available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/funeco

Habitat preferences, distribution, and temporal persistence of a novel fungal taxon in Alaskan boreal forest soils

Daniel J. GLASS[a,1], *Andrew D. TAYLOR*[b], *Ian Charold HERRIOTT*[a], *Roger W. RUESS*[a], *D. Lee TAYLOR*[a,*]

[a]*Institute of Arctic Biology, 311 Irving I Building, University of Alaska, Fairbanks, AK 99775, USA*
[b]*Department of Biology, University of Hawaii, Honolulu, HI 96822, USA*

ARTICLE INFO

Article history:
Received 2 December 2013
Revision received 3 February 2014
Accepted 21 February 2014
Available online 29 April 2014
Corresponding editor:
Nicole Hynson

Keywords:
Bonanza Creek
Habitat preference
Novel lineage
NS1
Soil horizon
Spatial autocorrelation
Temporal persistence
White spruce

ABSTRACT

Prior work in the boreal forest unearthed a novel ITS-LSU sequence (NS1) that falls outside known fungal phyla. Here we performed a targeted PCR survey to investigate the ecology of NS1. NS1 was found in three out of 99 soil cores at one black spruce (*Picea mariana*) site, and two cores were from nearby subplots, but clumping could not be demonstrated statistically. However, NS1 was detected 6 yr later in the same subplot, and in an adjacent subplot, with a join count probability of 0.0073. NS1 was not found in other lowland black spruce sites, but was detected in several upland mixed hardwood/white spruce (*Picea glauca*) sites and was correlated with presence of white spruce ($p = 0.0011$). It was also found in the same upland sites sampled in consecutive years. Our results provide clues concerning the ecology of NS1 and suggest that rare, divergent sequences should not necessarily be discarded from environmental sequence datasets.

© 2014 Elsevier Ltd and The British Mycological Society. All rights reserved.

Introduction

Fungi fulfill many crucial ecological functions as decomposers and plant symbionts (O'Brien et al., 2005). They play prominent roles in boreal forests due to their ability to function at low temperatures, low pH, and in nutrient poor environments (Taylor et al., 2010). Nevertheless, until relatively recently our understanding of boreal forest microbial ecological processes has largely been focused on functionality with little regard given to which species were fulfilling these functions (Taylor et al., 2010).

Beginning in 2003, a large-scale project to characterize soil fungal communities in the major stages of boreal forest succession using molecular methods was undertaken at the Bonanza Creek Long Term Ecological Research Site (BNZ-LTER) near Fairbanks, Alaska, USA (Taylor et al., 2010). These

* *Corresponding author*. Current address: Department of Biology, MSC03 2020, 1 University of New Mexico, Albuquerque, NM 98107, USA. Tel./fax: +1 011 505 277 0577.
E-mail addresses: fflt@unm.edu, ltaylor@iab.alaska.edu, dltaylor@alaska.edu (D.L. Taylor).
[1] Author deceased.
1754-5048/$ – see front matter © 2014 Elsevier Ltd and The British Mycological Society. All rights reserved.
http://dx.doi.org/10.1016/j.funeco.2014.03.001

sites are representative of the boreal forests of Interior Alaska; several lowland sites were dominated by 100–300 yr old black spruce (*Picea mariana*), while upland sites contained white spruce (*Picea glauca*) and either trembling aspen (*Populus tremuloides*) or Alaskan paper birch (*Betula neoalaskana*) ranging from 25 to ~300 yr old. These sites have been extensively studied as part of the BNZ-LTER program over the last 40 yr. Tens of thousands of soil clone sequences were analyzed and the majority of sequences appeared to be representatives of previously described fungal species, genera or families (Taylor et al., 2014). Of those that did not fall into known fungal clades, one sequence in particular, hereafter referred to as novel sequence one (NS1), appeared particularly divergent. Both BLAST (Altshul et al., 1997) and phylogenetic analyses place the sequence within the kingdom Eumycota but do not place it within any known phylum (Glass et al., 2013).

We conducted a variety of analyses to assess whether the NS1 sequence might be some sort of artifact such as a chimera or pseudogene. We found no evidence that it is a chimera using both Chimera Checker (Nilsson et al., 2010) and Uchime (Edgar et al., 2011). We then modeled the rRNA secondary structure of the NS1 sequence and compared its folding to that of other fungi to evaluate the presence of conserved motifs, domains, and compensatory base changes. The results of these analyses were consistent with NS1 representing functional rRNA rather than a biological or lab artifact (Glass et al., 2013). The original NS1 sequence was found in a soil clone library originating from a floodplain black spruce (*P. mariana*) stand within the BNZ-LTER (site code FP5C). A highly similar ~1 200 bp sequence, varying at only eight sites, was found in another clone library from a mid-successional upland site (mixed white spruce & paper birch, UP2A) amplified using different primers. These independent detections further support the authenticity of the sequence and suggest that it might occur in multiple boreal forest habitats in Interior Alaska.

Patterns of occurrence of DNA sequences in environmental samples can reveal valuable information regarding the ecologies of both well-known and unknown or novel species. The discovery of novel archael and bacterial 16S sequences in thermal hot springs ushered in the era of environmental genomics (Pace et al., 1985). Much can be learned about the habitat preferences of unseen organisms by tracking their DNA. Subsequently, focused efforts at visualization and culturing can be mounted. These efforts have been particularly successful, for example, with marine bacterioplankton (Rappé et al., 2002; Stingl et al., 2007), but have also been applied to fungi (Collado et al., 2007). Within the fungi, a clade of Ascomycota initially called Schadt clade 1 dominated soil clone libraries above the treeline in a study site in Colorado (Schadt, 2003). Subsequent environmental sequencing efforts demonstrated that members of this lineage are diverse and widely distributed across the globe (Porter et al., 2008). Most recently, a member of this clade, now named the Archaeorhizomyces, was brought into cultivation and characterized in much more detail (Rosling et al., 2011).

With respect to the ecologies of unseen and uncultured organisms, analyses of spatial autocorrelation can reveal clumped, random or over-dispersed distributions, each of which provides unique insights into the natural history of the organism. For example, over-dispersion is often taken as circumstantial evidence for the role of intra-specific competition. Occurrence of a sequence across seasons can illuminate seasonal dynamics (Schadt, 2003; Taylor et al., 2010), while occurrences across forest types (Geml et al., 2009), soil horizons (Dickie et al., 2002; Taylor et al., 2014), pH gradients (Toljander et al., 2006), edaphic variation (Schechter and Bruns, 2008) etc. reveal additional aspects of habitat preferences. Hence, even for organisms that have never been seen, molecular ecology can provide important clues about their ecology and evolution.

Here we utilize PCR surveys to gain further insight into the ecology of NS1. In particular, we assess its: (1) within-site distribution; (2) occurrence among forest and habitat types; (3) soil horizon preferences; and (4) temporal persistence.

Materials and methods

Initial detection of NS1

As described in previous publications, we collected soil cores from several BNZ-LTER sites near Fairbanks, AK, USA between 2003 and 2005. Large-scale PCR, cloning and clone library sequencing were conducted to characterize the fungal communities present. Site descriptions, DNA extraction and amplification methods have been described previously (Geml et al., 2010, 2009; Taylor et al., 2014, 2010). Briefly, the gene-region (~1 200 bp) encompassing the ribosomal internal transcribed spacers (ITS) and a portion (~700 bp) of the ribosomal large subunit (LSU) were amplified from soil extracts using the fungal-specific PCR primers ITS1-F (Gardes and Bruns, 1993) and TW13 (Taylor and Bruns, 1999). Amplicons were cloned using a TOPO-TA pCR 4.0 kit (Invitrogen, Carlsbad, CA, USA) and libraries were sent to the Broad Institute of MIT and Harvard where transformations, automated clone-picking, and sequencing of clone libraries took place. We refer to this as our "non-targeted" survey. The NS1 sequence comprised one out of 384 clones from the site FP5C sampled in 2003, and a closely related sequence comprised one out of 931 clones from site UP2A sampled in 2004 (Glass et al., 2013). FP5C is a black spruce site located adjacent to the Tanana river in the floodplain, while UP2A is a mid-successional upland site located approximately 2 miles from the Tanana River and is dominated by paper birch and white spruce.

Sites and samples

Two different sets of soil samples, which have been described and analyzed previously, were used to explore the ecology of NS1. The first set of samples was collected at FP5C, the site in which NS1 was originally discovered.

A 200 × 200 m study area was established within BNZ-LTER site FP5C. Within this area, nine square plots 3 m × 3 m were arrayed in a stratified-random spatial design using randomly generated coordinates (Fig 1). Corners of plots were marked with PVC pipe. Each plot was divided into nine square subplots of 1 m × 1 m. In 2003, three soil cores (1.8 cm in diameter × 30 cm deep) were collected from each subplot within plot 1 (27 cores in total), while a single core was

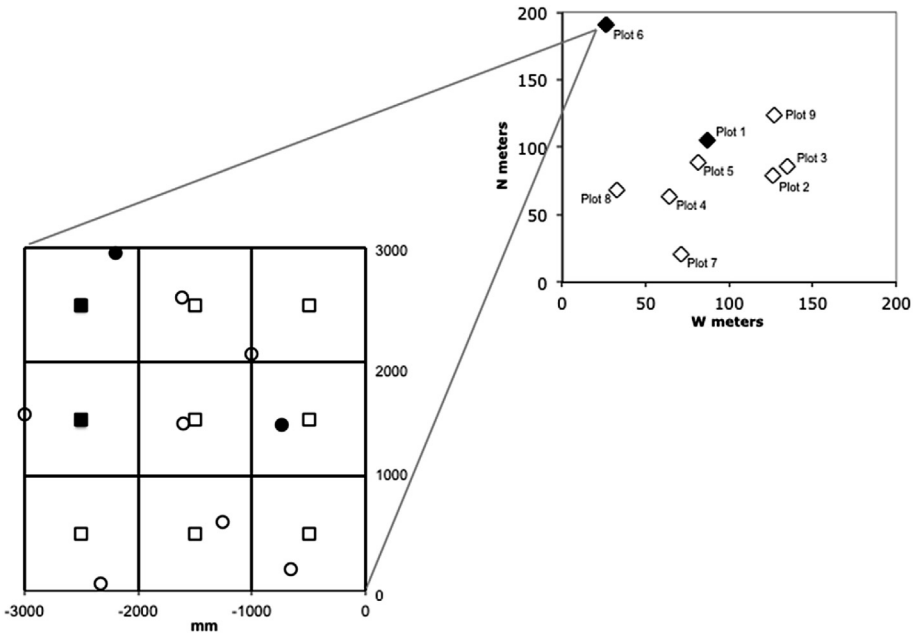

Fig 1 – Sampling design and occurrence of NS1 across site FP5C. The 200 × 200 m study area within FP5C is shown on the right, with the positions of the nine plots indicated with diamonds; open diamonds indicate that NS1 was absent, filled diamonds indicate the presence of NS1. On the left, plot 6 is blown up. Cores collected in 2003 are indicated with circles while cores collected in 2009 are indicated with squares. The four cores (across 2003 and 2009) in which NS1 was detected are shown with solid symbols. The additional cores from which NS1 was absent are shown with open symbols.

collected from each subplot within plots two through nine (72 additional cores, 99 in total). The X and Y coordinates of soil cores within the 1 m^2 subplots were selected using random number tables. We visually identified the organic, humic and mineral soil horizons and subsampled approximately 0.25 g of soil from each of these horizons. We resampled the nine subplots from plot 6 using the same protocol in 2009; this plot was chosen because NS1 was found here and it was impractical to resample the entire site.

To assess the occurrence of NS1 across major forest and habitat types within Bonanza Creek, we applied the same PCR-screening to an array of soil DNA extracts from samples collected in 2004 and 2005. The soil sampling, DNA extraction and analyses of total fungal communities were described in Taylor et al. (2010). Briefly, nine upland sites representing three stages of forest succession (early, mid and late) were screened. Fifty soil cores at 10 m intervals were collected along four parallel transects across each site. For each core, the organic and mineral soil horizons were separated and 0.25 g subsamples collected (Taylor et al., 2010). In contrast to the samples from site FP5C, subsamples were pooled by horizon before being frozen at −80 °C and lyophilized. DNAs from twelve previously studied mature, lowland black spruce sites representing a range of pH and moisture contents were also screened (Taylor et al., 2014).

Molecular diagnostics

Soil samples were frozen at −80 °C within 4 d of collection and then lyophilized. Prior to DNA extraction, samples were ground on a ball-mill with 8 mm steel beads in a walk-in freezer and using a vortex and garnet as specified in the kit protocols. For the 0.25 g soil samples collected from FP5C in 2003, we utilized the UltraClean-htp 96 Well Soil DNA kit (MoBio Laboratories Inc, Carlsbad, CA, USA). Soil DNA extractions from the pooled UP and lowland black spruce (labeled TKN) samples, as well as the 2009 samples from FP5C were conducted using the MoBio PowerMax Soil DNA Isolation kit. Approximately 1 g of organic soil or 2 g of mineral soil was used per extraction with the larger volume PowerMax kit.

We have previously developed a primer pair that is highly specific to NS1, L2F1 (5′CCCGGTCGATATATTTACGAGAAG 3′) and L2R2 (5′GGGCAGAGATGAATATGCTAACAC 3′) (Glass et al., 2013). The primers are situated in the ITS1 and ITS2 regions flanking the 5.8S and amplify a fragment of 290 bp. These primers were utilized to survey for additional forest sites in which NS1 might be present. Illustra PureTaq Ready-To-Go PCR beads (GE Healthcare Life Sciences, Pittsburgh, PA, USA) and an MJ Research PTC-225 Thermalcycler (Bio-Rad Laboratories, Hercules, CA, USA) were used in all PCRs. Negative controls containing only ultrapure water were included in each reaction mix. Amplification was carried out using the following program: 96 °C for 2 min, followed by 35 cycles 94 °C for 30 s, 55 °C for 40 s, 72 °C for 1 min, and a final elongation step of 72 °C for 10 min. Following thermocycling, we loaded 5 μl of each PCR product onto 1.5 % agarose gels, which were run at 85 V for 40 min and visualized using EtBr.

When amplification was detected, we attempted to sequence the amplicon to ensure that it originated from an NS1-like organism. In most cases, we simply sequenced the

amplicon directly. However, for a subset of samples, we cloned amplicons using the Invitrogen TOPO-TA pCR 4.0 kit in order to evaluate any hidden sequence variation. Amplicons were cycle-sequenced using Big Dye Terminator v3.1 (Applied Biosystems, CA, USA) chemistry. Reactions were purified with Sephadex in Centri-Sep columns (Princeton Seperations Inc., NJ, USA) then loaded on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Carlsbad, CA, USA) at the University of Alaska Fairbanks, Institute of Arctic Biology, Core Facility for Nucleic Acid Analysis. In some cases, samples were sent to McLab (South San Francisco, CA, USA) for PCR cleanup and sequencing. Clones were sequenced in the same way, but using M13 vector primers. We refer to PCR with the NS1-specific primers as our "targeted" survey. Note that this approach provides no information on the diversity or abundance of other fungi present in these soils, in contrast to the non-targeted surveys that yielded only two NS1 sequences out of thousands of sequences analyzed (Taylor et al., 2014, 2010).

Statistical analyses

We tested for spatial autocorrelation of the 2003 distribution of NS1 across soil cores within FP5C using a simple form of the join count method, which is appropriate for spatial point occurrences where the response at a point is a nominal data type (Sokal and Oden, 1978). With our binary presence/absence data, this method assesses the frequency of neighboring points which both have NS1 present (positive joins). We defined points as subplots, and for the subplots in plot 1 with multiple cores the subplot was scored as NS1 being present if it was detected in any of the cores. Because the plots were not contiguous, we could define neighborhoods at only two scales: adjacent subplots, and all subplots in a plot. The observed frequencies of positive joins were compared with random expectations given the overall observed frequency of subplots in which NS1 was present. Due to the low incidence of NS1, we did not use asymptotic methods assuming a normal distribution but instead calculated exact probabilities; the analysis using plots as neighborhoods is exactly equivalent to Fisher's exact test applied to a 2×9 table of the frequencies of positive and negative subplots in each of the nine plots.

We then tested for correlation between the 2003 and 2009 distributions of NS1 within FP5C as follows. Because only plot 6 was resampled in 2009, we could not directly compare the entire distributions across both years using standard methods. We instead performed an *ad hoc* analysis comparing the similarity between the 2003 and 2009 distributions in plot 6, under the assumption that the overall incidence of NS1 across the nine plots was the same in 2009 as in 2003. Because the details of this analysis depended on the distributions observed in 2003 and 2009, they are described in the Results section below.

We assessed soil horizon preferences by comparing counts of NS1 occurrence in the pooled organic vs mineral samples from the upland sites collected in 2005 using the Fisher's exact test in R ($n = 18$; Table 1; R Development Core Team, 2008). Lastly, we tested whether the occurrence of NS1 was related to the presence of white spruce within a site, again using Fisher's Exact test for a 2×2 contingency table. White spruce densities in trees per hectare (>2.5 cm dbh) were obtained from public data provided on the BNZ-LTER web site and are given in Table 2.

Table 1 – Presence of NS1 across soil DNA extracts

Description	NS1	Total
Non-pooled soil extracts		
FP5C (BS) 2003 organic	3	81
FP5C (BS) 2009 organic	2	9
Pooled soil extracts		
Upland 2004 organic	3	9
Upland 2005 organic[a]	3	9
Upland 2005 mineral[a]	1	9
TKN black spruce organic	0	12
Total	**12**	**129**

a Used in Fisher's exact test for soil horizon preferences.

Table 2 – Occurrence of NS1 in relation to density of white spruce in each site

Site	White spruce stems/hectare	NS1 present
UP1a	50	–
UP1b	0	–
UP1c	0	–
UP2a	1 416.667	+
UP2b	633.334	+
UP2c	191.666	–
UP3a	406.667	+
UP3b	306.667	+
UP3c	306.667	+
FP5C	>0	+
TKN0001	0	–
TKN0012	0	–
TKN0015	0	–
TKN0022	0	–
TKN0039	0	–
TKN0040	0	–
TKN0051	776.262	–
TKN0109	0	–
TKN0119	0	–
TKN0122	0	–
TKN0123	0	–
TKN0126	0	–

Results

Overall, a band of ~290 bp indicative of the presence of NS1 was detected in 12 of the 129 DNA extracts assayed (Table 1). In addition to the original two 1 200 bp sequences obtained from non-targeted soil clone libraries, 34 clones were sequenced from a single FP5C 290 bp amplicon; another five direct sequences were obtained from various upland plots. All sequences were highly similar to the original 1 200 bp NS1 sequences (Glass et al., 2013); in total, only 10 single-base differences were seen among the 41 sequences examined. Mutations were rarely shared across haplotypes (i.e. they were autapomorphies) and phylogenetic analysis produced a star phylogeny with shallow branches and no supported

groupings. Hence, we believe that the majority of these apparent differences arose from errors during PCR, cloning and sequencing.

Site FP5C

NS1 was found in three out of 81 subplots within site FP5C (Fig 1). None of the positive subplots were adjacent to each other, so there was no evidence of clustering at the smallest scale we could measure. Two of the positive subplots were in plot 6, located in the northwest corner of the site. Hence, there is some suggestion of a clumped distribution of NS1 at the scale of plots. However, the probability of at least this frequency of positive joins, if the three occurrences of NS1 were randomly distributed across the 81 subplots, was 0.28, providing no persuasive evidence for spatial autocorrelation.

We were only able to resample plot 6 in 2009. Strikingly, NS1 was still present in this plot 6 yr later. Furthermore, it was still present in one of the two subplots in which it was seen in 2003 (Fig 1) and was detected in an adjacent subplot as well. To test the statistical significance of this spatial correspondence between the two years, we calculated the probability of at least this degree of similarity, assuming that occurrences of NS1 in 2009 were distributed independently of those in 2003 and that the overall frequency of occurrence of NS1 across all nine plots was the same in 2009 as in 2003, i.e. NS1 was present in three out of 81 subplots. The total number of ways three occurrences could be distributed across the 81 subplots was first calculated. The number of these arrangements in which at least one of the three occurrences was in one of the three subplots in which NS1 had been present in 2003, and at least one of the other 2009 occurrences was in an adjacent subplot was then tabulated. The p-value for this test is the latter count of arrangements divided by the former, which gave $p = 0.0073$. (This test actually is conservative in that the count of arrangements in the numerator includes those in which the repeat occurrence and adjacent occurrence was in plot 1 rather than plot 6, an outcome that could not have been detected since plot 1 was not resampled in 2009). Hence, the repeated occurrence of NS1 in nearby locations sampled 6 yr apart is unlikely to have occurred by chance, and provides some initial suggestion of a persistent distribution.

Other sites

Analyses of our DNAs from pooled soil cores that represent major forest stand types within the BNZ-LTER provide additional insight into the distribution of NS1. Surprisingly, although found repeatedly in the floodplain black spruce site FP5C, NS1 was not found in any of the other 12 lowland black spruce sites examined (Table 1). However, NS1 was found in a number of the upland sites. This is in agreement with its presence in the original soil clone library, where NS1 was detected in site UP2A in both 2004 and 2005 sampling events. NS1 was also detected by targeted PCR in several additional sites where it was not found in non-targeted clone libraries. These sites included all of the mature white spruce stands, and two mid-succession stands composed of similar proportions of white spruce and Alaska paper birch. It was detected in these sites in both 2004 and 2005, again suggesting some degree of temporal persistence. NS1 was not found in site UP2C, a mid-succession stand dominated by aspen, with a much lower proportion of white spruce and paper birch. Nor was NS1 found in any UP1 sites. These are early succession stands dominated by a mixture of aspen, willow, and alder.

When the occurrence of NS1 was compared to the occurrence of white spruce across sites (Table 2), Fisher's Exact test showed a strong association ($p = 0.0011$, phi correlation coefficient = 0.74). Hence, NS1 is statistically more likely to be found in sites in which white spruce is present.

Soil horizon preferences

Overall, NS1 was present in 11 of 120 organic soil extracts and one out of nine mineral extracts (Table 1). To carry out a formal test, we only considered the UP sites collected in 2005 ($n = 18$) where NS1 was found in three out of nine organic samples and one out of nine mineral samples. Fisher's exact test was non-significant (df = 1, $p = 0.58$), providing no evidence for a horizon preference for NS1.

Discussion

The specificity of our primers for the NS1 lineage appears to be very high based on our sequencing results. Overall, NS1 was rare across the spectrum of soil DNAs tested, as expected given its low incidence in non-targeted surveys from the same sites. The targeted PCR results correspond fairly well with the clone libraries. For example, NS1 was detected in both 2004 and 2005 from the UP2A organic horizon DNAs, one of the two samples in which it was found in non-targeted surveys. NS1 was also detected in several upland sites (UP2B and the UP3 series) where it was not found in clone libraries. The UP2 and UP3 sites yielded 225-616 clones in our non-targeted survey (Taylor et al., 2010). The failure to find NS1 in these libraries is not surprising due to the stochastic nature of recovering rare clones. These results also suggest that our primers are quite sensitive, with a detection limit somewhere below 0.1 % of the total fungal community, if we assume that our non-targeted approach using ITS1-FL and TW13 is unbiased (Taylor et al., 2007).

Although originally detected in a black spruce dominated site, our targeted PCR surveys suggest that NS1 is actually more common in upland forests dominated by hardwoods and white spruce. While correlation is not causation, it is interesting that all the upland sites in which NS1 was detected have a substantial component of white spruce. It is possible that NS1 interacts directly with white spruce in some fashion. The Archaeorhizomyces are similar in that they are detected almost exclusively in rhizosphere and root tip samples (Rosling et al., 2013, 2011). However, in neither case have direct interactions been observed, and the correlations could be driven by indirect mechanisms. For example, NS1 could be a mycoparasite, with a host that interacts with white spruce. In this study, we failed to detect NS1 in any of the 12 lowland black spruce sites tested other than FP5C. It was also not detected in clone libraries from these sites (Taylor et al., 2014). FP5C differs somewhat from the other 12 black spruce sites in that it is the only riparian site. Furthermore, the FP5C site transitions from black to white spruce at its northern boundary. Plot 6, where

NS1 was most frequent, lies at the northeastern edge, and may be within the rooting zone of white spruce.

While some fungal taxa have very strong soil horizon preferences, others appear to be generalists with little selectivity for particular horizons. NS1, though rare, appears to fall into the latter category. It is tempting to draw inferences about the possible trophic niche of NS1 from the lack of horizon preference. However, fungi within particular guilds range from horizon generalists to specialists. For example, several species of the EMF genus *Cortinarius* displayed strong horizon preferences across the same set of black spruce sites in Interior Alaska (Taylor et al., 2014), yet most species within the EMF genus *Russula* had little or no horizon preference across these sites (Geml et al., 2010). Similarly, saprotrophic taxa range from specialists to generalists (Lindahl et al., 2006; Taylor et al., 2014).

The data from FP5C, though not statistically significant, are suggestive of a somewhat clumped distribution of NS1. Prior studies of both sporocarps and soil extracts have shown certain fungal taxa to have clumped distributions (Taylor and Bruns, 1999; Bergemann and Miller, 2002; Lilleskov et al., 2004). This could be due to microsite preferences or to the expansion of initial colonizing genets (Taylor and Bruns, 1999). Spatial clumping has been most often investigated in dominant taxa. It is noteworthy that NS1 appears to be extremely low in abundance yet may also have a clumped distribution.

Overall, NS1 appears to show little variability between two sampling events spaced 6 yr apart at site FP5C. Furthermore, it was detected from samples collected in consecutive years in several of the upland sites. Studies have shown that members of the Russulaceae with highly clumped distributions also demonstrated prolonged temporal persistence (Bergemann and Miller, 2002). In one case the same genets were found up to 11 yr later (Bergemann and Miller, 2002). Since NS1 persisted in the same locations for 6 yr, it does appear to be either long-lived or reproduces successfully on a local scale.

Few studies have analyzed entire soil fungal communities across a site. Far fewer have compared soil fungal communities at the same site across multiple years (Izzo et al., 2005). Izzo et al. (2005) found that the same dominant taxa were present on a coarse spatial scale (>25 m) from year to year but community structure varied on a finer scale (5 cm). Only 23 % of species were found in the same plot every year (Izzo et al., 2005). This high degree of community composition turnover typical of late successional forests makes the temporal persistence of NS1 particularly interesting.

This study provides only a superficial initial glimpse into the ecology of NS1. The two most similar discoveries within the Fungi are the Rozella/Cryptomycota clade and the Archaeorhizomycetes. Both of these clades appear to encompass hundreds to thousands of species (Jones et al., 2011; Rosling et al., 2013). Members of the Cryptomycota appear to be nearly ubiquitous in marine habitats (Jones et al., 2011). Similarly, members of the Archaeorhizomycetes have been found in soil samples all over the world, although they appear to be most common in boreal soils (Rosling et al., 2013). In some sites, they are highly abundant, e.g. comprising more than one quarter of the clones from some of the Colorado alpine samples (Schadt, 2003). In contrast, there is currently no evidence for diverse species related to NS1, nor for its occurrence outside of Interior Alaska. However, it does display temporal persistence and a preference for white spruce dominated habitats. Future studies might tease apart various substrata in soil (e.g. live vs dead roots, leaf litter, woody litter, fungal tissue, etc) in an effort to further delimit possible trophic niches. Similar efforts have led to pure culture isolation of members of major prokaryotic lineages previously known only from sequences, such as SAR11 (Rappé et al., 2002). Such efforts may be aided by the apparent temporal persistence of NS1. In tandem, efforts could be mounted to isolate and/or visualize the fungus. With respect to visualization, fluorescent *in situ* hybridization (FISH) has been used successfully in similar efforts with uncultured members of the Cryptomycota (Jones et al., 2011).

Though phylogenetically novel, and rare in samples studied to date, the ecological patterns emerging from our PCR surveys add evidence to the view that the NS1 sequences are derived from a real organism (see discussion in Glass et al., 2013). While the enormous depth and breadth of sequencing that can now be utilized for biodiversity studies holds tremendous promise for filling major gaps (Bik et al., 2012), these methods also bring new issues and constraints. Initial studies likely overinflated diversity estimates due to sequencing error and other artifacts (Ashelford et al., 2006; Kunin et al., 2010; Sogin et al., 2006). Once these issues became apparent, efforts have focused on avoiding overestimation (Engelbrektson et al., 2010; Nilsson et al., 2012). In addition, due to the size of recent datasets (millions to billions of reads), streamlined and simple strategies are often used for taxonomic identification. Two approaches are now commonplace: (1) removal of all OTUs below some abundance threshold (e.g. singletons or OTUs that account for <0.1 % of total reads – Nilsson et al., 2011; Tedersoo et al., 2010), and (2) removal of all OTUs that cannot be 'mapped' to reference taxa (e.g. "closed reference OTU picking" in QIIME; Caporaso et al., 2012). Application of either of these methods would have resulted in removal of NS1 sequences from our datasets. If the goal of a study is to uncover novel diversity or to estimate total species richness, these commonplace OTU removal practices become problematic. Hopefully, further advances in error correction, artifact detection and automated phylogenetic analysis will obviate these drastic OTU removal strategies.

Acknowledgments

We dedicate this paper to the memory of Daniel J Glass, a creative young scientist and kind soul, who passed away on September 22, 2013. This material is based upon work supported by the National Science Foundation through awards EF-0333308, ARC-0632332, DEB-0620579 and by the USDA Forest Service, Pacific Northwest Research Station, through agreement No. RJVA-PNW-01-JV-11261952-231. We thank Link Olson and Naoki Takebayashi for advice on analyses and earlier drafts of the manuscript and Teresa Hollingsworth for help with data on tree abundances.

REFERENCES

Altshul, S., Madden, T., Schäffer, A., Zhang, Z., Zheng, Z., Miller, W., Lipman, D., 1997. Gapped BLAST and PSI_BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25, 3389–3402.

Ashelford, K., Chuzhanova, N., Fry, J., Jones, A., Weightman, A., 2006. New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Applied and Environmental Microbiology* 72, 5734.

Bergemann, S.E., Miller, S.L., 2002. Size, distribution, and persistence of genets in local populations of the late-stage ectomycorrhizal basidiomycete, *Russula brevipes*. *New Phytologist* 156, 313–320.

Bik, H.M., Porazinska, D.L., Creer, S., Caporaso, J.G., Knight, R., Thomas, W.K., 2012. Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in Ecology and Evolution* 27, 233–243.

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., others, 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal* 6, 1621–1624.

Collado, J., Platas, G., Paulus, B., Bills, G.F., 2007. High-throughput culturing of fungi from plant litter by a dilution-to-extinction technique. *FEMS Microbiology Ecology* 60, 521–533.

Dickie, I.A., Xu, B., Koide, R.T., 2002. Vertical niche differentiation of ectomycorrhizal hyphae in soil as shown by T-RFLP analysis. *New Phytologist*, 527–535.

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.

Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H., Hugenholtz, P., 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *The ISME Journal* 4, 642–647.

Gardes, M., Bruns, T., 1993. ITS primers with enhanced specificity for Basidiomycetes – application to the identification of mycorrhizae and rusts. *Molecular Ecology* 2, 113–118.

Geml, J., Laursen, G.A., Herriott, I.C., McFarland, J.M., Booth, M.G., Lennon, N., Chad Nusbaum, H., Lee Taylor, D., 2010. Phylogenetic and ecological analyses of soil and sporocarp DNA sequences reveal high diversity and strong habitat partitioning in the boreal ectomycorrhizal genus Russula (Russulales; Basidiomycota). *New Phytologist* 187, 494–507.

Geml, J., Laursen, G.A., Timling, I., McFarland, J.W., Booth, M.G., Lennon, N.J., Nusbaum, C., Taylor, D.L., 2009. Molecular phylogenetic biodiversity assessment of arctic and boreal ectomycorrhizal *Lactarius* Pers. (Russulales; Basidiomycota) in Alaska, based on soil and sporocarp DNA. *Molecular Ecology* 18, 2213–2227.

Glass, D.J., Takebayashi, N., Olson, L.E., Taylor, D.L., 2013. Evaluation of the authenticity of a highly novel environmental sequence from boreal forest soil using ribosomal RNA secondary structure modeling. *Molecular Phylogenetics Evolution* 67, 234–245.

Izzo, A., Agbowo, J., Bruns, T., 2005. Detection of plot-level changes in ectomycorrhizal communities across years in an old-growth mixed-conifer forest. *New Phytologist*, 619–629.

Jones, M.D., Forn, I., Gadelha, C., Egan, M.J., Bass, D., Massana, R., Richards, T.A., 2011. Discovery of novel intermediate forms redefines the fungal tree of life. *Nature* 474, 200–203.

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

Lilleskov, E., Bruns, T., Horton, T., Taylor, D., Grogan, P., 2004. Detection of forest stand-level spatial structure in ectomycorrhizal fungal communities. *Fems Microbiology Ecology* 49, 319–332.

Lindahl, B.D., Ihrmark, K., Boberg, J., Trumbore, S.E., Högberg, P., Stenlid, J., Finlay, R.D., 2006. Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. *New Phytologist* 173, 611–620.

Nilsson, R., Abarenkov, K., Veldre, V., Nylinder, S., De Witt, P., Brosché, S., Alfredsson, J.F., Ryberg, M., Kristiansson, E., 2010. An open source chimera checker for the fungal ITS region. *Molecular Ecology Resources* 10, 1076–1081.

Nilsson, R.H., Tedersoo, L., Abarenkov, K., Ryberg, M., Kristiansson, E., Hartmann, M., Schoch, C.L., Nylander, J.A.A., Bergsten, J., Porter, T.M., others, 2012. Five simple guidelines for establishing basic authenticity and reliability of newly generated fungal ITS sequences. *MycoKeys* 4, 37–63.

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

O'Brien, H., Parrent, J., Jackson, J., Moncalvo, J., Vilgalys, R., 2005. Fungal community analysis by large-scale sequencing of environmental samples. *Applied and Environmental Microbiology* 71, 5544–5550.

Pace, N.R., Stahl, D.A., Lane, D.J., Olsen, G.J., 1985. Analyzing natural microbial populations by rRNA sequences. *ASM American Society for Microbiology News* 51, 4–12.

Porter, T., Schadt, C., Rizvi, L., Martin, A., Schmidt, S., Scott-Denton, L., Vilgalys, R., Moncalvo, J., 2008. Widespread occurrence and phylogenetic placement of a soil clone group adds a prominent new branch to the fungal tree of life. *Molecular Phylogenetics and Evolution* 46, 635–644.

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Rappé, M.S., Connon, S.A., Vergin, K.L., Giovannoni, S.J., 2002. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* 418, 630–633.

Rosling, A., Cox, F., Cruz-Martinez, K., Ihrmark, K., Grelet, G.A., Lindahl, B.D., Menkis, A., James, T.Y., 2011. Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. *Science* 333, 876–879.

Rosling, A., Timling, I., Taylor, D.L., 2013. Archaeorhizomycetes: patterns of distribution and abundance in soil. In: Genomics of Soil-and Plant-Associated Fungi. Springer, pp. 333–349.

Schadt, C., 2003. Seasonal dynamics of previously unknown fungal lineages in tundra soils. *Science* 301, 1359–1361.

Schechter, S., Bruns, T., 2008. Serpentine and non-serpentine ecotypes of *Collinsia sparsiflora* associate with distinct arbuscular mychorrizal fungal assemblages. *Molecular Ecology* 17, 3198–3210.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.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 Sciences of the United States of America* 103, 12115.

Sokal, R.R., Oden, N.L., 1978. Spatial autocorrelation in biology: 1. methodology. *Biological Journal of the Linnean Society* 10, 199–228.

Stingl, U., Tripp, H.J., Giovannoni, S.J., 2007. Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. *The ISME Journal* 1, 361–371.

Taylor, D.L., Bruns, T.D., 1999. Community structure of ectomycorrhizal fungi in a *Pinus muricata* forest: minimal overlap between the mature forest and resistant propagule communities. *Molecular Ecology* 8, 1837–1850.
Taylor, D.L., Herriott, I.C., Long, J., O'Neill, K., 2007. TOPO TA is A-OK: a test of phylogenetic bias in fungal environmental clone library construction. *Environmental Microbiology* 9, 1329–1334.
Taylor, D.L., Herriott, I.C., Stone, K.E., McFarland, J.W., Booth, M.G., Leigh, M.B., 2010. Structure and resilience of fungal communities in Alaskan boreal forest soils. *Canadian Journal of Forest Research* 40, 1288–1301.
Taylor, D.L., Hollingsworth, T.N., McFarland, J., Lennon, N.J., Nusbaum, C., Ruess, R.W., 2014. A first comprehensive census of fungi in soil reveals both hyperdiversity and fine-scale niche partitioning. *Ecological Monographs* 84, 3–20.
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.
Toljander, J., Eberhardt, U., Toljander, Y., Paul, L., Taylor, A., 2006. Species composition of an ectomycorrhizal fungal community along a local nutrient gradient in a boreal forest. *New Phytologist* 170, 873–883.