ORIGINAL ARTICLE

Distinct Succession Patterns of Abundant and Rare Bacteria in Temporal Microcosms with Pollutants

Shuo Jiao, Yantao Luo, Mingmei Lu, Xiao Xiao, Yanbing Lin, Weimin Chen#, Gehong Wei#

State Key Laboratory of Crop Stress Biology in Arid Areas, College of Life Sciences,
Northwest A&F University, Yangling, Shaanxi 712100, P. R. China

Running Head: Bacterial Succession in Temporal Microcosms

#Address correspondence to Gehong Wei and Weimin Chen

Postal address: State Key Laboratory of Crop Stress Biology in Arid Areas, College of Life
Sciences, Northwest A&F University, Yangling Shaanxi 712100, P. R. China

E-mail: weigehong@nwsuaf.edu.cn, chenwm029@nwsuaf.edu.cn

Tel: +86 29 87091175

Fax: +86 29 87091175

Word count for the abstract: 213 words

Word count for the text: 4253 words (without reference list and figure legends)

ABSTRACT

Elucidating the driving forces behind the temporal dynamics of abundant and rare microbes is essential for understanding the assembly and succession of microbial communities. Here, we explored the successional trajectories and mechanisms of abundant and rare bacteria via soil-enrichment subcultures in response to various pollutants (phenanthrene, *n*-octadecane, and CdCl₂) using time-series Illumina sequencing datasets. The results reveal different successional patterns of abundant and rare sub-communities in eighty pollutant-degrading consortia and two original soil samples. A temporal decrease in α -diversity and high turnover rate for β -diversity indicate that deterministic processes are the main drivers of the succession of the abundant sub-community; however, the high cumulative species richness indicates that stochastic processes drive the succession of the rare sub-community. A functional prediction showed that abundant bacteria contribute primary functions to the pollutant-degrading consortia, such as amino acid metabolism, cellular responses to stress, and hydrocarbon degradation. Meanwhile, rare bacteria contribute a substantial fraction of auxiliary functions, such as carbohydrate-active enzymes, fermentation, and homoacetogenesis, which indicates their roles as a source of functional diversity. Our study suggests that the temporal succession of microbes is mainly associated with abundant bacteria rather than the high proportion of rare taxa. The major forces driving microbial succession (i.e., stochastic or deterministic processes) are dependent on the low- or high-abundance community members.

Keywords: Rare bacteria; Microbial succession; Temporal turnover; Species-time relationships; Stochasticity

INTRODUCTION

The investigation of community successional dynamics has been a central theme of ecological research; however, it has mainly been focused on animal and plant communities. Microbial communities have a high taxonomic and metabolic diversity (1, 2), and perform important ecological functions (3). Studies focused on microbial succession have become more common with the development of high-throughput sequencing (4), and to this point have mainly explored the dynamics of microbial communities in response to environmental changes in diverse systems (5-8). The temporal variability in microbial communities is directional according to environmental conditions and is consistent across similar environments (9). However, elucidating the successional trajectories and mechanisms of microbial communities remains a major challenge.

There is an ongoing debate regarding how stochastic and deterministic processes influence microbial community assembly over time (10). Deterministic processes are associated with ecological selection, where abiotic and biotic factors determine the community assembly (11). The succession of communities driven by deterministic processes is directional and converges towards a stable state, with decreasing dissimilarities amongst patches and successional rates over time (12). In contrast, stochastic processes reflect random changes in the relative abundance of species (i.e., ecological drift) and are not associated with environmentally derived fitness (13, 14). Communities structured by stochastic processes should remain divergent, with β -diversity being maintained or even increasing over time. The lack of convergence can be related to stochastic events, such as stochastic dispersal, colonization, and local extinction (15). Using time-series datasets, it is possible to estimate the

directionality, convergence, and rate of successional change to provide insights into community assembly.

In nature, most microbial communities are comprised of a large number of species, the distribution of which is often unbalanced. A few of these species are abundant; meanwhile, a large number of species, often called the “rare biosphere”, are represented by only a few individuals (16, 17). The rare biosphere is typically defined as the long tail in the rank-abundance curves used to depict diversity (17). Explorations of the rare microbial biosphere have occurred in the Arctic Ocean (18) and coastal Antarctic lakes (19), as well as coastal marine (20) and salt marsh sediments (21). These studies have shown that abundant taxa have a log-normal species-abundance distribution and rare taxa have a log-series species-abundance distribution (18). Additionally, the dispersal probability of rare bacteria is limited compared with that of abundant taxa (22). Typically, the abundant species occupy core niche positions and are strongly impacted by deterministic filtering (23); meanwhile, rare species are typically impacted more by demographic stochasticity due to their small population sizes (24). However, previous studies have typically focused on the biogeography of communities and have not examined the role of abundant and rare species in the context of successional trajectories; this lack of knowledge hinders our ability to fully understand the temporal mechanisms of microbial community assembly.

Rare taxa provide a nearly limitless reservoir of genetic and functional diversity, and serve to ensure the maintenance of microbial diversity (25). They may also hold significant roles as keystone species in microbial consortia (26) or in maintaining biogeochemical functions (27, 28). When “passing through” favorable conditions, some rare microorganisms

can become abundant community members (29). A study in the Western English Channel showed that rare microorganisms were persistent over time and some occasionally bloomed (30). Some rare community members are hypothesized to be less important for the community; for example most members of the rare biosphere in coastal sands did not become abundant at any point over the course of a year (20). The rare archaeal biosphere shows seasonal patterns similar to those of abundant microbes and cannot be considered a seed bank (31). It is clear that we still lack a complete understanding of the dynamics of abundant and rare microorganisms, as well as the forces driving their general incidence. Additionally, bio-prospecting for members of the rare biosphere may provide novel products and processes (32); however, the significance of the rare biosphere remains obscure.

In natural environments, the temporal dynamics of microbial communities are often influenced by multiple environmental factors simultaneously, which makes it difficult to determine the effect of individual factors on microbial succession (8). Microcosms are a closed system with tightly controlled conditions, and are suitable for estimating the influence of single environmental factors on microbial succession (8, 33). Our previous studies indicated that bacterial succession in microcosms is influenced by the directionality of pollutants, and exhibits temporal decays in community similarity (8). Based on the existing time-series datasets from temporal microcosms, we aimed to reveal the successional trajectories and mechanisms of abundant and rare community members. The datasets examined the response of the microbial community composition to various pollutants via high-throughput sequencing of the 16S rRNA gene. Specifically, we addressed the following questions: (1) do abundant and rare taxa show distinct succession patterns and converge

towards a similar community state over time in microcosms; (2) what are the main forces (i.e., stochastic or deterministic) driving the microbial succession of abundant and rare sub-communities, and; (3) what are the potential functions of abundant and rare sub-communities?

MATERIALS AND METHODS

Datasets. The microbial time series used in this study have been previously published (8). The original soils were collected from two oil refineries (108°46'09" E and 34°21'35" N; 108°50'10" E and 37°35'35" N) in the northwestern area of China. The sampling sites have been continuously contaminated with crude oil for the last 20 years. The total petroleum hydrocarbons were as high as $5.69 \pm 0.01 \text{ g kg}^{-1}$ of dry soil, and the physicochemical properties of the two soil samples were markedly different (data not shown).

Illumina V4-V5 16S rRNA datasets were obtained from eighty pollutant-degrading consortia (8); these were generated by ten batches of soil-enrichment microcosms from two original soils corresponding to four treatments: (1) 500 mg/L phenanthrene (PHE); (2) 500 mg/L *n*-octadecane (C18); (3) 250 mg/L phenanthrene and 250 mg/L *n*-octadecane (PC), and; (4) 250 mg/L phenanthrene, 250 mg/L *n*-octadecane, and 50 mg/L CdCl₂ (PCC). Successive subcultures were performed at 10-day intervals over a 100-day incubation period. Cells from the consortium of each subculture were collected for DNA extraction.

Both phenanthrene and *n*-octadecane are components of petroleum and can be used by bacteria as carbon source in these microcosms. Cadmium (Cd) was used because it is a heavy metal that is highly toxic to animals, plants, and microorganisms (34). The pollutant

concentrations used in this experiment are common in oil-contaminated soils, as reported in previous studies (34-36). The tightly regulated microcosms were used as a model to understand the complex interactions between environmental factors and microbial communities, and ensured that succession was caused by the pollutants (8).

Sequence analyses. The bacterial community in each of eighty pollutant-degrading consortia (two original soils $\times$ ten batches $\times$ four treatments) and two original soils samples was characterized by sequencing a portion of V4-V5 hypervariable region of the 16S rRNA gene (37) on an Illumina MiSeq (250-bp paired-end reads). The data were derived from PCR amplification of environmental DNA using primer pairs designed to amplify the gene region from all, or nearly all, known bacterial taxa. The acquired sequences were quality filtered (38) and chimeras were removed in USEARCH using the UCHIME algorithm (39). The high quality sequences from all samples were clustered into OTUs at 97% sequence similarity using the “uclust” model, which searches and clusters orders of magnitude faster than BLAST. The OTUs that contained <2 reads were not used to avoid possible biases. The representative sequences for each OTU were assigned to taxonomic groups using the RDP classifier at an 80% confidence threshold.

Data analyses. The same number of sequences from each sample (15,000) was used to avoid potential side effects based on sample size. Rarefaction curves for each individual samples were performed to estimate the sequence coverage. In this study, we defined OTUs as “abundant” when they had relative abundances above 0.1% of total sequences or “rare” when their abundances were below 0.1% (B. J. Campbell, et al. (40), L. Alonso-Saez, et al. (41)). OTU richness and Shannon index were calculated using three iterations (step size = 100

sequences per sample) for analysis of α -diversity. The Bray–Curtis similarity matrix is considered to be one of the most robust similarity coefficients for ecological studies (42), and was used to analyze the relative abundance of bacterial OTUs. A NMDS ordination was used to investigate differences between the whole, abundant, and rare bacterial communities. Mantel’s test was used to calculate the correlations between the similarity matrices of the whole, abundant, and rare bacterial communities; this test also determines significance.

For each treatment, STRs were constructed based on the degree of richness using the moving window (43). Briefly, each time series was partitioned into as many subsets (windows) as possible based on the number of observations and the STR model (i.e., the power law relationship between time and richness) was fitted for each window. We estimated the STRs for the whole, abundant, and rare bacterial communities. Changes in β -diversity between adjacent batches along the succession for the whole, abundant and rare bacterial communities were calculated using the Bray–Curtis distance matrices.

To investigate the potential functions of abundant and rare taxa, 16S rRNA gene amplicon data were analyzed using Tax4Fun to predict functional capabilities of bacterial communities (44). Identified microbial gene families (specified by KEGG Orthology groups) were screened to the FOAM database and grouped into different FOAM levels (45). To show the functional differences between abundant and rare taxa, we combined the FOAM level 1 matrices of the abundant and rare sub-communities and performed redundancy analysis by constraining the category of abundant and rare samples. The functional differences in FOAM level 1 and level 2 between the abundant and rare sub-communities were estimated by Wilcoxon rank-sum tests (46).

All statistical analyses were performed in the R environment (<http://www.r-project.org>) using the vegan (47), gplots (48), agricolae (49), car (50), or Tax4Fun (44) packages.

RESULTS

Distributions of abundant and rare bacterial taxa in the pollutant-degrading consortia. After the quality filtering and the removal of chimeric sequences, 3,281,843 high quality sequences from the entire sequencing dataset of all the eighty microcosm consortia and the two original soil samples were clustered into 41,354 operational taxonomic units (OTUs, 97% similarity) (see **Table S1** in supplementary material). The number of sequences per sample ($n = 82$) ranged from 70,838 to 15,006. Rarefaction curves suggest that the majority of the bacterial taxa were recovered (see **Fig. S1** in supplementary material). The abundant taxa generally accounted for a significantly lower proportion of OTUs (mean = 2.8%) and a larger proportion of average relative abundance (mean = 80.5%) in each sample compared with the rare taxa (mean = 97.2% and 19.5%, respectively).

To explore persistent bacterial taxa, any OTUs present in ten consortia samples from each treatment were denoted as generalists. There were between 86–109 (average proportion: 5.6%) generalist OTUs in each treatment, with average relative abundances between 70.2–85.5%; the abundance of the generalist OTUs significantly increased with time (see **Fig. S2** in supplementary material). Among the generalists, there were between 14–28 abundant taxa and 21–80 rare taxa (**Fig. 1**); this demonstrates their persistence during the enrichments. These generalist and abundant taxa accounted for 56.9% of sequences; meanwhile, the generalist and rare taxa accounted for only 0.9% of sequences. We also estimated the local occupancy of

abundant and rare taxa across ten batches of each microcosm consortia (see **Fig. S3** in supplementary material). The number of rare taxa OTUs decreased with increasing local occupancies; meanwhile, the number of abundant taxa OTUs did not change with local occupancies. This indicates that the distribution patterns of abundant and rare taxa were distinct.

The composition of the abundant and rare sub-communities differed: *Gammaproteobacteria*, *Actinobacteria*, *Spartobacteria*, and *Planctomycetia* were most prevalent abundant taxa, and; *Sphingobacteriia*, *Alphaproteobacteria*, *Anaerolineae*, and *Bacilli* were most prevalent rare taxa (see **Fig. S4** in supplementary material). The rare taxa contained more taxonomic groups than the abundant taxa.

Succession of abundant vs. rare sub-communities. The rare sub-community had higher OTU richness and Shannon-Weiner values than the abundant sub-community ($P < 0.01$; Wilcoxon rank-sum). This suggests that rare biosphere contained a more diverse taxonomy. The α -diversity of the abundant taxa significantly decreased with time (**Fig. 2A and 2B**); meanwhile, the α -diversity of the rare taxa was not correlated with time (**Fig. 2C and 2D**). During enrichment, the proportion of OTUs and the relative abundance of the rare taxa significantly increased and decreased, respectively (see **Fig. S5** in supplementary material); the opposite trends were observed for the abundant taxa. This indicates that a decreasing number of taxa were enriched and dominated the consortia during enrichment.

Nonmetric dimensional scaling (NMDS) revealed that the structure of the abundant bacterial sub-community was more similar to the overall (i.e., whole) community than was the rare sub-community (**Fig. 3**); this was confirmed by mantel tests (abundant vs. whole: $r =$

0.9934, $P < 0.01$; rare vs. whole: $r = 0.7629$, $P < 0.01$). The β -diversity was calculated based on Bray-Curtis distance for different sub-communities. The results showed that the β -diversity of the rare sub-community was significantly higher than that of the abundant sub-community. The whole community had the lowest β -diversity (see **Fig. S6** in supplementary material).

Species richness was estimated by species-time relationships (STRs). The STRs were significant for different sub-communities of all microcosm consortia ($P < 0.05$), indicating an increase in the cumulative species richness over time (**Table 1**). For each treatment, the STR exponents of the rare sub-community were higher than those of the abundant sub-community; they were also marginally higher than those of the whole community. This might indicate that rare taxa contributed most of species accumulation to the whole community.

Moreover, plotting the changes in community structure from one batch to the next revealed that the turnover of different sub-communities significantly decreased ($P < 0.001$) with time (**Fig. 4**). The slope of temporal turnover (w) of the abundant sub-community was steeper than that of the rare sub-community, which indicates a more stable community structure over time for the abundant sub-community. Additionally, the slope of temporal turnover (w) of the whole community was closer to that of the abundant sub-community than the rare sub-community, suggesting a similar succession rate between the whole and abundant sub-communities.

To further explore the succession of abundant and rare sub-communities, the temporal dynamics of the top five most abundant phyla were estimated for different sub-communities (see **Fig. S7** in supplementary material). Based on our previous study(8), the ten-stage enrichment process can be divided into three phases: I (subcultures 1–3), II (subcultures 4–7),

and III (subcultures 8–10). We found similar dynamics of the phyla in the abundant and whole communities, which differed from the dynamics of the phyla in the rare sub-community. This might suggest that the abundant sub-community was the largest driver of succession and temporary turnover for the whole community.

Potential functions of abundant vs. rare taxa. A functional ontology provided by FOAM (Functional Ontology Assignments for Metagenomes) was used to classify gene functions relevant to environmental microorganisms. Based on FOAM level 1, the abundant and rare taxa formed distinct clusters in a redundancy analysis (RDA), suggesting that they possessed different functional profiles (**Fig. 5A**). The functions related to amino acid metabolism, cellular response to stress, hydrocarbon degradation, and fatty acid oxidation were more prevalent in the abundant taxa; meanwhile, carbohydrate-active enzymes, fermentation, homoacetogenesis, methylotrophy, Embden Meyerhof–Parnos, and gluconeogenesis were more prevalent in the rare taxa. These differences were significant between the abundant and rare taxa ($P < 0.05$, Wilcoxon rank-sum test) (**Fig. 5B**).

To further investigate the functions associated with environmental pollution, we selected a functional subset (FOAM level 2) of cellular response to stress and hydrocarbon degradation, and examined the differences between the abundant and rare sub-communities (see **Fig. S8** in supplementary material). These traits, including cellular response to stress, response to oxidative stress, polycyclic aromatic hydrocarbon degradation, and aromatic carboxylic acid degradation, were significantly more prevalent in the abundant sub-community. The results indicate that primary features associated with environmental pollution and nutrition metabolism were more active in the abundant taxa compared with the

264 rare taxa.

265 To explore the functional succession of the abundant and rare sub-communities, we
266 estimated the dynamics of functional traits based on FOAM level 2 for different
267 sub-communities. The Spearman correlations of every functional trait were calculated
268 between the abundant and rare sub-communities, as well as with the whole community (see
269 **Fig. S9** in supplementary material). The proportions of functional traits with significant
270 correlations ($P < 0.01$), as well the correlation coefficients, were larger in the abundant
271 sub-community compared with the rare sub-community; this suggests that the abundant
272 sub-community was more similar to the whole community in regards to functional succession.

274 **DISCUSSION**

275 Revealing the mechanisms of microbial succession is critical for understanding the assembly
276 of microbial communities, which is a central issue in microbial ecology (9). Furthermore,
277 elucidating the temporal dynamics of abundant and rare microbes is helpful in understanding
278 successional trajectories, convergence, and temporal turnover in community assembly. In an
279 effort to determine the successional trajectories and mechanisms of the abundant and rare
280 microbial biospheres, we conducted a meta-analysis of bacterial community structure in
281 temporal microcosms with pollutants. The results indicate that the succession of abundant
282 sub-community is mainly driven by deterministic processes, while stochastic processes drive
283 the succession of the rare sub-community. The temporal succession of bacterial community in
284 microcosms is mainly associated with abundant bacteria rather than rare taxa. This finding
285 suggests that the forces driving microbial succession (i.e., stochastic or deterministic

processes) are dependent on the low or high abundance community members.

New molecular tools and increased sampling efforts have confirmed the existence of the “rare biosphere” (17, 19, 31, 51). The abundance of a given bacterial species is the result of a balance between its growth rate and loss factors (17). In the present study, the high proportion and α -diversity of rare taxa suggest that the rare biosphere is an important contributor to α -diversity (52). Within the microcosms, the organic and inorganic pollutants appear to generate a more complex environment with a stronger selection pressure. The amount of persistent rare taxa in each consortium indicates that they can survive at low abundances under selection pressure; this may be because rare species can benefit from active losses via both viral lysis and predation (53). Rare taxa consist of dormant microorganisms that can be resuscitated under different environmental conditions (53). R. Logares, et al. (19) found a substantial amount of metabolically active lineages in the rare biosphere; this suggests that the rare sub-community constitutes a diversity reservoir that can respond rapidly to environmental change. In the current study, we found that between 14–28 abundant OTUs were persistent in each pollutant-degrading consortium and accounted for a substantial fraction of sequences.

The persistent abundant taxa may play important roles in the degradation of organic pollutants that serve as the sole energy and carbon source in the microcosms. In particular, several other persistent taxa were abundant in one batch but rare in other batches, indicating large oscillations in abundance during culture enrichment. This is similar to the fluctuation patterns seen in a study examining the temporal dynamics of microbial communities (12). The local distributions of the abundant and rare taxa were different (see **Fig. S4** in supplementary

material), which may be a result of their distinct life strategies; various species have different life habits and occupy different ecological niches (54). For instance, *Gammaproteobacteria*, which are reported to be oligotrophic (55), and *Spartobacteria*, an anaerobic bacterium (56), were most often abundant within the microcosms; meanwhile, *Alphaproteobacteria* and *Bacilli*, which are phototrophic (57) and aerobic thermophilic (58) bacterium, respectively, were most often rare.

The decrease in α -diversity of the abundant sub-community during the succession indicated that selection drove the predictable assembly of the abundant taxa, which was determined by deterministic processes (10). Meanwhile, the α -diversity of rare taxa did not change temporally, suggesting that random succession (stochastic process) drove this sub-community. Abundant species can occupy a wide variety of niches, competitively utilize an array of resources, and effectively adapt to the environment; this supports the notion that they may be strongly influenced by deterministic filtering (23). On the other hand, rare species may be more strongly influenced by demographic stochasticity due to their small population sizes (24). Our previous study showed that α -diversity of the whole bacterial community decreased during enrichment subculture (8), suggesting that deterministic processes are the primary drivers of whole community succession. Furthermore, the similar trends in α -diversity variation between the abundant sub-community and the whole community indicate that the microbial succession of the whole community can be primarily attributed to the abundant bacteria.

In the microcosms, the rare sub-community possessed higher α - and β -diversity than the abundant sub-community. The high β -diversity of this sub-community can be attributed to

new microbial inputs (i.e., immigration) (59). The higher rate of stochastic immigration in the rare sub-community may be due to the limited capacity of each community. Limited community sizes could effectively make local competition more neutral and less predictable, leading to more divergent community structures (24). Additionally, exponents of STRs can estimate the rate at which new taxa are observed in a community over time; a higher exponent means that a higher number of newly introduced taxa are present. In the tightly controlled microcosms, the newly detected taxa may be derived from taxa that are present at low levels in the original inoculum. A previous study demonstrated that the variations in community composition are due to changes in the relative abundance of taxa, rather than extinction and recolonization (30). The higher rate of cumulative species richness in the rare sub-community could also indicate intrinsically high and stochastic microbial inputs (**Table 1**). Likewise, these results support the effect of stochastic processes on the succession of rare bacteria. Moreover, the higher rate of temporary turnover in the abundant sub-community than in the rare sub-community indicates a more stable community structure in the abundant bacteria as succession proceeds (**Fig. 4**). This also suggests that deterministic processes contribute more to the succession of the abundant sub-community than they do to the rare sub-community.

The succession of the microbial community can be further elucidated by examining the partitioned temporal dynamics of the abundant and rare sub-communities. We found that the exponents of STRs for rare sub-communities were similar to those of whole communities, indicating that rare taxa accounted for most of the species accumulation across each whole community. This could be because rare bacteria are an important contributor to α -diversity (52) and act as a seed bank for community stability that can respond rapidly to environmental

change (19). The structure of the abundant sub-community was also similar to that of the whole community (**Fig. 3**). Additionally, the abundant sub-community accounted for most of the succession and temporal turnover in the whole community. Studies have suggested that abundant species can occupy wide niches and competitively utilize an array of resources (23), and that community succession is dependent on species abundance (60). The variations in dominant taxa are responsible for a large fraction of overall community β -diversity (29); thus, our results confirm that abundant and rare bacteria play distinct roles in the succession and species accumulation of microbial communities within microcosms.

Microbial communities are the core contributors to ecosystem function (61), especially in the pollutant-degrading consortia in this study. The functions related to cellular response to stress and hydrocarbon degradation were prevalent in the abundant sub-community. These functions can be enriched and become abundant within the consortia, as organic pollutants (i.e., *n*-octadecane and phenanthrene) serve as sole energy and carbon source in the microcosms and inorganic pollutants (e.g., cadmium) exert selection pressures on bacteria. Additionally, vital functions associated with energy synthesis and metabolism, such as amino acid metabolism and fatty acid oxidation, were prevalent in the abundant sub-community. These results suggest that the abundant bacteria are likely responsible for primary ecosystem functions in the consortia. This is also supported by the fact that the abundant sub-community was more similar to the whole community in functional succession compared with the rare sub-community (see **Fig. S9** in supplementary material). In contrast, the rare sub-community contributed a substantial fraction of auxiliary functions, such as carbohydrate-active enzymes, fermentation, and homoacetogenesis. This could contribute to the functional redundancy of

the community (i.e., the insurance hypothesis) and enable the ecosystem to respond to and counteract disturbances (62). Studies have revealed that the rare microorganisms may mediate ecosystem function and stability. In the Arctic Ocean, the rare microorganisms exhibit biogeography, which indicates that some rare taxa, similar to more abundant taxa, are distributed based on their ecological requirements (18). Additionally, nitrification-mediated pathways are facilitated by ammonia-oxidizing archaea through aerobic ammonia oxidation (63), and methane production occurs through an anaerobic process (i.e., methanogenesis) by methanogenic archaea (64); these archaea are often present in low abundance. In this case, the rare microbial biosphere may act as a diverse pool and play a role in maintaining ecosystem functions.

CONCLUSION

Our results show the distinct succession patterns of abundant and rare sub-communities in microcosms in response to various pollutants. The succession of the abundant sub-community is mainly driven by deterministic process; meanwhile, the intrinsically high and stochastic microbial inputs of the rare sub-community indicate the major roles of stochastic processes in succession. In particular, the abundant bacteria are the largest contributors to temporary microbial turnover, while the rare taxa are primarily responsible for species accumulation during microbial succession. Abundant bacteria likely contribute primary functions in the consortia, while the rare microbial biosphere may act as a source of functional diversity; therefore, both the abundant and rare sub-communities are essential to microbial community succession, and should be distinguished in future studies.

396

397 **CONFLICT OF INTEREST**

398 The authors declare no conflicts of interest.

399

400 **FUNDING INFORMATION**

401 This work was funded by the National Science Foundation of China (31570493 and 31270529)
402 and the Cheung Kong Scholars Program (T2014208).

403

REFERENCE

1. **Schloss PD, Handelsman J.** 2007. The last word: books as a statistical metaphor for microbial communities. *Annu Rev Microbiol* **61**:23-34.
2. **Prosser JI, Bohannon BJ, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, Green JL, Green LE, Killham K, Lennon JJ.** 2007. The role of ecological theory in microbial ecology. *Nat Rev Microbiol* **5**:384-392.
3. **Falkowski PG, Fenchel T, Delong EF.** 2008. The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**:1034-1039.
4. **Gonzalez A, King A, Robeson II MS, Song S, Shade A, Metcalf JL, Knight R.** 2012. Characterizing microbial communities through space and time. *Curr Opin Biotech* **23**:431-436.
5. **Kazemi S, Hatam I, Lanoil B.** 2016. Bacterial community succession in a high-altitude subarctic glacier foreland is a three-stage process. *Mol Ecol* **25**:5557-5567.
6. **Brown SP, Jumpponen A.** 2014. Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. *Mol Ecol* **23**:481-497.
7. **Mikkonen A, Lappi K, Wallenius K, Lindström K, Suominen L.** 2011. Ecological inference on bacterial succession using curve-based community fingerprint data analysis, demonstrated with rhizoremediation experiment. *FEMS Microbiol Ecol* **78**:604-616.
8. **Jiao S, Chen W, Wang E, Wang J, Liu Z, Li Y, Wei G.** 2016. Microbial succession in response to pollutants in batch-enrichment culture. *Sci Rep* **6**.
9. **Shade A, Caporaso JG, Handelsman J, Knight R, Fierer N.** 2013. A meta-analysis of changes in bacterial and archaeal communities with time. *ISME J* **7**:1493-1506.
10. **Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF.** 2015. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc Natl Acad Sci USA* **112**:E1326-E1332.
11. **Vellend M.** 2010. Conceptual synthesis in community ecology. *The Quarterly review of biology* **85**:183-206.
12. **Anderson KJ.** 2007. Temporal patterns in rates of community change during succession. *The American Naturalist* **169**:780-793.
13. **Chase JM, Myers JA.** 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philos Trans R Soc Lond B Biol Sci* **366**:2351-2363.
14. **Hubbell SP.** 2001. *The unified neutral theory of biodiversity and biogeography (MPB-32)*, vol 32. Princeton University Press.
15. **Chase JM.** 2003. Community assembly: when should history matter? *Oecologia* **136**:489-498.
16. **Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, Arrieta JM, Herndl GJ.** 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci USA* **103**:12115-12120.
17. **Pedrés-Alió C.** 2012. The rare bacterial biosphere. *Annu Rev Mar Sci* **4**:449-466.
18. **Galand PE, Casamayor EO, Kirchman DL, Lovejoy C.** 2009. Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci USA* **106**:22427-22432.
19. **Logares R, Audic S, Bass D, Bittner L, Boutte C, Christen R, Claverie J-M, Decelle J, Dolan JR, Dunthorn M.** 2014. Patterns of rare and abundant marine microbial eukaryotes. *Curr Biol* **24**:813-821.
20. **Gobet A, Böer SI, Huse SM, van Beusekom JE, Quince C, Sogin ML, Boetius A, Ramette A.** 2012. Diversity and dynamics of rare and of resident bacterial populations in coastal sands. *ISME J* **6**:542-553.
21. **Bowen JL, Morrison HG, Hobbie JE, Sogin ML.** 2012. Salt marsh sediment diversity: a test of the variability of the rare biosphere among environmental replicates. *ISME J* **6**:2014-2023.

- 448 22. **Nemergut DR, Costello EK, Hamady M, Lozupone C, Jiang L, Schmidt SK, Fierer N, Townsend AR,**
449 **Cleveland CC, Stanish L.** 2011. Global patterns in the biogeography of bacterial taxa. *Environ Microbiol*
450 **13**:135-144.
- 451 23. **Umaña MN, Zhang C, Cao M, Lin L, Swenson NG.** 2015. Commonness, rarity, and intraspecific
452 variation in traits and performance in tropical tree seedlings. *Ecol Lett* **18**:1329-1337.
- 453 24. **Orrock JL, Watling JL.** 2010. Local community size mediates ecological drift and competition in
454 metacommunities. *Proceedings of the Royal Society of London B: Biological Sciences* **277**:2185-2191.
- 455 25. **Jones SE, Lennon JT.** 2010. Dormancy contributes to the maintenance of microbial diversity. *Proc Natl*
456 *Acad Sci USA* **107**:5881-5886.
- 457 26. **Sauret C, Séverin T, Vétion G, Guigue C, Goutx M, Pujo-Pay M, Conan P, Fagervold SK, Ghiglione J-F.**
458 2014. 'Rare biosphere' bacteria as key phenanthrene degraders in coastal seawaters. *Environ Pollut*
459 **194**:246-253.
- 460 27. **Leininger S, Urich T, Schloter M, Schwark L, Qi J, Nicol G, Prosser J, Schuster S, Schleper C.** 2006.
461 Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**:806-809.
- 462 28. **Sjöstedt J, Koch-Schmidt P, Pontarp M, Canbäck B, Tunlid A, Lundberg P, Hagström Å, Riemann L.**
463 2012. Recruitment of members from the rare biosphere of marine bacterioplankton communities after
464 an environmental disturbance. *Appl Environ Microb* **78**:1361-1369.
- 465 29. **Shade A, Jones SE, Caporaso JG, Handelsman J, Knight R, Fierer N, Gilbert JA.** 2014. Conditionally rare
466 taxa disproportionately contribute to temporal changes in microbial diversity. *MBio* **5**:e01371-01314.
- 467 30. **Caporaso JG, Paszkiewicz K, Field D, Knight R, Gilbert JA.** 2012. The Western English Channel contains
468 a persistent microbial seed bank. *ISME J* **6**:1089-1093.
- 469 31. **Hugoni M, Taib N, Debroas D, Domaizon I, Dufournel IJ, Bronner G, Salter I, Agogué H, Mary I, Galand**
470 **PE.** 2013. Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface
471 coastal waters. *Proc Natl Acad Sci USA* **110**:6004-6009.
- 472 32. **Reid A, Buckley M.** 2011. The rare biosphere: a report from the American Academy of Microbiology.
473 Washington, DC: American Academy of Microbiology **436**.
- 474 33. **Viñas M, Sabaté J, Guasp C, Lalucat J, Solanas AM.** 2005. Culture-dependent and-independent
475 approaches establish the complexity of a PAH-degrading microbial consortium. *Can J Microbiol*
476 **51**:897-909.
- 477 34. **Thavamani P, Megharaj M, Naidu R.** 2012. Multivariate analysis of mixed contaminants (PAHs and
478 heavy metals) at manufactured gas plant site soils. *Environ Monit Assess* **184**:3875-3885.
- 479 35. **Hassanshahian M, Emtiazi G, Cappello S.** 2012. Isolation and characterization of crude-oil-degrading
480 bacteria from the Persian Gulf and the Caspian Sea. *Mar Pollut Bull* **64**:7-12.
- 481 36. **Yang Y, Wang J, Liao J, Xie S, Huang Y.** 2015. Abundance and diversity of soil petroleum
482 hydrocarbon-degrading microbial communities in oil exploring areas. *Appl Microbiol Biot*
483 **99**:1935-1946.
- 484 37. **Yu X, Yang J, Wang E, Li B, Yuan H.** 2015. Effects of growth stage and fulvic acid on the diversity and
485 dynamics of endophytic bacterial community in *Stevia rebaudiana* Bertoni leaves. *Frontiers in*
486 *microbiology* **6**.
- 487 38. **Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R.**
488 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl*
489 *Acad Sci USA* **108**:4516-4522.
- 490 39. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R.** 2011. UCHIME improves sensitivity and speed of
491 chimera detection. *Bioinformatics* **27**:2194-2200.

40. **Campbell BJ, Yu L, Heidelberg JF, Kirchman DL.** 2011. Activity of abundant and rare bacteria in a coastal ocean. *Proc Natl Acad Sci USA* **108**:12776-12781.
41. **Alonso-Saez L, Diaz-Perez L, Moran XAG.** 2015. The hidden seasonality of the rare biosphere in coastal marine bacterioplankton. *Environ Microbiol* **17**:3766-3780.
42. **Kent M.** 2011. *Vegetation description and data analysis: a practical approach.* John Wiley & Sons.
43. **P White E, B Adler P, K Lauenroth W, A Gill R, Greenberg D, M Kaufman D, Rassweiler A, A Rusak J, D Smith M, R Steinbeck J.** 2006. A comparison of the species–time relationship across ecosystems and taxonomic groups. *Oikos* **112**:185-195.
44. **Aßhauer KP, Wemheuer B, Daniel R, Meinicke P.** 2015. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* **31**:2882-2884.
45. **Prestat E, David MM, Hultman J, Taş N, Lamendella R, Dvornik J, Mackelprang R, Myrold DD, Jumpponen A, Tringe SG.** 2014. FOAM (Functional Ontology Assignments for Metagenomes): a hidden Markov model (HMM) database with environmental focus. *Nucleic Acids Res* **42**:e145.
46. **Wilcoxon F, Katti S, Wilcox RA.** 1970. Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test. *Selected tables in mathematical statistics* **1**:171-259.
47. **Oksanen J, Blanchet FG.** 2015. *Vegan: Community Ecology Package* 2.3-0.
48. **Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S.** 2009. *gplots: Various R programming tools for plotting data.* R package version 2.
49. **De Mendiburu F.** 2014. *Agricolae: statistical procedures for agricultural research.* R package version 1:1-6.
50. **Fox J, Weisberg S.** 2010. *An R companion to applied regression.* Sage.
51. **Liu L, Yang J, Yu Z, Wilkinson DM.** 2015. The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. *ISME J.*
52. **Lynch MD, Neufeld JD.** 2015. Ecology and exploration of the rare biosphere. *Nat Rev Microbiol* **13**:217-229.
53. **Pedrés-Alió C.** 2006. Marine microbial diversity: can it be determined? *Trends Microbiol* **14**:257-263.
54. **Potapov AA, Semenina EE, Korotkevich AY, Kuznetsova NA, Tiunov AV.** 2016. Connecting taxonomy and ecology: Trophic niches of collembolans as related to taxonomic identity and life forms. *Soil Biol Biochem* **101**:20-31.
55. **Cho J-C, Giovannoni SJ.** 2004. Cultivation and growth characteristics of a diverse group of oligotrophic marine Gammaproteobacteria. *Appl Environ Microb* **70**:432-440.
56. **Qiu Y-L, Kuang X-z, Shi X-s, Yuan X-z, Guo R-b.** 2014. *Terrimicrobium sacchariphilum* gen. nov., sp. nov., an anaerobic bacterium of the class ‘Spartobacteria’ in the phylum Verrucomicrobia, isolated from a rice paddy field. *Int J Syst Evol Micr* **64**:1718-1723.
57. **Srinivas T, Kumar PA, Sasikala C, Ramana CV, Imhoff JF.** 2007. *Rhodobacter vinaykumarii* sp. nov., a marine phototrophic alphaproteobacterium from tidal waters, and emended description of the genus *Rhodobacter*. *Int J Syst Evol Micr* **57**:1984-1987.
58. **Nazina T, Tourova T, Poltarau A, Novikova E, Grigoryan A, Ivanova A, Lysenko A, Petrunyaka V, Osipov G, Belyaev S.** 2001. Taxonomic study of aerobic thermophilic *bacilli*: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenuatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. th.* *Int J Syst Evol Micr* **51**:433-446.

59. **Dini-Andreote F, e Silva MdCP, Triado-Margarit X, Casamayor EO, Van Elsas JD, Salles JF.** 2014. Dynamics of bacterial community succession in a salt marsh chronosequence: evidences for temporal niche partitioning. *ISME J* **8**:1989-2001.
60. **Li Sp, Cadotte MW, Meiners SJ, Pu Z, Fukami T, Jiang L.** 2016. Convergence and divergence in a long-term old-field succession: the importance of spatial scale and species abundance. *Ecol Lett* **19**:1101-1109.
61. **Miki T, Ushio M, Fukui S, Kondoh M.** 2010. Functional diversity of microbial decomposers facilitates plant coexistence in a plant–microbe–soil feedback model. *Proc Natl Acad Sci USA* **107**:14251-14256.
62. **Yachi S, Loreau M.** 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc Natl Acad Sci USA* **96**:1463-1468.
63. **Kool D, Wrage N, Zechmeister - Boltensern S, Pfeffer M, Brus D, Oenema O, Van Groenigen JW.** 2010. Nitrifier denitrification can be a source of N₂O from soil: a revised approach to the dual - isotope labelling method. *Eur J Soil Sci* **61**:759-772.
64. **Aydin S, Ince B, Ince O.** 2015. Application of real-time PCR to determination of combined effect of antibiotics on Bacteria, Methanogenic Archaea, Archaea in anaerobic sequencing batch reactors. *Water Res* **76**:88-98.

FIGURE LEGENDS

FIG 1 The number and relative abundance of the generalist operational taxonomic units

(OTUs) categorized into: always abundant (i.e., > 0.1% abundance in all batches), always rare

(i.e., < 0.1% in all batches), or oscillating (i.e., not falling in the other categories) for eighty

pollutant-degrading consortia and two original soils samples.

FIG 2 Changes in α -diversity of the abundant (A, B) and rare taxa (C, D) between the original soils and the successive batches. Linear regressions of operational taxonomic unit (OTUs) richness (A, C) and Shannon index (B, D).

FIG 3 Nonmetric multidimensional scaling ordination for bacterial communities from eighty pollutant-degrading consortia and two original soils samples. Whole: all bacterial taxa; Abundant: abundant taxa, and; Rare: rare taxa.

FIG 4 Changes in β -diversity, as well as the succession of the whole, abundant, and rare bacterial communities. Variations in community dissimilarity were calculated with Bray-Curtis distance between two adjacent batches. Whole: all bacterial taxa; Abundant: abundant taxa; Rare: rare taxa, and; w: the slope of regression line.

FIG 5 The differences in potential functions of abundant and rare taxa in temporal microcosms. Redundancy analysis performed on FOAM level 1 matrix of abundant and rare sub-communities by constraining the category of abundant and rare samples (A). The functional differences in FOAM level 1 between abundant and rare sub-communities were estimated by Wilcoxon rank-sum test (B). *: $P < 0.05$; Abundant: abundant taxa. Rare: rare taxa.

583 **TABLE 1** The exponents of species-time relationships for the whole, abundant, and rare
584 bacterial communities in different consortia.

Community	J-C18	J-PHE	J-PC	J-PCC	X-C18	X-PHE	X-PC	X-PCC
Whole	0.7903	0.7607	0.7983	0.7984	0.7612	0.7157	0.7701	0.7454
Abundant	0.246	0.4103	0.3687	0.4344	0.4009	0.3096	0.3794	0.2596
Rare	0.799	0.7683	0.8053	0.8042	0.7672	0.7216	0.7857	0.7514

585 “J” means consortia enriched from the soils of 108°50'10" E and 37°35'35" N; “X” means
586 consortia enriched from the soils of 108°46'09" E and 34°21'35" N; C18 = 500 mg l⁻¹
587 *n*-octadecane; PHE = 500 mg l⁻¹ phenanthrene; PC = 250 mg l⁻¹ phenanthrene + 250 mg l⁻¹
588 *n*-octadecane; and PCC = 250 mg l⁻¹ phenanthrene + 250 mg l⁻¹ *n*-octadecane + 50 mg l⁻¹
589 CdCl₂.