

# A quantitative $^1\text{H}$ NMR approach for evaluating the metabolic response of *Saccharomyces cerevisiae* to mild heat stress

Francesc Puig-Castellví<sup>1</sup> · Ignacio Alfonso<sup>2</sup> · Benjamí Piña<sup>1</sup> · Romà Tauler<sup>1</sup>

Received: 23 January 2015 / Accepted: 20 May 2015 / Published online: 26 May 2015  
© Springer Science+Business Media New York 2015

**Abstract** Effect of growth temperature on the yeast (*Saccharomyces cerevisiae*) metabolome has been analysed by one-dimensional proton NMR spectroscopy ( $^1\text{H}$  NMR). Potential biomarkers have been first identified by a non-targeted chemometric evaluation of the spectra, followed by a comprehensive analysis of bayesian estimated concentrations of target metabolites in extracts of cells growth either at 30 or 37 °C. Tentative identification of metabolites whose concentrations were affected by this mild heat-shock stress was attempted by partial least squares-discriminant analysis (PLS-DA) on  $^1\text{H}$  NMR data, combined with Statistical Total Correlation Spectroscopy, and further confirmed with empirical data. An extensive assignment for most of the detected NMR signals was performed, with a total number of 38 identified metabolites. Concentrations estimated using automatic BATMAN modelling revealed that bayesian integration is a sufficient approach for obtaining relevant concentration changes of metabolites and biological information of interest. In contrast to when it is applied directly on spectral data, the application of PLS-DA on BATMAN recovered metabolite concentration estimates allowed for a better overview of the investigated

samples, since more metabolites were highlighted in the discriminatory model. Observed changes in metabolite concentrations were consistent with the expected process of temperature acclimation, showing alterations in amino acid cellular pools, nucleotide metabolism and lipid composition. The strategy described in this work can thus be proposed as a powerful and easy tool to investigate complex biological processes, from biomarker screening and discovery to the study of metabolite network changes in biological processes.

**Keywords**  $^1\text{H}$  NMR · BATMAN · Yeast · Temperature · OSC-PLS-DA

## 1 Introduction

Metabolomics is a field of ‘omics’ research that is primarily focused on the identification and characterisation of small molecule metabolites in cells, tissues, organs and organisms (German et al. 2005). Commonly, nuclear magnetic resonance (NMR) (Griffin 2003) and mass spectrometry (MS) (Dettmer et al. 2007) are used for metabolomics studies. NMR studies mostly include  $^1\text{H}$  NMR (Mazzei et al. 2013) and bidimensional heteronuclear proton-carbon (HSQC) (Kang et al. 2012) whereas MS studies include direct infusion (Højer-Pedersen et al. 2008) and hyphenated to chromatography techniques (Farrés et al. 2015). Less frequently, LC- $^1\text{H}$  NMR studies are also reported, mostly in the phytochemistry area (Wolfender et al. 2005).

The choice between NMR and MS lies in the evaluation of the pros and cons for both techniques. First, MS sensitivity is higher than that of NMR, which usually allows identification of only 30–50 metabolites. Second, molecule

**Electronic supplementary material** The online version of this article (doi:10.1007/s11306-015-0812-9) contains supplementary material, which is available to authorized users.

✉ Romà Tauler  
Roma.Tauler@idaea.csic.es

<sup>1</sup> Department of Environmental Chemistry, IDAEA-CSIC, Jordi Girona 18-26, 08034 Barcelona, Catalonia, Spain

<sup>2</sup> Department of Biological Chemistry and Molecular Modelling, IQAC-CSIC, Jordi Girona 18-26, 08034 Barcelona, Catalonia, Spain

assignment capacity in MS depends directly on the instrument accuracy and, without MS<sup>[n]</sup>, compounds can only be tentatively assigned, as molecular isomers would have identical molecular masses. For hyphenated chromatographic methods with MS, metabolite assignment can be performed by their metabolite retention time and  $m/z$  values. However, the power of resolving peaks depends strictly on the capacity of the chromatographic method, usually the limiting step. This problem can be circumvented by the use of chemometric approaches like multivariate curve resolution (MCR) (Farrés et al. 2015). Traditionally, metabolite identification in NMR spectra depended heavily on the NMR-spectrometrist's previous knowledge and in the ability of identifying the spectral patterns of each metabolite. Several NMR databases (Jewison et al. 2012; Wishart et al. 2013) with  $^1\text{H}$  and  $^{13}\text{C}$  spectra of biological compounds have been recently implemented, significantly helping in speeding-up the assignment process. Resolution of overlapping signals in NMR spectrum is not always as crucial as in (MS) chromatographic methods. If an NMR spectrum region cannot be resolved, it is possible that the resonance signals comprised in that region come from metabolites with equivalent signals in other regions of the spectrum, making the related metabolites equally estimable. And third, signals in NMR spectra have inherently absolute quantitative information, whereas MS requires a calibration curve for each compound, due to the variable ionisation rates of the different metabolites.

*Saccharomyces cerevisiae* is a key organism in both traditional (wine, beer, bread) and technological (bioethanol) fermentation processes. Growth temperature is a key parameter determining the yeast metabolism. For most strains, growth is optimal between 25 and 30 °C, whereas ethanol production is favoured by somewhat higher temperatures (Barnett et al. 2000; Mensonides et al. 2013). However, high temperatures affect negatively other aspects of yeast metabolism, including changes in membrane lipid composition to adapt its fluidity to the growth temperature (Arthur and Watson 1976). The compromise between ethanol production and yeast survival is solved in industrial procedures by using thermotolerant yeast strains (Nonklang et al. 2008).

In a previous study combining transcriptomics and metabolomics studies (Strassburg et al. 2010), upregulated genes during heat stress in yeast were mostly related to primary metabolism processes. Knowing in advance that the majority of compounds detectable by NMR belong to primary metabolites, and due to the fact that any metabolic approach is tightly dependent on the metabolomics strategy employed (Dunn et al. 2005), in this study we propose to use an NMR-based metabolomics approach to study the yeast thermal response to a mild heat stress.

There are two possible approaches to analyse NMR spectra and both are presented in this work. The first strategy is the non-targeted analysis of the whole set of NMR spectra through chemometric methods, like principal component analysis (PCA) (Bro and Smilde 2014) and partial least squares discriminant analysis (PLS-DA) (Wold et al. 2001), with the aim of identifying discriminatory peaks that can be related to particular metabolites, and if required, their corresponding peak areas from the raw spectrum can be compared to estimate the relative changes in their concentrations. The second approach is the assignment of the NMR spectrum peaks to a set of target metabolites (prior to know whether they are relevant or not in the studied effect) and the integration of their corresponding signals, followed by the application of chemometric methods to the resulting peak-assigned metabolite areas data matrix.

Whilst the direct non-target approach is faster and brings the information to discriminate among tested classes, the traditional targeted approach allows a better interpretation of the tested samples. However, metabolites assignment and confirmation is time-consuming and estimation of their area can be sometimes complicated due to overlapping signals. Also, manual peak integration might produce misleading results caused by human factor. To overcome with these possible problems of both type of approaches, in this work we tested the Batman approach (Hao et al. 2012), currently available in open source in R (R Core Team 2013), which uses a Bayesian approach to estimate the concentration of metabolites from previously assigned NMR peaks. This approach has already been confirmed to provide better and more accurate results than traditional manual integration approaches (Aistle et al. 2012).

## 2 Materials and methods

### 2.1 Experimental

#### 2.1.1 Yeast growth

*Saccharomyces cerevisiae* BY4741 (MATa; *his3Δ1*; *leu2Δ0*; *met15Δ0*; *ura3Δ0*) cells were pre-cultured in YPD medium on an orbital shaker (150 rpm) at 30 °C overnight. A 500 ml of YPD medium was inoculated with the pre-culture to an optical density at 600 nm (OD600) of 0.1 and divided into eight aliquots of 50 ml. Aliquots were grown at either 30 or 37 °C (four aliquots each, shaking at 150 rpm). After 6 h of cultivation, at 1.4–1.6 of OD600, cultures were arrested on ice and 4 fractions of 10 ml were collected from each culture for metabolite extraction. Growth was similar for cultures grown at both temperatures (not shown). Cell harvesting was performed by

centrifugation of each 10-ml fraction at  $4000\times g$  for 5 min and discarding the supernatant. Cells were washed afterwards with 1 ml of 100 mM sodium phosphate buffer (pH 7.0). The resulting pellets were stored at  $-80\text{ }^{\circ}\text{C}$  and lyophilised. For both studied temperatures, 15 pellets were analysed.

### 2.1.2 Metabolite extraction

Metabolites were extracted by using a slight modification of the chloroform–methanol extraction protocol (Palomino-Schätzlein et al. 2013). Pellets were resuspended in 1800  $\mu\text{l}$  of a cold ( $4\text{ }^{\circ}\text{C}$ ) methanol–chloroform (1:2) solution by vigorous vortexing and submerged into liquid nitrogen for 1 min, and afterwards thawed in ice for 2 min. The process was repeated by a total of five times. Next, 400  $\mu\text{l}$  of water were added to create a biphasic system and homogenised by vortexing. Organic and aqueous phases were separated by centrifugation at 16,500 rpm (3 min,  $4\text{ }^{\circ}\text{C}$ ). The upper aqueous phase was collected, and the process was repeated once. The combined aqueous phases were afterwards lyophilised.

### 2.1.3 NMR sample preparation

Extracts were dissolved in 700  $\mu\text{l}$  of deuterated phosphate buffer ( $\text{Na}_2\text{DPO}_4$  100 mM, pH 7.0) in  $\text{D}_2\text{O}$  with DSS 0.2 mM as internal standard. The resulting solution is placed into the NMR tube.

### 2.1.4 NMR experiments

Spectra were recorded in a 400 MHz Varian spectrometer, using a spectrometer frequency of 400.14 MHz with a OneNMR Probe and a ProTune System (Agilent). Spectral size range covered from  $-2$  to 10 ppm, consisting of 65,536 data points. The number of scans was 512 and the relaxation delay was 5 s. For bidimensional TOCSY experiments, 8 scans and 512 t1 increments were used.

### 2.1.5 NMR spectra preprocessing

Spectra were preprocessed with MestreNova v.9.0 (Mestrelab Research, Spain). Spectra preprocessing consisted in an exponential apodization of 0.5 Hz, a manual phasing and a baseline correction with Bernstein polynomial of 3rd order. After adjusting the reference to DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), regions of water (4.7–5.1 ppm), methanol (3.30–3.37 ppm) and chloroform (7.64–7.69 ppm) were removed. Data points with chemical shift higher than 9.45 ppm or lower than 0.8 ppm were also removed. The final NMR dataset consisted on a data matrix of 30 spectra (rows) having 38,213 ppm values (columns)

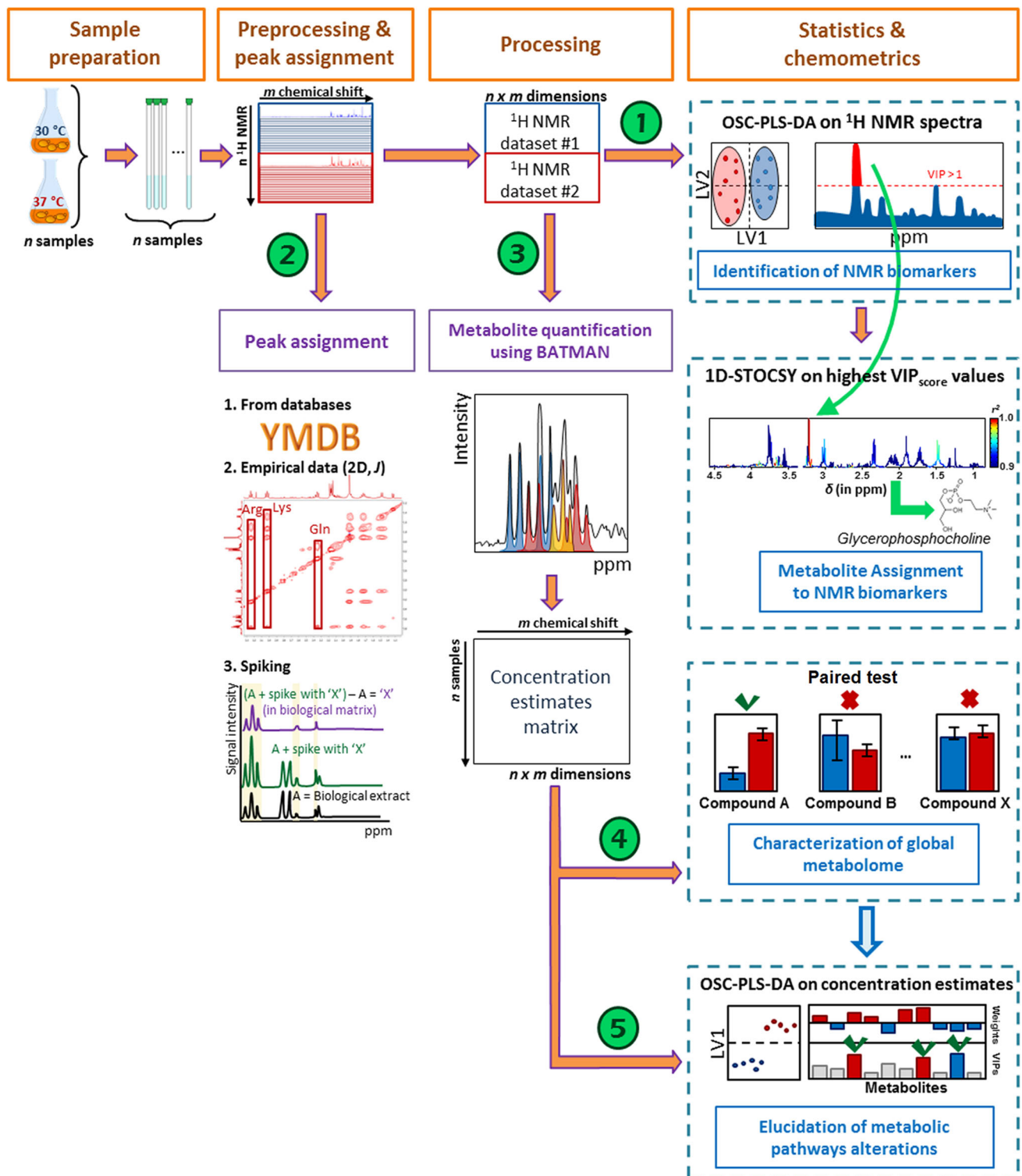
each one. This data matrix was stored in an ASCII file format.

### 2.1.6 Metabolite identification

A preliminary 1D-STOCSY (Cloarec et al. 2005) detection of correlated peaks with significant variation between the two temperatures was performed on the most relevant proton resonances indicated by VIPs score plot from the OSC-PLS-DA (see OSC-PLS-DA in Sect. 2.2) of the raw  $^1\text{H}$  NMR spectral data matrix (step 1 in Fig. 1). This step does not only serve as a biomarker discovery strategy, but it can be also applied to do a tentative metabolite assignment since intramolecular signals are likely to be shown as highly correlated in the 1D-STOCSY. Metabolite assignment (step 2 in Fig. 1) was performed by a detailed targeted metabolite profiling analysis of the  $^1\text{H}$  NMR and 2D-TOCSY yeast extract samples, using the  $^1\text{H}$  NMR spectra library developed during this work and also using the Yeast Metabolome Data Base library (Jewison et al. 2012) (YMDB). Some of the assigned metabolites were further confirmed by spiking using commercial standards. Online Resource 1 presents the list of methods used for each identification.

### 2.1.7 Metabolite quantification

Relative metabolite quantification was performed using BATMAN R-package (step 3 in Fig. 1). BATMAN implements a procedure based on the use of a bayesian model to deconvolute  $^1\text{H}$  NMR peaks and automatically assigns these peaks to metabolites from a target list, whose concentration estimate expressed in arbitrary units is obtained. BATMAN uses the modelled resonances of each assigned proton to reconstruct the empirical NMR spectrum, whereas the noise signal and unassigned resonances are modelled by wavelets. Proton signals that showed first-order couplings patterns were modelled using their spectrometric parameters. In other cases, multiplets were modelled as a ‘raster’ multiplet, using in-house  $^1\text{H}$  NMR spectra, or as a set of singlets. The list of parameters used to model each proton signal is given in Online Resource 2. BATMAN allows to interpret data directly from spectroscopic parameters of weak coupling spin systems ( $\text{AX}$ ,  $\text{A}_2\text{X}$ ,...). However, for strongly coupled spin systems, alternative input parameters have to be chosen. Thus, most of the signals in Online Resource 2 with non-integer proton intensities correspond to  $\text{AB}$  or to  $\text{AA}'\text{XX}'$  proton systems modelled as two  $\text{AX}$  systems. D-Glucose proton intensities are referenced as the sum of the two existing anomeric forms. All these non-integer proton intensities, including trehalose and glycerol proton intensities as well, were calculated by manual integration



**Fig. 1** Workflow of the NMR metabolomics data analysis procedure used in this work. Different steps are identified by the numbers inside the circles

with MestreNova software (Mestrelab Research, Spain). Finally, in order to let the fitting less restrictive and to obtain a better performance, Glutamine signal **23** (in

Online Resource 2) was defined as a sum of singlets, each one containing the weighted intensity of the whole multiplet. Additional parameters related to the bayesian

algorithm were also introduced. Some of these input parameters are summarised in Online Resource 3, while for the rest of input parameters their default value was used. More details of the procedure are described elsewhere (Hao et al. 2014). After introducing the input data, each NMR metabolite pattern should be correctly fitted in its corresponding  $^1\text{H}$  NMR experimental spectrum from the yeast extract sample. By performing this step, we concluded that the modelling of every proton signal is not necessary for achieving a good estimate. For instance, the use of overlapped signals when the same metabolite has isolated signals within the NMR spectra implies that more calculations are required for obtaining the same concentration estimate.

Running time depends directly on the width of the chemical shift region studied. To shrink dataset dimension, only 1 out of each 10 values from the 38,213 spectral set points were used. In addition, instead of using the whole spectral dimension, the analysis was performed by selecting those spectral regions comprising a number of multiplets and removing regions containing only noise. In Fig. 2b, d, f and g, the individual spectral regions used are shown. Thus, although all  $^1\text{H}$  NMR spectra were always processed simultaneously, metabolite concentrations were estimated sequentially depending on the region analysed each time.

Metabolite estimation performance from overlapped signals was evaluated carefully by observing the residual wavelet and contrasting whether non-overlapped signals from the same metabolites have been affected or not by wavelet penalisation. Therefore, additional regions containing peak signals from the same metabolite were always fitted if possible. For instance, when peaks from region of 1.15–1.35 ppm are fitted, both L-lactic acid and L-threonine signals were cross-checked by also adding into the fitting model their respective resonances from the proton at  $\text{C}_\alpha$  and  $\text{C}_\beta$ , respectively, which are located above 4 ppm in the NMR spectrum. In Fig. 2a, c, e and g, a stack plot from a representative sample showing the sum of deconvoluted areas and the residual part is presented.

### 2.1.8 Pathway analysis and transcriptome data mining

Metabolites highlighted in the OSC-PLS-DA (see Sect. 2.2) of the concentration estimates were correlated with known yeast metabolic pathways using the KEGG database (Kanehisa et al. 2012). Transcriptomic data comparison between yeast cell growing at 29 and 36 °C was obtained from <http://genome-www5.stanford.edu> (reference GSM1046). Metabolic maps were also obtained from the KEGG web page (<http://www.genome.jp/kegg/kegg2.html>).

## 2.2 Chemometric data preprocessing and analyses

### 2.2.1 Theory

A brief description of the different statistical and modelling methods used in this work to analyse  $^1\text{H}$  NMR spectra is given below.

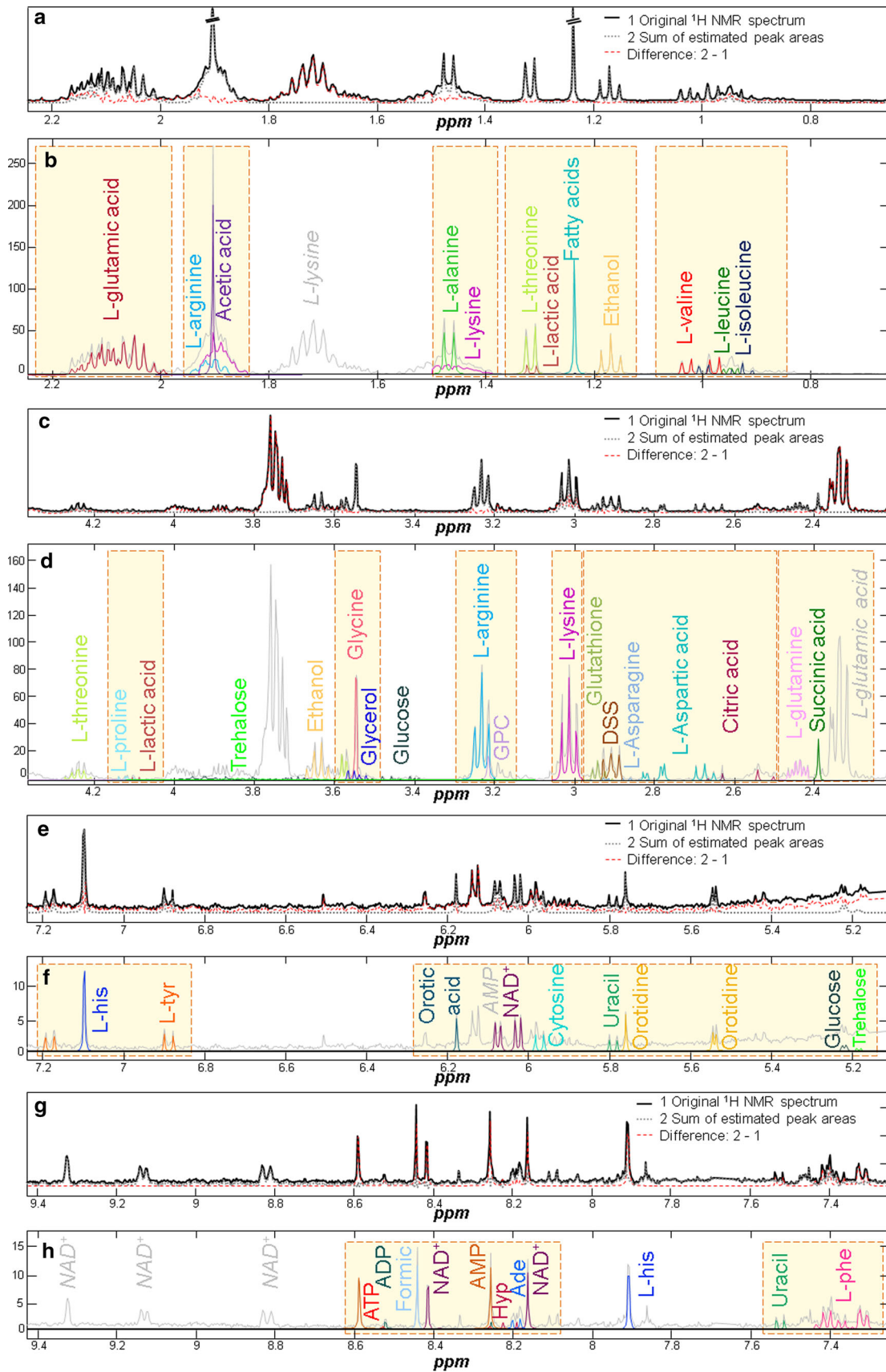
**2.2.1.1 Statistical total correlation spectroscopy (STOCSY)** In STOCSY (Cloarec et al. 2005), the autocorrelation matrix for a set of 1D NMR spectra of samples containing mixtures of metabolites is calculated. High correlations values might reveal which proton signals belong to the same molecule. STOCSY serves as a helpful tool for pattern recognition in complex mixtures (Li et al. 2014). In NMR spectra from biological extracts, if two or even more metabolites are strongly correlated, these correlations might be an evidence of biochemically connected metabolites. However, we have to be aware that these correlation values might be also distorted by spectral overlap and noise interferences.

**2.2.1.2 OSC-PLS-DA** PLS-DA (Barker and Rayens 2003) is a PLS regression method (Geladi and Kowalski 1986) variant that allows to correlate a set of discriminant response  $\mathbf{y}$ -variables to a set of correlated  $\mathbf{X}$ -variables (in this studied case, the temperature and the NMR spectral matrix, respectively). As a result, a reduced number of new linear combination of the independent original  $\mathbf{X}$ -values (called latent variables, LV) is obtained. This LV optimally correlate with the variation in  $\mathbf{y}$ -variables. For every LV, a vector of weight coefficients is acquired, showing which  $\mathbf{X}$ -variables affect most on  $\mathbf{y}$ .

Orthogonal signal correction (OSC) (Wold et al. 1998) is a signal pretreatment method that removes any variation within the  $\mathbf{X}$  dataset uncorrelated (orthogonal) to  $\mathbf{y}$ . In this work, the variable importance on projection (VIP) scores (Wold et al. 2001), which give a weighted sum of squares of PLS weights for each variable, were also calculated in order to determine the most influent variables in the model.  $\mathbf{X}$ -variables associated with VIP scores greater than one are considered to be relevant on the definition of the PLS-DA model (Chong and Jun 2005), since the average of squared VIP score is equal to 1.

### 2.2.2 Multivariate analysis of the NMR spectra

The spectra from the previous preprocessing steps were imported to Matlab R2013a (The Mathworks Inc. Natick,



**Fig. 2**  $^1\text{H}$  NMR 400 MHz spectrum of the aqueous cell extracts of *S. cerevisiae* grown at 30 °C. **a**, **c**, **e** and **g** stack plots represent the fit for the modelled resonances obtained. In *black* it is shown the empirical  $^1\text{H}$  NMR spectrum; in *dotted grey*, the sum of the modelled resonances; and in *dashed red*, the difference between the empirical and the modelled spectra. On the other hand, **b**, **d**, **f** and **h** stack plots show the individual models of each one of the used proton resonances. The *orange-dashed boxes* represent the different regions used. Coloured peak signals out of the *box* represent those signals used only as a cross-check validation of the fitting performance. Metabolite names in *grey* represent some of the identified but not used proton resonances. Abbreviated metabolites: *ADP* adenosine diphosphate, *Ade* adenine, *AMP* adenosine monophosphate, *ATP* adenosine triphosphate, *DSS* 4,4-dimethyl-4-silapentane-1-sulfonic acid, *Formic* formic acid, *GPC* glycerophosphocholine, *Hyp* hypoxanthine, *L-his* L-histidine, *L-phe* L-phenylalanine, *L-tyr* L-tyrosine (Color figure online)

MA, USA) and analysed with the PLS toolbox 7.3.1 (Eigenvector Research Inc., Wenatchee, WA, USA). Principal component analysis (PCA) was applied to the NMR spectral data matrix after standard normal variate (SNV) (Dong et al. 2011) normalisation and mean-centering pre-processing performed in this order. Although dried pellets have similar weights ( $6.9 \pm 0.3$  mg for pellets collected from samples cultured at 30 °C, and  $6.7 \pm 0.8$  mg for 37 °C cultured samples), and according to *p* value significance level they are statistically identical (*p* value = 0.32), if spectra were normalised by dry weight, data would be biased. In order to overcome this difficulty, spectral data (with solvent and DSS regions already removed) was normalised with SNV. PLS-DA of the SNV, plus OSC-normalised spectra and mean-centered data has been also performed. In both cases, PCA and PLS-DA, cross-validation with Leave-One-Out system has been applied to check the reliability of the obtained models. Additionally, two permutation tests and a leave-one-culture-out cross-validation were executed. In the two permutation tests applied to the OSC-PLS-DA model, the evaluated parameters were the cumulative predicted residual sums of squares and the number of misclassifications. With leave-one-culture-out cross-validation test, the similarity among samples from different cultures was tested. In this test, an OSC-PLS-DA model was built using all samples except those from one specific culture as a training set, and these excluded samples were used afterward as a validation set. This process was repeated for all the cultures.

### 2.2.3 Analysis on relative concentration estimates

Concentration changes between classes (low and high culture temperature) were evaluated using a Mann–Whitney–Wilcoxon test and their corresponding levels of significance were calculated (step 4 in Fig. 1). In order to

contrast this information, a Kruskal–Wallis one-way analysis of variance was applied to the same dataset.

PCA and OSC-PLS-DA were also applied to these autoscaled concentration estimates (step 5 in Fig. 1). Missing concentration estimates were imputed by PCA as implemented in PLS Toolbox. The same permutation tests and leave-one-culture out cross-validation as in the NMR spectral data were performed to the OSC-PLS-DA model.

## 3 Results and discussions

### 3.1 Workflow

A scheme summarising the complete data analysis workflow of the acquired  $^1\text{H}$  NMR data in the analysis of yeast samples at the two temperatures is shown in Fig. 1.

### 3.2 Preliminary study of NMR spectral data

Before performing OSC-PLS-DA, NMR spectra were analysed by PCA. PCA scores plot of the mean-centered  $^1\text{H}$  NMR spectra from the yeast extract samples shows clearly that yeast metabolomes obtained for the two growth temperatures are distinguishable. The separation of the two sample classes was accomplished with the first component (PC1), containing 54.28 % of the explained variance (Online Resource 4a). PCA model using 3 components explained 86.58 % of the X-variance.

In PCA plot of the 2 first components, all tested samples laid inside the 95 % confidence level. Although two samples have Q residuals over the 95 % of confidence and one more is over the 95 % of confidence of Hotelling  $T^2$  (Online Resource 4b), they were not considered outliers, as their removal did not show any significant improvement on explained variance nor contributed significantly to the residuals. Scores of the yeast culture batches were randomly distributed, implying that none of the cultures had an outlying response at the growth temperature. This lack of any distribution pattern was also observed for colour-coded scores according to their fractional sampling order (data not shown). Therefore, the fractioning step did not either constitute any source of variance nor produced any sample bias.

### 3.3 Initial NMR signal assignment of potential biomarkers

When OSC-PLS-DA was applied to the mean-centered  $^1\text{H}$  NMR spectra from the yeast extract samples, one latent variable (LV) was sufficient to discriminate between the two sample classes. 53.36 % of the X-variance was already

able to explain the 99.99 % of the y-variance (Online Resource 4c). Scores were very constant within each sample class. Predictive capacity was totally achieved, with a 100 % of sensitivity and selectivity. If cross-validation is performed, correct classification is obtained for the whole sample dataset as well. Low RMSEC and RMSECV values were obtained (Online Resource 5). The permutation tests confirmed the reliability of the model with a  $p < 0.001$ , and the leave-one-culture-out cross-validation shown a total predictive capacity for each one of the cultures used as a validation set.

From the VIP score plot (Online Resource 4d), 83 lorentzian-shaped signals were observed whose intensity corresponds to  $\text{VIP} > 1$ , all located below  $\delta = 5$  ppm except for one at  $\delta = 8.443$  ppm. This fact explains why only half of the total X variance was needed to explain the y variance.

Since VIPs were associated to chemical shifts and not to individual NMR signals, chemical shifts associated to overlapped signals would give different VIP scores than those from isolated signals. Another appreciation from the direct examination of VIP scores plot is that while some of the VIP scores were close to 1, other reach values of some orders of magnitude higher, even greater than 1000 (Online Resource 4e), indicating significant unequal fold-changes of some resonance signals occurring during the stress situation.

A preliminary assignment of those metabolites whose concentrations were more affected by the temperature can be achieved by representing the corresponding 1D-STOCSY plots of the list of lorentzian-shaped NMR peaks selected by the OSC-PLS-DA VIPs. In STOCSY plots, spectrum intensities are differently coloured depending on their  $r^2$  correlation with the selected chemical shift. An example of 1D-STOCSY for the peak at 3.218 ppm, which is the chemical shift where the maximum VIP value was found, is represented in Fig. 3.

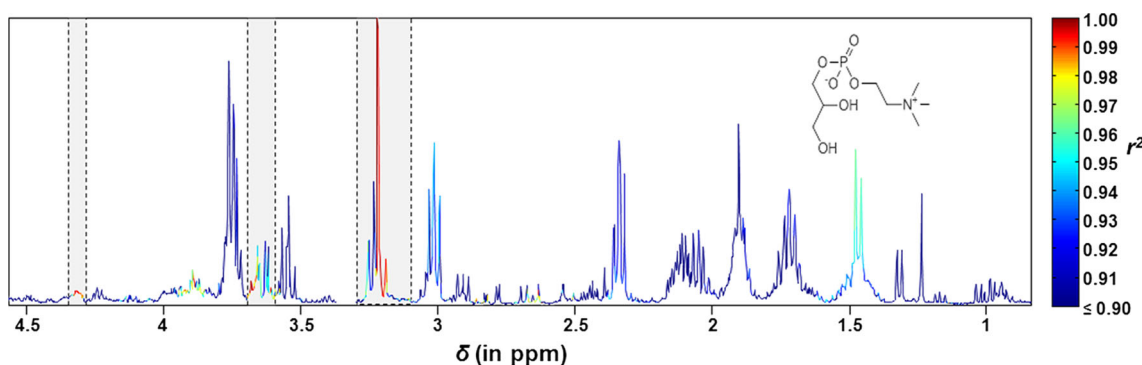
As observed in Fig. 3, two  $^1\text{H}$  NMR regions (around 3.67 and 4.31 ppm) were candidates of containing signals

from intramolecular protons connected to the one associated to the target resonance at 3.218 ppm. Using the YMDB, we tentatively assigned the peak of 3.218 ppm to be from L-glycerophosphocholine, since all three selected signals are also present in the  $^1\text{H}$  NMR spectrum of this metabolite. Using the same strategy, ethanol, DSS, L-threonine, L-lysine, glycerol and also glycine were tentatively assigned. Also, peaks from L-leucine, L-valine and L-isoleucine are found to be highly correlated, which agrees with the fact that they are biochemically linked.

### 3.4 Metabolite identification in $^1\text{H}$ NMR spectra

Metabolite assignment is not straightforward. Although STOCSY is a useful tool for identifying compounds in NMR spectra, total assignment of the resonances cannot be achieved because correlation coefficient values are negatively affected by spectral overlapping signals and noise interferences. From the interpretation of bidimensional NMR spectra, additional assignment of correlated signals can be obtained, but the total assignment is hardly obtained due to the overlapping problem. NMR databases represent a convenient tool to assign metabolite names to resonances, but slight chemical shifts from the ones reported in library databases are expected. Even if pH is controlled, other parameters such as differences in ionic strength may produce this variation. These peaks can be confirmed from spectroscopic data, although the absolute confirmation experiment is by metabolite spiking. However, in practice, this is not always possible. Therefore, a good knowledge on the interpretation of NMR spectroscopic data is required to do a proficient metabolite assignment.

In Fig. 2b, d, f and h, the metabolite identification in a  $^1\text{H}$  NMR 400 MHz experiment of a yeast cell extract is summarised. A wide range of biomolecules were identified, comprising  $\alpha$ -amino acids, sugars, nucleotides, organic acids and osmolytes. Online Resource 1 presents the list of methods used for each peak assignment. The number of



**Fig. 3** 1D-STOCSY plot at 3.218 ppm. Regions with segments with  $r^2$  correlations above 0.99 are encapsulated in the grey regions (Color figure online)



assigned metabolites in this study is on the average range of most NMR metabolomics studies, despite the fact that they use NMR spectrometers of higher magnetic fields, with higher sensitivity and resolution. In our case, however, the deconvolution process with BATMAN allowed to obtain each metabolite signal, regardless of the overlapping degree. In manual integration, some of the peaks are so small that artifacts might be produced because of the signal noise. By using BATMAN, the noise is modelled by wavelets and the signals from the compounds are more precisely estimated. This principle is also applied for peaks located in the shoulder of water signal, such as trehalose or glucose.

### 3.5 Metabolite quantification of yeast cell extracts

In this study we propose metabolite quantification by using a procedure based on bayesian integration. NMR metabolomics software based on Bayesian algorithms have been presented in previous literature, as the R-Package BQuant (Zheng et al. 2011). However, this software does not take profit of all the spectroscopic data from each signal. Other softwares can be used alternatively, such as the NMR suite-derived Chenomx NMR suite (Chenomx Inc., Canada). It has been previously stated that BATMAN “gave very comparable results to Chenomx, with  $r^2 > 0.996$  when templates were optimised” (Hao et al. 2014). In contraposition to Chenomx, BATMAN is a free and open source software. In this work the efficiency of BATMAN to obtain estimates of biologically interpretable data is tested.

Resonances used to estimate each metabolite concentration are shown in Online Resource 4, whereas their corresponding spectroscopic parameters are detailed in Online Resource 2.

Relative concentration estimates of the identified metabolites are presented in Online Resource 6. Due to deformations in peak shape, three concentration estimates of histidine were not estimated. Twenty eight metabolites showed significant concentration changes with  $p < 0.05$  due to temperature stress: at 37 °C, 22 metabolites (including cytosine, glutathione, ATP, formic acid and acetic acid, among others) had their concentration diminished when compared to 30 °C cultured samples, whereas six metabolites (including glycerophosphocholine and trehalose, among others) had increased their concentration. The same list of affected metabolites are significant with  $p < 0.05$  when the statistical test performed was the Kruskal–Wallis one-way analysis of variance.

In contrast with OSC-PLS-DA in the spectra matrix, more metabolites are highlighted now as being influenced by changes in temperature. Metabolite quantification revealed that glycerophosphocholine, one of the metabolites that had been pinpointed with STOCYSY (and associated

with a  $VIP_{score} = 1075.7$  in the OSC-PLS-DA), has low concentration in the sample (it is the 28th in the list of most abundant estimated metabolites). In the studied system, glycerophosphocholine concentration is increased by a factor of 5.57, but due to the fact that this molecule has 9 magnetically and chemically equivalent protons, its intensity value in the  $^1H$  NMR spectra is increased by approximately 50-fold, which explains its high VIP score obtained in the OSC-PLS-DA.

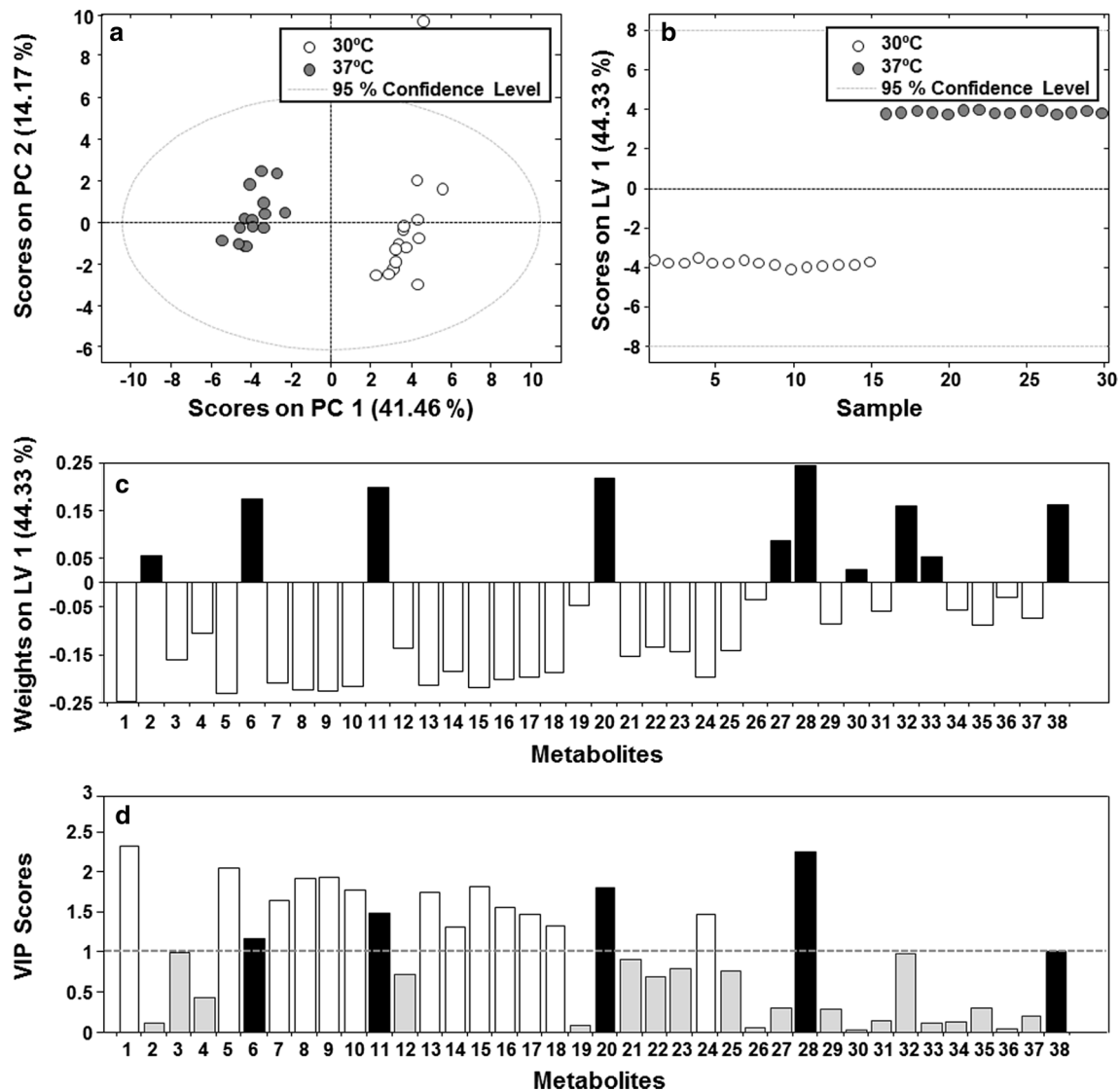
The tentative biomarker discovery strategy from the raw spectra matrix presented before is a convenient method to perform a tentative biomarker discovery, but should not be used for obtaining conclusions about general sample composition, since peak overlapping distorts the interpretation of the signals, and both OSC-PLS-DA and STOCYSY are negatively affected because of this. On the opposite, to perform a biomarker discovery strategy focused on concentration estimates is time-consuming in comparison to direct screening by PLS-DA on spectral data.

However, in our dataset, yeast metabolome is too stable within samples of the same class (Online Resource 4c), provoking that when the proposed strategy focused on the spectral data is applied, too many metabolic changes are highlighted instead of only those more intense. Since biomarkers should not be a large list of metabolites, by increasing the VIP score threshold in two orders of magnitude (dotted line in Online Resource 4d) the list of biomarkers was restricted to three (glycerophosphocholine, fatty acids and acetic acid). Since two of these metabolites are related to the lipid fraction, an alteration of this fraction due to the growth temperature is suggested.

### 3.6 Metabolic alterations observed by multivariate analysis of their concentration estimates

The first principal component in the preliminary PCA of the concentration estimates already explained 41.46 % of the total variance, and allowed the separation among the two classes of samples (those at 30 °C from those at 37 °C). PC2 (14.17 % of the variance) scores were homogeneously distributed within the 95 % of confidence limits, except for one sample. The residual value associated to that sample was below the limit but it had a high leverage (large hotelling  $T^2$  value). Variance explained by the model without this sample did not improve substantially (55.92 %). Scores projection for the 2 first PCs is represented in Fig. 4a.

OSC-PLS-DA LV1 on the autoscaled concentration estimates explained a 99.92 % of the y-variance using 44.33 % of the X-variance. All samples were correctly classified when cross-validation is performed. Low values of RMSEC and RMSECV were obtained (Online Resource 5). The permutation tests confirmed the reliability of the



**Fig. 4** **a** PCA scores projection on PC1 and PC2 subspace. *White* and *grey dots* correspond to 30 and 37 °C, respectively. **b** LV1 PLS-DA scores against each sample. **c** LV1 weights barplot. Negative weights are coloured in *white*, whereas positive weights are coloured in *black*. **d** VIP scores barplot. Bars with VIP values below threshold of 1 are coloured in *grey*. Other bars are coloured as their corresponding weight bars. Metabolites represented are the following: **1** L-threonine, **2** ethanol, **3** fatty acids, **4** L-glutamic acid, **5** L-glutathione, **6** L-histidine,

**7** formic acid, **8** NAD<sup>+</sup>, **9** L-valine, **10** AMP, **11** L-lysine, **12** L-arginine, **13** L-isoleucine, **14** acetic acid, **15** L-leucine, **16** L-phenylalanine, **17** cytosine, **18** L-glutamine, **19** orotidine, **20** L-alanine, **21** orotic acid, **22** L-lactic acid, **23** glycine, **24** L-aspartic acid, **25** uracil, **26** L-tyrosine, **27** glycerol, **28** glycerophosphocholine, **29** ADP, **30** L-proline, **31** D-glucose, **32** citric acid, **33** L-asparagine, **34** succinic acid, **35** ATP, **36** adenine, **37** hypoxanthine, **38** trehalose. *p* value significance levels: \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001 (Color figure online)

model with a *p* < 0.001, and the leave-one-culture-out cross-validation shown a total predictive capacity for each one of the cultures used as a validation set.

Score values distribution can be observed in Fig. 4b. In Fig. 4c, LV1 weights for each metabolite are displayed. According to LV1, negative weights (in white) show those metabolites with larger concentrations at 30 °C, whereas positive weights (in black) indicate the opposite (lower concentrations at this temperature). In Fig. 4d, VIP scores are given. They indicate those metabolites whose

concentrations have changed when culture temperature changed significantly (defined by a threshold of 1). Metabolites more important in the definition of the 37 °C class were **6** (L-histidine), **11** (L-lysine), **20** (L-alanine), **28** (glycerophosphocholine) and **38** (trehalose). For the 30 °C class, the metabolites of interest are **13**: **1** (L-threonine), **5** (glutathione), **7** (formic acid), **8** (NAD<sup>+</sup>), **9** (L-valine), **10** (AMP), **13** (L-isoleucine), **14** (acetic acid), **15** (L-leucine), **16** (L-phenylalanine), **17** (cytosine), **18** (L-glutamine) and **24** (L-aspartic acid).

Only a fraction of the metabolites identified as affected by growth temperature by the univariate statistical analysis (either at  $p < 0.05$  or  $p < 0.01$  confidence levels) was also detected as such by OSC-PLS-DA. This multivariate statistical analysis appeared thus to be more restrictive, probably because it models the whole set of metabolites for a given temperature at the same time, including their inter-replicate variability. Therefore, only those metabolites with concentration changes that are varying at unison within classes (but not among them) will be considered important for the model and highlighted as significant. On the contrary, in univariate statistical analysis, metabolites are evaluated one-by-one, whether or not their concentration change significantly, and coordinated metabolic variations cannot be detected.

Trehalose is one of the metabolites marked, both in the univariate and multivariate analysis, but it is not highlighted in the OSC-PLS-DA on NMR spectra. This divergence can be due to the effect of the applied pretreatment. Since NMR spectra were mean-centered instead of autoscaled, and trehalose is the identified metabolite with lowest concentration (Online Resource 6), its significance in the OSC-PLS-DA model was masked underneath other more concentrated significant metabolites. Therefore, in order to accomplish a biological interpretation from NMR data, it is preferred to perform the quantification of the effects on (autoscaled) metabolite concentrations rather than a direct interpretation from the spectrum.

### 3.7 Pathway analysis and biological interpretation

From a total of 28 metabolites whose concentrations varied between 30 and 37 °C and were present in the KEGG database (see below), 25 of them fell into four functional categories according to KEGG: amino acid metabolism (14 metabolites), nucleotide (purine and pyrimidine) metabolism (7 metabolites), respiration (TCA cycle and oxidative phosphorylation, 4 metabolites) and Pyruvate metabolism (4 metabolites). Some metabolites may belong to more than one category (Table 1). A comparison of these results with the transcriptomic analysis of yeast cells grown at 36 °C (GSM1046, <http://genome-www5.stanford.edu/>) revealed a rather good correlation between the metabolite changes and gene expression data (Table 1).

The figure in Online Resource 7 showed a summary of both metabolic and transcriptomic data. Note that affected metabolites and pathways are concentrated in particular areas of the total map, particularly at the top right (nucleotide metabolism), central-bottom right (amino acid metabolism) and at the centre of the map (energy gain, from sugar catabolism at the top, to oxidative phosphorylation at the bottom). In addition, most of affected metabolites (red and blue dots) are implicated in at least

one of the enzymatic steps catalyzed by a deregulated gene product, others lie shortly upstream or downstream of the same metabolic pathway (Online Resource 7).

Transcriptomic analysis showed deregulated genes in the fatty acid metabolism pathways (Online Resource 7). Although no KEGG-coded metabolite identified by NMR data analysis can be directly assigned to this pathway, it is like that the fatty acids singlet signal may be related to the alteration of the fatty acid pathway. Similarly, glycerophosphocholine concentration changes may be related to the deregulation of nine genes codifying enzymes of the Glycerophospholipid metabolism (Online Resource 7). In fact, different lipid composition changes are expected because of the type of experimental methodology applied to the analysis of yeast samples. The enhanced lipid biosynthesis allows compensating fluidity instabilities due to the changes in temperature (Sakamoto and Murata 2002; Torija et al. 2003).

Finally, glutathione showed a reduced concentration in 37 °C grown yeast cells, an effect probably related to the deregulation of different peroxidases and reductases involved in its metabolism (Table 1). Similarly, concentrations of  $\text{NAD}^+$ , related to the redox cell machinery, are notably diminished in cells grown at 37 °C, which may be related to a decreased mitochondrial function (McConnell et al. 1990). This putative redox potential alteration may be related to the low acetate/ethanol ratio observed at 37 °C relative to cells grown at 30 °C, as a consequence of the interconversion between ethanol and acetate controlled by the  $\text{NAD}^+$ -dependent aldehyde dehydrogenase (Racker 1949).

Trehalose, one of the few metabolites with elevated concentration at 37 °C, is not metabolically linked to the other altered metabolites, although the enzymatic reactions leading to its production and degradation are also catalyzed by deregulated gene products (Online Resource 7). Trehalose has been identified as a cellular protector to different stresses, including heat shock (Elbein et al. 2003; Estruch 2000). Our results (and others) suggest that trehalose concentration changes may be also part of the yeast acclimation process to high temperatures (Farrés et al. 2015; Strassburg et al. 2010).

Globally, our data showed a general decrease of many amino acid and nucleotide-related metabolites, whereas gene expression data of the same pathways indicate an overexpression of many enzyme-codifying genes. There is no easy way to translate changes in gene expression to changes in concentrations of metabolites, but it should be remembered that amino acids are rather strictly regulated, in part by negative feed-back mechanisms in which the absence of a given metabolite triggers transcription of the genes involved in the corresponding enzymatic pathway (Hahn and Young 2011; Hinnebusch 2005). Amino acid

**Table 1** KEGG classification of metabolites and related genes affected by temperature

KEGG pathway	Metabolites		Genes <sup>a</sup>	
	#	KEGG name	#	Official gene name
Biosynthesis of amino acids	14	L-Glutamate, glycine, L-alanine, L-lysine, L-aspartate, L-arginine, L-glutamine, L-phenylalanine, L-leucine, L-histidine, citrate, L-valine, L-threonine, L-isoleucine	18	<i>ALT1, ARO1, BAT2, CHA1, ERR1, ERR3, GLT1, GLY1, HIS3, IDH1, IDH2, LEU4, LYS20, LYS21, LYS9, NQM1, PYC1, SHM2</i>
Purine and pyrimidine metabolism <sup>b</sup>	7	ADP, AMP, cytosine, glycine, L-glutamine, orotate, uracil	19	<i>ADE17, ADE2, ADE8, ADO1, APA2, DAL1, GUK1, PDE1, PGM2, RNR1, RNR1, RPA190, RPA43, RPB11, RPO31, URA1, URA10, URA4, URA8</i>
Oxidative phosphorylation and TCA cycle <sup>b</sup>	4	NAD <sup>+</sup> , ADP, succinate, citrate	10	<i>FUM1, IDH1, IDH2, PDA1, PYC1, SDH2, SDH4, QCR2, RIP1, VMA8</i>
Pyruvate metabolism	4	Acetate, succinate, formate, (S)-lactate	10	<i>ACCI, ALD4, FUM1, GLO2, GLO4, LEU4, LYS20, LYS21, PDA1, PYC1</i>

<sup>a</sup> Data from dataset GSM1046, <http://genome-www5.stanford.edu>

<sup>b</sup> Combination of two KEGG categories

pools are known to decisively determine the capacity of de novo synthesis of proteins in yeast, and their composition may vary in processes leading to changes in the cell protein composition (Onodera and Ohsumi 2005; Suzuki 2013). At this point, it is important to point out that in our study, cells were allowed to acclimate to both temperatures for two to three cell doublings, and therefore the observed changes would not reflect a typical heat shock, but rather reflect an acclimation process i.e., the compensative reaction to the actual effects of the temperature, including changes in the global content in sugars, lipids and proteins, which in turn are reflected by changes in both the transcriptome and the metabolome.

#### 4 Concluding remarks

<sup>1</sup>H NMR combined with statistical multivariate data analysis tools has been able to produce a robust and reliable interpretation of the effects of temperature on yeast metabolism. Detection of significant changes on the concentration of some metabolites using OSC-PLS-DA of <sup>1</sup>H NMR spectra, in 1D-STOCSY format, is a feasible and rapid strategy.

By using Mann–Whitney–Wilcoxon's statistical test on the comparison between metabolite concentration estimates at the two temperatures, 28 metabolites were detected as significantly affected ( $p < 0.05$ ) by the temperature. From those, only 18 were highlighted on the corresponding PLS-DA, which appears to be more restrictive, probably because it includes the inter-replicate sample variability.

We consider that changes in glycerophosphocholine concentration may indicate more general changes in

membrane composition, likely related to the different growth temperature.

Apart from alterations on lipid membrane composition, other metabolic changes, such as the increase in the concentrations of some stress indicators like trehalose, were also detected. It is concluded that the combination of <sup>1</sup>H NMR metabolomics and multivariate data analysis tools are by themselves sufficient to obtain data interpretable from a biological point of view.

**Acknowledgments** The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement No. 320737.

**Conflict of interest** Francesc Puig-Castellví, Ignacio Alfonso, Benjamí Piña, and Romà Tauler declare that they have no conflict of interest.

**Ethical standard** This article does not contain any studies with human participants or animals performed by any of the authors.

#### References

- Arthur, H., & Watson, K. (1976). Thermal adaptation in yeast: Growth temperatures, membrane lipid, and cytochrome composition of psychrophilic, mesophilic, and thermophilic yeasts. *Journal of Bacteriology*, *128*, 56–68.
- Astle, W., De Iorio, M., Richardson, S., Stephens, D., & Ebbels, T. (2012). A Bayesian model of NMR spectra for the deconvolution and quantification of metabolites in complex biological mixtures. *Journal of American Statistical Association*, *107*, 1259–1271. doi:10.1080/01621459.2012.695661.
- Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, *17*, 166–173. doi:10.1002/cem.785.
- Barnett, J. A., Payne, R. W., & Yarrow, D. (2000). *Yeasts: Characteristics and identification* (3rd ed.). Cambridge: Cambridge University Press.

- Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods*, 6, 2812–2831. doi:10.1039/C3AY41907J.
- Chong, I.-G., & Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, 78, 103–112. doi:10.1016/j.chemolab.2004.12.011.
- Cloarec, O., Dumas, M.-E., Craig, A., et al. (2005). Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic <sup>1</sup>H NMR data sets. *Analytical Chemistry*, 77, 1282–1289. doi:10.1021/ac048630x.
- Core Team, R. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Detmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews*, 26, 51–78. doi:10.1002/mas.20108.
- Dong, J., Cheng, K.-K., Xu, J., Chen, Z., & Griffin, J. L. (2011). Group aggregating normalization method for the preprocessing of NMR-based metabolomic data. *Chemometrics and Intelligent Laboratory Systems*, 108, 123–132. doi:10.1016/j.chemolab.2011.06.002.
- Dunn, W. B., Bailey, N. J. C., & Johnson, H. E. (2005). Measuring the metabolome: Current analytical technologies. *Analyst*, 130, 606–625. doi:10.1039/B418288J.
- Elbein, A. D., Pan, Y. T., Pastuszak, I., & Carroll, D. (2003). New insights on trehalose: A multifunctional molecule. *Glycobiology*, 13, 17R–27R. doi:10.1093/glycob/cwg047.
- Estruch, F. (2000). Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. *FEMS Microbiology Reviews*, 24, 469–486. doi:10.1111/j.1574-6976.2000.tb00551.x.
- Farrés, M., Piña, B., & Tauler, R. (2015). Chemometric evaluation of *Saccharomyces cerevisiae* metabolic profiles using LC–MS. *Metabolomics*, 11, 210–224. doi:10.1007/s11306-014-0689-z.
- Geladi, P., & Kowalski, B. R. (1986). Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, 185, 1–17. doi:10.1016/0003-2670(86)80028-9.
- German, J. B., Hammock, B., & Watkins, S. (2005). Metabolomics: building on a century of biochemistry to guide human health. *Metabolomics*, 1, 3–9. doi:10.1007/s11306-005-1102-8.
- Griffin, J. L. (2003). Metabonomics: NMR spectroscopy and pattern recognition analysis of body fluids and tissues for characterisation of xenobiotic toxicity and disease diagnosis. *Current Opinion in Chemical Biology*, 7, 648–654.
- Hahn, S., & Young, E. T. (2011). Transcriptional regulation in *Saccharomyces cerevisiae*: Transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics*, 189, 705–736. doi:10.1534/genetics.111.127019.
- Hao, J., Astle, W., De Iorio, M., & Ebbels, T. M. D. (2012). BATMAN—an R package for the automated quantification of metabolites from nuclear magnetic resonance spectra using a Bayesian model. *Bioinformatics*, 28, 2088–2090. doi:10.1093/bioinformatics/bts308.
- Hao, J., Liebeke, M., Astle, W., De Iorio, M., Bundy, J. G., & Ebbels, T. M. D. (2014). Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN. *Nature Protocols*, 9, 1416–1427. doi:10.1038/nprot.2014.090.
- Hinnebusch, A. G. (2005). Translational regulation of GCN4 and the general amino acid control of yeast. *Annual Review of Microbiology*, 59, 407–450. doi:10.1146/annurev.micro.59.031805.133833.
- Højer-Pedersen, J., Smedsgaard, J., & Nielsen, J. (2008). The yeast metabolome addressed by electrospray ionization mass spectrometry: Initiation of a mass spectral library and its applications for metabolic footprinting by direct infusion mass spectrometry. *Metabolomics*, 4, 393–405. doi:10.1007/s11306-008-0132-4.
- Jewison, T., Knox, C., Neveu, V., et al. (2012). YMDB: The yeast metabolome database. *Nucleic Acids Research*, 40, D815–D820. doi:10.1093/nar/gkr916.
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, 40, D109–D114. doi:10.1093/nar/gkr988.
- Kang, W. Y., Kim, S. H., & Chae, Y. K. (2012). Stress adaptation of *Saccharomyces cerevisiae* as monitored via metabolites using two-dimensional NMR spectroscopy. *FEMS Yeast Research*, 12, 608–616.
- Li, M., Wang, J., Lu, Z., Wei, D., Yang, M., & Kong, L. (2014). NMR-based metabolomics approach to study the toxicity of lambda-cyhalothrin to goldfish (*Carassius auratus*). *Aquatic Toxicology*, 146, 82–92. doi:10.1016/j.aquatox.2013.10.024.
- Mazzei, P., Spaccini, R., Francesca, N., Moschetti, G., & Piccolo, A. (2013). Metabolomic by <sup>1</sup>H NMR spectroscopy differentiates “Fiano di Avellino” white wines obtained with different yeast strains. *Journal of Agriculture and Food Chemistry*, 61, 10816–10822.
- McConnell, S. J., Stewart, L. C., Talin, A., & Yaffe, M. P. (1990). Temperature-sensitive yeast mutants defective in mitochondrial inheritance. *Journal of Cell Biology*, 111, 967–976. doi:10.1083/jcb.111.3.967.
- Mensonides, F. I. C., Hellingwerf, K. J., de Mattos, M. J. T., & Brul, S. (2013). Multiphasic adaptation of the transcriptome of *Saccharomyces cerevisiae* to heat stress. *Food Research International*, 54, 1103–1112. doi:10.1016/j.foodres.2012.12.042.
- Nonklang, S., Abdel-Banat, B. M. A., Cha-aim, K., et al. (2008). High-temperature ethanol fermentation and transformation with linear DNA in the thermotolerant yeast *Kluyveromyces marxianus* DMKU3-1042. *Applied and Environmental Microbiology*, 74, 7514–7521. doi:10.1128/aem.01854-08.
- Onodera, J., & Ohsumi, Y. (2005). Autophagy is required for maintenance of amino acid levels and protein synthesis under nitrogen starvation. *Journal of Biological Chemistry*, 280, 31582–31586. doi:10.1074/jbc.M506736200.
- Palomino-Schätzlein, M., Molina-Navarro, M., Tormos-Pérez, M., Rodríguez-Navarro, S., & Pineda-Lucena, A. (2013). Optimised protocols for the metabolic profiling of *S. cerevisiae* by <sup>1</sup>H-NMR and HRMAS spectroscopy. *Analytical and Bioanalytical Chemistry*, 405, 8431–8441.
- Racker, E. (1949). Aldehyde dehydrogenase, a diphosphopyridine nucleotide-linked enzyme. *Journal of Biological Chemistry*, 177, 883–892.
- Sakamoto, T., & Murata, N. (2002). Regulation of the desaturation of fatty acids and its role in tolerance to cold and salt stress. *Current Opinion in Microbiology*, 5, 206–210. doi:10.1016/S1369-5274(02)00306-5.
- Strassburg, K., Walther, D., Takahashi, H., Kanaya, S., & Kopka, J. (2010). Dynamic transcriptional and metabolic responses in yeast adapting to temperature stress. *OMICS: A Journal of Integrative Biology*, 14, 249–259. doi:10.1089/omi.2009.0107.
- Suzuki, K. (2013). Selective autophagy in budding yeast. *Cell Death and Differentiation*, 20, 43–48.
- Torija, M. J., Beltran, G., Novo, M., et al. (2003). Effects of fermentation temperature and *Saccharomyces* species on the cell fatty acid composition and presence of volatile compounds in wine. *International Journal of Food Microbiology*, 85, 127–136. doi:10.1016/S0168-1605(02)00506-8.
- Wishart, D. S., Jewison, T., Guo, A. C., et al. (2013). HMDB 3.0—The human metabolome database in 2013. *Nucleic Acids Research*, 41, D801–D807. doi:10.1093/nar/gks1065.
- Wold, S., Antti, H., Lindgren, F., & Öhman, J. (1998). Orthogonal signal correction of near-infrared spectra. *Chemometrics and*

- Intelligent Laboratory Systems*, 44, 175–185. doi:[10.1016/S0169-7439\(98\)00109-9](https://doi.org/10.1016/S0169-7439(98)00109-9).
- Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58, 109–130. doi:[10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1).
- Wolfender, J.-L., Queiroz, E. F., & Hostettmann, K. (2005). Phytochemistry in the microgram domain—a LC–NMR perspective. *Magnetic Resonance in Chemistry*, 43, 697–709. doi:[10.1002/mrc.1631](https://doi.org/10.1002/mrc.1631).
- Zheng, C., Zhang, S., Ragg, S., Raftery, D., & Vitek, O. (2011). Identification and quantification of metabolites in  $^1\text{H}$  NMR spectra by Bayesian model selection. *Bioinformatics*, 27, 1637–1644. doi:[10.1093/bioinformatics/btr118](https://doi.org/10.1093/bioinformatics/btr118).