Main Manuscript for

A continuous epistasis model for predicting growth rate given combinatorial variation in gene expression and environment

Ryan M. Otto[1], Agata Turska-Nowak[2‡], Philip M. Brown[1], Kimberly A. Reynolds[1,2*]

[1]Green Center for Systems Biology - Lyda Hill Department of Bioinformatics, The University of Texas Southwestern Medical Center, Dallas, TX 75230 USA; [2]Department of Biophysics, The University of Texas Southwestern Medical Center, Dallas, TX 75230 USA

[‡]Current affiliations: Polish Stem Cells Bank, Warsaw, Poland; FamiCordTx, Warsaw, Poland
*Correspondence: kimberly.reynolds@utsouthwestern.edu

Preprint: BioRxiv. https://doi.org/10.1101/2022.08.19.504444, CC BY-NC-ND

Summary

The ability to predict changes in cellular growth rate given variation in gene expression is critical to engineer biosynthetic pathways and understand evolutionary constraints on mRNA abundance. However, the relationship between gene expression and growth rate is nonlinear and shaped by environmental and genetic context. To address these challenges, we examined the relationship between enzyme expression and *E. coli* growth rate for 36 metabolic gene pairs by measuring growth following 4,404 pairwise, titrated changes in enzyme abundance under varied environmental contexts. Using 20% of these data, we trained an interpretable epistatic model to predict growth rate following simultaneous expression and environmental perturbations. The model's predictions were robust to significant genetic and environmental epistasis and at times approached the accuracy of biological replicates. Based on these results, we propose a strategy using sparsely sampled, low-order measurements to quantify genetic interaction landscapes on the pathway-wide or genomic scale in a single assay.

Keywords: Genotype-to-phenotype, gene-by-environment, epistasis, gene expression, CRISPRi

Introduction

Variation in gene expression tunes organism-wide traits like biomass yield, cell division, and differentiation, but predicting these broad phenotypes based on molecular data remains challenging. In the context of bacterial infection, expression changes can lead to immune evasion[1], trigger a switch between commensal and pathogenic behavior[2], or confer antibiotic tolerance or resistance[3,4]. Similarly, dysregulated transcription in human cells can initiate tumorigenesis and metastasis[5,6] and transcriptional diversity within tumors confers differential susceptibility to chemotherapeutics and radiation[7,8]. If transcriptional molecular markers can be leveraged to predict these phenotypes, we could drastically improve patient prognosis. More generally, the relationship between gene expression and cellular growth rate represents a well-defined instantiation of the broader problem of mapping genotype to phenotype.

However, two complications stand in the way of making such predictions. First, expression changes occur over a continuous range and are often non-linearly related to growth rate[9-11]. Thus, a simple on/off mechanism is rarely sufficient to describe the continuum of expression levels observed in a population. While “endpoint” studies such as gene knockouts or near-complete knockdowns have classically proved useful for annotating gene function[12], discovering new regulatory networks[13], and revealing the organization of the cell[14-17], they do not offer sufficient information to predict the effects of intermediate changes in gene expression. Sampling a range of perturbation strengths is required to quantify this continuous, nonlinear response. Second, expression-growth rate relationships are context-dependent[9,10,18]. An organism’s genetic background and environment may alter the effect of gene expression changes, a phenomenon called epistasis[19,20]. The amount of data required to densely sample these context-dependent relationships scales exponentially with the complexity of the system. As a consequence, these experiments quickly become intractable. A new approach is required to approximate these landscapes using a reasonable amount of input data.

Interestingly, prior work on a related problem – predicting the effects of antibiotic and chemotherapeutic combination therapies on cell growth – has considered continuous and/or context-dependent models[21-28]. Of particular relevance is the dose-response model introduced in

Zimmer et al., 2016[28]. The authors examined cases where the presence of one drug altered the effect of a second drug. This interaction resulted in a change in the "effective dose" of the second drug and was summarized with two coupling constants – the effect of drug "A" on "B", and the effect of drug "B" on "A". The authors found that a model constructed using only these pairwise coupling terms could predict the effects of higher-order combinations of drugs. This model was unit-agnostic, meaning the framework could be trained on any quantitative input, not just drug treatment. However, to our knowledge, it has never been extended to make predictions of growth rate given variation in gene expression or nutrient abundance. Following the rationale that a reduction in enzyme expression is functionally analogous to inhibiting an enzyme with a drug, as well as studies identifying transcription as the main determinant of protein abundance in bacteria[29], we hypothesized that a similar coupling-based approach could capture context-dependence between genes and predict the impact of multiple gene expression changes on bacterial growth rate.

In this work, we considered continuous expression-growth rate relationships for *E. coli* metabolic enzymes drawn from diverse metabolic pathway architectures. We measured the growth rate effects of 4,404 pairwise combinations of titrated gene knockdowns. Using 20% of these data, we fit a continuous, coupling-sensitive model that was able to predict growth rates for the remaining 80% of pairwise perturbations. We then tested the model's ability to predict the effect of third-order gene knockdowns without being trained on high-order data and found that sparse pairwise interaction measurements were sufficient to recapitulate the phenotypes observed in this high-order study. Finally, we extended this model to predict the combined effect of gene expression and varied media additives. Growth rate predictions for third- and fourth-order genetic and environmental perturbations remained robust. Throughout these studies, this interpretable approach provided biological insight by identifying asymmetrical genetic and environmental interactions as well as hinting at the existence of high-order interactions in regimes where experimental data diverged from the model's predictions. Taken together, our experimental and computational methods provide a strategy to incorporate sparsely sampled epistasis measurements into a predictive model. This approach can be used to predict bacterial growth rate following numerous intermediate gene expression changes in a variety of environments and is

57 scalable to quantify expression-growth rate coupling across entire metabolic pathways or
58 potentially even genomes.
59
60 **Results**
61 **Growth rate titrates with gene expression following CRISPR interference**
62 Bacterial gene expression is a key determinant of cellular growth rate, and gene expression levels
63 are constrained at both absolute and relative levels. Within these constraints, however, tuning of
64 gene expression can modulate, but not eliminate, bacterial growth (Fig 1A). Classic genetic
65 interaction measurements consider pairwise interactions between genetic knockouts or strong
66 knockdowns and do not permit prediction of growth rates following intermediate changes in
67 expression or gene product activity (Fig 1B). We sought to capture this intermediate variation by
68 generating a continuous expression-growth rate model for a group of metabolic genes in *E. coli*
69 (Fig 1C). However, the amount of data required to directly map these epistatic landscapes scales
70 exponentially as you consider additional expression levels or new genes. Thus, our modeling
71 strategy needs to predict these continuous epistatic landscapes using only sparsely sampled
72 pairwise measurements.
73 Training and testing an expression-growth rate model requires high-throughput growth rate
74 measurements across cells harboring distinct, intermediate variations in gene expression. To
75 generate these data, we used mismatch CRISPR interference (CRISPRi). CRISPRi uses a
76 catalytically dead Cas9 endonuclease (dCas9) and a single guide RNA (sgRNA) to target a specific
77 DNA locus, sterically repressing transcription at the target site[30]. In standard CRISPRi, the target
78 of the sgRNA is specified by a 20-nucleotide homology region designed to be fully complementary
79 to the target gene. In contrast, mismatch CRISPRi uses sgRNAs containing nucleotide mismatches
80 in the target homology region[11,31-33]. These mismatches disrupt sgRNA-DNA interactions,
81 decreasing gene repression strength. By designing a group of sgRNAs targeting the same DNA
82 locus with different numbers of mismatches, we can tune a single gene's transcription to numerous
83 intermediate levels (Fig 1D). Growth rate defects caused by these knockdowns can be quantified
84 at high throughput using CRISPRi-seq[18,31,34-36]. In our implementation of this approach, we

85 transform a library of sgRNAs into *E. coli* containing a chromosomally encoded dCas9 under control
86 of a tet-inducible promoter. To control population size and environmental conditions, we grow the
87 transformed cells in a turbidostat[37]. We then induce each cell's CRISPRi machinery and use next-
88 generation sequencing to monitor changes in the frequency of individual sgRNAs over time. A linear
89 regression fit to the logarithm of each sgRNA's frequency over time allows us to estimate the growth
90 rate effects of thousands of sgRNAs in a single experiment (Fig 1E). By plotting growth rate as a
91 function of sgRNA repression strength, we can define continuous expression-growth rate functions
92 for genes of interest (Fig 1F).

93 In this study, we selected nine genes to target with mismatch CRISPRi, with an eye to
94 sampling distinct metabolic pathways and architectures (Fig 2A-D). Included were genes encoding
95 enzymes catalyzing consecutive metabolic steps (*dapA*/*dapB*, *purN*/*purL*), parallel reactions
96 (*gdhA*/*gltB*), and a three-enzyme loop (*folA*/*glyA*/*thyA*). All genes except *purN*, *gdhA*, and *gltB* were
97 previously shown to be essential in our experimental conditions[38] (OD_{600} < 0.01 after 24 hours in
98 minimal media). Double knockout of *gdhA*/*gltB* was shown to be synthetically lethal[39]. We designed
99 9-12 distinct sgRNAs targeting each gene as well as a non-targeting sgRNA control to generate a
100 96 sgRNA library (Table S1). We quantified the effect of each CRISPRi perturbation on the mRNA
101 concentration of its target gene using reverse transcription-quantitative PCR (RT-qPCR). As
102 expected, our library of sgRNAs created a titrated range of knockdowns for every gene in our data
103 set.
104

105 **Pairwise CRISPRi yields an expression-dependent epistatic landscape**

106 To investigate expression-growth rate coupling between genes, we used Golden Gate cloning to
107 assemble a library of all pairwise combinations of sgRNAs[40]. We constructed this library in three
108 distinct barcoded plasmid vectors to generate internal replicates. Sequencing of our library prior to
109 selection showed that it was complete and well-distributed across the designed sgRNAs (Fig S1A).
110 We quantified the growth rate effect of each CRISPRi knockdown (both singles and pairwise) on
111 exponentially growing *E. coli* cells in M9 minimal media with glucose using continuous culture and
112 next-generation sequencing. Growth rate was well-correlated between internal replicates (Fig

113 S1B). In agreement with earlier work[30], changing the sgRNA order within pairwise constructs
114 (sgRNA1-sgRNA2 vs. sgRNA2-sgRNA1) did not significantly alter growth rate (Fig S1C). Thus,
115 after pooling our three internal replicates for each of two sgRNA orders, we had six independent
116 measurements for each pairwise CRISPRi treatment. As in prior work[33], we used these replicates
117 to identify and filter out "escapers": replicates whose growth rate was significantly faster than all
118 other replicates of the same CRISPRi treatment, likely due to evasion of the CRISPRi machinery
119 (Fig S1D). Escaper correction removed 182 measurements (0.7% of the data, Fig S1E-F). Finally,
120 we averaged the remaining replicates and empirically rescaled our growth rate data, setting the
121 minimum growth rate observed to 0 and the non-targeting CRISPRi control to 1. In total, we
122 quantified the growth rate effect of 4,404 unique CRISPRi perturbations. We found that growth
123 rates from this study correlated well with measurements from our previous work investigating
124 single-gene knockdowns[33] (Fig S1G).
125 We quantified single gene expression-growth rate relationships by relating target gene repression
126 strength to growth rate for sgRNA constructs consisting of one targeting and one non-targeting
127 control sgRNA (Fig 2E-M). All genes showed monotonic expression-growth rate relationships
128 except *gdhA* and *gltB*, which were known to be nonessential in our growth condition. Notably, our
129 approach to sgRNA library design reliably sampled the steep, intermediate growth regime of each
130 gene, allowing us to precisely identify the knockdown level where each gene's transcription became
131 growth-limiting. We saw that the general shape of repression-growth rate curves was similar for
132 biochemically related genes – for example *dapA*/*dapB* (Fig 2E-F) or *purN*/*purL* (Fig 2J-K), in line
133 with previous findings[11].
134 In addition to these single gene relationships, we densely sampled pairwise expression-
135 growth rate landscapes (Fig 3, lower triangle). In this matrix, every row and column represents a
136 unique sgRNA, sorted by target gene and repression strength, and each pixel represents the
137 growth rate effect of a given sgRNA pair. To quantify interactions between CRISPRi perturbations
138 cooccurring within a cell, we quantified epistasis between every pair of CRISPRi knockdowns using
139 a multiplicative (Bliss) model, where epistasis between sgRNAs 1 and 2 is equal to $gr_{1,2} - gr_1 * gr_2$
140 (Fig 3, upper triangle). For the majority of sgRNA pairs, we observed minimal epistasis. This is

141 consistent with expectation given that: 1) many of the gene pairs are drawn from unrelated
142 metabolic pathways and 2) many sgRNAs yielded only modest changes in gene expression. Across
143 the library, however, we observed both positive and negative epistasis, particularly for sgRNAs with
144 the strongest effects on gene expression. Our data also recapitulated the synthetic lethality
145 previously reported for *gdhA* and *gltB* as negative epistasis[39]: we observed minimal growth rate
146 defects when these genes were repressed individually, but severe growth rate defects when both
147 genes were repressed together. However, even for gene pairs where epistasis was prevalent, the
148 intensity of epistasis varied drastically within different expression regimes. Given how widely gene-
149 gene epistasis varied based on the specific perturbations explored, we sought to discover a new
150 way to summarize and model gene-gene epistasis across entire pairwise expression-growth rate
151 landscapes.
152

153 **A continuous, coupling-sensitive model of expression-growth rate relationships**

154 In developing the model, we sought to move beyond a matrix of many discrete epistasis
155 measurements to a continuous, predictive description of expression-growth rate relationships that
156 would be amendable to future studies of biological data. Under the assumption that densely
157 measuring single perturbation-phenotype relationships is efficient (scaling linearly as the number
158 of perturbations explored increases), we used well-constrained single perturbation functions as the
159 backbone of our modeling strategy. To begin, we modeled the effect of single gene knockdowns
160 on growth rate. We found that a sigmoidal model relating relative repression (R) to growth rate (g)
161 using two parameters, steepness (n) and repression at half-maximal growth rate (R_0) effectively
162 modeled all single gene knockdown data (Equation 1, Fig 2E-M, Table S2).

$$g(R) = \frac{1}{1 + e^{n(R-R_0)}} \tag{1}$$

163

164 These sigmoidal fits enabled continuous prediction of the effects of gene repression on
165 growth rate, improving upon our discrete experimental measurements. We used bootstrap
166 resampling of the data to estimate 95% confidence intervals (CIs) for the sigmoidal fit and each
167 parameter (Table S2, Methods). These fit parameters were well constrained for all genes except

168 *gdhA* and *gltB* – two nonessential genes without significant growth rate defects. Although these
169 two expression-growth rate functions can be fit by a wide range of n and R_0 values, we will show
170 that our model's performance is insensitive to this variation and thus robust for essential and
171 nonessential genes. We wanted to retain this continuous predictive power as we expanded our
172 growth rate predictions to pairwise CRISPRi knockdowns. Assuming independence between
173 CRISPRi perturbations, we constructed a Null model for pairwise knockdowns by multiplicatively
174 combining the growth rate effects of two single gene knockdown functions (Equation 2).

$$g(R_1, R_2) = g_1(R_1) * g_2(R_2) \tag{2}$$

175

176 This initial model enabled continuous prediction of growth rate following pairwise gene
177 knockdown but did not explicitly account for the possibility of gene-gene epistasis. To introduce
178 epistasis into this modeling framework, we fit coupling constants a_{ij} and a_{ji} to modulate the effective
179 repression of a given gene (R_{ieff}) based on the intensity of a secondary perturbation (R_{jeff}) and vice
180 versa, as in Zimmer et al. 2016[28] (Equation 3). These effective repression values are then used in
181 place of true repression values in equations 1 & 2.

$$R_{ieff} = R_i \prod_{j \neq i} \left(1 + a_{ij} \frac{R_{jeff}/R_{0j}}{1 + R_{jeff}/R_{0j}}\right)^{-1} \tag{3}$$

182

183 Based on the assumption of sparse gene-gene coupling, we introduced a regularization
184 term to penalize the absolute values of a_{ij} and a_{ji} (Methods). Using mild regularization, we reduced
185 the spread of coupling constant values while retaining our model's predictive power (Fig S2, Table
186 S3).

187 To investigate these pairwise expression-growth rate landscapes further, we highlight
188 several example gene pairs. Firstly, for gene pair *dapA*/*purN* (Fig 4A), we observed that pairwise
189 knockdowns (lower right pixels) had less severe growth rate defects than would be expected from
190 single knockdown measurements alone (upper right and lower left). Our continuous epistasis model
191 captured this interaction with a pair of positive gene-gene coupling terms (a_{ij} = 0.061, 95% CI
192 [0.042, 0.099], a_{ji} = 0.197, [0.132, 0.238], Fig 4B), while the Null model erroneously underestimated

growth (Fig 4C). Notably, the couplings between *dapA* and *purN* are asymmetric: in this instance, it reflects the fact that *purN* repression more strongly modulates the effect of *dapA* repression than the converse. In contrast, the synthetic lethal *gdhA* and *gltB* show negative epistasis, where the double knockdown phenotype was more severe than expected (Fig 4D). In our continuous epistasis model, this was captured with one negative gene-gene coupling term and one zero coupling term (a_{ij} = 0, [-1.06, 0], a_{ji} = -0.425, [-1.19, 0], Fig 4E), but this phenotype was again missed by the Null model (Fig 4F). In this case, the apparent coupling directionality for *gdhA*/*gltB* is the result of our regularization strategy and, as the confidence intervals indicate, either coupling constant can be negative to recapitulate synthetic lethality. The gene pair *dapA*/*dapB* shows positive coupling with near-symmetrical coupling constants, which we infer is because their gene products catalyze consecutive metabolic reactions (a_{ij} = 0.361, [0.237, 0.449], a_{ji} = 0.327, [0.267, 0.387], Fig S3A-C). Metabolic flux is limited by only a single knockdown, and the effect of an additional knockdown is dampened by this first flux restriction. However, we see almost no interaction between *purN* and *purL*, even though the products of these genes are also consecutive (a_{ij} = 0.053, [0.034, 0.095], a_{ji} = 0.009, [0.002, 0.016], Fig S3D-F). Thus, in the presence of symmetrical or asymmetrical coupling, the continuous epistasis model summarizes gene-gene coupling across entire pairwise expression-growth rate landscapes without incorporating bias from prior expectations.

The resulting epistatic model effectively predicted growth rates following knockdown for every gene pair, as quantified by the root-mean-square deviation (RSMD) between the model's predictions and experimentally determined growth rates (Fig 4G). These decreases in RMSD (visualized as a point falling below y = x) illustrated that our model outperformed the nonepistatic Null model for most gene pairs. Indeed, when we calculated each model's overall RMSD by pooling data from every gene pair together, our model (RMSD = 0.123, [0.122, 0.139]) significantly outperformed the Null model (RMSD = 0.158, [0.151, 0.173]). We wanted to ensure that these improvements were sufficient to warrant the additional complexity introduced by fitting coupling constants. Thus, we computed the Akaike Information Criterion (AIC) for each model, which supported the inclusion of these additional parameters (our model AIC = -15,716, [-15,785,

220 -14,743], Null AIC = -13,958, [-14,299, -13,262]). We refer to this regularized, coupling-sensitive
221 model as the “continuous epistasis” model.

222 Next, we sought to understand the relationship between the couplings learned by our
223 approach and more standard couplings computed from “binary” genetic interaction measurements.
224 We compared the gene-gene coupling values (a_{ij} and a_{ji}) fit using our model to epistatic couplings
225 calculated from a Bliss model considering only maximal-strength knockdown, mimicking a classical
226 double-knockout study (Fig 4H-I). We found that coupling, when considered over an entire pairwise
227 expression-growth rate landscape, is vastly less common than what is indicated by end-point
228 perturbations alone. This insight highlights how seemingly disparate epistasis measurements can
229 be concisely summarized when treated as a continuous landscape.

230 Finally, we tested two alterations to our model. For the first, we tried an alternate method
231 with data requirements between the binary approach illustrated above and our continuous modeling
232 technique by rigorously fitting single gene repression-growth rate curves, but fitting coupling
233 constants using only the most extreme double-knockdown measurement. This allowed us to
234 evaluate the utility of training on growth rate data following intermediate, pairwise knockdowns.
235 This approach was prone to overfitting on the double-knockdown growth rate and failed to recreate
236 intermediate phenotypes or accurately quantify epistasis (full model RMSD = 0.194, [0.174, 0.206]).
237 As extreme double-knockdowns cause the most severe growth rate defects, and experimental
238 noise in our system increases as growth rate decreases, this overfit model performs worse than
239 the nonepistatic Null (Fig S3G). Second, to test whether the varying asymmetries we noticed in
240 epistatic coupling constants were meaningful and indicative of true biological directionality, we
241 transposed the coupling constants fit from our initial model, swapping a_{ij} and a_{ji} terms within each
242 gene pair, and reassessed model performance. This transposed model performed significantly
243 worse than the model with correct coupling directionalities, although it still outperformed our Null
244 model as some epistatic coupling constants are truly symmetric (Fig S3H, RMSD = 0.137, [0.131,
245 0.154]).
246

247 **The expression-growth rate model can be reliably inferred using a small number of**
248 **measurements**
249 A key strength of this modeling strategy is its ability to represent an entire pairwise expression-
250 growth rate landscape using a short list of parameters (n and R_0 for each of two genes, a_{ij}, and a_{ji}).
251 However, if the pairwise landscape must be densely sampled to reliably fit these parameters, a
252 predictive model will be of little additional use. To quantify the minimal sampling requirements to
253 accurately fit coupling constants, we evaluated the effect of subsampling our pairwise expression
254 data and refitting the model on a smaller training set. As we wanted this analysis to inform library
255 design strategies for future experiments, we tested three distinct subsampling patterns:
256 subsampling a fraction of sgRNA pairs from across our entire library (Fig S4A), subsampling a
257 fraction of sgRNA pairs for each gene pair in our library (Fig S4B), and subsampling a fraction of
258 sgRNAs from our initial oligo pool, then constructing all pairwise combinations of the remaining
259 sgRNAs (Fig S4C). The first approach would be implemented by constructing a large pairwise
260 library and then bottlenecking prior to growth rate selection and sequencing, while the last is akin
261 to this study with a smaller starting library. The middle approach captures an intermediate study
262 (and may be more difficult to realize experimentally). Using each approach, we randomly split our
263 pairwise expression-growth rate data into a training and test set, fit epistatic coupling constants for
264 each gene pair using the training data, and evaluated the model's growth rate predictions on our
265 test data over 100 unique subsampling iterations. By varying the relative sizes of our training and
266 test sets, we determined that the model is reliably fit with 20% of our original input data when
267 subsampling using either of the first two strategies (Fig S4D-E). The third approach, subsampling
268 sgRNAs from the starting library, was robust to subsampling down to 40% of the initial library size.
269 Given these data, we recommend bottlenecking a diverse CRISPRi library as a simple but robust
270 strategy to sparsely sample epistatic landscapes (RMSD = 0.128, [0.125, 0.144], Fig S5).
271
272 **High-order predictions can be made using low-order training data**
273 We next wanted to evaluate whether a model trained on only single and pairwise expression-growth
274 rate data could predict the effects of higher-order CRISPRi perturbations. This would vastly dampen

275 the effect of the combinatorial explosion problem, enabling more complex expression-growth rate
276 predictions. As such, we constructed a third-order sgRNA library and quantified growth rates
277 following these third-order CRISPRi perturbations (Fig 5A). The library contained 911 distinct
278 constructs consisting of 4 single, 76 pairwise, and 830 triple gene knockdowns and a non-targeting
279 control (Fig 5B). Four additional non-targeting sgRNA controls were added to the library to ensure
280 the presence of low-order sgRNA constructs in the final library. These alternative non-targeting
281 sgRNAs phenocopy our original non-targeting control (Fig S6A). As each sequencing experiment
282 measures growth rates relative to the overall culture and bulk culture doubling time varies based
283 on library contribution, we rescaled relative growth rates from this experiment such that our single
284 and pairwise CRISPRi knockdowns were well-correlated to those from our pairwise data set
285 (RMSD = 0.118, Fig S6B).

286 We used the gene-gene coupling values fit on our pairwise data to predict growth rates
287 following all third-order CRISPRi perturbations. These predictions outperformed the coupling-
288 insensitive Null model (RMSD = 0.174 vs RMSD = 0.234 for the epistatic and Null models,
289 respectively, Fig 5C-D). Both models underestimate growth rate near wildtype, as these minor
290 perturbations seem to combine sub-additively. However, when considering moderate and severe
291 growth rate defects (where ~75% of the data lie), the models' predictions diverge. The continuous
292 epistasis model recapitulates growth rates observed in our data, while the Null model continues
293 underestimating growth. The accuracy of the Null model is also highly variable based on the
294 presence or absence of epistasis between sgRNAs in a given construct, while the continuous
295 epistasis model accounts for these interactions in its predictions and performs comparably across
296 all epistatic regimes.

297 We then compared our model's predictions to ones made using two alternative methods
298 of estimating high-order phenotypes from low-order data: the Isserlis and regression-based
299 models. These models lack interpretability, as they do not explicitly account for epistasis using
300 learned parameters, and instead predict growth rates directly from single- and pairwise perturbation
301 measurements. The models are also not inherently continuous, so we utilized "smoothed" versions
302 of these models[28]. The Isserlis model performed comparably to our own, outperforming our model

303 for growth rates near wildtype at the cost of increased uncertainty under slower growth conditions
304 (RMSD = 0.160, Fig S6C). The regression-based model performed poorly due to its sensitivity to
305 variance in low-order measurements (RMSD = 0.238, Fig S6D). Our model's ability to extrapolate
306 from pairwise data to predict the effects of higher-order gene knockdowns suggests that future
307 studies of genetic epistasis across metabolic or transcriptional networks could be truncated at
308 pairwise measurements, dampening the effect of combinatorial complexity.
309
310 **The continuous epistasis model captures gene-by-environment interactions**
311 Importantly, the continuous epistasis model is unit-agnostic – because the model considers
312 perturbation intensities on a relative scale, couplings need not be limited to gene-gene epistasis,
313 but might more generally describe interactions between genes, environments, and other
314 continuous perturbations. To assess the ability of the model to capture gene-by-environment
315 interactions, we used mismatch CRISPRi to repress transcription of the folate metabolic genes
316 *folA* and/or *thyA* while supplementing culture growth media with two products of folate
317 metabolism: thymidine and/or methionine (Figure 6A, red text). Previous studies demonstrated
318 that these perturbations have higher-order growth rate interactions[41-43]. In our own prior work, we
319 observed that *folA* and *thyA* form an environmentally-dependent synthetic rescue. That is, growth
320 rate defects due to loss of function mutations to DHFR (the gene encoded by *folA*) are rescued by
321 loss of function mutations to TYMS (the gene encoding *thyA*) in the presence (but not the
322 absence) of thymidine[41]. Moreover, the balance between thymidine starvation and amino acid
323 limitation influences whether DHFR inhibition by the antibiotic trimethoprim results in growth
324 inhibition or cell death[42,43]. Thus, these combined genetic and environmental perturbations serve
325 as an excellent test of our model's ability to leverage pairwise couplings to predict high-order
326 phenotypes.
327 To build the foundation of our gene-by-environment model, we first quantified single
328 perturbation-growth curves by measuring the effect of media supplementation on bacterial growth
329 rate in the absence of expression perturbations as well as the growth effect of expression
330 perturbations in the absence of media supplementation. We measured growth rates for *E. coli*

331 harboring a non-targeting CRISPRi control in a titrated range of thymidine and methionine
332 concentrations. For wildtype *E. coli* in the absence of gene repression, the addition of either
333 supplement to growth media slightly accelerated growth (Fig S7A-B, Table S4). These data were
334 modeled with a 4-parameter sigmoid (using additional parameters g_{min} and g_{max} to represent the
335 minimum and maximum growth rate effects, respectively) to quantify the growth rate effect of each
336 environmental shift. We also refit repression-growth rate functions for *folA* and *thyA* using
337 repression-growth rate data gathered in this experiment and bootstrapped all single perturbation-
338 growth functions to estimate uncertainty in model performance and coupling strengths (Fig S7C-D,
339 Table S5). To investigate gene-by-environment interactions within folate metabolism, we created a
340 compact mismatch CRISPRi library containing 110 pairwise sgRNA combinations targeting *folA*
341 and *thyA*. We transformed this library into our CRISPRi strain and grew the transformed population
342 in a continuous culture device that allowed for parallel control of multiple supplemented media
343 conditions[37,44]. We again quantified relative growth rates following CRISPRi by next-generation
344 sequencing and rescaled these measurements such that the growth rate of *E. coli* harboring the
345 non-targeting CRISPRi control in unsupplemented media was equal to one. We then identified and
346 removed escapers (125 measurements, 0.9% of data) and averaged across the remaining growth
347 rate replicates, yielding 2,394 unique growth rate measurements across 22 growth conditions (Fig
348 6B). This generated a rich dataset to train and evaluate the continuous epistasis model on gene-
349 by-environment interactions.

350 We isolated our pairwise (second-order) perturbation data and fit pairwise couplings
351 between perturbations (Table S5). We found that our model effectively recapitulated growth rates
352 for this independent data set, with an epistatic model RMSD of 0.145, [0.144, 0.176] and a Null
353 RMSD of 0.273, [0.260, 0.293]. As with our gene-gene model, subsampling as little as 20% of our
354 gene-by-environment data was sufficient to reliably fit the continuous epistasis model (RMSD =
355 0.151, [0.147, 0.194]). We performed subsampling using the same regularization strength used in
356 our pairwise study, supporting the generalizability of this mild regularization. Thus, we see that our
357 unit-agnostic model can capture not only gene-gene interactions, but also gene-by-environment
358 and environment-by-environment interactions.

359 Next, we evaluated the ability of our model – trained exclusively on first- and second-order
360 perturbations – to predict the growth rate effect of 2,035 third- and fourth-order perturbations. The
361 continuous epistasis model outperformed the Null model on these higher-order predictions:
362 including pairwise couplings yielded a prediction RMSD of 0.275, [0.252, 0.351] between the model
363 and the data, while Null model predictions had an RMSD of 0.584, [0.555, 0.612]. High-order
364 predictions remained robust when subsampling pairwise training data (RMSD = 0.300, [0.237,
365 0.412]). Combining all perturbations, the continuous epistasis model significantly outperformed the
366 Null model on gene-by-environment predictions in terms of RMSD (continuous epistasis: 0.260,
367 [0.239, 0.330], Null: 0.547, [0.521, 0.575]) and AIC (continuous epistasis: -6,416, [-6,817, -5,278],
368 Null: -2,868, [-3,108, -2,637] (Fig 6C-D). We again compared our model to the Isserlis and
369 regression models, and all models performed similarly (Isserlis RMSD = 0.300, regression RMSD
370 = 0.254). However, only the continuous epistasis model yielded summarized coupling values
371 between genes and environments in addition to growth rate predictions.
372 We can visualize why the continuous epistasis model so drastically outperformed the Null
373 model using select environmental case-studies (Fig 6E). The top left matrix (red outline) contains
374 growth rate measurements following CRISPRi knockdown in unsupplemented M9-glucose media,
375 and as expected, knockdown of either gene is deleterious to growth. Both the continuous epistasis
376 and Null models accurately recapitulated this experimental data as there is not strong pairwise
377 epistasis between *folA* and *thyA* in this environment. However, upon addition of thymidine into the
378 media (bottom left, green outline), the models' predictions diverge. The addition of thymidine to the
379 media decouples *thyA* expression from growth rate. We observe this phenotype in the first few rows
380 of these matrices, where thymidine knockdown does not cause a growth defect. This phenotype is
381 captured by the continuous epistasis model, but the Null model's predictions do not change
382 following nutrient supplementation except to uniformly increase growth rate in the richer media
383 condition. We also note that a phenotype present in our experimental data is missing from both the
384 continuous epistasis and Null models – when thymidine is present in the media and *folA* is strongly
385 repressed, *thyA* repression actually *benefits* growth. As mentioned above, this third-order gene-by-
386 gene-by-environment interaction was expected and is not captured by our pairwise model (Fig 6C-

387 D, orange points). This highlights the need for careful evaluation of model performance, as a
388 systematic failure in a pairwise model may be indicative of biologically interesting high-order
389 interactions. Lastly, we observed a “stripe” of slightly elevated growth in the last row of our pairwise
390 expression-growth rate matrices. This apparent nonmonotonicity likely arises because matrix rows
391 are ordered by sgRNA repression strength, and the strongest knockdowns (bottommost rows) are
392 nearly indistinguishable by qPCR and thus may be misordered. This sensitivity highlights the
393 importance of the bootstrapping approach we used when fitting single perturbation-growth
394 functions, as we know the model’s overall performance is generally insensitive to these variations
395 at maximal knockdown.

396 To quantitatively, rather than qualitatively, investigate where our model’s accuracy gains
397 emerged, we separated our data by genetic and environmental perturbation types and plotted each
398 model’s error separately for each perturbation subset (Fig 6F). The continuous epistasis model
399 outperformed the Null model the most when *thyA* was repressed and thymidine was added to the
400 media. As noted previously, thymidine replaces the key metabolite made by the *thyA* gene product,
401 rescuing growth and resulting in a positive pairwise coupling (a_{ij} = 0, [0, 0], a_{ji} = 0.73, [0.20, 1.29]).
402 This asymmetric coupling is indicative of a hierarchy between thymidine and *thyA* expression,
403 where the addition of thymidine to growth media fully decouples growth rate from *thyA* expression
404 but not vice versa. These results indicate that the continuous epistasis model can be applied
405 beyond low-order perturbation studies to predict the combined effects of complex gene-by-
406 environment and environment-by-environment interactions. This now opens the door to creating
407 continuous, coupling sensitive pharmacogenomic models that describe the interaction between
408 drug, genotype, and environment.

409

410 **Discussion**

411 In this work, we generated and trained a continuous, coupling-sensitive model relating gene
412 expression to growth rate using sparsely-sampled single and pairwise perturbation measurements.
413 The model’s parameters are interpretable and relevant across gene expression levels, and this
414 unitless approach enables the investigation of context-dependence between disparate

415 perturbations, such as CRISPRi knockdown and metabolite supplementation. We observed that
416 gene-gene epistasis, when inferred across a continuous expression landscape, is less common
417 than epistasis inferred at the level of knockout or extreme knockdown. In addition, while it is
418 possible to infer why some gene pairs show positive or negative coupling based on prior
419 knowledge, we see that biochemical pathway structure is not sufficient to predict all gene-gene
420 interactions, and it does not seemingly offer perspective on coupling strength. Certainly, the
421 relationship between coupling strength, asymmetry, and functional interactions between enzymes
422 will be an interesting avenue for further study.

423 Based on our results, we propose a hybrid experimental/modeling approach to streamline
424 epistatic measurements, generate quantitative insight into the strength and directionality of
425 epistasis, and predict the outcome of perturbations not explored within a given study. First,
426 precisely quantify the effect of each single perturbation at varied intensities (Fig 7A). This can be
427 accomplished for gene knockdown measurements using mismatch CRISPRi or for drug treatment
428 by sampling around a known IC_{50}. Then, sparsely sample pairwise perturbation landscapes (exact
429 sampling density may vary; this study used 10s of measurements per combination) (Fig 7B). Using
430 the continuous epistasis model, fit pairwise couplings to quantify the magnitude and directionality
431 of pairwise epistasis between all perturbation combinations (Fig 7C). Finally, utilize these epistatic
432 couplings and single perturbation fits to continuously model each pairwise perturbation landscape
433 and predict the effect of high-order perturbation combinations (Fig 7D). For reference, if one wished
434 to perform an experiment with the same sampling density (20 measurements per pair), replicates
435 (6), and sequencing depth (≥70x coverage) as the study discussed here – but for 1,000 genes of
436 interest (totaling ~1 million epistatic interactions and half a million pairwise landscapes) – the
437 sequencing data required would fit on a single NovaSeq X 25B flow cell.

438 While we limited our experiments to transcriptional and environmental perturbations in this
439 work, one can imagine integrating proteomic, metabolomic, and other measurements into this
440 unitless modeling framework. We modeled the growth rate effect of each perturbation with a
441 monotonic function, however there are known examples of nonmonotonic growth rate relationships,
442 limiting some applications of the model as it is currently constructed[10]. Finally, while we showed

443 that accounting for pairwise coupling alone captured most variation in our third- and fourth-order
444 perturbation landscapes, sufficiently strong high-order epistasis in a system could require a
445 different approach.

446 To expand this workflow, the limiting experimental step is quantifying the repression
447 strength of individual sgRNAs. RT-qPCR is relatively low-throughput, precluding genome-level
448 studies. An alternative, single-cell approach to quantify CRISPRi repression, such as microSPLiT[45]
449 (bacterial cells), PETRI-seq[46] (bacterial cells) or perturb-seq[47] (mammalian cells) alongside next-
450 generation sequencing-based growth rate measurements would remove these throughput
451 restrictions and permit significantly larger CRISPRi studies. Alternatively, one could utilize a
452 predictive model relating sgRNA homology sequence to CRISPRi repression strength[11,32].

453 Applications of this continuous, epistatic model are numerous. The reduced sampling
454 requirements make pathway-level epistatic analyses feasible in single experiments. Prior work has
455 indicated that metabolic enzymes are often produced in evolutionarily conserved stoichiometric
456 ratios[48], and the experiments presented here could be used to better understand the constraints
457 on enzyme abundance in metabolic systems that shape these ratios. The continuous epistasis
458 model could also be applied to optimize biosynthetic pathway performance and/or cell growth
459 conditions. We hope that interpretable, continuous, and coupling-sensitive models like the one
460 described in this work can continue to improve computational predictions to extract more
461 information from the data-rich -omics experiments we now regularly perform.

462

463 **Acknowledgments**

464 The authors would like to thank Olivier Rivoire for insightful discussions and data analysis ideas.
465 We are grateful to Prashant Mishra for thoughtful comments on our manuscript. We thank Andrew
466 Mathis for feedback in the early stages of this work. We also thank the Reynolds lab for feedback
467 on the project and manuscript. Research reported in this publication was supported by the National
468 Institute of General Medical Sciences of the National Institutes of Health under award number
469 R01GM136842. R.O. was also partly supported by a National Institutes of Health training grant

5T32GM007062-46. This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

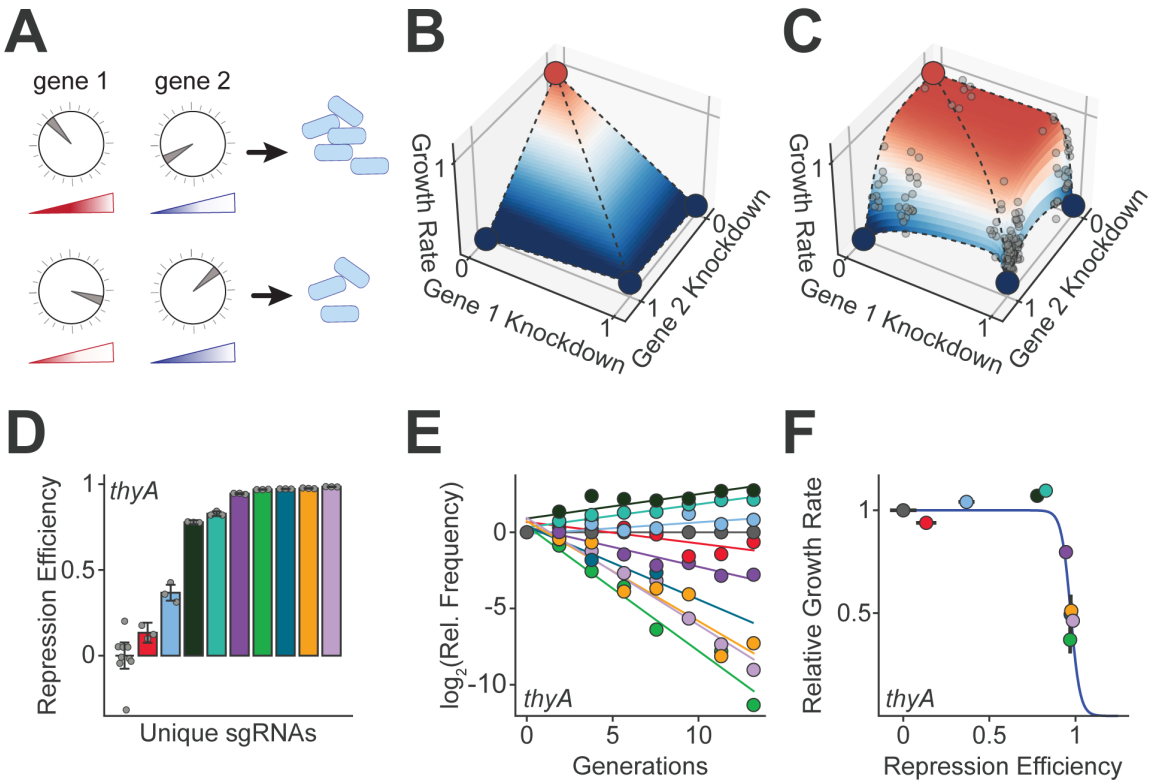
Author Contributions

Conceptualization, R.M.O. and K.A.R.; Formal Analysis – R.M.O., A.T.N., P.M.B., and K.A.R.; Investigation – R.M.O., A.T.N., and P.M.B.; Writing – Original Draft, R.M.O., P.M.B., and K.A.R.; Writing – Review & Editing, R.M.O., A.T.N., P.M.B., and K.A.R.; Supervision, K.A.R.; Funding Acquisition, K.A.R.

Declaration of Interests

The authors declare no competing interests.

483 **Main Text Figures**



484
485 **Figure 1. Mismatch CRISPRi enables characterization of gene expression-growth rate**
486 **relationships**. (A) Tuning the expression of *E. coli* genes (dials) and thus relative protein
487 abundances (gradients) alters cellular growth rate. (B) A linear interpolation from experimental data
488 exploring the growth rate effects of complete knockdown of one or two genes, akin to a classic
489 genetic interaction measurement. Growth rate of wildtype cells (x = y = 0, red) linearly decreases
490 with either or both knockdowns of genes 1 and 2 (dotted lines). (C) A computationally modeled,
491 continuous, pairwise expression-growth rate model, fit from intermediate experimental data (gray
492 dots). Growth rate no longer scales linearly with knockdown and is far more robust to expression
493 perturbation than a linear model would suggest. (D) The concentration of mRNA in log-phase *E.*
494 *coli* for each titrating sgRNA was quantified by RT-qPCR. A repression efficiency of 1 corresponds
495 to complete knockdown. Individual colored bars represent different titrating sgRNAs targeting *thyA,*
496 with error bars describing the standard error of the mean (SEM) across technical replicates.
497 Individual measurements are shown as gray dots. (E) Growth rate was quantified by next-
498 generation sequencing. We measured the $\log_2$ relative (Rel.) frequency of each sgRNA relative to
499 a non-targeting control (gray dots along y = 0) over eight time points spanning 14 hours. Color
500 coding of sgRNAs is identical to panel D. The slope of the line of best fit represents the growth rate
501 of each knockdown relative to the non-targeting control. (F) The measurements of (rescaled)
502 relative growth rate and repression efficiency show a sigmoidal relationship (blue line), with
503 significant growth rate deficits emerging at severe repression levels. Color coding of sgRNAs is
504 identical to panel D; individual points reflect average relative growth rate across n ≥ 4 replicates,
505 and error bars represent SEM of growth rate measurements (vertical) and RT-qPCR
506 measurements (horizontal).

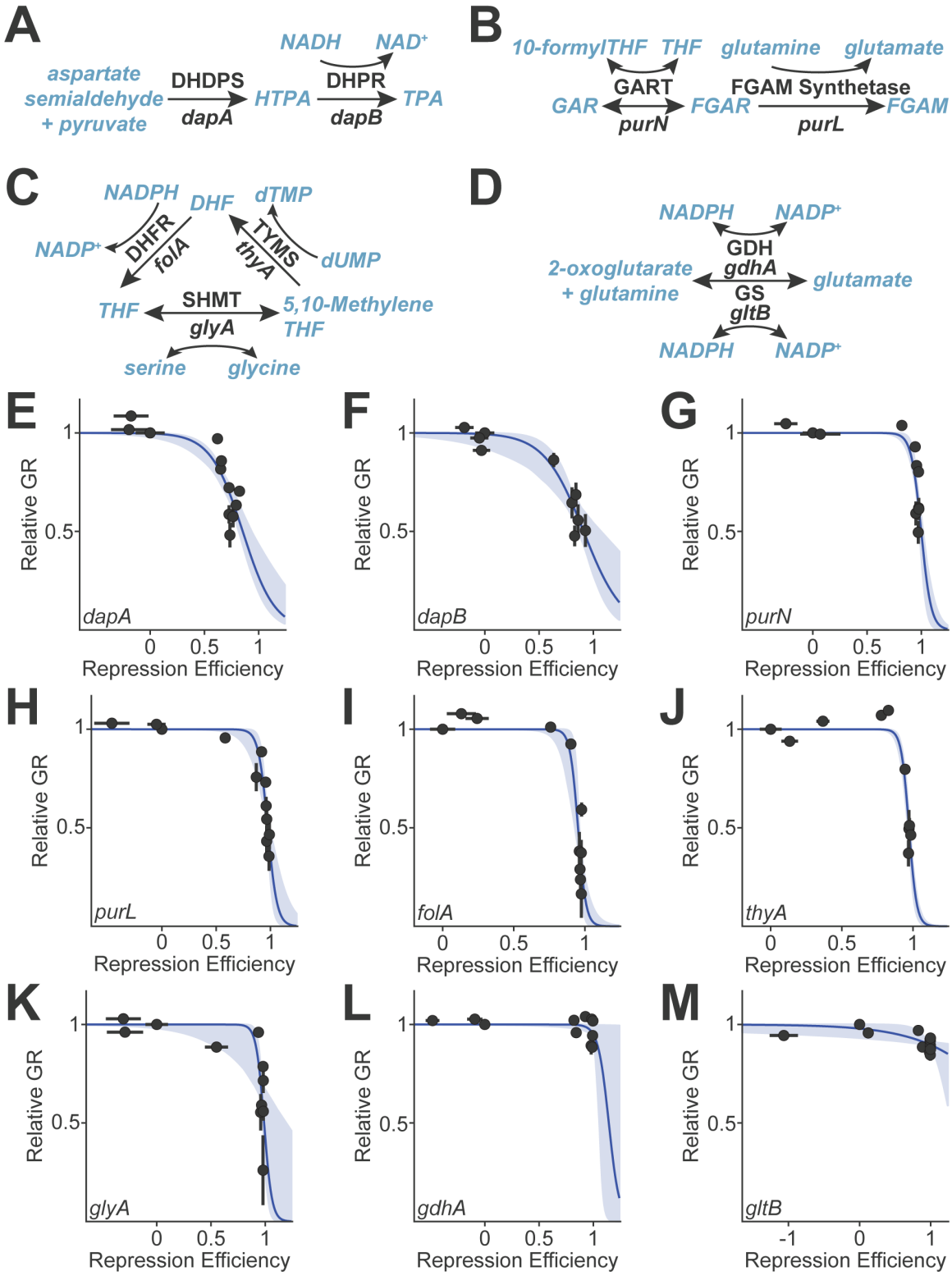


Figure 2. Mismatch CRISPRi for a diverse set of metabolic enzymes. sgRNAs were designed to target 9 *E. coli* genes from (A) lysine, (B) purine, (C) folate, and (D) glutamate metabolism. Gene names are listed in black italic text underneath the corresponding metabolic enzyme (bold black text). Metabolite names are in blue. (E-M) Repression efficiency of each CRISPRi sgRNA, as

512 measured by RT-qPCR, correlated with relative growth rate (GR), as measured by next-generation
513 sequencing. Error bars represent SEM of growth rate measurements between n ≥ 4 replicates
514 (vertical) and RT-qPCR measurements across n ≥ 3 technical replicates (horizontal). Blue lines
515 represent the two-parameter logistic fit to each data set. Blue shaded area represents a pointwise
516 95% confidence interval of the logistic fit estimated by bootstrapping. The genes *gdhA* and *gltB* are
517 nonessential in our experimental conditions and show no significant growth rate defect following
518 repression.

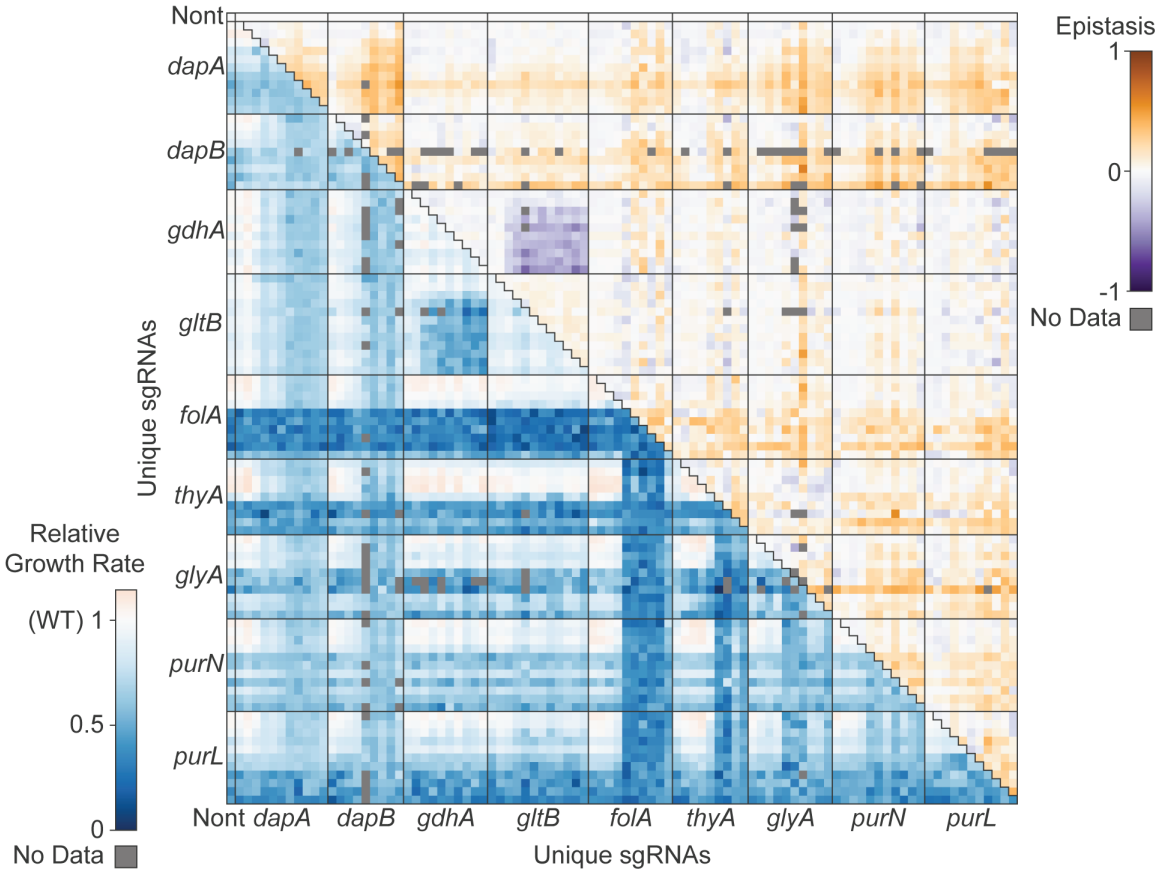


Figure 3. Pairwise CRISPRi growth rate and epistasis measurements. Each column and row represents a unique sgRNA perturbation. Gene names denote groups of sgRNAs targeting a given gene, and sgRNAs are sorted within each group by increasing CRISPRi repression strength (top-to-bottom and left-to-right). Nont indicates the non-targeting control sgRNA. The lower triangle of the matrix describes pairwise relative growth rate measurements calculated relative to the non-targeting control across 14 hours, averaged across n ≥ 4 experimental replicates, where wildtype-like (WT) growth is equal to 1. The upper triangle of the matrix describes pairwise growth rate epistasis. Epistasis was calculated as the difference between the observed growth rate and the multiplicative (Bliss) growth rate expectation of pairwise sgRNA perturbations.

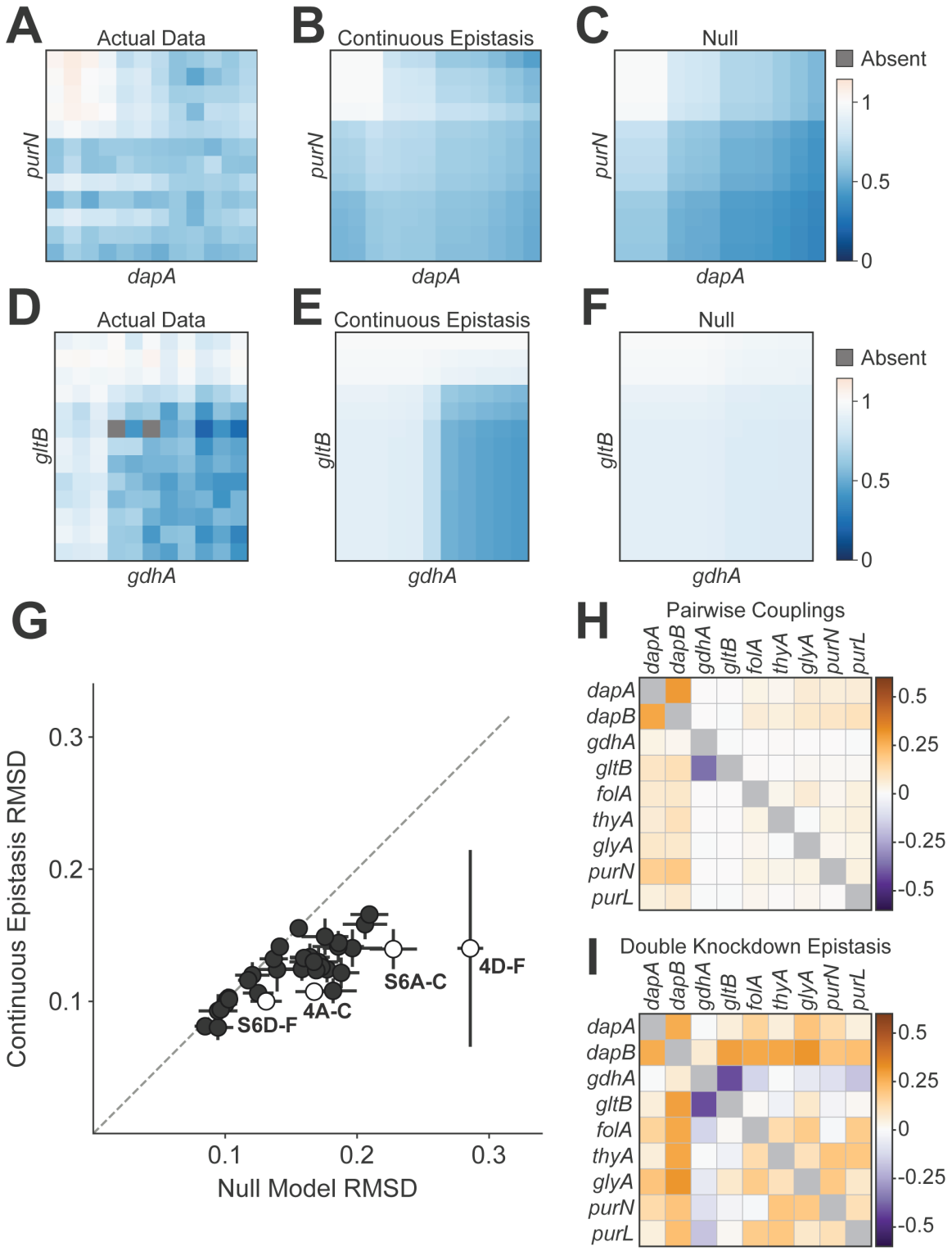


Figure 4. The continuous epistasis model captures gene-gene coupling at all expression levels. (A) Pairwise expression-growth rate data following CRISPRi knockdown of both *dapA* and *purN*, which show positive coupling. Each row and column represents a unique sgRNA, and pixels represent the growth rate effect of a given sgRNA pair. Rows and columns are sorted by knockdown intensity, ranging from wildtype-like expression (top left) to maximal double-

535 knockdown (bottom right). (B) Predicted expression-growth rate data utilizing the coupling-
536 sensitive continuous epistasis model. (C) Predicted expression-growth rate data utilizing the
537 coupling-insensitive Null model. (D-F) Same as A-C, but for the *gdhA*/*gltB* gene pair, which shows
538 negative coupling and synthetic lethality. (G) The performance of the coupling-sensitive model
539 compared to the Null model. Each data point represents model RMSD across all pairwise sgRNA
540 combinations for a gene pair. Full expression-growth rate data is shown for white, annotated dots
541 in this figure and Fig S3. Error bars are standard deviations of models fit on bootstrapped single
542 perturbation-growth curves. Dotted gray line is y = x. (H) Coupling constants fit across all pairwise
543 growth rate data, as calculated from the continuous epistasis model. (I) Growth rate epistasis of
544 the most severe double knockdown for each gene pair, calculated as the difference between the
545 experimentally determined growth rate and the product of each single-gene knockdown growth
546 rate.

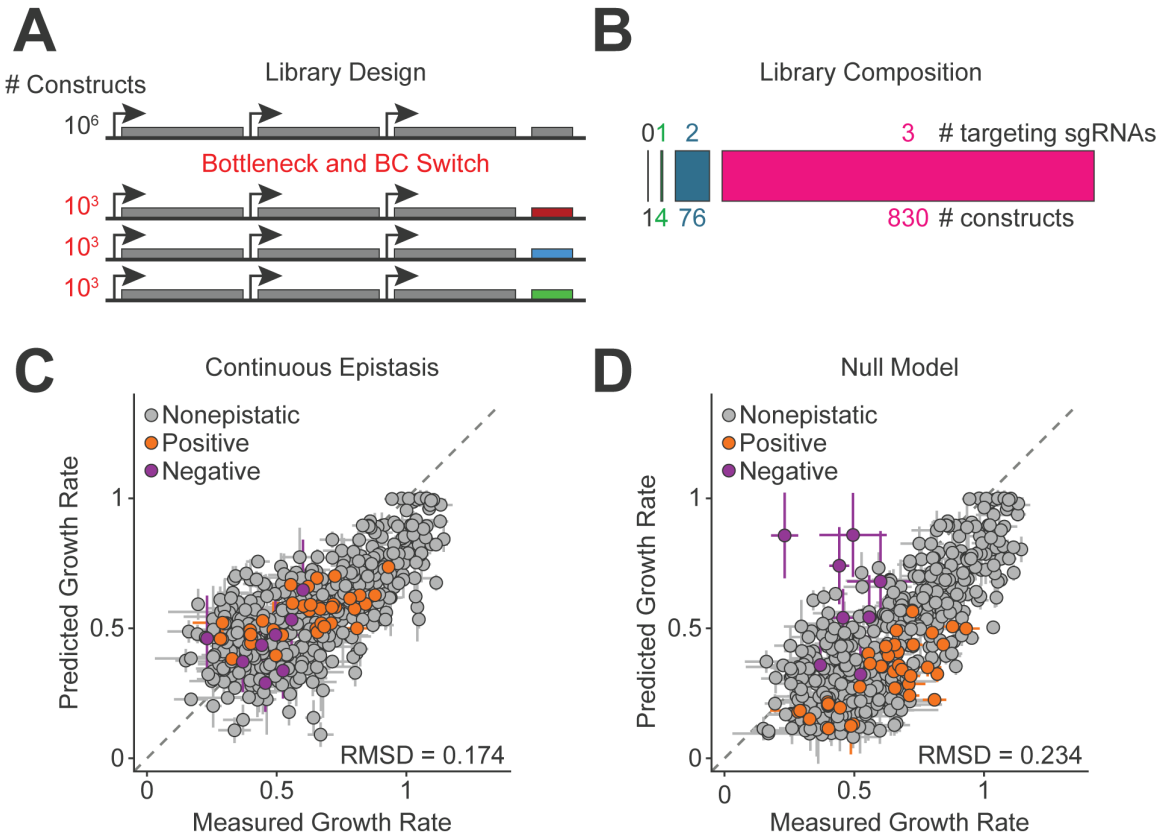


Figure 5. A model trained on pairwise measurements predicts growth rate following third-order gene knockdowns. (A) Diagram depicting the design of the third-order sgRNA library. After generating a diverse three-sgRNA library with >10^6 possible constructs, we bottlenecked our library to a diversity of ~10^3 constructs. We then used inverse PCR to generate five copies of the library with identical sgRNA constructs but distinct barcodes (BC). (B) Distribution of targeting sgRNAs in the library. The library contained zero- through third-order targeting sgRNA combinations, with bars scaled to illustrate each order's representation in the library. (C-D) Model performance on third-order sgRNA combinations, where experimentally determined growth rates are plotted against predicted growth rates for the (C) continuous epistasis and (D) Null model. Data is colored based on the presence of pairwise positive or negative epistasis (orange and purple points, respectively) within a construct. For a construct to be considered epistatic, two sgRNAs must target genes with a magnitude of epistasis (a_{ij} or a_{ji}) > 0.15 with fewer than 9 sgRNA mismatches. Dotted gray line is y = x.

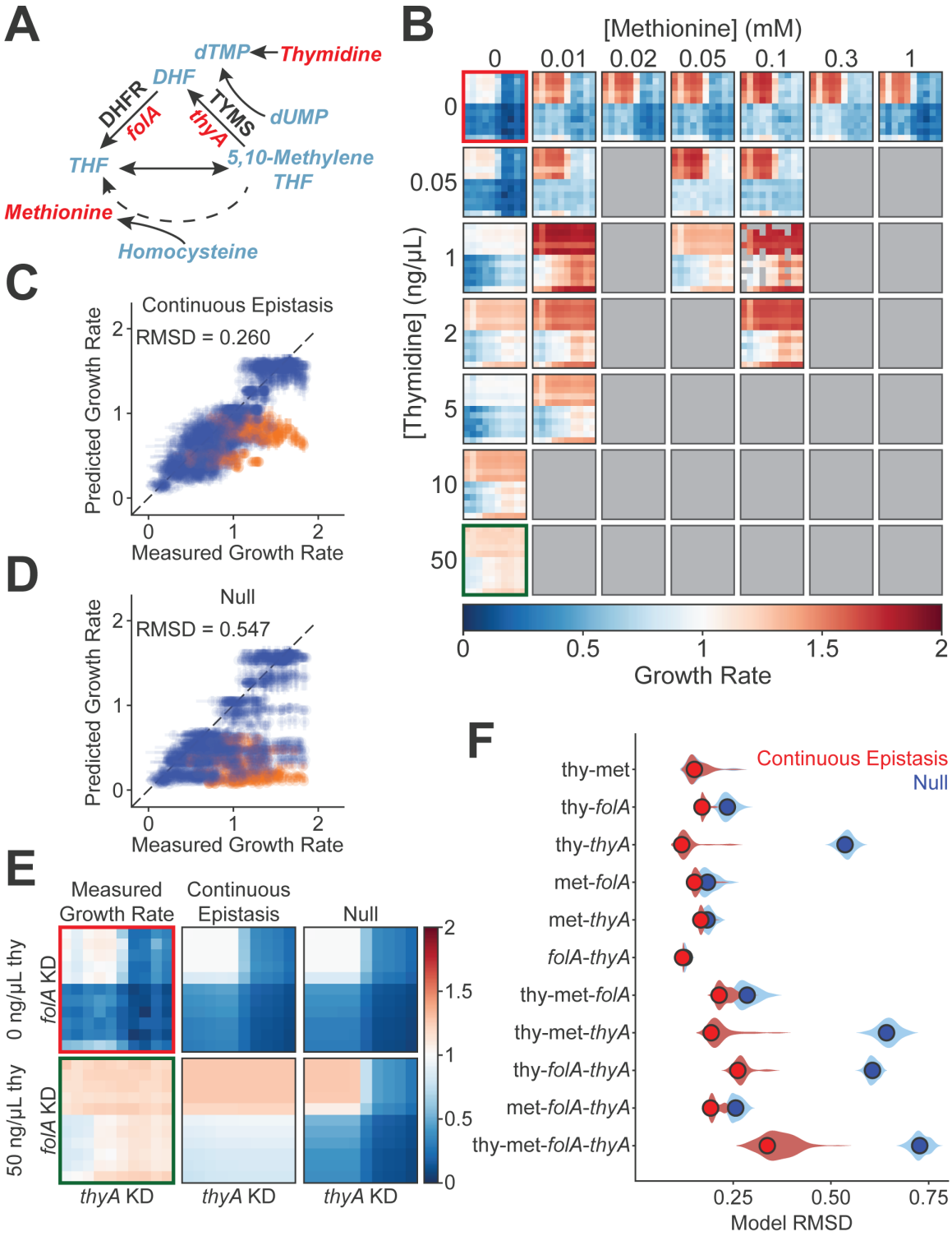


Figure 6. The continuous epistasis model captures gene-by-environment interactions. (A) Simplified schematic of a portion of *E. coli* folate metabolism. Gene names are listed in italics under each arrow, and enzyme abbreviations are listed in bold, black text above each arrow. Red text indicates a gene or metabolite perturbed over the course of the experiment. (B) Gene-by-gene-by-environment-by-environment relative growth rates, normalized to the growth rate of unsupplemented *E. coli* harboring a non-targeting control sgRNA. Within each heatmap, rows

568 represent distinct *folA*-targeting sgRNAs, and columns represent distinct *thyA*-targeting sgRNAs.
569 Thymidine and methionine supplementation varies between heat maps, and gray boxes represent
570 pairwise media combinations not explored in this study. (C-D) Growth rate predictions for all data
571 generated using the continuous epistasis (C) and Null model (D). Dotted line is y = x. The combined
572 effects of *folA* knockdown, *thyA* knockdown, and thymidine supplementation result in a third-order
573 interaction (orange points). (E) Expression-growth rate data and predictions following pairwise
574 *folA*/*thyA* knockdown in unsupplemented media (top row) and media supplemented with 50 ng/μL
575 thymidine (bottom row). (F) Continuous epistasis model performance, split by perturbation
576 combination. For each second- and higher-order combination of expression and environmental
577 perturbations (y-axis labels), the RMSD of growth rate measurements for the continuous epistasis
578 model (red) and Null model (blue) are indicated. Violin plots represent model performance over 100
579 bootstrapped replicates.

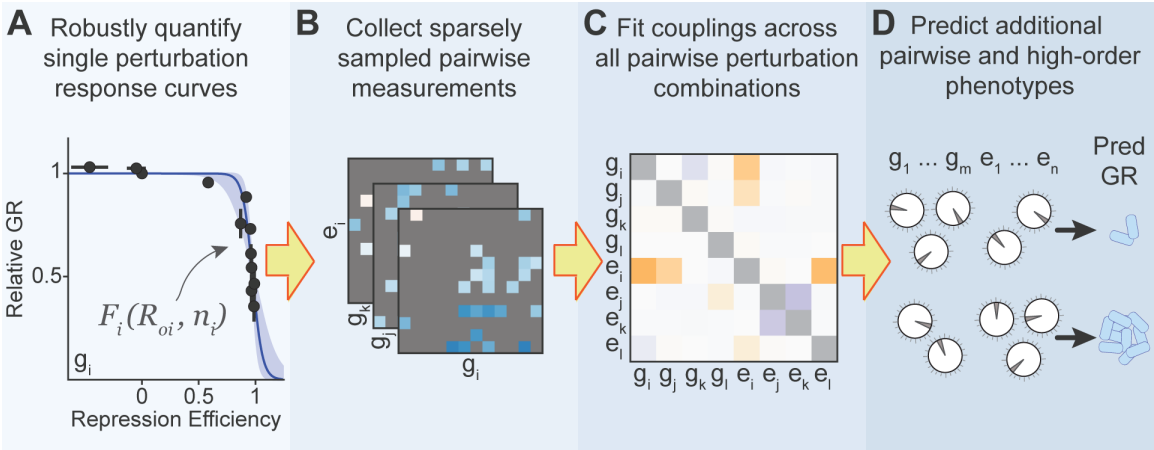


Figure 7. An overview of epistasis studies using the continuous epistasis model. (A) Densely sample single perturbation response functions using individual measurements, quantifying R_0 and n for each individual perturbation. (B) Collect sparse pairwise measurements between all individual perturbations, randomly sampling the pairwise landscape. g_i: gene i; e_i: environment i. (C) Use the single perturbation response functions and sparse pairwise data to fit pairwise couplings between all perturbation combinations. (D) Use the continuous epistasis model to model continuous single, pairwise, and higher-order landscapes and predict phenotypes for unmeasured perturbation combinations.

589 **Methods**

590 **Pairwise CRISPRi Library Construction**

591 We designed a pool of CRISPRi sgRNAs targeting nine *E. coli* K-12 MG1655 genes. For each
592 gene, we selected 9-12 distinct sgRNAs to provide a titrated range of mRNA knockdown. We
593 previously characterized CRISPRi knockdown of seven of these genes[33] (*dapA*, *dapB*, *folA*, *thyA*,
594 *glyA*, *purN*, and *purL*), giving us an initial list of sgRNAs for each gene with mutually distinguishable
595 growth rate effects. We supplemented these sgRNAs with up to two fully on-target sgRNAs per
596 gene and added sgRNAs by hand to populate underrepresented growth regimes. The final two
597 genes (*gdhA* and *gltB*) were previously uncharacterized by our lab. We designed two on-target
598 sgRNAs for each gene using previously published design principles[30]. In addition, we introduced 4,
599 6, 8, 10, and 14 mismatches into the homology region of each on-target sgRNA[33]. For a full list of
600 sgRNA sequences, see Table S1.

601 All sgRNAs were synthesized as a mixed oligo pool by Integrated DNA Technologies. We
602 resuspended the library to a final concentration of 50 nM in MB H_2O and added BsaI flanking sites
603 using PCR. We performed two separate PCRs to add two distinct sets of flanking sites onto the
604 library for Golden Gate cloning. We used the GG_UnivF/GG_R1 primers to generate the insert 1
605 library and the GG_F2/GG_UnivR primers to generate the insert 2 library (for primer sequences,
606 see Table S6). The PCR mix contained 10.75 µL MB H_2O, 5 µL Q5 Reaction Buffer (NEB), 5 µL
607 High GC Enhancer (NEB), 0.5 µL 10 mM dNTPs, 1.25 µL 10 µM forward primer, 1.25 µL 10 µM
608 reverse primer, 0.25 µL Q5 High-Fidelity DNA Polymerase (NEB, #M0491), and 1 µL of the 50 nM
609 oligo pool. Initial denaturation was carried out at 98°C for 30 seconds, followed by 25 cycles of
610 98°C for 30 seconds, 58°C for 10 seconds, and 72°C for 20 seconds. A final elongation was
611 performed at 72°C for 60 seconds, followed by an indefinite hold at 4°C. We ran the PCR product
612 on a 1% agarose 1X TAE gel with EtBr staining at 80 V for 45 minutes and excised the 176 bp
613 band. We extracted the band using a DNA Clean & Concentrator-5 Kit (Zymo, #D4014), eluted in
614 30 µL MB H_2O, and quantified dsDNA concentration using a PicoGreen assay (Thermo, #P7581).
615 To ensure representation of the non-targeting sgRNA sequence (Nont) in the final library, we
616 isolated this control sgRNA using the same protocol illustrated above, replacing the 50 nM oligo

617 pool with 1 μL of purified 20-50 ng/μL plasmid containing the Nont sgRNA[33] (pCRISPR2). We
618 spiked Nont inserts into each of the mixed libraries at a 1:100 Nont:library ratio.

619 We performed three Golden Gate cloning reactions to ligate the CRISPRi inserts into three
620 distinct barcoded CRISPRi vectors (pCRISPR3, AddGene #: 191856-191861). The reaction mix
621 contained 75 ng of barcoded vector, 6 ng of insert 1 mix, 6 ng of insert 2 mix, 1.5 μL of 10X T4
622 DNA Ligase Buffer (NEB), 1 μL T4 DNA Ligase (2,000,000 U/mL, NEB, #M0202M), 1 μL BsaI-HF
623 v2 (NEB, #R3733), and MB H_2O to 15 μL. Reaction conditions were 40 cycles of 37°C for 2 minutes
624 and 16°C for 3 minutes. These cycles were followed by 50°C for 10 minutes, 80°C for 20 minutes,
625 and an indefinite 4°C hold. A no-insert control, where both inserts were left out of the reaction, was
626 performed identically.

627 We cleaned and concentrated each reaction using a DNA Clean & Concentrator-5 Kit
628 (Zymo, #D4014), eluting in 20 μL MB H_2O. We transformed 2 μL of each purified library into 100
629 μL of XL1 Blue electrocompetent cells followed by a 1-hour recovery at 37°C with shaking. We
630 plated 50 μL of a 1:10 and 1:100 dilution to quantify library coverage and added the remaining cells
631 to 9 mL SOB + 35 μg/mL kanamycin. We incubated this culture overnight at 37°C with shaking. We
632 also transformed and plated the no-insert control and a no DNA control and adjusted library
633 coverage estimates accordingly. Library coverage ranged from 10-14x for each barcoded vector.
634 In addition to these mixed CRISPRi libraries, we synthesized constructs containing two Nont
635 sgRNAs in six distinct barcoded vectors using the same protocol described above, substituting
636 each of the mixed insert libraries for 6 ng of purified Nont sgRNA insert.

637 We isolated plasmids using the GeneJET Plasmid Miniprep Kit (Thermo, #K0503), eluted
638 DNA in 50 μL MB H_2O, and quantified dsDNA concentration with a PicoGreen assay (Thermo,
639 #P7581). We combined the three barcoded CRISPRi libraries in an equimolar ratio. We also mixed
640 the Nont+Nont controls in an equimolar ratio, then spiked this control into the library at a 1:5,000
641 Nont+Nont:library ratio. We transformed 3 ng of this final mixture into 45 μL MG1655 +
642 chromosomal dCas9 electrocompetent cells (gift from the Bikard lab) using the same
643 transformation protocol as above, plating 50 μL of 1:100 and 1:1,000 dilutions. We performed a no

644 DNA control with zero breakthrough colonies. Coverage exceeded 200x, and we grew the
645 remaining cells in 9 mL SOB + 35 µg/mL kanamycin overnight at 37°C with shaking.
646 The pairwise *folA*/*thyA* CRISPRi library was constructed as described above, with only
647 Nont, *folA*-targeting, and *thyA*-targeting sgRNAs included in the Integrated DNA Technologies-
648 synthesized oligo pool. Library coverage calculated after the first transformation event in the three
649 barcoded vectors ranged from 14x to 89x. We mixed three barcoded Nont+Nont CRISPRi
650 constructs in equimolar ratios and spiked them into the *folA/thyA* library at a 1:400
651 Nont+Nont:library ratio. As two continuous culture experiments were required to investigate all
652 environments of interest, we freshly transformed the library two separate times. For each
653 transformation, we added 1 ng of mixed library to 45 µL MG1655 + dCas9 electrocompetent cells
654 and achieved 583x and 127x library coverage.
655
656 **Continuous culture**
657 Following transformation of the MG1655 dCas9 electrocompetent *E. coli*, cells were grown
658 overnight in LB + 35 µg/mL kanamycin. We spun saturated outgrowth cultures at 1,000 rcf for 5
659 minutes, decanted media, and resuspended cells in 1 mL M9 minimal media pH 6.5, 0.4% glucose,
660 35 µg/mL kanamycin (M9+Kan). We repeated this wash, then resuspended 10 µL of concentrated
661 cells in 8 mL M9+Kan. We incubated this culture for 12 hours at 37°C with shaking.
662 Following outgrowth, we spun the culture at 1,000 rcf for 5 minutes, decanted, and
663 resuspended cells in 1 mL fresh M9+Kan. We repeated this wash, diluted the culture to OD_{600} =
664 0.05 in 15 mL M9+Kan, and grew these cells in a continuous culture device at 37°C, clamping the
665 optical density at 0.15. Our continuous culture device follows the design specifications from Toprak
666 et al., 2012 and 2013[37,44]. Following an overnight adaptation, we added anhydrotetracycline
667 (Cayman Chemical Company, #10009542) to a final concentration of 50 ng/mL to induce CRISPRi.
668 We took the first timepoint (T_0) after 3 hours of CRISPRi by extracting 1 mL of cells from the culture
669 and took additional timepoints every 2 hours until T_{14}. We centrifuged timepoints at 3,000 rcf for 5
670 minutes to pellet cells, decanted by pipetting, and stored them at -20°C.

671 We performed continuous culture for the third-order CRISPRi and environmental variation
672 experiments with slight modifications. We terminated these experiments after T_{10}, as the final two
673 timepoints were shown to be unnecessary for growth rate quantification. We inoculated the third-
674 order CRISPRi library into the continuous culture device at a starting OD_{600} = 0.2. Media conditions
675 were altered during the *folA*/*thyA* gene-by-environment experiment. Thymidine (Sigma, #T1895)
676 and/or methionine (Sigma, #M5308) were supplemented into M9+Kan to create the following
677 conditions: No additive, 0.05 ng/μL thymidine, 1 ng/μL thymidine, 2 ng/μL thymidine, 5 ng/μL
678 thymidine, 10 ng/μL thymidine, 50 ng/μL thymidine, 0.01 mM methionine, 0.02 mM methionine,
679 0.05 mM methionine, 0.1 mM methionine, 0.3 mM methionine, 1 mM methionine, 0.05 ng/μL
680 thymidine + 0.01 mM methionine, 0.05 ng/μL thymidine + 0.05 mM methionine, 0.05 ng/μL
681 thymidine + 0.1 mM methionine, 1 ng/μL thymidine + 0.01 mM methionine, 1 ng/μL thymidine +
682 0.05 mM methionine, 1 ng/μL thymidine + 0.1 mM methionine, 2 ng/μL thymidine + 0.01 mM
683 methionine, 2 ng/μL thymidine + 0.1 mM, and 5 ng/μL thymidine + 0.01 mM methionine.
684
685 **Sample preparation and next-generation sequencing**
686 We resuspended each frozen cell pellet (each corresponding to a sample – one time point and
687 environmental condition) in 100 μL MB H_2O with vortexing and lysed them at 95°C for 3 minutes.
688 Following a 10 minute, 20,000 rcf centrifugation to pellet debris, we took 50 μL of the supernatant
689 as a working stock for use in downstream reactions.
690 We then added adaptors for Illumina sequencing in two consecutive PCR steps. First, we
691 added flanking sequences to each timepoint or environmental sample. The reaction mix was
692 composed of 10.5 μL MB H_2O, 5 μL Q5 Reaction Buffer (NEB), 5 μL 50% glycerol, 0.5 μL 10 mM
693 dNTPs, 1.25 μL 10 μM TruSeqF primer, 1.25 μL 10 μM TruSeqR_trunc primer, 0.5 μL Q5 High-
694 Fidelity DNA Polymerase (NEB, #M0491), and 1 μL timepoint working stock. Reaction conditions
695 were 98°C for 30 seconds, 7 cycles of 98°C for 30 seconds, 61°C for 10 seconds, and 72°C for 15
696 seconds, followed by a final elongation at 72°C for 45 seconds and an indefinite hold at 4°C.
697 In the second PCR, we added Illumina sequencing adapters and i5/i7 barcoding sequences to each
698 sample. The reaction mix contained 10.5 μL MB H_2O, 5 μL Q5 Reaction Buffer (NEB), 5 μL 50%

699 glycerol, 0.5 µL 10 mM dNTPs, 1.25 µL 10 µM i5 primer, 1.25 µL 10 µM i7 primer, 0.5 µL Q5 High-
700 Fidelity DNA Polymerase (NEB, #M0491), and 1 µL template from the previous reaction. Reaction
701 conditions were 98°C for 30 seconds, 20 cycles of 98°C for 30 seconds, 61°C for 10 seconds, and
702 72°C for 20 seconds, followed by a final elongation at 72°C for 60 seconds and an indefinite hold
703 at 4°C.
704 Note: While the glycerol was used in the preparation of timepoints for the pairwise CRISPRi library
705 experiment and the *folA*/*thyA* sublibrary experiment, we have observed future amplification
706 reactions that failed in the presence of glycerol.
707 We quantified dsDNA concentration from each reaction using a PicoGreen assay (Thermo,
708 #P7581), and mixed timepoints in an equimolar ratio to create the mixed sequencing library. We
709 ran the library on a 1% agarose 1X TAE gel with EtBr staining at 90 V for 60 minutes and excised
710 the 486 bp construct. We performed gel extraction using the DNA Clean & Concentrator-5 Kit
711 (Zymo, #D4014) and eluted in 20 µL MB H_2O. Following a Qubit Assay (Thermo, #Q32851) to
712 quantify final dsDNA quantification, we submitted the library to Azenta for Illumina HiSeq
713 Sequencing on a 300-cycle paired-end run.
714
715 **Isolation of single CRISPRi knockdown strains**
716 To quantify the concentration of mRNA following each individual CRISPRi perturbation, we isolated
717 plasmids expressing a single targeting sgRNA paired with the Nont sgRNA. For this purpose, we
718 constructed a CRISPRi:Nont library using Golden Gate cloning as described previously, adding the
719 Nont sgRNA insert in place of the mixed library in the second sgRNA position. We screened
720 individual colonies from this mixed library to isolate the majority of CRISPRi:Nont constructs.
721 We individually synthesized plasmids not isolated during this screen using inverse PCR
722 (iPCR). First, we phosphorylated iPCR primers with a PNK reaction. The reaction mix consisted of
723 6.5 µL MB H_2O, 1 µL 10X T4 DNA Ligase Buffer (NEB), 1 µL 100 µM iPCR_sgRNA_F primer, 1 µL
724 100 µM iPCR_UnivR primer, and 0.5 µL T4 PNK (NEB, #M0201S). Thermocycler conditions were
725 37°C for 30 minutes, 65°C for 20 minutes, and an indefinite 4°C hold. Then, we performed iPCR
726 with a reaction mix of 8.25 µL MB H_2O, 5 µL Q5 Reaction Buffer (NEB), 0.5 µL 10 mM dNTPs, 0.5

727 μL Q5 High-Fidelity DNA Polymerase (NEB, #M0491), 10 μL primer phosphorylation reaction, and
728 1 μL 30-50 ng/μL template plasmid (pCRISPR2). Reaction conditions consisted of 98°C for 30
729 seconds, followed by 25 cycles of 98°C for 30 seconds, 58°C for 30 seconds, and 72°C for 90
730 seconds. A final 3-minute elongation at 72°C preceded an indefinite hold at 4°C. We cleaned and
731 concentrated the reaction product using a DNA Clean & Concentrator-5 Kit (Zymo, #D4014) and
732 eluted in 25 μL MB H_2O. We removed the template plasmid using a DpnI digestion. Reaction
733 conditions were 13 μL cleaned and concentrated product, 2.5 μL 10x Cutsmart Buffer (NEB,
734 #B6004), 8.25 μL MB H_2O, and 1.25 μL DpnI (NEB, #R0176). The reaction was held at 37°C for 1
735 hour, then held at 4°C. Finally, we circularized the PCR product in a ligation reaction consisting of
736 8 μL DpnI product, 1 μL 10X T4 DNA Ligase Buffer (NEB), and 1 μL T4 DNA Ligase (400,000
737 U/mL, NEB, #M0202S). Ligation occurred at 25°C for 2 hours, followed by 10 minutes at 65°C and
738 a hold at 4°C. We transformed this product into XL1 Blue chemically competent cells (1 μL DNA,
739 45 μL competent cells) with 1 hour recovery at 37°C with shaking. We purified plasmids using a
740 GeneJET Plasmid Miniprep Kit (Thermo, #K0503) and transformed into CRISPRi MG1655 + dCas9
741 cells as described previously.
742
743 **RNA Extraction and RT-qPCR**
744 To collect RT-qPCR data following CRISPRi, we grew CRISPRi-Nont strains from glycerol stocks
745 in 4 mL LB + 35 μg/mL kanamycin for 6-8 hours at 37°C with shaking. We washed these cultures
746 into M9+Kan as described previously and incubated them for 8-14 hours at 37°C with shaking. We
747 then washed cultures into M9+Kan+50 ng/mL anhydrotetracycline (Cayman Chemical Company,
748 #10009542) and incubated them for 5 hours at 37°C with shaking at a starting OD_{600} = 0.05. After
749 treatment, we isolated RNA using the Qiagen RNeasy Protect Bacteria Mini Kit (QIAGEN, #74524),
750 with on-column DNase digestion (QIAGEN, #79254), as described in their protocol. We extracted
751 RNA from control Nont+Nont strains using the same procedure. We performed RT-qPCR on a CFX
752 Opus 384 Real-Time PCR System (Bio-Rad, #12011452) using the Luna Universal One-Step RT-
753 qPCR Kit (NEB, #E3005). No template and genomic DNA controls were included in all experiments.
754 We analyzed raw results using the CFX Maestro Software (Bio-Rad), and calculated changes in

755 mRNA concentration using the $\Delta\Delta C_t$ method with the *hcaT* gene as reference in a custom Python
756 script.
757
758 **Third-order sgRNA library construction and next-generation sequencing**
759 We constructed the third-order sgRNA library in a single barcoded vector (pCRISPR3, AddGene #:
760 191856-191861) using Golden Gate cloning as described previously with minor modifications.
761 Three sgRNA inserts were used instead of two, using primers GG_UnivF/GG_R1 for insert 1,
762 GG_F2/GG_R2 for inserts 2, and GG_F3/GG_UnivR for insert 3 (Table S6). During Golden Gate
763 cloning, insert concentrations were doubled to 12 ng per insert per reaction, and reaction conditions
764 were adjusted to 60 cycles of 2 minutes at 37°C and 3 minutes at 16°C. These cycles were followed
765 by 10 minutes at 50°C, 20 minutes at 80°C, and an indefinite 4°C hold. Transformed cells were
766 serially diluted and plated, and we selected a library dilution containing ~1,000 colony forming units
767 to form our bottlenecked library.
768 This library was used as a template in an iPCR reaction to replace the 6-nucleotide barcode with a
769 new barcode sequence while maintaining sgRNA diversity. iPCR was carried out as described
770 above using primers BC_iPCR_F_BC[1-6] and BC_iPCR_UnivR_v2. These libraries were
771 sequence verified and transformed at sufficient efficiency to avoid further bottlenecking. These
772 sublibraries were quantified by Qubit and combined in an equimolar ratio to generate the final third-
773 order CRISPRi library with five barcodes. pCRISPR3 barcode 4 was excluded from the reaction.
774 Finally, to ensure that we could normalize sgRNA constructs to a wildtype-like subpopulation,
775 CRISPRi-sensitive *E. coli* harboring a Nont+Nont+Nont construct (in all five barcodes) were spiked
776 into the final population at a 1:500 ratio before CRISPRi treatment.
777 We altered the sampled preparation protocol to accommodate the increased amplicon size
778 of the third-order sgRNA CRISPRi library. For the first PCR, we used a reverse primer with 4
779 additional base pairs of homology (TruSeqR). The reaction conditions were changed to 30 seconds
780 at 98°C, 17 cycles of 30 seconds at 98°C and 60 seconds at 76°C, followed by 180 seconds at
781 72°C and an indefinite 4°C hold. In the second PCR, we altered thermal cycler conditions to 30
782 seconds at 98°C, 10 cycles of 30 seconds at 98°C and 75 seconds at 72°C, followed by 180

783 seconds at 72°C and an indefinite 4°C hold. We quantified dsDNA from these reactions using a
784 Qubit Assay (Thermo, #Q32851), mixed them in an equimolar ratio, and concentrated DNA using
785 a DNA Clean & Concentrator-5 Kit (Zymo, #D4014) with a 30 µL MB H_2O elution. We ran this
786 concentrated library on a 2% agarose 1X TAE gel with an EtBr stain at 90 V for 85 minutes and
787 extracted the resulting 636 bp band using a DNA Clean & Concentrator-5 Kit (Zymo, #D4014) with
788 a 20 µL MB H_2O elution. We quantified dsDNA with a Qubit Assay (Thermo, #Q32851) and
789 sequenced in-house using two Illumina MiSeq Nano v2 500-cycle paired-end runs and one Illumina
790 MiSeq v2 500-cycle paired-end run.
791
792 **Plate reader growth rate assay**
793 We cultured *E. coli* strains harboring a non-targeting CRISPRi control overnight in 4 mL LB + 35
794 µg/mL kanamycin at 37°C with shaking. The next day, we washed these cultures twice into M9+Kan
795 as described previously. We diluted cultures 1:200 into each of the 22 media conditions described
796 in Continuous Culture and adapted cells for 4 hours at 37°C with shaking. Following adaptation,
797 we washed cells into 4 mL of their respective supplemented media + 50 ng/mL anhydrotetracycline
798 at OD_{600} = 0.005. Three 200 µL replicates of each culture were grown in a Synergy Neo2 Hybrid
799 Multi Mode Reader (Agilent, #BTNEO2) overnight at 37°C while taking regular OD_{600}
800 measurements. Growth rates were fit across replicates using a log-linear fit of exponential phase
801 growth and averaged across triplicate measurements.
802
803 **Calculating CRISPRi growth rates**
804 We wrote all analysis code in Python 3.9.12 using Jupyter Notebook (available on github at
805 https://github.com/reynoldsk/ContinuousEpistasis). We extracted sgRNA sequences from Illumina
806 FASTQ files using regular expressions, searching for a 20 bp region between flanking sequences
807 CTAGCTCTAAAAC and A. Up to 3 mismatches and no gaps were permitted in the flanking
808 sequences. The identified homology region was required to match a desired sequence from the
809 CRISPRi library. In addition, to increase confidence, we imposed a filter requiring two specific base
810 calls in each sequence to have a Q-score above 30. The filtered bases were the final “on-target”

811 and first “mismatch” base in the homology region, as this two-nucleotide combination is unique to
812 a single sgRNA within each sgRNA family. We identified the plasmid barcode similarly, with flanking
813 sequences GTACAGCGAGGCAAC and ACGGATCCCCAC, imposing a maximum of 3
814 mismatches and allowing no gaps. The barcode sequence needed to match a target barcode
815 exactly, and no Q-score filter was imposed.

816 For the three-sgRNA library, the following modifications were made. As sequencing quality
817 decreases for a 500-cycle kit, the sgRNA flanking sequences were extended to CTAGCTCTAAAAC
818 and ACTAGTATTATAC. These flanking sequences were permitted to have up to 10 mismatches
819 for sgRNAs 1 and 2, and 6 mismatches for sgRNA 3. Flanking sequences for the barcode were the
820 same as above and were permitted to have up to 6 mismatches. The barcode itself was required
821 to match a target exactly, however the cutoff to call a specific sgRNA homology region was lowered.
822 Sequences in position 1 were allowed to have as many as 2 mismatches, and sequences in position
823 2 were allowed to have as many as 10 mismatches. This permissive filter was required due to the
824 sharp decrease in sequencing confidence at the end of the sequencing run. However, if a
825 sequenced sgRNA ambiguously mapped to multiple library sequences with the same number of
826 mismatches, no identity was called. We performed next-generation sequencing analysis from
827 HiSeq data on a high-performance computing cluster, while all subsequent analysis was performed
828 locally.

829 We calculated relative growth rates following each CRISPRi treatment using the change in
830 sgRNA counts over time. At each timepoint (t), we normalized counts for a specific barcoded
831 CRISPRi construct (sgRNA) to the control construct (Nont) and our initial timepoint (t0) using
832 Equation 4.

$$Relative\ frequency_{sgRNA,t} = \frac{counts_{sgRNA,t}/counts_{Nont,t}}{counts_{sgRNA,t0}/counts_{Nont,t0}} \qquad (4)$$

833

834 We then $\log_2$-transformed this relative frequency prior to growth rate fitting. For the pairwise sgRNA
835 libraries, if fewer than 10 counts were identified for a construct at a given timepoint, no future
836 timepoints were considered for that construct. Growth rates were calculated for CRISPRi constructs
837 with sufficient counts over at least the first three timepoints. For pairwise and third-order CRISPRi

838 experiments, the time axis was rescaled from hours to generations based on the overall culture
839 generation time in each continuous culture vial. We then fit a line to our $\log_2$(Relative frequency)
840 vs. generation data using scipy.stats.linregress and defined the slope of the best fit line as the raw
841 growth rate. At this step, two sgRNAs (gdhA_1_42_B_MM14 and gdhA_3_216_B_MM8) were
842 determined to have off-target effects and were removed from future analysis (see Jupyter Notebook
843 3_Pairwise_Growth_Rates.ipynb for details). For gene-by-environment analysis, continuous
844 culture growth rates for each media and sgRNA combination were rescaled using plate reader
845 absolute growth data using Equation 5.

$$Relative\ GR_{media,sgRNA} = \frac{Raw\ GR_{sgRNA,media} + Absolute\ GR_{media}}{Absolute\ GR_{M9}} \quad (5)$$

846

847 We then removed escapers, CRISPRi constructs that abnormally lacked a growth rate
848 deficit, using a one-way Dixon Q-test for outliers at 95% confidence[33]. If at least four growth rates
849 were successfully calculated across barcodes and sgRNA orders after removing escapers, the
850 replicates were averaged to determine a mean relative growth rate. Finally, we normalized growth
851 rates so the fully Nont control had a growth rate of 1 and the most severe growth rate perturbation
852 observed (glyA_1_26_C + thyA_1_60_B_MM2, Min GR) had a growth rate of 0 using Equation 6.

$$Normalized\ GR = \frac{Relative\ GR - Min\ GR}{abs(Min\ GR)} \quad (6)$$

853

854

855 **Fitting the Coupling Model**

856 We fit single-gene expression-growth rate curves using scipy.optimize.least_squares, bounding R_0
857 to non-negative values with an initial guess of 0.5 for R_0 and 0 for n. We fit gene-gene coupling
858 values for each gene pair using the strategy described in Zimmer et al., 2016[28], with some
859 modifications. We optimized coupling constants a_{ij} and a_{ji} using scipy.optimize.least_squares,
860 bounding between -1 and 10 with an initial guess of 0 for both. To limit the spread of coupling
861 constant values, we calculated the RMSD between growth rate predictions and experimental
862 measurements with an added regularization term $\lambda * \Sigma abs(a_{ij})$. We used a regularization value of λ

863 = $10^{-1.25}$ after evaluating a range of values on subsampled data for both the pairwise and *folA*/*thyA*
864 libraries (Fig S2). We calculated pairwise and third-order predicted growth rates using each gene's
865 individual repression-growth rate parameters and relevant coupling constants.
866
867 **Estimating Uncertainty with Bootstrapping**
868 We used bootstrapping to computationally evaluate how sensitive each model's predictions were
869 to changes in single perturbation-growth rate fits. For each single perturbation curve, we generated
870 a new training set of the same size as our input data set by resampling data with replacement from
871 individual perturbation-growth rate measurements. For cases where both perturbation intensity and
872 growth rate measurements had associated uncertainties (e.g., qPCR-based repression
873 measurements and NGS-based growth rate measurements), we resampled these measurements
874 independently. In all cases, we refit perturbation-growth curves using resampled data 100 times
875 and used the resulting curves to estimate variance and establish 95% confidence intervals on our
876 model's fit parameters and accuracy (Table S2).

References

1. Bernard, Q., Smith, A.A., Yang, X., Koci, J., Foor, S.D., Cramer, S.D., Zhuang, X., Dwyer, J.E., Lin, Y.-P., Mongodin, E.F., et al. (2018). Plasticity in early immune evasion strategies of a bacterial pathogen. Proc. Natl. Acad. Sci. U.S.A. *115*. 10.1073/pnas.1718595115.

2. Abt, M.C., McKenney, P.T., and Pamer, E.G. (2016). Clostridium difficile colitis: pathogenesis and host defence. Nat Rev Microbiol *14*, 609–620. 10.1038/nrmicro.2016.108.

3. Depardieu, F., Podglajen, I., Leclercq, R., Collatz, E., and Courvalin, P. (2007). Modes and modulations of antibiotic resistance gene expression. Clin Microbiol Rev *20*, 79–114. 10.1128/CMR.00015-06.

4. Woodford, N., and Ellington, M.J. (2007). The emergence of antibiotic resistance by mutation. Clinical Microbiology and Infection *13*, 5–18. 10.1111/j.1469-0691.2006.01492.x.

5. Sager, R. (1997). Expression genetics in cancer: Shifting the focus from DNA to RNA. Proc. Natl. Acad. Sci. U.S.A. *94*, 952–955. 10.1073/pnas.94.3.952.

6. Chiang, A.C., and Massagué, J. (2008). Molecular basis of metastasis. N Engl J Med *359*, 2814–2823. 10.1056/NEJMra0805239.

7. Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., et al. (2017). A pathology atlas of the human cancer transcriptome. Science *357*, eaan2507. 10.1126/science.aan2507.

8. Magen, A., Das Sahu, A., Lee, J.S., Sharmin, M., Lugo, A., Gutkind, J.S., Schäffer, A.A., Ruppin, E., and Hannenhalli, S. (2019). Beyond Synthetic Lethality: Charting the Landscape of Pairwise Gene Expression States Associated with Survival in Cancer. Cell Reports *28*, 938-948.e6. 10.1016/j.celrep.2019.06.067.

9. Dykhuizen, D.E., Dean, A.M., and Hartl, D.L. (1987). Metabolic Flux and Fitness. Genetics *115*, 25–31. 10.1093/genetics/115.1.25.

10. Keren, L., Hausser, J., Lotan-Pompan, M., Vainberg Slutskin, I., Alisar, H., Kaminski, S., Weinberger, A., Alon, U., Milo, R., and Segal, E. (2016). Massively Parallel Interrogation of the Effects of Gene Expression Levels on Fitness. Cell *166*, 1282-1294.e18. 10.1016/j.cell.2016.07.024.

11. Hawkins, J.S., Silvis, M.R., Koo, B.-M., Peters, J.M., Osadnik, H., Jost, M., Hearne, C.C., Weissman, J.S., Todor, H., and Gross, C.A. (2020). Mismatch-CRISPRi Reveals the Co-varying Expression-Fitness Relationships of Essential Genes in Escherichia coli and Bacillus subtilis. Cell Systems *11*, 523-535.e9. 10.1016/j.cels.2020.09.009.

12. Price, M.N., Wetmore, K.M., Waters, R.J., Callaghan, M., Ray, J., Liu, H., Kuehl, J.V., Melnyk, R.A., Lamson, J.S., Suh, Y., et al. (2018). Mutant phenotypes for thousands of bacterial genes of unknown function. Nature *557*, 503–509. 10.1038/s41586-018-0124-0.

13. Metzstein, M.M., Stanfield, G.M., and Horvitz, H.R. (1998). Genetics of programmed cell death in C. elegans: past, present and future. Trends in Genetics *14*, 410–416. 10.1016/S0168-9525(98)01573-X.

14. 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. Nat Methods *5*, 781–787. 10.1038/nmeth.1240.

15. Butland, G., Babu, M., Díaz-Mejía, 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. Nat Methods *5*, 789–795. 10.1038/nmeth.1239.

16. Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L.Y., Toufighi, K., Mostafavi, S., et al. (2010). The Genetic Landscape of a Cell. Science *327*, 425–431. 10.1126/science.1180823.

17. Costanzo, M., VanderSluis, B., Koch, E.N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S.D., et al. (2016). A global genetic interaction network maps a wiring diagram of cellular function. Science *353*, aaf1420–aaf1420. 10.1126/science.aaf1420.

18. Jaffe, M., Dziulko, A., Smith, J.D., St.Onge, R.P., Levy, S.F., and Sherlock, G. (2019). Improved discovery of genetic interactions using CRISPRiSeq across multiple environments. Genome Res. *29*, 668–681. 10.1101/gr.246603.118.

19. Poelwijk, F.J., Krishna, V., and Ranganathan, R. (2016). The Context-Dependence of Mutations: A Linkage of Formalisms. PLoS Comput Biol *12*, e1004771. 10.1371/journal.pcbi.1004771.

20. Domingo, J., Baeza-Centurion, P., and Lehner, B. (2019). The Causes and Consequences of Genetic Interactions (Epistasis). Annu. Rev. Genom. Hum. Genet. *20*, 433–460. 10.1146/annurev-genom-083118-014857.

21. Wood, K., Nishida, S., Sontag, E.D., and Cluzel, P. (2012). Mechanism-independent method for predicting response to multidrug combinations in bacteria. Proc. Natl. Acad. Sci. U.S.A. *109*, 12254–12259. 10.1073/pnas.1201281109.

22. Wood, K.B., Wood, K.C., Nishida, S., and Cluzel, P. (2014). Uncovering Scaling Laws to Infer Multidrug Response of Resistant Microbes and Cancer Cells. Cell Reports *6*, 1073–1084. 10.1016/j.celrep.2014.02.007.

23. Zimmer, A., Tendler, A., Katzir, I., Mayo, A., and Alon, U. (2017). Prediction of drug cocktail effects when the number of measurements is limited. PLoS Biol *15*, e2002518. 10.1371/journal.pbio.2002518.

24. Katzir, I., Cokol, M., Aldridge, B.B., and Alon, U. (2019). Prediction of ultra-high-order antibiotic combinations based on pairwise interactions. PLoS Comput Biol *15*, e1006774. 10.1371/journal.pcbi.1006774.

25. Tendler, A., Zimmer, A., Mayo, A., and Alon, U. (2019). Noise-precision tradeoff in predicting combinations of mutations and drugs. PLoS Comput Biol *15*, e1006956. 10.1371/journal.pcbi.1006956.

26. Pandey, S., Gruenbaum, A., Kanashova, T., Mertins, P., Cluzel, P., and Chevrier, N. (2020). Pairwise Stimulations of Pathogen-Sensing Pathways Predict Immune Responses to Multi-adjuvant Combinations. Cell Systems *11*, 495-508.e10. 10.1016/j.cels.2020.10.001.

27. Larkins-Ford, J., Greenstein, T., Van, N., Degefu, Y.N., Olson, M.C., Sokolov, A., and Aldridge, B.B. (2021). Systematic measurement of combination-drug landscapes to predict in vivo treatment outcomes for tuberculosis. Cell Systems *12*, 1046-1063.e7. 10.1016/j.cels.2021.08.004.

28. Zimmer, A., Katzir, I., Dekel, E., Mayo, A.E., and Alon, U. (2016). Prediction of multidimensional drug dose responses based on measurements of drug pairs. Proc. Natl. Acad. Sci. U.S.A. *113*, 10442–10447. 10.1073/pnas.1606301113.

29. Balakrishnan, R., Mori, M., Segota, I., Zhang, Z., Aebersold, R., Ludwig, C., and Hwa, T. (2022). Principles of gene regulation quantitatively connect DNA to RNA and proteins in bacteria. Science *378*, eabk2066. 10.1126/science.abk2066.

30. 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. 10.1016/j.cell.2013.02.022.

31. Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H.S., Koo, B.-M., Marta, E., et al. (2016). A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. Cell *165*, 1493–1506. 10.1016/j.cell.2016.05.003.

32. Jost, M., Santos, D.A., Saunders, R.A., Horlbeck, M.A., Hawkins, J.S., Scaria, S.M., Norman, T.M., Hussmann, J.A., Liem, C.R., Gross, C.A., et al. (2020). Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs. Nat Biotechnol *38*, 355–364. 10.1038/s41587-019-0387-5.

33. Mathis, A.D., Otto, R.M., and Reynolds, K.A. (2021). A simplified strategy for titrating gene expression reveals new relationships between genotype, environment, and bacterial growth. Nucleic Acids Research *49*, e6–e6. 10.1093/nar/gkaa1073.

34. Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat Protoc *8*, 2180–2196. 10.1038/nprot.2013.132.

35. Rousset, F., Cui, L., Siouve, E., Becavin, C., Depardieu, F., and Bikard, D. (2018). Genome-wide CRISPR-dCas9 screens in E. coli identify essential genes and phage host factors. PLoS Genet *14*, e1007749. 10.1371/journal.pgen.1007749.

36. Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z., Zhang, C., and Xing, X.-H. (2018). Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. Nat Commun *9*, 2475. 10.1038/s41467-018-04899-x.

37. 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. Nat Protoc *8*, 555–567. 10.1038/nprot.2013.021.

38. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol *2*. 10.1038/msb4100050.

39. Côté, J.-P., French, S., Gehrke, S.S., MacNair, C.R., Mangat, C.S., Bharat, A., and Brown, E.D. (2016). The Genome-Wide Interaction Network of Nutrient Stress Genes in Escherichia coli. mBio *7*, e01714-16. 10.1128/mBio.01714-16.

40. Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol *31*, 233–239. 10.1038/nbt.2508.

41. Schober, A.F., Mathis, A.D., Ingle, C., Park, J.O., Chen, L., Rabinowitz, J.D., Junier, I., Rivoire, O., and Reynolds, K.A. (2019). A Two-Enzyme Adaptive Unit within Bacterial Folate Metabolism. Cell Reports *27*, 3359-3370.e7. 10.1016/j.celrep.2019.05.030.

42. Kwon, Y.K., Higgins, M.B., and Rabinowitz, J.D. (2010). Antifolate-Induced Depletion of Intracellular Glycine and Purines Inhibits Thymineless Death in *E. coli*. ACS Chem. Biol. *5*, 787–795. 10.1021/cb100096f.

43. Sangurdekar, D.P., Zhang, Z., and Khodursky, A.B. (2011). The association of DNA damage response and nucleotide level modulation with the antibacterial mechanism of the anti-folate drug Trimethoprim. BMC Genomics *12*, 583. 10.1186/1471-2164-12-583.

44. 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. Nat Genet *44*, 101–105. 10.1038/ng.1034.

45. Kuchina, A., Brettner, L.M., Paleologu, L., Roco, C.M., Rosenberg, A.B., Carignano, A., Kibler, R., Hirano, M., DePaolo, R.W., and Seelig, G. (2021). Microbial single-cell RNA sequencing by split-pool barcoding. Science *371*, eaba5257. 10.1126/science.aba5257.

46. Blattman, S.B., Jiang, W., Oikonomou, P., and Tavazoie, S. (2020). Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. Nat Microbiol *5*, 1192–1201. 10.1038/s41564-020-0729-6.

47. Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C.P., Jerby-Arnon, L., Marjanovic, N.D., Dionne, D., Burks, T., Raychowdhury, R., et al. (2016). Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. Cell *167*, 1853-1866.e17. 10.1016/j.cell.2016.11.038.

48. Taggart, J.C., Lalanne, J.-B., and Li, G.-W. (2021). Quantitative Control for Stoichiometric Protein Synthesis. Annu. Rev. Microbiol. *75*, 243–267. 10.1146/annurev-micro-041921-012646.