Received: 17 June 2021 | Revised: 17 August 2021 | Accepted: 20 August 2021

DOI: 10.1111/mec.16161

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Diversity begets diversity during community assembly until ecological limits impose a diversity ceiling

Magdalena San Roman[1,2] | Andreas Wagner[1,2,3,4]

[1]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

[2]Swiss Institute of Bioinformatics, Lausanne, Switzerland

[3]The Santa Fe Institute, Santa Fe, NM, USA

[4]Stellenbosch Institute for Advanced Study, Wallenberg Research Centre at Stellenbosch University, Stellenbosch, South Africa

Correspondence
Andreas Wagner, Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.
Email: andreas.wagner@ieu.uzh.ch

Funding information
University Priority Research Program in Evolutionary Biology; H2020 European Research Council, Grant/Award Number: 739874; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A_172887

Abstract

Microbial communities are hugely diverse, but we do not yet understand how species invasions and extinctions drive and limit their diversity. On the one hand, the ecological limits hypothesis posits that diversity is primarily limited by environmental resources. On the other hand, the diversity-begets-diversity hypothesis posits that such limits can be easily lifted when new ecological niches are created by biotic interactions. To find out which hypothesis better explains the assembly of microbial communities, we used metabolic modelling. We represent each microbial species by a metabolic network that harbours thousands of biochemical reactions. Together, these reactions determine which carbon and energy sources a species can use, and which metabolic by-products—potential nutrients for other species—it can excrete in a given environment. We assemble communities by modelling thousands of species invasions in a chemostat-like environment. We find that early during the assembly process, diversity begets diversity. By-product excretion transforms a simple environment into one that can sustain dozens of species. During later assembly stages, the creation of new niches slows down, existing niches become filled, successful invasions become rare, and species diversity plateaus. Thus, ecological limitations dominate the late assembly process. We conclude that each hypothesis captures a different stage of the assembly process. Species interactions can raise a community's diversity ceiling dramatically, but only within limits imposed by the environment.

KEYWORDS
community assembly, metabolic modelling, microbial diversity, species–species interactions

1 | INTRODUCTION

Bacterial life on Earth is extraordinarily diverse. Our planet is inhabited by an estimated 1.4–1.9 million bacterial lineages (Louca et al., 2018), and every gram of soil hosts between 2000 and 18,000 distinct such lineages (Daniel, 2005). Understanding how Earth holds all this biodiversity is a fundamental challenge to ecology and evolution. Part of the difficulty in meeting this challenge is that planet Earth offers limited resources, which may impose a ceiling on biodiversity. In this vein, the prominent "ecological limits" hypothesis of biodiversity (Rabosky & Hurlbert, 2015; Schluter & Pennell, 2017) posits that rates of diversification should decrease as diversity increases and fills available ecological niches. In contrast, the "diversity begets diversity" hypothesis posits that diversity could stimulate further diversification (Calcagno et al., 2017; Whittaker, 1972) as species–species interactions get more complex and novel niches are created (Erwin, 2008; Laland et al., 1999), or as existing niches are partitioned more and more finely as a result of competition (Bailey et al., 2013; Dieckmann & Doebeli, 1999).

Evidence for either hypothesis in bacteria, animals and plants is mixed. For example, a combination of field work and phylogenetic analysis of Himalayan songbirds suggests that an observed slowdown in their speciation rate is explained by niche filling. Their species distributions are well explained by resource abundance (Price et al., 2014). Two evolution experiments with *Pseudomonas fluorescens* strains also support the ecological limits hypothesis (Brockhurst et al., 2007; Gómez & Buckling, 2013). In these experiments a strain's diversification slowed down as more strains were included in a coculture. Conversely, similar experiments with different strains of *P. fluorescens* support the “diversity begets diversity” perspective. These experiments tracked the diversification of a focal lineage of *P. fluorescens* in bacterial communities that harboured from one to eight lineages of *P. fluorescens*. More novel morphotypes evolved when communities were initially more diverse (Jousset et al., 2016). Further support for the “diversity begets diversity” hypothesis comes from experiments with crops and weeds, which show that weed diversity is high whenever crop diversity is high (Palmer & Maurer, 1997). It also comes from data on plant and arthropod diversity on the Canary and Hawaiian islands, where the proportion of endemic species increases with increasing species numbers (Emerson & Kolm, 2005). It also comes from inferences of microbial diversification rates from several taxonomic ratios, such as the species to genus ratio (Madi et al., 2020). In addition, theoretical modelling using adaptive dynamics also suggests that some initial diversity can facilitate further diversification (Calcagno et al., 2017).

The diversity of a community may increase both through ecological processes, as species disperse and invade a community, or through evolutionary processes, as existing member species diversify and speciate. Pertinent evidence on the ecological limits and diversity-begets-diversity hypotheses comes in part from experiments or field work that only quantify diversification rates, making it difficult to disentangle ecological from evolutionary limitations. However, the hypotheses should ideally be distinguished by considering both evolutionary and ecological processes. In a scenario where one can control the appearance of new species in a community, either because they evolve in the community or disperse into the community, will environmental resources constrain community diversity, or will species–species interactions promote ever-increasing diversity? This is the question we aim to answer here.

To this end, we used a very simple community assembly strategy inspired by the environmental filter metaphor (Levy & Borenstein, 2013; Thakur & Wright, 2017) and by trait-based models (Mcgill et al., 2006). Such models organize ecological processes around traits—species properties that impact species survival. Typically, species show trade-offs in these traits, and a trait that gives a species an advantage in one environment might be disadvantageous in another. The environment selects (filters) species according to their traits. This modelling strategy has been widely successful, for example to predict community composition along environmental gradients (Allison, 2012; Laughlin et al., 2012; Litchman & Klausmeier, 2008; Mcgill et al., 2006; Thakur & Wright, 2017). Here, we use it to model the assembly of microbial communities, where a trait refers to a species' ability to thrive on a specific source of carbon and energy.

We simulated the assembly of microbial communities in which we represent each individual species by a metabolic network that comprises thousands of metabolic reactions needed for an organism's survival in a given chemical environment. The advantage of this approach is that it allows fundamental metabolic traits to emerge from first biochemical principles, which are embodied in the metabolic reaction network of an organism. Especially important traits for our purpose include the ability to survive on a given source of carbon and energy, and on the ability to excrete specific by-products of metabolism. Traits like these can be computationally predicted with flux balance analysis (FBA; Orth et al., 2010), an experimentally validated (Orth et al., 2011; Varma & Palsson, 1994) computational method that can determine the flux of matter through every reaction in a metabolic network when cells are in a metabolic steady-state, and when they grow their biomass is at the maximum possible rate given the biochemical reactions they can catalyse. FBA has been successfully used to predict growth and by-product secretion of bacteria in different media (Edwards et al., 2001; Ibarra et al., 2002; Varma & Palsson, 1994).

Metabolic networks have been characterized for hundreds of organisms (Gu et al., 2019a). They are highly valuable to understand metabolic biology and evolution because they reflect an organism's evolutionary history. For the same reason, however, they are of limited use for studies like ours, which aim to understand how community assembly may be affected by specific metabolic properties of the assembled species, such as the number of biochemical reactions any one species harbours, and the number of carbon sources it can utilize. For this purpose, one needs to vary these properties systematically, but in the network of any one organism, they are fixed. To circumvent this limitation, we started our community assembly not from previously characterized metabolic networks of microbial species. Instead, we randomly sampled (with a Markov chain Monte Carlo [MCMC] method) thousands of metabolic networks from a much larger “universe” or “pan-metabolism” of biochemical reactions, such that each network fulfilled specific requirements. These requirements include the ability to sustain life on specific carbon sources such as glucose, while containing an otherwise random complement of chemical reactions. We refer to such metabolic networks as random viable networks. For the purpose of our analysis, each such network represents a different “species,” and we used these “species” to simulate the assembly of a community in a well-mixed chemostat-like environment where resources are supplied by the environment at a constant rate.

Our community assembly procedure consists of three iterated steps (Figure 1). First, a random species invades the environment/community. Second, the environment acts as a filter that selects those species that persist in the community. Third, the persisting

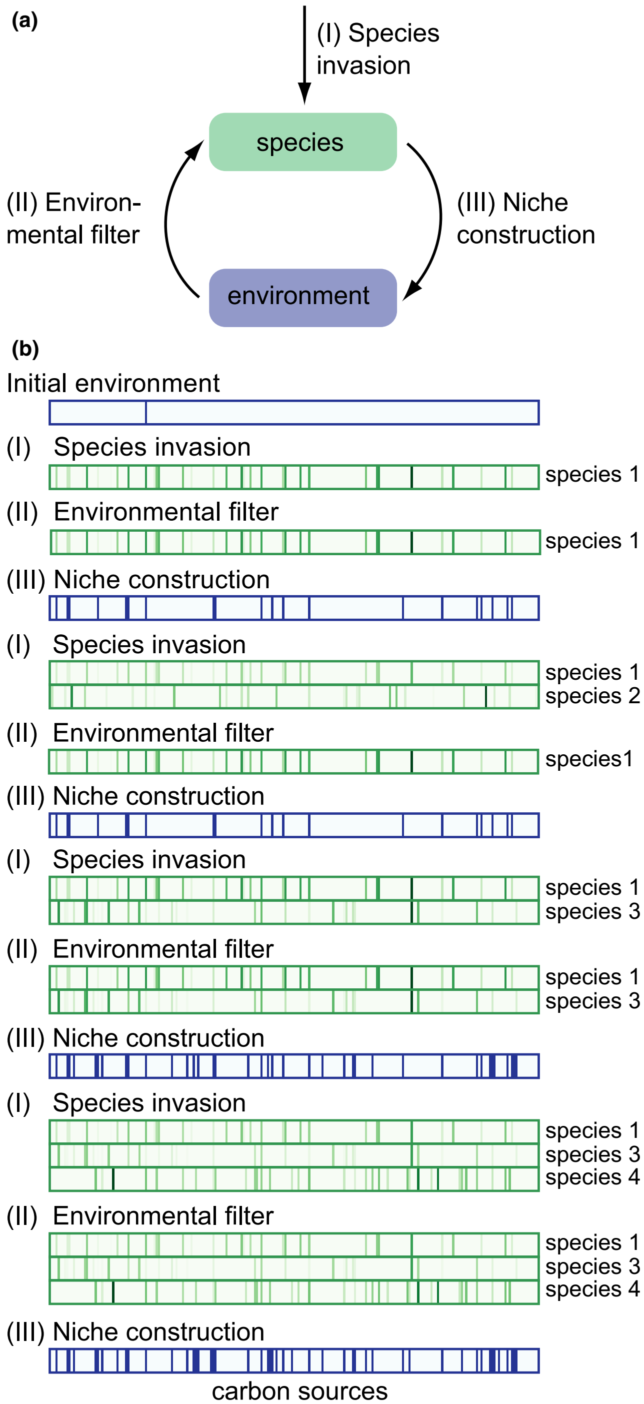

FIGURE 1 Modelling community assembly. (a) The three steps we used to model community assembly. (I) Species (modelled as random viable metabolisms) are added at random to a standing community. (II) The environment acts as a filter, selecting successful species, that is those best at growing on at least one of the carbon sources available in the environment. Successful species persist in the community. (III) The species that comprise the community modify the environment through the excretion of metabolites that can serve as carbon sources for other species. (b) Example assembly trajectory of one community up to the invasion of the fourth species. Simulations begin by initializing the composition of the environment, shown as a purple rectangle at the top. The presence of each of the 223 potential carbon sources is represented as a purple vertical bar, such that each location on the horizontal *x*-axis corresponds to one carbon source. In this example and in most of our simulations (unless otherwise stated), the initial environment contains glucose as the sole carbon source, which is indicated in the top rectangle by the single vertical purple line. After establishing the initial environment, we perform the first invasion with "species 1" chosen at random from our MCMC-derived sample of random viable networks. The species' ability to grow on each potential carbon source (whether or not the carbon source is present in the environment) is shown in the green rectangle by a green line in the *x*-position that corresponds to the specific carbon source. Darker green lines indicate higher growth. Species 1 can grow on the only available carbon source (glucose), and because no other species is yet present, we consider species 1 successful. Species 1 modifies the environment with the excretion of by-products of its growth on glucose, as shown in the second purple rectangle from the top. These by-products are available as potential nutrients for the second round of assembly. In this round, a randomly chosen species 2 invades the community. Its growth rate on each carbon source is shown below that of species 1. Species 2 cannot grow faster than species 1 on any of the available carbon sources. It therefore goes extinct and only species 1 persists in the community. The carbon sources it excretes are available for the third assembly step. Each such step consists of a new round of (I) species invasion, (II) environmental filtering and (III) niche construction

species may change the environment by excreting metabolic by-products. We repeated these three assembly steps for thousands of species invasions in any one community, and simulated hundreds of community assemblies in environments with different amounts of resources and varying strengths of competition.

During community assembly, we found that many persisting species excrete metabolic by-products, which can sustain species that invade later through cross-feeding interactions in which one organism consumes the excretion of another. In other words, species create opportunities for cross-feeding, which begets further diversity. Further, we find that communities grow more diverse when competition among species is stronger. Together, these two observations show that biotic interactions between species are critical to establishing diverse communities. At the same time, communities show limits in the number of species they accommodate. As community richness increases, the probability that a new species successfully invades a community decreases. The reason is that by-product excretion cannot create new niches ad infinitum. Eventually all new niches have been created, and subsequent invasions mostly fill the existing niches. At this late stage, the assembly dynamics enters an "ecological limits" regime.

2 | METHODS

2.1 | Flux balance analysis

FBA is a computational method to predict metabolic flux—the rate at which chemical reactions convert substrates into products—through all biochemical reactions that make up the organism's metabolic network (Orth et al., 2010). FBA requires information about the stoichiometry of all chemical reactions in the metabolic network, which is mathematically represented in the stoichiometric matrix S of size $m \times r$. The integer m denotes the number of metabolites, and the integer r denotes the number of reactions in the network. Each entry S_{ij} of the stoichiometric matrix contains the stoichiometric coefficient with which metabolite i participates in reaction j. FBA makes two central assumptions. The first is that cells effectively optimize some metabolic property such as growth rate (v_{growth}). The second is that cells (metabolisms) are in a metabolic steady state. Additional constraints can be incorporated into the optimization problem that FBA solves by setting lower bounds l_i and upper bounds u_i to individual reaction fluxes v_i, in order to account for the thermodynamic and enzymatic properties of a biochemical reaction (Orth et al., 2010). The optimization problem that FBA solves can be formalized as a linear programming problem (Orth et al., 2010; Varma & Palsson, 1994) in the following way:

$$\text{Maximize}\, v_{\text{growth}}$$
$$\text{s.t.}\, Sv = 0$$
$$l_i \leq v_i \leq u_i$$

We performed FBA with COBRAPY (Ebrahim et al., 2013).

2.2 | Modelling species with random viable networks

We sampled random viable networks with a computational variant of MCMC sampling, which permits efficient sampling of the large space containing all possible metabolic reaction, that is, the pan-metabolism. For our work, we used a pan-metabolic network comprising 5625 metabolites and 7222 biochemical reactions (Barve & Wagner, 2013; San Roman & Wagner, 2018). To sample random viable networks, we began with a specification of desired metabolic characteristics. For example, for many of our analyses, we wanted to create networks that are at least viable on glucose as a sole carbon source, and that comprise as many reactions as *Escherichia coli* (i.e., 2583 reactions in the widely used *E. coli* metabolic model *i*JO1366). To identify a single initial network fulfilling these requirements, we first performed FBA on the pan-metabolic network (Barve & Wagner, 2013; San Roman & Wagner, 2018) in a chemically minimal environment with glucose as the only source of carbon. Of all reactions in this pan-metabolic network, 1263 reactions showed nonzero metabolic flux. We included all these reactions in the initial network, which ensured the network's viability on glucose. We then chose the remaining (1320) reactions needed to arrive at an equal number of reactions as the *E. coli* metabolism at random (with a uniform distribution) from the pan-metabolic network.

This initial network was the starting point of MCMC sampling from the space of all possible metabolic networks. The MCMC procedure performs a long random walk through this space. Each step in this walk consists of a reaction swap, in which a reaction from the current network is randomly chosen for deletion, while a randomly chosen reaction from the pan-metabolism but absent in the current network is added to the current network (without changing reaction reversibility or stoichiometry). If the modified network remains viable after this reaction swap, the swap is accepted, and the network is modified with a further reaction swap. In contrast, if the reaction swap disrupts viability on the desired carbon source(s), such as glucose, the swap is rejected and a new swap is tried. Modifying metabolic networks through reaction swaps ensures that the number of reactions in the network remains constant and equal to the number of reactions in the initial network.

This MCMC procedure creates a long sequence of metabolic networks, all of them viable on the desired carbon sources. As the number of reaction swaps increases, the number of reactions that the altered network shares with the initial network becomes smaller and smaller, until its complement of reactions becomes effectively randomized, which occurs after ~5000 successful swaps (Samal et al., 2010). We stored such a randomized network for further analysis after 5000 successful swaps, and repeated this procedure from the initial network to create 1000 random networks viable on specific carbon sources, such as glucose. During this process we did not alter the exchange reactions of the starting network, which ensures that in the randomized networks the same metabolites can be exchanged with the environment as in *E. coli*. Furthermore, we used the biomass reaction from *E. coli i*JO1366 to assess the viability of networks during MCMC sampling. The random viable networks used in this study are available in GitHub at https://doi.org/10.5281/zenodo.5206841.

2.3 | Potential carbon sources

We performed FBA to examine all metabolites that the pan-metabolic network could use as a sole carbon source (San Roman & Wagner, 2018). To this end, we considered those carbon-containing metabolites which have an exchange reaction in the *E. coli* model *i*JO1366, and we used the biomass reaction of the same *E. coli* model in the pan-metabolic network. We performed FBA with the pan-metabolism in a simulated environment in which one candidate carbon source was available at a time and assumed that ammonium, calcium, chloride, cobalt, copper, iron, magnesium, manganese, molybdate, nickel, oxygen, phosphate, potassium, protons, sodium, sulphate and zinc were available in the environment in nonlimiting amounts. We considered a metabolite a "potential" carbon source if it allowed biomass production of the pan-metabolism. We identified 223 potential carbon sources. Any metabolism sampled from the pan-metabolic network will be viable on a subset of these carbon sources.

2.4 | Identifying species traits with FBA

For every random viable metabolic network ("species") we considered, we used FBA to determine its viability on all 223 potential carbon sources. To this end, we performed FBA in a minimal environment where only one potential carbon source was available at a time, as explained in the previous section. The result was a predicted biomass growth flux for each species on each of the 223 potential carbon sources. We used these predicted growth rates in the second step of our simulation of community assembly (i.e., in the environmental filtering step).

We also used FBA to analyse the niche construction potential of every species. We were interested in those metabolites that could be excreted by one species, and thus potentially serve as carbon sources for some other species. Therefore, we identified those metabolites from the list of potential carbon sources that could be produced as a by-product of growth when a species consumes a given carbon source. To this end, we allowed consumption of this carbon source (with a maximal uptake rate of 10 mmol $gDW^{-1} h^{-1}$) in an otherwise minimal environment, and maximized the production of each potential carbon source, while constraining the species' biomass production to be greater than zero. If a potential carbon source could be produced at a rate greater than zero under this constraint, we considered that the carbon source was excreted into the environment.

This modelling decision is motivated by two kinds of experimental observations. First, microbes modify their environment by secreting multiple metabolites. For instance, targeted metabolomics experiments show that *Saccharomyces cerevisiae*, *E. coli*, *Bacillus licheniformis* and *Corynebacterium glutamicum* excrete 30–40 metabolites due to overflow metabolism (Paczia et al., 2012). (Untargeted metabolomics might reveal an even greater diversity of secreted metabolites.) Second, metabolic modelling tends to underestimate secretions found experimentally, regardless of whether it maximizes microbial growth rate (Pinu et al., 2018).

3 | RESULTS

3.1 | Modelling community assembly

We first created a pool of 1000 species, from which we sample individual species for community assembly. Each species is represented by a metabolic network that we require to be viable—it can synthesize all essential biomass molecules such as amino acids, lipids and nucleotides (see Section 2)—on a specific carbon source, but that contains an otherwise random complement of biochemical reactions from a much larger universe ("pan-metabolism") of biochemical reactions known to take place in the biosphere. To create this species pool, we used an MCMC sampling procedure, which allows us to create random viable metabolic networks with specific characteristics (see Section 2). For most of our analysis these characteristics are that, first, all species (metabolic networks) contain the same number of 2583 metabolic reactions as the well-established and widely used metabolic model *i*JO1366 of *Escherichia coli* (Orth et al., 2011). Second, all species are permeable to the same 330 metabolites that *E. coli* cells can import or export (Orth et al., 2011). Third, we require all species to be viable on the same carbon source, glucose. We note that as a result of metabolism's complex and reticulate structure, species required to be viable on one carbon source are usually also viable on multiple others (Barve & Wagner, 2013). (We relax these constraints in later analyses (Figures S4 and S6)

We used FBA (see Methods) to determine viability for each of our 1000 species on 223 potential carbon sources (see Section 2), and found that one species is viable on an average of 32 ± 10 different carbon sources (Figure S1). We consider the carbon sources on which any one species is viable as the species' fundamental ecological niche. Together, all 1000 species are viable on 220 different carbon sources. On average, 15 ± 13% of the species are viable on each carbon source. In addition to viability, we also determined the metabolic by-products of growth that any one species excretes when growing on any one carbon source. We found that on average each of our species excretes 23 ± 4 metabolic by-products when growing on glucose, which is comparable to the number of secretions found in microbes such as *E. coli* or *Saccharomyces cerevisiae* (Methods; Paczia et al., 2012).

We began our community assembly process with a deliberately simple chemical environment, which supplies only one carbon source—glucose—at a constant rate. In other words, in this initial environment only one metabolic niche exists. Community assembly then consisted of multiple iterations of the following three steps (Figure 1):

1. Species invasion: a species is introduced into the environment/community. This step simulates the invasion of the community by a species, or the evolutionary origin of a new species within a community. Each such species is selected at random from our species pool with equal probability. The same species can be selected multiple times in the course of assembling one community.
2. Environmental filtering: the environment serves as a filter selecting those species from the community that can persist in the environment. Specifically, we assume that those species that persist achieve the highest biomass growth among all present species, when consuming at least one of the available carbon sources in the environment. All other species go extinct. This modelling decision is based on the competitive exclusion principle (Hardin, 1960), which states that species cannot coexist if they compete for the same limiting resources. This second step embodies competition between species in the community (Levy & Borenstein, 2013). We apply this filtering method because simulating the population dynamics of all species in each community is computationally infeasible, given that we simulate thousands of invasion steps in each of the hundreds of communities we assemble for each analysis.
3. Niche construction: all species that make up the community at a given time modify the environment by excreting by-products of metabolizing the carbon sources they use for growth.

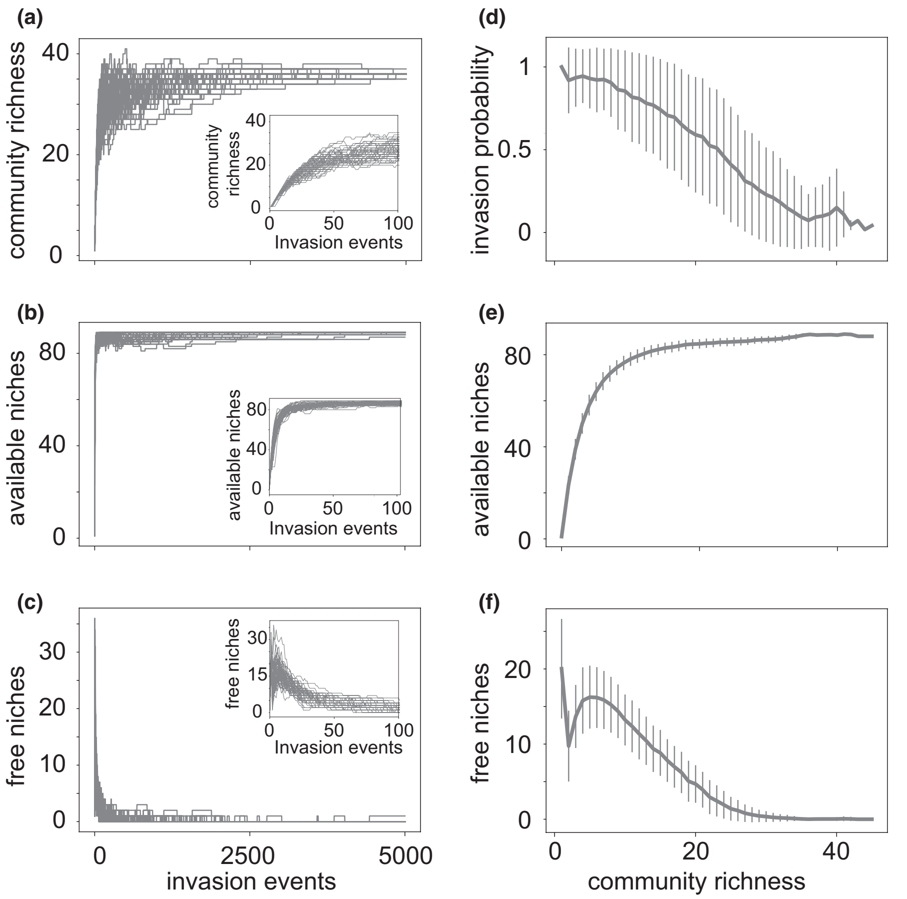

FIGURE 2 Species and niche dynamics during community assembly. (a–c) The horizontal axis shows time (number of species invasion events) during community assembly. The vertical axes show (a) community richness (number of species), (b) number of available niches (carbon sources in the environment) and (c) number of free niches (carbon sources present in the environment and not consumed by any species in the community). The insets show the same data but for up to 100 invasions. (d–f) The horizontal axis shows community richness. The vertical axes show (d) invasion probability (number of trials before a successful invasion), (e) the number of available niches (carbon sources) and (f) the number of free niches as a function of community richness. The plots show data from 50 (a–c) or 500 (d–f) assembled communities. During assembly, each community was subject to 5000 species invasion events. Error bars in (d)–(f) indicate one standard deviation

We repeated this three-step process for 5000 invasions, at which point the number of species in the communities no longer increases. In addition, we repeated the entire process of community assembly 500 times.

Figure 2a shows the assembly dynamics for 50 randomly chosen communities out of a total of 500 assembled communities, where each community had experienced up to 5000 individual species invasion events. Figure 2a shows that the number of persisting species (community richness, akin to alpha-diversity; Krebs, 2014) increases rapidly until it plateaus before 5000 species invasions. A final assembled community harbours on average 35.9 ± 0.4 species (based on 500 assembled communities). Final species richness varies somewhat, and ranges from 33 to 38 species. Collectively, the final 500 communities harbour 64 different species, which implies extensive overlap in the species making up different communities. This suggests that the order in which species invade the community has little effect on the community composition.

Figure 2b shows that the number of available metabolic niches (i.e., the number of carbon sources present in the environment) also increases rapidly. This increase results from the excretion of an increasing number of metabolic by-products as a community comes to host more and more species. At the same time, an increasing number of species also consume an increasing number of carbon sources. This is why the number of free niches (i.e., the number of carbon sources that are not consumed by any species in the community) only rises early during community assembly (Figure 2c). It reaches a peak after approximately six species invasions, and then declines slowly to zero. In other words, late in the assembly process, newly invading species are more likely to consume existing carbon sources (which are excreted by resident species) than they are to lead to the excretion of new carbon sources.

We determined the probability that a new species invasion is successful (i.e., that the invading species persists) as the number of invasion events that take place before a successful event occurs. Figure 2d shows that this probability declines monotonically with the number of species in a community. Figure 2e,f show the number of available and free niches as a function of community richness, averaged over 500 communities. The underlying data show that merely 12 ± 2 species are needed to construct 90% of the maximal number of available niches (Figure 2e). When a community hosts more than 6 ± 2 species, a successful species invasion begins to create fewer niches than it fills (i.e., the number of newly excreted carbon sources becomes smaller than the additional number of carbon sources consumed). As a consequence, the number of free niches stops increasing and starts decreasing (see peak in Figure 2f).

Together, the data in Figure 2 demonstrate how important biotic interactions and in particular the construction of new metabolic niches is to the assembly process we studied. In the absence of niche construction, our environment could have only accommodated the single species that is best suited for growth on the only carbon source initially available in the environment (Hardin, 1960). However, the construction of new niches does not continue throughout the assembly process. Towards the end of this process, newly invading species are more likely to fill existing niches than to help create new ones.

3.2 | More complex environments lead to more diverse communities

Thus far, we had assembled communities starting from the simplest possible environment, which contains only a single carbon and energy source. We next wanted to find out how abiotic environmental complexity (i.e., the number of resources in the environment) affects community assembly. We hypothesized that more complex environments sustain richer communities, because they harbour more niches that are independent of any one species' metabolic excretions. To validate this hypothesis, we assembled communities in environments that contained one, five, 10, 15 or 20 carbon sources. All these environments contained glucose, and we selected the remaining carbon sources at random from the list of 223 potential carbon sources (see Section 2).

Indeed, environments offering more resources lead to more diverse final communities (Figure 3). However, the complexity of the environment does not translate one-to-one into final community richness. For example, an environment that offers 20 times more resources than the simplest environment results in communities that sustain only 4.0 ± 0.1 additional species (Figure 3). We also found that the invasion probability as a function of community richness is modestly but consistently higher for the most complex environment (Figure S3a). We also observed more available and free niches in the initially complex environment than in the simple one (Figure S3b,c).

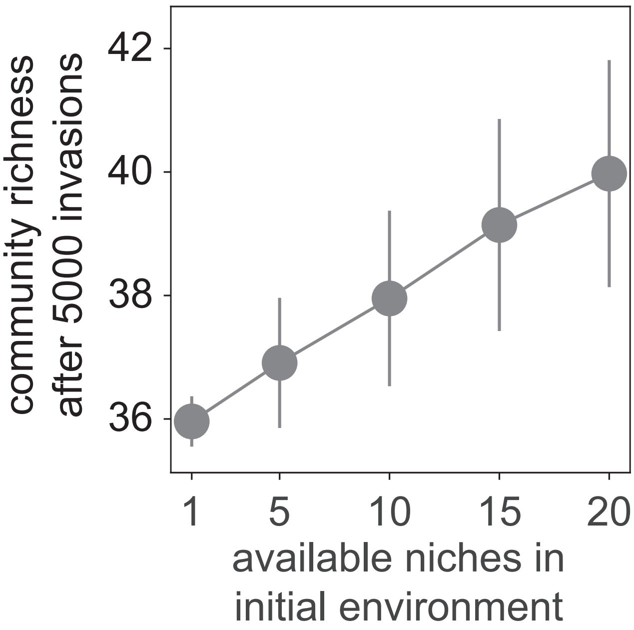

FIGURE 3 Richer initial environments lead to richer final communities. We simulated the assembly of communities in environments that offered between one and 20 carbon sources (horizontal axis) at the beginning of community assembly. The vertical axis shows the number of species (community richness) at the end of a community assembly process consisting of 5000 species invasions. Circles show means obtained from the assembly of 500 communities. Bars indicate one standard deviation

3.3 | Competition narrows the realized niche of species

Classical competition theory (Hutchinson, 1959) argues that competition is the major biological process controlling the structure of natural communities. Quantifying competition is generally difficult (Krebs, 2014) but our method allows us to quantify and manipulate competition strength in two ways. In this section, we explore how competition affects community richness. Specifically, we model exploitative competition (Tilman, 1982) for limiting carbon sources, and do so in two complementary ways. First, we change the number of species that compete for the available resources. Second, we change the species' fundamental niche width. We next describe both approaches in more detail.

We first varied the strength of competition by changing the number of species that can invade a community (i.e., by changing the size of the species pool). In the previous section, we had assembled communities through iterative species invasions, where we had chosen invading species from a pool of 1000 species. To strengthen competition, we now changed the pool size to 250, 500 or 750 species, sampled at random from the original set of 1000 species.

As the size of a species pool becomes larger, the likelihood that any one species is the fastest-growing among all species on any one carbon source is lower, which means that competition for resources gets stronger. In Figure 4a we show how increasing the species pool size does indeed decrease the likelihood that any one species from the pool is the fastest-growing. To calculate this likelihood we first computed the average number of resources on which a species showed the highest growth. This quantity varies between zero (if the species grows worse than any other species from the pool on all carbon sources) and one (if the species grows better than all other species on all carbon sources). We repeated this calculation for every species in a pool, and calculated the average for all species in the pool. We then repeated this analysis 500 times, each time for a different random pool of species (that comprised 250, 500 or 750 species). Figure 4a reports the average over these 500 replicates.

The stronger competition embodied in larger species pools results in more diverse communities (Figure 4b). This higher community richness is explained by two factors. First, as competition gets stronger, we observe a modest increase in the total number of niches available after 5000 species invasions (86.9 ± 1.6 and 88.9 ± 0.2 available niches for species pools of 250 and 1000 species, respectively, Figure S4b) while the number of free niches after 5000 invasions is not affected (all available niches are filled, Figure S4c). In other words, more niches are filled when competition is stronger. Second, we observe that stronger competition causes a reduction of the realized niche (Figure 4c). That is, individual persisting species

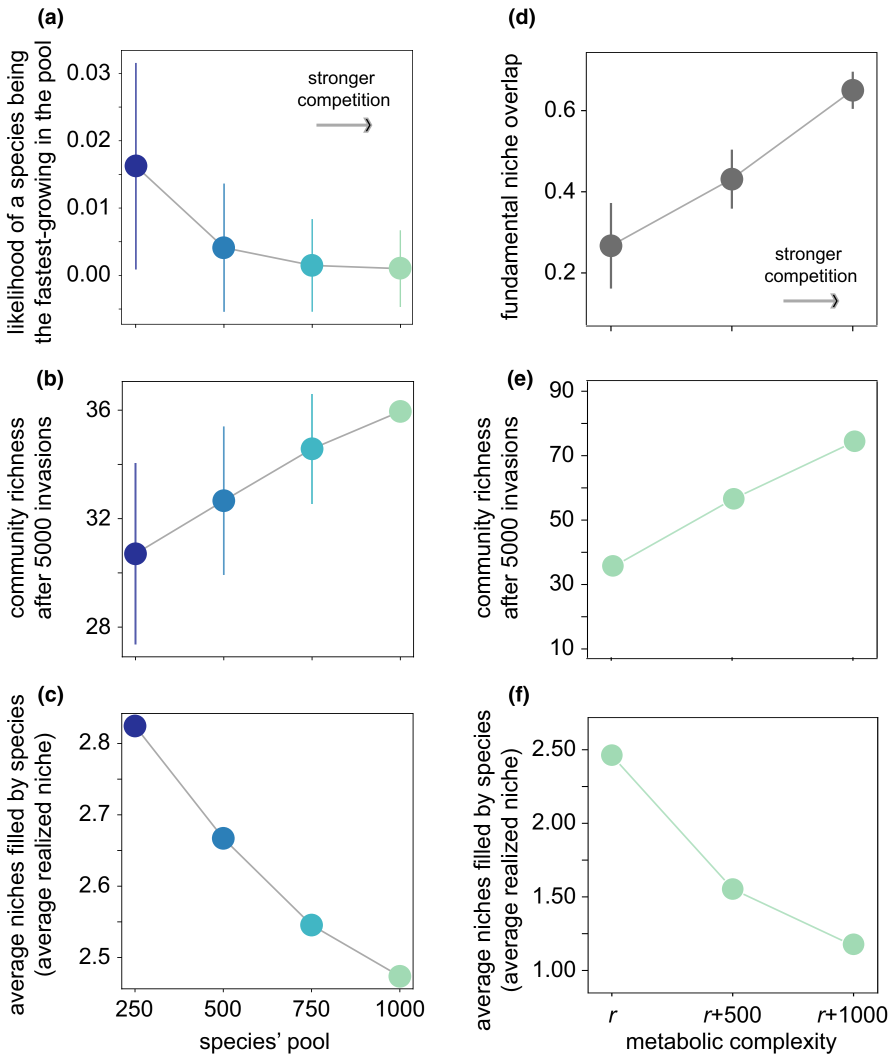

FIGURE 4 Competition causes a narrower division of available niches. (a–c) The horizontal axis shows the number of species in the species pool from which communities are assembled. Larger species pools imply higher competition among species. Vertical axes show (a) the probability of a species to be the fastest growing on a carbon source, averaged across carbon sources and across the whole pool, (b) community richness (number of species) and (c) average number of niches filled by each species (average realized niche, quantified as the number of available niches minus the number of free niches and divided by the total number of species in a community). (d–f) The horizontal axis shows the metabolic complexity of the species used for community assembly, that is their total number of biochemical reactions relative to the reference $r = 2583$ of *Escherichia coli* model iJO1366. (d) The vertical axis shows the average overlap in fundamental niche between two species pairs, quantified as the number of carbon sources on which both species can grow, divided by the union of carbon sources on which either species can grow. Higher metabolic complexity resulted in higher niche overlap between species, which increases the potential for competitive interactions between species. (e) Community richness and (f) average realized niche breadth (quantified as in c). Arrows in (a) and (d) indicate the direction in which competition gets stronger. Circles in (b), (c), (e) and (f) show averages over 500 assembled communities after simulating 5000 invasion events per community. Circle colours indicate the size of the species pool used to assemble the communities (a colour change from blue to green corresponds to an increase in the pool size from 250 to 1000 species). Bars show one standard deviation

consume, on average, fewer carbon sources. (To determine the average realized niche we divided the number of filled niches by community richness.) Altogether, the larger community richness observed when competition is stronger results from more niches being available and filled, and from the fact that each species occupies fewer niches (consumes fewer resources).

A second way to vary competition strength is to alter the degree of overlap in the fundamental niches that two species occupy (i.e., the number of carbon sources that they can thrive on). Species with more similar fundamental niches show more intense competition (Bailey et al., 2013). We can manipulate this niche overlap by altering the metabolic complexity of the species that we assemble into communities, that is by altering the number of biochemical reactions in each metabolic network, which is possible with our sampling approach (see Section 2). To this end, we created pools of metabolic networks that had 500 or 1000 more reactions than the *E. coli* model we used as a reference ("*r*", iJO1366; Orth et al., 2011). We designated members of these pools as having metabolic complexity "*r*+500" and "*r*+1000," respectively. We used species of equal metabolic complexity to assemble communities, selecting species from a pool of 1000 species. (Figure S4 shows analogous results for pool sizes of 250, 500 and 750 species.)

Species in these pools are on average viable on more carbon sources—they have a broader fundamental niche—than species with lower complexity (Figure S1). In addition, species pairs in these pools also overlap to a greater extent in their fundamental niches. We quantified this niche overlap by enumerating the number of carbon sources on which both species can grow, and divided this number by the total number of carbon sources on which at least one of the species can grow (Figure 4d). In sum, increased metabolic complexity increases competition among species by increasing their niche overlap.

We found that communities assembled from species with high metabolic complexity are more species-rich (Figure 4e). However, the number of available or free niches does not change when increasing metabolic complexity above that of *r* ("*r*+500" and "*r*+1000" in Figure S4b,c). Instead, the larger community diversity observed at high metabolic complexity results from a narrower partition of niche space among the species (Figure 4f).

We also explored the consequences of reducing metabolic complexity below that of *r*. We assembled communities with species that contain 500 or 1000 fewer reactions than the reference model (pools designated "*r*+500" and "*r*+1000"). The results are more difficult to interpret than those for metabolic complexity higher than *r* (see Figure S4 and Text S2). However, just as in the analysis where we changed the species' pool, a combination of two factors explains community richness. These are the total number of niches filled and the average realized niche (see Text S2 for a more detailed explanation).

In sum, we observe that strong competition results in a division of niche space into more and narrower niches that are occupied by the resident species. Everything else being equal, this results in more diverse communities.

4 | DISCUSSION

The "ecological limits" and "diversity begets diversity" hypotheses were proposed to explain the forces that shape biodiversity on Earth. Several analyses aiming to validate these hypotheses have examined the rate of evolutionary diversification to provide support for one or other hypothesis (Brockhurst et al., 2007; Gómez & Buckling, 2013; Jousset et al., 2016). However, evolution is not the only means by which community diversity can increase, because new species can also increase diversity by invading a community. We therefore feel that the hypotheses should be distinguished by considering a scenario that considers the contributions of both ecological and evolutionary processes to community diversity. A modelling approach like ours is well suited for this, because the species we simulate could be thought as being new evolved species or species that invade a community as they disperse from elsewhere. With these considerations in mind, we here simulate the assembly of thousands of communities in which we can control the number of new species allowed to invade a community and monitor the community's response to each invasion.

Our results suggest that the dichotomy between the diversity-begets-diversity and the ecological limits hypothesis is a false one. Both hypotheses capture important aspects of diversification. At initial stages of community assembly, species–species interactions in the form of cross-feeding and competition facilitate diversification. Species "construct" new niches through their metabolic excretions, which creates opportunities for the invasion of new species that thrive on these excretions (Figure 2). Competition helps partition niche space into more and narrower niches, which can result in larger diversity (Figure 4). At later stages in the assembly process, no new niches can be created, niche space cannot be partitioned further, successful species invasions become rare and diversity reaches a ceiling. In sum, the diversity-begets-diversity and the ecological limits scenarios best capture different stages—early vs. late—of community assembly. Community diversity has limits, but these limits are not just externally imposed. They can be modified by the species themselves.

Two other conclusions emerge from our results. First, even when all niches have become filled during community assembly, diversity may not have reached its ceiling. Competition can further increase community richness even in this case, because it can help narrow the realized niche of individual species (Figure S2). Second, our conclusions reinforce the argument that it is best to consider ecological processes separately from evolutionary processes when studying diversification. The reason is that the probability of successful invasions decreases even early during community assembly, while species are still creating new niches during the "diversity begets diversity" regime (Figure 2d; Figures S3, S5). This observation, which agrees with existing theory (Case, 1990; Tilman, 2004; Vila et al., 2019), experimental observations (Fargione et al., 2003; Stachowicz et al., 1999) and the empirical observation that diversification also slows down on evolutionary timescales (Condamine et al., 2019), implies that a slowdown in diversification (or speciation) is by itself not sufficient proof for the ecological limits hypothesis.

Competition is commonly thought to stimulate the diversification of species (Abrams, 2006; Bailey et al., 2013; Dieckmann & Doebeli, 1999; Meyer et al., 2016; Meyer & Kassen, 2007; Yoder et al., 2010), although it can also prevent such diversification (Bailey et al., 2013; Brockhurst et al., 2007; Gómez & Buckling, 2013). In our work, a successful invasion increases competition, because available resources have to be shared between more species. After such an invasion, communities may show an increase, no change or a net decrease in diversity if multiple species go extinct as a result of the invasion (Figure 2a), in agreement with empirical observations from the literature (Abrams, 2006; Bailey et al., 2013; Brockhurst et al., 2007; Dieckmann & Doebeli, 1999; Gómez & Buckling, 2013; Meyer et al., 2016; Meyer & Kassen, 2007; Yoder et al., 2010). However, such varying short-term consequences of competition for individual invasions need to be distinguished from the long-term consequences for an entire assembly process. At the end of this process (once community richness no longer increases), we find that communities with stronger competition have become more diverse (Figure 4). In other words, our work predicts that increased competition results in a net increase of diversity.

In an organism, many traits are interconnected. Traits that are advantageous for one aspect of fitness may trade-off with traits that affect other aspects. Such trade-offs are thought to be important for determining community structure (Begon et al., 2006; Johnson et al., 2012; Litchman & Klausmeier, 2008; Tilman, 1982). Our results also suggest—albeit indirectly—that trade-offs in resource use are important for the communities we assemble. If there were no such trade-offs, we would expect that one or a few species consume all available resources. Instead, we find a great diversity of species, which suggests the presence of trade-offs in resource use. It is important to point out that our modelling framework does not assume such trade-offs to exist. They arise from elementary biochemical principles embodied in the set of biochemical reactions that constitute a metabolic reaction network.

Our work has several limitations. Addressing them offers opportunities for future research. First, we study a very constrained subspace of Hutchinsonian niche space, the multidimensional space comprising all environmental factors required for the survival of a species (Hutchinson, 1957). Specifically, we only consider the sources of carbon and energy that sustain heterotrophic organisms, because they are frequently cross-fed between bacteria (D'Souza et al., 2018; McNally & Borenstein, 2018). Our modelling framework does not account for other abiotic factors, such as temperature and pH, nor does it account for non-nutrient-mediated species interactions, such as those that arise from cell-to-cell communication, or the secretion of toxic molecules. Community richness could increase above the limit we observe if such interactions help create even more niches. In addition, our knowledge about pan-metabolism is increasing, and newly characterized metabolic reactions could further increase a community's predicted potential to create more niches and diverse communities.

Second, we observe that different species grow on similar sets of carbon sources–their fundamental niches are similar—which may result from the method we used to create metabolic networks for community assembly. Specifically, we constrained these networks to be viable on glucose, which may have increased their niche similarity, because these networks must share subsets of biochemical reactions required for growth on glucose (Barve et al., 2012). To explore how this choice may have affected our analysis, we performed two complementary analyses. In a first analysis, we compared how the fundamental niche overlap of random viable networks compares to that in real organisms. For that, we used bacteria from the human gut (Magnúsdóttir et al., 2017). Specifically, we identified in the metabolic networks of 818 species those metabolites that could serve as sources of carbon, and quantified fundamental niche overlap based on the results. We found that the fundamental niche overlap between these species can range from zero (no overlap) and up to one (full overlap), like the range of overlap explored with random viable networks (see Text S1, Figures S7, S8). In a second analysis, we created random metabolic networks required to be viable on carbon sources other than glucose (acetate, pyruvate, serine, alanine and lactose) (Figure S6). Communities assembled using these networks differed little in community diversity, even when we assembled communities from species required to be viable on different carbon sources. Therefore, this limitation is not likely to affect our observations dramatically.

Third, we observed that communities can be composed of up to 80 species whose persistence depended, directly or indirectly, on the species consuming the initial resource, glucose. Unfortunately, our current modelling approach cannot answer how abundant these species are, and how much glucose would be required to sustain a highly diverse community. To answer these questions remains an important task for future work.

Fourth, we use the growth rate of species on single carbon sources to predict growth in a complex environment with multiple carbon sources. In such an environment, a microbial species growing in isolation may either consume a single (preferred) carbon source, several carbon sources, or all available carbon sources. Which of these strategies it pursues depends on multiple factors, including the species, the cultivation method (batch or chemostat) and the similarity between the carbon sources (Aidelberg et al., 2014; Egli et al., 1993; Kovárová-Kovar & Egli, 1998). The picture is even more complex in cocultures, where species interactions may affect consumption patterns. Our decision to use maximal growth on a single carbon source to model a species' persistence is motivated by recent experiments which show that the assembly of communities in mixtures of nutrients can be predicted from assembly in single-nutrient environments (Estrela, Sanchez-Gorostiaga et al., 2020; Fu et al., 2020). For example, the taxonomic structure of a community assembled in glucose medium (Goldford et al., 2018) can be explained by the growth of its constituent species in isolation, either on the externally provided glucose, or on the cross-fed metabolites acetate, succinate and lactate (Estrela, Vila, et al., 2020).

Metabolic modelling is a powerful modelling framework. It has been successfully used to study microbial physiology (Notebaart et al., 2008; Segrè et al., 2002; Shlomi et al., 2005; Varma & Palsson, 1994), to design strains for industrial and medical applications (Gu et al., 2019b; Lun et al., 2009; Mishra et al., 2018), and to explore questions in the field of evolution (Bajić et al., 2018; Barve & Wagner, 2013; William Harcombe et al., 2013; Ibarra et al., 2002; Notebaart et al., 2014; San Roman & Wagner, 2018, 2020; Sandberg et al., 2017) and ecology (Estrela, Sanchez-Gorostiaga et al., 2020; Harcombe et al., 2014; Levy & Borenstein, 2013; Machado et al., 2021; McNally & Borenstein, 2018; Zelezniak et al., 2015), as reviewed previously (García-Jiménez et al., 2021; Gu et al., 2019b; Mardinoglu & Nielsen, 2012). Though outside the scope of our work, the modelling framework used here can in principle also be used to study other ecological phenomena, such as successional dynamics (Bell & Pascual, 2020; Chase, 2003; Dini-Andreote et al., 2014; Lockwood et al., 1997;

Nemergut et al., 2007), and the role of historical contingency (or priority effects) in the assembly process (Chase, 2003; Fukami, 2015). What is more, whole-community genome sequencing, together with semi-automatic methods for metabolic reconstruction, facilitate the creation of genome-scale metabolic models for not just one organism, but for multiple organisms in a community (DeJongh et al., 2007; Dias et al., 2015; Mendoza et al., 2019; Wang et al., 2018). Thus, it may soon be possible to conduct an analysis like ours with metabolic networks characterized in a community from the wild. While such an analysis cannot control all the variables that our computational work can control, it can go beyond a proof of principle, explain actual limits on community diversity and identify rules of community assembly important in nature. It will be especially suited for the analysis of those complex communities where cross-feeding interactions are important, such as that of the human gut (Magnúsdóttir et al., 2017), the soil (Baran et al., 2015) or planktonic organisms (Enke et al., 2019).

ACKNOWLEDGEMENTS

We thank Owen Petchey for valuable discussions. This project received funding from the European Research Council under Grant Agreement No. 739874. We would also like to acknowledge support by Swiss National Science Foundation grant 31003A_172887 and by the University Priority Research Program in Evolutionary Biology. Open access funding provided by Universitat Zurich.

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

Conceived and designed the analyses: M.S.R., A.W. Performed the analyses: M.S.R. Analysed the data: M.S.R., A.W. Wrote the paper: M.S.R., A.W.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in GitHub at https://doi.org/10.5281/zenodo.5206841.

ORCID

Magdalena San Roman https://orcid.org/0000-0002-2593-7727
Andreas Wagner https://orcid.org/0000-0003-4299-3840

REFERENCES

Abrams, P. A. (2006). Adaptive change in the resource-exploitation traits of a generalist consumer: The ceolution and coexistence of generalists and specialists. *Evolution*, *60*(3), 427–439. https://doi.org/10.1111/j.0014-3820.2006.tb01124.x

Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., & Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia coli*. *BMC Systems Biology*, *8*(1), 133. https://doi.org/10.1186/s12918-014-0133-z

Allison, S. D. (2012). A trait-based approach for modelling microbial litter decomposition. *Ecology Letters*, *15*(9), 1058–1070. https://doi.org/10.1111/j.1461-0248.2012.01807.x

Bailey, S. F., Dettman, J. R., Rainey, P. B., & Kassen, R. (2013). Competition both drives and impedes diversification in a model adaptive radiation. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1766), 20131253. https://doi.org/10.1098/rspb.2013.1253

Bajić, D., Vila, J. C. C., Blount, Z. D., & Sánchez, A. (2018). On the deformability of an empirical fitness landscape by microbial evolution. *Proceedings of the National Academy of Sciences*, *115*(44), 11286–11291. https://doi.org/10.1073/pnas.1808485115

Baran, R., Brodie, E. L., Mayberry-Lewis, J., Hummel, E., Da Rocha, U. N., Chakraborty, R., Bowen, B. P., Karaoz, U., Cadillo-Quiroz, H., Garcia-Pichel, F., & Northen, T. R. (2015). Exometabolite niche partitioning among sympatric soil bacteria. *Nature Communications*, *6*(1), 8289. https://doi.org/10.1038/ncomms9289

Barve, A., Rodrigues, J. F. M., & Wagner, A. (2012). Superessential reactions in metabolic networks. *Proceedings of the National Academy of Sciences*, *109*(18), 6810. https://doi.org/10.1073/pnas.1113065109

Barve, A., & Wagner, A. (2013). A latent capacity for evolutionary innovation through exaptation in metabolic systems. *Nature*. https://doi.org/10.1038/nature12301

Begon, M., Townsend, C. R., & Harper, J. L. (2006). *Ecology: From individuals to ecosystems*. , 4th ed. Blackwell Pub.

Bell, T., & Pascual-García, A. (2020). Community-level signatures of ecological succession in natural bacterial communities. *Nature Communications*.

Brockhurst, M. A., Colegrave, N., Hodgson, D. J., & Buckling, A. (2007). Niche occupation limits adaptive radiation in experimental microcosms. *PLoS One*, *2*(2), e193. https://doi.org/10.1371/journal.pone.0000193

Calcagno, V., Jarne, P., Loreau, M., Mouquet, N., & David, P. (2017). Diversity spurs diversification in ecological communities. *Nature Communications*, *8*, 15810. https://doi.org/10.1038/ncomms15810

Case, T. J. (1990). Invasion resistance arises in strongly interacting species-rich model competition communities. *Proceedings of the National Academy of Sciences*, *87*(24), 9610. https://doi.org/10.1073/pnas.87.24.9610

Chase, J. M. (2003). Community assembly: When should history matter? *Oecologia*, *136*(4), 489–498. https://doi.org/10.1007/s00442-003-1311-7

Condamine, F. L., Rolland, J., & Morlon, H. (2019). Assessing the causes of diversification slowdowns: Temperature-dependent and diversity-dependent models receive equivalent support. *Ecology Letters*, *22*(11), 1900–1912. https://doi.org/10.1111/ele.13382

D'Souza, G., Shitut, S., Preussger, D., Yousif, G., Waschina, S., & Kost, C. (2018). Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural Product Reports*, *35*(5), 455–488. https://doi.org/10.1039/C8NP00009C

Daniel, R. (2005). The metagenomics of soil. *Nature Reviews Microbiology*, *3*(6), 470–478. https://doi.org/10.1038/nrmicro1160

DeJongh, M., Formsma, K., Boillot, P., Gould, J., Rycenga, M., & Best, A. (2007). Toward the automated generation of genome-scale metabolic networks in the SEED. *BMC Bioinformatics*, *8*, 139. https://doi.org/10.1186/1471-2105-8-139

Dias, O., Rocha, M., Ferreira, E. C., & Rocha, I. (2015). Reconstructing genome-scale metabolic models with merlin. *Nucleic Acids Research*, *43*(8), 3899–3910. https://doi.org/10.1093/nar/gkv294

Dieckmann, U., & Doebeli, M. (1999). On the origin of species by sympatric speciation. *Nature*, *400*(6742), 354–357. https://doi.org/10.1038/22521

Dini-Andreote, F., de Cássia Pereira e Silva, M., Triadó-Margarit, X., Casamayor, E. O., van Elsas, J. D., & Salles, J. F. (2014). Dynamics of bacterial community succession in a salt marsh chronosequence: Evidences for temporal niche partitioning. *The ISME Journal*, *8*(10), 1989–2001. https://doi.org/10.1038/ismej.2014.54

Ebrahim, A., Lerman, J. A., Palsson, B. O., & Hyduke, D. R. (2013). COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, *7*(1), 74. https://doi.org/10.1186/1752-0509-7-74

Edwards, J. S., Ibarra, R. U., & Palsson, B. O. (2001). In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature Biotechnology*, *19*(2), 125–130. https://doi.org/10.1038/84379

Egli, T., Lendenmann, U., & Snozzi, M. (1993). Kinetics of microbial growth with mixtures of carbon sources. *Antonie Van Leeuwenhoek*, *63*(3–4), 289–298. https://doi.org/10.1007/BF00871224

Emerson, B. C., & Kolm, N. (2005). Species diversity can drive speciation. *Nature*, *434*(7036), 1015–1017. https://doi.org/10.1038/nature03450

Enke, T. N., Datta, M. S., Schwartzman, J., Cermak, N., Schmitz, D., Barrere, J., Pascual-García, A., & Cordero, O. X. (2019). Modular assembly of polysaccharide-degrading marine microbial communities. *Current Biology*, *29*(9), 1528–1535.e6. https://doi.org/10.1016/j.cub.2019.03.047

Erwin, D. H. (2008). Macroevolution of ecosystem engineering, niche construction and diversity. *Trends in Ecology & Evolution*, *23*(6), 304–310. https://doi.org/10.1016/j.tree.2008.01.013

Estrela, S., Sanchez-Gorostiaga, A., Vila, J. C. C., & Sanchez, A. (2020). Nutrient dominance governs the assembly of microbial communities in mixed nutrient environments. *BioRxiv*. https://doi.org/10.1101/2020.08.06.239897

Estrela, S., Vila, J. C. C., Lu, N., Bajic, D., Rebolleda-Gomez, M., Chang, C.-Y., & Sanchez, A. (2020). Metabolic rules of microbial community assembly. *BioRxiv*. https://doi.org/10.1101/2020.03.09.984278

Fargione, J., Brown, C. S., & Tilman, D. (2003). Community assembly and invasion: An experimental test of neutral versus niche processes. *Proceedings of the National Academy of Sciences*, *100*(15), 8916. https://doi.org/10.1073/pnas.1033107100

Fu, H., Uchimiya, M., Gore, J., & Moran, M. A. (2020). Ecological drivers of bacterial community assembly in synthetic phycospheres. *Proceedings of the National Academy of Sciences*, *117*(7), 3656. https://doi.org/10.1073/pnas.1917265117

Fukami, T. (2015). Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 1–23. https://doi.org/10.1146/annurev-ecolsys-110411-160340

García-Jiménez, B., Torres-Bacete, J., & Nogales, J. (2021). Metabolic modelling approaches for describing and engineering microbial communities. *Computational and Structural Biotechnology Journal*, *19*, 226–246. https://doi.org/10.1016/j.csbj.2020.12.003

Goldford, J. E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P., & Sanchez, A. (2018). Emergent simplicity in microbial community assembly. *Science*, *361*(6401), 469. https://doi.org/10.1126/science.aat1168

Gómez, P., & Buckling, A. (2013). Real-time microbial adaptive diversification in soil. *Ecology Letters*, *16*(5), 650–655. https://doi.org/10.1111/ele.12093

Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., & Lee, S. Y. (2019a). Current status and applications of genome-scale metabolic models. *Genome Biology*, *20*(1), 121. https://doi.org/10.1186/s13059-019-1730-3

Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., & Lee, S. Y. (2019b). Current status and applications of genome-scale metabolic models. *Genome Biology*, *20*(1), 121. https://doi.org/10.1186/s13059-019-1730-3

Harcombe, W. R., Delaney, N. F., Leiby, N., Klitgord, N., & Marx, C. J. (2013). The ability of flux balance analysis to predict evolution of central metabolism scales with the initial distance to the optimum. *PLoS Computational Biology*, *9*(6), e1003091. https://doi.org/10.1371/journal.pcbi.1003091

Harcombe, W. R., Riehl, W. J., Dukovski, I., Granger, B. R., Betts, A., Lang, A. H., Bonilla, G., Kar, A., Leiby, N., Mehta, P., Marx, C. J., & Segre, D. (2014). Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Reports*, *7*(4), 1104–1115. https://doi.org/10.1016/j.celrep.2014.03.070

Hardin, G. (1960). The competitive exclusion principle. *Science*, *131*(3409), 1292. https://doi.org/10.1126/science.131.3409.1292

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, *22*, 415–427. https://doi.org/10.1101/SQB.1957.022.01.039

Hutchinson, G. E. (1959). Homage to Santa Rosalia or why are there so many kinds of animals? *The American Naturalist*, *93*(870), 145–159.

Ibarra, R., Edwards, J., & Palsson, B. O. (2002). Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*, *420*(6912), 186–189 https://doi.org/10.1038/nature01149

Johnson, D. R., Goldschmidt, F., Lilja, E. E., & Ackermann, M. (2012). Metabolic specialization and the assembly of microbial communities. *The Isme Journal*, *6*, 1985. https://doi.org/10.1038/ismej.2012.46

Jousset, A., Eisenhauer, N., Merker, M., Mouquet, N., & Scheu, S. (2016). High functional diversity stimulates diversification in experimental microbial communities. *Science Advances*, *2*(6), e1600124. https://doi.org/10.1126/sciadv.1600124

Kovárová-Kovar, K., & Egli, T. (1998). Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed-substrate kinetics. *Microbiology and Molecular Biology Reviews*, *62*(3), 646–666. https://doi.org/10.1128/MMBR.62.3.646-666.1998

Krebs, C. J. (2014). *Ecology: The experimental analysis of distribution and abundance*. , 6th ed. Pearson.

Laland, K. N., Odling-Smee, F. J., & Feldman, M. W. (1999). Evolutionary consequences of niche construction and their implications for ecology. *Proceedings of the National Academy of Sciences*, *96*(18), 10242–10247. https://doi.org/10.1073/pnas.96.18.10242

Laughlin, D. C., Joshi, C., van Bodegom, P. M., Bastow, Z. A., & Fulé, P. Z. (2012). A predictive model of community assembly that incorporates intraspecific trait variation. *Ecology Letters*, *15*(11), 1291–1299. https://doi.org/10.1111/j.1461-0248.2012.01852.x

Levy, R., & Borenstein, E. (2013). Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proceedings of the National Academy of Sciences*, *110*(31), 12804–12809. https://doi.org/10.1073/pnas.1300926110

Litchman, E., & Klausmeier, C. A. (2008). Trait-based community ecology of phytoplankton. *Annual Review of Ecology, Evolution, and Systematics*, *39*(1), 615–639. https://doi.org/10.1146/annurev.ecolsys.39.110707.173549

Lockwood, J. L., Powell, R. D., Nott, M. P., & Pimm, S. L. (1997). Assembling ecological communities in time and space. *Oikos*, *80*(3), 549–553. https://doi.org/10.2307/3546628

Louca, S., Shih, P. M., Pennell, M. W., Fischer, W. W., Parfrey, L. W., & Doebeli, M. (2018). Bacterial diversification through geological time. *Nature Ecology & Evolution*, *2*(9), 1458–1467. https://doi.org/10.1038/s41559-018-0625-0

Lun, D. S., Rockwell, G., Guido, N. J., Baym, M., Kelner, J. A., Berger, B., Galagan, J. E., & Church, G. M. (2009). Large-scale identification of genetic design strategies using local search. *Molecular Systems Biology*, *5*(1), 296. https://doi.org/10.1038/msb.2009.57

Machado, D., Maistrenko, O. M., Andrejev, S., Kim, Y., Bork, P., Patil, K. R., & Patil, K. R. (2021). Polarization of microbial communities between competitive and cooperative metabolism. *Nature Ecology & Evolution*, *5*(2), 195–203. https://doi.org/10.1038/s41559-020-01353-4

Madi, N., Vos, M., Murall, C. L., Legendre, P., & Shapiro, B. J. (2020). Does diversity beget diversity in microbiomes? *Elife*, *9*, e58999. https://doi.org/10.7554/eLife.58999

Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D. A., Bauer, E., Noronha, A., Greenhalgh, K., Jäger, C., Baginska, J., Wilmes, P., Fleming, R. M. T., & Thiele, I. (2017). Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature Biotechnology*, *35*(1), 81–89. https://doi.org/10.1038/nbt.3703

Mardinoglu, A., & Nielsen, J. (2012). Systems medicine and metabolic modelling: Key Symposium: Systems medicine and metabolic modelling. *Journal of Internal Medicine*, *271*(2), 142–154. https://doi.org/10.1111/j.1365-2796.2011.02493.x

Mcgill, B., Enquist, B., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, *21*(4), 178–185. https://doi.org/10.1016/j.tree.2006.02.002

McNally, C. P., & Borenstein, E. (2018). Metabolic model-based analysis of the emergence of bacterial cross-feeding via extensive gene loss. *BMC Systems Biology*, *12*(1), 69. https://doi.org/10.1186/s12918-018-0588-4

Mendoza, S. N., Olivier, B. G., Molenaar, D., & Teusink, B. (2019). A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biology*, *20*(1), 158. https://doi.org/10.1186/s13059-019-1769-1

Meyer, J. R., Dobias, D. T., Medina, S. J., Servilio, L., Gupta, A., & Lenski, R. E. (2016). Ecological speciation of bacteriophage lambda in allopatry and sympatry. *Science*, *354*(6317), 1301–1304. https://doi.org/10.1126/science.aai8446

Meyer, J. R., & Kassen, R. (2007). The effects of competition and predation on diversification in a model adaptive radiation. *Nature*, *446*(7134), 432–435. https://doi.org/10.1038/nature05599

Mishra, P., Lee, N.-R., Lakshmanan, M., Kim, M., Kim, B.-G., & Lee, D.-Y. (2018). Genome-scale model-driven strain design for dicarboxylic acid production in Yarrowia lipolytica. *BMC Systems Biology*, *12*(S2), 12. https://doi.org/10.1186/s12918-018-0542-5

Nemergut, D. R., Anderson, S. P., Cleveland, C. C., Martin, A. P., Miller, A. E., Seimon, A., & Schmidt, S. K. (2007). Microbial community succession in an unvegetated. *Recently Deglaciated Soil. Microbial Ecology*, *53*(1), 110–122. https://doi.org/10.1007/s00248-006-9144-7

Notebaart, R. A., Szappanos, B., Kintses, B., Pal, F., Gyorkei, A., Bogos, B., Lazar, V., Spohn, R., Csorg, B., Wagner, A., Ruppin, E., Pal, C., & Papp, B. (2014). Network-level architecture and the evolutionary potential of underground metabolism. *Proceedings of the National Academy of Sciences*, *111*(32), 11762–11767. https://doi.org/10.1073/pnas.1406102111

Notebaart, R. A., Teusink, B., Siezen, R. J., & Papp, B. (2008). Co-regulation of metabolic genes is better explained by flux coupling than by network distance. *PLOS Computational Biology*, *4*(1), e26. https://doi.org/10.1371/journal.pcbi.0040026

Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., & Palsson, B. Ø.(2011). A comprehensive genome-scale reconstruction of Escherichia coli metabolism–2011. *Molecular System Biology*, *7*, 535.

Orth, J. D., Thiele, I., & Palsson, B. O. (2010). What is flux balance analysis? *Nature Biotechnology*, *28*(3), 245–248. https://doi.org/10.1038/nbt.1614

Paczia, N., Nilgen, A., Lehmann, T., Gätgens, J., Wiechert, W., & Noack, S. (2012). Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms. *Microbial Cell Factories*, *11*(1), 122. https://doi.org/10.1186/1475-2859-11-122

Palmer, M. W., & Maurer, T. A. (1997). Does diversity beget diversity? A case study of crops and weeds. *Journal of Vegetation Science*, *8*(2), 235–240. https://doi.org/10.2307/3237352

Pinu, F. R., Granucci, N., Daniell, J., Han, T.-L., Carneiro, S., Rocha, I., Nielsen, J., & Villas-Boas, S. G. (2018). Metabolite secretion in microorganisms: The theory of metabolic overflow put to the test. *Metabolomics*, *14*(4), 43. https://doi.org/10.1007/s11306-018-1339-7

Price, T. D., Hooper, D. M., Buchanan, C. D., Johansson, U. S., Tietze, D. T., Alström, P., Olsson, U., Ghosh-Harihar, M., Ishtiaq, F., Gupta, S. K., Martens, J., Harr, B., Singh, P., & Mohan, D. (2014). Niche filling slows the diversification of Himalayan songbirds. *Nature*, *509*(7499), 222–225. https://doi.org/10.1038/nature13272

Rabosky, D. L., & Hurlbert, A. H. (2015). Species Richness at Continental Scales Is Dominated by Ecological Limits. *The American Naturalist*, *185*(5), 572–583. https://doi.org/10.1086/680850

Samal, A., Rodrigues, J. F. M., Jost, J., Martin, O. C., & Wagner, A. (2010). Genotype networks in metabolic reaction spaces. *BMC Systems Biology*, *4*(1). https://doi.org/10.1186/1752-0509-4-30

San Roman, M., & Wagner, A. (2018). An enormous potential for niche construction through bacterial cross-feeding in a homogeneous environment. *PLOS Computational Biology*, *14*(7), e1006340. https://doi.org/10.1371/journal.pcbi.1006340

San Roman, M., & Wagner, A. (2020). Acetate and glycerol are not uniquely suited for the evolution of cross-feeding in *E. coli*. *PLOS Computational Biology*, *16*(11), e1008433. https://doi.org/10.1371/journal.pcbi.1008433

Sandberg, T. E., Lloyd, C. J., Palsson, B. O., & Feist, A. M. (2017). Laboratory evolution to alternating substrate environments yields distinct phenotypic and genetic adaptive strategies. *Applied and Environmental Microbiology*, *83*(13), e00410–e417. https://doi.org/10.1128/AEM.00410-17

Schluter, D., & Pennell, M. W. (2017). Speciation gradients and the distribution of biodiversity. *Nature*, *546*, 48. https://doi.org/10.1038/nature22897

Segrè, D., Vitkup, D., & Church, G. M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences*, *99*(23), 15112. https://doi.org/10.1073/pnas.232349399

Shlomi, T., Berkman, O., & Ruppin, E. (2005). Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(21), 7695. https://doi.org/10.1073/pnas.0406346102

Stachowicz, J. J., Whitlatch, R. B., & Osman, R. W. (1999). Species diversity and invasion resistance in a marine ecosystem. *Science*, *286*(5444), 1577. https://doi.org/10.1126/science.286.5444.1577

Thakur, M. P., & Wright, A. J. (2017). Environmental filtering, niche construction, and trait variability: The missing discussion. *Trends in Ecology & Evolution*, *32*(12), 884–886. https://doi.org/10.1016/j.tree.2017.09.014

Tilman, D. (1982). *Resource competition and community structure*. Princeton University Press.

Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(30), 10854. https://doi.org/10.1073/pnas.0403458101

Varma, A., & Palsson, B. O. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Applied and Environment Microbiology*, *60*(10), 3724–3731. https://doi.org/10.1128/aem.60.10.3724-3731.1994

Vila, J. C. C., Jones, M. L., Patel, M., Bell, T., & Rosindell, J. (2019). Uncovering the rules of microbial community invasions. *Nature Ecology & Evolution*, *3*(8), 1162–1171. https://doi.org/10.1038/s41559-019-0952-9

Wang, H., Marcišauskas, S., Sánchez, B. J., Domenzain, I., Hermansson, D., Agren, R., Nielsen, J., & Kerkhoven, E. J. (2018). RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*. *PLOS Computational Biology*, *14*(10), e1006541. https://doi.org/10.1371/journal.pcbi.1006541

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, *21*(2/3), 213–251. https://doi.org/10.2307/1218190

Yoder, J. B., Clancey, E., Des Roches, S., Eastman, J. M., Gentry, L., Godsoe, W., Hagey, T. J., Jochimsen, D., Oswald, B. P., Robertson, J., Sarver, B. A. J., Schenk, J. J., Spear, S. F., & Harmon, L. J. (2010). Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, *23*(8), 1581–1596. https://doi.org/10.1111/j.1420-9101.2010.02029.x

Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D. R., Bork, P., & Patil, K. R. (2015). Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences*, *112*(20), 6449–6454. https://doi.org/10.1073/pnas.1421834112

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: San Roman, M., & Wagner, A. (2021). Diversity begets diversity during community assembly until ecological limits impose a diversity ceiling. *Molecular Ecology*, 30, 5874–5887. https://doi.org/10.1111/mec.16161