Functional universality in slow-growing microbial communities arises from thermodynamic constraints

Ashish B. George[1,2], Tong Wang[1,3,4], and Sergei Maslov[1,3,5*]

[1] *Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.*
[2] *Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.*
[3] *Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.*
[4] *Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA.*
[5] *Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.*

* maslov@illinois.edu

(Dated: March 17, 2022)

The dynamics of microbial communities is incredibly complex, determined by competition for metabolic substrates and cross-feeding of byproducts. Species in the community grow by harvesting energy from chemical reactions that transform substrates to products. In many anoxic environments, these reactions are close to thermodynamic equilibrium and growth is slow. To understand the community structure in these energy-limited environments, we developed a microbial community consumer-resource model incorporating energetic and thermodynamic constraints on an interconnected metabolic network. The central ingredient of the model is product inhibition, meaning that microbial growth may be limited not only by depletion of metabolic substrates but also by accumulation of products. We demonstrate that these additional constraints on microbial growth cause a convergence in the structure and function of the community metabolic network—independent of species composition and biochemical details—providing a possible explanation for convergence of community function despite taxonomic variation observed in many natural and industrial environments. Furthermore, we discovered that the structure of community metabolic network is governed by the thermodynamic principle of maximum heat dissipation. Overall, the work demonstrates how universal thermodynamic principles may constrain community metabolism and explain observed functional convergence in microbial communities.

INTRODUCTION

Over half of earth's prokaryotic biomass resides in ocean sediments and deep soil[1,2]. In many of these natural environments and anaerobic bioreactors, communities grow under energy-limited conditions due to the absence of strong electron acceptors, like oxygen or nitrate, and strong electron donors, like glucose[3–8]. Hence species in these communities are forced to harvest energy required for biomass growth from low free energy reactions[8–13]. Growth in these environments can be very slow—division times are measured in weeks and years—instead of minutes and hours used for growth in the lab[4,13–16]. These long timescales and difficulty in culturing these species in a lab environment make experimental investigation of these communities challenging. Models and theory of energy-limited microbial communities can help uncover their organizational principles and shared emergent properties and drive understanding.

Models and theories inspired by MacArthur's consumer-resource model have helped understand many of the principles and properties governing microbial communities structured by nutrient-limitation[17–23]. However, models of communities in low-energy environments, structured by energy-limitation, need to incorporate important additional details. First, microbial growth in low-energy environments is determined by energy assimilation rather than nutrient uptake. Second, the low free energy reactions utilized to drive growth are subject to thermodynamic constraints. And third, reactions need to be explicitly modeled to track their thermodynamic feasibility, rather than just species preference for resources. Here we propose and study a minimal model of microbial communities that accounts for these crucial details to understand the emergent properties of slow-growing communities in energy-limited environments.

Recently, a few studies have attempted to incorporate thermodynamics into community models[24,25]. These studies focused on community diversity, showing that thermodynamics enables communities to overcome the competitive exclusion principle and stabilize high diversity on a single substrate. In contrast, here we focus on the emergent properties of the communities in low-energy environments, such as the metabolic network structure and function of the community.

Simulations of our model generated communities that shared many features of natural communities. Strikingly, communities assembled from separate species pools in replicate environments realized the same metabolic network and performed similar metabolic functions at the community-level despite stark differences at the species level. This resembles the observed convergence in community function despite taxonomic divergence, seen in anaerobic bioreactors, oceans, and other environments[26–29].

Analyzing the model to understand how functional

convergence arises, we discovered that thermodynamic principles govern community-level metabolic network structure and function in slow-growing, energy-limited communities. We derived the thermodynamic principle of maximum dissipation, which, we show, determines the community metabolic network selected by ecological competition in energy-limited communities. This thermodynamic principle, along with physical constraints, shape the metabolic environment created by the community and the metabolic functions it performs. Further, the derivation of the maximum dissipation principle provides a concrete example in which communities are structured by thermodynamic optimization, an idea that has been conjectured in ecology but never proven[30].

Overall, our results highlight the crucial role of thermodynamic constraints in defining community metabolism and function of energy-limited microbial communities. The results provide an explanation for the observed convergence in microbial community function, serve as a concrete example of thermodynamic optimization in community ecology, and make predictions for the community metabolic network and resource environment.

RESULTS

Model of microbial communities in energy-poor environments

To understand slow-growing microbial communities in energy-limited environments, we develop a model that incorporates crucial features of growth in these environments: energetic-limitation of growth, thermodynamic constraints on reactions, and explicit modeling of reaction fluxes.

In the model, species grow utilizing energy harvested from chemical reactions that convert a higher energy resource (substrate) to a lower energy resource (product) (Fig. 1A). A reaction can be understood as a coarse-grained description of catabolism[31,32]. The resources are labeled in the order of decreasing standard state energy as $R_0, R_1, R_2, \ldots$. A species can utilize reactions connecting any pair of resources; the allowed reactions make a directed, fully-connected network. The product of one reaction can act as the substrate of another, facilitating cross-feeding interactions between microbes.

Species utilize a reaction by producing the corresponding, reaction-specific enzyme. A species allocates a portion of its total enzyme budget to any number of the possible reactions. We allow total enzyme budgets to vary, as in real species[33]; this also avoids potential degenerate behavior seen in community models with fixed enzyme budgets[18]. Thus a species is characterised by its enzyme allocation strategy and total enzyme budget (Fig. 1B). Through the explicit connection of species to reactions, substrate consumption can create different products depending on the enzyme used, like in real microbes[34]. This contrasts with microbial consumer-resource models, where substrate consumption by a species leads to a fixed mixture of products, i.e., all species use the same reactions and differ only in their consumption preferences[21,35,36].

From a reaction converting substrate S to product P with standard-state energies $\mathcal{E}_S^0$ and $\mathcal{E}_P^0$, a species assimilates $\mathcal{E}^{\mathrm{ATP}}$ energy. This energy, stored in ATP or other forms, is constrained by reaction stoichiometry, and so is assumed to be species-independent. The harvested energy drives the proportional growth of new biomass. The remaining energy is dissipated as heat Q_{SP}. Q_{SP} is is defined as

$$Q_{SP} = \mathcal{E}_S^0 - \mathcal{E}^{\mathrm{ATP}} - \mathcal{E}_P^0, \tag{1}$$

which can also be understood as the negative free energy difference under standard conditions, ΔG^0.

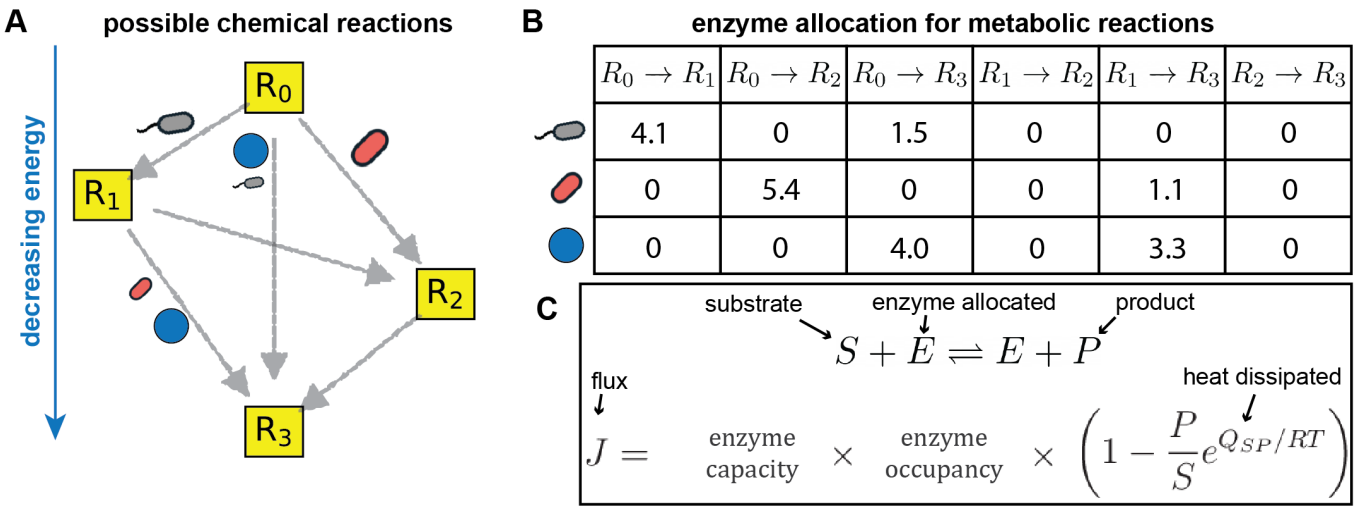
The flux through the reaction is determined by reversible Michaelis-Menten kinetics[37] :

$$J_{S\to P} = k_{\mathrm{cat}} E \frac{\frac{S}{K_S}}{1+\frac{S}{K_S}+\frac{P}{K_P}} \left(1 - \frac{P}{S}\exp\left[-\frac{Q_{SP}}{RT}\right]\right), \tag{2}$$

where E is the enzyme concentration, $k_{\mathrm{cat}}, K_S, K_P$ are enzyme-specific parameters describing the chemical kinetics, T is the temperature, and R is the gas constant. In thermodynamic parlance, $\frac{P}{S}\exp\left[-\frac{Q_{SP}}{RT}\right]$ represents the free energy difference after accounting for resource concentration gradient, $\Delta G' = \Delta G^0 + RT\log\frac{P}{S}$. Note that our results will apply to a general stoichiometric ratio of substrate and product beyond $1:1$ as well (see SI); we omit discussion of this scenario in the main text for clarity.

Eq. 2 can be understood as a product of three terms[38]. The first term specifies the maximal reaction rate, which is proportional to amount of enzyme allocated to the reaction, E. The second term describes the saturation level of the enzyme; the flux increases linearly at low substrate concentrations and saturates at high substrate concentrations. The last (and most crucial for our purposes) term describes thermodynamic inhibition, i.e., how the net rate of a reversible reaction decreases as products accumulate and the ratio P/S increases. The strength of the thermodynamic inhibition is controlled by the heat dissipated in the reaction. This thermodynamic inhibition of microbial growth has been quantitatively validated in the lab and in situ conditions[6,39–41]. If product accumulation is severe enough to reverse the reaction direction, we assume that species down-regulate enzyme production to prevent energy loss from catalyzing the reverse reaction.

The dynamics of species and resources are driven by these reaction fluxes. Species grow in proportion to the energy they assimilate from the reactions catalyzed by their enzymes. The concentration of a resource decreases due to consumption in a reaction and increases due to production in a reaction. Additionally, the entire system is diluted at rate δ, which represents the dilution rate of the bioreactor or the dilution by sedimentation



	$R_0 \to R_1$	$R_0 \to R_2$	$R_0 \to R_3$	$R_1 \to R_2$	$R_1 \to R_3$	$R_2 \to R_3$
	4.1	0	1.5	0	0	0
	0	5.4	0	0	1.1	0
	0	0	4.0	0	3.3	0

FIG. 1. **Model of communities in energy-limited environments. A)** Species harvest energy from reactions (grey arrows) that consume higher energy resource (substrate) to produce a lower energy resource (product). Reactions connecting any pair of the four resources R_0, R_1, R_2, R_3 can be utilized by a species, making the space of all possible reactions a fully connected network. **B)** A species can utilize a reaction by allocating a part of its overall enzyme budget to the corresponding, reaction-specific enzyme. Species differ in both the manner of enzyme allocation as well as their overall enzyme budgets. **C)** The reactions in low-energy environments are thermodynamically reversible. The flux J in a reversible, enzyme-catalyzed reaction between substrate S and product P is determined by three factors: 1) Enzyme capacity, which is the amount of enzyme available. 2) Enzyme occupancy, which measure the fraction of free enzyme available to catalyze the reaction, determined by the concentration of substrates and products. 3) Thermodynamic inhibition due to accumulation of products. Since low-energy reactions are reversible, the net flux in the forward direction will decrease as products accumulate. The strength of this inhibition is controlled by the heat dissipated in the reaction, Q.

in oceans[42]. This dilution of species biomass can also be extended to account for the maintenance energy requirements of a species (see SI). The precise equations describing species and resource dynamics are detailed in Methods. We also obtained similar results in a model without explicit dilution of resources and only species maintenance costs(see SI); we omit discussion of this model in the main text for clarity.

Communities in natural and industrial environments are the result of years of ecological competition and succession. To model such communities that emerge from ecological competition, we studied community assembly from a large and diverse species pool in an environment consisting of 6 resources, labelled in the order of decreasing standard-state energies as $R_0, R_1 ..., R_5$. The environment was supplied with the most energy-rich resource R_0 at high concentrations compared to the typical K_S. The other lower energy resources were generated as products in some of the 15 reactions that could be utilized by species for growth. The energy assimilated from a reaction was chosen to be a random fraction (between 15% and 85%) of the energy gap between product and substrate of that reaction. Enzymatic parameters describing reaction kinetics, were chosen from lognormal distributions, motivated by empirical observations[43]. Since we are interested in slow-growing communities, we chose a small dilution rate so as to not drive any slow-growing species to extinction. After introducing species from the pool into the environment, many species went extinct from competition and simulations settled down to a steady-state community comprised of the survivors.

Emergent metabolic structure and function in slow-growing, energy-limited communities

To study the features shared across slow-growing communities in energy-limited environments, we repeated the community assembly experiment from five separate pools of 600 species each. Each species allocated its enzyme budget, chosen from a lognormal distribution, to a random subset of the 15 possible reactions. Species were not shared across pools. We also varied enzymatic parameters (chemical kinetic parameters) to vary between pools. The environment and its energetics was shared across the experiments, i.e., the resource supply, resource energies, and energy uptake in each reaction (which is constrained by stoichiometry) was the same across experiments. Simulations of all pools converged to a steady-state maintained by a community of the surviving species.

We analyzed the final steady-state communities to find emergent properties of energy-limited communities that were shared across pools that differed in species and enzyme content. We identified three properties of the communities that resemble observations in natural communities:

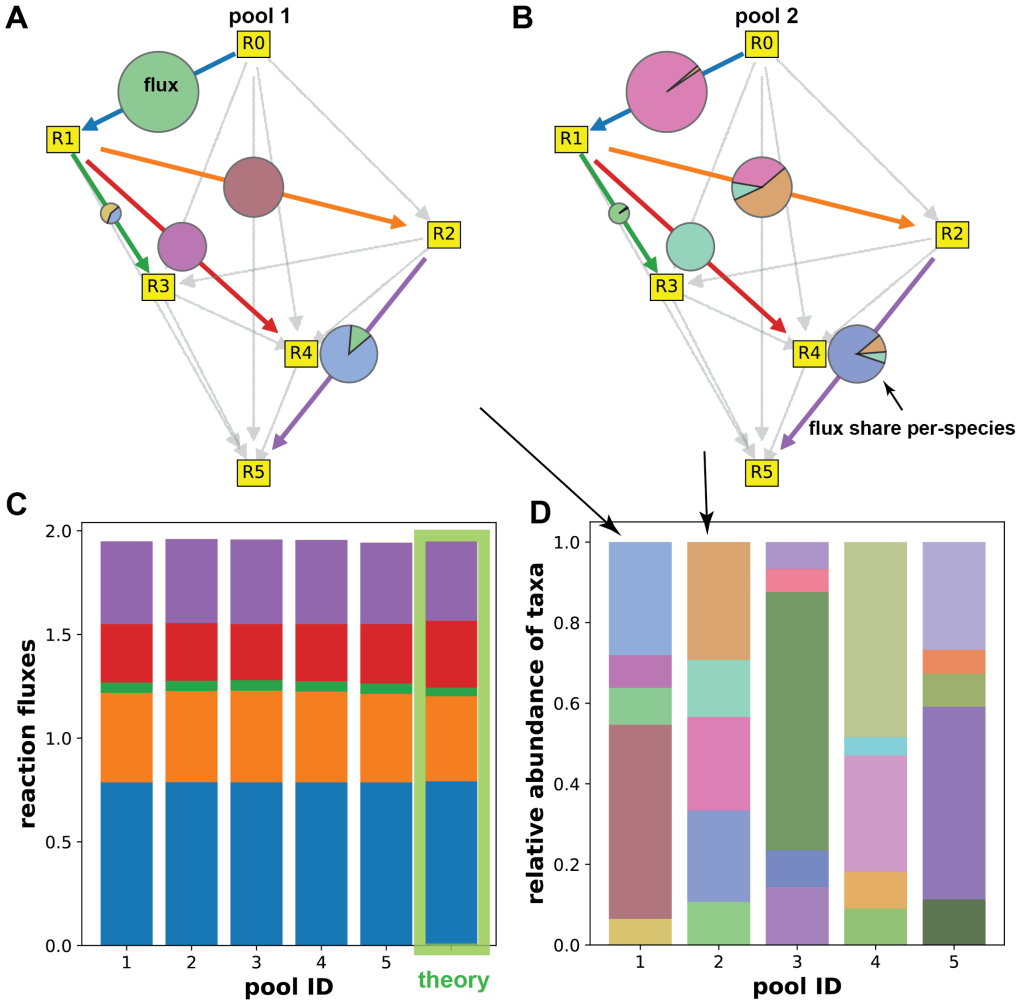


FIG. 2. **Functional convergence despite taxonomic variation.** We simulated the model to obtain the final community that arose as the outcome of competition between a large number of species. We repeated this community assembly experiment from five separate pools of 600 species each. **A,B)** The metabolic networks in the final communities obtained from two separate species pools are similar. From all possible reactions (grey arrows), the same subset of reactions are active (i.e., carry a positive flux) in the final communities (colored arrows). On each reaction arrow, the size of the circles show the flux carried in the reaction and the pie-chart on shows how this flux is shared between species. The manner in which species contribute to the flux differs between pools. **C)** The flux through the community metabolic network is similar across communities from different species pools. The theoretical prediction (Eq. (6)) matches simulations. The reactions in the bar plot are colored according to the arrow colors in panels A,B (pools 1 and 2). **D)** The relative abundance of species in the final communities. The species in the bar plot colors are colored according to the pie-charts in panels A,B (pools 1 and 2). The relative abundance is not proportional to the flux catalyzed by the species because the energy-yield per unit flux differs between reactions. For example, the green species from pool 1 (panel A) and the pink species from pool2 (panel B) catalyze a large flux but have a low relative abundance due to the low energy yield of the reaction (panel C).

Concurrent community metabolic networks: Only a subset of the 15 possible reactions were active (i.e., carried nonzero flux) in the final communities (colored arrows in Fig. 2A). This subset of reactions was shared across communities, despite starting from different species pools (Fig. 2A,B). Hence, the reactions utilized by the surviving species were determined by the environment. Furthermore, the metabolic network structure was constrained. The in-degree of each node was 1, i.e., each resource was produced by only a single reaction.

Functional convergence: We compared the metabolic functioning of the communities obtained from the different species pools. The function performed by the community was defined as the flux through the active reactions in the final community. These fluxes were strikingly similar across pools (radius of circles in Fig. 2A,B and Fig. 2C). Thus communities converged in the metabolic function they performed despite having been assembled from different species pools.

Taxonomic divergence: Bolstered by the convergence in metabolic network structure and function across pools, we measured the species abundances and func-

tional roles in the communities. The functional role of a species was quantified as the reaction flux catalyzed by it. We found that species contributed to the reaction fluxes idiosyncratically across pools (pie charts in Fig. 2A,B). Some reactions were performed by multiple species while others were performed by only one, and this breakup of reaction fluxes between the species varied between pools. The relative abundance of the species is determined by product of the metabolic flux catalyzed by a species and the energy assimilated per unit flux, which differed between reactions. The relative abundance is shown in Fig. 2D. Thus the relative abundance and functional roles of the species did not converge across the different pools.

Taken together, these observations mirror the 'functional convergence despite taxonomic divergence' observed in many natural communities[26–28]. We investigated the model analytically to understand why these slow-growing, energy-limited communities share these emergent properties. We uncovered fundamental thermodynamic principles governing functional convergence in these energy-limited communities, which we explain in the next section.

Principle of maximum heat dissipation explains selected community metabolic network

We first sought to understand why the same metabolic reactions are realized in the final community across species pools. We were able to make analytical progress by focusing our attention on slow-growing communities growing on low-energy reactions. At steady-state, the biomass growth in communities replaces loss by dilution, δ. In slow-growing communities, the growth at steady-state is much smaller than the maximum growth rate of the species $g_{\max}$. Motivated by conditions in ocean sediments and bioreactors, the substrate concentration in the resource supply in simulations was larger than the typical K_S[3,26,42,44]. The high substrate levels in the supply mean that growth was primarily slowed by the accumulation of products.

Since growth is slowed down by product accumulation, we can relate the concentrations of substrates and products of an active reaction in the final steady-state community. Fig. 3A shows an example of a reaction $R_0 \to R_1$ in a community with four resources. If the reaction is active in the final steady-state community, then we have:

$$\frac{R_1^*}{R_0^*} = e^{Q_{01}/RT} + \mathcal{O}\left(\frac{\delta}{g_{\max}}\right), \tag{3}$$

where R_1^*, R_0^* are the steady state concentrations of the resources and Q_{01} is the heat dissipated in the reaction $R_0 \to R_1$. This condition implies that reactions are slowed down to near thermodynamic equilibrium due to product accumulation. The first (and leading order) term depends only on the thermodynamics of the reaction. The second term, $\frac{\delta}{g_{\max}}$, which contains all the prop-

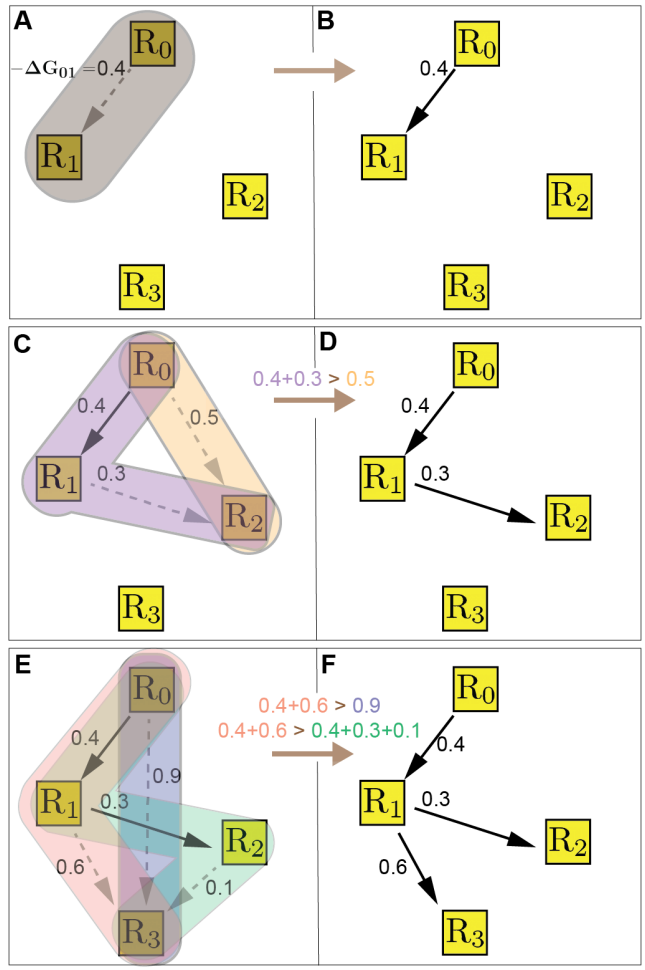


FIG. 3. **The maximum dissipation principle determines the community metabolic network.** In a system with four resources, we assemble the community metabolic network by considering the production of each resource sequentially to understand the governing thermodynamic principles. **A)** The sole path from the supplied resource, R_0, to the next-highest energy resource R_1. The path dissipates $Q_{01} = 0.4$ units of energy as heat. **B)** The steady-state community with reaction $R_0 \to R_1$ maintains resource concentrations such that $\frac{R_1^*}{R_0^*} \approx \exp(0.4/RT)$. **C)** The two possible paths from R_0 to the next resource, R_2, that could be added to the community. The paths $R_0 \to R_2$ and $R_0 \to R_1 \to R_2$ dissipate 0.5 and 0.7 units of energy as heat. **D)** The path that dissipates more heat, $R_0 \to R_1 \to R_2$, is selected in the final community because species utilizing this path will drive concentration of R_2 higher (such that $\frac{R_2^*}{R_0^*} \approx \exp(0.7/RT)$) than species utilizing $R_0 \to R_2$ **E)** The three possible paths from R_0 to R_3, that could be added to the community. We do not consider the path $R_0 \to R_2 \to R_3$ because we have already found that the maximally dissipative path to R_2 is $R_0 \to R_1 \to R_2$ and not $R_0 \to R_2$. **F)**The path that dissipates most heat, $R_0 \to R_1 \to R_3$, is selected in the final community because it can drive the concentration gradient R_3/R_0 highest (such that $\frac{R_3^*}{R_0^*} \approx \exp(1.0/RT)$). Thus, the active reactions in the final community network will be determined by the paths to each resource that dissipates the most heat.

erties specific to the species and enzyme, is only a small correction.

We now consider the production of the second resource, R_2, which can be produced via two possible reaction paths (Fig. 3C). To analyze this scenario, we derive a crucial extension to Eq.(3) that applies to reaction paths, composed of multiple reactions, that are active in the final community (see SI for derivation). If a reaction path proceeding from R_i to R_j is active in the final community, we have:

$$\frac{R_j^*}{R_i^*} = e^{Q_{i\Rightarrow j}/RT} + \mathcal{O}\left(\frac{\delta}{g_{\max}}\right), \tag{4}$$

where $Q_{i\Rightarrow j}$ refers to the sum of the heat dissipated in the reactions along the path from R_i to R_j. Again, note that the species- and enzyme- specific properties are only a small correction. This equation provides us different values for $\frac{R_2^*}{R_0^*}$ for each of the paths considered in Fig. 3C. Since both values cannot hold simultaneously, only one of the two paths can be realized in the final community. Thus the in-degree of any node in the network is always 1, as seen in Fig. 2A,B.

To understand which of the two paths is selected in the final community, we can consider the competition of two species utilizing reactions $R_1 \to R_2$ and $R_0 \to R_2$ added to the community with reaction $R_0 \to R_1$ shown in Fig. 3B. The species utilizing the reaction $R_1 \to R_2$ (and hence path $R_0 \to R_1 \to R_2$) will drive product accumulation until $\frac{R_2^*}{R_0^*} \approx e^{0.7/RT}$, while the species utilizing the reaction $R_0 \to R_2$ will drive product accumulation until $\frac{R_2^*}{R_0^*} \approx e^{0.5/RT}$. Since $e^{0.7/RT} > e^{0.5/RT}$, the former species can invade and grow in the steady-state environment of the latter, driving the product accumulation to levels where the latter cannot grow. Thus the species utilizing the path dissipating more heat is able to invade and displace its competitor, and create conditions that cannot be invaded by the other. Hence, the path to R_2 that dissipates more heat is realized in the final community (Fig. 3D).

The same argument can be extend to resources with multiple paths leading to them, such as R_3. We consider three paths to R_3 in Fig. 3E. Note that we exclude the path $R_0 \to R_2 \to R_3$ since we previously found the maximally dissipative path to R_2; this is an application of Prim's algorithm to find the maximal spanning tree in a network[45]. Comparing the heat dissipated among the three paths, we find the maximally dissipative path, $R_0 \to R_1 \to R_3$. A species utilizing this path is able to raise the concentration of R_3 high enough so as to make the other reactions infeasible, driving competitors relying on the paths extinct. Thus the principle of maximum dissipation governs the selection of the community metabolic network in slow-growing, energy-limited communities.

Functional convergence from thermodynamic principles

The principle of maximum heat dissipation explains why the same reactions are realized in the final communities across pools. But it does not explain the convergence in reaction fluxes across pools.

To understand how this convergence in fluxes arises, we build on results in Eq. (4) to derive the concentration at steady state of R_0,

$$R_0^* = \frac{h_0}{1 + \sum_{j=1}^{M} e^{Q_{0\Rightarrow j}/RT}} + \mathcal{O}\left(\frac{\delta}{g_{\max}}\right), \tag{5}$$

where h_0 is the steady state concentration of R_0 in absence of any consumption and $Q_{0\Rightarrow j}$ is the heat dissipated in the reaction path to R_j realized in the community (see SI). Using Eq. (4), we can obtain the the steady-state concentrations of all resources:

$$R_i^* = \frac{h_0}{1 + \sum_{j=1}^{M} e^{Q_{0\Rightarrow j}/RT}} e^{Q_{0\Rightarrow i}/RT} + \mathcal{O}\left(\frac{\delta}{g_{\max}}\right). \tag{6}$$

These equations imply that the resource concentrations at steady-state are determined to leading order by thermodynamics alone. The surviving species grow to abundances that allow them to maintain the environmental resource concentrations at these thermodynamically-determined values. Note that these equations resemble the Boltzmann-Gibbs distribution in statistical physics, now realized in an ecological context.

The fluxes in the steady-state community can be calculated by conserving resource flux along the chemical reaction network. The outflow from a terminal resource without any outgoing reactions, R_T, is solely determined by dilution and equal to δR_T^*. At steady state, this outflow is balanced by the reaction flux into the terminal resource. This can be used to calculate the flow from resources producing a terminal resource. Iterating this procedure, we can calculate the total flux into any resource R_i, I_i, as

$$I_i^* = R_i^*\delta + \sum_{j \text{ downstream of } i} R_j^*\delta + \mathcal{O}\left(\frac{\delta}{g_{\max}}\right), \tag{7}$$

where the sum is over all resources, R_j, that appear downstream of R_i in the reaction network. This theoretical prediction matches simulations (Fig. 2 A). Importantly, the steady-state resource concentrations and reaction fluxes are independent of enzyme budgets, enzyme allocation, and details of the chemical kinetics—none of the associated parameters (E, k_{cat}, K_M, K_S) appear in the equations to leading order (Eq.6). Hence there is strong selection for choosing the maximally dissipative network and converging to the thermodynamically constrained fluxes; selection on the species-specific properties is weaker. Thus, we observe functional convergence, driven by the shared reaction thermodynamics across experiments, despite taxonomic divergence, in terms of how species conspire to maintain the community function.

Note that a community utilizing the maximally dissipative reactions can still experience species turnover. For example, an invading species that allocates more enzyme to the maximally dissipative reactions can displace a resident. However, this invasion will have only a very small effect (of order $\frac{\delta}{g_{\max}}$) on the community metabolic function and steady-state resource environment. Hence communities can continuously experience species turnover while maintaining their metabolic functions, as seen in bioreactors[26,29].

DISCUSSION

Microbial community models have successfult exposed the organizational principles and universal, emergent behavior of microbial communities structured by nutrient limitation[17–23]. However, in many natural and industrial environments, microbial communities grow under energy-limited conditions and are forced to use low free energy reactions to drive their growth.[3–13]. Here, by explicitly modeling the thermodynamic inhibition of microbial catabolism and energy-limitation of growth, we have developed a minimal model of microbial communities in low-energy environments. We used the model to investigate the emergent properties of slow-growing microbial communities assembled in low-energy environments. Strikingly, model simulations displayed the 'functional convergence despite taxonomic divergence' observed in many communities[26–28]. We demonstrated that functional convergence in our model originates from the fundamental thermodynamic constraints on microbial growth; previous approaches have shown functional convergence in models of nutrient-limited communities due to shared metabolic constraints between species[36]. Further, we demonstrated that community metabolic network structure in slow-growing communities assembled by ecological competition over long periods of time is determined by the thermodynamic principle of maximum dissipation. Importantly, network selection by maximum dissipation manifests through ecological competition and is not expected in isolated microbes.

For simplicity, we demonstrated our main results using a minimal model of energy-limited communities. In the supplementary text, we show that our results hold more generally. In particular, we demonstrate that our results apply in the following scenarios: First, if reaction stoichiometry varies beyond 1 : 1 in the model (see SI Sec.6). Second, if maintenance energy requirements of the cell are incorporated in the model; the maintenance energy flux requirement resembles biomass dilution, and so our results apply as long as this requirement is small (see SI Sec.5). And third, a model without explicit dilution of resources, where cells uptake resources for anabolic processes in addition to catabolism (see SI Sec.7).

An important assumption in our theoretical derivation of the maximal dissipation principle and thermodynamic explanation of functional convergence was that communities were slow-growing. But what happens when growth rate increases? Our theoretical analysis predicts that differences in species' enzyme budgets and allocation, and in enzymatic parameters, will start to matter as growth rate increase, and will cause greater differences in the community metabolic network structure and function of communities from different pools. To study community metabolic network structure and function as the steady-state growth rate increases, we simulated community assembly from 10 separate pools of 600 species at different dilution rates δ. Since the growth rate of a community at steady-state is equal to the dilution rate, these simulations provided us with 10 different final communities at each steady-state growth rate.

Fig. 4A,B,C depicts three final communities obtained, at three different dilution rates from the same species pool. The maximum dissipation principle predicts all 5 reactions correctly in the slow-growing community, which grows at 1% of the maximum possible growth rate (dashed lines indicate reactions predicted correctly). At an intermediate growth rate, the maximum dissipation principle predicted 3 out of the 5 reactions correctly. Finally, for the fast-growing community, which grew at 30% of the maximum possible growth rate, 2 out of 5 reactions were predicted correctly. The accuracy of the predictions by the maximum dissipation principle, defined as the average fraction of reactions predicted correctly across the 10 species pools, is shown in Fig. 4D. Communities deviate from the predictions at higher dilution rates due to a combination of factors. First, due to the low energy yields of reactions on the maximally dissipative path, species utilizing these reactions need to catalyze a larger flux to achieve high growth rates than species utilizing reactions with higher energy yields on competing paths. Second, due to the variation in enzyme budgets and enzyme allocation strategies, a species utilizing reactions on a maximally dissipative path can get out-competed by a species with a larger enzyme budget allocated to reactions on a competing path. Third, due to the variation in enzymatic parameters, a reaction on a competing path might be preferred because it has better enzymes, (e.g. with higher k_{cat}).

As the idiosyncratic properties of species pools and enzymes start to play a bigger role in determining the community metabolic network, functional convergence weakens. Fig. 4E shows that the average distance between community fluxes, or functional divergence, increases in fast-growing communities. The dashed line shows the expected distance between two random flux vectors; this distance is higher than the most functionally distant communities because the random flux vectors are not subject to ecological and flux constraints. Together, the observations in Fig. 4 show that the maximum dissipation principle and functional convergence will be stronger in low-energy communities that grow slowly. Intriguingly some published results hint at this loss of functional convergence with faster growth. The functional stability of communities in anaerobic bioreactors was analyzed in three

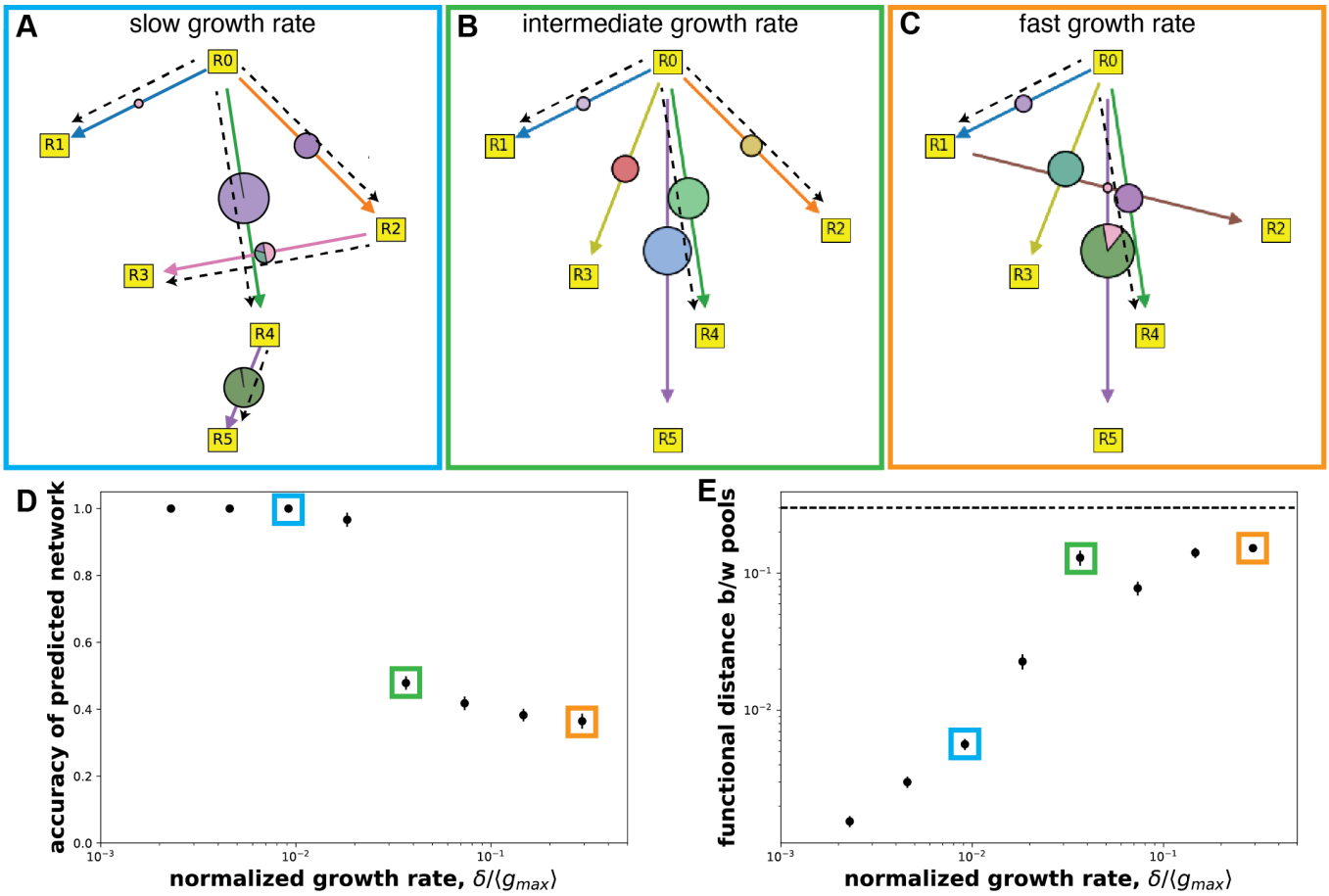


FIG. 4. **Functional convergence and the strength of thermodynamic constraints decreases in fast growing communities. A),B),C)** The realized community metabolic networks at slow growth rate, intermediate growth rate, and fast growth rate respectively. All three communities were obtained from the same species pool, and the steady-state growth rate was varied by changing the dilution rate, δ. The dashed lines indicate that the active reaction in the community agrees with the prediction based on maximum dissipation principle. In the slow-growing community, the active reactions match the five reactions predicted from the maximum dissipation principle; in the middle community 3 of the active reactions match the prediction; in the fast-growing community only 2 of the active reactions match the prediction. **D)** The fraction of reactions correctly predicted by the maximum dissipation principle, averaged across 10 different species pools in the same environment. The accuracy decreases as the growth rate increases. **E)** The functional distance between two communities was quantified by the Jensen-Shannon distance between the two community flux vectors. The average functional distance increases with the growth rate. Thus, the observed functional convergence decreases as the communities grow faster. The dashed line indicates the average distance between two random flux vectors. A random flux vector had the same length (15), and had each element picked from a uniform distribution between 0 and 1. The colored boxes in panels D and E correspond to the growth rates shown in the panels in A,B,C. The maximum growth rate, g_{max}, was used to normalize the steady-state growth rate. g_{max} was defined as the growth rate obtained on an irreversible reaction between consecutive resources, without any energy lost as heat, by a species investing its entire budget in the reaction (see Methods).

different studies[26,29,46], operating at dilution times of 10 days, 15 days and 1 day. While the two studies at the smaller dilution rates observed functional convergence, the study at the highest dilution rate did not.

Microbes in low-energy environments face a trade-off between harvesting energy efficiently, to increase biomass yield per unit reaction flux, and dissipating more heat, to increase reaction flux. A microbe cannot grow at either extreme—assimilating all the energy curbs growth because the reaction flux will be zero, while dissipating all the energy curbs growth because energy assimilated will be zero. Thus, microbes have to find a balance between these opposing forces. Our results suggest that, at small dilution rates, the effect of ecological competition will push microbes to evolve towards dissipating more energy. However, in practice this evolutionary drive to decrease energy assimilation will not continue forever. It will probably be limited by biochemical constraints[47–50], such as the smallest energy unit that can be assimilated by a species (one electron, one ATP, etc.), or when reactions become limited by substrate depletion rather than thermodynamic inhibition. In our simulations, to focus

on the ecological properties of energy-limited communities, a species at the lower-limit of energy assimilation could survive at the smallest dilution rates considered. An exploration of the evolutionary consequences of the selective pressures on energy assimilation is reserved for future work.

Optimization principles grounded in thermodynamics have been used in microbial metabolism to explain behavior at the cellular scale[51–53]. At the community level, optimization of thermodynamic quantities has been conjectured to be a mechanism by which communities self-organize, motivated mostly by verbal arguments and analogies with physical systems[54,55]. However, the question of which quantity should be maximized, and under what conditions, remains actively disputed in the literature[30,56]. By incorporating ecological and thermodynamic principles, our model connects these two disparate approaches toward understanding microbial communities. We derive the thermodynamic optimization principle of maximum dissipation, which structures slow-growing, energy-limited communities, providing a concrete example of thermodynamic optimization in ecology.

Our results make predictions for the metabolic structure and resource environment created by the communities in low-energy environments, based on the thermodynamics of microbial metabolism. This requires knowledge of the free energy differences and metabolomic data of the resource environment. Although calculating the free energies of the many metabolites present in a real environment is challenging, recent advances in computational methods make the outlook promising[57–59]. Combining these methods with metabolomic data and knowledge of microbial metabolism to predict metabolite composition in real environments shaped by microbial communities is an exciting prospect for future research.

Acknowledgements. The authors thank Avi Flamholz for valuable comments and feedback on the manuscript. This research was supported in part by NSF Grant No. PHY-1748958 and the Gordon and Betty Moore Foundation Grant No. 2919.02.

METHODS

Species and resource dynamics

Species dynamics are determined the amount of reaction flux they can utilize, the energy the assimilate from this to create biomass, and biomass loss due to dilution. The dynamics of species α with abundance $N^{(\alpha)}$ is described by

$$\frac{dN_\alpha}{dt} = N_\alpha \left[\sum_{i,j\in\mathcal{X}} E_{ij}^{(\alpha)} \mathcal{F}_{ij} \mathcal{E}_{ij}^{\mathrm{ATP}} Y - \delta \right], \tag{8}$$

where F_{ij} is the flux of reaction $R_i \to R_j$ catalyzed per unit enzyme, $E_{ij}^{(\alpha)}$ is the enzyme allocated to reaction $R_i \to R_j$ by species α, $\mathcal{X}$ is the set of all allowed reactions, and δ is the dilution rate. The maintenance energy requirements of a species can be accounted for by replacing δ with $\delta + m$ in Eq. 8, where m is the energy flux required for cell maintenance. Our results apply in this scenario as well (see SI).

Resource dynamics are determined by their consumption and production via the reactions catalyzed by the species, external supply, and dilution. The dynamics of resource i with concentration R_i are given by

$$\frac{dR_i}{dt} = h(R_i) - \sum_{\alpha,j} N_\alpha E_{ij}^{(\alpha)} \mathcal{F}_{ij} + \sum_{\alpha,k} N_\alpha E_{ki}^{(\alpha)} \mathcal{F}_{ki} - \delta R_i \tag{9}$$

where $h(R_i)$ is the resource supply. Only the top resource R_0 is supplied, so $h(R_0) = \delta h_0$ and zero otherwise. h_0 is the concentration of resource in the supply.

Simulation procedure

All species in a pool were introduced at the same initial density (=1) in an environment with resource concentration of the supplied resource $R_0 = h_0$ and all other resources at a very small concentration 10^{-10}. The species, reaction, and resource dynamics were simulated till they reached a steady-state. At regular intervals, species that fell below an abundance threshold had their concentration set to zero. The minimum allowed resource concentration was 10^{-10} to prevent division by zero in Eq. (2). A steady-state was reached if the maximum value of the vector $\frac{1}{N_\alpha}\frac{dN_\alpha}{dt}$, measuring the logarithmic growth rates, fell below 10^{-5}.

Simulation parameters

The enzyme budget of a species was chosen from a lognormal distribution with lognormal parameters $\mu = 1, \sigma = 0.1$. A species utilized a random subset of the 15 possible reactions. So species in the pool were not biased towards extreme generalists or specialists. The enzyme budget was allocated to the selected reactions uniformly, by using a Dirichlet distribution with the dirichlet parameter, $\alpha = 1$, for the selected reactions. There were 600 species in each of the 5 pools in Fig.2 and 10 pools in Fig. 4; species were not shared across pools.

The energies of resources R_0, R_1, R_2, R_3, R_4, and R_5 were $5RT, 4RT, 3RT, 2RT, 1RT$, and $0RT$. For each reaction, a random fraction between 15% and 85% of the standard energy gap between product and substrate $(\mathcal{E}_S - \mathcal{E}_P)$ was assimilated and the rest was dissipated as heat. This assimilated energy, along with resource energies, was fixed across the experiments with different species pools. The concentration of R_0 in the supply, $h_0 = 80$. The chemical kinetic parameters K_P,K_S, and k_{cat} were chosen from lognormal distributions with lognormal parameters $\mu = 0., \sigma = 0.01$, and so, their

mean was close to 1. The energy to biomass yield factor, $Y = 1$. The dilution rate, $\delta = 0.01$ in Fig. 2; this rate was small enough that a typical species would survive if it was alone.

The Jensen-Shannon distance, used to measure functional distance between two flux vectors in Fig. 4. It is the symmetrized analog of the Kullback-Leibler divergence, and was defined as

$$\mathrm{JSD}(f^{(a)}, f^{(b)}) = \frac{1}{2}\sum_i f^{(a)} \log \frac{f^{(a)}}{f^{(m)}} + \frac{1}{2}\sum_i f^{(b)} \log \frac{f^{(b)}}{f^{(m)}}, \tag{10}$$

where $f^{(a)}, f^{(b)}$ are the two normalized flux vectors and $f^{(m)}$ is an average flux vector defined as $(f^{(a)} + f^{(b)})/2$. We report the JSD averaged over all pairs of communities in Fig. 4.

g_{max} in Fig. 4 was defined as the growth rate obtained from an irreversible reaction where a species invested its entire enzyme budget into one reaction. Since this varies between reactions because reactions different energy yields, we chose the energy gap between consecutive resources. It was calculated as the product of mean k_{cat}, mean enzyme budget, energy gap between consecutive resources ($= RT$), and biomass yield factor Y.

[1] Falkowski, P. G., Fenchel, T. & Delong, E. F. The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science* **320**, 1034–1039 (2008). URL https://www.sciencemag.org/lookup/doi/10.1126/science.1153213.

[2] Whitman, W. B., Coleman, D. C. & Wiebe, W. J. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences* **95**, 6578–6583 (1998). URL http://www.pnas.org/cgi/doi/10.1073/pnas.95.12.6578.

[3] Jessen, G. L. *et al.* Hypoxia causes preservation of labile organic matter and changes seafloor microbial community composition (Black Sea). *Science Advances* **3**, e1601897 (2017). URL https://www.science.org/doi/10.1126/sciadv.1601897.

[4] Jørgensen, B. B. & Boetius, A. Feast and famine — microbial life in the deep-sea bed. *Nature Reviews Microbiology* **5**, 770–781 (2007). URL http://www.nature.com/articles/nrmicro1745.

[5] Leng, L. *et al.* A review on the bioenergetics of anaerobic microbial metabolism close to the thermodynamic limits and its implications for digestion applications. *Bioresource Technology* **247**, 1095–1106 (2018). URL https://linkinghub.elsevier.com/retrieve/pii/S0960852417316681.

[6] Hoh, C.-Y. & Cord-Ruwisch, R. Experimental evidence for the need of thermodynamic considerations in modelling of anaerobic environmental bioprocesses. *Water Science & Technology* **36**, 7 (1997).

[7] Nobu, M. K. *et al.* Catabolism and interactions of uncultured organisms shaped by eco-thermodynamics in methanogenic bioprocesses. *Microbiome* **8**, 111 (2020). URL https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-020-00885-y.

[8] Bradley, J. A. *et al.* Widespread energy limitation to life in global subseafloor sediments. *Science Advances* **6**, eaba0697 (2020). URL https://www.science.org/doi/10.1126/sciadv.aba0697.

[9] Conrad, R., Schink, B. & Phelps, T. Thermodynamics of H $_2$ -consuming and H $_2$ -producing metabolic reactions in diverse methanogenic environments under in situ conditions. *FEMS Microbiology Letters* **38**, 353–360 (1986). URL https://academic.oup.com/femsec/article-lookup/doi/10.1111/j.1574-6968.1986.tb01748.x.

[10] Hickey, R. F. & Switzenbaum, M. S. Thermodynamics of Volatile Fatty Acid Accumulation in Anaerobic Digesters Subject to Increases in Hydraulic and Organic Loading. *Research Journal of the Water Pollution Control Federation* **63**, 141–144 (1991).

[11] Jackson, B. E. & McInerney, M. J. Anaerobic microbial metabolism can proceed close to thermodynamic limits. *Nature* **415**, 454–456 (2002). URL http://www.nature.com/articles/415454a.

[12] LaRowe, D. E. & Amend, J. P. The Energetics of Fermentation in Natural Settings. *Geomicrobiology Journal* **36**, 492–505 (2019). URL https://www.tandfonline.com/doi/full/10.1080/01490451.2019.1573278.

[13] LaRowe, D. E., Dale, A. W., Amend, J. P. & Van Cappellen, P. Thermodynamic limitations on microbially catalyzed reaction rates. *Geochimica et Cosmochimica Acta* **90**, 96–109 (2012). URL https://linkinghub.elsevier.com/retrieve/pii/S0016703712002864.

[14] Lomstein, B. A., Langerhuus, A. T., D'Hondt, S., Jørgensen, B. B. & Spivack, A. J. Endospore abundance, microbial growth and necromass turnover in deep subseafloor sediment. *Nature* **484**, 101–104 (2012). URL http://www.nature.com/articles/nature10905.

[15] Park, H. *et al.* Impact of inocula and growth mode on the molecular microbial ecology of anaerobic ammonia oxidation (anammox) bioreactor communities. *Water Research* **44**, 5005–5013 (2010). URL https://linkinghub.elsevier.com/retrieve/pii/S0043135410004902.

[16] Sottile, W. & Zabransky, R. J. Comparative Growth Rates of Selected Anaerobic Species in Four Commonly Used Broth Media. *Antimicrobial Agents and Chemotherapy* **11**, 482–490 (1977). URL https://journals.asm.org/doi/10.1128/AAC.11.3.482.

[17] MacArthur, R. Species packing and competitive equilibrium for many species. *Theoretical Population Biology* **1**, 1–11 (1970). URL https://linkinghub.elsevier.com/retrieve/pii/0040580970900390.

[18] Posfai, A., Taillefumier, T. & Wingreen, N. S. Metabolic Trade-Offs Promote Diversity in a Model Ecosystem. *Physical Review Letters* **118** (2017). URL https://link.aps.org/doi/10.1103/PhysRevLett.118.028103.

[19] Tikhonov, M. & Monasson, R. Collective Phase in Resource Competition in a Highly Diverse Ecosystem. *Physical Review Letters* **118** (2017). URL https://link.aps.org/doi/10.1103/PhysRevLett.118.048103.

[20] Goldford, J. E. *et al.* Emergent simplicity in microbial community assembly. *Science* **361**, 469–474 (2018). URL http://www.sciencemag.org/lookup/doi/10.1126/science.aat1168.

[21] Marsland, R., Cui, W. & Mehta, P. The Minimum En-

vironmental Perturbation Principle: A New Perspective on Niche Theory. *The American Naturalist* **196**, 291–305 (2020). URL https://www.journals.uchicago.edu/doi/10.1086/710093.

[22] Dubinkina, V., Fridman, Y., Pandey, P. P. & Maslov, S. Multistability and regime shifts in microbial communities explained by competition for essential nutrients. *eLife* **8** (2019). URL https://elifesciences.org/articles/49720.

[23] Wang, Z. *et al.* Complementary resource preferences spontaneously emerge in diauxic microbial communities. *Nature Communications* **12**, 6661 (2021). URL https://www.nature.com/articles/s41467-021-27023-y.

[24] Großkopf, T. & Soyer, O. S. Microbial diversity arising from thermodynamic constraints. *The ISME Journal* **10**, 2725–2733 (2016). URL http://www.nature.com/articles/ismej201649.

[25] Cook, J., Pawar, S. & Endres, R. G. Thermodynamic constraints on the assembly and diversity of microbial ecosystems are different near to and far from equilibrium. *PLOS Computational Biology* **17**, e1009643 (2021). URL https://dx.plos.org/10.1371/journal.pcbi.1009643.

[26] Fernández, A. *et al.* How Stable Is Stable? Function versus Community Composition. *Applied and Environmental Microbiology* **65**, 3697–3704 (1999). URL https://AEM.asm.org/content/65/8/3697.

[27] Louca, S., Parfrey, L. W. & Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**, 1272–1277 (2016). URL https://www.sciencemag.org/lookup/doi/10.1126/science.aaf4507.

[28] Louca, S. *et al.* Function and functional redundancy in microbial systems. *Nature Ecology & Evolution* **2**, 936–943 (2018). URL http://www.nature.com/articles/s41559-018-0519-1.

[29] Peces, M., Astals, S., Jensen, P. & Clarke, W. Deterministic mechanisms define the long-term anaerobic digestion microbiome and its functionality regardless of the initial microbial community. *Water Research* **141**, 366–376 (2018). URL https://linkinghub.elsevier.com/retrieve/pii/S0043135418304007.

[30] Dewar, R. C. & Maritan, A. A Theoretical Basis for Maximum Entropy Production. In Dewar, R. C., Lineweaver, C. H., Niven, R. K. & Regenauer-Lieb, K. (eds.) *Beyond the Second Law*, 49–71 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2014). URL http://link.springer.com/10.1007/978-3-642-40154-1_3. Series Title: Understanding Complex Systems.

[31] Jin, Q. & Bethke, C. M. The thermodynamics and kinetics of microbial metabolism. *American Journal of Science* **307**, 643–677 (2007). URL http://www.ajsonline.org/cgi/doi/10.2475/04.2007.01.

[32] Wachtel, A., Rao, R. & Esposito, M. Thermodynamically consistent coarse graining of biocatalysts beyond Michaelis–Menten. *New Journal of Physics* **20**, 042002 (2018). URL https://iopscience.iop.org/article/10.1088/1367-2630/aab5c9.

[33] Milo, R. & Phillips, R. *Cell biology by the numbers* (Garland Science, Taylor & Francis Group, New York, NY, 2016).

[34] Krieger, C. J. MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Research* **32**, 438D–442 (2004). URL https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkh100.

[35] Marsland, R. *et al.* Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities. *PLOS Computational Biology* **15**, e1006793 (2019). URL http://dx.plos.org/10.1371/journal.pcbi.1006793.

[36] Fant, L., Macocco, I. & Grilli, J. Eco-evolutionary dynamics lead to functionally robust and redundant communities. preprint, Ecology (2021). URL http://biorxiv.org/lookup/doi/10.1101/2021.04.02.438173.

[37] Cornish-Bowden, A. *Fundamentals of enzyme kinetics* (Wiley-Blackwell, Weinheim, Germany, 2012), 4th edn.

[38] Noor, E., Flamholz, A., Liebermeister, W., Bar-Even, A. & Milo, R. A note on the kinetics of enzyme action: A decomposition that highlights thermodynamic effects. *FEBS Letters* **587**, 2772–2777 (2013). URL http://doi.wiley.com/10.1016/j.febslet.2013.07.028.

[39] Jin, Q. & Bethke, C. M. A New Rate Law Describing Microbial Respiration. *Applied and Environmental Microbiology* **69**, 2340–2348 (2003). URL https://AEM.asm.org/content/69/4/2340.

[40] Hoehler, T. M., Alperin, M. J., Albert, D. B. & Martens, C. S. Thermodynamic control on hydrogen concentrations in anoxic sediments. *Geochimica et Cosmochimica Acta* **62**, 1745–1756 (1998). URL https://linkinghub.elsevier.com/retrieve/pii/S0016703798001069.

[41] Westermann, P. The effect of incubation temperature on steady-state concentrations of hydrogen and volatile fatty acids during anaerobic degradation in slurries from wetland sediments. *FEMS Microbiology Ecology* **13**, 295–302 (1994). URL https://academic.oup.com/femsec/article-lookup/doi/10.1111/j.1574-6941.1994.tb00076.x.

[42] Arndt, S. *et al.* Quantifying the degradation of organic matter in marine sediments: A review and synthesis. *Earth-Science Reviews* **123**, 53–86 (2013). URL https://linkinghub.elsevier.com/retrieve/pii/S0012825213000512.

[43] Bar-Even, A. *et al.* The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry* **50**, 4402–4410 (2011). URL https://pubs.acs.org/doi/10.1021/bi2002289.

[44] Mentges, A., Feenders, C., Deutsch, C., Blasius, B. & Dittmar, T. Long-term stability of marine dissolved organic carbon emerges from a neutral network of compounds and microbes. *Scientific Reports* **9**, 17780 (2019). URL http://www.nature.com/articles/s41598-019-54290-z.

[45] Prim, R. C. Shortest Connection Networks And Some Generalizations. *Bell System Technical Journal* **36**, 1389–1401 (1957). URL https://ieeexplore.ieee.org/document/6773228.

[46] Zhou, J. *et al.* Stochastic Assembly Leads to Alternative Communities with Distinct Functions in a Bioreactor Microbial Community. *mBio* **4** (2013). URL https://journals.asm.org/doi/10.1128/mBio.00584-12.

[47] Schink, B. & Stams, A. J. M. Syntrophism among Prokaryotes. In Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H. & Stackebrandt, E. (eds.) *The Prokaryotes: Volume 2: Ecophysiology and Biochemistry*, 309–335 (Springer New York, New York, NY, 2006). URL https://doi.org/10.1007/0-387-30742-7_11.

[48] Schink, B. Energetics of syntrophic cooperation in methanogenic degradation. *MICROBIOL. MOL. BIOL. REV.* **61**, 19 (1997).

[49] Lever, M. A. *et al.* Life under extreme energy lim-

itation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiology Reviews* **39**, 688–728 (2015). URL https://academic.oup.com/femsre/article/39/5/688/531037.

[50] Hoehler, T. M. Biological energy requirements as quantitative boundary conditions for life in the subsurface. *Geobiology* **2**, 205–215 (2004). URL https://onlinelibrary.wiley.com/doi/10.1111/j.1472-4677.2004.00033.x.

[51] Henry, C. S., Broadbelt, L. J. & Hatzimanikatis, V. Thermodynamics-Based Metabolic Flux Analysis. *Biophysical Journal* **92**, 1792–1805 (2007). URL https://linkinghub.elsevier.com/retrieve/pii/S0006349507709876.

[52] Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W. & Milo, R. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proceedings of the National Academy of Sciences* **110**, 10039–10044 (2013). URL http://www.pnas.org/cgi/doi/10.1073/pnas.1215283110.

[53] Goldford, J. E., George, A. B., Flamholz, A. I. & Segrè, D. Protein cost minimization promotes the emergence of coenzyme redundancy. preprint, Systems Biology (2021). URL http://biorxiv.org/lookup/doi/10.1101/2021.05.17.444566.

[54] Schneider, E. & Kay, J. Life as a manifestation of the second law of thermodynamics. *Mathematical and Computer Modelling* **19**, 25–48 (1994). URL https://linkinghub.elsevier.com/retrieve/pii/0895717794901880.

[55] Vallino, J. J. Ecosystem biogeochemistry considered as a distributed metabolic network ordered by maximum entropy production. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 1417–1427 (2010). URL https://royalsocietypublishing.org/doi/10.1098/rstb.2009.0272.

[56] Meysman, F. J. R. & Bruers, S. Ecosystem functioning and maximum entropy production: a quantitative test of hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 1405–1416 (2010). URL https://royalsocietypublishing.org/doi/10.1098/rstb.2009.0300.

[57] Noor, E., Haraldsdóttir, H. S., Milo, R. & Fleming, R. M. T. Consistent Estimation of Gibbs Energy Using Component Contributions. *PLoS Computational Biology* **9**, e1003098 (2013). URL https://dx.plos.org/10.1371/journal.pcbi.1003098.

[58] Jinich, A. *et al.* Quantum chemistry reveals thermodynamic principles of redox biochemistry. *PLOS Computational Biology* **14**, e1006471 (2018). URL https://dx.plos.org/10.1371/journal.pcbi.1006471.

[59] Jinich, A., Sanchez-Lengeling, B., Ren, H., Harman, R. & Aspuru-Guzik, A. A Mixed Quantum Chemistry/Machine Learning Approach for the Fast and Accurate Prediction of Biochemical Redox Potentials and Its Large-Scale Application to 315 000 Redox Reactions. *ACS Central Science* **5**, 1199–1210 (2019). URL https://pubs.acs.org/doi/10.1021/acscentsci.9b00297.