Syntrophic exchange in synthetic microbial communities

Michael T. Mee[a,b], James J. Collins[a,c,d,e], George M. Church[b,c,1], and Harris H. Wang[f,1]

[a]Department of Biomedical Engineering, Boston University, Boston, MA 02215; [b]Department of Genetics, Harvard Medical School, Boston, MA 02115; [c]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115; [d]Howard Hughes Medical Institute and [e]Center of Synthetic Biology, Boston University, Boston, MA 02215; and [f]Department of Systems Biology, Columbia University, New York, NY 10032

Contributed by George M. Church, April 1, 2014 (sent for review January 30, 2014)

Metabolic crossfeeding is an important process that can broadly shape microbial communities. However, little is known about specific crossfeeding principles that drive the formation and maintenance of individuals within a mixed population. Here, we devised a series of synthetic syntrophic communities to probe the complex interactions underlying metabolic exchange of amino acids. We experimentally analyzed multimember, multidimensional communities of *Escherichia coli* of increasing sophistication to assess the outcomes of synergistic crossfeeding. We find that biosynthetically costly amino acids including methionine, lysine, isoleucine, arginine, and aromatics, tend to promote stronger cooperative interactions than amino acids that are cheaper to produce. Furthermore, cells that share common intermediates along branching pathways yielded more synergistic growth, but exhibited many instances of both positive and negative epistasis when these interactions scaled to higher dimensions. In more complex communities, we find certain members exhibiting keystone species-like behavior that drastically impact the community dynamics. Based on comparative genomic analysis of >6,000 sequenced bacteria from diverse environments, we present evidence suggesting that amino acid biosynthesis has been broadly optimized to reduce individual metabolic burden in favor of enhanced crossfeeding to support synergistic growth across the biosphere. These results improve our basic understanding of microbial syntrophy while also highlighting the utility and limitations of current modeling approaches to describe the dynamic complexities underlying microbial ecosystems. This work sets the foundation for future endeavors to resolve key questions in microbial ecology and evolution, and presents a platform to develop better and more robust engineered synthetic communities for industrial biotechnology.

synthetic ecosystem | amino acid exchange | population modeling

Microbes are abundantly found in almost every part of the world, living in communities that are diverse in many facets. Although it is clear that cooperation and competition within microbial communities is central to their stability, maintenance, and longevity, there is limited knowledge about the general principles guiding the formation of these intricate systems. Understanding the underlying governing principles that shape a microbial community is key for microbial ecology but is also crucial for engineering synthetic microbiomes for various biotechnological applications (1–3). Numerous such examples have been recently described including the bioconversion of unprocessed cellulolytic feedstocks into biofuel isobutanol using fungal–bacterial communities (4) and biofuel precursor methyl halides using yeast–bacterial cocultures (5). Other emerging applications in biosensing and bioremediation against environmental toxins such as arsenic (6) and pathogens such as *Pseudomonas aeruginosa* and *Vibrio cholerae* have been demonstrated using engineered quorum-sensing *Escherichia coli* (7, 8). These advances paint an exciting future for the development of sophisticated multispecies microbial communities to address pressing challenges and the crucial need to understand the basic principles that enables their design and engineering.

An important process that governs the growth and composition of microbial ecosystems is the exchange of essential metabolites, known as metabolic crossfeeding. Entomological studies have elucidated on a case-by-case basis the importance of amino acids in natural interkingdom and interspecies exchange networks (9–11). Recent comparative analyses of microbial genomes suggest that a significant proportion of all bacteria lack essential pathways for amino acid biosynthesis (2). These auxotrophic microbes thus require extracellular sources of amino acids for survival. Understanding amino acid exchange therefore presents an opportunity to gain new insights into basic principles in metabolic crossfeeding. Recently, several studies have used model systems of *Saccharomyces cerevisiae* (12), *Saccharomyces enterica* (13), and *E. coli* (14–16) to study syntrophic growth of amino acid auxotrophs in coculture environments. Numerous quantitative models have also been developed to describe the behavior of these multispecies systems, including those that integrate dynamics (17, 18), metabolism (19–21), and spatial coordination (22). Although these efforts have led to an improved understanding of the dynamics of syntrophic pairs and the energetic and benefits of cooperativity in these simple systems (23), larger more complex syntrophic systems have yet to be explored.

Here, we use engineered *E. coli* mutants to study syntrophic crossfeeding, scaling to higher-dimensional synthetic ecosystems of increasing sophistication. We first devised pairwise syntrophic communities that show essential and interesting dynamics that can be predicted by simple kinetic models. We then increased the complexity of the interaction in three-member synthetic consortia involving crossfeeding of multiple metabolites. To further increase the complexity of our system, we devised a 14-member community to understand key drivers of population dynamics over short and evolutionary timescales. Finally, we provide evidence for widespread trends of metabolic crossfeeding based on comparative genomic analysis of amino acid biosynthesis across

Significance

Metabolic exchange between microbes is a crucial process driving the development of microbial ecosystems. The exchange of essential amino acids presents an opportunity to investigate the guiding principles underlying microbial trade in nature. In this study, we devised synthetic communities of *Escherichia coli* bacteria of increasing complexity to measure general properties enabling metabolic exchange of amino acids. We identified numerous syntrophic interactions that enable cooperative growth, which exhibited both positive and negative epistasis with increasing community complexity. Our results suggest that amino acid auxotrophy may be an evolutionarily optimizing strategy to reduce biosynthetic burden while promoting cooperative interactions between different bacteria in the microbiome.

Author contributions: M.T.M., J.J.C., G.M.C., and H.H.W. designed research; M.T.M. and H.H.W. performed research; M.T.M. and H.H.W. contributed new reagents/analytic tools; M.T.M. and H.H.W. analyzed data; and M.T.M., J.J.C., G.M.C., and H.H.W. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: gchurch@genetics.med.harvard.edu or hw2429@c2b2.columbia.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1405641111/-/DCSupplemental.

thousands of sequenced genomes. Our large-scale and systematic efforts represent an important foray into forward and reverse engineering synthetic microbial communities to gain key governing principles of microbial ecology and systems microbiology.

Results

Our overall goal is to develop and understand a simple microbial model of metabolic crossfeeding that can be scaled in a tractable stepwise manner, toward reconstituting the complexity and dynamics exhibited by natural ecosystems (Fig. 1*A*). To this end, we devised a series of syntrophic microbial communities of increasing diversity and complexity using the simple model bacterium *E. coli*. Our system is based on the syntrophic behavior of amino acid exchange between auxotrophic mutants to facilitate coculture growth. We first investigated the energetics involved in amino acid utilization and exchange. Starting from a prototrophic *E. coli* derivative MG1655, we generated 14 strains, each containing a gene knockout that lead to an auxotrophic phenotype of 1 of 14 essential amino acids (*Materials and Methods*). The remaining six essential amino acids were left out of our study because they either did not have single-gene targets that would render them auxotrophic or the resulting mutants carried significant growth defects even in richly supplemented media. By convention, we designate each auxotrophic strain by the amino acid they need—for example, the methionine Δ*metA* auxotroph is strain *M*. The 14 auxotrophs (*C*, *F*, *G*, *H*, *I*, *K*, *L*, *M*, *P*, *R*, *S*, *T*, *W*, and *Y*) were confirmed to show no growth in M9-glucose minimal media after 4 d and growth only when supplemented with each amino acid needed. Using a microplate spectrophotometer, we performed kinetic growth curve analysis for each of the 14 auxotrophs grown in M9-glucose supplemented with varying initial amino acid levels. Under these amino acid-limiting conditions, an auxotrophic strain will grow exponentially until the amino acid supplementation is exhausted (Fig. S1*B*). Saturating cell densities (i.e., carrying capacities) plotted against initial seeding amino acid levels show a strong linear relationship (Fig. S1*B*) indicating that external amino acid levels can determine cell growth in a linear and predictable manner. We estimated the number of extracellular amino acids needed to generate a cell for each of the 14 amino acids (Fig. S1*A* and Dataset S1). The estimated biosynthetic cost to produce each amino acid (24) shows a strong inverse relationship with the amount of amino acids needed to produce a cell (Fig. 1*B*), suggesting that the *E. coli* proteome has been optimized for amino acid use to maximize energetic efficiency. Inexpensively produced amino acids (e.g., *S*, *G*, and *P*) are used more frequently than expensive ones (e.g., *M*, *H*, aromatics). Although previous computational approaches have suggested such relationships (24), this is the first experimental measurement we are aware of regarding this important property. Because auxotrophic genotypes are prevalently observed in nature, our results suggest that many microbes may be subjected to this energetic optimization and that amino acid exchange may be an essential link guiding the evolution of these microbial ecosystems.

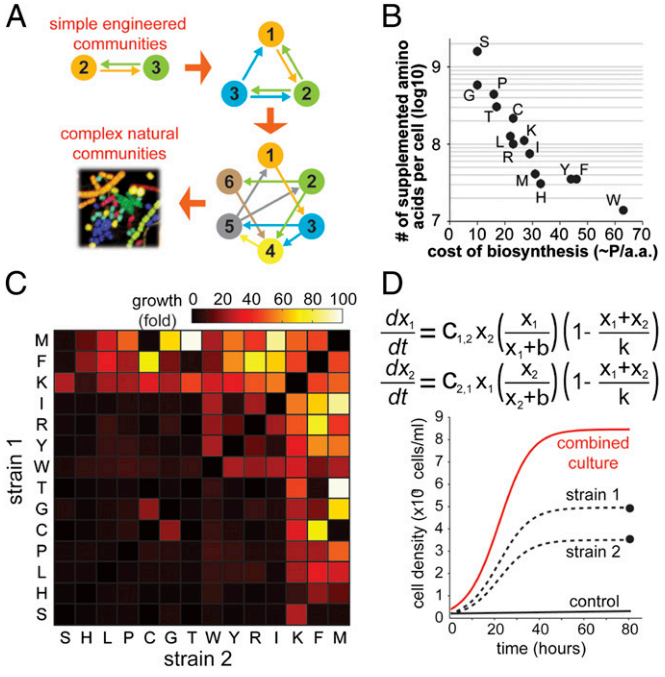

Fig. 1. Metabolic crossfeeding in syntrophic communities. (*A*) An illustration of engineered syntrophic interactions between microbial communities of increasing complexity toward network hierarchies matching those of natural systems. (*B*) Relationship between number of supplemented amino acid needed to make one *E. coli* cell in $\log_{10}$ units versus biosynthetic cost to produce each amino acid. (*C*) Syntrophic growth yield after 84 h between 14 single-KO auxotrophs (strain 1) and all pairwise combinations (strain 2). Color intensity indicated in the color bar denotes fold growth after 84 h over initial population. (*D*) Simple two-equation dynamic model that captures the essential features of the pairwise consortium. Cooperativity coefficients $C_{1,2}$ and $C_{2,1}$ can be determined through the total coculture growth curve (solid red line), the end point cell density of each strain (solid dots), and the simulated growth profile of each strain (black dotted lines). Control populations of only strain 1 or strain 2 (respectively ΔM and ΔF in this example) separately show no growth (solid black line).

To more deeply investigate the properties associated with metabolite exchange of amino acids in microbial communities, we developed synthetic ecosystems using the characterized auxotrophic *E. coli*. Studying synthetic communities using *E. coli* offers the benefits of robust and fast-growing cells, tractable genetics, and well-developed in silico models while maintaining a standardized and reproducible genetic background. We first developed a two-member syntrophic community composed of two different species from our 14 characterized auxotrophs. Each auxotroph is unable to grow in M9-glucose but could potentially grow as a coculture when paired with a different auxotroph. We probed all 91 possible pairwise syntrophic interactions in M9-glucose minimal media. In agreement with results from previous efforts (14), we observed significant synergistic and cooperative growth in a subset of these pairwise cocultures after 84 h (Fig. 1*C*). Pairings that involved crossfeeding of *M*, *F*, or *K* were highly cooperative with most of the 14 partners, whereas *I*, *R*, *Y*, and *W* had moderate cooperativity. *M*'s high crossfeeding productivity highlights the intuitive rule that partnering with direct biosynthesis precursors (*C* and *S* in this case) is unproductive. Crossfeeding between *T*, *G*, *C*, *P*, *L*, *H*, or *S* was generally nonproductive. Using quantitative PCR (qPCR) against the unique knockout chromosomal junction in each auxotrophic strain, we determined the relative abundance of each member of the pair (Dataset S2). In general, we did not find a significant relationship between coculture fold growth and the ratio between the pairwise consortium members, in contrast to previous results (14) (Fig. S2). Furthermore, our results suggest that expensive and rarer amino acids tended to crossfeed better than cheaper and more common amino acids. A set of simple dynamic equations was used to model these syntrophic interactions (Fig. 1*D*). In our model, growth of strain 1 is linearly proportional to the amount of strain 2 through the cooperativity coefficient term (c) for each pairing. The coculture reaches saturation at carrying capacity (k), and a buffer term (β) is used for low-density growth. We determined the cooperativity coefficients that best fit our experimental dataset, which provided the foundation for modeling higher-dimensional interactions (Dataset S2).

To further explore the scalability of our synthetic ecosystems, we turned to higher-order syntrophic interactions. We devised three-member synthetic consortia where each member is auxo-

trophic for two amino acids. Growth of each member can only occur when both amino acids are provided by partner strains. Member 1 (e.g., double-auxotroph *MF*) does not grow with member 2 (e.g., double-auxotroph *MK*) because both are auxotrophic for *M*, methionine, but they can potentially coculture together with a third member (e.g., double-auxotroph *FK*) that can provide *M*. At the same time, member 3 can only grow in the presence of members 1 and 2, thereby forming a syntrophic community. We experimentally probed these three-member syntrophies by first generating all 91 double-amino acid auxotrophic derivatives based on the 14 single-amino acid auxotrophs (*Materials and Methods*). All 91 strains showed no growth in the absence of extracellular supplementation of both needed amino acids. Using these 91 double auxotrophs, we systematically measured the growth profiles of all 364 possible three-member consortia. After coculture for 84 h, we find a significant number of three-member consortia to grow synergistically (Fig. 2). As a control, growth was not observed when only two of the three members are cocultured. Given that many microbes in nature are unable to synthesize multiple amino acids, our results demonstrated that higher-dimensional crossfeeding can yield productive syntrophic groups and are likely relevant also in natural microbial ecosystems.

The overall growth profiles of the doubly auxotrophic three-member consortia match those of the monoauxotrophic two-member consortium (Fig. 3*A*). Cocultures involving amino acids *M*, *F*, and *K* tend to exhibit strong cooperative growth in contrast to poorly syntrophic amino acids *H*, *C*, and *S*. To assess the predictability of syntrophic interactions when scaled to higher-dimensional communities, we compared the observed fold growth

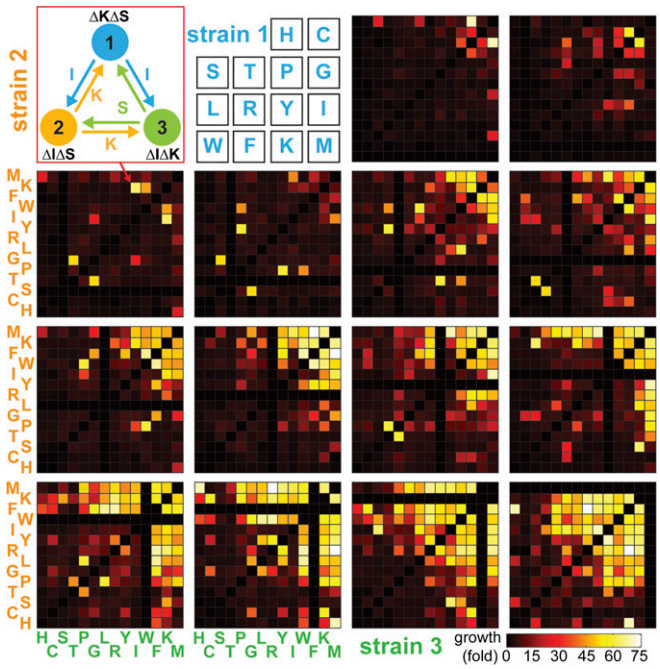

Fig. 2. Three-member syntrophic consortia with each strain being auxotrophic for two amino acids. All combinations of 14 × 14 × 14 three-way interactions are measured after 84 h of growth. Fourteen 14 × 14 panels are presented showing the growth yield of each three-member group. Each 14 × 14 panel correspond to a fixed strain 1 (blue color) against all combination of strains 2 and 3. The ordinate axis denote different strain 2 (orange color), and the abscissa axis denote different strain 3 (green color). The key for strain 1 designation is shown in the second panel. The first panel illustrates an example consortium of KS-IS-IK with the crossfeeding amino acids shown by the correspondingly colored arrows. Color intensity indicated in the color bar denotes fold growth after 84 h over initial population.

of all 364 three-member consortia with the fold growth of each of their two-member subset—for example, comparing doubly auxotrophic triplets MF-MK-FK versus each of the monoauxotrophic pairs M-F, M-K, or F-K. We found four general classes of observations (Fig. 3*B*, zones 1–4). First, as expected, a majority of nonproductive two-member crossfeeding interactions when scaled to three-members also yield nonproductive growth (Fig. 3*B*, zone 1). Conversely, three-member interactions where all two-member subsets are productively crossfeeding also generate highly productive cocultures (Fig. 3*B*, zone 2). Third, three-member interactions where one of the two-member subsets is nonproductive generally resulted in nonproductive three-way cocultures (Fig. 3*B*, zone 3). For example, the RMT triplet (RM-RT-MT) has a very limited growth of threefold even though the individual two-member R-M pair yields 43-fold growth and the M-T pair yields 98-fold growth. The limiting group is the two-member pair R-T, which yields less than onefold growth in coculture. Thus, the RMT triplet does not grow due to limited R-T crossfeeding despite robust crossfeeding by R-M and M-T. Finally, we find that a small group of nonproductive two-member coculture when placed together showed positive epistatic synergy, resulting in more productive three-way cocultures (Fig. 3*B*, zone 4). Fig. 3*C* summarizes the top consortia that exhibited this higher-dimensional synergy (Dataset S2). For example, the PYT triplet culture (PY-PT-TY) results in 58-fold growth, in comparison with pairwise cultures P-Y, P-T, and T-Y that only grew by eightfold, twofold, and twofold, respectively. These interactions highlight the surprising potential for both positive and negative epistatic interactions that exist for higher-dimensional syntrophic communities.

To assess whether a naïve model can predict the observed growth profile of our three-member consortia, we used the cooperativity coefficients (c) derived from the pairwise two-member interaction to build a simple three-member dynamic model, which is described by the following system of equations:

$$\frac{dX_1}{dt} = \frac{X_1}{X_1+\beta}\cdot \min\left(c_{1,2}X_2, c_{1,3}X_3\right)\left(\frac{1-(X_1+X_2+X_3)}{k}\right),$$

$$\frac{dX_2}{dt} = \frac{X_2}{X_2+\beta}\cdot \min\left(c_{2,1}X_1, c_{2,3}X_3\right)\left(\frac{1-(X_1+X_2+X_3)}{k}\right),$$

$$\frac{dX_3}{dt} = \frac{X_3}{X_3+\beta}\cdot \min\left(c_{3,1}X_1, c_{3,2}X_2\right)\left(\frac{1-(X_1+X_2+X_3)}{k}\right),$$

where X_1, X_2, and X_3 are the population sizes of the three members, β is the buffer term, and k is the carrying capacity of the population. Growth of strain 1 (dX_1/dt) is dictated either by the amount of strain 2 (X_2) times its cooperative coefficient ($c_{1,2}$) or the amount of strain 3 (X_3) times its cooperative coefficient ($c_{1,3}$), whichever is the limiting value. The $X_i/(X_i+\beta)$ term is used such that at very low X_i levels, growth is proportional to X_i/β as with standard exponential growth, but $X_i/(X_i+\beta)$ becomes 1 at moderate X_1 values. The last term $(1-(X_1+X_2+X_3))/k$ is used to limit the density of the saturating culture. Using this simple dynamic model, we find that the predicted fold-growth profile showed statistically significant correlation with the observed three-member fold growth (Pearson coefficient $r = 0.51$, P value $= 5.2\times10^{-25}$). When the zones representative of positive and negative epistatic interactions are removed (3 and 4), the fit of the model is greatly improved (Pearson coefficient $r = 0.655$, P value $= 1.2\times10^{-35}$; see Dataset S3 for fits for all zones). This suggests that our three-member model can capture a significant fraction of syntrophic interactions based on the two-member interactions (Fig. S3 and *SI Materials and Methods*). Discordance between the model and observed results highlight the potentially nonlinearly synergistic or antagonistic interactions of certain syntrophic consortia worthy of follow-up studies.

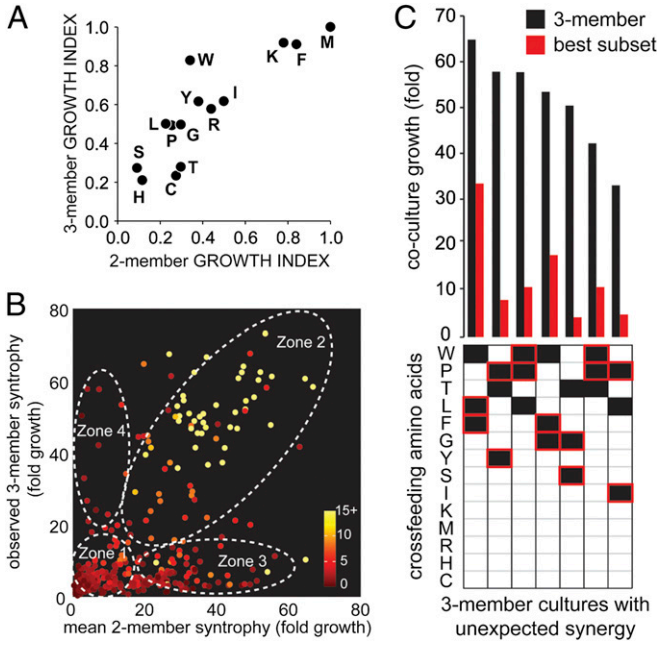

Fig. 3. Comparison of three-member syntrophies composed of double auxotrophs against two-member composed of single auxotrophs. (*A*) The sums of the final OD values for all two- or three-member communities containing a given auxotroph are normalized to the highest value (ΔM for both two- and three-member systems) to represent the syntrophic exchange growth potential of a given amino acid. This is termed the growth index for the three-member or two-member scenarios and shows consistent relationship when crossfeeding is scaled to higher dimensions. (*B*) Observed three-member growth (ordinate axis) for all 364 triplets versus the mean growth of their three corresponding two-member subsets (abscissa axis). Each point corresponds to a specific three-member group. Color intensity of each point designates the lowest growth yield of the three two-member subset and is mapped based on color bar using numerical values of fold growth (0–15+). Zones 1–4 are designated in the dotted regions (see text for detail). (*C*) Growth yield of top three-member consortia that grow better than their corresponding two-member subsets. The black bar indicates growth of three member. The red bar is the highest growth yield of the three two-member subsets. The corresponding three amino acids are shown in the bottom panel (read vertically) for each triplet. Two red boxes for each triplet designate the best two-member subset.

The microbial communities explored thus far are "strictly syntrophic (ss)" because all strains rely on each other to grow and no subset grouping can grow alone. However, most natural ecosystems are likely composed of "partially syntrophic (ps)" interactions where one or more of the subset grouping can grow by themselves. For example, in contrast to the strictly syntrophic three-member group MF-MK-FK that requires all members present, the partially syntrophic group MF-MK-HF contains a two-member subset (MK-HF) that can potentially grow without the third member (MF). At longer timescales, one might predict that all "partially syntrophic (*n*)-member" interactions are reduced to a more minimal strictly syntrophic consortium of fewer members. Thus, the evolution of partial syntrophy to strict syntrophy is of great importance to formation of sustainable crossfeeding consortia.

We tested the predicted reduction from partial to strict syntrophy by devising a synthetic consortium using the 14 monoauxotrophs that had been characterized. Combining all 14 auxotrophs into one pool produces a partially syntrophic 14-member consortium (Fig. 4*A*). In principle, each auxotroph requires only one other partner to survive and thus may reduce over time to a strictly syntrophic two-member population. Pairs that are able to grow the fastest are likely enriched. This system highlights the scenario where balancing between competitive growth and maintenance of cooperativity is key. Based on our prior two-member crossfeeding results, we first constructed a 14-member interaction map (Fig. 4*B*) and devised the following dynamic model to predict the possible outcomes of this 14-member consortium:

$$\dot{X}_i = \frac{X_i}{X_i + \beta}\left(\sum_{j=1, j\neq i}^{14} c_{ij}X_j\right)\left(\frac{1-\sum_{j=1}^{14} X_j}{k}\right).$$

Here, X_i is the population size for strain i and c_{ij} is the cooperative coefficient between each two-member pair i and j that we determined previously, and terms β and k are as described previously. Several general predictions are noted. First, the resulting population is likely to be dominated by strains *M*, *F*, *K*, *I*, *R*, *T*, and *W* strains based on their cooperativity profiles (Fig. S4). Strain *M* is likely to sustain *I*, *T*, and *K* due to the directionality of the cooperativity and thus may be a hub for any syntrophic interaction. Strain *F* is likely to derive benefit from multiple sources including *Y*, *C*, *I*, and *K*, while only contributing to the growth of *R*. Strain *K* is likely to benefit from many strains including *M*, *T*, *R*, *C*, and *Y*. Strains *L*, *W*, and *P* are expected to have modest contributions. Finally, we expect strains *H*, *G*, *C*, *Y*, and *S* to not be major components of the consortium because they do not generally yield productive pairwise crossfeeding. Further discussions of these predictions are provided in *SI Materials and Methods*.

Starting from a 10^7 cells per mL population composed of equal number of each of the 14 monoauxotrophic strain, we experimentally passaged two identical replicate populations in M9-glucose for 50 d (~400 generations) and tracked the population abundance using qPCR (Fig. 4*A*). At the end of each 24-h period, the population reached saturation and was diluted 100-fold for the next passage. At short timescales, we find that the 14-member systems undergo a drastic population shift toward a consortium dominated by four members (*R*, *K*, *M*, and *T*) after only two to three daily passages (Fig. 4*C*). Eventually, *R* is replaced by *I* and the population ratio varied from time to time but the member composition remained stable over the course of 50 d. The experiment was terminated at 400 generations when mutants were discovered in both replicates that confounded basic interpretations. We thus focused on the short-term dynamics of the system that are not subject to mutational events at evolutionary timescales. Satisfactorily, the *K*, *M*, *R*, *I*, and *T* strains dominant in the passaged population are also predicted to be predominant by our dynamic model. However, aromatic auxotrophs (*F*, *W*) were not seen despite their capabilities to crossfeed with others during individual pairwise matchups (Fig. 1*B*). These pairwise interactions almost always benefited the *F* or *W* strain more than the partner strain, which may be in part responsible for their absence in the total mixture. Nonetheless, the model was able to capture a majority of the basic and important features of this otherwise complex community.

To further probe the structure of the syntrophic network, we systematically tested 13-member consortia where one of the four initial dominant strains (*R*, *K*, *M*, *T*) was left out of the population. The composition of the mixture was tracked over seven daily passages (Fig. 4*C*). A number of interesting observations were seen. In contrast to the 14-member population, a 13-member consortium absent of the *K* auxotroph resulted in stable dominance of the *M*-*T* coculture only (Fig. 4*C*, panel 2). We note that the biosynthesis pathways of methionine and threonine converge upstream at a single common precursor, L-homoserine, and speculate the *M* auxotroph (Δ*metA*) could result in shunting biosynthetic flux in this branched synthesis pathway toward increasing *T* production. Conversely, the T auxotroph (Δ*thrC*) could result in the opposite shunt leading to increased *M* biosynthesis and has been shown to lead to increased M excretion

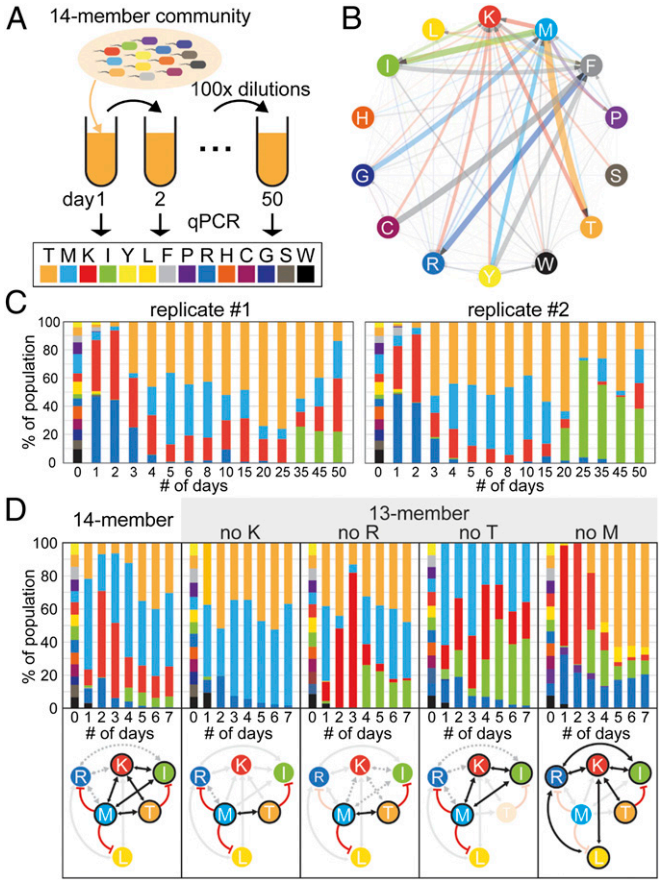

Fig. 4. Dynamics of a 14-member syntrophic consortium. (*A*) Fourteen different single-amino acid auxotrophs where combined in equal ratios to form a pooled mixture and passaged daily in minimal media over 50 d. Samples of the population were periodically measured to determine the absolute and relative abundance of each of the 14 auxotrophs. (*B*) Syntrophic interaction map generated from all measured 91 pairwise crossfeeding experiments. Each auxotroph is designated by a circle and a different color. The arrowed lines correspond to the directional interaction from each strain to all 13 corresponding partners. Lines are color-coded according to the directional benefit the receiving strain is gaining from the donor strain (e.g., all incoming lines to the *K* auxotroph are red, designating the benefit gained by *K* from each donor). Increased thickness and opacity of the lines quantitatively denote increased cooperative benefits. (*C*) Population distribution of two biological replicate 14-member populations over 50 daily passages. Each colored bar section denotes the fractional composition of each auxotroph in the population. Color coding is the same as that of *A* and *B*. (*D*) Subsequent short-term 7-d experiments of the 14-member population as well as 13-member populations that excluded one of four dominant amino acids (K, R, T, or M) from the initial population. The syntrophic interaction network is shown below each panel. We denoted cooperative interactions with bidirectional black arrows and competitive (seemingly inhibitory) effects by directional blunted red arrows. Each auxotroph dropout and their associated interactions are shown in faded colors. Transient cooperative interactions are shown as dotted gray arrowed lines. The black circles around each amino acid designate final fixation to a stable community of two to five members. Values are derived from the average of three to four biological replicates. Time series data are included in Dataset S2.

(25). Indeed, we find that the *M-T* pair yielded the highest fold growth among all 91 pairwise cocultures (Fig. 1*C*), suggesting that *M* and *T* could provide additional cooperative benefit to one another specifically for this pairing. Furthermore, the presence of characterized exporters for L-methionine and L-threonine could further facilitate syntrophy in this subgroup (26).The synergistic effect of branched pathway shunt is also seen for the M-I, K-I, K-T, and G-C pairs, although to a lesser extent. Lower synergism in these pairs may be due to increasing numbers of intermediates and potential for branch down-regulation as previously suggested (14). Additionally, nonexclusive shunting could diminish the effect size for cases such as G-C where ΔglyA and ΔcysE can redirect metabolic flux to a third parallel pathway, toward W in this case (Fig. S5). Similarly, we found that the *R*-absent population also led to the dominant *M-T* coculture, although the dynamics of the population was very different (Fig. 4*C*, panel 3). In the *R*-absent population, the *I* auxotroph bloomed, making up 20% of the population at day 4, but was eventually outcompeted in subsequent days. In contrast, the *I* auxotroph was not detected in the *K*-absent population. Interestingly, a *T*-absent population (Fig. 4*C*, panel 4) resulted in expansion of the *I* subpopulation as well as the maintenance of the *K* subpopulation, thus producing a stable *M-K-I* consortium. Together, these results suggest that the *T* and *I* auxotrophs are competing for similar cooperating partners such that the presence of *T* limits the growth capability of *I*, but absence of *T* allows growth of *I*. *T*'s competitive advantage over the *I* strain in this mixed environment could be explained by the fact that biosynthesis of *I* requires *T* as an essential precursor, and therefore *I* is likely sharing an additional costly metabolite (*T*) with the other strains. Finally in *M*-absent population, *L* is present in addition to *K*, *I*, and *R*. In contrast to the other mixtures, the *R* subpopulation was stably sustained over the course of 7 d. We integrated all these subtractive experiments to form a reconstructed topology of the syntrophic network (Fig. 4*C*). This interaction network recapitulates the important properties of the partially syntrophic community and highlights the role of transient dynamics in the development of syntrophic community.

The general patterns of amino acid crossfeeding in *E. coli* led us to hypothesize that amino acid exchange may be an important property across many microbial communities in the natural biosphere. To evaluate this hypothesis, we compiled the frequencies at which bacterial genomes were predicted to contain complete and intact biosynthetic pathways for each of the 20 essential amino acids (*Materials and Methods*). We analyzed 6,120 sequenced genomes through the Integrated Microbial Genomes (IMG) database and pipeline (27) (Fig. 5*B*). When we plotted the amino acid biosynthetic potential, we find that most bacteria are able to produce amino acids *E*, *G*, *N*, and *Q*, whereas only a surprisingly small fraction of bacteria are able to produce amino acids *K*, *H*, *F*, and *Y*. When the biosynthetic potential is plotted against the estimated metabolite cost to produce each amino acid, we find that the more expensive amino acids (e.g., *F*, *W*, *Y*) tend to be made less prevalently than inexpensive ones (e.g., *E*, *G*, *N*, *Q*) (Fig. 5*C*). This result suggests the microbial biosphere has been optimized such that costly but essential resources (i.e., amino acids) are made by only a small fraction of the members. To understand the structure of amino acid biosynthesis at the biosphere scale, we mapped the prototrophy prediction distribution across a phylogenetic tree of ∼2,000 bacteria from diverse environments (*Materials and Methods*). Interestingly, we find that biosynthesis capabilities are highly structured when clustered at the phylum and class level (Fig. 5*A*). For example, production of the most costly amino acids (e.g., *F*, *Y*) is concentrated in the closely related *Betaproteobacteria* and *Gammaproteobacteria*, whereas production of lysine is highly enriched in *Deltaproteobacteria* and *Cyanobacteria*. Similarly, the closely related intracellular parasitic phyla *Tenericutes* and *Chlamidiae* have similar biosynthesis profiles, making almost uniquely the amino acids *N*, *G*, and *Q*. Similar structure is apparent when looking at biosynthesis distribution on the species level. Within the observed *Bacteroidetes* (Fig. 5*D*), amino acid production profiles closely follow species relationships with certain clades predicted to have broad production abilities (several clades within the *Bacteroidia* class), whereas others have similar limited capabilities (*Flavobacteria*). Taken together, these results provide evidence suggesting that microbial genomes optimize their metabolic potential to reduce bio-

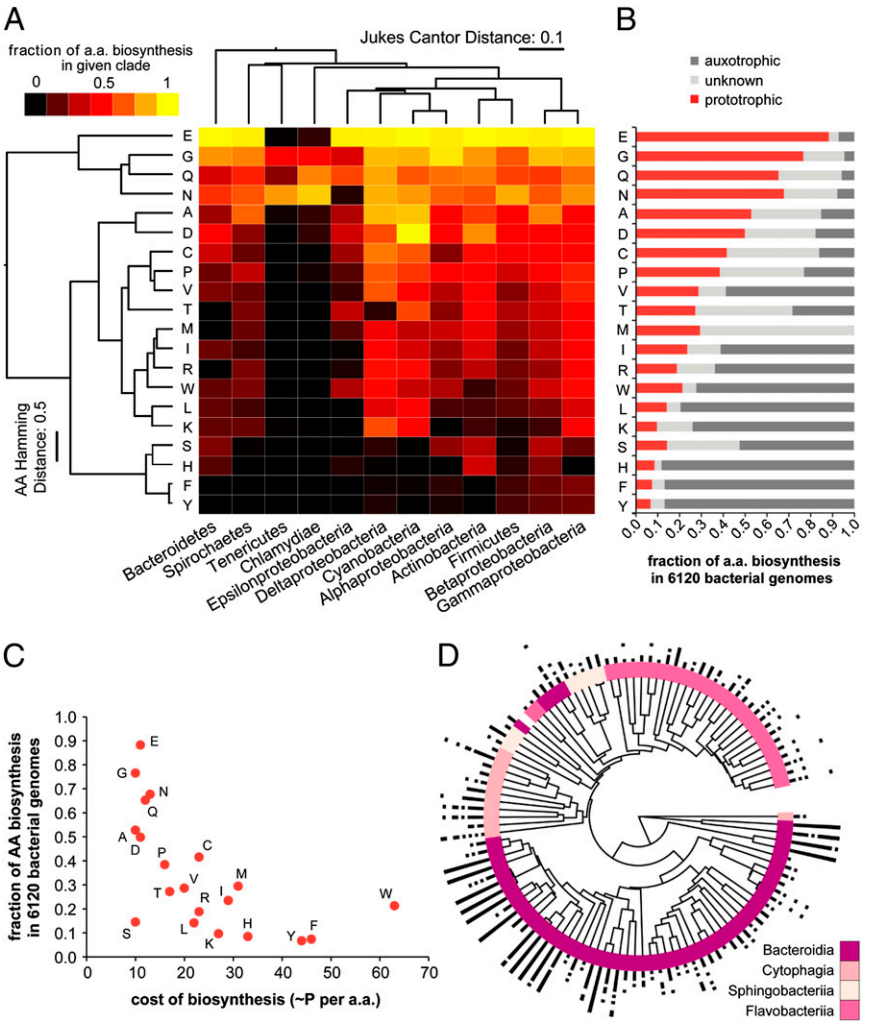

Fig. 5. Amino acid biosynthesis in the microbiome. (*A*) Heat map of amino acid biosynthetic capabilities of the indicated phyla and classes. Prototrophy predictions for each amino acid are averaged within groups. Phyla and classes with fewer than 20 of the 2,099 sequenced bacterial are excluded from the analysis. Value of 1/0 indicates all/none of the species in the group are prototrophic. Dendograms represent clustering of both phylogenetic distribution (based on median 16S sequence from each clade) and amino acid production profiles. Phylum/class leaf branches are not to scale to enhance higher-order relationships. (*B*) Distribution of amino acid biosynthetic capability of 6,120 sequenced bacteria. The red bars indicate complete pathway present. Incomplete or unknown pathways are denoted in black and gray bars, respectively. (*C*) Amino acid biosynthesis distribution plotted against metabolic cost of synthesizing each amino acid in terms of number of phosphates required. (*D*) Prototrophy distribution in *Bacteroidetes*. The black rings indicate biosynthesis of each amino acid in increasing order of prevalence from inner to outer rings (E, G, N, D, Q, C, A, S, V, I, P, K, L, W, H, M, R; T, F, Y not present).

synthetic burden, and that microbes may be tactically leveraging the specialized biosynthetic capabilities of their neighbors while reciprocating through mutualistic trade of essential metabolites.

Discussion

In natural ecological settings, metabolite exchange between cooperative organisms is often difficult to identify, the interactions are hard to quantitatively measure, and the resulting significance are challenging to interpret. Here, we systematically probed the syntrophic interactions driving metabolite exchange between amino acid auxotrophs to determine the general principles of cooperative crossfeeding of essential resources in a synthetic microbial community. Our results lead us to speculate that microbial syntrophy may contribute significantly to the development and persistence of auxotrophic phenotypes of otherwise essential metabolites in the biosphere including environments such as marine, terrestrial, plant roots, and human-associated surfaces.

Pairwise amino acid crossfeeding experiments showed that these simplified interactions can be quantitatively modeled. When these interactions are scaled to three-membered communities requiring cooperative behavior by all individuals involved, we also found many examples of syntrophic growth. The general trends in cooperative exchange were conserved in these higher-dimensional communities. Furthermore, we observed positive epistasis within a small subset of three-member consortia that performed better than their two-member constituents. It is feasible to further increase the complexity of the system by developing triply auxotrophic strains that interact as four-member strictly syntrophic communities. These cocultures demonstrate that microbes with multiauxotrophic phenotypes can rely on direct crossfeeding for survival.

Our 14-member consortium showed that crossfeeding interactions can often be quite complex and the system may not necessarily converge to an expected simpler two-member community through hundreds of generations. Various resulting three- or four-member populations were often stable over evolutionary timescales, but removal of certain keystone members (e.g., K or R) from the initial population resulted in convergence to the best pair M-T. Although

our dynamic models were able to capture general features of this system, specific quantitative predictions were less accurate, thus highlighting the current limitations to model ecosystems of even moderate complexity (28).

It is important to note that microbial crossfeeding studied here relies on export of amino acids into the shared extracellular environment. Such membrane transport systems have been recently characterized and many more are being found with the help of metagenomic sequencing. Our *E. coli* genome encodes several amino acid exporters for excretion of different amino acids including L-threonine (RhtA and RhtC) (26), L-leucine (YeaS) (29), L-aromatic amino acids (YddG) (30), L-arginine (YggA) (31), L-alanine (alaE) (32), and L-homoserine (RhtB) (26). Other exporters have been documented in related organisms including lysE for L-lysine export and brnFE for L-isoleucine and L-methionine export in *Corynebacterium glutamicum* (33, 34). In addition to active transport, some level of passive transport may also be involved because hydrophobic amino acids such as I, F, Y, and W have membrane permeability that is 100 times greater than hydrophilic amino acids (35). Recently, the YddG aromatic amino acid exporter has been exploited to tune microbial crossfeeding in an *E. coli* Y-W syntrophic system (15). These systems present an opportunity to study long-term evolution of microbial ecosystems and the enhancement of cooperative phenotypes. Although metabolic crossfeeding could be exploited by selfishly cheating phenotypes, the formation of spatial architectures such as biofilms and aggregates may help to prevent such scenarios. Furthermore, quorum sensing and response to the presence of cooperators may further help to drive the development of multispecies syntrophic growth.

We believe that the multidimensional syntrophic system presented here provide a useful foundation for studying and engineering microbial communities of increasing sophistication. These synthetic approaches can be used to study natural microbial communities such as the human microbiome in specific ways toward unraveling the complex interactions at play (36). Advances in synthetic and systems ecology will offer new avenues to explore and exploit natural and defined microbiota to develop sustainable solutions to global health, energy, and environmental issues.

Materials and Methods

Strain Construction and Verification. All strains used were based on the EcNR1 *E. coli* derivative of MG1655, which carried an integrated, temperature-inducible, λ-Red prophage for recombineering (37). Each of the 14 amino acid auxotroph was generated by Red-recombineering as previous described (38) of a chloramphenicol resistance cassette into each of the following targets separately: *argA* (R), *cysE* (C), *glyA* (G), *hisB* (H), *ilvA* (I), *leuB* (L), *lysA* (K), *metA* (M), *pheA* (F), *proA* (P), *serA* (S), *thrC* (T), *trpC* (W), and *tyrA* (Y). Double-amino acid auxotrophs were generated by introducing a kanamycin resistance cassette into the same set of targets in the 14 single-amino acid auxotrophs to generate the 91 double-auxotroph strains. All single-amino acid knockout strains were confirmed to be auxotrophic as evidenced by the lack of growth in minimal M9-glucose media after 4 d. All double-amino acid knockout strains were confirmed to not grow in M9-glucose supplemented with only one of the two amino acids. Presence of the inserted antibiotic cassettes was additionally verified in each strain via allele-specific PCR. All primers used are provided in Dataset S4.

Coculture Growth Conditions. Strains were first picked from an overnight colony into LB–Lennox medium (10 g/L bacto tryptone, 5 g/L NaCl, 5 g/L yeast extract) with selective antibiotics as appropriate (chloramphenicol, 20 μg/mL; kanamycin, 50 μg/mL). Late–exponential-phase cells were harvested and washed twice in M9 salts (6 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 1 g/L NH_4Cl, 0.5 g/L NaCl) by centrifugation at 17,900 × *g*. Cell concentrations were determined based on OD_{600} readings from a spectrophotometer. Before coculture experiments, all cell concentrations were adjusted to 10^7 cells per mL using M9 media. Coculture growth was performed by equal-volume inoculation of each strain at a seeding density of 10^7 cells per mL. All two-member and three-member cocultures were grown in 200 μL of M9-glucose media [M9 salts supplemented with 1 mM $MgSO_4{\cdot}7H_2O$, 0.083 nM thiamine, 0.25 μg/L D-biotin, and 0.2% (wt/vol) glucose] in 96-well microtiter plate format in an incubator or a plate reader at 30 °C to maintain λ-Red prophage repression. Microtiter plates were shaken at >500 rpm to maintain aerobic growth. Growth of 13- and 14-member cocultures was done in 3-mL cultures in a 30 °C rotating drum and passaged without washing by 100-fold dilution every 24 h as the cultures reach saturation. The growth and fold-growth metrics mentioned throughout the text refer to the yield of the community calculated by final cell density/initial density. Biological replicates were performed by splitting a single well-mixed initial seeding population.

Kinetic Growth Assays and Strain Identification. For precise determination of the cell density at amino acid-limited conditions, OD_{600} readings were taken every 5 min during exponential growth at 30 °C in a spectrophotometer (M5; Molecular Devices) with >500 rpm orbital mixing. Calibration between OD_{600} measurement and actual cell density was determined by resolving the colony-forming units of each auxotrophic strain plated on solid media at different concentrations. All growth experiments were performed in M9-glucose media in the absence of antibiotics. Proportional strain abundance was determined via qPCR. All population samples were frozen at −20 °C and assayed simultaneously to reduce run-to-run variations. qPCRs were performed in 20-uL reactions with 10 μL of KAPA SYBR Fast Universal 2× MasterMix (KAPA KK4600), 4 μL of a 10× dilution of frozen cells, and 6 μL of primer pairs resulting in the following final primer concentrations: R, K, M, P, T, W, Y (200 nM); G, I (150 nM); H, S (100 nM); C, L, F (50 nM). PCR conditions were based on manufacturer's recommendations (40 cycles of combined annealing extension at 60 °C for 20 s) and performed using a thermal cycler (Bio-Rad CTX96). The corresponding number of cycles was determined at a relative fluorescence unit of 150. The half-max values of each qPCR curve was calibrated to actual cell density by serial dilution plating of each auxotrophic strains and determination of colony-forming units on solid agar plates. The 13- and 14-member qPCR control populations were performed using equimolar mixtures of all auxotrophic strains. Relative proportion of each strain the population as determined by qPCR was further verified by plating using colorimetric assays. For two-member and three-member communities, each strain carried deletions Δ*malK*, Δ*lacZ*, or Δ*malK*/Δ*lacZ*, which could be visually distinguished on MacConkey-maltose plates (BD Difco) supplemented with XGAL-IPTG (Growcells).

Phylogenetic Analysis. Predicted amino acid biosynthesis phenotype was extracted from the IMG database (27) for all available sequenced bacterial genomes. Phylogenetic linkage between these strains is determined using the corresponding aligned 16S sequences from the SILVA database (39). Distances between sequences are determined with the standard Jukes–Cantor metric and hierarchical clustering performed with unweighted average distances. Phylogenetics trees were constructed using the iTol web application (40). Amino biosynthesis profiles were clustered at the species level using Hamming distance and complete linkage. Biosynthesis profiles at the phylum and class level were determined by averaging prototrophy predictions for each amino acid across all represented species. Only the 12 clades with more than 20 species were clustered with Euclidian distance and complete linkage.

ACKNOWLEDGMENTS. H.H.W. acknowledges funding from National Institutes of Health Director's Early Independence Award 1DP5OD009172-01. M.T.M. is supported through funding from Department of Energy (DOE) Grant DE-FG02-02ER63445 and a Doctoral Study Award from the Canadian Institutes of Health Research. J.J.C. acknowledges funding from the Howard Hughes Medical Institute. G.M.C. acknowledges funding from the DOE (as above) and National Science Foundation Grant SA5283-11210.

1. Brenner K, You L, Arnold FH (2008) Engineering microbial consortia: A new frontier in synthetic biology. *Trends Biotechnol* 26(9):483–489.
2. Mee MT, Wang HH (2012) Engineering ecosystems and synthetic ecologies. *Mol Biosyst* 8(10):2470–2483.
3. Shong J, Jimenez Diaz MR, Collins CH (2012) Towards synthetic microbial consortia for bioprocessing. *Curr Opin Biotechnol* 23(5):798–802.
4. Minty JJ, et al. (2013) Design and characterization of synthetic fungal-bacterial consortia for direct production of isobutanol from cellulosic biomass. *Proc Natl Acad Sci USA* 110(36):14592–14597.
5. Bayer TS, et al. (2009) Synthesis of methyl halides from biomass using engineered microbes. *J Am Chem Soc* 131(18):6508–6515.
6. Prindle A, et al. (2012) A sensing array of radically coupled genetic "biopixels.". *Nature* 481(7379):39–44.
7. Saeidi N, et al. (2011) Engineering microbes to sense and eradicate *Pseudomonas aeruginosa*, a human pathogen. *Mol Syst Biol* 7:521.
8. Duan F, March JC (2010) Engineered bacterial communication prevents *Vibrio cholerae* virulence in an infant mouse model. *Proc Natl Acad Sci USA* 107(25):11260–11264.

9. McCutcheon JP, von Dohlen CD (2011) An interdependent metabolic patchwork in the nested symbiosis of mealybugs. *Curr Biol* 21(16):1366–1372.
10. Wu D, et al. (2006) Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. *PLoS Biol* 4(6):e188.
11. Russell CW, Bouvaine S, Newell PD, Douglas AE (2013) Shared metabolic pathways in a coevolved insect-bacterial symbiosis. *Appl Environ Microbiol* 79(19):6117–6123.
12. Shou W, Ram S, Vilar JM (2007) Synthetic cooperation in engineered yeast populations. *Proc Natl Acad Sci USA* 104(6):1877–1882.
13. Harcombe W (2010) Novel cooperation experimentally evolved between species. *Evolution* 64(7):2166–2172.
14. Wintermute EH, Silver PA (2010) Emergent cooperation in microbial metabolism. *Mol Syst Biol* 6:407.
15. Kerner A, Park J, Williams A, Lin XN (2012) A programmable *Escherichia coli* consortium via tunable symbiosis. *PLoS One* 7(3):e34032.
16. Pande S, et al. (2013) Fitness and stability of obligate cross-feeding interactions that emerge upon gene loss in bacteria. *ISME J*, 10.1038/ismej.2013.211.
17. Bull JJ, Harcombe WR (2009) Population dynamics constrain the cooperative evolution of cross-feeding. *PLoS One* 4(1):e4115.
18. Estrela S, Gudelj I (2010) Evolution of cooperative cross-feeding could be less challenging than originally thought. *PLoS One* 5(11):e14121.
19. Stolyar S, et al. (2007) Metabolic modeling of a mutualistic microbial community. *Mol Syst Biol* 3:92.
20. Mazumdar V, Amar S, Segrè D (2013) Metabolic proximity in the order of colonization of a microbial community. *PLoS One* 8(10):e77617.
21. Klitgord N, Segrè D (2010) Environments that induce synthetic microbial ecosystems. *PLoS Comput Biol* 6(11):e1001002.
22. Nadell CD, Foster KR, Xavier JB (2010) Emergence of spatial structure in cell groups and the evolution of cooperation. *PLoS Comput Biol* 6(3):e1000716.
23. Wintermute EH, Silver PA (2010) Dynamics in the mixed microbial concourse. *Genes Dev* 24(23):2603–2614.
24. Akashi H, Gojobori T (2002) Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc Natl Acad Sci USA* 99(6): 3695–3700.
25. Usuda Y, Kurahashi O (2005) Effects of deregulation of methionine biosynthesis on methionine excretion in *Escherichia coli*. *Appl Environ Microbiol* 71(6):3228–3234.
26. Zakataeva NP, Aleshin VV, Tokmakova IL, Troshin PV, Livshits VA (1999) The novel transmembrane *Escherichia coli* proteins involved in the amino acid efflux. *FEBS Lett* 452(3):228–232.
27. Markowitz VM, et al. (2012) IMG: The Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res* 40(Database issue):D115–D122.
28. Klitgord N, Segrè D (2011) Ecosystems biology of microbial metabolism. *Curr Opin Biotechnol* 22(4):541–546.
29. Kutukova EA, et al. (2005) The yeaS (leuE) gene of *Escherichia coli* encodes an exporter of leucine, and the Lrp protein regulates its expression. *FEBS Lett* 579(21): 4629–4634.
30. Doroshenko V, et al. (2007) YddG from *Escherichia coli* promotes export of aromatic amino acids. *FEMS Microbiol Lett* 275(2):312–318.
31. Nandineni MR, Gowrishankar J (2004) Evidence for an arginine exporter encoded by yggA (argO) that is regulated by the LysR-type transcriptional regulator ArgP in *Escherichia coli*. *J Bacteriol* 186(11):3539–3546.
32. Hori H, et al. (2011) Inducible L-alanine exporter encoded by the novel gene ygaW (alaE) in *Escherichia coli*. *Appl Environ Microbiol* 77(12):4027–4034.
33. Vrljic M, Sahm H, Eggeling L (1996) A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. *Mol Microbiol* 22(5):815–826.
34. Kennerknecht N, et al. (2002) Export of L-isoleucine from *Corynebacterium glutamicum*: A two-gene-encoded member of a new translocator family. *J Bacteriol* 184(14):3947–3956.
35. Chakrabarti AC (1994) Permeability of membranes to amino acids and modified amino acids: Mechanisms involved in translocation. *Amino Acids* 6(3):213–229.
36. Faith JJ, McNulty NP, Rey FE, Gordon JI (2011) Predicting a human gut microbiota's response to diet in gnotobiotic mice. *Science* 333(6038):101–104.
37. Wang HH, et al. (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460(7257):894–898.
38. Yu D, et al. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA* 97(11):5978–5983.
39. Yilmaz P, et al. (2014) The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* 42(Database issue):D643–D648.
40. Letunic I, Bork P (2011) Interactive Tree of Life v2: Online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* 39(Web Server issue):W475–W478.

Supporting Information

Mee et al. 10.1073/pnas.1405641111

SI Materials and Methods

Dynamic Model of Three-Member Consortia. The three-member kinetic model was simulated using the Matlab environment. The model was initially seeded with 10^7 cells of each strain. Cooperativity coefficients were taken from pairwise cooperativity coefficients (c_{12}, c_{21}) provided in Dataset S2. Predicted fold growth was calculated by dividing time step t = 5,500 population values by the initial seeding value. The carrying capacity was set to 10^9 cells in all simulations. The β value was set to 1. Comparison between the predicted fold growth and observed fold growth is shown in Fig. S3*A*. In general, we find high predictive power in a significant fraction of the consortia (Pearson coefficient r = 0.51, P value = 5.2×10^{-25}). Differences between the predicted and observed fold growth highlights possible positive and negative epistatic interactions (Fig. S3*B*).

Dynamic Model of 14-Member Consortia. The 14-member kinetic model was simulated using the pairwise cooperativity coefficients (c_{12}, c_{21}) provided in Dataset S2. The population is seeded with equal amount of all 14 auxotrophs totaling 10^7 cells. The carrying capacity was set to 10^9 cells in all simulations. The β value was set to 1. The resulting steady-state population abundance was determined as the percentage of the total population dominated by each auxotroph type (Fig. S4*A*). For each 13-member dropout simulations, each dropout auxotroph was set to an initial seeding value of 0, and all other aspects of the simulation were unchanged. The difference in population abundance for each auxotroph in the 13-member groups compared with the 14-member group is shown in Fig. S4*B*.

Biosynthesis Predictions from the Joint Genome Institute Integrated Microbial Genomes Database. Amino acids biosynthetic predictions are pulled from the Integrated Microbial Genomes (IMG) database "phenotype function" system (1). This bioinformatic analysis pipeline first predicts the enzymatic capabilities of a given organism. Next, logical testing is used for prototrophy calls to ensure that for each reaction step in amino acid biosynthesis from a common precursor (e.g., pyruvate for leucine biosynthesis), there is a predicted enzyme that can catalyze the reaction. Similarly for auxotroph calls, an organism must be lacking enzymes catalyzing at least one of the steps in the biosynthesis pathway. Attempts have been made to take into account alternate biosynthesis pathways (e.g., methionine synthesis via homocysteine or methanethiol intermediates), although undoubtedly not every exception will be captured and these computational approaches will not have 100% accuracy. Nonetheless, this is a powerful tool to make general trend observations. For details on prediction rule, please see the IMG website.

1. Markowitz VM, et al. (2012) IMG: The Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res* 40(Database issue):D115–D122.

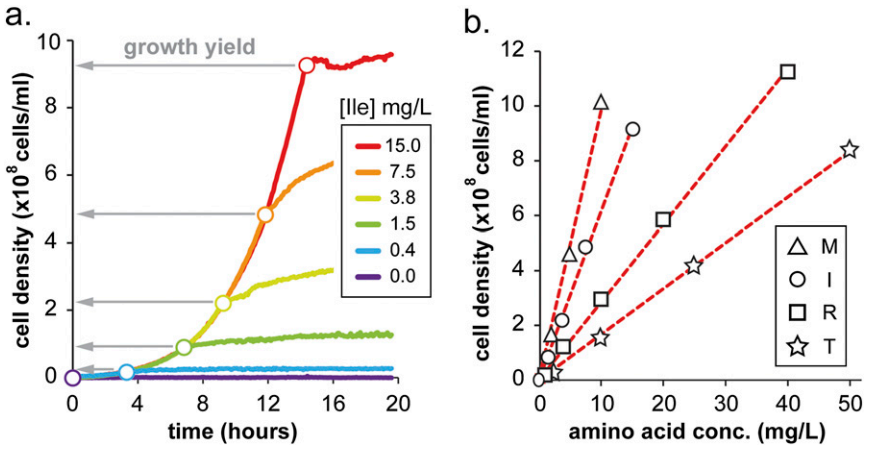

Fig. S1. Calculation of amino acid utilization during growth. (*A*) Example growth curves of *Escherichia coli* isoleucine auxotroph (Δ*ilvA*) in M9-glucose supplemented with varying initial concentrations of free L-isoleucine. The open circles denote the end of exponential phase resulting from supplemented amino acid exhaustion and marks final growth yield. (*B*) Linear relationship between amounts of supplemented amino acid and the observed growth yield for auxotrophs M, I, R, and T. Beyond the highest indicated amino acid concentrations, the yield response saturated due to other metabolites becoming growth limiting. The slope of the linear relationship between amino acid concentration and cell density is used to calculate the amino acids required per auxotrophic cell in M9 media presented in Dataset S1.

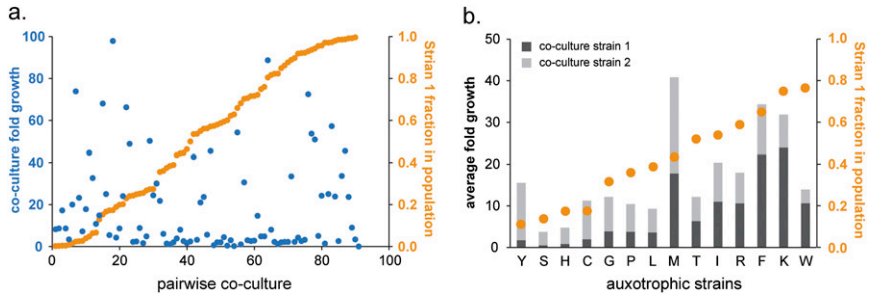

Fig. S2. Growth yield of 91 pairwise cocultures shows lack of relationship between relative population abundance and syntrophic growth. (*A*) Coculture growth of each of 91 pairwise two-member consortia. Left ordinate axis (in blue) shows whole coculture fold growth after 84 h for each of 91 consortia (data in blue dots). Right ordinate axis (in orange) shows fraction of strain 1 in population for each consortia (data in orange dots). Consortia are ranked from lowest to highest based fraction of population dominated by strain 1. The pairwise coculture index number indicated on the *x* axis refers to the coculture ID presented in Dataset S2. (*B*) Calculated average coculture growth for each auxotroph paired with all 13 other possible auxotrophs. The left ordinate axis (in black) shows whole coculture fold growth after 84 h for each auxotroph, which is designated strain 1. The color stacked bars indicates average fold growth contributed by strain 1 (in black) and strain 2 (in gray). The right ordinate axis (in orange) shows average fraction of strain 1 in population for each auxotroph group (data in orange range).

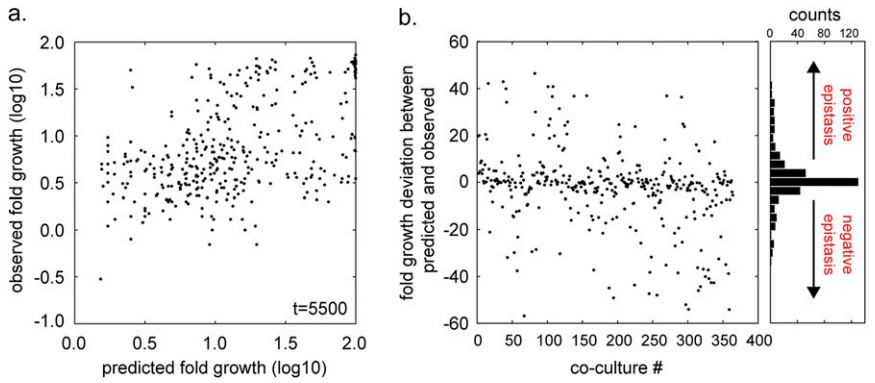

Fig. S3. Evaluation of dynamic growth model of three-member consortia. (*A*) Scatter plot comparing predicted fold growth of three-member consortia (in $\log_{10}$ units) versus observed fold growth of three-member consortia (in $\log_{10}$ units). Each black dot corresponds to each of the 364 possible unique consortia. (*B*) Scatter plot showing the difference between predicted and expected fold growth for all 364 consortia. Histogram on the *Right* shows the distribution across all consortia indicating general agreement between predictions and observations. Positive values indicate potential presence of positive epistasis, whereas negative values indicate potential negative epistasis.

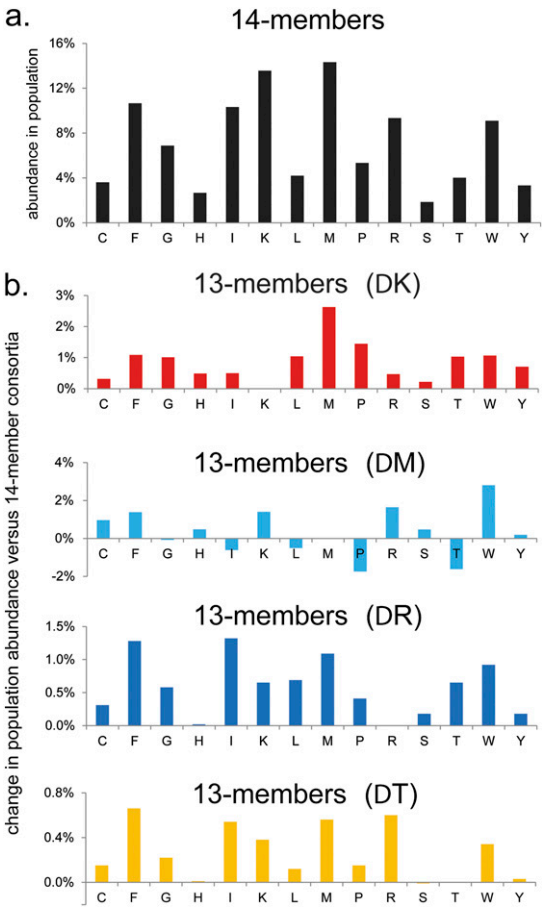

Fig. S4. Predictions of dynamic growth model of 14-member consortia. (*A*) Bar graph showing predicted steady-state population abundance of 14-member monoauxotrophs. Values are percentage of whole population. (*B*) Bar graph showing predicted change in steady state population abundance of four 13-member dropout consortia (no K, no M, no R, or no T) versus the 14-member consortium. Values are percentage of whole population.

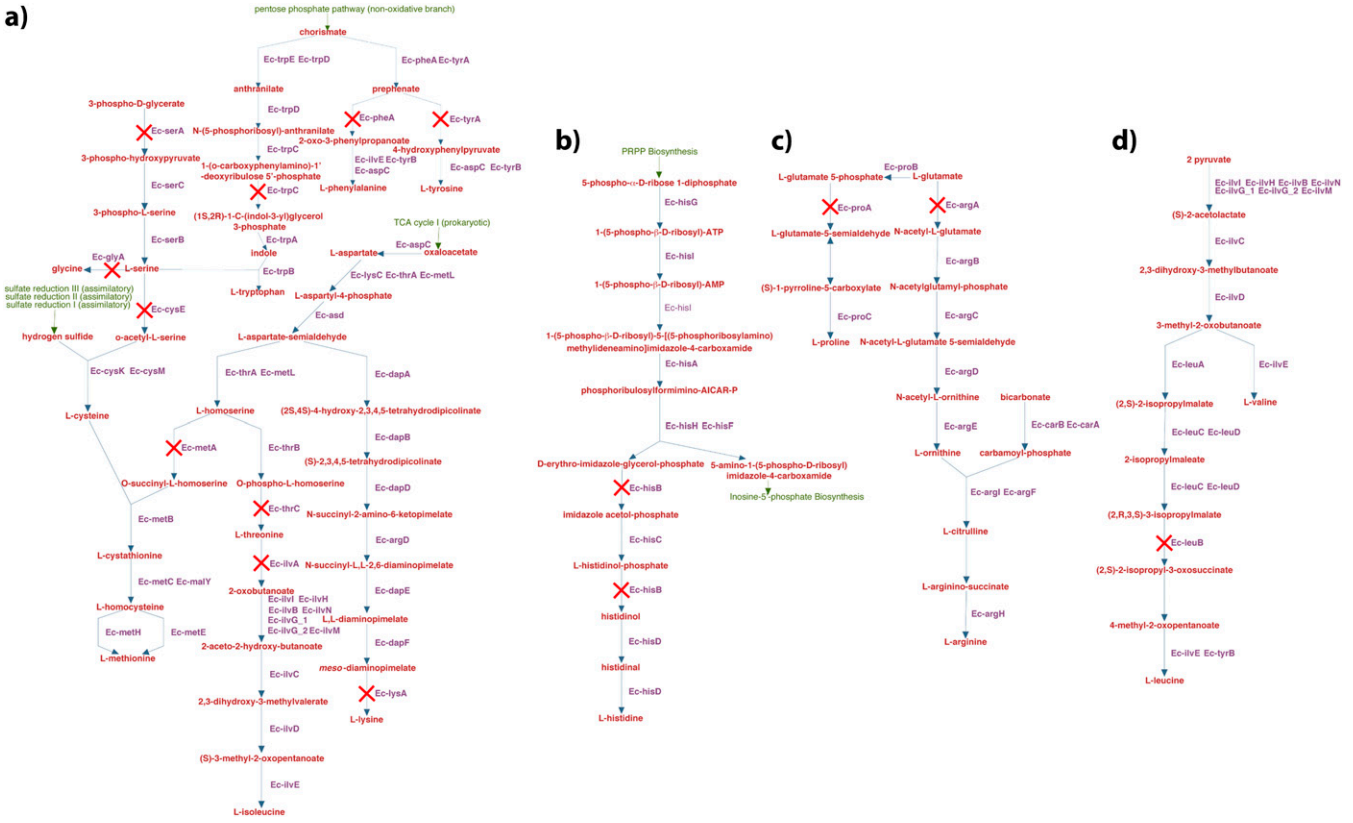

Fig. S5. Metabolic network diagrams. Biosynthesis reactions disrupted to generate auxotrophies are indicated with red crosses. Figures modified from Metacyc (1). Biosynthesis pathways for (*A*) C, F, G, I, K, M, S, T, W, Y; (*B*) H; (*C*) P, R; (*D*) L.

1. Caspi R, et al. (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* 42(Database issue):D459–D471.

Dataset S1. List of amino acid auxotrophs, the amount of supplemented amino acids needed to reproduce one cell, and the metabolic cost to biosynthesize each molecule of amino acid in units of phosphate bonds used

Dataset S1

Amino acid requirement per cell is calculated using the slope from Fig. S1*B* as follows: $(1/(\text{slope} \times \text{amino acid F.W. in g/mol} \times 10^6)) \times 6.022 \times 10^{23}$.

Dataset S2. Fold growth values of two-member and three-member consortia

Dataset S2

Growth values for two-member consortia are given for each strain as determined by quantitative PCR (*Materials and Methods*) and the determined cooperative coefficients are also shown. Total coculture growth values for each three-member consortia are given along with the three corresponding two-member subsets based on the pairwise data. Average cell concentrations for three or four replicates of the 14- and 13-member cultures are provided for the 7 d of passaging. Relative abundance for the 50-d passaging of the 14-member consortia is shown for the two replicates displayed in Fig. 4*C*.

Dataset S3. Detailed results from the 3- and 14-member dynamic modeling

Dataset S3

Fit between the three-member dynamic model and experimental results are listed for removal of certain groups of data points as defined by the zones in Fig. 3*B*. Sensitivity analysis for the initial cell concentration for the 14-member dynamic growth simulation is included. Sensitivity analysis 1 shows the result of stochastically seeding all strains with a normal distribution centered at 1e7 cells and SDs ranging from 10% to 100% of the mean as indicated. Similarly, sensitivity analysis 2 shows the results of introducing stochasticity (seeding density pulled from a normal distribution centered at 1e7 cells and SD of 50% of the mean) to each strain in the mixture independently. This is to assess whether any single strain has more of an effect than the others. The stability of the model demonstrated by these analyses suggests that any divergence with the experimental results is unlikely to be due to path dependence effects of the model.

Dataset S4. List of primers used in this study

Dataset S4