Proteome allocation and the evolution of metabolic cross-feeding

Florian J. F. Labourel[1,2,3]✉, **Vincent Daubin**[1], **Frédéric Menu**[1], **Etienne Rajon**[1]

[1]Univ Lyon, Universite Lyon 1, CNRS, Laboratoire de Biometrie et Biologie Evolutive UMR5558, Villeurbanne, France; [2]Milner Centre for Evolution, University of Bath, Bath BA2 7AY, United Kingdom; [3]Department of Life Sciences, University of Bath, Bath BA2 7AY, United Kingdom

Key words: Metabolism | Specialisation | Proteome optimisation | Gene expression | Mechanistic fitness

Significance statement: Can two species thrive on a single energetic resource? While the competitive exclusion principle predicts that one in the pair should go extinct, it may occur that an organism releases partly metabolised molecules in the environment, securing an ecological niche for a second organism in a specialisation process called metabolic cross-feeding. Here we investigate how evolution may favor the waste of a useful resource using a model that considers how a cell packed with proteins may be less efficient, hence favoring a shortening of metabolic pathways in order to reduce cell packing. Our model indicates that such specialisation only occurs under restricted conditions. Incidentally, this makes the signatures of cross-feeding, such as which metabolites are preferentially involved, quite predictable.

✉ **For correspondence:** fjfl20@bath.ac.uk

Data availability: All the scripts generated for this project are available at: https://github.com/FloTuzoLab/Scripts-Evolution-CF-proteome-allocation.

Competing interests: The authors declare no competing interests.

Abstract

Metabolic cross-feeding (MCF) is a widespread type of ecological interaction where organisms share nutrients. In a common instance of MCF, an organism incompletely metabolises sugars and releases metabolites that are used by another as a carbon source to produce energy. Why would the former waste edible food, and why does this preferentially occur at specific locations in the sugar metabolic pathway (acetate and glycerol are preferentially exchanged) have challenged evolutionary theory for decades. Addressing these questions requires to model the cellular features involved; to this end, we built an explicit model of metabolic reactions, including their enzyme-driven catalysis and the cellular constraints acting on the proteome that may incur a cost to expressing all enzymes along a pathway. After showing that cells should in principle prioritise upstream reactions when metabolites are restrained inside the cell, we investigate how the diffusivity of these metabolites may trigger the emergence of MCF in a population. We find that the occurrence of MCF is rare and requires that an intermediate metabolite be extremely diffusive: indeed, up to high membrane permeability coefficients, the expected evolutionary outcome is not a diversification that resembles MCF but a single genotype that instead overexpresses downstream enzymes. Only at very high levels of membrane permeability and under distinctive sets of parameters should the population diversify and MCF evolve. These results help understand the origins of simple microbial communities, and may later be extended to investigate how evolution has progressively built up today's extremely diverse communities.

Introduction

Genetic diversification [1, 2] may occur when different ecological niches are encountered [3–5], for instance when different carbon sources are available in the environment [6, 7]. What may at first glance sound puzzling – why not using all the available nutrients? – finds an explanation in physiological and environmental constraints or even absolute incompatibilities that make specialists of each resource outperform generalists [8–13]. Even more bewildering is the observation that diversification occurs in the homogeneous presence of a single energetic resource [14–17]. One finds a clear example in chemostats or controlled experimental systems in which glucose is continuously injected, where glucose consumers may evolve that release metabolites for others to use as carbon sources [14, 18, 19]. This unidirectional by-product process is a form of metabolite cross-feeding [20, 21], and its evolutionary underpinnings are still blurry [20, 22].

In particular, the reasons why specific metabolites are more likely involved in cross-feeding remain unclear. Indeed, a large number of metabolites produced by a glucose-reliant organism may constitute viable single carbon sources for others [for example, a glucose-reliant strain of *Es-*

cherichia coli can theoretically produce up to 58 useful metabolites, 22]. Yet only two metabolites are commonly reported – from experiments – as being traded in such cross-feeding interactions, namely acetate and glycerol [14, 19, 23, 24]. In line with the fact that some of these lineages may have been predisposed to use these metabolites [23], San-Roman and Wagner [25] have hypothesized that this preferential evolution could be due to shorter mutational paths. Accordingly, modifying the metabolic network to produce these interacting strains would require fewer and less destabilising mutations and could thence arise more readily. But their conclusion is that acetate or glycerol trades are no more likely than others to appear through mutation, making the mystery about their involvement – and potential predispositions of strains – in the evolution of metabolic cross-feeding even deeper.

Adaptation is often incomprehensible without considering the ability of an organism to perform a task as being dependent on internal constraints [26–29]. For example, fully expressing all enzymes along a metabolic pathway may incur a fitness cost [30, 31], to such an extent that sacrificing a part of a pathway becomes beneficial [32, 33]. Cells whose cytoplasm gets crowded with proteins actually pay a two-fold cost [34]. First, producing enzymes incurs a direct energetic expense, approximately proportional to the sum of enzyme concentrations in the cell [35–37]. Second, cell packing eventually compromises the diffusion of proteins, thereby hindering metabolic efficiency [38, 39].

In a previous study, we have shown that the evolution of enzyme kinetic parameters and concentrations is contingent on their competition with other processes for their substrate [34]. One of these competing processes may be leakage through the cell membrane, such that highly diffusive metabolites should be processed by more efficient or concentrated enzymes. The combination of this requirement for high concentrations, and the cost of an abundant proteome, could make these metabolites the preferential breakpoints in a metabolic pathway. Very few metabolites can diffuse through membranes, either because of their size or due to their electronic properties [40]. Such diffusion may be direct, as is the case for glycerol, or indirect when a non-diffusive metabolite can spontaneously transform into a diffusive one, as is the case with acetate [41, 42].

In this work – see Figure 1 for an overview of the metabolic model – we first determine how cells should allocate their proteome when metabolite diffusion is limited. We find that upstream reactions should be favored when selective pressures are similar along the pathway. We then assess the hypothesis that cross-feeding evolves in response to the high diffusion rate of an intermediate metabolite, due to the incurred selection on downstream enzymes. Indeed, a genotype producing the diffusive metabolite must also be very efficient at metabolising it to prevent its loss, and thus pays a high cost over-expressing downstream enzymes that, at some point, may become insurmountable. Interestingly, a second genotype feeding on the intermediate metabolite and only expressing downstream enzymes would thrive in this context where the metabolite has become a resource. This is because, on top of saving on the expression of upstream enzymes, extensive over-expression downstream is no longer a requirement as the high permeability coefficient of the metabolite actually helps its uptake. Yet, the overall evolutionary process must be continuous, instead of the schematic two-steps sequence presented here, making it difficult to predict when the evolution of cross-feeding should occur.

In order to embrace this continuity, we use Adaptive Dynamics [43, 44] to model the evolution of the pattern of enzyme expression along the pathway. This framework is particularly suited to model the complex ecological interactions that may arise as the genotype(s) composing the population shape their environment [32, 45], by controlling the equilibrium frequencies of both the nutrient and the diffusing metabolite. We find that MCF sometimes evolves, characterised by an evolutionary diversification giving rise to the coexistence between a genotype only expressing the enzymes upstream the diffusing metabolite, and another one expressing the enzymes downstream. This occurs at very high membrane permeability levels of the diffusing metabolite, only

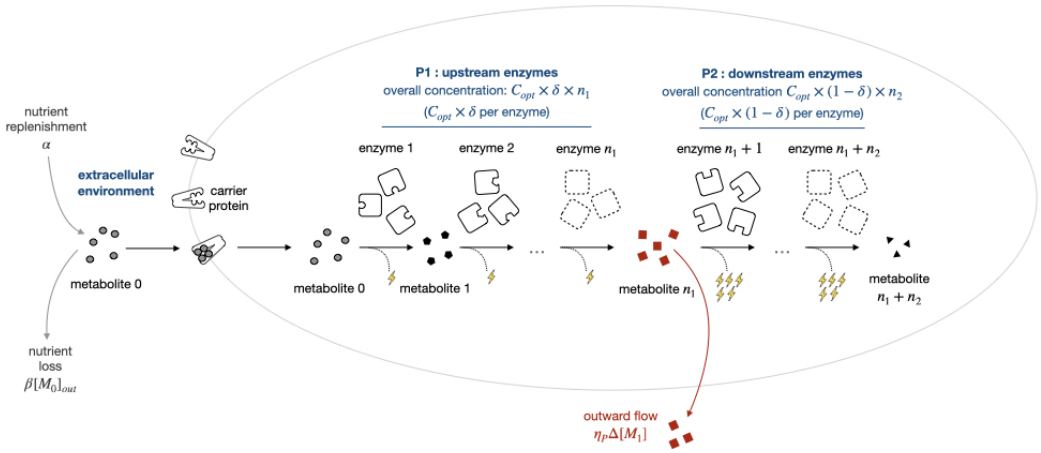

Figure 1. Overview of the model: the pathway is initiated by a carrier protein and comprised of n_1 upstream enzymes defining the sub-pathway P_1 and n_2 downstream enzymes defining the sub-pathway P_2. Allocation of the proteome is driven by two parameters: (i) C_{opt} coincides with the adaptive proteome fraction dedicated to one of the reaction of the full pathway $P_1 + P_2$ when all enzyme concentrations are identical (the case of no differential allocation – see section Optimal metabolic allocation); (ii) the parameter δ (*resp.* $1-\delta$) corresponds to the fraction dedicated to the first sub-pathway P_1 (*resp.* P_2). The extracellular dynamic of the nutrient (metabolite 0) is based on a constant replenishment-degradation process – see Materials and Methods section. Within the pathway, each reaction follows Briggs Haldane kinetics (with or without reversibility) where enzyme efficiency is constant and studied as a parameter. These reactions each provide a fitness yield, shown above as lightning, proportional to the amount of product produced and to the specific yield of energy gain set for reactions (1 energy unit per reaction per nutrient for upstream enzymes and 5 for downstream enzymes in the above example) – see section on overexpression. Fitness is simultaneously impeded by the cost of expression – production and crowding – and, depending on the subsection again, by the toxicity resulting from the total concentration of metabolites. The metabolite produced at the end of the upstream pathway P_1 is susceptible to leak through the membrane according to a permeability η_P – studied as a parameter – and the gradient of concentration between the extra and intracellular environments. Once in the extracellular environment, it is available to other cells but may also be degraded (like metabolite 0)

compatible with diffusion rates reported for acetate or glycerol. We also find that MCF requires moderate to high levels regarding the intracellular selective constraints acting on metabolites along the pathway.

Results

Optimal metabolic allocation and cell constraints

Evolution of the overall expression of metabolic enzymes

We first assume that all enzymes have an equal concentration and consider its evolution. Increasing concentration enhances the efficiency of catalysis and thus the production of energy, but with diminishing returns [34, 46]. It also incurs costs, firstly due to the actual energy cost of making proteins, and secondly because high protein concentrations in the cell decrease the efficiency of reactions due to cell packing. The former is captured in our model by a linear cost inflicted to extra production, and the latter through a penalty on k_f, whose effect has been estimated [39, 47] and modelled in previous studies [38] – [see Model and 34].

To approach the case of reactions involved in the carbon cycle [48], we consider a pathway comprised of 40 reactions and initiated by facilitated diffusion through a transporter [49], where energy is produced at each step in the process – see Figure 1 – and assessed the effect of reaction yield – see SM Figure S3. Reactions follow irreversible Briggs-Haldane kinetics [50], with kinetic constants set by default half an order of magnitude higher than the median observed for enzymes involved in the central metabolism [51] ($k_f = 10^{6.25} M^{-1} s^{-1}$, $k_{cat} = 10^{2.25} s^{-1}$) – see SM Table S2 for the

full set of parameters we used; besides, we also consider the case where reactions are reversible – see SM Text S1.2. Setting these efficiencies above the average found for enzymes involved in central carbon metabolism is conservative as it diminishes the selective pressure that promotes higher enzyme expression, and thus ensures that we are not examining an implausible range of values (*eg.* because the whole set would be biased by some enzymes that need not be efficient due to specific metabolic features). The local context of reactions – including reaction reversibility and metabolite toxicity – may also affect metabolic efficiency, as captured by a linear degradation rate of each metabolite η in this instance of the model [34] – see SM Text S1.2 for more complex selective pressures. Nutrients are added to the environment at a constant rate α and degraded at a linear rate β – the latter also applying to metabolites released by cells in the medium (see below) – whose levels are such that without cells, substrate concentration equilibriates at $1M$.

For all combinations of the parameters above considered, the evolutionarily expected concentrations of the 40 enzymes in the pathway sum up to $15-20$ % of the whole proteome (see SM – section Text S1). As we have shown previously, adaptive outcomes depend on a complex interaction between cell constraints, enzymes concentrations and kinetic efficiencies that cannot be captured through the influence of classical experimenters' parameters (such as K_M or the activity constant $k_{cat}[E_{tot}]/K_M$, for instance) alone [34]. The highest fraction is self-evidently obtained in conditions where selection for the rates of reactions is acute, such that increasing concentration becomes beneficial, to a certain extent, despite amplifying intracellular crowding and production costs. These predictions are consistent with estimates among unicellular species [48]: in most cases, enzymes involved in the carbon cycle constitute approximately 20 % of the proteome. In the remaining of this study, the overall concentration of enzymes in the pathway is considered fixed at its evolutionary expectation by default, *i.e.* that obtained for the specific combination of parameters studied when assuming an identical allocation all along the pathway.

Overexpression in upstream reactions

We then studied how cells should distribute enzyme expression along a pathway split into two parts of equal lengths. This is a proteome allocation problem, since the overall concentration is fixed to an (adaptive) optimum as just described; we study the evolution of the part of this overall concentration allocated to the first half P_1 of the metabolic pathway, δ, which we assume can change by mutation in the range $[0,1]$ – $\delta=0$ coinciding with no investment in P_1 such that all of the resources are allocated to P_2, while $\delta=1$ corresponds to all resources being allocated to P_1 (and none to P_2).

The evolution of δ is modelled using adaptive dynamics [32, 43, 44], as is appropriate when the fate of a mutant can depend on the environment shaped by one (or several) resident genotype(s). Here the resident strategy may impact the equilibrium concentrations of the nutrient and of the metabolites produced along the pathway. Nonetheless at this stage, each metabolite is considered to be unable to diffuse across the membrane, such that their concentration in the environment remains zero. Therefore in this case, the evolutionary outcome is always a single allocation strategy δ, as exemplified in Fig. 2. In this figure, whatever the initial resident strategy in place in the population, evolutionary trajectories will converge to $\delta\approx 0.6$, and, once in place in the population (as resident) this latter strategy will be stable against the invasion by mutants with any other δ. These features make an enhanced expression of upstream reactions – with $\delta\approx 0.6$ – the evolutionarily expected outcome, also described as a convergent stable strategy – CSS, hereafter.

In the presence of degradation, the adaptive investment in the first part of the pathway is generally above 0.5 : remarkably, δ evolves to more than 0.6 even at very low degradation rates where the resulting loss in metabolites is less than 1% along the pathway (see Table S3 of SM for an analysis of the influence of degradation on the loss of metabolites and Figure S5 of SM for more details). One factor that pushes upstream expression towards higher adaptive values is the selective pressure imposed by transporters – see Figure S5-B of SM for the influence of transporters. Yet, the process also holds when this influence is set apart. A plausible explanation is that within

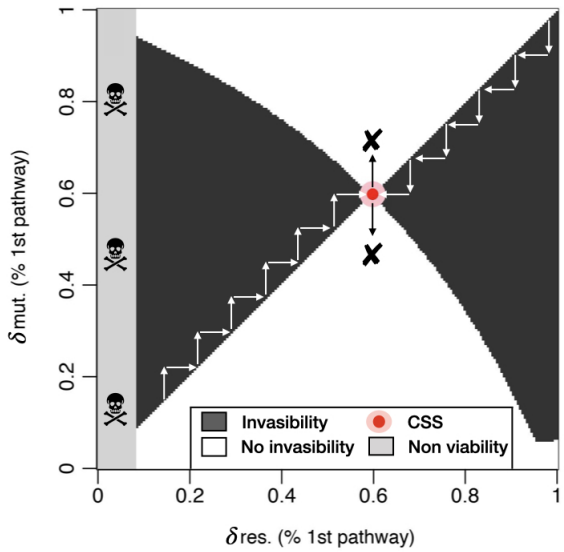

Figure 2. Example of the outcome of Adaptive Dynamics shown through a pairwise invasibility plot – PIP, hereafter – where black (resp. white) areas stand for positive (resp.negative) invasive fitness. The grey area represents an area where no strategy is viable because it cannot produce enough energy to reach the level set as corresponding to the demographic steady-state even when only one individual is present. When a mutant arrives in the population of resident strategies, it takes the place of the resident strategy if its invasion fitness is positive, a process represented by the white arrows. Here, below the CSS, mutants with higher values (black area above the left lower corner to right upper corner bisector) than the resident invades and evolution pushes resident to converge towards the CSS, while above the CSS, it is the other way around. This PIP shows how cells should spread their content, according to the respective fractions δ and $1-\delta$, between upstream and downstream enzymes and exemplifies that under the specific set of parameters chosen, they should prioritise upstream reactions since $\delta > 0.5$, as was generally found – see text for more details.

an irreversible pathway as the one modelled, upstream enzymes not only concur to fitness directly through the energy generated by their respective reaction, but also through their indirect contributions to downstream reactions – see SM Text S2 and Figure S6 for the analysis of this phenomenon through a more tractable model showing that the optimal strategy should always be to prioritise upstream reactions. This is also consistent with the fact that this unequal allocation holds when downstream reactions produce far more energy than their upstream counterparts so that upstream reactions contribute mostly indirectly to fitness (and is significantly heightened in the opposite case): even when increasing the yield of the reactions in the second half of the pathway tenfold as described for the carbon cycle [52], δ remains close to 0.6. Besides, the irreversible loss of metabolites caused by an increase in the degradation rate increases the asymmetry in fitness contributions further and thereby tends to increase the adaptive ratio of upstream to downstream enzyme expression – see Figures S5 and S15 of SM.

Such asymmetries in fitness contributions might help explain why enzymes catalysing reactions more upstream tend to face stronger selection [53–55]. Overexpression of upstream enzymes can nonetheless be – at least, partly – counteracted by selection for homogeneity in metabolite concentrations, as is the case when toxicity is high (Figure S7-B of SM) and equally spread, and it also depends on reactions reversibility, including that of the transporter, inasmuch as reversibility also shifts the balance between direct and indirect contributions of each subpathway: if downstream reactions are more reversible than upstream ones, cells should prioritise them, and vice versa – see Figure S7-A of SM. Be that as it may, considering realistic combinations of these pressures – moderate toxicity and degradation rates as well as the average reversibility found in central carbon metabolism [56] tends to corroborate the need for upstream overexpression, as shown in Section

Text S2.2 of SM. One constraint that can reshuffle the deck of cards is the permeability of the cell membrane, which was hitherto not considered.

Membrane permeability and cross-feeding

Membrane permeability impacts proteome allocation

Membranes are only permeable to a few metabolites, owing to their unique chemical features [40]. In our model, we allowed the metabolite produced by the last reaction of the first half of the metabolic pathway to diffuse – see the model overview in Figure 1 and the Model section for details – with permeability rates ranging from $P = 10^{-12}$ dm.s^{-1} to 10^{-4} dm.s^{-1}, in line with empirical estimates [40]. Metabolites making their way across the membrane have two consequences. First, they are lost for the cell that has produced them, which may act as a selective pressure for limiting diffusion. Second, they may accumulate in the external environment and be available to other individuals as a resource.

A cell has little leverage to limit outward diffusion; the most obvious solution is to use the metabolite before it is lost, which in our model is possible through an increase in the concentration of enzymes acting on the second part of the pathway. This is indeed what happens: the optimal allocation shifts from a higher concentration in the first part of the pathway – owing to the aforementioned factors – to a higher concentration in the second part as permeability increases, as shown for instance in Fig. 3 in the case of a low degradation rate (grey dots). The results presented in Fig. 3 are for proteins with kinetic parameters being slightly higher than the median reported for enzymes in the carbohydrate and energy pathway [51]. Higher efficiencies consistently produce a qualitatively similar result of a downward shift in allocation to the first part of the metabolic pathway, P_1, as permeability increases (see Figure S12 in SM). This outcome is also robust to the possible existence of different energetic yields along a pathway – see Figure S8 for the influence of yield on the optimal allocation – and most often holds when introducing other mechanistic constraints – see Figures S13 and S14.

High permeability coefficients can promote cross-feeding

The overinvestment in the second part of pathway, P_2, limits the loss of the leaky metabolite. As permeability increases, the investment in P_2 must increase further, but this strategy becomes less efficient such that the external concentration of the metabolite rises. This, in turn, may give a selective advantage to genotypes that give up preventing the metabolite's leakage by increasing their contribution to the first or second part of the pathway. From that point on, a new ecological niche may emerge, that can ultimately result in an evolutionary diversification between two genotypes. These situations can be identified by a specific type of pairwise invasibility plots where the singular strategy is convergent but evolutionarily unstable – as shown in Fig. 3-C for the case of a high degradation rate, the singular strategy at $\delta \approx 0.65$ can be invaded by nearby mutants – called a branching point. An adaptive diversification may occur at a branching point, requiring that we study the fate of mutants invading a pair of coexisting strategies, instead of a single one, through trait evolution plots (SM Figure S9).

Branching points indeed lead to an adaptive diversification as the diffusion of the intermediate metabolite increases above 10^{-6}dm.s^{-1}, with the two most extreme strategies, $\delta = 0$ and $\delta = 1$, evolving and forming a stable coalition (see SM - Text S3.1.2). This is characteristic of a complete metabolic cross-feeding, where a genotype transforms a nutrient into a metabolite released in the environment, and a second genotype feeds exclusively on that metabolite. For moderately high permeability rates – around $P = 10^{-6} dm \cdot s^{-1}$ – the evolutionary outcome is more dubious as invasion is only possible by distant mutants; the result, in this case, will depend on the distribution of mutational effects and cannot be studied using classic adaptive dynamics methods.

It should be noted that the results in Figure 3 correspond to a metabolic pathway with unequal contributions of reactions to the energy needs of the cell (downstream reactions produce more

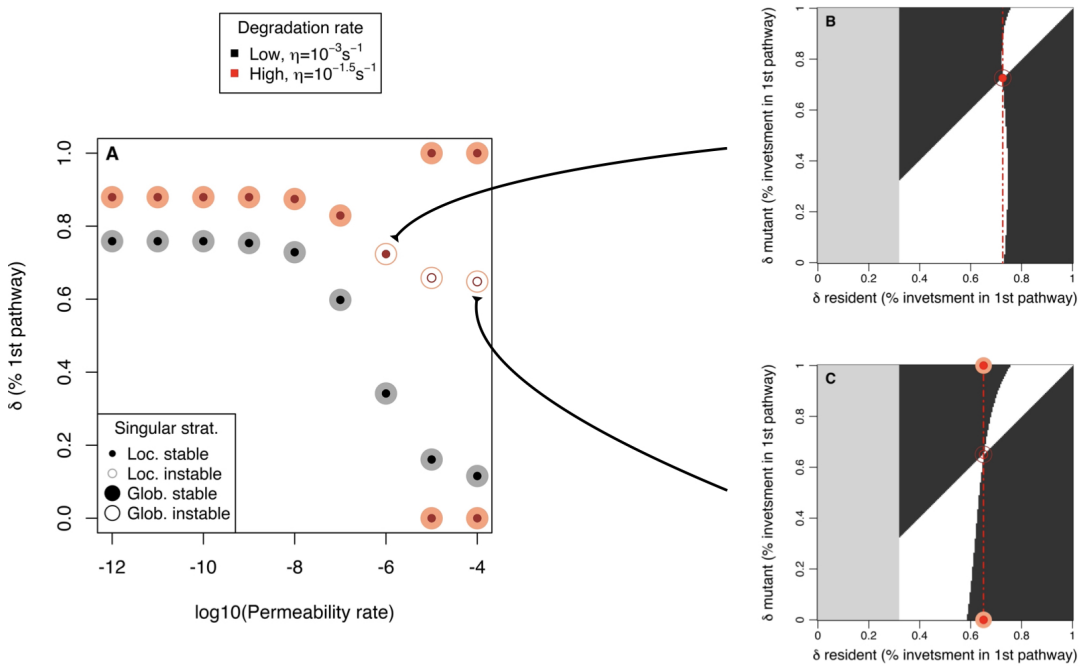

Figure 3. Permeability influences the evolution of strategies of enzyme allocations along a metabolic pathway, with the occurrence of cross-feeding at high permeability coefficients depending on the degradation rate. Two degradation rates are considered here (two other ones in Fig. S8 of SM) : low $\eta = 10^{-3}$ s^{-1} in grey and moderate $\eta = 10^{-1.5}$ s^{-1} in red. The enzymes catalysing reactions in either of two parts of the metabolic pathway, with their relative concentrations represented by δ (% 1st pathway), have kinetic parameters typical of those intervening in the carbon cycle ($k_f = 10^{6.25}$ M^{-1}. s^{-1} and $k_{\text{cat}} = 10^{2.25}$ s^{-1}); we also assume here that reactions in the second part P_2 of the pathway produce 10 times more energy than reactions in the first part P_1. For low permeability rates (below 10^{-6} dm.s^{-1}), the evolutionarily expected strategy is always globally and locally stable – as in Figure 2 – and consists in investing more in the first part of the pathway as discussed in the text – and described in further details through Fig. S5 and S15 (see the latter for the same enzyme efficiencies). Increasing permeability coefficients above 10^{-7} dm.s^{-1} results in a decrease in the investment in the first part of the pathway, δ, with different consequences for low and high degradation rates. At low degradation rates (in grey on panel A), lowering δ, *i.e.* increasing the investment in the second part, remains the most efficient strategy. At high degradation rates (in red on panel A), however, high permeability coefficients result in singular strategies that are both convergent (they evolve from any starting δ) and evolutionarily unstable (they can be invaded by mutants with close (locally unstable) or distant (globally unstable) allocation strategy δ – see panel C). This can lead to adaptive diversification, resulting in a stable coalition of strategies, that is, a population made of coexisting genotypes with different values of δ. The coalition can be determined (see SM - Figure S9) ; it is comprised of genotypes with $\delta = 1$ (expressing only the first part of the pathway P_1) and of genotypes with $\delta = 0$ (expressing only the second part P_2), corresponding to metabolic cross-feeding. Panel B represents a permeability $P = 10^{-6}$ dm.s^{-1} for which the singular strategy is globally instable but locally stable, so that the evolutionary outcome should be contingent to the mutational landscape.

than upstream ones, as observed on average in the carbon cycle). Equalling the energy contributions of all reactions often prevents the occurrence of cross-feeding. Yet, cross-feeding emergence was less sensitive to sub-pathway theoretical yields when we introduced the altogether realistic possibility of transporters and upstream enzymes co-expression, as exemplified on Figure S11. Reversibility and toxicity may also be involved in defining the critical point where cross-feeding coalitions can appear – see Section Text S3.3. Overall, the tipping point of $P = 10^{-6}$ dm $\cdot$ s^{-1} seems robust across all internal selective environments, albeit not being a sufficient condition as it also depends on the specific combination of constraints associated with these environments.

More generally, the evolution of cross-feeding appears contingent on many factors. For instance, selection on the overall concentration of enzymes contributing to the pathway is impor-

tant (*eg.* because it can reduce diffusive constraints), such that allowing higher concentrations precludes the occurrence of cross feeding. This is because increasing downstream concentrations, which may efficiently deal with metabolite diffusion, comes at a lower cost under these circumstances. This illustrates that the occurrence of cross-feeding may critically depend on the other tasks – and their contributions to fitness, which was not directly considered here, though the degradation rate may partly capture the behaviour – performed by cells and thereby on selection acting on their dedicated proteome, as the global proteome concentration may only vary to a small extent [40]. No less important should be the size of cells since smaller ones mechanically come with higher relative leakiness due to larger surface-to-volume ratios, which could in turn favor the occurrence of cross-feeding. Cells can nonetheless adapt their size to their content, at least to a certain extent [37], which may limit the costs incurred by an increase in concentration and prevent the occurrence of cross-feeding.

Finally, while in our model the efficiency of a reaction may only be changed through enzyme concentrations, kinetic parameters may also evolve and thus represent a relevant alternative in some specific parts of the metabolic network. The evolution of cross-feeding could thus also be contingent on the relative availability of mutations changing concentration *versus* kinetic parameters. Therefore, while this model includes much of the available information about enzyme kinetics and the selective constraints acting on the proteome, actually predicting how and when cross-feeding should evolve will require more efforts to better understand the building of global epistasis along metabolic networks and how critically this depends on the environment.

Discussion

The fact that few metabolites – acetate and glycerol, noticeably – are more likely involved in the evolution of cross-feeding has been a conundrum for as long as experiments have revealed this phenomenon [25]. Here, we have put forward an explanation based on the necessity for a cell to optimise its proteome allocation, accounting for incurred costs and physical constraints like the differential permeability of a cell's membrane to metabolites. Indeed, contrary to other metabolites, acetate is in constant chemical equilibrium with the highly diffusive acetic acid [42], and glycerol readily leaks towards the environment [40], which may predispose them to be involved in cross-feeding.

That metabolites rather upstream in a metabolic pathway – and therefore of potential use to generate more energy – will create a novel ecological niche when released is straightforward. But whether evolution will take this path when some genotypes have the potential to reduce the leak is worthy of a careful examination. In a number of cases that we have considered and that fall into the realistic range of parameters, proteome allocation will evolve in such a way that it prevents, or at least limits, the diffusion of the molecule. Only under some restricted conditions – that we have shown to coincide with the features of the two aforementioned metabolites – will cross-feeding evolve, characterised by a functional specialisation between a part of the population that transforms the nutrient into the diffusive metabolite, and another part that uses the metabolite as a carbon source, echoing work on digital evolution that similarly pointed to the possible contingency of cross-feeding [57, 58].

Beyond the existence of intrinsic cellular leakiness, one process that may facilitate the emergence of cross-feeding, or that may be co-opted to foster its efficacy when it is in place, is the cheap uptake and secretion of metabolites (*eg.* through facilitated transport) [42, 59, 60]. Even a slight leakage may for instance be enough to sustain a cross-feeder with high affinity transporters such as the one reported in *E.coli* for acetate [61]. This metabolic strategy, where cells actively give up a metabolite even though it still has the potential to bring fitness contributions, is often known as overflow metabolism [62, 63] and also frequently involves acetate. We did not consider the existence of a specific cost to the second part of the pathway, as has been documented in the past

to explain overflow [64, 65]; nor did we account for other possible costs such as the existence of a localised toxicity [66] or the extra membrane occupancy involved in cellular respiration [67, 68]. These processes may promote the advent of cross-feeding, for they could bring extra fitness to an organism expressing one or the other part of the pathway.

From a broader ecological perspective, nutrients are hardly ever constantly renewed in nature; they are subject at least to some stochasticity, and most often than not occur through unpredictable periods of feast and famine. Insight into how proteome allocation may evolve in this context, and whether it should likely involve cross-feeding, can be gained from the existence of so-called diauxic shifts, where the phenotype can switch from the production of acetate to its consumption when the medium is enriched in this nutrient [59, 69]. In this case, it appears that the environment could be used to store intermediate metabolites, both increasing a genotype's ability to uptake and use glucose when it is present, and its ability to await its renewal otherwise. Such plasticity in expressing different parts of a metabolic pathway should prevent, to some extent, the evolution of fixed specialists (as considered here) and further hinder the evolution of cross-feeding. We thus postulate that cross-feeding might only evolve in environments stable enough to prevent the evolution of plasticity.

Biology is actually teeming with interactions and emergent properties that complicate the big picture [70]. And, if functional simplicity may exist at the ecological scale [71], the ultimate underpinnings behind community assemblies are still blurred by the combination between a long and tumultuous evolutionary history and the wide prevalence of high order interactions, both within and between organisms, and in relation to the environment [57, 72, 73]. Addressing these community-level questions cannot bypass the existence of lower level features such as the epistatic relationships stemming from the joint effect of enzyme kinetic parameters [74] and their expression along pathways [75]; nor can we study metabolic evolution without a careful examination of their impact on the cellular micro-environment. We believe that the present study, by accounting for these interactions at a primordial ecological stage, will help explain how (some of) these communities appeared in the first place and how this may have then fuelled the scaffolding process underlying the building of microbial societies.

Models

Metabolic Model of fitness

Cell fitness results from the biomass and energy produced along a metabolic pathway (*eg.* ATP). The pathway is initiated by carrier proteins with $V_{Tm} = 1mM/s$ and $K_T = 10mM$ (see SM Text S1 for details and justification), passively transporting nutrients inside cells and whose features are based on those for glucose in yeasts, as detailed elsewhere [34]. Nutrients are added (resp. eliminated) at a constant rate α (resp. β) in the external environment. The metabolic pathway is linear, comprising N_r (40 in the paper, but see SM for other values) reactions catalysed by enzymes whose levels of expression may evolve. The product of the j^{ieth} reaction is used as the substrate of the next one. These metabolites are constantly degraded at a rate η or fuelling toxicity T according to processes justified elsewhere [34] and also detailed at the beginning of the SM (introduction of section Text S1 and subsection Text S1.2). Each reaction (j) follows either reversible or irreversible Briggs Haldane kinetics [76]:

$$\mathrm{E}_j + \mathrm{S}_j \underset{\mathrm{k}_{r,j}}{\overset{\mathrm{k}_{\mathrm{f,act,j}}}{\rightleftharpoons}} \mathrm{ES}_j \underset{\mathrm{k}_{\mathrm{inh,act,j}}}{\overset{\mathrm{k}_{\mathrm{cat,j}}}{\rightleftharpoons}} \mathrm{E}_j + \mathrm{P}_j \tag{1}$$

Irreversible reactions are modeled by setting $k_{inh} = 0$ – in this simple setting, the system can be solved for each pair of successive reactions since they do not feedback on upstream reactions, so that one has to solve N quadratic equations of the form:

$$Vm_j \frac{[S_j]}{[S_j] + K_{M,j}} = Vm_{j+1} \frac{[S_{j+1}]}{[S_{j+1}] + K_{M,j+1}}, \tag{2}$$

where V_m and K_M depends on microscopic parameters k_f and k_{cat}, as well as enzyme concentrations while η is the degradation rate (leakiness is modelled in a similar manner than the degradation rate but depends on the gradient between internal and external concentrations so that it also impacts the environment through the latter). Otherwise, reversibility is spread equally between backward parameters k_r and $k_{inh,act}$ (*eg.* if $K_{rev} = 1/9$, $k_r = k_{cat}/3$ and $k_{inh,act} = k_{f,act}/3$) and the system of N equations is solved using Broyden's method [77]. The actual values for $k_{inh,act}$ and $k_{f,act}$ depend on the cellular constraints – see next subsection of Model.

Each reaction produces energy. This is a simplification as some reactions in the carbon cycle do not, unimportant as we consider the global expression of large portions of a pathway. We consider the case where contributions are equal along the pathway as well as other more realistic setups.

Cellular constraints

Cell proteomes face two intrinsic constraints: (i) the cost of protein expression and (ii) the burden entailed by molecular crowding. We model (i) through a cost linear with protein expression, proportional to constant c [36]. We considered values of c such that the whole cytoplasmic proteome – the enzymes in the pathway and other free enzymes – costs 5 to 50% of the whole cell budget. We model molecular crowding (ii) through a non-linear decrease of diffusion [38, 39] that changes the affinity constant k_{f} to $k_{\mathrm{f,act}}$ in the model with irreversible reactions, according to equation:

$$k_{\mathrm{f,act}} = k_{\mathrm{f}}.10^{-([E_{\mathrm{other}}]+\sum_{i=1}^{40}[E_{tot,i}])/[M_b]}, \tag{3}$$

where $[E_{tot,i}] = [E_i] + [ES_i], [M_b] = 3 \cdot 10^{-3} M$ represents the scaling factor for the effect of diffusion, while $[E_{\mathrm{other}}]$ denotes the sum of the concentrations of other cytoplasmic proteins than the 40 under consideration. In the model with reversible reactions, k_{inh} is also affected by crowding, which is a conservative assumption as reverse reactions might be less sensitive to the diffusive process owing to the preexisting co-localisation between substrate molecules and enzymes.

Our model includes three processes involving the metabolites produced that select for the enhancement of enzyme activity, drawing a complex trade-off on the coexpression of enzymes: (1) metabolites can be lost, either because they are involved in parasitic reactions or because they are subject to targeted degradation [78], modelled through a linear degradation rate η; (2) metabolites can be toxic for the cell, for they engage in parasitic reactions, for instance through promiscuous interactions [79]; (3) highly reversible reactions within a pathway may also require efficient enzymes to maintain a high net flux [80]. We considered these three processes in various instances of our model, as described in SM Text S1; the results presented in the paper mainly comprise the action of a linear degradation rate, which provides a good qualitative understanding of how processes impacting the metabolites also impact selection on enzymes. Finally, the permeability of cell membrane to a given metabolite also acts as a constraint, which is introduced here by considering that one metabolite in the pathway diffuses passively at a rate η_d – on Figure 1, we show where this process occurs.

Ecological equilibria

As to perform the adaptive dynamics analysis, we need to determine the ecological equilibrium set by a given genotype or coalition. To this aim, we first compute the value of the net flux (number of energy units per individual per unit of time) $\Phi_{net,N}$ for a given population size N. As long as this flux is lower (or higher) than an arbitrary value $\Phi_{net} = 10^{-4} M \cdot s^{-1}$, N is increased (or decreased) at the next iteration (by one unit when the flux gets close enough from this value). The algorithm is stopped, either when the difference between these fluxes is lower than $10^{-6}\Phi_{net}$, or when it oscillates between two neighbouring values (N_{eq} is then set to the average between these values). The ecological equilibrium, is defined by the nutrient and metabolite concentrations in the environment at this demographic equilibrium. Fitness values of strategies are then calculated in these conditions of equilibrium.

Adaptive Dynamics of enzyme expression

We use Adaptive Dynamics [32, 43, 44] to model the evolution of enzyme expression along the metabolic pathway. This framework consider rare mutations, such that a resident "strategy" – corresponding to a given expression pattern – is assumed to have reached its demographic equilibrium before a mutant strategy appears in the population. At this demographic equilibrium, births compensate for deaths in the population, resulting in a concentration of nutrients specific of the resident (see previous section). The fitness of any mutant strategy is then determined for each resident equilibrium, which enables the drawing of Pairwise Invasibility Plots (PIPs) representing for each pair the ability of a rare mutant to invade the resident strategy, based on a comparison between the growth rate of the mutant and that of the resident. These plots are used to identify singular strategies and their properties, as defined in [43]. A particular type of singular strategies, branching points, may be indicative of a diversification in the population, which we further study by drawing areas of mutual invasibility (where both strategies invade when rare), before computing the ecological equilibrium for each coalition – composed of two resident strategies instead of one – in that space. We then calculate the growth rate of mutants nearby each strategy in the coalition to identify coalitions that are stable (they cannot be invaded by any of the nearby mutants) and convergent (there exist evolutionary trajectories towards them), hence identifying evolutionarily expected communities after diversification has occurred [81, 82]. This latter process is summarised on trait evolution plots – TEPs – that we have shown in SM.

PIPs were generally drawn for 250 strategies, unless lower resolution was sufficient to capture the trend. In order to determine the optimal allocation, we set the total proteome content to its optimal value as determined without the influence of permeability. An individual whose strategy is to invest as much in the first part of the pathway than in the second part corresponds precisely to this case. Notice that, besides using a two step process, solving the systems to find nutrient and metabolite concentrations required to use R package 'nleqslv' with Broyden's method [77].

Settings for the models

A list of the basic settings can be found at the end of section Text S1.1 - SM. We varied them within their biological realistic ranges. This allowed us to identify key drivers of the diversification process that eventually result in cross-feeding, as discussed in the results section. The extensive analysis of parameters can be found in SM, especially in Texts S1 and S3.

Acknowledgment

We thank T.Lenormand for helpful discussions that led us to consider the evolution of cross-feeding. This preprint was created using the LaPreprint template (https://github.com/roaldarbol/lapreprint) by Mikkel Roald-Arbøl.

Author contributions

FL, ER, FM and VD designed the research. FL and ER developed the models and analyzed the results. FL and ER wrote the first draft while FM and VD contributed further improvements.

References

[1] Sewall Wright. "The genetical structure of populations". In: *Annals of Eugenics* 15.1 (1949), pp. 323–354. DOI: https://doi.org/10.1111/j.1469-1809.1949.tb02451.x. URL: https://doi.org/10.1111/j.1469-1809.1949.tb02451.x.

[2] Francisco J Ayala et al. "Genetic differentiaiton during the speciation process in Drosophila". eng. In: *Evolution* 28.4 (1974), pp. 576–592. ISSN: 1558-5646 (Electronic); 0014-3820 (Linking). DOI: 10.1111/j.1558-5646.1974.tb00791.x.

[3] G. F. Gause. "Experimental Studies on the Struggle for Existence : I. Mixed Population of Two Species of Yeast". In: *Journal of Experimental Biology* 9.4 (1932), pp. 389–402. DOI: 10.1242/jeb.9.4.389. URL: https://doi.org/10.1242/jeb.9.4.389.

[4] G. Hardin. "The competitive exclusion principle." In: *Science* 131 3409 (1960), pp. 1292–7.

[5] Bruce R. Levin. "Coexistence of Two Asexual Strains on a Single Resource". In: *Science* 175.4027 (1972), pp. 1272–1274. URL: http://www.jstor.org/stable/1733726.

[6] Rutger Hermsen et al. "A growth-rate composition formula for the growth of E. coli on co-utilized carbon substrates". In: *Molecular Systems Biology* 11.4 (2015), p. 801. DOI: https://doi.org/10.15252/msb.20145537. URL: https://doi.org/10.15252/msb.20145537.

[7] Xin Wang et al. "Growth strategy of microbes on mixed carbon sources". In: *Nature Communications* 10.1 (2019), p. 1279. DOI: 10.1038/s41467-019-09261-3. URL: https://doi.org/10.1038/s41467-019-09261-3.

[8] G. E. Hutchinson. "The Paradox of the Plankton". In: *The American Naturalist* 95.882 (1961), pp. 137–145. DOI: 10.1086/282171. URL: https://doi.org/10.1086/282171.

[9] S. C. Stearns. "Trade-Offs in Life-History Evolution". In: 3.3 (1989), pp. 259–268. DOI: 10.2307/2389364. URL: http://www.jstor.org.inee.bib.cnrs.fr/stable/2389364.

[10] Richard E. Michod. "On the transfer of fitness from the cell to the multicellular organism". In: *Biology and Philosophy* 20.5 (2005), pp. 967–987. DOI: 10.1007/s10539-005-9018-2. URL: https://doi.org/10.1007/s10539-005-9018-2.

[11] Hila Sheftel et al. "The geometry of the Pareto front in biological phenotype space". In: *Ecology and Evolution* 3.6 (2013), pp. 1471–1483. DOI: https://doi.org/10.1002/ece3.528. URL: https://doi.org/10.1002/ece3.528.

[12] Aleksej Zelezniak et al. "Metabolic dependencies drive species co-occurrence in diverse microbial communities". In: *Proceedings of the National Academy of Sciences* 112.20 (2015), pp. 6449–6454. DOI: 10.1073/pnas.1421834112. URL: https://doi.org/10.1073/pnas.1421834112.

[13] Fernando Baquero et al. "The Origin of Niches and Species in the Bacterial World". In: *Frontiers in microbiology* 12 (2021), pp. 657986–657986. DOI: 10.3389/fmicb.2021.657986. URL: https://pubmed.ncbi.nlm.nih.gov/33815348.

[14] Robert B. Helling, Christopher N. Vargas, and Julian Adams. "Evolution of Escherichia coli During Growth in a Constant Environment". In: *Genetics* 116.3 (1987), p. 349. URL: http://www.genetics.org/content/116/3/349.abstract.

[15] Richard E. Lenski et al. "Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations". In: 138.6 (1991), pp. 1315–1341. URL: http://www.jstor.org/stable/2462549.

[16] Daniel E. Rozen and Richard E. Lenski. "Long-Term Experimental Evolution in Escherichia coli. VIII. Dynamics of a Balanced Polymorphism." In: *The American Naturalist* 155.1 (2000), pp. 24–35. DOI: 10.1086/303299. URL: https://doi.org/10.1086/303299.

[17] Benjamin H. Good et al. "The dynamics of molecular evolution over 60,000 generations". In: *Nature* 551.7678 (2017), pp. 45–50. DOI: 10.1038/nature24287. URL: https://doi.org/10.1038/nature24287.

[18] D S Treves, S Manning, and J Adams. "Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of Escherichia coli." In: *Molecular Biology and Evolution* 15.7 (1998), pp. 789–797. DOI: 10.1093/oxfordjournals.molbev.a025984. URL: https://doi.org/10.1093/oxfordjournals.molbev.a025984.

[19] Richard E Lenski. "Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations". In: *The ISME Journal* 11.10 (2017), pp. 2181–2194. DOI: 10.1038/ismej.2017.69. URL: https://doi.org/10.1038/ismej.2017.69.

[20] Glen D'Souza et al. "Ecology and evolution of metabolic cross-feeding interactions in bacteria". In: *Natural Product Reports* 35.5 (2018), pp. 455–488. DOI: 10.1039/C8NP00009C. URL: http://dx.doi.org/10.1039/C8NP00009C.

[21] Nick W. Smith et al. "The Classification and Evolution of Bacterial Cross-Feeding". In: *Frontiers in Ecology and Evolution* 7 (2019), p. 153. ISSN: 2296-701X. DOI: 10.3389/fevo.2019.00153. URL: https://www.frontiersin.org/article/10.3389/fevo.2019.00153.

[22] Magdalena San Roman and Andreas Wagner. "An enormous potential for niche construction through bacterial cross-feeding in a homogeneous environment". In: *PLOS Computational Biology* 14.7 (2018), e1006340–. URL: https://doi.org/10.1371/journal.pcbi.1006340.

[23] Margie Kinnersley et al. "Ex Uno Plures: Clonal Reinforcement Drives Evolution of a Simple Microbial Community". In: *PLOS Genetics* 10.6 (2014), e1004430–. URL: https://doi.org/10.1371/journal.pgen.1004430.

[24] Tobias Großkopf et al. "Metabolic modelling in a dynamic evolutionary framework predicts adaptive diversification of bacteria in a long-term evolution experiment". In: *BMC Evolutionary Biology* 16.1 (2016), p. 163. DOI: 10.1186/s12862-016-0733-x. URL: https://doi.org/10.1186/s12862-016-0733-x.

[25] Magdalena San Roman and Andreas Wagner. "Acetate and glycerol are not uniquely suited for the evolution of cross-feeding in E. coli". In: *PLOS Computational Biology* 16.11 (2020), e1008433–. URL: https://doi.org/10.1371/journal.pcbi.1008433.

[26] Douwe Molenaar et al. "Shifts in growth strategies reflect tradeoffs in cellular economics." eng. In: *Mol Syst Biol* 5 (2009), p. 323. ISSN: 1744-4292 (Electronic); 1744-4292 (Linking). DOI: 10.1038/msb.2009.82.

[27] Robert E. Beardmore et al. "Metabolic trade-offs and the maintenance of the fittest and the flattest". In: *Nature* 472.7343 (2011), pp. 342–346. DOI: 10.1038/nature09905. URL: https://doi.org/10.1038/nature09905.

[28] Romain Gallet et al. "The evolution of bacterial cell size: the internal diffusion-constraint hypothesis". In: *The Isme Journal* 11 (2017), 1559 EP -. URL: http://dx.doi.org/10.1038/ismej.2017.35.

[29] Torsten Held, Daniel Klemmer, and Michael Lässig. "Survival of the simplest in microbial evolution". In: *Nature Communications* 10.1 (2019), p. 2472. DOI: 10.1038/s41467-019-10413-8. URL: https://doi.org/10.1038/s41467-019-10413-8.

[30] Meike T. Wortel et al. "Evolutionary pressures on microbial metabolic strategies in the chemostat". In: *Scientific Reports* 6.1 (2016), p. 29503. DOI: 10.1038/srep29503. URL: https://doi.org/10.1038/srep29503.

[31] Meike T. Wortel et al. "Metabolic enzyme cost explains variable trade-offs between microbial growth rate and yield". In: *PLOS Computational Biology* 14.2 (2018), e1006010–. URL: https://doi.org/10.1371/journal.pcbi.1006010.

[32] M. Doebeli. "A model for the evolutionary dynamics of cross-feeding polymorphisms in microorganisms". In: *Population Ecology* 44.2 (2002), pp. 59–70. DOI: 10.1007/s101440200008. URL: https://doi.org/10.1007/s101440200008.

[33] Thomas Pfeiffer, Sebastian Bonhoeffer, and Associate Editor: Peter J. Morin. "Evolution of Cross-Feeding in Microbial Populations". In: 163.6 (2004), E126–E135. DOI: 10.1086/383593. URL: http://www.jstor.org/stable/10.1086/383593.

[34] Florian Labourel and Etienne Rajon. "Resource uptake and the evolution of moderately efficient enzymes". In: *Molecular Biology and Evolution* (2021). DOI: 10.1093/molbev/msab132. URL: https://doi.org/10.1093/molbev/msab132.

[35] Andreas Wagner. "Energy Constraints on the Evolution of Gene Expression". In: *Molecular Biology and Evolution* 22.6 (2005), pp. 1365–1374. ISSN: 0737-4038. DOI: 10.1093/molbev/msi126. eprint: https://academic.oup.com/mbe/article-pdf/22/6/1365/3879110/msi126.pdf. URL: https://doi.org/10.1093/molbev/msi126.

[36] Michael Lynch and Georgi K Marinov. "The bioenergetic costs of a gene". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.51 (2015), pp. 15690–15695. DOI: 10.1073/pnas.1514974112. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4697398/.

[37] Moshe Kafri et al. "The Cost of Protein Production". In: *Cell Reports* 14.1 (2016), pp. 22–31. DOI: https://doi.org/10.1016/j.celrep.2015.12.015. URL: https://www.sciencedirect.com/science/article/pii/S221112471501428X.

[38] Ken A. Dill, Kingshuk Ghosh, and Jeremy D. Schmit. "Physical limits of cells and proteomes". In: *Proceedings of the National Academy of Sciences* 108.44 (2011), p. 17876. DOI: 10.1073/pnas.1114477108. URL: http://www.pnas.org/content/108/44/17876.abstract.

[39] Pablo M. Blanco et al. "Macromolecular diffusion in crowded media beyond the hard-sphere model". In: *Soft Matter* 14 (16 2018), pp. 3105–3114. DOI: 10.1039/C8SM00201K. URL: http://dx.doi.org/10.1039/C8SM00201K.

[40] R. Milo, R. Phillips, and N. Orme. "Cell Biology by the Numbers". In: (2016). Ed. by New York: Garland Science.

[41] Mario Orsi, Wendy E. Sanderson, and Jonathan W. Essex. "Permeability of Small Molecules through a Lipid Bilayer: A Multiscale Simulation Study". In: *The Journal of Physical Chemistry B* 113.35 (2009), pp. 12019–12029. DOI: 10.1021/jp903248s. URL: https://doi.org/10.1021/jp903248s.

[42] Stéphane Pinhal et al. "Acetate Metabolism and the Inhibition of Bacterial Growth by Acetate". In: *Journal of bacteriology* 201.13 (2019), e00147–19. DOI: 10.1128/JB.00147-19. URL: https://pubmed.ncbi.nlm.nih.gov/30988035.

[43] S. A. H. Geritz et al. "Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree". In: *Evolutionary Ecology* 12.1 (1998), pp. 35–57. DOI: 10.1023/A:1006554906681. URL: https://doi.org/10.1023/A:1006554906681.

[44] J. Hofbauer and K. Sigmund. "Adaptive dynamics and evolutionary stability". In: *Applied Mathematics Letters* 3.4 (1990), pp. 75–79. DOI: https://doi.org/10.1016/0893-9659(90)90051-C. URL: https://www.sciencedirect.com/science/article/pii/089396599090051C.

[45] Thomas Pfeiffer and Stefan Schuster. "Game-theoretical approaches to studying the evolution of biochemical systems". In: *Trends in Biochemical Sciences* 30.1 (2005), pp. 20–25. DOI: https://doi.org/10.1016/j.tibs.2004.11.006. URL: https://www.sciencedirect.com/science/article/pii/S0968000404002968.

[46] D L Hartl, D E Dykhuizen, and A M Dean. "Limits of adaptation: the evolution of selective neutrality". In: *Genetics* 111.3 (1985), pp. 655–674. URL: https://pubmed.ncbi.nlm.nih.gov/3932127.

[47] Steven S Andrews. "Effects of surfaces and macromolecular crowding on bimolecular reaction rates." eng. In: *Phys Biol* 17.4 (2020), p. 045001. ISSN: 1478-3975 (Electronic); 1478-3967 (Linking). DOI: 10.1088/1478-3975/ab7f51.

[48] Wolfram Liebermeister et al. "Visual account of protein investment in cellular functions". In: *Proceedings of the National Academy of Sciences* 111.23 (2014), p. 8488. DOI: 10.1073/pnas.1314810111. URL: http://www.pnas.org/content/111/23/8488.abstract.

[49] Arnošt Kotyk. "Mobility of the free and of the loaded monosaccharide carrier in Saccharomyces cerevisiae". In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 135.1 (1967), pp. 112–119. DOI: https://doi.org/10.1016/0005-2736(67)90013-2. URL: http://www.sciencedirect.com/science/article/pii/0005273667900132.

[50] G E Briggs and J B Haldane. "A Note on the Kinetics of Enzyme Action". In: *The Biochemical journal* 19.2 (1925), pp. 338–339. DOI: 10.1042/bj0190338. URL: https://pubmed.ncbi.nlm.nih.gov/16743508.

[51] Arren Bar-Even et al. "The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters". In: *Biochemistry* 50.21 (2011), pp. 4402–4410. DOI: 10.1021/bi2002289. URL: https://doi.org/10.1021/bi2002289.

[52] Michael Cox and David Nelson. *Lehninger Principles of Biochemistry*. Vol. 5. 2000. DOI: 10.1007/978-3-662-08289-8.

[53] Anthony J. Greenberg, Sarah R. Stockwell, and Andrew G. Clark. "Evolutionary Constraint and Adaptation in the Metabolic Network of Drosophila". In: *Molecular Biology and Evolution* 25.12 (2008), pp. 2537–2546. DOI: 10.1093/molbev/msn205. URL: https://doi.org/10.1093/molbev/msn205.

[54] Kevin M. Wright and Mark D. Rausher. "The Evolution of Control and Distribution of Adaptive Mutations in a Metabolic Pathway". In: *Genetics* 184.2 (2010), p. 483. DOI: 10.1534/genetics.109.110411. URL: http://www.genetics.org/content/184/2/483.abstract.

[55] Carrie F. Olson-Manning et al. "Evolution of Flux Control in the Glucosinolate Pathway in Arabidopsis thaliana". In: *Molecular Biology and Evolution* 30.1 (2012), pp. 14–23. DOI: 10.1093/molbev/mss204. URL: https://doi.org/10.1093/molbev/mss204.

[56] Xin Li et al. "A database of thermodynamic properties of the reactions of glycolysis, the tricarboxylic acid cycle, and the pentose phosphate pathway". In: *Database* 2011 (2011). DOI: 10.1093/database/bar005. URL: https://doi.org/10.1093/database/bar005.

[57] Nuno M. Oliveira, Rene Niehus, and Kevin R. Foster. "Evolutionary limits to cooperation in microbial communities". In: *Proceedings of the National Academy of Sciences* 111.50 (2014), p. 17941. DOI: 10.1073/pnas.1412673111. URL: http://www.pnas.org/content/111/50/17941.abstract.

[58] Jeroen Meijer, Bram van Dijk, and Paulien Hogeweg. "Contingent evolution of alternative metabolic network topologies determines whether cross-feeding evolves". In: *Communications Biology* 3.1 (2020), p. 401. DOI: 10.1038/s42003-020-1107-x. URL: https://doi.org/10.1038/s42003-020-1107-x.

[59] Brice Enjalbert et al. "Acetate fluxes in Escherichia coli are determined by the thermodynamic control of the Pta-AckA pathway". In: *Scientific Reports* 7.1 (2017), p. 42135. DOI: 10.1038/srep42135. URL: https://doi.org/10.1038/srep42135.

[60] Alan R. Pacheco, Mauricio Moel, and Daniel Segrè. "Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems". In: *Nature Communications* 10.1 (2019), p. 103. DOI: 10.1038/s41467-018-07946-9. URL: https://doi.org/10.1038/s41467-018-07946-9.

[61] Joana Sá-Pessoa et al. "SATP (YaaH), a succinate–acetate transporter protein in Escherichia coli". In: *Biochemical Journal* 454.3 (2013), pp. 585–595. DOI: 10.1042/BJ20130412. URL: https://doi.org/10.1042/BJ20130412.

[62] G N Vemuri et al. "Overflow metabolism in Escherichia coli during steady-state growth: transcriptional regulation and effect of the redox ratio." eng. In: *Appl Environ Microbiol* 72.5 (2006), pp. 3653–3661. ISSN: 0099-2240 (Print); 1098-5336 (Electronic); 0099-2240 (Linking). DOI: 10.1128/AEM.72.5.3653-3661.2006.

606 [63] Olga Ponomarova et al. "Yeast Creates a Niche for Symbiotic Lactic Acid Bacteria through
607 Nitrogen Overflow". In: *Cell Systems* 5.4 (2017), 345–357.e6. DOI: https://doi.org/10.1016/j.cels.
608 2017.09.002. URL: https://www.sciencedirect.com/science/article/pii/S2405471217303903.

609 [64] Karl Peebo et al. "Proteome reallocation in Escherichia coli with increasing specific growth
610 rate." eng. In: *Mol Biosyst* 11.4 (2015), pp. 1184–1193. ISSN: 1742-2051 (Electronic); 1742-2051
611 (Linking). DOI: 10.1039/c4mb00721b.

612 [65] Markus Basan et al. "Overflow metabolism in Escherichia coli results from efficient proteome
613 allocation". In: *Nature* 528.7580 (2015), pp. 99–104. DOI: 10.1038/nature15765. URL: https:
614 //pubmed.ncbi.nlm.nih.gov/26632588.

615 [66] Harry Kemble et al. "Flux, toxicity, and expression costs generate complex genetic interac-
616 tions in a metabolic pathway". In: *Science Advances* 6.23 (), eabb2236. DOI: 10.1126/sciadv.
617 abb2236. URL: https://doi.org/10.1126/sciadv.abb2236.

618 [67] Kai Zhuang, Goutham N Vemuri, and Radhakrishnan Mahadevan. "Economics of membrane
619 occupancy and respiro-fermentation". In: *Molecular Systems Biology* 7.1 (2011), p. 500. DOI:
620 https://doi.org/10.1038/msb.2011.34. URL: https://doi.org/10.1038/msb.2011.34.

621 [68] Mariola Szenk, Ken A. Dill, and Adam M. R. de Graff. "Why Do Fast-Growing Bacteria Enter
622 Overflow Metabolism? Testing the Membrane Real Estate Hypothesis". In: *Cell Systems* 5.2
623 (2017), pp. 95–104. DOI: https://doi.org/10.1016/j.cels.2017.06.005. URL: https://www.
624 sciencedirect.com/science/article/pii/S2405471217302338.

625 [69] Pierre Millard et al. "Control and regulation of acetate overflow in Escherichia coli". In: *eLife*
626 10 (2021), e63661. DOI: 10.7554/eLife.63661. URL: https://doi.org/10.7554/eLife.63661.

627 [70] Joy Bergelson et al. "Functional biology in its natural context: A search for emergent simplic-
628 ity". In: *eLife* 10 (2021), e67646. DOI: 10.7554/eLife.67646. URL: https://doi.org/10.7554/eLife.
629 67646.

630 [71] Joshua E. Goldford et al. "Emergent simplicity in microbial community assembly". In: *Science*
631 361.6401 (2018), p. 469. DOI: 10.1126/science.aat1168. URL: http://science.sciencemag.org/
632 content/361/6401/469.abstract.

633 [72] Djordje Bajic and Alvaro Sanchez. "The ecology and evolution of microbial metabolic strate-
634 gies". In: *Current Opinion in Biotechnology* 62 (2020), pp. 123–128. DOI: https://doi.org/10.1016/
635 j.copbio.2019.09.003. URL: https://www.sciencedirect.com/science/article/pii/S0958166919300722.

636 [73] Sylvie Estrela et al. "Metabolic rules of microbial community assembly". In: *bioRxiv* (2020),
637 p. 2020.03.09.984278. DOI: 10.1101/2020.03.09.984278. URL: http://biorxiv.org/content/early/
638 2020/03/11/2020.03.09.984278.abstract.

639 [74] David Heckmann, Daniel C. Zielinski, and Bernhard O. Palsson. "Modeling genome-wide en-
640 zyme evolution predicts strong epistasis underlying catalytic turnover rates". In: *Nature Com-*
641 *munications* 9.1 (2018), p. 5270. DOI: 10.1038/s41467-018-07649-1. URL: https://doi.org/10.
642 1038/s41467-018-07649-1.

643 [75] Sergey Kryazhimskiy. "Emergence and propagation of epistasis in metabolic networks". In:
644 *eLife* 10 (2021), e60200. DOI: 10.7554/eLife.60200. URL: https://doi.org/10.7554/eLife.60200.

645 [76] J. B. S. Haldane. *Enzymes*. London, New York: Longmans, Green, 1930, vii, 235 p. URL: //catalog.
646 hathitrust.org/Record/001580201.

647 [77] Berend Hasselman. *Solve Systems of Nonlinear Equations - package 'nleqslv'*. Tech. rep. Version
648 3.3.2. 2018.

649 [78] Alessio Peracchi. "The Limits of Enzyme Specificity and the Evolution of Metabolism". In:
650 *Trends in Biochemical Sciences* 43.12 (2018), pp. 984–996. DOI: https://doi.org/10.1016/j.
651 tibs.2018.09.015. URL: https://www.sciencedirect.com/science/article/pii/S0968000418302007.

[79] Thomas D. Niehaus and Katie B. Hillmann. "Enzyme promiscuity, metabolite damage, and metabolite damage control systems of the tricarboxylic acid cycle". In: *The FEBS Journal* 287.7 (2020), pp. 1343–1358. DOI: https://doi.org/10.1111/febs.15284. URL: https://doi.org/10.1111/febs.15284.

[80] R Heinrich, S Schuster, and H G Holzhütter. "Mathematical analysis of enzymic reaction systems using optimization principles." eng. In: *Eur J Biochem* 201.1 (1991), pp. 1–21. ISSN: 0014-2956 (Print); 0014-2956 (Linking). DOI: 10.1111/j.1432-1033.1991.tb16251.x.

[81] Ulf Dieckmann and Richard Law. "The dynamical theory of coevolution: a derivation from stochastic ecological processes". In: *Journal of Mathematical Biology* 34.5 (1996), pp. 579–612. DOI: 10.1007/BF02409751. URL: https://doi.org/10.1007/BF02409751.

[82] Ulf Dieckmann and Michael Doebeli. "On the origin of species by sympatric speciation". In: *Nature* 400.6742 (1999), pp. 354–357. DOI: 10.1038/22521. URL: https://doi.org/10.1038/22521.