Microbes use transporters to regulate the release of metabolites based on value

Snorre Sulheim[1,2*], Gunn Broli[3], Peter F. Doubleday[4], Alisson Gillon[1], Julien S. Luneau[1], Prajwal Padmanabha[1], Andrew Quinn[1], Eric Ulrich[1], Margaret A. Vogel[1], Nicola Zamboni[4], 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
4. Institute of Molecular Systems Biology, ETH Zurich

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

Abstract

Microbial interactions are shaped by the exchange of metabolites, ranging from intermediates in central carbon metabolism to amino acids, vitamins and fermentation by-products[1]. Yet, the underlying reasons for metabolite release, and the mechanisms and factors influencing 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. By measuring metabolite release from 66 *E. coli* mutants lacking individual metabolite-transport genes, we find that transporters both cause and counteract net metabolite release. These findings show that metabolite release constitutes a fitness cost that microbes have evolved to reduce and reveal the underlying mechanism. This new conceptual explanation for metabolite release reshapes our understanding of how microbial interactions emerge and persist.

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, vitamins and organic acids[11–15]. 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[4,15–17]. 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[18–22], mechanisms of pathway regulation and maintenance of homeostasis[23], optimal catabolic pathway lengths in different environments[24,25] and redox balancing[26,27]. 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[28–31].

Here, we approach the problem by asking: does metabolite release generally confer a benefit or impose a cost on the producer? While recent work has explored how metabolite release can be beneficial, e.g. by alleviating costs associated with molecular noise[31,32], 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)[33–35]. 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[36,37]. The additional enzyme constraints account for the total enzyme pool limitation in cells[19], enabling prediction of relevant trade-offs and temporal dynamics, e.g. acetate overflow in *E. coli*[34]. Various measures of metabolite value have been proposed in the past[38–40]. 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[41].

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[42]. 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]. 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*[42]. 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

113 during model reconstruction. This conclusion was also supported by consistent results when
114 metabolite values were estimated using four other *E. coli* GEMs (Fig. S11).
115

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

118 The consistent negative correlation between metabolite release rates and metabolite values
119 encouraged us to ask how well this factor explains release variability compared to other metabolic
120 or physiochemical properties previously associated to metabolite release[32,43]. We focused on the
121 data from bioreactor batch cultures of *E. coli* with glucose, galactose, L-malate or L-alanine as the
122 carbon source, where conditions were well-defined, the measured metabolites were diverse and
123 reference values for intracellular metabolite concentrations were available[44–47]. We first built
124 univariate linear models to quantify the importance of these factors: metabolite value,
125 intracellular concentration, compound class, solubility in oil vs water (log P), molecular weight,
126 topological polar surface area, charge, hydrogen bond acceptor and donor count, rotatable bond
127 count, turnover and carbon source (Fig. 2A). Metabolite value explained 43% of the variability,
128 more than any other factor (Fig. 2B). The structure-based classification of compound class also
129 had good predictive power (29% explained), primarily because carboxylic and keto/hydroxy acids
130 were released at significantly higher rates than amino acids and other compounds (Fig. 2A).
131 However, these two factors are clearly confounded as class-level differences mirrored significant
132 differences in metabolite values (Fig. 2C). The number of rotatable bonds, solubility in oil vs water
133 (log P) and the number of hydrogen bonds, each related to membrane permeability[48,49], were
134 also predictive of metabolite release, suggesting passive diffusion as a relevant mechanism.
135 Metabolite turnover is predicted to correlate negatively with release rates, according to the noise-
136 averaging cooperation hypothesis (low turnover metabolites are more susceptible to noise and
137 therefore more useful to share within the population)[32], but our dataset shows little support for
138 this.
139

140 When we asked how well the best linear model could predict metabolite release rates, we found
141 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[30,31]. 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[44–47]. 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

170 accounted for by cell lysis, due to high intracellular concentrations and moderate to low
171 extracellular levels (Fig. S18).
172
173 Degradation of cell debris may also contribute to extracellular metabolite pools[50]. In *E. coli,* this
174 degradation is dependent on the presence of Lon protease in the cell lysate[51]. However, the
175 benefit of cell debris degradation by Lon was only significant well beyond exponential phase, and
176 only 19 ± 6% of oligopeptides (6-10 amino acids) in the lysate were catabolised by Lon over 20
177 hours[51]. Thus, in our dataset where 90% of the samples were collected before 26 hours (67% in
178 exponential phase), we expect only a fraction of the proteome from lysed cells to contribute to
179 the extracellular concentrations of amino acids. When we included 10% proteome degradation
180 in our analysis, only the extracellular levels of tryptophan and threonine additionally became fully
181 explained by cell lysis (Fig. S19). However, apart from amino acids, the metabolites in Fig. 3B are
182 not typical degradation products. Overall, these results suggest that cell lysis and proteome
183 degradation are insufficient to explain extracellular levels of most metabolites – consistent with
184 previous work[12,30].
185

186 ## Metabolic disruption and evolution change metabolite release rates

187 Given the effect of gene deletions on intracellular metabolite levels[52], we assumed that the lack
188 of negative correlation between metabolite value and release rate in *C. glutamicum* (Fig. 1D)
189 could be a result of its engineered metabolism[41]. To test the sensitivity of metabolite release
190 patterns to modifications in core metabolism, we measured the exometabolome of eight *E. coli*
191 knockout (KO) strains in late exponential phase in galactose medium (Fig. 4A). These gene
192 knockouts target different parts of core metabolism, were previously shown to affect intracellular
193 metabolite pools[52], and were predicted to change intracellular flux patterns (Fig. S20). In line with
194 our expectations, 14-38% (16-42 of 111) of the extracellular metabolite concentrations differed
195 significantly from the wild type across the KO strains (Welch's t-test, FDR < 0.05, Fig. 4B). In
196 contrast, only 2.7% of metabolite concentrations were different in the negative control strain
197 Δ*lacA*. Each KO strain displayed distinct exometabolome patterns (Fig. 4C), and we observed the
198 most pronounced differences for metabolites in proximity of the deleted reaction (Figs. S21-S23),

consistent with intracellular effects of enzyme-deletions[52]. The WT reference for these KO strains from the KEIO collection[53] - *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).

228 To estimate the fitness contribution of reducing metabolite release, we used ecGEMs to predict
229 relative growth rate penalties associated with the observed release rates, using the ancestors'
230 mean growth rates as references. In the Δ*aceE* ancestor, lactate release is predicted to reduce
231 fitness by 23% whereas pyruvate release has negligible impact (0.02%). The reduced release of
232 lactate in the isolates from chemostat M2 and M4 is predicted to reduce the negative fitness
233 effect to only 16% and 7%, respectively. These contributions are small compared to the observed
234 increase (94-146%). The pattern is similar for the Δ*sucB* isolates: the estimated net fitness
235 reduction of citrate, isocitrate and cis-aconitate in the ancestor is only 0.37%, far from sufficient
236 to explain the improved growth rates.
237
238 To better understand the reasons for increased growth and changed release rates, we identified
239 sequence variants in the selected fast-growing strains and mutations that had fixed in the
240 chemostat populations (Table S1-S4). Aside from the Δ*aceE*-M4 isolate, which likely
241 hypermutated due to a *mutT* mutation, the evolved Δ*aceE* isolates have few sequence variants,
242 and either related to galactose uptake (e.g. *galS* in M2, M4 and M7, regulating galactose uptake)
243 or modulating fluxes near the deleted reaction (e.g. *poxB* in M2, *ldhA* in M4, and *icd, aceK* and
244 *sdhA* in M5, M6 & M7 and M7, respectively). Fixed non-synonymous variants in the chemostat
245 populations support these observations, e.g. *ldhA* and *galS* were fixed in M2 and *ilvG* was fixed in
246 M4. The Δ*sucB* populations have many more fixed mutations, but among the ones shared across
247 replicates we find again *galS* and *aceK*. Together, these results show that metabolite release rates
248 change as microbes evolve, but in our experiments, these changes were likely secondary to other,
249 more dominant drivers of fitness – leading to inconsistent changes in metabolite release across
250 replicates.
251

252 Transporters cause and counteract metabolite release

253 Our findings so far suggest that microbes can modulate metabolite release rates. To assess the
254 underlying mechanism, we profiled the exometabolomes of 66 *E. coli* KO mutants from the KEIO
255 collection, each strain devoid of a gene associated with active or passive transport across the
256 inner or outer membrane (Table S5). These genes were selected to span the diversity of

257 transporters, covering major functional classes (ABC, PTS, symporter, antiporter, porin, and
258 channel), a variety of substrates (amino acids, sugars, nucleosides, carboxylic acids, and others),
259 and expected direction (import, export, or both). We cultivated each KO mutant and the WT in
260 glucose minimal medium and quantified changes in extracellular metabolite concentrations over
261 the course of batch growth using flow-injection analysis time-of-flight mass spectrometry (FIA-
262 TOF MS)[54]. The final dataset comprises 2184 unique samples collected at 10 timepoints, yielding
263 more than 8700 mass spectrometry measurements. We could annotate *m/z* ion peaks to 333
264 metabolites in the *E. coli* metabolome. Three strains (Δ*tolC,* Δ*dppB,* and Δ*ptsG*) showed severely
265 reduced growth and were excluded from downstream analyses. We quantified the *effect* of each
266 transporter on each metabolite as the ratio between the mean KO-WT difference and the within-
267 strain variability (Fig. 5A). Statistical significance for all 15,848 strain-metabolite *effects* was
268 assessed using a permutation test (Methods). The dataset is made available for interactive
269 exploration at keio.unil.ch.
270
271 A positive effect implies that removing a transporter increases the net release of the respective
272 metabolite in the KO strain compared to the WT, while a negative effect implies the opposite. We
273 found a strong positive effect in 2.2% (343/15848) of the strain-metabolite pairs (*Effect* > 3*, P* <
274 0.005). In many cases, the strongest effect matched the annotated function of the transporter
275 (Fig. 5B). For example, deletion of the high-affinity proline/sodium symporter gene *putP* strongly
276 increased proline release, suggesting that PutP counteracts proline release in the WT by re-
277 importing proline lost via other mechanisms. This mechanism has already been described for *E.*
278 *coli* strains selected for proline secretion[55]. Similarly, deletion of the oxoglutarate importer *kgtP*
279 most strongly increased oxoglutarate release. For glutamate, three of the four annotated
280 transporters linked to positive effects, suggesting that they collectively counteract its release (*P* <
281 0.005, Fig. 5C). However, we also observed strong, and sometimes diverse effects extending
282 beyond the metabolites annotated to the respective transporters in EcoCyc[56]. For example,
283 deletion of the glutamine importer *glnP* strongly increased the extracellular concentration of 20
284 metabolites (*Effect* > 3, *P* < 0.005), including valine, hydroxybutanoic acid, aminobutanoic acid,
285 isoleucine/leucine, alanine, and glutamine (Fig. S30).

286

287 Strong negative effects were found in 2.4% (379/15848) of the strain-metabolite pairs (*Effect* < -
288 3, $P < 0.005$), showing that transporters frequently cause metabolite release. Some of these
289 effects were linked to transporters primarily associated with metabolite export, such as the valine
290 exporter YgaZ or RhtC, which mediates threonine efflux (Fig. 5D)[57,58]. However, we observed
291 equally strong effects for transporters that are important for metabolite uptake such as the purine
292 and adenosine transporters PunC and AdeP, respectively (Fig. 5E)[59–61]. Similarly, the "exporters"
293 YgaZ and RhtC showed positive effects comparable in magnitude to their negative ones (Fig. S30).
294

295 To test whether the measured effects of transporter deletion on metabolite release are
296 ecologically relevant, we co-cultivated the WT and 2-4 KO strains with a proline, methionine or
297 isoleucine auxotroph. With one exception, the Δ*ygaZ* strain which supported the growth of the
298 Δ*ilvE* isoleucine auxotroph better than the WT despite the measured negative effect on
299 isoleucine/leucine, strains with a measured positive effect on amino acid release supported the
300 growth of the corresponding auxotrophs better than the WT (Fig. S31).
301

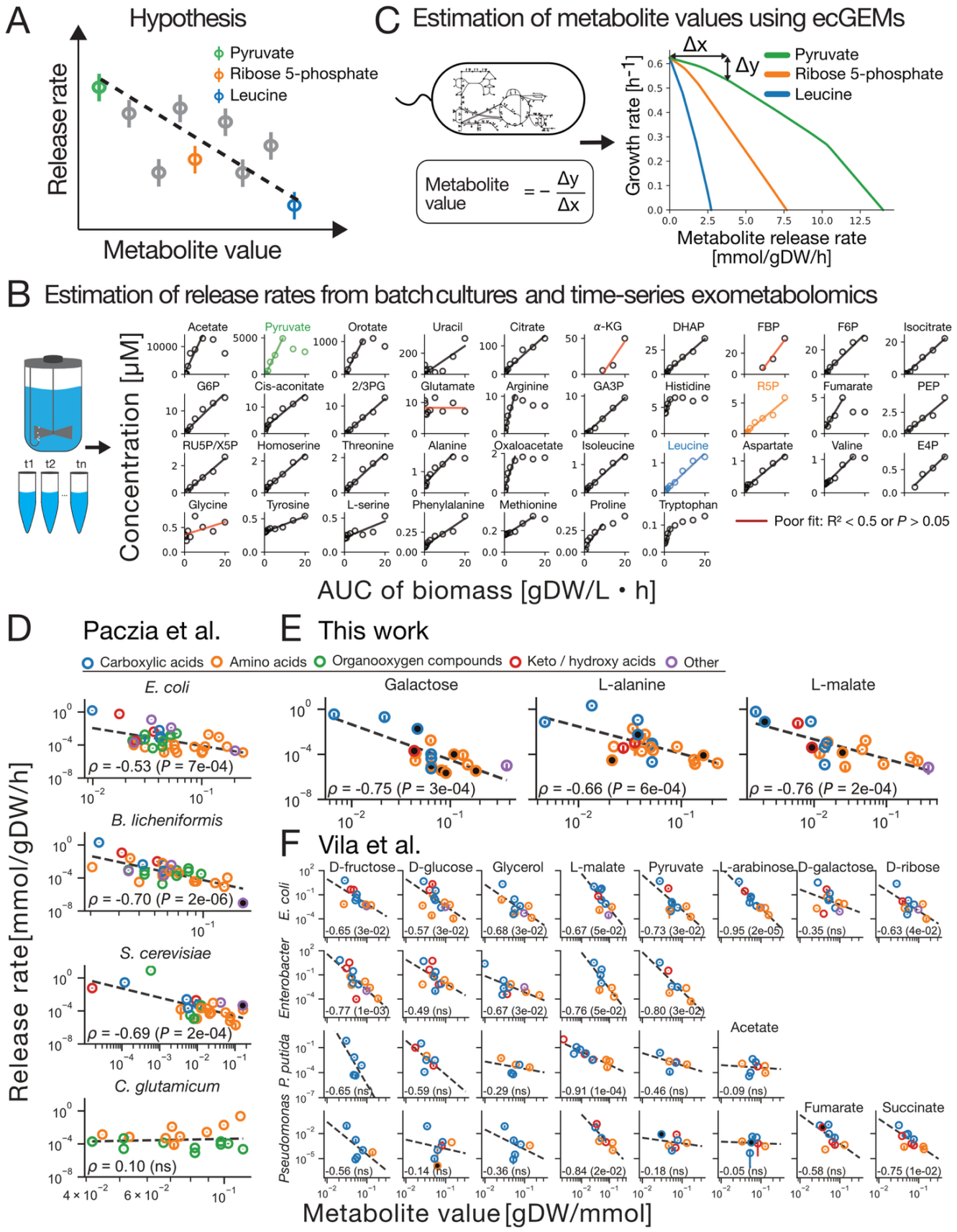
302 Of the 63 KO strains with growth comparable to the WT, 55 strains strongly influenced one or
303 more metabolite (|*Effect*| > 3, $P < 0.005$, Fig. 5F). Twenty-nine strains strongly influenced ten or
304 more metabolites, and 37 showed both positive and negative effects. While the balance between
305 positive and negative effects differed between strains, the overall dataset showed a striking
306 symmetry with almost equally strong and numerous positive and negative effects (Fig. 5G). From
307 the metabolite perspective, 122 metabolites were affected by one or more strains, and 57 of
308 these were affected in both directions (|*Effect*| > 3, $P < 0.005$, Fig. 5F). We found no significant
309 differences in the distribution of effect sizes across metabolite classes, nor any correlation
310 between the value of a metabolite and how frequently it was affected by the deletion of a
311 transporter (Fig. S31). Likewise, effect sizes were similarly distributed across transporter classes,
312 and neither the number of affected metabolites nor the maximum effect size correlated with
313 average transporter expression levels (Fig. S31)[62]. Taken together, our results demonstrate that
314 transporters expressed in *E. coli* are both a cause of metabolite release and a mechanism for *E.*

315 *coli* to counteract it. The complexity of strain-metabolite connections suggests that trade-offs are
316 common.

317 Discussion

318 To better understand metabolic interactions between microbes, it is necessary to develop and
319 test hypotheses for the fundamental principles that govern them. Here we propose that the broad
320 release of intracellular metabolites is in general a loss of energy and that for each metabolite
321 there is a trade-off between the cost of its release and the cost of preventing it. This hypothesis
322 connects to previous work linking fitness costs to microbial traits, including amino acid uptake
323 rates[14], proteome composition[10,38], properties of extracellular enzymes[63] and metabolic
324 strategies[64].
325
326 We find consistent support for our hypothesis: (i) metabolite values are negatively correlated with
327 metabolite release rates across different species and nutrient environments, (ii) metabolite value
328 explains the most variability in release rates among all tested factors, (iii) release rates cannot be
329 explained by cell lysis alone, (iv) knock-out of metabolite-transport genes frequently change net
330 metabolite release, providing a mechanistic explanation for the observed release and for how
331 microbes can act to reduce the net release of valuable metabolites. The complexity of KO strain–
332 metabolite interactions (Fig. 5F) suggests why eliminating metabolite release entirely may be
333 challenging, given the apparent trade-off between release and recapture, and the need to express
334 transporters at basal levels to sense and adapt to environmental changes[65–67]. The high frequency
335 of metabolites with both significant positive and negative effects suggests that futile cycles might
336 be common and raises the question of whether such cycles contribute to intracellular
337 homeostasis, extending beyond their previously demonstrated role in metal ion regulation[68].
338 Furthermore, we find that release rates are sensitive to changes in intracellular metabolism, both
339 caused by gene knockouts and subsequent evolution. Yet on the timescale of our evolutionary
340 experiments, changes in release rates were inconsistent across replicates and did not always
341 match our expectations. This may reflect alternative evolutionary paths that improved fitness
342 independently of metabolite release, or selection favouring higher growth rate over higher yield.

The large-scale exometabolomics screen of *E. coli* transporter knockout strains provides a unique resource for refining transporter annotations in databases and metabolic models. In parallel, our finding that simple linear models based on metabolite value can predict order-of-magnitude release rates offers a new avenue for such models to forecast microbial interactions. More broadly, our findings reshape 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 or from the release of waste products[2,69,70], our results suggest another alternative: Microbes release metabolites because using, retaining or immediately recapturing them is not worth the investment. This increases the scope for commensal cross-feeding to occur, which in this scenario can arise without the need to invoke the evolution of cooperation[71] or being vulnerable to cheaters[72].


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]

355

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[42]. 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.*

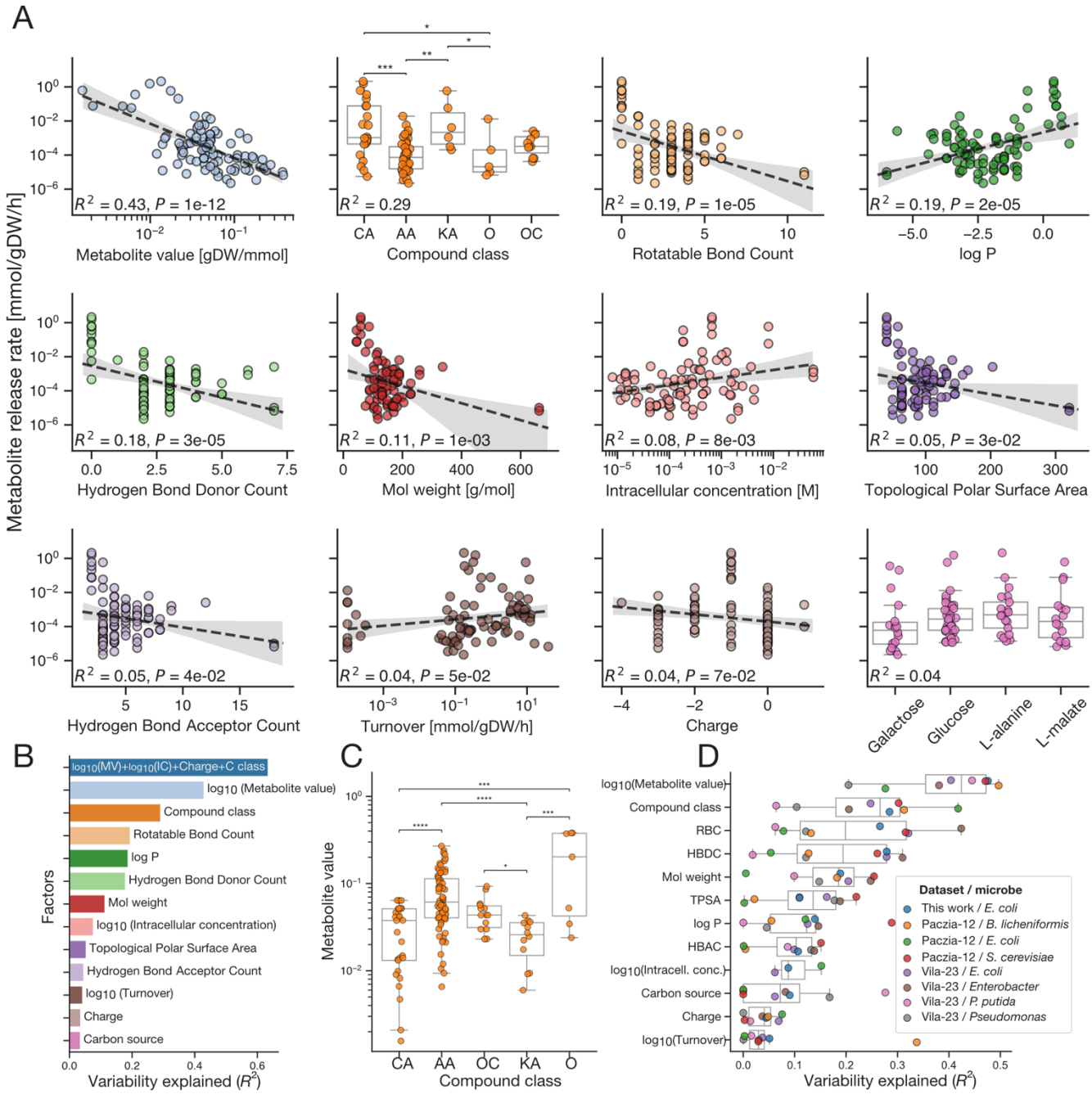


Figure 2: The importance of different factors in explaining metabolite release rates. *A) Among the tested chemical and metabolic factors, log-scaled metabolite value explains the most variability in log-scaled* E. coli *release rates. The tested factors are shown in order of how much variability (R^2) 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.*

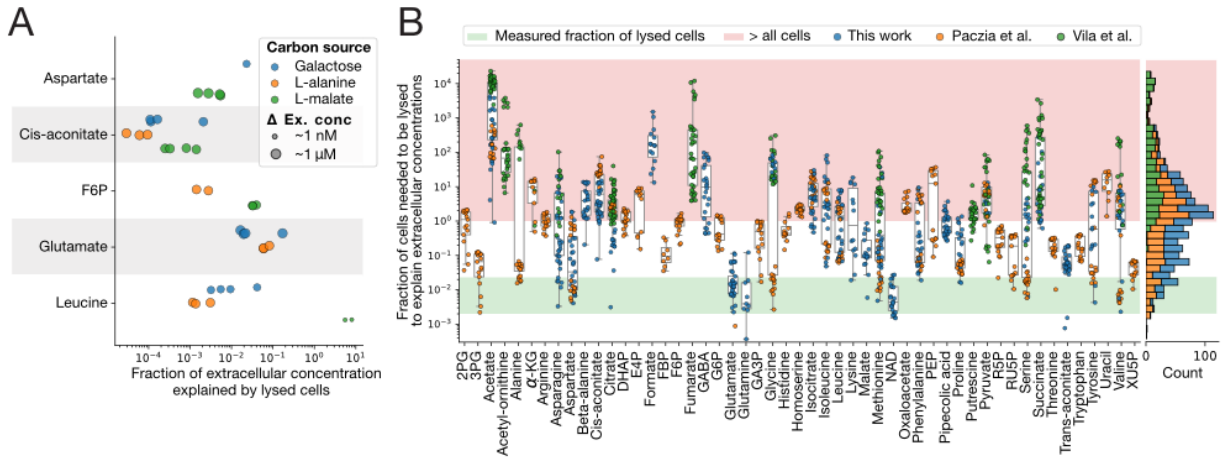


Figure 3: Contribution from cell lysis and associated release of intracellular metabolites to extracellular concentrations. *A) Using paired intra- and extracellular metabolite samples from the exponential phase of* E. coli *batch cultures we quantify what fraction of extracellular metabolite concentrations 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 ($OD_{600} \approx 1.1$), 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.*

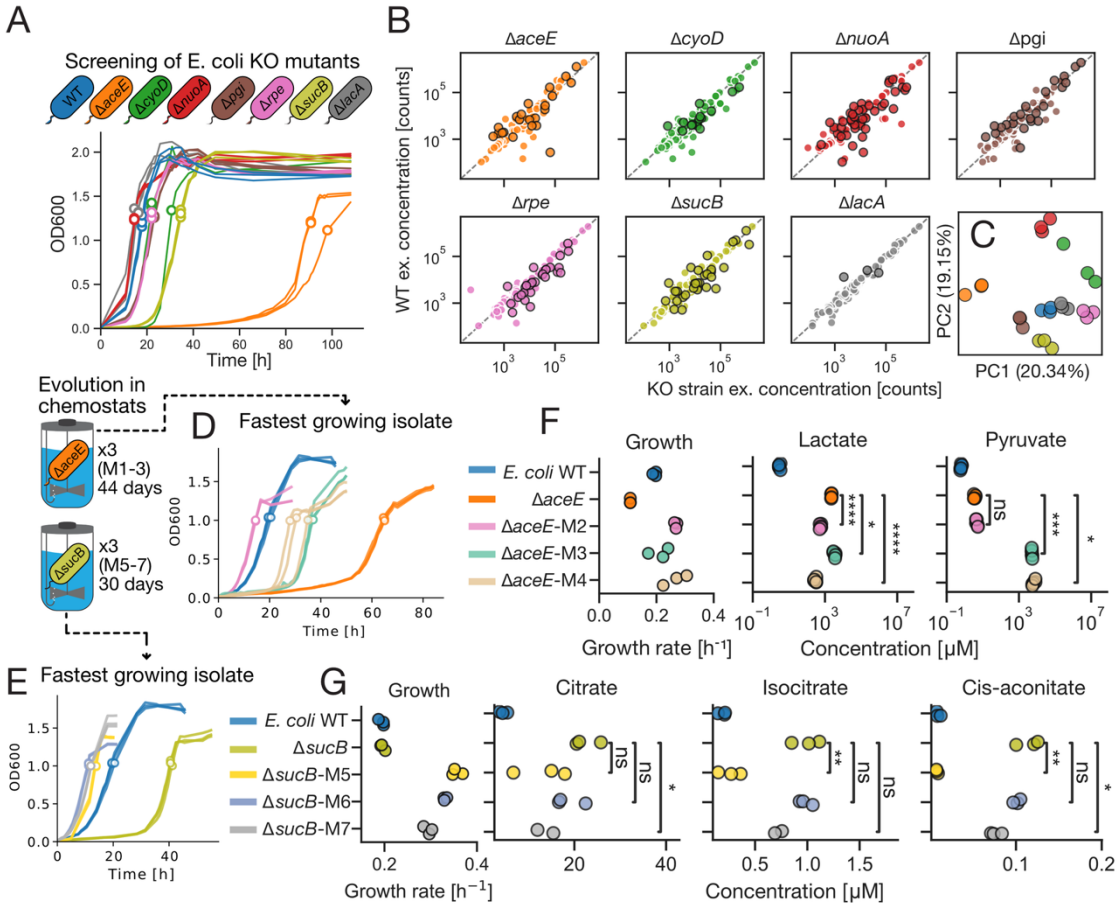


Figure 4: Effect of gene knockouts on metabolite release rates and the effect of subsequent evolution to higher fitness. *A) Growth curves of selected KO mutants, with sampling time-points indicated by open circles. B) Comparison of extracellular metabolite concentrations in late exponential (OD_{600}~1.2-1.3) 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.*

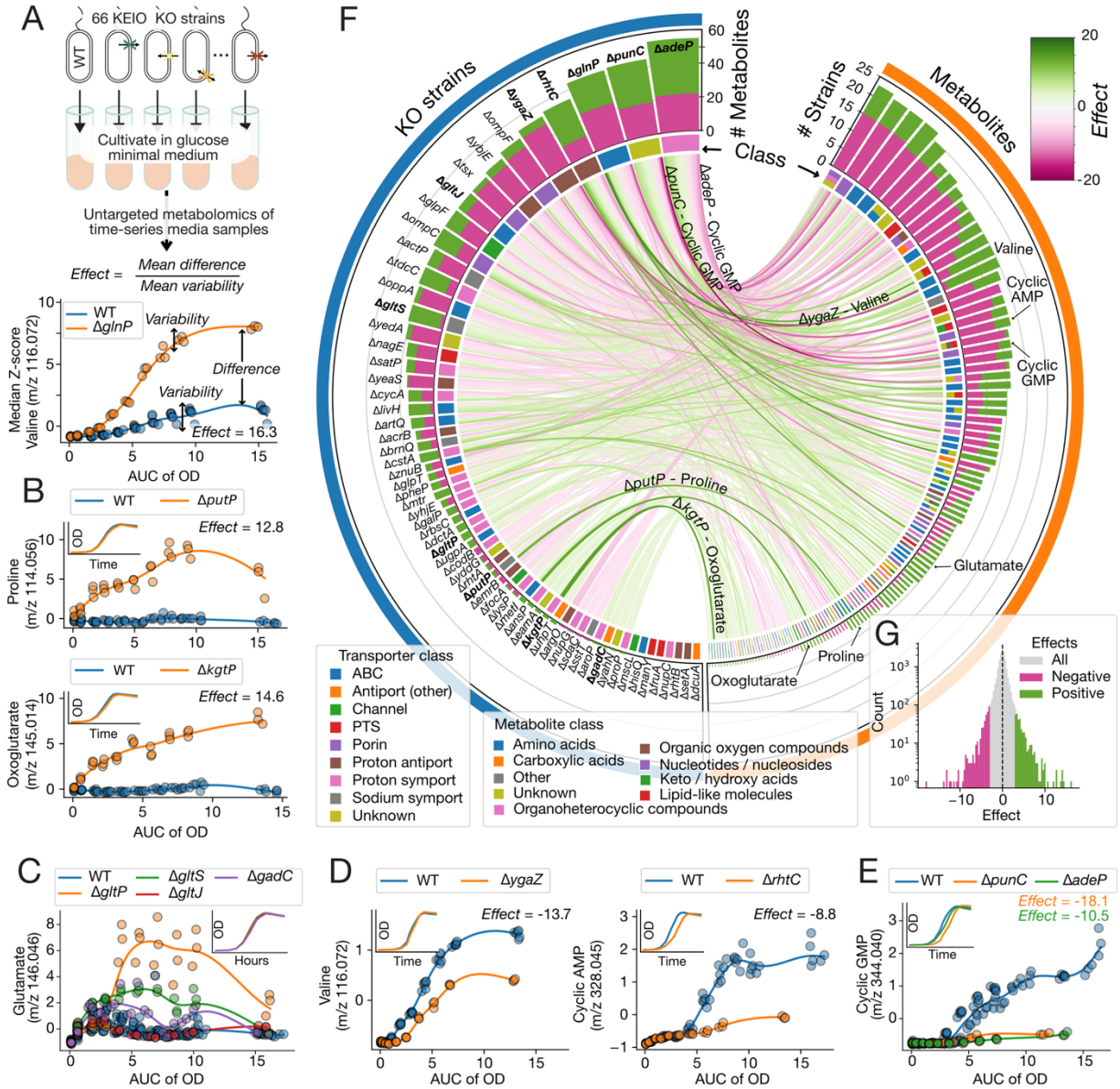


Figure 5: Transporters cause and counteract metabolite release. *A) Approach: We screened the temporal dynamics of extracellular metabolites in batch cultures for 66 different E. coli KO strains using untargeted metabolomics. We quantified the effect of each strain on each metabolite by comparing for each metabolite the ratio between the mean KO-WT difference and the within-strain variability (Methods). B-E) Each panel shows the extracellular concentration of one metabolite (Z-score, median of four technical replicates) for E. coli BW25113 (WT) and 1-4 KO mutants across a batch culture, with the area-under-the-curve of the optical density (AUC of OD) on the x-axis. Small inset panels compare the growth curves of the respective strains (mean OD, three technical replicates). F) Overview of significant, large effects (|Effect| > 3, P < 0.005, permutation test). The left bar chart displays the number of positive (green) and negative (pink) large effects across all 66 KO strains, and the color label indicates the class of the deleted transporter. The right bar chart displays the number of positive (green) and negative (pink) effects on each of the metabolites affected in one or more of the KO strains (122 in total, |Effect| > 3, P < 0.005, permutation test). The color label indicates the classification of all metabolites in E. coli mapping to the corresponding m/z value, which in some cases maps to two different classes. The chord diagram displays which metabolites that are affected (and the strength) in each KO strain. Metabolites / strains displayed in panel B-E annotated / highlighted in bold, and the strongest effects are also annotated on the links in the chord diagram. G) Distribution of all effects across all KO strains and metabolites.*

433 ## References

434 1. Kost, C., Patil, K. R., Friedman, J., Garcia, S. L. & Ralser, M. Metabolic exchanges are
435 ubiquitous in natural microbial communities. *Nat. Microbiol.* 1–9 (2023).
436 2. Culp, E. J. & Goodman, A. L. Cross-feeding in the gut microbiome: Ecology and mechanisms.
437 *Cell Host Microbe* **31**, 485–499 (2023).
438 3. Douglas, A. E. The microbial exometabolome: ecological resource and architect of microbial
439 communities. *Philos. Trans. R. Soc. B Biol. Sci.* **375**, 20190250 (2020).
440 4. D'Souza, G. *et al.* Ecology and evolution of metabolic cross-feeding interactions in bacteria.
441 *Nat. Prod. Rep.* **35**, 455–488 (2018).
442 5. Solden, L. M. *et al.* Interspecies cross-feeding orchestrates carbon degradation in the rumen
443 ecosystem. *Nat. Microbiol.* **3**, 1274–1284 (2018).
444 6. Pontrelli, S. *et al.* Metabolic cross-feeding structures the assembly of polysaccharide
445 degrading communities. *Sci. Adv.* **8**, eabk3076 (2022).
446 7. Gralka, M., Szabo, R., Stocker, R. & Cordero, O. X. Trophic Interactions and the Drivers of
447 Microbial Community Assembly. *Curr. Biol.* **30**, R1176–R1188 (2020).
448 8. Daniels, M., van Vliet, S. & Ackermann, M. Changes in interactions over ecological time scales
449 influence single-cell growth dynamics in a metabolically coupled marine microbial
450 community. *ISME J.* **17**, 406–416 (2023).
451 9. Chuang, Y.-C. *et al.* Bacterial adenine cross-feeding stems from a purine salvage bottleneck.
452 *ISME J.* **18**, wrae034 (2024).
453 10. Mee, M. T., Collins, J. J., Church, G. M. & Wang, H. H. Syntrophic exchange in synthetic
454 microbial communities. *Proc. Natl. Acad. Sci.* **111**, E2149–E2156 (2014).

11. Campbell, K. *et al.* Self-establishing communities enable cooperative metabolite exchange in a eukaryote. *eLife* **4**, e09943 (2015).

12. Paczia, N. *et al.* Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms. *Microb. Cell Factories* **11**, 122 (2012).

13. Yu, J. S. L. *et al.* Microbial communities form rich extracellular metabolomes that foster metabolic interactions and promote drug tolerance. *Nat. Microbiol.* **7**, 542–555 (2022).

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

15. Gregor, R. *et al.* Vitamin auxotrophies shape microbial community assembly on model marine particles. *ISME J.* **19**, wraf184 (2025).

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

17. Yousif, G. *et al.* Obligate cross-feeding of metabolites is common in soil microbial communities. 2025.01.29.635426 Preprint at https://doi.org/10.1101/2025.01.29.635426 (2025).

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

19. Basan, M. *et al.* Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).

476 20. Huelsmann, M., Schubert, O. T. & Ackermann, M. A framework for understanding collective
477 microbiome metabolism. *Nat. Microbiol.* **9**, 3097–3109 (2024).
478 21. Malina, C., Yu, R., Björkeroth, J., Kerkhoven, E. J. & Nielsen, J. Adaptations in metabolism and
479 protein translation give rise to the Crabtree effect in yeast. *Proc. Natl. Acad. Sci.* **118**,
480 e2112836118 (2021).
481 22. Pfeiffer, T., Schuster, S. & Bonhoeffer, S. Cooperation and Competition in the Evolution of
482 ATP-Producing Pathways. *Science* **292**, 504–507 (2001).
483 23. Reaves, M. L., Young, B. D., Hosios, A. M., Xu, Y.-F. & Rabinowitz, J. D. Pyrimidine homeostasis
484 is accomplished by directed overflow metabolism. *Nature* **500**, 237–241 (2013).
485 24. Heinrich, R., Schuster, S. & Holzhütter, H.-G. Mathematical analysis of enzymic reaction
486 systems using optimization principles. *Eur. J. Biochem.* **201**, 1–21 (1991).
487 25. Costa, E., Pérez, J. & Kreft, J.-U. Why is metabolic labour divided in nitrification? *Trends*
488 *Microbiol.* **14**, 213–219 (2006).
489 26. van Hoek, M. J. & Merks, R. M. Redox balance is key to explaining full vs. partial switching to
490 low-yield metabolism. *BMC Syst. Biol.* **6**, 22 (2012).
491 27. Vemuri, G. N., Eiteman, M. A., McEwen, J. E., Olsson, L. & Nielsen, J. Increasing NADH
492 oxidation reduces overflow metabolism in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.*
493 **104**, 2402–2407 (2007).
494 28. Bar-Even, A., Noor, E., Flamholz, A., Buescher, J. M. & Milo, R. Hydrophobicity and Charge
495 Shape Cellular Metabolite Concentrations. *PLoS Comput. Biol.* **7**, (2011).

496 29. Campbell, K., Herrera-Dominguez, L., Correia-Melo, C., Zelezniak, A. & Ralser, M.
497 Biochemical principles enabling metabolic cooperativity and phenotypic heterogeneity at
498 the single cell level. *Curr. Opin. Syst. Biol.* **8**, 97–108 (2018).
499 30. McKinlay, J. B. Are Bacteria Leaky? Mechanisms of Metabolite Externalization in Bacterial
500 Cross-Feeding. *Annu. Rev. Microbiol.* **77**, 277–297 (2023).
501 31. Gude, S., Pherribo, G. J. & Taga, M. E. Emergence of Metabolite Provisioning as a By-Product
502 of Evolved Biological Functions. *mSystems* **5**, 10.1128/msystems.00259-20 (2020).
503 32. Lopez, J. G. & Wingreen, N. S. Noisy metabolism can promote microbial cross-feeding. *eLife*
504 **11**, e70694 (2022).
505 33. Sánchez, B. J. *et al.* Improving the phenotype predictions of a yeast genome-scale metabolic
506 model by incorporating enzymatic constraints. *Mol. Syst. Biol.* **13**, 935 (2017).
507 34. Bekiaris, P. S. & Klamt, S. Automatic construction of metabolic models with enzyme
508 constraints. *BMC Bioinformatics* **21**, 19 (2020).
509 35. Mao, Z. *et al.* ECMpy 2.0: A Python package for automated construction and analysis of
510 enzyme-constrained models. *Synth. Syst. Biotechnol.* **9**, 494–502 (2024).
511 36. Orth, J. D., Thiele, I. & Palsson, B. Ø. O. What is flux balance analysis? *Nat Biotech* **28**, 245–
512 248 (2010).
513 37. Gu, C., Kim, G. B., Kim, W. J., Kim, H. U. & Lee, S. Y. Current status and applications of genome-
514 scale metabolic models. *Genome Biol.* **20**, 121 (2019).
515 38. Akashi, H. & Gojobori, T. Metabolic efficiency and amino acid composition in the proteomes
516 of Escherichia coli and Bacillus subtilis. *Proc. Natl. Acad. Sci.* **99**, 3695–3700 (2002).

517 39. Mori, M., Cheng, C., Taylor, B. R., Okano, H. & Hwa, T. Functional decomposition of
518 metabolism allows a system-level quantification of fluxes and protein allocation towards
519 specific metabolic functions. *Nat. Commun.* **14**, 4161 (2023).
520 40. Pacheco, A. R., Moel, M. & Segrè, D. Costless metabolic secretions as drivers of interspecies
521 interactions in microbial ecosystems. *Nat. Commun.* **10**, 103 (2019).
522 41. Georgi, T., Rittmann, D. & Wendisch, V. F. Lysine and glutamate production by
523 Corynebacterium glutamicum on glucose, fructose and sucrose: Roles of malic enzyme and
524 fructose-1,6-bisphosphatase. *Metab. Eng.* **7**, 291–301 (2005).
525 42. Vila, J. C. C. *et al.* Metabolic similarity and the predictability of microbial community
526 assembly. 2023.10.25.564019 Preprint at https://doi.org/10.1101/2023.10.25.564019
527 (2023).
528 43. Wendering, P. & Nikoloski, Z. COMMIT: Consideration of metabolite leakage and community
529 composition improves microbial community reconstructions. *PLOS Comput. Biol.* **18**,
530 e1009906 (2022).
531 44. Bennett, B. D., Yuan, J., Kimball, E. H. & Rabinowitz, J. D. Absolute quantitation of
532 intracellular metabolite concentrations by an isotope ratio-based approach. *Nat. Protoc.* **3**,
533 1299–1311 (2008).
534 45. Park, J. O. *et al.* Metabolite concentrations, fluxes and free energies imply efficient enzyme
535 usage. *Nat. Chem. Biol.* **12**, 482–489 (2016).
536 46. Sajed, T. *et al.* ECMDB 2.0: A richer resource for understanding the biochemistry of E. coli.
537 *Nucleic Acids Res.* **44**, D495–D501 (2016).

538 47. Thorfinnsdottir, L. B., García-Calvo, L., Bø, G. H., Bruheim, P. & Røst, L. M. Optimized Fast
539 Filtration-Based Sampling and Extraction Enables Precise and Absolute Quantification of the
540 Escherichia coli Central Carbon Metabolome. *Metabolites* **13**, 150 (2023).
541 48. Falcón-Cano, G., Molina, C. & Cabrera-Pérez, M. Á. Reliable Prediction of Caco-2 Permeability
542 by Supervised Recursive Machine Learning Approaches. *Pharmaceutics* **14**, 1998 (2022).
543 49. Liu, X., Testa, B. & Fahr, A. Lipophilicity and Its Relationship with Passive Drug Permeation.
544 *Pharm. Res.* **28**, 962–977 (2011).
545 50. Pontrelli, S. *et al.* Degradation of extracellular polymeric substances shapes microbial
546 community diversity. *PLOS Biol.* **23**, e3003287 (2025).
547 51. Gibson, S. E. R. *et al.* Bacteria encode post-mortem protein catabolism that enables altruistic
548 nutrient recycling. *Nat. Commun.* **16**, 1400 (2025).
549 52. Fuhrer, T., Zampieri, M., Sévin, D. C., Sauer, U. & Zamboni, N. Genomewide landscape of
550 gene–metabolome associations in Escherichia coli. *Mol. Syst. Biol.* **13**, 907 (2017).
551 53. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants:
552 the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
553 54. Fuhrer, T., Heer, D., Begemann, B. & Zamboni, N. High-throughput, accurate mass
554 metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry.
555 *Anal. Chem.* **83**, 7074–7080 (2011).
556 55. Rancourt, D. E., Stephenson, J. T., Vickell, G. A. & Wood, J. M. Proline excretion by Escherichia
557 coli K12. *Biotechnol. Bioeng.* **26**, 74–80 (1984).
558 56. Karp, P. D. *et al.* The EcoCyc database (2025). *EcoSal Plus* **0**, eesp-0019-2024 (2025).

559 57. Park, J. H., Lee, K. H., Kim, T. Y. & Lee, S. Y. Metabolic engineering of Escherichia coli for the
560 production of l-valine based on transcriptome analysis and in silico gene knockout
561 simulation. *Proc. Natl. Acad. Sci.* **104**, 7797–7802 (2007).
562 58. Zakataeva, N. P., Aleshin, V. V., Tokmakova, I. L., Troshin, P. V. & Livshits, V. A. The novel
563 transmembrane Escherichia coli proteins involved in the amino acid efflux. *FEBS Lett.* **452**,
564 228–232 (1999).
565 59. Burton, K. Adenine transport in Escherichia coli. *Proc. R. Soc. Lond. B Biol. Sci.* **255**, 153–157
566 (1997).
567 60. Papakostas, K., Botou, M. & Frillingos, S. Functional Identification of the
568 Hypoxanthine/Guanine Transporters YjcD and YgfQ and the Adenine Transporters PurP and
569 YicO of Escherichia coli K-12. *J. Biol. Chem.* **288**, 36827–36840 (2013).
570 61. Rodionova, I. A. *et al.* Identification of a transcription factor, PunR, that regulates the purine
571 and purine nucleoside transporter punC in E. coli. *Commun. Biol.* **4**, 991 (2021).
572 62. Mori, M. *et al.* From coarse to fine: the absolute Escherichia coli proteome under diverse
573 growth conditions. *Mol. Syst. Biol.* **17**, e9536 (2021).
574 63. Smith, D. R. & Chapman, M. R. Economical Evolution: Microbes Reduce the Synthetic Cost
575 of Extracellular Proteins. *mBio* **1**, e00131-10 (2010).
576 64. Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W. & Milo, R. Glycolytic strategy as a
577 tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci.* **110**, 10039–10044
578 (2013).
579 65. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol.*
580 *Biol.* **3**, 318–356 (1961).

581 66. Giuliani, S. E. *et al.* Environment sensing and response mediated by ABC transporters. *BMC*
582 *Genomics* **12**, S8 (2011).
583 67. Martínez-Antonio, A., Janga, S. C., Salgado, H. & Collado-Vides, J. Internal-sensing machinery
584 directs the activity of the regulatory network in Escherichia coli. *Trends Microbiol.* **14**, 22–
585 27 (2006).
586 68. Nies, D. H., Schleuder, G., Galea, D. & Herzberg, M. A flow equilibrium of zinc in cells of
587 Cupriavidus metallidurans. *J. Bacteriol.* **206**, e00080-24 (2024).
588 69. Fritts, R. K., McCully, A. L. & McKinlay, J. B. Extracellular Metabolism Sets the Table for
589 Microbial Cross-Feeding. *Microbiol. Mol. Biol. Rev.* **85**, 10.1128/mmbr.00135-20 (2021).
590 70. Lopez, J. A., Liu, B., Li, Z., Donia, M. S. & Wingreen, N. S. When is microbial cross-feeding
591 evolutionarily stable? 2025.05.16.654511 Preprint at
592 https://doi.org/10.1101/2025.05.16.654511 (2025).
593 71. Preussger, D., Giri, S., Muhsal, L. K., Oña, L. & Kost, C. Reciprocal Fitness Feedbacks Promote
594 the Evolution of Mutualistic Cooperation. *Curr. Biol.* **30**, 3580-3590.e7 (2020).
595 72. Zomorrodi, A. R. & Segrè, D. Genome-driven evolutionary game theory helps understand the
596 rise of metabolic interdependencies in microbial communities. *Nat. Commun.* **8**, 1563
597 (2017).
598 73. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia
599 coli K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**, 6640–6645 (2000).
600 74. Lewis, N. E. *et al.* Omic data from evolved E. coli are consistent with computed optimal
601 growth from genome-scale models. *Mol. Syst. Biol.* **6**, 390 (2010).

602 75. Campello, R. J. G. B., Moulavi, D. & Sander, J. Density-Based Clustering Based on Hierarchical
603 Density Estimates. in *Advances in Knowledge Discovery and Data Mining* (eds Pei, J., Tseng,
604 V. S., Cao, L., Motoda, H. & Xu, G.) 160–172 (Springer Berlin Heidelberg, Berlin, Heidelberg,
605 2013).
606 76. Gallart-Ayala, H. *et al.* A global HILIC-MS approach to measure polar human cerebrospinal
607 fluid metabolome: Exploring gender-associated variation in a cohort of elderly cognitively
608 healthy subjects. *Anal. Chim. Acta* **1037**, 327–337 (2018).
609 77. van der Velpen, V. *et al.* Systemic and central nervous system metabolic alterations in
610 Alzheimer's disease. *Alzheimers Res. Ther.* **11**, 93 (2019).
611 78. Myers, J. A., Curtis, B. S. & Curtis, W. R. Improving accuracy of cell and chromophore
612 concentration measurements using optical density. *BMC Biophys.* **6**, 4 (2013).
613 79. Ren, Q., Henes, B., Fairhead, M. & Thöny-Meyer, L. High level production of tyrosinase in
614 recombinant Escherichia coli. *BMC Biotechnol.* **13**, 18 (2013).
615 80. Goelzer, A. *et al.* Quantitative prediction of genome-wide resource allocation in bacteria.
616 *Metab. Eng.* **32**, 232–243 (2015).
617 81. Zhang, Y. *et al.* A new genome-scale metabolic model of Corynebacterium glutamicum and
618 its application. *Biotechnol. Biofuels* **10**, 169 (2017).
619 82. McGill, S. L. *et al.* Pseudomonas aeruginosa reverse diauxie is a multidimensional, optimized,
620 resource utilization strategy. *Sci. Rep.* **11**, 1457 (2021).
621 83. Kuepper, J. *et al.* Metabolic Engineering of Pseudomonas putida KT2440 to Produce
622 Anthranilate from Glucose. *Front. Microbiol.* **6**, 1310 (2015).

623 84. Niu, J. *et al.* Construction and Analysis of an Enzyme-Constrained Metabolic Model of
624 Corynebacterium glutamicum. *Biomolecules* **12**, 1499 (2022).
625 85. Wu, K. *et al.* ecBSU1: A Genome-Scale Enzyme-Constrained Model of Bacillus subtilis Based
626 on the ECMpy Workflow. *Microorganisms* **11**, 178 (2023).
627 86. Guo, J., Zhang, H., Wang, C., Chang, J.-W. & Chen, L.-L. Construction and analysis of a
628 genome-scale metabolic network for *Bacillus licheniformis* WX-02. *Res. Microbiol.* **167**, 282–
629 289 (2016).
630 87. Nogales, J. *et al.* High-quality genome-scale metabolic modelling of Pseudomonas putida
631 highlights its broad metabolic capabilities. *Environ. Microbiol.* **22**, 255–269 (2020).
632 88. Ebrahim, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRApy: COnstraints-Based
633 Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).
634 89. Varma, A., Boesch, B. W. & Palsson, B. O. Stoichiometric interpretation of Escherichia coli
635 glucose catabolism under various oxygenation rates. *Appl. Environ. Microbiol.* **59**, 2465–
636 2473 (1993).
637 90. Reznik, E., Mehta, P. & Segrè, D. Flux Imbalance Analysis and the Sensitivity of Cellular
638 Growth to Changes in Metabolite Pools. *PLOS Comput. Biol.* **9**, e1003195 (2013).
639 91. Ram, Y. *et al.* Predicting microbial growth in a mixed culture from growth curve data. *Proc.*
640 *Natl. Acad. Sci.* **116**, 14698–14707 (2019).
641 92. Reed, J. L., Vo, T. D., Schilling, C. H. & Palsson, B. O. An expanded genome-scale model of
642 Escherichia coli K-12 (iJR904 GSM/GPR). *Genome Biol.* **4**, R54 (2003).
643 93. Feist, A. M. *et al.* A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655
644 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* **3**, 121 (2007).

645 94. Orth, J. D. *et al.* A comprehensive genome-scale reconstruction of Escherichia coli
646 metabolism—2011. *Mol. Syst. Biol.* **7**, 535 (2011).
647 95. Monk, J. M. *et al.* iML1515, a knowledgebase that computes Escherichia coli traits. *Nat.*
648 *Biotechnol.* **35**, 904–908 (2017).
649 96. Djoumbou Feunang, Y. *et al.* ClassyFire: automated chemical classification with a
650 comprehensive, computable taxonomy. *J. Cheminformatics* **8**, 61 (2016).
651 97. Seabold, S. & Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. in
652 92–96 (Austin, Texas, 2010). doi:10.25080/Majora-92bf1922-011.
653 98. Lukačišinová, M., Fernando, B. & Bollenbach, T. Highly parallel lab evolution reveals that
654 epistasis can curb the evolution of antibiotic resistance. *Nat. Commun.* **11**, 3105 (2020).
655 99. Chen, S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication
656 using fastp. *iMeta* **2**, e107 (2023).
657 100. Simon Andrews. FastQC: A Quality Control Tool for High Throughput Sequence Data.
658 Babraham Institute (2010).
659 101. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–
660 3100 (2018).
661 102. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
662 103. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide
663 polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-
664 2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
665 104. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing.
666 Preprint at https://doi.org/10.48550/arXiv.1207.3907 (2012).

105. Burz, S. D. *et al.* A Guide for Ex Vivo Handling and Storage of Stool Samples Intended for Fecal Microbiota Transplantation. *Sci. Rep.* **9**, 8897 (2019).

106. Bennett, B. D. *et al.* Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* **5**, 593–599 (2009).

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

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

109. Simensen, V. *et al.* Experimental determination of Escherichia coli biomass composition for constraint-based metabolic modeling. *PLOS ONE* **17**, e0262450 (2022).

110. Saier, M. H. *et al.* The Transporter Classification Database (TCDB): 2021 update. *Nucleic Acids Res.* **49**, D461–D467 (2020).

111. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

112. Pignon, E., Holló, G., Steiner, T., van Vliet, S. & Schaerli, Y. Uptake and leakage rates differentially shape community arrangement and composition of microbial consortia. *ISME J.* **19**, wraf122 (2025).

113. Pignon, E. Engineering metabolic interactions to control spatial patterns and composition in auxotrophic microbial communities. https://iris.unil.ch/handle/iris/264404 (2025).

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, Milton Saier and Uwe Sauer for providing feedback, 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, Estelle Pignon for providing *E. coli* auxotrophs 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, PP)

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

718

719 ### Competing interests

720 Authors declare that they have no competing interests.

721

722 ### Data and materials availability

723 Strains and isolates used in this study are available on request. Formatted data and code for data
724 analyses and visualization are available on GitHub at https://github.com/Mitri-lab/metabolite-
725 release. A permanent archive of this repository will be deposited to Zenodo upon publication.
726 The raw sequencing data and metadata have been deposited to the NCBI SRA database under the
727 BioProject ID PRJNA1270783. The raw FIA-TOF MS data is available on MassIVE under the dataset
728 identifier MSV000097105.

729 Extended Data and Supplementary Information

730 Supplementary Figures S1-S32

731 Supplementary Tables S1-S7

732

733 Methods

734 ### Strains

735 All experiments were conducted with either *E. coli* K-12 MG1655, *E. coli* K-12 BW25113 (wild-type
736 ancestor for KEIO KO strains) or one of the 75 KO mutants selected from the KEIO collection[53]
737 (Table S6). The KEIO KO strains were verified by PCR with one primer binding inside the Kan^R
738 cassette inserted during gene knock-out (Keio_K1_Rev for genes on the (-) strand;
739 Keio_KT_RvComp_Fw for genes on the (+) strand) and one gene-specific primer binding
740 downstream of the knocked-out gene (primers are listed in Table S6)[53,73]. For all experiments, the
741 KEIO KO strains were used with the kanamycin resistance cassette in place.

742

743 ## Precultures and growth media

744 For all experiments detailed below, the strains were precultured as described here unless
745 otherwise stated: From glycerol stocks stored at -70 °C each strain was streaked out onto LB agar
746 plates and incubated overnight (37 °C). Then, one colony of each strain was picked and used to
747 inoculate liquid precultures (10 mL LB in 50 mL Erlenmeyer flasks) that were incubated overnight
748 (37 °C, 200 rpm). The kanamycin-resistant KEIO KO strains and strains evolved from these were
749 always precultured in liquid LB and streaked out onto LB agar plates with 25 µg/mL kanamycin,
750 while *E. coli* K-12 MG1655 and *E. coli* K12 BW25113 were always precultured in LB and streaked
751 onto LB agar plates without selection. Unless otherwise noted, precultures were washed three
752 times by centrifugation, with the supernatant removed after each step, and the cells resuspended
753 in M9 medium with no carbon source to eliminate residual LB before inoculation. Unless
754 otherwise stated, M9 medium was always prepared by mixing (per litre): 200 mL 5X M9 salts
755 (M6030, Sigma-Aldrich, St. Louis, MO, USA), 2 mL 1M $MgSO_4$, 0.1 mL 1M $CaCl_2$, 1 mL 0.5 g/L
756 $FeSO_4 \cdot 7H_2O$, MQ water and appropriate amounts of a carbon source stock solution. The final M9
757 medium was then pH-adjusted to pH = 7.0, filter sterilized (0.22 µm) and stored at 4°C until use.

758

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

760 To complement the existing exometabolome data from batch cultures of *E. coli* on glucose[12], we
761 aimed to select three carbon sources that differed in chemical class and caused variation in
762 metabolic flux patterns. We used the ecGEM eciJO1366 of *E. coli* K-12 MG1655 without any
763 modifications and parsimonious FBA (pFBA) to simulate optimal flux patterns across 104 different
764 growth supporting metabolites[34,74]. We then performed a PCA analysis of the predicted fluxes (as
765 Boolean values) to compare flux patterns between the carbon sources (Fig. S5). We used
766 HDBSCAN[75] to identify clusters and selected from the different clusters in total 10 compounds
767 that also differed in chemical class (sugars, amino acids and organic acids) and their entry point
768 into central carbon metabolism (Fig. S5). We then screened the growth of the selected carbon
769 sources in triplicates over 48 hours in 96-well plates (Fig. S5). For this experiment, we prepared
770 M9 medium with each of the 10 different carbon sources to a concentration scaled to 120 mM
771 carbon atoms. From washed LB precultures, *E. coli* K-12 MG16155 was inoculated to OD_{600} = 0.05

772 in 200 µL of media and cultivated with continuous shaking (double orbital, 425 cpm) in a BioTek
773 Synergy H1 (Agilent Technologies, Winooski, VT, USA) plate reader at 37 °C for 48 hours. We
774 ultimately chose galactose, L-alanine and L-malate, as they supported rapid growth, differed in
775 chemical class, entered central metabolism at different locations and were grouped into different
776 clusters based on flux patterns (Fig. S5).
777

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

779 *E. coli* K-12 MG1655 was cultivated in lab-scale bioreactors using three select carbon sources.
780 Their concentrations were normalized by the number of carbon atoms to a carbon atom
781 concentration of 120 mM, corresponding to 20 mM galactose (G0625, Sigma-Aldrich), 30 mM L-
782 malic acid (02290, Sigma-Aldrich) and 40 mM L-alanine (A7627, Sigma-Aldrich). M9 medium with
783 these three carbon sources were prepared as described above using 25X carbon source stock
784 solutions.
785

786 The strain was streaked from a frozen glycerol vial onto an LB agar plate and incubated at 37 °C
787 overnight. One colony was transferred to a 250 mL baffled shake flask containing 40 mL LB
788 medium. After 7 hours, a culture volume corresponding to a start OD_{600} = 0.05 was inoculated to
789 a 500 mL baffled shake flask containing 100 mL M9 + 40 mM L-alanine. This flask was incubated
790 for 36 h (200 rpm, 37 °C). For the two other carbon sources, the strain was incubated in LB for 17
791 h before inoculating a culture volume corresponding to OD_{600} = 0.05 in a 500 mL baffled shake
792 flask containing 100 mL M9 + 20 mM L-galactose or M9 + 30 mM L-malic acid, respectively. The
793 flasks were incubated for 24 h (200 rpm, 37 °C). Before inoculation to bioreactors, M9 precultures
794 were centrifuged (3220 rcf, 10 min, 4 °C), supernatant removed, and pellet resuspended in M9
795 medium without carbon.
796

797 Batch cultivations were performed in 1 L DASGIP bioreactors (Eppendorf DASGIP, Jülich, Germany)
798 with an initial volume of 600 mL M9 medium supplemented with either 20 mM galactose, 30 mM
799 L-malic acid or 40 mM L-alanine, three replicates per condition. The reactors were equipped with
800 two Rushton impellers and probes for measurements of dissolved oxygen (DO) and pH.

801 Submerged aeration was maintained at 0.5 vvm, and agitation was cascaded between 400 and
802 800, controlled by a DO set point of 30 %. pH set point was 7.0 and controlled using 2 M NaOH
803 and 2 M HCl. Start OD_{600} was 0.05.
804
805 Sampling was performed regularly during the exponential growth phase with additional samples
806 collected during stationary phase. OD_{600} was measured at all time points and supernatant stored:
807 2 mL culture was centrifuged (3220 rcf, 10 min, 4 °C), and supernatant frozen in -20 °C. At the end
808 of exponential phase and at the end of the experiment, samples were collected for cell dry weight
809 (CDW). 50 mL culture was centrifuged (3220 rcf, 10 min, 4 °C), supernatant removed, cell pellet
810 resuspended in 50 mL MQ water, centrifuged, supernatant removed, cell pellet resuspended in
811 5-10 mL MQ water and transferred to a pre-weighed aluminium beaker and dried at 105 °C for 24
812 hours. Beakers were then weighed again and CDW calculated.
813

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

816 Four samples from each of the nine batch cultures (three replicates of three different media) and
817 media samples were analysed with LC-MS and GC-MS to measure the extracellular concentration
818 of metabolites at different time points throughout the experiment. The time points were selected
819 based on time-series cultivation data to cover the exponential growth phase (three samples) and
820 one stationary phase sample (Fig. S6). Samples from the different replicates were aligned based
821 on the timing of the end of the exponential growth phase (Fig. S6).
822
823 LC-MS of selected bioreactor samples was conducted by the Metabolomics facility at UNIL using
824 their method "high-coverage targeted analysis of polar metabolites", following this procedure:
825 Media samples were thawed on ice. An aliquot (20 µL) was extracted with the addition of ice-cold
826 methanol (80 µL), vortexed (30 sec) and centrifuged (15000 rpm, 15 min, 4°C). The supernatants
827 were transferred to vials for liquid chromatography – tandem mass spectrometry (LC-MS/MS)
828 analyses. The extracts were analyzed by hydrophilic interaction chromatography coupled to
829 tandem mass spectrometer (6496 iFunnel, Agilent Technologies) using multiple reaction

830 monitoring - MRM approach in both, positive and negative ionization modes, to maximize the
831 polar metabolome coverage. The analytical conditions have been described in detail
832 elsewhere[76,77]. Raw LC-MS/MS metabolome data were processed using the Mass Hunter
833 Quantitative analysis software (Agilent Technologies). The peak areas (or extracted ion
834 chromatograms (EICs) for the monitored MRM transitions) were used for relative comparison of
835 metabolite levels between different conditions or groups of samples. Data quality assessment,
836 including signal drift correction, was performed using pooled quality control (QC) samples
837 analyzed periodically throughout the entire batch.
838
839 Metabolite quantification was performed using stable isotope labeled internal standards and
840 calibration curves following the signal drift correction. Data processing was done using
841 MassHunter Quantitative analysis. Peak area integration was manually curated, and
842 concentrations were reported by selecting a 6-point portion of the linear standard curves relevant
843 for the observed concentrations.
844
845 The relative LC-MS data was standardized before subsequent analyses and data processing. Both
846 the relative and absolute LC-MS data were cleaned of obvious outliers. For the absolute data the
847 outliers were caused by carry-over from previous samples leading to too high malate (six samples)
848 and succinate (one sample) values. For the relative LC-MS data there were additional outliers of
849 unknown reasons (39 values in total) related to in total 14 metabolites. For several metabolites,
850 there were more missing values (not detected/not quantified) in the absolute data than in the
851 relative data. For a few metabolites (phenylalanine, proline, creatine), we leveraged good linear
852 relationships between relative and absolute data (R^2>0.8) to estimate the missing absolute values
853 using a simple linear transformation.
854
855 GC-MS analysis of the same samples was conducted in the following way: Analytical standards
856 were prepared as M9 medium with increasing concentrations of acetate, lactate, pyruvate,
857 formate and propionate and stored at -70°C until use. Samples from the *E. coli* bioreactor batch
858 cultures and analytical standards were thawed on iced. Then, 100 µL was transferred to an

859 Eppendorf tube and kept on a cold block (-20 °C) while adding 5 µL 11% HCl and 500 µL diethyl
860 ether. Tubes were capped, vortexed on a thermoblock (2000 rpm, 10 min, 1°C) and centrifuged
861 (13 000 rcf, 5 min, 4°C). The top organic phase layer was pipetted to a glass vial and derivatized
862 with 20 µL N-(t-butyldimethylsilyl)-N-methyltrifluoroacetamide (MTBSTFA) and vortexed briefly.
863 Samples were then placed on a heat block (90 min, 35°C) and kept at 16°C until analysis. The
864 samples were injected (1 µL) by a Pal3 autosampler onto an Agilent 8890-5977B GC-MSD (Agilent
865 Technologies) with a VF-5MS (30 m x 0.25 mm x 0.25 mm) column. The samples were injected
866 with a split ratio of 15:1, helium flow rate of 1 mL/min and inlet temperature of 230 °C. The
867 temperature was held for 2 min at 50 °C, raised at 25 °C/min to 175 °C, 30 °C/min to 280 °C and
868 held for 3.5 min. The MSD was run in scan mode from 40-500 Da. Analyte abundances were
869 calculated using the MassHunter Quantitative Analysis software (Agilent Technologies). Absolute
870 concentrations were calculated using standard curves.
871
872 Galactose concentrations were measured using the L-Arabinose/D-Galactose assay kit (K-ARGA,
873 Megazyme, Bray, Ireland), following their rapid protocol for analyses in 96-well plates.
874

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

877 Time-series data on biomass (in OD_{600}) and absolute extracellular metabolite concentrations from
878 batch cultivations of *E. coli, B. licheniformis, C. glutamicum* and *S. cerevisiae* in high glucose
879 concentrations (10-20 g/L)[12] was kindly provided by the authors (mean and standard deviations).
880 For metabolites reported as a sum of two metabolites, we attributed equal amounts to those
881 metabolites to enable direct comparison with metabolite values estimated for individual
882 metabolites (described below). This includes 2-phosphoglycerate/3-phosphoglycerate and
883 ribulose 5-phosphate/xylulose 5-phosphate. To estimate the metabolite release and uptake rates
884 of each species, we first converted the OD_{600} values to gDW/L. For *E. coli* we used the conversion
885 factor we estimated from our own cell dry weight measurements in M9 galactose medium (0.346
886 gDW/L/OD_{600}), which is comparable, but on the lower end of values found in the literature (0.36-
887 0.515 gDW/L/OD_{600})[78,79]. For the three other species, we used literature values: B. licheniformis

(data from B. subtilis) 0.48 gDW/L/OD_{600}[80]; *C. glutamicum* 0.27 gDW/L/OD_{600}[81]; *S. cerevisiae* (mean of literature values) 0.746 gDW/L/OD_{600}[78]. 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, the concentration $C_i(t)$ can be 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

915 the biomass area under the curve used to estimate metabolite release and uptake rates further
916 described below.
917
918 We estimated the metabolite uptake or release rate by performing for each metabolite and each
919 carbon source a linear regression of measured concentration (relative or absolute) vs biomass
920 AUC using the three samples (timepoints T1-T3) acquired from the three replicate batch cultures
921 during exponential phase plus the initial M9 medium (T0; Figs. S7-S9). To avoid underestimating
922 release rates for metabolites that potentially were being re-consumed at the end of the
923 exponential phase, we discarded the last timepoint from the exponential phase in cases where
924 the rate was positive (indicating metabolite release) but there was a significant reduction in
925 extracellular concentration from T2-T3 ($P < 0.05$, one-sided t-test). For 17 (of 126) metabolites,
926 we chose to not include T0 for at least one of the three conditions in the linear regression, either
927 because of very large variability of T0 or because the trend from T0 to T1 was clearly different
928 from the general trend from T1-T3. We only estimated the rate for metabolites where we had at
929 least three data points from at least two different time points. For the metabolites where we had
930 both relative and absolute data, we used the absolute data if there were enough absolute data
931 points to get adequate slope estimates. For seven metabolites (glutamine, alpha-aminoadiapate,
932 serine, lysine, lactate, NAD and isocitrate) where both absolute and relative quantification was
933 performed, we could estimate the rate more accurately on the relative data set (because of more
934 missing data points in the absolute data set). However, there were sufficient absolute
935 measurements to estimate the absolute spread (standard deviation) of these data points,
936 enabling translation of the slope value from Z-score to µmol/gDW/h.
937

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

940 The dataset from Vila et al. (2023) was kindly provided by the authors[42]. The dataset contains
941 exometabolome data from three timepoints (16, 28, and 48 hours) of *E. coli* MG1655, *P. putida*
942 KT2440, and two environmental isolates of the genera *Enterobacter* and *Pseudomonas,*
943 respectively[42]. All species were cultivated in M9 medium with one of five carbon sources (D-

944 glucose, D-fructose, glycerol, pyruvate or L-malate). Additionally, the dataset includes
945 exometabolome data from two timepoints (28 and 48 hours) of *E. coli* and the *Enterobacter* sp.
946 on D-ribose, L-arabinose and D-galactose, and of *P. putida* and the *Pseudomonas* sp. on acetate,
947 fumarate and succinate. To estimate metabolite release rates, we first used the plate reader
948 growth curves to calculate biomass AUC after 16, 28 and 48 hours and to identify the end of
949 exponential growth of each organism in each condition. The end of exponential phase was
950 automatically identified as the first peak in the second derivative of the smoothed growth curve
951 with an OD_{600} above 60% of maximum. We converted the AUC from OD_{600} to gDW/L using the
952 same conversion factor for all the different carbon source environments: 0.36 gDW/L/ OD_{600} for
953 *E. coli* and the *Enterobacter* sp., mean of measured values and literature values[78,79]; 0.476 gDW/L/
954 OD_{600} for *P. putida* and the *Pseudomonas* sp, mean of literature values[82,83]. Finally, we performed
955 a linear regression between extracellular metabolite concentrations and AUC on both duplicates
956 of each strain/carbon source combination including all datapoints before the end of exponential
957 growth + 2 hours (a buffer to account for some inaccuracy in the automatic detection). We only
958 included rate estimates for metabolites that were quantified at least twice for each strain. As the
959 dataset did not include an early timepoint or media measurement, we assumed the initial
960 concentration of each metabolite to be 0. If a metabolite was detected more at more than one
961 timepoint, we discarded the estimated rate if there was a qualitative discrepancy in slope value
962 when the slope was estimated with or without the T0 concentration assumed to be 0. If the
963 carbon source was among the measured metabolites, this metabolite was excluded from rate
964 estimation in those specific conditions.
965

966 ## Genome-scale metabolic models

967 For *E. coli, S. cerevisiae and C. glutamicum* we used the previously developed enzyme-constrained
968 genome-scaled metabolic models (GEMs) eciJO1366[34], ecCGL1[84] and ecYeastGEM 8.3.4[33],
969 respectively. For *B. licheniformis* there was no available enzyme-constrained GEM, and we
970 therefore opted to use the enzyme-constrained GEM ecBSU1 of the closely related species *B.*
971 *subtilis*[85], which was available at github.com/tibbdc/ecBSU1. A normal GEM for *B. licheniformis* is
972 available[86], but with this model we could not achieve any feasible flux balance solution. To

973 simulate the environmental isolate of the genus *Enterobacter* we used the *E. coli* GEM eciJO1366.
974 For the environmental isolate of the genus *Pseudomonas* and for *P. putida* we used the enzyme-
975 constrained version of the *P. putida* GEM iJN1463[87] automatically generated by the ECMpy2
976 method[35], which we here refer to as eciJN1463.
977
978 We made a few minor changes to the published GEMs before predicting metabolite values. In the
979 *E. coli* GEM eciJO1366 the proteome weight of the forward reaction *GALKr* (galactokinase) was
980 extremely high (0.016), limiting the growth rate on galactose to much less than the growth rate
981 we measured in M9 galactose medium. To enable eciJO1366 to reach the measured growth rate
982 with galactose as the only carbon source, we replaced the proteome weight of the forward *GALKr*
983 reaction with the proteome weight of the reverse *GALKr* reaction which was much smaller
984 (0.00037). For the ecYeastGEM we only merged exchange reactions that were previously split to
985 simplify its use. Both ecBSU1 and eciJN1463 were translated from json to sMOMENT format in
986 xml[34], modified by merging split exchange reactions and fixing metabolite compartment names
987 and annotations. For ecBSU1 we also removed the enzyme constraint on the reaction
988 "PPAm_num1" (inorganic diphosphatase) as it caused unrealistic uptake and release rates of
989 phosphate and diphosphate, respectively. The *C. glutamicum* GEM ecCGL1 was also converted
990 from the ECMpy format[35] to the sMOMENT format[34] but otherwise kept unchanged. All GEMs
991 were used with their default biomass equations as objective when maximizing growth in all FBA
992 or pFBA analyses. Reframed 1.5.3 (github.com/cdanielmachado/reframed) or COBRApy 0.29.1[88]
993 were used to load, manipulate and run constraint-based analyses with Gurobi 10 (Gurobi
994 Optimization, LLC) as the solver.
995

996 ## Estimating metabolite value and turnover using genome-scale metabolic models

997 In constraint-based modelling a shadow price is an estimate of how sensitive the objective
998 function in a flux balance analysis (FBA) is to increased influx or efflux of a metabolite[89,90]. Hence,
999 one can interpret the shadow price as *metabolite value* as it quantifies the impact on growth
1000 caused by the release of a metabolite (to have positive values we here define metabolite value
1001 as the shadow price multiplied by -1). A similar approach to measure the cost of metabolite

1002 release has been used to identify costless secretions[40]. The metabolite value can be predicted for
1003 a specific organism in a specific context by using a GEM of the corresponding organism (or closely
1004 related) and then constraining this GEM to the specific context (i.e. primarily defining metabolite
1005 uptake rates). While shadow prices are usually estimated directly by the solver used to run FBA[36],
1006 the robustness of these values is often questioned because of numerical issues. To get more
1007 robust shadow price estimates we solved for each metabolite in each context a separate FBA,
1008 where we constrained the GEM to have an (additional) release of 0.01 mmol/gDW/h of that
1009 metabolite and quantified the change in the objective function, i.e. the change Δy in growth rate
1010 caused by the forced metabolite release. The metabolite value was then quantified as $\Delta y/0.01$. If
1011 the metabolite was predicted to be released by FBA, we added 0.01 mmol/gDW/h to the FBA-
1012 predicted value to estimate the rate. This was only relevant for acetate for *E. coli* in the L-malate
1013 condition, for ethanol, acetate, formate and pyruvate for *S. cerevisiae*, and for pyruvate for *P.*
1014 *putida* in the L-malate condition.
1015
1016 Using this approach, we estimated metabolite values for *E. coli* in batch cultures with glucose,
1017 galactose, L-malate or L-alanine as the carbon source by constraining the uptake of the
1018 corresponding exchange reaction in eciJO1366 to the uptake rate of the carbon source as
1019 estimated from the experimental data. The same approach was used for *C. glutamicum, B.*
1020 *licheniformis* and *S. cerevisiae* using the corresponding models. For *E. coli*, *P. putida* and the two
1021 environmental isolates of the genera *Enterobacter* and *Pseudomonas* we first extracted maximum
1022 growth rates for each organism in each environment from growth curves (using curveball[91]).
1023 Then, we used the respective GEMs to calculate the uptake rates of the corresponding carbon
1024 sources required to sustain the maximum growth rates. These uptake rates were then used to
1025 constrain the GEMs when estimating the metabolite values as described above. The turnover (i.e.
1026 the sum of positive fluxes producing an intracellular metabolite) was estimated by running
1027 parsimonious FBA[74], using the same GEMs constrained in the same way as described for above
1028 for metabolite values.
1029

1030 To validate that the negative correlation between log-transformed metabolite values and release
1031 rates, we compared estimates from eciJO1366 with estimates from the 4 benchmark *E. coli* K-12
1032 MG1655 GEMs iJR904[92], iAF1260[93], iJO1366[94] and iML1515[95] that are of increasing complexity.
1033 We then compared estimated metabolite values and estimated release rates using data from *E.*
1034 *coli* in bioreactors with glucose, galactose, L-malate or L-alanine as the sole carbon source, the
1035 same data used in Fig. 2A (Fig. S11). In the L-malate condition for the iJR904 model, we discarded
1036 the metabolite value for formate as an outlier because it was extremely small ($<10^{-12}$), below the
1037 solver tolerance. Note that no metabolite values were negative and hence no other values were
1038 discarded upon log-transformation.
1039

1040 ## Metabolite classification, chemical properties and intracellular concentrations

1041 The charge and molecular weight of metabolites were obtained from eciJO1366[34]. All other
1042 chemical properties and InChIKeys were obtained from PubChem using the python interface
1043 PubChemPy (github.com/mcs07/PubChemPy). The InChIKeys were then used to classify
1044 metabolites into a chemical taxonomy using ClassyFire[96]. To get an appropriate level of detail of
1045 metabolic classification, we used the "Class" level category. However, we separated the class
1046 "Carboxylic acids and derivatives" into "Amino acids" and "Carboxylic acids" based on the
1047 "Subclass" level. We also merged the classes "Hydroxy acids and derivatives" and "Keto acids and
1048 derivatives" into the joint class "Keto / hydroxy acids". Finally, metabolites that did not fall into
1049 either of the abovementioned categories nor the class "Organooxygen compounds" were joined
1050 into the category "Other".
1051

1052 ## Linear models for evaluating the importance of different factors for predicting metabolite
1053 release rates

1054 Linear models were created and analysed using statsmodels[97]. Compound class and carbon
1055 source was incorporated as categorical variables. Metabolites predicted to have zero turnover
1056 was given a value of 10^{-4}, less than the minimum predicted turnover. Release rates where
1057 corresponding values were missing in any of the compared factors were discarded before model
1058 fitting. To evaluate the statistical significance of the out-of-sample predictions, we performed 10^4

1059 permutations of model fitting and out-of-sample estimations where the labels (either metabolite
1060 or condition) was shuffled on each permutation. We then calculated the fraction of out-of-sample
1061 R^2 values from these permutations below the observed R^2 value to estimate the *P* value.
1062

1063 ## Cultivation and exometabolome analyses of KEIO knockout strains

1064 One objective was to study how disruption of key metabolic fluxes would affect extracellular
1065 metabolite concentrations and how this would change if these strains were allowed to evolve in
1066 a simple nutrient environment. The KEIO collection is a collection of non-essential *E. coli* KO
1067 strains that we could use for this purpose[53]. To identify relevant KEIO KO strains, we used
1068 eciJO1366 and parsimonious FBA[74] to predict the effect of a gene knockout on optimal metabolic
1069 fluxes, and previous metabolomics data[52] of these strains to see the effect on intracellular
1070 concentrations (Fig. S20A). We eventually chose seven strains (Δ*aceE*, Δ*cyoD*, Δ*nuoA*, Δ*pgi*, Δ*rpe*,
1071 Δ*sdhB*, Δ*sucB*) that were predicted to have different optimal flux distributions, had many
1072 significantly changed intracellular metabolite concentrations, and targeted different part of key
1073 metabolic pathways (Fig. S20). We included Δ*lacA* as a negative control expected to behave like
1074 the WT as done previously[98].
1075

1076 The selected strains from the KEIO collection were precultured on LB agar and liquid LB as
1077 previously described, and glycerol stocks used for experiments described below were prepared
1078 by adding glycerol to a final concentration of 25 %.
1079

1080 For exometabolome sampling, the KEIO strains and *E. coli* BW25113 were precultured as detailed
1081 above. They were then re-inoculated to OD_{600} = 0.05 in 10 mL M9 + 40 mM galactose (+ 25 µg/mL
1082 kanamycin for KEIO strains) in 50 mL Erlenmeyer flasks (200 rpm, 37 °C, 31 h). 1 mL culture was
1083 sampled into Eppendorf tubes and washed twice with M9 without carbon source by
1084 centrifugation (6000 rcf, 6 min, 4°C). Ultimately, samples were resuspended in 0.5 mL M9 without
1085 carbon source and used to inoculate 200 µL M9 + 40 mM galactose (+25 µg/mL kanamycin for
1086 KEIO strains) in a flat bottom 96-well plate to OD_{600} = 0.05 and 6 mL M9 + 40 mM galactose (+25
1087 µg/mL kanamycin for KEIO strains) to straight glass tubes (15.8 cm tall, 1.5 cm diameter) to OD_{600}

1088 = 0.005. The 96-well plate was incubated in a Synergy H1 plate reader with continuous shaking
1089 (double orbital, 425 cpm, 37 °C), and the glass tubes in a shaking incubator (200 rpm, 37 °C, 45°
1090 angle), both for 108 hours. The experiment in the well plate was used to monitor more closely
1091 the growth (as OD_{600}) of these strains and to help with the timing of the exometabolome
1092 sampling. All strains were cultivated in three replicates. Two samples were collected during mid-
1093 exponential growth phase for all cultures, mostly to ensure that at least one sample from each
1094 replicate was obtained before the end of exponential phase. 1 mL culture volume was centrifuged
1095 (14 000 rpm, 10 min, 4 °C), and supernatant stored in -70 °C for exometabolome analysis. The
1096 second sample for each replicate (Fig. 4A) was analysed using LC-MS at the UNIL Metabolomics
1097 facility, following their protocol described above. Additionally, we also analysed three samples of
1098 pooled inoculums to verify that any carry-over with the inoculum was much lower (or below
1099 detection limit) of than the sample metabolite concentrations. Three replicate pooled samples
1100 were obtained by inoculating 6 mL M9 + 20 mM galactose medium with all the strains, each strain
1101 to an OD_{600} = 0.005, and subsequently handled as described above for the culture samples.
1102

1103 ### Chemostat evolution experiment with Δ*aceE* and Δ*sucB* in bioreactors

1104 Two of the KEIO KO strains, Δ*aceE* and Δ*sucB*, were selected for the evolution experiment with
1105 continuous cultivation (chemostats) in bioreactors based on their reduced growth compared to
1106 their ancestor and difference in exometabolome patterns (Figs. 4A and S21-S23). The strains were
1107 streaked from glycerol vials onto LB + 25 µg/mL kanamycin agar plates and incubated at 37 °C
1108 overnight. 5-6 colonies were transferred to 500 mL baffled shake flasks containing 75 mL LB + 25
1109 µg/mL kanamycin and incubated for 22 h (200 rpm, 37 °C).
1110

1111 Chemostat cultivations were performed in 1 L DASGIP bioreactors with a constant volume of 300
1112 mL M9 + 20 mM galactose medium + 25 µg/mL kanamycin (prepared as previously described, but
1113 without pH adjustment because of practical challenges with the large volumes required, so pH ≈
1114 7.1). Cultivation was performed at 37 °C with a submerged airflow maintained at 0.5 vvm, and a
1115 cascaded agitation of 400-700 rpm, controlled by a DO set point of 30 %. pH control was started
1116 at day 6 (Δ*aceE*) or day 20 (Δ*sucB*) with a set point of 7.1, controlled using 2 M NaOH (Fig. S25A).

1117

1118 The reactors (three replicates per strain) were inoculated to OD_{600} = 0.1. During the late stage of
1119 exponential growth, pumps were started to feed fresh culture medium and remove used medium.
1120 Medium was fed using a pre-calibrated peristaltic pump and used medium was removed by a
1121 steel pipe placed right above the culture surface. The steel pipe was coupled to a peristaltic pump
1122 operating at a flow rate approximately 1.5x of the medium in flow, to avoid volume accumulation.
1123 The set dilution rate was adjusted several times during the experiment to compensate for
1124 (sometimes surprising) changes in growth dynamics (from 0-0.1 h^{-1} for Δ*aceE*, 0-0.3 for Δ*sucB,*
1125 Fig. S25B). The actual dilution rate was monitored by keeping the flasks with fresh culture medium
1126 on scales. OD_{600} was measured daily, and weekly, larger samples were collected. Culture was
1127 centrifuged (3220 rcf, 10 min, 4 °C), and supernatant and pellet stored at -70 °C. In addition,
1128 cultures were preserved at -70 °C as glycerol stocks. The number of generations were estimated
1129 from the measured dilution rates across the experiment plus the generations during the initial
1130 batch culture phase before the culture dilution was started.

1131

1132 Cultivation of single colonies from evolution experiment and sampling of exometabolome

1133 Culture samples from different bioreactors and timepoints collected in the chemostat experiment
1134 were streaked out onto LB + 25 µg/mL kanamycin agar plates and incubated overnight at 37 °C.
1135 Three colonies were picked and individually inoculated in 20 mL LB + 25 µg/mL kanamycin in 100
1136 mL Erlenmeyer flasks and incubated (37 °C, 200 rpm, 16 h). 12 mL of the culture was then
1137 centrifuged (3220 rcf, 6 min, RT), and the pellets stored at –20 °C for sequencing. The isolates
1138 were stored as glycerol stocks at -70 °C made from these cultures.

1139

1140 2 mL samples from the precultures were washed as previously described, resuspended in M9
1141 without carbon source, and used to inoculate a flat bottom 96-well plate to OD_{600} = 0.05 with 200
1142 µL M9 + 25 mM galactose + 25 µg/mL kanamycin. The M9 medium was prepared as previously
1143 described, but to a pH of 7.1 as in the chemostats. The well plate was incubated in a Synergy H1
1144 plate reader with continuous shaking (double orbital, 425 cpm, 37 °C, 120 h) and the OD_{600} was
1145 read every 20 minutes.

1146

1147 Based on these growth measurements we selected the fastest growing isolates from the last
1148 timepoint of each of the six bioreactors (three bioreactors per strain). These isolates, the initial
1149 KEIO KO strains Δ*aceE* and Δ*sucB*, and the KEIO ancestor *E. coli* BW25113 were precultured and
1150 subsequently washed as described above. The washed preculture were used to inoculate 10 mL
1151 M9 + 20 mM galactose to OD_{600} = 0.01 in straight glass tubes (15.8 cm tall, 1.5 cm diameter) in
1152 triplicates and incubated for 20-84 hours (200 rpm, 37 °C, 45° angle). 1 mL samples for
1153 exometabolome analysis were collected at $OD_{600} \approx 1$ (Fig. 4, D and E), filtered using a 0.22 µm
1154 Millex-GV PVDF syringe filter (SLGVR33RS, MilliporeSigma, Darmstadt, Germany) into a 1.5 mL
1155 Eppendorf tube stacked in a cold block (-20 °C), and then stored at -70 °C until it was analysed.
1156 The exometabolome LC-MS analyses were conducted by the UNIL Metabolomics facility, following
1157 the procedure described above.

1158

1159 ## DNA extraction of samples from selected isolates and chemostat culture samples

1160 The DNA extraction of 14 different samples from the chemostat cultures (#37A D44 M2 Pellet,
1161 #38A D44 M3 Pellet, #39A D44 M4 Pellet, #28B D30 M5 Pellet, #29B D30 M6 Pellet, #30B D30 M7
1162 Pellet) and the selected isolates (sucB-M5-D30-4, sucB-M6-D30-6, sucB-M7-D30-4, sucB-
1163 Ancestor, aceE-M2-D44-2, aceE-M3-D44-3, aceE-M4-D44-1, aceE-Ancestor) was performed by
1164 adapting the protocol from the FastPure Bacteria DNA Isolation Mini Kit (DC103-01, Vazyme,
1165 Nanjing, China). 1 mL GA buffer was added to the thawed pellet samples, and 230 µL of these
1166 resuspensions were kept in Eppendorf tubes for DNA extraction. The rest was centrifuged (4000
1167 rpm, 6 min, RT) and the pellet stored at –20 °C. Genomic DNA (gDNA) was extracted from the 230
1168 µL sample following the manufacturer's protocol. In the final step, elution of the DNA was
1169 performed twice with fresh elution buffer, eluting a total of 100 µL DNA. Extracted gDNA was
1170 stored at –70 °C until sequencing.

1171

1172 ## Genomic sequencing processing

1173 Genomic DNA from the selected isolates and chemostat samples was sequenced by Novogene
1174 (Planegg, Germany) using the Illumina NovaSeqX PE 150. The isolates and the bioreactor samples

1175 were sequenced to 1 Gb and 10 Gb sequencing depth, respectively. Good quality of both datasets
1176 was ensured using fastqc v. 0.12.1 before trimming the reads with fastp v. 0.24.0 (--
1177 qualified_quality_phred 20 --cut_front --cut_tail --average_qual 20 --length_required 50)[99,100].
1178 Both read sets were mapped against the *E. coli* BW25113 reference genome (NCBI RefSeq
1179 assembly GCF_000750555.1) with minimap2's v. 2.28 default "sr" parameters[101]. The resulting
1180 alignments were filtered with samtools view v. 1.20 (-b -f 3 -q 60) to filter out reads not properly
1181 mapped in pairs or mapped to the reference with a MAPQ lower than 60[102]. Average coverage
1182 was assessed for both datasets showing a minimum average coverage for the chemostat culture
1183 samples of 2017 and 190 for the sequenced isolates. For both datasets, variants were identified
1184 with freebayes v. 1.3.6 (--min-alternate-count 3 -p 1 --min-alternate-fraction 0.05 --pooled-
1185 continuous --haplotype-length 0) and annotated with snpEff v. 5.0 using the BW25113 reference
1186 genome[103,104]. A custom python script was used to filter the annotated variants. For the isolates,
1187 variants with a frequency lower than 0.95 and a coverage lower than 100 were filtered out. For
1188 the chemostat culture samples, variants with a frequency lower than 0.05 and coverage lower
1189 than 100 were filtered out. A variant was considered fixed if its alle frequency was at least 0.9.

1190 ## Paired intracellular and extracellular metabolomics and live/dead staining

1191 M9 medium with 20 mM galactose, 30 mM malate or 40 mM L-alanine was prepared as detailed
1192 above. *E. coli* K-12 MG1655 was precultured on LB agar plates and in liquid LB as detailed above.
1193 1 mL preculture was washed as described above and used to start a second preculture in 20 mL
1194 M9 + 40 mM L-alanine, inoculated to OD_{600} = 0.05. Another LB preculture of *E. coli* K-12 MG1655
1195 was used to inoculate a second M9 + 20 mM galactose, and a M9 + 30 mM malate preculture in
1196 the same way 12 hours later. After 36/24 hours, 1 mL was collected from each of the three
1197 precultures and washed twice as detailed above. These washed precultures were used to
1198 inoculate 10 mL of the same media (M9 + either 20 mM galactose, 30 mM malate or 40 mM L-
1199 alanine) in straight glass tubes (15.8 cm tall, 1.5 cm diameter). These glass tubes were incubated
1200 for two days (200 rpm, 37 °C, 45° angle). Growth was monitored by measuring OD_{600} directly in
1201 the glass tubes (NANOCOLOR VIS II, MACHEREY-NAGEL, Düren, Germany) at 21 timepoints. The
1202 glass tubes were vortexed briefly before any OD reading or sampling.
1203

1204 A reference exometabolome 0.5 mL sample was collected after 2 hours following the procedure
1205 detailed below. Then, 2 mL samples for paired intra- and extracellular metabolome and live/dead
1206 analyses were collected in the late exponential phase (at $OD_{600} \approx 1.1$, Fig. S32). Exometabolome
1207 samples were obtained by filtering 0.5 mL culture using a 0.22 µm Millex-GV PVDF syringe filter
1208 (SLGVR33RS, MilliporeSigma) into a 1.5 mL Eppendorf tube stacked in a cold block (-20 °C). The
1209 samples for intracellular metabolomics, i.e. cells, were obtained immediately after following this
1210 procedure: A 25 mm 0.22 µm Durapore PVDF filter (GVWP02500, MilliporeSigma) was placed in
1211 a new filter holder. A 5 mL syringe was connected to the inlet of the filter holder, and a vacuum
1212 pump (250 mbar) was connected to the outlet of the filter holder. Then, 2 mL of 37 °C MQ water
1213 was added to pre-wet the filter. Immediately after, 1 mL of culture sample was added to the
1214 syringe and allowed to pass through the filter. If the culture did not pass through solely based on
1215 the vacuum pump, light additional pressure was added by using the syringe itself to push the
1216 culture through the filter. This process took 10-30 seconds. Still using the same set-up, the filter
1217 was finally rinsed with 5 mL of 37 °C MQ water as previously recommended[47]. Two tweezers were
1218 quickly rinsed in MQ water and used to carefully transfer the filter from the filter holder to 2 mL
1219 lysis tubes filled with 1.4 mL of 80:20 methanol:water (v/v), pre-chilled to -20 °C and kept on a
1220 cold block also pre-chilled to -20 °C. All samples were then quickly transferred to -70 °C for
1221 storage. Live/dead staining was performed using a similar approach as previous work[105]: Briefly,
1222 the *live* sample was made by diluting 10 µL culture in 990 µL PBS and the *dead* sample was made
1223 by mixing 100 µL culture with 1 mL 70% isopropanol. Both samples were vortexed and incubated
1224 at room temperature for 1 h. Then, the isopropanol was removed from the dead sample by
1225 centrifugation (8000 rcf, 4 min, RT), pipetting off the supernatant, resuspending the pellet in 1 mL
1226 PBS and vortexing. This procedure was repeated once. Then, both the live and the dead sample
1227 were stained with propidium iodide (PI; P4170, Sigma-Aldrich) and SYBR green (Invitrogen S7563,
1228 Thermo Fisher Scientific, Carlsbad, CA, USA) with the following procedure: 48 µL sample (live or
1229 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
1230 and incubated for 15 min in the dark. Then, a 0.1X samples of both the stained live and the stained
1231 dead samples were created by diluting 10 µL of each sample in 90 µL of PBS. The stained live/dead
1232 samples were then analyzed immediately on a CytoFLEX S flow cytometer (Beckman Coulter,

1233 Brea, CA, USA) and the CytExpert software (v2.4.0.28). The dead samples were used to make the
1234 gatings required to quantify the fraction of dead cells in the live samples on the PE-H and FITC-h
1235 signals (Table S7, Fig. S32). The extra- and intracellular metabolomics samples were analysed by
1236 the metabolomics facility at UNIL to quantify the five metabolites leucine, glutamate, aspartate,
1237 fructose 6-phosphate and cis-aconitate. These metabolites were chosen because they were
1238 chemically and metabolically diverse and likely to be quantifiable both intracellularly and
1239 extracellularly based on previous experience with similar samples. Also, as glutamate is the most
1240 abundant intracellular metabolite[47,106], it serves as a best-case scenario in our attempt to test
1241 how much of the extracellular metabolite levels that can be explained by cell lysis and the
1242 associated release of intracellular metabolites. Intracellular metabolite samples were first
1243 normalized by the amount of protein detected in the same sample and subsequently converted
1244 to absolute intracellular concentrations using the estimated protein density in *E. coli* ($13.5 \cdot$
1245 $10^{-8}\ \mu g/\mu m^3$)[107]. The fructose 6-phosphate value for sample 1B, timepoint 1, was discarded as
1246 an outlier as it was one order of magnitude larger than any other fructose 6-phosphate
1247 measurements.
1248

1249 ## Contribution from intracellular metabolites and cell lysis to extracellular concentrations

1250 Approximate cell volume was estimated from measured OD_{600} and a previous estimate of OD to
1251 cell volume for *E. coli* of 3.6 µL/OD_{600}/mL[108]. From approximate cell volume, measured absolute
1252 intracellular concentrations and cell lysis fractions we then calculated the corresponding change
1253 in extracellular concentrations. This potential contribution was then compared to the net change
1254 in extracellular concentrations between the T0 reference sample (after 2-hours) and the late
1255 exponential phase sample. If all T0 values were below the detection limit, we assumed the initial
1256 concentration of this metabolite to be 0. If only some values at T0 were below the detection limit
1257 we imputed these values to the mean of the detected values, giving conservative estimates for
1258 the net change in extracellular concentrations.
1259 To identify typical biomass degradation products, we used the metabolites consumed by the
1260 biomass equation of the E. coli model iML1515[95], although excluding the soluble pool (as
1261 categorized in[93]) since this should correspond to the intracellular concentrations already

1262 accounted for. Of the metabolites in our dataset (Fig. 3B), only the amino acids were among the
1263 biomass components and only protein depolymerization is therefore covering the contribution
1264 from biomass degradation in this case. To estimate a potential contribution from protein
1265 depolymerization we estimated cell density in gDW/L from OD_{600} (0.346 gDW/L/OD_{600}) and from
1266 this specific amino acid density (in g/L) from recent estimates of *E. coli* amino acid mass
1267 fractions[109]. When mass fractions were attributed to pairs of amino acids (glutamate/glutamine
1268 and aspartate/asparagine), we assumed equal contributions. Amino acid density was converted
1269 from g/L to µM by dividing by the amino acids' molecular weight subtracted the weight of a water
1270 molecule to account for the condensation reaction in protein polymerization. A similar procedure
1271 as described above was followed when we extended this analysis by using literature values for
1272 intracellular concentrations. To map extracellular concentrations in our dataset to literature value
1273 for intracellular concentrations, we used the following procedure: We first tried to map to
1274 literature values from cultivations in minimal medium with the same growth phase
1275 (exponential/stationary) and carbon source. If this wasn't possible, we prioritized mapping to
1276 data from minimal medium conditions of the correct growth phase. If concentrations from
1277 minimal medium conditions were not available, we mapped to data from complex growth media.
1278 When multiple intracellular concentrations mapped to our extracellular data point, we used the
1279 median of the intracellular concentrations. When extracellular concentrations were reported as
1280 two indistinguishable molecules, we attributed equal amounts to the two metabolites. This
1281 includes 2-phosphoglycerate/3-phosphoglycerate and ribulose 5-phosphate/xylulose 5-
1282 phosphate.
1283

1284 ## Large-scale screening of metabolite release from transporter KO mutants

1285 We selected sixty-seven strains from the KEIO collection[53] to assess if transporters (or
1286 porins/channels) play a role in metabolite release (Table S5). Each of these strains lacks a gene
1287 associated with metabolite transport. Functional annotations of these genes were obtained from
1288 EcoCyc[56] and TCDB[110]. All strains were precultured and washed as described above and then
1289 inoculated to OD_{600}=0.01 in 10 mL of 20 mM M9 glucose medium (pH=7.4) in straight glass tubes
1290 (15.8 cm tall, 1.5 cm diameter) for 20 hours (37 °C and 200 rpm). The glass tubes were angled 45°

1291 in the shaking incubator to improve mixing and aeration. The whole experiment was conducted
1292 in 5 separate batches (conducted on separate days). The KEIO reference strain *E. coli* BW25113
1293 was included in 6 technical replicates in each batch. All other strains cultivated in three technical
1294 replicates.
1295
1296 During the experiment, growth was monitored by measuring OD_{600} directly in the glass tubes on
1297 1- or 2-hour intervals (NANOCOLOR VIS II, MACHEREY-NAGEL, Düren, Germany). The
1298 exometabolomes were sampled at 4,8,10,11,12,13,14,15,16 and 20 hours (+ at 23 and 24 hours
1299 for Δ*tolC*). These timepoints were selected to cover the whole exponential and stationary phase,
1300 with highest temporal resolution in the last part of the exponential phase and during the
1301 transition to stationary phase where expected the largest changes in extracellular metabolite
1302 concentrations to occur. The exometabolome sampling was carried out by first vortexing each
1303 glass tube and then immediately collecting 150 µL into a 96-well filter plate (Millipore MultiScreen
1304 HTS GV Filter Plate, 0.22 µm Hydrophilic PVDF membrane, MSGVN22). The filter plate was stacked
1305 on top of a receiver plate and briefly centrifuged (1 min, 1000 rcf) to filter the cultures. The
1306 receiver plate and the filter plates were both immediately placed on an aluminum plate pre-
1307 chilled and kept on dry ice to cool down the samples quickly before the samples were moved to
1308 a -70°C freezer.
1309
1310 The samples were measured by flow-injection analysis time-of-flight mass spectrometry (FIA-TOF
1311 MS) using an Agilent 6550 quadrupole time of flight mass spectrometer as previously described[54].
1312 Samples were injected online utilizing an Agilent 1100 Series HPLC system (Agilent) coupled to a
1313 Gerstel MPS 3 autosampler (Gerstel). The mobile phase flow rate was set at 0.15 mL/min, with
1314 the isocratic phase composed of 60:40 (v/v) isopropanol and water buffered to a pH of 9 with 4
1315 mM ammonium fluoride. The instrument was run in 4 GHz full-scan mode, collecting spectra
1316 between 50 and 1000 *m/z*. Online mass axis correction was performed with taurocholic acid and
1317 Hexakis (1H, 1H, 3H-tetrafluoropropoxy–phosphazene). For these samples, masses were
1318 annotated by exact mass to within 1.0 ppm of the theoretical monoisotopic mass against
1319 *Escherichia coli* KEGG metabolites via in-house software as previously described[54]. Processed

1320 mzML files for this dataset can be found at MASSIVE: MSV000097105. Samples from all batches
1321 were initially analyzed in one continuous run (randomized within sampling plates) with four
1322 injections (technical replicates) of each sample. A standard was analyzed at regular intervals
1323 between samples. However, because of the large number of injections (>10,000) instrument
1324 performance declined considerably during the analysis of batch 5, which was therefore re-
1325 analyzed later in a separate run. FIA-MS data were annotated against KEGG identifiers to within 1ppm
1326 of each metabolite's theoretical mass. In some cases, the complexity of MS data (i.e. close neighboring
1327 peaks) yielded improperly centroided mass values which subsequently failed annotation to this stringent
1328 and highly conservative mass accuracy. This led to some differences in annotated metabolites between
1329 batch 1-4 and 5.
1330
1331 TIC-normalized data was used for the analyses described below. First, we filtered the dataset by
1332 removing *m/z* scores where none of the KEGG annotations mapped to the *E. coli* metabolite
1333 database (ECMDB)[46]. SMILES from ECMDB were then used to classify metabolites with
1334 ClassyFire[96]. We used primarily the superclass level to group metabolites, but we separated the
1335 class "Organic acids and derivatives" into "Amino acids", "Carboxylic acids", "Keto / hydroxy acids"
1336 and "Other organic acids". To reduce complexity, we further merged the classes that mapped to
1337 few metabolites into a class named "Other", i.e. "Benzenoids", "Other organic acids",
1338 "Phenylpropanoids and polyketides" and "Organic nitrogen compounds". We then removed
1339 severe outliers, focusing on occurrences where multiple consecutive injections were affected, as
1340 these errors could propagate despite using the median rather than the mean of 4 injections in
1341 further data analyses. We standardized the data from batch 1-4 and batch 5 separately after
1342 removing outliers and before calculating the median value across the 4 injections from each
1343 sample.
1344
1345 To quantify the effect of each KO on each metabolite, we compared the time-series median
1346 standardized Z-score between the KO mutant and the WT, only using the data from the same
1347 batch. To calculate the spread of the data for the WT and KO mutant we first fitted a smoothing
1348 spline (using make_smoothing_spline in Scipy[111], with the regularization parameter = 10) to all

1349 median Z-score values of each strain with the area-under-the curve of the OD600 readings on the
1350 x-axis (Fig. 5A), with weights inversely proportional to the standard deviation of each data point.
1351 We then calculated the spread as the mean absolute deviation from this smoothing spline. We
1352 calculated the difference between the KO strain and the WT as the mean distance between the
1353 two smoothing splines across x-axis locations corresponding to the underlying datapoints and
1354 within the x-range of both datasets. We then calculated the *Effect* as the ratio between the mean
1355 difference and the mean spread. To assess the significance of the estimated effects, we repeated
1356 this procedure over 1000 permutations where we shuffled the KO strain / WT label of all data
1357 points in each permutation. We then counted the number of permutations where we observed a
1358 higher / lower effect than in the actual data (for positive and negative effects).
1359

1360 ## Amino acid auxotroph co-cultures

1361 *E. coli* strains auxotrophic for proline (*E. coli* K-12 MG1655 Δ*proC* mCherry), isoleucine (*E. coli* K-
1362 12 MG1655 Δ*ilvE*) and methionine (*E. coli* K-12 MG1655 Δ*metB* GFP) were kindly provided by
1363 Estelle Pignon[112,113]. The strains were transformed with the plasmid pEP28 providing
1364 chloramphenicol resistance[112]. We then cultivated each auxotroph in monoculture, with *E. coli*
1365 BW25113 and with 2-4 relevant transporter KO strains (proline: Δ*proP*, Δ*putP*; methionine:
1366 Δ*brnQ,* Δ*cstA,* Δ*metI,* Δ*mtr*; isoleucine: Δ*glnP*, Δ*tsx*, Δ*yeaS*, Δ*ygaZ*). Strains were precultured (with
1367 25 µg/mL chloramphenicol / kanamycin for auxotrophs / KEIO KO strains) and washed as
1368 previously described, inoculated to OD_{600}=0.001 in 20 mM glucose M9 medium (200 µL) and
1369 cultured in BioTek Synergy H1 platereader (24 hours, 37 °C) as previously described. After 24
1370 hours, the auxotrophs abundances were quantified as CFUs/mL on LB+25 µg/mL chloramphenicol
1371 agar plates. Because methionine was not quantified in the batch 5 data, we had no information
1372 about the effect of Δ*metI* on methionine from the main dataset. However, not all metabolites
1373 were affected by the deteriorating mass-spec quality in the first run of batch 5. We therefore went
1374 back to the first run of batch 5, verified that the distribution of methionine peak intensities was
1375 similar in batch 5 as in batch 1-4, and then estimated the effect of Δ*metI* on methionine from this
1376 data (Effect = 2.56, $P < 0.005$).
1377