

## A two-enzyme adaptive unit in bacterial folate metabolism

Andrew F. Schober<sup>1,2</sup>, Christine Ingle<sup>1</sup>, Junyoung O. Park<sup>3,4</sup>, Li Chen<sup>3,5</sup>, Joshua D. Rabinowitz<sup>3,5</sup>, Ivan Junier<sup>6</sup>, Olivier Rivoire<sup>7</sup> and Kimberly A. Reynolds<sup>1,2\*</sup>

<sup>1</sup>*The Green Center for Systems Biology, and* <sup>2</sup>*Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA*

<sup>3</sup>*Lewis-Sigler Institute for Integrative Genomics,* <sup>4</sup>*Department of Chemical and Biological Engineering, and* <sup>5</sup>*Department of Chemistry, Princeton University, Princeton, NJ 08544, USA*

<sup>6</sup>*Centre National de la Recherche Scientifique (CNRS), Université Grenoble Alpes, TIMC-IMAG, F-38000 Grenoble, France*

<sup>7</sup>*Center for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS, INSERM, PSL Research University, F-75005 Paris, France*

\*Please address correspondence to:

Kimberly Reynolds  
Green Center for Systems Biology  
UT Southwestern Medical Center  
6001 Forest Park Rd.  
Dallas, TX 75390-8597

Ph: 214-645-6157  
[kimberly.reynolds@utsouthwestern.edu](mailto:kimberly.reynolds@utsouthwestern.edu)

**The activity of a gene may be influenced or modified by other genes in the genome. Here, we show that co-evolution can be used to identify quasi-independent gene groups inside of larger cellular systems. Using folate metabolism as a case study, we show that co-evolution indicates a sparse architecture of interactions, with three small groups of genes co-evolving in the midst of others that evolve independently. For one such module – dihydrofolate reductase (DHFR) and thymidylate synthase (TYMS) – we use epistasis measurements and forward evolution to demonstrate both internal functional coupling and independence from the remainder of the genome. Mechanistically, the coupling is driven by a constraint on their relative activities, which must be balanced to prevent accumulation of a metabolic intermediate. Applying co-evolution analyses genome-wide reveals a number of other gene pairs with statistical signatures similar to DHFR/TYMS, suggesting that small adaptive units are a general feature of cellular systems.**

**Keywords:**

co-evolution, comparative genomics, folate metabolism, dihydrofolate reductase (DHFR), epistasis, experimental evolution

# Introduction

The collective action of enzymes in metabolism provides the basic materials for cells to grow and divide. Functional coupling between metabolic enzymes in a pathway can constrain enzyme activity, binding affinity, and expression level (Barenholz et al., 2017; Salvador and Savageau, 2003, 2006). Because the pattern of constraints within and between proteins remains largely unknown, our ability to rationally engineer new systems (Kim and Copley, 2012; Michener et al., 2014a; Michener et al., 2014b), understand how the cell adapts to perturbations (Kim et al., 2010; Long et al., 2018), and quantify the relationship between mutations and disease (Kondrashov et al., 2002; Zuk et al., 2012) is limited. In all of these instances, one needs to be able to answer the following basic question: if the activity or expression level of a particular enzyme is perturbed, which (if any) of the other enzymes in metabolism would require compensatory modifications? More generally, an ability to globally map such coupling between proteins and to identify groups of enzymes within a pathway that adapt and function as a unit would help render cellular systems more tractable and predictable.

To begin to address this problem, we used comparative genomics to study the pattern of selective constraints between enzymes in bacterial folate metabolism. The central premise of this approach is that functional couplings between proteins drive co-evolution of the associated genes, leaving behind a set of detectable statistical signatures in extant genomes. Co-evolution can be manifested in different ways, and here we use two measures: conservation of proximity on the chromosome (synteny, (Huynen et al., 2000; Janga et al., 2005; Junier and Rivoire, 2013)) and the coordinated loss and gain of genes across species (sometimes called phylogenetic profiling, (Kim and Price, 2011; Pellegrini et al., 1999; Rivoire, 2013)). Our analyses indicate three small groups of genes within folate metabolism that behave as independent co-evolutionary

units: the genes within a group co-evolve with each other, but each group appears to evolve relatively independently from the remainder of folate metabolism (and the genome).

If functional constraints between enzymes indeed underlie co-evolutionary patterns, then co-evolutionary maps may provide an approach to understanding how pathways adapt to perturbation. The prediction is simple: when part of a co-evolutionary unit is perturbed, we expect compensatory mutations within genes in that unit, but not elsewhere in the pathway or genome. To test this hypothesis, we chose the most strongly co-evolving enzyme pair, consisting of dihydrofolate reductase (DHFR) and thymidylate synthase (TYMS). Using quantitative measurements of epistasis and forward evolution, we demonstrate that these two enzymes are (1) epistatically coupled through a shared metabolite and (2) adapt independently from the remainder of the genome. In particular, we observe that mutations in these two enzymes are sufficient for *E. coli* to acquire resistance to the inhibition of DHFR by trimethoprim. Extending our statistical analyses of co-evolution genome wide reveals additional gene pairs that strongly co-evolve with one another yet are mostly independent from the rest of the genome, suggesting that the finding of small adaptive units embedded within larger cellular systems may be a general feature of metabolic networks.

## Results

### An evolution-based map of functional coupling in folate metabolism

The core one-carbon folate metabolic pathway consists of thirteen enzymes that interconvert various folate species and produce methionine, serine, thymidine and purine nucleotides in the process (Fig. 1A and Table S1) (Green and Matthews, 2013). The input of the pathway is 7,8-dihydrofolate (DHF), produced by the bifunctional enzyme dihydrofolate synthase/folylpolyglutamate synthetase (FPGS) through the addition of L-glutamate to

dihydropteroate. Once DHF is formed, it is reduced to 5,6,7,8-tetrahydrofolate (THF) by the enzyme dihydrofolate reductase (DHFR) using NADPH as a co-factor. THF can then be modified by a diversity of one-carbon groups at both the N-5 and N-10 positions, and subsequently serves as a one-carbon donor in several critical reactions, including the synthesis of serine and purine nucleotides (Fig. 1A, bottom square portion of pathway). The only step that oxidizes THF back to DHF is catalyzed by the enzyme thymidylate synthase (TYMS), which modifies uridine monophosphate (dUMP) to thymidine monophosphate (dTMP) in the process.

This pathway is well conserved across bacteria, ensuring good statistics for our analysis. Due to the central role of these metabolites in cell growth and division, the function of folate metabolism can be readily assessed through quantitative growth rate measurements (Reynolds et al., 2011). Further, folate metabolism is a target of several well-known antibiotics (trimethoprim and sulfamethoxazole), and chemotherapeutics (methotrexate and 5-fluorouracil) (Ducker and Rabinowitz, 2016; Gangjee et al., 2007). These factors enable experimental strategies to measure gene function, adaptation, and epistatic coupling *in vivo*. Thus, folate metabolism provides a good model system to examine the use of co-evolution in identifying adaptive units, and to study the functional mechanisms underlying evolutionary coupling in a metabolic pathway.

To study co-evolution in this pathway, we computed two different statistical measures across 1445 bacterial genomes: 1) synteny, the conservation of chromosomal proximity between genes and 2) co-occurrence, the coordinated loss and gain of genes across species. These measures have the advantage of being relatively straightforward to compute, provide two separate measures of evolutionary coupling, and were previously established as reliable indicators of protein functional relationships (Huynen et al., 2000; Janga et al., 2005; Junier and Rivoire, 2016; Kim and Price, 2011; Pellegrini et al., 1999; Snel et al., 2002). Following from

previous work, we used clusters of orthologous groups of proteins (COGs) to define orthologs across species (Galperin et al., 2015). For all COG pairs, we then computed the strength of co-evolution (as relative entropy) either by synteny or co-occurrence (see Supplemental Experimental Procedures for details).

These analyses reveal a sparse pattern of evolutionary coupling in which most folate metabolic genes are relatively independent from each other (Fig. 1B-C). Consistent with expectation, we observe coupling between physically interacting gene products: the glycine cleavage system proteins H, P and T (*gcvH*, *gcvP*, and *gcvT* in *E. coli*). Together with lipoamide dehydrogenase (*lpdA*), these enzymes form the glycine decarboxylase complex (GDC), a macromolecular complex that reversibly catalyzes either the degradation or biosynthesis of glycine (Okamura-Ikeda et al., 1993). Notably, lipoamide dehydrogenase also functions as part of both the 2-oxoglutarate dehydrogenase and pyruvate dehydrogenase multienzyme complexes in addition to the GDC (Carothers et al., 1989). The more general functional role of this protein in multiple complexes may explain the weaker co-evolution of *lpdA* with *gcvH*, *gcvP*, and *gcvT*.

Interestingly, we also see evolutionary coupling of enzyme pairs with no evidence for physical interaction: 1) DHFR/TYMS and 2) methionine synthase (MS) and methionine tetrahydrofolate reductase (MTHFR). DHFR and TYMS comprise the most strongly coupled gene pair in the folate cycle, and are de-coupled at the genome scale (Fig. 1). Both pairs of enzymes catalyze consecutive reactions in folate metabolism, suggesting that a biochemical mechanism might underlie their functional coupling. Yet, we note that most enzymes that catalyze neighboring reactions do not show statistical correlation (Fig. 1). Instead, co-evolution provides a different representation of folate metabolism in which most genes are independent and a few interact to form evolutionarily constrained units. Due to the limited sampling of

genomes, some false negatives are expected - that is, gene pairs which are functionally coupled but for which we do not see a strong co-evolutionary signal. Thus, the capacity for evolutionary statistics to predict adaptive units within a given organism requires thorough experimental testing.

To accomplish this, we focused on the DHFR and TYMS enzyme pair in *E. coli*. These proteins are the most strongly co-evolving in the pathway and are not known to physically interact, providing an opportunity to better understand how biochemical constraints might drive co-evolution. Though the genes encoding DHFR and TYMS are proximal along the chromosome of many bacterial species, they are approximately 2.9 megabases apart in *E. coli*. So, our experiments will determine whether coupling between genes found in synteny persists even when they are distant on the chromosome.

### **Coupling between DHFR and TYMS depends on enzyme activity**

Does the co-evolution of DHFR and TYMS correspond to epistatic coupling in the folate metabolic network? To address this, we conducted quantitative measurements of genetic epistasis for a library of ten DHFR mutants in the background of either WT TYMS or TYMS R166Q, a catalytically inactive variant (twenty constructs in total). Here, we quantify epistasis as non-additivity in the mutational effect on fitness for pairs of mutations in DHFR and TYMS (Fig 2A). The ten DHFR mutants were selected to span a range of catalytic activities; in prior work we measured the catalytic parameters ( $k_{cat}$  and  $K_m$ ) of these mutants *in vitro* (Reynolds et al., 2011). We used a previously validated next-generation sequencing based assay to measure the relative fitness of all possible mutant combinations (20 total) in a single internally controlled experiment (Reynolds et al., 2011). In this system, DHFR and TYMS are expressed from a

single plasmid that contains two DNA barcodes – one associated with DHFR and one with TYMS – that uniquely encode the identity of each mutant (Fig. 2B). The full library of mutants is transformed into the auxotroph strain *E. coli* ER2566  $\Delta folA \Delta thyA$ , and grown as a mixed population in a turbidostat. The turbidostat allows us to maintain the cell population at a fixed density in exponential phase for the duration of the experiment, with excellent control over media conditions. We then sampled time points over a twelve-hour period, and used next-generation sequencing to compute allele frequencies at each time point (Fig. 2C). By fitting a slope to the plot of allele frequencies versus time, we obtain a relative growth rate for each mutant in the population (Fig. 2C and Fig. S1). While many epistasis measurements rely on the extreme perturbation of total gene knock out (Babu et al., 2011b; Typas et al., 2008), this approach permits a more complete examination of how coupling varies with the magnitude of the functional perturbation.

The TYMS R166Q mutant is non-viable in the absence of exogenous thymidine. Epistasis was thus measured across two supplementation conditions: one in which TYMS R166Q produces a growth defect (5  $\mu$ g/ml thymidine), and one that constitutes a full rescue of TYMS activity (50  $\mu$ g/ml). Analysis of the DHFR mutants in the background of WT TYMS (grey points, Fig. 2D-E) shows that growth rate is relatively insensitive to ~10-fold decrease in DHFR activity. Tolerance to a small reduction of DHFR activity is preserved in a TYMS R166Q background, resulting in near additive epistasis. As DHFR activity is diminished further, the relationship between catalytic power and growth rate becomes monotonic: decreases in activity result in slower growth. In addition, a pattern of positive, or buffering epistasis emerges (upper bar plots in Fig. 2D-E); TYMS R166Q has the counter-intuitive consequence of partly or even fully (at 50  $\mu$ g/ml thymidine) restoring growth rate for low-activity DHFR mutants (Fig. 2D-E).



The presence of epistasis – in which loss of function in TYMS buffers decreases in the catalytic activity of DHFR – is consistent with the evolutionary coupling of this enzyme pair.

### **Mechanism of DHFR/TYMS coupling**

Given no evidence for the physical association of DHFR and TYMS in bacteria, what is the mechanism underlying their coupling? The finding that epistasis between the two genes depends on enzyme activity suggests a simple hypothesis: their coupling arises from the need to balance the concentration of key metabolites in the folate metabolic pathway. Support for this idea comes from prior work showing that treatment of *E. coli* with the DHFR inhibitor trimethoprim results in intracellular accumulation of DHF, which inhibits the upstream enzyme folylpoly- $\gamma$ -glutamate synthetase (FP- $\gamma$ -GS) (Kwon et al., 2008). FP- $\gamma$ -GS catalyzes the polyglutamylation of reduced folates, an important modification that increases folate retention in the cell and promotes the use of reduced folates as substrates in a number of downstream reactions (McGuire and Bertino, 1981). Thus, DHF accumulation results in off-target enzyme inhibition and cellular toxicity, a likely explanation for the growth rate defect observed in hypomorphic DHFR alleles (Fig. 2D-E). Because DHF is a product of TYMS, it is logical that loss-of-function mutations in TYMS might rescue growth in DHFR hypomorphs by preventing the accumulation of DHF.

To test this hypothesis, we carried out liquid chromatography-mass spectrometry (LC-MS) profiling of folate pathway metabolites in DHFR/TYMS mutant combinations. Specifically, we selected five DHFR variants that span a range of catalytic activities (WT, G121V, F31Y.L54I, M42F.G121V, and F31Y.G121V), and measured the relative abundance of intracellular folates in the background of either wild-type or R166Q TYMS. The experiment was

carried out for log-phase cultures in M9 glucose media supplemented with 0.1% ampicillin and 50  $\mu\text{g/ml}$  thymidine, conditions in which the selected DHFR mutations display significant growth defects individually, but in which the corresponding DHFR/TYMS double mutants are restored to near wild-type growth. Current mass spectrometry methods allow discrimination between the full diversity of folate species, which differ in oxidation, one-carbon modification, and polyglutamylation states, permitting a broad metabolic study of the effects of mutations (Lu et al., 2007).

The data confirm that for DHFR loss-of-function mutants, intracellular DHF concentration increases (Fig. 3A, bottom four rows). In addition, we find evidence for a depletion of reduced polyglutamated folates ( $\text{Glu} \geq 3$ ), while several mono- and di-glutamated THF species accumulate (particularly for THF, Methylene THF and 5-Methyl THF). This pattern of changes in the reduced folate pool is consistent with inhibition of  $\text{FP-}\gamma\text{-GS}$  by DHF accumulation (Fig. 3A, Fig. S2). It is also consistent with the observed growth rate defects in the DHFR loss-of-function mutants (Fig. 3B). How does the metabolite profile look in the background of the corresponding TYMS loss-of-function mutant? As predicted, we find clear evidence that the metabolite profile is corrected in the background of TYMS R166Q. Indeed, the concentrations of the reduced polyglutamated folates are restored to near-wild-type levels for most of the DHFR alleles (Fig 3A, first four rows below WT/R166Q). These data show that coordinated decreases in the activity of DHFR and TYMS maintain balance in key intracellular metabolites, a condition associated with optimal growth. Thus the coupling of DHFR and TYMS can be explained by a joint constraint on their catalytic activities, a biochemical mechanism that may have driven co-evolution of the DHFR/TYMS gene pair.

# **Forward evolution reveals adaptive independence of DHFR and TYMS from the rest of the genome**

The analysis of co-evolution presented in figure 1 goes beyond the prediction of epistatic coupling between DHFR and TYMS. The finding that this enzyme pair co-evolves relatively weakly with the remainder of the pathway suggests that they might represent an independent adaptive unit within folate metabolism. To test this, we conducted a global suppressor screen in which we inhibited DHFR with the common antibiotic trimethoprim and then examined the pattern of compensatory mutations with whole genome sequencing. If DHFR and TYMS indeed represent a quasi-independent adaptive unit, then suppressor mutations should be found within these two genes (*E. coli folA* and *thyA*) with minimal contribution from other sites.

We used a morbidostat, a specialized device for continuous culture, to facilitate forward evolution in the presence of trimethoprim (Toprak et al., 2012; Toprak et al., 2013) (Fig. 4A-C). The morbidostat dynamically adjusts the trimethoprim concentration in response to bacterial growth rate and total optical density, thereby providing steady selective pressure as resistance levels increase (see Fig. 4 legend for details). The basic principle is that cells undergo regular dilutions with fresh media until they hit a target optical density (OD = 0.15); once this density is reached, they are diluted with media containing trimethoprim until growth rate is decreased. This approach makes it possible to obtain long trajectories of adaptive mutations in the genome with sustained phenotypic adaptation (Toprak et al., 2012). For example, in a single 13-day experiment, we observe resistance levels in our evolving bacterial populations that approach the trimethoprim solubility limit in minimal (M9) media.

TYMS loss-of-function mutations have been observed in trimethoprim resistant clinical isolates for multiple genera of bacteria (King et al., 1983; Kriegeskorte et al., 2014). To emulate

the natural environment in which TYMS function is dispensable, we conducted evolution in three media conditions sampling various concentrations of exogenous thymidine (5, 10, and 50  $\mu\text{g/ml}$ ) that culminate in a full rescue of TYMS activity. All conditions were also supplemented with 0.2% ampicillin, a source of free amino acids. By alleviating selective pressure on the entire pathway, we aim to expose a larger range of adaptive mutations without biasing the pathway towards a particular result; this is similar to the common practice of conducting second site suppressor screens for essential genes under relatively permissive conditions (Forsburg, 2001). In this context, an absence of mutation outside of the two-gene pair becomes more significant.

Over the 13 days of evolution, we estimated the trimethoprim resistance of each evolving population by computing the median drug concentration in the culture vial beginning with the first dilution with drug to the end of the day (Fig 4D and Fig. S3). To identify the mutants causally related to trimethoprim resistance, we selected 10 single colonies from the endpoint of each of the three experimental conditions for phenotypic and genotypic characterization (30 strains in total, Fig. 5, Fig. S4 and Table S2). For each strain, we measured the trimethoprim  $\text{IC}_{50}$ , growth rate dependence on thymidine, and conducted whole genome sequencing. All 30 strains were confirmed to be highly trimethoprim resistant (Fig. 5A, Table S2). We were unable to measure  $\text{IC}_{50}$  values for strains 4, 5 and 10 from the 50  $\mu\text{g/ml}$  thymidine condition: these three strains grew very slowly but were completely insensitive to trimethoprim. Further, strains from all three thymidine supplemented conditions now depend on exogenous thymidine for growth, indicating a loss of function in the *thyA* gene that encodes TYMS (Fig. 5B and Fig. S4). We confirmed that this loss of function is not a simple consequence of neutral genetic variation in the presence of thymidine; cells grown in 50  $\mu\text{g/ml}$  thymidine in the absence of trimethoprim retain TYMS function over similar time scales (Fig. S5A-B).

Whole genome sequencing for all 30 strains reveals a striking pattern of mutation (Fig. 5C and Tables S3-S4). Isolates from all three conditions acquire coding-region mutations in both DHFR and TYMS, or even just in TYMS. In particular, strains 4, 5, 7, and 10 in the 50  $\mu$ g/ml thymidine condition contain mutations in TYMS but not DHFR – showing that one route to resistance is the acquisition of mutations in a gene not directly targeted by antibiotic. All mutations identified in DHFR reproduce those observed in the earlier morbidostat study of trimethoprim resistance (Toprak et al., 2012). The mutations in TYMS – two insertion sequence elements, a frame shift mutation, loss of two codons, and a non-synonymous active site mutation - are consistent with loss of function. Thus, the mutations in DHFR and TYMS are compatible with the proposed mechanism of coupling: reduced TYMS activity can buffer inhibition of DHFR.

In support of the adaptive independence of the DHFR/TYMS pair, we observe no other mutations in folate metabolism genes (Fig. 5C and Table S4). More generally, few other mutations occur elsewhere in the genome, and the majority of these are not systematically observed across clones. This suggests that they may be spurious variations not associated with the adaptive phenotype. One of the evolved strains contains only mutations in DHFR and TYMS (strain 1 in 50  $\mu$ g/ml thymidine, Fig. 5C), indicating that variation in the DHFR/TYMS genes is sufficient to produce resistance. To verify this, we introduced several of the observed DHFR and TYMS mutations into a clean wild-type *E. coli* MG1655 background and measured the IC<sub>50</sub>. These data show that the DHFR/TYMS mutations are sufficient to reproduce the resistance phenotype measured for the evolved strains (Fig. S5). DHFR and TYMS show a capacity for adaptation through compensatory mutation that is contained within the two-gene unit.

## Genome-wide analyses of co-evolution identifies additional adaptive pairs

Our data for the DHFR/TYMS pair provide confidence that co-evolution might reasonably detect adaptive units - small groups of genes that collectively evolve with each other but adapt more independently from the remainder of the cell. To examine the prevalence of other pairs like DHFR/TYMS genome wide, we extended our analyses of gene synteny and co-occurrence to include all genes represented in *E. coli*. To ensure good statistics, we filtered the orthologs analyzed to those that co-occur in a sufficiently large number of genomes (2095 COGs, ~500,000 pairs in total) (see also Supplemental Experimental Procedures). We compared our analysis to: (1) metabolic annotations from KEGG (Kanehisa et al., 2012) and (2) the set of high-confidence binding interactions in *E. coli* reported by the STRING database (Szklarczyk et al., 2015). Consistent with intuition and prior work, evolutionarily coupled gene pairs show enrichment for physical complexes, enzymes in the same metabolic pathway, and more specifically, enzymes with a shared metabolite (Fig. 6A,B).

Overall, these data provide encouraging validation of the ability of co-evolutionary methods to detect interactions. But to investigate the possibility of adaptive units like DHFR/TYMS, we must identify groups of genes that are strongly coupled to each other, and comparatively decoupled from the remainder of the genome. Here, we focus on the simplest case of a modular gene pair. Accordingly, in Figure 6C-D, we show scatterplots of all gene pairs, indicating the strength of coupling within each pair (as a relative entropy, along the x-axis) versus the strongest coupling outside of the pair (along the y-axis). In this representation, gene pairs below the diagonal are more tightly coupled to each other than any other gene in the dataset. Thus, we predict that gene pairs in the lower right corner may act as adaptive units. This simple plot is limited to the identification of two gene adaptive units, and our analysis will need

to be extended to identify communities within the complete weighted network of co-evolutionary relationships (Newman, 2010). However, this analysis clearly indicates the presence of other pairs like DHFR and TYMS, and suggests interesting candidates for initial experimental testing.

For illustrative purposes, we consider a simple definition for the prediction of adaptive units based on empirical cutoffs for internal and external coupling ( $D_{ij}^{\text{intra}} > 1.0$  and  $D_{ij}^{\text{exter}} < 0.5$ , Table S1). In this regime of the synteny analysis, we identify the gene pair *accB/accC* which encodes two of the four subunits of acetyl-CoA carboxylase, the first enzymatic step in fatty acid biosynthesis. Overexpression of either *accB* or *accC* individually causes reductions in fatty acid biosynthesis, but overexpressing the two genes in stoichiometric amounts rescues this defect (Abdel-Hamid and Cronan, 2007; Janssen and Steinbuchel, 2014). Constraints on relative expression have also been noted for the *selA/selB* and *tatB/tatC* gene pairs (Bolhuis et al., 2001; Rengby et al., 2004). The *tatB/tatC* genes encode components of the TatABCE twin-arginine translocation complex, while *selA/selB* are not known to bind but are both involved in selenoprotein biosynthesis. Thus, these co-evolving units may generally indicate groups of genes whose activity or expression is constrained relative to each other. This interpretation is consistent with previous work that has shown that synteny identifies stretches of genes larger than single operons that are functionally co-expressed (Junier and Rivoire, 2016). In this sense, the bifunctional fused form of DHFR/TYMS found in protozoa and plants could be regarded as an extreme case that guarantees stoichiometric expression (Beverley et al., 1986; Lazar et al., 1993). Taken together, these data support the idea that co-evolution analyses can be used to identify small adaptive units within cellular systems more generally.

## Discussion and Conclusions

### DHFR and TYMS represent an adaptive unit in folate metabolism

Our data indicate that DHFR and TYMS are epistatically coupled to each other through a constraint on relative activity mediated by metabolite concentration. Namely, prior work has shown that inhibition of DHFR with TMP leads to rapid accumulation of DHF and inhibition of the upstream enzyme FP- $\gamma$ -GS (Kwon et al., 2008). Using DHFR mutants of known catalytic activity, we show that reducing TYMS activity alleviates DHF accumulation and restores growth rate. This result is recapitulated in our forward evolution experiments, which show that inhibition of DHFR with TMP promotes TYMS loss of function mutations. These results provide a mechanistic explanation for the well-known clinical result that trimethoprim resistance is often accompanied by loss of function mutations in TYMS (King et al., 1983). Further, mutations within the two enzymes are sufficient to accomplish TMP resistance, without a need for compensatory mutations elsewhere in the pathway or genome. Our comparative genomics analyses show that this functional coupling and relative adaptive independence is reflected in the statistics of both synteny and gene co-occurrence across thousands of bacterial species: DHFR and TYMS strongly co-evolve with one another, but less so with the remainder of the genome.

These findings are consistent with several pieces of prior work. The first comes from examining TYMS ortholog variation across species. For instance, the bacterium *R. capsulatus* lacks both *thyA* (TYMS) or *folA* (DHFR) homologs, and instead produces thymidine via *thyX*, a thymidylate synthase that generates THF (rather than DHF) in the process of thymidine production. When *thyX* is deleted from *R. capsulatus*, growth can only be complemented by the addition of both *thyA* and *folA* from *R. sphaeroides*; the *thyA* gene alone is insufficient. Computational simulation shows that in the absence of a high-activity DHFR (*folA*), *thyA* rapidly



depletes reduced (THF) folate pools by converting them to DHF (Leduc et al., 2007). These data thus also suggest a fitness constraint on the relative activities of *folA* and *thyA*.

Further, mathematical models of the eukaryotic folate cycle indicate that standard biochemical kinetics might result in modularity of the DHFR/TYMS pair (Leduc et al., 2007; Nijhout et al., 2004). In eukaryotic cells, thymidine synthesis is the rate-limiting step for DNA synthesis, and transcription of the TYMS and DHFR genes is greatly upregulated (via a common transcription factor) at the G<sub>1</sub>/S cell cycle transition (Bjarnason et al., 2001). Computationally increasing the activities of DHFR and TYMS 100-fold in the context of a model based on known biochemical parameters results in increased thymidine synthesis but only modestly changes the concentration of other folates, indicating that biochemical rates and metabolic network structure give rise to modularity in different metabolite pools. This suggests that decoupling of the DHFR/TYMS pair from the remainder of folate metabolism might provide one route for metabolic homeostasis (Nijhout et al., 2004).

Taken together, our data indicate that DHFR/TYMS co-evolution is driven by a shared metabolic constraint. This is broadly consistent with the idea that co-evolution is not limited to detecting physical interactions but instead reflects the coupling of gene activities regardless of mechanism (Huynen et al., 2000; Snel et al., 2002). A toxic metabolic intermediate represents one case in which protein stoichiometry is important, but in principle any functional constraint on relative expression levels or biochemical flux might lead to similar evolutionary signatures. Recent work examining the potential small molecule regulatory network of metabolism in both *E. coli* and humans indicates many metabolic intermediates may have some inhibitory and/or regulatory capacity (Alam et al., 2017; Reznik et al., 2017). This suggests that constraints on intermediate formation could be a common driver of co-evolution between metabolic enzymes.

## Comparative genomics as a strategy for identifying adaptive units in cellular systems

Co-evolution analyses are often evaluated by their ability to recapitulate known interactions or predict new ones (Huynen et al., 2000; Janga et al., 2005; Kim and Price, 2011; Pellegrini et al., 1999; Szklarczyk et al., 2015). For example, in Figure 6 we show that our two measures of evolutionary coupling successfully enrich for physical complexes and proteins in shared metabolic pathways. But in analogy to recent work in proteins (Halabi et al., 2009) and gene expression (Junier and Rivoire, 2016), these methods might be further developed to identify quasi-independent adaptive units within cellular systems (Li et al., 2014; Reynolds, 2014). Our case study for the DHFR/TYMS pair provides the first experimental test that signatures of co-evolution might indeed be reasonably used to identify independent adaptive units at the systems scale. As in Figure 6, the pattern of (evolutionary) interactions can be examined to look for groups of two or more genes that are statistically coupled to one another but more decoupled from the rest of the genome with the hypothesis that these co-evolving units might represent cooperative units of function and adaptation within extant species.

Experimentally testing the functional meaning of these units requires new methods to not only evaluate interactions, but the independence of a group of genes from the rest of the genome. The global suppressor screen shown here establishes one general experimental criterion for establishing adaptive independence. The logic behind this approach is two-fold: 1) we are directly examining if past evolutionary statistics across prokaryotes relate to laboratory-timescale forward evolution in *E. coli* and 2) it permits screening for suppressor mutants genome wide. However, other experimental strategies for evaluating independence are possible and should be considered. One obvious criterion is modularity in genetic interactions: how do pairwise epistasis measurements compare to statistically identified adaptive units? At present, genome-wide

measurements of epistasis have not yet been made for *E. coli*. However, high-throughput approaches for generating double mutants in different genes (e.g. eSGA and GIANT coli) have yielded large data sets of pairwise epistasis for several pathways including cell wall biosynthesis, translation, and the DNA damage response (Babu et al., 2014; Babu et al., 2011a; Babu et al., 2011b; Butland et al., 2008; Gagarinova et al., 2016; Kumar et al., 2016; Typas et al., 2008). These datasets provide an interesting basis for further study, but face several limitations: 1) epistasis is often indirectly inferred – single mutant growth rates are not always explicitly measured and 2) fitness effects are often measured in the extreme limit of complete gene knockout. CRISPR-interference (CRISPR-i) overcomes many of these technical difficulties and is an interesting future avenue for introducing genetic knockdowns of varying magnitude in a high-throughput way (Bikard et al., 2013; Qi et al., 2013). The results presented here provide a starting place for now conducting these and other experiments comprehensively.

We detect co-evolution between DHFR and TYMS by both synteny and co-occurrence, but can coupling through a shared metabolite result in detectable amino acid sequence co-variation? Direct Coupling Analysis (DCA) and Statistical Coupling Analysis (SCA) represent two distinct approaches for the analysis of amino acid sequence co-evolution. The former was developed with the goal of detecting local, physical contacts between amino acids (Hopf et al., 2014; Morcos et al., 2011; Ovchinnikov et al., 2014; Skerker et al., 2008), while the latter has identified co-evolving residue networks associated with function and allostery (McLaughlin et al., 2012; Raman et al., 2016; Reynolds et al., 2011; Suel et al., 2003). DCA has proven powerful in predicting physical protein interactions and specific contacts in macromolecular complexes (Feinauer et al., 2016; Hopf et al., 2014; Ovchinnikov et al., 2014). SCA has not yet been applied to inter-protein interactions but has been used to study co-evolution and allostery between

domains (Lee et al., 2008; Reynolds et al., 2011; Smock et al., 2010). We used two web-server implementations of DCA to examine the DHFR/TYMS pair. One implementation (EVcomplex, (Hopf et al., 2014)) predicts no high-scoring contacts between the two proteins (EVcomplex score  $\geq 0.8$ ), while the other (GREMLIN, (Ovchinnikov et al., 2014)) predicts that the two proteins may interact, on the basis of a single high-scoring inter-protein residue pair. In either case, the finding of only one or zero couplings between DHFR and TYMS is consistent with the interpretation of DCA couplings as physical contacts because these enzymes are not known to physically interact in bacteria. In contrast, application of SCA to the DHFR/TYMS pair reveals a number of co-evolving positions, but also detects co-evolution between DHFR/TYMS and several less obviously related proteins (Fig. S6). Thus, more work is needed to understand the significance of the inter-protein SCA correlations. More generally, these results point to a clear need to develop appropriate statistical tools for examining co-evolution between functionally coupled proteins. The DHFR/TYMS pair can provide an experimentally powerful test case to carefully study the amino acid sequence constraints imposed between non-binding but epistatically coupled proteins.

Finally, our comparative genomics analysis has the capacity to identify co-evolving units distinct from expectations based on metabolic proximity or physical interaction maps. In many cases, the co-evolution observed between genes can be rationalized by involvement in a shared complex or metabolic pathway. However, the finding of independence is less intuitive – for example, DHFR and TYMS co-evolve with one another but not other enzymes in the pathway with a shared metabolite. We also observe the *accB* and *accC* gene pair as an independent co-evolving unit even though they are only two of the four proteins involved in the acetyl-CoA carboxylase physical complex. As mentioned earlier, experiments suggest the expression level

of the *accB/C* gene pair are constrained relative to each other, but seemingly less so with the remainder of the complex (Abdel-Hamid and Cronan, 2007). Thus, co-evolutionary analyses have the potential to reveal new units of adaptation and function within larger cellular systems and complexes. Though comprehensive testing remains a matter for extensive future work, this inspires the possibility that metabolism might be decomposed into small multi-gene units using evolutionary information. If so, these new units would assist in focusing mechanistic work, suggest strategies for the engineering of cellular systems, and provide a path towards predictive modeling of cellular phenotypes.

**Conflict of Interest** The authors declare they have no conflict of interest.

**Acknowledgements** We thank members of the Reynolds lab for review of the manuscript, E. Toprak for extensive advice on morbidostat construction and operation, S. Benkovic for the ER2566 *ΔfolA ΔthyA* strain, T. Kuhlman for molecular biology reagents used in genome editing, T. Bergmiller for the GFP/Chloramphenicol resistance marker incorporated into our founder strain, and R. Ranganathan for discussions. This research was funded in part by the Gordon and Betty Moore Foundation's Data-Driven Discovery Initiative through Grant GBMF4557 to K.R. and by the Green Center for Systems Biology at UT Southwestern Medical Center. A.S. was supported in part by NIH training grant 5T32GM8203-28. I.J. is supported by an ATIP-Avenir grant (Centre National de la Recherche Scientifique).

## Figure Legends

**Figure 1. Biochemical and statistical representations of folate metabolism.** A, Biochemical pathway map of folate metabolism. Black text indicates enzymes, light blue text indicates

metabolic intermediates. See Table S1 for a more complete description of each enzyme. **B-C**, Heatmaps of evolutionary coupling between gene pairs in folate metabolism as evaluated by gene synteny and co-occurrence. Pixel color shows the extent of coupling between genes, indicated as a relative entropy,  $D_{ij}^{intra}$ . Enzyme names are indicated on the left and top of the matrix in black text, the corresponding gene names are given at right in grey italics. In *E. coli*, a single gene (*folD*) encodes a bifunctional enzyme that catalyzes both the methylene tetrahydrofolate dehydrogenase (MTD) and methenyltetrahydrofolate cyclohydrolase (MTCH) reactions in the biochemical pathway as shown in A. The majority of gene pairs show little co-evolution by either measure (dark purple pixels).

**Figure 2 Epistatic coupling between DHFR and TYMS.** **A**, Schematic for computing pairwise epistasis using an additive model. Each corner of the square denotes a different genotype, where genes A and B take either a wild-type (capitalized) or mutant (lowercase) state. The symbol  $\Delta m$  represents a change in fitness (in our case relative growth) between states, and  $\Delta\Delta m_{ab}$  is epistasis. **B**, Genetic barcoding scheme for deep sequencing. Each plasmid contains two barcodes uniquely encoding the identity of *folA* and *thyA* genes. Sequencing of both barcodes enables determination of relative allele frequencies within a population as they vary with time and experimental condition. **C**, Relative allele frequency versus time for a growth competition assay carried out in 50  $\mu\text{g/ml}$  thymidine. The relative fitness of each allele pair is given by the linear slope  $m$ . See Fig. S1 for all growth rate fits. **D, E**, Scatter plots of relative growth rate for DHFR mutants spanning a range of catalytic specificities ( $k_{cat}/K_m$ ), and either a wild-type (WT, grey points) or catalytically dead (R166Q, red points) TYMS. Error bars correspond to standard error across triplicate measurements. Epistasis, shown as a bar chart above, is calculated using

the additive model in **A**. For low activity DHFR, we observe positive or “buffering” epistasis in both 5 and 50 µg/ml thymidine. In this case, the cost of reducing the activity of one enzyme is partly or totally mitigated by reducing activity in the other.

**Figure 3 A loss-of-function mutation in TYMS buffers metabolic changes from decreased DHFR activity.** **A**, Liquid chromatography-mass spectrometry profiling of intracellular folate species in M9 media supplemented with 0.1% ampicillin and 50 µg/ml thymidine. Rows reflect mutant DHFR/TYMS combinations, columns correspond to metabolites. Each folate species can be modified by the addition of 1-5 glutamates. The color of each square denotes the log<sub>2</sub> abundance of each species relative to wild-type. The data show that mutations reducing DHFR activity (G121V, F31Y.L54I, M42F.G121V, and F31Y.G121V) cause an accumulation of DHF and depletion of reduced folate species (THF) (bottom four rows). This effect is partly compensated by an inactivating mutation in TYMS (rows two-five). **B**, The corresponding doubling time for each mutant, as measured in batch culture (conditions identical to panel A). See also Fig. S2.

**Figure 4 Evolution of trimethoprim (TMP) resistance in *MG1655* cells using the morbidostat.** **A**, Schematic of the continuous culture tube. Dilutions are made through the inlet tubes labeled “media A,” “media B,” and “media C.” A constant volume of 15 ml is maintained by the outlet line, labeled “waste,” which aspirates extra medium after mixing. The optical density at 600 nm (OD<sub>600</sub>) of each culture is monitored by an LED-detector pair near the bottom of the tube. Drug concentration is dynamically varied to promote evolution of increased resistance. **B**, Control strategy for the addition of trimethoprim. Once the OD<sub>600</sub> exceeds 0.06,

dilutions of 3 ml were made every 20 minutes. For  $OD_{600}=0.06-0.15$ , “media A” is added, which contains no TMP. Above an  $OD_{600}$  of 0.15, drug was introduced through dilution with “media B,” which contains a lower amount of TMP. Once the TMP concentration in the culture tube exceeds 60% of the “media B” stock, then “media C,” which contains 5X more TMP, is used. Following a decrease in growth rate in response to drug, dilutions resume with “media A”. If “media C” was used in a particular day, then the TMP concentration in media “B & C” were incremented by a factor of 5X to enable further adaptation in the following day. **C**, A representative growth trajectory, color-coded by TMP concentration (day 7, 50  $\mu\text{g/ml}$  thymidine). See Figure S3 for full growth trajectories over 13 days. **D**, The trajectory of estimated TMP resistance (as measured by the median TMP concentration on each day) versus number of generations for each experimental condition.

**Figure 5 Measurements of phenotype and genotype.** Ten single colonies (strains) were selected at the endpoint of each forward evolution condition for phenotyping and genotyping (30 in total). **A**, Trimethoprim (TMP)  $IC_{50}$  measurements for each experimental strain. Error bars represent standard error over triplicate measurements. See also Table S2. **B**, Thymidine dependence for each experimental strain, as determined by area under the  $\log(OD_{600})$  curve in 0  $\mu\text{g/ml}$  thymidine over 10 hours. See Fig. S4 for growth rates across a range of thymidine concentrations. Evolved strains are no longer viable in the absence of extracellular thymidine, indicating a loss-of-function mutation in TYMS. **C**, Mutations acquired by each strain during forward evolution. Genes that were mutated in two or fewer strains across all conditions are excluded, as are synonymous mutations (see Table S3 for sequencing statistics, and Table S4 for a complete list of mutants). Gene names are labeled along the left edge of the map, with the



corresponding residue or nucleotide change(s) denoted along the right. If a strain acquires any mutation in a particular gene, the column section corresponding to that gene is shaded blue. Nearly all strains acquired mutations in both *folA* and *thyA*, encoding DHFR and TYMS, with a few exceptions lacking a *folA* mutation. A small red arrow indicates one strain with mutations in only DHFR and TYMS. See Fig. S5 for measurements of the trimethoprim resistance of four strains engineered to include only the *folA*/*thyA* mutations in a clean genetic background.

**Figure 6. Genome-wide analysis of co-evolution in *E. coli*.** **A**, Enrichment of physical and metabolic interactions as a function of synteny coupling. **B**, Enrichment of physical and metabolic interactions as a function of co-occurrence coupling. **C**, A scatterplot of synteny-based coupling for all analyzed gene pairs. Each point represents a pair of orthologous genes (COGs); coupling within the pair is shown on the x-axis, and the strongest coupling outside of the pair is shown on the y-axis. Color-coding reflects annotations from the STRING database (physical interactions) or KEGG database (metabolic pathways): green gene pairs bind, while pairs in dark blue or light blue do not interact but are found in the same metabolic pathway. Dark blue gene pairs share a metabolic intermediate. The DHFR/TYMS pair is highlighted in red. See Table S5 for an annotated list of gene pairs below the diagonal. **D**, Scatterplot of coupling by co-occurrence for all analyzed gene pairs (same format as C).

## Materials and Methods

**Comparative genomics analyses.** Synteny analysis was conducted using a simplified version of the methods described in (Junier and Rivoire, 2013, 2016). See the Supplemental Experimental Procedures for a detailed description of the synteny and co-occurrence calculations.

# **Forward evolution of trimethoprim resistance in the morbidostat.**

The morbidostat/turbidostat apparatus was constructed as described by Toprak and colleagues (Toprak et al., 2013). The founder strain for the forward evolution experiment was *E. coli* MG1655 modified by phage transduction to encode green fluorescent protein (*egfp*) and chloramphenicol resistance (*cat*) at the P21 attachment site. The goal of this modification was to prevent and detect contamination with other strains. Throughout the forward evolution experiment, cells were grown at 30°C in M9 media supplemented with 0.4% glucose and 0.2% ampicase (Sigma); 30 µg/ml of chloramphenicol (Cam) was added for positive selection.

To begin the experiment, the founder strain was cultured overnight at 37°C in Luria Broth (LB) + 30 µg/ml Cam. This culture was washed twice with M9, and back diluted into M9 + 30 µg/ml Cam supplemented with 5, 10, or 50 µg/ml thymidine (thy) for overnight adaptation in culture tubes at 30°C. The next day (henceforth referred to as day 0; day 1 is the end of the first day of adaptation), these overnight cultures were streaked onto LB agar plates: two colonies per condition were chosen for whole genome sequencing (WGS) in order to obtain an accurate sequence for the founder strain. The remainder of the overnight cultures was used to inoculate four morbidostat tubes at containing M9 media with varying thymidine supplementation (5, 10, and 50µg/ml thy). The starting optical density was approximately 0.005. Initial antibiotic concentrations were 0, 11.5 and 57.5 µg/ml trimethoprim for media stocks A, B, and C respectively. Each culture grew unperturbed until it surpassed an OD<sub>600</sub> of 0.06, at which point it underwent periodic dilutions with fresh media. The dilution rate is given by the formula  $r_{dil} = f \ln \frac{V}{V+\Delta V}$ , where  $V = 15ml$  is the culture volume, and  $\Delta V = 3ml$  is volume added. We chose a dilution frequency  $f = 3 \text{ h}^{-1}$ , to give  $r_{dil} = 0.55$ . Above the OD<sub>600</sub>=0.15, these dilutions are used to introduce TMP into the culture (see also Fig. 5B). This allows controlled inhibition of

DHFR activity in response to growth rate. Cycles of growth and dilution continued for a period of ~22 hours, at which point the run was paused to make glycerol stocks, replenish media, and update TMP stock concentrations. Culture vials for the next day of evolution were filled with fresh media and inoculated using 300µl from the previous culture. Complete trajectories of OD<sub>600</sub> and drug concentration are shown in Fig. S4. Endpoint cultures were streaked onto LB agar plates supplemented with 30 µg/ml of Cam and 50 µg/ml thymidine to obtain isolated colonies for whole genome sequencing.

**Whole genome sequencing.** Two isolates were selected from each adapted day 0 culture, and ten clonal isolates (colonies) were randomly selected from the endpoint of each evolution condition, totaling 36 strains. Isolation of genomic DNA was performed using the QIAamp DNA Mini Kit (Qiagen). The Nextera XT DNA Library Prep Kit (Illumina) was used to fragment and label each genome for paired-end sequencing using a v2 300-cycle MiSeq kit (Illumina). Average read length and coverage can be found in Table S3. Genome assembly and mutation prediction was performed using *breseq* (Deatherage and Barrick, 2014). The reference sequence was a modification of the *E. coli* MG1655 complete genome (accession no. NC\_000193), edited to include the GFP marker and chloramphenicol resistance cassette in our founder strain. The modified reference sequence and all complete genome sequences from the beginning and endpoint of forward evolution are available in the NCBI BioProject database (accession number: PRJNA378892).

**Measurements of thymidine dependence.** All strains were grown overnight in LB + 5µg/ml thy, with the exception of the strains evolved in the 50µg/ml thy, which were supplemented with

50 $\mu$ g/ml thy to ensure viability. Cultures were then washed twice in M9 media without thymidine, and inoculated at an OD<sub>600</sub>=0.005 in 96-well plates containing M9 media supplemented with 10-fold serial dilutions of thymidine, ranging from 0.005  $\mu$ g/ml to 50  $\mu$ g/ml (in singlicate). OD<sub>600</sub> was monitored in a Victor X3 plate reader at 30°C over a period of 20 hours. Growth was quantified using the positive integral of OD<sub>600</sub> over time. This measure captures mutational or drug-induced changes in the duration of lag phase as well as perturbations in growth rate (Toprak et al., 2012). For each strain, we identified a start-time ( $t_0$ ) at the end of lag-phase for the fully-rescued 50 $\mu$ g/ml thy condition. We chose each  $t_0$  computationally as the last point before monotonic growth above the limit of detection. The log(OD<sub>600</sub>) versus time curves for all conditions are then vertically shifted ('background-subtracted'), such that the function value at this start-time is zero. This curve is then numerically integrated from  $t_0$  to  $t_0$ +10 hours using the trapezoid method.

**Measurements of trimethoprim resistance (IC<sub>50</sub>).** All strains were grown overnight in LB + 5 $\mu$ g/ml thy, with the exception of the strains evolved in the 50 $\mu$ g/ml thy, which were supplemented with 50 $\mu$ g/ml thy to ensure viability. Each strain was then washed into media conditions corresponding to the strain's forward evolution condition, and adapted for 5.5 hours at 30°C. The recovery cultures were used to inoculate 96-well plates containing M9 media sampling serial dilutions of TMP (in triplicate), with a starting OD<sub>600</sub> = 0.005. OD<sub>600</sub> was monitored using a Tecan Infinite M200 Pro microplate reader and Freedom Evo robot at 30°C over a period of at least 12 hours. The trimethoprim resistance of each strain was quantified by its absolute IC<sub>50</sub>, the drug concentration ( $\mu$ g/ml) at which growth is half-maximal. The

relationship between growth and trimethoprim inhibition is modeled using the four parameter logistic function:

$$Y = \frac{a - d}{1 + (X/c)^b} + d$$

where  $Y$  is growth,  $X$  is TMP concentration,  $a$  is the asymptote for uninhibited growth,  $d$  is the limit for inhibited growth,  $c$  provides the concentration midway between  $a$  and  $d$ , and  $b$  captures sensitivity (Sebaugh, 2011). Growth was quantified using the positive integral of OD<sub>600</sub> data over a 10h period of growth (see also the methods for measurement of thymidine dependence). For each strain, we identify a start-time ( $t_0$ ) at the end of lag-phase for the uninhibited 0µg/ml TMP condition. Growth versus TMP concentration was fit to the above model using MATLAB. IC<sub>50</sub> was calculated as the concentration  $X^*$  for which growth  $Y(X^*) = a/2$ .

# **Growth without trimethoprim selection in 50µg/ml thymidine using the turbidostat**

The founder strain for this experiment was identical to that used for evolution of trimethoprim resistance. Throughout the experiment, cells were grown at 30°C in M9 media supplemented with 0.4% glucose and 0.2% ampicase (Sigma); 30 µg/ml of chloramphenicol (CAM) was added for positive selection. To begin the experiment, the founder strain was cultured overnight at 37°C in Luria Broth (LB) + 30 µg/ml Cam. This culture was washed twice with M9, and back diluted into M9 supplemented with 50µg/ml thymidine (thy) for overnight adaptation in culture tubes at 30°C. The next day (henceforth referred to as day 0; day 1 is the end of the first day of continuous culture), the overnight culture was used to inoculate three turbidostat tubes containing 17ml of M9 supplemented with 50 thy. The starting optical density was approximately 0.005. Each culture grew unperturbed until it reached an OD<sub>600</sub> of 0.15, at which point it was diluted with 2.4 ml of fresh media. These cycles of growth and dilution persisted for

a period of ~22 hours, at which point the run was paused to make glycerol stocks and replenish media. Culture vials for each following day of evolution were filled with fresh media and inoculated using 300µl from the previous culture.

**Epistasis Measurements.** All relative growth rate measurements were performed in the *E. coli* folate auxotroph strain ER2566  $\Delta folA \Delta thyA$  (Lee et al., 2008). DHFR (*folA*) and TYMS (*thyA*) are provided on the plasmid pACYC-Duet1 (in MCS1 and MCS2, respectively) and are each under control of a T7 promoter. For these experiments, we use leaky expression (no IPTG induction). Each mutant plasmid (20 in total) is marked with a genetic barcode in a non-coding region between the two genes. Plasmids were transformed into the auxotroph strain, and each mutant was grown overnight in separate LB +30µg/ml Cam +50µg/ml thy cultures. Then, cultures were washed 2x in M9 media supplemented with 0.4% glucose and 0.2% ampicase and 30µg/ml Cam, and adapted overnight at 30°C. All mutants were mixed in equal ratios based on OD<sub>600</sub> and inoculated at a starting OD<sub>600</sub> = 0.1 in the turbidostat. Growth rates were measured under two conditions: 5 thy and 50 thy, with three replicates each. The turbidostat clamps the culture to a fixed OD<sub>600</sub> = 0.15 by adding fresh dilutions of media. Every 2 hours over the course of 12 hours a 1ml sample was removed, pelleted and frozen for next-generation sequencing. Amplicons containing the barcoded region with appropriate sequencing adaptors (350 basepairs in total size) were generated by two sequential rounds of PCR with Q5 polymerase. The barcoded region was sequenced with a single-end MiSeq run using a v2 50 cycle kit (Illumina). We obtained 14,348,937 reads. Data analysis was performed using a series of custom python scripts to count barcodes, and MATLAB to fit relative growth rates.

**Constructing DHFR/TYMS mutants in a clean genetic background.** We followed the protocol for scarless genome integration using the modified  $\lambda$ -red system developed by Tas et al. (Tas et al., 2015). In this method, a tetracycline (Tet) resistance cassette (“landing pad”) is first integrated at the site targeted for mutagenesis. Then, the landing pad is excised by the endonuclease I-SceI, and replaced with the desired mutation by  $\lambda$ -red mediated recombination.  $\text{NiCl}_2$  is used to counterselect against cells that retain the tetracycline cassette. Tas et al. provides a detailed protocol; here we give the specifics necessary for our experiments. For the  $\lambda$ -red machinery, we transformed the plasmid pTKRED (Genbank accession number GU327533) into electrocompetent *E. coli* MG1655 with a genomic *egfp/cat* resistance cassette (the forward evolution founder strain). For the  $\Delta 25$ -26 TYMS mutation, we introduced the *tetA* landing pad between genome positions 2,964,900 and 2,965,201 (genome NC000913) corresponding to the N-terminus of the *thyA* gene. For the DHFR mutations (L28R, W30R, and P21L), the landing pad was recombined between genome positions 49,684 and 49,990 (genome NC000913). In order to replace the Tet cassette, cells were induced with 2mM IPTG and 0.4% arabinose, and then transformed with 100ng of dsDNA PCR product containing the mutation of interest (with appropriate homology arms). This reaction experienced 3 days of outgrowth at 30°C in rich defined media (RDM, Teknova) with glucose substituted for 0.5% v/v glycerol. The media was supplemented with 6 mM or 4mM  $\text{NiCl}_2$  for counterselection against *tetA* at the *thyA* locus or *folA* locus respectively. The outgrowth culture was streaked onto agar plates and screened daily for the mutant of interest using LB supplemented with 50  $\mu\text{g/ml}$  thy, 30  $\mu\text{g/ml}$  spectinomycin, and +/- 5-10  $\mu\text{g/ml}$  Tet. All mutations were confirmed by Sanger sequencing of the complete *folA* and *thyA* open reading frame; for *folA* the promoter region was also sequenced.

**LC-MS Metabolite Measurements.** Cells were cultured in M9 0.2% glucose media containing 0.1% ampicillin, 50 µg/ml thymine, and 30 µg/ml Cam at 30°C for metabolite analysis. In mid-log phase at OD<sub>600</sub> ~0.2, *E. coli* culture (3 ml for nucleotide measurement and 7 ml for folate measurement) was filtered on a nylon membrane (0.2 µm), and the residual medium was quickly washed away by filtering warm saline solution (200 mM NaCl at 30°C) over the membrane loaded with cells to exclude non-desirable extracellular metabolites from LC-MS analysis. The membrane was immediately transferred to a 6 cm Petri dish containing 1 ml cold extraction solvent (-20°C 40:40:20 methanol/acetonitrile/water; for folate stability, 2.5 mM sodium ascorbate and 25 mM ammonium acetate in folate extraction solvent (Lu et al., 2007)) to quench metabolism. After washing the membrane, the cell extract solution was transferred to a microcentrifuge tube and centrifuged at 13000 rcf for 10 min. The supernatant was transferred to a new microcentrifuge tube. Folate samples were prepared with an additional extraction: the pellet was resuspended in the cold extraction solvent and sonicated for 10 min in an ice bath. After the second extraction and centrifugation, the supernatant was combined with the initial supernatant. The metabolite extracts were dried under nitrogen flow and reconstituted in HPLC-grade water for LC-MS analysis. Metabolites were measured using stand-alone orbitrap mass spectrometers (ThermoFisher Exactive and Q-Exactive) operating in negative ion mode with reverse-phase liquid chromatography (Lu et al., 2010). Exactive chromatographic separation was achieved on a Synergy Hydro-RP column (100 mm×2 mm, 2.5 µm particle size, Phenomenex) with a flow rate of 200 µL/min. Solvent A was 97:3 H<sub>2</sub>O/MeOH with 10 mM tributylamine and 15 mM acetic acid; solvent B was methanol. The gradient was 0 min, 5% B; 5 min, 5% B; 7 min, 20% B; 17 min, 95% B; 20 min, 100% B; 24 min, 5% B; 30 min, 5% B. Q-Exactive chromatographic separation was achieved on an Poroshell 120 Bonus-RP column (150 x 2.1 mm,



689 2.7  $\mu$ m particle size, Agilent) with a flow rate of 200  $\mu$ L/min. Solvent A is 10mM ammonium  
690 acetate + 0.1% acetic acid in 98:2 water:acetonitrile and solvent B is acetonitrile. The gradient  
691 was 0 min, 2% B; 4 min, 0% B; 6 min, 30% B; 11 min, 100% B; 15 min, 100% B; 16 min, 2%  
692 B; 20 min, 2% B. LC-MS data were analyzed using the MAVEN software package (Clasquin et  
693 al., 2012)

## References

- Abdel-Hamid, A.M., and Cronan, J.E. (2007). Coordinate expression of the acetyl coenzyme A carboxylase genes, *accB* and *accC*, is necessary for normal regulation of biotin synthesis in *Escherichia coli*. *Journal of bacteriology* *189*, 369-376.
- Alam, M.T., Olin-Sandoval, V., Stincone, A., Keller, M.A., Zelezniak, A., Luisi, B.F., and Ralser, M. (2017). The self-inhibitory nature of metabolic networks and its alleviation through compartmentalization. *Nature communications* *8*, 16018.
- Babu, M., Arnold, R., Bundalovic-Torma, C., Gagarinova, A., Wong, K.S., Kumar, A., Stewart, G., Samanfar, B., Aoki, H., Wagih, O., *et al.* (2014). Quantitative genome-wide genetic interaction screens reveal global epistatic relationships of protein complexes in *Escherichia coli*. *PLoS genetics* *10*, e1004120.
- Babu, M., Diaz-Mejia, J.J., Vlasblom, J., Gagarinova, A., Phanse, S., Graham, C., Yousif, F., Ding, H., Xiong, X., Nazarians-Armavil, A., *et al.* (2011a). Genetic interaction maps in *Escherichia coli* reveal functional crosstalk among cell envelope biogenesis pathways. *PLoS genetics* *7*, e1002377.
- Babu, M., Gagarinova, A., and Emili, A. (2011b). Array-based synthetic genetic screens to map bacterial pathways and functional networks in *Escherichia coli*. *Methods in molecular biology* *781*, 99-126.
- Barenholz, U., Davidi, D., Reznik, E., Bar-On, Y., Antonovsky, N., Noor, E., and Milo, R. (2017). Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points. *eLife* *6*.
- Beverley, S.M., Ellenberger, T.E., and Cordingley, J.S. (1986). Primary structure of the gene encoding the bifunctional dihydrofolate reductase-thymidylate synthase of *Leishmania major*. *Proceedings of the National Academy of Sciences of the United States of America* *83*, 2584-2588.
- Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L.A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic acids research* *41*, 7429-7437.
- Bjarnason, G.A., Jordan, R.C., Wood, P.A., Li, Q., Lincoln, D.W., Sothorn, R.B., Hrushesky, W.J., and Ben-David, Y. (2001). Circadian expression of clock genes in human oral mucosa and skin: association with specific cell-cycle phases. *The American journal of pathology* *158*, 1793-1801.
- Bolhuis, A., Mathers, J.E., Thomas, J.D., Barrett, C.M., and Robinson, C. (2001). TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. *The Journal of biological chemistry* *276*, 20213-20219.

Butland, G., Babu, M., Diaz-Mejia, J.J., Bohdana, F., Phanse, S., Gold, B., Yang, W., Li, J., Gagarinova, A.G., Pogoutse, O., *et al.* (2008). eSGA: E. coli synthetic genetic array analysis. *Nature methods* 5, 789-795.

Carothers, D.J., Pons, G., and Patel, M.S. (1989). Dihydrolipoamide dehydrogenase: functional similarities and divergent evolution of the pyridine nucleotide-disulfide oxidoreductases. *Archives of biochemistry and biophysics* 268, 409-425.

Clasquin, M.F., Melamud, E., and Rabinowitz, J.D. (2012). LC-MS data processing with MAVEN: a metabolomic analysis and visualization engine. *Current protocols in bioinformatics* / editorial board, Andreas D Baxevanis [et al] *Chapter 14*, Unit14 11.

Deatherage, D.E., and Barrick, J.E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods in molecular biology* 1151, 165-188.

Ducker, G.S., and Rabinowitz, J.D. (2016). One-Carbon Metabolism in Health and Disease. *Cell metabolism*.

Feinauer, C., Szurmant, H., Weigt, M., and Pagnani, A. (2016). Inter-Protein Sequence Co-Evolution Predicts Known Physical Interactions in Bacterial Ribosomes and the Trp Operon. *PloS one* 11, e0149166.

Forsburg, S.L. (2001). The art and design of genetic screens: yeast. *Nature reviews Genetics* 2, 659-668.

Gagarinova, A., Stewart, G., Samanfar, B., Phanse, S., White, C.A., Aoki, H., Deineko, V., Beloglazova, N., Yakunin, A.F., Golshani, A., *et al.* (2016). Systematic Genetic Screens Reveal the Dynamic Global Functional Organization of the Bacterial Translation Machinery. *Cell reports* 17, 904-916.

Galperin, M.Y., Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2015). Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic acids research* 43, D261-269.

Gangjee, A., Jain, H.D., and Kurup, S. (2007). Recent advances in classical and non-classical antifolates as antitumor and antiopportunistic infection agents: part I. Anti-cancer agents in medicinal chemistry 7, 524-542.

Green, J.M., and Matthews, R.G. (2013). Folate Biosynthesis, Reduction, and Polyglutamylation and the Interconversion of Folate Derivatives. *EcoSal Plus*.

Halabi, N., Rivoire, O., Leibler, S., and Ranganathan, R. (2009). Protein sectors: evolutionary units of three-dimensional structure. *Cell* 138, 774-786.

Hopf, T.A., Scharfe, C.P., Rodrigues, J.P., Green, A.G., Kohlbacher, O., Sander, C., Bonvin, A.M., and Marks, D.S. (2014). Sequence co-evolution gives 3D contacts and structures of protein complexes. *eLife* 3.

Huynen, M., Snel, B., Lathe, W., 3rd, and Bork, P. (2000). Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. *Genome research* 10, 1204-1210.

Janga, S.C., Collado-Vides, J., and Moreno-Hagelsieb, G. (2005). Nebulon: a system for the inference of functional relationships of gene products from the rearrangement of predicted operons. *Nucleic acids research* 33, 2521-2530.

Janssen, H.J., and Steinbuchel, A. (2014). Fatty acid synthesis in *Escherichia coli* and its applications towards the production of fatty acid based biofuels. *Biotechnology for biofuels* 7, 7.

Junier, I., and Rivoire, O. (2013). Synteny in Bacterial Genomes: Inference, Organization and Evolution. *arXiv:13074291*.

Junier, I., and Rivoire, O. (2016). Conserved Units of Co-Expression in Bacterial Genomes: An Evolutionary Insight into Transcriptional Regulation. *PloS one* 11, e0155740.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic acids research* 40, D109-114.

Kim, J., and Copley, S.D. (2012). Inhibitory cross-talk upon introduction of a new metabolic pathway into an existing metabolic network. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2856-2864.

Kim, J., Kershner, J.P., Novikov, Y., Shoemaker, R.K., and Copley, S.D. (2010). Three serendipitous pathways in *E. coli* can bypass a block in pyridoxal-5'-phosphate synthesis. *Molecular systems biology* 6, 436.

Kim, P.J., and Price, N.D. (2011). Genetic co-occurrence network across sequenced microbes. *PLoS computational biology* 7, e1002340.

King, C.H., Shlaes, D.M., and Dul, M.J. (1983). Infection caused by thymidine-requiring, trimethoprim-resistant bacteria. *Journal of clinical microbiology* 18, 79-83.

Kondrashov, A.S., Sunyaev, S., and Kondrashov, F.A. (2002). Dobzhansky-Muller incompatibilities in protein evolution. *Proceedings of the National Academy of Sciences of the United States of America* 99, 14878-14883.

Kriegeskorte, A., Block, D., Drescher, M., Windmuller, N., Mellmann, A., Baum, C., Neumann, C., Lore, N.I., Bragonzi, A., Liebau, E., *et al.* (2014). Inactivation of *thyA* in *Staphylococcus aureus* attenuates virulence and has a strong impact on metabolism and virulence gene expression. *mBio* 5, e01447-01414.

Kumar, A., Beloglazova, N., Bundalovic-Torma, C., Phanse, S., Deineko, V., Gagarinova, A., Musso, G., Vlasblom, J., Lemak, S., Hooshyar, M., *et al.* (2016). Conditional Epistatic Interaction Maps Reveal Global Functional Rewiring of Genome Integrity Pathways in *Escherichia coli*. *Cell reports* 14, 648-661.

Kwon, Y.K., Lu, W., Melamud, E., Khanam, N., Bognar, A., and Rabinowitz, J.D. (2008). A domino effect in antifolate drug action in *Escherichia coli*. *Nature chemical biology* 4, 602-608.

Lazar, G., Zhang, H., and Goodman, H.M. (1993). The origin of the bifunctional dihydrofolate reductase-thymidylate synthase isogenes of *Arabidopsis thaliana*. *The Plant journal : for cell and molecular biology* 3, 657-668.

Leduc, D., Escartin, F., Nijhout, H.F., Reed, M.C., Liebl, U., Skouloubris, S., and Myllykallio, H. (2007). Flavin-dependent thymidylate synthase ThyX activity: implications for the folate cycle in bacteria. *Journal of bacteriology* 189, 8537-8545.

Lee, J., Natarajan, M., Nashine, V.C., Socolich, M., Vo, T., Russ, W.P., Benkovic, S.J., and Ranganathan, R. (2008). Surface sites for engineering allosteric control in proteins. *Science (New York, NY)* 322, 438-442.

Li, Y., Calvo, S.E., Gutman, R., Liu, J.S., and Mootha, V.K. (2014). Expansion of biological pathways based on evolutionary inference. *Cell* 158, 213-225.

Long, C.P., Gonzalez, J.E., Feist, A.M., Palsson, B.O., and Antoniewicz, M.R. (2018). Dissecting the genetic and metabolic mechanisms of adaptation to the knockout of a major metabolic enzyme in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 115, 222-227.

Lu, W., Clasquin, M.F., Melamud, E., Amador-Noguez, D., Caudy, A.A., and Rabinowitz, J.D. (2010). Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. *Analytical chemistry* 82, 3212-3221.

Lu, W., Kwon, Y.K., and Rabinowitz, J.D. (2007). Isotope ratio-based profiling of microbial folates. *Journal of the American Society for Mass Spectrometry* 18, 898-909.

McGuire, J.J., and Bertino, J.R. (1981). Enzymatic synthesis and function of folylpolyglutamates. *Molecular and cellular biochemistry* 38 Spec No, 19-48.

McLaughlin, R.N., Jr., Poelwijk, F.J., Raman, A., Gosal, W.S., and Ranganathan, R. (2012). The spatial architecture of protein function and adaptation. *Nature* 491, 138-142.

Michener, J.K., Camargo Neves, A.A., Vuilleumier, S., Bringel, F., and Marx, C.J. (2014a). Effective use of a horizontally-transferred pathway for dichloromethane catabolism requires post-transfer refinement. *eLife* 3.

Michener, J.K., Vuilleumier, S., Bringel, F., and Marx, C.J. (2014b). Phylogeny poorly predicts the utility of a challenging horizontally transferred gene in *Methylobacterium* strains. *Journal of bacteriology* 196, 2101-2107.

Morcos, F., Pagnani, A., Lunt, B., Bertolino, A., Marks, D.S., Sander, C., Zecchina, R., Onuchic, J.N., Hwa, T., and Weigt, M. (2011). Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proceedings of the National Academy of Sciences of the United States of America* 108, E1293-1301.

Newman, M.E.J. (2010). *Networks : an introduction* (Oxford ; New York: Oxford University Press).

Nijhout, H.F., Reed, M.C., Budu, P., and Ulrich, C.M. (2004). A mathematical model of the folate cycle: new insights into folate homeostasis. *The Journal of biological chemistry* *279*, 55008-55016.

Okamura-Ikeda, K., Ohmura, Y., Fujiwara, K., and Motokawa, Y. (1993). Cloning and nucleotide sequence of the *gcv* operon encoding the *Escherichia coli* glycine-cleavage system. *European journal of biochemistry / FEBS* *216*, 539-548.

Ovchinnikov, S., Kamisetty, H., and Baker, D. (2014). Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife* *3*, e02030.

Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D., and Yeates, T.O. (1999). Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proceedings of the National Academy of Sciences of the United States of America* *96*, 4285-4288.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* *152*, 1173-1183.

Raman, A.S., White, K.I., and Ranganathan, R. (2016). Origins of Allostery and Evolvability in Proteins: A Case Study. *Cell* *166*, 468-480.

Rengby, O., Johansson, L., Carlson, L.A., Serini, E., Vlamis-Gardikas, A., Karsnas, P., and Arner, E.S. (2004). Assessment of production conditions for efficient use of *Escherichia coli* in high-yield heterologous recombinant selenoprotein synthesis. *Applied and environmental microbiology* *70*, 5159-5167.

Reynolds, K.A. (2014). Finding a common path: predicting gene function using inferred evolutionary trees. *Developmental cell* *30*, 4-5.

Reynolds, K.A., McLaughlin, R.N., and Ranganathan, R. (2011). Hotspots for allosteric regulation on protein surfaces. *Cell* *147*, 1564-1575.

Reznik, E., Christodoulou, D., Goldford, J.E., Briars, E., Sauer, U., Segre, D., and Noor, E. (2017). Genome-Scale Architecture of Small Molecule Regulatory Networks and the Fundamental Trade-Off between Regulation and Enzymatic Activity. *Cell reports* *20*, 2666-2677.

Rivoire, O. (2013). Elements of coevolution in biological sequences. *Physical review letters* *110*, 178102.

- Salvador, A., and Savageau, M.A. (2003). Quantitative evolutionary design of glucose 6-phosphate dehydrogenase expression in human erythrocytes. *Proceedings of the National Academy of Sciences of the United States of America* *100*, 14463-14468.
- Salvador, A., and Savageau, M.A. (2006). Evolution of enzymes in a series is driven by dissimilar functional demands. *Proceedings of the National Academy of Sciences of the United States of America* *103*, 2226-2231.
- Sebaugh, J.L. (2011). Guidelines for accurate EC50/IC50 estimation. *Pharmaceutical statistics* *10*, 128-134.
- Skerker, J.M., Perchuk, B.S., Siryaporn, A., Lubin, E.A., Ashenberg, O., Goulian, M., and Laub, M.T. (2008). Rewiring the specificity of two-component signal transduction systems. *Cell* *133*, 1043-1054.
- Smock, R.G., Rivoire, O., Russ, W.P., Swain, J.F., Leibler, S., Ranganathan, R., and Gierasch, L.M. (2010). An interdomain sector mediating allostery in Hsp70 molecular chaperones. *Molecular systems biology* *6*, 414.
- Snel, B., Bork, P., and Huynen, M.A. (2002). The identification of functional modules from the genomic association of genes. *Proceedings of the National Academy of Sciences of the United States of America* *99*, 5890-5895.
- Suel, G.M., Lockless, S.W., Wall, M.A., and Ranganathan, R. (2003). Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nature structural biology* *10*, 59-69.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., *et al.* (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research* *43*, D447-452.
- Tas, H., Nguyen, C.T., Patel, R., Kim, N.H., and Kuhlman, T.E. (2015). An Integrated System for Precise Genome Modification in *Escherichia coli*. *PloS one* *10*, e0136963.
- Toprak, E., Veres, A., Michel, J.B., Chait, R., Hartl, D.L., and Kishony, R. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nature genetics* *44*, 101-105.
- Toprak, E., Veres, A., Yildiz, S., Pedraza, J.M., Chait, R., Paulsson, J., and Kishony, R. (2013). Building a morbidostat: an automated continuous-culture device for studying bacterial drug resistance under dynamically sustained drug inhibition. *Nature protocols* *8*, 555-567.
- Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L., *et al.* (2008). High-throughput, quantitative analyses of genetic interactions in *E. coli*. *Nature methods* *5*, 781-787.

Zuk, O., Hechter, E., Sunyaev, S.R., and Lander, E.S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences of the United States of America* *109*, 1193-1198.





Figure 2  
Schober et al.

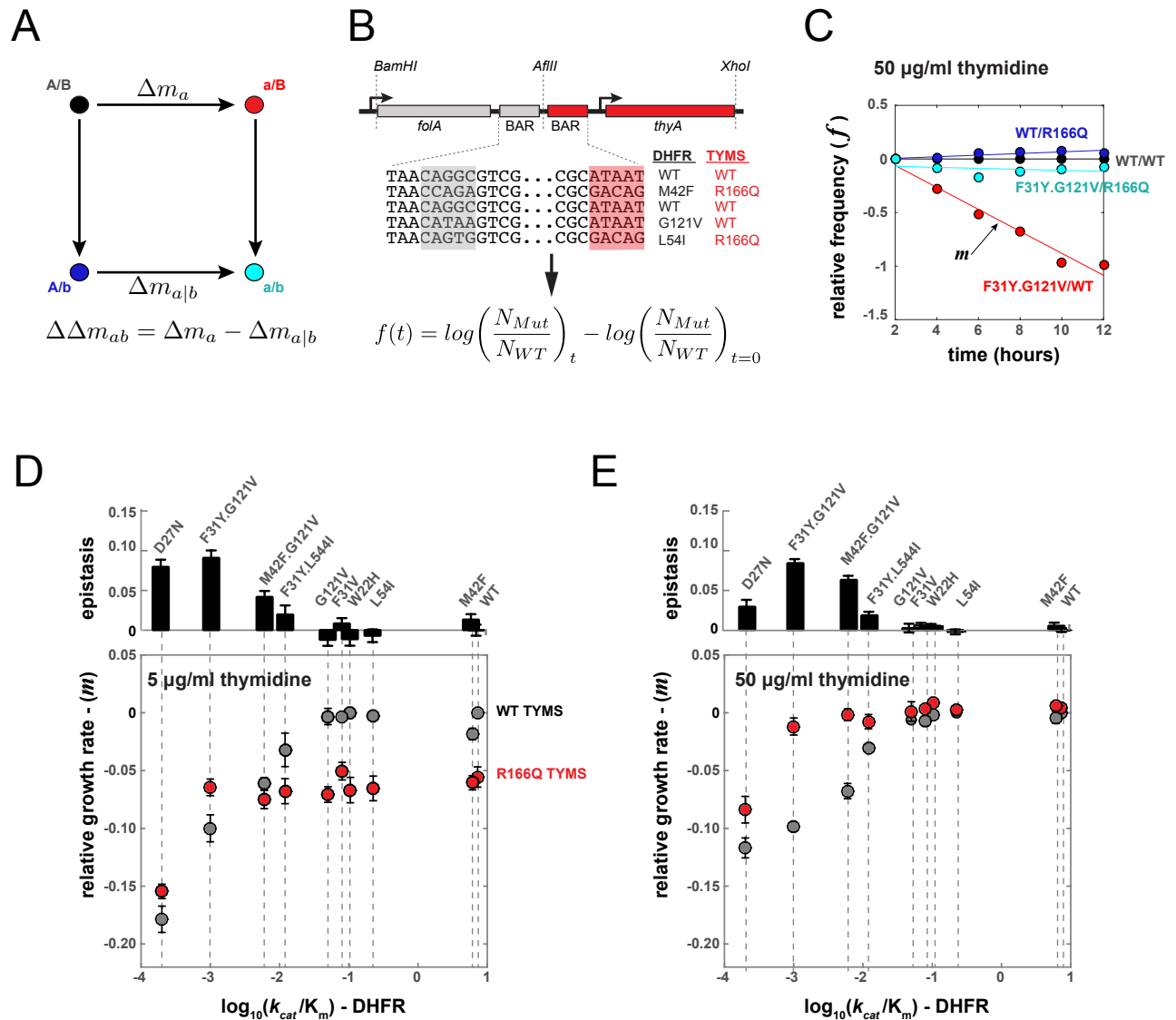


Figure 3  
Schober et al.

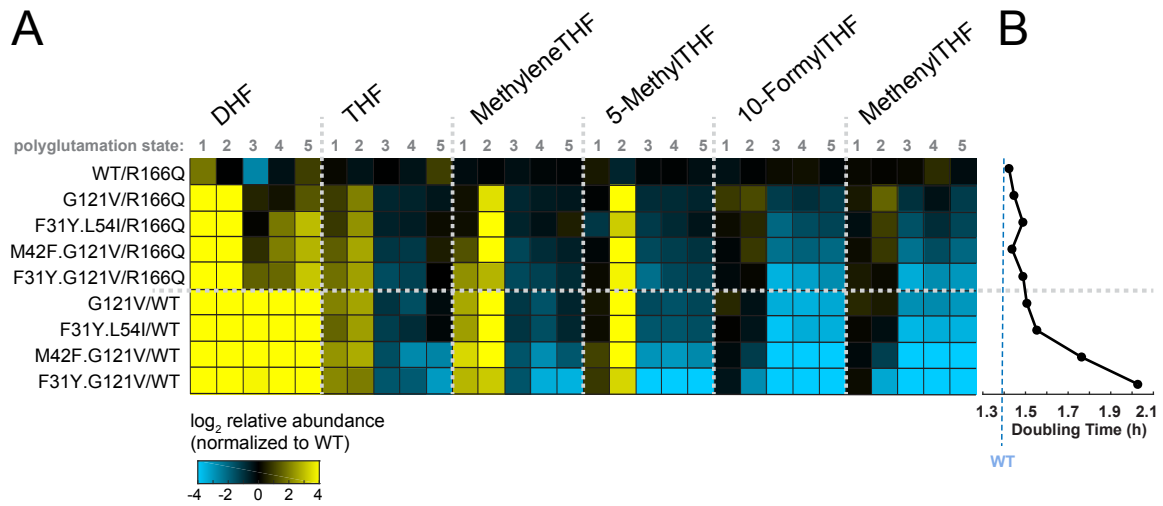


Figure 4  
Schober et al.

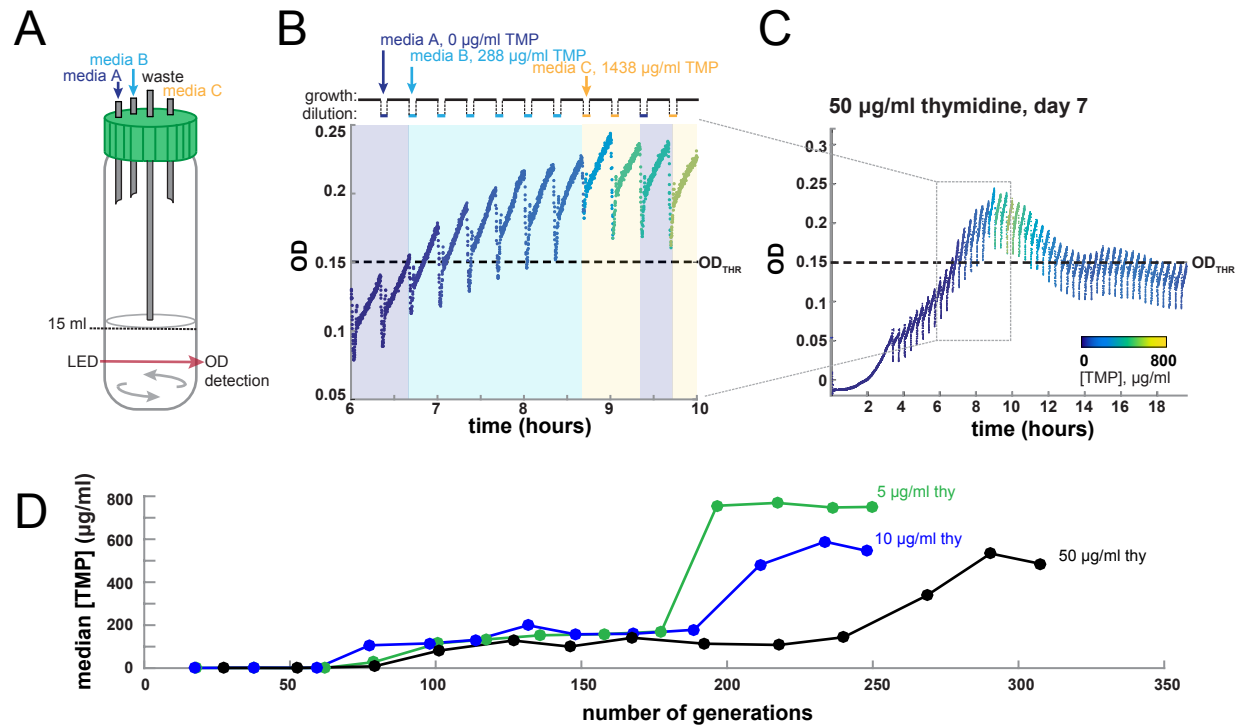




Figure 6  
Schober et al.

