

Fungal community analysis using environmental genomics

O'Brien, HE(1), Jackson, JA(1), Johnson, JE(2), Parrent, JL(1), Moncalvo, J-M(1), Vilgalys, R(1)

(1) Dept. of Biology, Duke University, Durham NC 27708 USA (2) Dept. of Biological Sciences, Cameron University, Lawson OK 23505 USA

Introduction

Fungal diversity is thought to greatly exceed the numbers of currently described species (Hawksworth 2001). While much of this diversity is likely to be found in under-explored habitats in the tropics, diversity in well-characterized, temperate habitats may also be underestimated. The use of culture-independent methods involving sequencing of ribosomal DNA fragments directly amplified from the substrate has revolutionized our knowledge of bacterial diversity over the last ten years (Hugenholz et al. 1998). Recent findings suggest that similar cryptic diversity may be found in fungi using this same approach (Vandenkoornhuyse et al. 2002). We used this sequence-based approach to examine the diversity of fungi and other microbial Eukaryotes inhabiting the soil and litter of a mixed deciduous forest in the Southeastern United States (The Duke Forest, Durham NC). Sampling was focused on the internal transcribed spacer (ITS) region of the ribosomal DNA array, which permits fine-scale taxonomic resolution and which allowed us to use the extensive databases available for this region to identify environmental isolates to the genus or even the species level in most cases. We also sequenced Eukaryotic (18S) and Prokaryotic (16S) small subunit rDNA to verify the results from ITS. In total, 832 ITS sequences and 300 SSU sequences were generated as part of this study.

The goals of this project were to:

1. Estimate the total diversity of soil and litter fungi in our samples
2. To compare the diversity and composition of fungal communities in different forest types and different soil strata
3. To determine the relative representation of different taxonomic groups in the soil Biota

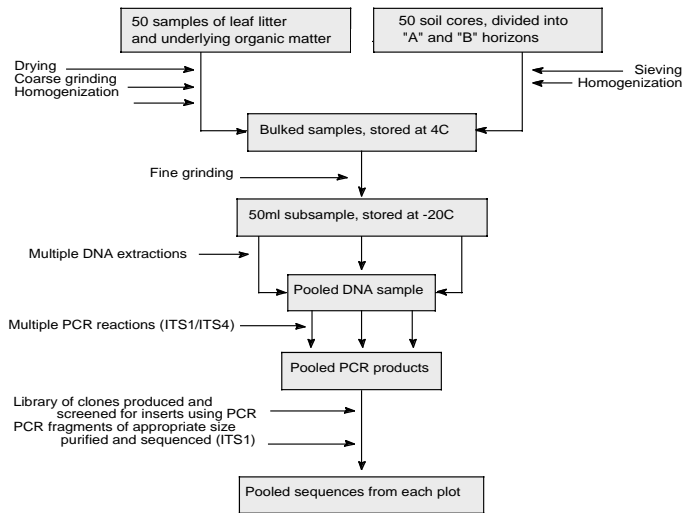
Sampling

Sampling was conducted in two permanent sample plots in the Duke Forest. One plot was located in an alluvial hardwood stand (578), the other in an upland site dominated by loblolly pine (39/40). DNA was extracted from four pooled soil samples from each site: leaf litter (L), the underlying organic layer (O), topsoil (A), and mineral soil (B).

Analysis

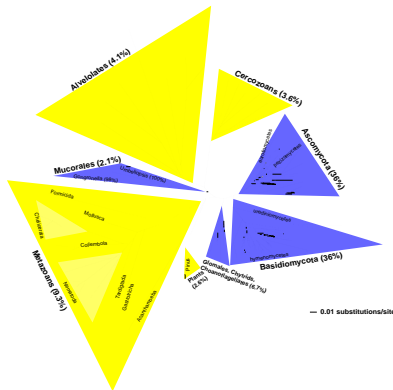
Sequences from all plots were pooled and assembled at 98% minimum match using Sequencher (GeneCodes Corp.). Consensus sequences from each contig compared to genbank and unpublished data using batch BLAST searches. Chimeric sequences identified by comparing the BLAST results for ITS1 and ITS2. Sequences with >95% identity in ITS (BLAST match greater than 200 bp) were assigned to major taxonomic clades. Sequences with >98% ITS identity were considered identifiable to genus level. Sequences with >80% ITS identity or >95% 5.8S identity were considered identifiable at the level of phylum. Samples with close matches to mycorrhizal taxa or to mycorrhizal isolates were identified as mycorrhizal. Diversity was estimated using the program EstimateS (R. Colwell) with 50 randomizations.

Methods



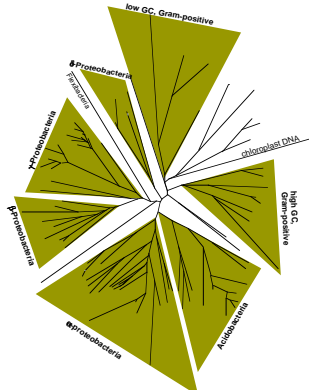
1. Relative diversity of major taxonomic groups

18S Diversity - Eukarya



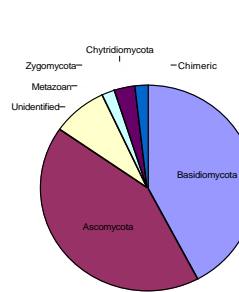
NJ tree based on HKY distances for 200 18S rDNA sequences

16S diversity - Eubacteria

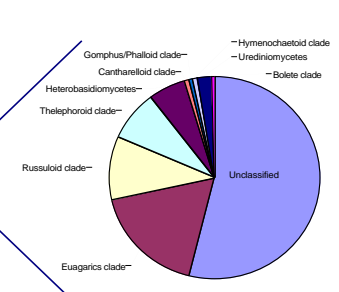


NJ tree based on HKY85 distances for 100 16S rDNA sequences

ITS diversity -Eukarya

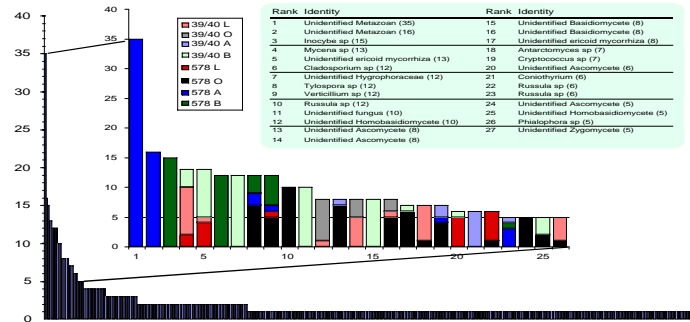


ITS diversity -Basidiomycota

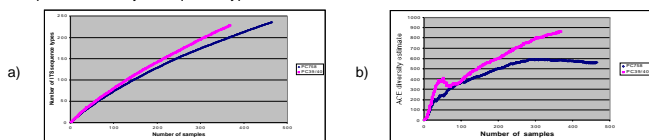


Pie charts representing the proportion of (a) eukaryotic phyla (441 sequence types) and (b) major clades of basidiomycetes (187 sequence types) based on ITS rDNA

2. Diversity estimates



Rank abundance graph showing the number of hits on each ITS sequence type. Inset: distribution across samples and identify of sequence types recovered five or more times



a) species accumulation curve showing number of ITS sequence types found as a function of number of sequences sampled and (b) graph showing the effect of sampling on the ACE diversity estimator

3. Comparison of plots and soil strata

Sample	# of sequences	# of sequence types	Simpsons diversity	Shannon diversity
39/40 L	96	70	68.06	4.03
39/40 O	73	54	65.70	3.82
39/40 A	81	60	90.00	3.96
39/40 B	118	58	33.84	3.68
39/40 pooled	368	236	162.43	5.14
578 L	52	26	28.83	3.07
578 O	206	135	148.14	4.71
578 A	113	46	8.56	2.59
578 B	93	48	21.39	3.43
578 pooled	464	227	84.31	5.00

Data summary

- High diversity: 441 ITS sequence types, 301 unique sequences.
- High abundance of fungi in soil: -97% of identifiable ITS sequences are fungal. -10% of sequences cannot be identified to kingdom.
- Roughly equal proportion of Basidiomycete and Ascomycete sequences.
- Large proportion of ectomycorrhizal taxa: 27% of identifiable Ascomycete sequences and 55% of Basidiomycetes.
- Most identifiable Basidiomycete sequences fall into Euagarics clade (38%), Russuloid clade (22%) and Theleporoid clade (17%).

Challenges

- Many chimeric sequences produced by PCR recombination across the highly conserved 5.8S region. Chimeras were identified by comparing BLAST results for ITS1 and ITS2. PCR recombination can be minimized by reducing the number of PCR cycles and increasing extension time.
- PCR bottlenecks: 51 of 113 sequences from plot 578A formed two closely related ITS sequence types that appear to be metazoan in origin. This problem can be minimized by constructing clone libraries from multiple PCR reactions and from multiple DNA extractions.

Conclusions

- Sequencing of environmental ITS sequences is a useful approach for investigating the diversity of Eukaryotic microorganisms.
- The extremely high diversity of soil communities presents a challenge for studying microbial ecology using this method.
- Improvements in taxonomic sequence databases and bioinformatics tools for analyzing high-throughput sequence data will enable better characterization of microbial communities.

References

- Hawksworth, D. L. 2001. *Mycological Research*, 105: 1422-1432.
- Hugenholz, P., Goebel, B. M., and Pace, N. R. 1998. *Journal of Bacteriology*, 180: 4765-4774.
- Vandenkoornhuyse, P., Baldauf, S. Let, C. al. 2002. *Science*, 295: 2051

Research funded by:

National Science Foundation Grant #NSF-MCB-00-84207