Mycorrhiza (2009) 19:283–285
DOI 10.1007/s00572-008-0210-5

NEWSCORNER

FESIN workshops at ESA—the mycelial network grows

Thomas R. Horton · A. Elizabeth Arnold · Thomas D. Bruns

Received: 24 September 2008 / Accepted: 22 October 2008 / Published online: 11 November 2008

Keywords FESIN · ESA · Fungi · Environmental sampling

The first large public meeting of the Fungal Environmental Sampling and Informatics Network (FESIN: http://www.bio.utk.edu/fesin/title.htm) took place at the Ecological Society of America (ESA) meetings in Milwaukee this summer. Last year's meetings at the Mycological Society of America (MSA) and ESA meetings were mostly informational, highlighting future activities of this NSF-sponsored research coordination network. A joint meeting between FESIN and the User-friendly Nordic ITS Ectomycorrhizal group (UNITE) also occurred last year to spearhead research proposals, and discuss pros and cons of current methods for analyzing fungal communities (Bruns et al. 2008). This year's meeting included 2 days of preconference workshops that are summarized below. In addition, Jeri Parrent and Betsy Arnold organized a FESIN-sponsored oral session entitled 'Frontiers in the Ecology of Plant–Fungal Interactions,' held during the main ESA meeting. This session brought together mycologists and ecologists in a standing-room-only series of eleven talks (see authors and abstracts at http://eco.confex.com/eco/2008/techprogram/S2894.HTM).

The 2008 FESIN meeting was a big success judged by level of attendance, composition of attendees, and direct feedback that the organizers received. Ninety people attended the workshops; of these, 58% were either graduate students or postdocs. A show of hands revealed that over half of those attending had never been to an ESA meeting before; an equal number had never attended a MSA meeting. This was great news given that one of the goals of FESIN is to increase connections between ecologists and mycologists and to bring new people and ideas into the growing field of fungal ecology.

Jeri Parrent, Kitty Gehring, Karen Hughes, Betsy Arnold, and Tom Bruns organized the preconference workshops. The goal of Saturday's workshop was to introduce appropriate molecular techniques to ecologists who are just beginning to work on fungi, and to introduce new methodological advances to those already familiar with the basics. The first set of talks by Mary Berbee, Jeri Parrent, Ian Dickie, and Shannon Schechter introduced Fungi as a group, the suite of major molecular techniques used to study them, and specialized problems associated with studying arbuscular mycorrhizal fungi. Then, Ari Jumpponen, Cedar Hesse, and Jason Stajich described high-throughput sequencing for studying fungal communities, and provided an up-to-date perspective on bioinformatics tools. Brendan Bohannan and Ian Alexander addressed the problems of sampling hyper-diverse communities that exhibit high spatial–temporal variation, and Chris Schadt discussed the use of functional gene arrays to study the "process" level of communities.

Sunday's workshop focused on the strengths, weaknesses, and opportunities offered by public sequence

T. R. Horton (✉)
SUNY—College of Environmental Science and Forestry,
Syracuse, NY, USA
e-mail: trhorton@esf.edu

A. E. Arnold
University of Arizona,
Tucson, AZ, USA

T. D. Bruns
University of California at Berkeley,
Berkeley, CA, USA

databases, which are an indispensable tool in the study of fungal communities. Rytas Vilgalys, Betsy Arnold, and Tom Bruns discussed deficiencies of current data, the need for more standardized metadata, and ideas for improved curation and classification of sequences. Next, David Hibbett and Jim Cole presented advances in automated classification systems and other analytical tools available on the ribosomal database and elsewhere. Finally, Noah Fierer and Kathleen Treseder presented new estimates of fungal diversity and new ways to assess ecological function, respectively.

Sunday's presentations lead to a general discussion of ways that FESIN could facilitate improvements in public databases, and committees were assembled and charged with working on specific aspects of these problems. Some of the future directions of FESIN are highlighted below.

Because GenBank does not allow third-party annotation, sequences remain listed with incorrect or incomplete taxonomic information unless authors return to update their files (Bidartondo et al. 2008). As researchers use GenBank for diagnostic purposes, problems are perpetuated—and become more difficult to trace. Recent discussions with FESIN and David Lipman of NCBI (National Center for Biotechnology Information) have led to the idea of a curated collection of reference sequences to be used specifically for fungal identification. While the exact form needs to be worked out, a solution discussed at the workshop is to generate a third-party database with an automated classification system similar to the Ribosomal Database used by bacteriologists (http://rdp.cme.msu.edu/; Cole et al. 2007). This would be initiated with a core set of sequences from GenBank that would be vetted for sequence quality and taxonomic accuracy. Third-party annotation in this database would be encouraged, with appropriate author notification, peer-review, and full transparency. Once built, this database would grow with high-quality sequences and currently accepted taxonomic concepts. As a first step, a fungal sequence "annotation jamboree" was proposed to bring experts together to compile and vet the first batch of sequences for the database. Phylogenetic treatments for various genera would be a good starting point, as would the inclusion of sequences from type and vouchered specimens. Initially, efforts will likely focus on the ITS region given its utility for diagnostic purposes in many Fungi and the fact that GenBank and EMBL are already populated with over 90,000 fungal ITS sequences. Other regions will be incorporated as the database develops. This third-party database will include a BLAST algorithm and be a more powerful diagnostic tool than GenBank.

Anyone who has used GenBank to identify an unknown sequence recognizes that the output is ambiguous with respect to applying a species name. This happens when the sequenced region has low resolution for resolving species in a fungal lineage, the database is not populated by identified sequences from a particular genus, or the sample is collected from a new research area. Still, taxonomic information is there and other researches may recover nearly identical sequences in their own studies. A solution discussed at the workshop is to develop a position paper on rules for applying names to sequences. The core ideas include automatically classifying sequences by phylogenetic methods and assigning unique numbers to terminal "species-sized" clades. In some cases, these clades will correspond well with existing species concepts, whereas in other cases, they may correspond to a group of closely related species, or to a genetically distinct group within a recognized species. In many cases, correspondence between species and clades will initially not be known. Nevertheless, assigning unique clade numbers would provide a common vocabulary so that researchers who encounter these organisms have a way to access biologically relevant data on them from prior studies.

Under this scenario, species names would be attached to clades when sufficient information warranted it, but those names would be treated as hypotheses that could be annotated as new results become available. For example, a new sequence generated from root tips might be identified as a *Russula* in "ITS clade 142." This clade may have no sequences that are associated with a named species, and so the new sequence would simply be classified as *Russula* 142. This is not dramatically different from what is used in current studies; the difference is that future studies would use a common name for sequences in this clade. If identified sequences in that clade were all *R. silvicola* then that name could be applied as *Russula* 142 (*silvicola*). Putting the name in parentheses indicates the current hypothesis and is analogous to using *cf.* when naming a voucher collection to indicate it is comparable to a named species. This name could change as new information becomes available, but the number would remain. The details of this system and exact form of the automated classifier were not resolved, and issues such as the degree to which single-locus phylogenies can be reliable and insightful in this regard are yet to be agreed upon. The urgency of producing a workable solution, however, was clear. All workshop participants realized that high-throughput environmental sequencing with pyrosequencing methodology will soon swell GenBank with hundreds of thousands or millions of unidentified sequences.

The discussion on automated classification tied directly into a discussion of "metadata"—the information associated with a sequence that provides context such as where, when and from what it was derived. Such data are analogous to data presented on labels for well-documented herbarium specimens, information mycologists have found useful for hundreds of years. Recent work by Ryberg et al. (2008), for example, explored the ecology and global distribution

patterns of the ectomycorrhizal genus *Inocybe* by "mining" metadata from both identified and environmental *Inocybe* sequences deposited in GenBank. This would not have been possible if the authors of the submitted sequences had not included at least some ecological data with their submissions. However, gleaning metadata from GenBank is currently a laborious task because there is no requirement to enter many common metadata elements. This point led to discussion regarding a standard set and format for data such as GPS coordinates, host species, collection date, substrate, etc. While the details are being worked out to keep the submission of such data simple, straightforward, and consistent, we encourage researchers to submit all the basic metadata elements that provide a context for their submissions.

Over the next 3 years FESIN will help coordinate development of infrastructure for the field of fungal ecology. Its meetings are open to all interested participants and will alternate between ESA and MSA meetings. Next year's meeting will be in July at MSA (Snowbird, Utah—joint meeting with the Botanical Society of America). There will likely be a hands-on workshop on ecological statistics for mycologists at this meeting, and other workshop and symposium ideas for this and future meetings are in the planning stages. To be kept informed of future developments and learn more about FESIN activities and products, sign on as a new member at the FESIN website.

References

Bidartondo MI et al (2008) Preserving accuracy in GenBank. Science 319:1616 doi:10.1126/science.319.5870.1616a

Bruns TD, Arnold AE, Hughes KW (2008) Fungal networks made of humans: UNITE, FESIN, and frontiers in fungal ecology. New Phytol 177:586–588

Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM, Bandela AM, Cardenas E, Garrity GM, Tiedje JM (2007) The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res 35:D169–D172 doi:10.1093/nar/gkl889

Ryberg M, Nilsson RH, Kristiansson E, Topel M, Jacobsson S, Larsson E (2008) Mining metadata from unidentified ITS sequences in GenBank: a case study in *Inocybe* (Basidiomycota). BMC Evol Biol 8:50 doi:10.1186/1471-2148-8-50