Unravelling the interplay of ecological processes structuring the bacterial rare biosphere

Xiu Jia (xibeihenai@gmail.com)
University of Groningen
Francisco Dini-Andreote
Pennsylvania State University
Joana Falcão Salles
University of Groningen

Research Article

Keywords: Community assembly, Species rarity, Soil, Selection, Dispersal

Posted Date: February 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1356351/v1

Unravelling the interplay of ecological processes structuring the bacterial rare biosphere

Xiu Jia[1*], Francisco Dini-Andreote[2,3], Joana Falcão Salles[1*]

[1]Microbial Ecology Cluster, Genomics Research in Ecology and Evolution in Nature (GREEN), Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Groningen, 9747AG, The Netherlands. [2]Department of Plant Science, The Pennsylvania State University, Pennsylvania, University Park, PA 16801, USA. [3]Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA 16801, USA.

* Correspondence: xibeihenai@gmail.com; j.falcao.salles@rug.nl

Abstract

Background: Although biological communities harbor many rare species, i.e., the rare biosphere, relatively little is known about how distinct ecological processes structure their existence. Here, we combined null modelling with spatiotemporal turnover in taxa rarity to model assembly processes of the rare biosphere.

Results: We found that an overall influence of homogeneous selection (i.e., spatiotemporally constant variables) in driving the assembly of the rare biosphere whereas the common biosphere is governed by a broader array of processes, from which variable selection (i.e., spatiotemporally fluctuating variables) plays the primary role. By partitioning rarity types, we further found that homogeneous selection explained the prevalence of permanently rare taxa, suggesting their persistent low abundances to be restrained by physiological traits. Variable selection explained the dynamics of conditionally rare taxa, indicating their ability to switch between rarity and commonness by responding to spatiotemporal variation.

Conclusion: This study contributes to the understanding of ecological mechanisms that maintain microbial diversity from the rare biosphere perspective. It also provides a new perspective on how to predict dynamic changes in the rare biosphere through space and time.

Keywords: Community assembly, Species rarity, Soil, Selection, Dispersal

Background

Communities are generally composed of a few highly abundant species and numerous low abundance ones, in which the latter is defined as the rare biosphere [1, 2]. In microbiomes, rare microbial species potentially play crucial roles in ecosystem functioning, for instance, by preventing pathogen spread, controlling nutrient cycling, and contributing to pollutant degradation [3-6]. However, it remains unclear how distinct ecological processes interplay in determining the assembly and successional dynamics of the different types of microbial rarity.

Species assumed to have low fitness with small populations and might undergo local extinction through ecological drift. Different mechanisms have been developed to explain the existence of the highly diverse rare species in plant and animal communities, including rarity is caused by either less competitive ability combined with frequency-dependent selection, narrow niche requirement, or limited dispersal ability [7-10]. These mechanisms are likely to apply to microorganisms [11]. For example, protozoan predation and viral lysis are thought to suppress the common bacteria and 'kill the winner' explains the existence of low abundance populations [12]. Lower competition ability, such as slow growth rate, can also lead to rarity, such as bacterial specialists occupying niches that are not explored by generalists and thus persist at low abundances [13]. However, Kurm *et al.* found that low competitive ability often cannot explain the existence of rare bacterial species, which indicates that physiological constraints do not always drive rarity [14]. Limited dispersal or species-specific (a)biotic constraints have also been reported as essential mechanisms governing bacterioplankton rarity [15]. Although these mechanisms are diverse, they are not in conflict and can operate simultaneously. Their relative importance might vary, thus contributing to different types of rarity within and across community types.

Ecological processes such as selection, dispersal, diversification, drift, and their interactions, drive the overall patterns observed in ecological communities, also known as community assembly [16-18]. We have recently proposed that by identifying the ecological processes that drive the dynamics of rare microbial communities over space and time, one could sort out the mechanisms leading to the composition of the rare microbial biosphere and the contribution of each type of rarity (see ref. [19]). The relative influence of each ecological process can be achieved stepwise by combining phylogenetic and taxonomic community structures with ecological null models (see Stegen's approach in ref. [20, 21]). This approach assumes that phylogenetically related species have more similar ecological niches than those drawn randomly from the phylogenetic tree [22]. After differentiating common from rare species, which can be realized through several methods, the first step consists of determining whether the community structure of each one of these components (rare or common) is driven by selection (either variable or homogeneous) or not (i.e., stochastic assembly) [19, 23]. This information can be obtained by comparing the phylogenetic structure of observed communities with the distribution of the phylogenetic community structures of null models (Supplementary Figure 1, Step 1). When the phylogenetic turnover is significantly higher than that of the null expectation, community assembly is

primarily mediated by variable selection, which indicates that community turnover is regulated by shifting environmental conditions over space or time. However, if the phylogenetic turnover of the community is significantly lower than that of the null expectation, community assembly is mediated by homogeneous selection, suggesting that community turnover is governed by environmental filters that are not undergoing spatiotemporal changes [21, 24]. When the phylogenetic distance between a given pair of communities is not significantly different from the null distribution, selection is absent, indicating that community turnover is primarily governed by stochastic processes such as dispersal and/or drift. These processes can be further disentangled by calculating the dissimilarity of taxonomic composition between two communities [Bray-Curtis (BC) distance] relative to the null distribution, using a Raup-Crick matrix [20, 25] (Supplementary Figure 1, Step 2). Here the supervised assumptions are the following: (*i*) high dispersal rates lead to homogeneous species distribution across communities (i.e., homogenizing dispersal; observed BC is lower than null BC); (*ii*) low dispersal causes high species turnover (i.e., dispersal limitation; observed BC is higher than null BC); and (*iii*) when both phylogenetic and taxonomic community structures do not differ from the null expectations, neither selection nor dispersal contributes significantly to community assembly. In this scenario, ecological drift and the negligible selection and dispersal effects drive community turnover and are collectively termed 'undominated processes' [21].

The abovementioned approach can be used to better understand what ecological processes lead to different types of rarity based on the abundance dynamics across space and/or time: 'permanently rare,' 'permanently rare with periodic distribution,' 'transiently rare,' 'periodic recruitment from the rare biosphere' and 'occasional recruitment from the rare biosphere' [23]. We hypothesized that distinct ecological processes influence different types of rarity, which are described below. Specifically, conditionally rare species can be found in the rare biosphere under unfavorable physicochemical or biological conditions but become occasionally or periodically abundant once local conditions become favorable [26]. Therefore, we hypothesize that conditionally rare species that are either periodic or occasionally recruited from the rare biosphere are driven by variable selection, caused by their fitness changes along environmental (spatial or temporal) gradients [19]. In contrast, permanently rare species always remain at low abundance regardless of the changing environmental factors [27]. Our hypothesis is that the existence of these species is structured by homogeneous selection [19], as the low abundances and fitness of permanently rare species might be associated with K-strategies. For example, specialized oligotrophic bacteria displaying a narrow ecological niche (and/or lower competitive ability) are expected to show minimal response to seasonal fluctuations of the labile carbon pool through the year, as their metabolisms are primarily responding to recalcitrant carbon resources that are more constant over time [28]. Permanently rare taxa that show periodic fluctuations in abundance to be determined by homogenizing dispersal process operating alongside ecological drift [19]. This might occur because high dispersal rates replenish the loss of individuals that can occasionally go extinct. We thus hypothesize that for those transiently rare taxa, dispersal limitation and local extinction through stochastic demographic processes might result in their brief appearance in a community since they

might migrate from another habitat by limited dispersal process but cannot indefinitely sustain their low levels of abundance [19]. Finally, the diversification process can also account for new species/genotypes within the rare biosphere (either by mutation and/or horizontal gene transfer) [19]. If this process results in a higher fitness, organisms will persist over time (permanently or conditionally rare); otherwise, they will be either eliminated by natural selection or drift and then constrained as transiently rare [19]. It is important to note that our analysis cannot quantitatively account for the diversification process due to inherent methodological hurdles associated with identifying new species/genotypes in natural ecosystems.

In the present study, we empirically quantified the extent to which these ecological processes govern the assembly and successional dynamics of the distinct rarity types (see ref. [19]) and the abundant subcommunities in soil bacterial communities. To test the abovementioned hypotheses, we used a dataset from an ecological gradient of primary soil succession spanning over a century on the island Schiermonnikoog, the Netherlands, which provides variation in biotic and abiotic conditions while under the influence of same species pools [29]. The profiling of the rare biosphere along this gradient at different points in time allowed us to cover enough variation of species turnover and identify the distinct types of rarity across both spatial and temporal scales. Moreover, the characterization of the common microbial biosphere allowed us to identify the mechanisms driving the assembly of common species. The putatively 'active' bacterial rare and common biospheres were characterized by sequencing the reverse transcribed bacterial 16S rRNA transcripts obtained from soil bacterial communities [29]. In this study, we first partitioned the common and rare biosphere components using three rarity cutoff values (i.e., 0.2%, 0.1%, and 0.05%) and applied the null model approach to quantify the relative influence of ecological processes [20, 21]. We focused on the results from the cutoff of 0.1% per sample as it can better represent the rare biosphere by showing the tails of rank abundance curves. As there is still a lack of a golden standard to define the rare biosphere, another two cutoffs (0.2% and 0.05%) were used to check whether a more relaxed or strict cutoff will influence the interpretation of the results. Then, we categorized and quantified the distinct types of rarity and commonness (Supplementary Figure 2) [19, 23]. Last, we provide a conceptual overview of how community assembly processes regulate different components of bacterial communities, which will serve as the basis for further hypotheses development in this emergent field of science.

Results

Study system and community profiling

The spatial and temporal changes in the composition of potentially active bacterial communities were examined by sequencing the reverse-transcribed 16S rRNA transcripts in 60 soils samples collected across five successional stages at the island Schiermonnikoog, the Netherlands, in May, July, September, and October 2017 (Fig. 1a) [29]. We obtained a rarefied feature table containing 31,500 cDNA sequences per sample, encompassing a total of 22,490 potentially active ASVs. Rarefaction curves of most of the samples reached a steady plateau. As a consequence, the estimated ASVs richness using the Chao1 index was mostly

equivalent to the observed richness, indicating a good representation of ASV richness in our dataset (Supplementary Figures 3 and 4). In accordance with previous observations [30], we showed that the structure of bacterial communities significantly changed along the successional chronosequence (PERMANOVA $R^2 = 0.45$, $P < 0.01$ for Jaccard distance and $R^2 = 0.61$, $P < 0.01$ for Bray-Curtis distance; Supplementary Figure 5 and Table 1).

Defining rare and common biospheres

To determine the most suitable cutoff value, we fitted the frequently used rarity cutoffs, i.e., 1.0%, 0.1%, and 0.01% [31-35], to the rank abundance curves in our dataset (Supplementary Figure 6B). These plots show that using 1.0% or 0.01% would either overestimate or underestimate the size of the rare biosphere in our samples (Supplementary Figure 6B). The cutoff value of 0.1% contained proper tails of rank abundance curves in the data, thus showing a good representation of the rare biosphere.

In addition, to complement this fixed threshold, we applied a "sample-specific rarity cutoff" approach, in which rarity is defined using a similar algorithm to calculate the h-index [36], where species are defined as rare when their abundances are not higher than their individual ranks in the rank abundance curve (Supplementary Figure 7). We recalibrated the cutoff by the sampling depth value, i.e., the ratio of the number of observed species (S_{obs}) to the number of estimated species (S_{chao1}), to balance the difference of sampling effort between samples. Here we used the Chao1 index to indicate estimated richness. Worth noticing, the estimated species number (Chao1) can be higher than the true species richness, as it can overestimate species numbers when the sample size is small [37]. However, since rarefaction curves seem to have reached a plateau steadily, the observed number of species was close to estimated species in most samples (Supplementary Figure 3). Although we believe the Chao1 index did not overestimate richness substantially in our dataset, the use of this index in future studies should be carefully evaluated. Using the "sample-specific rarity cutoff" approach, we identified the average of sample-specific rarity cutoffs at 0.2% (Supplementary Figure 8, 6A). Furthermore, to avoid the arbitrary definition of the rare biosphere, we tested two additional rarity cutoffs, i.e., 0.2% and 0.05%, which aligned with the tails of the relative abundance curves and represented interval deviations from the usually fixed threshold of 0.1%.

We split the dataset into two community components (the rare and common biospheres) using the abovementioned criteria, which entailed the tails of rank abundance curves in our dataset (0.2%, 0.1%, and 0.05%; Supplementary Figure 6A). Below we show the results based on cutoff 0.1% and briefly discuss whether a more relaxed (0.2%) or strict cutoffs (0.05%) provide similar results as the cutoff of 0.1%.

Bacterial composition in the rare and common biospheres

Based on a rarity cutoff of $\leq 0.1\%$ of total relative abundance per community, the richness of the rare biosphere (encompassing 22,368 ASVs) was roughly 10-fold higher than that of the common biosphere (total

of 2,319 ASVs, Fig. 1b, Supplementary Figure 9). Moreover, the rare and common biospheres differ in phylum composition: the rare biosphere encompassed a larger number of bacterial phyla (total of 44) compared with the common biosphere (total of 22, Supplementary Figure 10). Also, 22 bacterial (candidate) phyla were exclusively found within the rare biosphere. The composition of ASVs was significantly different between the rare and common biospheres (PERMANOVA $R^2 = 0.12$, $P < 0.01$; Fig. 1c, Supplementary Table 2,), and there was even distinct community composition across each successional stage (PERMANOVA $R^2 = 0.27$, $P < 0.01$). As succession proceeds, the level of divergence between the rare and common biosphere increases progressively (Fig. 1c). In line with this observation, PERMANOVA analysis revealed a significant interaction between successional stages and the community component (rare and common biospheres; PERMANOVA $R^2 = 0.20$, $P < 0.01$). It is worth mentioning that similar patterns of β-diversity were obtained using the rarity cutoff values of 0.2% and 0.05% (see Supplementary Figure 11 and Table 2).

The influence of assembly processes structuring the rare and common biospheres

Our results revealed that distinct processes mediate the structure of the bacterial rare and common biospheres. Overall, selection influenced the assembly of the rare and common biospheres, which accounted for 86.61% and 66.83%, respectively. However, the taxa turnover of the rare biosphere was governed mainly by homogeneous selection (66.67%, Fig. 2b). Variable selection and dispersal limitation also had moderate influences on the rare biosphere, i.e., 19.94% and 13.39%, respectively. On the other hand, the taxa turnover of the common biosphere was mostly governed by variable selection (47.06%). In contrast, variable selection, dispersal limitation, undominated processes, and homogenizing dispersal were found to exert less important roles (19.77%, 14.24, 13.67, and 5.25%, respectively; Fig. 2a).

The data obtained from community assembly models were further used to investigate patterns in both temporal (within each stage of succession) and spatial (across stages of succession) variations of the rare and common biospheres. Homogeneous selection dominated the temporal assembly of the rare biosphere at all soil successional stages (100%, Fig. 2b). The relative influences of homogeneous selection were greater for temporal rather than spatial turnover of both the rare and common biospheres (Fig. 2a, b). In contrast, the relative influences of variable selection were greater for spatial rather than temporal turnover of both the rare and common biospheres (Fig. 2a, b). This result illustrates the selection exerted by the different soil conditions across different successional stages, in contrast to similar conditions within each stage. Interestingly, homogenizing dispersal and undominated processes had a greater influence on temporal rather than spatial community turnover of the common biosphere (Fig. 2a). While homogenizing dispersal was more pronounced at late successional stages and displayed a progressive increase in importance as succession progressed, undominated processes had the opposite pattern (Fig. 2a). Notably, here, the ‘homogenizing

dispersal' contributing to the temporal changes at the same stage should be interpreted as the species composition was highly consistent at different months rather than migration among locations.

Similar assembly processes were identified when we used the rarity cutoff values of 0.2% and 0.05% to define the rare biosphere (Supplementary Figure 12). Specifically, the dominant roles of homogeneous selection in structuring the rare biosphere and variable selection in structuring the common biosphere were observed at all the tested rarity cutoffs, i.e., from 0.2% to 0.05% (Supplementary Figure 12). For both the rare and common biospheres, the relative importance of homogeneous selection gradually decreased as rarity cutoff value was reduced from 0.2% to 0.05%, whereas the extent of variable selection and dispersal limitation slightly increased (the relative influence of variable selection changed from 19.04% to 25.03%; the relative influence of dispersal changed from 11.53% to 16.05%).

Types of rarity and commonness

We identified a total of 11,838 permanently rare ASVs (encompassing 38 phyla), 8333 transiently rare ASVs (42 phyla), 2197 conditionally rare/common ASVs (21 phyla), and 122 permanently common ASVs (11 phyla) using the rarity cutoff of 0.1% (Supplementary Figure 14). By quantifying the abundance and richness of each type of rarity/commonness, we found that permanently rare and conditionally rare ASVs are the dominant types of rarity in the rare biosphere (Fig. 3a, b; Supplementary Table 3). In terms of relative abundance, more than half of the rare biosphere is represented by permanently rare ASVs (54.73 ± 0.85%, average ± standard error), followed by conditionally rare ASVs (41.13 ± 0.96%). This occurs even though the number of conditionally rare ASVs (437 ± 15 ASVs) is far lower than that of permanently rare ASVs (1,258 ± 53 ASVs). Last, transiently rare ASVs (139 ± 13 ASVs) only constituted 4.14 ± 0.42% of the total abundance of the rare biosphere. On the other hand, for the common biosphere, conditionally common ASVs (185 ± 4 ASVs) formed the dominant type of commonness, which encompassed 95.84 ± 0.41% of the total abundance of the common biosphere, whereas permanently common ASVs (5 ± 1 ASVs) only encompassed 4.16 ± 0.41% of the total abundance (Fig. 3a, b; Supplementary Table 3).

Notably, the composition of each type of rarity/commonness also changed when increasing the stringency of the rarity cutoff from 0.2% to 0.05%. Regarding the rare biosphere, the total relative abundance decreased as the rarity cutoff became stricter, i.e., 61.59%, 45.65%, and 28.36% at rarity cutoff values of 0.2%, 0.1%, and 0.05%, respectively. Moreover, as the rarity cutoff became stricter, the proportion of permanently and transiently rare ASVs decreased, while conditionally rare ASVs increased (Supplementary Figure 15, 16 and Table 3). Regarding the common biosphere, the proportion of conditionally common ASVs decreased, while

permanently common increased as the rarity cutoff became stricter (Supplementary Figure 15, 16 and Table 3).

Discussion

The microbial rare biosphere is crucial for maintaining ecosystem diversity and functioning. However, it remains unknown how this significant component of microbial communities is structured and which ecological processes such as selection, dispersal, drift, and diversification contribute to the dynamics of this large proportion of low abundant microbes. Using a primary succession system, where variation in soil physicochemical characteristics coexists with similar source communities, we showed that selection was the main driver of both rare and common biospheres (86.61% and 66.83%, respectively). This is in accordance with previous findings that both the rare and common biospheres show similar biogeographic patterns, i.e., species in these two sub-communities are sensitive to environmental conditions [38, 39]. Besides, this study further demonstrates that the rare and common biospheres are assembled through a distinct interplay of ecological processes.

Turnover of the rare biosphere is mainly driven by homogeneous selection

Our results reveal that an interplay of homogeneous selection, variable selection, and dispersal limitation led to higher ASV richness in the rare biosphere. Among these processes, homogeneous selection was the primary driver influencing the assembly of the rare biosphere (Fig. 4). According to our hypothesis that homogeneous selection is related to the existence of permanently rare species, we observed this type of rarity to represent the most abundant one in the rare biosphere. This suggests that a significant fraction of the rare biosphere consisting of permanently rare taxa is not driven by biotic and abiotic variation. Instead, the environmental conditions that primarily impose selection on the rare biosphere turnover can also be constant. Similarly, a study that targeted culturable low abundant bacteria found that the relative abundance of slow-growing rare taxa was not affected by changes in nutrient concentration [40]. Together, these results suggest that permanently rare ASVs with particular sets of ecological traits (e.g., low growth rate and competitive ability) can withstand environmental fluctuation. Therefore, permanently rare taxa in our system have potential selective advantages at low abundances ensuring their long-term persistence.

Variable selection was a secondary but essential process structuring the turnover of the rare biosphere (Fig. 4). Some soil physicochemical properties substantially changed throughout succession in this ecosystem or across sampling times, thus imposing variable selection [30]. For example, a significant change in the quality and quantity of soil organic matter has shown to be consistent with variable selection across successional stages [24]. In line with our hypothesis, variable selection operating through changes in environmental conditions was responsible for the dynamics of conditionally rare species, which were also a dominant type of rarity in the system. This idea is supported by Kurm *et al.,* who found that rare taxa with the ability to grow fast are affected by variation in nutrient concentration [40]. In another case, conditionally rare taxa

304 relying on labile nitrogen increased in abundance as nitrogen limitation was alleviated [41]. These examples
305 collectively confirm that species shifting between rare and common biospheres is related to environmental
306 fluctuations, i.e., variable selection can be responsible for the presence of conditionally rare species (see
307 discussion on balancing rarity and commonness and Fig. 4).
308
309 To verify the results obtained based on Stegen's approach, we further tested the dataset using Sloan's neutral
310 model [42]. However, since this model relies on species distributions, it does not work well for modeling
311 subsets of communities [19, 43]. Thereby, we only apply this model to the whole dataset rather than using it
312 to model the assembly of the rare biosphere. The model for the entire dataset explained 44% of the
313 spatiotemporal variation in the metacommunity, which is higher than that inferred from the phylogenetic-
314 based null model (<20%). By mapping rarity types on the species distribution plot, we found that distinct
315 types of rarity and commonness show different deviations from the neutral prediction (Supplementary Figure
316 13). In brief, permanently rare taxa were found more frequently than predicted by the neutral model
317 (Supplementary Figure 13), indicating the wide distribution of these taxa to not be restricted by dispersal.
318 Conditionally rare taxa were often found less frequently than expected by chance (Supplementary Figure 13),
319 suggesting they were selected by environmental conditions and only present at conditions that meet their
320 fitness. As such, results from the Sloan's model based on the entire dataset supported the results from the
321 Stegen's model, albeit the Sloan's neutral model cannot be directly applied to partition the rare biosphere.
322
323 Last, dispersal limitation also contributed to the turnover of the rare biosphere, albeit to a lower extent than
324 selection (Fig. 4). This is consistent with results from a previous study that reported that rare taxa are
325 geographically restricted [32]. Despite the differences found, the rare biosphere was governed mostly by
326 dispersal limitation [44]. As the increase of the proportion of transiently rare taxa was correlated with the rise
327 of dispersal limitation in the assembly of the rare biosphere when the rarity cutoff became stricter, we suggest
328 that dispersal limitation might be necessary for the persistence of a small fraction of the transiently rare
329 species. These transiently rare species are introduced via limited dispersal processes but cannot adapt to the
330 new abiotic or biotic conditions. For instance, immigrant bacteria are known to face colonization resistance
331 by the resident microbiome. In one study, transient food-borne bacteria were found to coexist with the
332 resident gut microbiome temporarily, but biotic competition hindered its long-term persistence within the
333 community [45]. In the studied salt marsh ecosystem, transiently rare bacteria likely resulted from restricted
334 dispersal processes, such as marine microbes carried by the tide, terrestrial microbes dispersed by the wind,
335 or microbes hitchhiking on eukaryotic organisms [46, 47]. These transiently rare species have often been
336 described in communities of macro-organisms [48]; however, the dynamics and ecological roles of this type
337 of rare species in the microbial communities remain mostly elusive. Understanding the ecological basis of
338 their assembly and dynamics will provide the initial step towards this endeavor.
339

By systematically partitioning the rare and common biospheres at three cutoff values (i.e., 0.05%, 0.1%, and 0.2%), we found that the results of community assembly are highly consistent. Notably, the relative influences of distinct ecological processes and the rarity types changed in a coordinated manner when increasing the stringency of the rarity cutoff. Specifically, the influence of homogeneous selection decreased congruously with the proportion of permanently rare ASVs in the rare biosphere, and the impact of variable selection and the proportion of conditionally rare ASVs in the rare biosphere increased in a similar fashion. The increased relative influence of dispersal limitation also aligned with the increased number of transiently rare ASVs. Taken together, these results further strengthen our hypothesis that a distinct interplay of ecological processes influences each type of rarity in a likely predictable manner.

Turnover of the common biosphere is governed mainly by variable selection

To contrast the processes structuring the rare biosphere, we also investigated the assembly of its counterpart, the common biosphere. Overall, the relative influences of community assembly processes changed in structuring the common biosphere as succession proceeded—an opposing pattern to that observed for the rare biosphere. Variable selection was the most dominant process across different successional stages, and sampling time, namely selection operates through environmental heterogeneity and/or biotic interactions [24]. We found the common biosphere mainly consisted of conditionally common species. These findings indicate that species in the common biosphere are likely more sensitive to biotic and abiotic fluctuations and can retract into the rare biosphere under harsh environmental conditions. For example, members of the phylum Cyanobacteria showed higher abundance in the common biosphere at early successional stages, i.e., 0 and 10 years of succession (Supplementary Figure 10). At these sites, the vegetation coverage is patchy, which provides appropriate conditions for oxygenic photosynthetic organisms [30]. While members of phylum Cyanobacteria mostly stay in low abundance at late successional stages as the environments do not favor them anymore (Supplementary Figure 10). This agrees with the finding that the abundance of dominant bacterial taxa fluctuates in response to disturbance treatments [49]. Similar results were also found in other ecosystems. For instance, depending on the bioavailability of organic pollutants, members of the conditionally rare taxa can become abundant and are responsible for degrading these compounds in the freshwater ecosystem [50]. In our study system, a dominant role of variable selection in structuring the common biosphere was observed at all the tested rarity cutoffs, i.e., from 0.2% to 0.05%, with the relative importance of selection increasing as rarity cutoff value was decreased. Specifically, the increased variable and homogeneous selection align with the increased number of conditionally and permanently common ASVs, respectively (Supplementary Figure 16). This further proved that different types of commonness are influenced by different ecological processes.

Undominated processes and homogenizing dispersal and also influenced the assembly of the common but not the rare biosphere. Previous data in communities of macro-organisms suggest that rare species are more subject to ecological drift than abundant species. But in our dataset, rare taxa with a relative abundance of

0.01% or one read per community could reflect as approximately 10^5-10^6 individuals in 2 grams of soil. With this considerable abundance, we would say that the theories applied for macro-organisms might not apply to microorganisms. The stronger influence of undominant processes in the common biosphere at the early successional stages was supported by the previous finding in the same system, which mostly focused on the members of the common biosphere due to less sequencing depth by pyrosequencing [51]. These results suggest microbial community composition at early successional stages is more prone to be influenced by ecological drift than late successional stages. Notice that since there are fewer species in the common biosphere, the null model results in Stegen's approach are more biased [52]. On the other hand, we found that although homogenizing process plays a minor role, it responded to the variation of community across successional stages. By analyzing microbial biogeographical patterns, Meyer *et al.* showed that abundant taxa are less restricted in distribution than low abundant ones [53]. A study of rare and dominant prokaryotic lineages in hydrothermal vent systems also reported that abundant lineages of archaea displayed a more cosmopolitan distribution, while rare lineages of archaea and almost all bacterial lineages are not widely dispersed [54]. For the temporal variation of communities, homogenizing dispersal found within each stage should be interpreted as species kept alive from previous times during temporal turnover. A stronger signal of homogenizing dispersal at late successional stages is actually community compositions in different sampling time points are more similar than random processes driven by ecological drift. This is in turn supported by the finding that undominant processes were less important at late successional stages. Together, these findings support the inference that abundant taxa tend to be influenced by homogenizing dispersal and display widespread distributional patterns.

Balancing rarity and commonness

We show that the rare biosphere corresponds to about 90% of the total bacterial diversity in these soils. Given that higher diversity helps to stabilize ecosystem processes in response to perturbation, decreases community susceptibility to invasion by non-native taxa, and increases nutrient cycling by resource partitioning [55-59], it is likely that the rare biosphere accounts for a fundamental but underestimated role in the functioning and stability of soil ecosystems. We demonstrate in this study that the ecological processes structuring the rare biosphere differ from that of the common biosphere, leading to distinct species distribution and composition. Moreover, the relative importance of different ecological processes reveals the dynamics between these components of the soil bacterial communities, as exemplified in Fig. 4. Specifically, the common biosphere consists primarily of conditionally common species, which might retract into the rare biosphere under harsh environmental conditions. Part of the rare biosphere can serve as a reservoir for these conditionally common species. Thus, they can stay at low abundances to avoid extreme conditions and respond once favorable environmental factors emerge, thus building up their population sizes (Fig. 4). When they become abundant, these conditionally rare taxa are expected to be related to the changes in the ecosystem functioning. They can also replace the previous abundant species, perform their corresponding functions and keep the resilience of ecosystems. However, the remaining part (about 2/3 of the rare biosphere) is driven mainly by homogeneous

selection, restraining these species to constant lower abundances. As such, these permanently rare species are more prone to perish in response to environmental disturbances, leading to important effects on the resistance of ecosystems.

Potential caveats and limitations

Here, we applied a quantitative approach to investigate how these processes govern the spatiotemporal turnover of the actively rare and common biospheres of bacterial communities across a natural soil chronosequence. However, this approach has intrinsic limitations for quantifying the influence of each assembly process [60-62]. First, it represents the overall influence of ecological processes at the meta-community level rather than the relative influence of these processes on microbial groups within each community [62]. Therefore, we can only quantify the relative importance of dominant processes among communities rather than within a community or a taxonomic group. Second, sequencing clusters also influence the inference of community assembly. The ASVs based dataset generates less phylogenetic and taxonomic turnover while higher stochastic signals than operational taxonomic units (OTUs) based dataset [62]. We delineated ASVs from raw sequences based on the DADA2 approach rather than sequencing clustering based on 97% similarity. By controlling errors, ASVs are more sensitive to single-nucleotide variations than OTUs [63]. Therefore, we stick to the results from ASVs, which reflect actual microbes better than OTUs. Third, dispersal processes might also be overestimated when the phylogenetic turnover cannot reflect the selection. Hence, dispersal limitation can play a less critical role in the successional chronosequence (< 10 km) than large spatial scales. In addition, dispersal limitation or homogenizing dispersal cannot be the process responsible for the temporal turnover of communities at the same location when the model shows taxonomic differences. Last, ecological drift can also be inflated when the number of species is small in a community or sub-community [52]. Despite these limitations, our approach provides the first attempt to distinguish the different ecological processes structuring various types of rarity. Further experiments and data are needed to verify the findings from this study.

Considering methodology, we have opted for an RNA-based approach to profile the bacterial rare biosphere rather than the usual DNA-based. The latter has the risk of misestimating rare species due to the pervasive existence of relic DNA in soil [64]. Although RNA-based approaches also have limitations, such as a variable number of ribosomal transcripts at different activities, we argue that this approach is more relevant for rare biosphere studies, as it controls for pseudo-rare species that can be potentially detected by DNA-based methods. Even though the RNA approach might wrongly classify rare species with high copy numbers of rRNA as abundant, it provides certainty that all active rare species are real and alive. Since the overall assembly processes did not differ significantly between DNA- and RNA-based approaches [29], we expected to draw similar conclusions for the rare biosphere when using the DNA-based approach.

Methods

Soil sampling and sequencing

The 16S rRNA sequence data analyzed in this study have been deposited at the Sequence Read Archive of the National Center for Biotechnology Information and are accessible through the accession numbers PRJNA546612 [29]. Details of soil sampling and sequencing are provided in SI Methods and ref. [29]. Briefly, soil samples were collected along a well-characterized soil chronosequence located on the island of Schiermonnikoog, the Netherlands (53°30' N, 6°10' E). This chronosequence covers over 100-years of primary succession in a developing salt marsh ecosystem since the sedimentation of particles carried out by the tide/wind causes the island to progressively extend eastwards [65]. Soil physicochemical properties and community composition (both macro- and micro-organisms) sequentially change over time along with the gradient [30, 65-67]. For example, this chronosequence presents a transition from sandy to clay soil. In addition, the overall soil nutrient status (i.e., organic carbon content, total nitrogen, ammonium, nitrate, and sulfate) increases over time, whereas soil pH decrease from ca. 8.7 to 7.4, as the succession advanced [30]. We collected soil samples across five successional stages (i.e., 0, 10, 40, 70, and 110 years of development from 1809 to 2017) to capture the variation in the rare biosphere across these sites (Figure 1A). Three replicates of composites soil from each stage (details see ref. [29]). Sampling took place in May, July, September, and November 2017 to include the temporal dynamics within each site. In total, we analyzed the assembly of the bacterial rare biosphere from 60 composited soil samples. To capture the putatively 'active' bacteria from soil, the V4 region of bacterial 16S rRNA was amplified based on reverse-transcribed total soil RNA using the primer set 515F and 806R [68, 69]. Sequencing was performed on a paired-end Illumina MiSeq (2 × 151bp) run (Illumina, USA) at the Argonne National Laboratory using the Version 2 chemistry sequencing reagent kit [68].

Sequence Processing

Sequence data analysis was performed using the open-source QIIME2/2018.2 pipeline [70]. Samples were demultiplexed, resulting in a total of 9,852,975 raw reads across the 60 samples. We applied the Divisive Amplicon Denoising Algorithm (DADA2) to infer exact/amplicon sequence variants (ASVs) [63]. The DADA2 method was applied to denoise paired-end sequences trimmed at 150bp, paired, and subjected to chimera removal using the default settings. The obtained feature table (site-to-species matrix) contained a total of 24,172 ASVs. The total frequency of these ASVs across all samples was 7,958,654. Taxonomic information for the representative sequences per ASV was obtained using the SILVA database (Silva 119 Naive Bayes 515F/806R taxonomy classifier) [71]. A phylogenetic tree was constructed by aligning representative sequences per ASV using the FastTree plugin in QIIME2 [72].

Defining the rare and common biospheres

After compiling the feature and taxonomy tables, community analyses and statistics were performed in the R environment (R version 3.5.0) [73, 74]. Figures were generated using the ggplot2 and venn.diagram

packages [75, 76]. The feature and taxonomy tables were combined, and all ASVs affiliated to archaea, chloroplasts, and mitochondria were removed. The feature table was rarefied to a depth of 31,500 sequences per sample using the 'rarefy' function in the package vegan [77]. The rarefied feature table was used for downstream analyses. We define rarity as per sample rather than through the entire dataset, which can obstruct observing the dynamics of the rare species. This allows us to depict sample-by-sample variations in rarity properly.

Define distinct types of rarity and commonness

After defining the rare and common biospheres, we further classified rare and common ASVs into distinct types of rarity and commonness (Supplementary Figure 2), as follows: 'conditionally rare/common'—rare ASVs in one or few samples that are occasionally common in other samples; 'transiently rare'—ASVs that appear only once in the rare biosphere across all samples (stage of succession and sampling time); 'permanently rare'—ASVs that are only present in the rare biosphere and appear more than once across all the samples; 'permanently common'—ASVs consistently present above the rare threshold across all samples. It should be mentioned that properly defining permanently rare ASVs is challenging. In this study, we sampled bacterial communities across a spatial chronosequence (from early 'sandy' to mature 'clay' soils), considering within-stage temporal variations (from May, July, September, and November). Such efforts allowed a thorough depiction of soil bacterial communities in this system, thus supporting a valid representation of the distinct types of rarity, particularly with respect to permanently rare ASVs.

Community analysis

Principal coordinate analysis (PCoA, 'pcoa' function in the package ape) was used to explore and visualize community dissimilarities using Bray-Curtis distances ('vegdist' function in the vegan package) [77, 78]. To assess whether biosphere component, stage of succession, and sampling time had significant effects on the bacterial community structure, permutational multivariate analysis of variance (PERMANOVA, 'adonis' function in the vegan package) was performed based on Bray-Curtis distances using 9,999 random permutations.

Quantifying the relative influences of distinct community assembly processes

To quantify community assembly processes structuring the common and rare biosphere, we used a previously developed approach [20]. Namely, phylogenetic community turnover was inferred based on the β-nearest taxon index (βNTI) across all samples. We calculated βNTI using the package picante and the script created by Stegen *et al.* [20, 79], which is the standardization of between-community mean nearest taxon distance (βMNTD), as follows:

$$\beta\text{MNTD} = 0.5\left[\sum_{i_k=1}^{n_k} f_{ik}\,\min(\Delta i_k j_m) + \sum_{i_m=1}^{n_m} f_{im}\,\min(\Delta i_m j_k)\right],$$

$$\beta NTI = \frac{\beta\text{MNTD}_{obs} - \overline{\beta\text{MNTD}_{null}}}{sd(\beta\text{MNTD}_{null})},$$

where f_{ik} is the relative abundance of ASV i in community k, n_k is the number of ASVs in community k, f_{im} is the relative abundance of ASV i in community m, n_m is the number of ASVs in community m, and $\min(\Delta i_k j_m)$ is the phylogenetic distance among closest ASVs occurring in community k and community m. The distribution of βMNTD from null models was built by shuffling ASVs among the tips of the phylogenetic tree with 999 permutations. βNTI values above +2 indicate a higher phylogenetic turnover in observed communities than in the null model distribution. This indicates a strong influence of variable selection on community turnover. βNTI values below -2 indicate lower phylogenetic community turnover in observed communities than in the null model distribution, which further indicates the influence of homogeneous selection on community assembly [20, 21, 24]. If $-2 < \beta NTI < +2$, community turnover does not significantly deviate from null expectation, and is thus governed mostly by stochastic processes, such as dispersal limitation, homogenizing dispersal, or undominated processes. In a follow-up analytical step, the Raup-Crick matrix based on the standardized Bray-Curtis matrix (referred to as RC_{bray}) was used to test whether the observed degree of turnover deviates from the expectation [25]. We applied the script created by Stegen *et al.* to compute the RC_{bray} matrix for all communities [20]. A null distribution of the Raup-Crick matrix was built by simulating 999 times for each pair of communities. The RC_{bray} was calculated by the deviation between observed values and the null distribution and rescaled to a range from -1 to +1. $RC_{bray} > +0.95$ indicates dispersal limitation coupled with drift that leads to community turnover greater than expected, whereas $RC_{bray} < -0.95$ indicates that community turnover is primarily governed by homogenizing dispersal. Last, $-0.95 < RC_{bray} < +0.95$ is interpreted as indicating undominated processes. We quantified the relative influences of each of these processes by calculating the percentage βNTI and RC_{bray} values that fulfill the above criteria across all the pairwise comparisons per each community compartment (Supplementary Figure 1, Step 3).

Supplementary information

Please see an additional word file for the supplementary information accompanied this paper.

Acknowledgments

We thank Leonel Herrera Alsina for his help in writing the R script used to define the distinct types of rarity, and Esther Chang for helpful comments on this manuscript. We acknowledge the Center for Information Technology at the University of Groningen for their support and access to the Peregrine high-performance computing cluster.

Authors' Contributions

559 XJ, FDA and JFS designed the study. XJ performed data analysis, and wrote the draft manuscript. All authors
560 contributed substantially to manuscript revisions and approved the final manuscript.
561
562 **Funding**
563 This research was funded by a personal grant to JFS. XJ was supported by a scholarship from the China
564 Scholarship Council and the University of Groningen scholarship program.
565
566 **Availability of data and materials**
567 The 16S rRNA amplicon data analyzed during the current study are available in the Sequence Read Archive
568 of the National Center for Biotechnology Information with the accession numbers PRJNA546612
569 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA546612).
570 All R codes used in this study are available on GitHub (https://github.com/Jia-
571 Xiu/rare_biosphere_assembly_2020).
572
573 **Ethics approval and consent to participate**
574 Not applicable
575
576 **Consent for publication**
577 Not applicable
578
579 **Competing interests**
580 The authors declare that they have no competing interests.
581
582 **References**
583
584 1. Pedros-Alio C. The rare bacterial biosphere. Ann Rev Mar Sci. 2012;4:449-66; doi:
585 10.1146/annurev-marine-120710-100948.
586 2. Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, et al. Microbial diversity in
587 the deep sea and the underexplored "rare biosphere". Proceedings of the National Academy of
588 Sciences. 2006;103(32):12115-20; doi: 10.1073/pnas.0605127103.
589 3. Hausmann B, Pelikan C, Rattei T, Loy A, Pester M. Long-Term Transcriptional Activity at Zero
590 Growth of a Cosmopolitan Rare Biosphere Member. mBio. 2019;10(1):e02189-18; doi:
591 10.1128/mBio.02189-18.
592 4. Pester M, Bittner N, Deevong P, Wagner M, Loy A. A 'rare biosphere'microorganism contributes
593 to sulfate reduction in a peatland. ISME J. 2010;4(12):1591-602.
594 5. Rivett DW, Bell T. Abundance determines the functional role of bacterial phylotypes in complex
595 communities. Nature Microbiology. 2018;3(7):767-72; doi: 10.1038/s41564-018-0180-0.
596 6. van Elsas JD, Chiurazzi M, Mallon CA, Elhottova D, Kristufek V, Salles JF. Microbial diversity
597 determines the invasion of soil by a bacterial pathogen. Proceedings of the National Academy of
598 Sciences of the United States of America. 2012;109(4):1159-64; doi: 10.1073/pnas.1109326109.
599 7. Magurran AE, Henderson PA. Explaining the excess of rare species in natural species abundance
600 distributions. Nature. 2003;422(6933):714-6.

8. Rabinowitz D, Rapp JK, Dixon PM. Competitive Abilities of Sparse Grass Species: Means of Persistence or Cause of Abundance. Ecology. 1984;65(4):1144-54; doi: 10.2307/1938322.
9. Reinhardt K, Köhler G, Maas S, Detzel P. Low dispersal ability and habitat specificity promote extinctions in rare but not in widespread species: the Orthoptera of Germany. Ecography. 2005;28(5):593-602; doi: 10.1111/j.2005.0906-7590.04285.x.
10. Yenni G, Adler PB, Ernest S. Strong self‐limitation promotes the persistence of rare species. Ecology. 2012;93(3):456-61.
11. Jousset A, Bienhold C, Chatzinotas A, Gallien L, Gobet A, Kurm V, et al. Where less may be more: how the rare biosphere pulls ecosystems strings. ISME J. 2017;11:853-62; doi: 10.1038/ismej.2016.174.
12. Thingstad TF. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnology and Oceanography. 2000;45(6):1320-8; doi: https://doi.org/10.4319/lo.2000.45.6.1320.
13. Szekely AJ, Langenheder S. The importance of species sorting differs between habitat generalists and specialists in bacterial communities. FEMS Microbiol Ecol. 2014;87(1):102-12; doi: 10.1111/1574-6941.12195.
14. Kurm V, van der Putten WH, de Boer W, Naus-Wiezer S, Hol WHG. Low abundant soil bacteria can be metabolically versatile and fast growing. Ecology. 2017;98(2):555-64; doi: 10.1002/ecy.1670.
15. Mo Y, Zhang W, Yang J, Lin Y, Yu Z, Lin S. Biogeographic patterns of abundant and rare bacterioplankton in three subtropical bays resulting from selective and neutral processes. ISME J. 2018;12(9):2198-210; doi: 10.1038/s41396-018-0153-6.
16. Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, et al. Patterns and processes of microbial community assembly. Microbiology and molecular biology reviews : MMBR. 2013;77(3):342-56; doi: 10.1128/MMBR.00051-12.
17. Vellend M. Conceptual Synthesis in Community Ecology. The Quarterly Review of Biology. 2010;85(2):183-206; doi: 10.1086/652373.
18. Vellend M. The Theory of Ecological Communities. Princeton University Pres. 2016:61-7.
19. Jia X, Dini-Andreote F, Falcao Salles J. Community Assembly Processes of the Microbial Rare Biosphere. Trends Microbiol. 2018;26(9):738-47; doi: 10.1016/j.tim.2018.02.011.
20. Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, et al. Quantifying community assembly processes and identifying features that impose them. ISME J. 2013;7(11):2069-79; doi: 10.1038/ismej.2013.93.
21. Stegen JC, Lin X, Fredrickson JK, Konopka AE. Estimating and mapping ecological processes influencing microbial community assembly. Front Microbiol. 2015;6:doi: 10.3389/fmicb.2015.00370; doi: 10.3389/fmicb.2015.00370.
22. Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. Phylogenies and Community Ecology. Annual Review of Ecology and Systematics. 2002;33(1):475-505; doi: 10.1146/annurev.ecolsys.33.010802.150448.
23. Lynch MDJ, Neufeld JD. Ecology and exploration of the rare biosphere. Nat Rev Micro. 2015;13(4):217-29; doi: 10.1038/nrmicro3400.
24. Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(11):E1326-E32; doi: 10.1073/pnas.1414261112.
25. Chase JM, Kraft NJB, Smith KG, Vellend M, Inouye BD. Using null models to disentangle variation in community dissimilarity from variation in α-diversity. Ecosphere. 2011;2(2):art24; doi: 10.1890/es10-00117.1.
26. Shade A, Jones SE, Caporaso JG, Handelsman J, Knight R, Fierer N, et al. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. mBio. 2014;5(4):e01371-14; doi: 10.1128/mBio.01371-14.
27. Strous M, Heijnen JJ, Kuenen JG, Jetten MSM. The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms. Applied Microbiology and Biotechnology. 1998;50(5):589-96; doi: 10.1007/s002530051340.

28. Goldfarb KC, Karaoz U, Hanson CA, Santee CA, Bradford MA, Treseder KK, et al. Differential growth responses of soil bacterial taxa to carbon substrates of varying chemical recalcitrance. Front Microbiol. 2011;2:94; doi: 10.3389/fmicb.2011.00094.
29. Jia X, Dini-Andreote F, Falcão Salles J. Comparing the Influence of Assembly Processes Governing Bacterial Community Succession Based on DNA and RNA Data. Microorganisms. 2020;8(6); doi: 10.3390/microorganisms8060798.
30. Dini-Andreote F, Silva M, Triado-Margarit X, Casamayor EO, van Elsas JD, Salles JF. Dynamics of bacterial community succession in a salt marsh chronosequence: evidences for temporal niche partitioning. ISME J. 2014;8(10):1989-2001; doi: 10.1038/ismej.2014.54.
31. Yamamoto K, Hackley KC, Kelly WR, Panno SV, Sekiguchi Y, Sanford RA, et al. Diversity and geochemical community assembly processes of the living rare biosphere in a sand-and-gravel aquifer ecosystem in the Midwestern United States. Scientific Reports. 2019;9(1); doi: 10.1038/s41598-019-49996-z.
32. Galand PE, Casamayor EO, Kirchman DL, Lovejoy C. Ecology of the rare microbial biosphere of the Arctic Ocean. Proceedings of the National Academy of Sciences. 2009;106(52):22427-32; doi: 10.1073/pnas.0908284106.
33. Reveillaud J, Maignien L, Murat Eren A, Huber JA, Apprill A, Sogin ML, et al. Host-specificity among abundant and rare taxa in the sponge microbiome. ISME J. 2014;8(6):1198-209; doi: 10.1038/ismej.2013.227.
34. Logares R, Audic S, Bass D, Bittner L, Boutte C, Christen R, et al. Patterns of rare and abundant marine microbial eukaryotes. Current biology : CB. 2014;24(8):813-21; doi: 10.1016/j.cub.2014.02.050.
35. Campbell BJ, Yu L, Heidelberg JF, Kirchman DL. Activity of abundant and rare bacteria in a coastal ocean. Proceedings of the National Academy of Sciences. 2011;108(31):12776-81; doi: 10.1073/pnas.1101405108.
36. Hirsch JE. An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(46):16569; doi: 10.1073/pnas.0507655102.
37. Haegeman B, Hamelin J, Moriarty J, Neal P, Dushoff J, Weitz JS. Robust estimation of microbial diversity in theory and in practice. ISME J. 2013;7(6):1092-101; doi: 10.1038/ismej.2013.10.
38. Jiao S, Lu Y. Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. Environ Microbiol. 2020;22(3):1052-65; doi: 10.1111/1462-2920.14815.
39. Logares R, Lindström ES, Langenheder S, Logue JB, Paterson H, Laybourn-Parry J, et al. Biogeography of bacterial communities exposed to progressive long-term environmental change. ISME J. 2012;7(5):937-48; doi: 10.1038/ismej.2012.168.
40. Kurm V, van der Putten WH, Weidner S, Geisen S, Snoek BL, Bakx T, et al. Competition and predation as possible causes of bacterial rarity. Environ Microbiol. 2019;21(4):1356-68; doi: 10.1111/1462-2920.14569.
41. Aanderud ZT, Saurey S, Ball BA, Wall DH, Barrett JE, Muscarella ME, et al. Stoichiometric Shifts in Soil C:N:P Promote Bacterial Taxa Dominance, Maintain Biodiversity, and Deconstruct Community Assemblages. Front Microbiol. 2018;9:1401; doi: 10.3389/fmicb.2018.01401.
42. Sloan WT, Woodcock S, Lunn M, Head IM, Curtis TP. Modeling taxa-abundance distributions in microbial communities using environmental sequence data. Microb Ecol. 2007;53(3):443-55; doi: 10.1007/s00248-006-9141-x.
43. Magurran AE, McGill BJ. Biological diversity: frontiers in measurement and assessment. Oxford University Press; 2011.
44. Richter-Heitmann T, Hofner B, Krah FS, Sikorski J, Wust PK, Bunk B, et al. Stochastic Dispersal Rather Than Deterministic Selection Explains the Spatio-Temporal Distribution of Soil Bacteria in a Temperate Grassland. Front Microbiol. 2020;11:1391; doi: 10.3389/fmicb.2020.01391.
45. Ivanov, II, Honda K. Intestinal commensal microbes as immune modulators. Cell Host Microbe. 2012;12(4):496-508; doi: 10.1016/j.chom.2012.09.009.
46. van Veelen HPJ, Falcao Salles J, Tieleman BI. Multi-level comparisons of cloacal, skin, feather and nest-associated microbiota suggest considerable influence of horizontal acquisition on the

709 microbiota assembly of sympatric woodlarks and skylarks. Microbiome. 2017;5(1):156; doi:
710 10.1186/s40168-017-0371-6.
711 47. Warmink JA, Nazir R, Corten B, van Elsas JD. Hitchhikers on the fungal highway: The helper
712 effect for bacterial migration via fungal hyphae. Soil Biology and Biochemistry. 2011;43(4):760-
713 5; doi: 10.1016/j.soilbio.2010.12.009.
714 48. Snell Taylor SJ, Evans BS, White EP, Hurlbert AH. The prevalence and impact of transient
715 species in ecological communities. Ecology. 2018;99(8):1825-35; doi: 10.1002/ecy.2398.
716 49. Kurm V, Geisen S, Gera Hol WH. A low proportion of rare bacterial taxa responds to abiotic
717 changes compared with dominant taxa. Environ Microbiol. 2019;21(2):750-8; doi: 10.1111/1462-
718 2920.14492.
719 50. Wang Y, Hatt JK, Tsementzi D, Rodriguez RL, Ruiz-Perez CA, Weigand MR, et al. Quantifying
720 the Importance of the Rare Biosphere for Microbial Community Response to Organic Pollutants in
721 a Freshwater Ecosystem. Appl Environ Microbiol. 2017;83(8):e03321-16; doi:
722 10.1128/AEM.03321-16.
723 51. Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF. Disentangling mechanisms that mediate the
724 balance between stochastic and deterministic processes in microbial succession. Proc Natl Acad
725 Sci U S A. 2015;112(11):E1326-32; doi: 10.1073/pnas.1414261112.
726 52. Cao J, Jia X, Pang S, Hu Y, Li Y, Wang Q. Functional structure, taxonomic composition and the
727 dominant assembly processes of soil prokaryotic community along an altitudinal gradient. Applied
728 Soil Ecology. 2020;155; doi: 10.1016/j.apsoil.2020.103647.
729 53. Meyer KM, Memiaghe H, Korte L, Kenfack D, Alonso A, Bohannan BJM. Why do microbes
730 exhibit weak biogeographic patterns? ISME J. 2018;12(6):1404-13; doi: 10.1038/s41396-018-
731 0103-3.
732 54. Anderson RE, Sogin ML, Baross JA. Biogeography and ecology of the rare and abundant
733 microbial lineages in deep-sea hydrothermal vents. FEMS Microbiol Ecol. 2015;91(1):1-11; doi:
734 10.1093/femsec/fiu016.
735 55. Mallon CA, Le Roux X, van Doorn GS, Dini-Andreote F, Poly F, Salles JF. The impact of failure:
736 unsuccessful bacterial invasions steer the soil microbial community away from the invader's niche.
737 ISME J. 2018;12(3):728-41; doi: 10.1038/s41396-017-0003-y.
738 56. Langenheder S, Bulling MT, Solan M, Prosser JI. Bacterial biodiversity-ecosystem functioning
739 relations are modified by environmental complexity. PLoS One. 2010;5(5):e10834; doi:
740 10.1371/journal.pone.0010834.
741 57. Bardgett RD, Van Der Putten WH. Belowground biodiversity and ecosystem functioning. Nature.
742 2014;515(7528):505.
743 58. Griffiths B, Ritz K, Wheatley R, Kuan H, Boag B, Christensen S, et al. An examination of the
744 biodiversity–ecosystem function relationship in arable soil microbial communities. Soil Biology
745 and Biochemistry. 2001;33(12-13):1713-22.
746 59. Hooper DU, Chapin F, Ewel J, Hector A, Inchausti P, Lavorel S, et al. Effects of biodiversity on
747 ecosystem functioning: a consensus of current knowledge. Ecological monographs. 2005;75(1):3-
748 35.
749 60. Logares R, Tesson SVM, Canback B, Pontarp M, Hedlund K, Rengefors K. Contrasting
750 prevalence of selection and drift in the community structuring of bacteria and microbial
751 eukaryotes. Environ Microbiol. 2018;20(6):2231-40; doi: 10.1111/1462-2920.14265.
752 61. Zhou J, Ning D. Stochastic Community Assembly: Does It Matter in Microbial Ecology?
753 Microbiology and Molecular Biology Reviews. 2017;81(4):e00002-17.
754 62. Logares R, Deutschmann IM, Junger PC, Giner CR, Krabberod AK, Schmidt TSB, et al.
755 Disentangling the mechanisms shaping the surface ocean microbiota. Microbiome. 2020;8(1):55;
756 doi: 10.1186/s40168-020-00827-8.
757 63. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational
758 taxonomic units in marker-gene data analysis. ISME J. 2017;11(12):2639-43; doi:
759 10.1038/ismej.2017.119.
760 64. Carini P, Marsden PJ, Leff JW, Morgan EE, Strickland MS, Fierer N. Relic DNA is abundant in
761 soil and obscures estimates of soil microbial diversity. Nature Microbiology. 2016;2:16242; doi:
762 10.1038/nmicrobiol.2016.242

https://www.nature.com/articles/nmicrobiol2016242#supplementary-information.
65. Olff H, De Leeuw J, Bakker JP, Platerink RJ, van Wijnen HJ. Vegetation Succession and Herbivory in a Salt Marsh: Changes Induced by Sea Level Rise and Silt Deposition Along an Elevational Gradient. Journal of Ecology. 1997;85(6):799-814; doi: 10.2307/2960603.
66. Dini-Andreote F, Pylro VS, Baldrian P, van Elsas JD, Salles JF. Ecological succession reveals potential signatures of marine–terrestrial transition in salt marsh fungal communities. ISME J. 2016;10:1984–97.
67. Schrama M, Berg Mp Fau - Olff H, Olff H. Ecosystem assembly rules: the interplay of green and brown webs during salt marsh succession. Ecology. 2012;93(11)(0012-9658 (Print)):2353–64.
68. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences. 2011;108(Supplement 1):4516-22; doi: 10.1073/pnas.1000080107.
69. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012;6(8):1621-4.
70. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, Al-Ghalith GA, et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. PeerJ Preprints. 2018;6:e27295v2; doi: 10.7287/peerj.preprints.27295v2.
71. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Res. 2014;42(Database issue):D643-8; doi: 10.1093/nar/gkt1209.
72. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Molecular biology and evolution. 2009;26(7):1641-50.
73. R Core Team: R: A language and environment for statistical computing. In. Vienna, Austria: R Foundation for Statistical Computing; 2017.
74. RStudio Team: RStudio: integrated development for R. In., vol. 42. Boston, MA: RStudio, Inc.; 2015.
75. Wickham H. ggplot2: elegant graphics for data analysis. J Stat Softw. 2010;35(1):65-88.
76. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC bioinformatics. 2011;12(1):35.
77. Dixon P. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science. 2003;14(6):927-30; doi: 10.1111/j.1654-1103.2003.tb02228.x.
78. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20(2):289-90.
79. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, et al. Picante: R tools for integrating phylogenies and ecology. Bioinformatics. 2010;26(11):1463-4.
80. Dini-Andreote F, Brossi MJ, van Elsas JD, Salles JF. Reconstructing the Genetic Potential of the Microbially-Mediated Nitrogen Cycle in a Salt Marsh Ecosystem. Front Microbiol. 2016;7(902):902; doi: 10.3389/fmicb.2016.00902.

804 **Figures legends**

805

806 **Fig. 1.** The distribution of bacterial species in the common and rare biospheres. **a** Sampling sites are shown
807 on the map of the island of Schiermonnikoog, the Netherlands. Distinct colours represent different
808 successional stages along the soil chronosequence that gradually develops from the east to the west. Dots
809 represent the sampling sites at five successional stages, namely 0, 10, 40, 70, and 110 years of succession
810 (from 1809 to 2017). Modified version from Dini-Andreote et al. [80]. This image was originally generated
811 using ArcGIS. **b** The number of rare and common ASVs and the number of shared ASVs in the rare and
812 common biospheres. **c** Principal Coordinate Analysis (PCoA) derived from Bray-Curtis distance of bacterial
813 composition for the rare and common biospheres represented by different shapes. Colors represent
814 successional stages (0, 10, 40, 70, and 110 years in succession). Percentage in the axis labels shows the
815 variation of species composition explained by each ordinate.

816

817 **Fig. 2.** The relative influence of each assembly process structuring the species turnover of the common (**a**)
818 and rare (**b**) biospheres. Pie plots show the relative influence of each assembly process shapes the species
819 turnover of the rare and common biospheres across all samples. Bar plots in the middle and right show the
820 interplay of assembly processes governing the temporal variation of communities within each successional
821 stage (i.e., 0, 10, 40, 70, and 110 years in succession), and the assembly processes driving the spatial variation
822 of communities across successional stages at four sampling times (i.e., May, July, September, and November).
823 The βNTI and RC_{bray} values were used to quantify the relative importance of each assembly process. Colors
824 represent different assembly processes, i.e., variable selection, homogeneous selection, dispersal limitation,
825 homogenizing dispersal, and uncommon processes. The asterisk denotes the impact of homogenizing
826 dispersal and undominated processes for the turnover of the whole community, which are 0.56% and 0.06%,
827 respectively.

828

829 **Fig. 3**. Types of rarity and commonness change over time and space. The relative abundance (**a**) and number
830 of ASVs (**b**) of each type of rarity and commonness in the rare and common biospheres across four sampling
831 times (M-May, J-July, S-September, N-November) and five successional stages (0, 10, 40, 70 and 110 years
832 in succession). Colors represent different types of rarity or commonness. The relative abundances and number
833 of ASVs correspond to the average values among three replicates.

834

835 **Fig. 4.** Conceptual figure displaying the relative influences of distinct ecological processes on the dynamics
836 of the common and rare microbial biospheres. This figure depicts a rank-abundance curve in which species
837 abundance (purple line) is ordered from high (common biosphere) to low (rare biosphere). In this study, we

demonstrated that the majority of the rare species constantly remain at low abundances and are thus being structured by homogeneous selection (red arrow). In contrast, dispersal limitation (brown arrow) plays a reduced role. Due to variable selection, a substantial fraction of the common species and the rare species shift between them across space and time (green arrows). In contrast with the rare biosphere, the abundance of common species is also driven by homogenizing dispersal (brown arrow) and drift (blue arrow), and to a lesser extent, by homogeneous selection (red arrow).

Figures

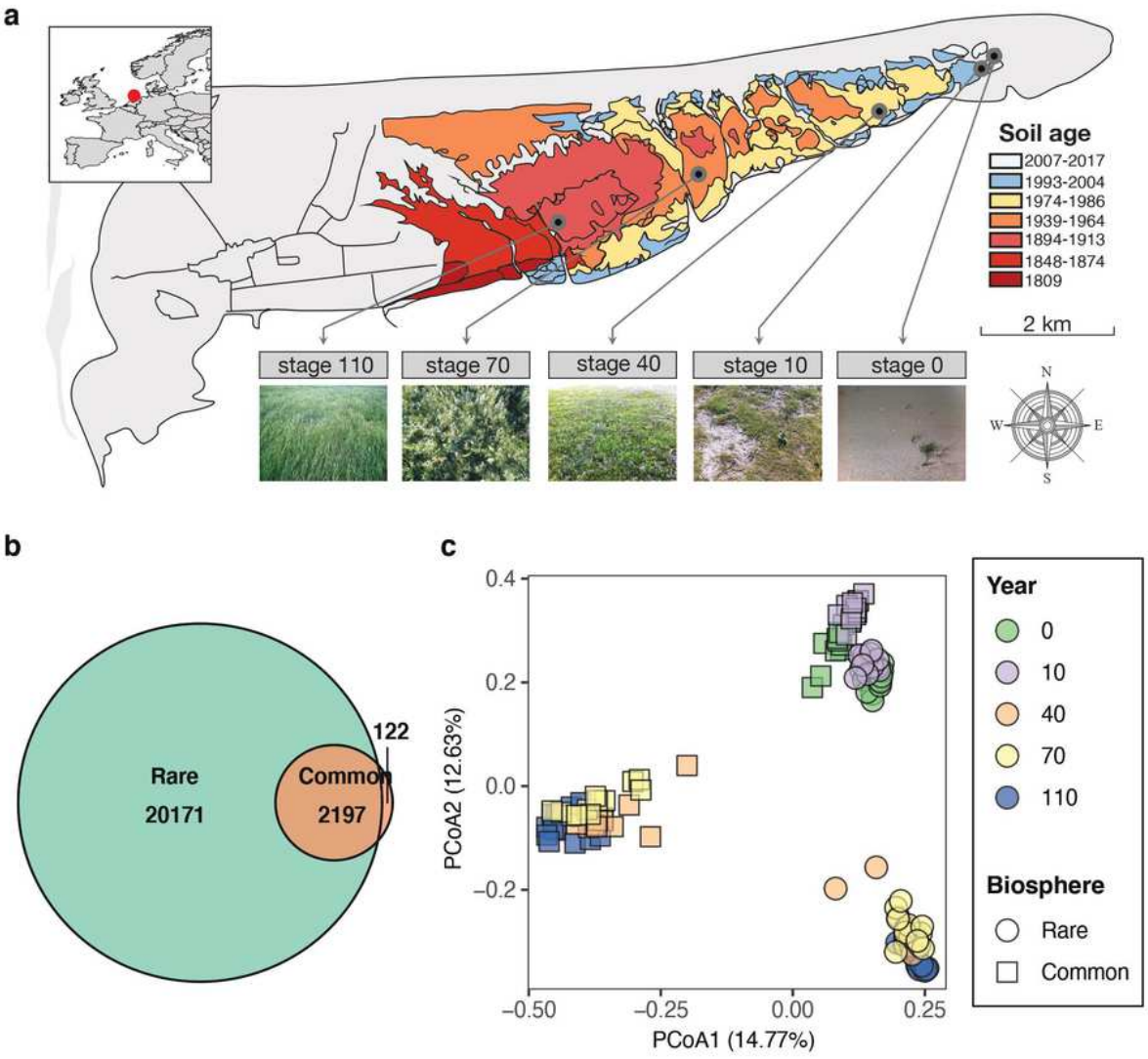

Fig. 1. The distribution of bacterial species in the common and rare biospheres. **a** Sampling sites are shown on the map of the island of Schiermonnikoog, the Netherlands. Distinct colours represent different successional stages along the soil chronosequence that gradually develops from the east to the west. Dots represent the sampling sites at five successional stages, namely 0, 10, 40, 70, and 110 years of succession (from 1809 to 2017). Modified version from Dini-Andreote et al. [80]. This image was originally generated using ArcGIS. **b** The number of rare and common ASVs and the number of shared ASVs in the rare and common biospheres. **c** Principal Coordinate Analysis (PCoA) derived from Bray-Curtis distance of bacterial composition for the rare and common biospheres represented by different shapes. Colors represent successional stages (0, 10, 40, 70, and 110 years in succession). Percentage in the axis labels shows the variation of species composition explained by each ordinate.

Figure 1

Please See image above for figure legend.

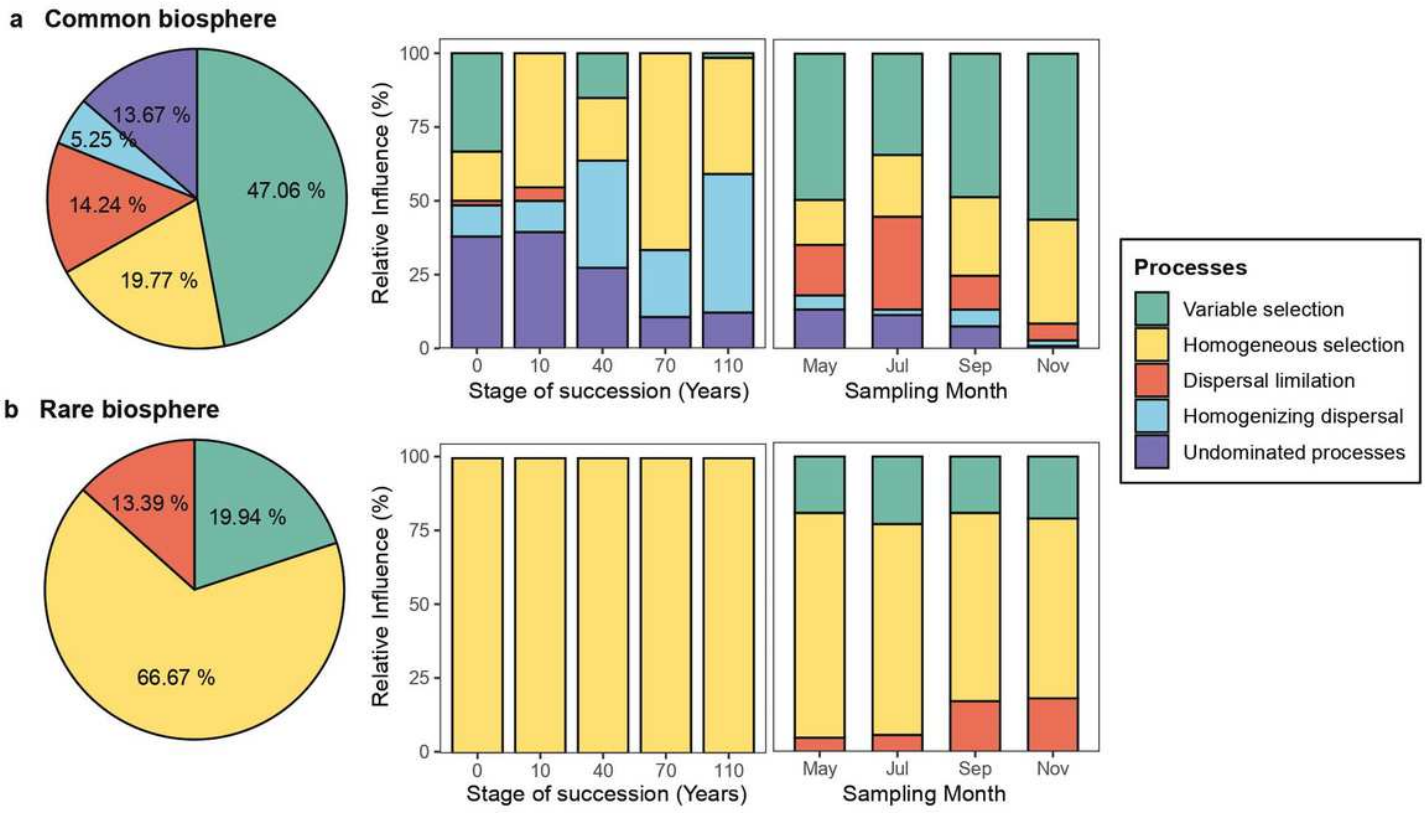

Fig. 2. The relative influence of each assembly process structuring the species turnover of the common (**a**) and rare (**b**) biospheres. Pie plots show the relative influence of each assembly process shapes the species turnover of the rare and common biospheres across all samples. Bar plots in the middle and right show the interplay of assembly processes governing the temporal variation of communities within each successional stage (i.e., 0, 10, 40, 70, and 110 years in succession), and the assembly processes driving the spatial variation of communities across successional stages at four sampling times (i.e., May, July, September, and November). The βNTI and RC_{bray} values were used to quantify the relative importance of each assembly process. Colors represent different assembly processes, i.e., variable selection, homogeneous selection, dispersal limitation, homogenizing dispersal, and uncommon processes. The asterisk denotes the impact of homogenizing dispersal and undominated processes for the turnover of the whole community, which are 0.56% and 0.06%, respectively.

Figure 2

Please See image above for figure legend.

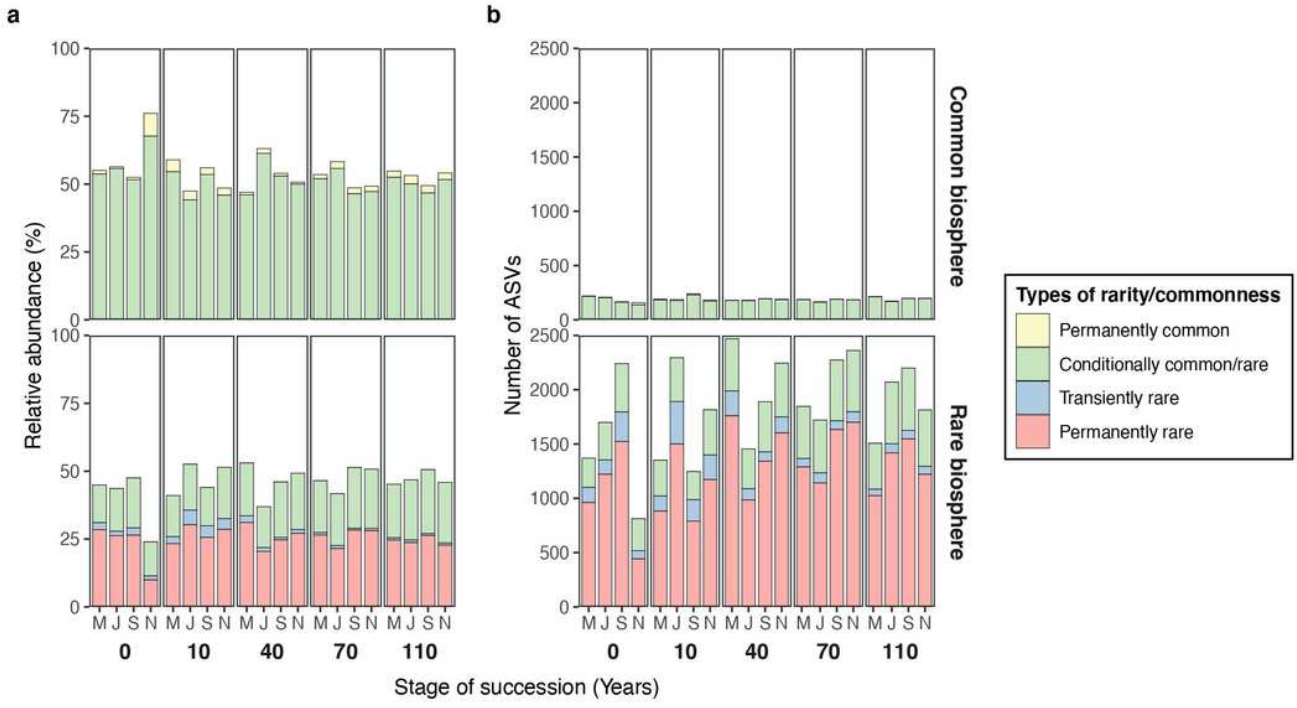

Fig. 3. Types of rarity and commonness change over time and space. The relative abundance (**a**) and number of ASVs (**b**) of each type of rarity and commonness in the rare and common biospheres across four sampling times (M-May, J-July, S-September, N-November) and five successional stages (0, 10, 40, 70 and 110 years in succession). Colors represent different types of rarity or commonness. The relative abundances and number of ASVs correspond to the average values among three replicates.

Figure 3

Please See image above for figure legend.

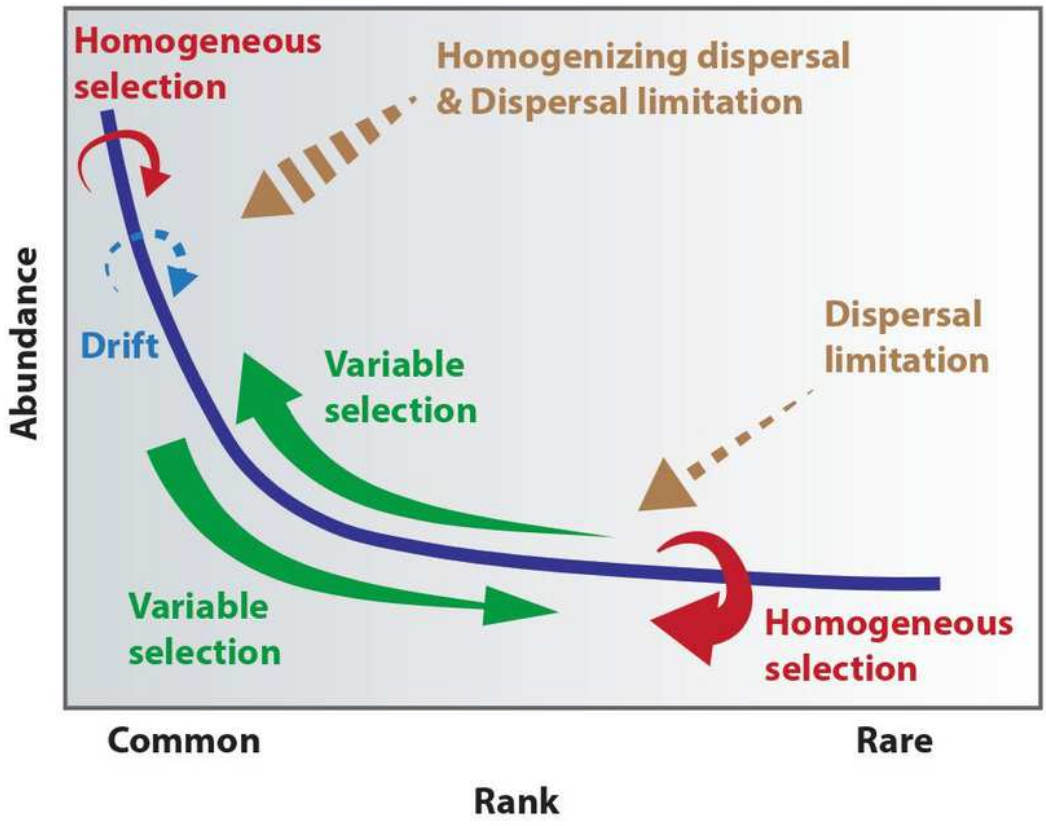

Fig. 4. Conceptual figure displaying the relative influences of distinct ecological processes on the dynamics of the common and rare microbial biospheres. This figure depicts a rank-abundance curve in which species abundance (purple line) is ordered from high (common biosphere) to low (rare biosphere). In this study, we demonstrated that the majority of the rare species constantly remain at low abundances and are thus being structured by homogeneous selection (red arrow). In contrast, dispersal limitation (brown arrow) plays a reduced role. Due to variable selection, a substantial fraction of the common species and the rare species shift between them across space and time (green arrows). In contrast with the rare biosphere, the abundance of common species is also driven by homogenizing dispersal (brown arrow) and drift (blue arrow), and to a lesser extent, by homogeneous selection (red arrow).

Figure 4

Please See image above for figure legend.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterialJiaetal14022022.pdf