RULES AND PATTERNS OF MICROBIAL COMMUNITY ASSEMBLY

by

SHAWN PAUL BROWN

B.S., University of Oregon, 2009

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Abstract

Microorganisms are critically important for establishing and maintaining ecosystem properties and processes that fuel and sustain higher-trophic levels. Despite the universal importance of microbes, we know relatively little about the rules and processes that dictate how microbial communities establish and assemble. Largely, we rely on assumptions that microbial community establishment follow similar trajectories as plants, but on a smaller scale. However, these assumptions have been rarely validated and when validation has been attempted, the plant-based theoretical models apply poorly to microbial communities. Here, I utilized genomics-inspired tools to interrogate microbial communities at levels near community saturation to elucidate the rules and patterns of microbial community assembly. I relied on a community filtering model as a framework: potential members of the microbial community are filtered through environmental and/or biotic filters that control which taxa can establish, persist, and coexist. Additionally, I addressed whether two different microbial groups (fungi and bacteria) share similar assembly patterns. Similar dispersal capabilities and mechanisms are thought to result in similar community assembly rules for fungi and bacteria. I queried fungal and bacterial communities along a deglaciated primary successional chronosequence to determine microbial successional dynamics and to determine if fungal and bacterial assemblies are similar or follow trajectories similar to plants. These experiments demonstrate that not only do microbial community assembly dynamics not follow plant-based models of succession, but also that fungal and bacterial community assembly dynamics are distinct. We can no longer assume that because fungi and bacteria share small propagule sizes they follow similar trends. Further, additional studies targeting biotic filters (here, snow algae) suggest strong controls during community assembly, possibly because of fungal predation of the algae or because of fungal utilization of algal exudates. Finally, I examined various technical aspects of sequence-based ecological investigations. These studies aimed to improve microbial community data reliability and analyses.

RULES AND PATTERNS OF MICROBIAL COMMUNITY ASSEMBLY

by

SHAWN PAUL BROWN

B.S., University of Oregon, 2009

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Approved by:

Major Professor
Dr. Ari. M. Jumpponen

Copyright

SHAWN PAUL BROWN

2014

Abstract

Microorganisms are critically important for establishing and maintaining ecosystem properties and processes that fuel and sustain higher-trophic levels. Despite the universal importance of microbes, we know relatively little about the rules and processes that dictate how microbial communities establish and assemble. Largely, we rely on assumptions that microbial community establishment follow similar trajectories as plants, but on a smaller scale. However, these assumptions have been rarely validated and when validation has been attempted, the plant-based theoretical models apply poorly to microbial communities. Here, I utilized genomics-inspired tools to interrogate microbial communities at levels near community saturation to elucidate the rules and patterns of microbial community assembly. I relied on a community filtering model as a framework: potential members of the microbial community are filtered through environmental and/or biotic filters that control which taxa can establish, persist, and coexist. Additionally, I addressed whether two different microbial groups (fungi and bacteria) share similar assembly patterns. Similar dispersal capabilities and mechanisms are thought to result in similar community assembly rules for fungi and bacteria. I queried fungal and bacterial communities along a deglaciated primary successional chronosequence to determine microbial successional dynamics and to determine if fungal and bacterial assemblies are similar or follow trajectories similar to plants. These experiments demonstrate that not only do microbial community assembly dynamics not follow plant-based models of succession, but also that fungal and bacterial community assembly dynamics are distinct. We can no longer assume that because fungi and bacteria share small propagule sizes they follow similar trends. Further, additional studies targeting biotic filters (here, snow algae) suggest strong controls during community assembly, possibly because of fungal predation of the algae or because of fungal utilization of algal exudates. Finally, I examined various technical aspects of sequence-based ecological investigations. These studies aimed to improve microbial community data reliability and analyses.

Table of Contents

vii

List of Figures

List of Tables

Acknowledgements

Dedication

I dedicate this dissertation to my wife (Mary Brown) and children (Vera, Lyndon, and Millie Brown). Without them, I would be wandering aimlessly through the proverbial forest. Thank you for keeping me grounded.

Preface

All chapters for this dissertation were written and formatted to specific journals as laid out in the footnotes for each chapter title page. Therefore, in-text citation and reference formats were followed for said journal and differ slightly for each chapter. Chapter 1 serves as an introduction and was not submitted for publication but follows the same formatting as Chapter 2 for consistency. Chapters 2, 4, and 5 have citations and references formatted for *Molecular Ecology*. Chapter 3 is formatted following requirements for *The ISME Journal*. Chapters 6 and 7 are formatted following requirements for *Fungal Ecology*.

Chapter 1 - Introduction

Using a basic definition of an ecological community as an assemblage of species that occur within the same space and time (Begon *et al.* 2006), it is easy think of communities as something stationary rather than the outcome of a dynamic and intense struggle that shape and structure what we see presently. In practice, the study of community ecology is the study of all contemporary and historical interactions that structure what we observe in the field. For microorganisms, for most who's natural histories, functionality, and metabolic processes remain unclear, determining current interactions is problematic let along historic factors.

What are these historical events that structure and shape observable microbial communities? Although a multitude of factors can influence microbial community structure, a conceptual framework in which we envision community establishment and persistence dictated by availability of propagules filtered by a host filter (if applicable), environmental filter and a biotic filter is useful (*see* Diamond 1975, Keddy 1992, Jumpponen & Egerton-Warburton 2005, Koide *et al.* 2011).

Here, the initial pool of all available propagules (including spores, hyphal fragments, airborne bacteria, and any other viable organismal fragments) is the locally available propagating particles that have deposited onto a given substrata. This initial propagule bank only consists of readily available propagules. That is, dispersal capability/restrictions of any organism coupled with the current range of a given taxa will dictate if that particular organism belongs to the initial propagule pool at a given location. This represents a dispersal/colonization filter prior to any further community filtration. Of course, just because these propagating units are in the area does not mean that they will establish. Filters that limit which propagules can establish and persist functionally reduce this initial pool. Potential propagules can be filtered based on host preference. That is, for some microbes (*e.g.* mycorrhizal fungi, symbiotic bacteria such as *Rhizobia*, obligate endophytes, and obligate fungal biotrophs (mildews, smuts, and rusts)) the presence of a suitable host is paramount for establishment and without a proper host, these propagules cannot colonize. These host filters thereby reduce the size of the potential community. The remaining potential propagules can be further restricted through an environmental filter. The physiological requirements of microbes may not be met at the location of the colonization event. If the location or colonized substrate does not meet the physiological

restrictions (*e.g.* suitable habitat availability (aquatic vs. terrestrial, regional climatic conditions, narrow pH range, nitrogen availability, etc.) of potential colonizers, then those microbes will not survive and the potential community is further reduced. Finally, further reduction in the communities can occur through a biotic filter. In this instance, organisms that remain are already physiologically adapted to survive in the location but the community may be further or structured reduced through the outcome of competitive interactions or facilitation associations. In addition to competitive or facilitation pressures, this biotic filter can be manifested as predation pressures as well as through allelopathy. Essentially, any biotic or ecological interaction that can impact the ability of any organism to persist can act as a filter that assists in structuring community composition. What remains after these filtration events is the "climax community" (Clements 1916), an antiquated and inaccurate term suggesting that these communities are no longer in flux. Of course the notion of a 'stable climax community' is a *non sequitur* and can be better thought of as an extant or sampled community as communities are constantly bombarded with external and internal stressors that help refine and shape community dynamics and are always at risk of severe large-scale perturbations that may potentially reset communities along a secondary successional trajectory. It is largely these assembly filters that drive my research interests for this dissertation. What are the physiological restrictors that help filter microbial communities? How do plants and other organisms influence the community structure of microbes? How are microbial communities shaped by geographic distance? These are a few of the guiding questions that influence my research.

Much of the early ecological theory about community assembly was derived from plant communities (Matthews 1914; Clements 1916; Pearsall 1918; Tansley 1920). In fact, until very recently, microbial communities were assumed to follow very similar patterns as plants as there was an assumed coupling between plants and microbes in the nature of positive feedback loops (Reynolds *et al.* 2003) that directly shape microbial communities based on plant communities. These earlier theories dominated thinking about not only plant communities but often community ecology as a whole. It was not until more recently that the divergence between plants and microbial ecology began with the realization that ecological patterns of microbes are often divergent from plants (*e.g.* Bryant *et al.* 2008; Fierer *et al.* 2011).

Rapid technological advances are allowing for deeper and more detailed interrogation of microbial communities. The emergence of 454-sequencing (Margulies *et al.* 2005) and

subsequent next generation sequencing (NGS) technologies has allowed for an unprecedented examination of microbial communities and allows for near sampling saturation of community members. This deep coverage has lead to many great advances in our understanding of microbial biodiversity and the near saturation allows for relevant comparisons across taxa with respect to ecological patterns. A large part of my research program is dedicated exploring if assembly patterns of plants are similar to microbes or if fungi and bacteria possess similar assembly patterns. That is, are there fundamental rules that govern the way in which organisms assemble; or, do we need to focus on taxa-specific rules of assembly as plants, animals, bacteria, and fungi likely follow different assembly trajectories? The question of "are plants equal to microbes" is a simple question but answering such a question is multifaceted and complex. To answer this question, I rely on high-throughput environmental sequencing to query the microbial community within the constraints of an experimental design that allows for detailed examination of assembly filters to elucidate community assembly patterns.

My first research chapter (Chapter 2) focuses most directly on the similarity of plant assembly and microbial assembly. Here (*see* Brown & Jumpponen 2014), we examined if bacteria and fungi follow primary successional assembly patterns similar to those or vegetation across a well-documented deglaciated chronosequence. Lyman Glacier is located in Washington, USA and has receded at documented rates over the past *ca*. 120 years. For this reason, a well-documented (Jumpponen *et al.* 1998) space for time substitution along the glacier forefront has been developed. This newly exposed substrate has become open for colonization and establishment of plants, microbes, and animals. Plant establishment has been documented over the last 20 years (reviewed in Jumpponen *et al*. 2012) as well as mycological examination utilizing everything from sporocarp examination (Jumpponen *et al.* 1999), PLFAs (Ohtonen *et al.* 1999), to rDNA sequence examination (Jumpponen 2003). At the Lyman Glacier basin, like many documented primary successional locations, plant communities increase in richness and diversity and display great taxon replacement as successional age increases (del Moral & Wood 1993; Walker & del Moral 2003). At Lyman, early plant communities are dominated by few species that largely lack mycorrhizal symbionts (namely *Saxifraga ferruginea*) and as substrates age, other plant species that rely on mycorrhizal associations come to prevalence (see Fig. 1 in Jumpponen *et al.* 2012). To investigate if fungi and bacterial follow assembly patterns and trajectories similar to plants, we utilized environmental sequencing by targeting taxonomically

informative loci (Internal Transcribed Spacer 1 [ITS1] of the rRNA ribosomal gene complex for fungi and variable regions V1 and V2 of the 16S rRNA gene complex for bacteria) to assess community assembly both across successional age and with vegetation as we sampled within the rhizosphere of different plants with different mycorrhizal association types as well as non-vegetated bare soil. This experimental design allows for not only successional shifts of organisms that are physiologically adapted to life on a harsh alpine glacier forefront [environmental filter] but also with interactions with different vegetation types [host filter/biotic filter].

Chapter 3 also takes advantage of the space for time substitution across Lyman Glacier. The results of Chapter 2 suggest fungi and bacteria exhibit different patterns of community filtering evidenced by vegetation establishment being more important in structuring bacterial communities than fungi. Additionally, results hinted that bacteria and fungi differ in the successional patterns. Bacteria converged with respect to community structure over time (*see* Fig. 6 in Brown & Jumpponen 2014) whereby bacterial communities further from the glacier terminus (longer substrate exposure) were more similar to each other than younger communities. Fungi neither converged nor diverged with respect to community structure but may converge with respect to genetic relatedness (*see* Fig. 7 in Jumpponen *et al.* 2012). That is, older fungal communities on the Lyman Glacier forefront may be more closely related to each other than earlier communities. It was this finding that prompted the investigation into whether fungal and bacterial primary succession are similar with respect to how communities are phylogenetically related to each other. This allows for elucidation of microbial successional dynamics coupled with the integration of evolutionary history to see if primary successional establishment and assembly dynamics are constrained by how related organisms are. To do this, I utilized various metrics of phylogenetic diversity to assess how the genetic similarity of fungal and bacterial communities changes over successional age. The results from this further solidify that fungi and bacteria cannot and should not be assumed to behave similarly. To understand these distinct taxon groups, it is not enough to take fungi and bacteria at face value by considering all bacteria of fungi as the unit of ecology but one must examine deeper and lower taxonomic levels to grasp fully the way in which certain fungi and bacteria respond during active community assembly.

The first two research chapters focused on primary successional dynamics, which by definition includes the incorporation of both the environmental and biotic filters. The next two

research chapters focus the effects of biotic filtering. One focuses on direct effects of biotic filtration and one examines the population dynamics of the biotic filter itself. These chapters take advantage of a naturally occurring biotic filtering agent. Snow algae are cosmopolitan and a diverse group of Chlorophyta that can be found anywhere snow packs persist late enough into the growing season to fuel massive autotrophic blooms of algae yet often can be found in discrete algae patched with a definite visual boundary (Remias *et al.* 2010; Fujii *et al.* 2010). It is the occurrence of these discrete patches of snow algae that allows for examination of snow algae as a biotic filter. Snow is not a passive repository of aerially deposited fungal spores, hyphal fragments and bacteria, rather it is a unique ecosystem that it teeming with metabolically active organisms despite the harsh environment (Carpenter *et al.* 2000; Hell *et al.* 2013). Because snow houses an abundant microbial community, even though it is of a lower complexity and richness that other aquatic systems, we can test if the presence of snow algae influences and acts as a biotic filter structuring the microbial community. To do this, paired snow samples were collected from snow visibly colonized by algae and adjacent (3m away) non-colonized snows in two consecutive years (2011 and 2012) across a semi-continental scale (Washington and Colorado) to examine biotic structuring of fungal communities locally, regionally, semi-continentally, and temporally. This fungal community structuring and filtering is the focus of Chapter 4. During these analyses, it became apparent that the fungus-specific primers that were utilized to amplify the ITS2 (Internal Transcribed Spacer 2) region of the fungal rRNA gene repeat are prone to non-fungal off-target amplification. Fortuitously, the majority of the off-target amplification belonged to snow algae. This allowed for an in-depth analysis of snow algae communities (Chapter 5). Snow algae patches are complex and diverse communities that consist of many taxa. The two most abundant taxa (delineated by 97% sequence similarity as binned into operational taxonomic units) were found in great abundance in all samples and were deemed abundant enough to conduct population analyses. All sequences from these two Operational Taxonomic units (OTUs) were extracted from the sequence file and reanalyzed using a haplotype analysis approach. This resulted in the intriguing finding that even though snow algae patches contain many algal taxa, within each discrete patch each taxon is highly clonal and spatially structured.

The final two chapters are more technical in nature. Next generation sequencing (NGS) technologies have allowed for rapid and cost-effective generation of sequence data and this

democratization of sequencing has resulted in many new researchers undertaking the study of microbial community ecology using these sequence data. However, despite the explosion of environmental sequencing, these NGS technologies are still less than a decade old (Margulies *et al.* 2005). There is still much that we do not understand about the NGS data, especially with regard to the generation of potentially spurious sequences. If diligent sequence quality control is not undertaken, these spurious sequences may be retained throughout all analyses and confound and skew ecological results. Bacterial sequence processing is more advanced. This is mainly attributable to broader and better curated databases that facilitate screening for erroneous sequences, partly because of the simplicity of bacteria and partly because of the fact that bacterial environmental sequencing using NGS tools is several years older than fungal sequencing. Fungal community analysis using NGS tools is only ~5 years old (see Buee *et al.* 2009, Jumpponen & Jones 2009). We are still learning how best to analyze these sequences and struggle to comprehend the most appropriate way to approach sequencing projects and bioinformatic processing and analyses (gene region choice, how to handle rare OTUs, appropriate thresholds for OTU calling, etc.). Traditionally, the Internal Transcribed Spacer (ITS) regions have been targeted for fungal community analyses and have been labeled the official fungal barcode for biodiversity (Schoch *et al.* 2012). This is in part due to the highly variable nature of ITS sequences that are flanked on both sides with highly conserved gene regions; this permits species level distinction across the Kingdom Fungi. The ITS regions also have larger and more complete database representatives that may aid in labeling sequences to taxa, however this expanded database often lacks support and curation that may lead to many erroneous and inaccurate identifications (Nilsson *et al.* 2006). Despite the selection of ITS as the fungal barcode, ITS regions do have a downside (Kiss 2012). They cannot be aligned except for closely related taxa and any forced alignment results in concatenated gaps and resultant alignments do not recapitulate phylogeny. For this reason, it is impossible to incorporate any alignment-based analyses into broader fungal community ecology, including but not limited to evolutionary ecological analysis. In contrast to ITS regions, the Large Subunit (LSU) region of the rRNA gene repeat can be aligned and still maintain highly variable regions flanked by conserved regions. Although there is much less database support for LSU, it may be a viable alternative to ITS that allows integration of evolutionary and community ecology. The overarching question for Chapter 6 is: do community-wide analyses that use ITS regions and the

LSU regions of the rRNA gene repeat differ in their community ecological measures? That is, can we sequence either region and retrieve similar ecological results? It turns out that by and large, the answer is yes; sequencing either region provides congruent community-wide ecological results. This now allows for great flexibility for any researcher as now they are not limited to a gene region that excludes global alignments and precludes detection of any phylogenetic signal.

Chapter 7 focuses on rare OTUs. One of the early 454-based examinations of the fungal community (Tedersoo *et al*. 2010) suggested that global sequence singletons (OTUs that only contain one sequence) are most likely artifacts and should therefore be eliminated from the dataset prior to downstream analyses. Since this time, the effect of rare OTUs has largely been ignored for fungal community analysis. However, since this important paper, several bioinformatic measures (*e.g.* better chimera screening, preclustering to reduce sequencing errors, and better database representation) have become the standard for sequence processing. The question is: have we become so good in filtering and screening sequence data for erroneous sequences that we need not worry about singleton OTUs anymore? To answer this I reanalyzed global singletons from 12 different sequencing projects (ranging from ITS1, ITS2, LSU as well as a range of NGS platforms: 454-FLX, 454-Titanium, Illumina MiSeq). I found that most singletons are not artifacts but this largely relies on gene region. However, even though the average singleton OTU is likely valid, there are still many who's validity we cannot be assured. This, coupled with the biological truism that rare taxa (especially true for fungi) are unlikely to be contributing a great deal to overall ecosystem processes, means we should continue to cull singleton OTUs per Tedersoo *et al.* (2010). I extend this recommendation from Tedersoo et al. (2010) to remove all rare taxa from analysis as inclusion of rare and non-important taxa that may be artifactual likely inflates richness estimators may skew measures of community dynamics.

References

Begon M, Townsend CR, Harper JL. 2006. *Ecology: from individuals to ecosystems.* Blackwell Publishing, Oxford, UK.

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

Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL. 2008. Microbes on mountainsides: contrasting elevational patterns of bacteria and plant diversity. *Proceedings of the National Academy of Science, USA* 105: Supplement 1 11505-11511.

Buee M, Reich M, Murat C, Morin E, Nilsson RH, Uroz S, Martin F. 2009. 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist* 184: 449-456.

Carpenter EJ, Lin SJ, Capone DG. 2000. Bacterial activity in South Pole snow. *Applied and Environmental Microbiology,* 66: 4514-4517.

Clements FE. 1916. *Plant succession: an analysis of the development of vegetation.* Carnegie Institution of Washington, Washington, D.C., USA

del Moral R, Wood DM. 1993. Early primary succession on a barren volcanic plain at Mount St. Helens, Washington. *American Journal of Botany* 80: 981-991.

Diamond JM. 1975. Assembly of species communities. In: Cody ML, Diamond JM (Eds.) *Ecology and evolution of communities.* Harvard University Press, Cambridge, MA, USA, pp. 342-344.

Fierer N, McCain CM, Meir P, Zimmermann M, Rapp JM, Silman MR, Knight R. 2011. Microbes do not follow the elevational diversity patterns of plants and animals. *Ecology* 92: 797-804.

Fujii M, Takano Y, Kojima H, Hoshino T, Tanaka R, Fukui M. 2010. Microbial community structure, pigment composition, and nitrogen source of red snow in Antarctica. *Microbial Ecology.* 59: 466-475.

Hell K, Edwards A, Zarsky J, Posmirseg SM, Girdwood S, Pachebat JA, Insam H, Sattler B. 2013. The dynamic bacterial communities of a melting High Arctic glacier snowpack. *The ISME Journal,* 7: 1814-1826.

Jumpponen A. 2003. Soil fungal community assembly in a primary successional glacier forefront ecosystem as inferred from analyses of rDNA sequences. *New Phytologist* 158: 569-578.

Jumpponen A, Mattson K, Trappe JM, Ohtonen R. 1998. Effects of established willows on primary succession on Lyman Glacier forefront, North Cascade Range, Washington, U.S.A.: evidence for simultaneous canopy inhibition and soil facilitation. *Arctic and Alpine Research* 30: 31-39.

Jumpponen A, Trappe JM, Cázares E. 1999. Ectomycorrhizal fungi in Lyman Lake basin: a comparison between primary and secondary successional sites. *Mycologia* 91: 575-582.

Jumpponen A, Egerton-Warburton LM. 2005. Mycorrhizal fungi in successional environments – a community assembly model incorporating host plant, environmental and biotic filters. In: Dighton J, White JF, Oudemans P (Eds.) *The fungal community.* CRC Press, New York, NY, USA, pp. 139-180.

Jumpponen A, Jones KL. 2009. Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytologist* 184: 438-448.

Jumpponen A, Brown SP, Trappe JM, Cázares E, Strömmer R. 2012. Twuenty years of research on fungal-plant interactions on Lyman Glacier forefront – lessons learned and questions yet unanswered. *Fungal Ecology* 5: 430-442.

Keddy PA. 1992. Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science* 3: 157-164.

Kiss L. 2012. Limits of nuclear ribosomal DNA internal transcribed spacer (ITS) sequences as species barcodes for Fungi. *Proceeding of National Academy of Science, USA* 109: E1811.

Koide RT, Fernandez C, Petprakob K. 2011. General principles in the community ecology of ectomycorrhizal fungi. *Annals of Forest Science* 68: 45-55.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, *et al.* 2005. Genome sequencing in microfabricated high-density picoliter reactors. *Nature* 437: 376-380.

Matthews JR. 1914. The white moss loch: a study in biotic succession. *New Phytologist* 13: 134-148.

Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson K, Kõljalg U. 2006. Taxonomic reliability of DNA sequences in public sequence databases : a fungal perspective. *PLOS One* 1: e59. doi:10.1371/journal.pone.0000059

Ohtonen R, Fritze H, Pennanen T, Jumpponen A, Trappe JM. 1999. Ecosystem properties and microbial community changes in primary succession on a glacier forefront. *Oecologia* 199: 239-246.

Pearsall WH. 1918. On the classification of aquatic plant communities. *Journal of Ecology* 6: 75-84.

Remias D, Lütz-Meindl U, Lütz C. 2005. Photosynthesis, pigments and ultrastructure of the alpine snow alga *Chlamydomonas nivalis*. *European Journal of Phycology*. 40: 259-268.

Reynolds HL, Packer A, Bever JD, Clay K. 2003. Grassroots ecology: plant-microbe-soil interactions as drivers of plant community structure and dynamics. *Ecology* 84: 2281-2291.

Schoch CL, Seifert KA, Huhdorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences, USA,* 109: 6241-6246.

Tansley AG. 1920. The classification of vegetation and the concept of development. *Journal of Ecology* 8: 118-149.

Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, Saar I, Bahram M, Bechem E, Chuyong G, Kõljalg U. 2010. 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist* 188: 271-301.

Walker LR, del Moral R. 2003. *Primary succession and ecosystem rehabilitation.* Cambridge University Press, New York, NY, USA.

Chapter 2 - Contrasting primary successional trajectories of fungi and bacteria in retreating glacial soils[1]

Abstract

Early community assembly of soil microbial communities is essential for pedogenesis and development of organic legacies. We examined fungal and bacterial succession along a well-established temperate glacier forefront chronosequence representing ~70 years of deglaciation to determine community assembly. As microbial communities may be heavily structured by establishing vegetation, we included non-vegetated soils as well as soils from underneath four plant species with differing mycorrhizal ecologies (*Abies lasiocarpa,* ectomycorrhizal; *Luetkea pectinata,* arbuscular mycorrhizal; *Phyllodoce empetriformis,* ericoid mycorrhizal; *Saxifraga ferruginea,* non-mycorrhizal). Our main objectives were to contrast fungal and bacterial successional dynamics and community assembly as well as to decouple the effects of plant establishment and time since deglaciation on microbial trajectories using high throughput sequencing. Our data indicate that distance from glacier terminus has large effects on biomass accumulation, community membership, and distribution for both fungi and bacteria. Surprisingly, presence of plants rather than their identity was more important in structuring bacterial communities along the chronosequence and played only a very minor role in structuring the fungal communities. Further, our analyses suggest that bacterial communities may converge during assembly supporting determinism whereas fungal communities show no such patterns. Although fungal communities provided little evidence of convergence in community structure, many taxa were nonrandomly distributed across the glacier foreland; similar taxon-level responses were observed in bacterial communities. Overall, our data highlight differing drivers for fungal and bacterial trajectories during early primary succession in recently deglaciated soils.

Keywords: *bacteria, fungi, glacier forefront, primary succession, 454 sequencing, rhizosphere*

Introduction

Primary succession has been studied in a number of North American temperate ecosystems and, consequently, much is known about plant establishment in volcanic (del Moral & Bliss 1993; del Moral *et al*. 1995) and glacier foreland systems (Chapin *et al*. 1994; Cázares *et al*. 2005; Jumpponen *et al*. 2012). In contrast, studies on microbial successional dynamics in these systems remain relatively few. With the growing concern of accelerating glacier recession (Barry 2006) and potential changes in global biogeochemical patterns (Cramer *et al*. 2001) it is becoming increasingly important to understand the fundamentals of microbial successional dynamics. Microbial primary succession has often been studied via rRNA community fingerprinting or cloning and sequencing (Sigler & Zeyer 2002; Jumpponen 2003), ectomycorrhizal root tip analysis (Trowbridge & Jumpponen 2004, Nara et al, 2003), and biochemical assays of soils and associated organisms (Tscherko *et al*. 2005; Ohtonen *et al*. 1999). With the advancement of high-throughput sequencing, these microbial communities in primary successional environments can now be queried in greater depth for both fungi (Blaalid *et al*. 2012) and bacteria (Schütte *et al*. 2010). Such studies have elucidated some general patterns of community dynamics over successional age. However, these patterns have not been explored simultaneously for both bacteria and fungi using high-throughput sequencing assays. Consequently, comparisons of fungal and bacterial successional dynamics are rare and preclude elucidation of universal patterns of microbial succession in soils. Concurrent examination of fungi and bacteria is mandatory to understand if these microbial guilds adhere to similar rules. Additionally, such studies permit evaluation of whether general conclusions or conceptual frameworks derived from plant community succession apply also to microbial communities.

Microbial succession does not occur in a vacuum. Rather, microbial communities interact with allochthonous substrates as well as early establishing autotrophs. Although essential to our greater appreciation of the microbial succession, the effects of substrate age and plant establishment remain insufficiently decoupled. This is especially true for plants whose mycorrhizal habits can be presumed to select particular fungal symbionts from the available propagule pools (Jumpponen & Egerton-Warburton 2005). These organismal interactions likely have precipitous effects across the establishing soil communities outside symbiotic partnerships.

Previous studies on fungal community dynamics in early primary successional environments have often focused solely on one mycorrhizal habit: ericoid (Tejesvi *et al.* 2010), ecto- (Helm *et al.* 1996; Trowbridge & Jumpponen 2004; Ashkannejhad & Horton 2006; Muhlmann & Peintner 2008), or arbuscular (Oehl *et al.* 2011, Sikes *et al.* 2012) mycorrhizas. As a result, a greater understanding of how vegetation may drive the microbial community succession can be gained through broader inclusion of mycorrhizal habits. Similarly, analyses of bacterial primary succession have focused on successional age (Schütte *et al.* 2010); or to a limited extent, vegetation influencing microbial communities (Knelman *et al.* 2012) and biogeochemical processes (Schmidt *et al.* 2008). Few previous studies investigated joint effects of successional age and vegetation establishment on microbial dynamics (Ohtonen *et al.* 1999; Zumsteg *et al.* 2012).

To our knowledge, this study represents one of the very first concurrent analyses of bacterial and fungal primary successional dynamics while also accounting for the effects of substrate age and establishment of plants that represent a broad selection of mycorrhizal habits utilizing deep interrogation of the soil microbial communities afforded by next-generation locus targeted sequencing. In the studies reported here, we evaluated the fungal and bacterial community composition using direct 454-pyrosequencing of ribosomal RNA (rRNA) gene targets. Our overall goal was to simultaneously analyze community dynamics of bacteria and fungi in a retreating glacier forefront system that is currently undergoing primary succession. We tested hypotheses on how time since deglaciation and established plants with differing mycorrhizal habits affect the fungal and bacterial communities during early primary succession. Specifically, we hypothesized that: 1) microbial communities increase in richness and diversity as substrate ages and becomes more heterogeneous as a result of plant establishment, 2) microbial communities associated with plants with different mycorrhizal habits are distinct, 3) fungal and bacterial biomass increase with successional age and shift from a bacteria-dominated early successional system to a fungus-dominated late successional system as plant establishment becomes of increasing importance, 4) communities exhibit successional trajectories as substrate ages indicating deterministic processes over successional time with specific plant associated mycorrhizal ecologies differentially influencing such deterministic patterns. These hypotheses are based partly on those derived for plant communities and summarized for comparable systems in Jumpponen *et al*. (2012).

Material and Methods

Study site

Lyman Glacier (ca. 1,900m a.s.l) is located within the Wenatchee National Forest in the North Cascade Mountain range in Washington State, U.S.A. (48°10'14"N, 120°53'44"W). Its receding forefront is characterized by a well-documented chronosequence of approximately a century with the terminal moraine located 1,100m north of the glacier terminus and suited well for a space for time substitution (Jumpponen *et al.* 1998). Plant (Cázares 1992; Jumpponen *et al.* 2012) and ectomycorrhizal fungus (Jumpponen *et al.* 1999; 2002; 2012) communities as well as root colonization of various plant hosts (Trowbridge & Jumpponen 2004, Cázares *et al.* 2005) have been previously characterized at this site.

Sampling

Topsoil (0-5cm) samples were collected from non-vegetated, bare soils as well as from the rhizospheres of four plant species differing in their mycorrhizal habits along the chronosequence at 150m intervals ranging from 0 to 750m from the glacier terminus (representing 0 to ~70 years since deglaciation). In all, we analyzed 72 samples (3 at 0m, 9 at 150m, 15 each at 300m, 450m, 600m and 750m from the glacier terminus. The target plant species occur commonly along the chronosequence and included ectomycorrhizal (EcM) *Abies lasiocarpa* (Hook.) Nutt., arbuscular mycorrhizal (AM) *Luetkea pectinata* Kuntze, ericoid mycorrhizal (ErM) *Phyllodoce empetriformis* D.Don, and non-mycorrhizal (NM) *Saxifraga ferruginea* Graham. Plant root colonization according to their mycorrhizal habits was earlier confirmed at this site (Cázares *et al.* 2005). From here on, these plants will be referred to only by their genus. Plant-associated soils were dug from the center of vegetation patch or directly underneath *Abies* stem; *Saxifraga* associated soils were sampled by excavating the whole plant and the associated soil collected. The three replicate samples for each plant or barren soil within each 150m interval were collected at least 5m apart. Non-vegetated soils were collected at least 1m from any established vegetation to ensure absence of roots in the substrate. Soils were sieved through a 5mm mesh to remove large rocks and root fragments, manually homogenized, and two sub-samples (~350 µL; 0.57g ± 0.15) collected directly into two DNA extraction bead tubes

(UltraClean Soil DNA Isolation kit; MoBio, Carlsbad, CA). The extraction tubes were placed on ice in collapsible coolers, and shipped to the laboratory at Kansas State University within 72 hours, where frozen at -20°C upon arrival. The two sub-samples were treated independently through DNA extraction and PCR amplification. A third soil sample was collected into a 2 ml collection tube and used to determine soil dry weight. All samples were collected on September 8th, 2009.

DNA extractions and analysis

DNA was extracted using UltraClean Soil DNA Isolation kits following the manufacturer's standard protocol. Extracted DNA was quantified using a ND1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE) and template DNA for each sample was aliquoted into 96-well plates at a working concentration of 2.5 ng μL^{-1}.

PCR-amplicons were generated for 454 sequencing using Lib-L unidirectional emPCR kit (Roche Applied Science, Indianapolis, Indiana, USA). Fungal amplicons of the Internal Transcribed Spacer 1 and 2 (ITS1, ITS2) were generated using an 8-bp DNA-bar-coded forward primer A-ITS1f (A-MIDs-CTTGGTCATTTAGAGGAAGTAA) and reverse primer of B-ITS4 (B-TCCTCCGCTTATTGATATGC) as described in Jumpponen *et al*. (2010). Bacterial amplicons were generated using a forward B-9F (B-GAGTTTGATCMTGGCTCAG) primer and 8bp DNA-bar-coded (see Table A.1 for complete list of MID sequences) reverse primer A-541R (A-MIDs-WTTACCGCGGCTGCTGG; Muyzer *et al.* 1993) to amplify V1 and V2 regions of the 16s rRNA (Baker *et al.* 2003). Each sub-sample was amplified in three independent 25μL PCR reactions for technical replication. PCR conditions for fungi were: 10 μM forward and reverse primers, 5ng template DNA, 200 μM of each dioxynucleotide, 25mM $MgCl_2$, 5 μL 5x Green GoTaq® Flexi Buffer (Promega, Madison, WI), 7.8 μL molecular biology grade water, and 1 U GoTaq® Hot Start Polymerase (Promega, Madison, WI). PCR cycle parameters consisted of a 94° initial denaturing for 3 min, four cycles of step down PCR at 94° denaturing for 1 min, 57-54° of annealing for 1 min, and 72° extension for 2 min followed by 29 cycles of 94° denaturing for 1 min, 54° annealing for 1 min, and 72° extension for 2 min followed by 8 min 72° final extension. PCR conditions for bacteria were: 10 μM forward and reverse primers, 5 ng template DNA, 12.5 μL AmpliTaq Gold® Master Mix (Applied Biosystems, Foster City,

CA), 5.5 µL molecular biology grade water with PCR cycle parameters with an initial denaturing step of 95° for 4 min, followed by 27 cycles of 95° denaturing for 30 sec, 54° annealing for 1 min, and 72° extension for 2 min, followed by a final extension step of 72° for 8 min 30 sec. All PCR reactions were performed on MasterCyclers (Eppendorf, Hamburg, Germany). Negative controls for DNA extractions and PCR reactions were included to ensure absence of contamination; no contamination was detected on PCR products visualized via an agarose gel electrophoresis.

Each separate PCR reaction (3 technical replicates, 2 sub-samples) was visualized on a 1.5% agarose (*w:v*) gel to ensure presence of PCR products. One fungal sample was omitted from further analysis because no product could be obtained. The remaining volume (20µl) of the 6 replicate PCR amplicons per experimental unit were pooled and cleaned with Agencourt® AmPure® cleanup kit using a SPRIplate 96-ring magnet (Beckman Coulter, Beverly, Massachusetts, USA) following the manufacturer's protocol with the exception that we used a 1:1 ratio of bead solution to reaction volume to discriminate against non-target small DNA fragments. Purified amplicons were quantified with NanoDrop and amplicons pooled at equal molarity (336 ng per fungal sample and 294 ng per bacterial sample) to equally represent each sample in subsequent sequencing reaction. Pooled samples were visualized on a low-melt agarose gel and the target-sized amplicon (~600-800 bp for fungi and ~500 bp for bacteria) excised for purification with UltraClean™ GelSpin™ DNA Extraction Kit (MoBio, Carlsbad, CA). The gel-excised fungal (1109.5 ng in 100 µL) and bacterial (1305 ng in 100 µL) amplicons were 454-pyrosequenced (GS FLX Titanium, Roche Applied Science, Indianapolis, Indiana, USA) at the Integrated Genomics Facility at Kansas State University (Manhattan, Kansas, USA). All sequences (.fastq for each experimental unit for fungi and bacteria) are deposited in NCBI Sequence Read Archive under the accession numbers (SRR943164-SRR943301).

The acquired sequence (.fasta) and quality (.qual) files were processed using the PyroTagger pipeline (Kunin & Hugenholtz 2010), where sequences that lacked an exact match to the MID-Primer sequence or were of poor quality (≥3% of bases with Q-values <27), and/or were of insufficient length (<350bp) were culled and sequences that passed quality control were truncated to 350bp. Operational taxonomic units (OTUs) were assigned at a 97% similarity threshold (UCLUST; Edgar 2010). Singleton OTUs were omitted from analysis as potential artifacts following Tedersoo and collaborators' (2010) recommendation. A representative

sequence for each fungal OTU was BLAST-n (nr/nt) (NCBI) queried for taxonomic affinity and matches with largest maximum identity were chosen after omission of uncultured/environmental accessions. Bacterial OTU taxon affinities were determined using Ribosomal Database Project's (RDP) Naïve Bayesian rRNA Classifier (Wang *et al.* 2007) using an 80% bootstrap confidence threshold. Putative mitochondrion and chloroplast sequences were removed from further analysis. The OTUs considered potentially chimeric in PyroTagger were further examined by manually checking a representative sequence of each OTU in BLAST-n (i.e. if the 5'- and 3'-ends of the queries aligned to different matches, the OTUs were culled); 19 fungal OTUs and 135 bacterial OTUs were determined to be putatively chimeric and omitted.

Ribosomal copy number estimates were determined as a proxy for bacterial and fungal biomass by quantitative PCR (qPCR) (Fierer *et al.* 2005). Briefly, plasmid standards were generated by amplifying *Escherichia coli* with primers EUB 338F (Lane, 1991) and 518R (Muyzer *et al.*, 1993) and *Agaricus bisporus* with primers ITS1F (Gardes and Bruns, 1993) and 5.8S (Vilgalys & Hester, 1990). Amplified products were cloned using One Shot® TOP10 Chemically Competent Cells and TOPO TA cloning kit (Invitrogen, Carlsbad, CA) and plasmids from positive transformants purified with UltraClean® plasmid prep kit (MoBio Laboratories, Carlsbad, CA). Triplicate qPCR reactions were performed in an iCycler iQ™ (Bio-Rad, Hercules, CA) RT-PCR detection system and the program iCycler (v. 3.1.7050) in 25μL reactions consisting of 12.5μL iQ™ SYBR® Green Supermix (Bio-Rad, Hercules, CA), 1.25μL (12.5 μM) forward primer, 1.25μL (12.5 μM) reverse primer, 5ng template DNA, and 8μL molecular biology grade H_2O. PCR cycle parameters were: initial denaturation at 95° C for 5 minutes, and 55 cycles of 95°C denaturation for 1 minute, 53°C annealing step for 30 seconds, and 72° extension step for 1 minute. Standard curves for bacteria and fungi were generated using a 10-fold dilution series (4×10^{-7} to 4×10^{-3} ng of the plasmid DNA per reaction). Copy numbers per gram dry soil were calculated (Pfaffl 2001) from the standard curves taking into account the length of target region and assuming an average molecular mass of 660 g mol^{-1} for double-stranded DNA.

Diversity indices

OTU richness (S) per sample was calculated by summing the number of OTUs. Complement of the Simpson's diversity index ($1-D=1-\Sigma\,p_i^2$) was calculated for each sample where p_i is the frequency that each OTU occurred in each sample. Shannon-Weaver's H' was also calculated but was omitted from analyses since the probabilistic nature of Simpson's index is more informative and the two indices were congruent. We also calculated evenness ($E_D =(1/D)/S$), where S is the OTU richness in each sample and D is Simpson's diversity index. To explore sequence depth across age of succession and vegetation type, rarefaction curves were generated using EstimateS (version 8.0.0; Colwell 2006).

Statistical analysis

To equalize sequencing effort per sample, we used an outlier analysis, in which a sample was omitted from analysis if the number of sequences was outside the range of $\pm$ 2SD (approximates a 95% CI assuming normal distribution across the data set) from mean sequence number. Subsequently, three bacterial samples were omitted from all subsequent analyses whereas all fungal samples were included. Simpson's diversity and evenness were arcsine transformed prior to analyses to meet the normality assumptions for ANOVA. To address the effects of distance from the glacier terminus and the effects of plant rhizospheres on OTU richness, Simpson's diversity and evenness, linear regression and ANOVA as well as subsequent pairwise comparisons (Tukey's HSD) were performed. In these analyses, we treated distance from the glacier terminus either as a continuous (linear regression) or categorical variable (ANOVA). All statistical analyses were conducted using JMP® (version 7.0.2; SAS Institute, Cary, NC). To examine community composition and clustering of the treatments, Nonmetric Multidimensional Scaling (NMS) multivariate analysis was performed using the Sørensen (Bray-Curtis) dissimilarity matrix in PC-ORD (version 4.1; McCune & Mefford 1999). The optimal number of dimensions (k) was selected based on Monte Carlo tests using empirical data compared to 100 randomized runs with a decrease in dimensionality from six dimensions to one. Based on the decline in stress, we chose k=3 dimensions. The ordination scores for each of the three axes were analyzed by linear regression and ANOVA across distance and vegetation as described above. Additionally, axes score standard deviations of vegetation treatments by

distance from glacier terminus were tested (ANOVA) to determine convergence or divergence of the communities (del Moral *et al.* 1995, Jumpponen *et al.* 2012).

The abundance of each OTU was tested for random distribution across the experimental design matrix. If species distribution in a successional environment is dictated purely by random allochthonous propagule input, then it stands to reason that species distribution would follow a discrete probability function where rarity is expected at any given location, in other words following a distribution such as Poisson. Consequently, for each OTU we tested if its distribution differed from an expected Poisson distribution using a χ^2 test (df=4) on a contingency table of observed sequence frequencies per sample and site. Significance was determined after Bonferroni correction (critical value α =0.05/n) where n is the number of taxa tested (n = 4,114 for bacteria and n = 310 for fungi). We used frequency categories for Poisson analysis following an increasing logarithmic scale (0, 1-10, 11-100, 101-1000, 1000+).

Ecosystem processes and successional dynamics are likely driven largely by community members that occur most frequently. For this reason, we identified core taxa (defined as present in $\geq$ 50% of samples; see Unterseher *et al.* 2011) for the glacier forefront system as a whole as well as for each of the different treatment categories. We tested these core taxa for changes in occurrence (square root + 0.5 transformed counts) across distance and rhizosphere environments using a combination of linear regression and ANOVA as described above.

Results

Sequence data

To characterize microbial communities in the Lyman Glacier forefront, we 454-sequenced 180,421 fungal and 194,513 bacterial amplicons. After excluding short (350bp threshold) and poor quality reads, 39,509 fungal and 97,716 bacterial sequences remained resulting in an average sequencing depth of 539±387 (mean±SD) per sample for fungi and 1454±391 per sample for bacteria. The number of fungal sequences was invariable across distance and vegetation ($F_{5,65}$=0.26 and 0.42, *P*=0.611 and 0.794 respectively) and the number of bacterial sequences was invariable across distance ($F_{5,59}$=1.33, *P*=0.254) but bare soil samples show lower sequence number compared to other vegetation treatments ($F_{5,59}$=3.15, *P*=0.020). After clustering at 97% sequence similarity, there were a total of 310 fungal and 4110 bacterial

non-singleton OTUs and 247 and 2972 singleton OTUs, respectively. Note that all singletons were omitted from the further analyses.

Estimation of soil microbial biomass by qPCR

Ribosomal copy numbers as a proxy for bacterial biomass estimated by qPCR (copy numbers per g soil dry weight) did not differ among the sampled rhizosphere environments. However, there was a near significant trend of increasing bacterial copy numbers with distance from the glacier terminus ($t = 1.96$, $P = 0.055$; Fig. 2.1). In contrast to bacteria, fungal copy numbers unequivocally increased with distance from the glacier terminus ($t = 2.33$, $P = 0.023$; Fig. 1), although the rates differed among the sampled vegetation; fungal biomass increased more slowly in *Abies* associated soils than in the non-vegetated soils ($t = -2.12$, $P = 0.038$) and increased at a greater rate in *Luetkea* associated soils than in the non-vegetated soil ($t = 3.98$, $P < 0.001$). Fungi to Bacteria (F:B) biomass ratio was stable along this primary successional chronosequence suggesting similar rates of fungal and bacterial biomass accumulation (Fig. 2.1).

Fungal communities

Soil fungal communities were strongly dominated by diverse Ascomycota with 192 OTUs (61.9%) and 27,962 sequences (70.8%) followed by Basidiomycota with 84 OTUs (27.1%) and 3,111 sequences (7.9%) (Fig. 2.2a, 2.2b). We also detected some Chytridiomycota with 10 OTUs (3.2%) and 66 sequences (0.2%), and Glomeromycota with 6 OTUs (1.9%) and 39 sequences (0.1%). We encountered an additional 18 OTUs (5.8%) comprised of 8,331 sequences (21.1%) representing other basal fungi. Of these basal fungi, 13 OTUs showed best BLAST match to sub-phylum Mucoromycotina and 5 OTUs to sub-phylum Entomophthoromycotina. Of the 44 orders of Fungi, the most abundant were Hypocreales (9,809 sequences, 24.8%) followed by Mortierellales (6,740 sequences, 17.1%) and Helotiales (5,620 sequences, 14.2%). On a family level, we observed 75 families, dominated by Mortierellaceae (6,740 sequences, 17.1%), Cordycipitaceae (3,750 sequences, 9.5%) and Clavicipitaceae (3,553 sequences, 9.0%). Lastly, of the 154 genera classified, *Mortierella* (6,740 sequences, 17.1%), *Trichocladium* (3,042 sequences, 7.70%), and *Articulospora* (2,944 sequences, 7.45%) were the most abundant.

The three most abundant OTUs were BLAST-assigned to *Mortierella alpina* (5,858 sequences, Accession AJ878532.1, 93% query coverage, 99% max identity, family Mortierellaceae), *Cudoniella clavus* (3,042 sequences, Accession DQ491502.1, 100% query coverage, 96% max identity, family Helotiaceae), and *Cordyceps bassiana* (2,799 sequences, Accession EU673367.1, 100% query coverage, 100% max identity, family Cordycipitaceae). It is important to note that these described taxonomic abundances represent only sequences that we were able to classify at high levels of taxonomic resolution through BLAST-n. There were many sequences that could not be classified (for example, 100 OTUs were not classified at the family level) and may thus represent novel OTUs.

Bacterial communities

Bacterial OTUs were distributed taxonomically as follows (Fig. 2.2c, 2.2d): the most abundant phyla were Proteobacteria with 1486 OTUs (36.2%) and 31,055 sequences (31.8%), followed by Actinobacteria with 612 OTUs (14.9%) and 18,995 sequences (19.4%) and Acidobacteria with 401 OTUs (9.8%) and 12,544 sequences (12.8%). Of particular note are the 68 OTUs representing the photosynthetic Cyanobacteria (1.7%) with 2,712 sequences (2.8%). Of the 159 classified orders, Actinomycetales (11,415 sequences, 11.7%), Rhizobiales (7,885 sequences, 8.1%) and Sphaerobacterales (6,879 sequences, 7.0%) were the most abundant. Of the 198 OTUs classified to a family, those assigned to Sphaerobacteraceae (6,879 sequences, 7.0%), Acetobacteraceae (5,107, 5.2%), and Ktedonobacteraceae (4,247, 4.4%) were the most abundant. Finally, of the 450 observed and taxonomically assigned genera, the most abundant were *Sphaerobacter* (6,879 sequences, 7.0%), Gp1 (undefined Acidobacteria genus, 4419 sequences, 4.5%), and *Ktedonobacter* (4,242 sequences, 4.3%). Many of the observed OTUs may represent novel taxa, as 1,757 OTUs could not be assigned to a family and 1,539 to a genus.

Microbial richness and diversity

Our rarefaction analyses suggest that although a large proportion of the fungal and bacterial richness was captured in our sampling, a greater sampling effort is needed for complete saturation (Fig. 2.3). Richness and diversity estimators (S, 1-D, E_D) showed distinct patterns that differed between fungal and bacterial communities (Fig. 2.4). Fungal OTU richness (2.4a),

diversity (2.4b) and evenness (2.4c) estimators did not respond to distance from glacier terminus, plant association, or their interaction (Table 2.1). In contrast, bacterial OTU richness (2.4d) increased over distance from glacier terminus ($F_{1,63}$=7.76, P=0.007) and differed between plant associated environments ($F_{4,59}$=5.99, P<0.001). There was no significant interaction between the distance from the glacier terminus and the plant associated environments, suggesting that these differences were additive and stable across the sampled glacier forefront. Although the complement of Simpson's diversity estimates (4e) for bacteria did not differ among the environments, bacterial community evenness (4f) decreased over distance from the glacier terminus ($F_{1,63}$=4.75, P=0.034). However, it did not differ among the plant-associated environments (Table 2.1).

Analyses of OTU distribution

Based on Bonferroni-corrected analyses testing whether OTUs are randomly distributed following a Poisson distribution across Lyman glacier forefront, a total of 240 bacterial OTUs (9.5 % of all bacterial OTUs) deviated from the presumed Poisson distribution of randomness as expected if allochthonous propagule sources dominated the community assembly across the forefront. It is notable that nearly twice as great a proportion of fungal OTUs (59 or 18.79%) deviated from the Poisson distribution of randomness. These data strongly suggest that fungi - and to a lesser degree also bacteria - are non-randomly distributed across this glacier forefront. Such non-random distributions suggest predictable community trajectories along the glacier forefront (Table A.2 for Poisson test statistics) and/or effects of plant establishment.

Analyses of core taxa

Fungi

Core OTUs (present in ≥50% of samples) across the entire landscape or associated with different plants had affinities within subphyla Mucoromycotina and Pezizomycotina, two common and ubiquitous subphyla (see Table 2.2 for complete list of core taxa responding to distance and associated test statistics). By our definition, there were six core fungal OTUs across the entire dataset. Of these, five changed in frequency over the distance from the glacier terminus (linear regression with Bonferroni correction) suggesting their correlation with

changing community and ecosystem structure. Two of these core OTUs also responded to vegetation: OTU9 [*Cudoniella clavus* - saprobic] occurred less frequently in *Saxifraga* than in *Abies*; OTU104 [*Phialocephala sphaeroides* – a putative dark septate endophyte] occurred less frequently in *Saxifraga* and bare soil than in *Abies*). Core OTUs associated with *Abies* consisted of ten OTUs, one of which one increased significantly in frequency with distance (OTU142 [*Serea difformis* – a putative saprobe]). *Luetkea* core consisted of twelve OTUs, six of which increased in abundance over distance from the glacier terminus (OTUs 3, 5, 13, 18, 24, 109 [*Pochonia bulbillosa* – insect parasite, *Lecythophora sp.* – saprobe, *Satchmopsis brasiliensis* - saprobe, *Articulospora tetracladia* – saprobe, *Cryptococcus skinneri* – wood pathogen, *Cladophialophora minutissima* – saprobe]). *Phyllodoce* core consisted of eleven OTUs, none of which changed significantly with distance. *Saxifraga* core consisted of eight OTUs, one of which decreased in frequency across the forefront (OTU81 [*Penicillium citreonigrum* – saprobe]). Non-vegetated soils had six core OTUs, two of which increased with distance from the glacier terminus (OTUs 1, 27 [*Mortierella alpina* – saprobe, *Mortierella elongata* – saprobe]).

Core taxon analyses focusing on the age of the substrate underlined the specificity in core membership with respect to distance from the glacier terminus. In all, 70 OTUs were considered as core OTUs in at least one of the distances – interestingly, 49 OTUs were core to only one distance. Of note is the abundance of an OTU (OTU 1) with affinity to *Mortierella alpina* (AJ878532.1). This OTU is a core taxon in all but samples collected adjacent to the glacier terminus. Remarkably, this OTU represents the most abundant sequence in all samples across the glacier forefront except those closest to the glacier where it was completely absent.

Bacteria

We identified 137 bacterial core OTUs that occurred in ≥50% of all samples. A total of twenty core OTUs (14.6%) changed in frequency (linear regression with Bonferroni correction): 18 increased and 2 decreased in frequency with distance from glacier terminus (Table 2.3). Overall, when the core OTUs were analyzed for response to distance from the glacier terminus, the number of core bacterial OTUs increased ($F_{1,4}$=18.12, *P*=0.013).

Core OTUs from *Abies* samples consisted of 182 OTUs but none of them changed in frequency with distance from the glacier terminus (Table 2.3). Similarly, *Phyllodoce* samples included 244 and *Saxifraga* samples 203 core OTUs, none of which changed in frequency over

the chronosequence. *Luetkea* samples included 140 core OTUs with only two (1.4%) changing with distance from glacier, both of which increased. Non-vegetated soils consisted of 107 core OTUs, five of which (4.7%) increased in frequency over distance from glacier terminus, while none decreased. Of the 137 bacterial core OTUs across all treatments, 14 show a response to vegetation indicating some specific plant-bacterial associations. Of these 14 responding OTUs, 9 differed in frequency between plant-associated and bare soils (Table 3). OTUs 1, 44, 53, 68, 119, and 203 were less frequent in bare than in plant-associated soils, whereas their frequencies did not differ among the plant-associated soils (OTUs 1, 53, 68 were placed with strong bootstrap support to the putatively nitrogen fixing order Rhizobiales; OTUs 44, 119, 203 placed to order Xanthomonadales, and classes Gp6 and Gp3 respectively with unknown function). In contrast, OTUs 60 (100% bootstrap support to likely photosynthetic Cyanobacteria), 75 (Gp1 with unknown function), and 486 (100% bootstrap support to phylum Chloroflexi [45% bootstrap support to genus *Sphaerobacter*], a putative aerobic thermophile adapted to the harsh UV-exposure characteristic of the bare soils) were more frequent in bare soil than plant-associated samples. These data suggest that plant establishment primarily controls core bacterial communities and their assembly. We emphasize the contrast between fungal and bacterial communities: while many bacteria responded to plant presence, the core fungal OTUs rarely responded to the presence or the taxon identity of the sampled plants, even though the plant species were specifically selected to represent differing root-symbioses with fungi.

Analyses of community composition

Fungi

While only few fungal OTUs seemed responsive to changes in environments along the successional chronosequence, analyses of NMS axes scores indicated strong and clear successional trajectories for fungal communities (Fig. 2.5a). Axis 1 (representing 23.2% of the variability) scores decreased drastically with increasing distance from glacier terminus ($t = -4.81$, $P < 0.001$). Treating distance from the glacier as a categorical variable, Axis 1 scores differed across distance (ANOVA: $F_{1,69} = 5.59$, $P < 0.001$). This was mainly attributable to Axis 1 scores in young substrates (at 0, 150, and 300m distances from the glacier terminus) that were distinct from those in more developed substrates (at 450, 600, and 750m from the glacier terminus; Tukey's HSD pairwise comparison at $\alpha = 0.05$). Axis 2 scores (representing 24.5% of

the variability) also tended to increase ($t = 2.41$, $P = 0.019$) across distance from the glacier terminus. However, *Abies* communities were distinct from others and their Axis 2 scores decreased with increasing distance from the glacier terminus (t = –2.22, P =0.012). These results suggest fungal community trajectories that are potentially distinct among establishing vegetation types during ecosystem development. Axis 3 (15.4% variation) showed neither a distinct trend across the distance along the glacier forefront nor any distinctions on communities across the distance from the glacier terminus as inferred from ANOVA.

In addition to the patterns with distance from glacier terminus, Axes 1 and 2 scores also distinguished fungal communities in the establishing vegetation (ANOVA: $F_{9,61} = 5.59$, P <0.001; $F_{9,61} = 3.20$, $P = 0.019$ respectively). Communities in *Luetkea* samples were distinct from bare and *Saxifraga* soils along Axis 1 and from *Abies* along Axis 2. Other soils remained indistinguishable based on our NMS analyses. These results suggest that this AM plant may strongly select fungal communities, whereas plants with other mycorrhizal habits do so to a lesser degree. Analyses of standard deviations of axis scores showed no change across distance ($F_{1,23}<1.22$, $P>0.2$ for all three axes), or vegetation ($F_{4,19}<1.4$ $P>0.3$ for all three axes) providing no support for fungal community convergence or divergence as a result of plant establishment or with distance from glacier terminus.

Bacteria

Soil bacterial communities showed strong trends with distance from the glacier terminus along all three NMS axes. Axis 1 (representing 19.1% of variation; $t = 3.67$, $P < 0.001$) and 3 (36.3% of variation; $t = 5.53$, $P < 0.001$) scores increased with distance from glacier terminus (Fig. 2.5b), whereas axis 2 (representing 31.6% of variation) decreased with distance (t=-3.58, P < 0.001). Axes 2 and 3 also clearly distinguished the communities between the sampled plant species ($F_{9,55} = 3.46$, $P = 0.014$; $F_{9.55} = 12.60$, $P < 0.001$, respectively), whereas Axis 1 scores did not differ among them. Axis 2 distinguished bacterial communities between *Phyllodoce* and non-vegetated soils, whereas axis 3 separated non-vegetated soils from all plant-associated soils (Fig. 2.5c) suggesting community filtering by vegetation regardless of the plants' mycorrhizal habits. Analyses of standard deviations for Axes 1 and 2 provided no support for community convergence along the chronosequence. In contrast, standard deviations of Axis 3 (36.3% of variation) decreased over distance from the glacier terminus (t=-2.49, P=0.021) suggesting

bacterial community convergence along Axis 3 (Fig. 2.6) but this pattern vanished after accounting for different plant environments.

Discussion

We sequenced bacterial and fungal rRNA gene amplicons from a primary successional glacier forefront soils to analyze soil microbial community assembly along a deglaciated chronosequence. To our knowledge, this is the first study to address community assembly of bacteria and fungi simultaneously in a primary successional system using high-throughput sequencing tools allowing for an unprecedented analysis of microbial community dynamics. Our analyses revealed three important and novel points about this early primary successional system that has continued to deglaciate for a century. *First*, the concurrent analyses of fungal and bacterial communities emphasize establishment and building organic legacies as a result of coinciding allochthonous resource arrival (see Hodkinson *et al*. 2001, 2002) as well as autochthonous microbial processes (see Kaštovská *et al.* 2007). *Second*, these analyses permitted comparisons between bacterial and fungal communities in early succession and identified differences either on the temporal scale or the trajectories between the two. *Third*, surprisingly and despite our choice of plants with differing mycorrhizal habits, we observed that bacterial communities are more strongly structured by established vegetation than fungal communities.

Our analyses highlighted diverse microbial communities in a successional system void of vegetation, corroborating conclusions of previous studies (Freeman *et at.* 2009a, 2009b; Strauss *et al.* 2012; Zumsteg *et al.* 2012) and emphasizing the importance of microbial ecosystem functions in early primary succession to establish a pool of organic carbon and nitrogen in soil. Hodkinson and co-workers (2001, 2002) have argued that heterotrophs (particularly arthropods and arthropod-vectored inocula), not autotrophs, are the principal early drivers of primary successional dynamics. Both autochthonous (on site carbon and nitrogen accumulation) and allochthonous (arrival of organisms and debris from out of site sources) organic inputs likely contribute to the early establishment of ecosystem services. Three observations in our data support the importance of allochthonous inputs argued by Hodkinson et al. (2001, 2002). *First*, our data included five common fungal OTUs representing the entomopathogenic taxa in the subphylum Entomophthoromycotina. *Second*, the third most common fungal OTU (9.5% of all

fungal sequences) was assigned to family Cordycipitaceae (*Cordyceps*) that includes a number of known entomopathogens. *Third*, OTU3 (with affinity to *Pochonia bulbillosa*, a known insect pathogen) was a core OTU across all samples and increased in frequency with successional age, suggesting a likely increase in insects colonized by this pathogen over the glacier chronosequence. In sum, our findings, combined with those of others, suggest that both autotrophic and heterotrophic microbial constituents play important roles in the accumulation of organic legacies in early successional systems.

In contrast commonly observed increase in plant community complexity and deterministic patterns of plant establishment during primary succession (del Moral *et al.* 1995; Jumpponen *et al.* 2012), fungal richness and diversity estimators were static across the Lyman glacier chronosequence. However, bacterial and plant communities seem to follow similar successional patterns characterized by increasing richness and declining evenness over successional time. These results suggest increasing heterogeneity in bacterial communities with increasing distance from glacier terminus. It is unclear whether plant establishment increases environmental heterogeneity leading to concomitant increases in bacterial community heterogeneity or bacterial communities facilitate plant establishment by expediting pedogenesis. Our findings on the community dynamics are unlikely universal: for example, Blaalid *et al.* (2012) observed an increasing fungal OTU richness over successional time on a glacier forefront in Norway. Reasons for this incongruence are uncertain. However, it is important to bear in mind that our glacier forefront chronosequence represents less than a century of deglaciation, whereas that studied by Blaalid *et al*. (2012) spanned across centuries. While it is attractive to argue that fungal early successional trajectories are unique, the successional dynamics may be strongly influenced by regional and local factors. Furthermore, while fungal OTU richness and diversity estimates at our site were rather stable, fungal community structure and distribution were dynamic over the chronosequence – an observation congruent with other glacier systems (Blaalid *et al.* 2012).

Our analyses clearly show that distributions of many OTU are non-random. However, our analyses fell short in determining the drivers that structure these communities. The non-random distribution challenges the hypothesized random draw of microbial communities from an allochthonous propagule rain (Jumpponen 2003), possibly highlighting the difficulties of finding universalities, as multiple different processes may govern the OTU establishment. Our data

suggest that some OTUs have distinct preferences for established vegetation or substrate age and that autochthonous propagation may be equally important in explaining the microbial establishment in successional landscapes. Moreover, compared to fungi, a smaller proportion of bacterial OTUs were nonrandom, suggesting that fungi likely have specific habitat requirements in early successional systems. However, our experimental design did not account for the primary drivers for this as indicated by lack of vegetation effect on the fungal community composition and the observed bacterial association with established vegetation. It is of particular note that many non-randomly distributed OTUs were also core taxa (found in at least 50% of treatments) which exhibited shifts in frequency across the glacier forefront suggesting the potential importance of these core taxa on influencing patterns of spatial heterogeneity and community dynamics. Interestingly, of these nonrandomly distributed core fungal OTUs, all but one are putative saprobes and one was a potential insect pathogen. In addition, many core OTUs were unique to one distance category, which we interpret to suggest rapid turnover and dynamics in microbial communities during early succession. The number of core bacterial OTUs increased with successional age, which is suggestive of decreasing importance of stochastic allochthonous deposition in bacterial community establishment in this system. The core fungi and bacteria likely influence the community structure throughout the landscape and strongly contribute to overall differences in community dynamics over successional age.

Our NMS analyses clearly differentiated fungal and bacterial communities along the chronosequence. Additionally, bacteria were strongly differentiated among the plant-associated environments. Similarly to Trowbridge and Jumpponen (2004), we interpret these results to indicate organismal niche preferences. Trowbridge and Jumpponen (2004) argued that it is indeed the successional age as well as the pedogenesis in the early successional environment that defines the fungal communities and selects the members that may successfully establish and survive. We extend this argument to account for bacterial successional trajectories. Unfortunately, our study - similarly to Trowbridge and Jumpponen (2004) - fails to provide adequately robust environmental data matrix to permit elucidation of those soil parameters that most strongly correlate with the observed communities. However, it is of note that previous studies in this system show that the soil organic matter, as well as carbon and nitrogen contents increase rather predictably as a result of increasing successional age and plant establishment (Cazarés 1992; Jumpponen *et al*. 1998). As a result, the observed shifts in soil fungal and

bacterial communities are likely correlated with accumulation of these organic legacies and the resultant shifts in soil chemical and physical properties.

One goal of our study was to evaluate microbial community divergence in a primary successional system similar to the dynamics observed in plant communities (de Moral 2009). We predicted that sampling plant hosts with distinct mycorrhizal habits would filter specific communities from the available propagule pool (Jumpponen & Egerton-Warburton 2005). However, despite the evident community turnover, our analyses did not support fungal community convergence. In contrast to the fungal communities, the presence of vegetation strongly influenced bacterial communities (with no observable difference in bacterial communities among different vegetation treatments) with bare soil bacterial communities being discrete from vegetation-associated communities. Taken together, this indicates a greater importance of stochastic processes in non-vegetated soils, *i.e.,* allochthonous propagule rain, whereas plant establishment partly drives deterministic processes in bacterial communities. Furthermore, in contrast to fungi, bacteria show evidence of community convergence as seen by the decreasing NMS axis 3 score standard deviations suggesting that communities may become similar to each other across the glacier forefront. We argue that this is evidence that early successional bacterial communities that are dictated by stochastic colonization with increasing determinism as a result of ecosystem development and/or pedogenesis.

Consistently with earlier studies, both fungal and bacterial biomasses (as proxied by qPCR assays) increased with distance along the Lyman glacier chronosequence (Ohtonen *et al.* 1999). Fungal and bacterial biomasses appeared to accumulate at similar rates as indicated by the stable F:B ratio. This is in contrast with previous work at this site reporting decreasing F:B ratio (Ohtonen *et al.*, 1999). Those results were interpreted as a shift from bacteria to fungus dominated system coinciding with the vegetation establishment. The authors argued that as vegetation structure increases in complexity, fungi can more efficiently utilize available carbon sources compared to bacteria leading to the observed shift in biomass ratio. The reasons for this disparity remain unclear but may include use of different biomass measurement tools as estimating biomass with qPCR may be influenced by many factors including variable copy numbers (Strickland and Rousk 2010), or the more than decade long time lag between the two studies.

Conclusions

Our study demonstrates that fungal and bacterial communities are dynamic along a primary successional chronosequence. Surprisingly, vegetation had a stronger effect on bacterial than fungal community dynamics even when the plant species were selected based on their mycorrhizal habit. Mycorrhizal habit should lead to a deterministic fungal community assembly, but our findings suggest other community controlling mechanisms in early succession. More importantly, our data clearly indicated that microbial community dynamics are strongly influenced by distance along the Lyman glacier chronosequence and that these communities exhibit rapid turnover. While the communities overall may not differ among the sampled plants, the plant-associated microbial (fungal and bacterial) communities are enriched for certain community members. Furthermore, the core taxon frequency shifts with successional age suggest niche distinction and increasing importance of autochthonous inputs over the chronosequence. The bacterial and fungal communities differed in responses to establishing vegetation and exhibited dramatic differences in successional trajectories across the chronosequence. Taken together, our data highlight distinct successional dynamics between fungi and bacteria, but also provide insight into patterns that may be universal. Our findings warrant further investigation to pinpoint the primary drivers for observed similarities and differences.

Acknowledgements

References

Ashkannejhad S, Horton TR (2006) Ectomycorrhizal ecology under primary succession on coastal sand dunes: interactions involving *Pinus contorta*, suilloid fungi and deer. *New Phytologist* **169**, 345-354.

Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* **55**, 541-555.

Barry RG (2006) The status of research on glaciers and global glacier recession: a review. *Progress in Physical Geography* **30**, 285-306.

Blaalid R, Carlsen T, Kumar S, Halvorsen R, Ugland KI, Fontana G, Kauserud H (2012) Changes in the root-associated fungal communities along a primary successional gradient analysed by 454 pyrosequencing. *Molecular Ecology* **21**, 1897-1908.

Cazarés E (1992) Mycorrhizal fungi and their relationship to plant succession in subalpine habitats. Ph.D. Dissertation, Oregon State University, Corvallis, Oregon. 126 pp.

Cazarés E, Trappe JM, Jumpponen A (2005) Mycorrhiza-plant colonization patterns on a subalpine glacier forefront as a model system of primary succession. *Mycorrhiza* **15**, 405-416.

Chapin FS, Walker LR, Fastie CL, Sharman LC (1994) Mechanisms of primary succession following deglaciation at Glacier Bay, Alaska. *Ecological Monographs* **64**, 149-175.

Colwell RK (2006) EstimateS: Statistical estimation of species richness and shared species from samples. Version 8.0.0. User's Guide and application published at: http://viceroy.eeb.uconn.edu/estimates/

Cramer W, Bondeau A, Woodward FI, Prentice IC, Betts RA, Brovkin V, Cox PM, Fisher V, Foley JA, Friend AD, Kucharik C, Lomas MR, Ramankutty N, Sitch S, Smith B, White A, Young-Molling C (2001) Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models. *Global Change Biology* **7:** 357-373.

delMoral R, Bliss LC (1993) Mechanisms of primary succession – insights resulting from the eruption of Mount St. Helens. *Advances in Ecological Research, Vol 24* **24**, 1-66.

delMoral R, Titus JH, Cook AM (1995) Early primary succession on Mount St. Helens, Washington, USA. *Journal of Vegetation Science* **6**, 107-120.

delMoral R (2009) Increasing deterministic control of primary succession on Mount St. Helens, Washington. *Journal of Vegetation Science* **20**, 1145-1154.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26,** 2460-2461.

Fierer N, Jackson JA, Vilgalys R, Jackson RB (2005) Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Applied and Environmental Microbiology* **71**, 4117-4120.

Freeman KR, Martin AP, Karki D, Lynch RC, Mitter MS, Meyer AF, Longcore JE, Simmons DR, Schmidt SK (2009) Evidence that chytrids dominate fungal communities in high-elevation soils, *Proceedings of the National Academy of Sciences of the United States of America* **106,** 18315-18320.

Freeman KR, Pescador MY, Reed SC, Costello EK, Robeson MS, Schmidt SK (2009) Soil CO_2 flux and photoautotrophic community composition in high-elevation, 'barren' soil. *Environmental Microbiology* **11,** 674-686.

Gardes M, Bruns TD (1993) ITS primers with enhanced specificity for Basidiomycetes –
Application to the identification of mycorrhizae and rusts. *Molecular Ecology* **2**, 113-118.

Helm DJ, Allen EB, Trappe JM (1996) Mycorrhizal chronosequence near Exit Glacier, Alaska.
Canadian Journal of Botany **74,** 1496-1506.

Hodkinson ID, Coulson SJ, Harrison J, Webb NR (2001) What a wonderful web they weave:
spiders, nutrient capture and early ecosystem development in the high Arctic – some
counter-intuitive ideas on community assembly. *Oikos* **95**, 349-352.

Hodkinson ID, Webb NR, Coulson SJ (2002) Primary community assembly on land – the
missing stages: why are the heterotrophic organisms always there first? *Journal of
Ecology* **90**, 569-577.

Jumpponen A (2003) Soil fungal community assembly in a primary successional glacier
forefront ecosystem as inferred from rDNA sequence analyses. *New Phytologist* **158**,
569-578.

Jumpponen A, Egerton-Warburton LM (2005) Mycorrhizal fungi in successional environments –
a community assembly model incorporating host plant, environmental and biotic filters.
In: The Fungal Community, ed. Dighton J, White JF, Oudemans P. CRC Press, New
York 139-180.

Jumpponen A, Mattson K, Trappe JM, Ohtonen R (1998) Effects of established willows on
primary succession on Lyman Glacier forefront, North Cascade Range, Washington,
USA: Evidence for simultaneous canopy inhibition and soil facilitation. *Arctic and Alpine
Research* **30**, 31-39.

Jumpponen A, Trappe JM, Cazares E (1999) Ectomycorrhizal fungi in Lyman Lake Basin: a
comparison between primary and secondary successional sites. *Mycologia* **91**, 575-582.

Jumpponen A, Trappe JM, Cazares E (2002) Occurrence of ectomycorrhizal fungi on the
forefront of retreating Lyman Glacier (Washington, USA) in relation to time since
deglaciation. *Mycorrhiza* **12**, 43-49.

Kaštovská K, Stibal M, Šabacká M, Černá B, Šantrůčková H, Elster J (2007) Microbial
community structure and ecology of subglacial sediments in two polythermal Svalbard
glaciers characterized by epifluorescence microscopy and PLFA. *Polar Biology* **30,** 227-
287.

Knelman, JE, Legg, TM, O'Neill, SP, Washenberger, CL, Gonzalez, A, Cleveland, CC,
Nemergut DR (2012) Bacterial community structure and function change in association
with colonizer plants during early primary succession in a glacier forefield. *Soil Biology
& Biochemistry* **46**, 172-180.

Kunin V, Hugenholtz P (2010) PyroTagger: A fast, accurate pipeline for analysis of rRNA
amplicon pyrosequence data. *The Open Journal* 1. Feb 2010.

Lane D (1991) 16s/23s rRNA sequencing, p. 115-175. *In* E. Stackebrant and M Goodfellow
(ed.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, West Sussex,
United Kingdom.

McCune B, Mefford JM (1999) PC-ORD: Multivariate analysis of ecological data. Version 4.
MjM Software, Gleneden Beach, Oregon, USA.

Muyzer G, de Wall EC, Uitterlinden AG (1993) Profiling of complex microbial populations by
denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified
genes coding for 16S rRNA. *Applied Environmental Microbiology* **59:** 695-700.

Nara K, Nakaya H, Wu BY, Zhou ZH, Hogetsu T (2003) Underground primary succession of Ectomycorrhizal fungi in a volcanic desert on Mount Fuji. *New Phytologist* **159:** 734-756.

Oehl F, Schneider D, Sieverding E, Burga CA (2011) Succession of arbuscular mycorrhizal communities in the foreland of the retreating Morteratsch glacier in the Central Alps. *Pedobiologia* **54**, 321-331.

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

Schmidt SK, Reed SC, Nemergut DR, Grandy AS, Cleveland CC, Weintraub MN, Hill AW, Costello EK, Meyer AF, Neff JC, Martin AM (2008) The earliest stages of ecosystem succession in high-elevation (5000 meters above sea level), recently deglaciated soils. *Proceedings of the Royal Society B-Biological Sciences*. **275:** 2793-2802.

Schutte UME, Abdo Z, Foster J*, et al.* (2010) Bacterial diversity in a glacier foreland of the high Arctic. *Molecular Ecology* **19**, 54-66.

Sigler WV, Zeyer J (2002) Microbial diversity and activity along the forefields of two receding glaciers. *Microbial Ecology* **43**, 397-407.

Sikes BA, Maherali H, Klironomos JN (2012) Arbuscular mycorrhizal fungal communities change among three stages of primary sand dune succession but do not alter plant growth. *Oikos* **121**, 1791-1800.

Strauss SL, Garcia-Pichel F, Day TA (2012) Soil microbial carbon and nitrogen transformations at a glacial foreland on Anvers Island, Antarctic Peninsula. *Polar Biology* **35**, 1459-1471.

Strickland MS, Rousk J (2012) Considering fungal:bacterial dominance in soils – methods, controls, and ecosystems implications. *Soil Biology and Biochemistry* **42,** 1385-1395.

Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, Saar I, Bahram M, Bechem E, Chuyong G, Koljalg U (2010) 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist* **188**, 291-301.

Tejesvi MV, Ruotsalainen AL, Markkola AM, Pirttila AM (2010) Root endophytes along a primary succession gradient in northern Finland. *Fungal Diversity* **41**, 125-134.

Trowbridge J, Jumpponen A (2004) Fungal colonization of shrub willow roots at the forefront of a receding glacier. *Mycorrhiza* **14**, 283-293.

Tscherko D, Hammesfahr U, Zeltner G, Kandeler E, Bocker R (2005) Plant succession and rhizosphere microbial communities in a recently deglaciated alpine terrain. *Basic and Applied Ecology* **6**, 367-383.

Unterseher M, Jumpponen A, Öpik M, Tedersoo L, Moora M, Dormann CF, Schnittler M (2011) Species abundance distributions and richness estimations of fungal metagenomics – lessons learned from community ecology. *Molecular Ecology* **20,** 275-285.

Vilgalys R, Hester M (1990) Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. *Journal of Bacteriology* **172:** 4238-4246.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**, 5261-5267.

Zumsteg A, Luster J, Göransson H, Smittenberg RH, Brunner I, Bernasconi SM, Zeyer J, Frey B (2012) Bacterial, archaeal, and fungal succession in the forefield of a receding glacier. *Microbial Ecology* **63**, 552-564.

Tables and Figures

Table 2.1 Multiple linear regression analyses of OTU richness, diversity and evenness estimators across plant-associated soils. Non-vegetated Bare Soil is used as a reference. Plant-associated intercept and slope estimates indicate difference in relation to Bare Soil. OTU richness estimates are reported as raw OTU counts whereas diversity and evenness are reported as arcsin($\sqrt{}$(calculated values)). Significant values are in *Bold and Italics* and test the null hypotheses (H$_0$: Intercept _Abies, Luetkea, Phyllodoce, or Saxifraga_ − Intercept _Ref Bare Soil_ = 0; and H$_0$: Slope _Abies, Luetkea, Phyllodoce, or Saxifraga_ − Slope _Ref Bare Soil_ = 0). In other words, significant _P_-values here indicate that there is a difference between intercept or slope terms for treatments *Abies, Luetkea, Phyllodoce,* or *Saxifraga* compared to Bare Soil. Level of significance indicated by asterisks, * refers to 0.01 < _P_ ≤ 0.05, ** refers to 0.001 < _P_ ≤ 0.01, *** refers to _P_ ≤0.001

Treatment	Intercept ± SE	Slope ± SE
OTU Richness (Fungi)		
Bare Soil[§]	***25.42 ± 3.16****	$2.98 \times 10^{-4} \pm 5.91 \times 10^{-3}$
Abies	-0.94 ± 2.46	$1.66 \times 10^{-2} \pm 1.35 \times 10^{-2}$
Luetkea	2.85 ± 2.16	***-2.24 x 10 $^{-2}$ ± 1.08 x 10 $^{-2}$****
Phyllodoce	4.07 ± 2.66	$-1.22 \times 10^{-2} \pm 1.45 \times 10^{-2}$
Saxifraga	-0.96 ± 2.17	$1.70 \times 10^{-2} \pm 1.03 \times 10^{-2}$
OTU Richness (Bacteria)		
Bare Soil[§]	***357.4 ± 34.4*****	***0.18 ± 0.06*****
Abies	26.4 ± 26.4	-0.04 ± 0.15
Luetkea	11.3 ± 21.5	0.08 ± 0.1
Phyllodoce	***61.7 ± 29.5****	-0.31 ± 0.16
Saxifraga	10.9 ± 21.1	0.16 ± 0.1
Diversity (1-D, Fungi)		
Bare Soil[§]	***0.788 ± 0.088******	$-1.20 \times 10^{-4} \pm 1.64 \times 10^{-4}$
Abies	-0.03 ± 0.069	$-9.75 \times 10^{-5} \pm 3.76 \times 10^{-4}$
Luetkea	0.077 ± 0.06	$-2.46 \times 10^{-4} \pm 3.01 \times 10^{-4}$

Phyllodoce	0.015 ± 0.074	$8.13 \times 10^{-5} \pm 4.05 \times 10^{-4}$
Saxifraga	0.0037 ± 0.06	$4.15 \times 10^{-4} \pm 2.87 \times 10^{-4}$
Diversity (1-D, Bacteria)		
Bare Soil[§]	***1.44 ± 0.016***	$-3.80 \times 10^{-6} \pm 2.94 \times 10^{-5}$
Abies	0.001 ± 0.012	$8.90 \times 10^{-8} \pm 7.18 \times 10^{-5}$
Luetkea	-0.002 ± 0.01	$2.76 \times 10^{-6} \pm 4.73 \times 10^{-5}$
Phyllodoce	0.03 ± 0.014	$-1.42 \times 10^{-4} \pm 7.45 \times 10^{-5}$
Saxifraga	-0.01 ± 0.009	$7.11 \times 10^{-5} \pm 4.82 \times 10^{-5}$
Evenness (E_D, Fungi)		
Bare Soil[§]	***0.189 ± 0.028***	$-7.48 \times 10^{-5} \pm 5.21 \times 10^{-5}$
Abies	-0.009 ± 0.022	$-3.06 \times 10^{-5} \pm 1.19 \times 10^{-4}$
Luetkea	0.005 ± 0.019	$8.67 \times 10^{-5} \pm 9.53 \times 10^{-5}$
Phyllodoce	-0.029 ± 0.023	$7.62 \times 10^{-5} \pm 1.28 \times 10^{-4}$
Saxifraga	0.0143 ± 0.019	$1.16 \times 10^{-6} \pm 9.09 \times 10^{-5}$
Evenness (E_D, Bacteria)		
Bare Soil[§]	***0.499 ± 0.038***	***$-1.58 \times 10^{-4} \pm 7.01 \times 10^{-5}$****
Abies	-0.028 ± 0.029	$6.83 \times 10^{-5} \pm 1.71 \times 10^{-4}$
Luetkea	-0.024 ± 0.024	$8.01 \times 10^{-5} \pm 1.13 \times 10^{-4}$
Phyllodoce	0.058 ± 0.033	$-3.48 \times 10^{-4} \pm 1.78 \times 10^{-4}$
Saxifraga	-0.018 ± 0.023	$8.06 \times 10^{-5} \pm 1.15 \times 10^{-4}$
§ = Treatment Bare Soil was selected as a reference level to emphasize the contrast between vegetation non-vegetated soils		

Table 2.2 Core fungi (found in at least 50% of samples) that change in frequency with successional age (distance from the glacier terminus) using Bonferroni-corrected liner regression for combined vegetation sampling and different plant-associated soils. Direction of change indicates if the frequency of OTUs increased (⬆) or decreased (⬇) across distance from the glacier. Species affinity refers to the best BLASTn match (nr/nt) with the exclusion of uncultured/environmental samples.

OTU	t-ratio	*P*-value	Direction of change	Species Affinity	Phylum
All Vegetation					
OTU1	2.0437	0.04479	⬆	*Mortierella alpina*	"Zygomycota"
OTU13	-2.3303	0.02272	⬇	*Satchmopsis brasiliensis*	Ascomycota
OTU27	2.8387	0.00594	⬆	*Mortierella elongata*	"Zygomycota"
OTU3	5.4349	7.77E-07	⬆	*Pochonia bulbillosa*	Ascomycota
OTU5	-4.0741	0.00012	⬆	*Lecythophora sp.*	Ascomycota
OTU1	2.0437	0.04479	⬆	*Mortierella alpina*	"Zygomycota"
Abies					
OTU142	2.3488	0.04071	⬆	*Sarea difformis*	Ascomycota
Luetkea					
OTU109	-3.3111	0.00562	⬆	*Cladophialophora minutissima*	Ascomycota
OTU13	-3.7416	0.00246	⬆	*Satchmopsis brasiliensis*	Ascomycota
OTU18	-3.1925	0.00706	⬆	*Articulospora tetracladia*	Ascomycota
OTU24	-3.2331	0.00653	⬆	*Cryptococcus skinneri*	Basidiomycota
OTU3	3.4869	0.00401	⬆	*Pochonia bulbillosa*	Ascomycota
OTU5	-3.4717	0.00413	⬆	*Lecythophora sp.*	Ascomycota
Saxifraga					

OTU81	-3.2925	0.00583	↓	*Penicillium citreonigrum*	Ascomycota
Phyllodoce					
No significant frequency changes for core OTUs					
Bare Soil					
OTU1	2.3021	0.03509	↑	*Mortierella alpina*	"Zygomycota"
OTU27	2.6516	0.01741	↑	*Mortierella elongata*	"Zygomycota"

Table 2.3 Core bacteria (found in at least 50% of samples) that change in frequency with successive age (distance from the glacier terminus) using Bonferroni-corrected liner regression for combined vegetation sampling and different plant-associated soils. Direction of change indicates if the frequency of OTUs increased (⬆) or decreased (⬇) across distance from the glacier. Species affinity refers to the best Blastn match (nr/nt) with the exclusion of uncultured/environmental samples.

OTU	t-ratio	*P*-value	Direction of change	Genus Affinity	Phylum
All Vegetation					
OTU3	5.6094	4.84E-07	⬆	*Sphaerobacter*	Chloroflexi
OTU11	5.7832	2.47E-07	⬆	*Thermoflavimicrobium*	Firmicutes
OTU21	4.5428	2.57E-05	⬆	*Gemmatimonas*	Gemmatimonadetes
OTU23	4.9367	6.14E-06	⬆	*Humicoccus*	Actinobacteria
OTU25	5.4843	7.82E-07	⬆	*Acidisphaera*	Proteobacteria
OTU45	5.6434	4.25E-07	⬆	*Gemmata*	Planctomycetes
OTU63	-5.0748	3.68E-06	⬇	*Caulobacter*	Proteobacteria
OTU71	3.9549	0.000196	⬆	*Gp1*	Acidobacteria
OTU102	3.9867	0.000176	⬆	*Rhodopseudomonas*	Proteobacteria
OTU105	3.8837	0.000249	⬆	*Rhodopseudomonas*	Proteobacteria
OTU176	5.0846	3.55E-06	⬆	*Gp2*	Acidobacteria
OTU215	5.8533	1.88E-07	⬆	*Nitrospira*	Nitrospira
OTU300	-4.8361	8.89E-06	⬇	*Solirubrobacter*	Actinobacteria
OTU327	3.8681	0.000262	⬆	*Saxeibacter*	Actinobacteria
OTU382	4.7272	0.000013	⬆	*Roseomonas*	Proteobacteria
OTU501	5.4177	1.01E-06	⬆	*Ktedonobacter*	Bacteria_incertae_sedis
OTU508	4.6598	0.000016	⬆	*Herbaspirillum*	Proteobacteria
OTU560	6.4705	1.66E-08	⬆	*Gemmatimonas*	Gemmatimonadetes
OTU686	4.1925	0.000087	⬆	*Flavisolibacter*	Bacteroidetes
OTU762	4.2865	6.34E-05	⬆	*Zavarzinella*	Planctomycetes
Saxifraga					
No significant frequency changes for core OTUs					
Luetkea					
OTU538	5.3365	0.00013	⬆	*Kozakia*	Proteobacteria

OTU572	5.6116	8.46E-05	↑	*Nitriliruptor*	Actinobacteria
Abies					
No significant frequency changes for core OTUs					
Phyllodoce					
No significant frequency changes for core OTUs					
Bare Soil					
OTU16	5.8904	7.36E-05	↑	*Blastococcus*	Actinobacteria
OTU23	4.7931	0.000439	↑	*Humicoccus*	Actinobacteria
OTU35	5.5558	0.000125	↑	*Elioraea*	Proteobacteria
OTU215	4.8953	0.000369	↑	*Nitrospira*	Nitrospira
OTU1390	5.1931	0.000224	↑	*Conexibacter*	Actinobacteria

Figure 2.1 Linear regression of the natural log of copy number per gram soil dry weight of bacteria, fungi and the fungi:bacteria across distance from Lyman glacier terminus. Biomass of fungi (*P* = 0.023) and bacteria (*P* = 0.055) increase with successional age, indicating increase in biomass as the vegetation establishes and substrate becomes more heterogeneous. The ratio of fungal to bacteria biomass remains stable (*P* = 0.998) across the chronosequence indicating similar rates of biomass accumulation.

ln(copy number/g soil Bacteria)
ln Copy number= 0.00263(Distance)+8.875, t=1.95, P=0.055

ln(copy number/g soil Fungi)
ln Copy number= 0.00264(Distance)+8.199, t=2.33, P=0.023

ln(Fungi:Bacteria/g soil)
ln Copy number= 4.19E-06(Distance)-0.677, t=0.003, P=0.998

ln(rRNA copy number)

Distance (m)
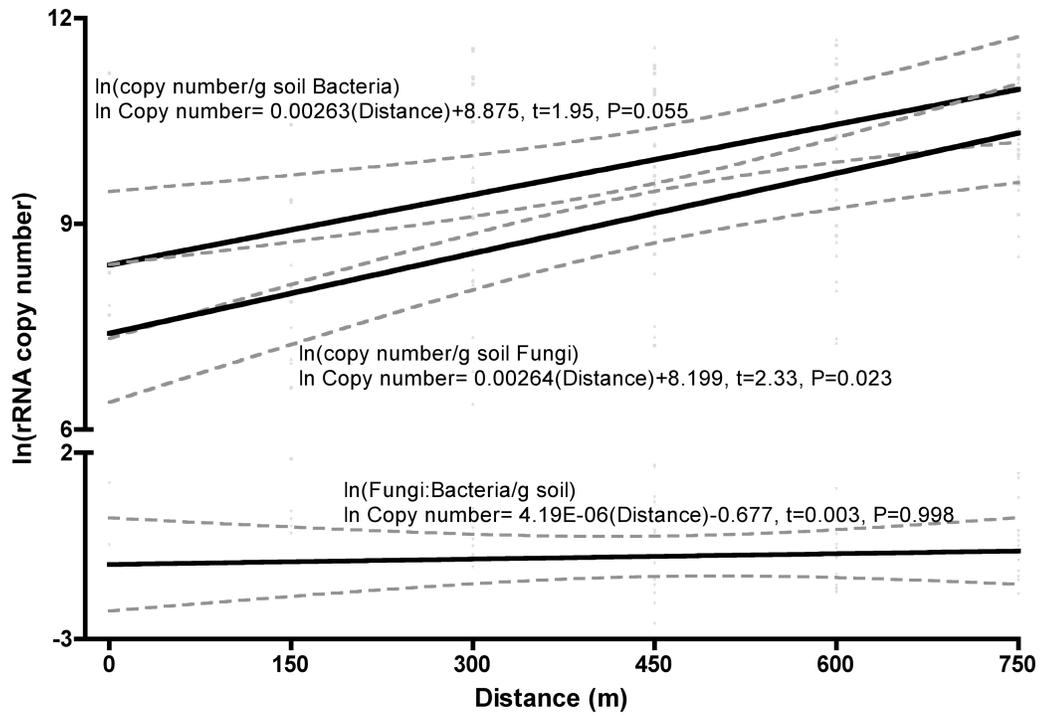
Figure 2.2 Taxonomic distribution of fungal and bacterial OTUs and sequences on the Lyman glacier forefront. Fungal OTUs (a) are dominated by diverse Ascomycota and Basidiomycota. Fungal sequence counts (b) are dominated by Ascomycota and include a significant proportion of basal fungal lineages. Bacterial OTUs (c) and sequences (d) are dominated by Proteobacteria, Acidiobacteria, and Acidiobacteria, but include large numbers of representatives from other phyla.

OTU (Fungi)

a

Glomeromycota 6 OTUs 1.9%
Chytridiomycota 10 OTUs 3.2%
Basal Fungi 18 OTUs 5.8%
Basidiomycota 84 OTUs 27.1%
Ascomycota 192 OTUs 61.9%

Sequences (Fungi)

b

Chytridiomycota 66 Sequences 0.2%
Glomeromycota 39 Sequences 0.1%
Basidiomycota 3111 Sequences 7.9%
Basal Fungi 8311 Sequences 21.0%
Ascomycota 27962 Sequences 70.8%

OTU (Bacteria)

c

Acidobacteria 104 OTUs 2.5%
Actinobacteria 612 OTUs 14.9%
Proteobacteria 1486 OTUs 36.2%
Other 1908 OTUs 46.4%

Sequences (Bacteria)

d

Acidobacteria 12544 Sequnces 12.8%
Actinobacteria 18995 Sequences 19.4%
Proteobacteria 18995 Sequences 31.8%
Other 35122 Sequences 35.9%

Figure 2.3 Rarefaction analysis indicates that the fungal (a) and bacterial (b) communities approach saturation. Note the difference in axis scales.
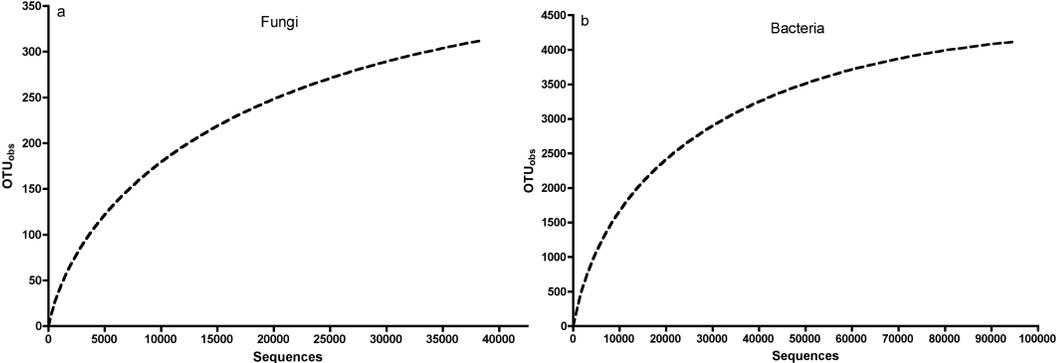
Figure 2.4 Diversity indices for fungal and bacterial communities along Lyman Glacier chronosequence. Fungal community OTU richness (a), diversity [1-D] (b), and evenness (c) show are stable and do not change with distance from glacier terminus (both linear regression and ANOVA). In contrast, bacterial communities differ in richness (d) across distance and vegetation. Bacterial diversity [1-D] (e) does not change with distance from glacier or with vegetation whereas bacterial evenness (f) decreases with distance from the glacier terminus. We provide the t-statistics for the slope terms and F-statistics for distance for the complete model with distance as a continuous variable (df=1) as well as F-statistics for vegetation and vegetation*distance interaction terms.
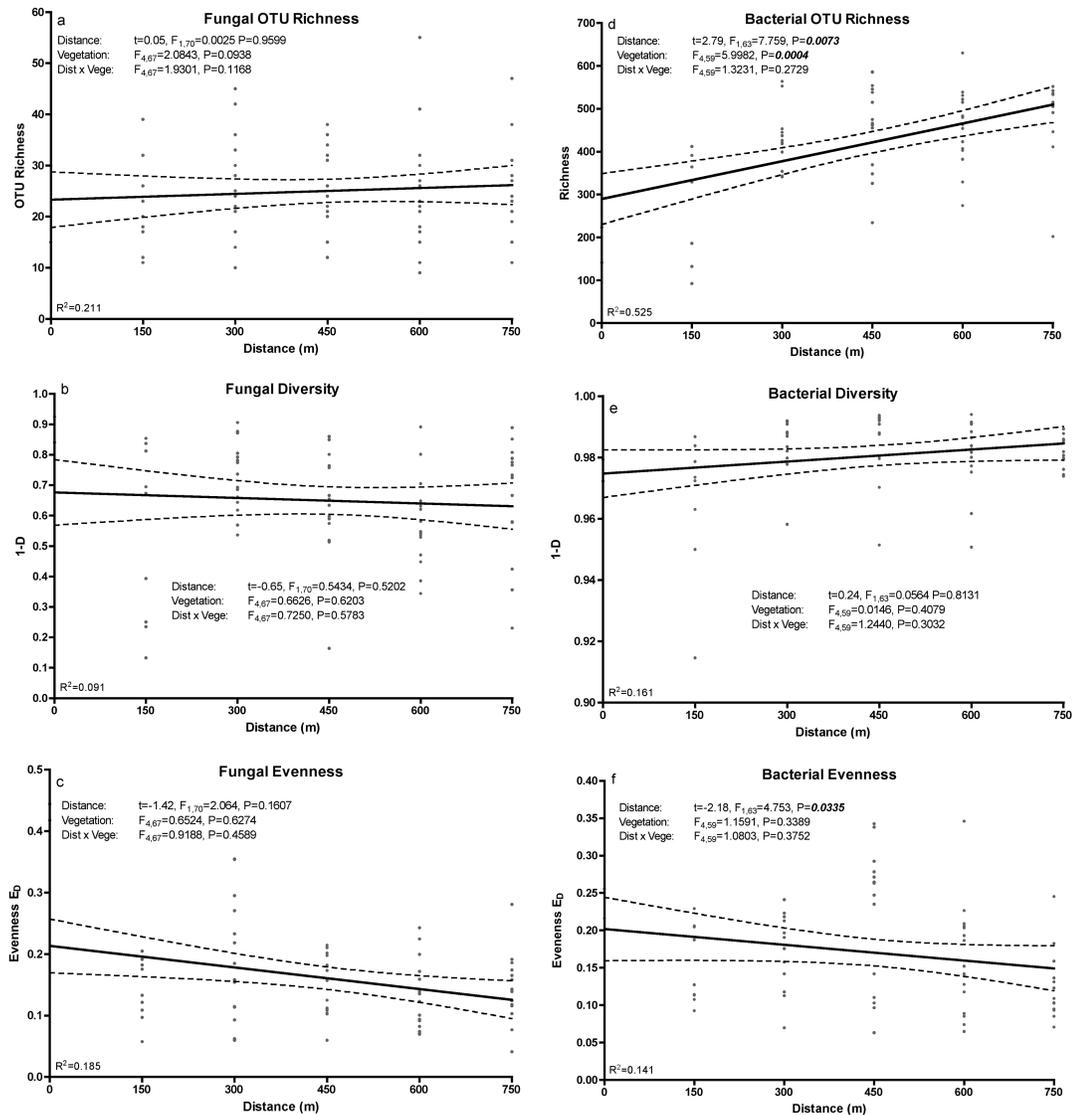
Figure 2.5 Nonmetric Multidimensional Scaling (NMS) representation of fungal and bacterial communities along Lyman glacier chronosequence indicating community trajectories with increasing successional age. NMS of fungal (a) and bacterial (b) show strong successional trajectories with dashed arrows representing directionality of community shifts as indicated by significant linear regression statistics. Bacterial communities (c) in vegetated and bare soils are distinct along Axis 3 (different letters indicate significant differences in Tukey's HSD post-hoc analysis at $\alpha < 0.05$).
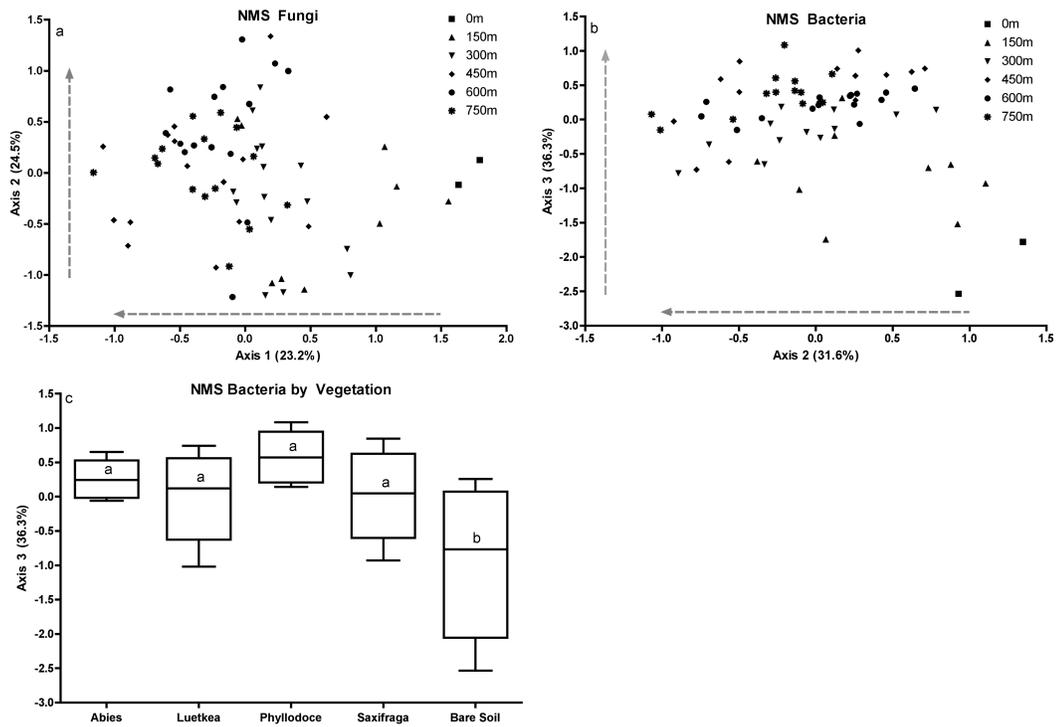
Figure 2.6 Nonmetric Multidimensional Scaling (NMS) Axis 3 score standard deviations for bacterial communities along the Lyman Glacier chronosequence decrease with distance from the glacier terminus (successional age). Points represent standard deviations of vegetation treatments for each distance category. When the two points (0m and 150m above 0.5) are removed from the analysis, the slope of the regression line no longer differs from zero. The analysis suggests that bacterial communities may converge with increasing successional age.

45
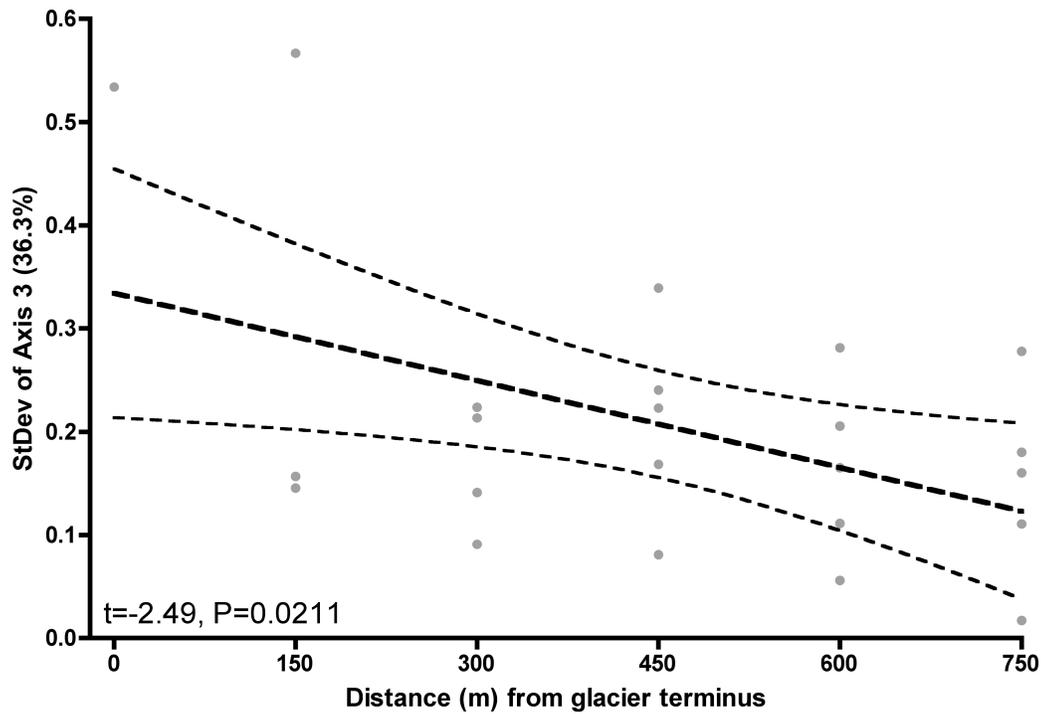
Chapter 3 - Phylogenetic diversity analyses reveal disparity between fungal and bacterial communities during primary succession[2]

Abstract

The early community assembly of fungi and bacteria differ in many important ways, including contrasting successional trajectories. Here, we examine fungal and bacterial succession across a well established temperate glacier chronosequence to determine if microbial succession is constrained by phylogenetic relatedness of the establishing microbial community. Fungal and bacterial communities were queried across a recently deglaciated forefront (150 m, 300 m, 450 m, 600 m, 750 m from glacier terminus) using a variety of phylogenetic diversity metrics: Net Relatedness Index (NRI), Nearest Taxon Index (NTI) and Faith's Index of phylogenetic diversity (FI). Additionally, organismal grouping for these analyses consisted of an all-inclusive grouping of all fungi and all bacteria as well as lower, less-inclusive taxonomic hierarchies at the Phylum and Class level as well as select, well-defined functional groups (N-fixation and photosynthetic bacteria). Our analyses suggest that fungi and bacteria differ in their phylogenetic clustering across succession age; fungi are generally unresponsive across successional age whereas bacteria are strongly structured phylogenetically with successional age. Importantly, our results highlight that analyses on Kingdom or Domain levels are not adequate by themselves to allow for understanding successional dynamics. Consequently, investigations should include both broad level (Kingdom or Domain) as well as lower, less-inclusive groups (perhaps even metabolically known taxa) to dissect successional drivers and community assembly dynamics.

Keywords: *Primary succession, Fungi, Bacteria, Phylogenetic diversity, NTI, NRI*

Introduction

Primary successional dynamics of microbial communities are crucial for understanding community assembly rules, nutrient transformations and pedogenesis that facilitate the successful colonization by vascular plants (Fierer *et al.* 2010). It is these early colonizing microorganisms that drive the establishment of ecosystem processes (Schmidt *et al.* 2008) that allow the establishment of vegetation and higher-trophic level food webs (Walker & del Moral 2003). Molecular studies on microbial primary succession have been conducted via rRNA community fingerprinting (Sigler & Zeyer 2002), cloning and sequencing (Jumpponen 2003), PLFAs (Ohtonen *et al.* 1999; Tscherko *et al.* 2005), and T-RFLPs (Zumsteg *et al.* 2012). Studies on early microbial primary succession using next generation sequencing (NGS) have largely focused on shifting community member abundances as a mechanism to understand community dynamics associated with succession (Schütte *et al.* 2010; Blaalid *et al.* 2012; Brown & Jumpponen 2014). Additionally, these early NGS studies often only consider a component of the total microbial constituents: root-associated fungi (Blaalid *et al.* 2012) or bacteria associated with plant colonization (Knelman *et al.* 2012). While targeting specific microbial community constituents allows for detailed interrogation into part of the overall successional story, focusing solely on bacteria, fungi, or archaea can limit our understanding of the microbial community as a whole. It is the cross-domain investigations that allow a more complete understanding of microbial successional dynamics integrating microbial interactions. The first deep-sequencing (NGS) queries into joint fungal and bacterial successional dynamics are beginning to emerge (Brown & Jumpponen 2014; Cutler *et al.* 2014). One of the most important results from these pioneering studies is that fungal and bacterial communities do not respond similarly across a primary successional landscape. These findings are in contrast with earlier studies (Zumsteg *et al.* 2012) possibly due to the limited depth of inquiry that T-RFLPs allow compared to NGS. Interestingly, both Brown & Jumpponen (2014) and Cutler *et al.* (2014) found that bacteria and fungi have contrasting successional trends but in different ways. Brown & Jumpponen focused on early successional dynamics (0-80 years) and found a tight link between plant establishment and bacterial communities but the fungal communities responded little to vegetation and were mainly influenced by time since deglaciation. In contrast, Cutler and coauthors' queried a more

extensive successional timeframe (165-852 years) and found that fungi are closely linked to plant establishment but bacteria are less so. The reason for this discrepancy is uncertain but may be a result of differing successional ages, substrate (glacier forefront vs. volcanic deposition) associated nutrient limitations, or that these observed successional patterns are location specific and therefore context dependent.

There is evidence that early bacterial communities may converge during primary succession (Brown & Jumpponen 2014; see Fig. 6). Similarly, fungal communities may converge phylogenetically (Jumpponen *et al.* 2012; see Figs. 7 and 8). This convergence was suggested based on decreasing standard deviations of ordination loading scores over successional time and can be considered circumstantial as the underlying reasons for this observation remained poorly understood. In this contribution, we explicitly test if the observed compositional convergence is phylogenetically constrained. That is, does phylogenetic diversity differ between fungi and bacteria during primary succession?

Colonization of newly developed substrate presents a challenge as the opportunities to colonize are hindered by strong environmental and biotic filtering. For a glacier forefront, the abiotic filters include strong UV irradiation, drastic diurnal temperature fluctuations, long-lasting snow cover, and extreme nutrient limitations (Jumpponen et al., 2012). The ability to establish in such harsh environments (or any environment) is likely related to evolutionary histories that determine the traits mandatory for successful establishment (and dispersal). That is, microbes that are suited for given environmental conditions likely share an evolutionary history that accounts for those traits determining the success in establishing as seen with bacteria (Philippot *et al.* 2010). Such a phylogenetic signal may be even more evident for fungi (Maherali & Klironomos 2007). It is in this light that investigations into the phylogenetic diversity of microbial succession can elucidate details of community assembly rules and successional dynamics.

Evolutionary frameworks have been widely used to understand community ecology (see Table 1 in Vamosi *et al.* 2009). Some examples include communities of trees (Webb 2000), shrubs and grasses (Purschke *et al.* 2013), and mammals (Cardillo 2011). Despite the common use of phylogenetically conserved gene regions for bacterial and fungal amplicon libraries, studies of bacterial (Horner-Devine & Bohannan 2006; Newton *et al.* 2007) and fungal (Anderson *et al.* 2004; Merckx *et al.* 2012; Rincón *et al.* 2014; Rämä *et al.* 2014) phylogenetic

diversity are rare and largely focus on specific groups of microbes at lower taxonomic levels. Even more rare are investigations that target microbial phylogenetic divergence across environmental gradients. Most studies investigating phylogenetic patterning across environmental gradients have focused on altitudinal changes (Bryant *et al.* 2008). Some bacterial phyla may shift in phylogenetic diversity with altitude (Wang *et al.* 2012), whereas Sebacinoid fungi did not change phylogenetically with altitude (Garnica *et al.* 2012). To our knowledge, concurrent phylogenetic diversity analyses of fungi and bacteria are lacking. Additionally, phylogenetic diversity has yet to be studied in a primary successional environment. Integration of evolutionary history into primary community assembly rules is crucial for an improved understanding of successional dynamics.

Materials and Methods

Sampling Location

The forefield of Lyman Glacier is located within the Wenatchee National Forrest in Glacier Peaks Wilderness Area in Washington State, USA (48°10'14"N, 120°53'44"W; ~1900 m a.s.l). The glacier has receded more than 1 kilometer over the past ~120 years (Jumpponen *et al.* 1998), is currently fragmented, and lays on top of the bedrock of the north slope of Chiwawa Mountain. Plant (Jumpponen et al. 2012) successional dynamics as well as bacterial and fungal community dynamics over successional age and in response to plant establishment (Brown & Jumpponen 2014) have been previously characterized at this site.

Sampling and Sequence Generation

Detailed sampling protocols are available in Brown & Jumpponen (2014); we used those samples for the current analyses. Briefly, topsoil samples were collected along the degleciated chronosequence (150 m, 300 m, 450 m, 600 m, and 750 m from the Lyman glacier terminus northwards toward the terminal moraine representing *circa* 90 years of successional time). Samples differing in vegetation were also collected: non-vegetated bare soil as well as rhizosphere samples from four plants with differing mycorrhizal habits (*Abies Lasiocarpa*

(Hook.) Nutt [ectomycorrhizal], *Luetkea pectinata* Kuntze [arbuscular mycorrhizal], *Phyllodoce empetriformis* D.Don [ericoid mycorrhizal], and *Saxifraga ferruginea* Graham [non-mycorrhizal]) as determined earlier for this forefront in Cázares *et al.* (2005). Previously, 454-amplicon libraries were generated for bacteria (primers B-9F and A-MID-541R where A is the sequencing adaptor, MID is 8-basepair unique molecular identifier tag, and B is the emPCR bead adhering adaptor) and fungi (primers A-MID-ITS1f and B-ITS4; see Table B.1 for MID sequences). For detailed description of amplicon preparation, see Brown & Jumpponen (2014). For this study, we reanalyzed the bacterial dataset generated from Brown & Jumpponen (2014) and generated a new fungal library from the same DNA extracts using primers that amplify the 5' end of the Large Subunit (LSU) of the ribosomal RNA gene repeat (our previous ITS1 and ITS2 analyses are not suitable for alignments required for analyses of phylogenetic diversity).

LSU amplicons were generated using three technical replicates utilizing a two-step PCR to minimize MID-induced PCR biases (Berry et al. 2011) in 25µL reaction volumes. Primary PCR conditions were: 10 µM forward and reverse primers (LR0R and LR3; see Amend et al. 2010), 5ng template DNA, 200 µM of each deoxynucleotide, 2.5mM $MgCl_2$, 5 µL 5x Green GoTaq® Flexi Buffer (Promega, Madison, WI), 7.8 µL molecular biology grade water, and 1 U GoTaq® Hot Start Polymerase (Promega, Madison, WI). Primary PCR cycle parameters consisted of a 94° initial denaturing for 4 min, 25 cycles of 94° denaturing for 1 min, 53° annealing for 45 sec. and 72° extension for 2 min followed by a 72° final extension for 8 min. Resultant PCR product was used as a template DNA for the secondary PCR (with primer fusion constructs that included 8-bp multiplexing tags (MID) and 454-sequencing specific linkers). Conditions of secondary PCR were: 5 µL primary PCR product, 5 µM forward and reverse primers (A-MID-LR0R and B-LR3 respectively), 2.5mM $MgCl_2$, 200 µM of each deoxynucleotide, 5 µL 5x Green GoTaq® Flexi Buffer (Promega, Madison, WI), 7.3 µL molecular biology grade water and 1 U GoTaq® Hot Start Polymerase with PCR parameters of 94° initial denaturing for 4 min, 5 cycles of 94° denaturing for 1 min, 54° annealing for 1 min and 72° extension for 2 min followed by a 72° final extension for 10 min. Positive and negative controls were included and samples remained free of contamination. The three technical replicates of each experimental unit were pooled and cleaned using Agencourt® AmPure® cleanup kit using a SPRIplate 96-ring magnet (Beckman Coulter, Beverly, MA) following the manufacturer's protocol with the exception that we used a 1:1 bead solution to amplicon volume

for better discrimination against small DNA fragments. Resultant amplicons were pooled equamolarly and the pooled amplicons cleaned once more with AmPure® and 454-pyrosequenced (GS FLX-Titanium, Roche Applied Science, Indianapolis, IN) at the Integrated Genomics Center at Kansas State University (Manhattan, KS).

Bioinformatic Processing and Analyses

The obtained bacterial and fungal sequences (accessioned in the Sequence Read Archive at NCBI as individual fastq files per experimental unit; BioProject **PRJNA201483**, biosample accessions **SRR943235-SRR973301** for bacteria and **SRR1016610-SRR1016625**, **SRR1016734-SRR1016736**, **SRR1016744-SRR1016777**, **SRR1016785-SRR1016788**, and **SRR1016859-SRR1016865** for fungi) were processed using the program MOTHUR (v. 1.31.2; Schloss *et al.* 2009). The obtained sequencing flowgrams were denoised with the MOTHUR embedded PyroNoise program (Quince *et al.* 2009) and the remaining high-quality sequences aligned to either the SILVA 16S reference alignment (bacteria) or to a modified LSU alignment (see Brown *et al.* 2014a) based on James' *et al.* (2006) All Fungi Tree of Life (AFTOL) LSU alignment. Putative chimeric sequences were identified using UCHIME (Edgar *et al.* 2011) and removed. Operational taxonomic units (OTUs) were defined using a 97% similarity threshold (average neighbor algorithm). The generated OTU x sample matrix (shared file in MOTHUR) was filtered to remove singleton OTUs (Tedersoo *et al.* 2010; Brown *et al.* 2014b) and subsampled to equal depth per experimental unit (600 sequences per bacterial experimental unit and 200 per fungal experimental unit) to limit biases from unequal library sizes (Gihring *et al.* 2012). OTUs were classified using the MOTHUR implemented Naïve Bayesian Classifier (Wang *et al.* 2007) against Ribosomal Database Project's reference files (16S v. 9 or 28S v. 7 rRNA reference) and all non-target OTUs were removed. Each OTU was manually annotated for putative ecology where possible. Using these classifier-based identities, a comprehensive list of sequence names for all fungi and all bacteria as well as sequence names for the most abundant phyla (Basidiomycota and Ascomycota for fungi and Acidobacteria and Actinobacteria for bacteria) as well as the bacterial classes α-Proteobacteria, β-Proteobacteria, and γ-Proteobacteria) were generated (sequencing provided names for each sequences that were positively screened to belong to the selective groups). Remaining fungal or bacteria phyla were either too rare, too

skewed in distribution, or were dominated by only few members that would confound phylogenetic diversity analyses. Additionally, sequence lists of all members of diazotrophs and photosynthetic bacteria representing defined functional groups were generated based on annotations. Initially mycorrhizal and entomopathogenic fungi were intended as fungal functional groups but these were too skewed in OTU abundance and omitted from further analyses.

The sequence lists for each group were used to harvest sequences from the original sff files and new fasta and group files were generated with only quality-controlled sequences of these selected groups. These newly derived fasta files were aligned as above. This method ensured the presence of only informative gaps for each group. For each queried group, sequences were subsampled (randomly without replacement) such that there were equal numbers of sequences per experimental unit. Where subsampling a given group would lead to elimination of 25% or more of experimental units at a depth of 100 sequences per experimental unit, the experimental units from each distance from the glacier terminus (150 m, 300 m, 450 m, 600 m, 750 m) were collapsed into one resulting of 5 functional experimental units - one for each distance category. This lead to a subsample depth of: All Fungi – 150 sequences per sample, Ascomycota – 150 sequences, Basidiomycota – 700 sequences (collapsed samples), All Bacteria – 300 sequences, Acidobacteria – 125 sequences, Actinobacteria – 150 sequences, α-Proteobacteria – 100 sequences, β-Proteobacteria – 850 sequences (collapsed), γ-Proteobacteria – 750 sequences (collapsed), Diazotrophs – 70 sequences (collapsed), and Photosynthetic Bacteria – 46 sequences (collapsed). Relaxed Neighbor-Joining trees were generated for each subsampled and aligned sequence list using the program CLEARCUT (Sheneman *et al.* 2006).

Indices of phylogenetic diversity were calculated in R (v. 2.10.1; R Development Core Team 2007) using package *Picante* (V. 1.1-1; Kembel *et al.* 2010). Nearest Taxon Index (NTI) and Net Relatedness Index (NRI) (Webb 2000) were calculated using the null model 'independent swap' (Connor & Simberloff 1979; Gotelli & Entsminger 2003) with 999 randomizations runs with 1000 iterations (see Webb *et al.* 2008); Faith's index of phylogenetic diversity (FI; Faith 1992) was non-iterative. NTI is a measure of phylogenetic clustering focusing on terminal nodes calculated as the average branch length distance between a sequence and its closest relative standardized by the maximum possible values for the tree. Thus NTI is less sensitive to deeper topologies, whereas NRI is based on mean pairwise distance of the

terminal nodes across the whole tree and is a more sensitive to deeper topological branching (Lozupone 2007). Higher values of NTI and NRI indicate stronger phylogenetic clustering whereas lower values represent no phylogenetic clustering or overdispersion in the communities. FI is the summative branch length between all members in a sample and is positively correlated with species or OTU richness (Vamosi *et al.* 2009). Communities were considered significantly phylogenetic clustered (NTI and NRI) if observed phylogeny was more extreme than randomized trees (based on 1000 Monte Carlo simulations).

Changes in phylogenetic diversity over successional age were tested using regression analyses. As reported previously, vegetation type had little effect on bacterial (bare soil samples were distinct from vegetated samples but vegetation types were similar) or fungal succession but both were strongly structured by successional age (Brown & Jumpponen 2014). As a result, only distance from the glacier terminus (age of substrate exposure) was used in these analyses. Metrics of phylogenetic diversity were regressed against distance from the glacier terminus using linear models. The model fit was tested using a Lack of Fit test and where significant at $\alpha=0.10$, further polynomial (quadratic) regression analyses were explored. If this model fit the data better, the polynomial regression models are reported. For the collapsed samples, polynomial regression was only used if the data were much better explained using such models (based on Adjusted R^2 values). All regression analyses were conducted using JMP® (v. 7.0.2; SAS Institute, Cary, NC, USA).

Results

All tested taxonomic or functional groups were phylogenetically clustered to some degree. However, this clustering differed between fungi and bacteria. Fungi consistently had many clustered samples when compared to randomly generated trees. Approximately half of the experimental units were significantly clustered in analyses that included fungi or only Ascomycota (see Table 3.1) whereas all communities within the phylum Basidiomycota (collapsed into five samples – 150 m, 300 m, 450 m, 600 m, and 750 m) were clustered. Furthermore, the proportion of fungal experimental units that were significantly phylogenetically clustered for NTI and NRI differed (NTI > NRI; Fisher's Exact Test P=0.0015). Similarly to the analyses focusing on fungi, bacteria in many samples were clustered phylogenetically and the

frequency of clustered samples based on NTI was greater than that based on NRI (P<0.0001; see Table 3.1). This disparity between NTI clustering and NRI clustering was sometimes drastic; over half of the queried α-Proteobacteria samples were clustered with respect to NTI and none appeared clustering using NRI.

Regression analyses of NTI, NRI and FI provide some interesting and contrasting results between fungi and bacteria. When phylogenetic diversity metrics for all fungi were regressed with distance from the glacier terminus, the slope estimates for NTI (t=0.00, P=0.999; Fig 3.1a), NRI (t=0.58, P=0.566; Fig 3.1b), or FI (t=-0.19, P=0.859; Fig 3.1c) did not differ from zero. This is in contrast with the bacterial analyses: both NTI and NRI declined with distance from the glacier, but more so for NTI than NRI (t=-5.37, P<0.0001, see Fig 3.1d; t=-1.84, P=0.0712, Fig 3.1e; respectively). This may be suggestive that for bacteria, there is an increase of niche diversity as the forefront develops. Further, the bacterial analyses indicated strong increases in phylogenetic branch length (FI; t=5.93, P<0.0001; Fig 3.1f) with distance from the glacier within each sample.

In the analyses of the lower (less inclusive) fungal taxonomic groups, a slightly different story emerged. Regression of Basidiomycota indicated no relationship with successional age for NTI (t=-0.81, P=0.479; Fig 3.2a), NRI (t=-0.62, P=0.578; Fig 3.2b) or FI (t=0.77, P=0.497; Fig 3.2c). Analysis of Ascomycota showed a slightly different story yet; regression models that best fit the NTI data were unimodal with a peak between 450 m and 600 m from the glacier terminus (df=48, P=0.0096; Fig 3.2d) whereas NRI and FI are unchanging (t=0.79, P=0.435; Fig 3.2e and t=-1.28, P=0.207; Fig 3.2f - respectively).

The different taxonomic levels and functional groups of bacteria showed some intriguing differences. The results for Acidobacteria followed those of the combined bacterial groups: NTI (t=-7.50, P<0.0001; Fig 3.3a) and NRI (t=-7.45, P<0.0001; Fig 3.3b) strongly and highly significantly decreased with distance from the glacier terminus, whereas FI (t=8.65, P<0.0001; Fig 3.3c) increased. Actinobacteria results were slightly different: similarly to Acidobacteria, NTI decreased (t=-4.10, P=0.0001; Fig 3.3d) and FI increased (t=2.60, P=0.0116; Fig 3.3f) with distance, but in contrast, NRI did not change (t=-1.30, P=0.1994; Fig 3.3e). These Actinobacterial data appear increasingly heteroskedastic with distance from the glacier but regression analysis of the absolute value of residuals indicate that the apparent increase in variance was not significant (t=1.19, P=0.2394). Even within the Proteobacteria, there were clear

contrasts. The NTI estimates for α-Proteobacteria decreased (t=-3.28, P=0.0019; Fig 3.3g) with distance, whereas NRI remained stable (t=-0.38, P=0.7025; Fig 3.3h) and FI increased (t=3.77, P=0.0004; Fig 3.3i). The β-Proteobacteria analyses suggested that for both NTI and NRI a quadratic regression was the most appropriate (NTI – df=4, P=0.0475, Fig 3.3j; NRI – df=4, P=0.0160, Fig 3.3k) responding strongly to distance from the glacier terminus but in opposite directions. The best-fit line for NTI was concave with a local minimum around 600 m, whereas the best-fit line for NRI was convex with a maximum around 450 m from the glacier terminus. The relationships of NTI and NRI within β-Proteobacteria seemingly complemented each other, perhaps representing taxon replacement from more closely related terminal nodes to taxa from the same deep branch but more dissimilar at the branch termini. However, total branch length (FI) increased linearly with distance (t=4.80, P=0.0172; Fig 3.3l) suggesting an increase in β-Proteobacterial richness. The γ-Proteobacterial NTI was best explained using quadratic regression and responded with distance from the glacier terminus (df=4, P=0.0079; Fig 3.3m) with minimum around 550 m whereas there were no good linear or quadratic models for NRI (df=4, P=0.1291; Fig 3.3n). The γ-Proteobacterial FI followed a convex quadratic regression with distance from the glacier (df=4, P=0.0105; Fig 3.3o) with a maximum around 550 m. Finally, the bacterial functional groups were rarely responsive to distance from the glacier. The diazotrophic communities were non-responsive for NTI (t=-1.03, P=0.3779; Fig 3.3p), NRI (t=0.81, P=0.4763; Fig 3.3q) or FI (t=0.12, P=0.9120; Fig 3.3r). In contrast, NTI (t=-0.46, P=0.6777; Fig 3.3s) and NRI (t=0.03, P=0.9771; Fig 3.3t) estimates of the photosynthetic bacteria were unresponsive, but FI estimates responded positively (t=3.49, P=0.0396; Fig 3.3u) to distance from the glacier terminus indicative of an increase on carbon fixing bacterial richness that establish independent of phylogenetic constraints.

Discussion

The early community assembly of fungi and bacteria in primary successional systems share many similarities including the stochastic propagule arrival of allochthonous propagules onto virgin substrates. The propagating species can either establish or fail to initiate metabolic activity, largely driven by environmental and physiological restrictions (Jumpponen & Egerton-

Warburton 2005; Koide *et al.* 2011). Additionally, these burgeoning communities are further filtered via biotic interactions and priority effects can control the success of competing species with similar niche requirements. Our results indicate that one should not consider the assembly of highly inclusive groups (Kingdom or Domain) as homogeneous and uniform. Rather, to understand assembly dynamics, lower and less inclusive groups must be considered. Considering subgroups, functional groups, or perhaps even individual species as the level of inquiry allows for a deeper and more complete understanding of the assembly of microbial communities.

In our previous characterization of fungal and bacterial succession at this site (Brown & Jumpponen 2014), fungal and bacterial dynamics were differently influenced by the plant establishment and possessed differing successional trajectories. This study further establishes that fungi and bacteria, as well as groups within them, differ in their successional dynamics. Fungi across Lyman Glacier basin were largely unresponsive to time since deglaciation in their phylogenetic structuring. This contrasts the strong fungal community turnover at this site (see Brown & Jumpponen 2014) suggesting that species turnover and the species replacement are by members that are similarly distributed across the phylogeny. In contrast, bacterial phylogenetic diversity responded more strongly. Bacteria shift from phylogenetically clustered communities toward over-dispersed 'terminal' communities that include greater numbers of phylogenetically more distant taxa. This shift in phylogenetic structure suggests that early successional substrate has fewer niches that selects for a more genetically similar bacterial consortia and as the forefront develops, more disparate niches open allowing for more dissimilar occupying bacteria.

Analyses of fungi as an all-inclusive single group showed that approximately half the samples were phylogenetically clustered compared to randomly generated trees under the independent swap null model and that the level of clustering was unchanging over the chronosequence. Similarly, fungi within the phylum Basidiomycota did not shift in phylogenetic clustering over the chronosequence but were consistently phylogenetically clustered. These results coupled with our previous observations (Brown & Jumpponen 2014) indicating that fungal communities shift over succession strongly suggest that fungi in primary successional environments are prone to frequent species turnover and species replacement by closely related taxa (perhaps con-generics). This suggests that the basidiomycetes across Lyman Glacier forefield posses highly similar niche attributes and the substrate has low niche diversity for the Basidiomycota allowing only similar and closely related taxa to occupy these niches.

Interestingly, only Ascomycota had any discernible patterns with successional age. The ascomycete Nearest Taxon Index (NTI) was best represented by curvilinear models (see Fig. 3.2a) over the chronosequence. The early successional ascomycetes were randomly dispersed and their clustering increased with successional age until ~500 m from the glacier terminus after which point the clustering become less strong. The quadratic regression model implies that there is an initial phylogenetic convergence of the ascomycete community (NTI) followed by a later divergence. Reasons for this pattern remain unclear but may be best explained by low lying areas of the Lyman Glacier forefield at ~500m from the glacier terminus that also possess distinct plant communities (see Jumpponen *et al.* 2012). These depressions may alter both hydrology and microclimate at this site. Basidiomycetes, like the fungal community as a whole, were not influenced by this topographical change. This highlights how different components of the fungal community follow different rules during early community assembly. Whilst basidiomycetes, and fungi in general, neither converge nor diverge, ascomycetes were phylogenetically structured during succession and initially converged. This distinction between the Ascomycota and other fungi may in part be explained by the greater diversity and sequence abundance within the Ascomycota.

When all bacteria were queried for phylogenetic diversity, clear contrasts emerged between bacteria and fungi. Whereas fungi had no discernable changes in phylogenetic diversity over successional time, bacteria did so strongly. The overall bacterial community NTI and FI responded strongly to distance from the glacier terminus; these trends were also marginally significant for NRI. In general, bacterial NTI and NRI decreased with a related increase in FI over time. The increasing FI is in accordance with previous investigation of bacterial communities in the Lyman Glacier Basin that showed a linear increase of bacterial richness with distance from the glacier terminus (Brown & Jumpponen 2014). Interestingly, the observed clustering was much more obvious for the overall bacterial communities as well as many of the less-inclusive phylogenetic groups at the terminal nodes (NTI) than for the deeper nodes (NRI). The stronger phylogenetic clustering at the terminal nodes suggests that these bacterial communities are very closely related to each other, perhaps at the genus level, further evidence for conserved niche space early on the chronosequence with niche divergence with successional age.

Perhaps the most interesting observations emerge when successional dynamics are

queried at a lower phylogenetic resolution (phyla or classes). The bacterial successional dynamics seem as diverse as bacteria themselves. The strong response of the Acidobacteria that shadows the patterns found for all bacteria coupled with the great abundance of Acidobacteria in soils suggest strongly that these bacteria are driving much of the overall community dynamics. Much about the Acidobacteria remains unknown, but these common soil-borne bacteria are generally thought to be oligotrophic (Fierer *et al.* 2007) and may be proficient in acquiring recalcitrant limiting nutrients from substrata (Jangid *et al.* 2013). These are characteristics that are likely crucial for survival in early primary successional systems where available nutrients are extremely limiting. Additionally, the high relative NTI and NRI in early succession for Acidobacteria align well with previous characterization of bacteria in this system: several Acidobacteria increased in abundance over the chronosequence (Brown & Jumpponen 2014). As a more diverse Acidobacterial community establishes in this successional environment, NTI and NRI predictably drop with a concomitant increase in FI. Interestingly, Actinobacteria and α-Proteobacteria also responded similarly to the complete bacterial community: while NTI declines and FI increased, NRI for these bacteria did not change. Unlike Acidobacteria that are generally considered oligotrophic and thrive in low nutrient conditions similar to those found in glacier forefronts, the copiotroph-oligotroph spectrum may not be applicable to the Actinobacteria and α-Proteobacteria (Ferrer *et al.* 2007) because these taxa did not shift in abundances with carbon availability. We interpret this to mean that the Actinobacteria and α-Proteobacteria are strongly structured early in the chronosequence. Actinobacteria and α-Proteobacteria are ubiquitous and the negative regression of NTI with distance from the glacier can perhaps be best explained by arrival and establishment of closely related Actinobacterial or α-Proteobacterial taxa that diverge as a result of pedogenesis and ecosystem transformations that open new niches for more distantly related Actinobacteria or α-Proteobacteria. Contrastingly, both β-Proteobacteria and γ-Proteobacteria ha curvilinear relationship with distance from the glacier terminus with maxima or minima at around 600 m. This intriguing similarity with ascomycotetes may suggest that similar subtle shifts environmental conditions drive the community dynamics within these groups.

We also attempted to gain further insights into ecosystem processes by specifically targeting two fundamentally important functional groups: the diazotrophs and photosynthetic bacteria. Primary successional environments suffer from extreme nutrient limitations, primarily

nitrogen. Vitousek *et al.* (1987) demonstrated nitrogen-fixing plants often invade young successional soils and this natural increase in nitrogen inputs alters the trajectory of ecosystem development in volcanic soils. Experimental nutrient amendments of the same volcanic soils demonstrated that nitrogen is the most limiting to plant establishment and concurrent ecosystem development (Vitousek *et al.* 1993). Similar dynamics are likely to occur for primary successional glacier systems. Nemergut *et al.* (2007) demonstrated that bacteria established ecosystem processes prior to any visible establishment of vegetation. Additionally, nutrient amendments to unvegetated recently deglaciated soils cause a shift in bacterial community structure to resemble those during later succession (Knelman *et al.* 2014). While important observations from the perspective of ecosystem process establishment, these analyses provided little additional insight to the overall successional dynamics of bacterial community composition. In our analyses, the diazotrophic and photosynthetic functional groups remained stable over time with the exception of FI for the photosynthetic bacteria. These results suggest that the selected functional traits are little structured during succession, despite their importance in early development of ecosystem processes (Schmidt *et al*. 2008). This is particularly interesting given the tight phylogenetic conservation of nitrogen fixation; this suggests that these functional groups that are crucial for initial establishment of ecosystem processes are compositionally and phylogenetically stable during early ecosystem development.

Our data strongly suggest that in order to better understand successional dynamics and ecosystem development in general, analyses if less inclusive levels of taxonomic hierarchy can provide far greater insight into the ecosystem and community development. Comparing Kingdoms or Domains, such as Fungi and Eubacteria, as a whole does not allow decoupling which community members drive the observed dynamics if such become evident. We argue that greater insights can be gained from analyses that target lower phylogenetic groupings, functional groups, or perhaps even target taxa with well-defined functional capabilities. While we may be far from fully understanding the metabolic roles of individual species in complex microbial systems, the phylogenetically conserved metabolic traits (N-fixation, photosynthetic activity, wood decomposition, or utilization of other recalcitrant substrates) may yield greatest insights into functional aspects of community assembly during early primary succession. The commonly observed increases/decreases in NTI (particularly within the Ascomycota, Acidiobacteria, Actinobacteria and α-Proteobacteria) and lacking strong responses in NRI suggest common

community reordering with new but closely related community members. This is like fine-tuning or optimizing the communities for harsh but slightly shifting environmental conditions.

Acknowledgements

References

Amend AS, Seifert KA, Samson R, Bruns TD. (2010). Indoor fungal composition is geographically pattered and more diverse in temperate zones than in the tropics. Proceedings of the National Academy of Sciences 107: 13748-13753.

Anderson TM, Lachance M, Starmer WT. (2004). The relationship of phylogeny to community structure: the cactus yeast community. The American Naturalist 164: 709-721.

Berry D, Mahfoudh KB, Wagner M, Loy A. (2011). Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Applied and Environmental Microbiology 77: 7846-7849.

Blaalid R, Carlsen T, Kumar S, Halvorsen R, Ugland KI, Fontana G, Kauserud H. (2012). Changes in the root-associated fungal communities along a primary succession gradient analyzed by 454 pyrosequencing. Molecular Ecology 21: 1897-1908.

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

Brown SP, Rigdon-Huss AR, Jumpponen A. (2014a). Analyses if ITS and LSU gene regions provide congruent results on fungal community responses. Fungal Ecology 9: 65-68.

Brown SP, Veach AM, Rigdon-Huss AR, Grond K, Lickteig SK, Lothamer K, Oliver AK, Jumpponen A. (2014b). Scrapping the bottom of the barrel: are rare high throughput sequences artifacts? Fungal Ecology (in review).

Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL. (2008). Microbes on mountainsides: contrasting elevational patterns of bacteria and plant diversity. Proceedings of the National Academy of Science, USA 105(s1): 11505-11511.

Cardillo M. (2011). Phylogenetic structure of mammal assembleges at large geographical scales: linking phylogenetic community ecology with macroecology. Phil. Trans. R. Soc. B 366: 2545-2553.

Cázares E, Trappe JM, Jumpponen A. (2005). Mycorrhiza-plant colonization patterns on a subalpine glacier forefront as a model system of primary succession. Mycorrhiza 15: 405-416.

Connor EF, Simberloff D. (1979). The assembly of species communities: chance or competition? Ecology 60: 1132-1140.

Cutler NA, Chaput DL, van der Gast, CJ. (2014). Long-term changes in soil microbial communities during primary succession. Soil Biology & Biochemistry 69: 359-370.

Edgar RC, Hass BJ, Clemete JC, Quince C, Knight R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27: 2194-2220.

Faith DP. (1992). Conservation evaluation and phylogenetic diversity. Biological Conservation 61: 1-10.

Fierer N, Bradford MA, Jackson RB. (2007). Toward and ecological classification of soil bacteria. Ecology 88:1354-1364.

Fierer N, Nemergut D, Knight R, Craine JM. (2010). Changes through time: integrating microorganisms into the study of succession. Research in Microbiology 161:635-642.

Garnica S, Riess K, Bauer R, Oberwinkler F, Weiß M. (2013). Phylogenetic diversity and structure of sebacinoid fungi associated with plant communities along an altitudinal gradient. FEMS Microbiology Ecology 83: 265-278.

Gihring TM, Green AJ, Schady CW. (2012). Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. Environmental Microbiology 14: 285-290.

Gotelli NJ, Entsminger GL. (2003). Swap algorithms in null model analysis. Ecology 84: 532-535.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, et al. (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443: 818-822.

Jangid K, Whiteman WB, Condron LM, Turner BL, Williams MA. (2013). Progressive and retrogressive ecosystem development coincide with soil bacterial community change in a dune system under lowland temperate rainforest in New Zealand. Plant and Soil 367: 235-247.

Jumpponen A. (2003). Soil fungal community assembly in a primary successional glacier forefront ecosystem as inferred from rDNA sequence analyses. New Phytologist 158: 569-578.

Jumpponen A, Brown SP, Trappe JM, Cázares E, Strömmer R. (2012). Twenty years of research of fungal-plant interactions on Lyman Glacier forefront – lessons learned and questions yet unanswered. Fungal Ecology 5: 430-442.

Jumpponen A, Egerton-Warburton LM. (2005). Mycorrhizal fungi in successional environments – a community assembly model incorporating host plant, environment and biotic filters. In: Dighton J, White JF, Oudemans P (eds). The Fungal Community. CRC Press, New York, pp 139-180.

Jumpponen A, Mattson K, Trappe JM, Ohtonen R. (1998). Effects of established willows on primary succession on Lyman Glacier forefront, North Cascade Range, Washington, USA: Evidence for simultaneous canopy inhibition and soil facilitation. Arctic and Alpine Research 30: 31-39.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blonberg SP, Webb CO. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26: 1463-1464.

Knelman JE, Legg TM, O'Neill SP, Washenberger CL, González A, Cleveland CC, Nemergut DR. (2012). Bacterial community structure and function change in association with colonizer plants during early primary succession on a glacier forefield. Soil Biology & Biochemistry 46: 172-180.

Knelman JE, Schmidt SK, Lynch RC, Darcy JL, Castle SC, Cleveland CC, Nemergut DR. (2014). Nutrient addition dramatically accelerates microbial community succession. PLOS one 9: e102609.

Koide RT, Fernandez C, Petprakob K. (2011). General principles in the community ecology of ectomycorrhizal fungi. Annals of Forest Science 68: 45-55.

Lozupone CA. (2007). Global patterns of bacterial diversity. Ph.D. Dissertation, University of Colorado, Boulder, Colorado. 182 pp.

Maherali H, Klironomos JN. (2007). Influence of phylogeny on fungal community assembly and ecosystem functioning. Science 316: 1746-1748.

Merckx VSFT, Hanssens SB, Hynson NA, Specht CD, Bruns TD, Smets EF. (2012). Mycoheterothophic interactions are not limited to a narrow phylogenetic range of arbuscular mycorrhizal fungi. Molecular Ecology 21: 1524-1532.

Nemergut DR, Anderson AP, Cleveland CC, Martin AP, Miller AE, Seimon A, Schmidt SK. (2007). Microbial community succession in an unvegetated, recently deglaciated soil.

Microbial Ecology 53: 110-122.

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

Philippot L, Andersson SGE, Battin TJ, Prosser JI, Schimel JP, Whitman WB, Hallin S. (2010). The ecological coherence of high bacterial taxonomic ranks. Nature Reviews Microbiology 8: 523-529.

Purschke O, Schnid BC, Sykes MT, Poschlod P, Michalski SG, Durka W, Kühn I, Winter M, Prentice HC. (2013). Contrasting changes in taxonomic, phylogenetic and functional diversity during a long-term succession: insights into assembly processes. Journal of Ecology 101: 857-866.

Quice C, Lanzén A, Curtis TP, Davenport RJ, Hall N, Head IM, Read LF, Sloan WT. (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. Nature Methods 6: 639-641.

R Development Core Team. (2007) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rämä T, Nordén J, Davey ML, Mathiassen GH, Spatafora JW, Kauserud H. (2014). Fungi ahoy! Diversity on marine wooded substrata in the high North. Fungal Ecology 8: 46-58.

Rincón A, Santamaría BP, Ocaña L, Verdú M. (2014). Structure and phylogenetic diversity of post-fire ectomycorrhizal communities of maritime pine. Mycorrhza 24: 131-141.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley, BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75: 7537-7541.

Schmidt AK, Reed AC, Nemergut DR, Grandy AS, Cleveland CC, Weintraub MN, Hill AW, Costello EK, Meyer AF, Neff JC Martin AM. (2008). The earliest stages of ecosystem succession in high-elevation (5000 metres above sea level), recently deglaciated soils. Proceedings of the Royal Society B 275: 2793-2802.

Schütte UME, Abdo Z, Foster J, Ravel J, Bunge J, Solheim B, Forney LJ. (2010). Bacterial diversity in a glacier foreland of the high Arctic. Molecular Ecology 19(s1): 54-66.

Sheneman L, Evens J, Foster JA. (2006). Clearcut: a fast implementation of relaxed neighbor joining. Bioinformatics 22: 2823-2824.

Sigler WV, Zeyer J. (2002). Microbial diversity and activity along the forefields of two receding glaciers. Microbial Ecology 43: 397-407.

Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, Saar I, Bahram M, Bechem E, Chuyong G, Kõljalg U. (2010). 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. New Phytologist 188: 291-301.

Tscherko D, Hammesfahr U, Zeltner G, Kandeler E, Böcker R. (2005). Plant succession and rhizosphere microbial communities in a recently deglaciated alpine terrain. Basic and Applied Ecology 6: 367-383.

Vamosi SM, Heard B, Vamosi JC, Webb CO. (2009). Emerging patterns in the comparative analysis of phylogenetic community structure. Molecular Ecology 18: 572-592.

Vitousek PM, Walker LR, Whiteaker LD, Mueller-Dombois D, Matson AP. (1987). Biological invasion my *Myrica faya* alters ecosystem development in Hawaii. Science 238: 802-804.

Vitousek PM, Walker LR, Whiteaker LD, Matson AP. (1993). Nutrient limitations to plant
    growth during primary succession in Hawaii Volcanoes National Park. Biogeochemistry
    23: 197-215.
Walker LR, de Moral R. (2003). Primary succession and ecosystem rehabilitation. Cambridge
    University Press, Cambridge, UK.
Wang J, Soininen J, He J, Shen J. (2012). Phylogenetic clustering increases with elevation for
    microbes. Environmental Microbiology Reports 4: 217-226.
Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naïve Bayesian classifier for rapid
    assignment of rRNA sequences into the new bacterial taxonomy. Applied and
    Environmental Microbiology 73: 5261-5267.
Webb CO. (2000). Exploring the phylogenetic structure of ecological communities: an example
    for rain forest trees. The American Naturalist 156: 145-155.
Webb CO, Ackerly DD, Kembel SW. (2008). Phylocom: software for the analysis of
    phylogenetic community structure and trait evolution. Bioinformatics 24: 2098-2100.
Zumsteg A, Luster J, Göransson H, Smittenberg RH, Brunner I, Bernasconi SM, Zeyer J, Frey B.
    (2012). Bacterial, archaeal and fungal succession in the forefield of a receding glacier.
    Microbial Ecology 63: 552-564.

Tables and Figures

Table 3.1 Proportion of samples that are significantly phylogenetically clustered compared in 1000 iterations against randomly generated trees (null model – independent swap). NTI – Nearest Taxon Index and NRI – Net Relatedness Index tested using a 2 x 2 contingency table with Fisher's Exact Test. Non-significant P-values are presented parenthetically.

	NTI (significant experimental units)	NRI (significant experimental units)	Fisher's Exact Test P-value
All Fungi	67.3%	44.2%	0.0015
Ascomycota	51.0%	57.1%	NS (0.685)
Basidiomycota	100%	100%	NS (1.00)
All Bacteria	57.1%	6.3%	<0.0001
Acidobacteria	66.7%	33.3%	0.0005
Actinobacteria	51.6%	12.9%	<0.0001
α-Proteobacteria	62.3%	0.0%	<0.0001
β-Proteobacteria	100%	40.0%	NS (0.167)
-Proteobacteria	100%	80.0%	NS (1.00)
Diazotrophs	60.0%	20.0%	NS (0.524)
Photosynthetic	40.0%	60.0%	NS (1.00)

Figure 3.1 Regression analyses of phylogenetic diversity of fungi and bacteria across the primary successional Lyman Glacier basin (distance from glacier terminus). The slope of the lines of best fit are not different from zero with distance from Lyman glacier for fungi for (a) Nearest Taxon Index (NTI); (b) Net Relatedness Index (NRI); or (c) Faith's index of phylogenetic diversity (FI). These contrast with Bacteria that show a decrease in NTI over distance (d), marginal decrease in NRI (e) and increase in FI (f) over successional distance. Regression models are inserted within each panel and significant slope and intercept estimates (* = 0.05 ≥ P > 0.001; ** = 0.001 ≥ P > 0.0001; *** = P ≤ 0.0001). The dashed line for the NTI and NRI plots represents the significance threshold for phylogenetic clustering based on 1000 Monte Carlo simulations such that points above the dashed line are significantly clustered phylogenetically whereas those below are either not different from random or phylogenetically over-dispersed.
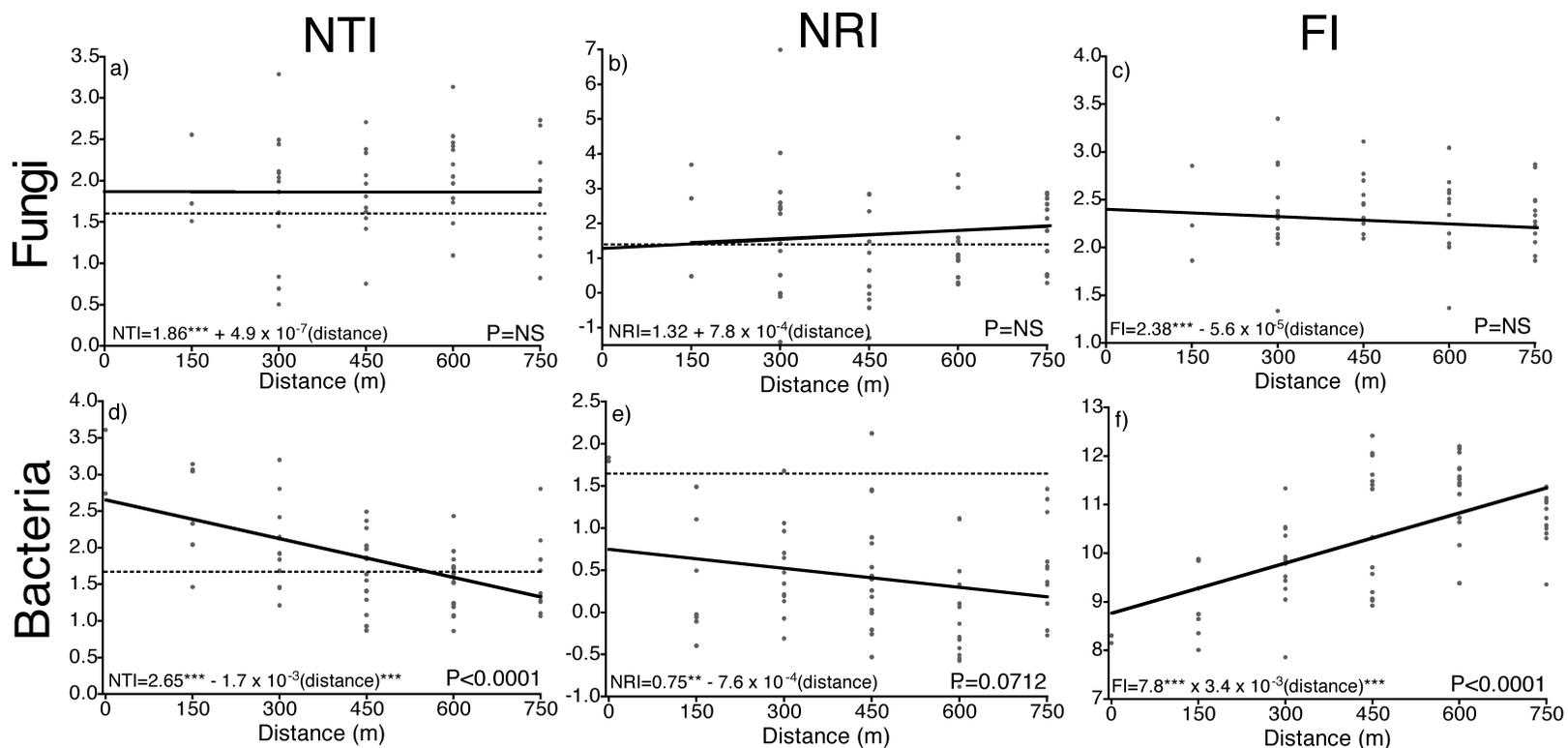
Figure 3.2 Regression analyses of phylogenetic diversity measures of the fungal phyla Ascomycota (a- NTI; b-NRI; c-FI) and Basidiomycota (d- NTI; e- NRI; f- FI) show largely unchanging phylogenetic diversity over successional age. Regression models are inserted within each panel and significant slope and intercept estimates (* = 0.05 ≥ P > 0.001; ** = 0.001 ≥ P > 0.0001; *** = P ≤ 0.0001). The dashed line for the NTI and NRI plots represents the significance threshold for phylogenetic clustering based on 1000 Monte Carlo simulations such that points above the dashed line are significantly clustered phylogenetically whereas those below are either not different from random or phylogenetically over-dispersed.
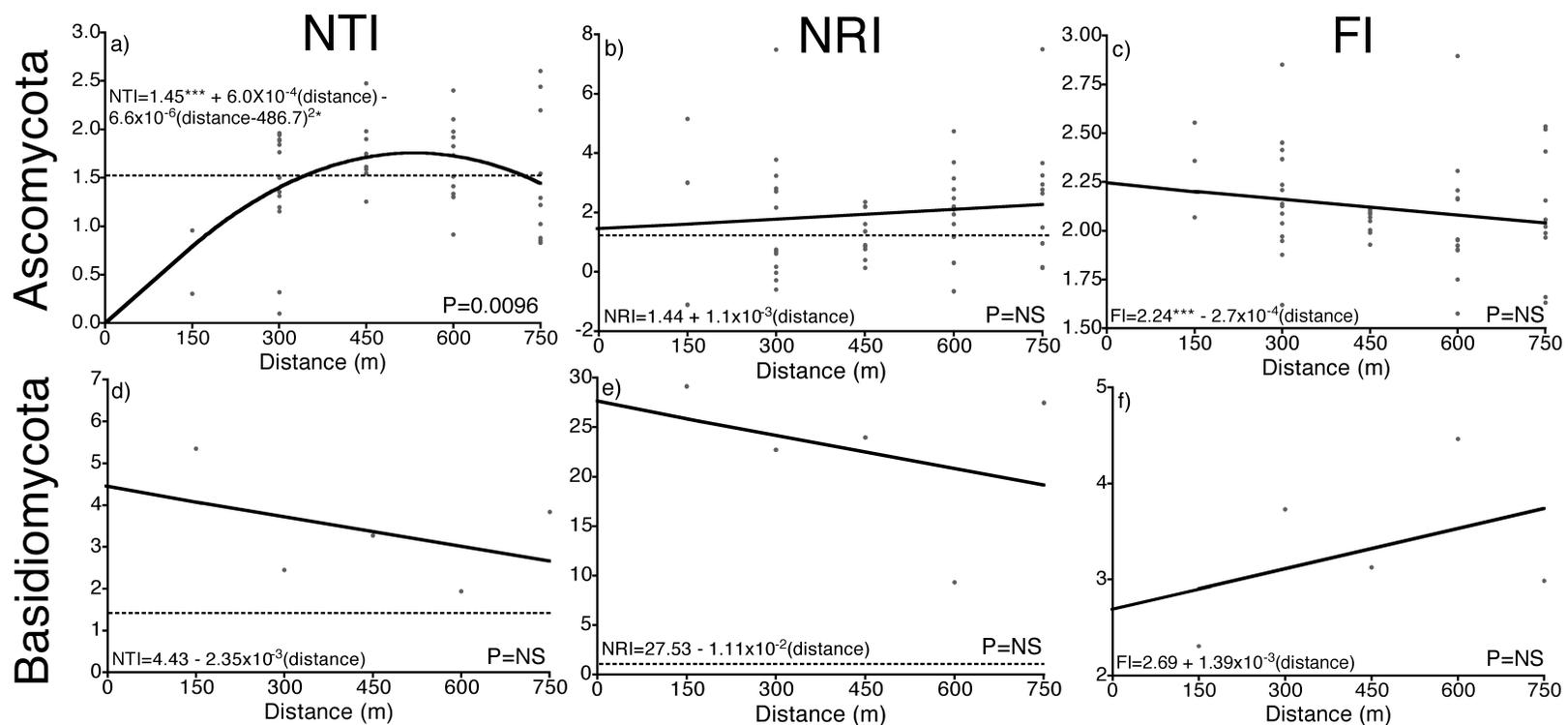
Figure 3.3 Regression analyses of phylogenetic diversity measures of bacterial phyla, classes, or functional groups with distance from glacier terminus. Shown are Acidobacteria (a- NTI; b- NRI; c- FI), Actinobacteria (d- NTI; e- NRI; f- FI), α-Proteobacteria (g- NTI; h- NRI; i- FI), β-Proteobacteria (j- NTI; k- NRI; l- FI), γ-Proteobacteria (m- NTI; n- NRI; o- FI), Diazotrophic bacteria (p- NTI; q- NRI; r- FI) and photosynthetic bacteria (s- NTI; t- NRI; u-FI). Regression models are inserted within each panel and significant slope and intercept estimates (* = $0.05 \geq P > 0.001$; ** = $0.001 \geq P > 0.0001$; *** = $P \leq 0.0001$). The dashed line for the NTI and NRI plots represents the significance threshold for phylogenetic clustering based on 1000 Monte Carlo simulations such that points above the dashed line are significantly clustered phylogenetically whereas those below are either not different from random or phylogenetically over-dispersed. Significance line (dashed line) for NTI for both β-Proteobacteria and γ-Proteobacteria are not depicted as all points fall well above threshold and are consistently clustered.

NTI | **NRI** | **FI**

Acidobacteria

a) $NTI = 3.78^{***} - 3.2 \times 10^{-3}(distance)^{***}$ $P < 0.0001$

b) $NRI = 3.98^{***} - 5.9 \times 10^{-3}(distance)^{***}$ $P < 0.0001$

c) $FI = 1.35^{***} + 2.2 \times 10^{-3}(distance)^{***}$ $P < 0.0001$

Actinobacteria

d) $NTI = 2.54^{***} - 1.4 \times 10^{-3}(distance)^{**}$ $P = 0.0001$

e) $NRI = 0.76^{*} - 8.1 \times 10^{-4}(distance)$ $P = NS$

f) $FI = 2.73^{***} + 5.9 \times 10^{-4}(distance)^{*}$ $P = 0.0116$

α-Proteobacteria

g) $NTI = 2.39^{***} - 1.4 \times 10^{-3}(distance)^{*}$ $P = 0.0019$

h) $NRI = 0.572^{*} - 1.5 \times 10^{-4}(distance)$ $P = NS$

i) $FI = 2.17^{***} + 6.0 \times 10^{-4}(distance)^{*}$ $P = 0.0004$

β-Proteobacteria

j) $NTI = 12.71^{*} - 0.01(distance)^{*} + 4.3 \times 10^{-5}(distance-450)^{2}$ $P = 0.0475$

k) $NRI = 1.54 + 2.3 \times 10^{-3}(distance) - 4.4 \times 10^{-5}(distance-450)^{2*}$ $P = 0.0160$

l) $FI = 7.98^{*} + 6.6 \times 10^{-3}(distance)^{*}$ $P = 0.0172$

γ-Proteobacteria

m) $NTI = 7.79^{*} - 5.9 \times 10^{-3}(distance)^{*} + 3.07 \times 10^{-5}(distance-450)^{2*}$ $P = 0.0079$

n) $NRI = 1.99 - 3.2 \times 10^{-3}(distance) + 4.5 \times 10^{-5}(distance-450)^{2}$ $P = NS$

o) $FI = 8.07^{*} + 4.7 \times 10^{-3}(distance)^{*} - 2.5 \times 10^{-5}(distance-450)^{2*}$ $P = 0.0105$

Diazotrophs

p) $NTI = 2.56 - 2.6 \times 10^{-3}(distance)$ $P = NS$

q) $NRI = -0.93 + 2.5 \times 10^{-3}(distance)$ $P = NS$

r) $FI = 1.54 + 1.7 \times 10^{-4}(distance)$ $P = NS$

Photosynthetic

s) $NTI = 1.55 - 1.3 \times 10^{-3}(distance)$ $P = NS$

t) $NRI = 1.79 + 9.9 \times 10^{-5}(distance)$ $P = NS^{*}$

u) $FI = 1.54^{**} + 1.8 \times 10^{-4}(distance)^{*}$ $P = 0.0396$

Distance (m)

Chapter 4 - Snow algae as the control of snow fungal community assembly[3]

Abstract

Late season snows are often colonized by psychrophilic snow algae that may provide a source of nutrients for microbes. Such late season snows are a harsh environment, but support a diverse and complex fungal community. We used culture independent methods (Illumina® MiSeq) to test if the presence of snow algae influences fungal communities in snow. We compared algae-colonized snows to adjacent (3m distant) non-colonized snows in a paired experimental design. Our data indicate that the presence of snow algae locally select for several fungi that are enriched in algae colonized snows. Although most of the fungi were basidiomycetous yeasts, our analyses identified a large number of snow-borne members of phylum Chytridiomycota. While the ecology and function of these snow Chytridiomycetes remain unclear, we hypothesize that their enrichment in the algal patches suggests that they depend on algae for nutrition. We propose three competing, non-exclusive hypotheses: saprobic, pathogenic, and mutualistic syntrophic relationships. We propose that these snow Chytridiomycetes are important components in these snow ecosystems, highlighting the under-estimation of the diversity and importance of this phylum. Taken together, our data strongly indicate that snow and cryosphere maintain fungal communities that are selected by (a-)biotic drivers that partially control fungal community assembly on seemingly homogenous solid water substrates.

Keywords: *Fungi, Snow Algae, Snow, Illumina MiSeq, Community Assembly, Snow Chytridiomycetes, Biotic Filter*

---

[3] Submitted to *FEMS Microbiology Ecology*: Brown SP, Olson BJSC, and Jumpponen A (2014). Snow algae as the control of snow fungal community assembly. *FEMS Microbiology Ecology* (in review)

Introduction

Earth's cryosphere is comprised of solid water persisting for more than one month (Fountain *et al.* 2012) and includes all snow, permafrost, sea ice, freshwater ice, glaciers and ice shelves. It plays important roles in global climate (Walsh *et al.* 2005) and terrestrial energy balance by influencing surface albedo (Prestrud 2007). With the predicted increase in the mean global annual temperature, the cryosphere is declining and becoming an endangered ecosystem (Derksen *et al.* 2012). The drastic decrease in the annual persistence of late season snow has exceeded climate model projections in the Northern Hemisphere (Derksen & Brown 2012). These changes likely impact local watershed dynamics and global biogeochemical cycles (Fountain *et al.* 2012). In addition to the consequences in local and global hydrological cycles, the cryosphere decline sets constraints on the distribution and diversity of organisms that depend on these unique habitats (Hoham & Duval 2001). The transience of the late season snow coupled with its changing volume highlights the importance of understanding the endemic biodiversity residing in these late season snow packs.

Snow packs are a harsh environment characterized by low temperatures and intense ultraviolet irradiation that may act as mutagenic stressors. Yet, there is evidence of a rich and diverse metabolically active microbial life in snow (Carpenter *et al.* 2000; Harding *et al.* 2011; Hell *et al.* 2013). These metabolically active communities contrast the commonly held perception that snow is but a passive recipient of aerially dispersed propagules that may remain viable but inactive in this substrate (Warren & Hudson 2003). This view is further highlighted by our present rudimentary understanding of microbial diversity in snow, particularly so for eukaryotic microbes. Consequently, snow can arguably be considered as a vast unexplored and undocumented ecosystem for microbial diversity (Larose *et al.* 2013).

Snow packs that persist through or linger late into summer often house complex snow-borne algal communities, frequently dominated by the red-pigmented algae *Chlamydomonas nivalis* or *Chloromonas nivalis,* both of whose taxonomies are currently unresolved. These algae often produce algal colonies on snow that are visible to the naked eye as a result of their characteristic red color. The red color of these algae is a result of the secondary carotenoid astaxanthin (Müller *et al.* 1998) and its fatty acid ester derivatives (Gorton *et al.* 2001) produced

in large quantities during their diploid, zygotic stage that is also characterized by thickened cell walls that are resistant to harsh environmental conditions including repeated freezing temperatures (Hoham 1980; Remias *et al.* 2005; Remias *et al.* 2010). During the warm season in small melt water pools, the zygotes undergo meiosis producing haploid offspring that are green, metabolically active and multiply asexually. Late in the season, when nitrogen is more limiting, these organisms develop into flagellated sexual gametes that mate producing new zygotes that can survive the next season's cold temperatures and snow (Müller *et al.* 1998).

Although the red colored colonies of *C. nivalis* may be visually dominant, snow algal communities often consist of many algal species representing Chlorophyceae (Fujii *et al.* 2010; Remias *et al.* 2010). In addition, snow colonized by algae house a broad range of fungi and bacteria that may be specifically adapted to grow in such environments (Hodson *et al.* 2008; Naff *et al.* 2013). Some evidence suggests syntrophic relationships between the snow algae and bacteria or fungi. In the most comprehensive microscopic examination to date, Weiss (1983) described the ultra-structure of the snow alga *C. nivalis* and repeatedly found encapsulated gram-negative bacteria on the surface of the zygotic resting stage in both Yellowstone National Park and the Sierra Nevada Mountains of California. Weiss (1983) posited that the microscopic observations suggested syntrophy as no similar bacteria were present in the adjacent snow without algal colonization. Similar syntrophisms have been suggested for snow algae and fungi (Kol 1968; Hoham *et al.* 1993). In such syntrophic or loose symbiotic relationships, algae-associated microorganisms may utilize dissolved organic carbon (DOC) excreted by the algae. This is indirectly supported by recent studies of *Chlamydomonas reinhardtii* suggesting that algae excrete large amounts of carbon (Find *et al.* 2012). Algae, in turn, may benefit from the "shade" provided by the microorganisms, thus buffering the algae against the harsh environmental conditions (Light & Belcher 1968; Hoham & Duval 2011; Remais *et al.* 2005). However, Kol (1968) argued that some of these fungi might simply parasitize the algae rather than form mutualistic associations.

To our knowledge, studies of snow-borne fungi and bacteria are rare and limited primarily to select fungi such as "snow molds." Generally, these "snow molds" are filamentous and thrive on organic substrates in the snow-soil interphase, but are not active in the snow itself (Robinson 2001; Matsumoto 2009). Recently, Naff *et al.* (2013) suggested that Chytridiomycetes might be abundant in snow. Additional broad inquiries of snow-borne microbial communities

indicate an abundance of microbes in snow (Harding *et al.* 2011) and suggest that these snow/ice inhabiting microbes are physiologically adapted to psychrophilic environments (Gunde-Cimerman *et al.* 2003). Our current understanding of the general ecology of fungi in snow remains rudimentary and is based primarily on soils liberated by thawing snow (de Garcia 2007) or periglacial soils (Freeman *et al.* 2009; Schmidt *et al.* 2012). Studies of fungal communities associated with snow algae are usually motivated by the great abundance of fungi observed in the course of microscopic examination of snow algae (Stein & Amundsen 1967; Kol 1968; Hoham *et al.* 1993).

Here we present the first high-throughput investigation of fungal communities associated with the snow algae by targeted ITS2 Illumina® MiSeq sequencing. The Internal Transcribed Spacer (ITS) regions are the *de facto* fungal barcode of life (Schoch *et al.* 2012). The ITS2 region in particular has been more frequently utilized because its length allows for sequencing the entire region even when using more recent sequencing technologies (Ion Torrent and Illumina) that provide relatively short reads. We utilized a community-filtering framework (Diamond 1975; Keddy 1992) and explicitly test hypotheses on algal biotic filtering of snow fungi communities (see Jumpponen & Egerton-Warburton 2005). We compared paired adjacent samples with and without visible snow algae. Specifically, we addressed following three questions: 1) do snow algae enrich the communities with specific fungi; 2) do snow algae shift fungal community composition; and, 3) will such shifts be consistent across larger geographic scales? Our results indicate fungal community filtering and enrichment of specific fungi by the snow algae. While these fungal communities appear locally and temporally stable, they differ on larger regional scales.

Materials and Methods

Sampling Sites

We sampled late season snows at six paired locations in September, 2011 and August, 2012 in the Glacier Peak Wilderness area, Wenatchee National Forest, Washington, USA (see Table 4.1 for locations and dates). Samples from each of the locations in Washington State were within 150m of each other between 2011 and 2012. Additionally, in Colorado we sampled two

paired sampling locations within Niwot Ridge Long Term Ecological Research Site (http://culter.colorado.edu/NWT) in August, 2011 and in July, 2012 at two paired sampling locations within nearby Indian Peaks Wilderness area, Arapahoe and Roosevelt National Forest. Colorado sites at Niwot Ridge LTER could not be sampled in 2012 due to lack of snow in areas where 2011 samples were collected. The sampling locations were still adjacent: the maximum distance between the 2011 and 2012 Colorado sampling locations was 5.7km. In all, we sampled all the accessible algae colonized snow patches within each respective landscape resulting in 16 paired samples (32 total). We selected sites so that the following conditions applied: 1) there was no signs of foot traffic or other anthropogenic disturbance and 2) there was an adjacent patch of uncolonized snow (based on visual assessment) within 3m from the boundary of algae colonized patch. We confirmed the absence of algae in the uncolonized snows using flow cytometry.

Sampling Protocol

Visible algal colonies and paired controls without any visible algal colonization 3m away from the edge of the colony were selected and a total of five 85cm^3 volumetric surface subsamples were collected using a steel cylinder and composited into samples representing the colonized or non-colonized snows. The pooled snow samples were placed into clean 1-gallon zip-top plastic bags and allowed to melt at ambient temperature. Once melted, one 100 ml sample was drawn with a sterile syringe (BD 30ml Syringe, Becton, Dickinson and Company, Franklin Lakes, NJ, USA) and passed through a 1.0μm Nuclepore Track-Etch Membrane filter (47mm diam) encased in a 47mm Swin-Lok Plastic Filter Holder (Whatman®, Kent, UK) to collect large cells (mainly fungi and algae). The flow through was collected into a field sterilized container (sterilized with denatured alcohol) and passed through 0.22μm Nuclepore Track-Etch Membrane filter (47mm dia) to collect bacterial cells (data not reported here). Collection receptacles and flow through collection containers were all field sterilized with denatured alcohol between sample collections. Filter holders and syringes were used only once to minimize cross contamination between samples. After filtration, filters were stored in MoBio UltraClean® Soil DNA Isolation Kit bead tubes (Carlsbad, CA, USA) with reagents S1 and IRS added to aid in sample preservation. In 2012, additional unfiltered samples were collected into sterile 15ml

falcon tubes for flow cytometric cell counts. Samples were overnight shipped to Kansas State University, where stored at -20C until further processing. Total algae counts from the 2012 samples were estimated using flow cytometry (Guava Technologies Inc. PCA-96, Hayward, CA) equipped with a 532 nm (green) excitation laser. Triplicate 10 µL samples were diluted in 250 mL 1X PBS. A combination of fluorescence emission at 675 nm (PM2, measure of chloroplast autofluorescence) and Forward Scatter (FCS) directly related to particle size were used to generate dot-plots. Using Flowing Software (v.2.5), boundaries segregating autoflorescent cells along the FCS axis were manually generated based on visual clustering. The proportions of snow algae were estimated by the proportion of counts that autofluoresced but were larger than bacteria (to discriminate against cyanobacteria). We used a paired t-test to test if the proportion of algae in the visibly colonized and uncolonized snows differed. Algae were about 35 times more abundant within the algal colonized snows that the adjacent uncolonized snows ($t = 3.25$, $P = 0.007$) representing a functionally negligible algal presence in uncolonized snows.

DNA extraction and amplicon generation

Total genomic DNA was extracted using MoBio extraction kits according to the manufacture's protocol with the following modifications: 1) filters were sonicated for 10 minutes (FS20; ThermoFisher Scientific, Waltham, MA, USA) in DNA extraction tubes with the bead solution, S1 and IRS to dislodge any cells adhering to the filters; 2) the filter was removed and two 2.4mm zirconia beads (BioSpec Products Inc., Bartlesville, OK, USA) were added into the bead tubes; 3) particles were homogenized in a FastPrep instrument (FP120; ThermoFisher Scientific, Waltham, MA, USA) at setting 4.0 for 30 seconds two times; and, 4) sterile molecular grade water was used as the elution liquid (50µL). The extracts were quantitated (ND1000 spectrophotometer; NanoDrop Technologies, Wilmington, DE) and each sample was aliquoted to a 96-well plate at 2ng µL^{-1} concentration. PCR amplicons for Illumina® MiSeq sequencing were generated using fungus specific primers to amplify the Internal Transcribed Spacer 2 (ITS2) region of the fungal rRNA gene with primers fITS7 (Ihrmark *et al.* 2012) and ITS4 (White *et al.* 1990). Unique Molecular Identifier Tags (MIDs) were connected to the ITS4 primer (MID-ITS4). MIDs were selected from the published Illumina® MID list (Caporaso *et al.* 2012) and

each MID-ITS4 combination was tested *in silico* (OligoAnalyzer 3.1; Integrated DNA Technologies, Coralville, IA, USA, http://www.idtdna.com/analyzer/Applications/OligoAnalyzer) for possible hairpins and/or primer dimers at melting temperatures above 40°. Only primers that passed this rubric were synthesized. PCR amplicons were generated in 50μL reactions under the following conditions: 1 μM forward and reverse primers, 10ng template DNA, 200 μM of each dioxynucleotide, 2.5mM $MgCl_2$, 10μL 5x Green GoTaq® Flexi Buffer (Promega, Madison, WI), 14.6μL molecular biology grade water, and 2 U GoTaq® Hot Start Polymerase (Promega, Madison, WI). PCR cycle parameters consisted of 94° initial denaturing step for 4 minutes, followed by 30 cycles of 94° for 1 minute, 54° annealing temperature for 1 minute, and 72° extension step for 2 minutes, followed by a final extension step at 72° for 10 minutes. All PCR reactions were done in triplicate to control for PCR stochasticity and negative PCR controls (sterile molecular grade water in place of DNA template) were performed throughout and remained free of contamination. Triplicate PCR products were pooled by experimental unit (total of 32) and cleaned using Agencourt® AmPure® cleanup kit using a SPRIplate 96-ring magnet (Beckman Coulter, Beverly, Massachusetts, USA) using the manufacturer's protocol except we used a 1:1 bead solution to amplicon ratio to better discriminate against small non-target DNA. Barcoded samples were equimolarly combined so that each experimental unit was equally represented. This final pool was cleaned with Agencourt® AmPure® cleanup kit once more as above. Illumina® MiSeq adaptors A and B were ligated onto the library and paired end sequenced on a MySeq Personal Sequencing System (Illumina®, San Diego, CA) using a MiSeq Reagent Kit v2, with 500 cycles. Ligation and sequencing were performed at the Integrated Genomics Facility at Kansas State University.

Sequence analysis

Sequence data were processed using mothur (v. 1.29.1; Schloss *et al.* 2009). The two obtained fastq (bidirectional reads) were contiged and the resultant fasta and qual files used as inputs for further sequence processing that we briefly describe below. We screened contiged sequences and required the following for inclusion: exact match to the MIDs (see Table S1 for complete list of MID primer sequences for each experimental unit), at most 1 bp difference in

match (wobble of one basepair to primer match) to both forward and reverse primers, and with a quality score of >= 35 over a 50 bp sliding window for the sequence. In other words, a sequence was culled if the average quality score fell below 35 for any 50 bp window (with a one bp slide) or if the sequences did not match both primers or MID. Additionally, contiged sequences were culled that had homopolymers longer than 8 bp or contained any ambiguous nucleotides. This ensured that only high-quality full length ITS2 reads remained. ITS2 sequences were truncated to 250bp for further analysis removing any conserved 5.8S regions and sequences shorter than 250bp were culled, and putative chimeras removed (chimera.uchime, UCHIME, Edgar *et al.* 2011). Remaining sequences were pair-wise aligned and the resulting distance matrix was clustered at 97% similarity using the average-neighbor algorithm. The rare OTUs (OTUs not among the 200 most abundant fungal OTUs) were eliminated. The 200 most abundant OTUs represented >97% of all sequences. Randomly selected sequences representing each of the 200 most abundant OTUs were queried (BLASTn nr/nt with the exclusion of uncultured and environmental samples) and best BLAST matches were recorded with full taxonomic string (See Table C.2). Despite the use of fungal specific primers, many most abundant OTUs belonged to non-fungal phyla: Chlorophyta (15 OTUs), Streptophyta (3 OTUs), Ciliphora (1 OTU) and Supergroup Rhizaria (1 OTU). We omitted these non-fungal OTUs and appended our analyses to include 200 fungal OTUs. Of note is that the most abundant OTU was algal (best BLASTn match to *Coenochloris* sp.) and seven times more abundant than the next most abundant OTU, suggesting that the primer bias was not adequate to discriminate against algal targets in samples highly enriched with phylum Chlorophyta. Rarefaction, richness, and diversity estimates were calculated [OTU richness = S_{obs}, Good's coverage = 1- n_1/N where n_1 is number of local singletons and N is number of sequences is a sample, complement of Simpson's Diversity = 1-D = 1- $\sum p_i^2$ (where p_i is the proportion of individuals in the ith species), and Simpson's Evenness = E_d = (1/D)/S (where S is the OTU richness at each sample and D is the Simpson's diversity index)] using an iterative approach such that each experimental unit was subsampled at a depth of 1500 sequences per experimental unit for 1000 iterations and the average of each estimator used in our analyses. Nonmetric Multidimensional Scaling (NMDS), based on a subsampled Bray-Curtis dissimilarity matrix with 1000 iterations (at 1500 subsampling depth) was used to examine fungal community composition for the first three axes (73.3% community variation, 3D stress=0.198). To determine if our Washington sites possessed different fungal communities than

our Colorado sites or if the 2011 and 2012 Washington samples differed, we used Analysis of Molecular Variance (AMOVA; PERMANOVA in Anderson 2001). Across the resolved three dimensional NMDS space, linear (Euclidian) distances were calculated between paired algae colonized and non-algae colonizes samples and these values were analyzed using Student's t-test to test if Colorado and Washington paired samples differ in their community similarities. To test for OTU enrichment between paired algae-colonized and non-colonized snow samples, we used a paired test of count data (square root + 0.5 transformed). Because of our paired design, our richness, diversity, NMDS axes scores and OTU abundance were analyzed using a nonparametric two-tailed Wilcoxon signed-rank test (H_0: $M_{algae} = M_{non-algae}$, H_1: $M_{algae} \neq M_{non-algae}$) and any significant responses were corrected for multiple comparison effects using a liberal False Discovery Rate (FDR=0.50). Data were also analyzed using a parametric paired t-test. These analyses were congruent with the non-parametric tests. As a result, we only report the non-parametric test statistics, as those rely on no assumptions on data distribution or variance homogeneity. All statistics were performed using a combination of mothur and JMP (v.10.0.2, SAS Institute, Cary, NC)

The taxon assignment of the observed OTUs to phylum Chytridiomycota (26 – or 13% – of the top 200 OTUs) was challenging as their low similarity to any vouchered or uncultured/environmental accessions. Because the low coverage and low similarity made BLASTn assignments to Chytridiomycota tentative, we further explored these data through Maximum Likelihood (PhyML) analyses. The hypervariable ITS gene regions are inappropriate for discerning higher-level relationships in phylogenetic analyses, we utilized this approach only as a method of confirming the placement of the OTUs tentatively assigned within Chytridiomycota. ITS2 reads for representatives of our chytridiomycetous environmental reads (26 OTUs) were aligned with a number of accessioned and vouchered Chytridiomycota and two closely related basal phyla (Blastocladiomycota and Monoblepharidomycota) with full length ITS2 gene regions in GenBank. Reference sequences were selected across Phylum Chytridiomycota to include all orders of Chytridiomycota (orders Cladochytriales and Polychytriales were not included because no full-length ITS2 sequences from vouchered specimens were available in GenBank at the time of analysis). Sequences were aligned using MUSCLE (1,000 iterations) as implemented in GENEIOUS (v.5.3.4, BioMatters LTD, Auckland, NZ). The alignment was trimmed manually to only include full-length ITS2 regions

and a Maximum Likelihood tree was generated (100 bootstrap iterations, substitution model = HKY85) in GENEIOUS with *Saccharomyces cerevisiae* as an outgroup.

Results

Sequence data characterization

We obtained more than 11×10^6 sequences from our MiSeq library (paired-end fastq files are deposited in Sequence Read Archive at NCBI (SRR1104197). Of these, 1,466,702 full-length ITS2 sequences matched our MIDs, forward and reverse primers (maximum primer difference=1), had no greater than 8 bp homopolymers, had average quality scores ≥35 over a 50 base pair sliding window, and were at least 250 base pairs long after trimming the primers and MIDs. After subsampling (26,157 sequences per experimental unit) prior to clustering based on unequal sequence yields, 837,024 sequences remained. After the removal of non-target, mainly algal OTUs (363,005 sequences) and OTUs that were not among the 200 most abundant, 460,161 sequences remained with unequal sequence counts per experimental unit (range 1,971-24,964). Rare OTUs were exceedingly common (3686 non singleton rare OTUs) and likely represent dormant organisms, aerially deposited spores that do not contribute to ecosystem functions in our snow samples, or artifactual OTUs with uncertain origin. For this reason, we limited our analysis to the 200 most abundant OTUs, (see Fig. C.1), representing more than 97% of all fungal sequences in our dataset. In all, the sequences represented a total of 7,835 OTUs with 3,949 global singletons. To avoid biases stemming from unequal sequence yields (Gihring *et al.* 2011), each experimental unit was subsampled iteratively to an equal depth of 1,500 sequences for estimating the richness and diversity metrics as well as for the ordination analyses.

Taxonomic distribution

The fungal communities were dominated by Basidiomycota (see Table C.2 for complete taxonomic assignments for the 200 most abundant OTUs). Sequence and OTUs counts of the most abundant Orders and Families with possible ecological roles are given in Table 4.2. The Basidiomycetes were dominated by cosmopolitan polyphyletic yeasts in the genus *Rhodotorula*

(46 OTUs – 23% of all OTUs and 255,911 sequences – 53.30% of total sequences), a distribution not uncommon in other snow and glacier fungi surveys (de Garcia et al., 2007; de Garcia et al., 2012). In fact, the four most abundant OTUs were classified as *Rhodotorula* and the next most abundant OTU was *Cryptococcus saitoi*, a commonly encountered snow dwelling yeast with uncertain ecology. OTUs assigned to Lyophyllaceae were common (8 OTUs) and included two OTUs with close BLASTn affinities to accessioned sequences of *Asterophora* – a genus with members known to be parasites of fungi. The remaining six OTUs in the family were only marginally similar to genus *Lyophyllum*, described as a soil borne saprobic or mycorrhizal macrofungi. The presence of these macrofungi in snow is unlikely and may be due to allochthonous introduction of spores from surrounding forests, or alternatively, these may not represent macrofungi at all, rather these OTUs may be undescribed/unknown taxa with unknown ecologies considering the low sequence similarity to previously accessioned sequences in the queried databases. Chytridiomycetes were a surprisingly high proportion of OTUs (26 OTUs – 13%; ~10% of total sequences) in our environmental sequencing. Of these OTUs none had high similarity to any accessioned sequences based on BLASTn alignments (coverage ranged from 15% to 60% with very low total BLAST scores). Because of the low similarity for known accessioned Chytridiomycota, we re-analyzed these data in queries that also included uncultured/environmental sequences. These analyses consistently failed to match closely any accessioned sequences. Despite their low similarity to sequences in the combined global genetic databases, our Maximum Likelihood (PhyML) analysis confirmed placement of the 26 OTUs with 99% bootstrap support within Chytridiomycota. Additionally, our environmental OTUs were closely related but distinct to soil borne Chytridiomycetous Order Lubulomycetales (Fig. C.2). Yet, bootstrap support within the environmental Chytridiomycota clade was low a likely result of using hypervariable ITS2 that is an inappropriate region for delineating higher-level phylogenies. As a result, the placement of these OTUs below phylum cannot be deduced from our PhyML analysis. Yet, most of our snow Chytridiomycete OTUs form a distinct clade suggesting that these taxa may represent a monophyletic group of snow-borne Chytridiomycetes with very little ITS2 sequence similarity to anything known.

Richness, diversity, evenness

Coverage and rarefaction (see Fig. C.3) estimators indicated that the fungal richness was adequately captured (Good's Coverage for algae colonized and non-colonized snow 0.972 ± 0.003 and 0.986 ± 0.001, respectively). Observed OTU richness in the algae colonized snow (74.98 ± 11.72; mean $\pm$ 1SD) was greater than in the adjacent, non-colonized snow (64.54 ± 11.72) across both sites and years (Fig. 4.1; paired two-tailed Wilcoxon Sign-Rank test; W=58, P=0.0013). In contrast, neither diversity (complement of Simpson's Diversity, 1-D: 0.890 ± 0.051 for the algae colonized snow and 0.865 ± 0.067 for non-colonized snow; W=30, P=0.130) nor evenness (E_D: 0.146 ± 0.053 for algae-colonized snow and 0.145 ± 0.073 for non-colonized snow; W=10, P=0.632) differed between the colonized and non-colonized snows.

Community differences

The fungal communities were resolved optimally on three NMDS axes (stress=0.198, r^2=0.733). Our community-wide AMOVA (PERMANOVA in Anderson 2001) analyses using the Bray-Curtis dissimilarity matrix indicated regional (Colorado vs. Washington) but not temporal differences in the snow fungal communities. These analyses failed to distinguish fungal communities in the colonized and non-colonized snow ($F_{1,30}$=1.247, P=0.227) on the whole. However, paired analyses of the Axis 2 loading scores (representing 57.16% of community variability) indicated that fungal communities in the colonized snow were distinct from those in the non-colonized snow (Fig. 4.2; W = 48, P = 0.011). In contrast, neither Axis 1 (r^2=6.38%) nor Axis 3 (r^2 = 9.85%) distinguished between the paired algal and non-algal snows (W=21, P=0.298; W=23, P=0.252, respectively). Further, our AMOVA suggested no difference across years ($F_{1,30}$=1.135, P=0.286) in the Washington sites but clearly distinguished the snow inhabiting fungal communities from Colorado and Washington ($F_{1,30}$ = 8.654, P<0.001). These analyses suggest that although our sites located in Colorado and Washington are distinct, the communities remain fairly stable over time – likely as a result of local fungal propagule inputs.

We also analyzed the Euclidian distances between paired colonized and non-colonized snows as a way to provide a metric of how different the paired samples were. The Euclidian distances across the three resolved ordination axes between paired algae-colonized and non-colonized snow-fungal communities was greater in the Washington snows than in Colorado (t=2.47, P=0.0267; Fig. C.4). This observed difference may be the result of a mitigated shift in

the fungal community associated with sampling locations. For instance, Colorado snow samples were at very high elevations (~3500 m) at or above the tree line whereas Washington snows were much lower in elevation (~1900 m), likely resulting in a stronger physiological constraint. Taken together, this indicates biotic community filtering coinciding with snow algae but the directionality of these community shifts is not conserved over long distances.

In all, 6.5% or a total of 13 of the 200 most abundant OTUs were enriched in algae-colonized snow (see Table 4.3), whereas none were enriched in the non-colonized snow as determined by Wilcoxon Sign-Rank analyses after correction for multiple comparisons using a liberal false discovery rate of 0.5. An additional 3 OTUs differed between algal colonized snow and non-algal colonized snow but were no longer significant after controlling for multiple comparisons. We expected that these enriched OTUs would provide the most valuable clues on the ecology of the fungi that inhabit the snow colonized by algae. We initially hypothesized that the algal cells and/or their nutrient-rich exudates provide substrates that potentially facilitate syntrophy or fungi that are opportunistic saprobes or algal pathogens. The colonized snow was enriched for saprobic and putatively pathogenic OTUs. Several *Rhodotorula* OTUs were enriched in the algal colonized snows suggesting the opportunistic utilization of increased organic matter associated with these snow algae. Particularly interesting are OTUs 9, 37, and 40 that - based on our PhyML analyses - represent novel Chytridiomycetes whose functions remain unknown (Fig. C.2). Also of note are OTUs 48 and 163, which are enriched in algal colonized snow but are extremely dissimilar to any known accessioned fungi (see Table C.2 for full BLAST scores). Given the great dissimilarity to any accessioned taxa, these OTUs most likely represent novel taxa and/or taxa that are underrepresented in the global nucleotide repositories for our locus of interest. Thus, further and more detailed investigation is needed to better understand the fungal communities in this environment. Best BLASTn matches identify these OTUs within the families Boletaceae (ectomycorrhizal) and Lyophyllaceae (saprobic or parasitic) respectively. Thriving ectomycorrhizal fungi are unlikely in alpine snow without host plants; the presence of such taxa most probably represents allochthonous deposition of spores and highlights the caveats of analyzing relatively rare OTUs or inferring ecologies of OTUs on the basis of poor BLAST matches. Because an OTU is present and differs in frequency across treatments, does not mean that it is influential in the ecosystem. Thus, reported results using rare OTUs should be taken with caution.

Discussion

We present the first study on how snow algae may influence the fungal community assembly and one of the very first deep-sequencing studies of snow-borne fungi. The late season snow-packs are a declining ecosystem; climate change predictions suggest that the earth's cryosphere will dramatically decline in volume (Derksen & Brown 2012). As a result, assessment of the endemic biodiversity in these "endangered" ecosystems is timely and critical. Snow has often been viewed as a passive collector of allochthonous propagules that maintains little if any biological activity. This view has all but vanished in recent years (Carpenter *et al.,* 2000; Harding *et al.,* 2011; Hell *et al.,* 2013). Yet, there is a dearth of knowledge about snow-borne life or microbial communities and their function in snow, particularly for fungi. Our analyses suggest that the algae-colonized snow selects fungal communities from the local propagule pool and enriches the community for fungi that are favored by the substrate. These results provide evidence for a broad biotic community filtering by snow algae.

Our data indicate that snow supports diverse fungal community dominated by Basidiomycetous yeasts whose ecologies and taxonomies are poorly understood. Yeast dominated systems have been reported on glacier surfaces (de Garcia *et al*. 2012) and in periglacial soils (Schmidt *et al*. 2012, Brown & Jumpponen 2014). In our analyses, the most common OTU was assigned to genus *Rhodotorula* (46 OTUs in all; 6 of which are enriched an algal snows). These yeasts are polyphyletic, understudied, and their generic delineation is historically morphological, shown not to recapitulate inferred phylogenies (Toome *et al*. 2013). Additionally, many OTUs were placed into taxa grouped as black meristematic fungi (BMF), another polyphyletic grouping based on anamorphic phenotypes. Although BMFs are overall poorly understood, they are often suspected to play a large role in mineral transformations and often are resilient in harsh environments (Onofri *et al*. 2007). These cosmopolitan BMFs likely utilize allochthonous organic matter such as wind-blown particulate matter common on the snow surface. These results reiterate that there is a dearth of information on psychrophilic/tolerant fungi and highlights the importance for future studies into these systems. Our analyses do not permit assessing whether these yeasts or other observed fungi are metabolically active in this substrate. Thus, further investigation into the activity of these snow fungi is required to gain a more complete picture of these fungal communities.

The high abundance of Chytridiomycetes in our study was surprising. This suggests that these often enigmatic fungi may be more abundant and diverse than previously thought. We have just recently begun to appreciate the hidden diversity of Chytridiomycetes. Freeman *et al*. (2009) demonstrated that Chytridiomycetes may dominate high altitude periglacial soils. The snow Chytridiomycetes in this study were highly abundant and dissimilar to any sequences accessioned to the nucleotide repositories. Our confirmatory phylogenetic analyses suggested that these fungi likely represent a novel clade of snow Chytridiomycetes (Fig. C.2). Placement of these novel Chytridiomycetes from this study at levels below Phylum remains uncertain but they may belong to the early divergent snow Chytridiomycetes identified as 'Snow Clade 1' or 'Snow Clade 2' from North American and European snows (Naff *et al*. 2013). However, this cannot be determined because different rRNA gene regions were used in these two studies. Naff *et al*. (2013) posited that these snow Chytridiomycetes may be parasitic to snow algae because they were common in clone libraries from algal snows. However, these hypotheses were neither explicitly tested, nor is parasitism the only reasonable nutritional hypothesis. The present study also differs from Naff *et al.* (2013) in a very important way, Naff and co-authors only collected snow that was colonized by snow algae and sequenced shallowly, whereas we test a community filtering utilizing paired adjacent samples (with and without colonization). Given that Chytridiomycetes also were abundant in the snows free of algae, these Chytridiomycetes may also possess saprobic or syntrophic life strategies. As a result, further and more detailed investigation of the snow Chytridiomycete ecology is needed to better understand their function in the cryosphere. Nevertheless, to capture a high abundance of Chytridiomycetes is striking; Chytridiomycetes tend to be infrequent in locus-targeted community sequencing studies. This may be a result of the highly divergent and difficult to amplify ITS regions of these basal fungi (Schoch *et al.* 2012). Thus, even the relatively high estimates of the Chytridiomycete abundance observed here may be an underrepresentation. Relying on OTU abundance as a proxy for organismal abundance may not result in a 1:1 relationship. Different fungal lineages may be under/overrepresented in sequence data due to a myriad of factors including primer bias and differential ITS copy number (Pukkila *et al.* 1993; Amend *et al.* 2010; Porter & Golding 2012). Yet, despite the potential poor amplification of Chytridiomycetes, we found that Chytrids in high

abundance and diverse, highlighting that Chytridiomycota is both common and important in snows.

It is tempting to speculate on whether these Chytridiomycetes parasitize or are pathogenic to snow algae (Naff *et al.* 2013) as such associations are common in algae dominated freshwater systems (Hoffman *et al,* 2008; Gutman *et al.* 2009; Rasconi *et al.* 2011). These snow Chytridiomycetes may also act as facultative mutualists or have an obligate syntrophic relationships; there is a precedence of such relationships in other aquatic systems (Picard *et al.* 2013). Although our data suggest the enrichment of these communities with such fungi, they do not allow specific statements about the ecology or the life strategies of the detected fungi. However, it is most likely that these novel Chytridiomycetes are major players in snow-borne fungal communities. It is also clear that snow fungi are a product of establishment from local propagule pools as the snow-borne fungal communities were compositionally distinct in Washington and Colorado (however, there are many OTUs that were found in all samples including two Chytridiomycetes). In contrast, both locations had stable fungal communities over two sampling years. It is probable that snow fungi initially establish from local propagules and the presence of snow algae facilitate their growth and metabolic activity, effectively acting as a biotic fungal community filter. Our data strongly suggest that snow is not a homogeneous passive recipient of arriving microbial propagules, but a substrate that selects for communities that maintain metabolic activity. Many small-scale biotic and/or abiotic factors that vary spatially have a substantial potential to select cryotolerant communities differing in their ecological and functional attributes.

Conclusions

Overall, our results indicate that snow algae act as an environmental filter altering fungal community assembly and resultant community composition. This community filtering is potentially facilitated by enrichment of saprobic and pathogenic fungi that are able to utilize snow alga directly of indirectly through their exudates. Alternatively, the enrichment of specific fungal community constituents may be an outcome of syntrophic associations between algae and fungi that are engaged in loose symbioses. Further in depth studies on the life history strategies and ecology of the snow-inhabiting fungi is mandatory to shed light into these unresolved

questions. Interestingly and congruently with studies by others (Naff *et al*. 2013), our data identified potentially novel groups of Chytridiomycetes, some of which are enriched in algae colonized snow and of undetermined functions. From these studies, it is clear that we are barely scratching the surface of the nearly unexplored cryosphere. To put it simply and boldly, snow is not a simple and passive receptacle or storage of deposited propagules, but an ecosystem that maintains unique communities that may vanish with the declining cryosphere before we have an opportunity to understand them.

Acknowledgements

References

Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Molecular Ecology* **19**, 5555-5565.

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecology* **26**, 32-46.

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

Carpenter EJ, Lin SJ, Capone DG (2000) Bacterial activity in South Pole snow. *Applied and Environmental Microbiology* **66**, 4514-4517.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal* **6**, 1621-1624.

de Garcia V, Brizzio S, Libkind D, Buzzini P, Broock M (2007) Biodiversity of cold-adapted yeasts from glacial meltwater rivers in Patagonia, Argentina. *FEMS Microbiology Ecology* **59**, 331-341.

de Garcia V, Brizzio S, Broock M (2012) Yeasts from glacial ice of Patagonian Andes, Argentina. *FEMS Microbiology Ecology* **82**, 540-550.

Derksen C, Brown R (2012) Spring snow cover extent reductions in the 2008-2012 period exceeding climate model projections. *Geophysical Research Letters* **39**, L19504, doi:10.1029/2012GL053387.

Derksen C, Smith SL, Sharp M, Brown L, Howell S, Copland L, Mueller DR, Gauthier Y, Fletcher CG, Tivy A, Bernier M, Bourgeois J, Brown R, Burn CR, Duguay C, Kushner P, Langlois A, Lewkowicz AG, Royer A, Walker A (2012) Variability and change in the Canadian cryosphere. *Climate Change* **115**, 59-88.

Diamond JM (1975) Assembly of species in communities. In *Ecology and Evolution of Communities* (eds Cody ML, Diamond JM), pp. 342-344. Harvard University Press, Cambridge, MA.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioimformatics* **27**, 2194-2200.

Fountain AG, Campbell JL, Schuur EAG, Stammerjohn SE, Williams MW, Ducklow HW (2012) The disappearing cryosphere: Impacts and ecosystem responses to rapid cryosphere loss. *BioScience* **62**, 405-415.

Freeman KR, Martin AP, Karki D, Lynch RC, Mitter MS, Meyer AF, Longcore JE, Simmons DE, Schmidt SK (2009) Evidence that chytrids dominate fungal communities in high-elevation soils. *Proc. Natl. Acad. Sci. USA* **106**, 18315-18320.

Fujii M, Takano Y, Kojima H, Hoshino T, Tanaka R, Fukui M (2012) Microbial community structure, pigment composition, and nitrogen source of red snow in Antarctica. *Microbial Ecology* **59**, 466-475.

Gihring TM, Zhang GX, Brandt CC, Brooks SC, Campbell JH, Carroll S, Criddle CS, Green SJ, Jardine P, Kostka JE, Lowe K, Mehlhorn TL, Overholt W, Watson DB, Yang ZM, Wu WM, Schadt CW (2011) A limited microbial consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. *Applied and Environmental Microbiology* **77**, 5955-5965.

Gorton HL, William WE, Vogelmann TC (2001) The light environment and cellular optics of the snow alga *Chlamydomonas nivalis* (Bauer) Wille. *Photochemistry and Photobiology* **73**, 611-620.

Gunde-Cimerman N, Sonjak A, Zalar P, Frisvad JC, Diderichsen B, Plemenitaš A (2003) Extremophilic fungi in arctic ice: a relationship between adaptation to low temperature and water activity. *Physics and Chemistry of the Earth* **28**, 1273-1278.

Gutman J, Zarka A, Boussiba S (2009) The host-range of *Paraphysoderma sedebokerensis,* a chytrid that infects *Heamatococcus pluvialis*. *European Journal of Phycology* **44**, 509-514.

Harding T, Jungblut AS, Lovejoy C, Vincent WF (2011) Microbes in high arctic snow and implications for the cold biosphere. *Applied and Environmental Research* **77**, 3234-3243.

Hell K, Edwards A, Zarsky J, Posmirseg SM, Girdwood S, Pachebat JA, Insam H, Sattler B (2013) The dynamic bacterial communities of a melting High Arctic glacier snowpack. *The ISME Journal* **7**, 1814-1826.

Hoffman Y, Aflalo C, Zarka A, Gutman J, James TY, Boussiba S (2008) Isolation and characterization of a novel chytrid species (phylum Blastocladiomycota) parasitic on the green alga *Hearmatococcus*. *Mycological Research* **112**, 70-81.

Hodson A, Anesio AM, Tranter M, Fountain A, Osborn M, Priscu J, Laybourn-Parry J, Sattler B (2008) Glacier Ecosystems. *Ecological Monographs* **78**, 41-67.

Hoham RW (1980) Unicellular chlorophytes – snow algae. In: *Phytoflagellates* (ed Cox ER), Elsevier-North, Holland.

Hoham RW, Laursen AE, Clive SO, Duval B (1993) Snow algae and other microbes in several alpine areas in New England. In: *Proc. 50th Annual Eastern Snow Conference*. pp 165-173.

Hoham RW, Duval B (2001) Microbial ecology of snow and freshwater ice with emphasis on snow algae. In: *Snow Ecology* (eds Johns HG, Pomeroy JW, Walker DA, Hoham RS), pp. 186-203. Cambridge University Press, Cambridge, UK.

Ihrmark K, Bodecker ITM, Cruz-Martinez K, Friberg H, Kubartova A, Schenck J, Strid Y, Stenlid J, Brandstrom-Durling M, Clemmensen KE, Lindahl BD (2012) New primers to amplify the fungal ITS2 region – evaluation by 454-sequencing or artificial and natural communities. *FEMS Microbiology Ecology* **82**, 666-677.

Jumpponen A. and Egerton-Warburton L. (2005) Mycorrhizal fungi in successional environments – a community assembly model incorporating host plant, environmental and biotic filters. In: *The Fungal Community* (eds Dighton J, White JF, Oudemans P), pp. 139-180. CRC Press, New York.

Keddy PA (1992) Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetative Science* **3**,157–164.

Kind T, Meissen JK, Yang D, Nocito F, Vaniya A, Cheng Y, VanderGheynst JS, Fiehn O (2012) Qualitative analysis of algal secretions with multiple mass spectrometric platforms. *Journal of Chromatography A* **1244**, 139-147.

Kol E (1968) *Kryobiologie, Biologie und Limnologie des Schnees and Eises 1. Kryovegetation*. Stuttgart, Germany: E. Schweizerbart'sche Verlagsbuchhandlung.

Larose C, Dommergue A, Vogel TM (2013) The dynamic arctic snow pack: an unexplored environment for microbial diversity and activity. *Biology* **2**, 317-330.

Light JJ, Belcher JH (1968) A snow microflora in the Cairngorm Mountains, Scotland. *British Phycology Bulletin* **3**, 471-473.

Matsumoto N (2009) Snow molds: a group of fungi that prevail under snow. *Microbes and Environments* **24**, 14-20.

Müller T, Bleiß W, Martin CD, Rogaschewski S, Fuhr G (1998) Snow algae from northwest Svalbard: their identification, distribution, pigment and nutrient content. *Polar Biology* **20**, 14-32.

Naff CS, Darcy JL, Schmidt SK (2013) Phylogeny and biogeography of an uncultured clade of snow chytrids. *Environmental Microbiology* **15**, 2672-2680.

Onofri S, Selbmann L, de Hoog GS, Grube M, Barreca D, Suisi S, Zucconi L (2007) Evolution and adaptation of fungi at boundaries of life. *Advances in Space Research* **40**, 1657-1664.

Picard KT, Letcher PM, Powell MJ (2013) Evidence for a facultative mutualist nutritional relationship between the green coccoid alga *Bracteacoccus* sp. (Chlorophyceae) and the zoosporic fungus *Rhizidium phycophilum* (Chytridiomycota). *Fungal Biology* **117**, 319-328.

Porter TM, Golding CB (2012) Factors that affect Large Subunit Ribosomal DNA amplicon sequencing studies of fungal communities: classification method, primer choice, and error. *PLoS ONE* **7**, e35749.

Prestrud P (2007) Why are snow and ice important to us? Chapter 2. In: *UNEP Global Outlook for Ice & Snow* (ed Eamer J), pp. 19-28. United Nations Environment Program.

Pukkila PJ, Skrzynia C (1993) Frequent changes in the number of reiterated ribosomal RNA genes throughout the life cycle of the Basidiomycete *Coprinus cinereus*. *Genetics* **133**, 203-211.

Rasconi S, Jobard M, Sime-Ngando T (2011) Parasitic fungi of phytoplankton: ecological roles and implications for microbial food webs. *Aquatic Microbial Ecology* **62**, 123-137.

Remias D, Karsten U, Lütz C, Leya T (2010) Physiological and morphological processes in the Alpine snow alga *Chloromonas nivalis* (Chlorophyceae) during cyst formation. *Protoplasma* **243**, 73-86.

Remias D, Lütz-Meindl U, Lütz C (2005) Photosynthesis, pigments and ultrastructure of the alpine snow alga *Chlamydomonas nivalis*. *European Journal of Phycology* **40**, 259-268.

Robinson CH (2001) Cold adaptation in Arctic and Antarctic fungi. *New Phytologist* **151**, 341-353.

Sampaio JS, Gadanho M, Bauer R, Weiß M (2003) Taxonomic studies in the Microbotryomycetidae: *Leucosporidium golubevii* sp. nov., *Leucosporidiella* gen. nov. and the new orders Leucosporidiales and Sporidiobolales. *Mycological Progress* **2**, 53-68.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: Open-source, platform independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**, 7537-7541.

Schmidt SK, Naff CS and Lynch RC (2012) Fungal communities at the edge: ecological lessons from high alpine fungi. *Fungal Ecology* **5**, 443-452.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Science, USA* **109**, 6241-6246.

Stein JR, Amundsen CC (1967) Studies on snow algae and fungi from the front range of Colorado. *Canadian Journal of Botany* **45**, 2033-2045.

Toome M, Roberson RW, Aime MC (2013) *Metedithblackwellia eburnea* gen. et sp. nov., Kriegeriaceae fam. nov. and Kriegeriales ord. nov.-toward resolving higher-level classification in Microbotryomycetes. *Mycologia* **105**, 486-495.

Walsh JE, Anisimov O, Hagen J, Jakobsson T, Oerlemans J, Prowse TD, Romanovsky V, Savelieva N, Serreze M, Shiklomanov I, Solomon S (2005) Cryosphere and hydrology. In: *Arctic Climate Impact Assessment* pp.183-242. Cambridge University Press, New York.

Warren SG, Hudson SR (2003) Bacterial activity in South Pole snow is questionable. *Applied and Environmental Microbiology* **69**, 6340-6341.

Weiss RL (1983) Fine-structure of the snow algal (*Chlamydomonas nivalis*) and associated bacteria. *Journal of Phycology* **19**, 200-204.

White TJ, Bruns T, Lee S. Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications* (eds Innis MA, Gelfand DH, Sininsky JJ, White TJ). Academic Press, Inc., New York.

Tables and Figures

Table 4.1 Sampling locations of paired algae-colonized and uncolonized snows across consecutive years. WA=Washington State, CO= Colorado. All Washington sampling locations were collected at or near the Lyman Glacier Basin.

State	Site	Landmark	Latitude	Longitude	Elevation (m)
2011					
CO	Niwot 1	Near Soddie Laboratory	40° 02' 56'' N	105° 34' 51'' W	3368
CO	Niwot 2	Saddle	40° 03' 30'' N	105° 35' 20'' W	3514
WA	Glacier Peak 1	Cloudy Pass	48° 12' 09'' N	120° 55' 28'' W	1961
WA	Glacier Peak 2	Terminal Moraine	48° 10' 59'' N	120° 54' 11'' W	1802
WA	Glacier Peak 3	Lyman Glacier	48° 10' 21'' N	120° 53' 50'' W	1880
WA	Glacier Peak 4	Spider Gap N	48° 10' 14'' N	120° 52' 55'' W	2135
WA	Glacier Peak 5	Spider Gap S	48° 10' 10'' N	120° 52' 53'' W	2123
WA	Glacier Peak 6	Lower Spider Snowfield	48° 09' 42'' N	120° 52' 42'' W	1897
2012					
CO	Indian Peaks 1	E of Shoshoni Peak	40° 04' 02'' N	105° 37' 44'' W	3407
CO	Indian Peaks 2	S Shore Lake Isabelle	40° 04' 01'' N	105° 04' 01'' W	3358
WA	Glacier Peak 1	Cloudy Pass	48° 12' 10'' N	120° 55' 27'' W	1966
WA	Glacier Peak 2	Terminal Moraine	48° 10' 58'' N	120° 54' 11'' W	1794
WA	Glacier Peak 3	Lyman Glacier	48° 10' 24'' N	120° 53' 49'' W	1866
WA	Glacier Peak 4	Spider Gap N	48° 10' 14'' N	120° 52' 55'' W	2173
WA	Glacier Peak 5	Spider Gap S	48° 10' 10'' N	120° 52' 53'' W	2137
WA	Glacier Peak 6	Lower Spider Snowfield	48° 09' 41'' N	120° 52' 35'' W	1893

Table 4.2 Taxonomic distributions of abundant Operational Taxonomic Units (OTUs). Frequencies of sequences that could be classified to Phyla, the most abundant Orders, and representative Families (Genus where *Incertae sedis* at the Family level are represented parenthetically). Taxonomic representations of OTUs based on best BLASTn matched across accessioned fungi deposited in GenBank. Purported ecologies at the Family (Genus) level are also reported. Number of OTUs of the 200 most abundant are given.

Phylum	Order	Families (Genus)	Ecology	Percentage of Sequences	Number of OTUs
Ascomycota				12.62%	67
	Dothideales			5.92%	13
		Dothiorceae	Biotrophic or Necrotrophic	4.88%	8
		(*Celosporium*)	Uncertain (BFM)	1.03%	5
	Chaetothyriales			2.31%	12
		Herpotrichiellaceae		0.98%	8
		(*Sarcinomyces*)	Uncertain (BFM)	1.33%	4
	Pleosporales			1.54%	7
		Pleosporaceae	Necrotrophic or Saprobic	1.45%	5
Basal Fungal Lineages				1.13%	6
	Mucorales			0.73%	2
		(*Umbelopsis*)		0.73%	2
	Mortierellales			0.40%	4
		Mortierellaceae	Saprobic	0.40%	4
Basidiomycota				77.36%	99
	Incertae sedis			32.31%	15
		(*Rhodotorula*)	Saprobic/(Pathogen?)	32.31%	15
	Kriegariales			20.50%	17
		Kriegeriaceae	Saprobic/(Pathogen?)	20.50%	17
	Leucosporidiales			7.38%	9
		(*Leucosporidiella*)	Uncertain (non-phytoparasitic)[1]	4.16%	4
		(*Leucosporidium*)	Uncertain (non-phytoparasitic)[1]	3.23%	5
Chytridiomycota‡				9.48%	26
	Rhizophydiales			6.69%	10
		Incertae sedis	Uncertain	4.33%	2
		Rhizophydiaceae	Saprobic	2.31%	7
	Polychytriales			1.10%	4
		(*Polychytrium*)	Saprobic/(Pathogen?)	1.10%	4
	Incertae sedis			0.20%	11
Glomeromycota‡				0.54%	2
	Glomerales			0.54%	2
		Glomeraceae		0.54%	2

‡: Best BLASTn match to phyla Chytridiomycota and Glomeromycota are extremely dissimilar to any accessioned taxa (Query Coverage ≤ 25% and BLAST score ≤ 90; see Table S2). Thriving Glomeromycetes are unlikely in absence of host, thus these likely represent unknown taxa.
BFM: Black Meristematic Fungi, polyphyletic group, primarily anamorphic, known to be resistant to harsh environments
1: These genera are defined as non-phytoparasitic (Sampaio *et al.* 2003) but ecologies remain uncertain.

Table 4.3 Fungal Operational Taxonomic Units that are enriched in algal colonized snow compared to paired non-algal colonized snow based on Wilcoxon Sign-Rank test after correction for multiple comparisons. Best BLASTn matches and putative Ecologies are also reported (EcM=Ectomycorrhizal). The symbol '‡' represents taxa whose best BLASTn match are extremely dissimilar to any accessioned taxa (Query Coverage ≤ 25% and BLAST score ≤ 90; see Table S2) that are likely novel fungal taxa whose ecologies remain uncertain.

OTU Number	Wilcoxon Sign-Rank Test Statistic W	P |W|	Ecology	Best BLASTn Match
2	47	0.0126	Saprobic	*Rhodotorula sp.*
4	41.5	0.03	Saprobic	*Rhodotorula psychrophenolica*
9	43	0.0248	Unknown	Rhizophydiales sp. ‡
37	30	0.0059	Unknown	Chytridiomycota sp. ‡
40	14	0.0156	Unknown	Chytridiomycota sp. ‡
43	23.5	0.0156	Saprobic	*Rhodotorula sp.*
45	32	0.0103	Saprobic	*Rhodotorula sp.*
48	10.5	0.0313	EcM (‡Unknown)	*Tylopilus formosus* ‡
59	47	0.0122	Pathogenic	*Ilyonectria macrodidyma*
119	22.5	0.0234	Saprobic	*Rhodotorula sp.*
163	14	0.0156	Saprobic (‡Unknown)	*Lyophyllum sp.* ‡
195	14	0.0156	Saprobic	*Rhodotorula sp.*
199	10.5	0.0313	Saprobic	*Leucosporidium scottii*

Figure 4.1 OTU richness is higher in algae colonized snow than in adjacent paired uncolonized snow (paired Wilcoxon Sign-Rank test). Uncolonized fungal OTU richness estimates (1000 iterations) are solid symbols and algae colonized richness estimates are open symbols. Dashed lines connect paired samples.
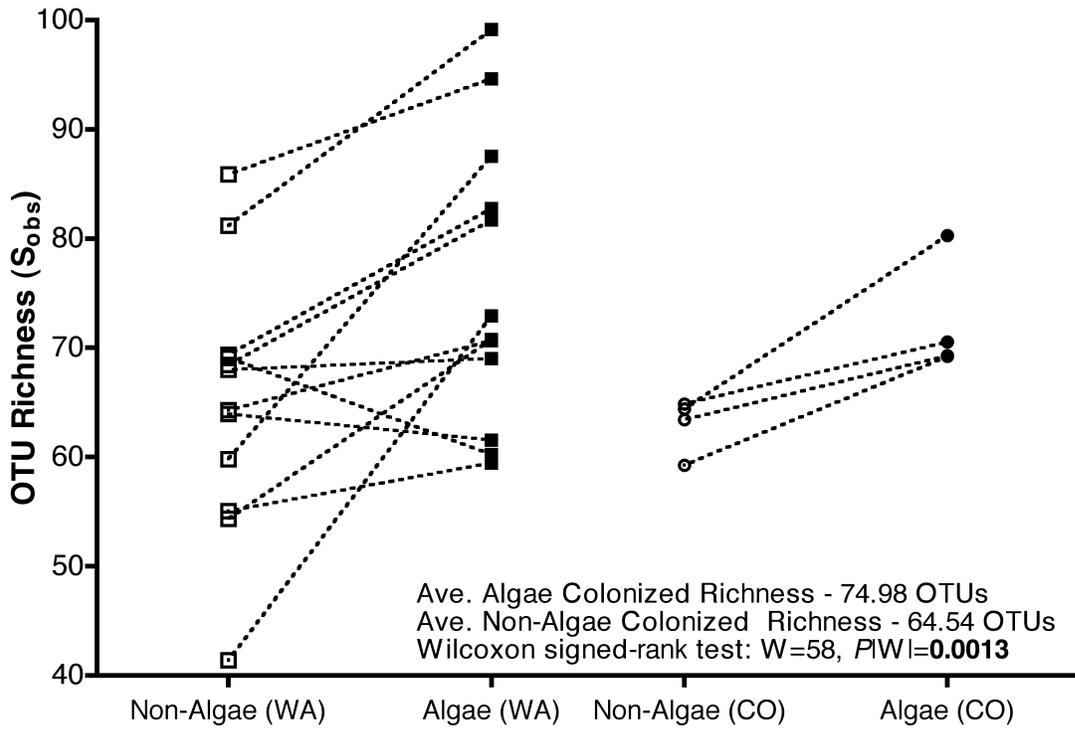
Figure 4.2 Non-Metric Multi-Dimensional Scaling (NMDS) plot of snow-borne fungi in algal colonized snow (solid symbols) and uncolonized snow (open symbols). Dashed lines connect paired algal colonized and adjacent uncolonized samples. AMOVA indicate that Colorado (circles) and Washington (squares) fungal communities are distinct. Insert represents paired Wilcoxon Signed-Rank test across Axis 2 (57.12% of community variability) and indicates that snow algae colonization shifts fungal communities (W=48, *P*=0.011), dashed lines connect paired samples.
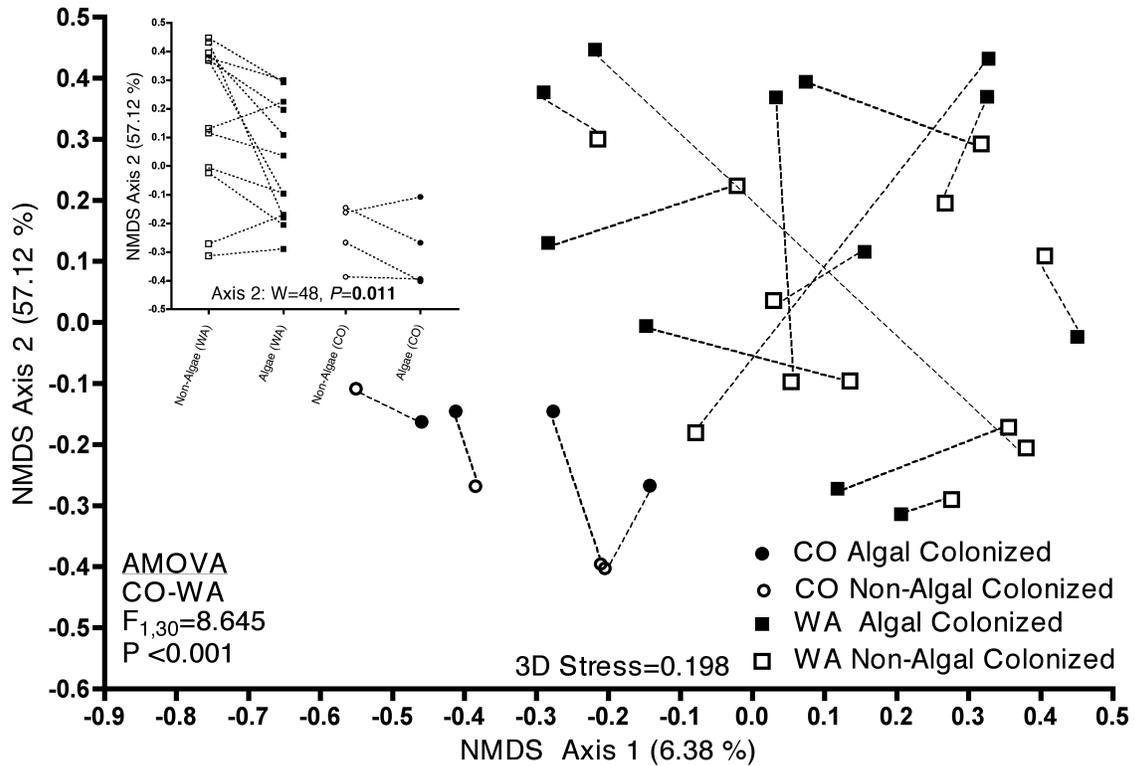
Chapter 5 - A community of clones: snow algae are diverse communities of spatially structured clones[4]

Abstract

Snow algae are cosmopolitan and often colonize late-season snow packs. These snow algae do not occur in isolation; rather, the visible algal patches consist of communities of algal species. Although several of these common snow algae have been characterized taxonomically, their inter- and intraspecific diversity remains unknown. Further, the phylogeographic and biogeographic structuring of snow algal species is poorly understood. Algal communities were censused through Illumina® MiSeq sequences. The results show that the communities are diverse and taxonomically broad [Orders: Chlamydomonadales (74% of operational taxonomic units – OTUs), Microthamniales (20% OTUs), and Chlorellales (6% OTUs)]. We further analyzed two of the most common and abundant OTUs for biogeographic haplotype diversity. We demonstrate that these two species (best BLASTn match to *Coenochloris* sp. and *Chlamydomonas* sp.) have distinct haplotype distributions both locally and regionally. Each sampled algae colony was dominated by one haplotype with negligible haplotype diversity. This suggests that these communities are highly clonal within a discrete patch. Further, these snow algae exhibit contrasting geographic structuring, evidencing different dispersal filtering.

Keywords: *snow algae, haplotype, biogeography, isolation by distance, Illumina MiSeq*

---

[4] To be submitted to *Molecular Ecology*: Brown SP, Ungerer MC, and Jumpponen A (2014). A community of clones: snow algae are diverse communities of spatially structured clones (in final preparation).

Introduction

Snows colonized by algae are visually striking with distinct boundaries separating colonized snows from non-colonized snows. Although it has been demonstrated that these algae-colonized snows house a diverse community of algae (Remias *et al.* 2010; Fujii *et al.* 2010), the importance of geographic structuring of these communities remains unknown. Previous queries into the associated ecologies of snow algae have been limited mainly to community composition. The increasing use of DNA metabarcoding tools has elucidated the hidden diversity of algae and presented a challenge to integrate these new discoveries with traditional taxonomy (De Clerck *et al.* 2013). The adoption of sequence based identification in snow algae community analyses, however, suffers from global nucleotide database deficiencies. Snow algae have been described as simple communities that consist of few species (Takeuchi 2001) and are generally considered to occur in closed ecosystems (Hoham & Duval, 2001). Most biogeographic studies of snow algae communities focus on altitudinal gradients and demonstrate that these simple communities exhibit drastically different altitudinal patterning associated with cell abundance (Yoshimura *et al.* 1997; Takeuchi 2001; Takeuchi & Kohshima 2004) suggesting that snow algae are strongly structured by differences in ice pack structure even at a local scale. It is uncertain if the observed differences between these snow algae result from geographic distance or differences in substrate quality. Additionally, to the best on our knowledge, investigations into the biogeography of snow algae using locus targeted sequencing integrating community ecology and phylogenetic structure are lacking.

The aim of this research was to investigate spatial and temporal dynamics of snow algae communities. Additionally, we explored genetic homogeneity and population dynamics of the two most common snow algae across regions and years as well as investigated evidence for isolation by distance suggestive of geographic structuring. We used population analyses to determine population haplotype structure and infer potential dispersal mechanisms of snow algae. Population analyses of two common snow algae sampled across two years across a semi-continental scale allowed for an in-depth investigation into the biogeography of snow algae. Our results provide new insights into the distribution of snow algae and address questions about

micro-evolutionary forces that drive the colonization of adjacent but discrete patches of snow by algae.

Materials and Methods

Sampling Locations and Sampling

Algae colonized snows were collected over two years (2011 and 2012) at sites in both Colorado, USA (Niwot Ridge Long-Term Ecological Research Site (2011) and Indian Peaks Wilderness Area within the Arapahoe and Roosevelt National Forest (2012)) and Washington State, USA (Glacier Peak Wilderness Area within the Wenatchee National Forest (2011 and 2012; see Table 5.1; Fig. 5.1). Detailed sampling locations and protocols were described previously in Brown *et al.* (2014).

Sequence Harvesting from Previous Data

In our previous investigations into the fungal communities associated with these algal patches (see Brown *et al.*, 2014), Illumina® MiSeq paired-end amplicons were generated using fungus specific primers for the Internal Transcribed Spacer 2 (ITS2) region of the rRNA gene repeat [fITS7 (Ihrmark *et al.*, 2012) and ITS4 (White *et al.*, 1990)]. Briefly, algae colonized snows were sampled by taking five ~85 cm^3 surface subsamples within visibly algae colonized snow patches from an area of about 2m^2 from each colonized patch. Snow with obvious anthropogenic interference was avoided. Collected snows were allowed to melt under ambient conditions, homogenized, and 100mL of the melt water filtered onto 2μm Nucleopore Track-Etch Membrane filters (Whatman®, Kent, UK). The filters were placed into MoBio UltraClean Soil DNA Isolation Kit (Carlsbad, CA, USA) bead tubes and transported to the laboratory. Total Genomic DNA was extracted and a locus-targeted amplicon library was generated and sequenced at the Integrated Genomics Facility (Manhattan, KS, USA). Despite the use of fungus specific primers, our libraries included a large proportion of plants and, fortuitously, algae (including a total of 369,651 algae sequences (~40% total sequences) after stringent sequence quality control). The initial investigation of fungi associated with snow algae provided a surprisingly high proportion of non-fungal targets of algal origin. In all, of the 200 most

abundant OTUs (see Brown et al. 2014), 15 OTUs likely represented snow algae. These results indicate that snow algae comprise a complex, diverse, and dynamic community as reported previously (Remias *et al.* 2010; Fujii *et al.* 2010; Fig. 5.1). The two most abundant non-fungal OTUs (using a 97% sequence similarity threshold for delineating OTU membership) had best BLASTn affinities to two snow algae (*Coenochloris* sp. [HQ404874.1, Query coverage=100%, Max Identity=82%] and *Chlamydomonas* sp. [HQ404867, Query coverage=100%, Max Identity=86%]). These algal OTUs were both abundant across all sampling sites and dominant in our snow-algae samples.

Population Analyses

Due to the fact that these two algae were highly abundant and widely distributed, they were deemed sufficiently abundant to allow population analyses. All sequences belonging to these two OTUs were harvested from the initial fasta file after paired contigs were generated. Given the highly skewed OTU abundance distribution for all locations (16 locations), we randomly selected 50 sequences without replacement from each sampling location for each of the two OTU (total of 800 sequences each for *Coenochloris* and *Chlamydomonas*). The OTU with affinity to *Chlamydomonas* sp. likely represents *Chlamydomonas nivalis* (Bau.) Wille (Division Chlorophyta)(Gradinger & Nürnberd 1996)*,* a red-pigmented chlorophyte that is visually the most dominant in red snow; we refer to this OTU as *Chlamydomonas* sp. from this point forward. However, we cannot place this putative *Chlamydomonas* sp. into *C. nivalis* with certitude as no ITS2 representative sequences for this alga exist in the combined global genetic databases. The identity of the second OTU with the best BLASTn affinity to *Coenochloris* is less clear. There is a dearth of information about these snow algae and their taxonomic identities are less than certain. However, we refer to this alga as *Coenochloris* sp. for the remainder of this text.

To confirm the putative taxonomic placement of our target algae, we conducted phylogenetic analyses. It is important to note that the hypervariable ITS regions are not the best for discerning higher-level phylogenetic relationships. However, it has a great potential for resolving algal phylogenies at the species or genus level (An *et al.* 1999). Taxonomy of most snow algae is uncertain and many groups remain poorly resolved at the present. To generate

Maximum Likelihood (ML) trees for confirmatory taxonomic placement, we harvested representative sequences of the delineated haplotypes and sequences from GenBank. For genus *Coenochloris*, we harvested all GenBank sequences deposited by the Culture Collection of Cryophilic Algae (CCCryo, www.cccryo.fraunhofer.de) that contained the full-length ITS2 region. Additionally, we included one full length ITS2 sequence for a freshwater lake *Coenochloris* as well as all full-length ITS2 sequences from the closely related genus *Sphaetocystis*. Currently, the relationship between *Coenochloris* and *Sphaetocystis* is unresolved as representatives of these genera are often moved between these genera. For genus *Chlamydomonas,* we harvested all full-length ITS2 sequences deposited by the CCCryo as well as two sequences from the closely related genus *Chloromonas,* whose placement is unresolved. Additional CCCryo generated Chlamydomonadaceae clones were included. Fasta formatted files of haplotype and reference sequences are provided in the supplementary information. These reference sequences were hand-edited to include only the ITS2 regions. Using the program GENEIOUS (v.5.3.4, Biomatters LTD., Auckland, NZ), representative haplotype sequences and reference ITS2 sequences were MAFFT aligned and a Maximum Likelihood Tree (PHYML) was generated using 100 iterations for each of the two snow algae.

Additionally, to determine if these algae are phylogenetically clustered across a continental scale (Colorado vs. Washington), a Neighbor Joining tree was generated (GENEIOUS, using a Jukes-Cantor distance matrix) using 100 iterations using all of the 800 sequences for each alga. To test if these consensus trees were clustered differently than expected by random, the trees were analyzed using UniFrac (Lozupone & Knight 2005; bmf.colorado.edu/unifrac) using 100 iterations with both weighted and unweighted options.

Haplotype Analyses

Using the program MOTHUR (v.1.31.1; Schloss *et al.*, 2009), the 800 con-OTU (con-specific) sequences for each alga (50 from each sampling location) were truncated to 325 base pairs after removal of primers and multiplex identification tags (MIDs). The truncated sequences were pairwise aligned (Needleman-Wunsch algorithm; Needleman & Wunsch 1970) to derive a pairwise distance matrix and the resultant matrix clustered at a 98.5% similarity threshold to delineate haplotypes. Because of the high variability within the ITS2 region, this threshold

allowed for calling haplotypes, whilst being conservative enough to not place all sequences into different haplotypes.

Haplotype diversity within and among populations was evaluated using an Analysis of Molecular Variance (AMOVA) in the program ARLEQUIN (v.3.5.1.3; Excoffier & Lischer 2010). To discern geographic patterns of snow algal distribution, a series of AMOVA models was used with samples grouped by: (i) sampling location (Colorado vs. Washington), (ii) Sampling year (2011 vs. 2012 in Washington only, as these were from same sampled Lyman Glacier basin), and (iii) Washington sites (the same locations were sampled within Washington for both 2011 and 2012). Models were run using 1000 permutations. Additionally, pairwise genetic distances for each alga between Washington State samples for both 2011 and 2012 were generated between the dominant haplotype sequences using both uncorrected genetic (pairwise distance – Needleman-Wunsch in MOTHUR) and corrected (Kimura two-parameter (K2P) distance in ARLEQUIN) distances. These pairwise distances were regressed with pairwise geographic distance to examine possible isolation by distance patterns. To further interrogate isolation by distance, correlation analyses between each algal pairwise geographic distances and genetic distance matrices using Mantel tests were performed (1000 iterations). All statistical tests were done with a combination of MOTHUR, ARLEQUIN, and JMP (v.7.0.2, SAS Institute Inc., Cary, NC).

Results

Snow Algae Diversity

The snow algae communities observed here included representatives from three orders and indicate diverse and dynamic snow algae communities. In all, algal sequences represented greater than 40% of the total ITS2 sequences of our previous fungal-targeted sequences (Brown *et al.* 2014). Of the 15 OTUs presented here, 11 are placed within Order Chlamydomonadales and likely represent *Chlamydomonas* and *Chloromonas* species commonly observed in visual surveys of snow algae. However, because of the poor sampling of representatives of these species in the combined global genetic databases, confirming these affinities is difficult. Additional taxa included three representatives of the Order Microthamniales (genus *Trebouxia*) and one of the Order Chlorellales (genus *Coenochloris*). The OTU identified as *Coenochloris* sp.

was most abundant by sequence count and over seven time more abundant than the next most abundant fungal OTU (Brown *et al.* 2014). It is important to note that these observed algal sequences were not specifically targeted in our initial study. Thus, our analyses may represent an underestimation of actual snow algae diversity.

Phylogenetic Analyses

Our ML analyses of representative ITS2 sequences from CCCyro, and other similar sequences from GenBank confirmed that sequences selected for our haplotype analyses belong to the clade with *Chlamydomonas* sp. and Chlamydomonadaceae sp. references. This clade does not include *Chloromonas,* another genus often found in snow. Despite the poor resolution among some species of snow-borne *Chlamydomonas* and *Chloromonas,* our queried haplotype sequences (*Chlamydomonas* sp.) shared a clade with Chlamydomonadaceae with strong bootstrap support (Fig. 5.2). We cannot be certain that our *Chlamydomonas* sequences represent *C. nivalis* because the combined global genetic databases have no complete ITS2 sequences of *C. nivalis.* Although it is likely that these environmental sequences represent *C. nivalis* as a result of sampling within red-snow patches, our sequences may represent different *Chlamydomonas* or Chlamydomonadaceae species altogether.

In contrast to our ML analyses of Chlamydomonadaceae*,* analyses of our putative *Coenochloris* haplotypes are not as clear. The environmental sequences of our putatively assigned *Coenochloris* sp. sequences are closely related to previously accessioned *Coenochloris* sequences with strong bootstrap support (Fig. 5.3). However, our environmental *Coenochloris* sequences are more closely related – not to snow-borne *Coenochloris* sequences from CCCryo (primarily from Svalbard, Norway) – but rather are more similar to the only other ITS2 *Coenochloris* sequence in GenBank (a freshwater isolate from Portugal). The CCCryo generated sequences from snow clustered with *Sphaerocystis* sequences evidencing the poor resolution of these two groups. This inconsistency may in part be due to the paucity of information available for *Coenochloris* and indicates the poor phylogenetic and taxonomic understanding of this genus.

Our UniFrac analyses indicated strong terminal node clustering between Colorado and Washington locations. Both *Chlamydomonas* and *Coenochloris* were geographically isolated as evidenced by significant UniFrac scores (P<0.01 for both alga) between Washington and

Colorado sites using either weighted or unweighted analyses after Bonferroni correction for multiple comparisons.

Haplotype Identification

Using a 98.5% similarity threshold, the 800 *Chlamydomonas* sequences were distributed into 27 haplotypes and the 800 *Coenochloris* sequences into 30 haplotypes (see Table D.1 for haplotype identification). The algae present within each discrete patch belonged, with only one exception, to non-overlapping haplotypes. Further, each algal patch was dominated by only one haplotype (sometimes exclusively) with a very minor proportion of sub-dominant haplotypes. In only one sampling location (Indian Peaks, 2012 Site 2) the dominant haplotype comprised a proportion < 80%: 54% of the sequences belonged to the dominant haplotype and 46% to a sub-dominant haplotype. Our data suggest that whilst the snow algal patches are comprised of multiple species, each species within a patch is highly clonal. It is likely that the observed clonality is primarily driven by strong priority effects that structure algal communities at small local scales, *i.e.* within each colonized patch.

Population-level Haplotype Diversity

Results of AMOVAs indicate that the majority of haplotype genetic diversity is explained by the among-population, within group variance component (Φ_{SC}). This component explains at least 86% of the genetic variation, no matter what the group delineation (Table 5.2). Interestingly, the within-population components are highly significant (Φ_{ST}), even though this component represented only a small proportion of the total variation. Somewhat surprisingly, the among-group variance components (Φ_{CT}) were negligible. This is particularly interesting given that our UniFrac tests indicated that Colorado and Washington sites differ in terminal node clustering of Neighbor-Joining trees for both algae.

Geographic Relationships

Chlamydomonas sp. and *Coenochloris* sp. differed in their relationships between pairwise genetic and geographic distances. Our analyses suggest different natural histories between the

two taxa. Uncorrected genetic distances show that the *Coenochloris* populations are strongly and positively correlated with geographic distances at the Washington State sites for both 2011 (R^2=0.443, t=3.22, P=0.0067; Fig. 5.4a) and 2012 (R^2=0.460, t=3.33, P=0.0054; Fig. 5.4b). In contrast, the genetic and geographic distances for *Chlamydomonas* populations were correlated neither in 2011 (R^2=0.016, t=-0.54, P=0.6015; Fig. 5.4c) nor in 2012 (R^2=0.004, t=0.22, P=0.8325; Fig. 5.4d). Mantel tests on matrix correlations corroborate our regression analyses. *Coenochloris* pairwise geographic and genetic matrices were positively correlated for both 2011(r=0.666, P=0.013) and 2012 (r=0.678, P=0.021) whereas *Chlamydomonas* was not correlated for 2011 (r=-0.147, P=0.624) or 2012 (r=0.060, P=0.434). However, when using the corrected genetic distance, the differences between the two algae were less clear. For 2011 and 2012, *Chlamydomonas* genetic distance remained uncorrelated with geographic distance (2011-R^2=0.0782, t=-1.05, P=0.3128; 2012 – R^2=0.0748, t=1.03, P=0.3237). The K2P corrected distances provided a slightly contrasting view of *Coenochloris* compared to uncorrected distances. The strong geographic response between genetic and geographic distance were congruent for 2011 (R^2=0.4366, t=3.17, P=0.0073). Interestingly, the 2012 *Coenochloris* populations still were still strongly positively trended, but this correlation was no longer significant (R^2=0.0850, t=1.10, P=0.2920).

Discussion

We present a first of its kind analysis of the genetic and geographic distribution of two taxa that dominate in snow algal communities. Similarly to other studies (Remias *et al.* 2010; Fujii *et al.* 2010), our data indicate a diverse and complex algal community even though the snows seem visually dominated by a single taxon. The algae co-occur in great numbers within these discrete patches, but are rare in adjacent uncolonized snows. The microhabitat attributes that allow for the establishment of these discrete algal patches are not yet understood. It is possible that establishment is stochastic and opportunistic. It may be that once an algal propagule successfully establishes under favorable conditions or in presence of available nutrients (most likely nitrogen), this alga then reproduces asexually in great numbers. This establishment may be aided by diazotrophs that establish a positive feedback loop further enriching the patch with nitrogen and carbon. The nutrient inputs may further facilitate establishment and population

growth of other algae and microbes including bacteria and fungi that may be enriched within the algal patch. The founder effects from few propagules potentially explain why only few haplotypes occur and dominate each patch.

One of the more interesting findings is the extreme site specificity of the haplotypes (see Table D.1; Figs. 5.2, 5.3). Both *Coenochloris* sp. and *Chlamydomonas* sp. within algal patches are strongly dominated by one haplotype with few, and often no sub-dominant representatives. The lack of sub-dominant haplotypes may be a result or our shallow sampling (50 sequences per algae patch per species) and deeper interrogation would likely show more haplotype diversity. However, it is unlikely that additional sub-dominant haplotypes would have proportions greater than the ~2% seen here. Further, the sub-dominant haplotypes are, with few exceptions, very closely related to the dominant haplotype in our ML analyses. Taken together, these observations suggest that the algal patches are highly clonal, driven by strong priority effects that inhibit successful establishment of subsequently arriving propagules. Alternatively, these patterns may be explained by competitive exclusion of one niche-occupying alga over another. This may also be a result of low propagule inputs or low propagule viabilities combined with stochastic establishment that controls the population/community assembly within a colonized patch.

The diploid zygotes of the algae often dominate the colonized snows. These zygotes likely propagate asexually locally producing clonal populations in extremely large number. Our analyses indicate that the sub-dominant haplotypes within a patch are closely related to the dominant haplotype, implying sexual recombination with close kin. Alternatively, although these algae are largely clonal, they may also reproduce unisexually explaining the distinct, but closely related, sub-dominant haplotype occurrence. Given that haplotypes that occupy the same snow-patch are very closely related, it is conceivable that these sub-dominant haplotypes arise from the parent population (dominant haplotype) of the major haplotype due to mutations (or unisexual reproduction). For the 325 bp sequences queried here, only five nucleotides are required to be distinct for inclusion into a new haplotype. These newly mutated, closely related haplotypes are then subjected to very strong kin competition that continues to suppress the sub-dominant haplotype frequencies. However, our data do not permit elucidating the reproductive strategies or the genesis of sub-dominant haplotypes; further investigation is required to adequately address these questions.

Our data also suggest contrasting life-history strategies of *Coenochloris* and *Chlamydomonas*. In our analyses, *Coenochloris* was strongly geographically structured as evidenced by the positive correlations between genetic and geographic distances (Figs. 5.3a and 5.3b). Although *Coenochloris* may undergo sexual reproduction within patches, our data suggest some isolation by distance. Such isolation may prevent crosspatch dispersal and suggests rampant near-neighbor mating with little long-distance propagule dispersal for reproduction. In contrast, snow-borne *Chlamydomonas* seems readily dispersed to regionally discrete patches during the reproductive cycle, as our data provide no evidence for geographic structure in the *Chlamydomonas* populations. This suggests that *Chlamydomonas* gametes may be readily dispersed across the landscape. The observed contrasts are likely attributable to different propagule sizes. The average *Chlamydomonas nivalis* spherical zygote is 14.9μm in diameter (Remias *et al.,* 2005), whereas the size of the adult cells of the type genus *Coenochloris* range from 5 to 11μm in diameter. However, up to 16 *Coenochloris* sp. cells may encapsulate into a mucilaginous envelope resulting in a dispersal unit up to 60μm in diameter (Fott 1974). In other words, despite the smaller *Coenochloris* cells*,* the dispersal propagules are actually more than 64 times more voluminous than *Chlamydomonas* sp. The larger dispersal propagules are unlikely to be effective in areal dispersal. It is important to note that our size estimates for the *Coenochloris* propagules are based on the type specimen for this genus, which is a freshwater organism. Interestingly, this contrasts our observations of the unique *Chlamydomonas* haplotypes within each algae patch. The lack of geographic structuring suggests that these algae can disperse to cross with other patches. Yet, our ML analyses provide no evidence for cross patch mating. Further experiments are required to conclusively elucidate the reproductive strategies of snow *Chlamydomonas*.

The contrasting correlations between the genetic by geographic distances indicate different dispersal potential of these two algae. Distant and discrete algae populations that share genetic structure often infer avian-mediated algal dispersal (Chiasson *et al*. 2003; Struneckí *et al*. 2012) and have been suggested to be important for snow algae dispersal. We find that there are no large-scale geographic similarities in haplotype genetic structure as would be expected in avian-vectored systems suggesting that avian vectored dispersal is unlikely the driver of snow algae dispersal. We conclude that these algae are likely dispersed aerially and that the avian vectors are less important. The *Chlamydomonas* sp. and *Coenochloris* sp. studied here seem to

have contrasting affinities for aerial dispersal. *Chlamydomonas* sp. is not limited for dispersal on a regional scale as evidenced by the lack of any correlation between the genetic and geographic distances across Washington State's Lyman Glacier basin in either 2011 or 2012. This indicates that *Chlamydomonas* sp. is readily transported on regional scales. In contrast, *Coenochloris* sp. populations were structured geographically and their genetic distances tended to be correlated with geographical distance. The haplotype specific patches support aerial dispersal in short distances. However, the mechanisms and the potential for the long-distance propagule dispersal remain uncertain.

Conclusions

Our study demonstrates that diverse and complex algal communities colonize discrete snow patches. While comprised of several species and thus possessing interspecific variability, each species within a snow patch seems to be highly clonal and with low intraspecific variability. Although our analyses indicated that both targeted taxa possessed similar low intraspecific variability, their populations on larger, regional scales were differently structured. These observations suggested contrasting dispersal abilities that may be primarily driven by the size of the dispersal propagules. Snow algae are communities of clones and these clonal communities are likely structured via intense kin competition as well as dispersal limitations. This makes snow algae an exceptional study system and one uniquely suited test theories of ecological persistence and selection pressures.

Acknowledgements

References

An, S.S., Friedl, T., & Hegewald, E. 1999. Phylogenetic relationships of *Scenedesmus* and *Scenedesmus*-like coccoid green algae as inferred from ITS-2 rDNA sequence comparisons. *Plant Biology.* 1:418-428.

Brown, S.P., Olson, B.J.S.C. & Jumpponen, A. 2014. Snow algae as the control of snow fungal community assembly. *Molecular Ecology* (in review).

Chiasson, W.B., Machesky, N.J. & Vis, M.L. 2003. Phylogeography of a freshwater red alga, *Batrachospermum helminthosum,* in North America. *Phycologia.* 42:654-660.

De Clerck, O., Guiry, M.D., Leliaert, F., Samyn, Y., & Verbruggen, H. 2013. Algae taxonomy: a road to nowhere? *Journal of Phycology.* 49:2415-225.

Excoffier, L. & Lischer, H.E.L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources.* 10:564-567.

Fott, B. 1974. Taxonomie der palmelloiden *Chlorococcales* (Familie *Palmogloeaceae*). *Preslia.* 46:1-31.

Fujii, M., Takano, Y., Kojima, H., Hoshino, T., Tanaka, R. & Fukui, M. 2010. Microbial community structure, pigment composition, and nitrogen source of red snow in Antarctica. *Microbial Ecology.* 59:466-475.

Gradinger, R. & Nürnberg D. 1996. Snow algae communities on Arctic pack ice floes dominated by *Chlamydomonas nivalis* (Bauer) Wille. *Proc. NIPR Symposium Polar Biology.* 9:35-43.

Gorton, H.L., William, W.E. & Vogelmann, T.C. 2001. The light environment and cellular optics of the snow alga *Chlamydomonas nivalis* (Bauer) Wille. *Photochemistry and Photobiology.* 73:611-620.

Hoham, R.W. 1980. Unicellular chlorophytes – snow algae. *In* Cox, E.R. [Ed.] *Phytoflagellates*. Elsevier-North, Holland.

Hoham, R.W. & Duval, B. 2001. Microbial ecology of snow and freshwater ice with emphasis on snow algae. In Jones, H.G, Pomeroy, J.W., Walker, D.A., & Hoham R.W. [Eds.] *Snow Ecology: An Interdisciplinary Examination of Snow-Covered Ecosystems*. Cambridge University Press, Cambridge, UK.

Ihrmark, K., Bödeker, I.T.M., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., Strid, Y., Stenlid, J., Brandstöm-Durling, M., Clemmensen, K.E. & Lindahl, B.D. 2012. New primers to amplify the fungal ITS2 region – evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiology Ecology.* 82:666-677.

Lozupone, C. & Knight, R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied Environmental Microbiology.* 71:8228-8235.

Needleman, S.B., & Wunsch, C.D. 1970. A general method applicable to the search for similarities in the amino acid sequences of two proteins. *Journal of Molecular Biology.* 48:443-453.

Remias, D., Lütz-Meindl, U. & Lütz, C. 2005. Photosynthesis, pigments and ultrastructure of the alpine snow alga *Chlamydomonas nivalis*. *European Journal of Phycology.* 40:259-268.

Remias, D., Karsten, U., Lütz, C. & Leya, T. 2010. Physiological and morphological processes in the Alpine snow alga *Chloromonas nivalis* (Chlorophyceae) during cyst formation. *Protoplasma.* 243:73-86.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M.,, Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G, Van Horn, D.J. & Weber, C.F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology.* 75:7537-7541.

Struneckí, O., Komárek, J. & Elster, J. 2012. Biogeography of *Phormidium autumnale* (Oscillatoriales Cyanobacteria) in western and central Spitsbergen. *Polish Polar Research.* 33:369-382.

Takeuchi, N. 2001. The altitudinal distribution of snow algae on an Alaska glacier (Gulkana Glacier in the Alaska Range). *Hydrological Processes.* 15:3447-3459.

Takeuchi, N. & Kohshima, S. 2004. A snow algal community on Tyndall Glacier in the Southern Patagonia Icefield, Chile. *Arctic, Antarctic, and Alpine Research.* 36:92-99.

Thomas, W.H. & Duval, B. 1995. Sierra Nevada, California, U.S.A., snow algae: snow albedo changes, algal-bacterial interrelationships, and ultraviolet radiation effects. *Arctic and Alpine Research.* 27:389-399.

White, T.J., Bruns, T., Lee, S. & Taylor, J.W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In Innis, M.A., Gelfand, D.H., Sininsky, J.J. & White, T.J. [Eds.]. *PCR Protocols: A Guide to Methods and Applications.* Academic Press, Inc., New York.

Yoshimura, Y., Kohshima, S. & Ohtani, S. 1997. A community of snow algae on a Himalayan glacier: change of algal biomass and community structure with altitude. *Arctic and Alpine Research.* 29:126-137.

Tables and Figures

Table 5.1 Sampling locations of the snow algae for 2011 and 2012. Niwot - Niwot Ridge LTER site, Colorado, USA; Indian Peaks – Indian Peaks Wilderness area, Colorado, USA; Lyman – Lyman Glacier basin within the Glacier Peak Wilderness area, Washington, USA

Year	State	Site	Landmark	Latitude	Longitude	Elevation (m)
2011	Colorado	Niwot 1	Near Soddie Laboratory	40° 02' 56" N	105° 34' 51" W	3368
2011	Colorado	Niwot 2	Saddle	40° 03' 30" N	105° 35' 20" W	3514
2011	Washington	Lyman 1	Cloudy Pass	48° 12' 09" N	120° 55' 28" W	1961
2011	Washington	Lyman 2	Terminal Moraine	48° 10' 59" N	120° 54' 11" W	1802
2011	Washington	Lyman 3	Lyman Glacier	48° 10' 21" N	120° 53' 50" W	1880
2011	Washington	Lyman 4	Spider Gap N	48° 10' 14" N	120° 52' 55" W	2135
2011	Washington	Lyman 5	Spider Gap S	48° 10' 10" N	120° 52' 53" W	2123
2011	Washington	Lyman 6	Lower Spider Snowfield	48° 09' 42" N	120° 52' 42" W	1897
2012	Colorado	Indian Peaks 1	E of Shoshoni Peak	40° 04' 02" N	105° 37' 44" W	3407
2012	Colorado	Indian Peaks 2	S Shore Lake Isabelle	40° 04' 01" N	105° 04' 01" W	3358
2012	Washington	Lyman 1	Cloudy Pass	48° 12' 10" N	120° 55' 27" W	1966
2012	Washington	Lyman 2	Terminal Moraine	48° 10' 58" N	120° 54' 11" W	1794
2012	Washington	Lyman 3	Lyman Glacier	48° 10' 24" N	120° 53' 49" W	1866
2012	Washington	Lyman 4	Spider Gap N	48° 10' 14" N	120° 52' 55" W	2173
2012	Washington	Lyman 5	Spider Gap S	48° 10' 10" N	120° 52' 53" W	2137
2012	Washington	Lyman 6	Lower Spider Snowfield	48° 09' 41" N	120° 52' 35" W	1893

Table 5.2 Results of AMOVA (phi-statistic) for three different grouping models

Groups	WA/CO		WA 2011/WA 2012		WA Sampling Sites	
Structure	*Coenochloris*	*Chlamydomonas* sp.	*Coenochloris*	*Chlamydomonas* sp.	*Coenochloris*	*Chlamydomonas* sp.
Φst	0.90376***(9.62)	0.86399***(13.60)	0.94482***(5.32)	0.88298***(11.70)	0.94409***(5.59)	0.88275***(11.73)
Φsc	0.90726***(94.15)	0.86610***(87.98)	0.94308***(91.43)	0.88199***(87.46)	0.94349***(93.35)	0.88035***(86.27)
Φct	-0.03776(-3.78)	-0.01579(-1.58)	0.03056(3.06)	0.00842(0.84)	0.01061(1.06)	0.02007(2.01)

Φ_{ST}, within-population variance; Φ_{SC}, among-population, within-groups variance; Φ_{CT}, among-groups variance
*** $P<0.001$

WA/CO = populations in Washington (n=12) vs. Colorado (n=4) combined across years (2011 and 2012); WA 2011/WA 2012 = Washington only samples for 2011 (n=6) and 2012 (n=6); WA sampling sites = the same sampling locations combined across years (2011 and 2012) as outlined in Table 1. Shown are the phi-statistics and the percentage of variation explained within each model (in parentheses).

Figure 5.1 Map of sampling locations in (top) Washington State and (bottom) Colorado. The pie charts represent the OTU abundance on an Order level of snow algae (see Brown *et al.* 2014) for 2011 and 2012. OTUs were widely spread and found in multiple years and across a semi-continental scale. For each sampling location, haplotype proportions are reported in Table D.1.
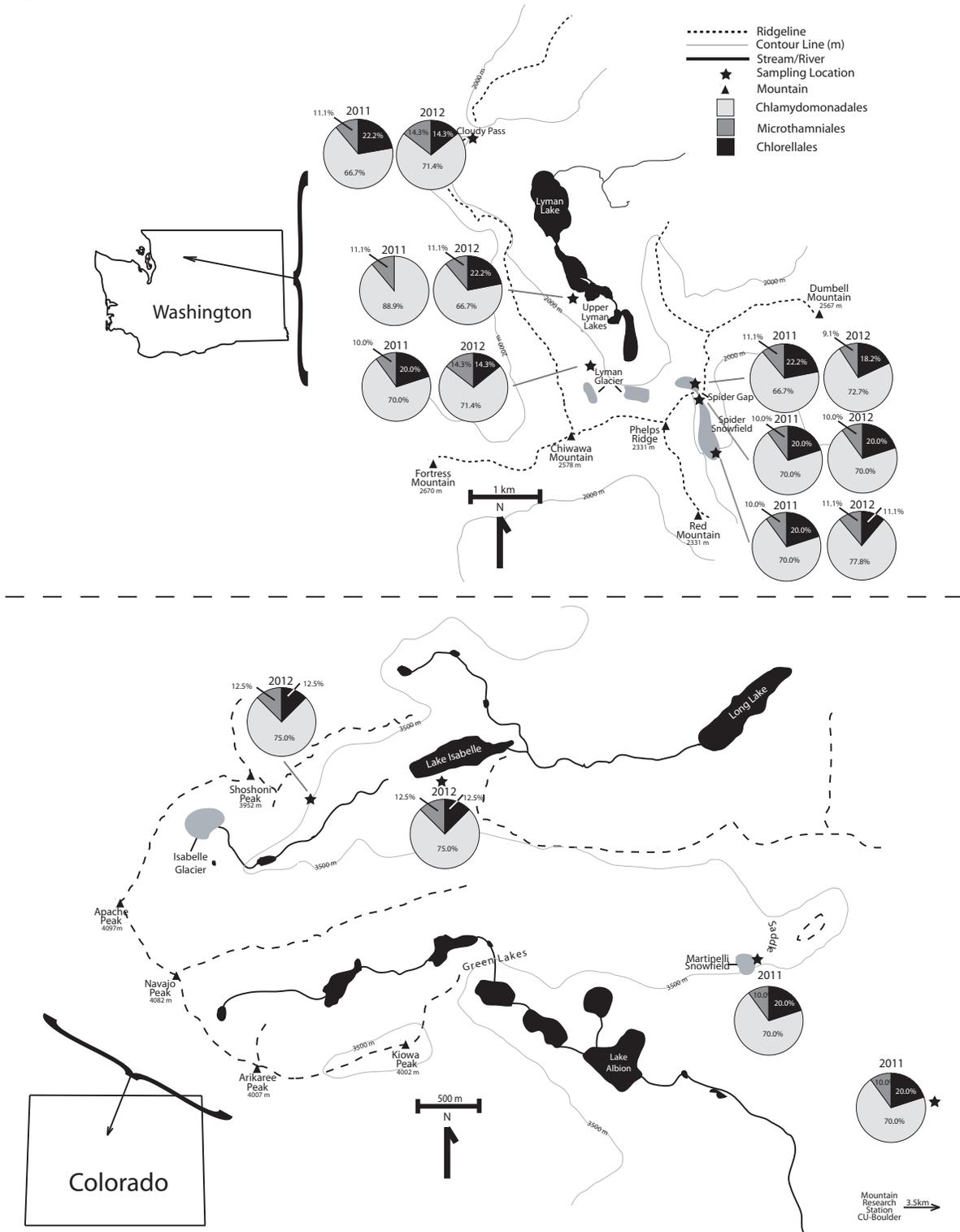
Figure 5.2 Maximum likelihood tree of environmental *Chlamydomonas* sp. haplotype sequences. The mafft-alignment includes representative sequences for each haplotype plus full length ITS2 sequences for snow-inhabiting *Chlamydomonas* or Chlamydomonadaceae and *Chloromonas* from GenBank. Numbers above the branches indicate bootstrap support for tree topology. Note that the obtained haplotypes fall within Chlamydomonadaceae with 100% bootstrap support. Representative haplotypes with their abundance (in parentheses) as well as sampling location (see Table 5.1) and reference sequences with their GenBank accession numbers (in parentheses) are indicated.
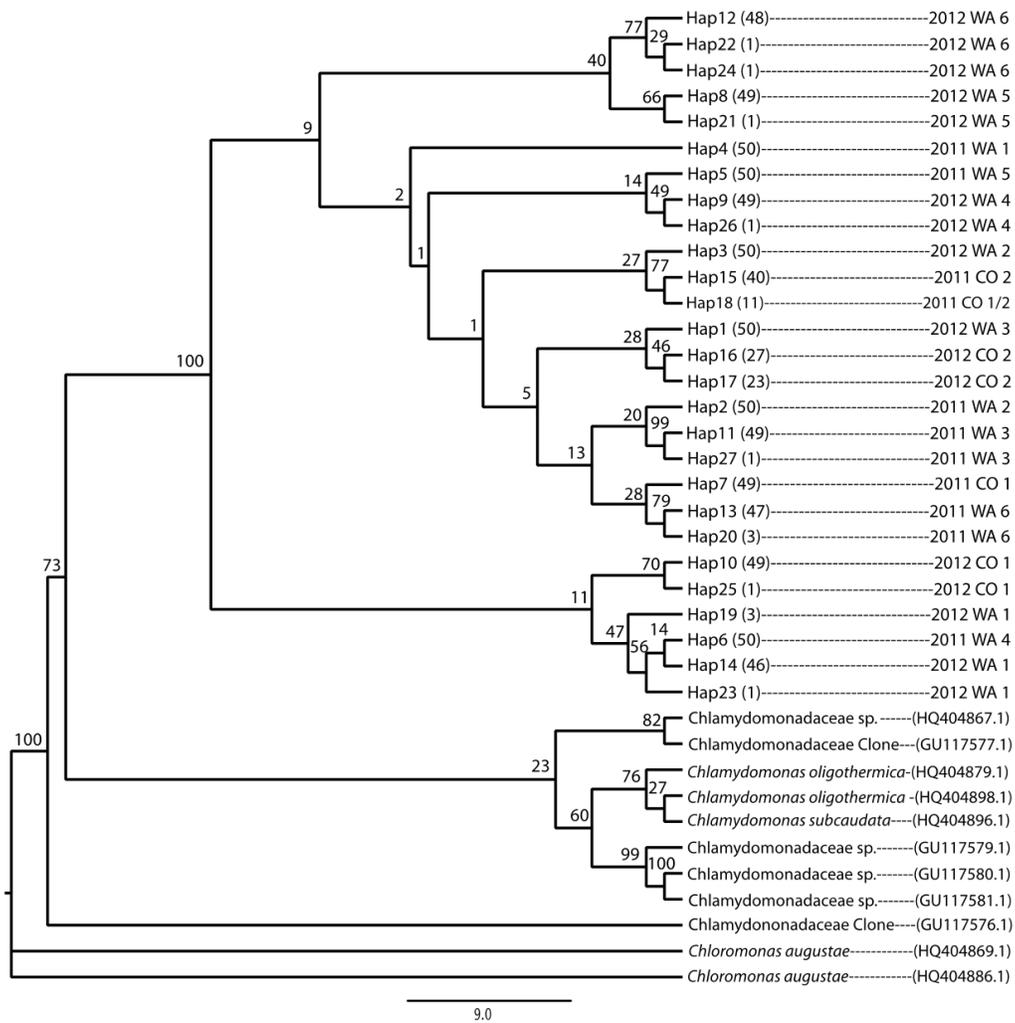
Figure 5.3 Maximum likelihood tree of environmental *Coenochloris* haplotype sequences. The mafft-alignment includes representative haplotypes sequences and all full length *Coenochloris* ITS2 sequences available in GenBank. Additionally, three full length ITS2 *Sphaerocystis* sequences were included as the two genera remain unresolved. Numbers above the branches indicate bootstrap support for tree topology. Note that the haplotype sequences group with one *Coenochloris* (freshwater isolate; GenBank accession GQ502288.1) with 100% bootstrap support, but not with the snow-inhabiting algae from CCCryo. Representative haplotypes with their abundance (in parentheses) as well as sampling location (see Table 5.1) and reference sequences with their GenBank accession numbers (in parentheses) are indicated.
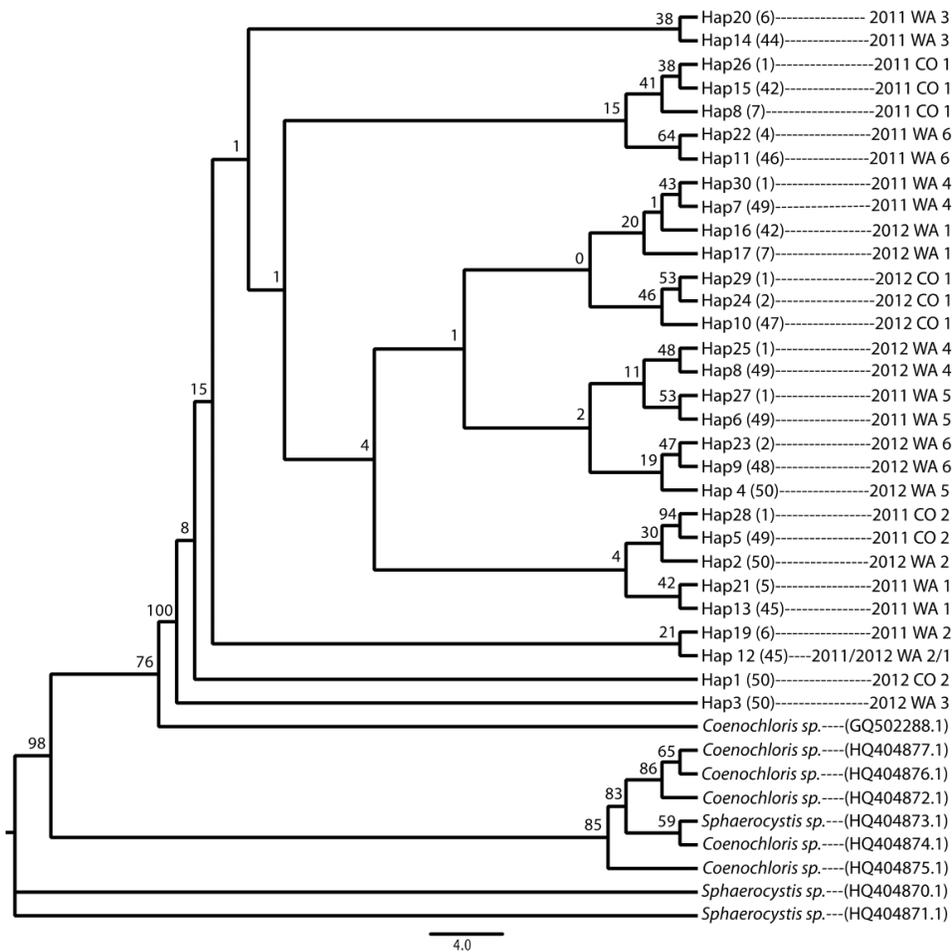
Figure 5.4 Regression between the sampling location geographic and haplotype genetic distances. Note that regression analyses of *Coenochloris* for 2011 (a) and 2012 (b) indicate positive correlation between geographic and genetic distances and geographic population structuring of this alga, i.e., potential dispersal limitation. In contrast, similar regression analyses of *Chlamydomonas* sp. indicated no geographic structuring either in 2011 (c) or 2012 (d) sampling.
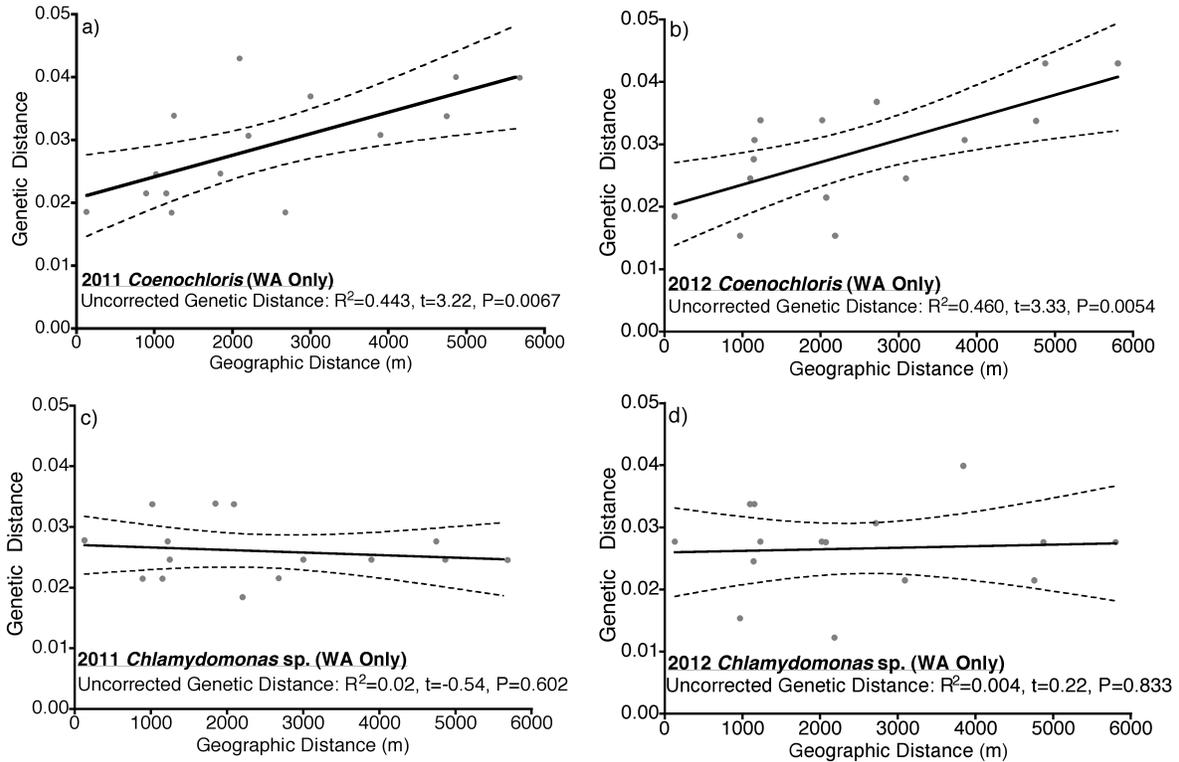
Chapter 6 - Analyses of ITS and LSU gene regions provide congruent results on fungal community responses[5]

Abstract

The Internal Transcribed Spacer (ITS) regions and the Large Subunit (LSU) of the nuclear ribosomal RNA (rRNA) gene complex are commonly used to elucidate questions in fungal community ecology. Here, we compared the congruence across these gene regions using two ecological experiments (primary successional dynamics at a receding glacier forefront and community dynamics in stored *Sorghum* biomass), in which both ITS1 and LSU were sequenced from the same DNA extracts. We analyzed richness, diversity and evenness estimators along with community shifts inferred from ordination analyses. Our analyses show that ITS and LSU provide similar results and consistent conclusions. Taken together, we conclude that either gene region is appropriate for testing ecological hypotheses as long as there are no a priori hypotheses that preclude the use of one gene region over the other.

Keywords: *ITS, LSU, Ecological Analysis, 454-Sequencing*

The Internal Transcribed Spacer (ITS) regions of the rRNA gene complex have been the primary targets for community ecology that relies on sequencing and sequence annotation for taxonomic information (Peay *et al,* 2008), so that ITS regions have been selected as the barcode for identification of environmental fungal sequences (Schoch *et al.*, 2012). There is little debate on whether ITS is a robust tool in fungal community ecology. ITS is also powerful in identifying fungi at species level taxonomic resolution. However, ITS regions are hyper-variable and great care must be exercised in ITS sequence data analyses to minimize user-generated biases (see Nilsson *et al.,* 2012). Often, novice investigators diving into next-generation sequencing targeting fungi are inclined to attempt global alignment of the ITS data to generate a distance matrix prior to binning sequences into OTUs as is commonly done with bacterial 16S sequences. It would seem reasonable to analyze fungal sequence data similarly, but this is usually inappropriate for filtering and clustering fungal ITS sequences because of the region's hyper-variability that precludes global alignments. Instead, ITS sequences can be pairwise aligned to generate a distance matrix for each pair of sequences. For larger datasets, however, this is often computationally expensive. The more conserved small subunit (SSU) or large subunit (LSU) sequences of the rRNA gene complex provide an alternative to ITS for testing ecological hypotheses, but may suffer from limited resolution when resolving obtained taxa. As a result of available reference alignments and the Naïve Bayesian Classifier (Wang *et al.,* 2007), the LSU region in particular has begun to gain some traction (Lothamer *et al.*, 2013; Porras-Alfaro *et al.*, 2014; Weber *et al.,* 2013). This is because these tools expedite analyses and provide access to a curated annotation tool for the sequence data, supporting the integration of evolutionary and ecological questions, similarly to bacterial systems. The available fungus specific primer sets for ITS and LSU regions that permit sequencing environmental fungal communities allow for generating relatively long sequences that can be analyzed using 454-pyrosequencing (Margulies *et al.,* 2005). It is of note that as of yet, there are no verified and minimally biased universal primer pairs that reliably and broadly amplify short enough gene regions within the fungal LSU to permit paired-end Illumina or Ion Torrent sequencing, although this is an area of current research. As little information is available whether the ITS or LSU provide congruent views of fungal ecology (see Kerekes *et al.,* 2013), we compared two pairs of ITS [primers ITS1f and ITS4] and LSU [primers LR0R and LR3; see Amend *et al.*, 2010] sequence data sets generated from the same DNA extracts to evaluate their reliability to describe basic fungal ecology.

We re-analyzed and compared ITS1 and LSU (with the variable region D1) 454-pyrosequencing libraries from two experiments that have been recently published or are currently in various stages of preparation; a total of four 454 sequence libraries. We selected a primary successional system at the forefront of a receding glacier (Brown and Jumpponen, 2014; BioProject **PRJNA201483;** SRA accessions for individual fastq files: ITS - **SRR934164-SRR943234**; LSU - **SRR1016610-SRR1016625**, **SRR1016734-SRR1016736**, **SRR1016744-SRR1016777**, **SRR1016785-SRR1016788**, **SRR1016859-SRR1016865**) and an experiment that evaluates effects of storage on fungal communities in stored *Sorghum* biomass (Rigdon-Huss *et al.,* unpublished; BioProject **PRJNA221342;** SRA accessions for sff files: ITS – **SRR1016401**; LSU – **SRR1016405**). The former evaluates community dynamics over time since deglaciation and as a result of plant establishment, whereas the latter follows fungal communities during six-month storage under various biomass covering treatments. All sequence data from the four libraries were processed identically with MOTHUR (v. 1.31.2; Schloss *et al.,* 2009) according to a modified standard operating protocol (Schloss *et al.,* 2011). Briefly, after denoising the .sff files and truncating all sequences to 250 bp in length, we removed all potentially chimeric reads. The sequence data for each library were pairwise aligned to generate a distance matrix and sequences clustered to OTUs at 97% sequence similarity using an average neighbor approach. After omission of singleton OTUs, we generated mean diversity and richness estimators based on 1000 iterations at a subsample depth of 200 sequences per experimental unit to minimize biases due to uneven sampling (Gihring *et al.* 2012): we compared OTU Richness (Sobs), Complement of Simpson's Diversity (1-D) and Simpson's Evenness (ED) using paired t-tests to test for dissimilarities between ITS and LSU sequence data sets on an experimental unit basis. As LSU permits global alignment, we also analyzed these data by alignment against a reference (James *et al.* 2006) in which all non-fungal sequences were removed and the remaining sequences were *de novo* aligned using MUSCLE. To analyze compositional shifts in the communities, we generated Bray-Curtis distance matrices and visualized communities using Non-metric Multi-Dimensional Scaling (NMDS) axes (k=3) with loading scores generated based on 1000 iterations for ITS and each of our LSU analysis methods at a 200 sequence subsampling depth. Two-way analysis of variance (ANOVA) was used to test if ecological indices and NMDS axes loading scores change with time, with treatment, or have time x treatment interactions on an experiment-wide basis across the ITS and LSU datasets (see Table E.1).

On an experimental unit basis, paired t-tests suggest significant but proportionally minor differences in diversity responses our two experiments between the two rRNA gene regions. However, it is the alignment of the LSU data to a reference that has a greater effect on results than the choice of a target region. The results from our glacier forefront study show that OTU richness, diversity and evenness differ between ITS and the reference aligned LSU data, but not between pairwise-aligned ITS and LSU data (Fig. 6.1a). This may be due to the relatively high abundance of early diverging taxa and broad taxonomic distribution of fungi present in the samples from the glacier forefront. Most reference alignments, including James' and coauthors' (2006), suffer from sparse sampling of unresolved early diverging clades. The biomass storage study provided a contrasting view. Regardless of how the LSU data were treated, the ITS and LSU data differed consistently in richness, diversity or evenness (Fig. 6.1b). Furthermore, pairwise- and reference-aligned LSU data were similar and consistently deflated in richness, diversity or evenness estimators compared to ITS. Again, this may be due to a phylogenetic distribution as well-characterized Ascomycota dominated the *Sorghum* biomass communities.

Although the paired t-tests provide interesting comparisons between the ITS and LSU data, experiment-wide ecological responses are usually more pertinent for an investigator deciding which gene region to use. In our glacier forefront study, richness, diversity and evenness behaved consistently and were unresponsive to the experimental conditions irrespective of the gene region or how the LSU data were analyzed. Similarly, in the stored biomass study, richness increased storage time across all three analyses. Diversity increased with storage time in LSU data sets (pairwise or reference aligned) with time, but did so only marginally significantly in the ITS data. Interestingly, the ITS data suggested clear treatment differences in evenness, whereas no such response was visible in the LSU data. In sum, experiment-wide, most diversity and richness metrics responded consistently in both systems regardless of the chosen gene region.

Finally, community analyses based on NMDS indicate that ITS and LSU also provide very similar results. Consistently with our original analyses (Brown and Jumpponen, 2014), axis loading scores show trajectories with distance from glacier in ITS and LSU (pairwise and reference aligned) data sets for at least one axis (Fig. 6.1c), but no effect of the plant establishment. Similarly to the glacier forefront study, the results from the *Sorghum* storage study are consistent regardless of the gene region or how the LSU data were treated: all analyses congruently suggest shifts in fungal communities over time (Fig. 6.1d) as well as distinguish one

of the four storage methods from others (Table E.1).

Overall, we conclude that ITS and LSU regions perform mainly similarly in estimating richness and diversity or in distinguishing treatments in community-wide ordination analyses. There was also congruence in taxa identified between ITS and LSU (see Fig. E.1) similarly to analyses by Porras-Alfaro *et al.* (2014), although obtained taxon affinities depend on inherent primer biases. This suggests that both gene regions are suitable for testing ecological hypotheses and provide comparable results with few exceptions. It is unclear what the underlying reasons for the observed differences are, but may be partly attributable to the variability between the gene regions, distribution of observed taxa, biases resulting from primer selection, or stochastic variability in the library production, sequencing, and analysis steps. Prior to deciding on a target gene region to sequence, the investigator must pay heed to the limitations of each gene region and how these may affect the downstream steps in sequencing or analysis. Furthermore, we wish to emphasize that the different sequencing targets may permit testing different hypotheses - a point to bear in mind when designing an experiment that relies on sequencing of variable regions. For example, if OTU identification to the level of species is paramount to the goals of a study, ITS regions are powerful and benefit from the greater number of available sequences in databases. In contrast, hypotheses addressing broad evolutionary ecological processes may be precluded by the inherently unalignable and hyper-variable nature of ITS regions. These caveats aside, we conclude that both ITS or LSU data are suited for analyses fungal community ecology.

References

Amend AS, Seifert KA, Samson R, Bruns TD. 2010. Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. *Proceedings on the National Academy of Sciences* **107**: 13748-13753.

Brown SP, Jumpponen A. 2014. Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. *Molecular Ecology* **23**: 481-197.

Gihring TM, Green AJ, Schadt CW. 2012. Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environmental Microbiology* **14:** 285- 290.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, *et al*. 2006. Resonstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443:** 818-822.

Kerekes J, Kaspari M, Stevenson B, Nilsson RH, Hartmann M, Amend A, Bruns TD. 2013. Nutrient enrichment increased species richness of leaf litter fungal assemblages in a tropical forest. *Molecular Ecology* **22:** 2827-2838**.**

Lothamer K, Brown SP, Mattox JD, Jumpponen A. 2013. Comparison of root-associated communities of native and non-native ectomycorrhizal hosts in an urban landscape. *Mycorrhiza* (in press) doi: 10.1007/s00572-013-0539-2.

Margulies M, Egholm M, Altman WE, Attiya A, Bader JS *et al.* 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376-380.

Nilsson RH, Tedersoo L, Abarenkov K, Ryberg M, Kristiansson E, Hartmann M, Schoch CL, Mylander JAA, Bergsten J, Porter TM, Jumpponen A, Vaishampayan P, Ovaskainen O, Hallenberg N, Bengtsson-Palme J, Eriksson KM, Larsson KH, Larsson E, Kóljalg U. 2012. Five simple guidelines for establishing basic authenticity and reliability of newly generated fungal ITS sequences. *MycoKeys* **4**: 37-63.

Peay KG, Kennedy PG, Bruns TD. 2008. Fungal community ecology: a hybrid beast with a molecular master. *BioScience* **58:** 799-810.

Porras-Alfaro A, Liu K, Kuske CR, Xie G. 2014. From genus to phylum: large-subunit and internal transcribed spacer rRNA operon regions show similar classification accuracies influenced by database composition. *Applied and Environmental Microbiology* **80**: 829-840.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mother: Open-source, platform-independent, community-suppoeted software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Schloss PD, Gevers D, Westcott SL. 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLOS ONE* **6:** e27310.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as

a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Science, USA* **109:** 6241-6246.

Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73:** 5261-5267.

Weber CF, Vilgalys R. Kuske CR. 2013. Changes in fungal community composition in response to elevated atmospheric CO_2 and nitrogen fertilization varies with soil horizon. *Frontiers in Microbiology* **4**: 78 doi: 10.3389/fmicb.2013.00078.

Tables and Figures

Figure 6.1 Paired t-tests of richness, diversity (1-d) and evenness (E_D) estimators derived from ITS, pairwise aligned LSU, and reference aligned LSU data in the Lyman Glacier Forefront (a) and Stored *Sorghum* Biomass (b) experiments. Bars indicate mean values across all experimental units and asterisks (*) represent significant differences at $P \leq 0.05$. Regression analyses of the Non-metric Multi-Dimensional Scaling (NDMS) loading scores of the first three resolved axes and distance from glacier terminus in the Lyman Glacier Forefront (c) and time of storage in the Stored *Sorghum* Biomass (d) experiments for ITS and LSU (pairwise and reference aligned). Asterisks indicate significant slopes ($P \leq 0.05$). Complete test statistics and *P*-values are in Appendix E.1.
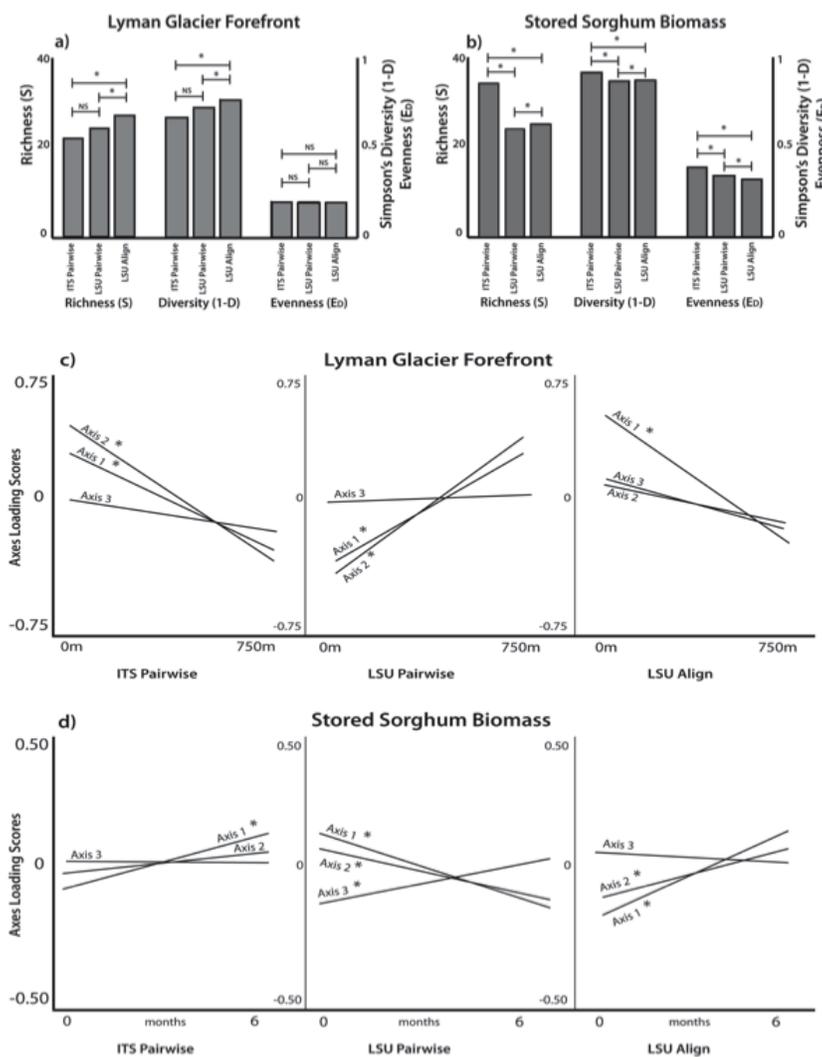
Chapter 7 - Scraping the bottom of the barrel: are rare high throughput sequences artifacts?[6]

Abstract

Metabarcoding data generated using next-generation sequencing (NGS) technologies are overwhelmed with rare taxa and skewed in Operational Taxonomic Unit (OTU) frequencies comprised of few dominant taxa. Low frequency OTUs comprise a rare biosphere of singleton and doubleton OTUs, which may include many artifacts. We present an in-depth analysis of global singletons across sixteen NGS libraries representing different ribosomal RNA gene regions, NGS technologies and chemistries. Our data indicate that many singletons (average of 38% across gene regions) are likely artifacts or potential artifacts, but a large fraction can be assigned to lower taxonomic levels with very high bootstrap support (~32% of sequences to genus with ≥ 90% bootstrap cutoff). Further, many singletons clustered into rare OTUs from other datasets highlighting their overlap across datasets or the poor performance of clustering algorithms. These data emphasize a need for caution when discarding rare sequence data *en masse*: such practices may result in throwing the baby out with the bathwater and underestimating the biodiversity. Yet, the rare sequences are unlikely to greatly affect ecological metrics. As a result, it may be prudent to err on the side of caution and omit rare OTUs prior to downstream analyses.

Keywords: *fungi, singleton, high-throughput sequencing, rare biosphere*

Next generation sequencing (NGS) permits deep interrogation of hyper-diverse fungal communities (Hibbett *et al*. 2009). Data generation has become expedient and sequence analysis/annotation more streamlined via available pipelines (e.g. MOTHUR, QIIME). Concurrently sequencing costs have declined, resulting in the democratization of sequencing in ecology (Caporaso *et al.* 2012). Many new investigators utilize NGS but are often uncertain how to handle rare operational taxonomic units (OTUs). These rarities are common - singletons alone often comprise half of all OTUs.

Rare OTUs may represent the 'rare biosphere' (Sogin *et al.* 2006) but their validity has been questioned; PCR/sequencing artifacts may lead to inflation of the 'rare biosphere' (Huse *et al.* 2010; Kunin *et al.* 2010; Quince *et al.* 2011). However, Zhan *et al.* (2013) sequenced aquatic communities and spiked the samples with known indicators to test sensitivity. They found that many singletons represented the spiked controls suggesting that not all singletons are artifacts.

To estimate the proportion of artifactual singletons and to test the origin of these singletons (NGS platform or PCR errors), we reanalyzed singletons from sixteen experiments that targeted three nuclear ribosomal RNA gene regions (LSU, ITS1, ITS2) from different sequencing technologies or chemistries (454-FLX, 454-Titanium, and Illumina-MiSeq; Table S1). These datasets included five ITS1 [454-FLX(3) and 454-Titanium(2)], six ITS2 (Illumina-MiSeq), and five Large Subunit variable region D1 (454-Titanium) libraries (see Table F.1 for primers and direction of sequencing). The datasets were analyzed using MOTHUR (v.1.32.1; Schloss *et al.* 2009), denoised (Quince *et al.* 2011), plus chimera- (UCHIME; Edgar *et al.* 2011) and sequencing-error screened (pre.cluster; Huse *et al.* 2010) prior to OTU binning at 97% similarity. After this quality control, ~ 50% of the OTUs were singletons, which we extracted into four fasta files (supplemental material) containing all comparable singleton sequences (ITS1-FLX, ITS1-Titanium, ITS2 and LSU). LSU libraries were aligned against a modified James *et al.* (2006) reference (Brown *et al.* 2014) and gaps removed prior to downstream analyses. Sequences were truncated to equal lengths and subsampled to equal numbers per library (Table F.1). Four MiSeq libraries were generated on split-reactions (EcM and Soil Fungi − Australia and EcM of Yellow Pine using two different polymerases) allowing differentiation among sequencing platform-generated artifacts from others.

Each singleton dataset was pairwise-aligned and resultant distance matrices clustered into OTUs at 97% similarity (using the MOTHUR implemented Average-Neighbor clustering

algorithm - UPGMA) to detect overlapping rare OTUs across libraries. It is important to note that the method of OTU binning can dramatically affect the generation of singletons: single-linkage clustering (nearest-neighbor in MOTHUR) produces fewer OTUs with higher average sequence dissimilarity within an OTU, whereas a complete-linkage clustering (furthest-neighbor in MOTHUR) produces more OTUs with higher sequence similarity within an OTU. Average-neighbor clustering (UPGMA) is a "middle ground" algorithm both in terms of OTU numbers and sequence similarity. After clustering, conserved regions (SSU, 5.8S, LSU) were removed from representative sequences for each ITS OTU (including singletons) using the online UNITE Phylogenetic Module ITSx using default online options with the exception that we set the minimal number of domains required to match for extraction to one (unite.ut.ee; Nilsson *et al.* 2010; Bengtsson-Palme *et al.* 2013). The extracted OTU sequences were assigned to taxa in MOTHUR using the Naïve Bayesian Classifier (Wang *et al.* 2007) with the RDP 28s rRNA reference (v.7) or with two ITS databases, Findley (ITS1; Findley *et al.* 2013) and UNITE plus INSD non-redundant ITS database (ITS1 and ITS2; Kõljalg *et al.* 2013). The Naïve Bayesian Classifier queries all non-overlapping 8-bp words (k-mers) against a reference dataset and provides bootstrap support estimates to taxonomic levels based on the number of times a queried sequence is placed in the same rank. OTUs were considered artifacts if: 1) OTUs were unclassified at a phylum level (many uncultured sequences may lack phylum level classification thus exaggerating proportion of artifact OTUs); 2) they did not classify to a phylum at 50% bootstrap support or higher; or, 3) the ITS sequences could not be mapped to ITS1 or ITS2 region (ITSx). Furthermore, sequences from the ITS1 libraries were considered artifacts if these conditions were met for taxonomy labels from both reference databases. Additionally, singletons were considered *potential artifacts* if they received < 50% bootstrap support at the family level. We report statistics on the proportion of singletons classified to all taxonomic levels at > 50%, 75%, and 90% bootstrap support (Table 7.1).

Many singletons from the sixteen libraries clustered at 97% with at least one other sequence at rates seemingly driven by gene region [LSU (Titanium) – 11.5%; ITS1 (FLX) – 0.83%; ITS1 (Titanium) – 0.43%; ITS2 (MiSeq) – 2.27%] reflecting variability of clustering efficiencies across gene regions. Singletons that clustered together often originated from within the same original library suggesting that they are a result of algorithm performance that provides

non-exact clustering solutions. The more conserved LSU likely performs better with these algorithms.

We queried our sequences against databases to estimate assignment robustness through bootstrapping. Overall, the proportion of artifact sequences (<50% support for phylum level classification) was much lower (12.94% - 19.10%; Table 7.1) than expected based on previous estimates suggesting that ~80% of singletons may be artifacts (Tedersoo *et al*. 2010). This is unexpected: our liberal inclusion of unclassified phyla as artifacts likely inflated the number of artifact singletons. The combined proportion of artifacts and *potential artifacts* was largely affected by region: LSU (54.80%) had a greater proportion of questionable sequences than ITS regions (Table 7.1). Interestingly, many sequences that were not considered artifacts or *potential artifacts* were assigned to lower taxonomic levels with high bootstrap support. The proportion of sequences with a genus-level affinity with ≥90% bootstrap support ranged from 10.53%-44.14%, a level of support unlikely for true artifacts.

Our analyses, similarly to Tedersoo *et al.* (2010), indicate that many singletons are likely artifacts. However, our estimates are less than half of the ~80% estimate of Tedersoo and coworkers. There are many underlying reasons for this discrepancy. The early 454-datasets explored how to analyze NGS data (*e.g.,* Buee *et al.* 2009; Jumpponen & Jones 2009; Tedersoo *et al.* 2010). Lessons from those analyses have led to recommendations on tools to utilize NGS data in fungal ecology (Nilsson *et al*. 2011; Lindahl *et al,* 2013), including adoption of denoising (Quince *et al.* 2011), standard chimera removal (Edgar *et al*. 2011) and preclustering (Huse *et al*. 2010). Noteworthy is that Tedersoo *et al.* included a BLAST-based post-hoc chimera check. However, this method is less sensitive as it relies on database accession quality, whereas pre-OTU binning methods (UCHIME; Edgar *et al.* 2011) rely on NGS-acquired data itself. Additionally, our study differs in other important ways; we neither had anchor taxa from the same samples nor performed the detailed phylogenetic analyses. Instead, we relied on the Naïve Bayesian Classifier, an approach that parallels the phylogenetic approach. Nonetheless, our results highlight the importance of appropriate quality controls to minimize artifacts.

Many 'global singleton' sequences clustered into new non-singleton OTUs. Whilst the underlying reasons remain unclear, we suggest two primary explanations. First, fungal communities are hyper-diverse (Jumpponen & Jones 2009), include large numbers of low frequency taxa, and are locally or regionally distinct (Meiser *et al.* 2014). Second, clustering

relies on imperfect heuristic algorithms that permit non-exact solutions for OTU membership, especially in large and complex datasets. This allows stochastic OTU memberships and sequences may be placed into different OTUs each time a dataset is clustered.

Our results suggest that half of the singletons may represent true target taxa. However, we cannot determine if artifact singletons result from sequencing platform errors. Singletons may also represent off-target amplification as evidenced by the common occurrence of sequences, from which ITS regions could not be extracted with ITSx. A surprisingly high proportion of queried sequences had no extractable ITS regions (5.33% for ITS1-FLX; 2.86% for ITS1-Titanium; 4.53% for ITS2-MiSeq; Table F.2). Similar proportions of non-target LSU sequences are likely but tools to evaluate this were not explored here. Interestingly, absence of extractable ITS regions were not solely due to non-target amplification: many discarded sequences were fungal, although no ITS regions could be excised using ITSx. More than 90% of our ITS1-FLX and all of our ITS1-Titanium sequences that failed to extract were fungal based on BLASTn analyses (see Table F.3 for complete list of sequences that failed to be extracted using ITSx and the best BLASTn taxonomic strings). ITS2 had the highest non-target amplification: 61.03% of the sequences that failed to be extracted were not fungal (Table F.3). Additional sequences failed to extract that were actually ITS2 fungal sequences. Peculiarly, all but two of the fungal sequences discarded because of failed ITS2 extraction belonged to Agaricomycetes (primarily Russulaceae and Thelephoraceae) suggesting that the Hidden Markov Models (HMM) in ITSx may fail to recognize this class fully. Alternatively, this could be explained by insufficient 5' LSU length upstream of the priming site causing the HMMs to fail for some Agaricomycetes. The remaining artifacts are likely PCR errors - polymerase mis-pairs, deletions, or insertions (Eckert & Kunkel 1991) and chimeras that evaded detection.

To investigate if these singletons represent true biological or artificial variability (platform specific variability, indels due to polymerase slippage, or homopolymeric reads), we aligned singletons against representative sequences of the 100 most abundant OTUs from the original datasets. The mismatches among singletons and the representative sequences of the common OTUs generated on 454 and Illumina platforms appeared stochastically distributed across the alignments suggesting that they were unlikely a result of poor read quality in the read termini. Singletons generated using 454 technologies differed from abundant OTUs frequently because of inconsistent homopolymer lengths and/or single nucleotide differences. In contrast to

454-sequencing, differences in the Illumina-generated singletons were most often nucleotide differences with no evidence of inconsistent homopolymer lengths. Based on these findings it is impossible to determine the source of the variability as polymerase slippage, suboptimal platform performance or true biological variability could result in similar outcomes.

Removal of rare sequences may underestimate observed and extrapolated richness (Unterseher *et al.* 2011). Rare taxa also affect community pairwise distances commonly visualized by ordination tools. Conversely, singleton exclusion may minimally affect community composition (Shade *et al.* 2013) or multivariate analyses (Gobet *et al.* 2010; Lindahl *et al.* 2013). Although removal of singletons may not substantially affect the visualization of community composition, rare sequences may be necessary for more accurate biodiversity estimates, if they represent real taxa but biodiversity estimates from sequence data are capricious (Haegeman *et al.* 2013).

We conclude that for most hypothesis-driven experiments that compare experimental conditions, rare taxa present a minor issue: excluding them unlikely sways strong community responses. However, if estimation of biodiversity is crucial, careful manual examination and annotation of the infrequent sequences is required. One must strike a balance: is it better to err on the side of caution and throw the baby out with the bathwater (exclude rare sequences) or to analyze the rare sequences and scrape the bottom of large pools of sequence data to account for every last unculturable fungus that occurs in the data if even only once? Due to the minimal effect these rare sequences have in community analyses, we concur with previous suggestions to remove all singletons and expand this recommendation to remove other highly rare (n=10) sequences in datasets as modern sequencing depth allows for such stringent practices.

Acknowledgements

References

Bengtsson-Palme J., Ryberg M., Hartmann M., Branco S., Wang Z., Godhe A., De Wit P.,
Sánchez-García M., Edersberger I., de Sousa F., Amend A.S., Jumpponen A., Unterseher
M., Kristiansson E., Abarenkov K., Bertrand Y.J.K., Sanli K., Eriksson K.M., Vik U.,
Veldre V., Nilsson R.H., 2013. Improved software detection and extraction of ITS1 and
ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of
environmental sequencing data. *Methods in Ecology and Evolution* **4**: 914-919.

Brown S.P., Rigdon-Huss A.R., Jumpponen A., 2014. Analyses of ITS and LSU gene regions
provide congruent results on fungal community responses. *Fungal Ecology* **9**: 65-68**.**

Buée M., Reich M., Murat C., Morin E., Nilsson R.H., Uroz S., Martin F., 2009. 454
Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity.
New Phytologist **184**: 449-456.

Caporaso J.G., Lauber C.L., Walters W.A., Berg-Lyons D., Huntley J., Fierer N., Owens S.M.,
Betley J., Fraser L., Bauer M., Gormley N., Gilbert J.A., Smith G., Knight R., 2012.
Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq
platforms. *The ISME Journal* **6**: 1621-1624.

Eckert K.A., Kunkel T.A., 1991. DNA polymerase fidelity and the polymerase chain reaction.
Genome Research **1**: 17-24.

Edgar R.C., Haas B.J., Clemente J.C., Quince C., Knight R., 2011. UCHIME improves
sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194-2200.

Findley K., Oh J., Yang J., Conlan S., Deming C., Meyer J.A., Schoenfeld D., Nomicos E., Park
M, NIH Intramural Sequencing Center Comparative Sequencing Program, Kong H.H.,
Segre J.A., 2013. Topographic diversity of fungal and bacterial communities in human
skin. *Nature* **498**: 367-370.

Gobet A., Quince C., Ramette A., 2010. Multivariate cutoff levels analysis (MultoCoLA) of
large community data sets. *Nucleic Acid Research* **38**(15): e155.

Haegeman B., Hamelin J., Moriarty J., Neal P., Dushoff J., Weitz J.S., 2013. Robust estimation
of microbial diversity in theory and practice. *The ISME Journal* **7**: 1092-1101.

Hibbett, D.S., Ohman, A., Kirk, P.M., 2009. Fungal ecology catches fire. *New Phytologist* **184**:
279-282.

Huse S.M., Welch D.M., Morrison H.G., Sogin M.L., 2010. Ironing out the wrinkles in the rare
biosphere through improved OTU clustering. *Environmental Microbiology* **12**: 1889-
1898.

Jumpponen A., Jones K.L., 2009. Massively parallel 454 sequencing indicated hyperdiverse
fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytologist*
184: 438-448.

Kõljalg U., Nilsson R.H., Abarenkov K., Tedersoo L., Taylor A.F.S., Bahram M., Bates S.T.,
Bruns T.D., Bengtsson-Palme J., Callaghan T.M., Douglas B., Drenkhan T., Eberhardt
U., Dueñas M., Grebenc T., Griffith G.W., Hartmann M., Kirk P.M., Kohout P., Larsson
E., Lindahl B.D., Lücking R., Nguyen N.H., Niskanen T., Oja J., Peay K.G., Peintner U.,
Peterson M., Põldmaa K., Saag L., Saar I., Schüßler A., Scott J.A., Senés C., Smith M.E.,
Suija A., Taylor D.L., Telleria M.T., Weiss M., Larsson K., 2013. Toward a unified
paradigm for sequence-based identification of fungi. *Molecular Ecology* **22**: 5271-5277.

Kunin V., Engelbrektson A., Ochman H., Hugenholtz P., 2010. Wrinkles in the rare biosphere: pyrosequencing errors lead to artificial inflation of diversity estimates. *Environmental Microbiology* **12**: 118-123.

Lindahl B.D., Nilsson R.H., Tedersoo L., Abarenkov K., Carlsen T., Kjøller R., Kõljalg U., Pennanen T., Rosendahl S., Stenlid J., Kauserud H., 2013. Fungal community analysis by high-throughput sequencing of amplified markers – a user's guide. *New Phytologist* **199**: 288-299.

Meiser A., Bálint M., Schmitt I., 2014. Meta-analysis of deep-sequenced fungal communities indicates limited taxon sharing between studies and the presence of biogeographic patterns. *New Phytologist* **201**: 623-635.

Nilsson R.H., Veldre V., Hartmann M., Unterseher M., Amend A., Bergsten J., Kristiansson E., Ruberg M., Jumpponen A., Abarenkov K., 2010. An open source software package for automated extraction of ITS1 and ITS2 from fungal ITS sequences for use in high-throughput community assays and molecular ecology. *Fungal Ecology* **3**: 284-287.

Nilsson R.H., Tedersoo L., Lindahl B.D., Kjøller R., Carlsen T., Quince C., Abarenkov K., Pennanen T., Stenlid J., Bruns T., Larsson K., Kõljalg U., Kauserud H., 2011. Towards standardization of the description and publication of next-generation sequencing datasets of fungal communities. *New Phytologist* **191**: 314-318.

Porras-Alfaro A., Liu K., Kuske C.R., Xie G., 2014. From genus to phylum: large-subunit and internal transcribed spacer rRNA operon regions show similar classification accuracies influenced by database composition. *Applied and Environmental Microbiology* **80**: 829-840.

Quince C., Lanzen A., Davenport R.J., Turnbaugh P.J., 2011. Removing noise from Pyrosequenced amplicons. *BMC Bioinformatics* **12**: 38.

Schloss P.D., Westcott S.L., Ryabin T., Hall J.R., Hartmann M., Hollister E.B., Lesniewski R.A., Oakley B.B., Parks E.B., Robinson C .J., Sahl J.W., Stres B., Thallinger G.G., Van Horn D.J., Weber C.F., 2009. Introducing mothur: open-source, platform independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**: 7537-7541.

Shade A., Hogan C.S., Klimowicz A.K., Linske M., McManus P.S., Handelsman J., 2012. Culturing captures members of the soil rare biosphere. *Environmental Microbiology* **14**: 2247-2252.

Sogin M.L., Morrison H.G., Huber J.A., Welch D.M., Huse A.M., Neal P.R., Arrieta J.M., Herndl G.J., 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proceedings of the National Academy of Science* **103**: 12115-12120.

Tedersoo L., Nilsson R.H., Abarenkov K., Jairus T., Sadam A., Saar I., Bahram M., Bechem E., Chuyong G., Kõljalg U., 2010. 454 pyrosequencing and sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist* 188: 291-301.

Unterseher M., Jumpponen A., Öpik M., Tedersoo L., Moora M., Dormann C.F., Schnittler M., 2011. Species abundance distributions and richness estimations in fungal metagenomics – lessons learned from community ecology. *Molecular Ecology* 20: 275-285.

Wang Q., Garrity G.M., Tiedje J.M. Cole J.R., 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into new bacterial taxonomy. *Applied and Environmental Microbiology* 73: 5267-5267.

Zhan A., Hulák M., Sylvester F., Huang X., Adebayo A.A., Abbot, C.L., Adamowicz S.J., Heath D.D., Cristescu M.E., Maclsaac H.J., 2013. High-sensitivity pyrosequencing of detection of rare species in aquatic communities. *Methods in Ecology and Evolution* 4: 58-565.

Tables and Figures

Table 7.1 Percentage of singletons that are artifacts and potential artifacts as well as the percentage of non-artifactual OTUs than are assigned to taxa above 50%, 75% and 90% bootstrap support on all levels of taxonomic levels.

	LSU-Titanium	ITS1-FLX	ITS1-Titanium	ITS2- MiSeq
Percentage of Artifacts	16.87%	12.94%	13.34%	19.10%
Percentage of Potential Artifacts	37.93%	21.67%	13.29%	17.20%
Percentage of Sequences Above Bootstrap Support Thresholds				
Phylum (90%)	67.80%	71.67%	74.00%	64.27%
Phylum (75%)	69.80%	80.17%	79.86%	69.50%
Phylum (50%)	73.60%	86.33%	86.29%	79.07%
Class (90%)	48.27%	62.67%	63.71%	58.60%
Class (75%)	55.60%	70.00%	71.57%	63.77%
Class (50%)	63.40%	76.83%	79.29%	70.23%
Order (90%)	32.73%	52.82%	58.86%	53.80%
Order (75%)	44.07%	61.33%	67.00%	60.33%
Order (50%)	56.53%	68.17%	77.00%	66.23%
Family (90%)	20.00%	48.00%	51.71%	47.40%
Family (75%)	32.40%	56.17%	60.71%	56.07%
Family (50%)	47.13%	65.50%	73.43%	64.13%
Genus (90%)	10.53%	39.17%	44.14%	37.30%
Genus (75%)	18.07%	51.17%	55.57%	48.97%
Genus (50%)	36.80%	61.33%	70.43%	61.33%

References

Chapter 1

Begon M, Townsend CR, Harper JL. 2006. *Ecology: from individuals to ecosystems.* Blackwell Publishing, Oxford, UK.

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

Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL. 2008. Microbes on mountainsides: contrasting elevational patterns of bacteria and plant diversity. *Proceedings of the National Academy of Science, USA* 105: Supplement 1 11505-11511.

Buee M, Reich M, Murat C, Morin E, Nilsson RH, Uroz S, Martin F. 2009. 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist* 184: 449-456.

Carpenter EJ, Lin SJ, Capone DG. 2000. Bacterial activity in South Pole snow. *Applied and Environmental Microbiology,* 66: 4514-4517.

Clements FE. 1916. *Plant succession: an analysis of the development of vegetation.* Carnegie Institution of Washington, Washington, D.C., USA

del Moral R, Wood DM. 1993. Early primary succession on a barren volcanic plain at Mount St. Helens, Washington. *American Journal of Botany* 80: 981-991.

Diamond JM. 1975. Assembly of species communities. In: Cody ML, Diamond JM (Eds.) *Ecology and evolution of communities*. Harvard University Press, Cambridge, MA, USA, pp. 342-344.

Fierer N, McCain CM, Meir P, Zimmermann M, Rapp JM, Silman MR, Knight R. 2011. Microbes do not follow the elevational diversity patterns of plants and animals. *Ecology* 92: 797-804.

Fujii M, Takano Y, Kojima H, Hoshino T, Tanaka R, Fukui M. 2010. Microbial community structure, pigment composition, and nitrogen source of red snow in Antarctica. *Microbial Ecology.* 59: 466-475.

Hell K, Edwards A, Zarsky J, Posmirseg SM, Girdwood S, Pachebat JA, Insam H, Sattler B. 2013. The dynamic bacterial communities of a melting High Arctic glacier snowpack. *The ISME Journal,* 7: 1814-1826.

Jumpponen A. 2003. Soil fungal community assembly in a primary successional glacier forefront ecosystem as inferred from analyses of rDNA sequences. *New Phytologist* 158: 569-578.

Jumpponen A, Mattson K, Trappe JM, Ohtonen R. 1998. Effects of established willows on primary succession on Lyman Glacier forefront, North Cascade Range, Washington, U.S.A.: evidence for simultaneous canopy inhibition and soil facilitation. *Arctic and Alpine Research* 30: 31-39.

Jumpponen A, Trappe JM, Cázares E. 1999. Ectomycorrhizal fungi in Lyman Lake basin: a comparison between primary and secondary successional sites. *Mycologia* 91: 575-582.

Jumpponen A, Egerton-Warburton LM. 2005. Mycorrhizal fungi in successional environments – a community assembly model incorporating host plant,

environmental and biotic filters. In: Dighton J, White JF, Oudemans P (Eds.) *The fungal community*. CRC Press, New York, NY, USA, pp. 139-180.

Jumpponen A, Jones KL. 2009. Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytologist* 184: 438-448.

Jumpponen A, Brown SP, Trappe JM, Cázares E, Strömmer R. 2012. Twuenty years of research on fungal-plant interactions on Lyman Glacier forefront – lessons learned and questions yet unanswered. *Fungal Ecology* 5: 430-442.

Keddy PA. 1992. Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science* 3: 157-164.

Kiss L. 2012. Limits of nuclear ribosomal DNA internal transcribed spacer (ITS) sequences as species barcodes for Fungi. *Proceeding of National Academy of Science, USA* 109: E1811.

Koide RT, Fernandez C, Petprakob K. 2011. General principles in the community ecology of ectomycorrhizal fungi. *Annals of Forest Science* 68: 45-55.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, *et al.* 2005. Genome sequencing in microfabricated high-density picoliter reactors. *Nature* 437: 376-380.

Matthews JR. 1914. The white moss loch: a study in biotic succession. *New Phytologist* 13: 134-148.

Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson K, Kõljalg U. 2006. Taxonomic reliability of DNA sequences in public sequence databases : a fungal perspective. *PLOS One* 1: e59. doi:10.1371/journal.pone.0000059

Ohtonen R, Fritze H, Pennanen T, Jumpponen A, Trappe JM. 1999. Ecosystem properties and microbial community changes in primary succession on a glacier forefront. *Oecologia* 199: 239-246.

Pearsall WH. 1918. On the classification of aquatic plant communities. *Journal of Ecology* 6: 75-84.

Remias D, Lütz-Meindl U, Lütz C. 2005. Photosynthesis, pigments and ultrastructure of the alpine snow alga *Chlamydomonas nivalis*. *European Journal of Phycology.* 40: 259-268.

Reynolds HL, Packer A, Bever JD, Clay K. 2003. Grassroots ecology: plant-microbe-soil interactions as drivers of plant community structure and dynamics. *Ecology* 84: 2281-2291.

Schoch CL, Seifert KA, Huhdorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences, USA,* 109: 6241-6246.

Tansley AG. 1920. The classification of vegetation and the concept of development. *Journal of Ecology* 8: 118-149.

Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, Saar I, Bahram M, Bechem E, Chuyong G, Kõljalg U. 2010. 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist* 188: 271-301.

Walker LR, del Moral R. 2003. *Primary succession and ecosystem rehabilitation.* Cambridge University Press, New York, NY, USA.

Chapter 2

Ashkannejhad S, Horton TR (2006) Ectomycorrhizal ecology under primary succession on coastal sand dunes: interactions involving *Pinus contorta*, suilloid fungi and deer. *New Phytologist* **169**, 345-354.

Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* **55**, 541-555.

Barry RG (2006) The status of research on glaciers and global glacier recession: a review. *Progress in Physical Geography* **30**, 285-306.

Blaalid R, Carlsen T, Kumar S, Halvorsen R, Ugland KI, Fontana G, Kauserud H (2012) Changes in the root-associated fungal communities along a primary successional gradient analysed by 454 pyrosequencing. *Molecular Ecology* **21**, 1897-1908.

Cazarés E (1992) Mycorrhizal fungi and their relationship to plant succession in subalpine habitats. Ph.D. Dissertation, Oregon State University, Corvallis, Oregon. 126 pp.

Cazarés E, Trappe JM, Jumpponen A (2005) Mycorrhiza-plant colonization patterns on a subalpine glacier forefront as a model system of primary succession. *Mycorrhiza* **15**, 405-416.

Chapin FS, Walker LR, Fastie CL, Sharman LC (1994) Mechanisms of primary succession following deglaciation at Glacier Bay, Alaska. *Ecological Monographs* **64**, 149-175.

Colwell RK (2006) EstimateS: Statistical estimation of species richness and shared species from samples. Version 8.0.0. User's Guide and application published at: http://viceroy.eeb.uconn.edu/estimates/

Cramer W, Bondeau A, Woodward FI, Prentice IC, Betts RA, Brovkin V, Cox PM, Fisher V, Foley JA, Friend AD, Kucharik C, Lomas MR, Ramankutty N, Sitch S, Smith B, White A, Young-Molling C (2001) Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models. *Global Change Biology* **7:** 357-373.

delMoral R, Bliss LC (1993) Mechanisms of primary succession – insights resulting from the eruption of Mount St. Helens. *Advances in Ecological Research, Vol 24* **24**, 1-66.

delMoral R, Titus JH, Cook AM (1995) Early primary succession on Mount St. Helens, Washington, USA. *Journal of Vegetation Science* **6**, 107-120.

delMoral R (2009) Increasing deterministic control of primary succession on Mount St. Helens, Washington. *Journal of Vegetation Science* **20**, 1145-1154.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26,** 2460-2461.

Fierer N, Jackson JA, Vilgalys R, Jackson RB (2005) Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Applied and Environmental Microbiology* **71**, 4117-4120.

Freeman KR, Martin AP, Karki D, Lynch RC, Mitter MS, Meyer AF, Longcore JE, Simmons DR, Schmidt SK (2009) Evidence that chytrids dominate fungal communities in high-elevation soils, *Proceedings of the National Academy of Sciences of the United States of America* **106,** 18315-18320.

Freeman KR, Pescador MY, Reed SC, Costello EK, Robeson MS, Schmidt SK (2009) Soil CO$_2$ flux and photoautotrophic community composition in high-elevation, 'barren' soil. *Environmental Microbiology* **11,** 674-686.

Gardes M, Bruns TD (1993) ITS primers with enhanced specificity for Basidiomycetes – Application to the identification of mycorrhizae and rusts. *Molecular Ecology* **2**, 113-118.

Helm DJ, Allen EB, Trappe JM (1996) Mycorrhizal chronosequence near Exit Glacier, Alaska. *Canadian Journal of Botany* **74,** 1496-1506.

Hodkinson ID, Coulson SJ, Harrison J, Webb NR (2001) What a wonderful web they weave: spiders, nutrient capture and early ecosystem development in the high Arctic – some counter-intuitive ideas on community assembly. *Oikos* **95**, 349-352.

Hodkinson ID, Webb NR, Coulson SJ (2002) Primary community assembly on land – the missing stages: why are the heterotrophic organisms always there first? *Journal of Ecology* **90**, 569-577.

Jumpponen A (2003) Soil fungal community assembly in a primary successional glacier forefront ecosystem as inferred from rDNA sequence analyses. *New Phytologist* **158**, 569-578.

Jumpponen A, Egerton-Warburton LM (2005) Mycorrhizal fungi in successional environments – a community assembly model incorporating host plant, environmental and biotic filters. *In:* The Fungal Community, ed. Dighton J, White JF, Oudemans P. CRC Press, New York 139-180.

Jumpponen A, Mattson K, Trappe JM, Ohtonen R (1998) Effects of established willows on primary succession on Lyman Glacier forefront, North Cascade Range, Washington, USA: Evidence for simultaneous canopy inhibition and soil facilitation. *Arctic and Alpine Research* **30**, 31-39.

Jumpponen A, Trappe JM, Cazares E (1999) Ectomycorrhizal fungi in Lyman Lake Basin: a comparison between primary and secondary successional sites. *Mycologia* **91**, 575-582.

Jumpponen A, Trappe JM, Cazares E (2002) Occurrence of ectomycorrhizal fungi on the forefront of retreating Lyman Glacier (Washington, USA) in relation to time since deglaciation. *Mycorrhiza* **12**, 43-49.

Kaštovská K, Stibal M, Šabacká M, Černá B, Šantrůčková H, Elster J (2007) Microbial community structure and ecology of subglacial sediments in two polythermal Svalbard glaciers characterized by epifluorescence microscopy and PLFA. *Polar Biology* **30,** 227-287.

Knelman, JE, Legg, TM, O'Neill, SP, Washenberger, CL, Gonzalez, A, Cleveland, CC, Nemergut DR (2012) Bacterial community structure and function change in association with colonizer plants during early primary succession in a glacier forefield. *Soil Biology & Biochemistry* **46**, 172-180.

Kunin V, Hugenholtz P (2010) PyroTagger: A fast, accurate pipeline for analysis of rRNA amplicon pyrosequence data. *The Open Journal* 1. Feb 2010.

Lane D (1991) 16s/23s rRNA sequencing, p. 115-175. *In* E. Stackebrant and M Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, West Sussex, United Kingdom.

McCune B, Mefford JM (1999) PC-ORD: Multivariate analysis of ecological data. Version 4. MjM Software, Gleneden Beach, Oregon, USA.

Muyzer G, de Wall EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied Environmental Microbiology* **59:** 695-700.

Nara K, Nakaya H, Wu BY, Zhou ZH, Hogetsu T (2003) Underground primary succession of Ectomycorrhizal fungi in a volcanic desert on Mount Fuji. *New Phytologist* **159:** 734-756.

Oehl F, Schneider D, Sieverding E, Burga CA (2011) Succession of arbuscular mycorrhizal communities in the foreland of the retreating Morteratsch glacier in the Central Alps. *Pedobiologia* **54**, 321-331.

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

Schmidt SK, Reed SC, Nemergut DR, Grandy AS, Cleveland CC, Weintraub MN, Hill AW, Costello EK, Meyer AF, Neff JC, Martin AM (2008) The earliest stages of ecosystem succession in high-elevation (5000 meters above sea level), recently deglaciated soils. *Proceedings of the Royal Society B-Biological Sciences*. **275:** 2793-2802.

Schutte UME, Abdo Z, Foster J*, et al.* (2010) Bacterial diversity in a glacier foreland of the high Arctic. *Molecular Ecology* **19**, 54-66.

Sigler WV, Zeyer J (2002) Microbial diversity and activity along the forefields of two receding glaciers. *Microbial Ecology* **43**, 397-407.

Sikes BA, Maherali H, Klironomos JN (2012) Arbuscular mycorrhizal fungal communities change among three stages of primary sand dune succession but do not alter plant growth. *Oikos* **121**, 1791-1800.

Strauss SL, Garcia-Pichel F, Day TA (2012) Soil microbial carbon and nitrogen transformations at a glacial foreland on Anvers Island, Antarctic Peninsula. *Polar Biology* **35**, 1459-1471.

Strickland MS, Rousk J (2012) Considering fungal:bacterial dominance in soils – methods, controls, and ecosystems implications. *Soil Biology and Biochemistry* **42,** 1385-1395.

Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, Saar I, Bahram M, Bechem E, Chuyong G, Koljalg U (2010) 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist* **188**, 291-301.

Tejesvi MV, Ruotsalainen AL, Markkola AM, Pirttila AM (2010) Root endophytes along a primary succession gradient in northern Finland. *Fungal Diversity* **41**, 125-134.

Trowbridge J, Jumpponen A (2004) Fungal colonization of shrub willow roots at the forefront of a receding glacier. *Mycorrhiza* **14**, 283-293.

Tscherko D, Hammesfahr U, Zeltner G, Kandeler E, Bocker R (2005) Plant succession and rhizosphere microbial communities in a recently deglaciated alpine terrain. *Basic and Applied Ecology* **6**, 367-383.

Unterseher M, Jumpponen A, Öpik M, Tedersoo L, Moora M, Dormann CF, Schnittler M (2011) Species abundance distributions and richness estimations of fungal metagenomics – lessons learned from community ecology. *Molecular Ecology* **20,** 275-285.

Vilgalys R, Hester M (1990) Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. *Journal of Bacteriology* **172:** 4238-4246.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**, 5261-5267.

Zumsteg A, Luster J, Göransson H, Smittenberg RH, Brunner I, Bernasconi SM, Zeyer J, Frey B (2012) Bacterial, archaeal, and fungal succession in the forefield of a receding glacier. *Microbial Ecology* **63**, 552-564.

Chapter 3

Amend AS, Seifert KA, Samson R, Bruns TD. (2010). Indoor fungal composition is geographically pattered and more diverse in temperate zones than in the tropics. Proceedings of the National Academy of Sciences 107: 13748-13753.

Anderson TM, Lachance M, Starmer WT. (2004). The relationship of phylogeny to community structure: the cactus yeast community. The American Naturalist 164: 709-721.

Berry D, Mahfoudh KB, Wagner M, Loy A. (2011). Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Applied and Environmental Microbiology 77: 7846-7849.

Blaalid R, Carlsen T, Kumar S, Halvorsen R, Ugland KI, Fontana G, Kauserud H. (2012). Changes in the root-associated fungal communities along a primary succession gradient analyzed by 454 pyrosequencing. Molecular Ecology 21: 1897-1908.

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

Brown SP, Rigdon-Huss AR, Jumpponen A. (2014a). Analyses if ITS and LSU gene regions provide congruent results on fungal community responses. Fungal Ecology 9: 65-68.

Brown SP, Veach AM, Rigdon-Huss AR, Grond K, Lickteig SK, Lothamer K, Oliver AK, Jumpponen A. (2014b). Scrapping the bottom of the barrel: are rare high throughput sequences artifacts? Fungal Ecology (in review).

Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL. (2008). Microbes on mountainsides: contrasting elevational patterns of bacteria and plant diversity. Proceedings of the National Academy of Science, USA 105(s1): 11505-11511.

Cardillo M. (2011). Phylogenetic structure of mammal assembleges at large geographical scales: linking phylogenetic community ecology with macroecology. Phil. Trans. R. Soc. B 366: 2545-2553.

Cázares E, Trappe JM, Jumpponen A. (2005). Mycorrhiza-plant colonization patterns on a subalpine glacier forefront as a model system of primary succession. Mycorrhiza 15: 405-416.

Connor EF, Simberloff D. (1979). The assembly of species communities: chance or competition? Ecology 60: 1132-1140.

Cutler NA, Chaput DL, van der Gast, CJ. (2014). Long-term changes in soil microbial communities during primary succession. Soil Biology & Biochemistry 69: 359-370.

Edgar RC, Hass BJ, Clemete JC, Quince C, Knight R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27: 2194-2220.

Faith DP. (1992). Conservation evaluation and phylogenetic diversity. Biological Conservation 61: 1-10.

Fierer N, Bradford MA, Jackson RB. (2007). Toward and ecological classification of soil bacteria. Ecology 88:1354-1364.

Fierer N, Nemergut D, Knight R, Craine JM. (2010). Changes through time: integrating microorganisms into the study of succession. Research in Microbiology 161:635-642.

Garnica S, Riess K, Bauer R, Oberwinkler F, Weiß M. (2013). Phylogenetic diversity and structure of sebacinoid fungi associated with plant communities along an altitudinal gradient. FEMS Microbiology Ecology 83: 265-278.

Gihring TM, Green AJ, Schady CW. (2012). Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. Environmental Microbiology 14: 285-290.

Gotelli NJ, Entsminger GL. (2003). Swap algorithms in null model analysis. Ecology 84: 532-535.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, et al. (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443: 818-822.

Jangid K, Whiteman WB, Condron LM, Turner BL, Williams MA. (2013). Progressive and retrogressive ecosystem development coincide with soil bacterial community change in a dune system under lowland temperate rainforest in New Zealand. Plant and Soil 367: 235-247.

Jumpponen A. (2003). Soil fungal community assembly in a primary successional glacier forefront ecosystem as inferred from rDNA sequence analyses. New Phytologist 158: 569-578.

Jumpponen A, Brown SP, Trappe JM, Cázares E, Strömmer R. (2012). Twenty years of research of fungal-plant interactions on Lyman Glacier forefront – lessons learned and questions yet unanswered. Fungal Ecology 5: 430-442.

Jumpponen A, Egerton-Warburton LM. (2005). Mycorrhizal fungi in successional environments – a community assembly model incorporating host plant, environment and biotic filters. In: Dighton J, White JF, Oudemans P (eds). The Fungal Community. CRC Press, New York, pp 139-180.

Jumpponen A, Mattson K, Trappe JM, Ohtonen R. (1998). Effects of established willows on primary succession on Lyman Glacier forefront, North Cascade Range, Washington, USA: Evidence for simultaneous canopy inhibition and soil facilitation. Arctic and Alpine Research 30: 31-39.

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blonberg SP, Webb CO. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26: 1463-1464.

Knelman JE, Legg TM, O'Neill SP, Washenberger CL, González A, Cleveland CC, Nemergut DR. (2012). Bacterial community structure and function change in association with colonizer plants during early primary succession on a glacier forefield. Soil Biology & Biochemistry 46: 172-180.

Knelman JE, Schmidt SK, Lynch RC, Darcy JL, Castle SC, Cleveland CC, Nemergut DR. (2014). Nutrient addition dramatically accelerates microbial community succession. PLOS one 9: e102609.

Koide RT, Fernandez C, Petprakob K. (2011). General principles in the community ecology of ectomycorrhizal fungi. Annals of Forest Science 68: 45-55.

Lozupone CA. (2007). Global patterns of bacterial diversity. Ph.D. Dissertation, University of Colorado, Boulder, Colorado. 182 pp.

Maherali H, Klironomos JN. (2007). Influence of phylogeny on fungal community assembly and ecosystem functioning. Science 316: 1746-1748.

Merckx VSFT, Hanssens SB, Hynson NA, Specht CD, Bruns TD, Smets EF. (2012). Mycoheterothophic interactions are not limited to a narrow phylogenetic range of arbuscular mycorrhizal fungi. Molecular Ecology 21: 1524-1532.

Nemergut DR, Anderson AP, Cleveland CC, Martin AP, Miller AE, Seimon A, Schmidt SK. (2007). Microbial community succession in an unvegetated, recently deglaciated soil. Microbial Ecology 53: 110-122.

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

119: 239-246.

Philippot L, Andersson SGE, Battin TJ, Prosser JI, Schimel JP, Whitman WB, Hallin S. (2010). The ecological coherence of high bacterial taxonomic ranks. Nature Reviews Microbiology 8: 523-529.

Purschke O, Schnid BC, Sykes MT, Poschlod P, Michalski SG, Durka W, Kühn I, Winter M, Prentice HC. (2013). Contrasting changes in taxonomic, phylogenetic and functional diversity during a long-term succession: insights into assembly processes. Journal of Ecology 101: 857-866.

Quice C, Lanzén A, Curtis TP, Davenport RJ, Hall N, Head IM, Read LF, Sloan WT. (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. Nature Methods 6: 639-641.

R Development Core Team. (2007) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rämä T, Nordén J, Davey ML, Mathiassen GH, Spatafora JW, Kauserud H. (2014). Fungi ahoy! Diversity on marine wooded substrata in the high North. Fungal Ecology 8: 46-58.

Rincón A, Santamaría BP, Ocaña L, Verdú M. (2014). Structure and phylogenetic diversity of post-fire ectomycorrhizal communities of maritime pine. Mycorrhza 24: 131-141.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley, BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75: 7537-7541.

Schmidt AK, Reed AC, Nemergut DR, Grandy AS, Cleveland CC, Weintraub MN, Hill AW, Costello EK, Meyer AF, Neff JC Martin AM. (2008). The earliest stages of ecosystem succession in high-elevation (5000 metres above sea level), recently deglaciated soils. Proceedings of the Royal Society B 275: 2793-2802.

Schütte UME, Abdo Z, Foster J, Ravel J, Bunge J, Solheim B, Forney LJ. (2010). Bacterial diversity in a glacier foreland of the high Arctic. Molecular Ecology 19(s1): 54-66.

Sheneman L, Evens J, Foster JA. (2006). Clearcut: a fast implementation of relaxed neighbor joining. Bioinformatics 22: 2823-2824.

Sigler WV, Zeyer J. (2002). Microbial diversity and activity along the forefields of two receding glaciers. Microbial Ecology 43: 397-407.

Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, Saar I, Bahram M, Bechem E, Chuyong G, Kõljalg U. (2010). 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. New Phytologist 188: 291-301.

Tscherko D, Hammesfahr U, Zeltner G, Kandeler E, Böcker R. (2005). Plant succession and rhizosphere microbial communities in a recently deglaciated alpine terrain. Basic and Applied Ecology 6: 367-383.

Vamosi SM, Heard B, Vamosi JC, Webb CO. (2009). Emerging patterns in the comparative analysis of phylogenetic community structure. Molecular Ecology 18: 572-592.

Vitousek PM, Walker LR, Whiteaker LD, Mueller-Dombois D, Matson AP. (1987). Biological invasion my *Myrica faya* alters ecosystem development in Hawaii. Science 238: 802-804.

Vitousek PM, Walker LR, Whiteaker LD, Matson AP. (1993). Nutrient limitations to plant growth during primary succession in Hawaii Volcanoes National Park. Biogeochemistry 23: 197-215.

Walker LR, de Moral R. (2003). Primary succession and ecosystem rehabilitation. Cambridge University Press, Cambridge, UK.

Wang J, Soininen J, He J, Shen J. (2012). Phylogenetic clustering increases with elevation for microbes. Environmental Microbiology Reports 4: 217-226.

Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73: 5261-5267.

Webb CO. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. The American Naturalist 156: 145-155.

Webb CO, Ackerly DD, Kembel SW. (2008). Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics 24: 2098-2100.

Zumsteg A, Luster J, Göransson H, Smittenberg RH, Brunner I, Bernasconi SM, Zeyer J, Frey B. (2012). Bacterial, archaeal and fungal succession in the forefield of a receding glacier. Microbial Ecology 63: 552-564.

Chapter 4

Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Molecular Ecology* **19**, 5555-5565.

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecology* **26**, 32-46.

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

Carpenter EJ, Lin SJ, Capone DG (2000) Bacterial activity in South Pole snow. *Applied and Environmental Microbiology* **66**, 4514-4517.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal* **6**, 1621-1624.

de Garcia V, Brizzio S, Libkind D, Buzzini P, Broock M (2007) Biodiversity of cold-adapted yeasts from glacial meltwater rivers in Patagonia, Argentina. *FEMS Microbiology Ecology* **59**, 331-341.

de Garcia V, Brizzio S, Broock M (2012) Yeasts from glacial ice of Patagonian Andes, Argentina. *FEMS Microbiology Ecology* **82**, 540-550.

Derksen C, Brown R (2012) Spring snow cover extent reductions in the 2008-2012 period exceeding climate model projections. *Geophysical Research Letters* **39**, L19504, doi:10.1029/2012GL053387.

Derksen C, Smith SL, Sharp M, Brown L, Howell S, Copland L, Mueller DR, Gauthier Y, Fletcher CG, Tivy A, Bernier M, Bourgeois J, Brown R, Burn CR, Duguay C, Kushner P, Langlois A, Lewkowicz AG, Royer A, Walker A (2012) Variability and change in the Canadian cryosphere. *Climate Change* **115**, 59-88.

Diamond JM (1975) Assembly of species in communities. In *Ecology and Evolution of Communities* (eds Cody ML, Diamond JM), pp. 342-344. Harvard University Press, Cambridge, MA.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioimformatics* **27**, 2194-2200.

Fountain AG, Campbell JL, Schuur EAG, Stammerjohn SE, Williams MW, Ducklow HW (2012) The disappearing cryosphere: Impacts and ecosystem responses to rapid cryosphere loss. *BioScience* **62**, 405-415.

Freeman KR, Martin AP, Karki D, Lynch RC, Mitter MS, Meyer AF, Longcore JE, Simmons DE, Schmidt SK (2009) Evidence that chytrids dominate fungal communities in high-elevation soils. *Proc. Natl. Acad. Sci. USA* **106**, 18315-18320.

Fujii M, Takano Y, Kojima H, Hoshino T, Tanaka R, Fukui M (2012) Microbial community structure, pigment composition, and nitrogen source of red snow in Antarctica. *Microbial Ecology* **59**, 466-475.

Gihring TM, Zhang GX, Brandt CC, Brooks SC, Campbell JH, Carroll S, Criddle CS, Green SJ, Jardine P, Kostka JE, Lowe K, Mehlhorn TL, Overholt W, Watson DB, Yang ZM, Wu WM, Schadt CW (2011) A limited microbial consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. *Applied and Environmental Microbiology* **77**, 5955-5965.

Gorton HL, William WE, Vogelmann TC (2001) The light environment and cellular optics of the snow alga *Chlamydomonas nivalis* (Bauer) Wille. *Photochemistry and Photobiology* **73**, 611-620.

Gunde-Cimerman N, Sonjak A, Zalar P, Frisvad JC, Diderichsen B, Plemenitaš A (2003) Extremophilic fungi in arctic ice: a relationship between adaptation to low temperature and water activity. *Physics and Chemistry of the Earth* **28**, 1273-1278.

Gutman J, Zarka A, Boussiba S (2009) The host-range of *Paraphysoderma sedebokerensis,* a chytrid that infects *Heamatococcus pluvialis*. *European Journal of Phycology* **44**, 509-514.

Harding T, Jungblut AS, Lovejoy C, Vincent WF (2011) Microbes in high arctic snow and implications for the cold biosphere. *Applied and Environmental Research* **77**, 3234-3243.

Hell K, Edwards A, Zarsky J, Posmirseg SM, Girdwood S, Pachebat JA, Insam H, Sattler B (2013) The dynamic bacterial communities of a melting High Arctic glacier snowpack. *The ISME Journal* **7**, 1814-1826.

Hoffman Y, Aflalo C, Zarka A, Gutman J, James TY, Boussiba S (2008) Isolation and characterization of a novel chytrid species (phylum Blastocladiomycota) parasitic on the green alga *Hearmatococcus*. *Mycological Research* **112**, 70-81.

Hodson A, Anesio AM, Tranter M, Fountain A, Osborn M, Priscu J, Laybourn-Parry J, Sattler B (2008) Glacier Ecosystems. *Ecological Monographs* **78**, 41-67.

Hoham RW (1980) Unicellular chlorophytes – snow algae. In: *Phytoflagellates* (ed Cox ER), Elsevier-North, Holland.

Hoham RW, Laursen AE, Clive SO, Duval B (1993) Snow algae and other microbes in several alpine areas in New England. In: *Proc. 50th Annual Eastern Snow Conference*. pp 165-173.

Hoham RW, Duval B (2001) Microbial ecology of snow and freshwater ice with emphasis on snow algae. In: *Snow Ecology* (eds Johns HG, Pomeroy JW, Walker DA, Hoham RS), pp. 186-203. Cambridge University Press, Cambridge, UK.

Ihrmark K, Bodecker ITM, Cruz-Martinez K, Friberg H, Kubartova A, Schenck J, Strid Y, Stenlid J, Brandstrom-Durling M, Clemmensen KE, Lindahl BD (2012) New primers to amplify the fungal ITS2 region – evaluation by 454-sequencing or artificial and natural communities. *FEMS Microbiology Ecology* **82**, 666-677.

Jumpponen A. and Egerton-Warburton L. (2005) Mycorrhizal fungi in successional environments – a community assembly model incorporating host plant, environmental and biotic filters. In: *The Fungal Community* (eds Dighton J, White JF, Oudemans P), pp. 139-180. CRC Press, New York.

Keddy PA (1992) Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetative Science* **3**,157–164.

Kind T, Meissen JK, Yang D, Nocito F, Vaniya A, Cheng Y, VanderGheynst JS, Fiehn O (2012) Qualitative analysis of algal secretions with multiple mass spectrometric platforms. *Journal of Chromatography A* **1244**, 139-147.

Kol E (1968) *Kryobiologie, Biologie und Limnologie des Schnees and Eises 1. Kryovegetation*. Stuttgart, Germany: E. Schweizerbart'sche Verlagsbuchhandlung.

Larose C, Dommergue A, Vogel TM (2013) The dynamic arctic snow pack: an unexplored environment for microbial diversity and activity. *Biology* **2**, 317-330.

Light JJ, Belcher JH (1968) A snow microflora in the Cairngorm Mountains, Scotland. *British Phycology Bulletin* **3**, 471-473.

Matsumoto N (2009) Snow molds: a group of fungi that prevail under snow. *Microbes and Environments* **24**, 14-20.

Müller T, Bleiß W, Martin CD, Rogaschewski S, Fuhr G (1998) Snow algae from northwest Svalbard: their identification, distribution, pigment and nutrient content. *Polar Biology* **20**, 14-32.

Naff CS, Darcy JL, Schmidt SK (2013) Phylogeny and biogeography of an uncultured clade of snow chytrids. *Environmental Microbiology* **15**, 2672-2680.

Onofri S, Selbmann L, de Hoog GS, Grube M, Barreca D, Suisi S, Zucconi L (2007) Evolution and adaptation of fungi at boundaries of life. *Advances in Space Research* **40**, 1657-1664.

Picard KT, Letcher PM, Powell MJ (2013) Evidence for a facultative mutualist nutritional relationship between the green coccoid alga *Bracteacoccus* sp. (Chlorophyceae) and the zoosporic fungus *Rhizidium phycophilum* (Chytridiomycota). *Fungal Biology* **117**, 319-328.

Porter TM, Golding CB (2012) Factors that affect Large Subunit Ribosomal DNA amplicon sequencing studies of fungal communities: classification method, primer choice, and error. *PLoS ONE* **7**, e35749.

Prestrud P (2007) Why are snow and ice important to us? Chapter 2. In: *UNEP Global Outlook for Ice & Snow* (ed Eamer J), pp. 19-28. United Nations Environment Program.

Pukkila PJ, Skrzynia C (1993) Frequent changes in the number of reiterated ribosomal RNA genes throughout the life cycle of the Basidiomycete *Coprinus cinereus*. *Genetics* **133**, 203-211.

Rasconi S, Jobard M, Sime-Ngando T (2011) Parasitic fungi of phytoplankton: ecological roles and implications for microbial food webs. *Aquatic Microbial Ecology* **62**, 123-137.

Remias D, Karsten U, Lütz C, Leya T (2010) Physiological and morphological processes in the Alpine snow alga *Chloromonas nivalis* (Chlorophyceae) during cyst formation. *Protoplasma* **243**, 73-86.

Remias D, Lütz-Meindl U, Lütz C (2005) Photosynthesis, pigments and ultrastructure of the alpine snow alga *Chlamydomonas nivalis*. *European Journal of Phycology* **40**, 259-268.

Robinson CH (2001) Cold adaptation in Arctic and Antarctic fungi. *New Phytologist* **151**, 341-353.

Sampaio JS, Gadanho M, Bauer R, Weiß M (2003) Taxonomic studies in the Microbotryomycetidae: *Leucosporidium golubevii* sp. nov., *Leucosporidiella* gen. nov. and the new orders Leucosporidiales and Sporidiobolales. *Mycological Progress* **2**, 53-68.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mother: Open-source, platform independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**, 7537-7541.

Schmidt SK, Naff CS and Lynch RC (2012) Fungal communities at the edge: ecological lessons from high alpine fungi. *Fungal Ecology* **5**, 443-452.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Science, USA* **109**, 6241-6246.

Stein JR, Amundsen CC (1967) Studies on snow algae and fungi from the front range of Colorado. *Canadian Journal of Botany* **45**, 2033-2045.

Toome M, Roberson RW, Aime MC (2013) *Metedithblackwellia eburnea* gen. et sp. nov., Kriegeriaceae fam. nov. and Kriegeriales ord. nov.-toward resolving higher-level classification in Microbotryomycetes. *Mycologia* **105**, 486-495.

Walsh JE, Anisimov O, Hagen J, Jakobsson T, Oerlemans J, Prowse TD, Romanovsky V, Savelieva N, Serreze M, Shiklomanov I, Solomon S (2005) Cryosphere and hydrology. In: *Arctic Climate Impact Assessment* pp.183-242. Cambridge University Press, New York.

Warren SG, Hudson SR (2003) Bacterial activity in South Pole snow is questionable. *Applied and Environmental Microbiology* **69**, 6340-6341.

Weiss RL (1983) Fine-structure of the snow algal (*Chlamydomonas nivalis*) and associated bacteria. *Journal of Phycology* **19**, 200-204.

White TJ, Bruns T, Lee S. Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications* (eds Innis MA, Gelfand DH, Sininsky JJ, White TJ). Academic Press, Inc., New York.

Chapter 5

An, S.S., Friedl, T., & Hegewald, E. 1999. Phylogenetic relationships of *Scenedesmus* and *Scenedesmus*-like coccoid green algae as inferred from ITS-2 rDNA sequence comparisons. *Plant Biology.* 1:418-428.

Brown, S.P., Olson, B.J.S.C. & Jumpponen, A. 2014. Snow algae as the control of snow fungal community assembly. *Molecular Ecology* (in review).

Chiasson, W.B., Machesky, N.J. & Vis, M.L. 2003. Phylogeography of a freshwater red alga, *Batrachospermum helminthosum,* in North America. *Phycologia.* 42:654-660.

De Clerck, O., Guiry, M.D., Leliaert, F., Samyn, Y., & Verbruggen, H. 2013. Algae taxonomy: a road to nowhere? *Journal of Phycology.* 49:2415-225.

Excoffier, L. & Lischer, H.E.L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources.* 10:564-567.

Fott, B. 1974. Taxonomie der palmelloiden *Chlorococcales* (Familie *Palmogloeaceae*). *Preslia.* 46:1-31.

Fujii, M., Takano, Y., Kojima, H., Hoshino, T., Tanaka, R. & Fukui, M. 2010. Microbial community structure, pigment composition, and nitrogen source of red snow in Antarctica. *Microbial Ecology.* 59:466-475.

Gradinger, R. & Nürnberg D. 1996. Snow algae communities on Arctic pack ice floes dominated by *Chlamydomonas nivalis* (Bauer) Wille. *Proc. NIPR Symposium Polar Biology.* 9:35-43.

Gorton, H.L., William, W.E. & Vogelmann, T.C. 2001. The light environment and cellular optics of the snow alga *Chlamydomonas nivalis* (Bauer) Wille. *Photochemistry and Photobiology.* 73:611-620.

Hoham, R.W. 1980. Unicellular chlorophytes – snow algae. *In* Cox, E.R. [Ed.] *Phytoflagellates.* Elsevier-North, Holland.

Hoham, R.W. & Duval, B. 2001. Microbial ecology of snow and freshwater ice with emphasis on snow algae. In Jones, H.G, Pomeroy, J.W., Walker, D.A., & Hoham R.W. [Eds.] *Snow Ecology: An Interdisciplinary Examination of Snow-Covered Ecosystems*. Cambridge University Press, Cambridge, UK.

Ihrmark, K., Bödeker, I.T.M., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., Strid, Y., Stenlid, J., Brandstöm-Durling, M., Clemmensen, K.E. & Lindahl, B.D. 2012. New primers to amplify the fungal ITS2 region – evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiology Ecology.* 82:666-677.

Lozupone, C. & Knight, R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied Environmental Microbiology.* 71:8228-8235.

Needleman, S.B., & Wunsch, C.D. 1970. A general method applicable to the search for similarities in the amino acid sequences of two proteins. *Journal of Molecular Biology.* 48:443-453.

Remias, D., Lütz-Meindl, U. & Lütz, C. 2005. Photosynthesis, pigments and ultrastructure of the alpine snow alga *Chlamydomonas nivalis*. *European Journal of Phycology.* 40:259-268.

Remias, D., Karsten, U., Lütz, C. & Leya, T. 2010. Physiological and morphological processes in the Alpine snow alga *Chloromonas nivalis* (Chlorophyceae) during cyst formation. *Protoplasma.* 243:73-86.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M.,, Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G, Van Horn, D.J. & Weber, C.F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology.* 75:7537-7541.

Strunecki, O., Komárek, J. & Elster, J. 2012. Biogeography of *Phormidium autumnale* (Oscillatoriales Cyanobacteria) in western and central Spitsbergen. *Polish Polar Research.* 33:369-382.

Takeuchi, N. 2001. The altitudinal distribution of snow algae on an Alaska glacier (Gulkana Glacier in the Alaska Range). *Hydrological Processes.* 15:3447-3459.

Takeuchi, N. & Kohshima, S. 2004. A snow algal community on Tyndall Glacier in the Southern Patagonia Icefield, Chile. *Arctic, Antarctic, and Alpine Research.* 36:92-99.

Thomas, W.H. & Duval, B. 1995. Sierra Nevada, California, U.S.A., snow algae: snow albedo changes, algal-bacterial interrelationships, and ultraviolet radiation effects. *Arctic and Alpine Research.* 27:389-399.

White, T.J., Bruns, T., Lee, S. & Taylor, J.W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In Innis, M.A., Gelfand, D.H., Sininsky, J.J. & White, T.J. [Eds.]. *PCR Protocols: A Guide to Methods and Applications.* Academic Press, Inc., New York.

Yoshimura, Y., Kohshima, S. & Ohtani, S. 1997. A community of snow algae on a Himalayan glacier: change of algal biomass and community structure with altitude. *Arctic and Alpine Research.* 29:126-137.

Chapter 6

Amend AS, Seifert KA, Samson R, Bruns TD. 2010. Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. *Proceedings on the National Academy of Sciences* **107**: 13748-13753.

Brown SP, Jumpponen A. 2014. Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. *Molecular Ecology* **23**: 481-197.

Gihring TM, Green AJ, Schadt CW. 2012. Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environmental Microbiology* **14:** 285- 290.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, *et al.* 2006. Resonstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443:** 818-822.

Kerekes J, Kaspari M, Stevenson B, Nilsson RH, Hartmann M, Amend A, Bruns TD. 2013. Nutrient enrichment increased species richness of leaf litter fungal assemblages in a tropical forest. *Molecular Ecology* **22:** 2827-2838**.**

Lothamer K, Brown SP, Mattox JD, Jumpponen A. 2013. Comparison of root-associated communities of native and non-native ectomycorrhizal hosts in an urban landscape. *Mycorrhiza* (in press) doi: 10.1007/s00572-013-0539-2.

Margulies M, Egholm M, Altman WE, Attiya A, Bader JS *et al.* 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376-380.

Nilsson RH, Tedersoo L, Abarenkov K, Ryberg M, Kristiansson E, Hartmann M, Schoch CL, Mylander JAA, Bergsten J, Porter TM, Jumpponen A, Vaishampayan P, Ovaskainen O, Hallenberg N, Bengtsson-Palme J, Eriksson KM, Larsson KH, Larsson E, Kóljalg U. 2012. Five simple guidelines for establishing basic authenticity and reliability of newly generated fungal ITS sequences. *MycoKeys* **4**: 37-63.

Peay KG, Kennedy PG, Bruns TD. 2008. Fungal community ecology: a hybrid beast with a molecular master. *BioScience* **58:** 799-810.

Porras-Alfaro A, Liu K, Kuske CR, Xie G. 2014. From genus to phylum: large-subunit and internal transcribed spacer rRNA operon regions show similar classification accuracies influenced by database composition. *Applied and Environmental Microbiology* **80**: 829-840.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mother: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Schloss PD, Gevers D, Westcott SL. 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLOS ONE* **6:** e27310.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal

Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Science, USA* **109:** 6241-6246.

Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73:** 5261-5267.

Weber CF, Vilgalys R. Kuske CR. 2013. Changes in fungal community composition in response to elevated atmospheric CO_2 and nitrogen fertilization varies with soil horizon. *Frontiers in Microbiology* **4**: 78 doi: 10.3389/fmicb.2013.00078.

Chapter 7

Bengtsson-Palme J., Ryberg M., Hartmann M., Branco S., Wang Z., Godhe A., De Wit P., Sánchez-García M., Edersberger I., de Sousa F., Amend A.S., Jumpponen A., Unterseher M., Kristiansson E., Abarenkov K., Bertrand Y.J.K., Sanli K., Eriksson K.M., Vik U., Veldre V., Nilsson R.H., 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution* **4**: 914-919.

Brown S.P., Rigdon-Huss A.R., Jumpponen A., 2014. Analyses of ITS and LSU gene regions provide congruent results on fungal community responses. *Fungal Ecology* **9**: 65-68**.**

Buée M., Reich M., Murat C., Morin E., Nilsson R.H., Uroz S., Martin F., 2009. 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist* **184**: 449-456.

Caporaso J.G., Lauber C.L., Walters W.A., Berg-Lyons D., Huntley J., Fierer N., Owens S.M., Betley J., Fraser L., Bauer M., Gormley N., Gilbert J.A., Smith G., Knight R., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal* **6**: 1621-1624.

Eckert K.A., Kunkel T.A., 1991. DNA polymerase fidelity and the polymerase chain reaction. *Genome Research* **1**: 17-24.

Edgar R.C., Haas B.J., Clemente J.C., Quince C., Knight R., 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194-2200.

Findley K., Oh J., Yang J., Conlan S., Deming C., Meyer J.A., Schoenfeld D., Nomicos E., Park M, NIH Intramural Sequencing Center Comparative Sequencing Program, Kong H.H., Segre J.A., 2013. Topographic diversity of fungal and bacterial communities in human skin. *Nature* **498**: 367-370.

Gobet A., Quince C., Ramette A., 2010. Multivariate cutoff levels analysis (MultoCoLA) of large community data sets. *Nucleic Acid Research* **38**(15): e155.

Haegeman B., Hamelin J., Moriarty J., Neal P., Dushoff J., Weitz J.S., 2013. Robust estimation of microbial diversity in theory and practice. *The ISME Journal* **7**: 1092-1101.

Hibbett, D.S., Ohman, A., Kirk, P.M., 2009. Fungal ecology catches fire. *New Phytologist* **184**: 279-282.

Huse S.M., Welch D.M., Morrison H.G., Sogin M.L., 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology* **12**: 1889-1898.

Jumpponen A., Jones K.L., 2009. Massively parallel 454 sequencing indicated hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytologist* **184**: 438-448.

Kõljalg U., Nilsson R.H., Abarenkov K., Tedersoo L., Taylor A.F.S., Bahram M., Bates S.T., Bruns T.D., Bengtsson-Palme J., Callaghan T.M., Douglas B., Drenkhan T., Eberhardt U., Dueñas M., Grebenc T., Griffith G.W., Hartmann M., Kirk P.M., Kohout P., Larsson E., Lindahl B.D., Lücking R., Nguyen N.H., Niskanen T., Oja J., Peay K.G., Peintner U., Peterson M., Põldmaa K., Saag L., Saar I., Schüßler A., Scott J.A., Senés C., Smith M.E., Suija A., Taylor D.L., Telleria M.T., Weiss M., Larsson K., 2013. Toward a unified paradigm for sequence-based identification of fungi. *Molecular Ecology* **22**: 5271-5277.

Kunin V., Engelbrektson A., Ochman H., Hugenholtz P., 2010. Wrinkles in the rare biosphere: pyrosequencing errors lead to artificial inflation of diversity estimates. *Environmental Microbiology* **12**: 118-123.

Lindahl B.D., Nilsson R.H., Tedersoo L., Abarenkov K., Carlsen T., Kjøller R., Kõljalg U., Pennanen T., Rosendahl S., Stenlid J., Kauserud H., 2013. Fungal community analysis by high-throughput sequencing of amplified markers – a user's guide. *New Phytologist* **199**: 288-299.

Meiser A., Bálint M., Schmitt I., 2014. Meta-analysis of deep-sequenced fungal communities indicates limited taxon sharing between studies and the presence of biogeographic patterns. *New Phytologist* **201**: 623-635.

Nilsson R.H., Veldre V., Hartmann M., Unterseher M., Amend A., Bergsten J., Kristiansson E., Ruberg M., Jumpponen A., Abarenkov K., 2010. An open source software package for automated extraction of ITS1 and ITS2 from fungal ITS sequences for use in high-throughput community assays and molecular ecology. *Fungal Ecology* **3**: 284-287.

Nilsson R.H., Tedersoo L., Lindahl B.D., Kjøller R., Carlsen T., Quince C., Abarenkov K., Pennanen T., Stenlid J., Bruns T., Larsson K., Kõljalg U., Kauserud H., 2011. Towards standardization of the description and publication of next-generation sequencing datasets of fungal communities. *New Phytologist* **191**: 314-318.

Porras-Alfaro A., Liu K., Kuske C.R., Xie G., 2014. From genus to phylum: large-subunit and internal transcribed spacer rRNA operon regions show similar classification accuracies influenced by database composition. *Applied and Environmental Microbiology* **80**: 829-840.

Quince C., Lanzen A., Davenport R.J., Turnbaugh P.J., 2011. Removing noise from Pyrosequenced amplicons. *BMC Bioinformatics* **12**: 38.

Schloss P.D., Westcott S.L., Ryabin T., Hall J.R., Hartmann M., Hollister E.B., Lesniewski R.A., Oakley B.B., Parks E.B., Robinson C .J., Sahl J.W., Stres B., Thallinger G.G., Van Horn D.J., Weber C.F., 2009. Introducing mothur: open-source, platform independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**: 7537-7541.

Shade A., Hogan C.S., Klimowicz A.K., Linske M., McManus P.S., Handelsman J., 2012. Culturing captures members of the soil rare biosphere. *Environmental Microbiology* **14**: 2247-2252.

Sogin M.L., Morrison H.G., Huber J.A., Welch D.M., Huse A.M., Neal P.R., Arrieta J.M., Herndl G.J., 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proceedings of the National Academy of Science* **103**: 12115-12120.

Tedersoo L., Nilsson R.H., Abarenkov K., Jairus T., Sadam A., Saar I., Bahram M., Bechem E., Chuyong G., Kõljalg U., 2010. 454 pyrosequencing and sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist* 188: 291-301.

Unterseher M., Jumpponen A., Öpik M., Tedersoo L., Moora M., Dormann C.F., Schnittler M.,
2011. Species abundance distributions and richness estimations in fungal metagenomics –
lessons learned from community ecology. *Molecular Ecology* 20: 275-285.

Wang Q., Garrity G.M., Tiedje J.M. Cole J.R., 2007. Naïve Bayesian classifier for rapid
assignment of rRNA sequences into new bacterial taxonomy. *Applied and Environmental
Microbiology* 73: 5267-5267.

Zhan A., Hulák M., Sylvester F., Huang X., Adebayo A.A., Abbot, C.L., Adamowicz S.J., Heath
D.D., Cristescu M.E., Maclsaac H.J., 2013. High-sensitivity pyrosequencing of detection
of rare species in aquatic communities. *Methods in Ecology and Evolution* 4: 58-565.

Appendix A - Supplemental Information for Chapter 2

Below are supplemental information discussed in Chapter 2. Additional supplemental information can be accessed at http://onlinelibrary.wiley.com/doi/10.1111/mec.12487.

Table A.2 454 adaptor, primer, and multiplex tag (MIDs) sequences

Sequencing adaptor A	5' -GCCTCCCTCGCGCCATCAG- 3'		
Sequencing adaptor B	5' -GCCTTGCCAGCCCGCTCAG- 3'		
Gene Region Primers			
Fungi	ITS1f	5' -CTTGGTCATTTAGAGGAAGTAA- 3'	
	ITS4	5' - GCTGCGTTCTTCATCGATGC -3'	
Bacteria	9F	5' -GAGTTTGATCMTGGCTCAG- 3'	
	541R	5' -WTTACCGCGGCTGCTGG- 3'	
Molecular Identifier (MID) tags			
Fungi	MIDs	Sample	Treatment
	CTCGCGTGTC	1	0m Bare Soil
	CGTGTCTCTA	2	0m Bare Soil
	CATACTCTAC	4	150m Saxifraga ferruginea
	CGACACTATC	5	150m Bare Soil
	AGACGCACTC	6	150m Bare Soil
	ACGCTCGACA	7	150m Saxifraga ferruginea
	ACGAGTGCGT	8	150m Bare Soil
	TCTACGTAGC	9	150m Saxifraga ferruginea
	CGTCTAGTAC	10	150m Luetkea pectinata
	TCACGTACTA	11	150m Luetkea pectinata
	CGTATGCGAC	12	150m Luetkea pectinata
	CGAGAGATAC	13	300m Luetkea pectinata
	CGAGACGCGC	14	300m Abies lasiocarpa
	TGATACGTCT	15	300m Abies lasiocarpa
	ACAGTCGTGC	16	300m Abies lasiocarpa
	ACATACGCGT	17	300m Luetkea pectinata
	TCGTCGCTCG	18	300m Luetkea pectinata
	CGTCGATCTC	19	300m Saxifraga ferruginea

	TACTCTCGTG	20	300m Saxifraga ferruginea
	TACGAGTATG	21	300m Phyllodoce empetriformis
	CGTAGACTAG	22	300m Saxifraga ferruginea
	ACGACTACAG	23	300m Phyllodoce empetriformis
	TGTACTACTC	24	300m Saxifraga ferruginea
	CACGCTACGT	25	300m Bare Soil
	ATAGAGTACT	26	300m Bare Soil
	AGTACGCTAT	27	300m Bare Soil
	AGCGTCGTCT	28	450m Bare Soil
	CTACGACTGC	29	450m Bare Soil
	ACTGTACAGT	30	450m Bare Soil
	ACTACTATGT	31	450m Abies lasiocarpa
	ACGCGAGTAT	32	450m Abies lasiocarpa
	CTAGTCACTC	33	450m Abies lasiocarpa
	TAGTGTAGAT	34	450m Luetkea pectinata
	TACGCTGTCT	35	450m Luetkea pectinata
	TACAGATCGT	36	450m Luetkea pectinata
	TACACGTGAT	37	450m Phyllodoce empetriformis
	TACACACACT	38	450m Phyllodoce empetriformis
	CGACGTGACT	39	450m Phyllodoce empetriformis
	CAGTAGACGT	40	450m Saxifraga ferruginea
	ACTAGCAGTA	41	450m Saxifraga ferruginea
	AGTAGTGATC	42	450m Saxifraga ferruginea
	CTCTACGCTC	43	600m Bare Soil
	TGTGAGTAGT	44	600m Bare Soil
	TGACGTATGT	45	600m Bare Soil
	TCTATACTAT	46	600m Abies lasiocarpa
	TCTAGCGACT	47	600m Abies lasiocarpa
	TCGCACTAGT	48	600m Abies lasiocarpa
	CGTACAGTCA	49	600m Luetkea pectinata
	CGCGTATACA	50	600m Luetkea pectinata
	CGCAGTACGA	51	600m Luetkea pectinata
	CGATCGTATA	52	600m Phyllodoce empetriformis
	AGTGCTACGA	53	600m Phyllodoce empetriformis
	AGTGTATGTC	54	600m Phyllodoce empetriformis
	AGTGACACAC	55	600m Saxifraga ferruginea
	AGCTCACGTA	56	600m Saxifraga ferruginea
	ATATAGTCGC	57	600m Saxifraga ferruginea

	TATGCTAGTA	58	750m Bare Soil
	TATATATACA	59	750m Bare Soil
	TAGTCGCATA	60	750m Bare Soil
	CTATAGCGTA	62	750m Abies lasiocarpa
	CTACGCTCTA	63	750m Abies lasiocarpa
	CGTACTCAGA	64	750m Abies lasiocarpa
	ACACATACGC	65	750m Luetkea pectinata
	TGTCGTCGCA	66	750m Luetkea pectinata
	TGTCACACGA	67	750m Luetkea pectinata
	TGTAGTGTGA	68	750m Phyllodoce empetriformis
	TGAGTCAGTA	69	750m Phyllodoce empetriformis
	TCTGACGTCA	70	750m Phyllodoce empetriformis
	TCGCTGCGTA	71	750m Saxifraga ferruginea
	TCGATAGTGA	72	750m Saxifraga ferruginea
	AGCGACTAGC	73	750m Saxifraga ferruginea
Bacteria	MIDs	Sample	Treatment
	CTCGCGTGTC	1	0m Bare Soil
	CGTGTCTCTA	2	0m Bare Soil
	ATCAGACACG	4	150m Saxifraga ferruginea
	AGCACTGTAG	5	150m Bare Soil
	AGACGCACTC	6	150m Bare Soil
	ACGCTCGACA	7	150m Saxifraga ferruginea
	TCTACGTAGC	9	150m Saxifraga ferruginea
	CGTCTAGTAC	10	150m Luetkea pectinata
	TCACGTACTA	11	150m Luetkea pectinata
	ATACGACGTA	12	150m Luetkea pectinata
	CGAGAGATAC	13	300m Luetkea pectinata
	CATAGTAGTG	14	300m Abies lasiocarpa
	TCTCTATGCG	16	300m Abies lasiocarpa
	ACATACGCGT	17	300m Luetkea pectinata
	TCGTCGCTCG	18	300m Luetkea pectinata
	TAGAGACGAG	19	300m Saxifraga ferruginea
	TACTCTCGTG	20	300m Saxifraga ferruginea
	TACGAGTATG	21	300m Phyllodoce empetriformis
	CGTAGACTAG	22	300m Saxifraga ferruginea
	ACGACTACAG	23	300m Phyllodoce empetriformis
	TGTACTACTC	24	300m Saxifraga ferruginea

	CACGCTACGT	25	300m Bare Soil
	ATAGAGTACT	26	300m Bare Soil
	AGTACGCTAT	27	300m Bare Soil
	AGCGTCGTCT	28	450m Bare Soil
	AGACTATACT	29	450m Bare Soil
	ACGCGAGTAT	30	450m Bare Soil
	ACTACTATGT	31	450m Abies lasiocarpa
	ACTGTACAGT	32	450m Abies lasiocarpa
	TCGATCACGT	33	450m Abies lasiocarpa
	TAGTGTAGAT	34	450m Luetkea pectinata
	TACGCTGTCT	35	450m Luetkea pectinata
	TACAGATCGT	36	450m Luetkea pectinata
	TACACGTGAT	37	450m Phyllodoce empetriformis
	CGACACTATC	38	450m Phyllodoce empetriformis
	CGACGTGACT	39	450m Phyllodoce empetriformis
	CAGTAGACGT	40	450m Saxifraga ferruginea
	ACTAGCAGTA	41	450m Saxifraga ferruginea
	ACGCGATCGA	42	450m Saxifraga ferruginea
	ACAGTATATA	43	600m Bare Soil
	TGTGAGTAGT	44	600m Bare Soil
	TGACGTATGT	45	600m Bare Soil
	TCTATACTAT	46	600m Abies lasiocarpa
	TCTAGCGACT	47	600m Abies lasiocarpa
	TCGCACTAGT	48	600m Abies lasiocarpa
	CGTACAGTCA	49	600m Luetkea pectinata
	CGCGTATACA	50	600m Luetkea pectinata
	CGAGACGCGC	51	600m Luetkea pectinata
	CGTATGCGAC	52	600m Phyllodoce empetriformis
	AGTGCTACGA	53	600m Phyllodoce empetriformis
	AGTCGAGAGA	54	600m Phyllodoce empetriformis
	AGTATACATA	55	600m Saxifraga ferruginea
	AGCTCACGTA	56	600m Saxifraga ferruginea
	TCACGCGAGA	57	600m Saxifraga ferruginea
	TATGCTAGTA	58	750m Bare Soil
	TATATATACA	59	750m Bare Soil
	CTATAGCGTA	62	750m Abies lasiocarpa
	CTACGCTCTA	63	750m Abies lasiocarpa
	CTACGACTGC	64	750m Abies lasiocarpa

	AGTGACACAC	65	750m Luetkea pectinata
	TGTCGTCGCA	66	750m Luetkea pectinata
	TGTCACACGA	67	750m Luetkea pectinata
	AGTGTATGTC	68	750m Phyllodoce empetriformis
	TGAGTCAGTA	69	750m Phyllodoce empetriformis
	TCTGACGTCA	70	750m Phyllodoce empetriformis
	TCGCTGCGTA	71	750m Saxifraga ferruginea
	TCGATAGTGA	72	750m Saxifraga ferruginea
	AGCGACTAGC	73	750m Saxifraga ferruginea

Table A.3 Table of bacterial and fungal OTU test statistics testing each OTU for a fit to a Poisson distribution

Table A.2 can be found at http://onlinelibrary.wiley.com/doi/10.1111/mec.12487

Appendix B - Supplemental Information for Chapter 3

Table B.4 Primer and multiplex tag (MIDs) sequences for fungal community analysis if the Large Subunit (LSU) as laid out in Chapter 3. Bacterial MIDs used for Chapter 3 are the same as in Table A.1

Primer LR0R	5' – CCGCTGAACTTAAGCATATCA - 3'	
Primer LR3	5' – CCGTGTTTCAAGACGGG - 3'	
MID Sequence	Sample	Distance from Glacier (m)
ACGAGTGCGT	1	0
ACGCTCGACA	2	0
AGACGCACTC	4	150
AGCACTGTAG	5	150
ATCAGACACG	6	150
ATATCGCGAG	7	150
CGTGTCTCTA	8	150
CTCGCGTGTC	9	150
ACGCGAGTAT	10	150
ACTACTATGT	11	150
ACTGTACAGT	12	150
AGACTATACT	13	300
AGCGTCGTCT	14	300
AGTACGCTAT	15	300
ATAGAGTACT	16	300
CACGCTACGT	17	300
AGCTCACGTA	18	300
AGTATACATA	19	300
AGTCGAGAGA	20	300
AGTGCTACGA	21	300
CGATCGTATA	22	300
CGCAGTACGA	23	300
CGCGTATACA	24	300
CGTACAGTCA	25	300
TCGATAGTGA	26	300
TCGCTGCGTA	27	300
TCTGACGTCA	28	450
TGAGTCAGTA	29	450
TGTAGTGTGA	30	450
TGTCACACGA	31	450
TGTCGTCGCA	32	450
ACACATACGC	33	450

CATACTCTAC	34	450
CGACACTATC	35	450
CGAGACGCGC	36	450
CGTATGCGAC	37	450
AGCTATCGCG	38	450
AGTCTGACTG	39	450
AGTGAGCTCG	40	450
ATAGCTCTCG	41	450
ATCACGTGCG	42	450
ATCGTAGCAG	43	600
ATCGTCTGTG	44	600
ATGTACGATG	45	600
ATGTGTCTAG	46	600
CACACGATAG	47	600
CACTCGCACG	48	600
CAGACGTCTG	49	600
CAGTACTGCG	50	600
CGACAGCGAG	51	600
CGATCTGTCG	52	600
CGCGTGCTAG	53	600
CGCTCGAGTG	54	600
CGTGATGACG	55	600
CTATGTACAG	56	600
CTCGATATAG	57	600
CTCGCACGCG	58	750
CTGCGTCACG	59	750
CTGTGCGTCG	60	750
TAGCATACTG	62	750
TATACATGTG	63	750
TATCACTCAG	64	750
TATCTGATAG	65	750
TCGTGACATG	66	750
TCTGATCGAG	70	750
TGACATCTCG	71	750
TGAGCTAGAG	72	750
TGATAGAGCG	73	750

Appendix C - Supplemental Information for Chapter 4

Table C.1 Primer and multiplex tag (MIDs) sequences

Gene Primers	
ITS7f	5'-GTGARTCATCGAATCTTTG-3'
ITS4	5'-TCCTCCGCTTATTGATATGC-3'
Molecular Identifier (MID) tags	
Tags	Sample
TCCCTTGTCTCC	Lyman-5-2011-Algae
ACGAGACTGATT	Lyman-1-2011-Algae
TACCGCTTCTTC	Niwot-3-2011-Algae
ATCACCAGGTGT	Lyman-3-2011-Algae
TGGTCAACGATA	Lyman-4-2011-Algae
ATCGCACAGTAA	Niwot-1-2011-Algae
GTCGTGTAGCCT	Lyman-6-2011-Algae
GATTATCGACGA	Niwot-2-2011-Algae
ATCCTTTGGTTC	Lyman-2-2011-Algae
GCCTAGCCCAAT	Lyman-6-2011-Non_Algae
ACCGGTATGTAC	Lyman-2-2011-Non_Algae
GATGTATGTGGT	Lyman-4-2011-Non_Algae
TGCATACACTGG	Lyman-5-2011-Non_Algae
AGTCGAACGAGG	Niwot-1-2011-Non_Algae
ACCAGTGACTCA	Niwot-2-2011-Non_Algae
GAATACCAAGTC	Lyman-3-2011-Non_Algae
GTAGATCGTGTA	Niwot-3-2011-Non_Algae
TAACGTGTGTGC	Lyman-1-2011-Non_Algae
ACTCCTTGTGTT	IndianPeaks-2-2012-Algae
CCAATACGCCTG	IndianPeaks-1-2012-Algae
ACTTGGTGTAAG	Lyman-3-2012-Algae
TCACCTCCTTGT	Lyman-4-2012-Algae
CAAACAACAGCT	Lyman-6-2012-Algae
GCAACACCATCC	Lyman-5-2012-Algae
GCACACCTGATA	Lyman-2-2012-Algae
CGAGCAATCCTA	Lyman-1-2012-Algae
AGTCGTGCACAT	IndianPeaks-1-2012-Non_Algae
GCGACAATTACA	IndianPeaks-2-2012-Non_Algae
CGAGGGAAAGTC	Lyman-2-2012-Non_Algae
TCATGCTCCATT	Lyman-6-2012-Non_Algae
AGATTGACCAAC	Lyman-3-2012-Non_Algae
AGTTACGAGCTA	Lyman-4-2012-Non_Algae
GCATATGCACTG	Lyman-1-2012-Non_Algae
CAACTCCCGTGA	Lyman-5-2012-Non_Algae

Table C.2 Taxonomic descriptions of the 200 most abundant fungal OTUs including sequence count, best BLASTn match, Max Score, Total Score, Query Coverage, E-Value, Max Identity and Accession numbers for the closet match.

OTU Number	Seq Count	Species	Phylum	Max Score	Total Score	Query Coverage	E value	Max Identity	Accession
2	49594	Rhodotorula sp.	Basidiomycota	315	315	100%	1.00E-82	89%	JF805370.1
3	45579	Rhodotorula sp.	Basidiomycota	407	407	100%	3.00E-110	96%	AB474394.1
4	41353	Rhodotorula psychrophenolica	Basidiomycota	446	446	100%	3.00E-122	99%	EF151247.1
5	39847	Rhodotorula sp.	Basidiomycota	297	297	100%	4.00E-77	87%	JF805370.1
6	32545	Cryptococcus saitoi	Basidiomycota	125	125	96%	1.00E-25	74%	JX976323.1
7	29737	Rhodotorula glacialis	Basidiomycota	302	302	100%	9.00E-79	88%	JQ857032.1
8	21673	Asterophora sp	Basidiomycota	374	374	100%	2.00E-100	93%	HM036644.1
9	20636	Rhizophydiales sp	Chytridiomycota	87.8	87.8	25%	3.00E-14	91%	EF634250.1
10	16371	Leucosporidiella fragaria	Basidiomycota	356	356	100%	4.00E-95	92%	JN400812.1
11	13141	Sydowia polyspora	Ascomycota	452	452	100%	8.00E-124	100%	JQ780656.1
12	12874	Leucosporidium sp.	Basidiomycota	273	273	100%	4.00E-70	85%	JX014242.1
15	7325	Rhodotorula glacialis	Basidiomycota	340	340	100%	3.00E-90	90%	JN400812.1
16	6726	Asterophora sp	Basidiomycota	266	266	100%	6.00E-68	85%	HM036644.1
17	6460	Rhodotorula sp.	Basidiomycota	385	385	100%	9.00E-104	94%	AB474394.1
18	6264	Rhodotorula psychrophenolica	Basidiomycota	210	210	95%	4.00E-51	82%	EF151247.1
19	5958	Sarcinomyces crustaceus	Ascomycota	452	452	100%	8.00E-124	100%	JN040515.1
20	5880	Stemphylium sp.	Ascomycota	446	446	100%	3.00E-122	99%	HQ622100.1
21	5839	Polyporoletus sublividus	Basidiomycota	64.4	64.4	18%	4.00E-07	91%	DQ389663.1
22	5764	Rhodotorula glacialis	Basidiomycota	298	298	100%	1.00E-77	88%	JQ857032.1
23	5057	Coniozyma leucospermi	Ascomycota	443	443	100%	4.00E-121	99%	EU552113.1
24	4644	Polychytrium aggregatum	Chytridiomycota	64.4	64.4	15%	4.00E-07	97%	NG_027613.1
25	4054	Rhizophydium sp	Chytridiomycota	89.7	89.7	21%	1.00E-14	96%	DQ485621.1
26	3238	Rhodotorula sp.	Basidiomycota	385	385	100%	9.00E-104	94%	AB474394.1
27	2914	Umbelopsis ramanniana	(Zygomycota)	66.2	66.2	28%	1.00E-07	83%	EU715662.1
28	2877	Rhizophydium chlorogonii	Chytridiomycota	64.4	64.4	15%	4.00E-07	97%	JN943815.1
29	2591	Aureobasidium pullulans	Ascomycota	421	421	100%	1.00E-114	98%	FJ744598.1
30	2523	Celosporium larixicola	Ascomycota	426	426	98%	3.00E-116	98%	FJ997287.1
31	2436	Glomus diaphanum	Glomeromycota	66.2	66.2	20%	1.00E-07	90%	AJ972462.1
32	2339	Rhodotorula sp.	Basidiomycota	347	347	100%	2.00E-92	91%	AB474394.1
33	2115	Rhizophydium sp	Chytridiomycota	80.5	80.6	60%	5.00E-12	74%	DQ485665.1
35	2040	cf.Sistotrema sp.	Basidiomycota	260	260	100%	3.00E-66	84%	FR838002.1
36	1908	Phaeococcomyces nigricans	Ascomycota	389	389	87%	8.00E-105	99%	AY843154.1
37	1739	Chytridiomycota sp.	Chytridiomycota	68	68	18%	3.00E-08	93%	EU873018.1

38	1646	Leucosporidium sp.	Basidiomycota	316	316	100%	4.00E-83	89%	JQ272411.1
39	1627	Leucosporidiella sp.	Basidiomycota	215	215	89%	4.00E-51	83%	JN197600.1
40	1621	Chytridiomycota sp.	Chytridiomycota	68	68	18%	3.00E-08	93%	EU873018.1
41	1532	Rhodotorula sp.	Basidiomycota	307	307	100%	2.00E-80	88%	AB474394.1
42	1464	Rhodotorula glacialis	Basidiomycota	288	288	100%	2.00E-74	87%	JQ857032.1
43	1457	Rhodotorula sp.	Basidiomycota	311	311	100%	2.00E-81	88%	JF805370.1
44	1394	Rhizophydium chlorogonii	Chytridiomycota	66.2	66.2	16%	1.00E-07	95%	JN943815.1
45	1338	Rhodotorula sp.	Basidiomycota	302	302	100%	9.00E-79	87%	AY474394.1
46	1212	Celosporium larixicola	Ascomycota	407	407	98%	3.00E-110	97%	FJ997287.1
47	1142	Rhodotorula glacialis	Basidiomycota	425	425	100%	1.00E-115	98%	JQ857037.1
48	1105	Tylopilus formosus	Basidiomycota	69.8	69.8	20%	9.00E-09	90%	HM060320.1
49	1066	Aureobasidium pullulans	Ascomycota	444	444	100%	1.00E-121	99%	JN400825.1
50	1036	Leucosporidiella sp.	Basidiomycota	232	232	89%	1.00E-57	85%	JN197600.1
52	996	Rhodotorula sp.	Basidiomycota	324	324	100%	3.00E-85	90%	AB474394.1
53	976	Cryptococcus podzollcus	Basidiomycota	452	452	100%	8.00E-124	100%	HF558652.1
54	967	Trichosporon gamsii	Basidiomycota	452	452	100%	8.00E-124	100%	NR_073247.1
55	934	Leucosporidiella muscorum	Basidiomycota	302	302	100%	9.00E-79	87%	FR717869.1
57	902	Rhodotorula sp.	Basidiomycota	437	437	100%	2.00E-119	99%	AB474394.1
58	898	Cladophialophora minutissima	Ascomycota	336	336	97%	4.00E-89	91%	EF016382.1
59	862	Ilyonectria macrodidyma	Ascomycota	452	452	100%	8.00E-124	100%	KC311505.1
60	858	Mycocentrospora cantuariensis	Ascomycota	398	398	100%	1.00E-107	95%	EU326864.1
61	840	Rhodotorula glacialis	Basidiomycota	266	266	98%	6.00E-68	86%	JQ857032.1
62	834	Capronia sp.	Ascomycota	349	349	99%	7.00E-93	92%	AF284129.1
63	802	Rhodotorula glacialis	Basidiomycota	293	293	100%	4.00E-76	87%	JQ857032.1
65	769	Rhodotorula glacialis	Basidiomycota	260	260	100%	3.00E-66	85%	JQ857032.1
67	709	Polychytrium aggregatum	Chytridiomycota	66.2	66.2	15%	1.00E-07	97%	NG_027613.1
68	692	Chytridiomycota sp.	Chytridiomycota	68	68	18%	3.00E-08	93%	EU873018.1
70	682	Mortierella kuhlmanii	(Zygomycota)	389	389	97%	8.00E-105	96%	HQ630294.1
71	661	Rhodotorula glacialis	Basidiomycota	232	232	100%	1.00E-57	81%	JQ857032.1
72	643	Polychytrium aggregatum	Chytridiomycota	66.2	66.2	15%	1.00E-07	97%	NG_027613.1
73	612	Mortierella verticillata	(Zygomycota)	452	452	100%	8.00E-124	100%	JN943798.1
74	608	Trichosporon coprophilum	Basidiomycota	452	452	100%	8.00E-124	100%	AB180199.1
75	600	Chytridiomycota sp.	Chytridiomycota	64.4	64.4	18%	4.00E-07	91%	EU873018.1
76	593	Rhodotorula glacialis	Basidiomycota	242	242	90%	7.00E-61	86%	JQ857032.1
77	593	Leucosporidium fellii	Basidiomycota	235	235	99%	1.00E-58	82%	NR_073276.1
78	577	Umbelopsis nana	(Zygomycota)	452	452	100%	8.00E-124	100%	KC489506.1
79	576	Capronia sp.	Ascomycota	446	446	100%	3.00E-122	99%	JQ354915.1
80	567	Dothidea berberidis	Ascomycota	437	437	100%	2.00E-119	99%	EU167601.1

81	566	Rhodotorula sp.	Basidiomycota	289	289	100%	5.00E-75	87%	JF805370.1
82	556	Chytridiomycota sp.	Chytridiomycota	64.4	64.4	18%	4.00E-07	91%	EU873018.1
83	551	Polychytrium aggregatum	Chytridiomycota	66.2	66.2	15%	1.00E-07	97%	NG_027613.1
86	509	Rhodotorula sp.	Basidiomycota	398	398	100%	1.00E-107	95%	AB474394.1
87	497	Aureobasidium sp.	Ascomycota	432	432	100%	7.00E-118	98%	HQ829153.1
89	492	Celosporium larixicola	Ascomycota	387	387	98%	3.00E-104	95%	FJ997287.1
91	486	Typhula variabilis	Basidiomycota	430	430	100%	2.00E-117	98%	AB267395.1
93	470	Rhodotorula glacialis	Basidiomycota	223	223	90%	7.00E-55	84%	JQ857032.1
94	453	Ulocladium chartarum	Ascomycota	452	452	100%	8.00E-124	100%	KC180717.1
95	450	Mortierella sp.	(Zygomycota)	446	446	100%	3.00E-122	99%	HQ608097.1
96	438	Malassezia sp.	Basidiomycota	439	439	100%	5.00E-120	99%	DQ347480.1
97	437	Trichoderma spirale	Ascomycota	452	452	100%	8.00E-124	100%	JZ076964.1
98	417	Rhodotorula sp.	Basidiomycota	293	293	100%	4.00E-76	86%	AB474394.1
99	408	Lactarius longisporus	Basidiomycota	208	208	100%	1.00E-50	80%	DQ421971.1
100	390	Rhodotorula glacialis	Basidiomycota	302	302	100%	9.00E-79	88%	JQ857032.1
101	380	Fibulobasidium murrhardtense	Basidiomycota	194	194	100%	3.00E-46	78%	GU327540.1
102	373	Celosporium larixicola	Ascomycota	360	360	98%	4.00E-96	92%	FJ997287.1
103	370	Ramariopsis laeticolor	Basidiomycota	293	293	100%	4.00E-76	88%	EU1186181
104	367	Celosporium larixicola	Ascomycota	340	340	98%	3.00E-90	91%	FJ997287.1
105	362	Rhodotorula sp.	Basidiomycota	360	360	100%	4.00E-96	93%	AB474394.1
106	358	Clavulinopsis miyabeana	Basidiomycota	168	168	100%	1.00E-38	77%	AB509666.1
107	350	Alternaria tenuissima	Ascomycota	452	452	100%	8.00E-124	100%	KC460834.1
108	344	Ganoderma sp.	Basidiomycota	452	452	100%	8.00E-124	100%	HM192933.1
109	319	Rhodotorula glacialis	Basidiomycota	443	443	100%	4.00E-121	99%	JQ857032.1
110	314	Rhodotorula glacialis	Basidiomycota	423	423	100%	4.00E-115	98%	JQ857032.1
111	314	Lyophyllum sp.	Basidiomycota	75.2	75.2	22%	2.00E-10	89%	DQ182502.1
112	314	Aureobasidium sp.	Ascomycota	235	235	100%	1.00E-58	82%	JX462675.1
113	308	Hyaloraphidium curvatum	Chytridiomycota	66.2	66.2	39%	1.00E-07	78%	AY997055.1
114	308	Polychytrium aggregatum	Chytridiomycota	66.2	66.2	15%	1.00E-07	97%	NG_027613.1
115	305	Fusarium acuminatum	Ascomycota	452	452	100%	8.00E-124	100%	KF181242.1
116	299	Tricholoma portentosum	Basidiomycota	80.6	80.6	20%	5.00E-12	94%	AB699672.1
117	296	Phaeosclera dermatioides	Ascomycota	233	233	87%	4.00E-58	86%	AJ244254.1
118	294	Chytridiomycota sp.	Chytridiomycota	68	68	18%	3.00E-08	93%	EU873018.1
119	287	Rhodotorula sp.	Basidiomycota	232	232	88%	1.00E-57	84%	KC455921.1
120	284	Dothideomycetes sp.	Ascomycota	430	430	100%	3.00E-117	98%	GQ153222.1
121	282	Rhodotorula sp.	Basidiomycota	372	372	100%	6.00E-100	94%	KC455921.1
122	274	Polychytrium aggregatum	Chytridiomycota	64.4	64.4	15%	4.00E-07	97%	NG_027613.1
123	268	Amandinea punctata	Ascomycota	408	408	100%	8.00E-111	96%	HQ650627.1

124	267	Endosporium aviarium	Ascomycota	232	232	90%	1.00E-57	85%	EU304350.1
125	267	Curvularia spicifera	Ascomycota	452	452	100%	8.00E-124	100%	KC897667.1
126	260	Lophium mytiilnum	Ascomycota	242	242	100%	7.00E-61	82%	EF596819.1
127	258	Verrucaria sp.	Ascomycota	206	206	100%	5.00E-50	80%	FJ664851.1
128	257	Alphamyces chaetifer	Chytridiomycota	82.4	82.4	18%	2.00E-12	100%	EF585633.1
129	251	Leucosporidium fellii	Basidiomycota	185	185	100%	2.00E-43	77%	NR_073276.1
130	244	Rhodotorula glacialis	Basidiomycota	248	248	89%	2.00E-62	86%	KC455919.1
131	240	Lyophyllum sp.	Basidiomycota	68	68	24%	3.00E-08	84%	JX966308.1
132	234	Clavulinopsis miyabeana	Basidiomycota	125	125	72%	1.00E-25	76%	AB509796.1
133	233	Zalerion arboricola	Ascomycota	434	434	100%	2.00E-118	98%	FR837917.1
134	230	Helicosporium gracile	Ascomycota	141	141	100%	2.00E-30	74%	AY916485.1
136	222	Phaeococcomyces eucalypti	Ascomycota	246	246	98%	6.00E-62	84%	KC005769.1
137	222	Presuuia sp.	Ascomycota	448	448	100%	1.00E-122	99%	KC333160.1
138	221	Dothichiza pityophili	Ascomycota	389	389	87%	8.00E-105	99%	AJ244242.1
139	221	Venturiaceae sp.	Ascomycota	134	134	100%	3.00E-28	73%	JQ272465.1
140	219	Rhodotorula sp.	Basidiomycota	379	349	100%	4.00E-102	94%	AB474394.1
141	217	Rhizophydium sp	Chytridiomycota	84.2	84.2	18%	4.00E-13	100%	EF585662.1
142	216	Paecilomyces sp.	Ascomycota	347	347	96%	2.00E-92	93%	GU108582.1
143	213	Botrysphaeriacese sp.	Ascomycota	235	235	10%	1.00E-58	82%	HM176528.1
144	210	Clavulinopsis sp.	Basidiomycota	134	134	100%	3.00E-28	71%	JN569120.1
145	207	Phialocephala fluminis	Ascomycota	383	383	100%	3.00E-103	93%	NR_103569.1
146	206	Rhizophydium sp	Chytridiomycota	86	86	22%	1.00E-13	95%	EF585662.1
147	205	Rhizophydium chlorogonii	Chytridiomycota	64.4	64.4	15%	4.00E-07	97%	JN943815.1
148	205	Tricholoma dulciolens	Basidiomycota	141	141	100%	2.00E-30	74%	AB738883.1
150	203	Rhodotorula sp.	Basidiomycota	230	230	96%	5.00E-57	82%	KC455921.1
151	198	Trechispora confinis	Basidiomycota	143	143	92%	5.00E-31	75%	AF347081.1
152	191	Cryptococcus magnus	Basidiomycota	446	446	100%	3.00E-122	99%	KC455883.1
153	191	Polychytrium aggregatum	Chytridiomycota	66.2	66.2	15%	1.00E-07	97%	NG_027613.1
154	188	Sarcinomyces crustaceus	Ascomycota	372	372	100%	6.00E-100	94%	JN040515.1
155	188	Dothideomycetes sp.	Ascomycota	425	425	100%	1.00E-115	98%	GQ153222.1
156	185	Trichoderma hamatum	Ascomycota	452	452	100%	8.00E-124	100%	KC884769.1
157	184	Mrakiella sp.	Basidiomycota	452	452	100%	8.00E-124	100%	JN400824.1
158	183	Entoloma minutum	Basidiomycota	199	199	100%	8.00E-48	80%	JX454829.1
159	179	Rhodotorula sp.	Basidiomycota	381	381	100%	1.00E-102	94%	KC455921.1
160	175	Tomentella sp.	Basidiomycota	68	68	20%	3.00E-08	90%	JN129414.1
161	172	Geminibasidium hirsutum	Basidiomycota	334	334	88%	2.00E-88	94%	JX242880.1
163	170	Lyophyllum sp.	Basidiomycota	75.2	75.2	22%	2.00E-10	89%	DQ182502.1
164	169	Mortierella sp.	(Zygomycota)	443	443	100%	4.00E-121	99%	GQ302682.1

165	168	Donadinia nigrella	Ascomycota	325	325	92%	8.00E-86	91%	JX669836.1
166	168	Thysanophora penicillioides	Ascomycota	203	203	100%	6.00E-49	79%	AB213266.1
167	167	Epicoccum nigrum	Ascomycota	452	452	100%	8.00E-124	100%	KC164754.1
168	166	Helicoma isiola	Ascomycota	271	271	100%	2.00E-69	85%	DQ341099.1
169	160	Trechispora subsphaerospora	Basidiomycota	253	253	100%	4.00E-64	84%	AF347080.1
170	160	Tuber mexiusanum	Ascomycota	446	446	100%	3.00E-122	99%	JX030294.1
171	159	Polychytrium aggregatum	Chytridiomycota	66.2	66.2	15%	1.00E-07	97%	NG_027613.1
172	159	Phialocephala fluminis	Ascomycota	374	374	100%	2.00E-100	93%	NR_103569.1
173	158	Trichosporon dermatis	Basidiomycota	68	68	24%	3.00E-08	85%	KC254108.1
174	158	Rhodotorula sp.	Basidiomycota	334	334	100%	2.00E-88	91%	KC333170.1
175	155	Postia alni	Basidiomycota	82.4	82.4	70%	2.00E-12	74%	KC595931.1
176	154	Sarcinomyces crustaceus	Ascomycota	307	307	100%	2.00E-80	89%	JN040515.1
177	152	Lyophyllum sp.	Basidiomycota	71.6	71.6	26%	3.00E-09	85%	DQ182502.1
178	150	Lyophyllum sp.	Basidiomycota	75.2	75.2	22%	2.00E-10	89%	DQ182502.1
179	150	Ascomycota sp.	Ascomycota	452	452	100%	8.00E-124	100%	JQ775574.1
180	149	Glomus sp.	Glomeromycota	66.2	66.2	18%	1.00E-07	91%	AJ504633.1
181	147	Metarhizium anisopliae	Ascomycota	446	446	100%	3.00E-122	99%	JN256671.1
182	143	Cryptococcus sp.	Basidiomycota	389	389	100%	8.00E-105	95%	FJ873574.1
183	143	Mrakia gelida	Basidiomycota	120	120	77%	6.00E-24	75%	JQ857036.1
184	143	Paecilomyces carneus	Ascomycota	304	304	100%	3.00E-79	89%	KC180711.1
185	143	Lyophyllum sp.	Basidiomycota	75.2	75.2	22%	2.00E-10	89%	DQ182502.1
186	139	Rhizophydiales sp	Chytridiomycota	82.4	82.4	23%	2.00E-12	92%	FR670788.1
187	139	Hygrocybe irrigata	Basidiomycota	116	116	54%	7.00E-23	81%	FM208881.1
188	138	Sistotrema sp.	Basidiomycota	233	233	100%	4.00E-58	82%	FR838002.1
189	138	Rhodotorula sp.	Basidiomycota	300	300	98%	3.00E-78	89%	KC455921.1
190	137	Rhodotorula sp.	Basidiomycota	309	309	100%	6.00E-81	88%	AB474394.1
191	137	Rhodotorula sp.	Basidiomycota	273	273	100%	4.00E-70	85%	AB474394.1
192	136	Penicillium sp.	Ascomycota	452	452	100%	8.00E-124	100%	KF305753.1
193	131	Exophiala sp.	Ascomycota	452	452	100%	8.00E-124	100%	JX243973.1
195	129	Rhodotorula sp.	Basidiomycota	309	309	100%	6.00E-81	88%	KC333170.1
196	129	Lophiostoma sp.	Ascomycota	396	396	100%	5.00E-107	96%	HQ914838.1
197	128	Rhodotorula glacialis	Basidiomycota	269	269	100%	5.00E-69	96%	KC455919.1
198	127	Saccharomyces cerevisiae	Ascomycota	466	466	100%	3.00E-122	99%	KC183729.1
199	125	Leucosporidium scottii	Basidiomycota	347	347	100%	2.00E-92	91%	JX014242.1
200	124	Chytridiomycota sp.	Chytridiomycota	69.8	69.8	22%	1.00E-08	88%	EU873018.1
201	124	Clavulinopsis miyabeana	Basidiomycota	163	163	77%	6.00E-37	80%	AB509666.1
202	123	Clavulinopsis sp.	Basidiomycota	109	109	82%	1.00E-20	71%	JN569120.1
204	122	Capronia pulcherrima	Ascomycota	298	298	100%	1.00E-77	86%	AF050256.1

205	120	Phaeosclera sp.	Ascomycota	96.9	96.9	40%	7.00E-17	82%	AY843195.1
206	119	Rhodotorula sp.	Basidiomycota	311	311	100%	2.00E-81	88%	AB474394.1
207	117	Clavulinopsis miyabeana	Basidiomycota	199	199	100%	8.00E-48	78%	AB509666.1
208	114	Capronia sp.	Ascomycota	360	360	99%	4.00E-96	93%	AF284129.1
209	112	Alternaria infectoria	Ascomycota	452	452	100%	8.00E-124	100%	HG324079.1
210	112	Mycena oregonenis	Basidiomycota	255	255	100%	1.00E-64	82%	JF908409.1
211	111	Capronia villosa	Ascomycota	210	210	100%	4.00E-51	81%	AF050261.1
213	110	Saccharata sp.	Ascomycota	239	239	100%	9.00E-60	82%	JN225922.1
214	109	Sarcinomyces crustaceus	Ascomycota	289	289	100%	6.00E-75	86%	JN040515.1
215	108	Trichoderma strigosellum	Ascomycota	446	446	100%	3.00E-122	99%	EU280139.1
216	108	Phaeomoniella prunicola	Ascomycota	342	342	100%	1.00E-90	91%	GQ154588.1
217	107	Lophium mytiilnum	Ascomycota	452	452	100%	8.00E-124	100%	EF596819.1
218	105	Holtermanniella watticus	Basidiomycota	452	452	100%	8.00E-124	100%	JQ857031.1
219	105	Rhodotorula sp.	Basidiomycota	385	385	100%	1.00E-103	94%	KC455921.1
220	103	Rhodotorula sp.	Basidiomycota	262	262	98%	8.00E-67	85%	KC455921.1

Figure C.1 Ranked OTU abundance distribution plot of all OTUs. The dashed line represents our cutoff off 200 OTUs for analyses and represents greater than 97% of all fungal sequences. Insert represents the first 30 OTUs that were extremely abundant.
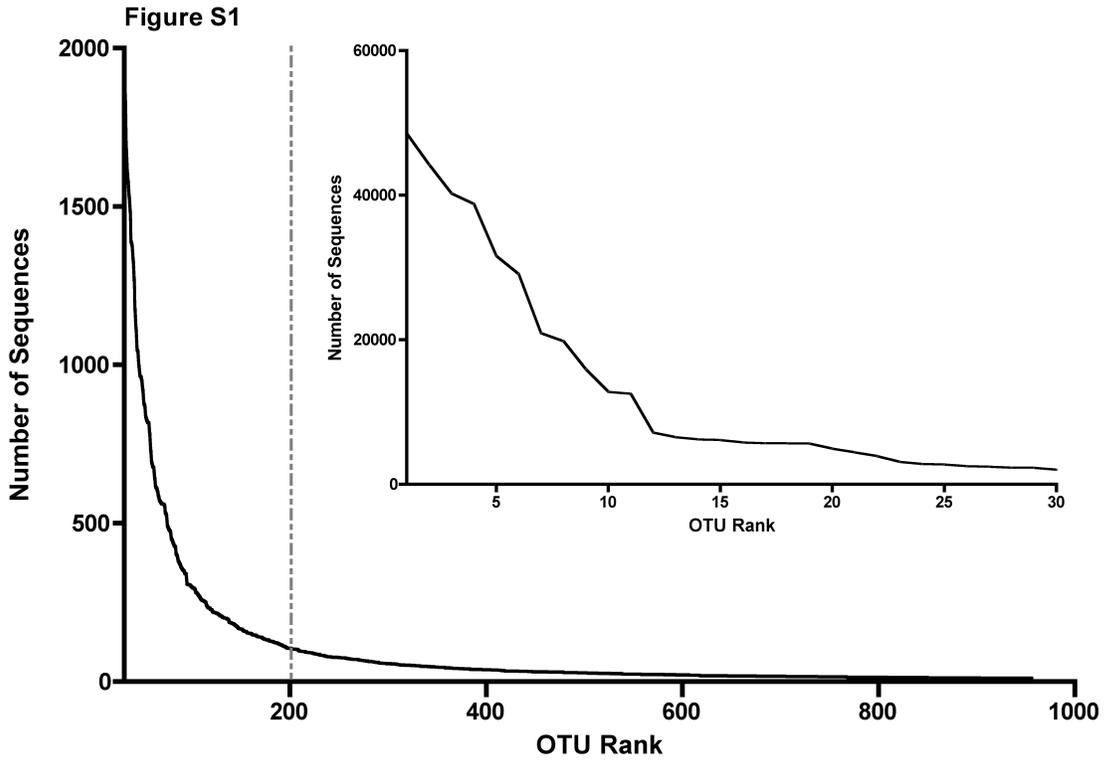
Figure C.2 Phylogenetic analysis (Maximum Likelihood) of putative Chytrid OTUs with vouchered representative ITS2 sequences within Phyla Chytridiomycota, Blastocladiomycota and Monoblepharidomycota indicate that observed novel OTUs are nested within Phylum Chytridiomycota with 99% bootstrap support.
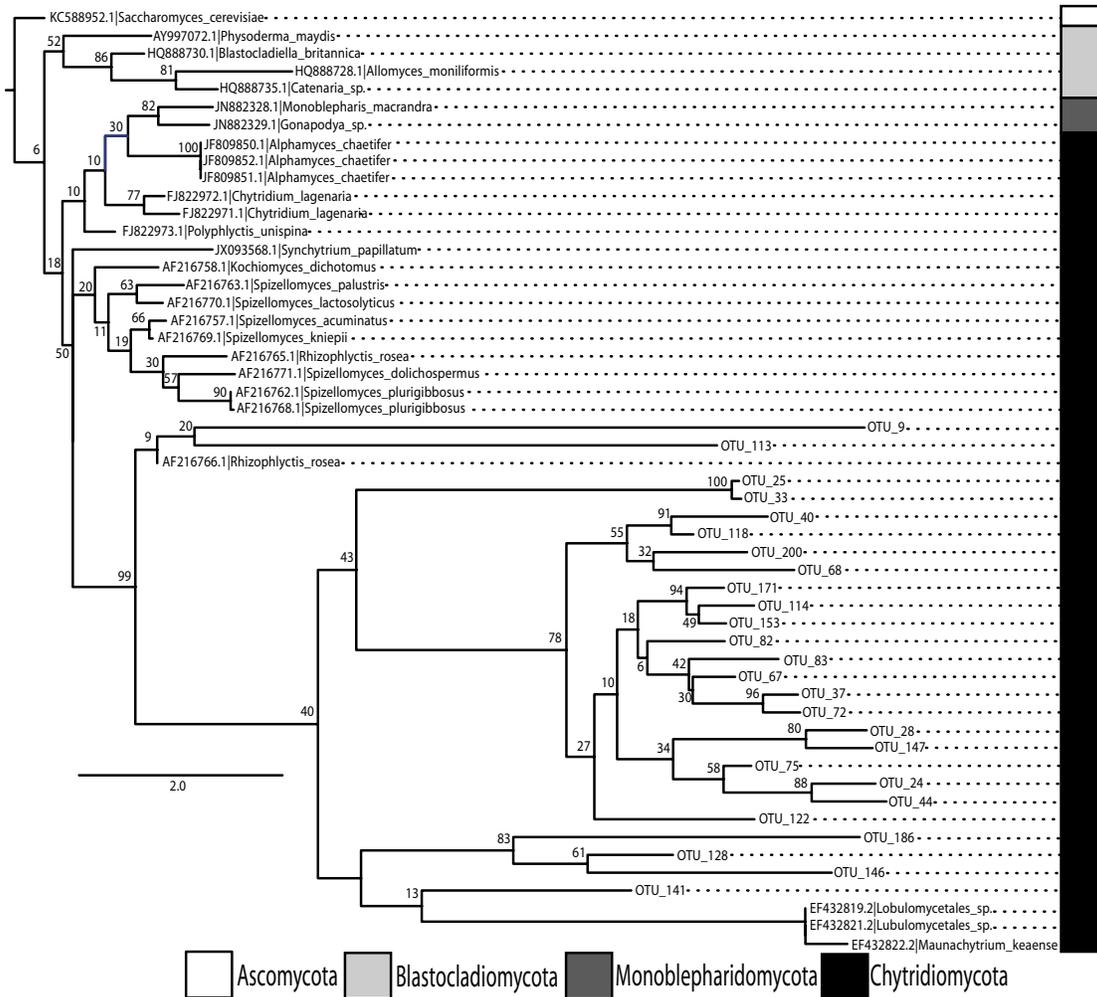
Figure C.3 Rarefaction analysis of observed OTUs for algal colonized and non-algal colonized snow fungi indicate that at the 1500 sequence subsampling point (dashed line), the majority of community members have been observed as this subsample value is well past the inflection point of the curves.
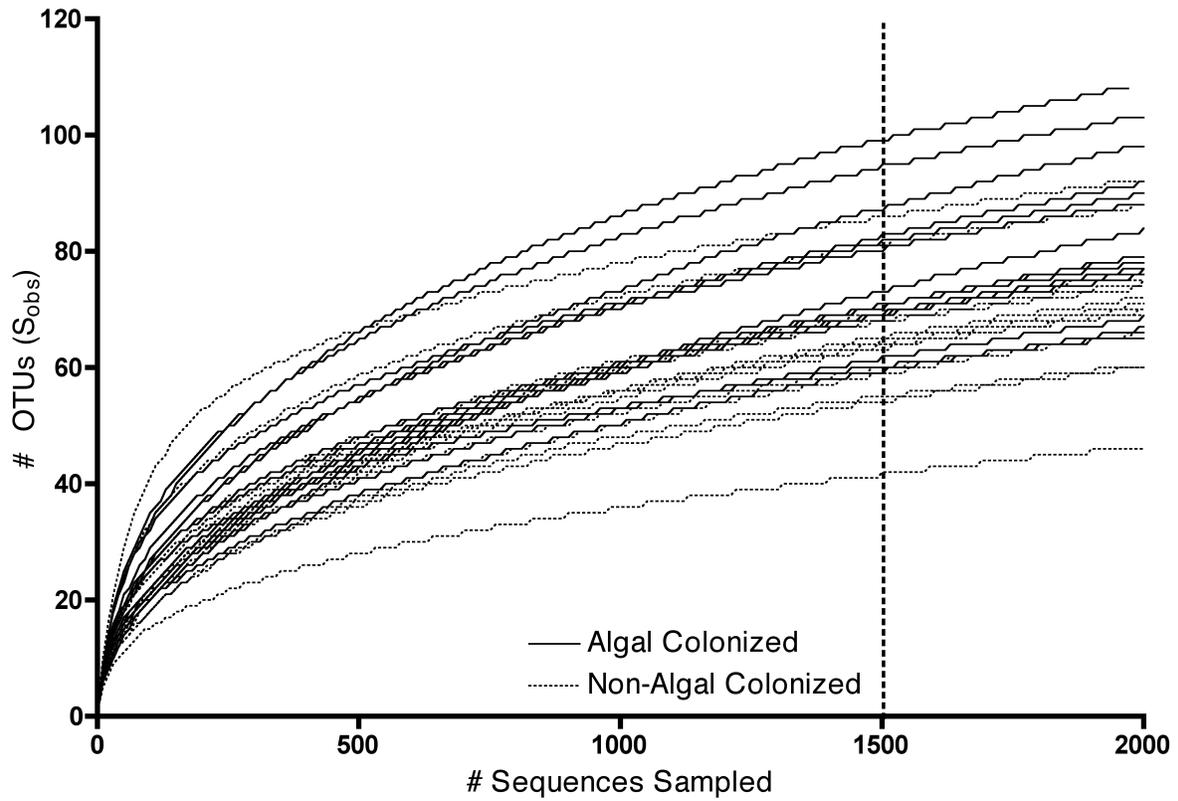
Figure C.4 Distances of paired algal colonized and non-algal colonized axes loading score across three-dimensional space indicate that Colorado (CO) fungal communities are more similar between paired samples than Washington (WA) paired fungal communities (Euclidian distance between paired samples based on t-test).
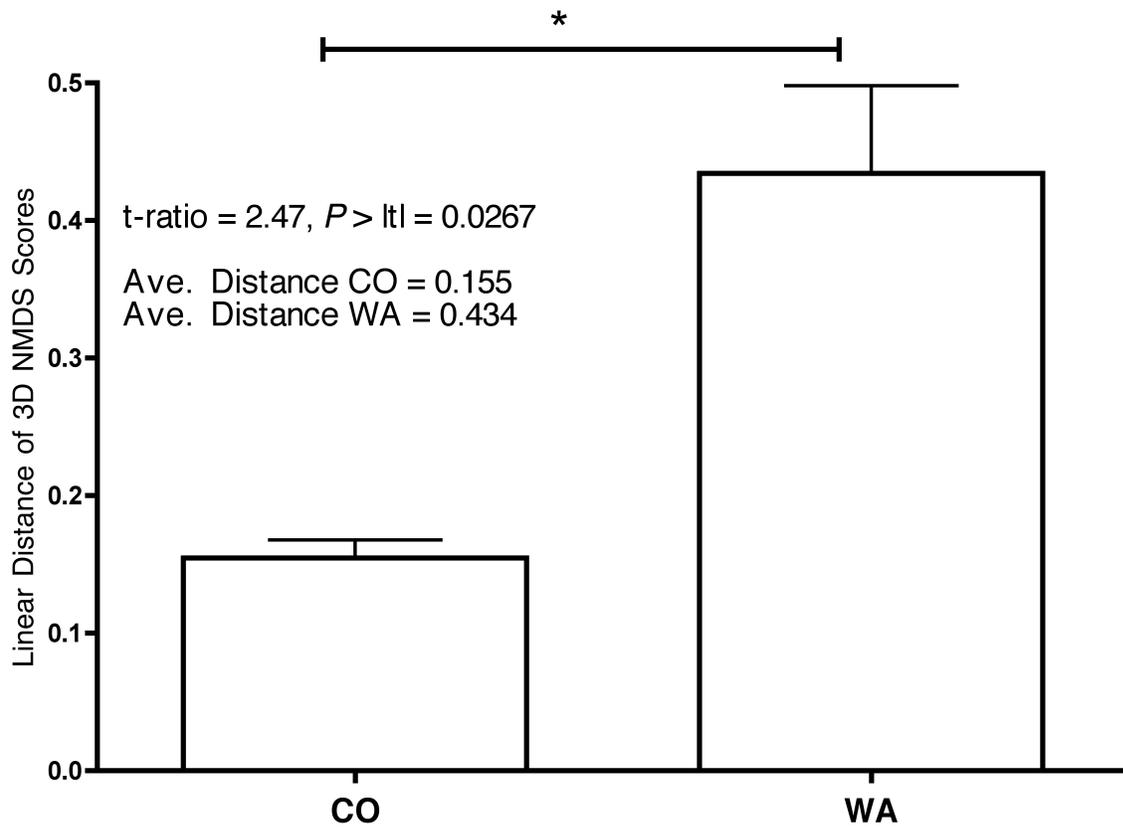
Appendix D - Supplemental Information for Chapter 5

Table D.1 Snow algae haplotype identification and frequencies. Haplotype distributions of *Coenochloris* (top) and *Chlamydomonas* (bottom) as described in Chapter 5. Lyman Samples are from the Lyman Glacier basin in Washington State and the Niwot and Indian Peaks samples are from Colorado. Sampling years are represented parenthetically. Haplotype frequency represents the proportion of a given haplotype from the same sampling locations and year (haplotype frequency out of 50 analyzed sequences per sample).

Haplotype distribution of *Coenochloris*			
Location/Sample	Haplotype ID	Haplotype Count	Haplotype Frequency (within sample)
Lyman 1 (2011)	13	45	90%
Lyman 1 (2011)	21	5	10%
Lyman 2 (2011)	12	44	88%
Lyman 2 (2011)	19	6	12%
Lyman 3 (2011)	14	44	88%
Lyman 3 (2011)	20	6	12%
Lyman 4 (2011)	7	49	98%
Lyman 4 (2011)	30	1	2%
Lyman 5 (2011)	6	49	98%
Lyman 5 (2011)	27	1	2%
Lyman 6 (2011)	11	46	92%
Lyman 6 (2011)	22	4	8%
Lyman 1 (2012)	12	1	2%
Lyman 1 (2012)	16	42	84%
Lyman 1 (2012)	17	1	2%
Lyman 2 (2012)	2	50	100%
Lyman 3 (2012)	3	50	100%
Lyman 4 (2012)	8	49	98%
Lyman 4 (2012)	25	1	2%
Lyman 5 (2012)	4	50	100%
Lyman 6 (2012)	9	48	96%
Lyman 6 (2012)	23	2	4%
Niwot 1 (2011)	15	42	84%
Niwot 1 (2011)	18	8	16%
Niwot 1 (2011)	26	1	2%
Niwot 2 (2011)	5	49	98%
Niwot 2 (2011)	28	1	2%
Indian Peaks 1 (2012)	10	47	94%

Indian Peaks 1 (2012)	24	2	4%
Indian Peaks 1 (2012)	29	1	2%
Indian Peaks 2 (2012)	1	50	100%
Haplotype distribution of *Chlamydomonas*			
Lyman 1 (2011)	4	50	100%
Lyman 2 (2011)	2	50	100%
Lyman 3 (2011)	11	49	98%
Lyman 3 (2011)	27	1	2%
Lyman 4 (2011)	6	50	100%
Lyman 5 (2011)	5	50	100%
Lyman 6 (2011)	13	47	94%
Lyman 6 (2011)	20	3	6%
Lyman 1 (2012)	14	46	92%
Lyman 1 (2012)	19	3	6%
Lyman 1 (2012)	23	1	2%
Lyman 2 (2012)	3	50	100%
Lyman 3 (2012)	1	50	100%
Lyman 4 (2012)	9	49	98%
Lyman 4 (2012)	26	1	2%
Lyman 5 (2012)	8	49	98%
Lyman 5 (2012)	21	1	2%
Lyman 6 (2012)	12	48	96%
Lyman 6 (2012)	22	1	2%
Lyman 6 (2012)	24	1	2%
Niwot 1 (2011)	7	49	98%
Niwot 1 (2011)	18	1	2%
Niwot 2 (2011)	15	40	80%
Niwot 2 (2011)	18	10	20%
Indian Peaks 1 (2012)	10	49	98%
Indian Peaks 1 (2012)	25	1	2%
Indian Peaks 2 (2012)	16	27	54%
Indian Peaks 2 (2012)	17	23	46%

Appendix E - Supplemental Information for Chapter 6

Table E.1 Results from paired *t*-tests from Lyman Glacier Forefront and Stored *Sorghum* Biomass experiments comparing diversity estimators derived using ITS and pairwise- or reference-aligned LSU. Additionally, results of two-way ANOVA testing if Richness, Diversity (1-D), evenness (E_D) and NMDS axes loading scores change with time, differ across treatments, or interact between the main effects. Lyman glacier forefront study included soils sampled across glacier forefront under plants with different mycorrhizal ecologies – *Abies lasiocarpa*, *Luetkia pectinata*, *Phyllodoce empetriformis*, *Saxifraga ferruginea*, and non-vegetated bare soil – referred here by genus. The stored biomass experiment included treatments of different biomass covering – No Plastic/No Tarp (NN), No Plastic/Tarp (NT), Plastic/No Tarp (PN), Plastic/Tarp (PT) – referred here by their abbreviation. Significant treatment effects are denoted in bold**; upon a significant treatment effect, descriptions of community shifts are explained under the *Change* column.**

Paired T-Test					
Lyman Glacier Forefront	**Average Per Experimental Unit**	**Test**	**t-statistic**	**DF**	**Prob >|t|**
Richness (Sobs)					
ITS	23.286	ITS vs. LSU pairwise	0.480821	48	0.6328
LSU pairwise	24.133	ITS vs. LSU aligned	-2.02429	48	**0.0485**
LSU aligned	27.141				
Diversity (1-D)					
ITS	0.697	ITS vs. LSU pairwise	1.184744	48	0.242
LSU pairwise	0.728	ITS vs. LSU aligned	2.616595	48	**0.0118**
LSU aligned	0.763				
Evenness (E_D)					
ITS	0.196	ITS vs. LSU pairwise	0.480821	48	0.6328
LSU pairwise	0.191	ITS vs. LSU aligned	-2.02429	48	**0.0485**
LSU aligned	0.194				
Stored Sorghum Biomass	**Average Per Experimental Unit**	**Test**	**t-statistic**	**DF**	**Prob >|t|**
Richness (Sobs)					
ITS	34.245	ITS vs. LSU pairwise	-28.9788	95	**<.0001**
LSU pairwise	24.086	ITS vs. LSU aligned	-23.7532	95	**<.0001**
LSU aligned	25.699				
Diversity (1-D)					
ITS	0.924	ITS vs. LSU pairwise	-30.708	95	**<.0001**

LSU pairwise	0.876	ITS vs. LSU aligned	-27.888	95	**<.0001**
LSU aligned	0.878				
Evenness (E$_D$)					
ITS	0.395	ITS vs. LSU pairwise	-14.2719	95	**<.0001**
LSU pairwise	0.345	ITS vs. LSU aligned	-19.7925	95	**<.0001**
LSU aligned	0.330				

Two-way ANOVA			
Lyman Glacier Forefront	**F statistic (9,39 df)**	**P value**	**Change**
Richness (S$_{obs}$)			
ITS	0.881	0.5501	
LSU pairwise	0.618	.07747	
LSU aligned	0.985	0.468	
Diversity (1-D)			
ITS	0.707	0.6986	
LSU pairwise	0.566	0.8161	
LSU aligned	0.946	0.6304	
Evenness (ED)			
ITS	1.164	0.344	
LSU pairwise	1.062	0.4112	
LSU aligned	0.786	0.6304	
NMDS ITS1			
Axis 1	0.387	0.9344	
Axis 2	4.170	**0.0008**	Decrease with distance
Axis 3	3.197	**0.0054**	Bare soil and *Luetkia* different from *Abies*
NMDS LSU pairwise			
Axis 1	1.365	1.3654	Increase with distance
Axis 2	2.019	0.0631	Increase with distance
Axis 3	1.099	0.386	
NMDS LSU aligned			
Axis 1	4.705	**0.0003**	Decrease with distance
Axis 2	3.955	**0.0012**	Bare soil different from *Luetkia, Saxifraga, Phyllodoce*
Axis 3	1.139	0.3601	
Stored Sorghum Biomass	**F statistic (7,88 df)**	**P value**	**Change**
Richness (S$_{obs}$)			
ITS	2.160	**0.0455**	Increase with time
LSU pairwise	3.271	**0.0039**	Increase with time, NN larger than NT
LSU aligned	5.699	**<0.0001**	Increase with time, NN larger than NT, PN, PT
Diversity (1-D)			
ITS	2.005	0.0633	Increase with time
LSU pairwise	2.480	**0.0227**	Increase with time
LSU aligned	3.794	**0.0012**	Increase with time
Evenness (ED)			
ITS	3.032	**0.0067**	NN lower than NT, PN, PT
LSU pairwise	0.619	0.7391	
LSU aligned	0.575	0.7744	
NMDS ITS1			
Axis 1	5.140	**<0.0001**	Increase with time
Axis 2	1.893	0.0801	
Axis 3	16.440	**<0.0001**	NN different from NT, PN, PT

NMDS LSU pairwise			
Axis 1	29.801	**<0.0001**	Decrease with time; NN different from NT, PN, PT
Axis 2	3.165	**0.0049**	Decrease with time; NN different from NT, PN, PT
Axis 3	7.107	**<0.0001**	Increase with time; NN different from NT, PN, PT
NMDS LSU aligned			
Axis 1	8.074	**<0.0001**	Increase with time
Axis 2	12.389	**<0.0001**	Increase with time; NN different from NT, PN, PT
Axis 3	6.214	**<0.0001**	NN different from NT, PN, PT

Figure E.5 Order level taxonomic affinities for ITS and LSU gene regions for Lyman Glacier forefront and stored *Sorghum* biomass experiments. Labeled order percentages represent proportion of total OTUs that were classified to order-level for the ten most abundant orders for each experiment. The taxon affinities are largely but not exclusively congruent.
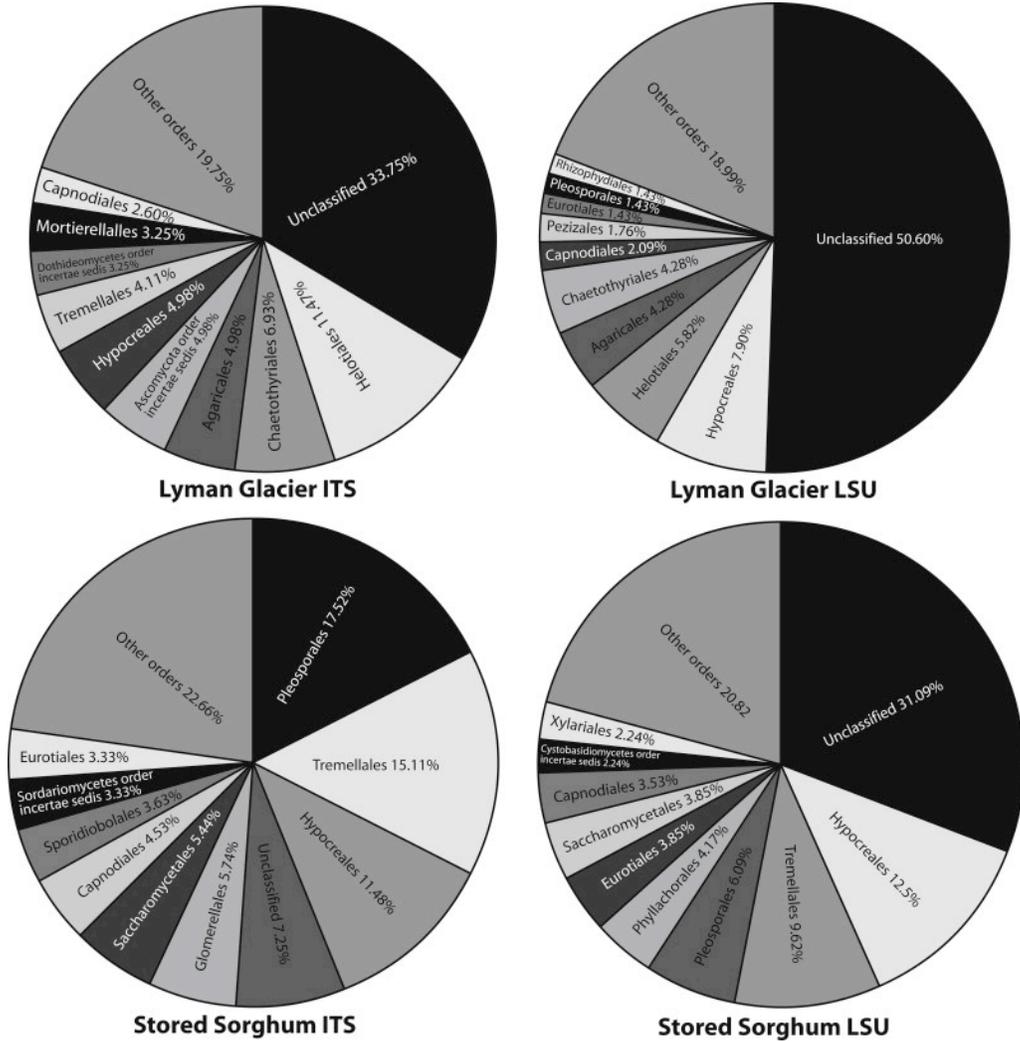
Appendix F - Supplemental Information for Chapter 7

Table F.1 Description of experiments from which singleton sequences were harvested from including region sequenced, sequence length used, the number of sequences used from each experiment, primers used, and references for those experiments.

Region	Description	Technology	Sequence Length	Number of sequences	Primers	References
ITS2	Snow Fungi, Colorado and Washington, USA	Illumina MiSeq	250	500	ITS4→fITS7	Brown *et al.*, 2014a
ITS2	Oak Bait, Kansas, USA	Illumina MiSeq	250	500	ITS4→fITS7	Lothamer (unpublished)
ITS2	EcM, Australia	Illumina MiSeq	250	500	ITS4→fITS7	Jumpponen (unpublished)
ITS2	Soil Fungi, Australia	Illumina MiSeq	250	500	ITS4→fITS7	Jumpponen (unpublished)
ITS2	EcM Yellow Pine, Georgia, USA, Phusion	Illumina MiSeq	250	500	ITS4→fITS7	Oliver (unpublished)
ITS2	EcM Yellow Pine, Georgia, USA, Phire	Illumina MiSeq	250	500	ITS4→fITS7	Oliver (unpublished)
ITS1	Periglacial Soil, Washington, USA	454-Titanium	250	350	ITS1f→ITS4	Brown and Jumpponen, 2014
ITS1	Sorghum Storage, Kansas, USA	454-Titanium	250	350	ITS1f→ITS4	Brown *et al.*, 2014b
ITS1	Oak Phyllosphere, Kansas, USA	454-FLX	200	200	ITS2→ITS1f	Jumpponen and Jones, 2009
ITS1	EcM, Oak, Kansas, USA	454-FLX	200	200	ITS2→ITS1f	Jumpponen *et al.*, 2010
ITS1	Oak Phyllosphere, Kansas, USA	454-FLX	200	200	ITS2→ITS1f	Jumpponen and Jones, 2010
LSU	Soil Fungi, Tallgrass Prairie June, Kansas, USA	454-Titanium	225	300	LR0R→LR3	Jumpponen (unpublished)
LSU	Soil Fungi, Tallgrass Prairie September, Kansas, USA	454-Titanium	225	300	LR0R→LR3	Jumpponen (unpublished)
LSU	Periglacial Soil, Washington, USA	454-Titanium	225	300	LR0R→LR3	Brown *et al.*, 2014b
LSU	Sorghum Storage, Kansas, USA	454-Titanium	225	300	LR0R→LR3	Ridgon-Huss *et al.*, 2014
LSU	Oak, Endophyte and Spore, Kansas, USA	454-Titanium	225	300	LR0R→LR3	Lickteig (unpublished)

Supplemental References

Brown S.P., Jumpponen A., 2014. Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. *Molecular Ecology* **23**: 481-497.

Brown S.P., Olson B.J.S.C., Jumpponen A., 2014a. Snow algae as the control of snow fungal community assembly. *Molecular Ecology* (in review).

Brown S.P., Rigdon-Huss A.R., Jumpponen A., 2014b. Analyses of ITS and LSU gene regions provide congruent results on fungal community responses. *Fungal Ecology* **9**: 65-68.

Jumpponen A., Jones K.L., 2009. Massively parallel 454 sequencing indicated hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytologist* **184**: 438-448.

Jumpponen A., Jones K.L., 2010. Seasonally dynamic fungal communities in the *Quercus macrocarpa* phyllosphere differ between urban and nonurban environments. *New Phytologist* **186**: 496-513.

Jumpponen A., Jones K.L., Mattox J.D., Yaege C., 2010. Massively parallel 454-sequencing of fungal communities in *Quercus* spp. ectomycorrhizas indicates seasonal dynamics in urban and rural sites. *Molecular Ecology* **19(s1)**: 41-53.

Rigdon-Huss, A., Brown, S.P., Maier, D.E., and A. Jumpponen. Fungal communities in stored lignocelulosic biomass respond strongly to storage conditions and length of storage. *Fungal Ecology* (submitted).

Table F.2 Proportion of singletons for each region clustered with other singleton sequences, the proportion of newly clustered sequences that are traced back to cross-experiment samples, and the non-ITS target amplification.

	LSU Titanium	ITS1 FLX	ITS1 Titanium	ITS2 MiSeq
Clustered Singletons	11.47%	0.83%	0.43%	2.27%
Cross-Experiment Clusters	21.51%	0.84%	0.43%	1.06%
Non-ITS Singletons	N/A	5.33%	2.86%	4.53%

Table F.3 Complete taxonomic identities of sequences that could not be traced to ITS regions using ITSx based on best BLASTn match while excluding environmental sequences with associated best BLASTn match accession number.

Region and Chemistry	Direction of sequencing				
ITS1 FLX	**ITS2-->ITS1F**				
32 Failed Extractions using ITSx					
Sequence Name	**Accession**	**Region Matched**	**Phylum**	**Genus**	**Species**
>E6TGR3Z08EP0ME63|1	KC800861.1	5.8S	Ascomycota	Retroconis	
>E622HTN11GYHUM118|1	JX427053.1	ITS1	Ascomycota	Preussia	minima
>E6TGR3Z07D8TWW154|1	KF850373.1	ITS1	Ascomycota	Trichocladium	opacum
>FJSKC4L04JPQFT304|1	CP001713.1	Bacteria	Chlamydiae	Chlamydia	pneumoniae
>FJSKC4L04IMQ45308|1	GU214539	SSU (3')	Ascomycota	Mycosphaerella	graminicola
>FJSKC4L04JRRRQ335|1	HQ634845	SSU (3')	Ascomycota	Meria	laricis
>FJSKC4L04IIJFJ340|1	GU180622.1	SSU (3')	Ascomycota		
>FJSKC4L04JW4Z8343|1	XM_002868301.1	Hypothetical Protein	Streptophyta	Arabidopis	lyrata
>FJSKC4L04H9X7T357|1	AY654886.1	SSU (3')	Basidiomycota	Laccaria	ochropurpurea
>FJSKC4L04IFEL9358|1	KF155200.1	SSU (3')	Ascomycota	Exophiala	equina
>FJSKC4L04JJ510366|1	EU940028.1	SSU (3')	Ascomycota	Epibryon	diaphanum
>FJSKC4L04JJ0VJ368|1	JF836023.1	SSU (3')	Ascomycota	Archaeorhizomyces	
>FJSKC4L04IFDKU369|1	EU940081.1	SSU (3')	Ascomycota	Hymenoscyphus	fructigenus
>FJSKC4L04IPU1G372|1	AB521040.1	SSU (3')	Ascomycota	Pseudigymnoascus	
>FJSKC4L04J0JYP390|1	GU180609.1	SSU (3')	Ascomycota	Kylindria	peruamazonensis
>FJSKC4L04JG797395|1	GQ280420.1	SSU (3')	Ascomycota	Neoscytalidium	dimidiatum
>FJSKC4L04IVJWC400|1	KF381538.1	SSU (3')	Basidiomycota	Lichenomphalia	umbellifera
>FJSKC4L04JHPRK401|1	JQ742003.1	Retrotransposon	Streptophyta	Castanea	mollissona
>FJSKC4L04IDN6C402|1	DQ834911.1	SSU (3')	Basidiomycota	Skortzovia	furfurella
>FJSKC4L04J337P405|1	GU187618.1	SSU (3')	Basidiomycota	Anomoloma	albolutescens
>FJSKC4L04JUBYK414|1	AJ437205.1	SSU (3')	Glomeromycota		
>FJSKC4L04IXP8G416|1	KF297951.1	SSU (3')	Basidiomycota	Auricularia	polytricha
>FJSKC4L04JOARJ422|1	JF836020.1	SSU (3')	Ascomycota	Archaeorhizomyces	
>FJSKC4L04IVRQS424|1	DQ898724.1	SSU (3')	Basidiomycota	Sistotrema	anthelioides
>FJSKC4L04H7LL6530|1	EU940065.1	SSU (3')	Ascomycota	Bryoscyphus	
>FJSKC4L04H989M531|1	KC969085.1	SSU (3')	Ascomycota	Saccharomyces	cerevisiae
>FJSKC4L04H6LBF535|1	AF518571.1	SSU (3')	Basidiomycota	Asterostroma	andinum
>FJSKC4L04I69RR541|1	FR750376.1	SSU (3')	Glomeromycota	Glomus	macrocarpum
>FJSKC4L04H41FY542|1	AJ716309.1	ITS1	Glomeromycota		
>FJSKC4L04I6PKY546|1	EU940080.1	SSU (3')	Ascomycota	Hyaloscypha	vitreola
>FJSKC4L04H61HY548|1	KC290944.1	ITS2	Ascomycota	Candida	aquae-textoris
>FJSKC4L04I7ZZP560|1	EU940046.1	SSU (3')	Ascomycota	Pleostigma	jungermannicola
Region and Chemistry	**Direction of sequencing**				
ITS1 Titanium	**ITS1F-->ITS4**				
20 Failed Extractions using ITSx					
Sequence Name	**Accession**	**Region Matched**	**Phylum**	**Genus**	**Species**
>HH1STK303GRAZH139|1	EU040233.1	ITS1	Ascomycota	Ochrocladosporium	elatum
>HH1STK303FHRV9166|1	EU552104.1	ITS1	Ascomycota	Camarosporium	brabeji
>GQT2G8G02D3LXI369|1	JQ272347.1	ITS1	Ascomycota		
>GQT2G8G02DGJF7382|1	AM292049.1	ITS1	Ascomycota	Trichocladium	opacum
>GQT2G8G02DHWBJ383|1	GU581246.1	ITS1			
>GQT2G8G02DXYRP391|1	KF156117.1	ITS1	Ascomycota	Paraconiothyrium	sporulosum

>GQT2G8G02DNPJ2417|1	JF340244.1	ITS1	Ascomycota		
>GQT2G8G02ERL36419|1	HQ601953.1	ITS1			
>GQT2G8G02DFC9Y440|1	JX967274.1	SSU	Cryptomycota	Amoeboaphelidium	protococcarum
>GQT2G8G02DSCLX472|1	EF093147.1	ITS1	Ascomycota		
>GQT2G8G02DXD8J485|1	EF093147.1	ITS1	Ascomycota		
>GQT2G8G02DTYVT526|1	KF212324.1	ITS1			
>GQT2G8G02DZGML567|1	EF585637.1	5.8S/ITS2	Chytridiomycota		
>GQT2G8G02D05UP585|1	GU581246.1	ITS1			
>GQT2G8G02CYP94587|1	FJ210518.1	ITS1	Ascomycota	Preussia	
>GQT2G8G02EIH6N612|1	AB691991.1	ITS1	Basidiomycota	Entoloma	
>GQT2G8G02DZKU9656|1	KC009773.1	ITS1	Ascomycota	Penicillium	chrysogenum
>GQT2G8G02D9CMG668|1	FJ008678.1	ITS1	Ascomycota		
>GQT2G8G02DZ46G688|1	KC592278.1	ITS2	Ascomycota	Trypetghliaceae	Trypethelium
>GQT2G8G02DHNA9692|1	JQ272347.1	ITS1	Ascomycota		
Region and Chemistry	**Direction of sequencing**				
ITS2 MiSeq	**ITS4-->fITS7**				
136 Failed Extractions using ITSx					
Sequence Name	**Accession**	**Region Matched**	**Phylum**	**Genus**	**Species**
>OTU44_2	HQ318283.1	ITS2	Basidiomycota	Lactarius	clarkeae
>OTU51_2	AF058495.1	ITS2	Streptophyta	Eucalyptus	sieberi
>OTU55_2	HM596063.1	ITS2	Streptophyta	Eucalyptus	regnans
>OTU76_1	FM213350.1	ITS2	Basidiomycota	Scleroderma	polyrhizum
>OTU154_1	JF273556.1	ITS2	Basidiomycota		
>OTU155_1	AF190360.1	ITS2	Streptophyta	Corymbia	eximia
>OTU176_1	KF245520.1	ITS2	Basidiomycota	Russula	subfoetens
>OTU192_1	KF218970.1	ITS2	Basidiomycota	Sistotrema	oblongisporum
>OTU423_1	AL844518.5	Zebrafish	Chordata	Danie	rerio
>OTU425_1	KF944462.1	ITS2		Mortierella	zonata
>OTU431_1	JX976020.1	ITS2		Mortierella	minutissima
>OTU438_1	JX975905.1	ITS2		Mortierella	verticillata
>OTU442_1	JX975905.1	ITS2		Mortierella	verticillata
>OTU445_1	HQ318283.1	ITS2	Basidiomycota	Lactarius	clarkeae
>OTU448_1	JX425382.1	ITS2	Basidiomycota	Russula	
>OTU463_1	AK406449.1	Silkworm	Arthropoda	Bonbycinae	
>OTU468_1	GU222317.1	ITS2	Basidiomycota	Clavulina	samuelsii
>OTU498_1	DQ411539.1	ITS2	Basidiomycota	Trechispora	alnicola
>OTU504_1	JX030195.1	ITS2	Basidiomycota	Scleroderma	polyrhizum
>OTU508_1	AB458893.1	ITS2	Basidiomycota	Lactarius	piperatus
>OTU509_1	HQ176328.1	ITS2	Cercoza	Thaumatomonas	
>OTU517_1	KC581351.1	ITS2	Basidiomycota	Hygrocybe	singeri
>OTU531_1	KC581351.1	ITS2	Basidiomycota	Hygrocybe	singeri
>OTU545_1	HQ176329.1	ITS2	Cercoza	Thaumatomonas	
>OTU556_1	FJ797471.1	ITS2	Cercoza	Eocercomonas	
>OTU560_1	JF273556.1	ITS2	Basidiomycota		
>OTU565_1	GU559984.1	ITS2		Mortierella	
>OTU566_1	NR_073311.1	ITS2	Basidiomycota	Kurtzmanomyces	tardis
>OTU582_1	HQ176331.1	ITS2	Cercoza	Allas	diplophysa
>OTU646_1	JQ950570.1	ITS2	Streptophyta	Viola	dissecta
>OTU685_1	AC104455.2	Human	Chordata	Homo	Sapiens
>OTU696_1	JX425382.1	ITS2	Basidiomycota	Russula	
>OTU725_1	JQ920362.1	ITS2	Chlorophyta	Neocystis	brevis
>OTU733_1	JQ281874.1	ITS2	Chlorophyta		
>OTU751_1	GQ922556.1	ITS2		Mortierella	alpina
>OTU752_1	KF944458.1	ITS2		Mortierella	elongata
>OTU763_1	EU019915.1	ITS2	Basidiomycota	Russula	neerimea

>OTU766_1	GU214810.1	ITS2	Basidiomycota	Tomentella	fuscocinerea
>OTU1160_1	AB634256.1	ITS2	Basidiomycota		
>OTU1387_1	AF466301.1	ITS2	Ascomycota	Neophaeosphaeria	barrii
>OTU1514_1	CP002403.1	Bacteria	Firmicutes	Ruminococcus	albus
>OTU1611_1	CP000473.1	Bacteria	Acidiobacteria	Candidatus Solibacter	usitatus
>OTU1658_1	AB634256.1	ITS2	Basidiomycota		
>OTU1808_1	AB634256.1	ITS2	Basidiomycota		
>OTU1834_1	AF272942.1	ITS2	Basidiomycota	Tomentella	fusco-cinerea
>OTU1947_1	GQ499379.1	LSU	Chytridiomycota		
>OTU1954_1	JQ711813.1	ITS2	Basidiomycota	Tomentella	
>OTU2002_1	HF679027.1	Hypothetical Protein	Ascomycota	Fusarium	fujikuroi
>OTU2023_1	CP001965.1	Bacteria	Proteobacteria	Sideroxydans	lithotrophicus
>OTU2024_1	AB634256.1	ITS2	Basidiomycota		
>OTU2084_1	CP003683.1	Bacteria	Proteobacteria	Klebsiella	oxytoca
>OTU2228_1	HM596061.1	ITS2	Streptophyta	Eucalyptus	pulverulenta
>OTU2231_1	HM596061.1	ITS2	Streptophyta	Eucalyptus	pulverulenta
>OTU2233_1	U83482.1	ITS2	Basidiomycota	Tomentella	
>OTU2242_1	FJ494728.1	ITS2	Streptophyta	Eucalyptus	rossii
>OTU2254_1	HM116969.1	ITS2	Streptophyta	Eucalyptus	porosa
>OTU2255_1	EF051503.1	ITS2	Streptophyta	Eucalyptus	triflora
>OTU2261_1	HM116969.1	ITS2	Streptophyta	Eucalyptus	porosa
>OTU2263_1	AM882810.2	ITS2	Basidiomycota	Inocybe	hystrix
>OTU2265_1	HM116970.1	ITS2	Streptophyta	Eucalyptus	siderophloia
>OTU2266_1	HM116970.1	ITS2	Streptophyta	Eucalyptus	siderophloia
>OTU2268_1	JX178490.1	ITS2	Basidiomycota	Russula	umerensis
>OTU2280_1	AY178408.1	ITS2	Streptophyta	Callitris	rhomboidea
>OTU2281_1	AF390533.1	ITS2	Streptophyta	Eucalyptus	pilularis
>OTU2289_1	JX178634.1	ITS2	Basidiomycota	Gallacea	eburnea
>OTU2293_1	KF245520.1	ITS2	Basidiomycota	Russula	subfoetens
>OTU2294_1	AF058495.1	ITS2	Streptophyta	Eucalyptus	sieberi
>OTU2317_1	EF694713.1	ITS2	Streptophyta	Eucalyptus	kitsoniana
>OTU2325_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2330_1	AF058482.1	ITS2	Streptophyta	Eucalyptus	radiata
>OTU2337_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2340_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2342_1	AY864901.1	ITS2	Streptophyta	Eucalyptus	tereticornis
>OTU2361_1	AF058482.1	ITS2	Streptophyta	Eucalyptus	radiata
>OTU2364_1	HM116969.1	ITS2	Streptophyta	Eucalyptus	porosa
>OTU2385_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2388_1	AF058482.1	ITS2	Streptophyta	Eucalyptus	radiata
>OTU2394_1	HM116969.1	ITS2	Streptophyta	Eucalyptus	porosa
>OTU2409_1	FJ494709.1	ITS2	Streptophyta	Eucalyptus	arenacea
>OTU2410_1	KC152242.1	ITS2	Basidiomycota	Thelephora	caryophyllea
>OTU2426_1	EF694713.1	ITS2	Streptophyta	Eucalyptus	kitsoniana
>OTU2436_1	AF058495.1	ITS2	Streptophyta	Eucalyptus	sieberi
>OTU2442_1	GU222292.1	ITS2	Basidiomycota	Russula	
>OTU2446_1	AY864901.1	ITS2	Streptophyta	Eucalyptus	tereticornis
>OTU2452_1	HM596065.1	ITS2	Streptophyta	Eucalyptus	rubiginosa
>OTU2460_1	KF218964.1	ITS2	Basidiomycota	Sistotrema	pistilliferum
>OTU2463_1	AJ716314.1	ITS2	Glomeromycota		
>OTU2471_1	HQ318283.1	ITS2	Basidiomycota	Lactarius	clarkeae
>OTU2477_1	HM596065.1	ITS2	Streptophyta	Eucalyptus	rubiginosa
>OTU2483_1	GQ366376	5.8S	Streptophyta	Corymbia	zygophylla
>OTU2491_1	HM596061.1	ITS2	Streptophyta	Eucalyptus	pulverulenta
>OTU2493_1	KC152242.1	ITS2	Basidiomycota	Thelephora	caryophyllea
>OTU2496_1	GQ366375.1	ITS2	Streptophyta	Corymbia	xanthope
>OTU2501_1	FM955848.1	ITS2	Basidiomycota	Tomentella	

>OTU2502_1	AF390477.1	ITS2	Streptophyta	Eucalyptus	
>OTU2503_1	AF058495.1	ITS2	Streptophyta	Eucalyptus	sieberi
>OTU2514_1	HM596065.1	ITS2	Streptophyta	Eucalyptus	rubiginosa
>OTU2522_1	HM116969.1	ITS2	Streptophyta	Eucalyptus	porosa
>OTU2531_1	AF058464.1	ITS2	Streptophyta	Eucalyptus	gunnii
>OTU2532_1	HM596061.1	ITS2	Streptophyta	Eucalyptus	pulverulenta
>OTU2536_1	HM116969.1	ITS2	Streptophyta	Eucalyptus	porosa
>OTU2538_1	KF668297.1	ITS2	Basidiomycota	Gymnopus	confluens
>OTU2539_1	AF058495.1	ITS2	Streptophyta	Eucalyptus	sieberi
>OTU2549_1	HQ318283.1	ITS2	Basidiomycota	Lactarius	clarkeae
>OTU2559_1	FJ494726.1	ITS2	Streptophyta	Eucalyptus	platydisca
>OTU2569_1	HM596065.1	ITS2	Streptophyta	Eucalyptus	rubiginosa
>OTU2574_1	HQ318284.1	ITS2	Basidiomycota	Lactarius	clarkeae
>OTU2587_1	AF058495.1	ITS2	Streptophyta	Eucalyptus	sieberi
>OTU2787_1	HM116969.1	ITS2	Streptophyta	Eucalyptus	porosa
>OTU2794_1	GQ366351.1	ITS2	Streptophyta	Corymbia	aparrerinja
>OTU2799_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2800_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2801_1	HQ176331.1	ITS2	Cercoza	Allas	diplophysa
>OTU2804_1	HM116969.1	ITS2	Streptophyta	Eucalyptus	porosa
>OTU2824_1	AY864901.1	ITS2	Streptophyta	Eucalyptus	tereticornis
>OTU2828_1	AF390521.1	ITS2	Streptophyta	Eucalyptus	deglupta
>OTU2832_1	HM596061.1	ITS2	Streptophyta	Eucalyptus	pulverulenta
>OTU2847_1	HM596061.1	ITS2	Streptophyta	Eucalyptus	pulverulenta
>OTU2848_1	AF390533.1	ITS2	Streptophyta	Eucalyptus	pilularis
>OTU2849_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2850_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2858_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2859_1	JX178490.1	ITS2	Basidiomycota	Russula	
>OTU2860_1	HM116971.1	ITS2	Streptophyta	Eucalyptus	vicina
>OTU2863_1	AY864901.1	ITS2	Streptophyta	Eucalyptus	tereticornis
>OTU2865_1	EF051503.1	ITS2	Streptophyta	Eucalyptus	triflora
>OTU2866_1	AY864901.1	ITS2	Streptophyta	Eucalyptus	tereticornis
>OTU2885_1	HM596065.1	ITS2	Streptophyta	Eucalyptus	rubiginosa
>OTU2886_1	AF058482.1	ITS2	Streptophyta	Eucalyptus	radiata
>OTU2911_1	AF058503.1	ITS2	Streptophyta	Eucalyptus	dives
>OTU2918_1	AY864901.1	ITS2	Streptophyta	Eucalyptus	tereticornis
>OTU2921_1	AB634273.1	ITS2	Basidiomycota		
>OTU2926_1	AY615672.1	ITS2	Streptophyta	Eucalyptus	
>OTU2928_1	AF390521.1	ITS2	Streptophyta	Eucalyptus	deglupta
>OTU2930_1	FN669278.1	ITS2	Basidiomycota		
>OTU2932_1	JX266623.1	ITS2	Basidiomycota	Russula	cheelii