Mycorrhiza (2003) 13:227–231
DOI 10.1007/s00572-003-0249-2

COMMENT

Anne Pringle · Jean-Marc Moncalvo · Rytas Vilgalys

Revisiting the rDNA sequence diversity of a natural population of the arbuscular mycorrhizal fungus *Acaulospora colossica*

Received: 2 February 2002 / Accepted: 13 May 2003 / Published online: 30 July 2003

Abstract In 1999, the diversity of a field population of the arbuscular mycorrhizal (AM) fungus *Acaulospora colossica* was characterized using DNA sequence data. Since 1999, AM fungal sequences have accumulated rapidly within public databases. Moreover, novel phylogenetic tools have been developed and can be used to interpret the data. A second analysis of those sequences collected in 1999 demonstrates that while the majority of the sequences are, in fact, sequences of *A. colossica*; a minority of the sequences still cannot be identified with confidence. Those sequences identified as *A. colossica* can be used to show that (1) the nuclear rDNA ITS regions are remarkably diverse, and (2) sequences isolated from different spores of the same site may be more closely related to each other than to sequences of other sites, so that the genetic diversity of an AM fungal field population may be spatially structured; however, identical sequences can also be recovered from different sites.

Keywords Internal transcribed spacer · Glomales · Glomeromycota · NCBI · GenBank nucleotide database

Recently, a Comment was published by Clapp et al. in *Mycorrhiza* entitled "Glomales rRNA gene diversity: all that glistens is not necessarily glomalean?" in which they explored the difficulties associated with gathering genetic data from arbuscular mycorrhizal (AM) fungi (Clapp et al. 2002). The authors stated that "many papers continue to be published without adequate critical examination of the data", and that "many of these errors result from the 'publish or else' attitude in short-term research projects". Finally, the authors imply that data may be corrupted by the "mislabeling of samples".

One of the publications used as an example of problematic data is Pringle et al. (2000). Although Clapp et al. (2002) also state that they are "not criticizing the authors" of the cited works, we believe that anyone reading this Comment would conclude that the publication of Pringle et al. (2000) is seriously flawed.

We would like to address the issues raised by Clapp et al. and to discuss (1) why it was difficult to assign a taxonomic rank to the nucleotide sequence data we collected in 1999, (2) the inclusion of outgroups in our original analyses, (3) our subsequent decision to label the data as Glomalean, (4) how the sequences might be reanalyzed in 2003, and finally, (5) the contribution that these data make to the literature of Glomalean population genetics.

In Pringle et al. (2000), seven types of rDNA sequences were recovered from a natural population of the AM fungus *Acaulospora colossica* (Shultz et al. 1999). Pringle et al. (2000) was one of the first publications in which nucleotide sequence data were used to explore the rDNA diversity of a single Glomalean spore. In 1999, as the manuscript was written and as stated in Pringle et al. (2000), the NCBI nucleotide database included approximately 100 ITS sequences of AM fungi but no ITS sequences of the genus *Acaulospora*. At that time and as stated in the publication, we were unable to find matches for our data within the NCBI database. In contrast, at this present time, a comparison of one of our *A. colossica* sequences to the database finds matches with other *A. colossica* sequences, both those identified in Pringle et al. (2000) and others, and also with sequences from *A. denticulata*, *A. laevis*, *A. mellea*, and *A.*

A. Pringle (✉)
Department of Plant and Microbial Biology,
University of California,
111 Koshland Hall, Berkeley, CA 94720-3102, USA
e-mail: apringle@uclink.berkeley.edu
Fax: +1-510-6424995

J.-M. Moncalvo
Centre for Biodiversity and Conservation Biology,
Royal Ontario Museum, Toronto, Canada

J.-M. Moncalvo
Department of Botany, University of Toronto,
100 Queen's Park, Toronto, Ontario, M5S 2C6, Canada

R. Vilgalys
Department of Biology, Duke University,
Durham, NC 27708, USA

morrowiae. Matches are also found to other genera of AM fungi, including *Archaeospora trappei* and *Gigaspora margarita*.

At that time, D. Redecker allowed us to compare our sequences to his personal 5.8S rDNA database for fungi (now public in Redecker et al. 1999). This analysis indicated that our type 1 sequences were in fact monophyletic with *Acaulospora laevis*, hypothesized to be a close relative of *A. colossica* (Shultz et al. 1999). The type 1 sequences constitute the majority of sequences (24 of 39) reported in Pringle et al. (2000).

But the identities of the other types were less clear. Clapp et al (2002) incorrectly state that we did not use outgroup sequences in our analyses. In fact, we wrote (Pringle et al. 2000, p 262): "An analysis of our 39 sequences with other 5.8S sequences from a broad spectrum of fungi (not available through the NCBI database but kindly supplied to us by D. Redecker), including 46 Ascomycetes, 18 Basidiomycetes, and 22 Glomalean taxa [showed that] the type 1 sequences are monophyletic with *A. laevis* [and that] only type 5 was found apart from any Glomales [...]". This analysis, which included both Ascomycetes and Basidiomycetes as outgroups, allowed us to discuss the question of whether or not sequences of types 2–7 were contaminants (Pringle et al. 2000, pp 263–264; see especially our discussion of the type 5 sequence). At that time, the analysis and several additional lines of evidence suggested that the sequences were not contaminants. We made a decision to enter our sequences into the NCBI database as sequences isolated from *A. colossica* as this was, in fact, where the sequences were isolated. We had no conclusive evidence that the sequences belonged to other taxa.

There are uncertainties associated with the use of nucleotide sequence data from the 5.8S rDNA gene in phylogenetic analyses. This gene is short (ca. 157 bp in fungi), and provides few informative characters. In addition, tree topologies produced from 5.8S sequence data are difficult to support statistically (see Fig. 1 and Redecker et al. 1999), and may vary significantly with both the taxon sampling and the method used for phylogenetic reconstruction (data not shown). For instance, in the 5.8S trees depicted in both Redecker et al. (1999) and Clapp et al. (2002), the *Glomus occultum* group does not cluster with other Glomalean taxa. Redecker et al. (1999) explained the nesting of the *G. occultum* group within the Basidiomycetes of an NJ tree as the result of long-branch attraction. If so, then in the NJ tree depicted in Clapp et al. (2002), one of our sequences (the type 7 sequence) that clusters with *G. occultum* cannot be rejected as being of Glomalean origin. [More recently, *G. occultum* was renamed *Paraglomus occultum*, and the nomenclature of other taxa in this group also updated (Morton and Redecker 2001). The nomenclature reflects the molecular divergence of these species from other Glomalean taxa, but the species are still considered Glomalean].

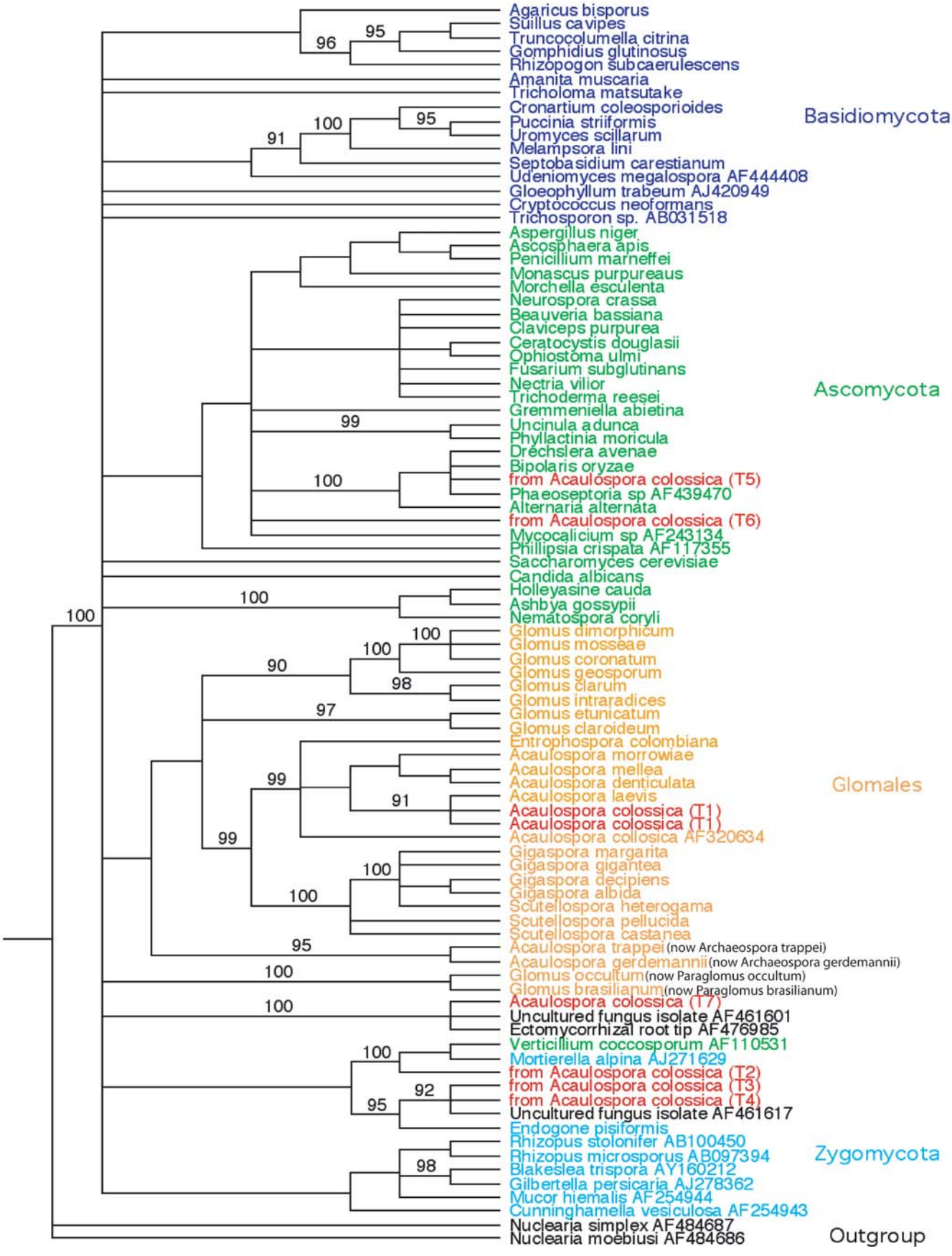
Several empirical and theoretical studies have demonstrated that when the amount of sequence data is limited and/or long-branch attraction is a potential problem, maximum-likelihood methods (ML) should be chosen over distance or parsimony methods for phylogenetic reconstruction (Cunningham et al. 1998; Steinbachs et al. 2001). Recently, it has been argued that the combination of ML and Bayesian statistics will provide a powerful tool with which to infer phylogenies (Huelsenbeck et al. 2001). In particular, this approach may be less conservative than the more traditional bootstrap or jackknife techniques in evaluating statistical confidence for branches in a phylogeny, especially when there are few characters for nodal support (Alfaro et al. 2003).

Here we use Bayesian analyses to reexamine the phylogenetic affinities of the seven sequence types discovered by Pringle et al. (2000). The aim of this analysis is to better identify these sequences by taking advantage of the much larger database currently available (in contrast to the data available at the time of the original study) in order to update the taxonomic status of these sequences in GenBank. We do agree with Clapp et al. (2002) that mislabeled sequences in public databases can "lead to much confusion".

We constructed our 5.8S data set as follows: (1) at least one of each of the sequence types reported in Pringle et al. (2000) was included in the fungal 5.8S data set of Redecker et al. (1999); (2) sets of nearly identical sequences (mostly congeneric sequences) were removed from that data set; (3) sequences found in a BLAST search in GenBank (as of 7 February 2003) to have the highest matches to the seven types of sequences discovered in Pringle et al. (2000) were added to the data set; (4) recently published sequences from the Zygomycetes were added; (5) two sequences from the genus *Nuclearia* were used as an outgroup to root the fungal phylogeny. *Nuclearia* has been shown recently to be a putative sister group to the fungi (Amaral-Zettler et al. 2001; Allen G. Collins, personal communication); (6) a small region of an ambiguous alignment near the 3' end of the gene (corresponding to positions 126–145 in the alignment of Redecker et al. 1999) was excluded from the analyses.

The Bayesian analysis is shown in Fig. 1. With respect to the seven sequence types of Pringle et al. (2000), this analysis and comparisons of the ITS region of the various types to the ITS sequences now available in GenBank demonstrate that:

Fig. 1 Bayesian inference of fungal 5.8S rDNA phylogeny. The tree depicted is a 50% majority-rule consensus tree obtained in MrBayes vs. 2.01 (Huelsenbeck and Ronquist, http://morphbank.ebc.uu.se/mrbayes/) using the GTR6+gamma ML model of evolution and five MCMC runs; each chain examined 200,000 trees (the first 50,000 trees were discarded as "burn in"), and one of every 100 examined trees was sampled. We used the data set in Redecker et al. (1999), to which we added the *Acaulospora colossica* types 1 to 7 (T1-T7) sequences from Pringle et al. (2000) and sequences obtained from GenBank (indicated by an accession number following the taxon name). Two sequences of *Nuclearia* (a protist) were used as outgroup sequences to root the fungal phylogeny


Agaricus bisporus
Suillus cavipes
Truncocolumella citrina
Gomphidius glutinosus
Rhizopogon subcaerulescens
Amanita muscaria
Tricholoma matsutake
Cronartium coleosporioides
Puccinia striiformis
Uromyces scillarum
Melampsora lini
Septobasidium carestianum
Udeniomyces megalospora AF444408
Gloeophyllum trabeum AJ420949
Cryptococcus neoformans
Trichosporon sp. AB031518
Basidiomycota
Aspergillus niger
Ascosphaera apis
Penicillium marneffei
Monascus purpureaus
Morchella esculenta
Neurospora crassa
Beauveria bassiana
Claviceps purpurea
Ceratocystis douglasii
Ophiostoma ulmi
Fusarium subglutinans
Nectria vilior
Trichoderma reesei
Gremmeniella abietina
Uncinula adunca
Phyllactinia moricula
Drechslera avenae
Bipolaris oryzae
from Acaulospora colossica (T5)
Phaeoseptoria sp AF439470
Alternaria alternata
from Acaulospora colossica (T6)
Mycocalicium sp AF243134
Phillipsia crispata AF117355
Saccharomyces cerevisiae
Candida albicans
Holleyasine cauda
Ashbya gossypii
Nematospora coryli
Ascomycota
Glomus dimorphicum
Glomus mosseae
Glomus coronatum
Glomus geosporum
Glomus clarum
Glomus intraradices
Glomus etunicatum
Glomus claroideum
Entrophospora colombiana
Acaulospora morrowiae
Acaulospora mellea
Acaulospora denticulata
Acaulospora laevis
Acaulospora colossica (T1)
Acaulospora colossica (T1)
Acaulospora colloscia AF320634
Gigaspora margarita
Gigaspora gigantea
Gigaspora decipiens
Gigaspora albida
Scutellospora heterogama
Scutellospora pellucida
Scutellospora castanea
Acaulospora trappei (now Archaeospora trappei)
Acaulospora gerdemannii (now Archaeospora gerdemannii)
Glomus occultum (now Paraglomus occultum)
Glomus brasilianum (now Paraglomus brasilianum)
Glomales
Acaulospora colossica (T7)
Uncultured fungus isolate AF461601
Ectomycorrhizal root tip AF476985
Verticillium coccosporum AF110531
Mortierella alpina AJ271629
from Acaulospora colossica (T2)
from Acaulospora colossica (T3)
from Acaulospora colossica (T4)
Uncultured fungus isolate AF461617
Endogone pisiformis
Rhizopus stolonifer AB100450
Rhizopus microsporus AB097394
Blakeslea trispora AY160212
Gilbertella persicaria AJ278362
Mucor hiemalis AF254944
Cunninghamella vesiculosa AF254943
Zygomycota
Nuclearia simplex AF484687
Nuclearia moebiusi AF484686
Outgroup

- Type 1 sequences cluster with *Acaulospora laevis* (posterior probability, or pp, is 91%) in the core Glomales clade, but do not appear to be monophyletic with another sequence of *A. colossica* available from GenBank (AF320634).
- Type 2 sequences cluster with one Ascomycete sequence (*Verticillium coccosporum*) and one Zygomycete sequence (*Mortierella alpina*) with high statistical support (pp is 100%). This relationship is also suggested by some low degree of similarity in the ITS 1 and ITS 2 regions, as revealed by a BLAST search (data not shown). However, because the taxonomy and phylogenetic affinities of this clade are unclear, we are still unsure of how to best annotate type 2 sequences.
- Type 3 and 4 sequences cluster together with an uncultured, unidentified fungal isolate and *Endogone pisiformis* (pp is 95%). In a strict consensus tree of 10,000 equally parsimonious trees, all of these branches collapse to the root of the fungal phylogeny (data not shown). There is no detectable similarity between our types 3 and 4 sequences and *Endogone* sequences in the ITS 1 or ITS 2 region; however, BLAST searches indicate a good match between these sequences and several unidentified sequences isolated from uncultured fungi. For these reasons, we cannot confidently label this sequence with any available taxonomic rank, nor can we fully exclude the possibility that it is of Glomalean origin.
- Type 5 sequences cluster with several Ascomycete taxa with high statistical support (pp is 100%), including the hyphomycete *Altenaria*: a possible affiliation between the type 5 sequences and *Alternaria* was discussed in our original paper (Pringle et al. 2000, p 262). Interestingly, sequences originally reported from *Scutellospora castanea* (Hijri et al. 1999) and later discovered to be of Ascomycete origin (Redecker et al. 1999) are also in this clade, according to Clapp et al. (2002). With the data available at this time, we follow Redecker et al. (1999) and Clapp et al. (2002) and recognize that this sequence is of Ascomycete origin.
- The single type 6 (AF133790) sequence nests in an isolated position in a statistically poorly supported clade composed of Ascomycete taxa. An Ascomycete origin of this sequence appears likely, but cannot be determined with certainty.
- The single type 7 sequence clusters with strong support (pp is 100%) with unidentified and recently deposited sequences labeled as either uncultured fungi or ectomycorrhizal root tips. A close relationship between these sequences is further supported by the high level of ITS similarity, as revealed in a BLAST search. However, no statistically significant relationship was found between these and the other sequences of our analysis. In the tree depicted in Clapp et al. (2002), our type 7 sequence (AF133791) clusters with *Glomus occultum*. We cannot reject a Glomalean origin for this sequence. In fact, the latter result would support (and predate) a recent finding by Rodriguez et al. (2001), which demonstrated that a single spore of a Glomalean morphospecies can include rDNA sequences related to different Glomalean families.

In light of the analyses discussed here and in Clapp et al. (2002), we can confidently state that our type 1 sequences are characteristic of the genome of the morphospecies *Acaulospora colossica*. However, despite the exponential accumulation of data in the NCBI nucleotide database since the deposition of our sequences (on 10 March 1999), we are still unable to identify many of the sequences isolated from *A. colossica* spores. In order to avoid further confusion, we have annotated the sequences of types 2–7 within the NCBI database as "uncultured fungi" or "environmental samples" isolated "from spores of *A. colossica*" instead of "*A. colossica*".

Finally, we note that the conclusions of our original paper are still current and true, even if only type 1 sequences are analyzed (see Fig. 3 of Pringle et al. 2000): (1) the nuclear rDNA ITS regions are remarkably diverse (but the organization of that diversity, either within or between nuclei, is unclear); (2) sequences isolated from different spores of the same site may be more closely related to each other than to sequences of other sites, suggesting that the genetic diversity of an AM fungal field population is spatially structured; however, identical sequences can also be recovered from different sites.

In contrast to Clapp et al. (2002), we do not think that the population genetics of AM fungi should focus exclusively on cultures available from BEG or INVAM. Pringle et al. (2000) remains one of only a few publications to explore the genetic diversity of a natural population of Glomalean fungi.

References

Alfaro ME, Zoller S, Lutzoni F (2003) Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol Biol Evol 2:255–66

Amaral-Zettler LA, Nerad TA, O'Kelly CJ, Sogin ML (2001) The nucleariid amoebae: more protists at the animal-fungal boundary. J Eukaryot Microbiol 48:293–297

Clapp JP, Rodriguez A, Dodd JC (2002) Glomales rRNA gene diversity — all that glistens is not necessarily glomalean? Mycorrhiza 12:269–270

Cunningham CW, Zhu H, Hillis DM (1998) Best fit maximum-likelihood models for phylogenetic inference: empirical tests with known phylogenies. Evolution 52:978–987

Hijri M, Hosny M, van Tuinen D, Dulieu H (1999) Intraspecific ITS polymorphism in *Scutellospora castanea* (Glomales, Zygomycota) is structured within multinucleate spores. Fungal Genet Biol 26:141–151

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294: 2310–2314

Morton JB, Redecker D (2001) Two new families of Glomales, Archaeosporaceae and Paraglomaceae, with two new genera *Archaeospora* and *Paraglomus*, based on concordant molecular and morphological characters. Mycologia 93:181–195

Pringle A, Moncalvo JM, Vilgalys R (2000) High levels of variation in ribosomal DNA sequences within and among spores of a natural population of the arbuscular mycorrhizal fungus *Acaulospora colossica*. Mycologia 92:259–268

Redecker D, Hijri M, Dulieu H, Sanders IR (1999) Phylogenetic analysis of a dataset of fungal 5.8S rDNA sequences shows that highly divergent copies of internal transcribed spacers reported from *Scutellospora castanea* are of Ascomycete origin. Fungal Genet Biol 28:238–244

Rodriguez A, Dougall T, Dodd JC, Clapp JP (2001) The large subunit ribosomal RNA genes of *Entrophospora infrequens* comprise sequences related to two different glomalean families. New Phytol 152:159–167

Schultz P, Bever JD, Morton J (1999) *Acaulospora colossica* sp. nov. from an old field in North Carolina and morphological comparisons with similar species, *A. laevis* and *A. koskei*. Mycologia 91:676–683

Steinbachs JE, Schizas NV, Ballard JWO (2001) Efficiencies of genes and accuracy of tree-building methods in recovering a known *Drosophila* genealogy. Pac Symp Biocomput 6:606–617