Environmental Microbiology (2003) **5**(1), 36–47

Potential bias of fungal 18S rDNA and internal transcribed spacer polymerase chain reaction primers for estimating fungal biodiversity in soil

Ian C. Anderson,[1,2]* Colin D. Campbell[1] and James I. Prosser[2]

[1]*The Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK.*

[2]*Department of Molecular and Cell Biology, University of Aberdeen, Institute of Medical Sciences, Foresterhill, Aberdeen AB25 2ZD, UK.*

Summary

Four fungal 18S rDNA and internal transcribed spacer (ITS) polymerase chain reaction (PCR) primer pairs were tested for their specificity towards target fungal DNA in soil DNA extracts, and their ability to assess the diversity of fungal communities in a natural grassland soil was compared. Amplified PCR products were cloned, and ≈ 50 clones from each library were sequenced. Phylogenetic analysis and database searches indicated that each of the sequenced cloned DNA fragments was of fungal origin for each primer pair, with the exception of the sequences generated using the 18S rDNA primers nu-SSU-0817 and nu-SSU-1196, where 35 of the 50 sequenced clones represented soil invertebrates. Although some of the primers have previously been suggested to be biased towards certain fungal taxonomic groups, the ratio of sequences representing each of the four main fungal phyla, *Ascomycota*, *Basidiomycota*, *Chytridiomycota* and *Zygomycota*, was similar for each of the primer pairs, suggesting that primer bias may be less significant than previously thought. Collector's curves were plotted to estimate the coverage obtained for each of the clone libraries after clustering the sequences into operational taxonomic units at a level of 99% sequence similarity. The curves indicated that good coverage of diversity was achieved, with the exception of the clone library constructed using primers nu-SSU-0817 and nu-SSU-1196, on account of the high number of non-fungal sequences obtained. The work demonstrates the usefulness of 18S rDNA and ITS PCR primers for assessing fungal diversity in environmental samples, and it also highlights some potential limitations of the approach with respect to PCR primer specificity and bias.

Received 2 August, 2002; revised 30 September, 2002; accepted 30 September, 2002. *For correspondence at the first address. E-mail i.anderson@macaulay.ac.uk; Tel. (+44) 1224 498 200; Fax (+44) 1224 498 207.

Introduction

Fungi play fundamentally important and diverse roles in terrestrial ecosystems, being involved in many of the key processes required for ecosystem functioning. They are important as pathogens of plants and animals, as mycorrhizal symbionts of plants (Smith and Read, 1997) and as the main agents for the decomposition of organic material. Fungi therefore possess the ability to control nutrient fluxes in natural ecosystems that may be aided through extensive below-ground mycelial networks. Despite their importance in terrestrial ecosystems, little is known of the diversity of natural fungal populations. Recent estimates suggest that 1.5 million fungal species are present in natural ecosystems, but only 5–10% have been described formally (Hawksworth, 1991; 2001; Hawksworth and Rossman, 1997). The primary reason for this apparent ignorance has been the dependence on cultivation-based techniques for the characterization of fungal diversity. The major limitations of these techniques are the inability to separate biomass from particulate material and a lack of growth media and cultivation conditions suitable for all members of the community.

The development and application of molecular techniques, particularly those based on the analysis of 16S rRNA genes amplified from extracted DNA, have transformed studies of bacterial diversity in natural environments. Similar techniques, however, have only recently been used to characterize fungal diversity (Kowalchuk *et al*., 1997; Kowalchuk, 1998; Smit *et al*., 1999; Borneman and Hartin, 2000; Vainio and Hantula, 2000; van Elsas *et al*., 2000; Lowell and Klein, 2001; Möhlenhoff *et al*., 2001; Pennanen *et al*., 2001; Schabereiter-Gurtner *et al*., 2001). A major challenge in applying these techniques to fungal communities is the design of suitable polymerase chain reaction (PCR) primers with specificity for fungal DNA, while reducing co-amplification of similar target DNA from non-fungal sources. Numerous PCR primers have been described that amplify fungal rDNA from a wide range of taxonomic groups (White *et al*., 1990), although few were designed for use with DNA

extracted from mixed communities, and their lack of selectivity may lead to inaccurate estimates of fungal diversity when applied to complex environmental samples. Gardes and Bruns (1993) designed ITS1-F and ITS4-B primers specifically to amplify fungal internal transcribed spacer (ITS) regions, without the co-amplification of plant or other eukaryotic DNA. Furthermore, several PCR primers have been designed recently to amplify partial 18S rRNA gene sequences from species belonging to the four major fungal phyla (*Ascomycota*, *Basidiomycota*, *Chytridiomycota* and *Zygomycota*) from DNA extracted directly from soil (Smit *et al*., 1999; Borneman and Hartin, 2000). Although a critical factor in designing PCR primers to fungal 18S rRNA genes is their specificity towards the target fungal DNA, achieving that specificity may ultimately bias the view that we obtain by attempting to avoid amplification of non-fungal DNA. This is because some regions of fungal 18S rRNA gene sequences share high similarity with other eukaryotes. This has indeed been suggested to be the case with recently designed fungal 18S rDNA primer pairs EF4/EF3 and EF4/fung5 (Smit *et al*., 1999). There have also been conflicting reports about the specificity of newly designed fungal 18S rDNA primers (Smit *et al*., 1999; Borneman and Hartin, 2000). For example, although the PCR primer pairs EF4/EF3 and EF4/fung5 were shown in one study to amplify only fungal 18S rDNA sequences from wheat rhizosphere soil (Smit *et al*., 1999), other investigators have shown that the same primers can also amplify some non-fungal template (Borneman and Hartin, 2000). In addition, 18S rRNA gene sequences are generally only able to resolve taxonomic groups to the level of genus, and the taxonomic resolution of fungal 18S rDNA and ITS sequences is limited by the current availability of information held within databases. The aim of the current study was to assess the specificity of previously published fungal 18S rDNA and ITS primers for estimating fungal diversity in environmental samples and to use them to characterize the diversity of fungal communities in a natural grassland soil. This was achieved by the construction of clone libraries from PCR products amplified using published primer sets and comparison with respect to the relative abundances of sequences falling within the major fungal groups.

Results

Amplification of fungal 18S rDNA and ITS fragments

PCR amplification of DNA extracted from soil was successful with all four of the primer pairs tested. In each case, PCR products of the expected sizes appeared as intense bands on agarose gels (data not shown). The approximate sizes of the amplified PCR products were: 400–500 bp (nu-SSU-0817 and nu-SSU-1196); 700–800 bp (nu-SSU-0817 and nu-SSU-1536); 1.4–1.5 kb (EF4 and EF3); 500–800 bp (ITS1-F and ITS4). PCRs were performed in both the presence and the absence of bovine serum albumin (BSA) for each primer pair. In each case, the PCR product yield was considerably increased in the presence of BSA.

Identification and phylogenetic analysis of nu-SSU-0817 and nu-SSU-1196 (A) clones

A phylogenetic analysis and FASTA database search were conducted for the sequences of 50 random clones from the nu-SSU-0817 and nu-SSU-1196 (A) clone library. Of the 50 clones sequenced, 35 were found to have 89–95% sequence similarity to *Crossodonthina koreana* or *Hypogastrura dolsana*, both of which are species of invertebrates that are commonly found in soil. The only exceptions were clone A30, which is most closely related to *Rhinosporidium seeberi* (89% sequence similarity), an aquatic protistan parasite, and clone A1, which is most closely related to *Diplolaimelloides meyli* (86% sequence similarity), a soil arthropod. All these sequences had 93.1–100% sequence similarity to each other over the entire sequence length, and they clustered together in the phylogenetic tree (Fig. 1). The remaining 15 clones clustered in a single clade of the phylogenetic tree and were separated from the non-fungal sequences by a strongly supported branch (92% bootstrap support). Each of these cloned sequences had a high sequence identity to fungal 18S rDNA sequences in the GenBank database (Fig. 1). A further phylogenetic analysis was carried out using the 15 fungal sequences to determine more accurately the relationships between sequences within this group (Fig. 2). All four major fungal phyla (*Ascomycota*, *Basidiomycota*, *Chytridiomycota* and *Zygomycota*) were represented among the 15 fungal clones.

Identification and phylogenetic analysis of nu-SSU-0817 and nu-SSU-1536 (B) clones

In contrast to the nu-SSU-0817 and nu-SSU-1196 (A) clone library, the FASTA results for the sequences of 48 clones analysed from the nu-SSU-0817 and nu-SSU-1536 (B) clone library revealed that all sequences were of fungal origin (Fig. 2). Thirty-two were most closely related to fungal basidiomycetes, whereas 15 were most closely related to species of ascomycetes. All 32 basidiomycete clones had 94.2–99.0% sequence similarity to basidiomycete sequences in GenBank and EMBL over the entire sequence length. In contrast, although some of the ascomycete sequences had high similarity to GenBank sequences (as high as 99.3%), sequence identities of several were as low as 88.0%, preventing reliable identification to the species and, in some instances, the genus

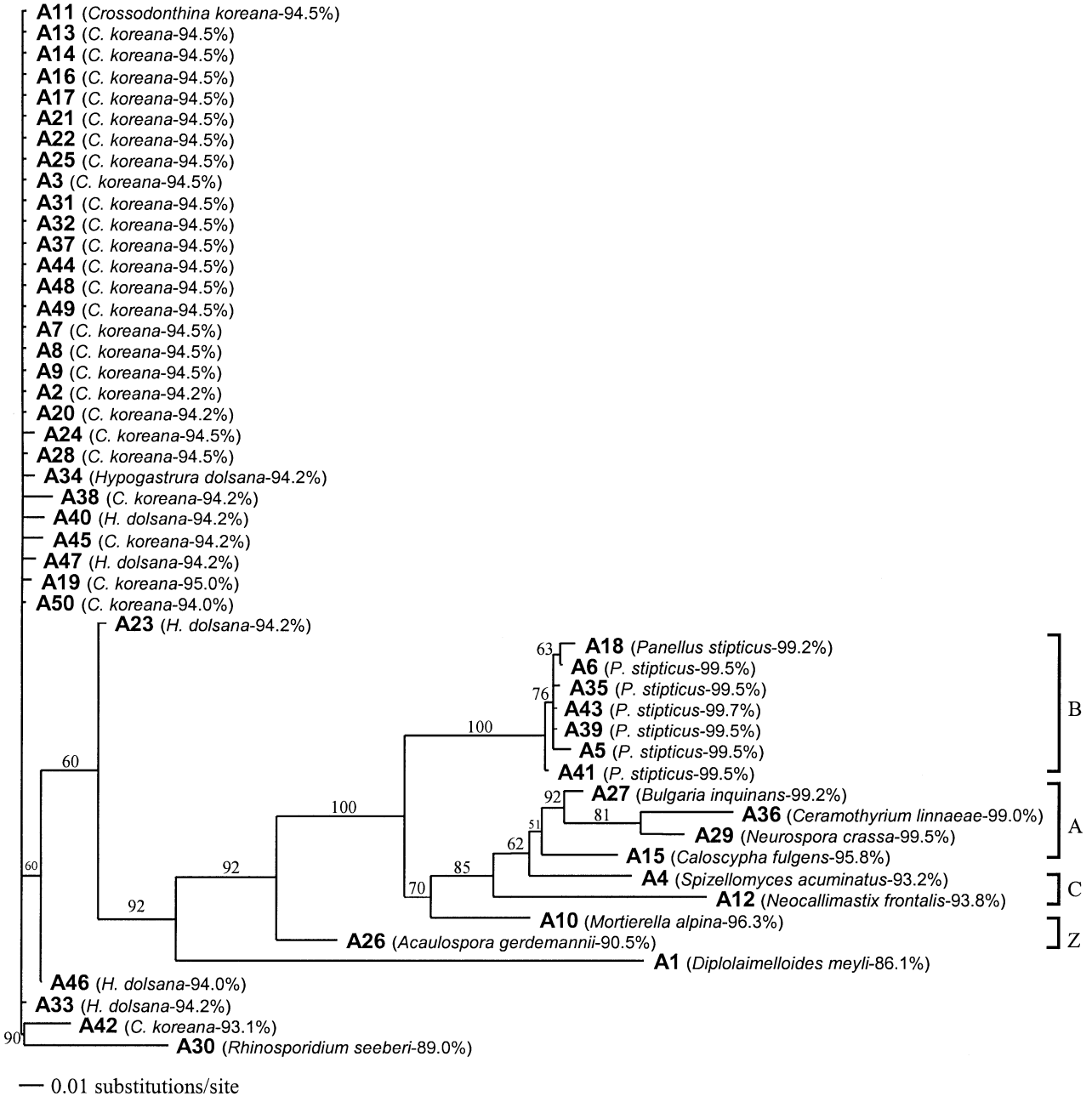

Fig. 1. Neighbour-joining tree showing the relationship between 50 18S rDNA clones obtained from DNA extracted from grassland soil using the primers nu-SSU-0817 and nu-SSU-1196 (1) (A clone library). A, ascomycete; B, basidiomycete; C, chytridiomycete; and Z, zygomycete sequences based on the results of a FASTA database search (closest sequence matches are shown in brackets). Sequences ranged in size from 376 to 385 nucleotides. The scale bar represents the number of base substitutions per site. Bootstrap values are shown for those branches that had >50% support in a bootstrap analysis of 1000 replicates.

level. A single clone (B13) was most closely related to a chytridiomycete (*Spizellomyces acuminatus*), whereas none of the 48 clones sequenced in this clone library represented zygomycetes (Fig. 3). As in the A clone library, the divergence between the majority of basidiomycete sequences was low, and the sequences of seven clones (B2, B20, B27, B35, B43, B49 and B6) were identical.

Identification and phylogenetic analysis of EF4 and EF3 (C) clones

Results using the EF4 and EF3 primer pair (clone library C) were similar to those observed for clone library B, in that all 46 clones analysed were most closely related to fungal 18S rDNA sequences in the GenBank and EMBL databases. Thirty-five clones were most closely related to

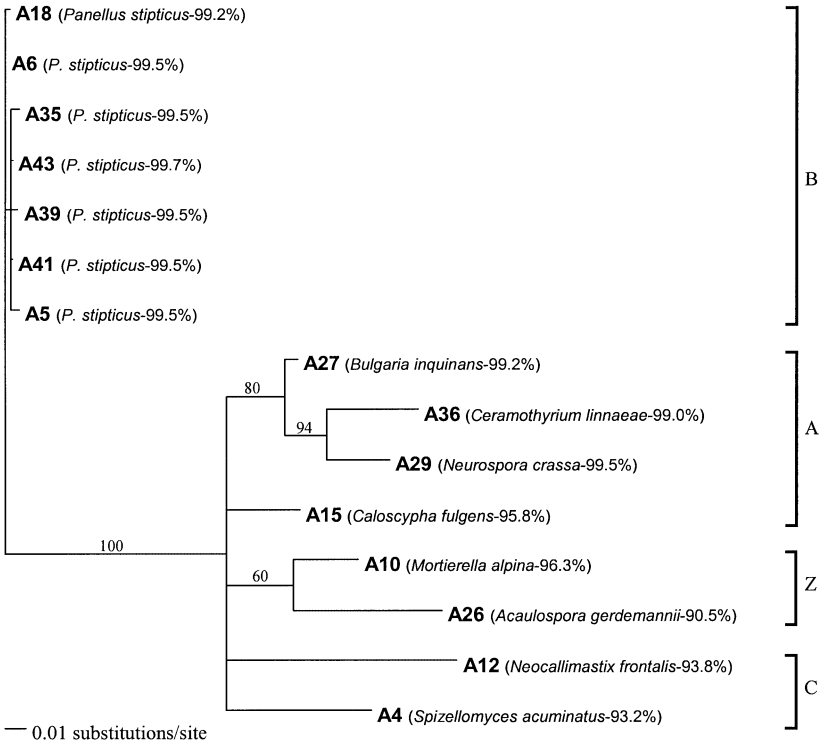

Fig. 2. Neighbour-joining tree showing the relationship between the 15 fungal 18S rDNA clones obtained from DNA extracted from grassland soil using the primers nu-SSU-0817 and nu-SSU-1196 (1) (A clone library). Sequences ranged in size from 378 to 385 nucleotides. Abbreviations and other information as described in the legend to Fig. 1.

basidiomycetes, each having 92.5–99.2% sequence similarity to basidiomycete sequences (Fig. 4). As observed for clone library B, the divergence within the basidiomycete sequences was low. All basidiomycete sequences had 98.3–100% identity over the entire sequence, which is indicated by the short branch lengths in the basidiomycete clade of the phylogenetic tree (Fig. 4), with the exception of clones C15, C29, C43 and C47, which were more distantly related to other basidiomycete sequences. Of the remaining 11 clones in this library, nine were most closely related to ascomycete sequences and two to zygomycete sequences. Each of these 11 remaining clones had a 94.2–100% match with a fungal 18S rDNA sequence in the database. Although the ascomycete and basidiomycete groups clustered together in a single clade of the phylogenetic tree, the zygomycete sequences fell into a clade separated from the ascomycete/basidiomycete groups by a strongly supported branch (100% bootstrap support from 1000 replicates) (Fig. 4).

Identification and phylogenetic analysis of ITS1-F and ITS4 (D) clones

All 50 clones sequenced from the ITS1-F and ITS4 (D) clone library were most closely related to fungal ITS sequences within the GenBank and EMBL databases. Phylogenetic analysis of the 50 ITS sequences placed them in two major groups representing ITS sequences of ascomycete and basidiomycete origin. These groups were separated by a strongly supported branch of the phylogenetic tree (100% bootstrap support from 1000 replicates) (Fig. 5). FASTA results and phylogenetic analysis indicated that none of the 50 ITS sequences generated using this primer pair represented a chytridiomycete or zygomycete ITS sequence. The ascomycete group consisted of 10 clone sequences with high similarity (76.4–89.7%) to fungal ascomycete ITS sequences in the databases (Fig. 5). With the exception of clone D54, the remaining 40 sequences had the highest sequence similarity to ITS sequences of basidiomycetes. These sequences clustered further into three smaller clades within the basidiomycete group (Fig. 5). The level of sequence divergence between the clone sequences within the basidiomycete clades I and II was relatively low, as indicated by relatively small branch lengths. In contrast, the remaining basidiomycete sequences (clade III) were placed on much longer branches in the tree, suggesting greater divergence than that exhibited by the other basidiomycete and ascomycete sequences analysed here (Fig. 5). A similar phylogenetic analysis was conducted for all D clone sequences, using the portion of the sequence encoding the 5.8S rRNA gene (Fig. 6), as entire ITS sequences were difficult to align because of significant differences in sequence length and base composition. Clustering of the 50 cloned 5.8S rRNA gene sequences was similar to that of the ITS phylogenetic tree, but with small rearrangements in some groups (Fig. 6). The sequences in the basidiomycete clade III were as diverse over the 5.8S rRNA gene sequence as

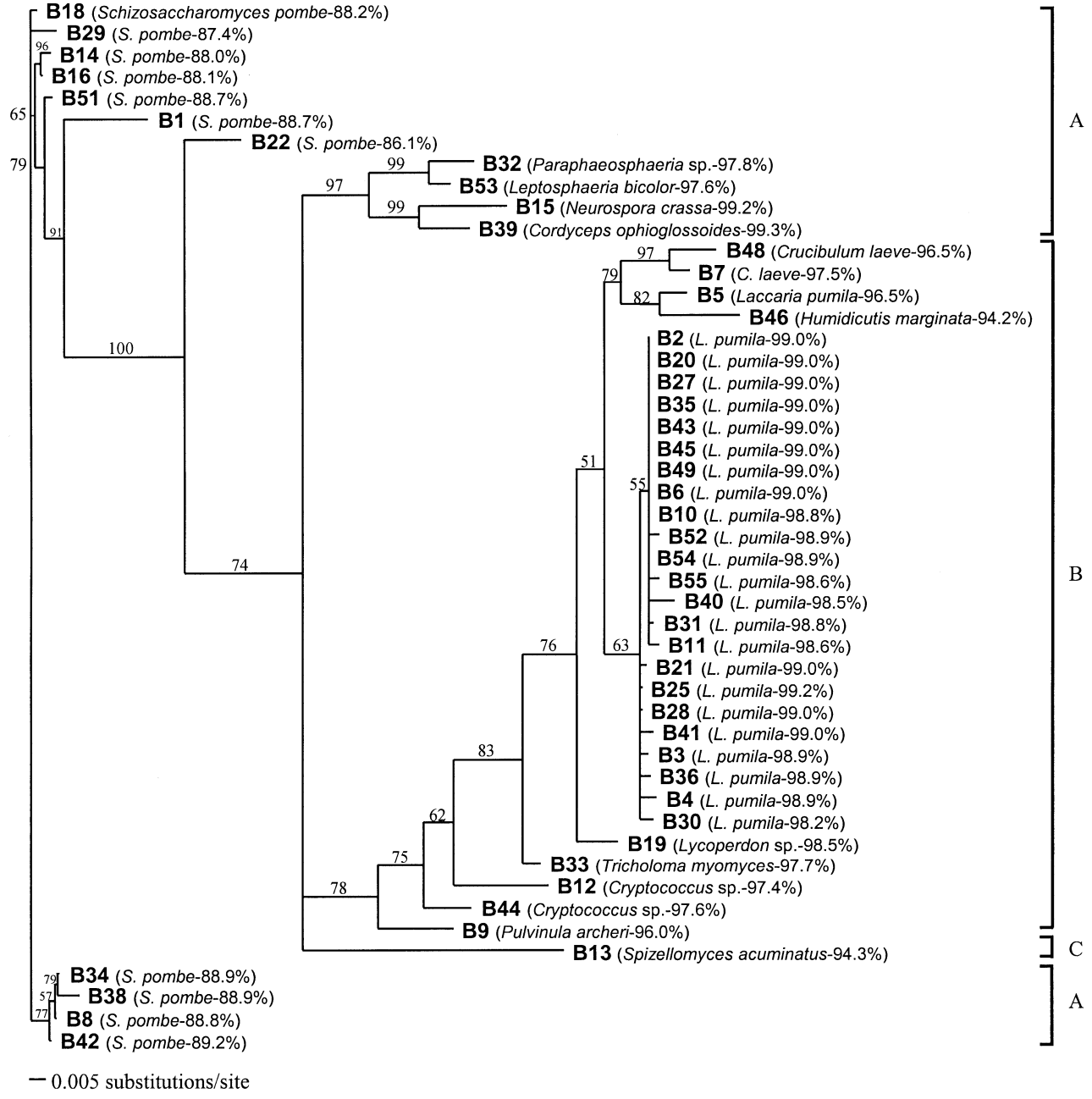

Fig. 3. Neighbour-joining tree showing the relationship between 48 fungal 18S rDNA clones obtained from DNA extracted from grassland soil using the primers nu-SSU-0817 and nu-SSU-1536 (1) (B clone library). Sequences ranged in size from 718 to 744 nucleotides. Abbreviations and other information as described in the legend to Fig. 1.

the whole ITS region. Although clone D54 falls within the basidiomycete group in this phylogenetic analysis, the closest match to this sequence in the database was with ITS sequences from ascomycete fungi.

Collector's curves

Collector's curves for each of the four clone libraries A, B, C and D were constructed as plots of the number of operational taxonomic units (OTUs) versus the number of clones sampled for each clone library (Fig. 7). Sequences with a similarity >99% were considered to belong to the same OTU. Percentage coverage was also determined, with full coverage of the diversity in any library indicated by the production of a plateau-shaped curve. Collector's curves for clone libraries B, C and D reached a plateau, suggesting that samples of 50 clones from these clone libraries provided reasonable coverage of the fungal diver-

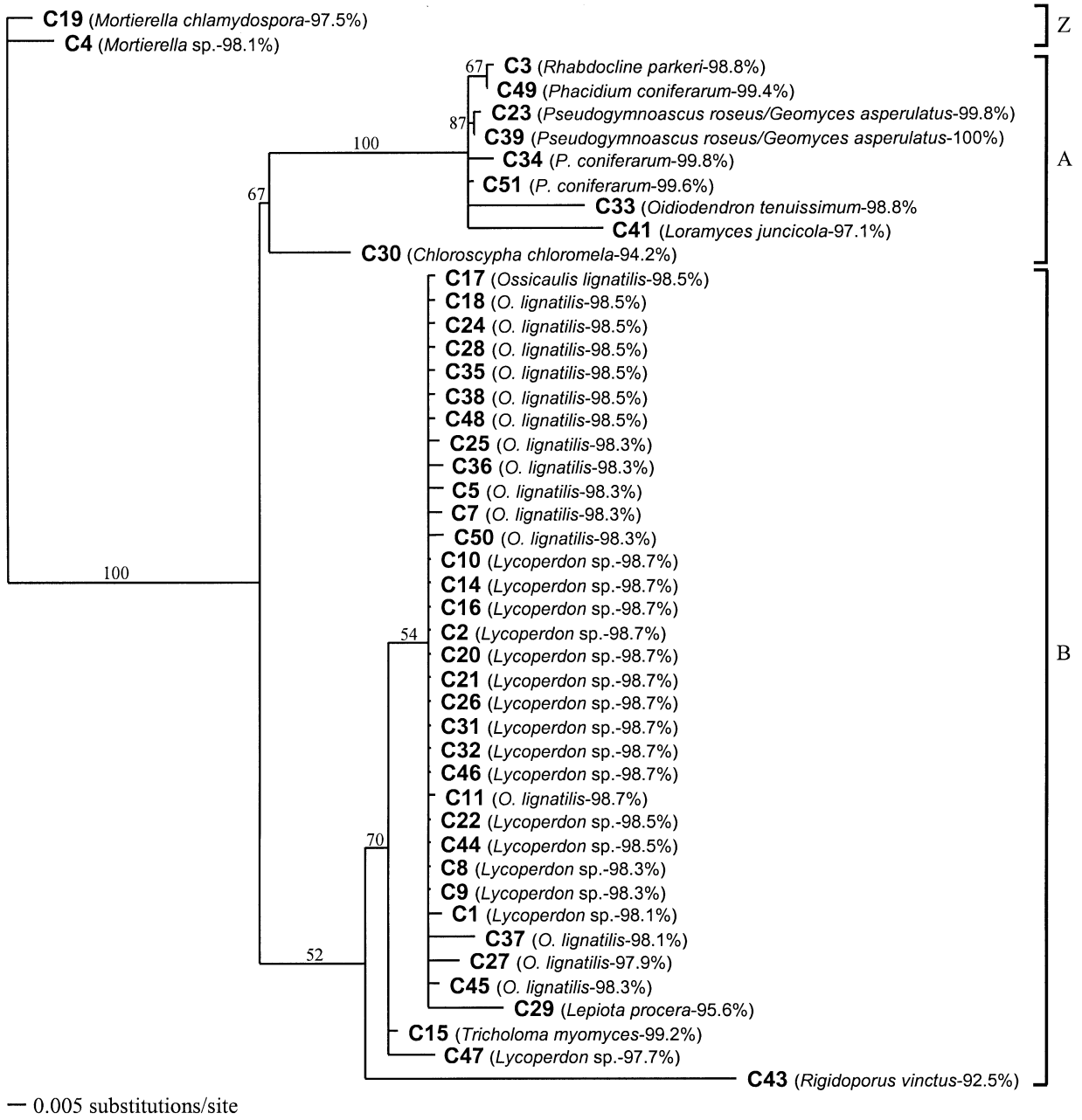

Fig. 4. Neighbour-joining tree showing the relationship between 46 fungal 18S rDNA clones obtained from DNA extracted from grassland soil using the primers EF4 and EF3 (19) (C clone library). Sequences ranged in size from 515 to 521 nucleotides. Abbreviations and other information as described in the legend to Fig. 1.

sity. The extent of coverage varied for each of the libraries, with values for percentage coverage of 58%, 74% and 46% for clone libraries B, C and D respectively. Coverage for clone library A was 33%, and the collector's curve for this library did not reach a plateau, probably because only 15 of the 50 clones sequenced were of a fungal origin (see Fig. 1).

A greater level of diversity was observed in the sequences from the ITS clone library (D), and this is reflected in a larger number of OTUs observed (Fig. 7). Collector's curves were also plotted for the ITS clone library (D) for OTU sequence similarity values of >70%, 80%, 90%, 95%, 98% and 99% (data not shown). Reducing the percentage sequence similarity to 98% or 95% did not significantly alter the collector's curve, but further reductions significantly reduced the number of sampled clones required for the collector's curve to plateau. Decreasing sequence similarity values to 90%, 80% and

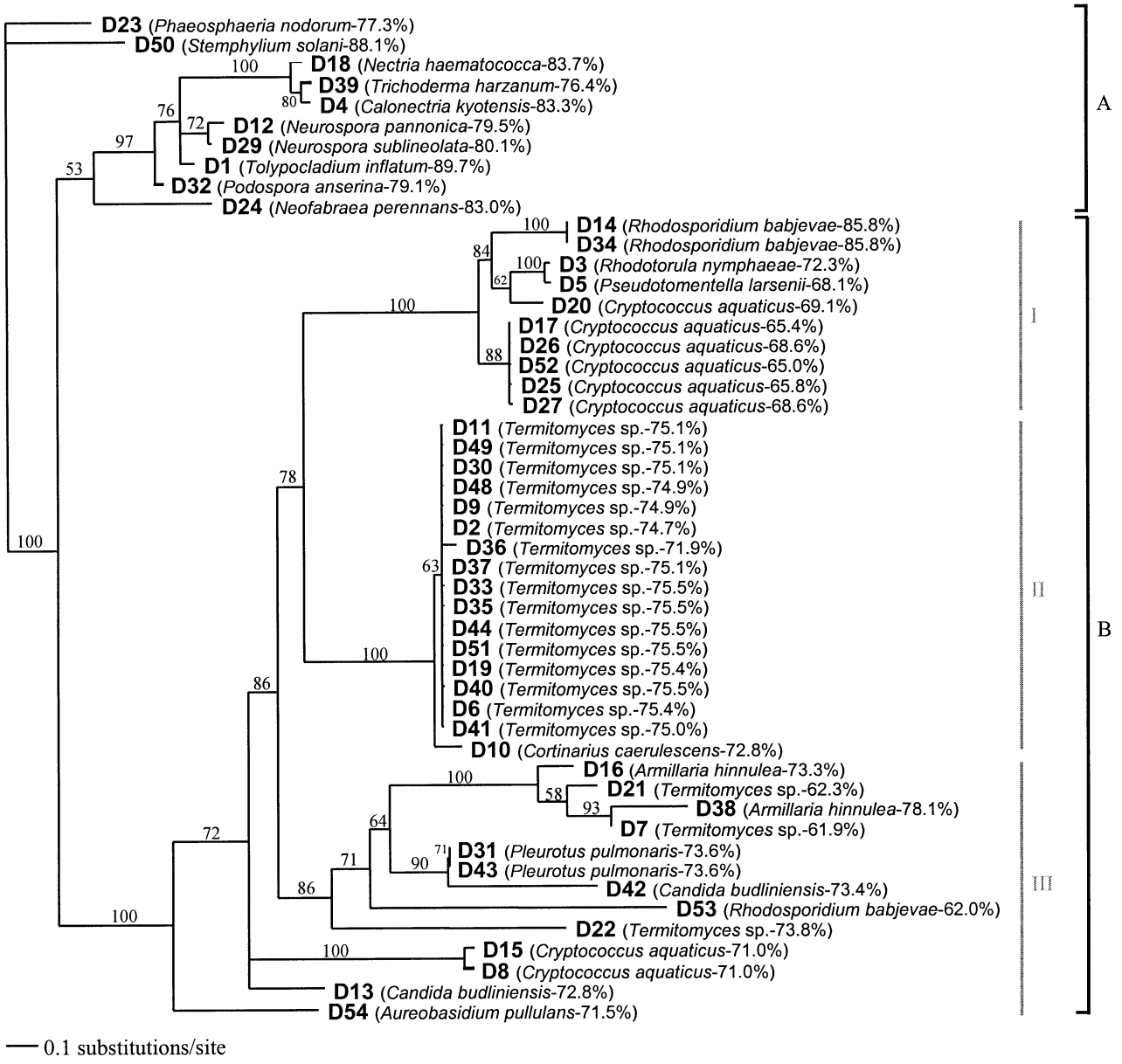

Fig. 5. Neighbour-joining tree showing the relationship between 50 ITS clones obtained from DNA extracted from grassland soil using the primers ITS1-F (3) and ITS4 (26) (D clone library). Sequences ranged in size from 467 to 742 nucleotides. Abbreviations and other information as described in the legend to Fig. 1.

70% also increased the values for percentage coverage of the diversity in the clone library to 54%, 60% and 72% respectively.

Discussion

The initial aim of this study was to assess the specificity of previously published 18S rDNA and ITS PCR primers in estimating fungal biodiversity in environmental samples. A major factor in designing PCR primers for environmental studies is the ability of the primers to amplify all members of the fungal community without bias, while also excluding the co-amplification of other eukaryote sequences. This is particularly difficult when using the relatively highly conserved 18S rDNA region (Bruns *et al*., 1991; Page and Holmes, 1998). Phylogenetic analysis indicated that all sequences amplified using the primer set nu-SSU-0817 and nu-SSU-1536 (clone library B) were of fungal origin. However, there is conflicting evidence for the specificity of the other two 18S rDNA primer pairs investigated. Borneman and Hartin (2000) reported the primer pair nu-SSU-0817 and nu-SSU-1196 (clone library A) to

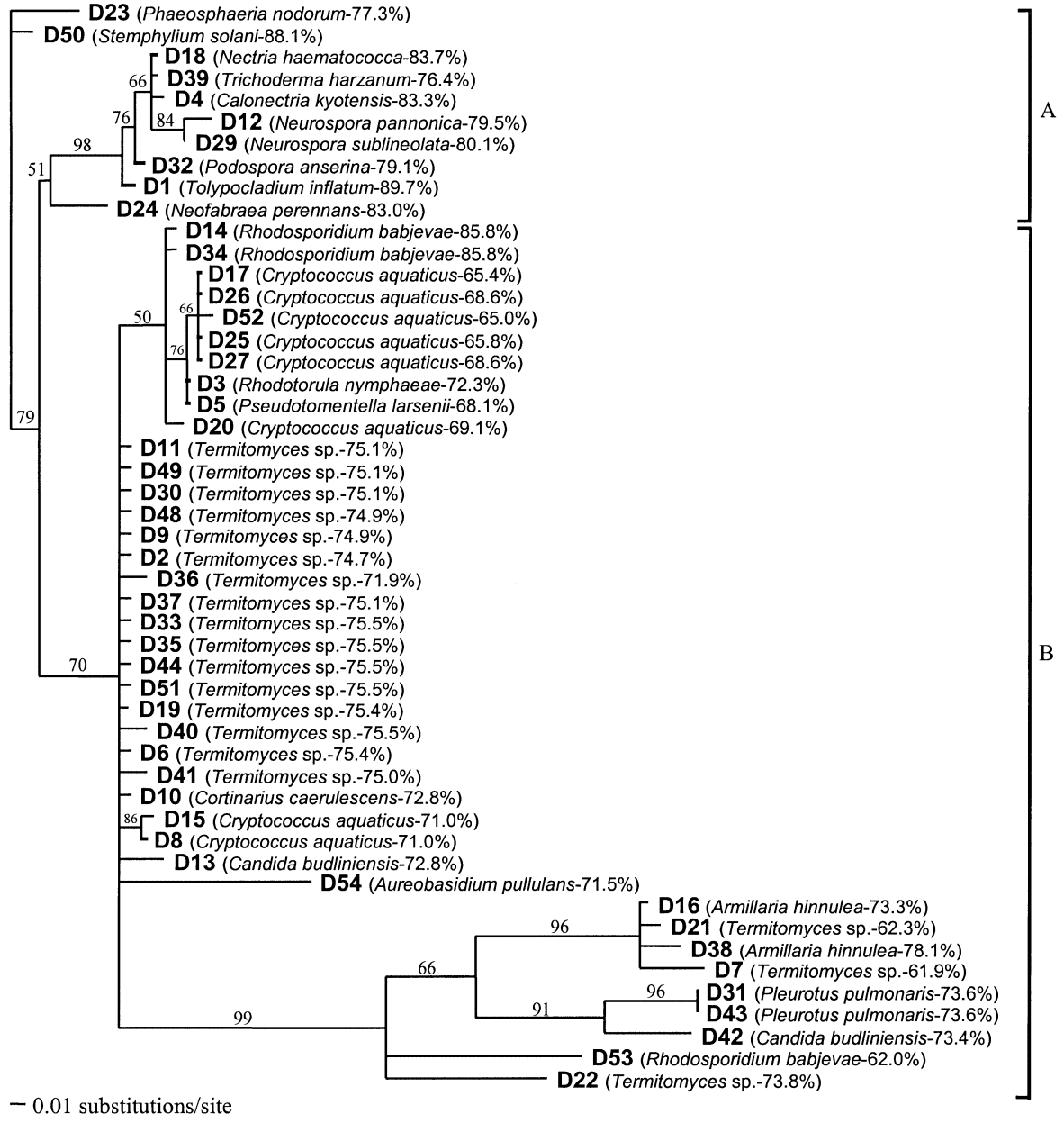

Fig. 6. Neighbour-joining tree showing the relationship between 50 ITS clones obtained from DNA extracted from grassland soil using the primers ITS1-F (3) and ITS4 (26) (D clone library) based on 5.8S rDNA sequences. Sequences ranged in size from 164 to 168 nucleotides. Abbreviations and other information as described in the legend to Fig. 1.

be specific for fungal 18S rDNA sequences. In contrast, of the 50 clones obtained using this primer set that were sequenced in our study (clone library A), only 15 were of fungal origin. The majority of the remaining 35 sequences were most closely related to species of soil invertebrates that clustered together in the phylogenetic analysis. Not surprisingly, when these non-fungal sequences were removed from the analysis, small changes in the topology of the remaining fungal sequences were observed (see Fig. 2). For the EF4 and EF3 primer pair, Smit *et al.* (1999) demonstrated selective amplification of fungal 18S rDNA sequences from wheat rhizosphere soil, whereas Borneman and Hartin (2000) reported amplification, by the same primers, of some non-fungal templates. The 46 clones sequenced from the library obtained using the EF4 and EF3 primer pair in our study (clone library C) were all of fungal origin. These apparent discrepancies highlight the need to evaluate primers carefully when amplifying genes from environmental samples, where much of the community may be poorly represented by cultured organ-

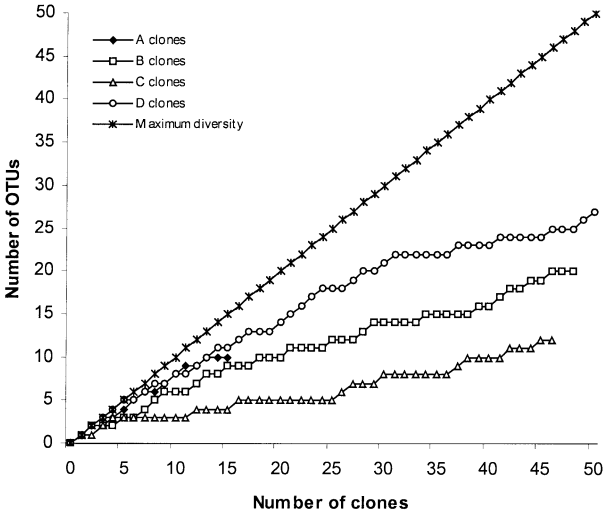

Fig. 7. Collector's curves for the A, B, C and D clone libraries. Clones were grouped into OTUs at a level of 99% sequence similarity. The maximum level of diversity that could be observed in a clone library (i.e. each clone being a different OTU) is also shown for comparison.

isms (Hawksworth and Rossman, 1997; Hawksworth, 2001) and where coverage in databases may be poor. It is possible that discrepancies in primer specificity may become more apparent over time as the primers are tested further, particularly with samples collected from different environments containing a diversity of eukaryotic communities. Recently, molecular techniques have been used to assess fungal communities in various sample types including painted art objects and window surfaces (Möhlenhoff *et al*., 2001; Schabereiter-Gurtner *et al*., 2001). Therefore, although the primer pair nu-SSU-0817 and nu-SSU-1196 lacks the desired specificity for characterizing fungal communities in soil, it may be useful for analysing fungal communities in other sample types, particularly as they appear to amplify 18S rDNA sequences from each of the four main fungal phyla.

A further constraint on the design of fungal-specific 18S rDNA primers is that of potential bias towards certain taxonomic groups, which is likely to increase with increasing specificity. Smit *et al*. (1999) suggested that primers EF4 and EF3 were slightly biased towards the amplification of basidiomycetes and zygomycetes after analysis of cultured organisms and environmental samples. Indeed, 35 of the 46 sequences obtained with these primers in our study were of basidiomycete origin (Fig. 4). The remaining 11 sequences represented nine ascomycetes and two zygomycetes. These results may be caused by primer bias, but the ratio of basidiomycete to ascomycete sequences was similar for each of the four primer pairs used, and the diversity (as indicated by branch lengths) observed within the ascomycete and basidiomycete groups in the phylogenetic trees was similar for each primer pair. Diversity within the ascomycete sequences was always greater than that within the basidiomycete sequences, the only exception being the basidiomycete clade III in the ITS phylogenetic tree (D clones). In addition, with the exception of the EF4 and EF3 (C clone) phylogenetic tree, the majority of amplified basidiomycete sequences in the clone libraries were most closely related to fungal species of the same family. For example, a large number of basidiomycete sequences had high (>98%) sequence similarity to *Panellus stipticus* and *Laccaria pumila* in the 18S rDNA clone libraries A and B, and high (> 75%) sequence similarity to *Termitomyces* sp. in the ITS (D) clone library. Each of these three fungal basidiomycete species belongs to the family Tricholomataceae. Therefore, although the findings may result from some level of primer bias, similar proportions of basidiomycete sequences within clone libraries targeting different regions of the fungal genome suggest that bias may be less significant than previously considered and that the results provide a true reflection of the natural fungal community.

Sequences representing chytridiomycetes were only detected in the A and B clone libraries, which was expected, as they are considered to have little importance in soil ecosystems (Smit *et al*., 1999). However, the majority of fungi forming arbuscular mycorrhizal symbioses with grass roots belong to the *Zygomycota* (Smith and Read, 1997) and would be expected to constitute a significant proportion of fungal biomass in pasture soils. Surprisingly, however, zygomycete sequences were only detected in the A and C clone libraries. This may reflect primer bias, as has been suggested previously (Smit *et al*., 1999), but may have resulted from other factors, including bias associated with the DNA extraction procedure. Although the bead beating DNA extraction procedure used in this study has been used previously in the analysis of both bacterial and fungal communities (Smit *et al*., 1999; Griffiths *et al*., 2000), the extent to which fungal mycelium and spores from different fungal taxa are lysed by this technique is unknown.

A major limitation in investigating fungal diversity in environmental samples is the lack of reference sequence information in databases. The majority of fungal sequences generated with the 18S rDNA primer pairs (A, B and C clone libraries) had relatively high similarity to fungal 18S rDNA database sequences. However, many of the clones sequenced from the ITS clone library clustering in the basidiomycete clade III matched database sequences over a very small stretch of the sequence (150–250 bp), with most similarity over the region of sequence corresponding to the 5.8S rRNA gene, which is thought to be highly conserved in fungi (Hibbett, 1992). Sequences in the basidiomycete clade III of the ITS phylogenetic tree (Fig. 5) appear to be considerably more diverse than the other groups of basidiomycetes, as indi-

cated by long branch lengths, but their placement within the basidiomycete group is well supported (100% bootstrap value). This suggests that they are more closely related to basidiomycetes than to other fungal phyla, although the lack of reference ITS sequences prevents identification of these sequences to a level higher than phylum. The sequence of clone D54 clustered within the basidiomycete clade III, and placement within the basidiomycete clade was supported by phylogenetic analysis of the 5.8S rRNA gene sequences. However, the closest match using a FASTA search was with an ascomycete, suggesting possible placement close to the point of ascomycete and basidiomycete divergence. Support for this hypothesis would require analysis with greater resolution. Therefore, although the limit of resolution of the ITS region is better than that of the more conserved 18S rDNA region, identification of sequences from previously undescribed fungal species, or species poorly represented in the database, is impossible. Although the 18S rDNA database information is also lacking, the conserved nature of this region enables identification of the majority of fungal sequences from environmental samples to a higher taxonomic level than phylum. The ITS sequences in the basidiomycete clade III may represent previously undescribed fungal taxa or species that do not have representative sequences in the database.

Estimating the coverage of fungal diversity was attempted by the construction of collector's curves for each clone library by grouping sequences into OTUs at a similarity of >99%. This similarity level was chosen because of the high level of conservation in fungal 18S rRNA genes, and plotted curves began to plateau for each clone library except clone library A. As the ITS region is not as highly conserved as the rRNA genes, the sequences from clone library D were clustered into OTUs at several levels of sequence similarity, but no change was observed in the collector's curve until values <90% were used, when curves reached a plateau more rapidly. Interestingly, even at a 99% sequence similarity, 46% of the diversity contained within the clone library was sampled in the 50 sequenced clones, although this may reflect the low level of sequence divergence within the basidiomycete sequences. The collector's curve data appear to demonstrate that reasonable coverage of the fungal diversity within the 18S rDNA and ITS clone libraries was detected in the 50 clones sequenced. This contrasts with a similar analysis of bacterial 16S rDNA clone libraries from the same soil type (McCaig *et al*., 1999), in which coverage was not achieved by sequencing 45 clones.

The data presented here demonstrate the requirement for rigorous testing of the specificity of PCR primers designed for the amplification of target organisms or taxa and the need to include DNA sequencing as a tool in this validation process. However, it also demonstrates the usefulness of both 18S rDNA and ITS primers in assessing the biodiversity of fungal communities in complex environmental samples and indicates possible limitations of the approach that require consideration. The lack of fungal DNA sequence information available in databases and the possibility of primer bias towards certain fungal taxonomic groups are likely to preclude the use of a single primer pair with universal application to studies of fungal biodiversity, and future analyses may require a combination of approaches.

Experimental procedures

DNA extraction

A 3 cm diameter × 5 cm soil core was collected from a natural grassland field site at Sourhope Research Station (national grid reference NT 854196) in the Borders Region of Scotland, UK. DNA was extracted from duplicate 0.5 g soil subsamples using the method previously described by Griffiths *et al*. (2000). Briefly, this involved the addition of 0.5 ml of hexadecyltrimethylammonium bromide (CTAB) extraction buffer and 0.5 ml of phenol–chloroform–isoamyl alcohol (25:24:1; pH 8.0) to 0.5 g of soil in a Bio-101 Multimix matrix tube. The sample was then lysed for 30 s at a speed of 5.5 ms^{-1} in a FastPrep bead beating system (Bio-101), and the aqueous phase was separated by centrifugation. The sample was extracted further with an equal volume of chloroform–isoamyl alcohol (24:1) before precipitating the nucleic acids with two volumes of 30% (w/v) polyethylene glycol 6000 in 1.6 M NaCl for 2 h at room temperature followed by centrifugation. Pelleted nucleic acids were then washed with ice-cold 70% (v/v) ethanol and air dried before resuspension in 50 μl of Tris-EDTA buffer (pH 7.4).

PCR amplification

PCR amplification of fungal 18S rDNA was carried out using the primer pairs: nu-SSU-0817 and nu-SSU-1196; nu-SSU-0817 and nu-SSU-1536 (Borneman and Hartin, 2000); and EF4 and EF3 (Smit *et al*., 1999), whereas fungal ITS sequences were amplified using primers ITS1-F (Gardes and Bruns, 1993) and ITS4 (White *et al*., 1990) (Table 1). PCRs were carried out on an Omn-E thermal cycler (Hybaid) using 50 μl reaction volumes containing: ≈ 50 ng of template DNA, 20 pmol of each primer, 2.0 mM $MgCl_2$, 250 μM each dATP, dCTP, dGTP and dTTP, 10× buffer [20 mM Tris-HCl (pH 7.5), 100 mM KCl, 1 mM dithiothreitol (DTT), 0.1 mM EDTA, 0.5% Tween 20 (v/v), 0.5% Nonidet P40 (v/v), 50% glycerol (v/v)], 1 μl of a BSA solution (20 mg ml^{-1}) and 2.5 U of Expand High Fidelity DNA polymerase (Roche). Cycling parameters were 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 30 s with a final extension at 72°C for 5 min for each primer pair. Reactions were performed in duplicate, and negative controls (containing no DNA) were included in each PCR. Amplification products were electrophoresed in 1.5% (w/v) agarose gels, stained with ethidium bromide and visualized under UV light.

Table 1. 18S rDNA and ITS PCR primers.

Primer pair	Target genomic region	Primer sequence (5′–3′)	Approx. fragment size (bp)	Reference
nu-SSU-0817	18S rDNA	TTAGCATGGAATAATRRAATAGGA	420	Borneman and Hartin (2000)
nu-SSU-1196		TCTGGACCTGGTGAGTTTCC		Borneman and Hartin (2000)
nu-SSU-0817	18S rDNA	TTAGCATGGAATAATRRAATAGGA	760	Borneman and Hartin (2000)
nu-SSU-1536		ATTGCAATGCYCTATCCCCA		Borneman and Hartin (2000)
EF4	18S rDNA	GGAAGGGRTGTATTTATTAG	1500	Smit *et al.* (1999)
EF3		TCCTCTAAATGACCAAGTTTG		Smit *et al.* (1999)
EF4[a]	18S rDNA	GGAAGGGRTGTATTTATTAG	550	Smit *et al.* (1999)
fung5[a]		GTAAAAGTCCTGGTTCCCC		Smit *et al.* (1999)
ITS1-F	ITS	CTTGGTCATTTAGAGGAAGTAA	450–750	Gardes and Bruns (1993)
ITS4		TCCTCCGCTTATTGATATGC		White *et al.* (1990)

a. Primers used for sequencing EF4 and EF3 clones.

Cloning and sequencing of 18S rDNA and ITS PCR products

PCR amplification products were purified using the Qiaquick PCR purification kit (Qiagen) before cloning with the pGEM-T Easy vector system (Promega) using a molar vector–insert ratio of 1:1. Ligations were transformed into JM109 high efficiency competent cells (Promega). Fifty white colonies from each of the four clone libraries (A, nu-SSU0817 and nu-SSU1196; B, nu-SSU0817 and nu-SSU1536; C, EF4 and EF3; and D, ITS1-F and ITS4) were picked at random and screened for positive inserts by performing colony PCR with the M13 forward and reverse vector primers (Promega) using the conditions outlined above, except that BSA was omitted from the reactions. Positive PCR products were visualized by agarose gel electrophoresis and subsequently purified using the Qiaquick PCR purification kit (Qiagen). Purified DNA was then sequenced using the BigDye Terminator cycle sequencing kit with an automated DNA sequencer (ABI model 377; Applied Biosystems). Sequencing reactions were performed with the vector primers T7 and SP6 (Promega) in each case, except for products from the EF4 and EF3 clone library, which were sequenced using primers EF4 and fung5 (Smit *et al.*, 1999).

Sequence analysis

DNA sequences were edited, and consensus sequences were obtained, using the SEQUENCHER software package (version 3.0; Gene Codes Corporation). All sequenced clones were analysed for the presence of chimeras using the CHIMERA CHECK program (version 2.7) from the Ribosomal Database Project (RDP). Sequences suspected of being chimeric were not included in further analyses. Final sequences were then aligned using CLUSTALW (version 1.8.2; Thompson *et al.*, 1994) for each of the four clone libraries. The transition/transversion ratio and the γ distribution parameter were estimated for each data set using TREE-PUZZLE (version 5.0; Strimmer and von Haeseler, 1996) before conducting a neighbour-joining analysis using the F84 model in PAUP (version 4.0b10; Swofford, 2002) with 1000 bootstrap replicates. Additionally, all sequences were analysed using the FASTA 3.0 program (Pearson and Lipman, 1998) to determine the closest matches in the GenBank and EMBL nucleotide databases. The level of sequence similarity between clones was also assessed for each clone library using the HOMOLOGIES program (within the EGCG extensions to the Wisconsin Package, version 8.1.0; Rice, 1996). Clones from the three 18S rDNA clone libraries (i.e. A, B and C) were clustered into operational taxonomic units (OTUs) at a level of >99% sequence similarity, whereas the clones from the ITS clone library (D) were clustered into OTUs at levels of >70%, 80%, 90%, 95%, 98% and 99% sequence similarity. Collector's curves (number of OTUs detected versus number of clones sequenced) were plotted to compare the diversity observed in each of the clone libraries. The percentage coverage was also calculated for each clone library [i.e. 1–(number of OTUs/number of clones sampled) $\times$ 100].

Nucleotide sequence accession numbers

All nucleotide sequences were submitted to GenBank under accession numbers AF504687–AF504736 for clones A1–A50, AF504737–AF504784 for clones B1–B55, AF504785–AF504830 for clones C1–C51 and AF504831–AF504880 for clones D1–D54.

Acknowledgements

This work was funded by The Macaulay Development Trust as part of The Macaulay Institute/University of Aberdeen Soil Health Initiative. C.D.C. is funded by the Scottish Executive Environment and Rural Affairs Department. The authors gratefully acknowledge Martin Embley for guidance with the phylogenetic analyses.

References

Borneman, J., and Hartin, R.J. (2000) PCR primers that amplify fungal rRNA genes from environmental samples. *Appl Environ Microbiol* **66:** 4356–4360.

Bruns, T.D., White, T.J., and Taylor, J.W. (1991) Fungal molecular systematics. *Annu Rev Ecol Syst* **22:** 525–564.

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. *J Microbiol Methods* **43:** 133–151.

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

Griffiths, R.I., Whiteley, A.S., O'Donnell, A.G., and Bailey, M.J. (2000) Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl Environ Microbiol* **66:** 5488–5491.

Hawksworth, D.L. (1991) The fungal dimension of biodiversity: magnitude, significance, and conservation. *Mycol Res* **95:** 641–655.

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

Hawksworth, D.L., and Rossman, A.Y. (1997) Where are all the undescribed fungi? *Phytopathology* **87:** 888–891.

Hibbett, D.S. (1992) Ribosomal RNA and fungal systematics. *Trans Mycol Soc Japan* **33:** 533–556.

Kowalchuk, G.A. (1998) Fungal community analysis using denaturing gradient gel electrophoresis. In *Molecular Microbial Manual 3.4.6*. Akkermans, A.D.L., van Elsas, J.D., and de Bruijn, F.J. (eds). Dordrecht: Kluwer, pp. 1–16.

Kowalchuk, G.A., Gerards, S., and Woldendorp, J.W. (1997) Detection and characterisation of fungal infections of *Ammophila arenaria* (Marram grass) roots by denaturing gradient gel electrophoresis of specifically amplified 18S rDNA. *Appl Environ Microbiol* **63:** 3858–3865.

Lowell, J.L., and Klein, D.A. (2001) Comparative single-strand conformation polymorphism (SSCP) and microscopy-based analysis of nitrogen cultivation interactive effects on the fungal community of a semiarid steppe soil. *FEMS Microbiol Ecol* **36:** 85–92.

McCaig, A.E., Glover, L.A., and Prosser, J.I. (1999) Molecular analysis of bacterial community structure and diversity in unimproved and improved upland grass pastures. *Appl Environ Microbiol* **65:** 1721–1730.

Möhlenhoff, P., Müller, 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 Microbiol Lett* **195:** 169–173.

Page, R.D.M., and Holmes, E.C. (1998) *Molecular Evolution: a Phylogenetic Approach*. Oxford: Blackwell Science Publishing.

Pearson, W.R., and Lipman, D.J. (1998) Improved tool for biological sequence analysis. *Proc Natl Acad Sci USA* **85:** 2444–2448.

Pennanen, T., Paavolainen, L., and Hantula, J. (2001) Rapid PCR-based method for the direct analysis of fungal communities in complex environmental samples. *Soil Biol Biochem* **33:** 697–699.

Rice, P. (1996) *Program Manual for the EGCG Package*. Cambridge: The Sanger Centre.

Schabereiter-Gurtner, C., Piñar, G., Lubitz, W., and Rölleke, S. (2001) Analysis of fungal communities on historical church window glass by denaturing gradient gel electrophoresis and phylogenetic 18S rDNA sequence analysis. *J Microbiol Methods* **47:** 345–354.

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. *Appl Environ Microbiol* **65:** 2614–2621.

Smith, S.E., and Read, D.J. (1997) *Mycorrhizal Symbiosis*. London: Academic Press.

Strimmer, K., and von Haeseler, A. (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol Biol Evol* **13:** 964–969.

Swofford, D.L. (2002) *paup*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sunderland, MA: Sinauer Associates.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22:** 4673–4680.

Vainio, E.J., and Hantula, J. (2000) Direct analysis of wood-inhabiting fungi using denaturing gradient gel electrophoresis of amplified ribosomal DNA. *Mycol Res* **104:** 927–936.

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 *PCR Protocols: a Guide to Methods and Applications.* Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. (eds). New York: Academic Press, pp. 315–322.