Report

Current Biology

Modular Assembly of Polysaccharide-Degrading Marine Microbial Communities

Highlights

- Complex marine microbial communities are composed of functional groups
- Functional groups occupy trophic levels connected by metabolic cross-feeding
- Communities assemble by recruiting these functional groups in a modular fashion
- Community composition can be predicted based on substrate composition

Authors

Tim N. Enke, Manoshi S. Datta, Julia Schwartzman, ..., Julien Barrere, Alberto Pascual-García, Otto X. Cordero

Correspondence

ottox@mit.edu

In Brief

Enke et al. show that particle-attached marine microbial communities assemble by recruiting functional groups of taxa in an additive manner. Specialist groups degrade specific polysaccharides, whereas generalist byproduct utilizers invade independently of particle substrate. This simple organization allows prediction of community structure.

Enke et al., 2019, Current Biology *29*, 1528–1535
May 6, 2019
https://doi.org/10.1016/j.cub.2019.03.047

CellPress

CellPress

Current Biology
Report

Modular Assembly of Polysaccharide-Degrading Marine Microbial Communities

Tim N. Enke,[1,2,4] Manoshi S. Datta,[1,4] Julia Schwartzman,[1] Nathan Cermak,[3] Désirée Schmitz,[2] Julien Barrere,[2] Alberto Pascual-García,[2] and Otto X. Cordero[1,5,*]

[1]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Department of Environmental Systems Science, ETH Zurich, 8093 Zurich, Switzerland
[3]Program in Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]These authors contributed equally
[5]Lead Contact
*Correspondence: ottox@mit.edu
https://doi.org/10.1016/j.cub.2019.03.047

SUMMARY

Understanding the principles that govern the assembly of microbial communities across earth's biomes is a major challenge in modern microbial ecology. This pursuit is complicated by the difficulties of mapping functional roles and interactions onto communities with immense taxonomic diversity and of identifying the scale at which microbes interact [1]. To address this challenge, here, we focused on the bacterial communities that colonize and degrade particulate organic matter in the ocean [2–4]. We show that the assembly of these communities can be simplified as a linear combination of functional modules. Using synthetic polysaccharide particles immersed in natural bacterioplankton assemblages [1, 5], we showed that successional particle colonization dynamics are driven by the interaction of two types of modules: a first type made of narrowly specialized primary degraders, whose dynamics are controlled by particle polysaccharide composition, and a second type containing substrate-independent taxa whose dynamics are controlled by interspecific interactions—in particular, cross-feeding via organic acids, amino acids, and other metabolic byproducts. We show that, as a consequence of this trophic structure, communities can assemble modularly—i.e., by a simple sum of substrate-specific primary degrader modules, one for each complex polysaccharide in the particle, connected to a single broad-niche range consumer module. Consistent with this model, a linear combination of the communities on single-polysaccharide particles accurately predicts community composition on mixed-polysaccharide particles. Our results suggest that the assembly of heterotrophic communities that degrade complex organic materials follows simple design principles that could be exploited to engineer heterotrophic microbiomes.

RESULTS AND DISCUSSION

Community Dynamics and Functional Groups

Particle-attached bacteria mediate carbon cycling in aquatic environments by degrading insoluble forms of organic matter that concentrate on particles [2–4, 6]. To degrade particles, bacteria attach to their surfaces, forming dense multi-species communities with high enzymatic activity [7–10]. From a community ecology standpoint, the assembly of these particle-attached communities is controlled by strong environmental filters that allow only a small subset of organisms to colonize, out of the vast diversity present in the water column. Laboratory and field studies indicate that the assembly process follows fast successional dynamics driven by a rapid attachment-detachment process and cross-feeding interactions [5, 8]. This process leads to the recruitment of hundreds of different taxa with diverse metabolic potentials, and it is unclear whether simple rules exist that could help us predict the composition and dynamics of their communities. A first step toward finding the rules that guide microbial community assembly is to group taxa into functional units, encompassing species with similar metabolism and dynamics, thereby reducing the complexity of taxonomically diverse communities [11–13]. If such functional units exist, one could ask to what extent can community assembly be described as a simple recruitment of functional groups to meet the metabolic demands of the environment. To the extent that communities assemble in this simple, modular fashion, their composition should be predictable in terms of linear combinations of the less complex communities that respond to the individual metabolic demands.

To test whether particle-attached marine bacterial communities can assemble in a modular fashion, we studied their assembly dynamics across model marine particles made out of one of four defined linear polysaccharides, chosen for their structural and compositional diversity and their abundance in marine environments—chitin, alginate, agarose, and carrageenan (Figure 1A)—or with a combination of these substrates. Chitin is frequently found in the shells of crustaceans such as copepods as well as on the cell walls of diatoms [14, 15]. It is a composed of β1,4-linked subunits of the amino sugar *N*-acetylglucosamine and can align in crystalline form to assemble in different conformations. Alginate is a structural component of

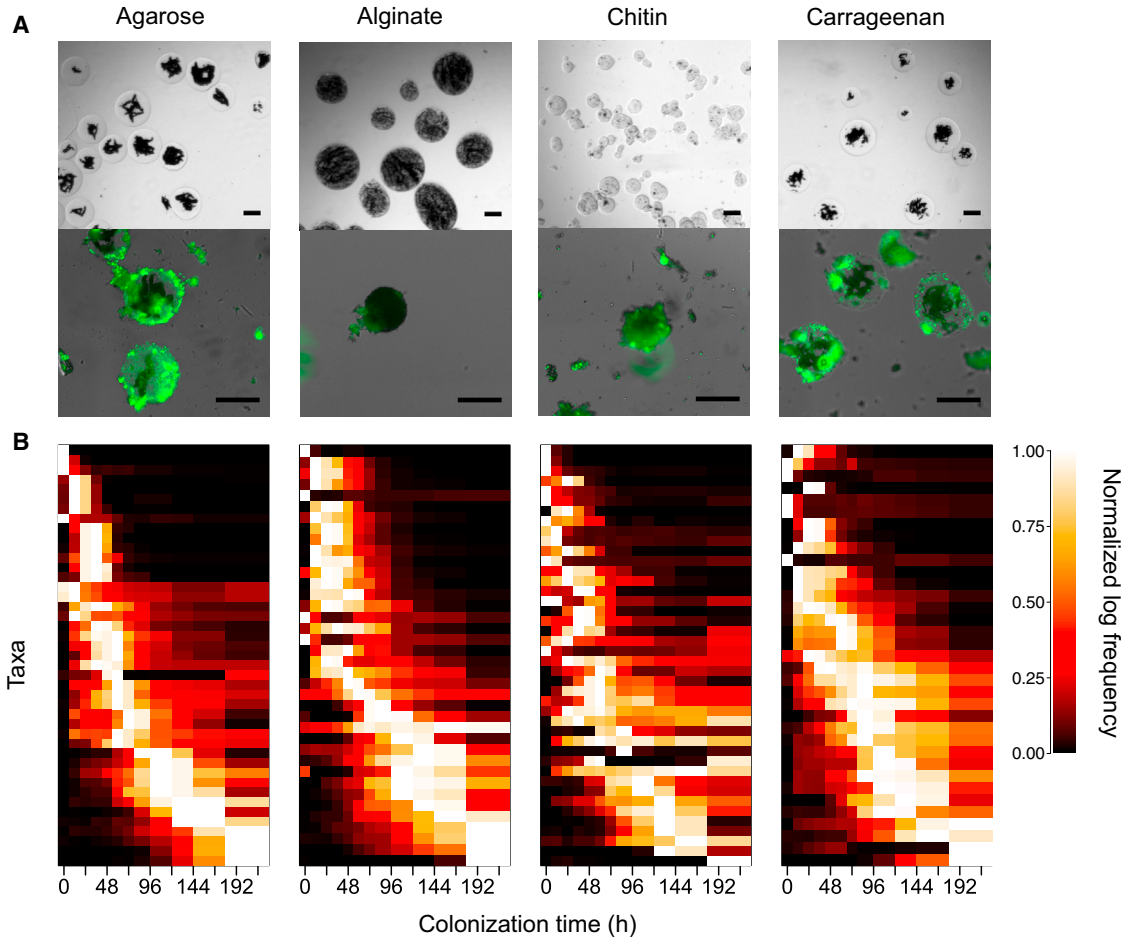

Figure 1. Rapid Successional Dynamics on Four Different Marine Polysaccharides

(A) Paramagnetic hydrogel beads made of agarose, alginate, chitin, or carrageenan are incubated in natural, unfiltered coastal seawater. Top: phase contrast images of the particles (with magnetite cores in black). Bottom: fluorescence microscopy images of particles stained with SYTO9 after 136 h of incubation, revealing dense microbial communities on particle surfaces. Scale bar corresponds to 100 μm.

(B) Successional dynamics on each particle type. For each particle type, heatmaps show the dynamics of ASVs with fractional abundance >1%, ordered by time of maximal fractional abundance. Despite differences in ASV composition (analyzed in Figure 2), the successional dynamics are nearly identical across different particle types. Community dynamics on beads and in seawater are further analyzed in Figure S1.

the cell walls of brown algae and is heteropolymer composed of the uronic acids mannuronate and its epimer guluronate. Agarose and carrageenan are enriched in red algae [16] and are structurally similar. The former is a copolymer of alternating units of D-galactose and its modified epimer 3,6-anhydro-L-galactopyranose, whereas the latter is a sulfated polymer composed of alternating D-galactose and 3,6-anhydrogalactose residues, which requires additional removal of sulfate groups to be utilized by cells.

Model marine particles ranging from 50 to 200 μm in diameter were manufactured using polysaccharide hydrogels containing paramagnetic powder embedded in the gel matrix to enable sampling from the seawater suspension. The use of hydrogels as model substrates removed the potential structural complexity in which polysaccharides are presented in nature (e.g., crystalline structures), while keeping the chemical composition constant. This compromise ensured that any potential difference in the assembly process on different particle types is driven primarily by the specificity of the interaction between polysaccharide and bacteria, and not by factors such as the physical structure of the surface (e.g., porosity). In order to study the community assembly dynamics as a function of the four different particle substrates, we suspended the different particle types in separate incubations with coastal seawater subsampled from the same ~20 L (STAR Methods), in order to allow access to the same species pool.

Previous work with chitin model particles has shown that, when colonized from a species pool in seawater, community assembly proceeds in a reproducible succession whereby early colonizers degrade chitin and facilitate the invasion of secondary consumers that lack enzymes required to hydrolyze chitin [5]. Consistent with this model, we found that community assembly proceeded via rapid successional dynamics across the four single-substrate particle types, indicating that the dynamics of community assembly are not specific to chitin. To characterize these dynamics, we collected samples of ~1,000 particles and

surrounding seawater at each of 12 time points, from 0 to 204 h and sequenced their communities using 16S rRNA gene amplicon sequencing (STAR Methods). On all four particle types tested, most taxa present at high fractional abundance in the first 12 h declined substantially in fractional abundance by 72–96 h (Figure 1B), suggesting a remarkably similar and rapid community succession. In general, these dynamics were different from the dynamics in the surrounding seawater (Figure S1), showing that the composition of the particle-attached community was controlled by processes specific to particles, such as cell-cell and cell-particle interactions.

Despite the overall similarity in colonization dynamics across particle types, the abundance and dynamics of individual amplicon sequence variants or ASVs [17] on different particle types was not necessarily conserved. To quantify differences in ASV abundance across particle types, we calculated a realized niche width index for the ASVs (hereafter called "niche width index" for purposes of brevity). To calculate this score, for each ASV, i, and each particle type, j, we computed the geometric mean frequency over time, f_{ij}, renormalized the mean frequencies so $\sum_j f_{ij} = 1$ and calculated the entropy of the mean ASV abundance over all single-substrate particle types, $-\sum_j f_{ij} log_2(f_{ij})$ (STAR Methods). The entropy represents an index that describes how uniformly the ASV was distributed over the four substrates. ASVs that appeared only on one particle type had a niche width score = 0, whereas ASVs that were equally prevalent across all particle types ($f_{ij} = 0.25$ for all j) had a niche width score index of 2.

We found that within particle-associated communities the distribution of the niche width indexes was bimodal (top histogram in Figure 2A). Using a Gaussian mixture model to cluster ASVs by distribution mode (STAR Methods), we found that 36% of the ASVs shown in Figure 2A grouped into a cluster of narrow-range taxa (niche width score <0.18) and 42% into a cluster of broad-range taxa (niche width score >1.52). Moreover, an unsupervised hierarchical clustering of ASVs based on their temporal dynamics across particle types allowed us to further partition narrow-range taxa by the substrate they appeared on (heatmap in Figure 2A). The best partitioning of the data divided ASVs into five natural groups, one for the broad-range taxa present on all particle types and four separate groups of narrow-range taxa for each of the four particle types (Figures 2B and 2C; STAR Methods). The broad-range group encompassed organisms that were not only highly prevalent across all particle types, but whose dynamics were highly correlated across substrates (average Spearman correlation = 0.54 across four particle types, Figure S2). This broad-range group was more diverse than the narrow-range ones and, on average, increased in frequency toward later time points, leading to a net increase in diversity during the course of succession (Figure S3). The invasion of the broad-range group at later stages of succession caused community composition across particle types to first diverge due to the colonization of narrow-range species (reaching maximum divergence at ~24 h) before converging to a set of broad-range taxa (Figure S4).

Metabolic Roles of Functional Groups

To study whether narrow and broad range groups were taxonomically cohesive, we mapped the functional groups onto an ASV phylogeny (Figure 2D). The phylogeny of the ASVs showed that narrow- and broad-range groups were associated with distinct taxonomic groups and distinct metabolic potentials, suggesting that there are consistent associations between taxonomic groups and trophic levels in the marine environment. Narrow-range groups mapped primarily to known particle-associated carbohydrate degraders in the ocean, such as the family *Flavobacteriaceae* or the genus *Sacharophagus* [18–21], leading us to hypothesize that these narrow-range taxa are specialized primary degraders. This potential specialization of taxa toward polysaccharide degradation is consistent with reports of this type of substrate being able to stimulate specialized bacterioplankton taxa in a highly specific manner [21].

To experimentally confirm the polysaccharide-degrading ability of narrow-range groups and to gain insight into the role of the less taxonomically defined broad-range group, we cultured 874 bacterial isolates from particles and sequenced their 16S rRNA V4 region (STAR Methods). Out of these, 247 isolates had a 100% identity match to 12 broad-range ASVs. Only 2, however, mapped to 2 narrow-range ASVs. We focused our efforts on one of these narrow-range isolates, which we named psychB3M02, and belonged to the genus *Psychromonas* in the alginate-specific group (marked with a red arrow in Figure 2D). In agreement with its specific association with alginate particles, psychB3M02 was able to grow on alginate as sole carbon source (Figure 3A). Moreover, HMM-based searches of glycosyl hydrolase (GH) and polysaccharide lyase (PL) families against its genome identified multiple copies of alginate lyases (PL7, 8 copies) and oligoalginate lyases (PL15, PL17, 4 copies) but found no other genes coding secreted enzymes for degrading other marine polysaccharides such as chitin (GH18, GH19, GH20) or agarose (GH16) (Table S1), in agreement with the notion that this organism is a polysaccharide specialist.

By contrast, hydrolytic enzymes were typically absent from the genomes of broad-range isolates. None of three sequenced isolates of the *Rhodobacteraceae* (α-proteobacteria), a clade exclusively found in the broad-range group (Figure 2D, clade 1), encoded genes to produce hydrolytic enzymes (Table S1). Two members of this clade, however, a *Loktanella*, loktaD2R18, and a *Ruegeria*, ruegeA3M17, had the machinery to import and utilize oligosaccharides of alginate and chitin, respectively, suggesting a potential role as "free-riders." The third organism, phaeoC3M10, classified as a *Phaeobacter*, had no genes to convert cytoplasmic intermediates into central metabolic substrates, indicating that this strain cannot harvest oligosaccharides and instead relies on metabolic intermediates released by other members of the community. These observations suggest that growth of the broad-range taxa may be facilitated by narrow-range taxa through simple metabolic intermediates.

To confirm this hypothesis, we collected spent media from psychB3M02 grown to peak cell density on alginate as the sole carbon source and asked whether this media would support growth of a panel of five broad range taxa that were unable to degrade and grow on alginate by themselves. To expand the diversity of the panel, we tested the three *Rhodobacteraceae* discussed above plus a *Marinobacter* and a *Vibrio* isolate from broad-range clades (Figure 3A). In accordance with our expectation, all five broad-range taxa were able to grow on

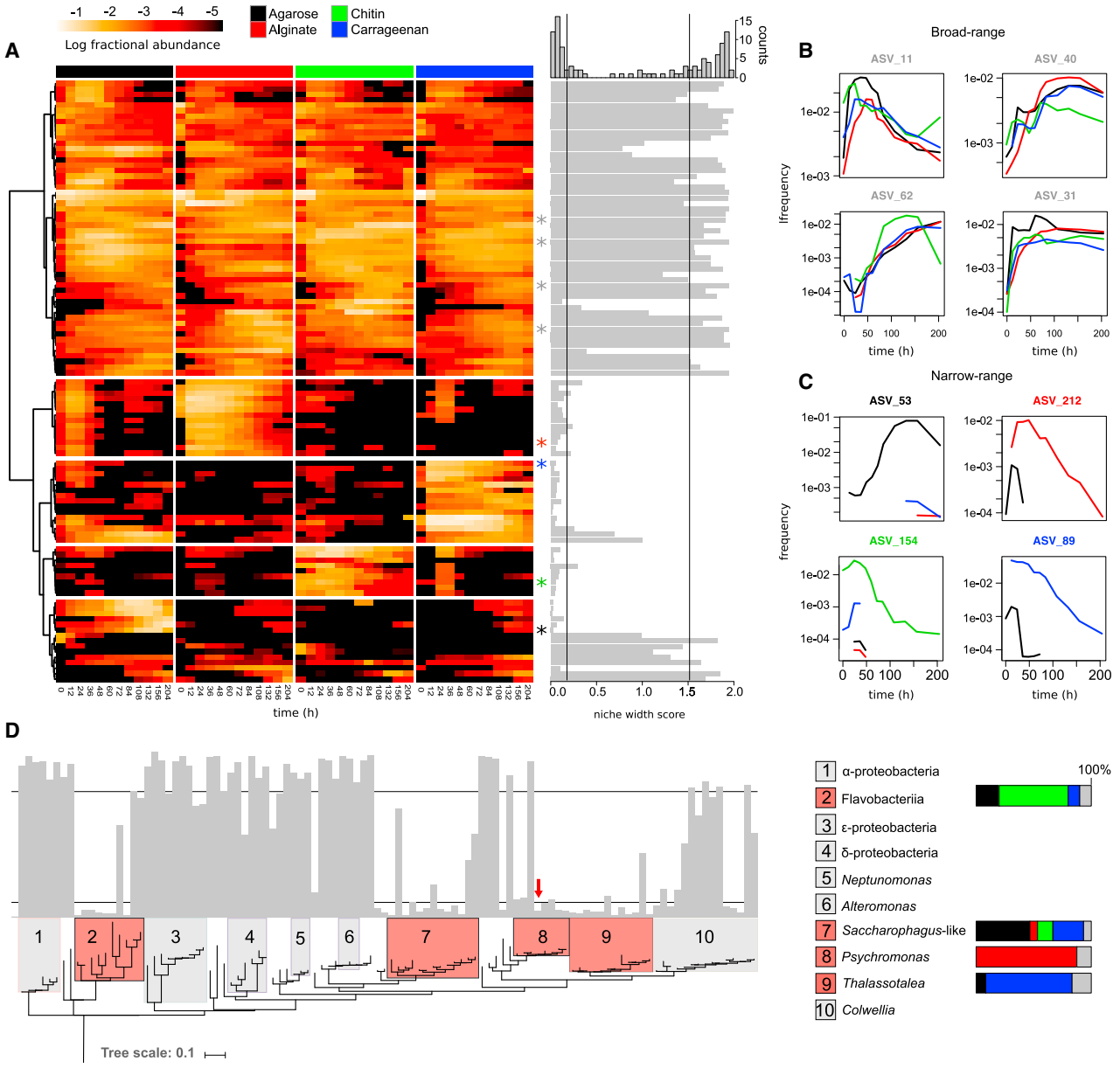

Figure 2. Unsupervised Detection of Narrow-Range and Broad-Range Functional Groups

(A) Clustering of taxa by occurrence across four particle types. The data show that taxa can be divided in a large cluster of broad-range taxa (top part of heatmap) and smaller clusters of narrow-range taxa (bottom part of heatmap). The realized niche width score (gray bars) shows that the distribution of niche widths is bimodal (see histogram on top).

(B and C) Dynamics of broad-range ASVs (B) and narrow-range ASVs (C). The positions of these specific ASVs are marked with an asterisk in the heatmap in (A). The color of the asterisk corresponds to the color-coding of substrates (legend in A), with the broad-range ASVs colored in gray. Figure S2 shows the dynamics of all broad-range (Figure S2A) and narrow-range (Figure S2B) taxa, respectively, and Figures S3 and S4 show overall metrics of community diversity and convergence, respectively.

(D) Phylogenetic distribution of ASVs based on 16S rRNA gene sequence similarity. Clades marked with numbers correspond to the largest monophyletic clades, defined at the class level for groups 1–4, and at the genus level for groups 5–10. In red are those monophyletic clades with a high incidence of narrow-range taxa (>50%). For those clades, the distribution of ASVs across the four substrates is shown on the right. Tree rooted with Sulfolobus as outgroup (not shown). Red arrow points to the position of psychB3M02.

the spent media, even without supplementing it with additional nutrients (Figure 3A). This confirms that, in an environment where alginate is the sole carbon source, narrow-range alginate degraders can facilitate the growth of broad-range, non-degrading taxa.

To learn more about the exact mechanisms of facilitation and its apparent non-specific nature, we performed a targeted metabolomic analysis [22] of psychB3M02's spent media before and after growth of non-degrading broad range taxa (STAR Methods). The data showed that non-degraders support their

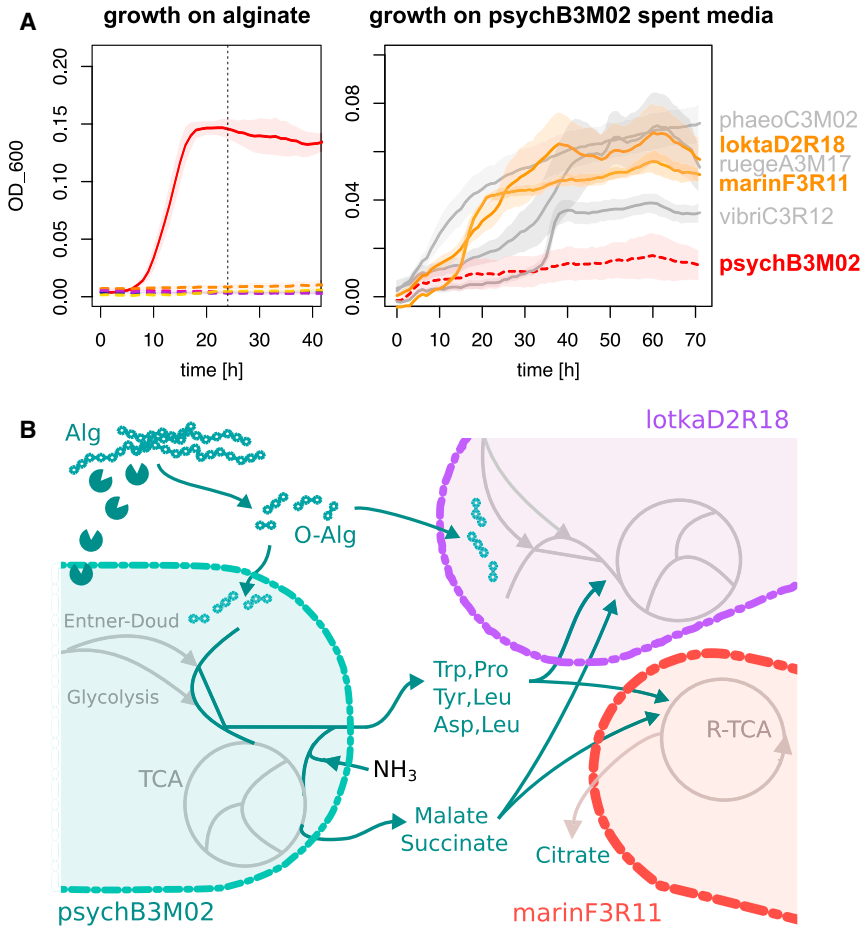

Figure 3. Facilitation of the Broad-Range Module Is Generic and Mediated by Multiple Amino Acid and Organic Acid Excretions

(A) Growth curves of a narrow-range degrader, psychB3M02, and 5 broad-range non-degraders on alginate (left) and on spent media of psychB3M02. Mean and SE of the OD measurements are calculated for three replicates.

(B) Model of possible cross-feeding pathways inferred from full genomes of psychB3M02, loktaD2R18, and marinF3R11, as well as from targeted metabolomics data (see metabolic pathway predictions in Table S1 and metabolites consumed Table S2; a full list of metabolites characterized is found in the STAR Methods).

growth by taking up multiple small metabolic byproducts. For this analysis, we focused on two non-degrading strains whose genomes suggested divergent metabolic capabilities: the *Loktanella* loktaD2R18 and the *Marinobacter* marinF3R11. We identified compounds that were produced by psychB3M02 and consumed by one of the non-degraders in at least two out of three replicates. Out of 82 possible compounds, we detected 11 compounds that fulfilled this criterion: these included six amino acids (Figure 3B); the amino acid precursor 3-methyl-2-oxopentanoic acid; tricarboxylic acid (TCA) cycle intermediates malate and succinate; nucleosides; and nucleotides (Table S2; STAR Methods). This general consumption of multiple metabolic intermediates was observed for both marinF3R11 and loktaD2R18. Some metabolites that could support growth of non-degraders were also released to the medium by non-degraders (Figure 3B). In particular, marinF3R11 secreted citrate, consistent with the prediction that this organism uses a reductive TCA cycle (Table S1).

Modular Assembly on Mixed-Substrate Particles

Having identified five distinct functional groups and their metabolic roles, we asked whether communities capable of degrading multiple polysaccharides could be assembled in a modular fashion, that is, by a simple aggregation of functional groups. To test this hypothesis, we examined community assembly dynamics on particles made of substrate mixtures and compared these dynamics to the one observed on the corresponding single substrate particles. In particular, we tested two polysaccharide mixtures: agarose-alginate and agarose-carrageenan (50% of each substrate by mass), which were incubated in the same seawater and conditions used for single substrate particles.

Consistent with the hypothesis of modularity, a simple linear combination of the species abundances on each single substrate accurately predicted the composition of communities assembled on mixed particles (Figure 4A). To quantify this, we fitted the vector of ASV geometric mean frequencies on the mixed particles with a linear combination of the vectors of the corresponding single substrate particles (STAR Methods). The best-fitting linear model (lm) for the agarose- alginate mixture, $\overrightarrow{ASV}_{agarose-alginate} = \alpha\overrightarrow{ASV}_{agarose} + \beta\overrightarrow{ASV}_{alginate}$, had an R^2 of 0.84, and the corresponding model for agarose-carrageenan an R^2 of 0.74, showing that a linear combination had high-predictive power (Figure 4A). To rule out the possibility that the result was driven by broad-range taxa, we calculated the Spearman correlation coefficient between model and data only for the relevant narrow-range ASVs, finding values of 0.75 and 0.83 for the agarose-alginate and agarose-carrageenan communities, respectively. Furthermore, we fitted a model with an explicit interaction term to test whether this would improve the results. We found that for the agarose-alginate such a nonlinear model had an inferior goodness of fit compared to the simple linear combination (STAR Methods). In the case of agarose-carrageenan particles, the nonlinear model (nlm) improves the fit relative to the lm, but only marginally (R^2 = 0.76 versus 0.74 in the lm) and the model is only weakly nonlinear (nlm $\sim$lm$^{0.98}$) (STAR Methods). When we considered ASV dynamics, we found these were also highly correlated between single and mixed particles, in a manner consistent with a model of community assembly by simple linear aggregation of ecological modules (Figures 4B and 4C). Across all narrow-range ASVs, the median Spearman correlation between the single- and mixed-substrate

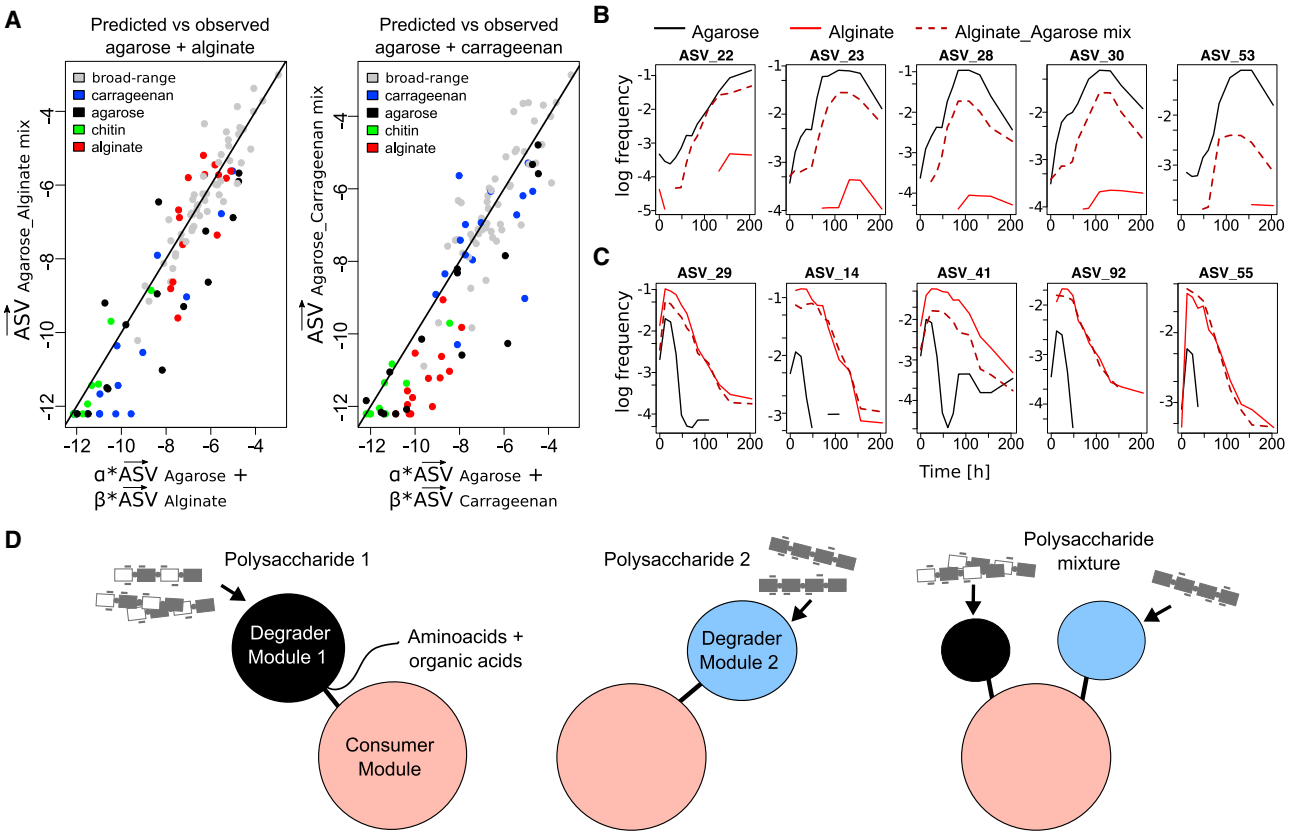

Figure 4. Communities Assemble by Linear Combinations of Functional Groups

(A) ASV frequencies in mixed particles plotted in log-log scale against the predicted ASV frequencies, based on a linear combination of single substrate vectors. The fitted coefficients are $\alpha = 0.67$, $\beta = 0.40$ for agarose-alginate, and $\alpha = 0.89$, $\beta = 0.11$ for agarose-carrageenan.

(B and C) Similar ASV trajectories in mixed versus single substrate particles for agarose (B) and alginate (C) specific ASVs. Solid lines depict trajectories in single substrate particles and dashed lines in mixed particles. The median Spearman correlation between the dynamics of agarose-specific ASVs on single and mixed substrate particles is 0.86 (B), and for the alginate-specific ASVs 0.96 (C) (Table S3).

(D) Model of modular assembly, which mirrors the structure of metabolic pathways. Peripheral, narrow-range modules perform the degradation of complex biopolymers, whereas the core, broad-range module processes simple metabolic intermediates.

time dynamics ranges between 0.65 and 0.96 (Table S3). Overall, these results show that communities on substrate mixtures could be linearly composed from the communities on single substrates. Taken together, our results suggest that at the level of functional groups, microbial communities have a simple trophic structure that allows functionally complex communities to be assembled in a modular fashion (Figure 4D).

Our study shows that, although specific functional groups are recruited depending on polysaccharide identity, community assembly on marine particles follows rapid successional dynamics that are independent of the substrate. Although metabolic cross-feeding appears to be the main driver of successional dynamics, establishing the flow of energy from narrow-range degraders to broad-range consumers, it remains unclear what drives the fast and seemingly generic species turnover observed on particles. In particular, it is unclear why species leave the particle at early stages before fully degrading it. One tantalizing hypotheses is that frequent detachment is a programmed foraging behavior that allows bacteria to explore new particles [23]. In addition, detachment can also be partly driven by biotic interactions, such as phage predation, which is likely to take place on particles given the fast growth rates of particle-attached bacteria, or cell-cell communication such as quorum sensing. Further research is needed to understand both how foraging strategies on patchy landscapes may affect the dynamic of bacteria-particle interactions and the potential role of biotic interactions in modulating dispersal.

The ability to decompose or compose interconnected systems by simple combinations of functional groups is a common property of many complex biological systems, such as metabolic networks or multi-domain proteins [24, 25]. Here, we have shown that the assembly of microbial communities can also follow a similar logic. We suggest this simple logic provides us with potential design principles that can be exploited to create synthetic consortia and controllable by way of substrate supply.

STAR★METHODS

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

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2019.03.047.

ACKNOWLEDGMENTS

We thank Gabriel E. Leventhal, Shaul Pollak, and Elizabeth Kujawinski for their comments and careful reading of the manuscript. We also thank all members of the Cordero lab and, in particular, José Saavedra and Matthew Metzger for their support. This project was supported by Simons Early Career Award 410104, the Alfred P Sloan fellowship FG-20166236, NSF grant OCE-1658451, and the Simons Collaboration: Principles of Microbial Ecosystems (PriME) award number 542395.

AUTHOR CONTRIBUTIONS

M.S.D., N.C., and O.X.C. designed the study. M.S.D., N.C., T.N.E., J.S., D.S., and J.B. executed the study. M.S.D., T.N.E., J.S., D.S., and O.X.C. analyzed the results. T.N.E., J.S., and O.X.C. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 11, 2019
Revised: February 27, 2019
Accepted: March 21, 2019
Published: April 25, 2019

REFERENCES

1. Cordero, O.X., and Datta, M.S. (2016). Microbial interactions and community assembly at microscales. Curr. Opin. Microbiol. *31*, 227–234.
2. Alldredge, A.L., and Silver, M.W. (1988). Characteristics, dynamics and significance of marine snow. Prog. Oceanogr. *20*, 41–82.
3. Azam, F., and Malfatti, F. (2007). Microbial structuring of marine ecosystems. Nat. Rev. Microbiol. *5*, 782–791.
4. Azam, F., and Long, R.A. (2001). Sea snow microcosms. Nature *414*, 495, 497–498.
5. Datta, M.S., Sliwerska, E., Gore, J., Polz, M.F., and Cordero, O.X. (2016). Microbial interactions lead to rapid micro-scale successions on model marine particles. Nat. Commun. *7*, 11965.
6. Smith, D.C., Simon, M., Alldredge, A.L., and Azam, F. (1992). Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. Nature *359*, 139–142.
7. Fontanez, K.M., Eppley, J.M., Samo, T.J., Karl, D.M., and DeLong, E.F. (2015). Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. Front. Microbiol. *6*, 469.
8. Pelve, E.A., Fontanez, K.M., and DeLong, E.F. (2017). Bacterial Succession on Sinking Particles in the Ocean's Interior. Front. Microbiol. *8*, 2269.
9. Ganesh, S., Parris, D.J., DeLong, E.F., and Stewart, F.J. (2014). Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. ISME J. *8*, 187–211.
10. Stewart, F.J. (2013). Where the genes flow. Nat. Geosci. *6*, 688–690.
11. Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. Proc. Natl. Acad. Sci. USA *108*, 14288–14293.
12. Louca, S., Polz, M.F., Mazel, F., Albright, M.B.N., Huber, J.A., O'Connor, M.I., Ackermann, M., Hahn, A.S., Srivastava, D.S., Crowe, S.A., et al. (2018). Function and functional redundancy in microbial systems. Nat. Ecol. Evol. *2*, 936–943.
13. Louca, S., Parfrey, L.W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. Science *353*, 1272–1277.
14. Durkin, C.A., Mock, T., and Armbrust, E.V. (2009). Chitin in diatoms and its association with the cell wall. Eukaryot. Cell *8*, 1038–1050.
15. Jeuniaux, C., and Voss-Foucart, M.F. (1991). Chitin biomass and production in the marine environment. Biochem. Syst. Ecol. *19*, 347–356.
16. Usov, A.I. (2011). Polysaccharides of the red algae. Adv. Carbohydr. Chem. Biochem. *65*, 115–217.
17. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583.
18. Weiner, R.M., Taylor, L.E., 2nd, Henrissat, B., Hauser, L., Land, M., Coutinho, P.M., Rancurel, C., Saunders, E.H., Longmire, A.G., Zhang, H., et al. (2008). Complete genome sequence of the complex carbohydrate-degrading marine bacterium, Saccharophagus degradans strain 2-40 T. PLoS Genet. *4*, e1000087.
19. Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schüler, M., Teeling, H., Amann, R.I., and Schweder, T. (2014). Functional characterization of polysaccharide utilization loci in the marine Bacteroidetes 'Gramella forsetii' KT0803. ISME J. *8*, 1492–1502.
20. Cottrell, M.T., and Kirchman, D.L. (2000). Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. Appl. Environ. Microbiol. *66*, 1692–1697.
21. Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., Kassabgy, M., Huang, S., Mann, A.J., Waldmann, J., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. Science *336*, 608–611.
22. Longnecker, K., Futrelle, J., Coburn, E., Kido Soule, M.C., and Kujawinski, E.B. (2015). Environmental metabolomics: Databases and tools for data analysis. Mar. Chem. *177*, 366–373.
23. Fernandez, V.I., Yawata, Y., and Stocker, R. (2019). A Foraging Mandala for Aquatic Microorganisms. ISME J. *13*, 563–575.
24. Wagner, G.P., Pavlicev, M., and Cheverud, J.M. (2007). The road to modularity. Nat. Rev. Genet. *8*, 921–931.
25. Nash, P.D. (2012). Why modules matter. FEBS Lett. *586*, 2572–2574.
26. Turner, S., Pryer, K.M., Miao, V.P., and Palmer, J.D. (1999). Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. J. Eukaryot. Microbiol. *46*, 327–338.

27. Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., and Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. USA *108* (*Suppl 1*), 4516–4522.

28. Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. *73*, 5261–5267.

29. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

30. Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. *25*, 1043–1055.

31. Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., et al. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics *9*, 75.

32. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. *42*, D490–D495.

33. Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P.K., Xu, Y., and Yin, Y. (2018). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. *46* (W1), W95–W101.

34. Chambers, M.C., Maclean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., et al. (2012). A cross-platform toolkit for mass spectrometry and proteomics. Nat. Biotechnol. *30*, 918–920.

35. Clasquin, M.F., Melamud, E., and Rabinowitz, J.D. (2012). LC-MS Data Processing with MAVEN: A Metabolomic Analysis and Visualization Engine. Current Protocols in Bioinformatics (John Wiley & Sons, Inc.).

36. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

37. Pruesse, E., Peplies, J., and Glöckner, F.O. (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics *28*, 1823–1829.

38. Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS ONE *5*, e9490.

39. Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. *44*, W242–W245.

40. Wolters, G.H.J., Fritschy, W.M., Gerrits, D., and van Schilfgaarde, R. (1991). A versatile alginate droplet generator applicable for microencapsulation of pancreatic islets. J. Appl. Biomater. *3*, 281–286.

41. Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G. (2008). A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnol. Oceanogr. Methods *6*, 230–235.

42. Longnecker, K. (2015). Dissolved organic matter in newly formed sea ice and surface seawater. Geochim. Cosmochim. Acta *171*, 39–49.

43. Kido Soule, M.C., Longnecker, K., Johnson, W.M., and Kujawinski, E.B. (2015). Environmental metabolomics: Analytical strategies. Mar. Chem. *177*, 374–387.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Loktanella sp. D2R18	This study	SAMN09522135
Marinobacter sp. F3R11	This study	SAMN09522136
Phaeobacter gallaeciensis C3M10	This study	SAMN09522133
Psychromonas sp. B3M02	This study	SAMN09522130
Ruegeria sp. A3M17	This study	SAMN09522134
Tenacibaculum sp. E3R01	This study	SAMN09522131
Vibrionales bacterium C3R12	This study	SAMN09522132
Chemicals, Peptides, and Recombinant Proteins		
Artificial seawater	Sigma-Aldrich	Cat#S9883
Bacto Agar	BD Diagnostics	Cat#214010
Marine Broth 2216	Difco	Cat#279110
Low viscosity alginate	Sigma-Aldrich	Cat #A1112
Carrageenan, predominantly κ, lesser amounts λ	Sigma-Aldrich	Cat#C1013
Omnipur agarose	EMD Millipore	Cat#2120
Synthetic black iron oxide	Alpha Chemicals	Cat#725410978229
Mineral oil	Sigma-Aldrich	Cat#330779
SPAN-80	Sigma-Aldrich	Cat#85548
Medium viscosity alginate	Sigma-Aldrich	Cat#A2033
Chitin magnetic beads	New England Biolabs	Cat#E8036L
SYTO9 nucleic acid stain	Thermo Fisher	Cat#S34854
Phusion Polymerase with 5x HF Buffer	New England Biolabs	Cat#M0530S
SYBR Green	Thermo Fisher	Cat#S7567
Ampure XP SPRI beads	Beckman Coulter	Cat#A63880
Critical Commercial Assays		
MasterPure DNA Purification Kit	Epicenter	Cat#MCD85201
Agincourt DNA Advance Kit	Beckman Coulter	Cat#A48705
Nextera XT DNA Library Preparation Kit	Illumina	Cat#FC-131-1096
Nextera XT Index Kit (24 indexes, 96 samples)	Illumina	Cat#FC-131-1001
Deposited Data		
Whole genome shotgun sequence assembly C3M10	This study	GCA_003318155.1
Whole genome shotgun sequence assembly B3M02	This study	GCA_003318165.1
Whole genome shotgun sequence assembly C3R12	This study	GCA_003318175.1
Whole genome shotgun sequence assembly E3R01	This study	GCA_003318185.1
Whole genome shotgun sequence assembly D2R18	This study	GCA_003318235.1
Whole genome shotgun sequence assembly A3M17	This study	GCA_003318255.1
Whole genome shotgun sequence assembly F3R11	This study	GCA_003318275.1
16 s rRNA short read sequencing	This study	PRJNA478695
Oligonucleotides		
Primer: 8F: AGAGTTTGATCCTGGCTCAG	[26]	N/A
Primer: 1492R: GGTTACCTTGTTACGACTT	[26]	N/A
Primer: 515F GTGCCAGCMGCCGCGGTAA	[27]	http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/
Primer: 806R GGACTACHVGGGTWTCTAAT	[27]	http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
RDP 16	[28]	https://rdp.cme.msu.edu/classifier/classifier.jsp
EVOS Software revision 31201	Thermo-Fisher	N/A
Fiji version 2.0.0-rc-68/1.52 g	[29]	N/A
CLC Genomics Workbench 11	QIAGEN	N/A
CheckM	[30]	N/A
RAST	[31]	N/A
CAZY	[32]	N/A
dbCAN2	[33]	N/A
msConvert	[34]	N/A
MAVEN	[35]	N/A
DADA2 Big Data:Paired end workflow	This paper, also based on http://benjjneb.github.io/dada2/bigdata.html	https://github.mit.edu/josephe/dada2_pipeline
R Bioconductor dada2 package	[17]	https://www.bioconductor.org/packages/release/bioc/html/dada2.html
NCBI BLAST 2.7.1+.	[36]	N/A
SILVA SINA aligner 1.2.11	[37]	http://www.arb-silva.de/aligner/
FastTree 2.1	[38]	N/A
iTOL	[39]	N/A
R	https://www.r-project.org/about.html	N/A
R package 'cluster' 2.0.7-1	https://cran.r-project.org/web/packages/cluster/index.html	N/A
Other		
Falcon 100 μm cell strainer	Corning	Cat#352360
Millex-SV 5μm filter unit	EMD Millipore	Cat#SLSV025LS
Whatman GF/D filter	Sigma-Aldrich	Cat# WHA1823042
4-40 tap	McMaster-Carr	Cat#2522A715
1" 4-40 screws	McMaster-Carr	Cat#94613A115
30 gauge blunt needle	McMaster-Carr	Cat#75165A31
EVOS FL Auto Imaging System	Thermo-Fisher	Cat# AMAFD1000
4x Plan Fluor objective NA 0.13, WD 19.7 mm	Thermo-Fisher	Cat#AMEP4622
20x LPlan Ph2 objective NA 0.40, WD 3.1 mm	Thermo-Fisher	Cat#AMEP4682
Bond Elut PPL cartridges	Agilent	Cat#12255002
Acquity HSS T3 column	Waters	Cat# 186005614

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Otto X. Cordero (ottox@mit.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Seawater sampling and incubation

Coastal ocean surface water samples were collected in 2015 from Canoe Beach, Nahant, Massachusetts, USA; 42° 25'11.5''N, 70° 54'26.0''W. For each particle type, we set up triplicate 800 mL seawater incubations with model particles, using 1 L wide-mouth Nalgene bottles. Particles stored in artificial sea water (ASW; a sterile-filtered solution of 40 g/L Sea salts) with 20% ethanol, were washed twice with ASW to remove the ethanol and inoculated at a concentration of 100 particles per mL. Bottles were rotated overhead at room temperature and a speed of 7.5 rpm for 10 days. At t = 0, 12, 24, 36, 48, 60, 72, 108, 132, 156, 180, 204 h, 10 mL (~1000 particles) were sampled from each replicate incubation. Particles collected by magnetic separation for DNA sequencing and isolation.

Isolation of bacteria attached to particles

After 1.5, 3.5 and 6.5 days of incubation, particles were sampled, separated from the sea water and washed as described above and split into 1:1, 1:10 and 1:100 dilutions in ASW. Dilutions were vortexed for 20 s and plated using sterile glass beads on 1.5% Bacto agar plates with Marine Broth 2216 or Tibbles-Rawling minimal media as described previously [5] with carbon sources specific for the particle type: 0.05% low viscosity alginate, 0.04% carrageenan, 0.1% glucosamine, or with no added carbon source (to screen for agarose-degrading microbes). Following two days of incubation at room temperature (defined as 21-23°C), at least 16 colonies per particle and plate carbon source type were picked and re-streaked twice on Marine Broth 2216 1.5% Bacto agar plates for purification. To obtain stocks, purified isolates were grown in deep-well plates with liquid Marine Broth 2216 for 48 h, shaking at 300 rpm at room temperature. The liquid culture was frozen at −80°C for further characterization. Isolates were genotypes using the 16S rRNA gene (see below for details). Taxonomic classification was done using the 16S rRNA classifier algorithm hosted by the RDP database [28].

Crossfeeding experiments

The alginate-degrading strain psychB3M02 was streaked on Marine Broth 2216 1.5% Bacto agar plates and incubated at 25°C. After 48 h single colonies were picked and grown in liquid Marine Broth at 25°C. After 48 h, cells were pelleted and washed with Tibbles-Rawling minimal media twice. PsychB3M02 cells were then transferred at a starting OD of 0.005 to Tibbles-Rawling minimal media with 0.15% low viscosity alginate as the sole carbon source, and incubated in 10 mL volumes at 20°C and with overhead rotation. After 24h, the spent media was harvested by gently pelleting the cells (3000 *rcf.* for 10 min) and filtering the supernatant through a 0.2 μm syringe filter. The five alginate non-degraders were pre-grown and harvested in a similar manner and transferred to fresh raw spent media at a starting OD of 0.005 in 200μl volumes. Growth was measured using OD600 on a BioTek Synergy2 microplate reader.

METHOD DETAILS

Particle fabrication

Agarose particles

1.5 g Omnipur agarose were added to 100 mL Milli-Q deionized water in autoclaved glassware and stirred with a magnetic stirrer while being heated on a hot plate until boiling. Immediately upon complete dissolution of the Omnipur agarose, 10mL of this solution was rapidly mixed with 0.5 mL 10% magnetite solution (1 g synthetic black iron oxide washed 3X water with 50 mL Milli-Q water and resuspended in 10mL Milli-Q water) and then added to a 125 mL Erlenmeyer flask containing 50 mL of hot (100°C), rapidly stirring (500 rpm 1'' Teflon stir bar) mineral oil with 1% SPAN-80. After 2 min of rapid stirring, an emulsion formed and the stir bar was rapidly removed at the same time the flask was removed from the hot plate and the transferred to ice, where it was gently agitated for the following 5 min to keep the droplets from settling and merging before the agarose hardened. The flask was then transferred to a 4°C refrigerator for 30 min. Excess oil was then decanted, taking care to retain the particles at the bottom of the flask. The remaining oil and particles (~20 mL) were transferred to a 50 mL conical centrifuge tube. ASW was added to bring the volume to 50 mL and the conical tube was mixed by vortexing and then gently centrifuged at 250 *rcf.* for 5 min. The oil was again decanted and ASW was added to bring the volume up to 50mL. This process was repeated twice more. Sometimes the particles would clump and settle at the water-oil interface; when this happened, the clump was re-dispersed by a combination of vortexing and rapidly pipetting up and down. Once the oil was entirely removed, particles were collected on top of a Falcon 100 μm cell strainer, washed with ASW, and the filtrate was discarded. Particles were then resuspended in ASW with 20% EtOH and stored at 4°C until use.

Alginate particles

1.5 g medium viscosity alginate was dissolved in 100 mL Milli-Q water. This solution was filtered first with a Whatman GF/D filter and then a Millex-SV 5μm filter unit. 0.5 mL of 10% magnetite solution (as described in the agarose particle section above) was mixed with 10mL filtered medium viscosity alginate solution. This solution was then sprayed through a self-manufactured droplet sprayer (see ''Droplet sprayer'' section below) into a bath of sterile-filtered 100mM $CaCl_2$ over several hours. As with agarose particles, alginate particles were retained and washed on a 100 μm filter, and then stored in ASW with 20% EtOH at 4°C until use.

Carrageenan particles

1.5 g Carrageenan and 0.5 g KCl were added to 100 mL Milli-Q deionized water in autoclaved glassware and heated on a hot plate until boiling. Particles were then synthesized as with agarose, except that a 2% KCl solution was used instead of ASW to break the emulsion and wash the beads. Carrageenan particles were stored as with agarose particles in ASW with 20% EtOH at 4°C.

Mixed-type particles

Agarose-alginate and agarose-carrageenan particles were made using the aqueous solution in mineral oil emulsion as described for the agarose particles above. 5 mL of hot agarose solution were mixed with 5 mL of pre-warmed agarose and carrageenan solution, respectively, and 0.5 mL 10% magnetite solution (same reagents and concentrations as for the pure substrate particles). Excess mineral oil was removed and particles were stored as described above for the agarose particles. Particles were stored in ASW with 20% EtOH at 4°C prior to use.

Chitin particles

Chitin magnetic beads were commercially obtained. The concentration of beads was estimated as roughly $2.5x10^5$ beads per mL. The storage buffer for the beads was changed to ASW containing 20% ethanol. To change the buffer, particles were washed three times with ASW by using a magnetic pulldown. Beads in 20% EtOH were stored at 4°C prior to use.

Droplet Sprayer

We constructed a droplet sprayer by modifying a previously published design [40]. A small hole was drilled in the base of a 50 mL Falcon conical centrifuge tube, and two holes were drilled in the lid. Four 1/16" holes were drilled at 90° angles to one another in the sides of the tube at both the 20 mL and 40 mL marks. These side holes were tapped with a 4-40 tap and 1" 4-40 screws were screwed into them. A separate 15 mL Falcon conical centrifuge tube was cut at approximately the 11 mL mark and the top discarded. A small hole was drilled in the tip of the 15 mL tube. A 30 gauge blunt needle attached via a luer-to-barb connector to 1/16" ID flexible tubing was glued (using a hot glue gun) inside the cut off 15 mL tube such that the needle protruded from the hole in the tip of the Falcon tube as far as possible. The 15 mL Falcon tube was then placed inside the 50 mL Falcon tube, such that the tip of the needle was roughly at the center of the hole drilled in the tip of the 50 mL tube. The screws in the 50 mL Falcon tube were then tightened to affix the 15 mL Falcon tube (and thus the needle) in place. The lid of the 50 mL Falcon tube was threaded over the 1/16" ID flexible tubing leading to the needle, screwed in place (sealing the 50 mL Falcon tube), and hot glue was applied around the hole to hold the tube in place. A barb-to-screw adaptor was threaded into the second hole in the lid. A tube connecting this barb to a pressure source, an air pump, was used to provide air flow around the needle at the nozzle.

Particle characterization

Particles were counted and characterized using phase contrast micrographs in a 96-well plate and an EVOS FL Auto Imaging System. We used the 'scan' feature from the EVOS software and a 4x Plan Fluor objective. To obtain shape and size characteristics, we manually outlined particles in micrographs in Fiji [29] as ellipsoids and calculated the diameter of a circle with equal area. Concentrations of the particle stocks were adjusted with artificial sea water with 20% ethanol to be approximately 10^4 particles per mL.

Staining and microscopy

Colonized particles of t = 132 h were stained after fixing with formalin by adding SYTO9, 500 nM final concentration and incubating for 1 h at room temperature in the dark. Microscopy was performed on an EVOS FL Auto Imaging System using a GFP light cube with emission 525/50 and excitation 470/22, a 20x LPlan Ph2 objective and EVOS software.

DNA Extraction

DNA was extracted from the frozen particle samples with the MasterPure DNA purification kit with a few adjustments to the manufacturer's protocol: each sample (500 μl) was mixed with 250 μL of T&C Lysis Solution and 1 μL Proteinase K and incubated at 65°C for 1h before being put on ice for 5 min. Lysate was mixed with 250 μL of MPC Protein Precipitation Reagent and treated as in the original protocol. Supernatant was transferred into a clean tube containing 0.5 μl glycogen. 500 μl 100% isopropanol were added and samples kept at −80°C for 12 h. After precipitation and ethanol washing, the pellet was resuspended in 30 μL ddH2O.

16S rRNA amplicon library preparation

Relative DNA concentrations were quantified by performing qPCR with primers 515F and 806R, which are designed to amplify the V4 region of 16S rRNA. 20 μl reactions contained 10.6 μl ddH_2O, 4 μl 5x HF buffer, 0.4 μl 10 mM dNTPs, 1 μl 3μM 515F, 1 μl 3μM 806R, 0.8 μl 20 μg/μl BSA, 2 μl DNA, 0.2 μl Phusion polymerase, and 0.02 μl of 200x SYBR Green. The cycling conditions were a denaturation (30 s at 98°C) followed by 40 amplification cycles (10 s 98°C, 60 s at 50°C, 90 s at 72°C), and a final 10 min extension at 72°C. Samples were diluted so that all concentrations were equivalent to a Ct of 22. The normalized samples were then used to prepare 16S rRNA amplicon libraries. A 10 cycle pre-amplification was performed. 20 μl reactions were prepared as above, omitting SYBR Green. The pre-amplification cycling conditions the same as those used for qPCR above, but only 10 amplification cycles were used. The pre-amplified rRNA was purified using Ampure XP SPRI beads. After the initial amplification of the 16rRNA gene, the barcoding of samples and subsequent amplification was performed by the Whitehead Institute for Biomedical Research (MIT, Cambridge, MA, U.S.A.). After library preparation, all samples were sequenced on a HiSeq 2500 with 2x250 paired-end format.

Genome sequencing

For selected isolates from our collection, genomic DNA was extracted from a liquid overnight culture in Marine Broth 2216 using the Agincourt DNA Advance kit. Genomes were sequenced using the Nextera XT DNA Library Preparation and index kits. Sequencing was performed on an Illumina HiSeq 2500 (250x250 bp paired-end reads) at the Whitehead Institute for Biomedical Research (MIT, Cambridge, MA, U.S.A.). Genomes were assembled using CLC Genomics Workbench 11, and curated using CheckM [30]. Open reading frames were annotated using the RAST pipeline [29] and additional annotation of carbohydrate-associated genes was performed using the CAZY database [32], run from the dbCAN2 server [33].

Metabolomics

Metabolomics was performed at the Microbial Biogeochemistry Group at the Woods Hole Oceanographic Institution. To extract the metabolites from the spent media, the filtrate was acidified to a pH ~3 using 12 M hydrochloric acid and the extracellular organic compounds extracted using Bond Elut PPL cartridges (1 g/6 mL sized cartridges) following the protocol of Dittmar [41] as modified by Longnecker [42]. Dissolved organic matter was eluted from the cartridges using 100% methanol. The resulting organic matter extracts were analyzed using targeted mass spectrometry. Briefly, the extracts for targeted analysis were re-dissolved in 95:5 (v/v) water:acetonitrile with deuterated biotin (final concentration 0.05 mg ml^{-1}). Samples were then analyzed by ultra-performance

liquid chromatography using an Accela Open Autosampler and Accela 1250 Pump coupled to a heated electrospray ionization source (H-ESI) and a TQS Vantage triple quadrupole mass spectrometer operated under selected reaction monitoring (SRM) mode. Chromatographic separation was performed on a Waters Acquity HSS T3 column (2.1 × 100 mm, 1.8 μm) equipped with a Vanguard pre-column and maintained at 40°C. The column was eluted with (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile at a flow rate of 0.5 mL min^{-1}. The gradient was programmed as follows: start 1% B for 1 min, ramp to 15% B from 1-3 min, ramp to 50% from 3-6 min, ramp to 95% B from 6-9 min, hold until 10 min, ramp to 1% from 10-10.2 min, and a final hold at 1% B (total gradient time 12 min). Separate autosampler injections of 5 μl each were made for positive and negative ion modes.

The samples were analyzed in a random order with a pooled sample run after every six samples. A set of target compounds was analyzed: 2-3-dihydroxybenzoic acid, 2-3-dihydroxypropane1sulfonate, 3-mercapto proprionate, 3, methyl-2-oxopentanoic acid, 4-aminobenzoic acid, 4-hydroxybenzoic acid, 4-methyl-2-oxopentanoic acid, 5′-Deoxy-5′(methylthio)adenosine, 6-phosphogluconic acid, D-glucosamine 6-phosphate, DMSP, GABA, HET, HMP, NAD pos, NADP, acetyltaurine, adenine pos, adenosine, adenosine 5′-monophosphate pos, alpha-ketoglutaric acid, arginine, aspartic acid pos, betaine, biotin pos, caffeine, choline, ciliatine, citric acid, citrulline, cyanocobalamin, cytosine, desthiobiotin pos, dihydroxyacetone phosphate, ectoine, folic acid neg, glucose6P, glutamic acid, glutamine, glyphosate, guanine, guanosine neg, hemin I, indole 3-acetic acid, inosine 5′-monophosphate neg, isethionic acid, isoleucine, l-glutathione, l-glutathione oxidized pos, l-kynurenine, l-leucine, l-tyrosine, malic acid, methionine, muramic acid, n-acetyl glutamic acid, n-acetyl muramic acid neg, orotic acid, pantothenic acid neg, phenylalanine, phycocyanobilin, proline, pyridoxine, riboflavin, s-(5′-adenosyl)-L-homocysteine neg, sarcosine, serine, sn-glycerol 3-phosphate neg, sn-glycerol 3-phosphate pos, sodium taurocholate, succinic acid, syringic acid, taurine, thiamine, threonine (isom. homoserine), thymidine, tryptamine, tryptophan, uracil pos, valine, xanthine neg, and xanthosine neg. Metabolites detected are listed in Table S2. The mass spectrometer was operated in selected reaction monitoring (SRM) mode; optimal SRM parameters (s-lens, collision energy) for each target compound were optimized individually using an authentic standard [43]. Two SRM transitions per compound were monitored for quantification and confirmation. Eight-point external calibration curves based on peak area were generated for each compound. The resulting data were converted to mzML files using the msConvert tool [33] and processed with MAVEN [35].

16S amplicon data analysis

We developed a pipeline based on the DADA2 Big Data:Paired end workflow. Briefly, a parametric error model is learned from the sequencing data, using a subset of two million reads drawn randomly from all those sequenced. Then, this error model is used to "denoise" samples by identifying erroneous sequence variants and combining them with the sequence variant from which they most likely originated. All other read processing steps–including merging paired-end reads, trimming primer sequences, and dereplicating reads–were performed with functions from the R Bioconductor "dada2" package [17].

For our analysis, we focused on the abundant ASVs, defined as those with a frequency > 1% in at least one sample across all samples, including replicates, time points and single substrate particle types. The resulting 107 ASVs were used throughout our analysis. Replicates were combined by calculating the weighted average frequency for every ASV, using the read counts of that sample as weights. We smoothed the data with a running median filter, window size = 3 and renormalized to work with mean frequencies.

Mapping isolates to amplicon sequencing data

To map our cultured isolates to the culture independent amplicon sequencing data, we matched the 250 bp exact sequence variants (ASV) (16S rRNA-gene bp 515 – 806, 250 bp read) to the 16S sequences of our isolates (16S rRNA gene bp 27 – 1492, ~1000 bp read). We made a NCBI BLAST 2.7.1+ database [36] from our ASV table as produced by DADA2 using *makeblastdb* and *-dbtype nucl*. Next, we BLASTed all isolates against the ASV database using *blastn* with options *-max_target_seqs 1 -perc_identity 100 -outfmt 6.* We searched for ASVs with a 100% identity for each isolate.

Mixed substrate particle linear model fitting

We fit a linear model to predict the composition of the community on the mixed substrate particle ASVs uing linear combinations of the single substrate particle compositions. We used the R function *lm*. The model has the form $Y = \alpha^*X_1 + \beta^*X_2$, where Y is the vector of ASV geometric mean abundances on the mixed polysaccharide particle, X_1 and X_2 are the vector of geometric mean abundances on the single polysaccharide particles and α, β are the fitted coefficients. The following tables contain the fitting summary statistics.

Model Agarose – Alginate

α * geometric mean of ASVs on agarose + β * geometric mean of ASVs on alginate
Residuals: Min = −0.0082277; 1Q: −0.0003486, Median = −0.0000122; 3Q = 0.0001198; Max = 0.0172305
Coefficients for α *:* Estimate = 0.67365; Standard Error = 0.05366; t-value = 12.554; Pr(> |t|) = < 2e-16
Coefficients for β: Estimate = 0.39864; Standard Error = 0.06638; t-value = 6.005; *Pr(> |t|)* = 2.78e-08
Residual standard error: 0.003012 on 105 degrees of freedom
Multiple R-squared: 0.8439, Adjusted R-squared: 0.8409
F-statistic: 283.7 on 2 and 105 DF, p value: < 2.2e-16

Model Agarose – Carrageenan

α * geometric mean of ASVs on agarose + β * geometric mean of ASVs on carrageenan
Residuals: Min = −0.0180190; 1Q = −0.0003712; Median = −0.0000297; 3Q = 0.0001741, Max = 0.0202486
Coefficients for α: Estimate = 0.89530; Standard Error = 0.06171; t-value = 14.507; Pr(> |t|) = < 2e-16
Coefficients for β: Estimate = 0.11393; Standard Error = 0.04755; t-value = 2.396; Pr(> |t|) = 0.0183
Residual standard error: 0.004318 on 105 degrees of freedom
Multiple R-squared: 0.742, Adjusted R-squared: 0.7371
F-statistic: 151 on 2 and 105 DF, p value: < 2.2e-16

We compared these linear fits with a model containing an interaction term to establish to what extent simple linear models where sufficient to describe the data. The new model reads $Y = \alpha^*X_1 + \beta^*X_2 + \gamma^* X_1 *X_2$, where γ is the fitted interaction coefficient. We used model selection based on the Akaike Information Criterion (AIC) to substantiate the choice of model. For agarose-alginate particles the simple model without interaction had a similar AIC (−932.8) but the p value of the fitted coefficient (γ) was not significant (0.665). For the agarose-carrageenan case the model with an interaction term had a lower AIC (−864) than the model without interaction (−857) and the interaction term was significant (p value 0.00261). However, this model introduces an extremely weak non-linearity, as the nonlinear model (nlm) fits the linear model (lm) according to $nlm = 0.16^*lm^{0.986}$, with R^2 0.9995. Therefore, the composition of the community on mixed particles is well described by a simple linear model.

QUANTIFICATION AND STATISTICAL ANALYSIS

Phylogenetic tree of ASVs

To create a phylogenetic tree of the top 1% ASVs, we first aligned the 16S V4V5 sequences on Silva's SINA alignment server with standard settings, the option *Search and Classify* enabled with *minimum identity with query sequence* = 0.9 and *classification: rdp*. After removing non-informative positions from the alignment we used FastTree [38] with the options -gtr -n) to infer an approximate maximum-likelihood phylogenetic tree which we formatted using iTol [38].

Niche width index

To study the prevalence of each ASV across different particle types, we devise a niche width index. We calculated the geometric mean frequency ASV on a particle type, $f_{ij} = e^{<\log(f_{ij}(t))>}$, $f_i = e^{<\log(f(t))>}$ $f_i = e^{<\log(f(t))>}$ where $f_{ij}(t)$ is the frequency of ASV *i* at time *t* on particle type *j*. We added pseudo counts $(10^{-6})$to $f_{ij}(t)$ to account for zeroes. With the normalized geometric mean frequencies, $g_{ij} = f_{ij}/\sum f_{ij}$ we calculated a niche width index over *j* using the entropy: $-\sum g_i \log2(g_i)$. We use the R function Mclust to group our ASVs into three optimal groups according to their niche width index. The niche width index cutoff values for the groups are < 0.18 and > 1.52. The three resulting groups have 38, 24 and 45 members, respectively.

Hierarchical clustering of ASV trajectories

We cluster the most abundant ASVs based on their log-transformed frequencies across all time-points and all particle types. We used the R function *hclust* with the clustering method *'ward.D'* and Euclidean distances. To evaluate the best cutoff for our hierarchical clustering, we cut the tree generated by 'hclust' into 2-15 groups using the 'cutree' function in R. We used the silhouette function from the R package 'cluster' to evaluate the clusters generated. Our analysis shows that 5 clusters are the optimal partitioning of our data.

DATA AND SOFTWARE AVAILABILITY

The short read 16 s rRNA amplicon sequencing data (BioSample: SAMN11023523-SAMN11023755), and 7 genomes (GenBank: GCA_003318155.1, GeneBank: GCA_003318165.1, GeneBank: GCA_003318175.1, GeneBank: GCA_003318185.1, GeneBank: GCA_003318235.1, GeneBank: GCA_003318255.1, GeneBank: GCA_003318275.1) have been deposited in NCBI with the BioProject identifier PRJNA478695.

Current Biology, Volume *29*

Supplemental Information

Modular Assembly of Polysaccharide-Degrading Marine Microbial Communities

Tim N. Enke, Manoshi S. Datta, Julia Schwartzman, Nathan Cermak, Désirée Schmitz, Julien Barrere, Alberto Pascual-García, and Otto X. Cordero

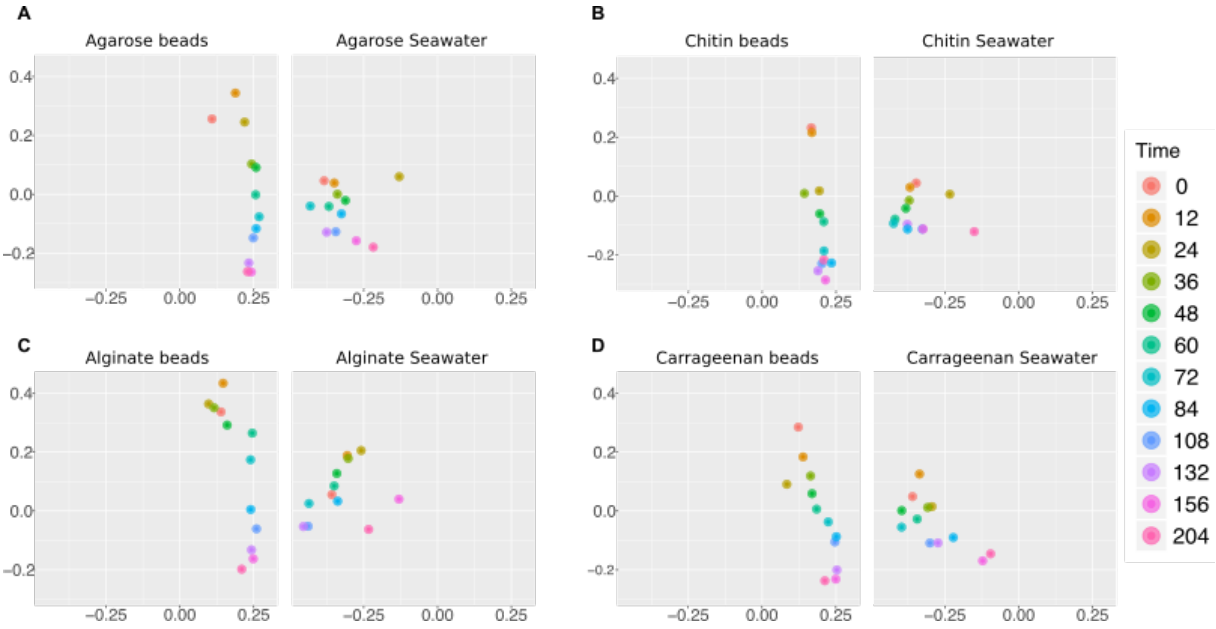

Figure S1. **Comparison of the community assembly trajectories on beads and on the seawater sampled around the corresponding beads, described by a PCoA of the Bray Curtis dissimilarity between timepoints, related to Figure 1B.** Data analyzed were those used for analyses in Figures 1B and 2A. The plot is divided in four panels, one for each particle type. The figure shows that the bead communities and the seawater communities follow very distinct trajectories, orthogonal in PCoA space to the trajectories followed by the particle attached communities. This indicates that although there is some coupling between particle-attached and free-living fractions, the particle-attached community has dynamics that are independent from the seawater.

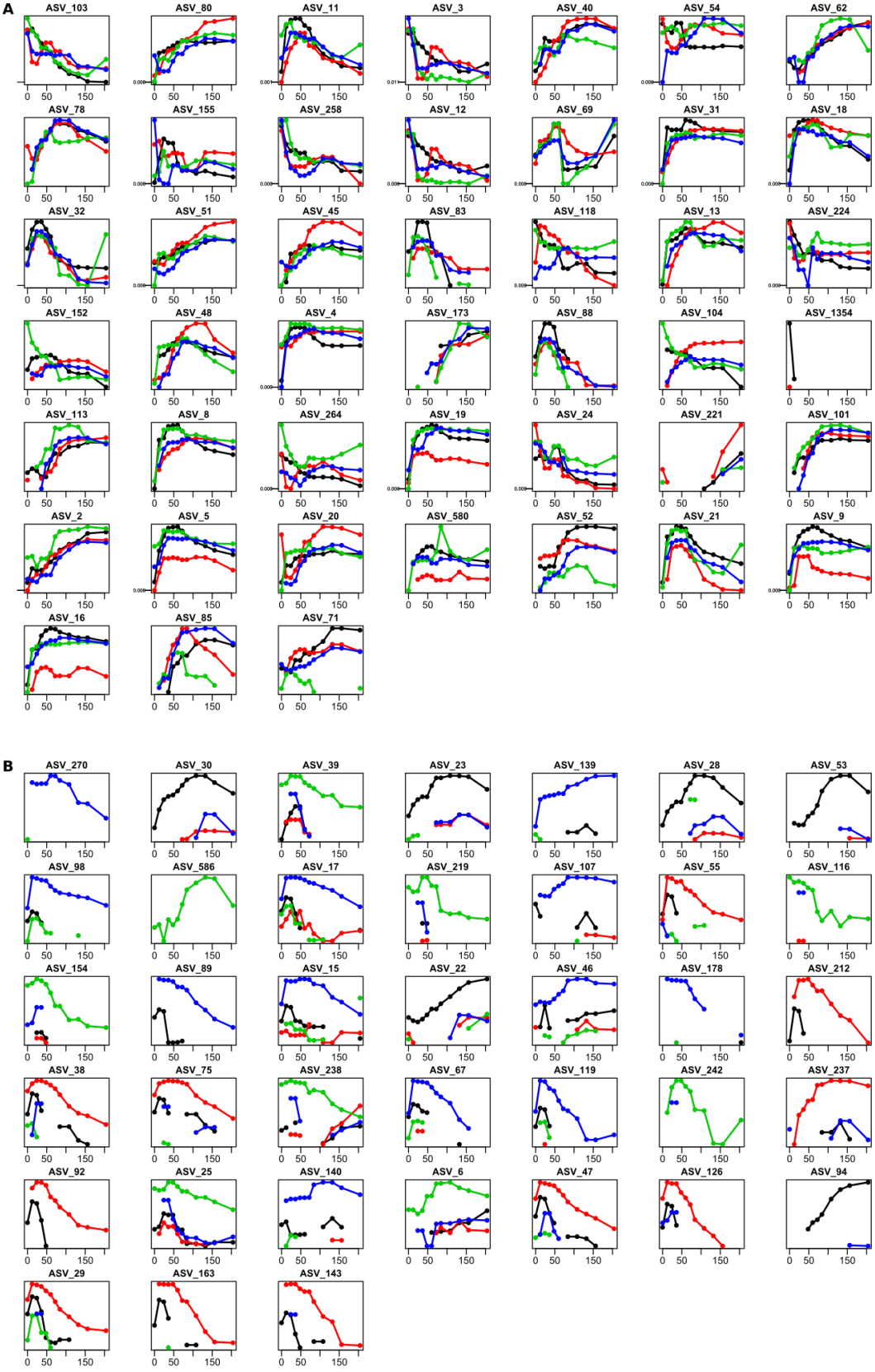

Figure S2 Time dynamics of most abundant broad-range (A) and narrow-range (B) taxa, related to Figure 2B-C. Trajectories are plotted for individual ASVs assigned to clusters in Figure 2A. Green: chitin, blue: carrageenan, red: alginate, black: agarose.

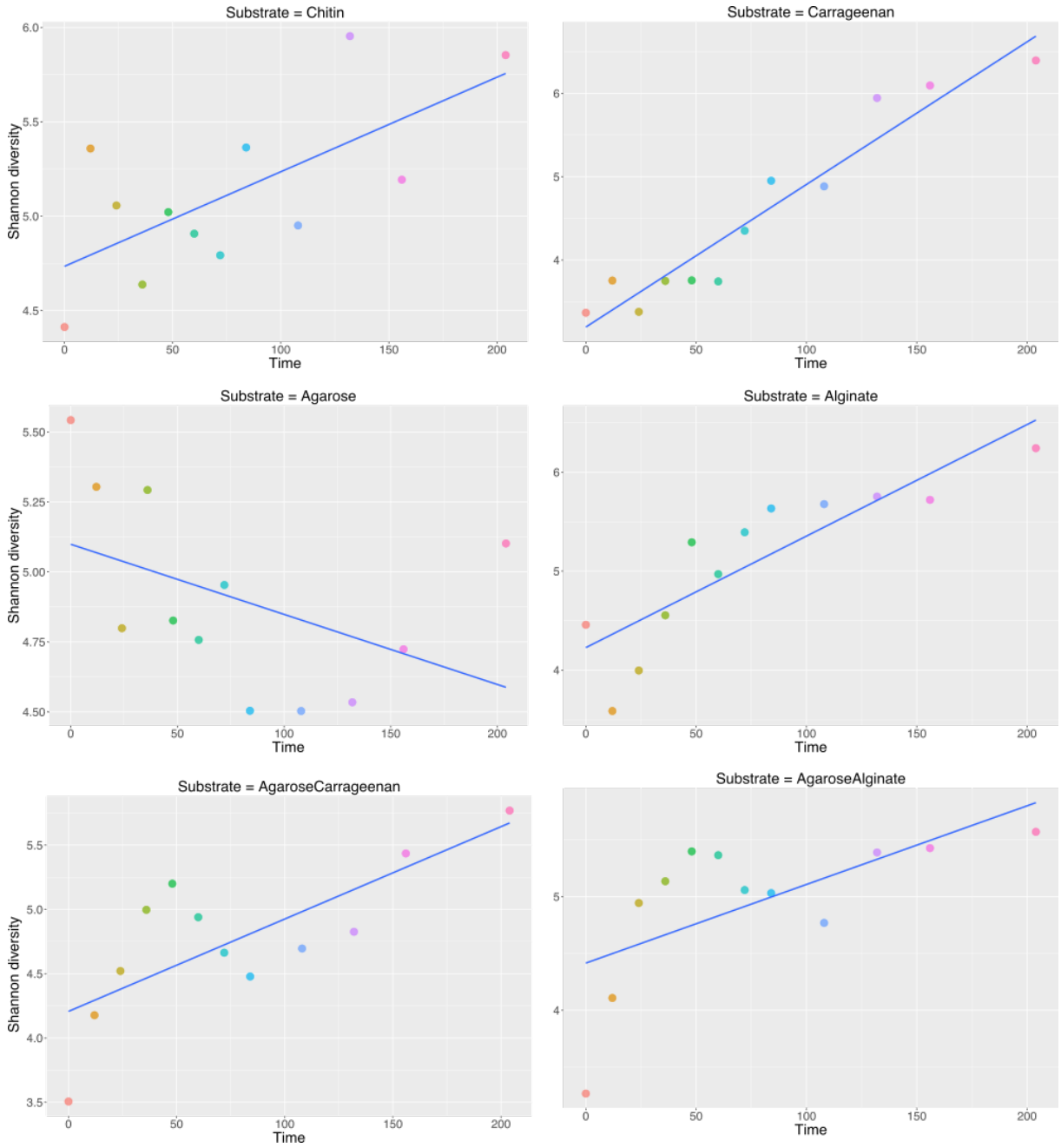

Figure S3. Shannon diversity during successional dynamics for each particle type, related to Figure 2B-C. To compute the biodiversity of the samples we first rarefied all samples to 1000 reads, in order to avoid potential artifacts arising from the different sampling depths. We aggregated the three replicates at each timepoint and computed the Shannon diversity index at each time point. The plots show a linear regression of the Shannon index against time to illustrate the mean trend, although in most cases (perhaps with the notable exception of carrageenan) the diversity trends were nonlinear suggesting different phases of succession. To facilitate comparisons across the different substrates, dots are colored according to timepoint.

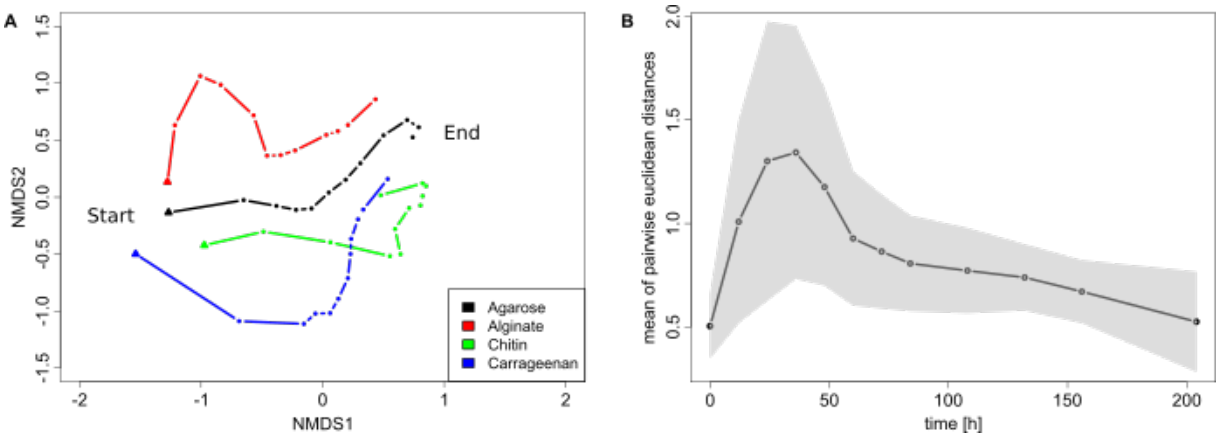

Figure S4. Convergence in community composition across particle types, related to Figure 2.B-C A. Non-dimensional scaling visualization of the community composition (log-transformed frequencies) across the different particles over time. The trajectories show an initial divergence due to the invasion of narrow range taxa, followed by a convergence due to the invasion of broad range taxa. This trend is shown also in panel B), which measures the mean pairwise distance between communities (most abundant ASVs) in NDMS space (2 dimensions). Shaded area shows one standard deviation around the mean.

Strain	Genus	Alginate			Carrageenan/Agar			Chitin		
		1	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
B3M02	*Psychromonas*	8 (PL7)	4 (2 PL15, 2 PL17)	+	-	-	+	-	2 (GH3)	+
E3R01	*Tenacibaculum*	-	-	-	-	-	+	2 (GH18, GH19)	1 (GH20)	+
C3R12	*Vibrio*	-	-	-	1 (GH16)	-	+	3 (2 GH18, 1 GH19)	1 (GH20)	+
C3M10	*Phaeobacter*	-	-	+	-	-	-	-	-	-
A3M17	*Rugeria*	-	-	+	-	-	-	-	2 (1 GH3, 1GH20)	+
D2R18	*Loktanella*	-	1 (PL15)	+	-	-	-	-	1 (GH20)	+
F3R11	*Marinobacter*	-	-	-	-	-	-	-	-	-

Table S1. Predicted pathways from sequenced genomes, related to Figure 3. Predictions are based on RAST subsystem maps EC #s are indicated. Extracellular enzymatic steps were predicted using the CAZY classification system. GH, glycosyl hydrolase; PL polysaccharide lyase. Alginate catabolism steps: 1, alginate lyase; 2, oligoalginate lyase; 3 KDG to pyruvate (EC 4.1.2.14, EC 2.7.1.45). Carageenan and Agar catabolism steps: 1, agarase (3.2.1.81, 3.2.1.158) and carrageenase (EC 3.2.1.83, EC 3.2.162, EC 3.2.1.157); 1, AHG 3,6-anhydrogenase to galactose (EC 1.2.1.92, EC 5.5.1.); 3, galactose utilization (EC 5.1.3.3, EC 2.7.1.6, EC 2.7.7.10, EC 5.1.3.2). Chitin catabolism steps: 1, secreted chitinase; 2, periplasmic oligo-chitinase; 3 GlcNAc to fructose 6-phosphate (EC 3.5.99.6, EC 3.5.1.25).

	F3R11			D2R18		
	mean % consumed	**sd % consumed**	**replicates* (out of 3)**	**mean % consumed**	**sd % consumed**	**replicates* (out of 3)**
tryptophan	100	0	3	66.67	57.73	3
succinic acid	60.92	53.46	3	63.37	17.97	2
proline	86.68	11.68	3	79.77	18.53	3
malic acid	100	0	2	93.26	9.54	2
l.tyrosine	78.57	30.30	2	85.71	20.20	2
l.leucine	50	70.71	3	100	0	2
ectoine	100	0	2	100	0	2
choline	91.67	14.43	3	100	0	3
aspartic acid	100	0	2	100	0	2
adenosine	0	0	3	100	0	2
3-methyl-2-oxopentanoic acid	100	0	2	50	70.710678	2

Table S2. Metabolites produced by psychB3M02 and consumed by F3R11 and D2R18 in at least two out of three biological replicates, related to Figure 3B. The column replicates indicate the number of replicates in which the metabolite was produced. The mean and standard deviation are calculated over those cases. A list of the metabolites targeted in this analysis is found in the STAR Methods.

Agarose-Alginate	**Mean correlations**	**Median correlations**
narrow aga vs. mix	0.780	0.865
narrow alg vs. mix	0.945	0.964
broad on aga vs. mix	0.681	0.810
broad on alg vs. mix	0.674	0.748
Agarose-Carrageenan		
narrow aga vs. mix	0.492	0.652
narrow car vs. mix	0.671	0.720
broad on aga vs. mix	0.673	0.741
broad on car vs. mix	0.676	0.864

Table S3. Spearman correlations between the temporal dynamics of ASVs on single substrate and mixed substrate particles, related to Figure 4. The data is divided by narrow and broad range taxa, showing that both groups have similar dynamics on single and mixed substrate particles.