Division of labor in metabolic regulation by transcription, translation, acetylation and phosphorylation

Sriram Chandrasekaran[1,2,*]

[1] - Department of Biomedical Engineering, [2] - Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA, 48109

* - Correspondence: csriram@umich.edu

Abstract

The metabolism of most organisms is controlled by a diverse cast of regulatory processes, including transcriptional regulation and post-translational modifications (PTMs). Yet how metabolic control is distributed between these regulatory processes is unknown. Here we present *Comparative Analysis of Regulators of Metabolism* (CAROM), an approach that compares regulators based on network connectivity, flux, and essentiality of their reaction targets. Using CAROM, we analyze transcriptome, proteome, acetylome and phospho-proteome dynamics during transition to stationary phase in *E. coli* and *S. cerevisiae*. CAROM uncovered that the targets of each regulatory process shared unique metabolic properties: growth-limiting reactions were regulated by acetylation, while isozymes and futile-cycles were preferentially regulated by phosphorylation. Reversibility, essentiality, and molecular-weight further distinguished reactions controlled through diverse mechanisms. While every enzyme can be potentially regulated by multiple mechanisms, analysis of context-specific datasets reveals a conserved partitioning of metabolic regulation based on reaction attributes.

Author summary

There are several ways to regulate an enzyme's activity in a cell. Yet, the design principles that determine when an enzyme is regulated by transcription, translation or post-translational modifications are unknown. Each control mechanism, such as transcription, comprises several regulators that control a distinct set of targets. So far, it is unclear if similar partitioning of targets occurs at a higher level, between different control mechanisms. Here we systematically analyze patterns of metabolic regulation in model microbes. We find that five key parameters can distinguish the targets of each mechanism. These key parameters provide insights on specific roles played by each mechanism in determining overall metabolic activity. This approach may help define the basic regulatory architecture of metabolic networks.

Introduction

A myriad control mechanisms regulate microbial metabolic adaptation to new environments [1–8]. Nevertheless, microbes deploy distinct regulatory mechanisms to regulate enzyme activity in response to specific environmental challenges. For example, *B. subtilis* cells primarily utilize transcriptional regulation when glucose is available, but post-transcriptionally regulate metabolic enzymes after malate addition [9]. In both *E. coli* and yeast, some pathways, such as glycolysis, are predominantly regulated by post-transcriptional regulation, while others, such as the TCA cycle, are regulated at the transcriptional level [1,3,10]. This suggest that apart from differences in response time, specific mechanisms are deployed for specialized regulatory tasks. Nevertheless, it is unclear why some enzymes are regulated using acetylation or via other PTMs such as phosphorylation [3,4].

Numerous advantages of regulation by PTMs have been proposed over the past five decades [11–13]. These include low energy requirements, rapid response, and signal amplification. Yet these

characteristics are not unique to PTMs, and these features also do not differentiate between PTMs such as acetylation and phosphorylation. The staggering complexity of each regulatory process has limited the comparative analysis of metabolic regulation at a systems level [3]. Existing studies have focused on a small set of metabolic pathways or on a single regulatory process [4,10,14–20]. Such studies have revealed reaction reversibility and metabolic network structure to be predictive of regulation [8,16,21–24]. Yet these studies do not shed light on the differences between each regulatory process. In sum, although some general network principles of regulation are known, how it is partitioned among various regulatory mechanisms is unclear.

We hence developed a data-driven approach, called *Comparative Analysis of Regulators of Metabolism* (CAROM), to identify unique features of each regulatory process. CAROM achieves this by comparing various properties of metabolic enzymes, including essentiality, flux, molecular weight and topology. It identifies those properties that are highly enriched among targets of each process than expected out of random chance.

Results and Discussion

Here we focus on four well-studied control mechanisms with available omics datasets - transcription, post-transcription, phosphorylation and acetylation. We analyzed the dynamics of metabolic regulation during a well-characterized process in yeast, namely, transition to stationary phase. We obtained RNA sequencing, time-course proteomics, acetylomics, and phospho-proteomics data from the literature [25–27]. Targets for each process were determined based on differential levels between stationary and exponential phase (Methods). We assumed that PTMs and other regulatory sites that are dynamic and conditionally regulated are likely to be functional [28].

The targets of diverse regulatory mechanisms were used as input to CAROM. CAROM analyzes the properties of the targets in the context of a genome-scale metabolic network model of yeast [29]. We hypothesized that differences in target preferences between diverse regulators can be inferred from the network topology and fluxes. Protein and gene targets of each process were mapped to corresponding metabolic reactions in the model. There was significant overlap among reactions regulated through changes in both the transcriptome and proteome, and transcriptome and acetylome (hypergeometric p-value = 5×10^{-25} and 1×10^{-15} respectively; S. Table 1). In contrast, there was little overlap between targets of phosphorylation with other mechanisms (p-value > 0.1; S. Table 1). While prior studies found higher overlap between targets of PTMs [30,31], they used all possible sites that can be acetylated or phosphorylated. However, only a fraction of PTM sites are likely to be active and functional in a single condition. Overall, each regulatory mechanism had a distinct set of targets (Figure 1A).

What are the common features of enzymes that are regulated by each mechanism? To answer this, we used CAROM to compare the regulation of enzymes that are essential for growth in minimal media. Essential enzymes in the yeast metabolic model were determined using Flux Balance Analysis (FBA) [32]. Surprisingly, this set of enzymes was highly enriched among those regulated by acetylation but not by other processes (ANOVA p-value < 10^{-16}; Figure 1B; S. Table 2). Since regulation can be optimized for fitness across multiple conditions [33], we identified enzymes that impact growth in 87 different nutrient conditions comprising various carbon and nitrogen sources using FBA. This set of essential enzymes was once again enriched for acetylation relative to other mechanisms (ANOVA p-value < 10^{-16}; S. Figure 1). This trend was observed using experimentally derived list of essential genes as well (hypergeometric p-value = 2×10^{-7} for acetylation). Interestingly, in contrast to acetylation, genes regulated at the proteomic level were significantly under-represented among the essential genes (hypergeometric p-value of depletion = 8×10^{-11}). Thus, essential enzymes are likely to be constitutively

expressed and their activity modulated through acetylation. This may explain why transcriptional regulation has minimal impact on fluxes in central metabolism, which contain several growth-limiting enzymes [3,10].

We next used CAROM to determine the impact of reaction position in the network on its regulation. We counted the number of pathways each reaction is involved in, along with other topological metrics, such as the closeness, degree and page rank. We found that the regulation of enzymes differed significantly based on network topology (Figure 1C). First, reactions with low connectivity, measured through any of the topological metrics, were highly likely to be unregulated. In contrast, highly connected enzymes linking multiple pathways were more likely to be regulated by PTMs. Interestingly, reactions regulated by both the PTMs had the highest connectivity (S. Figures 2, 3). Several key hubs, such as acetyl-CoA acetyltransferase, hexokinase and phosphofructokinase are regulated by at least 2 different mechanisms (S. Table 3).

We next assessed how regulation differs based on the magnitude and direction of flux through the network. We inferred the full range of fluxes possible through each reaction using flux variability analysis (FVA) [34]. Since yeast cells may not optimize their metabolism for biomass synthesis during transition to stationary phase, we also performed FVA without assuming biomass maximization. We found that irreversible reactions were highly likely to be regulated (S. Figure 4). A recent study found the same trend for allosteric regulation as well [21]. However, reversibility alone did not differentiate between regulatory mechanisms.

Interestingly, reactions that have the potential to carry high fluxes were predominantly regulated by phosphorylation (Figure 1D; ANOVA p-value < 10^{-16}). This set of phosphorylated reactions comprise several kinase-phosphatase pairs, enzymes that are part of loops that consume energy ("futile cycles"), or reactions that have isozymes in compartments such as vacuoles or nucleus (S. Table 4). Thus, phosphorylation in this condition selectively regulates reactions to avoid futile cycling between antagonizing reactions or those operating in different compartments. Using data from experimentally constrained fluxes from Hackett *et al* study [21] revealed similar patterns of regulation (S. Figure 5). Reactions with the highest flux, such as ATP synthase, phosphofructokinase, and nucleotide kinase, were also regulated by multiple mechanisms.

Finally, we compared regulation based on fundamental enzyme properties: catalytic activity and molecular weight. While catalytic activity was similar across the targets of all mechanisms, targets of phosphorylation had the highest molecular weight (p-value < 10^{-16}) (S. Figures 6,7). There is a weak correlation between molecular weight and maximum flux (Pearson's correlation R = 0.02), suggesting that both maximum flux and molecular weight are likely to be independent predictors of regulation by phosphorylation.

To check if this pattern of regulation is observed in other conditions, using CAROM, we analyzed data from nitrogen starvation response and the cell cycle in yeast, where both phospho-proteomics and transcriptomics data are available [35–38]. A similar trend of regulation was observed in these conditions with phosphorylation regulating isozymes and enzymes that can carry high fluxes (futile cycles) (Figure 2). Since isozymes arise frequently from gene duplication, our results may explain the observation that duplicated genes are more likely to be regulated by phosphorylation [39].

Since many mechanisms of metabolic regulation are evolutionarily conserved, we next analyzed data from *E. coli* cells during stationary phase [40–42]. By analyzing transcriptomics, proteomics, acetylomics and phosphoproteomics data using the *E. coli* metabolic network model, CAROM uncovered that the pattern of regulation observed in yeast was also observed in *E. coli* (Figure 3).

Reactions that were regulated in *E. coli* had higher topological connectivity compared to those that were unregulated. Further, essential reactions were enriched for regulation by acetylation, and reactions with high maximum flux or large enzyme molecular weight were enriched for regulation by phosphorylation. However, in contrast to yeast, phosphorylation impacted very few metabolic genes in *E. coli,* and may play a relatively minor role in this specific context. Phosphorylation had 20-fold fewer targets compared to other mechanisms, and its targets overlapped significantly with other processes (S. Tables 5-6).

In sum, our analysis reveals a unique distribution of regulation within the metabolic network (Figure 4). Within each process, it is well known that individual regulators such as transcription factors or kinases have their own unique set of targets. Here we find that similar specialization occurs at a higher scale, involving diverse processes. Reaction properties identified by CAROM to be associated with distinct regulatory mechanisms may be related to specific functions performed by each regulator. For example, phosphorylation may represent a mechanism of feedback regulation to control futile cycles and high flux reactions that consume ATP [6,43]. Finally, this pattern of regulation is context specific – predictive features such as reaction flux or essentiality can change between conditions and influence regulation. Further, while most essential reactions were regulated, a small subset (14%) were not found to be regulated by any mechanism. These enzymes could be sites of allosteric regulation or other regulatory mechanisms not covered here due to the lack of context specific datasets (S. Table 7). Overall, these results are robust to the thresholds used for finding differentially regulated sites, using data from different sources, and other modeling parameters (S. Tables 8-12).

Since microbes exhibit a wide range of metabolic behaviors, it is not possible to uncover regulation in each condition through experiments. We need tools like CAROM to identify factors that determine the deployment of regulatory mechanisms in a metabolic context. Although flux balance analysis of metabolic models can accurately forecast optimal flux distribution, it does not provide insights on how the flux rewiring is achieved. Our analysis predicts regulatory mechanisms that will likely orchestrate flux adjustments based on reaction attributes. This can guide drug discovery and metabolic engineering efforts by identifying regulators that are dominant in different parts of the network [44]. CAROM can be applied to uncover target specificities of other regulators such as non-coding RNAs and PTMs, and help understand the architecture of metabolic regulation in a wide range of organisms.

Methods

CAROM

The CAROM approach takes as input a list of genes that are the targets of one or more regulatory processes. It compares the properties of the targets and identifies significant differences in target properties between mechanisms using ANOVA. Overall, CAROM compares the following 13 properties:

- Impact of gene knockout on biomass production, ATP synthesis, and viability across 87 different conditions
- Flux through the network measured through Flux Variability analysis and PFBA, reaction reversibility
- Enzyme molecular weight and catalytic activity
- The total pathways each reaction is involved in, its Degree, Closeness and PageRank

The CAROM source-code is available from the Synapse bioinformatics repository https://www.synapse.org/CAROM

Processing omics data

We used RNA-sequencing data from Treu *et al* 2014 that compared the expression profile of S. cerevisiae between mid-exponential growth phase with early stationary phase [27]. A 2-fold change threshold was used to identify differentially expressed genes. Lysine acetylation and protein phosphorylation data were obtained from the Weinert *et al* 2014 study that compared PTM levels between exponentially growing and stationary phase cells using *stable isotope labeling with amino acids in cell culture* (SILAC) [26]. A 2-fold change threshold of the protein-normalized PTM data was used to identify differentially expressed PTMs. Proteomics data was taken from Murphy *et al* time-course proteomics study [25]. The hoteling T2 statistic defined by the authors was used to identify proteins differentially expressed during diauxic shift; the top 25% of the differentially expressed proteins were assumed to be regulated. Proteomics data from Weinert *et al* was also used as an additional control and we observed the same trends using this data as well (S. Table 10). Further, we repeated the analysis after removing genes that were not expressed during transition to stationary phase; the transcripts for a total of 12 genes out of the 910 in the model were not detected by RNA-sequencing in the Treu *et al* study [27]. Removing the 12 genes did not impact any of the results (S. Table 9).

As additional validation, we used periodic data from the yeast cell cycle. Time-course SILAC phospho-proteomics data was obtained from Touati *et al* [37]. Phospho-sites whose abundance declined to less than 50% or increased by more than 50% at least two consecutive timepoints were considered dephosphorylated or phosphorylated respectively as defined by the authors. Transcriptomics data was taken from Kelliher *et al* study that identified 1246 periodic transcripts using periodicity-ranking algorithms [38].

The phospho-proteomics and transcriptome data during nitrogen shift was obtained from Oliveira *et al* [35,36]. The nitrogen shift studies compared the impact of adding glutamine to yeast cells growing on a poor nitrogen source (proline alone or glutamine depletion) with cells growing on a rich nitrogen source (glutamine plus proline). A 2-fold change threshold was used to identify differentially expressed transcripts and phospho-sites.

E. coli acetylation data was taken from the Weinert *et al* study comparing actively growing exponential phase cells to stationary phase cells [42]. Proteomics and transcriptomics were from Houser *et al* study of *E. coli* cells in early exponential phase and stationary phase [41]. Phospho-proteomics data for exponential and early stationary phase *E. coli* cells was taken form Soares *et al* [40]. We used a 2-fold change ($p < 0.05$) threshold for all studies.

The results are robust to the thresholds used for identifying differentially expressed genes or proteins (S. Table 11). In all studies, genes and proteins that are either up or down regulated were considered to be regulated. The final data set table used for all comparative analyses is provided as a supplementary material (S. Table 13).

Genome scale metabolic modeling

We used the yeast metabolic network reconstruction (Yeast 7) by Aung *et al*, which contains 3,498 reactions, 910 genes and 2,220 metabolites [29]. The analysis of *E. coli* data was done using the IJO1366 metabolic model [45]. All analyses were performed using COBRA toolbox for MATLAB [46].

The impact of gene knockouts on growth was determined using flux balance analysis (FBA). FBA identifies an optimal flux through the metabolic network that maximizes an objective, usually the production of biomass. A minimal glucose media (default condition) was used to determine the impact of gene knockouts. Further, gene knockout analysis was repeated in a set of 87 different minimal nutrient conditions to identify genes that impact growth across diverse conditions; these conditions span all carbon and nitrogen sources that can support growth in the Yeast 7 model. The number of times each gene was found to be lethal (growth < 0.01 units) across all conditions was used as a metric of essentiality.

To infer topological properties, a reaction adjacency matrix was created by connecting reactions that share metabolites. We used the Centrality toolbox function in MATLAB to infer all network topological attributes including centrality, degree and PageRank.

Flux Variability Analysis (FVA) was used to infer the range of fluxes possible through every reaction in the network. Two sets of flux ranges were obtained with FVA – the first with optimal biomass and the latter without assuming optimality. In the second case, the fluxes are limited by the availability of nutrients and energetics alone, thus it reflects the full range of metabolic activity possible in a cell. Reactions with maximal flux above 900 units were assumed to be unconstrained and were excluded from the analysis, as they are likely due to thermodynamically infeasible internal cycles [47]; the choice of this threshold for flagging unconstrained reactions did not impact the distribution between regulators over a wide range of values (S. Table 12).

For fitting experimentally derived flux data from Hackett *et al* [21], reactions were fit to the fluxes using linear optimization and the flux through remaining reactions that do not have experimentally derived flux data were inferred using FVA. Analysis using a related approach for inferring fluxes – PFBA, did not reveal any significant difference as PFBA eliminates futile cycles and redundancy by minimizing total flux through the network while maximizing for biomass [48] (S. Figure 5).

Reaction reversibility was determined directly from the model annotations. We also used additional reversibility annotation from Martinez *et al* based on thermodynamics analysis of the Yeast metabolic model [49]. Pathway annotations, enzyme molecular weight and catalytic activity values were obtained

from Sanchez *et al* [50]. The comparative analysis of regulatory mechanisms was also repeated using the updated Yeast 7.6 model and yielded similar results (S. Table 8) [50].

The comparative analysis of target properties was done using gene-reaction pairs rather than genes or reactions alone; the gene-reaction pairs accounts for regulation involving all possible combinations of genes and associated reaction, including isozymes that may involve different genes but the same reaction or multi-functional enzymes involving same the gene associated with different reactions. The 910 genes and 2310 gene-associated reactions resulted in 3375 unique gene-reaction pairs in yeast.

All statistical tests were performed using MATLAB. Significance of overlap between lists was estimated using the hypergeometric test. Significance of the differences in distribution of target properties between mechanisms were determined using ANOVA, the non-parametric Kruskal-Wallis test, and after multiple hypothesis correction (S. Table 8).

Figures

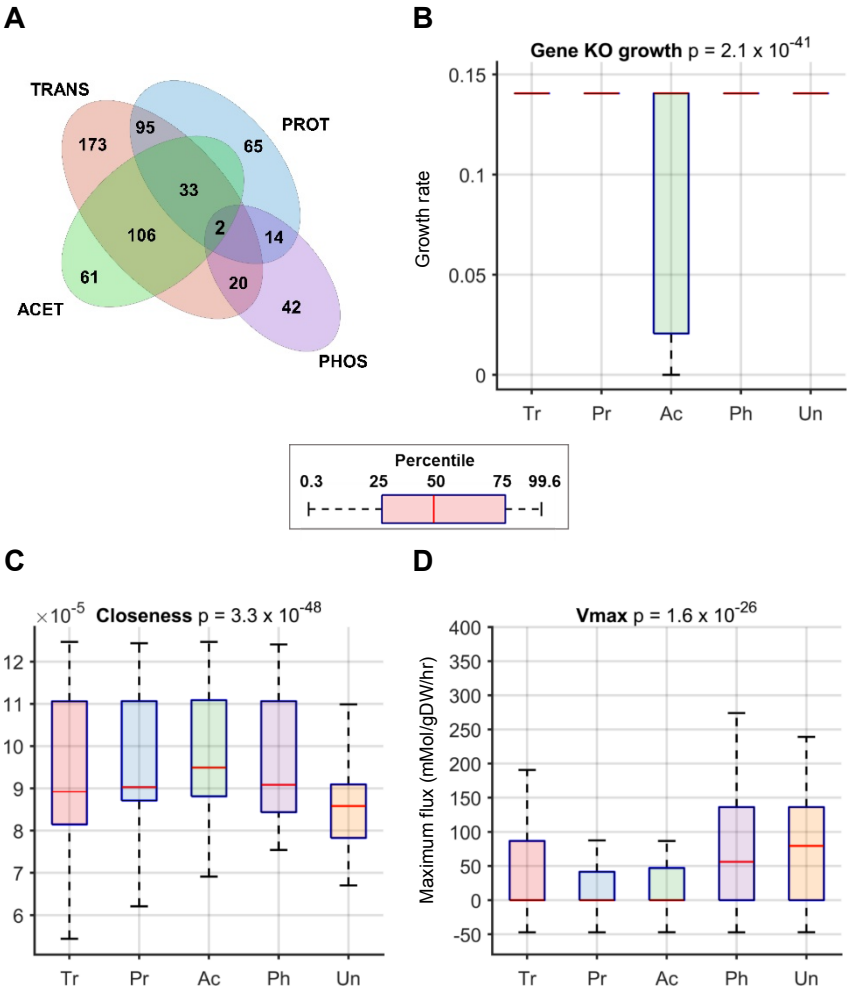


Figure 1. Comparison of the properties of the targets of regulation in yeast during transition to stationary phase. A. The Venn diagram shows the extent of overlap between targets of each process. Only 2 genes were found to be regulated by all four mechanisms. Targets of phosphorylation did not show any significant overlap with other mechanisms, while transcriptome and proteome showed the highest overlap (S. Table 1). **B.** Enzymes that impact growth when knocked out are highly likely to be acetylated. **C.** Enzymes with poor connectivity, as measured through the network connectivity metric - closeness, are more likely to be Unregulated. **D.** Enzymes catalyzing reactions with high maximum flux are likely to be either regulated through phosphorylation or to be unregulated. The Anova p-value comparing the differences in means is shown in the title. (Abbreviation: transcription (Tr), post-transcription (Pr), acetylation (Ac), phosphorylation (Ph) or Unregulated (Un)).

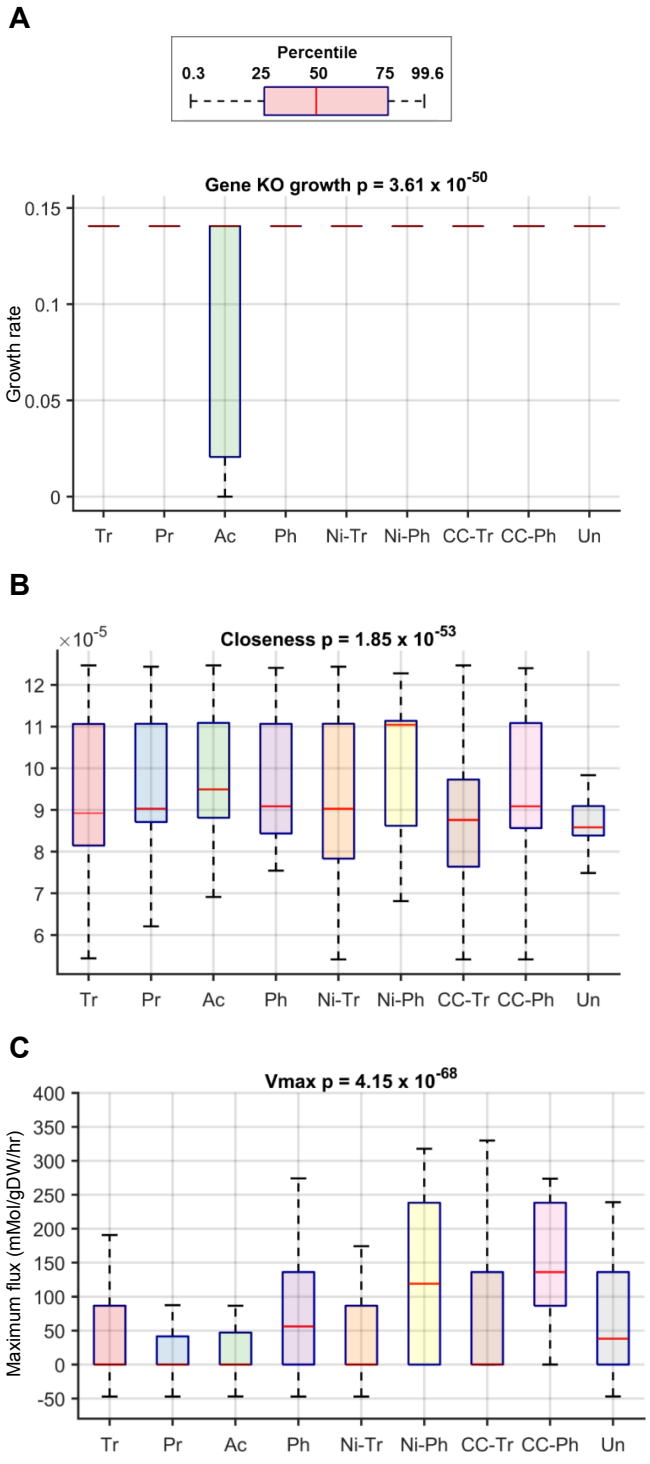


Figure 2. Comparison of the properties of enzymes in yeast regulated by each mechanism during the cell cycle (CC-Tr, CC-Ph) and nitrogen starvation (Ni-Tr, Ni-Ph). Data from stationary phase conditions (transcription (Tr), post-transcription (Pr), acetylation (Ac), phosphorylation (Ph) or Unregulated (Un)) are shown for comparison. Similar to stationary phase, enzymes that impact growth when knocked out are likely to be acetylated **(A),** enzymes that are highly connected are likely to be regulated by one of the four mechanisms **(B)** and those that catalyze reactions with high flux are likely to be regulated through phosphorylation in all three conditions **(C)**. The Anova p-value comparing the differences in means is shown in the title.

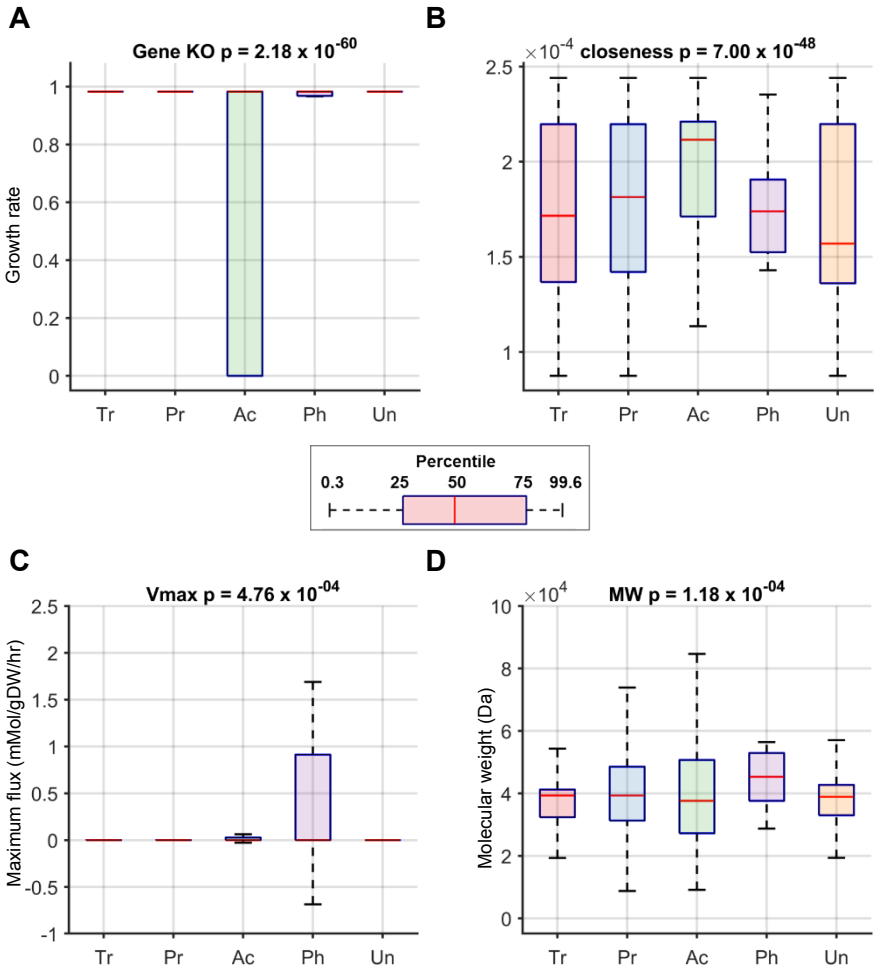


Figure 3. Comparison of the properties of enzymes in *E. coli* regulated by transcription (Tr), post-transcription (Pr), acetylation (Ac), phosphorylation (Ph) or Unregulated (Un) during transition to stationary phase. Similar to yeast, reaction essentiality **(A)**, connectivity **(B)**, maximum flux **(C)** and molecular weight (**D)** are predictive of regulation by acetylation, all four mechanisms, and phosphorylation (Vmax, MW) respectively.

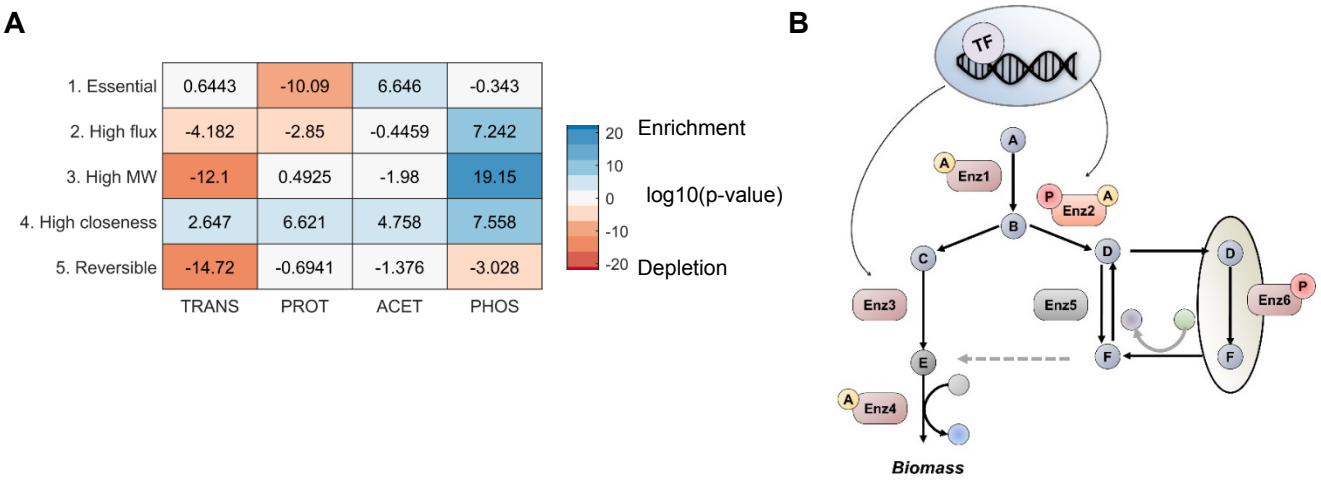


Figure 4. Reaction attributes predictive of regulation by each process in yeast. **A.** The heatmap shows the statistical enrichment and depletion of the targets of each process among reactions that are - (1) essential, (2) have high maximum flux (Vmax > 75th percentile), (3) catalyzed by enzymes with high molecular weight (MW > 75th percentile), (4) highly connected (Closeness > 75th percentile), and (5) reversible. The log-transformed p-values from hypergeometric test are shown with a positive sign for enrichment and negative sign for depletion. **B.** A schematic pathway summarizing the division of labor in metabolic regulation. Essential reactions (Enz1 and Enz4) are preferentially acetylated; reactions in futile cycles and in different compartments (Enz6) are phosphorylated; non-essential enzymes with low connectivity are regulated through transcriptional regulation (Enz3), and reactions with high connectivity are regulated through multiple mechanisms (Enz2). Reversible reactions are predominantly unregulated (Enz5).

Acknowledgments: Funding: This work was supported by faculty start-up funds from the University of Michigan to SC. **Author contributions:** SC conceived the study, designed and performed research, and wrote the manuscript. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All datasets are available in the supplementary materials

References

1. Nielsen J. Systems Biology of Metabolism. Annu Rev Biochem. 2017; doi:10.1146/annurev-biochem-061516-044757
2. Cho BK, Zengler K, Qiu Y, Park YS, Knight EM, Barrett CL, et al. The transcription unit architecture of the Escherichia coli genome. Nat Biotechnol. 2009; doi:10.1038/nbt.1582
3. Chubukov V, Gerosa L, Kochanowski K, Sauer U. Coordination of microbial metabolism. Nature Reviews Microbiology. 2014. doi:10.1038/nrmicro3238
4. Heinemann M, Sauer U. Systems biology of microbial metabolism. Curr Opin Microbiol. 2010/03/12. 2010;13: 337–343. doi:10.1016/j.mib.2010.02.005
5. Aebersold R, Agar JN, Amster IJ, Baker MS, Bertozzi CR, Boja ES, et al. How many human proteoforms are there? Nature Chemical Biology. 2018. doi:10.1038/nchembio.2576
6. Kochanowski K, Sauer U, Noor E. Posttranslational regulation of microbial metabolism. Curr Opin Microbiol. 2015;27: 10–17.
7. Ihmels J, Levy R, Barkai N. Principles of transcriptional control in the metabolic network of Saccharomyces cerevisiae. Nat Biotechnol. 2004; doi:10.1038/nbt918
8. Stadtman ER. Mechanisms of Enzyme Regulation in Metabolism. Enzymes. 1970; doi:10.1016/S1874-6047(08)60171-7
9. Buescher JM, Liebermeister W, Jules M, Uhr M, Muntel J, Botella E, et al. Global network reorganization during dynamic adaptations of Bacillus subtilis metabolism. Science (80-). 2012; doi:10.1126/science.1206871
10. Daran-Lapujade P, Rossell S, van Gulik WM, Luttik MAH, de Groot MJL, Slijper M, et al. The fluxes through glycolytic enzymes in Saccharomyces cerevisiae are predominantly regulated at posttranscriptional levels. Proc Natl Acad Sci. 2007; doi:10.1073/pnas.0707476104
11. Holzer H, Duntze W. Metabolic Regulation by Chemical Modification of Enzymes. Annu Rev Biochem. 1971; doi:10.1146/annurev.bi.40.070171.002021
12. Fell D, Cornish-Bowden A. Understanding the control of metabolism. Portland press London; 1997.
13. Stadtman ER, Chock PB. Interconvertible Enzyme Cascades in Metabolic Regulation. Current Topics in Cellular Regulation. 1978. doi:10.1016/B978-0-12-152813-3.50007-0
14. Zhao S, Xu W, Jiang W, Yu W, Lin Y, Zhang T, et al. Regulation of cellular metabolism by protein lysine acetylation. Science (80-). 2010;327: 1000–1004.

15. Oliveira AP, Ludwig C, Picotti P, Kogadeeva M, Aebersold R, Sauer U. Regulation of yeast central metabolism by enzyme phosphorylation. Mol Syst Biol. 2012; doi:10.1038/msb.2012.55

16. Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, et al. Just-in-time transcription program in metabolic pathways. Nat Genet. 2004; doi:10.1038/ng1348

17. Lee JM, Gianchandani EP, Eddy JA, Papin JA. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. PLoS Comput Biol. 2008/05/17. 2008;4: e1000086. doi:10.1371/journal.pcbi.1000086

18. Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO. Integrating high-throughput and computational data elucidates bacterial networks. Nature. 2004/05/07. 2004;429: 92–96. doi:10.1038/nature02456nature02456 [pii]

19. Shen F, Boccuto L, Pauly R, Srikanth S, Chandrasekaran S. Genome-scale network model of metabolism and histone acetylation reveals metabolic dependencies of histone deacetylase inhibitors. Genome Biol. 2019;20: 49. doi:10.1186/s13059-019-1661-z

20. Chandrasekaran S, Price ND. Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in Escherichia coli and Mycobacterium tuberculosis. Proc Natl Acad Sci. National Acad Sciences; 2010;107: 17845–17850.

21. Hackett SR, Zanotelli VRT, Xu W, Goya J, Park JO, Perlman DH, et al. Systems-level analysis of mechanisms regulating yeast metabolic flux. Science (80-). 2016; doi:10.1126/science.aaf2786

22. Almaas E, Kovács B, Vicsek T, Oltvai ZN, Barabási AL. Global organization of metabolic fluxes in the bacterium Escherichia coli. Nature. 2004; doi:10.1038/nature02289

23. Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED. Metabolic network structure determines key aspects of functionality and regulation. Nature. 2002; doi:10.1038/nature01166

24. Stelling J, Sauer U, Szallasi Z, Doyle FJ, Doyle J. Robustness of cellular functions. Cell. 2004. doi:10.1016/j.cell.2004.09.008

25. Murphy JP, Stepanova E, Everley RA, Paulo JA, Gygi SP. Comprehensive Temporal Protein Dynamics during the Diauxic Shift in Saccharomyces cerevisiae . Mol Cell Proteomics. 2015; doi:10.1074/mcp.m114.045849

26. Weinert BT, Iesmantavicius V, Moustafa T, Schölz C, Wagner SA, Magnes C, et al. Acetylation dynamics and stoichiometry in Saccharomyces cerevisiae. Mol Syst Biol. 2014;10: 716.

27. Treu L, Campanaro S, Nadai C, Toniolo C, Nardi T, Giacomini A, et al. Oxidative stress response and nitrogen utilization are strongly variable in Saccharomyces cerevisiae wine strains with different fermentation performances. Appl Microbiol Biotechnol. 2014; doi:10.1007/s00253-014-5679-6

28. Beltrao P, Bork P, Krogan NJ, Van Noort V. Evolution and functional cross-talk of protein post-translational modifications. Molecular Systems Biology. 2013. doi:10.1002/msb.201304521

29. Aung HW, Henry SA, Walker LP. Revising the representation of fatty acid, glycerolipid, and glycerophospholipid metabolism in the consensus model of yeast metabolism. Ind Biotechnol. 2013;9: 215–228. doi:10.1089/ind.2013.0013

30. Oliveira AP, Sauer U. The importance of post-translational modifications in regulating Saccharomyces cerevisiae metabolism. FEMS Yeast Research. 2012. doi:10.1111/j.1567-1364.2011.00765.x

31. Beltrao P, Albanèse V, Kenner LR, Swaney DL, Burlingame A, Villén J, et al. Systematic functional prioritization of protein posttranslational modifications. Cell. 2012; doi:10.1016/j.cell.2012.05.036

32. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? Nat Biotechnol. Nature Research; 2010;28: 245–248. doi:10.1038/nbt.1614

33. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. Multidimensional optimality of microbial metabolism. Science (80-). 2012; doi:10.1126/science.1216882

34. Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng. 2003;5: 264–276. doi:10.1016/j.ymben.2003.09.002

35. Oliveira AP, Dimopoulos S, Busetto AG, Christen S, Dechant R, Falter L, et al. Inferring causal metabolic signals that regulate the dynamic TORC1-dependent transcriptome. Mol Syst Biol. 2015; doi:10.15252/msb.20145475

36. Oliveira AP, Ludwig C, Zampieri M, Weisser H, Aebersold R, Sauer U. Dynamic phosphoproteomics reveals TORC1-dependent regulation of yeast nucleotide and amino acid biosynthesis. Sci Signal. 2015; doi:10.1126/scisignal.2005768

37. Touati SA, Kataria M, Jones AW, Snijders AP, Uhlmann F. Phosphoproteome dynamics during mitotic exit in budding yeast. EMBO J. 2018; doi:10.15252/embj.201798745

38. Kelliher CM, Leman AR, Sierra CS, Haase SB. Investigating Conservation of the Cell-Cycle-Regulated Transcriptional Program in the Fungal Pathogen, Cryptococcus neoformans. PLoS Genet. 2016; doi:10.1371/journal.pgen.1006453

39. Amoutzias GD, He Y, Gordon J, Mossialos D, Oliver SG, Van de Peer Y. Posttranslational regulation impacts the fate of duplicated genes. Proc Natl Acad Sci. 2010; doi:10.1073/pnas.0911603107

40. Soares NC, Spät P, Krug K, MacEk B. Global dynamics of the Escherichia coli proteome and phosphoproteome during growth in minimal medium. J Proteome Res. 2013; doi:10.1021/pr3011843

41. Houser JR, Barnhart C, Boutz DR, Carroll SM, Dasgupta A, Michener JK, et al. Controlled Measurement and Comparative Analysis of Cellular Components in E. coli Reveals Broad Regulatory Changes in Response to Glucose Starvation. PLoS Comput Biol. 2015; doi:10.1371/journal.pcbi.1004400

42. Weinert BT, Iesmantavicius V, Wagner SA, Schölz C, Gummesson B, Beli P, et al. Acetyl-phosphate is a critical determinant of lysine acetylation in E. coli. Mol Cell. 2013;51: 265–272.

43. Humphrey SJ, James DE, Mann M. Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. Trends in Endocrinology and Metabolism. 2015. doi:10.1016/j.tem.2015.09.013

44. Choi KR, Jang WD, Yang D, Cho JS, Park D, Lee SY. Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. Trends in Biotechnology. 2019. doi:10.1016/j.tibtech.2019.01.003

45. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, et al. A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011. Mol Syst Biol. 2011/10/13. 2011;7: 535. doi:10.1038/msb.2011.65

46. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. Nat Protoc. 2007/04/05. 2007;2: 727–738. doi:nprot.2007.99 [pii]10.1038/nprot.2007.99

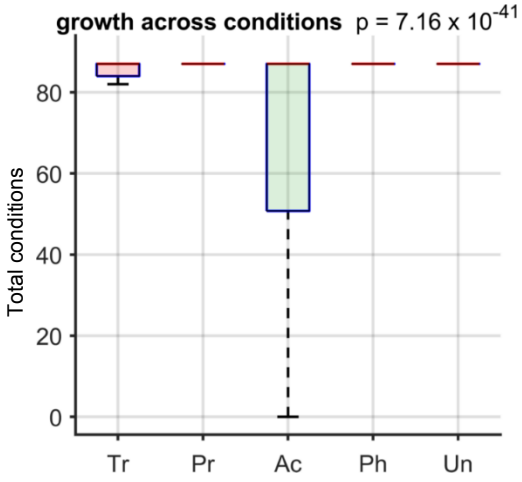
47. Schellenberger J, Lewis NE, Palsson B. Elimination of thermodynamically infeasible loops in steady-state metabolic models. Biophys J. 2011; doi:10.1016/j.bpj.2010.12.3707

48. Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, et al. Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. Mol Syst Biol. 2010;6: 390.

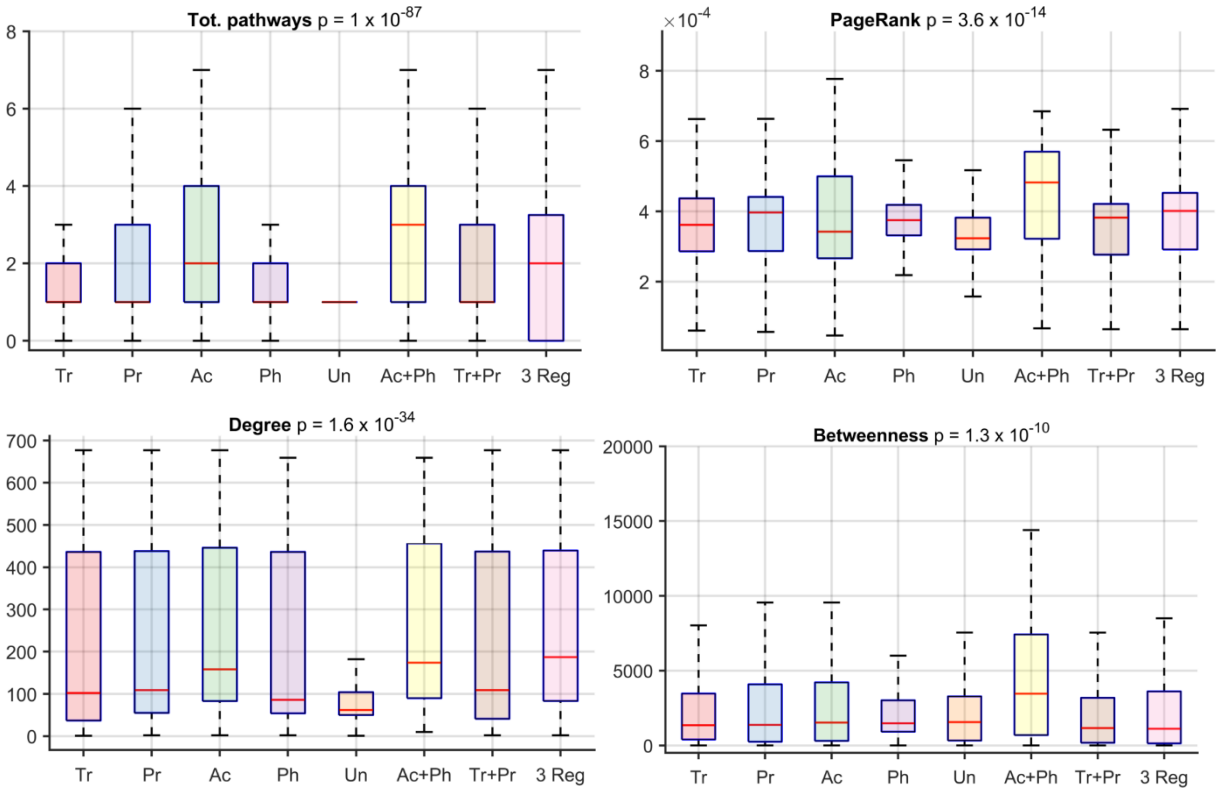
49. Martínez VS, Quek LE, Nielsen LK. Network thermodynamic curation of human and yeast genome-scale metabolic models. Biophys J. 2014; doi:10.1016/j.bpj.2014.05.029

50. Sánchez BJ, Zhang C, Nilsson A, Lahtvee P, Kerkhoven EJ, Nielsen J. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. Mol Syst Biol. 2017; doi:10.15252/msb.20167411

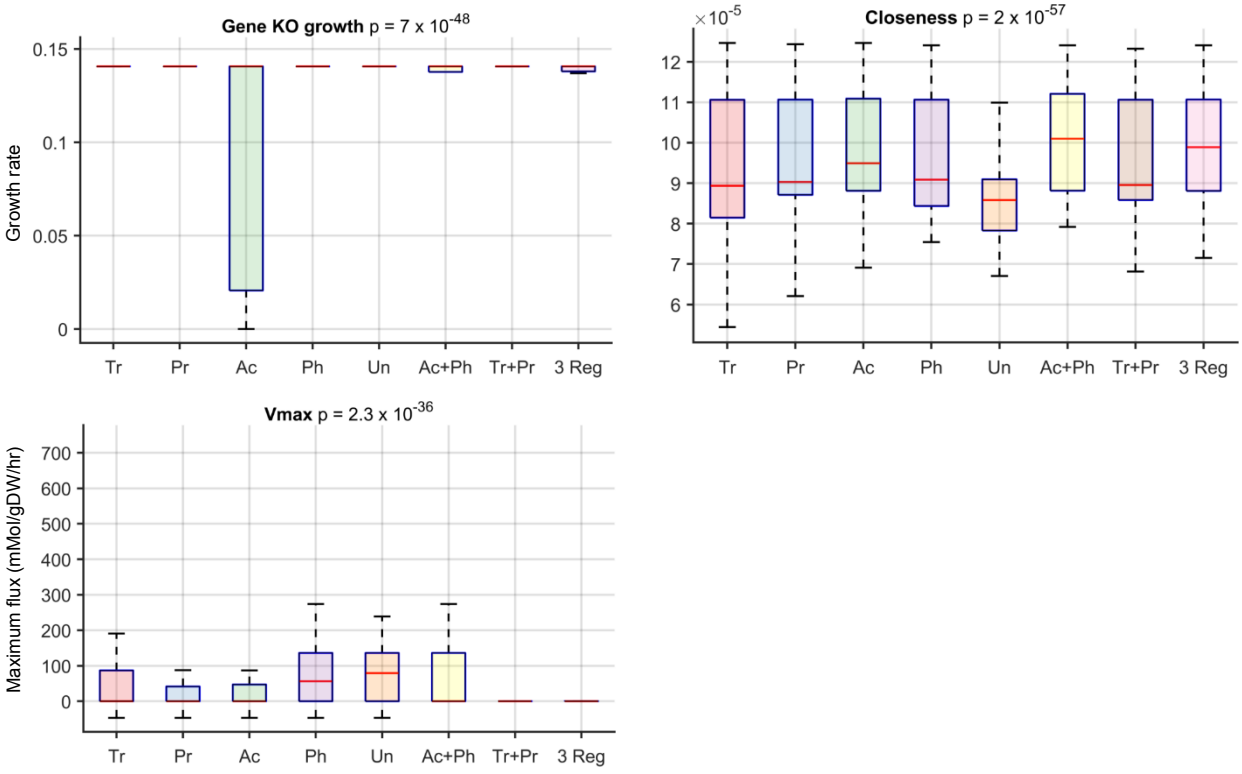
Supplementary Figures



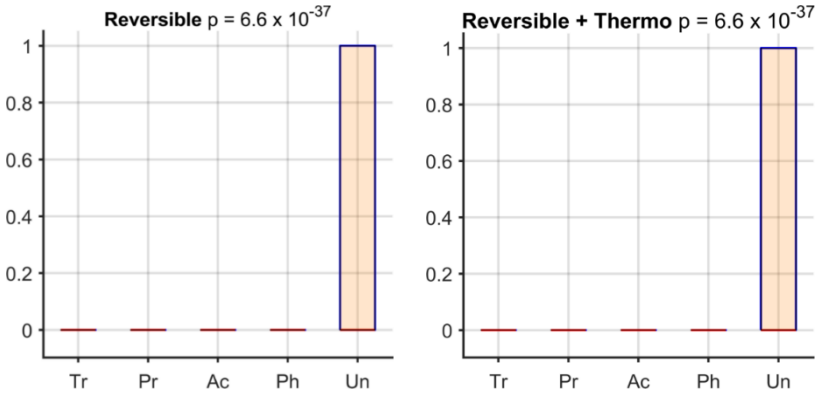
S. Figure 1. Distribution of regulation based on gene essentiality across 87 different conditions. These conditions comprise 56 different carbon sources including glucose, and 31 different nitrogen sources including ammonium ions. The total number of conditions in which each gene deletion was viable was calculated. This total number was then compared between targets of each regulatory mechanism. The box plots show that acetylation preferentially regulates the genes that impact growth across the 87 conditions. The box plot whiskers extend to the 99.3rd percentile of each distribution. The ANOVA p-value comparing the means is 7.1 x 10^{-41}.



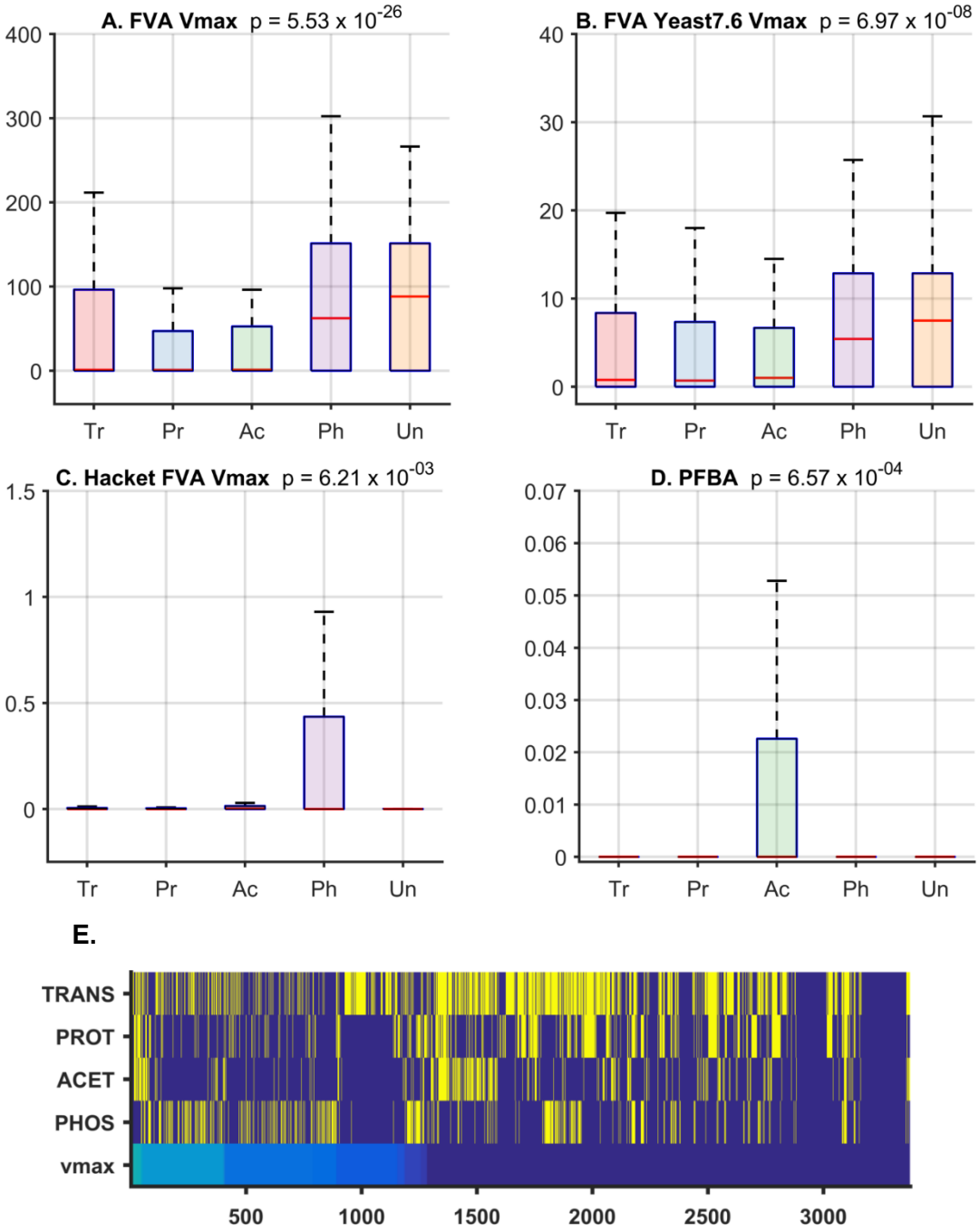
S. Figure 2. Distribution of regulation based on topological properties of each reaction. Four different topological properties are shown in the box plots - the total number of annotated pathways each reaction participates (Tot. pathways), the number of times each reaction is traversed during a random walk between reactions in the network (Pagerank), the total number of connected reactions (Degree) and the number of times each reaction appears on a shortest path between two reactions (Betweenness). These show that reactions that are regulated by any mechanism have a higher connectivity compared to those that are unregulated. Furthermore, reactions regulated by both acetylation and phosphorylation had the highest connectivity across all metrics. The ANOVA p-value comparing the means is provided in the title. (Abbreviations: regulation by both transcription and post-transcription (Tr + Pr), both acetylation and phosphorylation (Ac + Ph), at least 3 regulators (3 Reg), and Unregulated (Un)).



S. Figure 3. Properties of reactions regulated by multiple mechanisms. The box plots compare the properties of enzymes regulated by transcription, post-transcription, acetylation, phosphorylation with those regulated by both transcription and post-transcription (Tr + Pr), both acetylation and phosphorylation (Ac + Ph), or at least 3 regulators (3 Reg). This set of combinations among regulators was chosen as both acetylation and phosphorylation are PTMs, and the transcriptome and proteome of yeast cells show significant correlation. Reactions regulated by both acetylation and phosphorylation had the highest connectivity as measured by the inverse sum of the distance from a reaction to all other reactions in the network (Closeness). Apart from connectivity, reactions regulated by two different mechanisms did not share properties of reactions regulated by each individual mechanism. For example, reactions regulated by acetylation and phosphorylation were not likely to be essential or have high maximum flux. The ANOVA p-value comparing the means is provided in the title.

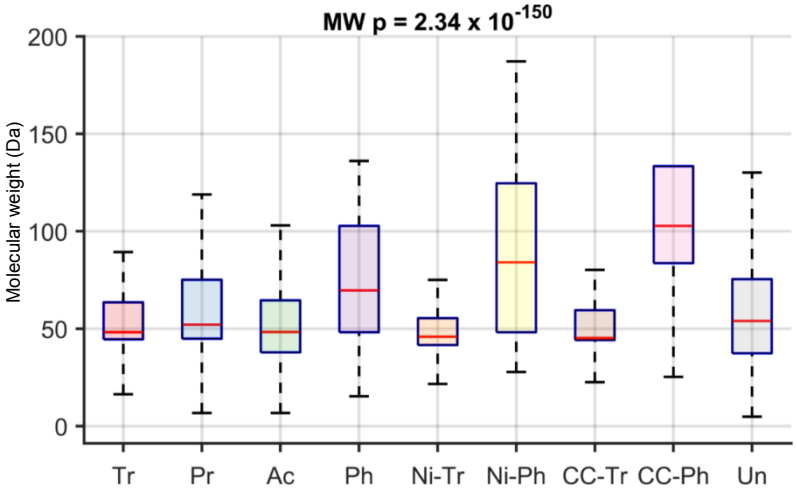


S. Figure 4. The box plots show the distribution of regulation based on reaction reversibility. Reversible reactions were highly likely to be not regulated by any of the four mechanisms. The left panel compares the distribution of regulation of reversible reactions based on the annotation from the Yeast 7 model (reversible reactions are set to 1 and irreversible reactions are set to 0). The panel on the right uses an updated list based on thermodynamic analysis of the Yeast metabolic model by Martinez *et al* [49].

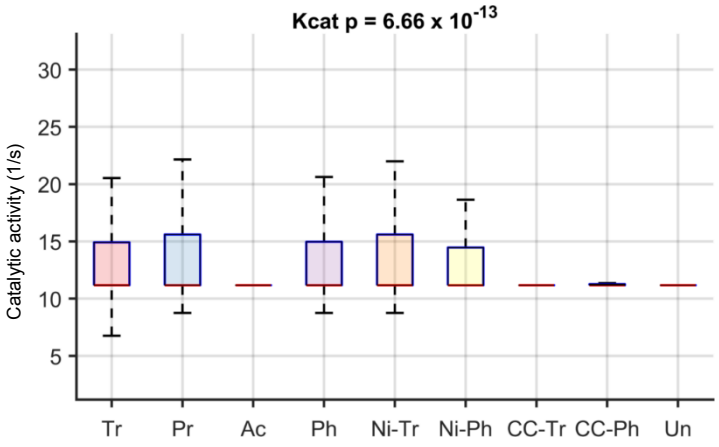


S. Figure 5. Distribution of regulation based on magnitude of maximum possible flux (mmol/gDW/hr) through each reaction. The plots compare the distribution of regulation using flux calculated using various methods and models. The ANOVA p-value comparing the means is provided in the panel title of each plot. These results show that phosphorylated reactions are highly enriched among those reactions with high maximum flux. **A.** Maximum flux through each reaction was calculated using FVA using the Yeast 7 model without assuming that cells maximize their biomass (the default objective in FVA and FBA). The box plots compare the maximum flux value of reactions regulated by each mechanism. **B.** Maximum flux through each reaction was calculated using FVA without assuming that cells maximize their biomass using the Yeast 7.6 model (Yeast 7 model was used for all analyses). **C.** The flux through

the model was first fit to the experimentally inferred flux data from Hackett *et al*[21]. The maximum flux through all reactions was then determined using FVA. **D.** The flux through each reaction was inferred from Parsimonious FBA (PFBA). Note that PFBA does not provide the maximum flux but the flux value that minimizes the sum of flux through all reactions while maximizing the biomass objective. Hence it does not reveal any futile cycles or redundancy in the network. **E.** The heatmap shows the distribution of regulation based on magnitude of maximum possible flux (Vmax) through of each reaction. Reactions are sorted based on Vmax inferred from FVA. The columns correspond to each reaction-gene pair. Those that are regulated by each mechanism are shown in yellow, while those that are not regulated by a specific mechanism are in blue.



S. Figure 6. The box plots show the distribution of regulation in Yeast based on enzyme molecular weight. Enzymes regulated by phosphorylation on average tended to have high molecular weight. Data for targets of phosphorylation and transcriptional regulation in Nitrogen starvation (denoted by 'Ni_' prefix) and Cell cycle (denoted by 'CC_' prefix) conditions are also shown for comparison.



S. Figure 7. The box plots show the distribution of regulation based on enzyme catalytic activity (kcat) in Yeast (data from Sanchez *et al* [50]). No consistent difference across datasets was observed in regulation based on the catalytic activity of the target enzyme.

Supplementary Tables

A

Regulatory mechanisms		Reaction Overlap	p-value
TRANS	PROT	421	4.12×10^{-30}
TRANS	ACET	285	2.36×10^{-19}
TRANS	PHOS	266	0.241723
PROT	ACET	133	8.53×10^{-05}
PROT	PHOS	117	0.925481
ACET	PHOS	89	0.420549

B

Regulatory mechanisms		Gene Overlap	p-value
TRANS	PROT	153	0.010941
TRANS	ACET	157	0.001552
TRANS	PHOS	61	0.931005
PROT	ACET	69	0.925509
PROT	PHOS	42	0.291789
ACET	PHOS	34	0.860463

C

Total regulators	Percentage among those regulated
2 or more	47.8%
3 or more	8.7%
All 4	0.08%

S. Table 1. Overlap between targets of various mechanisms - transcription (TRANS), post-transcription (PROT), acetylation (ACET), phosphorylation (PHOS). This reveals low overlap between targets of regulation by phosphorylation and other mechanisms. **A.** Overlap between target reactions **B.** Overlap between target genes. **C.** Percentage of reactions regulated by multiple mechanisms. Overall, 69% of the gene-associated reactions in the model were regulated; among those regulated, 47.8% were regulated by more than one mechanism.

S. Table 2. Essential reactions regulated by acetylation (Spreadsheet file)

S. Table 3. Top 50 reactions sorted based on topological connectivity (Spreadsheet file)

S. Table 4. Top 50 reactions with maximum reaction flux regulated by phosphorylation (Spreadsheet file)

Regulation	*E. coli*	S. cerevisiae
Transcription	469	468
Post-transcription/Proteomic	372	266
Acetylation	460	265
Phosphorylation	17	133

S. Table 5. Comparison of total genes regulated by each process in *E. coli* with *S. cerevisiae* shows that phosphorylation plays a relatively minor role in *E. coli* metabolic regulation during stationary phase.

Regulatory mechanisms		p-value	Reaction Overlap
TRANS	PROT	3.77×10^{-23}	590
ACET	TRANS	0.042022	442
ACET	PROT	0.192068	379
ACET	PHOS	5.60×10^{-11}	28
PHOS	TRANS	0.004853	22
PHOS	PROT	0.95148	9

S. Table 6. Overlap between targets of various mechanisms in *E. coli* - transcription (TRANS), post-transcription (PROT), acetylation (ACET), phosphorylation (PHOS).

S. Table 7. Gaps in regulation – Essential genes that are unregulated. One representative reaction is shown for each gene in case there are multiple reactions associated with it (Spreadsheet file)

Model	Yeast 7 (default)		Yeast 7.6	
p-value	ANOVA	Kruskal-Wallis	ANOVA	Kruskal-Wallis
Growth rate	2.07×10^{-41}	1.24×10^{-29}	7.82×10^{-43}	1.37×10^{-47}
Closeness	3.33×10^{-48}	1.74×10^{-55}	1.66×10^{-39}	3.61×10^{-51}
Vmax (without max. biomass)	5.53×10^{-26}	9.25×10^{-13}	6.97×10^{-8}	1.30×10^{-11}

S. Table 8. Robustness of the results comparing the difference in distribution of properties between targets of various regulatory mechanisms using the Yeast 7.6 model. Significance of results using the non-parametric Kruskal-Wallis test is also shown. The p-values for the key reaction features shown in Figure 1 using the Yeast 7 model is provided as comparison. All p-values are significant at FDR < 0.01 using both Bonferroni adjustment and Benjamin-Hochberg multiple hypothesis correction.

Model	All genes (default)		All expressed genes	
p-value	ANOVA	Kruskal-Wallis	ANOVA	Kruskal-Wallis
Growth rate	2.07×10^{-41}	1.24×10^{-29}	3.3×10^{-41}	2.1×10^{-29}
Closeness	3.33×10^{-48}	1.74×10^{-55}	2.2×10^{-49}	3.5×10^{-56}
Vmax	1.59×10^{-26}	2.51×10^{-21}	8.1×10^{-27}	1.3×10^{-21}

S. Table 9. Robustness of the results after removing genes that are not-expressed (i.e. not detected in RNA-seq data) in both exponential and stationary phase cultures. The p-values reported in Figure 1 using all the metabolic genes in the Yeast 7 model is provided as comparison.

Model	Murphy *et al* (default)		Weinert *et al*	
p-value	ANOVA	Kruskal-Wallis	ANOVA	Kruskal-Wallis
Growth rate	2.07×10^{-41}	1.24×10^{-29}	1.59×10^{-37}	2.63×10^{-33}
Closeness	3.33×10^{-48}	1.74×10^{-55}	3.10×10^{-44}	1.62×10^{-48}
Vmax	1.59×10^{-26}	2.51×10^{-21}	1.71×10^{-22}	1.94×10^{-19}

S. Table 10. Comparison of results using proteomics data from Weinert *et al* instead of Murphy *et al*. The ANOVA p-value comparing the means are provided. The p-values reported in Figure 1 using Murphy *et al* data is provided as comparison.

Fold change	2 (default)	1.5	3	4
Growth rate	2.07×10^{-41}	3.72×10^{-35}	1.12×10^{-34}	5.82×10^{-26}
Closeness	3.33×10^{-48}	2.09×10^{-47}	1.13×10^{-32}	6.95×10^{-20}
Vmax (with max. biomass)	1.59×10^{-26}	6.97×10^{-24}	1.26×10^{-23}	1.01×10^{-19}

Top Percentile	25 (default)	50	15	5
Growth rate	2.07×10^{-41}	1.42×10^{-42}	2.49×10^{-39}	2.71×10^{-40}
Closeness	3.33×10^{-48}	1.86×10^{-41}	7.15×10^{-55}	9.66×10^{-48}
Vmax (with max. biomass)	1.59×10^{-26}	6.07×10^{-18}	5.66×10^{-31}	8.98×10^{-36}

S. Table 11. Comparison of thresholds used for identifying differentially expressed genes and proteins. These show that our results are robust to the thresholds for identifying the targets of various regulatory mechanisms. The ANOVA p-value comparing the means are provided. Note that the first table uses fold change thresholds for transcriptomics, acetylation and phospho-proteomics data alone. Since the

608 proteomics data uses a percentile cut off, the robustness analysis for this data was performed
609 separately.

610

611

Threshold for unconstrained reactions	ANOVA p-value for Vmax
100	5.07×10^{-31}
200	1.08×10^{-57}
300	5.27×10^{-50}
400	3.53×10^{-27}
500	2.24×10^{-25}
600	2.24×10^{-25}
700	2.24×10^{-25}
800	1.59×10^{-26}
900	1.59×10^{-26}
1000	4.68×10^{-150}

612

613 **S. Table 12.** Comparison of thresholds used for identifying unconstrained reactions from FVA.
614 Reactions with maximal flux above the threshold listed in the table were assumed to be unconstrained
615 and were excluded from the analysis, as they are likely due to thermodynamically infeasible internal
616 cycles. The ANOVA p-value comparing the means of the maximum flux through the target reactions of
617 different regulatory mechanisms is provided. The default value (900 mmol/gDW/hr) for eliminating
618 unconstrained reactions is highlighted and was used for all analyses. These show that our results are
619 robust to the thresholds for identifying unconstrained reactions.

620

621 **S. Table 13.** Raw dataset containing all yeast genes and associated reactions, the corresponding
622 regulators, and the reaction properties (Spreadsheet file).

623