Soil bacterial community assembly during succession

Jia, Xiu

DOI:
10.33612/diss.156586683

CHAPTER 1

General introduction

Improved understanding of microbial communities

The Earth is estimated to harbour around 10^{12} microbial species (Locey and Lennon 2016), $4–6\times10^{30}$ prokaryotic cells with microbial cell numbers estimated to be approximately 1.2×10^{29} in aquatic environments and $4–5\times10^{30}$ in terrestrial systems (Whitman et al. 1998b, Singh et al. 2009). These microbes constitute about 60% of the Earth's biomass and allocate 350–550 Pg of carbon in total (Whitman et al. 1998b, Singh et al. 2009). Phylogenetic analysis of prokaryotic genomes has pointed that the last universal common ancestor (LUCA) of archaea and bacteria is likely the pioneer organisms that enable exploring the hydrothermal environments on the young Earth (Weiss et al. 2016). Since the emergence of microbes around 3.8–3.5 billion years ago, the evolution of microbes has tightly intertwined with Earth's environmental evolution (Arndt and Nisbet 2012). For instance, the photosynthetic activity of Cyanobacteria was responsible for the substantial change of the atmosphere of our planet in the Proterozoic Eon, as they assimilate CO_2 and release O_2, leading to the great oxygenation event (Holland 2006, Lyons et al. 2014). This event is perceived to be responsible for the big extinction of microbes, but also paved the way for the evolution of early eukaryotes (Knoll et al. 2016). After witnessing the extinction of more than 99.9 % of all evolutionary lines that used to live on Earth (Mayr 1997), bacteria and archaea are still ubiquitous on Earth and live together with other organisms in this Anthropocene. These microbes with a wide range of ecological breadths occupy any living system on Earth, including soil, sediments, river, lake, ocean, glacier, hydrothermal vents, etc. They can be characterized as either free-living organisms, or host-associated organisms, such as plant-associated or animal-associated ones (Thompson et al. 2017).

Our knowledge about microbes tracks back to the 17th century when Antonie van Leeuwenhoek developed a microscope to observe bacteria together with other 'animalcules' (O'Malley 2007). Later on, identifying microbial species using the culturing approach in the 19th century, we began to acquire knowledge of microbes, such as their classification, cellular properties and physiology (e.g. thermal tolerances, the substrates they consume, or the enzymes they produce) (Blevins and Bronze 2010). But the majority of microorganisms from natural environments are unculturable under laboratory conditions, leading to the 'great plate count anomaly', namely the order of magnitude of environmental cells that formed in culture media is different from the number of them detected by microscopic examination (Staley and Konopka 1985). Until today, researchers are in the dilemma to study environmental microorganisms in the culturable state, even though a lot of endeavours have been made and the implications of new approaches currently being proposed (Stewart 2012, Browne et al. 2016).

After the discovery of nucleic acids as genetic materials and the advance of molecular biological techniques, microbiologists began to endeavour on identifying microorganisms by directly sequencing nucleic acids – DNA or RNA – instead of based on culturable cell characteristics. A revolution in microbiology occurred when Carl Woese applied the 16S ribosomal RNA molecule as a biomarker to analyse the phylogenetic relationship of microbes in 1977 (Woese and Fox 1977), which was developed as the three-domain system, bacteria, archaea and eukaryote in 1990 (Woese et al. 1990). At the same time,

genetic markers started to be used to examine the distribution of marine bacteria in the Sargasso Sea using nucleic acid hybridization probes (Giovannoni et al. 1990). Later on, the development of low-resolution DNA fingerprinting, such as DGGE (Denaturing gradient gel electrophoresis), TRFLP (terminal restriction fragment length polymorphism) and ARISA (automated ribosomal intergenic spacer analysis), has enabled the understanding of patterns of microbial communities (Nocker et al. 2007). In the mid-2000s, advances in high-throughput DNA sequencing of amplicons have improved the identification of different microbial communities by targeting the 16S rRNA or 18S rRNA genes, the internal transcribed spacer (ITS), or specific functional genes (Logares et al. 2012, Knight et al. 2018). Besides, high-throughput sequencing of untargeted shotgun metagenomics of environmental samples, together with other 'omics' approaches, such as metatranscriptome, metaproteome and metabolome, have been providing high-resolution information on the composition and functional characteristics of microbial communities (Handelsman 2004, Franzosa et al. 2015a, Hultman et al. 2015). Using these approaches, a few large projects were initiated aiming at unveiling the factors structuring microbiomes and their functions across different habitats, such as the Earth Microbiome Project (Thompson et al. 2017), the Tara Oceans Project (Sunagawa et al. 2020) and the Human Microbiome Project (Turnbaugh et al. 2007).

The rapid advance of sequencing platforms has revolutionized our understanding of microbial communities. This led to the emergence of the term 'microbiome', i.e. a collection of all microorganisms, such as bacterial, archaea, fungi, viruses, protists or other eukaryotes, that live in the same habitat and tightly interact with each other and their local environment. Other researchers prefer the term 'microbiota', while a combination of microbiota, their genomes and their environments refer to the microbiome (Chase et al. 2020). Due to difficulties associated with profiling all types of microbes in a given study, most studies focus on targeting a specific group of organisms within the microbiome. In these cases, it would be better to use the original terms, such as bacterial, archaeal or fungal communities rather than the term 'microbiome' [see (Berg et al. 2020)]. Throughout this thesis, I set a focus on studying the ecological processes and mechanisms mediating the assembly and successional dynamics of **bacterial communities.** Worth mentioning, since most of my findings relate to principles of ecology, rather than specific organismal processes, it is likely that they are valuable and can be applied and/or extended to other organismal groups.

With the advance of technologies and the efforts of researchers across countries, we have been improving our knowledge on the physiology of microbes, their phylogeny, ecological interactions with other organisms, their distribution and functionality (Parks et al. 2018). Microbes play a fundamental role in ecosystem processes and provide essential ecosystem services, such as the maintenance of nutrient cycling, and the promotion of health of their associated hosts (e.g. *Homo sapiens*, domestic animals and crops). To better understand their role in ecosystem functioning and predict how the functions they performed changes in response to environmental fluctuations, it is important to identify the ecological mechanisms mediating their assembly, especially by partitioning species with different abundances within a given community.

Mechanisms of species coexistence

Microbial populations rarely live alone in natural habitats, in contrast, a collection of potentially interacting microbial populations usually occur together, thus forming microbial communities (Prosser et al. 2007). A key focus in community ecology is to understand the processes maintaining the coexistence of species and the building-up of species abundances from local to global spatiotemporal scales (Chesson 2000, Ovaskainen et al. 2017). Below I introduce some bases from both contemporary ecological processes and the historical contingency perspectives.

Ecological theories that were introduced to explain species coexistence originated in plant and animal communities, which has been debated along two distinct lines of reasoning: the niche and neutral theories. The ecological niche of a species was first proposed by George E. Hutchinson in the 1970s, which defined the dynamics of populations as a result of a set of environmental conditions (Holt 2009). This theory highlights the importance of deterministic processes that emphasizes ecological selection through abiotic factors (environmental filtering) or species interactions (antagonistic and synergistic interactions) [see (Chase and Leibold 2003)]. On the other hand, the neutral theory developed by Stephen P. Hubbell primarily focused on the tropical rain forest with high diversity (Hubbell 2001). The neutral theory emphasizes stochastic processes, such as random demography, probabilistic dispersal, speciation and unpredictable disturbances, as the main determinants of species abundance in a community. Recent characterizations of microbial communities provide evidence of deterministic processes in structuring community dynamics, such as environmental filtering or interspecies competition (Koeppel and Wu 2014). In other cases, stochastic processes were also found to be important drivers influencing microbial community assembly (Zhou et al. 2013). To date, mounting evidence suggests these two processes influence community assembly simultaneously, and their relative influences change over time and across ecosystem types (Stegen et al. 2012, Fillinger et al. 2019, Jiao et al. 2020). The challenge lies in quantifying their relative contributions structuring community assembly and the mechanisms underpinning their variations at different temporal and spatial scales (Tilman 2004, Gravel et al. 2006, Leibold and McPeek 2006, Adler et al. 2007).

In 2010, Mark Vellend introduced a framework unifying different concepts used to explain the patterns of species distribution in ecological communities. He did that by proposing that any given community is structured by four high-level ecological and evolutionary processes, namely speciation (or diversification), dispersal, selection and ecological drift (Vellend 2010, Vellend 2016). This conceptual synthesis avoids the dichotomy discrepancy between deterministic and stochastic processes, diminishes the redundant theories in community ecology, and provides a coherent and unified framework.

In addition to these contemporary processes, historical contingency also plays an important role in influencing community assembly, also known as ‘priority effect’ (Fukami 2015). Priority effects were reported mostly in plant and animal communities. For instance, it was shown that the initial establishment of plant species in an area

exerted an effect on newly assembled plant communities during grassland restoration and postfire conifer recruitment (Werner et al. 2020). Nowadays, studies in microbial ecology started to value the importance of the priority effect influencing microbiome assembly. For example, in a transplant experiment, the initial inoculated bacterial community reduced the success of taxa establishment from the later arriving community, thus indicating the assembly to depend on the order and timing of species arrival (i.e. dispersal) (Svoboda et al. 2018). Priority effect was also suggested to affect the interplay of diversification, drift and selection in structuring the infant gut microbiota, therefore, determining long-lasting consequences for infant health (Sprockett et al. 2018). Given that, a better understanding of priority effects in microbial ecology can offer new insights into predicting and directing the future states of microbiomes across divergent systems.

Together, understanding community assembly from both contemporary and historical perspectives will facilitate predicting the dynamics of microbial communities, such as community turnover during succession or recovery following disturbance events. Besides, this can also assist the estimation of ecosystem functioning, as the assembly processes and ecosystem functioning are tightly related (Leibold et al. 2017, Bannar-Martin et al. 2018, Mori et al. 2018). Ecosystem functioning is most likely maximized at intermediate levels of dispersal because dispersal limitation can constrain ecosystem functioning by reducing the arriving of better-suited species, while excess dispersal among heterogeneous habitats can reduce ecosystem functioning by introducing maladapted species with relative poorly environmental performance (Graham and Stegen 2017, Leibold et al. 2017).

Rare microbes matter

With the advance of sequencing technologies, studies often find a tremendous number of taxa in microbial communities, which are generally composed of a few abundant taxa co-occurring with numerous low abundance ones (Pedros-Alio 2012). Sogin et al. (2006) were the first to extensively describe the low abundance of specific microbial taxa and how they account for most of the phylogenetic diversity in deep sea water. In this study, a collection of low abundance taxa was first coined as the ‘rare biosphere’. In comparison to macroorganisms, rare bacteria and archaea might be less constrained by reproduction, being mostly limited by nutrient availability and their oligotrophic metabolism. Hence, the prokaryotic communities often show longer tails in the rank abundance curves in comparison to eukaryotic communities (Shade et al. 2018).

After the term was coined in 2006, the rare biosphere has been substantially investigated across diverse habitats, mostly using high-throughput sequencing approaches. However, the definition of the rare biosphere has not yet been unified. Most studies use amplicon sequencing to characterize the rare biosphere, even though it is challenging to distinguish real living taxa from relic DNA and artificial biases (e.g. chimeras generated by PCR and sequencing artefacts). Some studies defined the rare biosphere within communities, while others define it using the metacommunity approach (i.e. the whole dataset).

Furthermore, the thresholds used to partition the rare from the abundant ('common') biosphere are either based on relative abundances (e.g. <1%, <0.1% or <0.01%), or sequence number (e.g. < two sequences in a data set) (Galand et al. 2009, Campbell et al. 2011, Logares et al. 2014, Reveillaud et al. 2014).

Species in the rare biosphere do not necessarily show similar distribution patterns across space and time. Lynch and Neufeld (2015) have proposed five types of rarity, which presents the temporal abundance profiles of different rare groups. (*i*) The *r*-selected microbes that are **periodically recruited from the rare biosphere** can switch between abundant and rare states depending on periodic environmental conditions, such as daily changes and seasonality. (*ii*) Similar to the periodic dynamic group, the *r*-selected microbes that persist at relatively low abundances can be **occasionally recruited from the rare biosphere**, i.e. responding to occasional episodic or stochastic events, such as precipitation and stress release. (*iii*) The K-selected microbes that **periodically increase in abundance but are permanently rare** are adapted to exist at low relative abundance and respond with shifts in abundance to avoid predation; these taxa occupy narrow niches with increased susceptibility to starvation. (*iv*) Similar to above, the K-selected microbes that are **permanently rare with consistent abundance** persist at low-abundances by occupying narrow niches and escaping predation, although they might have increased susceptibility to starvation. (*v*) **Transiently rare** taxa are occasionally existent due to limited immigration, i.e. a recently immigrated species is rare when it first enters a community by chance via dispersal. Most current studies have been accounting for the rare biosphere as a unique group of taxa, however, to better predict the patterns and consequences of the rare biosphere in a given ecosystem, partitioning across these different types of rarity offers a promising approach.

For some time, the rare biosphere was believed to serve merely as a 'seed bank' of microbial taxa, whereas only abundant species were assumed to contribute to ecosystem functioning (Pedros-Alio 2006, Lennon and Jones 2011). However, recent studies revealed a significant proportion of the rare biosphere to be active, such as rare bacteria in marine ecosystems (Campbell et al. 2011) and rare bacteria, archaea and protists in freshwater ecosystems (Debroas et al. 2015, Inceoglu et al. 2015). A study in the plant rhizosphere has found 91% of plant-specific positive response bacteria belong to the rare biosphere (bacteria of <0.1% relative abundance) (Dawson et al. 2017). Some rare microbes, such as nitrite oxidizers, maintained at low abundance by high mortality rates were shown to transport a large portion of nutrients for other organisms when they lyse (Kitzinger et al. 2019). To be noticed, rare taxa sometimes even make disproportionate contributions to ecosystem functioning, such as a bacterium that makes up 0.006% in 16S rRNA gene amplicon library drives the major sulphate reduction (Pester et al. 2010). Besides the functionality performed directly by active rare taxa, the rare biosphere also serves as a genetic reservoir of species that allow the rapid colonization of new habitats or buffer community functioning against environmental fluctuations or species invasion (van Elsas et al. 2012, Nuccio et al. 2016, Liang et al. 2020). Given the importance of rare taxa, a better understanding of the ecological

processes mediating their distribution and long-term persistence in the environment is key to advance our ability to monitor, predict and manipulate distinct functional aspects of microbial communities across divergent systems.

Aim of this thesis

This thesis aims to investigate how distinct ecological processes interplay in mediating soil bacterial community assembly during succession. Greater attention is given to members of the microbial rare biosphere due to their high diversity and relevance for ecosystem functioning. This thesis focuses on disentangling the underlying mechanisms governing the assembly and successional dynamics of distinct fractions of soil bacterial communities based on organismal relative abundance distribution. I do that by answering the following questions:

- What is the contribution of distinct ecological processes to the structuring of microbial communities during succession?
- Does the use of different biomarkers (RNA or DNA) reflect distinct responses of the processes structuring bacterial community turnover during succession?
- Is the interplay of ecological processes structuring the rare and common bacterial biospheres different from each other? To what extent do the relative influences of ecological processes contribute to the distinct types of rarity in bacterial communities?
- How does dispersal by inundation influence the soil bacterial community structure at different stages of succession? Are these effects related to the selective pressure imposed on local bacterial communities, their diversity and/or historical contingency? Which groups of species are more affected by inundation, rare or common?

Outline of the thesis

The aforementioned questions are sequentially addressed throughout this thesis in three parts. The first part focuses on aspects of the microbial community succession, and biomarkers choice in sequencing approaches (**Chapter 2 and 3**). The second part (**Chapter 4 and 5**) explores the ecological processes structuring the rare biosphere of microbial communities. Because rare species constitute a large proportion of the diversity of microbial communities, unveiling the ecological mechanisms maintaining their existence helps to understand the rate of microbial community turnover during succession. Since dispersal is a crucial process that influences soil bacterial communities in salt marshes, the last part (**Chapter 6**) focuses on quantifying the influence of dispersal by inundation on soil bacterial communities coming from contrasting successional stages. All these questions were tested on a barrier island in the Wadden Sea, Schiermonnikoog, the Netherlands. This island provides a chronosequence of salt marshes that has been developing for over a century, providing the perfect scenario for

studying space per time replacement. Figure 1.1 provides a schematic overview of the topics addressed in the individual chapters, which are briefly introduced below.

Chapter 2 gives an overview of recent studies in microbial community succession using high-throughput sequencing approaches. In this chapter, I provide a general synthesis of the patterns and processes structuring microbial community assembly and turnover during both primary and secondary succession. Additionally, this chapter presents previous findings from studies performed at the island of Schiermonnikoog, the Netherlands, showing the primary succession of free-living soil bacterial and fungal communities from both taxonomic and functional perspectives and the bacterial communities associated with the rhizosphere of *Limonium vulgare*, a plant species that is found in all successional stages having plant cover.

Most of the studies profiling microbial communities are based on environmental DNA and do not distinguish between active and dormant organisms/populations, or even account for the existence of relic DNA in soil. **Chapter 3** compares environmental DNA and RNA based approaches for studying patterns of community assembly and turnover in soil bacterial communities. This is achieved by sampling soils across five successional stages along a primary succession chronosequence at Schiermonnikoog. Bacterial communities were characterized by sequencing both 16S rRNA amplicons from the extracted DNA and RNA transcripts of this gene. The relative influences of assembly processes were estimated using a combination of modelling approaches, i.e. phylogenetic and taxonomic community structure null model analyses.

Chapter 4 provides a perspective on community assembly processes structuring the microbial rare biosphere. Together with co-authors, I first examined the different definitions of the microbial rare biosphere and hypothesized that different types of rarity are driven by the distinct interplay of ecological processes. This chapter provides an overview of the framework used to study how the microbial rare biosphere is structured using a community assembly approach, as in **Chapter 3**.

Chapter 5 is an initial test of the hypotheses conceptualized in **Chapter** 4. Using the 16S rRNA data acquired in **Chapter 3**, I access how ecological processes structure the bacterial rare biosphere in space (five successional stages) and time (four sampling time points), thus leading to different types of rarity.

Salt marsh with constant influence by tidal inundation represents a critical interface between marine and terrestrial ecosystems (Mitsch and Gosselink 2000). However, dispersal processes of bacteria from marine to terrestrial ecosystems have not yet been rigorously/extensively studied. In **Chapter 6**, I focus on exploring the effect of dispersal promoted by sea water inundation on soil bacterial community assembly at different stages of succession. Together with co-authors, I performed a microcosm experiment where we quantified how dispersal by inundation influences soil bacterial communities from two distinct successional stages. A sterile sea water treatment was also used to distinguish the influence of inundation between abiotic (changes in soil characteristics) and biotic factors (bacterial dispersal). Bacterial communities were

characterized through 16S rRNA gene sequencing over a treatment period of 20 days. By doing so, we evaluate the influence of dispersal by inundation on soil bacterial communities from the early and late stages of succession, and how the influence is affected by diversity, selection and historical contingency.

Finally, results obtained throughout this thesis are synthesized and further discussed in **Chapter 7**. In particular, I discuss how assembly processes structuring different fractions of bacterial communities and maintaining taxa coexistence operate. I do that by integrating my findings with contemporary literature. In the end, I discuss current challenges and potential future directions in this field of research.

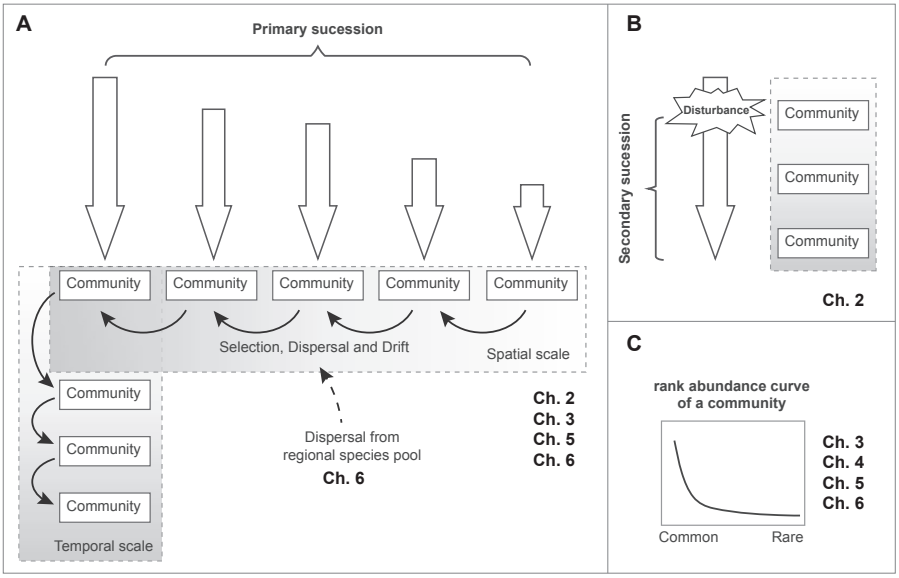

Figure 1.1 Schematic overview of the main research aspects explored in this thesis. The lengths of open arrows represent the successional time of microbial communities developed from non-colonized habitat. (**A**) A primary successional gradient is presented by a space-for-time substitution approach. (**B**) A secondary successional gradient is presented by a series of communities developed after a disturbance event. (**C**) A rank abundance curve represents the species distribution in a community based on the rank of species abundance.