Molecular Systems Biology 6; Article number 349; doi:10.1038/msb.2010.3
Citation: ***Molecular Systems Biology*** **6:349**

www.molecularsystemsbiology.com

molecular systems biology

Revealing a signaling role of phytosphingosine-1-phosphate in yeast

L Ashley Cowart[1,3], Matthew Shotwell[2], Mitchell L Worley[1], Adam J Richards[1], David J Montefusco[1], Yusuf A Hannun[1,*] and Xinghua Lu[1,*]

[1] Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, SC, USA, [2] Department of Medicine, Medical University of South Carolina, Charleston, SC, USA and [3] Ralph H. Johnson Veteran's Affairs Medical Center, Charleston, SC, USA
* Corresponding authors. X Lu or YA Hannun, Department of Biochemistry and Molecular Biology, 174 Ashley Avenue, Charleston, SC 29425, USA.
Tel.: +1 843 876 1111; Fax: +1 843 876 1126; E-mail: lux@musc.edu or Tel.: +1 843 792 9318; Fax: +1 843 792 4322;
E-mail: hannun@musc.edu

Received 16.7.09; accepted 28.12.09

Sphingolipids including sphingosine-1-phosphate and ceramide participate in numerous cell programs through signaling mechanisms. This class of lipids has important functions in stress responses; however, determining which sphingolipid mediates specific events has remained encumbered by the numerous metabolic interconnections of sphingolipids, such that modulating a specific lipid of interest through manipulating metabolic enzymes causes 'ripple effects', which change levels of many other lipids. Here, we develop a method of integrative analysis for genomic, transcriptomic, and lipidomic data to address this previously intractable problem. This method revealed a specific signaling role for phytosphingosine-1-phosphate, a lipid with no previously defined specific function in yeast, in regulating genes required for mitochondrial respiration through the HAP complex transcription factor. This approach could be applied to extract meaningful biological information from a similar experimental design that produces multiple sets of high-throughput data.
Molecular Systems Biology **6**: 349; published online 16 February 2010; doi:10.1038/msb.2010.3
Subject Categories: functional genomics; signal transduction
Keywords: information integration; lipidomics; signal transduction; sphingolipids; transcriptomics

Introduction

Sphingolipids, a class of lipids found in all cell types across eukaryotic species, include bioactive molecules such as sphingosine-1-phosphate and ceramide (Zheng *et al*, 2006; Hannun and Obeid, 2008b). In budding yeast, the synthesis of sphingolipids increases acutely on heat stress and mediates cell survival at high temperature (Dickson *et al*, 1997; Jenkins *et al*, 1997). Several sphingolipid-dependent subprograms of the cell response to heat stress have been identified including cell cycle arrest, regulation of protein synthesis and degradation, and transcriptional reprogramming (Chung *et al*, 2000; Jenkins and Hannun, 2001; Cowart *et al*, 2003; Cowart and Hannun, 2005; Meier *et al*, 2006); however, these findings derive largely from studies wherein all sphingolipid synthesis is blocked, and thus, biological roles for specific lipid species remain unknown. Although specific genetic manipulations can be designed to determine the effects of deletion or over-expression of enzymes of sphingolipid metabolism, the interconnectedness of the sphingolipid metabolic network (as in many metabolic pathways) leads to widespread changes in lipid levels on single enzyme mutation, thus preventing attribution of the observed effects to specific lipid species. Here, we illustrate this problem in the yeast heat stress response. We then present a systems biology approach designed to reveal potential lipid-specific signaling pathways by deconvoluting a body of heterogeneous biological information derived from genomic, transcriptomic, lipidomic, and functional annotation data (Figure 1). The analysis indicates that phytosphingosine-1-phosphate (PHS1P), a sphingolipid with no previously known biological function in *Saccharomyces cerevisiae,* regulates the expression of genes involved in cellular respiration in a manner that requires the HAP2/3/4/5 transcription factor (TF) complex (Bonander *et al*, 2008). Biological validation of these findings indicated the systems analysis successfully identified a biological pathway mediated by PHS1P. Importantly, the methods and approaches developed in this study may be applicable to other metabolomic and genomic studies that generate high-throughput data sets across multiple platforms and where the interest is in the function of specific metabolites.

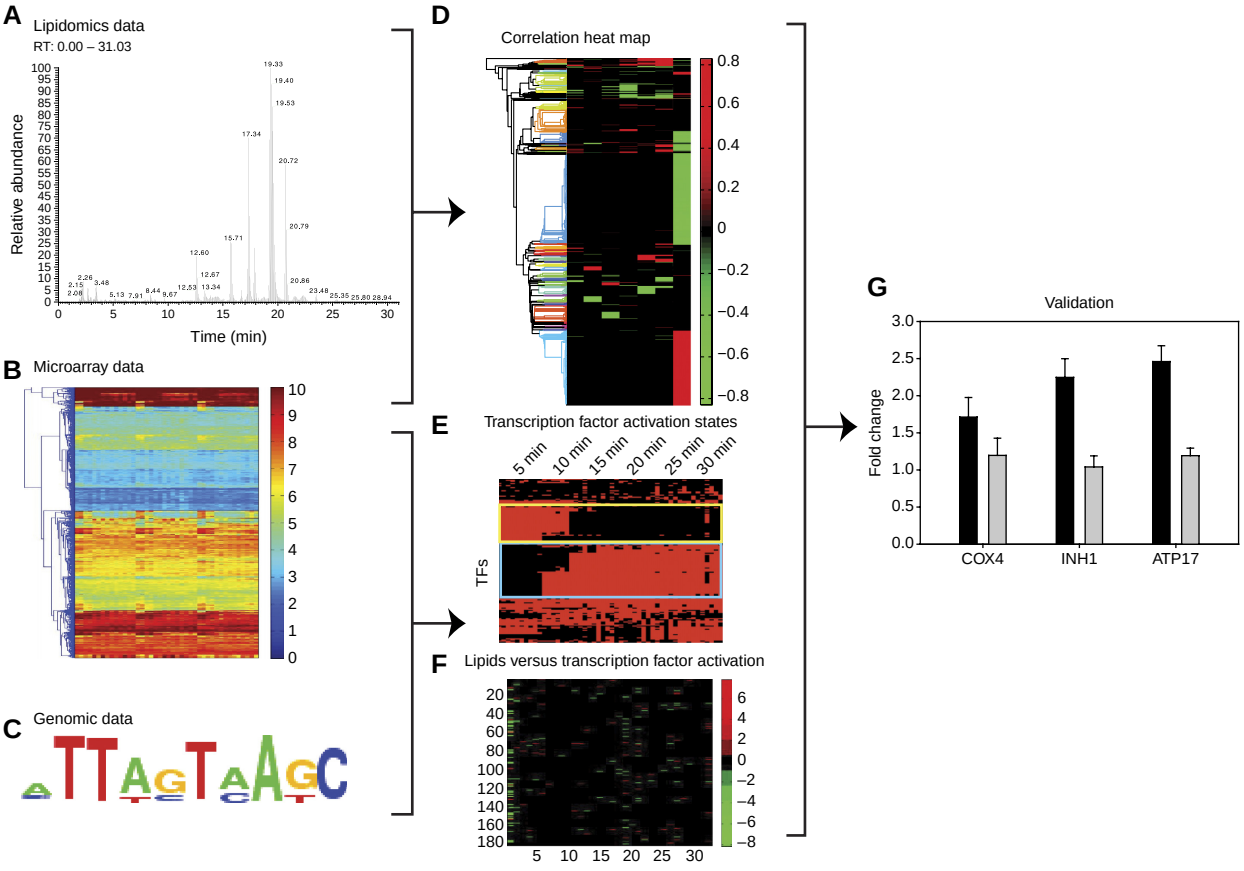

Figure 1 Overview of the integrative systems approach. The lipidomics, transcriptomic, and genomic data were collected from experiments and databases; (**A, B, C**) example data points or data matrix. Integrating the matching lipidomic and transcriptomic data in a correlation analysis lead to a gene-versus-lipid correlation coefficient matrix shown as a heat map shown in (**D**). Genomic and transcriptomic data were combined to infer the activation states of TFs under each experiment, shown as a TF-versus-condition heat map representing the activation states in (**E**). The inferred activation sates of TFs from (E) were combined with lipidomic data (A) to model the relationship between lipid mass and activation of TFs, shown as a heat map representing the significant logistic parameters in (**F**). The results from (E) and (F) resulted in the hypothesis that PHS1P mediated regulation of a subset of genes through activation of the HAP complex, which was tested in a series of genetic and pharmacological experiments (**G**).

Results

Sphingolipid metabolism

The initial metabolites resulting from sphingolipid biosynthesis include the sphingoid bases dihydrosphingosine (DHS) and phytosphingosine (PHS), which in turn serve as metabolic precursors for synthesis of an array of chemically diverse species including ceramides and sphingoid base phosphates (Figure 2). Though the homologous mammalian lipid sphingosine-1-phosphate activates a variety of signaling events through receptor-mediated modulation of characterized signaling pathways (Alvarez *et al*, 2007), specific cell functions for sphingoid base phosphates in yeast remain unknown. PHS1P was shown earlier to transiently increase during heat stress (Skrzypek *et al*, 1999), suggesting a potential role for this lipid in the heat stress response. Moreover, mutant yeast strains that accumulate PHS1P exhibited poor growth and heat stress resistance (Skrzypek *et al*, 1999; Kim *et al*, 2000), though a specific role for this lipid in the heat stress response remains unidentified. We hypothesized that PHS1P may mediate subprograms of the heat stress response, and, therefore, we designed experiments to perturb PHS1P metabolism using a conventional gene deletion approach. Deletion of LCB4 and LCB5 genes, encoding the yeast sphingoid base kinases (Nagiec *et al*, 1998), attenuates PHS1P production whereas deletion of DPL1, encoding the sphingoid base phosphate lyase (Saba *et al*, 1997), blocks its degradation and results in accumulation of PHS1P (Figure 2). Thus, mutating the above genes allows a 'clamp' on PHS1P in cells at either low or high levels, respectively. Log-phase cultures of the *lcb4*Δ/*lcb5*Δ, the *dpl1*Δ, and *wt* strains were subjected to heat stress, and samples were collected at 5-min intervals over a time course of 30 min in two duplicate experiments. To collect multiple '-omics' data reflecting cellular signaling systems, each sample was divided into two portions, thus allowing mRNA and sphingolipid extraction from each sample for transcriptomic and lipidomic analyses (Bielawski *et al*, 2006), generating a total of 42 transcriptomics measurements paired with 42 corresponding lipidomics measurements.

Transcriptomic responses to heat stress

As previously reported (Gasch *et al*, 2000; Gasch and Werner-Washburne, 2002; Cowart *et al*, 2003), heat stress significantly

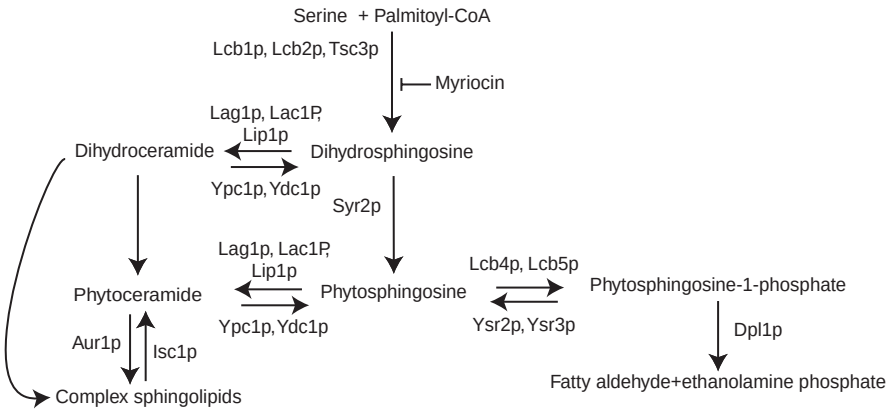

Figure 2 Summary of major sphingolipid biosynthetic pathways in *Saccharomyces cerevisiae*.

induced or repressed the expression of over a thousand genes in the *wt* strain, and a significant portion of this program was conserved in both the *lcb4*Δ*/lcb5*Δ and the *dpl1*Δ strains. However, we further identified the genes that are differentially expressed among the strains under comparable conditions. This group included 687 probe sets (corresponding to 441 genes with gene names) that were differentially expressed after heat stress in the *lcb4*Δ*/lcb5*Δ mutant when compared to *wt* strains. Analysis of this group using GOStat (Beissbarth and Speed, 2004) revealed a broad range of functional categories including DNA repair, translational regulation, post-translational modification of proteins, cell wall organization and biogenesis, and others (Figure 3A).

Lipidomics responses to heat stress

Recognizing the connectedness of sphingolipid metabolism (Alvarez-Vasquez *et al*, 2005; Hannun and Obeid, 2008a), we hypothesized that the mutations might cause broad changes in sphingolipid metabolism and that only a subset of these genes spanning broad functional categories actually resulted from the lack of PHS1P, whereas many of these genes' aberrant regulation might result from modulation of other sphingolipids in this mutant. Therefore, we further evaluated the lipidomic data to determine the impact of heat stress on lipid species and to determine whether the effects of mutations generate a metabolic 'ripple effect' leading to changes in other metabolites. Indeed, out of 40 distinct sphingolipid species measured, 28 demonstrated changes at least at one time point of heat stress in wild-type cells (Figure 3B). Heat stress induced an increase in C18 PHS1P from nearly undetectable levels to 0.02 pmol/nmol phosphate (see Supplementary Table 1), peaking around 20 min and returning to basal levels by 30 min. These kinetics were similar to previously published data that demonstrated an eight-fold increase in PHS1P that peaked around 10 min and returned to basal levels around 20 min (Skrzypek *et al*, 1999). As expected, these changes did not occur in the *lcb4*Δ*/lcb5*Δ mutant strain. On the other hand, the *dpl1*Δ strain showed constitutively elevated PHS1P, which further increased during heat stress (Figure 3B, bottom row). More importantly, the data revealed widespread differences in lipid profiles between the three strains, both basal and under heat stress. For example, the *lcb4*Δ*/lcb5*Δ strain demonstrated significant elevation of PHS, increased α-hydroxylated ceramides of 20 and 22-carbon *N*-acyl chain length, increased phytoceramides of 24 carbon chain length, and decreased dihydroceramide of 26 carbon chain length. In fact, out of the 28 lipid species showing changes in the wild type over the time course, at least 19 species showed differences between *wt* and the *lcb4*Δ*/lcb5*Δ strain at least one time point.

Lipidomics profiles of the *dpl1*Δ strain also demonstrated widespread differences as compared to the parental strain (Figure 3B, right panel). With the exception of the expected accumulation of PHS1P, lipid measurements in this strain revealed fewer and more subtle variation from the wild-type strain at basal temperature; however, time course measurements indicated that deletion of DPL1 significantly altered the heat stress sphingolipid response in that at least 20 of the 28 measured lipids exhibited differences from the parental strain in at least one point of the time course. Moreover, changes in this mutant were partially distinct from changes observed in the *lcb4*Δ*/lcb5*Δ mutant strain.

In summary, these mutations not only had the expected effects on PHS and PHS1P (substrate and product), but also caused widespread changes in many sphingolipid species. Therefore, any changes in gene regulation observed in the *lcb4*Δ*/lcb5*Δ or *dpl1*Δ mutant strains could not be readily attributed to PHS1P (or any single lipid species). To circumvent the limitations of 'gene-centric' approach, an alternative approach was devised to integrate information from lipidomic and transcriptomic data in a manner that allows inferring lipid-mediated events.

Indentify potential sphingolipid-regulated genes

To identify and quantify the information connecting lipidomic and transcriptomic changes, we performed covariance analysis (DeGroot and Schervish, 2002) between all lipid-probe pairs, which led to a *genes*-versus-*lipids* matrix in which an element contained a correlation coefficient (ρ) for a lipid–gene pair if the ρ is statistically significant or a zero otherwise. It was noted that a row of the matrix constituted a correlation (information) pattern demonstrated by a gene with respect to all lipids, and similarly a column encoded an information

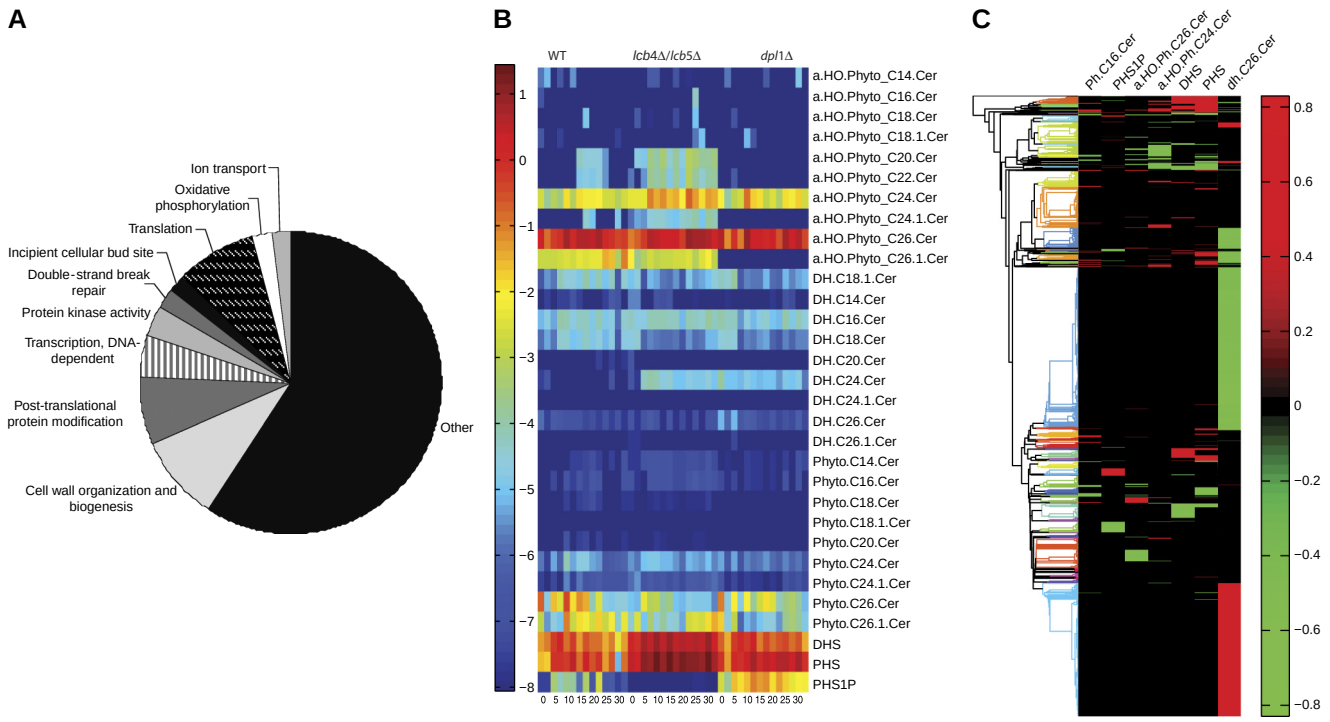

Figure 3 Effects of mutations in specific sphingolipid metabolic genes on gene expression and total sphingolipid profiles. (**A**) Overrepresented Gene Ontology annotations for genes aberrantly regulated during heat stress in the *lcb4*Δ*/lcb5*Δ mutant strain. (**B**) Heat map depicting changes in sphingolipid profiles over a time course of heat stress. Data are shown as a pseudo-colored heat map reflecting the logarithms of lipid mass measurements normalized to total phospholipid content of the sample. Log values of normalized measurements are color coded as indicated in the scale to the left of the heat map. (**C**) A double-sided clustering map depicting relationships between specific lipid–gene pairs over heat stress. A statistically significant (P-value $\leqslant 0.05$ and q-value < 0.1) positive correlation coefficient between a gene and a lipid is shown as a red bar; a significant negative one is shown as a green bar; the value of correlation coefficient is pseudo-color coded. In the map, rows represent genes, and columns represent lipids. The clustering tree on the left side of the map indicate gene clusters; a block across rows in the map represents a group of genes sharing similar information with respect to lipids; lipids with similar information with respect to gene expression (columns with similar color pattern) are grouped close to each other.

pattern demonstrated by a lipid with respect to all genes. We sought to identify the shared information patterns among genes and lipids by applying a double-sided hierarchical clustering analysis, which grouped genes (and lipids) sharing similar information patterns into clusters. Selected results focusing on key sphingolipids are shown in Figure 3C as a heat map (also see Supplementary Table 2). The results indicate that lipid species demonstrated distinct correlation patterns with respect to modules of genes, suggesting potential regulatory roles of the lipids on the modules. The double-sided clustering revealed clusters of genes sharing similar information with respect to lipid (blocks across rows) and clusters of lipids sharing similar information with respect to gene expression data (columns with similar correlation coefficient patterns). For example, DHS and PHS, two closely related metabolites in the metabolic network, were grouped together because of their shared information with respect to clusters of genes at the top region Figure 3C, but they also showed distinct information with other gene clusters.

Importantly, this procedure identified subsets of genes that were significantly correlated to PHS1P, with 44 positively correlated probe sets (mapped to 23 named genes) from the microarrays and 61 negatively correlated ones (mapped to named 54 genes), which was significantly less than the genes identified in the microarray analysis of the *lcb4*Δ*/lcb5*Δ mutant (441 genes). Among these 77 PHS1P-sensitive genes, 40 genes were also deemed differentially expressed according to differential expression analysis, whereas the other 33 genes were not. The results indicate a tentative statistical advantage of correlation analysis over the differential expression analysis in that the former can use all samples (42 samples for a lipid-versus-gene pair) whereas the latter can only use samples from specific conditions (12 *lcb4*Δ*/lcb5*Δ microarrays versus 12 *wt* microarrays after heat stress).

We applied a method referred to as GO Steiner Tree (GOSteiner, 2009) to analyze the functional coherence of the gene sets and to visualize their functional relationships. The method represents the genes and their associated Gene Ontology terms as a graph, finds a subgraph (a Steiner tree) connecting all genes and their annotations with a shortest total functional semantic distance, and finally evaluates the statistical properties of the tree as metrics of functional coherence. Supplementary Figure 1 shows the GO Steiner tree for genes that demonstrated failure to induce on heat stress in the *lcb4*Δ*/lcb5*Δ strain when compared to *wt*. Although the analysis showed that the *lcb4*Δ*/lcb5*Δ-sensitive genes had diverse functions, the visualization of the GO Steiner tree revealed clusters of genes with coherently related functions, including genes involved in mitochondrial metabolism, for example, COX4, INH1, and ATP17, and vesicular transport,

for example, APS2, SVP26, TPM2, and TVP23. The results indicate that the deletion of LCB4 and LCB5 led to aberrant regulation of several distinct groups of genes, among which genes showing high correlation to PHS1P represented only a subset of those genes (highlighted in Supplementary Figure 1). Importantly, this subset included genes in specific sub-categories (e.g. the mitochondrial metabolism), but not others (e.g. vesicular transport) of the LCB4/LCB5-regulated genes.

The above results led to the hypothesis that, although both sets of genes were *lcb4*Δ*/lcb5*Δ sensitive, only the genes showing strong correlation to PHS1P levels were truly PHS1P sensitive, whereas dysregulation of other genes in this mutant strain could result from confounding changes in other sphingolipids which had become apparent from the lipidomics analysis (Figure 3B). We tested this hypothesis by treating wild type cells (*wt)* with exogenous PHS1P in at non-heat stress temperature and monitored the expression of sample genes from the putative PHS1P-dependent set involved in the mitochondrial metabolism (COX4, INH1, and ATP17) and a putative non-PHS1P-dependent set involved in vesicular transport (APS2, SAR1, SVP26, TPM2, and TVP23). Indeed, expression of the genes involved in vesicular transport failed to be induced by PHS1P (Figure 4A), whereas the mitochondrial metabolism genes demonstrated ~2.4–3-fold increases on treatment (Figure 4B). To further test whether the effects were specific for PHS1P, *wt* and *lcb4*Δ*/lcb5*Δ mutant cells were treated with PHS, the metabolic precursor of PHS1P, which undergoes conversion to PHS1P in the wild type but not in the *lcb4*Δ*/lcb5*Δ mutant. Indeed, in wild-type cells, PHS produced a significant upregulation of the genes in the respiration set, which was totally lacking in the *lcb4*Δ*/lcb5*Δ mutant (Figure 4C). Thus, the integrated analysis enabled the identification of the PHS1P-dependent subset of genes within the larger set of genes that showed failed regulation in the *lcb4*Δ*/lcb5*Δ mutant.

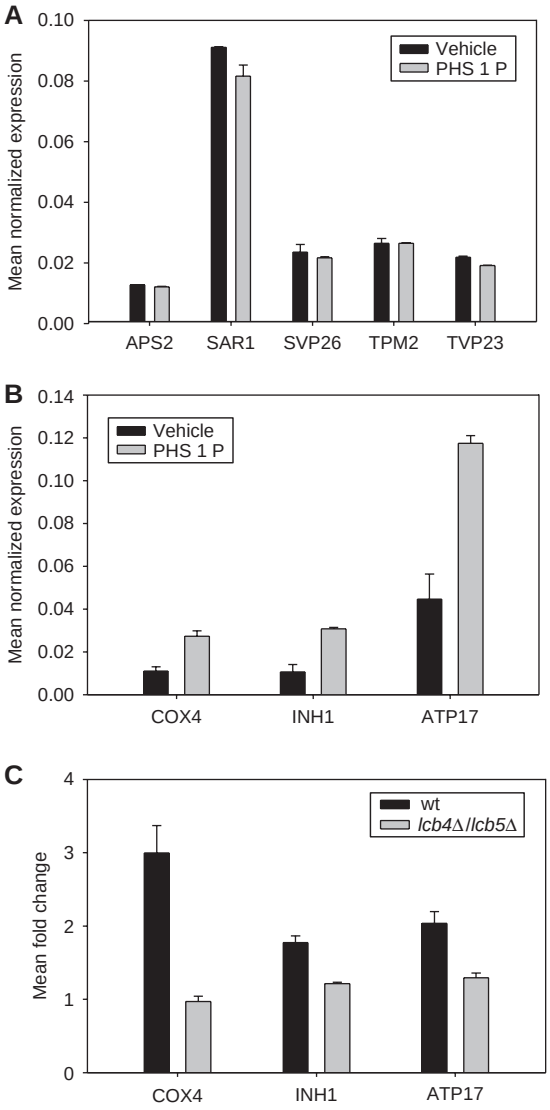

Figure 4 PHS1P-mediated gene expression regulation. (**A**) Treatment with PHS1P in the absence of heat stress did not induce all genes aberrantly regulated in the *lcb4*Δ*/lcb5*Δ strain. (**B**) Treatment with PHS1P in the absence of heat stress induces gene expression of the putative PHS1P-dependent genes identified by the integromics analysis. (**C**) The metabolic precursor of PHS1P, PHS, upregulated PHS1P-dependent genes in the wild-type strain, but not in the *lcb4*Δ*/lcb5*Δ mutant, which cannot phosphorylate PHS. Experiments were performed two to three times in triplicate and represented as mean ± s.e.m.

Identifying candidate TFs

Defining specific lipid-mediated responses, beyond the overall gene-mediated responses manifested by mutations, is important in that it allows studying the distinct roles and mechanisms of specific lipids—the putative functional mediators—that contribute to the overall responses. Highly selective PHS1P-mediated expression of specific genes led to the hypothesis that PHS1P regulates a specific pathway by activating some downstream TFs, which then mediate the regulation of these genes in response to PHS1P. To identify such putative TFs, a transcription factor binding site (TFBS) matrix was constructed, in which an element contains a binary variable indicating if a gene (g) can be bound by a TF (t). AN element, b_{tg}, of the matrix is set to 1 if the analysis of the chromatin immunoprecipitation experiments (Lee *et al*, 2002; MacIsaac *et al*, 2006) indicates that the TF t is capable of binding to the promoter sequence of gene g, or if there is documentations of the interaction between the TF and the gene according to the yeast TF database, YEASTRACT (Monteiro *et al*, 2008). The knowledge of TFBSs enabled us to evaluate whether any TFBS was significantly enriched in the promoters of a gene set. Using a hypergeometric-distribution-based model, we assessed enrichment of TFBSs in the promoters of the genes that demonstrated positive correlation to PHS1P and identified 22 TFBSs as significantly 'enriched' (see Supplementary information). This led to the hypothesis that the activation states of some of the candidate TFs were sensitive to changes in PHS1P, thus transmitting the signal from PHS1P to genes. To test the hypothesis, we further developed a two-stage Bayesian information integration model to reveal information flow from bioactive lipids → TFs → gene expression.

Inferring activation states of TFs

We developed a novel Bayesian model to infer the activation states of the TFs under each experimental condition. This

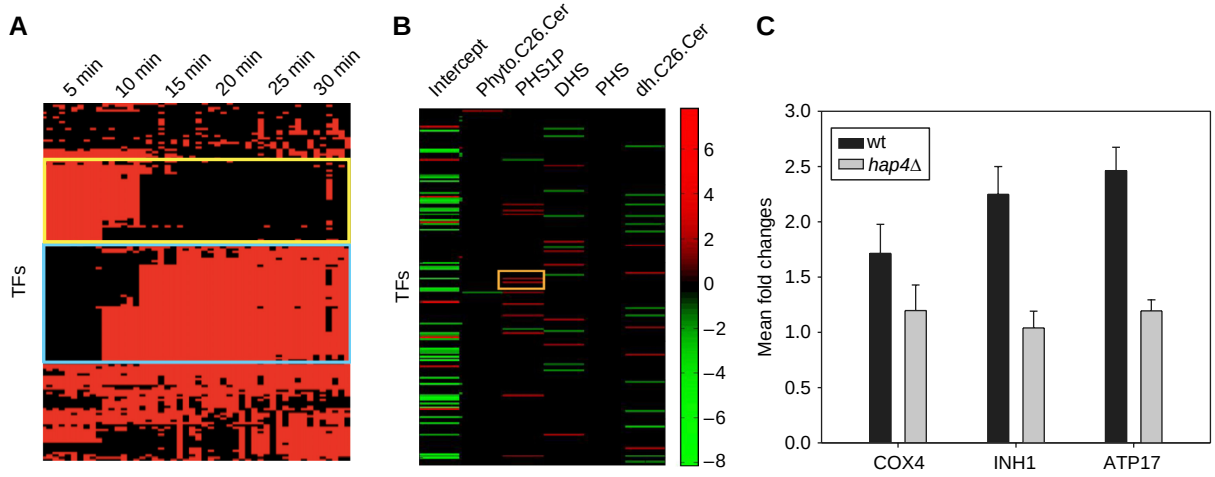

Figure 5 Modeling information flow from lipid, to TFs, and to gene expression. (**A**) Inferred TF activation states through integrating genomic and transcriptomic data. Red color indicates activated state and black denotes inactivate states. The TFs were grouped according to their state across experiment; the yellow block indicates a group of TFs 'turned off' after heat stress; the purple box outlines the TFs 'turned on' after heat stress. (**B**) Logistic regression modeling of the relationship between sphingolipids and TF states. Statistically significant regression parameters are shown as a TF-versus-lipid heat map. The orange box indicates the significant parameters associated with PHS1P with respect to Hap2p and Hap4p. (**C**) The ability of PHS1P treatment to induce PHS1P-dependent genes in the absence of HAP4 was determined. The experiment was performed three times in triplicate and represented as mean ± s.e.m.

model extends our previous published method (Lu *et al*, 2004) such that the genomic data of TFBSs can be integrated with transcriptomic data to infer the activation state of TFs. In this model, the activation state of a TF (t) under a specific condition (a) is represented as a binary variable, s_{at}, such that s_{at}=1 indicates the TF is active and s_{at}=0 otherwise. Representing activation states of TFs as binary variables has two advantages: (1) it provides a more intuitive representation of activation (active/inactive) state of a TF, in comparison to possible negative activation state allowed by some other models (Lee and Batzoglou, 2003; Liao *et al*, 2003; Gao *et al*, 2004; Ochs *et al*, 2004; Battle *et al*, 2005; Sun *et al*, 2006); (2) it renders the mathematical convenience to model the relationship between lipids and activation states of TFs using logistical regression, in which the sigmoid function mimics dose–response curves commonly observed in signal transduction pathways.

The model specifies that the expression value of a gene from a specific experiment (e_{ga}) is influenced by three factors: (1) TFs that bind to its promoter, indicated by $b_{gt}\forall t\in\{1,.....,T\}$; (2) the states of each of the TFs under this specific condition, represented by $s_{at}\forall t\in\{1,.....,T\}$; (3) and the strength and direction (induction or repression) of an activated TF on its expression, represented by $w_{gt}\forall t\in\{1,.....T\}$. We define the probabilistic relationship between the above parent variables and the gene expression value as follows:

$$e_{ga}=\sum_{t=1}^{T} b_{gt}s_{at}w_{ga}+\varepsilon \text{ or}$$

$$e_{ga}|b_{gt},\ s_{at},\ w_{ga},\sim N\left(\sum_{t=1}^{T} b_{gt}s_{at}w_{ga},\ \tau^{-1}\right)$$

where ε and τ represent the noise of the system and N stands for the Gaussian distribution. It is of interest to note that the product of two binary variables, $b_{gt}s_{at}$, encodes a logic AND relationship between the two variables in the equation, such that the equation can be interpreted as follows: TF t influences the expression value of gene g under the condition a if and only if it has a binding site in the promoter of the gene (b_{gt}=1) AND it is activated under the condition (s_{at}=1). The equation also reflects coordinated influences of multiple activated TFs on a gene's expression in a linear form (usually in logarithmic scales), an assumption widely used in modeling of expression systems (Liao *et al*, 2003; Gao *et al*, 2004; Lu *et al*, 2004; Battle *et al*, 2005). Given TFBS matrix and expression data, the model probabilistically infers the state of each TF under a specific condition using a variational Bayes technique (Lu *et al*, 2004) (see Supplementary information for detailed description). Figure 5A shows that many TFs switch states during heat stress, among which many are well-documented stress-responding TFs.

Modeling the role of lipids in TF activation

On the basis of the estimated binary TF states from the previous section, a Bayesian logistic regression model was applied to investigate the relationship between the levels of sphingolipids and the putative TF activation states. In this model, the states of a TF under each experimental condition were modeled as a sigmoid function of the concentrations of sphingolipids, where the probability that a TF (t) is active, conditioning on observed lipid profiles ($l_{1,\ldots,L}$) under an experiment condition (a), is defined as follows: $\text{logit}(p(s_{at}=1|l_1,l_2,...,l_L))=\beta_0+\beta_1 l_1+\beta_2 l_2+...+\beta_L l_L$. In this model, a parameter β_{lt} reflects the strength and direction of influence that the sphingolipid l has on the activation state of the TF t. Using a Gibb's sampling-based Bayesian logistic regression, we identified all statistically significant parameters, and we interpret a significant parameter β_{lt} with respect to a TF t as an indication that the TF

is regulated by lipid *l*, and the selected results are shown in Figure 5B. A total of 13 TFs showed with significant coefficient with respect to PHS1P, with 11 positively influenced and 2 negatively influenced. Among the positive TFs, Hap2p, Hap4p, NRG1 and MGA1 were also among the 22 TFs deemed significantly enriched in the PHS1P postively correlated gene set. Hap2p and Hap4p are members of the HAP complex (Chodosh *et al*, 1988; Buschlen *et al*, 2003), whose binding sites were dominant in the PHS1P postively correlated gene set, particularly in the subset involved in the cellular respiration. Thus, modeling of the relationship between sphingolipids and TF states, in combination with the static information inferred from promoter analysis, provided evidence at a mechanistic level that gave rises to the hypothesis that PHS1P regulates expression of respiratory genes through modulating the activity of the HAP complex.

To test this hypothesis, we evaluated the ability of PHS1P to regulate PHS1P-dependent genes in a mutant strain deleted for the gene encoding Hap4p (*hap4*Δ cells), an essential component of the HAP complex. *Wt* cells or *hap4*Δ cells were treated with exogenous PHS1P added into the culture media. RNA was isolated from each culture, and gene expression was determined by real time RT–PCR. Indeed, as compared with the parental background strain, the *hap4*Δ strain demonstrated complete loss of PHS1P-mediated induction of the target genes (Figure 5C). The result indicated that regulation of these genes by PHS1P required a functional HAP complex.

Discussion

Although there are publications (Fischer, 2005; Hirai *et al*, 2005; Ippolito *et al*, 2005) that simultaneously analyze metabolomic and transcriptomic data, these studies mainly concentrate on the relationships between the expression levels of the enzymes and their metabolites. To our best knowledge, this study represents a novel approach to integrating multiple '-omics' data to infer signal transduction pathways involving bioactive lipids at a mechanistic level. By integrating information from heterogeneous types of data in the principled probabilistic framework, the approaches developed in this study overcame the difficulties associated with conventional gene deletion/overexpression experiments commonly used in studies to delineate signaling pathways. The problem that single gene mutation leads to system-wide perturbation is likely to be a general case in many biological systems, particularly in metabolic networks; the interconnectedness of sphingolipid metabolism demonstrated in this study is likely an example rather than an exception of such systems. From a systems biology point of view, genetic manipulation of genes is a powerful tool to perturb systems, which provides opportunities to study the systems at mechanistic level but it is not necessarily sufficient to derive a causal relationship between metabolites and their effectors. By collecting and assimilating information at systems level, our study transcends the 'gene-centric' framework, leading to the identification of the signaling role of PHS1P in cellular stress responses in yeast. Progressive analyses generated specific hypotheses at different mechanistic levels: (1) PHS1P specifically regulates the expression of a set of genes involved in cellular respiration; and (2) this regulation requires the HAP TF complex; both findings received support from experimental validation. The study demonstrates the utility of integromic approaches in studying cellular signaling systems. This should be of great value in the study of bioactive lipids and other metabolic pathways, where the need arises to dissect functions of specific metabolites.

Materials and methods

Yeast culture and treatment

Strains of *S.cerevisiae* used are listed in Supplementary information. Yeasts were routinely cultured in Yeast Proteose Dextrose media at 30°C. Working cultures were seeded from overnight 5 ml cultures of a single colony and allowed to grow with 200–250 r.p.m. shaking to mid-logarithmic phase (OD=0.4–0.8). For sample treatment, cells were shifted to 39°C over a time course of 5, 10, 15, 20, 25, and 30 min. Cells were collected by centrifugation at 3500 *g* for 3 min, and pellets were snap-frozen in an ethanol/dry ice bath. Frozen pellets were stored at −80°C. For treatment with exogenous compounds, cells were maintained at 30°C and compounds were added as solutions in dimethylsulphoxide (DMSO), or DMSO alone was added as a negative control. Treatments with exogenous compounds were for 15 min.

RNA preparation and microarray hybridization

RNA was prepared from snap-frozen pellets using the RNeasy kit from Qiagen according to manufacturer's directions. Preparation of target for microarray hybridization to the Affymetrix YG-S98 chip was performed according to manufacturer's instructions. Microarray hybridization was performed at the Microarray Core Facility at the Medical University of South Carolina. Microarray analyses were performed using R packages of Bioconductor suite (Bioconductor, Gentleman *et al*, 2004). Microarrays were normalized using the robust microarray averaging package, and differential expression was assessed using the linear model for microarray data package from Bioconductor; false discovery rate was assess with Q-value package (Storey and Tibshirani, 2003). The threshold for differential expression and significant correlation coefficient was set at P-value <0.05 and q-value $\leqslant 0.1$. The microarray data set is publicly available at the Gene Expression Omnibus database (Barrett *et al*, 2009), with an accession number of GSE18121.

Lipidomics analysis

Lipids were extracted from snap-frozen yeast pellets as described (Bielawski *et al*, 2006) and subjected to high-throughput LC/MS analysis as described earlier (Bielawski *et al*, 2006). Quantification was based on comparison of peak intensity to internal standards as described earlier.

Real Time RT–PCR

Cells were grown in YPD media to mid-logarithmic phase and treated with compounds dissolved in DMSO as indicated or vehicle alone for 15 min in a 30°C water bath with 200–250 r.p.m. shaking. Cells were collected by centrifugation at 3500 *g* for 3 min, decanted, and immediately frozen at −80°C. RNA was extracted from frozen pellets using the RNeasy kit from Qiagen. cDNA was prepared from RNA using Superscript II or Superscript III (Invitrogen) according to manufacturer's directions. cDNA was diluted 10–20-fold before real-time PCR using the SybrGreen solution and protocols (Bio-Rad) and primers as indicated in the Supplementary information. Gene intensity signals were normalized to levels of RNA for either ribosomal 18S subunit or actin. Reactions were conducted in a Bio-Rad iCycler.

Two-staged Bayesian information flow model

Detailed mathematical and computational descriptions of the Bayesian latent variable model and Bayesian logistic regression are available as online Supplementary information at Molecular Systems Biology.

Supplementary information

Supplementary information is available at the *Molecular Systems Biology* website (www.nature.com/msb).

Acknowledgements

We acknowledge the following grant supports: P20 RR017677-07 (to LAC and XL). 1R01LM10144 to XL, 5R01GM63265 to YAH, the Department of Veterans' Affairs Merit award to LAC, 5R01LM009153, 3T15LM07438 to XL and AJR, and 5T32GM074934 to MS. We thank Alan Wilder and Jason Gandy for technical assistance. We also acknowledge the MUSC Lipidomics Core Facility of the COBRE in Lipidomics and Pathobiology for sample analysis, the MUSC Proteogenomics Core Facility for microarray assays, the MUSC department of Art Services for assistance with figure preparation, and Dr Hiroko Hama for critical reading of the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Alvarez SE, Milstien S, Spiegel S (2007) Autocrine and paracrine roles of sphingosine-1-phosphate. *Trends Endocrinol Metab* **18:** 300–307

Alvarez-Vasquez F, Sims KJ, Cowart LA, Okamoto Y, Voit EO, Hannun YA (2005) Simulation and validation of modelled sphingolipid metabolism in Saccharomyces cerevisiae. *Nature* **433:** 425–430

Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muertter RN, Edgar R (2009) NCBI GEO: archive for high-throughput functional genomic data. *Nucleic Acids Res* **37:** D885–D890

Battle A, Segal E, Koller D (2005) Probabilistic discovery of overlapping cellular processes and their regulation. *J Comput Biol* **12:** 909–927

Beissbarth T, Speed TP (2004) GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* **20:** 1464–1465

Bielawski J, Szulc ZM, Hannun YA, Bielawska A (2006) Simultaneous quantitative analysis of bioactive sphingolipids by high-performance liquid chromatography-tandem mass spectrometry. *Methods* **39:** 82–91

Bonander N, Ferndahl C, Mostad P, Wilks MD, Chang C, Showe L, Gustafsson L, Larsson C, Bill RM (2008) Transcriptome analysis of a respiratory Saccharomyces cerevisiae strain suggests the expression of its phenotype is glucose insensitive and predominantly controlled by Hap4, Cat8 and Mig1. *BMC Genomics* **9:** 365

Buschlen S, Amillet JM, Guiard B, Fournier A, Marcireau C, Bolotin-Fukuhara M (2003) The *S. cerevisiae* HAP complex, a key regulator of mitochondrial function, coordinates nuclear and mitochondrial gene expression. *Comp Funct Genomics* **4:** 37–46

Chodosh LA, Olesen J, Hahn S, Baldwin AS, Guarente L, Sharp PA (1988) A yeast and a human CCAAT-binding protein have heterologous subunits that are functionally interchangeable. *Cell* **53:** 25–35

Chung N, Jenkins G, Hannun YA, Heitman J, Obeid LM (2000) Sphingolipids signal heat stress-induced ubiquitin-dependent proteolysis. *J Biol Chem* **275:** 17229–17232

Cowart LA, Hannun YA (2005) Using genomic and lipidomic strategies to investigate sphingolipid function in the yeast heat-stress response. *Biochem Soc Trans* **33:** 1166–1169

Cowart LA, Okamoto Y, Pinto FR, Gandy JL, Almeida JS, Hannun YA (2003) Roles for sphingolipid biosynthesis in mediation of specific programs of the heat stress response determined through gene expression profiling. *J Biol Chem* **278:** 30328–30338

DeGroot MH, Schervish MJ (2002) *Probability and Statitstics*, 3rd edn. Addison-Wesley: Boston

Dickson RC, Nagiec EE, Skrzypek M, Tillman P, Wells GB, Lester RL (1997) Sphingolipids are potential heat stress signals in Saccharomyces. *J Biol Chem* **272:** 30196–30200

Fischer HP (2005) Towards quantitative biology: integration of biological information to elucidate disease pathways and to guide drug discovery. *Biotechnol Annu Rev* **11:** 1–68

Gao F, Foat BC, Bussemaker HJ (2004) Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. *BMC Bioinformatics* **5:** 31

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* **11:** 4241–4257

Gasch AP, Werner-Washburne M (2002) The genomics of yeast responses to environmental stress and starvation. *Funct Integr Genomics* **2:** 181–192

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G *et al* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5:** R80

GOSteiner (2009) http://projects.dbbe.musc.edu/gosteiner (accessed on May 4 2009)

Hannun Y, Obeid L (2008a) Principles of bioactive lipid signalling: lessons from sphingolipids. *Nat Rev Mol Cell Biol* **9:** 139–150

Hannun YA, Obeid LM (2008b) Principles of bioactive lipid signalling: lessons from sphingolipids. *Nature Rev* **9:** 139–150

Hirai MY, Klein M, Fujikawa Y, Yano M, Goodenowe DB, Yamazaki Y, Kanaya S, Nakamura Y, Kitayama M, Suzuki H, Sakurai N, Shibata D, Tokuhisa J, Reichelt M, Gershenzon J, Papenbrock J, Saito K (2005) Elucidation of gene-to-gene and metabolite-to-gene networks in arabidopsis by integration of metabolomics and transcriptomics. *J Biol Chem* **280:** 25590–25595

Ippolito JE, Xu J, Jain S, Moulder K, Mennerick S, Crowley JR, Townsend RR, Gordon JI (2005) An integrated functional genomics and metabolomics approach for defining poor prognosis in human neuroendocrine cancers. *Proc Natl Acad Sci USA* **102:** 9901–9906

Jenkins GM, Hannun YA (2001) Role for *de novo* sphingoid base biosynthesis in the heat-induced transient cell cycle arrest of Saccharomyces cerevisiae. *J Biol Chem* **276:** 8574–8581

Jenkins GM, Richards A, Wahl T, Mao C, Obeid L, Hannun Y (1997) Involvement of yeast sphingolipids in the heat stress response of Saccharomyces cerevisiae. *J Biol Chem* **272:** 32566–32572

Kim S, Fyrst H, Saba J (2000) Accumulation of phosphorylated sphingoid long chain bases results in cell growth inhibition in Saccharomyces cerevisiae. *Genetics* **156:** 1519–1529

Lee SI, Batzoglou S (2003) Application of independent component analysis to microarrays. *Genome Biol* **4:** R76

Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne JB, Volkert TL, Fraenkel E, Gifford DK *et al* (2002) Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science* **298:** 799–804

Liao JC, Boscolo R, Yang YL, Tran LM, Sabatti C, Roychowdhury VP (2003) Network component analysis: reconstruction of regulatory signals in biological systems. *Proc Natl Acad Sci USA* **100:** 15522–15527

Lu X, Hauskrecht M, Day RS (2004) Modeling cellular processes with variational Bayesian cooperative vector quantizer. *Pac Symp Biocomput* 533–544

MacIsaac KD, Wang T, Gordon DB, Gifford DK, Stormo GD, Fraenkel E (2006) An improved map of conserved regulatory sites for Saccharomyces cerevisiae. *BMC Bioinformatics* **7:** 113

Meier KD, Deloche O, Kajiwara K, Funato K, Riezman H (2006) Sphingoid base is required for translation initiation during heat stress in Saccharomyces cerevisiae. *Mol Biol Cell* **17:** 1164–1175

Monteiro PT, Mendes ND, Teixeira MC, d'Orey S, Tenreiro S, Mira NP, Pais H, Francisco AP, Carvalho AM, Lourenco AB, Sa-Correia I, Oliveira AL, Freitas AT (2008) YEASTRACT-DISCOVERER: new tools to improve the analysis of transcriptional regulatory associations in Saccharomyces cerevisiae. *Nucleic Acids Res* **36:** D132–D136

Nagiec MM, Skrzypek M, Nagiec EE, Lester RL, Dickson RC (1998) The LCB4 (YOR171c) and LCB5 (YLR260w) genes of Saccharomyces encode sphingoid long chain base kinases. *J Biol Chem* **273:** 19437–19442

Ochs MF, Moloshok TD, Bidaut G, Toby G (2004) Bayesian decomposition: analyzing microarray data within a biological context. *Ann N Y Acad Sci* **1020:** 212–226

Saba JD, Nara F, Bielawska A, Garrett S, Hannun YA (1997) The BST1 gene of Saccharomyces cerevisiae is the sphingosine-1-phosphate lyase. *J Biol Chem* **272:** 26087–26090

Skrzypek MS, Nagiec MM, Lester RL, Dickson RC (1999) Analysis of phosphorylated sphingolipid long-chain bases reveals potential roles in heat stress and growth control in Saccharomyces. *J Bacteriol* **181:** 1134–1140

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* **100:** 9440–9445

Sun N, Carroll RJ, Zhao H (2006) Bayesian error analysis model for reconstructing transcriptional regulatory networks. *Proc Natl Acad Sci USA* **103:** 7988–7993

Zheng W, Kollmeyer J, Symolon H, Momin A, Munter E, Wang E, Kelly S, Allegood JC, Liu Y, Peng Q, Ramaraju H, Sullards MC, Cabot M, Merrill Jr AH (2006) Ceramides and other bioactive sphingolipid backbones in health and disease: lipidomic analysis, metabolism and roles in membrane structure, dynamics, signaling and autophagy. *Biochim Biophys Acta* **1758:** 1864–1884