Microbes release lower-value metabolites at higher rates

Snorre Sulheim[1,2*], Gunn Broli[3], Alisson Gillon[1], Julien S. Luneau[1], Andrew Quinn[1], Eric Ulrich[1], Margaret A. Vogel[1], Philipp Engel[1] & Sara Mitri[1*]

Affiliations

1. Department of Fundamental Microbiology, University of Lausanne
2. Department of Biological Sciences, University of Bergen
3. Department of Biotechnology and Nanomedicine, SINTEF Industry

*Corresponding author. Email: snorre.sulheim@unil.ch; sara.mitri@unil.ch

Abstract

Microbes release a wide range of metabolites into their environment, yet the reasons for this release and the factors that influence release rates are not well understood. Here, using a combination of computational and experimental approaches on seven different microbes we show that release rates are negatively correlated with the value of metabolites, and that this relationship explains more variability in release rates than other frequently associated factors. These findings are in accordance with the idea that metabolite release generally constitutes a fitness cost that microbes have evolved to reduce. This new conceptual explanation for why microbes release metabolites has important implications for how we think about the emergence of cross-feeding interactions.

Main

Introduction

Metabolites are frequently exchanged between microbes in natural communities (*1*). Such cross-feeding plays a key role in maintaining community diversity and function (*2*–*6*). While certain occurrences of cross-feeding arise from extracellular metabolism such as polymer degradation by extracellular enzymes (*6*, *7*), many observations of cross-feeding stem from the

release of intracellular metabolites into the environment by microbial community members (*8–10*). Recent *in vitro* quantifications of microbial cultures have demonstrated that this release includes high-energy glycolytic intermediates, amino acids and organic acids (*11–14*). The ubiquity of metabolite release may help explain the high frequency of auxotrophs isolated from natural environments that may be supported by co-cultivated partners (*15–18*). However, why so many diverse intracellular metabolites are being released and what explains differences in release rates remains unanswered.

Previous efforts to answer why microbes release intracellular metabolites have revealed fundamental principles of metabolism, including rate-yield trade-offs and constraints on proteome allocation (*19–23*), mechanisms of pathway regulation and maintenance of homeostasis (*24*), optimal catabolic pathway lengths in different environments (*25*, *26*) and redox balancing (*27*, *28*). These explanations are context-, metabolite- or species-specific. However, to fully understand metabolite release, we need to consider broader principles that apply across diverse microbial taxa and environments. Cell lysis and passive diffusion across the cell membrane are two such commonly proposed broad mechanisms (*29–32*).

Here we move beyond these mechanistic explanations and ask whether metabolite release generally confers a benefit or imposes a cost on the source organism. While recent work has explored how metabolite release can be beneficial, e.g. by alleviating costs associated with molecular noise (*32*, *33*), we here propose and test an alternative hypothesis that metabolite release comes with a cost that microbes have evolved to reduce.

Metabolite release rates are negatively correlated with metabolite value

If metabolite release generally imposes a fitness cost, and natural selection acts to minimize such costs, then the release rates of specific metabolites should be negatively correlated with their relative importance, or value, to the microbe (Fig. 1A). Release of more valuable metabolites – that require more energy to produce – should have a more severe negative impact on fitness and be more strongly selected against.

To quantify metabolite release rates across different species, we first analysed an existing time-series exometabolome dataset with high temporal resolution and absolute quantification of 19 to 37 metabolites collected from batch cultures of *E. coli*, *B. licheniformis, S. cerevisiae* and *C.*

glutamicum in high glucose concentrations (*12*). We estimated net uptake and release rates for all metabolites using simple linear regression, modelling extracellular metabolite concentration as a linear function of the area under the growth curve (AUC), until the point where the extracellular concentration saturated or the exponential growth phase ended (Methods, Fig. 1B, and Figs. S1-S4). The initial dynamics of 44-95% of metabolites were well described by this simple linear model (two-sided Wald test, FDR < 0.05, R^2 > 0.5, Fig. S1-S4).

To estimate metabolite value, we used previously reconstructed enzyme-constrained genome-scale metabolic models (ecGEMs) (*34–36*). Genome-scale metabolic models (GEMs) are mathematical representations of a species' metabolic capacity that can for example be used to predict optimal metabolic phenotypes across various environments (*37*, *38*). The additional enzyme constraints account for the total enzyme pool limitation in cells (*20*), enabling prediction of relevant trade-offs and temporal dynamics, e.g. acetate overflow in *E. coli* (*35*). Various measures of metabolite value have been proposed in the past (*39–41*). Here we define metabolite value as the negative change in growth rate on metabolite release close to optimal growth as predicted by enzyme-constrained GEMs (Fig. 1C). This measure explicitly reflects the fitness penalty on growth rate associated with the marginal release of a metabolite, and it can easily be used for different species across different environmental contexts.

With these two measures we could now test our hypothesis that release rate should be negatively correlated with metabolite value. Indeed, this holds for *E. coli, S. cerevisiae, B. licheniformis* in glucose medium (Pearson ρ, P < 0.001, Fig. 1D). This did not hold for *C. glutamicum* (Fig. 1D), possibly because the engineering of its central carbon metabolism for enhanced lysine production has also affected the release of other metabolites (*42*).

To test whether our hypothesis held beyond simple glucose environments, we conducted new well-controlled batch cultivations of *E. coli* in minimal media with galactose, L-malate or L-alanine as the single carbon source. These carbon sources were chosen to span the variability in *E. coli* metabolism to avoid biasing the results towards a particular class of compounds or metabolic flux patterns (Methods, Fig. S5), since this has been shown to produce similar exometabolome patterns (*43*). Triplicate batch bioreactor experiments were performed under strictly controlled aerobic conditions with active regulation of pH and dissolved oxygen (Fig. S6). Relative extracellular concentrations of 126 metabolites were then quantified at three exponential-phase samples and one sample from stationary phase. Of these metabolites, 41-

56% were released at significant rates during exponential phase (two-sided Wald test, FDR < 0.05, Figs. S7-S10), consistent with previous findings showing that microbes generate rich exometabolomes (*12*, *13*, *44*). From this set, we were able to estimate absolute release rates for 34 metabolites that we compared with estimated metabolite values. Again, we found significant negative correlations for all three conditions (Pearson ρ, $P < 0.001$, Fig. 1E).

Finally, we analysed a third dataset that in addition to *E. coli* and *Pseudomonas putida* contained environmental isolates from the genera *Enterobacter* and *Pseudomonas* (*43*). Unlike previous datasets, these experiments were conducted without pH and oxygen control, offering an opportunity to test the hypothesis under more variable conditions. Despite these differences, we again found consistent negative correlations (Fig. 1F), suggesting that the relationship between metabolite value and release rate is robust across species, nutrient conditions, and experimental platforms. Importantly, the ecGEMs used here were reconstructed by different research groups using different pipelines, suggesting that our results are not dependent on the choices made during model reconstruction. This conclusion was also supported by consistent results when metabolite values were estimated using four other *E. coli* GEMs (Fig. S11).

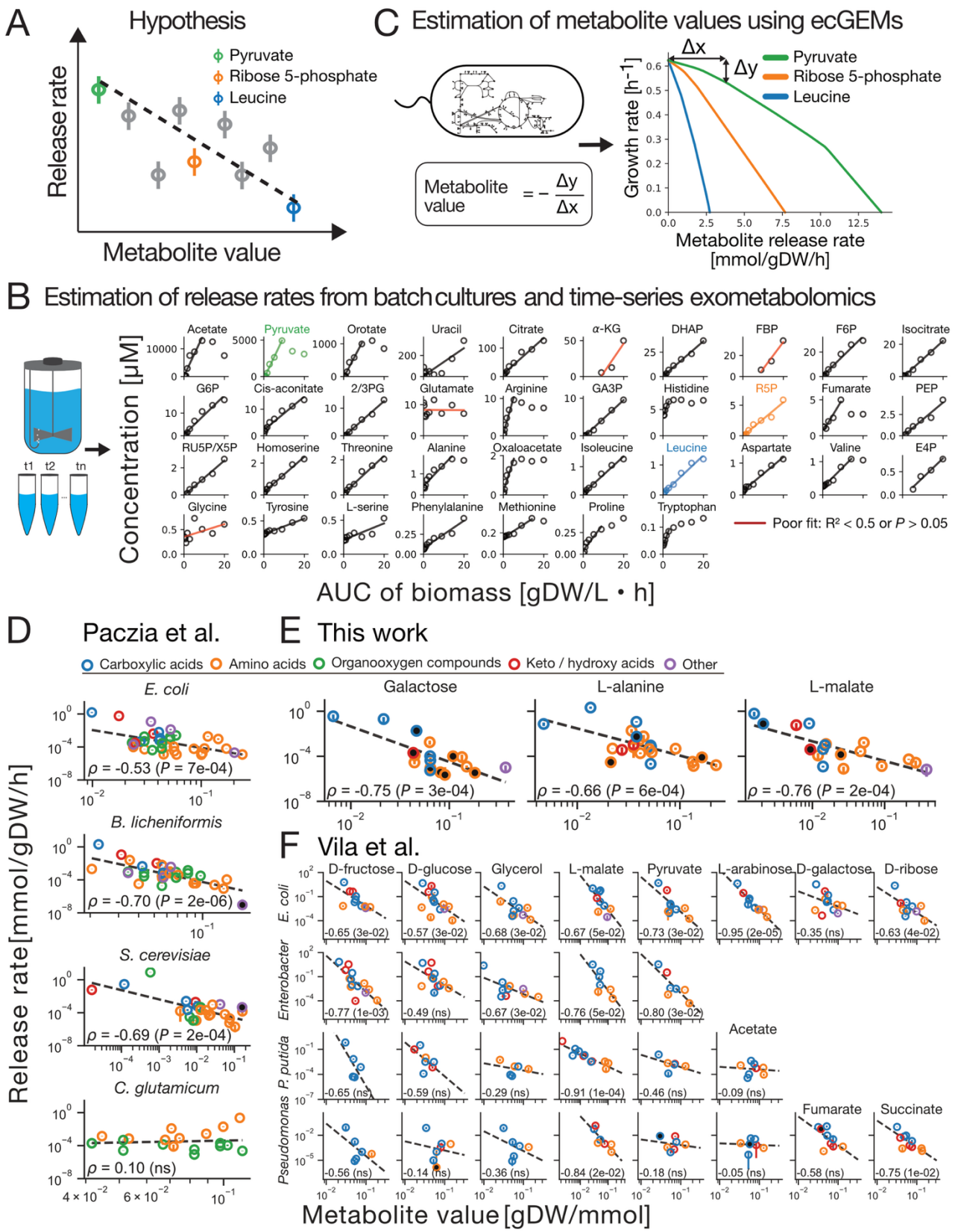
A
Hypothesis
Pyruvate
Ribose 5-phosphate
Leucine
Release rate
Metabolite value
C
Estimation of metabolite values using ecGEMs
Metabolite value = − Δy/Δx
Δx
Δy
Growth rate [h⁻¹]
Metabolite release rate [mmol/gDW/h]
B
Estimation of release rates from batch cultures and time-series exometabolomics
t1 t2 tn
Concentration [μM]
Acetate
Pyruvate
Orotate
Uracil
Citrate
α-KG
DHAP
FBP
F6P
Isocitrate
G6P
Cis-aconitate
2/3PG
Glutamate
Arginine
GA3P
Histidine
R5P
Fumarate
PEP
RU5P/X5P
Homoserine
Threonine
Alanine
Oxaloacetate
Isoleucine
Leucine
Aspartate
Valine
E4P
Glycine
Tyrosine
L-serine
Phenylalanine
Methionine
Proline
Tryptophan
Poor fit: R² < 0.5 or P > 0.05
AUC of biomass [gDW/L • h]
D
Paczia et al.
E
This work
Carboxylic acids
Amino acids
Organooxygen compounds
Keto / hydroxy acids
Other
E. coli
ρ = -0.53 (P = 7e-04)
B. licheniformis
ρ = -0.70 (P = 2e-06)
S. cerevisiae
ρ = -0.69 (P = 2e-04)
C. glutamicum
ρ = 0.10 (ns)
Release rate [mmol/gDW/h]
Galactose
ρ = -0.75 (P = 3e-04)
L-alanine
ρ = -0.66 (P = 6e-04)
L-malate
ρ = -0.76 (P = 2e-04)
F
Vila et al.
D-fructose
D-glucose
Glycerol
L-malate
Pyruvate
L-arabinose
D-galactose
D-ribose
E. coli
-0.65 (3e-02)
-0.57 (3e-02)
-0.68 (3e-02)
-0.67 (5e-02)
-0.73 (3e-02)
-0.95 (2e-05)
-0.35 (ns)
-0.63 (4e-02)
Enterobacter
-0.77 (1e-03)
-0.49 (ns)
-0.67 (3e-02)
-0.76 (5e-02)
-0.80 (3e-02)
Acetate
P. putida
-0.65 (ns)
-0.59 (ns)
-0.29 (ns)
-0.91 (1e-04)
-0.46 (ns)
-0.09 (ns)
Fumarate
Succinate
Pseudomonas
-0.56 (ns)
-0.14 (ns)
-0.36 (ns)
-0.84 (2e-02)
-0.18 (ns)
-0.05 (ns)
-0.58 (ns)
-0.75 (1e-02)
Metabolite value [gDW/mmol]

Figure 1: ***Correlation between microbial metabolite release rates and metabolite values.*** *A) Hypothesis: if metabolite release generally constitutes a fitness cost that microbes have evolved to reduce, metabolite release rates should be negatively correlated with metabolite values. Three metabolites are used to conceptually illustrate their relative position based on their estimated release rate and metabolite value, see panels B and C. B) Metabolite release rates were estimated from time-series exometabolomics data from batch cultures by linear regression of extracellular concentrations vs area under the curve (AUC) of biomass. This panel shows extracellular concentrations of 37 metabolites measured during exponential phase of E. coli cultivated in glucose medium by Paczia et al. (2012), and the linear regressions used to estimate net release rates. Linear regressions with an* $R^2 < 0.5$ *or* $P > 0.05$ *(two-sided Wald test) are drawn in red. Full time-series, including stationary phase and linear regression for E. coli, B. subtilis, C. glutamicum and S. cerevisiae are shown in Figs. S1-S4. Metabolite abbreviations: α-KG: alpha-ketoglutarate, DHAP: dihydroxyacetone phosphate, FBP: fructose 1,6-bisphosphate, F6P: fructose 6-phosphate, G6P: glucose 6-phosphate, 2/3PG: 2/3-phosphoglycerate, GA3P: glyceraldehyde 3-phosphate, R5P: ribose 5-phosphate, PEP: phosphoenolpyruvate, RU5P/X5P: ribulose 5-phosphate/xylulose 5-phosphate, E4P: erythrose-4-phosphate. C) We used enzyme-constrained genome-scale metabolic models to estimate metabolite values, quantified as the negative change in growth rate (Δy) on metabolite release (Δx) close to optimal growth. D-F) Three different datasets were used to test our hypothesis: D) Paczia et al. (2012), including S. cerevisiae, B. licheniformis, C. glutamicum and E. coli on glucose (12), E) E. coli on three other carbon sources (this work), and F) Vila et al. (2023) covering a smaller number of metabolites on E. coli, P. putida and two environmental strains of the genera Enterobacter and Pseudomonas on up to 8 different carbon sources (43). The scatter plots show mean estimated rates +- standard errors vs estimated metabolite values. Black-filled circles are datapoints where the lower error bar is omitted because mean minus standard error is less than 0 and cannot be included on a log-scale. Axis labels and colour legend are shared for panel D-F. Annotations in panel F show the Pearson correlation and associated P value in parentheses.*

Metabolite value is the most important factor explaining variability in metabolite release rates

The consistent negative correlation between metabolite release rates and metabolite values encouraged us to ask how well this factor explains release variability compared to other metabolic or physiochemical properties previously associated to metabolite release (*33*, *45*). We focused on the data from bioreactor batch cultures of *E. coli* with glucose, galactose, L-malate or L-alanine as the carbon source, where conditions were well-defined, the measured metabolites were diverse and reference values for intracellular metabolite concentrations were available (*46*–*49*). We first built univariate linear models to quantify the importance of these factors: metabolite value, intracellular concentration, compound class, solubility in oil vs water (log P), molecular weight, topological polar surface area, charge, hydrogen bond acceptor and donor count, rotatable bond count, turnover and carbon source (Fig. 2A). Metabolite value explained 43% of the variability, more than any other factor (Fig. 2B). The structure-based classification of compound class also had good predictive power (29% explained), primarily because carboxylic and keto/hydroxy acids were released at significantly higher rates than amino acids and other compounds (Fig. 2A). However, these two factors are clearly confounded as class-level differences mirrored significant differences in metabolite values (Fig. 2C). The number of rotatable bonds, solubility in oil vs water (log P) and the number of hydrogen bonds, each related to membrane permeability (*50*, *51*), were also predictive of metabolite release, suggesting passive diffusion as a relevant mechanism. Metabolite turnover is predicted to

correlate negatively with release rates, according to the noise-averaging cooperation hypothesis (low turnover metabolites are more susceptible to noise and therefore more useful to share within the population) (*33*), but our dataset shows little support for this.

When we asked how well the best linear model could predict metabolite release rates, we found that 63% of the variability could be explained by metabolite value, compound class, intracellular concentration and charge (Fig. 2B). However, predictive accuracy declined when applied to out-of-sample conditions (45%, $P = 0.04$, 52.9 ± 3.8% expected from random model, Fig. S12) or out-of-sample metabolites (50%, $P < 0.001$, 55.2 ± 0.7% expected from random model, Fig. S13), likely due to correlations among the factors, which warrants caution when interpreting factors' relative importance (Fig. S14). Nevertheless, excluding metabolite value as a factor reduced model performance on out-of-sample conditions (10% decline in median R^2, $P = 0.003$, one-sided Mann-Whitney U, N = 20, Fig. S15), and model quality was largely dependent on including metabolite value or compound class (Fig. S16). Furthermore, we found differences in metabolite value between conditions to be negatively correlated with differences in release rate (Pearson ρ, P = 2e-3, Fig. S17). This not only supports the relevance of this factor in predicting out-of-sample rates but also raises the question of whether microbes adapt their release rates to context-dependent costs. Finally, when we estimated each factor's importance separately for each microbe in each dataset (except *C. glutamicum*), metabolite value still emerged as the most robust explanatory factor (Fig. 2D).

Cell lysis is commonly considered a key mechanism for metabolite release (*31*, *32*). To measure its importance in our experiments, we used *E. coli* batch cultures with paired sampling of intra- and extracellular metabolites, as well as flow cytometry with live/dead staining to quantify the fraction of lysed cells. In general, less than 10% of extracellular concentrations could be explained by the release of intracellular metabolites following cell lysis, and often much less (Fig. 3A). We then extended this analysis by comparing our collection of *E. coli* extracellular metabolite concentrations to published intracellular concentrations (*46–49*). To avoid asserting equal lysis rates across conditions, we asked how large a fraction of the cell population would need to be lysed to account for the extracellular levels, considering only the contribution from intracellular metabolite pools (Fig. 3B). Only 10.4% (157/1508) of the data points fall within the range of measured lysis fraction (0.2-2.3%, mean ± std across conditions = 0.6 ± 0.6%) or below, and more than 56.3% (849/1508) are above the limit where more than the whole cell population is required. Glutamate, glutamine and NAD are the only metabolites where most

data points can be accounted for by cell lysis, due to high intracellular concentrations and moderate to low extracellular levels (Fig. S18).

Degradation of cell debris may also contribute to extracellular metabolite pools (*52*). In *E. coli,* this degradation is dependent on the presence of Lon protease in the cell lysate (*53*). However, the benefit of cell debris degradation by Lon was only significant well beyond exponential phase, and only 19 ± 6% of oligopeptides (6-10 amino acids) in the lysate were catabolised by Lon over 20 hours (*53*). Thus, in our dataset where 90% of the samples were collected before 26 hours (67% in exponential phase), we expect only a fraction of the proteome from lysed cells to contribute to the extracellular concentrations of amino acids. When we included 10% proteome degradation in our analysis, only the extracellular levels of tryptophan and threonine additionally became fully explained by cell lysis (Fig. S19). However, apart from amino acids, the metabolites in Fig. 3B are not typical degradation products. Overall, these results suggest that cell lysis and proteome degradation are insufficient to explain extracellular levels of most metabolites – consistent with previous work (*12*, *31*).

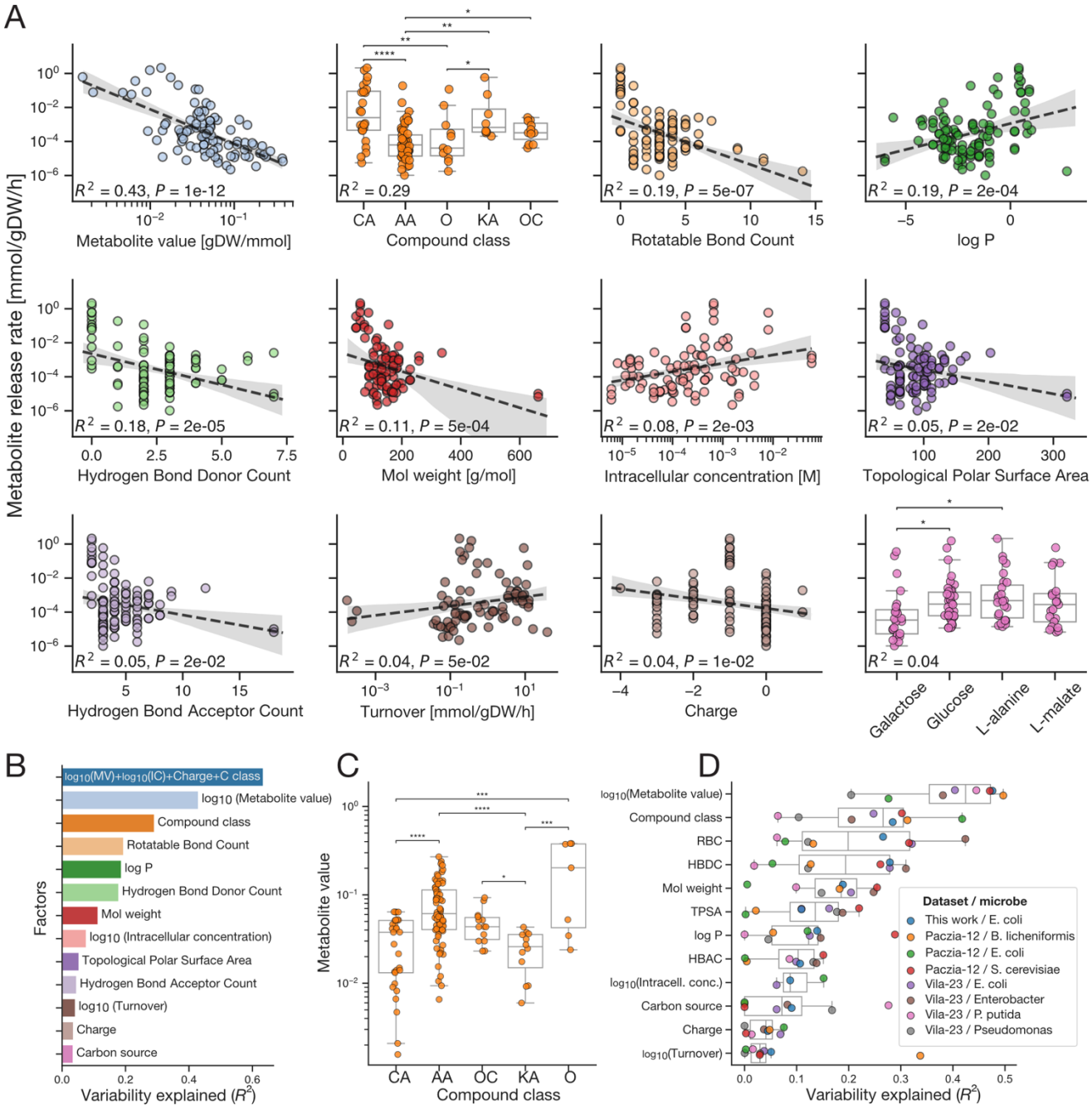

Figure 2: ***The importance of different factors in explaining metabolite release rates.*** *A) Among the tested chemical and metabolic factors, log-scaled metabolite value explains the most variability in log-scaled* E. coli *release rates. The tested factors are shown in order of how much variability (R^2) in log-scaled metabolite value they explain as a univariate linear model, shown in panel B. Compound class abbreviations: CA: Carboxylic acids, AA: Amino acids, KA: Keto/hydroxy acids, OC: Organooxygen compounds, O: Other. Statistical test for compound classes and carbon source: Kruskal-Wallis H-test followed by Conover's test with BH correction. For the other factors the reported P values are from Pearson correlations. B) The best linear model, ranked by BIC score, is a four-factor model that explains 63.4% of the variability in log-scaled release rates. C) Differences in metabolite values between compound classes. Abbreviations and statistical test as in Fig. 2A. D) Log-scaled metabolite value consistently ranks among the most important factors when evaluated separately for each species in each dataset. Factor abbreviations: RBC: Rotatable Bond Count, HBDC: Hydrogen Bond Donor Count, TPSA: Topological Polar Surface Area, HBAC: Hydrogen Bond Acceptor Count.*

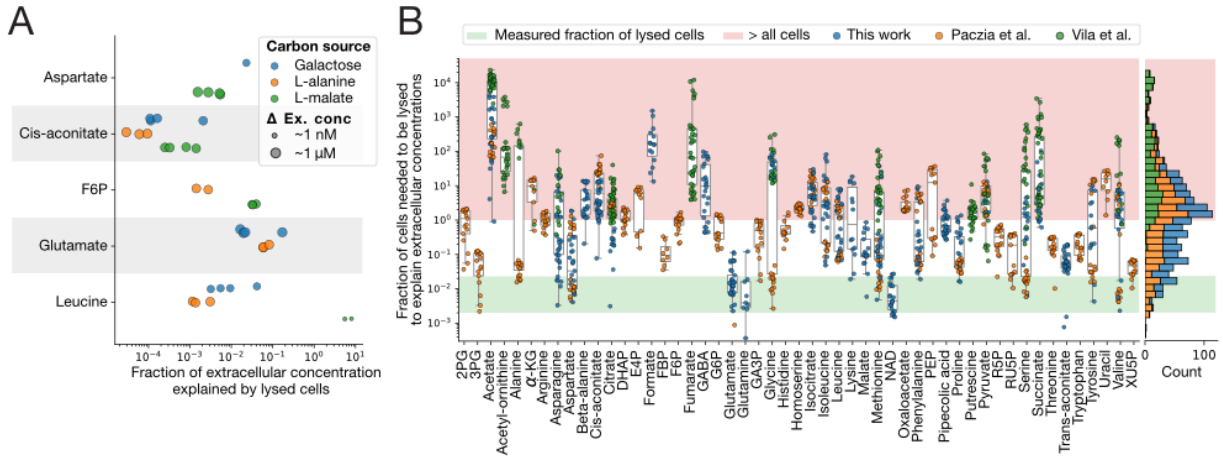

Figure 3: ***Contribution from cell lysis and associated release of intracellular metabolites to extracellular concentrations.*** *A) Using paired intra- and extracellular metabolite samples from the exponential phase of* E. coli *batch cultures we quantify what fraction of extracellular metabolite concentrations can be explained by cell lysis. Point size reflects order-of-magnitude increase in extracellular concentration from the 2-hour sample to the late-exponential phase sample (Fig. S30), and all values except leucine in the L-malate condition were in the μM range. B) We expanded this analysis to all metabolite measurements in our* E. coli *datasets by using literature values for intracellular concentrations (Methods). For each metabolite, we estimated the fraction of cell population that would need to be lysed for the intracellular concentrations released by cell lysis to explain the measured extracellular concentrations. The green region covers the range of measured fractions of lysed cells from A), and the red region is where more than the whole cell population is required to explain extracellular concentrations. Abbreviations: 2PG: 2-phosphoglycerate, 3PG: 3-phosphoglycerate, α-KG: Alpha-ketoglutarate, DHAP: Dihydroxyacetone phosphate, E4P: Erythrose 4-phosphate, FBP: Fructose 1,6-bisphosphate, F6P: Fructose 6-phosphate, G6P: Glucose 6-phosphate, PEP: Phosphoenolpyruvate, R5P: Ribose 5-phosphate, RU5P: Ribulose 5-phosphate, XU5P: Xylulose 5-phosphate.*

Metabolic disruption and evolution change metabolite release rates

Given the effect of gene deletions on intracellular metabolite levels (*54*), we assumed that the lack of negative correlation between metabolite value and release rate in *C. glutamicum* (Fig. 1D) could be a result of its engineered metabolism (*42*). To test the sensitivity of metabolite release patterns to modifications in core metabolism, we measured the exometabolome of eight *E. coli* knockout (KO) strains in late exponential phase in galactose medium (Fig. 4A). These gene knockouts target different parts of core metabolism, were previously shown to affect intracellular metabolite pools (*54*), and were predicted to change intracellular flux patterns (Fig. S20). In line with our expectations, 14-38% (16-42 of 111) of the extracellular metabolite concentrations differed significantly from the wild type across the KO strains (Welch's t-test, FDR < 0.05, Fig. 4B). In contrast, only 2.7% of metabolite concentrations were different in the negative control strain Δ*lacA*. Each KO strain displayed distinct exometabolome patterns (Fig. 4C), and we observed the most pronounced differences for metabolites in proximity of the deleted reaction (Figs. S21-S23), consistent with intracellular effects of enzyme-deletions (*54*).

The WT reference for these KO strains from the KEIO collection (*55*) - *E. coli* BW25113 - had largely comparable rates as *E. coli* MG1655 used in the experiments presented in Fig. 1 (Fig. S24).

Our central hypothesis for the negative correlation between metabolite value and release rate is that bacteria face a trade-off: the cost of preventing metabolite release versus the energy lost through that release. If this trade-off is under strong selection, then disrupting the balance (e.g. via a gene knockout) should trigger rapid evolutionary compensation to restore it. To test this, we evolved the Δ*aceE* and Δ*sucB* mutants – chosen for their severe growth defects and distinct exometabolome characteristics – for ~100 (replicates M2-M4) and ~200 generations (M5-M7) in galactose medium, respectively (Fig. S25). The strains evolved in chemostats, where byproducts are continuously removed to avoid accumulation and later uptake of metabolites. From each chemostat endpoint, we isolated three strains and selected from each replicate the fastest growing clone (Fig. 4, D and E). These clones showed 94-146% (Δ*aceE*) and 50-82% (Δ*sucB*) growth rate increases relative to their ancestors (Fig. 4, F and G). To assess changes in metabolite release, we measured extracellular concentrations of 5 metabolites at the end of exponential growth (Fig. 4, D and E). These metabolites were selected because their release was strongly disturbed in the ancestors (Figs. S21-S23), and either near the deleted reaction (pyruvate and lactate for the Δ*aceE*; citrate, isocitrate and cis-aconitate for Δ*sucB*) where we expected the strongest effects, or located elsewhere in the metabolic network and of different chemical classes to probe more widespread effects. Lactate levels were significantly reduced in 2/3 Δ*aceE* isolates (Fig. 4F), which also exhibited the highest growth rates. The reductions were, however, not sufficient to reach WT concentrations. In contrast, all three Δ*aceE* isolates showed increased pyruvate release, contrary to expectations. These results were qualitatively invariant to growth rate normalization (Fig. S26). For the Δ*sucB* isolates, we found in general small reductions in extracellular concentrations (Fig. 4G), but some of these differences change qualitatively upon growth rate normalization (Fig. S27). For the metabolites elsewhere in the metabolic network, the outcomes were variable (Figs. S28 and S29).

To estimate the fitness contribution of reducing metabolite release, we used ecGEMs to predict relative growth rate penalties associated with the observed release rates, using the ancestors' mean growth rates as references. In the Δ*aceE* ancestor, lactate release is predicted to reduce fitness by 23% whereas pyruvate release has negligible impact (0.02%). The reduced release of lactate in the isolates from chemostat M2 and M4 is predicted to reduce the negative fitness

effect to only 16% and 7%, respectively. These contributions are small compared to the observed increase (94-146%). The pattern is similar for the Δ*sucB* isolates: the estimated net fitness reduction of citrate, isocitrate and cis-aconitate in the ancestor is only 0.37%, far from sufficient to explain the improved growth rates.

To better understand the reasons for increased growth and changed release rates, we identified sequence variants in the selected fast-growing strains and mutations that had fixed in the chemostat populations (Table S1-S4). Aside from the Δ*aceE*-M4 isolate, which likely hypermutated due to a *mutT* mutation, the evolved Δ*aceE* isolates have few sequence variants, and either related to galactose uptake (e.g. *galS* in M2, M4 and M7, regulating galactose uptake) or modulating fluxes near the deleted reaction (e.g. *poxB* in M2, *ldhA* in M4, and *icd, aceK* and *sdhA* in M5, M6 & M7 and M7, respectively). Fixed non-synonymous variants in the chemostat populations support these observations, e.g. *ldhA* and *galS* were fixed in M2 and *ilvG* was fixed in M4. The Δ*sucB* populations have many more fixed mutations, but among the ones shared across replicates we find again *galS* and *aceK*. Together, these results show that metabolite release rates change as microbes evolve, but in our experiments, these changes were likely secondary to other, more dominant drivers of fitness – leading to inconsistent changes in metabolite release across replicates.

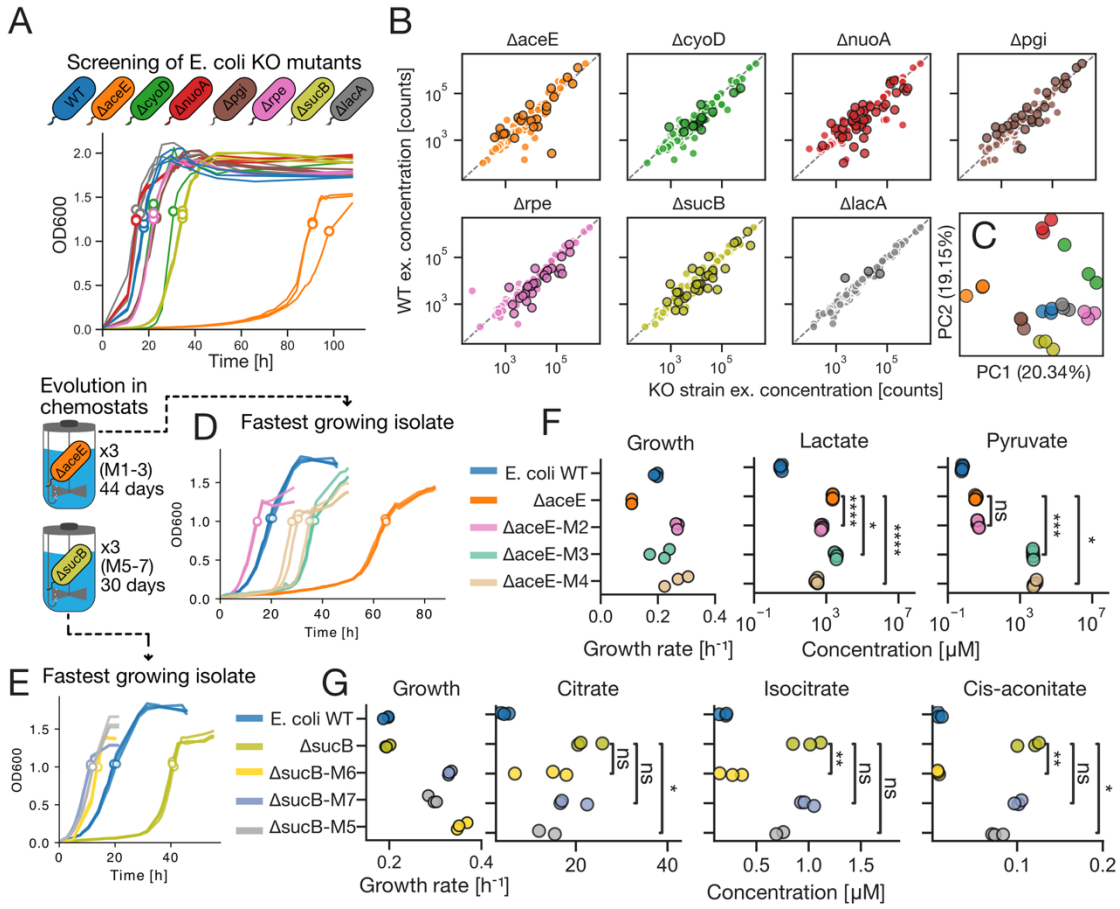

Figure 4: ***Effect of gene knockouts on metabolite release rates and the effect of subsequent evolution to higher fitness.*** *A) Growth curves of selected KO mutants, with sampling time-points indicated by open circles. B) Comparison of extracellular metabolite concentrations in late exponential (OD_{600}~1.2-1.3) between KO mutants and WT. Significant differences are marked by solid black outlines (Welch's t-test, FDR < 0.05). C) A PCA plot of exometabolome profiles, showing that each strain has a distinct pattern. D & E) Batch culture growth curves and sampling timepoints for isolates from the endpoint of the evolution experiment for ΔaceE and ΔsucB, respectively. F & G) Growth rates and extracellular metabolite concentrations for metabolites near the deleted reaction, for ΔaceE and ΔsucB, respectively. Statistical significance was assessed using Welch's t-test.*

Discussion

To better understand metabolic interactions between microbes, it is necessary to develop and test hypotheses for the fundamental principles that govern them. Here we propose that the broad release of intracellular metabolites is in general a loss of energy and that for each metabolite there is a trade-off between the cost of its release and the cost of preventing it. How microbes act to prevent release remains poorly understood, but one possibility is expressing

transporters that immediately re-consume released metabolites (*56*, *57*). This hypothesis connects to previous work linking fitness costs to microbial traits, including amino acid uptake rates (*14*), proteome composition (*10*, *39*), properties of extracellular enzymes (*58*) and metabolic strategies (*59*).

We find consistent support for our hypothesis: (i) Metabolite values are negatively correlated with metabolite release rates across different species and nutrient environments, (ii) Metabolite value explains the most variability in release rates among all tested factors, (iii) Release rates cannot be explained by cell lysis alone. We also find that release rates are sensitive to changes in intracellular metabolism, both caused by gene knockouts and subsequent evolution. But on the timescale of our evolutionary experiments, changes in release rates were inconsistent across replicates and did not always match our expectations. This may be because alternative paths towards increased fitness were favoured that were independent of metabolite release, or that increased growth rate was selected for more than increased yield.

Beyond enabling order-of-magnitude predictions of release rates using metabolic models, these findings impact how we think about metabolite release and the evolution of microbial cross-feeding. While it is commonly assumed that cross-feeding either emerges from extracellular enzymes, from the release of waste products or is an act that benefits other species at the expense of the releaser (*2*, *60*–*63*), our results suggest another alternative: Microbes release metabolites because using, retaining or recapturing them is not worth the investment. This increases the scope for commensal cross-feeding, which in this scenario can arise without the need to invoke the evolution of cooperation or being vulnerable to cheaters (*64*).

References

1. C. Kost, K. R. Patil, J. Friedman, S. L. Garcia, M. Ralser, Metabolic exchanges are ubiquitous in natural microbial communities. *Nat. Microbiol.*, 1–9 (2023).

2. E. J. Culp, A. L. Goodman, Cross-feeding in the gut microbiome: Ecology and mechanisms. *Cell Host Microbe* **31**, 485–499 (2023).

3. A. E. Douglas, The microbial exometabolome: ecological resource and architect of microbial communities. *Philos. Trans. R. Soc. B Biol. Sci.* **375**, 20190250 (2020).

4. G. D'Souza, S. Shitut, D. Preussger, G. Yousif, S. Waschina, C. Kost, Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Nat. Prod. Rep.* **35**, 455–488 (2018).

5. L. M. Solden, A. E. Naas, S. Roux, R. A. Daly, W. B. Collins, C. D. Nicora, S. O. Purvine, D. W. Hoyt, J. Schückel, B. Jørgensen, W. Willats, D. E. Spalinger, J. L. Firkins, M. S. Lipton, M. B. Sullivan, P. B. Pope, K. C. Wrighton, Interspecies cross-feeding orchestrates carbon degradation in the rumen ecosystem. *Nat. Microbiol.* **3**, 1274–1284 (2018).

6. S. Pontrelli, R. Szabo, S. Pollak, J. Schwartzman, D. Ledezma-Tejeida, O. X. Cordero, U. Sauer, Metabolic cross-feeding structures the assembly of polysaccharide degrading communities. *Sci. Adv.* **8**, eabk3076 (2022).

7. M. Gralka, R. Szabo, R. Stocker, O. X. Cordero, Trophic Interactions and the Drivers of Microbial Community Assembly. *Curr. Biol.* **30**, R1176–R1188 (2020).

8. M. Daniels, S. van Vliet, M. Ackermann, Changes in interactions over ecological time scales influence single-cell growth dynamics in a metabolically coupled marine microbial community. *ISME J.* **17**, 406–416 (2023).

9. Y.-C. Chuang, N. W. Haas, R. Pepin, M. G. Behringer, Y. Oda, B. LaSarre, C. S. Harwood, J. B. McKinlay, Bacterial adenine cross-feeding stems from a purine salvage bottleneck. *ISME J.* **18**, wrae034 (2024).

10. M. T. Mee, J. J. Collins, G. M. Church, H. H. Wang, Syntrophic exchange in synthetic microbial communities. *Proc. Natl. Acad. Sci.* **111**, E2149–E2156 (2014).

11. K. Campbell, J. Vowinckel, M. Mülleder, S. Malmsheimer, N. Lawrence, E. Calvani, L. Miller-Fleming, M. T. Alam, S. Christen, M. A. Keller, M. Ralser, Self-establishing communities enable cooperative metabolite exchange in a eukaryote. *eLife* **4**, e09943 (2015).

12. N. Paczia, A. Nilgen, T. Lehmann, J. Gätgens, W. Wiechert, S. Noack, Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms. *Microb. Cell Factories* **11**, 122 (2012).

13. J. S. L. Yu, C. Correia-Melo, F. Zorrilla, L. Herrera-Dominguez, M. Y. Wu, J. Hartl, K. Campbell, S. Blasche, M. Kreidl, A.-S. Egger, C. B. Messner, V. Demichev, A. Freiwald, M. Mülleder, M. Howell, J. Berman, K. R. Patil, M. T. Alam, M. Ralser, Microbial communities form rich extracellular metabolomes that foster metabolic interactions and promote drug tolerance. *Nat. Microbiol.* **7**, 542–555 (2022).

14. M. Zampieri, M. Hörl, F. Hotz, N. F. Müller, U. Sauer, Regulatory mechanisms underlying coordination of amino acid and glucose catabolism in Escherichia coli. *Nat. Commun.* **10**, 3354 (2019).

15. G. D'Souza, S. Waschina, S. Pande, K. Bohl, C. Kaleta, C. Kost, Less Is More: Selective Advantages Can Explain the Prevalent Loss of Biosynthetic Genes in Bacteria. *Evolution* **68**, 2559–2570 (2014).

16. R. Gregor, G. T. Vercelli, R. E. Szabo, M. Gralka, R. C. Reynolds, E. B. Qu, N. M. Levine, O. X. Cordero, Vitamin auxotrophies shape microbial community assembly in the ocean. bioRxiv [Preprint] (2024). https://doi.org/10.1101/2023.10.16.562604.

17. B. Ryback, M. Bortfeld-Miller, J. A. Vorholt, Metabolic adaptation to vitamin auxotrophy by leaf-associated bacteria. *ISME J.* **16**, 2712–2724 (2022).

18. G. Yousif, F. Zorrilla, S. Dash, L. Oña, A. Shekhar, S. Giri, R. Guan, S. Harshe, M. Itermann, D. Welter, V. Benes, K. R. Patil, C. Kost, Obligate cross-feeding of metabolites is common in soil microbial communities. bioRxiv [Preprint] (2025). https://doi.org/10.1101/2025.01.29.635426.

19. V. Baldazzi, D. Ropers, J.-L. Gouzé, T. Gedeon, H. de Jong, Resource allocation accounts for the large variability of rate-yield phenotypes across bacterial strains. *eLife* **12**, e79815 (2023).

20. M. Basan, S. Hui, H. Okano, Z. Zhang, Y. Shen, J. R. Williamson, T. Hwa, Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).

21. M. Huelsmann, O. T. Schubert, M. Ackermann, A framework for understanding collective microbiome metabolism. *Nat. Microbiol.* **9**, 3097–3109 (2024).

22. C. Malina, R. Yu, J. Björkeroth, E. J. Kerkhoven, J. Nielsen, Adaptations in metabolism and protein translation give rise to the Crabtree effect in yeast. *Proc. Natl. Acad. Sci.* **118**, e2112836118 (2021).

23. T. Pfeiffer, S. Schuster, S. Bonhoeffer, Cooperation and Competition in the Evolution of ATP-Producing Pathways. *Science* **292**, 504–507 (2001).

24. M. L. Reaves, B. D. Young, A. M. Hosios, Y.-F. Xu, J. D. Rabinowitz, Pyrimidine homeostasis is accomplished by directed overflow metabolism. *Nature* **500**, 237–241 (2013).

25. R. Heinrich, S. Schuster, H.-G. Holzhütter, Mathematical analysis of enzymic reaction systems using optimization principles. *Eur. J. Biochem.* **201**, 1–21 (1991).

26. E. Costa, J. Pérez, J.-U. Kreft, Why is metabolic labour divided in nitrification? *Trends Microbiol.* **14**, 213–219 (2006).

27. M. J. van Hoek, R. M. Merks, Redox balance is key to explaining full vs. partial switching to low-yield metabolism. *BMC Syst. Biol.* **6**, 22 (2012).

28. G. N. Vemuri, M. A. Eiteman, J. E. McEwen, L. Olsson, J. Nielsen, Increasing NADH oxidation reduces overflow metabolism in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.* **104**, 2402–2407 (2007).

29. A. Bar-Even, E. Noor, A. Flamholz, J. M. Buescher, R. Milo, Hydrophobicity and Charge Shape Cellular Metabolite Concentrations. *PLoS Comput. Biol.* **7** (2011).

30. K. Campbell, L. Herrera-Dominguez, C. Correia-Melo, A. Zelezniak, M. Ralser, Biochemical principles enabling metabolic cooperativity and phenotypic heterogeneity at the single cell level. *Curr. Opin. Syst. Biol.* **8**, 97–108 (2018).

31. J. B. McKinlay, Are Bacteria Leaky? Mechanisms of Metabolite Externalization in Bacterial Cross-Feeding. *Annu. Rev. Microbiol.* **77**, 277–297 (2023).

32. S. Gude, G. J. Pherribo, M. E. Taga, Emergence of Metabolite Provisioning as a By-Product of Evolved Biological Functions. *mSystems* **5**, 10.1128/msystems.00259-20 (2020).

33. J. G. Lopez, N. S. Wingreen, Noisy metabolism can promote microbial cross-feeding, *eLife* (2022). https://doi.org/10.7554/eLife.70694.

34. B. J. Sánchez, C. Zhang, A. Nilsson, P.-J. Lahtvee, E. J. Kerkhoven, J. Nielsen, Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol. Syst. Biol.* **13**, 935 (2017).

35. P. S. Bekiaris, S. Klamt, Automatic construction of metabolic models with enzyme constraints. *BMC Bioinformatics* **21**, 19 (2020).

36. Z. Mao, J. Niu, J. Zhao, Y. Huang, K. Wu, L. Yun, J. Guan, Q. Yuan, X. Liao, Z. Wang, H. Ma, ECMpy 2.0: A Python package for automated construction and analysis of enzyme-constrained models. *Synth. Syst. Biotechnol.* **9**, 494–502 (2024).

37. J. D. Orth, I. Thiele, B. Ø. O. Palsson, What is flux balance analysis? *Nat Biotech* **28**, 245–248 (2010).

38. C. Gu, G. B. Kim, W. J. Kim, H. U. Kim, S. Y. Lee, Current status and applications of genome-scale metabolic models. *Genome Biol.* **20**, 121 (2019).

39. H. Akashi, T. Gojobori, Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis. *Proc. Natl. Acad. Sci.* **99**, 3695–3700 (2002).

40. M. Mori, C. Cheng, B. R. Taylor, H. Okano, T. Hwa, Functional decomposition of metabolism allows a system-level quantification of fluxes and protein allocation towards specific metabolic functions. *Nat. Commun.* **14**, 4161 (2023).

41. A. R. Pacheco, M. Moel, D. Segrè, Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nat. Commun.* **10**, 103 (2019).

42. T. Georgi, D. Rittmann, V. F. Wendisch, Lysine and glutamate production by Corynebacterium glutamicum on glucose, fructose and sucrose: Roles of malic enzyme and fructose-1,6-bisphosphatase. *Metab. Eng.* **7**, 291–301 (2005).

43. J. C. C. Vila, J. Goldford, S. Estrela, D. Bajic, A. Sanchez-Gorostiaga, A. Damian-Serrano, N. Lu, R. Marsland, M. Rebolleda-Gomez, P. Mehta, A. Sanchez, Metabolic similarity and the predictability of microbial community assembly. bioRxiv [Preprint] (2023). https://doi.org/10.1101/2023.10.25.564019.

44. R. Baran, E. L. Brodie, J. Mayberry-Lewis, E. Hummel, U. N. Da Rocha, R. Chakraborty, B. P. Bowen, U. Karaoz, H. Cadillo-Quiroz, F. Garcia-Pichel, T. R. Northen, Exometabolite niche partitioning among sympatric soil bacteria. *Nat. Commun.* **6**, 8289 (2015).

45. P. Wendering, Z. Nikoloski, COMMIT: Consideration of metabolite leakage and community composition improves microbial community reconstructions. *PLOS Comput. Biol.* **18**, e1009906 (2022).

46. B. D. Bennett, J. Yuan, E. H. Kimball, J. D. Rabinowitz, Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. *Nat. Protoc.* **3**, 1299–1311 (2008).

47. J. O. Park, S. A. Rubin, Y.-F. Xu, D. Amador-Noguez, J. Fan, T. Shlomi, J. D. Rabinowitz, Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nat. Chem. Biol.* **12**, 482–489 (2016).

48. T. Sajed, A. Marcu, M. Ramirez, A. Pon, A. C. Guo, C. Knox, M. Wilson, J. R. Grant, Y. Djoumbou, D. S. Wishart, ECMDB 2.0: A richer resource for understanding the biochemistry of E. coli. *Nucleic Acids Res.* **44**, D495–D501 (2016).

49. L. B. Thorfinnsdottir, L. García-Calvo, G. H. Bø, P. Bruheim, L. M. Røst, Optimized Fast Filtration-Based Sampling and Extraction Enables Precise and Absolute Quantification of the Escherichia coli Central Carbon Metabolome. *Metabolites* **13**, 150 (2023).

50. G. Falcón-Cano, C. Molina, M. Á. Cabrera-Pérez, Reliable Prediction of Caco-2 Permeability by Supervised Recursive Machine Learning Approaches. *Pharmaceutics* **14**, 1998 (2022).

51. X. Liu, B. Testa, A. Fahr, Lipophilicity and Its Relationship with Passive Drug Permeation. *Pharm. Res.* **28**, 962–977 (2011).

52. S. Pontrelli, K. B. Villi, A. Sichert, J. Trouillon, A. Rutz, Z. C. Landry, S. H. Rüdisser, R. Stocker, U. Sauer, Degradation of extracellular polymeric substances shapes microbial community diversity. *PLOS Biol.* **23**, e3003287 (2025).

53. S. E. R. Gibson, I. Frost, S. J. Hierons, T. Moses, W. C. K. Poon, S. A. West, M. J. Cann, Bacteria encode post-mortem protein catabolism that enables altruistic nutrient recycling. *Nat. Commun.* **16**, 1400 (2025).

54. T. Fuhrer, M. Zampieri, D. C. Sévin, U. Sauer, N. Zamboni, Genomewide landscape of gene–metabolome associations in Escherichia coli. *Mol. Syst. Biol.* **13**, 907 (2017).

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

56. D. E. Rancourt, J. T. Stephenson, G. A. Vickell, J. M. Wood, Proline excretion by Escherichia coli K12. *Biotechnol. Bioeng.* **26**, 74–80 (1984).

57. B. Enjalbert, P. Millard, M. Dinclaux, J.-C. Portais, F. Létisse, Acetate fluxes in Escherichia coli are determined by the thermodynamic control of the Pta-AckA pathway. *Sci. Rep.* **7**, 42135 (2017).

58. D. R. Smith, M. R. Chapman, Economical Evolution: Microbes Reduce the Synthetic Cost of Extracellular Proteins. *mBio* **1**, e00131-10 (2010).

59. A. Flamholz, E. Noor, A. Bar-Even, W. Liebermeister, R. Milo, Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci.* **110**, 10039–10044 (2013).

60. R. K. Fritts, A. L. McCully, J. B. McKinlay, Extracellular Metabolism Sets the Table for Microbial Cross-Feeding. *Microbiol. Mol. Biol. Rev.* **85**, 10.1128/mmbr.00135-20 (2021).

61. J. A. Lopez, B. Liu, Z. Li, M. S. Donia, N. S. Wingreen, When is microbial cross-feeding evolutionarily stable? bioRxiv [Preprint] (2025). https://doi.org/10.1101/2025.05.16.654511.

62. S. Pande, H. Merker, K. Bohl, M. Reichelt, S. Schuster, L. F. de Figueiredo, C. Kaleta, C. Kost, Fitness and stability of obligate cross-feeding interactions that emerge upon gene loss in bacteria. *ISME J.* **8**, 953–962 (2014).

63. D. Preussger, S. Giri, L. K. Muhsal, L. Oña, C. Kost, Reciprocal Fitness Feedbacks Promote the Evolution of Mutualistic Cooperation. *Curr. Biol.* **30**, 3580-3590.e7 (2020).

64. A. R. Zomorrodi, D. Segrè, Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. *Nat. Commun.* **8**, 1563 (2017).

65. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**, 6640–6645 (2000).

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

67. R. J. G. B. Campello, D. Moulavi, J. Sander, "Density-Based Clustering Based on Hierarchical Density Estimates" in *Advances in Knowledge Discovery and Data Mining*, J. Pei, V. S. Tseng, L. Cao, H. Motoda, G. Xu, Eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013), pp. 160–172.

68. H. Gallart-Ayala, I. Konz, F. Mehl, T. Teav, A. Oikonomidi, G. Peyratout, V. van der Velpen, J. Popp, J. Ivanisevic, A global HILIC-MS approach to measure polar human cerebrospinal fluid metabolome: Exploring gender-associated variation in a cohort of elderly cognitively healthy subjects. *Anal. Chim. Acta* **1037**, 327–337 (2018).

69. V. van der Velpen, T. Teav, H. Gallart-Ayala, F. Mehl, I. Konz, C. Clark, A. Oikonomidi, G. Peyratout, H. Henry, M. Delorenzi, J. Ivanisevic, J. Popp, Systemic and central nervous system metabolic alterations in Alzheimer's disease. *Alzheimers Res. Ther.* **11**, 93 (2019).

70. J. A. Myers, B. S. Curtis, W. R. Curtis, Improving accuracy of cell and chromophore concentration measurements using optical density. *BMC Biophys.* **6**, 4 (2013).

71. Q. Ren, B. Henes, M. Fairhead, L. Thöny-Meyer, High level production of tyrosinase in recombinant Escherichia coli. *BMC Biotechnol.* **13**, 18 (2013).

72. A. Goelzer, J. Muntel, V. Chubukov, M. Jules, E. Prestel, R. Nölker, M. Mariadassou, S. Aymerich, M. Hecker, P. Noirot, D. Becher, V. Fromion, Quantitative prediction of genome-wide resource allocation in bacteria. *Metab. Eng.* **32**, 232–243 (2015).

73. Y. Zhang, J. Cai, X. Shang, B. Wang, S. Liu, X. Chai, T. Tan, Y. Zhang, T. Wen, A new genome-scale metabolic model of Corynebacterium glutamicum and its application. *Biotechnol. Biofuels* **10**, 169 (2017).

74. S. L. McGill, Y. Yung, K. A. Hunt, M. A. Henson, L. Hanley, R. P. Carlson, Pseudomonas aeruginosa reverse diauxie is a multidimensional, optimized, resource utilization strategy. *Sci. Rep.* **11**, 1457 (2021).

75. J. Kuepper, J. Dickler, M. Biggel, S. Behnken, G. Jäger, N. Wierckx, L. M. Blank, Metabolic Engineering of Pseudomonas putida KT2440 to Produce Anthranilate from Glucose. *Front. Microbiol.* **6**, 1310 (2015).

76. J. Niu, Z. Mao, Y. Mao, K. Wu, Z. Shi, Q. Yuan, J. Cai, H. Ma, Construction and Analysis of an Enzyme-Constrained Metabolic Model of Corynebacterium glutamicum. *Biomolecules* **12**, 1499 (2022).

77. K. Wu, Z. Mao, Y. Mao, J. Niu, J. Cai, Q. Yuan, L. Yun, X. Liao, Z. Wang, H. Ma, ecBSU1: A Genome-Scale Enzyme-Constrained Model of Bacillus subtilis Based on the ECMpy Workflow. *Microorganisms* **11**, 178 (2023).

78. J. Guo, H. Zhang, C. Wang, J.-W. Chang, L.-L. Chen, Construction and analysis of a genome-scale metabolic network for *Bacillus licheniformis* WX-02. *Res. Microbiol.* **167**, 282–289 (2016).

79. J. Nogales, J. Mueller, S. Gudmundsson, F. J. Canalejo, E. Duque, J. Monk, A. M. Feist, J. L. Ramos, W. Niu, B. O. Palsson, High-quality genome-scale metabolic modelling of Pseudomonas putida highlights its broad metabolic capabilities. *Environ. Microbiol.* **22**, 255–269 (2020).

80. Z. Mao, J. Niu, J. Zhao, Y. Huang, K. Wu, L. Yun, J. Guan, Q. Yuan, X. Liao, Z. Wang, H. Ma, ECMpy 2.0: A Python package for automated construction and analysis of enzyme-constrained models. *Synth. Syst. Biotechnol.* **9**, 494–502 (2024).

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

82. A. Ebrahim, J. A. Lerman, B. O. Palsson, D. R. Hyduke, COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).

83. A. Varma, B. W. Boesch, B. O. Palsson, Stoichiometric interpretation of Escherichia coli glucose catabolism under various oxygenation rates. *Appl. Environ. Microbiol.* **59**, 2465–2473 (1993).

84. E. Reznik, P. Mehta, D. Segrè, Flux Imbalance Analysis and the Sensitivity of Cellular Growth to Changes in Metabolite Pools. *PLOS Comput. Biol.* **9**, e1003195 (2013).

85. Y. Ram, E. Dellus-Gur, M. Bibi, K. Karkare, U. Obolski, M. W. Feldman, T. F. Cooper, J. Berman, L. Hadany, Predicting microbial growth in a mixed culture from growth curve data. *Proc. Natl. Acad. Sci.* **116**, 14698–14707 (2019).

86. J. L. Reed, T. D. Vo, C. H. Schilling, B. O. Palsson, An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR). *Genome Biol.* **4**, R54 (2003).

87. A. M. Feist, C. S. Henry, J. L. Reed, M. Krummenacker, A. R. Joyce, P. D. Karp, L. J. Broadbelt, V. Hatzimanikatis, B. Ø. Palsson, A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* **3**, 121 (2007).

88. J. D. Orth, T. M. Conrad, J. Na, J. A. Lerman, H. Nam, A. M. Feist, B. Ø. Palsson, A comprehensive genome-scale reconstruction of Escherichia coli metabolism—2011. *Mol. Syst. Biol.* **7**, 535 (2011).

89. J. M. Monk, C. J. Lloyd, E. Brunk, N. Mih, A. Sastry, Z. King, R. Takeuchi, W. Nomura, Z. Zhang, H. Mori, A. M. Feist, B. O. Palsson, iML1515, a knowledgebase that computes Escherichia coli traits. *Nat. Biotechnol.* **35**, 904–908 (2017).

90. Y. Djoumbou Feunang, R. Eisner, C. Knox, L. Chepelev, J. Hastings, G. Owen, E. Fahy, C. Steinbeck, S. Subramanian, E. Bolton, R. Greiner, D. S. Wishart, ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J. Cheminformatics* **8**, 61 (2016).

91. S. Seabold, J. Perktold, "Statsmodels: Econometric and Statistical Modeling with Python" (Austin, Texas, 2010; https://doi.curvenote.com/10.25080/Majora-92bf1922-011), pp. 92–96.

92. M. Lukačišinová, B. Fernando, T. Bollenbach, Highly parallel lab evolution reveals that epistasis can curb the evolution of antibiotic resistance. *Nat. Commun.* **11**, 3105 (2020).

93. S. Chen, Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**, e107 (2023).

94. Simon Andrews, FastQC: A Quality Control Tool for High Throughput Sequence Data, Babraham Institute (2010).

95. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

96. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).

97. P. Cingolani, A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, D. M. Ruden, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).

98. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. arXiv arXiv:1207.3907 [Preprint] (2012). https://doi.org/10.48550/arXiv.1207.3907.

99. S. D. Burz, A.-L. Abraham, F. Fonseca, O. David, A. Chapron, F. Béguet-Crespel, S. Cénard, K. Le Roux, O. Patrascu, F. Levenez, C. Schwintner, H. M. Blottière, C. Béra-Maillet, P. Lepage, J. Doré, C. Juste, A Guide for Ex Vivo Handling and Storage of Stool Samples Intended for Fecal Microbiota Transplantation. *Sci. Rep.* **9**, 8897 (2019).

100. B. D. Bennett, E. H. Kimball, M. Gao, R. Osterhout, S. J. Van Dien, J. D. Rabinowitz, Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* **5**, 593–599 (2009).

101. M. Mori, S. Schink, D. W. Erickson, U. Gerland, T. Hwa, Quantifying the benefit of a proteome reserve in fluctuating environments. *Nat. Commun.* **8**, 1–8 (2017).

102. B. Volkmer, M. Heinemann, Condition-Dependent Cell Volume and Concentration of Escherichia coli to Facilitate Data Conversion for Systems Biology Modeling. *PLoS ONE* **6**, e23126 (2011).

103. V. Simensen, C. Schulz, E. Karlsen, S. Bråtelund, I. Burgos, L. B. Thorfinnsdottir, L. García-Calvo, P. Bruheim, E. Almaas, Experimental determination of Escherichia coli biomass composition for constraint-based metabolic modeling. *PLOS ONE* **17**, e0262450 (2022).

Acknowledgements

We thank all members of the Mitri lab at the University of Lausanne for discussions and feedback. We also thank people at the Department of Biotechnology and Nanomedicine at SINTEF for technical support and advice on the experimental design, and Vassily Hatzimanikatis and the Laboratory of Computational Systems Biotechnology at EPFL for feedback on the computational analyses. Furthermore, we would like to thank Simon van Vliet, Olga Schubert, Guilhem Panneau, Jean C. C. Vila and Uwe Sauer for good discussions, Olga Schubert and Kim Schlegel for providing the KEIO strains, Sebastian Burz, You Zheng Teo and Clément Vulin for support with the live/dead staining protocol, and Vladimir Sentchilo for support with the flow cytometry and sampling for metabolomics analyses. We also thank the Metabolomics unit at UNIL, and specifically Hector G. Ayala and Julijana Ivanisevic, not only for conducting the LC-MS analyses, but also for providing guidance and recommendations on related matters. Special thanks to Florent Mazel and Daniel Segrè for feedback on the draft manuscript. Finally, we thank Lilja B. Thorfinnsdottir, Jean C. C. Vila and Stephan Noack for sharing of experimental data.

Funding

Swiss National Science Foundation Swiss Postdoctoral Fellowship TMPFP3_217172 (SS),
National Center of Competence in Research Microbiomes grant SNF 51NF40_180575 (SM, SS, AG, MAV, JL)
Faculty of Biology and Medicine, University of Lausanne (EU)
Swiss National Science Foundation Eccellenza grant PCEGP3_181272 (SM, MAV)
SINTEF Industry (GB)
Research Council of Norway, SFI Industrial Biotechnology, grant no 309558 (GB)

Author contributions

Conceptualization: SS, SM; Methodology: SS, AQ, JSL, GB, MAV, EU; Investigation: SS, GB, AG, EU; Visualization: SS; Funding acquisition: SS, SM, PE; Project administration: SS, SM; Supervision: SS, SM; Writing – original draft: SS; Writing – review & editing: SS, AQ, GB, JSL, PE, EU, MAV, SM, AG

Competing interests

Authors declare that they have no competing interests.

Data and materials availability

Strains and isolates used in this study are available on request. Formatted data and code for data analyses and visualization are available on GitHub at https://github.com/Mitri-lab/metabolite-release. A permanent archive of this repository will be deposited to Zenodo upon publication. The raw sequencing data and metadata have been deposited to the NCBI SRA database under the BioProject ID PRJNA1270783.

Materials and Methods

Strains

All experiments were conducted with either *E. coli* K-12 MG1655, *E. coli* K12 BW25113 (wild-type ancestor for KEIO KO strains) or the following KO mutants from the KEIO collection (*55*): Δ*aceE* (JW0110), Δ*cyoD* (JW0419), Δ*lacA* (JW0333), Δ*nuoA* (JW2283), Δ*pgi* (JW3985), Δ*rpe* (JW3349), Δ*sucB* (JW0716), Δ*zwf* (JW1841). The KEIO KO strains were verified by PCR with one primer binding inside the Kan^R cassette inserted during gene knock-out (Keio_K1_Rev for genes on the (-) strand; Keio_KT_RvComp_Fw for genes on the (+) strand) and one gene-specific primer binding downstream of the knocked-out gene (primers are listed in Table S5) (*55*, *65*). For all experiments, the KEIO KO strains were used with the kanamycin resistance cassette in place.

Precultures and growth media

For all experiments detailed below the strains were precultured as described here unless otherwise stated: From glycerol stocks stored at -70 °C each strain was streaked out onto LB agar plates and incubated overnight (37 °C). Then, one colony of each strain was picked and used to inoculate liquid precultures (10 mL LB in 50 mL Erlenmeyer flasks) that were incubated overnight (37 °C, 200 rpm). The kanamycin-resistant KEIO KO strains and strains evolved from these were always precultured in liquid LB and streaked out onto LB agar plates with 25 µg/mL kanamycin, while *E. coli* K-12 MG1655 and *E. coli* K12 BW25113 were always precultured in LB and streaked onto LB agar plates without selection. Unless otherwise noted, precultures were washed three times by centrifugation, with the supernatant removed after each step, and the cells resuspended in M9 medium with no carbon source to eliminate residual LB before inoculation. Unless otherwise stated, M9 medium was always prepared by mixing (per litre): 200 mL 5X M9 salts (M6030, Sigma-Aldrich, St. Louis, MO, USA), 2 mL 1M $MgSO_4$, 0.1 mL 1M $CaCl_2$, 1 mL 0.5 g/L $FeSO_4 \cdot 7H_2O$, MQ water and appropriate amounts of a carbon source stock

solution. The final M9 medium was then pH-adjusted to pH = 7.0, filter sterilized (0.22 µm) and stored at 4°C until use.

Selecting carbon sources for *E. coli* K-12 MG1655 batch cultures

To complement the existing exometabolome data from batch cultures of *E. coli* on glucose (*12*), we aimed to select three carbon sources that differed in chemical class and caused variation in metabolic flux patterns. We used the ecGEM eciJO1366 of *E. coli* K-12 MG1655 without any modifications and parsimonious FBA to simulate optimal flux patterns across 104 different growth supporting metabolites (*35*, *66*). We then performed a PCA analysis of the predicted fluxes (as Boolean values) to compare flux patterns between the carbon sources (Fig. S5). We used HDBSCAN (*67*) to identify clusters and selected from the different clusters in total 10 compounds that also differed in chemical class (sugars, amino acids and organic acids) and their entry point into central carbon metabolism (Fig. S5). We then screened the growth of the selected carbon sources in triplicates over 48 hours in 96-well plates (Fig. S5). For this experiment, we prepared M9 medium with each of the 10 different carbon sources to a concentration scaled to 120 mM carbon atoms. From washed LB precultures, *E. coli* K-12 MG155 was inoculated to OD_{600} = 0.05 in 200 µL of media and cultivated with continuous shaking (double orbital, 425 cpm) in a BioTek Synergy H1 (Agilent Technologies, Winooski, VT, USA) plate reader at 37 °C for 48 hours. We ultimately chose galactose, L-alanine and L-malate, as they supported rapid growth, differed in chemical class, entered central metabolism at different locations and were grouped into different clusters based on flux patterns (Fig. S5).

E. coli K-12 MG1655 batch cultures in bioreactors with different carbon sources

E. coli K-12 MG1655 was cultivated in lab-scale bioreactors using the three chosen carbon sources. Their concentrations were normalized by the number of carbon atoms to a carbon atom concentration of 120 mM, corresponding to 20 mM galactose (G0625, Sigma-Aldrich), 30 mM L-malic acid (02290, Sigma-Aldrich) and 40 mM L-alanine (A7627, Sigma-Aldrich). M9 medium with these three carbon sources were prepared as described above using 25X carbon source stock solutions.

The strain was streaked from a frozen glycerol vial onto an LB agar plate and incubated at 37 °C overnight. One colony was transferred to a 250 mL baffled shake flask containing 40 mL LB medium. After 7 hours, a culture volume corresponding to a start OD_{600} = 0.05 was inoculated

to a 500 mL baffled shake flask containing 100 mL M9 + 40 mM L-alanine. This flask was incubated for 36 h (200 rpm, 37 °C). For the two other carbon sources, the strain was incubated in LB for 17 h before inoculating a culture volume corresponding to OD_{600} = 0.05 in a 500 mL baffled shake flask containing 100 mL M9 + 20 mM L-galactose or M9 + 30 mM L-malic acid, respectively. The flasks were incubated for 24 h (200 rpm, 37 °C). Before inoculation to bioreactors, M9 precultures were centrifuged (3220 rcf, 10 min, 4 °C), supernatant removed, and pellet resuspended in M9 medium without carbon.

Batch cultivations were performed in 1 L DASGIP bioreactors (Eppendorf DASGIP, Jülich, Germany) with an initial volume of 600 mL M9 medium supplemented with either 20 mM galactose, 30 mM L-malic acid or 40 mM L-alanine, three replicates per condition. The reactors were equipped with two Rushton impellers and probes for measurements of dissolved oxygen (DO) and pH. Submerged aeration was maintained at 0.5 vvm, and agitation was cascaded between 400 and 800, controlled by a DO set point of 30 %. pH set point was 7.0 and controlled using 2 M NaOH and 2 M HCl. Start OD_{600} was 0.05.

Sampling was performed regularly during the exponential growth phase with additional samples collected during stationary phase. OD_{600} was measured at all time points and supernatant stored: 2 mL culture was centrifuged (3220 rcf, 10 min, 4 °C), and supernatant frozen in -20 °C. At the end of exponential phase and at the end of the experiment, samples were collected for cell dry weight (CDW). 50 mL culture was centrifuged (3220 rcf, 10 min, 4 °C), supernatant removed, cell pellet resuspended in 50 mL MQ water, centrifuged, supernatant removed, cell pellet resuspended in 5-10 mL MQ water and transferred to a pre-weighed aluminium beaker and dried at 105 °C for 24 hours. Beakers were then weighed again and CDW calculated.

Exometabolome analyses of time-series samples from *E. coli* K-12 MG1655 batch cultures with different carbon sources

Four samples from each of the nine batch cultures (three replicates of three different media) plus media samples were analysed with LC-MS and GC-MS to measure the extracellular concentration of metabolites at different time points throughout the experiment. The time points were selected based on time-series cultivation data to cover the exponential growth phase (three samples) plus a stationary phase sample (Fig. S6). Samples from the different

replicates were aligned based on the timing of the end of the exponential growth phase (Fig. S6).

LC-MS of selected bioreactor samples was conducted by the Metabolomics facility at UNIL using their method "high-coverage targeted analysis of polar metabolites", following this procedure: Media samples were thawed on ice. An aliquot (20 µL) was extracted with the addition of ice-cold methanol (80 µL), vortexed (30 sec) and centrifuged (15000 rpm, 15 min, 4°C). The supernatants were transferred to vials for liquid chromatography – tandem mass spectrometry (LC-MS/MS) analyses. The extracts were analyzed by hydrophilic interaction chromatography coupled to tandem mass spectrometer (6496 iFunnel, Agilent Technologies) using multiple reaction monitoring - MRM approach in both, positive and negative ionization modes, to maximize the polar metabolome coverage. The analytical conditions have been described in detail elsewhere (*68*, *69*). Raw LC-MS/MS metabolome data were processed using the Mass Hunter Quantitative analysis software (Agilent Technologies). The peak areas (or extracted ion chromatograms (EICs) for the monitored MRM transitions) were used for relative comparison of metabolite levels between different conditions or groups of samples. Data quality assessment, including signal drift correction, was performed using pooled quality control (QC) samples analyzed periodically throughout the entire batch.

Metabolite quantification was performed using stable isotope labeled internal standards and calibration curves following the signal drift correction. Data processing was done using MassHunter Quantitative analysis. Peak area integration was manually curated, and concentrations were reported by selecting a 6-point portion of the linear standard curves relevant for the observed concentrations.

The relative LC-MS data was standardized before subsequent analyses and data processing. Both the relative and absolute LC-MS data were cleaned of obvious outliers. For the absolute data the outliers were caused by carry-over from previous samples leading to too high malate (six samples) and succinate (one sample) values. For the relative LC-MS data there were additional outliers of unknown reasons (39 values in total) related to in total 14 metabolites. For several metabolites there were more missing values (not detected/not quantified) in the absolute than in the relative data. For a few metabolites (phenylalanine, proline, creatine) we leveraged good linear relationships between relative and absolute data (R^2>0.8) to estimate the missing absolute values using a simple linear transformation.

GC-MS analysis of the same samples was conducted in the following way: Analytical standards were prepared as M9 medium with increasing concentrations of acetate, lactate, pyruvate, formate and propionate and stored at -70°C until use. Samples from the *E. coli* bioreactor batch cultures and analytical standards were thawed on iced. Then, 100 µL was transferred to an Eppendorf tube and kept on a cold block (-20 °C) while adding 5 µL 11% HCl and 500 µL diethyl ether. Tubes were capped, vortexed on a thermoblock (2000 rpm, 10 min, 1°C) and centrifuged (13 000 rcf, 5 min, 4°C). The top organic phase layer was pipetted to a glass vial and derivatized with 20 µL N-(t-butyldimethylsilyl)-N-methyltrifluoroacetamide (MTBSTFA) and vortexed briefly. Samples were then placed on a heat block (90 min, 35°C) and kept at 16°C until analysis. The samples were injected (1 µL) by a Pal3 autosampler onto an Agilent 8890-5977B GC-MSD (Agilent Technologies) with a VF-5MS (30 m x 0.25 mm x 0.25 mm) column. The samples were injected with a split ratio of 15:1, helium flow rate of 1 mL/min and inlet temperature of 230 °C. The temperature was held for 2 min at 50 °C, raised at 25 °C/min to 175 °C, 30 °C/min to 280 °C and held for 3.5 min. The MSD was run in scan mode from 40-500 Da. Analyte abundances were calculated using the MassHunter Quantitative Analysis software (Agilent Technologies). Absolute concentrations were calculated using standard curves.
Galactose concentrations were measured using the L-Arabinose/D-Galactose assay kit (K-ARGA, Megazyme, Bray, Ireland), following their rapid protocol for analyses in 96-well plates.

Estimation of metabolite release rates of *E. coli*, *B. licheniformis*, *C. glutamicum* and *S. cerevisiae* in glucose medium

Time-series data on biomass (in OD_{600}) and absolute extracellular metabolite concentrations from batch cultivations of *E. coli, B. licheniformis, C. glutamicum* and *S. cerevisiae* in high glucose concentrations (10-20 g/L) (*12*) was kindly provided by the authors (mean and standard deviations). For metabolites reported as a sum of two metabolites we attributed equal amounts to those metabolites to enable direct comparison with metabolite values estimated for individual metabolites (described below). This includes 2-phosphoglycerate/3-phosphoglycerate and ribulose 5-phosphate/xylulose 5-phosphate. To estimate the metabolite release and uptake rates of each species we first converted the OD_{600} values to gDW/L. For *E. coli* we used the conversion factor we estimated from our own cell dry weight measurements in M9 galactose medium (0.346 gDW/L/OD_{600}), which is comparable, but on the lower end of values found in the literature (0.36-0.515 gDW/L/OD_{600}) (*70*, *71*). For the three other species we used literature

values: B. licheniformis (data from B. subtilis) 0.48 gDW/L/OD_{600} (*72*); *C. glutamicum* 0.27 gDW/L/OD_{600} (*73*); *S. cerevisiae* (mean of literature values) 0.746 gDW/L/OD_{600} (*70*). We then integrated the translated biomass time-series data to obtain biomass AUC data that we leveraged together with the mean absolute exometabolome data to estimate the uptake or release rates for each measured metabolite for each species using linear regression (Figs. S1-S4). Rates were only estimated for the exponential growth phase, and only until each metabolite either saturated or was clearly being re-consumed.

Theoretical justification of estimating uptake or release rates from extracellular concentrations

We start out from a simple assumption that the change in extracellular concentration C_i of metabolite *i* can be explained by the net release or uptake $a_i(t)$ (by any mechanism) by the microbes with biomass density $X(t)$ in the batch culture:

$$\frac{dC_i}{dt} = X(t) \cdot a_i(t)$$

The concentration $C_i(t)$ at a given timepoint *t* is thus:

$$C_i(t) = \int_{t0}^{t} X(t) \cdot a_i(t)dt + C_i(t_0)$$

If we assume that the uptake or release rate $a_i(t)$ is constant, we see that the concentration $C_i(t)$ is described by a linear function of the constant release rate and the integrated biomass (area under the curve, AUC):

$$C_i(t) = a_i \int_{t0}^{t} X(t)dt + C_i(t_0) = a_i \cdot X_{AUC}(t) + C_0$$

Hence, when metabolites are released at constant rates the dynamics should be well explained by a linear regression of extracellular metabolite concentrations on biomass AUC, and the specific net rate (uptake or release) is the slope of the curve.

Estimation of metabolite release rates of *E. coli* in galactose, L-malate and L-alanine

For our bioreactor batch cultivations of *E. coli* in galactose, L-malate and L-alanine, we used paired cell dry weight measurements and OD_{600} readings to calculate conversion factors used to translate OD_{600} values for all time points (galactose: 0.346 $\pm$ 0.016; L-malate: 0.279 $\pm$ 0.017; L-alanine: 0.296 $\pm$ 0.011; mean $\pm$ std in gDW/L/OD_{600}). The translated time-series data was used to calculate the biomass area under the curve used to estimate metabolite release and uptake rates further described below.

We estimated the metabolite uptake or release rate by performing for each metabolite and each carbon source a linear regression of measured concentration (relative or absolute) vs biomass AUC using the three samples (timepoints T1-T3) acquired from the three replicate batch cultures during exponential phase plus the initial M9 medium (T0; Figs. S7-S9). To avoid underestimating release rates for metabolites that potentially were being re-consumed at the end of the exponential phase we discarded the last timepoint from the exponential phase in cases where the rate was positive (indicating metabolite release) but there was a significant reduction in extracellular concentration from T2-T3 ($P < 0.05$, one-sided t-test). For 17 (of 126) metabolites we chose to not include T0 for at least one of the three conditions in the linear regression, either because of very large variability of T0 or because the trend from T0 to T1 was clearly different from the general trend from T1-T3. We only estimated the rate for metabolites where we had at least three data points from at least two different time points. For the metabolites where we had both relative and absolute data, we used the absolute data if there were enough absolute data points to get adequate slope estimates. For seven metabolites (glutamine, alpha-aminoadiapate, serine, lysine, lactate, NAD and isocitrate) where both absolute and relative quantification was performed, we could estimate the rate more accurately on the relative data set (because of more missing data points in the absolute data set). However, there were sufficient absolute measurements to estimate the absolute spread (standard deviation) of these data points, enabling translation of the slope value from Z-score to µmol/gDW/h.

Estimation of metabolite release rates of *E. coli*, *P. putida,* an *Enterobacter* sp. and a *Pseudomonas* sp. in various carbon source environments

The dataset from Vila et al. (2023) was kindly provided by the authors. The dataset contains exometabolome data from three timepoints (16, 28, and 48 hours) of *E. coli* MG1655, *P. putida* KT2440, and two environmental isolates of the genera *Enterobacter* and *Pseudomonas,* respectively (*43*). All species were cultivated in M9 medium with one of five carbon sources (D-glucose, D-fructose, glycerol, pyruvate or L-malate). Additionally, the dataset includes exometabolome data from two timepoints (28 and 48 hours) of *E. coli* and the *Enterobacter* sp. on D-ribose, L-arabinose and D-galactose, and of *P. putida* and the *Pseudomonas* sp. on acetate, fumarate and succinate. To estimate metabolite release rates, we first used the plate reader growth curves to calculate biomass AUC after 16, 28 and 48 hours and to identify the end of exponential growth of each organism in each condition. The end of exponential phase was

automatically identified as the first peak in the second derivative of the smoothed growth curve with an OD_{600} above 60% of maximum. We converted the AUC from OD_{600} to gDW/L using the same conversion factor for all the different carbon source environments: 0.36 gDW/L/ OD_{600} for *E. coli* and the *Enterobacter* sp., mean of measured values and literature values (*70*, *71*); 0.476 gDW/L/ OD_{600} for *P. putida* and the *Pseudomonas* sp, mean of literature values (*74*, *75*). Finally, we performed a linear regression between extracellular metabolite concentrations and AUC on both duplicates of each strain/carbon source combination including all datapoints before the end of exponential growth + 2 hours (a buffer to account for some inaccuracy in the automatic detection). We only included rate estimates for metabolites that were quantified at least twice for each strain. As the dataset did not include an early timepoint or media measurement we assumed the initial concentration of each metabolite to be 0. If a metabolite was detected more at more than one timepoint we discarded the estimated rate if there was a qualitative discrepancy in slope value when the slope was estimated with or without the T0 concentration assumed to be 0. If the carbon source was among the measured metabolites, this metabolite was excluded from rate estimation in those specific conditions.

Genome-scale metabolic models

For *E. coli, S. cerevisiae and C. glutamicum* we used the previously developed enzyme-constrained genome-scaled metabolic models (GEMs) eciJO1366 (*35*), ecCGL1 (*76*) and ecYeastGEM 8.3.4 (*34*), respectively. For *B. licheniformis* there was no available enzyme-constrained GEM, and we therefore opted to use the enzyme-constrained GEM ecBSU1 of the closely related species *B. subtilis* (*77*), which was available at github.com/tibbdc/ecBSU1. A normal GEM for *B. licheniformis* is available (*78*), but with this model we could not achieve any feasible flux balance solution. To simulate the environmental isolate of the genus *Enterobacter* we used the *E. coli* GEM eciJO1366. For the environmental isolate of the genus *Pseudomonas* and for *P. putida* we used the enzyme-constrained version of the *P. putida* GEM iJN1463 (*79*) automatically generated by the ECMpy2 method (*80*), which we here refer to as eciJN1463.

We made a few minor changes to the published GEMs before predicting metabolite values. In the *E. coli* GEM eciJO1366 the proteome weight of the forward reaction *GALKr* (galactokinase) was extremely high (0.016), limiting the growth rate on galactose to much less than the growth rate we measured in M9 galactose medium. To enable eciJO1366 to reach the measured growth rate with galactose as the only carbon source we replaced the proteome weight of the forward

GALKr reaction with the proteome weight of the reverse *GALKr* reaction which was much smaller (0.00037). For the ecYeastGEM we only merged exchange reactions that were previously split to simplify its use. Both ecBSU1 and eciJN1463 were translated from json to sMOMENT format in xml (*35*), modified by merging split exchange reactions and fixing metabolite compartment names and annotations. For ecBSU1 we also removed the enzyme constraint on the reaction "PPAm_num1" (inorganic diphosphatase) as it caused unrealistic uptake and release rates of phosphate and diphosphate, respectively. The *C. glutamicum* GEM ecCGL1 was also converted from the ECMpy format (*36*) to the sMOMENT format (*35*) but otherwise kept unchanged. All GEMs were used with their default biomass equations as objective when maximizing growth in all FBA or pFBA analyses (*81*). Reframed 1.5.3 (github.com/cdanielmachado/reframed) or COBRApy 0.29.1 (*82*) were used to load, manipulate and run constraint-based analyses with Gurobi 10 (Gurobi Optimization, LLC) as the solver.

Calculation of metabolite values and metabolite turnover using genome-scale metabolic models

In constraint-based modelling a shadow price is an estimate of how sensitive the objective function in a flux balance analysis (FBA) is to increased influx or efflux of a metabolite (*83*, *84*). Hence, one can interpret the shadow price as *metabolite value* as it quantifies the impact on growth caused by the release of a metabolite (to have positive values we here define metabolite value as the shadow price multiplied by -1). A similar approach to measure the cost of metabolite release has been used to identify costless secretions (*41*). The metabolite value (or the shadow price) can be predicted for a specific organism in a specific context by using a GEM of the corresponding organism (or closely related) and then constraining this GEM to the specific context (i.e. primarily defining metabolite uptake rates). While shadow prices are usually estimated directly by the solver used to run FBA (*37*), the robustness of these values is often questioned because of numerical issues. To get more robust shadow price estimates we solved for each metabolite in each context a separate FBA, where we constrained the GEM to have an (additional) release of 0.01 mmol/gDW/h of that metabolite and quantified the change in the objective function, i.e. the change Δy in growth rate caused by the forced metabolite release. The metabolite value was then quantified as $\Delta y/0.01$. If the metabolite was predicted to be released by FBA, we added 0.01 mmol/gDW/h to the FBA-predicted value to estimate the rate. This was only relevant for acetate for *E. coli* in the L-malate condition, for ethanol, acetate, formate and pyruvate for *S. cerevisiae*, and for pyruvate for *P. putida* in the L-malate condition.

Using this approach, we estimated metabolite values for *E. coli* in batch cultures with glucose, galactose, L-malate or L-alanine as the carbon source by constraining the uptake of the corresponding exchange reaction in eciJO1366 to the uptake rate of the carbon source as estimated from the experimental data. The same approach was used for *C. glutamicum, B. licheniformis* and *S. cerevisiae* using the corresponding models. For *E. coli*, *P. putida* and the two environmental isolates of the genera *Enterobacter* and *Pseudomonas* we first extracted maximum growth rates for each organism in each environment from growth curves (using curveball (*85*)). Then, we used the respective GEMs to calculate the uptake rates of the corresponding carbon sources required to sustain the maximum growth rates. These uptake rates were then used to constrain the GEMs when estimating the metabolite values as described above. The turnover (i.e. the sum of positive fluxes producing an intracellular metabolite) was estimated by running parsimonious FBA (*66*), using the same GEMs constrained in the same way as described for above for metabolite values.

To validate that the negative correlation between log-transformed metabolite values and release rates we compared estimates from eciJO1366 with estimates from the 4 benchmark *E. coli* K-12 MG1655 GEMs iJR904 (*86*), iAF1260 (*87*), iJO1366 (*88*) and iML1515 (*89*) that are of increasing complexity. We then compared estimated metabolite values and estimated release rates using data from *E. coli* in bioreactors with glucose, galactose, L-malate or L-alanine as the sole carbon source, the same data used in Fig. 2A (Fig. S11). In the L-malate condition for the iJR904 model we discarded the metabolite value for formate as an outlier because it was extremely small ($<10^{-12}$), below the solver tolerance. Note that no metabolite values were negative and hence no other values were discarded upon log-transformation.

Metabolite classification, chemical properties and intracellular concentrations

The charge and molecular weight of metabolites were obtained from eciJO1366 (*35*). All other chemical properties and InChIKeys were obtained from PubChem using the python interface PubChemPy (github.com/mcs07/PubChemPy). The InChIKeys were then used to classify metabolites into a chemical taxonomy using ClassyFire (*90*). To get an appropriate level of detail of metabolic classification we used the "Class" level category. However, we separated the class "Carboxylic acids and derivatives" into "Amino acids" and "Carboxylic acids" based on the "Subclass" level. We also merged the classes "Hydroxy acids and derivatives" and "Keto acids

and derivatives" into the joint class "Keto / hydroxy acids". Finally, metabolites that did not fall into either of the abovementioned categories nor the class "Organooxygen compounds" were joined into the category "Other".

Linear models for evaluating the importance of different factors for predicting metabolite release rates

Linear models were created and analysed using statsmodels (*91*). Compound class and carbon source was incorporated as categorical variables. Metabolites predicted to have zero turnover was given a value of 10^{-4}, less than the minimum predicted turnover. Release rates where corresponding values were missing in any of the compared factors were discarded before model fitting. To evaluate the statistical significance of the out-of-sample predictions we performed 10^4 permutations of model fitting and out-of-sample estimations where the labels (either metabolite or condition) was shuffled on each permutation. We then calculated the fraction of out-of-sample R^2 values from these permutations below the observed R^2 value to estimate the P value.

Cultivation and exometabolome analyses of KEIO knockout strains

One objective was to study how disruption of key metabolic fluxes would affect extracellular metabolite concentrations and how this would change if these strains were allowed to evolve in a simple nutrient environment. The KEIO collection is a collection of non-essential *E. coli* KO strains that we could use for this purpose (*55*). To identify relevant KEIO KO strains we used eciJO1366 and parsimonious FBA (*81*) to predict the effect of a gene knockout on optimal metabolic fluxes, and previous metabolomics data (*54*) of these strains to see the effect on intracellular concentrations (Fig. S20A). We eventually chose seven strains (Δ*aceE*, Δ*cyoD*, Δ*nuoA*, Δ*pgi*, Δ*rpe*, Δ*sdhB*, Δ*sucB*) that were predicted to have different optimal flux distributions, had many significantly changed intracellular metabolite concentrations, and targeted different part of key metabolic pathways (Fig. S20). We included Δ*lacA* as a negative control expected to behave like the WT as done previously (*92*).

The selected strains from the KEIO collection were precultured on LB agar and liquid LB as previously described, and glycerol stocks used for experiments described below were prepared by adding glycerol to a final concentration of 25 %.

For exometabolome sampling, the KEIO strains and *E. coli* BW25113 were precultured as detailed above. They were then re-inoculated to OD_{600} = 0.05 in 10 mL M9 + 40 mM galactose (+ 25 µg/mL kanamycin for KEIO strains) in 50 mL Erlenmeyer flasks (200 rpm, 37 °C, 31 h). 1 mL culture was sampled into Eppendorf tubes and washed twice with M9 without carbon source by centrifugation (6000 rcf, 6 min, 4°C). Ultimately, samples were resuspended in 0.5 mL M9 without carbon source and used to inoculate 200 µL M9 + 40 mM galactose (+25 µg/mL kanamycin for KEIO strains) in a flat bottom 96-well plate to OD_{600} = 0.05 and 6 mL M9 + 40 mM galactose (+25 µg/mL kanamycin for KEIO strains) to straight glass tubes (15.8 cm tall, 1.5 cm diameter) to OD_{600} = 0.005. The 96-well plate was incubated in a Synergy H1 plate reader with continuous shaking (double orbital, 425 cpm, 37 °C), and the glass tubes in a shaking incubator (200 rpm, 37 °C, 45° angle), both for 108 hours. The experiment in the well plate was used to monitor more closely the growth (as OD_{600}) of these strains and to help with the timing of the exometabolome sampling. All strains were cultivated in three replicates. Two samples were collected during mid-exponential growth phase for all cultures, mostly to ensure that at least one sample from each replicate was obtained before the end of exponential phase. 1 mL culture volume was centrifuged (14 000 rpm, 10 min, 4 °C), and supernatant stored in -70 °C for exometabolome analysis. The second sample for each replicate (Fig. 4A) was analysed using LC-MS at the UNIL Metabolomics facility, following their protocol described above. Additionally, we also analysed three samples of pooled inoculums to verify that any carry-over with the inoculum was much lower (or below detection limit) of than the sample metabolite concentrations. Three replicate pooled samples were obtained by inoculating 6 mL M9 + 20 mM galactose medium with all the strains, each strain to an OD_{600} = 0.005, and subsequently handled as described above for the culture samples.

Chemostat evolution experiment with Δ*aceE* and Δ*sucB* in bioreactors

Two of the KEIO KO strains, Δ*aceE* and Δ*sucB*, were selected for the evolution experiment with continuous cultivation (chemostats) in bioreactors based on their reduced growth compared to their ancestor and difference in exometabolome patterns (Figs. 4A and S21-S23). The strains were streaked from glycerol vials onto LB + 25 µg/mL kanamycin agar plates and incubated at 37 °C overnight. 5-6 colonies were transferred to 500 mL baffled shake flasks containing 75 mL LB + 25 µg/mL kanamycin and incubated for 22 h (200 rpm, 37 °C).

Chemostat cultivations were performed in 1 L DASGIP bioreactors with a constant volume of 300 mL M9 + 20 mM galactose medium + 25 µg/mL kanamycin (prepared as previously described, but without pH adjustment because of practical challenges with the large volumes required, so pH ≈ 7.1). Cultivation was performed at 37 °C with a submerged airflow maintained at 0.5 vvm, and a cascaded agitation of 400-700 rpm, controlled by a DO set point of 30 %. pH control was started at day 6 (Δ*aceE*) or day 20 (Δ*sucB*) with a set point of 7.1, controlled using 2 M NaOH (Fig. S25A).

The reactors (three replicates per strain) were inoculated to OD_{600} = 0.1. During the late stage of exponential growth, pumps were started to feed fresh culture medium and remove used medium. Medium was fed using a pre-calibrated peristaltic pump and used medium was removed by a steel pipe placed right above the culture surface. The steel pipe was coupled to a peristaltic pump operating at a flow rate approximately 1.5x of the medium in flow, to avoid volume accumulation. The set dilution rate was adjusted several times during the experiment to compensate for (sometimes surprising) changes in growth dynamics (from 0-0.1 h^{-1} for Δ*aceE*, 0-0.3 for Δ*sucB,* Fig. S25B). The actual dilution rate was monitored by keeping the flasks with fresh culture medium on scales. OD_{600} was measured daily, and weekly, larger samples were collected. Culture was centrifuged (3220 rcf, 10 min, 4 °C), and supernatant and pellet stored at -70 °C. In addition, cultures were preserved at -70 °C as glycerol stocks. The number of generations were estimated from the measured dilution rates across the experiment plus the generations during the initial batch culture phase before the culture dilution was started.

Cultivation of single colonies from evolution experiment and sampling of exometabolome

Culture samples from different bioreactors and timepoints collected in the chemostat experiment were streaked out onto LB + 25 µg/mL kanamycin agar plates and incubated overnight at 37 °C. Three colonies were picked and individually inoculated in 20 mL LB + 25 µg/mL kanamycin in 100 mL Erlenmeyer flasks and incubated (37 °C, 200 rpm, 16 h). 12 mL of the culture was then centrifuged (3220 rcf, 6 min, RT), and the pellets stored at –20 °C for sequencing. The isolates were stored as glycerol stocks at -70 °C made from these cultures.

2 mL samples from the precultures were washed as previously described, resuspended in M9 without carbon source, and used to inoculate a flat bottom 96-well plate to OD_{600} = 0.05 with 200 µL M9 + 25 mM galactose + 25 µg/mL kanamycin. The M9 medium was prepared as

previously described, but to a pH of 7.1 as in the chemostats. The well plate was incubated in a Synergy H1 plate reader with continuous shaking (double orbital, 425 cpm, 37 °C, 120 h) and the OD_{600} was read every 20 minutes.

Based on these growth measurements we selected the fastest growing isolates from the last timepoint of each of the six bioreactors (three bioreactors per strain). These isolates, the initial KEIO KO strains Δ*aceE* and Δ*sucB*, and the KEIO ancestor *E. coli* BW25113 were precultured and subsequently washed as described above. The washed preculture were used to inoculate 10 mL M9 + 20 mM galactose to OD_{600} = 0.01 in straight glass tubes (15.8 cm tall, 1.5 cm diameter) in triplicates and incubated for 20-84 hours (200 rpm, 37 °C, 45° angle). 1 mL samples for exometabolome analysis were collected at $OD_{600} \approx 1$ (Fig. 4, D and E), filtered using a 0.22 µm Millex-GV PVDF syringe filter (SLGVR33RS, MilliporeSigma, Darmstadt, Germany) into a 1.5 mL Eppendorf tube stacked in a cold block (-20 °C), and then stored at -70 °C until it was analysed. The exometabolome LC-MS analyses were conducted by the UNIL Metabolomics facility, following the procedure described above.

DNA extraction of samples from selected isolates and chemostat culture samples

The DNA extraction of 14 different samples from the chemostat cultures (#37A D44 M2 Pellet, #38A D44 M3 Pellet, #39A D44 M4 Pellet, #28B D30 M5 Pellet, #29B D30 M6 Pellet, #30B D30 M7 Pellet) and the selected isolates (sucB-M5-D30-4, sucB-M6-D30-6, sucB-M7-D30-4, sucB-Ancestor, aceE-M2-D44-2, aceE-M3-D44-3, aceE-M4-D44-1, aceE-Ancestor) was performed by adapting the protocol from the FastPure Bacteria DNA Isolation Mini Kit (DC103-01, Vazyme, Nanjing, China). 1 mL GA buffer was added to the thawed pellet samples, and 230 µL of these resuspensions were kept in Eppendorf tubes for DNA extraction. The rest was centrifuged (4000 rpm, 6 min, RT) and the pellet stored at –20 °C. Genomic DNA (gDNA) was extracted from the 230 µL sample following the manufacturer's protocol. In the final step, elution of the DNA was performed twice with fresh elution buffer, eluting a total of 100 µL DNA. Extracted gDNA was stored at –70 °C until sequencing.

Genomic sequencing processing

Genomic DNA from the selected isolates and chemostat samples was sequenced by Novogene (Planegg, Germany) using the Illumina NovaSeqX PE 150. The isolates and the bioreactor samples were sequenced to 1 Gb and 10 Gb sequencing depth, respectively. Good quality of

both datasets was ensured using fastqc v. 0.12.1 before trimming the reads with fastp v. 0.24.0 (--qualified_quality_phred 20 --cut_front --cut_tail --average_qual 20 --length_required 50) (*93*, *94*). Both read sets were mapped against the *E. coli* BW25113 reference genome (NCBI RefSeq assembly GCF_000750555.1) with minimap2's v. 2.28 default "sr" parameters (*95*). The resulting alignments were filtered with samtools view v. 1.20 (-b -f 3 -q 60) to filter out reads not properly mapped in pairs or mapped to the reference with a MAPQ lower than 60 (*96*). Average coverage was assessed for both datasets showing a minimum average coverage for the chemostat culture samples of 2017 and 190 for the sequenced isolates.
For both datasets, variants were identified with freebayes v. 1.3.6 (--min-alternate-count 3 -p 1 --min-alternate-fraction 0.05 --pooled-continuous --haplotype-length 0) and annotated with snpEff v. 5.0 using the BW25113 reference genome (*97*, *98*). A custom python script was used to filter the annotated variants. For the isolates, variants with a frequency lower than 0.95 and a coverage lower than 100 were filtered out. For the chemostat culture samples, variants with a frequency lower than 0.05 and coverage lower than 100 were filtered out. A variant was considered fixed if its alle frequency was at least 0.9.

Paired intracellular and extracellular metabolomics and live/dead staining

M9 medium with 20 mM galactose, 30 mM malate or 40 mM L-alanine was prepared as detailed above. *E. coli* K-12 MG1655 was precultured on LB agar plates and in liquid LB as detailed above. 1 mL preculture was washed as described above and used to start a second preculture in 20 mL M9 + 40 mM L-alanine, inoculated to OD_{600} = 0.05. Another LB preculture of *E. coli* K-12 MG1655 was used to inoculate a second M9 + 20 mM galactose and a M9 + 30 mM malate preculture in the same way 12 hours later. After 36/24 hours, 1 mL was collected from each of the three precultures and washed twice as detailed above. These washed precultures were used to inoculate 10 mL of the same media (M9 + either 20 mM galactose, 30 mM malate or 40 mM L-alanine) in straight glass tubes (15.8 cm tall, 1.5 cm diameter). These glass tubes were incubated for two days (200 rpm, 37 °C, 45° angle). Growth was monitored by measuring OD_{600} directly in the glass tubes (NANOCOLOR VIS II, MACHEREY-NAGEL, Düren, Germany) at 21 timepoints. The glass tubes were vortexed briefly before any OD reading or sampling.
A reference exometabolome 0.5 mL sample was collected after 2 hours following the procedure detailed below. Then, 2 mL samples for paired intra- and extracellular metabolome and live/dead analyses were collected in the late exponential phase (at $OD_{600} \approx 1.1$, Fig. S30). Exometabolome samples were obtained by filtering 0.5 mL culture using a 0.22 µm Millex-GV PVDF syringe filter (SLGVR33RS, MilliporeSigma) into a 1.5 mL Eppendorf tube stacked in a cold

block (-20 °C). The samples for intracellular metabolomics, i.e. cells, were obtained immediately after following this procedure: A 25 mm 0.22 µm Durapore PVDF filter (GVWP02500, MilliporeSigma) was placed in a new filter holder. A 5 mL syringe was connected to the inlet of the filter holder and a vacuum pump (250 mbar) was connected to the outlet of the filter holder. Then, 2 mL of 37 °C MQ water was added to pre-wet the filter. Immediately after, 1 mL of culture sample was added to the syringe and allowed to pass through the filter. If the culture did not pass through solely based on the vacuum pump, light additional pressure was added by using the syringe itself to push the culture through the filter. This process took 10-30 seconds. Still using the same set-up, the filter was finally rinsed with 5 mL of 37 °C MQ water as previously recommended (*49*). Two tweezers were quickly rinsed in MQ water and used to carefully transfer the filter from the filter holder to 2 mL lysis tubes filled with 1.4 mL of 80:20 methanol:water (v/v), pre-chilled to -20 °C and kept on a cold block also pre-chilled to -20 °C. All samples were then quickly transferred to -70 °C for storage. Live/dead staining was performed using a similar approach as previous work (*99*): Briefly, the *live* sample was made by diluting 10 µL culture in 990 µL PBS and the *dead* sample was made by mixing 100 µL culture with 1 mL 70% isopropanol. Both samples were vortexed and incubated at room temperature for 1 h. Then, the isopropanol was removed from the dead sample by centrifugation (8000 rcf, 4 min, RT), pipetting off the supernatant, resuspending the pellet in 1 mL PBS and vortexing. This procedure was repeated once. Then, both the live and the dead sample were stained with propidium iodide (PI; P4170, Sigma-Aldrich) and SYBR green (Invitrogen S7563, Thermo Fisher Scientific, Carlsbad, CA, USA) with the following procedure: 48 µL sample (live or dead) was mixed in a 96-well plate with 49 µL PBS, 2 µL 0.5 mg/mL PI and 1 µL 100X SYBR green and incubated for 15 min in the dark. Then, a 0.1X samples of both the stained live and the stained dead samples were created by diluting 10 µL of each sample in 90 µL of PBS. The stained live/dead samples were then analyzed immediately on a CytoFLEX S flow cytometer (Beckman Coulter, Brea, CA, USA) and the CytExpert software (v2.4.0.28). The dead samples were used to make the gatings required to quantify the fraction of dead cells in the live samples (Table S6). The extra- and intracellular metabolomics samples were analysed by the metabolomics facility at UNIL to quantify the five metabolites leucine, glutamate, aspartate, fructose 6-phosphate and cis-aconitate. These metabolites were chosen because they were chemically and metabolically diverse and likely to be quantifiable both intracellularly and extracellularly based on previous experience with similar samples. Also, as glutamate is the most abundant intracellular metabolite (*49*, *100*), it serves as a best-case scenario in our attempt to test how much of the extracellular metabolite levels that can be explained by cell lysis and the associated release of

intracellular metabolites. Intracellular metabolite samples were first normalized by the amount of protein detected in the same sample and subsequently converted to absolute intracellular concentrations using the estimated protein density in *E. coli* ($13.5 \cdot 10^{-8}\ \mu g/\mu m^3$) (*101*). The fructose 6-phosphate value for sample 1B, timepoint 1, was discarded as an outlier as it was one order of magnitude larger than any other fructose 6-phosphate measurements.

Calculation of contribution from intracellular metabolites and cell lysis to extracellular concentrations

Approximate cell volume was estimated from measured OD_{600} and a previous estimate of OD to cell volume for *E. coli* of 3.6 µL/OD_{600}/mL (*102*). From approximate cell volume, measured absolute intracellular concentrations and cell lysis fractions we then calculated the corresponding change in extracellular concentrations. This potential contribution was then compared to the net change in extracellular concentrations between the T0 reference sample (after 2-hours) and the late exponential phase sample. If all T0 values were below the detection limit, we assumed the initial concentration of this metabolite to be 0. If only some values at T0 were below the detection limit we imputed these values to the mean of the detected values, giving conservative estimates for the net change in extracellular concentrations. To estimate a potential contribution from protein depolymerization we estimated cell density in gDW/L from OD_{600} (0.346 gDW/L/OD_{600}) and from this specific amino acid density (in g/L) from recent estimates of *E. coli* amino acid mass fractions (*103*). Where mass fractions were attributed to pairs of amino acids (glutamate/glutamine and aspartate/asparagine) we assumed equal contributions. Amino acid density was converted from g/L to µM by dividing by the amino acids' molecular weight subtracted the weight of a water molecule to account for the condensation reaction in protein polymerization. A similar procedure as described above was followed when we extended this analysis by using literature values for intracellular concentrations. When extracellular concentrations were reported as two indistinguishable molecules, we attributed equal amounts to the two metabolites. This includes 2-phosphoglycerate/3-phosphoglycerate and ribulose 5-phosphate/xylulose 5-phosphate. To identify typical biomass degradation products we used the metabolites consumed by the biomass equation of the E. coli model iML1515 (*89*), although excluding the soluble pool (as categorized in (*87*)) since this should correspond to the intracellular concentrations already accounted for. Of the metabolites in our dataset (Fig. 3B), only the amino acids were among the consumed biomass components and therefore considered to be typical biomass degradation products.