Method or madness: does OTU delineation bias our perceptions of fungal ecology?

<table>
<thead>
<tr>
<th>Journal:</th>
<th>New Phytologist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>NPH-C-2014-17263</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>C - Commissioned Material - Commentary</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>13-Mar-2014</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Sikes, Benjamin; University of Kansas, Department of Ecology and Evolutionary Biology; Until Aug 2013: Lincoln University, Bio-Protection Research Centre Powell, Jeff; University of Western Sydney, Hawkesbury Institute for the Environment</td>
</tr>
<tr>
<td>Key Words:</td>
<td>OTU delineation, arbuscular mycorrhizal fungi, high-throughput sequencing, species concepts, niche conservatism, alpha diversity, beta diversity</td>
</tr>
</tbody>
</table>
Method or madness: does OTU delineation bias our perceptions of fungal ecology?

Jeff R. Powell¹ and Benjamin A. Sikes²

¹ Hawkesbury Institute for the Environment, University of Western Sydney, Sydney, Australia
² Department of Ecology and Evolutionary Biology and Kansas Biological Survey, University of Kansas, Lawrence, Kansas 66047

Keywords: OTU delineation, arbuscular mycorrhizal fungi, high-throughput sequencing, species concepts, niche conservatism, alpha diversity, beta diversity

Arbuscular mycorrhizal (AM) fungi are among the most widespread and studied terrestrial mutualisms (Smith & Read, 2008). As molecular tools have become widely available, the application of these tools to characterize communities of AM fungi has met with significant obstacles and complications. The multigenomic nature of individual spores has been hotly debated (Hijri & Sanders, 2005) and caused delays in sequencing the entire genome of the typological species, *Rhizophagus irregularis* (Tisserant et al., 2013; Lin et al., 2014). The low number of “known” AM fungal species is largely derived from morphology based on spore traits from culturable isolates, but molecular techniques are starting to reveal cryptic and unculturable diversity necessitating the creation of “virtual taxa”, currently without morphological analogues (Öpik et al., 2006). In the absence of an operational species concept, delineating "species" appropriately is essential for mycorrhizal ecologists seeking to compare the diversity of AM fungi both in nature and experiments, particularly for studies using high-throughput DNA sequencing. However, not all approaches for delineating environmental DNA sequences into operational taxonomic units (OTUs) are equivalent and each can result in substantially different descriptions of the fungal communities under study. In this issue of New Phytologist, Lekberg et al. (2014) compare two methods with high-throughput AM fungal DNA sequences and ask if the methodology used impacts our inferences about the processes driving AM fungal community structure.

The first, widely used method applies a universal cutoff value to produce OTUs containing members sharing a high level of sequence similarity (usually 97 %). This method originally derives from studies of DNA-DNA reassociation and 16S sequences for a number of isolated bacteria (Stackebrandt & Goebel, 1994). It assumes 1) similar rates of evolution across lineages in the taxon under consideration (clearly violated for fungi; Nilsson et al., 2008) and 2) the gene region used largely tracks this divergence rate, especially in known species previously described using morphological characters. The second method uses evolutionary based metrics that define OTUs as individual monophyletic clades with significant bootstrap support (Lekberg et
al., 2012), resulting in clades with a range of intraclade sequence similarities. This method is analogous to the virtual taxon approach developed by Öpik et al. (2008). Lekberg et al. find that the method employing a universal cutoff results in substantially greater numbers of OTUs (alpha diversity) than the method using the evolutionary metric, with the greatest disparity at the local scale (Figure 1A). However, both methods produced OTUs that exhibited consistent ecological sorting among different spatial scales and environmental gradients in their datasets (beta diversity). The range of spatial extents over which this pattern is observed is substantial, from a scale of metres to hundreds of kilometres, and while there may be variation in the differences between alpha and beta diversity estimates at these different scales (Figure 1A and B, with the largest differences in alpha and beta diversity being observed for the 'local' and 'regional' datasets, respectively), the statistical significance of different ecological variables did not appear to be affected by the OTU delineation method.

These results largely agree with a previous evolutionary model-based approach to OTU delineation (Powell et al., 2011), and together they suggest that ecologically significant OTU delineation may occur using a level of sequence similarity that is lower than commonly used universal cutoffs. These findings support AM fungal community assembly as a product of phylogenetic niche conservatism of functional traits that affect fitness along certain environmental gradients. Previous work has focused on deeply rooted clades of AM fungi that mainly differ in their capacity to produce extensive networks of intraradical or extramatrical hyphae, which can influence the outcome of interspecific interactions among closely related species (Maherali & Klironomos, 2007). Although we lack a full understanding of the autecology of AM fungi, species distribution modeling suggests some coherence in species responses within families to long gradients in edaphic properties (Veresoglou et al., 2013). Although encompassing multiple spatial scales, each of the datasets analysed by Lekberg et al. and Powell et al. encompass significant shifts in land management and vegetation types, as well as severe disturbance, so these conclusions are primarily being drawn when considering long gradients in the environment. In this case, and for taxa exhibiting this degree of phylogenetic niche conservatism, a clustering approach that results in fewer, monophyletic OTUs may be preferred to improve the power to detect correlations among environmental variables and the distribution of OTUs across samples. Although not statistically significant, this benefit is likely indicated by the improved ability of ‘MCA’ OTUs to explain environmental variation at the ‘regional’ spatial scale compared to the universal cutoff OTUs (Figure 1B). An additional benefit is that there is less opportunity for sequence variation (whether real variation within populations or occurring due to errors during sequencing) to generate noise in OTU-sample tables.

But could small trait differences among closely related taxa be functionally more important along short environmental gradients? While phylogenetic niche conservatism might be the
norm when comparing quite divergent environments, genetic variation within species of AM fungi might result in differential fitness of populations across small shifts in the environment. Delineating OTUs that detected these AM fungal ecotypes may require higher resolution than monophyletic taxa or the universal cutoff currently used. The supervised learning approach employed by Lekberg et al. (Figure S5 of their letter) may actually provide some indirect support for this; when assigning AM fungal assemblage to plant communities, Lekberg et al. observed that the 90% and 99% universal cutoffs produced similar error rates. However, the lowest error rate was observed using a universal cutoff of 98%, suggesting even better correspondence between delineated OTUs and this axis in the environment. While substantial intraspecific functional variation in some isolates of AM fungi has been observed (e.g., Munkvold et al., 2004), our understanding of the incidence of ecotypes in the natural environment lags behind that of other groups of fungi (e.g., Taylor et al., 2014).

Other community assembly mechanisms, including dispersal and local extinction may also overwhelm environmental filtering and biotic interactions within short environmental gradients. For example, Caruso et al. (2012) also used the 'sites' data analysed here by Lekberg et al. and observed differing patterns of community assembly under the different treatments included in the study. Colonisation of Plantago lanceolata roots, in the presence or absence of disturbance, resulted in highly divergent assemblages of AM fungi within roots, indicative of environmental filtering along an unidentified gradient or strong biotic interactions among colonisers. However, AM fungal assemblages associated with roots sampled from the surrounding plant species converged upon a common composition, indicating a deterministic assembly process, even though the identities of the host species likely changed among samples. These patterns were observed following the MCA approach for OTU delineation, so perhaps alternative approaches to OTU delineation could result in different, and possibly more intuitive, patterns. While researchers are often constrained to demonstrating statistically significant shifts in communities as a prerequisite to publish their work, there is a need to focus on AM fungal community assembly along small environmental gradients to determine at what level of environmental variation scientists need to concern themselves with comparing delineation methods (Figure 2). Alternatively, empirical demonstrations that ecological inferences are robust to OTU delineation approaches along small environmental gradients would help to confirm the claims by Lekberg et al. that the method chosen is relatively unimportant.

The final point made by Lekberg et al. (2014) is an important one: we need to keep in mind that the end goal of efforts to mass sequence DNA from the environment is the advancement of our understanding of AM fungal ecology and biogeography, and those of microorganisms in general. While there are very real issues to be dealt with regarding, for example, error correction and OTU clustering approaches, too much focus on what is the correct way to do this may impede progress in achieving this goal and evaluating the sensitivity of our inferences to...
employing multiple, imperfect approaches (including approaches that are not based on high-throughput sequencing) may provide even more insight into these processes.


Figure Legends:

Figure 1: Differences in alpha and beta diversity at different spatial scales (site, local and regional) based on data in Lekberg et al. 2014. Panel 1A shows the percentage increase in OTU numbers based on delineation using the 97% cutoff method relative to the MCA evolutionary method. Panel 1B shows the differences in variance explained at the different scales in both the ANOSIM and the BEST model.

Figure 2: Hypothesized relationships between OTU delineation approach and environmental variation (low, medium and high) on estimates of alpha and beta diversity. The box to the right of the figure indicates the environmental gradient over which Lekberg et al. (2014) and Powell et al. (2011) have likely evaluated this relationship. The y-axis (not-scaled) is similar to figure 1 indicating the predicted differences in the metrics based on OTU delineation method alone.
Fig. 1

- **A**: % increase in alpha diversity (97% relative to MCA)
- **B**: Absolute change in beta diversity (97% relative to MCA)

**Spatial Scale**: Site, Local, Regional

**Legend**:
- **R² from ANOSIM**
- **R² from BEST**

216x259mm (300 x 300 DPI)