Microbial community assembly in an evolving ecosystem

Dini Andreote, Francisco

CHAPTER 7

Microbial community assembly in soils: a synthesis

Francisco Dini-Andreote

Introduction

The study of community assembly and ecological succession yields fascinating knowledge on the mechanisms governing species interactions and temporal dynamics, from which principles and generalities can be drawn (e.g. Clements, 1916; Tilman, 1988). Despite being one of the oldest concepts in ecology, the fundaments of ecological succession have only recently been explored — to some extent — within microbial systems. Very often, microbial communities are assumed to simply be composed of a selection of organisms that collectively perform a single (or a combination of) function(s). Here we lack an understanding of the underlying mechanisms that allow these species to successfully co-exist, and to vary both in space and time in a systematic manner. To study such mechanisms, great strides have been made that reveal the role of environmental parameters, in particular pH and salinity, as drivers of soil microbial community divergences across local and continental scales (e.g. Fierer and Jackson, 2006; Lauber et al., 2009). However, there is, as yet, no definitive or complete picture of such systems. Clearly, microbial communities are living entities that possess both 'endogenous' (at the level of cell interactions) and 'exogenous' (at the level of local constraining factors) components that collectively modulate their spatiotemporal variations (see **chapter 1** for details) (Shade et al., 2013; Konopka et al., 2015). This thesis particularly focuses on the latter component; hence, the microscale patches where 'endogenous' mechanisms operate are beyond the scope of the study. Importantly, within the concept of community assembly applied here, it is intrinsically assumed that communities are formed through a dynamic interplay of several ecological processes (dispersal, selection, diversification and drift), which are underlined by the natural historical contingencies that the system experienced through succession.

This synthesis aims at providing a general discussion of the knowledge gathered throughout the chapters of this book in an intertwined manner. I start by describing an explicit consideration of the experimental design and how I made use of the study system to investigate patterns of community dynamics in a multiscalar manner. Next, I highlight the importance of merging macro ecological theory into studies of microbial ecology. These are often disconnected disciplines, and their merger offers unique opportunities for creative experimental designs and tests of ecological principles. Moreover, both perspectives raised in **chapter 1**; i.e. those from 'ecosystem' and from 'microbial' ecology viewpoints, are considered, as well as the answers to the research questions posed. I proceed in this synthesis by providing a perspective on the growing body of literature working towards the continuing integration of micro and macro ecological fields. Lastly, I describe the outlook of this thesis with emphasis on the general conclusions and potential future directions in this field of research.

The use of a soil chronosequence in a multiscalar approach to study microbial successions

Explicitly defining the ideal time frame to study microbial community succession in soils is not an easy task. In macro-organismal ecology, chronosequences have been used to enable

the study of temporal patterns in community succession over time periods that are longer than direct observation would permit (Walker et al., 2010). In these studies, an important aspect is that community dynamics is often investigated at temporal scales that represent 1–10 times the life span of the dominant species (Walker and del Moral 2003). Such an approach is clearly not applicable to this thesis or to other areas of microbiology. However, this raises a fundamental question with respect to the experimental design applied to the study system. It constitutes a critical point that lies back at the initial delineation upon which the chapters presented in this thesis were developed.

I choose to delineate the assessment of soil microbial communities in a multiscalar manner. I initially took into account previous macro-organismal research data in the Schiermonnikoog chronosequence that describe the patterns of above-ground community dynamics (Schrama, 2012). One implicit assumption used was that the rate of succession varies along the gradient, being high between the initial stages and progressively declining as succession proceeds. To account for these differences, the 'intervals' (in years of soil development) that define the stages vary intentionally along the chronosequence, i.e. relatively small intervals between the consecutive initial stages (stages 0 to 5, five years interval), and progressively increasing ones as succession proceeds (stages 5 to 35 and 35 to 65, thirty years interval; stages 65 to 105, forty years interval). In addition, to account for short-term temporal community dynamics within the local sites, I sampled each of the successional stages (along triplicated plots) on a bimonthly basis (i.e. in May, July, September and November) in the year 2012. With this design, the patterns of community dynamics were effectively investigated at both 'local' (within-stage) and 'regional' (across-stage) scales: the multiscalar approach.

Ecological theory as a foundation to study microbial community dynamics

"Theory is used to classify, interpret and predict the world around us. Without it, microbial ecology is merely the accumulation of situation-bound statements that are of limited predictive ability, providing microbiologists with few insights" (Prosser et al., 2007)

Perhaps one of the greatest challenges that microbial ecologists have been facing in the past decade is to work towards the continued integration of the vast knowledge gained throughout approximately one century of intense research (pioneered by the godfather of microbiology, Antonie van Leeuwenhoek 1632–1723, and later, around 1900-1920, picked up by the two great founders of microbial ecology Martinus Beijerinck and Sergei Winogradski) into general ecological theories. I see this quest for integration as a win-win situation: on the one side, microbial ecology profits from well-developed ecological theories aiming to understand and model species distributions (mostly designed upon macro-organismal research), whereas, on the other side, microbial systems constitute ideal systems to test and refine these general ecological principles (this thesis, Jessup et al., 2004; Cadotte et al., 2005). This is supported by the fact that microbes display much shorter generation times than any macro-organism. Moreover, the current knowledge in microbiology offers plenty of techniques that are convenient to grow and manipulate microbes under laboratory conditions. And given their ubiquitous occurrence in nature, microbial communities

across environmental gradients are often assumed (which is the case of this thesis) to behave like natural "laboratory" systems to advance fundamental ecological knowledge.

Along this thesis, I achieved the practical integration of both fields by addressing the processes that govern microbial community assembly and ecological succession in soils. A first example is given in **chapter 2**. Here, I found an unexpectedly high phylogenetic diversity of bacterial communities in the (bare sand) initial soil sites (stages 0 and 5), which progressively decreased as succession proceeds. This finding stood in striking contrast with general ecological expectations (for instance, the *diversity begets diversity* theorem; Whittaker, 1972). Also, late successional stage communities are likely to be significantly overdispersed (less phylogenetically clustered) than earlier-stage communities (Nemergut et al., 2016). It is expected that time allows for the progressive immigration of different species from the regional species pool, which results in the development of phylogenetically more diverse communities. This also corroborates the notion that the well-structured soils, at later stages, offer a multiplicity of nutrient and local conditions that support spatially-explicit niche diversities, to a greater extent than the initial (bare sand) soil stages. However, and in sharp contrast to these considerations, the joint application of short-term assessment of community turnover (in **chapter 2** termed 'within-stage phylogenetic β-diversity') and OTU co-occurrence network analyses revealed outcomes that feed alternative theories. Thus, the high turnover in species composition at the initial soil stages was probably linked to the natural turbance of the system, promoting temporally-driven niche partitioning. In this way, phylogenetically-diverse bacteria (at the level of OTU) coexisted, in a temporally dynamic manner, at the same local site. The high complexity of the OTU co-occurrence network found at the initial soil stages illustrated this point. This temporal niche partitioning, a concept borrowed from macro-organismal ecology, may have been the key driver, overwhelming the effect of spatial niche partitioning observed in the later stages, thus promoting biodiversity in the early stages of succession.

I continue the macro ecology – microbial ecology integration through **chapter 3**. Here, I addressed the knowledge gaps identified in **chapter 2**. In particular, I examined the relative influence of community assembly processes (i.e. stochastic and deterministic) through the systematic changes of bacterial communities along the successional gradient. For that, I started by designing a phased conceptual model based upon Vellend's synthesis in community ecology (Vellend, 2010), and further tested it using the empirical data from **chapter 2**. The knowledge gaps were two-fold, as expressed in the questions (*i*) and (*ii*):

(*i*) is the temporal turnover at the initial soil stages (termed 'temporal niche partitioning') mediated by stochastic or deterministic processes? To which degree does the relative influence of these processes vary along the salt marsh chronosequence? Here, the answers do not only have a direct implication for the mechanistic understanding of a previously observed phenomenon, but they may directly affect the predictability of the successional trajectories of bacterial communities in the system.

To address this question, I made use of a previously developed null-model approach that takes into account reconstructed phylogenetic trees of bacterial communities and uses permutation analyses to infer the relative influence of stochastic and deterministic processes (Stegen et al., 2012). Here, I found the high turnover at the initial soil stages to be mostly governed by stochastic processes (in this case, attributed to high rates of dispersal

and a relatively low level of selection). However, as succession proceeds, the relative influence of stochasticity declined, as that of deterministic selection progressively emerged (in this case, homogeneous selection — *sensu* Dini-Andreote et al., 2015, see below).

(*ii*) why, despite their 10- to 100-fold greater sizes, are bacterial communities at the late successional stages less phylogenetically diverse than those at the initial stages? How and to what degree does deterministic selection act within local sites along this system's progression?

I start answering these questions by showing that the progressive accretion of sodium in the system (concentration ranging from ~1.8–2.4% by weight at initial stages to ~13.8–14.4% at late stages) is the major selective force that increases in importance over time. This causes the emergence of a 'stringent' environmental filter (i.e. homogeneous selection) within local communities. As a result, the homogeneous selection imposed by salinity may have limited the successful establishment of a suite of (potentially dispersed) OTUs. It may also have acted as a filter on the resident OTUs through the course of succession, thus constraining community phylogenetic diversity. With respect to community size, soil nutrient status (in particular soil organic matter — SOM), revealing progressive increments, was a major driver. It was found to correlate significantly with increases in both bacterial and fungal community sizes over succession (**chapters 2** and **4**).

I posit that **chapter 3** stands out as a strong example of how a previously developed synthesis in community ecology (Vellend, 2010) can yield solid ground upon which adequate community assembly models can be conceptualized and empirically tested. Here, I successfully integrated the classes of processes raised in Vellend (2010) into a testable framework for microbial systems that links both primary and secondary (after a disturbance, community recovery) successions. As stated in **chapter 3**, the knowledge from this study constitutes a nexus point that enables multiple experimental designs aimed at validating the hypotheses raised in the developed conceptual models (Figures 3.1 and 3.4). Also, it highlights how an explicit consideration of spatial scale influences the results, the so-called 'scale-dependency' (see **chapter 3** and the discussion below for details).

Contrasting the successional patterns of bacterial and fungal communities

Do bacterial and fungal communities assemble differently along ecosystem successions? Despite the simplicity of this question, properly answering it requires a deep understanding of how ecological mechanisms operate in both of these organismal groups. This is further complicated by the fact that bacteria and fungi display very distinct cell sizes, ecological traits and ecophysiologies, which collectively determine their dynamics along an ecosystem gradient. Even though the data obtained in **chapter 2** (bacteria) and **4** (fungi) are not entirely conclusive, a close examination of the patterns obtained from the study system, and of the available literature, provides substantial novel material for a discussion on this subject.

The case of dispersal. Dispersal processes of microorganisms in terrestrial ecosystems have as-yet not being rigorously studied or quantified. Microbial dispersal over non-micro distances is typically a passive process that occurs through transport via wind, water

and 'hitchhiking' with mobile organisms (Nemergut et al., 2013). However, although passive dispersal if often considered to be stochastic, it is not entirely so: different microbial taxa vary in dispersal ability, making dispersal probability not entirely random among individual types. For example, Darcy and co-workers (2011) found Betaproteobacteria (in this case, *Polaromonas* spp.) to be globally distributed across high-altitude alpine environments. They suggested that the capacity to become dormant may enable these bacteria to travel, without a loss of viability, through the upper atmosphere. In contrast, other larger microbes, for example zoosporic fungi, may not display divergent geographic distributions at both the regional and global scales (an indication of dispersal limitation) (Naff et al., 2013). Furthermore, a theoretical treatise of how microbial cell sizes might affect dispersibility was brought to the front by Wilkinson et al. (2012). In this study, it was shown that microbial cells with dimensions <20 μm are more prone to passive dispersal than larger cells (>20 μm, the majority of fungi). Taken together, these empirical and theoretical arguments were later used in the literature (see Brown and Jumpponen, 2013; Schmidt et al., 2014) to hypothesize that the early-successional stages in environmental chronosequences might be more consistently colonized by bacteria than by fungi, due to differences in dispersal capabilities between these two organismal types. Adding to this argument is the fact that bacteria have a larger physiological breadth (e.g. photoautotrophs, heterotrophs, chemoautotrophs) than fungi (all heterotrophs), thus allowing them to successfully colonize a multiplicity of nutrient-poor sites. In contrast, all fungi depend on sources of fixed carbon and nitrogen that build up during succession.

Notwithstanding the above considerations, I found no significant shift from a bacterial- to a fungal-dominated system in the salt marsh chronosequence (like, for example, reported in Ohtonen et al., 1999). A potential explanation is that the early-successional stages in the salt marsh system are under marine influence and are therefore very dynamic, experiencing daily influences of the tides. A flooding event at these sites acts as a strong homogenising mechanism, shuffling the organisms spatially and contributing to the mass effect of immigration by originally marine-derived bacteria and fungi. As such, this system differs greatly from other ones (e.g. glacier forelands), where immigration occurs mostly passively at a smaller scale (e.g. through melt water or wind blows). Likewise, the effect of the tides also hints at an important mechanism by which organic carbon enters the system. For instance, marine water carries high loads of dissolved and particulate carbonaceous compounds that serve as prime resources based upon which microbial establishment ensues in the system. Later, SOM was shown to gradually accumulate along the chronosequence, thus explaining the microbial community size increases along the progression of the system (see meta-analysis at **chapter 4** for details).

The case of selection. No significant relation between the patterns of alfa-diversity of bacteria and fungi were found in the study system. Very likely, the mechanisms controlling OTU richness at local scale differed for the two groups. For example, whereas a homogeneous stringent bacterial filter was apparently provided by salinity (exerting a selection upon bacterial types), fungi displayed an erratic pattern that did not relate to any measured soil parameter at the local scale. However, at the regional scale, the pattern of fungal beta-diversity was effectively parameterized as driven by a combination of soil physical structure (clay and sand – **chapter 4**; as similarly observed for bacteria – **chapter 2**) and

SOM. Both beta-diversity patterns (bacterial and fungal) were later related to each other via Procrustes analyses, where they statistically segregate in a similar fashion along the succession (see meta-analyses at **chapter 4**). Thus, at a regional scale either these community structures are constrained by the same environmental filter(s) or by divergent ones that vary colinearly. A close examination of the data presented in **chapters 3** and **4** provides support for the following tenets. Whereas in **chapter 4**, SOM was hypothesized to drive fungal community dynamics at the regional scale, in **chapter 3** the relative influence of assembly mechanisms structuring bacterial communities at this scale was also linked to SOM dynamics. Thus, apart from other local drivers (i.e. salinity for bacteria), the organic compounds (derived either from marine or terrestrial sources), being heterogeneous, drive the structures of both bacterial and fungal communities at regional scales in this system. The environmental gradient reflects a signature from an initially 'brown' (marine) food web that progresses towards a 'green' (terrestrial) system (Schrama et al., 2012). However, a full appreciation of this tenet will depend on future experiments that more finely partition the SOM quantity and quality and their influences on both organismal types.

A functional trait based perspective on microbial community assembly

Whereas **chapters 2** to **4** focused on the phylogenetically-based distributions of the microbial communities in the study system, **chapters 5** and **6** investigated these at the level of their functional (genetic) potentials. The emphasis was on the ability of the different microbial communities to perform specified metabolic functions as well as on the inherent divergences in the trait distributions. A perspective on microbial community assembly is thus produced that tightly correlates the observed patterns and their implications for ecosystem functioning. Moreover, the focus on specific microbial traits provides insights into how local environmental selection operates in microbial populations, for instance by altering or maintaining trait frequencies and the genes underlying them (Martiny et al., 2015).

In **chapter 6**, by making use of metagenomic information, I illustrate how selected metabolic pathways can be reconstructed *in silico*. I thus used these as a comparative metric to access ecological divergences across the community types. Here, special attention was given to the key bacterial and archaeal genes that mediate N cycle transformations (see additional discussion below). Moreover, I also used the metagenomic datasets in a study that focused on the physiological assets of the different microbial communities (**chapter 5**). In doing so, I propose and use trait-based signatures that align with the marine-to-terrestrial transition. Thus, I reduced the complexity of the trait-based analyses to two distinct classes that confer adaptive value in the contrasting habitats (i.e. the marine versus terrestrial systems), here called the 'flight' or 'fight' response modus. The 'flight' modus — represented by the *in silico* reconstruction of bacterial chemotaxis and flagellar assembly pathways — prevailed at the early-successional stages as a reflection of the high diffusibility (connectivity) in these habitats. Under such conditions, motile chemosensory behaviour constitutes a key physiological asset that allows bacterial cells to perceive and exploit diverse microscale nutrient patches that occur in the system. On the other hand, in the more terrestrial (less connected) habitats, diffusibility becomes limited and chemically-

mediated mechanisms of biotic interaction very likely become progressively more dominant. This was illustrated by comparisons of the relative abundances of genes encoding different classes of antibiotic resistances (antibiotic resistance genes (ARGs) — used as proxies for chemical warfare across different populations — the 'fight' modus).

In summary, I focused on genes for processes that support the development of ecological theories, whereas other gene classes were not deeply studied. Truly, the analysis of microbial communities through metagenomics is limited by our ability to effectively annotate a large proportion of sequences and by the inadequacy of linking the relative abundance of a set of genes with microbial activities. However, metagenomic information is useful as a starting point to tackle the potential role of particular traits in community dynamics. The examples given above provide support for this contention, pointing at future avenues of research on the mechanisms underpinning the colonization of the system. Moreover, the microbial communities at all stages of the chronosequence were, as expected, metabolically highly redundant and the microbial traits varied considerably in the degree of conservation. These issues must be taken into account in future studies. For instance, by narrowing the variations of traits that are likely under selective pressure across different community types and environmental conditions. I foresee the development of an integrative framework that assesses the community variations based on trait relative abundances, allowing to manipulate these in natural settings to achieve the desired ecological functioning.

Microbially-mediated carbon and nitrogen dynamics in the chronosequence: an ecosystem description

One centrepiece of information in this thesis is the assessment of microbial genes involved in the processing of carbon and nitrogen in the salt marsh soils. The data obtained along the evolving ecosystem provide a unique microbial gene inventory to which that of other systems under differing stages of succession can be compared. I here posit that a microbial community represents '*more than the sum of its pieces*'. In other words, community status is the result of consecutive (successional) historical events that modulate the relative distribution of organisms (OTUs), in addition to the adaptive processes within these OTUs. This argument may sound obvious, but it has critical implications for the construction of ecological models that aim at monitoring salt marsh systems. We here need an explicit consideration as to how the complex eco-evolutionary mechanisms operates and allow different populations to successfully co-exist.

The case of carbon (C) in salt marsh soils was addressed in two chapters of this thesis. In **chapter 4**, I argued that salt marshes are thought to constitute effective C sinks in temperate climate zones (as in Chmura et al., 2003). As major players governing the C stocks in soil, fungi were assessed in the system at the level of composition and ecophysiological traits. Several insights were obtained. For instance, mycorrhizal fungi were rare at all successional stages (both ecto- and arbuscular mycorrhizal types), and saprotrophic fungi clearly varied in composition, constituting a niche partitioning signature along the marine-to-terrestrial gradient. Moreover, in **chapter 5**, the metagenomic analyses revealed patterns of genes for different CAZy (Carbohydrate-Active Enzymes) families and classes

along the gradient. Here, the initial soil stages were found to hold relatively high amounts of genes for CAZy proteins potentially involved in the degradation of marine-derived carbohydrates (e.g. laminarin and agar), whereas late-terrestrial-stage communities were found to mainly hold genes for CAZy proteins involved in the degradation of more recalcitrant substrates (e.g. chitin, hemicellulose, lignin).

In **chapter 6** I focused on the assessment and *in silico* reconstruction of microbial genes involved in the cycling of nitrogen (N). The motivation for this study was the current surge in the literature stressing the deleterious effects of exogenous N amendments leading to salt marsh degradation at a global scale (Deegan et al., 2012). Here, I explored the metagenomics data and used quantitative PCR assays to depict the systematic changes in the abundances of selected N-cycling genes in the chronosequence. I specifically used correlational analyses to infer the potential influence of soil parameters on these gene distributions. The data yielded examples of metabolic redundancy and niche partitioning, as, for instance, observed for the nitrifying communities (the case of the ammonia monooxygenase gene (*amoA*) allocated in ammonia-oxidizing bacteria (AOB) and archaea (AOA)). Similar findings were obtained for denitrifying genes (the nitrite reductase genes *nirS* and *nirK*, and the nitrous oxide reductase gene *nosZ* clades I and II). Importantly, chapter 6 provides several lines of evidence supporting the current theory on the ecology of N-cycling genes in soils across disparate ecosystems (e.g. Hallin et al., 2009; Jones et al., 2013).

Unifying microbial ecology and general ecological theory: a case for species richness*

**This section was written with basis in a perspective article 'in preparation' by James C Stegen et al. – used with permission.*

In the foregoing, I emphasized the need for a better integration of ecological theories into microbial ecology. However, there is a fundamental disconnection that separates microbial diversity patterns from biodiversity theory and the patterns derived from plant/animal ecology. This point has not yet been raised in the literature and, as such, was not considered during the development of this thesis. The discrepancy lies in how OTU/species richness (R) is measured in micro- versus macro-organismal research. In microbial ecology, OTU richness is estimated on a per-individual basis (R_{indiv}), where samples are rarefied to equal sequence depth per sample, in order to minimize any effects of sample sizes upon further analyses. In macro-organismal ecology however, species richness is quantified on a per-area basis (R_{area}).

Due to technical limitations imposed by the *modus operandi* in microbiology, setting an appropriate area to assign species richness is not a feasible or even sensible task. As such, to more closely align microbial to macro ecology theory, Stegen et al. (*in preparation*) proposed a novel approach that estimates OTU richness on a per-mass of substrate basis (R_{mass}). In brief, the rationality behind it is that sequencing depth is adjusted per community size/cell density (for instance, by using as a reference the bacterial 16S rRNA gene copy numbers per gram of substrate). In doing so, the use of R_{mass} as a metric is suggested as the analogue for R_{area} used in plant/animal research.

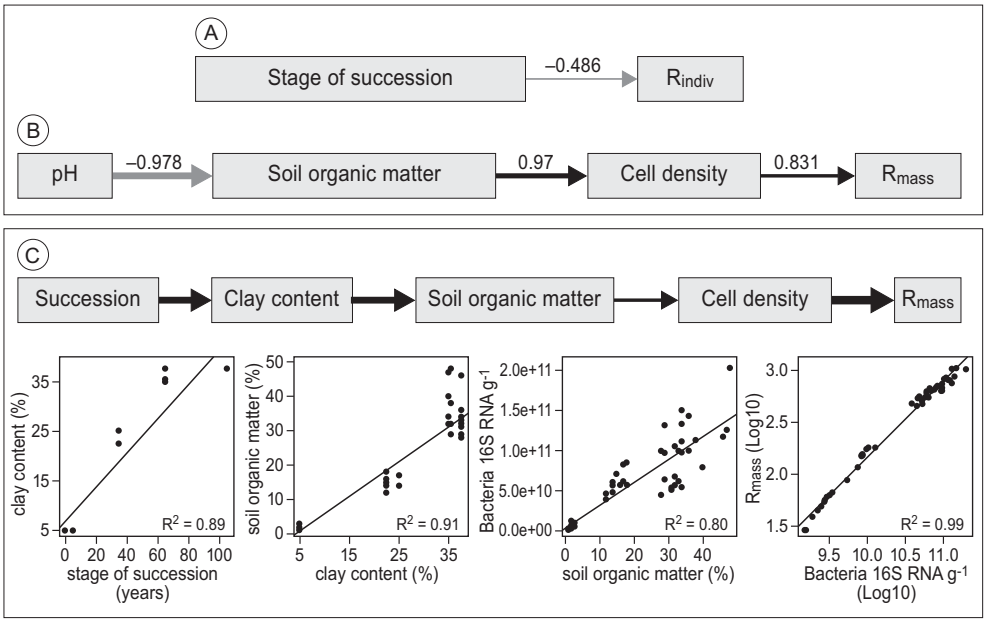

Figure 7.1. Structural equation models showing direct and indirect connections among variables that collectively explain variations in R_{indiv} (A) or R_{mass} (B) in the salt marsh chronosequence. (C) Univariate linear regressions relating shifts in soil physicochemical properties and how these ultimately relate to increments in community size — here estimated using as a proxy the bacterial 16S rRNA gene copies per gram of soil — influencing R_{mass}.

I tested this approach by making use of the bacterial community size data and the physicochemical parameters along the salt marsh chronosequence (**chapters 2** and **3**). Notably, the pattern of R_{mass} in our dataset was well explained by a structural equation model that depicts the influence of SOM on cell density, thus directly relating to R_{mass} (Figure 7.1B). This strongly supports a resource-based control over microbial R (in this case R_{mass}), a fact that was only weakly supported when R_{indiv} values were used in the model (Figure 7.1A). Moreover, the use of R_{mass} also allows for a more comprehensive overview of the factors controlling bacterial R in the chronosequence. This was achieved by correlating variables that relate to the natural progression of ecosystem formation, which results in a well-fitted linear regression that relates increments in community size with R_{mass} in the system (Figure 7.1C).

To finalize this short discussion, the example drafted above does not compromise the findings provided along this thesis or counters the current concepts in the literature. Instead, as suggested by Stegen et al. (*in preparation*), the R_{mass} approach aims to unify measurements of community richness taken across micro- and macro-organismal systems, in an effort towards rectifying this previously unrecognized disconnection. The conciliation of this feature between (micro- / macro-) systems will likely contribute towards the unification of general ecological theories. This is of particular importance for spatially explicit models and practical applications of theories built upon mechanisms controlling scale-dependent biodiversity patterns.

Outlook and future perspectives

A large body of literature assumes microbial community variation in space and time to be governed by environmental perturbations of different nature (type, magnitude, duration and periodicity). However, the mechanisms underpinning the shifts in community composition are often perceived to be idiosyncratic. Thus, a conceptual and testable framework is lacking, that takes into account ecological principles and quantitative metrics. Such a framework should integrate system divergences in a unified manner. Here I address this issue by parsing out the influence and type of environmental selection and stochastic processes in imparting community assembly and successional dynamics. In particular, in **chapter 3** I developed an experimentally testable model composed of *a priori* hypotheses that enable a systematic and quantitative understanding of community assembly and successional dynamics that is applicable to microbial communities across ecosystems.

This thesis also used different angles on the basis of which niche partitioning (niche specialization and differentiation) can be understood in microbial systems. The first example was shown in **chapter 2**, where the high dynamics in bacterial community turnover was raised as a mechanism that promotes species co-existence in time — temporal niche partitioning. Later, spatial niche specialization was shown to occur in both bacterial and fungal communities. This was investigated either at the level of individual OTUs, ecophysiologies (the case of fungi) and community-aggregated traits (metagenomics). Most importantly, I interrogated the system not uniquely with respect to constraining factors or soil nutrient status, but also regarding the contrasting nature of this gradient, which represents a shift from an initially 'brown' to a later 'green' system. The last example is based upon a study on the distribution of genes involved in N cycle transformations (**chapter 6**). Here, I reveal the patterns of metabolic redundancy and niche partitioning in the microbially-mediated N-cycle transformations in soils through the ecosystem formation process.

Finally, I here show that high-throughput sequencing technologies can be applied to natural communities beyond the common usage as a 'discovery-based' science. As defended extensively in this thesis, a change from reductionist approaches, emphasizing local organism distributions and environmental parameters, towards a more integrative ecological perspective, provides the basis of advances in fundamental microbial community ecology research. However, important gaps in our knowledge remain. Below, I summarize what I judge to be the most promising fields of research that this thesis may serve as a basis for:

(1) Experimental validation of microbial community assembly constraining factors and elucidation of how these operate through selection: along this study I showed evidence supporting the contention that the mechanisms controlling the relative influence, and the type, of deterministic selection are scale-dependent (**chapter 3**). Following this line of reasoning, I propose two main experimental follow-ups: (*i*) to assess the relationship between the (increases in) sodium (Na) in the system and the decreases in the relative influence of stochasticity (**chapter 3**). This issue was not fully addressed because the nature of the chronosequence did not allow the fine partitioning of Na increments as the ecosystem progressed. As such, it is unknown whether the relationship is truly linear or if there is a threshold at which small increases in Na concentration lead to a large decrease of stochasticity. Thus, the development of prospective experiments that

more finely manipulate Na concentrations and assess effects on these communities (for instance, by testing how 'homogeneous' selection emerges as Na concentrations in soils increases) is needed. This may also be applicable to terrestrial systems where other 'stringent' environmental filters have emerged, e.g. pH and temperature. (*ii*) The effect of SOM on the microbial (bacterial as well as fungal) communities was raised several times throughout this thesis. Particularly in this system, SOM consists of a heterogeneous mixture of organic compounds that are derived from marine and/or terrestrial sources and are at different stages of recalcitrance or degradation. Different from Na, the effect of SOM on microbial communities is more likely to be dependent on its composition rather than on its absolute concentration. Manipulative experiments that test this hypothesis can largely support the inferences made in the simulation model applied in **chapter 3** and the interpretations of the observed ecophysiologies and community trait shifts detailed in **chapter 4** and **5**. I propose these experiments to be accompanied by recently developed methods that allow for a detailed molecular characterization of SOM (e.g. Fourier transform ion cyclotron resonance mass spectrometry; FTICR-MS; see Tfaily et al., 2015). This endeavour should ideally disentangle the chemical complexity of SOM, thus allowing for a more precise determination of how it operates through selection of distinct populations within microbial communities.

(2) Testing and expanding conceptual models: in **chapter 3**, I provided two fundamentally linked conceptual models that integrate both primary and secondary successions in microbial systems. The empirical data from **chapter 2** lend support for the validation of one predicted scenario. Thus, the application of the quantitative approach, used to disentangle the relative influences of deterministic and stochastic processes (Stegen et al., 2012), across other successional trajectories, is necessary to test additional predictions. The defined framework constitutes a collection of straightforward hypotheses that are likely to be supported in a few systems and rejected in others. I judge both outcomes to be equally important, and data from system(s) that diverge(s) from the predictions must be taken as valid pieces of information for the further expansion of the models. Outcomes from this effort are key to achieve progress in our understanding of how community assembly processes and mechanisms underpinning their influence operate in disparate microbial systems.

(3) Integrating community genetic potentials with expression levels and activities in soils: Here, I based all functional and trait-based inferences raised along the chapters upon the genetic potential carried by the communities. This does not directly imply that these genes and functions are indeed active in the system. Hence, future studies should focus on the relative expression levels of genes marked as important for the selected processes or ecological traits. This must be conducted in fresh samples that realistically represent the status of the soil in the field. I propose, for instance, the quantification of gene expression related to chemotaxis/flagellar motility behaviour as opposed to genes encoding antibiotic resistance. These data would corroborate the arguments provided in **chapter 5**. Moreover, many other functional activities can be measured from fresh samples, in particular those of key importance for ecosystem functioning. Concerning that, I propose measurements in soil of nitrification, N_2 fixation and denitrification activities, which reliably discern the relative influence of each of these processes across

the chronosequence. Thus, such measurements would corroborate (or not) the descriptive study and metagenomics analyses provided in **chapter 6**.

(4) Integrating microbial communities into predictive ecosystem models: Microbes run the world, yet the discipline of microbial ecology is often neglected in studies aiming to mitigate and predict climate changes and anthropogenic impacts in soils. Along this thesis, I provided the arguments for the use of information on microbial communities and their successional trajectories as baselines for the development of comprehensive ecosystem models. Particularly in **chapter 4** to **6**, insightful considerations were given on organismal ecophysiologies, ecological traits and potential functions. The data gathered here can be better explored if incorporated into models that explicitly consider microbial populations and their functional differences, in order to improve ecosystem assessment, design and management. Advances in this area must take into account for not only the descriptive nature of the data, but also, most importantly, for the ecological mechanisms that underpin the observed patterns. Predictive ecosystem models will only be realistic if they have a strong ecological basis that supports the distribution, coexistence and desired activity of microorganisms in nature.

Concluding remarks

A myriad of mechanisms governing community assembly and turnover are discoverable, for instance by tracking the divergent microbial systems that occur in nature (Schrama, 2012; Stegen et al., 2012; Nemergut et al., 2013). The challenge lies in the design of strategies and concepts that unify such findings (Vellend, 2010). Moreover, when possible, one needs to integrate them conceptually, in order to make the complexity of microbial communities comprehensible (Prosser et al., 2007). Although exceptions do occur in biology, the conceptualization of fundamental ecological principles has often advanced the ability to predict the outcomes by bringing cross-system divergences into integrated frameworks (Nemergut et al., 2013). In this thesis, I took this contention as a motivation to 'dive into' the study system assuming it to represent a 'natural experimental setup' from which generalized ecological principles could be gleaned. The outcomes resulted in a vision of a collection of ecological mechanisms that operate across organismal types in an intertwined manner, underpinning the systematic changes in soil microbial community composition along ecosystem formation. The data and arguments forwarded throughout this thesis are at the forefront of the knowledge of ecological succession in soil formation gradients. The progress could only be achieved by infusing ecological principles and mechanisms into patterns of soil microbial community dynamics. The value of adopting such a strategy is reflected in the proposed broader applicability of the main findings brought up by this thesis. Despite being born out of a single ecosystem, the ecological mechanisms and models addressed here are applicable to any other system, ranging from miniaturized biofilms to gut systems, and/or to any other environmental/clinical setting. Finally, I posit that true scientific progress can only be achieved by expanding the scope upon which previous as well as new findings are applicable. As such, this thesis stands as an elegant example of how this can be achieved in soil community ecology.

References

Brown SP, Jumpponen A. (2014). Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. Molecular Ecology, 23: 481–497.

Cadotte MW, Drake JA, Fukami T. (2005). Constructing nature: laboratory models as necessary tools for investigating complex ecological communities. Advances in Ecological Research, 37: 333–353.

Chmura GL, Anisfeld SC, Cahoon DR, Lynch JC. (2003). Global carbon sequestration in tidal, saline wetland soils. Global Biogeochemical Cycles, 17: 1111–1133.

Clements FE. (1916). Plant succession: an analysis of the development of vegetation. Carnegie Institution of Washington: Washington, DC, USA.

Darcy JL, Lynch RC, King AJ, Robeson MS, Schmidt SK. (2011). Global distribution of *Polaromonas phylotypes* – evidence for a highly successful dispersal capacity. PLoS One, 6: e23742.

Deegan LA, Johnson DS, Warren RS, Peterson BJ, Fleeger JW et al. (2012). Coastal eutrophication as a driver of salt marsh loss. Nature, 490: 388–392.

Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. Proceedings of the National Academy of Sciences USA, 112: E1326–E1332.

Fierer N, Jackson RB. (2006). The diversity and biogeography of soil bacterial communities. Proceedings of the National Academy of Sciences USA, 103: 626–631.

Hallin S, Jones CM, Schloter M, Philippot L. (2009). Relationship between N-cycling communities and ecosystem functioning in a 50-year-old fertilization experiment. The ISME Journal, 1: 1–9.

Jessup CM, Kassen R, Forde SE, Kerr B, Buckling A et al. (2004). Big questions, small worlds: microbial model systems in ecology. Trends in Ecology & Evolution, 19: 189–197.

Jones CM, Graf DR, Bru D, Philippot L, Hallin S. (2013). The unaccounted yet abundant nitrous oxide-reducing microbial community: a potential nitrous oxide sink. The ISME Journal, 7: 417–426.

Konopka A, Lindemann S, Fredrickson J. (2015). Dynamics in microbial communities: unraveling mechanisms to identify principles. The ISME Journal, 9: 1488–1495.

Lauber CL, Hamady M, Knight R, Fierer N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community composition at the continental scale. Applied and Environmental Microbiology, 75: 5111–5120.

Martiny JB, Jones SE, Lennon JT, Martiny AC. (2015). Microbiomes in light of traits: a phylogenetic perspective. Science, 350: aac9323.

Naff CS, Darcy JL, Schmidt SK. (2013). Phylogeny and biogeography of an uncultured clade of Snow Chytrids. Environmental Microbiology, 15: 2672–2680.

Nemergut DR, Knelman JE, Ferrenberg S, Bilinski T, Melbourne B et al. (2016). Decreases in average bacterial community rRNA operon copy number during succession. The ISME Journal. *In press*.

Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM et al. (2013). Patterns and processes of microbial community assembly. Microbiology and Molecular Biology Reviews, 77: 342–356.

Ohtonen R, Fritze H, Pennanen T, Jumpponen A, Trappe J. (1999). Ecosystem properties and microbial community changes in primary succession on a glacier foreland. Oecologia, 119: 239–246.

Prosser JI, Bohannan BJM, Curtis TP, Ellis RJ, Firestone MK et al. (2007). The role of ecological theory in microbial ecology. Nature Reviews Microbiology, 5: 384–392.

Schmidt SK, Nemergut DR, Darcy JL, Lynch R. (2014). Do bacterial and fungal communities assemble differently during primary succession? Molecular Ecology, 23: 254–258.

Schrama M, Berg MP, Olff H. (2012). Ecosystem assembly rules: the interplay of green and brown webs during salt marsh succession. Ecology, 93: 2353–2364.

Schrama M. (2012). The assembly of a saltmarsh ecosystem: the interplay of green and brown webs. University of Groningen: Groningen, The Netherlands.

Shade A, Caporaso JG, Handelsman J, Knight R, Fierer N. (2013). A meta-analysis of changes in bacterial and archaeal communities with time. The ISME Journal, 7: 1493–1506.

Stegen JC, Hurlbert AH, Bond-Lamberty B, Chen X, Anderson CG et al. (2016). Aligning the measurement of microbial diversity with macroecological theory. *In preparation*.

Stegen JC, Lin X, Konopka AE, Fredrickson JK. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. The ISME Journal, 6: 1653–1664.

Tfaily MM, Chu RK, Tolić N, Roscioli KM, Anderton CR et al. (2015). Advanced solvent based methods for molecular characterization of soil organic matter by high-resolution mass spectrometry. Analytical Chemistry, 87: 5206–5215.
Tilman D. (1988). Plant strategies and the dynamics and structure of plant communities. Princeton University Press: Princeton, NJ, USA.
Vellend M. (2010). Conceptual synthesis in community ecology. The Quarterly Review of Biology, 85: 183–206.
Walker LR, del Moral R. (2003). Primary succession and ecosystem rehabilitation. Cambridge University Press: Cambridge, UK.
Walker LR, Wardle DA, Bardgett RD, Clarkson BD. (2010). The use of chronosequences in studies of ecological succession and soil development. Journal of Ecology, 98: 725–736.
Whittaker RH. (1972). Evolution and measurement of species diversity. Taxon, 21: 213–251.
Wilkinson DM, Koumoutsaris S, Mitchell EAD, Bey I. (2012). Modeling the effect of size on the aerial dispersal of microorganisms. Journal of Biogeography, 39: 89–97.