Supporting Appendix

1. Characterization and validation of metabolite essentiality

We initially constructed a genome-scale *in silico E. coli* model comprising 762 metabolites and 932 metabolic reactions (1–4), and by the constraints-based flux analysis (5), calculated the metabolic fluxes which assign the activation rates of the given reactions to produce or consume the metabolites.

To explore the robustness of *E. coli* metabolism from the metabolite perspective, we should identify the metabolites which are substantial in cellular functions. In this regard, all intracellular metabolites are classified into two categories, essential and non-essential metabolites. In addition, we can validate *in vivo* the essentiality of metabolites by means of the multiple gene disruptions around the considered metabolites (SI Fig. 4).

1.1 Constraints-based flux analysis

This analysis quantifies the cell growth by solving a biomass equation derived from the drain of biosynthetic precursors and relevant cofactors into *E. coli* biomass with their appropriate ratios (6). The stoichiometric relationships among all metabolites and reactions of the genome-scale *in silico E. coli* model are balanced under the stationary hypothesis. The resultant balanced reaction model is, however, almost always underdetermined in calculating the flux distribution due to insufficient measurements and/or constraints. Thus, the unknown fluxes within the metabolic reaction network are evaluated by maximizing the growth flux generating biomass, subject to the constraints pertaining to mass conservation, reaction thermodynamics, and capacity as follows:

$$\sum_{j \in J} S_{ij} v_j = b_i \,, \quad \alpha_j \le v_j \le \beta_j \,,$$

where S_{ij} represents the stoichiometric coefficient of metabolite i in reaction j, v_j the flux of reaction j, J the set of all reactions, and b_i the net transport flux of metabolite i. If this metabolite is an intermediate, b_i would be zero. α_j and β_j are the lower and upper bounds of the flux of reaction j, respectively. Herein, the flux of any irreversible reaction is considered to be positive: the negative flux signifies the reverse direction of the reaction. The intracellular fluxes were quantified to elucidate the robustness of *E. coli*

metabolism in response to genetic perturbations under a variety of environmental conditions (SI Table 2). Note that of diverse environmental conditions, the case of the glucose-minimal aerobic condition was presented as the representative result if similar results were observed for all other conditions.

1.2 Characterization of metabolite essentiality

The metabolite essentiality can be defined as the phenotypic effect on cell growth when the consumption rate of a given metabolite *M* is set to zero. All fluxes around the metabolite *M* should be restricted to only produce the metabolite, for which balancing constraint of mass conservation is relaxed to allow nonzero values of the incoming fluxes while all outgoing fluxes are limited to zero. As such, other metabolites linked to the reactions producing the metabolite *M* can be consistently taken into account, preventing the phenotypic effect irrelevant to the essentiality of the given metabolite *M*. We scaled the resultant change of cell growth rate relative to the growth rate of the wild type for calculating the essentiality of the metabolite. When all reactions around the metabolite are inactive for specific growth condition, we consider that metabolite as non-essential. Since the essentiality of all metabolites follows a clear bimodal distribution as in SI Fig. 5, an essential metabolite can be easily identified when its absence leads to decrease in cell growth rate at least one-half of that of the wild type, while the absence of a non-essential metabolite has minimal or no degrading effect on cell growth.

1.3 Construction of gene knock-out mutant strains

Strains were constructed by the one-step gene inactivation method (7). Plasmid pKD46 containing phage λ recombination system was transformed into the wild-type strain of *E. coli* K-12 W3110. The resultant recombinant *E. coli* K-12 W3110 cell with pKD46 was cultivated at 30°C, and λ recombinases were expressed by adding L-arabinose (10 mM) at the optical density (600 nm) of 0.4. Then electrocompetent cells were prepared by standard protocol (8). The PCR was performed using the plasmid pKD3 or pKD4 which contains the antibiotic resistance gene flanked by FRT sequence (FLP recognition target) as templates. Primers employed in this PCR are listed below.

(a) Oligonucleotides used for gene deletion

Name	Sequence (5' → 3') *
upkopurN	ATGAATATTGTGGTGCTTATTTCCGGCAACGGAAGTAATTTACAGGCAATGATTGCAGCATTACACGTCTTG
dokopurN	TTTCGTGCATTTTCAGACGACCATCGGCAAACCAGCTAATCACCAGTGGACACTTAACGGCTGACATGGGA
upkolpdA	CTGGCTGAACACGGTATCGTCTTCGGCGAACCGAAAACCGATATCGACAAGGATTGCAGCATTACACGTCTTG
dokolpdA	TCAACGCGGATGAAACCACGGTCGTCAACTTCCACGCCTGCTTTGCCTGCCACTTAACGGCTGACATGGGA
upkoglyA	CAGGAAAAAGTACGTCAGGAAGAGCACATCGAACTGATCGCCTCCGAAGATTGCAGCATTACACGTCTTG
dokoglyA	CTCGATAACGGCTTCATCATTGATGCTGTCCAGCACGTCACACATCCACACTTAACGGCTGACATGGGA
upkodxs	GAACTGACCGTGGCGCTGCACTATGTCTACAACACCCCGTTTGACCAATTAGATTGCAGCATTACACGTCT
dokodxs	TAAGGATCGCCAGTTTCTCGCCACGACGCTTCACAATGCCTTTGCCAATTCACTTAACGGCTGACATGGGA
upkoxylB	ATTACCGGCAACCTGATGATGCCCGGATTTACTGCGCCTAAATTGCTATGGGTAGATTGCAGCATTACACGTCT
dokoxylB	TGTTCTAACGGTAGTTGCGGCAACAATTCAATGAGCGATTTCTCTGGATTCGCACTTAACGGCTGACATGGGA
CuppurN	AATCTTCCTGGCAGTGGCCG
CdopurN	ATTCACGGGACACTGCCTGG
CpulpdA	CGGTGTTTGCCTGAACGTCG
CdolpdA	CGGTATAGGCGATGGACGGG
CupglyA	CCAACAGGACCGCCTATAAAGG
CdoglyA	GCGCAGATGTCGAGAACTTTACC
Cupdxs	TTACTCGACAGCGTGAGCCG
Cdodxs	GCGTACCAAAGTTAAGGATCGCC
CupxylB	TGGTGGCAGGCAACTGATCG
CdoxylB	TGGCTGATAAGCGGCATAACG

The PCR products were transformed into the electrocompetent *E. coli* W3110 harboring pKD46. Colonies were selected on LB agar plates containing ampicilin (50 μg/mL) or chloramphenicol (30 μg/mL). Gene replacement with the antibiotic marker was confirmed by PCR. A helper plasmid pCP20 was used to eliminate the antibiotic marker, which encodes the FLP recombinase. Plasmid pCP20 contains ampicilin and chloramphenicol resistant markers, and shows temperature-sensitive replication. pCP20 was transformed into the Cm^R or Km^R knockout mutants, and ampicillin-resistant transformants were subsequently selected on LB agar plates at 30°C. Several colonies were cultivated without antibiotic marker in LB liquid media at 42°C and then examined for the loss of all antibiotic resistances. The colony which lost both the FLP helper plasmid and the FRT-flanked resistance gene was selected. The elimination of antibiotic marker was verified by PCR.

1.4 *In vivo* responses of the growth capability to the gene disruptions for essential and non-essential metabolites

To detect the indispensability of essential metabolites in cell survival, we conducted the gene deletion experiments for the neighboring reactions of tetrahydrofolate, which is identified as an essential metabolite *in silico*, and involved in eight reactions. We blocked all the non-lethal outgoing reactions, three reactions out of eight, and observed the resultant change of the growth capability. For the comparison, 1-deoxy-D-xylulose 5-phosphate which is identified as a non-essential metabolite *in silico*, was also subjected to the neighboring gene deletion experiment. Throughout this experiment, we prepared the deletion strains as described in Section 1.3, and the growth conditions were maintained as the following:

(a) Composition of M9 minimal medium

glucose	$MgSO_4$	$CaCl_2$	thiamine	$Na_2HPO_4 \cdot 7H_2O$	KH_2PO_4	NaCl	NH_4CL
5g/L	2mM	0.1mM	1mg/L	12.8g/L	3g/L	0.5g/L	1g/L

(b) Culture conditions in Bioscreen C

temperature	shaking intensity	shaking interval	preheating time	total incubation time	volume of culture broth
37℃	medium	continuous	10min	36h	200μL

We determined the specific growth rate from the slope of the logarithm of optical density (600 nm) versus time curve by using Bioscreen C (Oy Growth Curves AB Ltd, Helsinki, Finland). All the strains were cultured in microplates in wells with 200 uL of M9 minimal medium supplemented with 5 g/L of glucose at an initial pH of 6.7 and 37°C. Cells were precultured overnight in test-tubes containing the same medium, and an aliquot (5 uL) of cell suspension was inoculated into each well containing fresh medium. All cell suspensions were adjusted to the same cell density before inoculation. Cell growth was monitored by optical density (600 nm) for 36 hours, taking measurements every 15 min with continuous shaking. Optical density (600 nm) was measured with triplicates. Wells without inocula were used as a negative control.

Gene deletions for tetrahydrofolate (essential)

We cultivated the combinatory derivatives for the deletions of the three genes, *purN*, *lpdA*, and *glyA*, associated with the three reactions around tetrahydrofolate, which are phosphoribosylglycinamide formyltransferase, glycine cleavage system, and glycine hydroxymethyltransferase, respectively.

Strain	**W3110**	**Δ*purN***	**Δ*lpdA***	**Δ*glyA***	**Δ*purN*Δ*lpdA***	**Δ*purN*Δ*lpdA*Δ*glyA***
growth rate (μ)	0.313h^{-1}	0.292h^{-1}	0.228h^{-1}	0.188h^{-1}	0.102h^{-1}	no growth

(The triple gene deletion mutant Δ*purN*Δ*lpdA*Δ*glyA* could not be even isolated in the LB cultures.)

Each single and double gene deletion mutant could still survive albeit with some growth rate changes, but simultaneous deletions of all the three genes did not allow the cell to grow at all, reflecting that the combinatory suppression of the tetrahydrofolate pool is indeed fatal to the cell.

Gene deletions for 1-deoxy-D-xylulose 5-phosphate (non-essential)

We disrupted all the genes relevant to the incoming reactions of 1-deoxy-D-xylulose 5-phosphate, *dxs* for 1-deoxy-D-xylulose 5-phosphate synthase and *xylB* for 1-deoxy-D-xylulose kinase, thereby blocking the production of the metabolite completely. Nonetheless, this strict perturbation did not affect the cell growth significantly (μ =0.349h^{-1}), quite contrary to the partial gene deletions around the essential metabolite, tetrahydrofolate, which induced the fatal cell damage.

Computational analysis

We also performed three different *in silico* analyses of the gene perturbations based on the genome-wide metabolic models of *E. coli* under the glucose-minimal aerobic condition. The first model A relies on a constraints-based flux analysis, while modified model B is with additional transcriptional regulatory constraints (9), and model C is for another optimization scheme, MOMA (10). Owing to each of the three models, we can evaluate and compare the *in silico* gene-deletion data to the experimentally verified ones as presented below. Here $W(\Delta X)$ represents the growth rate of strain ΔX relative to wild type.

	W(Δ*purN*)	*W*(Δ*lpdA*)	*W*(Δ*glyA*)	*W*(Δ*purN*Δ*lpdA*)	*W*(Δ*purN*Δ*lpdA*Δ*glyA*)	*W*(Δ*dxs*Δ*xylB*)
Exp	0.933	0.728	0.602	0.326	0.000	1.115
Model A	0.996	0.943	0.964	0.938	0.000	1.000
Model B	0.996	0.927	0.956	0.923	0.000	1.000
Model C	0.987	0.529	0.755	0.523	0.000	1.000

(Yellow: experimental data; gray: computational results consistent with experimental data.)

The experimental data and the outcomes of all models, A, B, C, commonly suggest the perfect synergistic epistasis between *purN*/*lpdA* and *glyA*, which are around the essential metabolite. The models uncover the individual and synthetic effects of the gene disruptions supported by the experimental data in which each growth rate relative to the wild type is consistent with at least one of the model outcomes. The slightly improved growth rate after the deletions of *dxs* and *xylB* can be understood by the hypothesis of antagonistic pleiotropy (11) where adaptations to different constitutions of the environment are likely to interfere with each other, and therefore loss of dispensable genes is potentially beneficial. Throughout this computational analysis, model A and model B provided the similar growth rates in their results, while model C gave the relatively small values. Some discrepancies between the model outcomes and the experimental data might be due to incomplete biochemical knowledge or inaccuracies in modeling the types and relative amounts of nutrient conditions and biosynthetic components. These caveats aside, our approach could be considered a step towards investigating the metabolite essentiality by the computational methods.

2. Metabolite flux-sum and its stability

2.1 Flux-sum and flux-sum fluctuation

To understand the robustness of the cellular metabolism from the metabolite perspective, it is necessary to quantify the usage of all relevant fluxes to a given metabolite. In this sense, we introduce the flux-sum (Φ) of the metabolite, which is defined as the summation of all incoming or outgoing fluxes as follows:

$$\Phi_i = \sum_{j \in P_i} S_{ij} v_j = -\sum_{j \in C_i} S_{ij} v_j = \frac{1}{2} \sum_j \left| S_{ij} v_j \right|,$$

where S_{ij} is the stoichiometric coefficient of metabolite i in reaction j, and v_j is the flux of reaction j. P_i denotes the set of reactions producing metabolite i, C_i the set of reactions consuming metabolite i. Under the stationary assumption, Φ_i is the mass flow contributed by all fluxes producing (consuming) metabolite i.

Based on this measure pertaining to the behavioral characteristic of metabolites, we can analyze the robustness of *E. coli* metabolism how the cells maintain their functions against the genetic mutations. The sensitivity to genetic perturbation for a given metabolite can be quantified by evaluating the relative fluctuation of Φ_i in response to each deletion of *active non-lethal* reactions: $\sqrt{\langle \Phi_i^2 \rangle - \langle \Phi_i \rangle^2} / \langle \Phi_i \rangle$ where $\langle \cdots \rangle$ denotes the average over the deletions of active non-lethal reactions. When determining the relative fluctuation values, we exclude an inactive metabolite ($\Phi_i = 0$) for every single-reaction deletion. As the flux-sum fluctuations decrease, essential metabolites are more likely to dominate non-essential ones in number, and the probability distribution of the fluctuations for essential metabolites turns out to be steeper to small fluctuations than that for non-essential ones (SI Fig. 6).

2.2 Relationship with structural and functional properties

To observe the correlation between the structural property and above flux-based quantity, we plot the number of reactions (degree) participated in by each metabolite (12) versus its flux-sum for essential and active non-essential metabolites (SI Fig. 7*a*). We also address the relationship between degree and flux-sum fluctuation for essential and non-essential metabolites (SI Fig. 7*b*). We observe that flux-sum of an essential metabolite tends to increase as the degree gets larger. For both essential and non-

essential metabolites, exceedingly large degree helps lower the flux-sum fluctuation under genetic perturbations.

The investigation on the flux-sum and flux-sum fluctuation according to the functional categories for essential/non-essential metabolites reveals the various function-specific properties. Here we exclude cofactors and inactive metabolites under every single reaction deletion to avoid any possible biases unrelated with the function-specific properties.

SI Figure 8 shows that the mean flux-sum of essential metabolites is apparently greater than those of non-essential metabolites for most categories. TCA cycle/glyoxylate pathway and oxidative phosphorylation important for ATP generation have the high values of mean flux-sums, while the membrane formation pathways and the amino acid synthesis show the relatively low values. The categories of high flux-sums are massively exploited for broad demands of the metabolic network, and their generated mass-flows are finely circulated into the specified categories of low flux-sums.

SI Figure 9 shows that the average of the flux-sum fluctuations of essential metabolites is apparently smaller than those of non-essential metabolites for most categories (compare the scales of two vertical axes in the figure). The mass flows of essential metabolites are evenly stabilized for all categories, while non-essential metabolites are expendably used under each different internal disturbance as manifested in the central pathways like glyoxylate pathway; this is probably due to the need of stabilizing the mass flows of essential metabolites by utilizing the redundant non-essential metabolites. In the membrane formation pathways and the amino acid synthesis, non-essential metabolites show the small fluctuations, thought to be rather drawn to the flows of essential metabolites. It is noticeable that many of the non-essential metabolites in the highest twenty ranks of the large fluctuations (13) participate in alternate carbon metabolism; they include the one having the largest fluctuation, 2-hydroxy-3-oxopropanoate, which also participates in glyoxylate pathway already mentioned.

3. Flux redistribution against strict perturbation

To determine what factors would contribute to the resistance of essential metabolites to the internal perturbation, we carried out the perturbation analysis on the individual flux values around essential metabolites under consideration. We evaluated flux-vector fluctuation $\sqrt{\left\langle \left|\overrightarrow{\Psi_i}\right|^2 \right\rangle - \left|\left\langle \overrightarrow{\Psi_i} \right\rangle\right|^2} \Big/ \left|\left\langle \overrightarrow{\Psi_i} \right\rangle\right|$ for each of essential/non-essential metabolites, where the flux-vector of metabolite i, $\overrightarrow{\Psi_i}$, is defined as $\overrightarrow{\Psi_i} = \{S_{ij}v_j\}$, $|\cdots|$ denotes the magnitude of a given vector, and $\langle\cdots\rangle$ the average over the deletions of active non-lethal reactions. Flux-vector fluctuation represents the relative deviation of the flux values around the given metabolite under the genetic perturbation. It turns out that flux-sum fluctuations of essential metabolites are not much affected by the increment of flux-vector fluctuations (Kendall τ = 0.55), compared with those of non-essential ones (Kendall τ = 0.86) (14). This indicates that flux-sums of essential metabolites are relatively insensitive to the flux variations around them.

To clarify such resistance of essential metabolites against the internal disturbance, the severe perturbation was conducted by removing the active non-lethal reaction j having the maximum flux contribution $|S_{ij}v_j|/\Phi_i$ to metabolite i. For many essential metabolites, the resultant flux loss is found to be mostly recovered by the fluxes of other remaining reactions, thereby leading to very small change of flux-sum Φ, contrary to the case of non-essential metabolites (15).

3.1 Examples of essential metabolites recovering the flux-sum under severe perturbations

Such essential metabolites include carbamoyl phosphate, dUMP, CMP, and L-glutamate 5-semialdehyde. In SI Figs. 10 to 13, thickness of each arrow represents the amount of flux, as shown below the name of the reaction.

Carbamoyl phosphate Carbamoyl phosphate (cbp) is used for the arginine and proline metabolism, or for the purine and pyrimidine biosynthesis (SI Fig. 10). Carbamoyl phosphate can be converted to N-carbamoyl-L-aspartate (cbasp) and L-citrulline (Citr-L) by aspartate carbamoyltransferase (ASPCT) and ornithine carbamoyltransferase (OCBT). The flux-sum Φ of carbamoyl phosphate is reproducible by other fluxes when the highest flux of carbamate kinase (CBMK) is eliminated.

98.9% of the flux-sum is recovered by carbamoyl-phosphate synthase (CBPS).

The analysis on the Affymetrix *E.coli* antisense oligonucleotide array data from Allen TE, *et al*. (16) actually reveals the negative correlation between the expression levels of the gene for carbamate kinase (*arcC*) and those of the gene cluster for carbamoyl-phosphate synthase (*carA*/*carB*): $r=-0.36$ for both between *arcC* and *carA*, and between *arcC* and *carB*, which is significantly different compared with $\langle r\rangle=0.04$, $\sigma_r=0.34$ over all gene pairs in the array. The negative correlation between *arcC* and *carA*/*carB* is consistently supported by the spotted cDNA array data published by Allen TE, *et al*. (16). This indicates probably the mutually-exclusive activation of carbamate kinase and carbamoyl-phosphate synthase against the perturbations, as claimed by our flux-based analysis.

dUMP/CMP dUMP (dump) and CMP (cmp) are the members of nucleotide salvage pathways which contribute to the rescue/resynthesis of nucleotides (SI Figs. 11 and 12). Metabolite dUMP is required for both breakdown and synthesis reactions in order to exchange the useful nucleotides. The flux-sums of both metabolites are conserved when relatively high fluxes are eliminated, indeed, 99.98 and 99.6% of the flux-sums are recovered for the cases of dUMP and CMP, respectively.

L-glutamate 5-semialdehyde L-glutamate 5-semialdehyde (glu5sa) participates in urea cycle, and also in arginine/proline metabolism (SI Fig. 13). L-glutamate 5-semialdehyde can be converted to 1-pyrroline-5-carboxylate (1pyr5c). In the case of the elimination of the highest flux, the flux-sum of L-glutamate 5-semialdehyde is conserved by other fluxes, 99.89% of the flux-sum is recovered.

3.2 Identification of the reactions to be activated under severe perturbations to recover the flux-sum

For most essential metabolites, the loss of the metabolite flux-sum can be compensated for by other reactions when the neighboring reaction with the highest flux is eliminated/perturbed. It would be beneficial if we could predict which candidate reaction will be turned on when severe perturbation is applied. The potential reaction of the highest flux after the severe perturbation around metabolite i can be simply identified *a priori* based on the stoichio-similarity Ω^i_{jk} between eliminated reaction j and another relevant reaction k:

$$\Omega_{jk}^{i} = n(\mathrm{P}_j \bigcap \mathrm{P}_k) + n(\mathrm{C}_j \bigcap \mathrm{C}_k), \ \ \{k \mid i \in (\mathrm{P}_j \bigcap \mathrm{P}_k) \bigcup (\mathrm{C}_j \bigcap \mathrm{C}_k)\},$$

where reaction j (k) produces the set of metabolites, P_j (P_k), and consumes C_j (C_k), and $n(\cdots)$ denotes the number of elements in the set. In other words, we score each of the neighboring reactions under consideration by the number of the connected metabolites which are also commonly produced or consumed by the eliminated reaction. We finally select the reaction with the highest score Ω, but if several reactions have the same similarity with each other, all of them are treated equally.

In the left panel of SI Fig. 14, reaction 1 provides the highest flux to metabolite A. To predict the reaction with the highest flux after eliminating reaction 1, we calculate the stoichio-similarity Ω with other relevant reactions as the following:

$$\Omega_{12}^{\mathrm{A}} = n(\mathrm{P}_1 \bigcap \mathrm{P}_2) + n(\mathrm{C}_1 \bigcap \mathrm{C}_2) = n(\{\mathrm{A}, a\}) + n(\{b\}) = 3 \ \ ,$$

$$\Omega_{13}^{\mathrm{A}} = n(\mathrm{P}_1 \bigcap \mathrm{P}_3) + n(\mathrm{C}_1 \bigcap \mathrm{C}_3) = n(\{\mathrm{A}\}) + n(\{c\}) = 2 \ \ .$$

Since reaction 1 has the highest Ω with reaction 2, we expect reaction 2 to provide the highest flux after the severe perturbation as in the right panel.

Example

One of essential metabolites, 2-oxoglutarate, is associated with amino acid synthesis, and participates in 24 reactions (excluding transport processes). The reaction of the highest flux, glutamate dehydrogenase ($NADP^+$), provides 65.5% of the flux-sum of 2-oxoglutarate. To predict the reaction with the highest flux after eliminating glutamate dehydrogenase ($NADP^+$), we need to calculate the stoichio-similarity Ω with the other relevant reactions. Without the elimination, glutamate dehydrogenase ($NADP^+$) converts the set of metabolites as the following:

$$\text{2-oxoglutarate} + H^+ + \text{NADPH} + NH_4^+ \ \rightarrow \ \text{L-glutamate} + NADP^+ + H_2O \ .$$

It has the highest value of stoichio-similarity with glutamate synthase (NADPH) among the other considered reactions:

$$\text{2-oxoglutarate} + H^+ + \text{NADPH} + \text{L-glutamine} \ \rightarrow \ 2 \ \text{L-glutamate} + NADP^+ \ ,$$

which produces/consumes 5 common metabolites, i.e., Ω=5. Each of the remaining reactions does not share more than 3 metabolites in common, thus glutamate synthase (NADPH) can be regarded as the potential candidate with the highest flux after eliminating glutamate dehydrogenase ($NADP^+$). Indeed, the originally inactivated reaction, glutamate synthase (NADPH), comes to provide 65.3% of the flux-sum after the perturbation, thus 92.6% of the basal flux-sum is recovered.

Result

For essential metabolites to be compensated for under severe perturbations, we can obtain the probability that our method predicts the potential reactions correctly, as the flux contribution ($|S_{ij}v_j|/\Phi_i$) of eliminated reactions varies (17). The results obtained by blind prediction are also compared (termed as 'random' in SI Fig. 15). As illustrated in SI Fig. 15, the method based on stoichio-similarity predicts correctly more than two fold of the blind prediction. If the eliminated reaction dominates the total amount of the flux-sum ($|S_{ij}v_j|/\Phi_i > 0.999$), then the accuracy of the prediction increases up to 73.1%.

The method of stoichio-similarity takes the advantage of not requiring the global information of metabolic networks, thus can be applied to the conditions/organisms in the absence of precise large-scale metabolic models. However, it should be noted the accuracy of the result is not simply the byproduct of the possibility that the reactions of large degree act as the substitutes after the severe perturbations. Scoring the reactions by their degree, not by the stoichio-similarity, actually gives no qualified result compared with the blind prediction. On the other hand, applying our method to non-essential metabolites does not work well compared with the blind prediction, due to the frequent occurrence of tied scores for considered reactions. Therefore, our stoichio-similarity provides the applicable measure to predict the recovery substitutes only for essential metabolites.

4. Attenuation of metabolite flux-sum

Essential metabolites play a pivotal role in the cell survival, steadily maintaining their flux-sum even against severe perturbation by actively redistributing the relevant fluxes. In other sense, attenuating the level of such maintenance can directly suppress the cell growth. The phenotypic effect caused by reduced flux-sum of essential metabolites demonstrates that even the failure of maintaining the flux-sum of a single essential metabolite can suppress the whole cellular growth drastically.

4.1 *In silico* realization of reducing the flux-sum for intracellular metabolites

To see the phenotypic effect when the flux-sum (Φ_i) of metabolite i is reduced down to desired value C, we constrain $\Phi_i \leq C$ where $\Phi_i = 1/2\Sigma_j|S_{ij}v_j|$, and maximize biomass yield subject to balance constraints. This nonlinear programming with discontinuous derivative problem can be recast as the linear programming problem (18), according to the following arguments:

(1) Given metabolite i, let $S_{ij}v_j = f_{ij} - g_{ij}$ where f_{ij} and g_{ij} are declared as positive variables. Then $|S_{ij}v_j| \leq f_{ij} + g_{ij}$, and $|S_{ij}v_j| = f_{ij} + g_{ij}$ if and only if $f_{ij} = 0$ or $g_{ij} = 0$.
(2) From (1), we obtain the relation, $\Phi_i \leq 1/2\Sigma_j(f_{ij}+g_{ij})$.
(3) From (2), we obtain $\Phi_i \leq C$ by constraining $1/2\Sigma_j(f_{ij}+g_{ij}) \leq C$. From (1), there exists $\{f_{ij}, g_{ij}\}$ satisfying $\Phi_i = 1/2\Sigma_j(f_{ij}+g_{ij})$ for an arbitrary value of Φ_i. In this regard, we can also find $\{f_{ij}, g_{ij}\}$ satisfying $\Phi_i = C = 1/2\Sigma_j(f_{ij}+g_{ij})$.

Following the above discussion, reducing the flux-sum of metabolite i down to C can be implemented as:

Maximize *Biomass*
Subject to:
$S_{ij}v_j = f_{ij} - g_{ij}$, where $0 \leq f_{ij}$ and $0 \leq g_{ij}$,
$1/2\Sigma_j(f_{ij}+g_{ij}) \leq C$,
$\Sigma_j S_{kj}v_j = b_k$, if metabolite k is an intermediate, $b_k = 0$,
$\alpha_j \leq v_j \leq \beta_j$.

4.2 Phenotypic effect caused by reducing the flux-sum of essential metabolites

Reducing the flux-sum of each essential metabolite exhibits the characteristic profile of cell growth. SI Figure 16 displays the cell growth rate as a function of reduced flux-sum, and the values are scaled relatively to the wild type. Essential metabolites of type A in SI Fig. 16 induce the cell growth proportionally to the flux-sum, while type B metabolites hardly affect the cell growth over the finite range of the flux-sum. Reducing the flux-sum of type C metabolites below the cut-off value does not allow the cell growth at all. For the three metabolites, ubiquinol-8, ubiquinone-8, and L-malate, the growth rate is found to be suppressed by reducing the flux-sum under some conditions, but takes nonzero values even when the flux-sum is set to zero. They are classified as type D metabolites.

Accordingly, we can categorize all essential metabolites as shown in SI Table 1, and find some interesting features:

(1) Regardless of growth conditions, lots of amino acids and lipids which participate directly in the cell growth as the elementary building blocks are proved to be type A. Of the metabolites in amino acid synthesis, the serine-related ones, nevertheless belong to type B metabolites. The sulfur and cysteine pathways relevant to the serine metabolism presumably contribute to this property, distinctive from those of other amino acids mostly along the linear pathways. From the central metabolism, for example, D-glucose 1-phosphate also belongs to type A metabolites.

(2) The cofactors, including H, H_2O, ATP, Pi, and ADP, are type C metabolites, which means the cell should utilize such metabolites more than the threshold minimally required for the viability.

(3) NAD^+ and NADH are type C for anaerobic/low-oxygen conditions, while type B for aerobic conditions unless subjected to arginine uptake (type C). But even on this arginine culture, they take the growth profile close to the type B. Oxygen-limited environment enforces that interchange of NAD^+ and NADH is only attributable to the presence of organic acids, thus the cell growth rate can respond sensitively to the availability of NAD^+ and NADH compared with the oxygen-rich case.

(4) 5-phospho-α-D-ribose 1-diphosphate for nucleotide biosynthesis comes to be type B on ammonia-limited culture without oxygen; for other conditions, it is type A. When subjected to arginine uptake, putrescine synthesized by the precursors of arginine becomes type B, otherwise it is type A.

(5) The set of the metabolites which share the identical types for every condition does show that the involved elements are also located proximately to each other in the metabolic networks. We list the sets as the following: {L-aspartate, L-homoserine, L-aspartate 4-semialdehyde, 4-phospho-L-aspartate}, {dihydroxyacetone phosphate, D-fructose 6-phosphate}, {reduced thioredoxin, α-D-ribose 5-phosphate}, {oxidized thioredoxin, D-ribulose 5-phosphate}, {N-(L-arginino)-succinate, L-citrulline}.

The essential metabolites of types A and C reveal the cellular-level fragility against the flux-sum attenuation. We expect such metabolites and their associated genes to be considered potential targets in antimicrobial strategy.

5. References

1. Reed JL, Vo TD, Schilling CH, Palsson BØ (2003) *Genome Biol* 4:R54.1–R54.12.
2. Kanehisa M, Goto S (2000) *Nucleic Acids Res* 28:27–30.
3. Ma H, Zeng A-P (2003) *Bioinformatics* 19:270–277.
4. Hou BK, et al. (2004) *Bioinformatics* 20:3270–3272.
5. Price ND, Reed JL, Palsson BØ (2004) *Nat Rev Microbiol* 2:886–897.
6. Neidhardt FC, et al. (1996) *Escherichia coli and Salmonella* (ASM Press, Washington DC).
7. Datsenko KA, Wanner BL (2000) *Proc Natl Acad Sci USA* 97:6640–6645. The pKD3, pKD4, pKD46, and pCP20 plasmids used for our gene-disruption experiments were kindly provided by B. L. Wanner.
8. Sambrook J, Russell DW (2001) *Molecular cloning: a laboratory manual*, 3rd ed (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY).
9. Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BØ (2004) *Nature* 429:92–96.
10. Segrè D, Vitkup D, Church GM (2002) *Proc Natl Acad Sci USA* 99:15112–15117.
11. Cooper VS, Lenski RE (2000) *Nature* 407:736–739.
12. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási A-L (2000) *Nature* 407:651–654.
13. Non-essential metabolites in the descending order of flux-sum fluctuations, also correspondent to the highest twenty ranks in total metabolites (* marked for alternate carbon metabolism): 2-hydroxy-3-oxopropanoate*, cytosine, reduced glutathione, hydroxypyruvate*, urea, agmatine, D-lactate, (R)-S-lactoylglutathione, methylglyoxal, (R)-glycerate*, 2-demethylmenaquinol 8*, 2-demethylmenaquinone 8*, 4-aminobutanoate, succinic semialdehyde, 2-dehydro-3-deoxy-D-gluconate 6-phosphate*, dUTP, O-phospho-L-homoserine, D-glucose*, formate, glyoxylate*
14. We exclude some metabolites with degree of 2 that topologically constrain flux-vector fluctuation and flux-sum fluctuation to be identical in the trivial way.
15. For the severe perturbations, we exclude some trivial cases including those with the metabolites having degree of 2.
16. Allen TE, et al. (2003) *J Bacteriol* 185:6392–6399.
17. We exclude the trivial cases where the number of considered reactions is less than 2, and the coin-toss cases where all neighboring reactions have the same score.
18. http://www.gams.com/dd/docs/solvers/conopt.pdf