

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>

available at www.sciencedirect.com

ScienceDirect

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

Short Communication

Analyses of ITS and LSU gene regions provide congruent results on fungal community responses

Shawn P. BROWN^{a,*}, Anne R. RIGDON-HUSS^b, Ari JUMPPONEN^a^aDivision of Biology, Kansas State University, USA^bDepartment of Grain Science and Industry, Kansas State University, USA

ARTICLE INFO

Article history:

Received 6 November 2013

Revision received 27 January 2014

Accepted 20 February 2014

Available online 26 March 2014

Corresponding editors:

Håvard Kauserud and Marie Louise Davey

Keywords:

Ecological analysis

ITS

LSU

454-sequencing

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.

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

Introduction

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 data sets, however, this is often

* Corresponding author. Division of Biology 116 Ackert Hall Kansas State University Manhattan, KS 66506, USA. Tel.: +1 785 532 3934; fax: +1 785 532 6653.

E-mail address: spbrown1@ksu.edu (S.P. Brown).

1754-5048/\$ – see front matter © 2014 Elsevier Ltd and The British Mycological Society. All rights reserved.

<http://dx.doi.org/10.1016/j.funeco.2014.02.002>

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; Porrás-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 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 on 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 6-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 (S_{obs}), Complement of Simpson's Diversity (1-D) and Simpson's Evenness (E_D) 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 \times treatment interactions on an experiment-wide basis across the ITS and LSU data sets (Table S1).

On an experimental unit basis, paired t-tests suggested significant ($P > 0.05$) but proportionally minor differences in diversity responses, in 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 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' et al. (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 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 ($P > 0.05$) 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

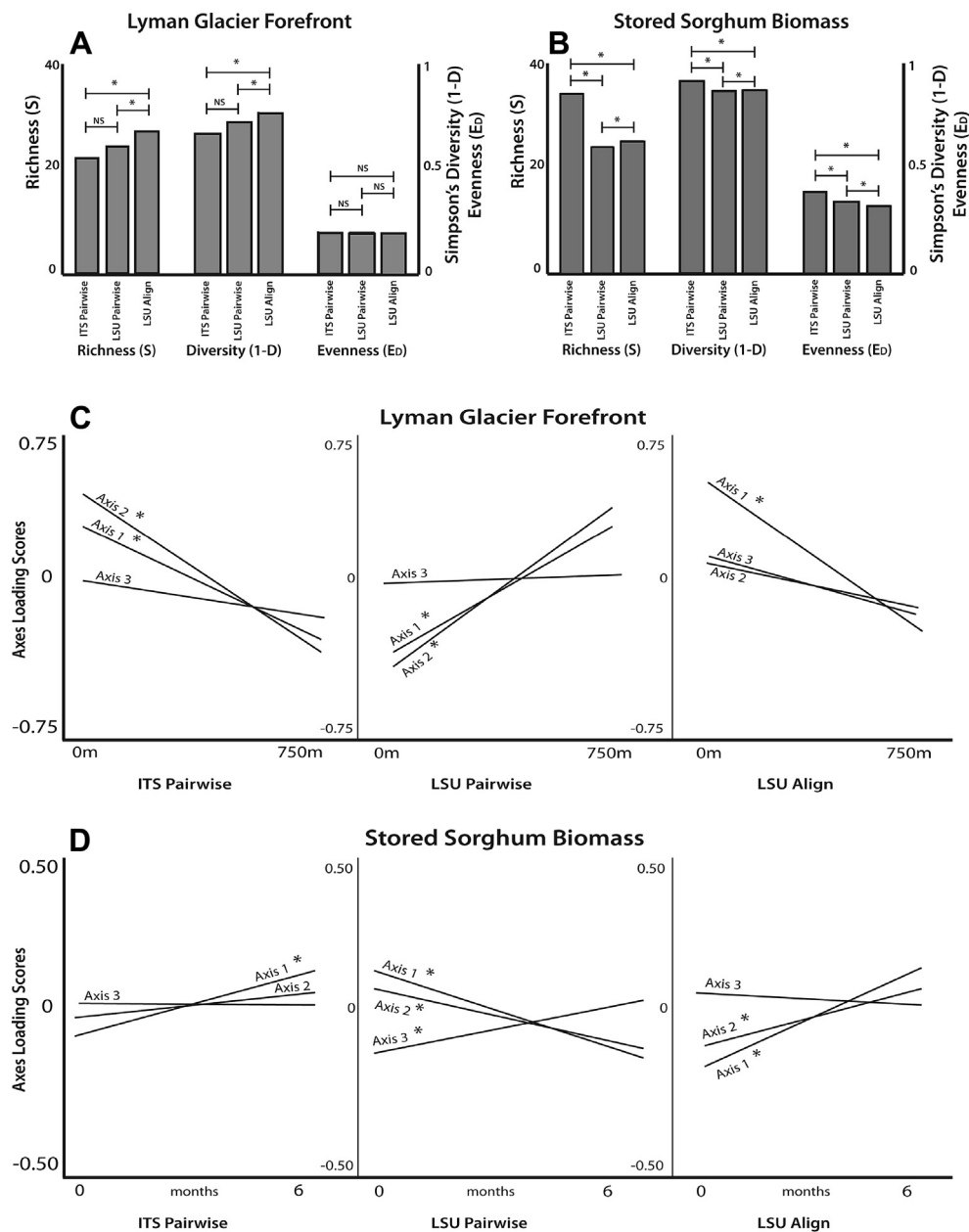


Fig 1 – Paired t-tests of richness, diversity (1-D) and evenness (E_D) estimators derived from ITS, pairwise aligned LSU, and reference aligned LSU data in the Lyman Glacier Forefront (A) and Stored Sorghum Biomass (B) experiments. Bars indicate mean values across all experimental units and asterisks (*) 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 Table S1.

least one axis (Fig 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 1D) as well as distinguishing one of the four storage methods from the others (Table S1).

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 S1), 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 of fungal community ecology.

Acknowledgments

This research was partly supported by US National Science Foundation Grant DEB-0516456. The authors' would like to express our great appreciation to Rauni Strömmer, Chad Fox and Laura Fox for assistance with sample collection. Shawn Brown is supported by the National Science Foundation GK-12 fellowship (NSF DGE-0841414, PI Ferguson). Anne Rigdon-Huss is supported by National Science Foundation IGERT (NSF 0903701 to KSU) I-STAR Bioenergy fellowship and Kansas Alliance for BioRefining and BioEnergy (KABB GAGS92152 to KSU – Grain Science and Industry).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.funeco.2014.02.002>.

REFERENCES

- Amend, A.S., Seifert, K.A., Samson, R., Bruns, T.D., 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, S.P., Jumpponen, A., 2014. Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. *Molecular Ecology* 23, 481–197.
- Gihring, T.M., Green, A.J., Schadt, C.W., 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, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., et al., 2006. Reconstructing the early evolution of fungi using a six-gene phylogeny. *Nature* 443, 818–822.
- Kerekes, J., Kaspari, M., Stevenson, B., Nilsson, R.H., Hartmann, M., Amend, A., Bruns, T.D., 2013. Nutrient enrichment increased species richness of leaf litter fungal assemblages in a tropical forest. *Molecular Ecology* 22, 2827–2838.
- Lothamer, K., Brown, S.P., Mattox, J.D., Jumpponen, A., 2013. Comparison of root-associated communities of native and non-native ectomycorrhizal hosts in an urban landscape. *Mycorrhiza*. <http://dx.doi.org/10.1007/s00572-013-0539-2>.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, A., Bader, J.S., et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380.
- Nilsson, R.H., Tedersoo, L., Abarenkov, K., Ryberg, M., Kristiansson, E., Hartmann, M., Schoch, C.L., Mylander, J.A.A., Bergsten, J., Porter, T.M., Jumpponen, A., Vaishampayan, P., Ovaskainen, O., Hallenberg, N., Bengtsson-Palme, J., Eriksson, K.M., Larsson, K.H., Larsson, E., Kõljalg, U., 2012. Five simple guidelines for establishing basic authenticity and reliability of newly generated fungal ITS sequences. *MycKeys* 4, 37–63.
- Peay, K.G., Kennedy, P.G., Bruns, T.D., 2008. Fungal community ecology: a hybrid beast with a molecular master. *BioScience* 58, 799–810.
- 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.
- 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 mother: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75, 7537–7541.
- Schloss, P.D., Gevers, D., Westcott, S.L., 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLOS One* 6, e27310.
- 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, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* 73, 5261–5267.
- Weber, C.F., Vilgalys, R., Kuske, C.R., 2013. Changes in fungal community composition in response to elevated atmospheric CO₂ and nitrogen fertilization varies with soil horizon. *Frontiers in Microbiology* 4, 78. <http://dx.doi.org/10.3389/fmicb.2013.00078>.