## ARTICLE

**OPEN**

Check for updates

# Contingent evolution of alternative metabolic network topologies determines whether cross-feeding evolves

Jeroen Meijer [1,2✉], Bram van Dijk[1,2] & Paulien Hogeweg[1]

Metabolic exchange is widespread in natural microbial communities and an important driver of ecosystem structure and diversity, yet it remains unclear what determines whether microbes evolve division of labor or maintain metabolic autonomy. Here we use a mechanistic model to study how metabolic strategies evolve in a constant, one resource environment, when metabolic networks are allowed to freely evolve. We find that initially identical ancestral communities of digital organisms follow different evolutionary trajectories, as some communities become dominated by a single, autonomous lineage, while others are formed by stably coexisting lineages that cross-feed on essential building blocks. Our results show how without presupposed cellular trade-offs or external drivers such as temporal niches, diverse metabolic strategies spontaneously emerge from the interplay between ecology, spatial structure, and metabolic constraints that arise during the evolution of metabolic networks. Thus, in the long term, whether microbes remain autonomous or evolve metabolic division of labour is an evolutionary contingency.

[1] Theoretical Biology and Bioinformatics, Department of Biology, Utrecht University, Padualaan 8, Utrecht 3584 CH, The Netherlands. [2] These authors contributed equally: Jeroen Meijer, Bram van Dijk. ✉email: j.meijer1@uu.nl

Natural microbial communities are typically complex, composed of many different taxonomic groups that stably coexist. The complexity of these communities might partly reflect the complexity of their environment, which can generate and maintain diversity by allowing specialisation on different niches[1–10]. Experimental evolution has shown how initially clonal populations can adaptively diversify into stably coexisting ecotypes, each specialised on pre-existing niches defined by available nutrients[1], spatial structure[2,11], temporal variability such as the feast and famine cycles in serial transfer[3–9], or combinations thereof[10]. However, even in constant, unstructured environments with a single limiting carbon source metabolic diversification routinely evolves[4,12–14]. Here, new niches are constructed by microbes themselves, as metabolic byproducts released by one become the growth substrate for another, allowing stable coexistence mediated by metabolic interactions. For example, initially clonal *Escherichia coli* populations grown in a glucose-limited chemostat genetically diversify into a lineage that rapidly but inefficiently grows on the provided resource, and lineages that specialise in using the overflow acetate produced by the first lineage[4,13,15]. Niche construction can thus lead to stable coexistence even with a single resource, which according to the competitive exclusion principle[16] would support only a single species.

Several mechanisms have been proposed that might explain both the evolution of cross-feeding in simple, constant environments as well as the prevalence of cross-feeding in natural communities[17–22] (see ref. [23] for a review). For example, cellular or metabolic trade-offs might favour metabolic specialisation[24,25], and division of labour might increase productivity of a community[26]. Alternatively, the "Black Queen Hypothesis" holds that gene loss can be adaptive for functions that are costly, leaky and essential—such as the production of essential building blocks that end up in the environment through diffusion or lysis—provided other community members compensate for the lost function[27–29]. Indeed, by removing two metabolic genes in *E. coli*, a recent study showed that in a synthetic community with engineered obligate dependencies for amino acids, strains grow up to 20% faster compared with the autonomous wild-type, suggesting positive selection for the loss of biosynthetic genes[18]. Cross-feeding might thus be caused by a gene loss ratchet in a community of initially autonomous microbes that, driven by escaping the burden of biosynthesis genes, evolve complementary metabolic networks and become dependent[30,31].

Theoretical studies on the evolution of metabolic dependency typically predefine a limited number of metabolic strategies[32], take cellular trade-offs for granted[25,32–36], or explicitly assume external forcing such as seasonality[37,38], and investigate how other conditions favour autonomous or cross-feeding strategies, or prevent the evolution of metabolic free-loaders[39]. For example, by constructing genome-scale metabolic models and limiting the number of reactions that can be performed by a single individual, under certain conditions multiple cross-feeding strategies together can outperform an autonomous strategy[26]. However, these trade-offs are not always necessarily an inescapable physical reality that must be faced, but are themselves evolved cellular properties[40], and little attention is given to the evolution *of* trade-offs compared with the evolution *on* trade-offs.

In addition, these models tend to assume ecological and evolutionary timescales are separated—the fate of each individual mutation is determined by playing out ecological dynamics before the next mutation arrives. We know from experimental studies[9,41–45] that microbial communities do not operate in such a mutation-limited domain, and instead, multiple mutations continuously emerge and cause complex evolutionary dynamics and contingency. Theoretical work[46,47] demonstrated that under such true "eco-evolutionary" conditions, evolutionary dynamics qualitatively change and result in more complex and diverse ecosystems.

Here, we investigate how different metabolic strategies can evolve using a bottom-up, mechanistic model of microbial eco-evolutionary dynamics that explicitly accounts for the high diversity in communities without assuming cellular trade-offs or external drivers for metabolic division of labour. We show that initially identical populations—when propagated under identical conditions—can reach two qualitatively different eco-evolutionary attractors: a community of cross-feeding metabolic specialists or, alternatively, a community formed by a single lineage of microbes that are metabolically autonomous, producing all metabolites they need. Which type of community evolves is dependent on, and can be predicted from, a "frozen metabolic accident": the topology of the metabolic network that typically fixes much earlier in evolution. Differences between these topologies appear to be neutral across populations at time of fixation, but have far reaching repercussions for subsequent evolution.

## Results

**Model overview**. See "Methods" for a detailed model description and parameters.

In the model, we explicitly incorporate a "chemical universe", metabolism, cell growth and division, genome evolution and a two-dimensional spatial environment that all co-evolve (Fig. 1). We do not predefine fitness (such as a target genotype or biomass reaction to be optimised), but set basic rules for cell growth, reproduction and death. This means whether a mutation is beneficial, neutral or deleterious depends on local environmental conditions and interactions, cellular state and genomic background, all of which are shaped by prior evolution in the model. As a consequence, metabolic or ecological strategies are not predefined, but emerge during evolutionary simulations as microbes evolve and explore the possibilities of the chemical universe and reshape their local environment by metabolite uptake and exchange. This approach allows us to de novo create microbial communities with their own evolutionary histories and study them with access to a perfect digital "fossil record".

Microbes compete for a single resource molecule $R$ and limited space on a 2d grid, reproducing locally into empty neighbouring sites (Fig. 1a, b). They require two essential (non-substitutable) building block metabolites $B_1$, $B_2$ for cell growth and expressing proteins that perform metabolic functions. Building blocks do not natively occur in the environment, but can be synthesised from the provided resource by expressing relevant metabolic proteins. In addition to building blocks, microbes require energy metabolite $E$ to operate transporter proteins to pump metabolites (such as the provided resource) in and out of the cell. The chemical universe available for evolution to meet these metabolic demands consists of a predefined set of nine metabolites ($R$, $B_1$, $B_2$, $M_{1-5}$ and $E$) connected by 59 reactions (43 conversion reactions and 16 transport reactions, see Methods), that contains many redundant pathways and provides many degrees of freedom to form functional metabolic networks.

Proteins catalyse individual reactions (e.g. $1R \rightarrow 1M_1 + 5E$) that can be combined to form metabolic pathways. They are coded on the microbe's genome, which typically covers only a small subset of all reactions. When cells reproduce the genome can mutate, allowing the metabolic networks to evolve by tuning the rates of individual reactions through point mutations (basal expression rate and kinetic parameters of the enzyme) and gene copy number (gene deletion, duplication). New pathways can be formed by discovering new genes or through horizontal gene transfer from nearby cells.
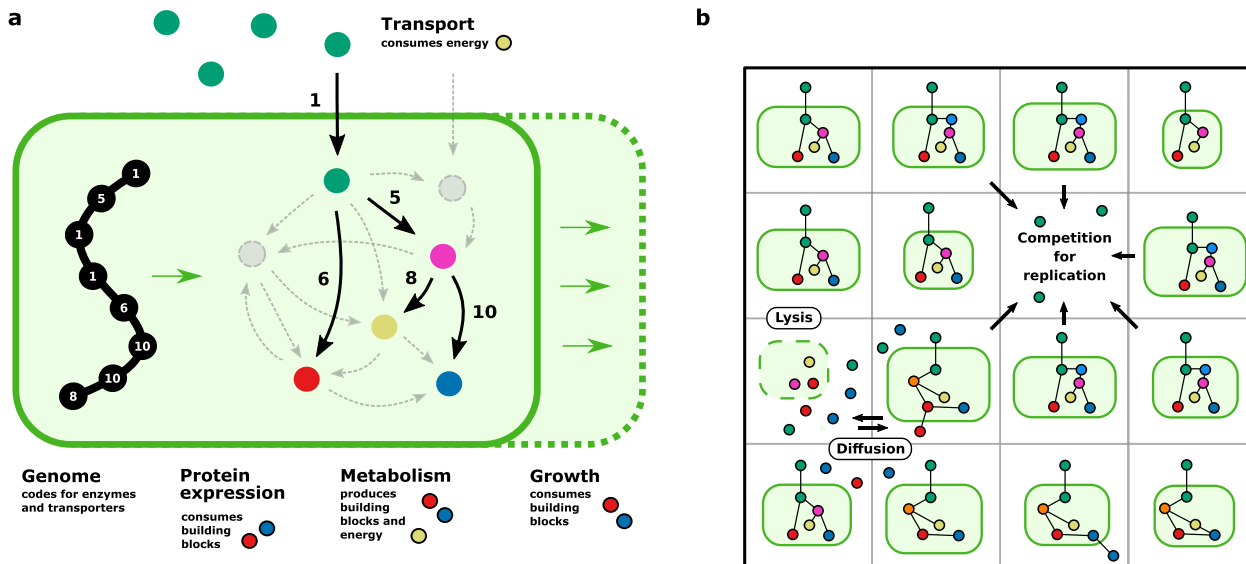
**Fig. 1 Model of microbial eco-evolutionary dynamics. a** Genes on a linear genome code for specific metabolic enzymes that catalyse individual reactions of the metabolic network. To express proteins and grow, microbes require two non-substitutable building block metabolites $B_1$ and $B_2$ (red, blue) that do not natively occur in their environment, but can be metabolised from the single provided resource $R$ (green) by expressing the right metabolic pathways. Active transport of metabolites across the cell membrane requires an energy metabolite $E$ (yellow). The genome of a single microbe typically covers a small subset of the complete "chemical universe" of 59 reactions (see Methods). **b** Microbes compete for space and metabolites on a $45 \times 45$ lattice. They can reproduce in an adjacent empty space if they meet the minimal division cell size. Here, microbes NE and W of the empty space are too small to reproduce. Upon replication, genomes can mutate through gene duplication and deletion, discovery of new genes, and point mutations that can change the expression rate and kinetic parameters of individual genes. New genes can also be acquired via horizontal gene transfer from nearby microbes. Active transport of metabolites and lysis changes the composition of a microbes' local environment.

Cells can reproduce in a neighbouring empty site if they meet a minimal division size (Fig. 1b), with competition biased towards larger cells when multiple cells are eligible. Cell death is modelled as a stochastic process with a basal death rate that is potentially elevated when internal metabolites reach toxic concentrations. Cell lysis releases all internal metabolites into the local environment, which then locally diffuse and become available for other nearby microbes to take up. In this way, microbes change the metabolite composition of their local environment through active transport, passive diffusion across the cell membrane and cell death (see Fig. 1b). Motivated by experimental work[11,48] that shows that microscale gradients quickly establish and influence microbial metabolism and community dynamics, we first consider evolution in a spatially structured environment with limited diffusion (mimicking biofilm conditions), and subsequently investigate evolution simulating a well-mixed medium.
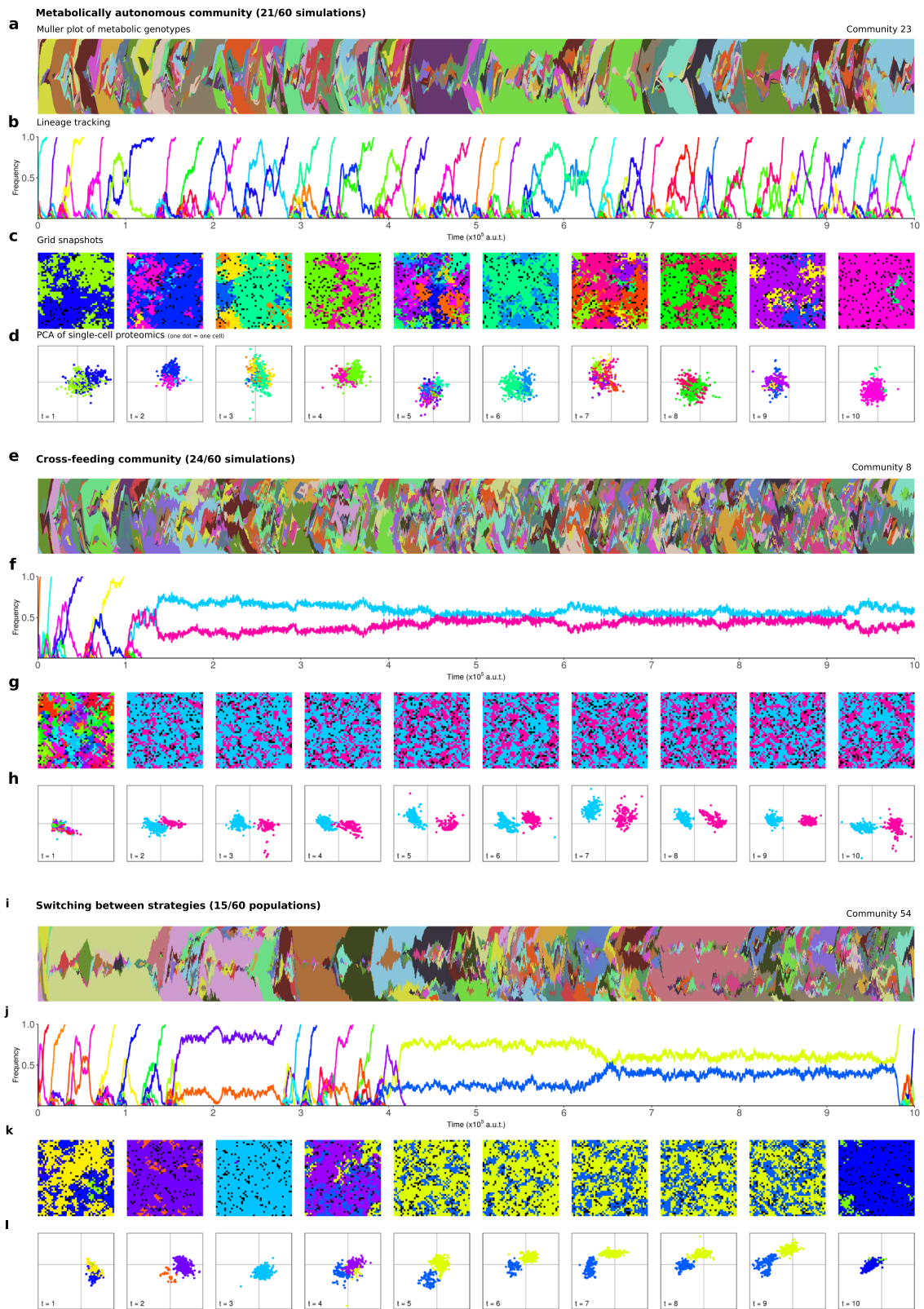
We constructed an initial population of "minimally viable" microbes by generating 2025 randomly parameterised genomes coding for metabolic networks that contain a food importer and randomly selected genes to produce both building blocks. We then evolved 60 identical copies of this population in parallel under the exact same conditions for $10^6$ time steps ($\sim 4 \times 10^5$ generations), while fluxing in food metabolite $R$ at a constant rate at all grid points. Using this model, we examine whether "ecosystem based" metabolic strategies evolve, i.e. cross-feeding species with complementary metabolic networks, or "individual-based" strategies in the form of autonomous microbes that produce all required building blocks.

**Diverse metabolic strategies evolve in a simple, constant environment**. We investigated the evolution of metabolic strategies with a mechanistic model, first focusing on the effect of contingency with a parallel evolution experiment. The ancestral

community consists of microbes with metabolic networks composed of a food importer and randomly selected genes to produce both building blocks, all of which have randomly sampled kinetic parameters and expression rates. During the simulations point mutations fix that tune fluxes through specific reactions, and metabolic networks are extended with reactions that are dedicated to producing energy—which allows increased food uptake—and reactions that process byproducts for more energy and/or building blocks. Furthermore, importers are recruited to recycle building blocks that accumulate in the environment through cell lysis. Thus eventually, all populations evolve efficient, closed metabolic networks that make use of all produced metabolites.

However, mutants with different metabolic repertoires continuously arise and compete for dominance within populations and all populations are highly diverse throughout the evolutionary simulations. As only very few genotypically identical individuals are present at any given time, we found it useful for interpretation and visualisation purposes to classify microbes based on their "metabolic genotype": a binary representation that indicates the presence or absence for each of the 59 metabolic genes and transporters in the genome. Tracking the abundances of these metabolic genotypes over time shows that the evolutionary dynamics are complex and characterised by clonal interference and frequent hitchhiking, leapfrogging and horizontal gene transfer (see Muller plots in Fig. 2).

To condense these complex dynamics, we use lineage tracking and analyse the cell proteomes of these lineages. This reveals that some of these heterogeneous communities are dominated by a single lineage that performs all the metabolic functions outlined above by itself (see Fig. 2a–d; Supplementary Fig. 1 for lineage markers for all 60 populations), while other communities diversify in two complimentary, cross-feeding lineages that specialise in producing one, and importing the other building block (Fig. 2e–h). These cross-feeding lineages form interleaved

**a** Metabolically autonomous community (21/60 simulations)
Muller plot of metabolic genotypes                                    Community 23



**b** Lineage tracking



**c** Grid snapshots



**d** PCA of single-cell proteomics (one dot = one cell)



**e** Cross-feeding community (24/60 simulations)
                                                                      Community 8



**f**



**g**



**h**



**i** Switching between strategies (15/60 populations)
                                                                      Community 54



**j**



**k**



**l**



patterns in the spatial environment, and quickly mix when separated from each other (see Supplementary Fig. 6 and and Supplementary Movie 1). In some communities, cross-feeding or autonomous epochs that last tens of thousands of generations change by quickly switching strategy (Fig. 2i–l, from here on we refer to "autonomous", "cross-feeding" and "switching" community types). Switching occurs only occasionally, and typically once either strategy is established in a community it lasts until the end of the simulation. So, even though new mutations continue to fix in the population and metabolic networks remain in flux, the metabolic strategy of a community is very stable.

**Fig. 2 Emergence of diverse metabolic strategies. a–d** Example of population dominated by a single autonomous lineage. **a** Muller plot showing relative frequencies and phylogenetic relationships of different metabolic genotypes throughout the experiment. Clades of microbes with different metabolic genotypes (colours) continuously evolve, resulting in complex evolutionary dynamics of competition and leapfrogging. **b** Tracking ancestral relationships with renewing lineage markers shows a continued turnover in markers, indicating that at any point during the simulation all microbes have a recent common ancestor. **c** Snapshots of spatial environment. **d** Principal component analysis of single-cell proteomes shows that in these communities all microbes express similar proteins. **e–h** Example of a population that diversifies in two lineages that cross-feed on essential building blocks. **g** Lineages form an interleaved pattern in the spatial environment (see Supplementary Fig. 6 and Supplementary Movie 1). **h** Single-cell proteomes show that these lineages express different metabolic enzymes. **i–l** Example of a population that switches between autonomous and cross-feeding strategies. Lineage markers are redistributed when a single marker fixes in the whole population. PCAs coloured for lineage markers, and composed per simulation on relative single-cell protein expressions, see "Methods" for details.

To investigate the nature of the cross-feeding interaction, we examined whether lineages could survive in absence of the other. Specifically, at different time points after the two lineages emerged, we tested metabolic dependencies in the standing diversity of cross-feeding populations by removing all microbes from one lineage and preventing further mutations to occur in the remaining lineage (Fig. 3a–c). We find that cross-feeding communities generally consist of a major lineage that produces both building blocks and can survive by itself, and a minor lineage that is obligately dependent on the major lineage for one of the building blocks and goes extinct when the major lineage is removed, barring a few rare mutants (Fig. 3c). These dependencies are not constant during the simulation, but can increase, decrease, completely switch direction and change to fully co-dependent, as reflected by large changes in size of the population bottleneck following lineage removal (Fig. 3c) and changes in the ratio that cross-feeding lineages occur in a community during the main experiment (Fig. 2; Supplementary Fig. 1). However, both major and minor lineage nearly always grow faster in the presence of their partner (98,8% of cases that survive, see Fig. 3d, e; Supplementary Fig. 2), supplementing their own metabolism with building blocks produced by the other lineage.

At any given time during a simulation mutants with the opposing strategy can be found within a community, but even though simulations last tens of thousands of generations communities only switch strategy a couple of times and most communities (45/60) do not switch at all (Fig. 2; Supplementary Fig. 1). For example, in cross-feeding populations autonomous mutants can easily evolve via horizontal gene transfer between lineages with complementary metabolic networks. When a cross-feeding lineage is removed, such mutants in the remaining lineage can successfully take over and found a new, completely autonomous population, but their growth rates are higher in the context of the original cross-feeding population where they exploit the environment created by the whole ecosystem (Fig. 3d, e; Supplementary Fig. 3). This explains why they cannot replace the resident cross-feeding community in the main evolutionary experiment, where they sometimes reach substantial fractions but are typically only transiently present. Similarly, in autonomous communities gene loss can produce mutants that specialise on producing or importing only one building block, but these fail to invade in the resident population that imports and produces both building blocks. Apparently, both cross-feeding and metabolic autonomy are eco-evolutionary attractors that are stable against invading mutants of the opposed strategy, with only occasional occurrences of populations switching between them. Since all simulations started from the same ancestral community, this shows evolutionary contingency determines what kind of community evolves. Before we further consider the consequences for predicting evolution, we first need to understand exactly what determines which strategy evolves.

**The evolution of cross-feeding is not explained by protein cost.** The Black Queen Hypothesis explains the evolution of cross-feeding

through the adaptive loss of costly biosynthetic genes for metabolites that are produced by community members and publicly available[27–29]. In our simulations, the evolution of cross-feeding is characterised by loss of genes for building block synthesis and/or transporters, and results in smaller genomes for cross-feeding compared with autonomous strategies. As we assume an explicit cost for protein expression and essential building blocks are an "inescapable public good" because they are released into the environment when cells die, evolution of cross-feeding could thus be driven by Black Queen dynamics.

To test this we study the effect of varying the cost for protein expression on the evolution of metabolic strategies. Surprisingly, the emergence of autonomous, cross-feeding and occasionally switching communities is robust to increasing or decreasing the costs of proteins expression an order of magnitude (Supplementary Fig. 4). Although some of the dynamics change (for example, lower expression costs allow larger genomes to evolve and increased expression costs cause the evolutionary dynamics to slow down), both strategies evolve under all conditions and are stable eco-evolutionary attractors. Thus, even though in our model the production of building blocks acts as a public good and protein expression has an explicit cost that can be reduced by gene loss, the evolution of cross-feeding is not driven by gene loss to escape this cost.

**Trade-offs emerge during the evolution of metabolic networks.** To look for signatures for cross-feeding and autonomous strategies, we further investigated the diversity of metabolic networks that evolved. First, we clustered the final evolved populations at the end of the simulation based on metabolic gene frequencies in each population (see Fig. 4a). This shows that all populations share a core set of five genes for the uptake of the food resource and production and uptake of both building blocks. In cross-feeding communities, the genes for production and uptake of building blocks are only present in subsets of the population, reflecting how these communities have a distributed metabolic network. Note that populations strongly differ in which reaction is recruited to produce energy, and how byproducts from this reaction are further metabolised. Typically, a single dedicated energy reaction fixes in a population. Although clustering is dominated by individual energy generating reactions which clusters autonomous and cross-feeding populations with a few exceptions (Fig. 4a), no single gene acts as a signature for either community type.

Next, as metabolic strategies are stable during long-term evolution even though metabolic networks continuously evolve, we consider how these gene frequencies change over the complete duration of the simulation by PCA (Fig. 5; Supplementary Fig. 5). We find that the major component separates cross-feeding and autonomous communities. Moreover, based on the energy reactions recruited by each strategy, the metabolic networks can be classified in two different topologies that associate exclusively with either strategy: networks that degrade the food metabolite
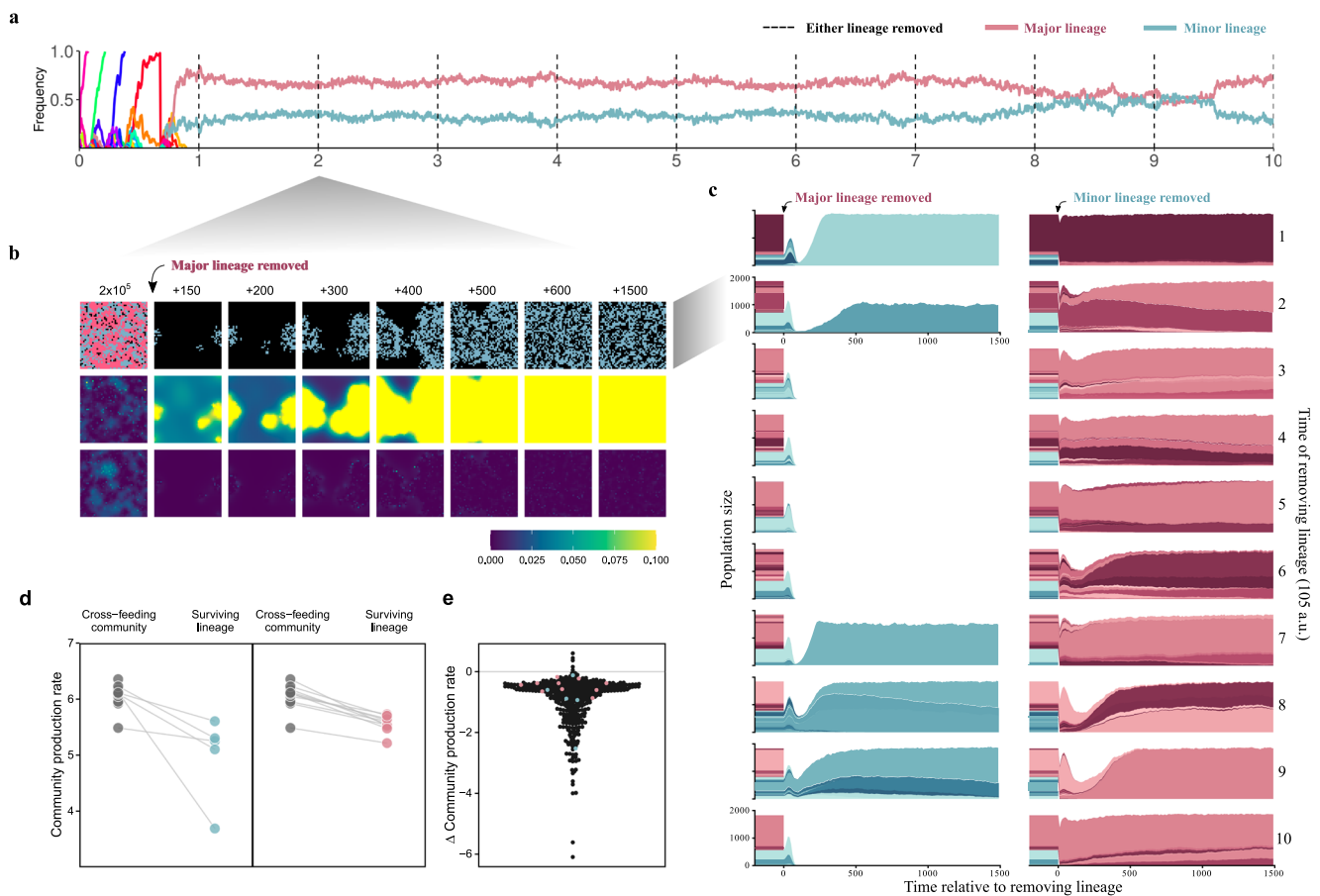
**Fig. 3 Metabolic dependencies in cross-feeding communities.** We tested metabolic dependencies in 29 selected populations by removing either lineage at different time points after cross-feeding evolved, and without allowing further mutations to occur. **a–d** Example of 2 × 10 tests of metabolic dependency in replicate population 23. Times indicated with dashed lines in **a**. **b**. When removing the major lineage (pink) at $t = 2 \times 10^5$, most microbes of the remaining lineage (blue) die out. However, a rare mutant is able to grow by itself, though it cannot import building block 1 and does not reach a high abundance. **c** Outcome of removing lineages for all time points in (**a**), with different metabolic genotypes within each lineage indicated with shades of the lineage colour. Typically, the minor lineage goes extinct or contains only few mutants that survive in isolation, reflecting obligate dependency on the major lineage. In contrast, microbes in the major lineage can mostly survive without the minor lineage. These dependencies are not constant over evolutionary time as metabolic genotypes that dominate within each lineage change. Note that directly following lineage removal, all remaining lineages can initially quickly grow on the limited store of building blocks that were produced by partner lineage and are still present in the environment. **d** Community production rates before and 1500 time steps after lineage removal. All surviving minor and major lineages have higher growth rates in the context of the original cross-feeding population. **e** Difference in community growth rate for surviving lineages in 484 tests of metabolic dependency in 29 populations. In total, 407 out of 412 (98.8%) tested cases that survive removal have reduced growth rates in isolation. Surviving lineages shown in **a–d** are highlighted.

for energy (i.e. R → energy + byproduct) are found in cross-feeding communities, and networks that degrade a building block for energy (i.e. B1 or B2 → energy + byproduct) in autonomous communities (Fig. 4b). Further support that links network topology to community strategy comes from communities that switched between strategies. Here, a switch from cross-feeding to autonomous (or vice versa) is accompanied by a simultaneous switch in network topology, and when a switch occurs communities move along the first principal component accordingly (Figs. 5d and 6a). Finally, communities can be composed of microbes with hybrid metabolic networks that degrade building blocks as well as food for energy (Figs. 4, 6). Interestingly, within a hybrid metabolic network one type of energy reaction appears dominant, as communities with such networks follow either a cross-feeding or autonomous strategy and do not mix different strategies within a community. Such communities can also switch strategy (and move accordingly in the PCA) without changing their network topology.

Why do these topologies determine metabolic strategies? The amount of energy available to a microbe is limited, and as a

consequence, importing more of one metabolite trades off with importing others, depending on what metabolite is used as an energy precursor. If microbes create energy by degrading the food resource, taking up other metabolites such as building blocks lowers the cell's energy budget. The amount of additional metabolites that can be imported is thus constrained for metabolic networks with this topology. As building blocks are produced and accumulate in the environment, this creates two niches (i.e. one for each building block) that can be exploited by different lineages. In contrast, when building blocks are degraded for energy, importing them increases the energy budget and does not trade-off with importing the food resource, allowing individual microbes to import both types of building blocks as well as retain their competitive ability for the food resource. Furthermore, as only one of the two building blocks is used for energy in autonomous communities, a cross-feeding scenario with this topology would be inherently asymmetrical and unstable, as one lineage would be dependent on the other for both energy and building blocks, while the other lineage only requires the complementary building block.
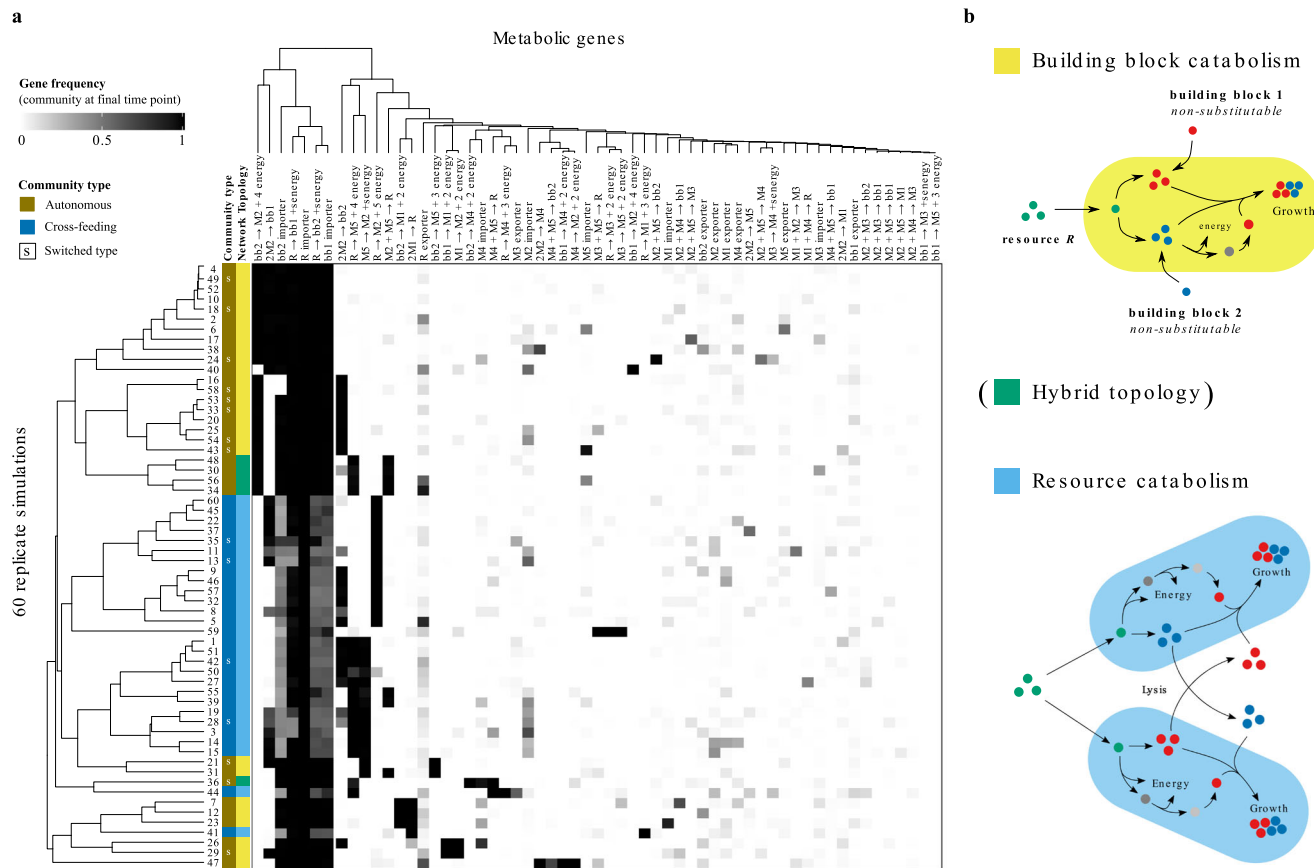
**Fig. 4 Emergent metabolic strategies differ in their energy metabolism. a** Heatmap showing the frequency of 59 metabolic genes (columns) in 60 evolved communities (rows) at the end of the simulation. Cross-feeding communities (dark blue label) and single-lineage autonomous communities (in mustard) cluster mostly together, but no single gene is associated with either metabolic strategy. Instead, the topology of the evolved metabolic network determines community strategy. **b** Examples of metabolic networks with different topologies. Topology is determined by the substrate of the energy reaction (resource or building block), and networks with the same topology may differ in the specific reaction used to produce energy and other reactions. All communities that degrade resource R for energy follow the cross-feeding strategy (light blue), while in contrast all autonomous communities degrade building block $B_1$ or $B_2$ for energy (yellow). 15 out of 60 communities switched strategy during the evolutionary simulation (marked with letter "S" in **a**) in most cases because a mutant with the opposing network topology invaded and replaced the resident population. Some communities are formed by microbes with hybrid metabolic networks that degrade both resource and building block for energy (indicated in green) and can switch strategy without changing network topology.

**The evolution of cross-feeding requires spatial structure.** Recent experimental and theoretical work[11,48–51] re-emphasised the importance of spatial structure and local interactions on eco-evolutionary dynamics, and metabolic division of labour in particular. In our evolutionary experiment microbes reshape the composition of the local environment through metabolic activity, and cross-feeding lineages self-organise into interleaved spatial patterns, locally enriching it for one of both building blocks (see Figs. 2, 3b; Supplementary Movie 1 and Supplementary Fig. 6). To test whether spatial structure was necessary for cross-feeding to evolve, we re-ran the experiment 18 times starting from the same ancestral population while simulating well-mixed but otherwise identical conditions. No cross-feeding lineages emerged, even though metabolic networks evolved that reliably associated with cross-feeding strategies in unmixed conditions. Moreover, when we stopped mixing, populations with the cross-feeding-associated topology quickly diversified in two cross-feeding lineages, while communities with an autonomous-associated topology remained metabolically autonomous, signifying that it is the interplay between environmental structure and evolved metabolic constraints that drives cross-feeding.

Finally, reasoning that long-term coexistence might result in increased robustness of the cross-feeding interaction, we tested

the ecological and evolutionary stability of cross-feeding communities from the original experiment by transfer to a well-mixed medium. Specifically, we subjected seven randomly chosen cross-feeding populations to well-mixed conditions at varying time steps after cross-feeding evolved, while either allowing or preventing further mutations to occur. In all "ecology-only" tests (i.e. without mutation), cross-feeding is stably maintained, and population size and community productivity increase. The increased productivity makes intuitive sense, as mutants that are less productive are outcompeted and cannot re-appear due to lack of mutations. Moreover, under unmixed conditions, local reproduction and metabolite diffusion limit the interface between both lineages and therefore reduce efficient exchange of building blocks. In contrast, when mutations are allowed under mixed conditions, all cross-feeding communities are quickly taken over by autonomous mutants. Strikingly, the resulting autonomous communities have smaller population sizes and productivity than their ancestral cross-feeding community. This shows that while spatial structure puts an upper limit to the efficiency of cross-feeding, it also protects against autonomous metabolic strategies. Consistent with previous results[32,51], we find that spatial structure is needed to evolve and maintain metabolic cross-feeding and also find that whether cross-feeding evolves or not depends on
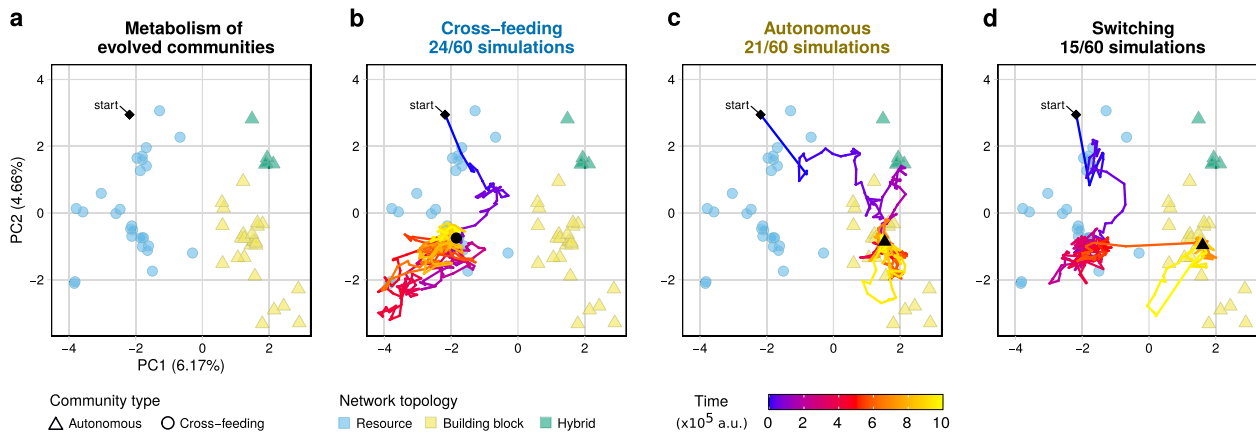
**Fig. 5 Evolutionary trajectories towards community attractors. a** PCA of gene frequencies of 59 metabolic genes in 58 communities over the whole duration of the experiment. One dot represents one community. For clarity, only the initial community and final time point of the simulations are shown. This separates communities by strategy along the first component, and reveals that topology of the evolved metabolic network determines metabolic strategy of the community: networks with reactions that degrade resource $R$ for energy cross-feed on building blocks, whereas networks with reactions that degrade building block $B_1$ or $B_2$ for energy remain metabolically autonomous and consume all building blocks from the environment. **b–d** Evolutionary trajectory showing all time points in the PCA for a community that **b** evolves cross-feeding (community 9), **c** metabolic autonomy (community 4) and **d** switches between strategies (community 18). For visualisation purposes outlier populations 59 and 47 were omitted from this analysis (see Supplementary Fig. 5 for analysis including these outliers).

constraints of previous metabolic adaptations. As microbes evolve to produce more energy from either the resource or one of the building blocks, importing one metabolite trades off with importing others. We find that the shape of this trade-off is an evolved property of the metabolic networks and the local environmental niches they construct.

**Metabolic strategies are an evolutionary contingency**. Our results demonstrate that the topology of the evolved metabolic network, combined with spatial structure, determines whether cross-feeding evolves or not. Which topology evolves in a population is arbitrary and often establishes early on. For simulations where the cost of protein expression is increased, this topology often fixes up to tens of thousands of generations before metabolic networks "mature" by making use of all building blocks that accumulate in the environment. What eco-evolutionary strategy will evolve when microbes finally evolve to tap into that source can be predicted from the evolved topology (see Fig. 6b), realising a fate already cemented earlier in its evolutionary history. However, it is interesting to note that exact prediction is limited by several factors. Firstly, evolution of the topology of the metabolic network is typically "founder controlled", where the energy reaction that establishes itself first in the community quickly accumulates more beneficial mutations and is never outcompeted by other energy reaction genes that are discovered later on. However, mutants with a different energy type occasionally do invade and replace the original population, changing community fate (Fig. 6a). Secondly, microbes that have a hybrid metabolism can switch between strategies as they evolve and different energy reactions dominate the metabolic network or are lost (Fig. 6a).

Concluding, a combination of contingency and predictability is manifest in our eco-evolutionary modelling experiment. Given the topology of the metabolic network that evolves, in the long run, and with intermittent metastable states, the type of community which evolves is predictable.

## Discussion
Previous experimental studies have shown repeated evolution of cross-feeding[8,13,14,43,52], and the emergence of cross-feeding is typically interpreted as driven by physiological constraints. By explicitly incorporating constraints such as trade-offs in resource uptake or production[25], or a limit in the maximum number of reactions within a single organism[26], theoretical approaches demonstrate how these constraints can force metabolic specialisation and drive the evolution of cross-feeding. Here, rather than pre-defining such physiological constraints in our model, we demonstrate how metabolic constraints can themselves evolve, and whether or not they are driving cross-feeding is a historic contingency.

Next to metabolic constraints, other factors important for the evolution of metabolic exchange have been previously identified. For example, in batch culture repeated cycles of feast and starvation generate temporal niches that explain stable coexistence of cross-feeding ecotypes[7,8,25,37,38]. In these studies, cross-feeding is unidirectional, i.e. one lineage feeds on metabolic byproducts of the other. In contrast, we here show bidirectional cross-feeding of two non-substitutable building blocks, and in a constant environment. This matches recent experimental observations of the evolution of cross-feeding on essential nutrients[43], and stable coexistence mediated by uptake of debris from lysed cells of a coexisting lineage[8].

Another suggested driver for the evolution of cross-feeding is that the collective exploitation of resources can be more efficient or productive than individual-based solutions[19,26,53,54]. In our simulations, autonomous mutants in cross-feeding communities are indeed less productive than the resident cross-feeders (Fig. 3; Supplementary Figs. 2, 3). However, communities that evolve a completely autonomous strategy are equally or more productive than cross-feeding communities, showing that productivity alone may not be sufficient to explain widespread metabolic-dependencies observed in nature. Thus, we should realise that community productivity itself is an evolved property of the evolved metabolism, and that division of labour is not necessarily more productive, as productivity depends on evolved metabolic constraints.

Metabolic dependencies can also arise due to non-adaptive gene loss caused by genetic drift in small populations[55]. In our model this seems not to be the case as cross-feeding has a clear adaptive benefit (Supplementary Fig. 2), and lost functions are easily regained by HGT.
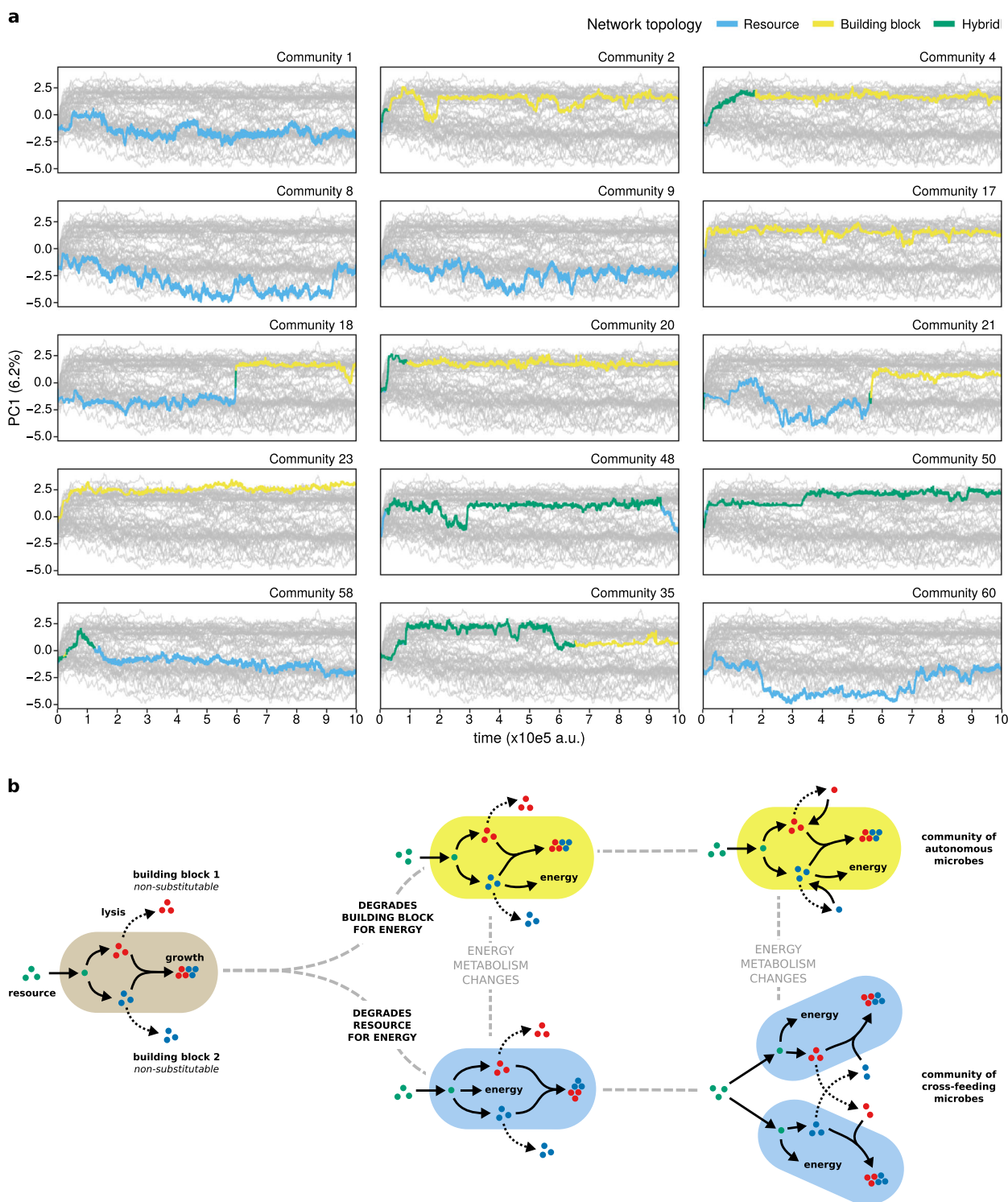
**Fig. 6 Prior metabolic adaptations constrain future ecological roles. a** Evolutionary trajectories of example communities (first component from PCA in Fig. 5 v.s. time) towards cross-feeding (negative *y*-value) or autonomous (positive *y*-value) strategy, coloured for topology of the metabolic network. Grey lines indicate trajectories of all other communities. Changes in dominant network topology cause a switch in community strategy. **b** Cartoon of evolutionary trajectories. Earlier metabolic adaptations that fix in the initial population dictate final eco-evolutionary attractor, but are an evolutionary contingency. However, prediction is limited because the duration of each depicted stage is unpredictable, and cases where mutants with an alternate network topology invade and replace the population (dashed arrow in **b**, population 18 and 21 in **a**) are possible.

Recent experimental work showed that initially clonal microbial communities quickly become genetically heterogeneous[9,44], and that in the resulting complex evolutionary dynamics the fate of individual mutations does not only depend on that specific mutation's fitness, but mainly on what new mutations accumulate in that background[44]. Likewise, we find that when no further mutations are allowed to occur, autonomous mutants already present in cross-feeding population never fixate in the well-mixed selection regime, but when they can acquire more mutations always take over and dominate evolutionary outcome. In other

words, the outcome of evolution (i.e. type of community that evolves) depends critically on interlocking evolutionary and ecological timescales, recapitulating earlier theoretical work[46,47].

It is interesting to note that in each cross-feeding lineage more mutants gain partial dominance than is the case in autonomous lineages (Fig. 2). The amount of mutations in synthetic cross-feeding versus autonomous strains was recently studied in[56]. Contrasting with our results, they found that although, under strong selection, the sum of the mutations in the two cross-feeding strains was larger and more varied than in the autonomous strain, the separate cross-feeding strains acquired less mutations than the autonomous strain. We conclude that the speed and trajectory of evolutionary dynamics in relation to co-evolving, interdependent lineages under various selection regimes is a promising avenue of future computational, theoretical and experimental research.

For computational feasibility and simplicity, we here considered a simple metabolism with only two essential building blocks, while in reality many metabolites such as carbon and nitrogen sources, amino acids, nucleotides and vitamins are required for growth, and are also exchanged in microbial communities[23]. It would be interesting to extend the model metabolism with more building blocks and investigate whether the here reported distinct community types (metabolic autonomy and cross-feeding) are maintained, or whether communities with intermediate levels of metabolic division of labour evolve.

Interestingly, the dichotomy in community types we report here matches a recent paper that investigated the metabolism of natural communities composed of up to 40 species. The authors reported that natural communities in a wide range of environments are either highly cooperative, with smaller genomes and diverse metabolic dependencies, or highly competitive, and dominated by species with larger genomes and overlapping nutritional needs[57]. This suggests the result of our computational experiment is a general property of microbial communities. Furthermore, it would be interesting to investigate whether the rapid switching between cross-feeding and autonomous community types we find also occurs in natural communities. It is exciting to see that patterns emerging in the eco-evolutionary dynamics of our multi-level virtual microbes model are now detectable through metagenomic analysis of natural communities. We look forward to future cross-feeding of the two approaches to enhance our understanding of the microbial world.

## Methods

**Metabolic genotypes and the construction of Muller plots**. We found it useful for interpretation and visualisation purposes to classify microbes based on their "metabolic genotype": a binary representation that indicates presence or absence of each of the 59 metabolic genes and transporters in the genome, ignoring copy number variations, promoter strength and kinetic parameters. During simulations, we created a phylogeny with clades consisting of related microbes with the same metabolic genotype by tracking gain and loss mutations that changed this metabolic genotype and collecting abundances and ancestral relationships. For creating Muller plots, we subsampled the data every 500 time steps, and used a cut-off of 5% of maximum population size to remove clades that never became abundant. Plots were created in R with the GGMuller package[58], which follows the convention to stack every new clade emerging from the centre of its parent clade in chronological order of appearance.

**Lineage tracking with renewing markers**. To allow tracking of ecological and evolutionary dynamics, we used a technique similar to barcoding methods used in experimental evolution[59]. We label each microbe in the initial population with a unique neutral marker that it passes on to their offspring. If at any point during the simulation all living microbes have the same marker—indicating that they share a common ancestor that was alive when markers were distributed—markers are renewed by giving each microbe a new unique marker. Note that these markers only track lineage dynamics and not mutations, and lineages do not correspond directly to the clades of metabolic genotypes shown in the Muller plots. Markers do not reflect genetic or phenotypic identity nor provide high-resolution insight in the

thousands of independent adaptive and non-adaptive mutations competing within a population.

**Diversity within replicate populations**. To investigate different metabolic strategies of individual microbes within a community, we measured differential investments across all metabolic reactions by normalising protein concentrations of each metabolic protein to the total protein concentration in the cell. We collected protein data from all cells for time points (1, 2, ..., 10) $\times 10^5$ and then performed a separate PCA for each simulation.

**Testing metabolic dependency**. To asses metabolic dependency of microbes in cross-feeding populations, we removed all microbes in one lineage (including internal metabolites) and set mutation rates to zero. Following removal, we continued the simulation for a maximum of 2000 time steps or population extinction. For 29 populations that evolved cross-feeding (all 21 cross-feeding populations and 8 randomly selected switching population), we tested both lineages at every multiple of $10^5$ time steps after cross-feeding evolved. In total, we performed 484 tests. We measured production rates of microbes at the time step just preceding lineage removal and, if populations survived, after 2000 time steps following removal.

**Diversity between replicate populations**. We performed a single principal component analysis on the frequencies of all 59 metabolic genes in all 60 replicate populations over all time points in the main experiment. For clarity of visualisation, we omitted two outliers (population 47 and 59) from this analysis, which did not impact the result. See Supplementary Fig. 5 for an analysis including these outliers.

**Model description**. VirtualMicrobes extends an earlier explicit model of genome evolution, metabolism and homoeostasis[60–62] by adding a structured environment and emergent microbe-microbe interactions. It is a highly configurable framework designed to study emergent, eco-evolutionary microbial dynamics under a wide range of possible conditions. For clarity, here we only describe the configuration and features that are relevant to this study. A full manual is available at https://bitbucket.org/thocu/virtualmicrobes and https://virtualmicrobes.readthedocs.io, and parts have previously been described in an earlier study[38].

**Overview**. VirtualMicrobes is an agent-based model, where each individual microbe occupies a grid point on a 2D plane. Population dynamics (i.e. competition, reproduction and death) play out every time step. Within a time step, cellular processes (cellular growth, gene expression, metabolic reactions including uptake, excretion and diffusion across the cell membrane, metabolite and protein decay) and environmental metabolite dynamics (food influx, degradation and diffusion between grid points) are updated using ordinary differential equations (ODEs).

**Metabolic universe**. The metabolic universe is an a priori defined set of all metabolites and possible reactions between them. We procedurally generated a universe consisting of nine metabolites (a single resource $R$, two designated building block molecules $B_1$, $B_2$, energy carrier $E$, and four intermediate metabolites $M_{1-5}$) with their own toxicity, degradation and diffusion parameters, and 43 metabolic reactions (see Table 1). This provides evolution considerable degrees of freedom for forming metabolic networks to produce the required building blocks and energy from the available resource.

**Transport**. For all metabolites, transporters exist that import or export the metabolite across the cell membrane. Transporters $\mathcal{T}$ catalyse the transport of substrate $S$ over the cell membrane by consuming energy metabolite $E$. Transport rate $v$ is then given by Michaelis–Menten kinetics:

$$v = v_{\max_{\mathcal{T}}} \cdot [\mathcal{T}] \cdot \frac{[S] \cdot [E]}{([S] + K_S) \cdot ([E] + K_E)}, \tag{1}$$

where $K_S$, and $K_E$ are the Michaelis–Menten constants for the substrate and energy molecule. Depending on the direction of transport (importing or exporting) $[S]$ is either the external or internal concentration of the substrate. $K_S$, $K_E$ and $v_{\max \mathcal{T}}$ are all freely evolvable parameters.

**Metabolism**. Metabolic enzymes catalyse reactions of the general form:

$$R_0 + R_1 + \dots \longrightarrow P_0 + \dots , \tag{2}$$

converting reactant metabolites $\{R_i, \dots\}$ to products $\{P_j, \dots\}$. The rate of catalysis $v$ is calculated with Michaelis–Menten kinetics as follows:

$$v = v_{\max_{\mathcal{E}}} \cdot [\mathcal{E}] \cdot \frac{\prod_{R \in \mathcal{R}} [R]}{\prod_{R \in \mathcal{R}} ([R] + K_R)}, \tag{3}$$

where $[\mathcal{E}]$ is the concentration of the enzyme catalysing the reaction, $\mathcal{R}$ the set of all reactant metabolites and $K_R$ and $v_{\max_{\mathcal{E}}}$ are evolvable kinetic parameters of enzyme $\mathcal{E}$.

**Table 1 Metabolic universe: Using VirtualMicrobes, we generated an artificial biochemistry comprised of 9 metabolites, 8 importers, 8 exporters and 43 conversion reactions, with two non-substitutable building blocks and a single energy molecule.**

| Metabolite | Class | Diffusion rate | Degradation rate | Toxicity | Mass | Influx rate |
|---|---|---|---|---|---|---|
| R | Resource | 0.011 | 0.0003 | 0.077 | 9 | 0.002 |
| B1 | Building block | 0.011 | 0.0100 | 0.058 | 8 | 0 |
| B2 | Building block | 0.012 | 0.0100 | 0.103 | 8 | 0 |
| E | Energy carrier | 0.050 | 0.1000 | 0.065 | 1 | 0 |
| M1 | – | 0.014 | 0.0006 | 0.105 | 6 | 0 |
| M2 | – | 0.019 | 0.0003 | 0.080 | 4 | 0 |
| M3 | – | 0.013 | 0.0006 | 0.079 | 7 | 0 |
| M4 | – | 0.014 | 0.0014 | 0.158 | 6 | 0 |
| M5 | – | 0.016 | 0.0008 | 0.047 | 5 | 0 |

**Reactions**

| | | | | |
|---|---|---|---|---|
| R $\rightarrow$ B2 + E | B1 $\rightarrow$ M4 + 2 E | M1 $\rightarrow$ M2 + 2 E | M2 + M5 $\rightarrow$ B1 | M3 + M5 $\rightarrow$ R |
| R $\rightarrow$ B1 + E | B1 $\rightarrow$ M5 + 3 E | M2 + M3 $\rightarrow$ B1 | M2 + M5 $\rightarrow$ M1 | M3 $\rightarrow$ M5 + 2 E |
| R $\rightarrow$ M1 + 3 E | B2 $\rightarrow$ M1 + 2 E | M2 + M3 $\rightarrow$ B2 | 2 M1 $\rightarrow$ R | M3 $\rightarrow$ M4 + E |
| R $\rightarrow$ M2 + 5 E | B2 $\rightarrow$ M2 + 4 E | M2 + M5 $\rightarrow$ M4 | 2 M2 $\rightarrow$ M1 | M4 $\rightarrow$ M2 + 2 E |
| R $\rightarrow$ M4 + 3 E | B2 $\rightarrow$ M3 + E | M2 + M4 $\rightarrow$ M3 | 2 M2 $\rightarrow$ B1 | M4 + M5 $\rightarrow$ B2 |
| R $\rightarrow$ M5 + 4 E | B2 $\rightarrow$ M4 + 2 E | M2 + M5 $\rightarrow$ M3 | 2 M2 $\rightarrow$ B2 | M4 + M5 $\rightarrow$ R |
| R $\rightarrow$ M3 + 2 E | B2 $\rightarrow$ M5 + 3 E | M2 + M4 $\rightarrow$ B2 | 2 M2 $\rightarrow$ M4 | M4 + M5 $\rightarrow$ B1 |
| B1 $\rightarrow$ M1 + 2 E | M1 + M4 $\rightarrow$ R | M2 + M5 $\rightarrow$ B2 | 2 M2 $\rightarrow$ M5 | M5 $\rightarrow$ M2 + E |
| B1 $\rightarrow$ M2 + 4 E | M1 + M2 $\rightarrow$ M3 | M2 + M5 $\rightarrow$ R | | |

Importers: one for each non-energy metabolite (8 total)
Exporters: one for each non-energy metabolite (8 total)

**Degradation and dilution of molecules**. Concentrations of all molecules (i.e. proteins, product and metabolites) are adjusted according to the change in cell volume d$Vol$/d$t$. In addition, all molecules degrade with a molecule-specific rate degradation.

**Biomass production budget**. Microbes convert building blocks $B_1$ and $B_2$ to a production budget $\mathcal{B}$ with rate *production*, favouring homeostasis of building block concentrations with:

$$production = \frac{[B_1] \cdot [B_2]}{1 + \left( |[B_1] - \frac{[B_1]+[B_2]}{2}| + |[B_2] - \frac{[B_1]+[B_2]}{2}| \right)}. \quad (4)$$

This budget is spent for increasing cell size $cost_{\text{growth}}$ and producing proteins by gene expression $cost_{\text{expr}}$. Note that to maintain sufficient selection pressure to select for adaptive mutations as building block synthesis rates of populations increase during evolutionary simulations, we scale the amount of budget spent with $\mathcal{B}_{\text{scaling}}$:

$$\mathcal{B}_{\text{scaling}} = \frac{\mathcal{B}}{\mathcal{B} + \mathcal{B}_{\text{pop}}}, \quad (5)$$

where $\mathcal{B}_{\text{pop}}$ is the time averaged population production value. The concentration of $\mathcal{B}$ then changes with a rate:

$$\frac{d\mathcal{B}}{dt} = production - cost_{\text{growth}} - cost_{\text{expr}} - degradation - dilution. \quad (6)$$

When cells grow, $\mathcal{B}$ is consumed proportional to the rate of growth:

$$cost_{\text{growth}} = Vol_{\text{growth}} \cdot \mathcal{B}_{\text{scaling}} \cdot \mathcal{B}. \quad (7)$$

Total consumption of product for gene expression is summed over all genes:

$$cost_{\text{expr}} = \sum_{i=1}^{N_{\text{genes}}} Pr_i \cdot \mathcal{B}_{\text{scaling}} \cdot \mathcal{B}. \quad (8)$$

**Cell volume growth**. We assume that if microbes grow without being able to divide due to lack of space, they approach a maximum cell size $Vol_{\text{max}}$ with rate $g$. We implement a cost for maintaining cell volume with a small shrinking rate $s$. Volume changes with a rate:

$$\frac{dVol}{dt} = Vol_{\text{growth}} - s \cdot Vol, \quad (9)$$

where $Vol_{\text{growth}}$ is given by:

$$Vol_{\text{growth}} = g \cdot Vol \cdot \frac{1 - Vol}{Vol_{\text{max}}} \cdot \mathcal{B}_{\text{scaling}}. \quad (10)$$

**Protein expression**. Protein concentrations $\mathcal{P}$ for any given gene are governed by the function:

$$\frac{d\mathcal{P}}{dt} = Pr \cdot \mathcal{B}_{\text{scaling}} - degradation - dilution, \quad (11)$$

where $Pr$ is the per gene evolvable parameter promoter strength. VirtualMicrobes features explicit transcription factor proteins (TF) that can modify protein transcription. TFs were not relevant for this study and not recruited during evolutionary simulations and are here omitted for clarity. See the full documentation for a description of TFs.

**Toxicity and death**. We model death as a stochastic process depending on an intrinsic death rate $r = 0.03$, which is potentially increased when internal metabolite concentrations reach a toxic threshold. This cumulative toxic effect $e_{\text{tox}}$ is computed over all internal metabolites $M$ and the current lifetime $\tau$ of a microbe as:

$$e_{\text{tox}} = \sum_{m \in M} \int_{t=0}^{\tau} f(m, t) dt, \quad (12)$$

with the toxic effect function $f(m, t)$ for the concentration of metabolite $m$ at time $t$ with metabolite-specific toxicity threshold $tox_m$:

$$f(m, t) = \max(0, \frac{[m_t] - tox_m}{tox_m}). \quad (13)$$

This toxic effect increases the death rate $d$ of microbes starting at the intrinsic death rate $r$:

$$d = \frac{e_{\text{tox}}}{s + e_{\text{tox}}} \cdot (1 - r) + r, \quad (14)$$

where $s$ scales the toxic effect. Microbes that survive after an update cycle retain the toxic level they accumulated so far. Cells can also die from starvation. When they are unable to produce enough biomass product to maintain the slowly decaying volume of the cell and cell size drops below a predefined threshold size, cells are automatically marked for death.

**Reproduction**. Microbes grow and reproduce on a $45 \times 45$ grid, competing for metabolites and available space. A microbe can reproduce if it meets the minimal division size and there is an empty grid site in any of its eight neighbouring sites. When multiple microbes are eligible for reproduction in the same empty site, reproduction chance depends on the microbe's production value $P$, with a dynamically scaled chance for a "no reproduction" event occurring. Cell volume is divided equally between parent and offspring (molecule and protein concentrations remain constant), and the genome is copied with possible mutations. Toxic effects built up during the parent's lifetime do not carry over to offspring.

**Table 2 Types of mutations and their probabilities.**

| Mutations | Description | Probability |
|---|---|---|
| Duplication | A stretch of 1 or more genes is duplicated. Stretch length is geometrically distributed with $P = 0.3$ | 0.001 |
| Deletion | A stretch of 1 or more genes is deleted | 0.001 |
| Inversion | A stretch of 1 or more genes is inverted in order | 0.001 |
| Translocation | A stretch of 1 or more genes is moved to a random location on the genome. | 0.001 |
| Gene discovery | Per time-step probability of discovering a new gene (randomly parameterised) | 0.0001 |
| Horizontal Gene Transfer | Per time-step probability of copying a gene from a cell in a neighbouring site | 0.0005 |
| Point mutation | Per gene per generation probability of modifying a single parameter of a gene | 0.02 |

| Evolvable gene parameters | Gene types | Value range |
|---|---|---|
| Promoter strength | Enzyme, transporter | [0.01, 8] |
| $K_S$ | Enzyme, transporter | [0.01, 8] |
| $K_E$ | Transporter | [0.01, 8] |
| $V_{\max}$ | Enzyme, transporter | [0.01, 8] |
| Exporting | Transporter | [True, False] |

**Genome and mutations**. The genome is a linear organisation of genes coding for enzymes and transporters. When a microbe successfully reproduces it divides and the offspring inherits a copy of the parent's genome, which can then be subject to various types of mutations. Stretches of genes can be duplicated, deleted, inverted or translocated to another position on the genome (Table 2). At the single gene level, all evolvable parameters can mutate individually (Table 2). Horizontal gene transfer can occur on every time step. Innovations are modelled as HGT from an external (off-grid) source by drawing a random reaction from the metabolic universe with random parameters, and can occur every time step.

**Initialisation and experimental setup**. We constructed an initial population of 2025 "minimally viable" microbes by generating genomes containing one importer gene for the food resource and 5 random enzymes, with the constraint that the resulting metabolic network was at least capable of producing the necessary building blocks and energy from the provided resource. Genes were randomly parameterised and ordered on the genome. In total, 60 identical copies of this population were propagated in parallel for $1 \times 10^6$ time steps (corresponding to roughly $4 \times 10^5$ generations) under the exact same conditions, differing only in the mutation seed.

**Statistics and reproducibility**. The experiments and analysis conducted in this paper can be reproduced with the publicly available code.

**Reporting summary**. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability
The data shown supporting this study is available for download from https://doi.org/10.5281/zenodo.3840463[63], with exception of the data used for creating the Muller diagrams in Fig. 2, Supplementary Fig. 6 and Supplementary Movie 1, which are available from the corresponding author upon request.

## Code availability
Full documentation, source code, and installation instructions for the VirtualMicrobes Python module are publicly available at https://bitbucket.org/thocu/virtual-microbes. This study used VirtualMicrobes version 0.2.5, which is archived at https://doi.org/10.5281/zenodo.3843609[64]. Custom scripts used for the analysis and creation of figures in this paper are available at https://github.com/meijer-jeroen/contingent-evolution-2020 and archived at https://doi.org/10.5281/zenodo.3841930[65].

## References
1. Herron, M. D. & Doebeli, M. Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli*. *PLoS Biol.* **11**, e1001490 (2013).
2. Rainey, P. B. & Travisano, M. Adaptive radiation in a heterogeneous environment. *Nature* **394**, 69–72 (1998).
3. Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am. Naturalist* **138**, 1315–1341 (1991).
4. Kinnersley, M. et al. Ex uno plures: clonal reinforcement drives evolution of a simple microbial community. *PLoS Genet.* **10**, e1004430 (2014).
5. Le Gac, M. et al. Metabolic changes associated with adaptive diversification in *Escherichia coli*. *Genetics* **178**, 1049–1060 (2008).
6. Rozen, D. E. & Lenski, R. E. Long-term experimental evolution in *Escherichia coli*. VIII. Dynamics of a balanced polymorphism. *Am. Naturalist* **155**, 24–35 (2000).
7. Gac, M. L., Plucain, J., Hindré, T., Lenski, R. E. & Schneider, D. Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **109**, 9487–9492 (2012).
8. Rozen, D. E., Philippe, N., Arjan de Visser, J., Lenski, R. E. & Schneider, D. Death and cannibalism in a seasonal environment facilitate bacterial coexistence. *Ecol. Lett.* **12**, 34–44 (2009).
9. Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E. & Desai, M. M. The dynamics of molecular evolution over 60,000 generations. *Nature* **551**, 45–50 (2017).
10. Behringer, M. G. et al. *Escherichia coli* cultures maintain stable subpopulation structure during long-term evolution. *Proc. Natl Acad. Sci. USA* **115**, E4642–E4650 (2018).
11. Dal Co, A., van Vliet, S. & Ackermann, M. Emergent microscale gradients give rise to metabolic cross-feeding and antibiotic tolerance in clonal bacterial populations. *Philos. Trans. R. Soc. B: Biol. Sci.* **374**, 20190080 (2019).
12. Helling, R. B., Vargas, C. N. & Adams, J. Evolution of *Escherichia coli* during growth in a constant environment. *Genetics* **116**, 349–358 (1987).
13. Rosenzweig, R. F., Sharp, R. R., Treves, D. S. & Adams, J. Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* **137**, 903–917 (1994).
14. Treves, D. S., Manning, S. & Adams, J. Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Mol. Biol. Evolution* **15**, 789–797 (1998).
15. Kinnersley, M. A., Holben, W. E. & Rosenzweig, F. E. Unibus plurum: genomic analysis of an experimentally evolved polymorphism in *Escherichia coli*. *PLoS Genet.* **5**, e1000713 (2009).
16. Gause, G. F. *The Struggle for Existence* (The Williams & Wilkins company, Baltimore, 1934).
17. Kwong, W. K., Engel, P., Koch, H. & Moran, N. A. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl Acad. Sci. USA* **111**, 11509–11514 (2014).
18. Pande, S. et al. Fitness and stability of obligate cross-feeding interactions that emerge upon gene loss in bacteria. *ISME J.* **8**, 953–962 (2014).
19. Rosenthal, A. Z., Matson, E. G., Eldar, A. & Leadbetter, J. R. RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. *ISME J.* **5**, 1133–1142 (2011).
20. Blasche, S. et al. Emergence of stable coexistence in a complex microbial community through metabolic cooperation and spatio-temporal niche partitioning. Preprint at https://www.biorxiv.org/content/10.1101/541870v1 (2019).
21. Zelezniak, A. et al. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc. Natl Acad. Sci. USA* **112**, 6449–6454 (2015).
22. Garcia, S. L. et al. Auxotrophy and intrapopulation complementary in the 'interactome' of a cultivated freshwater model community. *Mol. Ecol.* **24**, 4449–4459 (2015).
23. D'Souza, G. et al. Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Nat. Prod. Rep.* **35**, 455–488 (2018).
24. Johnson, D. R., Goldschmidt, F., Lilja, E. E. & Ackermann, M. Metabolic specialization and the assembly of microbial communities. *ISME J.* **6**, 1985–1991 (2012).

25. Großkopf, T. et al. Metabolic modelling in a dynamic evolutionary framework predicts adaptive diversification of bacteria in a long-term evolution experiment. *BMC Evolut. Biol.* **16**, 163 (2016).

26. Thommes, M., Wang, T., Zhao, Q., Paschalidis, I. C. & Segrè, D. Designing metabolic division of labor in microbial communities. *mSystems* **4**, e00263–18 (2019).

27. Morris, J. J., Lenski, R. E. & Zinser, E. R. The black queen hypothesis: evolution of dependencies through adaptive gene loss. *mBio* **3**, e00036–12 (2012).

28. D'Souza, G. et al. Less is more: selective advantages can explain the prevalent loss of biosynthetic genes in bacteria. *Evolution* **68**, 2559-2570 (2014).

29. Pande, S. & Kost, C. Bacterial unculturability and the formation of intercellular metabolic networks. *Trends Microbiol.* **25**, 349–361 (2017).

30. Dekel, E. & Alon, U. Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**, 588 (2005).

31. Tsoi, R. et al. Metabolic division of labor in microbial systems. *Proc. Natl Acad. Sc. USA* **115**, 2526–2531 (2018).

32. Stump, S. M., Johnson, E. C., Sun, Z. & Klausmeier, C. A. How spatial structure and neighbor uncertainty promote mutualists and weaken black queen effects. *J. Theor. Biol.* **446**, 33–60 (2018).

33. Josephides, C. & Swain, P. S. Predicting metabolic adaptation from networks of mutational paths. *Nat. Commun.* **8**, 685 (2017).

34. Gudelj, I., Beardmore, R. E., Arkin, S. S. & Maclean, R. C. Constraints on microbial metabolism drive evolutionary diversification in homogeneous environments. *J. Evolut. Biol.* **20**, 1882–1889 (2007).

35. Pfeiffer, T., Schuster, S. & Bonhoeffer, S. Cooperation and competition in the evolution of ATP-producing pathways. *Science* **292**, 504–507 (2001).

36. Roman, M. S. & Wagner, A. An enormous potential for niche construction through bacterial cross-feeding in a homogeneous environment. *PLoS Comput. Biol.* **14**, e1006340 (2018).

37. Rocabert, C., Knibbe, C., Consuegra, J., Schneider, D. & Beslon, G. Beware batch culture: seasonality and niche construction predicted to favor bacterial adaptive diversification. *PLoS Computational Biol.* **13**, e1005459 (2017).

38. van Dijk, B., Meijer, J., Cuypers, T. D. & Hogeweg, P. Trusting the hand that feeds: microbes evolve to anticipate a serial transfer protocol as individuals or collectives. *BMC Evolut. Biol.* **19**, 201 (2019).

39. Stump, S. M., Johnson, E. C. & Klausmeier, C. A. Local interactions and self-organized spatial patterns stabilize microbial cross-feeding against cheaters. *J. R. Soc. Interface* **15**, 20170822 (2018).

40. Roller, B. R. K., Stoddard, S. F. & Schmidt, T. M. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nat. Microbiol.* **1**, 1–7 (2016).

41. Kinnersley, M. et al. Evolutionary dynamics of de novo mutations and mutant lineages arising in a simple, constant environment. Preprint at https://doi.org/10.1101/540625 (2019).

42. Levy, S. F. et al. Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature* **519**, 181–186 (2015).

43. D'Souza, G. & Kost, C. Experimental evolution of metabolic dependency in bacteria. *PLoS Genet.* **12**, e1006364 (2016).

44. Ba, A. N. N. et al. High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. *Nature* **575**, 494–499 (2019).

45. Cvijović, I., Nguyen Ba, A. N. & Desai, M. M. Experimental studies of evolutionary dynamics in microbes. *Trends Genet.* **34**, 693–703 (2018).

46. Van Der Laan, J. & Hogeweg, P. Predator-prey coevolution: interactions across different timescales. *Proc. Royal Soc. London B: Biol. Sci.* **259**, 35–42 (1995).

47. Kotil, S. E. & Vetsigian, K. Emergence of evolutionarily stable communities through eco-evolutionary tunnelling. *Nat. Ecol. Evolution* **2**, 1644 (2018).

48. Co, A. D., van Vliet, S., Kiviet, D. J., Schlegel, S. & Ackermann, M. Short-range interactions govern the dynamics and functions of microbial communities. *Nat. Ecol. Evol.* **4**, 366–375 (2020).

49. Germerodt, S. et al. Pervasive selection for cooperative cross-feeding in bacterial communities. *PLoS Comput. Biol.* **12**, e1004986 (2016).

50. Marchal, M. et al. A passive mutualistic interaction promotes the evolution of spatial structure within microbial populations. *BMC Evolut. Biol.* **17**, 106 (2017).

51. Pande, S. et al. Privatization of cooperative benefits stabilizes mutualistic cross-feeding interactions in spatially structured environments. *ISME J.* **10**, 1413–1423 (2016).

52. Pfeiffer, T. & Bonhoeffer, S. Evolution of cross-feeding in microbial populations. *Am. Naturalist* **163**, E126–135 (2004).

53. Alessi, A. M. et al. Defining functional diversity for lignocellulose degradation in a microbial community using multi-omics studies. *Biotechnol. Biofuels* **11**, 166 (2018).

54. Yang, D.-D. et al. Fitness and productivity increase with ecotypic diversity among escherichia coli strains that coevolved in a simple, constant environment. *Appl. Environ. Microbiol.* **86**, e00051–20 (2020).

55. Kuo, C.-H., Moran, N. A. & Ochman, H. The consequences of genetic drift for bacterial genome complexity. *Genome Res.* **19**, 1450–1454 (2009).

56. Lilja, E. E. & Johnson, D. R. Substrate cross-feeding affects the speed and trajectory of molecular evolution within a synthetic microbial assemblage. *BMC Evolut. Biol.* **19**, 129 (2019).

57. Machado, D. et al. Polarization of microbial communities between competitive and cooperative metabolism. Preprint at https://www.biorxiv.org/content/10.1101/2020.01.28.922583v1 (2020).

58. Noble, R. ggmuller: create muller plots of evolutionary dynamics. https://CRAN.R-project.org/package=ggmuller (2019).

59. Blundell, J. R. & Levy, S. F. Beyond genome sequencing: Lineage tracking with barcodes to study the dynamics of evolution, infection, and cancer. *Genomics* **104**, 417–430 (2014).

60. Cuypers, T. D. & Hogeweg, P. Virtual genomes in flux: an interplay of neutrality and adaptability explains genome expansion and streamlining. *Genome Biol. Evolut.* **4**, 212–229 (2012).

61. Cuypers, T. D. & Hogeweg, P. A synergism between adaptive effects and evolvability drives whole genome duplication to fixation. *PLoS Comput. Biol.* **10**, e1003547 (2014).

62. Cuypers, T. D., Rutten, J. P. & Hogeweg, P. Evolution of evolvability and phenotypic plasticity in virtual cells. *BMC Evolut. Biol.* **17**, 60 (2017).

63. Meijer, J., van Dijk, B. & Hogeweg, P. Archived VirtualMicrobes data supporting: "Contingent evolution of alternate metabolic network topologies determines whether cross-feeding evolves". https://doi.org/10.5281/zenodo.3840463 (2010).

64. Cuypers, T. D. & van Dijk, B. VirtualMicrobes: a scientific Python package for in silico evolutionary experiments. https://bitbucket.org/thocu/virtual-microbes (2019).

65. Meijer, J. R scripts to analyze VirtualMicrobes data. https://github.com/meijer-jeroen/contingent-evolution-2020 (2020).

## Acknowledgements

## Author contributions
J.M. wrote the paper with input from B.v.D. and P.H.; J.M., B.v.D. and P.H. contributed equally to the experimental design, analysis and interpretation of the results.

## Competing interests
The authors declare no competing interests.

## Additional information
**Supplementary information** is available for this paper at https://doi.org/10.1038/s42003-020-1107-x.

**Correspondence** and requests for materials should be addressed to J.M.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.