Fingerprinting the fungal community

Nabla Kennedy and Nicholas Clipson

Department of Industrial Microbiology, University College Dublin, Belfield, Dublin 4, Ireland.

Nabla.kennedy@ucd.ie +35317161574

Nicholas.clipson@ucd.ie +35317161365

Fax +35317161183

Running title: Fingerprinting the fungal community.

Corresponding Author: Nabla Kennedy

Department of Industrial Microbiology, University College Dublin, Belfield, Dublin 4, Ireland

Phone +35317161574

Fax +35317161183

Email: nabla.kennedy@ucd.ie

Abstract

Fungi can be found in almost any environment, and play important roles in ecosystem processes such as nutrient cycling and degradation. Despite their importance, the vast majority of fungi have not yet been isolated and identified. Due to the difficulties inherent in culture-based methods, fungal ecologists have turned to community fingerprinting techniques, which utilize signal molecules to profile the fungal members of an environmental sample without culturing. Commonly used signal molecules include chitin, ergosterol, membrane lipids, and nucleic acids. Several DNA-based fingerprinting methods have been successfully applied to fungal communities, including D/TGGE (denaturing/temperature gradient gel electrophoresis), SSCP (single-stranded conformational polymorphism), RISA (ribosomal intergenic spacer analysis), and T-RFLP (terminal restriction length fragment polymorphism). These techniques allow the fungal ecologist to rapidly profile fungal populations in an ecosystem, without the need for laborious culturing or cloning.

Introduction

From the woodland floor to the mouldy shower curtain, fungi are found in almost every environment. Fungal ecologists need to assess the role of fungi in these ecosystems in terms of the fungal species present, their prevalence, and the processes they perform. At present we have a very limited view of fungal activity in any ecosystem studied. To assess fungal community structure one must ask questions about the fungal diversity present, the biomass and mycelial structure of the individual fungal species, and the biochemical functions of those species. Like a jigsaw puzzle, fungal community structure has to be pieced together from diverse data sets to answer these individual questions.

Fungal community structure has proved rather difficult to fully profile in most environments, as individual ecosystems are generally complex, with fungi only forming a component of wider community assemblages. Most fungal biology has concentrated on that part of the Fungal Kingdom that is culturable, visible to the naked eye, or discernible morphologically under the microscope. The very nature of physiological, and biochemical studies requires organisms that can be cultured, and consequently such studies have a tendency to revolve around a few model 'ideal' organisms. There have been enormous efforts to isolate fungi from diverse marine and terrestrial environments, and newly isolated species are routinely reported in mycological journals, which also contain countless articles assessing phylogenetic relationships, largely of cultivable strains. Nevertheless, to many fungal ecologists it is that vast number of unculturable and often unknown fungi that holds the most potential interest.

The extent of the Fungal Kingdom is unclear, and there have been many estimates of global fungal diversity, ranging from 0.5 to 9.9 million fungal species (see Hawksworth, 2001). Mycologists have currently identified and classified around 74,000 fungal species, representing only about 5% of estimated fungal diversity (based on the most commonly accepted estimate of around 1.5 million species). Fungal diversity probably remains underestimated both due to a lack of global exploration and research effort, and low fungal culturability. Although culture-based approaches have told us most of what we presently know about fungal ecology, methods used to isolate fungi tend to select for species able to grow on particular media, and are therefore quite limited. Most environments are oligotrophic, and very different nutritionally from standard (rich) mycological media. Therefore, most fungi isolated in this way are fast-growing species adapted to high substrate levels, and possibly not the dominant organisms of their ecosystem in nature. There has been

widespread concern within microbial ecology as a whole that isolation-based approaches have given a biased view of many microbial ecosystems. To overcome these problems, microbial ecologists have turned in recent years to the analysis of specific signal molecules to assess microbial community structure, and these culture-independent approaches are being taken up by fungal ecologists (Burnett, 2003).

Signal molecules

Fungal ecologists, when interested in complex habitats such as soils, sediments and waters, can rarely distinguish the types and amounts of organisms present. In many cases, fungi are not the dominant members of communities, and techniques are required that can differentiate them from other commonly found groups such as bacteria or arthropods. A frequently used approach is to measure the concentration of a signal molecule. Signal molecules are biological molecules that are generally common to most organisms, but show subtle differences in chemical structure between different species or groups. These differences may be quite broad allowing discrimination simply to group level (i.e. bacteria or fungi), or may be very specific, allowing discrimination to species level. Additionally the amount of a signal molecule in a sample can be an indication of the amount of that organism present.

Many commonly measured signal molecules are constituents of cell walls or membranes, with chitin and ergosterol being used as indicators of fungal presence and biomass. These molecules are utilised because they are broadly specific to the fungi, chitin being a constituent of hyphal walls, and ergosterol being a significant membrane sterol. However, interpretation of these measurements can be problematic as different groups of fungi have differing amounts of these constituents, concentrations can be affected by environmental parameters, and they are found in other groups of organisms such as crustacea, arthropods,

and microalgae (Bermingham, Maltby and Cooke, 1995). Another commonly analysed chemical group is the membrane lipids. Lipids are ubiquitous to all organisms and are structurally diverse, with each species having a characteristic lipid composition. Lipid profiles in environmental samples can be used to describe microbial community composition, with measurement of membrane phospholipids or fatty acids being commonly used. These are often given the acronyms PLFA (phospholipid fatty acid) or FAME (fatty acid methyl ester), respectively. For example, the relative abundance of active bacterial and fungal biomass in a grassland soil was calculated by Bardgett *et al.*, (1999) as a ratio between some PLFAs only found in bacteria, and the PLFA 18:2ω6, which was used to represent fungal biomass.

Although ergosterol and lipid analysis are still widely used, they are unable to distinguish between different fungal species in environmental samples, and thus give a limited view of fungal community structure. Increasingly, fungal ecologists exploit the variability in the genetic code to identify individual species in diverse ecosystems.

Why use DNA?

About 35 years ago the American microbiologist Carl Woese realised that differences in the genetic code could be used to differentiate between and classify microorganisms (Woese, 1967). In particular, he focused on the genes that code for ribosomal RNA (rRNA). Fungi (and most eukaryotes) contain 80S ribosomes, which consist of two subunits – the so-called large (60S) and small (40S) subunits. Each subunit consists of rRNA as a structural molecule and a number of associated proteins, with the large subunit containing the 28S, 5.8S and 5S rRNA molecules and the small subunit containing the18S rRNA molecule. Genes coding for rRNA are suitable signal molecules as the synthesis of ribosomes has been strongly

conserved over evolution, due to the central role of ribosomes in gene expression. The rRNA genes for the rRNA subunits, although not varying greatly in length, contain both strongly conserved and variable regions within their sequences. The genes for these rRNA molecules are also separated by the ITS (internally transcribed spacer) regions, which are highly variable both in length and sequence composition. Fig 1 shows the arrangement of the fungal rRNA genes, together with the ITS regions. This pattern of rRNA subunit genes interspersed with ITS regions continues along the chromosome, with each cluster of genes and ITS regions being separated by intergenic spacers (IGS). By exploiting the coding patterns within rRNA genes, or the ITS or IGS regions, phylogenetic relationships (how closely or distantly species are related) can be determined.

Although these approaches have been widely applied to single fungal isolates, they offer great potential to the fungal ecologist as they can profile fungal populations in samples taken directly from the environment. By extracting the total DNA present in an environmental sample, and using the polymerase chain reaction (PCR) with primers specific for fungal rRNA genes, sequences from the fungal species in the sample can be amplified. Knowledge of these rRNA gene sequences can be compared to a database of rDNA sequences, and depending upon sequence similarities, phylogenetic associations can be drawn. This allows the fungal ecologist to deduce the structure of the fungal community without having to culture or isolate organisms.

How can fungal ecologists exploit rRNA gene sequences?

The first step in a molecular evaluation of the fungal members of an ecosystem is to extract the DNA from samples that represent the environment of interest. These could be from any type of environment where fungal presence is suspected, including atmospheric, aquatic or

terrestrial samples. Methods used to extract DNA depend upon the nature of the sample and how many potentially contaminating substances may be present. For example, complex matrices such as soil contain substances such as humic acids, which can contaminate DNA extracts and make subsequent analysis difficult. Samples from simpler environments, such as water, are less likely to contain such contaminants. Extraction procedures must be optimised on a case-by-case basis, but in general extraction centres on breaking open the fungal cells and separating the released DNA from other cellular constituents such as lipids and proteins and from substances released from the sample matrix. Typical approaches use grinding or bead beating methods to break open fungal cells, followed by phenol/chloroform treatment to purify the DNA. The isolated DNA can then be precipitated using ethanol or isopropanol, and finally dissolved in a buffer system. Various kits are available for this procedure such as the Epicentre SoilMaster™ and MoBIO UltraClean™ DNA extraction kits. The resulting total community DNA will be a mixture containing not just fungal sequences, but also DNA from whatever other organisms are present in the sample.

At this point, several approaches can be taken to analyse the extracted mixed community DNA. The choice of approach depends upon the nature of the information required about the fungal community under investigation. This ranges from a complete inventory of the fungal species present, to general comparisons of fungal community structure. Important considerations for fungal ecologists in the choice of approach include cost, time required for analysis, and the number of samples involved.

Although methods such as cross-hybridisation (Griffiths *et al.*, 1997) and %G+C profiling (Holben and Harris, 1995) have been used in the past to analyse the extracted community DNA directly, PCR-based techniques are now almost universally used. The advantage of

PCR is that with the selection of appropriate primer combinations, rRNA genes of specific groups of organisms within the community DNA can be selectively amplified for subsequent analysis. A number of primer combinations has been developed which are thought to be specific to the Fungi, and allows the fungal ecologist to separate out fungal sequences from the background rRNA genes (see Table 1 for examples). In principle, rDNA primers can be used to discriminate not only between fungi and other kingdoms, but can be designed to differentiate between fungal groups and potentially to the species level. In practice, there may be overlap with sequences from other eukaryotic groups making it difficult to design primer sets specific to fungi (see Anderson, Campbell and Prosser, 2003 for a full discussion).

If a complete inventory of fungal species present is required, the fungal rRNA gene sequences within the community DNA need to be fully sequenced, with each nucleic acid sequence being compared to a computer database in order determine their homology to known sequences (using a sequence matching program such as BLAST, available at http://www.ncbi.nlm.nih.gov/blast/). To obtain sequence information, the community DNA is amplified using PCR with fungal specific primers. The resulting amplified sequences (amplicons) are then inserted into a vector (normally a bacteriophage or a plasmid), which is then entered into the genome of a competent bacterium (usually *E.coli*) by a process known as transfection. As the bacterial clones grow, multiple copies of the gene sequence result. A number of individual transfected colonies (normally 100 or more) is isolated, which make up a clone library. A clone library contains a range of rDNA sequences from the fungal population, and can be screened on the basis of restriction digests for individual rDNAs, which are then sequenced. By comparing these sequences with those in a database, the genetic relationship between an unknown and uncultured organism and known organisms can

be established (Down, 2002). This often allows an unknown organism to be assigned to a group (i.e. Basidomycetes, Oomycetes) which can even give fungal ecologists some idea as to the function of the organism in a community (Bruns and Bidartondo, 2002).

Although cloning and sequencing result in high-quality data detailing fungal members of an environmental sample, it is a relatively costly and time-consuming procedure. Microbial ecologists have come up with simpler, more rapid, methods of profiling fungal populations, collectively known as "community fingerprinting". Although these do not produce the same high-quality sequence data as cloning-based approaches, they have a much higher sample throughput and are especially useful for comparing differences in fungal communities between environments or samples.

Fingerprinting approaches start with mixtures of community PCR amplicons, which are then separated into individual amplicons (called ribotypes) to give a community profile. These DNA mixtures are separated on the basis of sequence composition and/or sequence base pair length. The mixtures are generally separated using electrophoresis, either manually in an electrophoresis gel tank using polyacrylamide gels, or automatically using a nucleic acid sequencer. The former method allows bands to be excised for subsequent sequencing and phylogenetic analysis, whereas the latter method is more sensitive in detecting amplicons and has high sample throughput. A number of community fingerprinting methods is shown in Fig 2.

The most commonly used electrophoretic separations based on sequence composition are denaturing- or temperature-gradient gel electrophoresis (DGGE or TGGE) (Muyzer, DeWaal and Uitterlinden, 1993). DGGE/TGGE are normally used to separate mixtures of amplicons

of the same length (up to a maximum sequence length of around 800bp). The amplified sequences are electrophoresed on a polyacrylamide gel either containing an increasing gradient of chemical denaturants in the case of DGGE, or using a temperature gradient (TGGE). As the amplicons migrate down the gel, they "melt", changing from double stranded to single stranded. The denaturant concentration or temperature at which they melt varies according to sequence composition. The stronger bond in DNA is the guanine + cytosine bond, thus those amplicons with a high G+C concentration melt later than those with more adenine + thymine bonds. Once the sequence melts, it sticks in the gel and does not migrate further. The completed gel is then stained with a nucleic acid dye (such as silver or ethidium bromide) and visualised. Once stained, gels will have one discrete band per amplicon, and the number of bands in the visualised profile indicates the number of ribotypes (putative species) present. Individual bands can then be excised and sequenced to identify species, although the short sequences typically used in D/TGGE are limited in their value for subsequent phylogenetic analysis. In addition, gel banding patterns can be analysed for band intensity in order to estimate the sizes of individual species populations. T/DGGE has been used successfully to profile fungal communities from environments as diverse as wheat rhizospheres (Smit *et al.*, 1999) and seventeenth century mural paintings (Möhlenhoff *et al.*, 2001). Similar to DGGE/TGGE is single-stranded conformational polymorphism (SSCP); amplicons of different sequences have different conformational structures which affects their migration through a polyacrylamide gel (Peters *et al.,* 2000).

Other fingerprinting techniques exploit the variability of the internal transcribed spacer genes (ITS) in fungi. These regions are very variable in sequence composition and also vary in length between species, and thus can be used to profile the number of ribotypes present in a community. DNA is amplified using fungal ITS primers that target the ribosomal DNA

region between the 18S and 28S genes containing the two internal transcribed spacers (ITS) and the 5.8S rRNA gene. The various sized amplicons can be separated conventionally on a polyacrylamide gel, a technique known as ribosomal intergenic spacer analysis (RISA), or automatically on a sequencer (automated ribosomal intergenic spacer analysis, or ARISA). Each ribotype will form one discrete band (on a gel) or peak (in a sequencer profile), revealing the number of species present. The intensity of the band or size of the peak height relative to the overall sample can be used as a crude estimate of the abundance of certain ribotypes in the community. For example, ARISA has been used to profile soil fungal communities (Ranjard *et al.*, 2001).

Restriction-fragment length polymorphism (RFLP) and its automated version, terminal-RFLP (T-RFLP), are similar to RISA and ARISA, but involve an intermediate restriction digest after PCR and before separation of fragments. The procedure can be applied to both the rRNA genes and ITS regions, with amplified fragments subjected to a digest with one or more restriction enzymes, followed by separation by electrophoresis on a polyacrylamide gel (RFLP). T-RFLP develops this by separating fragments using an automated sequencer, which detects amplicons which have been fluorescently labelled at one end through the use of a fluorescently tagged primer during PCR. RFLP and T-RFLP patterns are analysed in a similar way to RISA/ARISA, with each band or peak counted as an individual fungal ribotype. T-RFLP has recently been successfully used to profile ascomycete communities in a salt marsh (Buchan *et al.*, 2002).

Cautions

Fingerprinting techniques are dependent upon PCR, and profiles are affected by the drawbacks of this method (see von Wintzingerode, Gobel and Stackebrandt, 1997 for detailed

discussion). During PCR amplification, a number of problems can occur, including the formation of chimeras, mispriming, formation of heteroduplexes, and over or underamplification of particular sequences. Many of the primer sets commonly used for fungal community studies have also been shown to co-amplify genes from other eukaryotes, such as invertebrates, particularly those designed to amplify the 18S rDNA region (Anderson *et al.*, 2003). There is also some concern that some universal fungal primers may preferentially amplify sequences from certain taxonomic groups. Even if the genes amplified are of fungal origin, lack of variability between species within the 18S rDNA can reduce the taxonomic resolution to the level of genus or above. The ITS region exhibits higher sequence variability, and sample profiles generated with ITS primers show higher diversity than those amplified by 18S rDNA primers (Lord *et al.* 2002; Anderson *et al.*, 2003), but the limited nature of current fungal sequence databases, which tend to focus on medically and commercially important fungi, makes phylogenetic association difficult (Bridge *et al.*, 2003). In some studies, the intensity of amplicon bands or peak heights has been used as an indication of relative fungal biomass present within a sample (van Elsas *et al.*, 2000; Brodie, Edwards and Clipson, 2003). This approach must also be treated with some caution as different species have different copy numbers of rDNA (Bridge and Spooner, 2001), making it impossible to definitively quantify fungal populations using techniques based on standard PCR.

Concluding Remarks

Community fingerprinting techniques provide a powerful new weapon in the fungal ecologist's arsenal. Diversity profiles of fungal communities from complex substrates can be generated relatively rapidly, with the level of replication necessary to draw statistical

conclusions. Although DNA-based methods have certain limitations, molecular approaches to fungal ecology give the fullest view of fungal community structure yet possible.

References

Anderson, I. C., Campbell, C. D. & Prosser, J. I. (2003). Potential bias of fungal 18S rDNA and internal transcribed spacer polymerase chain reaction primers for estimating fungal biodiversity in soil. *Environmental Microbiology* **5**: 36-47.

Bardgett, R. D., Mawdsley, J. L., Edwards, S., Hobbs, P. J., Rodwell, J. S. and Davies, W. J. (1999). Plant species and nitrogen effects on soil biological properties of temperate upland grasslands. *Functional Ecology* **13**: 650-660.

Bermingham, S., Maltby, L. and Cooke, R. C. (1995). A critical assessment of the validity of ergosterol as an indicator of fungal biomass. *Mycological Research* **99**: 479-484.

Borneman, J. and Hartin, R. J. (2000). PCR primers that amplify fungal rRNA genes from environmental samples. *Applied and Environmental Microbiology* **66**: 4356-4360.

Bridge, P. and Spooner, B. (2001). Soil fungi: diversity and detection. *Plant and Soil* **232**: 147-154.

Bridge, P., Roberts, P. J., Spooner, B. M., and Panchal, G. (2003). On the unreliability of published DNA sequences. *New Phytologist* **160** (1): 43-48.

Brodie, E., Edwards, S. and Clipson, N. (2003). Soil fungal community structure in a temperate upland grassland soil. *FEMS Microbiology Ecology* **45**: 105-114.

Bruns, T. D. and Bidartondo, M. I. (2002). Molecular windows into the below-ground interactions of ectomycorrhizal fungi. *Mycologist* **16** (2): 47-50.

Buchan, A., Newell, S. Y., Moreta, J. I. L. and Moran, M. A. (2002). Analysis of internal transcribed spacer (ITS) regions of rRNA genes in fungal communities in a southeastern U.S. salt marsh. *Microbial Ecology* **43**: 329-340.

Burnett, J. H. (2003). *Fungal populations and species*. Oxford, Oxford University Press.

Down, G. (2002). Fungal family trees: finding relationships from molecular data. *Mycologist* **16** (2): 51-58.

Gardes, M. and Bruns, T.D. (1993). ITS primers with enhanced specificity for basidiomycetes: application to the identification of mycorrhiza and rusts. *Molecular Ecology* **2**: 113-118.

Griffiths, B. S., Diaz-Ravina, M., Ritz, K., McNicol, J. W., Ebblewhite, N., and Bååth, E. (1997). Community DNA hybridisation and %G+C profiles of microbial communities from heavy metal polluted soils. *FEMS Microbiology Ecology* **24**: 103-112.

Hawksworth, D. L. (2001). The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycological Research* **105**: 1422-1432.

Holben, W. E. and Harris, D. (1995). DNA-based monitoring of total bacterial community structure in environmental samples. *Molecular Ecology* **4**: 627-631.

Lord, N. S., Kaplan, C. W., Shank, P., Kitts, C. L. and Elrod, S. L. (2002). Assessment of fungal diversity using terminal restriction fragment (TRF) pattern analysis: comparison of 18S and ITS ribosomal regions. *FEMS Microbiology Ecology* **42**: 327-337.

Möhlenhoff, P., Muller, L., Gorbushina, A. A. and Petersen, K. (2001). Molecular approach to the characterisation of fungal communities: methods for DNA extraction, PCR amplification and DGGE analysis of painted art objects. *FEMS Microbiology Letters* **195**: 169-173.

Muyzer, G., DeWaal, E. C., and Uitterlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* **59** (3): 695-700.

Peters, S., Koschinsky, S., Schwieger, F. and Tebbe, C. C. (2000). Succession of microbial communities during hot composting as detected by PCR-single-strand-conformation polymorphism-based genetic profiles of small-subunit rRNA genes. *Applied and Environmental Microbiology* **66**: 930-936.

Ranjard, L., Poly, F., Lata, J.-C., Mougel, C., Thioulouse, J. and Nazaret, S. (2001). Characterization of bacterial and fungal soil communities by automated ribosomal intergenic spacer analysis fingerprints: biological and methodological variability. *Applied and Environmental Microbiology* **67**: 4479-4487.

Sequerra, J., Marmeisse, R., Valla, G., Normand, P., Capellana, A. and Moiroud, A. (1997). Taxonomic position and intraspecific variability of the nodule forming Penicillium nodositatum inferred from RLFP analysis of the ribosomal intergenic spacer and random amplified polymorphic DNA. *Mycological Research* **101**: 465-472.

Smit, E., Leeflang, P., Glandorf, B., van Elsas, J. D. and Wernars, K. (1999). Analysis of fungal diversity in the wheat rhizosphere by sequencing of cloned PCR-amplified genes encoding 18S rRNA and temperature gradient gel electrophoresis. *Applied and Environmental Microbiology* **65**: 2614-2621.

van Elsas, J. D., Duarte, G. F., Keijzer-Wolters, A., and Smit, E. (2000). Analysis of the dynamics of fungal communities in soil via fungal-specific PCR of soil DNA followed by denaturing gradient gel electrophoresis. *Journal of Microbiological Methods* **43**: 133-151.

von Wintzingerode, F., Gobel, U. B. and Stackebrandt, E. (1997). Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiology Reviews* **21**: 213-229.

White, T. J., Bruns, T. D., Lee, S. and Taylor, J. (1990). Analysis of phylogenetic relationships by amplification and direct sequencing of ribosomal RNA genes in M. A.

Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (Eds), *PCR protocols: a guide to methods and applications*, New York, Academic Press, 315-322.

Woese, C. R. (1967). *The genetic code: the molecular basis for genetic expression.* New York, Harper and Rowe.
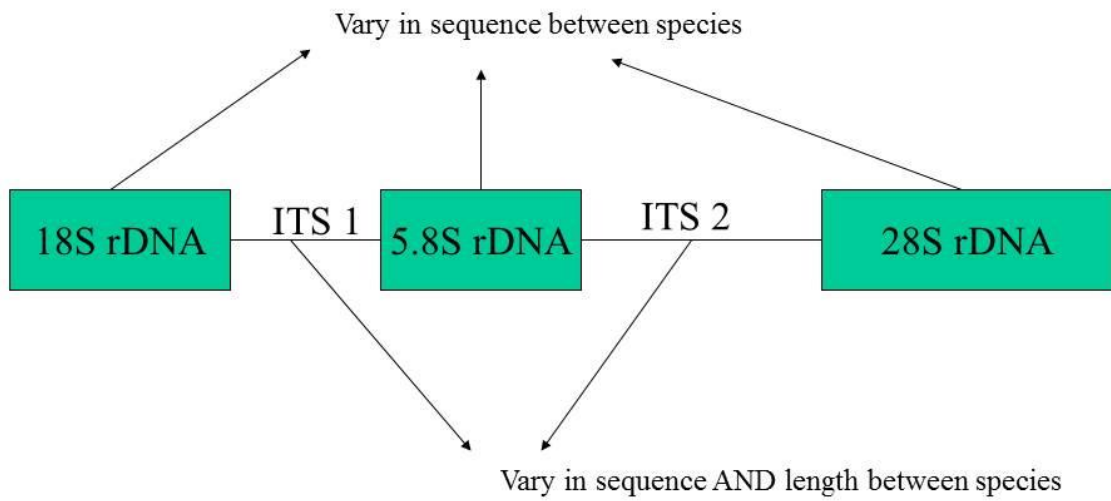
Figure and Table Captions

Fig 1. Graphic representation of ribosomal RNA gene structure in fungi.

Table 1. Examples of PCR primers commonly used for amplification of fungal sequences in community fingerprinting studies.

Fig 2. Flowchart detailing steps in common fungal community fingerprinting procedures. The DGGE, T-RFLP and ARISA profiles are all of fungal amplicons from the same upland grassland soil sample. Note the increase in fragments (bands and peaks) from DGGE through T-RFLP to the ARISA profile, reflecting the increased sensitivity of the approaches. T-RFLP and ARISA are generated from automated sequence analysis, giving both amplicon size (numbers on graphs – in base pairs) and fluorescence intensity.

Vary in sequence between species

18S rDNA | ITS 1 | 5.8S rDNA | ITS 2 | 28S rDNA

Vary in sequence AND length between species

Region Amplified	Name	Reference
18S	nu-SSU-817	Borneman and Hartin 2000
	nu-SSU-1196	Borneman and Hartin 2000
	nu-SSU-1536	Borneman and Hartin 2000
	EF4	Smit *et al.* 1999
	EF3	Smit *et al.* 1999
	Fung5	Smit *et al.* 1999
ITS	ITS1-F	Gardes and Bruns 1993
	ITS4	White *et al.* 1990
	2234-C	Sequerra *et al.* 1997
	3126	Sequerra *et al.* 1997

Environmental Sample

Nucleic Acid Extraction

Mixture of environmental DNA

Polymerase Chain Reaction
with fungal-specific primers

Separation of sequences
using D/TGGE

rDNA Fragments

Separation of different-sized ITS
fragments using ARISA

Restriction digest

Separation of terminal
fragments using T-RFLP