Report

Current Biology

A Single Community Dominates Structure and Function of a Mixture of Multiple Methanogenic Communities

Highlights

- A mix of communities has the same methane production as its best individual component
- The composition of the mixes resembles the composition of the best individual community
- The results suggest a modular co-selection of species in the mix
- The more communities are being mixed, the bigger the gas production of the mixes

Authors

Pawel Sierocinski, Kim Milferstedt, Florian Bayer, ..., Orkun S. Soyer, Jérôme Hamelin, Angus Buckling

Correspondence

p.sierocinski@exeter.ac.uk

In Brief

Coalescence of whole communities occurs frequently in the microbial world. Here Sierocinski et al. study coalescence of methanogenic communities. Both methane production and composition of the mixed communities resemble that of their best individual component, suggesting modular co-selection of species. This effect increases when more communities are used.

Sierocinski et al., 2017, Current Biology *27*, 1–6
November 6, 2017
https://doi.org/10.1016/j.cub.2017.09.056

CellPress

Please cite this article in press as: Sierocinski et al., A Single Community Dominates Structure and Function of a Mixture of Multiple Methanogenic Communities, Current Biology (2017), https://doi.org/10.1016/j.cub.2017.09.056

Current Biology
Report

CellPress

A Single Community Dominates Structure and Function of a Mixture of Multiple Methanogenic Communities

Pawel Sierocinski,[1,6,*] Kim Milferstedt,[2] Florian Bayer,[1] Tobias Großkopf,[3] Mark Alston,[4] Sarah Bastkowski,[4] David Swarbreck,[4] Phil J. Hobbs,[5] Orkun S. Soyer,[3] Jérôme Hamelin,[2] and Angus Buckling[1]
[1]Biosciences, University of Exeter, Penryn, Cornwall TR10 9FE, UK
[2]Laboratoire de Biotechnologie de l'Environnement (LBE), Institut National de la Recherche Agronomique (IRNA), 11100 Narbonne, France
[3]School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK
[4]Earlham Institute, Norwich Research Park, Norwich NR4 7UH, UK
[5]Anaerobic Analytics, Okehampton EX20 1AS, UK
[6]Lead Contact
*Correspondence: p.sierocinski@exeter.ac.uk
https://doi.org/10.1016/j.cub.2017.09.056

SUMMARY

The ecology of microbes frequently involves the mixing of entire communities (community coalescence), for example, flooding events, host excretion, and soil tillage [1, 2], yet the consequences of this process for community structure and function are poorly understood [3–7]. Recent theory suggests that a community, due to coevolution between constituent species, may act as a partially cohesive unit [8–11], resulting in one community dominating after community coalescence. This dominant community is predicted to be the one that uses resources most efficiently when grown in isolation [11]. We experimentally tested these predictions using methanogenic communities, for which efficient resource use, quantified by methane production, requires coevolved cross-feeding interactions between species [12]. After propagation in laboratory-scale anaerobic digesters, community composition (determined from 16S rRNA sequencing) and methane production of mixtures of communities closely resembled that of the single most productive community grown in isolation. Analysis of each community's contribution toward the final mixture suggests that certain combinations of taxa within a community might be co-selected as a result of coevolved interactions. As a corollary of these findings, we also show that methane production increased with the number of inoculated communities. These findings are relevant to the understanding of the ecological dynamics of natural microbial communities, as well as demonstrating a simple method of predictably enhancing microbial community function in biotechnology, health, and agriculture [13].

RESULTS AND DISCUSSION

We wanted to determine whether coalesced methanogenic communities were dominated by the community that used resources most efficiently in isolation. We used methanogenic communities primarily because methane production is a useful proxy for the ability of an anaerobic community to fully exploit available resources: methanogenesis results from the conversion of H_2, CO_2, and short-chain fatty acids produced by hydrolysis and fermentation of more complex organic material, and it is often the only thermodynamically feasible way of actively removing inhibitory end metabolites [12]. Moreover, methanogenic communities are characterized by complex cross-feeding interactions [12, 14, 15]; hence, the role of community cohesion in shaping community performance is likely to be particularly important [9]. To provide insight into the temporal dynamics of compositional and functional change after community mixing, we first measured the methane production and composition of two methanogenic communities derived from industrial anaerobic digesters (ADs) (Table 1) grown in isolation or as a mixture in laboratory-scale ADs. Both the individual communities and mixes were grown in four replicates. To remove any potentially confounding effects caused by differences in starting density of tested communities, we standardized microbial density based on qPCR-estimated counts of 16S rRNA gene copies. We found that the methane production of the mixed community was initially intermediate between the two individual communities, but after 5 weeks propagation, it started to produce gas at a rate indistinguishable from that of the more productive of the individual communities (Figure 1A). We examined both the starting-point and the endpoint composition of the single and mixed communities by Illumina sequencing 16 s rRNA gene amplicon libraries. Consistent with the phenotypic data, the composition of the mixture was much more similar to the better- than to the worse-performing community at the endpoint (Figure 1B). This was despite the single endpoint communities changing considerably from their ancestral composition over the 5 weeks (Figure 1B).

We next determined whether a single community dominated when multiple communities were mixed. To this end, we

Table 1. List of Individual Communities Used in This Analysis and Their Source

Sample Name	Feed and/or Type	Temperature	Used in Experiment
P01	silage and food waste AD	44°C–42.5°C	1, 2, 3
P02	silage and food waste AD	44°C–42.5°C	2, 3
P03	maize, cow slurry, and chicken manure AD	45°C	3
P04	maize, cow slurry, and chicken manure AD	45°C	2, 3
P05	sewage sludge AD	36°C	1, 2, 3
P06	raw sewage	ambient	2, 3
P08	thickened sewage sludge	ambient	2, 3
P09	sewage based AD	36°C	2, 3
P10	food waste AD	36°C	2, 3
P11	cow slurry	ambient	3
P12	silage, slurry, and manure pre-digestate	ambient	3
P13	silage, slurry, and manure AD	40°C	2, 3
P15	food waste AD	36°C	2

All anaerobic digester (AD) communities were derived from industrial ADs in the southwest of England. Specific locations cannot be provided because of commercial sensitivity. Note that experiment numbers correspond with figure numbers.

propagated ten single communities (from either industrial ADs or sewage or agricultural waste AD feedstocks, with each replicated three times), and ten replicates of a mixture of all ten communities (Table 1). The results were consistent with those from the two-community mixture. First, methane production in mixtures of ten communities was higher than the average of the individual communities. However, methane production of the mixtures did not differ from the best-performing single community, P13, (Figure 2A), which, like each of the single communities used, was a constituent of all of the mixtures. Second, the community composition of mixtures (which varied very little between replicates, presumably because they all had the same ten community-starting inocula) most closely resembled the best-performing community, P13 (Figure 2B). More generally, the more compositionally similar an individual community was to the replicated ten-community mixtures, the greater the gas production of the community when grown in isolation (Figure S1). Other community characteristics that positively correlated with methane production were bacterial cell densities and within-community (alpha) diversity, but not methanogen density (Figure S2). In summary, the results demonstrate that the community most efficient at using resources (which in these experiments was also the most diverse) dominates when multiple communities are mixed together, thus enhancing mixed-community productivity beyond the average of the component communities.

We next explored the ecological mechanism(s) underpinning the observed dominance by the community that produced the most methane. One explanation is that multiple taxa from the same community act as semi-cohesive units and are selected together. This might arise as a result of coevolved mutualistic (or unidirectional) cross-feeding interactions, notably between methanogenic Archaea and hydrogen and/or acetate producers, where each organism both provides essential resources and removes damaging waste products for the other [12, 15, 16]. Moreover, coevolved resource partitioning can result in taxa being selected together, because species are expected to coevolve to minimize competition with co-occurring taxa [17–19]. Note that the selection of multiple taxa together in these contexts does not require any form of group selection [11, 20], but simply selection of particular individuals from a key taxon whose presence provides an advantage for individuals from taxa they have coevolved with. This process can be described as ecological co-selection, equivalent to genetic co-selection, where a gene can hitchhike to high frequency purely as consequence of being linked to genes under positive selection [21].

An alternative explanation is that coevolved interactions within individual communities are relatively unimportant, and the dominant community simply contains more competitive taxa (for any functional task and/or interaction) than other communities. This does not imply that coevolved cross-feeding interactions are unimportant for methanogenic communities, but that these co-evolved interactions are no more specific for taxa isolated from within a community than for taxa isolated from different communities. In other words, functionally equivalent taxa are

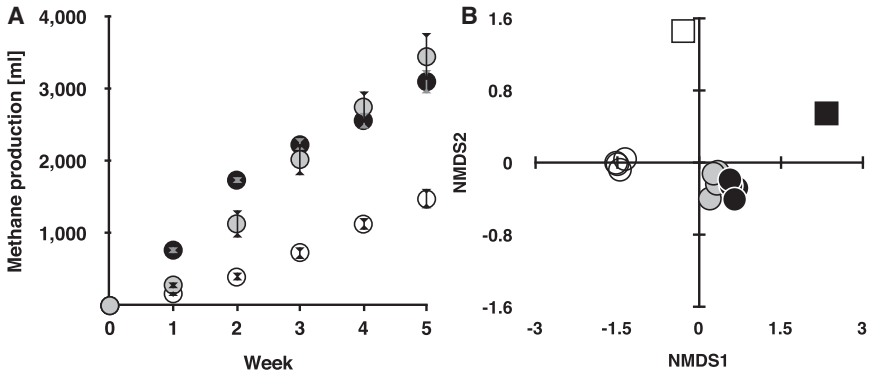

Figure 1. Temporal Dynamics of Methane Production and Composition When Two Communities Are Mixed

(A) Cumulative methane production in milliliters (±SEM) over time of community P01 (white circles), community P05 (black circles), and their mixes (gray circles). Cumulative methane production differed between treatments (ANOVA: $F_{2,9} = 23.2$, $p < 0.001$) but did not differ between the mixed community and P05 (Tukey-Kramer honest significant difference [HSD]: $p = 0.5$). P01 was lower than both other treatments (Tukey-Kramer HSD: $p < 0.001$ in both cases).

(B) Non-metric multidimensional scaling (NMDS) plot of unweighted UniFrac of communities P01 (white), P05 (black), and their mixes (gray). Ancestral samples are represented by squares, and samples from the endpoint of the experiment are represented by circles. At the endpoint, P05 was compositionally more similar to the mixtures than P01, based on both unweighted (t tests of mean distance to each mixture for each replicate single community: $t_6 = 8.3$, $p < 0.001$) and weighted ($t_6 = 2.3$, $p = 0.03$) UniFrac distances.

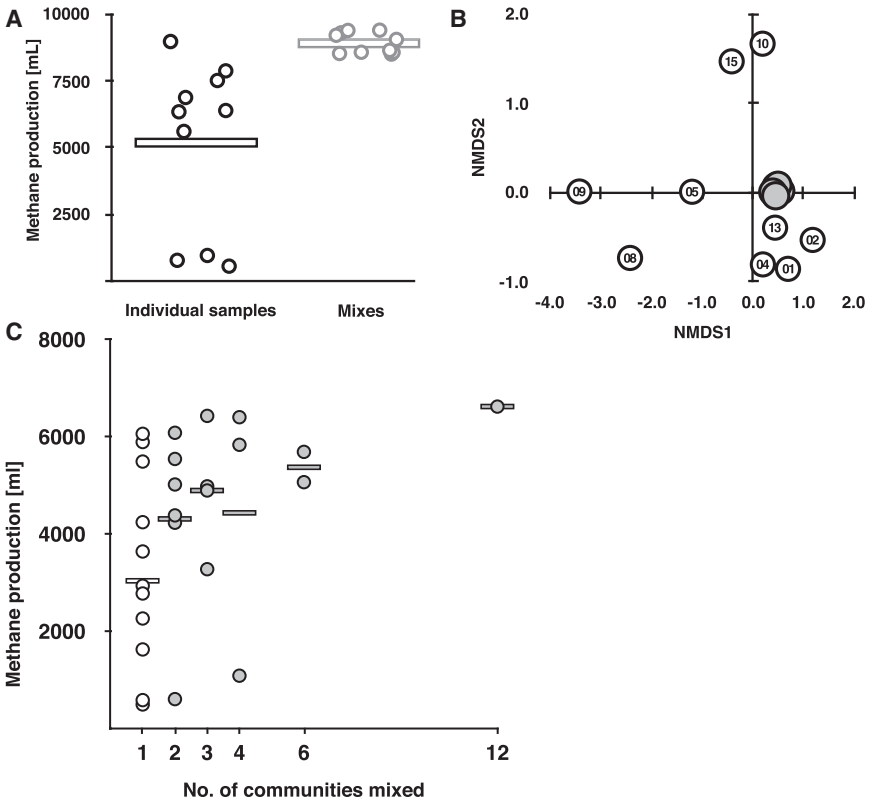

Figure 2. Methane Production and Community Composition When Multiple Communities Are Mixed

(A) Total methane production of mixed (gray) and individual (white) communities, with mean values shown as horizontal lines. Mean total methane production was greater for mixtures than for individual communities (t test: $p < 0.001$ in nine cases), except when measured against community P13 (the best performer).

(B) NMDS plot of unweighted UniFrac of ten mixtures (gray) and nine individual communities (white). Numbers in circles refer to individual community identifiers (Table 1). Community P13 was significantly closer in composition to the ten mixed communities than any other community (weighted and unweighted UniFrac distances; paired t tests: $p < 0.001$ in all cases). There was also a significant link between the community composition and the difference in gas production between the communities (see Figure S1). Note that DNA yield from community P06, which had the lowest gas production of all communities, was insufficient for sequencing; therefore, it is excluded from this and the following graphs.

(C) Individual communities (white circles) and their average methane production (white line), as well as mixes of communities (gray circles) and their averages (gray line). There was a monotonic increase in methane production with number of communities used (regression: $F_{1,26} = 5.4$, $p = 0.03$). For community composition of the mixes, see Tables 1 and S1.

interchangeable between communities. These different scenarios, selection for the best individual taxa and co-selection, are two extremes of a continuum. The distinction is important because dominance by a single community is necessarily a more likely consequence of community coalescence when co-selection operates. Figures 3A–3C provide an illustration of the two extreme scenarios, no co-selection and co-selection of the entire community, and an intermediate case in which there are two groups of interacting taxa, or modules, and co-selection occurs within each.

The most direct way to demonstrate a role of co-selection would be to show that the outcome of competition between single taxa from different communities does not predict the outcome of competition at the community level [11]. Unfortunately, this is not feasible for such complex communities, in which many taxa are very difficult to grow in isolation. However, there are other testable predictions associated with the operation of co-selection or otherwise. If the success of an individual taxon is independent of whether they are in the presence of taxa from the same community (i.e., co-selection does not occur), communities that use resources most efficiently and hence achieve the highest biomass per unit of time (productivity) will contain the highest number of the best-performing taxa. It then follows that there will be a positive relationship between the productivity of a community and the proportion of taxa it contributes to the mixture (Figure 3A). If instead taxa are co-selected as modules, the correlation between individual community contribution and productivity is likely to break down. This is best illustrated by the extreme scenario whereby all taxa within a mixed community are co-selected from a single community: the mixture will be entirely dominated by a single constituent community, and hence the contribution of all other communities will be independent of their individual productivity (i.e., they will contribute null to the mixture's composition, even though they have non-zero productivity individually; Figure 3B). The intermediate scenario, where co-selection occurs within two independent modules, can also break down this correlation if one module contributes much more to community productivity than the other (Figure 3C).

To determine whether co-selection contributed to our findings, we first estimated the contribution of each community to the ten-community mixtures using a non-negative least-squares (NNLS) approach. The community that had the most similar composition to the mixtures (and produced the most methane) contributed an estimated 40% of its taxa to the mixtures, with only two other communities contributing more than 10% of their taxa to the mixtures (Figure 3D). We then correlated the contribution each community made to the mixtures with two measures of community productivity: methane production and cell densities (based on 16 s rRNA gene copy number), which themselves were positively correlated (Figure S2A). We found no suggestion of a positive correlation between either measure of productivity and contribution to the community (Figures 3D and 3E). These results suggest that co-selection of taxa played an important role in dominance by the community that produced the most methane.

That community coalescence results in the most productive individual community dominating the mixed community has

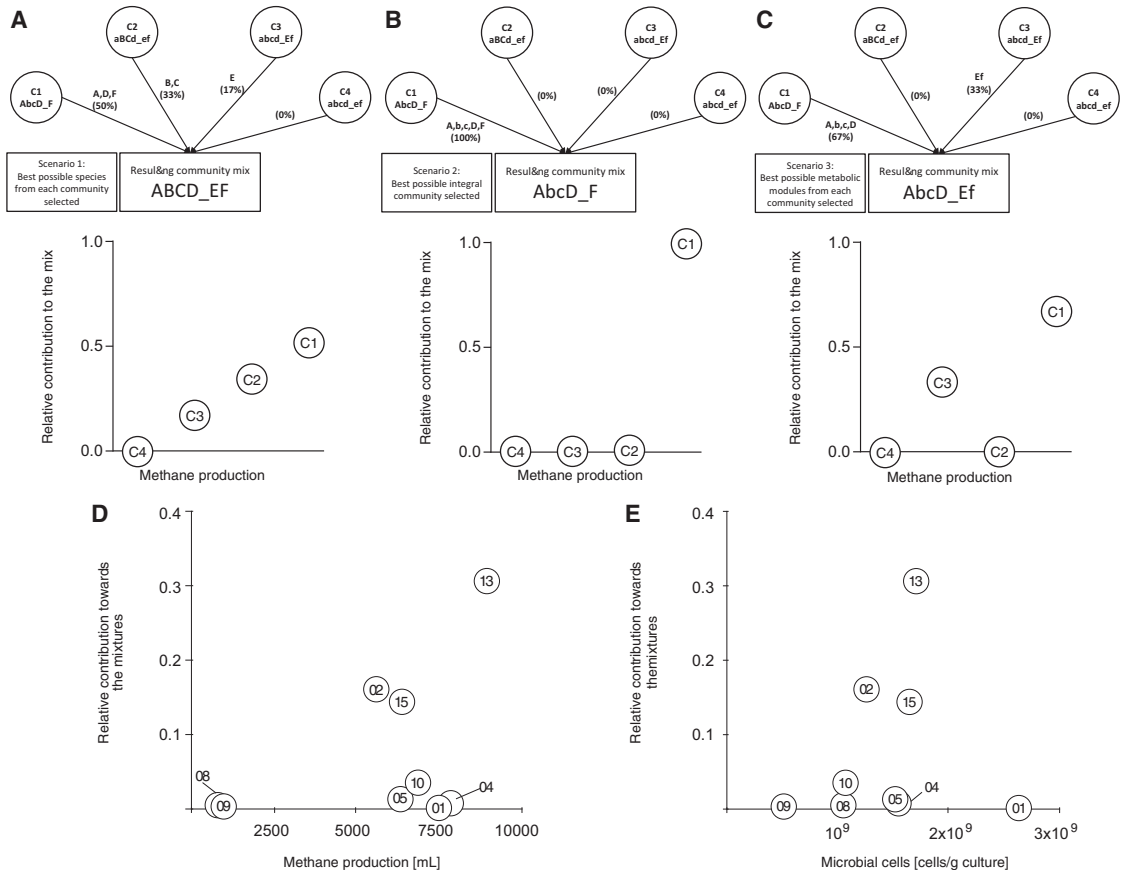

Figure 3. The Role of Co-selection in Explaining Dominance by a Single Community

(A–C) The top panels illustrate three hypothetical scenarios describing how communities contribute to a mixture of communities, and the bottom panels show the expected relationships between a community's contribution and its methane production. The letters within the top panels indicate taxa that drive two biochemical processes (abcd and ef); capitalized letters are the best representatives of a taxon among all of the communities. (A) No co-selection. (B) Co-selection of all taxa within a community. (C) Co-selection of taxa within two independent modules.

(D) Mean estimated relative contribution of each individual community (numbered) toward the ten coalesced communities calculated using the NNLS method, plotted against mean cumulative methane production for each community; there is no significant relationship (regression: $F_{1,7} = 1.7$, $p > 0.2$).

(E) As in (D), but with relative contribution plotted against number of bacterial and archaeal cells calculated based on the 16S rRNA gene copy number (regression: $F_{1,7} = 1.7$, $p > 0.5$). Note that the relative contribution is not a fractional contribution because some operational taxonomic units (OTUs) present in the mixture were not detected in the constituent communities. This is presumably because they only reached detected frequencies in the mixture, but we can't rule out that the community that we failed to get sufficient reads from contributed to the composition of the mixtures. Mind that the cell densities of Archaea and bacteria do not significantly correlate with the gas production (see Figure S2).

direct implications for biotechnological uses of microbial communities. Given that the best-performing community in isolation largely determined both the composition and performance of mixtures of communities, methane production should increase with increasing number of communities in a mixture. We therefore inoculated laboratory-scale anaerobic digesters with 1, 2, 3, 4, 6, or 12 communities, ensuring that each of the 12 starting communities was used an equal amount of times at each diversity level [22] (see Table S1). Cumulative methane production over a 5-week period increased with increasing number of communities used as an inoculum (Figure 2C). The positive correlation between community function and the number of inoculating communities is analogous to the commonly observed finding that community productivity increases with increasing species diversity [23]. In this case, the mechanism underlying this positive relationship between the number of communities and productivity is a "sampling effect": inoculating more communities increases the chance that the best-performing community will be present in the mix [24]. However, given that domination of mixtures by one community was not complete (Figures 3D and 3E), it is possible that mixing communities could increase performance beyond that of the maximum of single communities in some circumstances (transgressive over-yielding [25]).

Here, we have shown that coalescence of microbial communities results in dominance of a single community, the identity of which can be predicted from its efficiency of resource use in isolation. This dominance is likely to be driven in part by co-selection of interacting taxa within coevolved communities, which is likely to greatly increase the magnitude of dominance after mixing [11]. It is unclear whether such effects would be apparent for aerobic communities, in which cross-feeding interactions are less important for efficient resource use [26],

although studies to date [4] suggest that asymmetric outcomes, although less extreme, may be common. Our study has also identified a simple method to significantly improve methane yield during anaerobic digestion: inoculate digesters with a broad range of microbial communities, and the best-performing community will dominate. However, further work under a range of conditions is clearly required to determine the generality of these findings. Given that resource use efficiency is often a desirable property of microbial communities, this approach could be applied to a range of biotechnological processes driven by microbial communities, as well as to manipulate microbiomes in clinical and agricultural contexts [13].

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Methanogenic communities
- METHOD DETAILS
  - Cultivation details
  - Experiment structure
  - Measuring methane content of Biogas
  - DNA extraction and qPCR quantification
  - Amplicon library construction and sequencing
  - Analyses of sequenced samples
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures, one table, and supplemental text and can be found with this article online at https://doi.org/10.1016/j.cub.2017.09.056.

AUTHOR CONTRIBUTIONS

Methodology: P.S., K.M., F.B., O.S.S., P.J.H., J.H., and A.B.; Formal Analysis P.S., S.B., M.A., and A.B.; Investigation: P.S. and F.B.; Writing – Original Draft: P.S. and A.B.; Writing – Review and Editing: all authors; Visualization: P.S. and A.B.; Supervision: A.B.

ACKNOWLEDGMENTS

The work was funded by BBSRC project BB/K003240/1, the Royal Society, the AXA Research Fund, and the NERC. K.M. and J.H. were funded through the project ENIGME from the INRA Metaprogramme MEM (Meta-omics and Microbial Ecosystems). K.M. was additionally funded through an Institut Carnot 3BCAR international travel grant. We would also like to thank Edyta Wocial for her help with the figures.

Received: April 3, 2017
Revised: June 6, 2017
Accepted: September 26, 2017
Published: October 26, 2017

REFERENCES

1. Rillig, M.C., Antonovics, J., Caruso, T., Lehmann, A., Powell, J.R., Veresoglou, S.D., and Verbruggen, E. (2015). Interchange of entire communities: microbial community coalescence. Trends Ecol. Evol. *30*, 470–476.
2. Rillig, M.C., Lehmann, A., Aguilar-Trigueros, C.A., Antonovics, J., Caruso, T., Hempel, S., Lehmann, J., Valyi, K., Verbruggen, E., Veresoglou, S.D., et al. (2016). Soil microbes and community coalescence. Pedobiologia (Jena) *59*, 37–40.
3. Hausmann, N.T., and Hawkes, C.V. (2009). Plant neighborhood control of arbuscular mycorrhizal community composition. New Phytol. *183*, 1188–1200.
4. Livingston, G., Jiang, Y., Fox, J.W., and Leibold, M.A. (2013). The dynamics of community assembly under sudden mixing in experimental microcosms. Ecology *94*, 2898–2906.
5. Souffreau, C., Pecceu, B., Denis, C., Rummens, K., and De Meester, L. (2014). An experimental analysis of species sorting and mass effects in freshwater bacterioplankton. Freshw. Biol. *59*, 2081–2095.
6. Adams, H.E., Crump, B.C., and Kling, G.W. (2014). Metacommunity dynamics of bacteria in an arctic lake: the impact of species sorting and mass effects on bacterial production and biogeography. Front. Microbiol. *5*, 82.
7. Calderón, K., Spor, A., Breuil, M.-C., Bru, D., Bizouard, F., Violle, C., Barnard, R.L., and Philippot, L. (2017). Effectiveness of ecological rescue for altered soil microbial communities and functions. ISME J. *11*, 272–283.
8. Gilpin, M. (1994). Community-level competition: asymmetrical dominance. Proc. Natl. Acad. Sci. USA *91*, 3252–3254.
9. Toquenaga, Y. (1997). Historicity of a simple competition model. J. Theor. Biol. *187*, 175–181.
10. Wright, C.K. (2008). Ecological community integration increases with added trophic complexity. Ecol. Complex. *5*, 140–145.
11. Tikhonov, M. (2016). Community-level cohesion without cooperation. eLife *5*, e15747.
12. Schink, B. (1997). Energetics of syntrophic cooperation in methanogenic degradation. Microbiol. Mol. Biol. Rev. *61*, 262–280.
13. Rillig, M.C., Tsang, A., and Roy, J. (2016). Microbial community coalescence for microbiome engineering. Front. Microbiol. *7*, 1967.
14. Hillesland, K.L., and Stahl, D.A. (2010). Rapid evolution of stability and productivity at the origin of a microbial mutualism. Proc. Natl. Acad. Sci. USA *107*, 2124–2129.
15. Embree, M., Liu, J.K., Al-Bassam, M.M., and Zengler, K. (2015). Networks of energetic and metabolic interactions define dynamics in microbial communities. Proc. Natl. Acad. Sci. USA *112*, 15450–15455.
16. Großkopf, T., and Soyer, O.S. (2016). Microbial diversity arising from thermodynamic constraints. ISME J. *10*, 2725–2733.
17. Schluter, D. (2000). The Ecology of Adaptive Radiation (Oxford University Press).
18. Roughgarden, J. (1976). Resource partitioning among competing species–a coevolutionary approach. Theor. Popul. Biol. *9*, 388–424.
19. MacArthur, R. (1970). Species packing and competitive equilibrium for many species. Theor. Popul. Biol. *1*, 1–11.
20. Gardner, A., and Grafen, A. (2009). Capturing the superorganism: a formal theory of group adaptation. J. Evol. Biol. *22*, 659–671.
21. Baker-Austin, C., Wright, M.S., Stepanauskas, R., and McArthur, J.V. (2006). Co-selection of antibiotic and metal resistance. Trends Microbiol. *14*, 176–182.
22. Hodgson, D.J.D., Rainey, P.B., and Buckling, A. (2002). Mechanisms linking diversity, productivity and invasibility in experimental bacterial communities. Proc. Biol. Sci. *269*, 2277–2283.
23. Tilman, D., Lehman, C.L., and Thomson, K.T. (1997). Plant diversity and ecosystem productivity: theoretical considerations. Proc. Natl. Acad. Sci. USA *94*, 1857–1861.
24. Tilman, D. (1999). The ecological consequences of changes in biodiversity: a search for general principles. Ecology *80*, 1455–1474.
25. Harper, D. (1977). The Population Biology of Plants (Academic Press).

26. Morris, B.E.L., Henneberger, R., Huber, H., and Moissl-Eichinger, C. (2013). Microbial syntrophy: interaction for the common good. FEMS Microbiol. Rev. *37*, 384–406.
27. Einen, J., Thorseth, I.H., and Ovreås, L. (2008). Enumeration of Archaea and Bacteria in seafloor basalt using real-time quantitative PCR and fluorescence microscopy. FEMS Microbiol. Lett. *282*, 182–187.
28. Ruijter, J.M., Ramakers, C., Hoogaars, W.M.H., Karlen, Y., Bakker, O., van den Hoff, M.J., and Moorman, A.F. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res. *37*, e45.
29. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.
30. Brankatschk, R., Fischer, T., Veste, M., and Zeyer, J. (2013). Succession of N cycling processes in biological soil crusts on a Central European inland dune. FEMS Microbiol. Ecol. *83*, 149–160.
31. Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl. Environ. Microbiol. *79*, 5112–5120.
32. Eren, A.M., Maignien, L., Sul, W.J., Murphy, L.G., Grim, S.L., Morrison, H.G., and Sogin, M.L. (2013). Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. Methods Ecol. Evol. *4*, 1111–1119.
33. Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460–2461.
34. Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics *27*, 2194–2200.
35. McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. *6*, 610–618.
36. Mullen, K., and Van Stokkum, I. (2007). The Lawson-Hanson algorithm for non-negative least squares (NNLS). https://cran.r-project.org/package=nnls.
37. Soetaert, K., Meersche, K.V.D., and Oevelen, D. V (2009). Package limSolve, solving linear inverse models in R. https://cran.r-project.org/package=limSolve.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Brilliant III Ultra-Fast SYBR Green QPCR Master Mix	Agilent Technologies	600882-51
meat extract	Sigma-Aldrich	70164-500G
Xylose	Sigma-Aldrich	W360100-1KG
Cellulose	Sigma-Aldrich	C6288
Starch	Sigma-Aldrich	33615-1KG
Glucose	Sigma-Aldrich	G8270-1KG
Critical Commercial Assays		
QIAamp DNA Stool Mini Kit (QIAGEN)	QIAGEN	ID: 51504
FastDNA SPIN Kit for Soil	MP Biomedicals,	116560200
PowerLyzer PowerSoil DNA Isolation Kit	MO BIO Laboratories	12855-100
Deposited Data		
Sequencing data Experiment 1	European Nucleotide Archive	ENA: PREJEB21187; https://www.ebi.ac.uk/ena/data/view/PRJEB21193
Sequencing data Experiment 2	European Nucleotide Archive	ENA: PREJEB21187; https://www.ebi.ac.uk/ena/data/view/PRJEB21187
Experimental Models: Organisms/Strains		
Community P01	Silage and Foodwaste Anaerobic Digester (AD)	This paper
Community P02	Silage + Food waste AD	This paper
Community P03	Maize/Cow Slurry/Chicken Manure AD	This paper
Community P04	Maize/Cow Slurry/Chicken Manure AD	This paper
Community P05	Sewage Sludge AD	This paper
Community P06	Raw Sewage	This paper
Community P08	Thickened Sewage Sludge	This paper
Community P09	Sewage Based AD	This paper
Community P10	Food Waste AD	This paper
Community P11	Cow Slurry	This paper
Community P12	Silage, Slurry and Manure Pre-Digestate	This paper
Community P13	Silage, Slurry and Manure AD	This paper
Community P15	Food waste AD	This paper
Oligonucleotides		
338f - ACT CCT ACG GGA GGC AGC AG	[27]	N/A
518r - ATT ACC GCG GCT GCT GG	[27]	N/A
931f - AGG AAT TGG CGG GGG AGC A	[27]	N/A
m1100r - BGG GTC TCG CTC GTT RCC	[27]	N/A
Software and Algorithms		
StepOne Software v.2.3	life technologies	https://www.thermofisher.com/uk/en/home/technical-resources/software-downloads/StepOne-and-StepOnePlus-Real-Time-PCR-System.html
LinRegPCR version 2016.0	[28]	linregpcr.nl

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
R version 3.4.0	R Core Team (2013)	Mac: https://cran.r-project.org/bin/macosx/; PC: https://cran.r-project.org/bin/windows/base/old/
macQIIME	[29]	http://www.wernerlab.org/software/macqiime/download
Other		
NNLS Method for assessing community contribution in a mix	This paper	Method S1

CONTACT FOR REAGENT AND RESOURCE SHARING

The authors are happy to share any further resources linked to the research involved with qualified third parties. Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Pawel Sierocinski (p.sierocinski@exeter.ac.uk).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Methanogenic communities

The communities used were collected from anaerobic digesters (AD plants: communities P01,P02, P03, P04, P05, P09, P10, P13 and P15) and communities present in nature used to seed the AD plants (communities P06, P08, P11, P12, see the details in Table 1). All the communities have been collected in the South West area of United Kingdom from operating Anaerobic Digesters and the seeding communities they use for the reactors. The communities were operating at temperatures between ambient and 45°C in their places of origin. Communities were stored at 4°C prior to use.

METHOD DETAILS

Cultivation details

For all experiments, communities were grown in 500 mL bottles (600 mL total volume with headspace; Duran) using the commercially available Automated Methane Potential Test System (AMPTS, Bioprocess Control Sweden AB) to measure CO_2-stripped biogas production (referred to as methane in this paper). Samples were fed weekly with 25 mL of defined medium in a fed-batch mode using a defined medium (see below for media composition).

The communities used in experiment 1 were equalized in terms of bacterial cells per gram of sample before inoculation using M9 salts to dilute them to the community with the lowest cell density, based on qPCR enumeration of 16S rRNA gene copies. For experiments 2 and 3, starting 16S rRNA copy number was determined (but not equalized between communities) and did not correlate with methane production. The fermenters were inoculated with 275 g of sample and fed with 25 mL of defined medium: meat extract 111.1 gl^{-1}, cellulose 24.9 gl^{-1}, starch 9.8 gl^{-1} glucose 0.89 gl^{-1}, xylose 3.55 gl^{-1} (carbon to nitrogen ratio of 15:1) every week, starting with t_0. Before the start of the fermentation, 0.3 mL of 1000x Trace Metal stock (1 gl^{-1} $FeCl_2 \cdot 4H_2O$, 0.5 gl^{-1} $MnCl_2 \cdot 4H_2O$, 0.3 gl^{-1} $CoCl_2 \cdot 4H_2O$, 0.2 gl^{-1} $ZnCl_2$, 0.1 gl^{-1} $NiSO_4 \cdot 6H_2O$, 0.05 gl^{-1} $Na_2MoO_4 \cdot 4H_2O$, 0.02 gl^{-1} H_3BO_3, 0.008 gl^{-1} $Na_2WO_4 \cdot 2H_2O$, 0.006 gl^{-1} $Na_2SeO_3 \cdot 5H_2O$, 0.002 gl^{-1} $CuCl_2 \cdot 2H_2O$) was added to each fermenter.

Experiment structure

In Experiment 1 (results shown in Figure 1) we cultivated community P01 (four replicates) and community P05 (four replicates) and a 1:1 mix of the two. It ran for 5 weeks before samples were harvested for sequencing (see below). The initial community was sequenced at the same time. In Experiment 2 (results shown in Figure 2), we cultivated 10 individual communities (listed in Table 1), in triplicate, and 10 mixes of all 10 communities mixed in equal volumes, at the same total volume as the single communities. After 6 weeks samples were harvested for sequencing. In Experiment 3 (results shown in Figure 2C) we used 12 communities (detailed in Table 1). They were grown in isolation as well as pseudo-randomly combined to create a gradient of number of starting communities, with each community used only once for each number of communities. This resulted in 12 single communities, 6 pairs, 4 triplets, 3 quadruplets, 2 mixes of 6 and one mix of 12 communities. Specific details of mixing can be seen in Table S1. The cultures were propagated for 5 weeks.

Measuring methane content of Biogas

All resulting lab-scale reactors inoculated with the samples were run at 37°C using the Automatic Methane Potential Test System (AMPTS). The AMPTS is a setup of 15 simple fermenters using a 0.5L lab bottle as the vessel with its own stirring system provided

with a butyl rubber stopper and sampling ports. It is connected to an online gas measuring system to allow continuous gas measurements. The AMPTS system measures the volume of biogas produced following stripping of CO_2 (by passing the gas through 50 mL of 3M NaOH solution) from the produced gas. To reproduce our results, however, there is no need for a sophisticated setup, some pilot experiments yielding similar results in terms of gas production were conducted using anaerobic serum flasks. We confirmed that the measured biogas was > 95% methane using Gas Chromatography with Flame Ionization Detection optimized for methane detection.

DNA extraction and qPCR quantification

DNA for 16S rRNA gene amplicon sequencing was extracted using QIAamp DNA Stool Mini Kit (QIAGEN) or FastDNA SPIN Kit for Soil (MP), depending on the experiment. Note that DNA extraction for mixed community P06 from experiment 2 failed. The DNA for qPCR was extracted with the QIAamp DNA Stool Mini Kit (QIAGEN), protocol for pathogen detection with the 95°C incubation step and the Powerlyzer Powersoil DNA KIT (MOBIO). DNA from *Acinetobacter baylyi*, *Pseudomonas fluorescens* SBW25 for Bacteria and from *Halobacterium salinarum* DSM 669 for Archaea was used as standards. The primers [27] used to quantify Bacteria were 16S rRNA 338f - ACT CCT ACG GGA GGC AGC AG, 518r - ATT ACC GCG GCT GCT GG for Archaea: 931f - AGG AAT TGG CGG GGG AGC A, m1100r - BGG GTC TCG CTC GTT RCC. The reagents used were: 1x Brilliant III Ultra-Fast SYBR Green QPCR Master Mix; 150nM 338f and 300nM 518r or 300nM 931f and 300 nM m1100r; ROX 300nM; and BSA 100 ng/μL final concentration. All samples were run in triplicate on a StepOnePlus (Applied Biosystems) qPCR machine using a program with 3 min 95°C initial denaturation followed by 40 cycles of 5 s at 95°C and 10 s at 60°C, followed by a melting curve 95°C for 15 s; 60°C for 1 min ramping up to 95°C in steps of 0.3°C for 15 s each. The melting curve analysis and the confirmation of the negative controls was done using StepOne Software v.2.3 (life technologies). The Cq values and the efficiencies of the samples and standards was determined as previously using LinRegPCR version 2016.0 [28]. The quantities were calculated using the one point calibration method as described earlier [30].

Amplicon library construction and sequencing

16S rRNA gene libraries were constructed using primers designed to amplify the V4 region and multiplexed [31]. Amplicons were generated using a high-fidelity polymerase (Kapa 2G Robust) and purified using the Agencourt AMPure XP PCR purification system and quantified using a fluorometer (Qubit, life technologies). The purified amplicons were then pooled in equimolar concentrations by hand based on Qubit quantification. The resulting amplicon library pool was diluted to 2 nM with sodium hydroxide and 5 μL transferred into 995 μL HT1 (Illumina) to give a final concentration of 10 pM. 600 μL of the diluted library pool was spiked with 10% PhiX Control v3 and placed on ice before loading into Illumina MiSeq cartridge following the manufacturer's instructions. The sequencing chemistry utilized was MiSeq Reagent Kit v2 (500 cycles) with run metrics of 250 cycles for each paired end read using MiSeq Control Software 2.2.0 and RTA 1.17.28.

Analyses of sequenced samples

MiSeq amplicon reads were merged using Illumina-utils software [32]. We quality-filtered only the mismatches in the overlapping region between read pairs using a minimum overlap (–min-overlap-size) of 200 nt and a minimum quality Phred score (–min-qual-score) of Q20. We allowed no more than five mismatches per 100 nt (-P 0.05) over the 200 nt overlapping region.

Reads that fulfilled the quality criteria were analyzed using Quantitative Insights Into Microbial Ecology (QIIME v.1.7) [29]. We removed chimera using the *identify_chimeric_seqs.py* script, UCHIME reference 'Gold' database and USEARCH [33, 34], which we also used to select OTUs. We assigned the taxonomy of our reads with QIIME *pick_open_reference_otus.py* function, using the Greengenes database version v13_8 [35] as a reference with a minimum cluster size of 2 (i.e., each OTU must contain at least two sequences). We collapsed the technical replicates and filtered out the low abundance OTUs (< 0.01% total, *filter_otus_from_otu_table.py*) and samples rarefied to an even depth of 26702 for both experiments where sequencing data is available. QIIME was used to calculate alpha and beta diversity data and produce NMDS plots.

QUANTIFICATION AND STATISTICAL ANALYSIS

MacQiime was used to calculate alpha and beta diversity data and produce NMDS plots. Data obtained with MacQiime was later combined with the gas production data and analyzed using JMP Pro 13 software (SAP) as described in the Figure legends.

For the NNLS analysis, following removal of low abundance OTUs and cumulative sum scaling transformation, the resulting biom file was used to create a matrix $A \in \mathbb{Z}_{\geq 0}^{m \times n}$ (m rows of OTUs by n sample columns) for all of the single communities, and a column vector $b \in \mathbb{Z}_{\geq 0}^{m}$ for each of the mixed communities; both A and b hold non-negative integers of OTU abundances. Note that one of the individual samples contained a negligible number of reads and was discarded from the analysis. The contribution, or weight, of each seed sample to the pattern of OTUs observed in each of the mixed communities is given by the column vector $x \in \mathbb{R}^n$ when solving a system of linear equations $Ax = b$. Written out this equation $(Ax = b)$ looks like this for each mixture:

$$\begin{pmatrix} \mathrm{OTU}_{1,S1} & \cdots & \mathrm{OTU}_{1,Sn} \\ \vdots & \ddots & \vdots \\ \mathrm{OTU}_{m,S1} & \cdots & \mathrm{OTU}_{m,Sn} \end{pmatrix} \mathrm{x} \begin{pmatrix} x_{S1} \\ \vdots \\ x_{Sm} \end{pmatrix} = \begin{pmatrix} OTU_{mix\_1} \\ \vdots \\ OTU_{mix\_m} \end{pmatrix}$$

where S refers to each single community.

When modeling count data for environmental samples the fitted parameters of x will also be non-negative and the number of OTUs will usually exceed the number of samples ($m > n$). The task is to solve an over-determined system of linear equations where there are more equations than unknowns. It is likely that some of the linear equations will 'disagree' and there will be no exact solution. Geometrically, this may be interpreted as b not lying in the column space of A, a (hyper)plane holding the column vectors of A, or $Ax - b \neq 0$. A least-squares approach may find the non-negative vector $\overline{x} = (A^T A)^{-1} A^T b$ which is the projection of b back onto the column space of A that minimizes the least-squares error 'distance' $\|A\overline{x} - b\|$. For our study the non-negative least-squares (NNLS) solution, $\overline{x}$, and least-squares errors were computed via the R packages 'nnls' [36] and 'limSolve' [37] for each of the mixed samples, with the mean relative contribution of each community to the 10 mixtures shown in Figure 2C (relative to the gas production).

DATA AND SOFTWARE AVAILABILITY

The raw sequences obtained from our experiments are available online in the European Nucleotide Archive under accession numbers ENA: PRJEB21193 (Experiment 1) and ENA: PREJEB21187 (Experiment 2). We also included the R code that allows the user to calculate the contribution a single community has in a mix of communities (see Method S1). Using this code, a NNLS analysis can be conducted with the input of a pre-filtered OTU table.