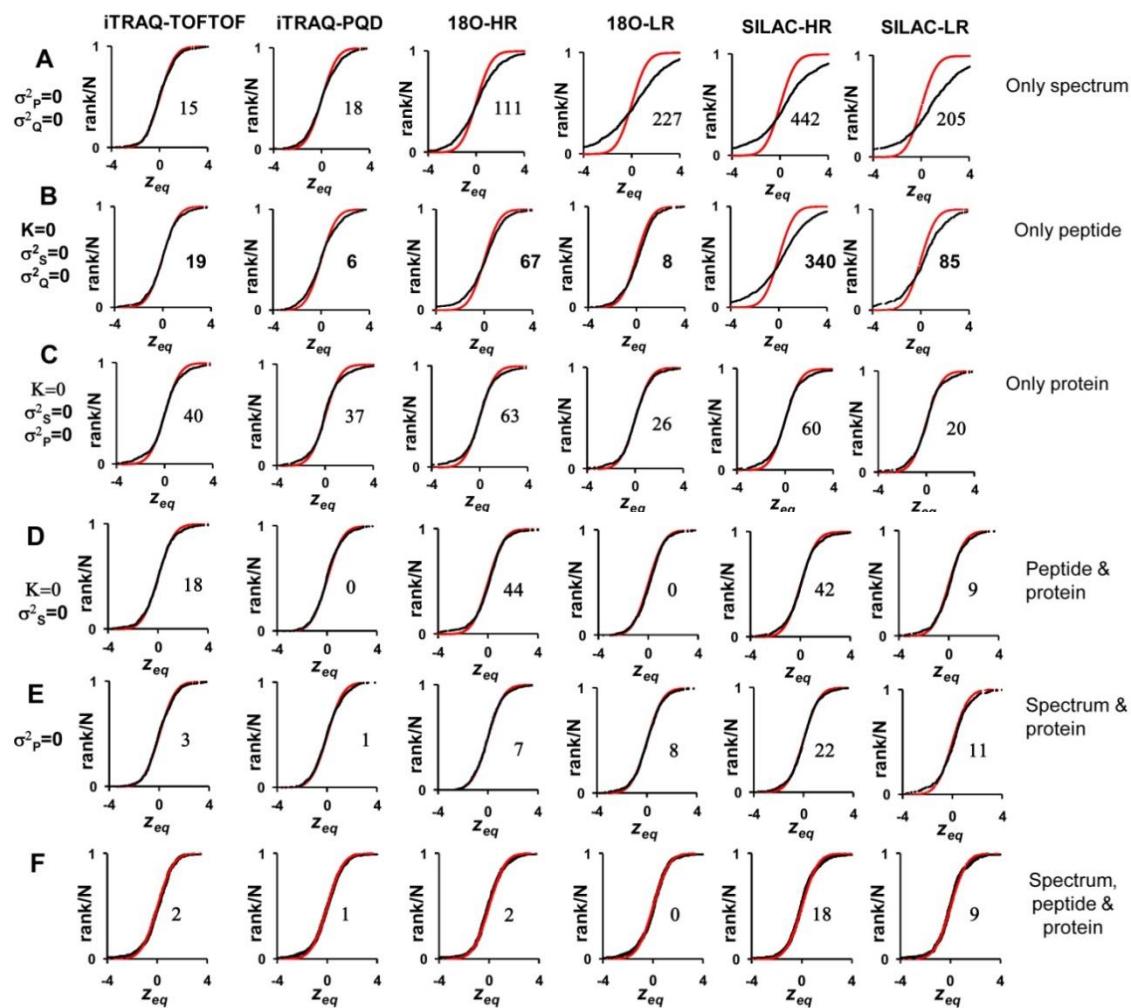
Supplementary Figure 1: Analysis of local normality. The plots represent the results of D'Agostino normality test¹⁴ for the different SIL experiments applied in sliding windows of 200 spectra ranked according to fitting weight; the values taken by the Y parameter are plotted as a function of the median fitting weight in each window. For comparison, the *arrow* on the *y* axis indicates the Y value obtained by applying the test to the entire population.



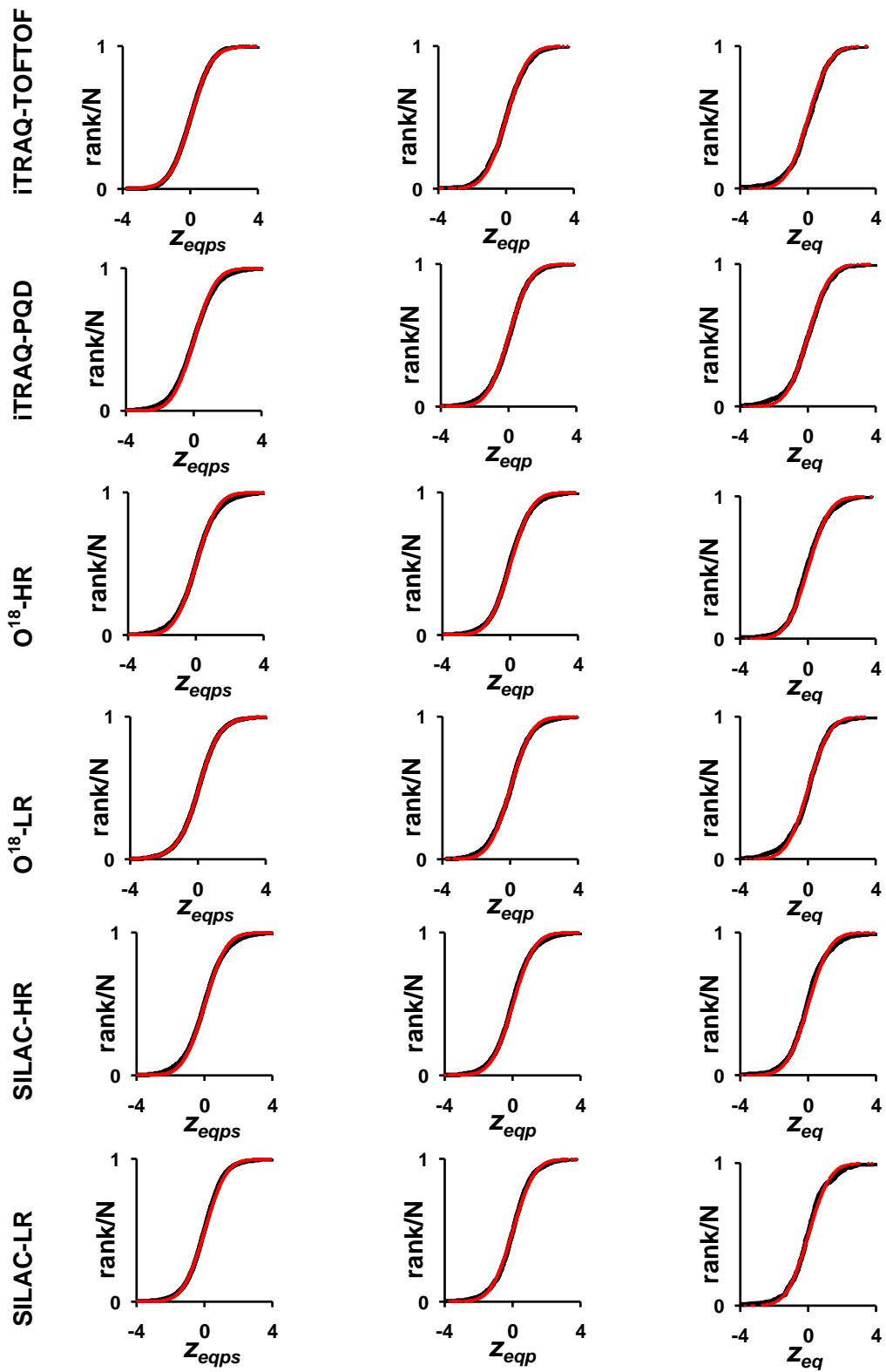
Supplementary Figure 2: Estimation of the weight constant (k_e) for the different SIL experiments. The red points represent the theoretical values calculated using the estimated k_e and σ_{Se}^2 values. For the sake of clarity, two kinds of plots, using the fitting weight (v_{qeps}) (left panels) or the fitting weight rank (right panels) in the x -axis, are presented. Details are described in Section 2.2.



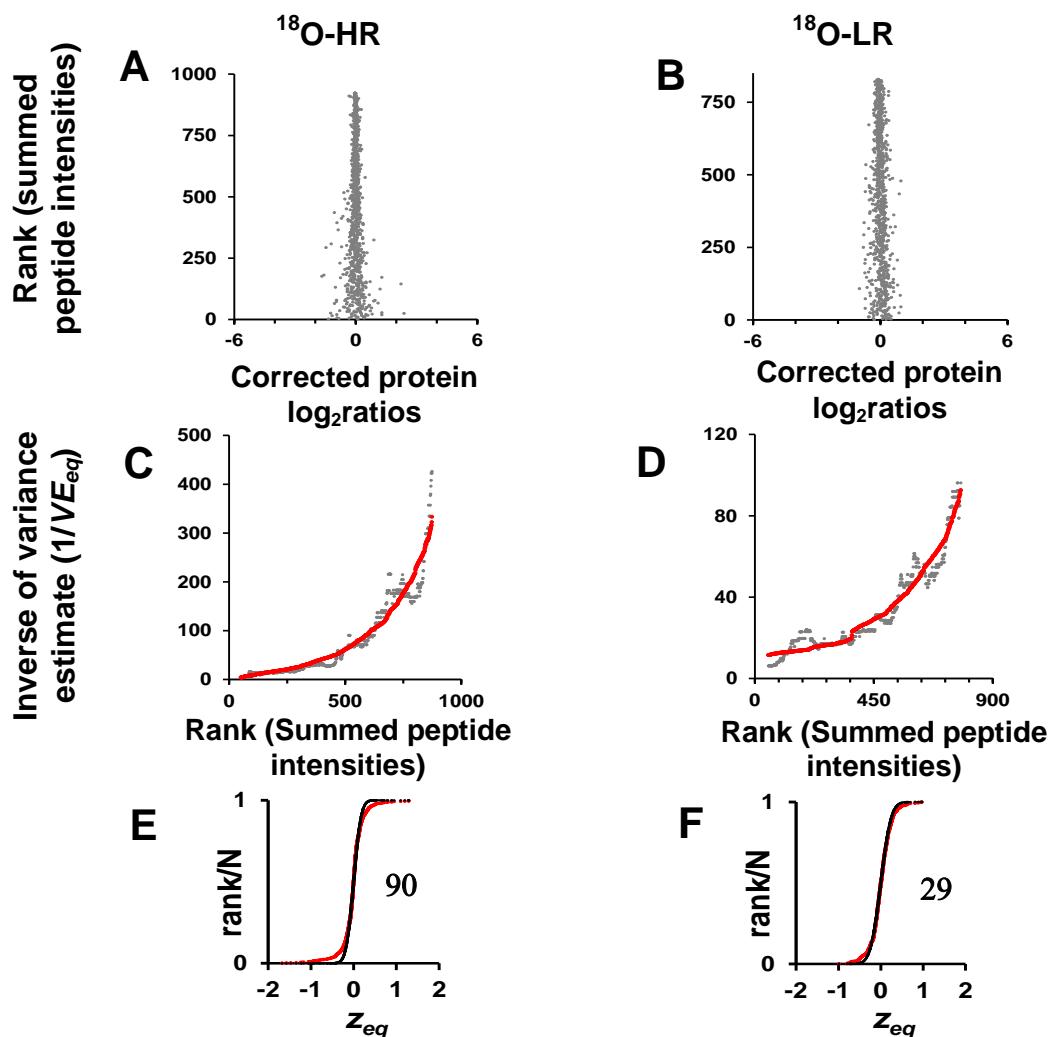
Supplementary Figure 3: Evidence that three separate sources of variance at the spectrum, peptide and protein level are needed to give a correct description of the different NH experiments. The cumulative distributions of the standardized variable at the protein level are shown; variance calculation and statistical analysis were performed considering only spectrum (A), only peptide (B), only protein (C), only peptide and protein (D), only spectrum and protein (E) or spectrum, peptide and protein (F) levels; red lines are drawn according to the theoretical normal distribution with zero mean and unit variance. The numbers indicate the amount of protein abundance changes detected as statistically significant at a 5% FDR as explained in section 2.5.



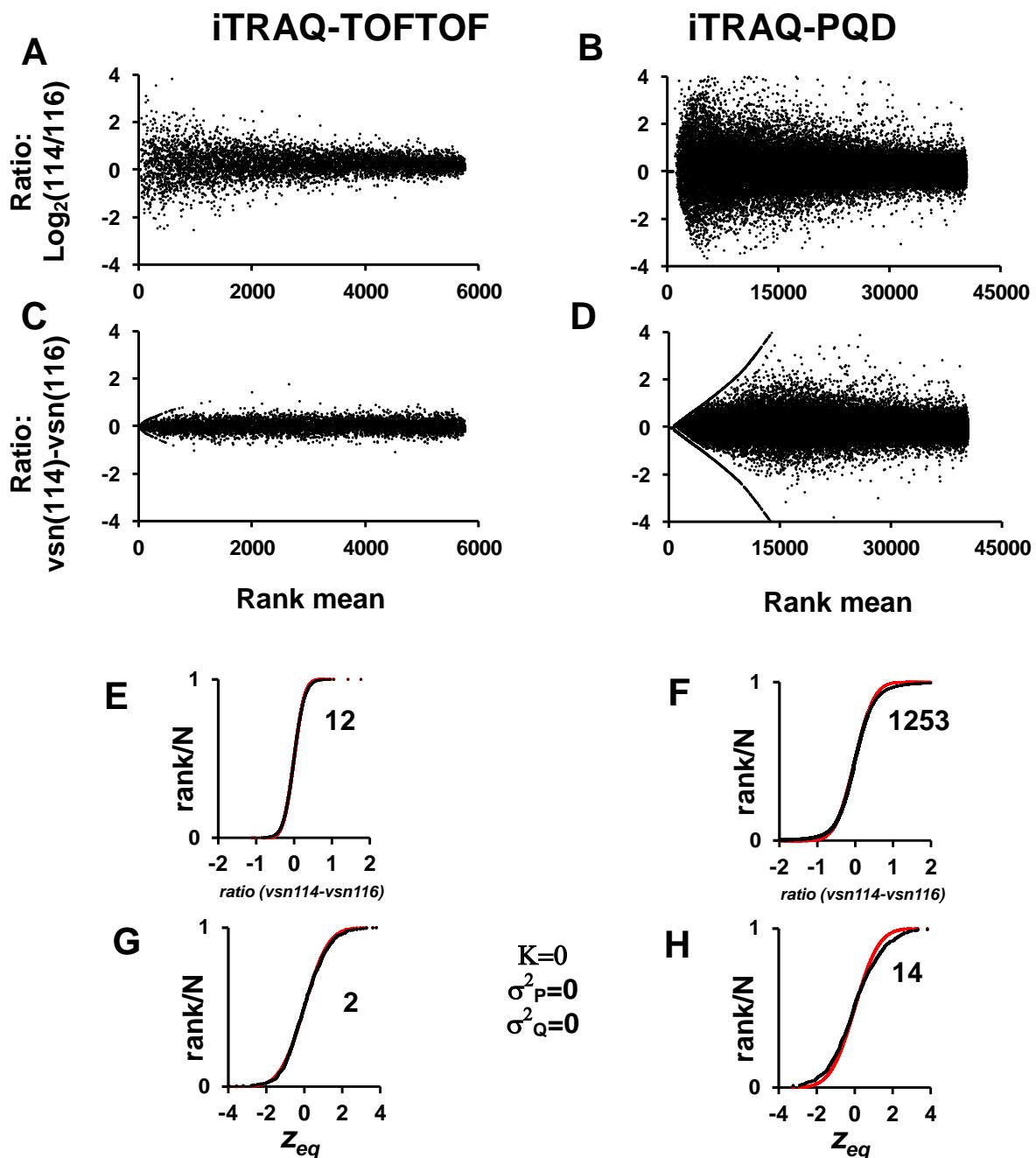
Supplementary Figure 4: Analysis at the spectrum, peptide and protein levels of the control vs H₂O₂-treatment experiments



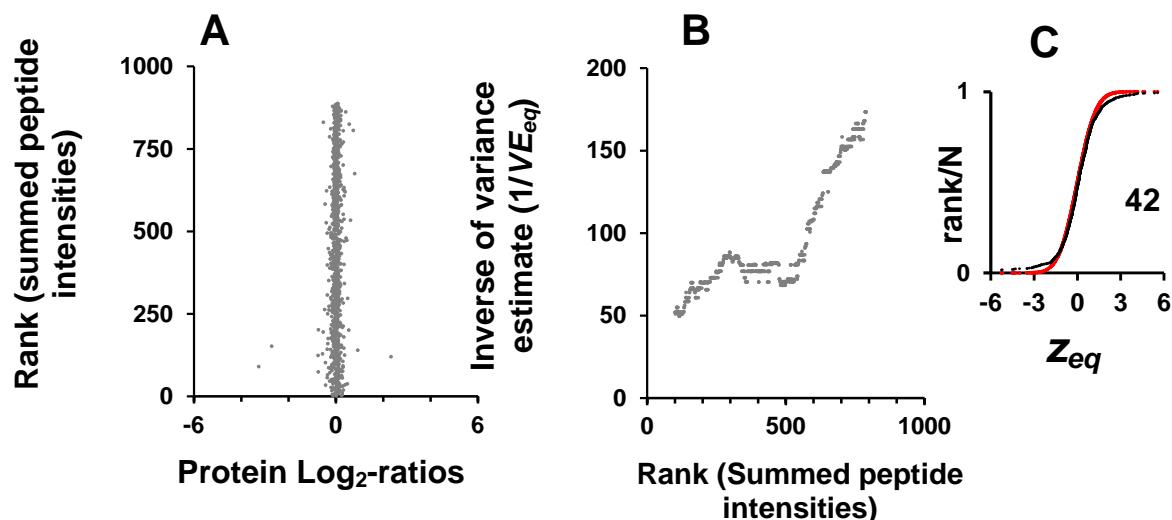
Supplementary Figure 5: Analysis of the high-resolution (A,C,E) and low-resolution (B,D,F) ^{18}O NH experiments using UNiquant algorithm. (A,B) The rank of summed peptide intensities is plotted against the corrected protein log₂-ratios. (C,D) The inverse of mean squared deviation applied in sliding windows of 200 points is plotted against the rank of summed peptide intensities. (E,F) These plots show the cumulative distribution of the standardized variable at the protein level; the red line is drawn according to the theoretical normal distribution with zero mean and unit variance. The numbers indicate the amount of protein abundance changes (false protein abundance changes) detected as statistically significant at a 5% FDR.



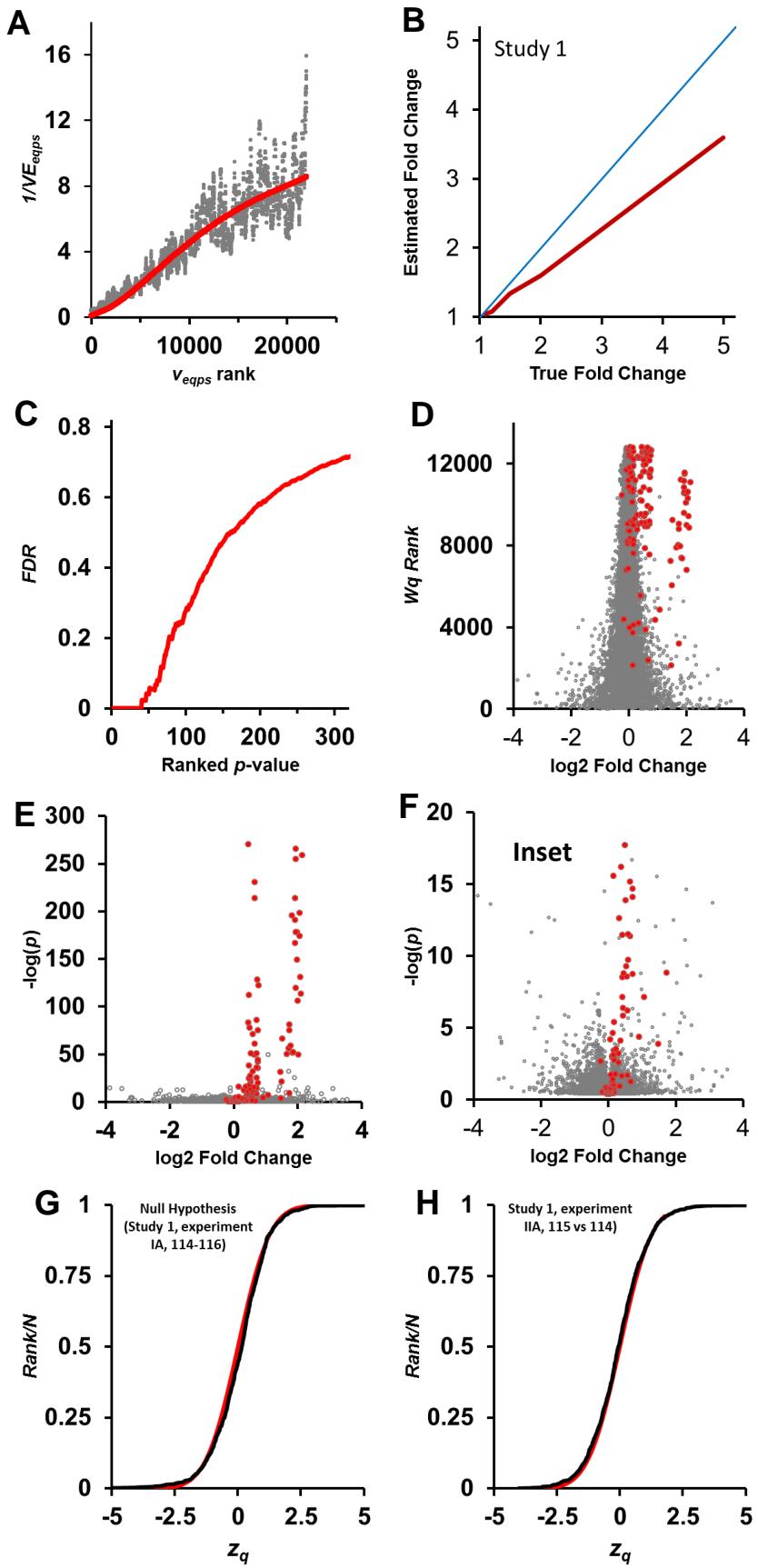
Supplementary Figure 6. Analysis of the TOFTOF and PQD iTRAQ NH experiments using the variance-stabilizing normalization (VSN) strategy. The upper graphs show the distribution of log2-ratios (A,B) and VSN-ratios (C,D) as a function of the rank of the average reporter ion signal. The lower graphs show the cumulative distributions of the VSN-transformed variable (E,F) and of the standardized protein variable (G,H), calculated by integrating the VSN-corrected spectrum data to the peptide and protein levels assuming that the local spectrum variance is constant and the peptide and protein variances are zero. The red line is drawn according to the theoretical normal distribution. The numbers indicate the amount of spectrum outliers (E,F) and of protein abundance changes (false protein abundance changes) (G,H) detected as statistically significant at a 5% FDR.



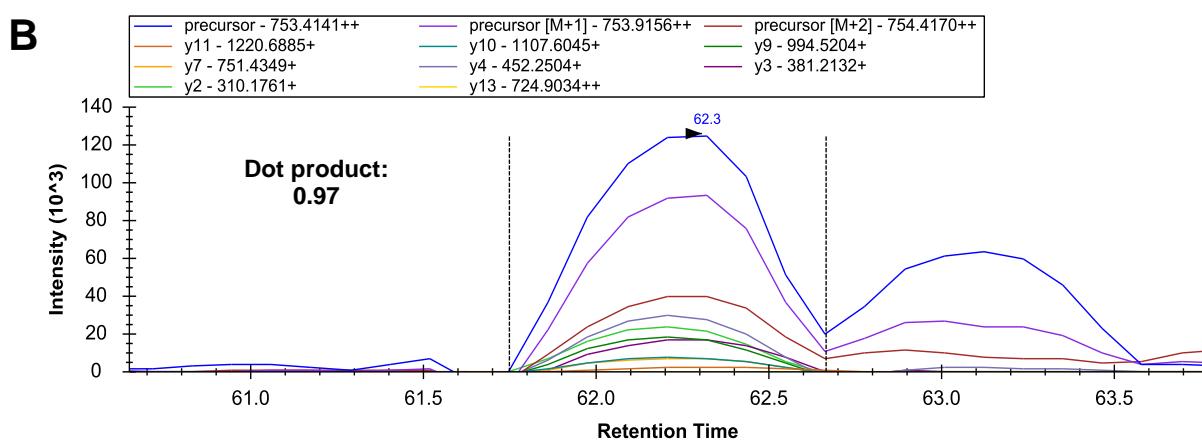
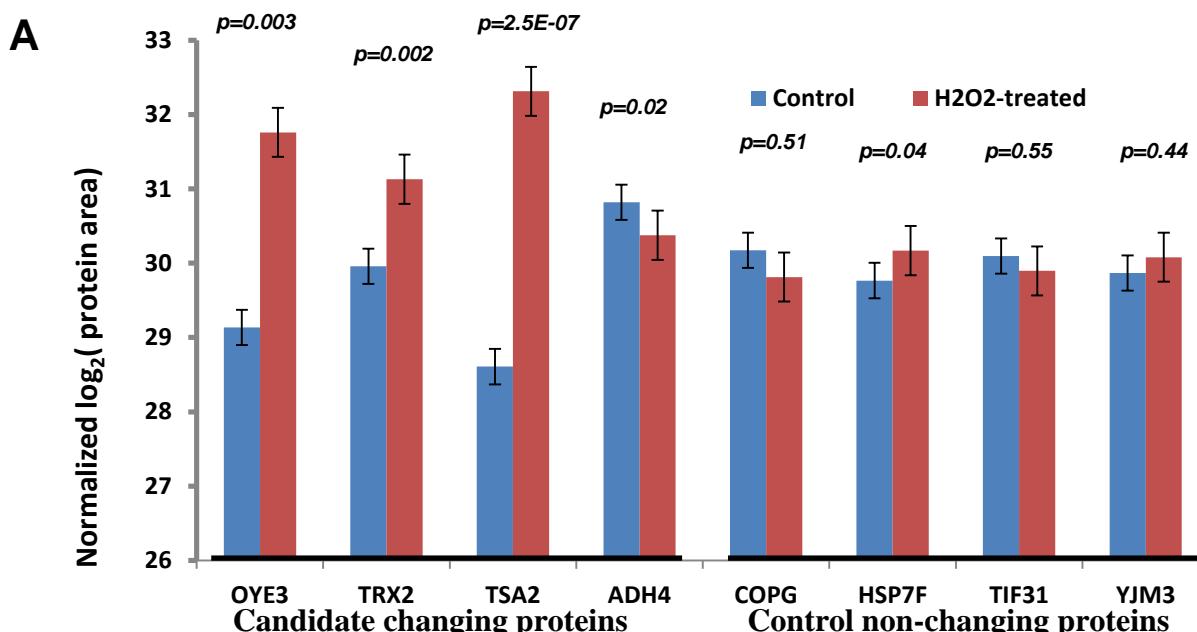
Supplementary Figure 7. Analysis of the SILAC -HR, pseudo-NH experiment using MaxQuant. The data from this experiment were analyzed using MaxQuant. (A) Protein log₂-ratios are plotted against the rank of summed peptide intensities. (B) The inverse of mean squared deviation applied in sliding windows of 200 points was plotted against the rank of summed peptide intensities. (C) The plot shows the cumulative distribution of the standardized variable at the protein level, which was calculated using the estimated local variance; the red line is drawn according to the theoretical normal distribution with zero mean and unit variance. The number indicates the protein abundance changes detected as statistically significant at a 5% FDR. MaxQuant allowed the quantification of 888 proteins and detected 42 significant abundance changes (C).

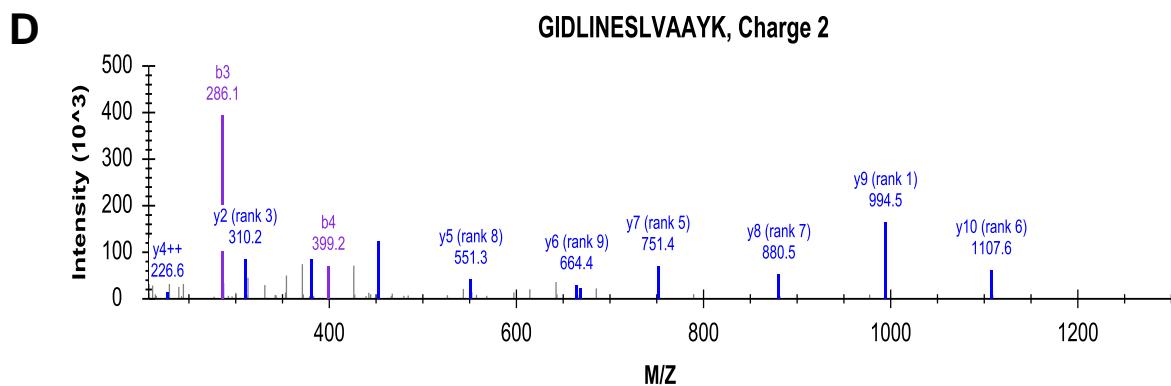
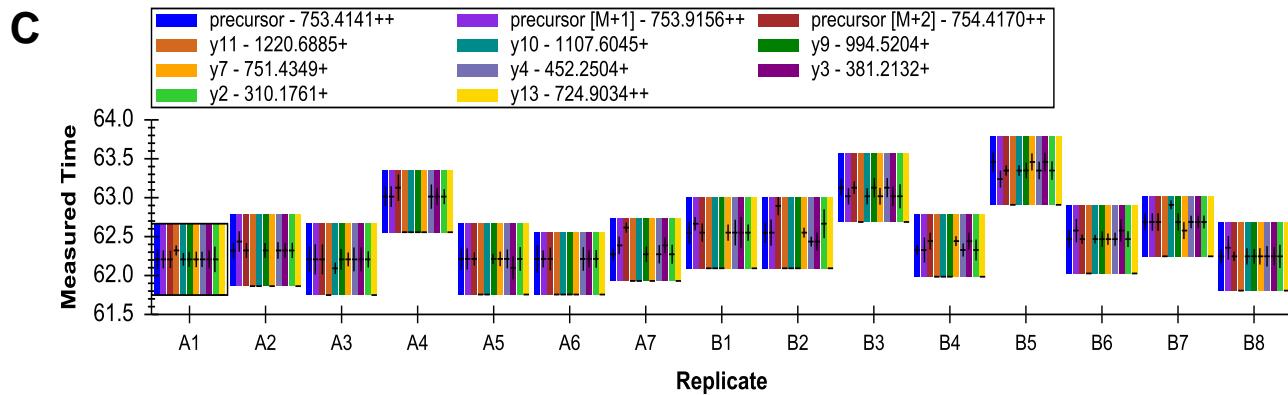


Supplementary Figure 8: Performance of the WSSP model in the analysis of spike-in mixtures of known protein concentrations in a complex background. iTRAQ 4-plex data obtained on a LTQ-Orbitrap instrument (Study 1) were taken from a previously published work¹³ and analyzed with the WSPP model. The dataset was obtained from the analysis of a constant 1:1:1:1 yeast background with two mixture types of 16 spiked-in non-yeast proteins: mixture A (1:,5:1:5) and mixture B (1,1:1:2:1,2); two relative quantities of each of the 16 spike-in proteins in a particular mixture (I and II); and two technical replicates (see Tables I and II from ¹³). (A) Estimation of the weight constant (k_e). Light grey points represent the real data, whereas the red line is drawn with the estimated k_e and σ_{se}^2 values. (B) Accuracy of quantifications. The graph represents the estimated fold changes versus the true fold changes for the 16 spike-in proteins. The blue line represents the line of identity when there is no bias. (C) Dependence of the False Discovery Rate (FDR) versus ranked *p-value*. All proteins were ranked by their *p-value* and the FDR was calculated as the percentage of yeast proteins in the list of proteins showing significant abundance change according to a *p-value* threshold. (D) Weight distributions of protein quantifications. Gray points indicate yeast proteins, whereas red points represent the spike-in proteins. Positive values correspond to an increase and negative values to a decrease in protein abundance. (E,F) Volcano plot obtained from the analysis of all possible pair-wise comparisons of relative abundance, and inset at higher p-values. The gray dots correspond to yeast proteins, whereas the red dots are from the spike-in proteins. (G,H) Cumulative distributions of the standardized variable at the protein level for the Null Hypothesis (1:1) experiment IA and for the 1:1,1 IIA experiment. Red lines are drawn according to the theoretical normal distribution with zero mean and unit variance.

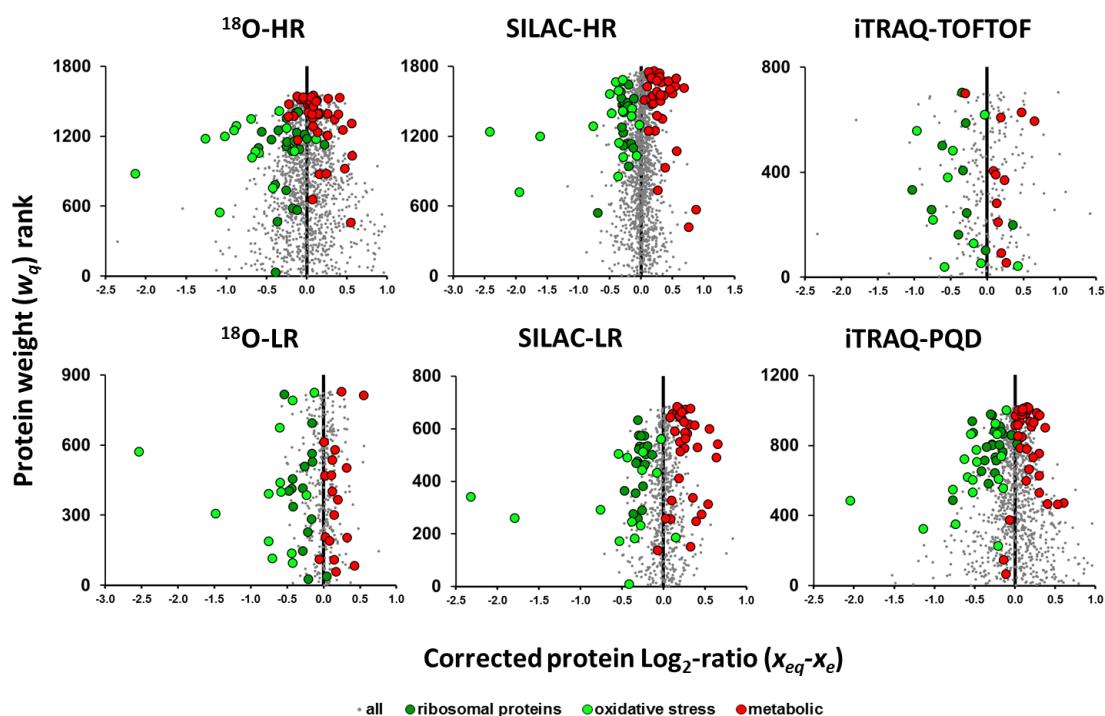


Supplementary Figure 9: Verification of some protein abundance changes in the control vs treatment experiments by label-free Parallel Reaction Monitoring (A). Extracted ion chromatogram for the precursor and fragment ions of the proteotypic peptide GIDLINESLVAAYK of ADH4 in a total yeast digest (B). The retention time at which this peptide was detected in each replicate (C), and the MS2 spectrum used to construct the library (D) are shown. GPX2 peptides were not detected in the unfractionated proteome.





Supplementary Figure 10: Quantitative distributions of ribosomal, oxidative stress and metabolic proteins from Table 2 in the control vs H₂O₂-treatment experiments. The weight distributions of protein quantifications around the grand mean obtained by each of the indicated SIL methods are shown. Only the proteins that belong to any of these three functional categories and that show a significant abundance change after the integration of all the experiments, as indicated in Table 2, are highlighted. In these plots negative values indicate an increase in protein abundance (toward the left) and positive values a decrease (toward the right). Ribosomal proteins are depicted in dark green, oxidative stress proteins in light green and metabolic proteins in red. These plots illustrate the consistency in the quantitative behavior of these proteins in all experiments.



5. References

1. Lopez-Ferrer, D.; Ramos-Fernandez, A.; Martinez-Bartolome, S.; Garcia-Ruiz, P.; Vazquez, J., Quantitative proteomics using $^{16}\text{O}/^{18}\text{O}$ labeling and linear ion trap mass spectrometry. *Proteomics* **2006**, *6 Suppl 1*, S4-11.
2. Martinez-Bartolome, S.; Navarro, P.; Martin-Maroto, F.; Lopez-Ferrer, D.; Ramos-Fernandez, A.; Villar, M.; Garcia-Ruiz, J. P.; Vazquez, J., Properties of average score distributions of SEQUEST: the probability ratio method. *Mol Cell Proteomics* **2008**, *7* (6), 1135-45.
3. Ramos-Fernandez, A.; Lopez-Ferrer, D.; Vazquez, J., Improved method for differential expression proteomics using trypsin-catalyzed ^{18}O labeling with a correction for labeling efficiency. *Mol Cell Proteomics* **2007**, *6* (7), 1274-86.
4. Peterson, A. C.; Russell, J. D.; Bailey, D. J.; Westphall, M. S.; Coon, J. J., Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics* **2012**.
5. Picotti, P.; Rinner, O.; Stallmach, R.; Dautel, F.; Farrah, T.; Domon, B.; Wenschuh, H.; Aebersold, R., High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat Methods* **2010**, *7* (1), 43-6.
6. Bonzon-Kulichenko, E.; Perez-Hernandez, D.; Nunez, E.; Martinez-Acedo, P.; Navarro, P.; Trevisan-Herraz, M.; Ramos Mdel, C.; Sierra, S.; Martinez-Martinez, S.; Ruiz-Meana, M.; Miro-Casas, E.; Garcia-Dorado, D.; Redondo, J. M.; Burgos, J. S.; Vazquez, J., A robust method for quantitative high-throughput analysis of proteomes by ^{18}O labeling. *Mol Cell Proteomics* **2011**, *10* (1), M110 003335.
7. MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010**, *26* (7), 966-8.
8. Chang, C. Y.; Picotti, P.; Huttenhain, R.; Heinzelmann-Schwarz, V.; Jovanovic, M.; Aebersold, R.; Vitek, O., Protein significance analysis in selected reaction monitoring (SRM) measurements. *Mol Cell Proteomics* **2012**, *11* (4), M111 014662.
9. Huang, X.; Tolmachev, A. V.; Shen, Y.; Liu, M.; Huang, L.; Zhang, Z.; Anderson, G. A.; Smith, R. D.; Chan, W. C.; Hinrichs, S. H.; Fu, K.; Ding, S. J., UNiquant, a program for quantitative proteomics analysis using stable isotope labeling. *J Proteome Res* **2011**, *10* (3), 1228-37.
10. Huber, W.; von Heydebreck, A.; Sultmann, H.; Poustka, A.; Vingron, M., Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **2002**, *18 Suppl 1*, S96-104.
11. Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **2008**, *26* (12), 1367-72.
12. Karp, N. A.; Huber, W.; Sadowski, P. G.; Charles, P. D.; Hester, S. V.; Lilley, K. S., Addressing accuracy and precision issues in iTRAQ quantitation. *Mol Cell Proteomics* **2010**, *9* (9), 1885-97.
13. Mahoney, D. W.; Therneau, T. M.; Heppelmann, C. J.; Higgins, L.; Benson, L. M.; Zenka, R. M.; Jagtap, P.; Nelsestuen, G. L.; Bergen, I., H Robert; Oberg, A. L., Relative Quantification: Characterization of Bias, Variability and Fold Changes in Mass Spectrometry Data from iTRAQ-Labeled Peptides. *Journal of Proteome Research* **2011**, *10* (9), 4325-4333.

14. D'Agostino, R. B., An Omnibus Test of Normality for Moderate and Large Size Samples. *Biometrika* **1971**, 58 (2), 341-348.