



Evaluation of predicted network modules in yeast metabolism using NMR-based metabolite profiling

Jacob G. Bundy, Balázs Papp, Rebecca Harmston, et al.

Genome Res. 2007 17: 510-519 originally published online March 5, 2007

Access the most recent version at doi:[10.1101/gr.5662207](https://doi.org/10.1101/gr.5662207)

References This article cites 35 articles, 14 of which can be accessed free at:
<http://genome.cshlp.org/content/17/4/510.full.html#ref-list-1>

Open Access Freely available online through the *Genome Research* Open Access option.

License Freely available online through the Genome Research Open Access option.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Copyright © 2007, Cold Spring Harbor Laboratory Press

Methods

Evaluation of predicted network modules in yeast metabolism using NMR-based metabolite profiling

Jacob G. Bundy,^{1,3} Balázs Papp,² Rebecca Harmston,¹ Roy A. Browne,¹
Edward M. Clayson,¹ Nicola Burton,² Richard J. Reece,² Stephen G. Oliver,²
and Kevin M. Brindle^{1,4}

¹Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, United Kingdom; ²Faculty of Life Sciences, The University of Manchester, Manchester M13 9PT, United Kingdom

Genome-scale metabolic models promise important insights into cell function. However, the definition of pathways and functional network modules within these models, and in the biochemical literature in general, is often based on intuitive reasoning. Although mathematical methods have been proposed to identify modules, which are defined as groups of reactions with correlated fluxes, there is a need for experimental verification. We show here that multivariate statistical analysis of the NMR-derived intra- and extracellular metabolite profiles of single-gene deletion mutants in specific metabolic pathways in the yeast *Saccharomyces cerevisiae* identified outliers whose profiles were markedly different from those of the other mutants in their respective pathways. Application of flux coupling analysis to a metabolic model of this yeast showed that the deleted gene in an outlying mutant encoded an enzyme that was not part of the same functional network module as the other enzymes in the pathway. We suggest that metabolomic methods such as this, which do not require any knowledge of how a gene deletion might perturb the metabolic network, provide an empirical method for validating and ultimately refining the predicted network structure.

[Supplemental material is available online at www.genome.org.]

Genome-scale metabolic modeling is one of the most promising approaches for predicting cellular function based on knowledge of the properties of individual cellular components (Borodina and Nielsen 2005). Constraint-based metabolic models (Price et al. 2004), or stoichiometric flux balance models, have been used successfully in the analysis of phenotypes resulting from gene deletion (Famili et al. 2003; Duarte et al. 2004; Kuepfer et al. 2005) and for guiding gene expression analysis (Patil and Nielsen 2005). For example, in the yeast *Saccharomyces cerevisiae*, flux balance analysis of the reconstructed metabolic network was shown to predict the qualitative effect of single-gene deletions with a combined accuracy of 81%–89%, with 96%–98% of viable and 68%–80% of lethal deletions correctly identified (Kuepfer et al. 2005).

A key step in generating these models is the reconstruction of the metabolic network. This usually starts with the assignment of reactions to annotated genes and is followed by validation, based largely on the published biochemical literature, and by iterative improvement through the analysis of high-throughput data. We have shown previously that multivariate statistical analysis of cellular metabolite profiles from gene deletion mutants of *S. cerevisiae* is capable of distinguishing between genes in different pathways and associating genes in the same pathway. Recognition of the similarity between the intra- or extracellular metabolite profiles of a strain deleted for a gene of unknown

function and the corresponding profiles from a strain deleted for a gene of known function makes it possible to deduce the function of the unknown gene (Raamsdonk et al. 2001; Allen et al. 2003). However, these studies did not establish whether it is possible to use metabolomic methods to distinguish gene knockouts of different steps of a single pathway from gene knockouts in a separate pathway(s).

We show here, using a series of mutants of *S. cerevisiae*, that the metabolic profiles from mutants within a pathway generally cluster and are separate from the profiles obtained from mutants in another pathway. More importantly, we also demonstrate that there are outliers, where the metabolite profile of a mutant that lacks the gene for an enzyme or protein, which is nominally associated with a particular pathway, is not clustered with the metabolite profiles of other deletants lacking genes normally associated with the same pathway. The presence of outliers raises the question as to whether intuitively defined pathways correspond to natural modules of biochemical reaction networks, where modules can be broadly defined as groups of enzymes working together for the same cellular function and, more specifically, as groups of reactions that always operate together in steady states of the metabolic network (i.e., those with coupled fluxes) (Papin et al. 2004). An efficient approach to identifying mathematically defined modules is to perform flux coupling analysis (FCA) (Burgard et al. 2004). We demonstrate that FCA (Burgard et al. 2004; Kuepfer et al. 2005) may be used to show that these experimentally determined outliers are not part of the same mathematically defined metabolic modules as the other steps in the pathways. Thus, gene deletion mutants within a module have indistinguishable metabolite profiles, whereas deletion of a gene for another enzyme in the pathway, which is not flux coupled to the module, produces a different metabolite profile, which allows the two to be distinguished. Therefore, the

³Present address: Department of Biomolecular Medicine, Division of Surgery, Oncology, Reproductive Biology, and Anaesthetics, Faculty of Medicine, Sir Alexander Fleming Building, London SW7 2AZ, UK.

⁴Corresponding author.

E-mail kmb@mole.bio.cam.ac.uk; fax 44-1223-766002.

Article published online before print. Article and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.5662207>. Freely available online through the *Genome Research* Open Access option.

enzyme deleted in an outlier is part of the pathway, but not part of the module, since it can participate in other pathways, and hence its flux is uncoupled from the module flux. We propose, therefore, that this approach could be used as an empirical experimental tool to validate modules within metabolic networks that have been identified by analysis of network structure.

Results

Resolution of the intracellular metabolite profiles from mutants in pyrimidine biosynthesis and proline utilization is compromised by an outlier in the pyrimidine pathway

We chose the proline utilization and pyrimidine biosynthesis pathways for study because they represented two closely linked metabolic pathways, one of which is catabolic and the other anabolic. These two pathways share a common intermediate, glutamate, which is the starting point for pyrimidine biosynthesis and the end point for proline utilization (Fig. 1). Mutants carrying deletions in genes encoding enzymes or regulators of these two interlinked pathways should provide a stringent test of metabolomic methods for distinguishing different functional groupings of genes.

Metabolomic analyses were performed using proton NMR (see Methods) on cell extracts prepared from yeast cultures grown on a minimal medium until they reached early stationary phase (24–25 h). The yeast strains studied were deleted for single genes (see Fig. 1; Supplemental Table 1) encoding enzymes or regulators for the proline utilization and pyrimidine biosynthesis pathways. It should be noted that the *ura3Δ0* allele, present in the genetic background of the BY series of deletion mutants (Winzeler et al. 1999; Giaever et al. 2002), was removed by replacing it with the wild-type *URA3* gene (see Methods). The expectation, based on our previous work, was that mutants in proline metabolism would give a distinct metabolite profile that was different from the profile produced by pyrimidine pathway mutants, and that these different profiles would allow the two different sets of pathway mutants to be distinguished (Raamsdonk et al. 2001; Allen et al. 2003).

Gene deletions that affect growth rate are likely to have profound effects on cellular metabolite profiles, which could obscure those metabolic changes more specifically associated with deletion of a particular gene. Therefore, in all of the studies described here, the effect of gene deletion on growth rate was determined. In this initial set of experiments, comparing intracellular metabolite profiles in proline and pyrimidine pathway mutants, the different gene deletions did not have a significant effect on the rate of exponential phase growth, although there was a slight growth delay of ~2 h in the case of the *ura7Δ* mutant (see Supplemental Material). The *ura7Δ* mutant had a growth rate in exponential phase of $0.22 \pm 0.03 \text{ h}^{-1}$ ($n = 3$), as compared with $0.232 \pm 0.003 \text{ h}^{-1}$ for all the other strains ($n = 12$ strains, where the means of three replicates were used for each strain; i.e., 36 measurements for each time point in the mean growth curve). All the strains reached an equivalent cell density after 24–25 h.

The overall similarity of the NMR-derived metabolite profiles (Fig. 2) was visualized by principal components analysis (PCA). This is an unsupervised data analysis technique that generates new axes from multivariate data that are linear combinations of the original variables. Each axis explains the greatest

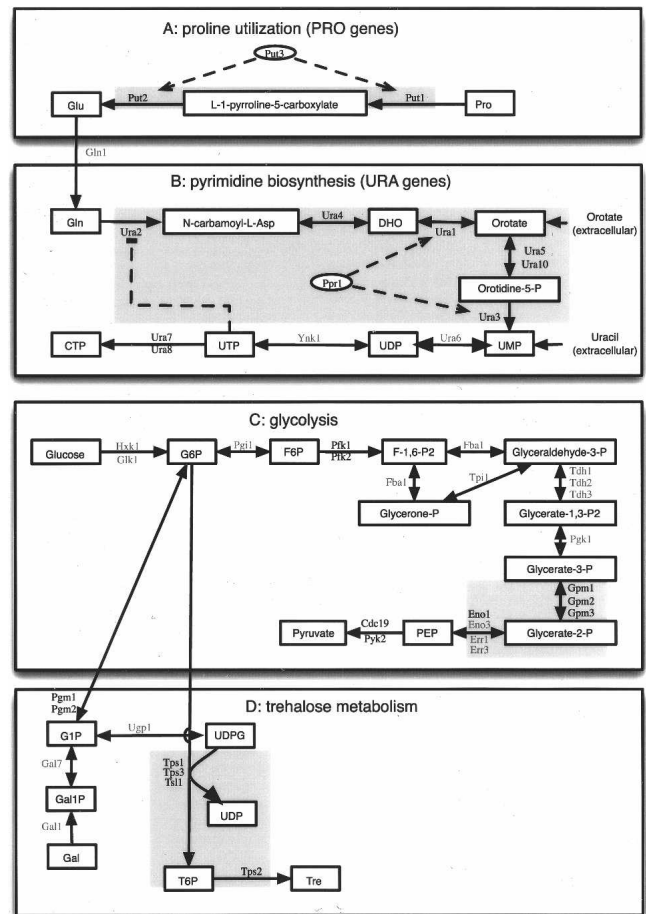


Figure 1. The metabolic pathways affected in the gene deletion mutants used in this study. (Bold) Deleted genes, (shaded) metabolic co-sets, identified using flux coupling analysis (see text). (A) Proline utilization (*PUT* genes). (B) Pyrimidine biosynthesis (*URA* genes). (DHO) Dihydroorotate. (C) Glycolysis. (G6P) Glucose 6-phosphate, (F6P) fructose 6-phosphate, (F-1,6-P₂) fructose-1,6-bisphosphate, (Glycerone-P) dihydroxyacetone phosphate, (Glyceraldehyde-3-P) glyceraldehyde 3-phosphate, (Glycerate-1,3-P₂) 1,3-bisphosphoglycerate, (Glycerate-3-P) 3-phosphoglycerate, (Glycerate-2-P) 2-phosphoglycerate, (PEP) phosphoenolpyruvate. (D) Trehalose synthesis. (G1P) glucose 1-phosphate, (Gal1P) galactose 1-phosphate, (Gal) galactose, (Tre) trehalose, (T6P) trehalose-6-phosphate.

amount of variance possible; axis 1 is fitted first (and therefore explains the single greatest amount of variance possible in one dimension), and each subsequent axis is orthogonal to all preceding axes. The aim of PCA is to reduce the dimensionality of the original data by fitting a small number of axes that explain the majority of the variance in the data, and it can be used to visualize inherent patterns within the data. The principal component (PC) 1 scores were correlated with the OD of the cultures at cell extraction ($r^2_{\text{adj}} = 0.47$, $P = 2.2 \times 10^{-10}$) and, therefore, were interpreted as representing small differences in cell density between individual flasks. Since each of the strains showed the same range of final ODs (the mean cell density after 24–25 h was the same for all strains), there was no strain separation along PC1. A plot of PC2 against PC3, which together represented the greatest amount of variance remaining in the data, showed no separation of the two different pathway mutants (Fig. 3A). However, there was an unexpected separation of the *ura7Δ* mutant

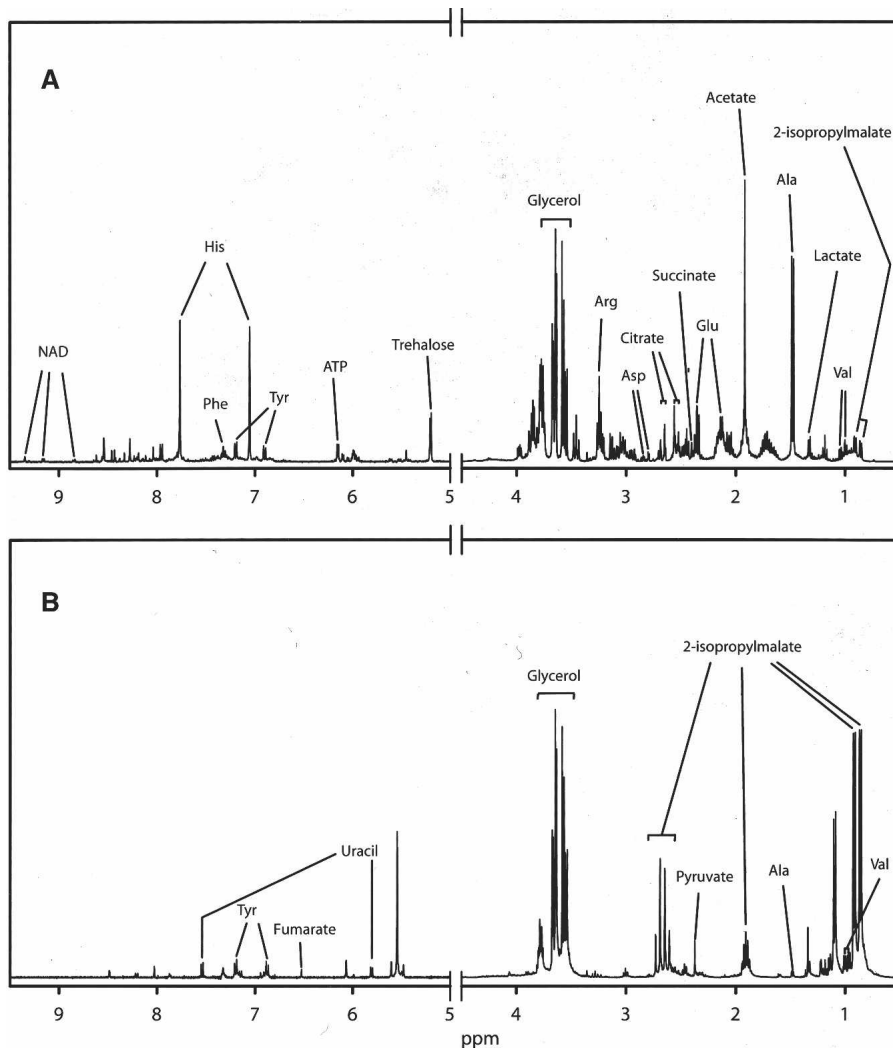


Figure 2. ¹H NMR spectra of a typical cell extract (A) and of cell growth medium (B).

(*URA7* encodes CTP synthase) in the pyrimidine biosynthetic pathway, which was clearly separated from all of the other mutants in both the proline and pyrimidine pathways. Separation of this mutant from the other mutants in the pyrimidine biosynthetic pathway could have been a consequence of its position in the pathway (Fig. 1), since it is the only enzyme, with the exception of the isoenzyme encoded by *URA8*, which is downstream of the uracil supplied in the culture medium. The failure of the *ura8Δ* mutant to cluster with *ura7Δ* may be explained by its much smaller influence on metabolic flux (~15% of the pathway flux is thought to be catalyzed by *Ura8p*) (Ozier-Kalogeropoulos et al. 1994). However, when the same strains were grown on orotate in place of uracil, *ura5Δ* and *ura10Δ* continued to cluster with the other strains, and *ura7Δ* remained an outlier (Fig. 3B). Had the source of exogenous pyrimidine been a factor in the separation of the *ura7Δ* mutant, then it might have been expected that the *ura5Δ*, *ura10Δ*, and *ura3Δ* mutants would cluster with *ura7Δ* and *ura8Δ* mutants when the cells were supplied with orotate in place of uracil (see Fig. 1).

Analysis of all the data in a single PCA model (cells grown on uracil and orotate) showed that the largest source of variance

was still the final culture OD (correlated with PC1), with the next largest source of variance being the choice of orotate or uracil, which gave a highly significant separation of samples along PC2 ($P = 4.6 \times 10^{-37}$, *t*-test). This separation along PC2 is not evident in Figure 3B, as we have presented averages for each of the strains grown on uracil and orotate. However, the *ura7Δ* mutant was still clearly separated from the other strains along PC3 (Fig. 3B). Moreover, the metabolites responsible for the separation of the *ura7Δ* mutant from the other strains were largely the same whether the cells were grown on orotate or uracil. The PCA loadings can be interpreted as showing which metabolites differed between the *ura7Δ* mutant and the remaining strains. The remarkable similarity between the loadings for PC2 in Figure 3A and PC3 in Figure 3B, shown in Figure 3C as the red and blue lines, respectively, demonstrates that the metabolic changes responsible for separating the *ura7Δ* mutant from the other strains were unaffected by the pyrimidine source (uracil or orotate) in the growth medium. The statistical significance of the metabolite differences between the *ura7Δ* mutant and the other strains was demonstrated using multiple *t*-tests (Fig. 3C). In summary, the metabolite profile of the *ura7Δ* mutant shows it to be completely different from the other mutants in the pyrimidine biosynthetic pathway, and that this profile persists regardless of the pyrimidine source (uracil or orotate) in the growth medium.

Extracellular metabolite profiles confirm the *ura7Δ* mutant as an outlier and demonstrate that the *ura8Δ* mutant is also an outlier

In order to confirm that the *ura7Δ* mutant was an outlier when compared with the other mutants in pyrimidine biosynthesis, spectra were also acquired from media taken after 72 h of cell culture. Acquisition of extracellular metabolite profiles, or “footprints,” has previously been shown to be a useful method for distinguishing the effects of gene deletion in both yeast and a bacterium (Allen et al. 2003; Kaderbhai et al. 2003), and for distinguishing the mechanism of action of different compounds (Allen et al. 2004). Footprinting not only has the advantages of simplicity and not requiring quenching of cell metabolism, but may also be more sensitive to the metabolic state of the cell. Extracellular metabolite concentrations could potentially vary more between experimental treatments than for intracellular metabolites, as these excreted products are not constrained by cellular homeostasis. Moreover, the footprint represents the accumulation of metabolic products across the entire growth curve (Kell et al. 2005).

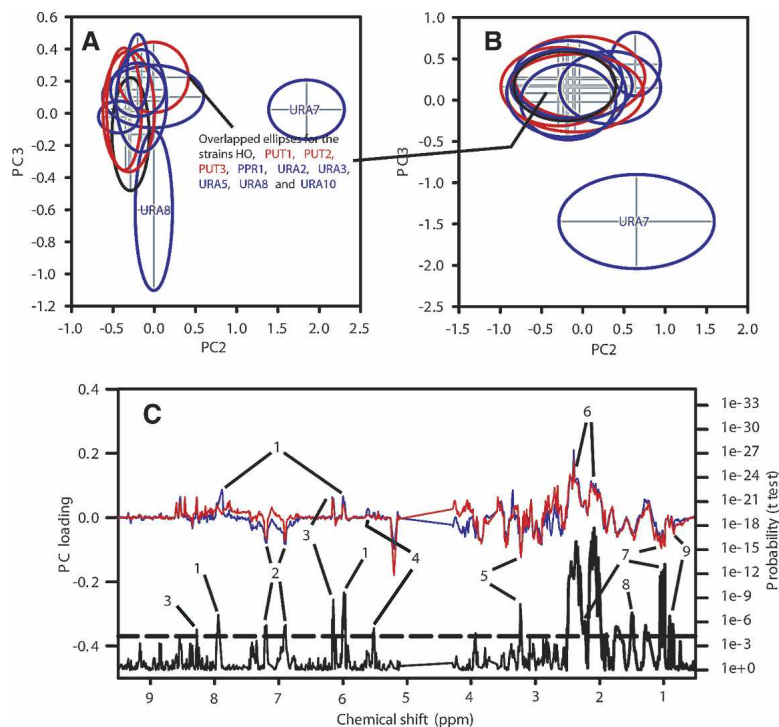


Figure 3. Principal Components Analysis (PCA) scores plots for cell extracts. The center of each ellipse indicates the mean, and the margin indicates one standard deviation. The cells were grown on minimal medium containing uracil. (A) PC2 vs. PC3 scores plot for three mutations in genes affecting proline utilization (*PUT1*, *PUT2*, and *PUT3*; red), seven mutations in genes affecting pyrimidine biosynthesis (*URA2*, *URA3*, *URA5*, *URA7*, *URA8*, *URA10*, *PPR1*; blue), and a control gene (*HO*, black) (four to six replicate samples for each strain). PC1 was correlated with sample OD and is not shown. (B) Scores plot for the same 11 strains grown on minimal medium containing uracil and on minimal medium containing orotate (eight to 13 replicate samples for each strain). PC1 was correlated with sample OD and is not shown. (C) Loadings plot showing the variables responsible for separation of the *ura7Δ* mutant. (Blue line) PC 2 loadings from the PCA model shown in panel A, (red line) PC 3 loadings from the PCA model shown in panel B, (black line) probability of significant difference between these variables for the *ura7Δ* and *hoΔ* samples (data from both the orotate and uracil cultures were used; *t*-test, where for the two groups $n = 24$ and 13 , respectively), (dashed line) threshold for statistical significance, adjusted in accordance with the Bonferroni criterion. Metabolites labeled directly on the plot: (1) uridine/cytidine nucleotides, (2) tyrosine, (3) adenine nucleotides, (4) possibly a phosphosugar, (5) choline metabolites, (6) glutamate, (7) valine, (8) alanine, (9) isopropylmalate.

The *ura7Δ* mutant remained a strong outlier in the analysis of these footprint data, and the *ura8Δ* mutant was also now an outlier (Fig. 4A). The loadings plot showed that one of the metabolites separating these mutants was glucose (Fig. 4B). However, the amount of glucose was low compared with the other metabolites present (see Supplemental Material), and the *ura7Δ* and *ura8Δ* mutants remained outliers even when the resonances from glucose were excluded from the analysis (Fig. 4C). This residual glucose cannot be ascribed to differences in the growth of *ura7Δ* and *ura8Δ* mutants from the other strains. The growth of the *ura8Δ* mutant was indistinguishable from that of the other strains (see Supplemental Data) and, although the *ura7Δ* mutants showed a slight growth delay, all the strains reached the same cell density and had been at this maximal cell density for 48 h when the media samples were taken for NMR analysis. The metabolic phenotype of the *ura7Δ* mutant could have been the result of some other mutation introduced during the production of the deletion strain (see Methods). Therefore, we confirmed a specific association

by transforming the *ura7Δ* mutant with a plasmid expressing *URA7* and demonstrating that this rescued the mutant metabolic phenotype (Fig. 4D). Expression of *URA7* in these plasmid-transformed cells was confirmed by demonstrating that it restored resistance to the growth inhibitory effects of Calcofluor white (Ram et al. 1994; data not shown). These data also provided the only instance in which a gene deletion could be trivially related to the observed metabolic changes: The *ura3Δ* strain showed an accumulation of orotate in the medium (singlet resonance at $\delta = 6.19$ ppm), which was not observed in any of the other strains.

The presence of outliers may mask more subtle variation in the remaining samples. Consequently, we reanalyzed the data with the *ura7Δ* and *ura8Δ* samples omitted. This showed, in accordance with our original expectations, that the mutants in proline utilization (*put* mutants) tended to form a separate cluster from the *ura* mutants, although there was not a complete separation of the metabolic profiles (Fig. 4E).

The proline pathway mutants separated from the pyrimidine pathway mutants when the cells were grown on proline, and the *put3Δ* mutant separated from the *put1Δ* and *put2Δ* mutants

The *PUT* genes are expected to be active only in the presence of proline; Put3p (encoded by *PUT3*) is a transcription factor that increases *PUT1* and *PUT2* expression following binding of proline (Brandriss and Magasanik 1979; Sellick and Reece 2003). When proline was used instead of ammonium sulfate as the nitrogen source, all three *put* mutants became outliers (Fig. 4F). *put1Δ* and *put2Δ* overlapped in principal components 1 and 2, whereas the *put3Δ* samples formed a separate cluster. The *ura7Δ* strain remained an outlier, although it was apparently similar to *put1Δ* and *put2Δ* (Fig. 4F). However, in this case there appeared to be some association between growth rate and PC1. The *put1Δ* and *put2Δ* mutants showed a marked growth delay, whereas the *put3Δ* and *ura7Δ* mutants showed similar growth profiles and a much smaller growth delay (see Supplemental Material). Nevertheless, despite the similarity in their growth profiles, the *put3Δ* and *ura7Δ* mutants are well separated along PC1, showing that variation in this component is not due to differences in growth alone. Furthermore, examination of lower components showed the *ura7Δ* mutant to be completely resolved, on a separate orthogonal axis (PC5), from the rest of the mutants ($P = 4.5 \times 10^{-4}$). A PLS-DA analysis of the same data also showed that *ura7Δ* was separated along axis 3 (data not shown). PLS-DA is a supervised technique

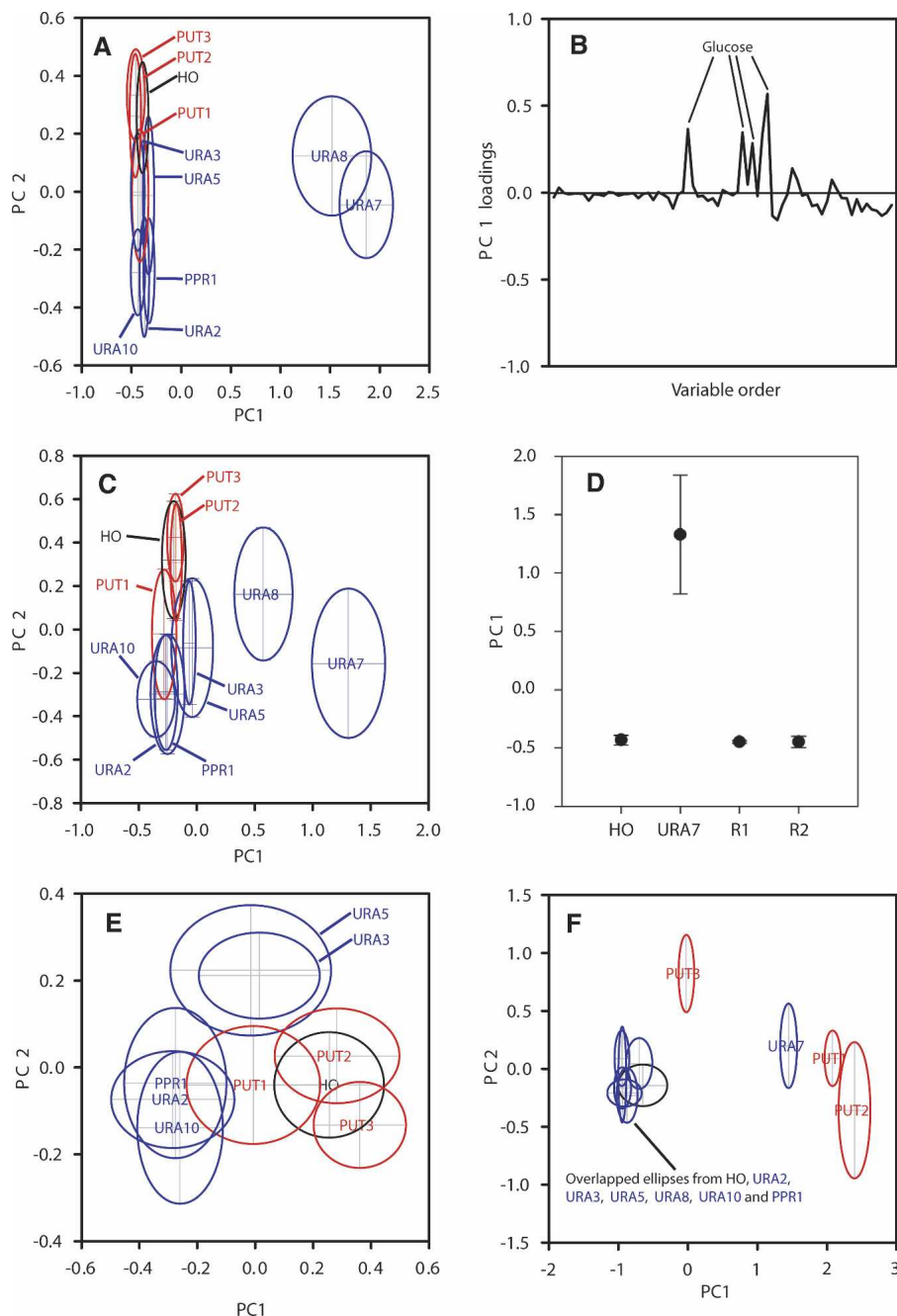


Figure 4. PCA scores plots for the extracellular metabolite profiles. (A) PC1 vs. PC2 scores plot for three mutations affecting genes in proline utilization (*PUT1*, *PUT2*, and *PUT3*; red), seven mutations affecting genes in pyrimidine biosynthesis (*URA2*, *URA3*, *URA5*, *URA7*, *URA8*, *URA10*, *PPR1*; blue), and a control gene (*HO*, black) (eight replicate samples for each strain). (B) Loadings plot for the PC1 axis shown in panel A. (C) PC1 vs. PC2 scores plot for the data shown in panel A, recalculated after removing variables containing NMR signals from glucose. (D) Scores plot (PC1) showing that reintroduction of functional *URA7* restores a wild-type (*hoΔ*) metabolic phenotype. (R1) *ura7Δ* strain with chromosomal reinsertion of *URA7* under control of its native promoter, (R2) *ura7Δ* strain with chromosomal reinsertion of *URA7* and its native terminator under control of the native promoter. (E) PC1 vs. PC2 scores plot for the data shown in panel A, recalculated after excluding the *ura7Δ* and *ura8Δ* samples. (F) PC1 vs. PC2 scores plot of the same strains grown on minimal medium containing proline rather than ammonium as a nitrogen source (five to six replicate samples for each strain).

that uses prior knowledge about class membership to find orientations of the PLS axes that increase the separation between groups.

and *PUT2* transcription, but not necessarily identical. This corresponds to our experimental findings, in which *put1Δ/put2Δ*, *put3Δ*, *ura7Δ*, and the remaining *ura* mutants all form separate

Identification of functional network modules using flux coupling analysis can explain why *ura7Δ* and *ura8Δ* are outliers in the pathway of pyrimidine biosynthesis

The “pyrimidine biosynthetic pathway” shown in Figure 1 is just one possible path through the metabolic network from glutamine to CTP, and not all reactions in that path might operate together in all physiological network states. To investigate whether the reactions of this intuitively defined pathway are part of the same mathematically defined metabolic module, we performed flux coupling analysis (FCA) (Burgard et al. 2004) (see Methods). FCA identifies coupled biochemical reactions in steady-state flux distributions of a network. Two reactions are fully coupled if non-zero flux through one reaction implies a non-zero and fixed flux through the other reaction and vice versa, and are defined as directionally coupled if non-zero flux through one reaction implies non-zero flux through the other, but not necessarily the reverse (Burgard et al. 2004). Fully coupled reactions are expected to have perfectly correlated flux levels, hence being members of the same co-set or module. Application of FCA showed that the enzymes of the pyrimidine biosynthesis pathway from *Ura2p* to *Ura3p* form a single fully coupled module in the form of a linear pathway (Fig. 1). Therefore, one would expect the metabolite profiles of strains deleted for their cognate genes to cluster as a single group. However, the “downstream” enzymes are not coupled to the rest of the pathway. This provides a possible explanation for why the *ura7Δ* and *ura8Δ* strains do not cluster with the rest of the pyrimidine pathway mutants. Uncoupling of enzymes in the lower part of the pathway from the *Ura2p*-*Ura3p* module is further supported by our observation that the *ura6Δ* strain possesses a severe growth defect (data not shown).

FCA also showed that *Put1p* and *Put2p* form a coupled module. Thus, the *put1Δ* and *put2Δ* deletants are expected to show similar metabolite profiles; *Put3p* is a transcription factor and hence is not explicitly considered in the FCA. We might expect that *put3Δ* would be similar, given its modulation of *PUT1*

clusters (Fig. 4F). We expected *put3Δ* to cluster with *put1Δ/put2Δ*, but only if the Put3p transcription factor regulated the expression of nothing but *PUT1* and *PUT2*. We observed that the *put3Δ* mutant is separated from the *put1Δ/put2Δ* mutants along both PCs 1 and 2. The separation along PC1 may be explained, at least in part, by differences in growth, but there was no correlation with growth along PC2. The fact that *put3Δ* does not cluster with *put1Δ/put2Δ* along either PCs 1 or 2 is consistent with data showing that Put3p binds to upstream regions of other genes besides *PUT1* and *PUT2* (Lee et al. 2002).

Testing the predictions of flux coupling analysis in trehalose biosynthesis and glycolysis

FCA was used to select another set of enzyme-gene deletions, in the pathways of glycolysis and trehalose biosynthesis (Fig. 1). The trehalose biosynthesis pathway was chosen because it enabled us to test the context dependency of coupling relationships. The reaction catalyzed by Tps1p, Tps3p, and Tsl1p and that catalyzed by Tps2p are fully coupled irrespective of the environmental conditions. However, the reaction catalyzed by Pgm1p and Pgm2p was predicted to be directionally coupled to the other two reactions when cells are grown on glucose (i.e., its activity is implied by the activity of the other two reactions, but not necessarily the reverse), but uncoupled when they are grown on galactose (Fig. 1). Thus, the *pgm1Δ* and *pgm2Δ* strains should be less similar to the other strains when grown on galactose than when grown on glucose. (Although, since Pgm2p is the major isoform and accounts for 80%–90% of the flux [Boles et al. 1994], a secondary prediction was that the effects of *pgm2Δ* deletion might be greater than for that of *PGM1*, as was the case for the *ura7Δ* and *ura8Δ* gene deletions.)

When the cells were grown on glucose, there was no clear distinction between the different mutants using PCA (Fig. 5A). However, using a supervised analysis method (PLS-DA), there was, as expected, increased strain separation. This was most marked for the *pgm2Δ* mutant, which now appeared to be resolved from the other mutants (data not shown). The *tps1Δ* mutant did not grow on glucose, presumably because of its role in regulating glycolysis (Teusink et al. 1998). When grown on galactose, the *tps1Δ* and *pgm2Δ* strains exhibited significant growth defects, with doubling times of 4.6 ± 0.2 h and 5.9 ± 0.5 h (SD, $n = 5$), respectively. The remaining mutants did not have significantly different growth rates (3.68 ± 0.07 h). This led to significant amounts of galactose remaining in the media samples from the *tps1Δ* and *pgm2Δ* strains, which were both separated from the other strains along PC 1; hence, we interpreted PC 1

as a “growth defect” axis (Fig. 5B). Examining lower components showed that both *pgm2Δ* and *pgm1Δ* were differentiated from the remaining mutants (data not shown). In the case of the *pgm1Δ* mutants, there was no difference in the growth profile when compared with the other strains. Although there was some overlap of points, the separation of these mutants along PC4 was significant ($P = 2 \times 10^{-9}$, *t*-test considering *pgm2Δ* and *pgm1Δ* data as the first group, and all other mutants as the second group), thus confirming the predictions of FCA.

This separation of the *pgm1Δ* and *pgm2Δ* strains from the others, while visible in the PCA plot as described, was demonstrated more clearly by using PLS-DA (Fig. 5C). Although this analysis was supervised, each strain was set as a separate level in a factor (categorical variable), and therefore the association of *pgm1Δ* and *pgm2Δ* along axis 3 (Fig. 5C) was not supervised. Four PLS axes were used for the model shown in Figure 5C. The cumulative q^2 value of 0.55 for the four axes was high in comparison to the cumulative r^2 value of 0.59, and a permutation test showed that the equivalent r^2 and q^2 values were near zero for randomized data, indicating that the model was valid and had not been overfitted (Eriksson et al. 1997). As a further test, we used canonical discriminant analysis (CDA), as an additional tool

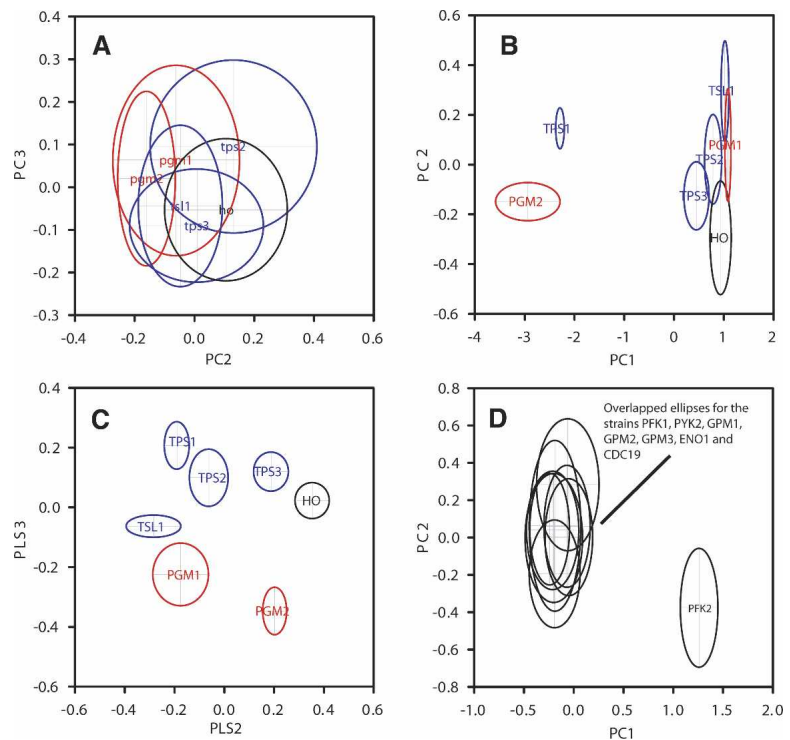


Figure 5. PCA and projection to latent structures (PLS) discriminant analysis scores plots for the extracellular metabolite profiles of strains deleted for genes involved in trehalose metabolism or in glycolysis. (A) PC2 vs. PC3 scores plot for six trehalose mutants grown on minimal medium with glucose as sole carbon source (*TPS2*, *TPS3*, and *TSL1*; blue); (*PGM1* and *PGM2*, defined by FCA to belong to a different coupled unit, red) (11–12 replicate samples for each strain). PC1 was correlated with total spectral intensity (prior to normalization) and is not shown. (B) PCA scores plot (PC1 vs. PC2) for trehalose mutants grown on galactose as sole carbon source, same strains as shown in panel A with the addition of *tps1Δ*. Reduced growth rates for the *tps1Δ* and *pgm2Δ* mutants are reflected in their separation from the other strains along PC1, which was correlated with total spectral intensity (prior to normalization) (seven to 12 replicate samples for each strain). (C) PLS scores plot (axis 2 vs. axis 3) for trehalose mutants grown on galactose as sole carbon source. PLS axis 1 separated the *tps1Δ* and *pgm2Δ* mutants from the other strains, and was correlated with total spectral intensity. (D) PCA scores plots (PC1 vs. PC2) for glycolysis mutants (nine to 12 replicate samples for each strain). *pfk2Δ* is clearly separated from the other strains along PC1 (no other differences between strains were found on lower components, either from PCA or from PLS discriminant analysis).

to PLS-DA, to predict class membership. Using the first four PCA scores as input (which explained 93% of the variance in the data), with a two-level factor (i.e., *pgm1Δ* and *pgm2Δ* as one group, and all other strains as the second group), the overall cross-validated classification success rate was 93% for leave-one-out cross-validation, with 13 out of 18 *pgmΔ* samples and 58 out of 58 of the other samples predicted successfully. However, when using leave-one-strain-out cross-validation (which best simulates whether a particular strain belongs to a particular co-set), the overall accuracy of classification dropped to 66% (weighted average), with only 29% of the *pgmΔ* samples and 84% of the other samples predicted successfully. The strain that was worst classified was *pgm1Δ*, with no correct predictions. As described above, a secondary prediction was that the *PGM1* deletion might have less of an effect on metabolite profiles than the *PGM2* deletion, as it is responsible for a smaller proportion of flux. Both the PLS-DA and CDA results support this hypothesis, suggesting that the *pgm1Δ* data partially overlap both classes. Hence, we recalculated the CDA, omitting the data for *pgm1Δ*. This resulted in 100% correct predictions for both leave-one-out and leave-one-strain-out cross-validation.

In addition to the trehalose biosynthesis pathway, we also tested some of the model's predictions about glycolysis. In glycolysis, FCA predicted that the profiles from the *pfk* mutants would be outliers when compared with those from mutants in the lower part of the pathway (Fig. 1). Metabolomic analysis showed complete separation of *pfk2Δ*, but not *pfk1Δ*, from the other strains along PC1 ($P = 5.5 \times 10^{-13}$). (*PFK1* and *PFK2* encode the α and β subunits of 6-phosphofructo-1-kinase, respectively [Heinisch 1986]) (Fig. 5D). The *pfk2Δ* strain also exhibited a slight growth delay (~6 h, see Supplemental Material), but reached the same final cell density as the other strains. The nonequivalence of the metabolic phenotypes produced by deletion of the *PFK1* or *PFK2* genes has been noted previously, where a *pfk2Δ* mutant was shown to produce higher levels of glucose 6-phosphate and glycogen than a *pfk1Δ* mutant (Huang et al. 1997). Unlike the *ura7Δ* and *ura8Δ* exometabolome spectra, which contained some residual glucose, there was no NMR-detectable glucose in the *pfk2Δ* samples, further demonstrating that the residual glucose observed with the *ura7Δ* mutants was not a consequence of the slight growth delay observed with this strain. Furthermore, the *pfk2Δ* spectra contained other regions that were clearly different from the *ura7Δ* spectra; i.e., the two strains do not just share a common growth-delay phenotype (data not shown). FCA also predicted that flux through Pyk2p can be uncoupled from the flux through Eno1p and Gpm1p, Gpm2p, Gpm3p. However, we failed to find evidence for this, which might be explained by a limited ability of alternative reactions to use phosphoenolpyruvate under the physiological conditions investigated, or the limited power of FCA to discriminate between different degrees of flux correlation.

Discussion

Deletion of a series of genes encoding enzymes and a transcription factor involved in the pathways of pyrimidine and trehalose biosynthesis, proline utilization, and glycolysis produced patterns of metabolites that were, with some notable exceptions, similar within pathways and different between them. Thus, these data strongly support the proposal that metabolite profiling can

be used to deduce gene function (Raamsdonk et al. 2001; Allen et al. 2003). More importantly, however, we have shown that this approach gives information on the topology of the metabolic network. These pathways were originally identified using well-established biochemical principles, in which the reactions were first elucidated and then related reactions grouped to form the pathway. However, the pathways, which are shown in Figure 1 in isolation, are embedded in a metabolic network in which there may be many enzyme-catalyzed routes that consume or produce the pathway intermediates. The choice of these particular pathways is, to some extent, intuitive and may neglect other functionally important routes (Papin et al. 2004).

The introduction of genome-scale metabolic network reconstructions has permitted an in silico approach to the identification of functional modules or correlated reaction sets (co-sets) within these networks. However, these unbiased network modules, which are derived mathematically from the structure of the network as a whole, require experimental validation. We have shown here that specific mutants in these pathways—*ura7Δ* and *ura8Δ* in the pyrimidine biosynthetic pathway, *pgm1Δ/pgm2Δ* in trehalose biosynthesis, *put3Δ* in proline utilization, and *pfk2Δ* in glycolysis—are outliers in terms of their metabolite profiles when compared with other members of the pathway to which they are conventionally assigned. Flux coupling analysis (FCA) (Burgard et al. 2004), which can be used to identify co-sets within a network, identified each of the steps mediated by the products of these genes as being uncoupled from the rest of their respective pathways. Thus, these data suggest that pattern recognition within NMR-based metabolite profiles can be used as an empirical tool to confirm modules or co-sets predicted by genome-scale metabolic models, and thus to validate experimentally and ultimately refine these models. These results along with recent phylogenetic approaches, which have demonstrated that enzymes of coupled reactions are often lost or gained together during evolution (Pal et al. 2005), suggest that co-sets might be natural modules of metabolic networks.

FCA assumes that the metabolic fluxes are constant, i.e., that cell metabolism is in a steady state. Although the cultures were not in a steady state, the metabolic network can be still be considered to fulfill a quasi-steady-state condition due to the separation in timescale between the dynamics of intermediary metabolism and cell growth. It is possible that acquiring intracellular metabolite profiles from cells that are in a steady state (e.g., when grown in a chemostat, or during exponential phase growth) might improve the separation of mutants observed in the multivariate statistical analysis, and this should be examined. Nevertheless, the separations that we have observed in the metabolite profiles of the mutants are robust and consistent with the predictions of FCA.

All of the mutants that produced outlier metabolite profiles, with the exception of *ura8Δ*, showed a growth defect (see Supplemental Material), and therefore it could be argued that the detection of uncoupled enzymes in a metabolic network might be more easily accomplished by measurements of growth rates. However, while growth-rate differences clearly contributed to the observed metabolic phenotypes in some of the experiments, this was not always the case. For example, in the data on the endometabolomes from the mutants in pyrimidine biosynthesis, PC1 was significantly correlated with culture OD, not with the strain. The *ura7Δ* strain did not reach an OD that was significantly different from those of the remaining strains, and the *ura8Δ* strain separated (Fig. 4A,C) even though it had no detectable growth

phenotype. Moreover, where differences in growth rate were evidently responsible for the separation observed in the metabolite profiles (e.g., the separation of *tps1Δ* and *pgm2Δ* mutants along PC1; Fig. 5B), there were components of the variance that characterized the mutations and were clearly independent of growth, for example *pgm1Δ* and *pgm2Δ* along PC4 (data not shown). In the case of proline biosynthesis, the *put3Δ* mutant was clearly separated from the other *put* mutants along PC2 (Fig. 4F), where there was no correlation with growth.

We have interpreted a lack of correlation in the metabolite profiles of strains deleted for genes encoding enzymes in the same pathway to indicate that these reactions are uncoupled. However, we cannot exclude that in some cases such a lack of correlation may also be due to the functional pleiotropy of the enzyme concerned. For example, mutation of both *PGM2* and *PFK2* has been shown to affect Ca^{2+} homeostasis (Aiello et al. 2002). Of course, an enzyme that is uncoupled from a pathway of which it is nominally a part might also be expected to have a greater probability of functional pleiotropy. Nevertheless, whether caused by uncoupling or by functional pleiotropy, the presence of a metabolite profile outlier indicates a lack of concordance with the predictions of the metabolic model and therefore the necessity for further biochemical investigation.

Although NMR detects relatively few metabolites, these seem to be amongst the most highly connected in the metabolic network, and therefore their levels can be expected to be sensitive to perturbations anywhere in the network (Brindle 2004). This connectivity argument has been advanced recently to explain why the detection of only ~100 metabolites in yeast by GC-MS is sufficient for functional analysis of a large number of different mutants (Villas-Bôas et al. 2005). The acquisition of more metabolite data, using GC- or LC-MS methods, may improve mutant discrimination, although this remains to be established from a rigorous comparison of NMR- and MS-based metabolite profiling. However, it will not necessarily give a better insight into exactly how the mutation perturbs the metabolic network. In an elegant study, Weckwerth et al. (2004) showed that a metabolite correlation analysis, using metabolite data obtained using GC-MS, was better able to discriminate between different plant mutants than other statistical methods, but that a mechanistic explanation was lacking because the precise factors that cause these correlations are unknown. The method presented here provides a relatively simple and empirical way to investigate metabolic network structure that does not require any mechanistic understanding of the link between the metabolite patterns resulting from different mutations. Further, it suggests a strategy in which measurements of growth rate could be used initially to identify uncoupled enzymes followed by metabolite profiling to identify those uncoupled enzymes where mutation has little (e.g., *ura7Δ*) or no (e.g., *ura8Δ*, *pgm1Δ*) effect on growth.

We found that, under rich media conditions, there are 131 fully coupled reaction sets (43 of which contained at least three reactions) in the yeast network, and ~45% of all metabolic reactions are members of fully coupled sets. Therefore, a considerable fraction of genes should, in principle, be classified by this method. Application of this methodology to a wider panel of mutants should both validate genome-scale metabolic models and give new insights into network structure under different metabolic conditions.

Methods

Yeast strains

Experiments were done with the strain BY4743 *MAT a/α* homozygous diploid BY4743 (*MATa/α orfΔ::kanMX4/orfΔ::kanMX4 his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 met15Δ0/MET15 ura3Δ0/ura3Δ0*), in which single genes of interest had been replaced with the kanMX deletion cassette (conferring resistance to geneticin). In the control strain, this cassette was inserted into the *HO* mating-type switching locus, which is a reportedly phenotypically neutral site (Baganz et al. 1997; Oliver et al. 1998). All of the mutant strains were obtained from the EUROFAN library. As we wished to examine the effects of single gene deletions in the biosynthetic pathway for pyrimidines, the function of the Ura3 protein was restored by transformation of the strains with the *URA3* gene amplified from S288C chromosomal DNA. To replace the *ura3Δ0* allele with a wild-type copy of the gene, the *URA3* fragment was transformed into the deletion strains using the Gietz transformation protocol (<http://www.umanitoba.ca/faculties/medicine/biochem/gietz/method.html>). Selection of transformants was made on YNB + histidine (20 mg/L) + leucine (100 mg/L), except for the *ura1Δ*, *ura2Δ*, and *ura4Δ* deletion strains, for which the selection medium also contained orotic acid (0.1%). The strains used in this study are listed in Supplemental Table 1.

Cell culture

Cells were grown on a synthetic defined minimal medium containing a cocktail of amino acids, bases, vitamins, and organic acids (Supplemental Table 2). Glucose (2%) was used as a carbon source, except where stated otherwise. Single colonies were then taken from the agar plates, inoculated into 20-mL precultures, and grown for 24 h before subculturing in 50-mL or 3-mL main cultures for analysis of intracellular or extracellular metabolites, respectively. The plates used were all between 72- and 96-h old and had never been refrigerated, as we observed subtle effects on metabolite profiles from cultures taken from either old or refrigerated plates, despite the use of an initial preculture (data not shown).

Metabolite extraction

For cell extracts ("fingerprints"), 50-mL cultures were inoculated at 5×10^5 cells/mL and grown between 24–25 h (30°C, 200 rpm), when they were in early stationary phase, reaching a typical OD of 2.5–3.5, which corresponded to a cell number of $5\text{--}7 \times 10^7$ cells/mL. The cultures were then cooled on ice, and the cells were pelleted by centrifugation (5 min, 4°C, 3000g), washed twice in ice-cold water, and extracted by a method modified from Gonzalez et al. (1997). Briefly, 5 mL of 75% ethanol v/v solution, at 80°C, was added directly to the cell pellet, together with 2 mL of 0.3 mm diameter glass beads. The HEPES buffer used by Gonzalez et al. was omitted, as it gives a number of resonances in the proton NMR spectrum that would obscure endogenous metabolite resonances. This mixture was vortexed for 30 sec, heated for 3 min to 80°C, and vortexed again for 30 sec. The supernatant was then decanted, the beads were washed with a further 2 mL of 75% ethanol v/v, and the combined extract + wash was centrifuged (10 min, 16,000g) to remove cell debris and then dried in a rotary vacuum concentrator at 30°C. This extraction procedure will not deliver a rapid arrest of cell metabolism, as has been developed for yeast cells by "quenching" them in cold methanol/water mixtures (de Koning and Van Dam 1992; Gonzalez et al. 1997; Castrillo et al. 2003). However, our aim was to achieve highly reproducible metabolite profiles, which reflected the

metabolic state of the cells, using a simple and rapid cell extraction method that allowed high throughput. For the majority of the experiments, however, extracellular metabolite profiles (“footprints”) were acquired. For these experiments, 3-mL cultures in six-well dishes were inoculated with 1% v/v ($\sim 5 \times 10^5$ cells/mL), sealed with Parafilm, and then incubated with gentle shaking (150 rpm) for 72 h at 30°C. The cultures were then cleared by centrifugation (10 min, 16,000g) and the supernatants were freeze-dried.

NMR spectroscopy

The dried media and cell extracts were dissolved in 0.8 mL of buffer (0.1 M phosphate buffer at pH 7.0, in $^2\text{H}_2\text{O}$, containing 1 mM sodium trimethylsilyl-2,2,3,3-tetradeuteropropionate; TSP) and then passed through 0.22- μm filters. The filtrate (0.6 mL) was transferred to a 5-mm NMR tube and analyzed using a Varian INOVA spectrometer interfaced to a 9.4 T magnet, giving a resonance frequency for ^1H of 400 MHz. The sample temperature was 303 K, and samples were spun at 16 Hz. One-dimensional spectra were acquired across a spectral width of 8 kHz into 32 K data points using the NOESY pulse sequence: Relaxation Delay (RD)-90- t_1 -90- t_m -90-acquire; RD = 1.5 sec, t_1 = 4 msec, t_m = 150 msec. The water resonance was suppressed by low-power irradiation during the t_m and RD periods. An additional delay of 1 sec was added prior to RD, which, with an acquisition delay of 1.95 sec, gave a total recycle time of 4.45 sec. Typically, 64 transients were acquired per sample. Spectral processing was carried out using ACDLabs NMR Processor v7.0. The summed transients were multiplied by a 0.6-Hz apodization factor prior to Fourier transformation; chemical shifts were referenced to the TSP resonance ($\delta = 0.0$), and baseline correction was performed using the Spectral Averaging routine, with typical settings of 25 points/noise = 3. Resonance assignment was made by comparison of their shifts with reported literature values, and confirmed by running authentic standards and by spiking cell extracts with these standards.

Data analysis

Data reduction was carried out by binning the spectra into 4-Hz (0.01-ppm) regions; i.e., each spectrum was divided into 0.01-ppm regions and the total signal within each region integrated between 0.50 ppm and 9.50 ppm. For each experiment, the spectra were overlaid and, for any cases where it was clear that resonances were shifted between spectra, all of the bins to which that resonance contributed were summed and treated as a single bin for subsequent data analysis. The region from 5.00 ppm to 4.50 ppm, which contained the residual signal from the suppressed water resonance, and variables corresponding to resonances from residual ethanol were excluded. This resulted in each spectrum being reduced to a vector of length 789. For the media samples, there was pronounced shifting of resonances between spectra, presumably because of larger variations in the pH of the extracellular medium. Consequently, for these samples, the spectra were not left as 0.01-ppm bins covering the entire spectral width, but rather the bins were summed to cover the signal envelope for each resonance, and bins with no resonances were excluded. This resulted in each spectrum being reduced to a vector of length 69–79 (depending on carbon source in the medium). The boundaries between integrals are given in Supplemental Table 3. The data were then normalized to the total spectral area in order to remove apparent separation between samples based solely on the total concentration of metabolites in the sample. The data were then transformed by $\log(n_i + c)$, in which c was chosen such that the variance was stabilized; i.e., the overall correlation between

intensity and standard deviation for a series of technical replicates was removed. The value of c was 0.3 for the fingerprint samples and 1 for the footprint samples. This is a simplified version of a variance-stabilizing transformation described by Purohit et al. (2004). The data were then mean-centered (subtraction of column mean from i^{th} data point) and analyzed in all cases by principal components analysis (PCA) and by partial-least-squares discriminant analysis (PLS-DA) using Simca-P 8.0 (Umetrics). Each strain was set as a different level in a factor (categorical variable) for PLS-DA. However, because all strains were treated equally in the analysis, the PLS-DA was *not* supervised with respect to classification into pathways/co-sets; i.e., in those cases where two or more strains are observed not to separate on a PLS axis, this clustering was not forced by overfitting. (We have not presented both sets of analyses for all of the data because of considerations of space; we have shown the wholly unsupervised PCA results in the majority of cases, and have included PLS-DA plots only where it clarifies a feature observed in PCA.) The number of significant axes for PLS-DA was initially judged by cross-validation: Model robustness was assessed by determination of the q^2 residual, which is the cross-validated equivalent of the r^2 summary statistic. The standard r^2 residual is given by:

$$1 - \text{RSS}/\text{TSS} \quad (1)$$

where RSS is the residual sum of squares, and TSS is the total sum of squares. The q^2 is given by:

$$1 - \text{RSS}^{\text{CV}}/\text{TSS} \quad (2)$$

where RSS^{CV} is the predictive residual sum of squares; i.e., when a point is omitted and then predicted in cross-validation, then the predictive residual is the difference between the observed and predicted values for the omitted point (or points). Thus, RSS^{CV} will be high when a number of points have high leverage, resulting in a low q^2 .

Flux coupling analysis

We used iLL672 (Kuepfer et al. 2005), a recently updated version of the genome-scale stoichiometric model of *S. cerevisiae* metabolism (Forster et al. 2003), for FCA. The model consists of 745 unique biochemical reactions (including transport processes) and 672 genes. The metabolic reconstruction gives accurate information on the stoichiometry and direction of enzymatic reactions, the presence of isoenzymes, and on enzymatic complexes. The model also includes a biomass reaction, which represents the relative contribution of major biosynthetic compounds to the biomass of yeast.

Stoichiometric network analyses start by assuming steady-state for each metabolite in the network, written in the form of $S \times v = 0$, where S is the $m \times n$ stoichiometric matrix, m is the number of metabolites, n is the number of reactions, and v is the vector of individual fluxes through the network. An individual element S_{ij} gives the contribution of the j^{th} reaction to metabolite i . Availability of nutrients and directions of individual reactions are included as boundary conditions (nutrient availability was set up to mimic the medium used in the experiments). FCA (Burgard et al. 2004) is based on fixing the flux through one reaction, and then maximizing and minimizing the flux through all other reactions in turn by linear programming (we used the optimization package Cplex 7.5 to solve these problems). This procedure gives the minimum and maximum of flux ratios for any two reactions in steady state. Two types of coupling relationships were considered here: (1) fully coupled reactions, which have fixed flux ratios (the minimum and maximum of flux ratios are equal) and are always used together; and (2) directionally coupled reactions,

where the activity of one reaction implies activity of the other, but not necessarily the reverse. In contrast, the flux ratios of uncoupled reactions can take up any value. Calculations were run without assuming a constant biomass composition to avoid coupling of a large set of fluxes to the biomass reaction (thus, all biomass components were allowed to be drained independently of one another) (Burgard et al. 2004).

Acknowledgments

We thank Evgeni Nikolaev for discussions on flux coupling analysis. B.P. is a Fellow of the Human Frontier Science Program. Work on metabolomics and systems biology in S.G.O.'s laboratory and on NMR-based metabolomics in K.M.B.'s laboratory is supported by the BBSRC and the Wellcome Trust.

References

- Aiello, D.P., Fu, L., Miseta, A., and Bedwell, D.M. 2002. Intracellular glucose 1-phosphate and glucose 6-phosphate levels modulate Ca²⁺ homeostasis in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277**: 45751–45758.
- Allen, J., Davey, H.M., Broadhurst, D., Heald, J.K., Rowland, J.J., Oliver, S.G., and Kell, D.B. 2003. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat. Biotechnol.* **21**: 692–696.
- Allen, J., Davey, H., Broadhurst, D., Rowland, J., Oliver, S., and Kell, D.B. 2004. Discrimination of modes of action of antifungal substances by use of metabolic footprinting. *Appl. Environ. Microbiol.* **70**: 6157–6165.
- Baganz, F., Hayes, A., Marren, D., Gardner, D.C.J., and Oliver, S.G. 1997. Suitability of replacement markers for functional analysis studies in *Saccharomyces cerevisiae*. *Yeast* **13**: 1563–1573.
- Boles, E., Liebetrau, W., Hofmann, M., and Zimmerman, F.K. 1994. A family of hexosephosphate mutases in *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **220**: 83–96.
- Borodina, I. and Nielsen, J. 2005. From genomes to in silico cells via metabolic networks. *Curr. Opin. Biotechnol.* **16**: 350–355.
- Brandriss, M.C. and Magasanik, B. 1979. Genetics and physiology of proline utilization in *Saccharomyces cerevisiae*: Mutation causing constitutive enzyme expression. *J. Bacteriol.* **140**: 504–507.
- Brindle, K.M. 2004. Magnetic resonance spectroscopy and imaging—Powerful tools for functional genomics. In *Metabolic engineering in the post genomic era*. (eds. B.N. Kholodenko and H.V. Westerhoff), pp. 69–87. Horizon Bioscience, UK.
- Burgard, A.P., Nikolaev, E.V., Schilling, C.H., and Maranas, C.D. 2004. Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Res.* **14**: 301–312.
- Castrillo, J., Hayes, A., Mohammed, S., Gaskell, S.J., and Oliver, S. 2003. An optimized protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry. *Phytochemistry* **62**: 929–937.
- de Koning, W. and Van Dam, K. 1992. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Anal. Biochem.* **204**: 118–123.
- Duarte, N.C., Herrgard, M.J., and Palsson, B.O. 2004. Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res.* **14**: 1298–1309.
- Eriksson, L., Johansson, E., and Wold, S. 1997. Quantitative structure–activity relationship model validation. In *Quantitative structure–activity relationships in environmental sciences*. (eds. F. Chen and G. Schuurmann), pp. 381–397. SETAC Press, Pensacola, FL.
- Famili, I., Forster, J., Nielson, J., and Palsson, B.O. 2003. *Saccharomyces cerevisiae* phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proc. Natl. Acad. Sci.* **100**: 13134–13139.
- Forster, J., Famili, I., Fu, P., Palsson, B.O., and Nielsen, J. 2003. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* **13**: 244–253.
- Giaever, G., Chu, A., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**: 387–391.
- Gonzalez, B., Francois, J., and Renaud, M. 1997. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast* **13**: 1347–1355.
- Heinisch, J. 1986. Construction and physiological characterization of mutants disrupted in the phosphofructokinase genes of *Saccharomyces cerevisiae*. *Curr. Genet.* **11**: 227–234.
- Huang, D., Wilson, W., and Roach, P. 1997. Glucose-6-P control of glycogen synthase phosphorylation. *J. Biol. Chem.* **272**: 22495–22501.
- Kaderbhai, N., Broadhurst, D., Ellis, D., Goodacre, R., and Kell, D. 2003. Functional genomics via metabolic footprinting: Monitoring metabolite secretion by *Escherichia coli* tryptophan metabolism mutants using FT-IR and direct injection electrospray mass spectrometry. *Comp. Funct. Genom.* **4**: 376–391.
- Kell, D., Brown, M., Davey, H., Dunn, W., Spasic, I., and Oliver, S. 2005. Metabolic footprinting and systems biology: The medium is the message. *Nat. Rev. Microbiol.* **3**: 557–565.
- Kuepfer, L., Sauer, U., and Blank, L.M. 2005. Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*. *Genome Res.* **15**: 1421–1430.
- Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., et al. 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**: 799–804.
- Oliver, S.G., Winson, M.K., Kell, D.B., and Baganz, F. 1998. Systematic functional analysis of the yeast genome. *Trends Biotechnol.* **16**: 373–378.
- Ozier-Kalogeropoulos, O., Adeline, M.T., Yang, W.L., Carman, G.M., and Lacroute, F. 1994. Use of synthetic lethal mutants to clone and characterize a novel CTP synthetase gene in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **242**: 431–439.
- Pal, C., Papp, B., and Lercher, M.J. 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* **37**: 1372–1375.
- Papin, J.A., Reed, J.L., and Palsson, B.O. 2004. Hierarchical thinking in network biology: The unbiased modularization of biochemical networks. *Trends Biochem. Sci.* **29**: 641–647.
- Patil, K.R. and Nielsen, J. 2005. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Natl. Acad. Sci.* **102**: 2685–2689.
- Price, N.D., Reed, J.L., and Palsson, B.O. 2004. Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nat. Rev. Microbiol.* **2**: 886–897.
- Purohit, P.V., Rocke, D.M., Viant, M.R., and Woodruff, D.L. 2004. Discrimination models using variance-stabilizing transformation of metabolomic NMR data. *OMICS* **8**: 118–130.
- Raamsdonk, L.M., Teusink, B., Broadhurst, D., Zhang, N., Hayes, A., Walsh, M.C., Berden, J.A., Brindle, K.M., Kell, D.B., Rowland, J.J., et al. 2001. Functional genomics via the metabolome: A FANCY approach to characterising mutations with a silent phenotype. *Nat. Biotechnol.* **19**: 45–50.
- Ram, A.F.J., Wolters, A., Tenhoopen, R., and Klis, F.M. 1994. A new approach for isolating cell-wall mutants in *Saccharomyces cerevisiae* by screening for hypersensitivity to calcofluor white. *Yeast* **10**: 1019–1030.
- Sellick, C.A. and Reece, R.J. 2003. Modulation of transcription factor function by an amino acid: Activation of Put3p by proline. *EMBO J.* **22**: 5147–5153.
- Teusink, B., Walsh, M.C., van Dam, K., and Westerhoff, H.V. 1998. The danger of metabolic pathways with turbo design. *Trends Biochem. Sci.* **23**: 162–169.
- Villas-Bôas, S.G., Moxley, J.F., Åkesson, M., Stephanopoulos, G., and Nielsen, J. 2005. High-throughput metabolic state analysis: The missing link in integrated functional genomics of yeasts. *Biochem. J.* **388**: 669–677.
- Weckwerth, W., Loureiro, M.E., Wenzel, K., and Fiehn, O. 2004. Differential metabolic networks unravel the effects of silent plant phenotypes. *Proc. Natl. Acad. Sci.* **101**: 7809–7814.
- Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., et al. 1999. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**: 901–906.

Received June 19, 2006; accepted in revised form January 3, 2007.