Core species and interactions prominent in fish-associated microbiome dynamics

Daii Yajima[1], Hiroaki Fujita[1], Ibuki Hayashi[1], Genta Shima[1], Kenta Suzuki[2] and Hirokazu Toju[1†]

[1]Center for Ecological Research, Kyoto University, Otsu, Shiga 520-2133, Japan

[2]Integrated Bioresource Information Division, BioResource Research Center, RIKEN, Tsukuba, Ibaraki 305-0074, Japan

[†]**Correspondence:** Hirokazu Toju (toju.hirokazu.4c@kyoto-u.ac.jp).

This article includes 6 Figures, 10 Extended Data Figures, and 1 Supplementary Table 1.

Abstract

In aquatic ecosystems, the health of fish depends greatly on the dynamics of microbial community structure in the background environment. Nonetheless, finding microbes with profound impacts on fish's performance out of thousands of candidate species remains a major challenge. We here show that time-series analyses of microbial population dynamics illuminate core components and structure of fish-associated microbiomes. By targeting eel aquaculture microbiomes as model systems, we reconstructed the population dynamics of 9,605 bacterial and 303 archaeal species/strains across 128 days. Due to the remarkable increase/decrease of constituent microbial populations, the taxonomic compositions of microbiomes changed drastically through time. We then found that some specific microbial taxa showed positive relationship with eels' activity level even after excluding cofounding effects of environmental parameters (pH and dissolved oxygen level) on population dynamics. In particular, a vitamin B_{12}-producing bacteria, *Cetobacterium somerae*, consistently showed strong positive associations with eels' activity level across the replicate time-series of the five aquaculture tanks. Network theoretical and metabolic modeling analyses further suggested that the highlighted bacterium formed compartments of close microbe-to-microbe interactions with some other bacterial taxa, forming potential core microbiomes with positive impacts on eels. Overall, these results suggest that integration of microbiology, ecological theory, and network science allows us to explore core species and interactions embedded within complex dynamics of fish-associated microbiomes.

Microbial communities are essential factors of the life of vertebrates[1–4], playing key roles in the development and homeostasis of their hosts[5–7]. Gut microbiomes, for example, play key roles in the nutrition and disease prevention of human and other mammal species[8,9]. Such physiological and ecological effects of gut microbes on hosts have been reported as well for fish[6,10,11]. Meanwhile, because fish are continuously exposed to numerous pathogenic and non-pathogenic microbial species in the water, their performance (or fitness) depends not only on gut-associated microbes[6,10] but also on the microbiomes of the background environment[12–14]. Therefore, finding key microbiome components whose dynamics determine fish's health or performance is of interdisciplinary interest spanning from microbiology to zoology and environmental science. However, due to the tremendous diversity of bacteria and archaea in aquatic ecosystems[15,16], exploring such core microbial species associated with fish health remains a challenge.

A starting point for finding fish-health-associated microbes in aquatic ecosystems is to track the dynamics of microbial community compositions. Nonetheless, we still have limited knowledge of the extent to which structure of fish-associated microbiomes change through time. Although time-series data of microbiomes have become available in pioneering projects of human-associated microbes[17,18], few attempts have been made to monitor microbiomes associated with other animals over tens of time points. Moreover, continuous sampling of fecal samples of targeted vertebrate individuals is generally much harder in aquatic environments than in terrestrial environments. Thus, developing model systems for time-series analyses of microbe–fish ecological interactions is a demanding but essential step for exploring core bacteria and/or archaea out of thousands of candidate species in microbial communities.

Despite the hardship in gaining time-series microbiome samples at the individual level, fish-associated microbiome dynamics can be monitored at the population or community level by sampling environmental water samples[19–21]. Because excrements of fish are released to water, samples of background water are expected to reflect gut microbiomes of fish populations or communities. Furthermore, as individual fish are continuously exposed to the background microbiomes, analyses of water samples provide essential insights into surrounding environmental conditions and potential sources of gut microbiomes[12–14]. In this respect, time-series analyses of aquaculture or aquarium systems offer an ideal opportunity for investigating relationship between microbial community structure, core microbial species, and vertebrate health.

By targeting a recirculating aquaculture system of the Japanese eel (*Anguilla japonica*), we herein integrate microbiology, community ecology, and network science for detecting key species and structure within fish-associated microbiomes. Based on the DNA metabarcoding of prokaryote (bacterial and archaeal) communities for the 128-day time-series, we revealed to what extent the compositions of aquaculture microbiomes fluctuate through time. We then reconstructed the population dynamics (i.e., increase/decrease) of the 9,908 microbial amplicon sequencing variants (ASVs) constituting the aquaculture microbiomes, screening bacteria or archaea whose abundance was tightly linked with the health condition of eels. We then found that several microbial ASVs showed positive associations with eel health consistently across the five replicate aquaculture tanks, even after controlling the effects of their environmental preference (e.g., preference to pH and dissolved oxygen level). With the approaches of network science and metabolic modeling, we further examined potential interactions between the core microbes. Overall, this study illustrates how core species and interactions are detected based on time-series datasets of microbiome dynamics.

Results

Microbiome dynamics. Monitoring of microbiome dynamics was conducted targeting the five water tanks of an aquaculture farm of the Japanese eel. In each water tank (diameter = 2.5 m; height = 1 m; volume = 20 m^3), 1,400–4,300 eel individuals (average weight = 80–130 g) had been kept. The pH and dissolved oxygen (DO) concentrations were recorded for each tank every day. In addition, as a measure of ecosystem-level functions of microbiomes, the health condition of eels was evaluated based on eight criteria, yielding eel activity scores on a scale of 0 to 40 (see Methods). For the analyses of microbiome dynamics, water was sampled from each aquaculture tank every 24 hours during 128 days. By applying a quantitative amplicon sequencing approach for estimating 16S ribosomal RNA gene (16S rRNA) copy concentrations of respective microbes[22,23], we obtained time-series datasets representing the increase/decrease of 9,605 bacterial and 303 archaeal ASVs representing 618 genera and 325 families (Fig. 1a). Thus, our data offered a novel opportunity to test synchronizations among microbial population dynamics, environmental factors (pH and DO), and vertebrate performance (eel activity level).

At the community level, drastic taxonomic turnover was observed in the timeseries of each aquaculture tank (Fig. 1; Extended Data Figs. 1-3). In Tanks 1 and 2, for example, the community structure characterized by the predominance of Fusobacteriaceae and

Microbacteriaceae was suddenly altered by a Flavobacteriaceae-dominated state around Day 45 (Fig. 1). Meanwhile, microbiomes of Tanks 3–5 displayed more complex dynamics represented by frequent shifts between Flavobacteriaceae-dominated and Chitinophagaceae-dominated states, although clear classification of community states was difficult (Fig. 1).

A multivariate analysis of the prokaryote community structure indicated that the community state characterized by dominance of Fusobacteriaceae and Microbacteriaceae was associated with high eels' activity (Fig. 2). In contrast, the Flavobacteriaceae-dominated and Chitinophagaceae-dominated states, which were observed in high-pH conditions, were associated with low eels' activity (Fig. 2). At the genus level, the high-eel-activity-related state of dominance by Fusobacteriaceae and Microbacteriaceae was characterized by high relative abundance of *Cetobacterium*, which includes species potentially contribute to fish physiological homeostasis[24]. On the other hand, the Flavobacteriaceae-dominated and Chitinophagaceae-dominated states associated with low eels' activity were represented by *Flavobacterium* and *Edaphobaculum*, respectively (Extended Data Fig. 4). Among the genera, *Flavobacterium* includes fish pathogens[25], while *Edaphobaculum*[26] has been poorly investigated in terms of their effects on fish physiology. These results suggest potential impacts of environmental microbiome dynamics on fish health/behavior in aquaculture systems.

Exploring microbes with key roles. We next evaluated how the population dynamics of each microbial ASV were associated with environmental variables and eels' activity level. Specifically, we examined how population size (absolute abundance) of each ASV varied with pH, DO, and eels' activity level (Fig. 3a) based on correlation analyses with twin-surrogate permutations[27] (Fig. 3b-c). The ASVs varied in their environmental preference for pH and DO conditions as well as in their associations with eels' activity level (Extended Data Fig. 5). We also found that ASVs' relationship with eels' activity level displayed tank-dependent complex associations with pH or DO preference (Fig. 3d-e). Thus, for each microbial ASV in each aquaculture tank, we calculated a partial correlation between absolute abundance and eels' activity scores through the time-series by controlling the effects of pH or DO. Because partial correlation coefficients were consistent between the pH-controlled and DO-controlled calculations (Fig. 3f), the pH-controlled partial correlation coefficients were used in the following analyses.

The partial correlation coefficients with eels' activity level varied greatly depending on prokaryote taxa (Fig. 2e). Nonetheless, ASVs belonging to some bacterial genera showed

consistently positive correlation with eels' activity scores across the five tanks (Fig. 2g; Extended Data Fig. 5). The list of those ASVs included bacteria belonging to the genera *Cetobacterium* (Fusobacteriaceae; Fusobacteriia; ASV ID = X_0002), *Plesiomonas* (Enterobacteriaceae; Gammaproteobacteria; X_0020), *Turicibacter* (Erysipelotrichaceae; Bacilli; X_0041), *Paraclostridium* (Clostridiaceae; Clostridia; X_0014), *Romboutsia* (Peptostreptococcaceae; Clostridia; X_0028), *Edwardsiella* (Hafniaceae; Gammaproteobacteria; X_0027), *Clostridium* (Clostridiaceae; Clostridia; X_0029), and an ASV belonging to Barnesiellaceae (Bacteroidia; X_0064) (Fig. 3g; Extended Data Fig. 5d).

An additional database search of the 16S rRNA sequences suggested that some of the ASVs with positive associations with eels' activity level belonged to bacterial species with potential physiological impacts on fish. For example, the *Cetobacterium* ASV, which showed strongest positive partial correlation with eels' activity level, was represented by the 16S rRNA sequences completely matching that of *Cetobacterium somerae* (formerly recognized as "*Bacteroides* type A") in the NCBI nucleotide database. This *Cetobacterium* species has been known to produce high concentrations of vitamin B_{12} and hence their potential contributions to fish's physiology have been anticipated. Meanwhile, the *Edwardsiella* ASV listed above was allied to the notorious fish pathogen *E. tarda*[28], illuminating paradoxical relationships with eels' health. However, our supplementary phylogenetic analysis based on the *sodB* gene marker[29] indicated that 95.1 % of *Edwardsiella* bacteria detected in the focal eel aquaculture system belonged to non-pathogenic clades[29,30] within the genus *Edwardsiella* (Extended Data Fig. 6).

In terms of negative impacts on eels' activity level, bacteria in the genera *Aeromonas* (Aeromonadaceae; Gammaproteobacteria), *Methylobacterium* (alternatively, *Methylorubrum*; Beijerinckiaceae; Alphaproteobacteria), and *Acinetobactor* (Moraxellaceae; Gammaproteobacteria) were highlighted (Fig. 3g). Among them, *Aeromonas* and *Acinetobactor* have been known to include fish pathogens[31,32]. At the ASV level, an ASV allied to the cvE6 clade within the order Chlamydiales (Chlamydiae; Verrucomicrobiota) showed strongest negative correlation with eels' activity scores (Extended Data Fig. 5d).

Although the above analysis controlling environmental preferences of respective bacteria allows high-throughput screening for species with potential positive/negative impacts on target biological functions, the simple statistical approach with partial correlation analyses precludes insights into the direction of causality. Specifically, it is important to consider the possibility that high/low abundance of an ASV is a consequence but not a cause of eels' high/low activity. Therefore, we performed an additional analysis introducing time lags into

eels' activity scores throughout the time-series. We then found that the abundance of the *Cetobacterium* ASV was positively correlated with eels' activity scores of the next day, while correlations between *Cetobacterium* abundance and past eels' activity scores were much lower than those with no time lags (Fig. 3h). Meanwhile, high correlation between 5-days-ago eels' activity level and present-day *Cetobacterium* abundance was observed in some tanks (Tanks 1 and 4; Fig. 3h), illuminating the importance of carefully interpreting the results of the time-series analysis.

Networks of interactions. We then reconstructed webs of potential microbe-to-microbe interactions to illuminate microbial groups or interactions positively associated with eels' health. We first applied the Meinshausen-Bühlmann (MB) method, which was designed to evaluate patterns of coexistence realized by the effects of microbe–microbe interactions as well as those of niche sharing between microbes. For each aquaculture tank, the reconstructed network of microbe–microbe coexistence (Extended Data Figs. 7–8; Supplementary Table 1) was compartmentalized into several modules, which differed in mean partial correlations with eels' activity scores (Fig. 4). We then found that each of the five networks included a module constituted by the abovementioned *Cetobacterium* ASV and several other ASVs with consistently positive associations with eels' activity level (Fig. 4; Extended Data Fig. 9). The bacteria consistently formed network modules of coexistence with the *Cetobacterium* ASV were *Plesiomonas* (X_0020), *Turicibacter* (X_0041), *Paraclostridium* (X_0014), *Romboutsia* (X_0028), *Edwardsiella* (X_0027), and *Clostridium* (X_0029) (Fig. 4).

To infer the presence/absence of direct interactions between these bacteria with positive relationship with eels' activity level, we conducted an additional network analysis based on the sparse and low-rank (SLR) decomposition method, which allowed us to remove latent effects of environmental conditions. In the networks reconstructed with the SLR method (Fig. 5), potential effects of niche sharing were controlled and hence the links between bacterial ASVs were expected to represent potential positive interactions. The estimated interaction coefficients were highly correlated between the MB and SLR methods (Extended Data Fig. 10). Meanwhile, in the SLR-based network, removing the effects of potential niche sharing (sharing of environmental preference) resulted in the simplification of network structure, in which estimated direct interactions between microbes were focused (Fig. 5). Despite the considerable difference between MB- and SLR-based network topology, the *Cetobacterium* ASV with the strongest associations with eels' activity level was, again, linked with the

Plesiomonas, *Turicibacter*, *Paraclostridium*, *Romboutsia*, *Edwardsiella*, and *Clostridium* ASVs within the SLR network (Fig. 5), suggesting positive interactions with these bacteria.

Potential metabolic interactions. To estimate functional interactions between microbes, we focused on genomic compositions of respective microbes within the aquaculture microbiomes. After retrieving the information of genomic compositions from reference databases, we analyzed the inferred gene repertoires (KEGG metabolic pathway/process profiles) of the microbial ASVs based on multivariate analysis. Along the principal component axes, *Cetobacterium*, which showed consistent correlations with eels' activity (Fig. 3g; Extended Data Fig. 5d), was located distantly from *Edwardsiella, Plesiomonas*, and *Turicibacter* (Fig. 6a). In contrast, *Romboutsia*, *Paraclostridium*, and *Clostridium* displayed similar metabolic gene repertories with *Cetobacterium* (Fig. 6a).

We next evaluated potential competitive and facilitative interactions between microbes based on a genome-scale metabolic modeling approach. In the analysis, reference genomic information was used to infer competition for available resources and exchanges of metabolites, yielding metabolic resource overlap and metabolic interaction potential scores for each pair of microbial ASVs. We then found that the *Romboutsia*, *Edwardsiella,* and *Plesiomonas* ASVs had relatively low metabolic resource overlap and relatively high metabolic interaction potential with the *Cetobacterium* ASV among the prokaryotes examined (Fig. 6b).

Discussion

Through the 128-day monitoring of thousands of microbial species/strains, we here found that aquatic microbiomes associated with fish could show drastic shifts of community structure through time. Such dynamical nature of community processes has been intensively investigated in human-associated microbiomes in light of potential influence on host status[17,18]. In particular, shifts (collapse) of microbial community structure to disease-related states (i.e., dysbiosis) have been considered as essential mechanisms determining human health[33,34]. Given the growing literature on microbiome dynamics in medical science, knowledge of shifts between alternative states of fish-related microbiomes[14] is expected to shed new lights on physiological and ecological processes of vertebrates.

The aquaculture microbiome dynamics were described as shifts among Fusobacteriaceae-abundant states, Flavobacteriaceae-dominated states, and Chitinophagaceae-dominated states, although intermediate states existed through the time-series (Fig. 1). Among them, Fusobacteriaceae-abundant states, which were characterized by high abundance of *Cetobacterium*, were designated as microbiome compositions positively associated with eels' activity level (Fig. 2; Extended Data Fig. 4). In fact, among the 9,908 microbial ASVs examined, the ASV representing *Cetobacterium somerae* showed the strongest associations with eel's activity level through the time-series even after controlling effects of environmental preference (Fig. 3; Extended Data Fig. 5). This *Cetobacterium* species has been reported from a broad taxonomic range of freshwater fish[35–37], especially from intestines of species that do not require dietary vitamin B_{12}[24]. Although vitamin B_{12} (cobalamin) plays essential roles in animal physiology (e.g., normal functioning of nervous systems and the maturation of red blood cells), they can be synthesized only by specific clades of bacteria and archaea[38,39]. Genomic studies have shown that *C. somerae* has a series of genes for anaerobic vitamin B_{12} biosynthesis[40]. Indeed, the bacterium produces highest concentrations of vitamin B_{12} compared to other culturable bacteria within freshwater fish-associated microbiomes[24,41]. Given the prevalence of *Cetobacterium* in freshwater fish species[35–37], our results suggest that maintaining microbiomes at *Cetobacterium*-abundant states is the key to build general platforms for stably keeping freshwater aquaculture/aquarium systems.

Further analyses based on network theory and metabolic modeling indicated the possibility that the *Cetobacterium* species form facilitative interactions with some other microbial species/ASVs (Figs. 4-6). Among the bacteria for which interactions with *Cetobacterium* were inferred from multiple analyses, *Edwardsiella tarda* has been known to include notorious pathogens of broad taxonomic groups of fish including eels[28,29,42]. However, we found that the *E. tarda* population of the investigated aquaculture system was dominated by non-pathogenic strains[29,30] of the species (Extended Data Fig. 6). Thus, the presence of microbial species/strains belonging to broadly-known taxa of pathogens do not necessarily result in negative impacts on fish. Rather, our analyses suggested that "seemingly pathogenic" microbes could be involved in core microbiome components (network modules) constituted by microbes contributing to the maintenance of fish health. Further studies are awaited to explore potential mechanisms such as competitive exclusion of pathogenic strains by non-pathogenic strains[43,44] or indirect negative impacts on pathogenic strains through the activation of fish immune systems[6,45] by non-pathogenic strains. In contrast to *E. tarda*, *Romboutsia* and *Plesiomonas*, which were inferred as microbes with facilitative interactions

with *C. somerae*, too (Figs. 4-6), have been poorly investigated in terms of their functions. Their potential roles in competitive exclusion of pathogens or activation of host immune systems deserve further investigations.

While the time-series dataset allowed us to highlight core species and interactions within microbial communities, more sophisticated statistical platforms beyond simple correlational approaches are necessary for confirming causative relationships between microbiome dynamics and vertebrate health/performance. In this respect, methods based on nonlinear mechanics, such as transfer entropy and empirical dynamic modeling[46,47], are expected to help us infer causative interactions among microbial population dynamics, environmental factors, and vertebrate performance. Albeit promising, these methods require substantial computational resources when we try to analyze microbiomes consisting of thousands of ASVs. Further methodological advances will deepen our understanding of the mechanisms by which microbiome dynamics and vertebrate performance are linked with each other.

As the analyses of microbiome dynamics extend from medical science to researches targeting other vertebrates, we will be more and more aware of overlooked roles of microbes in both terrestrial and aquatic ecosystems. Feedback between intestine and environmental microbiomes, for example, deserves future intensive research in terms of potential great impacts on ecosystem processes. In particular, given that aquatic vertebrates are continuously exposed to excrements of other individuals or species, their gut microbiome dynamics (and related health conditions) may be more likely to be synchronized at the population or community levels than those of terrestrial vertebrates. Therefore, simultaneous monitoring of intestine and background environmental microbiomes will provide platforms for uncovering such feedback and synchronization processes. Further insights into fish-associated microbiome dynamics will reorganize our basic understanding of aquatic ecosystem dynamics, advancing technologies for sustainable food production through stable aquaculture systems[48–50].

References

1. Ley, R. E. *et al.* Evolution of mammals and their gut microbes. *Science (1979)* **320**, 1647–1651 (2008).

2. Youngblut, N. D. *et al.* Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nat Commun* **10**, 1–15 (2019).

3. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174–180 (2011).

4. Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R. & Gordon, J. I. Worlds within worlds: Evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* **6**, 776–788 (2008).

5. McFall-Ngai, M. *et al.* Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci U S A* **110**, 3229–3236 (2013).

6. Yukgehnaish, K. *et al.* Gut microbiota metagenomics in aquaculture: factors influencing gut microbiome and its physiological role in fish. *Rev Aquac* **12**, 1903–1927 (2020).

7. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).

8. Barko, P. C., McMichael, M. A., Swanson, K. S. & Williams, D. A. The Gastrointestinal Microbiome: A Review. *J Vet Intern Med* **32**, 9–25 (2018).

9. Cho, I. & Blaser, M. J. The human microbiome: At the interface of health and disease. *Nat Rev Genet* **13**, 260–270 (2012).

10. Legrand, T. P. R. A., Wynne, J. W., Weyrich, L. S. & Oxley, A. P. A. A microbial sea of possibilities: current knowledge and prospects for an improved understanding of the fish microbiome. *Rev Aquac* **12**, 1101–1134 (2020).

11. Tarnecki, A. M., Burgos, F. A., Ray, C. L. & Arias, C. R. Fish intestinal microbiome: diversity and symbiosis unravelled by metagenomics. *J Appl Microbiol* **123**, 2–17 (2017).

12. Bartelme, R. P., Smith, M. C., Sepulveda-Villet, O. J. & Newton, R. J. Component microenvironments and system biogeography structure microorganism distributions in recirculating aquaculture and aquaponic systems. *mSphere* **4**, e00143-19 (2019).

13. Rud, I. *et al.* Deep-sequencing of the bacterial microbiota in commercial-scale recirculating and semi-closed aquaculture systems for Atlantic salmon post-smolt production. *Aquac Eng* **78**, 50–62 (2017).

14. Infante-Villamil, S., Huerlimann, R. & Jerry, D. R. Microbiome diversity and dysbiosis in aquaculture. *Rev Aquac* **13**, 1077–1096 (2021).

15. Venter, J. C. *et al.* Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science (1979)* **304**, 66–74 (2004).

16. Wang, Y. *et al.* Comparison of the levels of bacterial diversity in freshwater, intertidal wetland, and marine sediments by using millions of illumina tags. *Appl Environ Microbiol* **78**, 8264–8271 (2012).

17. Ravel, J. *et al.* Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome* **1**, 29 (2013).

18. Halfvarson, J. *et al.* Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* **2**, 1–7 (2017).

19. Stoeck, T. *et al.* Environmental DNA metabarcoding of benthic bacterial communities indicates the benthic footprint of salmon aquaculture. *Mar Pollut Bull* **127**, 139–149 (2018).

20. Klase, G. *et al.* The microbiome and antibiotic resistance in integrated fishfarm water: Implications of environmental public health. *Science of the Total Environment* **649**, 1491–1501 (2019).

21. Djurhuus, A. *et al.* Environmental DNA reveals seasonal shifts and potential interactions in a marine community. *Nat Commun* **11**, 254 (2020).

22. Ushio, M. Interaction capacity as a potential driver of community diversity. *Proceedings of the Royal Society B: Biological Sciences* **289**, 20212690 (2022).

23. Fujita, H. *et al.* Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *bioRxiv* https://doi.org/10.1101/2022.08.23.505041.

24. Sugita, H., Miyajima, C. & Deguchi, Y. The vitamin B12-producing ability of the intestinal microflora of freshwater fish. *Aquaculture* **92**, 267–276 (1991).

25. Cai, W., de La Fuente, L. & Arias, C. R. Biofilm formation by the fish pathogen flavobacterium columnare: Development and parameters affecting surface attachment. *Appl Environ Microbiol* **79**, 5633–5642 (2013).

26. Cao, M., Huang, J., Li, J., Qiao, Z. & Wang, G. *Edaphobaculum flavum* gen. Nov., sp. nov., a member of family Chitinophagaceae, isolated from grassland soil. *Int J Syst Evol Microbiol* **67**, 4475–4481 (2017).

27. Thiel, M., Romano, M. C., Kurths, J., Rolfs, M. & Kliegl, R. Twin surrogates to test for complex synchronisation. *Europhys Lett* **75**, 535 (2006).

28. Park, S. bin, Aoki, T. & Jung, T. S. Pathogenesis of and strategies for preventing *Edwardsiella tarda* infection in fish. *Vet Res* **43**, 67 (2012).

29. Yamada, Y. & Wakabayashi, H. Identification of fish-pathogenic strains belonging to the genus *Edwardsiella* by sequence analysis of *sodB*. *Fish Pathol* **34**, 145–150 (1999).

30. Li, G. Y., Li, J., Xiao, P., Guo, Y. H. & Mo, Z. L. Detection of type III secretion gene as an indicator for pathogenic *Edwardsiella tarda*. *Lett Appl Microbiol* **52**, 213–219 (2011).

31. Esteve, C., Biosca, E. & Amaro, C. Virulence of *Aeromonas hydrophita* and some other bacteria isolated from European eels *Anguilla anguilla* reared in fresh water. *Dis Aquat Organ* **16**, 15–20 (1993).

32. Kozińska, A., Paździor, E., Pękala, A. & Niemczuk, W. *Acinetobacter johnsonii* and *Acinetobacter lwoffii* - The emerging fish pathogens. *Bulletin of the Veterinary Institute in Pulawy* **58**, 193–199 (2014).

33. Carding, S., Verbeke, K., Vipond, D. T., Corfe, B. M. & Owen, L. J. Dysbiosis of the gut microbiota in disease. *Microb Ecol Health Dis* **26**, 26191 (2015).

34. Levy, M., Kolodziejczyk, A. A., Thaiss, C. A. & Elinav, E. Dysbiosis and the immune system. *Nat Rev Immunol* **17**, 219–232 (2017).

35. Kim, P. S. *et al.* Host habitat is the major determinant of the gut microbiome of fish. *Microbiome* **9**, 166 (2021).

36. Peng, M. *et al.* Dysbiosis of intestinal microbiota induced by dietary oxidized fish oil and recovery of diet-induced dysbiosis via taurine supplementation in rice field eel (*Monopterus albus*). *Aquaculture* **512**, 734288 (2019).

37. Huang, W. *et al.* Community composition, diversity, and metabolism of intestinal microbiota in cultivated European eel (*Anguilla anguilla*). *Appl Microbiol Biotechnol* **102**, 4143–4157 (2018).

38. Degnan, P. H., Taga, M. E. & Goodman, A. L. Vitamin B12 as a modulator of gut microbial ecology. *Cell Metab* **20**, 769–778 (2014).

394 39. Hazra, A. B. *et al.* Anaerobic biosynthesis of the lower ligand of vitamin B12. *Proc*
395 *Natl Acad Sci U S A* **112**, 10792–10797 (2015).

396 40. LaFrentz, B. R., LaFrentz, S. A., Beck, B. H. & Arias, C. R. Draft genome sequences
397 of *Cetobacterium somerae* 2G large and two novel *Cetobacterium Isolates* from
398 intestines of channel catfish (*Ictalurus punctatus*). *Microbiol Resour Announc* **9**,
399 e01006-20 (2020).

400 41. Tsuchiya, C., Sakata, T. & Sugita, H. Novel ecological niche of *Cetobacterium*
401 *somerae*, an anaerobic bacterium in the intestinal tracts of freshwater fish. *Lett Appl*
402 *Microbiol* **46**, 43–48 (2008).

403 42. Wakabayashi, H., Egusa, S. & Yabe, K. *Edwardsiella tarda* (*Paracolobactrum*
404 *anguillimortiferum*) associated with pond-cultured eel disease. *Nippon Suisan*
405 *Gakkaishi* **39**, 931–936 (1973).

406 43. De, B. C. *et al.* Probiotics in fish and shellfish culture: Immunomodulatory and
407 ecophysiological responses. *Fish Physiol Biochem* **40**, 921–971 (2014).

408 44. Moriarty, D. J. W. Control of luminous *Vibrio* species in penaeid aquaculture ponds.
409 *Aquaculture* **164**, 351–358 (1998).

410 45. Gomez, D., Sunyer, J. O. & Salinas, I. The mucosal immune system of fish: The
411 evolution of tolerating commensals while fighting pathogens. *Fish Shellfish Immunol*
412 **35**, 1729–1739 (2013).

413 46. Sugihara, G. *et al.* Detecting causality in complex ecosystems. *Science* **338**, 496–500
414 (2012).

415 47. Schreiber, T. Measuring information transfer. *Phys Rev Lett* **85**, 461–464 (2000).

416 48. Klinger, D. & Naylor, R. Searching for solutions in aquaculture: Charting a sustainable
417 course. *Annu Rev Environ Resour* **37**, 247–276 (2012).

418 49. Bostock, J. *et al.* Aquaculture: Global status and trends. *Philosophical Transactions of*
419 *the Royal Society B: Biological Sciences* **365**, 2897–2912 (2010).

420 50. Dawood, M. A. O., Koshio, S., Abdel-Daim, M. M. & van Doan, H. Probiotic
421 application for sustainable aquaculture. *Rev Aquac* **11**, 907–924 (2019).

422 51. Ushio, M. *et al.* Quantitative monitoring of multispecies fish environmental DNA
423 using high-throughput sequencing. *Metabarcoding Metagenom* **2**, 1–15 (2018).

52. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* **108**, 4516–4522 (2011).

53. Apprill, A., Mcnally, S., Parsons, R. & Weber, L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology* **75**, 129–137 (2015).

54. Lundberg, D. S., Yourstone, S., Mieczkowski, P., Jones, C. D. & Dangl, J. L. Practical innovations for high-throughput amplicon sequencing. *Nat Methods* **10**, 999–1002 (2013).

55. Stevens, J. L., Jackson, R. L. & Olson, J. B. Slowing PCR ramp speed reduces chimera formation from environmental samples. *J Microbiol Methods* **93**, 203–205 (2013).

56. Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J. & Knight, R. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* **5**, 235–237 (2008).

57. Tanabe, A. Claident v0.2.2018.05.29, a software distributed by author at http://www.fifthdimension.jp/. (2018).

58. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581–583 (2016).

59. R Core Team. R: A language and environment for statistical computing. *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/* (2020).

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

61. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590–D596 (2013).

62. Klappenbach, J. A., Saxman, P. R., Cole, J. R. & Schmidt, T. M. Rrndb: The ribosomal RNA operon copy number database. *Nucleic Acids Res* **29**, 181–184 (2001).

63. Oksanen, J. The vegan package available at https://cran.r-project.org/web/packages/vegan/index.html. (2007).

64. Kurtz, Z. D. *et al.* Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol* **11**, e1004226 (2015).

65. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal Complex Systems* **1695**, 1–9 (2006).

66. Kurtz, Z. D., Bonneau, R. & Müller, C. L. Disentangling microbial associations from hidden environmental and technical factors via latent graphical models. *bioRxiv* (2019) doi:10.1101/2019.12.21.885889.

67. Douglas, G. M. *et al.* PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* **38**, 685–688 (2020).

68. Fahimipour, A. K. & Gross, T. Mapping the bacterial metabolic niche space. *Nat Commun* **11**, 1–8 (2020).

69. Alneberg, J. *et al.* Ecosystem-wide metagenomic binning enables prediction of ecological niches from genomes. *Commun Biol* **3**, 1–10 (2020).

70. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**, D353–D361 (2017).

71. Zelezniak, A. *et al.* Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc Natl Acad Sci U S A* **112**, 6449–6454 (2015).

72. Machado, D., Andrejev, S., Tramontano, M. & Patil, K. R. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res* **46**, 7542–7553 (2018).

Methods

Sampling. Monitoring of microbiome dynamics was conducted targeting the five water tanks of the eel aquaculture system of A-Zero Inc. (Nishiawakura, Okayama Prefecture, Japan). In each water tank (diameter = 2.5 m; height = 1 m; volume = 20 m^3), 1,400–4,300 eel individuals (average weight = 80–130 g) had been kept. About 10 % of tank water was replaced with warmed fresh well water every day, and the water temperature in the tanks was kept at around 30 ºC. The drainage from the five tanks were mixed and processed in a series of filtration equipment. The filtered drainage was returned to each tank after being processed in another filtration equipment adjacent to each tank. The eels were fed with mixture of commercial artificial diets. The pH, dissolved oxygen (DO), and eels' activity level were recorded for each tank every day. The eels' activity level was evaluated based on the sum of the scores of the following eight criteria: initial responses to feeders, the proportion of eels responding to feeders, sharpness of movement, the proportion of eels eating the artificial diet, the level of splashes, the amount of scattered diet, the time to consume the diet, and the proportion of foraging eels at the end of feeding. For each of the criteria, scoring was done on a five-point scale (maximum point = 5) by an expert of eel aquaculture maintenance: thus, 40 (5 × 8 criteria) is the maximum point of the eels' activity score. Albeit subjective, the criteria evaluated continuously by a professional provide inferences of eel's health conditions throughout the time series. The water in the tanks were continuously mixed by the movement of eels.

From each aquaculture tank (Tank 1–5), ca. 1.5 mL of water was sampled in the morning every day during the 128 days from March 25 to July 30, 2020, except for 9 days (Days 102, 103, 120, 121, 122, 123, 124, 125, and 126): i.e., the samples of 119 days were available. In Tank 4, samples were unavailable on additional three days (Day 67–69) due to the cleaning and the entire replacement of water. Consequently, the number of collected samples were 592 (119 days × 5 tanks – 3 days in Tank 4). Water sample was collected in a 2.0 mL microtube and they were immediately stored at -20 ºC in a freezer until DNA extraction.

Quantitative 16S rRNA sequencing. To extract DNA from each sample, 250 µL of the collected water was mixed with mixed with 400 µL lysis buffer (0.0025 % SDS, 20 mM Tris (pH 8.0), 2.5 mM EDTA, and 0.4 M NaCl) and 250 µL 0.5 mm zirconium beads in a 2.0 mL microtube. The microtubes were then shaken at 25 Hz for 5 min using TissueLyser II (Qiagen, Venlo). After centrifugation, the aliquot was mized with proteinase K solution

(×1/100 of the total volume), being incubated at 40 ºC for 60 min followed by 95 ºC for 5 min.

We then performed PCR by applying a quantitative amplicon sequencing method[22,51]. Although most existing microbiome studies were designed to infer "relative" abundance of microbial amplicon sequence variants (ASVs) or operational taxonomic units (OTUs), information of "absolute" abundance provide additional insights into microbiome dynamics: i.e., insights into increase/decrease of the population size of each prokaryote ASV/OTU within a microbiome throughout a time-series[22]. The quantitative amplicon sequencing approach is based on the addition of artificial (standard) DNA sequences with defined concentrations into PCR master solutions. Therefore, even if compositions or concentrations of PCR inhibitor molecules in DNA extracts vary among time-series samples, potential bias caused by such inhibitors can be corrected based on the use of the internal standards (i.e., standard DNAs within PCR master solutions).

Prokaryote 16S rRNA region was PCR-amplified with the forward primer 515f[52] fused with 3–6-mer Ns for improved Illumina sequencing quality and the forward Illumina sequencing primer (5'- TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG- [3–6-mer Ns] – [515f] -3') and the reverse primer 806rB[53] fused with 3–6-mer Ns for improved Illumina sequencing quality[54] and the reverse sequencing primer (5'- GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G [3–6-mer Ns] - [806rB] -3') (0.2 μM each). To apply the quantitative amplicon sequencing, five standard DNA sequence variants with different concentrations of artificial 16S rRNA sequences (0.1, 0.05, 0.02, 0.01, and 0.005 nM) were added to PCR master mix solutions[22]. The buffer and polymerase system of KOD One (Toyobo) was used with the temperature profile of 35 cycles at 98 ºC for 10 s, 55 ºC for 30 s, 68 ºC for 30 s. To prevent generation of chimeric sequences, the ramp rate through the thermal cycles was set to 1 ºC/sec[55]. Illumina sequencing adaptors were then added to respective samples in the supplemental PCR using the forward fusion primers consisting of the P5 Illumina adaptor, 8-mer indexes for sample identification[56] and a partial sequence of the sequencing primer (5'- AAT GAT ACG GCG ACC ACC GAG ATC TAC AC - [8-mer index] - TCG TCG GCA GCG TC -3') and the reverse fusion primers consisting of the P7 adaptor, 8-mer indexes, and a partial sequence of the sequencing primer (5'- CAA GCA GAA GAC GGC ATA CGA GAT - [8-mer index] - GTC TCG TGG GCT CGG -3'). KOD One was used with a temperature profile: followed by 8 cycles at 98 ºC for 10 s, 55 ºC for 30 s, 68 ºC for 30 s (ramp rate = 1 ºC/s). The PCR amplicons of the samples were then pooled after a purification/equalization process with the AMPureXP Kit (Beckman Coulter). Primer dimers, which were shorter than 200 bp, were removed from the pooled library by supplemental

purification with AMpureXP: the ratio of AMPureXP reagent to the pooled library was set to 0.6 (v/v) in this process. Because the quality of forward sequences is generally higher than that of reverse sequences in Illumina sequencing, we optimized the MiSeq run setting in order to use only forward sequences. Specifically, the run length was set 271 forward (R1) and 31 reverse (R4) cycles to enhance forward sequencing data: the reverse sequences were used only for screening 16S rRNA sequences in the following bioinformatic pipeline.

Bioinformatics. In total, 16,298,203 sequencing reads were obtained in the Illumina sequencing. The raw sequencing data were converted into FASTQ files using the program bcl2fastq 1.8.4 distributed by Illumina. The raw sequencing data were converted into FASTQ files using the program bcl2fastq 1.8.4 distributed by Illumina. The output FASTQ files were demultiplexed using Claident v0.2. 2018.05.29[57]. The removal of low-quality sequences and ASV inferences were done using DADA2[58] v.1.22.0 of R 4.1.2[59](R Core Team, 2020). The taxonomy of the output ASVs was inferred based on the naive Bayesian classifier method[60] using the SILVA v.138 database[61]. Based on the calibration with the concentration gradients of the five standard DNAs, concentrations of respective ASVs were obtained for each sample (16S rRNA copy numbers per unit volume of tank water samples; copies/μL). As the number of 16S rRNA copies per genome generally varies among prokaryotic taxa[62], 16S rRNA copy concentration is not directly the optimal proxy of cell or biomass concentration. Meanwhile, in this study, estimates of 16S rRNA copy concentrations were used to observe increase/decrease of abundance (i.e., population dynamics) *within* the time-series of respective microbial ASVs. Thus, variation in the number 16S rRNA copy numbers among microbial taxa had no qualitative effects on the subsequent population- and community-ecological analyses. The samples in which Pearson's coefficients of correlations between sequencing read numbers and standard DNA copy numbers (i.e., correlation coefficients representing calibration curves) were less than 0.8 were removed as those with unreliable estimates. Samples with less than 1,000 reads were discarded as well. In total, microbiome data were successfully obtained from 577 out of 592 samples. For each aquaculture tank, we then obtained a sample × ASV matrix, in which a cell entry depicted the concentration of 16S rRNA copies of an ASV in a sample.

Community structure. For each aquaculture tank, Bray-Curtis β-diversity was calculated for all pairs of time points based on the matrix describing the relative abundance of prokaryote families using the vegan 2.6.2 package[63] of R. Based on the β-diversity estimates, the

community structure of all the samples across the five water tanks were visualized on the surface of non-metric multidimensional scaling (NMDS). The vectors representing the environmental variables (pH and DO) and eels' activity level were calculated with the "envfit" function of R and they were shown on the NMDS surface. The analysis was conducted as well based on the matrix describing the relative abundance of genera.

Environmental preference of ASVs. To evaluate environmental preference of each microbial ASV, Spearman's correlation between absolute abundance (in the metric of DNA copy numbers of 16S rRNA) and pH was calculated. For each tank, the ASVs that appeared in 30 or more samples were subjected to the analysis. For each ASV in each water tank, the statistical significance of the obtained correlation coefficient was examined with a randomization analysis obtained based on a twin-surrogate method for time-series data[27] (100,000 permutations). Correlation coefficients less than -0.3 and those larger than 0.3 tended to show statistically significant negative and positive correlations with pH, respectively, after Benjamini-Hochberg adjustment of *P* values in multiple testing [i.e., false discovery rate (FDR)]. Likewise, Pearson's correlation coefficients between respective ASVs' absolute abundance and DO concentrations were calculated.

ASV abundance and eel's activity. We explored microbial ASVs that potentially have profound impacts on eels' health. For each water tank, Spearman's correlation between absolute abundance and eels' activity score was calculated for the ASVs that appeared in 30 or more samples. However, because ASV abundance could be affected by pH or dissolved oxygen concentration, the use of such simple correlation coefficients might be misleading. Therefore, we controlled potential effects by environmental factors/conditions based on a partial correlation approach as follows:

$$r_{xy\cdot z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{1-r_{xz}^2}\sqrt{1-r_{yz}^2}},$$

where r_{xy}, r_{xz}, and r_{yz} were correlation between ASV abundance and eels' activity level, that between ASV abundance and an environmental factor (pH or dissolved oxygen concentration), and that between eels' activity level and an environmental factor, respectively. For each ASV, a randomization analysis was performed with the twin-surrogate method (100,000 permutations).

Time-lag analysis. We extended the analysis of partial correlation between microbial abundance and eels' activity level by introducing time lags between the two variables. Specifically, partial correlation between an ASV's abundance on Day x and eels' activity score on Day $x + l$ was calculated. The time lag l ranged from -5 to 5 in the analysis ($l = 0$ means no delay introduced to eels' activity level).

Pathogenic and non-pathogenic *Edwardsiella.* We performed an additional analysis to infer the proportion of pathogenic and non-pathogenic clades[29,30] of *Edwardsiella* bacteria in the aquaculture system. In a previous phylogenetic study based on an internal fragment of iron-cofactored superoxide dismutase gene (*sodB*), *Edwardsiella* species and strains have been classified into two major clades, which differ in the presence of pathogenicity to fish (hereafter, "pathogenic" and "non-pathogenic" clades). Therefore, we characterized *Edwardsiella* bacteria in the aquaculture tanks based on the illumina sequencing of the *Edwardsiella sodB* gene sequences. The fragment of the *sodB* region was PCR-amplified with the forward primer E1F[29] fused with 3–6-mer Ns for improved Illumina sequencing quality and the forward Illumina sequencing primer (5'- TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG- [3–6-mer Ns] – [E1F] -3') and the reverse primer 497R[29] fused with 3–6-mer Ns for improved Illumina sequencing quality[54] and the reverse sequencing primer (5'- GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G [3–6-mer Ns] - [497R] -3') (0.2 μM each). The buffer and polymerase system of KOD One (Toyobo) was used with the temperature profile of 35 cycles at 98 ºC for 10 s, 55 ºC for 5 s, 68 ºC for 30 s (ramp rate = 1 ºC/sec). The sequencing adaptors and sample identifier indexed were added to the amplicons, and the purification of the library and sequencing was performed as detailed above.

The output sequencing data were demultiplexed and processed with DADA2. The ASVs that were not aligned to the *sodB* sequences of *Edwardsiella*[29] were discarded. The neighbor-joining tree of the remaining ASVs and previously reported *Edwardsiella* sequences was reconstructed based on the maximum composite likelihood method with a bootstrap test (1,000 permutations). The ASVs belonging to the pathogenic clade and those belonging to the non-pathogenic clade of *Edwardsiella* were distinguished within the phylogeny.

Microbe–microbe interactions. Potential positive/negative interactions between microbial ASVs were inferred based on the framework of sparse inverse covariance estimation for

ecological associations (SPIEC-EASI[64]). For each water tank, patterns in the coexistence (co-occurrence) were examined with the Meinshausen-Bühlmann (MB) method as implemented in the SpiecEasi package[64] of R. The network inference based on coexistence patterns allowed us to detect pairs of microbial ASVs that potentially interact with each other in facilitative ways and/or those potentially sharing environmental preference. Because estimation of coexistence patterns was not feasible for rare nodes, the microbial ASVs that appeared in less than 30 samples were excluded from the input matrices of the network analysis. Network modules, within which closely associated ASVs were interlinked with each other, were identified with the algorithm based on edge betweenness using the igraph package[65] of R. For each network module in each water tank, mean partial correlation with eels' activity level was calculated across ASVs constituting the module.

In addition to the networks representing whole coexistence patterns, we reconstructed networks depicting direct interactions between microbial ASVs. To separate effects of direct microbe–microbe interactions from those of shared environmental preferences between microbes (i.e., shared niches), 10 latent components (latent variables) were included in the analysis based on the "sparse and low-rank" (SLR) model[66].

KEGG pathway/process profiles. To infer metabolic interactions between microbial ASVs, we performed a series of analysis based on reference genome information. We performed phylogenetic prediction of gene repertoires using PICRUSt2 v2.3.0-b[67] in order to gain the overview of the niche space defined with metagenomic information[68,69]. ASVs that appeared in 30 or more sample across the five tanks were subjected to the analysis. Based on the inferred KEGG metabolic pathway/process profiles[70], microbial ASVs were plotted on a two-dimensional surface of a principal coordinate analysis (PCoA) based on Bray-Curtis β-diversity of KEGG metabolic pathway/process profiles.

Metabolic modeling. To infer potential metabolic interactions between microbes, we performed the species metabolic interaction analysis[71]. For the ASVs that appeared in 30 or more samples (day) in at least one aquaculture tank, we explored NCBI RefSeq genome sequences whose 16S rRNA sequences matched those of query ASVs with ≥ 99 % identity. In the database exploration, reference genome information was available for 181 out of 417 ASVs examined. The reference genome information was subjected to genome-scale metabolic modeling as implemented in CarveMe[72] 1.5.0. Metabolic resource overlap (MRO) and

676 metabolic interaction potential (MIP) were then estimated for each pair of microbial ASVs as
677 implemented in SMETANA[71] 1.0.0.

678

679 ## Data availability

680 The 16S rRNA sequencing data are available from the DNA Data Bank of Japan (DDBJ) with
681 the accession number PRJDB14313. The microbial community data are deposited at the
682 GitHub repository (https://github.com/hiro-toju/EelMicrobiome128).

683

684 ## Code availability

685 All the R scripts used to analyze the data are available at the GitHub repository
686 (https://github.com/hiro-toju/EelMicrobiome128).

687

688 **Acknowledgements.** We thank Yuta Nogi and A Zero inc. for support in sampling. JSPS
689 Grant-in-Aid for Scientific Research (20K20586), and JST FOREST (JPMJFR2048) to H.T.,
690 JSPS Grant-in-Aid for Scientific Research (20K06820 and 20H03010) to K.S., and JSPS
691 Fellowship to H.F..

692

693 **Author Contributions.** H.T. designed the work with D.Y.. D.Y., I.H., G.S., and H.F.
694 performed experiments. D.Y., H.F., K.S., and H.T. analyzed the data. H.T. and D.Y. wrote the
695 paper with all the authors.

696

697 **Competing Interests.** The authors declare no competing interests.

698

699 **Correspondence and requests for materials** should be addressed to
700 toju.hirokazu.4c@kyoto-u.ac.jp.

701
702

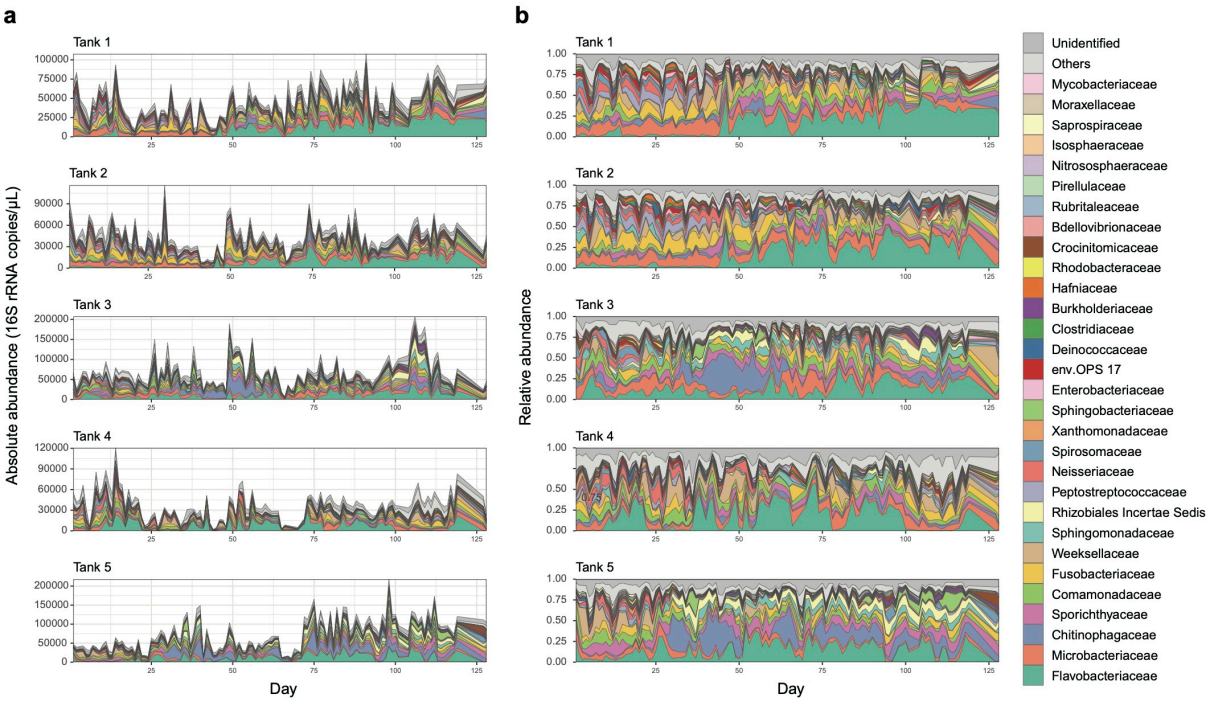
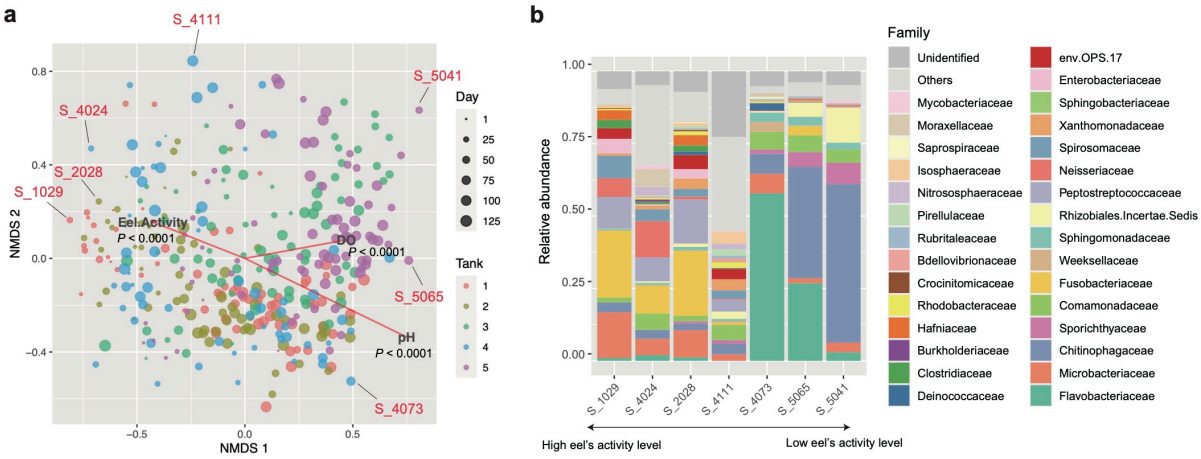


Fig. 1 | Microbiome dynamics in the eel aquaculture system. **a**, Dynamics of absolute abundance. For each water sample of each aquaculture tank, absolute abundance of prokaryotes was inferred as 16S rRNA gene copy concentration based on the quantitative amplicon sequencing approach with standard DNA gradients. **b**, Dynamics of relative abundance. The time-series of the family-level taxonomic compositions are shown for each aquaculture tank. See Extended Data Figures 1–3 for phylum-, order-, and genus-level taxonomic compositions.



712

713 **Fig. 2 | Multivariate analysis of community structure**. **a,** Community state space.
714 Community compositions of the samples are plotted on the two-dimensional surface defined
715 with non-metric multidimensional scaling (NMDS). The NMDS was performed based on the
716 Bray-Curtis β-diversity of family-level taxonomic compositions. The projections of the data
717 points onto the vectors have maximum correlation with the variables examined (pH, DO, and
718 eels' activity level). See Extended Data Figure 4 for an additional analysis based on genus-
719 level taxonomic compositions. **b,** Examples of community structure in the NMDS surface.
720 For several points within the NMDS surface (panel **a)**, family-level taxonomic compositions
721 are shown. The example points are ordered along the vector representing high eels' activity
722 level.
723

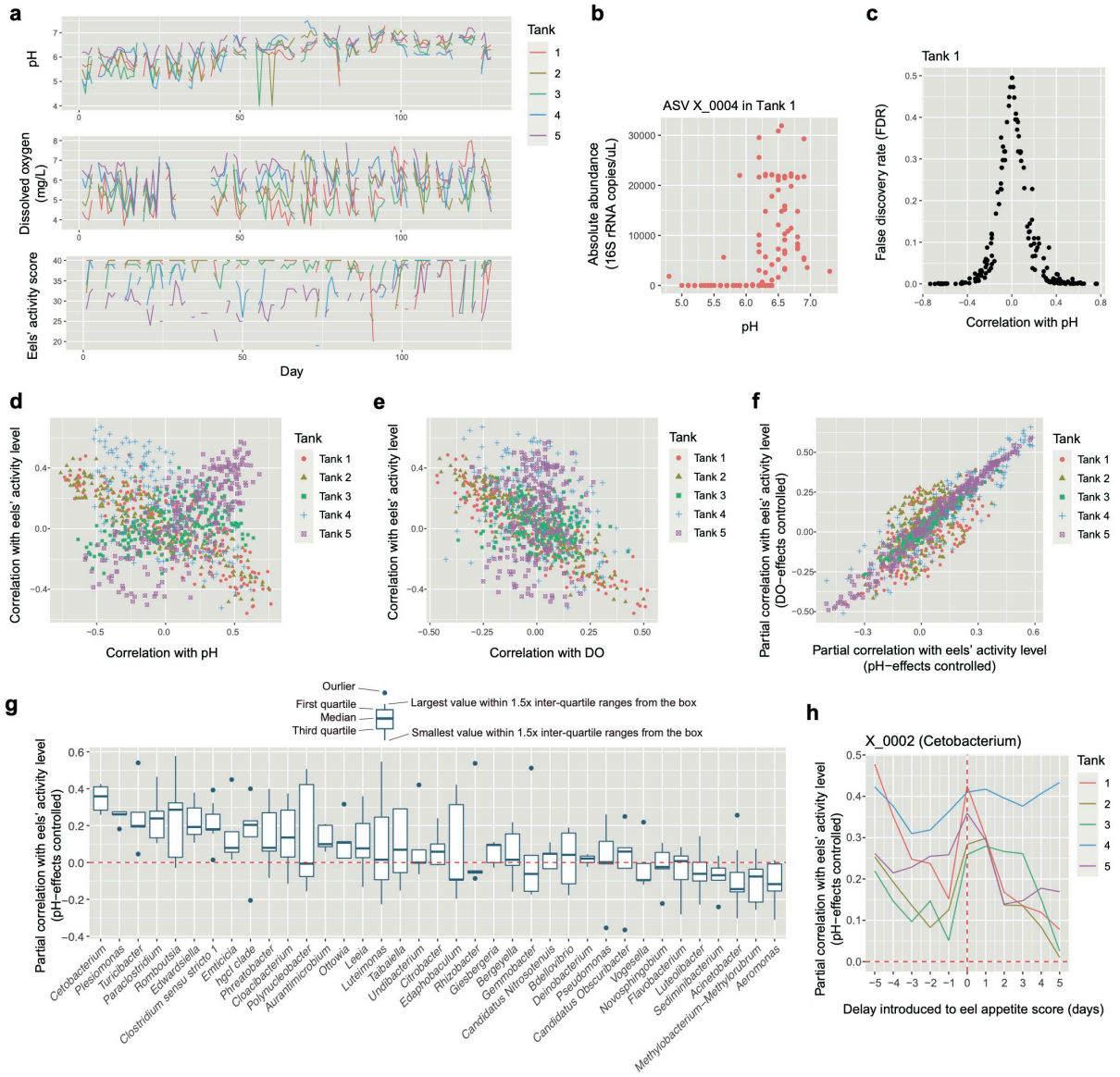


Fig. 3 | Microbes associated with eels' activity. **a**, Timeseries of pH, dissolved oxygen (DO) level, and eels' activity score are shown for each aquaculture tank. **b**, Example of the correlation analysis. For each variable shown in the panel **b**, Spearman's correlation with the absolute abundance of each ASV in each aquaculture tank was examined. **c**, Randomization analysis of correlation. Significance of correlation coefficients was examined based on a twin-surrogate randomization analysis of time-series data (100,000 permutations). Coefficients less than -0.3 and those larger than 0.3 roughly represent significant negative and positive correlations, respectively. **d**, Each ASV's correlation with pH and eels' activity level. **e**, Each ASV's correlation with DO and eels' activity level. **f**, Partial correlation with eels' activity level. To control the effects of pH or DO, partial correlation between absolute abundance and eels' activity scores was calculated for each ASV in each tank. **g**, Taxonomic comparison of relationship with eels' activity level. Partial correlation with eels' activity level is shown for

737 the genera that appeared in all the aquaculture tanks (shown in the decreasing order of mean
738 values). **h**, Time-lag analysis of correlations. In calculating partial correlation between eels'
739 activity level and the absolute abundance of the *Cetobacterium* ASV (X_0002), defined time-
740 lag was introduced to the eels' activity variable.
741

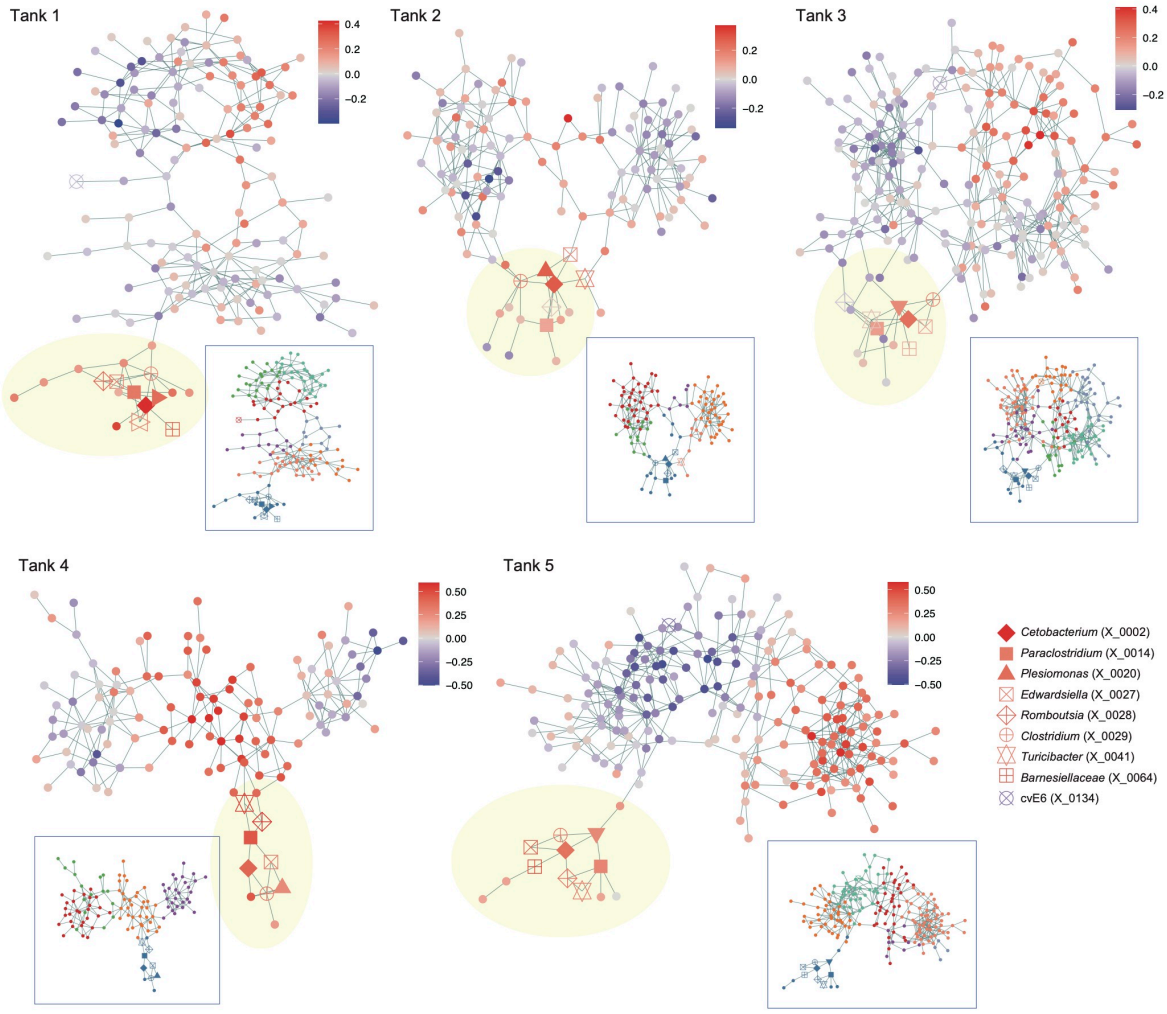


Fig. 4 | Microbe-to-microbe coexistence networks. For each aquaculture tank, patterns of coexistence were analyzed based on the sparse inverse covariance estimation for ecological associations with the Meinshausen-Bühlmann (MB) model. Only the ASVs that appeared in 30 or more samples were targeted in the analysis of each tank. Within the networks, pairs of microbial ASVs that may interact with each other in facilitative ways and/or those potentially sharing environmental preference are linked with each other. Network modules, which represent groups of densely linked ASVs, are shown for each network. The color of nodes indicates partial correlation between ASV abundance and eels' activity level (controlled variable = pH). The inferred network modules are shown by colors for each tank in a box. The ASVs that consistently displayed positive or negative correlation with eels' activity level (Extended Data Fig. 5) are highlighted with the defined symbols. See Extended Data Figures 6–8 for additional information of the nodes (ASVs) and modules within the network. ASVs included in minor sub-networks (number of nodes < 5) are not shown.

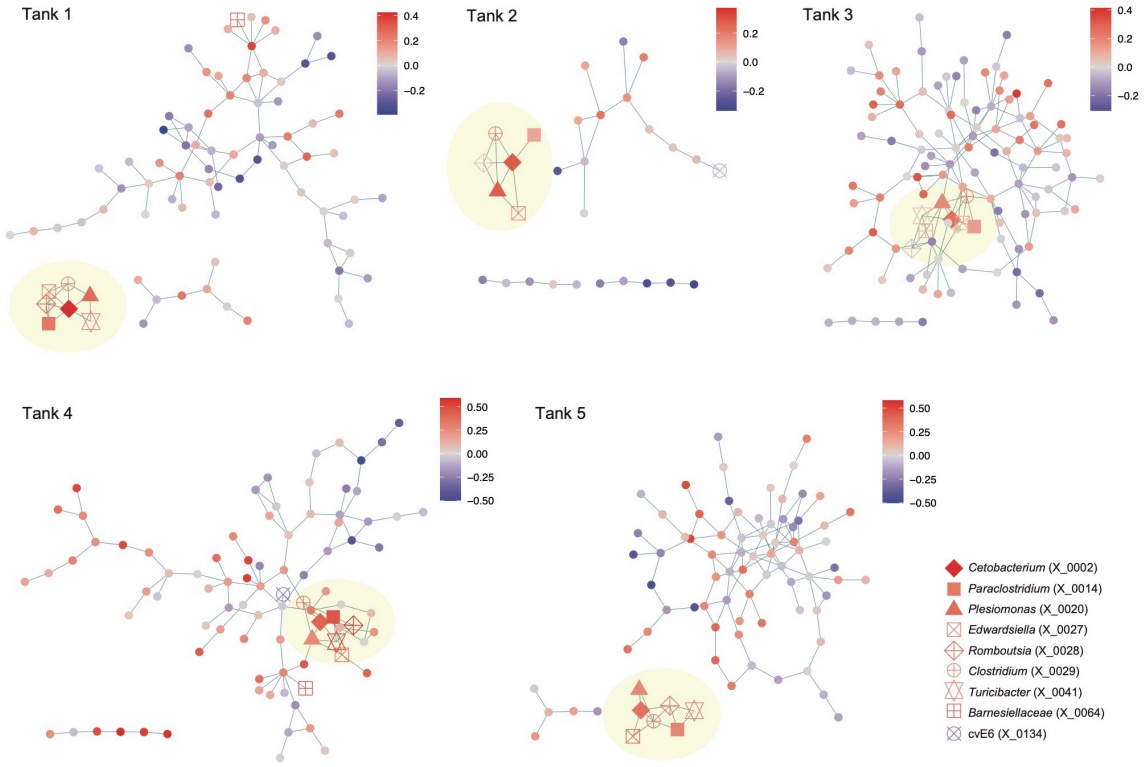


Fig. 5 | Inferred direct interactions between microbes. Based on the "sparse and low-rank" (SLR) model, direct interactions between microbial ASVs were inferred by controlling the effects of shared environmental preference. Only the ASVs that appeared in 30 or more samples were targeted in the analysis of each tank. The links between nodes represent potentially positive interactions between ASVs. The color of nodes indicates partial correlation between ASV abundance and eels' activity level (controlled variable = pH). ASVs included in minor sub-networks (number of nodes < 5) are not shown.

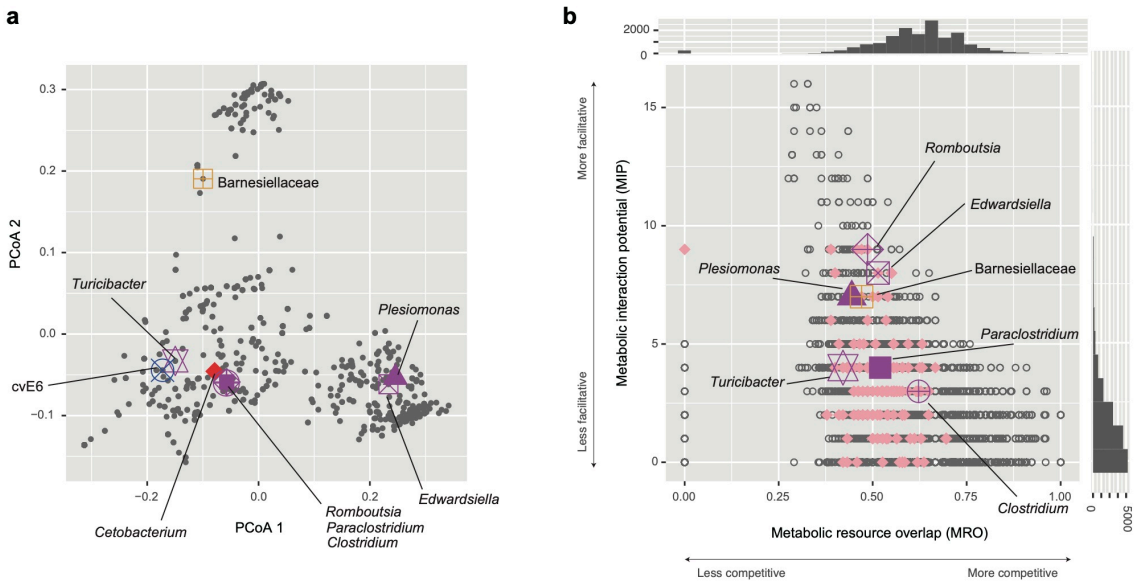
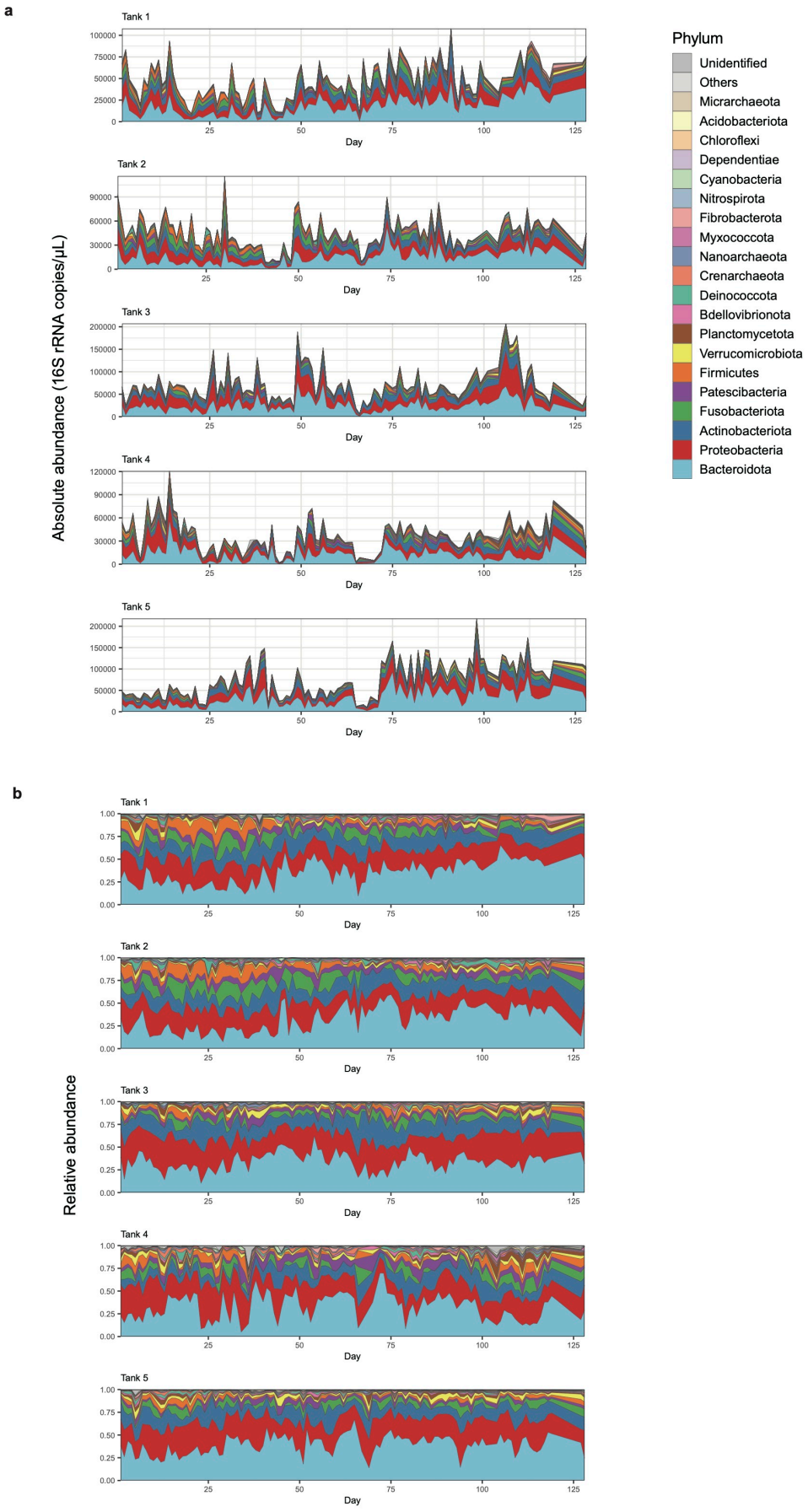


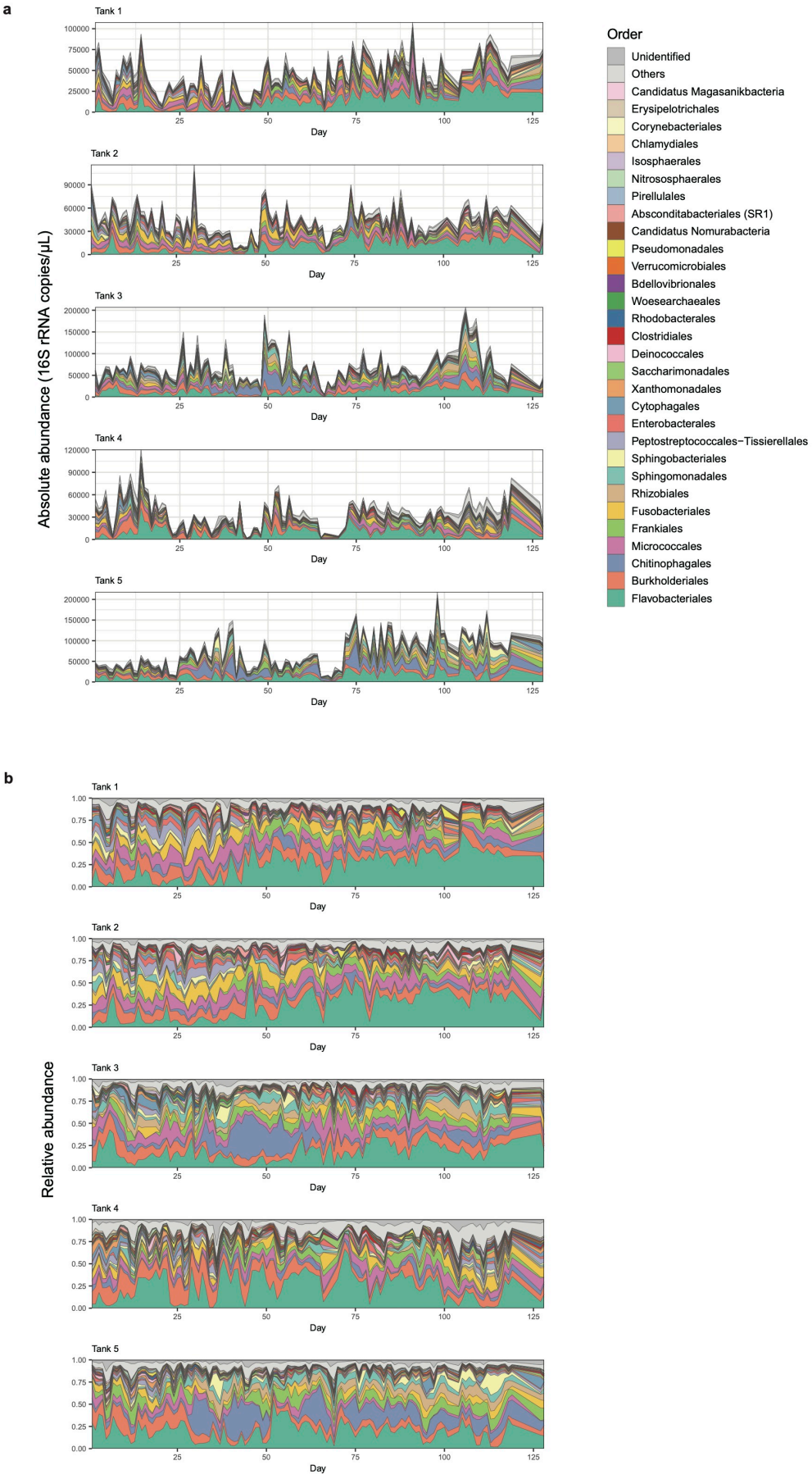
Fig. 6 | Metabolic interactions between microbes. **a**, Metagenomic niche space. Microbial ASVs are plotted on a two-dimensional surface of PCoA based on their KEGG metabolic pathway/process profiles inferred with a phylogenetic prediction of genomes. Microbial ASVs plotted closely within the surface are expected to have similar gene repertoires. The ASVs highlighted in Figures 4 and 5 are shown with large symbols. **b**, Potential competitive and facilitative interactions. Based on the NCBI RefSeq genome information, potential metabolic interactions between each pair of ASVs were inferred in terms of metabolic resource overlap (MRO) and metabolic interaction potential (MIP). Histograms of MRO and MIP are shown on the horizontal and vertical axes, respectively. ASV pairs including the *Cetobacterium* ASV, whose abundance were positively associated with eels' activity level in all the five water tanks (Fig. 3g; Extended Data Fig. 5), are shown in pink. Relationships between the *Cetobacterium* ASV and the ASVs highlighted in Figures 4 and 5 are indicated as well.

Extended Data Figures



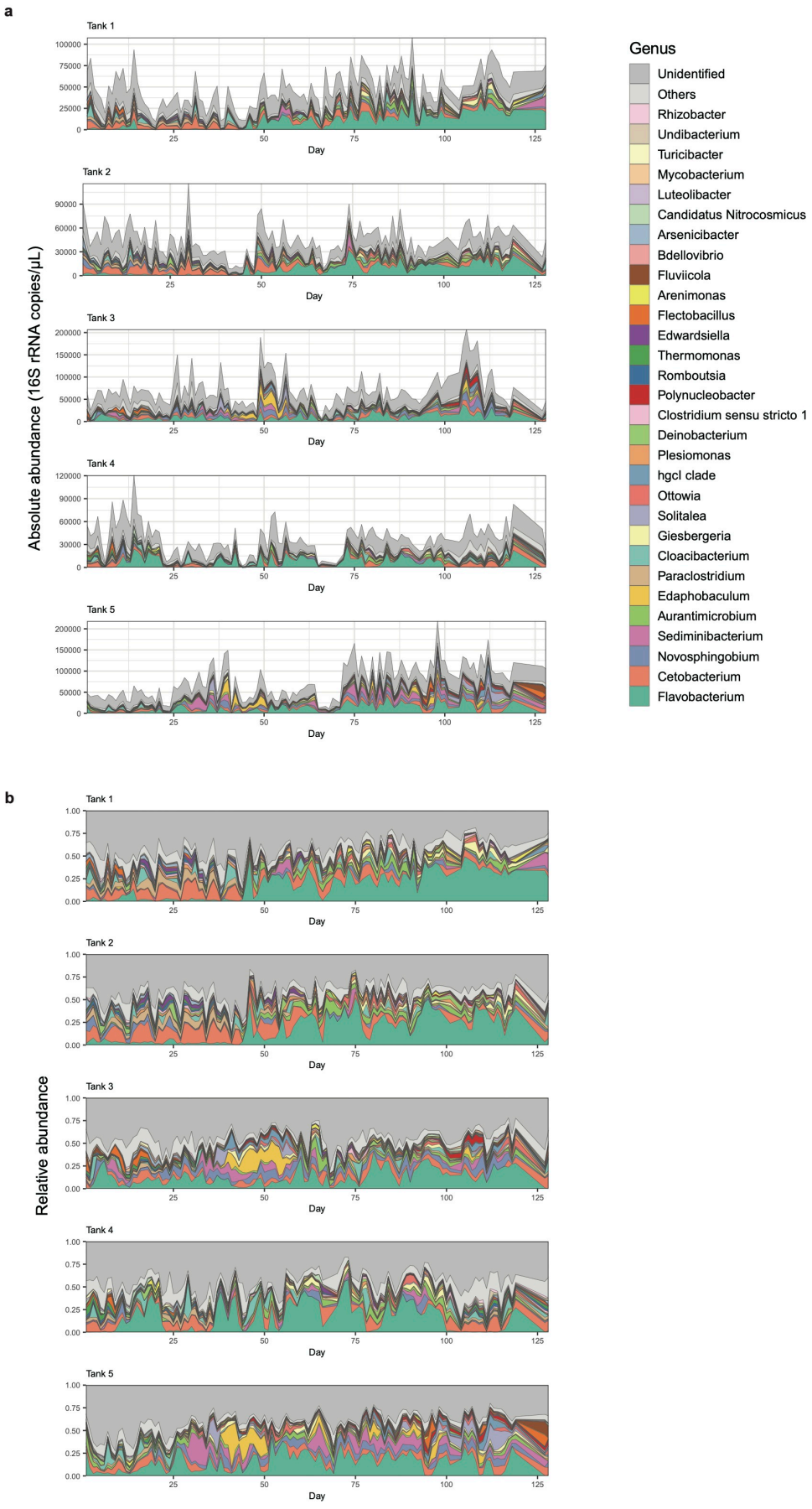
784

785 **Extended Data Fig. 1 | Phylum-level community structure**. **a**, Dynamics of absolute
786 abundance. For each water sample of each aquaculture tank, absolute abundance of
787 prokaryotes was inferred as 16S rRNA gene copy concentration based on the quantitative
788 amplicon sequencing approach with standard DNA gradients. **b**, Dynamics of relative
789 abundance. The time-series of the phylum-level taxonomic compositions are shown for each
790 aquaculture tank.
791



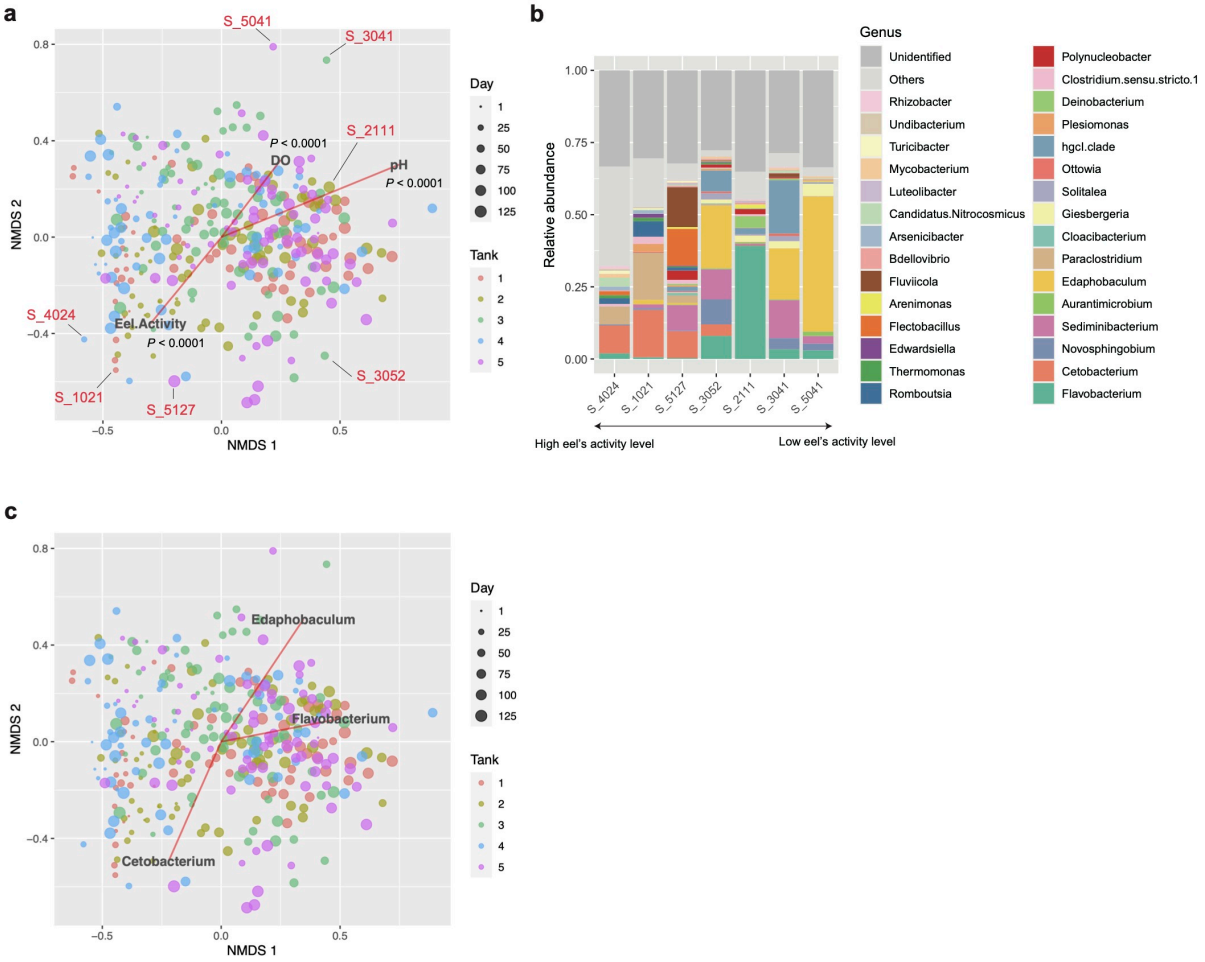
792

Extended Data Fig. 2 | Order-level community structure. **a**, Dynamics of absolute abundance. For each water sample of each aquaculture tank, absolute abundance of prokaryotes was inferred as 16S rRNA gene copy concentration based on the quantitative amplicon sequencing approach with standard DNA gradients. **b**, Dynamics of relative abundance. The time-series of the order-level taxonomic compositions are shown for each aquaculture tank.



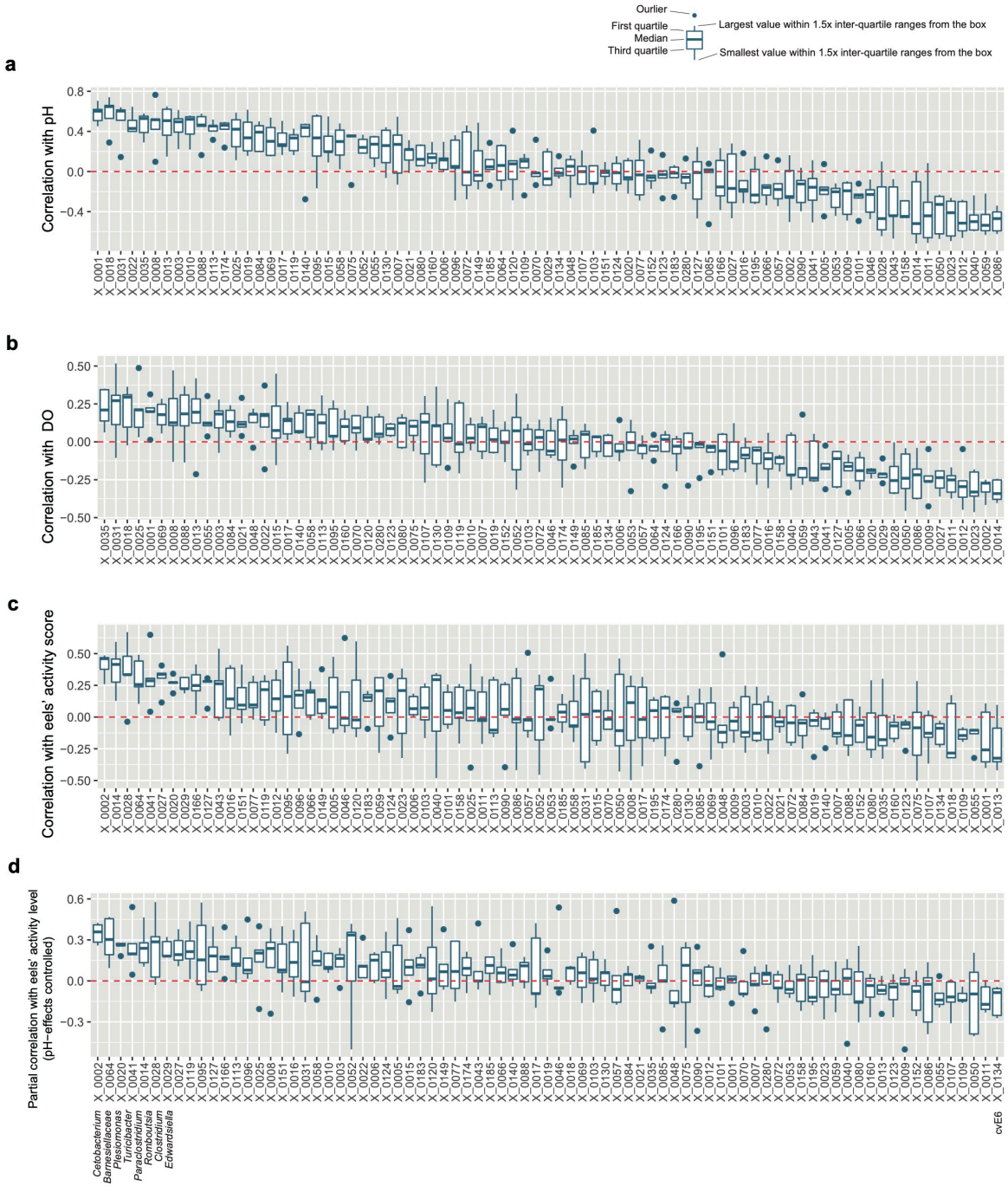
800

801 **Extended Data Fig. 3 | Genus-level community structure**. **a**, Dynamics of absolute
802 abundance. For each water sample of each aquaculture tank, absolute abundance of
803 prokaryotes was inferred as 16S rRNA gene copy concentration based on the quantitative
804 amplicon sequencing approach with standard DNA gradients. **b**, Dynamics of relative
805 abundance. The time-series of the genus-level taxonomic compositions are shown for each
806 aquaculture tank.
807



808

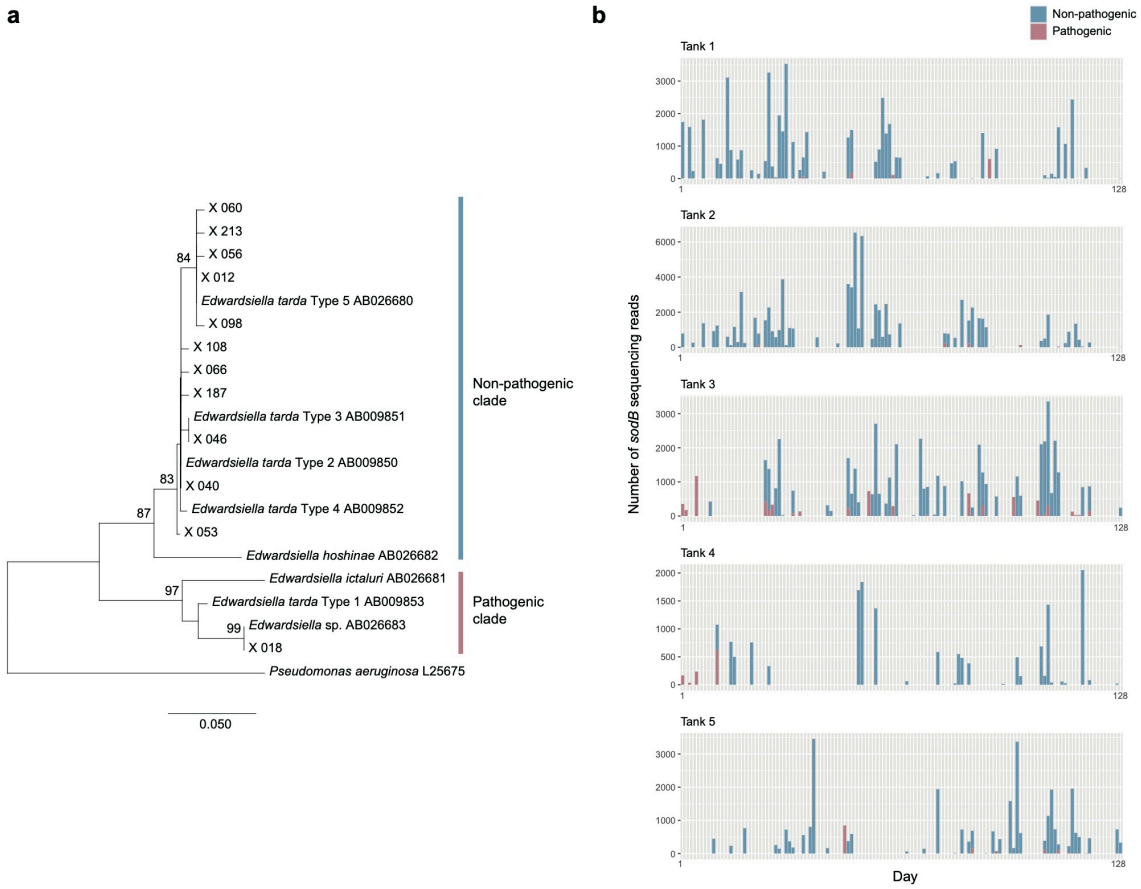
809 **Extended Data Fig. 4 | Multivariate analysis of community structure (genus level)**. **a,**
810 Community state space. Community compositions of the samples are plotted on the two-
811 dimensional surface defined with non-metric multidimensional scaling (NMDS). The NMDS
812 was performed based on the Bray-Curtis β-diversity of genus-level taxonomic compositions.
813 The projections of the data points onto the vectors have maximum correlation with the
814 variables examined (pH, DO, and eels' activity level). Examples of community structure in
815 the NMDS surface. For several points within the NMDS surface (panel **a)**, genus-level
816 taxonomic compositions are shown. The example points are ordered along the vector
817 representing high eels' activity level. **c,** Indicator genera. The vectors representing the relative
818 abundance of *Cetobacterium*, *Flavobacterium*, and *Edaphobaculum*, which were highlighted
819 in the main text, are shown.
820



821

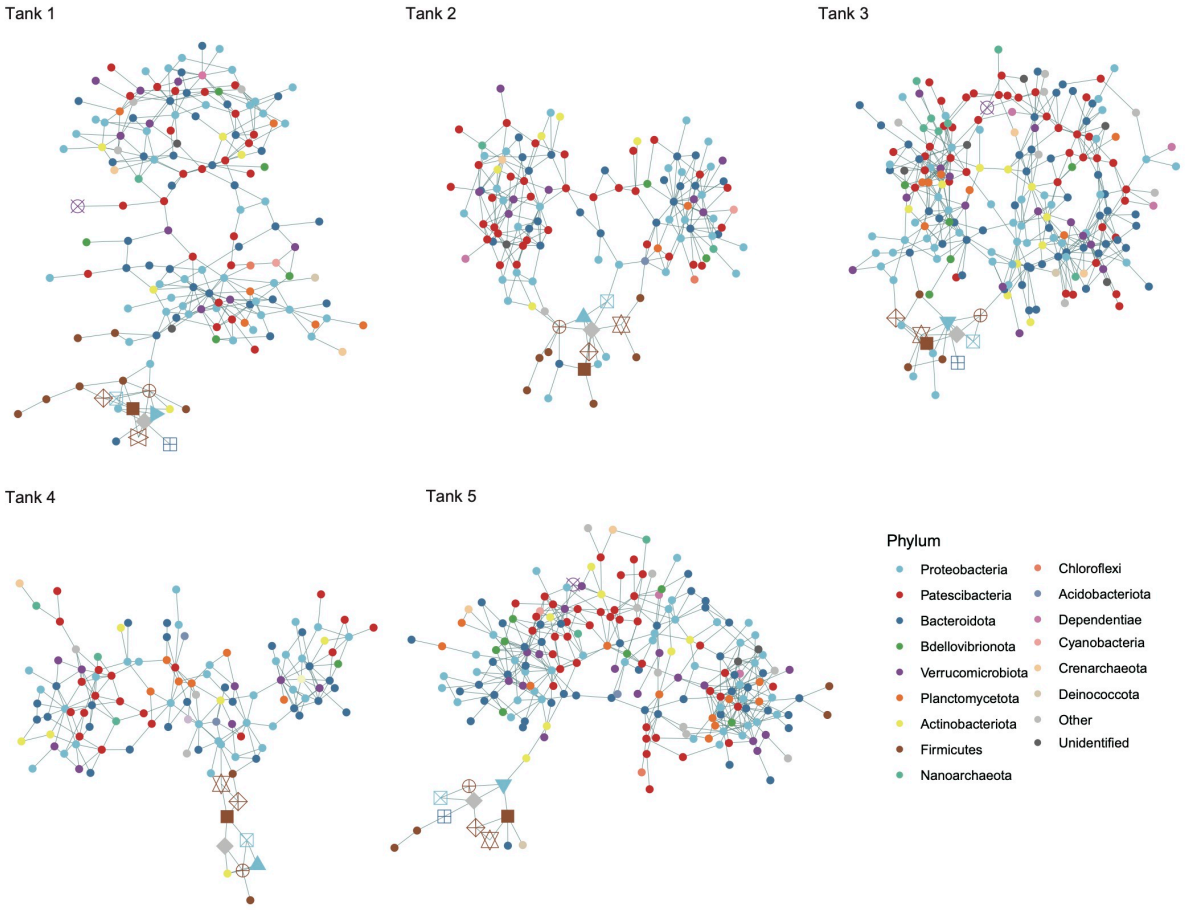
822 **Extended Data Fig. 5 | ASV-level comparison of correlation with environmental**
823 **variables and eels' activity level**. **a**, Correlation with pH. Correlation with eels' activity level
824 is shown for the ASVs that appeared in all the aquaculture tanks (shown in the decreasing
825 order of mean values). The boxes and bars represent variation across tanks. **b**, Correlation
826 with DO. **c**, Correlation with eels' activity level. **d**, Partial correlation with eels' activity level
827 (controlled variable = pH). Taxonomic information is shown for the ASVs discussed in the
828 main text.

829



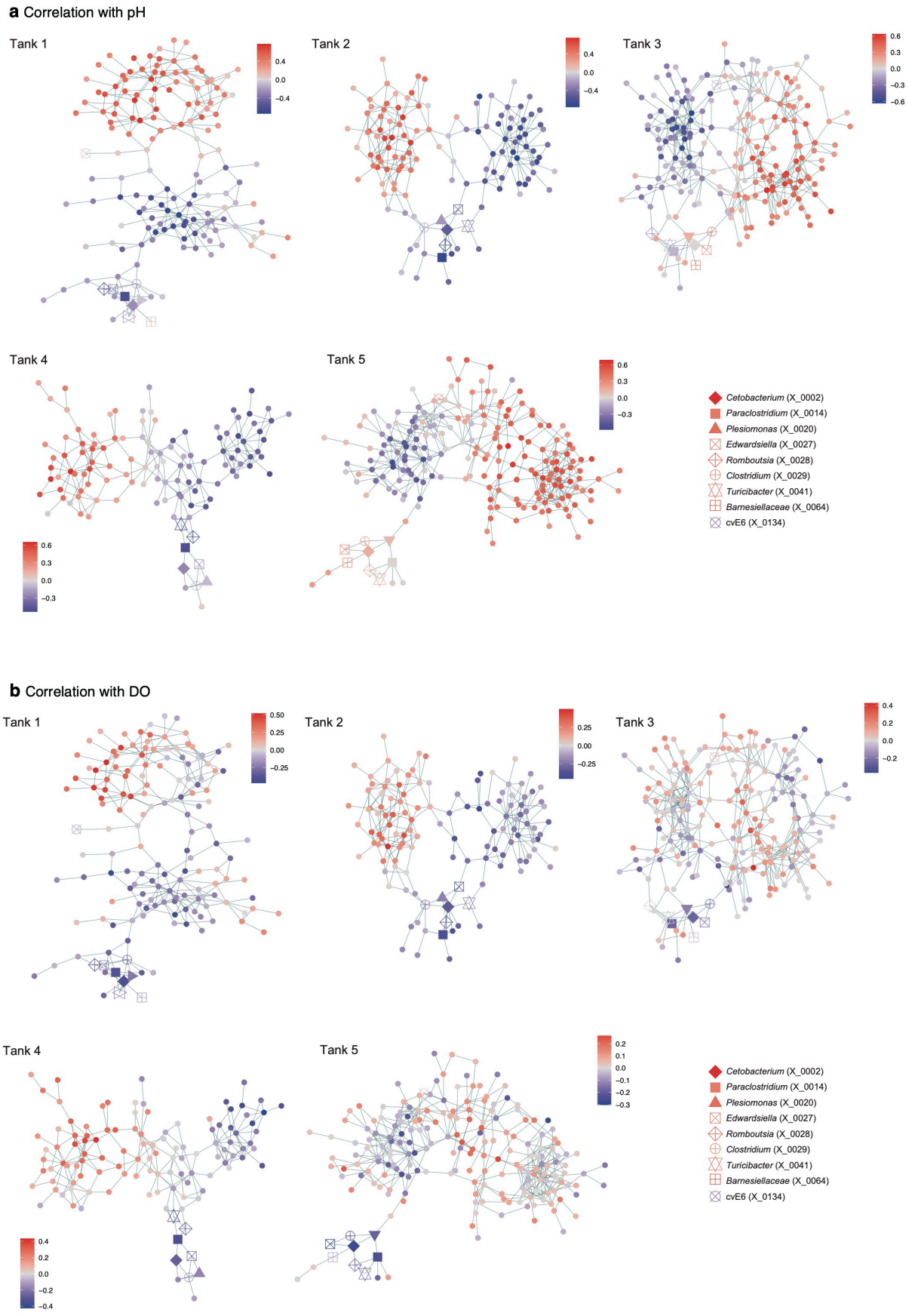
830

831 **Extended Data Fig. 6 | Phylogenetic analysis of *Edwardsiella*.** **a**, Phylogeny of
832 *Edwardsiella.* In an additional amplicon sequencing of the *sodB* gene, the neighbor-joining
833 tree of the *Edwardsiella* bacteria was reconstructed with the maximum composite likelihood
834 method. Bootstrap values larger than 70 % are shown on the nodes (1,000 permutations). The
835 pathogenic and non-pathogenic clades identified in a previous study[29] are indicated. **b**, Time-
836 series of pathogenic and non-pathogenic *Edwardsiella.* The number of detected sequencing
837 reads of the *sodB* fragment is across the time-series of each aquaculture tank.
838



839

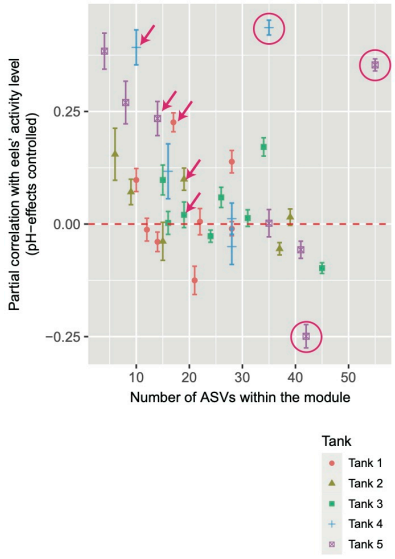
840 **Extended Data Fig. 7 | Taxonomy of the nodes within the coexistence networks.** Within
841 the coexistence networks shown in Figure 4, phylum-level taxonomy of the ASVs is shown.
842 ASVs included in minor sub-networks (number of nodes < 5) are not shown. Only the ASVs
843 that appeared in 30 or more samples were targeted in the analysis of each tank.
844



845

Extended Data Fig. 8 | Correlations with environmental variables. a, Correlation with pH. For each microbial ASV included within the coexistence network of each aquaculture tank (Fig. 4), correlation between absolute abundance and pH is shown. ASVs included in minor sub-networks (number of nodes < 5) are not shown. Only the ASVs that appeared in 30 or more samples were targeted in the analysis of each tank. **b**, Correlation with dissolved oxygen level. For each microbial ASV included within the coexistence network of each aquaculture tank (Fig. 4), correlation between absolute abundance and dissolved oxygen (DO) level is shown.

a



b

Modules including *Cetobacterium* (X_0002)

Tank	ASV ID	Phylum	Class	Order	Family	Genus	Score
1	X_0002	Fusobacteriota	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	*Cetobacterium*	0.425
	X_0016	Bacteroidota	Bacteroidia	Flavobacteriales	Weeksellaceae	*Cloacibacterium*	0.374
	X_0064	Bacteroidota	Bacteroidia	Bacteroidales	Barnesiellaceae	Unidentified	0.303
	X_0028	Firmicutes	Clostridia	Peptostreptococcales-Tissierellales	Peptostreptococcaceae	*Romboutsia*	0.286
	X_0020	Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Plesiomonas*	0.278
2	X_0002	Fusobacteriota	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	*Cetobacterium*	0.284
	X_0020	Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Plesiomonas*	0.274
	X_0272	Spirochaetota	Brevinematia	Brevinematales	Brevinemataceae	*Brevinema*	0.209
	X_0029	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	*Clostridium* sensu stricto 1	0.182
	X_0082	Bacteroidota	Bacteroidia	Chitinophagales	Saprospiraceae	*Haliscomenobacter*	0.168
3	X_0002	Fusobacteriota	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	*Cetobacterium*	0.259
	X_0020	Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Plesiomonas*	0.182
	X_0029	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	*Clostridium* sensu stricto 1	0.180
	X_0014	Firmicutes	Clostridia	Peptostreptococcales-Tissierellales	Peptostreptococcaceae	*Paraclostridium*	0.134
	X_0027	Proteobacteria	Gammaproteobacteria	Enterobacterales	Hafniaceae	*Edwardsiella*	0.107
4	X_0028	Firmicutes	Clostridia	Peptostreptococcales-Tissierellales	Peptostreptococcaceae	*Romboutsia*	0.577
	X_0041	Firmicutes	Bacilli	Erysipelotrichales	Erysipelotrichaceae	*Turicibacter*	0.540
	X_0014	Firmicutes	Clostridia	Peptostreptococcales-Tissierellales	Peptostreptococcaceae	*Paraclostridium*	0.465
	X_0005	Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	*Unidentified*	0.460
	X_0002	Fusobacteriota	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	*Cetobacterium*	0.410
5	X_0064	Bacteroidota	Bacteroidia	Bacteroidales	Barnesiellaceae	*Unidentified*	0.468
	X_0027	Proteobacteria	Gammaproteobacteria	Enterobacterales	Hafniaceae	*Edwardsiella*	0.379
	X_0002	Fusobacteriota	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	*Cetobacterium*	0.359
	X_0028	Firmicutes	Clostridia	Peptostreptococcales-Tissierellales	Peptostreptococcaceae	*Romboutsia*	0.323
	X_0029	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	*Clostridium* sensu stricto 1	0.288

c

Outlier modules with high mean partial correlation with eels' activity score

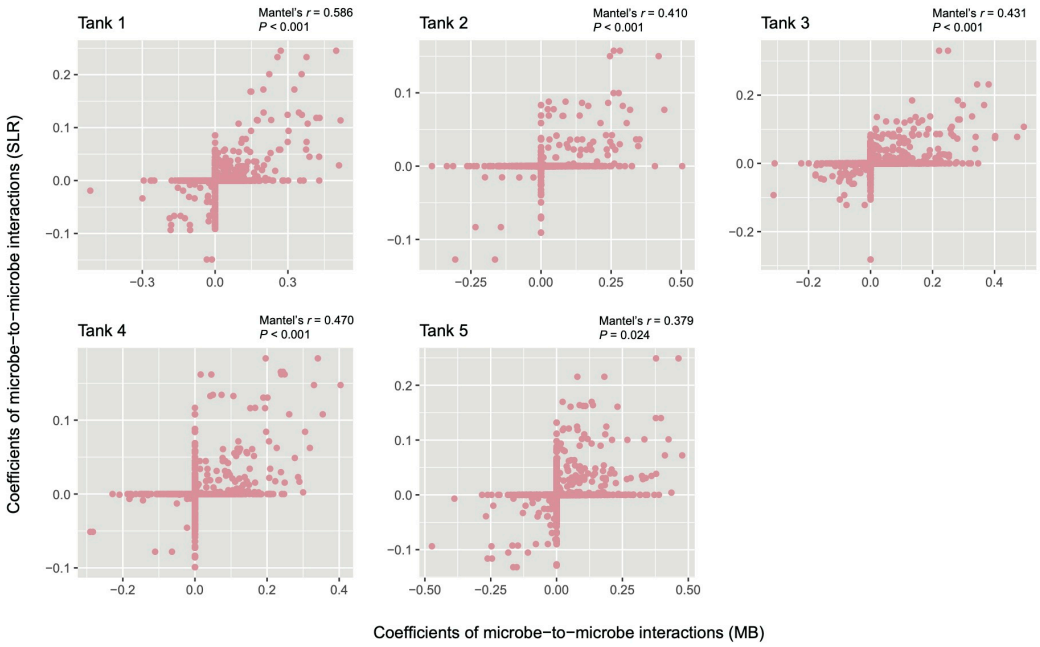
Tank	ASV.ID	Phylum	Class	Order	Family	Genus	Score
4	X_0159	Proteobacteria	Gammaproteobacteria	Burkholderiales	Comamonadaceae	*Rhizobacter*	0.588
	X_0048	Proteobacteria	Gammaproteobacteria	Steroidobacterales	Steroidobacteraceae	Unidentified	0.587
	X_0192	Bacteroidota	Bacteroidia	Cytophagales	Spirosomaceae	Unidentified	0.574
	X_0222	Nitrospirota	Nitrospiria	Nitrospirales	Nitrospiraceae	*Nitrospira*	0.570
	X_0120	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Luteimonas*	0.547
5	X_0169	Patescibacteria	Gracilibacteria	Candidatus Peribacteria	Unidentified	Unidentified	0.579
	X_0095	Patescibacteria	Saccharimonadia	Saccharimonadales	Unidentified	Unidentified	0.574
	X_0327	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	*Thermomonas*	0.518
	X_0031	Proteobacteria	Gammaproteobacteria	Burkholderiales	Burkholderiaceae	*Polynucleobacter*	0.507
	X_0087	Bacteroidota	Bacteroidia	Chitinophagales	Chitinophagaceae	Unidentified	0.498

d

Outlier module with low mean partial correlation with eels' activity score

Tank	ASV.ID	Phylum	Class	Order	Family	Genus	Score
5	X_0052	Patescibacteria	Parcubacteria	Unidentified	Unidentified	Unidentified	-0.501
	X_0075	Patescibacteria	ABY1	Candidatus Magasanikbacteria	Unidentified	Unidentified	-0.491
	X_0040	Patescibacteria	Parcubacteria	Candidatus Nomurabacteria	Unidentified	Unidentified	-0.460
	X_0297	Planctomycetota	Planctomycetes	Planctomycetales	Rubinisphaeraceae	Unidentified	-0.453
	X_0262	Patescibacteria	Microgenomatia	Candidatus Amesbacteria	Unidentified	Unidentified	-0.440

Extended Data Fig. 9 | Properties of network modules. **a**, Module size and mean partial correlation with eels' activity level. For each module within the coexistence network of each aquaculture tank (Fig. 4), the number of ASVs and mean partial correlation with eels' activity level are shown. The modules including the *Cetobacterium* ASV (X_0002) is indicated by arrows. The outlier modules with large numbers of constituent ASVs and low/high mean partial correlation with eels' activity level are highlighted by circles. **b**, Modules including the *Cetobacterium* ASV (X_0002). The top-five ASVs with the highest partial correlation with eels' activity level are shown for each module. **c**, Outlier modules with high mean partial correlation with eels' activity level. **d**, Outlier module with low mean partial correlation with eels' activity level.



867

868 **Extended Data Fig. 10 | Comparison of network reconstruction methods**. For each
869 aquaculture tank, the network links inferred with the MB method was compared with those
870 inferred with the SLR methods. The former is expected to represent interspecific interactions
871 as well as potential sharing of environmental preference (i.e., niches) between nodes (ASVs).
872 Meanwhile, the latter is expected to represent direct interactions between nodes. A
873 positive/negative value indicates a potentially positive/negative interaction between a pair of
874 microbial ASVs. The positive values (> 0) were used to draw networks of potential positive
875 interactions between microbes as shown in Figure 5.