Journal of proteome research

Article

pubs.acs.org/jpr

General Statistical Framework for Quantitative Proteomics by Stable Isotope Labeling

Pedro Navarro,[†,‡,◆] Marco Trevisan-Herraz,[†,§,◆] Elena Bonzon-Kulichenko,[†,§,◆] Estefanía Núñez,[†,§] Pablo Martínez-Acedo,[†,§] Daniel Pérez-Hernández,[†,§] Inmaculada Jorge,[†,§] Raquel Mesa,[†,§] Enrique Calvo,[§] Montserrat Carrascal,[∥] María Luisa Hernáez,[⊥] Fernando García,[#] José Antonio Bárcena,[▽] Keith Ashman,[#,○] Joaquín Abian,[∥] Concha Gil,[⊥] Juan Miguel Redondo,[§] and Jesús Vázquez*[,†,§]

[†]Centro de Biología Molecular Severo Ochoa, CSIC−UAM, 28049 Madrid, Spain

[‡]Institute of Immunology, University Medical Center of the Johannes Gutenberg University Mainz, 55131 Mainz, Germany

[§]Centro Nacional de Investigaciones Cardiovasculares, 28029 Madrid, Spain

[∥]Instituto de Investigaciones Biomédicas de Barcelona (IIBB-CSIC), 08036 Barcelona, Spain

[⊥]Facultad de Farmacia, Universidad Complutense de Madrid, 28040 Madrid, Spain

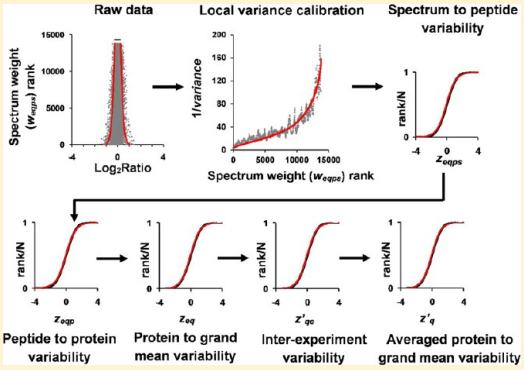
[#]Centro Nacional de Investigaciones Oncológicas, 28029 Madrid, Spain

[▽]Universidad de Córdoba and Córdoba Maimónides Institute for Biomedical Research (IMIBIC), 14004 Córdoba, Spain

[○]University of Queensland, Brisbane, St Lucia, Queensland 4072, Australia

Supporting Information

ABSTRACT: The combination of stable isotope labeling (SIL) with mass spectrometry (MS) allows comparison of the abundance of thousands of proteins in complex mixtures. However, interpretation of the large data sets generated by these techniques remains a challenge because appropriate statistical standards are lacking. Here, we present a generally applicable model that accurately explains the behavior of data obtained using current SIL approaches, including ^{18}O, iTRAQ, and SILAC labeling, and different MS instruments. The model decomposes the total technical variance into the spectral, peptide, and protein variance components, and its general validity was demonstrated by confronting 48 experimental distributions against 18 different null hypotheses. In addition to its general applicability, the performance of the algorithm was at least similar than that of other existing methods. The model also provides a general framework to integrate quantitative and error information fully, allowing a comparative analysis of the results obtained from different SIL experiments. The model was applied to the global analysis of protein alterations induced by low H_2O_2 concentrations in yeast, demonstrating the increased statistical power that may be achieved by rigorous data integration. Our results highlight the importance of establishing an adequate and validated statistical framework for the analysis of high-throughput data.



KEYWORDS: *Quantitative proteomics, stable isotope labeling, statistical analysis, yeast*

INTRODUCTION

Biological systems are dynamic and contain many molecular species (including DNA, RNA, proteins, carbohydrates, and lipids) whose interactions result in complex series of physicochemical, spatial, and temporal changes. Because proteins participate in all cellular processes, knowing how protein amounts and activities change over time provides important information about the state of the system. It is from such knowledge that the molecular mechanisms of disease and new pharmacological targets and biomarkers will emerge. Recent advances in mass spectrometry (MS)-based proteomics allow the identification and relative quantification of thousands of proteins in a single study. Despite these advances, the reproducibility of MS-based proteomics has been called into question.[1] It is an accepted fact that when the technology is properly applied, it is highly reproducible;[2] therefore, progress in the field will depend on a correct understanding of these techniques and their limitations.[3] The development of suitable statistical models is a critical step toward achieving this goal.

MS-based quantitative proteomics may be performed by direct quantitation of precursor or fragment ion peptide intensity in each of the samples (label-free approaches) or by using stable isotope labeling (SIL) techniques. In the most

Received: July 5, 2013
Published: January 31, 2014

dx.doi.org/10.1021/pr4006958 | J. Proteome Res. 2014, 13, 1234−1247

common setup, label-free quantitation involves analyzing several technical replicates of the same sample or samples from several subjects (biological replicates) belonging to two or more different conditions. Because of the multiplicative nature of the different factors (fixed effects) and error sources (random effects) involved, quantitative data are usually subjected to a logarithmic transformation that allows treating these effects as additive, providing a natural way for modeling the replicate structure of the data within the analysis of variance (ANOVA) framework.[4] A common feature of these ANOVA models is that the replicated structure are used to estimate the fixed effects and the variances associated with random errors, which are assumed to be normally distributed.[4−9] Although the analysis may be performed at the peptide feature level (one test per feature),[4,6] the quantitative results from different peptide features belonging to the same protein are usually integrated so that the analysis is at the protein level (one test per protein). Taking into account that the analysis is repeated for very large numbers of proteins (multiple hypothesis testing), statistically significant protein abundance changes under the different conditions are then detected by adjusting the *p*-value threshold to control for the false discovery rate (FDR).[10] In these models, all peptide features are considered to contribute equally to protein abundance, which may be estimated from the plain average of feature log-corrected intensities[4] or from features corrected by fixed effects[4,6,7] or scaled to the same level[9] before generating the protein average. In these approaches, random errors are assumed to derive from only one source so that the variance is a measure of the technical variability, but in some cases, biological variability is also taken into account by decomposing the total variance into the biological and the technical components.[4,6]

Stable isotope labeling (SIL) techniques, including stable isotope labeling with amino acids in cell culture (SILAC),[11,12] isobaric tagging for relative and absolute quantification (iTRAQ),[13] and enzymatic $^{16}O/^{18}O$ labeling,[14,15] currently offer the most accurate means of performing comparative quantitative proteomics studies.[16] Although the existence of separate technical, experimental,[4,6] and biological variations in iTRAQ analysis has been analyzed,[17] no statistical models were derived from these studies. The simplest model to analyze iTRAQ data is to calculate the protein value as an average of peptide ratios and to compare each of the protein values across several replicates (one test per protein) using an appropriate statistical test such as Student's *t* test. The protein average may be calculated using the mean,[18] the median,[19,20] or an average calculated by minimizing the square-root distance from the peptide readings from the log-transformed peptide ratios.[20] This kind of testing at the protein level has been extended using an ANOVA model, similar to those proposed to treat label-free data, with additive peptide effects and only one random effect that combines the biological error and the measurement noise,[21−24] which has been applied to the analysis of a case study involving four treatment groups with several replicates each.[8] All of these approaches have in common that all of the peptide readings originating from the same protein are equally considered, under the implicit assumption that they have the same variance. This assumption is based on original analysis showing that noncorrected ratios of peptides measured by iTRAQ follow approximately a log-normal distribution.[20] However, other analyses have demonstrated that not all of the peptides are quantified with the same accuracy, demonstrating a clear dependence of variance with ion intensity,[25−29] which may produce deviations from normality. Therefore, using intensity-weighted peptide averages of log ratios[26,30] to calculate protein averages has been proposed. Other approaches try to model or control the behavior of variance by a two-parameter modeling of the dependence of peptide variance with intensity[28] or by using a variance-stabilizing normalization (VSN) transformation[25,31] (similar to those employed in microarray approaches[32]), calculating the protein values as the median[31] or the trimmed average[25] of transformed peptide values. The transformed ratios at the spectrum level have been shown to have an apparently normal distribution, from which it is possible to detect statistically significant regulation of specific peptides (such as tyrosine-phosphorylated peptides) from the global distribution of transformed peptide ratios.[28] To test the significance of the changes at the protein level, most methods assume that the set of protein values follow a normal distribution and use a 0.05 probability threshold.[26,31] This test may be performed by direct fitting to a normal distribution[26] or by using estimates of the standard deviation,[31] whereas other approaches adjust the significance threshold so that 95% of experimental variation is encompassed.[25] Finally, some authors correct the significance value for multiple hypothesis testing.[31] Concerning SILAC data, the majority of studies use MaxQuant algorithm[33] to analyze quantitative results. In MaxQuant, protein ratios are calculated as the median of all SILAC peptide ratios, and the proteins are then grouped into bins according to their summed peptide intensities; in each bin, protein log -ratios are then assumed to be normally distributed, and the standard deviation is calculated using a robust estimate, from which a statistical significance is assigned to each protein. This procedure empirically takes into account the observed fact that high-abundance protein values have a lower variability than low-abundance ones.[33] Finally, we have recently proposed a statistical model to analyze quantitative data obtained by ^{18}O labeling, which decomposes the total technical variance into the spectrum, peptide, and protein variance components.[34] This approach models the heterogeneous variance at the spectrum level and integrates log ratios to the protein level using weighted averages according to error propagation theory. The validity of the model was demonstrated by showing that the distribution of protein values follows a normal distribution and by the very low percentage of outliers found at the spectrum, peptide, and protein levels.[34]

All of these studies show that, in spite of the efforts made in the field, a comprehensive statistical theory for the general analysis of quantitative data by SIL has not been developed yet. Existing models are highly specific to each SIL method and mass spectrometer, making them unsuitable for examining data from different laboratories, judging experimental quality on the basis of unified criteria, handling, comparing, and integrating multiple measurements, or interpreting the complete set of experimental results from different SIL approaches as a whole. Moreover, most models and statistical significance tests are based on normality assumptions that have not been tested despite the fact that heterogeneity of variance has been documented in all SIL methods.[25,33,34] These techniques are based on peptide-centric measurements, and the lack of general models leads to the subjective choice of a method for combining multiple peptide readings to estimate protein ratios.[25] This problem is further aggravated by the under-sampling that characterizes SIL-based MS analysis:[2] the number of peptides that quantify a protein is variable and cannot be

dx.doi.org/10.1021/pr4006958 | J. Proteome Res. 2014, 13, 1234−1247

Table 1. Statistical Parameters Estimated in the Null-Hypothesis Experiments

method	samples	weight constant (k_e)	spectrum variance (σ^2_S) (95% CI)	peptide variance (σ^2_P) (95% CI)	protein variance (σ^2_Q) (95% CI)	number of protein outliers ($FDR_q < 0.05$)
iTRAQ-TOFTOF	A vs A (114 vs 116)	182	0.029 (0.025−0.035)	0.012 (0.005−0.02)	0.002 (0−0.005)	0
iTRAQ-TOFTOF	B vs B (115 vs 117)	210	0.017 (0.013−0.021)	0.01 (0.005−0.015)	0.005 (0.002−0.009)	0
iTRAQ PQD	A vs A (114 vs 116)	14	0.110 (0.106−0.115)	0.010 (0.003−0.013)	0.003 (0.001−0.007)	1
iTRAQ PQD	B vs B (115 vs 117)	17	0.095 (0.091−0.100)	0.029 (0.022−0.037)	0.001 (0.000−0.005)	1
^{18}O-HR	A vs A	36	0.008 (0.0071−0.0085)	0.011 (0.009−0.012)	0.006 (0.0035−0.0064)	0
^{18}O-LR	A vs A	0.17	0.019 (0.018−0.021)	0.044 (0.038−0.050)	0.005 (0.002−0.009)	0
SILAC-HR	A* vs A	15	0.0001 (0.0001−0.0003)	0.0048 (0.004−0.005)	0.010 (0.009−0.012)	18
SILAC-LR	A* vs A	0.1	0.006 (0.0056−0.0063)	0.0082 (0.0069−0.0085)	0.0055 (0.005−0.0073)	9

Table 2. Statistical Parameters Estimated in the Control vs H_2O_2-Treatment Experiments

method	samples	weight constant (k_e)	spectrum variance (σ^2_S) (95% CI)	peptide variance (σ^2_P) (95% CI)	protein variance (σ^2_Q) (95% CI)	number of protein outliers ($FDR_q < 0.05$)
iTRAQ-TOFTOF	A vs B (114 vs 115)	236	0.018 (0.014−0.023)	0.0077 (0.003−0.013)	0.0269 (0.022−0.033)	1
iTRAQ-TOFTOF	A vs B (116 vs 117)	311	0.02 (0.014−0.028)	0.0073 (0.001−0.013)	0.0159 (0.010−0.022)	5
iTRAQ PQD	A vs B (114 vs 115)	15	0.11 (0.106−0.114)	0.017(0.012−0.021)	0.0058 (0.001−0.010)	12
iTRAQ PQD	A vs B (116 vs 117)	16	0.11 (0.105−0.113)	0.024 (0.019−0.028)	0.009 (0.005−0.015)	6
^{18}O-HR	A vs B	32	0.0076 (0.0060−0.0080)	0.042 (0.039−0.046)	0.035 (0.030−0.041)	12
^{18}O-LR	A vs B	0.17	0.003 (0.0025−0.0032)	0.021 (0.020−0.023)	0.011 (0.009−0.015)	8
SILAC-HR	A* vs B	9.4	0.0004 (0.0002−0.0006)	0.0037 (0.003−0.004)	0.018 (0.017−0.02)	39
SILAC-LR	A* vs B	0.16	0.0041 (0.0039−0.0043)	0.007 (0.0065−0.0074)	0.021 (0.019−0.023)	14

controlled between experiments, a nontrivial fact that complicates mathematical modeling.

Here, we describe the WSPP (weighted spectrum, peptide, and protein) model, a generally applicable statistical framework for the analysis of data generated with SIL-MS technologies. The model provides a detailed description of the behavior of technical variance, and by analyzing it independently at the spectrum, peptide, and protein levels, the model is able to capture separately the specific error sources of each SIL and MS method, demonstrating that error distributions are accurately modeled in all cases at the three levels. The model generates the integration at each level taking into account the separate variances according to error propagation theory so that the specific variance of each value at any level is accurately estimated. The model efficiently resolves the under-sampling problem, providing a framework to analyze the data at each level using unique normal distributions. In addition, the statistical framework also allows comparing and integrating results obtained using different SIL techniques so that full control over variance is maintained in the integrated data, opening the possibility of making further integrations at upper levels. The WSPP model provides a general framework for analyzing quantitative SIL data on the basis of a unique and validated statistical model.

MATERIALS AND METHODS

Yeast Culture, SILAC Labeling, and Sample Preparation

Cells from the lysine auxotrophic *Saccharomyces cerevisiae* strain YMJ38 (his3Δ1 leu2Δ0 ura3Δ0 arg4Δ::kanMX4 trp1Δ::kanMx4 lys2Δ) were cultured and SILAC-labeled with [$^{13}C_6$]L-lysine and [$^{13}C_6$]L-arginine as previously described.[35] SILAC light (A) and heavy (A*) media were prepared from minimal synthetic dextrose medium (0.5% (w/v) ammonium sulfate, 2% (w/v) glucose, 0.17% (w/v) nitrogen base without amino acids, and 20 mg/L of amino acids except for lysine and arginine).[35] Cells were grown in the corresponding medium at 30 °C and with shaking at 200 rpm until OD_{600} = 0.8. After adjusting both cell cultures to OD_{600} = 0.4 with the corresponding SILAC medium, culture A was split in two equal volumes, 20% (v/v) H_2O_2 (Panreac) was added to one of them (sample B) at a final concentration of 0.5 mM, and the other sample remained as a control (sample A). The three cell cultures (A, A*, and B) were grown at 30 °C with shaking at 200 rpm for 3 h until OD_{600} = 1.6 and were then centrifuged at 3000*g* at 4 °C for 10 min. Cell pellets were washed three times with cold distilled water, taken up in 500 μL of homogenization buffer (50 mM Tris-HCl, pH 7.5, 10% glycerol, 1% Triton X-100, 150 mM NaCl, 0.1% SDS, and 5 mM EDTA) supplemented with 1 mM PMSF and the protease inhibitors cocktail from Sigma-Aldrich (St. Louis, MO) and lysed using a homogenizer with glass Ballotini beads (0.45 μm) (Sigma). Cell lysates were centrifuged (13 000*g*) at 4 °C for 15 min, and protein concentration in the supernatant was determined by Bradford assay (Biorad). The cleared lysates were stored in aliquots of 100 or 500 μg of protein.

Experimental Design of Quantitative Techniques

The lysate aliquots from the three samples (A, non treated; B, treated; and A*, nontreated, SILAC-labeled) were subjected to high-throughput quantitation using the iTRAQ, ^{18}O, and SILAC-labeling techniques using both low- and high-resolution instruments in a total of 16 experiments (Tables 1 and 2) performed in different laboratories. Eight experiments were performed to test the null hypothesis associated with each technique (Table 1), and another eight experiments were performed to analyze the effect of H_2O_2 treatment (Table 2). Note that because of the use of four-plex iTRAQ reagents we could perform two replicate comparisons in the same experiment.

Protein Digestion and Labeling and Peptide Fractionation

In each experiment, the lysate aliquots were independently digested, and the resulting peptides were labeled (except in the SILAC cases) so that all of the replicates included all of the

dx.doi.org/10.1021/pr4006958 | J. Proteome Res. 2014, 13, 1234−1247

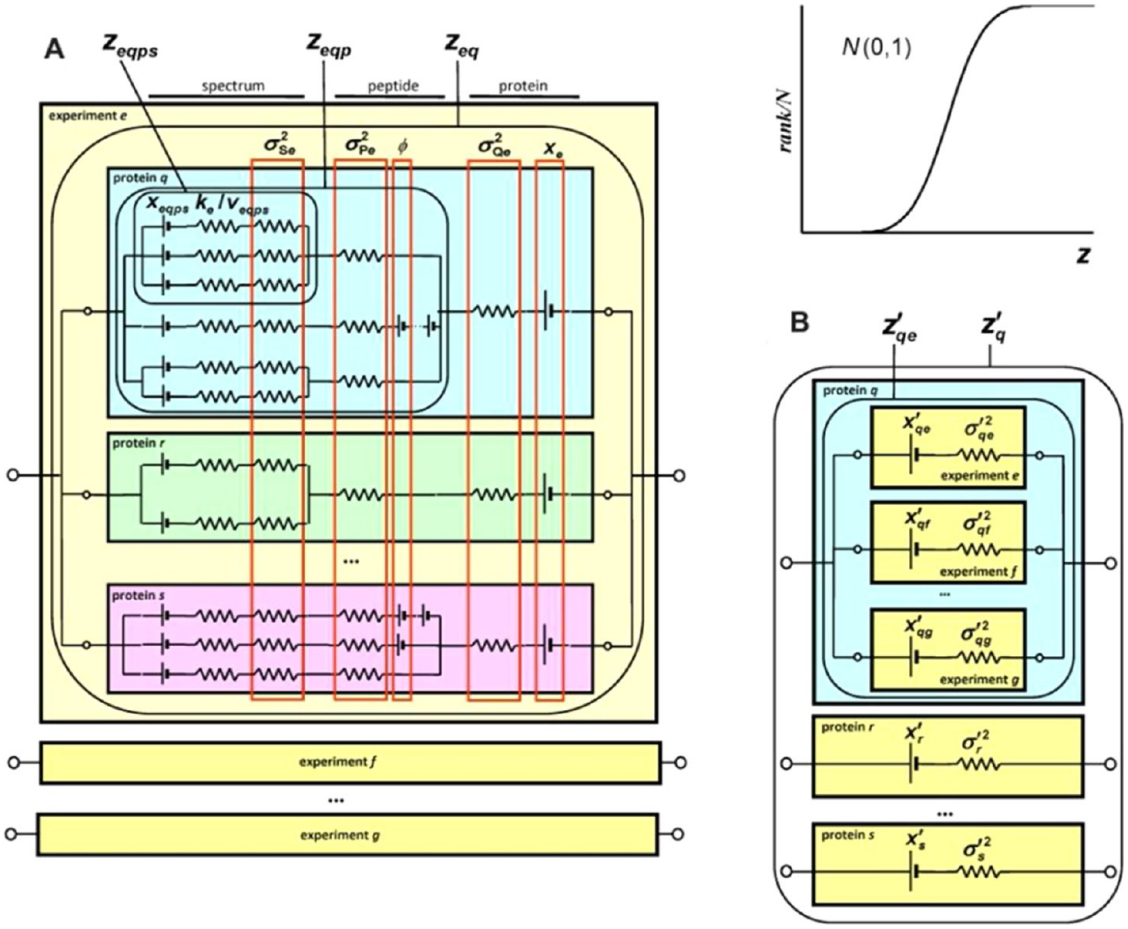


Figure 1. Representation of the WSPP model as an equivalent electric circuit. (A) In a quantitative SIL experiment, each protein is quantified by one or more peptides and each peptide is quantified in one or more spectra, represented by the branches of an electrical circuit. Resistances and batteries correspond to the variances and $\log_2$ ratios of each element, respectively. Within a given experiment, spectrum branches contain a spectrum-specific resistance (k_e/v_{eqps}) and a constant resistance (σ^2_{Se}), whereas peptide and protein branches have constant resistances (σ^2_{Pe} and σ^2_{Qe}). The correction for Arg to Pro conversion in SILAC experiments, estimated by a maximum likelihood method, is equivalent to a battery (ϕ) for each Pro residue in the peptide branches, and the systematic experiment error is equivalent to a constant battery (x_e) in the protein branches. (B) Integration of technical replicates is equivalent to setting the protein circuits from different experiments in parallel. In the null hypothesis, normal distributions with zero mean and unit variance are followed by the seven standardized parameters describing the variability in units of standard deviation among spectra within a peptide (z_{eqps}), among peptides within a protein (z_{eqp}), among proteins within a experiment (z_{eq}), among experiments within a protein (z'_{qe}), and among proteins (z'_q) (inset).

technical error sources. Each lysate aliquot was trypsin-digested separately using the whole proteome in-gel digestion protocol we described previously.[15] Prior to digestion, the gel bands were reduced with 10 mM DTT and alkylated with 50 mM iodoacetamide.[45] iTRAQ labeling was performed essentially according to the manufacturer's instructions; details of the procedure are described in the Supporting Information. ^{18}O labeling was performed following the robust protocol that we have previously described in detail.[15] In the experiments in Table 2, sample B was labeled with ^{18}O and sample A was labeled with ^{16}O. In all of the experiments and prior to MS analysis, peptides were IEF-separated into 24 fractions exactly as described previously.[15]

Mass Spectrometry Analysis

Low-resolution analysis of SILAC- and ^{18}O-labeled peptides was performed using a linear ion trap LTQ (Thermo-Finnigan) in the Cardiovascular Proteomics Laboratory at the Centro de Biología Molecular Severo Ochoa (Madrid). High-resolution analyses of SILAC- and ^{18}O-labeled peptides were performed using an LTQ-Orbitrap XL ETD (Thermo-Finnigan) in the Proteomics Unit of the Centro Nacional de Investigaciones Cardiovasculares (Madrid). iTRAQ-labeled samples were analyzed using either a linear ion trap LTQ (Thermo-Finnigan) working in the PQD scanning mode in the Proteomics Facility of the IDIBAPS (Barcelona) or a MALDI-TOF/TOF (Applied Biosystems) in the Proteomics Unit of the Centro Nacional de Investigaciones Oncológicas (Madrid). Details of the methods used are described in the Supporting Information.

Peptide Identification

Peptide identification was performed using either SEQUEST or Mascot. SEQUEST results were analyzed using the probability ratio method,[36] taking into account isoelectric points of peptides to improve peptide identification.[15] In all cases, false discovery rates (FDR) of peptide identifications were calculated from the search results against the inverted databases using the refined method.[15] Full details are described in the Supporting Information. The identification data were deposited at the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[37] under data set identifier PXD000325.

dx.doi.org/10.1021/pr4006958 | J. Proteome Res. 2014, 13, 1234−1247

Peptide Quantification and Statistics

For low-resolution MS data (LTQ) obtained by ^{18}O labeling, peptide quantification from ZoomScan spectra and calculation of labeling efficiencies of peptides were performed by fitting the spectra to a theoretical function as described previously[38,39] using QuiXoT, a program written in C# in our laboratory, which is available at http://www.cnic.es/en/inflamacion/proteomica/wiki/. Information about the meaning of the parameters used in QuiXoT is also available in the Supporting Information. For high-resolution MS data (LTQ-Orbitrap) obtained by ^{18}O labeling, quantification was performed directly in the full-scan spectrum with the same algorithm used for calculating labeling efficiency by taking into account the theoretical isotopic envelope of the identified peptide,[39] with the exception that mass peaks were not fitted to a Gaussian/double-exponential distribution but to the maximum intensity of each MS peak within a 0.02 m/z window around the theoretical m/z value. The same methods were used for low- and high-resolution SILAC data, respectively, except that the difference between the labeled and nonlabeled species was calculated from the peptide sequence and no correction for labeling efficiency was used. For iTRAQ data, only the intensity of the reporter ions within 0.4 Da windows around the theoretical values was taken into account for the quantification. Reporter intensities are corrected for isotopic contaminants by taking the information provided by the manufacturer into consideration. For each spectrum, a fitting weight was calculated upon quantification according to Supporting Information Table 2 by taking into account the intensity of peaks and the mean squared deviation between the theoretical curve and the experimental data. The data analysis workflow was performed as previously described.[34] The statistical model used to analyze the quantitative data is described in full detail in the Supporting Information and is schematized in Figure 1. Briefly, the $\log_2$ ratio of concentration in the two samples being compared, A and B, determined by spectrum s of peptide p derived from protein q in experiment e is expressed as $x_{eqps} = \log_2(A/B)$. The $\log_2$-ratio value associated with each peptide, x_{eqp}, is then calculated as a weighted average of the spectra used to quantify the peptide, and the value associated with each protein, x_{eq}, is similarly the weighted average of its peptides. In addition, a grand mean, x_e, is calculated in each experiment as a weighted average of the protein values. The statistical weights in all cases are the inverse of the local variances of each of the spectrum, peptide, and protein values. The local variance of each spectrum is modeled as a hyperbolic function of its fitting weight using two parameters, k_e and σ^2_{Se}. The local variances of the peptide and protein values are then calculated by error propagation theory by taking into account two additional constant variances: at the peptide, σ^2_{Pe}, and at the protein level, σ^2_{Qe}. The four constant parameters, k_e, σ^2_{Se}, σ^2_{Pe}, and σ^2_{Qe}, which describe the error distribution in each experiment, are estimated from all of the data together by an iterative method that uses robust approaches. The global distribution of values at each one of the levels is described using a standardized variable, z, that expresses the quantitative values in units of standard deviation and that in the null hypothesis is expected to follow a $N(0,1)$ normal distribution. Outliers at the scan and peptide levels and significant protein-abundance changes are detected from the z values by using a false discovery rate (FDR) threshold of 5%.

RESULTS

WSPP: A Statistical Model with a Multilayered Structure That Can Be Described by an Equivalent Electrical Circuit

The WSPP model separately considers the variances produced during (i) protein extraction and manipulation, (ii) peptide generation from their corresponding proteins and labeling, and (iii) generation of quantitative information from the mass spectra. We reasoned that this three-layered structure was the most appropriate for addressing the diverse sources of error generated by different labeling and MS approaches. The complete mathematical formulation is described in the Supporting Information, and the parameters used by the model are described in Supporting Information Table 1. The WSPP model takes into account under-sampling of data and uses robust algorithms to estimate variances. The model can be represented by a mathematically equivalent electrical circuit (Figure 1A) in which variances at the spectral, peptide, and protein levels, being independent, are therefore additive and can be represented as a series of resistances. Similarly, when replicates at the spectral or peptide levels are integrated to produce peptide or protein averages, their variances are treated following the same rule used for resistances located in parallel. Voltages represent abundance ratios in $\log_2$ scale. Branches with lower resistance contribute more to the final voltage, mirroring the fact that measurements with lower variance, which are more accurate, are weighted more in the averages. The equivalent circuit also shows how the addition of more branches at a given level decreases the resistance, reflecting the fact that the error of quantification diminishes when more values are averaged. However, the constant resistances set in series, which represents the constant error sources at each level, put a lower limit to this effect. This property reflects the fact that extensive averaging of a large number of spectra will never eliminate the error made at the peptide level, and averaging large number of peptides does not eliminate the error made at the protein level. Hence, averages are not affected much by very intense peaks or proteins with a very large number of peptides, as often happens in other weighting schemes.

To test the theory, we used a model system consisting of protein extracts from *S. cerevisae* cultures: sample A was obtained from untreated cells, sample A*, from untreated cells labeled by SILAC, and sample B, from cells treated with 0.5 mM H_2O_2. To ensure that data came from technical replicates, protein samples were prepared only once and stored in aliquots, and each aliquot was processed separately. The aliquots were digested with trypsin, and the resulting peptides were labeled (except for SILAC) and fractionated by isoelectric focusing using a robust protocol developed previously.[15] To analyze the null hypothesis (NH), A was compared with A or B was compared with B; for SILAC samples, pseudonull hypotheses were analyzed instead by comparing A* with A. To analyze the effect of H_2O_2 treatment, samples A and B were compared either by iTRAQ or ^{18}O labeling, and SILAC A* was compared with B. A total of 32 sample aliquots were processed for a total of 16 pairwise quantitative experiments (Tables 1 and 2). Peptide fractions were distributed among five laboratories and were analyzed on different mass spectrometers. SILAC and ^{18}O samples were analyzed using both low-resolution LTQ linear ion traps and a high-resolution hybrid LTQ-Orbitrap, whereas iTRAQ data were analyzed using both low-energy PQD fragmentation on an LTQ and high-energy

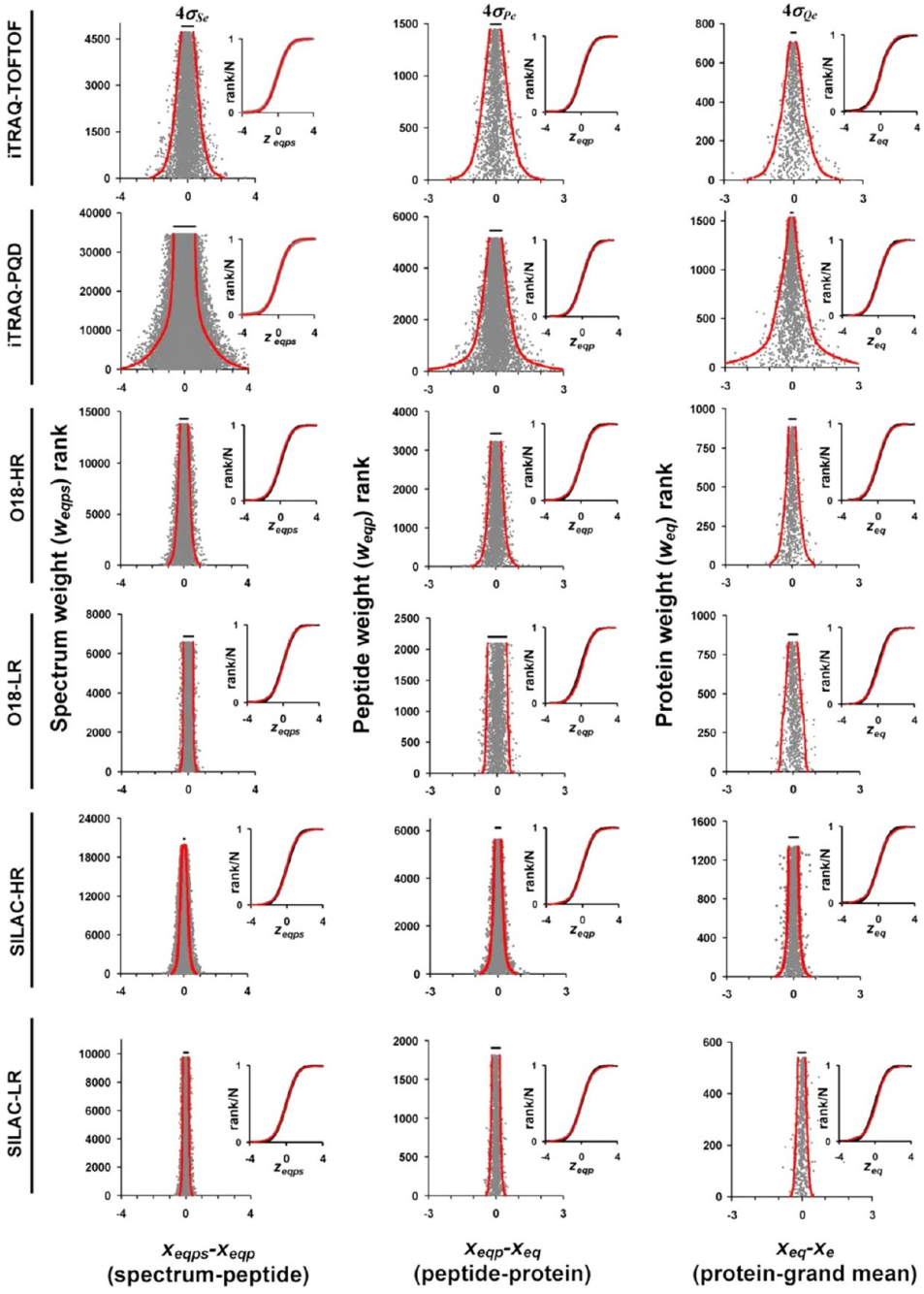


Figure 2. WSPP model gives a very accurate description of the distribution of quantitative errors under the null hypothesis at the spectral, peptide, and protein levels for all of the SIL methods analyzed. The plots show, for each one of the six SIL methods as indicated, the weight distributions of (left panels) $\log_2$ ratios of individual quantifications (spectra) around the corresponding peptides, (central panels) peptides around their corresponding proteins, and (right panels) proteins around the grand mean of the experiment. For better clarity, in the iTRAQ cases, only one of the two possible comparisons is shown. Red lines indicate local 95% confidence intervals (two standard deviations) in each case, as predicted by the model; horizontal black bars indicate the minimum (asymptotic) interval value. Insets show the cumulative distributions of the standardized variables described in Figure 1 (black points) at each level and SIL method; red lines are drawn according to the theoretical normal distribution with zero mean and unit variance, highlighting the excellent agreement between results and theory. In the true null-hypothesis experiments, only one false abundance change at the protein level (in iTRAQ-PQD) was detected as statistically significant at a 5% FDR among more than 1500 proteins (Tables 1 and 2).

CID fragmentation on a MALDI-TOF/TOF. The resulting MS data from all procedures were then gathered and analyzed.

Errors at the Spectrum, Peptide, and Protein Levels Can Be Modeled by Normal Distributions

Quantitative SIL experiments are peptide-centric approaches where ions are directly quantified in the MS detector; because not all spectra produce equally precise quantifications, each spectrum has a different variance and hence quantitative data cannot, in general, be treated as a whole by using a unique normal distribution.[34] The dependence of variance with ion intensity has been reported in several works.[25−29] The same trend was observed in the quantitative experiments performed to test the null hypothesis in this work; as shown in Figure 2,

dx.doi.org/10.1021/pr4006958 | J. Proteome Res. 2014, 13, 1234−1247

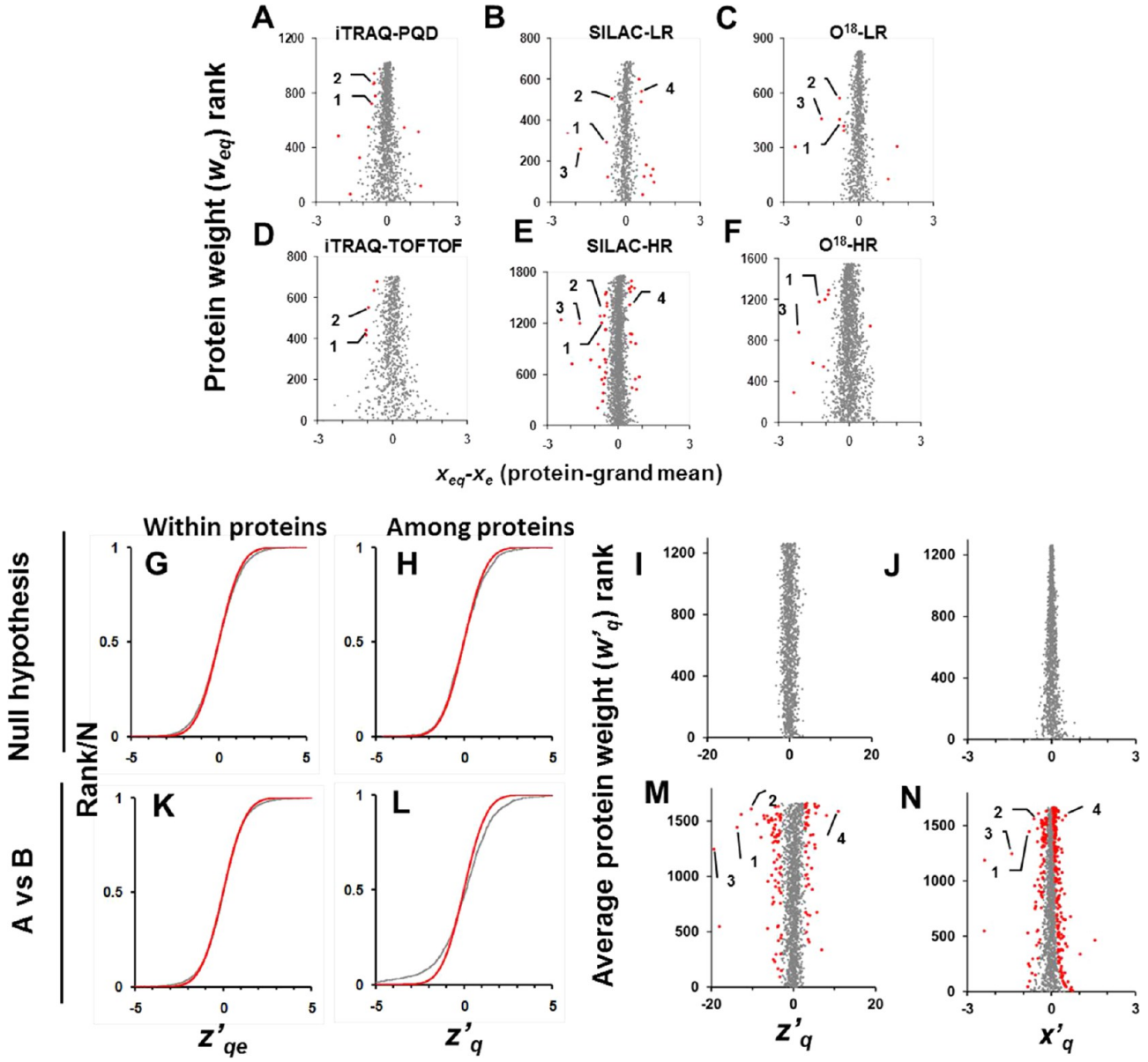


Figure 3. Detection of statistically significant protein-abundance changes produced by H_2O_2 treatment. (A–F) Weight distributions of protein quantifications around the grand mean obtained by each of the indicated SIL methods. In these plots, negative values signify an increase in protein abundance (toward the left) and positive values, a decrease (toward the right). Outliers at the protein level ($FDR_{eq} < 0.05$) are highlighted in red. Numbers indicate protein outliers that are consistently detected by most SIL methods: 1, cytochrome c peroxidase; 2, thioredoxin-2; 3, glutathione peroxidase 2; and 4, alcohol dehydrogenase 4. (G–L) Distribution of quantitative errors within and among proteins after integration of results from all SIL experiments. Upper panels correspond to the null-hypothesis experiments and lower panels, to the control vs H_2O_2-treatment study. The sigmoid plots show the distributions of the standardized variables describing variability of results from different experiments within the same protein (z'_{qe}, G, K) or among the different proteins (z'_q, H, L). The right panels show the weight distributions of z'_q (tornado plots I, M) and x'_q(J, N); red points indicate proteins showing a significant abundance change ($FDR'_q < 0.05$).

left column, the dispersion of quantitative data was not homogeneous but decreased with the statistical weight, a parameter that depended on ion intensity (see below). These data confirmed the heterogeneity of variance in all SIL and MS combinations analyzed.

The WSPP model addresses this issue by calculating a fitting weight (v_{eqps}) for each spectrum; this parameter considers quality features, such as peak intensity or goodness of fit, that are characteristic of the MS detector and SIL method used. The fitting weights were carefully designed (Supporting Information Table 2) to allow ranking by increasing accuracy of the entire collection of spectra in a given experiment so that quantifications from spectra having the same weight locally follow a normal distribution in all of the cases (Supporting Information Figure 1). The collection of fitting weights in a given experiment are then considered together to estimate the calibration constant (k_e) and the asymptotic spectrum variance (σ^2_{Se}) (Supporting Information Figure 2). These two parameters, which are characteristic of each experiment, are used to estimate separately the local variance of each one of the spectra from the fitting weights ($k_e/v_{eqps} + \sigma^2_{Se}$, Figure 1A). To study the distribution of $\log_2$ ratios at the spectrum level as a whole, we calculated a standardized variable, which expresses the $\log_2$-ratio deviation of the spectrum from the peptide average in

dx.doi.org/10.1021/pr4006958 | J. Proteome Res. 2014, 13, 1234–1247

Table 3. Integration of Protein-Abundance Changes in the Control vs Treated Experiments

Protein Description[a]	Corrected log_2-ratio: 18O HR	18O LR	iTRAQ PQD 114-115	iTRAQ PQD 116-117	iTRAQ TOFTOF 114-115	iTRAQ TOFTOF 116-117	SILAC HR	SILAC LR	x'_q	Fold change		z_{eq}: 18O HR	18O LR	iTRAQ PQD 114-115	iTRAQ PQD 116-117	iTRAQ TOFTOF 114-115	iTRAQ TOFTOF 116-117	SILAC HR	SILAC LR	z'_q	FDR'$_q$	Comments
sp|Q04120|TSA2 Peroxiredoxin TSA2	-3.09	-2.53	-2.05	-1.77			-2.42	-2.32	-2.38	5.21	up	-13.93	-13.41	-8.91	-7.83			-16.78	-14.14	-31.23	4.E-211	oxidative stress response, regulated by Yap1
sp|P38143|GPX2 Glutathione peroxidase	-2.13	-1.49	-0.59	-0.70			-1.61	-1.79	-1.39	2.63	up	-8.70	-9.54	-3.17	-3.43			-11.14	-10.39	-19.33	1.E-80	oxidative stress response, regulated by Yap1 and Skn7
sp|P41816|OYE3 NADPH dehydrogenase 3	-3.43						-1.94		-2.39	5.25	up	-14.21						-12.23		-18.03	3.E-70	oxidative stress response, regulated by Yap1
sp|P00431|CCPR Cytochrome c peroxidas	-1.26	-0.76	-0.63	-0.61	-0.74	-0.97	-0.77	-0.76	-0.78	1.72	up	-5.81	-4.85	-4.10	-3.63	-5.01	-5.76	-5.34	-4.48	-13.63	5.E-40	oxidative stress response, regulated by Yap1
sp|P22803|TRX2 Thioredoxin-2	-0.59	-0.75	-0.55	-0.73	-0.58	-0.74	-0.47	-0.54	-0.62	1.53	up	-2.62	-5.42	-4.54	-5.14	-5.28	-5.53	-3.29	-3.55	-12.68	1.E-34	oxidative stress response, regulated by Yap1
sp|P02365|RS6 40S ribosomal protein S	-0.25	-0.54	-0.52	-0.35	-0.48	-0.62	-0.25	-0.28	-0.45	1.37	up	-1.12	-3.41	-4.38	-2.49	-5.65	-5.93	-1.75	-1.88	-10.20	3.E-22	oxidative stress response, regulated by Skn7
sp|P29509|TRXB1 Thioredoxin reductase	-0.88	-0.58	-0.48	-0.29	-0.56	-0.47	-0.40	-0.44	-0.49	1.40	up	-4.19	-3.66	-3.47	-1.62	-3.80	-2.64	-2.91	-2.86	-8.85	1.E-16	oxidative stress response, regulated by Yap1 and Skn7
sp|P54114|ALDH3 Aldehyde dehydrogenas	-1.02	-0.60	-0.74	-0.46	-0.26	-0.54	-0.50	-0.35	-0.50	1.41	up	-4.73	-3.17	-2.52	-1.50	-2.02	-3.38	-3.56	-1.93	-7.88	3.E-13	oxidative stress response
sp|P05743|RL26A 60S ribosomal protein	-0.12	-0.44	-0.53	-0.35	-0.36	-0.26	-0.23	-0.24	-0.35	1.28	up	-0.58	-2.66	-5.04	-2.69	-2.95	-1.85	-1.59	-1.54	-7.16	8.E-11	ribosomal protein
sp|P38616|YGP1 Protein YGP1	-0.42	-0.42	-0.22	-0.38	-0.41	-0.59	-0.07	-0.38	-0.36	1.28	up	-1.63	-2.66	-1.85	-2.61	-3.57	-4.48	-0.46	-2.19	-7.01	2.E-10	oxidative stress response, regulated by Msn2/4
sp|P02406|RL28 60S ribosomal protein	-0.56	-0.16	-0.15	-0.38	-0.43	-0.28	-0.30	-0.31	-0.31	1.24	up	-2.64	-1.14	-1.09	-2.24	-4.12	-2.29	-2.14	-2.03	-6.34	2.E-08	oxidative stress response, regulated by Yap1
sp|P53829|CAF40 Protein CAF40	0.05						-1.17		-0.84	1.79	up	0.18						-7.42		-6.24	3.E-08	CCR4-NOT core complex, regulation of transcription
sp|Q96VH4|HBN1 Putative nitroreductas			-0.15	-0.03	1.24	-1.07		-0.42	-0.59	1.51	up			-0.72	-0.12	6.25	-4.56		-1.98	-6.17	4.E-08	oxidative stress response
sp|P05756|RS13 40S ribosomal protein	-0.09	-0.19	-0.37	-0.42	-0.35	-0.43	-0.25	-0.29	-0.32	1.24	up	-0.40	-1.27	-2.42	-2.47	-3.15	-3.08	-1.77	-1.86	-6.03	1.E-07	ribosomal protein
sp|P26781|RS11 40S ribosomal protein	-0.24	-0.42	-0.24	-0.48	-0.34	-0.33	-0.27	-0.26	-0.32	1.25	up	-1.11	-2.75	-2.16	-3.49	-1.88	-1.50	-1.89	-1.65	-5.93	2.E-07	ribosomal protein
sp|P05753|RS4 40S ribosomal protein S	-0.14	-0.25	-0.29	-0.33			-0.32	-0.31	-0.28	1.22	up	-0.68	-2.05	-2.96	-2.70			-2.28	-2.09	-5.43	3.E-06	ribosomal protein
sp|P09938|RIR2 Ribonucleoside-diphosp	-0.05	-0.33	-0.18	-0.18	-0.31	-0.61	-0.23	-0.25	-0.27	1.21	up	-0.25	-2.65	-1.50	-1.16	-2.21	-3.64	-1.67	-1.66	-5.28	6.E-06	DNA replication
sp|P05150|OTC Ornithine carbamoyltran	-0.17	-0.34	-0.71	-0.40	-0.01	0.83	-0.70		-0.47	1.39	up	-0.75	-1.64	-2.75	-1.24	-0.01	1.19	-4.80		-5.21	9.E-06	arginine synthesis
sp|P38061|RL32 60S ribosomal protein	-0.26	-0.43	-0.26	-0.39			-0.28	-0.31	-0.32	1.25	up	-0.98	-2.68	-2.21	-2.81			-1.55	-1.79	-5.18	1.E-05	ribosomal protein
sp|P14126|RL3 60S ribosomal protein L	0.25	-0.26	-0.30	-0.34	-0.32	-0.02	-0.32	-0.23	-0.26	1.20	up	1.19	-1.98	-3.14	-2.89	-0.86	-0.04	-2.32	-1.55	-5.00	2.E-05	ribosomal protein
sp|P33315|TKT2 Transketolase 2	-1.08		-1.14	-0.64			-0.28		-0.57	1.48	up	-3.66		-3.79	-1.96			-1.90		-4.99	2.E-05	oxidative stress response, regulated by Msn2/4
sp|Q03102|YMN1 Uncharacterized membra			-0.77	-0.46	-1.17	-0.30	-0.37		-0.54	1.45	up			-3.73	-1.51	-2.76	-0.50	-2.38		-4.98	2.E-05	oxidative stress response via the HOG pathway
sp|Q03104|MSC1 Meiotic sister chromat	-0.63	-0.31	-0.13	-0.12			-0.32	-0.42	-0.30	1.23	up	-3.08	-2.54	-0.99	-0.79			-2.32	-2.79	-4.96	3.E-05	enviornmental stress response
sp|P39938|RS26A 40S ribosomal protein	-0.36	-0.22	-0.25	0.03	-0.43	-1.03	-0.31	-0.47	-0.33	1.26	up	-1.19	-1.36	-1.38	0.15	-1.75	-3.73	-2.14	-2.89	-4.91	3.E-05	ribosomal protein
sp|P42846|KRI1 Protein KRI1	-2.35						-0.30		-0.81	1.75	up	-7.06						-1.56		-4.87	4.E-05	ribosome biogenesis
sp|P05738|RL9A 60S ribosomal protein	-0.03	-0.28	-0.25	-0.60	-0.03	-0.32	-0.26	-0.30	-0.27	1.21	up	-0.13	-2.26	-1.66	-3.57	-0.14	-1.34	-1.83	-1.97	-4.78	6.E-05	ribosomal protein
sp|P05754|RS8 40S ribosomal protein S	0.00	-0.29	-0.23	-0.43	-0.26	-0.35	-0.29	-0.30	-0.28	1.21	up	0.00	-2.32	-1.51	-2.41	-1.35	-1.40	-2.04	-1.99	-4.77	6.E-05	ribosomal protein
sp|P06115|CATT Catalase T	-0.25	-0.44	-0.52	-0.37	-0.29	0.42	-0.36	-0.28	-0.35	1.27	up	-1.17	-2.75	-2.47	-1.15	-0.71	0.85	-2.56	-1.57	-4.77	5.E-05	oxidative stress response, regulated by Msn2/4
sp|Q07653|HBT1 Protein HBT1	-0.97						-0.50	-0.44	-0.54	1.45	up	-3.06						-3.07	-2.47	-4.77	5.E-05	oxidative stress response, regulated by Msn4
sp|Q12068|GRE2 NADPH-dependent methyl	-0.90	-0.60	-0.47	0.01				0.15	-0.37	1.30	up	-4.27	-3.66	-2.99	0.04				0.83	-4.61	1.E-04	oxidative stress response via the HOG pathway
sp|P22943|HSP12 12 kDa heat shock pro	-0.03	-0.34	-0.14	-0.36			-0.40	-0.43	-0.29	1.22	up	-0.13	-1.63	-1.16	-2.50			-2.78	-2.71	-4.56	1.E-04	enviornmental stress response
sp|P05735|RL19 60S ribosomal protein	-0.35	-0.21	-0.29	-0.37	0.08	0.35	-0.31	-0.27	-0.26	1.20	up	-1.67	-1.64	-2.18	-2.37	0.23	0.91	-2.21	-1.79	-4.52	2.E-04	ribosomal protein
sp|P14796|RL40 60S ribosomal protein			-0.77	-0.76					-0.77	1.70	up			-3.37	-2.96					-4.49	2.E-04	ribosomal protein
sp|P22768|ASSY Argininosuccinate synt	0.30	-0.11	-0.17	-0.43			-0.40	-0.40	-0.25	1.19	up	1.39	-0.78	-1.64	-3.34			-2.84	-2.67	-4.41	3.E-04	oxidative stress response, regulated by Skn7
sp|P05030|PMA1 Plasma membrane ATPase	0.12	-0.13	-0.11	-0.25	-0.19	-0.23	-0.29	-0.25	-0.18	1.13	up	0.54	-1.10	-1.18	-2.20	-2.05	-2.02	-2.12	-1.62	-4.35	3.E-04	oxidative stress response
sp|P47137|YJ6G Uncharacterized oxidor	-0.64		-0.21	0.17			-0.35	-0.53	-0.41	1.33	up	-2.83		-0.56	0.43			-2.41	-2.93	-4.34	3.E-04	oxidative stress response
sp|P32827|RS23 40S ribosomal protein	-0.18	-0.47	-0.28	-0.23			-0.20	-0.26	-0.29	1.22	up	-0.61	-3.12	-2.14	-1.45			-1.31	-1.53	-4.33	3.E-04	ribosomal protein
sp|P05740|RL17A 60S ribosomal protein	-0.44	-0.16	-0.15	-0.27	-0.17	-0.34	-0.21	-0.21	-0.22	1.17	up	-2.03	-1.22	-1.26	-1.89	-1.22	-2.06	-1.47	-1.40	-4.27	4.E-04	ribosomal protein
sp|P14065|GCY Protein GCY	-0.72	-0.43	-0.53	0.10					-0.43	1.35	up	-3.15	-2.86	-2.43	0.37					-4.27	4.E-04	oxidative stress response
sp|P39078|TCPD T-complex protein 1 su	-0.24		-0.66	-1.06			-0.15	-0.39	-0.36	1.29	up	-1.12		-2.66	-3.44			-1.04	-2.36	-4.20	0.001	protein folding
sp|Q03246|RT17 37S ribosomal protein	-0.39						-0.69		-0.65	1.57	up	-0.86						-4.10		-4.14	0.001	ribosomal protein
sp|P00447|SODM Superoxide dismutase [	-0.68	-0.23	-0.20	-0.20	-0.16	-0.08	-0.29	-0.26	-0.26	1.20	up	-2.97	-1.70	-1.08	-0.96	-0.73	-0.29	-2.04	-1.67	-4.13	0.001	oxidative stress response
sp|P00445|SODC Superoxide dismutase [	-0.15	-0.40	-0.24	0.00	-0.16	-0.54	-0.02	-0.03	-0.20	1.15	up	-0.68	-2.69	-2.25	-0.03	-1.36	-3.62	-0.16	-0.20	-4.04	0.001	oxidative stress response, regulated by Yap1
sp|P47771|ALDH2 Aldehyde dehydrogenas	-0.34	-0.71	-0.53	-0.33	-0.04	-0.19	-0.14		-0.32	1.25	up	-1.70	-2.86	-2.77	-1.46	-0.13	-0.53	-1.01		-3.99	0.001	oxidative stress response
sp|P34227|PRX1 Mitochondrial peroxire	-0.69	-0.42	-0.17	-0.20	-0.11	-0.03	-0.15	-0.08	-0.25	1.19	up	-3.34	-3.05	-1.13	-1.06	-0.40	-0.09	-1.06	-0.51	-3.96	0.001	oxidative stress response
sp|P33442|RS3A 40S ribosomal protein	-0.11	-0.29	-0.19	-0.20	-0.05	-0.29	-0.20	-0.14	-0.19	1.14	up	-0.52	-2.26	-1.70	-1.53	-0.42	-1.99	-1.43	-0.90	-3.87	0.002	ribosomal protein
sp|P32337|IMB3 Importin subunit beta-	0.01						-0.59		-0.51	1.42	up	0.02						-4.14		-3.84	0.002	ribosome biogenesis
sp|P23248|RS3B 40S ribosomal protein	-0.39		-0.33	-0.16	-0.13	-0.40	-0.16	-0.36	-0.26	1.19	up	-1.54		-1.68	-0.73	-0.91	-2.30	-1.06	-2.13	-3.83	0.002	ribosomal protein
sp|P39741|RL35 60S ribosomal protein	-0.60	0.04	-0.35	-0.03			-0.29	-0.34	-0.25	1.19	up	-2.70	0.24	-2.60	-0.19			-1.95	-2.06	-3.79	0.003	ribosomal protein
sp|P38062|TVP23 Golgi apparatus membr	0.20						-0.87		-0.69	1.61	up	0.44						-4.24		-3.67	0.004	oxidative stress response, regulated by Msn2/4
sp|P36148|GPT2 Glycerol-3-phosphate O					-0.13	-0.55	-0.63		-0.48	1.40	up					-0.51	-1.83	-3.51		-3.66	0.004	phospholipid biosynthesis
sp|Q07896|NOC3 Nucleolar complex-asso	-0.36						-0.55		-0.52	1.44	up	-0.99						-3.53		-3.63	0.004	ribosome biogenesis
sp|P38013|AHP1 Peroxiredoxin type-2 O	-0.31	-0.22	-0.14	-0.19	-0.17	-0.34	0.16	0.20	-0.15	1.11	up	-1.40	-1.81	-1.44	-1.55	-2.04	-3.29	1.12	1.35	-3.62	0.005	oxidative stress response via the HOG pathway
sp|P49334|TOM22 Mitochondrial import		-0.18	-0.10	-0.15			-0.62		-0.35	1.27	up		-1.20	-0.34	-0.47			-4.06		-3.60	0.005	mitochondrial membrane organization
sp|Q04432|HSP31 Probable chaperone pr	-0.61	-0.29	-0.12	-0.14	-0.32	-0.02			-0.24	1.18	up	-2.65	-1.84	-0.73	-0.75	-2.51	-0.10			-3.59	0.005	oxidative stress response
sp|P04449|RL24A 60S ribosomal protein	-0.15	-0.22	-0.19	-0.23			-0.28	-0.24	-0.22	1.17	up	-0.70	-1.44	-1.47	-1.46			-1.95	-1.56	-3.57	0.005	ribosomal protein
sp|P16547|OM45 Mitochondrial outer me	-0.22	-0.50	-0.46	-0.15	0.08	1.34	-0.32	-0.12	-0.27	1.20	up	-0.92	-2.36	-2.54	-0.72	0.15	2.14	-2.08	-0.73	-3.51	0.007	mitochondrial membrane organization
sp|Q02753|RL21A 60S ribosomal protein	0.22	-0.15	-0.18	-0.33			-0.28	-0.33	-0.22	1.16	up	1.00	-0.80	-1.51	-2.36			-1.01	-2.12	-3.48	0.007	ribosomal protein
sp|P49631|RL43 60S ribosomal protein	-0.12	-0.29	0.02	-0.31	-0.37	-0.35	-0.12	-0.25	-0.20	1.15	up	-0.41	-1.53	0.18	-1.95	-2.21	-1.81	-0.84	-1.54	-3.46	0.008	ribosomal protein
sp|P53315|SOL4 Probable 6-phosphogluc	-1.07	-0.22			-0.83	0.24	-0.29		-0.35	1.27	up	-3.07	-1.34			-2.07	0.43	-1.89		-3.42	0.009	pentose phosphate pathway
sp|P06367|RS14A 40S ribosomal protein	-0.18	-0.28	-0.04	-0.09	-0.33	-0.11	-0.12	-0.19	-0.17	1.12	up	-0.79	-2.13	-0.34	-0.60	-2.71	-0.72	-0.83	-1.28	-3.39	0.009	ribosomal protein
sp|P04451|RL23 60S ribosomal protein	-0.25	-0.17	-0.42	0.17	0.04	-0.75	-0.21	-0.23	-0.21	1.15	up	-1.23	-1.31	-2.39	0.83	0.15	-2.74	-1.50	-1.46	-3.39	0.009	ribosomal protein
sp|P18544|ARGD Acetylornithine aminot	-0.04				-0.06	-0.80	-0.36	-0.48	-0.35	1.27	up	-0.13				-0.12	-1.09	-2.06	-2.54	-3.36	0.010	arginine synthesis
sp|Q06405|ATPK ATP synthase subunit f		-0.60	-0.43	0.33	-0.30	-0.31	-0.19		-0.30	1.23	up		-2.41	-0.99	0.73	-1.90	-1.74	-1.01		-3.34	0.011	mitochondrial proton-transporting ATP synthase complex
sp|P53858|BNI4 Protein BNI4			-1.13	-0.53					-0.77	1.71	up			-3.03	-1.74					-3.26	0.014	cell morphogenesis regulation
sp|P36139|PET10 Protein PET10	-0.47		-0.55	-0.03	-0.24	0.42	-0.23	-0.38	-0.28	1.21	up	-1.96		-1.94	-0.10	-0.76	1.11	-1.57	-2.15	-3.23	0.015	mitochondrial membrane organization
sp|P32288|GLNA Glutamine synthetase O	0.02	-0.50	-0.04	-0.39	0.26	-0.76	-0.09	-0.17	-0.20	1.14	up	0.08	-3.11	-0.26	-2.13	1.44	-3.68	-0.61	-1.13	-3.23	0.015	glutamine biosynthesis
sp|Q04304|YMY0 UPF0659 protein YMR090	-0.70		-0.16	-0.11	-0.21	-0.16	-0.28		-0.25	1.19	up	-2.79		-0.88	-0.55	-1.22	-0.73	-1.92		-3.23	0.015	oxidative stress response, regulated by Yap1
sp|P49723|RIR4 Ribonucleoside-diphosp	-0.20	-0.14	-0.12	-0.14	-0.14	-0.35	-0.21	-0.20	-0.17	1.12	up	-0.98	-1.22	-1.05	-1.01	-0.64	-1.46	-1.49	-1.36	-3.18	0.018	deoxyribonucleoside diphosphate metabolic process
sp|P41805|RL10 60S ribosomal protein	-0.13	-0.04	-0.25	-0.01	0.35	-0.31	-0.11	-0.08	-0.16	1.12	up	-0.58	-0.26	2.35	-0.07	2.29	-1.76	-0.74	-0.55	-3.12	0.021	ribosomal protein
sp|P38113|ADH5 Alcohol dehydrogenase	-0.01	0.36	0.27	-0.88			-0.78		-0.35	1.28	up	-0.03	1.38	0.70	-1.88			-4.84		-3.10	0.023	ethanol biosynthetic process
sp|Q01217|ARG5 Protein ARG5,6, mitoc	-0.17	0.02					-0.47	-0.29	-0.28	1.21	up	-0.83	0.11					-3.32	-1.44	-3.08	0.024	arginine synthesis
sp|P22217|TRX1 Thioredoxin-1	0.21	0.04	-0.26	-0.24	-0.40	-0.31	0.06	0.07	-0.16	1.11	up	0.95	0.28	-2.33	-1.78	-3.14	-1.99	0.42	0.36	-3.08	0.024	oxidative stress response
sp|P40515|FIS1 Mitochondria fission 1	-0.27						-0.49		-0.43	1.35	up	-0.95						-2.97		-3.05	0.026	mitochondrial membrane organization
sp|P47122|NPA3 GTPase NPA3	-0.09						-0.65		-0.48	1.39	up	-0.32						-3.43		-3.03	0.027	GTP binding
sp|Q07163|TY1AB Transposon TyH3 Gag-P	0.03	0.02	-0.08	-0.21	-0.22	-0.66	-0.15	-0.21	-0.15	1.11	up	0.17	0.18	-0.85	-1.79	-1.22	-3.21	-1.12	-1.43	-3.02	0.028	transposition
sp|P53221|RL26B 60S ribosomal protein	-0.15		-0.30	0.02			-0.36	-0.29	-0.25	1.19	up	-0.51		-1.67	0.09			-2.31	-1.66	-2.96	0.033	ribosomal protein
sp|P54861|DNM1 Dynamin-related protei	-0.23	-0.42					-0.29		-0.31	1.24	up	-1.04	-2.03					-1.96		-2.93	0.036	mitochondrial membrane organization
sp|P48361|ASK10 Protein ASK10	-1.04						-0.36		-0.56	1.48	up	-2.93						-1.57		-2.91	0.039	oxidative stress response
sp|Q12000|TMA46 Translation machinery	-0.63						-0.34		-0.40	1.32	up	-1.98						-2.21		-2.85	0.045	ribosome biogenesis

dx.doi.org/10.1021/pr4006958 | J. Proteome Res. 2014, 13, 1234−1247

Table 3. continued

Protein Description[a]	Corrected $\log_2$-ratio: 18O HR	18O LR	iTRAQ PQD 114-115	iTRAQ PQD 116-117	iTRAQ TOFTOF 114-115	iTRAQ TOFTOF 116-117	SILAC HR	SILAC LR	x'_q	Fold change		z_{eq}: 18O HR	18O LR	iTRAQ PQD 114-115	iTRAQ PQD 116-117	iTRAQ TOFTOF 114-115	iTRAQ TOFTOF 116-117	SILAC HR	SILAC LR	z'_q	FDR'_q	Comments
sp|P29952|MPI Mannose-6-phosphate iso	0.16	0.14	0.14	0.13	0.25	0.13	0.20	0.20	0.17	1.13	down	0.80	1.14	0.72	0.58	1.54	0.70	1.46	1.32	3.01	0.047	cell wall biosynthesis
sp|P25383|YC21A Transposon Ty2-C Gag							0.47	0.24	0.37	1.29	down							2.86	1.36	3.03	0.046	transposition
sp|P32775|GLGB 1,4-alpha-glucan-branc	0.30		0.55	-0.14			0.34	0.19	0.28	1.21	down	1.29		1.81	-0.42			2.40	1.10	3.04	0.044	glycogen metabolic process
sp|P40510|SER33 D-3-phosphoglycerate	0.45	0.32	0.17	-0.01			0.19		0.23	1.17	down	2.12	2.28	1.00	-0.06			1.39		3.08	0.040	serine family amino acid biosynthetic process
sp|P15019|TAL1 Transaldolase	0.02	0.01	0.04	0.21	0.15	0.24	0.24	0.23	0.13	1.10	down	0.12	0.09	0.50	1.88	1.35	1.73	1.74	1.54	3.08	0.041	pentose-phosphate shunt
sp|P11154|PYC1 Pyruvate carboxylase 1	0.35	0.20	0.00	0.07			0.25	0.41	0.17	1.13	down	1.68	1.55	0.00	0.56			1.82	2.71	3.09	0.040	gluconeogenesis
sp|P00925|ENO2 Enolase 2	-0.04	0.13	0.09	0.12	0.05	0.22	0.21	0.20	0.12	1.09	down	-0.22	1.16	0.90	0.95	0.66	2.38	1.53	1.36	3.11	0.039	gluconeogenesis
sp|P38934|BFR1 Nuclear segregation pr	0.07	0.31	0.08	0.21	0.22	0.14	0.15	0.19	0.16	1.11	down	0.36	1.38	0.84	1.72	1.51	0.81	1.06	1.25	3.11	0.040	mRNA metabolic process
sp|Q08224|THI20 Phosphomethylpyrimidi	0.49	0.14					0.38	0.36	0.32	1.25	down	1.10	0.69					2.47	1.90	3.20	0.030	mRNA metabolic process
sp|Q3E841|YN034 Uncharacterized prote	0.10	0.18	-0.24	0.40	0.39	0.01	0.48		0.23	1.17	down	0.40	1.17	-1.24	1.63	2.15	0.06	3.43		3.22	0.029	
sp|P33754|SEC66 Translocation protein							-0.26	1.16	0.44	1.36	down							-1.34	5.97	3.25	0.026	cell growth
sp|P37291|GLYC Serine hydroxymethyltr	0.15	0.15	0.09	0.12	0.15	0.16	0.32	0.20	0.16	1.11	down	0.72	1.20	0.99	1.07	0.91	0.85	2.33	1.35	3.28	0.025	serine family amino acid metabolic process
sp|P04807|HXK8 Hexokinase-2	0.01	0.13	0.07	0.21	0.41	0.38	0.24	0.20	0.18	1.13	down	0.08	1.09	0.50	1.32	2.00	1.58	1.78	1.35	3.29	0.024	hexose import
sp|P07149|FAS1 Fatty acid synthase su	0.08	0.08	0.10	0.16	0.08	0.21	0.21	0.19	0.13	1.10	down	0.43	0.73	1.24	1.50	0.89	1.79	1.52	1.29	3.29	0.025	fatty acid synthase activity
sp|P00560|PGK Phosphoglycerate kinase	-0.12	0.11	0.15	0.12	0.05	0.20	0.12	0.16	0.11	1.08	down	-0.62	0.99	1.93	1.13	0.82	2.52	0.86	1.13	3.32	0.023	gluconeogenesis
sp|P39003|HXT6 High-affinity hexose t		0.14	0.13	-0.11			0.74		0.30	1.23	down		0.62	0.74	-0.58			4.93		3.34	0.022	hexose transport
sp|P40043|YEP7 Uncharacterized protei	0.11	0.10	-0.22	0.12			0.53	0.45	0.28	1.21	down	0.39	0.60	-0.77	0.47			3.53	2.45	3.37	0.021	energy reserve metabolic process
sp|P11632|NHP6A Non-histone chromosom			0.44	0.21			0.50		0.43	1.34	down			1.78	0.77			3.01		3.46	0.015	polymerase III transcriptional preinitiation complex asser
sp|P07283|PMM Phosphomannomutase	0.07	0.22	0.20	0.24	0.00	-0.08	0.30	0.31	0.18	1.13	down	0.36	1.65	1.84	1.77	-0.02	-0.47	2.12	2.08	3.52	0.013	hexose biosynthetic process
sp|P46989|ATG27 Autophagy-related pro	-0.19				-0.01	0.89	0.72		0.46	1.38	down	-0.69				-0.03	1.90	4.28		3.54	0.012	vacuole organization
sp|P21524|RIR1 Ribonucleoside-diphosp	0.54	0.19	0.24	0.20			0.41		0.33	1.26	down	2.19	0.99	0.94	0.70			2.71		3.55	0.012	DNA replication
sp|Q03497|STE20 Serine/threonine-prot		0.70					0.37		0.51	1.43	down		3.24					1.92		3.59	0.011	positive regulation of protein kinase activity
sp|P19358|METK2 S-adenosylmethionine	0.26	0.15	0.15	0.18	0.19	0.09	0.31	0.26	0.20	1.15	down	1.22	1.02	1.08	1.15	1.16	0.43	2.23	1.75	3.62	0.010	sulfur compound biosynthetic process
sp|P46367|ALDH4 Potassium-activated a	0.05	0.00	0.03	0.06	0.25	0.40	0.44	0.38	0.18	1.13	down	0.25	0.00	0.32	0.44	1.85	2.48	3.17	2.54	3.68	0.008	alcohol metabolic process
sp|P47120|DOHH Deoxyhypusine hydroxyl	0.27		0.22	0.49	1.19	-0.64	0.17		0.29	1.22	down	1.17		1.58	2.99	2.92	-1.24	1.18		3.70	0.008	translation
sp|P04046|PUR1 Amidophosphoribosyltra	-0.01				1.29	-0.28	0.28	0.26	0.37	1.30	down	-0.06				2.16	-0.32	1.95	4.65	3.77	0.006	purine base metabolic process
sp|P54115|ALDH6 Magnesium-activated a	0.13	0.13	0.17	0.23	0.22	0.24	0.17	0.08	0.17	1.13	down	0.65	1.10	1.87	2.00	1.52	1.36	1.22	0.53	3.80	0.006	NADPH regeneration
sp|P01095|IPB2 Protease B inhibitors	0.24	-0.02	0.24	0.53	0.24	0.21	0.26	0.29	0.26	1.20	down	0.81	-0.08	1.51	2.74	0.98	0.66	1.85	1.76	3.83	0.006	vacuole organization
sp|Q06408|ARO10 Transaminated amino a	0.52						0.59		0.57	1.48	down	2.03						3.35		3.91	0.004	L-phenylalanine catabolic process
sp|P00812|ARGI Arginase	0.55		-0.07	0.32			0.77		0.51	1.42	down	1.78		-0.23	1.04			4.34		4.08	0.002	arginine catabolic process
sp|P39976|DLD3 D-lactate dehydrogenas	0.14	0.20	0.30	0.31	0.22	0.05	0.36	0.39	0.26	1.20	down	0.69	1.53	1.41	1.53	1.10	0.22	2.61	2.25	4.19	0.001	D-lactate dehydrogenase (cytochrome) activity
sp|Q04371|YMR7 UPF0364 protein YMR027	0.53		1.35	-0.17			0.04	0.57	0.39	1.31	down	1.97		6.24	-0.83			0.27	5.26	4.50	4.E-04	
sp|Q08911|FDH1 Formate dehydrogenase	0.24		0.40	0.52			0.57		0.47	1.39	down	0.99		1.67	2.01			3.87		4.62	2.E-04	NADH metabolic process
sp|P43616|CPGL Glutamate carboxypepti	0.02	0.38	0.23	0.35	-0.21	0.99	0.31	0.36	0.30	1.23	down	0.10	2.82	1.53	2.16	-0.58	2.64	2.25	2.16	4.88	7.E-05	dipeptidase activity
sp|P00924|ENO1 Enolase 1	-0.04	0.13	0.17	0.27	0.13	0.25	0.30	0.32	0.19	1.14	down	-0.21	1.21	2.05	2.46	1.79	2.75	2.15	2.22	5.22	1.E-05	gluconeogenesis
sp|P07991|OAT Ornithine aminotransfer	0.07	0.10	0.60	-0.03	1.42	0.83	0.55	0.64	0.36	1.29	down	0.33	0.73	2.54	-0.12	1.83	0.74	3.94	4.15	5.25	1.E-05	arginine catabolic process
sp|P16474|GRP78 78 kDa glucose-regula	0.06	0.27	0.10	0.27	0.29	0.22	0.30	0.22	0.22	1.16	down	0.29	2.27	1.14	2.51	3.01	1.86	2.14	1.51	5.31	9.E-06	response to unfolded protein
sp|P38009|PUR92 Bifunctional purine b	0.56	0.15	0.53	0.21			0.50	0.46	0.38	1.30	down	2.68	1.05	2.21	0.89			3.61	2.68	5.36	8.E-06	purine base metabolic process
sp|P17967|PDI Protein disulfide-isome	0.08	0.16	0.29	0.25	0.23	0.73	0.26	0.25	0.26	1.20	down	0.42	1.32	2.94	2.07	1.40	3.86	1.86	1.67	5.49	4.E-06	endoplasmic reticulum lumen
sp|Q03529|SCS7 Inositolphosphorylcera		0.69	-0.14	0.36			0.88		0.69	1.62	down		3.16	-0.35	0.83			5.33		5.75	1.E-06	lipid biosynthesis
sp|P11986|INO1 Inositol-3-phosphate s	0.27	0.24	0.27	0.15	0.35	0.12			0.25	1.19	down	1.38	2.17	2.93	1.29	4.37	1.21			5.79	1.E-06	lipid biosynthesis
sp|P07262|DHE4 NADP-specific glutamat	0.11	0.16	0.31	0.28	0.13	0.13	0.56	0.55	0.25	1.19	down	0.59	1.35	3.18	2.34	1.62	1.22	4.07	3.71	6.19	1.E-07	synthesis of α-Ketoglutarate
sp|P40989|FKS2 1,3-beta-glucan syntha		1.55	-0.14				0.25		1.04	2.06	down		3.23	-0.30				0.80		6.89	1.E-09	cell wall biosynthesis
sp|P40825|SYAC Alanyl-tRNA synthetase	0.41	0.32	0.23	0.43	0.40	0.65	0.46	0.53	0.40	1.32	down	2.13	2.49	2.18	3.40	2.97	3.94	3.32	3.20	8.16	0.E+00	translation
sp|P10127|ADH4 Alcohol dehydrogenase	0.39	0.42	0.38	0.54	0.56	0.47	0.69	0.65	0.51	1.42	down	1.94	3.33	3.39	3.99	5.20	3.68	4.93	4.30	10.99	0.E+00	alcohol biosynthetic process

[a]Proteins 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 scale at the top.

units of standardized deviation; using this variable, it was possible to analyze the joint behavior of all of the spectra in terms of a unique distribution (Figure 1, inset). When the quantitative data obtained from the analysis of the NH samples were processed using this procedure, the standardized variables describing the variability between different spectra within the same peptide (z_{eqps}, Figure 1A) closely followed a normal distribution with zero mean and unit variance (Figure 2, left column), demonstrating the accuracy of the model and the validity of the NH at the spectrum level for all SIL methods and MS machines.

Errors produced at the peptide and protein levels are treated in the WSPP model by assuming the simplest case (i.e., that they are normally distributed with zero mean and constant variance). This statement is reflected by the two fixed resistances in the peptide and protein branches (σ^2_{Pe} and σ^2_{Se}, respectively) in the electrical circuit (Figure 1). The validity of this double assumption was demonstrated by the accuracy with which the corresponding standardized variables (z_{eqp} and z_{eq}) follow the expected normal distribution with zero mean and unit variance in all SIL methods (Figure 2, center and right columns). In the two SILAC cases, because of the metabolic nature of this kind of labeling, it was not possible to compare identical replicates, and a pseudonull hypothesis (A* vs A) was used instead. Although, as expected, some minor abundance changes were detected (Table 1), the distribution of standardized variables at the peptide and protein levels from the SILAC experiments were also in excellent agreement with the predictions of the model (Figure 2, center and right columns).

Interestingly, the variances estimated at the three levels in the different SIL approaches (Tables 1 and 2) reflect the specific characteristics of each method. For instance, iTRAQ quantifications by PQD have a greater variance at the spectrum level than those obtained by TOF/TOF, consistent with the fact that the latter produce higher-intensity reporter ions that produce more accurate readings. However, their peptide variances are similar, as expected for methods that share a common peptide-preparation procedure. Likewise, SILAC variances at the peptide level are practically zero, as expected for a predigestion labeling method, whereas the other methods share a similar nonzero variance at the protein level. Finally, SILAC variances at the spectrum level were also the lowest among all SIL methods, reflecting the facts that SILAC quantification is directly performed in the MS spectrum and that the mass difference between labeled and unlabeled species is higher than that produced by ^{18}O labeling, thus minimizing quantification errors resulting from isotopomer superposition.

To test the WSPP algorithm further, we also analyzed whether it was possible to construct a statistical model containing fewer parameters. As described in Supporting Information section 3.1, the model had to take into account three sources of variance to be able to describe the results of all of the NH experiments, producing a negligible number of false protein abundance changes in all cases (Supporting Informa-

tion Figure 3). Finally, the performance of the WSPP model was compared with that of other commonly used statistical models specifically developed for ^{18}O[40], iTRAQ,[25,30] or SILAC[33] (Supporting Information section 3.2). We analyzed the accuracy of each model to describe the results obtained in the NH experiments (Supporting Information Figures 5−7), compared the performance of the WSPP model with that of a weighted least-squares method for iTRAQ[30] in terms of accuracy and FDR using a yeast background with spiked-in proteins at different concentrations and ratios (Supporting Information Figure 8), and performed a side-by-side comparison of WSPP and MaxQuant to determine protein-abundance changes produced by H_2O_2 treatment (Supporting Information Tables 5 and 6). Our results indicate that in spite of their general applicability the performance of the WSPP model was at least similar to that of these previous models.

WSPP Model Allows Comparison and Coherent Integration of Results Obtained from All Quantification Approaches, Increasing the Statistical Power of Protein Quantification

The analysis using the WSPP model of the relative quantification of control versus H_2O_2-treated samples produced a set of statistical parameters very similar to the NH experiments (compare Table 1 with Table 2), highlighting the robustness of the algorithm in a real quantitative experiment. The distribution of standardized variables at the spectral and peptide levels was also in good agreement with the null hypothesis (Supporting Information Figure 4), reinforcing the validity of the model. At the protein level, a significant number of quantifications (ranging from 5 to 39) behaved as outliers at a 5% FDR (Figure 3A−F and Tables 1 and 2), reflecting protein-abundance changes produced by the H_2O_2 treatment. Without an appropriate statistical hypothesis, these results could be only compared and validated for the 16 proteins that changed their abundance in at least two experiments. As shown in Supporting Information Table 3, there was a good agreement between the different experiments. These results were further confirmed by an independent analysis of some of the changing proteins by using label-free quantification (Supporting Information Table 4 and Supporting Information Figure 9). In spite of the consistency of these results, the observed changes provided very limited biological information, indicating only the activation of a first-line defense against oxidative conditions, probably through the H_2O_2-responsive transcriptional activators Yap1p, Skn7p, and Msn2/4p42, which induced expression of thiol homeostasis proteins such as cytochrome c peroxidase, thioredoxin 2, and glutathione peroxidase 2.

To judge the reproducibility and dispersion of results among different experiments and to integrate the quantitative information from all of them, the WSPP model introduces a fourth layer of analysis. In this layer, protein quantifications are averaged from the different experiments according to error propagation theory (Supporting Information), a procedure mathematically equivalent to setting up protein batteries in parallel (Figure 1B). Demonstrating the validity of this fourth layer, the distribution of the standardized variables describing the variability between different NH experiments within the same protein (Figure 3G) and between different proteins (Figure 3H) were exactly as expected from theory, and no significant protein-abundance changes were detected at a 5% FDR (Figure 3I). The interexperiment distribution of protein data from the control versus H_2O_2-treatment experiments also followed the expected trend (Figure 3K), demonstrating that quantifications from all of the experiments were, again, reproducible and therefore can be integrated (Table 3). The protein distribution, however, clearly deviated from the NH trend (Figure 3L), and 123 significant protein-abundance changes at 5% FDR were revealed in the tornado plot (Figure 3M). Hence, the number of significant abundance changes increased almost 1 order of magnitude when the results from the different experiments were integrated in the framework of the WSPP model, whereas there were no significant abundance changes when results from the NH experiments were integrated using the same approach. The consistency in the quantitative behavior of these significantly changing proteins in the original experiments is illustrated in Supporting Information Figure 10. This increase in statistical significance afforded by data integration allowed not only the detection of more upregulated proteins implicated in primary stress response but also detection of numerous proteins related to the ribosome and the mitochondrial membrane as well as downregulated proteins implicated mostly in metabolic functions such as glycolysis and gluconeogenesis (Table 3), a finding that agrees with other works.[41−44]

DISCUSSION

In this study, we present WSPP, a general statistical framework for the analysis of quantitative proteomics results, and we demonstrate that the statistical model very accurately describes the technical variability of data for a representative set that includes the most common SIL methods. In addition, the model allows a systematic comparison and integration of data from different experiments. The general validity of the model was demonstrated by the analysis of 16 quantitative experiments performed using six different combinations of SIL and MS approaches (Tables 1 and 2). The distribution of the data were analyzed at three levels (spectrum, peptide, and protein) so that a total of 48 experimental distributions were confronted against 18 different null hypotheses. This statistical framework efficiently resolves the problems of variance heterogeneity, data integration, and under-sampling and provides a statistically sound method for testing the quality of quantitative experiments and detecting experimental deviations. Moreover, our results reinforce the idea that quantitative SIL experiments, if properly performed and analyzed, are highly reproducible irrespective of the labeling technique or MS platform used.

At the spectral level, the model analyzes variance using a strategy different from variance-stabilization normalization procedures used to treat microarray[32] and iTRAQ data.[25,31] Instead of transforming the data, WSPP uses a two-parameter function to model the behavior of variance, and from this function, it directly assigns a variance to each one of the quantifications, keeping the original readings. This two-parameter modeling of variance is similar to that followed in a previous work to treat iTRAQ data,[28] although in that work, the final analysis was made at the spectral level and no integration of the data at the peptide or protein levels was performed. The general applicability of our model to any SIL and MS combination confirms in the most general case a fundamental property of MS-based peptide-centric quantifications: the error produced during the quantitative SIL analysis of peptides generating the same intensity at a given MS detector, irrespective of their sequence or molecular structure, is constant and normally distributed. We believe that this property is of paramount importance in the field and simplifies interpretation

dx.doi.org/10.1021/pr4006958 | J. Proteome Res. 2014, 13, 1234−1247

of quantitative data produced by MS. Finally, a model that analyzes specifically the variance at the spectral level has the advantage that it allows for separate control of the error produced during MS analysis and quantification from that produced at the time of peptide or protein preparation and thus may be used to check the proper functioning and calibration of the MS machines and even to detect chromatographic shifts resulting from incomplete coelution of peptide pairs. For instance, some of the MS conditions used in this work, such as the collision energy in PQD fragmentation, were optimized by selecting the ones that produced the minimum variance at the spectral level.

Our results also suggest that, at least using the protocol followed here, the error produced during peptide preparation can be considered constant and normally distributed, reinforcing previous results we obtained using other biological systems[15] and peptide-preparation methods.[34] In the general case, the null hypothesis formulated here provides the basis for testing the validity of this assumption for any peptide preparation method. Analysis of variance at the peptide level may be very useful in practice to detect deviations from the expected behavior during peptide preparation, such as those produced by artifacts like partial digestion or methionine oxidation.[15] However, it may also be used to assign a statistical significance level to abundance changes in postranslationally modified peptides, such as Cys sites subjected to oxidative modification, as we demonstrated in a recent work.[45] It could also be potentially used to detect effects because of the presence of SNPs or differential splicing. In addition, the protein layer can also be ignored in our model so that statistical analysis of abundance changes is directly performed at the peptide level, without grouping peptides into proteins.

Finally, the error at the protein level has been shown to follow the same trend in this and other biological systems,[15] making it a very convenient starting point from which to analyze other error sources, such as biological variance, which is highly dependent on experimental design and must be modeled separately in each case. Besides, because the variances at the protein level estimated in all of the SIL experiments performed in this work (Tables 1 and 2) are very similar to those calculated in previous studies using ^{18}O labeling in several biological models,[15,34] including tissue extracts, and in general are below 0.01, this value may be used as a reference to determine whether the protein variability in a given preparation is higher than that expected for a conventional protein-manipulation method. Thus, an increase in protein variance above the reference value not accompanied by a concomitant increase in peptide and spectrum variances indicates an increased heterogeneity in protein composition that is not related to peptide manipulation or MS quantification. This heterogeneity may indicate technical problems related, for instance, to a protein-preparation protocol involving too many steps, but it may also reflect a high biological variability between the samples that are compared, as we have found in a previous work[46] and also when comparing human samples extracted from different individuals (unpublished data). Note that the majority of existing models to analyze SIL data integrate the data intro protein averages, without taking into account the variance at the lower levels, and then analyze the distribution of protein quantifications as a whole;[25,26,31] in doing so, all of the technical error sources are comprised in only one random variable, and it is not possible to interpret results in terms of the variance at the protein level.

One of the most important characteristics of the WSPP model is that it provides a general framework to make a full integration of quantitative and error information from one level to a superior level. When several spectra are integrated into a peptide average, the model takes into account the variance associated with each one of the spectra according to error propagation theory so that the most accurate have a more significant contribution. This procedure simplifies the interpretation of data because quantifications of poor quality have a negligible effect on the peptide average and do not need to be eliminated from the analysis. However, the variance assigned to each one of the peptides takes into account not only the variance at the spectrum level, which is diminished because of averaging several spectra, but also the intrinsic variance associated with the process of peptide preparation. The same is done when peptides are integrated into proteins; although peptide averaging diminishes the variance carried out from the spectrum and peptide levels, the protein average includes the variance produced by protein manipulation. In this sense, the constant variances at the spectrum, peptide, and protein levels can be conceived as asymptotic errors because they reflect the lowest error that can be achieved at each one of the levels (see horizontal bars in Figure 2). This property of data integration is illustrated in the equivalent circuit (Figure 1A), where resistances set in series reflect the additive nature of the variances associated to independent events, whereas those set in parallel reflect the effect of averaging on variances. Most importantly, this concept of data integration is of general validity and can be extended to upper levels, where the effect of averaging is reflected when protein readings from different experiments are considered together. The decrease in resistance (i.e., variance) of the data integrated at upper levels increase the statistical power to detect deviations from the null hypothesis. This explains why only a dozen of the significant protein-abundance changes are detected when the experiments were considered separately, whereas the alteration of more than 100 proteins becomes evident when the data are integrated at upper levels.

The use of weighted averages to calculate protein ratios is not new and reflects the current view that not all quantitative measurements have the same accuracy.[25−29,33] However, the weighting scheme followed by our model is different from other approaches in several aspects. Although other methods have been proposed that separately estimate the biological and technical variance components,[4,6] to the best of our knowledge and with the exception of its predecessor,[34] no models have been formulated previously for quantitative proteomics data that decompose technical variance into two or more components. Besides, in our model, the averages are calculated following error propagation theory so that the statistical weight with which each value contributes to the average is exactly the inverse of its local variance, and the variances of each one of the averaged values are known with accuracy. A similar kind of weighting by the error made at the spectrum and peptide levels was proposed in one of the earliest models put forth to analyze quantitative data produced by using stable isotope-coded affinity tags (ICAT);[47] however, the approach that was followed in that work only propagates the error produced at the time of quantification from the MS spectrum and does not take into account the variance introduced by further errors produced by peptide generation and protein manipulation. Finally, to the best of our knowledge, our weighting method is

the first one that has been demonstrated to be of general validity for a wide range of different SIL and MS approaches.

The WSPP model also resolves the problem of undersampling and at the same time provides a framework to integrate data and a robust algorithm to estimate variances, which are of general applicability. This is accomplished by using standardized variables at each one of the integration levels. These variables express $\log_2$ ratios in units of standard deviation, introducing a bias correction for the number of degrees of freedom, which depends on the number of elements that are used to compute the average (i.e, the number of peptides that are used to estimate a protein value). Using the standardized variables, all of the elements at a given level of integration can be analyzed together in a unique distribution that, under the null hypothesis, is expected to be a normal distribution with unit variance (Figure 1, central inset). We take advantage of this general property to make robust estimates of variances by an iterative method that is of general applicability for all integration levels and quantification approaches. We should note here that our approach to estimate variance is similar to that used by other authors,[28,31,33] although in these works, only the total technical variance was estimated. Analysis of the standardized variable is also very useful to detect the presence of outliers at any integration level; in the WSPP model, this may be done at seven different levels (z parameters in Figure 1), and in each one, it provides specific information. Thus, at the spectrum and peptide levels, outliers are indicative of incorrect quantifications produced by a variety of causes (for instance, peak coelution, bad fitting, or methionine oxidation[15,34]), as commented above. At the protein level, outliers indicate statistically significant abundance changes. However, at other levels of integration, an outlier may also indicate that an experimental replicate gives quantitative results that deviate from the other replicates more than expected by chance alone and the absence of outliers indicate that all the replicates behave as expected by the null hypothesis (as observed in this work). We should note here that this global conception of variances, which considers together the whole wealth of data, is in contrast with other approaches commonly followed to detect outliers and artifacts, like Dixon's test,[47] which locally analyze the distribution of the elements used to calculate a particular average and which have a very limited utility when the number of elements is very low (i.e., when a protein is quantified by only two peptides). In the WSPP model, all of the elements, even single hits, are assigned a local variance, and this is done on the basis of only four parameters that are estimated from the analysis of the whole collection of data.

CONCLUSIONS

We present a statistical framework that explains the behavior of data obtained using the most common SIL-MS approaches. The WSPP model allows a systematic comparison and integration of data from different SIL experiments, and, in spite of its general applicability, its performance is at least similar to other commonly employed methods. We demonstrate the importance of performing rigorous data integration to uncover subtle but widespread protein changes taking place in a proteome, using *S. cerevisiae* exposed to a low dose of H_2O_2 as an example. Thus, comparison of the results at the protein level revealed activation only of first-line responses in response to the treatment, whereas data integration from different SIL experiments using the WSPP framework uncovered more subtle changes in groups of proteins related to the ribosome and mitochondria and also to metabolic pathways. These results highlight the importance of establishing an adequate and validated statistical framework for the analysis of high-throughput quantitative data. We believe that the WSPP model may contribute to developing a general standard for the analysis of quantitative proteomics data obtained by stable isotope labeling.

ASSOCIATED CONTENT

Supporting Information

Additional materials and methods including details of the WSPP statistical model; evidence that three independent sources of variance are necessary in the WSPP model; comparison of the WSPP model with other existing algorithms; elements of the WSPP statistical model; definition of fitting weights; comparison of significant protein abundance changes in the control versus H_2O_2-treatment experiments; list of proteins and peptides subjected to a label-free parallel reaction monitoring; side-by-side comparison of protein abundance changes in the control versus H_2O_2-treatment SILAC-HR experiment determined by MaxQuant and WSPP; comparison of protein abundance changes obtained in the control versus H_2O_2-treatment experiments using WSPP and in the SILAC-HR experiment using MaxQuant; analysis of local normality; estimation of the weight constant for the different SIL experiments; 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; analysis at the spectrum, peptide, and protein levels of the controls versus H_2O_2-treatment experiments; analysis of the high- and low-resolution 18O NH experiments using UNiquant algorithm; analysis of the TOFTOF and PQD iTRAQ NH experiments using the variance-stabilizing normalization strategy; analysis of the SILAC-HR, pseudo-NH experiment using MaxQuant; performance of the WSSP model in the analysis of spike-in mixtures of known protein concentrations in a complex background; verification of some protein abundance changes in the control versus treatment experiments by label-free parallel reaction monitoring; comparison of changing ribosomal, oxidative stress, and metabolic proteins in the control versus treatment experiments; complete list of proteins quantified by the different SIL-MS approaches for the null hypothesis; complete list of proteins quantified by the different SIL-MS approaches for the controls versus treatment experiments; complete list of protein quantifications integrated from the different SIL-MS approaches for the null hypothesis and for the controls versus treatment experiments; and information about the meaning of the parameters used in QuiXoT. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jvazquez@cnic.es. Phone: (+34) 91 4531200. Fax: (+34) 91 4531245.

Author Contributions

◆These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants BIO2009-07990, BIO2009-11735, BFU2009-08004, SAF 2009-07520, and BIO2012-37926 from the Spanish Ministry of Science and Education, CAM BIO/0194/2006 (Cardiovrep) from the Madrid regional government, and an institutional grant from the Fundación Ramón Areces to the CBMSO. Grants RD06/0014/0030, RD12/0042/0021, RD06/0014/0005, and RD12/0042/0022 from the Red Temática de Investigación Cooperativa en Enfermedades Cardiovasculares (RECAVA/RIC, Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III, Ministry of Health) supported the research of J.V. and J.M.R. P.M.-A. is recipient of a fellowship from the Madrid regional government supported by the European Social Fund. We thank S. Bartlett for language editing.

■ REFERENCES

(1) Aebersold, R. A stress test for mass spectrometry-based proteomics. *Nat. Methods* **2009**, *6*, 411−412.
(2) Nilsson, T.; Mann, M.; Aebersold, R.; Yates, J. R., 3rd; Bairoch, A.; Bergeron, J. J. Mass spectrometry in high-throughput proteomics: Ready for the big time. *Nat. Methods* **2010**, *7*, 681−685.
(3) Domon, B.; Aebersold, R. Options and considerations when selecting a quantitative proteomics strategy. *Nat. Biotechnol.* **2010**, *28*, 710−721.
(4) Clough, T.; Key, M.; Ott, I.; Ragg, S.; Schadow, G.; Vitek, O. Protein quantification in label-free LC-MS experiments. *J. Proteome Res.* **2009**, *8*, 5275−5284.
(5) 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*, M111.014662-1−M111.014662-13.
(6) Daly, D. S.; Anderson, K. K.; Panisko, E. A.; Purvine, S. O.; Fang, R.; Monroe, M. E.; Baker, S. E. Mixed-effects statistical model for comparative LC-MS proteomics studies. *J. Proteome Res.* **2008**, *7*, 1209−1217.
(7) Karpievitch, Y.; Stanley, J.; Taverner, T.; Huang, J.; Adkins, J. N.; Ansong, C.; Heffron, F.; Metz, T. O.; Qian, W. J.; Yoon, H.; Smith, R. D.; Dabney, A. R. A statistical framework for protein quantitation in bottom-up MS-based proteomics. *Bioinformatics* **2009**, *25*, 2028−2034.
(8) Oberg, A. L.; Vitek, O. Statistical design of quantitative mass spectrometry-based proteomic experiments. *J. Proteome Res.* **2009**, *8*, 2144−2156.
(9) Polpitiya, A. D.; Qian, W. J.; Jaitly, N.; Petyuk, V. A.; Adkins, J. N.; Camp, D. G., 2nd; Anderson, G. A.; Smith, R. D. DAnTE: A statistical tool for quantitative analysis of -omics data. *Bioinformatics* **2008**, *24*, 1556−1558.
(10) Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **1995**, *1*, 289−300.
(11) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002**, *1*, 376−386.
(12) Ong, S. E.; Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* **2006**, *1*, 2650−2660.
(13) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3*, 1154−1169.
(14) Mirgorodskaya, O. A.; Kozmin, Y. P.; Titov, M. I.; Korner, R.; Sonksen, C. P.; Roepstorff, P. Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1226−1232.
(15) 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*, M110.003335-1−M110.003335-14.
(16) Cox, J.; Mann, M. Is proteomics the new genomics? *Cell* **2007**, *130*, 395−398.
(17) Gan, C. S.; Chong, P. K.; Pham, T. K.; Wright, P. C. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J. Proteome Res.* **2007**, *6*, 821−827.
(18) Unwin, R. D.; Pierce, A.; Watson, R. B.; Sternberg, D. W.; Whetton, A. D. Quantitative proteomic analysis using isobaric protein tags enables rapid comparison of changes in transcript and protein levels in transformed cells. *Mol. Cell. Proteomics* **2005**, *4*, 924−935.
(19) Rodriguez-Suarez, E.; Gubb, E.; Alzueta, I. F.; Falcon-Perez, J. M.; Amorim, A.; Elortza, F.; Matthiesen, R. Virtual expert mass spectrometrist: iTRAQ tool for database-dependent search, quantitation and result storage. *Proteomics* **2010**, *10*, 1545−1556.
(20) Boehm, A. M.; Putz, S.; Altenhofer, D.; Sickmann, A.; Falk, M. Precise protein quantification based on peptide quantification using iTRAQ. *BMC Bioinf.* **2007**, *8*, 214-1−214-18.
(21) Hill, E. G.; Schwacke, J. H.; Comte-Walters, S.; Slate, E. H.; Oberg, A. L.; Eckel-Passow, J. E.; Therneau, T. M.; Schey, K. L. A statistical model for iTRAQ data analysis. *J. Proteome Res.* **2008**, *7*, 3091−3101.
(22) Oberg, A. L.; Mahoney, D. W.; Eckel-Passow, J. E.; Malone, C. J.; Wolfinger, R. D.; Hill, E. G.; Cooper, L. T.; Onuma, O. K.; Spiro, C.; Therneau, T. M.; Bergen, I.; Robert, H. Statistical analysis of relative labeled mass spectrometry data from complex samples using ANOVA. *J. Proteome Res.* **2008**, *7*, 225−233.
(23) Oberg, A. L.; Mahoney, D. W. Statistical methods for quantitative mass spectrometry proteomic experiments with labeling. *BMC Bioinf.* **2012**, *13*, S7-1−S7-18.
(24) Herbrich, S. M.; Cole, R. N.; West, J.; Keith, P.; Schulze, K.; Yager, J. D.; Groopman, J. D.; Christian, P.; Wu, L.; O'Meally, R. N.; May, D. H.; McIntosh, M. W.; Ruczinski, I. Statistical inference from multiple iTRAQ experiments without using common reference standards. *J. Proteome Res.* **2013**, *12*, 594−604.
(25) 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*, 1885−1897.
(26) Lin, W. T.; Hung, W. N.; Yian, Y. H.; Wu, K. P.; Han, C. L.; Chen, Y. R.; Chen, Y. J.; Sung, T. Y.; Hsu, W. L. Multi-Q: A fully automated tool for multiplexed protein quantitation. *J. Proteome Res.* **2006**, *5*, 2328−2338.
(27) Shadforth, I. P.; Dunkley, T. P.; Lilley, K. S.; Bessant, C. i-Tracker: For quantitative proteomics using iTRAQ. *BMC Genomics* **2005**, *6*, 145-1−145-6.
(28) Zhang, Y.; Askenazi, M.; Jiang, J.; Luckey, C. J.; Griffin, J. D.; Marto, J. A. A robust error model for iTRAQ quantification reveals divergent signaling between oncogenic FLT3 mutants in acute myeloid leukemia. *Mol. Cell. Proteomics* **2010**, *9*, 780−790.
(29) Bantscheff, M.; Boesche, M.; Eberhard, D.; Matthieson, T.; Sweetman, G.; Kuster, B. Robust and sensitive iTRAQ quantification on an LTQ orbitrap mass spectrometer. *Mol. Cell. Proteomics* **2008**, *7*, 1702−1713.
(30) Mahoney, D. W.; Therneau, T. M.; Heppelmann, C. J.; Higgins, L.; Benson, L. M.; Zenka, R. M.; Jagtap, P.; Nelsestuen, G. L.; Bergen, I.; Robert, H.; Oberg, A. L. Relative quantification: Characterization of bias, variability and fold changes in mass spectrometry data from iTRAQ-labeled peptides. *J. Proteome Res.* **2011**, *10*, 4325−4333.
(31) Arntzen, M. O.; Koehler, C. J.; Barsnes, H.; Berven, F. S.; Treumann, A.; Thiede, B. IsobariQ: Software for isobaric quantitative proteomics using IPTL, iTRAQ, and TMT. *J. Proteome Res.* **2011**, *10*, 913−920.

(32) 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*, S96−S104.
(33) 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*, 1367−1372.
(34) Jorge, I.; Navarro, P.; Martinez-Acedo, P.; Nunez, E.; Serrano, H.; Alfranca, A.; Redondo, J. M.; Vazquez, J. Statistical model to analyze quantitative proteomics data obtained by $^{18}O/^{16}O$ labeling and linear ion trap mass spectrometry: Application to the study of vascular endothelial growth factor-induced angiogenesis in endothelial cells. *Mol. Cell. Proteomics* **2009**, *8*, 1130−1149.
(35) Gruhler, A.; Olsen, J. V.; Mohammed, S.; Mortensen, P.; Faergeman, N. J.; Mann, M.; Jensen, O. N. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* **2005**, *4*, 310−327.
(36) 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*, 1135−1145.
(37) Vizcaino, J. A.; Cote, R. G.; Csordas, A.; Dianes, J. A.; Fabregat, A.; Foster, J. M.; Griss, J.; Alpi, E.; Birim, M.; Contell, J.; O'Kelly, G.; Schoenegger, A.; Ovelleiro, D.; Perez-Riverol, Y.; Reisinger, F.; Rios, D.; Wang, R.; Hermjakob, H. The PRoteomics IDEntifications (PRIDE) database and associated tools: Status in 2013. *Nucleic Acids Res.* **2013**, *41*, D1063−D1069.
(38) Lopez-Ferrer, D.; Ramos-Fernandez, A.; Martinez-Bartolome, S.; Garcia-Ruiz, P.; Vazquez, J. Quantitative proteomics using $^{16}O/^{18}O$ labeling and linear ion trap mass spectrometry. *Proteomics* **2006**, *6*, S4−S11.
(39) 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*, 1274−1286.
(40) 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*, 1228−1237.
(41) Herrero, E.; Ros, J.; Belli, G.; Cabiscol, E. Redox control and oxidative stress in yeast cells. *Biochim. Biophys. Acta* **2008**, *1780*, 1217−1235.
(42) Godon, C.; Lagniel, G.; Lee, J.; Buhler, J. M.; Kieffer, S.; Perrot, M.; Boucherie, H.; Toledano, M. B.; Labarre, J. The H_2O_2 stimulon in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **1998**, *273*, 22480−22489.
(43) Molina-Navarro, M. M.; Castells-Roca, L.; Belli, G.; Garcia-Martinez, J.; Marin-Navarro, J.; Moreno, J.; Perez-Ortin, J. E.; Herrero, E. Comprehensive transcriptional analysis of the oxidative response in yeast. *J. Biol. Chem.* **2008**, *283*, 17908−17918.
(44) McDonagh, B.; Ogueta, S.; Lasarte, G.; Padilla, C. A.; Barcena, J. A. Shotgun redox proteomics identifies specifically modified cysteines in key metabolic enzymes under oxidative stress in *Saccharomyces cerevisiae*. *J. Proteomics* **2009**, *72*, 677−689.
(45) Martinez-Acedo, P.; Nunez, E.; Gomez, F. J.; Moreno, M.; Ramos, E.; Izquierdo-Alvarez, A.; Miro-Casas, E.; Mesa, R.; Rodriguez, P.; Martinez-Ruiz, A.; Dorado, D. G.; Lamas, S.; Vazquez, J. A novel strategy for global analysis of the dynamic thiol redox proteome. *Mol. Cell. Proteomics* **2012**, *11*, 800−813.
(46) Bonzon-Kulichenko, E.; Martinez-Martinez, S.; Trevisan-Herraz, M.; Navarro, P.; Redondo, J. M.; Vazquez, J. Quantitative in-depth analysis of the dynamic secretome of activated Jurkat T-cells. *J. Proteomics* **2011**, *75*, 561−571.
(47) Li, X. J.; Zhang, H.; Ranish, J. A.; Aebersold, R. Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal. Chem.* **2003**, *75*, 6648−6657.