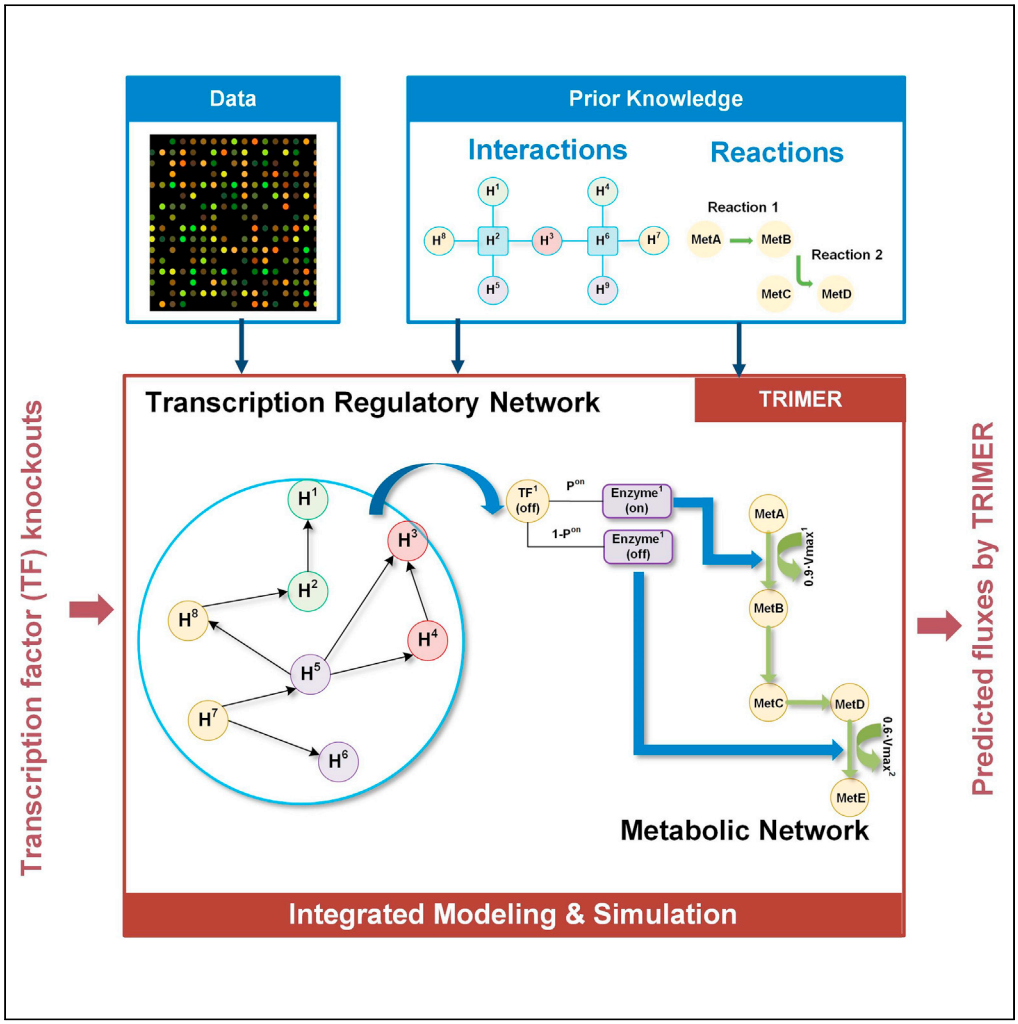
iScience

CellPress
OPEN ACCESS

Article

TRIMER: Transcription Regulation Integrated with Metabolic Regulation



Puhua Niu, Maria J. Soto, Byung-Jun Yoon, Edward R. Dougherty, Francis J. Alexander, Ian Blaby, Xiaoning Qian

ikblaby@lbl.gov (I.B.)
xqian@ece.tamu.edu (X.Q.)

Highlights

TRIMER models transcription-regulated metabolism using Bayesian network modeling;

TRIMER integrates prior knowledge (regulatory interaction) with data (expression);

TRIMER enables metabolic behavior prediction for general knockout strategies;

TRIMER includes a simulator as an evaluation platform for similar hybrid models;

TRIMER reliably predicts metabolite yields for both simulated and experimental data.

Niu et al., iScience *24*, 103218
November 19, 2021
https://doi.org/10.1016/j.isci.2021.103218

iScience

CellPress
OPEN ACCESS

Article

TRIMER: Transcription Regulation Integrated with Metabolic Regulation

Puhua Niu,[1] Maria J. Soto,[2] Byung-Jun Yoon,[1,3] Edward R. Dougherty,[1] Francis J. Alexander,[3] Ian Blaby,[2,4,*] and Xiaoning Qian[1,3,5,*]

SUMMARY

There has been extensive research in predictive modeling of genome-scale metabolic reaction networks. Living systems involve complex stochastic processes arising from interactions among different biomolecules. For more accurate and robust prediction of target metabolic behavior under different conditions, not only metabolic reactions but also the genetic regulatory relationships involving transcription factors (TFs) affecting these metabolic reactions should be modeled. We have developed a modeling and simulation pipeline enabling the analysis of Transcription Regulation Integrated with Metabolic Regulation: TRIMER. TRIMER utilizes a Bayesian network (BN) inferred from transcriptomes to model the transcription factor regulatory network. TRIMER then infers the probabilities of the gene states relevant to the metabolism of interest, and predicts the metabolic fluxes and their changes that result from the deletion of one or more transcription factors at the genome scale. We demonstrate TRIMER's applicability to both simulated and experimental data and provide performance comparison with other existing approaches.

INTRODUCTION

There has been extensive research in *in silico* modeling and prediction of genome-scale metabolic behavior, mostly focusing on mutant strain design with metabolic reaction network modeling (Varma and Palsson, 1994a; Edwards and Palsson, 2000; Barrett et al., 2006; Segre et al., 2002; Burgard et al., 2003; Shlomi et al., 2005; Lewis et al., 2010; Jensen et al., 2011; Ren et al., 2013; Palsson, 2015; Apaydin et al., 2017; Arkin et al., 2018). However, living systems involve complex and stochastic processes arising from interactions between different types of biomolecules. For more accurate and robust prediction of target metabolic behavior under different conditions or contexts, not only metabolic reactions but also the integration of genetic regulatory relationships involving transcription factors (TFs) that may regulate metabolic reactions, should be appropriately modeled. Due to the increasing computational complexity when considering multiple types of biomolecules in one computational system model, often transcription regulation has been integrated via "transcriptional regulatory constraints" with various heuristics for flux-balance analysis (FBA) of metabolic networks (Covert and Palsson, 2003; Shlomi et al., 2007; Covert et al., 2008; Shlomi et al., 2008; Fendt et al., 2010; Machado and Herrgård, 2014; Reiss et al., 2015; Reed, 2017; Motamedian et al., 2017; Yu and Blair, 2019). Many of these computational tools were often only validated for selected model organisms with curated data and network models. To generalize these integrated hybrid models for different organisms, the reproducibility of the results require careful validation, for example, starting from simulated ground truth models.

Probabilistic Regulation of Metabolism (PROM) (Chandrasekaran and Price, 2010) introduced probabilistic modeling of transcription regulation for better integration with condition-specific metabolism. PROM can be considered as one of the first integrated transcriptional-metabolic network models that take advantage of both existing prior knowledge and gene expression data. Specifically, conditional probabilities were inferred by microarray data analysis for annotated TF (target gene)-reaction interactions to incorporate transcriptional regulation information in genome-scale metabolic network analysis under different conditions or contexts. IDREAM (Wang et al., 2017), an updated version of PROM, additionally allowed modeling subtle growth defects to further improve metabolic flux predictions. Recently, an algorithm called OptRAM was developed based on IDREAM for designing optimized strains for ethanol overproduction in yeast (Shen et al., 2019).

[1]Texas A&M University, Department of Electrical and Computer Engineering, College Station, TX, 77843, USA

[2]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA

[3]Brookhaven National Laboratory, Computational Science Initiative, Upton, NY, 11973, USA

[4]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA

[5]Lead contact

*Correspondence: ikblaby@lbl.gov (I.B.), xqian@ece.tamu.edu (X.Q.)

https://doi.org/10.1016/j.isci.2021.103218

The essential idea of PROM and its extensions is to infer the TF-gene conditional probabilities of the form Pr(gene = ON/OFF| TF = ON/OFF) so that metabolic reactions regulated by specific genes—for example, through the specific enzymes manifested as gene-protein-reaction (GPR) rules—can be modeled dependent on either genotypic or environmental changes by adjusting the reaction flux constraints in the FBA formulation for metabolic modeling. Although it is computationally desirable to simplify the TF regulatory roles by introducing TF-gene conditional probabilities estimated by local frequentist estimates based on gene expression profiles, global TF-gene dependency structures may not be well captured. The existing models are also limited in the sense that only conditional probabilities based on univariate conditions were modeled. More flexible modeling that enables predictions with more complicated condition changes, for example, multiple TF knockouts when designing mutant strains, is still lacking in the literature.

The main contribution of this paper is to introduce a new flexible genome-scale simulation and analysis pipeline, *TRIMER*—Transcription Regulation Integrated with Metabolic Regulation, for integrative systems modeling of TF-regulated metabolism. Specifically, a Bayesian network (BN) (Koller and Friedman, 2009) is employed in TRIMER, instead of local TF-gene conditional probabilities or transcriptional regulatory constraints, thereby aiming at effectively capturing the global transcriptional regulatory relationships that may affect metabolism. Through this BN, the influence of transcription regulation (and its changes) on metabolic behavior under different conditions will be manifested more accurately via more flexible conditional probability inference which is linked to metabolism through the prior knowledge on TF-gene-reaction interactions. In the prediction mode of TRIMER for a given model organism, expression data, and prediction tasks, a BN will be first inferred based on gene expression profiles with the prior knowledge on TF-gene-reaction interactions. Based on the inferred Bayesian network, given a condition (for example, multiple TF knockouts), we can infer the corresponding probabilities of gene states and consequently flux predictions can be performed by corresponding *in silico* metabolic models.

In addition to the modeling and analysis functionalities in TRIMER, we have also developed a simulator that simulates the TF-regulated metabolic network, which can generate both gene expression states and metabolic fluxes from a given transcriptional-metabolic hybrid model. Such a simulator provides a fair performance evaluation platform to help better benchmark and validate new model inference and flux prediction methods in computational systems biology.

RESULTS

In this section, we first provide a brief overview of our hybrid TF-regulated metabolic network model, *TRIMER*: Transcription Regulation Integrated with MEtabolic Regulation. We then present the experimental results based on both simulated and experimental data to demonstrate the effectiveness and flexibility of TRIMER for metabolic flux prediction under different conditions.

TRIMER: Transcription Regulation Integrated with Metabolic regulation

TRIMER differs from the existing methods in the way of systematic prediction of effective intervention strategies when applied to the transcription regulatory network for regulation of metabolism. Specifically, TRIMER is based on a Bayesian Network (BN) for learning transcription regulation from gene expression data. Instead of utilizing simple conditional probabilities of the form Pr(gene = ON/OFF| TF = ON/OFF) as in PROM (Chandrasekaran and Price, 2010), the BN can be used to determine a probabilistic inference of the effect of alterations (e.g., gene deletions) of multiple TFs (or genes). While the framework presented is independent of the nature of TF engineering, we focus herein on gene deletions (i.e., knockouts (KO)). Furthermore, BN modeling enables intuitive incorporation of prior knowledge (e.g., pathways or pairwise regulatory relationship between genes) for learning the *TF-Regulated gene Network (TRN)*.

In TRIMER, a BN is trained from the gene expression data to model the joint distribution for all the relevant TFs and genes, where the resulting BN can be subsequently used to infer the steady-state conditional probabilities of the form Pr(gene(s)|TF(s)) = $p(\vec{g}\,|\,\vec{TF})$– i.e., the probability of gene states given the states of TFs of interest. For example, we can use the BN to estimate the probability that a target gene known to regulate a specific metabolic pathway is induced given that expression of one or more TFs is abolished by gene deletion. The estimated probabilities can be used to constrain the metabolic reaction fluxes of interest, based on which the flux changes of selected metabolites resulting from the genetic alteration (e.g., TF gene deletion) can be predicted via flux balance analysis (FBA). The gene-protein-reaction (GPR) rules,

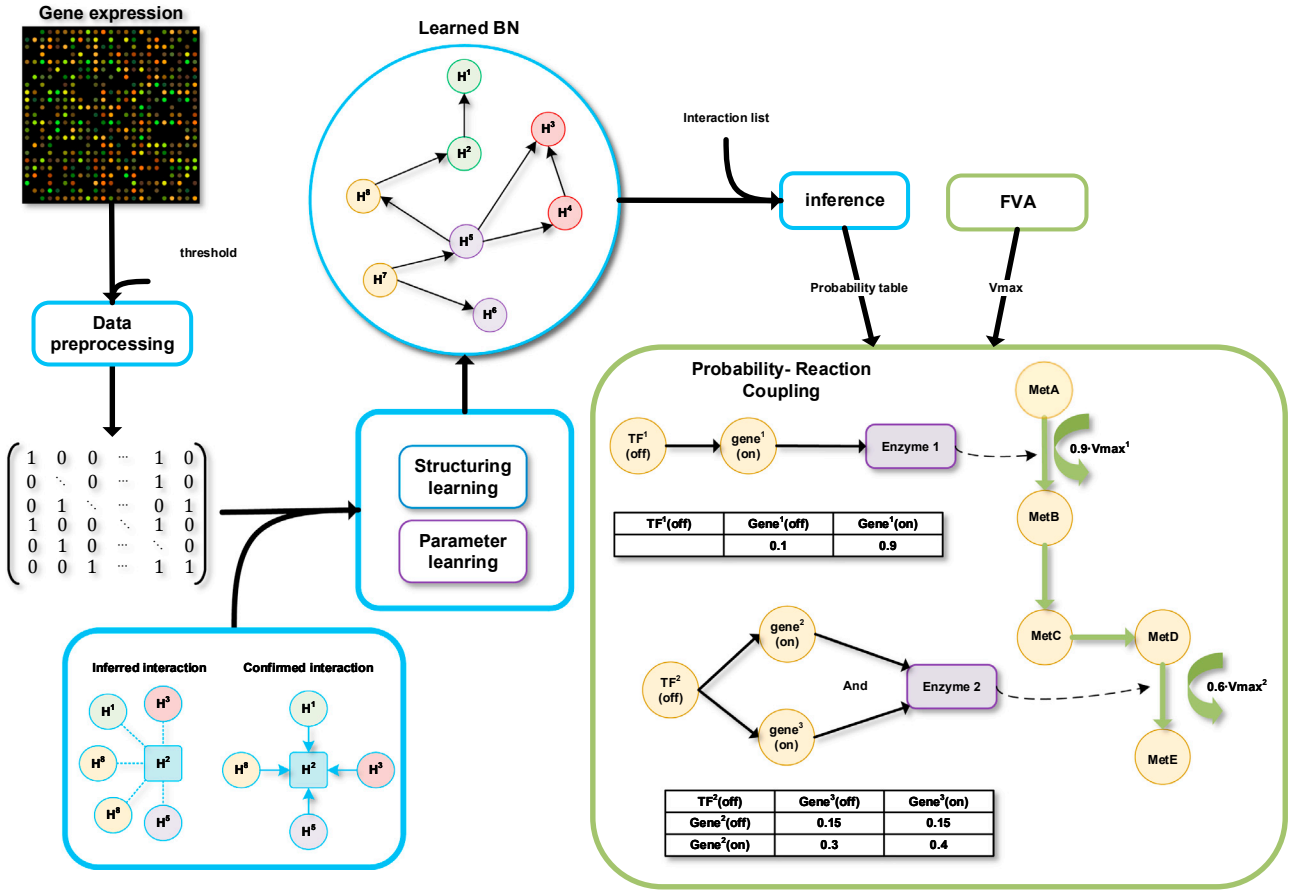


Figure 1. Illustrative overview of TRIMER

Gene expression data are used to infer the Bayesian network (BN) modeling the transcriptional regulations with the prior knowledge on molecular interactions. The impact of transcription factor knockout on downstream target genes that affect metabolic pathways are inferred using the BN. The estimated probability that a given target gene being turned on modulates the constraints in the flux variability analysis (FVA) resulting in probabilistic metabolic predictions. Each module component in TRIMER has the detailed explanations in the flowchart in Figure 2 with the matched box boundary colors for the corresponding transcription regulation and metabolic flux prediction modules.

which inform us of the respective metabolic pathways regulated by different genes, are used to link the translation regulation modeled by the BN with the metabolic regulation simulated by FBA.

TRIMER, which jointly models transcription regulation and metabolic regulation via the hybrid approach described above, allows us to assess the efficacy of potential TF engineering strategies and identify the optimal strategy for modulating the metabolic fluxes of interest. The desirability of a given genetic alteration can be assessed *in silico* using TRIMER, which can be validated through actual TF deletion and screening experiments in the laboratory.

Figure 1 provides a high-level overview of TRIMER, illustrating its main workflow. As shown in this diagram and detailed in Section STAR Methods, TRIMER consists of two main modules: (1) The BN module for modeling and inference of transcription regulation and (2) the metabolic flux prediction module for estimating the impact of alterations in the TRN on the metabolic outcomes. The two modules are linked to each other by the GPR rules. Furthermore, Figure 2 depicts the overall workflow of module components with explanations in TRIMER, including the interconnections among the computational modules that comprise TRIMER. As TRIMER involves multiple variables in the transcriptional-metabolic hybrid model, we provide the corresponding mathematics notations in Table 1 of Section STAR Methods. We also illustrate the implemented data structures in our TRIMER package in Figure 3 of Section STAR Method.

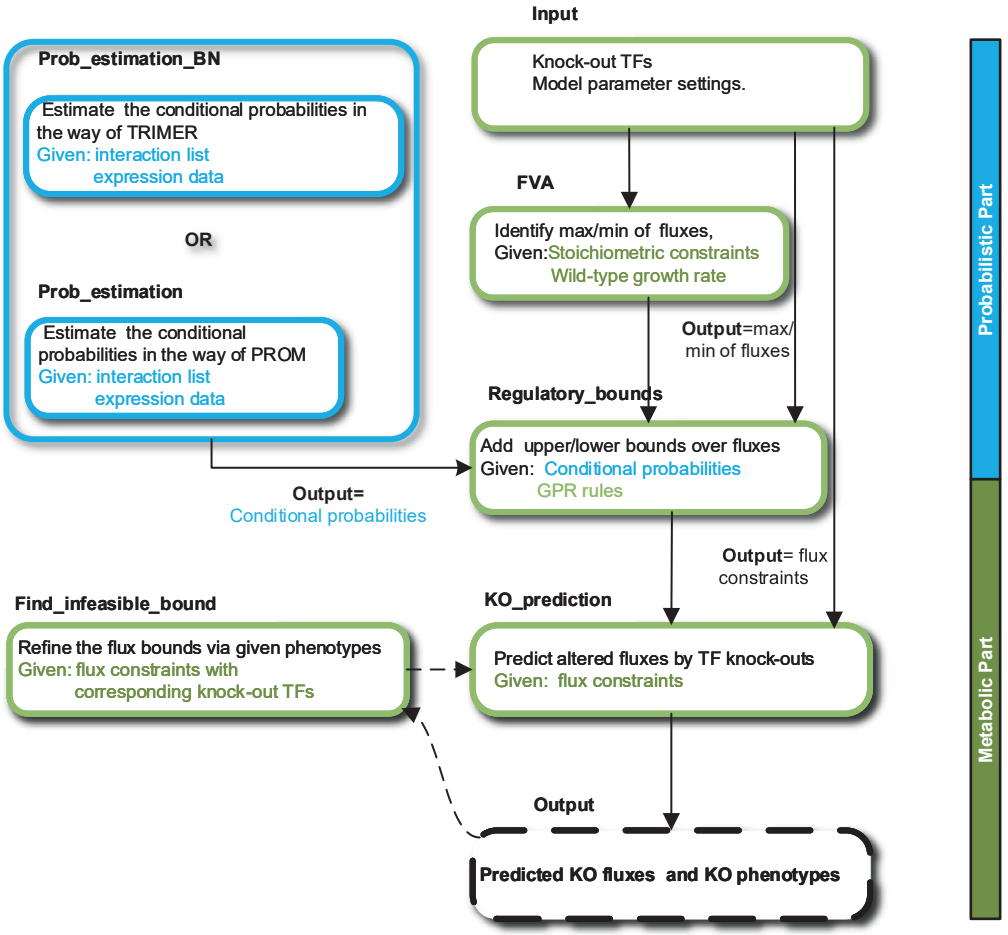


Figure 2. TRIMER flow-chart, where the explanations of the major computational module component for both transcription regulatory network modeling and metabolic flux prediction in TRIMER, together with their interconnections are illustrated

The blue boxes denote the module components for transcription regulation and the green boxes denote the metabolic reaction network model components.

TRIMER as a simulator

TRIMER can serve as a simulator to generate both gene expression and metabolic flux data. Specifically, given model organisms or specific pathways of interest, metabolic network model can be first extracted from the existing models in the COBRA toolbox Heirendt et al. (2019). Based on GPR rules and available knowledge on gene-gene interactions, we can simulate a BN (of different size if needed by growing from a set of metabolism-regulating genes to the whole genome for example) by randomly sampled genes and edges between the selected gene pairs. Given a simulated BN structure, BN model parameters characterizing the corresponding conditional probability tables can also be simulated. With the simulated BN connecting to the metabolic model through GPR rules, we can first sample the gene expression data based on the BN. At the same time, with the simulated conditional probabilities under different conditions, for example with TF knockouts, metabolic flux predictions can be computed by constructing the corresponding transcriptional constraints as described in Section STAR Methods. When serving as a simulator, TRIMER directly simulates a BN instead of learning the BN based on the regulatory prior knowledge and gene expression data as shown in Figure 2. Through this simulation procedure, TRIMER can serve as a fair benchmark platform for validation and comparison of different transcriptional-metabolic prediction methods as we have showcased in the following simulation experiments.

Table 1. Mathematical notations in TRIMER.

Module	Symbol	Value	Description
Transcription regulation	TF	$\{0, 1\}$	transcription factor expression state
	g	$\{0, 1\}$	gene expression state
	G	index set	set of genes
	$p(\overrightarrow{g}\|\overrightarrow{TF})$	$(0, 1)$	conditional probabilities of gene state given TF state
Metabolism regulation	r	index	metabolic reaction index
	R	index set	set of metabolic reactions
	$G(r)$	index set	set of genes that regulate reaction r
	$\overrightarrow{v}$	$[\overrightarrow{lb}, \overrightarrow{ub}]$	steady-state metabolic fluxes
	lb_r	$\mathbb{R}$	flux lower bound of reaction r
	ub_r	$\mathbb{R}^+$	flux upper bound of reaction r
	S	real-valued matrix	stoichiometric coefficients in the metabolic reaction network
	$v_{max}(r)$	$\mathbb{R}^+$	maximum flux of reaction r, estimated via flux variability analysis
	$\overrightarrow{v}^0$	$[\overrightarrow{lb}, \overrightarrow{ub}]$	wild-type steady-state metabolic fluxes

Simulation of *E. coli* transcription regulatory network

In this set of experiments, we first validate that the BN learning in TRIMER can capture the regulatory relationships, which thereafter leads to reliable metabolic flux predictions with TF knockouts, based on a simulated BN model as the ground truth.

We first describe the procedure to simulate the BN for *Escherichia. coli* TRN based on the TRIMER simulator, given the corresponding metabolic network model. We start with a simulated small-scale BN and then demonstrate the scalability and flexibility of TRIMER with multiple TF knockout results in a large-scale BN in this section. With the simulated ground truth BN, the metabolic network model is adopted to simulate reaction fluxes with the constraints based on the simulated conditional probabilities.

Simulating integrated transcription and metabolic regulations for a small-scale BN

For the experiments with the small-scale BN, we simulated the interactions between key genes related to indole production. Besides 12 transcription factors studied in the aforementioned TF knockout experiments using the Keio library, 32 corresponding target genes were also taken as the backbone nodes for the small-scale BN. To be specific, we selected these target genes by computing Pearson correlation coefficients (PCC) between the 12 indole-related TFs and the remaining genes in the gene expression data from EcoMAC (Carrera et al., 2014). The resulting 32 target genes were selected as each of these had PCC>0.65 with at least one of these 12 TFs.

We took these 12 TFs and 32 selected target genes as the backbone nodes (44 in total) to simulate the BN as the transcription regulatory network. When simulating edges in this small-scale BN, directed regulatory interactions of these nodes were initialized with the following restrictions: 1) Only the nodes corresponding to the TFs can serve as parent nodes in the simulated regulatory interactions; 2) the maximum number of edges between one TF parent node and all the other TF nodes was restricted to be half of the total number of TFs, and the maximum number of edges between one TF parent node and all the other target gene nodes was restricted to be half of the total number of target genes (This restriction prevents from simulating a very dense BN); 3) the edges were randomly generated between every valid pair of nodes with the corresponding values of conditional probability table (CPD) for each node being initialized randomly according to the uniform distribution Unif$(0, 1)$. The gene expression data were first generated by sampling based on the simulated BN. Ten sample sets of 2000 binary gene expression profiles were drawn via the forward sampling procedure on the simulated BN. For each sample set of 2000 generated samples, five subsets of 100,

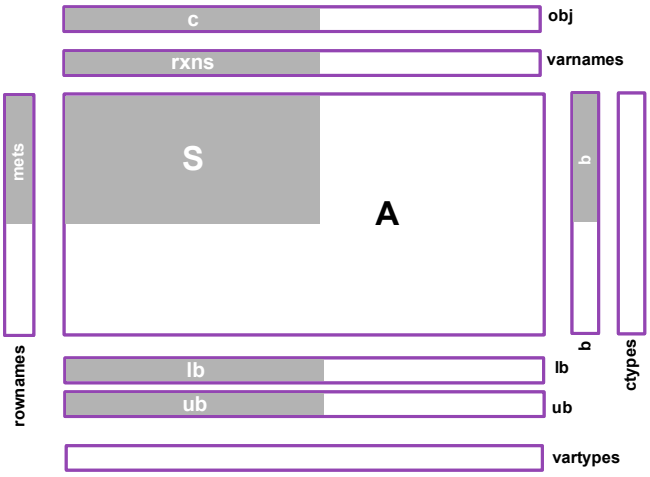


Figure 3. Metabolic models represented as MATLAB data structures: Boxes indicate size and orientation of the fields

Black text denotes the corresponding field names. Gray areas contain data from the metabolic model, with white text indicating the relevant field names.

200, 400, 800, and 1600 samples were randomly selected to construct the corresponding training sets for performance evaluation. In this way, 50 datasets in total with sizes ranging from 100 to 1600 were obtained.

On the other hand, regulating targets for each TF were found by the dependency between pairs of nodes, which can be obtained from the simulated BN structure. Based on these interactions, we can infer the probabilities of the corresponding gene states for different TF knockouts and the wild-type from the simulated BN. Based on the inferred probabilities and the gene-reaction relationships, the flux constraints in FBA can be adjusted to predict corresponding reaction fluxes of TF knockout mutants and the wild-type. For both the simulating ground truth BN and the inferred BNs based on the simulated expression data, the corresponding metabolic fluxes can be simulated based on this procedure for performance evaluation. Note that all of our simulation experiments are based on the *E. coli* iAF1260 metabolic network model for FBA.

Small-scale BN structure inference based on simulated gene expression data

Given the simulated gene expression data, the first task was to learn the BN structure that best fits the simulated expression data for performance evaluation of discovering the regulatory interaction between TFs and target genes. In our experiments, we used score-based structure learning methods for this task, where the quality of the learned BN structure was measured by the Bayesian Information Criterion (BIC) score. We tested two BN structures: Chow-Liu tree search algorithm for identifying the global optimal tree-based BN structure and Tabu search algorithm for more general BN structure learning. Tabu search only finds the local optimal structure. In order to guarantee the quality of the predicted solutions in our experiments, the Tabu length was set to be 100 where the best structural changes in every 100 iterations were iteratively updated as a reference for future search.

Once the BN structure was inferred based on the expression data, conditional probability tables for the BN were fit to the expression data by maximum likelihood estimates (MLEs). Finally, the regulated genes and associated conditional probabilities given TF states can be computed via examining dependency from the learned BN structure and performing the exact/approximate inference algorithm over the fitted BN. It should be noted that the original PROM estimates the conditional probabilities of gene states given TF states by MLE (relative frequencies) directly based on the expression data, while the authors stated that they adjusted FBA constraints by investigating only the "experimentally verified" TF-gene pairs. In our experiments, the underlying dependency between pairs of nodes in the simulated ground truth BN was considered as the actual TF-gene pairs for PROM, to some extent in favor of PROM since it did not learn the regulation network structure.

In Table 2, we have shown the average numbers of false positive and false negative edges in the inferred general BN models by Tabu search from different numbers of training gene expression profiles, compared

Table 2. False negative/positives for learned BN structures by Tabu search[a]

Training dataset size	100	200	400	800	1600
False positive (avg± std)	39.2±3.7	23.6±6.0	17.0±3.2	11.8±2.9	10.4±2.5
False negative (avg± std)	55.8±3.8	36.6±5.4	24.4±2.2	15.6±2.5	13.8±2.3

[a]The total number of edges in the ground-truth simulated Bayesian network is 137.

to the edges in the simulated BN. With the increasing number of training gene expression profiles, it is clear that the BN learning module in TRIMER can derive the BN model closer to the ground truth BN that is used to simulate the expression profiles as expected. Figure 4 shows the exemplar BN models learned with different numbers of training expression profiles. We have also checked the structure learning performance when we inferred tree-based BN models, whose results are provided in Table 3. As expected, due to the imposed constraints to allow optimal search for the tree-based BN models, the true negatives are much larger compared to the results in Table 2. Nevertheless, both false positives and false negatives decrease with the increasing number of training expression profiles.

Evaluation of flux prediction using TRIMER based on the small-scale inferred network

We further compare the flux prediction results by TRIMER-C with Chow-Liu tree (tree-TRIMER) and general BN structure (BN-TRIMER) to the results by PROM, based on simulated gene expression data. Note that we focused on applying the flux constraints based on Equation 1 (TRIMER-C) for fair performance comparison with PROM. We computed the PCC between the simulated biomass and indole fluxes based on the simulated ground truth network model and the predicted biomass and indole fluxes based on the inferred networks of both wild-type and the mutant strains deleted for TFs in the regulation network. For 10 simulated datasets of the same number of gene expression profiles, the average PCC and its standard deviation (std) were computed. Figures 5 and 6 summarize the performance comparison of TRIMER and PROM for biomass and indole flux prediction respectively, with the inferred BNs based on different numbers of simulated gene expression data.

As shown in Figure 5, from simulated expression data, BN-TRIMER consistently gives the closest biomass flux prediction to the simulated fluxes based on the ground truth model. It is clear that with more expression data, the predicted fluxes can get better and vary less with different simulated expression data. With small training expression data, PROM's flux prediction can have quite a weak correlation while with increasing number of expression profiles, the prediction can be improved. For tree-TRIMER, as the model class deviates from the ground truth model, the prediction performance saturates when the number of training expression profiles is 400. On the other hand, with small training sets, tree-TRIMER still performs better than PROM.

Comparison for the indole flux predictions are also provided in Figure 6. Note that the ground truth BN models were simulated based on the core subnetwork centering around indole-related reactions, identified by correlation analysis using EcoMAC gene expression data. We observe that both versions of TRIMER have better indole flux prediction performance, especially with small training data, compared to the results in Figure 4. The tree-TRIMER shows much better performance, which suggests that good prior knowledge on what to model for the TF regulation network may significantly enhance flux predictions. On the other hand, unlike TRIMER, PROM only models local dependency instead of global dependency, and its indole and biomass flux prediction performances are similar. In the Data S1, we also provide a table (Table A3) summarizing the quantitative results of the expected PCC values and their corresponding standard deviations accompanying Figures 5 and 6.

Simulating integrated transcription and metabolic regulations for a large-scale BN

We further simulate a large-scale BN with multiple TF knockouts to demonstrate the scalability and flexibility of the BN learning and metabolic flux prediction modules in TRIMER. To simulate a large-scale BN, we used all the genes included in the interaction list from EcoMAC and randomly selected 40% valid pairs of the interaction list as edges, resulting in a large-scale BN with 1591 genes and 1503 edges in this set of experiments.

One sample set of 2000 expression profiles was drawn via the forward sampling procedure as described for the small-scale BN. As done in the previous experiments, randomly selected interactions as edges in this

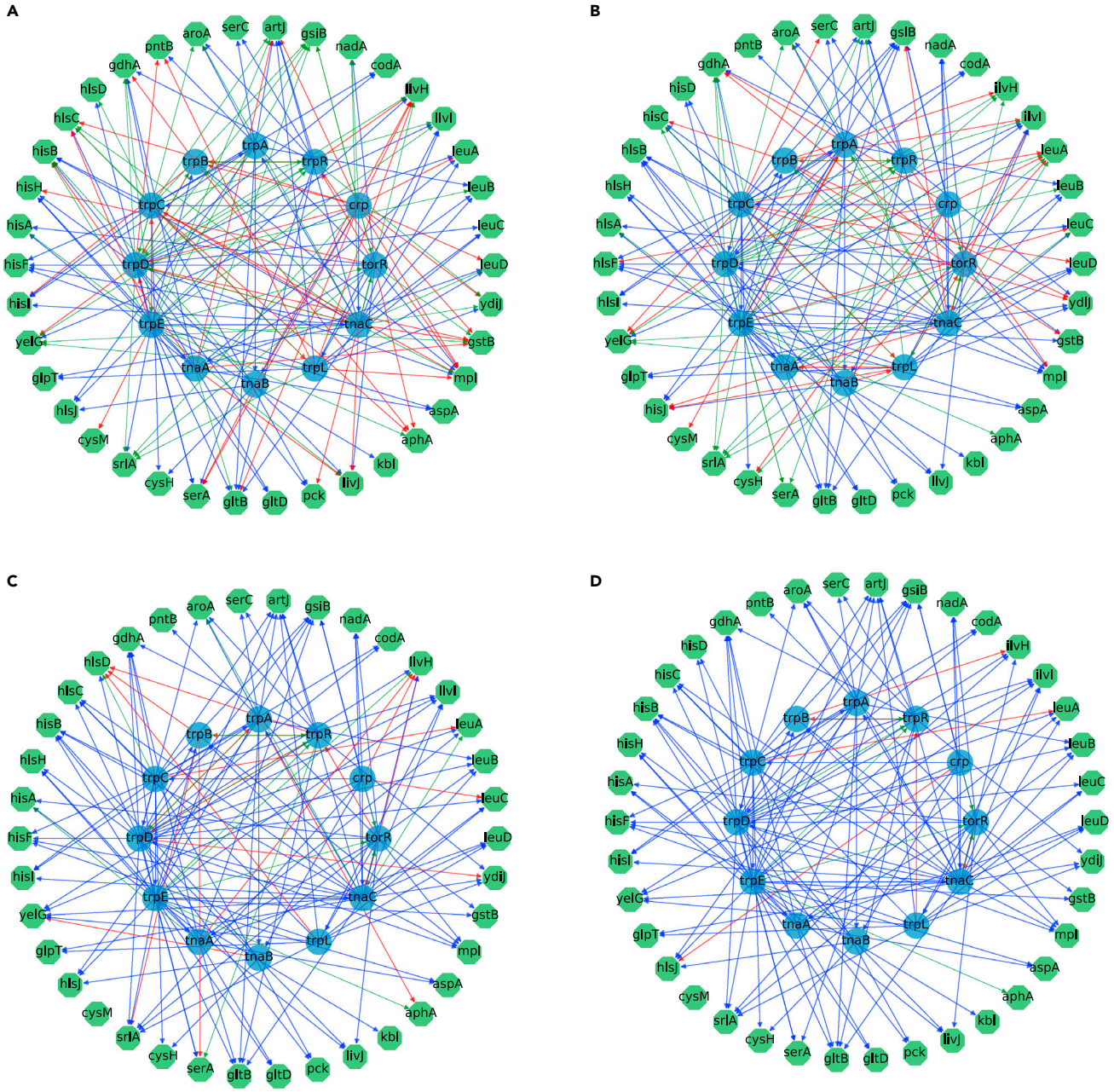


Figure 4. Bayesian network modeling of transcription regulation network

Examples of learned BNs from (A) 100, (B) 200, (C) 800, and (D) 1600 simulated expression profiles. The blue circled nodes represent 12 TFs while the green nodes are the corresponding target genes. The blue edges denote the accurately learned edges, the red edges are false positives where the regulatory edges were falsely added by BN structure learning, and the green edges are false negatives that BN learning was not able to identify.

simulated BN were taken as ground truth interaction list and corresponding conditional probabilities associated with them were simulated. We used TRIMER-B to simulate the fluxes when we knocked out two TFs at the same time in this set of experiments.

Evaluation of flux prediction using TRIMER based on the large-scale inferred network

We compared the flux prediction results by TRIMER-B with Chow-Liu tree (tree-TRIMER) and general BN structure (BN-TRIMER) with the results by PROM, based on PCC between the simulated and predicted fluxes for both biomass and indole production. Base on the sampled expression data from the simulated

Table 3. False negative/positives for learned tree-based BN structures[a]

Training dataset size	100	200	400	800	1600
False positive (avg± std)	21.8±3.4	16.6±3.1	13.5±2.1	10.1±1.0	9.4±1.2
False negative (avg± std)	107.8±3.4	102.6±3.1	99.5±2.1	96.1±1.0	95.4.±1.2

[a]The total number of edges in the ground truth simulated Bayesian network is 137.

ground truth BN model, a general-structure BN was inferred using our TRIMER package, resulting in 1377 edges, denoted as BN-TRIMER. When we restricted the BN to a Chow-Liu tree, the inferred BN had 1590 edges, denoted as tree-TRIMER. We used the simulated ground truth interaction list for both TRIMER and PROM to construct the corresponding transcriptional constraints for flux predictions. To demonstrate the capability of TRIMER modeling mutant strains with multiple knockout TFs at the same time, ten random sets of 50 TF pairs were selected according to the EcoMAC interaction list. The average PCC and standard deviation (std) were computed based on these ten sets of 50 double TF knockout mutants for comparison, as shown in Table A4 in the Data S1. Figure 7 shows the corresponding bar plots, from which it is clear that based on the simulated expression data, BN-TRIMER consistently gave the best biomass and indole flux prediction with respect to the simulated fluxes based on the ground truth model.

Experimental validation of metabolic flux predictions made by TRIMER

To further demonstrate the utility of TRIMER in *in silico* metabolic flux prediction for TF knockout mutants, we compared the prediction performance of TRIMER with PROM, IDREAM (Wang et al., 2017), and TR-FBA (Motamedian et al., 2017) for both biomass and indole flux prediction for *E. coli* TF-knockout mutants. We inferred the corresponding models based on the archived microarray gene expression data and the experimentally verified TF-gene interactions in EcoMAC (Carrera et al., 2014).

For PROM, we used all 3704 interactions in EcoMAC. A key parameter for PROM is the binarization threshold value. In our experiments, it was determined by searching for the value when PROM achieved the best performance from 0.01 to 0.9 with the step-size of 0.01 based on the normalized microarray expression values. IDREAM is an improved version of PROM as revised in the Data S1, whose performance relies

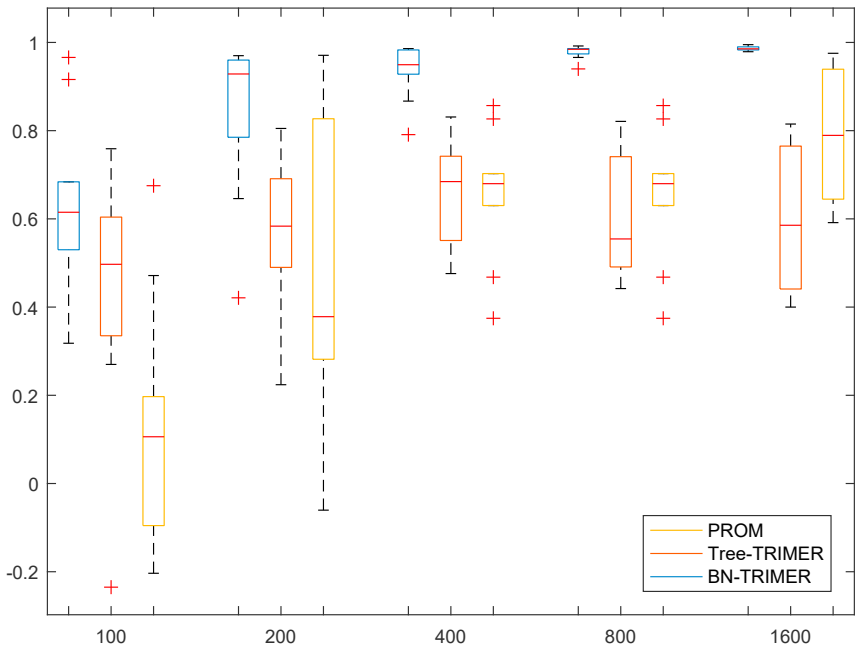


Figure 5. Biomass flux prediction comparison between TRIMER and PROM in the small-scale BN
The bar plot with the average and standard deviation (std) values is based on 10 randomly simulated datasets.

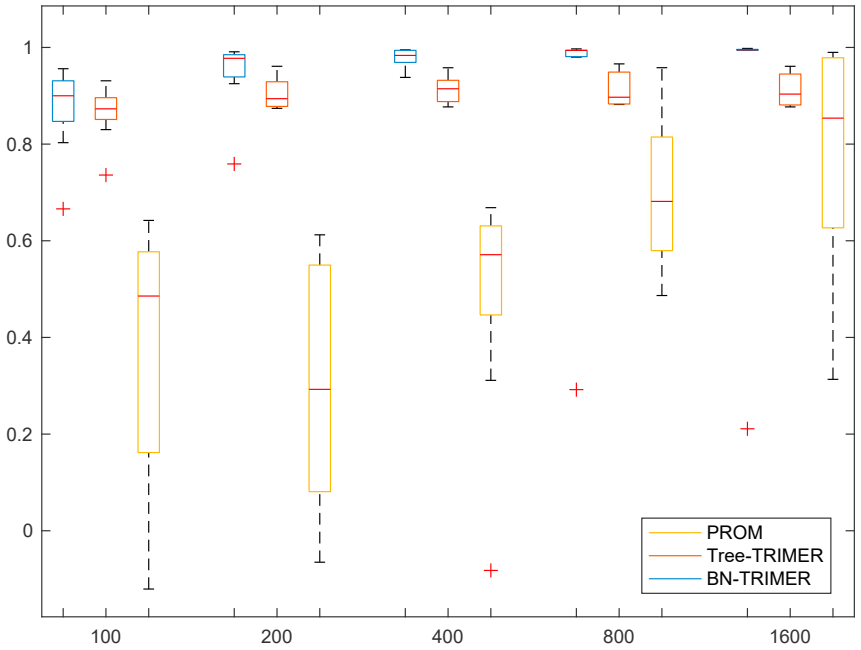


Figure 6. Indole flux prediction comparison between TRIMER and PROM in the small-scale BN
The bar plot with the average and standard deviation (std) values is based on 10 randomly simulated datasets.

heavily on the inferred interactions by EGRIN (Brooks et al., 2014). As neither IDREAM (Wang et al., 2017) nor EGRIN (Brooks et al., 2014) provided the inferred models for *E. coli*, we tried to derive the corresponding models using the data in EcoMAC but put the complete results in the Data S1.

In TRFBA, a constant parameter *C* is used to convert the expression levels to the upper bounds of the metabolic reactions. In our experiments, it was set to be the optimal value when TRFBA achieved the best performance. For TRFBA, we also used all the interactions in EcoMAC directly. In TRIMER, the binarized expression data are taken to infer the corresponding Bayesian network via Tabu search for modeling the TF regulation network using the general BN inference module of TRIMER. For fair comparison, we used the same binarization threshold value as PROM. With the search space restricted to 3704 EcoMAC-archived interactions, Tabu search was ran for one time based on all the expression data as we observed no significant change between learned BNs with or without bootstrapping. A BN with 1409 edges is learned from the EcoMAC expression data. Based on the inferred BN, the conditional probabilities of corresponding gene states when given TF knockouts were computed. Taking these inferred probabilities, the metabolic network flux prediction module with different implementations in TRIMER were adopted to predict biomass and indole fluxes for the corresponding TF knockout mutants.

Run-time

We ran our experiments on a PC with Intel Xeon 6248R processor. It should be noted that the BN structure learning part of TRIMER can be completed within ten minutes with the search space limited to the interaction list. To predict corresponding biomass or indole fluxes for each TF-knockout mutant, it took TRIMER 7.32 s on average on the same PC.

Biomass prediction

We first compared the *in silico* flux predictions by TRIMER, PROM, and TRFBA with the experimentally measured biomass productions in Covert et al. (2004). To compare the two ways of estimating conditional probabilities, we implemented both TRIMER models: we refer the TRIMER model with the conditional probabilities computed in the first or second way as TRIMER-C or TRIMER-B. For the biomass objective, we took Ec_biomass_iAF1260_core_59p81M in iAF1260 as done in PROM. Three FBA formulations, standard FBA, sFBA, and ROOM for TRIMER-C, TRIMER-B, and PROM were implemented, where PROM-sFBA is the original PROM model. In the original TRFBA implementation, the metabolic network flux

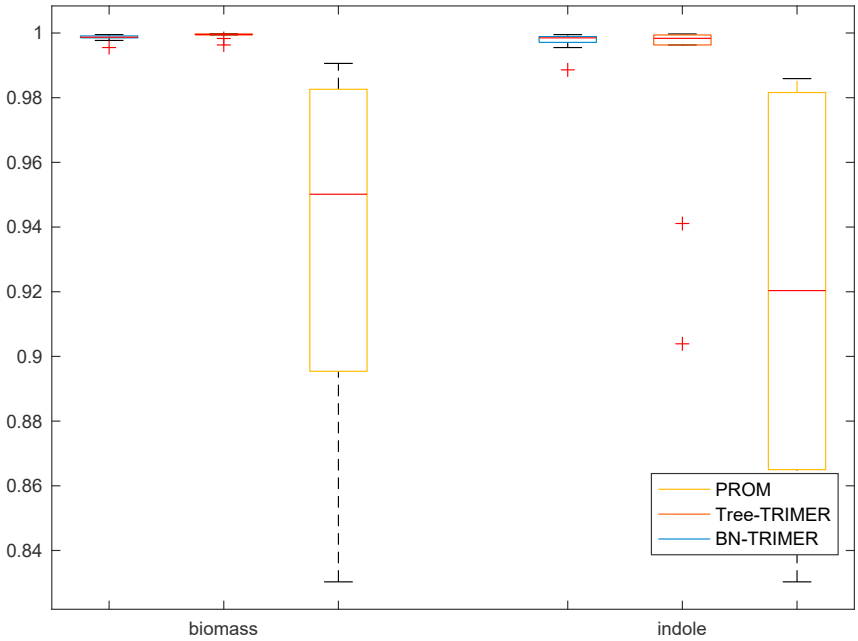


Figure 7. Flux prediction comparison between TRIMER and PROM for double TF knockouts in the simulated large-scale BN

The bar plot with the average and standard deviation (std) values is based on simulated knockout datasets with 50 randomly selected TF pairs.

prediction formulations are based on FBA. In our experiments, the parameters δ and ε in the ROOM formulation were set to be 0.05 and 0.001. Binarization threshold for PROM and TRIMER was set to 0.33, which is used in the original PROM paper. The *C* value in TRFBA is set to 2.30 by searching from 0 to 3 with the step-size of 0.05.

Table 4 provides the comparison of the experimental and predicted fluxes by TRIMER-C, TRIMER-B, PROM, and TRFBA for different TF knockout mutants. It should be pointed out that we used a fixed uptake rate for glucose and oxygen. In this way, simply using the metabolic model has no predictive capability for TF knockout mutants without integrating the change to the reaction regulations due to knockouts and the flux prediction will be the same as that of the wild-type. To illustrate how these hybrid models considering regulations improve over the simple metabolic network model, we have also included the results by simply running FBA, denoted as MET in the table. This makes it more straightforward to compare how these integrated regulatory-metabolic model improve the predictive capability of metabolic-only model denoted as MET. Pearson correlation coefficients (PCC) were computed for performance comparison based on flux predictions for both wild-type and knockout strains. As experimental fluxes were measured under two growth conditions, we also computed PCC between experimental flux ratio and predicted flux ratio, where the flux ratio was obtained by dividing fluxes by the wild-type experimental flux in the corresponding growth condition. The PCC computed with the flux ratios is denoted as rPCC, which can better illustrate how knockout fluxes deviate from wild-type fluxes.

As shown in Table 4, TRIMER-B consistently achieved the highest PCC with the experimentally measured fluxes when compared to PROM and TRIMER-C under three FBA formuations. With sFBA, both TRIMER-C and TRIMER-B performed better than PROM. This shows the superiority of TRIMER over PROM. We can ascribe the overall superiority to the effective modeling of the global dependency in TF regulations through the BN learning and inference in TRIMER, in contrast to using simple conditional probability estimates adopted in PROM. It is notable running FBA only without integrating transcription regulations always output wild-type flux predictions, which gave a high PCC, which did not provide meaningful evaluation. However, when normalized by the corresponding growth conditions, it led to a much lower rPCC. This explains why TRFBA achieved the highest PCC but the corresponding rPCC is only 0.425 since most of its knockout flux predictions were the same as its wild-type ones, which is clearly not desirable. By contrast,

Table 4. Predicted biomass flux comparison for the knockout experiments in Covert et al. (2004). The unit of fluxes is mmol/gDCW/hr.[a,b]

TF KO	Actual	TRIMER-C			TRIMER-B			PROM			MET	TRFBA
		FBA	sFBA	ROOM	FBA	sFBA	ROOM	FBA	sFBA	ROOM	sFBA	FBA
WT +O2	0.710	0.708	0.708	0.708	0.708	0.708	0.708	0.708	0.708	0.708	0.708	0.563
arcA + O2	0.686	0.123	0.631	0.122	0.378	0.610	0.356	0.197	0.272	0.053	0.708	0.563
fnr +O2	0.635	0.391	0.538	0.388	0.381	0.547	0.381	0.399	0.526	0.356	0.708	0.563
arcA fnr +O2	0.648	0.127	0.395	0.055	0.315	0.619	0.298	0.197	0.272	0.015	0.708	0.563
appY +O2	0.636	0.708	0.708	0.671	0.708	0.708	0.671	0.708	0.708	0.671	0.708	0.563
oxyR +O2	0.637	0.708	0.708	0.671	0.708	0.708	0.671	0.708	0.708	0.671	0.708	0.563
soxS +O2	0.724	0.653	0.707	0.652	0.650	0.707	0.650	0.649	0.707	0.649	0.708	0.563
WT -O2	0.485	0.481	0.481	0.481	0.481	0.481	0.481	0.481	0.481	0.481	0.481	0.407
arcA -O2	0.377	0.023	0.023	0.022	0.071	0.071	0.062	0.037	0.037	0.034	0.481	0.355
fnr -O2	0.410	0.266	0.366	0.266	0.259	0.371	0.259	0.139	0.271	0.139	0.481	0.353
arcA fnr -O2	0.301	0.024	0.024	0.019	0.160	0.160	0.154	0.037	0.037	0.023	0.481	0.356
appY -O2	0.476	0.481	0.481	0.456	0.481	0.481	0.456	0.481	0.481	0.456	0.481	0.354
oxyR -O2	0.481	0.481	0.481	0.456	0.481	0.481	0.456	0.481	0.481	0.456	0.481	0.357
soxS -O2	0.465	0.443	0.481	0.443	0.442	0.481	0.442	0.441	0.479	0.441	0.481	0.355
PCC	–	0.538	0.870	0.517	0.700	0.906	0.702	0.619	0.693	0.484	0.918	0.927
rPCC	–	0.684	0.851	0.679	0.723	0.841	0.732	0.725	0.770	0.653	0.282	0.425

[a]In the FBA formulations, substrate (glucose) and oxygen uptake rates for aerobic conditions are set to be 8.5 and 14.6 mmol/gDCW/hr, respectively. They are set to 20.8 and 0 mmol/gDCW/hr for anaerobic conditions.
[b]The optimization is by the CPLEX solver.

PCC and rPCC for TRIMER and PROM are consistent and TRIMER-sFBA achieved the highest rPCC. In summary, the predictions by TRIMER and PROM are more reliable and TRIMER performs the best for capturing the varying patterns of knockout fluxes.

Indole flux prediction

We further validated the predicted fluxes by TRIMER with our experimentally-generated data from TF-knockout experiments for indole production as described in Section STAR Methods. As we generated all these experimental data under the same growth condition, we only report PCC for performance comparison in this set of experiments. We used the same parameters as in the previous experiment for all the models and took TRPS3 in iAF1260 for indole flux prediction. Table 5 provides the comparison of the experimental and predicted fluxes by TRIMER-C, TRIMER-B, PROM, and TRFBA for different TF knockout mutants grown in M9 minimal media and the overall PCCs between experimental and predicted fluxes. In this set of experiments, TRIMER-B achieved consistently better correlation with the experimental results with all three formulations compared to TRIMER-C and PROM. It should be also noted that TRIMER with the ROOM formulation has achieved the highest correlation values, which were significantly better than the other FBA formulations for both TRIMER and PROM. TRFBA also achieved similar PCC as TRIMER with the ROOM formulation.

In the Data S1, we include additional results based on more comprehensive performance comparison with other existing hybrid modeling methods, including PROM (Chandrasekaran and Price, 2010), IDREAM (Wang et al., 2017), and TRFBA (Motamedian et al., 2017), tested on both *E. coli* and yeast models. Our results clearly show that TRIMER consistently outperforms other methods, where the trends are similar to those observed in the experiments above.

DISCUSSION

Based on the performance comparison results with the presented experiments, we have shown that BN modeling transcription regulations in TRIMER can capture regulation relationships better and improve metabolic flux predictions for knockout mutants compared to the relative frequency based conditional probability estimation in PROM and its extensions. In addition, TRIMER does not require significant tuning

Table 5. Predicted indole flux comparison for our TF knockout (KO) experiments in M9 minimal media. The unit of fluxes is mmol/gDCW/hr.[a,b]

TF KO	Actual	TRIMER-C			TRIMER-B			PROM			TRFBA
		FBA	sFBA	ROOM	FBA	sFBA	ROOM	FBA	sFBA	ROOM	FBA
Fnr	0.0427	0.0231	0.0293	0.0427	0.0216	0.0301	0.0427	0.0224	0.0295	0.0427	0.0100
soxS	0.0387	0.0366	0.0397	0.0386	0.0364	0.0397	0.0374	0.0365	0.0397	0.0367	0.0100
Crp	0.0397	0.0197	0.0197	0.0383	0.0193	0.0200	0.0367	0	0	0.0367	0.0100
lysR	0.0400	0.0372	0.0372	0.0392	0.0370	0.0370	0.0380	0.0370	0.0370	0.0370	0.0100
fucR	0.0390	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
Mali	0.0403	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
phoB	0.0390	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
cpxR	0.0393	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
creB	0.0383	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
trpB	0	0	0	0	0	0	0	0	0	0	0
trpD	0	0	0	0	0	0	0	0	0	0	0
trpE	0	0	0	0	0	0	0	0	0	0	0
paaX	0.0393	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
trpA	0	0	0	0	0	0	0	0	0	0	0
tnaA	0.0380	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
trpL	0.0393	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
tnaC	0.0397	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
tnaB	0.0400	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
dhaR	0.0403	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0397	0.0397	0.0377	0.0100
PCC	–	0.9270	0.9448	0.9988(7)	0.9203	0.9478	0.9988(8)	0.8305	0.8481	0.9987	0.9983

[a]In the FBA formulations, substrate (glucose) and oxygen uptake rates are set to be 9.5 mmol/gDCW/hr and 13.0 mmol/gDCW/hr, respectively.
[b]The optimization is by the CPLEX solver.

and is more user-friendly when being implemented for different model organisms, prediction tasks, and/or expression data. In contrast, many existing hybrid models often require careful tuning when being applied to different models, tasks and data, which can be challenging, time-consuming, and requiring both biology and modeling expertise. One of important challenges in applying developed computational systems biology packages to solve real-world problems is to develop more flexible and user-friendly frameworks and tools, for which we have tried to consider when developing the TRIMER package.

Limitations of the study

Due to the inherent stochasticity and complexity of living systems, accurate inference of the transcription regulatory network model as well as the metabolic network model is practically challenging, especially when studying non-model organisms other than *E. coli* or yeast studied in this paper. In order to better capture the potential model uncertainty and be able to make reliable predictions in the presence of substantial uncertainty, we may have to deal with an uncertainty class of network models – rather than a single best model – that are consistent with the available data and knowledge. This also enables closed-loop experimental design, where new experiments may be designed to reduce model uncertainty, the outcomes of the designed experiments may be used to update the uncertainty class, and where this experimental loop may be repeated. We leave this for our future research.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.103218.

ACKNOWLEDGMENTS

The materials presented in this paper are based upon the work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under contract number DE-SC0012704. P.N. and X.Q. are partially supported by the National Science Foundation under Grant CCF-1553281.

AUTHOR CONTRIBUTIONS

Conceptualization, B.J.Y., E.R.D., F.J.A., I.B., and X.Q.; Methodology, P.N., B.J.Y., and X.Q.; Investigation, P.N., M.J.S., B.J.Y., E.R.D., F.J.A., I.B., and X.Q.; Formal Analysis, P.N., B.J.Y., E.R.D., F.J.A., and X.Q.; Software: P.N.; Experimental validation, M.J.S. and I.B.; Writing - Original Draft, P.N., M.J.S., I.B., and X.Q.; Writing - Review & Editing, P.N., M.J.S., B.J.Y., E.R.D., F.J.A., I.B., and X.Q.; Funding Acquisition, B.J.Y., E.R.D., F.J.A., I.B., and X.Q.; Resources, F.J.A., I.B., and X.Q.; Supervision, I.B. and X.Q.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 19, 2021
Revised: August 22, 2021
Accepted: September 29, 2021
Published: November 19, 2021

REFERENCES

Apaydin, M., Xu, L., Zeng, B., and Qian, X. (2017). Robust mutant strain design by pessimistic optimization. BMC Genomics *18*, 677.

Arkin, A., Cottingham, R., Henry, C., Harris, N., Stevens, R., Maslov, S., et al. (2018). KBase: the United States department of Energy systems biology knowledgebase. Nat. Biotechnol. *36*, 566. https://doi.org/10.1038/nbt.4163.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K., Tomita, M., Wanner, B., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*, 2006.0008.

Bachmann, B.J. (1972). Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. *36*, 525–557.

Barrett, C.L., Kim, T.Y., Kim, H.U., Palsson, B.Ø., and Lee, S.Y. (2006). Systems biology as a foundation for genome-scale synthetic biology. Curr. Opin. Biotechnol. *17*, 488–492.

Bonneau, R., Reiss, D.J., Shannon, P., Facciotti, M., Hood, L., Baliga, N.S., and Thorsson, V. (2006). The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets *de novo*. Genome Biol. *7*, 1–16.

Brooks, A.N., Reiss, D.J., Allard, A., Wu, W.J., Salvanha, D.M., Plaisier, C.L., Chandrasekaran, S., Pan, M., Kaur, A., and Baliga, N.S. (2014). A system-level model for the microbial regulatory genome. Mol. Syst. Biol. *10*, 740.

Burgard, A.P., Pharkya, P., and Maranas, C.D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. *84*, 647–657.

Carrera, J., Estrela, R., Luo, J., Rai, N., Tsoukalas, A., and Tagkopoulos, I. (2014). An integrative, multi-scale, genome-wide model reveals the phenotypic landscape of *Escherichia coli*. Mol. Syst. Biol. *10*, 735. https://doi.org/10.1186/1752-0509-5-147.

Chandrasekaran, S., and Price, N.D. (2010). Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. *107*, 17845–17850. https://www.pnas.org/content/107/41/17845. https://www.pnas.org/content/107/41/17845.full.pdf.

Chandrasekaran, S., and Price, N.D. (2013). Metabolic constraint-based refinement of transcriptional regulatory networks. PLoS Comput. Biol. *9*, e1003370.

Chant, E.L., and Summers, D.K. (2007). Indole signalling contributes to the stable maintenance of *Escherichia coli* multicopy plasmids. Mol. Microbiol. *63*, 35–43.

Chow, C., and Liu, C. (1968). Approximating discrete probability distributions with dependence trees. IEEE Trans. Inf. Theor. *14*, 462–467.

Covert, M.W., Knight, E.M., Reed, J.L., Herrgard, M.J., and Palsson, B.Ø. (2004). Integrating high-throughput and computational data elucidates bacterial networks. Nature *429*, 92–96.

Covert, M.W., and Palsson, B.O. (2003). Constraints-based models: regulation of gene expression reduces the steady-state solution space. J. Theor. Biol. *221*, 309–325. https://doi.org/10.1006/jtbi.2003.3071.

Covert, M.W., Xiao, N., Chen, T.J., and Karr, J.R. (2008). Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. Bioinformatics *24*, 2044–2050. https://doi.org/10.1093/bioinformatics/btn352. https://academic.oup.com/bioinformatics/article-pdf/24/18/2044/29153528/btn352.pdf.

Edwards, J.S., and Palsson, B.Ø. (2000). The *Escherichia coli* MG1655 *in silico* metabolic

genotype: its definition, characteristics, and capabilities. Proc. Natl. Acad. Sci. *97*, 5528–5533.

Fendt, S.M., Oliveira, A.P., Christen, S., Picotti, P., Dechant, R.C., and Sauer, U. (2010). Unraveling condition-dependent networks of transcription factors that control metabolic pathway activity in yeast. Mol. Syst. Biol. *6*, 432.

Haury, A.C., Mordelet, F., Vera-Licona, P., and Vert, J.P. (2012). TIGRESS: trustful inference of gene regulation using stability selection. BMC Syst. Biol. *6*, 1–17.

Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S.N., Richelle, A., Heinken, A., Haraldsdóttir, H.S., Wachowiak, J., Keating, S.M., Vlasov, V., Magnusdóttir, S., Ng, C.Y., Preciat, G., Žagare, A., Chan, S.H.J., Aurich, M.K., Clancy, C.M., Modamio, J., Sauls, J.T., Noronha, A., Bordbar, A., Cousins, B., El Assal, D.C., Valcarcel, L.V., Apaolaza, I., Ghaderi, S., Ahookhosh, M., Ben Guebila, M., Kostromins, A., Sompairac, N., Le, H.M., Ma, D., Sun, Y., Wang, L., Yurkovich, J.T., Oliveira, M.A.P., Vuong, P.T., El Assal, L.P., Kuperstein, I., Zinovyev, A., Hinton, H.S., Bryant, W.A., Aragón Artacho, F.J., Planes, F.J., Stalidzans, E., Maass, A., Vempala, S., Hucka, M., Saunders, M.A., Maranas, C.D., Lewis, N.E., Sauter, T., Palsson, B.Ø., Thiele, I., and Fleming, R.M.T. (2019). Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nat. Biotechnol. *14*, 1750–2799. https://doi.org/10.1038/s41596-018-0098-2.

Højsgaard, S., et al. (2012). Graphical independence networks with the gRain package for R. J. Stat. Softw. *46*, 1–26.

Irrthum, A., Wehenkel, L., Geurts, P., et al. (2010). Inferring regulatory networks from expression data using tree-based methods. PLoS One *5*, e12776.

Jensen, P., Lutz, K., and Papin, J. (2011). TIGER: toolbox for integrating genome-scale metabolic models, expression data, and transcriptional regulatory networks. BMC Syst. Biol. *5*. https://doi.org/10.1186/1752-0509-5-147.

Koller, D., and Friedman, N. (2009). Probabilistic Graphical Models: Principles and Techniques (MIT press).

Lewis, N.E., Hixson, K.K., Conrad, T.M., Lerman, J.A., Charusanti, P., Polpitiya, A.D., Adkins, J.N., Schramm, G., Purvine, S.O., Lopez-Ferrer, D., Weitz, K.K., Eils, R., König, R., Smith, R.D., and Palsson, B.Ø. (2010). Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. Mol. Syst. Biol. *6*, 390. https://doi.org/10.1038/msb.2010.47. https://www.embopress.org/doi/abs/10.1038/msb.2010.47. https://www.embopress.org/doi/pdf/10.1038/msb.2010.47.

Machado, D., and Herrgård, M. (2014). Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. PLOS Comput. Biol. *10*, e1003580. https://doi.org/10.1371/journal.pcbi.1003580.

Mahadevan, R., and Schilling, C.H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab. Eng. *5*, 264–276. https://www.sciencedirect.com/science/article/pii/S1096717603000582.

Motamedian, E., Mohammadi, M., Shojaosadati, S.A., and Heydari, M. (2017). TRFBA: an algorithm to integrate genome-scale metabolic and transcriptional regulatory networks with incorporation of expression data. Bioinformatics *33*, 1057–1063.

Niu, P., Soto, M.J., Yoon, B.J., Dougherty, E.R., Alexander, F.J., Blaby, I., and Qian, X. (2021). TRIMER GitHub Repository. http://github.com/niupuhua1234/TRIMER.

Nagarajan, R., Scutari, M., and Lèbre, S. (2013). Bayesian Networks in R with Applications in Systems Biology 2013 (New York, NY: Springer-Verlag).

Palsson, B. (2015). Systems Biology (Cambridge University Press).

Reed, J.L. (2017). Genome-scale metabolic modeling and its application to microbial communities. In The Chemistry of Microbiomes: Proceedings of a Seminar Series (National Academies Press), pp. 85–91.

Reiss, D.J., Plaisier, C.L., Wu, W.J., and Baliga, N.S. (2015). cMonkey2: automated, systematic, integrated detection of co-regulated gene modules for any organism. Nucleic Acids Res. *43*, e87.

Ren, S., Zeng, B., and Qian, X. (2013). Adaptive bi-level programming for optimal gene knockouts for targeted overproduction under phenotypic constraints. BMC Bioinformatics *14*, S17.

Russell, S., and Norvig, P. (2002). Artificial Intelligence: A Modern Approach (Prentice Hall).

Santos-Zavaleta, A., Salgado, H., Gama-Castro, S., Sánchez-Pérez, M., Gómez-Romero, L., Ledezma-Tejeida, D., García-Sotelo, J.S., Alquicira-Hernández, K., Muñiz Rascado, L.J., Peña Loredo, P., Ishida-Gutiérrez, C., Velázquez-Ramírez, D.A., Moral-Chávez, V.D., Bonavides-Martínez, C., Méndez-Cruz, C.F., Galagan, J., and Collado-Vides, J. (2019). RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli K-12*. Nucleic Acids Res. *47*, D212–D220. https://doi.org/10.1093/nar/gky1077.

Scutari, M., and Nagarajan, R. (2013). Identifying significant edges in graphical models of molecular networks. Artif. Intelligence Med. *57*, 207–217. https://doi.org/10.1016/j.artmed.2012.12.006. https://www.sciencedirect.com/science/article/pii/S0933365712001546.

Segre, D., Vitkup, D., and Church, G.M. (2002). Analysis of optimality in natural and perturbed metabolic networks. Proc. Natl. Acad. Sci. *99*, 15112–15117.

Shen, F., Sun, R., Yao, J., Li, J., Liu, Q., Price, N.D., Liu, C., and Wang, Z. (2019). OptRAM: *in-silico* strain design via integrative regulatory-metabolic network modeling. PLOS Comput. Biol. *15*, 1–25. https://doi.org/10.1371/journal.pcbi.1006835.

Shlomi, T., Berkman, O., and Ruppin, E. (2005). Regulatory on/off minimization of metabolic flux changes after genetic perturbations. Proc. Natl. Acad. Sci. *102*, 7695–7700.

Shlomi, T., Cabili, M.N., Herrgård, M.J., Palsson, B.Ø., and Ruppin, E. (2008). Network-based prediction of human tissue-specific metabolism. Nat. Biotechnol. *26*, 1546–1696. https://doi.org/10.1038/nbt.1487.

Shlomi, T., Eisenberg, Y., Sharan, R., and Ruppin, E. (2007). A genome-scale computational study of the interplay between transcriptional regulation and metabolism. Mol. Syst. Biol. *3*, 101. arXiv. https://doi.org/10.1038/msb4100141. https://www.embopress.org/doi/abs/10.1038/msb4100141. https://www.embopress.org/doi/pdf/10.1038/msb4100141.

Varma, A., and Palsson, B.Ø. (1994a). Metabolic flux balancing: basic concepts, scientific and practical use. Bio/technology *12*, 994–998.

Varma, A., and Palsson, B.Ø. (1994b). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. Appl. Environ. Microbiol. *60*, 3724–3731.

Wang, Z., Danziger, S.A., Heavner, B.D., Ma, S., Smith, J.J., Li, S., Herricks, T., Simeonidis, E., Baliga, N.S., Aitchison, J.D., and Price, N.D. (2017). Combining inferred regulatory and reconstructed metabolic networks enhances phenotype prediction in yeast. PLOS Comput. Biol. *13*, 1–23. https://doi.org/10.1371/journal.pcbi.1005489.

Young, I.T. (1977). Proof without prejudice: use of the Kolmogorov-Smirnov test for the analysis of histograms from flow systems and other sources. J. Histochem. Cytochem. *25*, 935–941.

Yu, H., and Blair, R.H. (2019). Integration of probabilistic regulatory networks into constraint-based models of metabolism with applications to Alzheimer's disease. BMC Bioinformatics *20*, 386.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Crp	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100050	JW5702
tnaA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100051	JW3686
lldR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100052	JW3579
glcC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100053	JW2947
nadR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100054	JW5800
relB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100055	JW1556
fabR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100056	JW3935
arsR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100057	JW3468
acrR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100058	JW0453

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
fhlA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100059	JW2701
araC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100060	JK0063
marR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100061	JW5248
uxuR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100062	JW4287
metR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100063	JW3804
Cytr	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100064	JW3905
dsdX	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100065	JW2362
allR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100066	JW0494
exuR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100067	JW3065
soxR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100068	JW3024

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
gadX	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100069	JW3484
Fnr	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100070	JW1328
mprA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100071	JW2659
nanR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100072	JW3195
stpA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100073	JW2644
nhaR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100074	JW0019
kdgR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100075	JW1816
idnR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100076	JW4221
soxR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100068	JW3024
melR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100077	JW4079

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ada	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100078	JW2201
metJ	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100079	JW3909
mlrA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100080	JW2115
galS	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100081	JW2138
tyrR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100082	JW1316
ilvY	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100083	JW3746
xapR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100084	JW2396
zntR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100085	JW3254
rstA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100086	JW1600
nagC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100087	JW0662

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
csiR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100088	JW2639
hupA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100089	JW3964
trpA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100090	JW1252
leuO	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100091	JW0075
ebgR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100092	JW3046
lacI	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100093	JW0336
E. coli K-12 BW25113 (WT)	Bachmann B. J. (1972). Pedigrees of some mutant strains of Escherichia coli K-12. *Bacteriological reviews*, *36*(4), 525–557. https://doi.org/10.1128/br.36.4.525-557.1972	BW25113
soxS	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100095	JW4023
rtcR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100096	JW3385
rbsR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100097	JW3732

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
narL	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100098	JW1212
gntR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100099	JW5946
chbR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100100	JW1724
lysR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100101	JW2807
glnG	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100102	JW3839
Lrp	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100103	JW0872
CbI	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100104	JW1966
rhaS	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100105	JW3876
sgrR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100106	JW0068
yehT	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100107	JW5352

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
envR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100108	JW3232
glpR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100109	JW3386
fadR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100110	JW1176
xylR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100111	JW3541
uidR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100112	JW1610
torR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100113	JW0980
oxyR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100114	JW3933
rhaR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100115	JW3877
rpiR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100116	JW4050
appY	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100117	JW0553

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
creB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100118	JW4361
hcaR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100119	JW2521
slyA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100120	JW5267
prpR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100121	JW0322
feaR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100122	JW1379
srlR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100123	JW2676
dcuR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100124	JW4085
argP	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100125	JW2883
caiF	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100126	JW0033
Nac	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100127	JW1967

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
cynR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100128	JW5894
betI	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100129	JW0305
aidB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100130	JW5867
Fis	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100131	JW3248
adiY	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100132	JW4077
trpC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100133	JW1254
galR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100134	JW2805
fliZ	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100135	JW1906
argR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100136	JW3206
tdcA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100137	JW3089

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
evgA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100138	JW2366
gadE	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100139	JW3480
gadW	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100140	JW3483
gutM	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100141	JW2675
cdaR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100142	JW5013
ycfQ	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100143	JW5159
arcA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100144	JW4364
marA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100145	JW5249
bglJ	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100146	JW5955
trpL	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100147	JW1257

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
treR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100148	JW4200
phoP	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100149	JW116
iscR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100150	JW2515
paaX	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100151	JW1394
Fur	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100152	JW0669
tnaC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100153	JW3685
mngR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100154	JW0719
cueR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100155	JW0476
Rob	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100156	JW4359
tnaB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100157	JW5619

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
csgD	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100158	JW1023
asnC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100159	JW3721
cadC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100160	JW4094
mhpR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100161	JW0337
yeiL	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100162	JW2150
cspA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100163	JW3525
kdpE	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100164	JW5096
gatR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100165	JW2074
bolA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100166	JW5060
norR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100167	JW5843

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
sdiA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100168	JW1901
malI	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100169	JW1612
purR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100170	JW1650
lrhA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100171	JW2284
Zur	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100172	JW5714
narP	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100173	JW2181
basR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100174	JW4074
alaS	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100175	JW2667
atoC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100176	JW2214
envY	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100177	JW0555

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
phoB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100178	JW0389
uhpA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100179	JW3644
fucR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100180	JW2776
malT	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100181	JW3381
hdfR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100182	JW5607
pdhR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100183	JW0109
gcvA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100184	JW2779
zraR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100185	JW3968
trpR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100186	JW4356
cusR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100187	JW0560

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
hyfR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100188	JW2476
baeR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100189	JW2064
deoR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100190	JW0824
yqhC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100191	JW5849
pepA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100192	JW4217
ompR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100193	JW3368
yiaJ	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100194	JW3546
tdcR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100195	JW5525
yjiE	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100196	JW4290
cpxR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100197	JW3883

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
hipB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100198	
ascG	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100199	JW1501
putA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100200	JW0999
dinJ	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100201	JW0216
qseB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100202	JW2993
agaR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100203	JW3100
trpD	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100204	JW1255
trpE	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100205	JW1256
iclR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100206	JW3978
dhaR	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100207	JW5188

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
cysB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100208	JW1267
ihfA	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100209	JW1702
Hns	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100210	JW1225
flhC	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100211	JW1880
trpB	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100212	JW1253
yqjI	Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008. https://doi.org/10.1038/msb4100213	JW3042
Software and algorithms		
TRIMER	This paper	https://github.com/niupuhua1234/TRIMER
COBRA	Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S.N., Richelle, A., Heinken, A., Haraldsdottir, H.S., Wachowiak, J.,Keating, S.M., Vlasov, V., Magnusdottir, S., Ng, C.Y., Preciat, G., Zagare, A., Chan, S.H.J., Aurich, M.K., Clancy, C.M., Modamio, J., Sauls, J.T., Noronha, A., Bordbar, A., Cousins, B., El Assal, D.C., Valcarcel, L.V., Apaolaza, I., Ghaderi, S., Ahookhosh, M., Ben Guebila, M., Kostromins, A., Sompairac, N., Le, H.M., Ma, D., Sun, Y., Wang, L., Yurkovich, J.T., Oliveira, M.A.P., Vuong, P.T., El Assal, L.P., Kuperstein, I., Zinovyev, A., Hinton, H.S., Bryant, W.A., Aragon Artacho, F.J., Planes, F.J., Stalidzans, E., Maass, A., Vempala, S., Hucka, M., Saunders, M.A., Maranas, C.D., Lewis, N.E., Sauter, T., Palsson, B.Ø., Thiele, I., Fleming, R.M.T., 2019. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nature Biotechnology 14,1750–2799. https://doi.org/10.1038/s41596-018-0098-2, https://doi.org/10.1038/s41596-018-0098-2.	https://opencobra.github.io/cobratoolbox/stable/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bnlearn	Nagarajan, R., Scutari, M., Lebre, S. Bayesian Networks in R with Applications in Systems Biology 2013. New York, NY; Springer-Verlag.	https://www.bnlearn.com/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Xiaoning Qian (email: xqian@ece.tamu.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The developed package TRIMER is available online in an open-source GitHub repository (Niu et al., 2021). All the implemented functions in TRIMER are documented using MATLAB's help function. Code for the reported experiments in this paper can also be found in the referred GitHub repository.

Experimental data for total indole concentrations of the TF-knockout deletants and the parental strain in LB and M9 media are provided in Tables S1 and S2 as excel files, respectively.

METHOD DETAILS

We introduce the main components of TRIMER organized in two major modules—namely, the *transcription regulation* network module and the *metabolic regulation* module—that are integrated within a unified interacting framework (Figure 1). The proposed hybrid model enables condition-dependent transcriptomic and metabolic predictions for both wild-type and TF-knockout mutant strains, through general Bayesian network (BN) modeling of transcriptional regulations. We also provide the details of our TF knockout experiments from which the experimentally observed fluxes validate the *in silico* flux predictions made by TRIMER.

Notations

As TRIMER involves multiple variables in the transcriptional-metabolic hybrid model, we summarize the corresponding math notations in Table 1.

Transcription regulation inference in TRIMER

Gene expression data preprocessing. The gene expression data need to be discretized for BN learning as in TRIMER, the *TF-Regulated gene Network (TRN)* concerns 'ON/OFF' states of TFs and genes in the network. In our implementation, quantile normalization is first applied to raw data. Then the threshold for a given quantile value is computed and data is binarized according to the threshold. The choice of the quantile value can be either set manually or be similarly determined as in PROM (Chandrasekaran and Price, 2010). In other words, we search for the best value based on the prediction performance of the learned BN. Based on the results of our experiments, the suggested quantile value for thresholding is in the range of [0.3,0.4].

BN learning. The key component of TRIMER is to model the genome-scale TF-regulated gene network by a Bayesian network (BN) learned from discretized gene expression. This BN is expected to capture the interactions between regulators (TFs) and target genes. For this purpose, we have integrated bn-learn, a Bioconductor package for Bayesian network modeling of biological networks (Nagarajan et al., 2013). A naive way to learn a BN from available observed gene states is to search over the space of all possible directed acyclic graphs (DAGs) and identify the one that optimizes a given objective function evaluating the goodness of fit. However, the search space of BN model structures grows exponentially with the number of variables (nodes in the BN). Without restricting the BN structures, the BN learning can easily become

infeasible even when considering only a dozen variables.In our experiments, we implemented two structure learning strategies, tree-based search for learning tree-based BN in a restricted family of Chow-Liu trees (Chow and Liu, 1968) and Tabu search (Russell and Norvig, 2002), a greedy algorithm for learning general BNs that incorporate prior knowledge of gene-gene interactions. After finding the desired BN structure, BN model parameters are estimated by maximum likelihood estimates (MLEs).

In TRIMER, we have implemented Chow-Liu-tree-based BN learning for tree-based search. For learning general BNs, Tabu search, a modified greedy hill-climbing optimization strategy, is implemented in bn-learn as the search method based on a chosen score function, for example, either Bayesian information criterion (BIC) or Akaike information criterion (AIC). In our implementation, we further explore the proposed bootstrap resampling in Scutari and Nagarajan (2013) to learn a more robust structure. Specifically, we search for high-score BN structures by bootstrapping multiple expression samples from the given total samples (simulated or from expression databases). The inferred edges present in at least N% of the learned BNs are finally included in the final structure. N is a threshold value, which is determined automatically as described in Scutari and Nagarajan (2013). Such a model averaging strategy helps to establish the significance of the edges in the final "average" structure for robustness against the potential data uncertainty and scarcity.

We further note that in our experiments, to restrict the search space of general BN structure learning, only experimentally confirmed gene-gene interactions are considered as candidate edges in BNs. For *E. coli*, we have employed the interactions archived in RegulonDB Santos-Zavaleta et al. (2019). When needed, separate interaction inference and validation methods, such as GENIE3 (Irrthum et al., 2010), TIGRESS (Haury et al., 2012), orInferelator (Bonneau et al., 2006), can also serve as the prior knowledge to extend the search space for structure learning.

Gene state inference. Once we have learned a BN, we can infer all the relevant conditional probabilities in the form of Pr(gene(s)|TF(s)) = $p(\vec{g}\,|\,\vec{TF})$ that regulate the genome-scale metabolic network (iAF1260 for *E. coli* for example) so that for TF-knockout mutants, the conditional probabilities Pr(gene(s)|TF(s)) can model the effect of TF knockouts over the regulated target genes and, therefore, the corresponding metabolic fluxes at the genome scale. To do that in TRIMER without incurring high computational cost to exhaust all potential Pr(gene(s)|TF(s)) for metabolism regulation, we only focus on the TF-target interaction list to determine which genes can be affected (annotated as target genes) when one or multiple TFs are knocked out. Generally speaking, due to potential I-equivalent classes when learning BNs from data, determining the exact causal relationships from the learned BN structure is difficult. We rely on the annotated TF-gene interaction list (in RegulonDB for example). The Kolmogorov-Smirnov test (Young, 1977) is performed to select significantly coupled TF-target pairs in the interaction list. Then the filtered list is further pruned by removing the pairs that are d-separated in the learned BN. In cases when multiple TFs are knocked out at the same time, the list of the affected genes is the union of the target gene lists corresponding to each knockout TF in TRIMER. In addition, we only care about the probabilities that will affect metabolic reactions so that only the target genes that are associated with the metabolic reactions as described by the gene-protein reaction (GPR) rules will be considered. Given this pruned interaction list, TRIMER infers corresponding conditional probabilities by BN inference algorithms. In TRIMER, exact inference is performed by the integrated package gRain (Højsgaard et al., 2012) and approximate inference in bn-learn (Nagarajan et al., 2013) can also be directly utilized for computational efficiency.

Metabolic flux prediction in TRIMER

The workflow of connecting the BN inference and metabolic flux prediction modules is illustrated in the flowchart shown in Figure 2. We first briefly review the Flux Balance Analysis (FBA) framework and then detail the corresponding TRIMER implementations for the constituting module components for the metabolic flux prediction module.

Flux balance analysis (FBA). Since it has been proposed in Varma and Palsson (1994a, b); Palsson (2015), FBA, as a simplified network analysis model for metabolic flux analysis, has been widely adopted for steady-state flux analyses by assuming the balance of production and consumption fluxes of metabolic reaction network models. Mathematically, with the prior stoichiometry knowledge, FBA assumes that the weighted sum of reaction fluxes, denoted by the vector $\vec{v}$, based on calibrated stoichiometric coefficients

S, is 0: $S\vec{v} = 0$. Such a steady-state flux analysis can be performed by assuming that the corresponding wild-type microbial species always optimizes for its growth:

$$\begin{aligned} &\max_{\vec{v}} && \text{biomass}(\vec{v}) \\ &\text{s.t} && S\vec{v} = 0; \\ & && lb_i \le v_i \le ub_i, \quad \forall i \in \{1, \ldots, m\}, \end{aligned}$$

where v_i, $1 \le i \le m$ denotes the flux value for the ith metabolic reaction of the total m reactions in the metabolic network, and S an $m \times n$ stoichiometric matrix involving all the n metabolites in the given metabolic reaction network model. The biomass production flux: $\text{biomass}(\vec{v}) = \sum_{j \in I_{\text{biom}}} c_j v_j$ is based on the annotated set of reaction indices, I_{biom}, involving the metabolite precursors that contribute to the biomass production in FBA with the corresponding given weights c_j (Segre et al., 2002). Each reaction flux is bounded by the corresponding lower and upper bounds lb_i and ub_i.

For wild-type microbial strains, a common assumption is that their steady-state flux values follow an optimal distribution that maximizes the biomass production rate. The steady-state flux distribution can be approximately solved as a linear programming (LP) problem to maximize the biomass production flux subject to the FBA stoichiometry constraints as the above formulation.

However, when modeling mutant strains, the researchers found that the biomass maximization assumption for wild-type strains may not approximate the steady-state fluxes well. To achieve better agreement with experimental observations, approximation formulations of knockout metabolic fluxes undergoing a minimization of metabolic adjustment (MOMA) process (Segre et al., 2002) or by the regulatory on/off minimization (ROOM) (Shlomi et al., 2005) have been proposed to address the long-term post knockout metabolic flux distribution predication problem. We will detail the corresponding implementations in TRIMER in the following subsections.

Construct transcriptional constraints over flux variables. The first module component for metabolic flux prediction is to integrate transcriptional changes into metabolic network modeling. Metabolism regulation in TRIMER is achieved by integrating the inferred conditional probabilities under different conditions from the BN to construct constraints for the corresponding metabolic reaction fluxes according to the GPR rules. From the BN learning and inference module, we derive a list of conditional probabilities associated with the corresponding metabolic reactions in the metabolic network model. Similar as in PROM, these probabilities together with the fluxes bounds estimated via flux variability analysis (FVA) (Mahadevan and Schilling, 2003) are used to constrain the reaction flux bounds through GPR rules. FVA helps determine alternative optimal solutions for the constraint-based linear programming formulation of FBA by screening the corresponding polygon boundaries of the feasible solution space, which identifies the minimum and maximum possible fluxes through a reaction in the metabolic model. To integrate transcription regulation into the metabolic models, GPR rules are represented as Boolean expressions associated with corresponding reactions to describe the nonlinear relationships between genes and reactions. In TRIMER, we have implemented a general platform as in TIGER (Jensen et al., 2011) to convert the conditional probability values into linear constraints over flux variables and integrate them with the metabolic model in COBRA (Heirendt et al., 2019) for flux prediction.

We have adopted two ways to derive the updated reaction flux constraints according to the two ways of inferring conditional probabilities based on the learned BN.

The first way is the same as the one adopted in PROM. Suppose there are M genes, $G = \{g_1, \ldots, g_m, \ldots, g_M\}$, that are regulated by the corresponding TF(s). Then via the provided GPR rules in the COBRA model, we can find the corresponding affected reactions denoted as $R = \{r_1, \ldots, r_n, \ldots, r_N\}$. For each $r_n \in R$, we can find a subset of regulating genes in G, denoted as $G(r_n)$, based on the corresponding GPR rules. With the corresponding TF knockout mutants, the reaction flux bounds are then adjusted in the following way:

$$\begin{aligned} ub_{r_n} &= \min_{g \in G(r_n)} \{p(g = 1 | TF = 0)\} \times v_{max}(r_n); \\ lb_{r_n} &= \min_{g \in G(r_n)} \{p(g = 1 | TF = 0)\} \times (-v_{max}(r_n)), \end{aligned} \qquad \text{(Equation 1)}$$

where $v_{max}(r)$ is estimated by FVA for reaction r. An example is given in the Data S1 to illustrate this operation.

In TRIMER, we have also implemented a more general way for integrating both the probabilities and the GPR rules into the flux constraints, so we can obtain the joint probabilities of the states of multiple genes regulating the same reaction, instead of simply combining the conditional probabilities for individual genes in the heuristic manner as in the previous approach. The reaction flux bounds can be set by directly multiplying the maximum flux with the sum of all probabilities with the corresponding gene states that affect the corresponding reaction according to the GPR rules:

$$ub_{r_n} = \sum_{Bool(\pi)=1} p(G(r_n) = \pi | TF = 0) \times v_{max}(r_n);$$
$$lb_{r_n} = \sum_{Bool(\pi)=1} p(G(r_n) = \pi | TF = 0) \times (-v_{max}(r_n)), \quad \text{(Equation 2)}$$

where $Bool(\pi) = 1$ denotes that the corresponding GPR rules between the genes and the reaction are satisfied with the state profile π representing the corresponding states of genes. Note that the above flux constraints are directly derived based on the conditional joint probabilities of all the regulating genes for a given reaction r_n. One illustrative example is given in the Data S1.

Finally, we note that the above equations can be extended to experiments that involve multiple TF knockouts, enabled by flexible BN-based transcription regulation modeling. In the remaining content, we use TRIMER-C to denote the TRIMER implementations including the flux constraints computed in the first way and TRIMER-B for the second way.

Data structure for metabolic reaction network. TRIMER adopts a data structure organized in a similar way as that in the TIGER package (Jensen et al., 2011) to represent the TF-regulated metabolic reaction network. In this data structure, constraints, lower/upper bounds, variable types of the reaction flux variables provided in the model files from the COBRA toolbox, together with the corresponding information for additional variables are represented and stored in a unified framework. As shown in the data structure representation in Figure 3, fields obj, varnames, vartype, lb, and ub correspond to the coefficient vector used in the objective function of the corresponding metabolic network model formulations, such as FBA or ROOM (Shlomi et al., 2005); descriptive names of involved variables; variable types; and lower/upper bounds.Fields A, b, ctype store all the information about the constraints over variables, including the specific parameter setups in the corresponding metabolic model under given conditions. Stoichiometry constraints $S\vec{v} = 0$ for flux variables $\vec{v}$ and all the other additional linear constraints over the decision variables in the data structure specified by users are collected into the matrix A and vector b and represented as a single expression $A\vec{var}$ *op* b, where $\vec{var}$ denotes all the variables included in TRIMER and *op* is an operator vector constituting {'>', '<', '='}stored in the fieldctype. In TRIMER, build-in functions are implemented to provide a standardized way to build the aforementioned data structure. One example can be found in the Data S1.

Metabolic flux prediction. In TRIMER, we have implemented two variations of the FBA formulations for metabolic flux prediction in addition to the standard FBA formulation with biomass as the objective function as described earlier. When predicting corresponding reaction fluxes of knockout mutants for all these formulations, let $\vec{v}$, $\vec{v}^0$, ub, $lb \in \mathbb{R}^m$ and I, denote the flux variables, wild-type optimal flux vector (the fluxes obtained by performing the standard FBA with the initial flux bounds given by the COBRA toolbox), flux upper and lower bounds for all the m reactions, as well as the set of reactions affected by the corresponding TF knockout(s). For each affected reaction, the reaction flux bounds are modified as described previously. With that, the optimization formulation for mutants with the biomass objective and slack variables allowing violating flux bound constraints, denoted as **sFBA**, is as follows:

$$\max_{\vec{v},\vec{\alpha},\vec{\beta}} \quad \text{biomass}(\vec{v}) - \vec{\kappa}^{\mathrm{T}}\left(\vec{\alpha} + \vec{\beta}\right)$$
$$\text{s.t} \quad S\vec{v} = 0;$$
$$lb_i - \alpha_i \le v_i \le ub_i + \beta_i, \qquad \forall i \in \{1, \ldots, m\};$$
$$\kappa_i \begin{cases} = \dfrac{biomass(\vec{v}^0)}{max(|v_{max}(i)|, v_{thresh})}, & \forall i \in I; \\ = 0, & \text{otherwise}, \end{cases}$$

where $\vec{\alpha}$ and $\vec{\beta}$ can be considered as slack variables and $\vec{\kappa}_i$ is a coefficient vector controlling which reactions are allowed to exceed the upper/lower bounds and the penalty for exceeding the bounds.

We have also implemented ROOM, which is believed to better model mutant strains (Shlomi et al., 2005). In the ROOM formulation, the objective is to minimize the number of reactions with significant changes from the wild-type fluxes $\vec{v}^0$. TRIMER solves the following optimization problem:

$$\begin{aligned}
\min_{y} \quad & \sum_i y_i \\
\text{s.t} \quad & S\vec{v} = 0; \\
& lb_i \le v_i \le ub_i, && \forall\, i \in I; \\
& v_i - (ub_i - w_i)y_i \le w_i, && \forall\, i; \\
& v_i - (lb_i - w_i)y_i \ge w_i, && \forall\, i; \\
& w_i = v_i^0 + \delta\left|v_i^0\right| + \varepsilon, && \forall\, i,
\end{aligned}$$

where δ and ∈ are two hyperparameters used in the original ROOM formulation to define the allowed flux changes from the wild-type fluxes $\vec{v}^0$.

The corresponding lower and upper reaction flux bounds in these metabolic network models are modified based on the inferred conditional probabilities given transcriptional changes as described in the previous subsections.

Following TIGER (Jensen et al., 2011), TRIMER builds a customized MATLAB CMPI (Common Mathematical Programming Interface) for metabolic flux prediction based on the data structure detailed above. This CMPI defines a consistent structure for mathematical programming solvers, including CPLEX and GLPK.

Datasets and software packages

TRIMER integrates several existing packages. For the BN learning and inference module, bn-learn (Nagarajan et al., 2013) and gRain (Højsgaard et al., 2012) are adopted for Bayesian network learning and inference respectively. For the metabolic flux prediction module, TRIMER supports CPLEX and GLPK as solvers for the three aforementioned FBA formulations. In addition, TRIMER is also compatible with the CMPI module in that TIGER package (Jensen et al., 2011) to interface with the corresponding FBA solvers.

Microarray datasets. We have focused on the analyses with *E. coli* by TRIMER in the main text. To infer the TF regulation network and determine the 'ON/OFF' gene states, quantile normalization is performed over the archived microarray data in EcoMAC (Carrera et al., 2014; Chandrasekaran and Price, 2013) as described previously.

TF-gene interaction annotations. For *E. coli*, we have used the interaction set in EcoMAC (Carrera et al., 2014). These data comprise all archived interactions in RegulonDB v8.1 (Santos-Zavaleta et al., 2019) that were experimentally validated to support the existence of regulatory interactions, and we have 3,704 regulatory interactions in total. Serving as prior knowledge, those interaction pairs helped to learn the BN from microarray data and derive the TF-target list for metabolism regulation as detailed previously.

Metabolic model. In general, TRIMER can take any metabolic model in the COBRA format based on the organism under study. We have used the iAF1260 model for *E. coli* from the COBRA toolbox (Heirendt et al., 2019) throughout all the current experiments as the lab experimental data are collected from *E. coli* wild-type strains and knockout mutants.

GPR rules. In COBRA (Heirendt et al., 2019), the GPR rules are provided for most of the metabolic reactions, including iAF1260. TRIMER takes these GPR rules from COBRA directly.

Experimental data collection

E. coli mutants and validation. Strains deleted for genes encoding transcription factors used in this study were obtained from the Keio collection, an *E. coli* mutant library (Baba et al., 2006). All comparisons

were made to BW25113, the parent strain of the collection (Bachmann, 1972). Mutants were validated with internal gene-specific primers by colony PCR.

Kovac's assay for indole quantification. The amount of indole produced by each mutant of interest was quantified by Kovac's assay as described in Chant and Summers (2007). Briefly, total indole concentrations were determined by growing strains at 37°C overnight in LB or M9 minimal media, data for growth in each media is provided in Tables S1 and S2, and normalized to an OD_{600} of 0.3 the following morning. 60 μl of Kovac's reagent (comprised of 150 ml isoamyl alcohol (IAA), 50 ml concentrated hydrochloric acid (HCl) and 10 g of para-dimethylaminobenzaldehyde (DMAB)) was added per 200 μl of normalized culture and incubated for 2 minutes. 10 μl were subsequently removed and added to 200 μl of an HCl-IAA solution, and the absorbance measured at 540 nm. Indole concentrations were then calculated using an indole standard curve prepared in the same manner as described above.